Electronic Supplementary Information (ESI)

Design of Simple-Structured Conjugated Polymers for Organic Solar Cells by Machine Learning-Assisted Structural Modification and Experimental Validation

*Shogo Tadokoro[1], Ryosuke Kamimura[1], Fumitaka Ishiwari[1,2,3,4], Akinori Saeki[1,2]**

[1] Department of Applied Chemistry, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan.

[2] Innovative Catalysis Science Division, Institute for Open and Transdisciplinary Research Initiatives (ICS-OTRI), Osaka University, 1-1 Yamadaoka, Suita, Osaka 565-0871, Japan.3

[3] PRESTO, Japan Science and Technology Agency (JST), Kawaguchi, Saitama 332-0012, Japan.

[4] Graduate School of Urban Environmental Sciences, Tokyo Metropolitan University, 1-1 Minamiosawa, Hachioji, Tokyo 192-0297, Japan.

AUTHOR INFORMATION

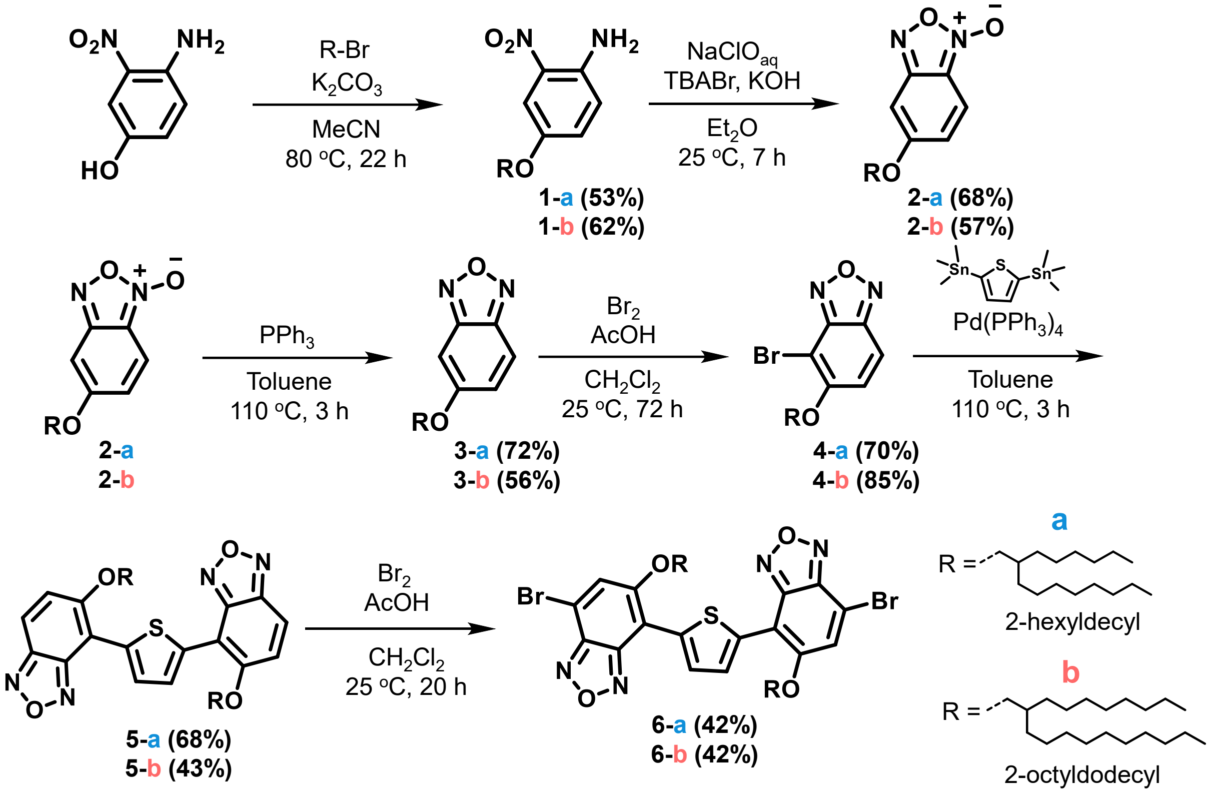
Corresponding Author

* saeki@chem.eng.osaka-u.ac.jp (A.S.)

This supporting information presents the following contents.

1. Synthesis of monomers and polymers

Synthesis of monomer



·Synthesis of 1-a

4-amino-3-nitrophenol (4.01 g, 26 mmol, 1.0 eq), K_2CO_3 (10.8 g, 78.1 mmol, 3.0 eq), 1-bromo-2-hexyldecane (19.3 mL, 61.9 mmol, 2.4 eq) and MeCN (150 mL) were added to a 200 mL two-necked flask under nitrogen atmosphere. Then the mixture was refluxed for 22 h at 80 ℃. After that it was quenched with water and extracted with CH_2Cl_2. After it was washed with brine and dried with anhydrous $MgSO_4$, the solvent was evaporated under reduced pressure. The residue was purified with silica-gel column chromatography (hexane/ chloroform = 1/1). As a result, **1-a** was obtained (5.22 g, 13.8 mmol, 53%). ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 7.53 (d, J = 1.6 Hz, 1H), 7.07 (dd, J = 1.6 Hz, 6.0 Hz, 1H), 6.75 (d, J = 6.0 Hz, 1H), 3.79 (d, J = 3.6 Hz, 2H), 1.75 (m, 1H), 1.37 (m, 24H), 0.88 (m, 6H). ^{13}C NMR (151 MHz) δ (ppm): 150.72, 139.88, 131.70, 127.29, 120.08, 107.20, 71.83, 38.03, 32.04, 31.98, 31.59, 31.48, 31.47, 30.14, 29.81, 29.71, 29.46, 26.97, 26.94, 22.92, 22.81, 14.25. FT-IR (KBr) ν(cm^{-1}): 3488, 3364, 2956, 2926, 2853, 1579, 1572, 1521, 1462, 1415, 1372, 1329, 1288, 1246, 1216, 1151, 1078, 1025, 839, 814, 767, 718, 684, 547, 478. HR ESI$^+$ MS: calced. For $C_{22}H_{38}N_2O_3Na$ $[M+Na]^+$: m/z = 401.27746; found: 401.27790.

·Synthesis of 2-a[30]

1-a (2.08 g, 5.50 mmol, 1.0 eq), Tetrabutylammonium Bromide (TBAB) (0.0301 g, 0.0935 mmol, 0.017 eq), KOH aq (2.16 mL, 23.1 mmol, 50 wt%, 4.2 eq) and diethyl ether (10 mL) were added to a 50 mL two-necked flask. To this mixture, a sodium hypochlorite solution (11 mL, 14.8 mmol, activated chlorine ≒ 10%, 2.69 eq) was added dropwise. After stirring at 25 ℃ for 7 h, the organic layer was separated. The aqueous layer was extracted with CH_2Cl_2, washed with brine and dried with anhydrous $MgSO_4$. The combined organic layers were evaporated under reduced pressure. As a result, **2-a** was obtained (1.42 g, 3.76 mmol, 68%) without further purification. FT-IR (KBr) ν(cm^{-1}): 2955, 2927, 2855, 1637, 1622, 1602, 1585, 1541, 1495, 1466, 1379, 1313, 1271, 1225, 1192, 1119, 1006, 808. HR ESI$^+$ MS: calced. For $C_{22}H_{36}N_2O_3Na$ $[M+Na]^+$: m/z = 399.26181; found: 399.26584.

・Synthesis of 3-a[30]

2-a (1.42 g, 3.76 mmol, 1.0 eq), triphenylphosphine (PPh_3) (1.20 g, 4.59 mmol, 1.22 eq), and toluene (60 mL) were added to a 100 mL flask. The mixture was refluxed for 3 h then cooled. The solvent was evaporated under reduced pressure. The residue was purified with silicagel column chromatography (hexane/ chloroform = 3/1). As a result, **3-a** was obtained (0.976 g, 3.76 mmol, 72%). ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 7.68 (d, J = 9.6 Hz, 1H), 7.10 (dd, J = 9.6 Hz, 2.0 Hz, 1H), 6.82 (d, J = 2.0 Hz, 1H), 3.92 (d, J = 5.6 Hz, 2H), 1.84 (m, 1H), 1.41 (m, 24H), 0.87 (m, 6H). ^{13}C NMR (151 MHz) δ (ppm): 160.84, 150.12, 146.82, 129.64, 116.90, 89.74, 71.61, 37.50, 31.79, 31.72, 31.23, 31.21, 29.85, 29.52, 29.45, 29.21, 26.71, 26.67, 22.57, 22.56, 13.40, 13.98. FT-IR (KBr) ν(cm^{-1}): 2955, 2927, 2856, 1635, 1544, 1490, 1467, 1412, 1367, 1349, 1317, 1256, 1216, 1134, 1010, 883, 840, 810, 750, 723, 628, 618, 438. HR EI$^+$ MS: calced. For $C_{22}H_{36}N_2O_2$ $[M]^+$: m/z = 360.2777; found: 360.2768.

・Synthesis of 4-a[31]

3-a (0.65 g, 1.81 mmol, 1.0 eq), Br_2 (0.40 mL, 7.69 mmol, 4.15 eq), AcOH (3.29 mL, 30 eq) and CH_2Cl_2 (40 mL) were added to a 100 mL flask. The mixture was stirred in the dark for 3 days at 25 ℃ and then poured into NaOH aq (2.5 g in 100 mL). A saturated sodium bisulfite aqueous solution was then added to remove excess bromine. The aqueous layer was extracted with CH_2Cl_2, washed with brine and dried with anhydrous $MgSO_4$. The combined organic layers were evaporated under reduced pressure. The residue was purified with silicagel column chromatography (hexane/ chloroform = 3/1). As a result, **4-a** was obtained (0.56 g, 1.27 mmol, 70%). ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 7.79 (d, J = 9.6 Hz, 1H), 7.33 (d, J = 9.6 Hz, 1H), 4.09 (d, J = 5.6 Hz, 2H), 1.84 (m, 1H), 1.45 (m, 24H), 0.87 (m, 6H). ^{13}C NMR (151 MHz) δ (ppm): 156.94, 150.42, 146.60, 124.24, 116.58, 91.66, 73.80, 38.35, 31.95, 31.86, 31.15, 31.14, 30.00, 29.67, 29.60, 29.37, 26.85, 26.83, 22.73, 22.71, 14.16, 14.15. FT-IR (KBr) ν(cm^{-1}): 2954, 2927, 2855, 1625, 1540, 1516, 1457, 1378, 1307, 1254, 1063, 1021, 961, 887, 852, 800, 757, 724. HR DART$^+$ MS: calced. For $C_{22}H_{35}N_2O_2BrH^+$ $[M+H]^+$: m/z = 439.19547; found: 439.19431.

・Synthesis of 5-a[31]

4-a (770 mg, 1.75 mmol, 2.0eq), 2,5-Bis(trimethylstannyl)thiophene (360 mg, 0.875 mmol, 1.0 eq), $Pd(PPH_3)_4$ (38 mg, 0.033 mmol, 5 mol%), and toluene (30 ml, degassed for 15 min) were added to a 50 mL two-necked flask under nitrogen atmosphere. Then the reaction mixture was refluxed for 16 h. After that, it was filtered by celite to remove Pd catalyst and the solvent was removed under reduced pressure. The residue was purified by silicagel column chromatography (hexane/chloroform = 4/1). As a result, **5-a** was obtained (476 mg, 0.595 mmol, 68%). ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.45 (s, 1H), 7.7 3(d, J = 9.6 Hz, 1H), 7.48 (d, J = 9.6 Hz, 1H), 4.21 (d, J = 6.0 Hz, 2H), 2.04 (m, 1H), 1.45 (m, 24H), 0.84 (m, 6H). ^{13}C NMR (151 MHz) δ (ppm): 154.41, 149.36, 147.06, 135.57, 130.17, 124.56, 115.16, 107.91, 74.37, 38.54, 32.00, 31.97, 31.29, 31.27, 30.14, 29.80, 29.73, 29.45, 26.92, 26.91, 22.78, 22.77, 14.21, 14.18. FT-IR (KBr) ν(cm^{-1}): 2954, 2926, 2855, 1699, 1684, 1654, 1616, 1558, 1542, 1521, 1507, 1472, 1465, 1457, 1447, 1439, 1341, 1238, 1010, 892, 795, 483, 425, 410. HR MALDI-TOF MS: calced. For $C_{48}H_{72}N_4O_4SNa^+$ $[M+Na]^+$: m/z = 823.51665; found: 823.51632.

・Synthesis of 6-a

5-a (528 mg, 0.66 mmol, 1.0 eq), Br_2 (0.068 mL, 1.32 mmol, 2.0 eq), AcOH (1.19 mL, 19.8 mmol, 30 eq) and CH_2Cl_2 (25 mL) were added to a 50 mL two-necked flask. The mixture was stirred in the dark for 20 h at 25 ℃ and then poured into NaOH aq (0.5 g in 50 mL). A saturated sodium bisulfite aqueous solution was then added to remove excess bromine. The aqueous layer was extracted with CH_2Cl_2, washed with brine and dried with anhydrous $MgSO_4$. The combined organic layers were evaporated under reduced pressure. The residue was purified with silicagel column chromatography (hexane/ chloroform = 4/1). As a result, **6-a** was obtained (271 mg, 0.28 mmol, 42%).^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.38 (s, 1H), 7.67 (s, 1H), 4.19 (d, J = 6.0 Hz, 2H), 2.03 (m, 1H), 1.45 (m, 24H), 0.83 (m, 6H). ^{13}C NMR (151 MHz) δ (ppm): 154.52, 148.82, 147.42, 135.45, 130.37, 126.93, 107.21, 107.18, 74.84, 38.50, 32.01, 31.97, 31.27, 31.26, 30.12, 29.79, 29.73, 29.47, 26.91, 26.90, 22.79, 22.78, 14.22, 14.19. FT-IR (KBr) ν(cm^{-1}): 2954, 2927, 2855, 1698, 1684, 1653, 1647, 1636, 1625, 1617, 1604, 1558, 1541, 1522, 1507, 1472, 1465, 1457, 1337, 1216, 1014, 876, 810. HR MALDI-TOF MS: calced. For $C_{48}H_{70}N_4O_4SBr_2$ [M] : m/z = 979.33767; found: 979.3383.

・Synthesis of 1-b

1-b was synthesized by the same procedure as **1-a** using 1-bromo-2-octyldodecane instead of 1-bromo-2-hexyldecane (7.24 g, 62%). ^{1}H NMR (600 MHz, $CDCl_3$) δ (ppm): 7.54 (d, J = 2.8 Hz, 1H), 7.07 (dd, J = 2.8 Hz, 9.2 Hz, 1H), 6.75 (d, J = 9.2 Hz, 1H), 3.79 (d, J = 5.6 Hz, 2H), 1.75 (m, 1H), 1.36 (m, 32H), 0.88 (m, 6H). ^{13}C NMR (151 MHz) δ (ppm): 150.69, 139.91, 131.65, 127.28, 120.07, 107.15, 71.80, 38.01, 32.05, 32.03, 31.45, 30.13, 29.79, 29.77, 29.75, 29.71, 29.48, 29.45, 26.96, 22.81, 22.80, 14.23. FT-IR (KBr) ν(cm^{-1}): 3490, 3371, 2956, 2925, 2854, 1587, 1571, 1518, 1467, 1417, 1378, 1332, 1283, 1251, 1222, 1155, 1088, 1031, 820, 763, 723, 687, 604, 581, 553, 436, 417. HR ESI$^+$ MS: calced. For $C_{26}H_{46}N_2O_3H^+$ $[M+H]^+$: m/z = 435.35812; found: 435.35923.

・Synthesis of 2-b

2-b was synthesized by the same procedure as **2-a** (2.90 g, 57%). FT-IR (KBr) ν(cm^{-1}): 2953, 2925, 2854, 1621, 1601, 1585, 1541, 1495, 1466, 1378, 1312, 1271, 1225, 1192, 1119, 1006, 808, 722, 625, 590, 557. HR EI$^+$ MS: calced. For $C_{26}H_{44}N_2O_3$ $[M]^+$: m/z = 432.3352; found: 433.3351.

・Synthesis of 3-b

3-b was synthesized by the same procedure as **3-a** (1.47 g, 56%). ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 7.68 (d, J = 9.6 Hz, 1H), 7.10 (dd, J = 9.6 Hz, 2.0 Hz, 1H), 6.82 (d, J = 2.0 Hz, 1H), 3.92 (d, J = 5.6 Hz, 2H), 1.83 (m, 1H), 1.39 (m, 32H), 0.88 (m, 6H). ^{13}C NMR (151 MHz) δ (ppm): 160.81, 150.09, 146.79, 129.60, 116.88, 89.74, 71.60, 37.49, 31.81, 31.79, 31.31, 31.21, 31.09, 29.85, 29.76, 29.55, 29.53, 29.49, 29.45, 29.31, 29.24, 29.21, 29.15, 26.81, 26.71, 26.58, 22.58, 22.56, 13.40, 13.99. FT-IR (KBr) ν(cm^{-1}): 2958, 2925, 2855, 1635, 1544, 1490, 1467, 1367, 1317, 1256, 1216, 1134, 1010, 883, 840, 809, 755, 722, 629, 582, 533, 477, 435, 420. HR DART$^+$ MS: calced. For $C_{26}H_{44}N_2O_2H^+$ $[M+H]^+$: m/z = 417.34756; found: 417.34794.

・Synthesis of 4-b

4-b was synthesized by the same procedure as **4-a** (1.13 g, 85%). ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 7.79 (d, J = 9.6 Hz, 1H), 7.33 (d, J = 9.6 Hz, 1H), 4.09 (d, J = 5.6 Hz, 2H), 1.84 (m, 1H), 1.41 (m, 32H), 0.87 (m, 6H). ^{13}C NMR (151 MHz) δ (ppm): 156.80, 150.20, 146.40, 124.02, 116.38, 91.37, 73.61, 38.27, 31.91, 31.90, 31.58, 31.08, 29.95, 29.67, 29.65, 29.59, 29.55, 29.36, 29.33, 26.78, 22.67, 22.66, 22.63, 14.05. FT-IR (KBr) ν(cm^{-1}): 2958, 2926, 2854, 1625, 1540, 1516, 1458, 1378, 1307, 1254, 1159, 1064, 1021, 963, 887, 852, 800, 757, 722, 643, 625, 564. HR DART$^+$ MS: calced. For $C_{26}H_{43}N_2O_2BrH^+$ $[M+H]^+$: *m*/*z* = 495.25807; found: 495.25791.

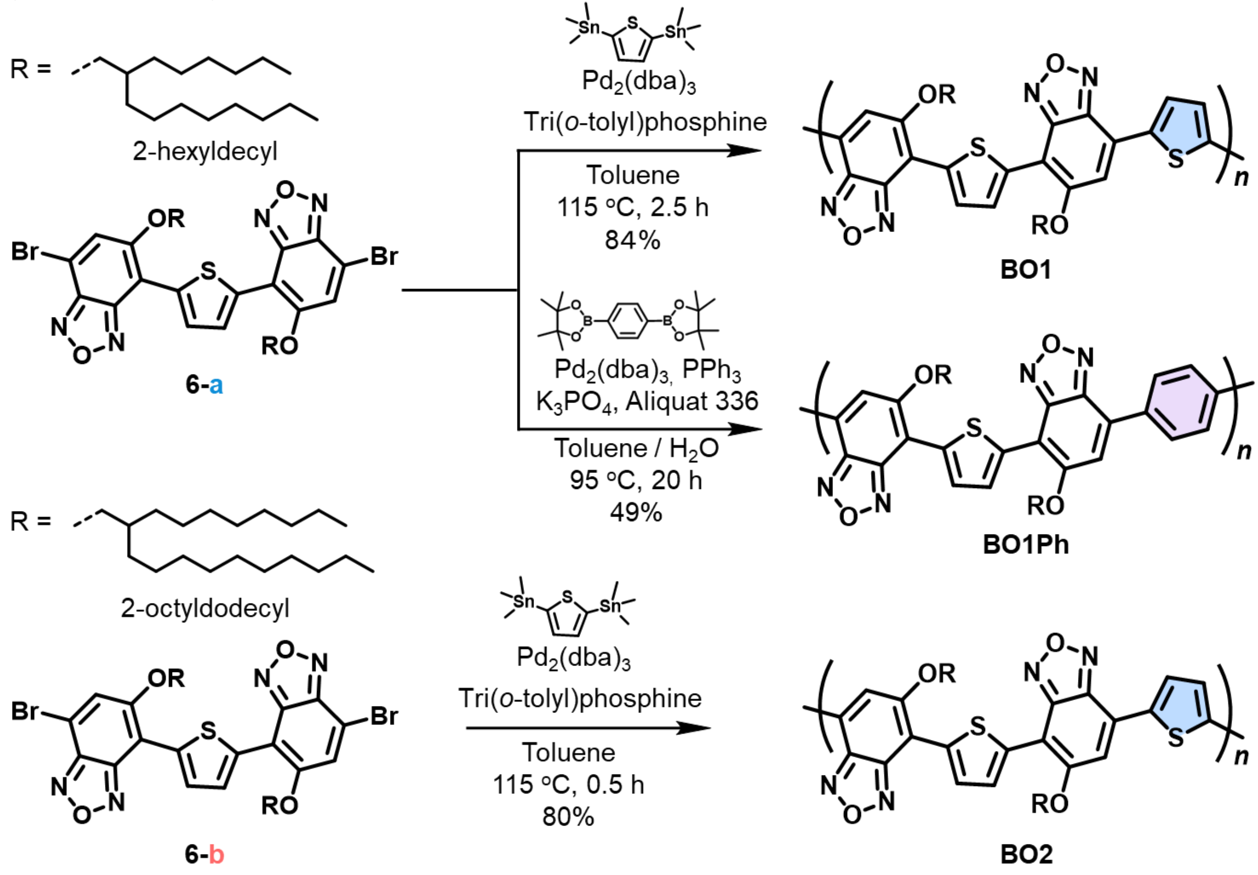
・Synthesis of 5-b

5-b was synthesized by the same procedure as **5-a** (385 mg, 43%). ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.45 (s, 1H), 7.73 (d, *J* = 6.4 Hz, 1H), 7.48 (d, *J* = 6.4 Hz, 1H), 4.21 (d, *J* = 4.0 Hz, 2H), 2.04 (m, 1H), 1.47 (m, 32H), 0.84 (m, 6H). ^{13}C NMR (151 MHz) δ (ppm): 154.37, 149.33, 147.04, 135.57, 130.16, 124.56, 115.16, 107.90, 74.35, 38.51, 32.03, 32.00, 31.26, 30.14, 29.80, 29.78, 29.76, 29.73, 29.47, 29.46, 26.92, 22.81, 22.78, 14.24, 14.21. FT-IR (KBr) ν(cm^{-1}): 2951, 2925, 2854, 1542, 1457, 1342, 1237, 1061, 1061, 1010, 891, 798, 644, 588, 576, 441. HR MALDI-TOF MS: calced. For $C_{56}H_{88}N_4O_4SNa^+$ $[M+Na]^+$: *m*/*z* = 935.64185; found: 935.64235.

・Synthesis of 6-b

6-b was synthesized by the same procedure as **6-a** (109 mg, 42%). ^{1}H NMR (600 MHz, $CDCl_3$) δ (ppm): 8.39 (s, 1H), 7.67 (s, 1H), 4.21 (d, *J* = 6.0 Hz, 2H), 2.04 (m, 1H), 1.47 (m, 32H), 0.86 (m, 6H). ^{13}C NMR (151 MHz) δ (ppm): 154.49, 148.81, 147.42, 135.45, 130.38, 126.91, 107.23, 107.18, 74.84, 38.48, 32.04, 32.02, 31.26, 30.13, 29.82, 29.79, 29.78, 29.74, 29.49, 29.48, 26.93, 22.82, 22.80, 14.25, 14.23. FT-IR (KBr) ν(cm^{-1}): 2954, 2926, 2854, 1604, 1536, 1466, 1433, 1379, 1358, 1333, 1299, 1233, 1216, 1176, 1095, 1014, 895, 875, 811, 722, 512. HR MALDI-TOF MS: calced. For $C_{56}H_{86}N_4O_4SBr_2$ [M] : *m*/*z* = 1068.47310; found: 1068.47521.

Synthesis of Polymers



・Synthesis of BO1

6-a (63 mg, 0.0657 mmol, 1.0eq), 2,5-Bis(trimethylstannyl)thiophene (26.9 mg, 0.0657 mmol, 1.0 eq), Tri(*o*-tolyl)phosphine (1.60 mg, 0.00526 mmol, 8 mol%), Pd_2dba_3 (1.23 mg, 0.00132 mmol, 2 mol%), and toluene (4 ml, degassed for 15 min) were added to a 10 mL two-necked flask under nitrogen atmosphere. Then the reaction mixture was stirred for 2.5 h at 110 ℃. After that, it was cooled and precipitated into the methanol. Then soxhlet extraction was carried out with methanol, hexane, acetone, and chloroform. The chloroform fraction was evaporated and precipitated from methanol to get **BO1** (44 mg, 84%). Molecular weight cannot be measured accurately due to the exclusion limit of GPC ($CHCl_3$ at 40℃) and aggregation of polymer. ^1H NMR (600 MHz, $CDCl_3$) δ (ppm): 7.71 (m), 7.52(m), 3.54(m), 1.33 (m), 0.88 (m).

・Synthesis of BO2

6-b (57 mg, 0.053 mmol, 1.0 eq), 2,5-Bis(trimethylstannyl)thiophene (21.8 mg, 0.053 mmol, 1.0 eq), Tri(*o*-tolyl) phosphine (1.30 mg, 0.00424 mmol, 8 mol%), Pd_2dba_3 (0.98 mg, 0.00106 mmol, 2 mol%), and toluene (6 ml, degassed for 15 min) were added to a 10 mL two-necked flask under nitrogen atmosphere. Then the reaction mixture was stirred for 0.5 h at 110 ℃. After that, it was cooled and precipitated into the methanol. The residue was purified by GPC ($CHCl_3$ at 40℃) to obtain **BO2** (42 mg, M_n= 32.7 kDa, M_n/M_w= 1.9, 80%). ^1H NMR (600 MHz, $CDCl_3$) δ (ppm): 7.71 (m), 7.52(m), 4.05(m), 1.45 (m), 0.94 (m).

・Synthesis of BO1Ph

6-a (83 mg, 0.0866 mmol, 1.0 eq), 1,4-Benzenediboronic acid bis(pinacol) Ester (28.6 mg, 0.0866 mmol, 1.0 eq), K_3PO_4 (40.4 mg, 0.191 mmol, 2.2 eq), PPh_3 (2.68 mg, 0.0103 mmol, 0.12 eq), Pd_2dba_3 (4.76 mg, 0.00520 mmol, 6 mol%) toluene (2.4 mL, degassed for 15 min) and water (0.6 mL, degassed for 15 min) were added to a 10 ml two-necked flask under nitrogen atmosphere. Then the reaction mixture was stirred for 16 h at 95 ℃. After that, it was cooled and precipitated into the methanol. The residue was purified by GPC ($CHCl_3$ at 40℃) to obtain **BO1Ph** (37 mg, M_n= 31.5 kDa, M_n/M_w= 2.6, 49%). ^1H NMR (600 MHz, $CDCl_3$) δ (ppm): 7.70 (m), 7.52(m), 4.30(m), 1.45 (m), 0.87 (m).

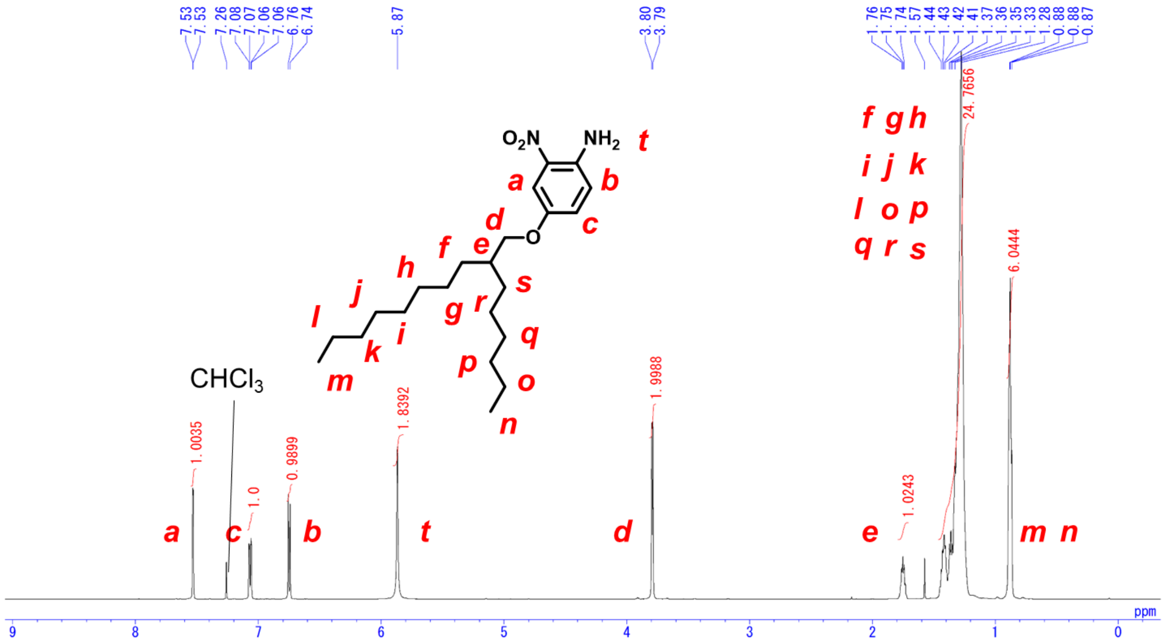


Figure S1. ^{1}H NMR spectrum of **1-a** (400 MHz, CDCl$_3$).

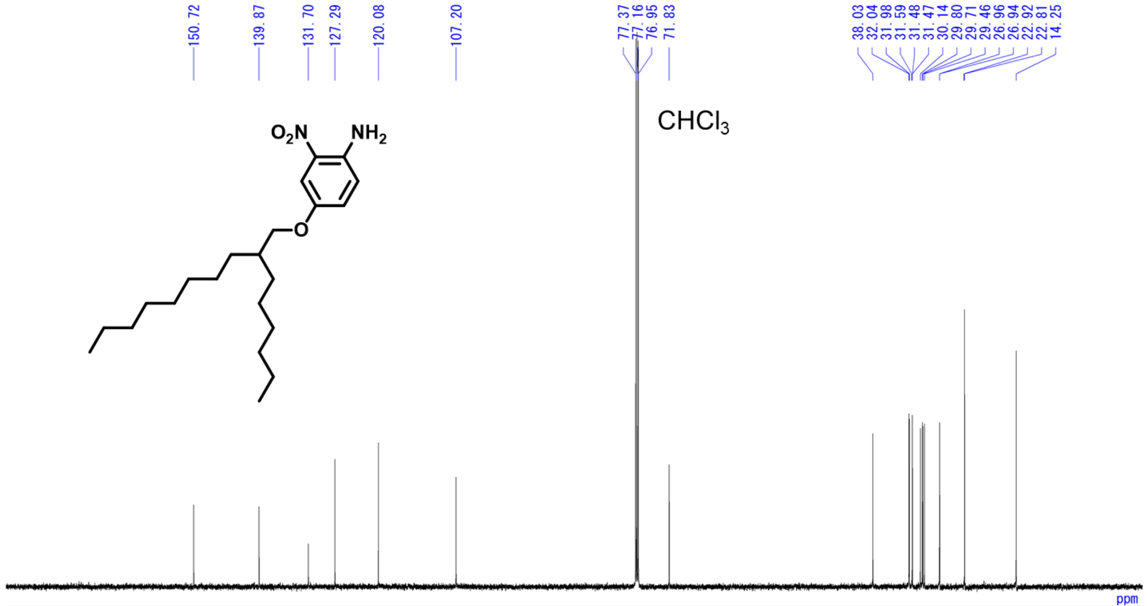


Figure S2. ^{13}C NMR spectrum of **1-a** (151 MHz, CDCl$_3$).

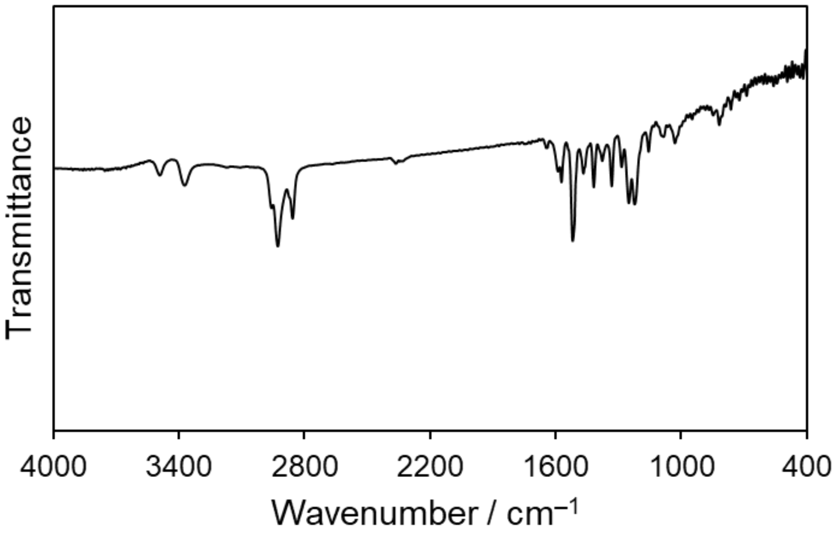
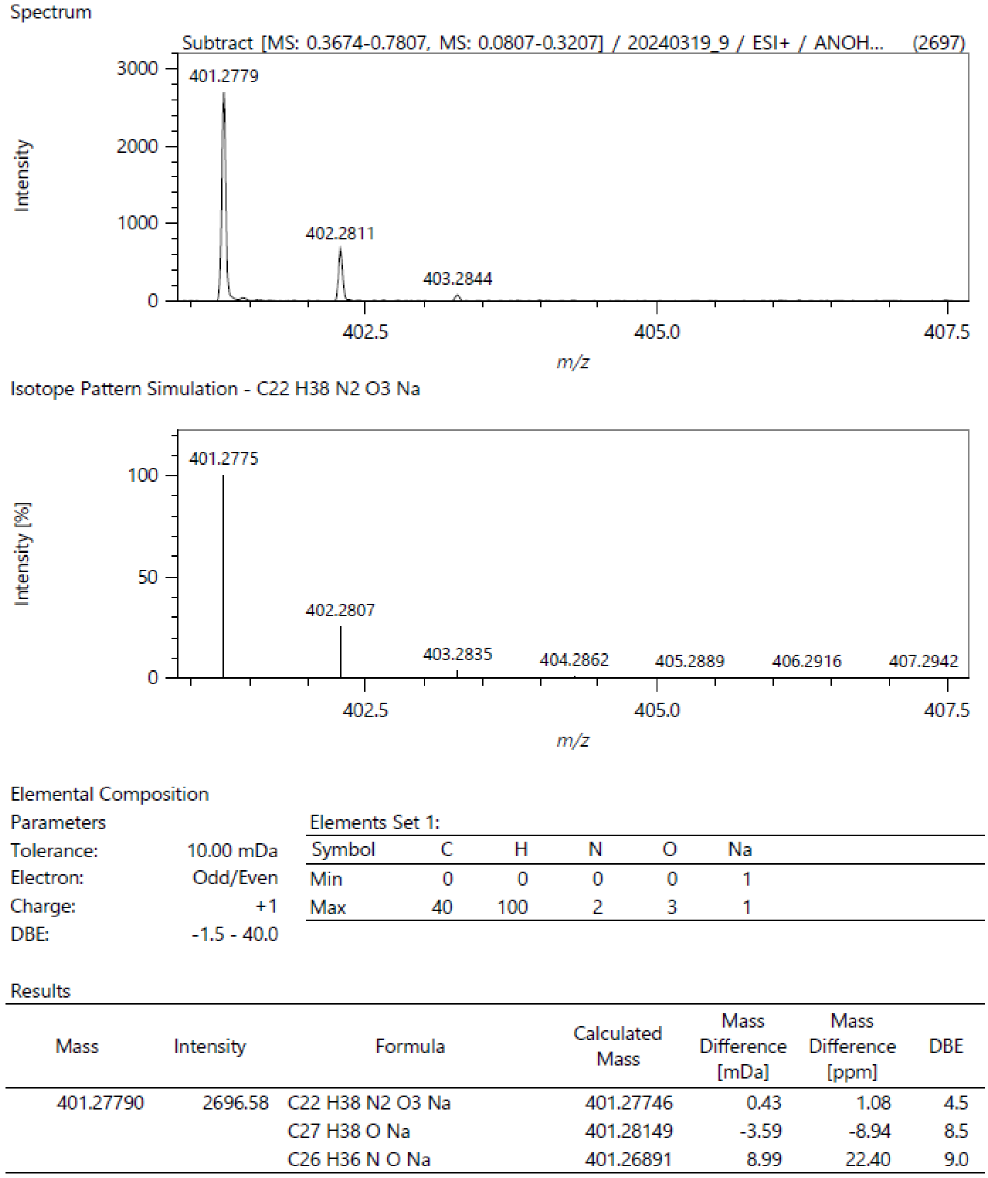


Figure S3. FT-IR spectrum of **1-a** (KBr).



Elemental Composition

Parameters

Tolerance:	10.00 mDa
Electron:	Odd/Even
Charge:	+1
DBE:	-1.5 - 40.0

Elements Set 1:

Symbol	C	H	N	O	Na
Min	0	0	0	0	1
Max	40	100	2	3	1

Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [mDa]	Mass Difference [ppm]	DBE
401.27790	2696.58	C22 H38 N2 O3 Na	401.27746	0.43	1.08	4.5
		C27 H38 O Na	401.28149	-3.59	-8.94	8.5
		C26 H36 N O Na	401.26891	8.99	22.40	9.0

Figure S4. ESI+ mass spectrum of **1-a**.

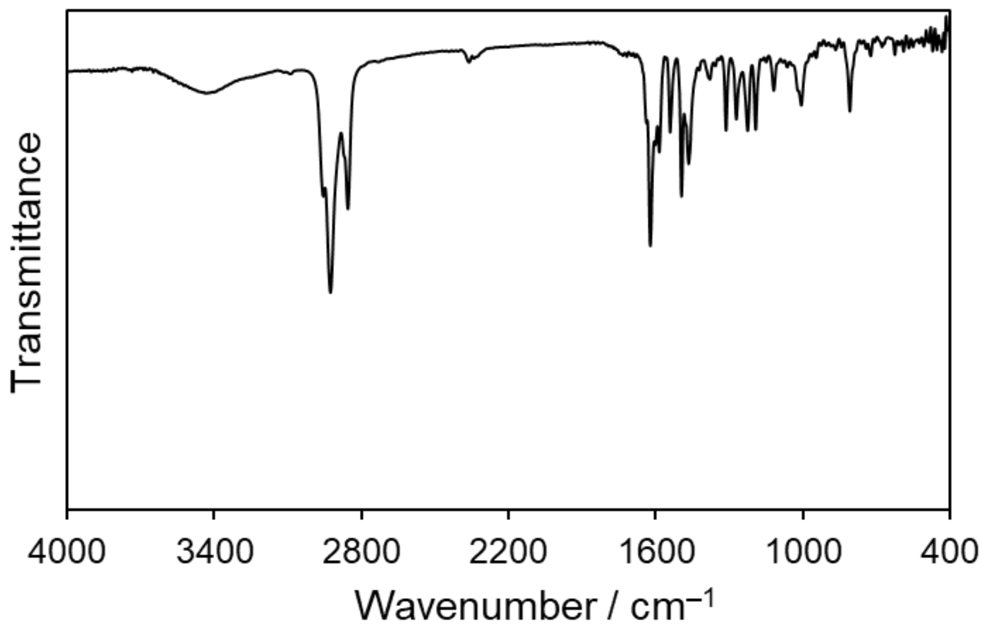
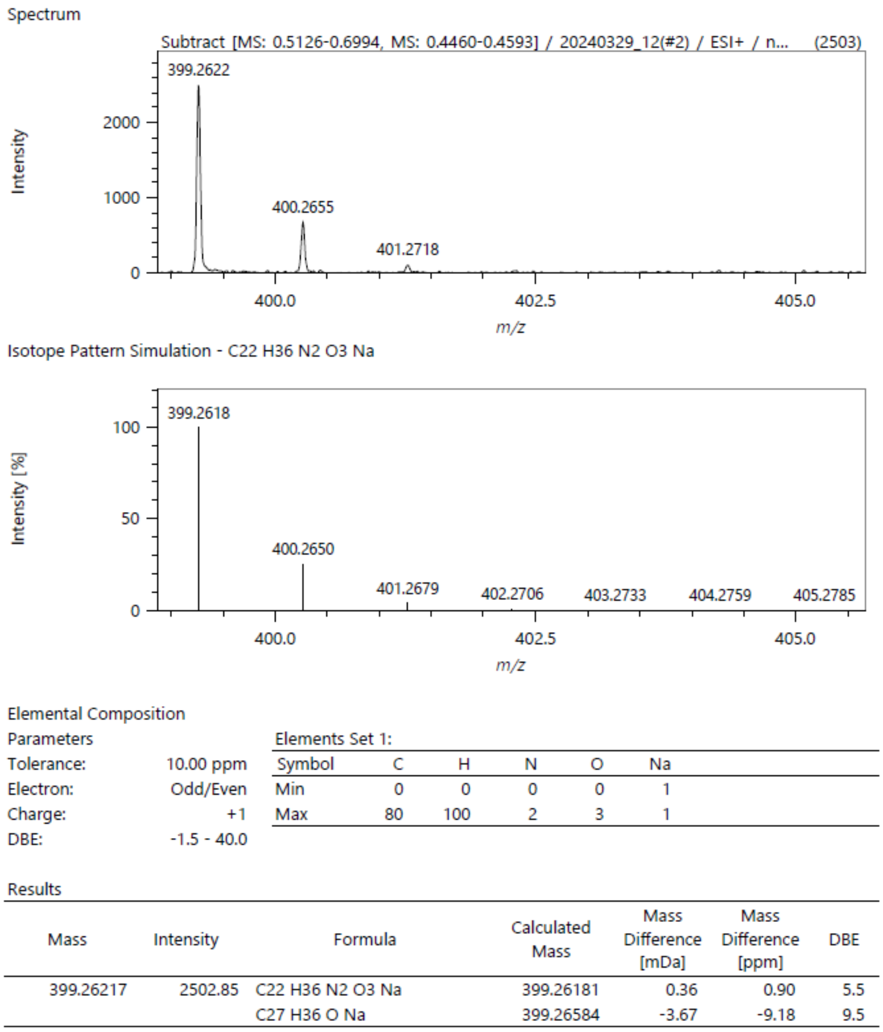


Figure S5. FT-IR spectrum of **2-a** (KBr).



Isotope Pattern Simulation - C22 H36 N2 O3 Na

Elemental Composition

Parameters

Tolerance:	10.00 ppm
Electron:	Odd/Even
Charge:	+1
DBE:	-1.5 - 40.0

Elements Set 1:

Symbol	C	H	N	O	Na
Min	0	0	0	0	1
Max	80	100	2	3	1

Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [mDa]	Mass Difference [ppm]	DBE
399.26217	2502.85	C22 H36 N2 O3 Na	399.26181	0.36	0.90	5.5
		C27 H36 O Na	399.26584	-3.67	-9.18	9.5

Figure S6. ESI+ mass spectrum of **2-a**.

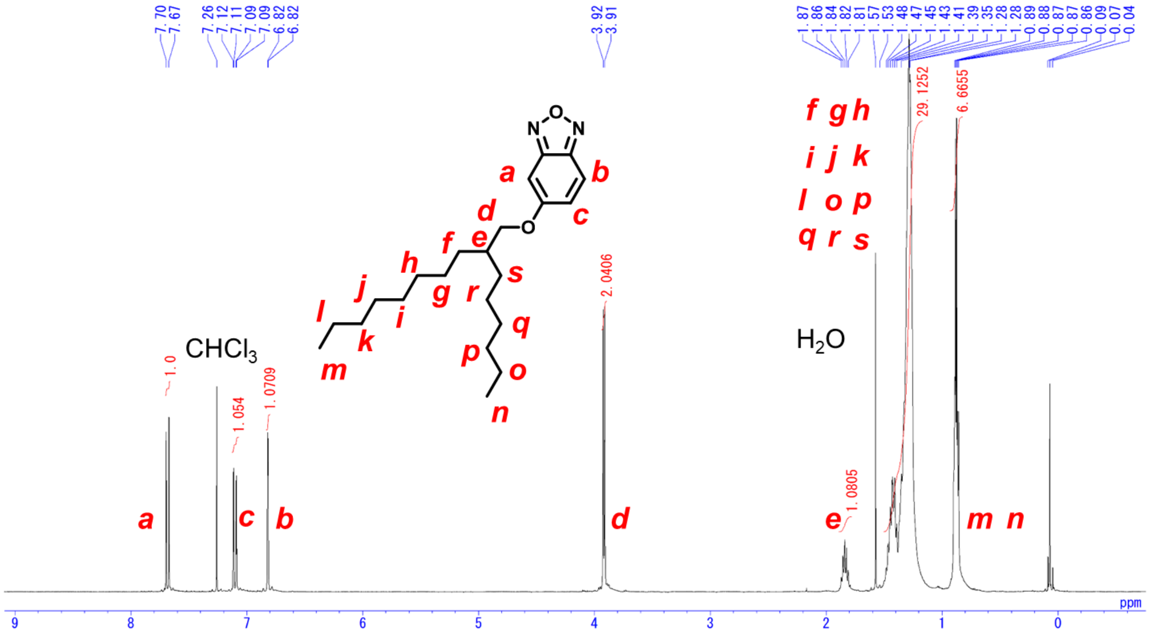


Figure S7. ^{1}H NMR spectrum of **3-a** (400 MHz, CDCl$_3$).

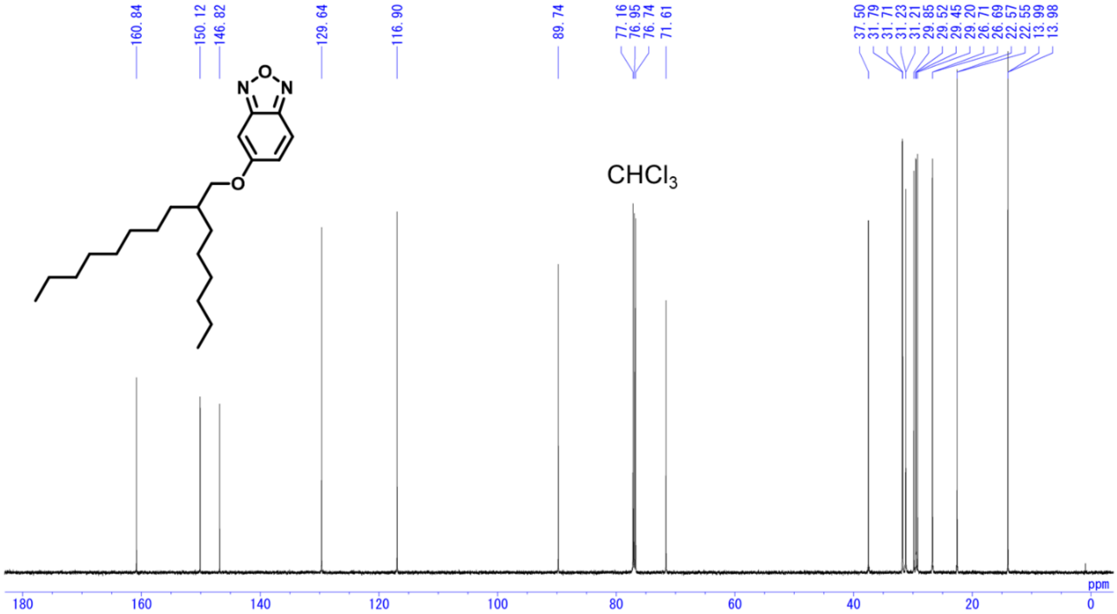


Figure S8. ^{13}C NMR spectrum of **3-a** (151 MHz, CDCl$_3$).

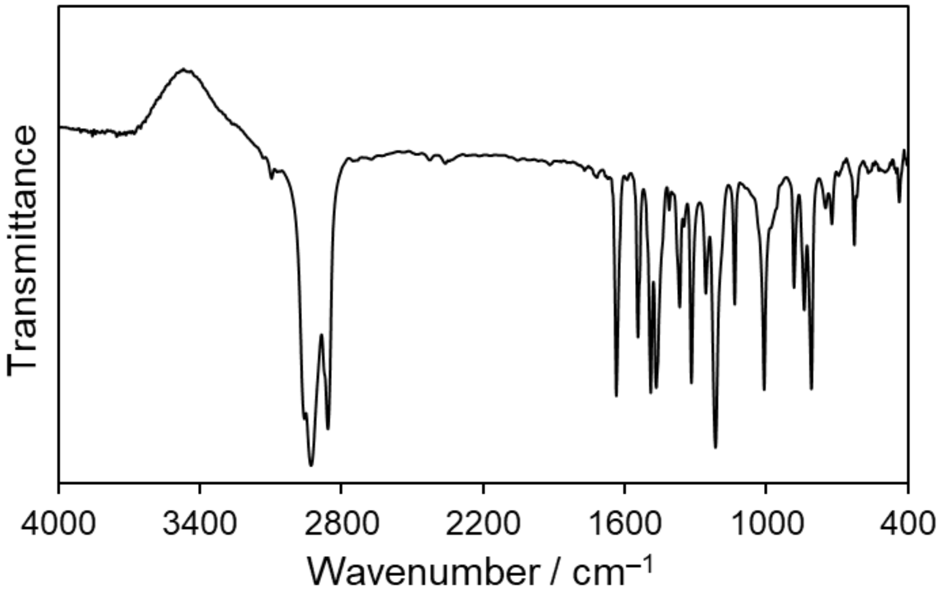
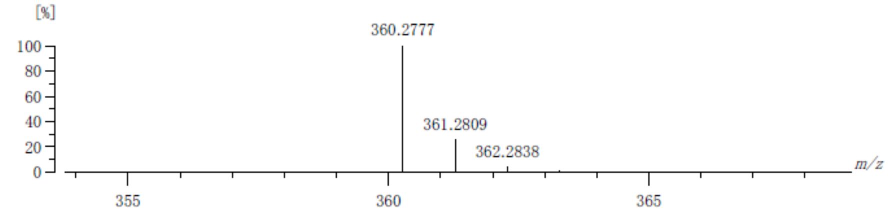
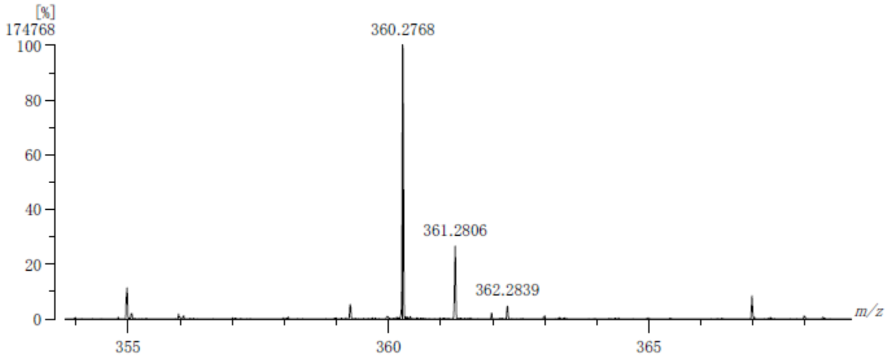


Figure S9. FT-IR spectrum of **3-a** (KBr).

```
[ Molecular Formula ]
Data : Saeki_lab HR-066      Date : 01-Nov-2024 15:42
Instrument : JMS-700
Sample : OHD-comp.3
Inlet : Direct       Ion Mode : EI+
RT : 0.90 min       Molecular Formula : C22 H36 N2 O2
Elements : C 80/0, H 100/0, N 2/0, O 2/0
Mass Tolerance       : 10mmu
Unsaturation (U.S.) : -0.5 - 15.0
```



```
[ Mass Spectrum ]
Data : Saeki_lab HR-066      Date : 01-Nov-2024 15:42
RT : 0.90 min       Scan# : (28,31)
```



```
 Observed m/z  Int%     Err[ppm / mmu]    U.S.  Composition
1    360.2768  100.00    -13.6 /   -4.9   10.0  C27 H36
2                        +21.3 /   +7.7   10.5  C26 H34 N
3                         -2.4 /   -0.9    6.0  C22 H36 N2 O2
```

Figure S10. EI+ mass spectrum of **3-a**.

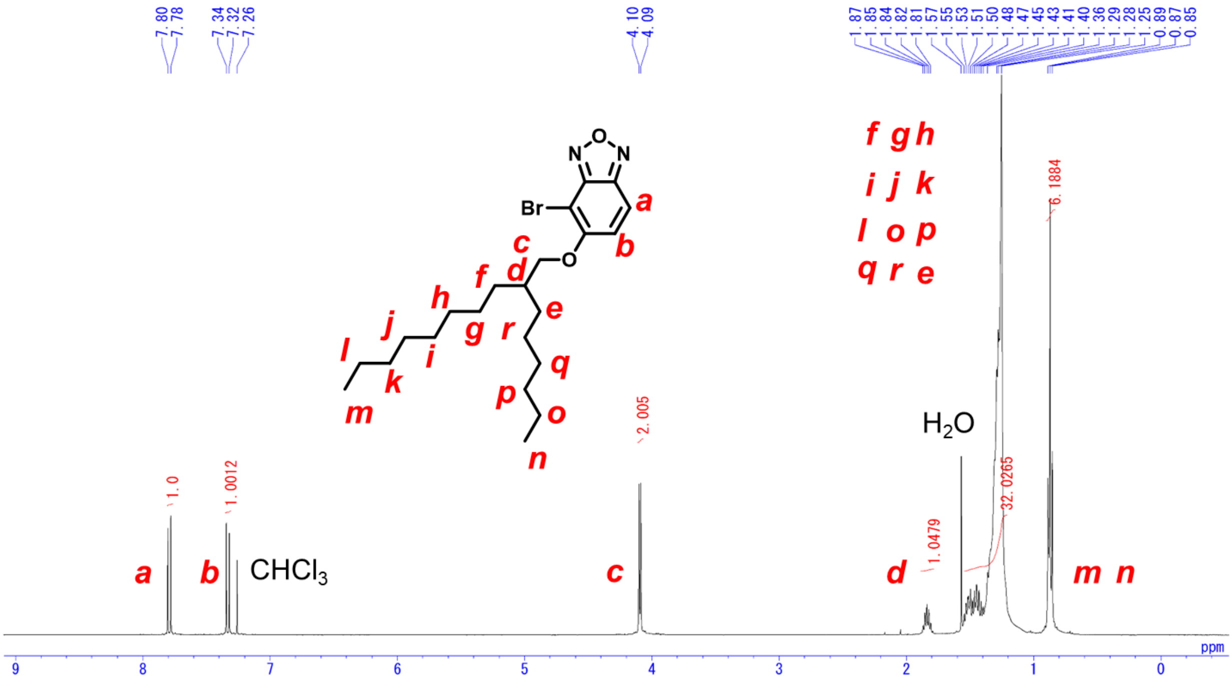


Figure S11. ^{1}H NMR spectrum of **4-a** (400 MHz, CDCl$_3$).

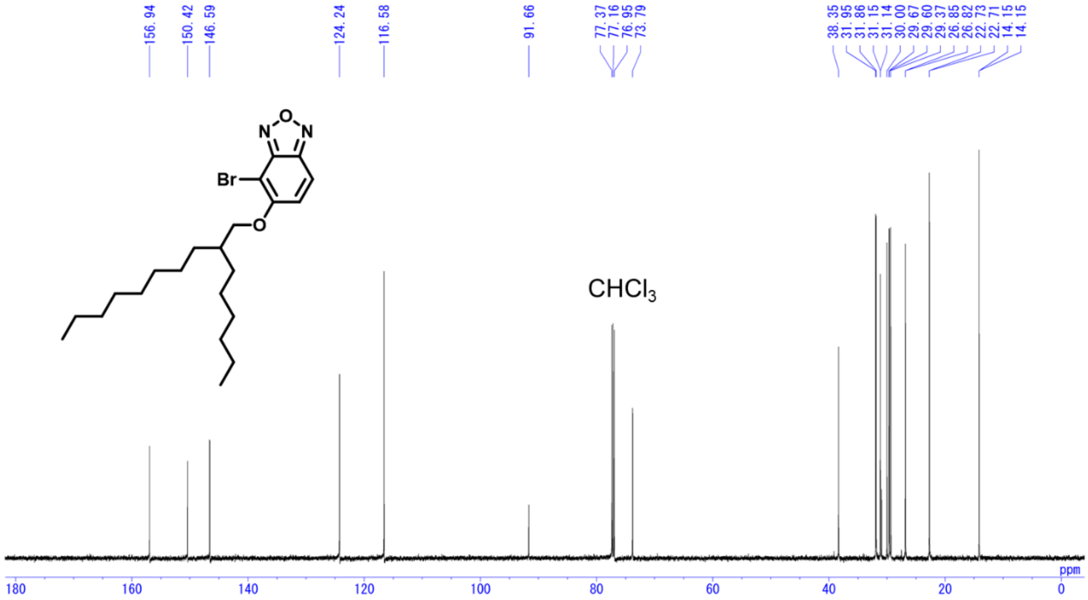


Figure S12. ^{13}C NMR spectrum of **4-a** (151 MHz, CDCl$_3$).

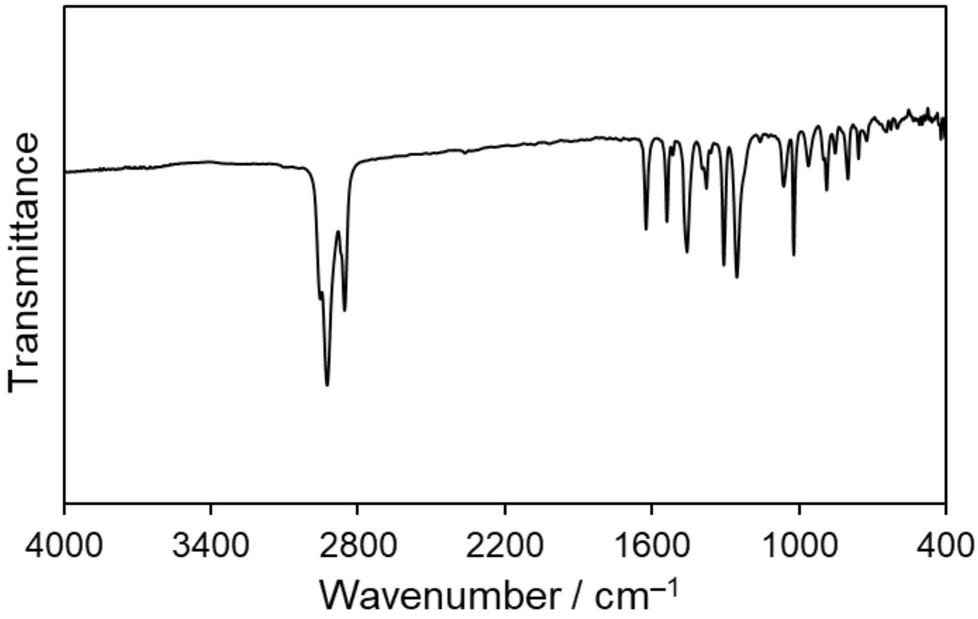
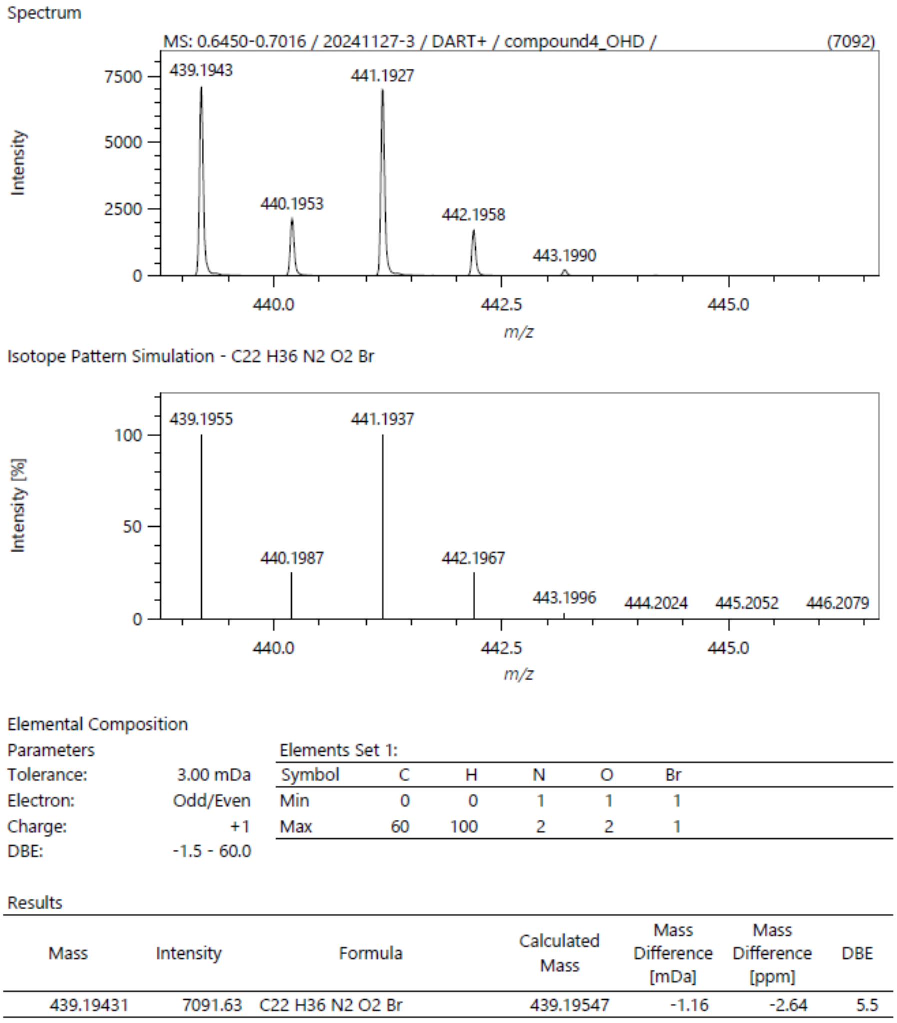


Figure S13. FT-IR spectrum of **4-a** (KBr).



Elemental Composition

Parameters	
Tolerance:	3.00 mDa
Electron:	Odd/Even
Charge:	+1
DBE:	-1.5 - 60.0

Elements Set 1:

Symbol	C	H	N	O	Br
Min	0	0	1	1	1
Max	60	100	2	2	1

Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [mDa]	Mass Difference [ppm]	DBE
439.19431	7091.63	C22 H36 N2 O2 Br	439.19547	-1.16	-2.64	5.5

Figure S14. DART+ mass spectrum of **4-a**.

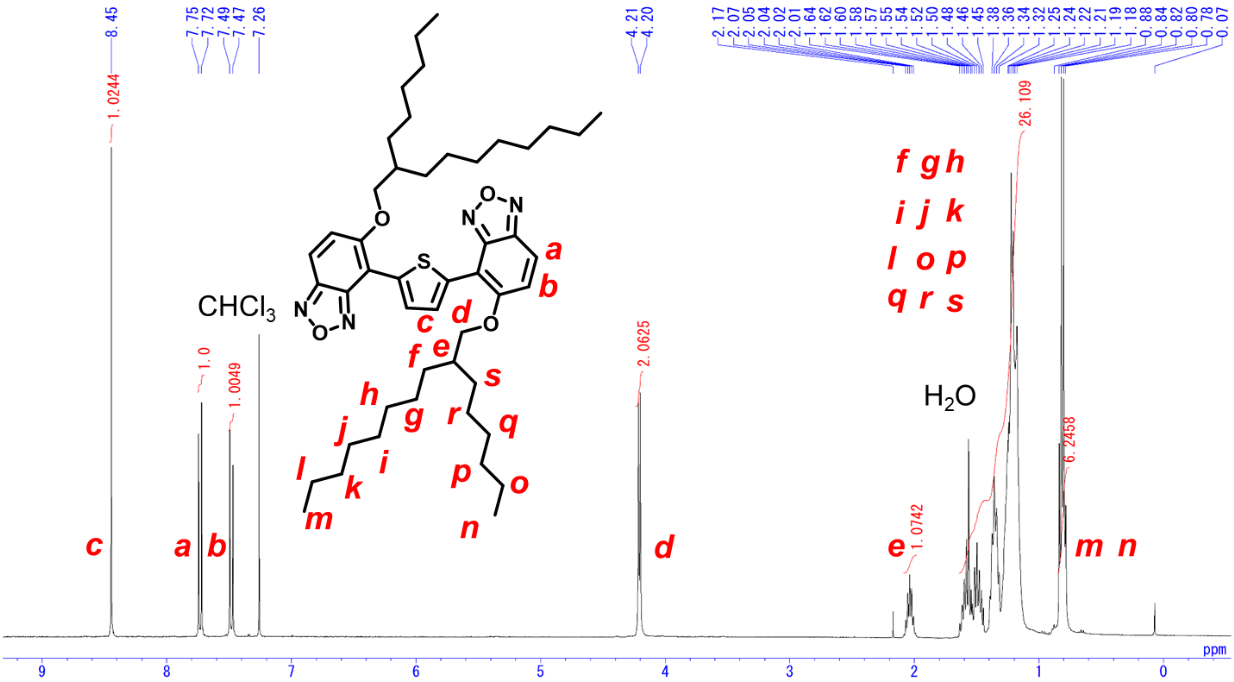


Figure S15. ^{1}H NMR spectrum of **5-a** (400 MHz, CDCl$_3$).

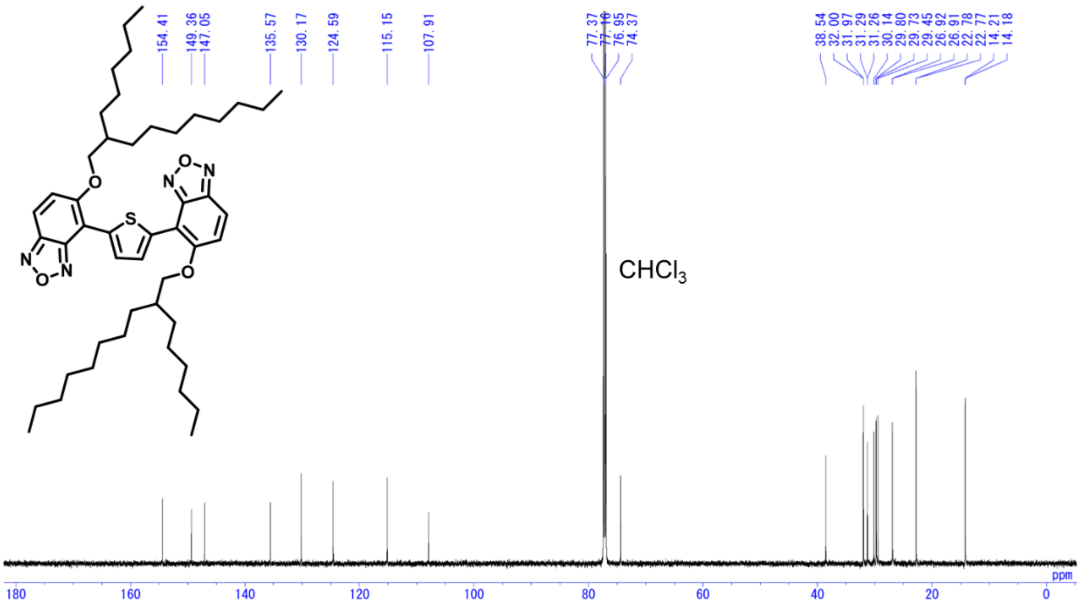


Figure S16. ^{13}C NMR spectrum of **5-a** (151 MHz, CDCl$_3$).

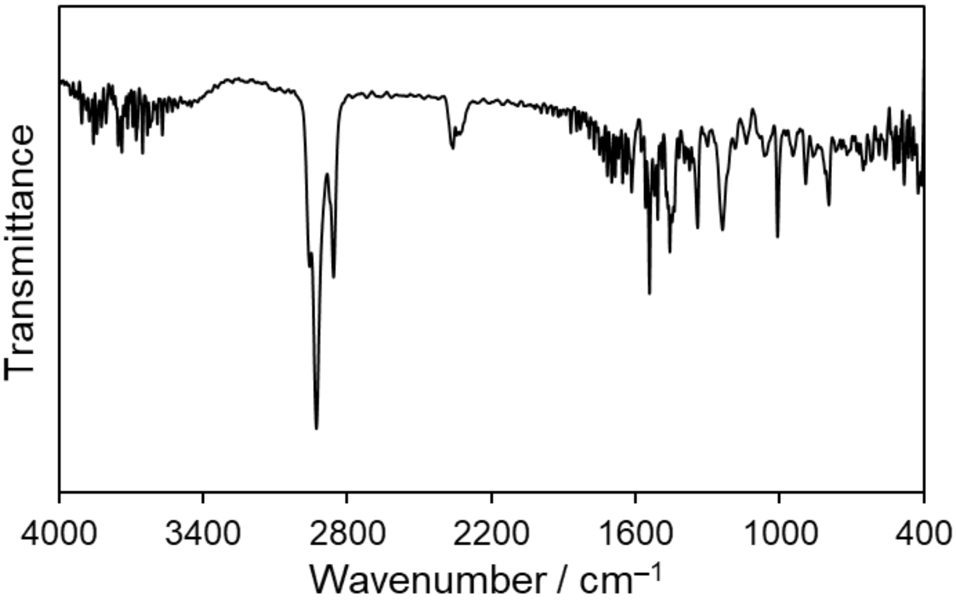
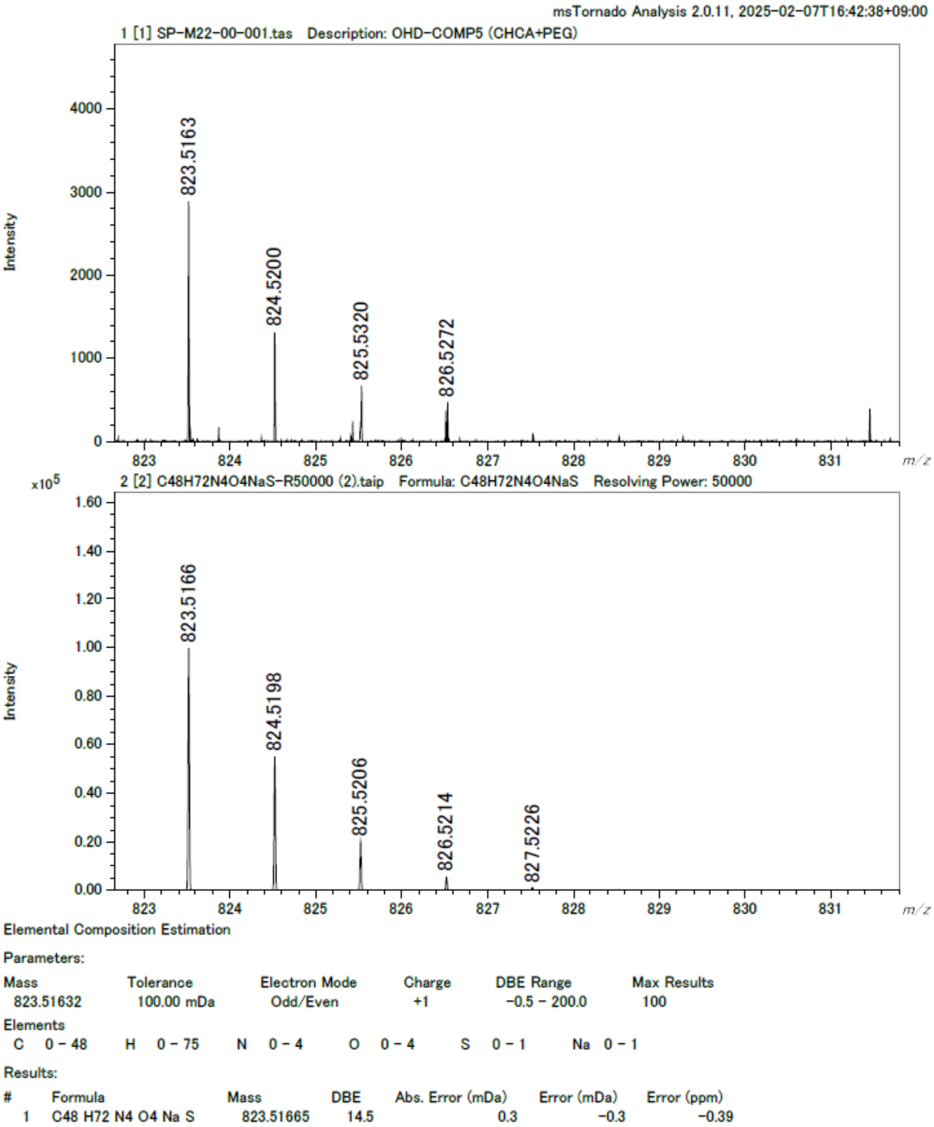


Figure S17. FT-IR spectrum of **5-a** (KBr).



```
Elemental Composition Estimation

Parameters:
Mass          Tolerance       Electron Mode    Charge    DBE Range        Max Results
 823.51632     100.00 mDa      Odd/Even          +1        -0.5 - 200.0     100
Elements
 C   0 - 48    H   0 - 75    N   0 - 4    O   0 - 4    S   0 - 1    Na  0 - 1

Results:
#   Formula                Mass        DBE    Abs. Error (mDa)   Error (mDa)   Error (ppm)
 1  C48 H72 N4 O4 Na S      823.51665   14.5               0.3          -0.3         -0.39
```

Figure S18. MALDI-TOF mass spectrum of **5-a**.

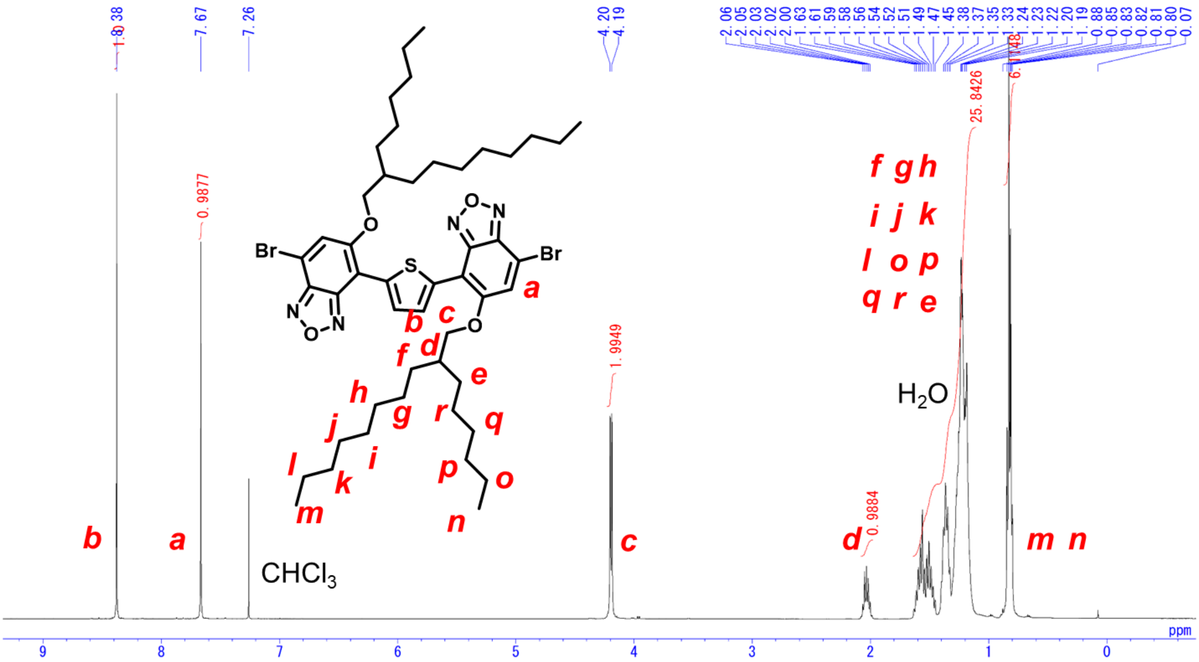


Figure S19. ^{1}H NMR spectrum of **6-a** (400 MHz, $CDCl_3$).

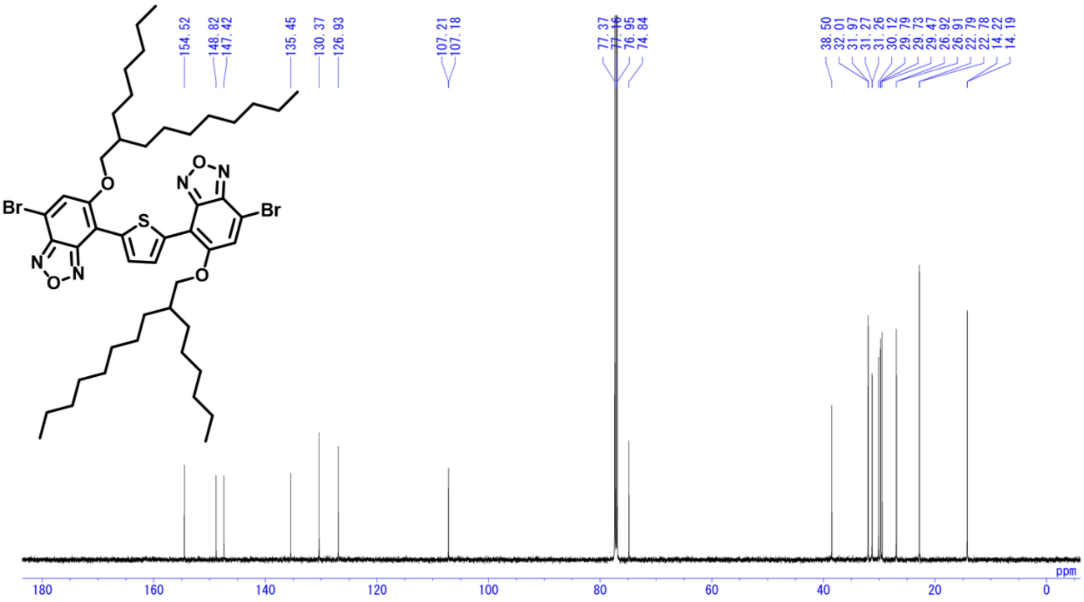


Figure S20. ^{13}C NMR spectrum of **6-a** (151 MHz, $CDCl_3$).

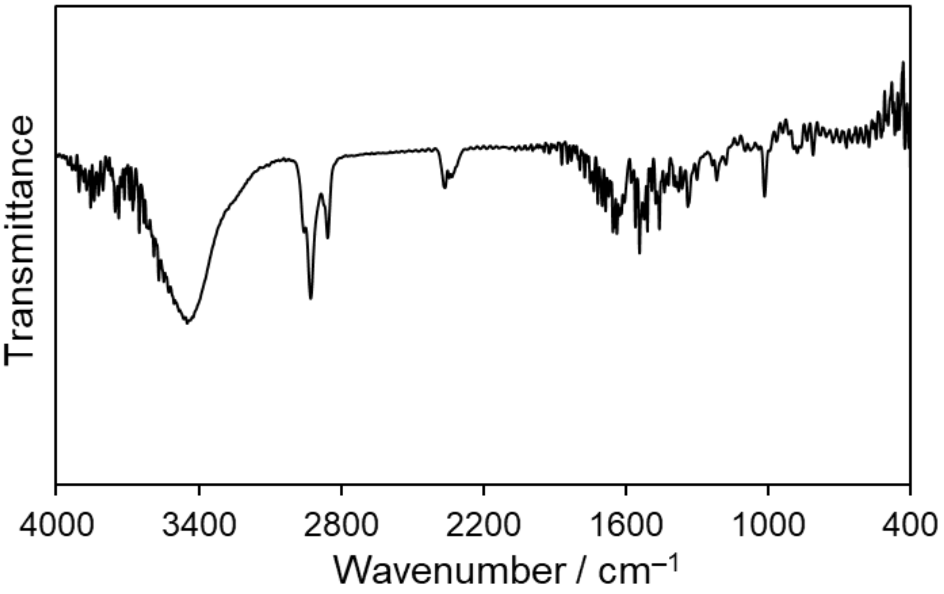
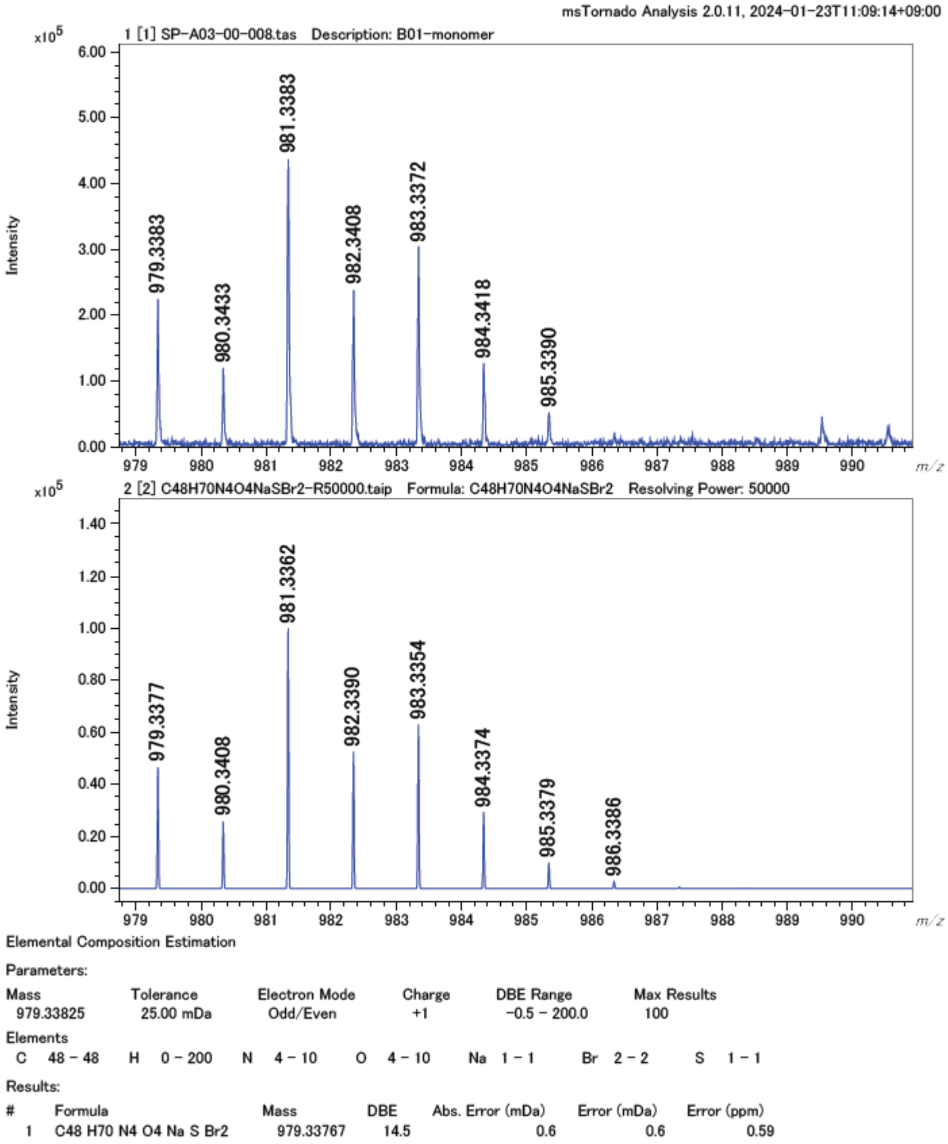


Figure S21. FT-IR spectrum of **6-a** (KBr).



```
Elemental Composition Estimation

Parameters:
Mass          Tolerance       Electron Mode     Charge     DBE Range        Max Results
 979.33825     25.00 mDa       Odd/Even          +1          -0.5 - 200.0      100

Elements
 C   48 - 48     H   0 - 200     N   4 - 10     O   4 - 10     Na   1 - 1     Br   2 - 2     S   1 - 1

Results:
#     Formula                          Mass        DBE    Abs. Error (mDa)    Error (mDa)    Error (ppm)
  1    C48 H70 N4 O4 Na S Br2           979.33767    14.5                 0.6            0.6            0.59
```

Figure S22. MALDI-TOF mass spectrum of **6-a**.

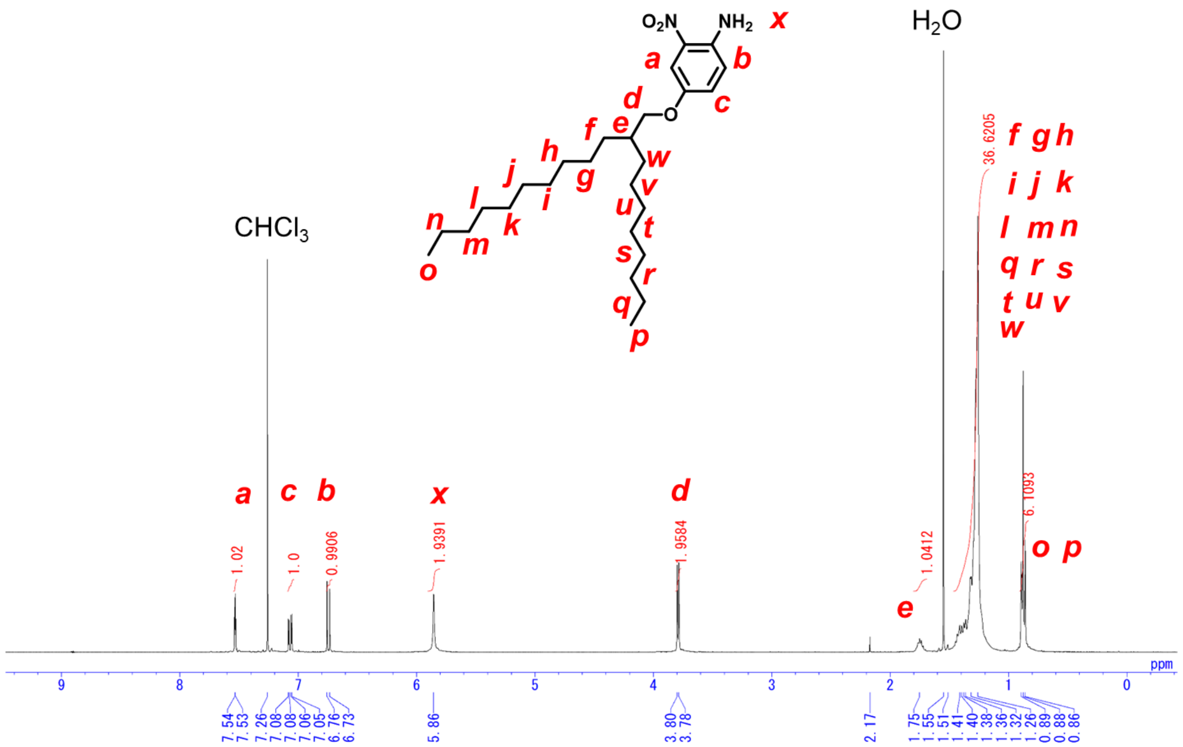


Figure S23. ^{1}H NMR spectrum of **1-b** (600 MHz, $CDCl_3$).

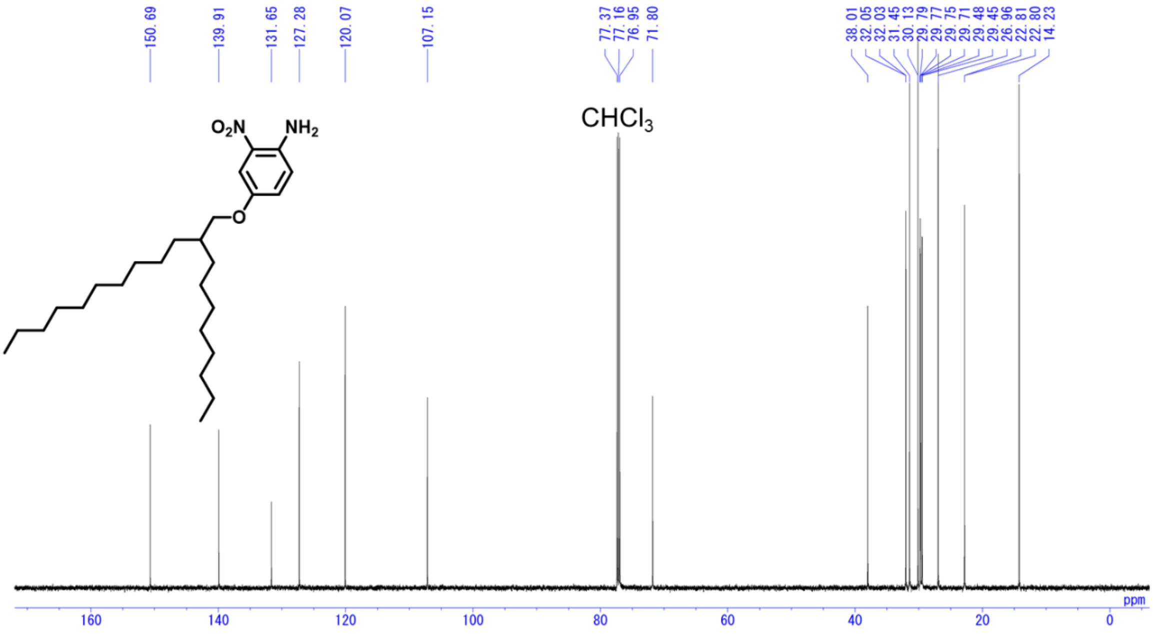


Figure S24. ^{13}C NMR spectrum of **1-b** (151 MHz, $CDCl_3$).

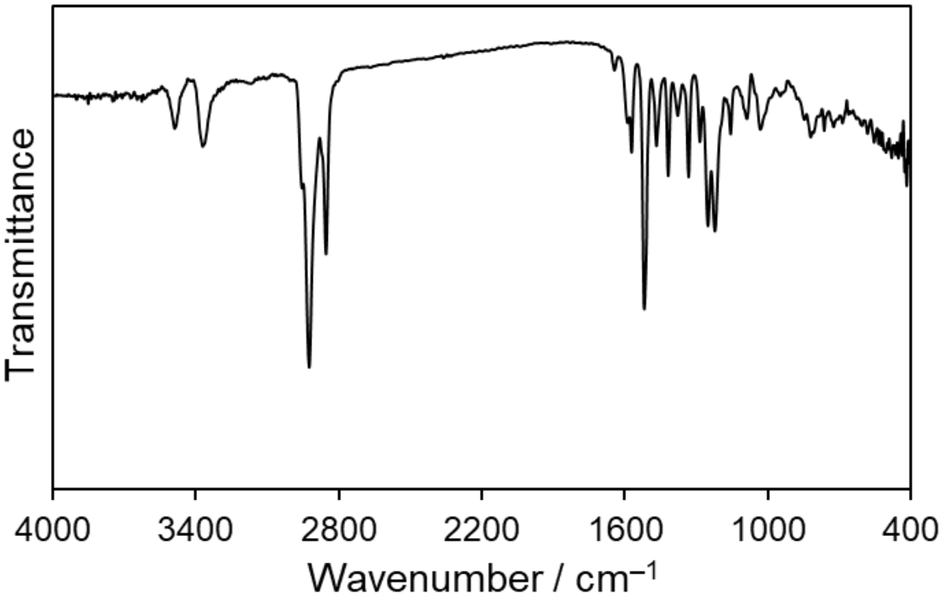
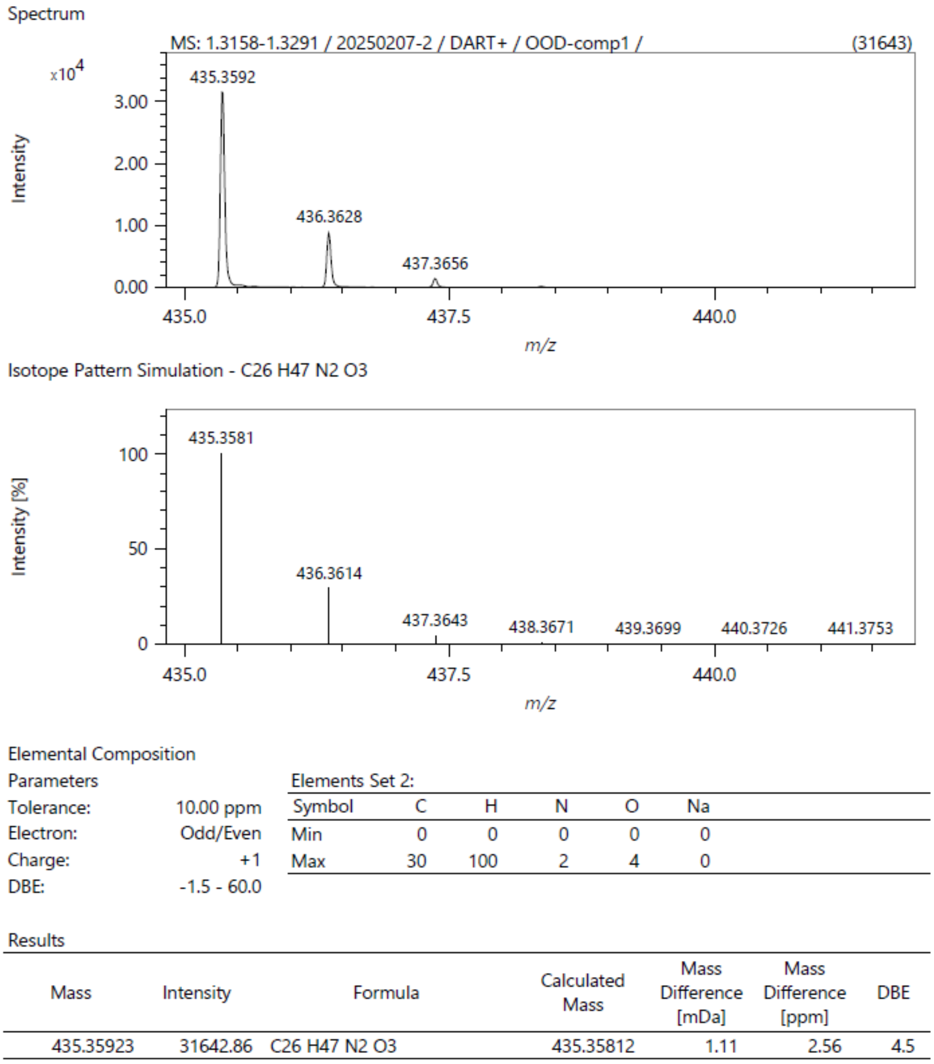


Figure S25. FT-IR spectrum of **1-b** (KBr).



Isotope Pattern Simulation - C26 H47 N2 O3

Elemental Composition

Parameters	
Tolerance:	10.00 ppm
Electron:	Odd/Even
Charge:	+1
DBE:	-1.5 - 60.0

Elements Set 2:

Symbol	C	H	N	O	Na
Min	0	0	0	0	0
Max	30	100	2	4	0

Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [mDa]	Mass Difference [ppm]	DBE
435.35923	31642.86	C26 H47 N2 O3	435.35812	1.11	2.56	4.5

Figure S26. DART+ mass spectrum of **1-b**.

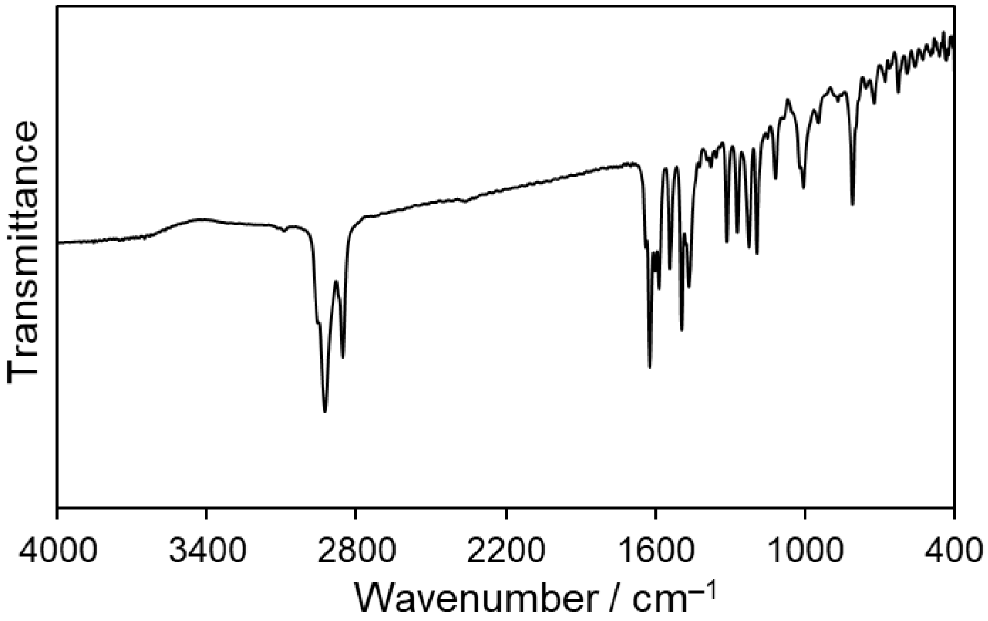
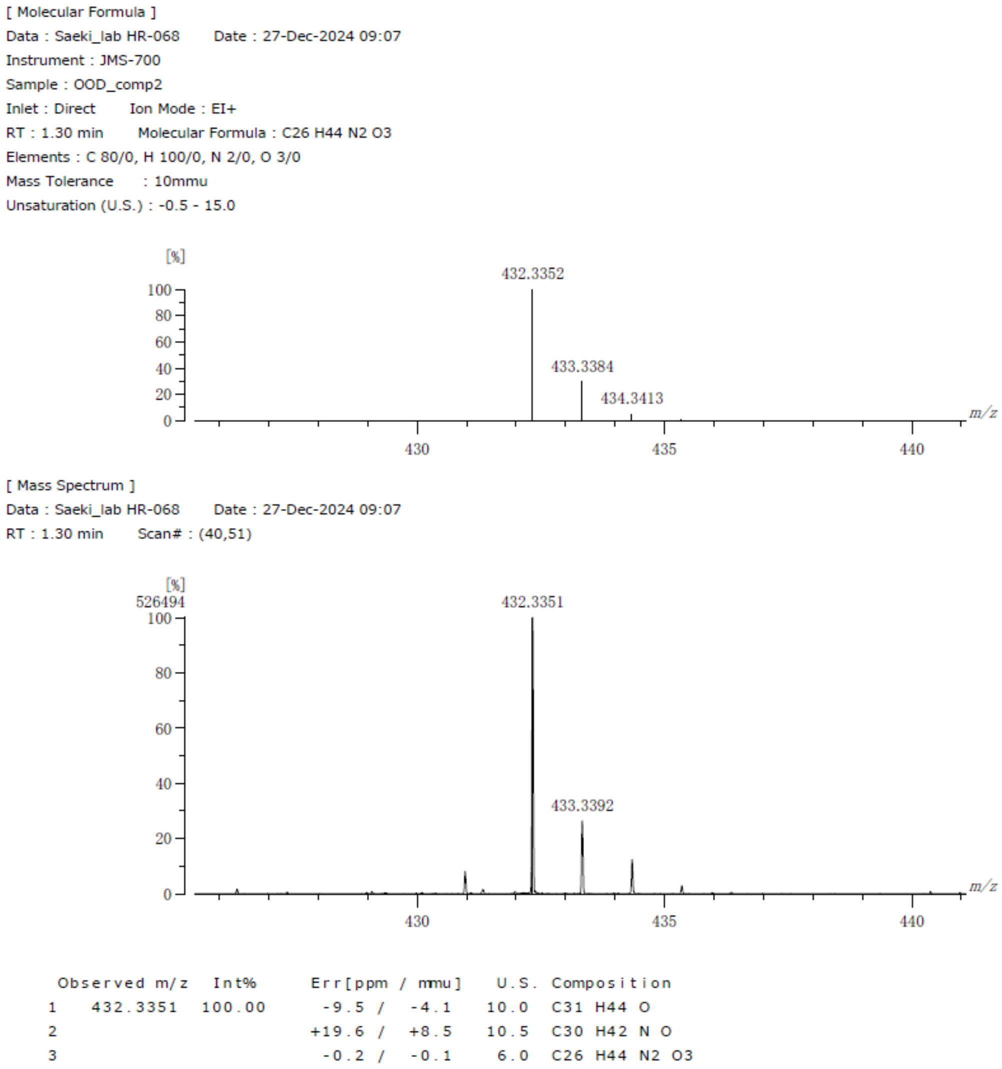


Figure S27. FT-IR spectrum of **2-b** (KBr).


```
[ Molecular Formula ]
Data : Saeki_lab HR-068      Date : 27-Dec-2024 09:07
Instrument : JMS-700
Sample : OOD_comp2
Inlet : Direct      Ion Mode : EI+
RT : 1.30 min      Molecular Formula : C26 H44 N2 O3
Elements : C 80/0, H 100/0, N 2/0, O 3/0
Mass Tolerance      : 10mmu
Unsaturation (U.S.) : -0.5 - 15.0
```

```
[ Mass Spectrum ]
Data : Saeki_lab HR-068      Date : 27-Dec-2024 09:07
RT : 1.30 min      Scan# : (40,51)
```

```
 Observed m/z   Int%      Err[ppm / mmu]    U.S.  Composition
1   432.3351   100.00      -9.5 /  -4.1    10.0  C31 H44 O
2                          +19.6 /  +8.5    10.5  C30 H42 N O
3                          -0.2 /  -0.1     6.0  C26 H44 N2 O3
```

Figure S28. EI+ mass spectrum of **2-b**.

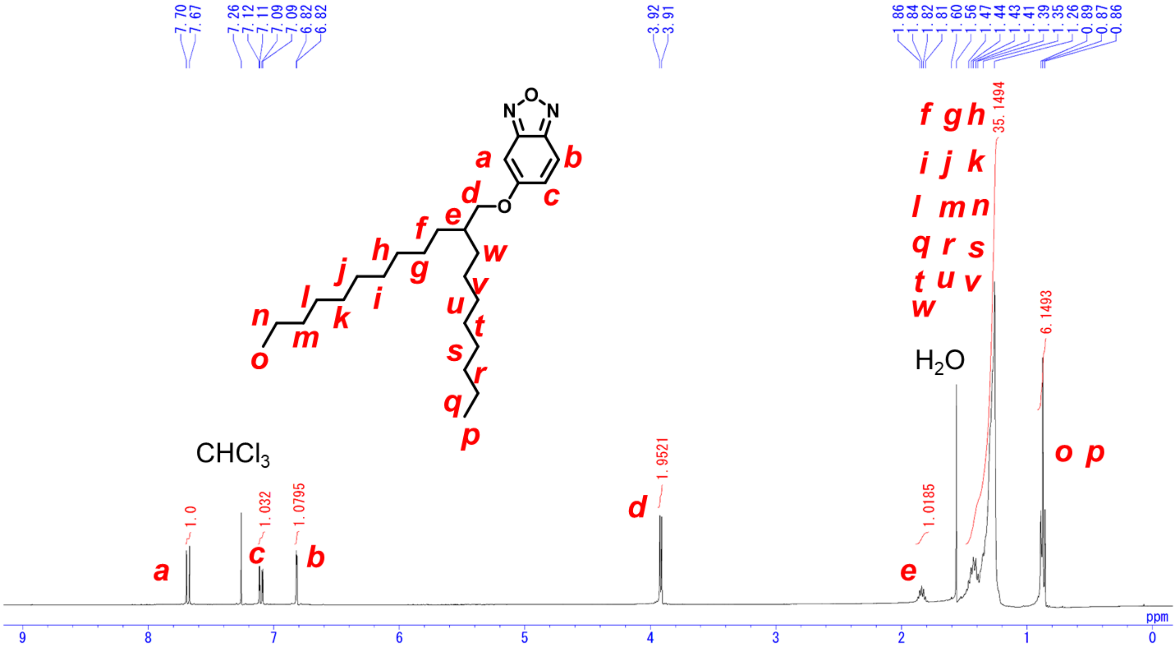


Figure S29. ^{1}H NMR spectrum of **3-b** (400 MHz, CDCl$_3$).

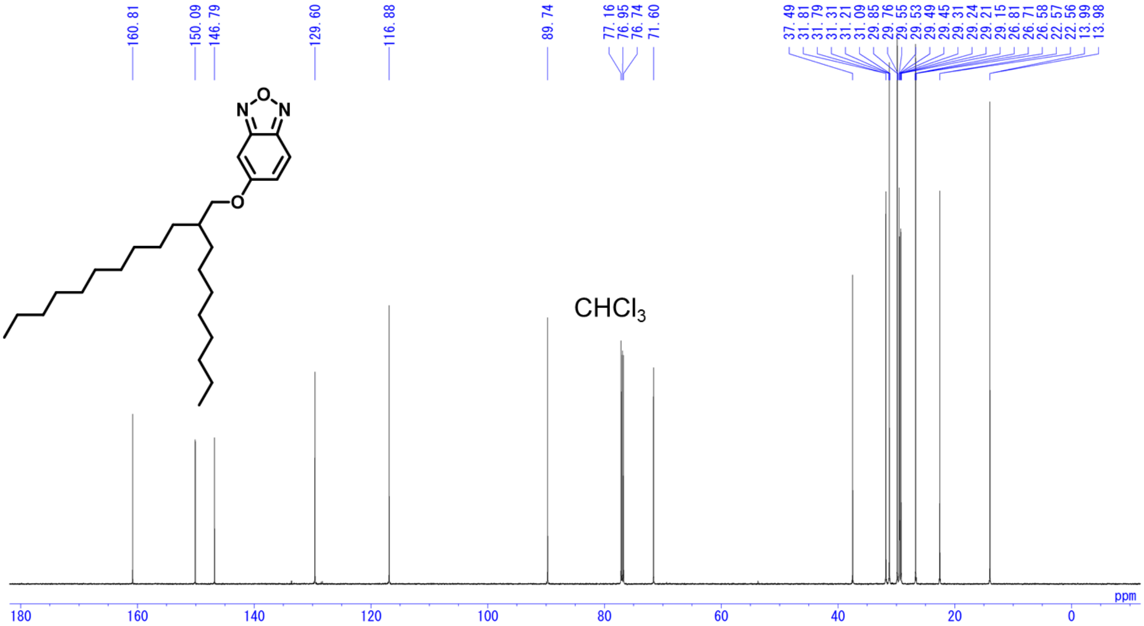


Figure S30. ^{13}C NMR spectrum of **3-b** (151 MHz, CDCl$_3$).

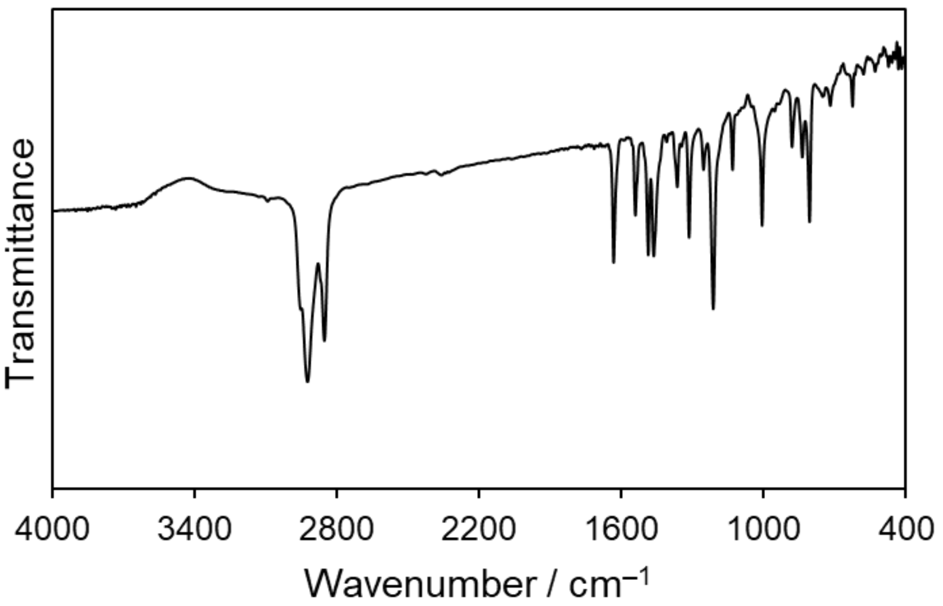
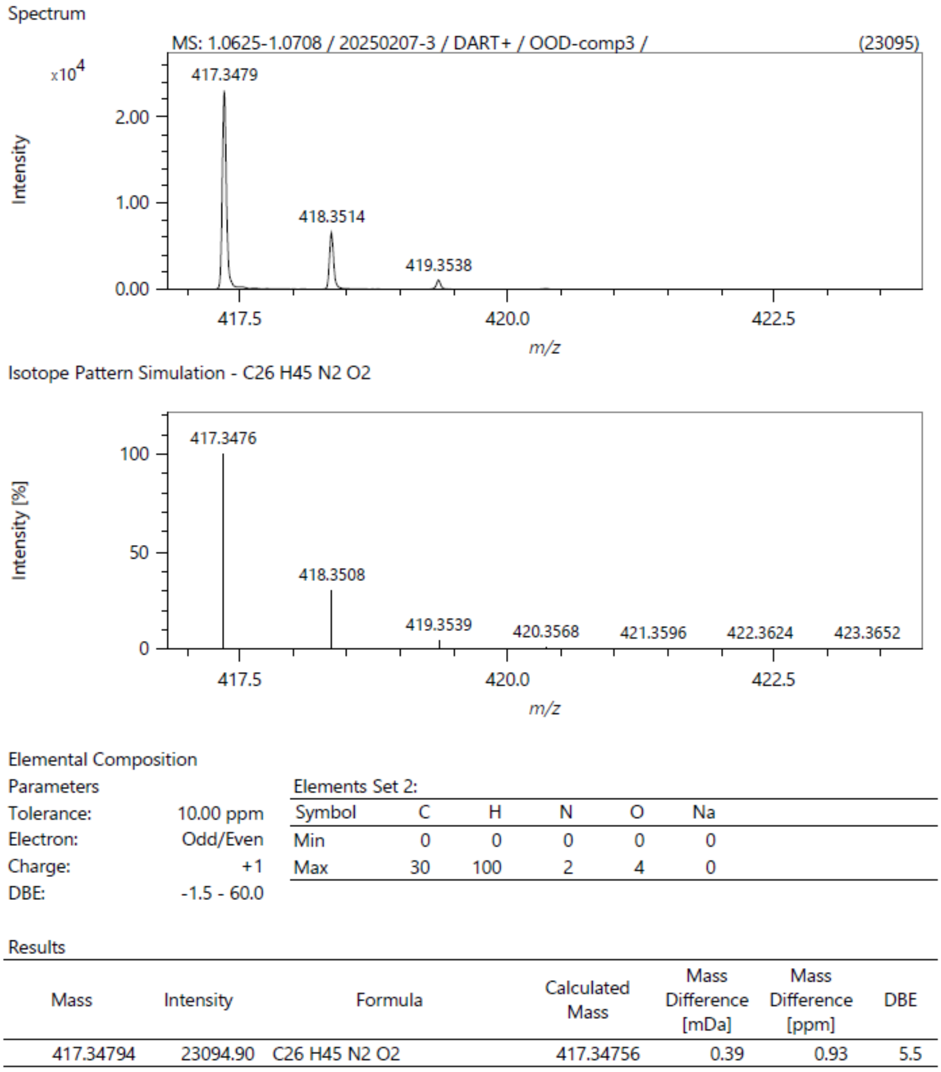


Figure S31. FT-IR spectrum of **3-b** (KBr).



Elemental Composition

Parameters

Tolerance:	10.00 ppm
Electron:	Odd/Even
Charge:	+1
DBE:	-1.5 - 60.0

Elements Set 2:

Symbol	C	H	N	O	Na
Min	0	0	0	0	0
Max	30	100	2	4	0

Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [mDa]	Mass Difference [ppm]	DBE
417.34794	23094.90	C26 H45 N2 O2	417.34756	0.39	0.93	5.5

Figure S32. DART+ mass spectrum of **3-b**.

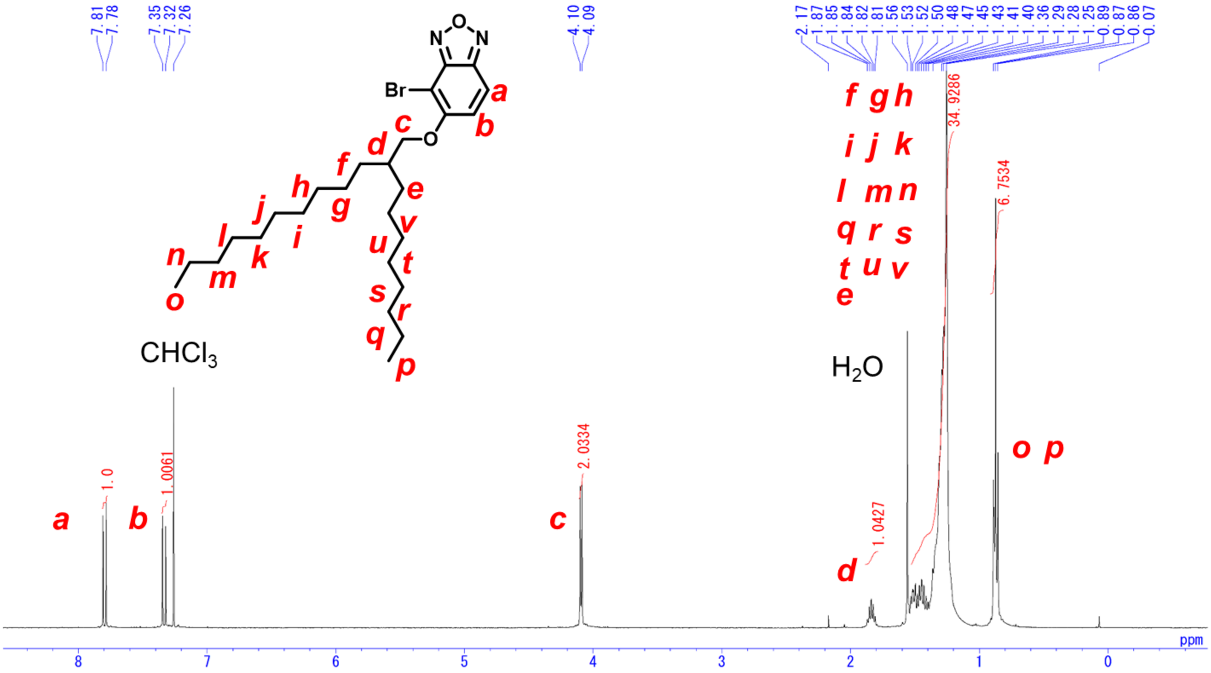


Figure S33. ^{1}H NMR spectrum of **4-b** (400 MHz, CDCl$_3$).

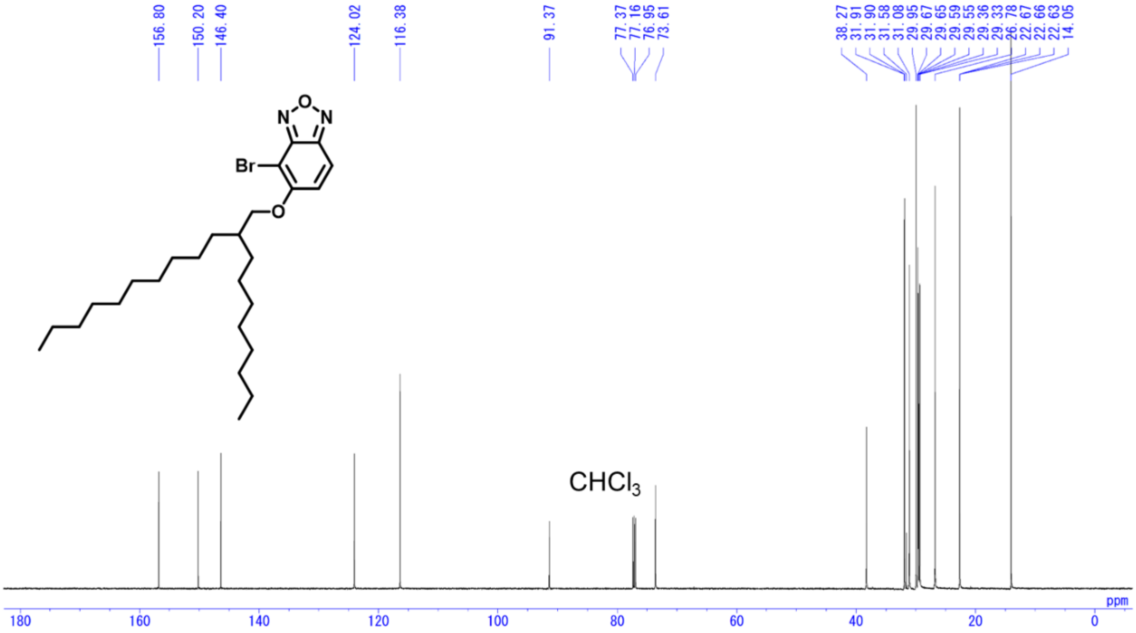


Figure S34. ^{13}C NMR spectrum of **4-b** (151 MHz, CDCl$_3$).

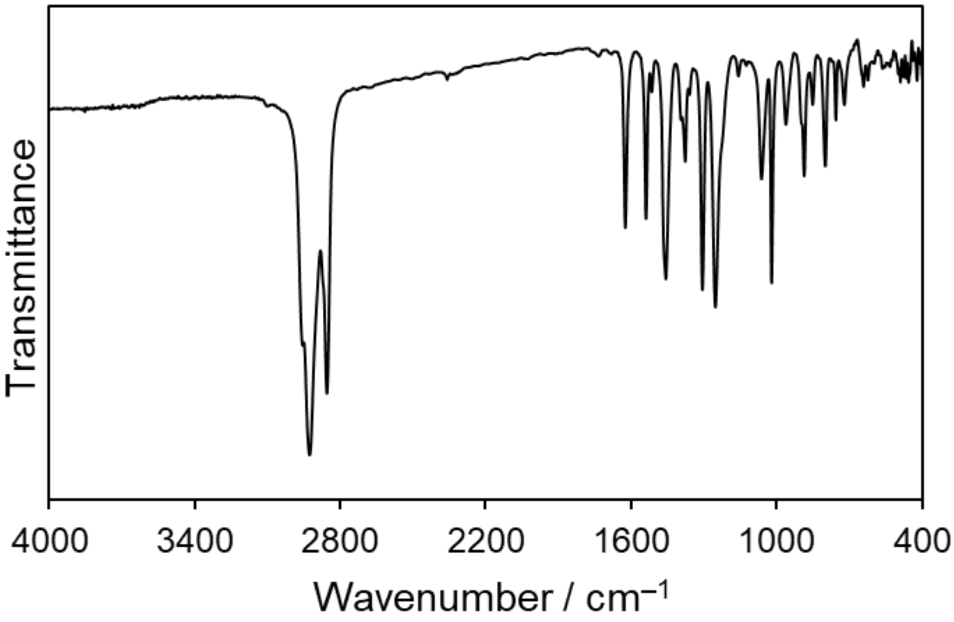
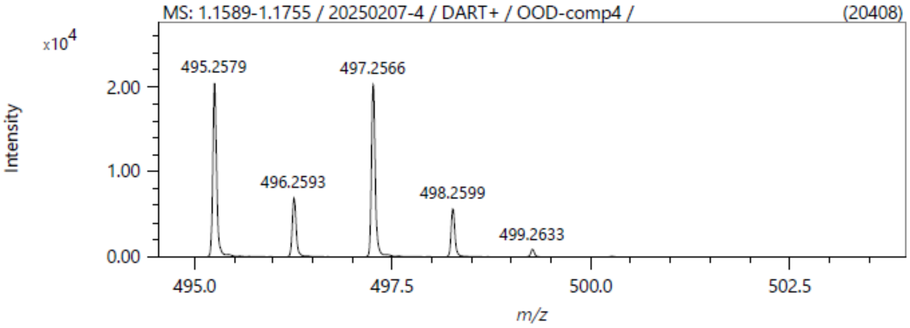
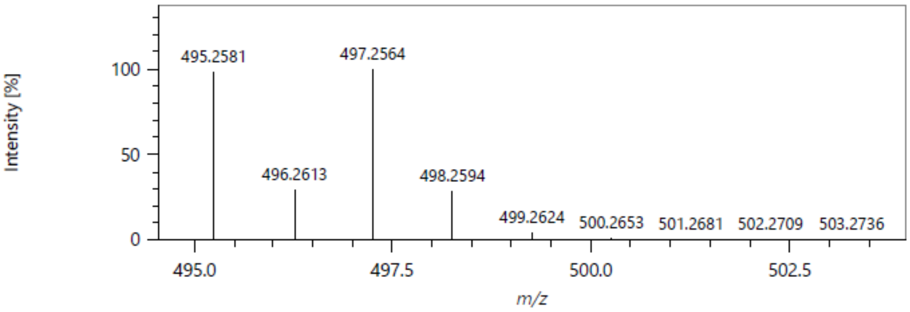


Figure S35. FT-IR spectrum of **4-b** (KBr).

Spectrum



Isotope Pattern Simulation - C26 H44 N2 O2 Br



Elemental Composition

Parameters

Tolerance:	10.00 ppm
Electron:	Odd/Even
Charge:	+1
DBE:	-1.5 - 60.0

Elements Set 2:

Symbol	C	H	N	O	Br
Min	0	0	0	0	0
Max	30	100	2	4	1

Results

Mass	Intensity	Formula	Calculated Mass	Mass Difference [mDa]	Mass Difference [ppm]	DBE
495.25791	20408.49	C26 H44 N2 O2 Br	495.25807	-0.15	-0.31	5.5

Figure S36. DART+ mass spectrum of **4-b**.

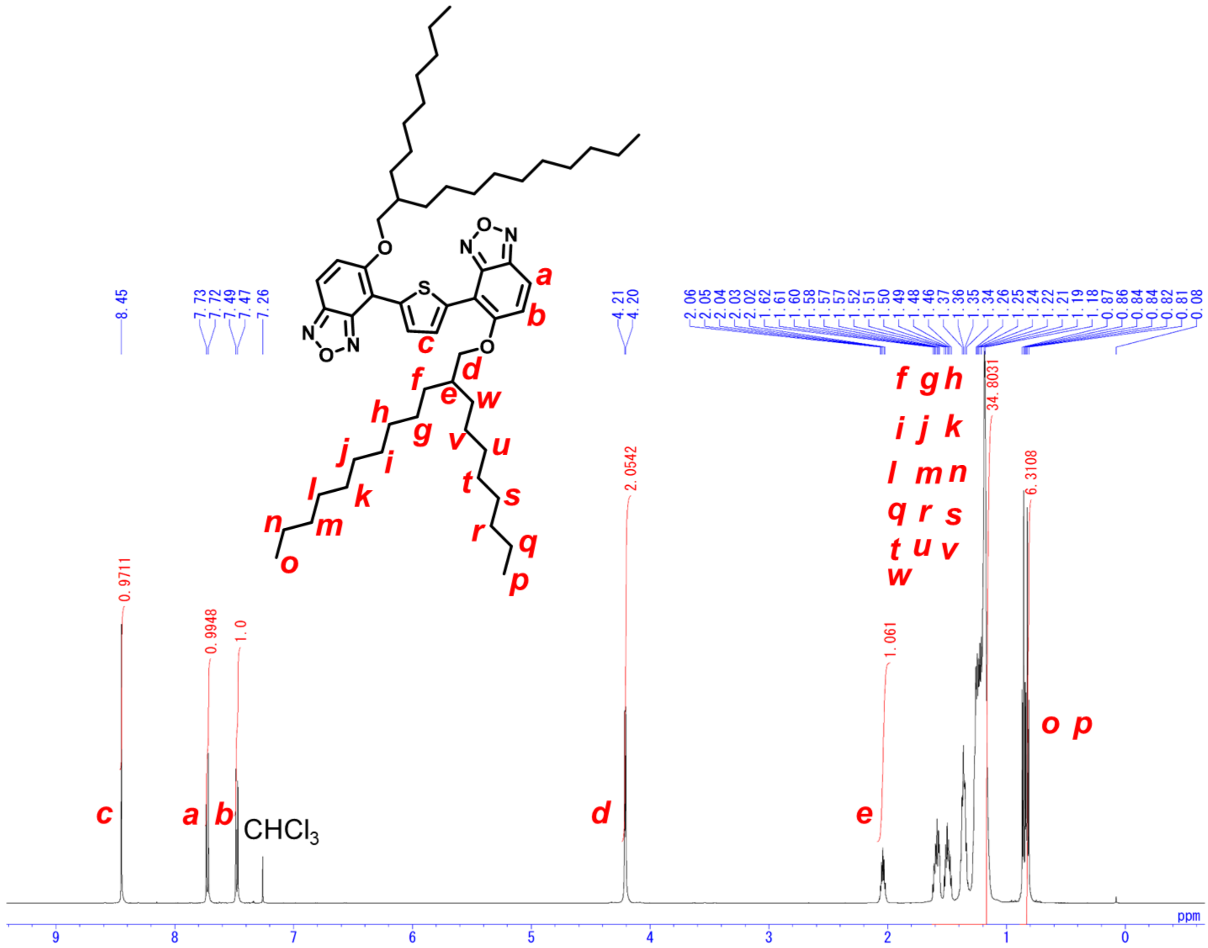


Figure S37. ^{1}H NMR spectrum of **5-b** (400 MHz, CDCl$_3$).

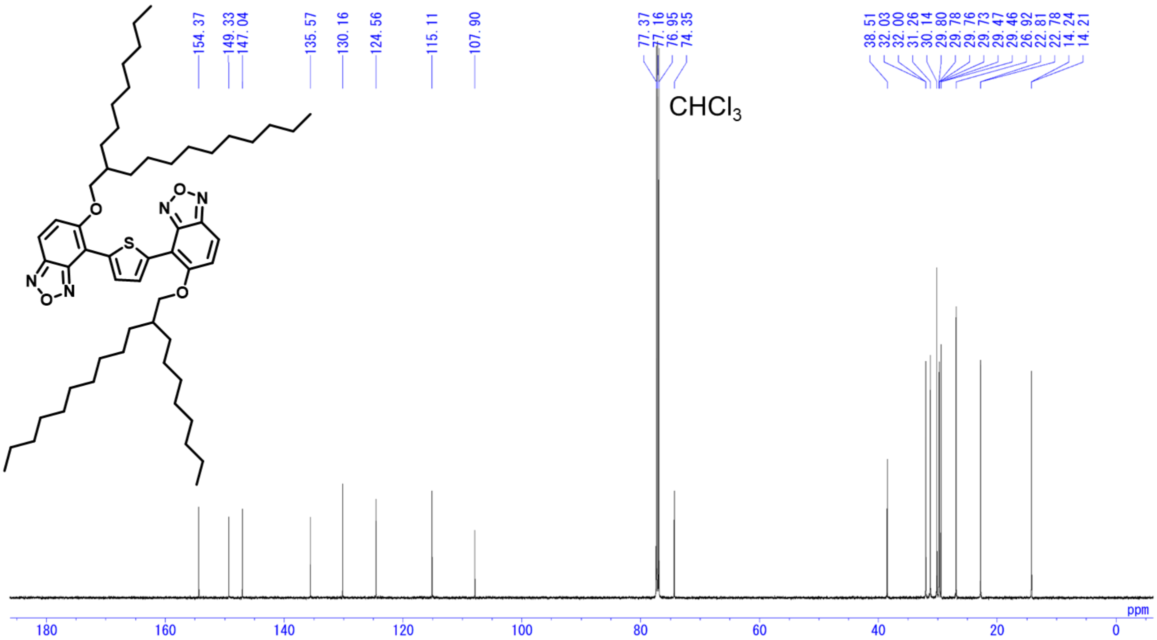


Figure S38. ^{13}C NMR spectrum of **5-b** (151 MHz, CDCl$_3$).

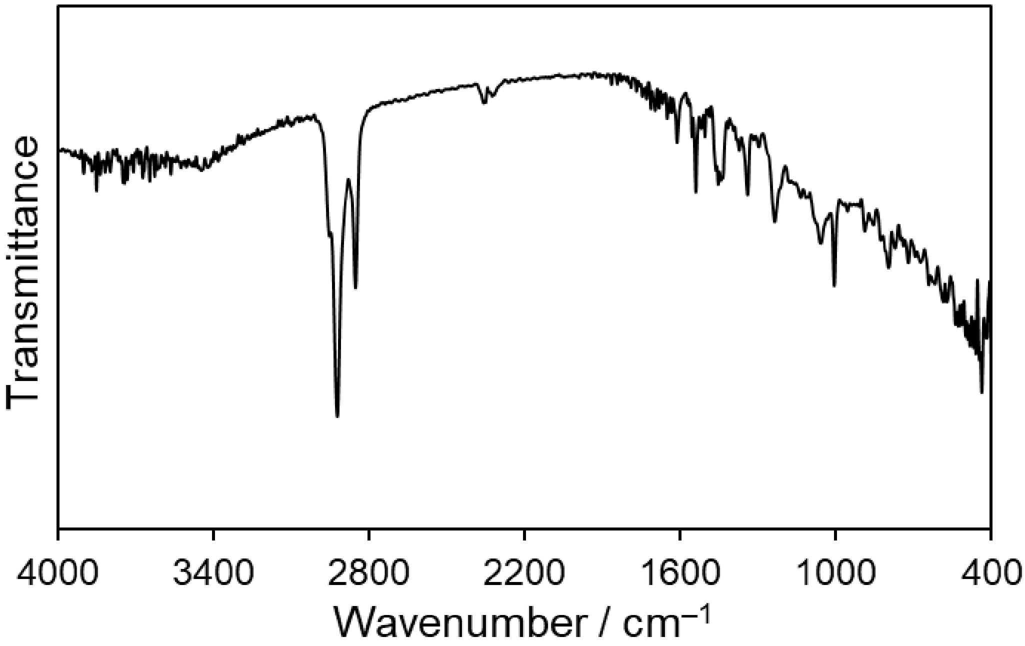
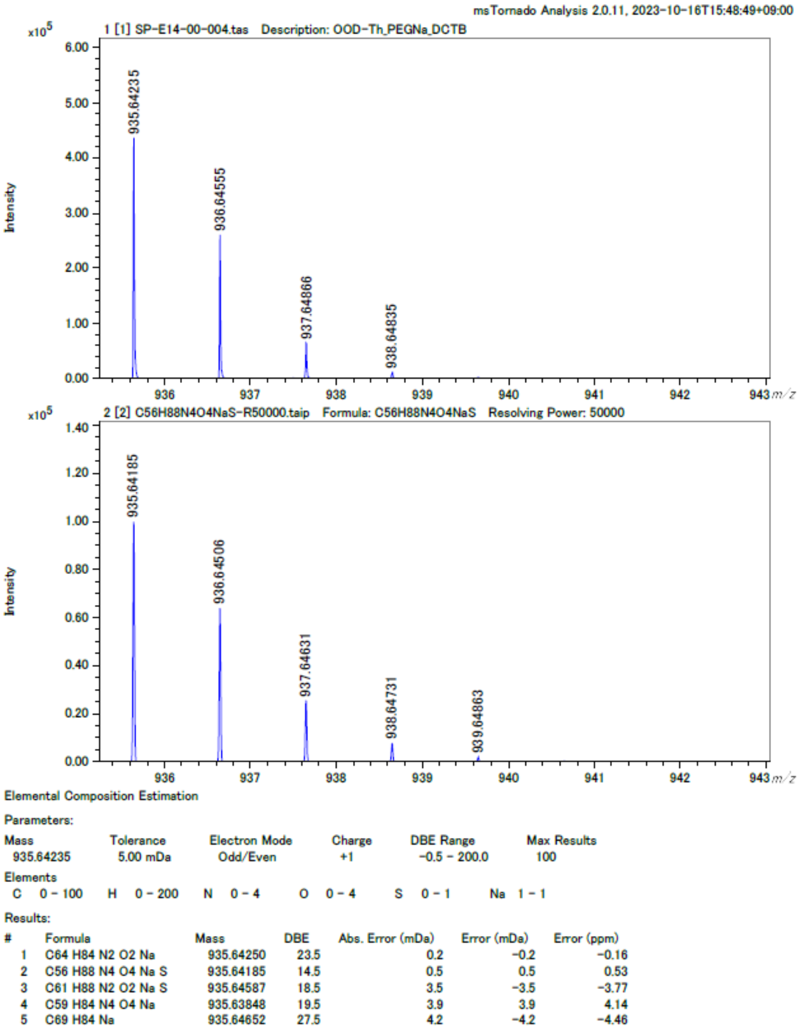


Figure S39. FT-IR spectrum of **5-b** (KBr).



msTornado Analysis 2.0.11, 2023-10-16T15:48:49+09:00

1 [1] SP-E14-00-004.tas Description: OOD-Th_PEGNa_DCTB

2 [2] C56H88N4O4NaS-R50000.taip Formula: C56H88N4O4NaS Resolving Power: 50000

Elemental Composition Estimation

Parameters:

Mass	Tolerance	Electron Mode	Charge	DBE Range	Max Results
935.64235	5.00 mDa	Odd/Even	+1	−0.5 − 200.0	100

Elements

C 0 − 100 H 0 − 200 N 0 − 4 O 0 − 4 S 0 − 1 Na 1 − 1

Results:

#	Formula	Mass	DBE	Abs. Error (mDa)	Error (mDa)	Error (ppm)
1	C64 H84 N2 O2 Na	935.64250	23.5	0.2	−0.2	−0.16
2	C56 H88 N4 O4 Na S	935.64185	14.5	0.5	0.5	0.53
3	C61 H88 N2 O2 Na S	935.64587	18.5	3.5	−3.5	−3.77
4	C59 H84 N4 O4 Na	935.63848	19.5	3.9	3.9	4.14
5	C69 H84 Na	935.64652	27.5	4.2	−4.2	−4.46

Figure S40. MALDI-TOF mass spectrum of **5-b**.

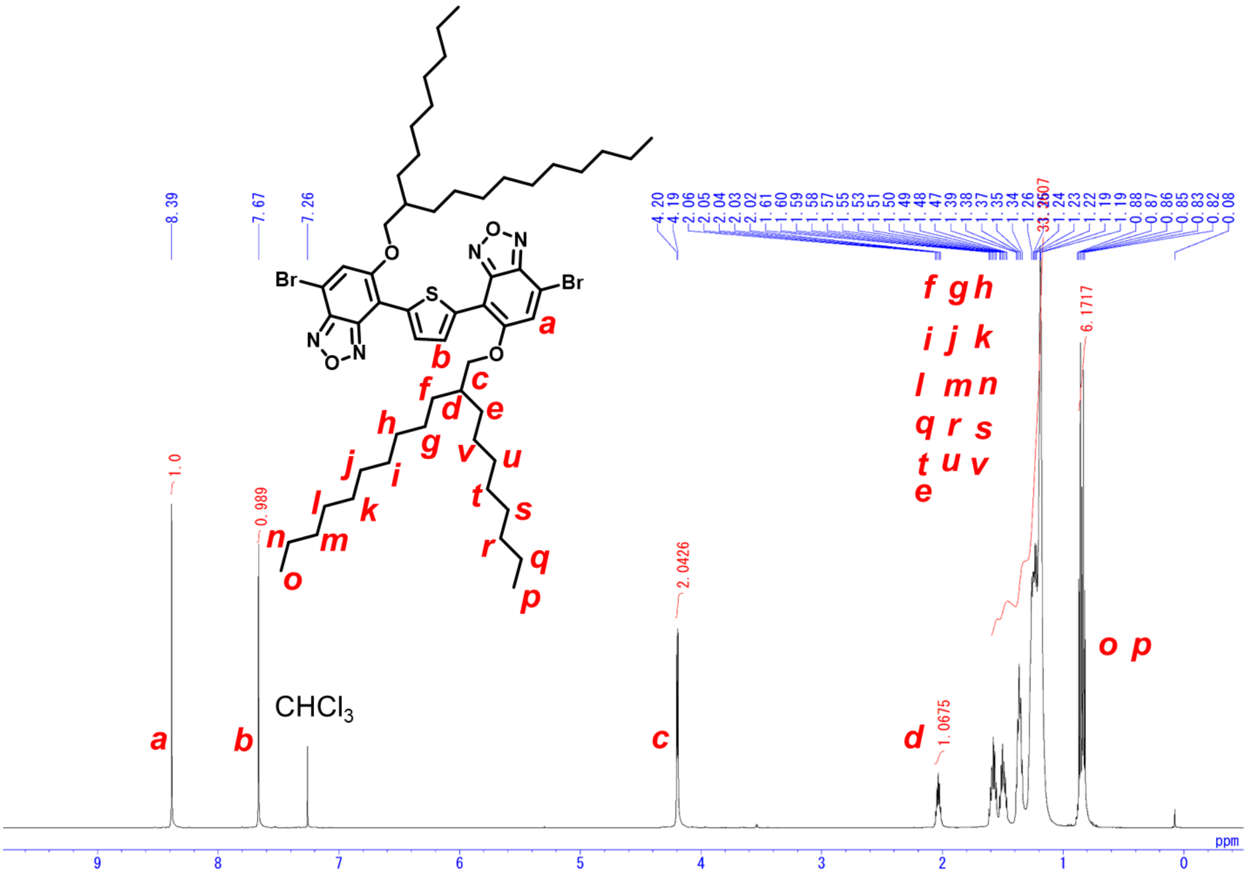


Figure S41. ^{1}H NMR spectrum of **6-b** (600 MHz, $CDCl_3$).

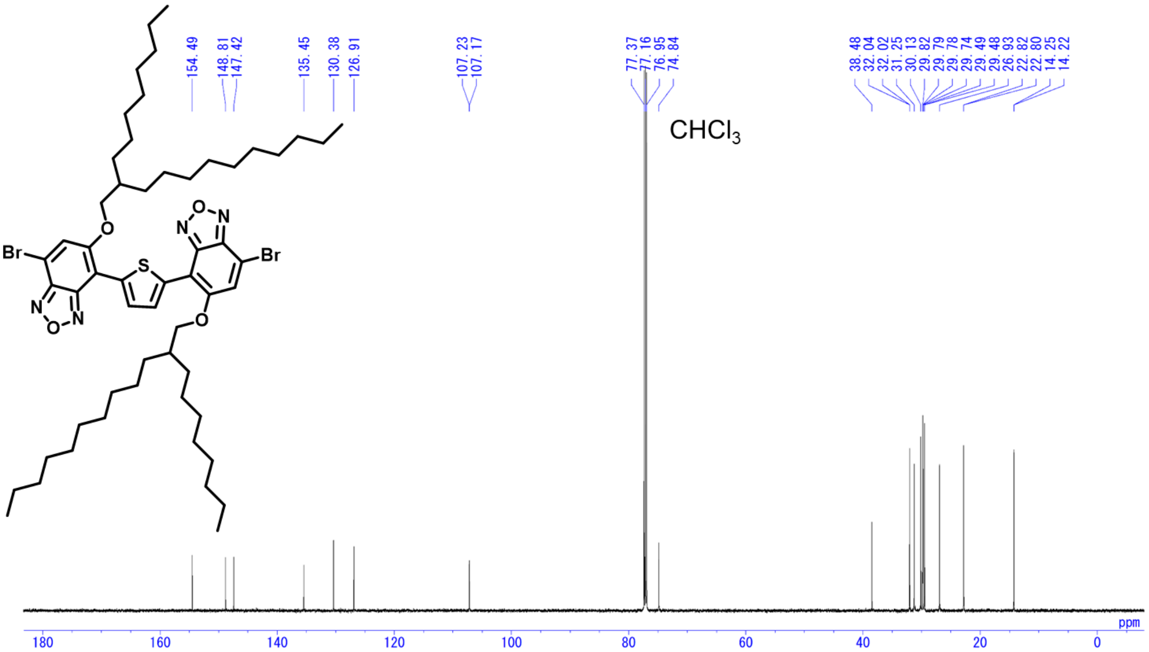


Figure S42. ^{13}C NMR spectrum of **6-b** (151 MHz, $CDCl_3$).

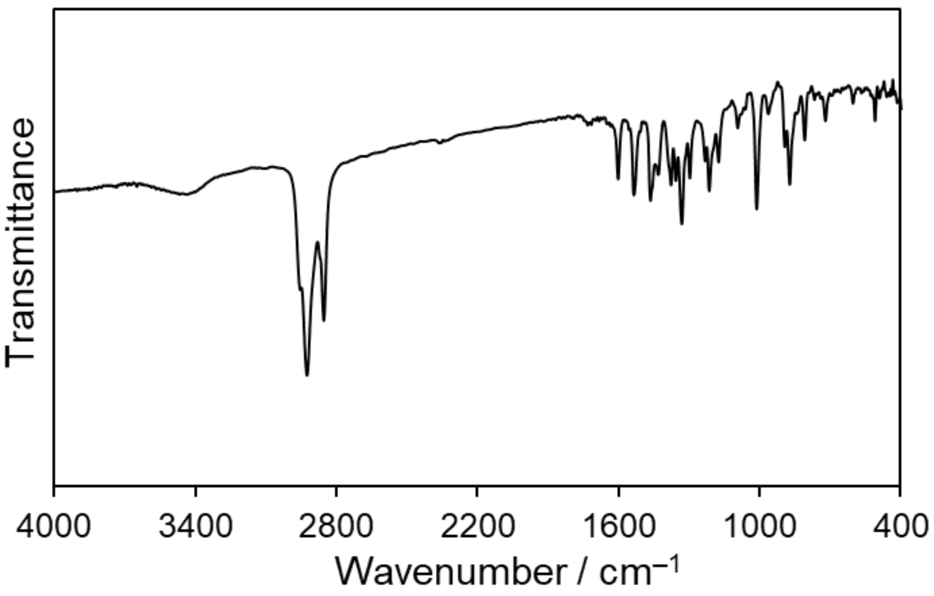
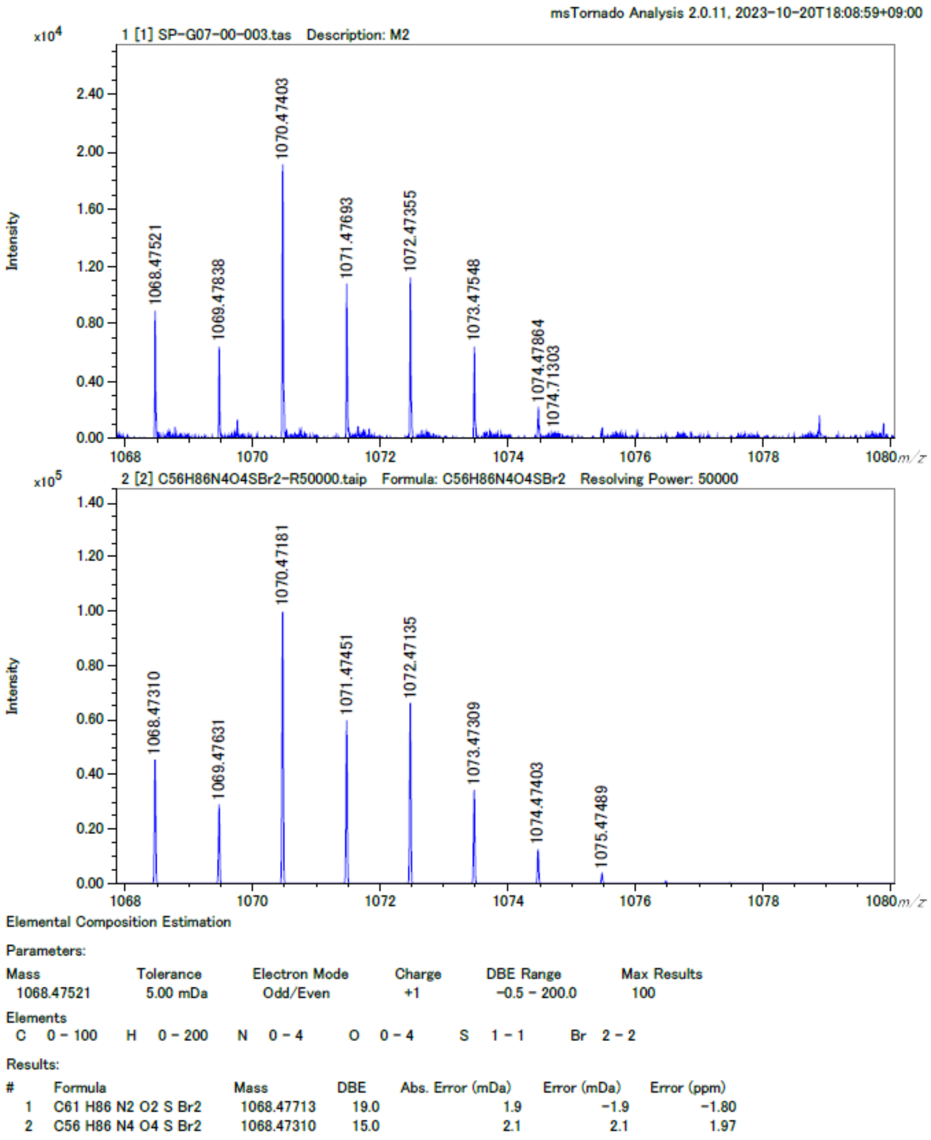


Figure S43. FT-IR spectrum of **6-b** (KBr).



```
Elemental Composition Estimation

Parameters:
Mass           Tolerance     Electron Mode    Charge    DBE Range       Max Results
 1068.47521     5.00 mDa      Odd/Even          +1       -0.5 - 200.0     100

Elements
 C   0 - 100    H   0 - 200    N   0 - 4     O   0 - 4     S   1 - 1     Br  2 - 2

Results:
#     Formula                  Mass          DBE     Abs. Error (mDa)    Error (mDa)    Error (ppm)
 1    C61 H86 N2 O2 S Br2      1068.47713    19.0                 1.9           -1.9          -1.80
 2    C56 H86 N4 O4 S Br2      1068.47310    15.0                 2.1            2.1           1.97
```

Figure S44. MALDI-TOF mass spectrum of **6-b**.

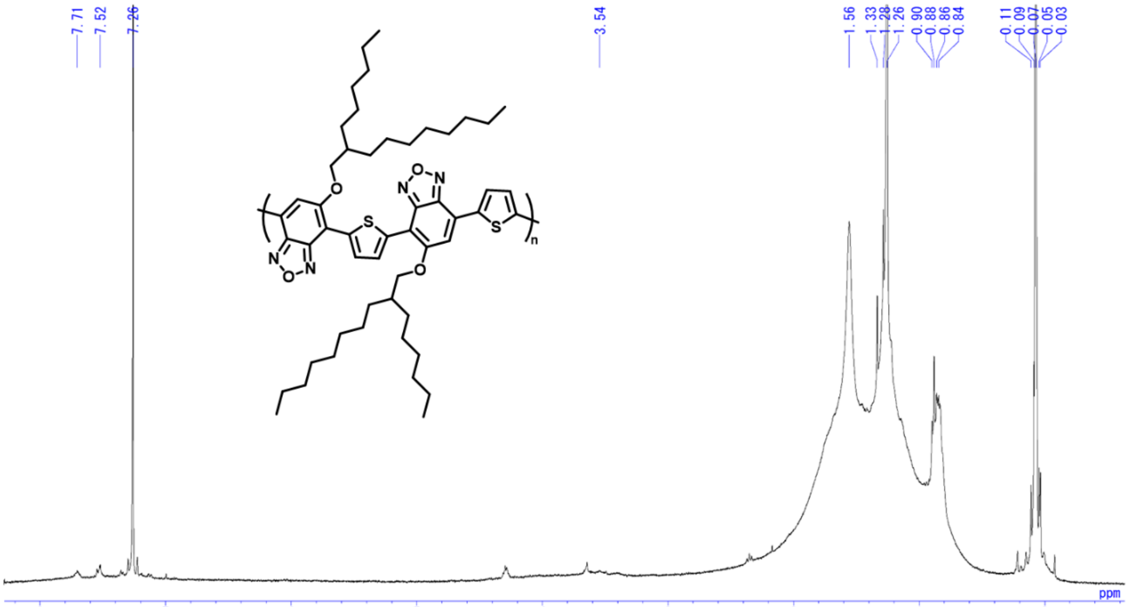


Figure S45. ^{1}H NMR spectrum of **BO1** (600 MHz, $CDCl_3$).

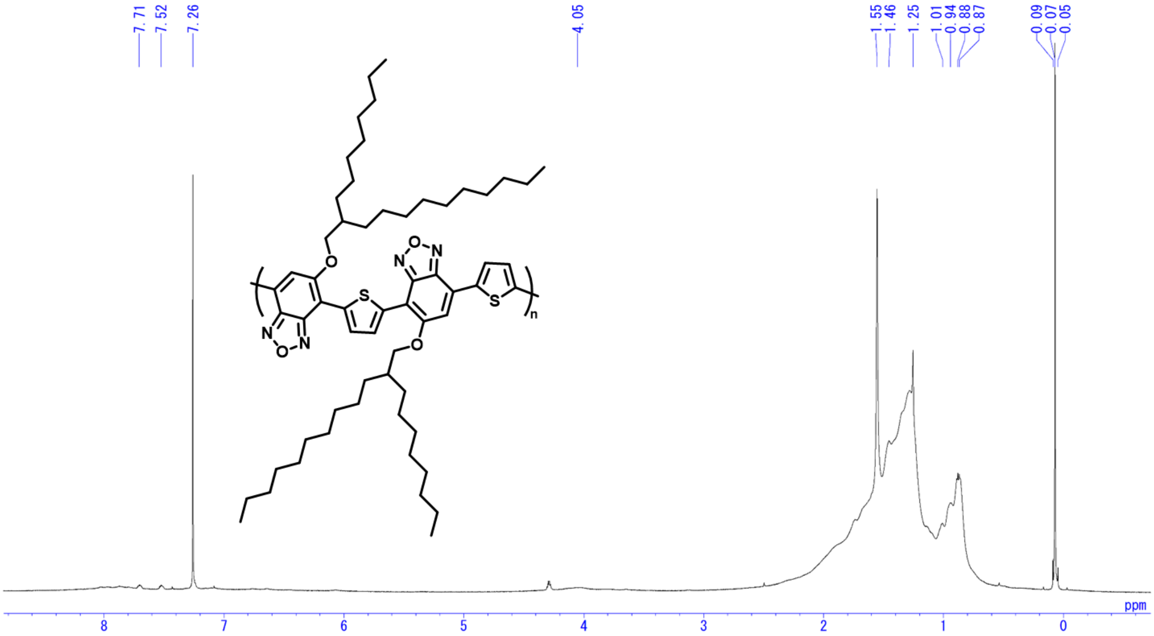


Figure S46. ^{1}H NMR spectrum of **BO2** (600 MHz, $CDCl_3$).

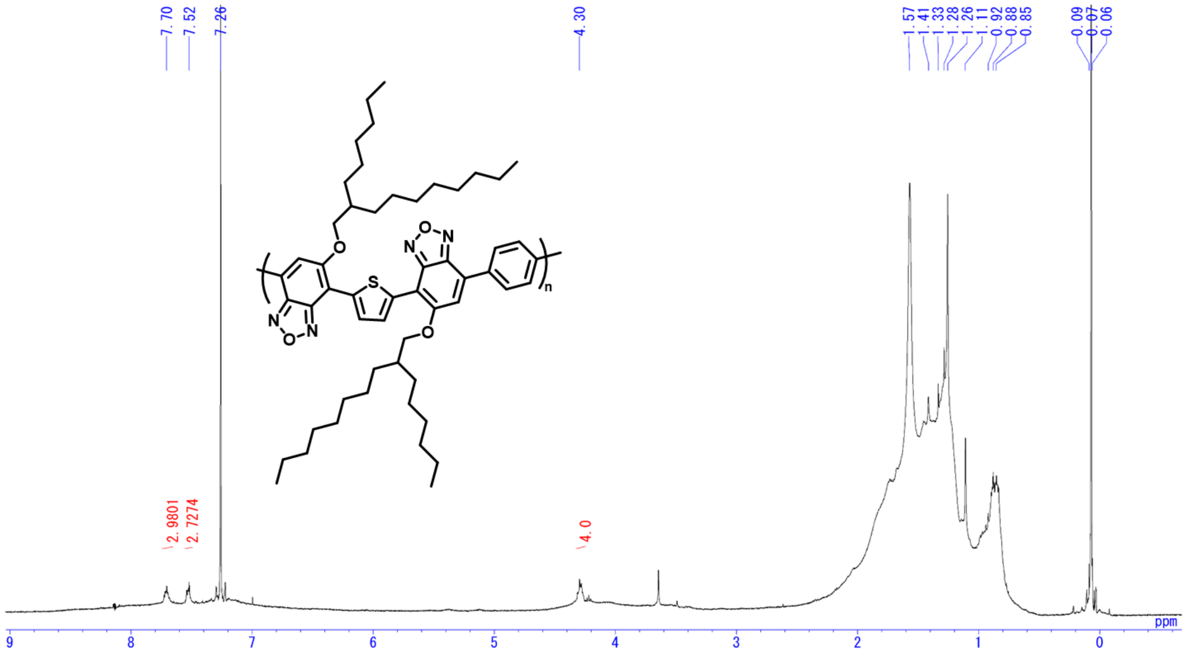


Figure S47. ^{1}H NMR spectrum of **BO1Ph** (600 MHz, $CDCl_3$).

2. Supporting Tables

Table S1. Molecular weight and M_w/M_n of BO2 and BO1Ph. [a]

Polymer	M_n [kDa]	M_w [kDa]	M_w/M_n
BO1	NA	NA	NA
BO2	32.7	62.1	1.9
BO1Ph	31.5	81.9	2.6

[a] SEC measurements in chloroform at 40 °C. BO1 was not available due to its low solubility.

Table S2. Summary of TRMC analysis.

BHJ	$\varphi\Sigma\mu$ [$\times 10^{-4}$ cm^2 V^{-1} s^{-1}]	τ_1 [μs] (intensity [%])	τ_2 [μs] (intensity [%])	τ_{ave}[a] [μs]
BO1 : Y6	8.16	0.218 (99.0)	6.63 (1.00)	0.794
BO2 : Y6	2.03	0.457 (69.3)	13.5 (30.7)	3.51
BO1Ph : Y6	1.56	0.658 (62.0)	15.1 (38.0)	4.63

[a] Average lifetime = $(A_1\tau_1 + A_2\tau_2)(A_1 + A_2)^{-1}$

Table S3. Summary of contact angle measurements.[a]

Polymer / NFA	Contact Angle [°] Water	Glycerol	γ_d [mJ m^{-2}]	γ_p [mJ m^{-2}]	γ [mJ m^{-2}]
BO1	90.9	83.6	12.0	7.61	19.6
BO2	104.4	89.6	20.2	0.79	21.0
BO1Ph	100.8	85.8	22.0	1.17	23.1
Y6[b]	106.8	94.9	14.5	1.29	15.8

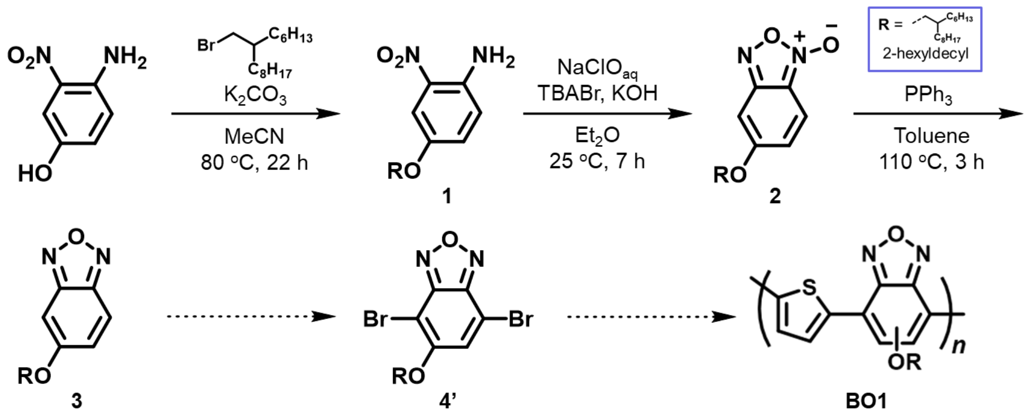
[a] γ_d: dispersion component of surface free energy, γ_p: polar component of surface free energy, γ: total surface free energy. [b] Taken from a reference. K. Kranthiraja, A. Saeki, *ACS Appl. Mater. Interfaces* **2022**, *14*, 28936.

Table S4. Summary of hole and electron mobilities of BHJ.

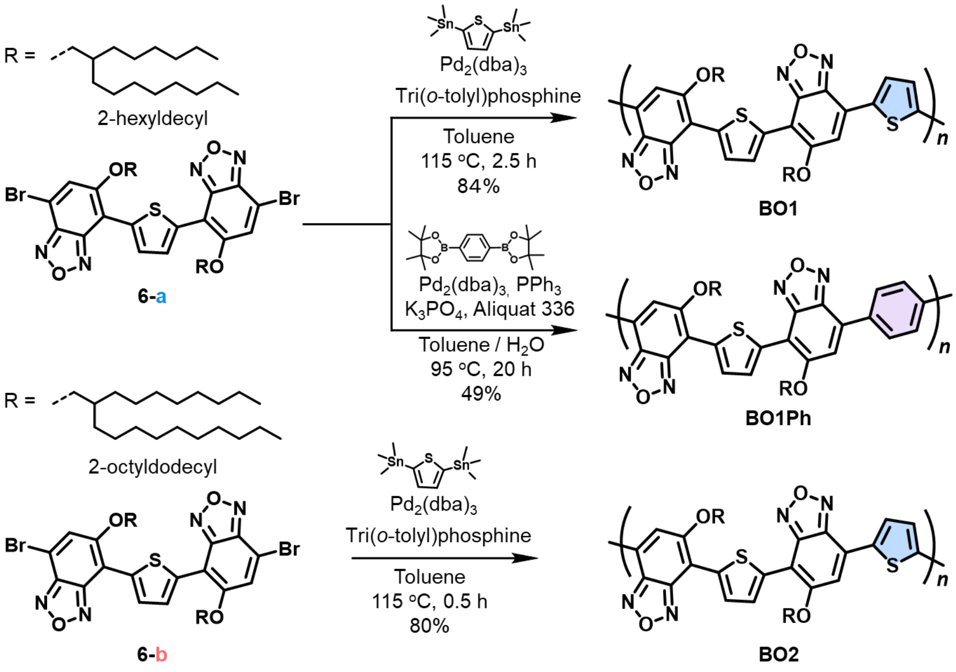
Polymer : NFA	μ_h [cm^2 V^{-1} s^{-1}]	μ_e [cm^2 V^{-1} s^{-1}]	μ_e/μ_h	L of HOD[a] [nm]	L of EOD[a] [nm]
BO1 : Y6	9.1×10^{-5}	8.6×10^{-4}	9.5	81 ± 8	102 ± 7
BO2 : Y6	2.9×10^{-5}	1.6×10^{-4}	5.5	92 ± 3	176 ± 14
BO1Ph : Y6	2.2×10^{-4}	1.4×10^{-5}	0.061	104 ± 11	58 ± 8

[a] Thickness (L) of hole-only device (HOD) and electron-only device (EOD).

3. Supporting Schemes



Scheme S1. Initial synthesis route of BO1.



Scheme S2. Synthesis route of BO1, BO2 and BO1Ph.

4. Supporting Figures

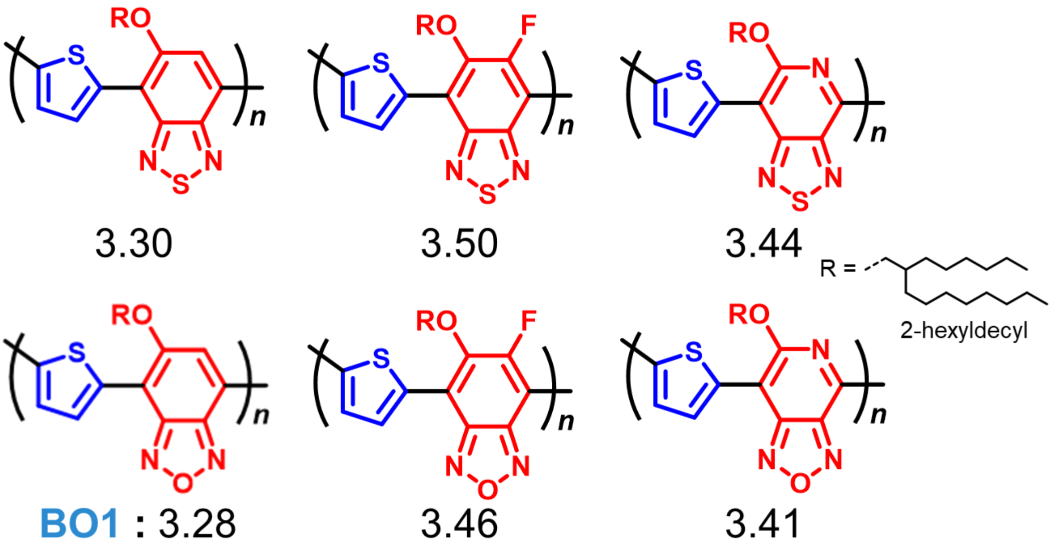


Figure S48. Chemical structure and SA score of candidate compounds.

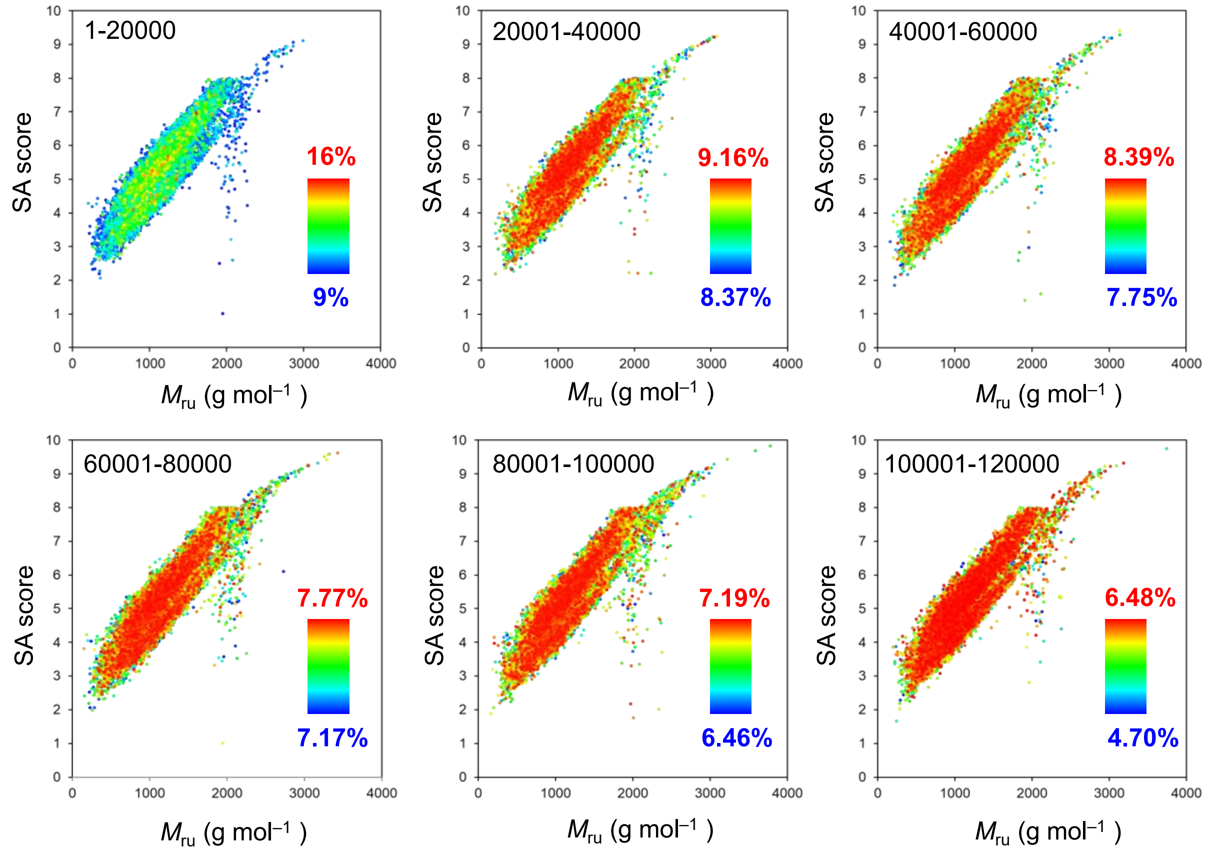
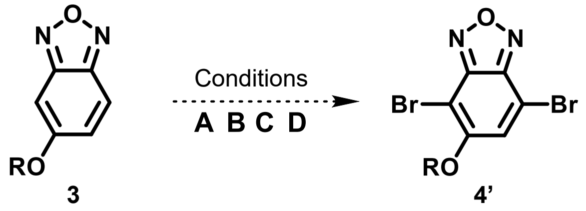


Figure S49. Synthetic accessibility (SA) score vs molecular weight of repeating unit (M_{ru}) of virtually generated polymers. The color represents ML-predicted PCE of polymers when they are combined with Y6. The panel from left top to right bottom are the polymers sorted by predicted PCE (e.g. 1-20000 is the ranking in PCE).



Entry	Conditions	Catalyst (eq)	Br_2/NBS (eq)	Temp (°C)	Time (h)	Mono-brominated	Di-brominated
1	A	-	3.0	25	72	70%	0%
2	B	0.2	3.0	100	3	88%	0%
3	C	-	3.0	125	24	72%	0%
4	D	0.05	3.5	60	23	49%	0%

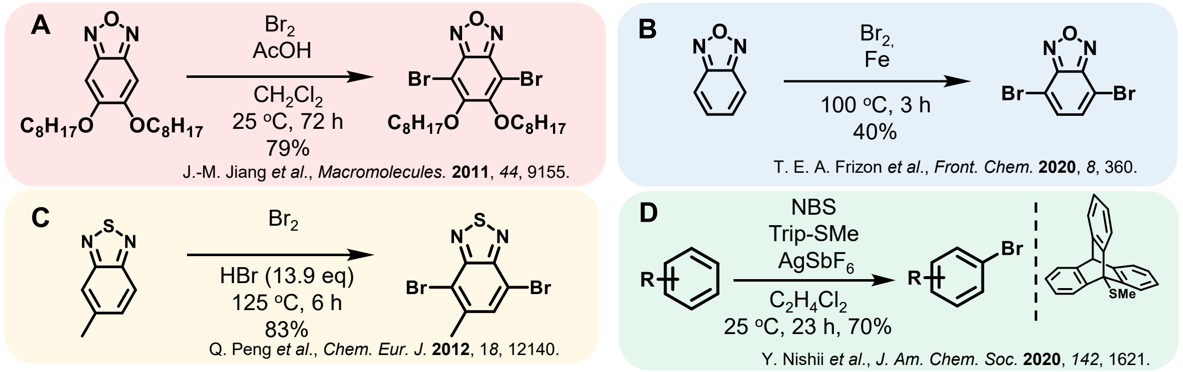


Figure S50. Bromination condition of **3**. All the reaction schema were unsuccessful (mostly mono brominated compound was obtained).

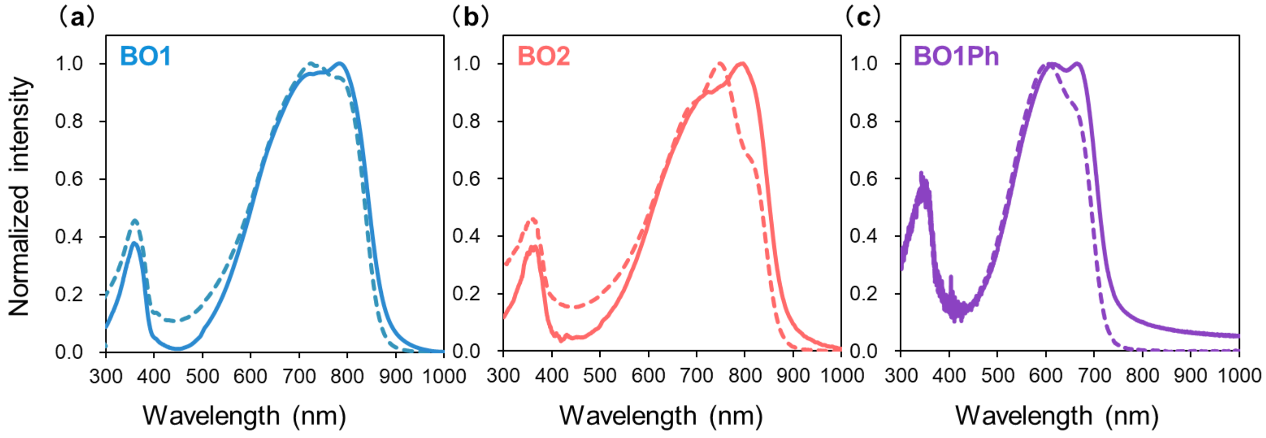


Figure S51. UV-Vis absorption spectra of polymers in the thin film (solid line) and chloroform solution (40 µg mL^{-1}, dotted line) (a) BO1, (b) BO2, and (c) BO1Ph.

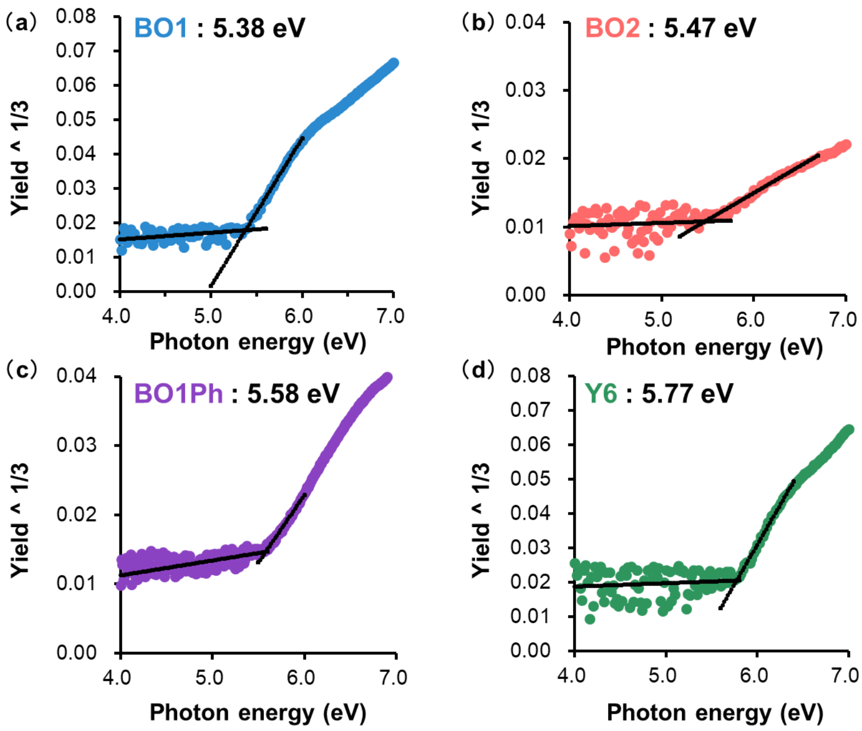


Figure S52. PYS data of polymers and Y6 in the thin film of (a) BO1, (b) BO2, (c) BO1Ph, and (d) Y6.

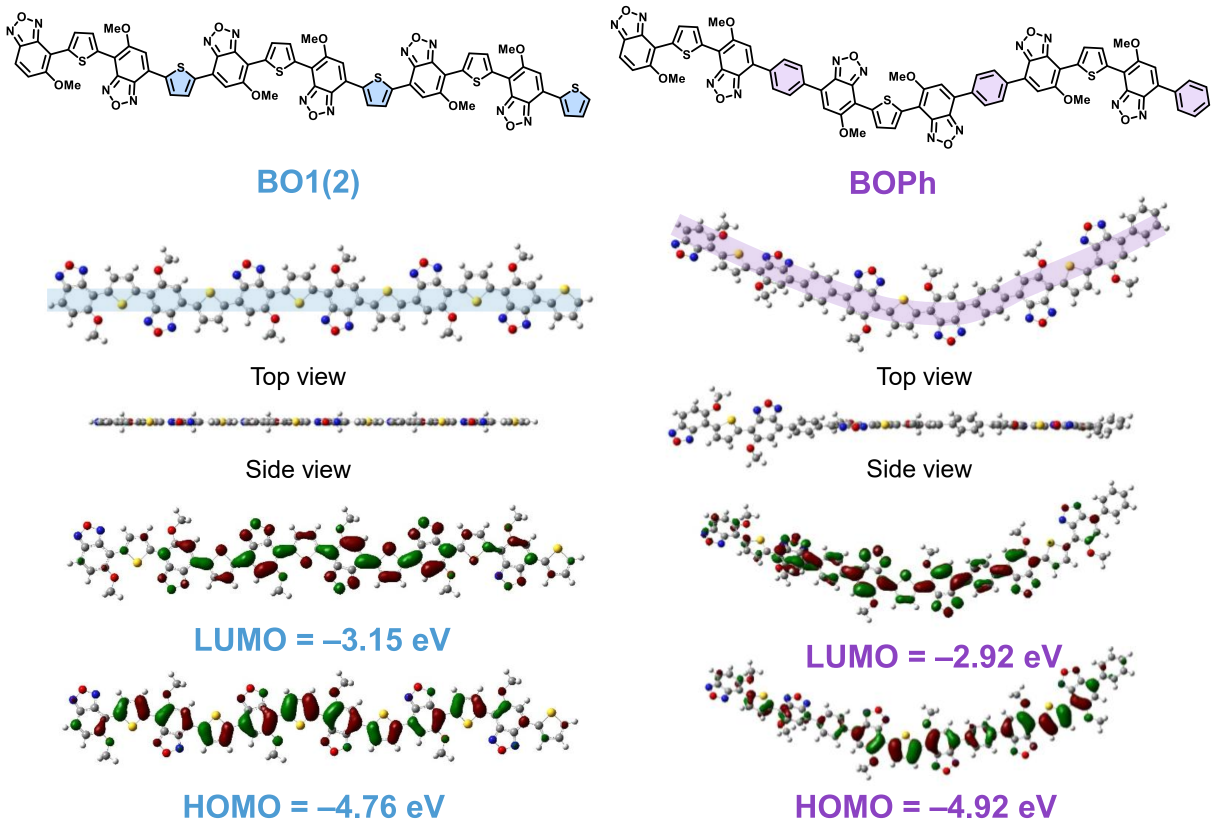


Figure S53. DFT (B3LYP/6-31G) calculation results of three repeating units of BO1(or BO2) and BOPh. The alkyl chains were replaced by methyl for simplicity.

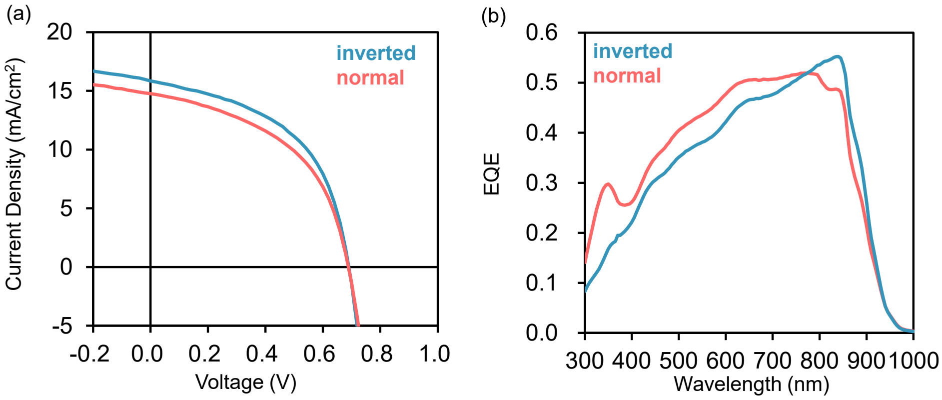


Figure S54. (a) *JV* curves and (b) EQE spectra of normal (red, ITO/PEDOT:PSS/BHJ/PDINO/Ag) and inverted (blue, ITO/ZnO/BHJ/MoOx/Ag) structures of BO2:Y6 (1:1.5).

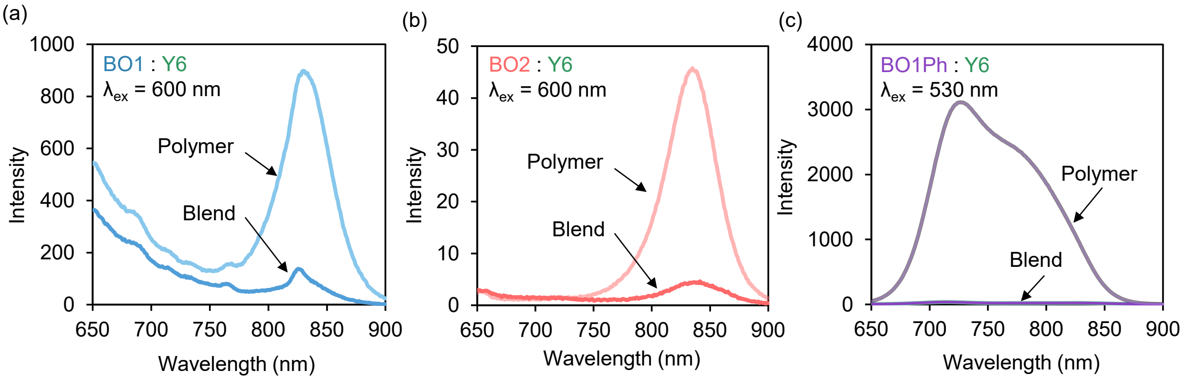


Figure S55. PL quench of each polymer (bright line) and BHJ with Y6 (dark line) for (a) BO1, (b) BO2, and (c) BO1Ph. The excitation wavelength is 600 nm for BO1 and BO2, and 530 nm for BO1Ph, at which mostly polymer is excited.

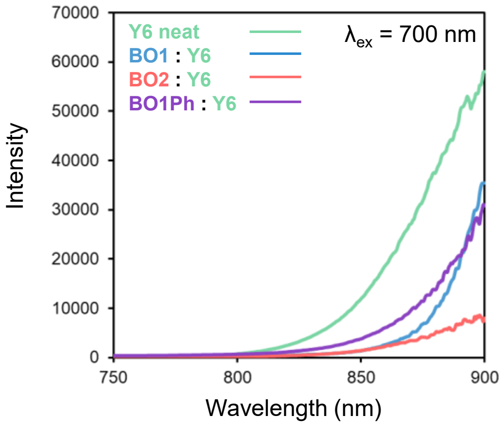


Figure S56. PL quench of BHJ (BO1:Y6, BO2:Y6, and BO1Ph:Y6) and Y6 pristine films excited at 700 nm. Note that 700 nm is not a perfect selective excitation of Y6. The sensitivity at each wavelength is calibrated.

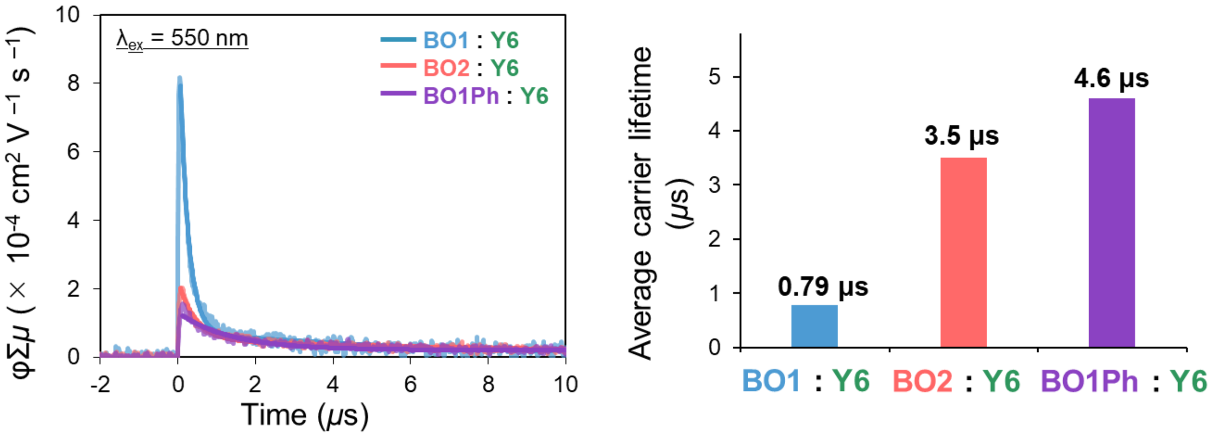


Figure S57. TRMC signal and average carrier lifetime of each BHJ (ex 550 nm, 7.1 × 10^{15} photons cm^{-2} pulse^{-1}).

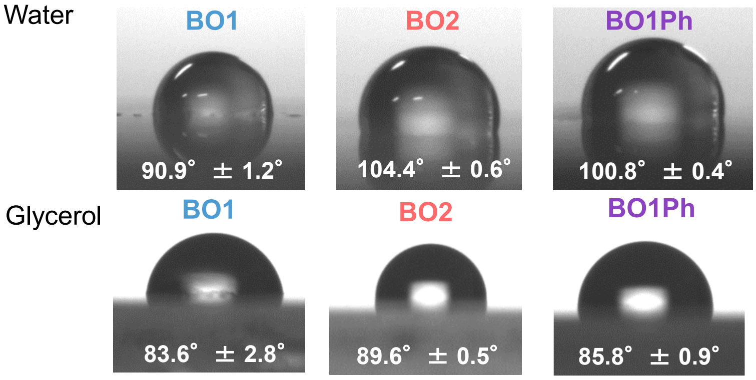


Figure S58. Contact angle (CA) measurements of pristine polymer film using water and glycerol droplets. The CAs of Y6 pristine film was 106.8° for water and 94.9° for glycerol, respectively (K. Kranthiraja, A. Saeki, *ACS Appl. Mater. Interfaces* **2022**, *14*, 28936).

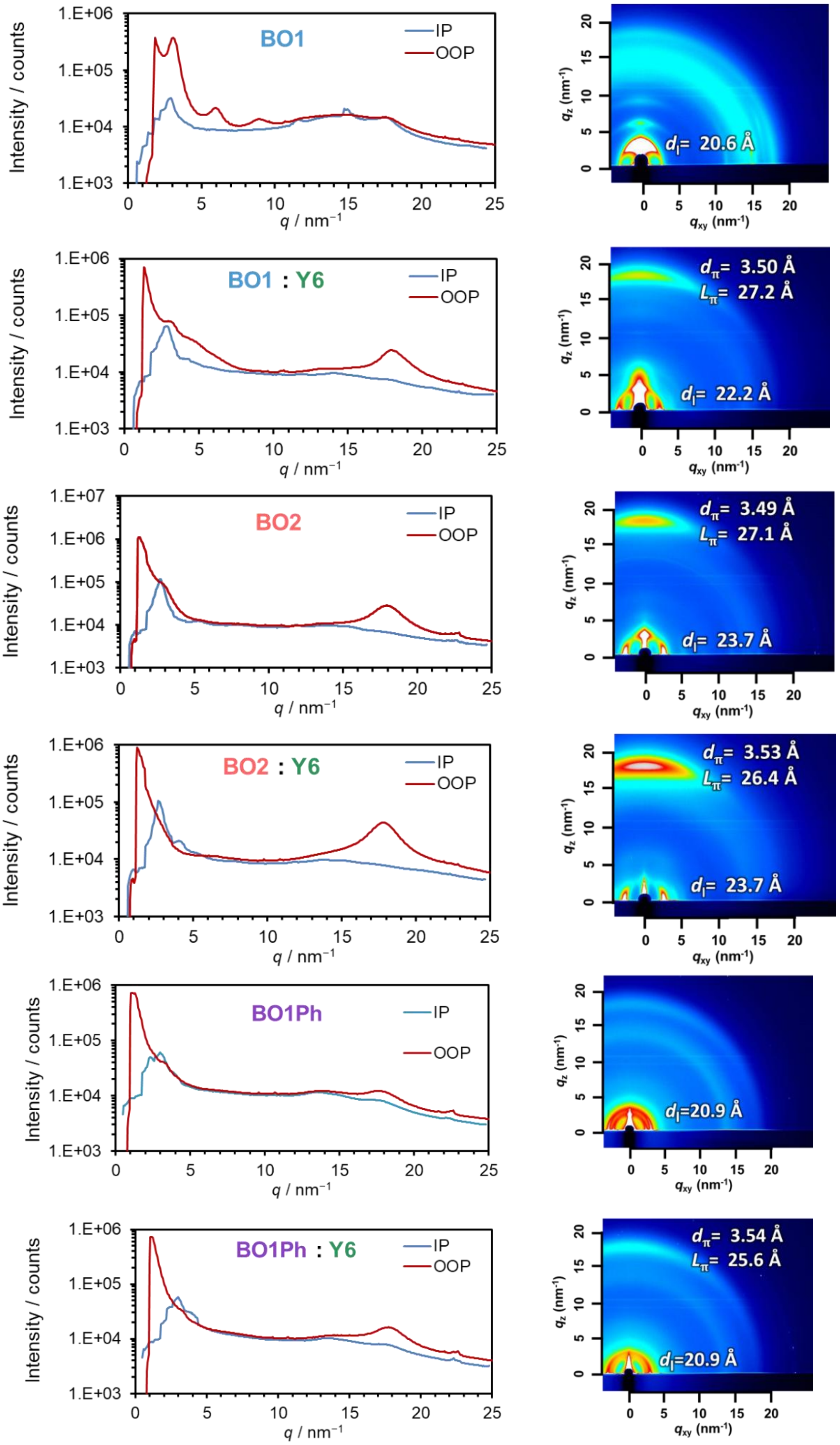


Figure S59. 1D profiles and 2D diffraction images of each polymer and BHJ. IP: in-plane, OOP: out-of-plane direction.

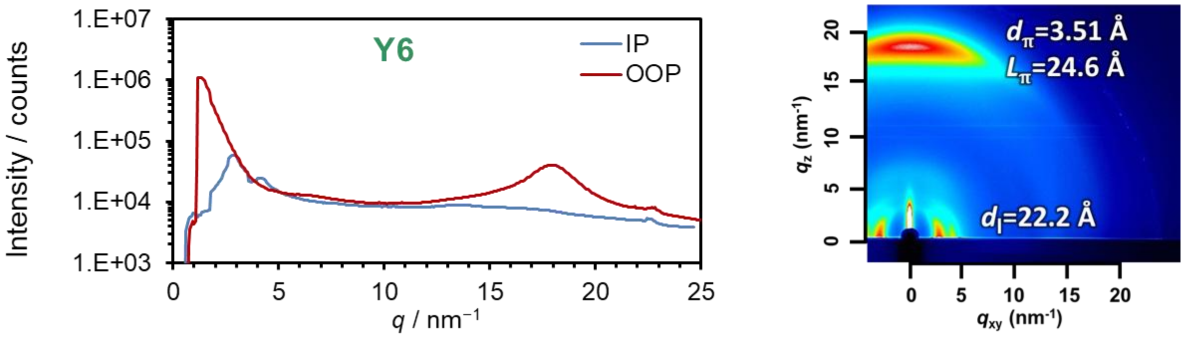


Figure S60. 1D profile and 2D diffraction image of Y6 pristine film. IP: in-plane, OOP: out-of-plane direction.

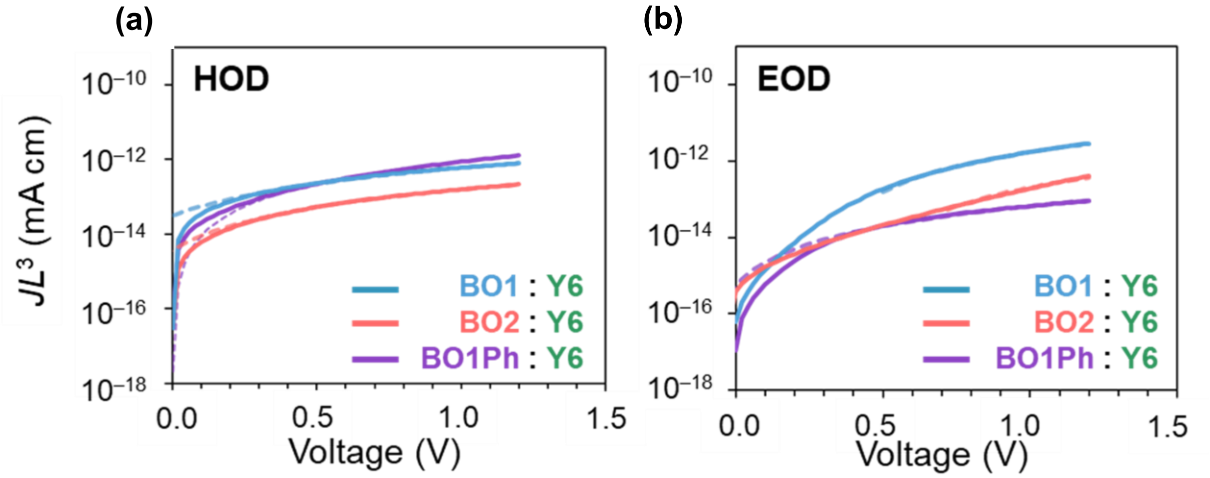
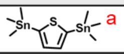
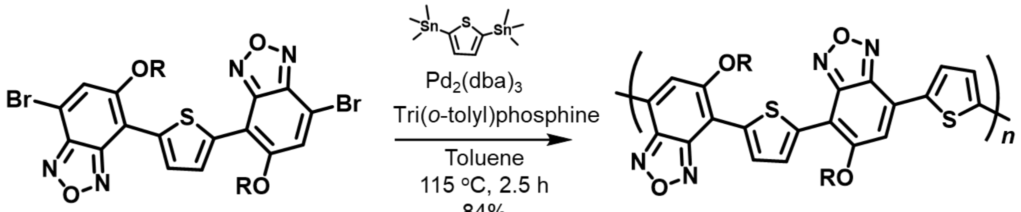


Figure S61. JL^3 vs V curves of (a) hole-only and (b) electron-only devices.

Each parameters of monomer [a] J. Xiao *et al.*, *Adv. Mater.* **2021**, *33*, 2008158

Monomer	Number of Synthetic steps	Yield (%)	Purification steps					Hazardous chemicals							
			①	②	③	④	⑤	a	b	c	d	e	f	g	h
 [a]	1	81	1	1	0	1	0	4							
M1	6	6.3	2	4	5	0	0	33							



6 **BO1**

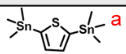
$RY_{polymer}$: $1/(0.81 \times \underline{0.84}) + 1/(0.063 \times \underline{0.84})$ $NHC_{polymer} = 4 + 33 + \underline{5} = 42$
$= 1.4697 + 18.896 = 20.4$

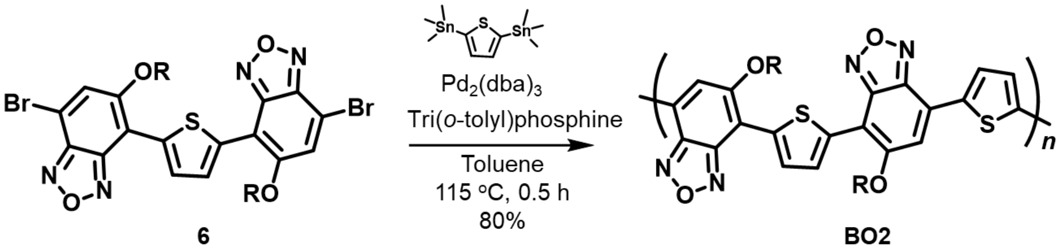
Absolute values					Relative Value[b]					SC
NSS	*RY*	*NUO*	*NCC*	*NHC*	NSS	RY(log)	NUO	NCC	NHC	
8	20.4	10	5	42	0.36	0.68	0.26	0.38	0.95	48.7

[b] Values used for the normalization: $NSS_{max} = 22$; $RY_{max} = 86.9$; $NUO_{max} = 39$; $NCC_{max} = 13$; $NHC_{max} = 44$.

Figure S62. Calculation log of SC value of BO1. The number of synthetic steps (NSS), the reciprocal yields of the monomers (RY), the number of unit operations required for the isolation/purification of the monomers (NUO), the number of column chromatographic purifications required by the monomers (NCC), and the number of hazardous chemicals used for their preparation (NHC).[S1]

Each parameters of monomer [a]) J. Xiao *et al.*, *Adv. Mater.* **2021**, *33*, 2008158

Monomer	Number of Synthetic steps	Yield (%)	Purification steps					Hazardous chemicals							
			①	②	③	④	⑤	a	b	c	d	e	f	g	h
 [a]	1	81	1	1	0	1	0	4							
M2	6	5.3	2	4	5	0	0	33							



6

BO2

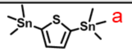
RY_{polymer}: $1/(0.81 \times \underline{0.80}) + 1/(0.053 \times \underline{0.80})$ $NHC_{\text{polymer}} = 4 + 33 + \underline{5} = 42$
$= 1.5432 + 23.585 = 25.1$

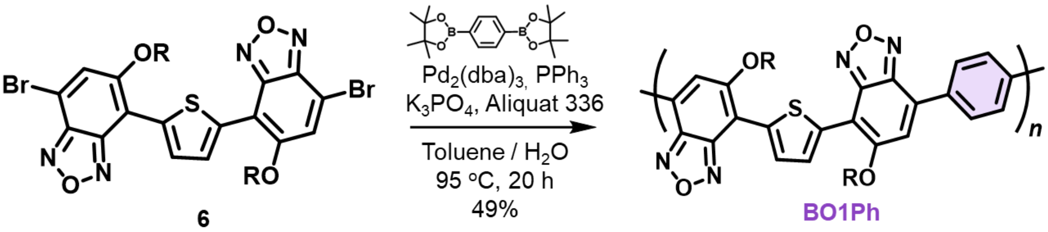
Absolute values					Relative Value[b]					SC
NSS	*RY*	*NUO*	*NCC*	*NHC*	NSS	RY(log)	NUO	NCC	NHC	
8	25.1	10	5	42	0.36	0.72	0.26	0.38	0.95	49.7

[a])Values used for the normalization: NSS_{max}= 22; RY_{max}= 86.9; NUO_{max}= 39; NCC_{max}= 13; NHC_{max}= 44.

Figure S63. Calculation log of SC value of BO2. The number of synthetic steps (NSS), the reciprocal yields of the monomers (RY), the number of unit operations required for the isolation/purification of the monomers (NUO), the number of column chromatographic purifications required by the monomers (NCC), and the number of hazardous chemicals used for their preparation (NHC).[S1]

Each parameters of monomer [a] J. Xiao *et al.*, *Adv. Mater.* **2021**, *33*, 2008158

Monomer	Number of Synthetic steps	Yield (%)	Purification steps					Hazardous chemicals							
			①	②	③	④	⑤	a	b	c	d	e	f	g	h
 [a]	1	81	1	1	0	1	0	4							
M1	6	6.3	2	4	5	0	0	33							



RY_{polymer}: $1/(0.81 \times \underline{0.49}) + 1/(0.063 \times \underline{0.49})$ $NHC_{\mathrm{polymer}} = 4 + 33 + \underline{9} = 46$
$= 2.5195 + 32.394 = 34.9$

Absolute values					Relative Value[b]					SC
NSS	*RY*	*NUO*	*NCC*	*NHC*	NSS	RY(log)	NUO	NCC	NHC	
8	34.9	10	5	46	0.36	0.80	0.26	0.38	1.05	52.7

[b] Values used for the normalization: $NSS_{\mathrm{max}} = 22$; $RY_{\mathrm{max}} = 86.9$; $NUO_{\mathrm{max}} = 39$; $NCC_{\mathrm{max}} = 13$; $NHC_{\mathrm{max}} = 44$.

Figure S64. Calculation log of SC value of BO1Ph. The number of synthetic steps (NSS), the reciprocal yields of the monomers (RY), the number of unit operations required for the isolation/purification of the monomers (NUO), the number of column chromatographic purifications required by the monomers (NCC), and the number of hazardous chemicals used for their preparation (NHC).[S1]

5. Supporting References

[S1] R. Po, G. Bianchi, C. Carbonera and A. Pellegrino, *Macromolecules*, 2015, **48**, 453–461.