JOUL, Volume *1*

Supplemental Information

Design Principles and Top Non-Fullerene Acceptor Candidates for Organic Photovoltaics

Steven A. Lopez, Benjamin Sanchez-Lengeling, Julio de Goes Soares, and Alán Aspuru-Guzik

SUPPORTING INFORMATION

CONTENTS

Supplementary Figures.

Figure S1. Histograms of the distributions of difference between calibrated and calculated values.
Figure S2. GP calibration of PCE values.
Figure S3. Histogram of distribution of calibrated PCE values.
Figure S4. Histograms of population difference between top and non-top molecules, grouped by fragment group.
Figure S5. Boxplot of PCE values for groups of molecules containing a particular fragment, colored based on Z-score and grouped by fragment group.
Figure S6. Distribution of counts of activated bits in all molecules.
Figure S7. Z-scores for a representative sample of 48 bits.
Figure S8. Boxplots for PCE values of molecules with activated bits.
Figure S9. Grid of example molecules of the representative sample of 48 bits.

Supplementary Tables.

Table S1. Summary table of references values utilized in the calibration set.
Table S2. Structures *n*-type materials in table S1.

Supplementary Datafiles

Datafile S1. 'Molecules.csv' – Comma separated value text file containing information on the entire dataset of acceptor molecules, 51280 in total. Data included is: InchI key, smiles string, HOMO energy, LUMO energy, Open Circuit Voltage (Voc), Short circuit current (Jsc), Fill Factor (FF) and power conversion efficiency (PCE). GP Calibrated values (_calib) and their uncertainty (_calib_std) are included for HOMO, LUMO and PCE values. The PCE calibrated values are PCE_calib_plus.

Datafile S2. 'Frags.csv' – Comma separated value text file containing information on the 112 fragments used to build the library of molecules. Data included is: Group classification for fragment (group), label, smiles string, representative PCE statistics (mean, SD, median, etc.) in python dictionary format for molecules that contain that fragment, PCE mean and SD, count of molecules containing a fragment for three categories: global, PCE < 8 and PCE > 8 (global_pop and global_pop_percent, etc.) and Z-score for that particular fragment.

Datafile S3. 'Bit.csv' – Comma separated value text file containing information on fingerprint analysis utilizing 8192 bits with a connectivity radius of 4. Data included is: Bit number (Bit), count (_count) and percentage of molecules (Pop %) with that particular activated bit percentage for two categories: entire data set (global) and top candidates (top), representative PCE statistics (mean, SD, median, etc.) in python dictionary format for molecules that have a activated bit, PCE mean and SD, mean and std of radius of the fingerprint for all molecules and Z-score for a particular activated bit.

Datafile S4. 'HOMO-LUMO_calibration.csv' – Comma separated value text file containing information on molecules used for HOMO-LUMO GP calibration. Contains experimental values and values calculated with TDDFT.

Datafile S5. 'PCE_calibration.csv' – Comma separated value text file containing information on molecules used for PCE GP calibration. Contains experimental values, values calculated with TDDFT, scharber model values and values obtained after calibration.

Figure S1. Histograms of the distributions of difference between calibrated and calculated values. *Related to Figure 2.* For some properties we also show the statistical measures for these distributions.

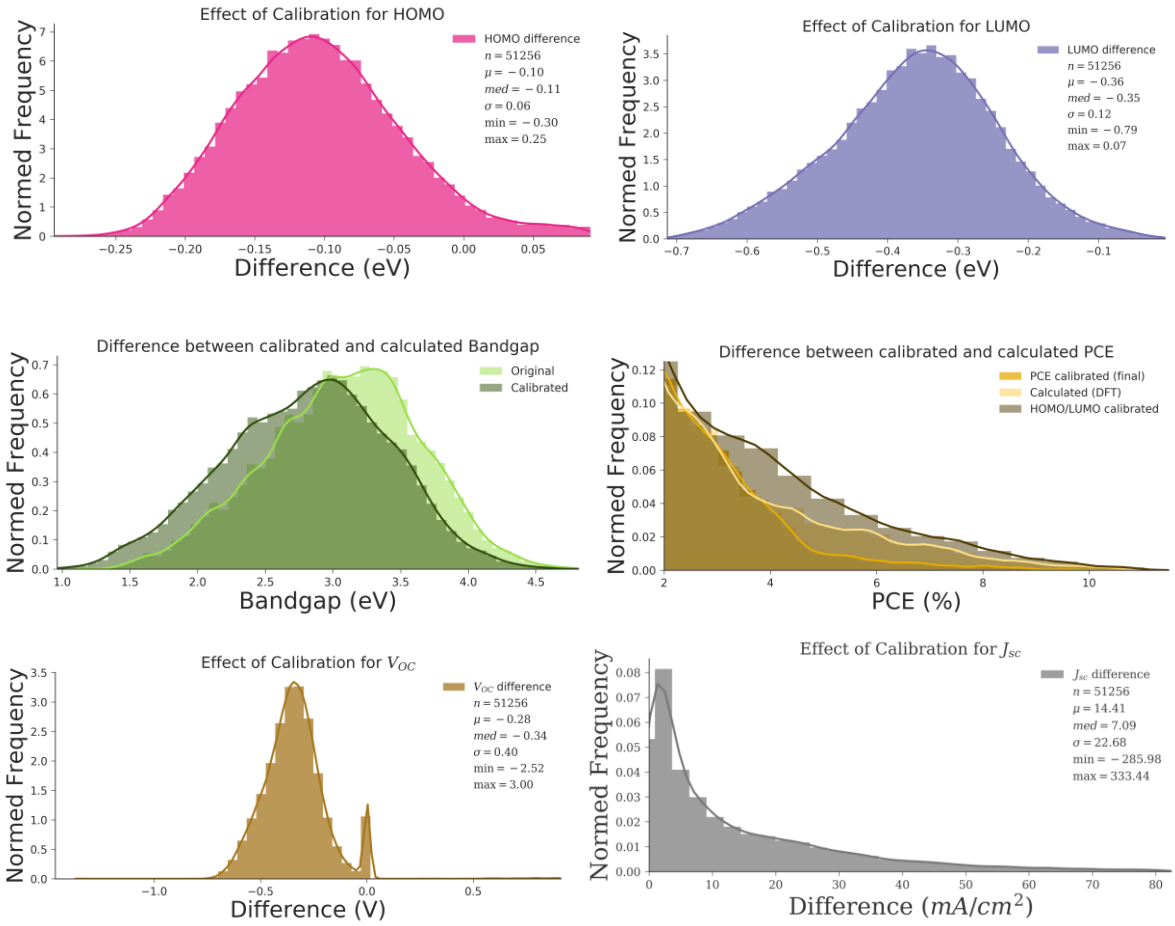


Figure S2. GP calibration of PCE values. Related to section *Machine learning: Gaussian Processes regression.* Experimental (EXP), calculated (DFT/Scharber model) and GP calibrated LOO-CV predictions (CALIB) are compared. Uncertainty increases with lightly shaded data points. Dotted line represents the ideal fit.

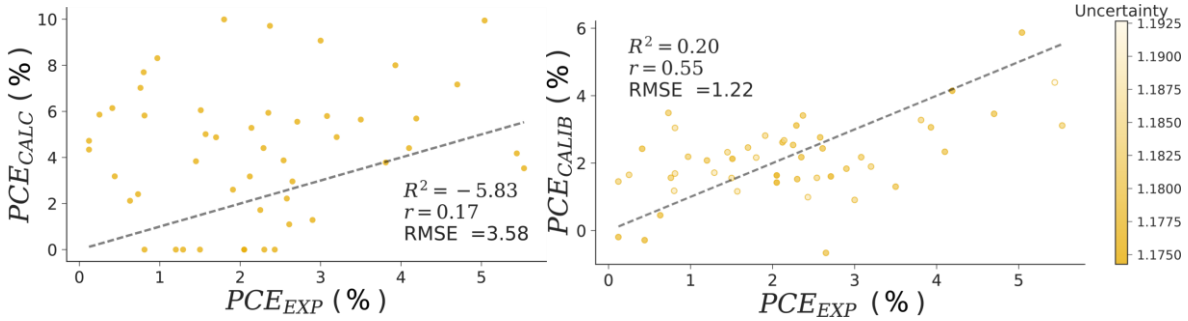


Figure S3. Histogram of distribution of calibrated PCE values. *Related to section Scharber Model: PCE calculations.* Top (green) and non-top (red) molecules are color coded, y-axis is using a log scale. Statistical meausures of the global distribuion are included.

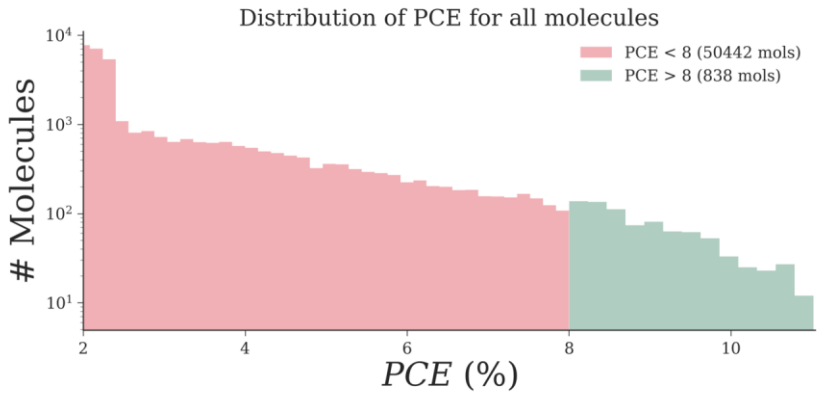
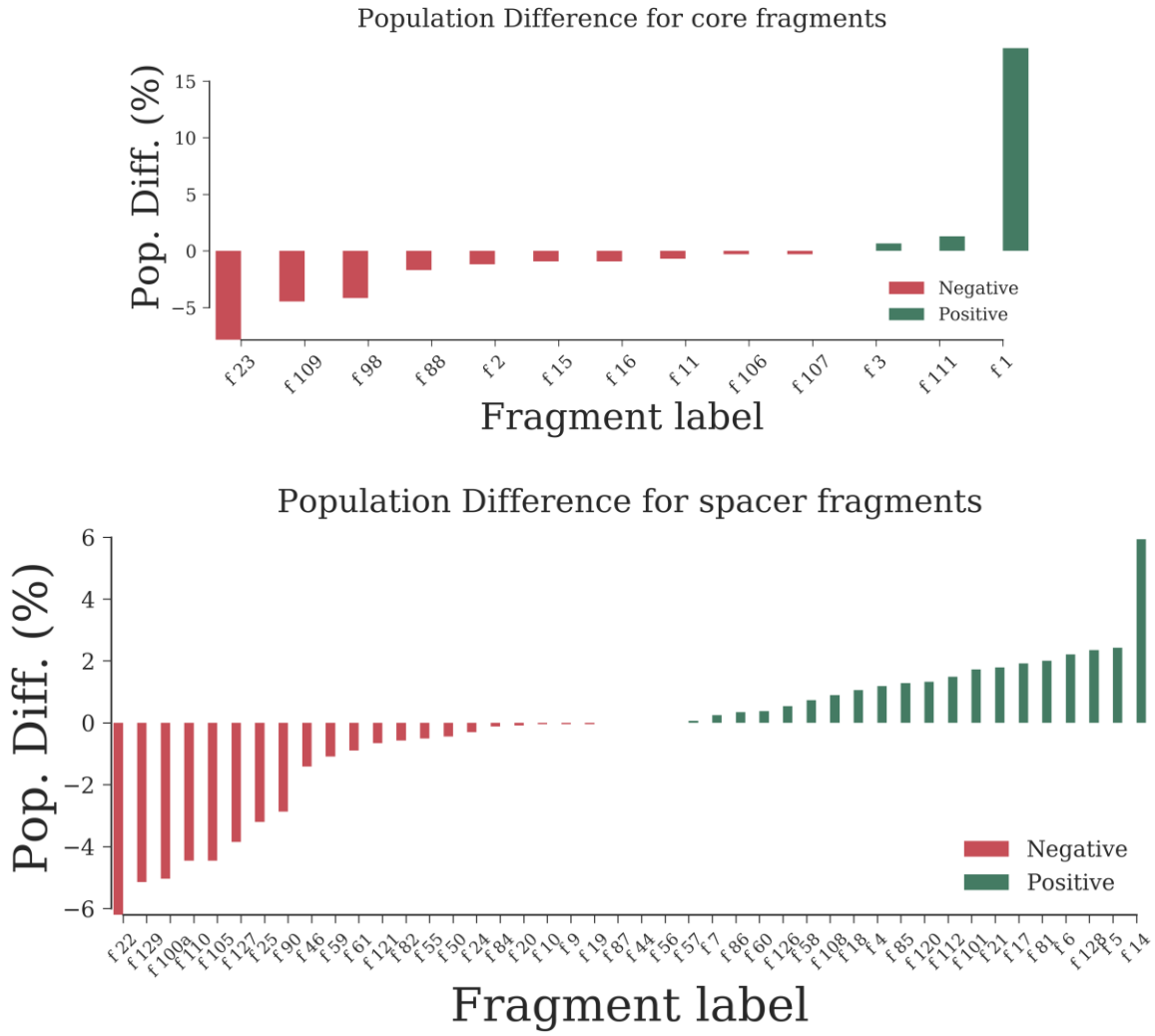


Figure S4. Barplots of population difference between top and non-top molecules, grouped by fragment group. *Related to section Scharber Model: PCE calculations.* Color-coded based on positive or negative representation. Positive differences indicate a fragment is more present in top candidate than non-top candidates.



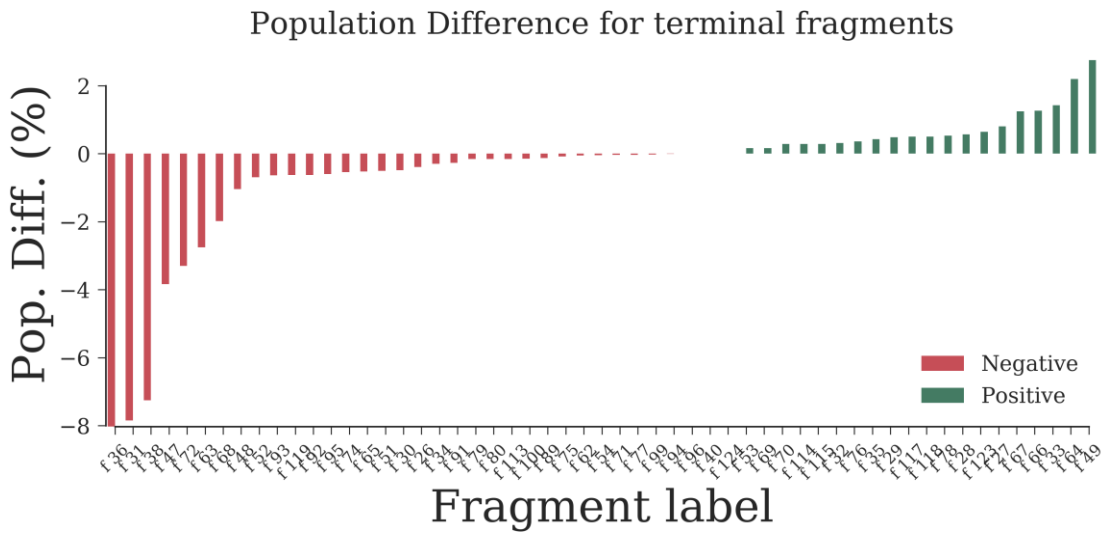


Figure S5. Boxplot of PCE values for groups of molecules containing a particular fragment. *Related to section Scharber Model: PCE calculations.* Colored based on Z-score and grouped by fragment group. Mean values are indicated as black lines, median as red dots. Each box represents the 25% to 75% quartile of the values with min and max values as whiskers.

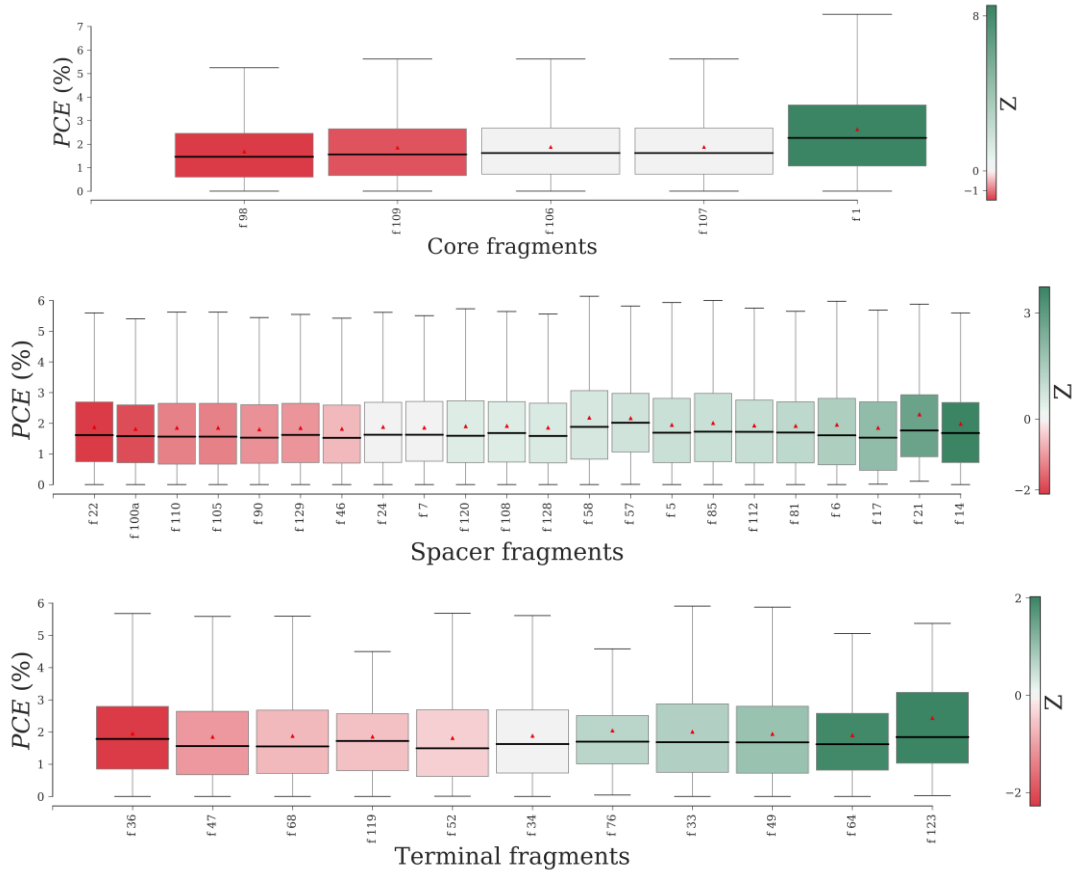


Figure S6. Distribution of counts of activated bits in all molecules. *Related to section Fingerprint statistics.*

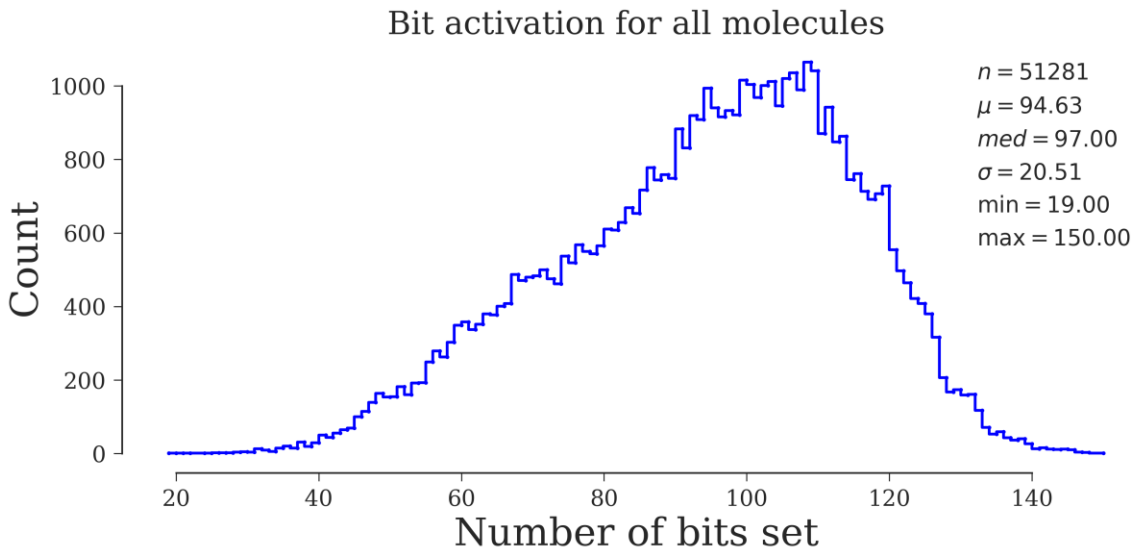


Figure S7. Z-scores for a representative sample of 48 bits. *Related to section Fingerprint statistics.* Representative sample was created using a histogram on the Z-scores and picking random bits from each bin. Binning was using the 'doane' estimator which works better with non-normal datasets and takes in to account outliers and data variability.

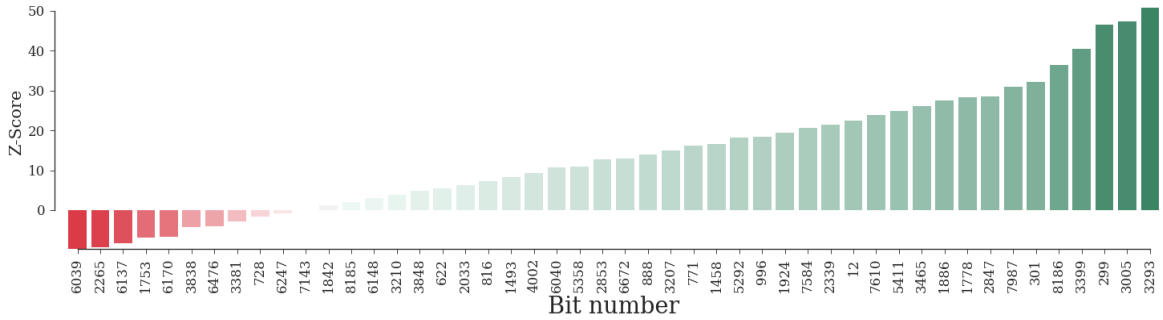


Figure S8. Boxplots for PCE values of molecules with activated bits. *Related to section Fingerprint statistics.* Using a representative sample of 48 bits. Boxes are colored based on their z-scores.

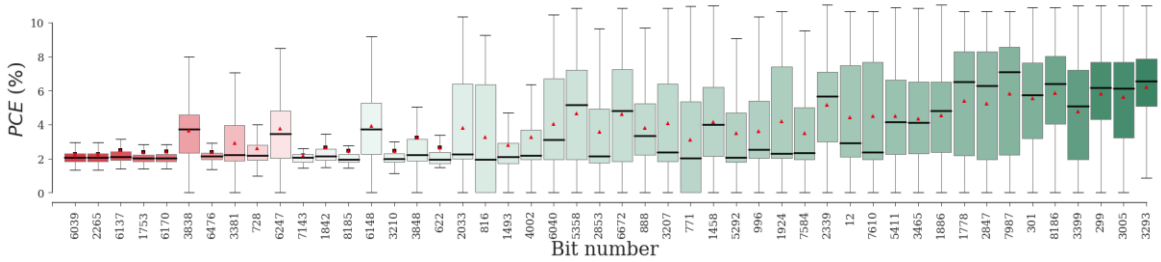


Figure S9. Grid of example molecules of the representative sample of 48 bits. *Related to section Fingerprint statistics.* Structure encoded in bit is highlighted in red.

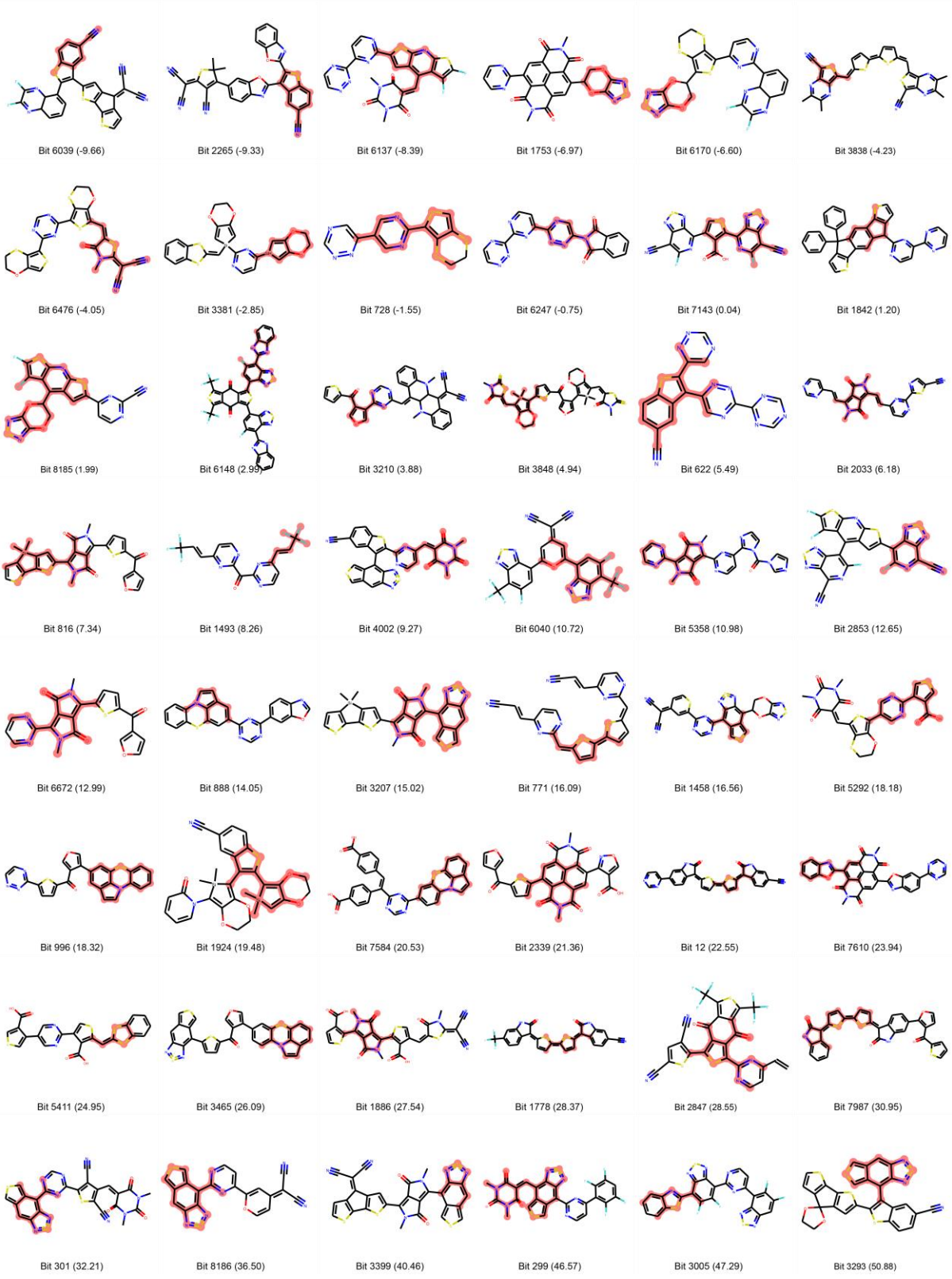
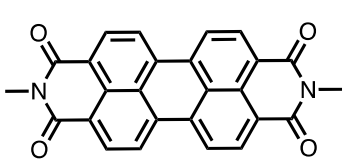

Bit 6039 (-9.66)
Bit 2265 (-9.33)
Bit 6137 (-8.39)
Bit 1753 (-6.97)
Bit 6170 (-6.60)
Bit 3838 (-4.23)
Bit 6476 (-4.05)
Bit 3381 (-2.85)
Bit 728 (-1.55)
Bit 6247 (-0.75)
Bit 7143 (0.04)
Bit 1842 (1.20)
Bit 8185 (1.99)
Bit 6148 (2.99)
Bit 3210 (3.88)
Bit 3848 (4.94)
Bit 622 (5.49)
Bit 2033 (6.18)
Bit 816 (7.34)
Bit 1493 (8.26)
Bit 4002 (9.27)
Bit 6040 (10.72)
Bit 5358 (10.98)
Bit 2853 (12.65)
Bit 6672 (12.99)
Bit 888 (14.05)
Bit 3207 (15.02)
Bit 771 (16.09)
Bit 1458 (16.56)
Bit 5292 (18.18)
Bit 996 (18.32)
Bit 1924 (19.48)
Bit 7584 (20.53)
Bit 2339 (21.36)
Bit 12 (22.55)
Bit 7610 (23.94)
Bit 5411 (24.95)
Bit 3465 (26.09)
Bit 1886 (27.54)
Bit 1778 (28.37)
Bit 2847 (28.55)
Bit 7987 (30.95)
Bit 301 (32.21)
Bit 8186 (36.50)
Bit 3399 (40.46)
Bit 299 (46.57)
Bit 3005 (47.29)
Bit 3293 (50.88)

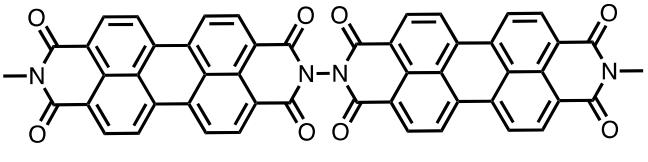
Table S1. Summary table of references values utilized in the calibration set. *Related to section Molecular generation and DFT calculations.* Experimental and computed frontier molecular orbitals used for training in the GP calibration. Includes source for reference.

		Experimental		Calculated		Calibrated	
SI Code	**Reference**	**HOMO (eV)**	**LUMO (eV)**	**HOMO (eV)**	**LUMO (eV)**	**HOMO (eV)**	**LUMO (eV)**
1	3	-5.48	-3.84	-5.50	-3.05	-5.57	-3.76
2	4	-5.71	-3.71	-5.65	-3.06	-5.74	-3.64
3	5	-5.94	-3.84	-5.71	-3.30	-5.85	-3.86
4	6	-5.81	-3.61	-5.76	-3.07	-5.83	-3.59
5	7	-5.26	-3.52	-5.27	-3.22	-5.37	-3.57
6	8	-5.90	-4.09	-5.71	-3.63	-5.84	-4.11
7	9	-6.12	-3.43	-6.24	-3.18	-6.03	-3.394
8	10	-5.57	-3.95	-5.32	-3.61	-5.44	-3.90
10	11	-6.02	-4.01	-6.07	-3.61	-5.95	-3.92
11	12	-5.65	-3.84	-6.04	-3.86	-5.8	-4.03
14	13	-6.04	-3.77	-6.25	-3.91	-5.98	-3.93
19	14	-5.71	-3.71	-5.35	-3.55	-5.65	-3.76
20	15	-5.90	-4.10	-6.00	-3.35	-5.95	-4.01
21	9	-6.27	-3.44	-6.17	-3.21	-6.28	-3.53
22	16	-5.80	-3.60	-6.11	-3.59	-5.97	-3.70
23	17	-6.00	-3.60	-6.09	-3.34	-5.99	-3.53
24	16	-5.80	-3.80	-5.59	-3.67	-5.55	-3.77
27	18	-5.48	-3.83	-5.66	-3.61	-5.59	-3.89
28	6	-5.81	-3.61	-6.53	-3.67	-6.12	-3.77
30	19	-5.18	-3.24	-5.03	-2.63	-5.19	-3.17
31	20	-5.69	-3.71	-5.41	-3.08	-5.61	-3.65
32	21	-5.53	-4.05	-5.89	-3.45	-5.89	-3.98
33	22	-5.88	-3.80	-6.05	-3.58	-5.97	-3.79
34	22	-5.46	-3.78	-5.25	-3.51	-5.38	-3.94
35	23	-6.74	-4.35	-6.72	-3.93	-6.60	-4.21
48	25	-6.28	-4.23	-6.17	-3.18	-6.27	-4.07
50	14	-6.40	-4.10	-6.17	-3.18	-6.38	-3.95

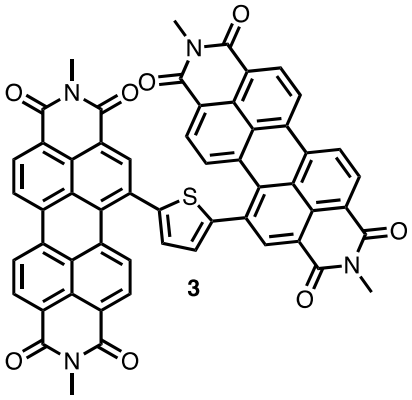
Table S2. Structures of *n*-type materials in table S1. Related to section *Molecular generation and DFT calculations.*



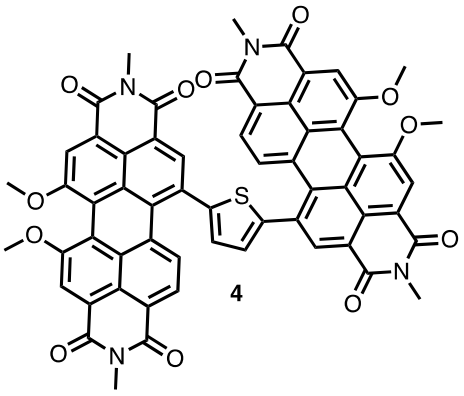
1



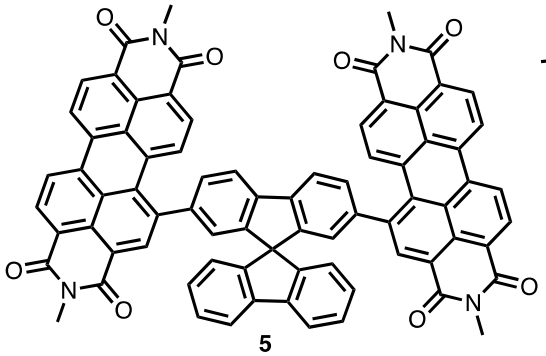
2



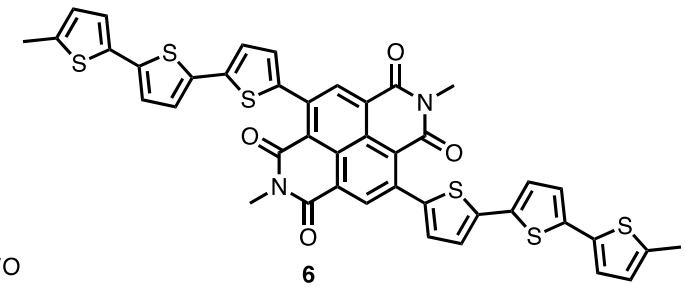
3



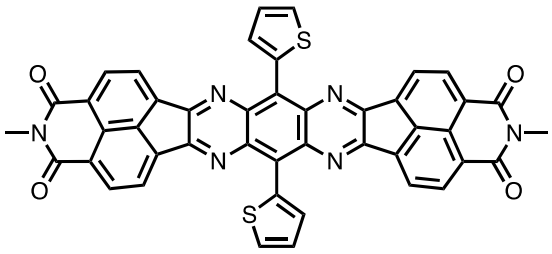
4



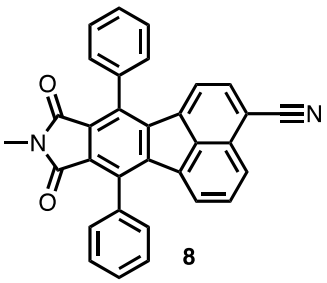
5



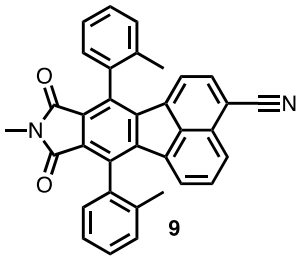
6



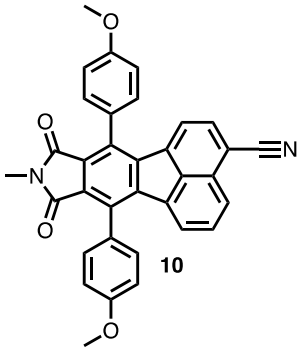
7



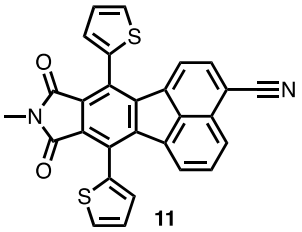
8



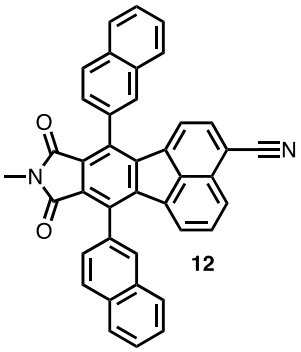
9



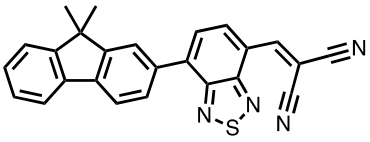
10



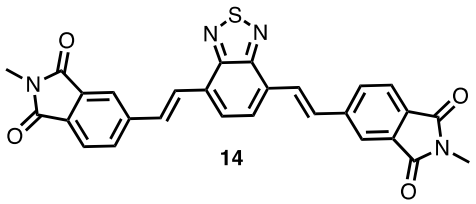
11



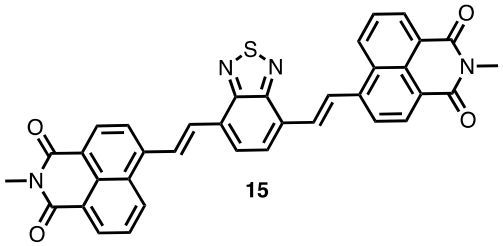
12



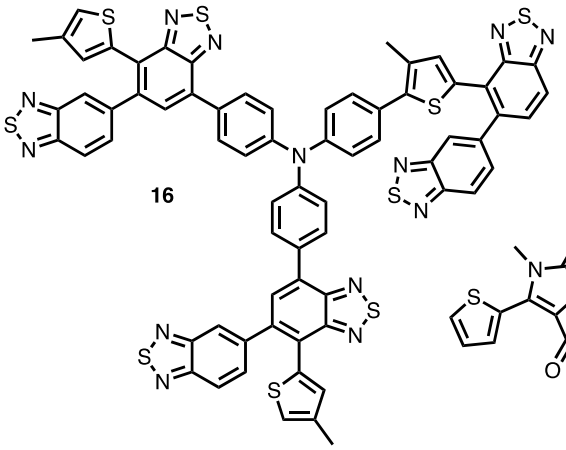
13



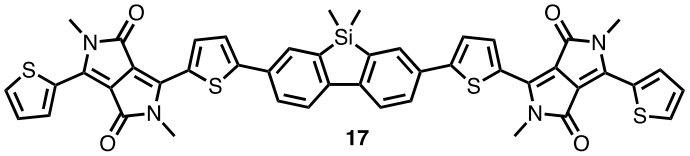
14



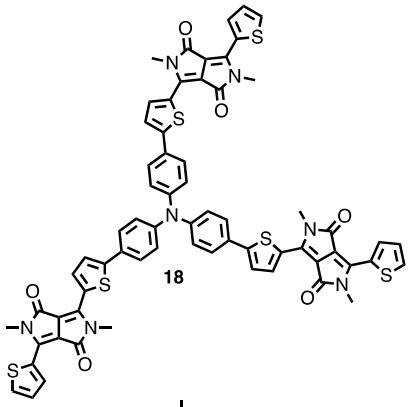
15



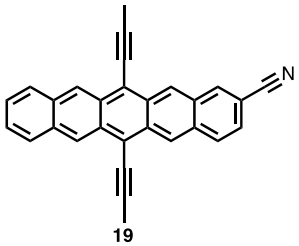
16



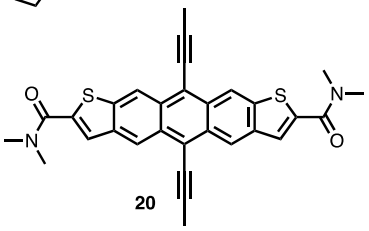
17



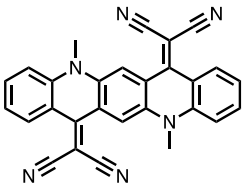
18



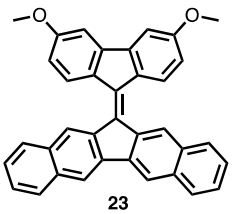
19



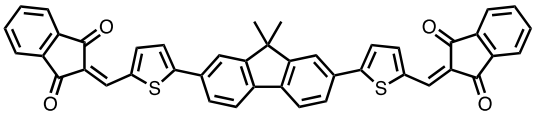
20



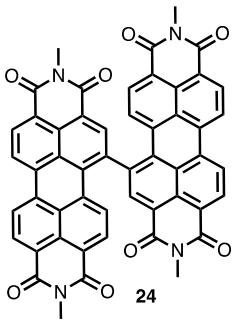
21



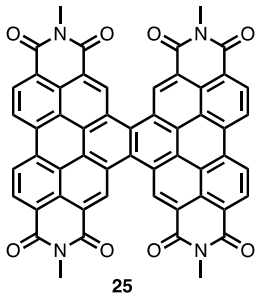
23



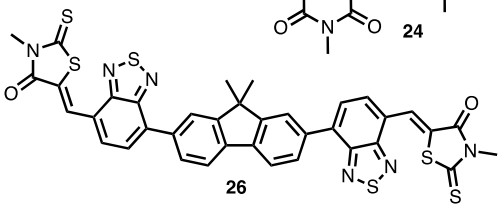
22



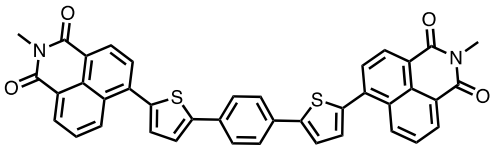
24



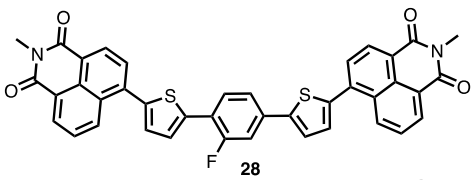
25



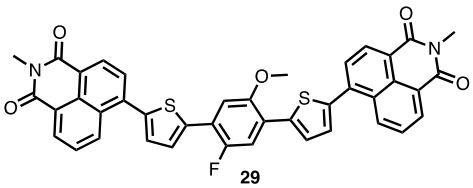
26



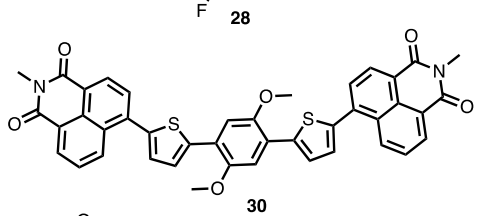
27



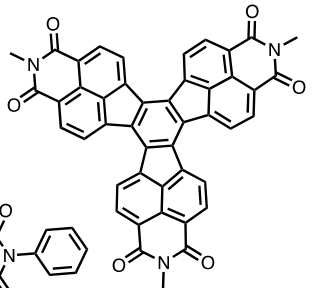
28



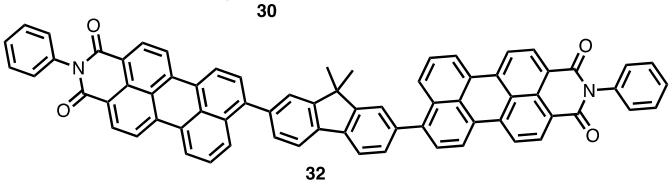
29



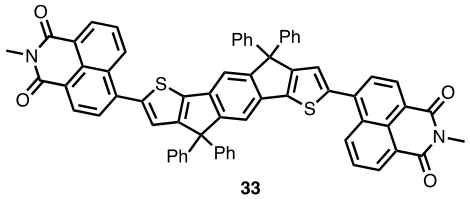
30



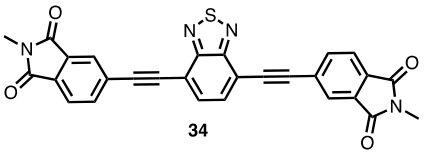
31



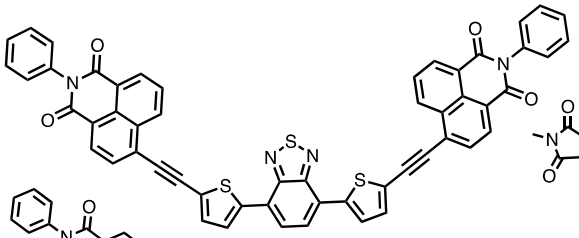
32



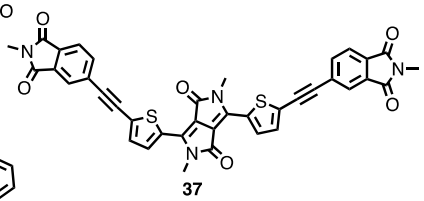
33



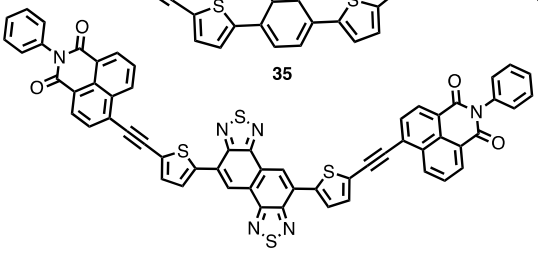
34



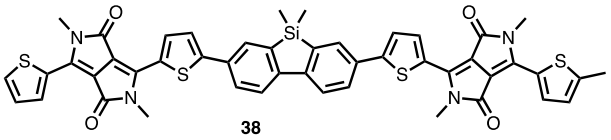
35



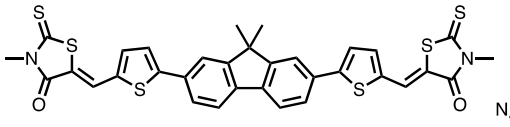
37



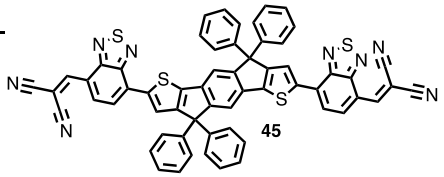
36



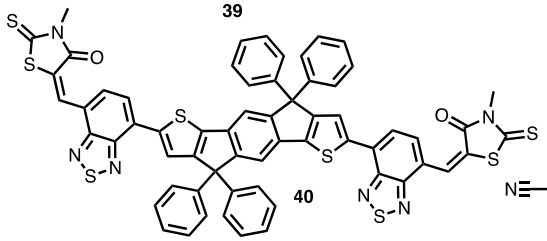
38



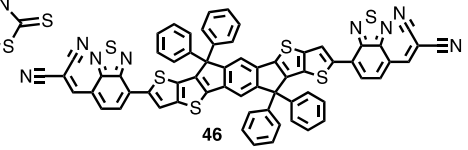
39



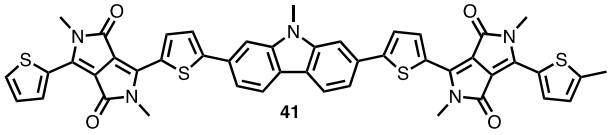
45



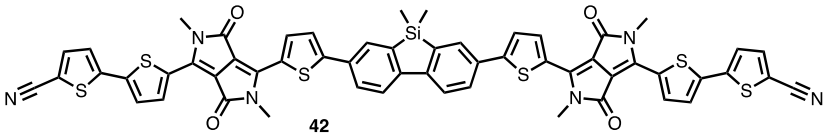
40



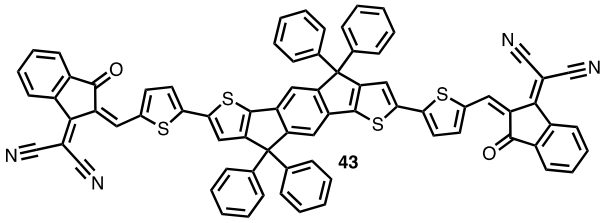
46



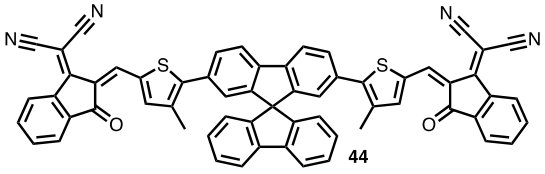
41



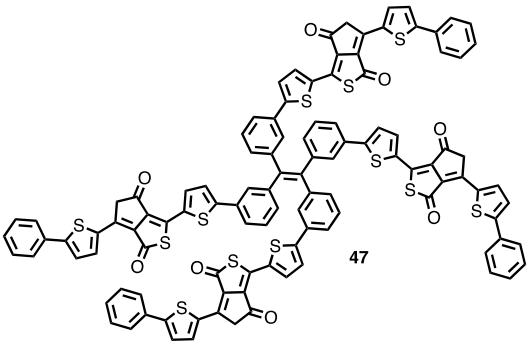
42



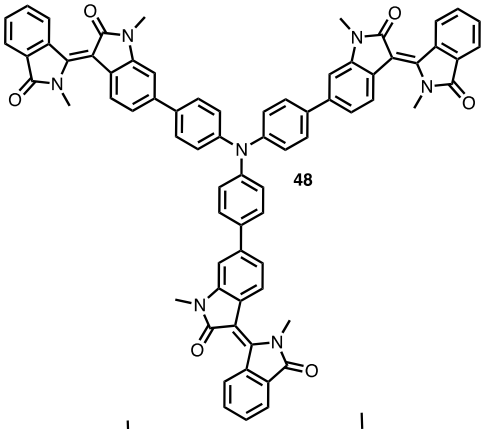
43



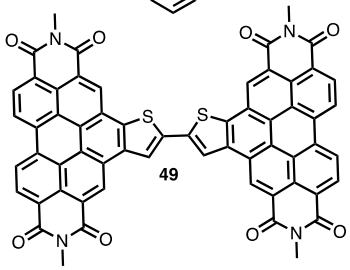
44



47



48



49

Supplementary References

[1] Lin, Y.; Zhan, X. “Non-Fullerene Acceptors for Organic Photovoltaics” *Mater. Horiz*. **2014**, *1*, 470–488.

[2] Zhan, C.; Zhang, X.; Yao, J. “New Advances in Non-Fullerene Acceptor Based Organic Solar Cells” *RSC. Adv.* **2015**, *5,* 93002-93026.

[3] Jiang, B.; Zhang, X.; Zhan, C.; Lu, Z.; Huang, J.; Ding, X.-L.; He, S.-G. ; Yao, J. “Benzodithiophene bridged dimeric perylene diimide amphiphiles as efficient solution-processed non-fullerene small molecules” *Polym. Chem*. **2013**, *4*, 4631-1638.

[4] Yan, Q.; Zhou, Y.; Zheng, Y.-Q. ; Pei, J.; Zhao, D. “Towards rational design of organic electron acceptors for photovoltaics: a study based on perylenediimide derivatives” *Chem. Sci.* **2013**, *4*, 4389-4394.

[5] Jiang, W.; Ye, L.; Li, X.; Xiao, C.; Tan, F.; Zhao, W.; Hou, J.; Wang, Z. “Bay-linked perylene bisimides as promising non-fullerene acceptors for organic solar cells” *Chem. Commun*. **2014**, *50*, 1024-1026.

[6] Pho, T. V.; Toma, F. M.; Chabinyc, M. L.; Wudl, F. “Self-Assembling Decacyclene Triimides Prepared through a Regioselective Hextuple Friedel–Crafts Carbamylation” *Angew. Chem.*, Int. Ed., **2013**, *52*, 1446-1451.

[7] Sonar, P.; Ng, G.-M.; Lin, T. T.; Dodabalapur, A.; Chen, Z.-K. “Solution processable low bandgap diketopyrrolopyrrole (DPP) based derivatives: novel acceptors for organic solar cells” *J. Mater. Chem*., **2010**, *20*, 3626-3636.

[8] Karsten, B. P.; Bijleveld, J. C.; Janssen, R. A. J. “Diketopyrrolopyrroles as acceptor materials in organic photovoltaics”*Macromol. Rapid Commun*., **2010**, *31*, 1554-1559.

[9] Zhou, Y.; Dai, Y.-Z.; Zheng, Y.-Q.; Wang, X.-Y.; Wang, J.-Y.; Pei, J. “Non-fullerene acceptors containing fluoranthene-fused imides for solution-processed inverted organic solar cells” *Chem. Commun.*, **2013**, *49*, 5802-5804.
[10] Zhang, X. L.; Jiang, B.; Zhang, X.; Huang, J. H.; Zhan, C. L.; Yao, J. N. “Cooperatively Tuning Phase Size and Absorption of Near IR Photons in P3HT: Perylene Diimide Solar Cells by Bay-Modifications on the Acceptor” *J. Phys. Chem. C*, **2014**, *118*, 24212–24220.
[11] Hartnett, P. E.; Timalsina, A.; Matte, H. S. S. R.; Zhou, N.; Guo, X.; Zhao, W.; Facchetti, A.; Chang, R. P. H.; Hersam, M. C.; Wasielewski, M. R.; Marks, T. J. “SlipStacked Perylenediimides as an Alternative Strategy for High Efficiency Nonfullerene Acceptors in Organic Photovoltaics” *J. Am. Chem. Soc.*, **2014**, *136*, 16345–16356.
[12] Lu, Z. H.; Jiang, B.; Zhang, X.; Tang, A. L.; Chen, L. L.; Zhan, C. L.; Yao, J. N. “Perylene-Diimide Based NonFullerene Solar Cells with 4.34% Efficiency through Engineering Surface D/A Compositions” *Chem. Mater.*, **2014**, *26*, 2907–2914.
[13] Zhong, Y.; Trinh, M. T.; Chen, R.; Wang, W.; Khlyabich, P. P.; Kumar, B.; Xu, Q.; Nam, C.-Y.; Sfeir, M. Y.; Black, C.; Steigerwald, M. L.; Loo, Y.-L.; Xiao, S.; Ng, F.; Zhu, X.-Y.; Nuckolls, C. “Efficient Organic Solar Cells with Helical Perylene Diimide Electron Acceptors” *J. Am. Chem. Soc.*, **2014**, *136*, 15215–15221.
[14] Nielsen, C. B.; Voroshazi, E.; Holliday, S.; Cnops, K.; Cheyns, D.; McCulloch, I. “Electron-deficient truxenone derivatives and their use in organic photovoltaics” *J. Mater. Chem. A*, **2014**, *2*, 12348–12354.
[15] Zhou, T.; Jia, T.; Kang, B.; Li, F.; Fahlman, M.; Wang, Y.; “Nitrile-Substituted QA Derivatives: New Acceptor Materials for Solution-Processable Organic Bulk Heterojunction Solar Cells” *Adv. Energy Mater.*, **2011**, *1*, 431–439.
[16] Li, H.; Earmme, T.; Ren, G.; Saeki, A.; Yoshikawa, S.; Murari, N. M.; Subramaniyan, S.; Crane, M. J.; Seki, S.; Jenekhe, S. A. “Beyond Fullerenes: Design of Nonfullerene Acceptors for Efficient Organic Photovoltaics” *J. Am. Chem. Soc.*, **2014**, *136*, 14589–14597.
[17] Li, H.; Earmme, T.; Subramaniyan, S.; Jenekhe, S. A. “Bis(Naphthalene Imide)diphenylanthrazolines: A New Class of Electron Acceptors for Efficient Nonfullerene Organic Solar Cells and Applicable to Multiple Donor Polymers” *Adv. Energy Mater.*, **2015**, 1402041.
[18] Lin, Y.; Wang, J.; Zhang, Z.-G.; Bai, H.; Li, Y.; Zhu, D.; Zhan, X. “An Electron Acceptor Challenging Fullerenes for Efficient Polymer Solar Cells” *Adv. Mater.*, **2015**, *27*, 1170–1174.
[19] Brunetti, F. G.; Gong, X.; Tong, M.; Heeger, A. J.; Wudl, F. “Strain and Huckel Aromaticity: Driving Forces for ¨ a Promising New Generation of Electron Acceptors in Organic Electronics” *Angew. Chem.*, Int. Ed., **2010**, *49*, 532– 536.
[20] Mao, Z.; Le, T. P.; Vakhshouri, K.; Fernando, R.; Ruan, F.; Muller, E.; Gomez, E. D.; Sauv´e, G. “Processing additive suppresses phase separation in the active layer of organic photovoltaics based on naphthalene diimide” *Org. Electron.*, **2014**, *15*, 3384–3391.
[21] Ahmed, E.; Ren, G.; Kim, F. S.; Hollenbeck, E. C.; Jenekhe, S. A. “Design of New Electron Acceptor Materials for Organic Photovoltaics: Synthesis, Electron Transport, Photophysics, and Photovoltaic Properties of Oligothiophene-Functionalized Naphthalene Diimides” *Chem. Mater.*, **2011**, *23*, 4563–4577.
[22] Wang, X.; Huang, J. H.; Niu, Z. X.; Sun, Y. X.; Zhan, C. L. “Dimeric Naphthalene Diimide Based Small Molecule Acceptors: Synthesis, Characterization, and Photovoltaic Properties” *Tetrahedron*, **2014**, *70*, 4726–4731.
[23] Liu, Y.; Zhang, L.; Lee, H.; Wang, H.-W. ; Santala, A.; Liu, F.; Diao, Y.; Briseno, A. L.; Russell, T. P. “NDI-Based Small Molecule as Promising Nonfullerene Acceptor for Solution-Processed Organic Photovoltaics” *Adv. Energy Mater.*, **2015**, 1500195.
[24] Morse, G. E.; Gantz, J. L.; Steirer, K. X. ; Armstrong, N. R.; Bender, T. P. “Pentafluorophenoxy Boron Subphthalocyanine (F5BsubPc) as a Multifunctional Material for Organic Photovoltaics” *ACS Appl. Mater. Interfaces*, **2014**, *6*, 1515–1524.
[25] Sullivan, P.; Collis, G. E.; Rochford, L. A.; Arantes, J. F.; Kemppinen, P.; Jones, T. S.; Winzenberg, K. N. “An N-ethylated barbituric acid end-capped bithiophene as an electron-acceptor material in fullerene-free organic photovoltaics” *Chem. Commun.*, **2015**, *51*, 6222– 6225.
[26] Bartynski, A. N.; Gruber, M. ; Das, S.; Rangan, S.; Mollinger, S.; Trinh, C.; Bradforth, S. E.; Vandewal, K.; Salleo, A.; Bartynski, R. A.; Bruetting, W.; Thompson, M. E. “Symmetry-Breaking Charge Transfer in a Zinc Chlorodipyrrin Acceptor for High Open Circuit Voltage Organic Photovoltaics” *J. Am. Chem. Soc.*, **2015**, *137*, 5397– 5405.