CHEMISTRY
A EUROPEAN JOURNAL

Supporting Information

Effect of β-*O*-Glucosylation on L-Ser and L-Thr Diamides: A Bias toward α-Helical Conformations

Francisco Corzana, Jesús H. Busto, Søren B. Engelsen, Jesús Jiménez-Barbero, Juan Luis Asensio, Jesús M. Peregrina and Alberto Avenoza*

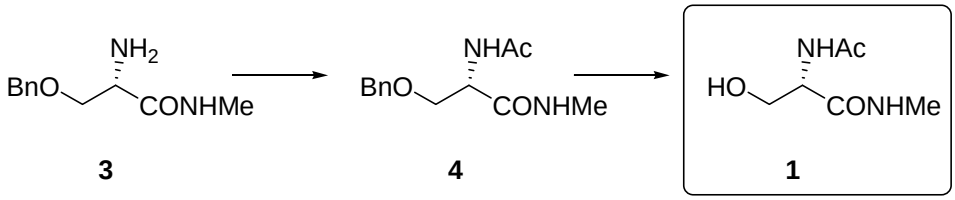
CONTENTS:

Abbreviations. NOE: Nuclear Overhauser Effect, NOESY: Nuclear Overhauser Effect Spectroscopy, ROESY: Rotating frame Overhauser Enhancement Spectroscopy, MD-tar: molecular dynamics with time-averaged restraints, ff99': modified ff99 force field[1]

Experimental procedures for new compounds and a full listing of ^{1}H and ^{13}C NMR data.

General Procedures. Solvents were purified according to standard procedures. Analytical TLC was performed using Polychrom SI F254 plates. Column chromatography was performed using Silica gel 60 (230–400 mesh). ^{1}H and ^{13}C NMR spectra were recorded on a Bruker ARX 300 and Bruker Avance 400 spectrometer. ^{1}H and ^{13}C NMR spectra were recorded in $CDCl_3$ with TMS as the internal standard and in D_2O with TMS as the external standard using a coaxial microtube (chemical shifts are reported in ppm on the δ scale, coupling constants in Hz). Melting points were determined on a Büchi B-545 melting point apparatus and are uncorrected. Optical rotations were measured on a Perkin-Elmer 341 polarimeter. Microanalyses were carried out on a CE Instruments EA-1110 analyser and are in good agreement with the calculated values.

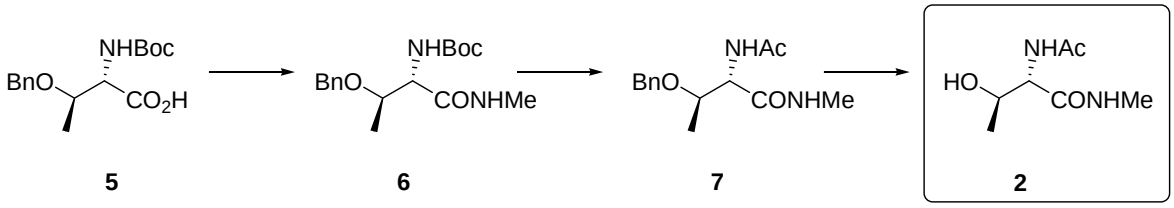
Synthesis of Ser diamide 1



Ac-L-Ser(OBn)-NHMe (4). To a suspension of L-Ser(OBn)-NHMe hydrochloride (**3**) (2.12 g, 8.7 mmol) in CH_2Cl_2 (100 mL), at 0 ºC under an inert atmosphere, triethylamine (TEA) (2.4 mL, 15.9 mmol) and acetyl chloride (1.1 mL, 13.7 mmol) were added. The mixture was allowed to warm up to rt and was stirred for 12 h. The reaction mixture was washed with saturated $NaHCO_3$ (2 × 15 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered and evaporated to give a residue that was purified by silica gel column chromatography eluting with methanol/ethyl acetate 1:9 to yield 1.51 g (70%) of Ac-L-Ser(OBn)-NHMe (**4**) as a white solid. Mp: 137-139 ºC. $[\alpha]^{24.8}_{D}$ (c = 1.1, CH_3OH): +10.7. Anal. calcd. for $C_{13}H_{18}N_2O_3$: C, 62.38; H, 7.25; N, 11.19. Found: C, 62.42; H, 7.32; N, 11.02. **^{1}H NMR (400 MHz, CD_3OD)** *δ*: 2.00 (s, 3H), 2.74 (d, 3H, *J*=9.2 Hz), 3.66 (dd, 1H, *J*=9.6 Hz, *J*=4.8 Hz), 3.74 (dd, 1H, *J*=5.6 Hz, *J*=9.6 Hz), 4.46-4.57 (m, 3H), 7.23-7.34 (m, 5H). **^{13}C NMR (100 MHz, CD_3OD)** *δ*: 22.5, 26.4, 55.0, 70.7, 74.1, 128.7, 128.8, 129.4, 139.2, 172.8, 173.4.

Ac-L-Ser-NHMe (1). A solution of Ac-L-Ser(OBn)-NHMe (**4**) (1.51 g, 6.0 mmol) in methanol (20 mL) was hydrogenated, using 30 mg of 10% palladium-carbon as a catalyst, at rt for 16 h. The catalyst and solvent were removed, and the residue was purified by silica gel column chromatography, eluting with methanol/ethyl acetate (15:85) to give 978 mg (95%) of Ac-L-Ser-NHMe (**1**) as a white solid. Mp: 113-115 ºC. $[\alpha]^{23.5}_{D}$ (c = 1.16, CH_3OH): -15.3. Anal. calcd. for $C_6H_{12}N_2O_3$: C, 44.99; H, 7.55; N, 17.49. Found: C, 44.89; H, 7.49; N, 17.52. **^{1}H NMR (400 MHz, CD_3OD)** δ: 2.02 (s, 3H), 2.74 (s, 3H), 3.72-3.80 (m, 2H), 4.36-4.38 (m, 1H). **^{13}C NMR (100 MHz, CD_3OD)** δ: 22.7, 26.4, 57.0, 63.0, 173.1, 173.5.

Synthesis of Thr diamide 2



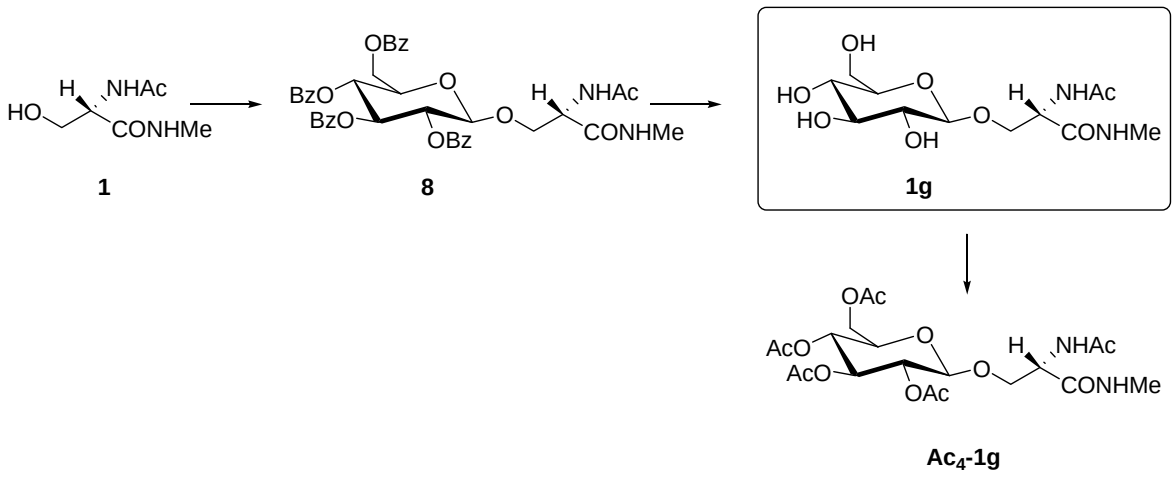
Boc-L-Thr(OBn)-NHMe (6). A solution of Boc-L-Thr(OBn)-OH (**5**) (1.00 g, 3.23 mmol) in acetonitrile (30 mL) was treated with diisopropilethylamine (DIEA) (2.24 mL, 12.93 mmol), methylamine hydrochloride (427 mg, 6.46 mmol) and benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (TBTU) (1.21 g, 3.88 mmol) under an inert atmosphere. The reaction mixture was stirred at rt for 10 h, then partitioned between brine (20 mL) and ethyl acetate (12 mL). The organic layer was washed with 0.1 N HCl (2 × 15 mL) and 5% $NaHCO_3$ (2 × 10 mL), dried over anhydrous Na_2SO_4, filtered and evaporated to give a residue that was purified by silica gel column chromatography, eluting with hexane/ethyl acetate (2:8) to yield 1.14 g (98%) of Boc-L-Thr(OBn)-NHMe (**6**) as a white solid: Mp: 118-120 ºC. $[\alpha]^{29.4}_{D}$ (c = 1.15, CH_3OH): +16.3. Anal. calcd. for $C_{17}H_{26}N_2O_4$: C, 63.33; H, 8.13; N, 8.69. Found: C, 63.45; H, 8.19; N, 8.58. **^{1}H NMR (400 MHz, $CDCl_3$)** δ: 1.16 (d, 3H, *J*=6.1 Hz), 1.45 (s, 9H), 2.82 (d, 3H, *J*=4.8 Hz), 4.19-4.25 (m, 2H), 4.53-4.62 (m, 2H), 5.50 (d, 1H, *J*=5.8 Hz), 6.50 (s, 1H), 7.27-7.37 (m, 5H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ: 15.6, 26.2, 28.3, 57.7, 71.6, 74.8, 80.1, 127.7, 127.8, 128.4, 138.0, 155.8, 170.4.

Ac-L-Thr(OBn)-NHMe (7). Boc-L-Thr(OBn)-NHMe (**6**) (924 mg, 2.87 mmol) was dissolved in 50 mL of 2 N HCl/THF (3:7) and the solution was stirred at 20 ºC for 24 h. The solvent was evaporated in vacuo to give 535 mg (72%) of L-Thr(OBn)-NHMe·HCl as a white solid and was used without any purification. To a suspension of L-Thr(OBn)-NHMe·HCl (189 mg, 0.73 mmol) in CH_2Cl_2 (15 mL), at 0 ºC under an inert atmosphere, TEA (0.2 mL, 1.3 mmol) and acetyl chloride (0.09 mL, 1.12 mmol) were added. The mixture was allowed to warm up to rt and was stirred for 24

h. The reaction mixture was washed with water (2 × 5 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered and evaporated to give a residue that was purified by silica gel column chromatography, eluting with dichloromethane/methanol (9:1) to yield 128 mg (57%) of Ac-L-Thr(OBn)-NHMe (**7**) as a white solid. Mp: 116-118 ºC. $[\alpha]^{29.4}_D$ (c = 1.30, CH_3OH): +10.6. Anal. calcd. for $C_{14}H_{20}N_2O_3$: C, 63.62; H, 7.63; N, 10.60. Found: C, 63.74; H, 7.55; N, 10.71. **^{1}H NMR (400 MHz, $CDCl_3$)** δ: 1.11 (d, 3H, *J*=6.4 Hz), 2.04 (s, 3H), 2.81 (d, 3H, *J*=4.9 Hz), 4.09-4.12 (m, 1H), 4.54-4.56 (m, 1H), 4.60-4.68 (dd, 2H, *J*=11.6 Hz, *J*=19.7 Hz), 6.50-6.60 (m, 1H), 6.64 (d, 1H, *J*=6.2 Hz), 7.31-7.38 (m, 5H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ: 15.2, 23.2, 26.2, 55.9, 71.5, 74.0, 127.8, 128.4, 129.6, 137.8, 169.9, 170.3.

Ac-L-Thr-NHMe (2). A solution of Ac-L-Thr(OBn)-NHMe (**7**) (306 mg, 1.2 mmol) in methanol (20 mL) was hydrogenated, using 61 mg of 10% palladium-carbon as a catalyst, at rt for 12 h. The catalyst and solvent were removed, and the residue was purified by silica gel column chromatography, eluting with dichloromethane/methanol (85:15) to give 186 mg (92%) of Ac-L-Thr-NHMe (**2**) as a white solid Mp: 156-158 ºC. $[\alpha]^{26.7}_D$ (c = 1.50, CH_3OH): +10.9. Anal. calcd. for $C_7H_{14}N_2O_3$: C, 48.26; H, 8.10; N, 16.08. Found: C, 48.31; H, 8.02; N, 16.06. **^{1}H NMR (400 MHz, H_2O/D_2O 9:1)** δ: 1.16 (d, 3H, *J*=8.0 Hz), 2.06 (s, 3H), 2.72 (d, 3H, *J*=4.0 Hz), 4.17-4.24 (m, 2H), 8.00 (s, 1H), 8.16 (d, 1H, J=7.4 Hz). **^{13}C NMR (100 MHz, $CDCl_3$)** δ: 18.0, 23.1, 26.1, 56.5, 66.3, 171.4, 171.9.

Synthesis of β-glucosylated Ser diamide 1g



β-D-Bz$_4$Glc-Ac-L-Ser-NHMe (8). Silver triflate (975 mg, 3.60 mmol) was added to a suspension of Ac-L-Ser-NHMe (**1**) (350 mg, 2.18 mmol) and powdered molecular sieves (4 Å, 1 g) in dichloromethane (5 mL) under an inert atmosphere. The mixture was stirred at -30 ºC and then 2,3,4,6-tetra-*O*-benzoyl-α-D-glucopyranosyl bromide (2 g, 3.03 mmol) in dichloromethane (2 mL)

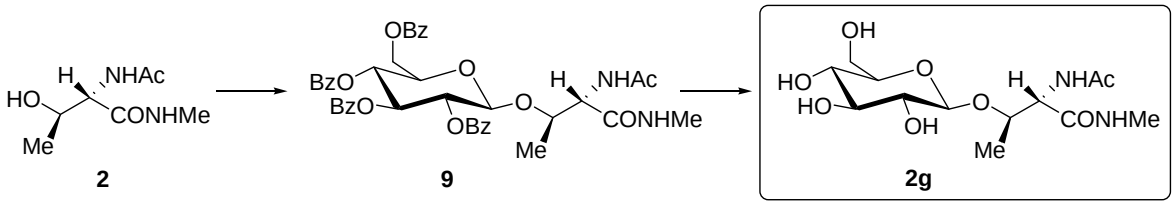
was added. The mixture was stirred at this temperature for 1 h and then was warmed at rt and stirred for another 14 h. The crude was filtered, concentrated and purified by silica gel column chromatography, eluting with dicloromethane/methanol (95:5) to yield 821 mg (51%) of β-D-Bz_4Glc-Ac-L-Ser-NHMe (**8**) as a white solid. Mp: 109-111 ºC. $[\alpha]^{26.5}_D$ (c = 1.35, $CDCl_3$): +62.1. Anal. calcd. for $C_{40}H_{38}N_2O_{12}$: C, 65.03; H, 5.18; N, 3.79. Found: C, 64.98; H, 5.13; N, 3.83. **^{1}H NMR (400 MHz, $CDCl_3$)** δ: 1.94 (s, 3H), 2.71 (d, 3H, *J*=4.8 Hz), 3.72 (dd, 1H, *J*=8.5 Hz, *J*=10.9 Hz), 4.15 (dd, 1H, *J*=3.9 Hz, *J*=10.9 Hz), 4.23 (ddd, 1H, *J*=2.7 Hz, *J*=4.5 Hz, *J*=9.9 Hz), 4.43 (dd, 1H, *J*=4.7 Hz, *J*=12.3 Hz), 4.66 (dt, 1H, *J*=3.9 Hz, *J*=8.1 Hz), 4.82 (dd, 1H, *J*=2.5 Hz, *J*=12.3 Hz), 5.03 (d, 1H, *J*=8.0 Hz), 5.50 (dd, 1H, *J*=8.1 Hz, *J*=9.8 Hz), 5.73 ('t', 1H, *J*=9.8 Hz), 5.94 ('t', 1H, *J*=9.7 Hz), 6.49-6.54 (m, 2H), 7.22-7.30 (m, 2H), 7.33-7.60 (m, 10H), 7.78-7.84 (m, 2H), 7.90-7.98 (m, 4H), 8.05-8.10 (m, 2H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ: 23.0, 26.4, 52.1, 62.1, 69.0, 70.2, 71.7, 72.6, 102.3, 128.3, 128.5, 128.6, 128.6, 128.7, 128.9, 129.3, 129.7, 129.8, 129.9, 133.3, 133.4, 133.5, 133.6, 165.1, 165.3, 165.6, 166.3, 169.9, 170.2.

β-D-Glc-Ac-L-Ser-NHMe (1g). A solution of β-D-Bz_4Glc-Ac-L-Ser-NHMe (**8**) (355 mg, 0.48 mmol) in methanol (10 mL) was treated with sodium methoxide/methanol (0.5 M) to pH=9. After stirring for 3 h, the mixture was neutralized with Dowex 50-X8, filtered and concentrated. Purification of the residue with C18 reverse-phase sep-pak cartridge gave 141 mg (91%) of β-D-Glc-Ac-L-Ser–NHMe (**1g**). $[\alpha]^{26.5}_D$ (c = 0.53, CH_3OH): +2.7. Anal. calcd. for $C_{12}H_{22}N_2O_8$: C, 44.72; H, 6.88; N, 8.69. Found: C, 44.64; H, 6.76; N, 8.56. **^{1}H NMR (400 MHz, D_2O)** δ: 2.04 (s, 3H), 2.71 (s, 3H), 3.25 (dd, 1H, *J*=8.1Hz, *J*=9.2Hz), 3.31-3.36 (m, 1H), 3.39-3.48 (m, 2H), 3.68 (dd, 1H, *J*=5.9 Hz, *J*=12.3 Hz), 3.81-3.90 (m, 2H), 4.19 (dd, 1H, *J*=5.2 Hz, *J*=10.6 Hz), 4.42 (d, 1H, *J*=7.9 Hz), 4.48 ('t', 1H, *J*=4.6 Hz). **^{13}C NMR (100 MHz, D_2O)** δ: 21.8, 26.0, 53.9, 60.7, 68.7, 69.5, 73.0, 75.6, 75.9, 102.2, 171.8, 174.5.

β-D-Ac_4-Glc-Ac-L-Ser-NHMe (Ac_4-1g). β-D-Glc-Ac-L-Ser–NHMe (**1g**) (15 mg, 0.05 mmol) was dissolved in pyridine/acetic anhydride (2:1, 6 mL) and stirred for 3 h. Removal of the volatiles and column chromatographic purification (ethyl acetate/methanol, 95:5) gave a residue (18 mg, 81%) corresponding to β-D-Ac_4-Glc-Ac-L-Ser-NHMe (**Ac_4-1g**). $[\alpha]^{25.4}_D$ (c = 0.96, $CDCl_3$): +14.5. Anal. calcd. for $C_{20}H_{30}N_2O_{12}$: C, 48.98; H, 6.17; N, 5.71. Found: C, 49.08; H, 6.05; N, 5.79. **^{1}H NMR (400 MHz, D_2O)** δ: 1.91 (s, 3H), 1.92 (s, 3H), 1.95 (s, 3H), 1.97 (s, 3H), 2.00 (s, 3H), 2.60 (s, 3H), 3.81 (dd, 1H, *J*=5.3 Hz, *J*=10.8 Hz), 3.89-3.96 (m, 2H), 4.09 (dd, 1H, *J*=1.5 Hz, *J*=12.8 Hz), 4.29 (dd, 1H, *J*=3.6 Hz, *J*=12.9 Hz), 4.34 ('t', 1H, *J*=5.9 Hz), 4.74 (d, 1H, *J*=8.0 Hz), 4.83 ('t', 1H, *J*=8.7 Hz), 4.99 ('t', 1H, *J*=9.7 Hz), 5.23 ('t', 1H, *J*=9.4 Hz). **^{13}C NMR (100 MHz, D_2O)** δ: 20.0, 20.1, 20.1, 21.7, 25.9, 53.7, 61.6, 68.0, 68.3, 71.1, 71.3, 72.9, 99.8, 171.3, 172.5, 172.7, 173.1,

173.6, 174.3. **^{1}H NMR (400 MHz, $CDCl_3$)** δ: 1.99 (s, 3H), 2.01 (s, 3H), 2.02 (s, 3H), 2.03 (s, 3H), 2.08 (s, 3H), 2.80 (d, 1H, *J*=4.8 Hz), 3.70 (dd, 1H, *J*=8.9 Hz, *J*=10.5 Hz), 3.75 (dd, 1H, *J*=3.4 Hz, *J*=10.0 Hz), 4.03 (dd, 1H, *J*=4.7 Hz, *J*=10.7 Hz), 4.21-4.24 (m, 2H), 4.55-4.60 (m, 1H), 4.62 (d, 1H, *J*=8.1 Hz), 4.97 (dd, 1H, *J*=8.2 Hz, *J*=9.6 Hz), 5.06 ('t', 1H, *J*=9.7 Hz), 5.19 ('t', 1H, *J*=9.5 Hz), 6.45-6.51 (m, 2H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ: 20.5, 20.6, 20.7, 23.1, 26.4, 52.1, 61.4, 68.0, 69.8, 71.1, 72.0, 72.4, 101.6, 169.4, 169.5, 169.9, 170.0, 170.2, 170.7.

Synthesis of β-glucosylated Thr diamide 2g



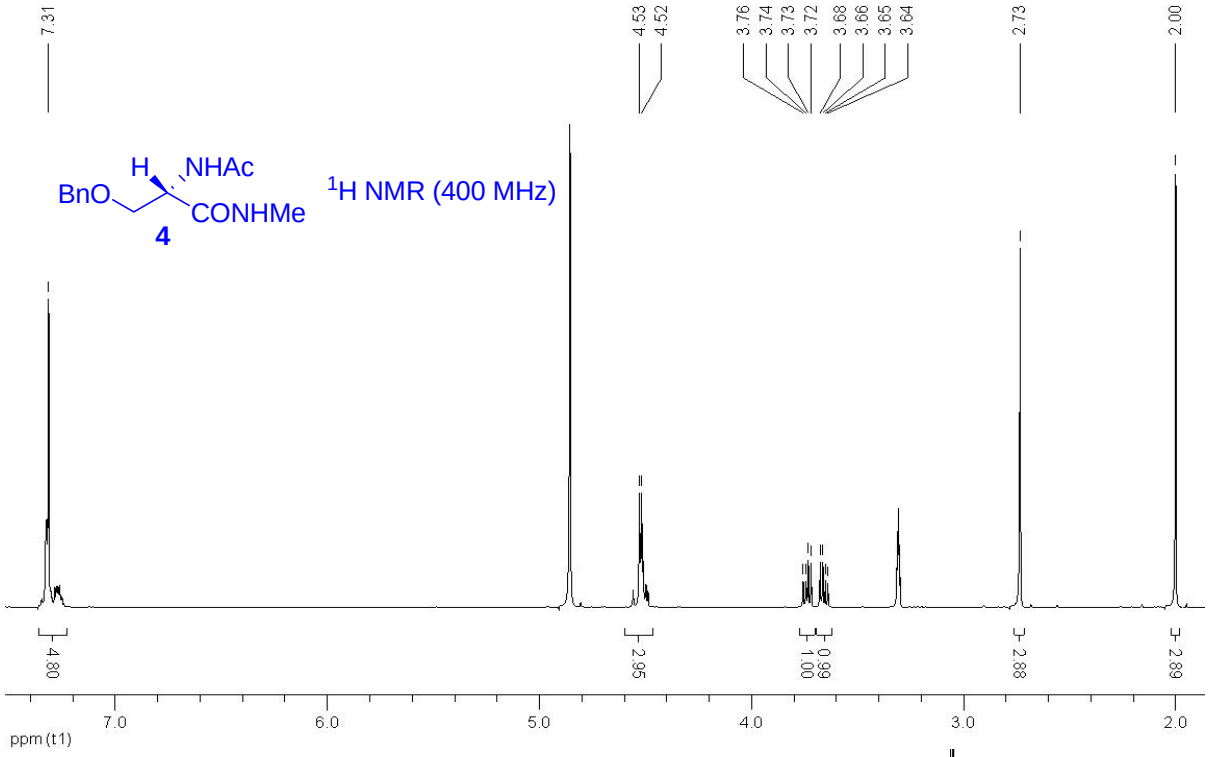
β-D-Bz$_4$-Glc-Ac-L-Thr-NHMe (9). Silver triflate (290 mg, 1.13 mmol) was added to a suspension of Ac-L-Thr-NHMe (**2**) (120 mg, 0.68 mmol) and powdered molecular sieves (4 Å, 1 g) in dichloromethane (4 mL) under an inert atmosphere. The mixture was stirred at -30 ºC and then 2,3,4,6-tetra-*O*-benzoyl-α-D-glucopyranosyl bromide (632 mg, 0.96 mmol) in dichloromethane (2 mL) was added. The mixture was stirred at this temperature for 1 h and was then warmed at rt and stirred for another 14 h. The crude was filtered, concentrated and purified by silica gel column chromatography, eluting with ethyl acetate/hexane (9:1) to yield 254 mg (30%) of β-D-Bz$_4$Glc-Ac-L-Thr-NHMe (**9**) as oil. $[\alpha]^{26.0}_D$ (c = 0.38, $CDCl_3$): +62.6. Anal. calcd. for $C_{41}H_{40}N_2O_{12}$: C, 65.42; H, 5.36; N, 3.72. Found: C, 65.32; H, 5.48; N, 3.66. **^{1}H NMR (300 MHz, $CDCl_3$)** δ: 0.93 (d, 3H, *J*=6.6 Hz), 1.99 (s, 3H), 2.79 (d, 1H, *J*=4.8 Hz), 4.21-4.34 (m, 2H), 4.43 (dd, 1H, *J*=4.9 Hz, *J*=12.3 Hz), 4.68 (dd, 1H, *J*=3.0 Hz, *J*=6.6 Hz), 4.80 (dd, 1H, *J*=2.5 Hz, *J*=12.3 Hz), 5.14 (d, 1H, *J*=8.1 Hz), 5.48 (dd, 1H, *J*=8.1 Hz, *J*=9.9 Hz), 5.72 ('t', 1H, *J*=9.7 Hz), 5.93 ('t', 1H, *J*=9.7 Hz), 6.50-6.60 (m, 2H), 7.24-7.62 (m, 12H), 7.80-7.85 (m, 2H), 7.90-7.98 (m, 4H), 8.02-8.10 (m, 2H). **^{13}C NMR (75 MHz, $CDCl_3$)** δ: 15.9, 23.1, 26.4, 53.4, 55.4, 61.9, 69.0, 71.8, 72.5, 72.7, 101.8, 128.3, 128.4, 128.5, 128.5, 128.6, 128.7, 129.0, 129.3, 129.7, 129.7, 129.8, 129.9, 133.3, 133.4, 133.4, 133.6, 165.2, 165.2, 165.6, 166.1, 168.7, 170.1.

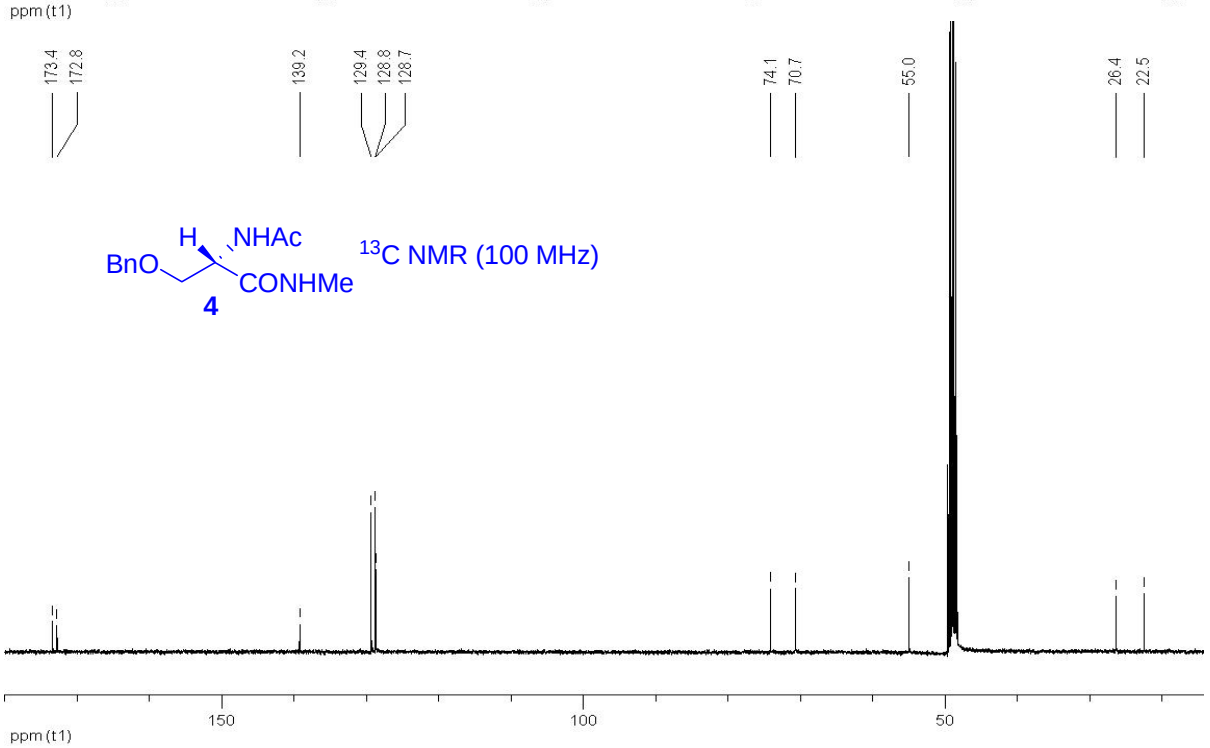
β-D-Glc-Ac-L-Thr–NHMe (2g). A solution of β-D-Bz$_4$Glc-Ac-L-Thr-NHMe (**9**) (254 mg, 0.34 mmol) in methanol (15 mL) was treated with sodium methoxide/methanol (0.5 M) to pH=9. After stirring 3 h, the mixture was neutralized with Dowex 50-X8, filtered and concentrated. Purification of the residue with C18 reverse-phase sep-pak cartridge gave 101 mg (89%) of β-D-Glc-Ac-L-Thr-

NHMe (**2g**). $[\alpha]^{23.4}_{D}$ (c = 1.35, CH_3OH): –2.8. Anal. calcd. for $C_{13}H_{24}N_2O_8$: C, 46.42; H, 7.19; N, 8.33. Found: C, 46.37; H, 7.10; N, 8.42. **^{1}H NMR (400 MHz, D_2O)** δ: 1.23 (d, 3H, *J*=6.3 Hz), 2.10 (s, 3H), 2.74 (s, 3H), 3.23 (dd, 1H, *J*=8.2 Hz, *J*=9.2 Hz), 3.34-3.50 (m, 3H), 3.71 (dd, 1H, *J*=5.4 Hz, *J*=12.3 Hz), 3.88 (dd, 1H, *J*=1.6 Hz, *J*=12.3 Hz), 4.35 (d, 1H, *J*=3.4 Hz), 4.39-4.46 (m, 1H), 4.50 (d, 1H, *J*=7.9 Hz) **^{13}C NMR (100 MHz, D_2O)** δ: 15.7, 21.7, 25.9, 58.4, 60.6, 69.5, 72.9, 73.3, 75.6, 75.7, 99.6, 172.2, 174.9.

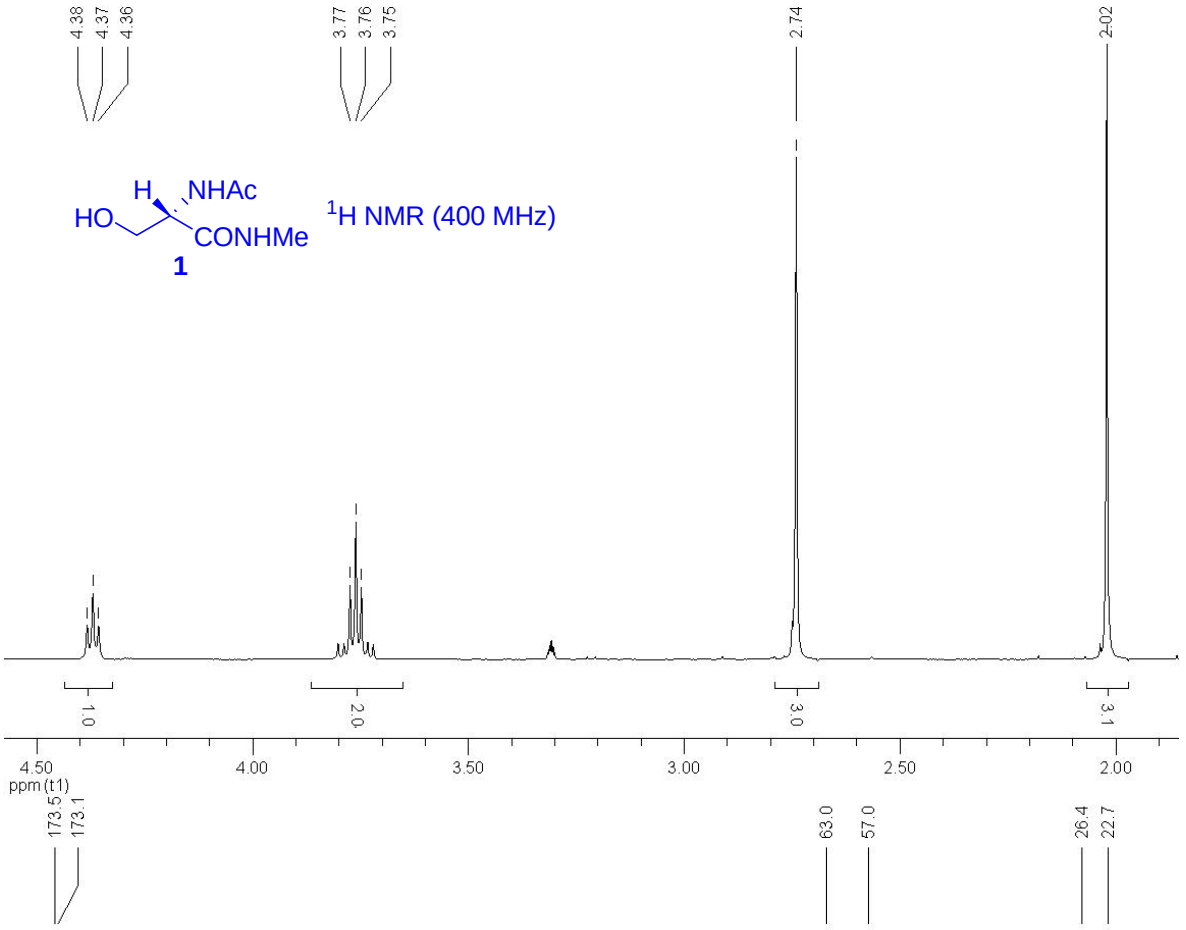
^{1}H and ^{13}C NMR spectra for compounds 1, 2, 1g, Ac_4-1g, 2g, 4, 6, 7, 8 and 9 as well as COSY and HSQC correlations for compounds 1g, Ac_4-1g, 2g and 8.

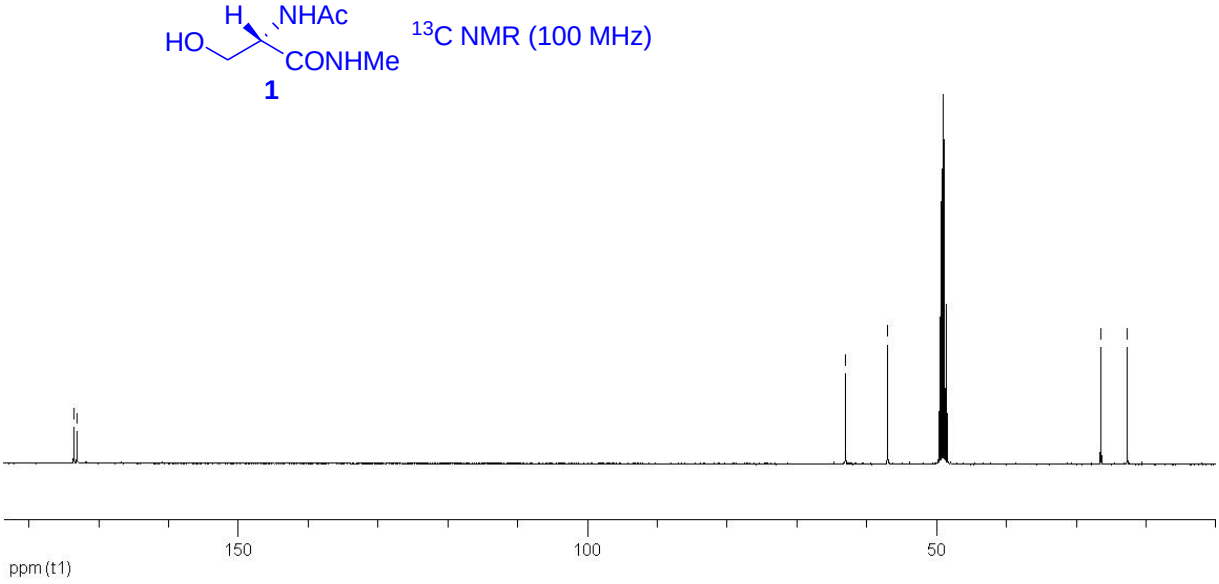
Ac-L-Ser(OBn)-NHMe (4)



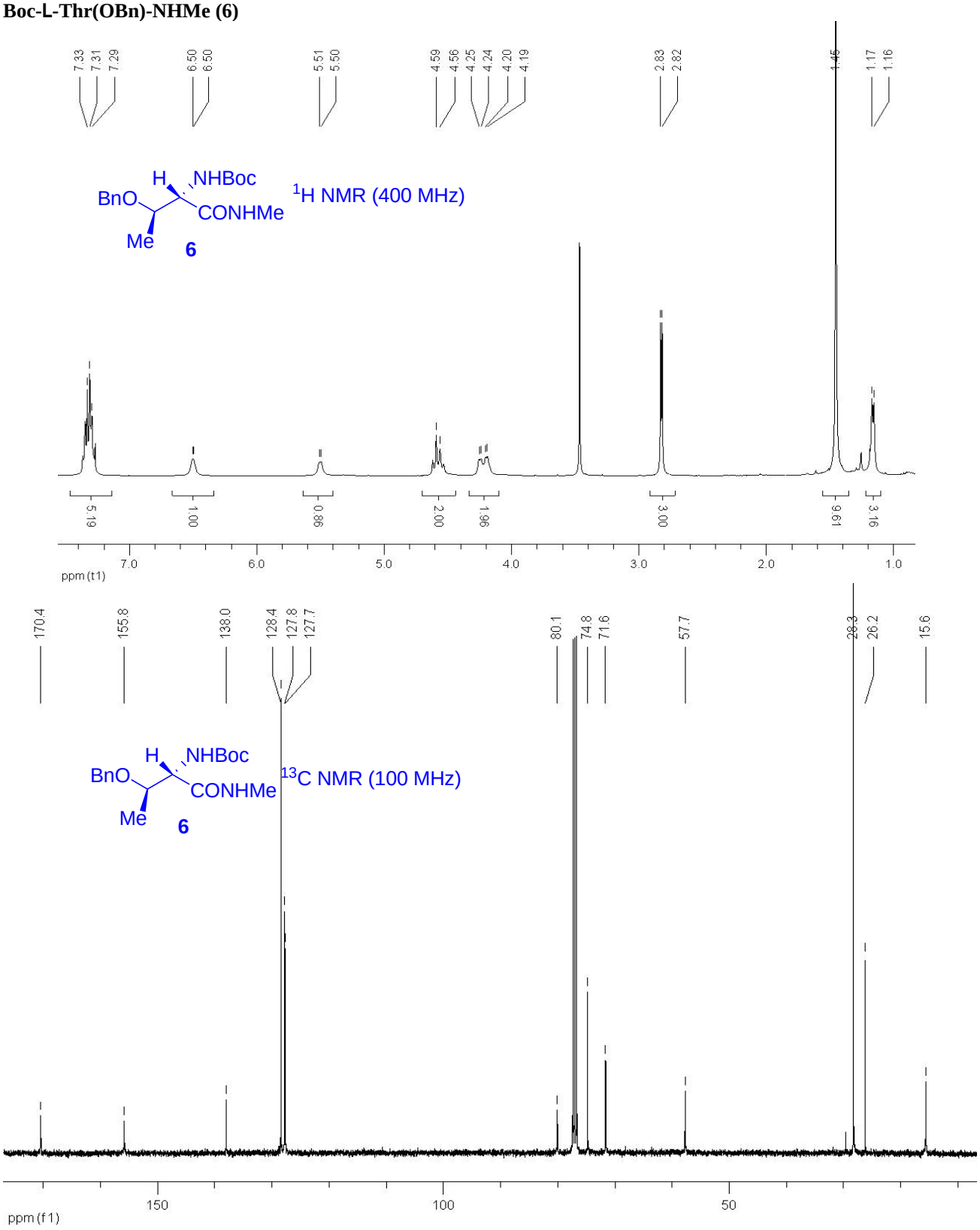


Ac-L-Ser-NHMe (1)

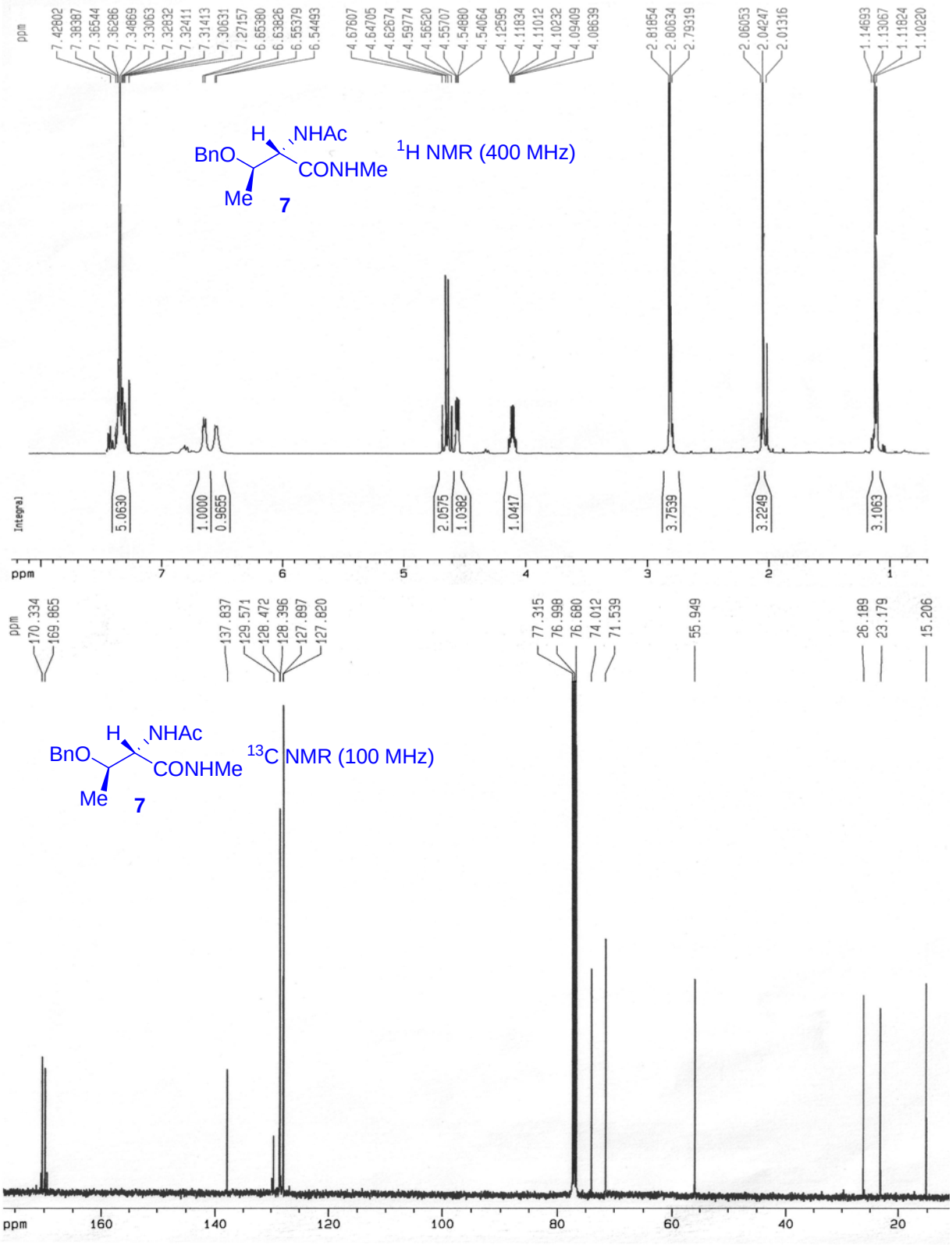




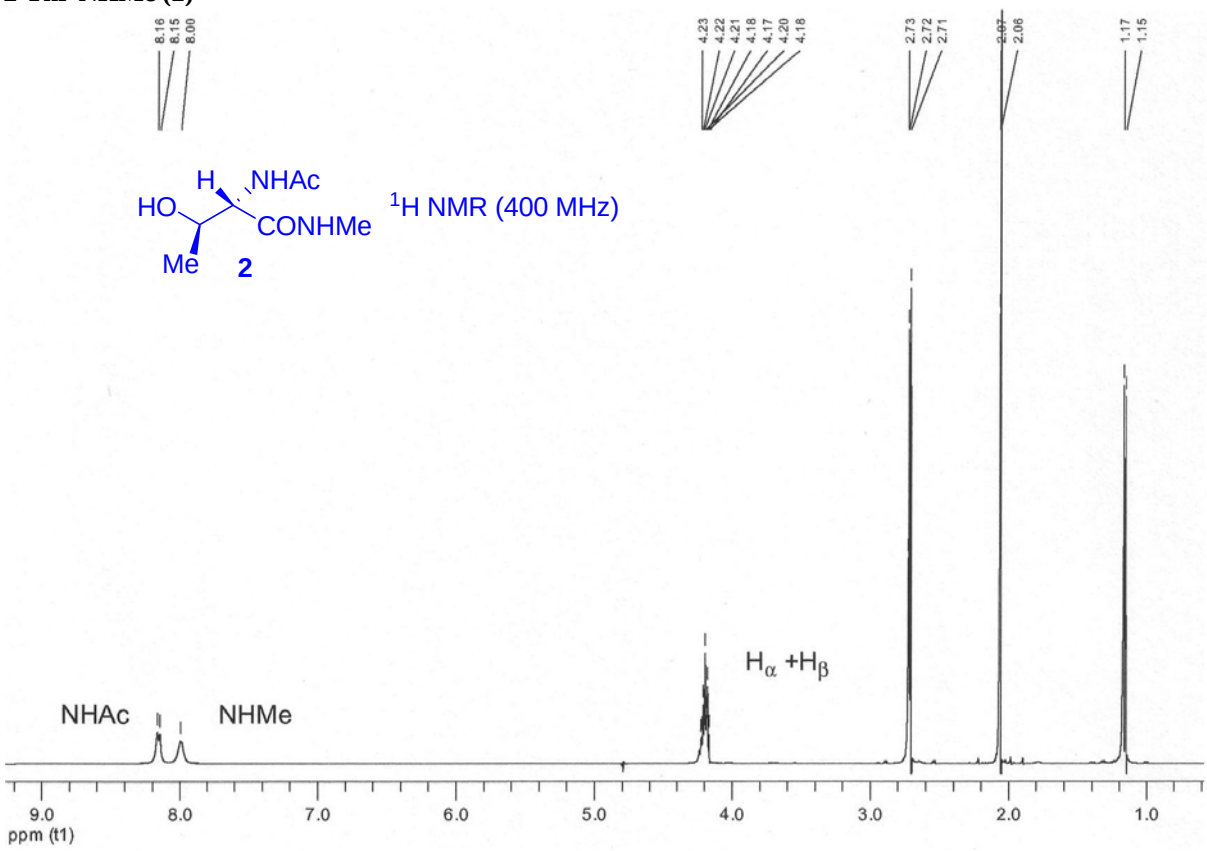
Boc-L-Thr(OBn)-NHMe (6)

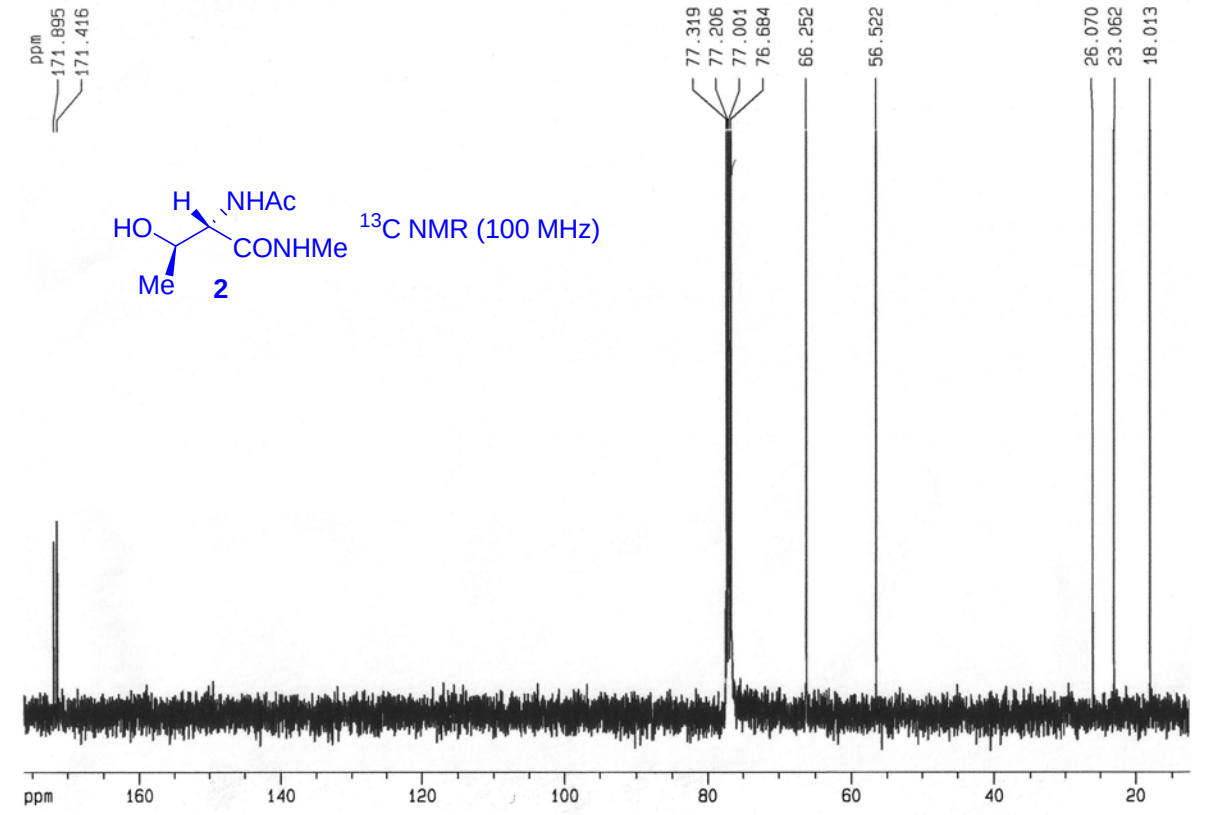


Ac-L-Thr(OBn)-NHMe (7)

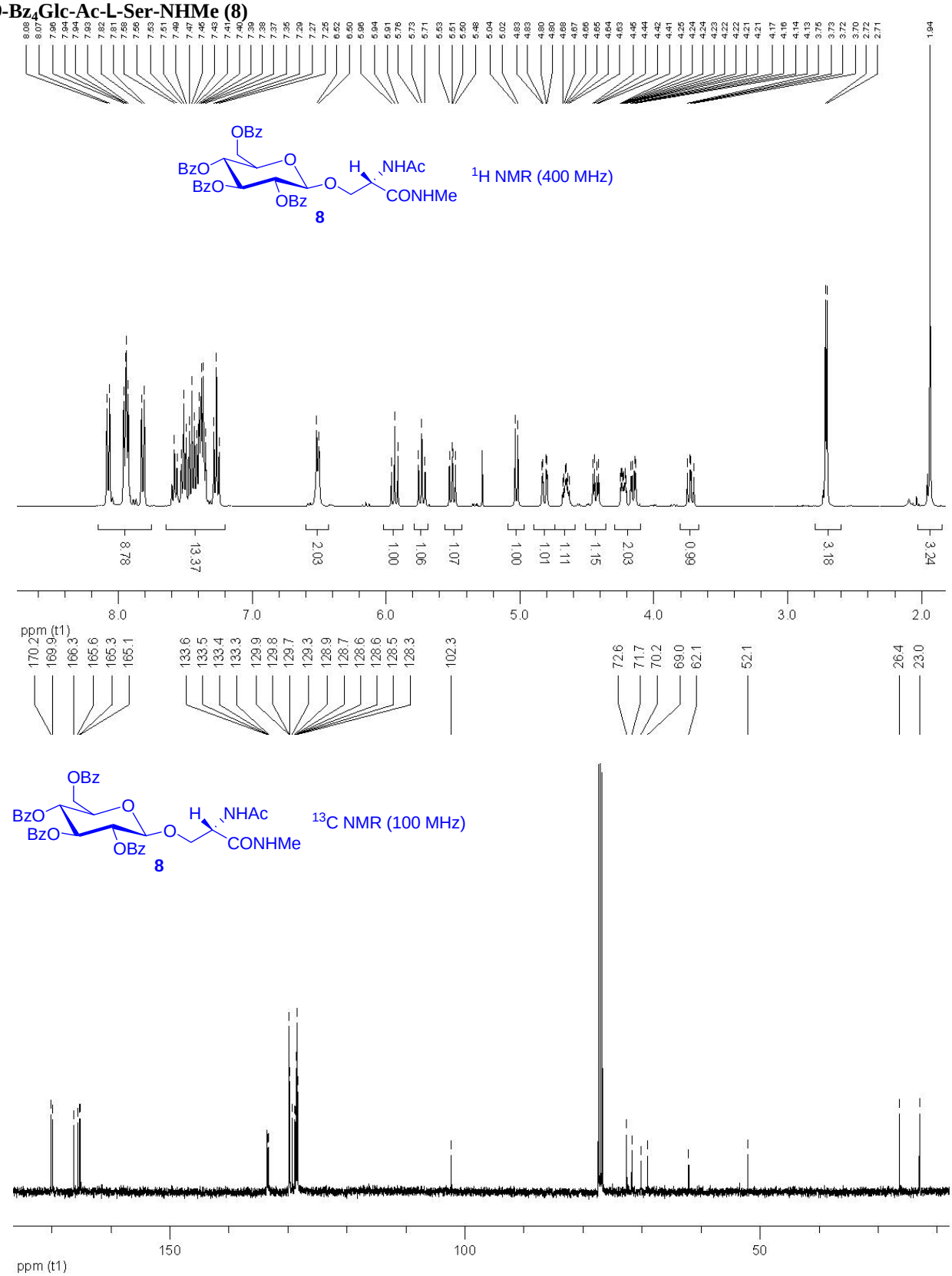


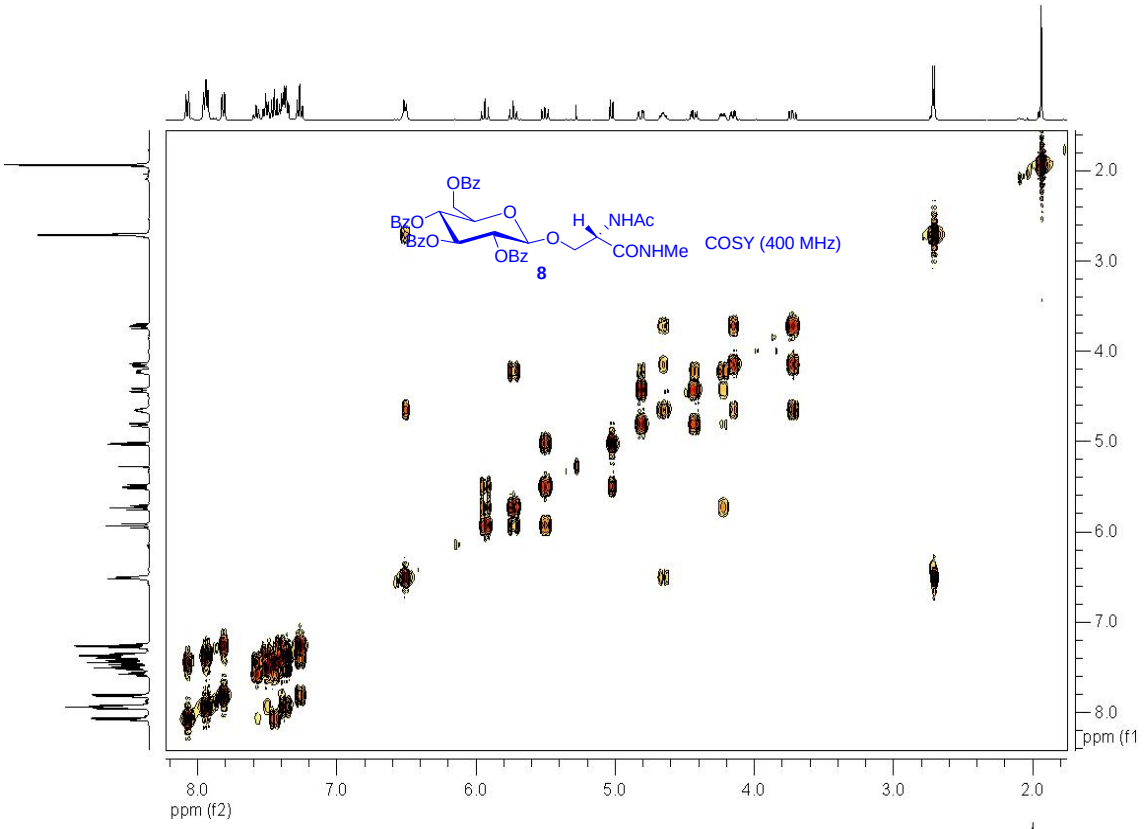
Ac-L-Thr-NHMe (2)

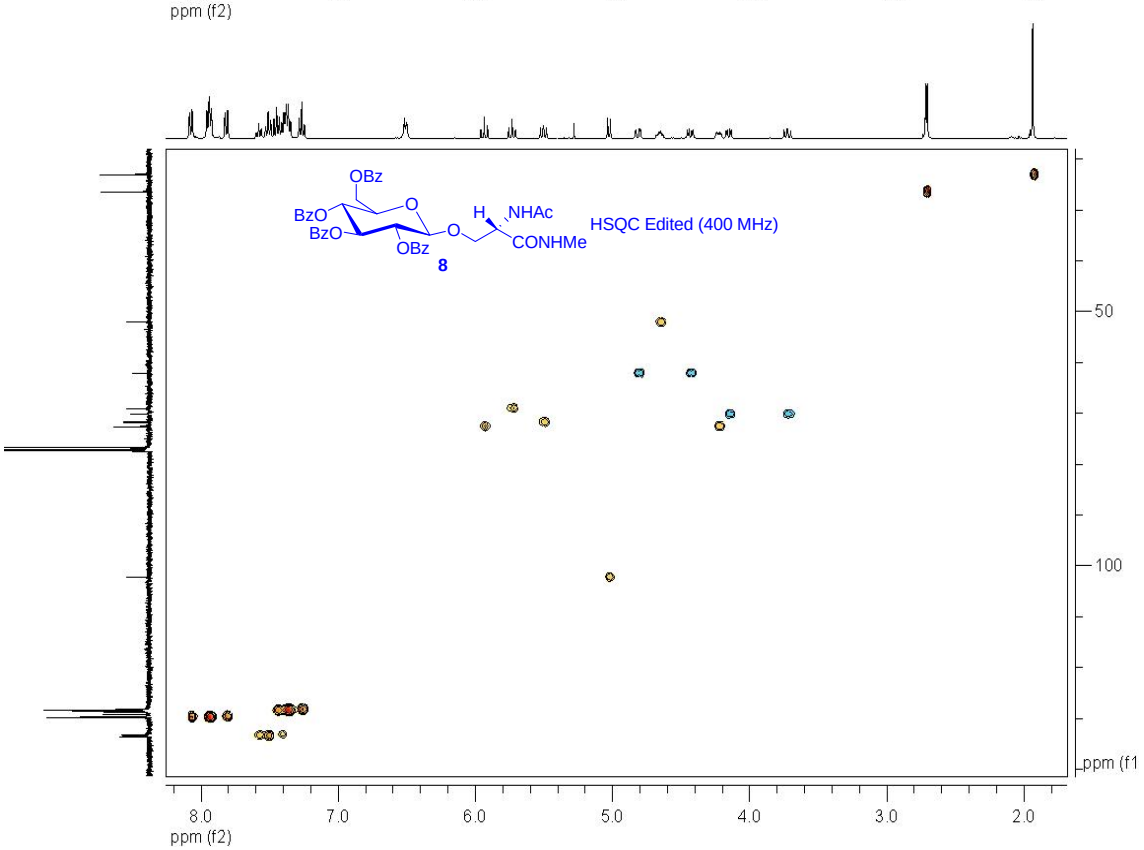




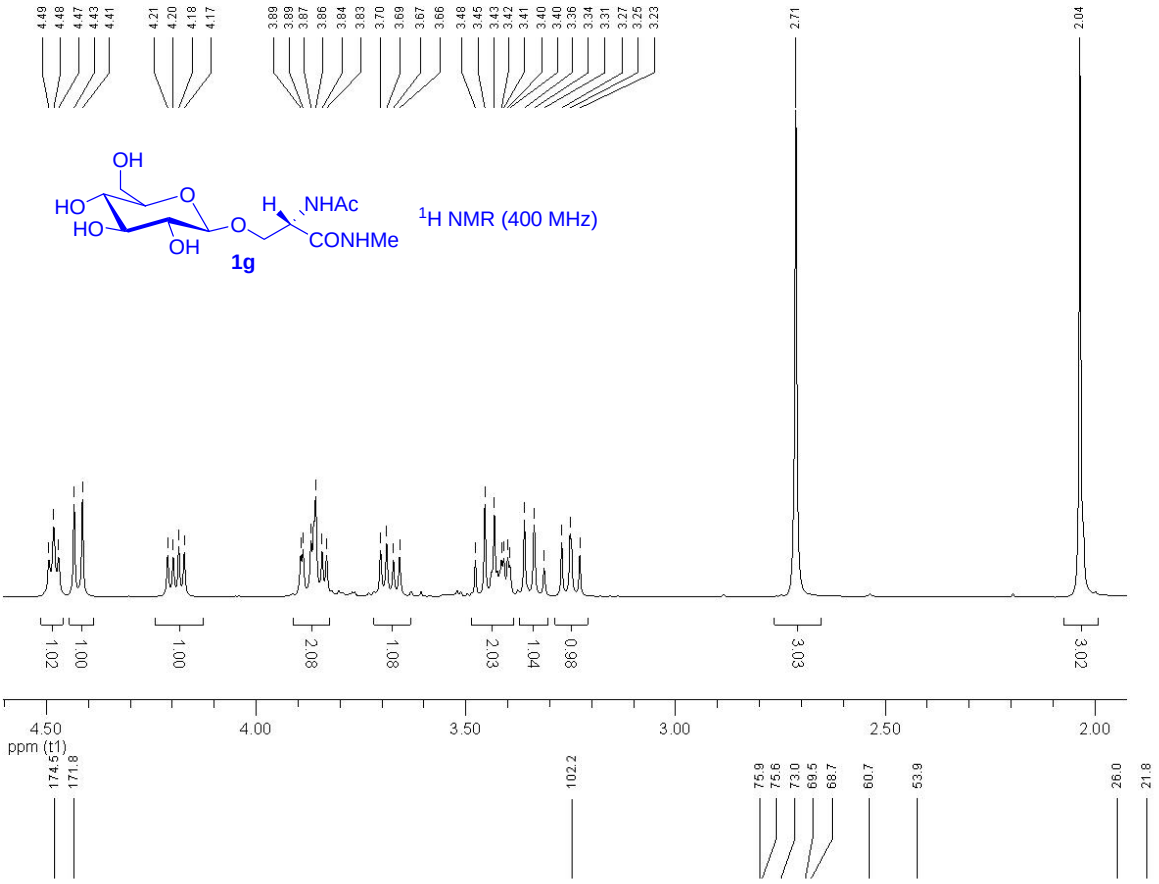
β-D-Bz$_4$Glc-Ac-L-Ser-NHMe (8)

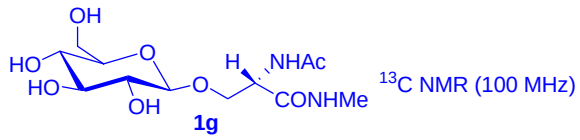


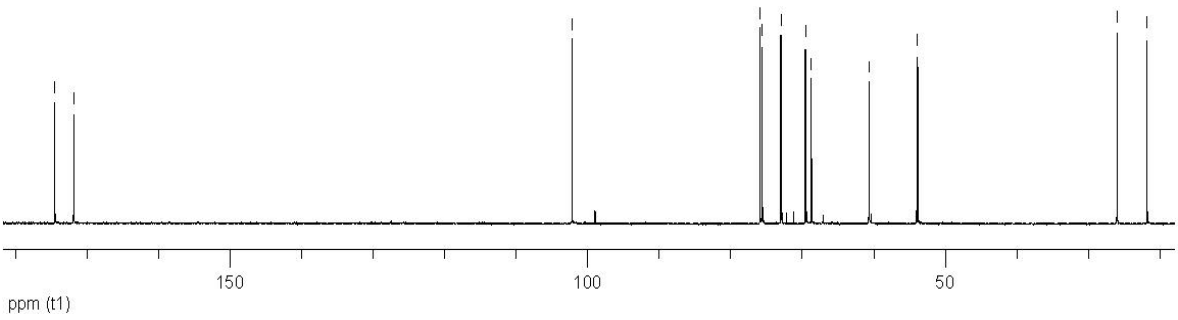

OBz
BzO
BzO
O
OBz
O
H
NHAc
CONHMe
8
COSY (400 MHz)
2.0
3.0
4.0
5.0
6.0
7.0
8.0
ppm (f1
8.0
7.0
6.0
5.0
4.0
3.0
2.0
ppm (f2)

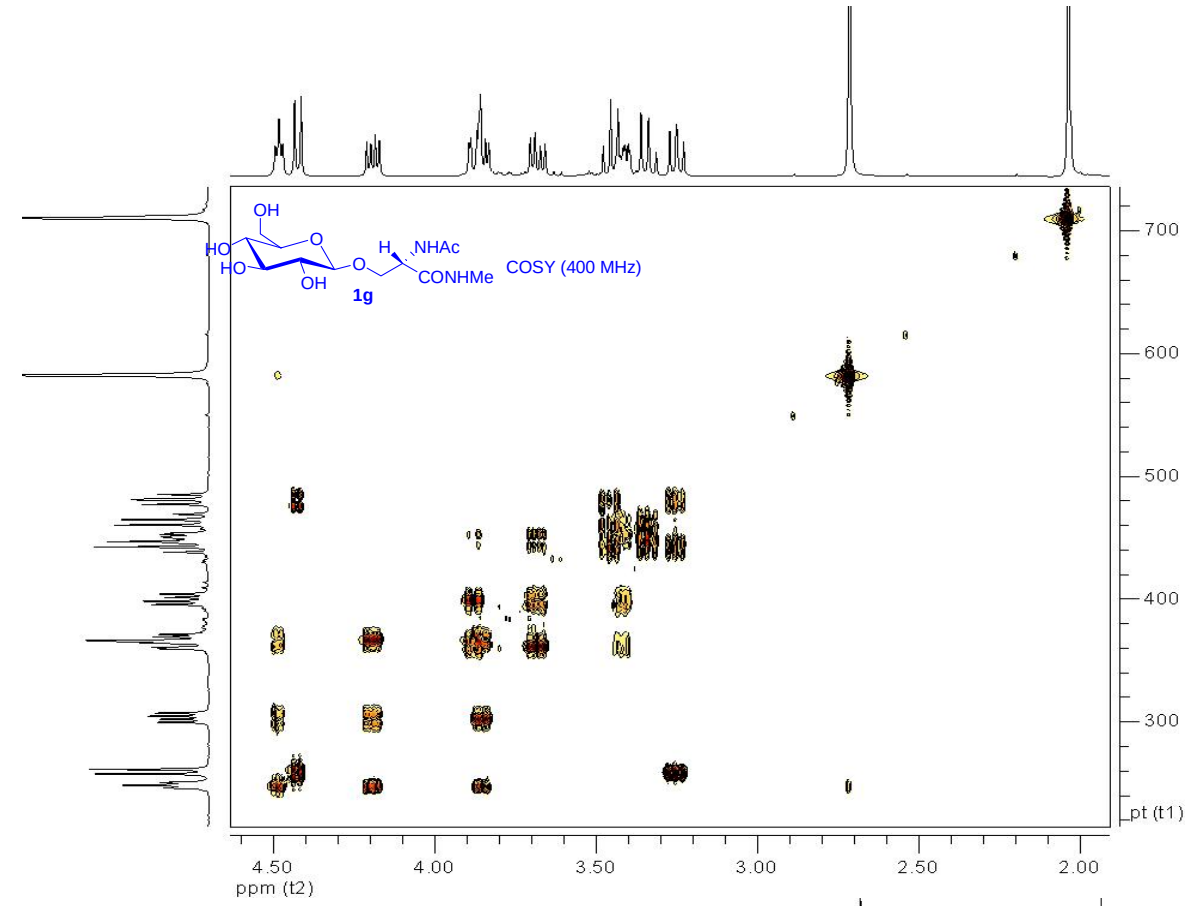

OBz
BzO
BzO
O
OBz
O
H
NHAc
CONHMe
8
HSQC Edited (400 MHz)
50
100
ppm (f1
8.0
7.0
6.0
5.0
4.0
3.0
2.0
ppm (f2)

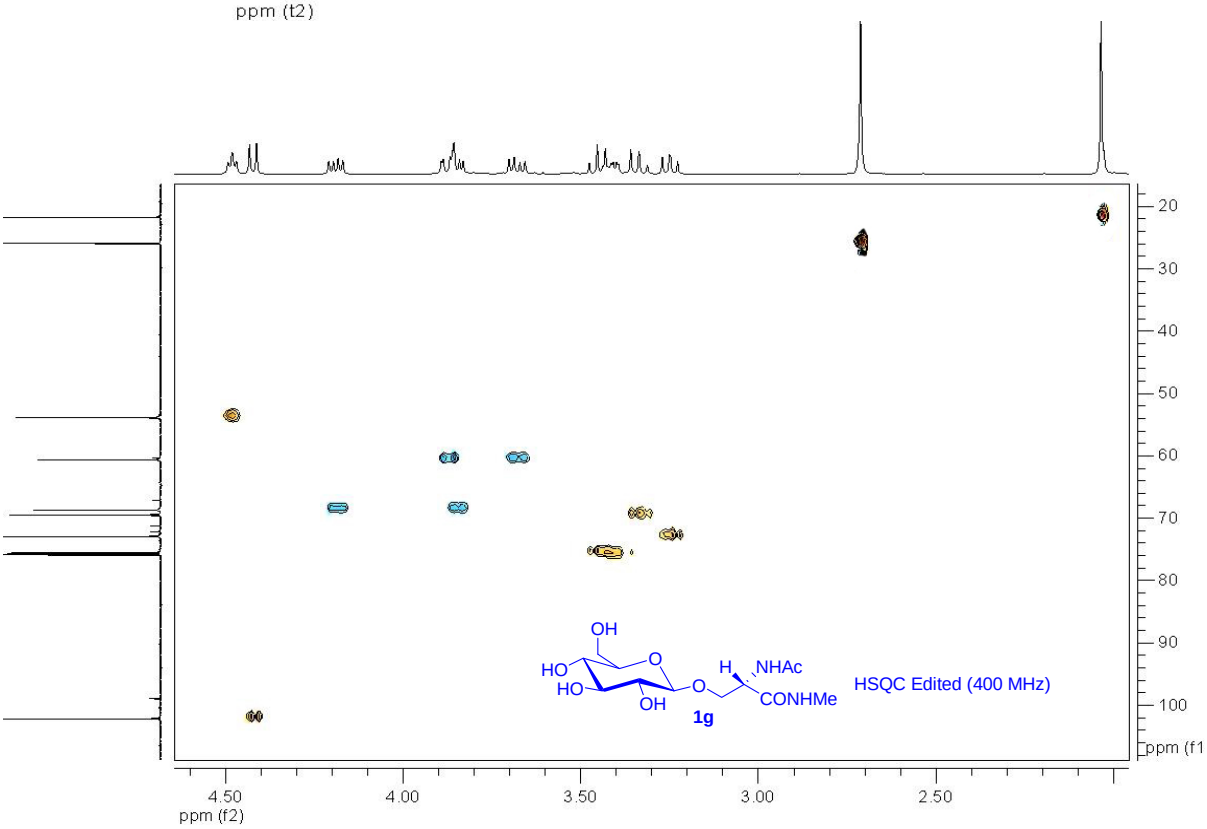
β-D-Glc-Ac-L-Ser-NHMe (1g)



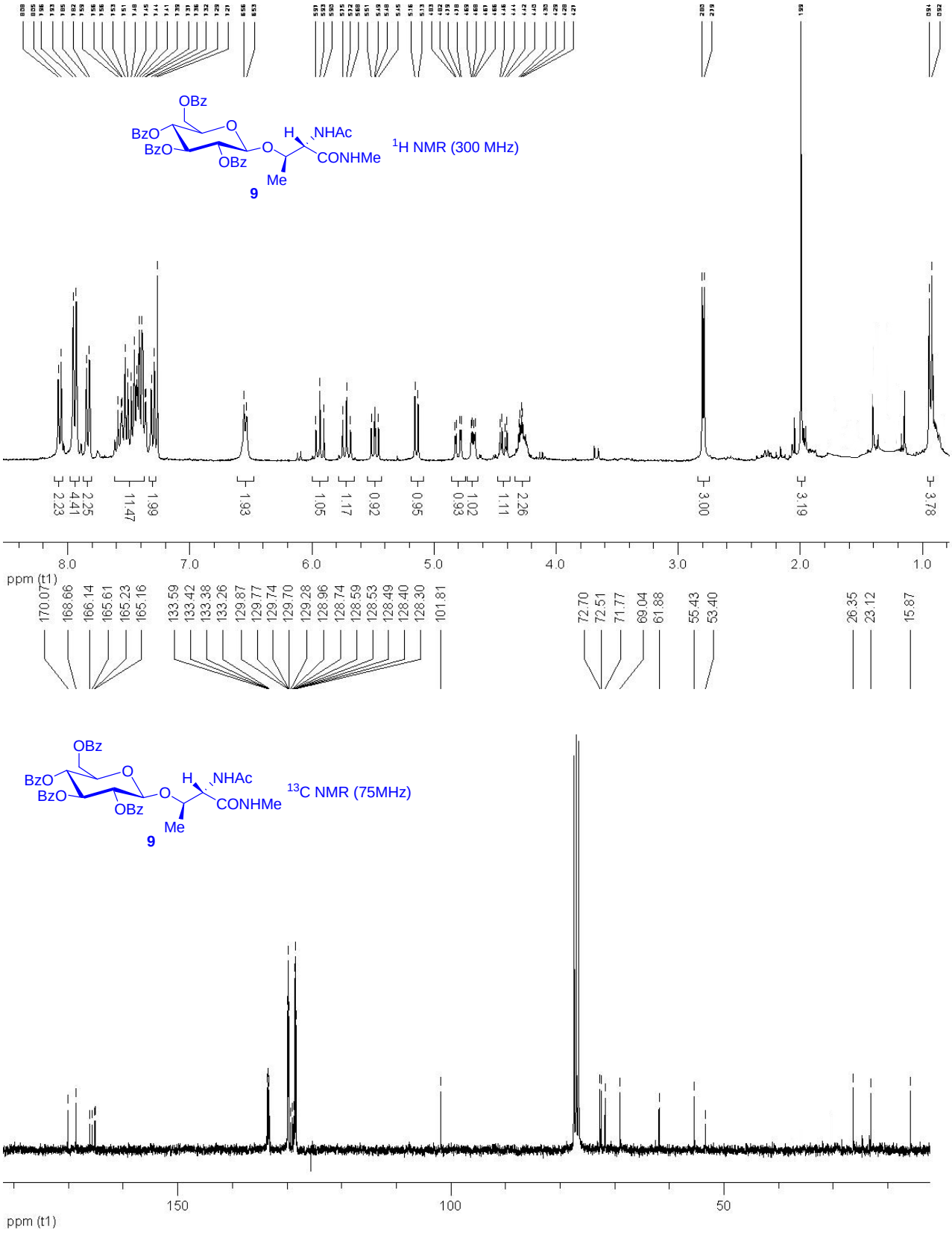




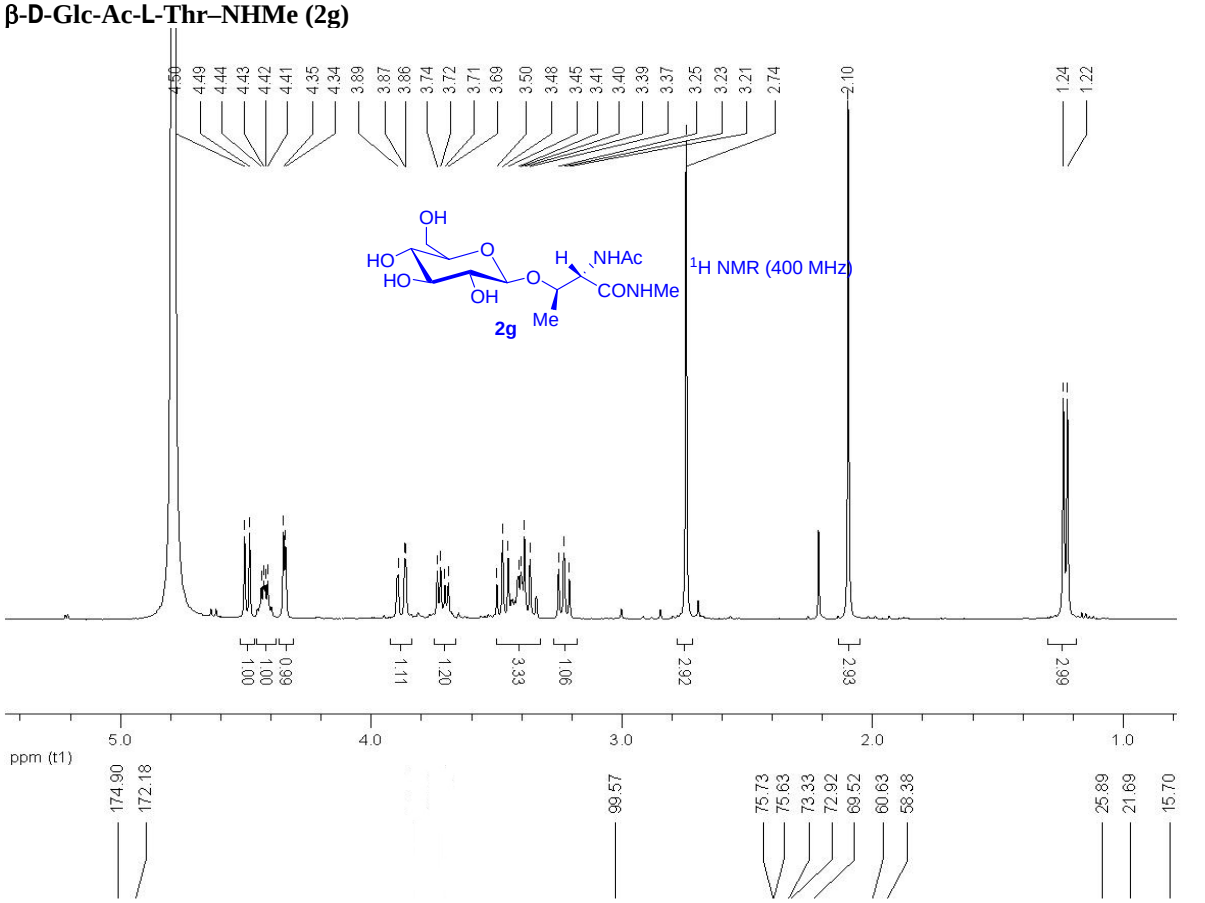

OH
HO
HO
O
OH
O
H
NHAc
CONHMe
1g
COSY (400 MHz)
700
600
500
400
300
pt (t1)
4.50
4.00
3.50
3.00
2.50
2.00
ppm (t2)

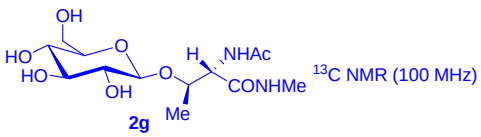

OH
HO
HO
O
OH
O
H
NHAc
CONHMe
1g
HSQC Edited (400 MHz)
20
30
40
50
60
70
80
90
100
ppm (f1
4.50
4.00
3.50
3.00
2.50
ppm (f2)

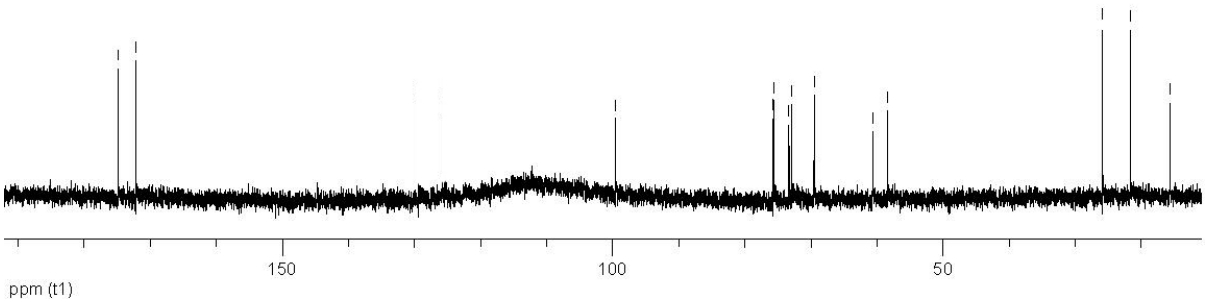
β-D-Bz$_4$-Glc-Ac-L-Thr-NHMe (9)

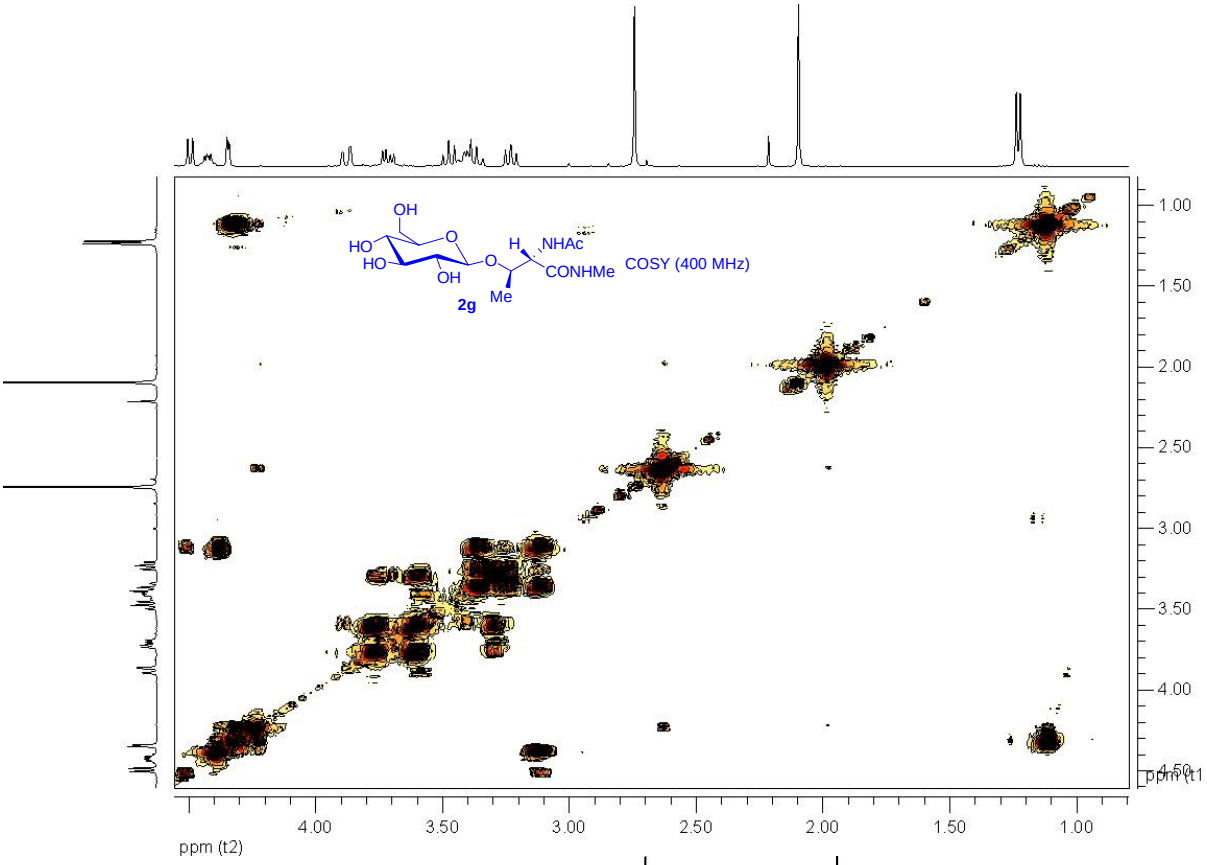


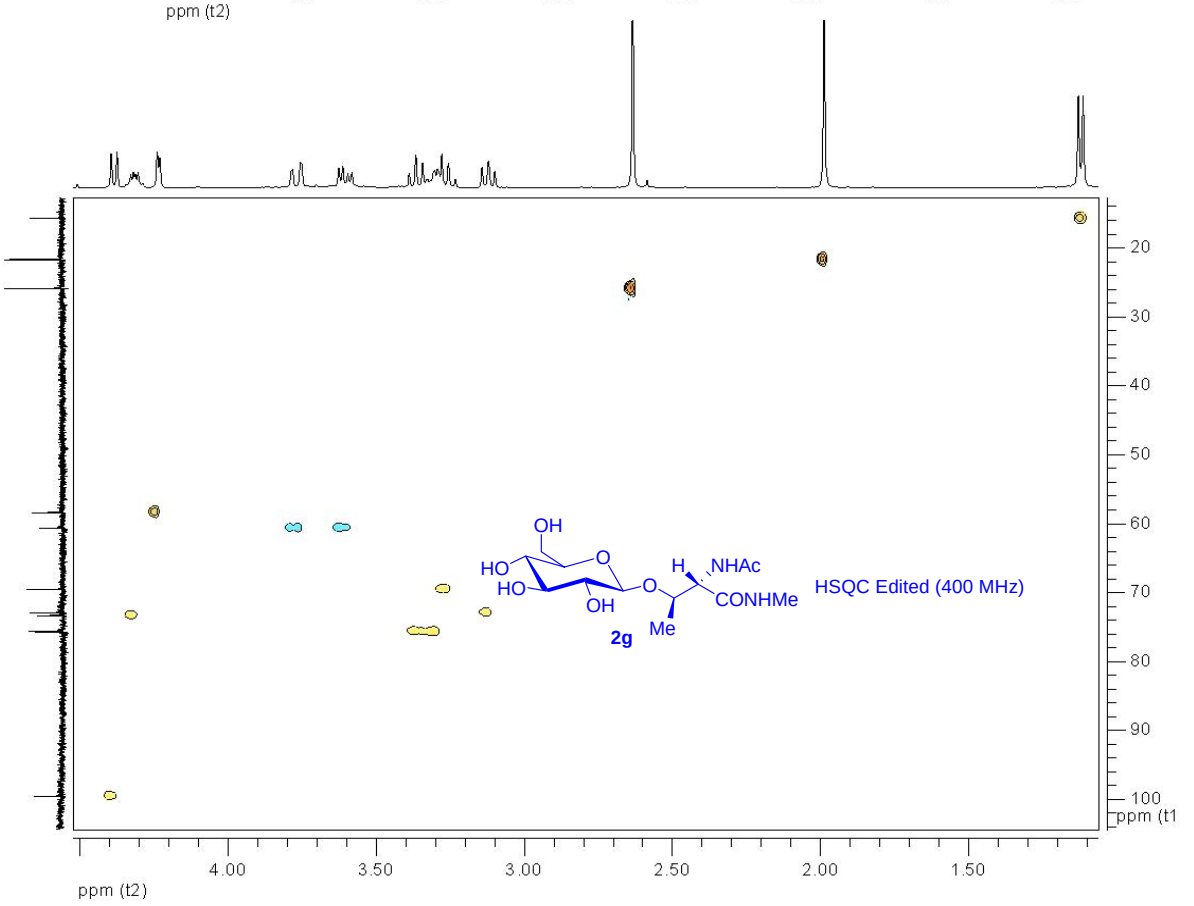
β-D-Glc-Ac-L-Thr–NHMe (2g)



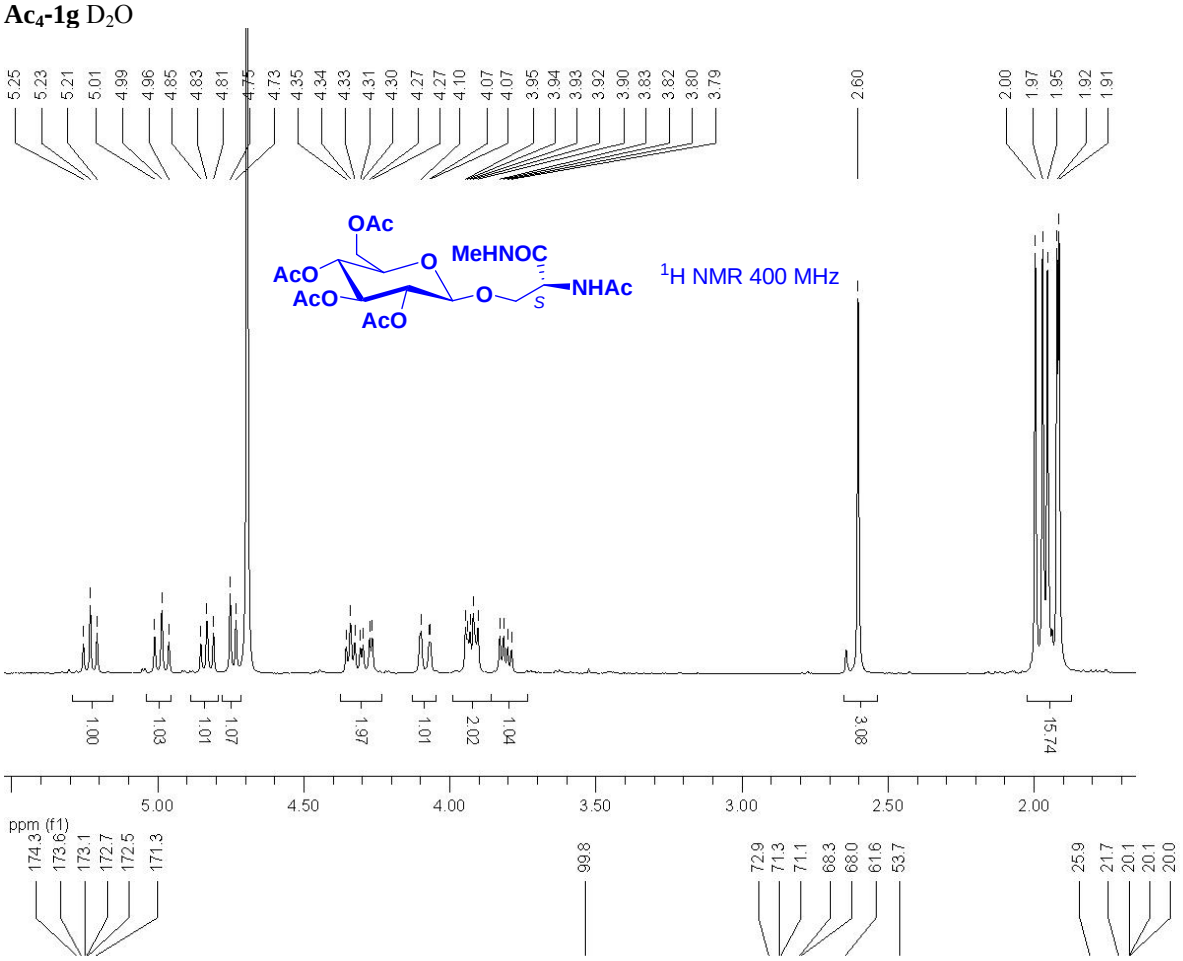


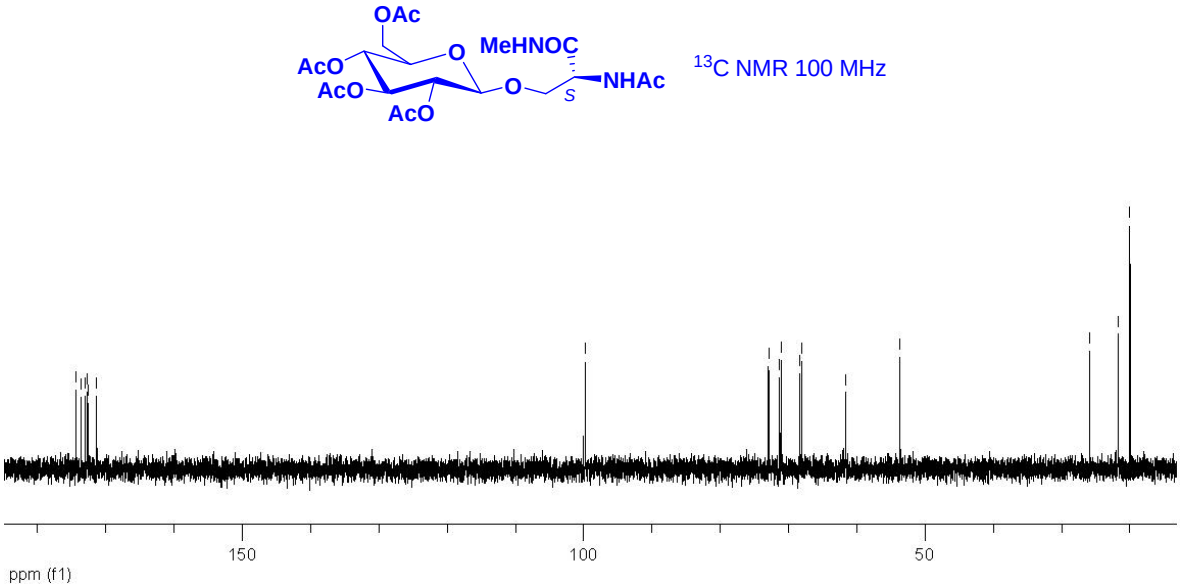


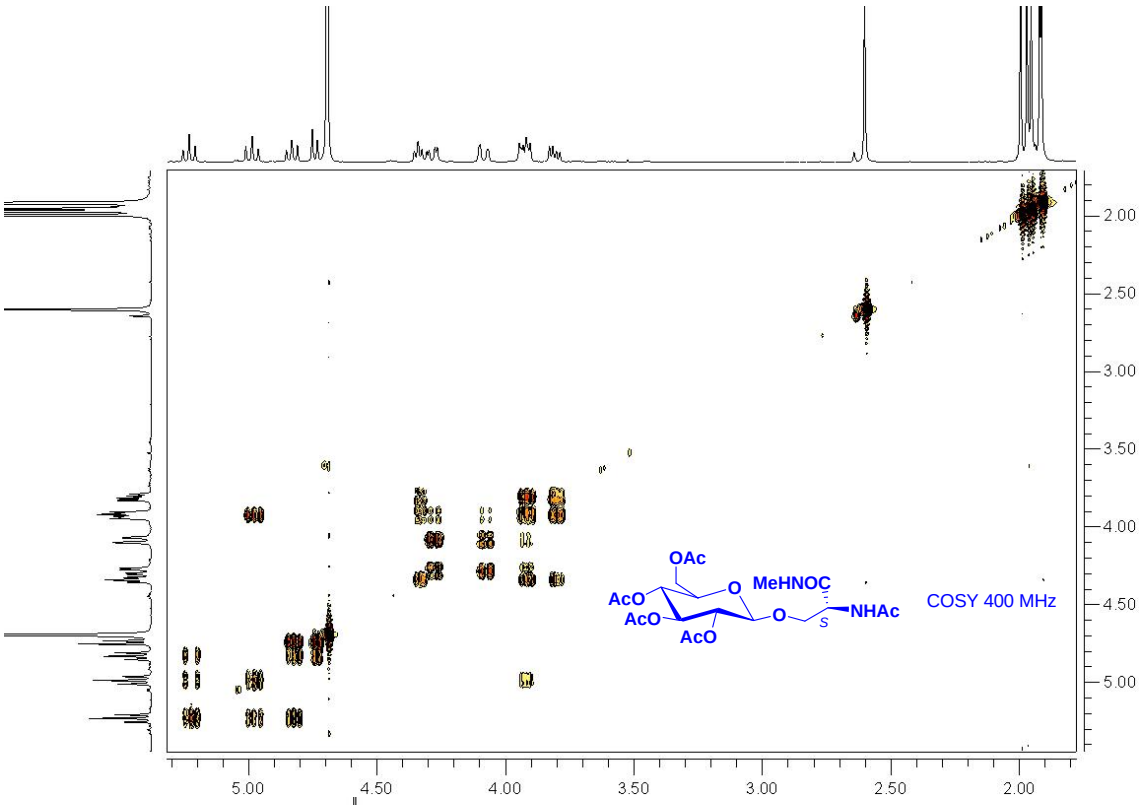

OH
HO
HO
O
OH
O
H
NHAc
CONHMe
Me
2g
COSY (400 MHz)
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
ppm (t1)
4.00
3.50
3.00
2.50
2.00
1.50
1.00
ppm (t2)

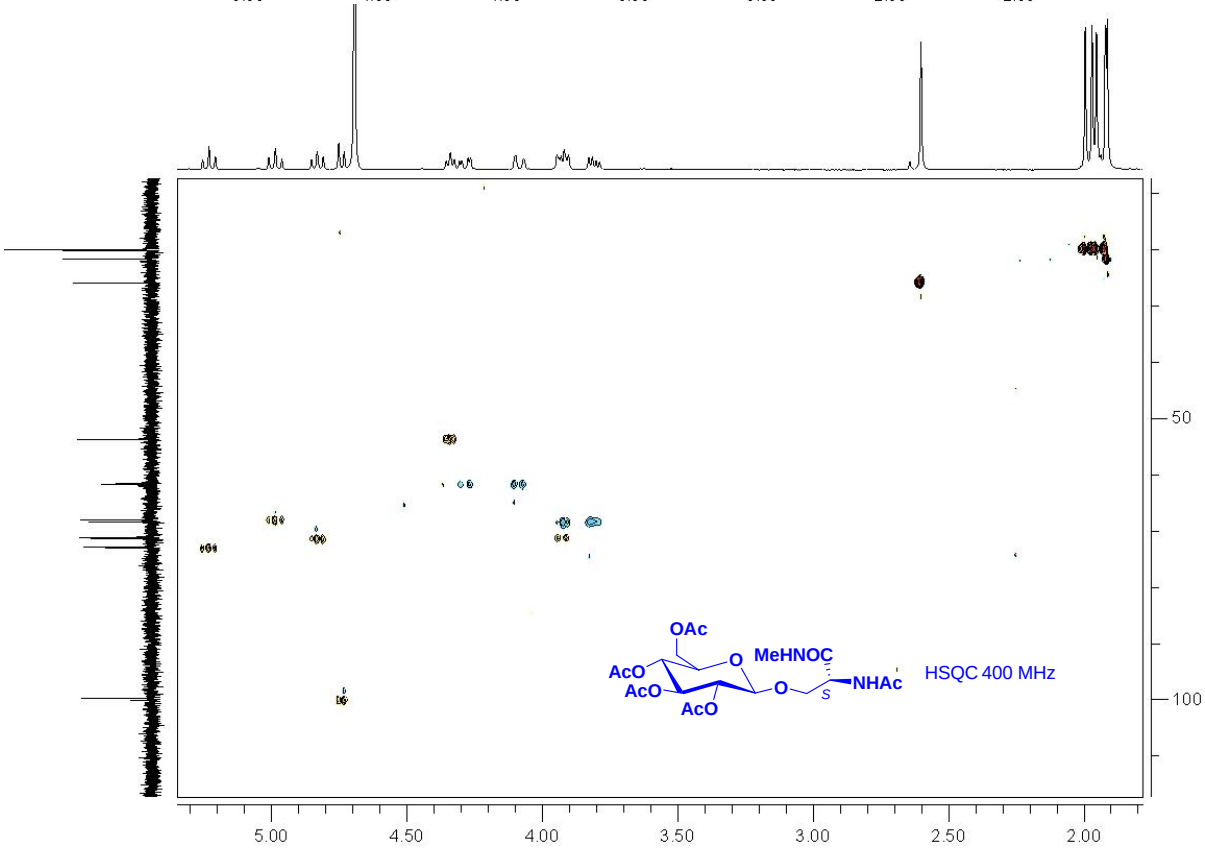

OH
HO
HO
O
OH
O
H
NHAc
CONHMe
Me
2g
HSQC Edited (400 MHz)
20
30
40
50
60
70
80
90
100
ppm (t1)
4.00
3.50
3.00
2.50
2.00
1.50
ppm (t2)

Ac_4-1g D_2O

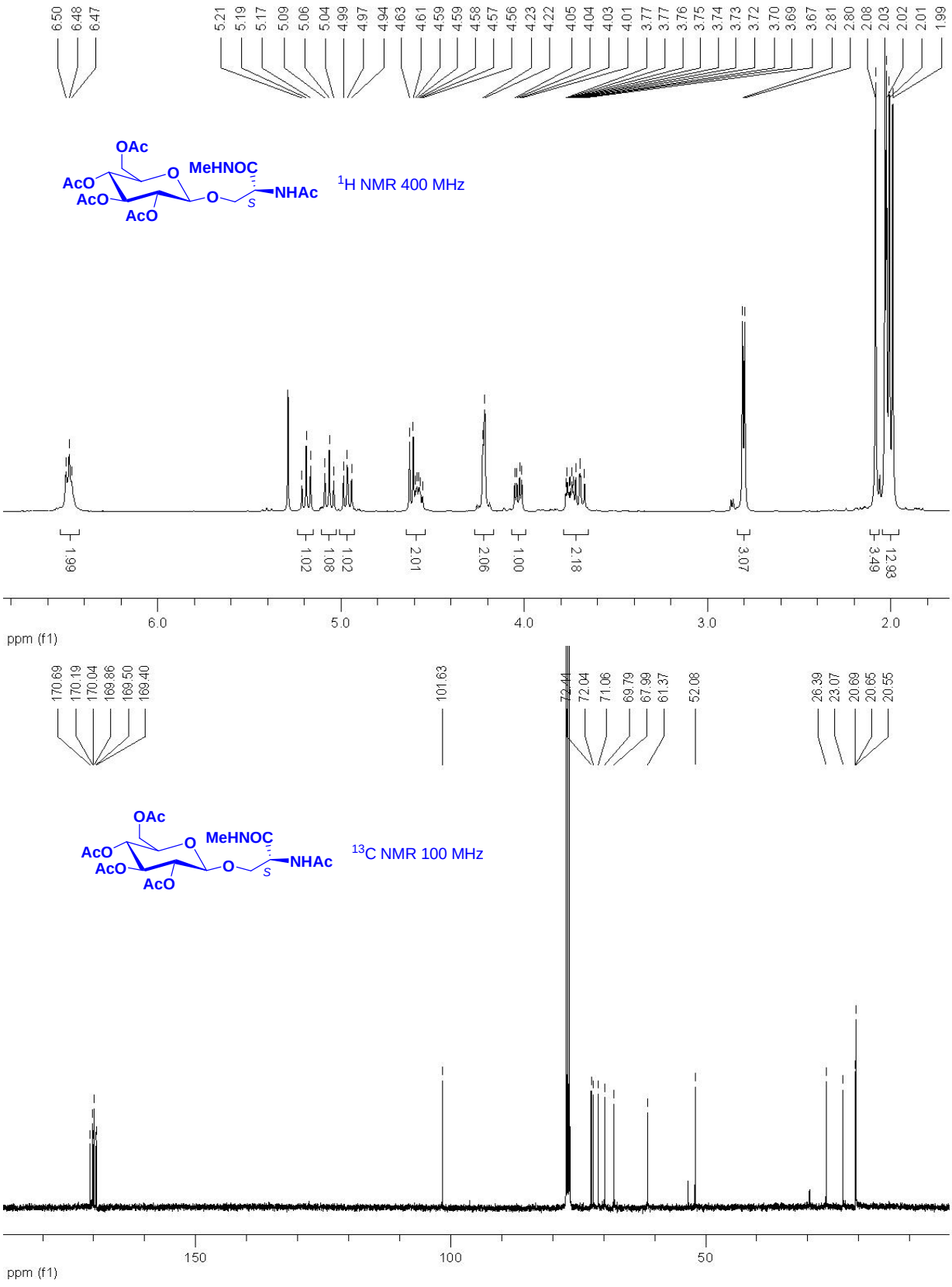

5.25 5.23 5.21 5.01 4.99 4.96 4.85 4.83 4.81 4.75 4.73 4.35 4.34 4.33 4.31 4.30 4.27 4.27 4.10 4.07 4.07 3.95 3.94 3.93 3.92 3.90 3.83 3.82 3.80 3.79
2.60
2.00 1.97 1.95 1.92 1.91
OAc
AcO
AcO
AcO
O
MeHNOC
O
S
NHAc
1H NMR 400 MHz
1.00 1.03 1.01 1.07 1.97 1.01 2.02 1.04 3.08 15.74
5.00 4.50 4.00 3.50 3.00 2.50 2.00
ppm (f1)

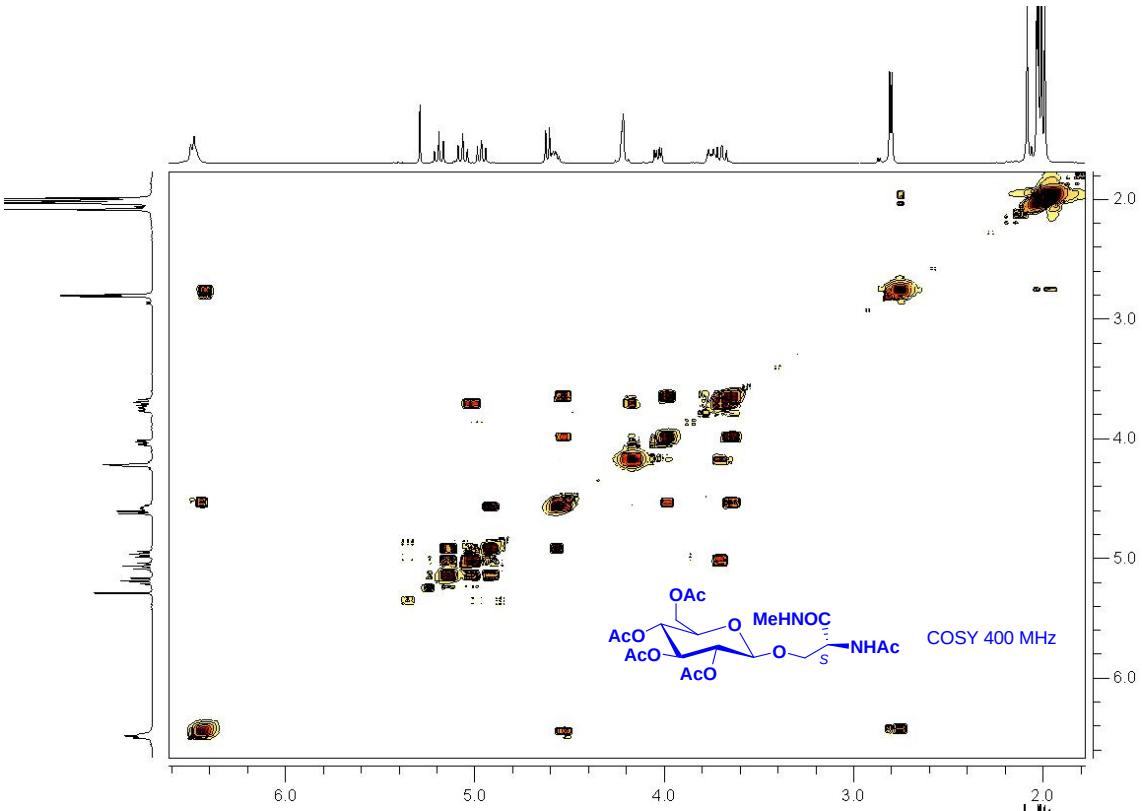

174.3 173.6 173.1 172.7 172.5 171.3
99.8
72.9 71.3 71.1 68.3 68.0 61.6 53.7
25.9 21.7 20.1 20.1 20.0
OAc
AcO
AcO
AcO
O
MeHNOC
O
S
NHAc
13C NMR 100 MHz
150 100 50
ppm (f1)

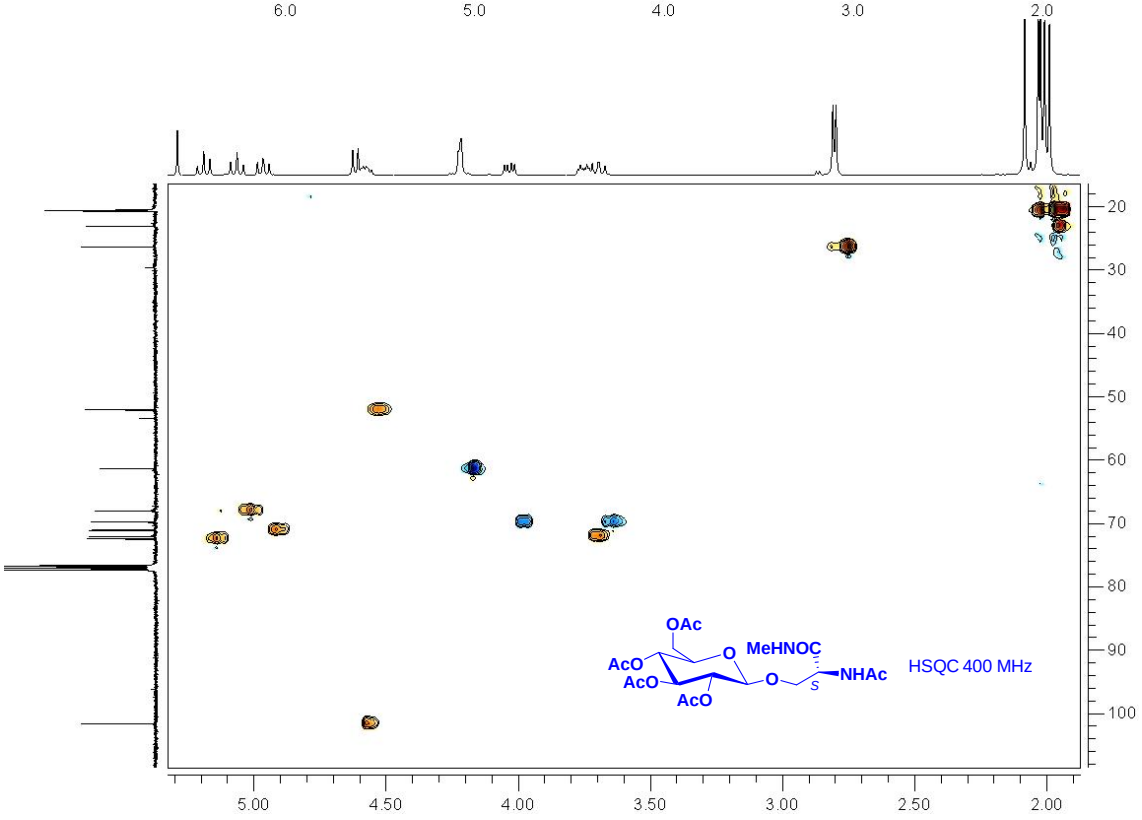

OAc
AcO
AcO
AcO
O
MeHNOC
O
S
NHAc
COSY 400 MHz
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.00
4.50
4.00
3.50
3.00
2.50
2.00


OAc
AcO
AcO
AcO
O
MeHNOC
O
S
NHAc
HSQC 400 MHz
50
100
5.00
4.50
4.00
3.50
3.00
2.50
2.00

Ac_4-1g $CDCl_3$


6.50 6.48 6.47
5.21 5.19 5.17 5.09 5.06 5.04 4.99 4.97 4.94 4.63 4.61 4.59 4.59 4.58 4.57 4.56 4.23 4.22 4.05 4.04 4.03 4.01 3.77 3.77 3.76 3.75 3.74 3.73 3.72 3.70 3.69 3.67 2.81 2.80 2.08 2.03 2.02 2.01 1.99
OAc
AcO
AcO
AcO
O
MeHNOC
O
S
NHAc
1H NMR 400 MHz
1.99 1.02 1.08 1.02 2.01 2.06 1.00 2.18 3.07 3.49 12.93
6.0 5.0 4.0 3.0 2.0
ppm (f1)
170.69 170.19 170.04 169.86 169.50 169.40
101.63
72.44 72.04 71.06 69.79 67.99 61.37
52.08
26.39 23.07 20.69 20.65 20.55
OAc
AcO
AcO
AcO
O
MeHNOC
O
S
NHAc
13C NMR 100 MHz
150 100 50
ppm (f1)

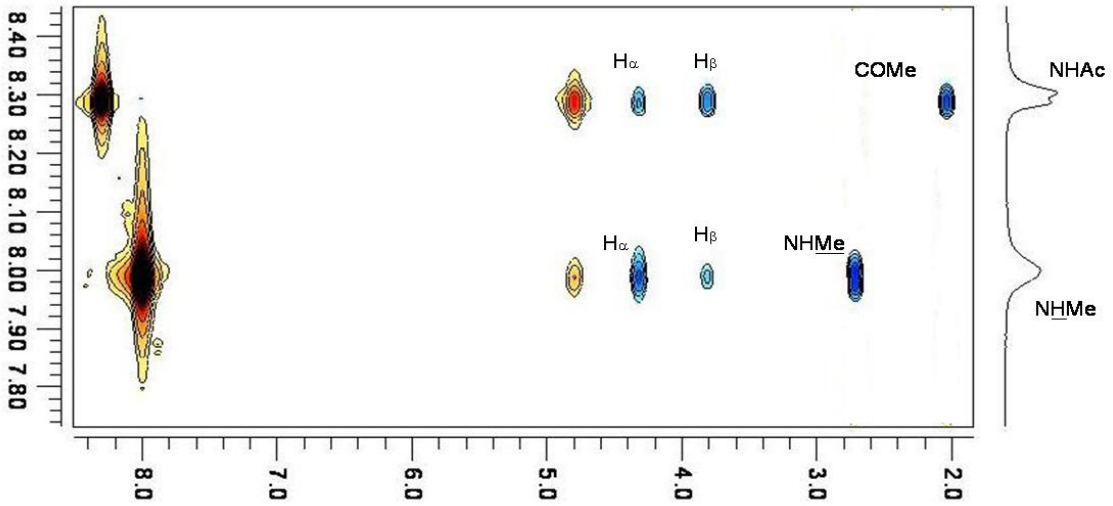

OAc
AcO
AcO
AcO
O
MeHNOC
O
S
NHAc
COSY 400 MHz
2.0
3.0
4.0
5.0
6.0
6.0
5.0
4.0
3.0
2.0


OAc
AcO
AcO
AcO
O
MeHNOC
O
S
NHAc
HSQC 400 MHz
20
30
40
50
60
70
80
90
100
5.00
4.50
4.00
3.50
3.00
2.50
2.00

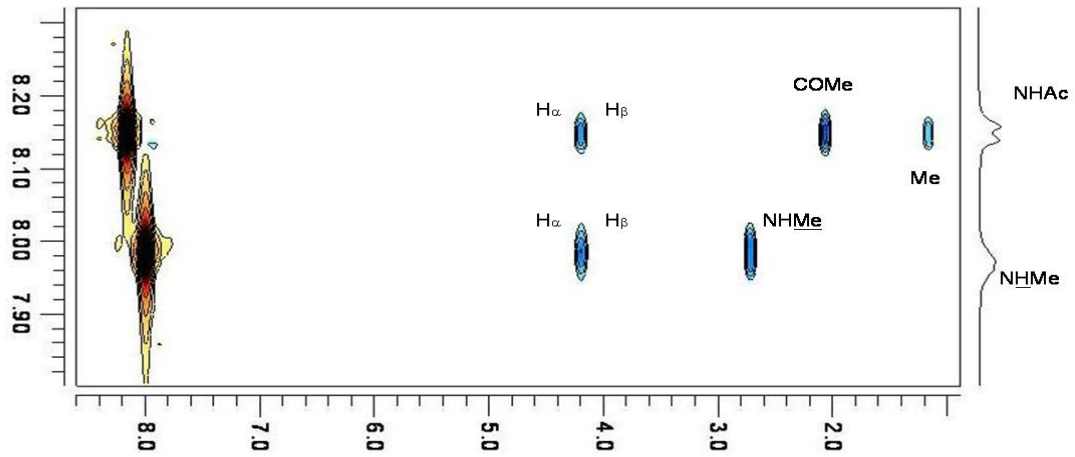
NOE build-up curves for 1g, Ac_4-1g and 2g and 2D NOESY for 1, 2, 1g, Ac_4-1g and 2g.

NOEs intensities were normalized with respect to the diagonal peak at zero mixing time. Experimental NOEs were fitted to a double exponential function, $f(t)=p0(e^{-p1t})(1-e^{-p2t})$ with p0, p1 and p2 being adjustable parameters.[1] The initial slope was determined from the first derivative at time t=0, $f'(0)=p0p2$. From the initial slopes, interproton distances were obtained by employing the isolated spin pair approximation. Selective ge-1D ROESY was carried out using the 1D-SPFGE sequence.

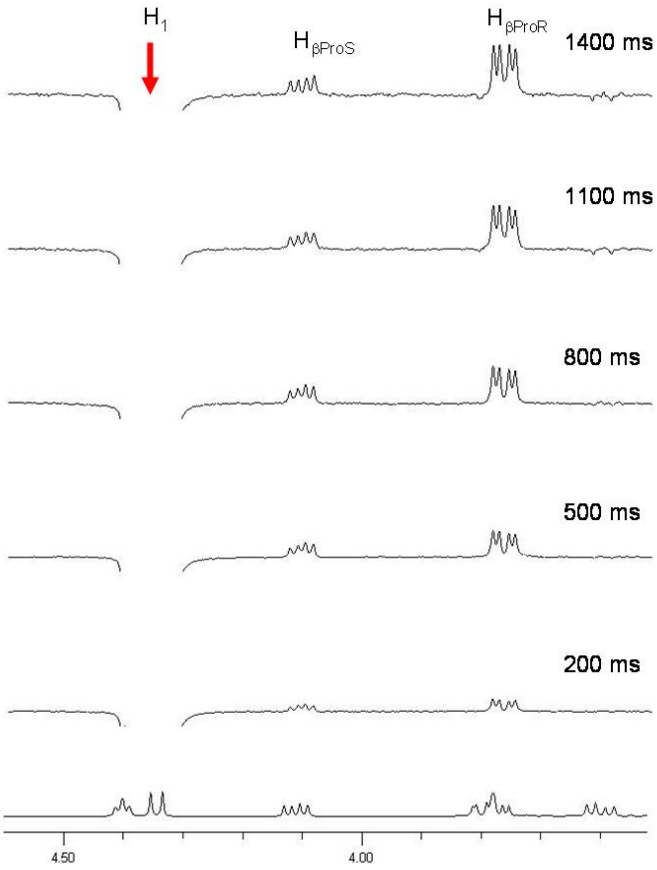
Amide-aliphatic cross-peaks NOESY (mixing time = 800 ms 400 MHz) for compound **1**



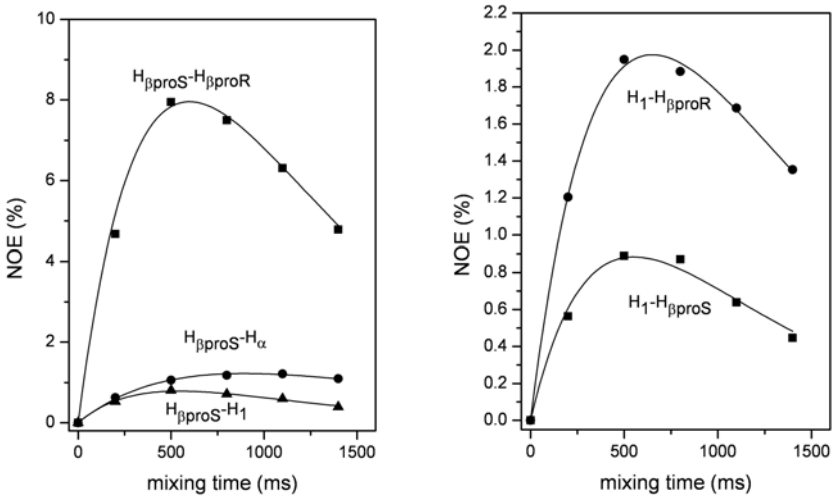
Amide-aliphatic cross-peaks NOESY (mixing time = 800 ms 400 MHz) for compound **2**

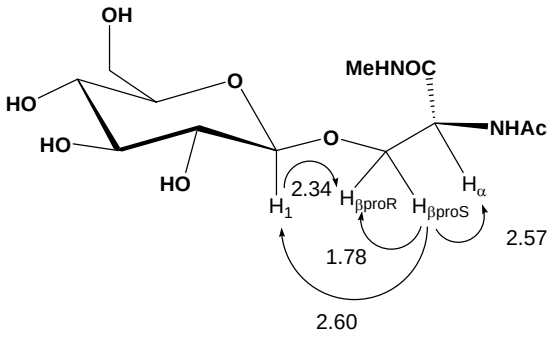


1D-NOESY experiments at different mixing times for compound **1g**

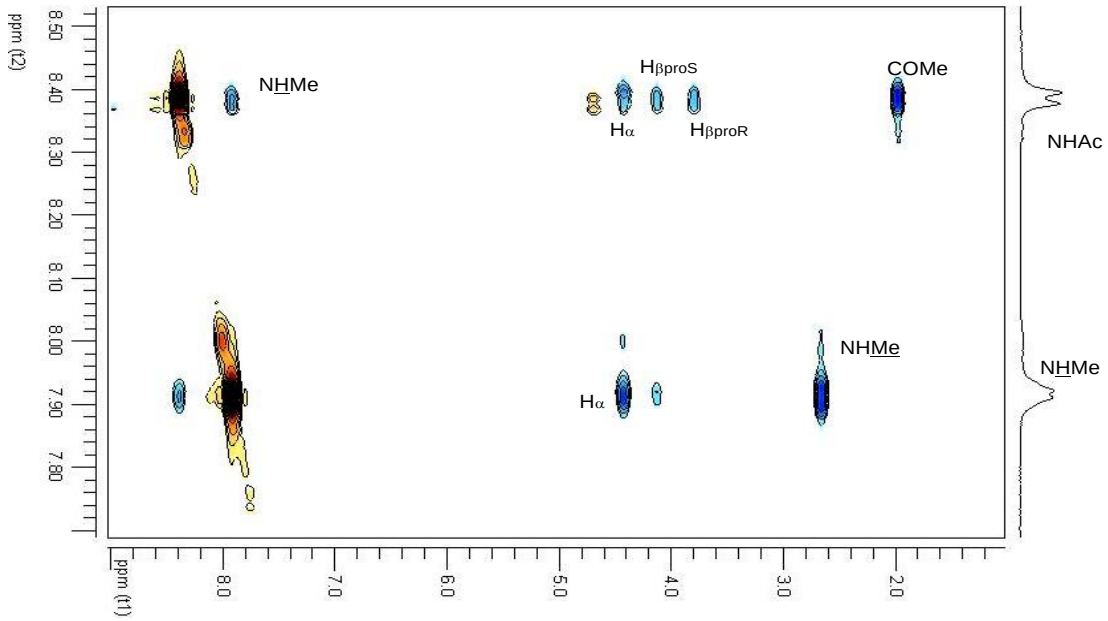


NOE Build-up curves for compound **1g**

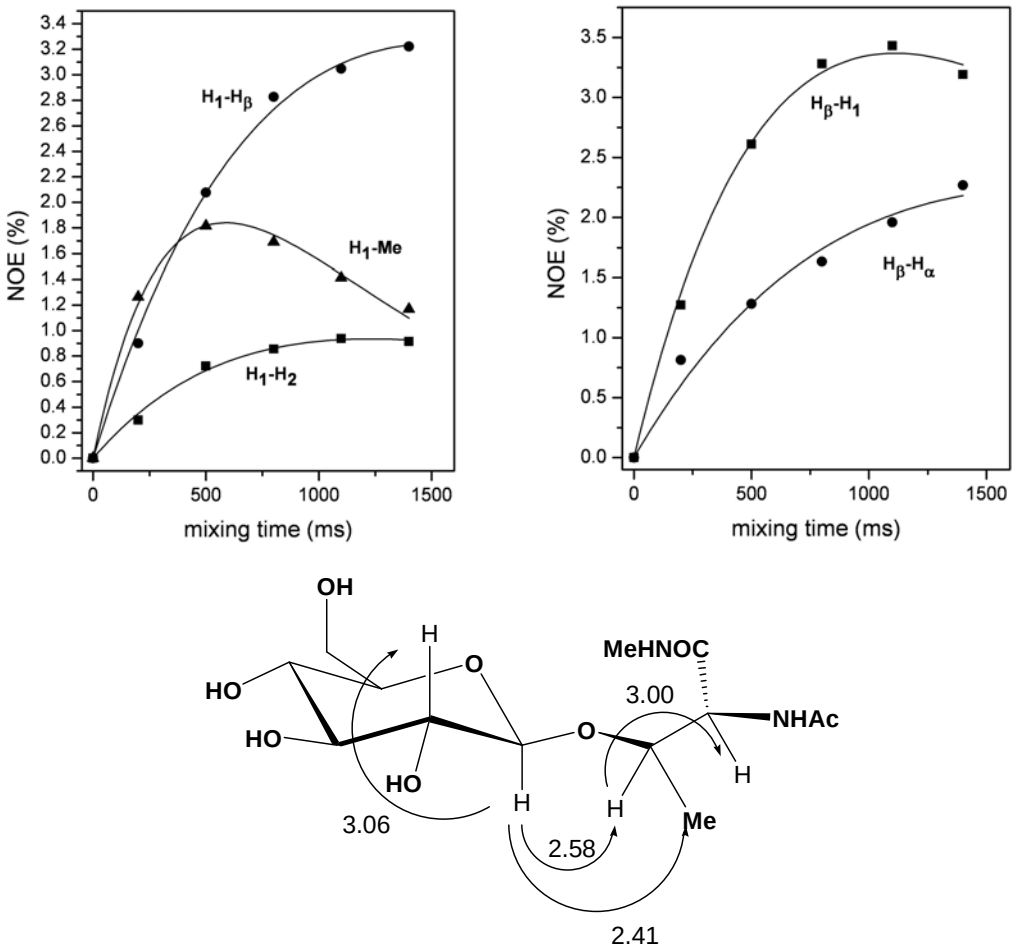




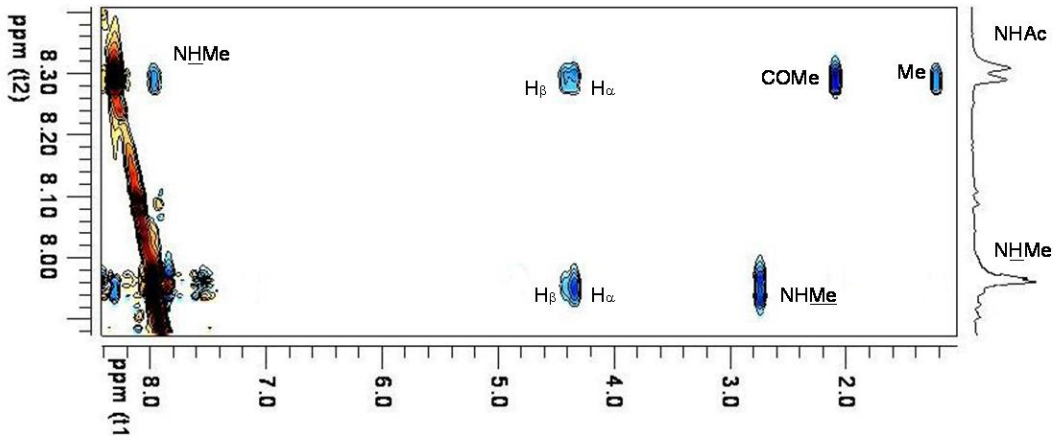
Amide cross-peaks NOESY (mixing time = 800 ms 400 MHz) for compound **1g**



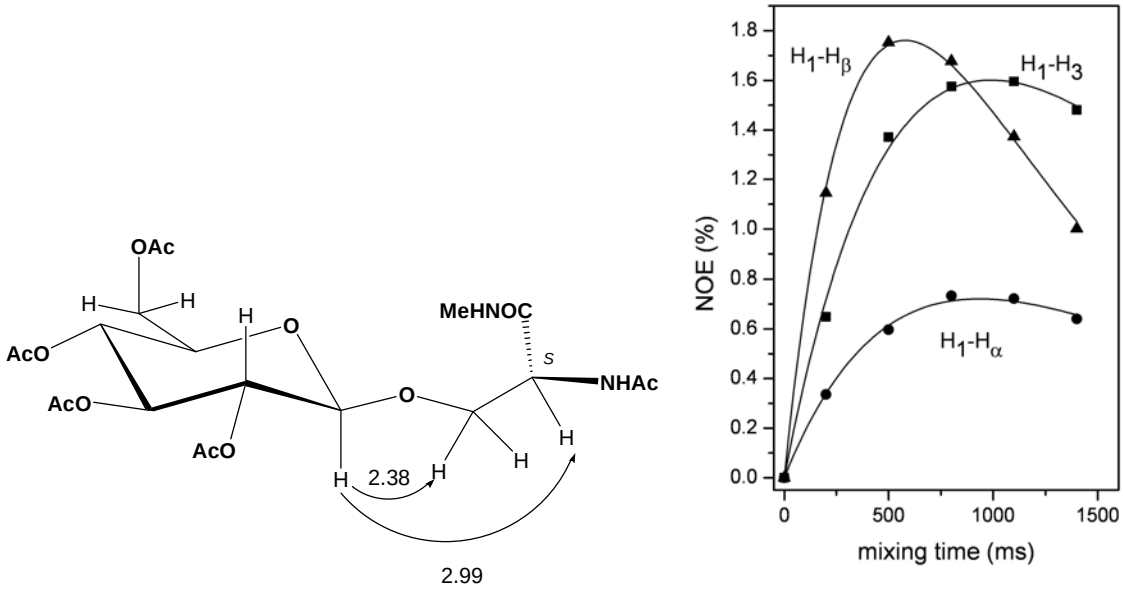
NOE Build-up curves for compound **2g**



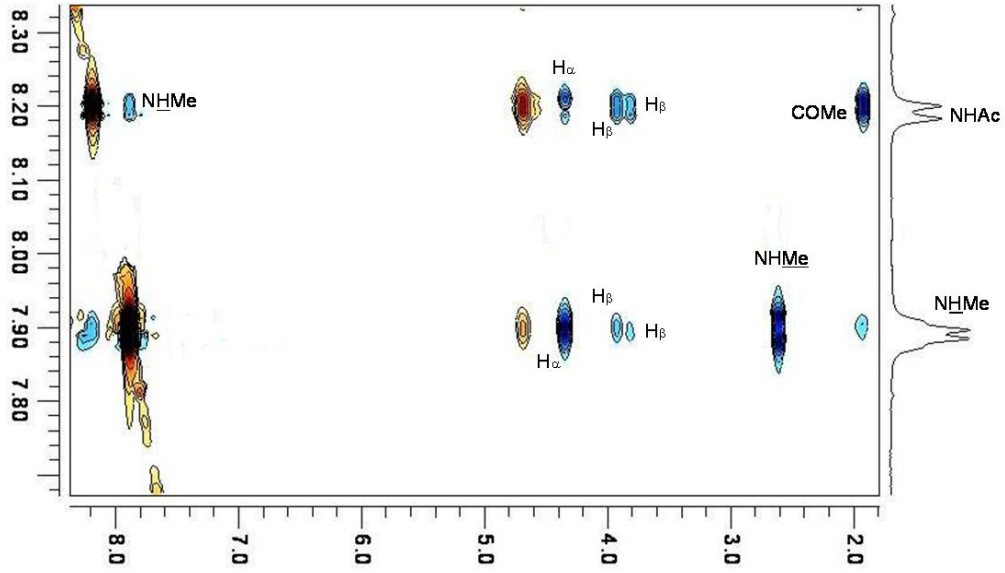
Amide cross-peaks NOESY (mixing time = 800 ms 400 MHz) for compound **2g**



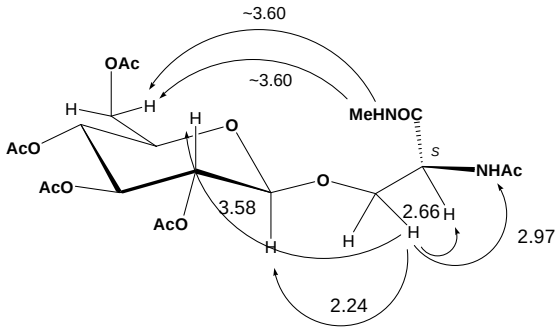
NOE Build-up curves for compound **Ac_4-1g** in D_2O

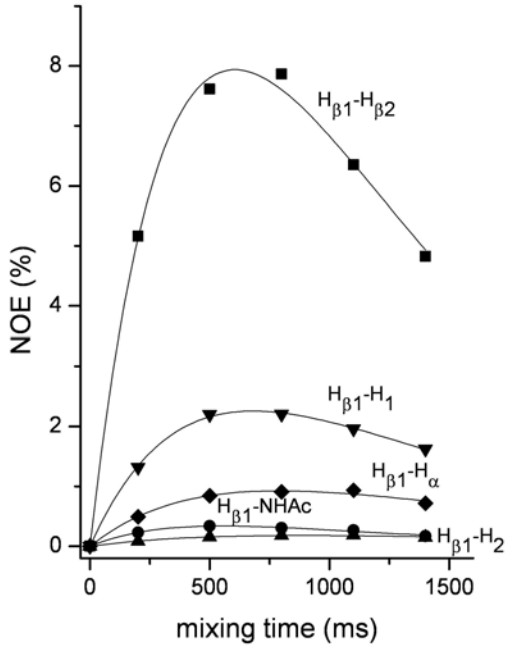


NOESY (mixing time = 800 ms 400 MHz) for compound **Ac_4-1g** in D_2O

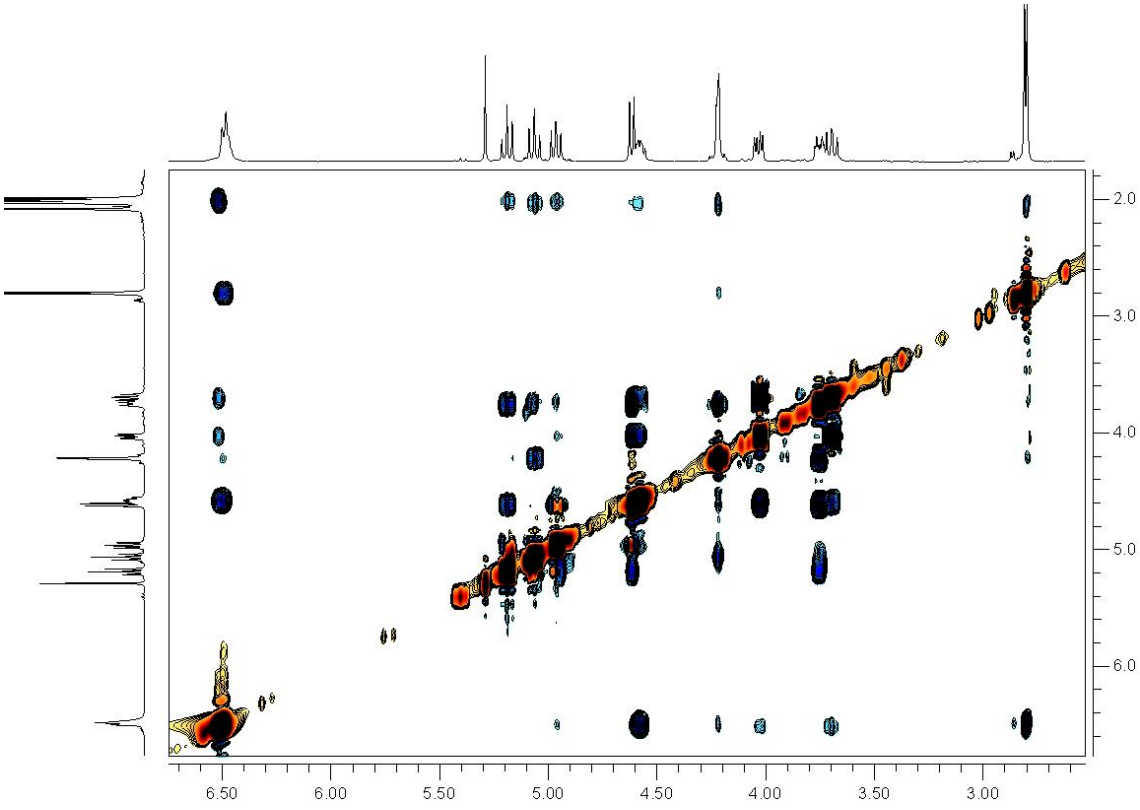


NOE Build-up curves for compound **Ac_4-1g** in $CDCl_3$

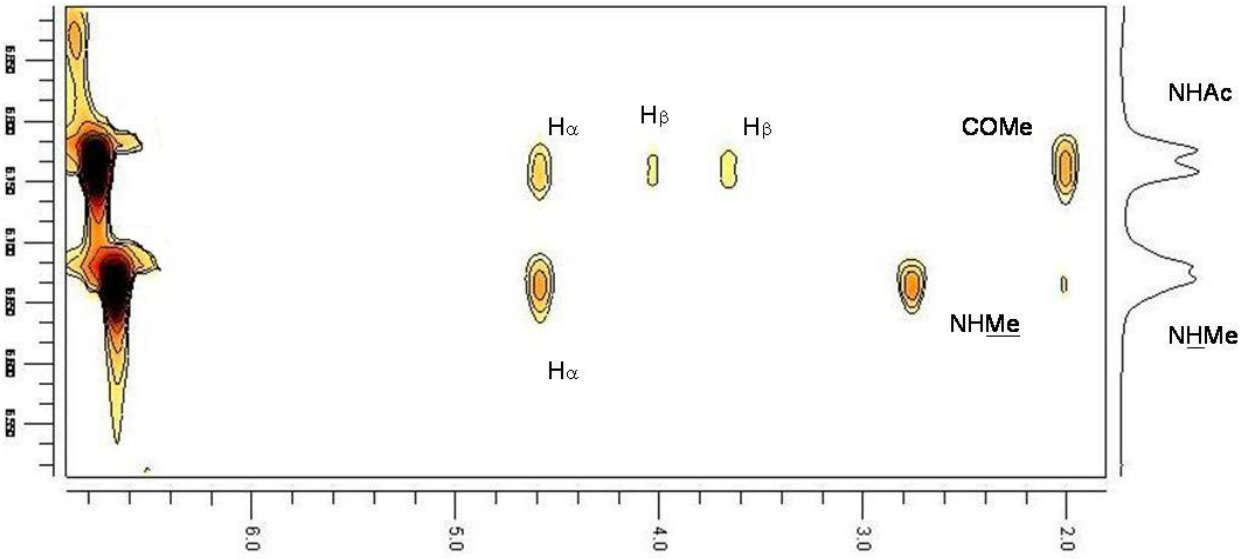




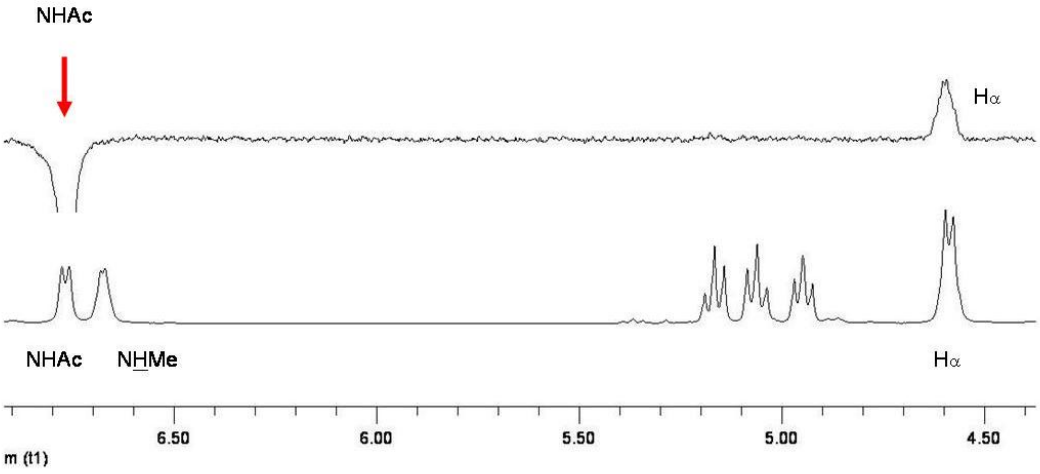
NOESY (mixing time = 800 ms 400 MHz) for compound **Ac_4-1g** in $CDCl_3$



NOESY (mixing time = 800 ms 400 MHz at –50 ºC) for compound **Ac_4-1g** in $CDCl_3$



Selective ROESY (mixing time = 800 ms 400 MHz at –50 ºC) for compound **Ac_4-1g** in $CDCl_3$



Analysis of molecular dynamics simulations using different force fields.

Calculations: *MD-tar simulations (dielectric constant=80).*– In order to obtain a NMR-derived ensemble, MD-tar simulations were performed for compounds **1**, **2**, **1g** and **2g**. NOE-derived distances (see Tables 1 and 2) were included as time averaged distance constraints, and scalar coupling constants *J* as time averaged coupling constraints. A $\langle r^{-6} \rangle^{-1/6}$ average was used for the distances and a linear average was used for the coupling constants. Final trajectories were run using an exponential decay constant of 8000 ps and a simulation length of 80 ns.

Molecular modelling in explicit water.– First, the solute molecule was immersed in a bath TIP3P water molecules[2] with the LEAP module.[3] The simulation was performed using periodic boundary conditions and the particle-mesh Ewald approach[4] to introduce long-range electrostatic effects. The SHAKE algorithm[5] for hydrogen atoms, which allows using a 2 fs time step, was employed. Finally, a 9 Å cutoff was applied to Lennard-Jones interactions. Equilibration of the system was carried out as follows; as a first step, a short minimization with positional restraints on solute atoms was run to remove any potentially bad contact. The force constant for the positional constraints was 500 Kcal mol^{-1} $Å^{-2}$. We ran then a 12.5 ps molecular dynamics calculation at 300 K maintaining positional restraints on the solute in order to equilibrate the water box. For these two steps, a 9 Å cutoff was used for the treatment of the electrostatic interactions. As a next step, the system was equilibrated using the mesh Ewald method, as water properties are slightly different with this treatment. With this purpose, a short MD simulation (12.5 ps) was run at 300 K, also using the Ewald approach for long-range electrostatic effects. Then, the system was subjected to several minimization cycles (each using 1000 steepest descent iterations) gradually reducing positional restraints on the solute from 500 Kcal mol^{-1} $Å^{-2}$ to 0. Finally, one unrestrained MD trajectory at constant pressure (1 atm) and temperature (300 K) was collected and analyzed using the CARNAL module.[6] The simulation length was 10 ns.

Table S1. Comparison of the experimental and MD-tar simulations derived distances and ^3J couplings obtained using different force fields for peptide **1**.[a]

	Exp[b]	ff94	ff99	ff99'
$d_{NH1,NH2}$	absent	3.0	3.2	3.3
$d_{H\alpha,NH1}$	s (2.3)	2.4	2.4	2.3
$d_{H\alpha,NH2}$	m (2.7)	2.9	2.9	2.9
$^3J_{H\alpha,H\beta}$[c]	5.7	5.4	5.0	5.3
$^3J_{NH2,H\alpha}$[d]	6.3	6.5	6.8	6.8

Table S2. Comparison of the experimental and MD-tar simulations derived distances and ^3J couplings obtained using different force fields for glycopeptides **1g**. [a]

	Exp[b]	ff94	ff99	ff99'
$d_{NH1,NH2}$	m (2.7)	2.5	2.5	2.5
$d_{H\alpha,NH1}$	s (2.3)	2.5	2.5	2.4
$d_{H\alpha,NH2}$	m (2.6)	2.9	2.9	2.9
$d_{H\beta proS,NH2}$	m (2.9)	2.7	2.9	2.9
$d_{H\beta proR,NH2}$	m (2.8)	2.8	3.0	2.9
$d_{H\beta proS,H\alpha}$	2.6	2.5	2.5	2.5
$d_{H\beta proR,H\alpha}$	2.6	2.5	2.5	2.5
$d_{H\beta proS,H1}$	2.6	2.6	2.4	2.4
$d_{H\beta proR,H1}$	2.3	2.4	2.6	2.6
$^3J_{H\alpha,H\beta}$[c]	4.6	5.1	5.3	5.4
$^3J_{NH2,H\alpha}$[d]	6.9	6.8	7.2	7.3

[a] Distances are given in Å and ^3J coupling in Hz. [b] w = weak, m = medium and s = strong NOE. [c] Estimated using the Karplus equation given in reference: A. Marco, M. Llinas and K. Wuthrich, *Biopolymers* **1978**, 17, 617–636. [d] Estimated using the Karplus equation given in reference: G. W. Vuister, A. Bax, *J. Am. Chem. Soc.* **1993**, 115, 7772–7777.

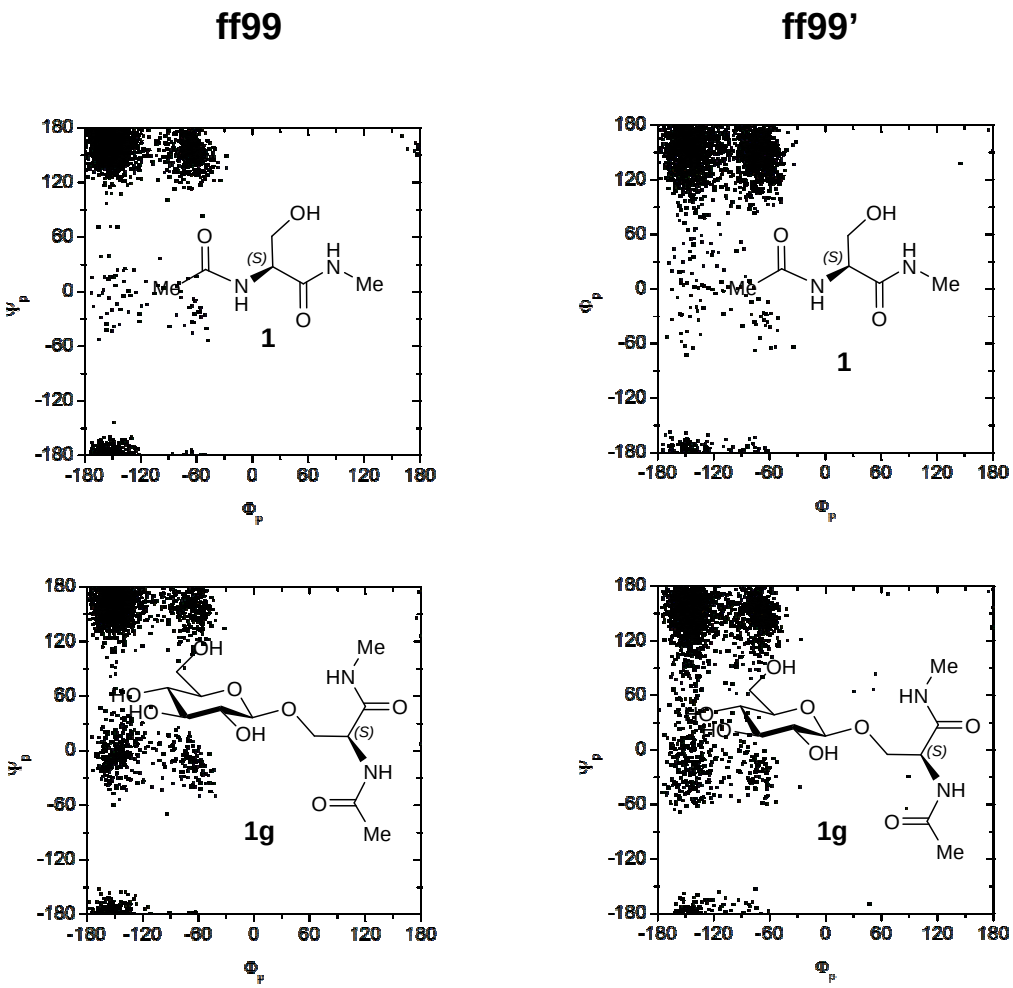


Figure S1 Comparison of the Φ_p/Ψ_p distributions obtained from the MD-tar simulations for **1** and **1g** using ff99 and ff99' force fields.

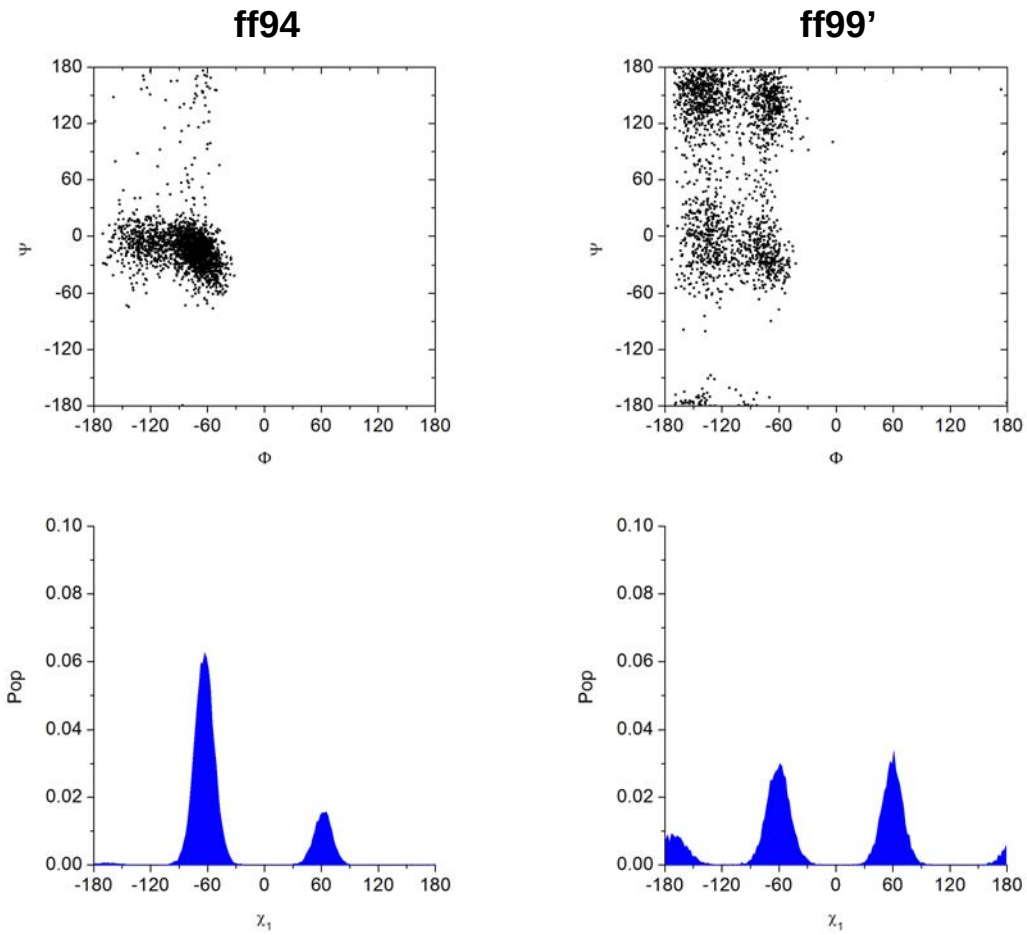


Figure S2 Comparison of the Φ_p/Ψ_p and χ^1 distributions obtained from the MD-free simulations in explicit water for **1** using ff94 and ff99' force fields.

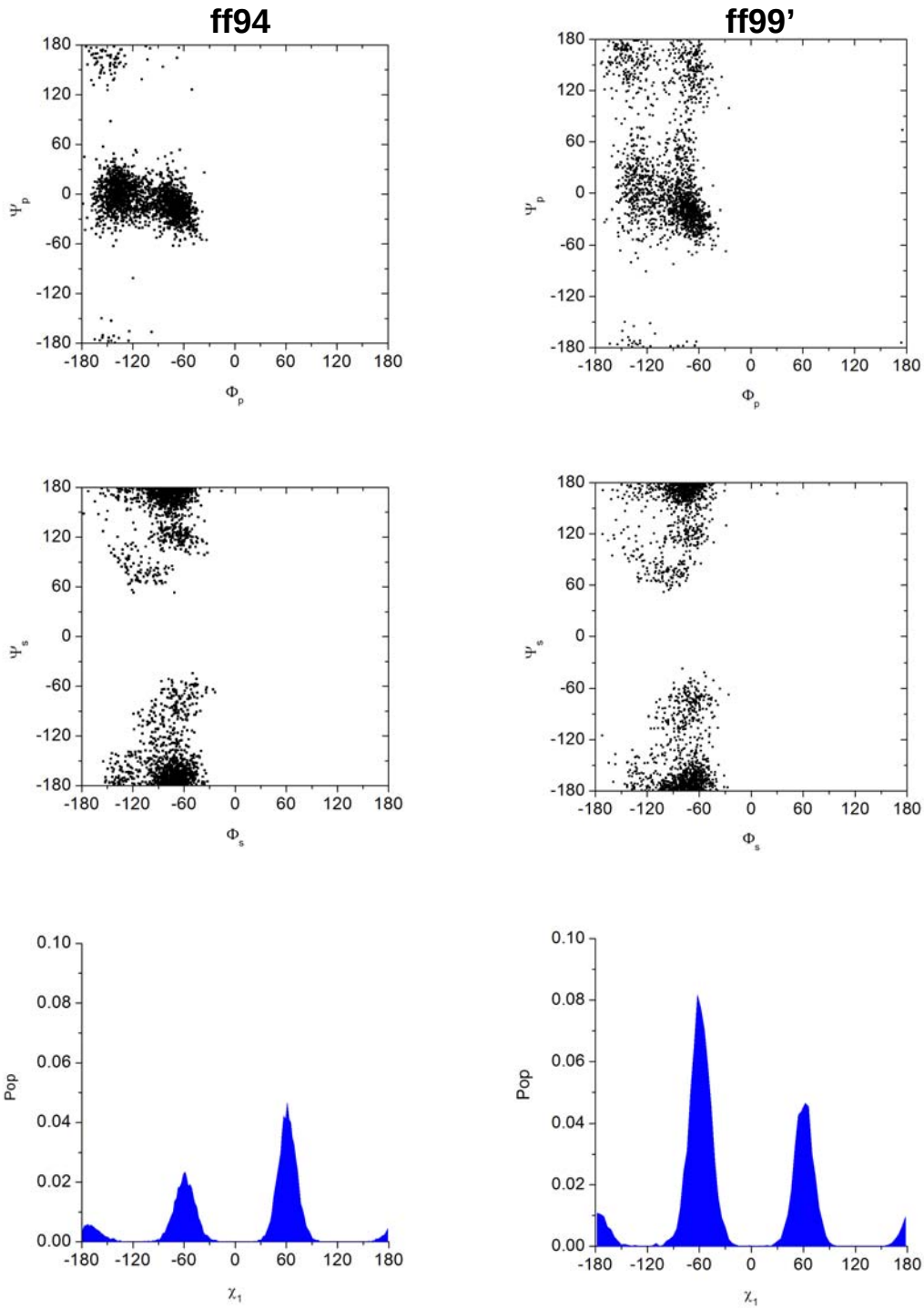


Figure S3 Comparison of the Φ_p/Ψ_p, Φ_s/Ψ_s and χ^1 distributions obtained from the MD-free simulations in explicit water for **1g** using ff94 and ff99' force fields.

Radial pair distribution function (RDF) of water and heteroatoms of glycopeptide 1g.

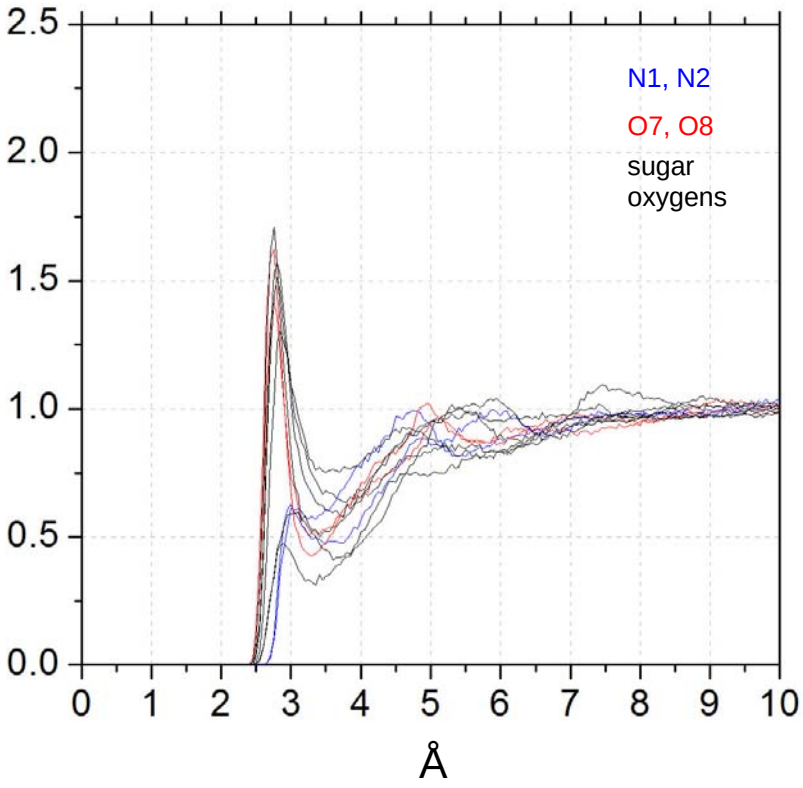


Figure S4 Radial pair distribution function (RDF) of water and heteroatoms of glycopeptide **1g** using ff94/GLYCAM04 force field.

Comparison of most relevant dihedrals obtained from the MD-tar simulations for Ac_4-1g in water and in chloroform solution.

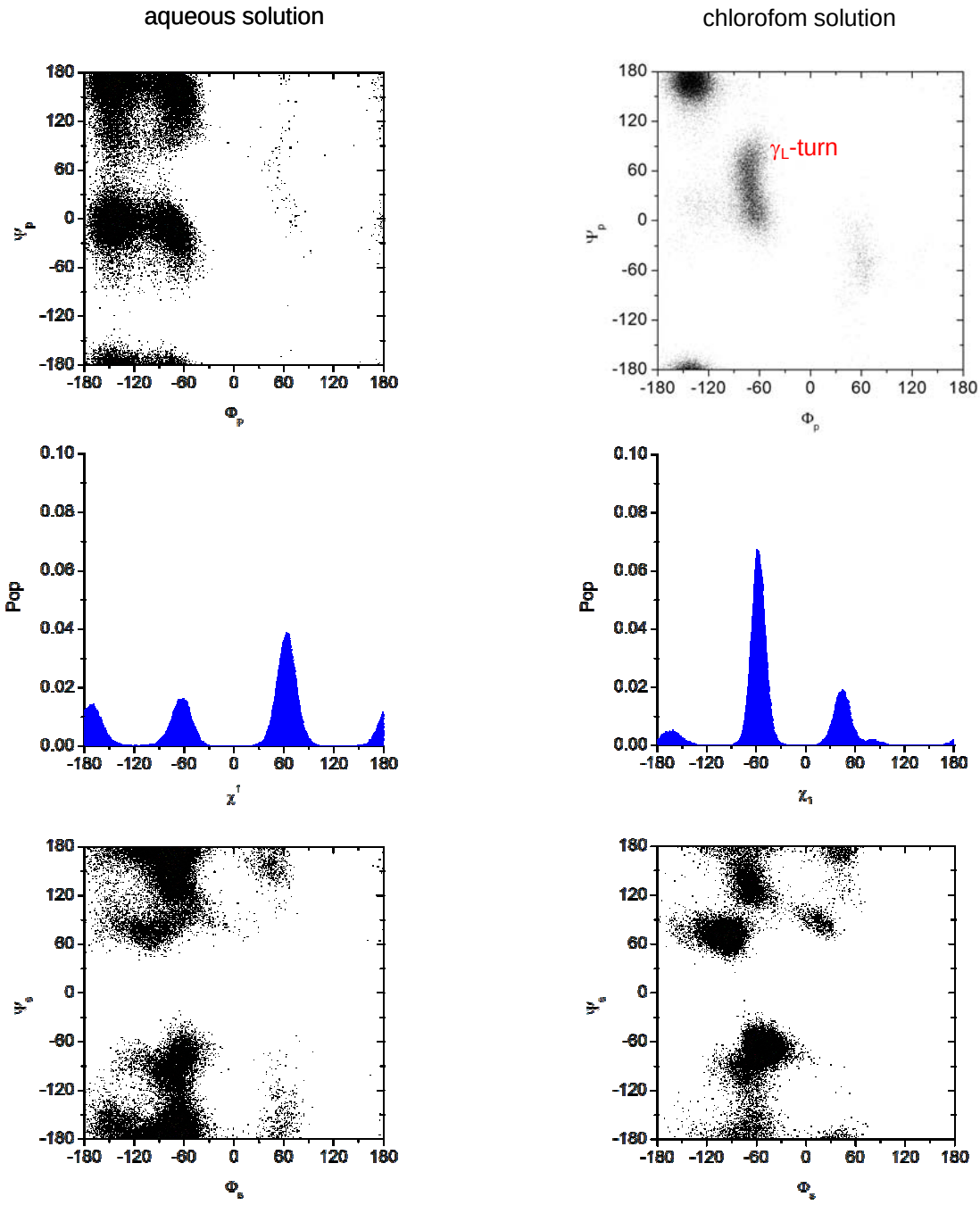


Figure S5 Comparison of most relevant dihedrals obtained from the MD-tar simulations (ff94/GLYCAM04) for **Ac4-1g** in water (ε=80) and in chloroform (ε=1) solution.

B3LYP/6-31G(d) energy, enthalpy, free energy, entropy and coordinates of the optimized structure of 1g.

All calculations were carried out by means of the B3LYP hybrid functional.[7] Full geometry optimizations were carried out with the 6-31G(d) basis set using the Gaussian 03 package.[8] BSSE corrections have not been considered in this work. Frequency analyses were carried out at the same level used in the geometry optimizations.

*glycopetide **1g***	E_0	H	S	G	Lowest freq.
	(Hartree)[a]	(Hartree)[b]	(cal mol^{-1} K^{-1})[b]	(Hartree)[b]	(cm^{-1})
	-1181.801003	-1181.409541	163.521	-1181.487235	34.88

[a] 1 Hartree = 627.5 Kcal mol^{-1}. [b] Thermal corrections at 298.15 K.

Cartesian coordinates of the optimized structure of **1g** (B3LYP/6-31G*)

C	2.74546100	2.32176000	-1.89098600
H	2.98385700	2.12434500	-2.94257400
H	3.51035100	2.97085600	-1.45981900
H	1.77984200	2.82899200	-1.83802500
N	2.69784200	1.08689300	-1.11369100
H	3.50037500	0.82408000	-0.54139300
C	1.85030800	0.08501300	-1.46937200
O	1.01852700	0.19151800	-2.36661200
C	2.04738800	-1.23737100	-0.67839100
N	2.18607300	-1.06718500	0.76273400
C	3.34441100	-0.72053200	1.37760600
O	4.36173700	-0.41788300	0.74485200
C	3.32153100	-0.72144200	2.89533100
H	2.34739000	-0.98347300	3.31819700
H	3.61861100	0.26775800	3.25659200
H	4.06823800	-1.43667300	3.25473000
H	1.33247100	-1.19297000	1.29270000
H	2.98861700	-1.66889300	-1.04546900
C	0.91834300	-2.24531500	-0.95037000
H	0.62058400	-2.17913500	-2.00060800
H	1.27638500	-3.25498000	-0.73543100

O	-0.21186900	-2.07860600	-0.09154300
C	-1.08152400	-1.02664900	-0.37791800
H	-1.05987800	-0.77731000	-1.44624300
O	-0.63900900	0.10635200	0.38602200
C	-1.42310800	1.29270900	0.19790600
C	-0.72421500	2.39733800	0.98199100
O	0.53937400	2.74195600	0.45468000
H	1.05762500	1.92483200	0.34896100
H	-1.33276300	3.30746300	0.93880500
H	-0.65212100	2.08391300	2.03804700
H	-1.43931500	1.56549700	-0.86794200
C	-2.85801300	1.01437200	0.67834800
O	-3.75773600	2.08997700	0.39439000
H	-3.66779800	2.76167900	1.08621800
H	-2.84442400	0.79324300	1.75625800
C	-3.40842400	-0.19809500	-0.06353900
O	-4.68790300	-0.58366200	0.42289800
H	-5.25226400	0.20344400	0.34532900
H	-3.47563000	0.06461400	-1.13261800
C	-2.49118800	-1.41172500	0.06996300
H	-2.43878200	-1.70964800	1.12932300
O	-2.95406300	-2.47583700	-0.73973200
H	-3.9070620	-2.54648200	-0.56140000

References

[1] T. Haselhorst, T. Weimar, T. Peters, *J. Am. Chem. Soc.* **2001**, *123*, 10705–10714.

[2] M. Chiansan, N. A. Baner, J. Simpson, J. A. McCammon, *J. Am. Chem. Soc.* **2002**, *124*, 1438–1442.

[3] C. E. A. F. Schafmeister, W. F. Ross, V. Romanovsky, University of California, San Francisco, 1995.

[4] C. S. A. T. Darden, *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155–179.

[5] J. P. Rykaert, G. Ciccote, J. C. Berendsen, *J. Comp. Phys.* **1977**, *23*, 327–341.

[6] W. S. Ross, Carnal: coordinate Analysis Program, University of California, San Francisco, Department of Pharmaceutical Chemistry.

[7] a) C. Lee, W. Yang, R. Parr, *Phys. Rev. B* **1988**, *37*, 785–789. b) A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[8] Gaussian 03, Revision B.05, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 2003.