Electronic Supplementary Information for:

Engineering of an Archaeal Phosphodiesterase to Trigger Aggregation-Induced Emission (AIE) of Synthetic Substrates

Yunlong Zhang, Zhiyuan Wu, Ippei Takashima, Nobuyuki Matsumoto, Kathy-Uyen Nguyen, and Jonathan S. Lindsey

Table of Contents **Page**

1. Further information concerning enzyme engineering procedures S1

2. Additional data concerning enzyme engineering

Figure S1, SDS-PAGE of PDE-MJ0936 monomer S3

Figure S2, MALDI-TOF-MS of PDE-MJ0936 monomer S3

Figure S3, Effects of pH to the enzymatic activity of PDE-MJ0936 S4

Figure S4, Mutation residues of PDE-MJ0936 S4

Figure S5, BLASP alignment of PDE-MJ0936 and YfcE S5

Table S1, List of the PDE mutants S5

Figure S6, Selection of PDE mutants with the substrates **1**, **2** and **3** S6

Figure S7, Initial rate of C141D at pH 7.0 and 20 °C S7

Figure S8, FPLC trace of two batches of C141D purification S8

Figure S9, Absorption spectra of **1–3** in ethanol S8

3. ^{1}H and ^{13}C NMR data for new synthetic compounds S9

1. Further information concerning enzyme engineering procedures

PCR for site-directed mutagenesis. With pure pLIC.B2.MJ0936, site-directed mutagenesis was carried out by using Q5 site-directed mutagenesis kit with the following PCR method. In a 25-μL PCR tube, 1 μL of DNA template, 9 μL of H_2O, 1.25 μL of forward primer (see Table 2), 1.25 μL of reverse primer (see Table 2) and 12.5 μL of Q5 Hot-Start Master Mix were mixed for PCR. Initial denaturation was at 98 °C for 30 s. Annealing and elongation (40 cycles) were as follows: 98 °C for 10 s, T_a (see Table 2) for 20 s and 72 °C for 3 min. Final extension was as follows: 72 °C for 3 min. Then the sample was cooled down to 4 °C for electrophoresis.

KLD reaction and transformation. A sample of 100 μL of frozen *E. coli* stock (Top10 for plasmid storage or BL21 for protein expression) was slowly thawed on ice. The PCR product

(1 μL) was mixed with 2× KLD buffer (5 μL), 10× KLD enzyme mix (1 μL) and ddH_2O (3 μL) and reacted at room temperature for 5 min. The reaction mixture (5 μL) was added to the freshly thawed *E. coli* stock and mixed by gentle finger agitation (flicking) the tube. Then the cell stock was incubated on ice for 30 min, followed by a quick heat-shock at 42 °C in a water bath for exactly 45 s. The cells were immediately cooled down on ice for 5 min after heat-shock. A sample of SOC outgrowth media (900 μL) was added to the cooled cells, and the mixture was incubated at 37 °C for 1 h. Afterwards, 100 μL of the cell culture was spread on an LB/Amp agar plate. The grown colonies contain the PCR-amplified plasmid. For protein expression, after colony selection (*vide infra*), the pure strain was frozen for a second round of transformation with the helper plasmid pSJS1240, which is used to increase the protein yield of archaeal proteins. To spread and grow *E. coli* containing pSJS1240, LB/Amp/Sp plate and media were used.

Protein fractionation and purification. TALON cobalt affinity resin was loaded into a Bio-Rad Econo-Pac® plastic column to prepare a 10-mL affinity column. The flow rate was 3 mL/min for all steps except 1 mL/min for elution. The column was washed and equilibrated with 50 mL of running buffer (20 mM Na_3PO_4, 0.5 M NaCl, 5 mM imidazole, pH 7.4). Then the cell lysate (~ 9 mL) was loaded and flowed through. 100 mL of running buffer was loaded to wash the column. Afterwards, 50 mL of elution buffer (20 mM Na_3PO_4, 0.5 M NaCl, 500 mM imidazole, pH 7.4) was used to elute the proteins bound to the column. The collection volume was 2 mL. 5 μL of each eluent was added to 195 μL of 2 mM bis-*p*NPP (in PDE buffer) for an incubation at 60 °C for 30 min to select the active PDE fractions.

The purity of the previous PDE fraction (crude PDE) was roughly ~60%, which was used for a qualitative assay. Purer protein was obtained by the following method for quantitative assays. Buffer A (20 mM Tris, 100 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, pH 7.5) and Buffer B (20 mM Tris, 700 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, pH 7.5) were freshly prepared for FPLC system. Crude PDE was buffer-exchanged with H_2O in dialysis tubing (molecular weight cut-off 3000 Da) and lyophilized. Then the lyophilized PDE was dissolved in 2 mL of cold buffer A and loaded to a Bio-Rad NGC medium-pressure chromatography system (FPLC). Two HiTrap Q columns (2 × 5 mL) were connected to the FPLC system and equilibrated with 100% buffer A. The flow rate was consistently 5 mL/min. Protein elution was monitored by absorption spectroscopy at 280 nm. FPLC procedures: (1) Equilibration, 2 column volumes (CV), 0% buffer B; (2) Sample application, 5 mL/min; (3) Elution, 10 CV, 0–100% buffer B; (4) Elution, 1 CV, 100% buffer B; (5) Column wash: 1 CV, 100–0% buffer B. The collection volume of eluents was 3 mL. 5 μL of each eluent was added to 195 μL of 2 mM bis-*p*NPP (in PDE buffer) for an incubation at 60 °C for 30 min to select the active PDE fractions. An FPLC trace is shown in Figure S1. PDE fractions were buffer-exchanged with H_2O by dialysis and then lyophilized. The purity of the PDE (or mutants) was >95% (by SDS-PAGE).

2. Additional data concerning enzyme engineering

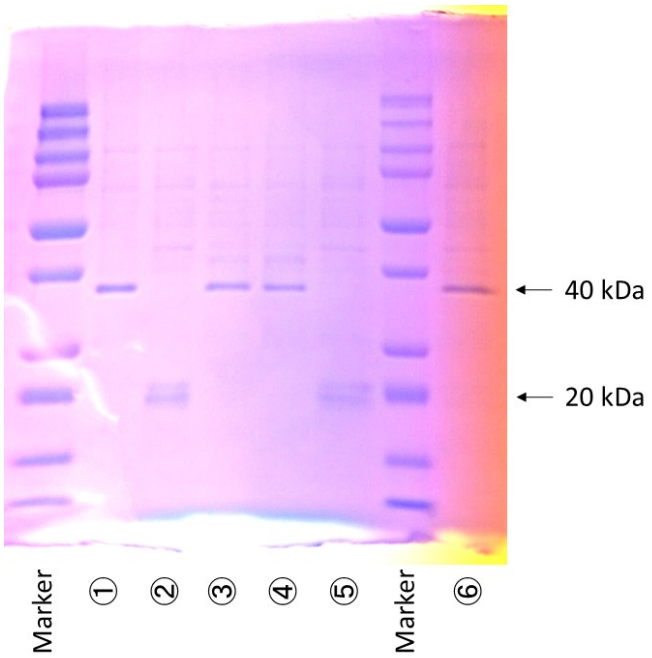


Figure S1. SDS-PAGE of PDE-MJ0936 using a 12% acrylamide gel under denaturing conditions: (1) Non-treated; (2) reductants (1 M β-mercaptoethanol + 30 mM dithiothreitol) and heating (99 °C × 3 min); (3) 6% Triton X-100 and heating; (4) 8 M urea and heating; (5) all additives in runs (2–4) + heating; and (6) all additives in runs (2–4) without heating. Loading: PDE-MJ0936 is 0.24 pmol.

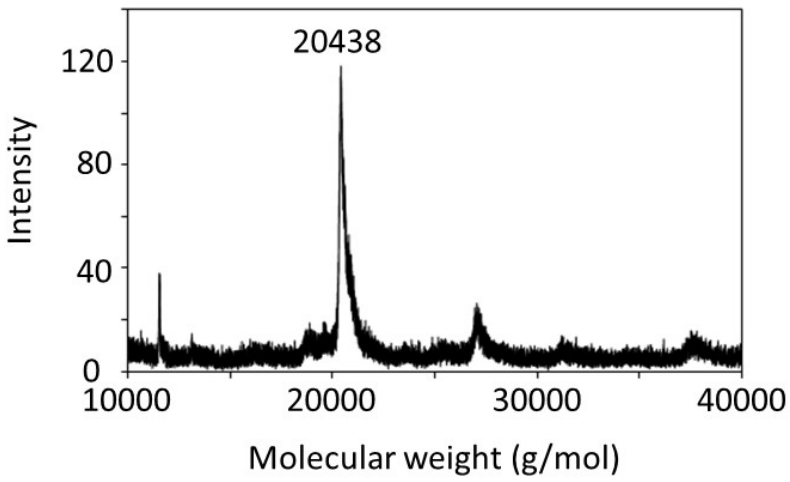


Figure S2. MALDI-TOF-MS spectrum of PDE-MJ0936 with the matrix sinapinic acid.

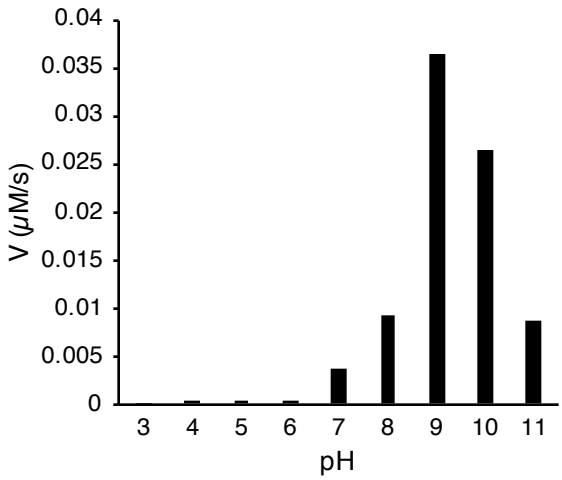


Figure S3. Effect of pH on the enzymatic activity of PDE-MJ0936 (all at 60 °C). The substrate was bis-*p*NPP.

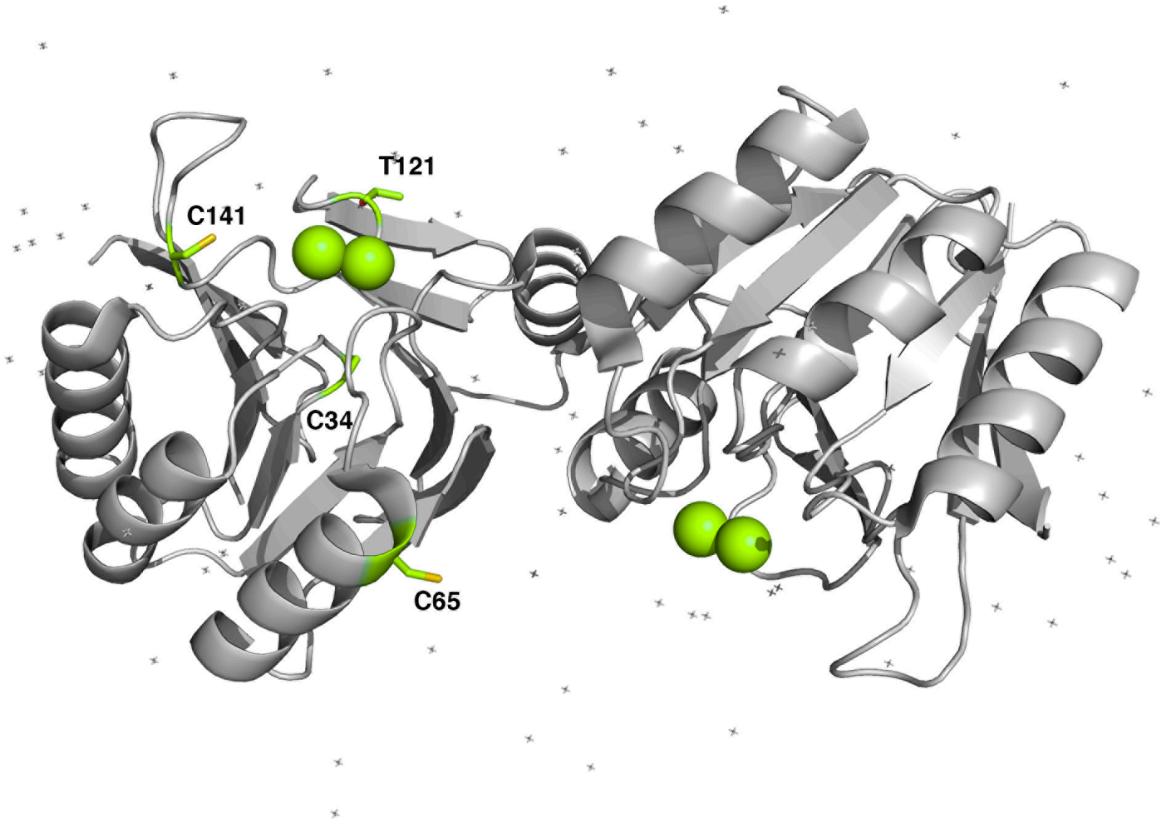


Figure S4. Crystal structure of PDE-MJ0936 with Ni^{2+} (PDB #: 1S3M). Mutation sites in a monomer are labeled. Each H_2O molecule is indicated by a cross.

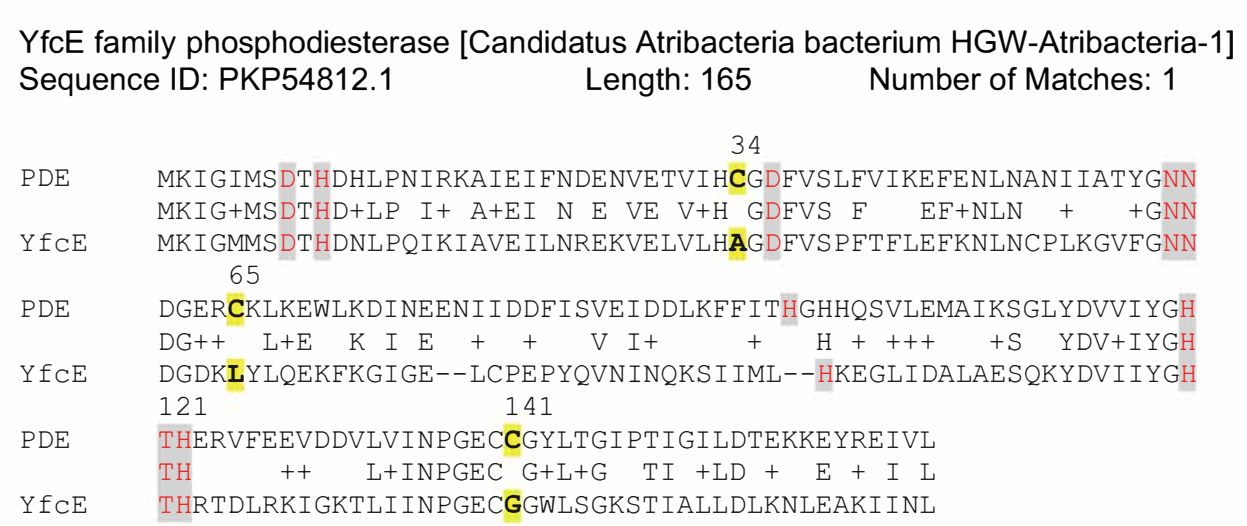


Figure S5. Protein alignment (BLASTP) of PDE-MJ0936 and YfcE (GenBank # PKP54812), a bacterial phosphodiesterase. All residues in the active site (red) are conserved (grey highlighted). All cysteines in PDE-MJ0936 and their aligned residues in YfcE are highlighted in yellow.

Table S1. Nine PDE mutants to obtain monomeric PDE and/or PDE with broader substrate scope.

Name	Amino acid	Feature	Substitution	Reason
C34A	Cys	Not surface	Ala	YfcE residue
C34D			Asp	negatively charged
C65D	Cys	Surface	Asp	negatively charged
C65L			Leu	YfcE residue
C141D	Cys	Surface	Asp	negatively charged
C141G			Gly	YfcE residue

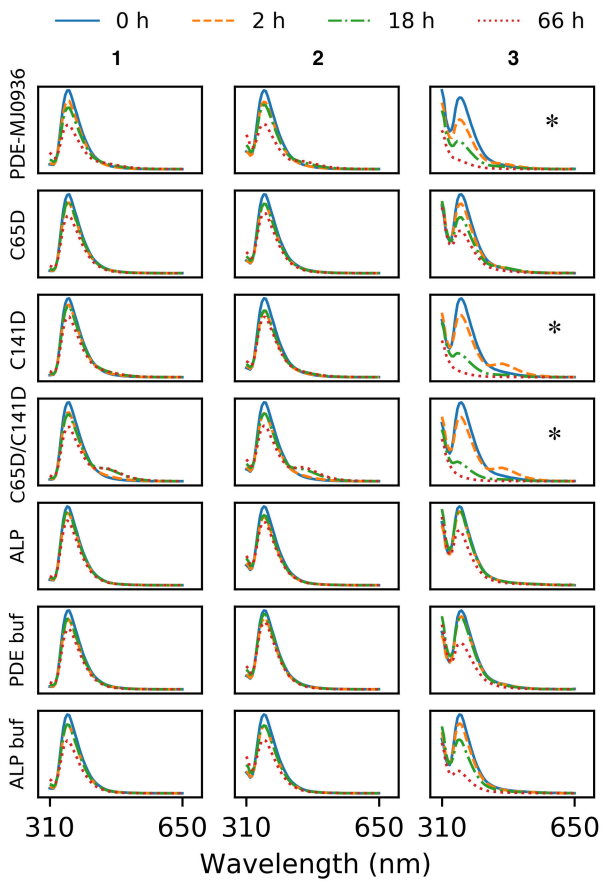


Figure S6. Fluorescence signal from PDE mutants with the substrates **1**–**3** based on AIE. The * indicates the observation of cloudiness (precipitates) during the assay.

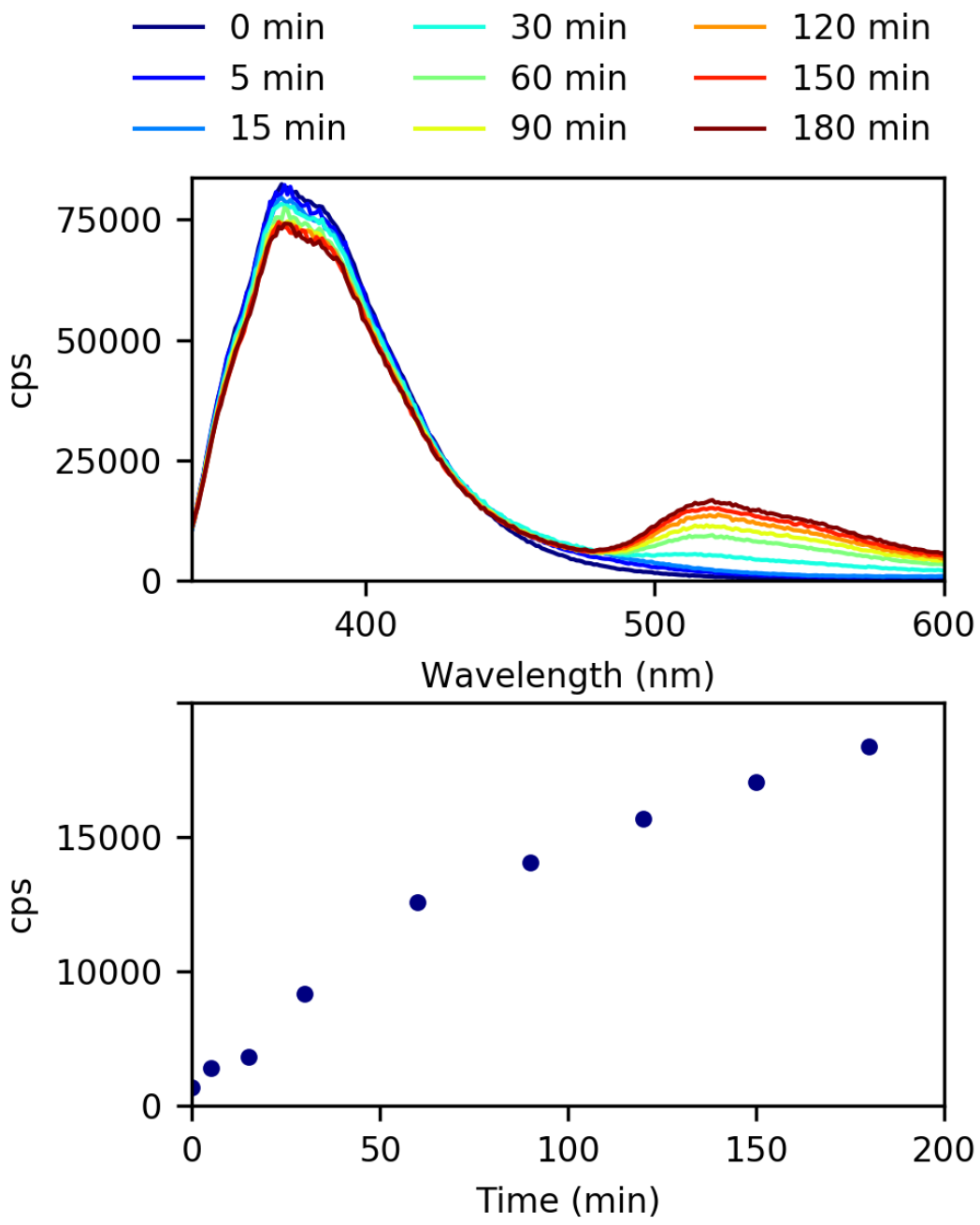


Figure S7. Initial rate of hydrolysis of **3** by C141D at pH 7.0 and room temperature (~20 °C). No incubation of the plate. Note that the fluorescence at 370 nm comes from the phosphodiester **3**, not the resulting liberated **12**. The fluorescence from **3** has high intensity due to the higher amount compared to that of the newly formed **12**.

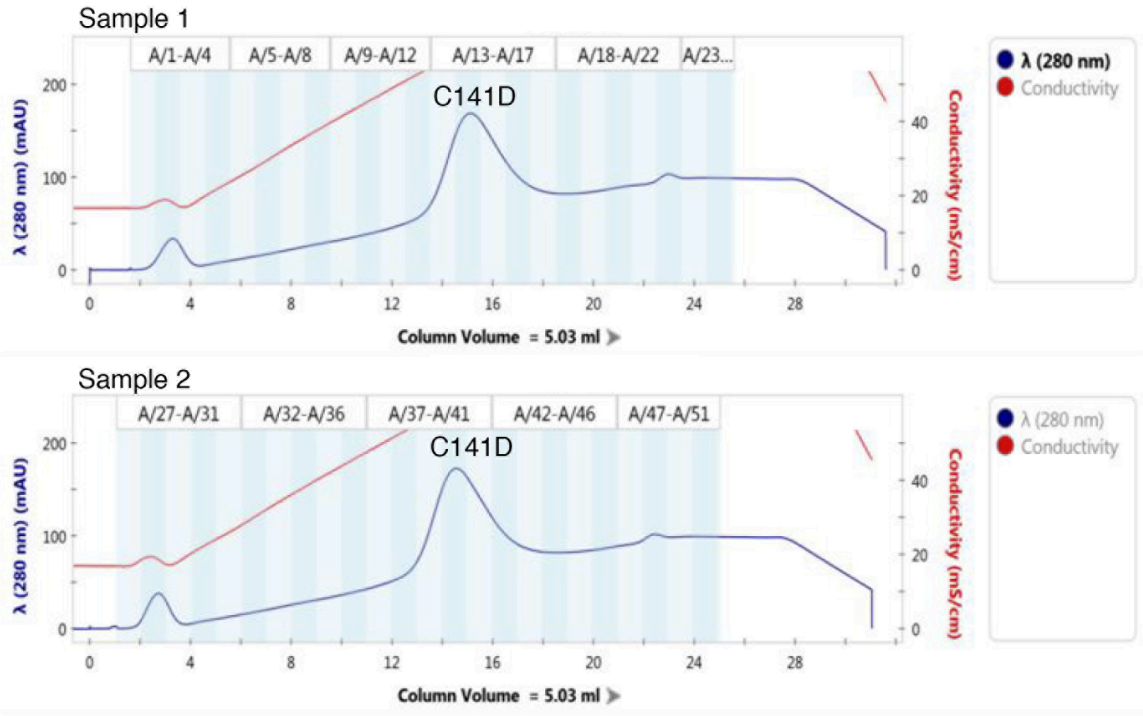


Figure S8. FPLC trace of two batches of C141D purification.

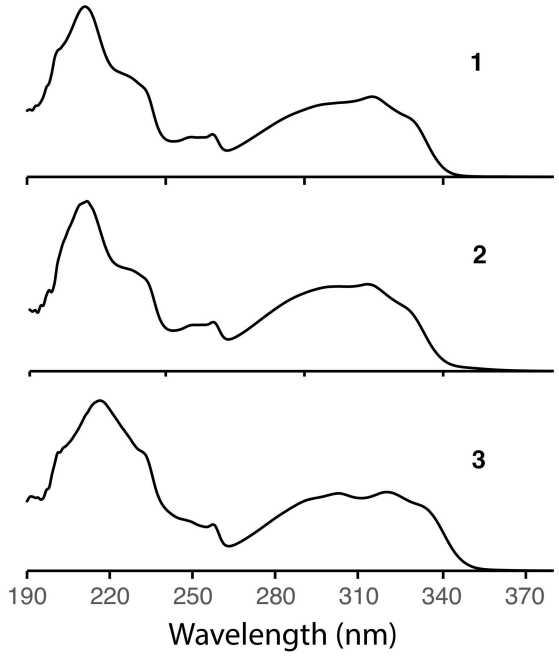
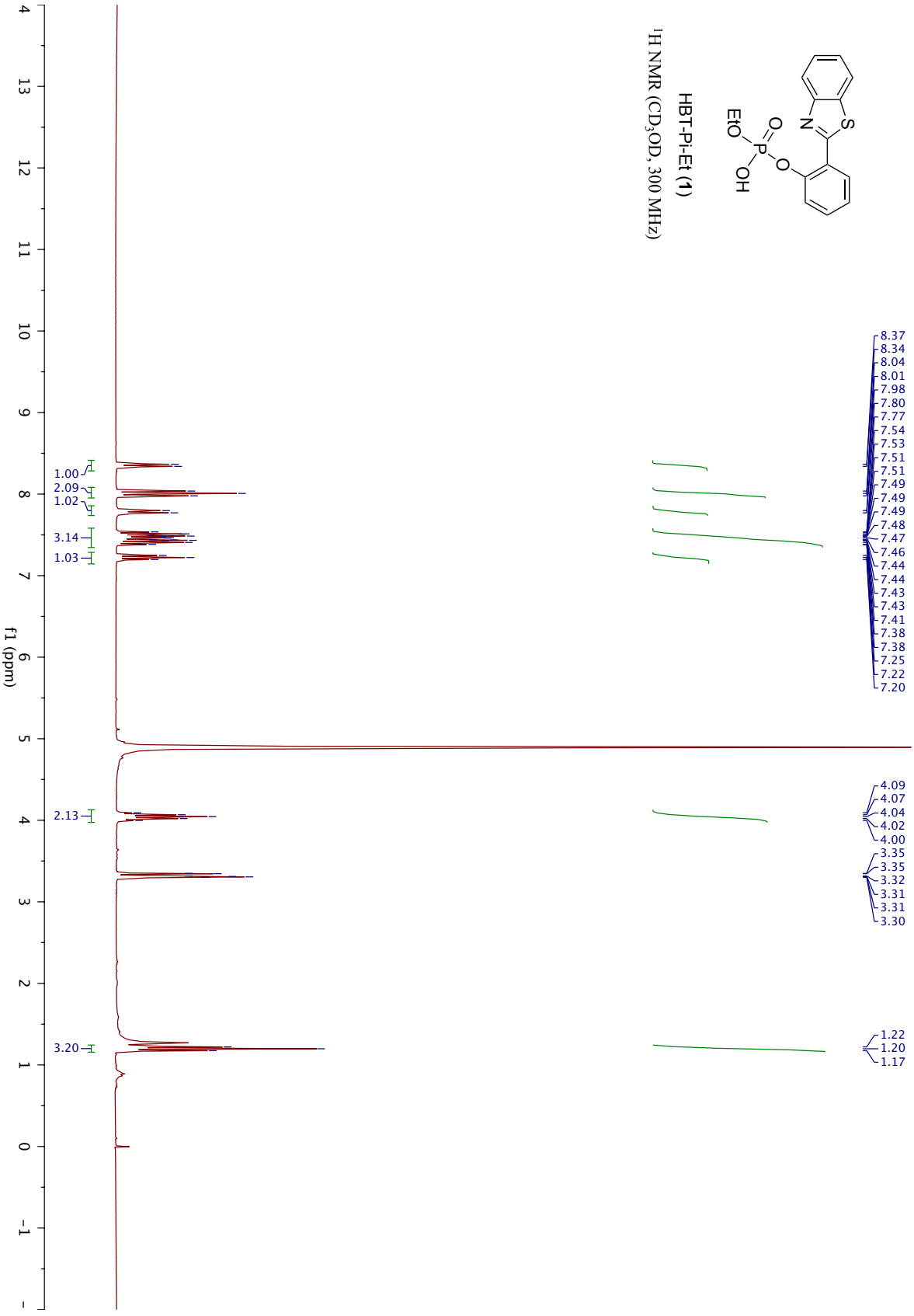
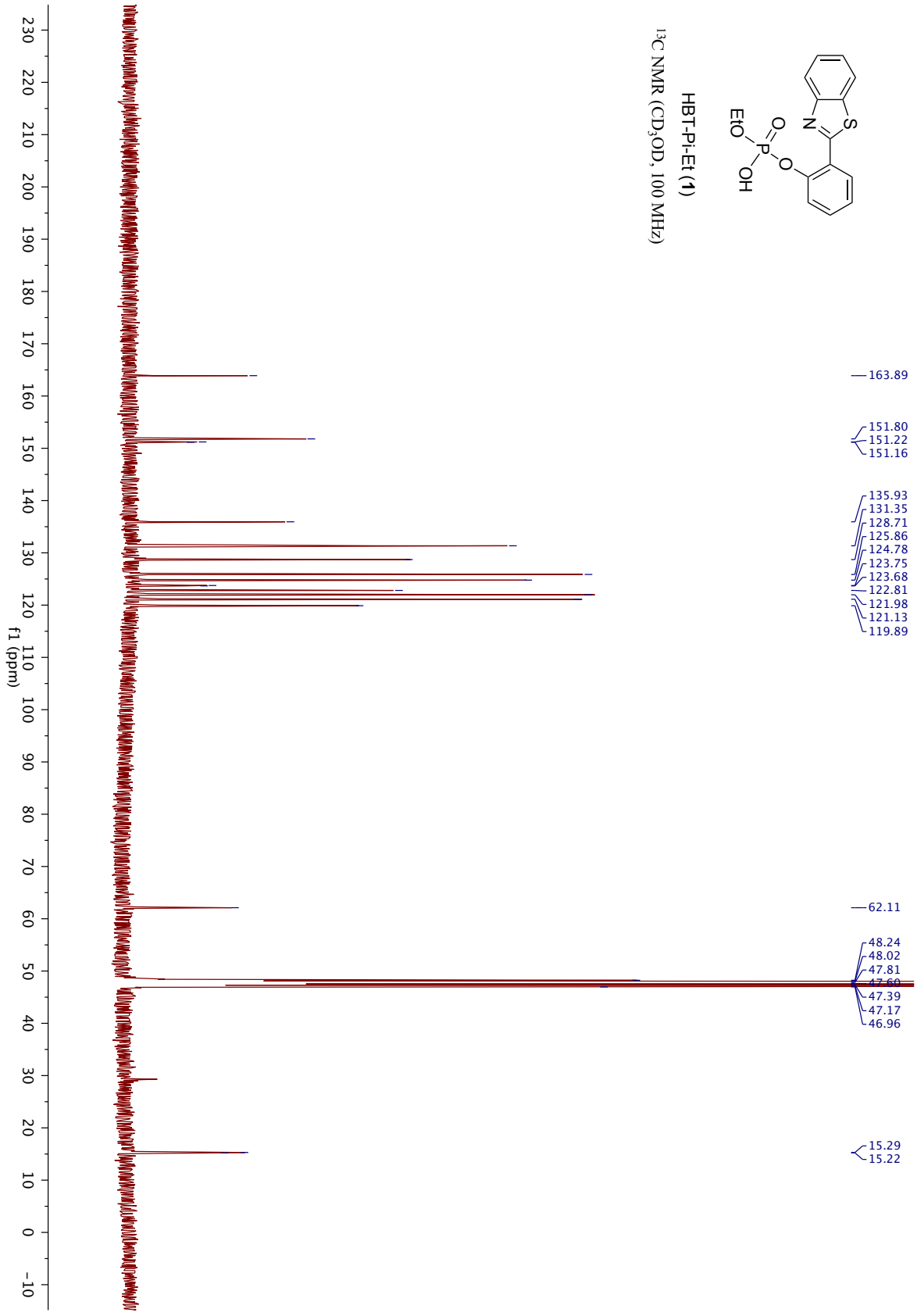
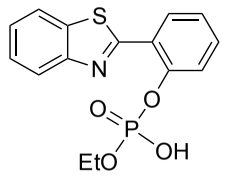


Figure S9. Absorption spectra of **1–3** in ethanol at room temperature (1-cm pathlength cuvette).

3. ^{1}H and ^{13}C NMR data for new synthetic compounds

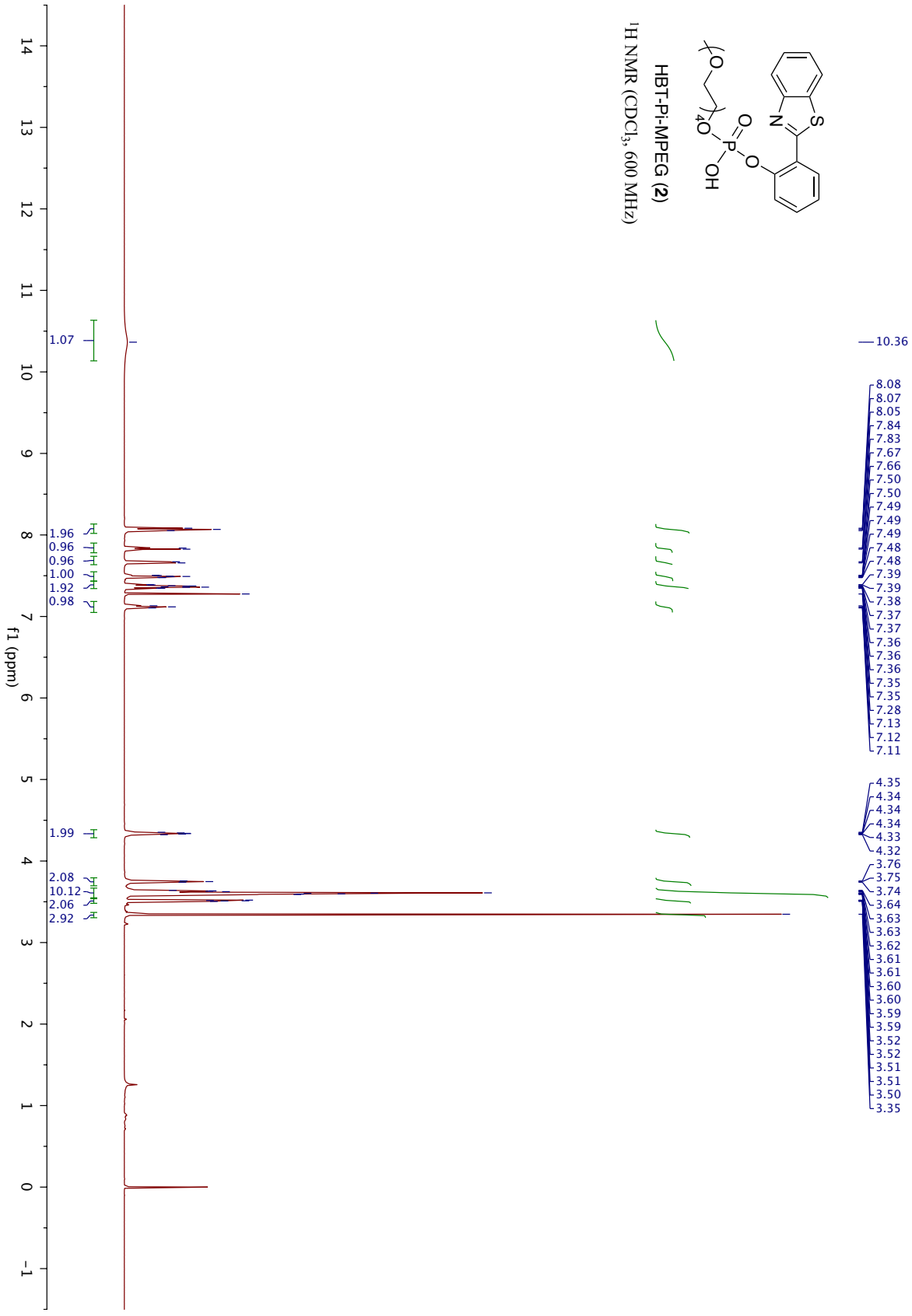

HBT-Pi-Et (1)
^{1}H NMR (CD_3OD, 300 MHz)
EtO
OH
O
P
O
N
S
f1 (ppm)
1.00
2.09
1.02
3.14
1.03
2.13
3.20
8.37
8.34
8.04
8.01
7.98
7.80
7.77
7.54
7.53
7.51
7.51
7.49
7.49
7.49
7.48
7.47
7.46
7.44
7.44
7.43
7.43
7.41
7.38
7.38
7.25
7.22
7.20
4.09
4.07
4.04
4.02
4.00
3.35
3.35
3.32
3.31
3.31
3.30
1.22
1.20
1.17

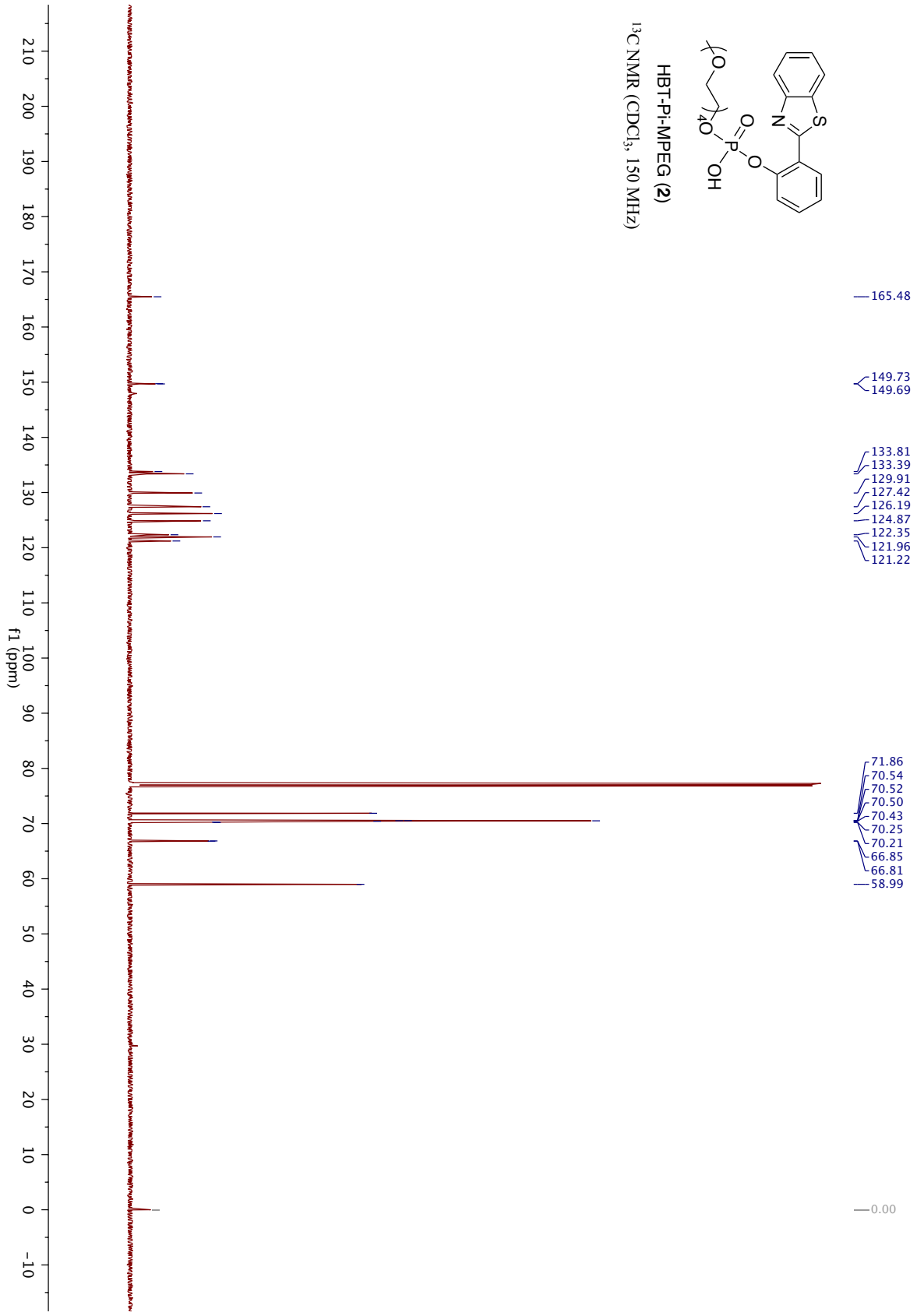

HBT-Pi-Et (1)
^{13}C NMR (CD_3OD, 100 MHz)
EtO
O
P
O
OH
N
S
163.89
151.80
151.22
151.16
135.93
131.35
128.71
125.86
124.78
123.75
123.68
122.81
121.98
121.13
119.89
62.11
48.24
48.02
47.81
47.60
47.39
47.17
46.96
15.29
15.22
f1 (ppm)

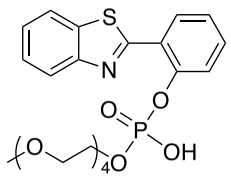


HBT-Pi-Et (**1**)

^{31}P NMR (CD_3OD, 120 MHz)

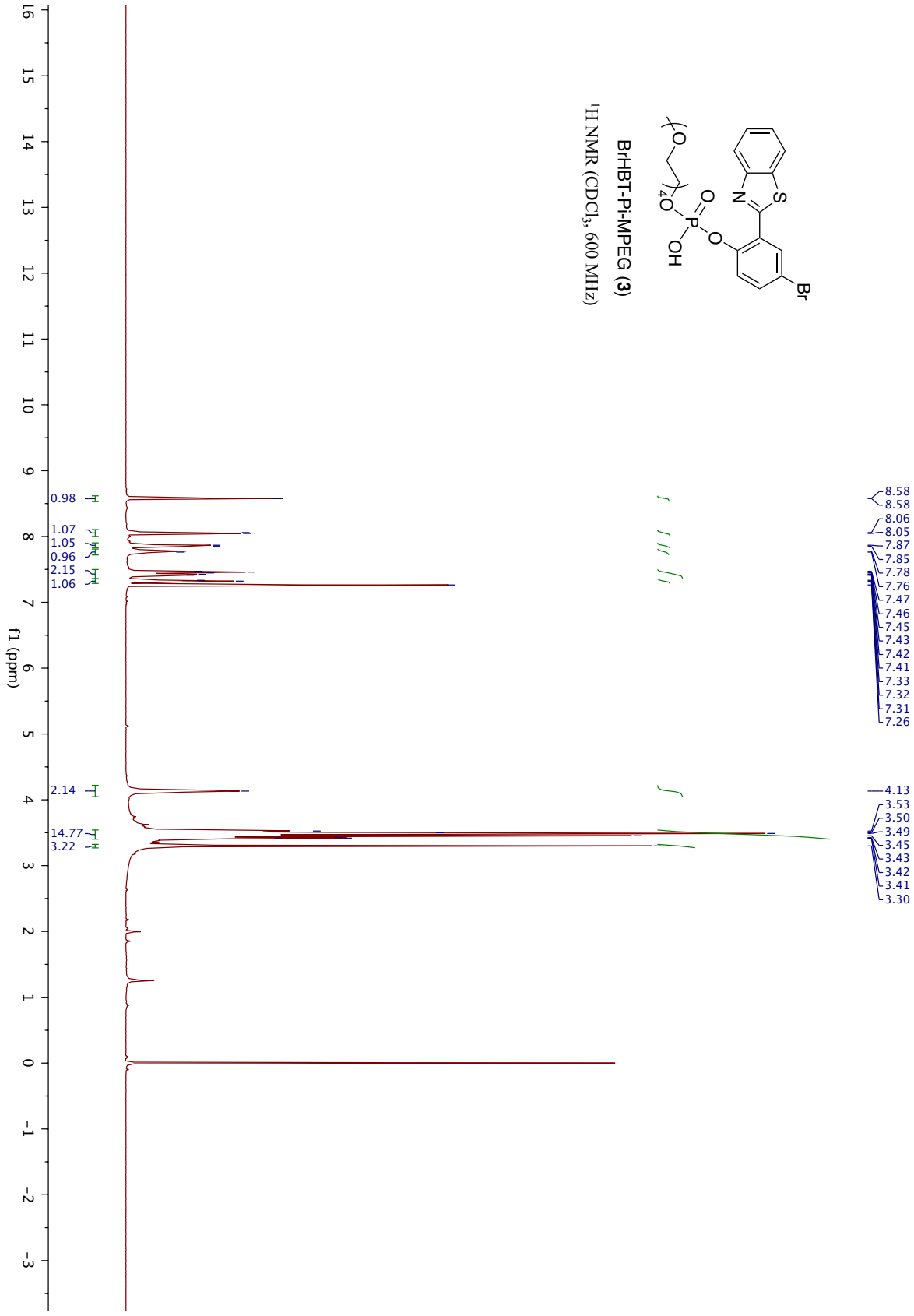

HBT-Pi-MPEG (2)
^{1}H NMR (CDCl$_3$, 600 MHz)
f1 (ppm)
1.07
1.96
0.96
0.96
1.00
1.92
0.98
1.99
2.08
10.12
2.06
2.92
10.36
8.08
8.07
8.05
7.84
7.83
7.67
7.66
7.50
7.50
7.49
7.49
7.49
7.48
7.48
7.39
7.39
7.38
7.37
7.37
7.36
7.36
7.36
7.35
7.35
7.28
7.13
7.12
7.11
4.35
4.34
4.34
4.34
4.33
4.32
3.76
3.75
3.74
3.64
3.63
3.63
3.62
3.61
3.61
3.60
3.60
3.59
3.59
3.52
3.52
3.51
3.51
3.50
3.35

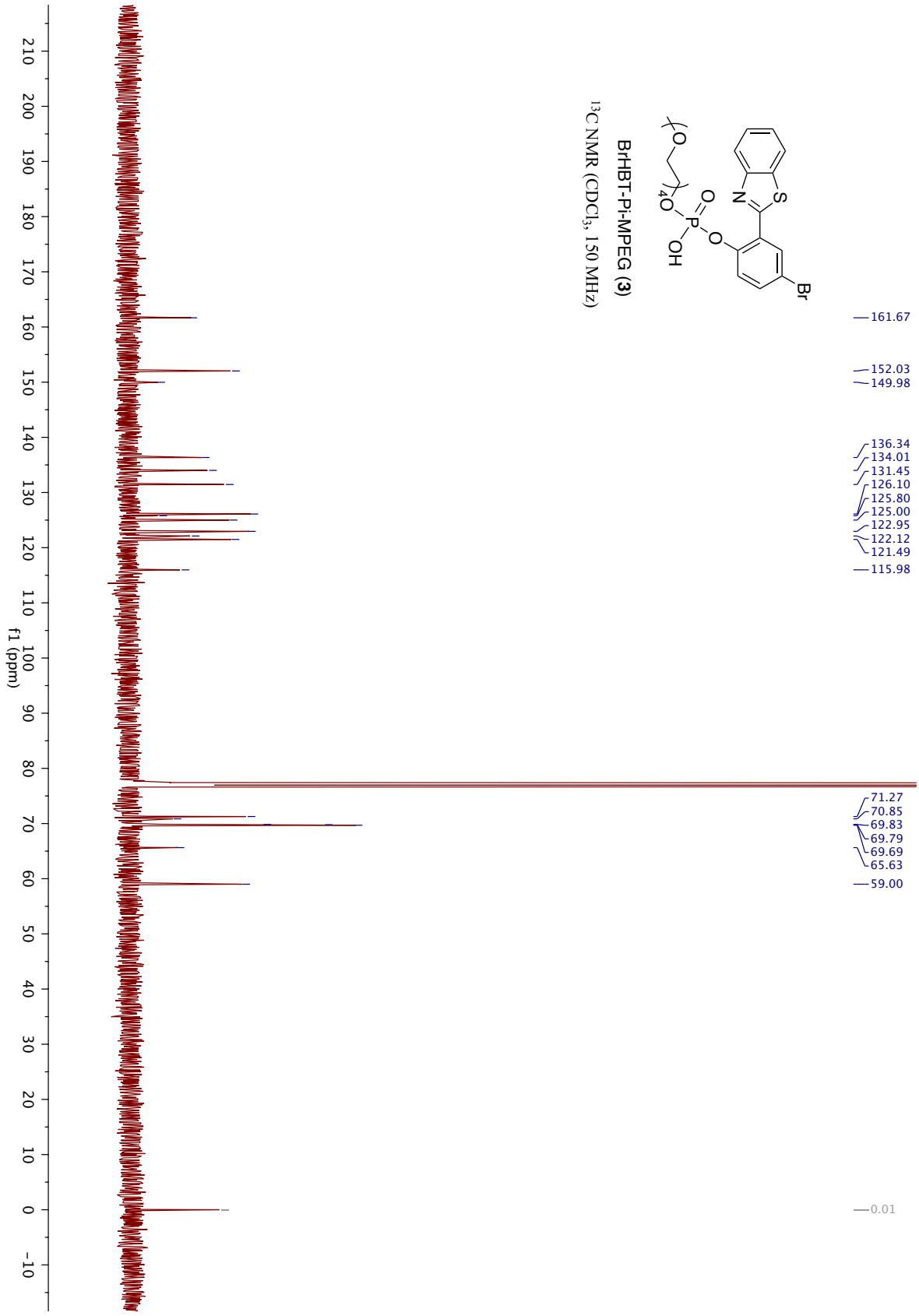

HBT-Pi-MPEG (2)
^{13}C NMR ($CDCl_3$, 150 MHz)
165.48
149.73
149.69
133.81
133.39
129.91
127.42
126.19
124.87
122.35
121.96
121.22
71.86
70.54
70.52
70.50
70.43
70.25
70.21
66.85
66.81
58.99
0.00
f1 (ppm)

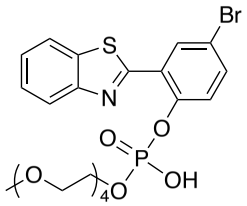


HBT-Pi-MPEG (**2**)

^{31}P NMR ($CDCl_3$, 160 MHz)

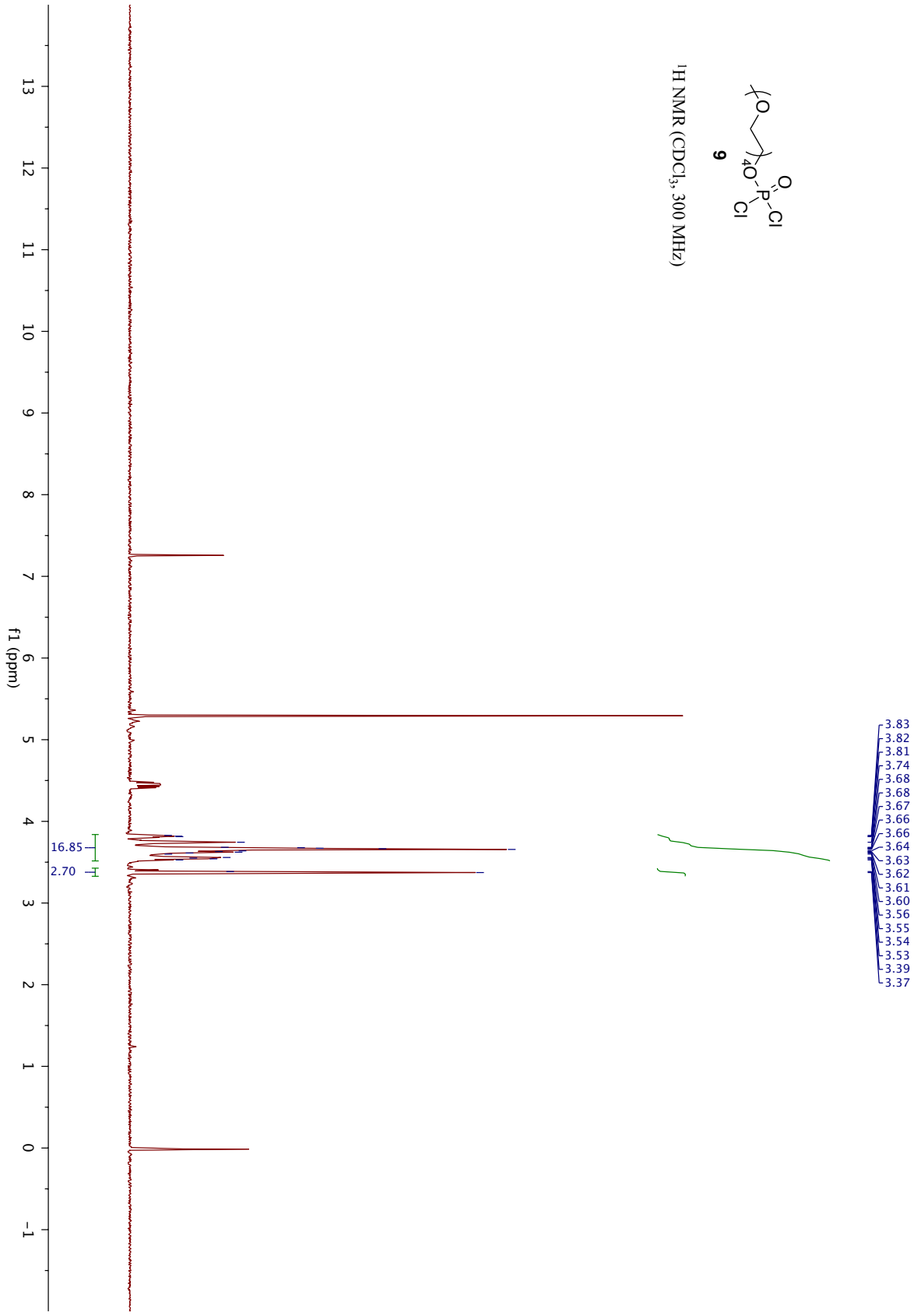

BrHBT-Pi-MPEG (3)
^{1}H NMR ($CDCl_3$, 600 MHz)
f1 (ppm)
0.98
1.07
1.05
0.96
2.15
1.06
2.14
14.77
3.22
8.58
8.58
8.06
8.05
7.87
7.85
7.78
7.76
7.47
7.46
7.45
7.43
7.42
7.41
7.33
7.32
7.31
7.26
4.13
3.53
3.50
3.49
3.45
3.43
3.42
3.41
3.30

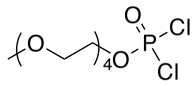

BrHBT-Pi-MPEG (3)
^{13}C NMR ($CDCl_3$, 150 MHz)
161.67
152.03
149.98
136.34
134.01
131.45
126.10
125.80
125.00
122.95
122.12
121.49
115.98
71.27
70.85
69.83
69.79
69.69
65.63
59.00
0.01
f1 (ppm)



BrHBT-Pi-MPEG (**3**)

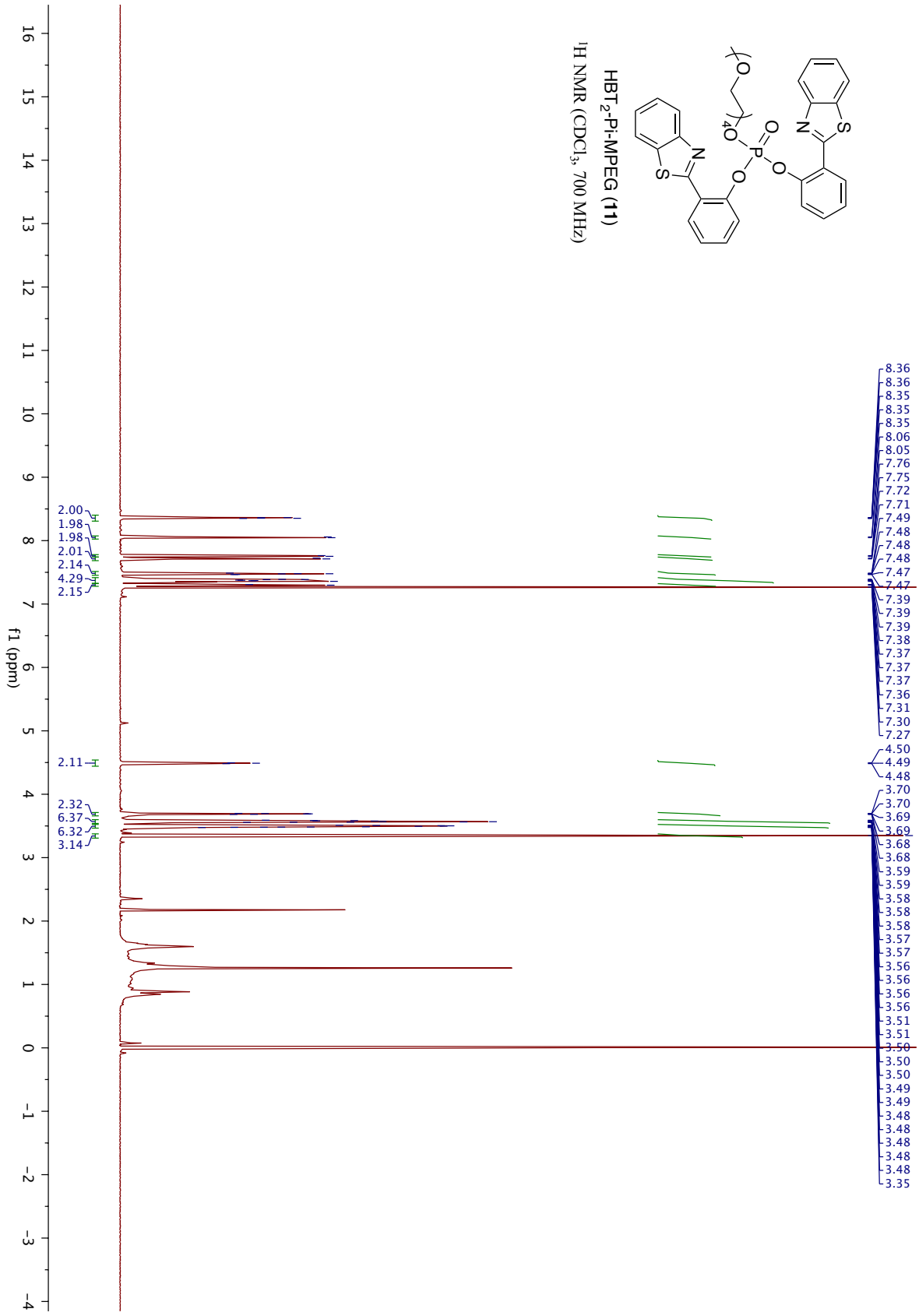
^{31}P NMR ($CDCl_3$, 120 MHz)

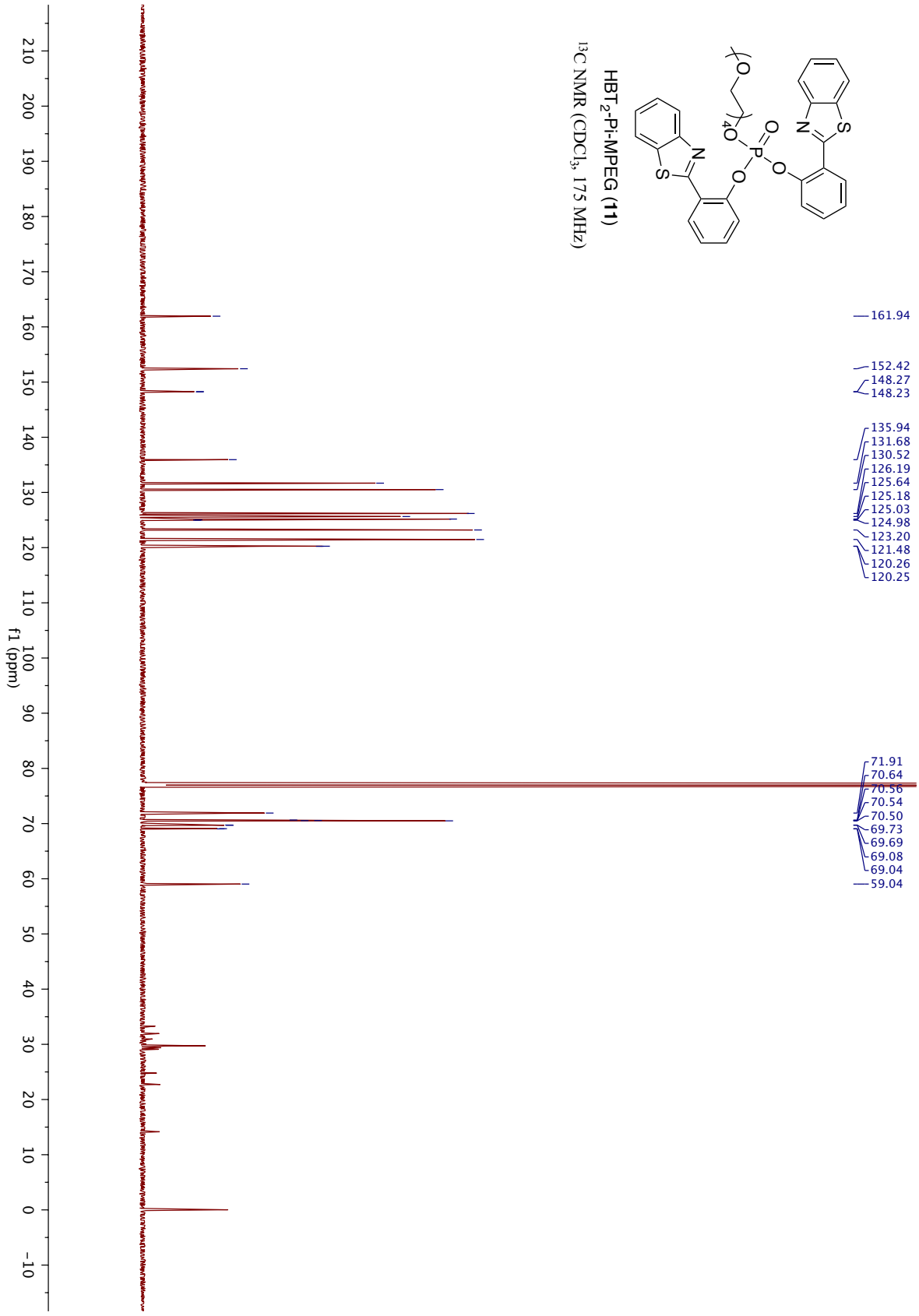

^{1}H NMR (CDCl$_3$, 300 MHz)
9
f1 (ppm)
16.85
2.70
3.83
3.82
3.81
3.74
3.68
3.68
3.67
3.66
3.66
3.64
3.63
3.62
3.61
3.60
3.56
3.55
3.54
3.53
3.39
3.37

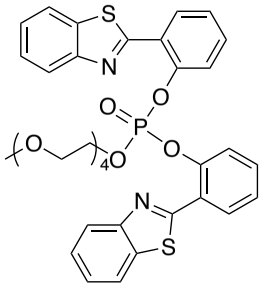


9

^{31}P NMR ($CDCl_3$, 120 MHz)

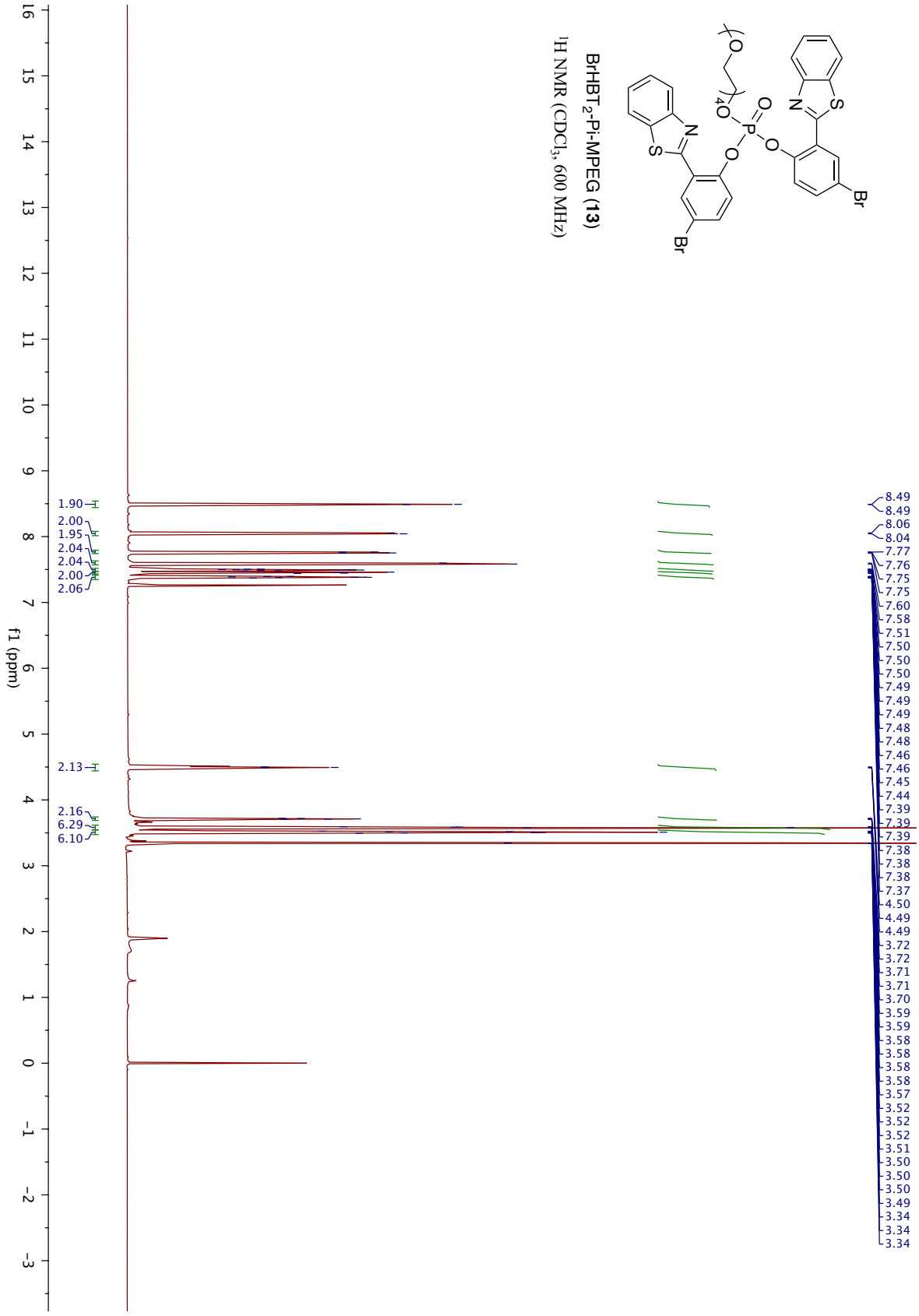

HBT2-Pi-MPEG (11)
1H NMR (CDCl3, 700 MHz)
f1 (ppm)
2.00
1.98
1.98
2.01
2.14
4.29
2.15
2.11
2.32
6.37
6.32
3.14
8.36
8.36
8.35
8.35
8.35
8.06
8.05
7.76
7.75
7.72
7.71
7.49
7.48
7.48
7.48
7.47
7.47
7.39
7.39
7.39
7.38
7.37
7.37
7.37
7.36
7.31
7.30
7.27
4.50
4.49
4.48
3.70
3.70
3.69
3.69
3.68
3.68
3.59
3.59
3.58
3.58
3.58
3.57
3.57
3.56
3.56
3.56
3.56
3.51
3.51
3.50
3.50
3.50
3.49
3.49
3.48
3.48
3.48
3.48
3.48
3.35

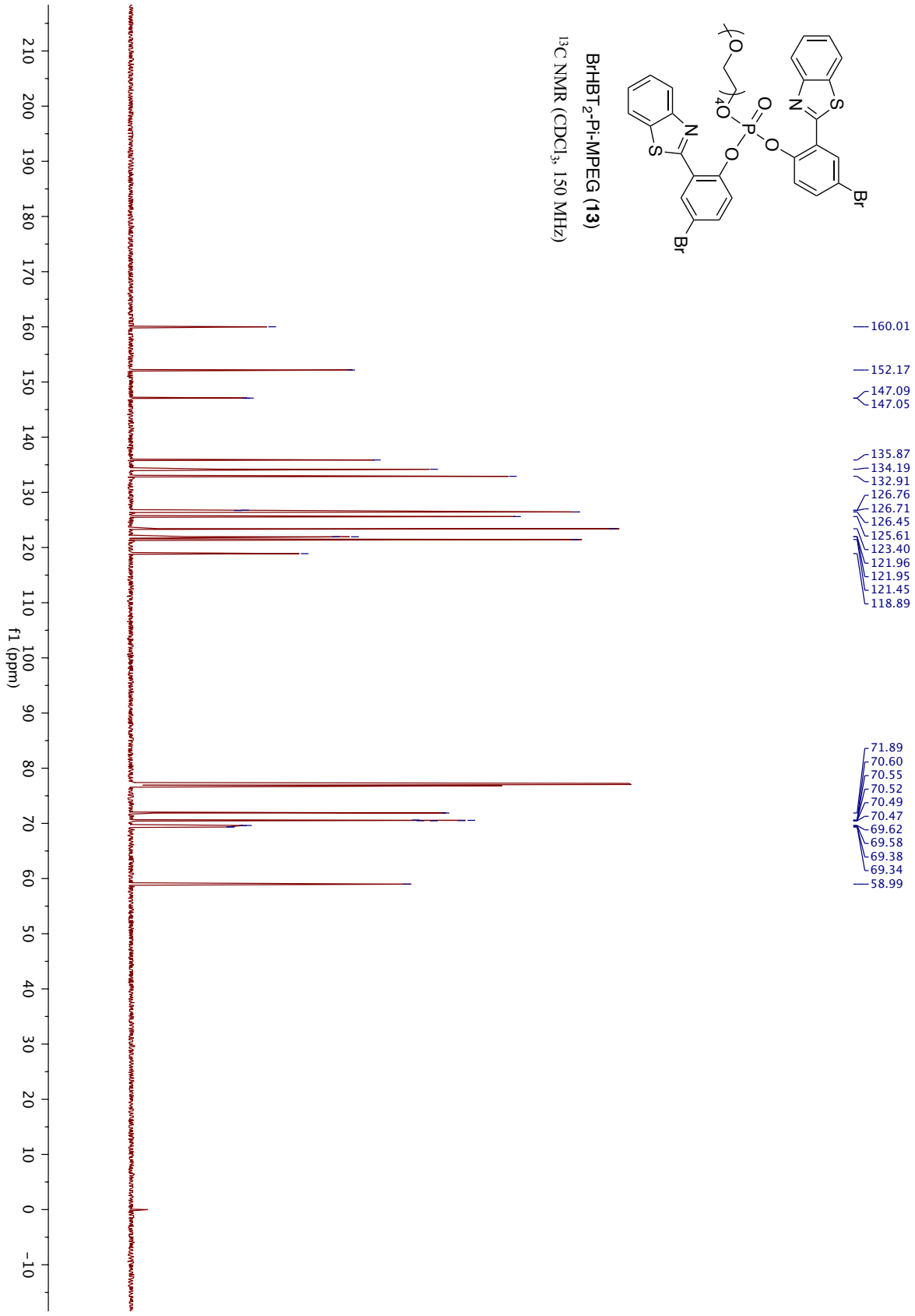

HBT2-Pi-MPEG (11)
13C NMR (CDCl3, 175 MHz)
161.94
152.42
148.27
148.23
135.94
131.68
130.52
126.19
125.64
125.18
125.03
124.98
123.20
121.48
120.26
120.25
71.91
70.64
70.56
70.54
70.50
69.73
69.69
69.08
69.04
59.04
f1 (ppm)

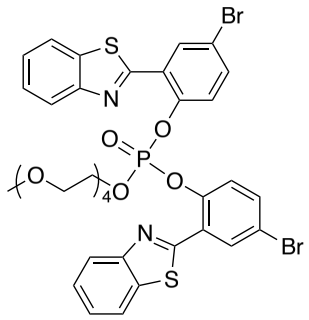


HBT$_2$-Pi-MPEG (**11**)

^{31}P NMR (CD$_3$OD, 120 MHz)


BrHBT2-Pi-MPEG (13)
1H NMR (CDCl3, 600 MHz)
f1 (ppm)
1.90
2.00
1.95
2.04
2.04
2.00
2.06
2.13
2.16
6.29
6.10
8.49
8.49
8.06
8.04
7.77
7.76
7.75
7.75
7.60
7.58
7.51
7.50
7.50
7.50
7.49
7.49
7.49
7.49
7.48
7.48
7.46
7.46
7.45
7.44
7.39
7.39
7.39
7.38
7.38
7.38
7.37
4.50
4.49
4.49
3.72
3.72
3.71
3.71
3.70
3.59
3.59
3.58
3.58
3.58
3.58
3.57
3.52
3.52
3.52
3.51
3.50
3.50
3.50
3.49
3.34
3.34
3.34


BrHBT2-Pi-MPEG (13)
13C NMR (CDCl3, 150 MHz)
160.01
152.17
147.09
147.05
135.87
134.19
132.91
126.76
126.71
126.45
125.61
123.40
121.96
121.95
121.45
118.89
71.89
70.60
70.55
70.52
70.49
70.47
69.62
69.58
69.38
69.34
58.99
f1 (ppm)



BrHBT$_2$-Pi-MPEG (**13**)

^{31}P NMR ($CDCl_3$, 150 MHz)