STRUCTURAL BIOLOGY

Volume 78 (2022)

Supporting information for article:

Mechanism-based cross-linking probes capture the *Escherichia coli* ketosynthase FabB in conformationally-distinct catalytic states

Aochiu Chen, Jefferey T. Mindrebo, Tony D. Davis, Woojoo E. Kim, Yohei Katsuyama, Ziran Jiang, Yasuo Ohnishi, Joseph P. Noel and Michael D. Burkart

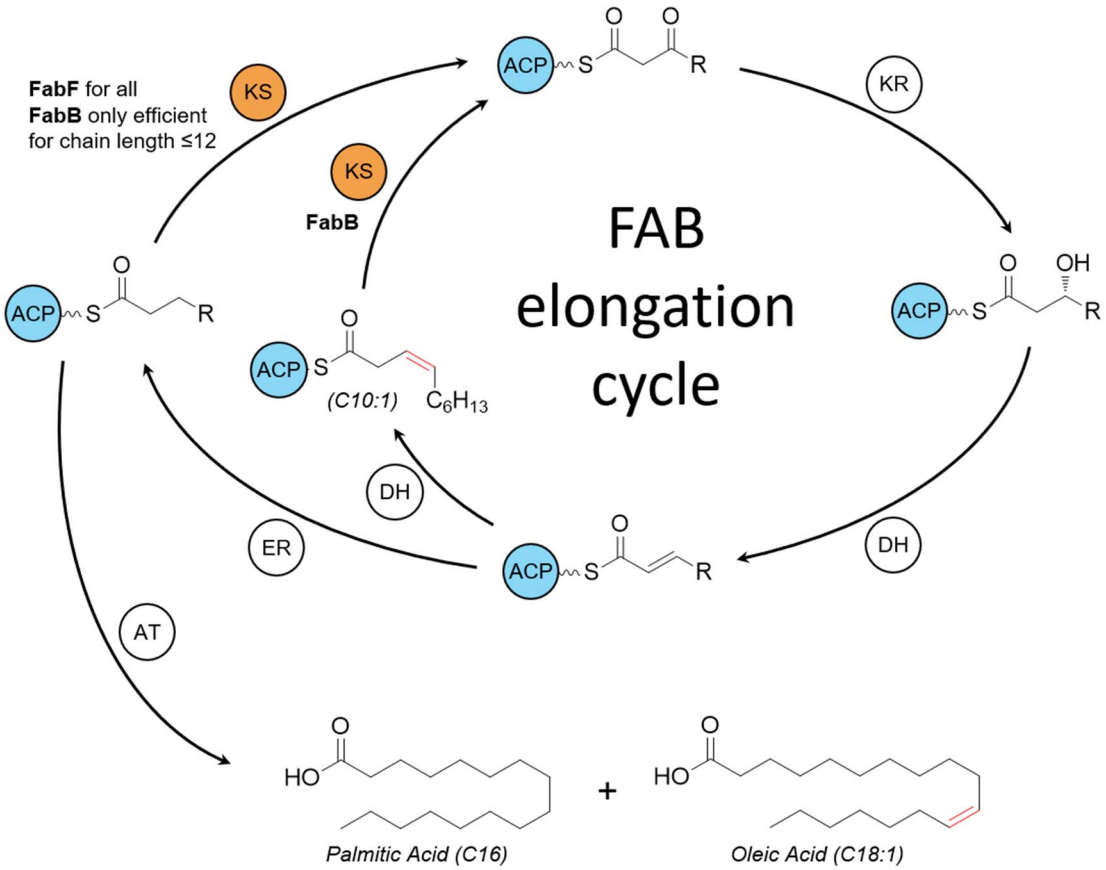


Figure S1 The elongation cycle of fatty acid biosynthesis in *E. coli*. While FabF has a broad substrate spectrum, FabB specializes in elongating *cis*-3-decenoyl acid (C10:1) and has reduced activity towards elongating acyl chain longer than C12. The two major products of *E. coli* FAB are listed at the bottom. Abbreviation: FAB = fatty acid biosynthesis, ACP = acyl carrier protein, KS = ketosynthase, KR = ketoreductase, DH = dehydratase, ER = enoylreductase, AT = acyltransferase.

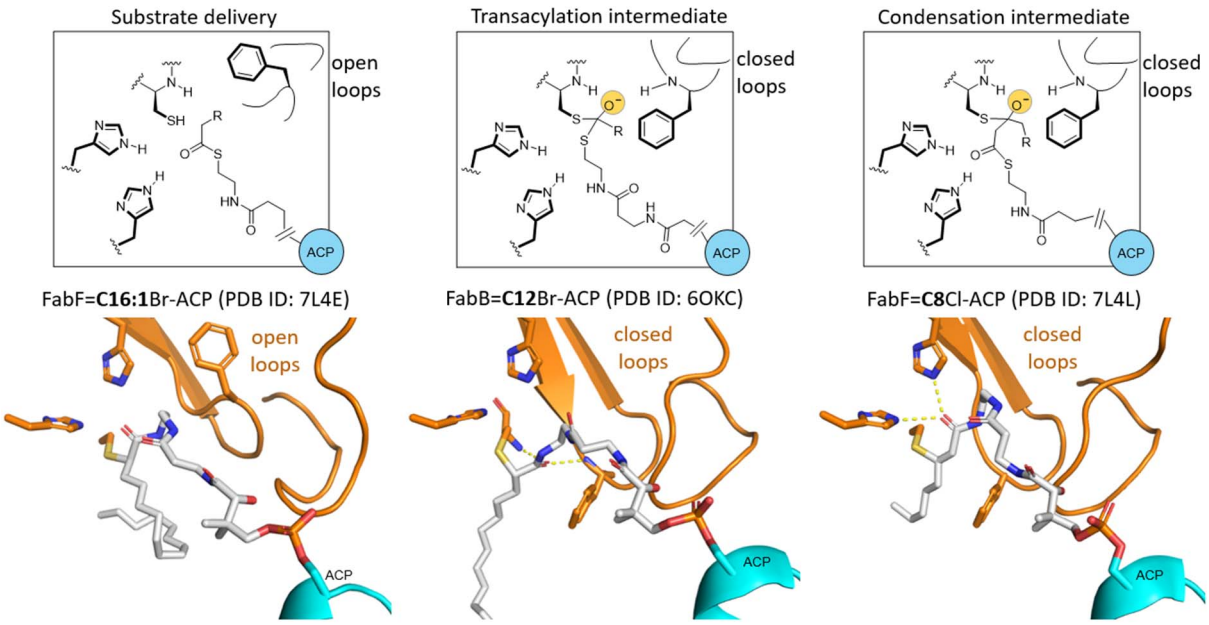


Figure S2 Previous published KS-ACP structures that mimic different stages of catalysis. The native states of KS are depicted in the first row, and the second row shows each of the crosslinked complexes that has similar arrangement with each state. The α-bromo probes crosslink FabF and ACP in mainly the open gating loops conformation (left column). In contrast, they crosslink FabB and ACP in the closed gating loops conformation (middle column) with the acyl carbonyl residing in the oxyanion hole (shown as yellow ball). The chloroacrylate probes trap both FabF-ACP (right column) and FabB-ACP (this work) complexes in the closed loops conformation with the acyl carbonyl coordinating with the two active site histidine residues, which mimic the condensation intermediate.

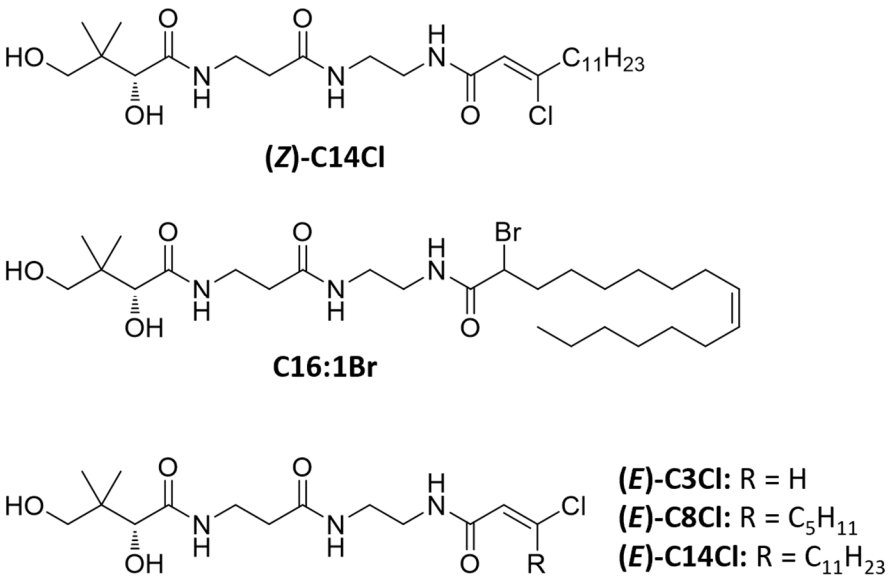


Figure S3 Crosslinkers used in this study. The top two crosslinkers were used to obtain the two FabB-ACP crosslinked structures while the three *E* form crosslinkers were used for crosslinking activity study.

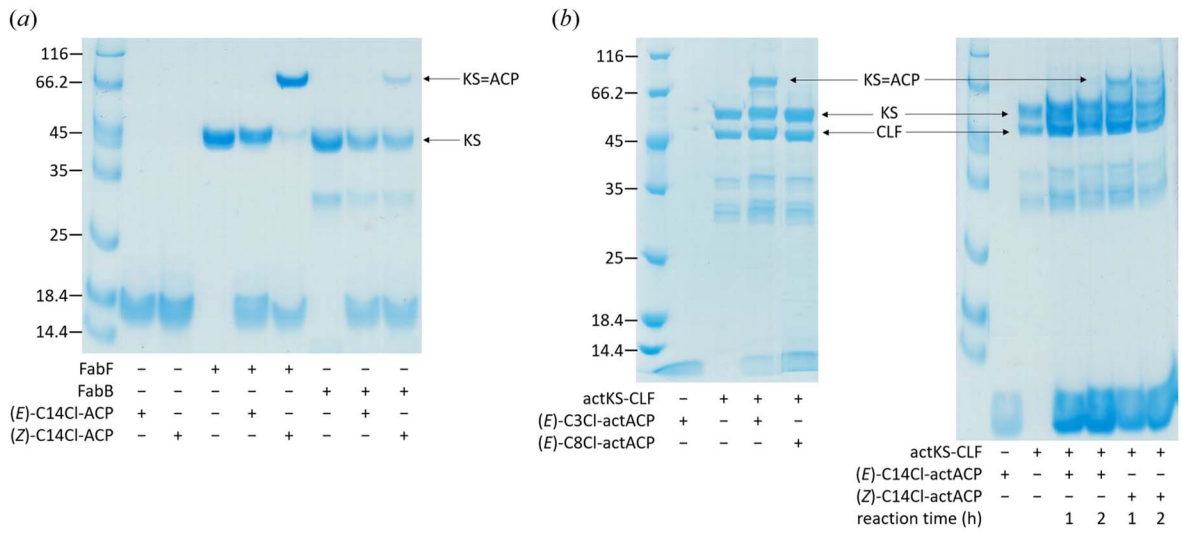


Figure S4 Crosslinking test on three different KSs shows stereospecificity toward *Z*-form chloroacrylate crosslinker. Samples were run on 12% SDS-PAGE and stained by Coomassie Brilliant Blue. (*a*) (*Z*)-C14Cl can crosslink ACP with FabF or FabB whereas (*E*)-C14Cl shows no crosslinking activity. (*b*) Actinorhodin ketosynthase (actKS-CLF) exists as a heterodimer, in which the chain length factor subunit lacks the active site residues and is slightly smaller than KS subunit in size. Only (*E*)-C3Cl and (*Z*)-C14Cl show crosslinking activity toward actKS-CLF and actACP.

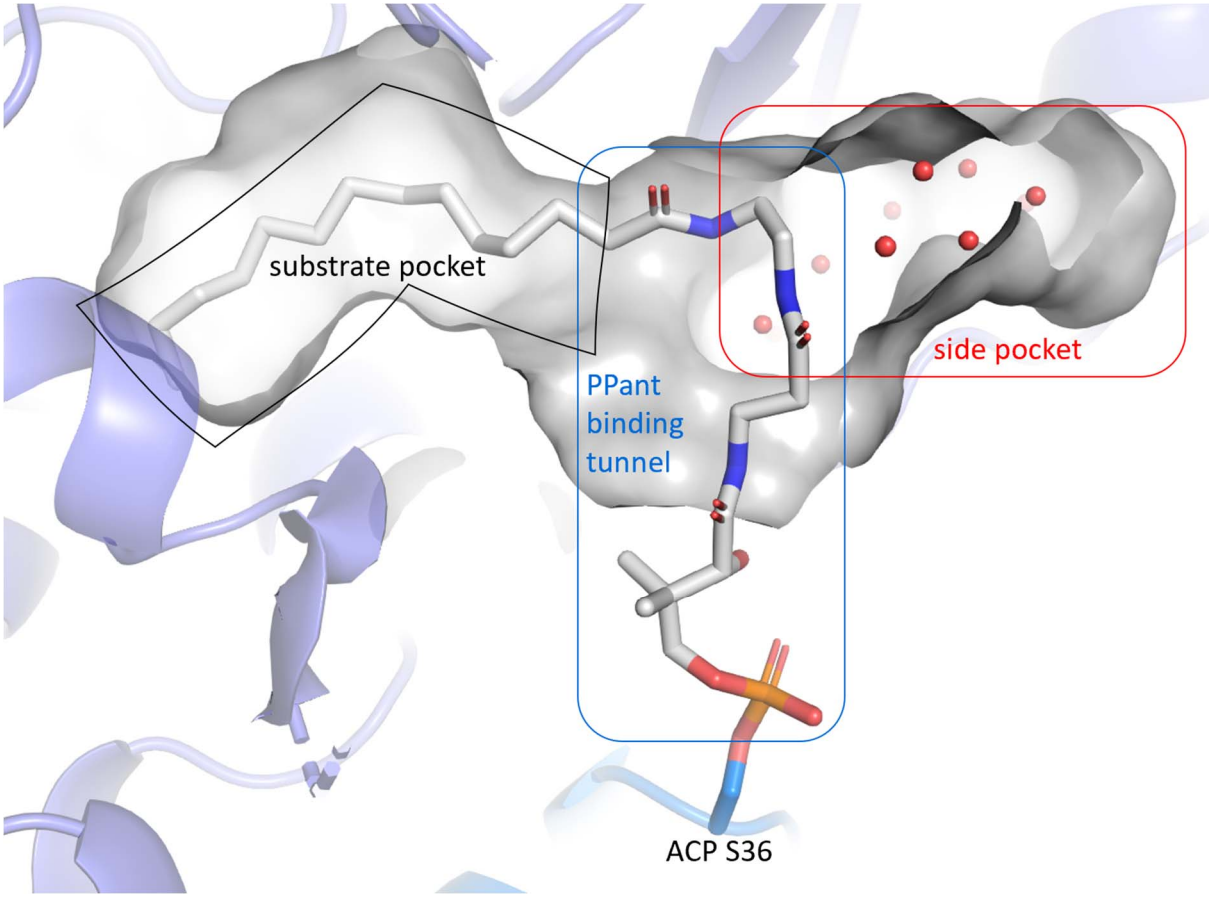


Figure S5 PPant binding tunnel, substrate pocket, and the side pocket form a "τ" shaped space inside FabB. This figure is generated from the FabB=C14-ACP (PDB ID: 7SQI) structure.

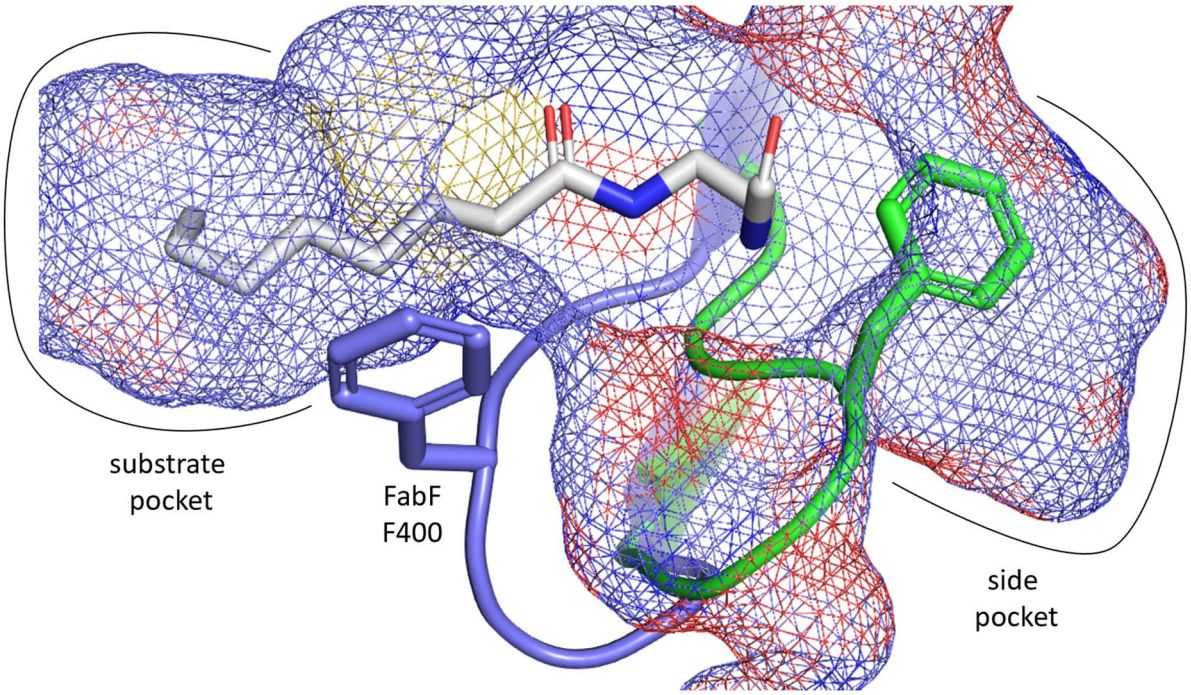


Figure S6 The side pocket accommodates the open gating loop 1. Structure of a KS in closed conformation (PDB ID: 7L4L) is shown in blue cartoon and mesh, depicting gating loop 1 and available space around the active site, respectively. Superimposition of a gating loop 1 in open conformation (green, PDB ID: 6OKG) fits the open loop into the space we termed the side pocket.

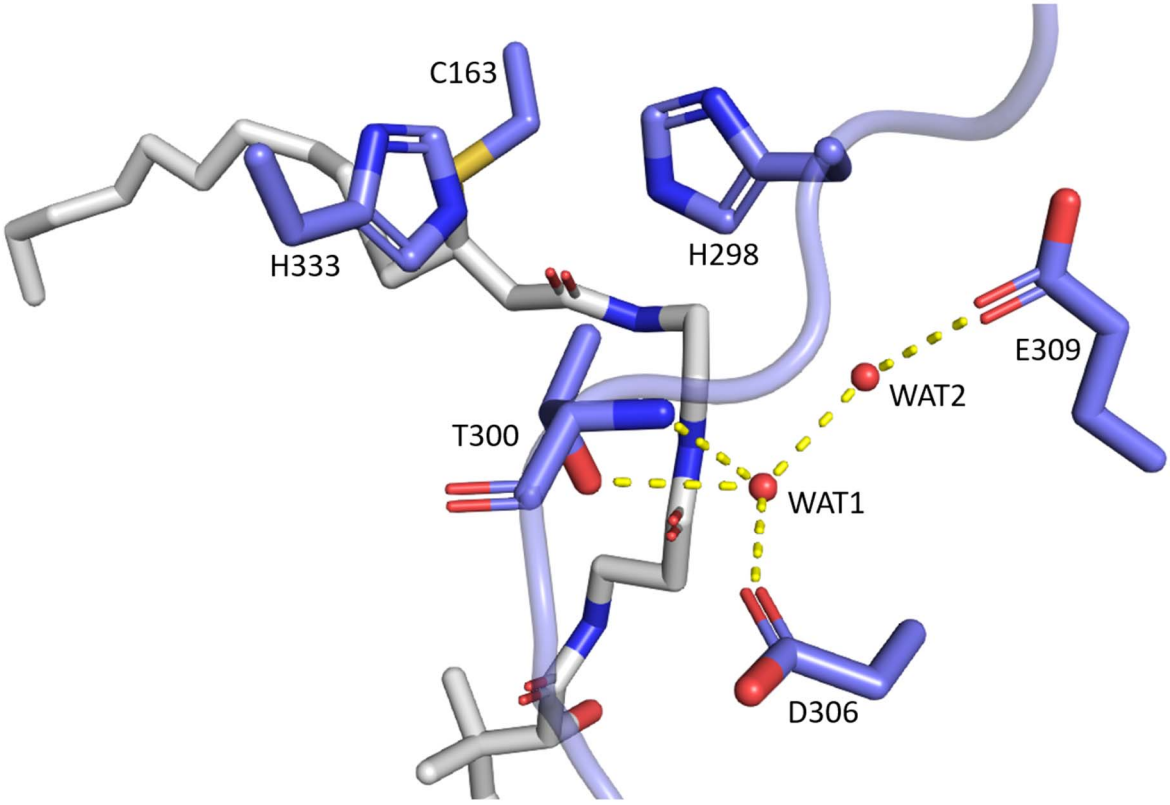


Figure S7 The two water molecules, WAT1 and WAT2, respectively, in proximity of the catalytic histidine residues (FabB=C14-ACP, PDB ID: 7SQI). Polar contacts are shown by yellow dashed lines. All residues shown are 100 % conserved over 461 FabB sequences aligned.

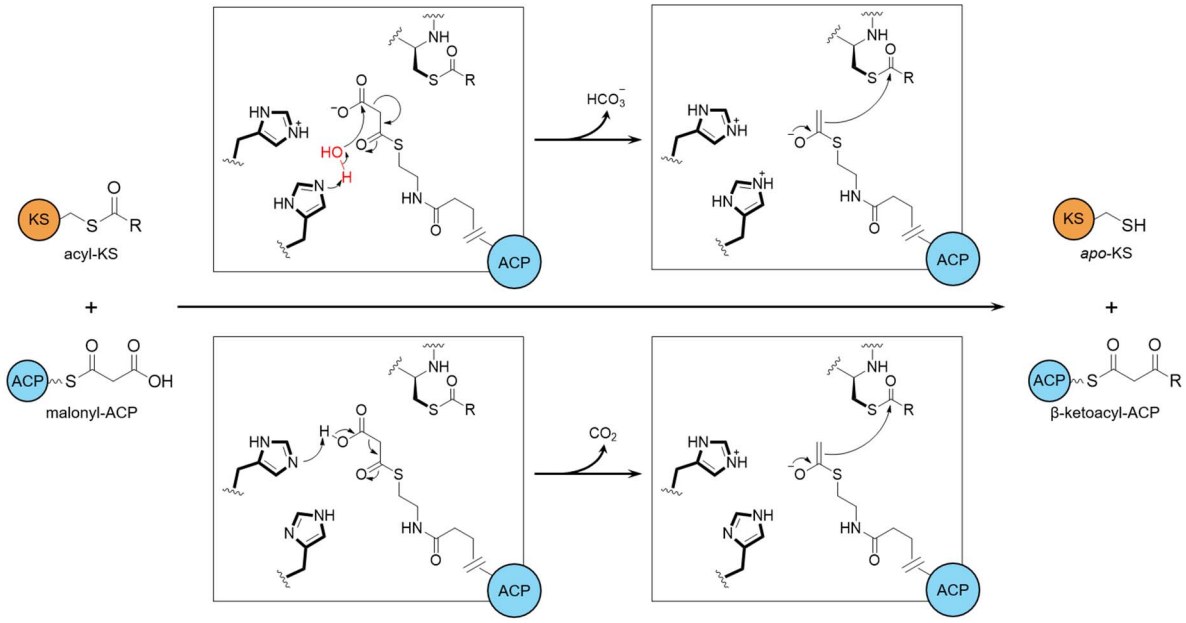


Figure S8 Two proposed condensation mechanisms that release bicarbonate (top) or carbon dioxide (bottom), respectively. The water molecule needed for bicarbonate formation is highlighted red. This figure is adapted from the review article written by Heil *et al*. (Heil *et al*. 2019)

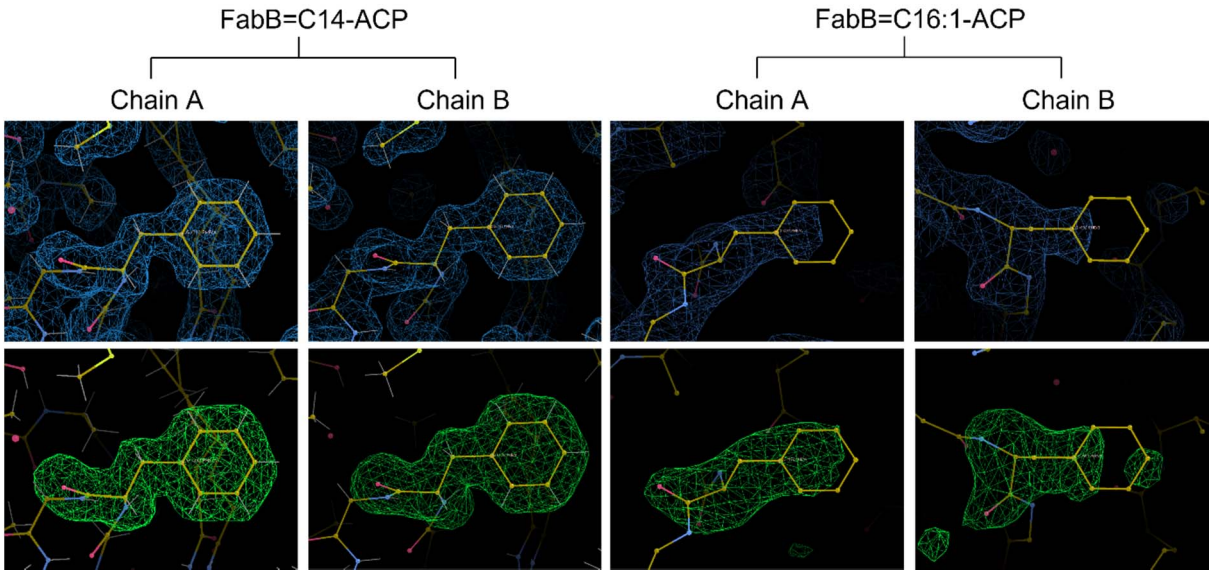
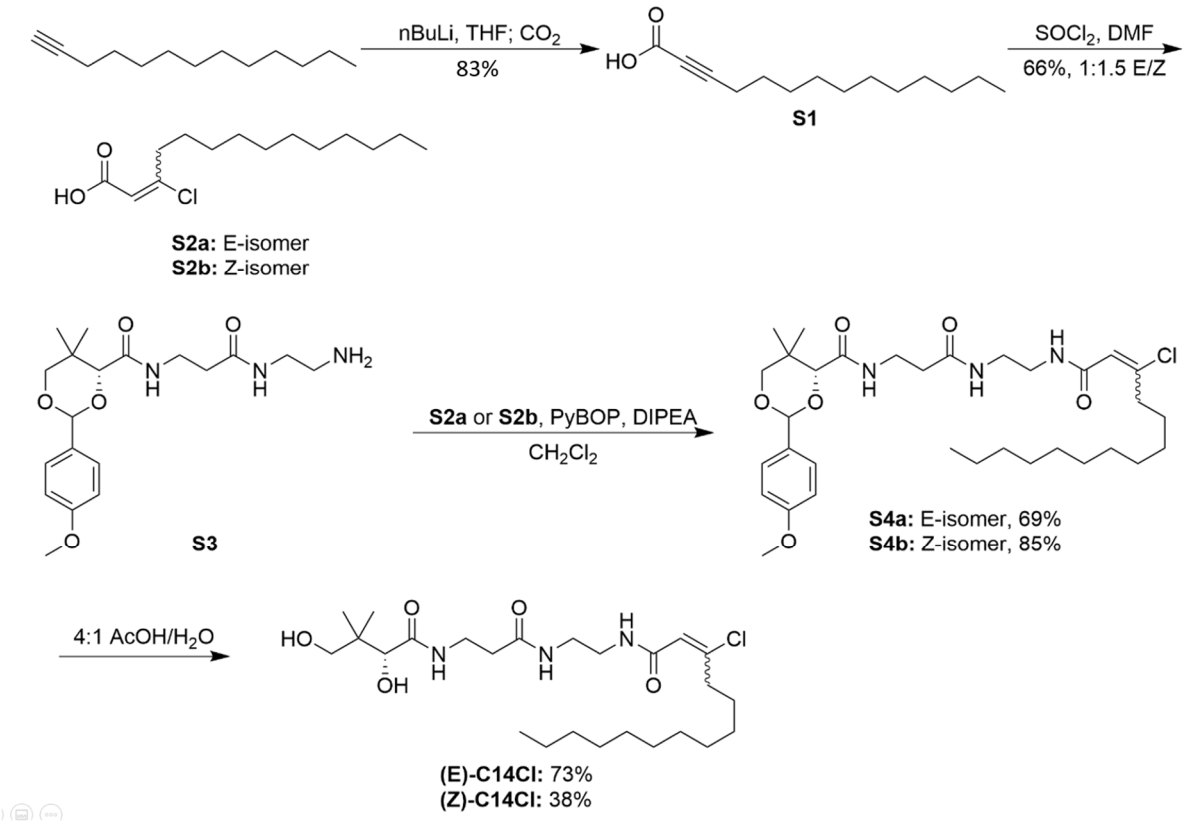


Figure S9 Density map of the gating residue F392. The electron density contoured at sigma value 1.0 is shown in blue mesh (top row). The Fo-Fc polder omit map calculated by omitting F392 is contoured at sigma value 4.5 and shown in green mesh (bottom row).

S1. Synthesis of crosslinking probes

Chemical reagents were purchased from Acros, Fluka, Sigma-Aldrich, or TCI. Deuterated NMR solvents were purchased from Cambridge Isotope Laboratories. All reactions were conducted with vigorously dried anhydrous solvents that were obtained by passing through a solvent column exposed of activated A2 alumina. All reactions were performed under positive pressure of argon in flame-dried glassware sealed with septa and stirred with Teflon coated stir bars using an IKAMAG TCT-basic mechanical stirrer (IKA GmbH). Analytical Thin Layer Chromatograpy (TLC) was performed on Silica Gel 60 F254 precoated glass plates (EM Sciences). Visualization was achieved with UV light and/or appropriate stain (I_2 on SiO_2, $KMnO_4$, bromocresol green, dinitrophenylhydrazine, ninhydrin, or ceric ammonium molybdate). Flash column chromatography was carried out with Geduran Silica Gel 60 (40–63 mesh) from EM Biosciences. Yield and characterization data correspond to isolated, chromatographically, and spectroscopically homogeneous materials. ^{1}H NMR spectra were recorded on Varian Mercury 400, Varian Mercury Plus 400, or JEOL ECA500 spectrometers. ^{13}C NMR spectra were recorded at 100 MHz on Varian Mercury 400 or Varian Mercury Plus 400 spectrometers. Chemical shifts for ^{1}H NMR and ^{13}C NMR analyses were referenced to the reported values (Gottlieb *et al.*, 1997) using the signal from the residual solvent for ^{1}H spectra, or to the ^{13}C signal from the deuterated solvent. Chemical shift δ values for the ^{1}H and ^{13}C spectra are reported in parts per millions (ppm) relative to these referenced values, and multiplicities are abbreviated as s=singlet, d=doublet, t=triplet, q-quartet, m=multiplet, b=broad. All ^{13}C NMR spectra were recorded with complete proton decoupling. FID files were processed using MestreNova 10.0 (MestreLab Research). Electrospray ionization (ESI) mass spectrometric analyses were preformed using a ThermoFinnigan LCQ Deca spectrometer. Spectral data and procedures are provided for all new compounds and copies of spectra have been provided.



Scheme S1 Synthesis of chloroacrylamide crosslinking probes. PyBOP= benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate; DIPEA=N,N-diisopropylethylamine; AcOH=acetic acid.

2-tetradecynoic acid (S1). In a 50 mL round-bottomed flask, 1-tridecyne (0.12 mL, 0.52 mmol, 1.0 equiv.) and 5 mL THF were added. The vessel was cooled to 0 °C before the slow, dropwise addition of nBuLi (1.6 M in hexanes, 0.40 mL, 0.64 mmol, 1.1 equiv.). The reaction was stirred for 2 hours at 0 °C before flushing with carbon dioxide gas. After stirring for 3 hours at 0 °C, the reaction was warmed to room temperature (20 °C) and stirred overnight for 17 hours. The reaction was quenched with saturated aqueous NH_4Cl (10 mL) and extracted with ethyl acetate (4x25 mL). The combined organic extracts were dried ($MgSO_4$), filtered, and concentrated by rotary evaporation. Purification by silica flash chromatography (9:1 hexanes/ethyl acetate → 3:2 hexanes/ethyl acetate + 1% acetic acid) afforded carboxylic acid **S1** (96.7 mg, 83%, white solid).

TLC: R_f 0.21 (4:1 hexanes/ethyl acetate + 1% acetic acid). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ10.98 (s, 1H), 2.34 (t, J = 7.0 Hz, 1H), 1.81– 1.55 (m, 2H), 1.39 (s, 3H), 1.25 (s, 14H), 0.87 (t, J = 5.6 Hz, 3H).**^{13}C-NMR** (100 MHz, $CDCl_3$): δ 158.83, 92.92, 72.77, 32.03, 29.71, 29.53, 29.13, 28.94, 27.51, 22.81, 19.01, 18.88, 18.74, 14.19.

3-chloro-2-tetradecenoic acid (S2). In a 10 mL pear-shaped flask, 2-tetradecynoic acid **S1** (126.1 mg, 0.5621 mmol, 1.0 equiv.) and 5 mL DMF were added. The vessel was cooled to 0 °C before the addition of $SOCl_2$ (48.9 µL, 0.675 mmol, 1.2 equiv.). After stirring for 1 hour at 0 °C, the reaction was warmed to room temperature (20 °C). After 3 hours at room temperature, additional $SOCl_2$ was added (20 µL) and stirring was continued overnight. After 15 hours, additional $SOCl_2$ was added (48.9 µL). After 2 hours, the reaction was concentrated by rotary evaporation, and the resulting liquid was poured over 12 grams of ice and extracted with diethyl ether (4x25 mL). The combined organic extracts were dried ($MgSO_4$), filtered, and concentrated by rotary evaporation. Purification by silica flash chromatography (9:1 hexanes/diethyl ether → 4:1 hexanes/diethyl ether) afforded (*E*)-3-chloro-2-tetradecenoic acid **S2a** (39.2 mg, yellow solid) and (*Z*)-3-chloro-2-tetradecenoic acid **S2b** (57.2 mg, brown solid)

(*E*)-3-chloro-2-tetradecenoic acid (S2a). TLC: R_f 0.61 (1:1 hexanes/diethyl ether). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 6.09 (s, 1H), 2.98 (t, 2H), 1.63 (m, 2H), 1.26 (s, 14H), 0.95 – 0.80 (t, 3H).**^{13}C-NMR** (100 MHz, $CDCl_3$): δ 170.11, 161.06, 118.44, 36.11, 32.08, 29.78, 29.62, 29.51, 29.45, 28.90, 27.83, 22.86, 14.27.

(*Z*)-3-chloro-2-tetradecenoic acid (S2b). TLC: R_f 0.45 (3:2 hexanes/diethyl ether). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 6.05 (s, 1H), 2.46 (t, J = 7.4 Hz, 1H), 1.63 (m, 2H), 1.26 (s, 15H), 0.88 (t, 3H).**^{13}C-NMR** (100 MHz, $CDCl_3$): δ 169.52, 154.24, 115.60, 41.74, 32.05, 29.73, 29.58, 29.47, 29.38, 28.67, 27.35, 22.83, 14.25.

(4*R*)-*N*-(3-((2-((*E*)-3-chlorotetradec-2-enamido)ethyl)amino)-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide (S4a). In a 10 mL pear-shaped flask, (4R)-N-(3-((2-aminoethyl)amino)-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide S3 (Meier & Burkart, 2009) (57.0 mg, 0.105 mmol, 1.0 equiv.), (E)-3-chloro-2-tetradecenoic acid S2a (39.2 mg, 0.150 mmol, 1.0 equiv.), and 5 mL CH2Cl2 were added. To the solution was added PyBOP (94.3 mg, 0.181 mmol, 1.7 equiv.) and DIPEA (78.5 µL, 0.451 mmol, 3.0 equiv.). After 20 hours, the volatiles were removed by rotary evaporation. Purification by silica flash chromatography (3:1 hexanes/ethyl acetate → 39:1 ethyl acetate/methanol) afforded S4a (64.8 mg, 69%, white solid).

TLC: R_f 0.31 (ethyl acetate). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 7.41 (d, J = 8.2 Hz, 2H), 7.02 (s, 1H), 6.89 (d, J = 8.3 Hz, 3H), 6.75 (s, 1H), 5.97 (s, 1H), 5.44 (s, 1H), 4.06 (s, 1H), 3.79 (s, 3H), 3.65 (q, J = 11.4 Hz, 2H), 3.50 (m, 2H), 3.31 (s, 3H), 2.97 (t, J = 7.0 Hz, 2H), 2.38 (bs, 2H), 1.58 (bs, 2H), 1.24 (s, 15H), 1.05 (d, J = 6.4 Hz, 6H), 0.84 (t, J = 6.3 Hz, 3H). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ 172.05, 169.96, 164.95, 160.37, 153.67, 130.15, 127.66, 120.97, 113.83, 101.46, 83.89, 55.42, 39.84, 39.59,

36.36, 35.42, 35.05, 33.17, 32.01, 29.75, 29.73, 29.65, 29.54, 29.44, 29.05, 27.87, 22.78, 21.94, 19.21, 14.23.

(4*R*)-*N*-(3-((2-((*Z*)-3-chlorotetradec-2-enamido)ethyl)amino)-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide (S4b). Prepared as described for S4a from S3 (Meier & Burkart, 2009) and S2b to afford S4b (116.3 mg, 85%, white solid).

TLC: R_f 0.11 (ethyl acetate). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 7.39 (d, *J* = 8.7 Hz, 2H), 7.06 (t, *J* = 6.2 Hz, 1H), 6.95 (s, 2H), 6.87 (d, *J* = 8.7 Hz, 2H), 5.96 (s, 1H), 5.43 (s, 1H), 4.04 (s, 1H), 3.77 (s, 3H), 3.67 (m, 3H), 3.48 (m, 2H), 3.33 (s, 4H), 3.10 (m, 1H), 2.36 (m, 3H), 1.53 (s, 2H), 1.37 (q, *J* = 6.7, 6.1 Hz, 7H), 1.23 (s, 15H), 1.06 (d, *J* = 10.2 Hz, 6H), 0.83 (t, *J* = 6.6 Hz, 3H). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ 172.19, 169.68, 165.22, 160.25, 144.34, 130.18, 127.61, 119.50, 113.75, 101.34, 83.84, 78.43, 55.36, 55.28, 43.24, 40.83, 39.53, 35.99, 35.05, 33.07, 31.93, 29.66, 29.64, 29.53, 29.37, 28.70, 27.28, 22.72, 21.86, 19.15, 18.56, 17.13, 14.16, 12.61.

(R,E)-3-chloro-N-(2-(3-(2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl)tetradec-2-enamide ((E)-C14Cl). In a 20 mL vial, **S4a** (64.8 mg, 0.104 mmol, 1.0 equiv.) and 1.0 mL 4:1 AcOH/H_2O were added. After 21 hours, the mixture was concentrated by rotary evaporation, and azeotroped from cyclohexane (3x10 mL) and benzene (3x10 mL). Purification by silica flash chromatography (ethyl acetate → 85:15 ethyl acetate/methanol) afforded **((*E*)-C14Cl** (38.2 mg, 73%, clear oil).

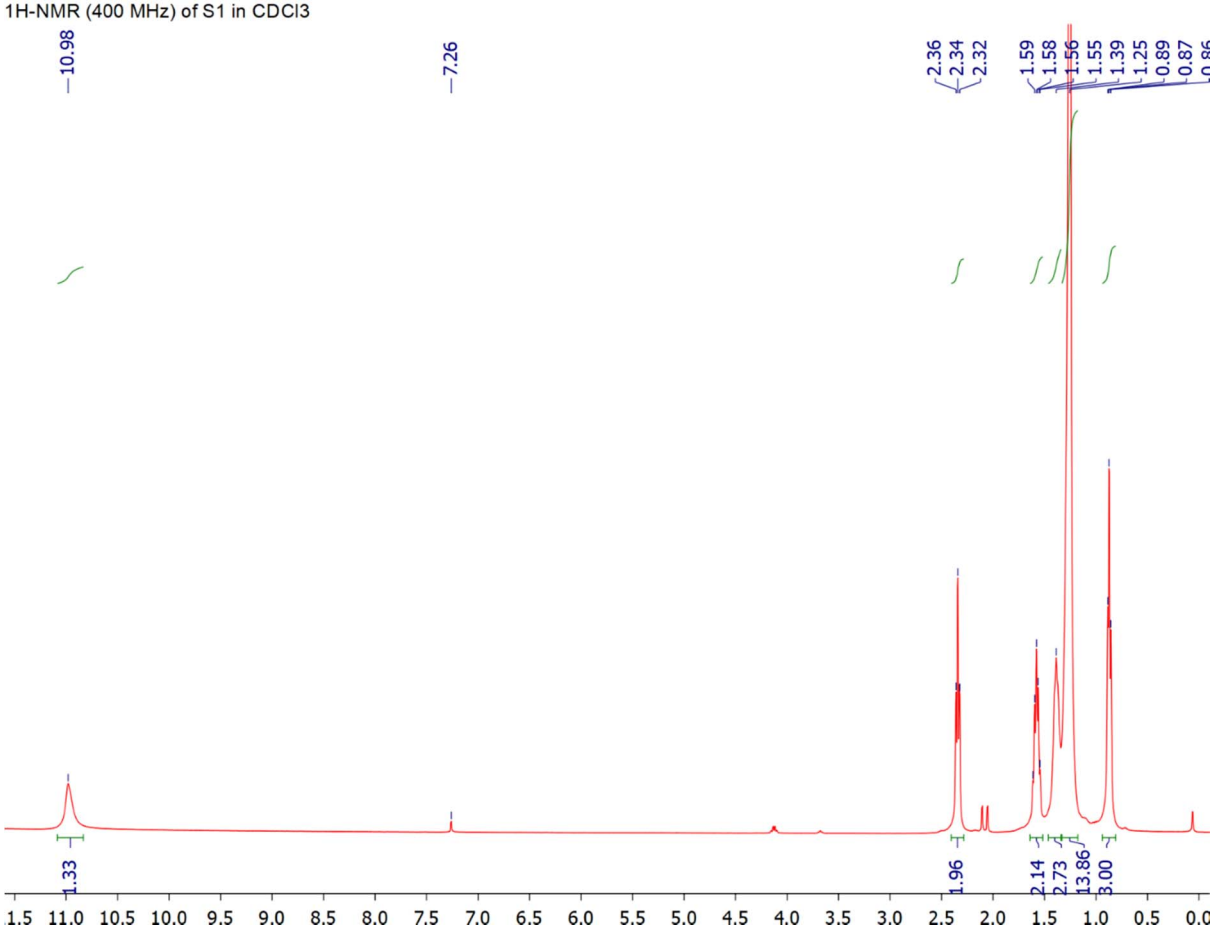
TLC: R_f 0.31 (9:1 ethyl acetate/methanol). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 7.54 (t, *J* = 5.1 Hz, 1H), 7.30 (s, 1H), 7.17 (s, 1H), 6.03 (s, 1H), 3.98 (s, 1H), 3.68–3.21 (m, 9H), 2.95 (t, 2H), 2.43 (s, 2H), 1.58 (s, 2H), 1.24 (s, 18H), 0.89 (d, *J* = 14.4 Hz, 6H), 0.84 (t, *J* = 6.6 Hz, 3H). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ 174.38, 172.55, 165.33, 153.90, 120.96, 77.58, 77.16, 76.74, 70.73, 39.86, 39.49, 39.37, 36.10, 35.50, 32.04, 29.79, 29.76, 29.68, 29.58, 29.47, 29.12, 27.92, 22.81, 21.34, 20.82, 14.25.

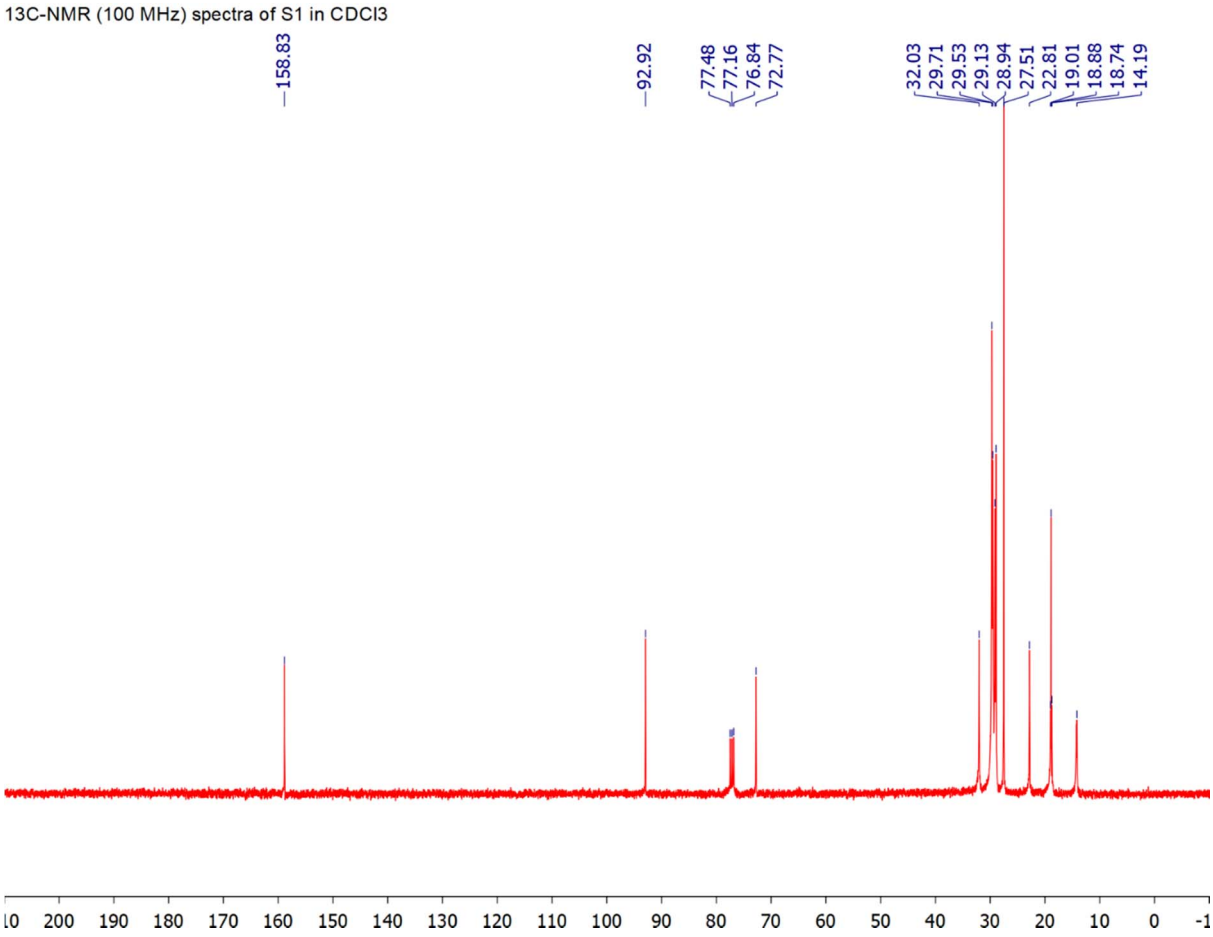
(*R*,*Z*)-3-chloro-*N*-(2-(3-(2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl)tetradec-2-enamide ((*Z*)-C14Cl). Prepared as described for **(*E*)-C14Cl** from **S4b** to afford **((*Z*)-C14Cl** (35.9 mg, 38%, white solid).

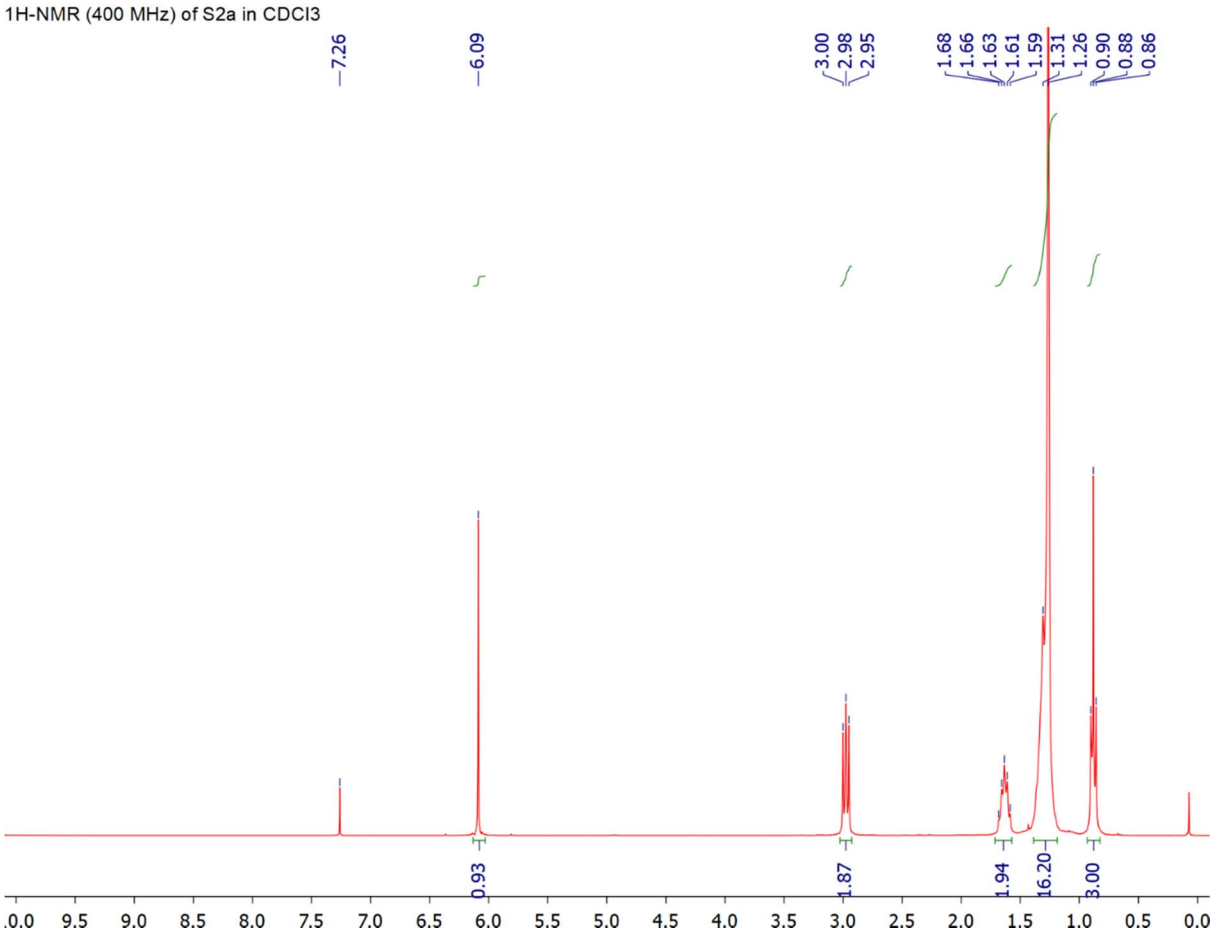
TLC: R_f 0.29 (9:1 ethyl acetate/methanol). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 7.33 (t, 1H), 6.70 (d, 2H), 5.97 (s, 1H), 4.01 (s, 1H), 3.68 (m, 1H), 3.50 (m, 5H), 3.30 (m, 1H), 2.42 (m, 4H), 1.59 (m, 2H), 1.28 (m, 15H), 1.03 (s, 3H), 0.93 (s, 3H), 0.88 (t, *J* = 7.3 Hz, 3H). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ 172.27, 165.72, 144.67, 119.80, 100.12, 41.09, 40.42, 39.74, 39.48, 36.35, 35.37, 32.12, 29.83, 29.69, 29.55, 29.50, 28.82, 27.40, 22.90, 21.92, 20.68, 14.35.

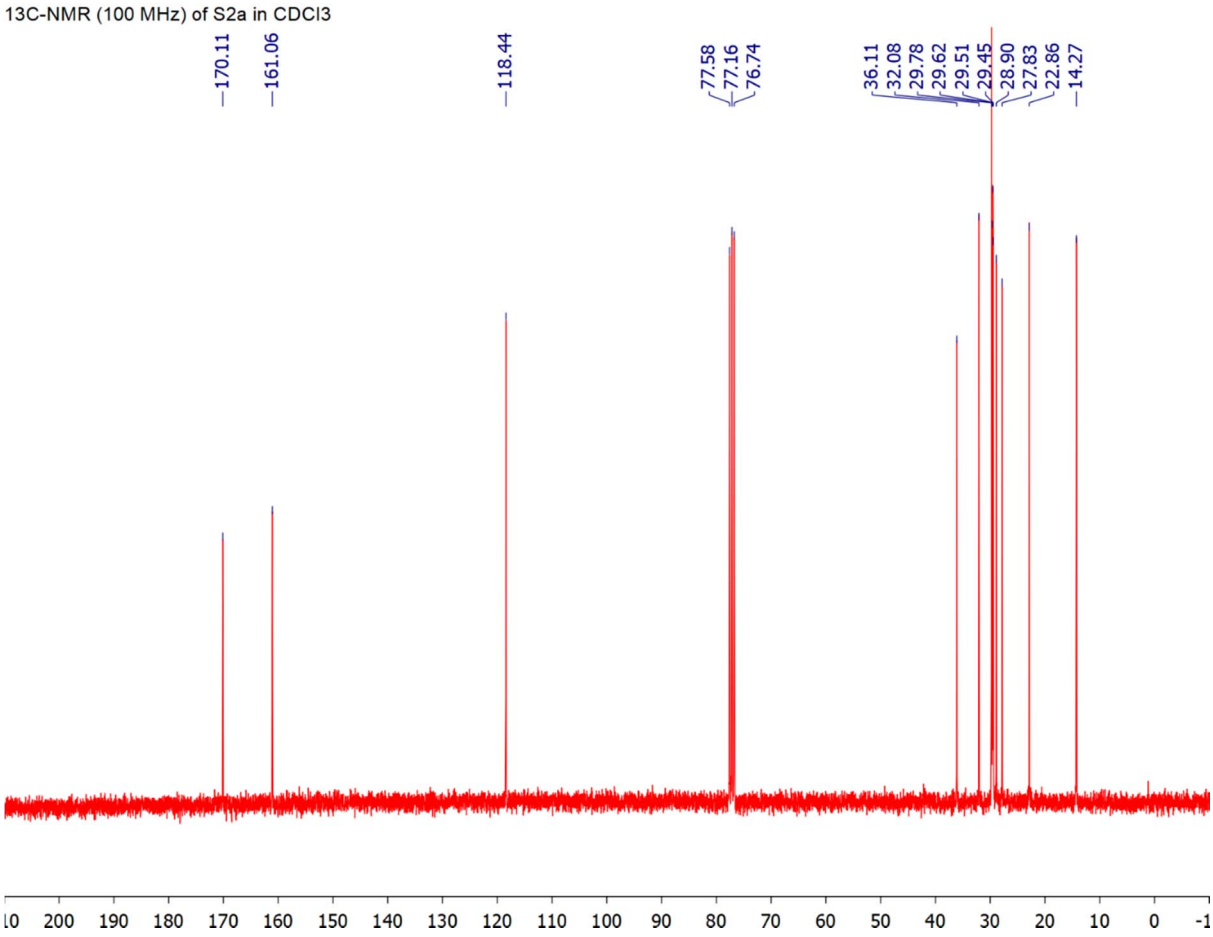
S2. NMR Spectra

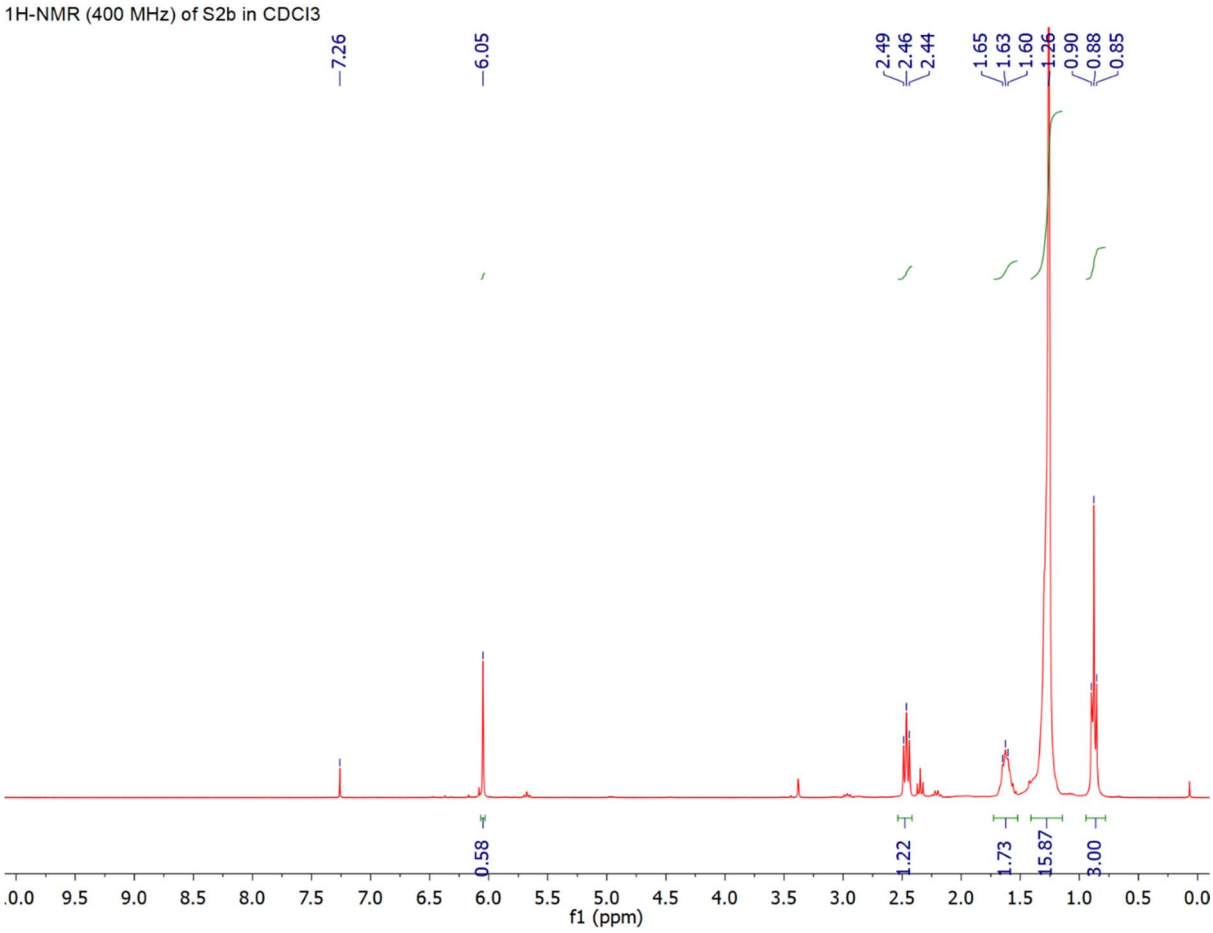
1H-NMR (400 MHz) of S1 in CDCl3

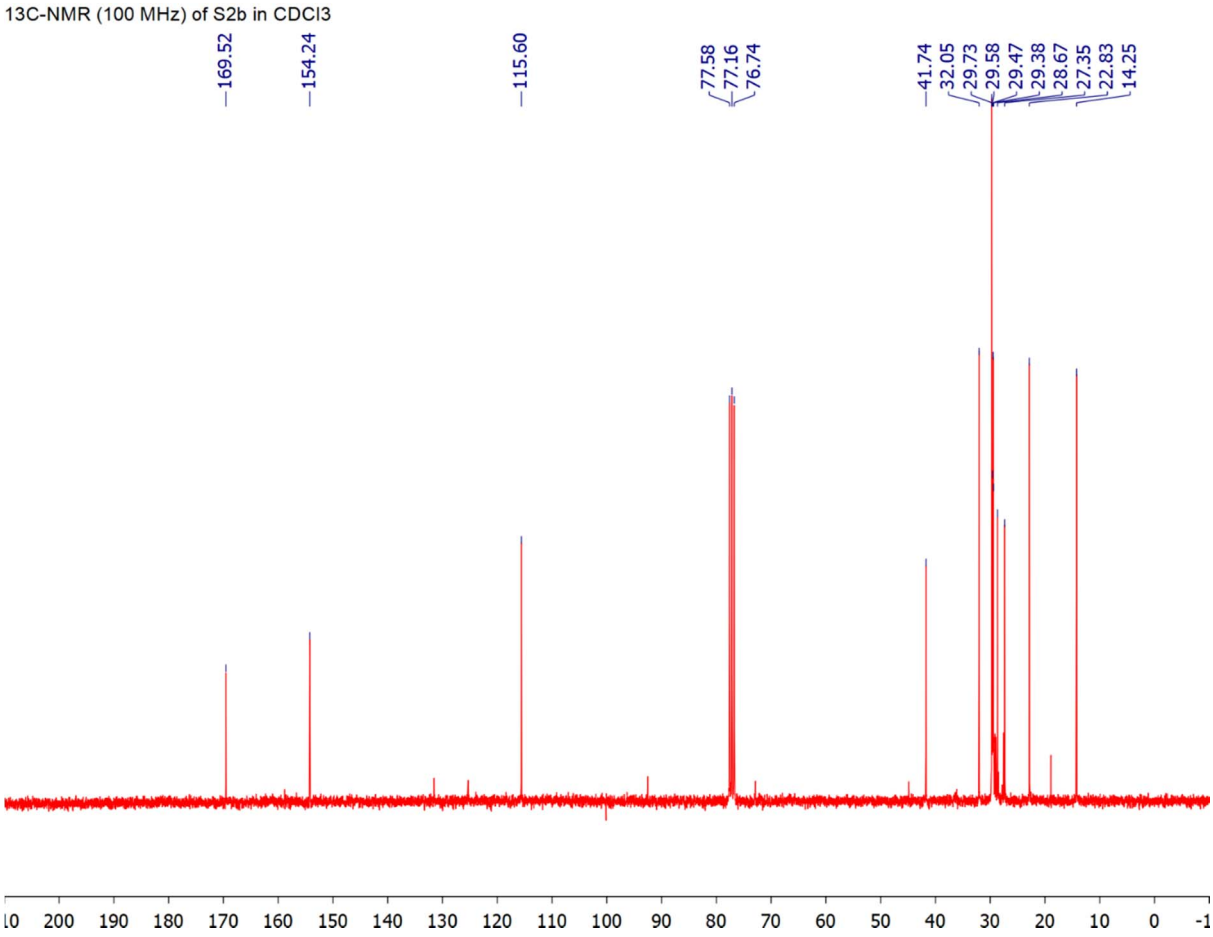


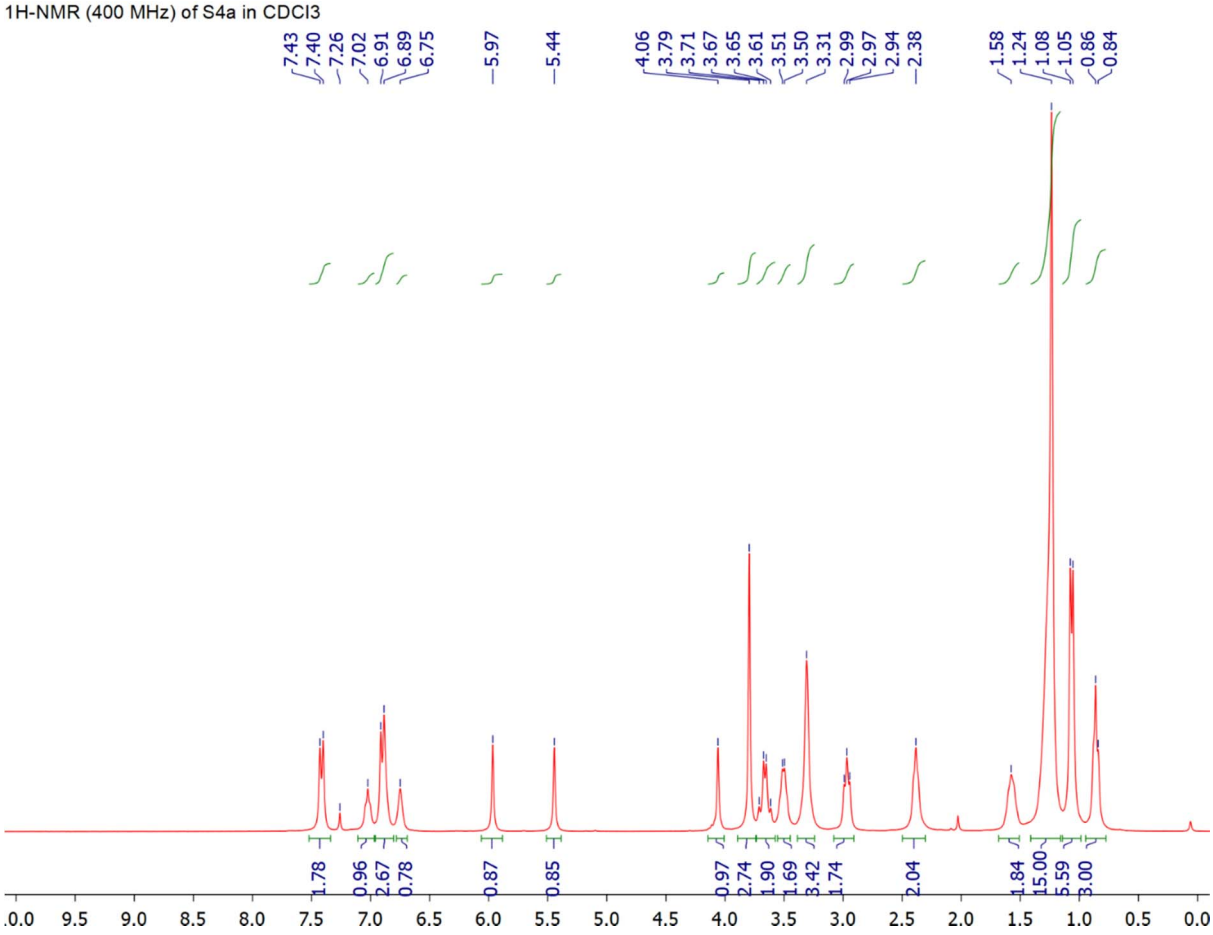

13C-NMR (100 MHz) spectra of S1 in CDCl3
158.83
92.92
77.48
77.16
76.84
72.77
32.03
29.71
29.53
29.13
28.94
27.51
22.81
19.01
18.88
18.74
14.19
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -1

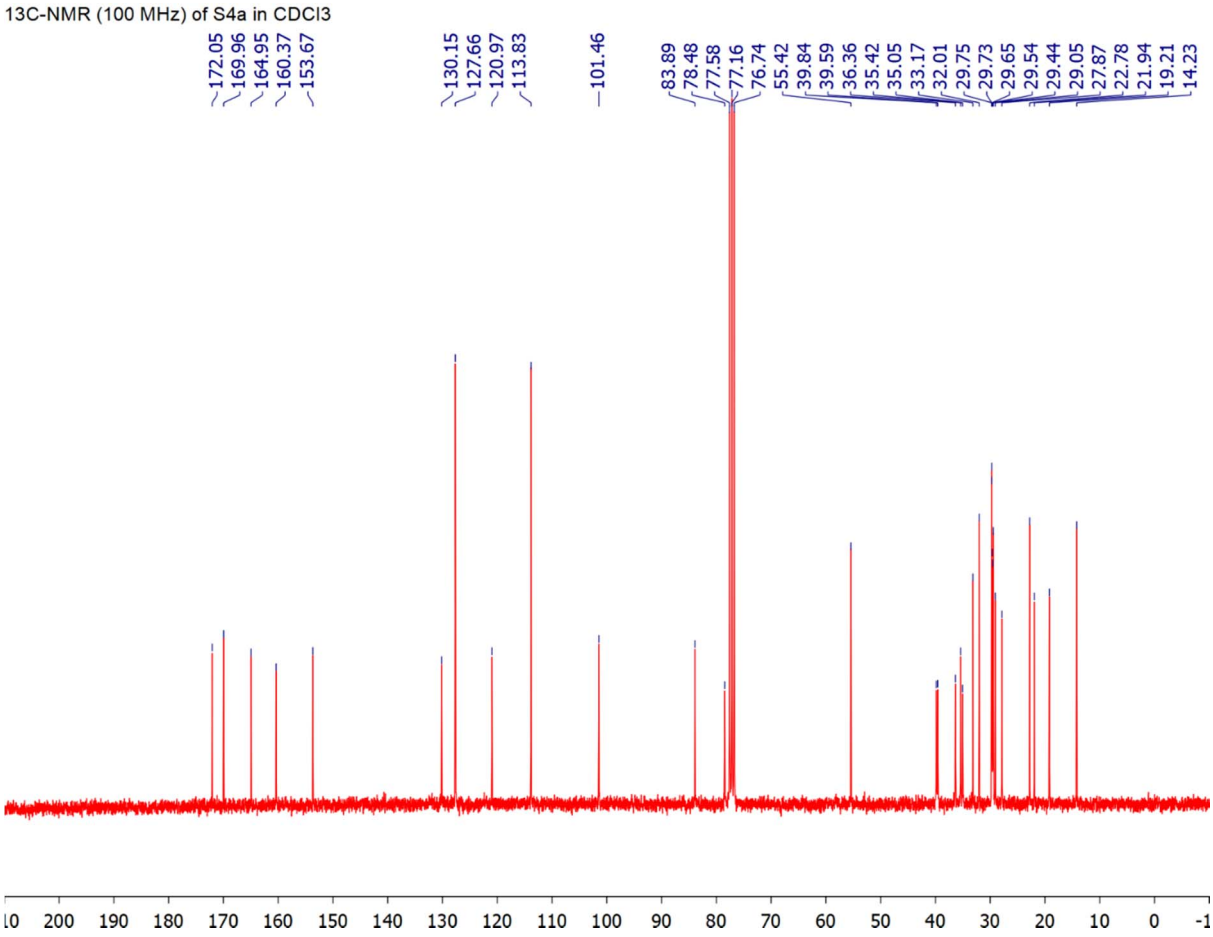

1H-NMR (400 MHz) of S2a in CDCl3
7.26
6.09
3.00
2.98
2.95
1.68
1.66
1.63
1.61
1.59
1.31
1.26
0.90
0.88
0.86
0.93
1.87
1.94
16.20
3.00
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0

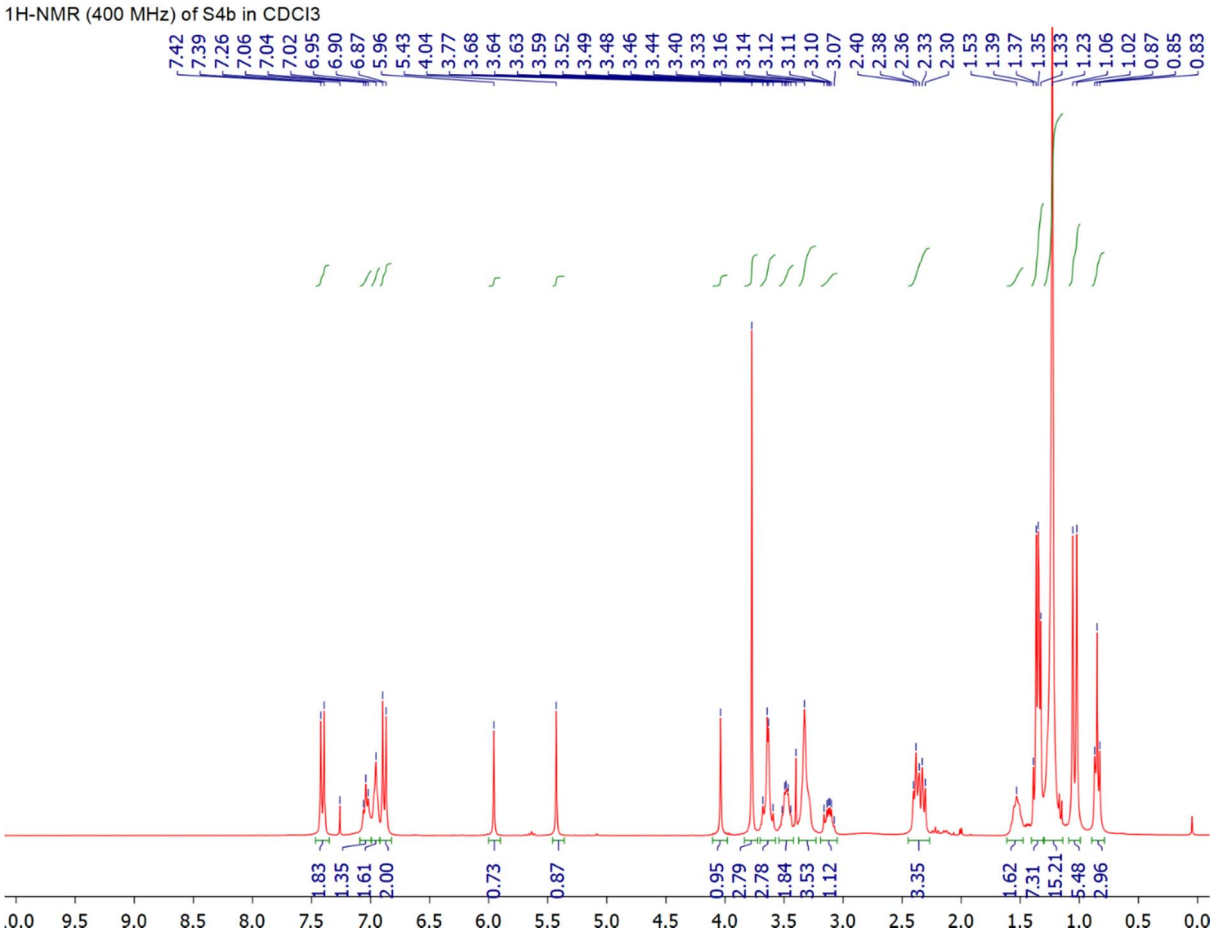

13C-NMR (100 MHz) of S2a in CDCl3
170.11
161.06
118.44
77.58
77.16
76.74
36.11
32.08
29.78
29.62
29.51
29.45
28.90
27.83
22.86
14.27
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -1

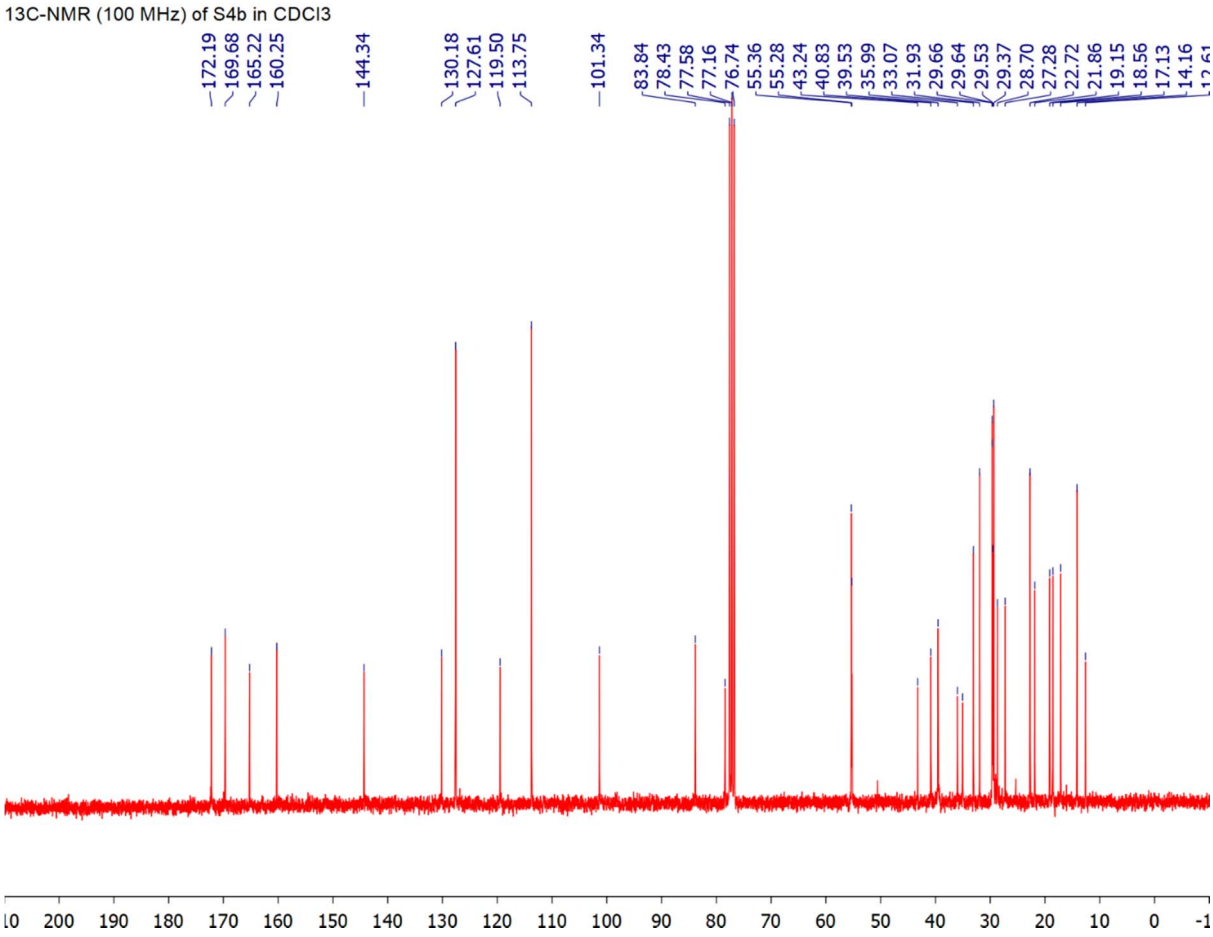

1H-NMR (400 MHz) of S2b in CDCl3
7.26
6.05
2.49
2.46
2.44
1.65
1.63
1.60
1.26
0.90
0.88
0.85
0.58
1.22
1.73
15.87
3.00
f1 (ppm)

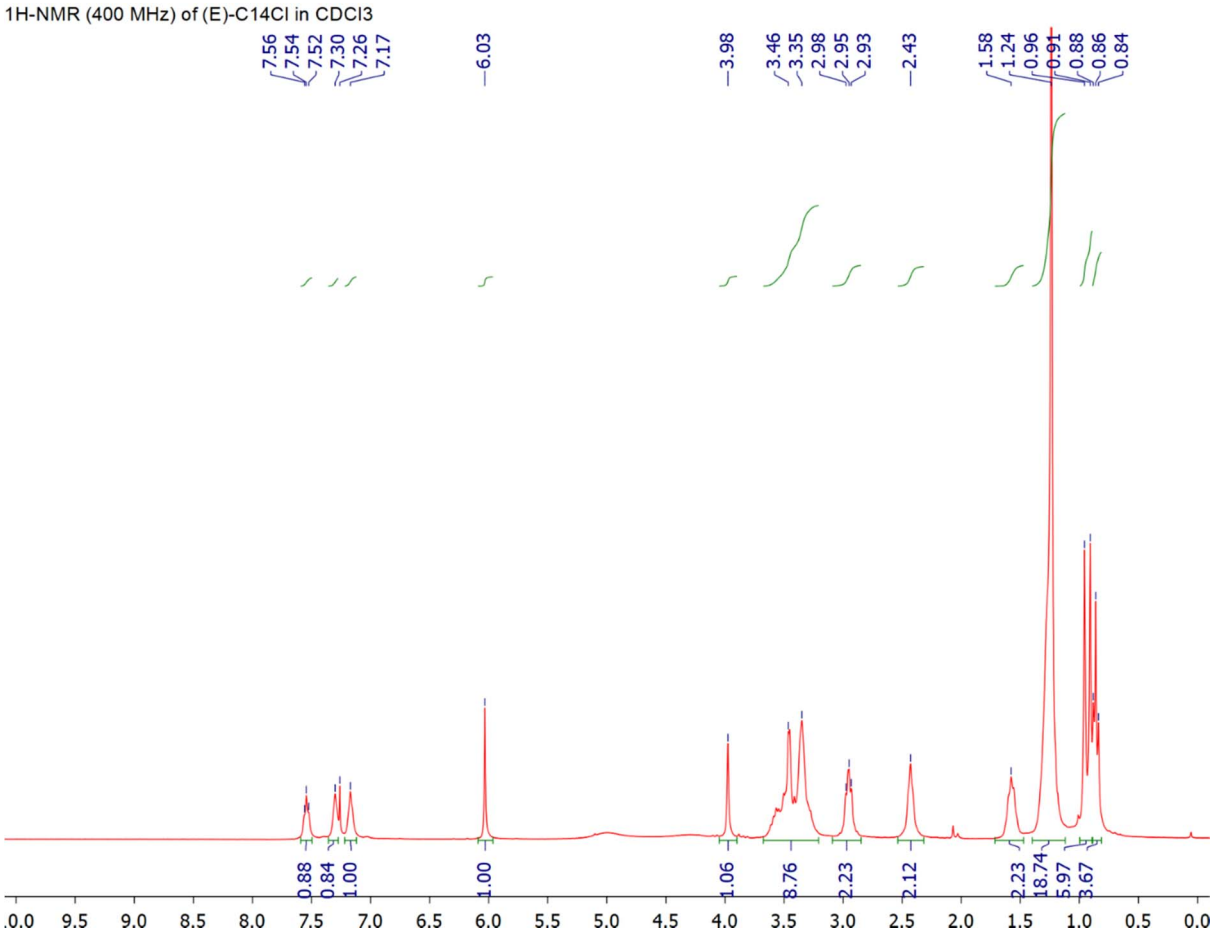

13C-NMR (100 MHz) of S2b in CDCl3
169.52
154.24
115.60
77.58
77.16
76.74
41.74
32.05
29.73
29.58
29.47
29.38
28.67
27.35
22.83
14.25
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

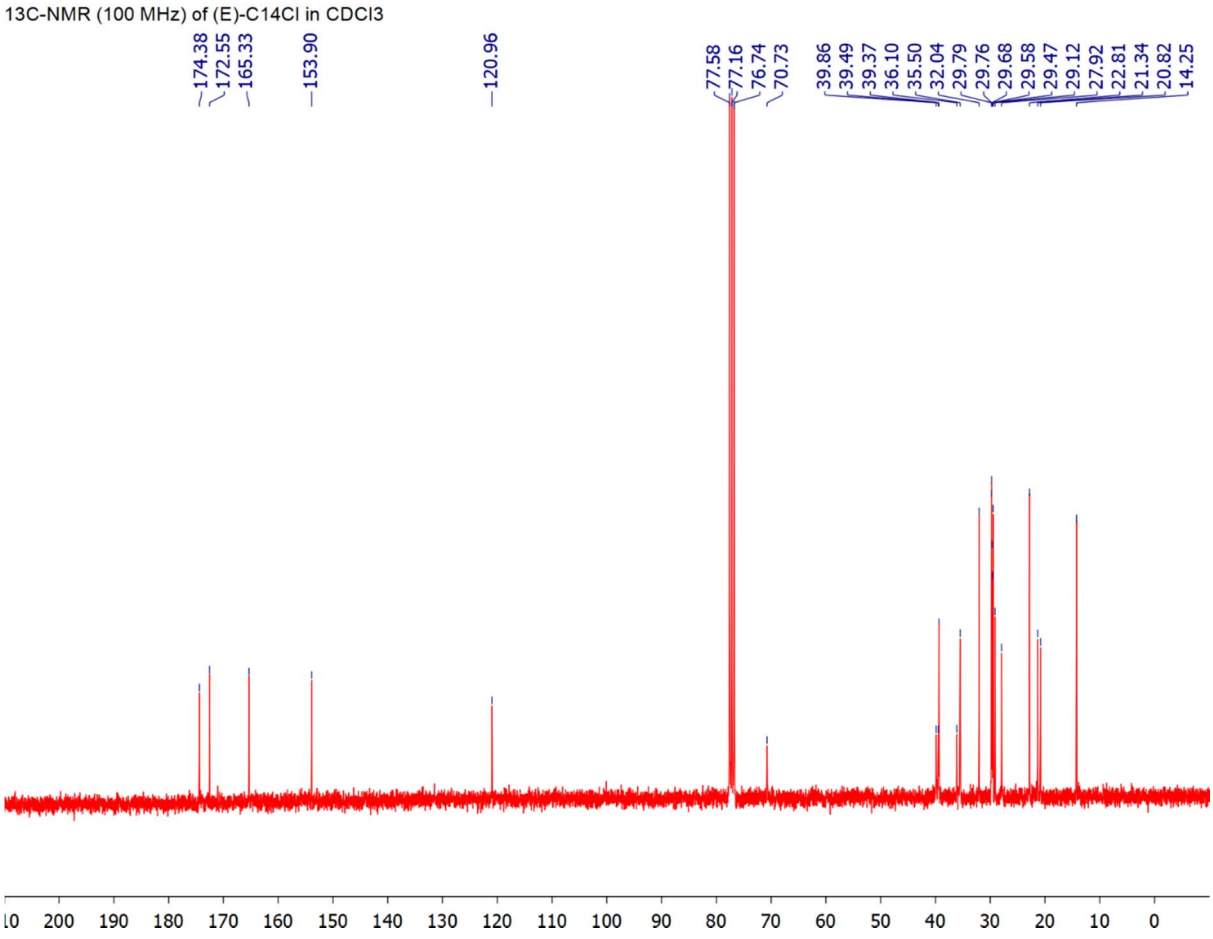

1H-NMR (400 MHz) of S4a in CDCl3
7.43
7.40
7.26
7.02
6.91
6.89
6.75
5.97
5.44
4.06
3.79
3.71
3.67
3.65
3.61
3.51
3.50
3.31
2.99
2.97
2.94
2.38
1.58
1.24
1.08
1.05
0.86
0.84
1.78
0.96
2.67
0.78
0.87
0.85
0.97
2.74
1.90
1.69
3.42
1.74
2.04
1.84
15.00
5.59
3.00
.0.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

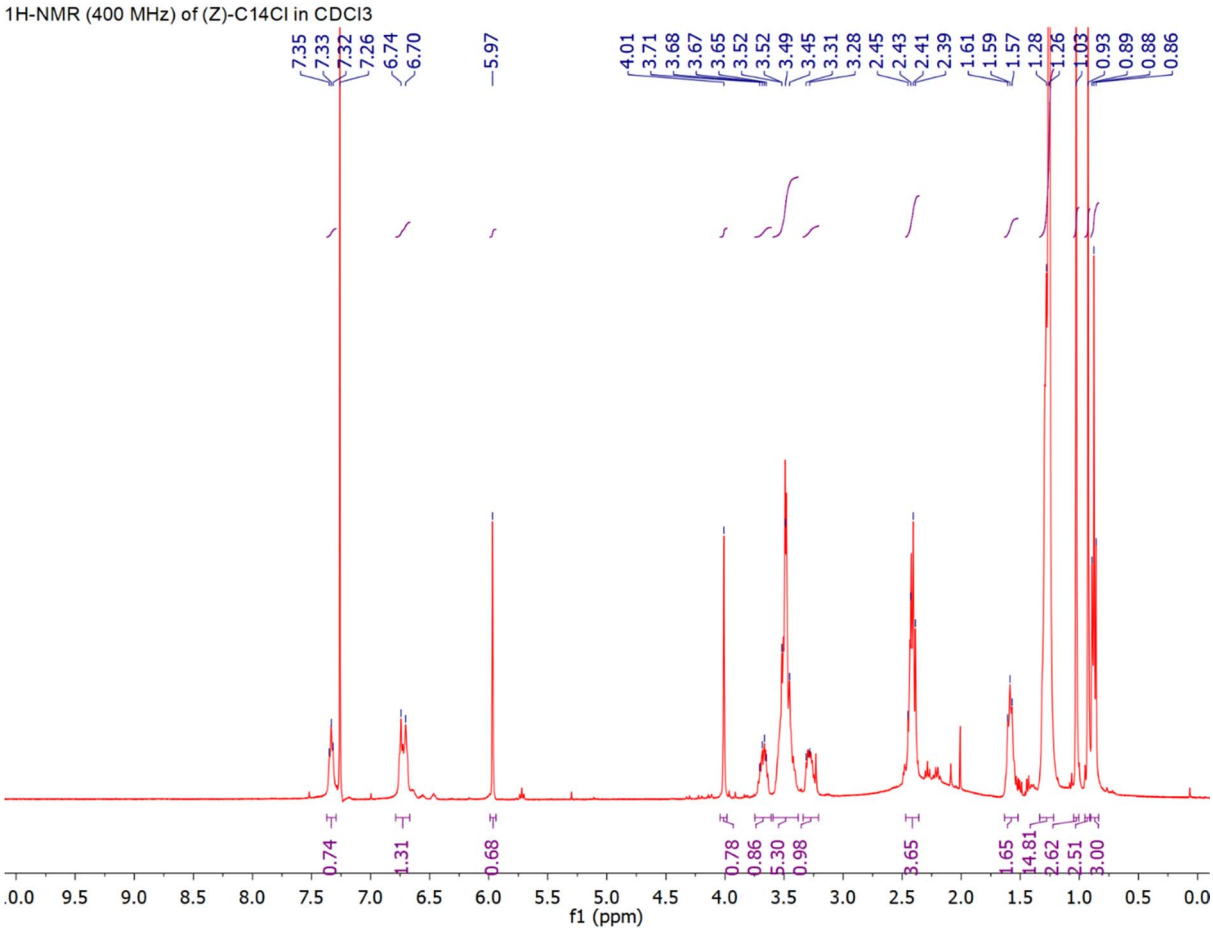

13C-NMR (100 MHz) of S4a in CDCl3
172.05
169.96
164.95
160.37
153.67
130.15
127.66
120.97
113.83
101.46
83.89
78.48
77.58
77.16
76.74
55.42
39.84
39.59
36.36
35.42
35.05
33.17
32.01
29.75
29.73
29.65
29.54
29.44
29.05
27.87
22.78
21.94
19.21
14.23
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -1

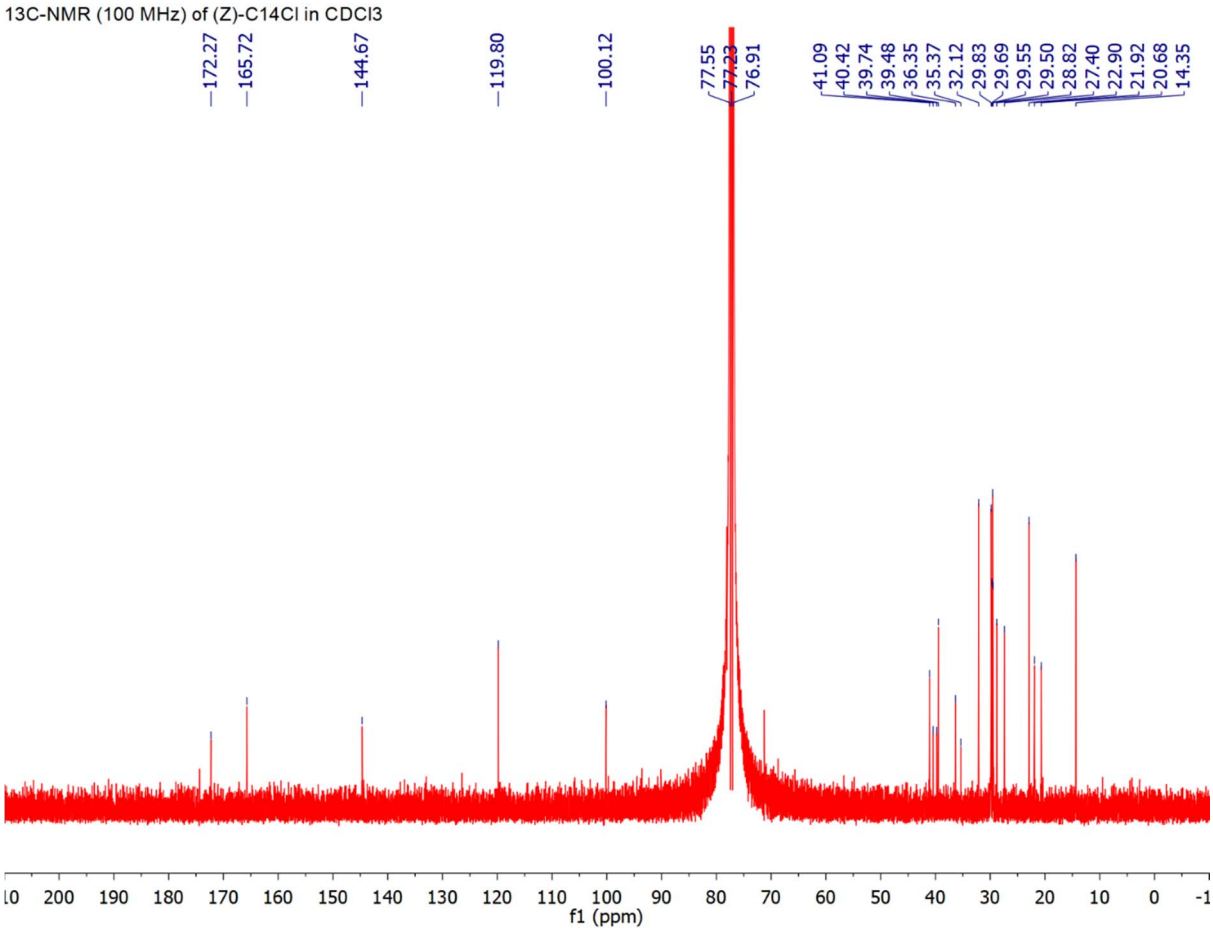

1H-NMR (400 MHz) of S4b in CDCl3
7.42
7.39
7.26
7.06
7.04
7.02
6.95
6.90
6.87
5.96
5.43
4.04
3.77
3.68
3.64
3.63
3.59
3.52
3.49
3.48
3.46
3.44
3.40
3.33
3.16
3.14
3.12
3.11
3.10
3.07
2.40
2.38
2.36
2.33
2.30
1.53
1.39
1.37
1.35
1.33
1.23
1.06
1.02
0.87
0.85
0.83
1.83
1.35
1.61
2.00
0.73
0.87
0.95
2.79
2.78
1.84
3.53
1.12
3.35
1.62
7.31
15.21
6.48
2.96
.0.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0


13C-NMR (100 MHz) of S4b in CDCl3
172.19
169.68
165.22
160.25
144.34
130.18
127.61
119.50
113.75
101.34
83.84
78.43
77.58
77.16
76.74
55.36
55.28
43.24
40.83
39.53
35.99
33.07
31.93
29.66
29.64
29.53
29.37
28.70
27.28
22.72
21.86
19.15
18.56
17.13
14.16
10 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -1


1H-NMR (400 MHz) of (E)-C14Cl in CDCl3
7.56
7.54
7.52
7.30
7.26
7.17
6.03
3.98
3.46
3.35
2.98
2.95
2.93
2.43
1.58
1.24
0.96
0.91
0.88
0.86
0.84
0.88
0.84
1.00
1.00
1.06
8.76
2.23
2.12
2.23
18.74
5.97
3.67
.0.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0


13C-NMR (100 MHz) of (E)-C14Cl in CDCl3
174.38
172.55
165.33
153.90
120.96
77.58
77.16
76.74
70.73
39.86
39.49
39.37
36.10
35.50
32.04
29.79
29.76
29.68
29.58
29.47
29.12
27.92
22.81
21.34
20.82
14.25
10 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0


1H-NMR (400 MHz) of (Z)-C14Cl in CDCl3
7.35
7.33
7.32
7.26
6.74
6.70
5.97
4.01
3.71
3.68
3.67
3.65
3.52
3.52
3.49
3.45
3.31
3.28
2.45
2.43
2.41
2.39
1.61
1.59
1.57
1.28
1.26
1.03
0.93
0.89
0.88
0.86
0.74
1.31
0.68
0.78
0.86
5.30
0.98
3.65
1.65
14.81
2.62
2.51
3.00
f1 (ppm)


13C-NMR (100 MHz) of (Z)-C14Cl in CDCl3
172.27
165.72
144.67
119.80
100.12
77.55
77.23
76.91
41.09
40.42
39.74
39.48
36.35
35.37
32.12
29.83
29.69
29.55
29.50
28.82
27.40
22.90
21.92
20.68
14.35
f1 (ppm)

Supplementary references

Gottlieb, H. E., Kotlyar, V. & Nudelman, A. J. (1997). *J. Org. Chem.* **62**, 7512-7515.

Meier, J. L. & Burkart, M. D. (2009). *Methods in Enzymology* **458**, 219-254.