Supplementary Information

Genetically Encoded Chemical Cross-linking of Carbohydrate

Shanshan Li[†], Nanxi Wang[†], Bingchen Yu[†], Wei Sun, Lei Wang*

University of California San Francisco, Department of Pharmaceutical Chemistry and the Cardiovascular Research Institute, 555 Mission Bay Blvd. South, San Francisco, California 94158, United States

[†] These authors contributed equally to this work.
* E-mail: Lei.Wang2@ucsf.edu

Table of Contents

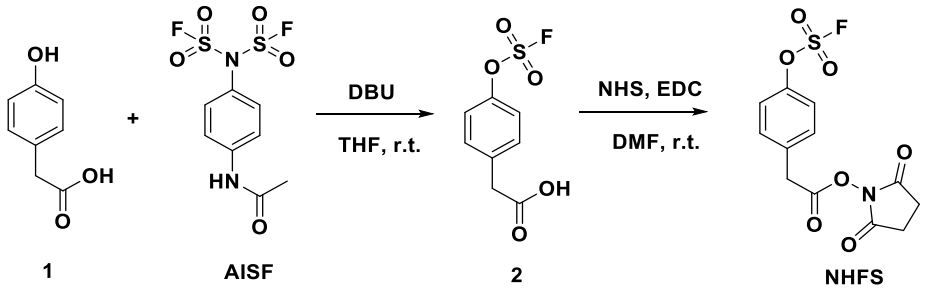
Chemo-enzymatic synthesis of azido-GD3

The scheme for chemo-enzymatic synthesis of azido-GD3 was shown in Supplementary Fig. 4. Azido-lac was chemically synthesized as described.[1] α-2,3 sialic acid transferase was used for the addition of sialic acid to get the azido-3'-Sialyllactose. Azido-GD3 was synthesized via enzymatic catalysis with azido-3'-Sialyllactose, *N*-acetylmannosamine, and pyruvate in the presence of aldolase, CMP-sialic acid synthetase and α-2,8 sialic acid transferase. The final product azido-GD3 was purified using HPLC and characterized with ESI-MS. [M+H], [M+Na] peaks of azido-GD3 were observed (Supplementary Fig. 5).

Chemical synthesis procedure

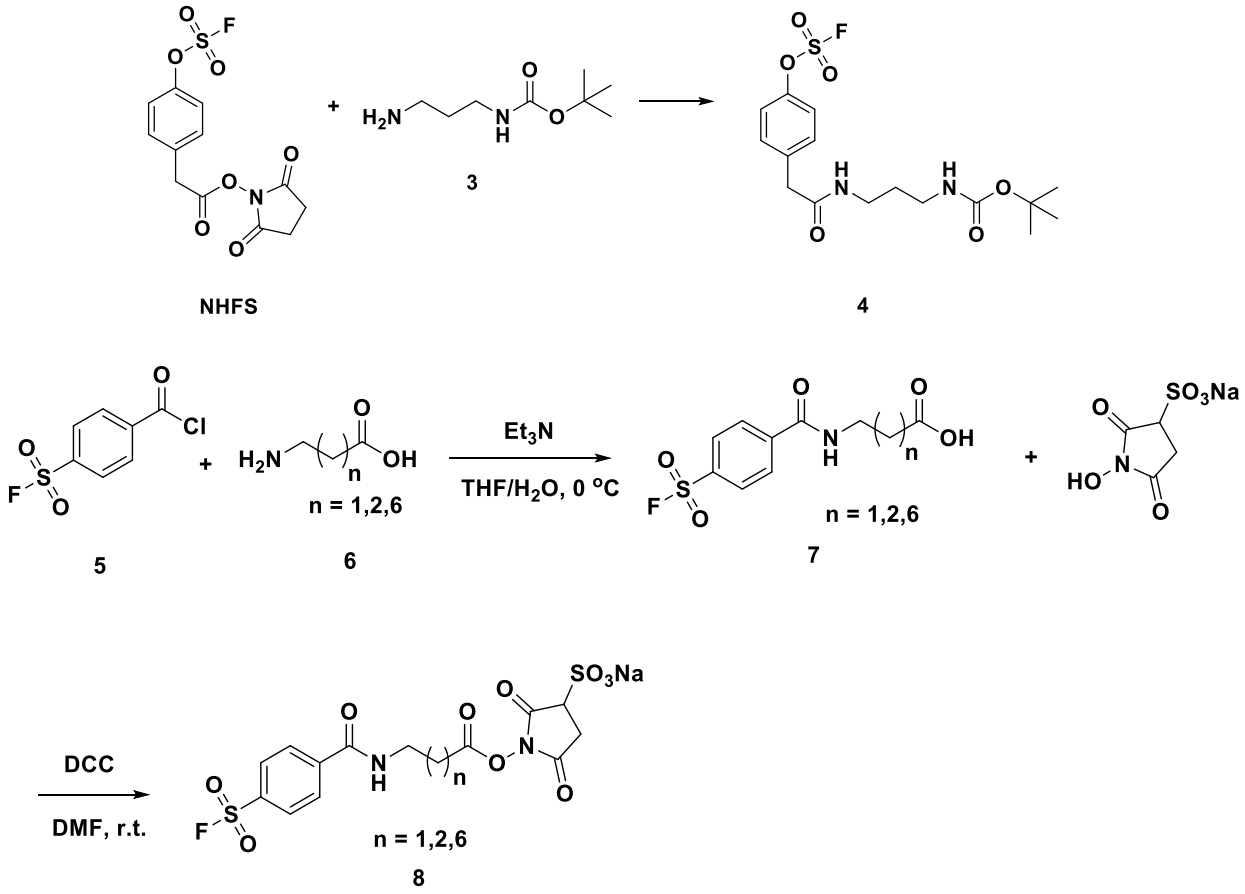
General information for chemical synthesis

All solvents were of reagent grade and were purchased from Fisher Scientific and Aldrich. Reagents were purchased from Aldrich and VWR. ^{1}H-NMR (400 MHz) and ^{13}C-NMR (100 MHz) spectra were recorded on a Bruker 400 MHz NMR spectrometer.



Synthesis of 2-(4-((fluorosulfonyl)oxy)phenyl)acetic acid (**2**). 2-(4-hydroxyphenyl)acetic acid (**1**) was converted to fluorosulfate using [4-(acetylamino)phenyl]imidodisulfuryl difluoride (AISF). [2] 1.5 g compound **1** (1.5 g, 9.9 mmol) and AISF (3.4 g, 10.8 mmol) was dissolved in 50 mL anhydrous THF. Then 1,8-Diazabicyclo[5.4.0]undec-7-ene (DBU, 3.2 g, 22 mmol) was added dropwise at room temperature (r.t.). The mixture was stirred at r.t. for 10 min. Then 200 mL EtOAc was added to dilute the reaction mixture and the organic phrase was washed sequentially by H_2O (100 mL) and brine (100 mL). The organic phase was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure to give the crude product, which was then purified by column chromatography (silica gel, DCM: MeOH=50:1) to give a white solid (1.2g, 53 %).

Synthesis of 2,5-dioxopyrrolidin-1-yl 2-(4-((fluorosulfonyl)oxy)phenyl)acetate (**NHFS**). To a stirred solution of compound **2** (500 mg, 2.1 mmol) and *N*-hydroxysuccinimide (NHS, 358 mg, 3.1 mmol) in 4 mL anhydrous DMF was added *N*-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (EDC· HCl, 611 mg, 3.2 mmol). The mixture was stirred at r.t. for 24 h. Then the reaction was quenched with the addition of H_2O (30 mL) and the mixture was extracted with EtOAc (2 × 30 mL). The combined organic phase was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure to give the crude product, which was purified by column chromatography (silica gel, DCM: EtOAc = 50:1) to give a white solid (452 mg, 65 %). ^{1}H NMR ($CDCl_3$): δ 7.46 (d, *J* = 8.8 Hz, 2H), 7.34 (d, *J* = 8.8 Hz, 2H), 3.98 (s, 2H), 2.84 (s, 4H). ^{13}C NMR ($CDCl_3$): δ 169.0, 166.2, 149.6, 132.4, 131.5, 121.5, 37.0, 25.7. **NHFS** itself have poor signal during mass spectrum analysis. **NHFS** was converted to compound **4** for mass analysis. Briefly, 20 mM compound **NHFS**, 20 mM *tert*-butyl (3-aminopropyl)carbamate (**3**) and 20 mM NaOH was incubated in H_2O at r.t. for 2 h. Then the solution was subject to mass spectrum analysis. HRMS calcd for $C_{16}H_{23}FN_2Na_2O_6S$ $[M+Na]^+$ 413.1153, found: 413.1158.

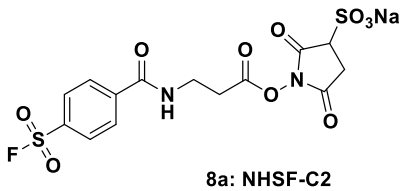


General synthetic procedure for compound 8.

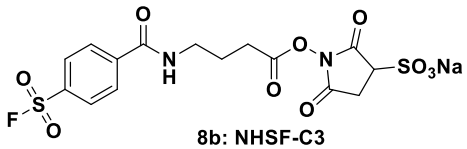
Compound **5** was synthesized from 4-(fluorosulfonyl)benzoic acid using $SOCl_2$ according to literature procedure. [3]

Synthesis of compound **7**. To a stirred solution of compound **6** (2.5 mmol) and triethylamine (Et_3N, 5.0 mmol) in H_2O (2 mL) was added dropwise compound **5** (2.5 mmol) in THF (4 mL) at 0 °C. The mixture was allowed to warm to room temperature and was stirred for 2 h. Then the reaction was quenched with the addition of H_2O (25 mL) and the mixture was extracted with EtOAc (2 × 25 mL). The combined organic phase was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure to give the crude product, which was purified by column chromatography (silica gel, DCM: MeOH = 20:1) to give a white solid (~ 70 %).

Synthesis of compound **8**. Compound **7** (0.72 mmol), *N*-Hydroxysulfosuccinimide sodium salt (0.72 mmol) and *N*,*N*'-Dicyclohexylcarbodiimide (DCC, 0.72 mmol) was dissolved 1.5 mL anhydrous DMF. The mixture was stirred at r.t. for 24 h under N_2. A white precipitate was formed during the reaction and was removed by filtration. 20 mL diethyl ether was added to the filtrate, and a white precipitate was formed and collected by centrifuge (10 min, 3,000 rpm). The white precipitate was redissolved in 4 mL MeOH and 20 mL diethyl ether was added, and a white precipitate was formed and collected by centrifuge (10 min, 3,000 rpm). The white precipitate was further purified by preparation HPLC (C18 column) using H_2O/ACN (0.05 % TFA) as mobile phrase (~ 65 %).

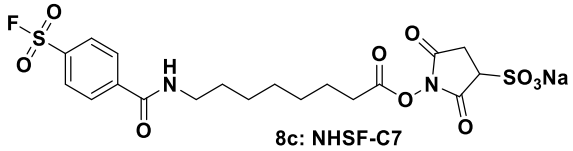


Sodium 1-((3-(4-(fluorosulfonyl)benzamido)propanoyl)oxy)-2,5-dioxopyrrolidine-3-sulfonate (**8a, NHSF-C2**). ^{1}H NMR (DMSO-d_6): δ 9.08 (t, *J* = 5.6 Hz, 1H), 8.26 (d, *J* = 8.4 Hz, 2H), 8.15 (d, *J* = 8.4 Hz, 2H), 3.94 (d, *J* = 8.0 Hz, 1H), 3.63-3.58 (m, 2H), 3.14-3.07 (m, 1H), 3.02 (t, *J* = 6.4 Hz, 2H), 2.85 (dd, *J* = 16.0 Hz, *J* = 2.4 Hz, 1H). ^{13}C NMR (DMSO-d_6): δ 168.7, 165.2, 164.7, 141.2, 133.6 (d, *J* = 23 Hz, C-F), 129.1, 128.7, 56.3, 35.2, 31.0, 30.2. HRMS calcd for $C_{14}H_{12}FN_2Na_2O_{10}S_2$ $[M+Na]^+$ 496.9707, found: 496.9716.



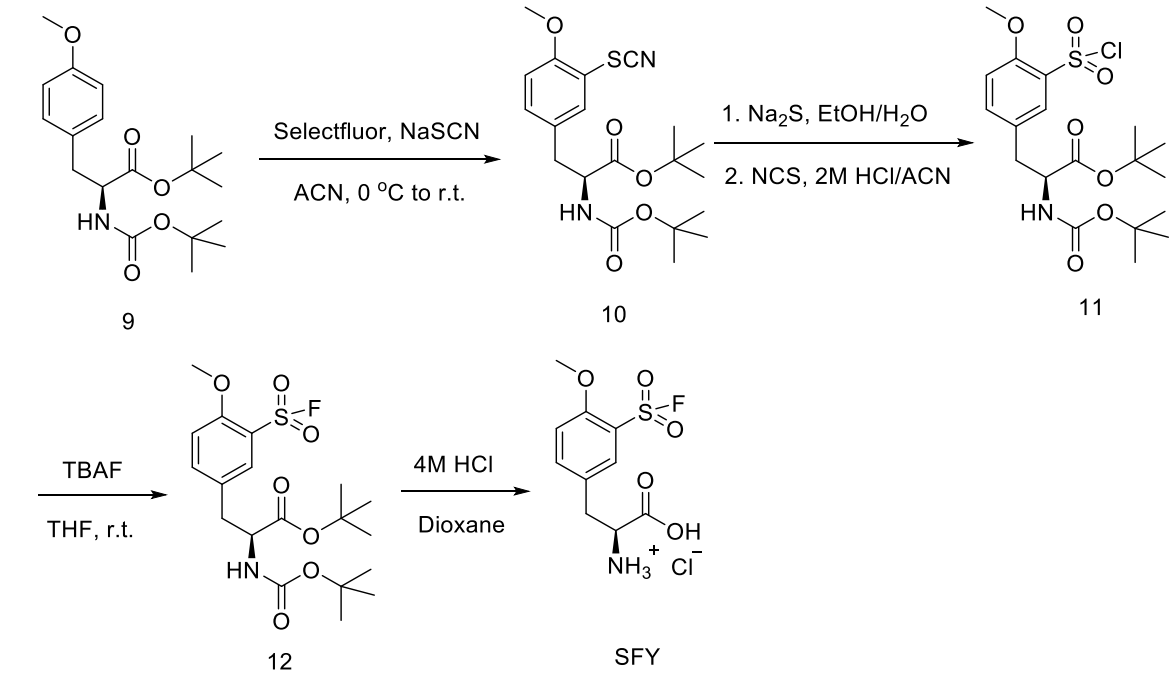
8b: NHSF-C3

Sodium 1-((4-(4-(fluorosulfonyl)benzamido)butanoyl)oxy)-2,5-dioxopyrrolidine-3-sulfonate. (**8b, NHSF-C3**). ^{1}H NMR (DMSO-d_6): δ 8.94 (t, *J* = 5.6 Hz, 1H), 8.24 (d, *J* = 8.4 Hz, 2H), 8.17 (d, *J* = 8.4 Hz, 2H), 3.95 (d, *J* = 8.8 Hz, 1H), 3.40 -3.35 (m, 2H), 3.20-2.83 (m, 2H), 2.78 (t, *J* = 7.6 Hz, 2H), 1.94-1.87 (m, 2H), ^{13}C NMR (DMSO-d_6): δ 168.8, 165.4, 164.2, 141.7, 133.3 (d, *J* = 24 Hz, C-F), 129.0, 128.6, 56.3, 38.6, 31.0, 27.9, 24.0. HRMS calcd for $C_{15}H_{14}FN_2Na_2O_{10}S_2$ $[M+Na]^+$ 510.9864, found: 510.9851.



8c: NHSF-C7

sodium 1-((8-(4-(fluorosulfonyl)benzamido)octanoyl)oxy)-2,5-dioxopyrrolidine-3-sulfonate (**8c, NHSF-C7**).
^{1}H NMR (DMSO-d_6): δ 8.84 (t, *J* = 5.6 Hz, 1H), 8.24 (d, *J* = 8.4 Hz, 2H), 8.15 (d, *J* = 8.4 Hz, 2H), 3.94 (d, *J* = 8.4 Hz, 1H), 3.30-3.26 (m, 2H), 3.15-2.82 (m, 2H), 2.65 (t, *J* = 7.6 Hz, 2H), 1.64-1.52 (m, 4H), 1.38-1.30 (m, 6H). ^{13}C NMR (DMSO-d_6): δ 168.8, 165.3, 164.6, 141.5, 133.4 (d, *J* = 24 Hz, C-F), 129.2, 128.6, 56.3, 31.0, 30.2, 28.8, 28.2, 28.0, 26.3, 24.3. HRMS calcd for $C_{19}H_{23}FN_2NaO_{10}S_2$ $[M+H]^+$ 545.0670, found: 545.0660.



9 10 11

12 SFY

Synthesis of tert-butyl (S)-2-((tert-butoxycarbonyl)amino)-3-(4-methoxy-3-thiocyanatophenyl)propanoate (**10**). To a stirred solution of Selectfluor (2.4g, 6.8 mmol) and

NaSCN (550 mg, 6.8 mmol) in ACN (20 mL) was added compound **9** (800 mg, 2.27 mmol) in ACN (5 mL) at 0 °C under N_2. The reaction mixture was allowed to stir at r.t. for overnight. Then the solvent was removed under reduced pressure and the residue was dissolved in 25 mL EtOAc. The organic phrase was washed sequentially by H_2O (25 mL) and brine (25 mL). The organic phase was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure to give the crude product, which was then purified by column chromatography (silica gel, Hexane: EtOAc = 5:1) to give a yellow solid (637 mg, 69 %). ^{1}H NMR ($CDCl_3$): δ 7.33 (d, *J* = 2.0 Hz, 1H), 7.18 (dd, *J* = 8.4 Hz, *J* = 2.0 Hz, 1H), 6.85 (d, *J* = 8.4 Hz, 1H), 5.04 (d, *J* = 8.0 Hz, 1H), 4.43 - 4.38 (m, 1H), 3.89 (s, 3H), 3.10-2.95 (m, 2H), 1.43 (d, 18H). ^{13}C NMR ($CDCl_3$): δ 170.7, 155.6, 155.1, 131.8, 131.1, 130.7, 111.5, 110.4, 82.6, 80.0, 56.4, 54.9, 37.7, 28.4, 28.2. HRMS calcd for $C_{20}H_{28}N_2NaO_5S$ $[M+Na]^+$ 431.1611, found: 431.1627.

Synthesis of tert-butyl (S)-2-((tert-butoxycarbonyl)amino)-3-(3-(fluorosulfonyl)-4-methoxyphenyl)propanoate (**12**). To a solution of **10** (620 mg, 1.52 mmol) in EtOH (3.5 mL) was added $Na_2S{\cdot}9H_2O$ (730 mg, 3.0 mmol) in H_2O (12 mL) at 60 °C. The reaction mixture was then heated at 85°C for 2 h. The reaction mixture was then allowed to cool down to r.t., and 10 mL H_2O was added. The mixture was then adjusted to pH 6.5 with acetic acid and extracted with EtOAc (10 mL × 3). The combined organic phase was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure to give the crude thiol product as yellow solid, which was immediately used for the next step. To a stirred solution of *N*-chlorosuccinimide (0.65 g, 4.9 mmol) in 2M HCl (0.6 mL) and acetonitrile (2.5 mL) was added dropwise crude thiol in acetonitrile (1 mL) dropwise at 0 °C. The mixture was stirred at 0 °C for another 30 min. The mixture was then diluted with EtOAc (10 mL), the organic phrase was washed sequentially by H_2O (10 mL) and brine (10 mL). The organic phase was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure to give the crude sulfonyl chloride (compound **11**) product as yellow oil (604 mg).

Half of the newly prepared crude sulfonyl chloride was used for the next step. To a stirred solution of Compound **11** (300 mg, 0.67 mmol) in anhydrous THF (2 mL) was added 1.3 mL 1M tetrabutylammonium fluoride (TBAF, 1.33 mmol) in THF. The mixture was stirred at r.t. for 1 h and the completion of reaction was monitored by mass spectrum. The mixture was then diluted with EtOAc (10 mL), the organic phrase was washed sequentially by H_2O (10mL) and brine (10 mL). The organic phase was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure to give the crude product, which was then purified by column chromatography (silica gel, DCM: EtOAc = 25:1) to give compound **12** as white solid (76 mg, 24 % for 3 steps). ^{1}H NMR ($CDCl_3$): δ 7.70 (d, *J* = 2.0 Hz, 1H), 7.51 (dd, *J* = 8.4 Hz, *J* = 2.0 Hz, 1H), 7.03 (d, *J* = 8.4 Hz, 1H), 5.07 (d, *J* = 8.4 Hz, 1H), 4.43 - 4.38 (m, 1H), 3.98 (s, 3H), 3.16-2.98 (m, 2H), 1.42 (d, 18H). ^{13}C NMR ($CDCl_3$): δ 170.3, 157.1, 155.1, 138.5, 132.1, 129.4, 121.1 (d, *J* = 23 Hz, C-F), 112.9, 83.0, 80.1, 56.7, 54.8, 37.3, 28.4, 28.1. HRMS calcd for $C_{19}H_{28}FNNaO_7S$ $[M+Na]^+$ 456.1463, found: 456.1473.

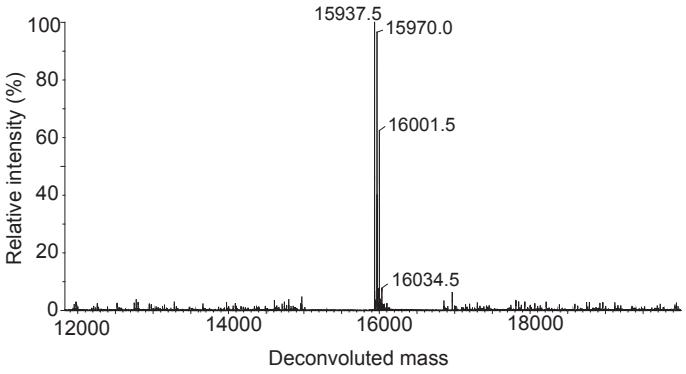
Synthesis of (S)-1-carboxy-2-(3-(fluorosulfonyl)-4-methoxyphenyl)ethan-1-aminium (SFY). Compound **12** (76 mg, 0.18 mmol) was stirred in 4 M HCl in dioxane (0.5 mL) at r.t. for 24 h. Then 5 mL diethyl ether was added to the reaction mixture, and a white precipitate was formed and collected by centrifuge (10 min, 3,000 rpm). The white solid was further dried under reduced pressure to give SFY in HCl salt form (52 mg, 92 %). ^{1}H NMR (D_2O): δ 7.87 (d, *J* = 2.4 Hz, 1H), 7.74 (dd, *J* = 8.8 Hz, *J* = 2.4 Hz, 1H), 7.32 (d, *J* = 8.8 Hz, 1H), 4.30 (t, *J* = 6.8 Hz, 1H), 4.01 (s, 3H), 3.37-3.24 (m, 2H). ^{13}C NMR (D_2O): δ 171.3, 157.5, 139.4, 131.6, 126.8, 119.5 (d, *J* = 21 Hz, C-F), 114.3, 56.7, 54.0, 34.4. HRMS calcd for $C_{10}H_{13}FNO_5S$ $[M+H]^+$ 278.0493, found: 278.0510.

Characterization of protein-carbohydrate linkage with mass spectrometry

We have tried to characterize the Siglec-7v(104SFY) cross-linked with azido-GD3 with mass spectrometry. When analyzing at the digested peptide level, proteolytic cleavage was often hindered by glycosylation (due to GD3 cross-linking), resulting in large, highly glycosylated peptide moieties, which drastically reduced the sensitivity in detecting the GD3-crosslinked peptide. Glycosylated peptides have been reported difficult to dissociate in MS

measurements.[4] In addition, while carbohydrate linked to natural amino acid residues through O/N-C bond has been characterized by MS,[4] the linkage in our work was on the unnatural amino acid SFY and was a different O-S bond, which may result in different fragmentation pattern and require a different dissociation method as well as MS data analysis software. When analyzing on the intact protein level, chemical properties of carbohydrate make it difficult to preserve both protein and carbohydrate intact during the process. For example, GD3 was unstable under the acidic separation conditions of LC. In the end, we succeeded in MS analysis on the protein level. A peak was observed at 17054 Da, matching exactly Siglec-7v(104SFY) cross-linked with azido-GD3 (expected 17054 Da). A large peak was also detected at 16051 Da, possibly corresponding to Siglec-7v with the cross-linked azido-GD3 hydrolyzed during the MS analysis procedures. The peak intensity for the cross-linked complex could be lowered by: 1) acidic separation condition of LC resulted in GD3 loss, and 2) glycan-crosslinked protein has lower signal than non-glycan modified protein.

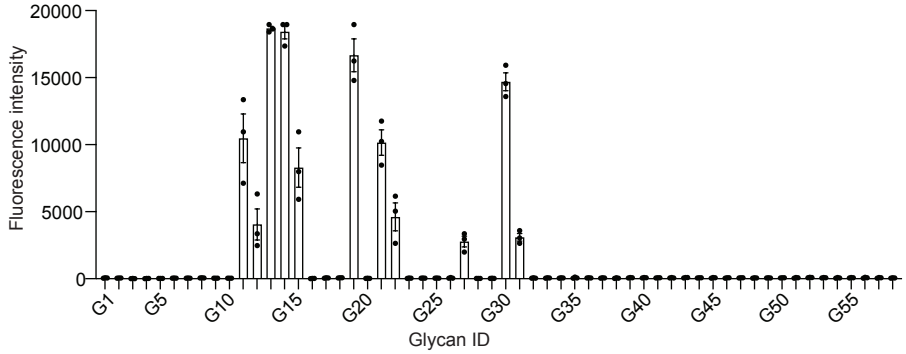
We have also confirmed the covalent linkage with various other methods ranging from fundamental chemical aspects to biological functions, such as gel separation, Western blot, binding with cells, as well as enhancement of NK killing. Since we already know the identity of the GD3 in the cross-linked complex, results obtained with these various methods provide sufficient evidence that a covalent linkage was formed between Siglec-7v(104SFY) and the azido-GD3.



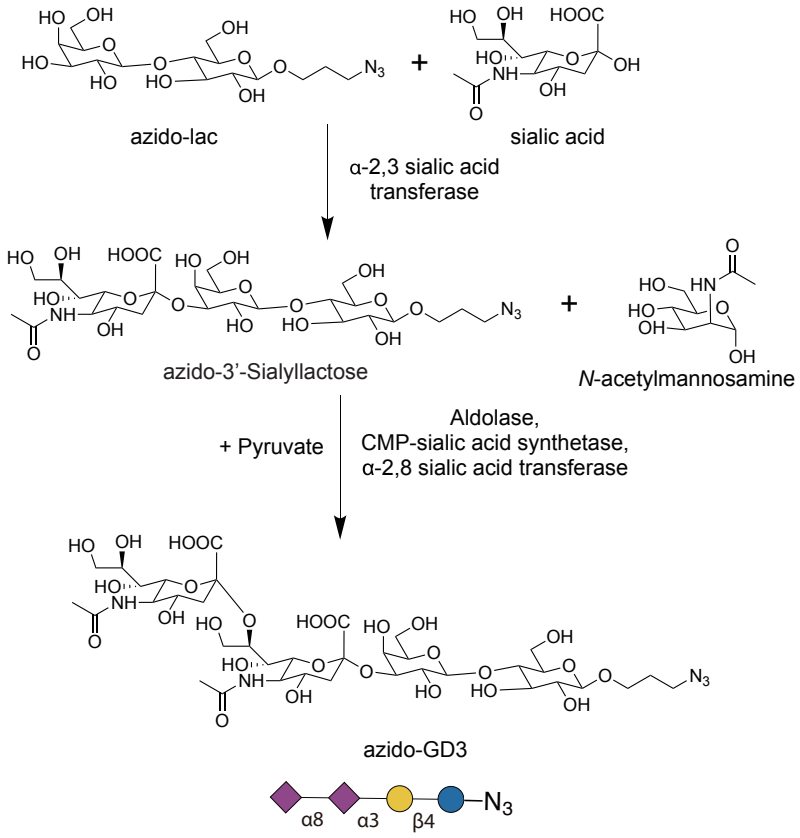
Supplementary Figure 1. Mass spectrum of the intact siglec-7v protein.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	G1	G1	G1	G1	G1	G1	G2	G2	G2	G2	G2	G2	G3	G3	G3	G3	G3	G3	G4	G4	G4	G4	G4	G4	G5	G5	G5	G5	G5	G5
2	G6	G6	G6	G6	G6	G6	G7	G7	G7	G7	G7	G7	G8	G8	G8	G8	G8	G8	G9	G9	G9	G9	G9	G9	G10	G10	G10	G10	G10	G10
3	G11	G11	G11	G11	G11	G11	G12	G12	G12	G12	G12	G12	G13	G13	G13	G13	G13	G13	G14	G14	G14	G14	G14	G14	G15	G15	G15	G15	G15	G15
4	G16	G16	G16	G16	G16	G16	G17	G17	G17	G17	G17	G17	G18	G18	G18	G18	G18	G18	G19	G19	G19	G19	G19	G19	G20	G20	G20	G20	G20	G20
5	G21	G21	G21	G21	G21	G21	G22	G22	G22	G22	G22	G22	G23	G23	G23	G23	G23	G23	G24	G24	G24	G24	G24	G24	G25	G25	G25	G25	G25	G25
6	G26	G26	G26	G26	G26	G26	G27	G27	G27	G27	G27	G27	G28	G28	G28	G28	G28	G28	G29	G29	G29	G29	G29	G29	G30	G30	G30	G30	G30	G30
7	G31	G31	G31	G31	G31	G31	G32	G32	G32	G32	G32	G32	G33	G33	G33	G33	G33	G33	G34	G34	G34	G34	G34	G34	G35	G35	G35	G35	G35	G35
8	G36	G36	G36	G36	G36	G36	G37	G37	G37	G37	G37	G37	G38	G38	G38	G38	G38	G38	G39	G39	G39	G39	G39	G39	G40	G40	G40	G40	G40	G40
9	G41	G41	G41	G41	G41	G41	G42	G42	G42	G42	G42	G42	G43	G43	G43	G43	G43	G43	G44	G44	G44	G44	G44	G44	G45	G45	G45	G45	G45	G45
10	G46	G46	G46	G46	G46	G46	G47	G47	G47	G47	G47	G47	G48	G48	G48	G48	G48	G48	G49	G49	G49	G49	G49	G49	G50	G50	G50	G50	G50	G50
11	G51	G51	G51	G51	G51	G51	G52	G52	G52	G52	G52	G52	G53	G53	G53	G53	G53	G53	G54	G54	G54	G54	G54	G54	G55	G55	G55	G55	G55	G55
12	G56	G56	G56	G56	G56	G56	G57	G57	G57	G57	G57	G57	G58	G58	G58	G58	G58	G58							NC	NC	NC	NC	NC	NC
13	PC1	PC1	PC1	PC1	PC1	PC1	PC2	PC2	PC2	PC2	PC2	PC2	PC3	PC3	PC3	PC3	PC3	PC3	PC4	PC4	PC4	PC4	PC4	PC4	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER

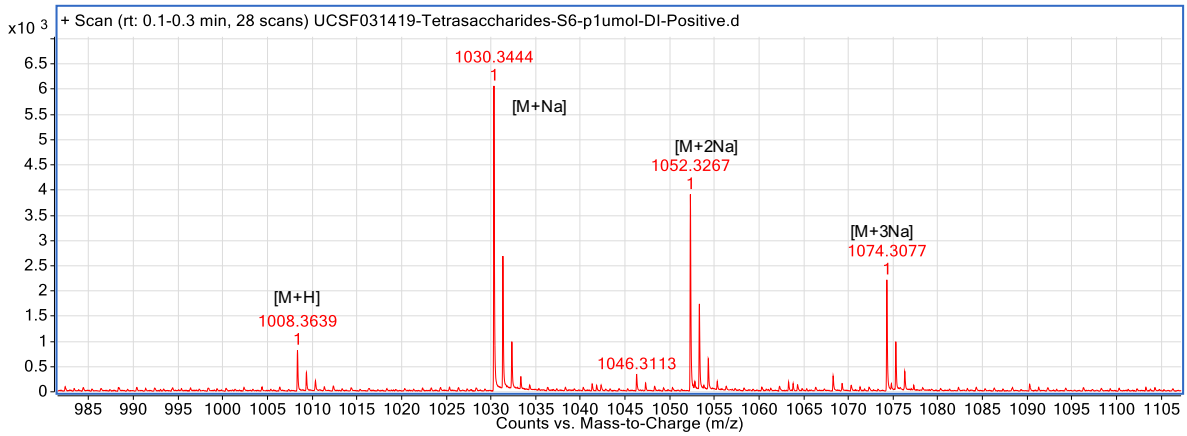
Supplementary Figure 2. List of glycosphingolipid glycans on the glycan microarray. The array was printed with 58 glycans, each with six repeats. The identity of each glycan is shown in Supplementary Table 2.



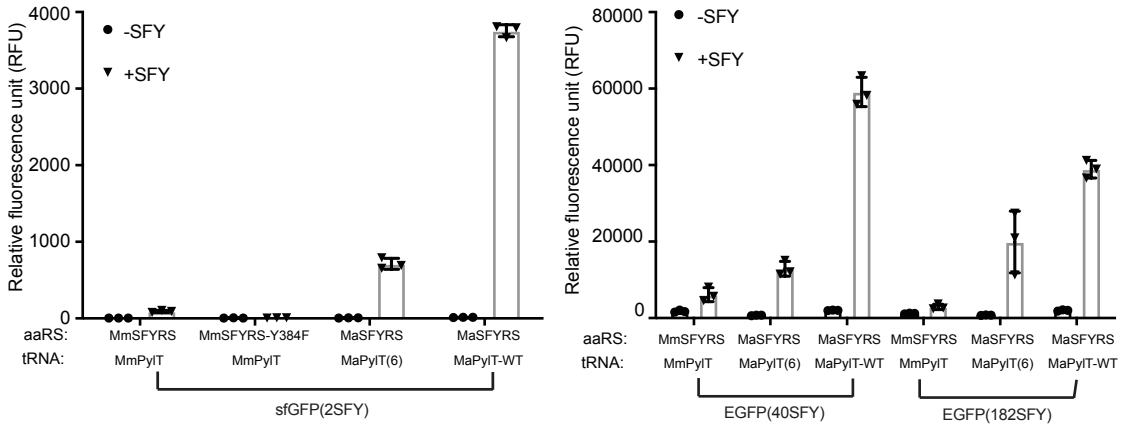
Supplementary Figure 3. Glycan microarray analysis of Fc-Siglec-7 commercially available from R&D. Siglec-7v purified from *E. coli* showed specific signals in binding with all sialoglycans that are known binders of Siglec 7, including G20 and G28 (Figure 1c). The commercial Fc-Siglec-7 expressed in mouse cells showed a similar binding pattern but had no signal for G20 and G28, possibly because of glycosylation effect and/or interference of the Fc tag. n=3 independent samples. The bar height and error bar represent mean ± SEM.



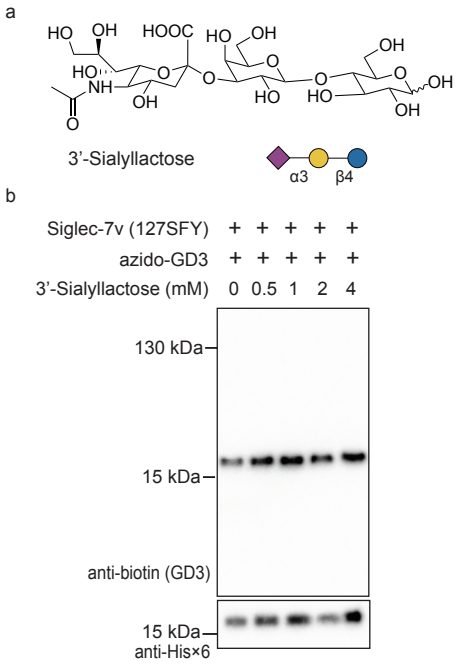
Supplementary Figure 4. Chemo-enzymatic synthesis of azido-GD3.



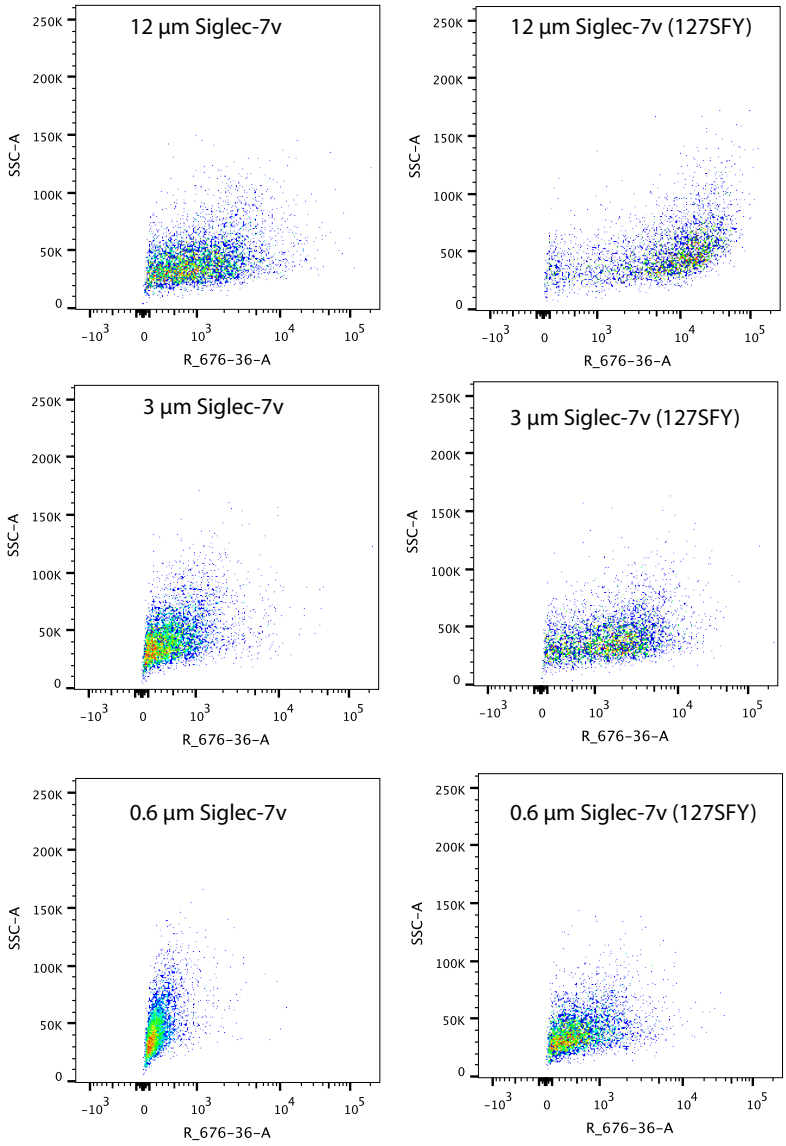
Supplementary Figure 5. ESI-MS of azido-GD3 glycan ligand.



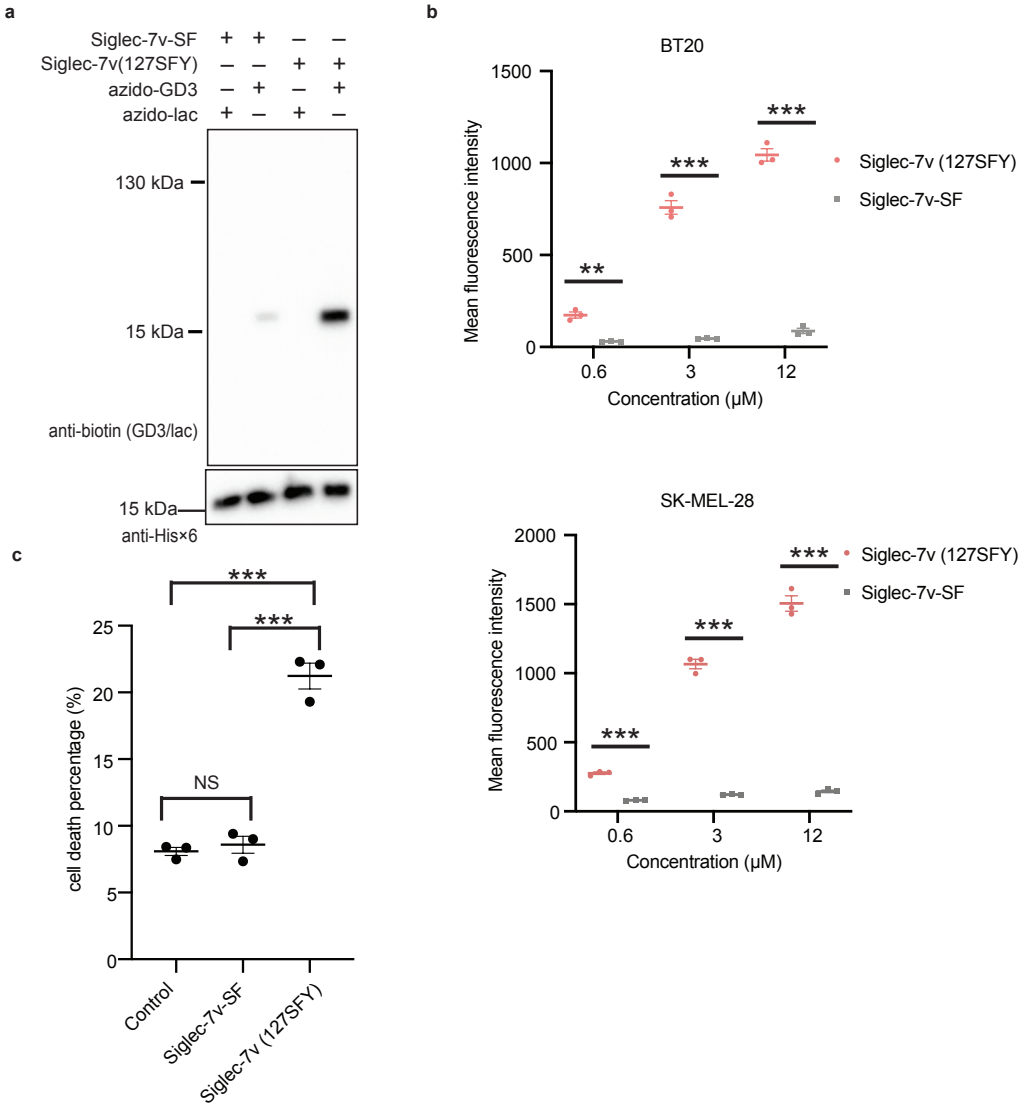
Supplementary Figure 6. Comparison of SFY incorporation into different sites of GFP using Mm-tRNA^{Pyl}/MmSFYRS and Ma-tRNA^{Pyl}/MaSFYRS in *E. coli*. Fluorescence intensities of the expressed sfGFP(2SFY), EGFP(40SFY), and EGFP(182SFY) in *E. coli* cells using the indicated tRNA^{Pyl} and SFYRS were quantified with flow cytometry. The MaPylT-WT tRNA afforded much higher incorporation efficiency than the mutant MaPylT(6) tRNA.[5] In all cases, the MaPylT-WT and MaSFYRS pair afforded the highest incorporation efficiency of SFY in *E. coli*. n = 3 independent samples. The bar height and error bar represent mean ± SEM.



Supplementary Figure 7. Addition of 3'-sialyllactose did not reduce the cross-linking of Siglec-7v(127FSY) with azido-GD3. a) Structure of 3'-Sialyllactose. **b)** The addition of 3'-Sialyllactose didn't reduce the cross-linking of Siglec-7v(127SFY) with azido-GD3. Siglec-7v(127SFY) (60 μM) was incubated with 2 mM azido-GD3, then supplemented without or with different concentrations of 3'-Sialyllactose. Samples are boiled and subjected for Western blot analysis.



Supplementary Figure 8. Representative cytomagrams for Figure 5e.



Supplementary Figure 9. A comparative study between NHSF pretreated Siglec-7v (Siglec-7v-SF) and Siglec-7v(127SFY). a) Siglec-7v(127SFY) cross-linked azido-GD3 efficiently, while Siglec-7v-SF could not. Siglec-7v(127SFY) or Siglec-7v-SF was incubated with azido-GD3 or azido-lac followed with Western blot detection. The azido group was click reacted with alkyne-biotin for Western blot detection of GD3/lac. **b)** Siglec-7v(127SFY) bound to the surface of BT20 (top panel) and SK-MEL-28 (bottom panel) cell lines in a dose-dependent manner, while Siglec-7v-SF could not bind with eithercells. Cells were treated with protein, washed, stained with a fluorescently labeled antibody specific for the Hisx6 tag appended at the C-terminus of Siglec-7v, and quantified with flow cytometry. **c)** Siglec-7v(127SFY) significantly enhanced NK cell killing of cancer cells, while Siglec-7v-SF could not. Pre-stained BT-20 cells were incubated with 12 μM Siglec-7v-SF or Siglec-7v(127SFY) for 2 h. Cells were then washed and incubated with NK-92 cells for 4 h. Cells were stained with propidium iodide and NK cytotoxicity was evaluated by flow cytometry. Control group: no protein treatment. The line and error bar represent mean ± SEM; n = 3 independent batches of Siglec-7v proteins. ** $p < 0.01$, *** $p < 0.001$; NS, not significant, two-sided t test. p value in panel **b** for BT20 cells: 0.0010 for 0.6 μM, 0.000042 for 3 μM, and 0.000013 for 12 μM. p value in panel **b** for SK-MEL-28 cells: 0.000031 for 0.6 μM, 0.000010 for 3 μM, and 0.000017 for 12 μM. p value in panel **c**: 0.52 for control and Siglec-7v-SF, 0.0002 for control and Siglec-7v(127SFY), and 0.0004 for Siglec-7v-SF and Siglec-7v-(127SFY).

Supplementary reference

1. Szunerits, S. *et al.* Label-free detection of lectins on carbohydrate-modified boron-doped diamond surfaces. *Anal. Chem.* **82,** 8203-8210 (2010).
2. Zhou, H. *et al.* Introduction of a crystalline, shelf-stable reagent for the synthesis of sulfur(vi) fluorides. *Org. Lett.* **20,** 812–815 (2018).
3. Yang, X. *et al.* Development of covalent ligands for g protein-coupled receptors: a case for the human adenosine A3 receptor. *J. Med. Chem.* **62,** 3539–3552 (2019).
4. Riley, N.M., Malaker, S.A., Driessen, M.D., Bertozzi, C.R. Optimal dissociation methods differ for N- and O-glycopeptides. *J. Proteome Res.* **19**, 3286-3301 (2020).
5. Willis, J. C. W. & Chin, J. W. Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs. *Nat. Chem.* **10,** 831–837 (2018).

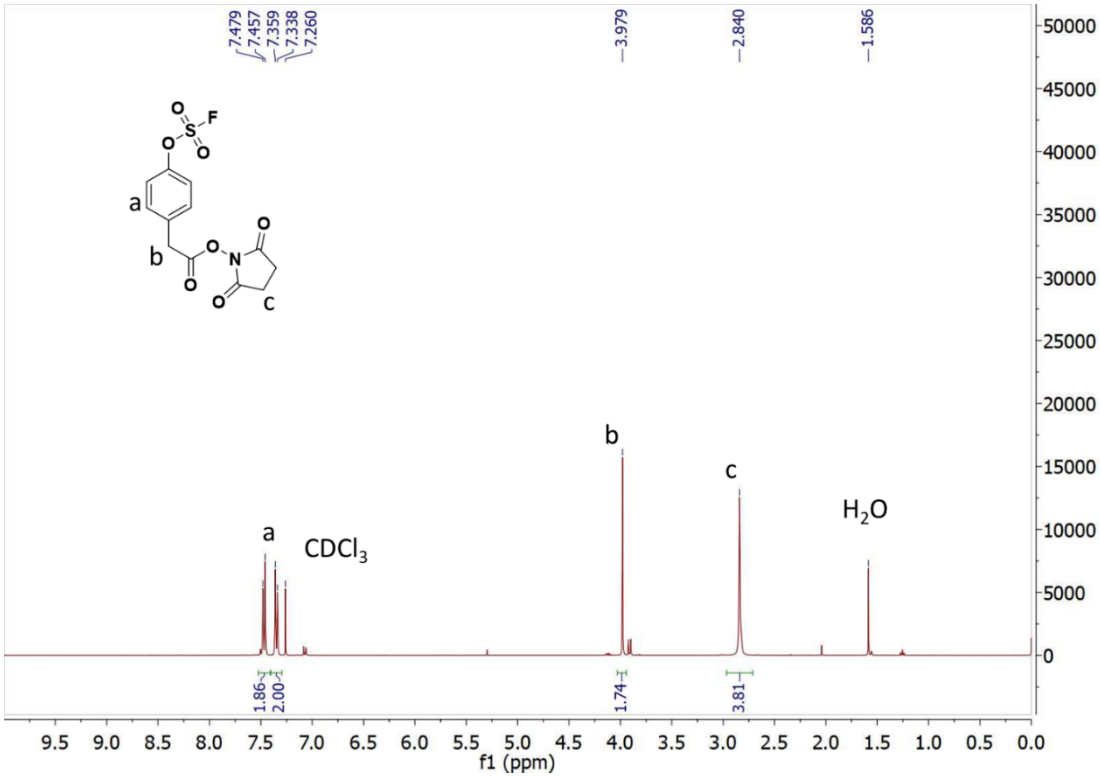
Supplementary Table 1: Primers for cloning

primer	oligonucleotide sequence (5'→3')
siglec7(20TAG)-For	ACATATGCAGTAGAGTAATCGTAAAGATTATAG
siglec7(20TAG)-Rev	ATATCTCCTTCTTAAAGTTAAAC
siglec7(24TAG)-For	AAGTAATCGTTAGGATTATAGCCTGAC
siglec7(24TAG)-Rev	TTCTGCATATGTATATCTCC
siglec7(75TAG)-For	TATTAGTTGGTAGGCCCCGGTTGC
siglec7(75TAG)-Rev	TCATTACCGGCGCGAAAC
siglec7(104TAG)-For	CCCGCAGACCTAGAATTGTACCC
siglec7(104TAG)-Rev	TCACCCAGCAGATGAAAG
siglec7(127TAG)-For	TCGCATGGAATAGGGTAATATCAAATGG
siglec7(127TAG)-Rev	AAGAAATAGCGGCCTGCA
siglec7(129TAG)-For	GGAAAAAGGTTAGATCAAATGGAATTAC
siglec7(129TAG)-Rev	ATGCGAAAGAAATAGCGG
siglec7(130TAG)-For	AAAAGGTAATTAGAAATGGAATTACAAGTAC
siglec7(130TAG)-Rev	TCCATGCGAAAGAAATAG
siglec7(131TAG)-For	AGGTAATATCTAGTGGAATTACAAGTAC
siglec7(131TAG)-Rev	TTTTCCATGCGAAAGAAATAG
siglec7(135TAG)-For	ATGGAATTACTAGTACGATCAGC
siglec7(135TAG)-Rev	TTGATATTACCTTTTTCCATG
siglec7(20Gly)-For	ACATATGCAGGGTAGTAATCGTAAAGATTATAG
siglec7(20Gly)-Rev	ATATCTCCTTCTTAAAGTTAAAC
siglec7(24Gly)-For	AAGTAATCGTGGTGATTATAGCCTGACC
siglec7(24Gly)-Rev	TTCTGCATATGTATATCTCC
siglec7(75Gly)-For	TATTAGTTGGGGTGCCCCGGTTGCC
siglec7(75Gly)-Rev	TCATTACCGGCGCGAAAC
siglec7(104Gly)-For	CCCGCAGACCGGTAATTGTACCC
siglec7(104Gly)-Rev	TCACCCAGCAGATGAAAG
siglec7(127Gly)-For	TCGCATGGAAGGTGGTAATATCAAATGGAATTACAAG
siglec7(127Gly)-Rev	AAGAAATAGCGGCCTGCA
siglec7(131Gly)-For	AGGTAATATCGGTTGGAATTACAAGTACGATC
siglec7(131Gly)-Rev	TTTTCCATGCGAAAGAAATAG
siglec7(135Gly)-For	ATGGAATTACGGTTACGATCAGCTGAGTGTTAATG
siglec7(135Gly)-Rev	TTGATATTACCTTTTTCCATG
MmSFYRS-SpeI-F	AACAGGAGGAATTACTAGTATGGATAAAAAGCCTTTG
MmSFYRS-SalI-R	GATGATGATGATGATGGTCGACTTACAGGTTAGTAGAA
MaPylRS-SpeI-F	TAACAGGAGGAATTACTAGTATGACCGTGAAGTAC
MaPylRS-SalI-R	TGATGATGATGATGGTCGACTTAATTGATTTTGGCACCA
MaPylT(6)-F	CTAGCATAGCGGGGTTCGACGCCCCGGTCTCTCGCCAAATTCGAAAAGCCTGC
MaPylT(6)-R	GTTTTAGAGACCCGCTGGTCGCCGGACCGTCCCCCAATGCGGGGCGCATCT
MaSFYRS-R1	CAGGTTTGGCGCCAGC
MaSFYRS-F2	GCTGGCGCCAAACCTGTTGAGCGTGGCTCGTGACCTGCG
MaSFYRS-R2	CAGCATGGTGAACTCC
MaSFYRS-F3	GGAGTTCACCATGCTGGCTCTGATGGATATGGGTCCGC
MaSFYRS-R3	CGGCTCATGCACGTC
MaSFYRS-F4	CGTGCATGAGCCGACAAGCGGTGCTGGTTTTG
MaPylT(wt)-F	AAAACCTAGCCAGCGGGGTTCGACG
MaPylT(wt)-R	AGAGACCCGCTGGTCGCC

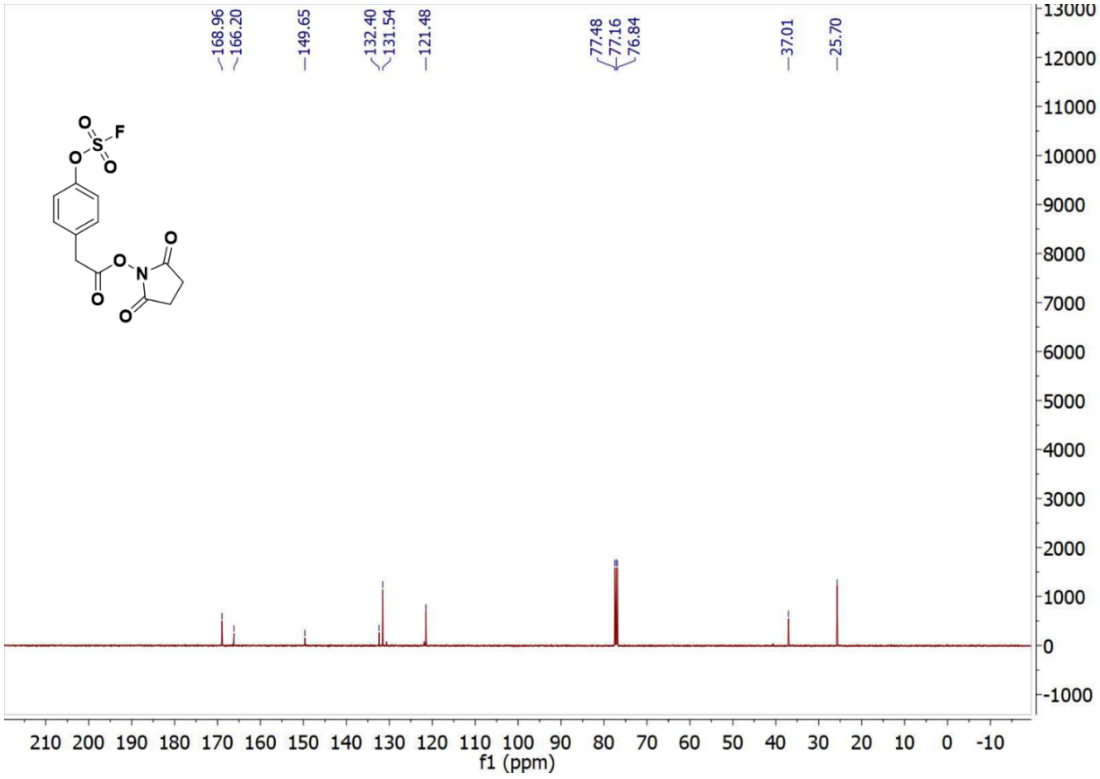
Supplementary Table 2: Name, sequence, and structure of the 58 glycans on the glycan microarray

G1	Neu5Acα2–3Galβ1–4Glc	**G30**	Neu5Acα2–8Neu5Acα2–3Galβ1–3GalNAcβ1–4(Neu5Aca2–3)Galβ1–4Glc
G2	Neu5Gcα2–3Galβ1–4Glc	**G31**	GalNAcβ1-4(Neu5Acα2–8Neu5Acα2–8Neu5Acα2–3)Galβ1–4Glc
G3	Kdnα2–3Galβ1–4Glc	**G32**	Neu5Acα2–8Neu5Acα2–8Neu5Acα2–3Galβ1–4Glc
G4	Neu5Ac8Meα2–3Galβ1–4Glc	**G33**	GlcNAcβ1–3Galβ1–4Glc
G5	Neu5Acα2–3(GalNAcβ1–4)Galβ1–4Glc	**G34**	Galβ1–3GlcNAcβ1–3Galβ1–4Glc
G6	Neu5Gcα2–3(GalNAcβ1–4)Galβ1–4Glc	**G35**	Galβ1–4GlcNAcβ1–3Galβ1–4Glc
G7	Kdnα2–3(GalNAcβ1–4)Galβ1–4Glc	**G36**	Galβ1–4(Fucα1–3)GlcNAcβ1–3Galβ1–4Glc
G8	Neu5Acα2–3(Galβ1–3GalNAcβ1–4)Galβ1–4Glc	**G37**	Neu5Acα2–3Galβ1–4GlcNAcβ1–3Galβ1–4Glc
G9	Neu5Gcα2–3(Galβ1–3GalNAcβ1–4)Galβ1–4Glc	**G38**	Neu5Gcα2–3Galβ1–4GlcNAcβ1–3Galβ1–4Glc
G10	Kdnα2–3(Galβ1–3GalNAcβ1–4)Galβ1–4Glc	**G39**	Kdnα2–3Galβ1–4GlcNAcβ1–3Galβ1–4Glc
G11	Neu5Acα2–8Neu5Acα2–3Galβ1–4Glc	**G40**	Neu5Ac8Meα2–3Galβ1–4GlcNAcβ1–3Galβ1–4Glc
G12	Neu5Acα2–8Neu5Gcα2–3Galβ1–4Glc	**G41**	Neu5Acα2–3Galβ1–3GlcNAcβ1–3Galβ1–4Glc
G13	Neu5Acα2–8Kdncα2–3Galβ1–4Glc	**G42**	Neu5Gcα2–3Galβ1–3GlcNAcβ1–3Galβ1–4Glc
G14	Neu5Gcα2–8Neu5Acα2–3Galβ1–4Glc	**G43**	Kdnα2–3Galβ1–3GlcNAcβ1–3Galβ1–4Glc
G15	Neu5Gcα2–8Neu5Gcα2–3Galβ1–4Glc	**G44**	Neu5Ac8Meα2–3Galβ1–3GlcNAcβ1–3Galβ1–4Glc
G16	Kdnα2–8Neu5Acα2–3Galβ1–4Glc	**G45**	Neu5Gcα2–3Galβ1–4(Fucα1–3)GlcNAcβ1–3Galβ1–4Glc
G17	Kdnα2–8Neu5Gcα2–3Galβ1–4Glc	**G46**	Kdnα2–3Galβ1–4(Fucα1–3)GlcNAcβ1–3Galβ1–4Glc
G18	Kdnα2–8Kdnα2–3Galβ1–4Glc	**G47**	Galα1–4Galβ1–4Glc
G19	Neu5Acα2–8Neu5Acα2–3(GalNAcβ1–4)Galβ1–4Glc	**G48**	Galα1–3Galβ1–4Glc
G20	Neu5Acα2–8Neu5Gcα2–3(GalNAcβ1–4)Galβ1–4Glc	**G49**	GalNAcβ1–3Galα1–4Galβ1–4Glc
G21	Neu5Gcα2–8Neu5Acα2–3(GalNAcβ1–4)Galβ1–4Glc	**G50**	GalNAcβ1–3Galα1–3Galβ1–4Glc
G22	Neu5Gcα2–8Neu5Gcα2–3(GalNAcβ1–4)Galβ1–4Glc	**G51**	Galβ1–3GalNAcβ1–3Galα1–4Galβ1–4Glc
G23	Kdnα2–8Neu5Acα2–3(GalNAcβ1–4)Galβ1–4Glc	**G52**	Galβ1–3GalNAcβ1–3Galα1–3Galβ1–4Glc
G24	Kdnα2–8Neu5Gcα2–3(GalNAcβ1–4)Galβ1–4Glc	**G53**	Fucα1-2Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glc
G25	Kdnα2–8Kdnα2–3(GalNAcβ1–4)Galβ1–4Glc	**G54**	Neu5Gcα2–3Galβ1–3GalNAcβ1–3Galα1–4Galβ1–4Glc
G26	Neu5Acα2–3Galβ1–3GalNAcb1–4(Neu5Aca2–3)Galβ1–4Glc	**G55**	Kdnα2–3Galβ1–3GalNAcβ1–3Galα1–4Galβ1–4Glc
G27	Neu5Acα2–8Neu5Acα2–3(Galβ1–3GalNAcβ1–4)Galβ1–4Glc	**G56**	Neu5Acα2–3Galβ1–3GalNAcβ1–3Galα1–3Galβ1–4Glc
G28	Neu5Gcα2–8Neu5Gcα2–3(Galβ1–3GalNAcβ1–4)Galβ1–4Glc	**G57**	Neu5Gcα2–3Galβ1–3GalNAcβ1–3Galα1–3Galβ1–4Glc
G29	Kdnα2–8Neu5Gcα2–3(Galβ1–3GalNAcβ1–4)Galβ1–4Glc	**G58**	Kdnα2–3Galβ1–3GalNAcβ1–3Galα1–3Galβ1–4Glc

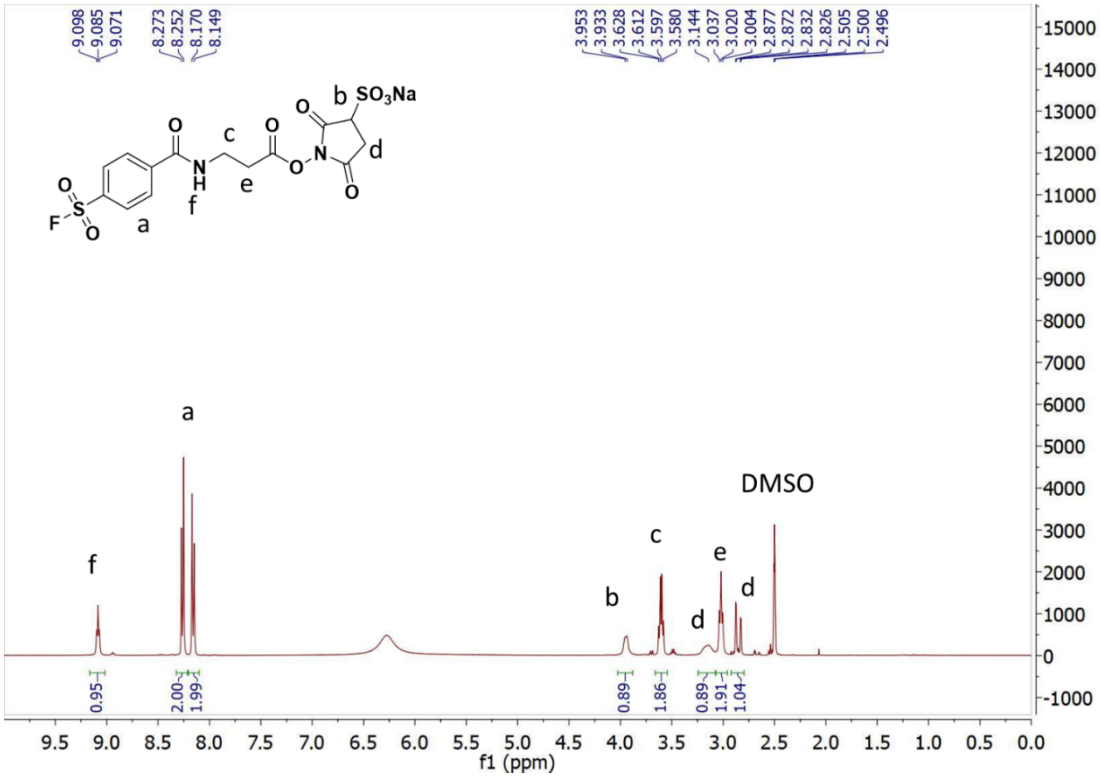
NMR spectra



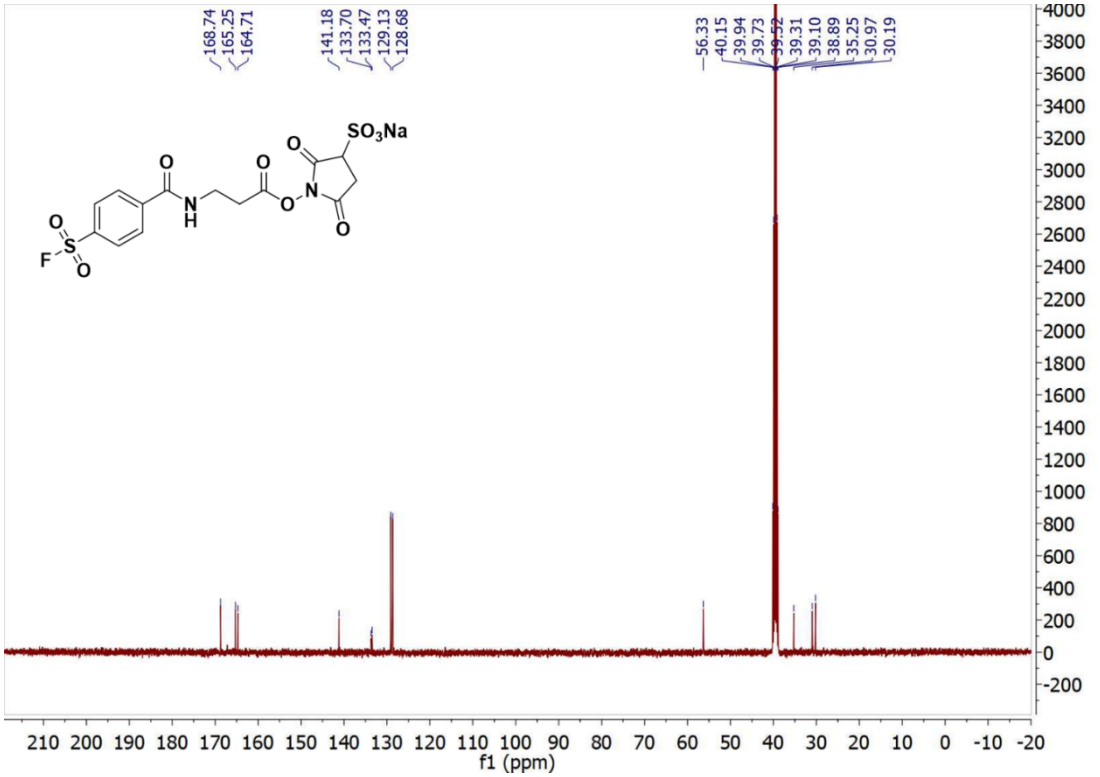
^{1}H NMR ($CDCl_3$) of NHFS.



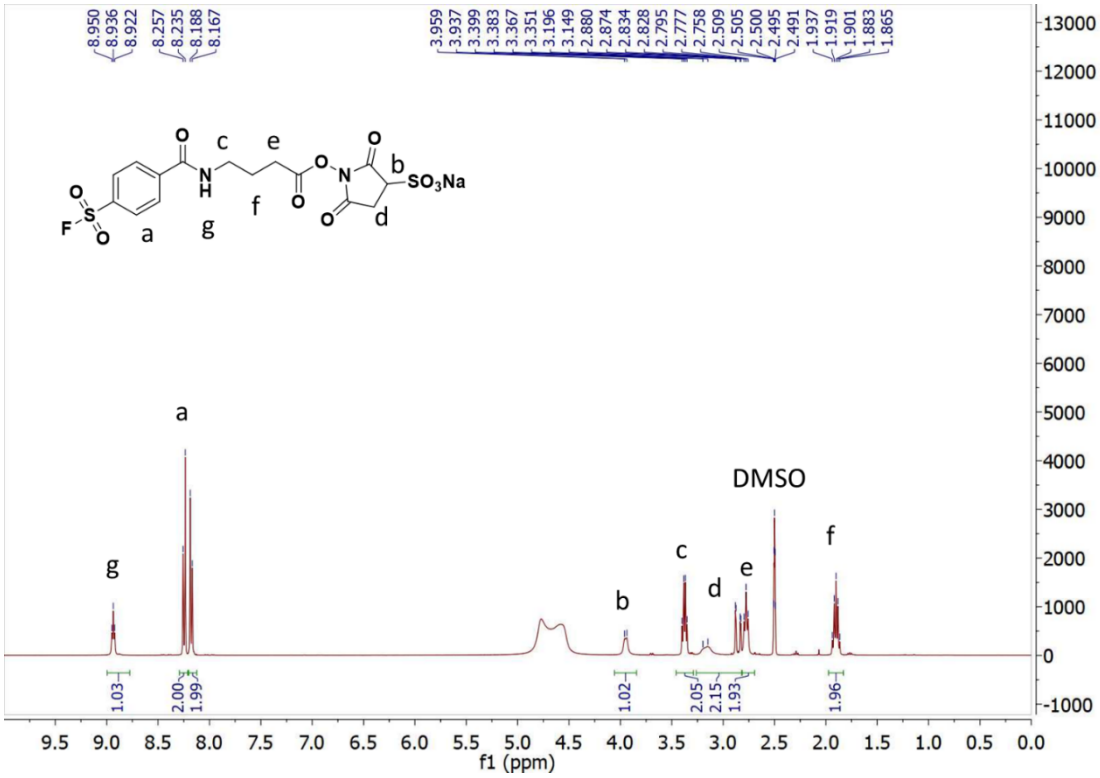
^{13}C NMR ($CDCl_3$) of NHFS.



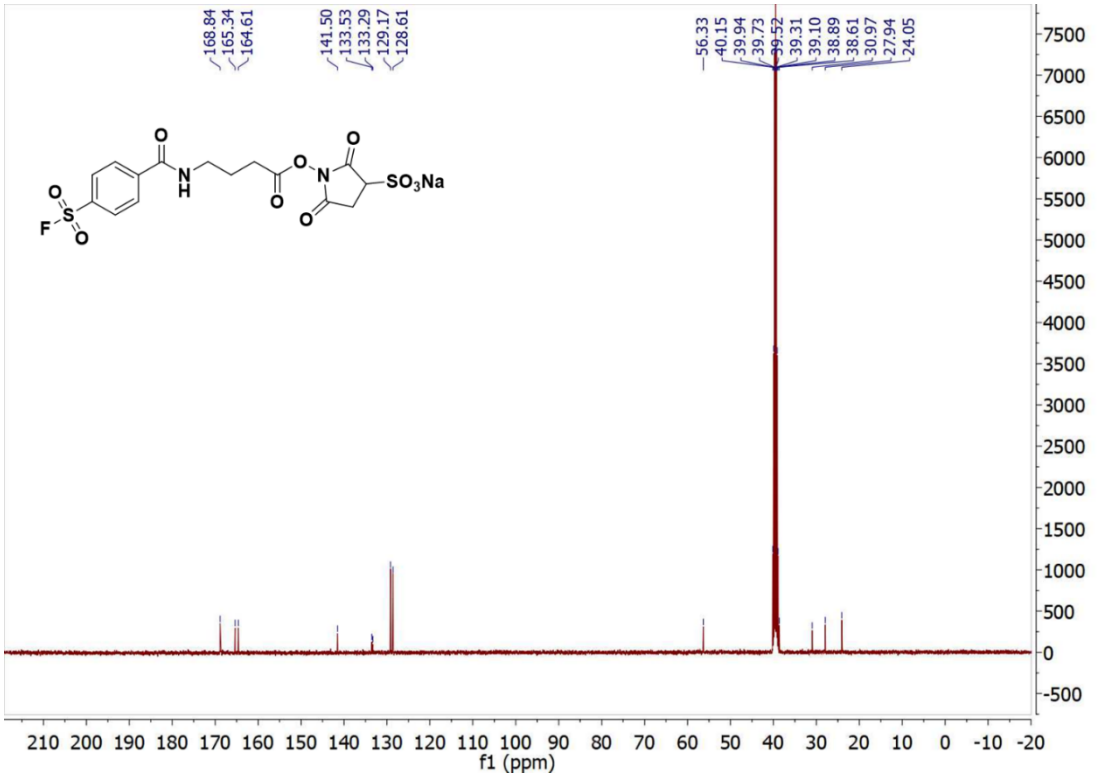
^{1}H NMR (DMSO-d_6) of NHSF-C2.



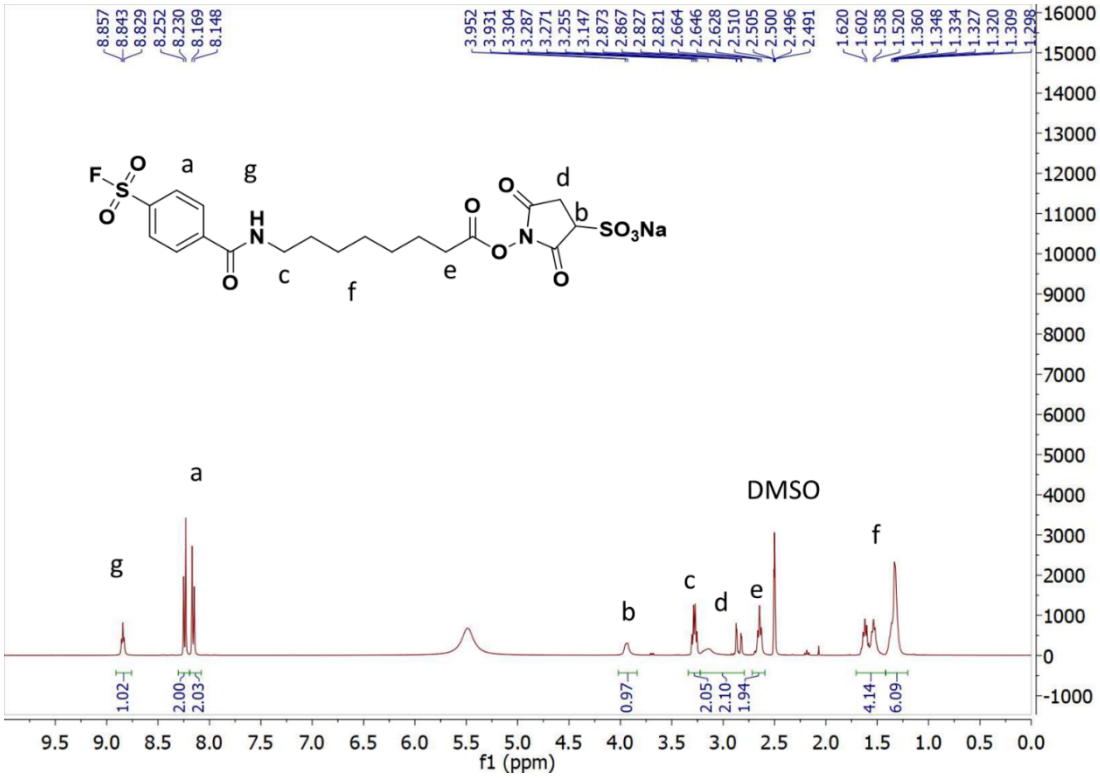
^{13}C NMR (DMSO-d_6) of NHSF-C2.



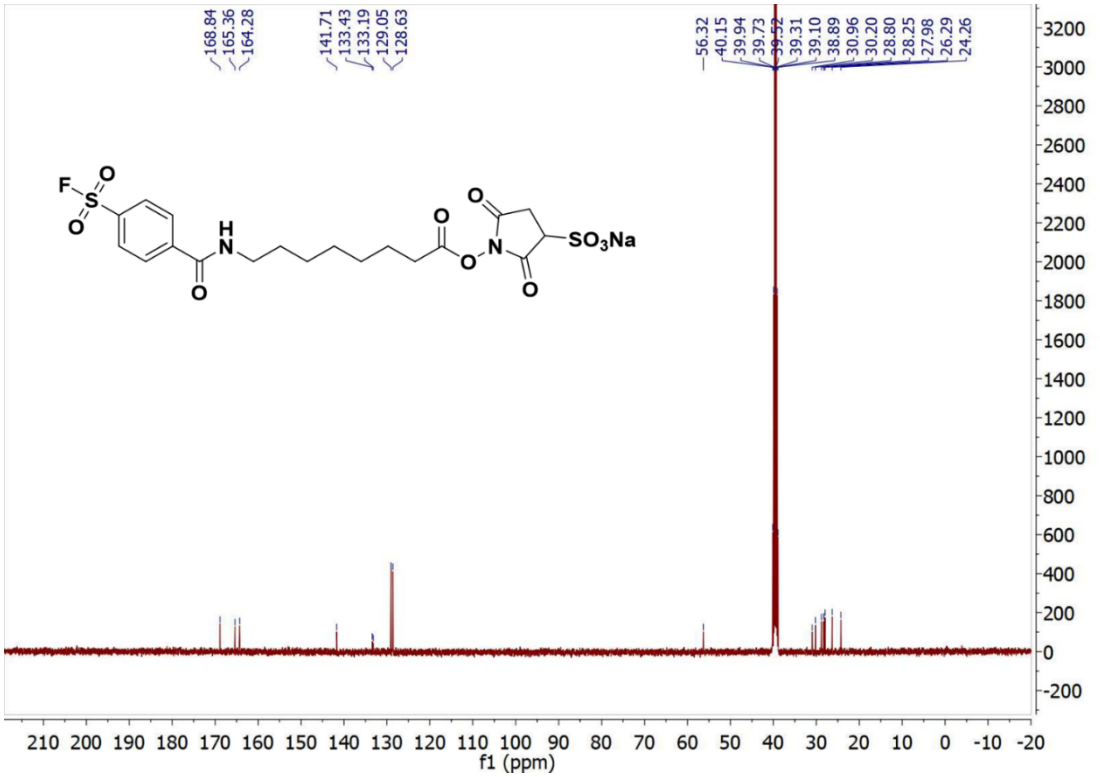
^{1}H NMR (DMSO-d_6) of NHSF-C3.



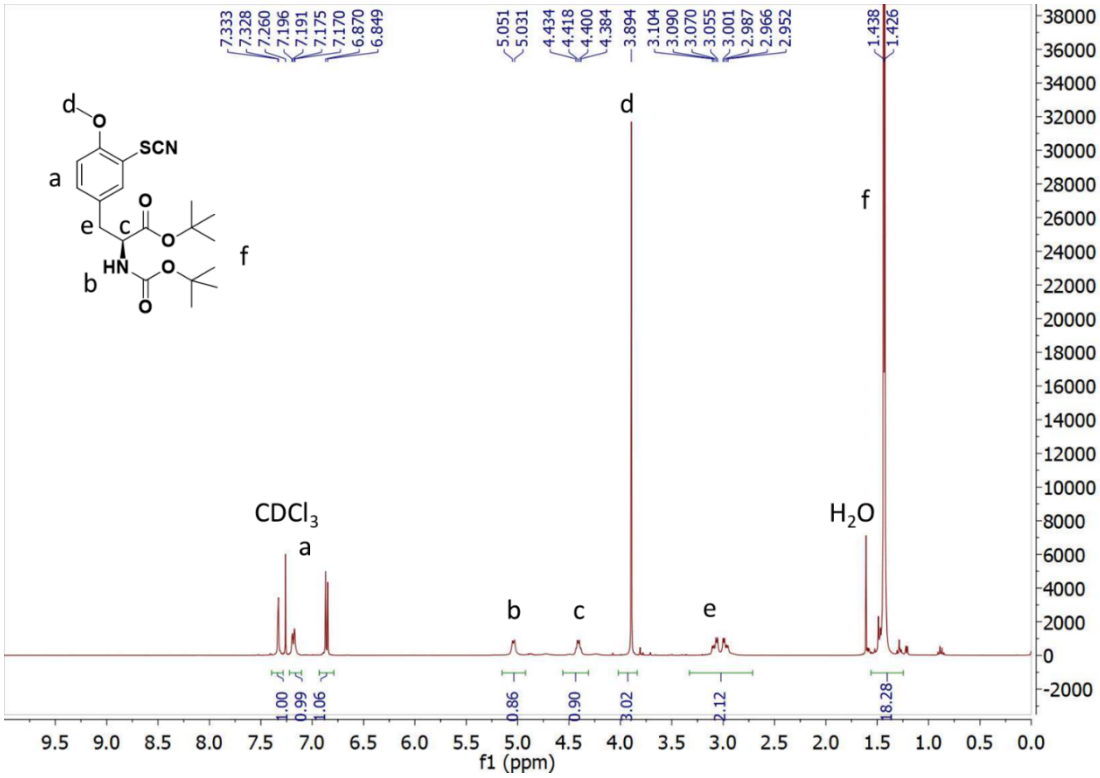
^{13}C NMR (DMSO-d_6) of NHSF-C3.



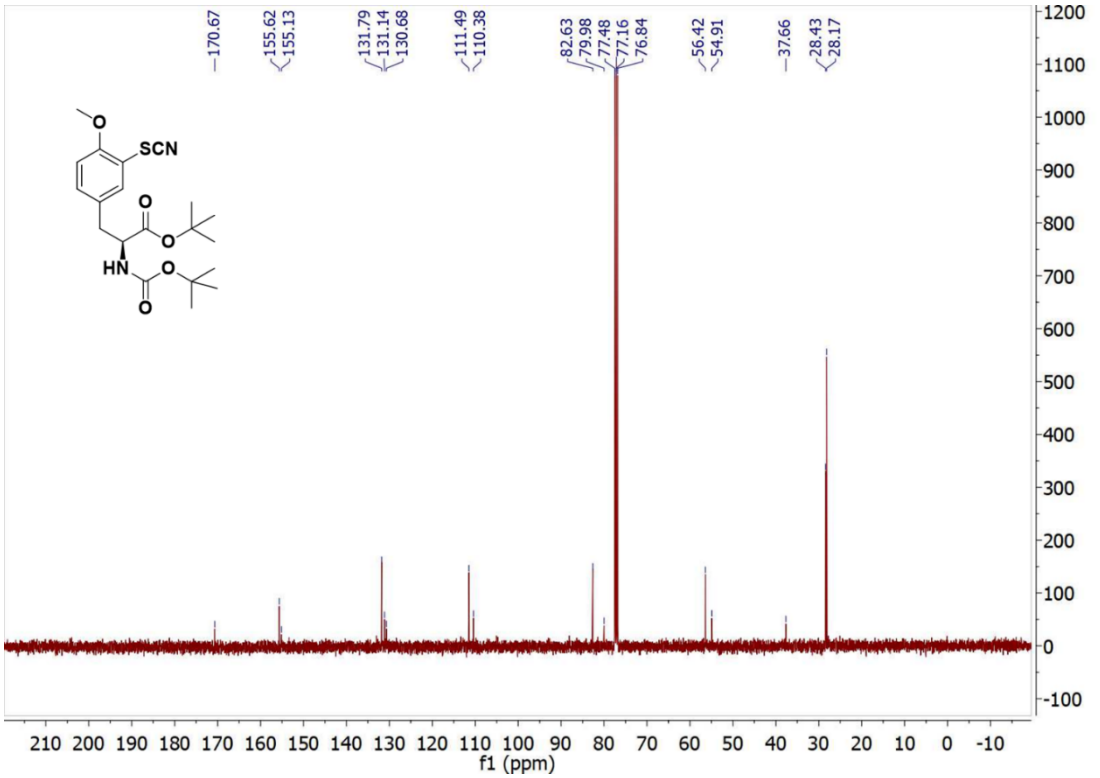
^{1}H NMR (DMSO-d_6) of NHSF-C7.



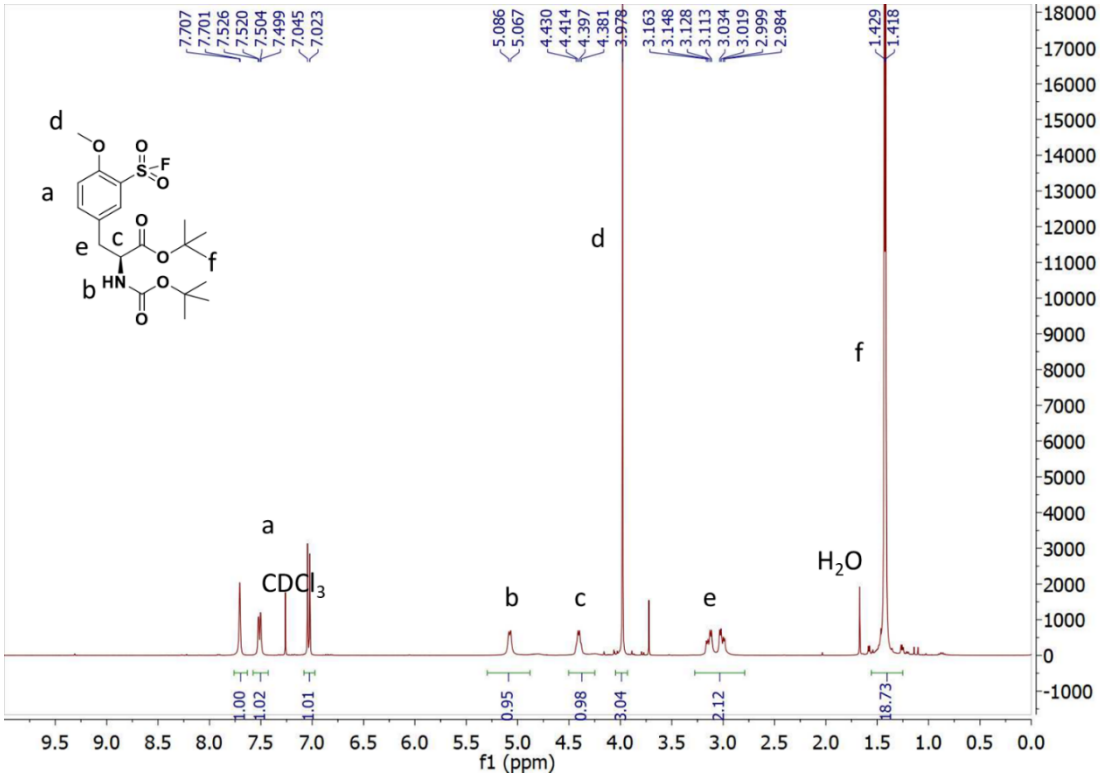
^{13}C NMR (DMSO-d_6) of NHSF-C7.



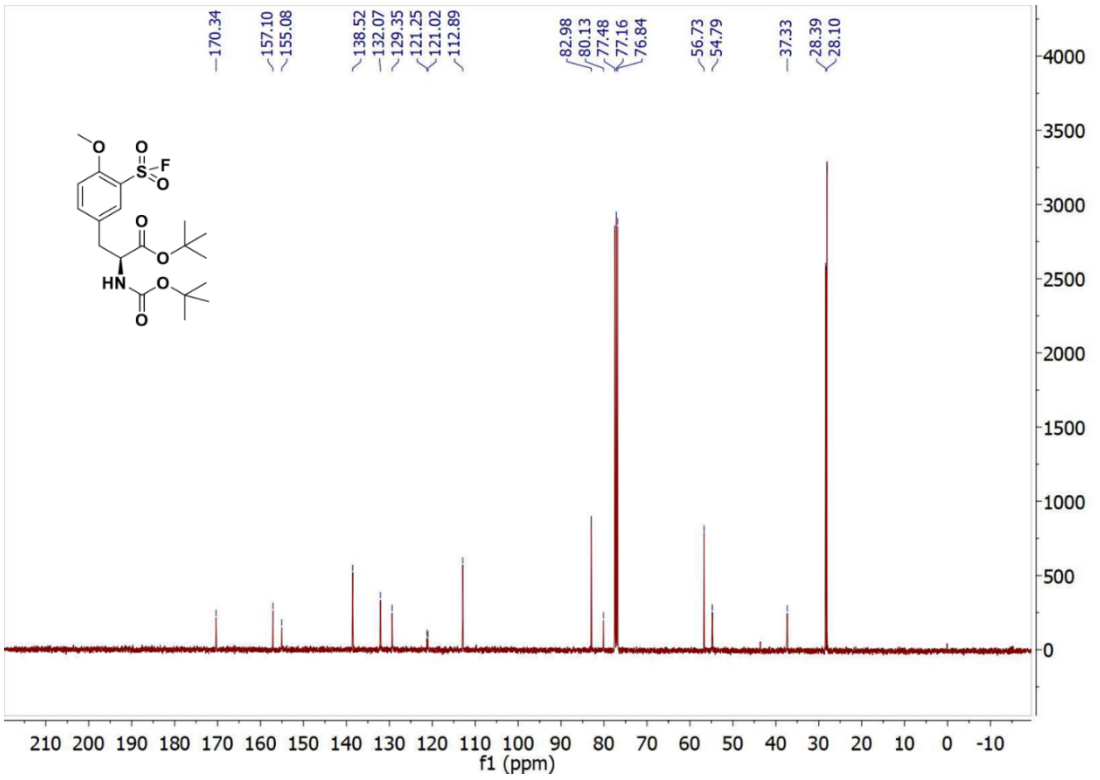
^{1}H NMR ($CDCl_3$) of compound 10.



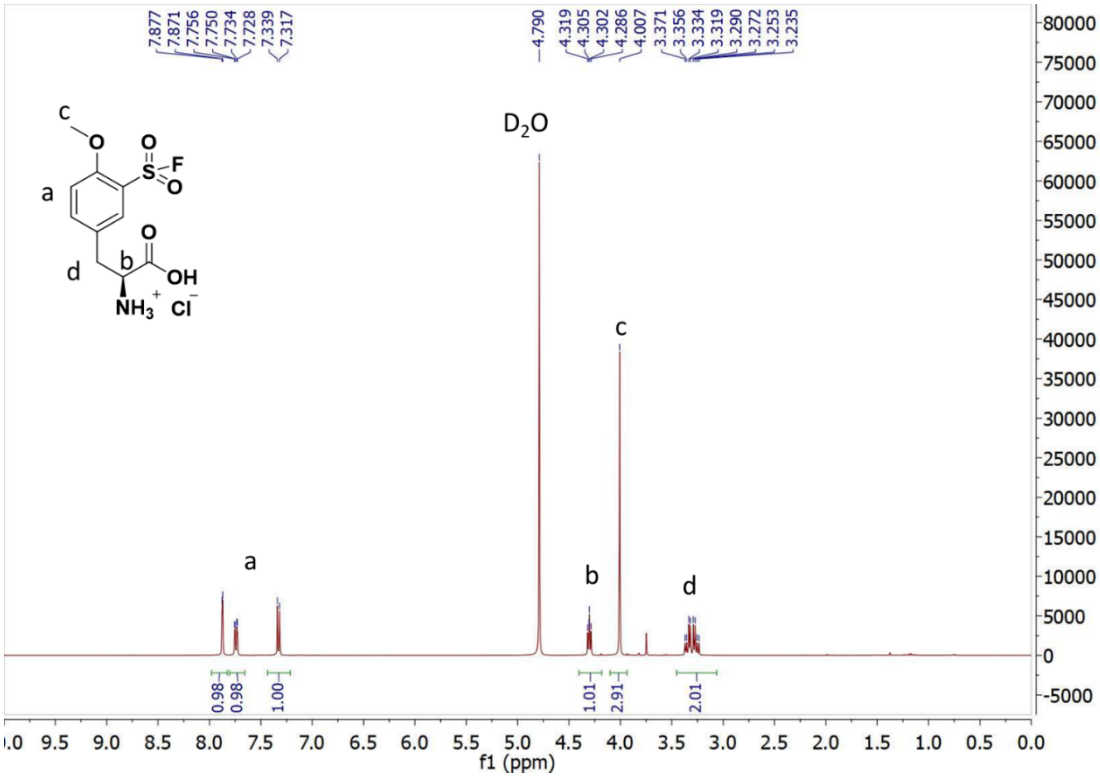
^{13}C NMR ($CDCl_3$) of compound 10.



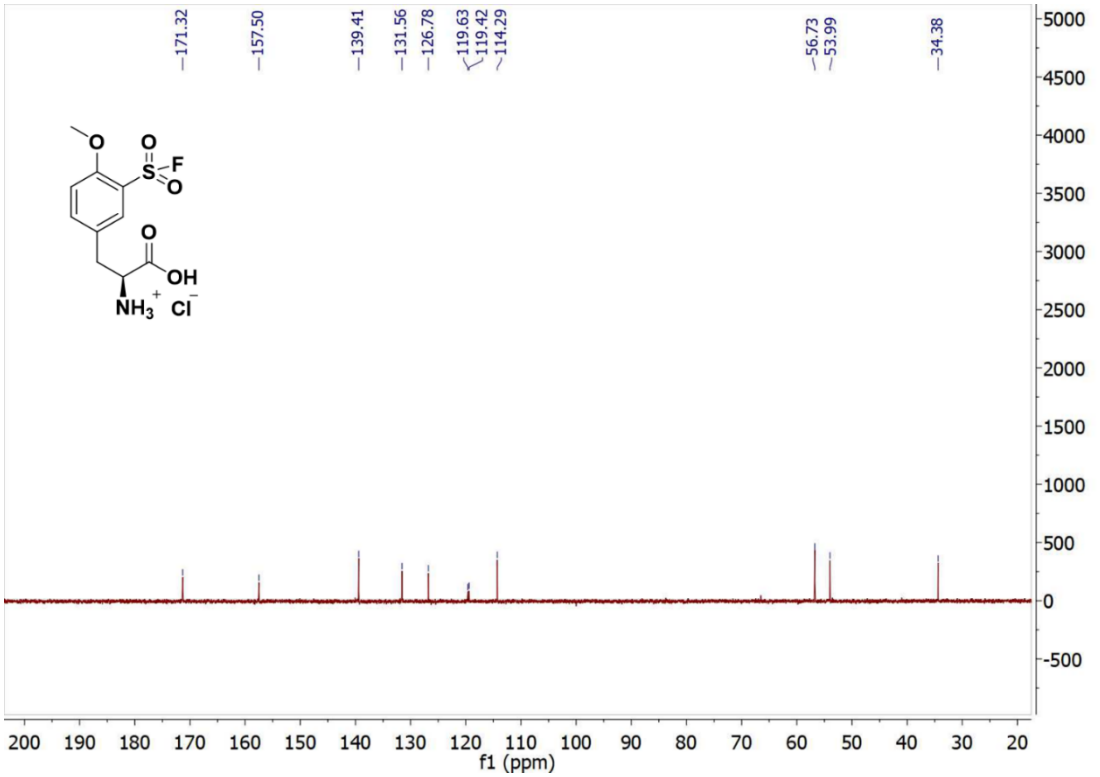
^{1}H NMR ($CDCl_3$) of compound 12.



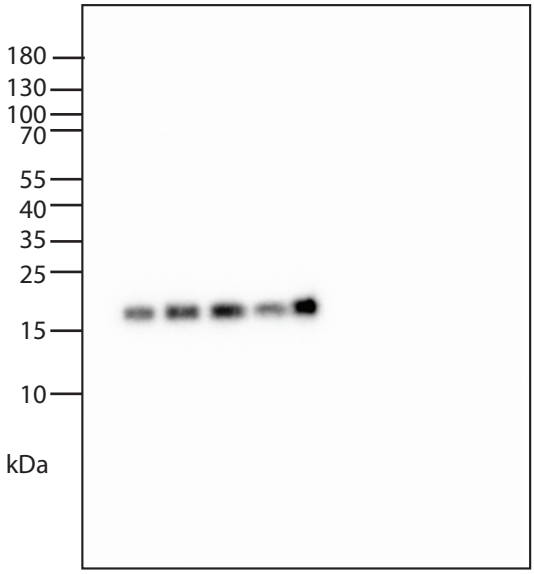
^{13}C NMR ($CDCl_3$) of compound 12.



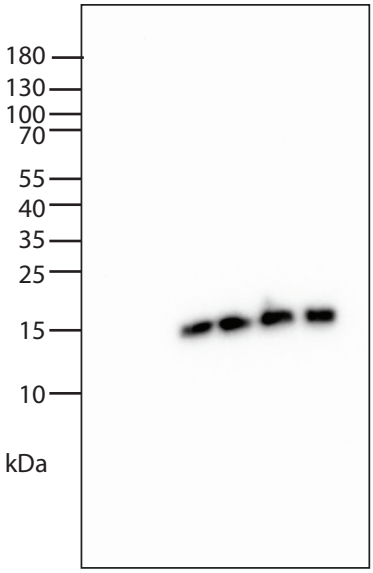
^{1}H NMR (D_2O) of compound SFY.



^{13}C NMR (D_2O) of compound SFY.



Supplementary Data File 1. Uncropped blot scan of Supplementary Figure 7b bottom panel.



Supplementary Data File 2. Uncropped blot scan of Supplementary Figure 9a bottom panel.