

Development of rare bacterial monosaccharide analogs for metabolic glycan labeling in pathogenic bacteria

Emily L. Clark,¹ Madhu Emmadi,² Katharine L. Krupp,¹ Ananda R. Podilapu,²
Jennifer D. Helble,¹ Suvarn S. Kulkarni,^{2,*} and Danielle H. Dube^{1,*}

¹Department of Chemistry & Biochemistry, Bowdoin College, 6600 College Station, Brunswick, Maine 04011; ²Department of Chemistry, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India

Supporting Information

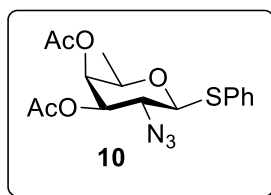
Table of Contents

Synthesis	s-2
General.....	s-2
Phenyl 3,4- <i>O</i> -diacetyl-2-azido-2,6-dideoxy-1-thio- β -D-galactopyranoside (10).....	s-2
Phenyl 3,4- <i>O</i> -diacetyl-2-(azidoacetyl amino)-2,6-dideoxy-1-thio- β -D-galactopyranoside (11).....	s-3
Acetyl 3,4- <i>O</i> -diacetyl-2-(azidoacetyl amino)-2,6-dideoxy- α / β -D-galactopyranoside (4).....	s-4
Phenyl 2,4-(diazidoacetyl amino)-3- <i>O</i> -benzoyl-2,4,6-trideoxy-1-thio- β -D-glucopyranoside (14).....	s-5
Acetyl 3- <i>O</i> -acetyl-2,4-(diazidoacetyl amino)-2,4,6-trideoxy- α / β -D-glucopyranoside (5).....	s-6
Phenyl 2,4-diazido-2,4,6-trideoxy-1-thio- β -D-galactopyranoside (16).....	s-7
Phenyl 2,4-(diazidoacetyl amino)-2,4,6-trideoxy-1-thio- β -D-galactopyranoside (17).....	s-7
Acetyl 3- <i>O</i> -acetyl-2,4-(diazidoacetyl amino)-2,4,6-trideoxy- α / β -D-galactopyranoside (6).....	s-8
Biology	s-9
General.....	s-9
Bacterial growth conditions.....	s-10
Mammalian cell growth and metabolic labeling.....	s-10
Metabolic labeling of bacterial strains.....	s-11
Probing for azides on cells and in lysates via Staudinger ligation.....	s-11
Western blot analysis of azide-labeled proteins.....	s-11
Flow cytometry experiments.....	s-12
Supplementary Figures	s-14
Supplementary Figure 1.....	s-14
Supplementary Figure 2.....	s-15
Supplementary Figure 3.....	s-16
Supplementary Figure 4.....	s-17
Supplemental References	s-18
NMRs	s-19

Synthesis

General. All reactions were conducted under a dry nitrogen atmosphere. Solvents (CH₂Cl₂ >99%, THF 99.5%, Acetonitrile 99.8%, DMF 99.5%) were purchased in capped bottles and dried under sodium or CaH₂. All other solvents and reagents were used without further purification. All glasswares used were oven dried before use. TLC was performed on pre-coated Aluminium plates of Silica Gel 60 F254 (0.25 mm, E. Merck). Developed TLC plates were visualized under a short-wave UV lamp and by heating plates that were dipped in ammonium molybdate/cerium (IV) sulfate solution. Silica gel column chromatography was performed using Silica Gel (100-200 mesh) and employed a solvent polarity correlated with TLC mobility. NMR experiments were conducted on 400 MHz and 500 MHz instruments using CDCl₃ (D, 99.8%) or (CD₃)₂CO (D, 99.9%) as solvents. Chemical shifts are relative to the deuterated solvent peaks and are in parts per million (ppm). ¹H-¹H COSY was used to confirm proton assignments. Mass spectra were acquired in the ESI mode. Melting points were determined by capillary apparatus. Specific rotation experiments were measured at 589 nm (Na) and 25 °C. IR spectra were recorded on an FT-IR spectrometer using CsCl plates. Ac₄GlcNAc (**1**), Ac₄GlcNAz (**2**), Ac₄GalNAz, Ac₃6Az-GlcNAc (**3**), and Phos-FLAG were synthesized as previously described.¹⁻³

Phenyl 3,4-O-diacetyl-2-azido-2,6-dideoxy-1-thio-β-D-galactopyranoside (10)

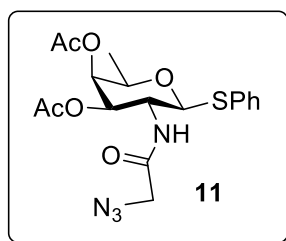


To a clear solution of **9** (0.12 g, 0.38 mmol) in CH₂Cl₂ (2 mL), was added AcCl (0.08 mL, 1.13 mmol) and pyridine (0.09 mL, 1.13 mmol). After 1 h, reaction mixture was concentrated in *vacuo* and chromatographed on silica gel (5% ethyl acetate: petroleum ether) to obtain compound **10**

Supporting Information, Clark *et al.*

as a pale yellowish liquid (0.13 g, 95%). ^1H NMR (400 MHz, CDCl_3) δ 7.65-7.62 (m, 2H, ArH), 7.38-7.33 (m, 3H, ArH), 5.19 (d, $J = 2.8$ Hz, 1H, H-4), 4.87 (dd, $J = 9.4, 2.8$ Hz, 1H, H-3), 4.99 (d, $J = 9.4$ Hz, 1H, H-1), 3.28 (q, $J = 6.2$ Hz, 1H, H-5), 3.62 (t, $J = 9.4$ Hz, 1H, H-2), 2.11 (s, 3H, CH_3), 2.01 (s, 3H, CH_3), 1.23 (d, $J = 6.2$ Hz, 3H, CH_3); ^{13}C NMR (100 MHz, CDCl_3) δ 170.5, 169.9, 133.4, 131.6, 129.1, 128.5, 86.5, 73.5, 73.2, 69.8, 59.4, 20.8, 20.7, 16.7; HR-ESI-MS (m/z): $[\text{M} + \text{Na}]^+$ calcd for $\text{C}_{16}\text{H}_{19}\text{NaN}_3\text{O}_5\text{S}$, 388.0938, found 388.0938.

Phenyl 3,4-O-diacetyl-2-(azidoacetyl amino)-2,6-dideoxy-1-thio- β -D-galactopyranoside (11)



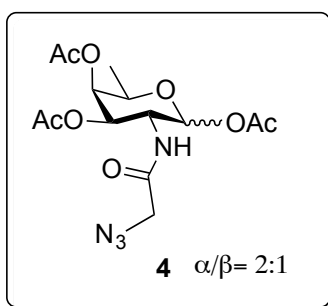
Palladium on carbon (0.1 g) was added to a clear solution of **10** (0.13 g, 0.35 mmol) in EtOH (8 mL) and the solution was stirred at room temperature under hydrogen (1 atm) for 3 h. After complete consumption of starting material the reaction mixture was filtered through celite pad and washed with EtOH. The filtrate was concentrated and the residue was used for the next step without any purification.

The residue obtained after solvents removal was dissolved in CH_3CN (2 mL), to this NaHCO_3 (84 mg, 1.0 mmol) and azido acetic acid (53 mg, 0.53 mmol) were added. After 5 min, EDC (0.1 g, 0.53 mmol) and HOBT (15 mg, 0.09 mmol) were added and the reaction was continued with stirring at rt for 2 h. After complete consumption of starting materials solvent was removed and the crude residue was purified on silica gel (25% ethyl acetate: petroleum ether) to obtain desired product **11** as a foam (0.13 g, 88%). ^1H NMR (500 MHz, CDCl_3) δ 7.54-7.51 (m, 2H, ArH), 7.35-7.30 (m, 3H, ArH), 6.41 (d, $J = 9.2$ Hz, 1H, NH), 5.28-5.24 (m, 2H, H-3 & H-4),

Supporting Information, Clark *et al.*

4.97 (d, $J = 10.0$ Hz, 1H, H-1), 4.19 (q, $J = 10.0$ Hz, 1H, H-2), 3.97 (ABq, $J = 16.6, 5.0$ Hz, 2H, CH₂), 3.86 (q, $J = 6.2$ Hz, 1H, H-5), 2.17 (s, 3H, CH₃), 2.01 (s, 3H, CH₃), 1.26 (d, $J = 6.2$ Hz, 3H, CH₃); ¹³C NMR (125 MHz, CDCl₃) δ 170.8, 170.7, 167.2, 133.1, 132.3, 129.1, 128.0, 86.8, 73.4, 71.4, 70.1, 52.8, 49.9, 20.9, 16.8; HR-ESI-MS (m/z): [M + Na]⁺ calcd for C₁₈H₂₂NaN₄O₆S, 445.1152, found 445.1154.

Acetyl 3,4-O-diacetyl-2-(azidoacetyl amino)-2,6-dideoxy- α/β -D-galactopyranoside (**4**)

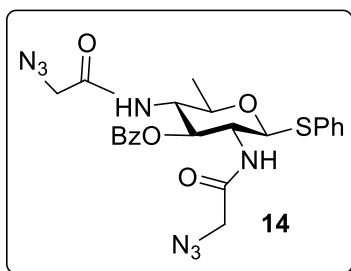


NBS (0.1 g, 0.5 mmol) was added at 0 °C to a cooled solution of **11** (0.07 g, 0.17 mmol) in THF: H₂O (2.5 mL, 4:1). After 10 min, reaction mixture was brought to rt and stirred for 30 min. After complete consumption of starting material reaction mixture was diluted with EtOAc and washed with aq. NaHCO₃. Separated organic layer dried over Na₂SO₄ and concentrated *in vacuo*. The crude product obtained in above step was dissolved in CH₂Cl₂ (1.5 mL) and to this clear solution Ac₂O (30 μ L, 0.3 mmol) and Et₃N (0.13 mL, 0.9 mmol) were added. After stirring at rt for 4 h, solvents were removed in *vacuo* and the crude product was purified by column chromatography on silica gel (50% ethyl acetate: pet ether) to afford the desired product **4** as a viscous liquid (50 mg, 76%). ¹H NMR (500 MHz, CDCl₃) δ 6.48 (d, $J = 9.4$ Hz, 1H, NH α), 6.33 (d, $J = 9.4$ Hz, 1H, NH β), 6.18 (d, $J = 3.7$ Hz, 1H, H-1 α), 5.74 (d, $J = 10.0$ Hz, 1H, H-1 β), 5.27-5.14 (m, 2H, H-3 & H-4), 4.68-4.63 (m, 1H, H-2 α), 4.39 (q, $J = 10.0$ Hz, 1H, H-2 β), 4.16 (q, $J = 6.2$ Hz, 1H, H-5 α), 3.93-3.89 (m, 3H, H-5 β & CH₂), 2.18 (s, 3H, CH₃), 2.16, 2.10 (s, 3H, CH₃), 2.01, 1.99 (s, 3H, CH₃), 1.21 (d, $J = 6.2$ Hz, 3H, CH₃ β), 1.14 (d, $J = 6.2$ Hz, 3H, CH₃ α); ¹³C NMR (125 MHz,

Supporting Information, Clark *et al.*

CDCl_3) δ 171.3, 170.8, 169.8, 169.4, 169.3, 167.2, 92.8, 91.4, 70.6, 69.9, 69.5, 68.2, 67.4, 52.7, 52.6, 49.9, 47.1, 29.7, 21.5, 21.0, 20.8, 16.2; HR-ESI-MS (m/z): $[\text{M} + \text{Na}]^+$ calcd for $\text{C}_{14}\text{H}_{20}\text{NaN}_4\text{O}_8$, 395.1173, found 395.1105.

Phenyl 2,4-(diazidoacetyl amino)-3-O-benzoyl-2,4,6-trideoxy-1-thio- β -D-glucopyranoside (14)



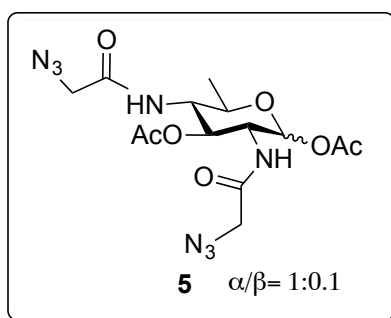
Palladium on carbon (0.2 g) was added to a clear solution of **13** (0.15 g, 0.36 mmol) in EtOH (10 mL) and the solution was stirred at room temperature under hydrogen (1 atm) for 10 h. After complete consumption of starting material the reaction mixture was filtered through celite pad and washed with EtOH. The filtrate was concentrated and the residue was used for the next step without any purification.

The residue obtained after solvents removal was dissolved in CH_3CN (2 mL), to this NaHCO_3 (0.18 g, 2.1 mmol) and azido acetic acid (0.11 g, 1.1 mmol) were added. After 5 min, EDC (0.21 g, 1.1 mmol) and HOBT (30 mg, 0.18 mmol) were added and the reaction was continued with stirring at rt for 2 h. After complete consumption of starting materials solvent was removed and the crude residue was purified on silica gel (30% ethyl acetate: petroleum ether) to obtain desired product **14** as a foam (0.12 g, 72%). ^1H NMR (500 MHz, CDCl_3) δ 7.94 (d, J = 8.2 Hz, 1H, ArH), 7.61-7.54 (m, 3H, ArH), 7.45-7.41 (m, 2H, ArH), 7.37-7.31 (m, 3H, ArH), 7.12 (d, J = 8.8 Hz, 1H, NH), 7.08 (d, J = 8.8 Hz, 1H, NH), 6.00 (t, J = 10.0 Hz, 1H, H-3), 5.26 (d, J =

Supporting Information, Clark *et al.*

10.0 Hz, 1H, H-1), 4.14-3.91 (m, 3H, H-2, H-4 & H-5), 3.89-3.70 (m, 4H, 2CH₂), 1.41 (d, $J = 6.2$ Hz, 3H, CH₃); ¹³C NMR (125 MHz, CDCl₃) δ 167.9, 167.5, 167.2, 134.2, 133.0, 132.6, 130.0, 129.2, 128.9, 128.7, 128.3, 85.7, 75.1, 73.8, 55.9, 54.4, 52.7, 52.6, 18.4; HR-ESI-MS (m/z): [M + Na]⁺ calcd for C₂₃H₂₄NaN₈O₅S, 547.1483, found 547.1475.

Acetyl 3-O-acetyl-2,4-(diazidoacetyl amino)-2,4,6-trideoxy- α/β -D-glucopyranoside (**5**)



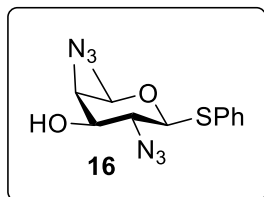
NaOMe (30 mg) was added to clear solution of **14** (0.3 g, 0.56 mmol) in MeOH (3 mL) and stirred at rt for 30 min. After complete consumption of starting material, reaction mixture was neutralized with Amberlite (H⁺, 0.3 g) and the reaction mixture was filtered and concentrated. NBS (0.3 g, 1.7 mmol) was added at 0 °C to a cooled solution of above obtained residue in THF: H₂O (10.0 mL, 4:1). After 10 min, reaction mixture was brought to rt and stirred for 30 min. After complete consumption of starting material, reaction mixture was diluted with EtOAc and washed with aq. NaHCO₃. Separated organic layer was dried over Na₂SO₄ and concentrated *in vacuo*. The crude product obtained in above step was dissolved in CH₂Cl₂ (1.5 mL) and to this clear solution Ac₂O and Et₃N were added. After stirring at rt for 6 h, solvents were removed *in vacuo* and the crude product was purified by column chromatography on silica gel (70% ethyl acetate: pet ether) to afford the desired product **5** as a white solid (0.16 g, 68%). ¹H NMR (400 MHz, CDCl₃) δ 6.33 (d, $J = 9.2$ Hz, 1H, NH), 6.29 (d, $J = 9.2$ Hz, 1H, NH), 6.16 (d, $J = 3.7$ Hz, 1H, H-1), 5.21 (t, $J = 10.0$ Hz, 1H, H-3), 4.47-4.41 (m, 1H, H-2), 3.99-3.90 (m, 6H, H-4, H-5 & 2CH₂), 2.08 (s, 3H, CH₃), 2.07 (s, 3H, CH₃), 1.23 (d, $J = 6.2$ Hz, 3H, CH₃); ¹³C NMR (125 MHz, CDCl₃) δ

Supporting Information, Clark *et al.*

172.2, 169.2, 166.9, 90.7, 70.1, 69.1, 55.0, 52.7, 52.6, 51.4, 21.1, 20.9, 18.0; HR-ESI-MS (m/z):

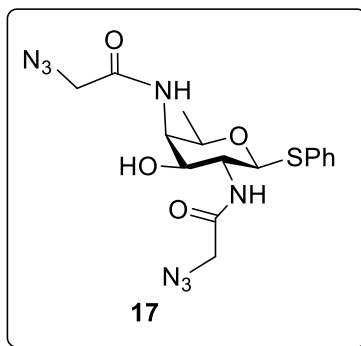
$[M + Na]^+$ calcd for $C_{14}H_{20}NaN_8O_7$, 435.1347, found 435.1362.

Phenyl 2,4-diazido-2,4,6-trideoxy-1-thio- β -D-galactopyranoside (**16**)



NaOMe (50 mg) was added to clear solution of **15** (0.5 g, 1.22 mmol) in MeOH (5 mL) and stirred at rt for 30 min. After complete consumption of starting material reaction mixture was neutralized with Amberlite (H^+ , 0.5 g) and the reaction mixture was filtered, concentrated and the crude product was purified by column chromatography on silica gel (15% ethyl acetate: pet ether) to afford the desired product **16** as a yellowish liquid (0.36 g, 96%). 1H NMR (500 MHz, $CDCl_3$) δ 7.62-7.61 (m, 2H, ArH), 7.36-7.28 (m, 3H, ArH), 4.40 (d, $J = 10.0$ Hz, 1H, H-1), 3.72-3.66 (m, 3H, H-3, H-4 & H-5), 3.54 (t, $J = 9.4$ Hz, 1H, H-2), 1.41 (d, $J = 6.2$ Hz, 3H, CH_3); ^{13}C NMR (125 MHz, $CDCl_3$) δ 133.2, 131.7, 129.2, 128.5, 86.9, 74.5, 74.1, 65.1, 62.8, 18.0; HR-ESI-MS (m/z): $[M + Na]^+$ calcd for $C_{12}H_{14}NaN_6O_2S$, 329.0791, found 329.0796.

Phenyl 2,4-(diazidoacetyl amino)-2,4,6-trideoxy-1-thio- β -D-galactopyranoside (**17**)

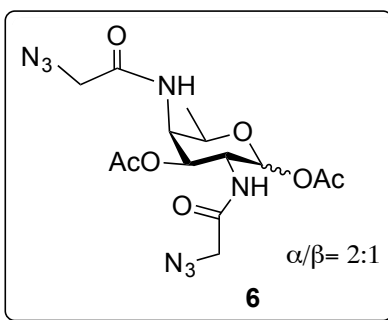


Palladium on carbon (0.35 g) was added to a clear solution of **16** (0.2 g, 0.65 mmol) in EtOH (14

mL) and the solution was kept for stirring at room temperature under hydrogen (1 atm) 10 h. After complete consumption of starting material, the reaction mixture was filtered through celite pad and washed with EtOH. The filtrate was concentrated and the residue was used for the next step without any purification.

The residue obtained after solvents removal was dissolved in CH₃CN (3 mL), to this NaHCO₃ (0.47 g, 5.6 mmol) and azido acetic acid (0.3 g, 2.9 mmol) were added. After 5 min, EDC (0.56 g, 2.93 mmol) and HOBT (75 mg, 0.49 mmol) were added and the stirring was continued at rt for 2 h. After complete consumption of starting material, solvent was removed under reduced pressure. The crude residue was dissolved in MeOH (2 mL) and to this NaOMe (20 mg) was added. After 20 min, reaction mixture was neutralized with Amberlite (H⁺, 0.2 g) and filtered. Filtrate was concentrated and the crude product was purified by column chromatography on silica gel (35% ethyl acetate: pet ether) to afford the desired product **17** as a yellowish liquid (0.20 g, 74%). ¹H NMR (500 MHz, CDCl₃) δ 7.54-7.51 (m, 2H, ArH), 7.36-7.29 (m, 3H, ArH), 6.72 (d, *J* = 8.4 Hz, 1H, NH), 6.56 (d, *J* = 8.4 Hz, 1H, NH), 4.80 (d, *J* = 10.0 Hz, 1H, H-1), 4.31 (dd, *J* = 10.0, 3.2 Hz, 1H, H-4), 4.10-3.98 (m, 5H, H-3 & 2CH₂), 3.84 (q, *J* = 6.4 Hz, 1H, H-5), 3.62 (q, *J* = 10.0 Hz, 1H, H-2), 1.23 (d, *J* = 6.4 Hz, 3H, CH₃); ¹³C NMR (125 MHz, CDCl₃) δ 169.4, 168.7, 132.9, 132.3, 129.4, 128.2, 86.5, 73.9, 72.2, 53.6, 52.6, 52.5, 52.3, 17.2; HR-ESI-MS (*m/z*): [M + Na]⁺ calcd for C₁₆H₂₀NaN₈O₄S, 443.1220, found 443.1220.

Acetyl 3-O-acetyl-2,4-(diazidoacetyl amino)-2,4,6-trideoxy- α/β -D-galactopyranoside (6)



Supporting Information, Clark *et al.*

NBS (0.11 g, 0.6 mmol) was added at 0 °C to a cooled solution of **17** (87 mg, 0.2 mmol) in THF:H₂O (2.5 mL, 4:1). After 10 min, reaction mixture was brought to rt and stirred for 30 min. After complete consumption of starting material, reaction mixture was diluted with EtOAc and washed with aq. NaHCO₃. Separated organic layer dried over Na₂SO₄ and concentrated *in vacuo*. The crude product obtained in above step was dissolved in CH₂Cl₂ (1.5 mL) and to this clear solution Ac₂O (60 μL, 0.6 mmol) and Et₃N (0.26 mL, 1.8 mmol) were added. After stirring at rt for 4 h, solvents were removed in vacuo and the crude product was purified by column chromatography on silica gel (70% ethyl acetate: pet ether) to afford the desired product **6** as a viscous liquid (62 mg, 72%). ¹H NMR (500 MHz, CDCl₃) δ 6.66 (d, *J* = 9.4 Hz, 1H, NH), 6.62(d, *J* = 9.4 Hz, 1H, NH), 6.38 (d, *J* = 9.4 Hz, 1H, NH), 6.32(d, *J* = 9.4 Hz, 1H, NH), 6.18 (d, *J* = 3.8 Hz, 1H, H-1 α), 5.67 (d, *J* = 10.0 Hz, 1H, H-1 β), 5.21 (d, *J* = 2.8, 10.0 Hz, 1H, H-3 α), 5.09 (d, *J* = 2.8, 10.0 Hz, 1H, H-3 β), 4.54-4.13 (m, 3H, H-2, H-4 & H-5), 4.09 (s, 2H, CH₂), 3.94 (s, 1H, CH₂), 3.92 (s, 1H, CH₂), 1.24 (d, *J* = 6.2 Hz, 3H, CH₃), 1.16 (d, *J* = 6.2 Hz, 3H, CH₃); ¹³C NMR (125 MHz, CDCl₃) δ 169.1, 167.8, 167.5, 167.3, 93.3, 90.8, 71.4, 70.8, 68.9, 67.1, 52.7, 52.66, 52.62, 50.7, 50.5, 50.0, 47.2, 29.8, 21.1, 20.9, 16.73, 16.70; HR-ESI-MS (*m/z*): [M + Na]⁺ calcd for C₁₄H₂₀NaN₈O₇, 435.1347, found 435.1364.

Biology

General. All biological reagents were obtained from commercial suppliers and used without further purification. *Helicobacter pylori* strain G27⁴ was a gift from Manuel Amieva (Stanford University). *Mycobacterium smegmatis* strain Mc²155::otsA treS treY⁵ was a gift from Peter Woodruff (University of Southern Maine). *Campylobacter jejuni* (ATCC 33560). *Burkholderia thailandensis* (ATCC 700388), *Ralstonia solanacearum* ATCC 33291* (*incorrectly cataloged as *C. jejuni*; species confirmed as *Ralstonia solanacearum* by 16S ribosomal RNA sequencing),

Supporting Information, Clark *et al.*

Bacteroides fragilis ATCC 23745, *Pseudomonas aeruginosa* ATCC 39324, and MDCK cells (ATCC CCL-34) were purchased from ATCC and grown as described below.

Bacterial growth conditions. Six bacterial species were used to examine incorporation of unnatural azide-containing sugars. *H. pylori* was grown on horse blood agar plates (4% Columbia agar, 5% horse blood, 10 µg/mL vancomycin, 5 µg/mL cefsulodin, 0.3 µg/mL polymixin B, 5 µg/mL trimethoprim, and 8 µg/mL amphotericin B) or in Brucella Broth (with 10% Fetal Bovine Serum (FBS) and 6 µg/mL vancomycin) for 3-5 days in 14% CO₂ at 37 °C. *M. smegmatis* was grown on Middlebrook 7H9 agar plates or in Middlebrook 7H9 broth (with 0.25% Tween 80) at 37 °C in an aerobic environment for 3-5 days. *R. solanacearum* and *C. jejuni* were grown for 3-5 days on 4% Müller-Hinton agar plates or in Müller-Hinton broth in 14% CO₂ at 37 °C. *B. thailandensis* was grown on 2.3% nutrient agar plates or in 2.5% nutrient broth at 30 °C in an aerobic environment for 3-5 days. *B. fragilis* was grown overnight on brain-heart infusion plates (1.5% Bacto agar, 3.7% brain-heart infusion broth, 0.5% yeast extract, and 15 µg/mL hematin porcine) or in brain-heart infusion broth (3.7% brain-heart infusion broth, 0.5% yeast extract, and 15 µg/mL hematin porcine). *B. fragilis* cultures were incubated at 37 °C under anaerobic conditions generated using a GasPak™ EZ Anaerobe Pouch System (Becton Dickinson).⁶

Mammalian cell growth and metabolic labeling. MDCK cells were grown in Dulbecco's Modified Eagle Medium (DMEM) (Life Technologies) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (P/S) in a 5% CO₂, water saturated incubator at 37°C. Cells were passaged every 3 days by lifting using trypsin-EDTA (Gibco) and plating at a density of 2 x 10⁵ cells/mL in supplemented DMEM. For metabolic labeling, cells were seeded into DMEM with 10% fetal bovine serum (FBS), penicillin/streptomycin, and 250 µM of

Supporting Information, Clark *et al.*

Ac₄GlcNAc (**1**), Ac₄GlcNAz (**2**), Ac₃6-AzGlcNAc (**3**), Ac₃FucNAz (**4**), Ac₂Bac-diNAz (**5**), Ac₂DATDH-diNAz (**6**), or Ac₄GalNAz, then grown for 3 days in a 5% CO₂ water saturated incubator at 37°C. After three days of metabolic labeling, cells were lifted using trypsin-EDTA, rinsed three times with PBS, then analyzed for azides via Staudinger ligation (see below).

Metabolic labeling of bacterial strains. Bacterial strains from a frozen stock were streaked onto agar plates using a sterile tip applicator and then incubated under appropriate conditions for the bacterial strain (see *Bacterial growth conditions*). The bacteria from plates were transferred with a sterile tip applicator to 3 mL of liquid growth containing 1 mM of Ac₄GlcNAc (**1**), Ac₄GlcNAz (**2**), Ac₃6-AzGlcNAc (**3**), Ac₃FucNAz (**4**), Ac₂Bac-diNAz (**5**), Ac₂DATDH-diNAz (**6**), or Ac₄GalNAz. Once cells reached stationary phase, they were centrifuged at 3500 rpm using a Sorvall Legend RT⁺ centrifuge (Thermo Scientific, Waltham, MA) and washed three times with PBS.

Probing for azides on cells and in lysates via Staudinger ligation. To probe for total azide-labeled glycoproteins in cells, metabolically labeled and rinsed cells were lysed in lysis buffer (20 mM Tris-HCl, pH 7.4, 1% Igepal, 150 mM NaCl, 1 mM EDTA) containing protease inhibitor cocktail (Sigma Aldrich, St. Louis, MO). Lysates were subsequently standardized to ~3 mg/mL, reacted 1:1 with 500 μM Phos-FLAG at 37°C for 7-12 hours, and analyzed by Western blot, as described below. Alternatively, the presence of cell surface azides on metabolically labeled bacteria was determined by performing Staudinger ligation with cell-impermeable Phos-FLAG (500 μM for 5-6 hours) on intact cells.⁷ Phos-FLAG reacted cells were then rinsed three times with PBS and lysed to facilitate analysis by Western blot.

Western blot analysis of azide-labeled proteins. Protein concentrations of samples were

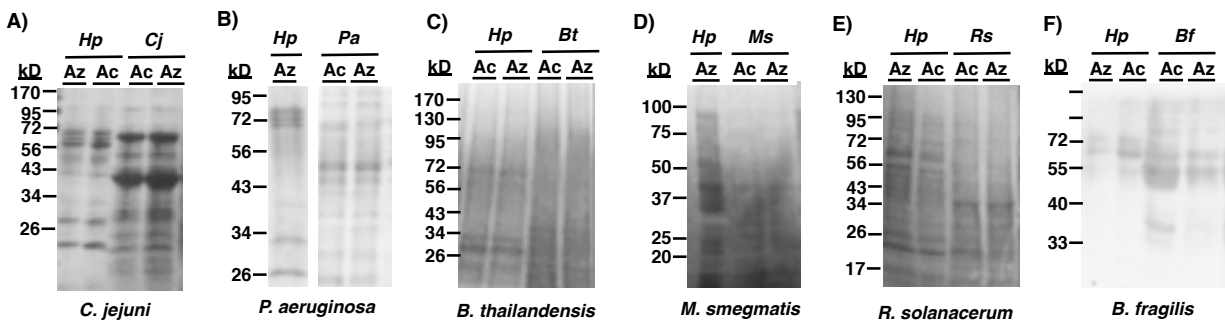
measured using the DC Protein Assay (Bio-Rad, Hercules, CA) and standardized to equal concentrations (~3 mg/mL). All samples were then combined in a 1:1 ratio with 2X SDS reducing loading buffer and boiled at 95°C for 5-10 minutes. Samples (20 µg for bacterial samples; 14 µg for MDCK samples), alongside a molecular weight ladder (EZ-Run Prestained Rec Protein Ladder, Fisher Scientific), were loaded onto a 4-15% Tris-HCl gradient gel (Bio-Rad) or a 12% Tris-HCl SDS-PAGE gel with a 4% stacking layer. After electrophoresis at 200 V for 40 minutes on ice, proteins were transferred to a nitrocellulose membrane (Bio-Rad) at 100 V for 1 hour or stained with Coomassie (Stain: 45% deionized water, 45% Methanol, 10% acetic acid, 0.25% Coomassie brilliant blue/Destain: 50% deionized water, 40% methanol, 10% acetic acid) or Zinc Stain (BioRad) to visualize equal protein loading. Immunoblots were blocked for 1 hour with 5% non-fat dried milk in 0.05% TBS-T buffer (5 mM Tris-HCl, 0.05% Tween-20 (BioRad), pH 7.4. Anti-FLAG-HRP (1:1000 dilution; Sigma Aldrich) was employed to visualize FLAG-tagged proteins via chemiluminescence with film (Thermo Scientific) or G:BOX Chemi XRQ gel doc system (Syngene).

Flow cytometry experiments. To complement Western blot analyses, the presence of azides on cell surfaces was also probed via flow cytometry. For these experiments, metabolically labeled cells were reacted with Phos-FLAG (500 µM in PBS) for 1 hour. After washing with FACS buffer (1X PBS containing 1% fetal bovine serum (FBS)), cells were immunostained with α-FLAG-FITC (1:100 dilution in FACS buffer; Sigma-Aldrich; ex: 488/em: 519) for 1 hour, rinsed with FACS buffer, and analyzed using a BD FACSCalibur flow cytometer. Alternatively, azide-containing sugars were detected using Click-IT Alexa Fluor 488 DIBO Alkyne for copper free click chemistry detection of azide (Thermo Fisher Scientific). In this case, whole cells metabolically labeled as described above (see “Metabolic Labeling”) were reacted with 20 µM Click-IT Alexa Fluor 488 DIBO Alkyne for copper free click chemistry detection of azide (ex:

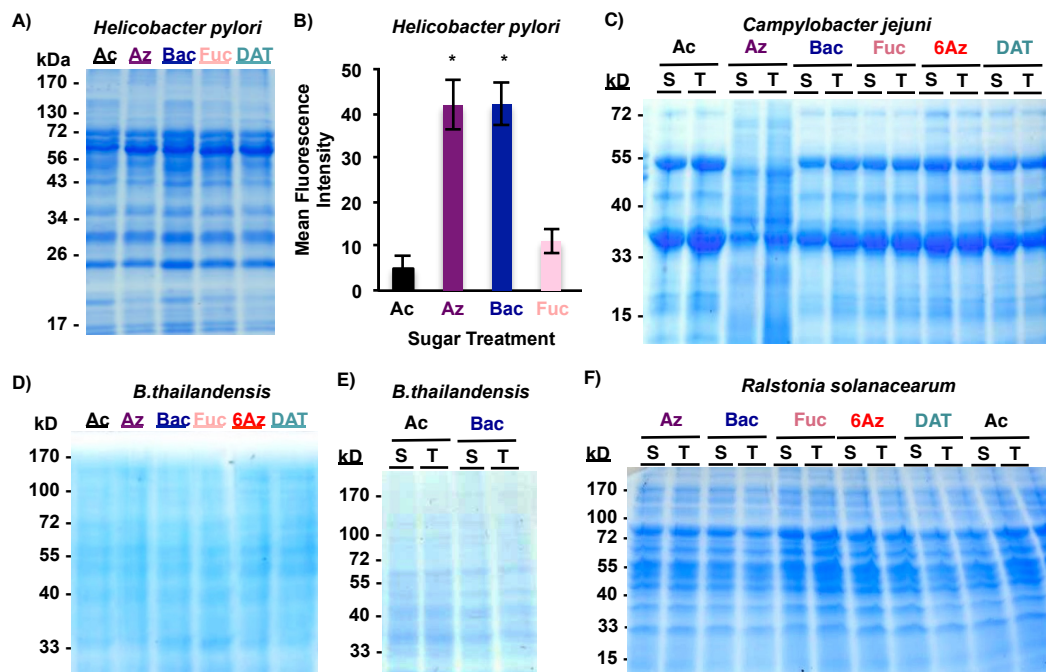
Supporting Information, Clark *et al.*

488/em: 519)⁸ under standard growth conditions for 5 hours in the dark. Cells were then rinsed with PBS supplemented with 3% BSA, and analyzed using a BD FACSCalibur flow cytometer. For this single color analysis, 20,000 live cells were gated on for each replicate. Labeling was performed in triplicate and is reported as number of cells versus fluorescence intensity. Alternatively, flow data are reported as the mean fluorescence intensity (MFI) of a population of cells from replicate experiments, as calculated using FlowJo software (Ashland, OR).

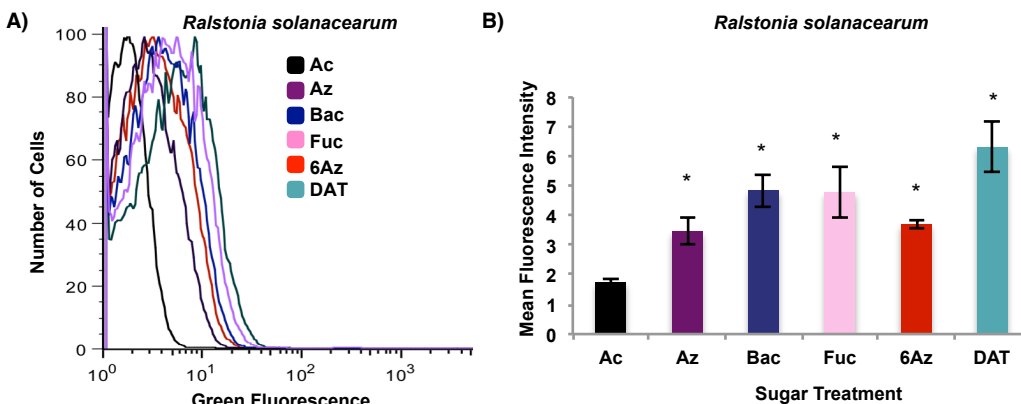
Supplemental Figures



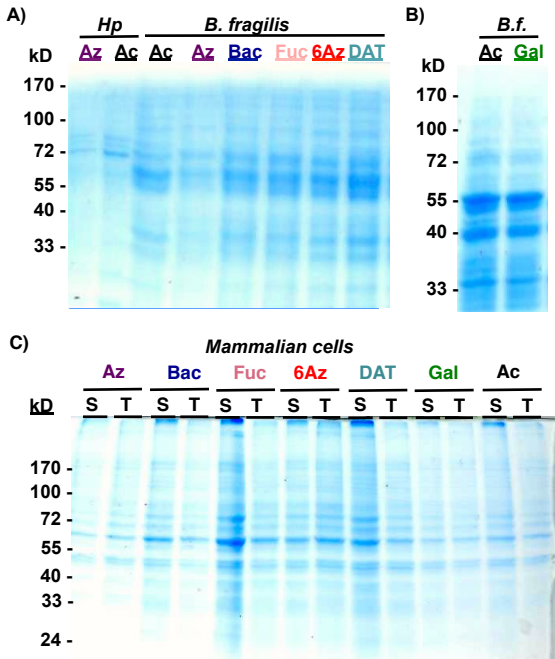
Supplementary Figure 1. Samples probed by Western blot in Figure 3 were electrophoresed and stained with Coomassie (A, B, F) or zinc stain (C, D, E) to reveal that samples contain protein in each experiment. For these experiments, lysates from bacterial cells treated with 1 mM Ac₄GlcNAz (**2**, Az) or the azide-free control sugar Ac₄GlcNAc (**1**, Ac) were reacted with Phos-FLAG (250 μM), then analyzed by SDS-PAGE with Coomassie or zinc stain (BioRad) to detect protein levels.



Supplementary Figure 2. Complementary data to those presented in Figure 4. (A, C, D, E, F) Coomassie staining of electrophoresed samples from Figure 4 reveal that all Western samples contain roughly equivalent protein levels. (B) Flow cytometry indicate that *H. pylori* robustly express surface-accessible azides after treatment with Ac₄GlcNAz (**2**, Az) and Ac₂Bac-diNAz (**5**, Bac). For this experiment, *H. pylori* were metabolically labeled in media containing 1 mM of the indicated sugars, probed for the presence of azides using Phos-FLAG and anti-FLAG-FITC, and analyzed by flow cytometry. Flow cytometry data are presented as mean fluorescence intensity from triplicate samples. Error bars represent standard deviations of triplicate samples. Asterisks indicate significantly elevated MFI relative to the azide-free control Ac₄GlcNAc (**1**, Ac) ($p < 0.05$ by Student's t-test).



Supplementary Figure 3. Flow cytometry data reveal that *R. solanacearum* robustly express surface-accessible azides after treatment with Ac₄GlcNAc (**1**, Ac), Ac₄GlcNAz (**2**, Az), Ac₂Bac-diNAz (**5**, Bac), Ac₃FucNAz (**4**, Fuc), Ac₃6Az-GlcNAc (**3**, 6Az), and Ac₂DATDG-diNAz (**6**, DAT). For this experiment, *R. solanacearum* were metabolically labeled in liquid media containing 1 mM of the indicated sugars, probed for the presence of azides using Phos-FLAG and anti-FLAG-FITC, and analyzed by flow cytometry. Flow cytometry data are presented as single parameter histograms (A) and mean fluorescence intensity from triplicate samples (B). Error bars represent standard deviations of triplicate samples. Asterisks indicate significantly elevated MFI relative to the azide-free control Ac₄GlcNAc (**1**, Ac) ($p < 0.05$ by Student's t-test).



Supplementary Figure 4. Samples probed by Western blot in Figure 5 were electrophoresed and stained with Coomassie to reveal that all samples contain roughly equivalent protein levels. (A, B) SDS-PAGE analysis of samples from *H. pylori* and *B. fragilis* treated with 1 mM Ac₄GlcNAc (1, Ac), Ac₄GlcNAz (2, Az), Ac₂Bac-diNAz (5, Bac), Ac₃FucNAz (4, Fuc), Ac₃6Az-GlcNAc (3, 6Az), and Ac₂DATDG-diNAz (6, DAT), then reacted with Phos-FLAG (250 μM). Proteins were visualized via Coomassie stain. (C) SDS-PAGE analysis of MDCK samples from cells cultured in media supplemented with unnatural sugars (0.25 mM), reacted with Phos-FLAG (250 μM), then electrophoresed and stained with Coomassie.

Supplemental References

- [1] Luchansky, S. J., Hang, H. C., Saxon, E., Grunwell, J. R., Yu, C., Dube, D. H., and Bertozzi, C. R. (2003) Constructing azide-labeled cell surfaces using polysaccharide biosynthetic pathways, *Methods in Enzymology* 362, 249-272.
- [2] Laughlin, S. T., and Bertozzi, C. R. (2007) Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation, *Nature Protocols* 2, 2930 - 2944.
- [3] Chuh, K. N., Zaro, B. W., Piller, F., Piller, V., and Pratt, M. R. (2014) Changes in metabolic chemical reporter structure yield a selective probe of O-GlcNAc modification, *Journal of the American Chemical Society* 136, 12283-12295.
- [4] Baltrus, D. A., Amieva, M. R., Covacci, A., Lowe, T. M., Merrell, D. S., Ottemann, K. M., Stein, M., Salama, N. R., and Guillemin, K. (2009) *The complete genome sequence of Helicobacter pylori* strain G27, *J. Bacteriol.* 191, 447-448.
- [5] Woodruff, P. J., Carlson, B. L., Siridechadilok, B., Pratt, M. R., Senaratne, R. H., Mougous, J. D., Riley, L. W., Williams, S. J., and Bertozzi, C. R. (2004) Trehalose is required for growth of *Mycobacterium smegmatis*, *Journal of Biological Chemistry* 279, 28835-28843.
- [6] Brigham, C., Caughlan, R., Gallegos, R., Dallas, M. B., Godoy, V. G., and Malamy, M. H. (2009) Sialic acid (N-acetyl neuraminic acid) utilization by *Bacteroides fragilis* requires a novel N-acetyl mannosamine epimerase, *Journal of Bacteriology* 191, 3629-3638.
- [7] Kaewsapsak, P., Esonu, O., and Dube, D. H. (2013) Recruiting the host's immune system to target *Helicobacter pylori's* surface glycans, *ChemBioChem* 14, 721-726.
- [8] Geva-Zatorsky, N., Alvarez, D., Hudak, J. E., Reading, N. C., Erturk-Hasdemir, D., Dasgupta, S., von Andrian, U. H., and Kasper, D. L. (2015) *In vivo* imaging and tracking of host-microbiota interactions via metabolic labeling of gut anaerobic bacteria, *Nature Medicine* 21, 1091-1100.

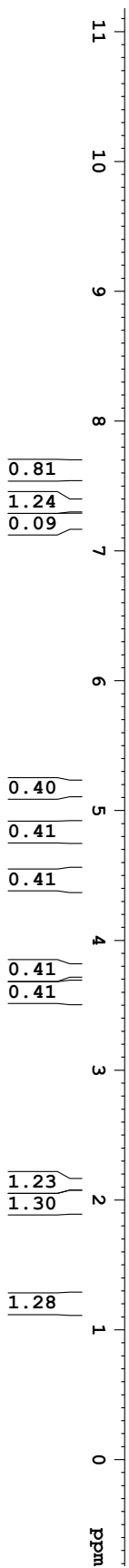
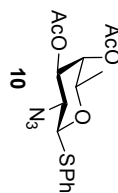
SSK-6-ME-1153-1H

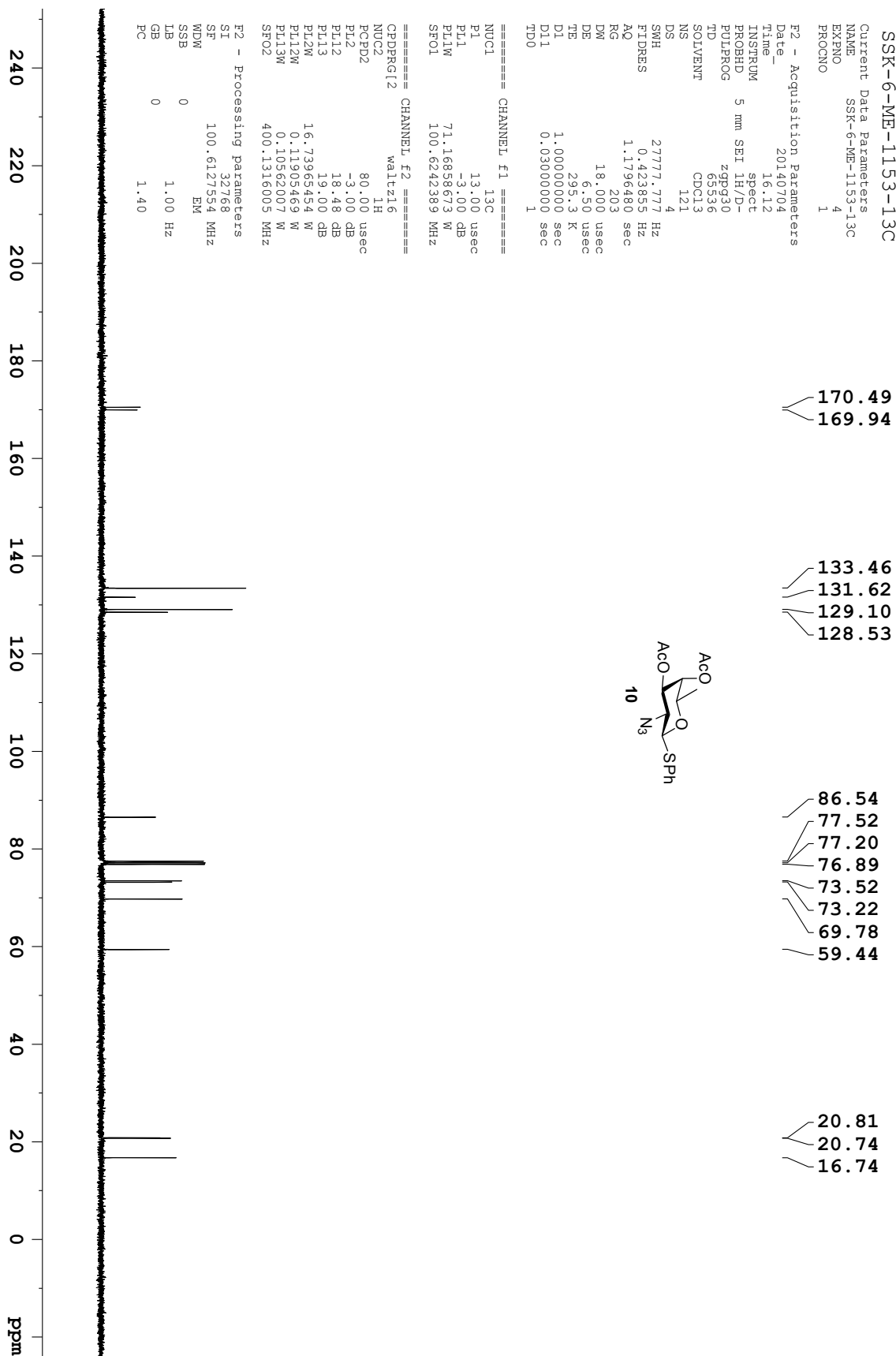
Current Data Parameters
 NAME SSK-6-ME-1153-1H
 EXPNO 3
 PROCNO 1

F2 - Acquisition Parameters
 Date_ 20140704
 Time 16.09
 INSTRUM spect
 PROBHD 5 mm SEI 1H/D-
 PULPROG zg30
 TD 49340
 SOLVENT CDCl3
 NS 31
 DS 0
 SWH 8223.685 Hz
 FIDRES 0.166674 Hz
 AQ 2.9998720 sec
 RG 57
 DW 60.800 usec
 DE 6.50 usec
 TE 295.2 K
 D1 1.0000000 sec
 TDO 1

==== CHANNEL f1 =====
 NUCl 1H
 P1 6.75 usec
 PL1 -3.00 dB
 PL1W 16.73965454 W
 SF01 400.1324710 MHz

F2 - Processing parameters
 SI 32768
 SF 400.1300105 MHz
 WDW EM
 SSB 0
 LB 0.30 Hz
 GB 0
 PC 1.00





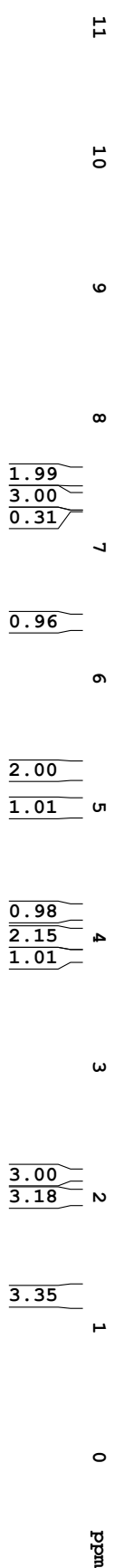
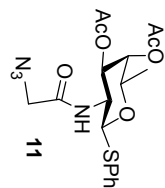
SSK-6-ME-1075-1H

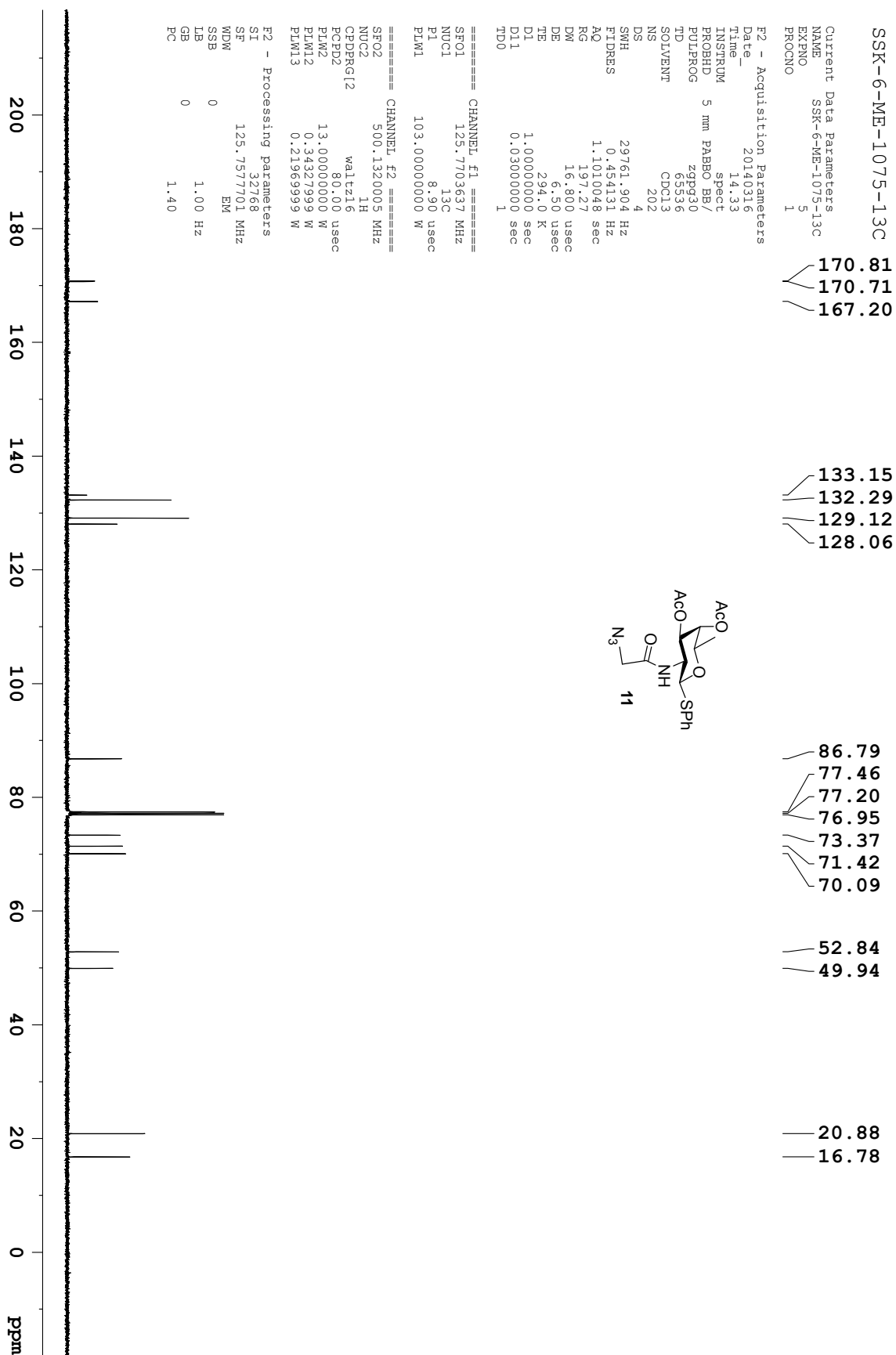
Current Data Parameters
 NAME SSK-6-ME-1075-1H
 EXPNO 4
 PROCNO 1

F2 - Acquisition Parameters
 Date_ 20140316
 Time_ 14.30
 INSTRUM spect
 PROBHID 5 mm PABBO BB/
 PULPROG zg30
 TD 65536
 SOLVENT CDCl3
 NS 16
 DS 2
 SWH 10000.000 Hz
 FIDRES 0.152388 Hz
 AQ 3.2767999 sec
 RG 30.72
 DM 50.000 usec
 DE 6.50 usec
 TE 293.5 K
 D1 1.00000000 sec
 TD0 1

==== CHANNEL f1 =====
 SF01 500.1330885 MHz
 NUC1 1H
 P1 13.00 usec
 PLM1 13.00000000 W

F2 - Processing parameters
 SI 65536
 SF 500.1300000 MHz
 WDM EM
 SSB 0
 LB 0.30 Hz
 GB 0
 PC 1.00





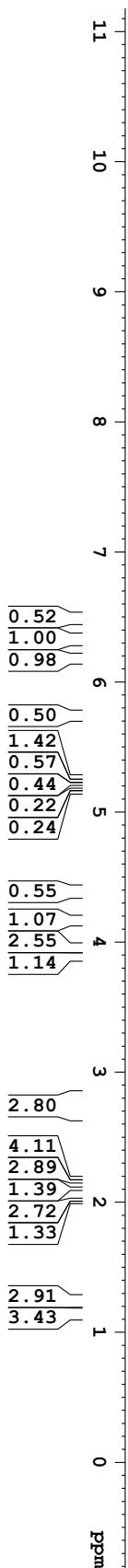
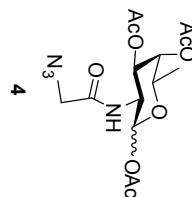
SSK-6-ME-1079-1H

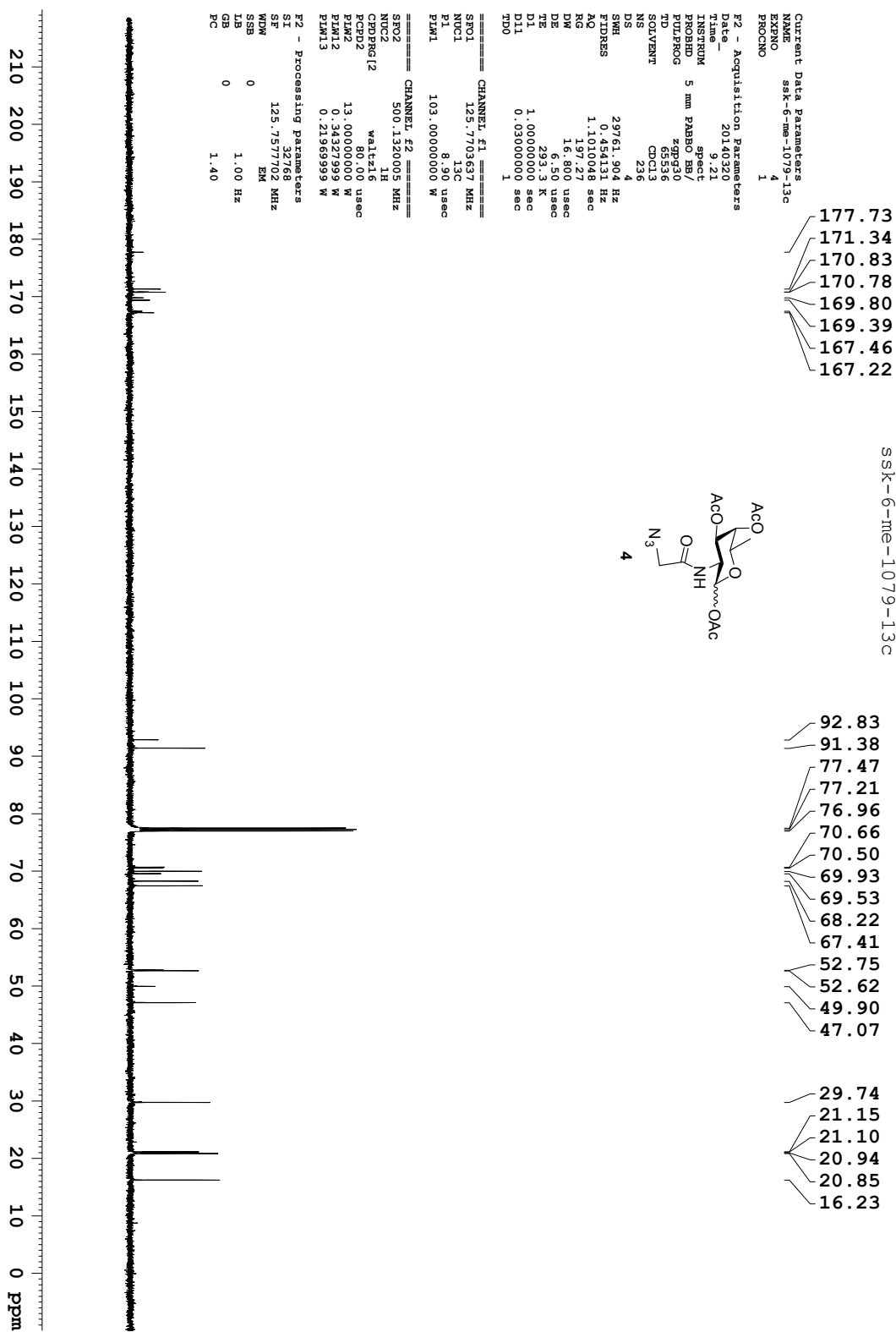
Current Data Parameters
 NAME SSK-6-ME-1079-1H
 EXPNO 2
 PROCNO 1

F2 - Acquisition Parameters
 Date_ 20140320
 Time 9.38
 INSTRUM spect
 PROBHD 5 mm PABBO BB/
 PULPROG zg30
 TD 59998
 SOLVENT CDCl3
 NS 42
 DS 0
 SWH 10000.000 Hz
 FIDRES 0.166672 Hz
 AQ 2.9999001 sec
 RG 61.42
 DW 50.000 usec
 DE 6.50 usec
 TE 292.4 K
 D1 1.00000000 sec
 TD0 1

==== CHANNEL f1 =====
 SFO1 500.1330885 MHz
 NUC1 1H
 P1 13.00 usec
 PLW1 13.00000000 W

F2 - Processing parameters
 SI 65536
 SF 500.1300124 MHz
 MDW EM
 SSB 0
 LB 0.20 Hz
 GB 0
 PC 1.00





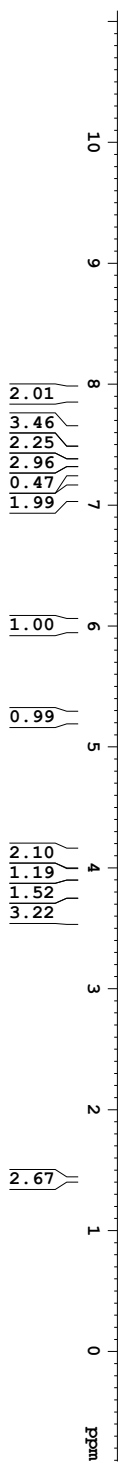
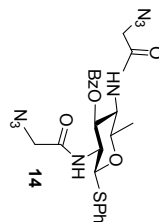
ssk-6-me-1069-1h

Current Data Parameters
 NAME ssk-6-me-1069-1h
 EXPNO 2
 PROCNO 1

F2 - Acquisition Parameters
 Date_ 20140310
 Time 20.00
 INSTRUM spect
 PROBHD 5 mm PABBO BB/
 PULPROG zg30
 TD 65536
 SOLVENT CDCl3
 NS 7
 DS 2
 SWH 10000.000 Hz
 FIDRES 0.152588 Hz
 AQ 3.276789 sec
 RG 87.91
 DW 50.000 usec
 DE 6.50 usec
 TE 296.5 K
 DI 1.0000000 sec
 TD0 1

==== CHANNEL f1 =====
 SF01 500.1330885 MHz
 NUCL1 1H
 P1 13.00 usec
 PLW1 13.00000000 W

F2 - Processing parameters
 SI 65536
 SF 500.1300000 MHz
 WDW EM
 SSB 0
 TB 0.30 Hz
 GB 0
 PC 1.00



SSK-6-ME-1069-13C

Current Data Parameters
 NAME SSK-6-ME-1069-13C
 EXPNO 8
 PROCNO 1

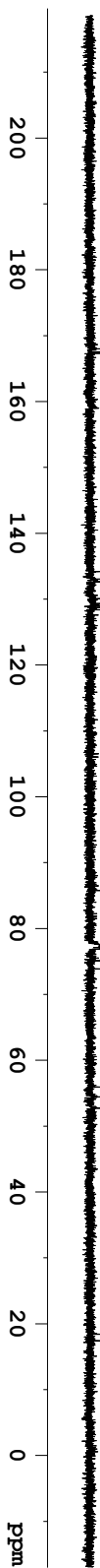
F2 - Acquisition Parameters
 Date_ 20140310
 Time 21.56

INSTRUM 5 mm PABBO BB/
 PROBHD zgpg30
 PULPROG zgpg30
 TD 65536
 SOLVENT CDC13
 NS 93
 DS 4
 SWH 29761.904 Hz
 FIDRES 0.454131 Hz
 AQ 1.1010048 sec
 RG 197.27
 DW 16.800 usec
 DE 6.50 usec
 TE 297.4 K
 D1 1.00000000 sec
 D11 0.03000000 sec
 TDO 1

===== CHANNEL f1 =====
 SF01 125.7703637 MHz
 NUC1 13C
 P1 8.90 usec
 PL1 103.00000000 W

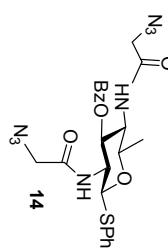
===== CHANNEL f2 =====
 SF02 500.1320005 MHz
 NUC2 1H
 CPDPRG2 waltz16
 PCPD2 80.00 usec
 PLM2 13.00000000 W
 PLM12 0.34327999 W
 PLM13 0.21969999 W

F2 - Processing parameters
 SI 32768
 SF 125.757692 MHz
 MWDW EM
 SSB 0
 LB 1.00 Hz
 GB 0
 PC 1.40



167.93
 167.50
 167.26

134.12
 132.99
 132.60
 130.01
 129.25
 128.95
 128.69
 128.33



85.73
 77.46
 77.20
 76.95
 75.08
 73.84

55.90
 54.41
 52.72
 52.66

18.38

ssk-6-me-1080-1h

Current Data Parameters
 NAME ssk-6-me-1080-1h
 EXPNO 4
 PROCNO 1

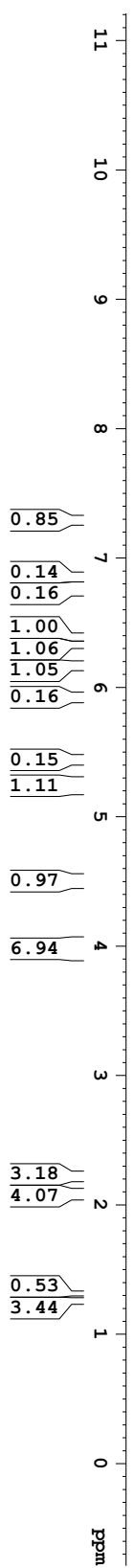
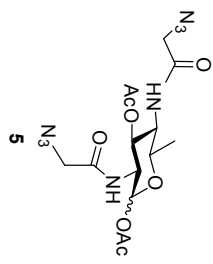
F2 - Acquisition Parameters
 Date_ 20140320
 Time 9.31

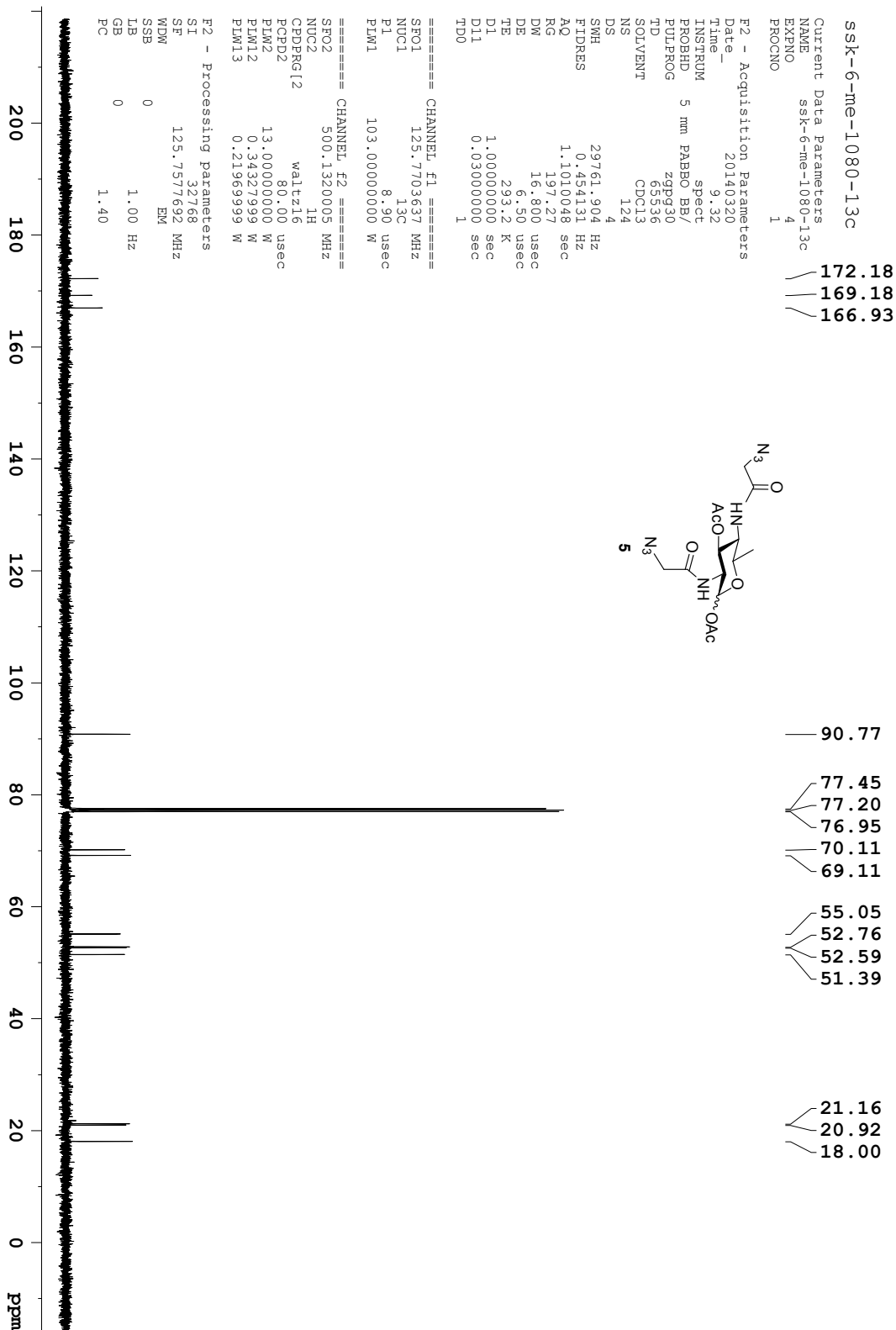
INSTRUM spect
 PROBHD 5 mm PABBO BB/
 PULPROG zg30
 TD 65536
 SOLVENT CDCl3

DS 8
 SMH 10000.000 Hz
 FIDRES 0.152588 Hz
 AQ 3.2767999 sec
 RG 119.07
 DW 50.000 usec
 DE 6.50 usec
 TE 292.6 K
 DI 1.00000000 sec
 TDO 1

==== CHANNEL f1 =====
 SFO1 500.1330885 MHz
 NUC1 1H
 P1 13.00 usec
 PLW1 13.00000000 W

F2 - Processing parameters
 SI 65536
 SF 500.130000 MHz
 WDW EM
 SSB 0
 LB 0.30 Hz
 GB 0
 PC 1.00





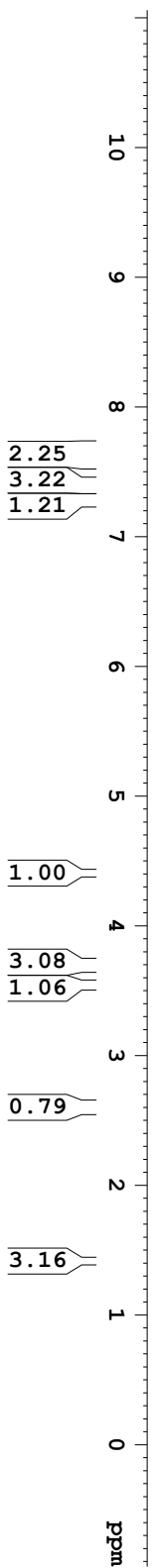
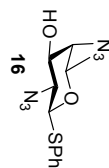
SSK-6-ME-474-3OH-1H

Current Data Parameters
 NAME SSK-6-ME-474-3OH-1H
 EXPNO 1
 PROCNO 1

F2 - Acquisition Parameters
 Date_ 20140708
 Time 20.30
 INSTRUM spect
 PROBD 5 mm PABO BB/
 PULPROG zg30
 TD 65536
 SOLVENT CDCl3
 NS 29
 DS 2
 SWH 10000.000 Hz
 FIDRES 0.152588 Hz
 AQ 3.2767999 sec
 RG 177.33
 DW 50.000 usec
 DE 6.50 usec
 TE 296.9 K
 D1 1.00000000 sec
 TD0 1

==== CHANNEL f1 =====
 SFO1 500.1330885 MHz
 NUC1 1H
 P1 13.00 usec
 PLW1 13.00000000 W

F2 - Processing parameters
 SI 65536
 SF 500.1300000 MHz
 WDW EM
 SSB 0
 LB 0.30 Hz
 GB 0
 PC 1.00



SSK-6-ME-474-3OH-13C

Current Data Parameters
 NAME SSK-6-ME-474-3OH-13C
 EXPNO 1
 PROCNO 1

F2 - Acquisition Parameters

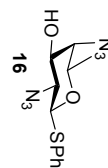
Date_ 20140708
 Time_ 20.46
 INSTRUM spect
 PROBHID 5 mm PABBO BB/
 PULPROG zgpg30
 TD 65536
 SOLVENT CDCl3
 NS 545
 DS 4
 SMH 29761.904 Hz
 FIDRES 0.454131 Hz
 AQ 1.1010048 sec
 RG 197.27
 DW 16.800 usec
 DE 6.50 usec
 TE 297.2 K
 D1 1.0000000 sec
 D11 0.0300000 sec
 TD0 1

==== CHANNEL F1 =====
 SFO1 125.7703637 MHz
 NU01 13C
 P1 8.90 usec
 PLW1 103.0000000 W

==== CHANNEL F2 =====
 SFO2 500.1320005 MHz
 NU02 1H
 CPDPRG12 waltz16
 PCPD2 80.00 usec
 PLW2 13.0000000 W
 PLW12 0.34327999 W
 PLW13 0.21969999 W

F2 - Processing parameters
 SI 32768
 SF 125.7577673 MHz
 WDW EM
 SSB 0
 LB 1.00 Hz
 GB 0
 PC 1.40

133.27
 131.68
 129.21
 128.48



86.88
 77.45
 77.20
 76.94
 74.51
 74.11
 65.08
 62.81

18.01



SSK-6-ME-1150-1H

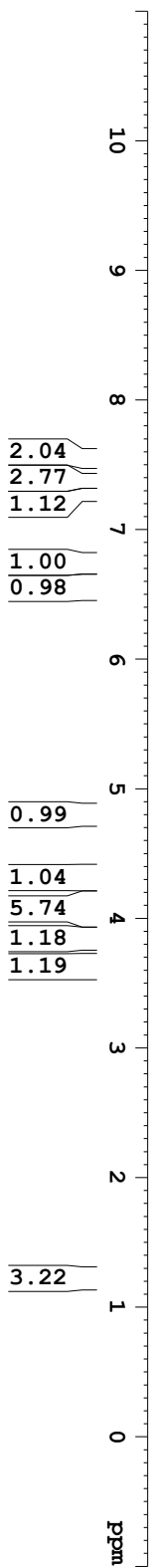
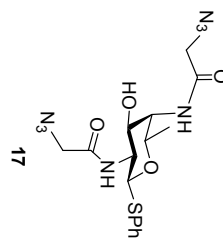
Current Data Parameters
 NAME SSK-6-ME-1150-1H
 EXPNO 5
 PROCNO 1

F2 - Acquisition Parameters

Date_ 20140714
 Time_ 21.42
 INSTRUM spect
 PROBHD 5 mm PABBO BB/
 PULPROG zg30
 TD 65536
 SOLVENT CDCl3
 NS 15
 DS 2
 SWH 10000.000 Hz
 FIDRES 0.152588 Hz
 AQ 3.2767999 sec
 RG 119.07
 DW 50.000 usec
 DE 6.50 usec
 TE 297.1 K
 D1 1.00000000 sec
 TDO 1

===== CHANNEL f1 =====
 SFO1 500.1350885 MHz
 NUC1 1H
 P1 13.00 usec
 PLW1 13.00000000 W

F2 - Processing parameters
 SI 65536
 SF 500.1300103 MHz
 WDW EM
 SSB 0
 LB 0.30 Hz
 GB 0
 PC 1.00



SSK-6-ME-1150-13C

Current Data Parameters
 NAME SSK-6-ME-1150-13C
 EXPNO 2
 PROCNO 1

169.43
 168.73

132.88
 132.35
 129.14
 128.27

86.48
 77.45
 77.20
 76.95
 73.90
 72.21

53.59
 52.64
 52.49
 52.35

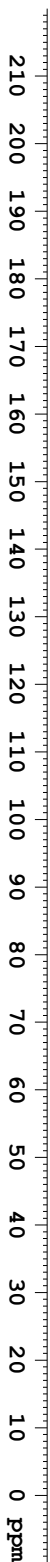
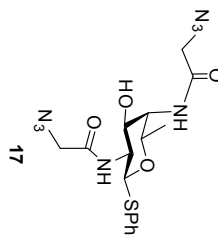
17.24

F2 - Acquisition Parameters
 Date_ 20140701
 Time 12.48
 INSTRUM spect
 PROBHD 5 mm PABBO BB/
 PULPROG zgpg30
 TD 65936
 SOLVENT CDCl3
 NS 227
 DS 4
 SWH 29761.904 Hz
 FIDRES 0.454131 Hz
 AQ 1.1010048 sec
 RG 197.27
 DW 16.800 usec
 DE 6.50 usec
 TE 297.2 K
 D1 1.00000000 sec
 D11 0.03000000 sec
 TD0 1

==== CHANNEL f1 =====
 SFO1 125.7703637 MHz
 NUC1 13C
 P1 8.90 usec
 PLW1 103.00000000 W

==== CHANNEL f2 =====
 SFO2 500.1320005 MHz
 NUC2 1H
 CPDPRG(2) waltz16
 PCPD2 80.00 usec
 PLM2 13.00000000 W
 PLM12 0.34327999 W
 PLM13 0.21969999 W

F2 - Processing parameters
 SI 32768
 SF 125.757719 MHz
 WDW EM
 SSB 0
 LB 1.00 Hz
 GB 0
 PC 1.40



SSK-6-ME-1151-1H

Current Data Parameters
 NAME SSK-6-ME-1151-1H
 EXPNO 3
 PROCNO 1

F2 - Acquisition Parameters
 Date_ 20140704
 Time 2.30
 INSTRUM spect
 PROBHD 5 mm SEI 1H/D-
 PULPROG zg30
 TD 65536
 SOLVENT CDCl3
 NS 16
 DS 0
 SWH 8223.685 Hz
 FIDRES 0.125483 Hz
 AQ 3.9845889 sec
 RG 114
 DM 60.800 usec
 DE 6.50 usec
 TE 295.5 K
 D1 1.00000000 sec
 TD0 1

===== CHANNEL f1 =====
 NUCL 1H
 P1 13.50 usec
 PL1 -1.00 dB
 PL1W 10.56200695 W
 SFO1 400.1324710 MHz
 F2 - Processing parameters
 SI 32768
 SF 400.1298146 MHz
 WDW EM
 SSB 0
 LB 0.30 Hz
 GB 0
 PC 1.00

