SUPPORTING INFORMATION

Tuning the Structural and Spectroscopic Properties of Donor-Acceptor-Donor Oligomers via Mutual X-Bonding, H-Bonding, and π-π Interactions

Asmerom O. Weldeab[a], April Steen[b], Daken J. Starkenburg[c], Jon Steven Dal Williams[b], Khalil A. Abboud[a], Jiangeng Xue[c], Nathan I. Hammer[b], Ronald K. Castellano[a]*, and Davita L. Watkins[b]*

[a]Department of Chemistry, University of Florida, Gainesville, FL 32611, USA

[b]Department of Chemistry and Biochemistry, University of Mississippi, University, MS 38677, USA

[c]Department of Materials Science and Engineering, University of Florida, Gainesville, FL 32611, USA

Table of Contents Pages

General Information

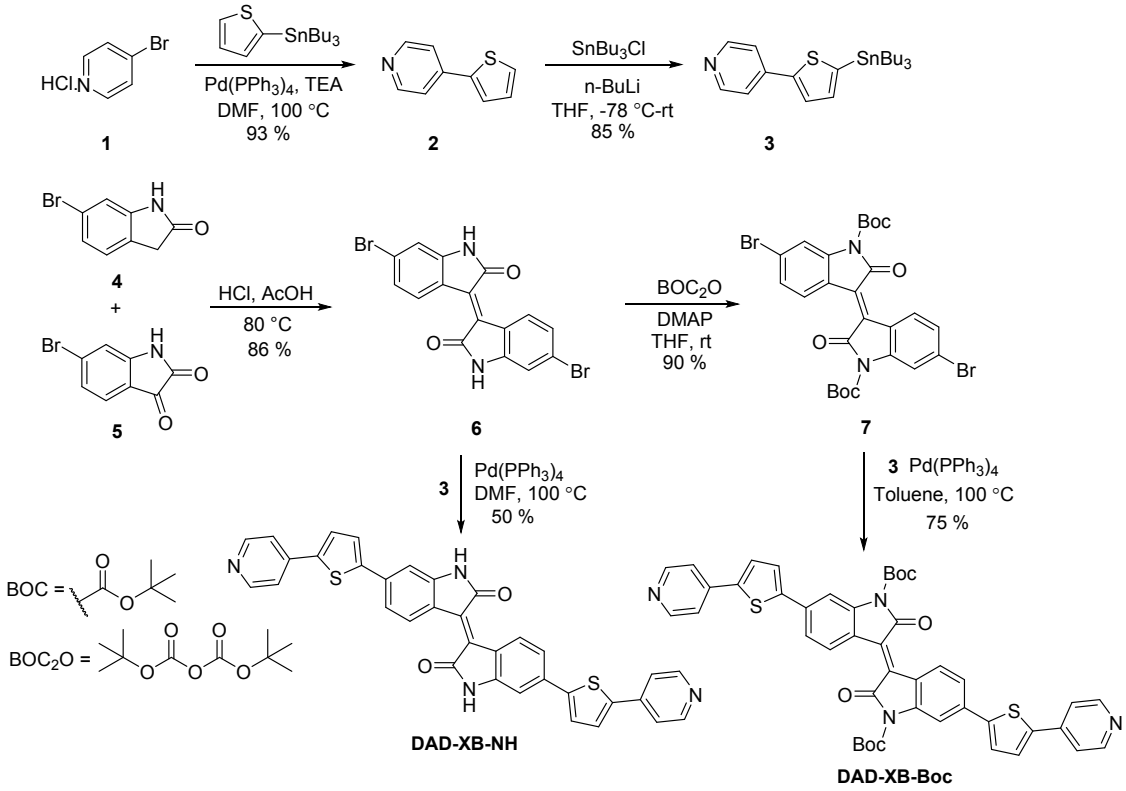
Reagents and solvents were purchased from commercial sources and used without further purification unless otherwise specified. THF and DMF were degassed in 20 L drums and passed through two sequential purification columns (activated alumina; molecular sieves for DMF) under a positive argon atmosphere. All synthetic procedures were carried out under argon atmosphere using standard schlenk line techniques unless otherwise stated. Thin layer chromatography (TLC) was performed on SiO_2-60 F_{254} aluminum plates with visualization by UV light. Flash column chromatography was performed using Silica gel technical grade, pore size 60 Å, 230−400 mesh particle size, 40−63 μm particle size from Sigma-Aldrich. ^{1}H (^{13}C) NMR were recorded on Mercury 300 or INOVA 500 spectrometer. Chemical shifts (δ) are given in parts per million (ppm) relative to TMS and referenced to residual protonated solvent purchased from Cambridge Isotope Laboratories, Inc. ($CDCl_3$: δH 7.26 ppm, δC 77.16 ppm; DMSO-d_6: δH 2.50 ppm, δC 39.52 ppm). Abbreviations used are s (singlet), d (doublet), t (triplet), q (quartet), quin (quintet), hp (heptet), b (broad), and m (multiplet). Electrospray ionization (ESI) high resolution mass spectra (HRMS) were recorded on an Agilent 6210 TOF spectrometer with MassHunter software. Absorption spectra were recorded on a UV-Vis-near-IR spectrophotometer. Emission spectra were recorded on a Photon Technology International (PTI) fluorimeter and collected 90^{o} relative to the excitation beam. Crystal (solid state) diffuse reflectance spectra were obtained using an integrating sphere, a Xenon lamp, and an Ocean Optics USB2000 spectrometer. Similarly crystal fluorescence emission spectra and excited state lifetimes were obtained using a Nikon TE2000U inverted microscope and CCD detection with a pulsed picosecond 405nm diode laser. TGA measurements were performed using ~2 mg of sample on TA Instruments TGA Q5000- 0121 (platinum pan, room temperature to 600 ^{o}C, ramp rate = 20 ^{o}C min^{-1} under nitrogen atmosphere) and analyzed on Universal Analysis

2000 4.4A software. Solid-state IR is recorded using a Perkin Elmer Spectrum one FT-IR spectrophotometer. Solid samples were dried under vacuum prior to measurement. All measurements were taken using Opus 7.0 software.

The structural geometry of the halogen bond acceptor (**DAD-XB-Boc**) was optimized at the B3LYP/6-31G* level of theory (as implemented in Gaussian 09), accessed through the University of Florida High-Performance Computing Center. Frequency calculations were performed at the same computational level, and no imaginary frequencies were found. Molecular orbital plots were made using Visual Molecular Dynamics (VMD) software from the Gaussian output files.

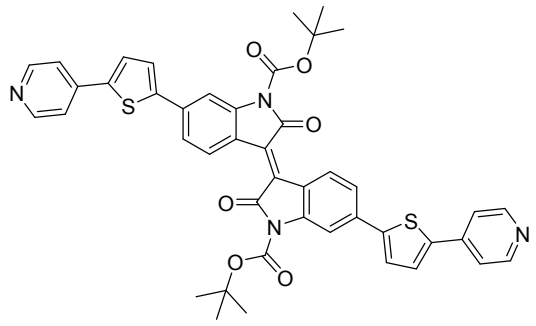
Synthesis Procedures

Intermediate compounds **2**, **3**, **6,** and **7** were prepared in accordance with literature procedures.[1-4] The ^{1}H NMR of all the intermediates agreed well with those available in the literature.



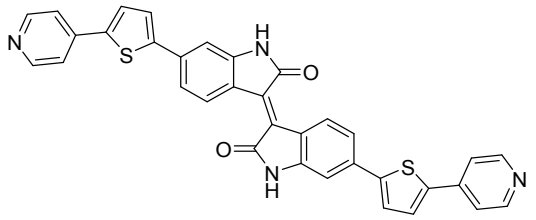
Scheme S1: Synthesis of **DAD-XB-Boc** and **DAD-XB-NH**

Di-tert-butyl (*E*)-2,2'-dioxo-6,6'-bis(5-(pyridin-4-yl)thiophen-2-yl)-[3,3'-biindolinylidene]-1,1'-dicarboxylate (DAD-XB-Boc):



Under argon, anhydrous toluene (7 mL) was added to a two-necked round-bottom flask containing **7** (0.10 g, 0.16 mmol), **3** (0.21 g, 0.48 mmol) and tetrakis (triphenylphosphine) palladium (0) (0.040 g, 0.032 mmol). The mixture was heated to 100°C for 24 h. The reaction mixture was cooled and poured into cold acetone. The precipitate obtained was filtered, washed with acetone and dried under vacuum. The solid was re-dissolved in chloroform, filtered to get rid of any insoluble particles and evaporated to give the pure product (0.090 g, 75 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.03 (d, J = 8.8 Hz, 1H), 8.65 (d, J = 7.8 Hz, 2H), 8.21 (s, 1H), 7.56 (d, 1H), 7.55 – 7.51 (m, 3H), 7.49 (d, 1H), 1.77 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.01, 150.53, 148.68, 145.46, 141.88, 141.83, 140.84, 137.77, 129.76, 128.44, 126.65, 126.05, 121.74, 121.09, 119.60, 111.21, 85.12, 28.25. DART-HRMS-ESI: *m/z* $[M+H]^+$ calcd: 781.2149; found: 781.2126.

Di-tert-butyl (*E*)-2,2'-dioxo-6,6'-bis(5-(pyridin-4-yl)thiophen-2-yl)-[3,3'-biindolinylidene]-1,1'-dicarboxylate (DAD-XB-NH):



Under argon, anhydrous DMF (15 mL) was added to a two-necked round-bottom flask containing **6** (0.10 g, 0.24 mmol), **3** (0.33 g, 0.72 mmol) and tetrakis (triphenylphosphine) palladium (0) (0.06 g, 0.048 mmol). The mixture was heated to 100°C for 24 h. The reaction mixture was cooled to room temperature and the precipitate obtained was filtered. The solid obtained was washed by refluxing in acetone and then chloroform: methanol = 1:1 to give the desired product (0.06 g, 50%). Attempts to further purify the compound was not successful due to the very poor solubility of the compound.

^{1}H NMR (500 MHz, DMSO-d_6): δ 11.05 (s, 1H), 9.14 (d, J = 8.7 Hz, 1H), 8.61 (d, J = 5.0 Hz, 2H), 7.90 (d, J = 3.7 Hz, 1H), 7.77 (d, J = 3.4 Hz, 1H), 7.72 (d, J = 4.7 Hz, 2H), 7.42 (d, J = 8.1 Hz, 1H), 7.14 (s, 1H). DART-HRMS-ESI: m/z $[M+H]^+$ calcd: 581.1100; found: 581.1099. Attempts to obtain ^{13}C NMR was not successful due to the poor solubility of the compound.

S1: Coplanarity Comparison of DAD-XB-Boc in Solid State and Gas Phase

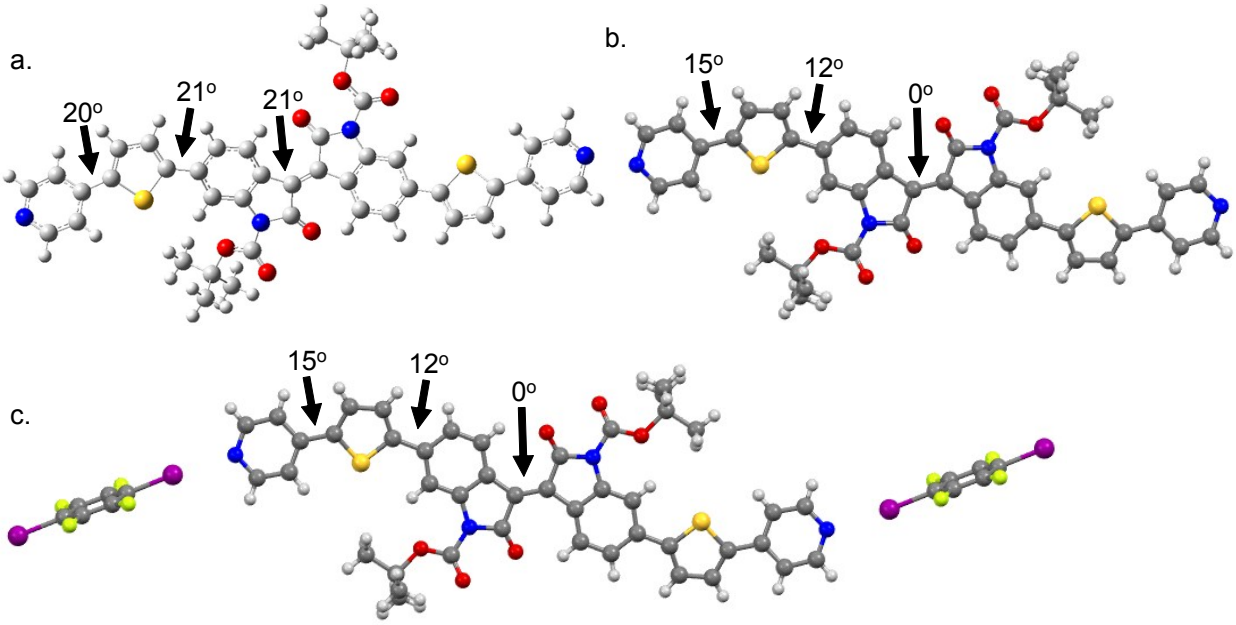


Figure S1: Comparison of the torsions of **DAD-XB-Boc** obtained from DFT calculation (a), mono-crystal (b), and co-crystal (c)

S2: UV-Vis General Information

Solution absorption measurements were taken on a Cary 100 Bio UV-Visible spectrophotometer controlled by Cary Win UV software and equipped with a Peltier 1 × 1 Cell Holder using 1 cm quartz cells. Spectroscopic grade DMF was purchased from Sigma-Aldrich. Solution emission spectra were recorded on a Photon Technology International (PTI) fluorimeter and collected 90° relative to the excitation beam.

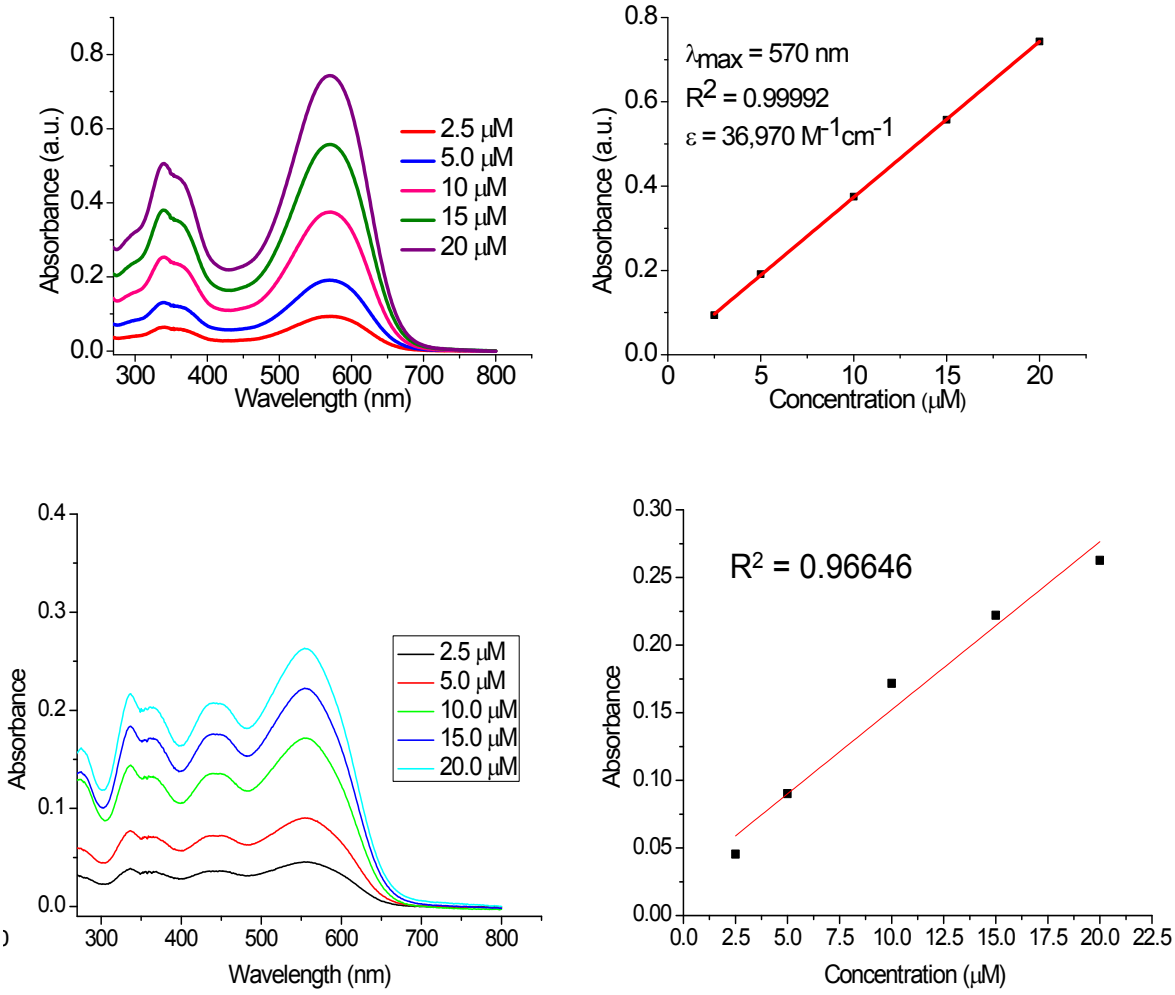


Figure S2: UV-Vis spectra (top, left) and Beer-Lambert plot (top, right) of **DAD-XB-Boc** in DMF; UV-Vis spectra (bottom, left) and Beer-Lambert plot (bottom, right) of **DAD-XB-NH** in DMSO. The distinctive non-linear plot indicates aggregation and poor solubility of **DAD-XB-NH.**

X-ray Crystallography

DAD-XB-Boc mono-crystal X-ray experimental: X-ray intensity data were collected at 100 K on a Bruker DUO diffractometer using MoKα radiation ($\lambda = 0.71073$ Å) and an APEXII CCD area detector.

Raw data frames were read by the program SAINT[1] and integrated using 3D profiling algorithms. The resulting data were reduced to produce hkl reflections and their intensities and

estimated standard deviations. The data were corrected for Lorentz and polarization effects and numerical absorption corrections were applied based on indexed and measured faces.

The structure was solved and refined in *SHELXTL2014*, using full-matrix least-squares refinement. The non-H atoms were refined with anisotropic thermal parameters and all of the H atoms were calculated in idealized positions and refined riding on their parent atoms. The asymmetric unit consists of a half molecule and a chloroform solvent molecule. There exist a Hydrogen bond between the chloroform proton and the N2 of the pyridyl group. In the final cycle of refinement, 5126 reflections (of which 3812 are observed with $I > 2\sigma(I)$) were used to refine 289 parameters and the resulting R_1, wR_2 and S (goodness of fit) were 3.61%, 8.62% and 0.987, respectively. The refinement was carried out by minimizing the wR_2 function using F^2 rather than F values. R_1 is calculated to provide a reference to the conventional R value but its function is not minimized. *SHELXTL2014* (2014). Bruker-AXS, Madison, Wisconsin, USA.

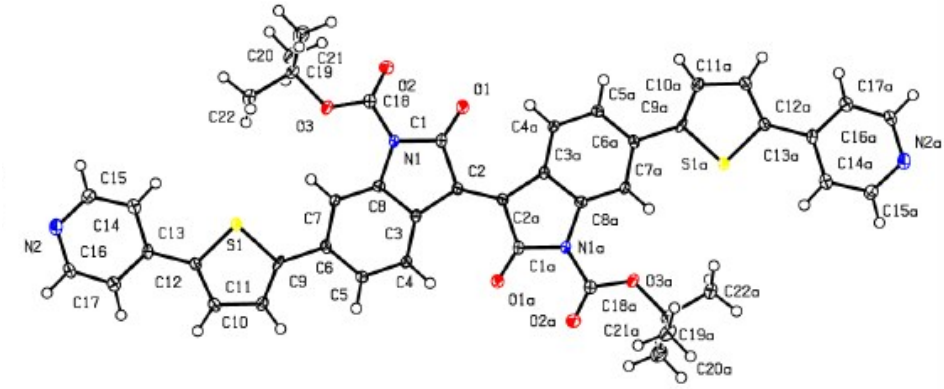


Table 1. Crystal data and structure refinement for **DAD-XB-Boc**.

Identification code	asme1
Empirical formula	C46 H38 Cl6 N4 O6 S2
Formula weight	1019.62
Temperature	100(2) K
Wavelength	0.71073 Å

Crystal system	Triclinic	
Space group	P-1	
Unit cell dimensions	a = 8.2494(3) Å	α= 87.1718(10)°.
	b = 9.3069(4) Å	β= 78.7559(10)°.
	c = 15.1499(7) Å	γ = 77.6043(9)°.
Volume	1114.20(8) $Å^3$	
Z	1	
Density (calculated)	1.520 Mg/m^3	
Absorption coefficient	0.535 mm^{-1}	
F(000)	524	
Crystal size	0.381 x 0.027 x 0.026 mm^3	
Theta range for data collection	1.370 to 27.499°.	
Index ranges	-10≤h≤10, -12≤k≤11, -19≤l≤19	
Reflections collected	16994	
Independent reflections	5126 [R(int) = 0.0354]	
Completeness to theta = 25.242°	100.0 %	
Absorption correction	Integration	
Max. and min. transmission	0.9902 and 0.9073	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	5126 / 0 / 289	
Goodness-of-fit on F^2	0.987	
Final R indices [I>2sigma(I)]	R1 = 0.0361, wR2 = 0.0862 [3812]	
R indices (all data)	R1 = 0.0584, wR2 = 0.1010	
Extinction coefficient	n/a	
Largest diff. peak and hole	0.581 and -0.262 $e.Å^{-3}$	

$R1 = \sum(||F_o| - |F_c||) / \sum|F_o|$ $wR2 = [\sum[w(F_o^2 - F_c^2)^2] / \sum[w(F_o^2)^2]]^{1/2}$

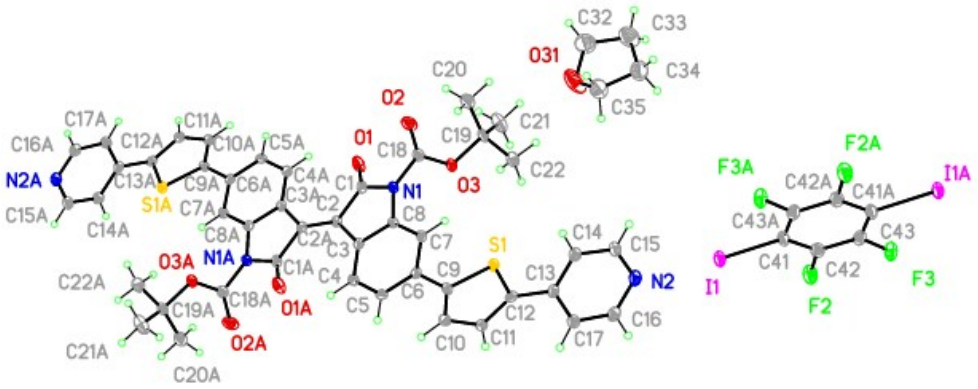
$S = [\sum[w(F_o^2 - F_c^2)^2] / (n-p)]^{1/2}$ $w = 1/[\sigma^2(F_o^2)+(m*p)^2+n*p]$, $p = [\max(F_o^2,0)+ 2* F_c^2]/3$, m & n are constants.

DAD-XB-Boc-TFDIB co-crystal X-ray experimental: X-ray intensity data were collected at 100 K on a Bruker DUO diffractometer using MoKα radiation (λ = 0.71073 Å) and an APEXII CCD area detector.

Raw data frames were read by program the SAINT[1] and integrated using 3D profiling algorithms. The resulting data were reduced to produce hkl reflections and their intensities and estimated standard deviations. The data were corrected for Lorentz and polarization effects and numerical absorption corrections were applied based on indexed and measured faces.

The structure was solved and refined in *SHELXTL2014*, using full-matrix least-squares refinement. The non-H atoms were refined with anisotropic thermal parameters and all of the H atoms were calculated in idealized positions and refined riding on their parent atoms. In the final cycle of refinement, 6157 reflections (of which 5049 are observed with $I > 2\sigma(I)$) were used to refine 352 parameters and the resulting R_1, wR_2 and S (goodness of fit) were 2.51%, 4.99% and 0.966, respectively. The refinement was carried out by minimizing the wR_2 function using F^2 rather than F values. R_1 is calculated to provide a reference to the conventional R value but its function is not minimized.

SHELXTL2014 (2014). Bruker-AXS, Madison, Wisconsin, USA.



A representation of the full molecular fragments (thermal ellipsoids = 40%).

Table 1. Crystal data and structure refinement for **DAD-XB-Boc-TFDIB**.

Identification code	asme3
Empirical formula	C58 H52 F4 I N4 O8 S2

Formula weight	663.48	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Triclinic	
Space group	Pī	
Unit cell dimensions	a = 8.1917(4) Å	α= 100.6726(9)°.
	b = 9.5688(5) Å	β= 91.6388(9)°.
	c = 18.0523(10) Å	γ= 104.9351(9)°.
Volume	1339.22(12) $Å^3$	
Z	2	
Density (calculated)	1.645 Mg/m^3	
Absorption coefficient	1.327 mm^{-1}	
F(000)	666	
Crystal size	0.253 x 0.026 x 0.023 mm^3	
Theta range for data collection	2.304 to 54.998°.	
Index ranges	$-10\leq h\leq 10$, $-12\leq k\leq 12$, $-23\leq l\leq 23$	
Reflections collected	31002	
Independent reflections	6157 [R(int) = 0.0441]	
Completeness to theta = 25.242°	100.0 %	
Absorption correction	Analytical	
Max. and min. transmission	0.9780 and 0.8410	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	6157 / 291 / 352	
Goodness-of-fit on F^2	0.966	
Final R indices [I>2sigma(I)]	R1 = 0.0251, wR2 = 0.0499 [5049]	
R indices (all data)	R1 = 0.0385, wR2 = 0.0525	
Extinction coefficient	n/a	
Largest diff. peak and hole	0.532 and -0.384 $e.Å^{-3}$	

$R1 = \sum(||F_o| - |F_c||) / \sum|F_o|$

$wR2 = [\sum[w(F_o^2 - F_c^2)^2] / \sum[w(F_o^2)^2]]^{1/2}$

$S = [\sum[w(F_o^2 - F_c^2)^2] / (n-p)]^{1/2}$

$w= 1/[\sigma^2(F_o^2)+(m*p)^2+n*p]$, $p = [\max(F_o^2,0)+ 2* F_c^2]/3$, m & n are constants.

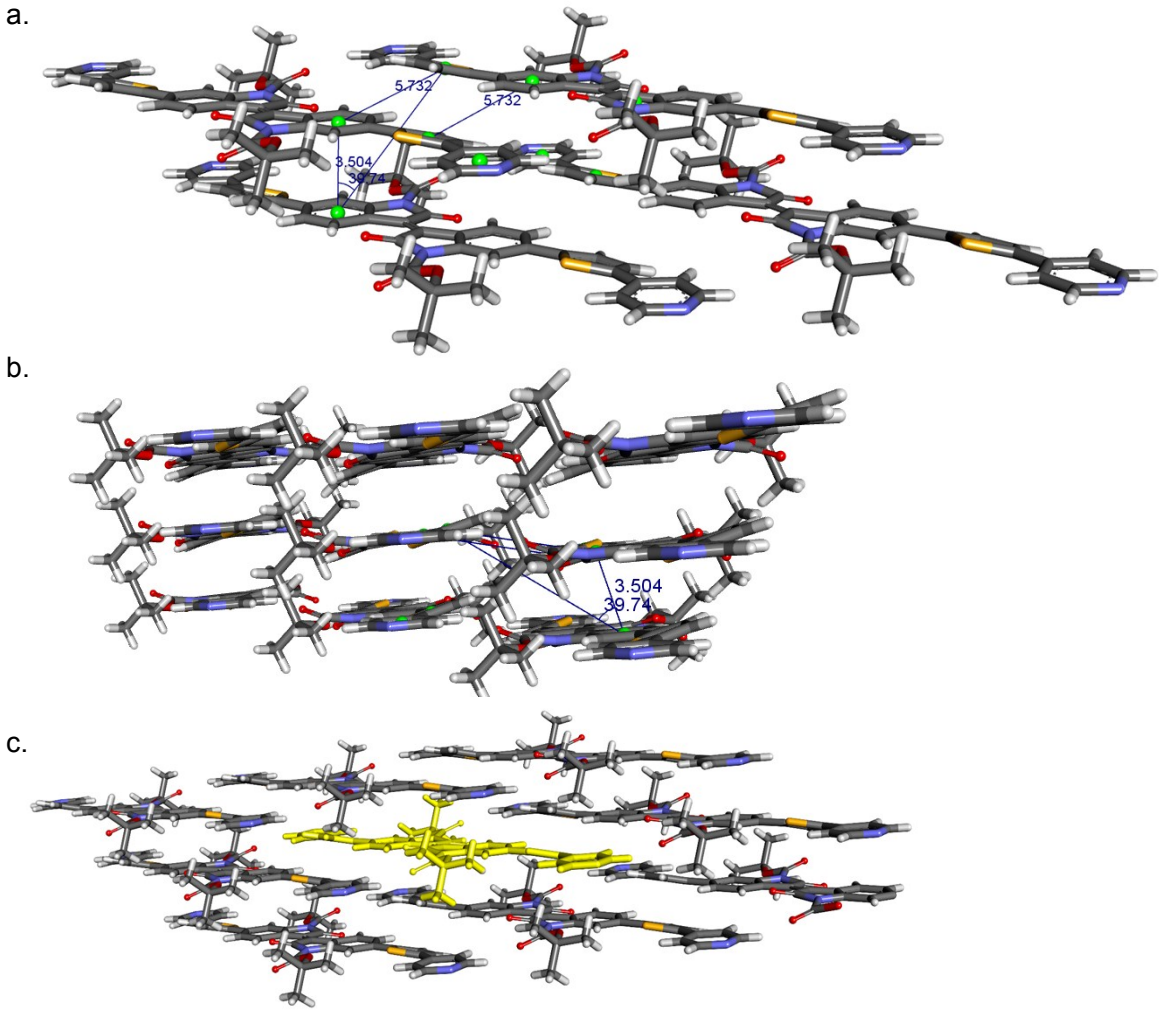


Figure S3: Crystal packing structure of **DAD-XB-Boc** viewed along the *b*-axis (a) and *a*-axis (b). Packing diagram (c) highlights nearest neighbor contacts. Distances and angles (blue) are included. Centroid (green spheres) of atoms were used to monitor distances. Images and characterization of crystals were carried out using CrystalMaker software (crystallographic information files) and Accelry's Discovery Studio software.

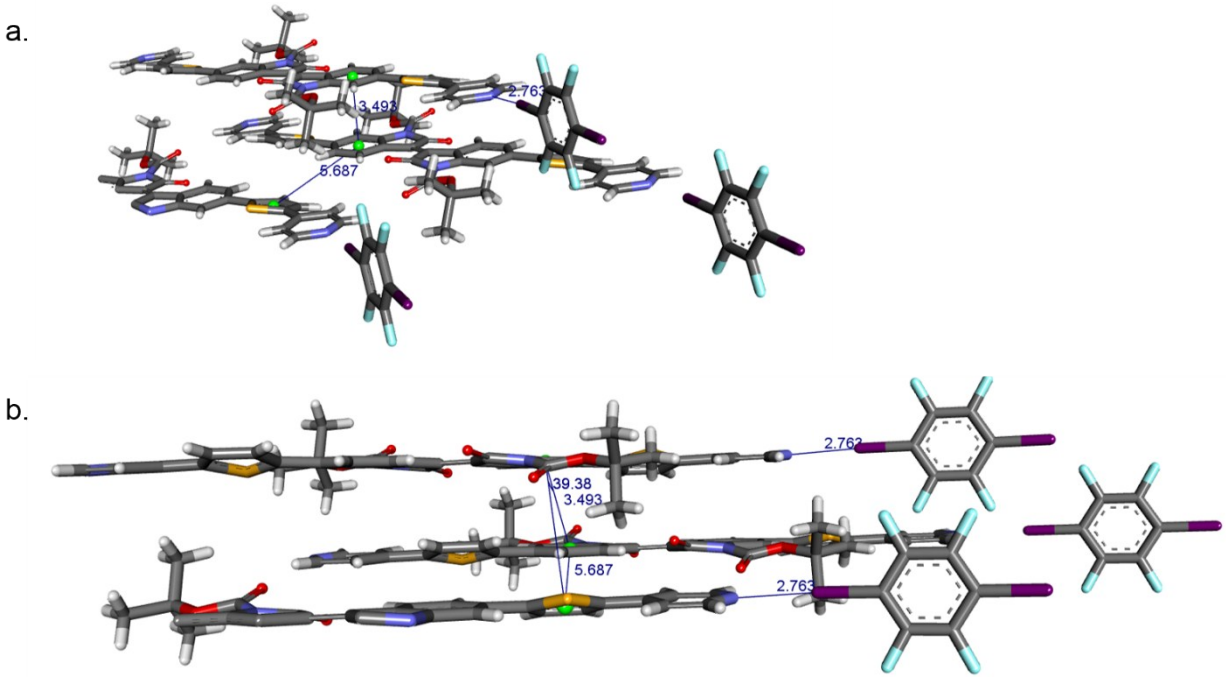


Figure S4: Crystal packing structure of **DAD-XB-Boc-TFDIB** viewed along the *a*-axis (a) and *b*-axis (b). Images highlight the nearest neighbor contacts for the XB acceptor, distances for π-stacking and XB, and angles (blue). Centroid (green spheres) of atoms were used to monitor distances. Images and characterization of crystals were carried out using CrystalMaker software (crystallographic information files) and Accelry's Discovery Studio software.

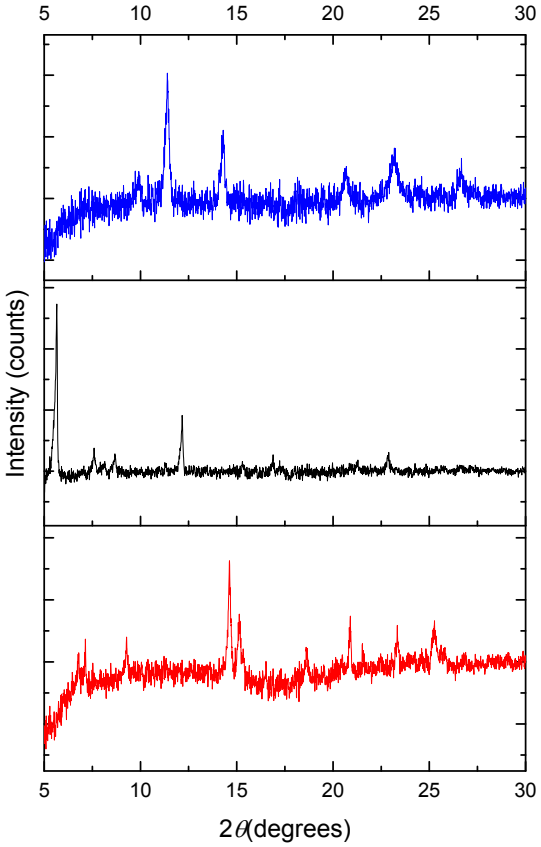


Figure S5. The powder XRD patterns **for DAD-XB-NH** (red), **DAD-XB-Boc** (black), and **DAD-XB-Boc** after thermal treatment (blue).

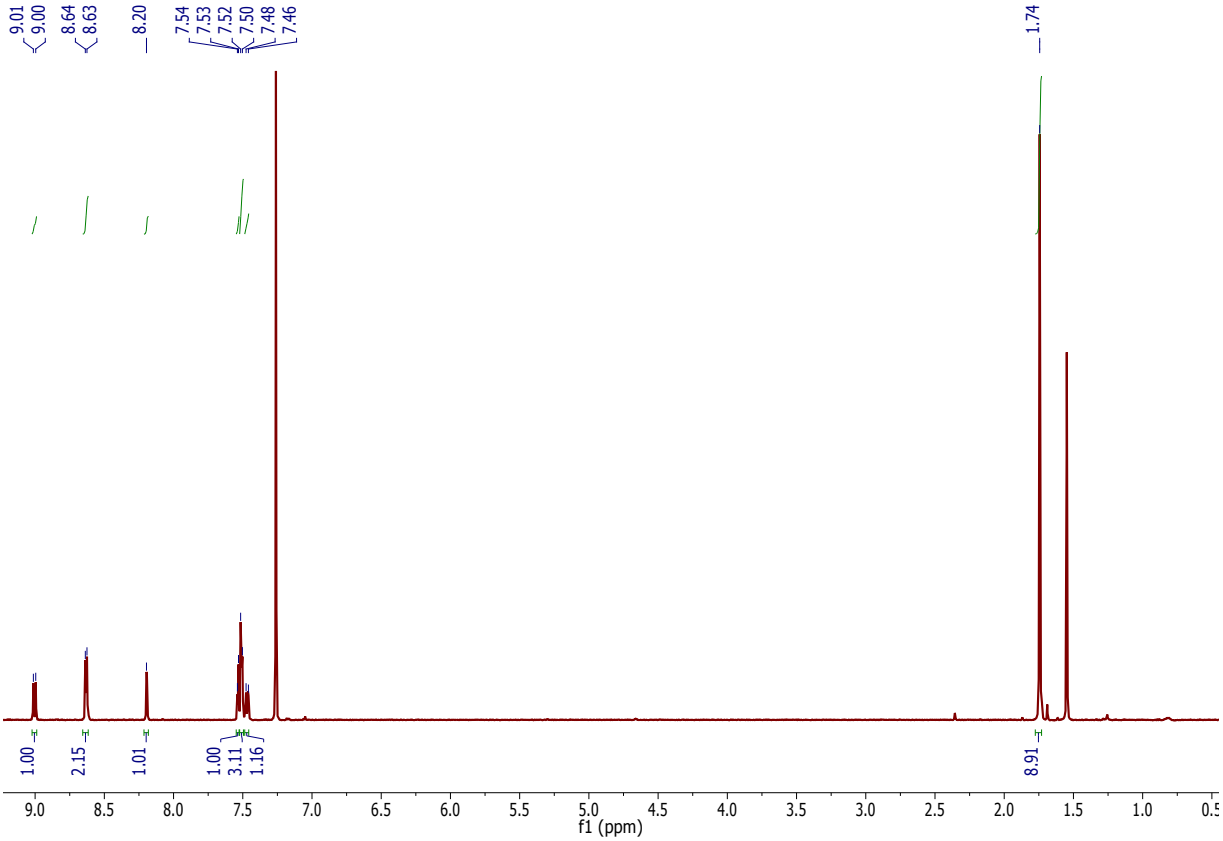


Figure S6: ^{1}H NMR of **DAD-XB-Boc** in $CDCl_3$

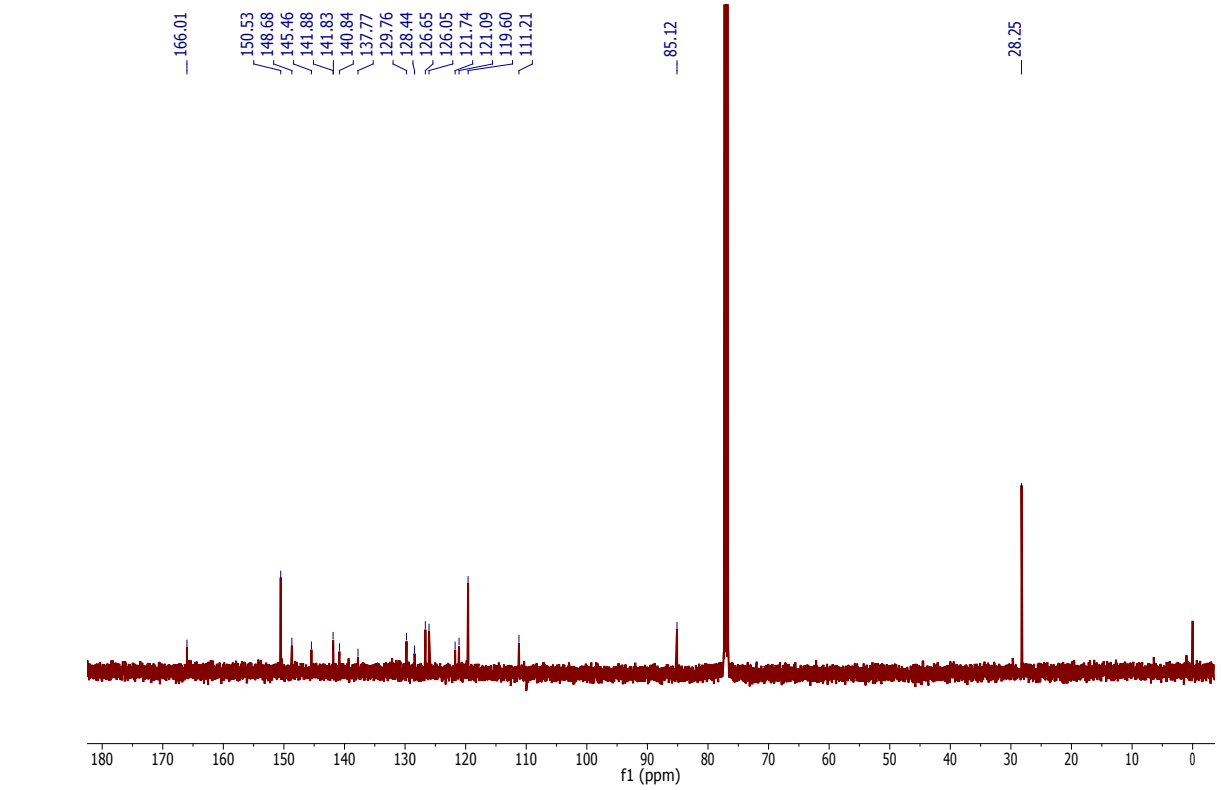


Figure S7: ^{13}C NMR of **DAD-XB-Boc** in $CDCl_3$

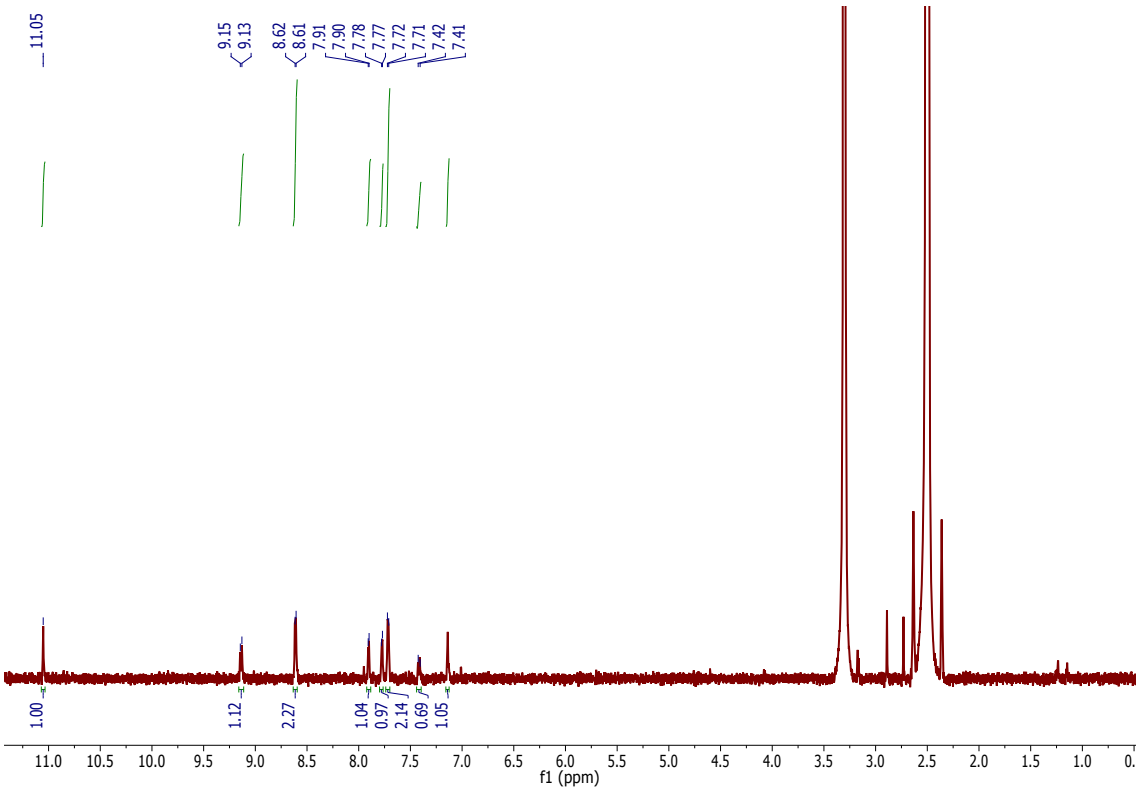


Figure S8: ^{1}H NMR of **DAD-XB-NH** in DMSO-d_6

Cartesian Coordinates of Optimized DAD-XB-Boc Structures

92

1	C	10.861899	1.819300	-0.508699	2	2	6	57	
2	C	9.805499	0.989700	-0.925200	2	1	3	7	
3	C	10.162999	-0.186100	-1.606600	2	2	4	58	
4	C	11.509000	-0.464500	-1.830699	2	3	5	59	
5	N	12.515900	0.327500	-1.441100	37	4	6		
6	C	12.173099	1.449000	-0.792599	2	1	5	60	
7	C	8.411899	1.349899	-0.658599	2	2	8	11	
8	C	7.887700	2.593800	-0.375699	2	7	9	61	
9	C	6.486099	2.592700	-0.178300	2	8	10	62	
10	C	5.905899	1.345900	-0.301700	2	9	11	12	
11	S	7.133800	0.152099	-0.677800	42	7	10		
12	C	4.497300	0.980800	-0.159699	2	10	13	17	
13	C	3.493200	1.959099	-0.307200	2	12	14	63	
14	C	2.149900	1.646899	-0.161699	2	13	15	64	
15	C	1.754300	0.330800	0.135600	2	14	16	20	
16	C	2.773400	-0.650300	0.260400	2	15	17	18	
17	C	4.117700	-0.346200	0.120699	2	12	16	65	
18	N	2.182100	-1.918299	0.460500	40	16	19	45	
19	C	0.765799	-1.792699	0.441599	3	18	20	48	
20	C	0.459200	-0.331399	0.227500	2	15	19	21	
21	C	-0.835699	0.154099	0.194399	2	20	22	25	
22	C	-2.118700	-0.501999	-0.012699	2	21	23	29	
23	C	-3.144900	0.473099	0.081100	2	22	24	26	
24	N	-2.571300	1.738799	0.357000	40	23	25	41	
25	C	-1.156699	1.606400	0.451199	3	21	24	44	
26	C	-4.479400	0.179299	-0.144100	2	23	27	66	

27	C	-4.838000	-1.140200	-0.486799	2	26	28	30	
28	C	-3.825699	-2.115400	-0.600700	2	27	29	67	
29	C	-2.492799	-1.809400	-0.373399	2	22	28	68	
30	C	-6.233600	-1.498999	-0.727000	2	27	31	34	
31	S	-7.553099	-0.528099	-0.105200	42	30	32		
32	C	-8.758799	-1.586900	-0.804300	2	31	33	35	
33	C	-8.146099	-2.631099	-1.467200	2	32	34	69	
34	C	-6.733399	-2.582399	-1.424100	2	30	33	70	
35	C	-10.190700	-1.325299	-0.651299	2	32	36	40	
36	C	-10.694600	-0.053699	-0.328299	2	35	37	71	
37	C	-12.069800	0.124600	-0.200200	2	36	38	72	
38	N	-12.974200	-0.847599	-0.373599	37	37	39		
39	C	-12.491900	-2.059100	-0.683100	2	38	40	73	
40	C	-11.138500	-2.349800	-0.825600	2	35	39	74	
41	C	-3.326400	2.918000	0.554800	3	24	42	50	
42	O	-2.552099	3.994299	0.493299	6	41	43		
43	C	-3.091299	5.348400	0.763399	1	42	51	52	53
44	O	-0.398499	2.496400	0.771999	7	25			
45	C	2.808299	-3.176300	0.648400	3	18	46	49	
46	O	4.029699	-3.006300	1.178500	6	45	47		
47	C	4.897800	-4.175599	1.472099	1	46	54	55	56
48	O	-0.000100	-2.703700	0.667000	7	19			
49	O	2.275499	-4.223299	0.372499	7	45			
50	O	-4.529200	2.890099	0.728100	7	41			
51	C	-1.839199	6.218199	0.634899	1	43	75	76	77
52	C	-4.128900	5.714799	-0.301199	1	43	78	79	80
53	C	-3.658000	5.407899	2.184500	1	43	81	82	83
54	C	5.211700	-4.931400	0.177999	1	47	84	85	86

55	C	6.155599	-3.516200	2.041399	1	47	87	88	89
56	C	4.223900	-5.061499	2.523600	1	47	90	91	92
57	H	10.667500	2.731200	0.046900	5	1			
58	H	9.404000	-0.869499	-1.976299	5	3			
59	H	11.792699	-1.373400	-2.359100	5	4			
60	H	12.996400	2.086200	-0.473400	5	6			
61	H	8.493200	3.492200	-0.329000	5	8			
62	H	5.920600	3.483299	0.070999	5	9			
63	H	3.769800	2.977100	-0.560600	5	13			
64	H	1.407099	2.421599	-0.270899	5	14			
65	H	4.864200	-1.115900	0.251499	5	17			
66	H	-5.223199	0.959400	-0.068300	5	26			
67	H	-4.092499	-3.139699	-0.838500	5	28			
68	H	-1.746700	-2.585400	-0.443599	5	29			
69	H	-8.698099	-3.396799	-2.000800	5	33			
70	H	-6.098700	-3.306600	-1.922100	5	34			
71	H	-10.027799	0.793000	-0.193900	5	36			
72	H	-12.466500	1.107200	0.050399	5	37			
73	H	-13.231399	-2.847099	-0.817299	5	39			
74	H	-10.825700	-3.364700	-1.049000	5	40			
75	H	-1.416600	6.140200	-0.371700	5	51			
76	H	-2.091800	7.266300	0.826500	5	51			
77	H	-1.076699	5.901599	1.352600	5	51			
78	H	-3.701800	5.604200	-1.303800	5	52			
79	H	-5.017999	5.087199	-0.222800	5	52			
80	H	-4.425600	6.761800	-0.173100	5	52			
81	H	-4.543700	4.778700	2.288499	5	53			
82	H	-2.903399	5.084999	2.909599	5	53			

83	H	-3.935000	6.441400	2.420000	5	53
84	H	5.946500	-5.716600	0.387699	5	54
85	H	4.316999	-5.394600	-0.240699	5	54
86	H	5.644900	-4.254299	-0.566500	5	54
87	H	6.629000	-2.863099	1.300700	5	55
88	H	5.913800	-2.918299	2.925600	5	55
89	H	6.878199	-4.286100	2.330799	5	55
90	H	3.974699	-4.474300	3.413999	5	56
91	H	3.313200	-5.519199	2.134600	5	56
92	H	4.915800	-5.856100	2.824000	5	56

References:

1. J. Wilson, J. S. Dal Williams, C. Petkovsek, P. Reves, J. W. Jurss, N. I. Hammer, G. S. Tschumper and D. L. Watkins, *RSC Advances*, 2015, **5**, 82544-82548.
2. J. Mei, K. R. Graham, R. Stalder and J. R. Reynolds, *Organic Letters*, 2010, **12**, 660-663.
3. M. Shaker, B. Park, J.-H. Lee, W. kim, C. K. Trinh, H.-J. Lee, J. w. Choi, H. Kim, K. Lee and J.-S. Lee, *RSC Advances*, 2017, **7**, 16302-16310.
4. E. C. Constable, C. E. Housecroft, M. Neuburger and C. X. Schmitt, *Polyhedron*, 2006, **25**, 1844-1863.