Reconstituting poly(glycerol phosphate) wall teichoic acid biosynthesis *in vitro* using authentic substrates

Robert T. Gale,[a] Edward W. Sewell,[a] Teresa A. Garrett[b] and Eric D. Brown[*a]

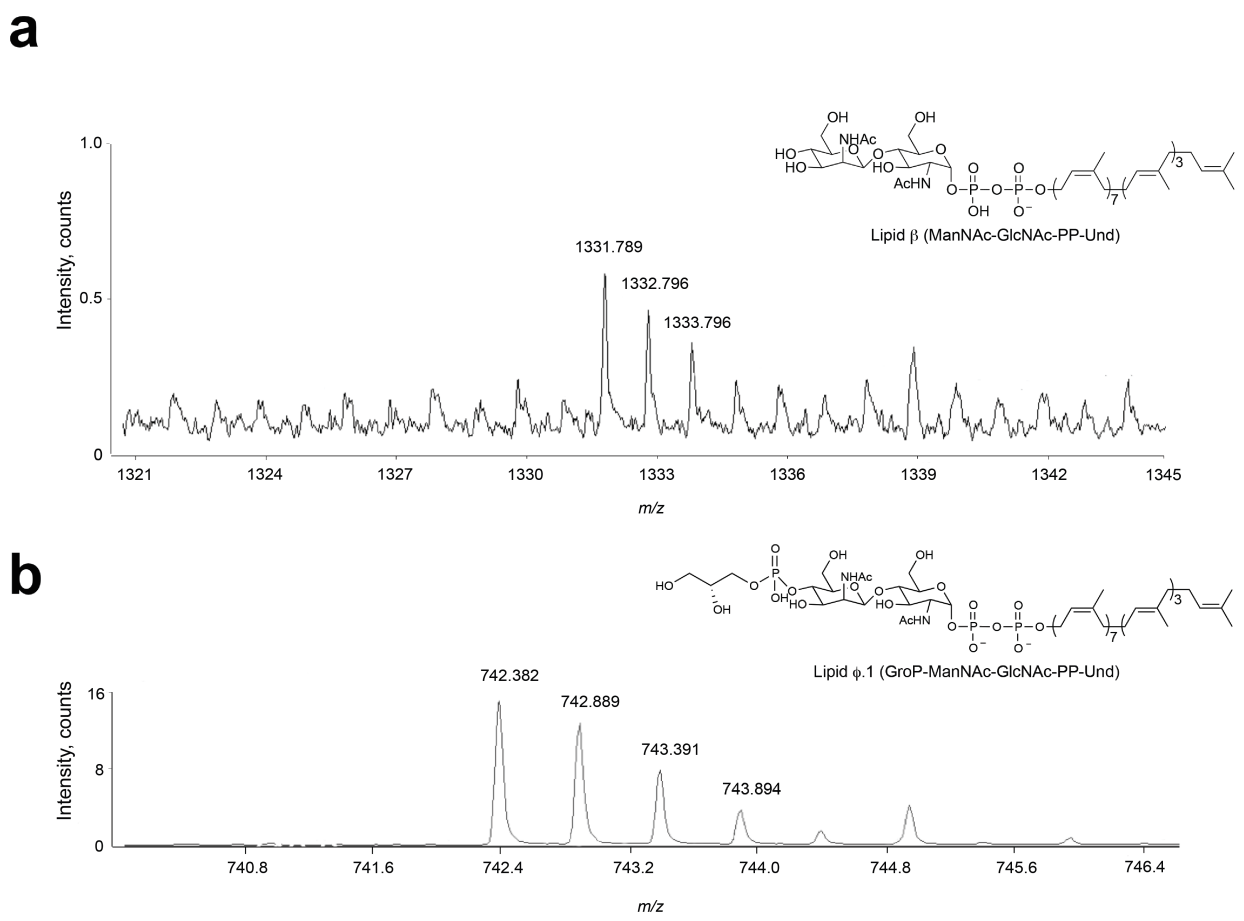
[a] Michael G. DeGroote Institute of Infectious Disease Research and Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario L8N 3Z5, Canada.

[b] Department of Chemistry, Vassar College, Poughkeepsie, New York 12604, United States.

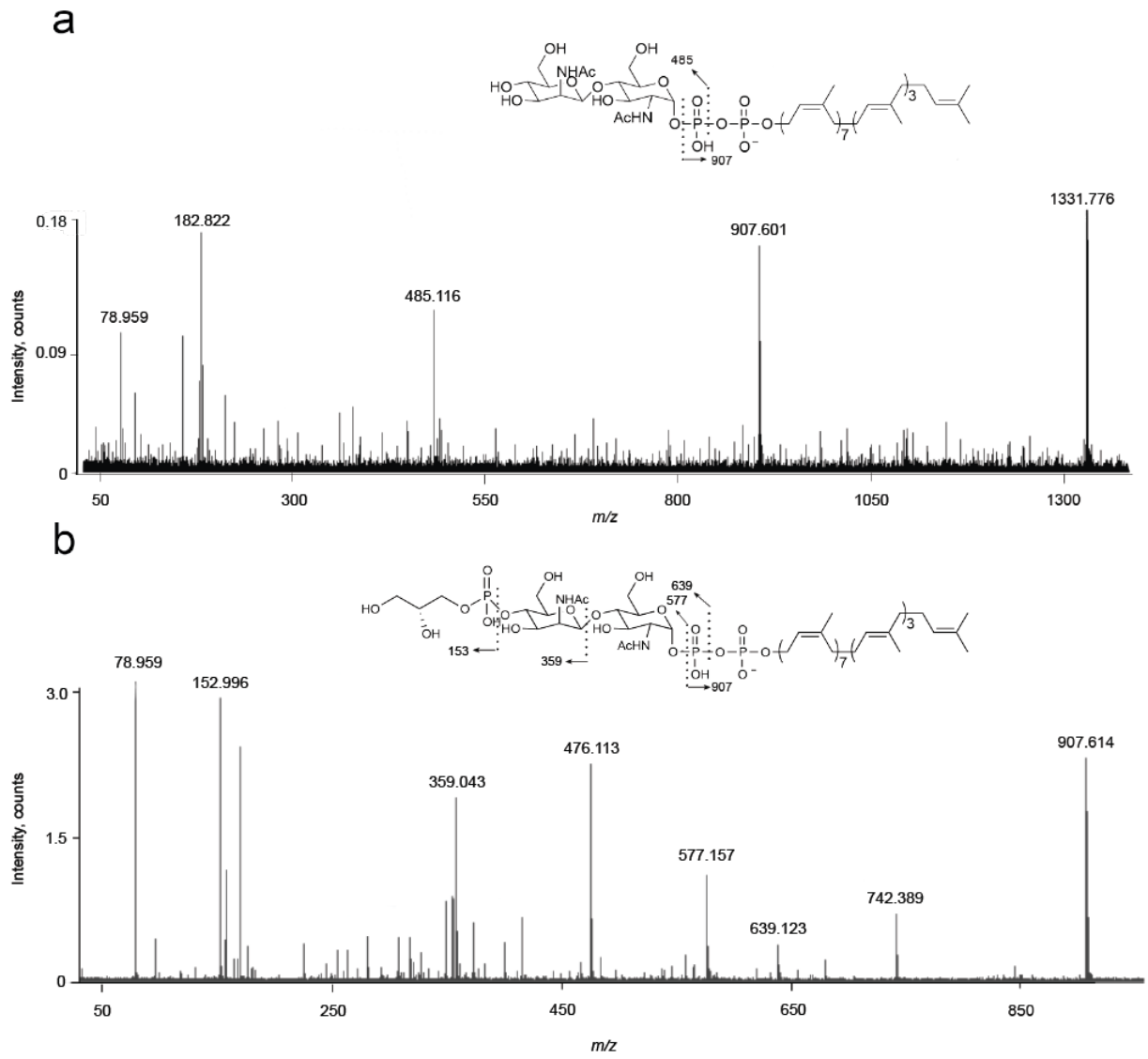
* To whom correspondence should be addressed.

Supplementary Material

1.0 Supplementary Figures

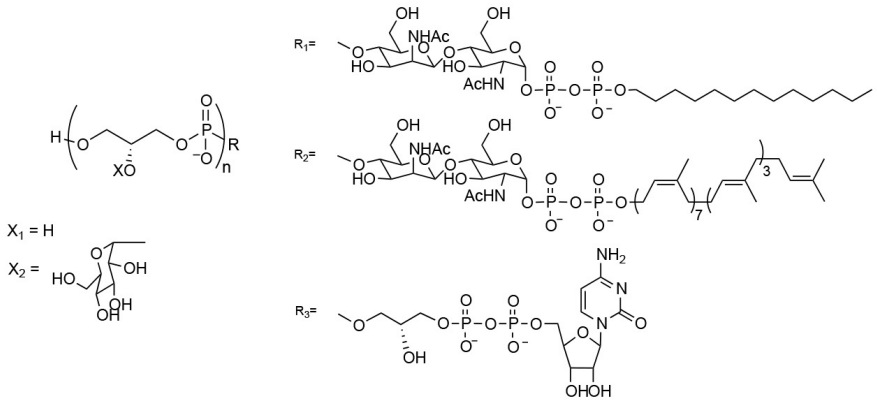


Supplementary Fig 1. HRMS of WTA intermediates produced *in vitro*. Negative-ion electrospray ionization mass spectrum (ESI-MS) of lipid extracted products from: a) Non-radioactive reaction involving TagA, semisynthesized Lipid α and UDP-ManNAc; b) Non-radioactive reaction involving TagA, TagB, semisynthesized Lipid α, UDP-ManNAc and CDP-glycerol. The exact *m/z* of the $[M\text{-}H^{+}]^{-}$ for Lipid β is 1331.783. The exact *m/z* of the $[M\text{-}2H^{+}]^{2-}$ for Lipid ϕ.1 is 742.389.

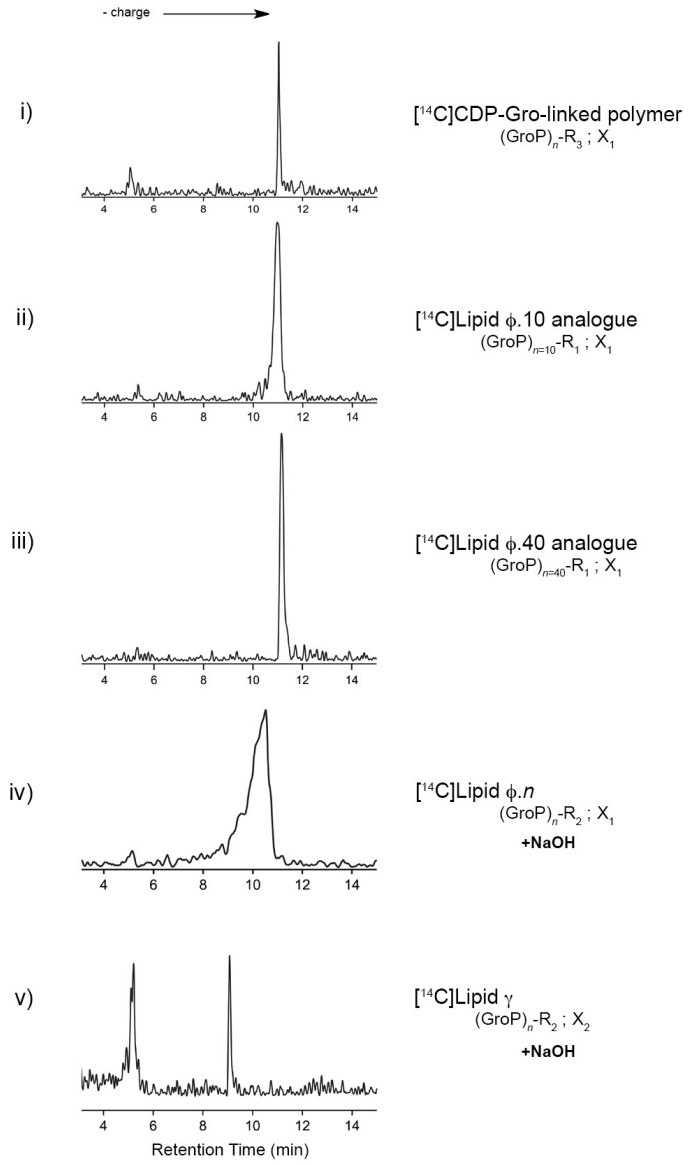


Supplementary Fig 2. Negative-ion collision-induced dissociation mass spectra (MS/MS) of WTA intermediates produced *in vitro*. MS/MS spectra were obtained for the ions corresponding to Lipid β (*m/z* 1331.783) and Lipid ϕ.1 (*m/z* 742.389). The inset shows the predicted product ions for each.

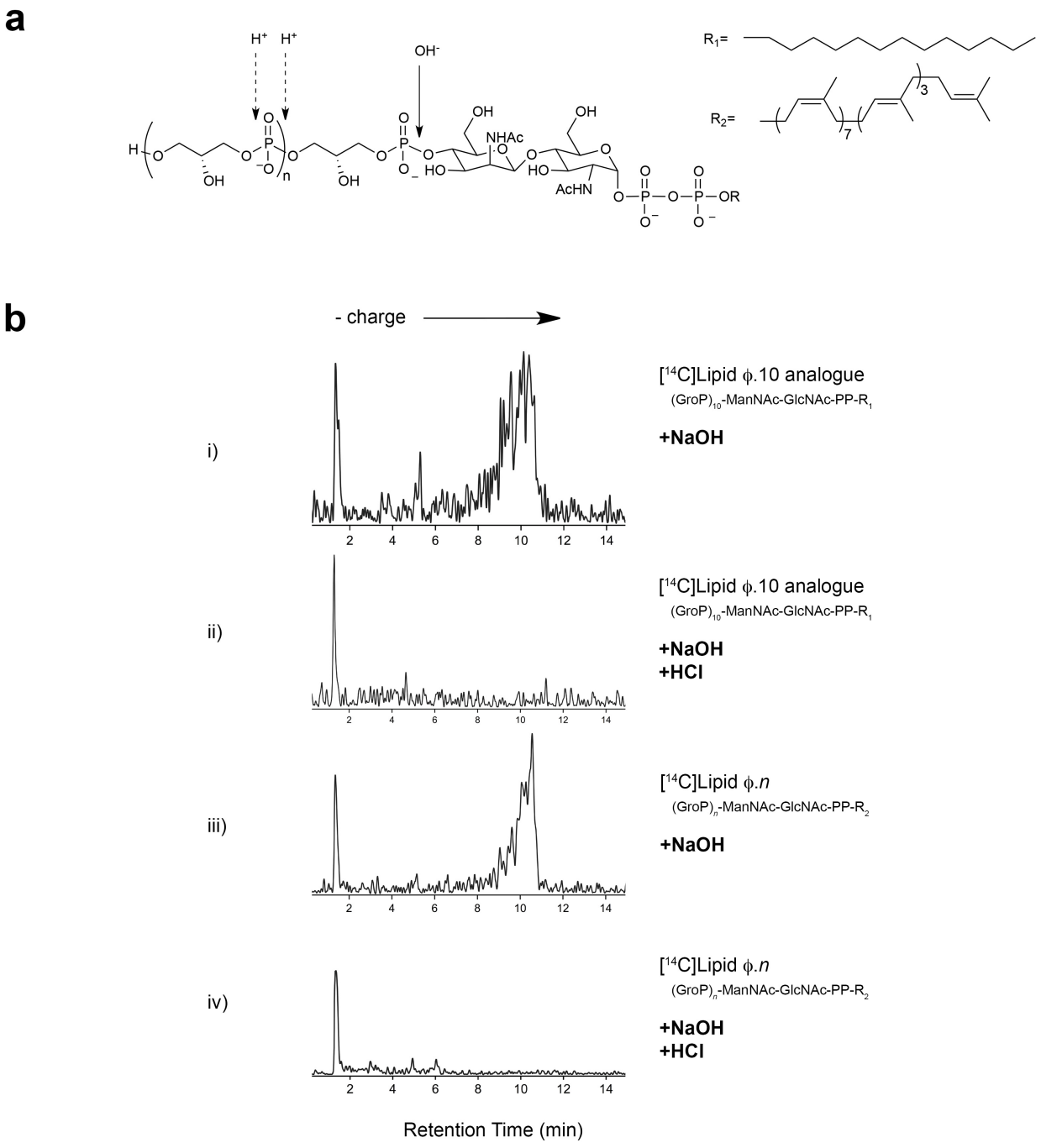
a


R1=
R2=
R3=
X1 = H
X2 =

b


- charge
i)
[14C]CDP-Gro-linked polymer
(GroP)n-R3 ; X1
ii)
[14C]Lipid ϕ.10 analogue
(GroP)n=10-R1 ; X1
iii)
[14C]Lipid ϕ.40 analogue
(GroP)n=40-R1 ; X1
iv)
[14C]Lipid ϕ.n
(GroP)n-R2 ; X1
+NaOH
v)
[14C]Lipid γ
(GroP)n-R2 ; X2
+NaOH
4 6 8 10 12 14
Retention Time (min)

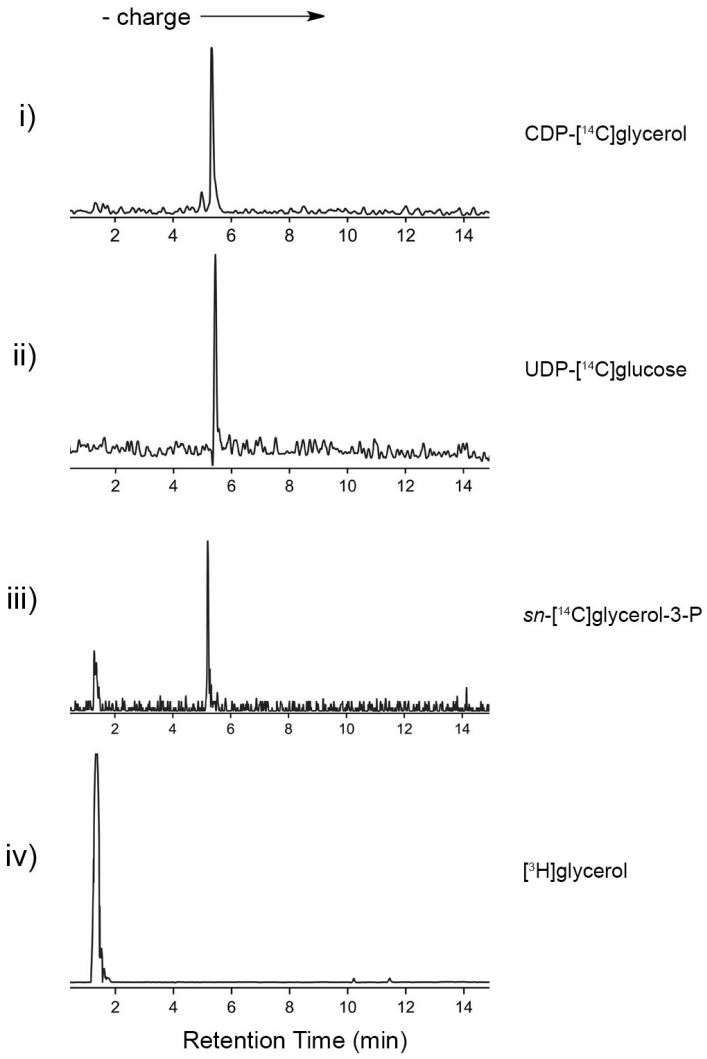
Supplementary Fig 3. TagF and TagE reaction products are polyanionic. (a) Depiction of poly(glycerol phosphate) polymers synthesized for anion exchange HPLC analysis. Polymers were built on tridecyl-linked substrates (R_1) bearing 10 and 40 glycerol phosphate repeats (n=10,40); native undecaprenyl-linked substrates (R_2) through enzymatic elaboration of semisynthetic Lipid α (**5**); and CDP-Gro (R_3). In addition, some polymers were chemically tailored with α-linked glucose (X_2). **(b)** Anion exchange HPLC profile of chemically distinct radioactive WTA polymers. Radioactivity was incorporated into polymers during synthesis from either CDP-[^{14}C]Gro (profiles i-iv) or UDP-[^{14}C]glucose (profile v) precursors. Polymer nomenclature and composition of reaction products following synthesis are indicated to the right of the frame. '+ NaOH' indicates reaction products further subjected to mild alkaline conditions (0.5M NaOH, 37°C, 25min) prior to anion exchange HPLC analysis. Elution profiles for the following are shown: i) CDP-Gro-linked polymer; ii) Lipid ϕ.10 analogue; iii) Lipid ϕ.40 analogue; iv) mild alkali-treated TagF reaction mixture after 4 hours of incubation with enzyme and Lipid ϕ.1; and v) mild alkali-treated TagE reaction mixture after 4 hours of incubation with enzyme and Lipid ϕ.*n*.



Supplementary Fig 4. The TagF reaction product shows typical WTA lability patterns.

(a) Chemical structure of WTAs prepared for anion exchange HPLC analysis. Poly(glycerol phosphate) polymers were built on tridecyl (R_1) and undecaprenyl (R_2) lipid chains. Relevant

linkages that are labile under mild acid (H^+) and mild alkali (OH^-) conditions during *in vitro* assays are shown with dashed and solid arrows respectively. **(b)** Anion exchange HPLC profile of hydrolyzed WTA polymers. Radioactivity was incorporated into synthesized polymers from a CDP-[^{14}C]Gro precursor. Polymer nomenclature and composition prior to acid/base treatment are indicated to the right of the frame. In addition, treatment of polymers to mild alkali (0.5 M NaOH, 37°C, 25 min) and/or mild acid (1N HCl, 100°C, 3 hours) conditions are indicated with '+NaOH' and '+HCl' labels respectively. Elution profiles for the following are shown: i) Lipid ϕ.10 analogue treated with NaOH; ii) Lipid ϕ.10 analogue treated with NaOH then HCl; iii) TagF reaction mixture after 4 hours of incubation with enzyme and Lipid ϕ.1 treated with NaOH; iv) same as iii) with subsequent HCl treatment.

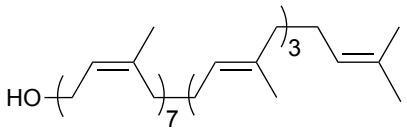


Supplementary Fig 5. Radioactive standards for anion exchange HPLC analysis. Standards subject to anion exchange HPLC analysis are indicated to the right of the frame. Elution profiles for the following compounds are shown: i) CDP-[U-^{14}C]glycerol; ii) UDP-[^{14}C]glucose; iii) *sn*-[U-^{14}C]glycerol-3-phosphate; and iv) [2-^{3}H]glycerol.

2.0 Preparation and Characterization of Lipid α (5) and Intermediates (1-4)

General. Chemicals and solvents were purchased from Sigma-Aldrich (Oakville, ON) or Fisher Scientific (Whitby, ON) unless otherwise stated. Flash column chromatography was performed using silica, C18, and SAX pre-packed columns from Teledyne Isco (Lincoln, NE) or Silicycle (Quebec City, QC) on a CombiFlash Rf system (Teledyne ISCO Inc.). Analytical thin layer chromatography (TLC) was carried out on silica gel 60 F254 aluminum-backed plates from EMD Chemicals (Gibbstown, NJ) or glass-backed C18 plates from Silicycle. TLC plates were visualized by exposure to ultraviolet light and/or exposure to iodine vapor (I_2) or an acidic solution of *p*-anisaldehyde. ^{1}H, ^{13}C and ^{31}P NMR spectra were obtained using a Bruker AVIII 700 MHz spectrometer. Spectra are reported in parts per million on the δ scale and are referenced to internal methanol (Methanol-d_4: ^{1}H, δ = 3.31 and ^{13}C, δ = 49.0). ^{1}H data are reported as follows: chemical shift (δ, ppm) (multiplicity, coupling constant (Hz), integration), ^{13}C and ^{31}P data are reported as follows: chemical shift (δ, ppm) (multiplicity, coupling constant (Hz)). Multiplicity abbreviations are as follows: s = singlet, d = doublet, dd = doublet of doublets, t = triplet, m = multiplet, b = broad. High-resolution mass spectra and collision-induced dissociation mass spectra (MS/MS) were collected on an Agilent 6520 quadrupole time-of-flight (Q-TOF) mass spectrometer. All compounds were lyophilized and resuspended in $CHCl_3/CH_3OH$ (2:1) prior to analysis following previously described methods.[11]

Extraction and isolation of undecaprenol (3)

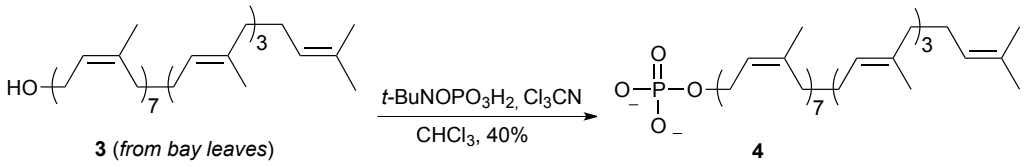


3 (*from bay leaves*)

Undecaprenol (**3**) was prepared loosely following published methodology for obtaining polyprenols from plant sources.[1-3] Our procedure was the following: Commercially available leaves of the *Laurus nobilis* tree (bay leaves) (42 g) were granulated and extracted in a Soxhlet apparatus with acetone/hexane (9:1) (500 mL) for 48 hours. The extract was set aside and the procedure was repeated. The combined extracts were evaporated to dryness, dissolved in 500 mL of hexane/EtOH/15% aqueous KOH (w/v) (3:15:2), and refluxed for 1 hour at 90°C. The mixture was cooled and extracted with H_2O/diethyl ether (1:1) (500 mL). The unsaponifiable extract was dried over Na_2SO_4, evaporated to dryness and purified by silica normal-phase chromatography (hexane/diethyl ether, 100:0 -> 95:5 -> 85:15). Polyisoprenol enriched fractions were identified by mass spectrometry, combined, evaporated to dryness and further purified by C18 reverse-phase chromatography (H_2O/Acetone, 5:95). Undecaprenol (**3**) enriched fractions were identified by mass spectrometry, combined, evaporated to dryness and used for phosphorylation reactions without further purification (158 mg, colourless oil).

HRMS (ESI) (*m/z*):	calc'd for $C_{55}H_{90}ClO$, $[M+Cl^-]^-$: 801.669 observed: 801.668
TLC (H_2O/Acetone, 5:95, Silica), R*f*:	0.29 (I_2)

Preparation of undecaprenyl phosphate (4)



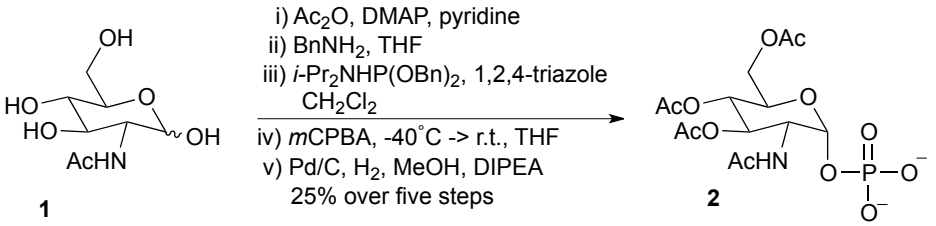
3 (*from bay leaves*) **4**

Undecaprenyl phosphate (**4**) was prepared from *Laurus nobilis* derived undecaprenol (**3**) (122 mg, 0.16 mmol) following previously described methodology for phosphorylation of polyprenols.[4] Our modifications to the procedure were the following: anion exchange chromatography was conducted using a pre-packed RediSep®Rf SAX column (5.7g) from Teledye ISCO. Compound elution was achieved using NH_4Ac (0 mM -> 30 mM -> 150 mM) in $CHCl_3$/MeOH (2:1). Lipid-linked products were extracted away from salts using a previously published CHCl3/MeOH/H_2O extraction technique.[5, 6] The lipid extract was concentrated under reduced pressure and further purified by C18 reverse-phase chromatography ($CHCl_3$/MeOH/H_20 in 0.1% aqueous NH_4OH, 97:3:0 -> 65:25:4) to afford **4** (54 mg, 40%) as a white powder.

^{1}H NMR (700 MHz, MeOH-d_4):	δ ppm 1.60 (s, 9H), 1.61 (s, 3H), 1.68 (s, 21H), 1.73 (s, 3H), 1.99-2.09 (m, 40H), 4.40 (t, J = 6.3 Hz, 2H), 5.10-5.16 (m, 11H), 5.42 (t, J = 6.3 Hz, 1H)
^{13}C NMR (176 MHz, MeOH-d_4):	δ ppm 16.17, 16.19, 17.81, 23.72, 23.76, 23.81, 23.86, 23.88, 25.94, 27.54, 27.58, 27.60, 27.61, 27.63, 27.67, 27.70, 27.75, 27.85, 32.92, 33.25, 33.26, 33.29, 33.33, 40.83, 40.88, 40.90, 62.69 (d, $J_{C\text{-}P}$ = 5.28 Hz), 124.13 (d, $J_{C\text{-}P}$ = 8.8 Hz), 125.46, 125.49, 125.50, 125.52, 125.92, 126.16, 126.18, 126.20, 132.03, 135.81, 135.82, 136.02, 136.19, 136.22, 136.24, 136.27, 136.29, 136.41, 139.67.
^{31}P NMR (283 MHz, MeOH-*d*4):	δ ppm 1.37
HRMS (ESI) (*m/z*):	calc'd for $C_{55}H_{90}O_4P$, $[M\text{-}H^+]^-$: 845.658
	observed: 845.661
TLC ($CHCl_3$/MeOH/H_2O, 65:25:5, Silica), R*f*:	0.38 (I_2)

*Our spectral characterization of **4** provides additional information to incomplete ^{1}H and ^{13}C spectra previously reported in the literature.[7]

Preparation of (2*R*,3*R*,4*R*,5*S*,6*R*)-3-acetamido-4,5-diacetoxy-6-(acetoxymethyl)tetrahydro-2*H*-pyran-2-yl phosphate (2)

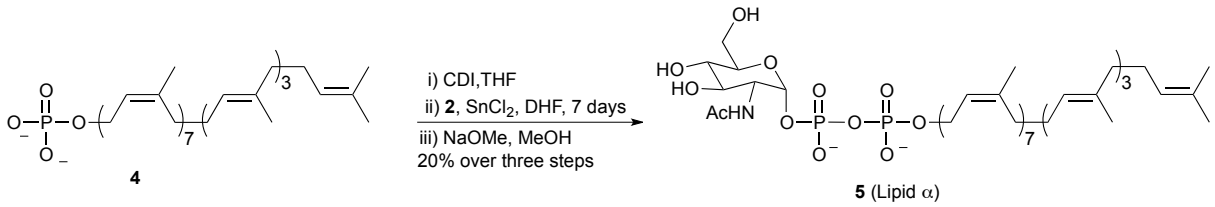


The α-phosphate (**2**) was synthesized from commercially available *N*-acetyl-D-glucosamine (**1**) according to known methods.[8,9,12,13] In addition to referenced methods, **2** was further purified by C18 reverse-phase chromatography (H_2O/MeCN in 0.1% aqueous NH_4OH, 95:5 -> 5:95) to afford **2** (25%) as a brown/yellow film.

^{1}H NMR (700 MHz, MeOH-d_4):	δ ppm 1.94 (s, 3H), 1.96 (s, 3H), 2.00 (s, 3H), 2.05 (s, 3H), 4.18 (d, J = 11.9 Hz, 1H), 4.24 (d, J = 10.5 Hz, 1H), 4.28 (d, J = 11.9 Hz, 1H), 4.33 (d, J = 9.8 Hz, 1H) 5.09 (t, J = 9.8 Hz, 1H), 5.32 (t, J = 9.8 Hz, 1H), 5.48 (bs, 1H)
^{13}C NMR (176 MHz, MeOH-d_4):	δ ppm 20.63, 20.69, 22.60, 53.48, 62.96, 69.46, 69.84, 73.00, 94.91 (d, $J_{C\text{-}P}$ = 5.63 Hz), 171.35, 172.01, 172.60, 173.61
^{31}P NMR (283 MHz, MeOH-*d*4):	δ ppm -0.80
HRMS (ESI) (*m/z*):	calc'd for $C_{14}H_{21}NO_{12}P$, $[M\text{-}H^+]^-$: 426.081
	observed: 426.080

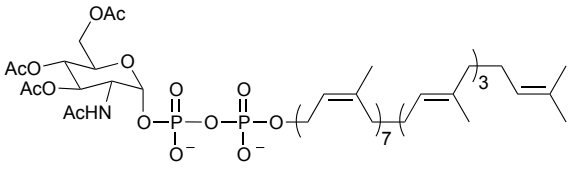
* Our spectral characterization of **2** provides additional information to ^{1}H spectra reported for the dibenzylated precursor of **2.**[9]

Preparation of GlcNAc-PP-undecaprenol (Lipid α) (5)



All steps of the synthesis were performed in oven-dried glassware under argon atmosphere. The ammonium salts of **2** (5 mg, 10.8 μmol , 1.0 equiv) and **4** (10 mg, 11.8 μmol, 1.1 equiv) were placed in separate vessels and co-evaporated three times with anhydrous toluene. The vessels were dried under high vacuum for 3 hours. The ammonium salt of **4** was dissolved in anhydrous THF (2.4 mL) and CDI (9.2 mg, 56 μmol, 5.5 equiv) was added to initiate activation. Reaction progress was monitored by MS and showed incomplete activation after 18 hours incubation at ambient temperature. Therefore, an additional amount of CDI (11.6 mg, 72 μmol, 6.1 equiv) was added. Reaction completion was observed after overnight incubation at ambient temperature. The reaction was quenched with anhydrous methanol (20 uL) and stirred for 30 min. The solvents were evaporated and the residue was dried under high vacuum for two hours. Activated **4** was dissolved in DHF (1.5 mL) and transferred to the vessel containing the α-phosphate **2,** after which anhydrous tin(II) chloride (4 mg, 22 μmol, 2 equiv) was added**.** The reaction mixture was stirred at ambient temperature for one week and quenched with ethylenediaminetetraacetic acid (EDTA) (10 mM, 2.2mL). Solvents were evaporated and the lipid-linked products were extracted using a previously published CHCl3/MeOH/H_2O extraction technique.[5, 6] The lipid extract was concentrated under reduced pressure and purified by silica normal-phase chromatography ($CHCl_3$/MeOH in 0.2% aqueous NH_4OH, 95:5 -> 50:50) to afford the peracetylated form of Lipid α (**5**) (3.2 mg, 22%) as a white powder.

Structure of peracetylated form:



Precursor to 5

^{1}H NMR (700 MHz, MeOH-d_4): δ ppm 1.60 (s, 9H), 1.61 (s, 3H), 1.65-1.70 (m, 21H), 1.74 (s, 3H), 1.95 (s, 3H), 1.97-1.99 (m, 6H), 1.99 (s, 3H), 1.99-2.17 (m, 40H) 4.20 (dd, J_a = 2.1 Hz, J_b = 12.6 Hz, 1H), 4.30 (d, J = 10.5 Hz, 1H), 4.33 (dd, J_a = 2.8 Hz, J_b = 12.6 Hz, 1H), 4.41 (d, J = 9.8 Hz, 1H), 4.53 (t, J = 6.3 Hz, 2H), 5.07-5.18 (m, 12H), 5.34 (t, J = 10.5 Hz, 1H), 5.47 (t, J = 6.3 Hz, 1H), 5.59 (dd, J_a = 3.5 Hz, J_b = 7 Hz, 1H)

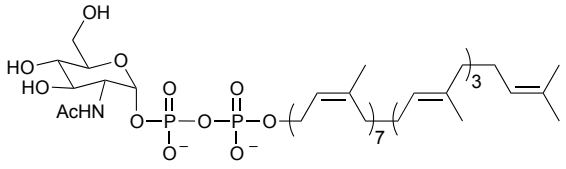
^{31}P NMR (283 MHz, MeOH-d4): δ ppm -10.17 (d, $J_{P\text{-}P}$ = 21.2 Hz), -13.21 (d, $J_{P\text{-}P}$ = 20.9 Hz)

HRMS (ESI) (*m/z*): calc'd for $C_{69}H_{109}NO_{15}P_2$, $[M\text{-}2H^+]^{2-}$: 626.864

observed: 626.864

TLC ($CHCl_3$/MeOH/H_2O, 65:25:5, Silica), R*f*: 0.36 (I_2)

The peracetylated form of **5** (3.2 mg, 2.6 μmol) was evaporated three times with anhydrous methanol and placed under high vacuum for two hours. The residue was then dissolved in anhydrous methanol (4 mL) and a 23% (v/v) $NaOCH_3$ in methanol solution (0.9 mL) was added dropwise to initiate deacetylation. The reaction mixture was stirred at ambient temperature for one hour, quenched with ammonium acetate (1.2 mL, 0.5 M in H_2O) and stirred again at ambient temperature for 30 min. The solvents were evaporated and the residue was dried under high vacuum for 3 hours. Lipid α (**5**) was isolated from reaction components following known lipid extraction methods[5,6] to afford **5** (2.7 mg, 92%) as a white powder. Lipid α (**5**) was thus obtained in 5% yield over the entire synthetic procedure when considering overall yields during the preparation of α-phosphate **2,** and over final coupling and deprotection steps.



5 (Lipid α)

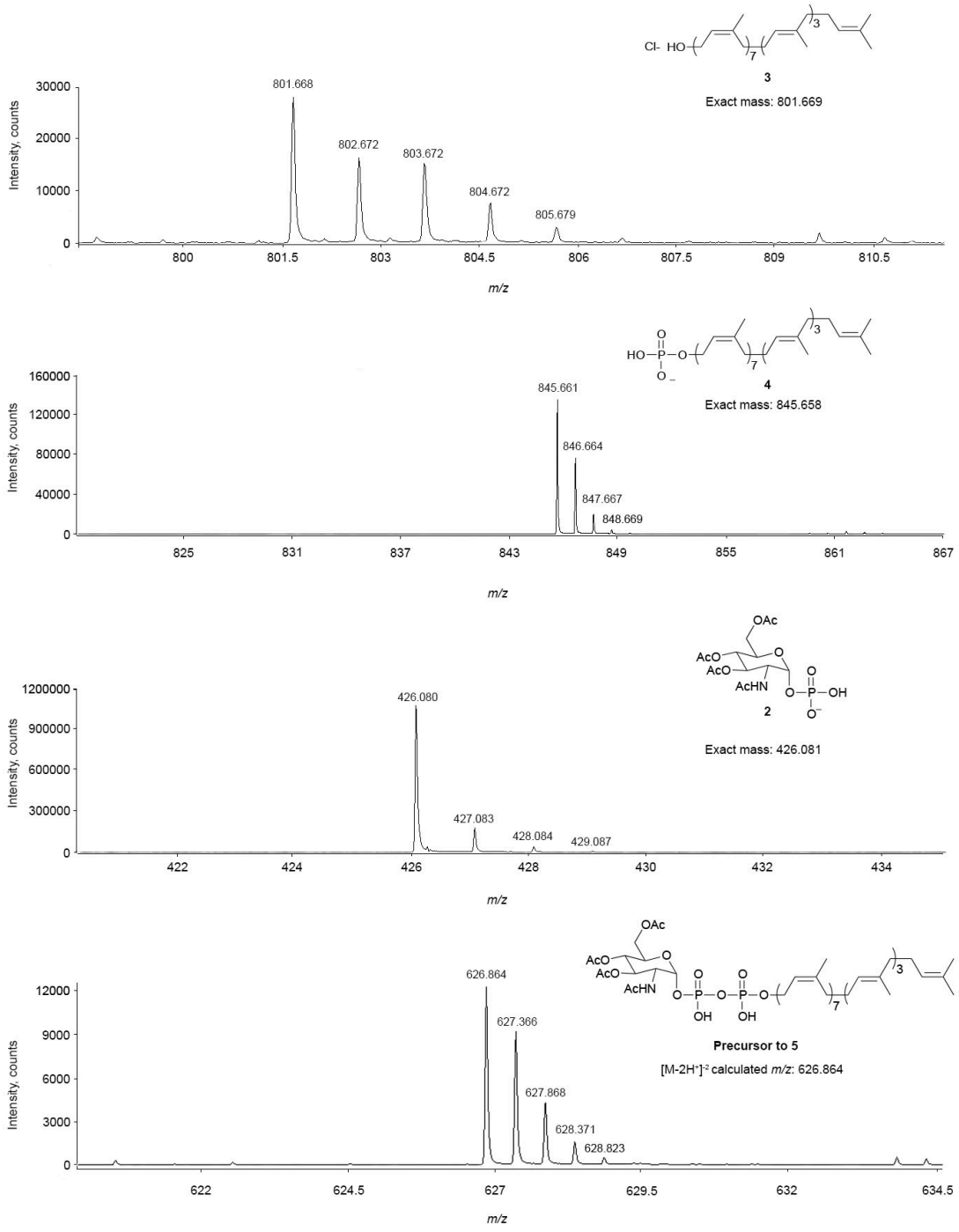
*****The final product proved too insoluble to obtain suitable spectral data, as has been reported for other undecaprenyl-linked glycosyl diphosphates.[8,10]

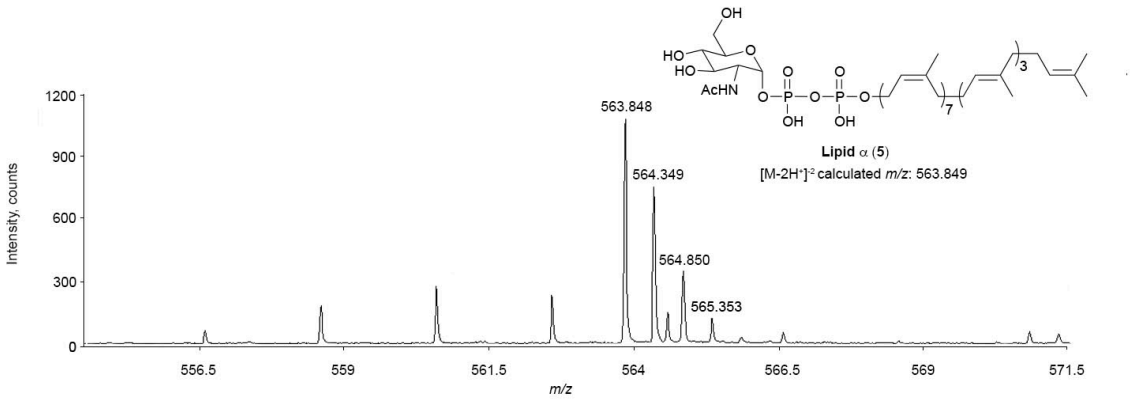
HRMS (ESI) (*m/z*): calc'd for $C_{63}H_{103}NO_{12}P_2$, $[M\text{-}2H^+]^{2-}$: 563.849

observed: 563.848

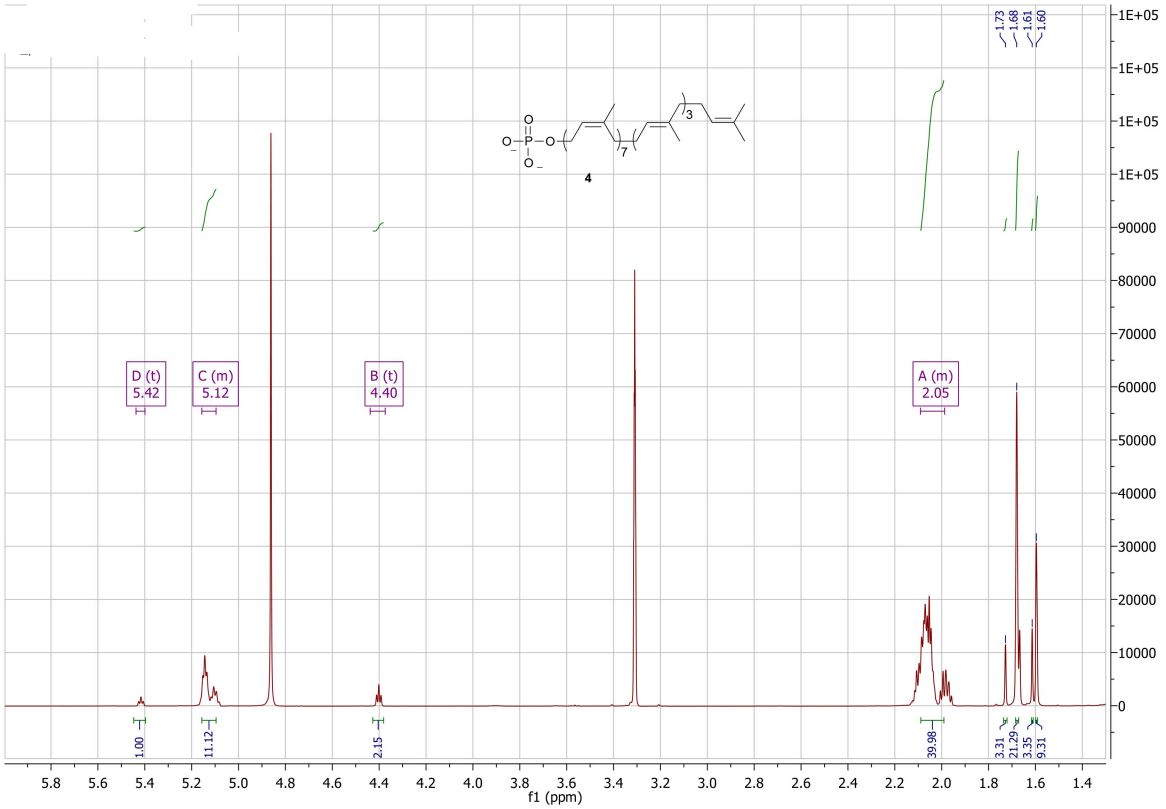
TLC ($CHCl_3$/MeOH/H_2O, 65:25:5, Silica), R*f*: 0.3 (I_2)

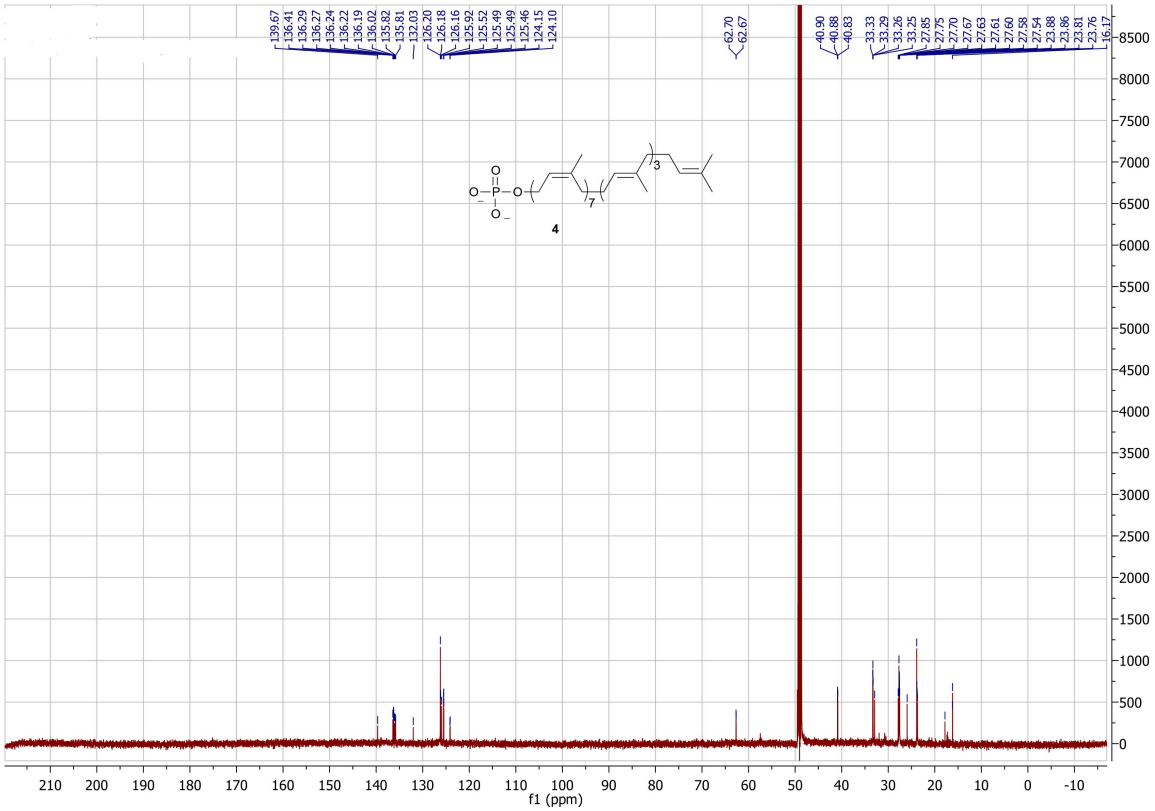
Mass spectra

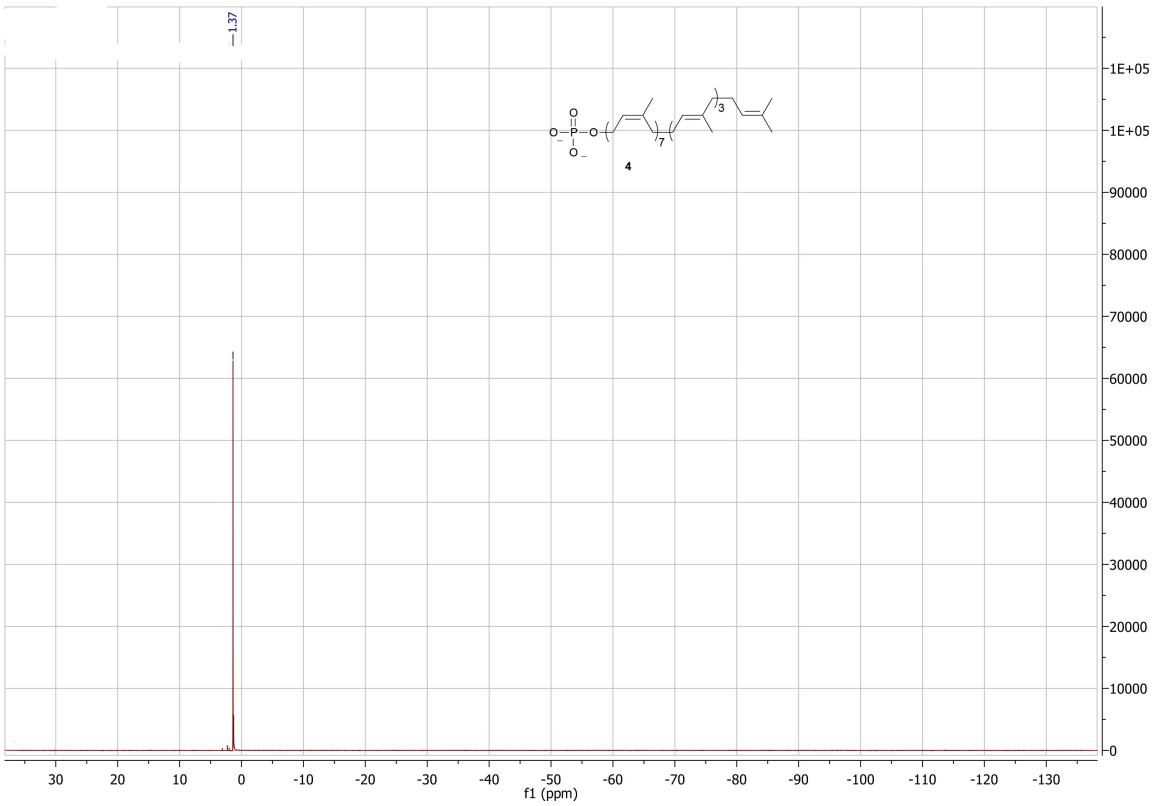


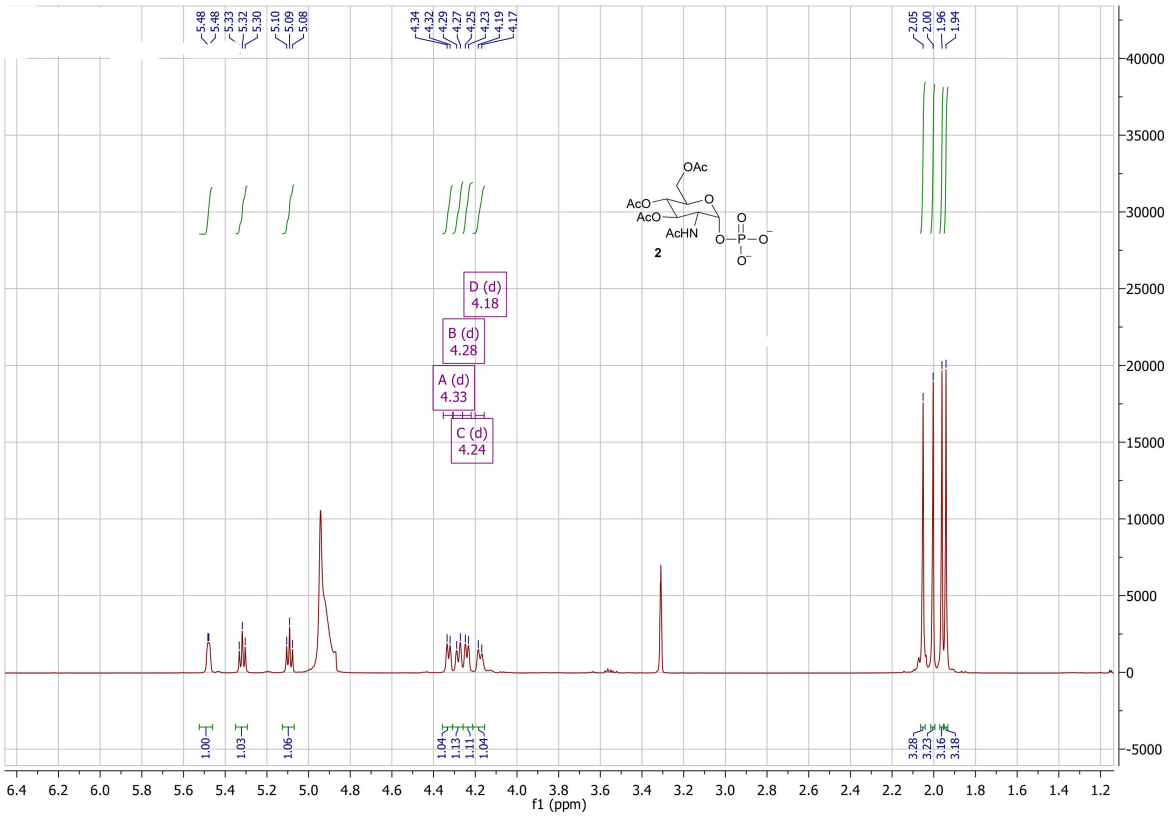

OH
HO
HO
AcHN
O
O
O
O
O
O
O
OH
OH
3
7
Lipid α (5)
[M-2H+]-2 calculated m/z: 563.849
563.848
564.349
564.850
565.353
1200
900
600
300
0
Intensity, counts
556.5
559
561.5
564
566.5
569
571.5
m/z

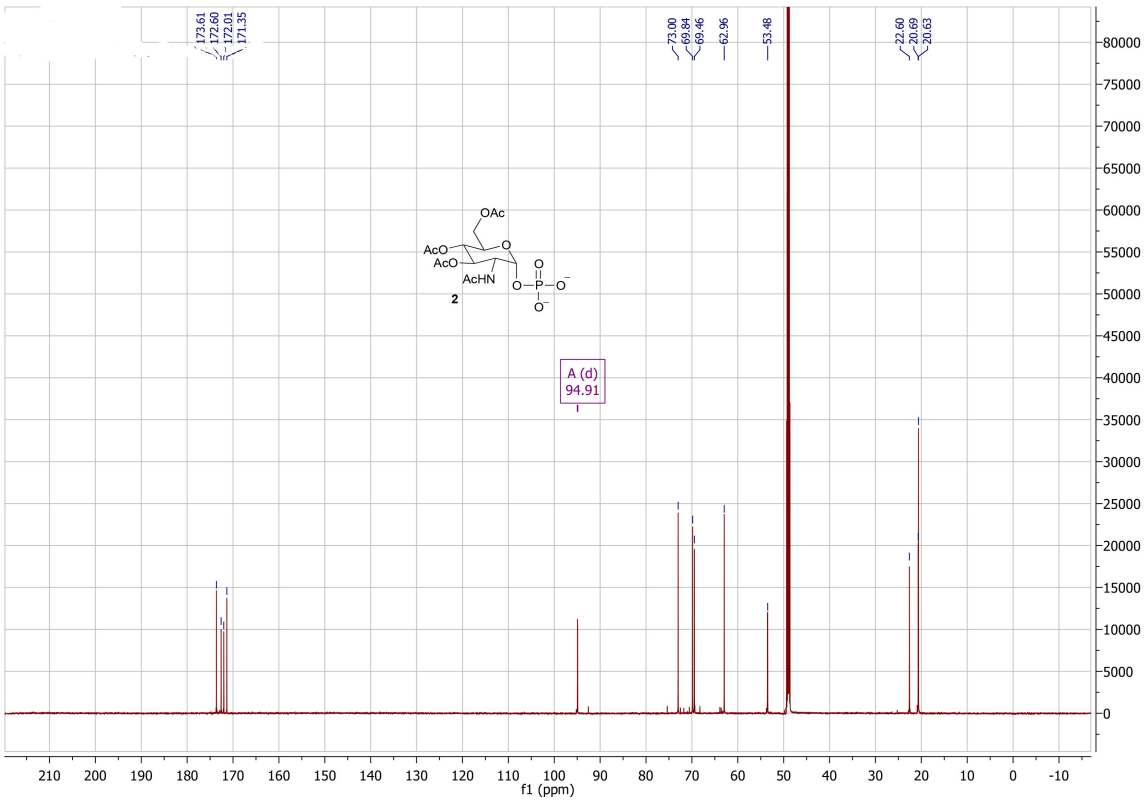
NMR spectra

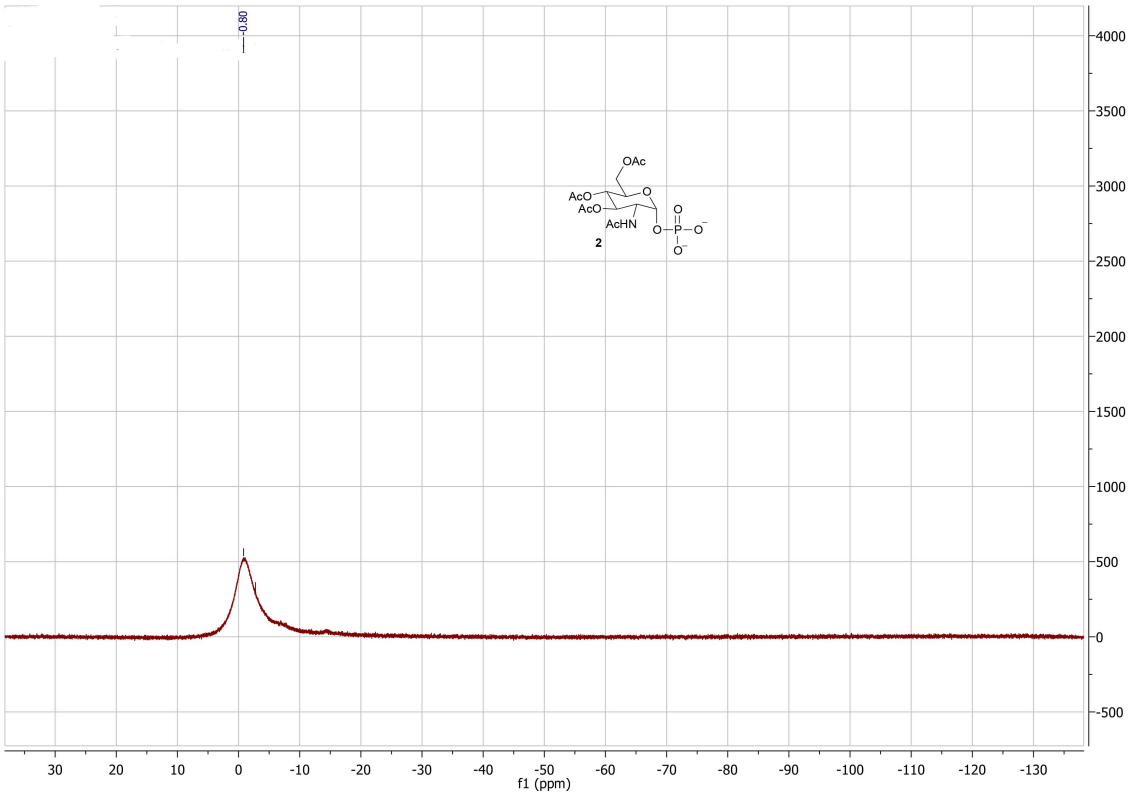


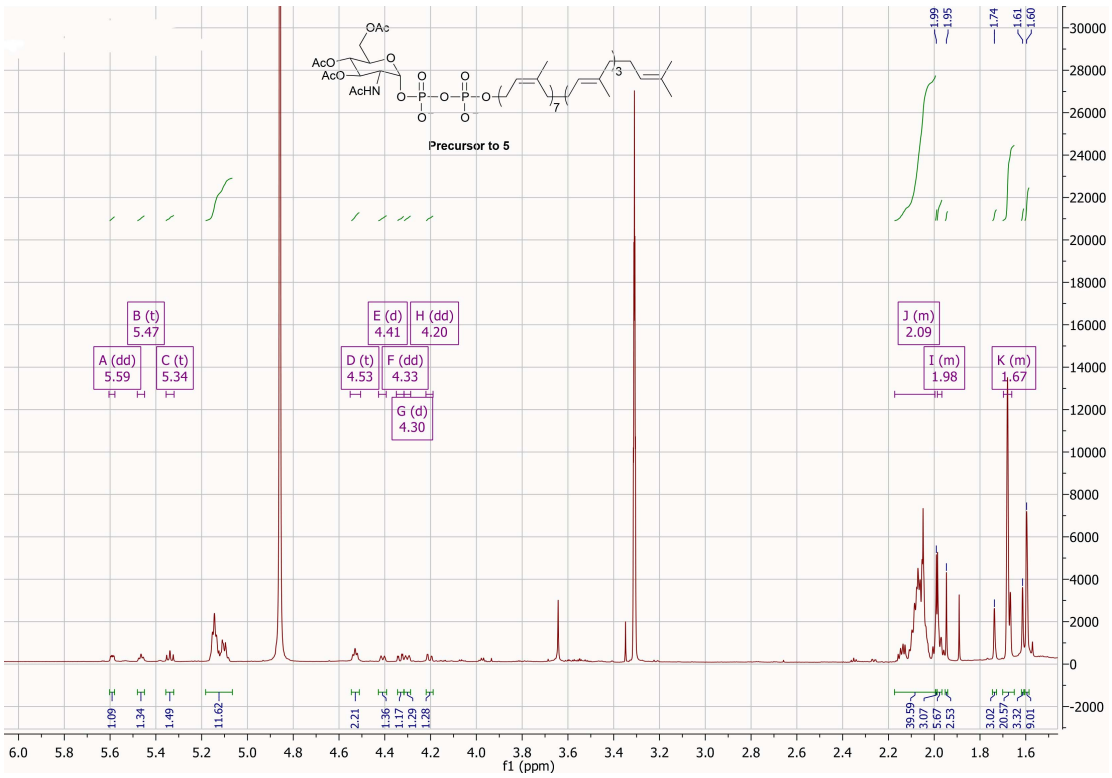


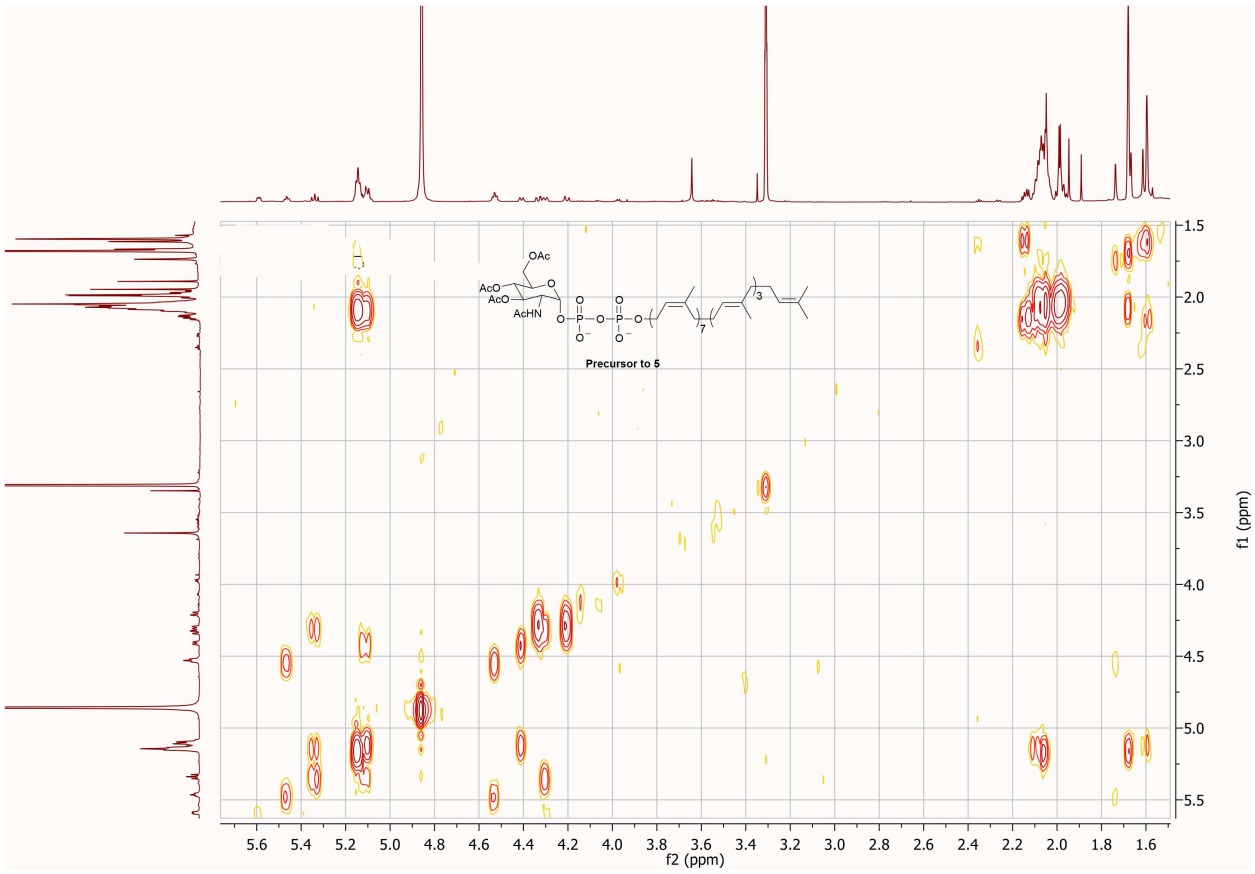

1.37
4
f1 (ppm)

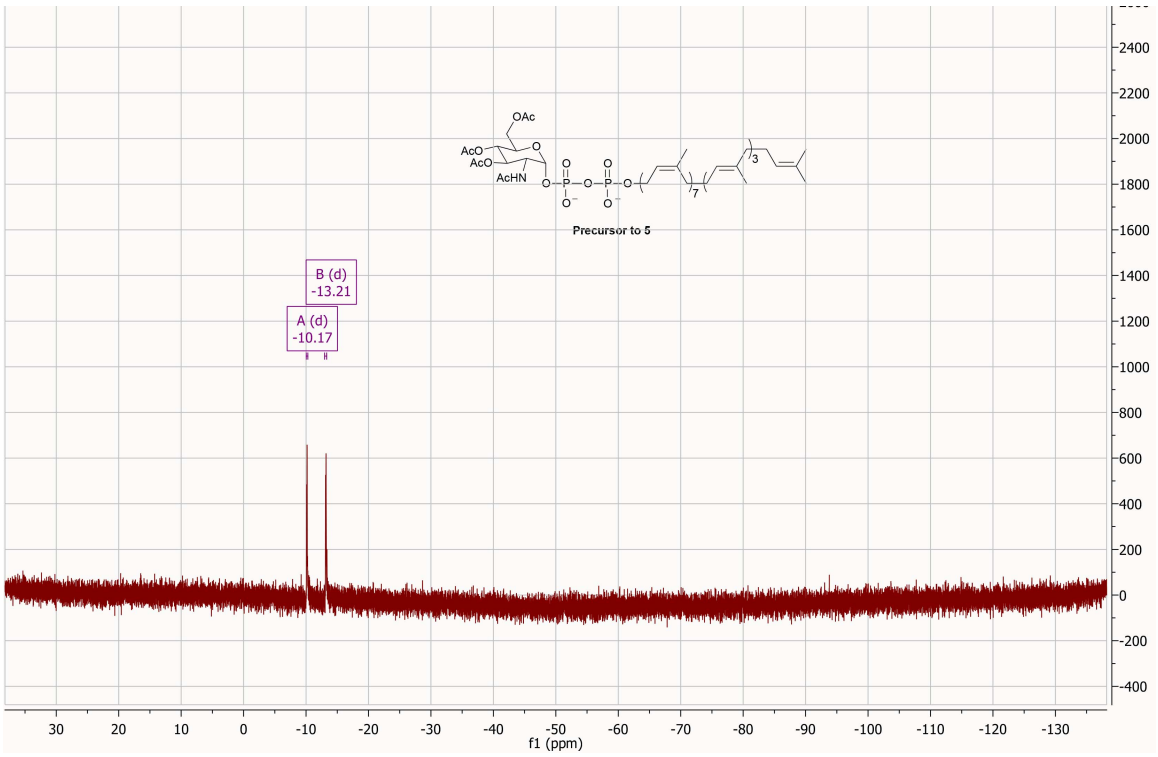

5.48
5.48
5.33
5.32
5.30
5.10
5.09
5.08
4.34
4.32
4.29
4.27
4.25
4.23
4.19
4.17
2.05
2.00
1.96
1.94
OAc
AcO
AcO
AcHN
2
D (d)
4.18
B (d)
4.28
A (d)
4.33
C (d)
4.24
1.00
1.03
1.06
1.04
1.13
1.11
1.04
3.28
3.23
3.16
3.18
f1 (ppm)


173.61
172.60
172.01
171.35
73.00
69.84
69.46
62.96
53.48
22.60
20.69
20.63
OAc
AcO
AcO
AcHN
O
O
O
O
O
O
2
A (d)
94.91
80000
75000
70000
65000
60000
55000
50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)


0.80
OAc
AcO
AcO
AcHN
O
O
O
O
O
O
2
4000
3500
3000
2500
2000
1500
1000
500
0
-500
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
-110
-120
-130
f1 (ppm)


Precursor to 5
B (t) 5.47
A (dd) 5.59
C (t) 5.34
E (d) 4.41
H (dd) 4.20
D (t) 4.53
F (dd) 4.33
G (d) 4.30
J (m) 2.09
I (m) 1.98
K (m) 1.67
f1 (ppm)


Precursor to 5
f2 (ppm)
f1 (ppm)


OAc
AcO
AcO
AcHN
Precursor to 5
B (d)
-13.21
A (d)
-10.17
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
-200
-400
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
-110
-120
-130
f1 (ppm)

3.0 References

1. E. Breukink, H. E. van Heusden, P. J. Vollmerhaus, E. Swiezewska, L. Brunner, S. Walker, A. J. R. Heck, and B. de Kruijff, *J. Biol. Chem.*, 2003, **278**, 19898–19903.
2. N. K. Khidyrova and K. M. Shakhidoyatov, *Chemistry of natural compounds*, 2002, **38**, 107-121.
3. E. Swiezewska, W. Sasak, T. Mankowski, W. Jankowski, T. Vogtman, I. Krajewska, J. Hertel, E. Skoczylas, and T. Chojnacki, *Acta Biochim. Pol.*, 1994, **41**, 221–260.
4. L. L. Danilov, T. N. Druzhinina, N. A. Kalinchuk, S. D. Maltsev and V. N. Shibaev, *Chem Phys Lipids*, 1989, **51**, 191-203.
5. C. Schäffer, T. Wugeditsch, P. Messner, and C. Whitfield, *Appl. Environ. Microbiol.*, 2002, **68**, 4722–4730.
6. M. A. Farha, A. Leung, E. W. Sewell, M. A. D'Elia, S. E. Allison, L. Ejim, P. M. Pereira, M. G. Pinho, G. D. Wright, and E. D. Brown, *ACS Chem. Biol.*, 2013, **8**, 226–233.
7. Y. J. Lee, A. Ishiwata, and Y. Ito, *Tetrahedron*, 2009, **65**, 6310–6319.
8. R. Woodward, W. Yi, L. Li, G. Zhao, H. Eguchi, P. R. Sridhar, H. Guo, J. K. Song, E. Motari, L. Cai, P. Kelleher, X. Liu, W. Han, W. Zhang, Y. Ding, M. Li, and P. G. Wang, *Nat. Chem. Biol.*, 2010, **6**, 418–423.
9. M. M. Sim, H. Kondo, and C. H. Wong, *J. Am. Chem. Soc.*, 1993, **115**, 2260–2267.
10. A. Holkenbrink, D. C. Koester, J. Kaschel, and D. B. Werz, *Eur. J. Org. Chem.*, 2011, **2011**, 6233–6239.
11. E. Bulat and T. A. Garrett, *J Biol Chem*, 2011, **286**, 33819–33831.
12. C. Ginsberg, Y.-H. Zhang, Y. Yuan, and S. Walker, *ACS Chem. Biol.*, 2006, **1**, 25–28.
13. M. P. Pereira, J. W. Schertzer, M. A. D'Elia, K. P. Koteva, D. W. Hughes, G. D. Wright, and E. D. Brown, *ChemBioChem*, 2008, **9**, 1385–1390.