Supporting Information

Transforming Flask Reaction into Cell-Based Synthesis: Production of Polyhydroxylated Molecules via Engineered *Escherichia coli*

Mohui Wei,[†,⊥] Zijie Li,[‡,⊥] Tiehai Li,[†] Baolin Wu,[†] Yunpeng Liu,[†] Jingyao Qu,[†] Xu Li,[†] Lei Li,[†] Li Cai,[§] Peng George Wang[*,†,∥]

[†]Center for Diagnostics and Therapeutics, Department of Chemistry, Georgia State University, Atlanta, Georgia 30303, United States, [‡]The Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214122, China, [§]Department of Chemistry, University of South Carolina Salkehatchie, Walterboro, South Carolina 29488, United States, [∥]State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy, Nankai University, Tianjin 300457, China. [⊥] Contributed equally to this work.

Corresponding author
[*]Email: pwang11@gsu.edu

Table of Contents

General information

^{1}H and ^{13}C NMR spectra were recorded on Bruker AV-400 MHz. Chemical shifts are expressed in ppm using residual $CDCl_3$ (7.26 ppm for ^{1}H NMR and 77 ppm for ^{13}C NMR) or CD_3OD (3.31 ppm for ^{1}H NMR and 49 ppm for ^{13}C NMR) or D_2O at 298 k as internal standard. High-resolution mass spectra were recorded under ESI-TOF Mass spectra conditions. Optical rotations were measured with JASCO P-1020 polarimeter. Analytical thin-layer chromatography (TLC) was performed on pre-coated plates (Silica Gel 60). Silica gel 60 (E. Merck) was employed for all flash chromatography. Reagents and starting materials obtained from commercial suppliers were used without further purification unless otherwise noted.

Bacterial strains and plasmids

Bacterial strains, plasmids and primers were summarized in Table S1. Pfx DNA polymerase was purchased from Invitrogen (CA, USA). Restriction enzymes and T4 ligase were purchased from Fermentas (MBI, Canada). *Thermus thermophiles* HB8 genomic DNA was purchased from ATCC (Manassas, VA). Plasmid pKKfda containing *fda* gene from *Staphylococcus carnosus* TM300 was a kind gift from professor Wolf-Dieter Fessner. Bio gel P-2 gel and Aminex HPX-87H column (300 × 7.8 mm) were purchased from Bio-Rad Laboratories, Inc. (Hercules, CA). XK column (100 × 2.6 cm) was purchased from GE Healthcare (Piscataway, NJ).

Table S1 Plasmids, strains and primers used in this study

Materials	Relevant genotype or primer sequence	Source
Plasmids		
pCDFDuet-1	CloDF13 ori lacI T7lac Str^r	Novagen
pCDF-Y	pCDFDuet-1 harboring *yqaB* gene from *E. coli* MG1655	This work
pCDF-fucA-Y	pCDFDuet-1 harboring *fucA* gene from *T. thermophiles* HB8 and *yqaB* gene from *E. coli* MG1655	This work
pKK-fda	Plasmid pKK223-3 harboring *fda* gene from *S. carnosus*	Fessner W.D et al.(1999)
pCDF-fda-Y	pCDFDuet-1 harboring *fda* gene from *S. carnosus* and *yqaB* gene from *E. coli* MG1655	This work
pCDF-rhuA-Y	pCDFDuet-1 harboring *rhuA* gene from *E.coli* MG1655 and *yqaB* gene from *E.coli* MG1655	This work
Strains		
DH5α	lacZΔM15 hsdR recA	Gibco-BRL
MG1655	F- λ- ilvG rfb-50 rph-1	Lab stock
BL21Star (DE3)	F- ompT hsdSB(rB- mB-) gal dcm rne131 (DE3)	Invitrogen
E.coli FucA-Y	BL21Star (DE3) harboring plasmid pCDF-fucA-Y	This work
E.coli FruA-Y	BL21Star (DE3) harboring plasmid pCDF-fda-Y	This work
E.coli RhuA-Y	BL21Star (DE3) harboring plasmid pCDF-rhuA-Y	This work
Primers		
pCDF-Y-F	5'-GCGCCATATGTACGAGCGTTATGCAGGTT-3'(*Nde*I)	
pCDF-Y-R	5'-TATACTCGAGCAGCAAGCGAACATCCACG-3'(*Xho*I)	
pCDF-fucA-F	5'-TATAGGATCCGCGCGCCCGGTTGTACG-3'(*BamH*I)	
pCDF-fucA-R	5'-TATAAAGCTTTCATTCCCCACCCCCCAAG-3'(*Hind*III)	
pCDF-fda-F	5'-GCGCGGATCCGAACCAAGAACAATT-3'(*BamH*I)	
pCDF-fda-R	5'-GCGCCTGCAGTTAAGCTTTGTTTACTGAA-3'(*Pst*I)	
pCDF-rhuA-F	5'-GCGTGGATCCGCAAAACATTACTCAGT-3'(*BamH*I)	
pCDF-rhuA-R	5'-TATAAAGCTTTTACAGCGCCAGCGCACT-3'(*Hind*III)	

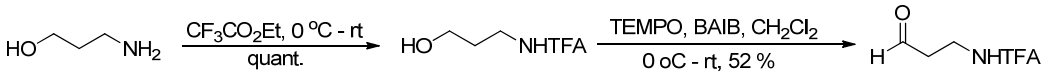
Construction of pCDF-fucA-Y, pCDF-fda-Y and pCDF-rhuA-Y plasmids

Primers pCDF-Y-F and pCDF-Y-R were used to amplify the *yqaB* gene by PCR using *E. coli* MG1655 genomic DNA as the template. The amplified *yqaB* gene was digested with *Nde*I and *Xho*I then inserted into the MCS-2 of pCDFDuet-1 plasmid with the same enzymes digested to generate plasmid pCDF-Y. Primers pCDF-fucA-F and pCDF-fucA-R were used to amplify *fucA* gene encoding *T. thermophiles* HB8

L-fuculose-1-phosphate aldolase by PCR using *T. thermophiles* HB8 genomic DNA as the template. The amplified *fucA* gene was digested with *BamH*I and *Hind*III then ligated into MCS-1 of plasmid pCDF-Y with the same enzymes digested to generate plasmid pCDF-fucA-Y. Primers pCDF-fda-F and pCDF-fda-R were used to amplify the *fda* gene encoding fructose-1, 6-bisphosphate aldolase from *S.carnosus* TM300 by PCR with plasmid pkk-fda as the template. The amplified *fda* was digested by *BamH*I and *Pst*I then inserted into MCS-1 of plasmid pCDF-Y with the same enzymes digested to generate plasmid pCDF-fda-Y. Primers pCDF-rhuA-F and pCDF-rhuA-R were used to amplify the gene *rhuA* encoding L-rhamnulose-1-phosphate aldolase with *E.coli* MG1655 as the template. The *rhuA* gene amplified was digested with *BamH*I and *Hind*III then inserted into MCS-1 of plasmid pCDF-Y with the same enzymes digested to generate plasmid pCDF-rhuA-Y. These recombinant plasmids were all transformed into DH5α strain for amplifying and sequencing. *pCDF-fucA-Y, pCDF-fda-Y and pCDF-rhuA-Y plasmids* were transformed into *E.coli* strain BL21Star (DE3) respectively, resulting in the recombinant strains *E.coli* FucA-Y, *E.coli* FruA-Y and *E.coli* RhuA-Y.

Synthesis of aldehyde acceptors

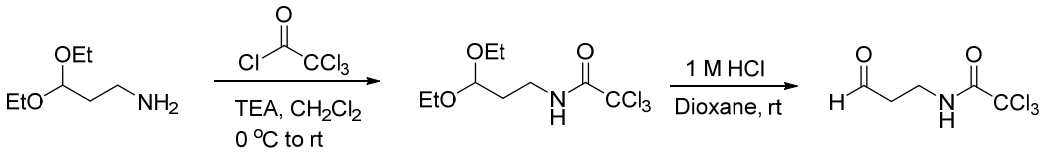
3-trifluoroacetamido propanal



To a stirred solution of 3-aminopropan-1-ol (75 g, 1 mol, 1 equiv), ethyl trifluoroacetate (177.5 g, 1.5 mol, 1.25 equiv) was added dropwise at 0 °C. After completion of the addition, the mixture was warmed to room temperature and stirred overnight.[1] When 3-aminopropan-1-ol was completely consumed, the resulting mixture was evaporated under reduced pressure to afford 3-trifluoroacetamido propan-1-ol, which was used for the TEMPO oxidation directly.

To a stirred solution of 3-trifluoroacetamido propan-1-ol (171 g, 1 mol, 1 equiv) in CH_2Cl_2 (1000 mL), TEMPO (15.6 g, 0.1 mol, 0.1 equiv) was added at 0 °C, followed by bis(acetoxy)iodobenzene (BAIB) (354 g, 1.1 mol, 1.1 equiv) in small portions. After the addition, the reaction mixture was warmed to room temperature[2] and stirred at rt for 3 days. The reaction mixture was extracted with water (200 mL x 3). The aqueous phase was washed with hexane, then saturated with NaCl and extracted with CH_2Cl_2 (500 mL x 3). The organic phase was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The residue was purified by vacuum distillation to give 3-trifluoroacetamido propanal 87.88 g with 52 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ: 9.82 (s, 1H), 6.91 (s, 1H), 3.65 (dt, J = 6.0, 5.6 Hz, 2H), 2.83 (t, J = 5.6 Hz, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ: 200.6, 157.3 (q, J = 37.0 Hz), 115.7 (q, J = 286.0 Hz), 42.5, 33.3; ^{19}F NMR (376 MHz, $CDCl_3$) δ: -76.1 (s, 3F).

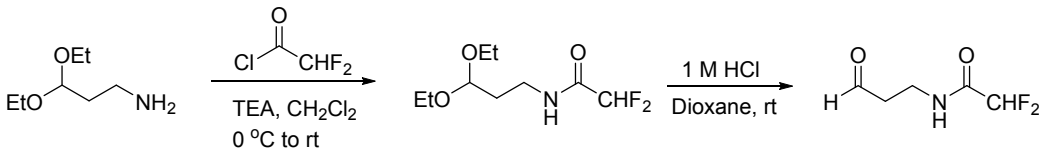
3-trichloroacetamido propanal



To the cooled (0 °C) solution of 3-aminopropionaldehyde diethylacetal (14.7 g, 100 mmol, 1 equiv) and TEA (11.13 g, 110 mmol, 1.1 equiv) in CH_2Cl_2 (150 mL), a solution of trichloroacetyl chloride (19.09 g, 105 mmol, 1.05 equiv) in CH_2Cl_2 (50 mL) was added dropwise. After completion of addition, the reaction mixture was allowed to warm to rt and stirred overnight.[3] The resulting mixture was washed with water, dried over Na_2SO_4, and concentrated to provide 3-trichloroacetamido propionaldehyde diethylacetal. The crude 3-trichloroacetamido propionaldehyde diethylacetal was sufficiently pure and was used without further purification.

The mixture of 3-trichloroacetamido propionaldehyde diethylacetal in 1 M HCl in dioxane (100 mL) was stirred at room temperature until completion of the reaction (TLC monitored). The resulting reaction mixture was concentrated and the residue was dissolved in ethyl acetate, then washed with water, dried over Na_2SO_4, and concentrated *in vacuo* to afford 14.64 g 3-trichloroacetamido propanal with 67 % yield, which was sufficiently pure and was used for fermentation without further purification. ^{1}H NMR (400 MHz, $CDCl_3$) δ: 9.82 (s, 1H), 7.30 (s, 1H), 3.64 (dt, J = 6.0, 6.0 Hz, 2H), 2.84 (t, J = 5.8 Hz, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ: 200.8, 162.0, 67.0, 42.6, 34.8.

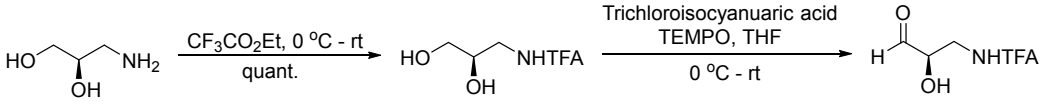
3-difluoroacetamido propanal



To the cooled (0 °C) solution of 3-aminopropionaldehyde diethylacetal (14.7 g, 100 mmol, 1 equiv) and TEA (11.13 g, 110 mmol, 1.1 equiv) in CH_2Cl_2 (150 mL), a solution of difluoroacetyl chloride (12.02 g, 105 mmol, 1.05 equiv) in CH_2Cl_2 (50 mL) was added dropwise. After completion of addition, the reaction mixture was allowed to warm to rt and stirred overnight. The resulting mixture was washed with water, dried over Na_2SO_4, and concentrated to provide 3-difluoroacetamido propionaldehyde diethylacetal. The crude 3-difluoroacetamido propionaldehyde diethylacetal was sufficiently pure and was used without further purification.

The mixture of 3-difluoroacetamido propionaldehyde diethylacetal in 1 M HCl in dioxane (100 mL) was stirred at room temperature until completion of the reaction (TLC monitored). The resulting reaction mixture was concentrated and the residue was dissolved in ethyl acetate, then washed with water, dried over Na_2SO_4, and concentrated *in vacuo* to afford 9.52 g 3-difluoroacetamido propanal with 63 % yield, which was sufficiently pure and was used for fermentation without further purification. ^{1}H NMR (400 MHz, $CDCl_3$) δ: 9.77 (s, 1H), 7.03 (s, 1H), 5.85 (t, J = 54 Hz, 1H), 3.62-3.57 (m, 2H), 2.78 (s, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ: 200.7, 162.7 (t, J = 25.3 Hz), 108.2 (t, J = 252.3 Hz), 42.8, 32.7.

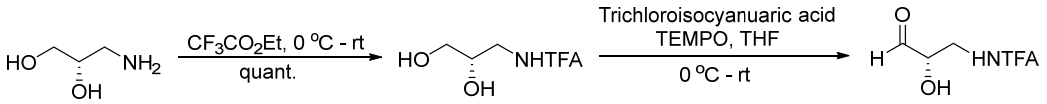
(*R*)-3-trifluoroacetamido-2-hydroxypropanal



To a stirred solution of (*R*)-3-aminopropane-1, 2-diol (5 g, 54.95 mmol, 1 equiv) in methanol (50 mL), ethyl trifluoroacetate (9.75 g, 68.68 mmol, 1.25 equiv) was added dropwise at 0 °C. After the addition, the mixture was warmed to room temperature and stirred overnight.[1] When (*R*)-3-aminopropane-1, 2-diol was completely consumed, the resulting mixture was evaporated under reduced pressure to afford (*R*)-3-trifluoroacetamido propan-1, 2-diol quantitatively, which was used for the IBX oxidation directly.

To a stirred solution of (*R*)-3-trifluoroacetamido propan-1, 2-diol (10.27 g, 54.95 mmol, 1 equiv) in EtOAc (500 mL), 2-iodoxybenzoic acid (76.3 g, 47 %, 82.5 mmol, 1.5 equiv) was added and refluxed overnight.[4] The reaction mixture was quenched by filtering through celite. DD water (50 mL) was added to the filtrate and the organic solvent was removed under reduced pressure. The pH of the aldehyde solution was adjusted to 5 with $NaHCO_3$. After sterilization by filtering through 0.2 μm, nylon, sterile membrane, the resulting aldehyde solution was used for fermentation directly. The estimated yield was around 50 %.

(*S*)-3-trifluoroacetamido-2-hydroxypropanal



The (*S*)-3-trifluoroacetamido-2-hydroxypropanal was prepared analogously from commercial (*S*)-3-aminopropane-1, 2-diol. The estimated yield was around 50 %.

Media and fermentation procedures

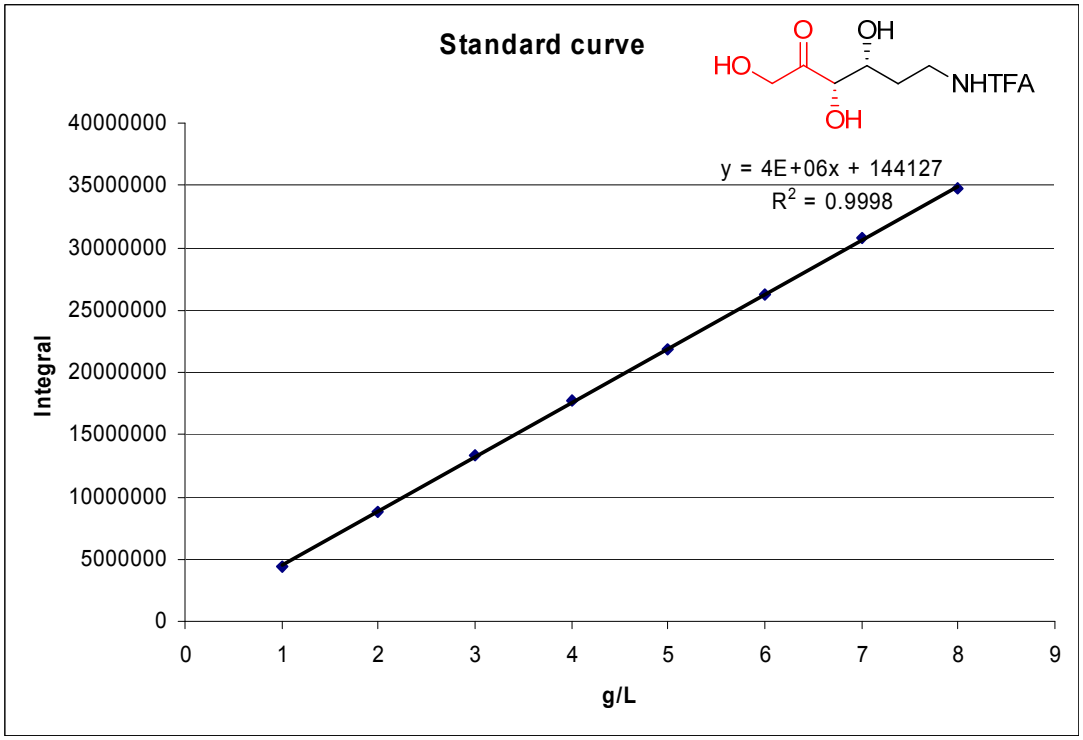
20 mL overnight cultured recombinant *E.coli* strain was inoculated to 1 L LB Broth medium (LB broth 25 g/L, $NaH_2PO_4·H_2O$ 10.78 g/L, $Na_2HPO_4·7H_2O$ 17.32 g/L, $MgSO_4$ 120.00 mg/L, $ZnSO_4$ 32.30 mg/L, $CaCl_2$ 11.10 mg/L, thiamine 10.00 mg/L, streptomycin 50.00 mg/L, glucose 4.00 g/L), or ECAM medium ($NaH_2PO_4·H_2O$ 10.78 g/L, $Na_2HPO_4·7H_2O$ 17.32 g/L, KCl 4.27 g/L, $(NH_4)_2SO_4$ 2.33 g/L, citric acid 1 g/L, $MgSO_4$ 1 g/L, $CaCl_2$ 40 mg/L, thiamine 10 mg/L, EDTA 5 mg/L, $FeSO_4·7H_2O$ 10 mg/L, $ZnSO_4·7H_2O$ 2 mg/L, $MnSO_4·H_2O$ 2 mg/L, $CoCl_2·6H_2O$ 0.2 mg/L, $CuSO_4·5H_2O$ 0.1 mg/L, $Na_2MoO_4·2H_2O$ 0.2 mg/L, H_3BO_3 0.1 mg/L, streptomycin 50.00 mg/L, glucose 4.00 g/L) or LB Broth/ECAM (1/1, v/v) medium (LB Broth 12.5 g/L, $NaH_2PO_4·H_2O$ 10.78 g/L, $Na_2HPO_4·7H_2O$ 17.32 g/L, KCl 2.14 g/L, $(NH_4)_2SO_4$ 1.17 g/L, citric acid 0.5 g/L, $MgSO_4$ 0.5 g/L, $CaCl_2$ 20 mg/L, thiamine 5 mg/L, EDTA 2.5 mg/L, $FeSO_4·7H_2O$ 5 mg/L, $ZnSO_4·7H_2O$ 1 mg/L, $MnSO_4·H_2O$ 1 mg/L, $CoCl_2·6H_2O$ 0.1 mg/L, $CuSO_4·5H_2O$ 0.05 mg/L, $Na_2MoO_4·2H_2O$ 0.1 mg/L, H_3BO_3 0.05 mg/L, streptomycin 50.00 mg/L, glucose 4.00 g/L) and grown aerobically at 37 ^oC, 220 rpm until the OD_{600} reached 1.80. Then the temperature was switched to 30 ^oC and isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added at a final concentration of 1 mM to induce the co-expression of aldolase and phosphatase for 12 h. Then aldehyde aqueous solution (sterilized by filtering through 0.2 μm, nylon, sterile membrane) was added into the medium with an initial concentration of 20 mM. The consumption of glucose, aldehyde and generation of aldol product in the medium were monitored by HPLC (HPX-87H Ion Exchange Column, column temperature 60

°C, 5 mM H_2SO_4, 0.5 mL/min). Aldehyde (totally 40 mmol) and glucose were fed when necessary. The fermentation was stopped when the concentration of the aldol product reached a plateau.

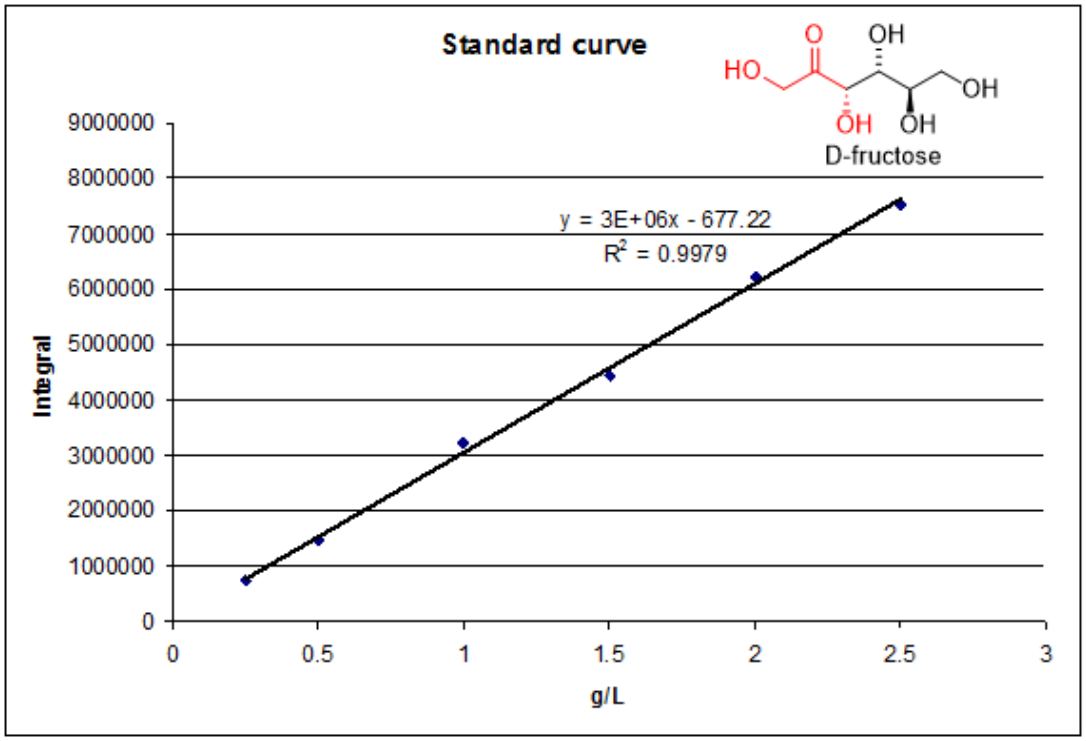
When 3-trifluoroacetamido propanal, 3-trichloroacetamido propanal, 3-difluoroacetamido propanal, 3-(methylthio)propanal, 4,4,4-trifluorobutanal, (*R*)-3-trifluoroacetamido-2-hydroxypropanal and (*S*)-3-trifluoroacetamido-2-hydroxypropanal were used as the acceptor, LB Broth/ECAM (1/1, v/v) medium was used; when D-glyceraldehyde was used as the acceptor, ECAM medium was used for easier purification of aldol products.

Analytical methods

The progress of the fermentation was monitored by HPLC. Samples were taken at regular intervals to monitor the consumption of aldehyedes and generation of aldol products. After centrifugation, the supernatant was applied to HPLC column (Aminex HPX-87H, 300 × 7.8 mm) with 5 mM sulfuric acid as mobile phase and detected with Refractive Index Detector. The flow rate was 0.5 mL/min, and the column temperature was 60 °C.



Standard curve of (3*S*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one.



Standard curve of D-fructose.

Product purification

The fermentation medium was centrifuged to remove *E. coli* cells and 2 L acetone was added to the supernatant to precipitate nucleic acid *etc*. When 3-trifluoroacetamido propanal, 3-trichloroacetamido propanal, 3-difluoroacetamido propanal, (*R*)-3-trifluoroacetamido-2-hydroxypropanal and (*S*)-3-trifluoroacetamido-2-hydroxypropanal were used as the acceptor, after removal of precipitate, the pH value of resulting supernatant was adjusted to 3 by concentrated hydrochloric acid, then silica gel was added and concentrated under reduced pressure. The resulting residue was loaded onto a pad of silica gel and washed with EtOAc/MeOH/HOAc (20/0.5/0.1, v/v/v). The filtrate was concentrated and the residue was purified by C-18 reverse phase silica gel column chromatography (water as the eluent), followed by silica gel column chromatography (CH_2Cl_2/MeOH, 10/1). For (*R*)-3-trifluoroacetamido-2-hydroxypropanal and (*S*)-3-trifluoroacetamido-2-hydroxypropanal, CH_2Cl_2/MeOH (6/1, v/v) was used to wash the residue. The filtrate was concentrated and the resulting residue was purified by preparative HPLC (C-18 reverse phase, mobile phase, 5 mM TFA).

When 3-(methylthio)propanal and 4,4,4-trifluorobutanal were used as the acceptor, after removal of precipitate, silica gel was added and the supernatant was concentrated under reduced pressure. The resulting residue was loaded onto a pad of silica gel and washed with EtOAc/MeOH/HOAc (20/0.5/0.1, v/v/v). The filtrate was concentrated and the residue was purified by C-18 reverse phase silica gel column

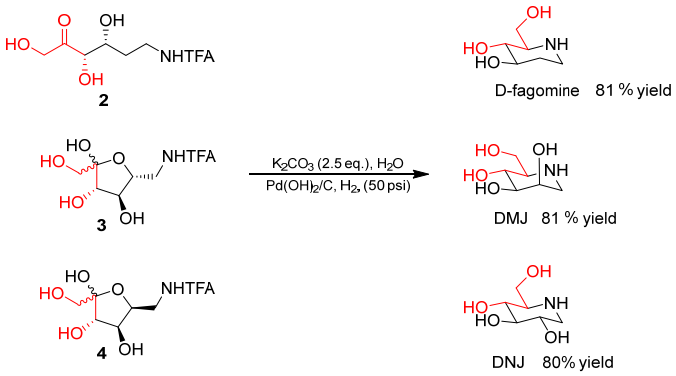
chromatography (water as the eluent), followed by silica gel column chromatography (CH_2Cl_2/MeOH, 15/1).

When D-glyceraldehyde was used as the acceptor, after removal of precipitate, silica gel was added and the supernatant was concentrated under reduced pressure. The resulting residue was purified by silica gel column chromatography (EtOAc/iPrOH/H_2O 9/3/1 (v/v/v)), followed by Bio gel P-2 column, a mixture of D-psicose and D-sorbose (or D-fructose and glucose) was obtained. This mixture was isolated by cation exchange resin column as described below.

Procedure for isolation of D-sorbose and D-psicose using cation exchange resin (Ca^{2+} form)

D-Sorbose and D-psicose mixture was dissolved in 3 mL ddH_2O and applied to a cation exchange resin (Ca^{2+} form, 100 × 2.6 cm) which was preheated to 65°C using a thermostatic jacket. The column was eluted with ddH_2O (flow rate ~1.5 mL/min) and the whole isolation process was performed at 65-70°C. Fractions were collected with an automatic fraction collector and identified by HPLC. D-Sorbose was eluted off first and D-psicose was eluted off in later fractions. Fractions containing pure D-sorbose or D-psicose were pooled and lyophilized to give pure D-sorbose and D-psicose. Pure D-fructose could also be isolated from glucose according to the procedure described above using cation exchange resin.

Synthesis of D-fagomine, DMJ, DNJ



D-fagomine:

To a solution of **2** (259 mg, 1 mmol) and K_2CO_3 (345 mg, 2.5 mmol) in 10 mL water, $Pd(OH)_2/C$ (10 %) was added. The mixture was hydrogenated (50 psi H_2) at room temperature overnight. The mixture was filtered through 0.45 µm nylon membrane filter. The filtrate was concentrated and the residue was purified by basic Al_2O_3 column chromatography (THF/MeOH/water/NH_4OH) to give 119 mg D-fagomine with 81 % yield.

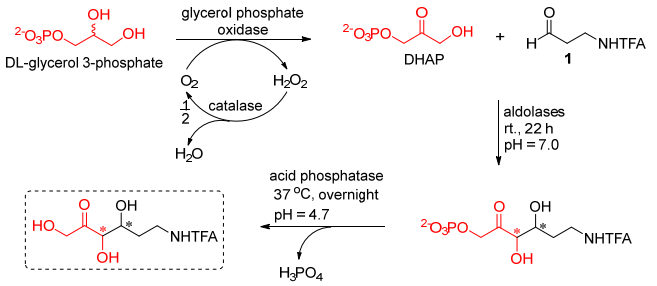
DMJ:

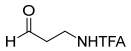
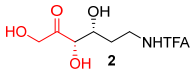
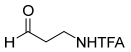
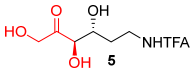
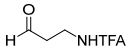
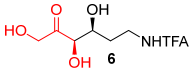
Following the same procedure, **3** (150 mg, 0.55 mmol) was hydrogenated to give 72 mg DMJ with 81 % yield.

DNJ:

Following the same procedure, **4** (150 mg, 0.55 mmol) was hydrogenated to give 71 mg DNJ with 81 % yield.

One-pot four-enzyme synthesis of 2, 5, 6

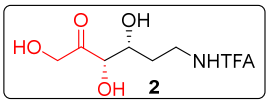


entry	aldolase	acceptor	product	isolated yield	*dr*[a]
1	FruA		 **2**	53 %	> 95:5
2	FucA		 **5**	48 %	87:13
3	RhuA		 **6**	42 %	89:11

To a solution of DL-glycerol 3-phosphate magnesium salt (457.3 mg, 2 mmol) in 10 mL ddH2O, aldehyde **1** (253.5 mg, 1.5 mmol) was added at pH 7.0, followed by glycerol phosphate oxidase (70 U, 2 mg), catalase (1000 U, 1.2 μL), and aldolase (FruA from *Staphylococcus carnosus*, FucA from *Thermus. thermophilus* HB8 or RhuA from *E.* coli, final concentration 0.5 mg/mL). The mixture was shaken at rt for 22 h and the reaction was monitored by TLC (developed by n-BuOH/AcOH/H2O: 2/1/1 (v/v/v) and stained with anisaldehyde sugar stain). The reaction mixture was heated for 10 min at 75 °C, then cooled to rt. Then, pH was adjusted to 4.7 with 6 M HCl and 11 μL acid phosphatase (18 U) was added and the mixture was shaken overnight at 37 °C. After cooling to rt, the pH was adjusted to 7.0 with 1 M NaOH and the mixture was diluted with methanol. The solution was filtered through celite and washed with methanol. The filtrate was concentrated under reduced pressure and the residue was purified by silica gel column chromatography (dichloromethane/methanol: 15/1 (v/v)) to afford product. After purification, 137.4 mg **2** was provided with 53 % yield, 124.4 mg **5** was provided with 48 % yield and 87:13 *dr*, and 108 mg **6** was provided with 42 % yield and 89:11 *dr*. Yields were calculated based on L-glycerol 3-phosphate (1 mmol).

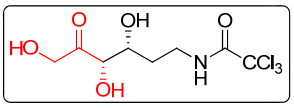
Product characterizations

(3*S*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one



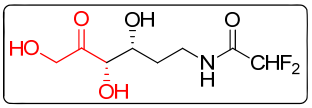
^{1}H NMR (400 MHz, CD_3OD) δ: 4.54 (d, *J* = 19.2 Hz, 1H), 4.44 (d, *J* = 19.6 Hz, 1H), 4.15 (s, 1H), 3.97 (t, *J* = 5.6 Hz, 1H), 3.46-3.38 (m, 2H), 1.85-1.80 (m, 2H); ^{13}C NMR (100 MHz, CD_3OD) δ: 213.3, 159.0 (q, *J* = 36.5 Hz), 117.5 (q, *J* = 284.7 Hz), 79.4, 71.1, 67.9, 37.9, 33.3; ^{19}F NMR (376 MHz, CD_3OD) δ: -77.4 (s, 3F); HRMS (ESI): $[M-H]^-$ calcd. for $C_8H_{11}F_3NO_5$, 258.0595; found, 258.0592.

(3*S*, 4*R*)-6-trichloroacetamido-1,3,4-trihydroxyhexan-2-one



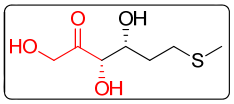
^{1}H NMR (400 MHz, CD_3OD) δ: 4.50 (d, *J* = 19.2 Hz, 1H), 4.40 (d, *J* = 19.2 Hz, 1H), 4.13 (s, 1H), 3.99-3.94 (m, 1H), 3.40 (t, *J* = 6.8 Hz, 2H), 1.82 (dt, *J* = 6.8, 6.8 Hz, 2H); ^{13}C NMR (100 MHz, CD_3OD) δ: 213.1, 164.2, 79.3, 71.2, 67.8, 39.5, 33.2; HRMS (ESI): $[M+Na]^+$ calcd. for $C_8H_{12}Cl_3NNaO_5$, 329.9673; found, 329.9668.

(3*S*, 4*R*)-6-difluoroacetamido-1,3,4-trihydroxyhexan-2-one



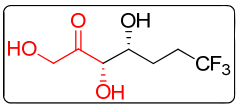
^{1}H NMR (400 MHz, CD_3OD) δ: 6.00 (t, *J* = 54.0 Hz, 1H), 4.51 (d, *J* = 19.3 Hz, 1H), 4.41 (d, *J* = 19.3 Hz, 1H), 4.13 (s, 1H), 3.98-3.93 (m, 1H), 3.36 (t, *J* = 6.8 Hz, 2H), 1.73 (dt, *J* = 6.8, 6.8 Hz, 2H); ^{13}C NMR (100 MHz, CD_3OD) δ: 213.3, 165.1 (t, *J* = 25.0 Hz), 109.9 (t, *J* = 240.7 Hz), 79.2, 71.1, 67.7, 37.3, 33.4; HRMS (ESI): $[M+Na]^+$ calcd. for $C_8H_{13}F_2NNaO_5$, 264.0654; found, 264.0648.

(3*S*, 4*R*)-1,3,4-trihydroxy-6-(methylthio)hexan-2-one



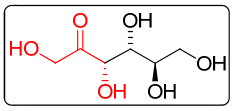
^{1}H NMR (400 MHz, CD_3OD) δ: 4.54 (d, *J* = 19.2 Hz, 1H), 4.44 (d, *J* = 19.2 Hz, 1H), 4.14 (d, *J* = 2 Hz, 1H), 4.07-4.04 (m, 1H), 2.67-2.52 (m, 2H), 2.09 (s, 3H), 1.88-1.81 (m, 2H); ^{13}C NMR (100 MHz, CD_3OD) δ: 213.4, 79.4, 72.2, 67.8, 33.7, 31.3, 15.3; HRMS (ESI): $[M+Na]^+$ calcd. for $C_7H_{14}NaO_4S$, 217.0505; found, 217.0500.

(3*S*, 4*R*)-7,7,7-trifluoro-1,3,4-trihydroxyheptan-2-one



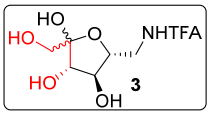
^{1}H NMR (400 MHz, CD_3OD) δ: 4.54 (d, *J* = 19.6 Hz, 1H), 4.45 (d, *J* = 19.6 Hz, 1H), 4.14 (s, 1H), 3.96 (s, 1H), 2.41-2.29 (m, 1H), 2.27-2.16 (m, 1H), 1.88-1.74 (m, 2H); ^{13}C NMR (100 MHz, CD_3OD) δ: 213.2, 128.9 (q, *J* = 273.4 Hz), 79.3, 72.0, 67.8, 31.2 (q, *J* = 28.6 Hz), 26.9; HRMS (ESI): $[M-H]^-$ calcd. for $C_7H_{10}F_3O_4$, 215.0537; found, 215.0536.

D-fructose



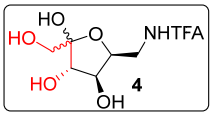
See reference 5.

3



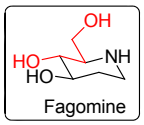
^{1}H NMR (400 MHz, D_2O) δ: 4.12-4.07 (m, 2H), 3.97-3.91 (m, 1H), 3.66-3.61 (m, 2H), 3.59-3.52 (m, 2H); ^{13}C NMR (100 MHz, D_2O) δ: 158.2 (q, *J* = 37.2 Hz), 114.8 (q, *J* = 284.1 Hz), 100.7, 77.0, 74.7, 74.1, 61.4, 41.0; ^{19}F NMR (376 MHz, CD_3OD) δ: -75.8 (s, 3F); HRMS (ESI): $[M-H]^-$ calcd. for $C_8H_{11}NO_6F_3$, 274.0540; found, 274.0538.

4



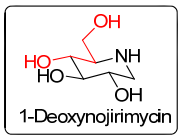
^{1}H NMR (400 MHz, D_2O) δ: 4.38-4.21 (m, 2H), 4.09-3.99 (m, 1H), 3.55-3.44 (m, 3H), 3.43-3.36 (m, 1H); ^{13}C NMR (100 MHz, D_2O) δ: 158.1 (q, *J* = 37.2 Hz), 114.9 (q, *J* = 284.1 Hz), 101.0, 74.6, 74.4, 74.0, 61.7, 39.1; ^{19}F NMR (376 MHz, CD_3OD) δ: -75.8 (s, 3F) ; HRMS (ESI): $[M-H]^-$ calcd for $C_8H_{11}NO_6F_3$, 274.0540; found, 274.0538.

D-fagomine



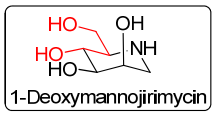
$[\alpha]_D^{25}$ = + 18.6 (c = 0.5 in H2O); 1H NMR (400 MHz, CD3OD) δ: 3.86 (dd, J = 11.0, 3.0 Hz, 1H), 3.60 (dd, J = 11.0, 6.6 Hz, 1H), 3.40 (ddd, J = 11.2, 8.6, 5.0 Hz, 1H), 3.10 (t, J = 9.2 Hz, 1H), 3.01 (ddd, J = 12.8, 4.4, 2.4 Hz, 1H), 2.62 (td, J = 12.8, 2.8 Hz, 1H), 2.47-2.42 (m, 1H), 1.96-1.89 (m, 1H), 1.52-1.42 (m, 1H); ^{13}C NMR (100 MHz, CD_3OD) δ: 75.1, 74.9, 63.2, 63.1, 44.5, 34.6; HRMS (ESI): $[M+H]^+$ calcd. for $C_6H_{14}NO_3$, 148.0968; found, 148.0967.

1-Deoxynojirimycin (DNJ)



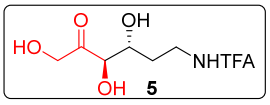
$[\alpha]_D^{25}$ = + 35.7 (c = 0.3 in H_2O); ^{1}H NMR (400 MHz, D_2O) δ: 3.85 (dd, *J* = 11.6, 3.0 Hz, 1H), 3.68 (dd, *J* = 11.6, 6.0 Hz, 1H), 3.44 (ddd, *J* = 10.8, 9.2, 5.2 Hz, 1H), 3.27 (t, *J* = 9.0 Hz, 1H), 3.18 (t, *J* = 9.4 Hz, 1H), 3.07 (dd, *J* = 12.2, 4.8 Hz, 1H), 2.68 (ddd, *J* = 9.6, 6.4, 2.8 Hz, 1H), 2.40 (t, *J* = 11.6 Hz, 1H),; ^{13}C NMR (100 MHz, D_2O) δ: 81.2, 74.1, 73.5, 64.0, 63.5, 51.4; HRMS (ESI): $[M+H]^+$ calcd. for $C_6H_{14}NO_4$, 164.0917; found, 164.0922.

1-Deoxymannojirimycin (DMJ)



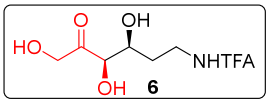
$[\alpha]_D^{25}$ = - 38.2 (c = 0.4 in H_2O); ^1H NMR (400 MHz, D_2O) δ: 3.96 (d, *J* = 1.5 Hz, 1H), 3.73 (d, *J* = 4.0 Hz, 2H), 3.56 (dd, *J* = 19.2, 9.5 Hz, 1H), 3.52 (dd, *J* = 9.5, 2.9 Hz, 1H), 2.97 (dd, *J* = 14.4, 2.7 Hz, 1H), 2.73 (dd, *J* = 14.0, 1.2 Hz, 1H), 2.45 (dt, *J* = 9.6, 4.0 Hz, 1H); ^{13}C NMR (100 MHz, D_2O) δ: 76.6, 71.2, 70.4, 62.8, 62.6, 50.3; HRMS (ESI): $[M+H]^+$ calcd. for $C_6H_{14}NO_4$, 164.0917; found, 164.0922.

(3*R*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one



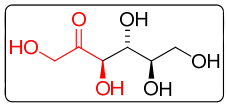
^1H NMR (400 MHz, CD_3OD) δ: 4.54 (d, *J* = 19.6 Hz, 1H), 4.44 (d, *J* = 19.2 Hz, 1H), 4.15 (d, *J* = 1.2 Hz, 1H), 3.97 (t, *J* = 5.6 Hz, 1H), 3.45-3.38 (m, 2H), 1.88-1.78 (m, 2H),; ^{13}C NMR (100 MHz, CD_3OD) δ: 213.3, 159.0 (q, *J* = 36.5 Hz), 117.5 (q, *J* = 284.7 Hz), 79.4, 71.1, 67.9, 37.9, 33.3; ^{19}F NMR (376 MHz, CD_3OD) δ: -77.4 (s, 3F); HRMS (ESI): $[M-H]^-$ calcd. for $C_8H_{11}F_3NO_5$, 258.0595; found, 258.0592.

(3*R*, 4*S*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one



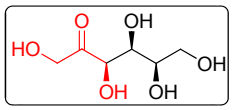
^1H NMR (400 MHz, CD_3OD) δ: 4.51 (d, *J* = 19.2 Hz, 1H), 4.41 (d, *J* = 19.6 Hz, 1H), 4.12 (s, 1H), 3.94 (t, *J* = 5.6 Hz, 1H), 3.41-3.36 (m, 2H), 1.84-1.74 (m, 2H); ^{13}C NMR (100 MHz, CD_3OD) δ: 213.3, 159.0 (q, *J* = 36.5 Hz), 117.5 (q, *J* = 284.7 Hz), 79.3, 71.1, 67.8, 37.9, 33.3; ^{19}F NMR (376 MHz, CD_3OD) δ: -77.4 (s, 3F); HRMS (ESI): $[M-H]^-$ calcd. for $C_8H_{11}F_3NO_5$, 258.0595; found, 258.0592.

D-psicose



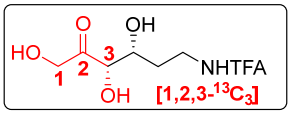
See references 6 and 7.

D-sorbose



See references 6 and 7.

[1, 2, 3-$^{13}C_3$] (3*S*,4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one



^{1}H NMR (400 MHz, CD_3OD) δ: 4.75-4.60 (m, 1H), 4.38-4.24 (m, 1.5H), 4.00-3.95 (m, 1.5H), 3.46-3.39 (m, 2H), 1.84-1.80 (m, 2H); ^{13}C NMR (100 MHz, CD_3OD) δ: 213.2 (dd, *J* = 42.7 41.4 Hz), 159.0 (q, *J* = 36.7 Hz), 117.5 (q, *J* = 284.8 Hz), 79.3 (dd, *J* = 43.1, 12.9 Hz), 71.1 (d, *J* = 39.6 Hz), 67.8 (dd, *J* = 40.9, 12.9 Hz), 37.9, 33.3; HRMS (ESI): $[M+Na]^+$ calcd. for $C_5{}^{13}C_3H_{12}F_3NO_5$, 285.0660; found, 285.0656.

References

1) Tang, R.; Ji, W.; Wang, C. *Polymer* **2011**, *52*, 921-932.

2) Einhorn, J.; Einhorn, C.; Ratajczak, F.; Pierre, J.-L. *J. Org. Chem.* **1996**, *61,* 7452-7454.

3) Kalisiak, J.; Trauger, S. A.; Kalisiak, E.; Morita, H.; Fokin, V. V.; Adams, M. W. W.; Sharpless, K. B.; Siuzdak, G. *J. Am. Chem. Soc.* **2009**, *131*, 378-386.

4) Concia, A. L.; Lozano, C.; Castillo, J. A.; Parella, T.; Joglar, J.; Clapés, P. *Chem. Eur. J.* **2009**, *15*, 3808-3816.

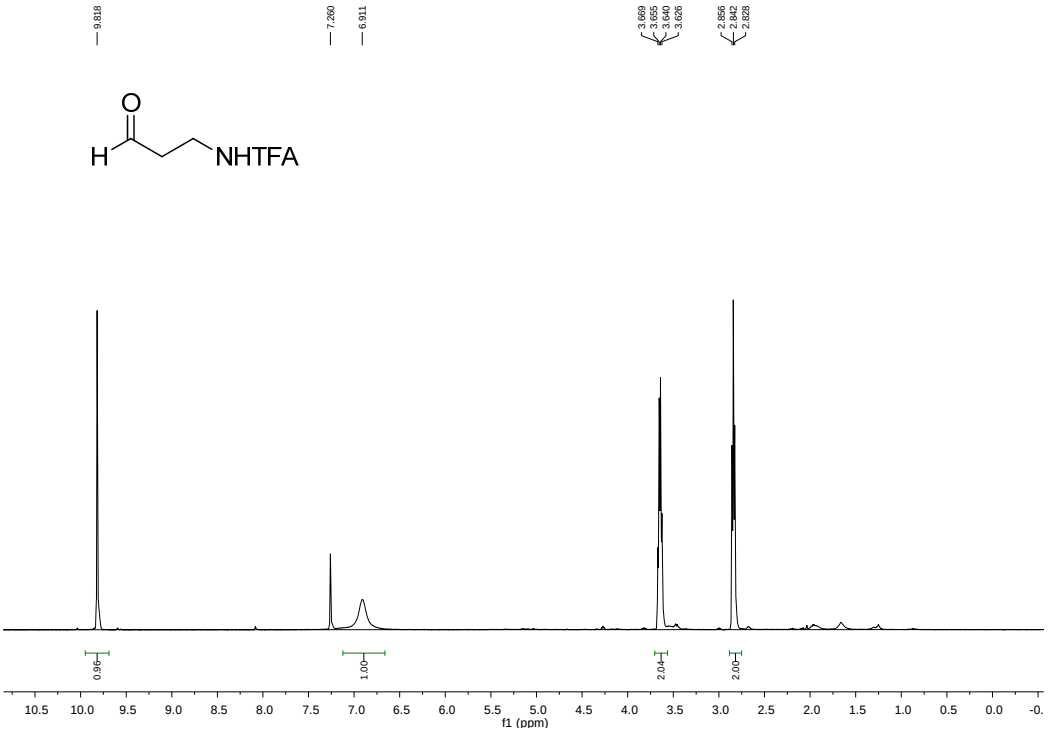
5) Barclay, T.; Ginic-Markovic, M; Johnston, M. R.; Cooper, P.; Petrovsky, N. *Carbohydrate Research* **2012**, *347*, 136-141.

6) Li, Z.; Cai, L.; Qi, Q.; Styslinger, T. J.; Zhao, G.; Wang, P. G. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 5084-5087.

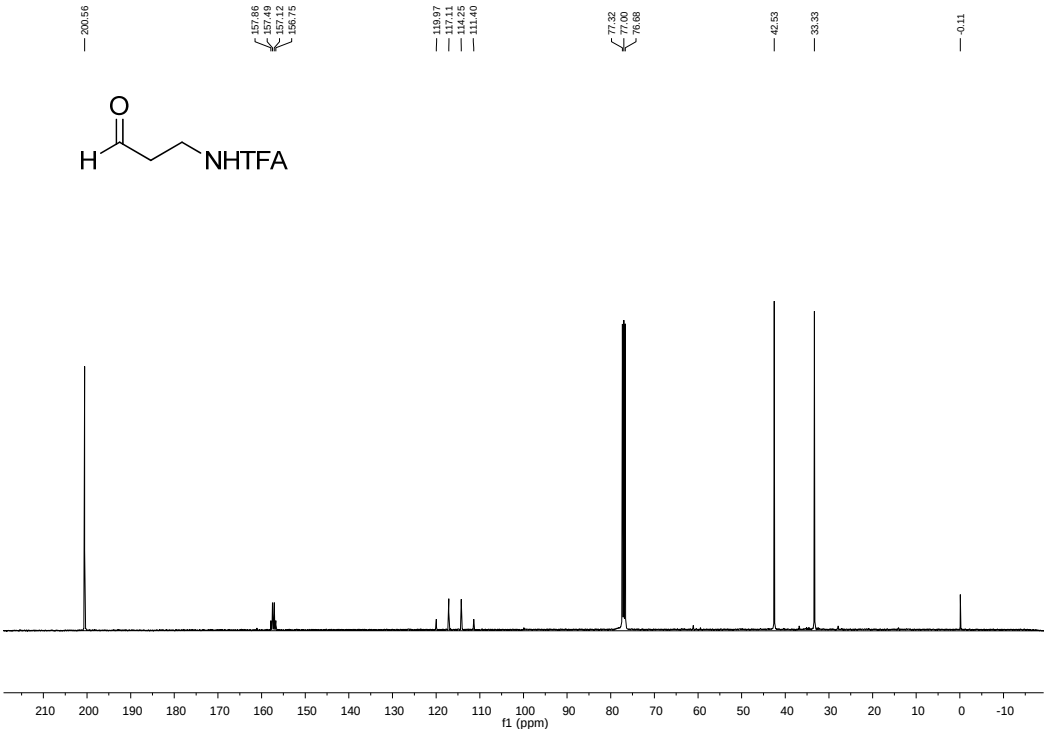
7) Li, Z.; Cai, L.; Qi, Q.; Wang, P. G. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 7081-7084.

NMR spectra

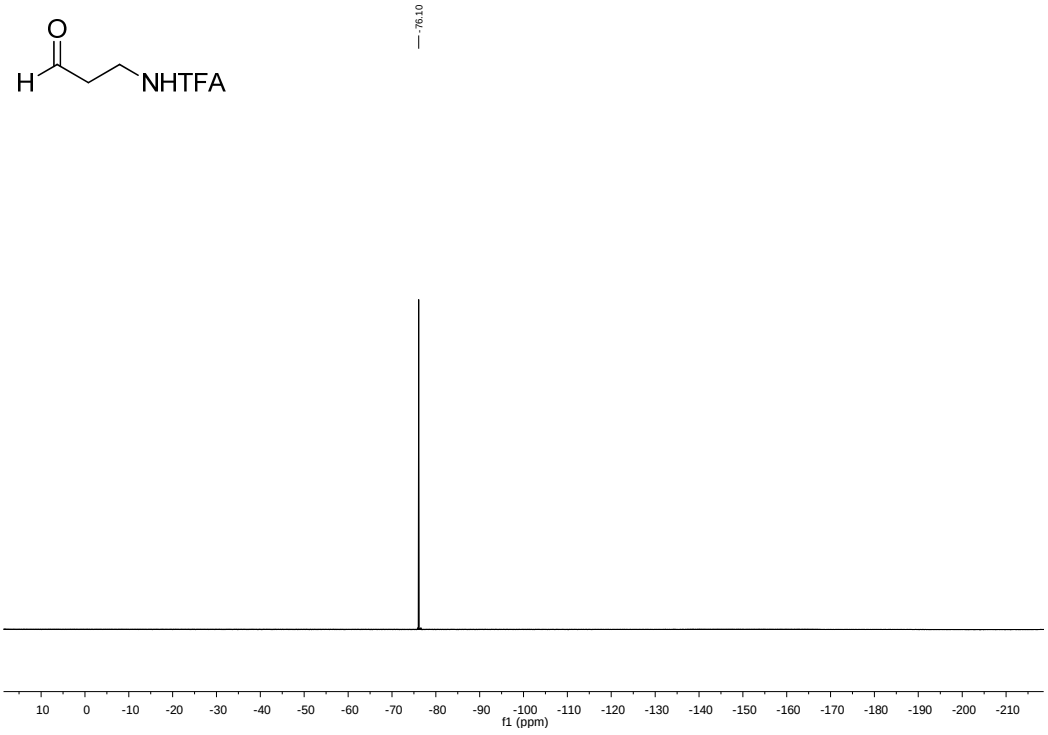
^{1}H-NMR spectrum of 3-trifluoroacetamido propanal in $CDCl_3$.



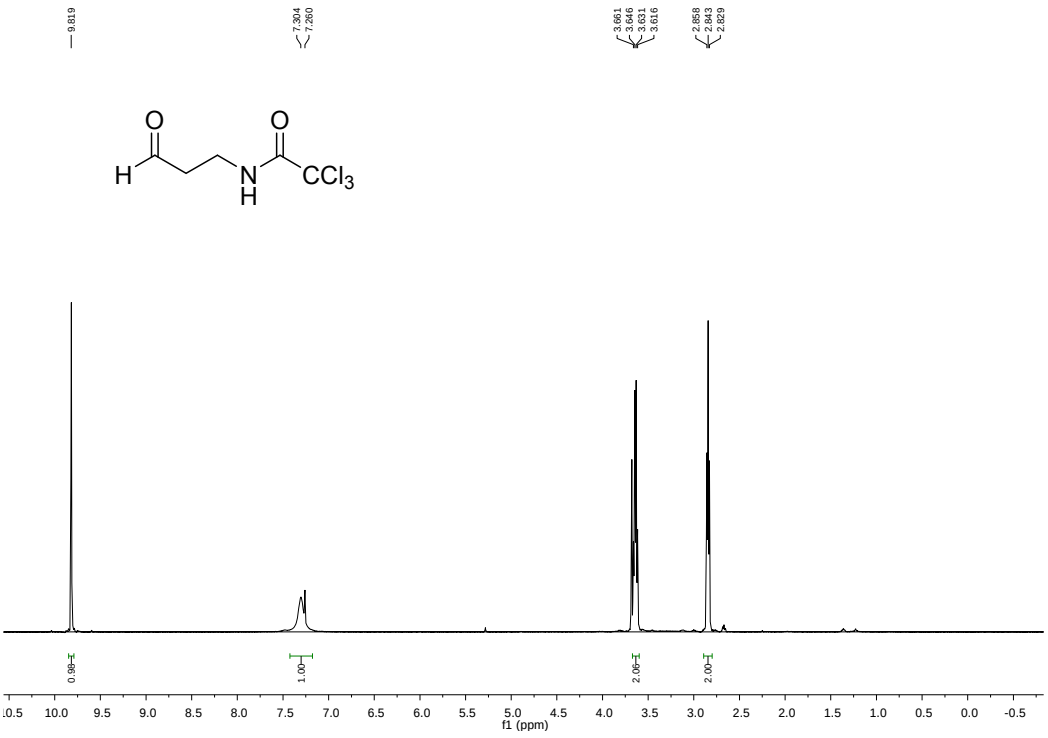
^{13}C-NMR spectrum of 3-trifluoroacetamido propanal in $CDCl_3$.



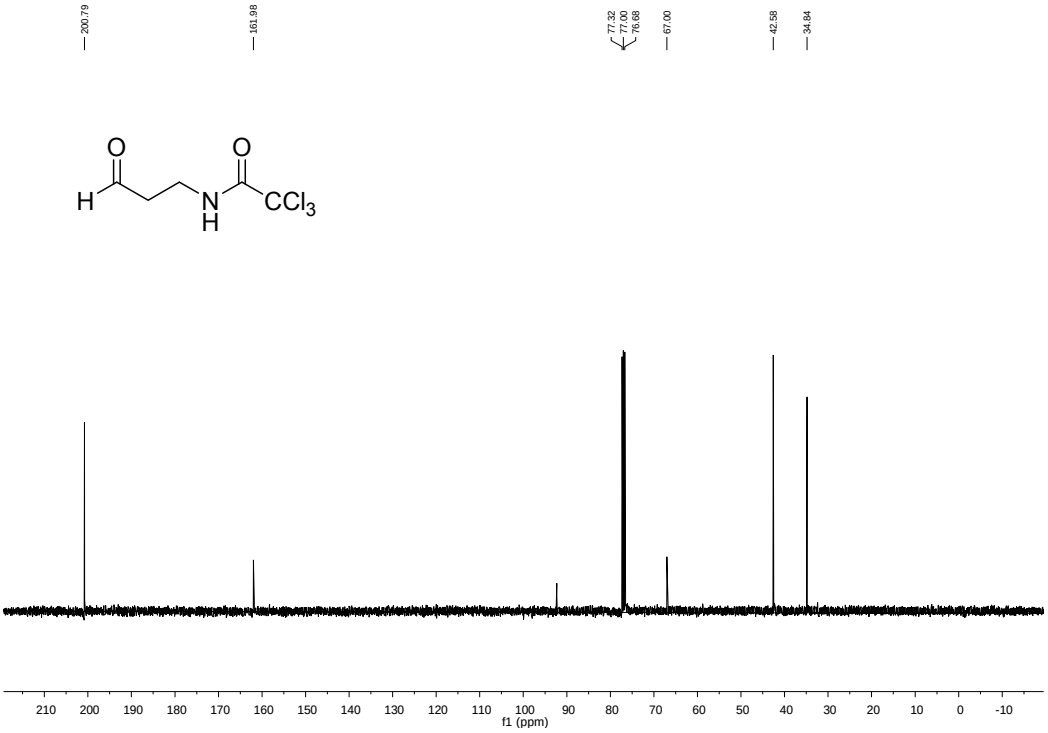
^{19}F-NMR spectrum of 3-trifluoroacetamido propanal in $CDCl_3$.



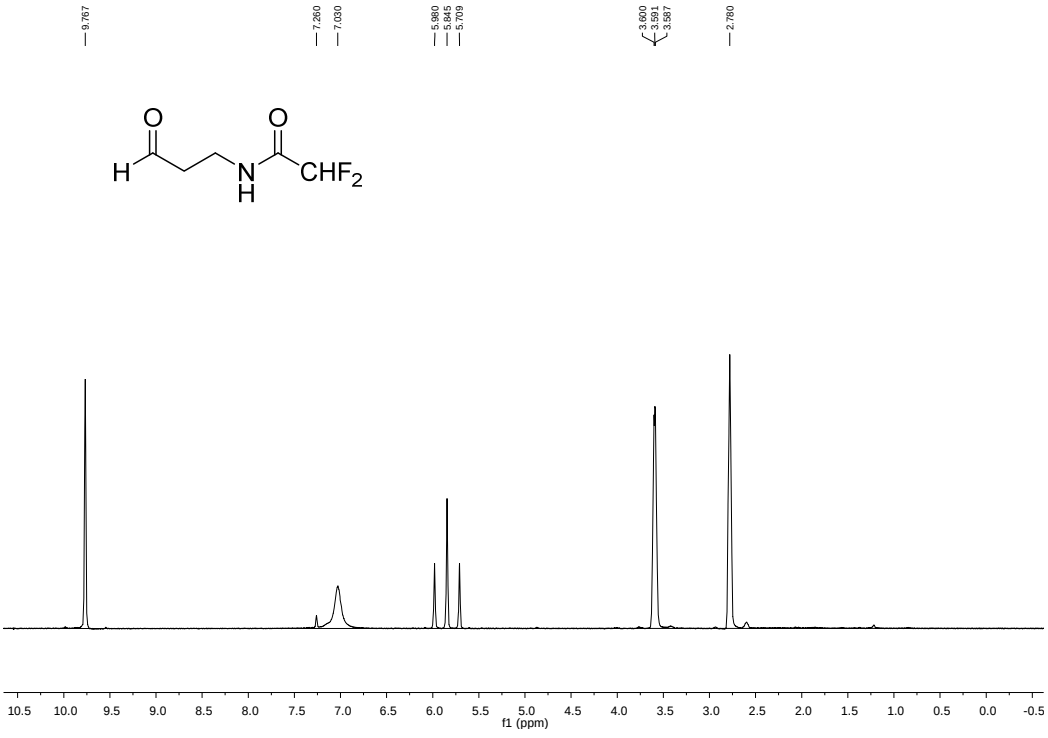
^{1}H-NMR spectrum of 3-trichloroacetamido propanal in $CDCl_3$.



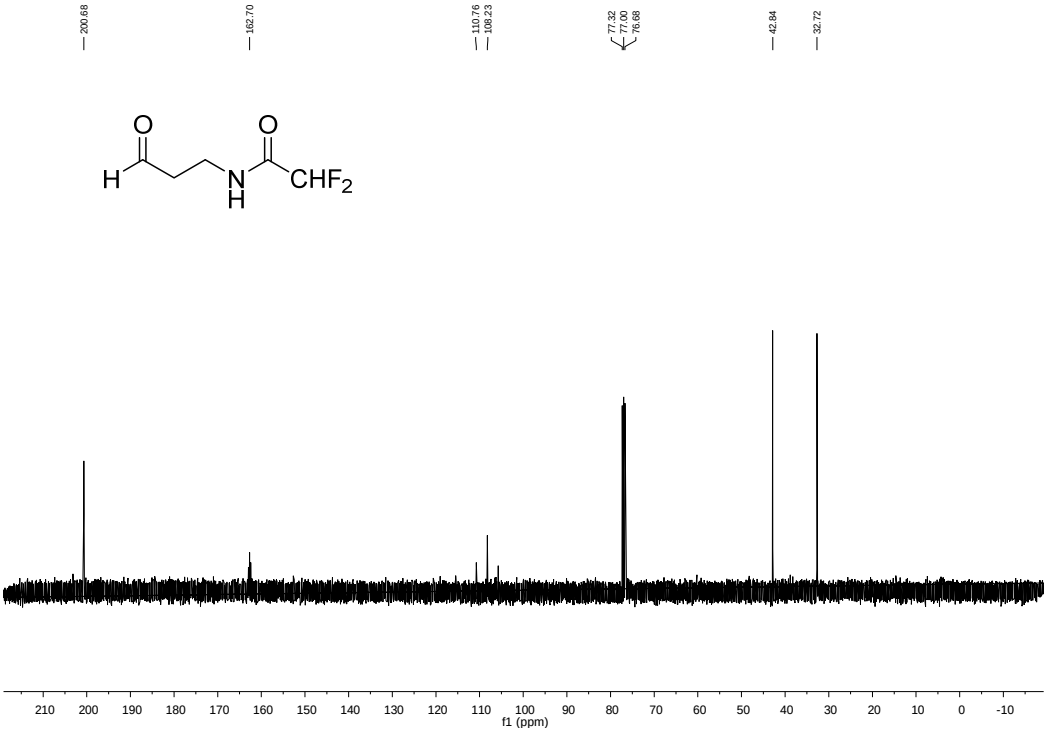
^{13}C-NMR spectrum of 3-trichloroacetamido propanal in $CDCl_3$.



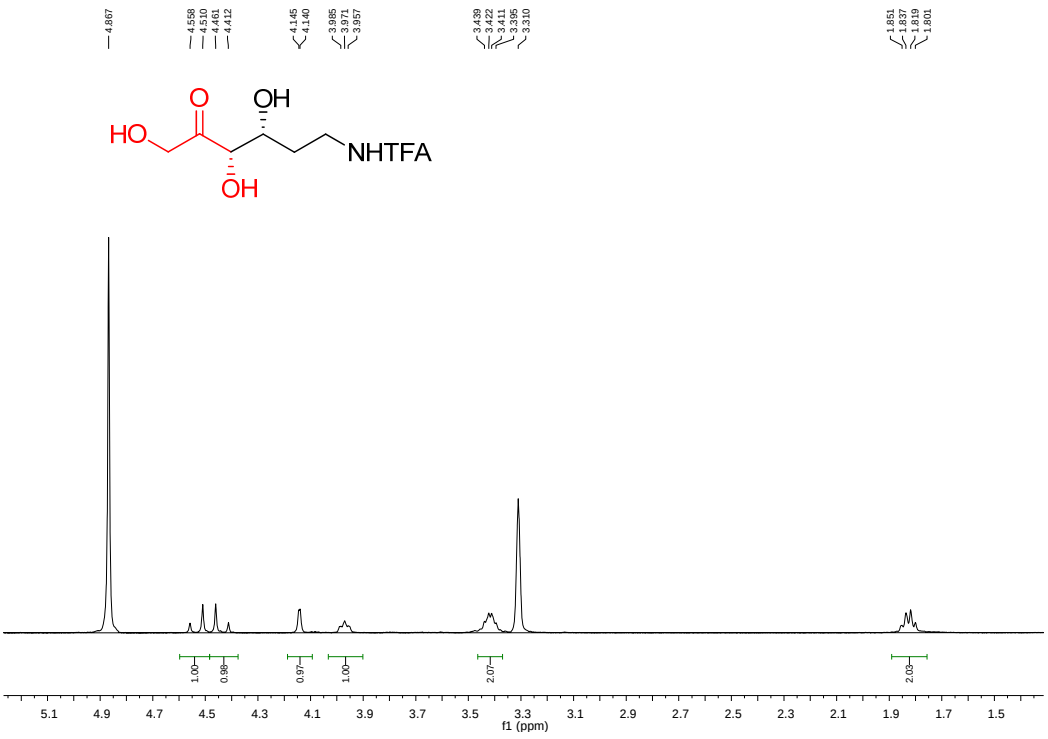
^{1}H-NMR spectrum of 3-difluoroacetamido propanal in $CDCl_3$.



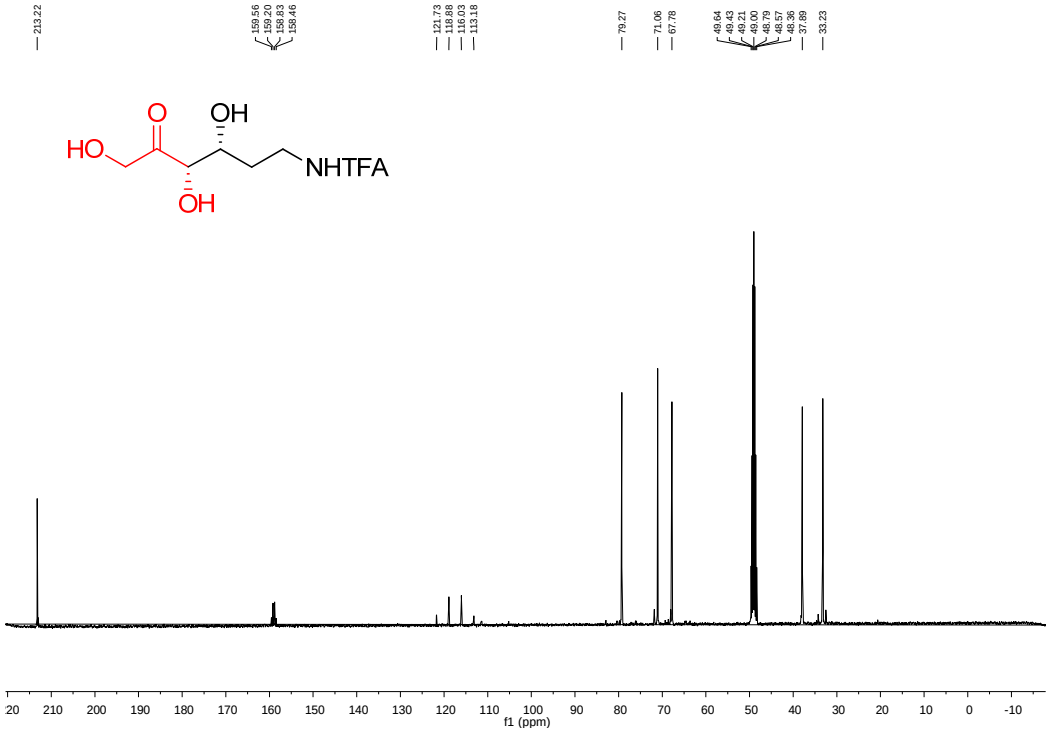
^{13}C-NMR spectrum of 3-difluoroacetamido propanal in $CDCl_3$.



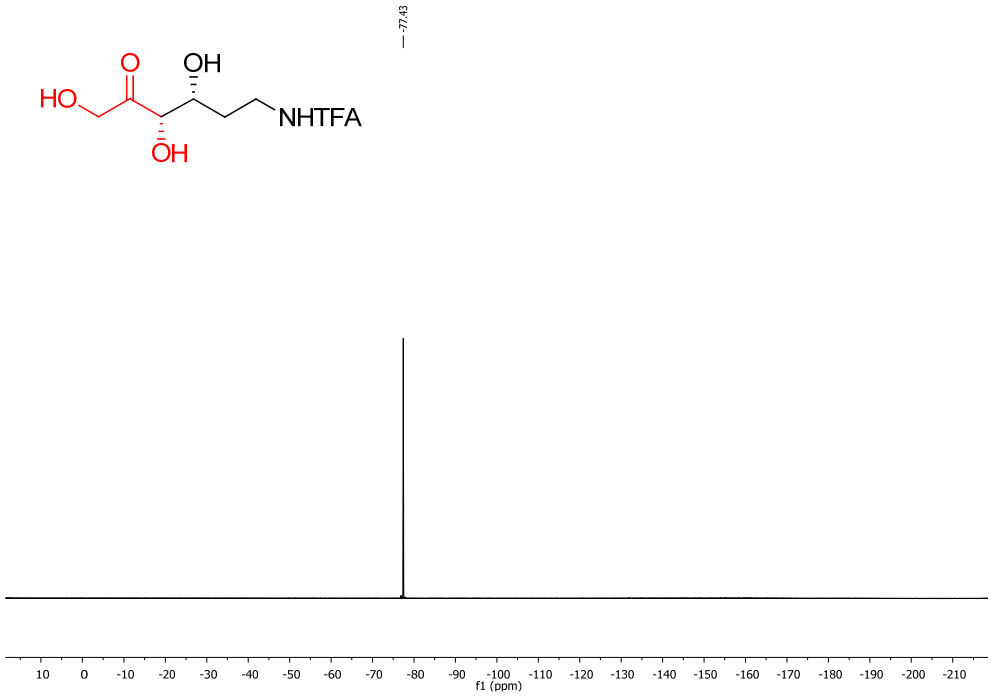
^{1}H-NMR spectrum of (3*S*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



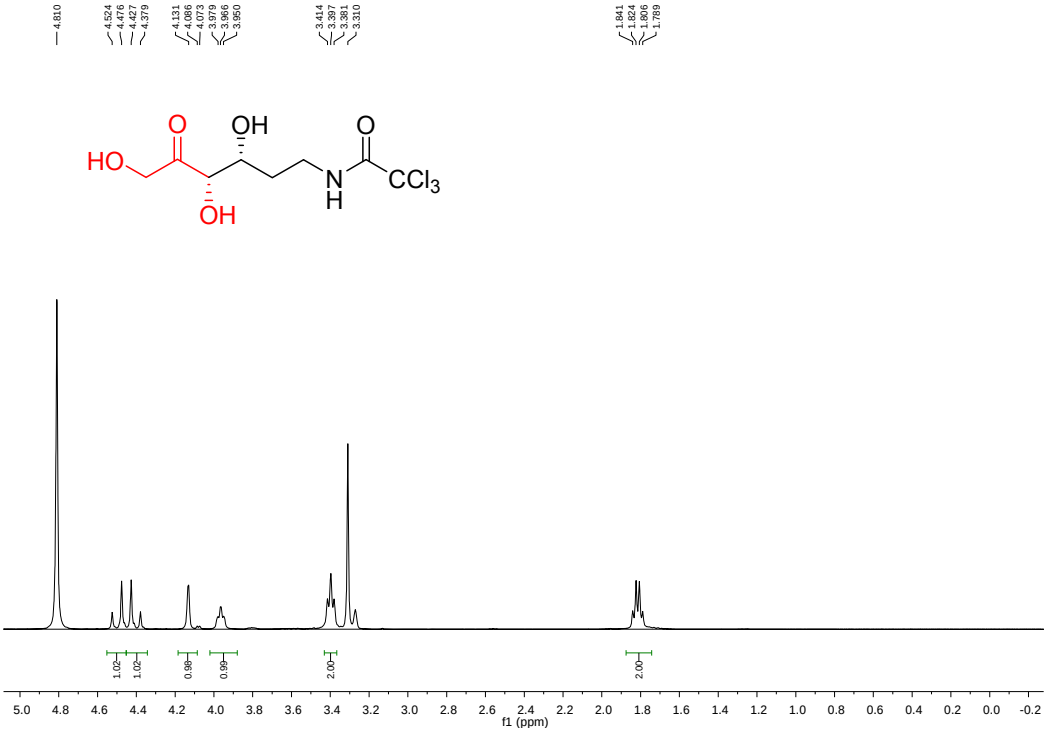
^{13}C-NMR spectrum of (3*S*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



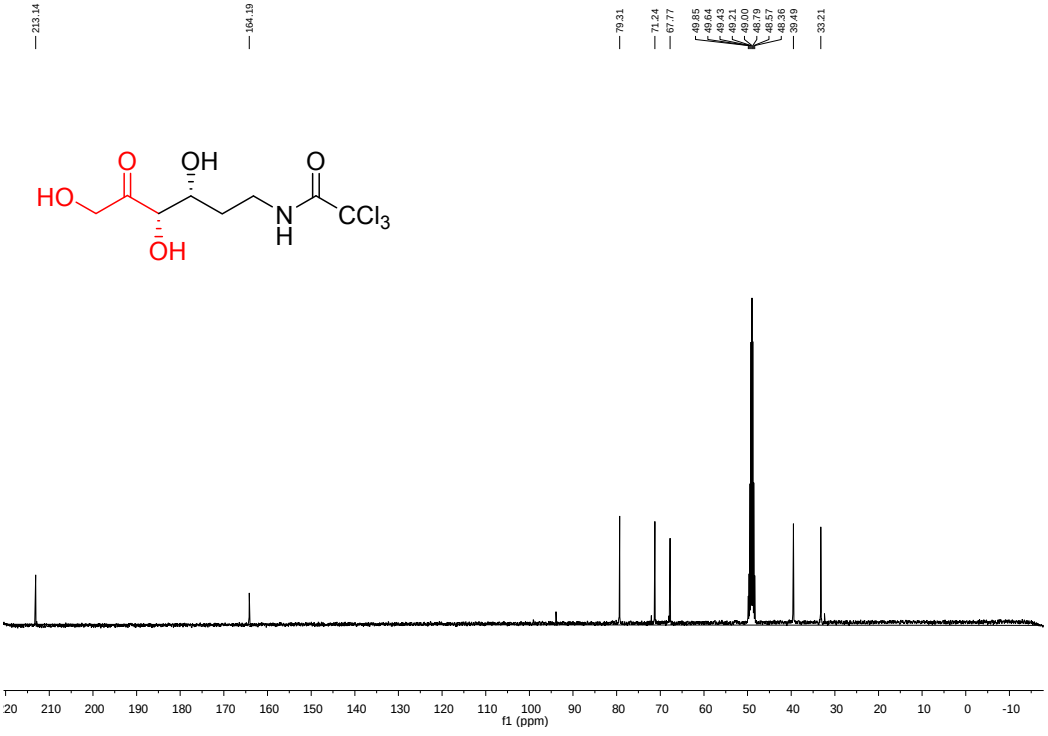
^{19}F-NMR spectrum of (3*S*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



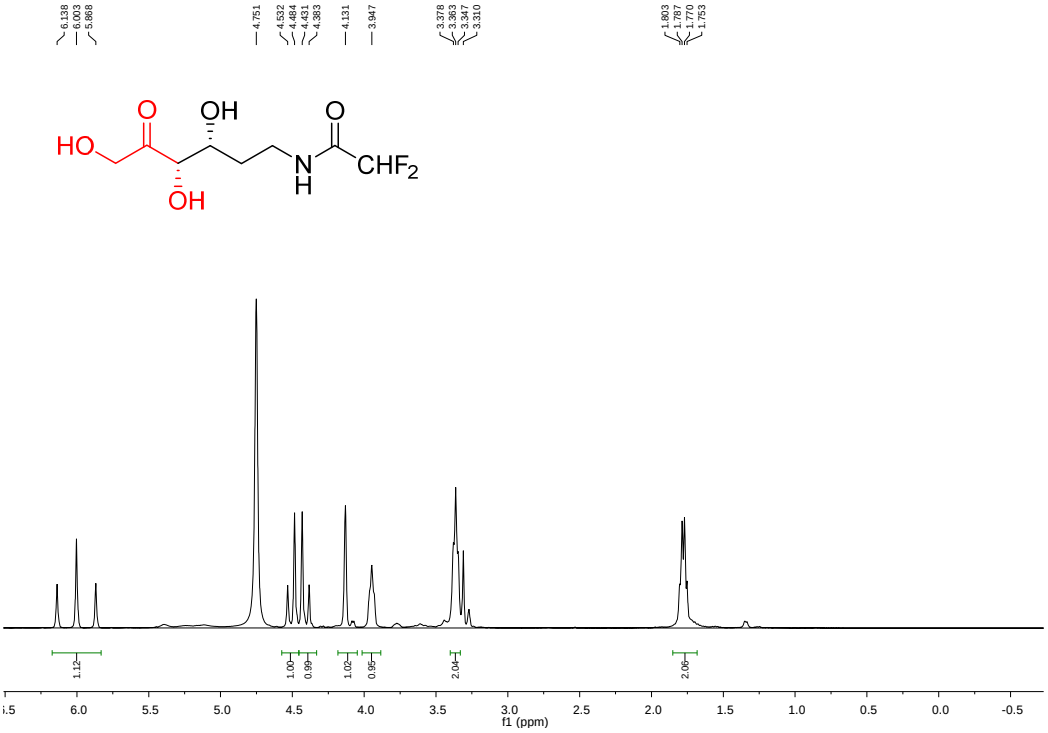
^{1}H-NMR spectrum of (3*S*, 4*R*)-6-trichloroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



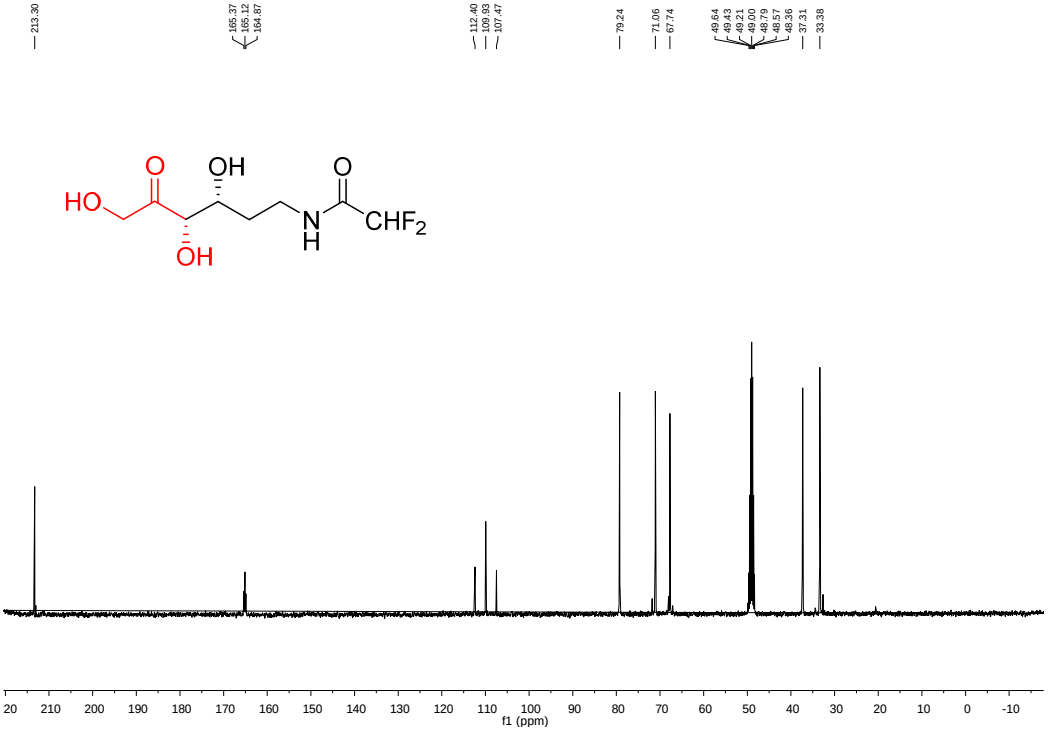
^{13}C-NMR spectrum of (3*S*, 4*R*)-6-trichloroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



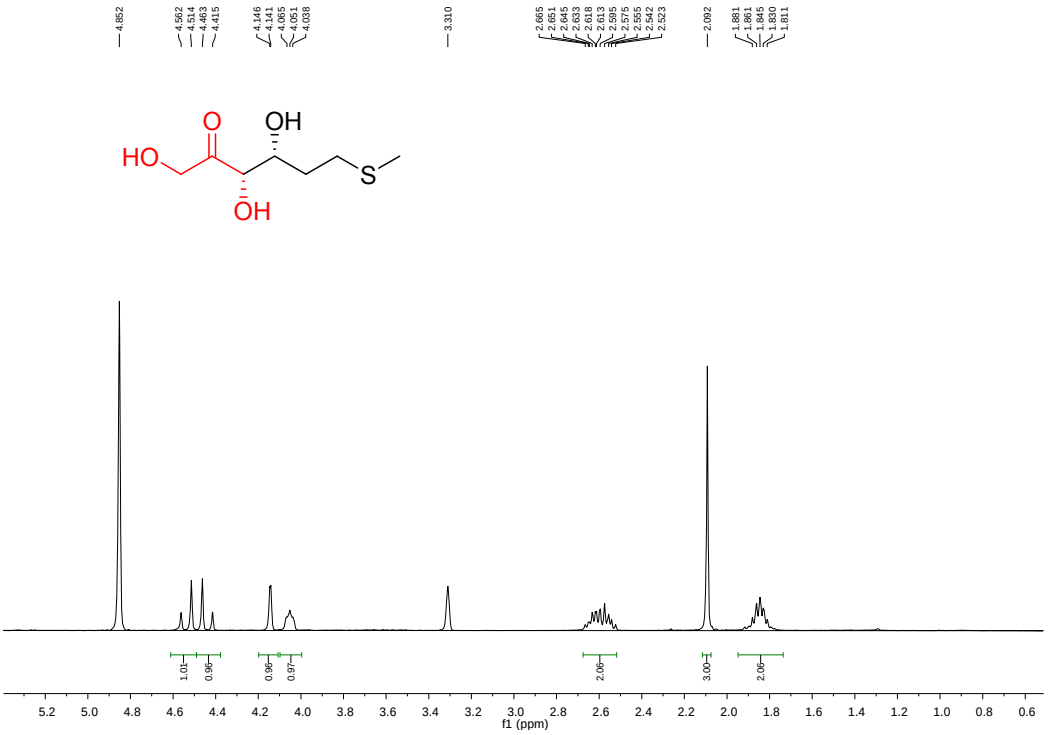
^{1}H-NMR spectrum of (3*S*, 4*R*)-6-difluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



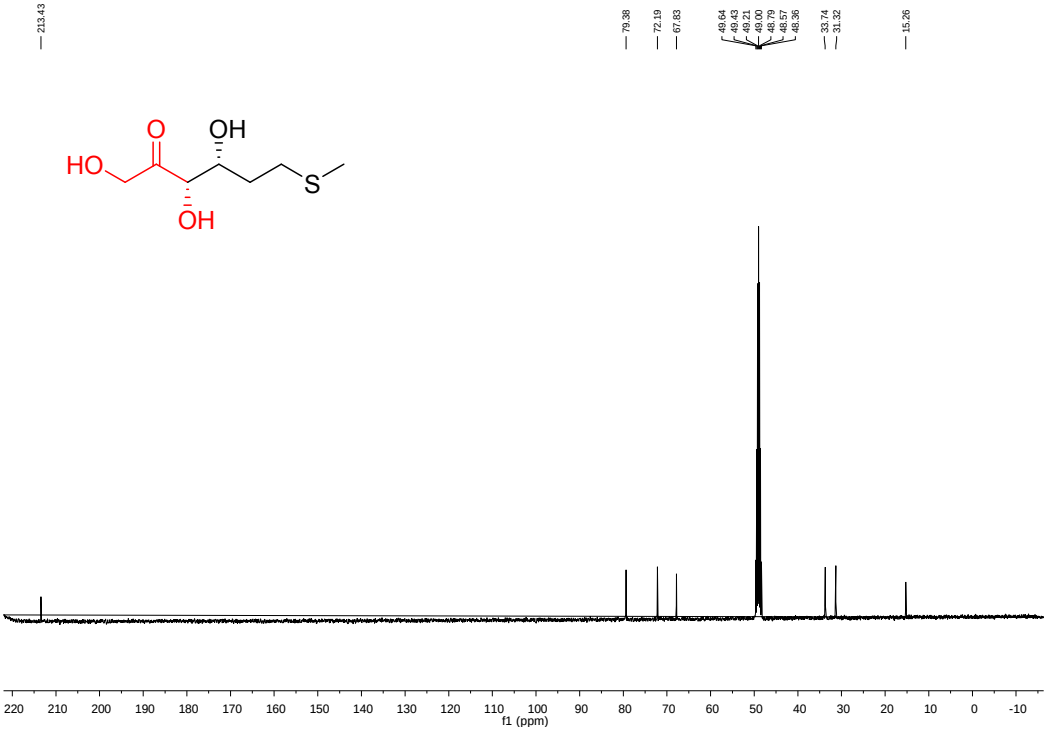
^{13}C-NMR spectrum of (3*S*, 4*R*)-6-difluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



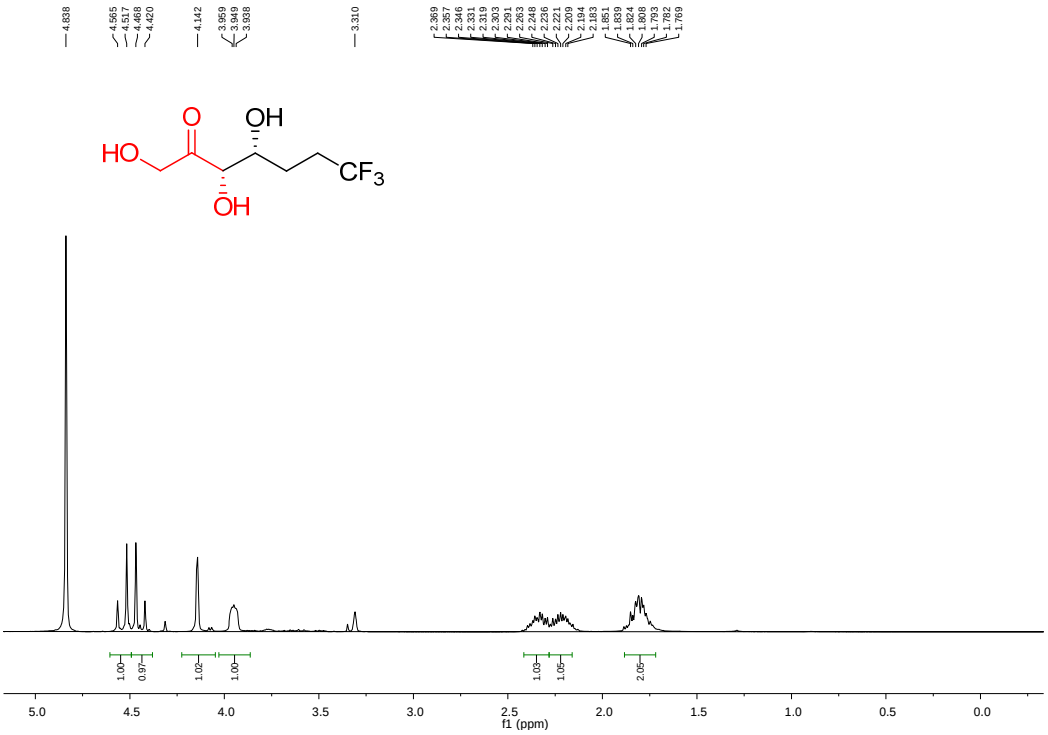
^{1}H-NMR spectrum of (3*S*, 4*R*)-1,3,4-trihydroxy-6-(methylthio)hexan-2-one in CD_3OD.



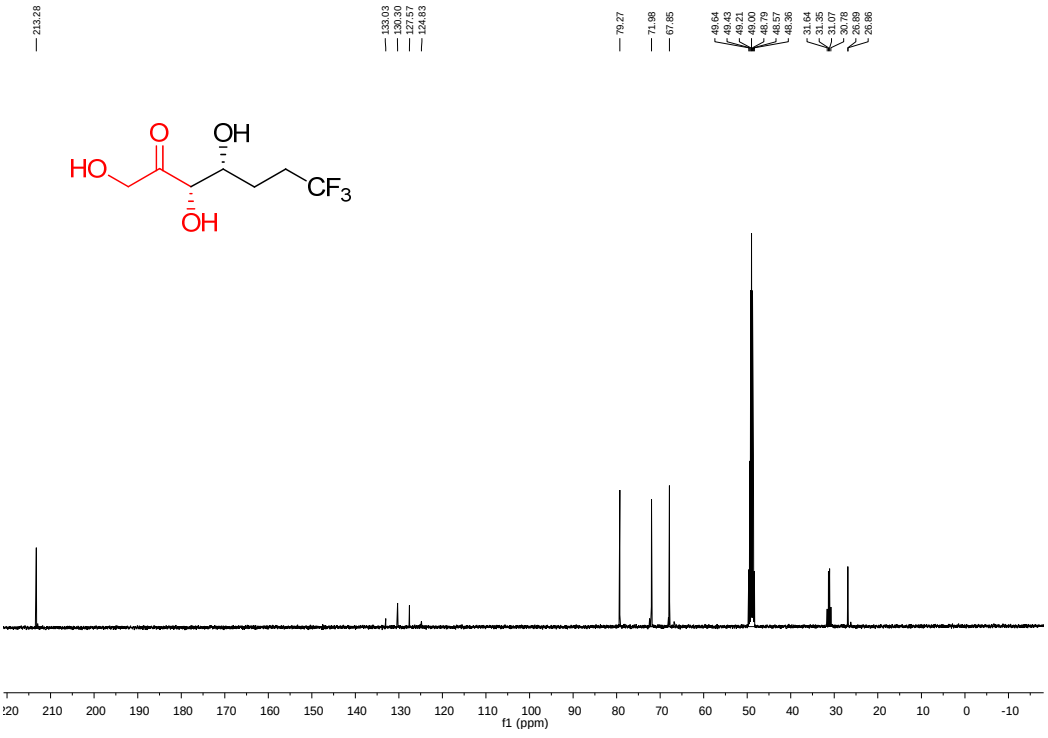
^{13}C-NMR spectrum of (3*S*, 4*R*)-1,3,4-trihydroxy-6-(methylthio)hexan-2-one in CD_3OD.



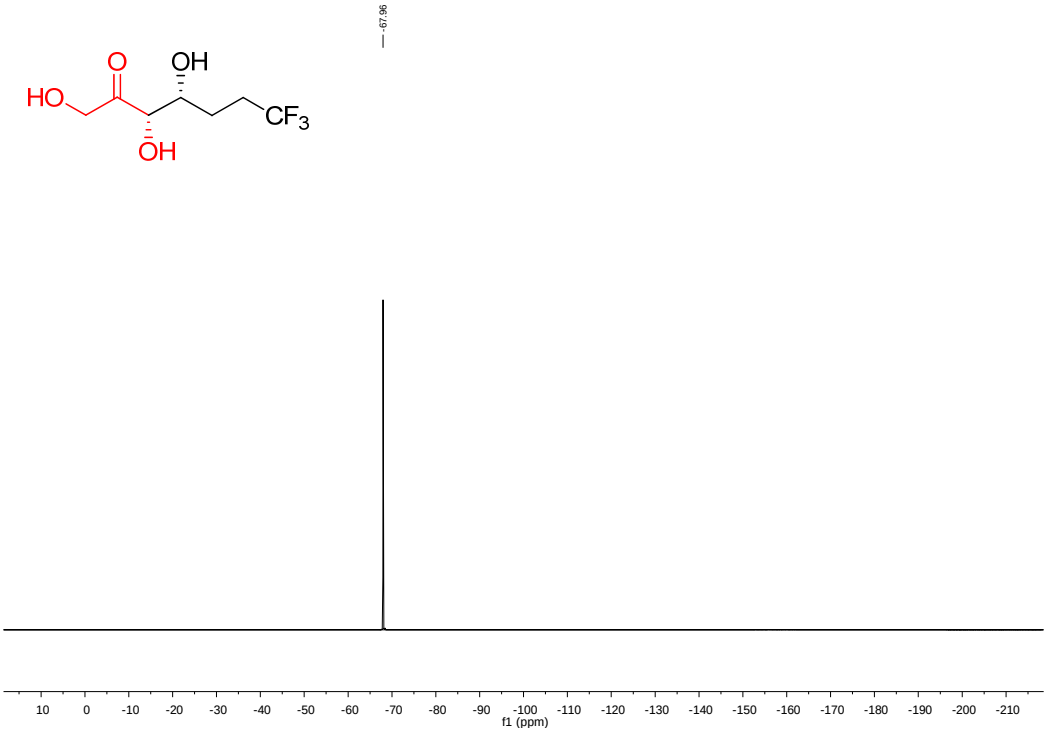
^{1}H-NMR spectrum of (3*S*, 4*R*)-7,7,7-trifluoro-1,3,4-trihydroxyheptan-2-one in CD_3OD



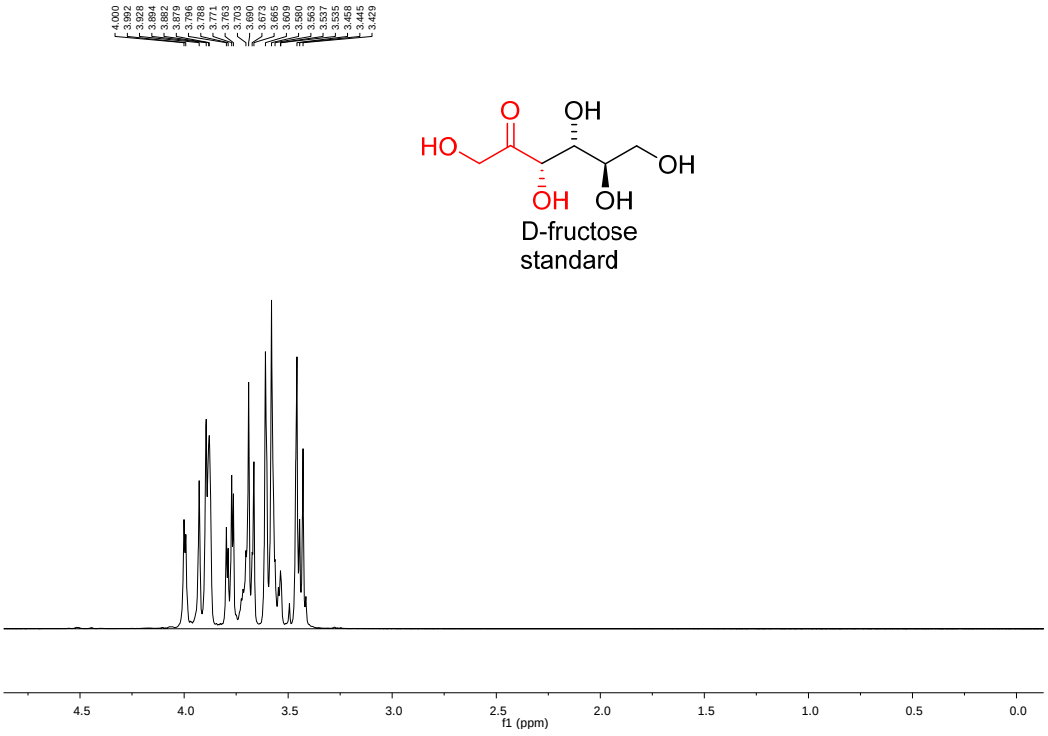
^{13}C-NMR spectrum of (3*S*, 4*R*)-7,7,7-trifluoro-1,3,4-trihydroxyheptan-2-one in CD_3OD



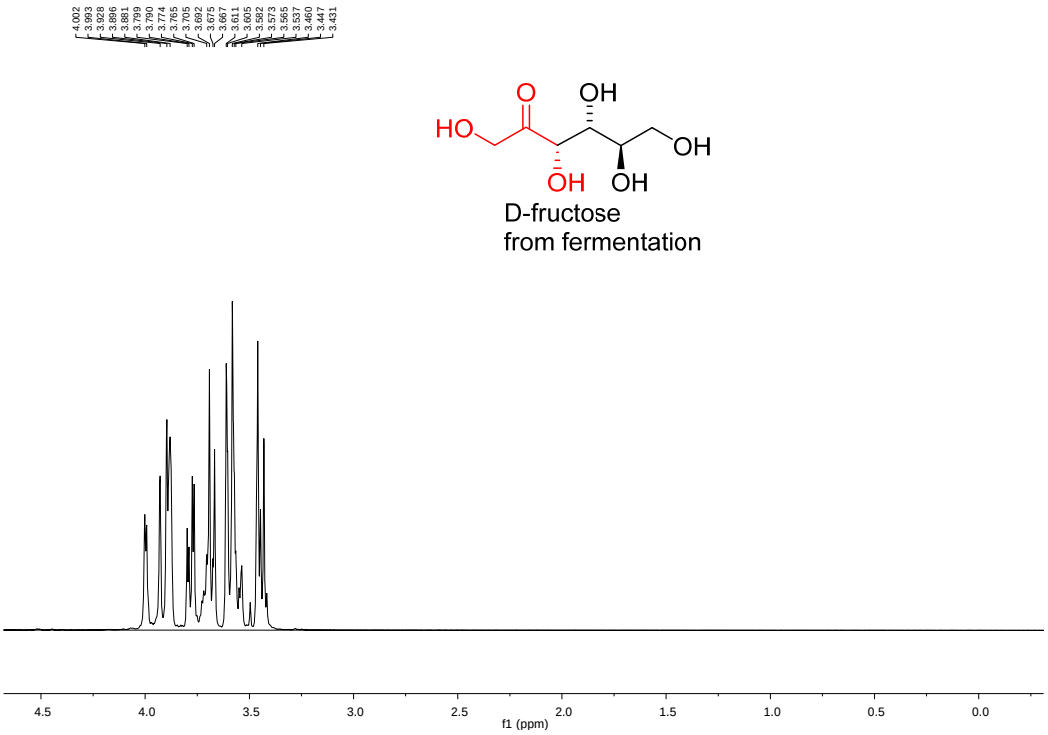
^{13}C-NMR spectrum of (3*S*, 4*R*)-7,7,7-trifluoro-1,3,4-trihydroxyheptan-2-one in CD_3OD.



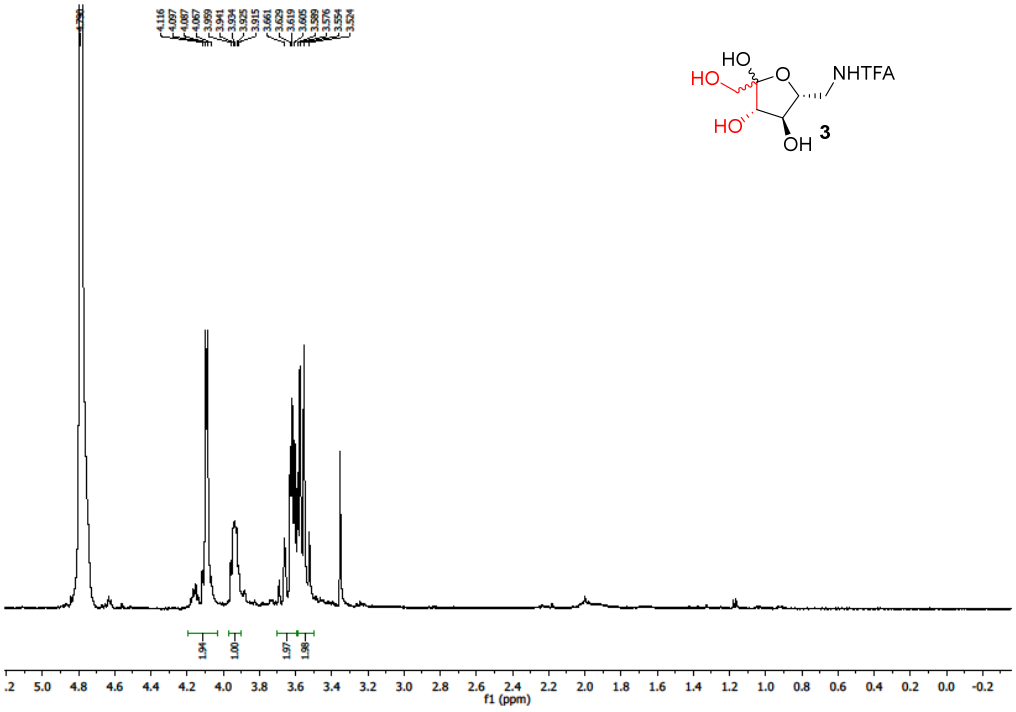
^{1}H-NMR spectrum of D-fructose standard in D_2O



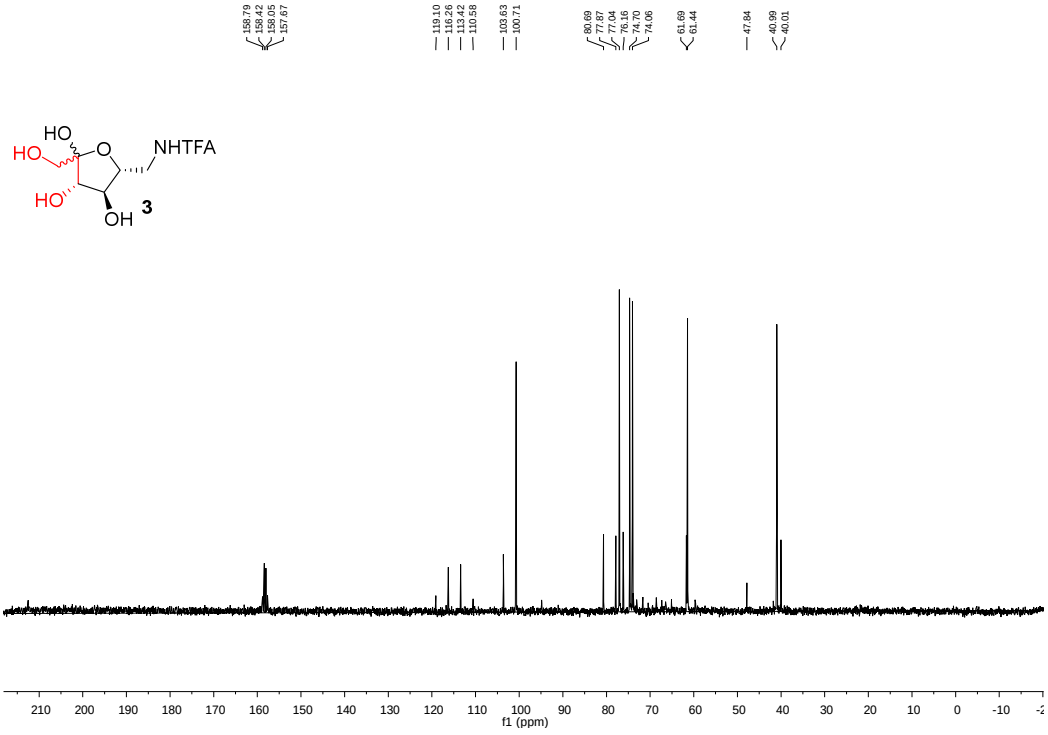
^{1}H-NMR spectrum of D-fructose from fermentation in D_2O



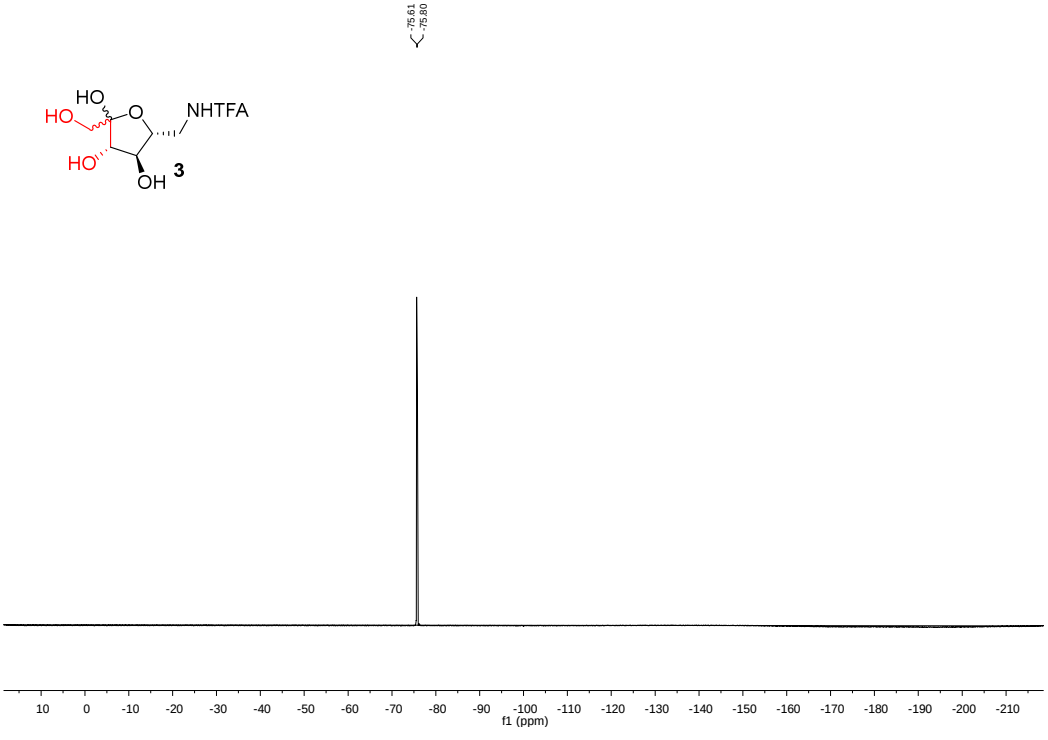
^{1}H-NMR spectrum of **3** in D_2O



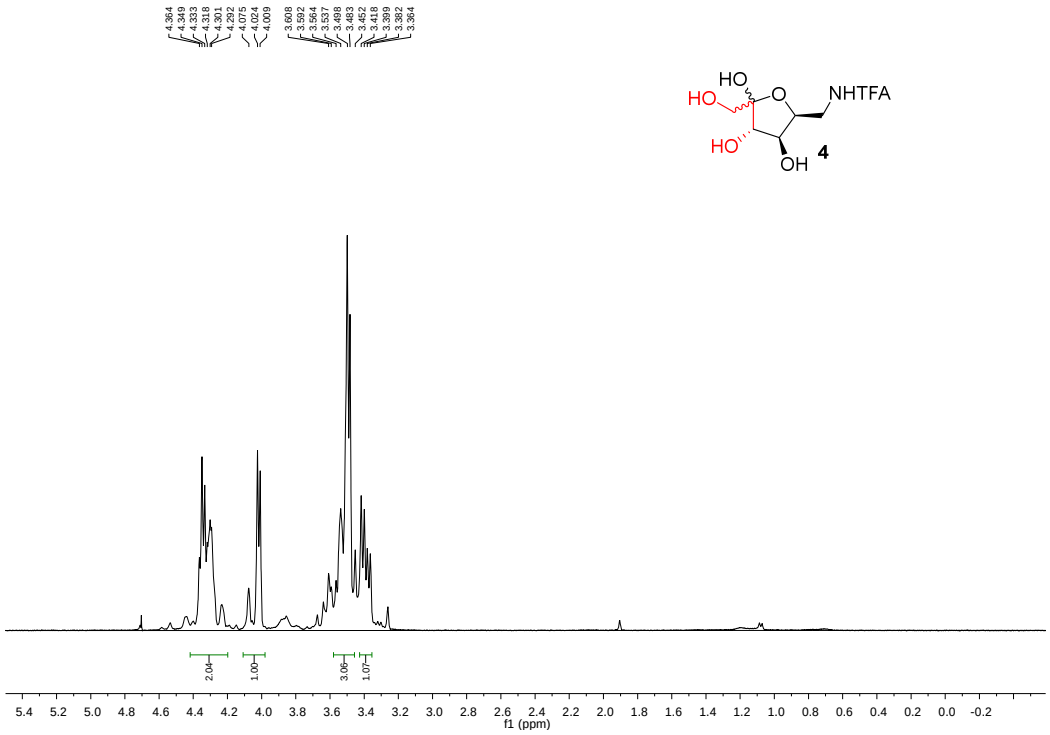
^{13}C-NMR spectrum of **3** in D_2O



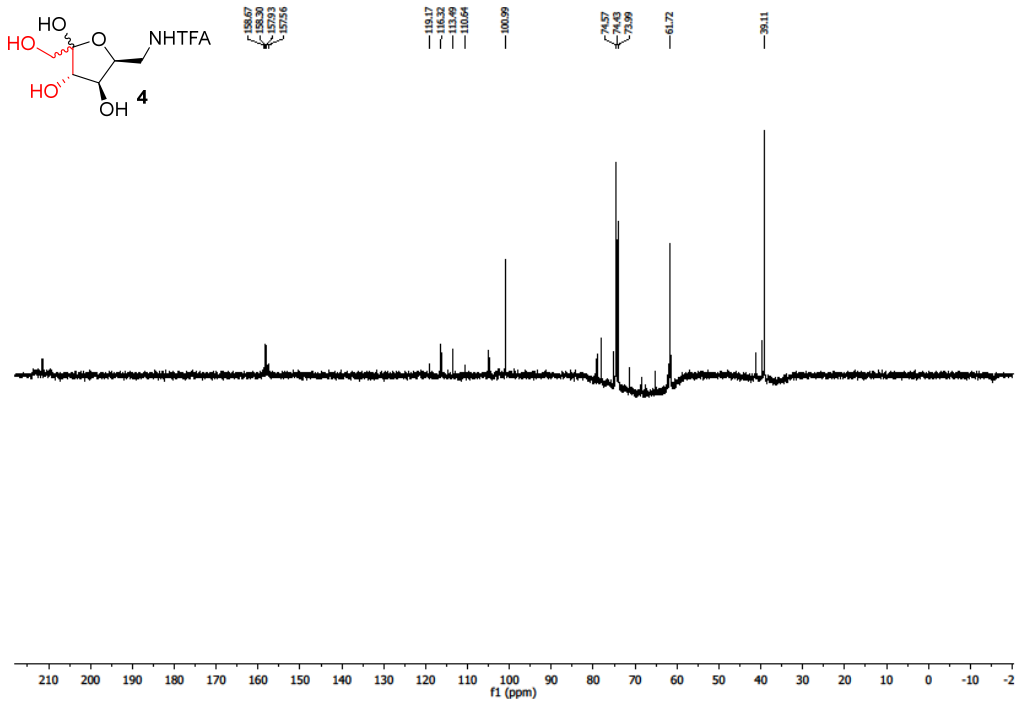
^{19}F-NMR spectrum of **3** in D_2O



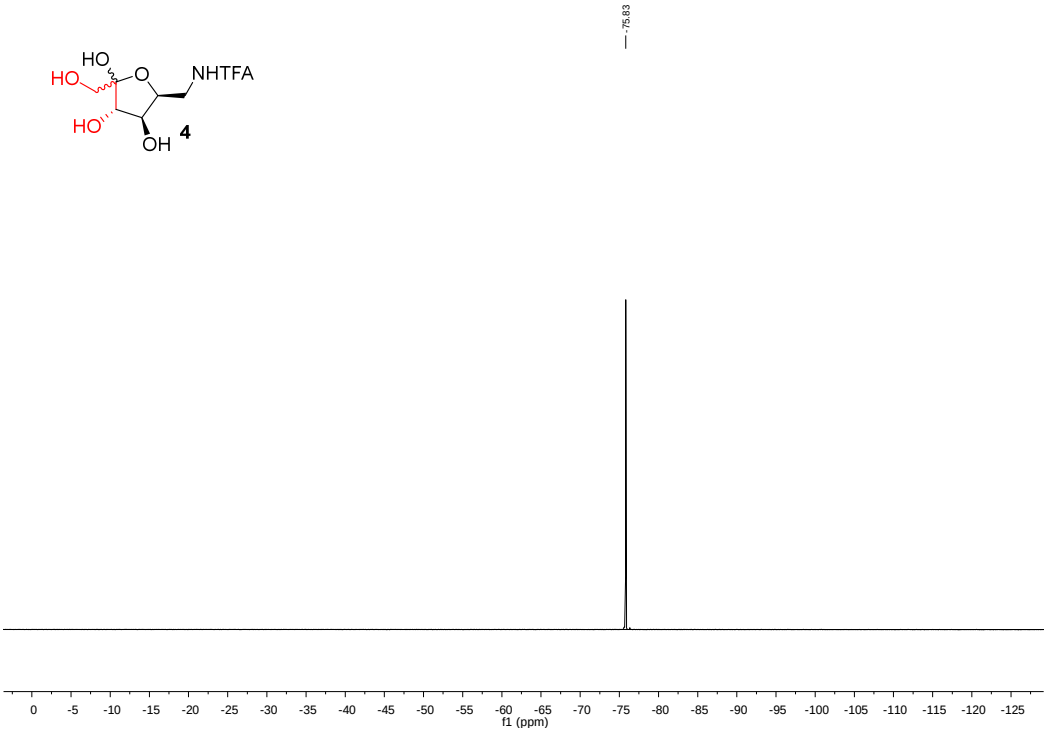
^{1}H-NMR spectrum of **4** in D_2O



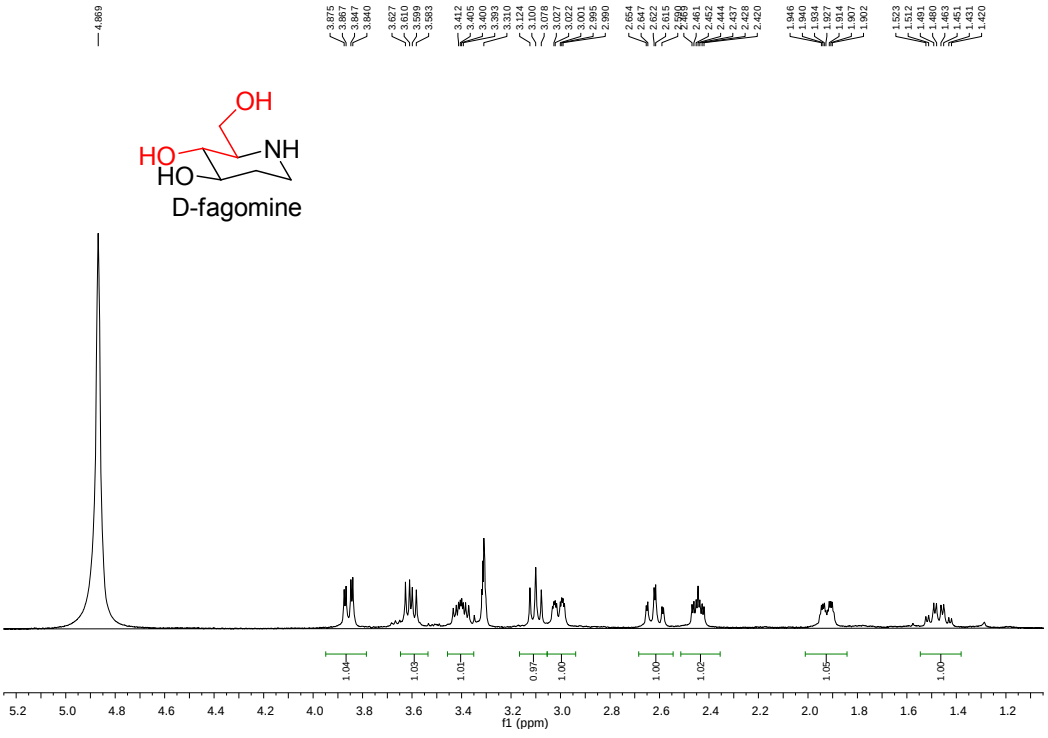
^{13}C-NMR spectrum of **4** in D_2O



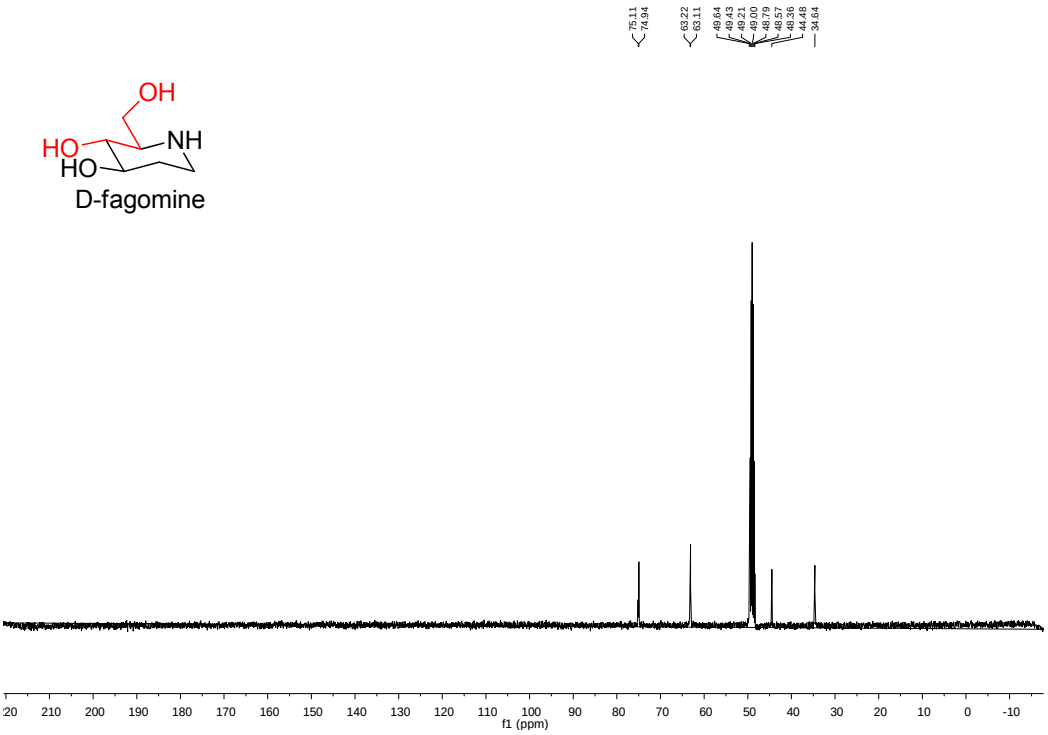
^{19}F-NMR spectrum of **4** in D_2O



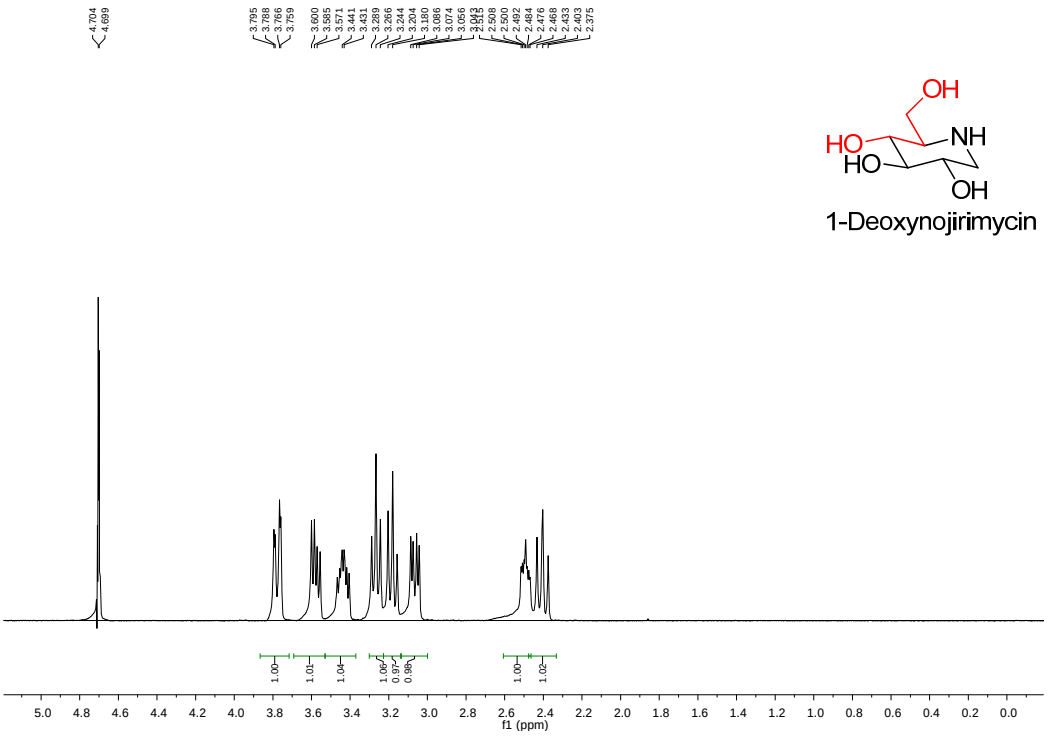
^{1}H-NMR spectrum of D-fagomine in CD_3OD.



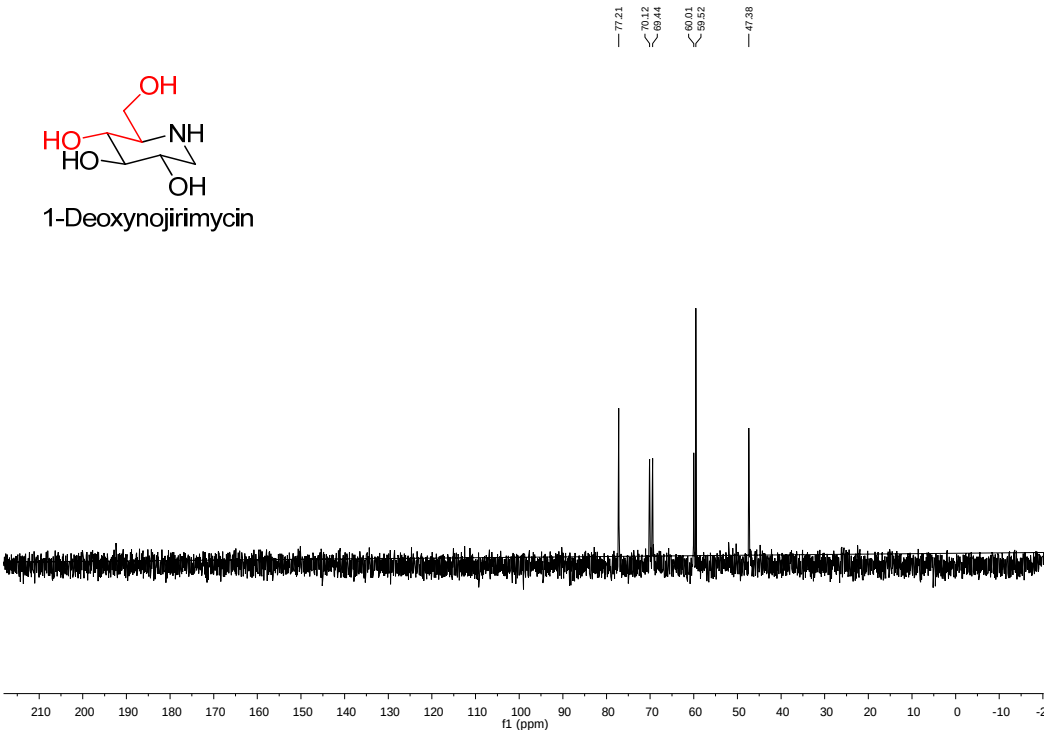
^{13}C-NMR spectrum of D-fagomine in CD_3OD.



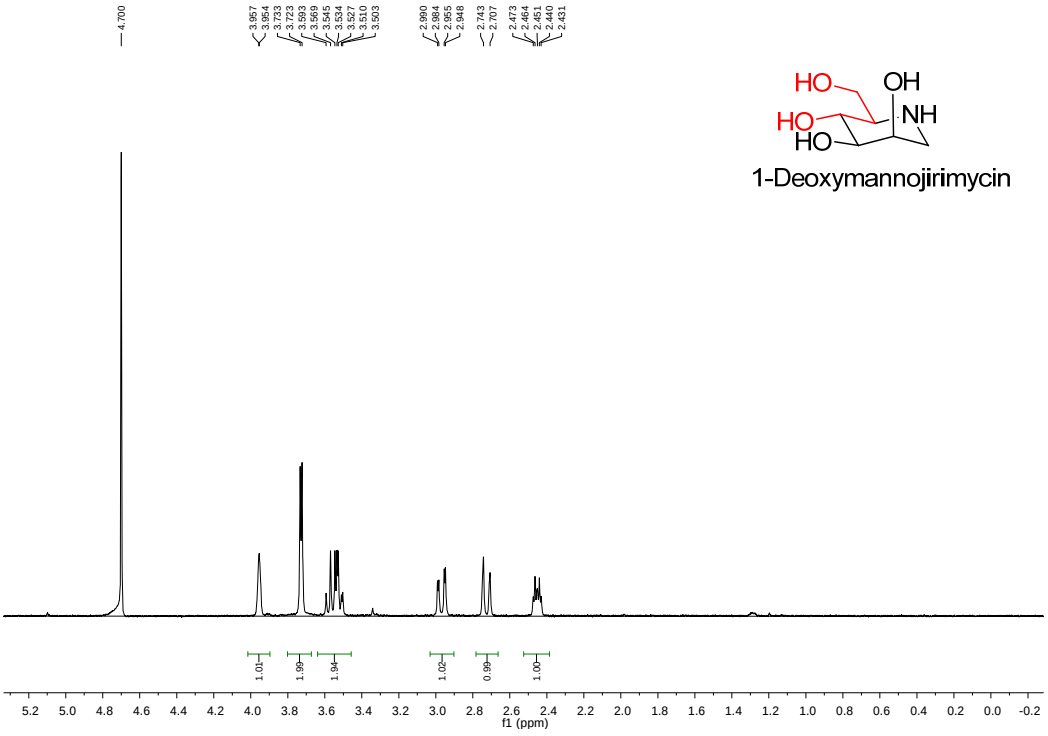
^{1}H-NMR spectrum of 1-Deoxynojirimycin in D_2O



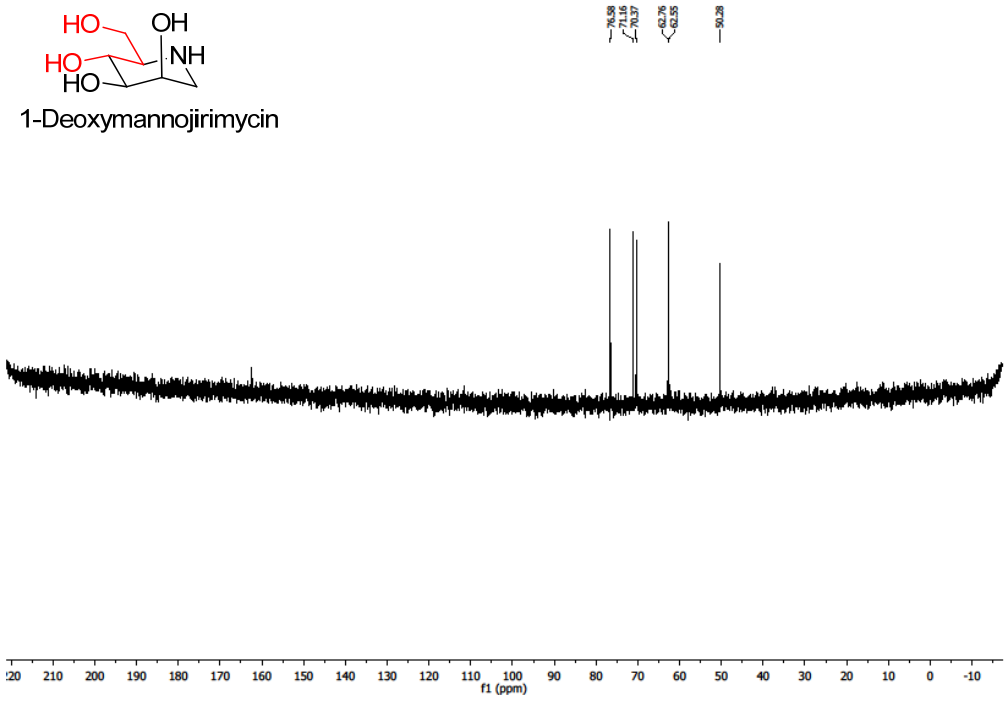
^{13}C-NMR spectrum of 1-Deoxynojirimycin in D_2O



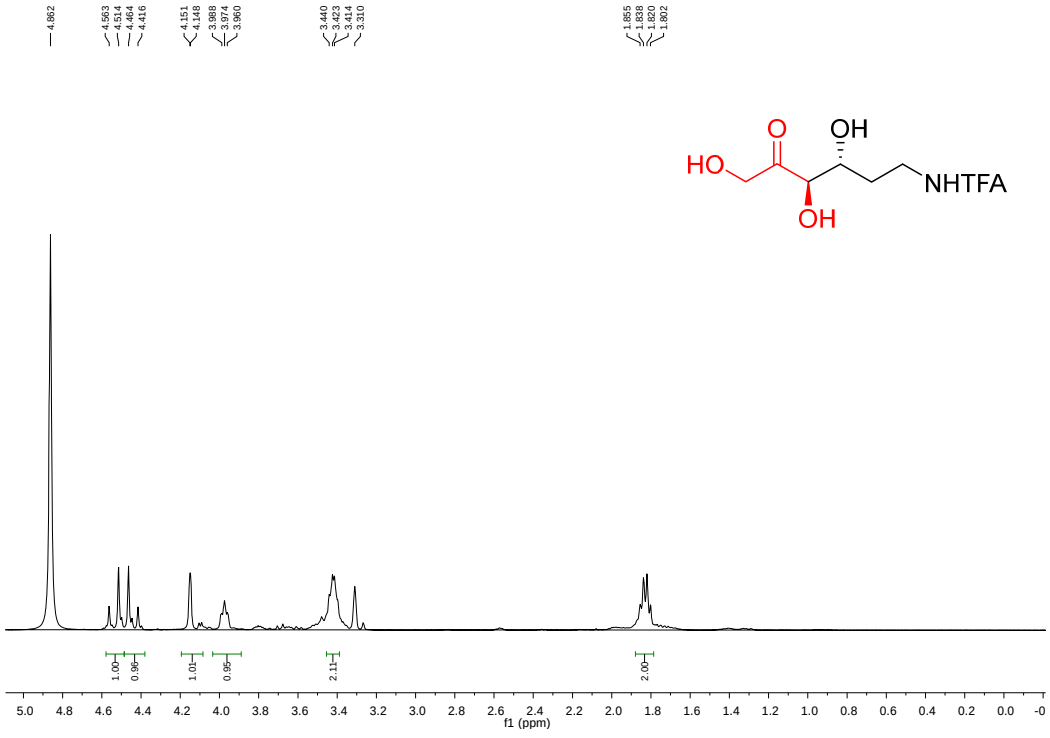
^{1}H-NMR spectrum of 1-Deoxymannojirimycin in D_2O



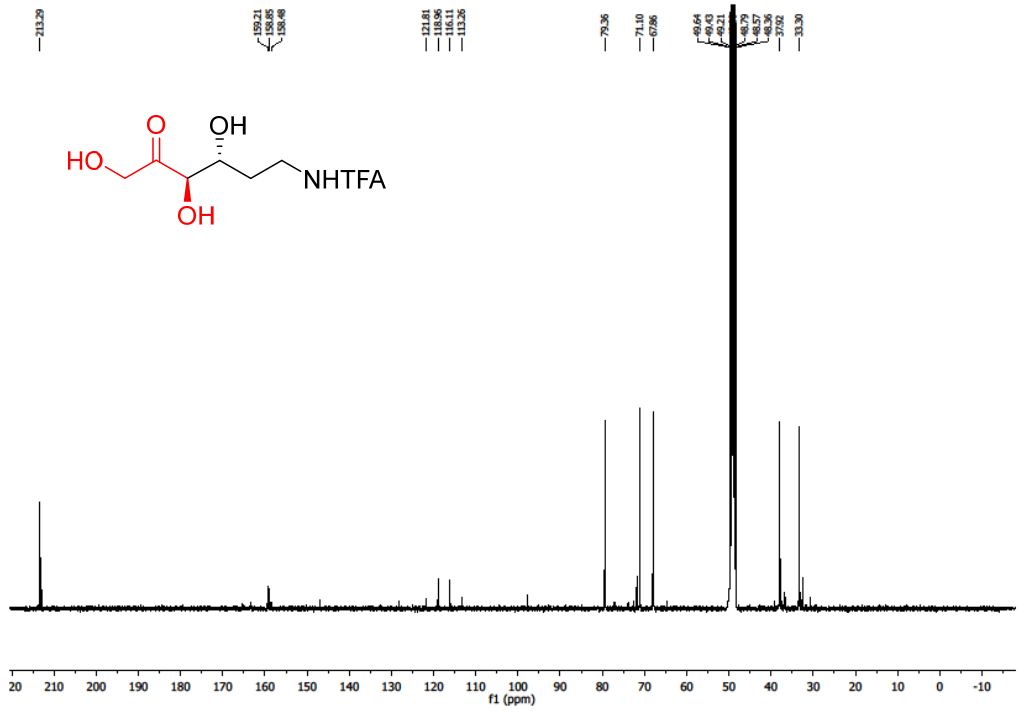
^{13}C-NMR spectrum of 1-Deoxymannojirimycin in D_2O



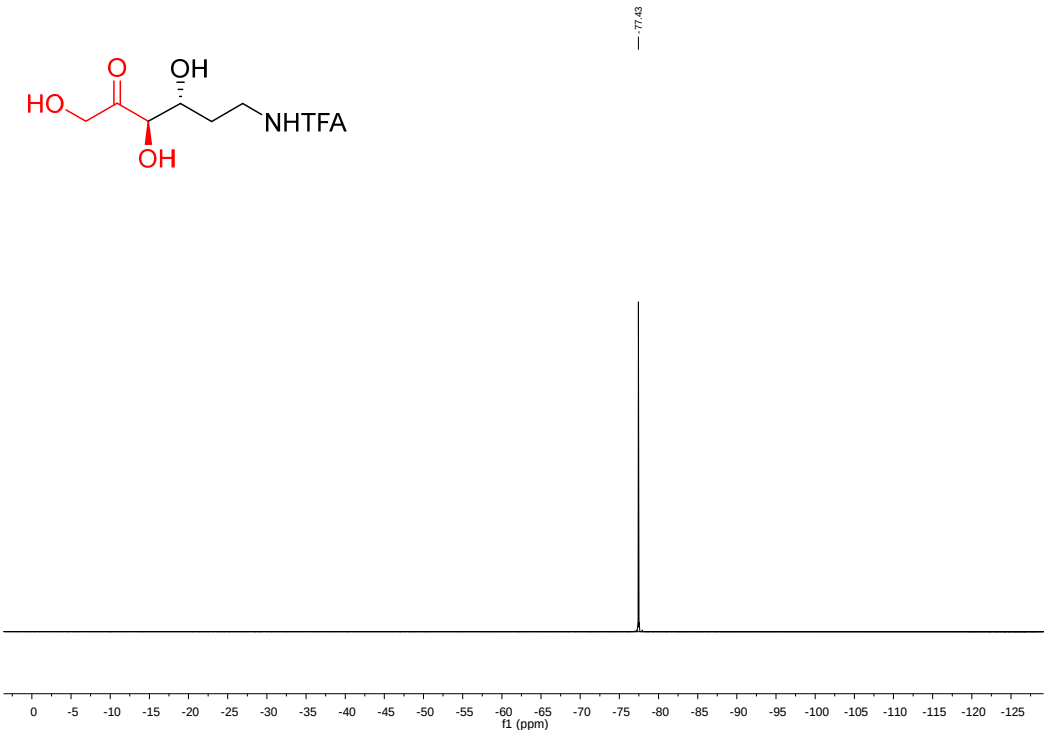
^{1}H-NMR spectrum of (3*R*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



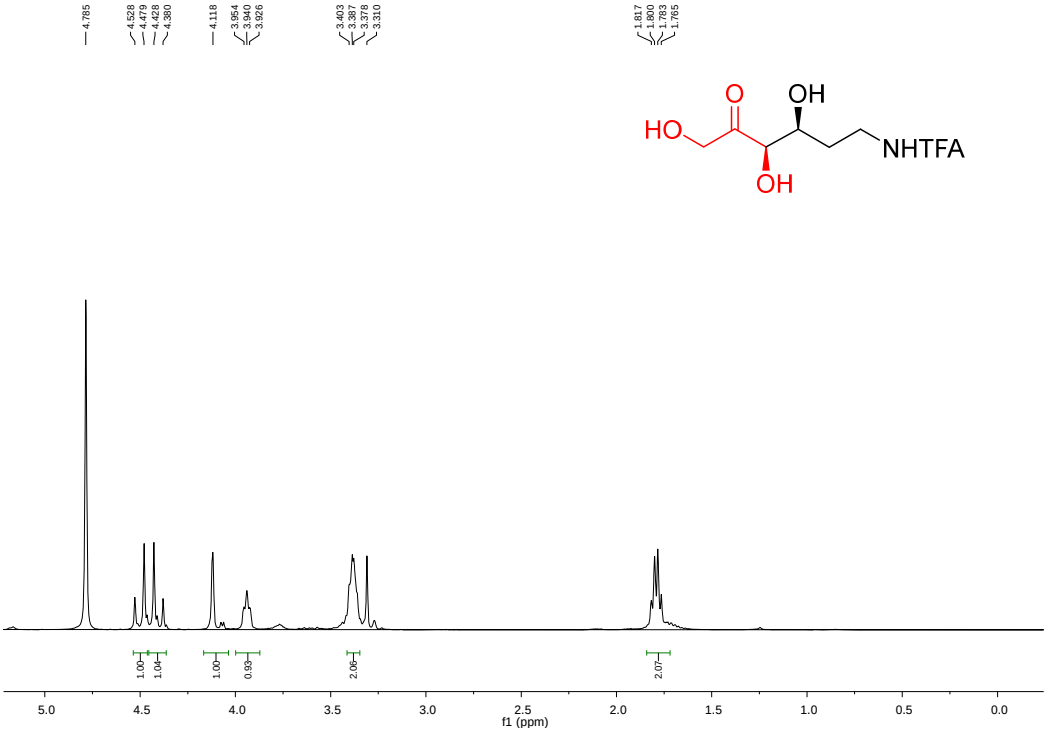
^{13}C-NMR spectrum of (3*R*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



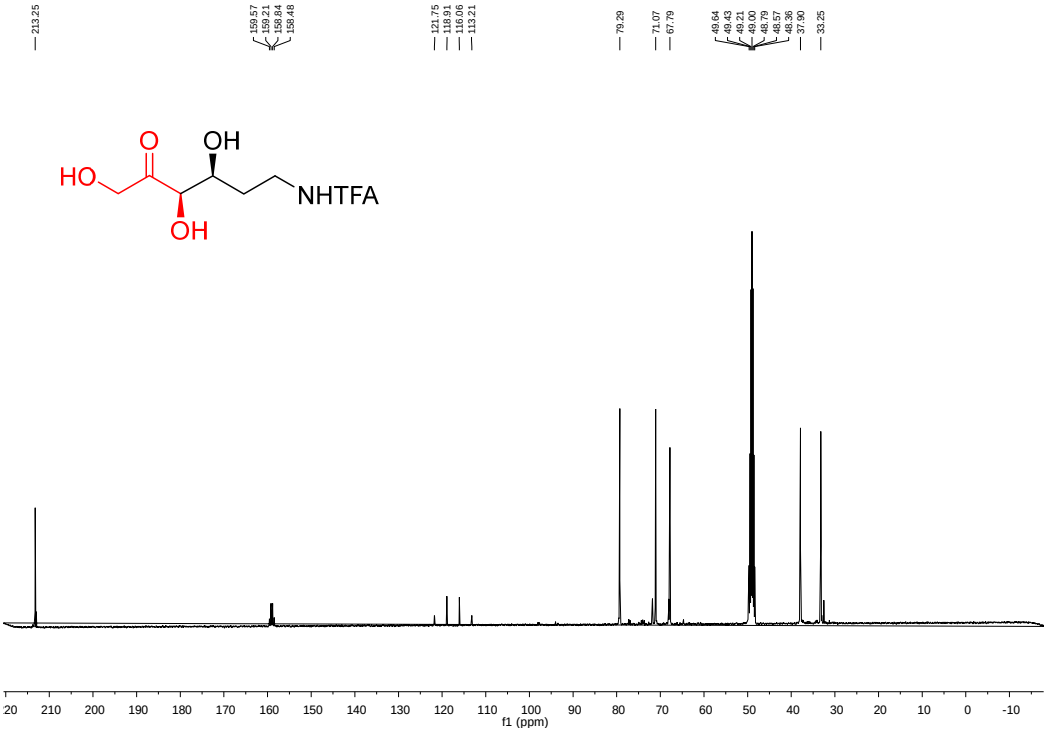
^{19}F-NMR spectrum of (3*R*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



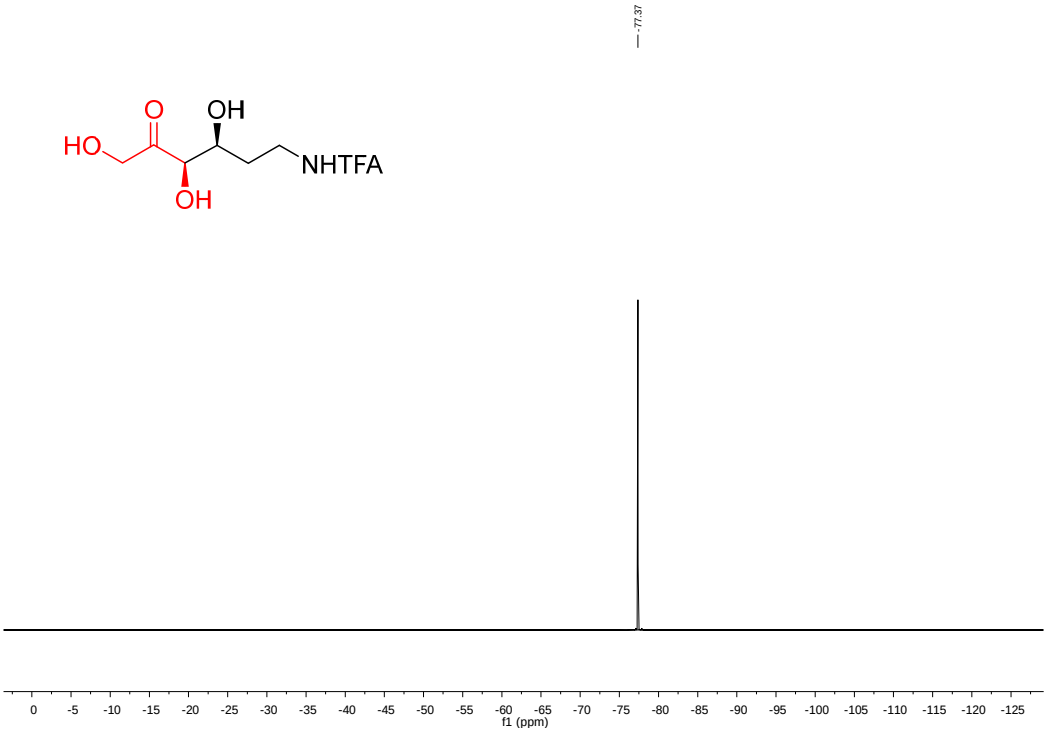
^{1}H-NMR spectrum of (3*R*, 4*S*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



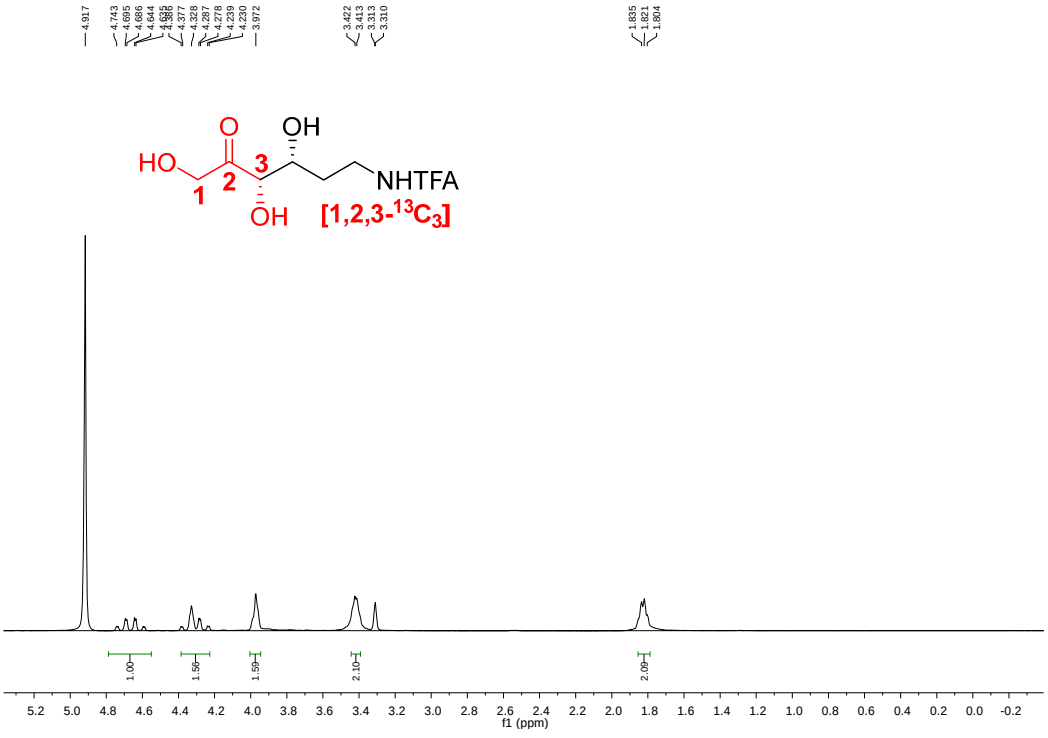
^{13}C-NMR spectrum of (3*R*, 4*S*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



^{19}F-NMR spectrum of (3*R*, 4*S*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



^{1}H-NMR spectrum of [1, 2, 3-$^{13}C_3$] (3*S*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.



^{13}C-NMR spectrum of [1, 2, 3-$^{13}C_3$] (3*S*, 4*R*)-6-trifluoroacetamido-1,3,4-trihydroxyhexan-2-one in CD_3OD.