Supporting Information

Sulfonyl Azide-Mediated Norbornene Aziridination for Orthogonal Peptide and Protein Labeling

Michael J. Gattner, Michael Ehrlich and Milan Vrabel*

Department of Chemistry, Ludwig-Maximilians-Universität München, Butenandtstrasse 5–13, 81377 Munich (Germany) and Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 16610, Prague (Czech Republic)

*E-mail: Vrabel@uochb.cas.cz

1. Content

1. Content 2
2. General methods 3
3. Synthesis of the compounds 4
   3.1. 4-aminobenzenesulfonyl azide 4
   3.2. 4-((4-(azidosulfonyl)phenyl)amino)-4-oxobutanoic acid (1) 4
   3.3. Dansyl benzenesulfonyl azide (6) 5
   3.4. Biotin benzenesulfonyl azide (5) 5
   3.5. Model reaction 6
4. Kinetic measurements 9
5. Peptide labeling experiments 11
6. Stability studies 14
7. Investigation of potential side reactions with endogenous amino acids 15
8. Cloning, expression and purification of norbornene–containing proteins 17
   8.1. Cloning and mutagenesis of IBA3_Trx N65amber 17
   8.2. Expression of norbornene–containing Trx 17
   8.3. Tryptic digestion and MS/MS of norbornene–containing Trx 19
   8.4. Mutagenesis of pACA_HCA H36amber 20
   8.5. Expression of norbornene containing HCA 20
   8.6. Tryptic digestion and MS/MS of norbornene–containing HCA 21
9. Protein labeling studies 22
10. NMR spectra 23
11. Literature 28

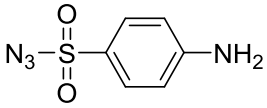
2. General methods

Chemicals were purchased from *Sigma-Aldrich*, *Fluka*, *ABCR* or *Acros Organics* and were used without further purification. Solvents used were of reagent grade. Reactions were monitored on *Merck* Silica 60 F254 TLC plates. Detection was done by irradiation with UV light (254 nm or 366 nm) and staining with *p*-anisaldehyde solution in ethanol, CAM or $KMnO_4$ staining solutions. Flash column chromatography was performed on Silica 60 (*Merck*, 230-400 mesh). NMR spectra were recorded on the following spectrometers: *Varian Oxford 200, Bruker AC 300, Varian XL 400* and *Bruker AMX 600*. The chemical shifts δ are in ppm and coupling constants in Hz. The spectrometers were calibrated using residual undeuterated solvents as internal reference. The structures were solved using 2D NMR techniques. Multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, qi = quintet, m = multiplet, br. = broad. Mass spectra were recorded on the following machines: *Finnigan MAT 95* (EI), *Bruker Autoflex II* (MALDI-TOF), *Thermo Finnigan LTQ-FT* (ESI-ICR) and *Thermo LTQ-Orbitrap XL*. For analytical HPLC separations of protein and peptide samples with subsequent MS a *Dionex Ultimate 3000 Nano* HPLC was used. Proteins were purified on an *ÄKTA purifier* system. The in-gel fluorescence imaging was performed on *LAS-3000* imaging system (*Raytest*). The kinetic measurements were performed on *Waters* HPLC Alliance equipped with Nucleosil C18 analytical column using a gradient of 100% A to 80% B in 20 min (A = 0.1 M triethylammonium acetate in water, B = 0.1 M triethylammonium acetate in 80% CH_3CN). Preparative HPLC was performed on *Waters 1525* with a *Waters 2487 Dual lambda absorbance detector* using *Nucleosil* columns (250 * 4 mm, C18ec, particle size 3 µm or 250 * 10 mm, C18ec, 5 µm) from *Machery-Nagel*. The purified final products were concentrated using a *Christ* alpha 2-4 LD plus lyophyllizer. Acetonitrile for HPLC-purification was purchased from *VWR*. Acetonitrile of LC-MS grade was purchased from *Carl Roth GmbH + Co. KG*. Water was purified by a Milli-Q Plus system from *Merck Millipore*.

3. Synthesis of the compounds

The *endo*-norbornene amino acid **4** and the norbornene–containing peptide AcHN-Ala-Phe-Asp-Lys(Norb)-Lys-Asp-Lys-Pro-Ala-Ala-$CONH_2$ were prepared as previously described.[1]

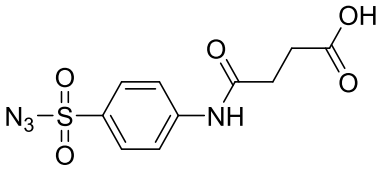
3.1. 4-aminobenzenesulfonyl azide



Commercially available 4-acetamidobenzenesulfonyl azide was deprotected according to literature procedure.[2]

Briefly: 4-acetamidobenzenesulfonyl azide (1.2 g, 5 mmol) was dissolved in 4.5 mL of conc. HCl (37%) and heated to 90 °C for 35 min. The solution was cooled down on ice and brought to pH~6 by slow addition of saturated solution of $NaHCO_3$ and solid $NaHCO_3$. The crude product was extracted with ether, the organic phase washed with brine and dried over Na_2SO_4. The solvent was removed under vacuum and the product was used without further purification in the next step.

3.2. 4-((4-(azidosulfonyl)phenyl)amino)-4-oxobutanoic acid (1)



4-Aminobenzenesulfonyl azide (355 mg, 1.8 mmol) and succinic anhydride (896 mg, 9 mmol, 5 equiv.) were dissolved in dry THF (10 mL) and heated to 70 °C until TLC (DCM/MeOH/AcOH = 9.5/5/0.01) showed complete conversion (ca. 4 h). The solvent was removed under vacuum and the residue was purified by silica gel column chromatography using DCM/MeOH/AcOH = 9.5/5/0.01 as eluent. The appropriate fractions were collected and the residue was isolated as a colorless solid (444 mg, 83%).

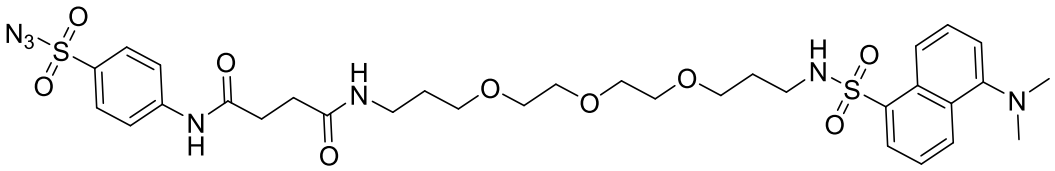
^{1}H-NMR (400 MHz, CD_3OD): δ = 7.90 (m, 4H), 2.70 (m, 4H).

^{13}C NMR (101 MHz, CD_3OD) δ = 176.3, 174.8, 146.5, 133.3, 130.0, 120.6, 32.5, 29.7.

HRMS (ESI^-): calcd. for $[C_{10}H_9N_4O_5S]^-$: 297.0294, found: 297.0299 $[M\text{-}H]^-$.

IR $\tilde{\nu}$ / cm^{-1}: 3378 (*m*), 3200 (*w, br*), 2360 (*w*), 2140 (*m*), 1765 (*m*), 1689 (*m*), 1560 (*m*), 1528 (*s*), 1395 (*m*), 1360 (*s*), 1254 (*m*), 1152 (*s*), 1086 (*w*), 971 (*w*), 828 (*m*), 843 (*m*), 763 (*s*), 668 (*m*).

3.3.Dansyl benzenesulfonyl azide (6)



4-((4-(azidosulfonyl)phenyl)amino)-4-oxobutanoic acid (50 mg, 0.17 mmol) and HATU (70 mg, 0.185 mmol, 1.1 equiv.) were dissolved in dry DMF (2 mL) and cooled down on ice. Solution of dansylamine[3] (76 mg, 0.17 mmol, 1 equiv) in Et_3N (70 µL, 3 equiv.) and DMF (0.5 mL) was added dropwise under Ar to this solution. The reaction mixture was stirred at room temperature overnight in the dark (TLC in DCM/AcOEt/MeOH = 7/3/0.25). The solvent was removed under vacuum and the residue was purified by silica gel column chromatography using a mixture of DCM/AcOEt/MeOH = 35/15/1.5 as eluent. The product was isolated as yellowish solid (32 mg, 26%).

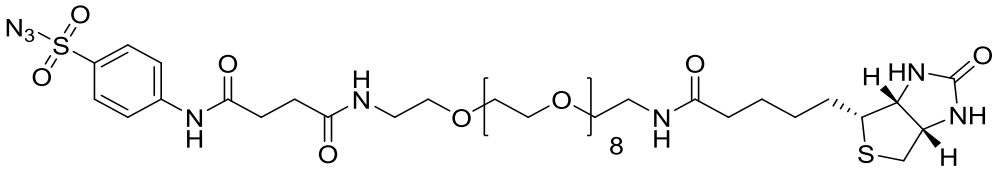
^{1}H-NMR (400 MHz, CD_3OD): δ = 8.55 (d, J = 8.0 Hz, 1H), 8.34 (d, J = 8.0 Hz, 1H), 8.19 (d, J = 8.0 Hz, 1H), 7.88 (m, 4H), 7.58 (dd, J = 7.5, 7.2 Hz , 2H), 7.27 (d, J = 8.0 Hz, 1H), 3.54 (m, 4H), 3.48 (m, 4H), 3.33 (m, 4H), 3.24 (t, J = 6.7 Hz, 2H), 2.93 (t, J = 6.7 Hz, 2H), 2.87 (s, 6H), 2.72 (t, J = 6.7 Hz, 2H), 2.55 (t, J = 6.7 Hz, 2H), 1.73 (m, 2H), 1.58 (m, 2H).

^{13}C NMR (101 MHz, CD_3OD) δ = 174.4, 173.5, 153.2, 146.4, 136.9, 133.2, 131.2, 131.1, 131.0, 130.3, 129.9, 129.1, 124.3, 120.6, 120.5, 116.4, 71.5, 71.4, 71.1, 71.0, 69.8, 69.3, 45.8, 41.3, 37.9, 33.0, 31.5, 30.5, 30.3.

HRMS (ESI$^-$): calcd.for $[C_{32}H_{42}N_7O_9S_2]^-$: 732.2485, found: 732.2486 $[M\text{-}H]^-$.

IR $\tilde{\nu}$ / cm^{-1}: 3280 (*w*, *br*), 2952 (*w*), 2360 (*w*), 1697 (*m*), 1673 (*m*), 1588 (*m*), 1530 (*m*), 1396 (*m*), 1305 (*s*), 1149 (*s*), 1090 (*m*), 974 (*w*), 908 (*m*), 864 (*m*), 836 (*s*), 775 (*m*), 707 (*m*).

3.4.Biotin benzenesulfonyl azide (5)



4-((4-(azidosulfonyl)phenyl)amino)-4-oxobutanoic acid (21 mg, 0.07 mmol, 1.2 eq.), *O*-(2-aminoethyl)-*O′*-[2-(biotinylamino)ethyl]octaethylene glycol (95%, 40 mg, 0.06 mmol, 1.0 eq.) and HATU (26 mg, 0.07 mmol, 1.2 eq.) were dissolved in DMF (2 mL). To the reaction mixture *N*,*N*-diisopropylethylamine (25 µL, 0.15 mmol, 2.5 eq.) was added and the reaction mixture was stirred for 5 h at ambient temperature. Subsequently the reaction mixture was concentrated under reduced pressure and the biotin reagent **5** was purified by preparative HPLC (buffer A: ddH_2O, buffer B: MeCN, gradient 100 % A, 0 % B → 0 % A, 100 % B in 45 min, retention time: 22 min *Nukleosil 100 7 –C18* column) and obtained as colorless solid (46 mg, 0.05 mmol, 83%).

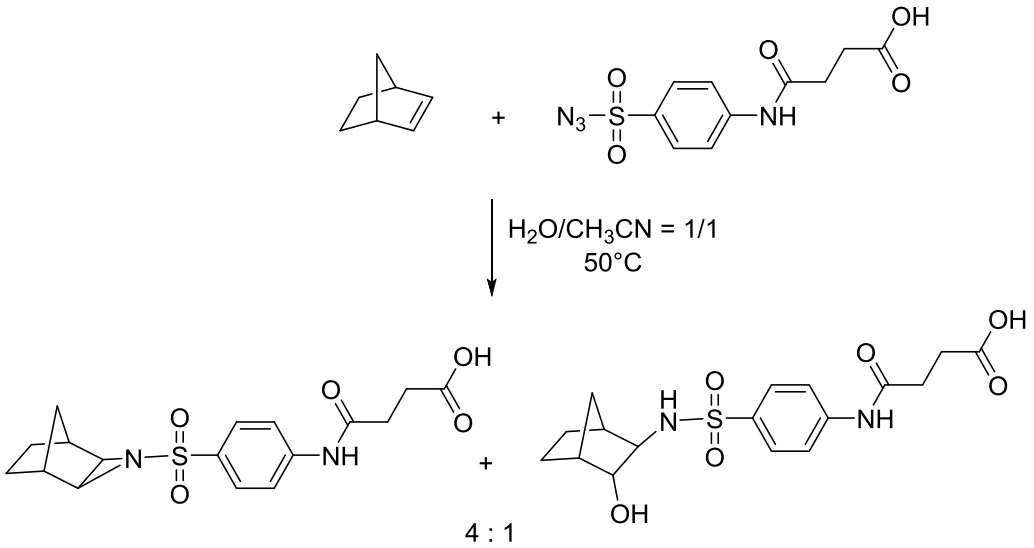
^{1}H-NMR (400 MHz, CD_3OD): δ = 7.91 (s, 4H), 4.49 (dd, J = 8.0, 5.0 Hz, 1H), 4.30 (dd, J = 7.9, 4.4 Hz, 1H), 3.67 – 3.58 (m, 32H), 3.54 (td, J = 5.5, 2.2 Hz, 4H), 3.36 (q, J = 5.4 Hz, 4H), 3.25 – 3.17 (m, 1H), 2.92 (dd, J = 12.7, 5.0 Hz, 1H), 2.78 – 2.66 (m, 3H), 2.60 (t, J = 6.7 Hz, 2H), 2.22 (t, J =7.4 Hz, 2H), 1.80 – 1.51 (m, 4H), 1.51-1.38 (m, 2H).

^{13}C NMR (101 MHz, CD_3OD) δ = 176.1, 174.6, 173.4, 166.1, 146.5, 133.2, 130.0, 120.6, 71.6 (16C), 71.6, 71.3, 70.57, 70.54, 63.4, 61.6, 57.0, 41.1, 40.5, 40.4, 36.7, 33.1, 31.4, 29.8, 29.5, 26.9.

HRMS (ESI^+): calcd. for $[C_{40}H_{67}N_8O_{15}S_2]^+$: 963.4162, found: 963.4149 $[M+H]^+$.

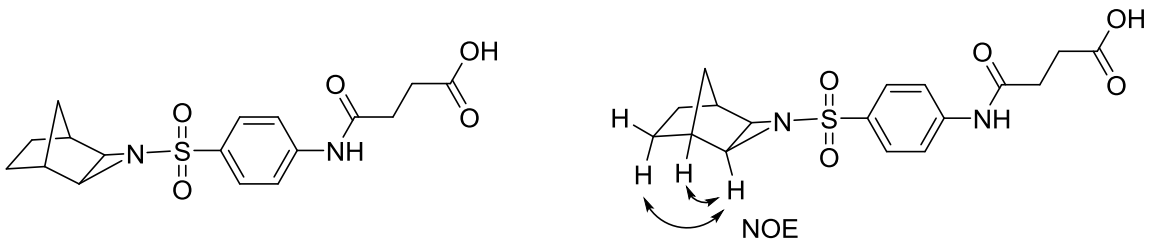
IR $\tilde{\nu}$ / cm^{-1}: 3294 (*w*, *br*), 2867 (*w*), 2125 (*w*), 1692 (*m*), 1587 (*m*), 1531 (*s*), 1442 (*m*), 1366 (*m*), 1306 (*m*), 1261 (*m*), 1148 (*s*), 1088 (*s*), 1035 (*m*), 974 (*w*), 908 (*m*), 864 (*m*), 844 (*m*), 754 (*m*), 703 (*m*).

3.5. Model reaction



Scheme S1: Scheme of the model reaction of norbornene with sulfonyl azide **1**.

4-((4-(azidosulfonyl)phenyl)amino)-4-oxobutanoic acid **1** (30 mg, 0.1 mmol) was stirred with norbornene (19 mg, 0.2 mmol, 2 equiv.) at 45-50 °C in a mixture of H_2O/CH_3CN = 1:1 (4 mL) overnight (Scheme S1). The crude reaction mixture was diluted with Mili Q water, filtered and purified by semipreparative HPLC using a gradient of 100% A to 50% B in 45 min (A = 0.1 M triethylammonium acetate in water, B = 0.1 M triethylammonium acetate in 80% CH_3CN) The products were isolated as Et_3N salt after lyophilization (aziridine **2**: 28 mg, 78%, sulfonamide **3**: 7 mg, 18%).

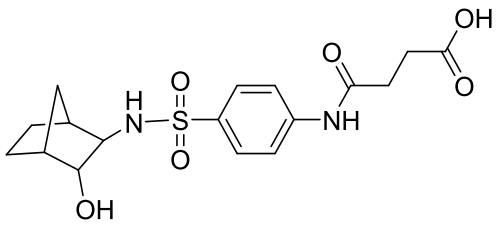


^{1}H-NMR (400 MHz, D_2O): δ = 7.87 (d, *J* = 9.1 Hz, 2H), 7.72 (d, *J* = 9.1 Hz, 2H), 3.03 (s, 2H), 2.67 (td, *J* = 7.4, 0.8 Hz, 2H), 2.52 (td, *J* = 7.3, 0.8 Hz, 2H), 2.45 (s, 2H), 1.46 (ddd, *J* = 8.9, 3.2, 1.5 Hz, 2H), 1.28 (m, 1H), 1.18 (dd, *J* = 7.7, 2.3 Hz, 1H), 0.82 (d, *J* = 10.3 Hz, 1H).

^{13}C NMR (101 MHz, D_2O) δ = 181.0, 174.8, 142.9, 130.5, 128.9, 120.6, 43.4, 35.4, 33.0, 32.3, 27.4, 24.6.

HRMS (ESI$^+$): calcd. for $[C_{17}H_{21}N_2O_5S]^+$: 365.1171, found: 365.1166 $[M+H]^+$.

IR $\tilde{\nu}$ / cm^{-1}: 3240 (*w*, *br*), 2953 (*w*), 2868 (*w*), 2359 (*m*), 1696 (*m*), 1672 (*m*), 1588 (*s*), 1530 (*s*), 1397 (*m*), 1305 (*s*), 1225 (*m*), 1150 (*s*), 1090 (*m*), 1035 (*m*), 974 (*m*), 908 (*s*), 865 (*s*), 845 (*m*), 836 (*m*), 773 (*m*), 707 (*m*), 668 (*m*).



^{1}H-NMR (400 MHz, D_2O): δ = 7.90 (d, *J* = 8.5 Hz, 2H), 7.74 (d, *J* = 8.5 Hz, 2H), 3.86 (d, *J* = 6.3 Hz, 1H), 3.22 (m, 1H), 2.74 ((t, *J* = 7.1 Hz, 2H), 2.62 (t, *J* = 7.1 Hz, 2H), 2.12 (bs, 1H), 2.02 (d, *J* = 4.0 Hz, 1H), 1.84 (dd, *J* = 13.9, 7.4 Hz,, 1H), 1.62 (m, 1H), 1.44 (m, 2H), 1.00 (m, 2H).

^{13}C NMR (101 MHz, D_2O) δ = 180.0, 174.5, 141.9, 133.4, 128.1, 120.9, 74.7, 61.3, 45.3, 39.1, 37.8, 32.5, 31.4, 25.0, 21.9

HRMS (ESI$^-$): calcd. for $[C_{17}H_{21}N_2O_6S]^-$: 381.1120, found: 381.1125$[M-H]^-$.

IR $\tilde{\nu}$ / cm^{-1}: 3248 (*w*, *br*), 2958 (*w*), 2360 (*w*), 1691 (*m*), 1590 (*s*), 1531 (*s*), 1400 (*m*), 1310 (*s*), 1255 (*m*), 1145 (*s*), 1092 (*m*), 976 (*m*), 910 (*m*), 837 (*s*), 800 (*m*), 697 (*m*).

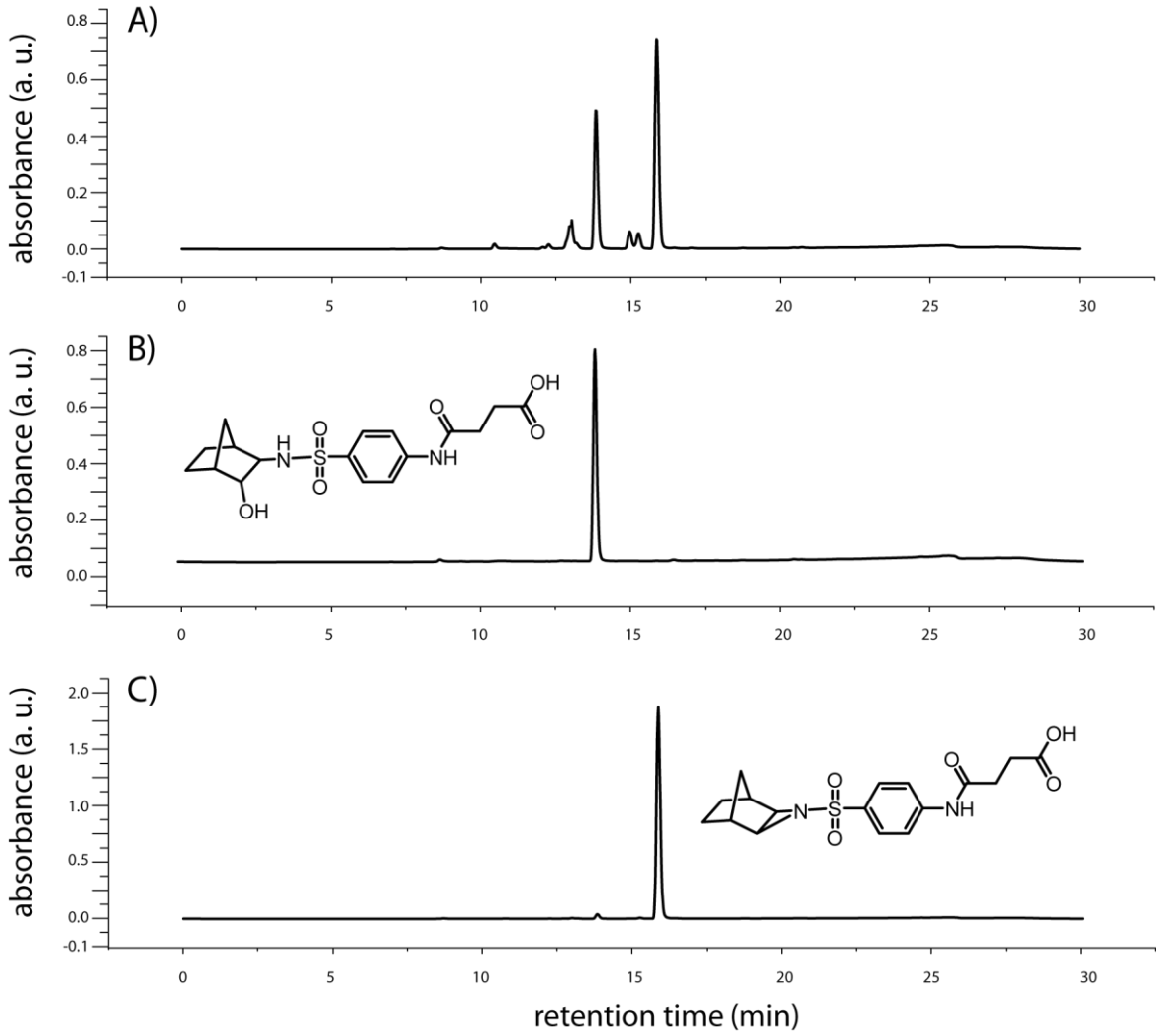


Figure S1: HPLC chromatograms (at 254 nm) of A) crude reaction mixture, B) hydrolysis product after purification, C) *exo*-aziridine product after purification. Retention times (*Nucleosil* C18 analytical column Machery-Nagel, gradient 100% A to 80% B in 20 min, where A = 0.1 M triethylammonium acetate in water, B = 0.1 M triethylammonium acetate in 80% CH_3CN): retention times: Sulfonyl azide **1**: 14.9 min, *exo*-aziridine **2**: 15.7 min and sulfonamide **3**: 13.7 min.

4. Kinetic measurements

The *endo*-5-norbornene-2-methanol was used for the kinetic measurements to ensure sufficient water solubility of the norbornene reagent. The kinetic measurements were performed in H_2O/CH_3CN (9:1) mixture at room temperature (21 °C) under pseudo first order conditions using an excess of norbornene and were performed in triplicate as follows:

The decay in the absorption of sulfonyl azide **1** at 260 nm in the presence of norbornene was followed in time. 1 mM final concentration of sulfonyl azide **1** and 10, 20, 30, 40 and 50 mM final concentrations of the norbornene corresponding to 10 to 50 equivalents were used. The appropriate solutions (final 10% CH_3CN in water, stock solution of the sulfonyl azide **1**: 10 mM in CH_3CN and norbornene: 100 mM in water) of the reagents were combined in HPLC vials and the data were recorded using a total 30 min gradient (from 100% A to 80% B in 20 min, where A = 0.1 M triethylammonium acetate in water, B = 0.1 M triethylammonium acetate in 80% CH_3CN). The decay in the absorption of the sulfonyl azide peak (15 min) at 260 nm was followed and the integral of the peak area was recorded. The respective concentrations of the sulfonyl azide at a specific time were then calculated from the slope of the plot obtained from the calibration curve: $y = a + bx$ where $a = -4.12 \times 10^5$ and $b = 2.75 \times 10^7$. The data were fit to a single exponential equation ($y = a + be^{-kx}$) to provide the observed rate constants k` for each norbornene concentration. The observed rate constants were plotted against the norbornene concentrations and the second-order rate constant k was obtained from the slope of this plot ($k = 0.0017 \pm 0.00021\ M^{-1}s^{-1}$).

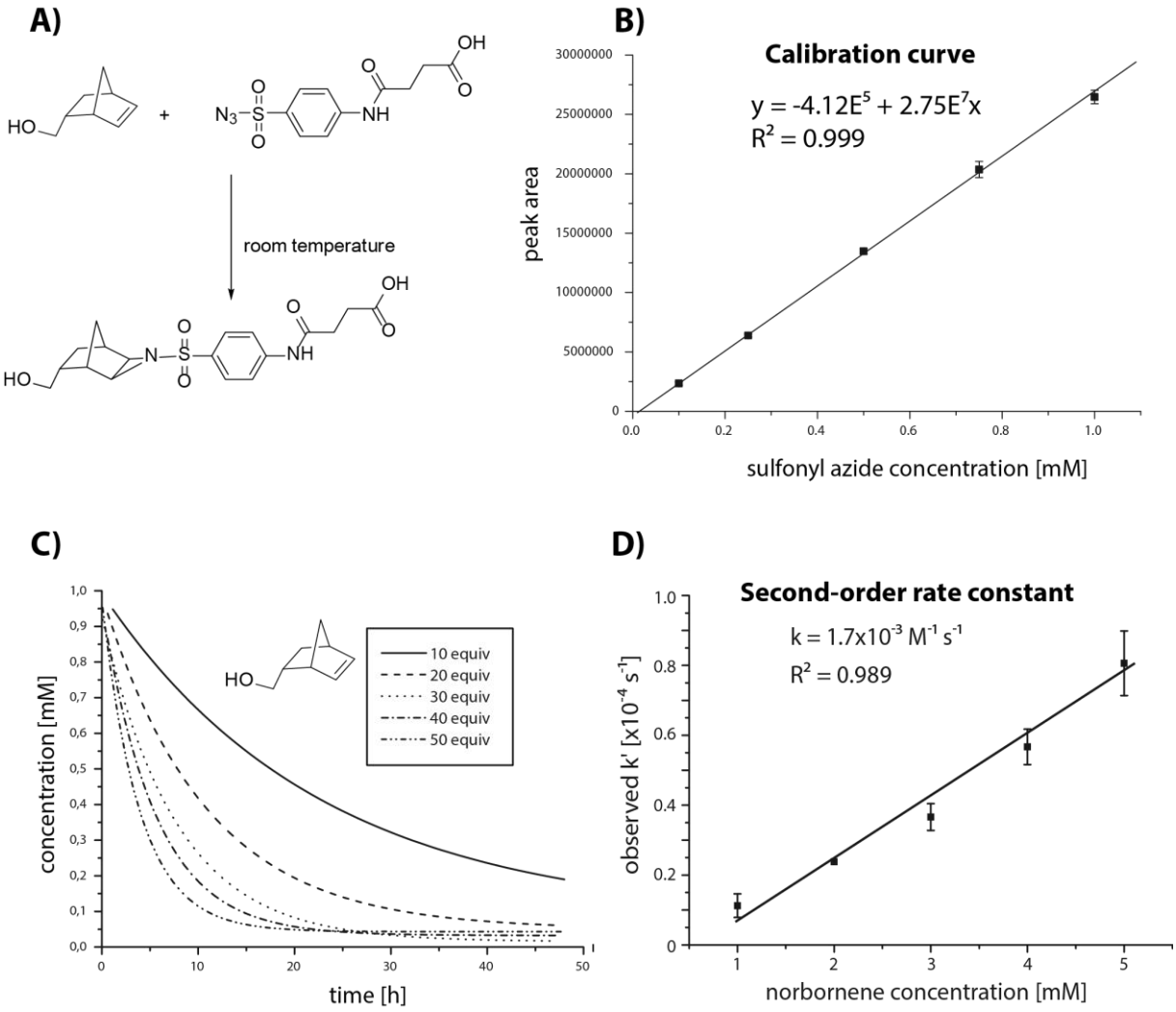


Figure S2: Depiction of the reaction scheme and plots from the kinetic measurements. A) reaction scheme for the measurements of the rate constant, B) plot showing the dependence of the peak area (from HPLC) on the sulfonyl azide concentration. The indicated function ($y = -4.12E^5 + 2.75E^7x$) obtained from the slope of this plot was used for the calculations of sulfonyl azide concentration throughout the HPLC kinetic measurements, C) plot showing the decay of sulfonyl azide concentration in time using 10-50 equiv. of norbornene alcohol, D) The second order rate constant k was calculated from the dependence of the observed rate constants k' on the norbornene concentration.

5. Peptide labeling experiments

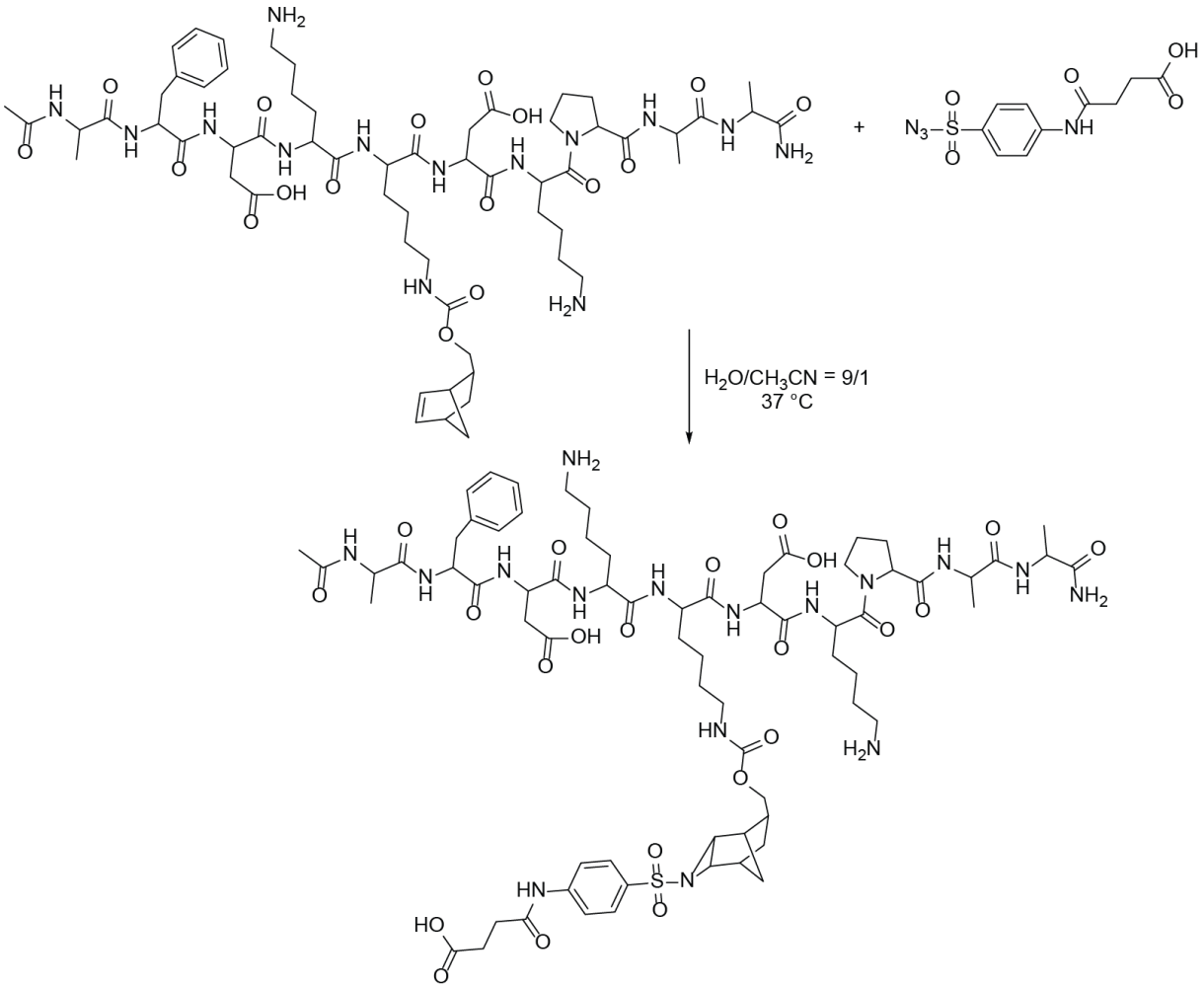
The experiments were performed as follows:

50 μM of the peptide solution was incubated at 37 °C with 2.5 mM 4-((4-(azidosulfonyl)phenyl)amino)-4-oxobutanoic acid **1** in H_2O/CH_3CN (9:1) mixture and the progress of the reaction was monitored using MALDI mass spectrometry. Probes of the reaction mixture were spotted onto MALDI target using sinapic acid solution as a matrix (SA in 50/50 H_2O/CH_3CN + 0.1%TFA) and the spectra were recorded in negative mode.

Calculated masses:

Norbornene–containing peptide: $[M-H]^-$: 1279.7 Da

Labeled peptide: $[M-H]^-$: 1549.7 Da



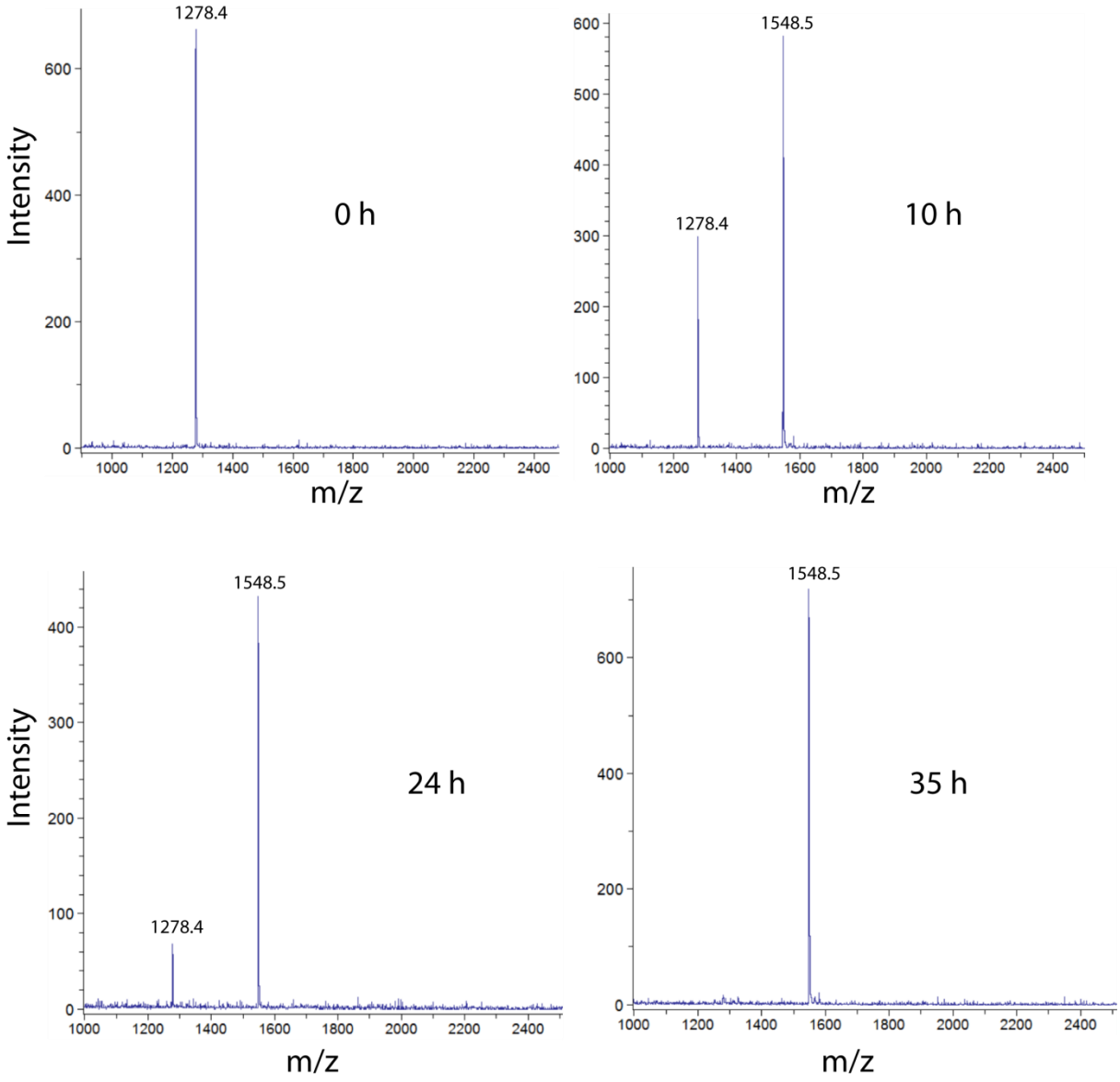


Figure S3: Peptide labeling experiments using 50 µM peptide and 2.5 mM sulfonyl azide **1**. MALDI-TOF spectra showing the progress of the peptide labeling reaction in time. Starting peptide calc. mass: 1279.7 Da and the product mass: 1549.7 Da; observed: 1278.4 Da and 1548.5 Da respectively. $\Delta M_{calc.}$= 270.0 Da; ΔM_{found} = 270.1 Da.

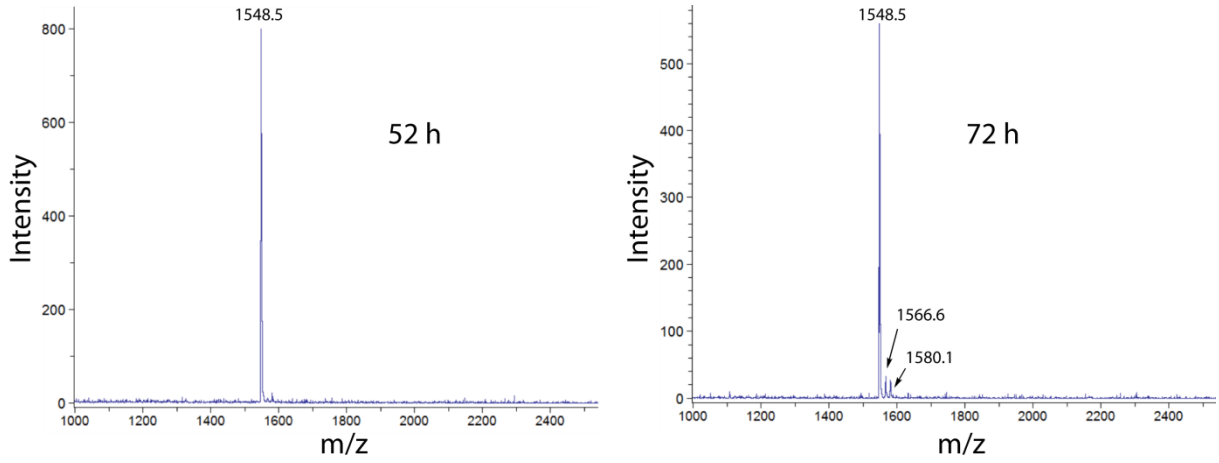


Figure S4: Peptide labeling experiments using 50 µM peptide and 2.5 mM sulfonyl azide **1**. MALDI-TOF spectra showing the formation of the hydrolyzed sulfonamide product after prolonged incubation time (calc. mass $[M\text{-}H]^-$: 1567.7 Da; observed: 1566.6 Da). The observed mass peak at 1580.1 could not be assigned. Desired aziridine product mass: 1549.7 Da; observed: 1548.5 Da.

Additional experiment using 50 µM peptide and 5 mM (100 equiv.) 4-((4-(azidosulfonyl)phenyl)amino)-4-oxobutanoic acid **1** in H_2O/CH_3CN (9:1) mixture showed full conversion of the starting peptide to the aziridinated product within 24 hours (Figure S5).

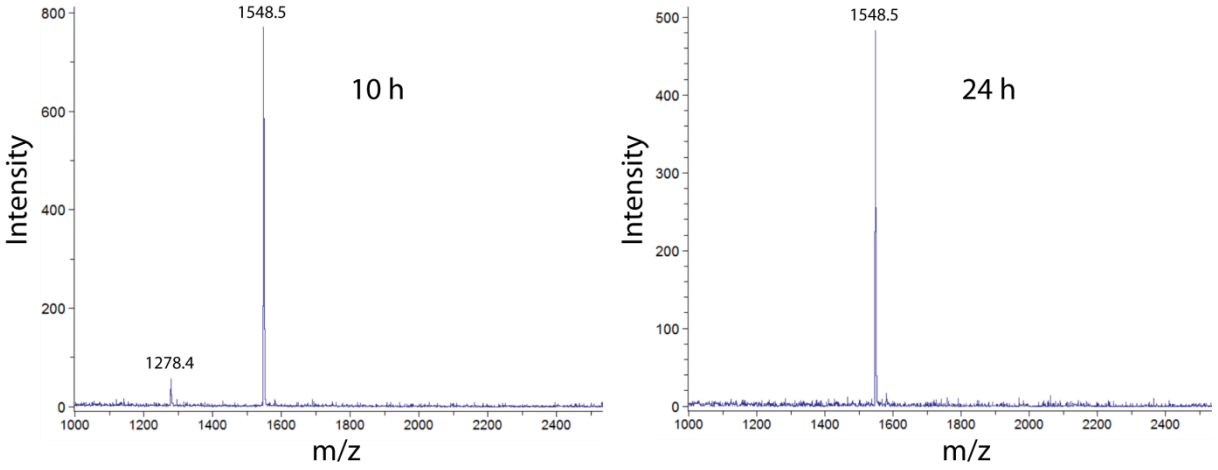
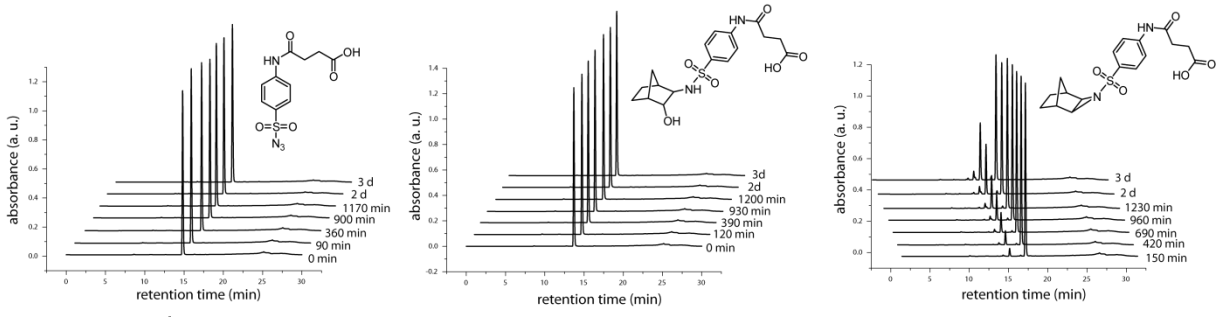


Figure S5: Peptide labeling experiments using 50 µM peptide and 5 mM sulfonyl azide **1**. MALDI-TOF showed full conversion within 24 hours in this case. Starting peptide calc. mass: 1279.7 Da and the product mass: 1549.7 Da; observed: 1278.4 Da and 1548.5 Da respectively. $\Delta M_{calc.}$= 270.0 Da; ΔM_{found} = 270.1 Da.

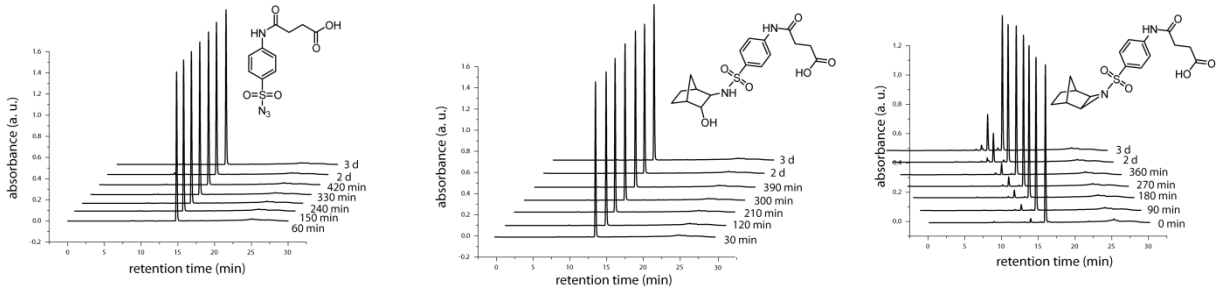
6. Stability studies

To evaluate the stability of sulfonyl azide **1** and the products formed in their reaction with norbornene (**2** and **3**) under various conditions, the corresponding purified compounds (0.5 mM final concentration) were incubated in 50 mM MES buffer pH = 5.5, in 50 mM TRIS buffer pH = 8.5 and in 50 mM cystein solutions at room temperature for 3 days and the solutions were analyzed using analytical HPLC. The results are summarized in Figure S6. (MES = 2-(*N*-morpholino)ethanesulfonic acid, TRIS = 2-Amino-2-hydroxymethyl-propane-1,3-diol)

A) In MES buffer pH = 5.5



B) In TRIS buffer pH = 8.5



C) In 50 mM cysteine solution

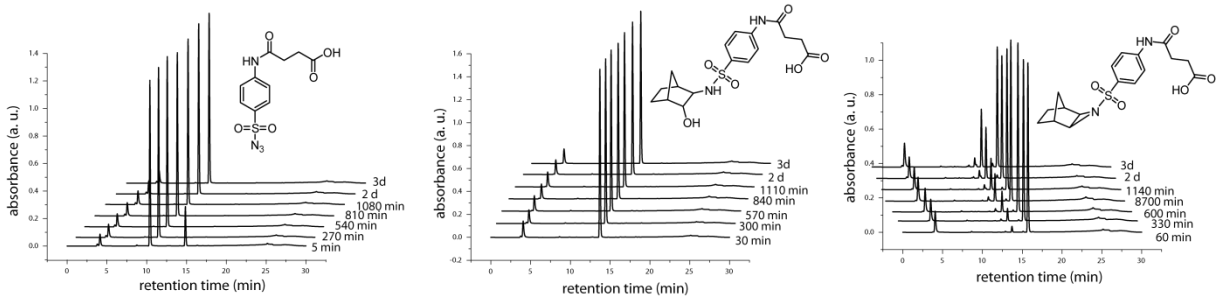


Figure S6: Stability of sulfonyl azide **1** and products formed in its reaction with norbornene (**2** and **3**). A) incubated in 50 mM MES buffer pH = 5.5; only slow hydrolysis of the aziridine **2** to the sulfonamide **3** was observed. B) incubated in 50 mM TRIS buffer pH = 8.5; again only slow hydrolysis of the aziridine **2** to the sulfonamide **3** was observed. C) incubated in 50 mM cystein solution; in this case the reduction of the sulfonyl azide to the corresponding sulfonamide was observed (see also part below and figure S7) and again slow hydrolysis of the aziridine **2** to the sulfonamide **3** was observed.

7. Investigation of potential side reactions with endogenous amino acids

To investigate potential side reactions of sulfonyl azides with amino acid side chains on peptides and proteins a series of experiments were performed.

Azides can be reduced by thiols to form the corresponding amines.[4,5] This reaction however, only leads to partial decrease in effective concentration of the sulfonyl azide reagent.

Nucleophilic substitution reaction on cysteine could in principle lead to the formation of sulfenamides that can be reversibly cleaved using reducing agents such as TCEP or DTT.[6-8]

Our experiments show that sulfonyl azides are indeed slowly reduced to the corresponding sulfonamides in the presence of an excess cysteine. However, no other side reactions regarding nucleophilic substitution (thiol group of the cysteine) were observed during the experiments (see Figure S7 and also stability studies above).

Reduction of sulfonyl azide with cysteine

Sulfonyl azide **1** (0.5 mM) was incubated with cysteine (2.5 mM, 5 equiv.) in 50 mM MES buffer pH=5.5 at room temperature. The progress of the reaction was followed by analytical HPLC and the results are shown in Figure S7.

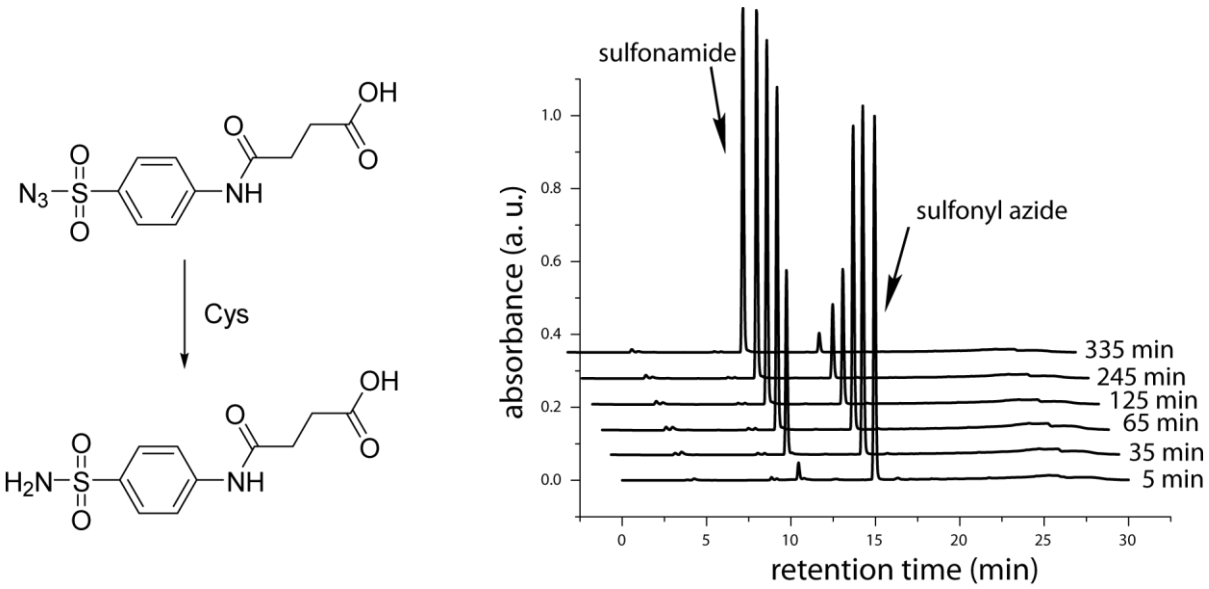


Figure S7: Reduction of sulfonyl azide **1** in the presence of cysteine. The new peak (at 10.3 min) was identified as the corresponding sulfonamide by LC-MS: Calc. mass for $C_{10}H_{11}N_2O_5S$ $[M\text{-}H]^-$: 271.0389, found: 271.0393.

Potential reaction with lysines

Nucleophilic substitution of ε-NH_2 group of lysines can also be excluded since incubation of the sulfonyl azide in TRIS buffer (pH = 8.5) that contains free amino group did not show any reaction even after prolonged time (see stability studies Figure S6 B).

Sulfonyl azides can react with nucleophilic double bonds including indol, N-methylindole or tetrahydropyridine.[7,9,10] However, we did not observe any side reaction on tryptophan under the conditions tested (see Figure S8).

*Potential reaction of sulfonyl azide **1** with tryptophan*

Sulfonyl azide **1** (0.5 mM) was incubated with tryptophan (0.5 mM) in MES buffer pH=5.5 at room temperature. As it can be seen from the chromatogram in Figure S8 there is no detectable reaction between tryptophan and sulfonyl azide **1** under the conditions used.

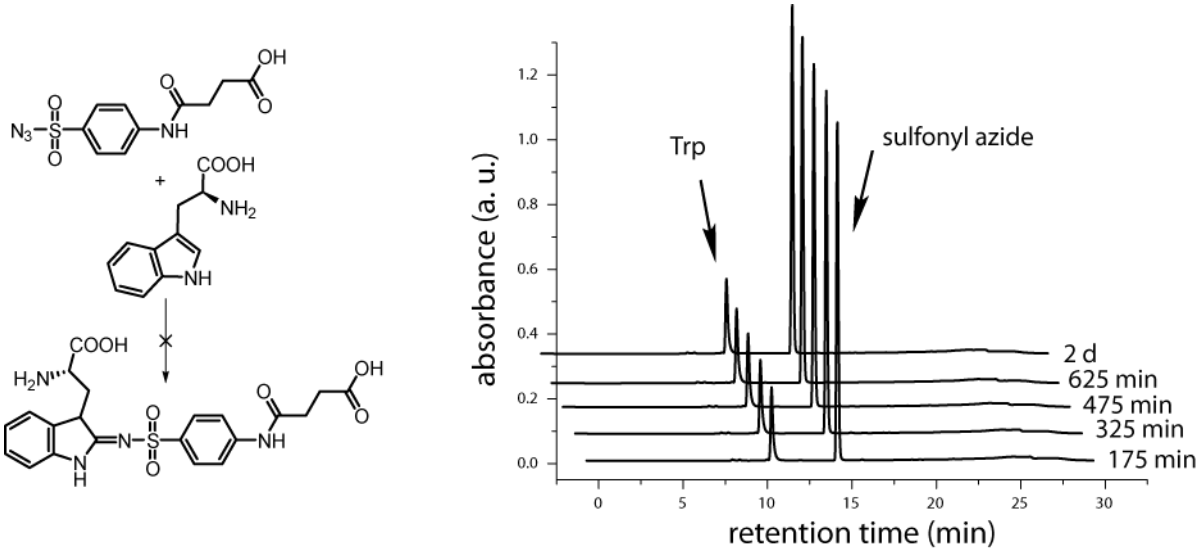


Figure S8: Incubation of sulfonyl azide **1** with tryptophan. The analysis showed no reaction after 2 days.

8. Cloning, expression and purification of norbornene–containing proteins

8.1. Cloning and mutagenesis of IBA3_Trx N65amber

The expression vector pPSG_IBA3_trx was created following the Stargate cloning protocol (IBA, Göttingen) using the primers forward *trxA pENTRY_IBA51* and reverse *trxA pENTRY_IBA51* (Table S1). The commercial vector pBAD202 (Life Technologies) was used as template for *trx*A gene amplification. The amber codon (TAG) at position Asn65 was introduced by blunt end site directed mutagenesis using primers *forward trxA N65amber* and *reverse trxA N65amber*.

Table S1: Sequences of the used primers for generation of expression vector IBA trx N65amber.

Name	Sequence
forward trxA pENTRY_IBA51	**AGCGGCTCTTC** AATG GGA TCT GAT AAA ATT ATT CAT CTG ACT GAT G
reverse trxA pENTRY_IBA51	**AGCGGCTCTTC** TCCC CAG GTT AGC GTC GAG GAA CTC TTT C
forward trxA N65amber	5'phosph CCG GGC ACT GCG CCG AAA TAT G
reverse trxA N65amber	5'phosph **CTA** GTG ATC GAT GTT CAG TTT TGC AAC GG

8.2. Expression of norbornene–containing Trx

E. coli strain BL21(DE3) was transformed with pACyc_pylRS Norb, 3xpylT[1] and pPSG_IBA3_trxA N65amber. An overnight culture was used for inoculation (to OD_{600} of 0.1) of 2 L LB medium containing 2 mM **4**, 50 mg/L carbenicillin and 34 mg/L chloramphenicol. Cells were grown at 37 °C until an OD_{600} of 0.6. Expression of trxA N65X was induced by addition 1 mM IPTG. The cells were shaken for further 16 h at 30 °C. After centrifugation (10.000 x g, 10 min, 4°C) the cells were stored at -20°C or directly purified.

All purification steps were carried out at 4°C. Cells were resuspended in StrepA buffer (10 mM Tris-HCl pH 8, 1 mM EDTA, 150 mM NaCl) and supplemented with protease inhibitor mix (*Roche*). The cells were lysed using a French press (*Thermo Scientific*). Cell debris was removed by centrifugation (38 000 x g, 30 min, 4 °C) and the supernatant was applied to a 5 mL Strep-Tactin column (*IBA*), equilibrated with StrepA buffer. Proteins were eluted from the column using the same buffer containing 2.5 mM desthiobiotin. The eluted fractions were pooled and concentrated. Typical protein yields of 25 mg per liter expression medium of **4** containing Trx were achieved. Purified Trx N65X was equilibrated with StrepA buffer, concentrated and stored at -80 °C. A raw intact MS spectrum of Trx N65X (X = norbornene amino acid **4**) is shown in Figure S9. The deconvoluted spectrum is shown in red in Figure 3A of the main text.

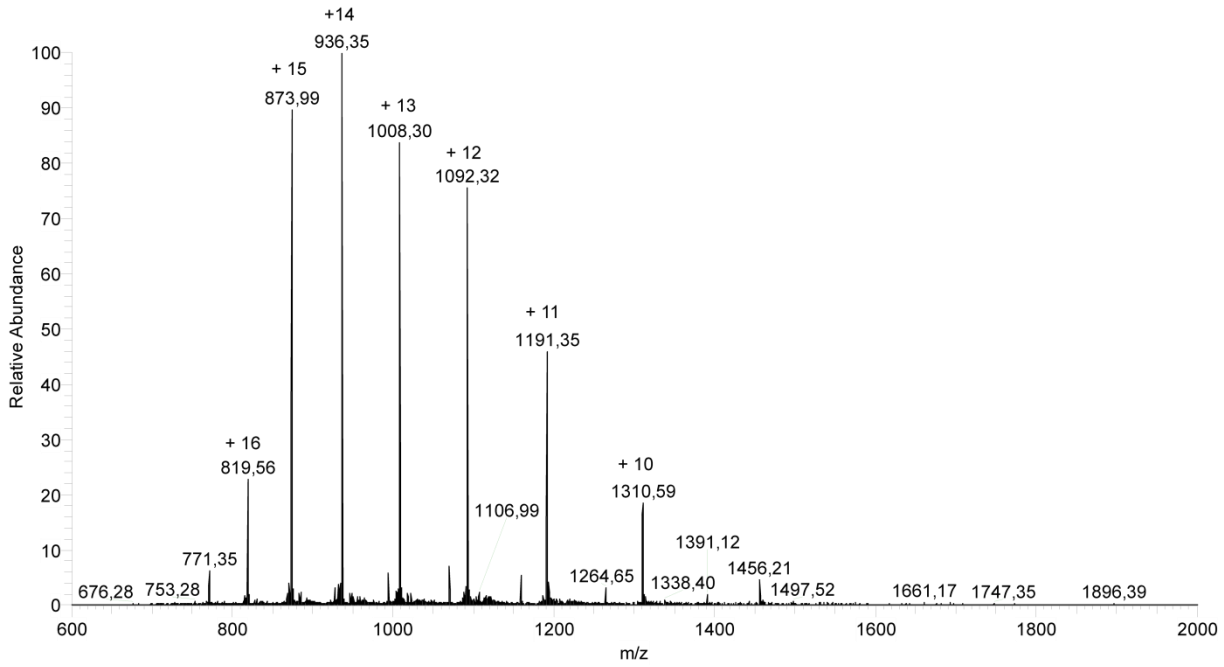


Figure S9: Raw intact MS spectrum of Trx N65X (X = norbornene amino acid **4**)

8.3. Tryptic digestion and MS/MS of norbornene–containing Trx

The sequence of Trx is shown in Table S2. Position Asn65 which was chosen for the incorporation of amino acid **4** is shown in red. The peptide generated after tryptic digestion is emphasized in bold letters. Figure S10 shows the corresponding MS/MS spectrum. Table S3 shows the expected and identified MS/MS fragments of the relevant tryptic peptide.

Table S2: Amino acid sequence of Trx

.

```
         10          20          30          40          50          60
MGSDKIIHLT DDSFDTDVLK ADGAILVDFW AHWCGPCKMI APILDEIADE YQGKLTVAKL

         70          80          90         100         110         120
NIDHNPGTAP KYGIRGIPTL LLFKNGEVAA TKVGALSKGQ LKEFLDANLG SAWSHPQFEK
```

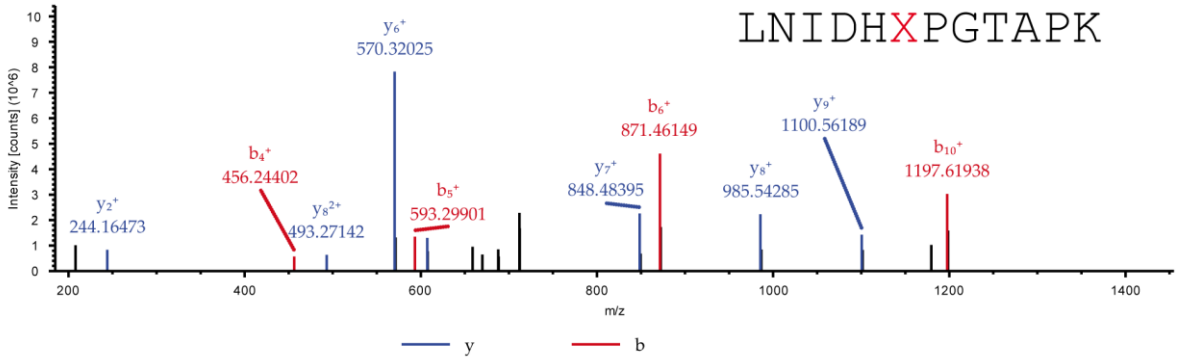


Figure S10: MS/MS spectrum of the tryptic peptide LNIDHXPGTARK (X = **4**). Parent ion: $[M+2H]^{2+}_{calc.}$ = 720.8961, $[M+2H]^{2+}_{obs.}$ = 720.8932 (ΔM = 4 ppm).

Table S3: Expected and identified MS/MS fragments of the tryptic peptide LNIDHXPGTARK (X = **4**). Identified fragments are shown in red for b ions and blue for y ions.

#1	b^+	Seq.	y^+	y^{2+}	#2
1	114.09135	L			12
2	228.13428	N	1327.70051	664.35389	11
3	341.21835	I	1213.65758	607.33243	10
4	456.24530	D	1100.57351	550.79039	9
5	593.30421	H	985.54656	493.27692	8
6	871.46724	**X**	848.48765	424.74746	7
7	968.52001	P	570.32462	285.66595	6
8	1025.54148	G	473.27185	237.13956	5
9	1126.58916	T	416.25038	208.62883	4
10	1197.62628	A	315.20270	158.10499	3
11	1294.67905	P	244.16558	122.58643	2
12		K	147.11281	74.06004	1

8.4. Mutagenesis of pACA_HCA H36amber

The amber codon (TAG) was introduced into the expression vector pACA_HCA [11] at position His36 of the human carbonic anhydrase II gene by blunt end site directed mutagenesis using the primers *forward HCA H36amber* and *reverse HCA H36amber*.

Table S4: Sequences of the used primers for the generation of expression vector pACA_HCA H36amber.

Name	**Sequence**
forward HCA H36amber	5'phosph GTT GAC ATC GAC ACT **TAG** ACA GCC AAG TAT GAC
reverse HCA H36amber	5'phosph AGG GGA CTG GCG CTC TCC CTT GG

8.5. Expression of norbornene containing HCA

The expression vector pACA_HCA G131amber was transformed together with pACyc_pylRS Norb, 3xpylT[1] which contains the genes of the triple mutant of PylRS and three copies of *pyl*T in *E. coli* BL21(DE3) cells (NEB). 1 L of LB medium containing 34 mg/L chloramphenicol, 100 mg/L carbenicillin and 2 mM norbornene amino acid **4** was inoculated with 10 mL of an overnight culture. The cells were stirred at 37 °C until an OD_{600} of 0.9. At this optical density 1 mM $ZnSO_4$ and 0.1 mM IPTG were added to induce the expression of the HCA H36amber gene. After further 10 h at 37 °C the cells were harvested and stored at -20 °C until further use. The harvested cells were resuspended in washing buffer (25 mM Tris; 50 mM Na_2SO_4; 50 mM $NaClO_4$; pH 8.8) and disrupted by French Press procedure. The supernatant of the centrifuged lysate was used for sulfonamide affinity protein purification using an ÄKTA purifier system. The self-packed 3 mL column of *p*-Aminomethylbenzenesulfonamide-Agarose resin (Sigma-Aldrich, A0796) was equilibrated with washing buffer. After binding (0.75 mL/min) of the protein solution, the column was washed with 7 column volumes of washing buffer. HCA was eluted by lowering the pH by elution buffer (100 mM NaOAc; 200 mM $NaClO_4$; pH 5.6). The protein containing fractions were combined, analyzed by SDS-PAGE, dialyzed against water and lyophilized. Typical yields of the pure norbornene amino acid **4** containing protein HCA H36X were 20 mg/L expression medium.

8.6. Tryptic digestion and MS/MS of norbornene-containing HCA

The sequence of HCA II is shown in Table S5. Position His36 which was chosen for the incorporation of amino acid **4** is shown in red. The peptide generated after tryptic digestion is emphasized in bold letters. Figure S11 shows the corresponding MS/MS spectrum. Table S6 shows the expected and identified MS/MS fragments of the relevant tryptic peptide.

Table S5: Amino acid sequence of HCA II.

```
         10         20         30         40         50         60
MAHHWGYGKH NGPEHWHKDF PIAKGERQSP VDIDTHTAKY DPSLKPLSVS YDQATSLRIL

         70         80         90        100        110        120
NNGHAFNVEF DDSQDKAVLK GGPLDGTYRL IQFHFHWGSL DGQGSEHTVD KKKYAAELHL

        130        140        150        160        170        180
VHWNTKYGDF CKAVQQPDGL AVLGIFLKVG SAKPGLQKVV DVLDSIKTKG KSADFTNFDP

        190        200        210        220        230        240
RGLLPESLDY WTYPGSLTTP PLLESVTWIV LKEPISVSSE QVLKFRKLNF NGEGEPEELM

        250        260
VDNWRPAQPL KNRQIKASFK
```

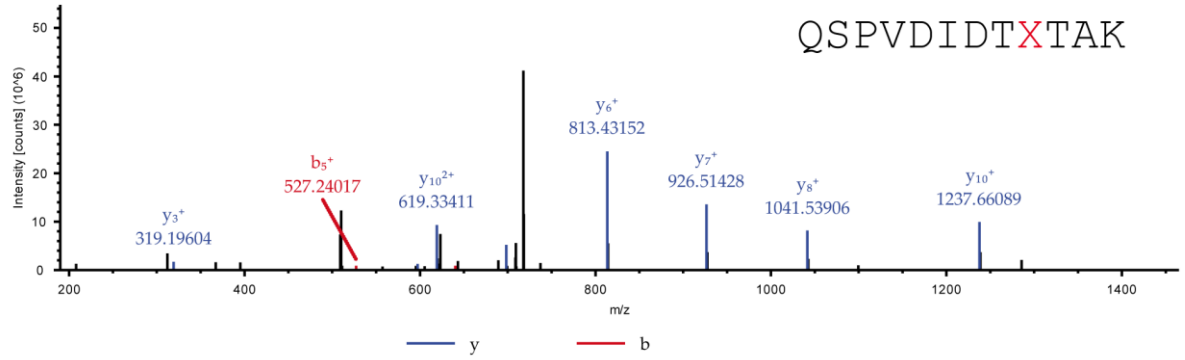


Figure S11: MS/MS spectrum of the tryptic peptide QSPVDIDTXTAK (X = **4**). Parent ion: $[M+2H]^{2+}_{calc.}$ = 726.8829, $[M+2H]^{2+}_{obs.}$ = 726.8807 (ΔM = 3 ppm).

Table S6: Expected and identified MS/MS fragments of the tryptic peptide QSPVDIDTXTAK (X = **4**). Identified fragments are shown in red for b ions and blue for y ions.

#1	b+	Seq.	y+	#2
1	129.06586	Q		12
2	216.09789	S	1324.69949	11
3	313.15066	P	1237.66746	10
4	412.21908	V	1140.61469	9
5	527.24603	D	1041.54627	8
6	640.33010	I	926.51932	7
7	755.35705	D	813.43525	6
8	856.40473	T	698.40830	5
9	1134.56774	**X**	597.36062	4
10	1235.61542	T	319.19761	3
11	1306.65254	A	218.14993	2
12		K	147.11281	1

9. Protein labeling studies

The purified Trx and HCA proteins were used for modification using different sulfonyl azide reagents. The final concentrations of the proteins were 40 µM (in 50 mM Tris buffer pH 7.5 for Trx or 50 mM MES buffer pH 5.5 for HCA). The sulfonyl azide derivatives were added to a final concentration of 2 mM. In case of dansyl sulfonyl azide **6** a 10 mM stock solution in DMF was prepared. The final DMF concentration in this case was therefore 20% (v/v). Biotin sulfonyl azide **5** is water soluble and was therefore added without any organic solvent. After overnight incubation at 37 °C the samples were used for SDS-PAGE and/or intact mass spectrometry analysis. To visualize the modification of HCA H36X with the fluorophore **6** the SDS gel was analyzed using an image reader (Fujifilm LAS-3000). The excitation wavelength was 312 nm. The SDS-gels are shown in the main text (Figure 3). The raw intact MS spectrum of Trx labeled with **5** is shown in Figure S12. The deconvoluted spectrum is shown in blue in Figure 3A in the main text.

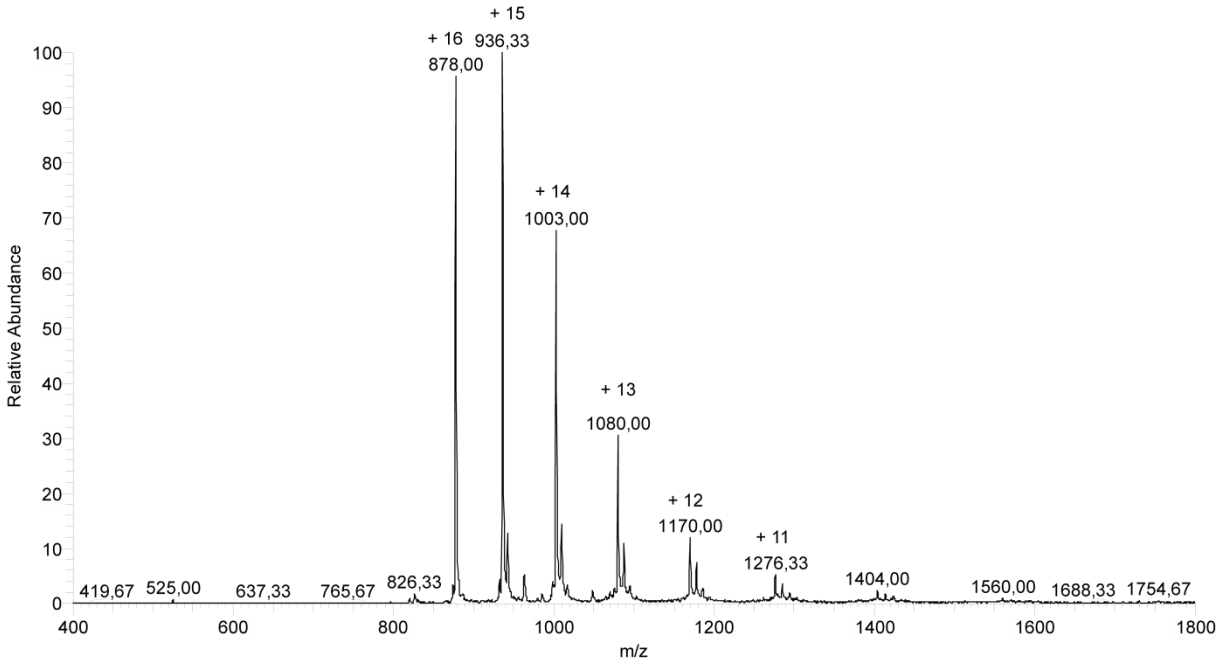
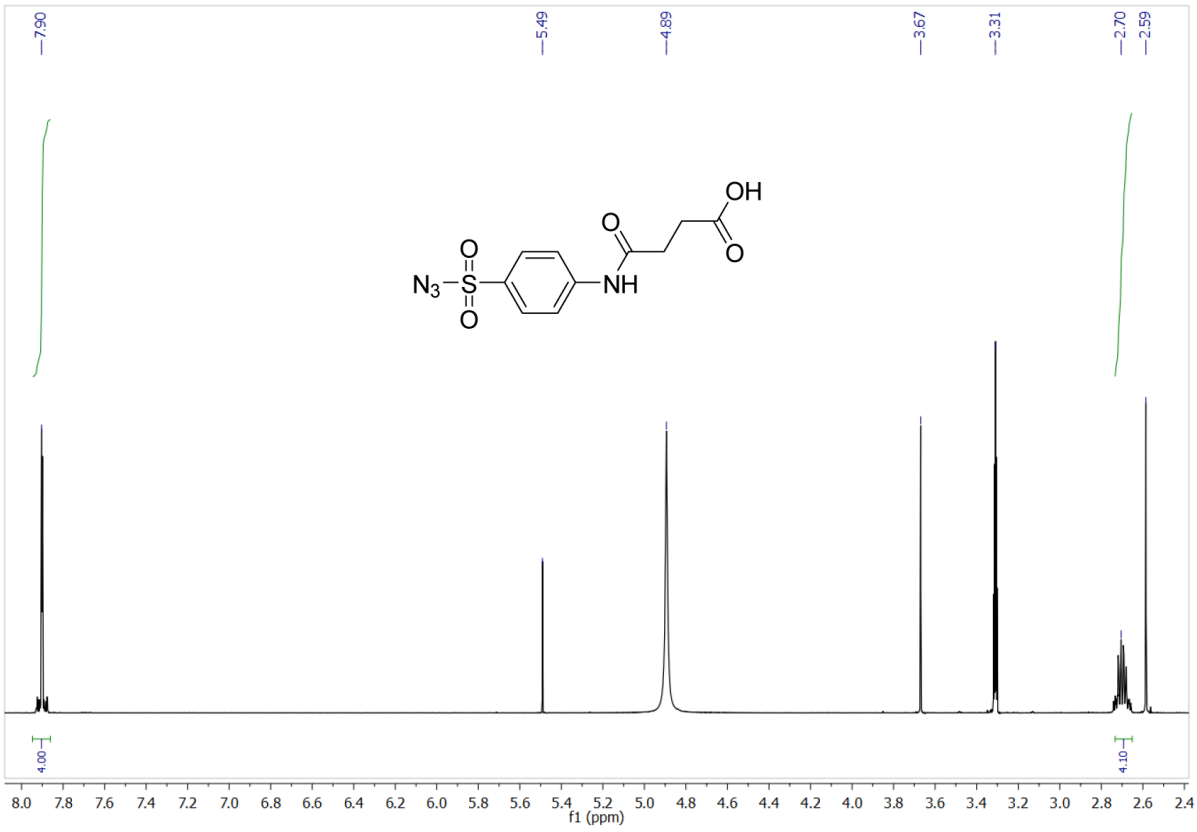
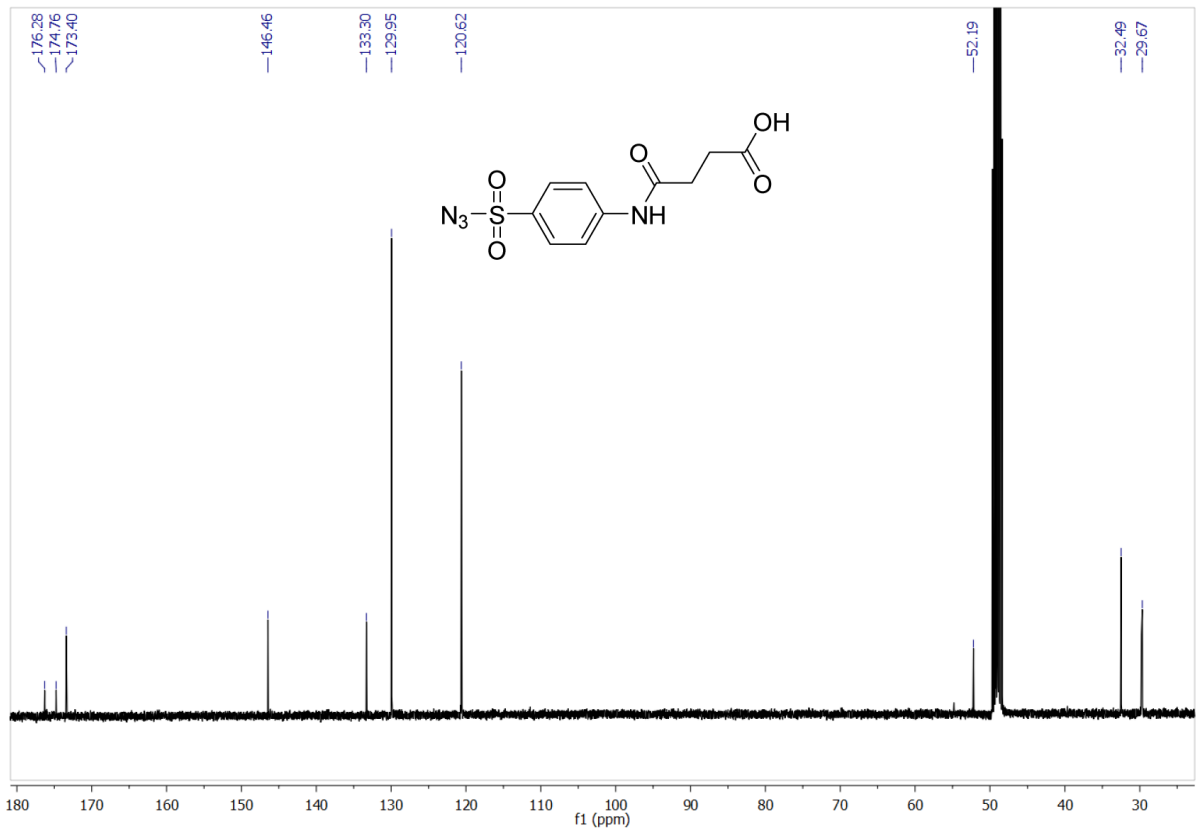
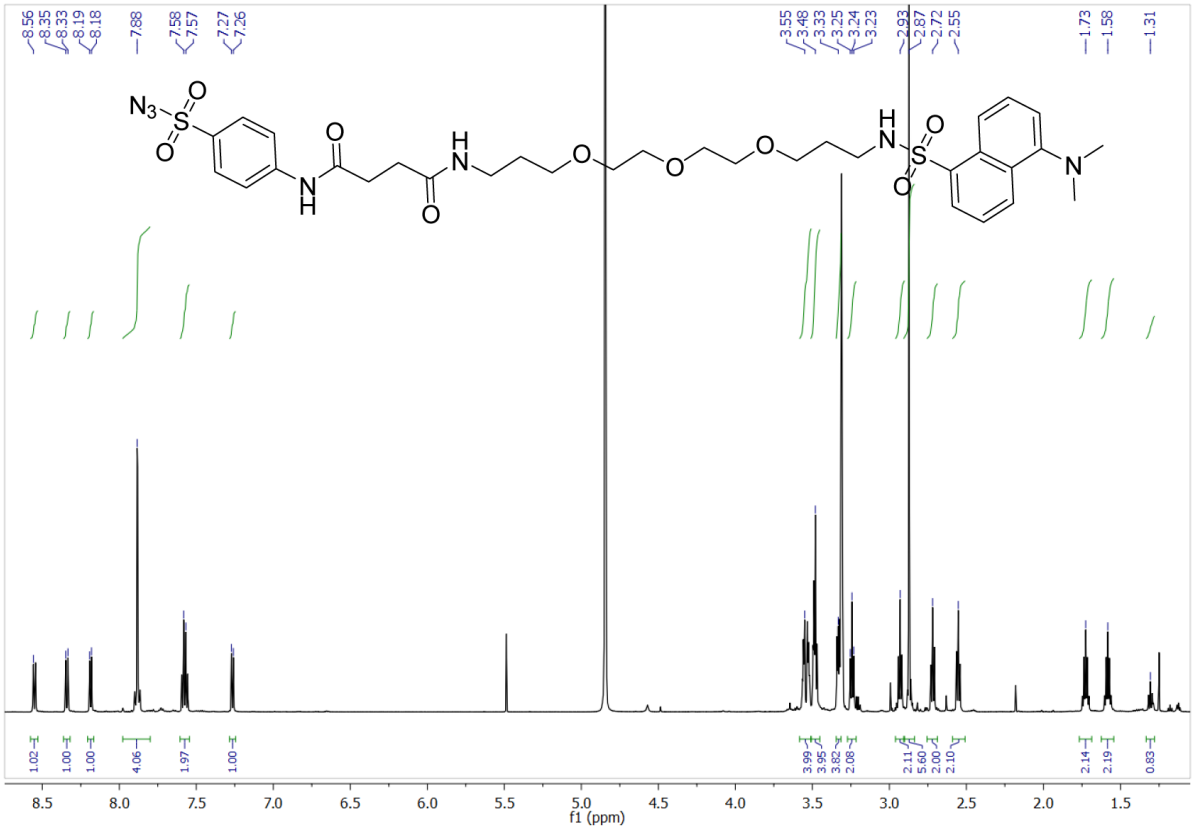


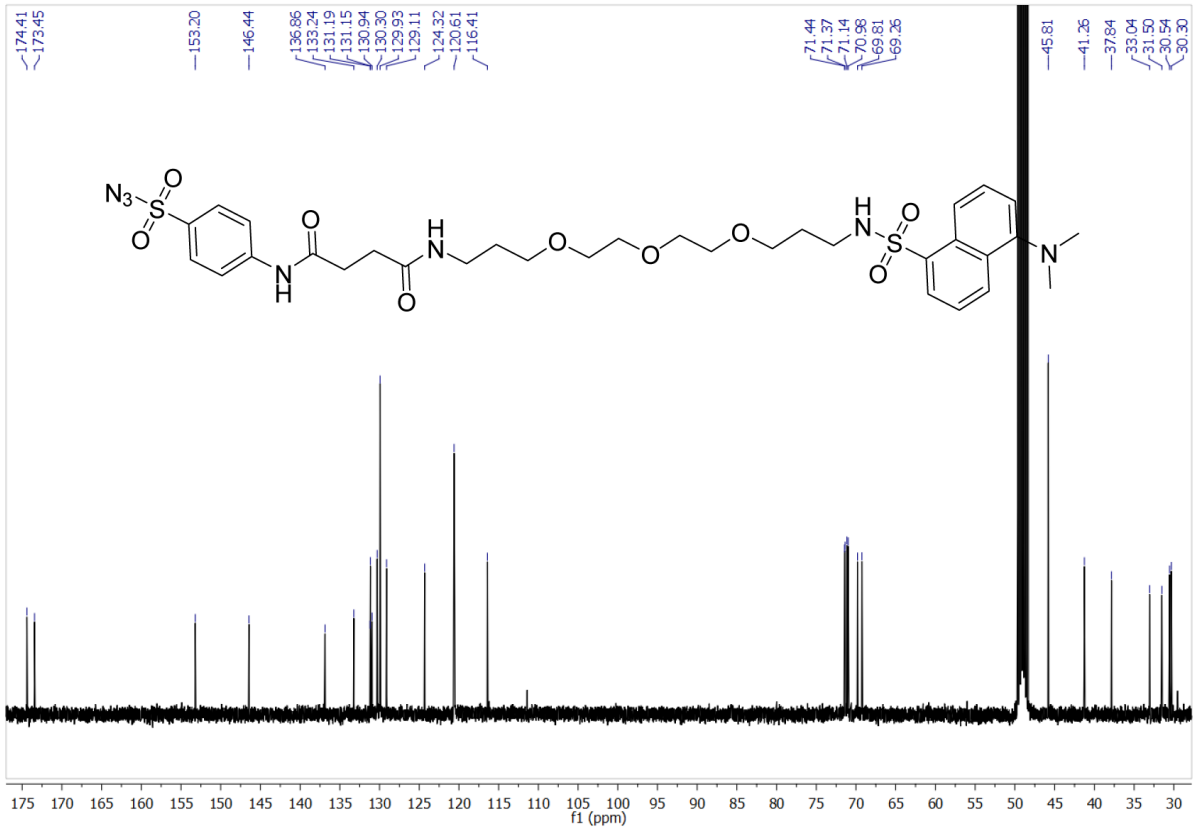
Figure S12: Raw intact MS spectrum of biotinylated Trx N65X (X = norbornene amino acid **4** aziridinated with **5**).

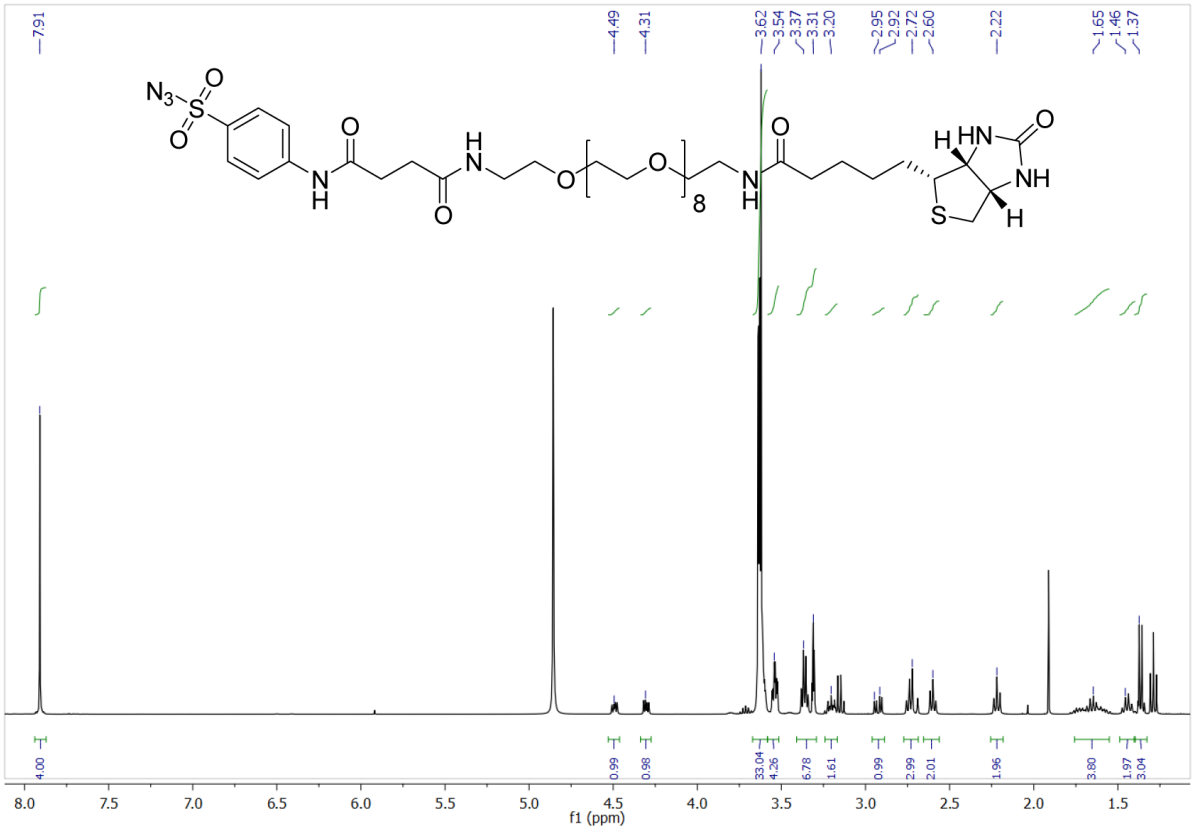
10. NMR spectra

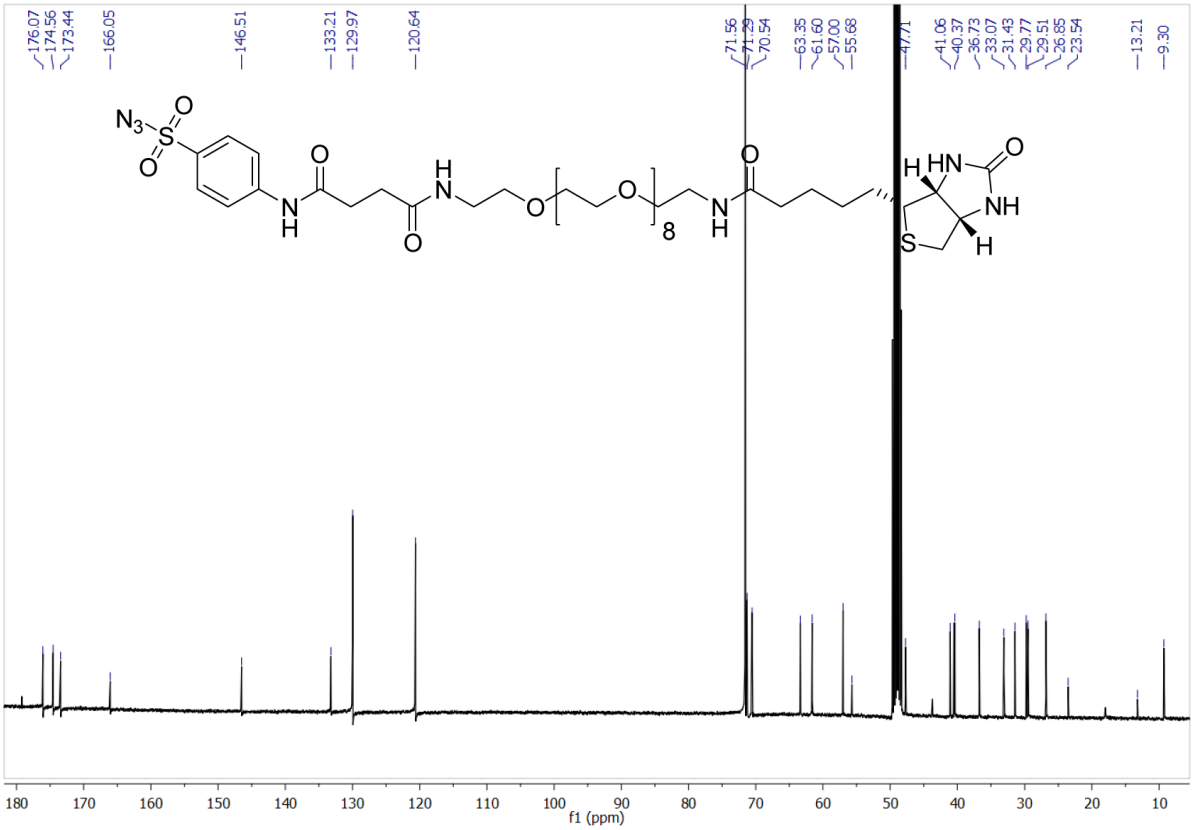


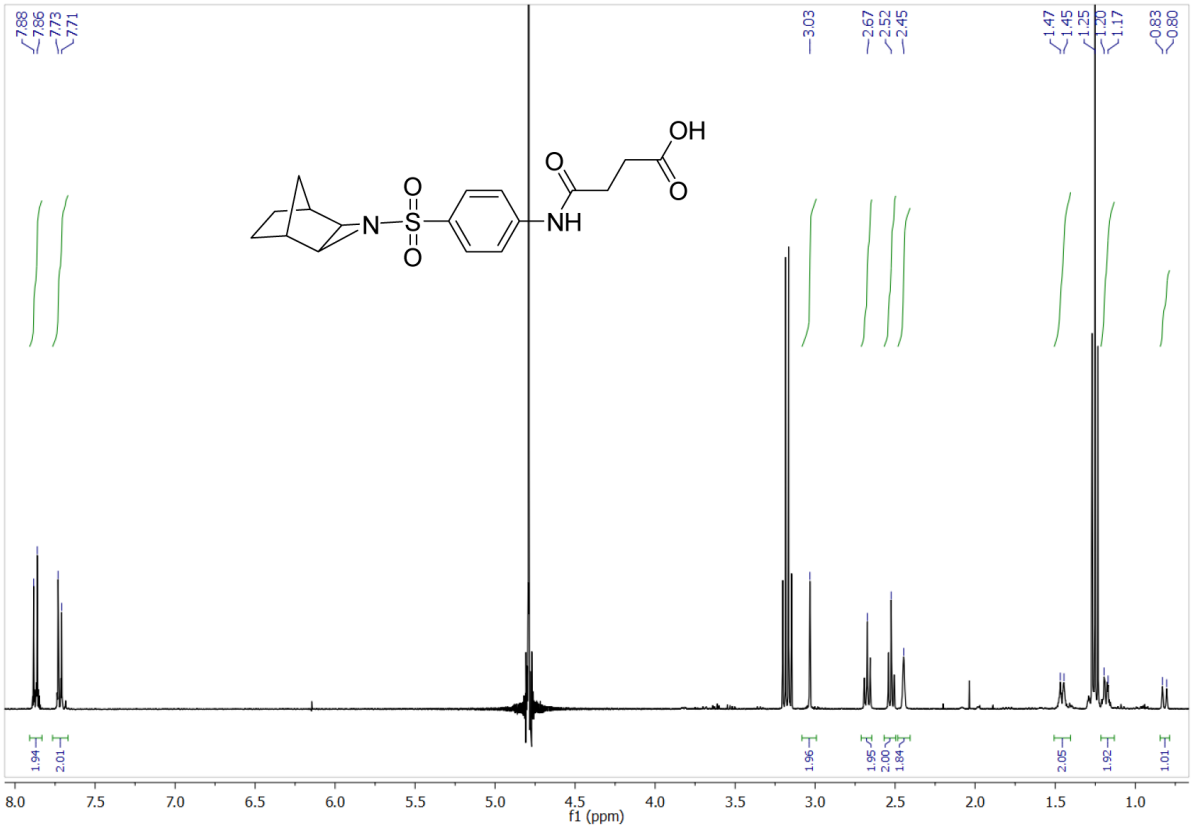


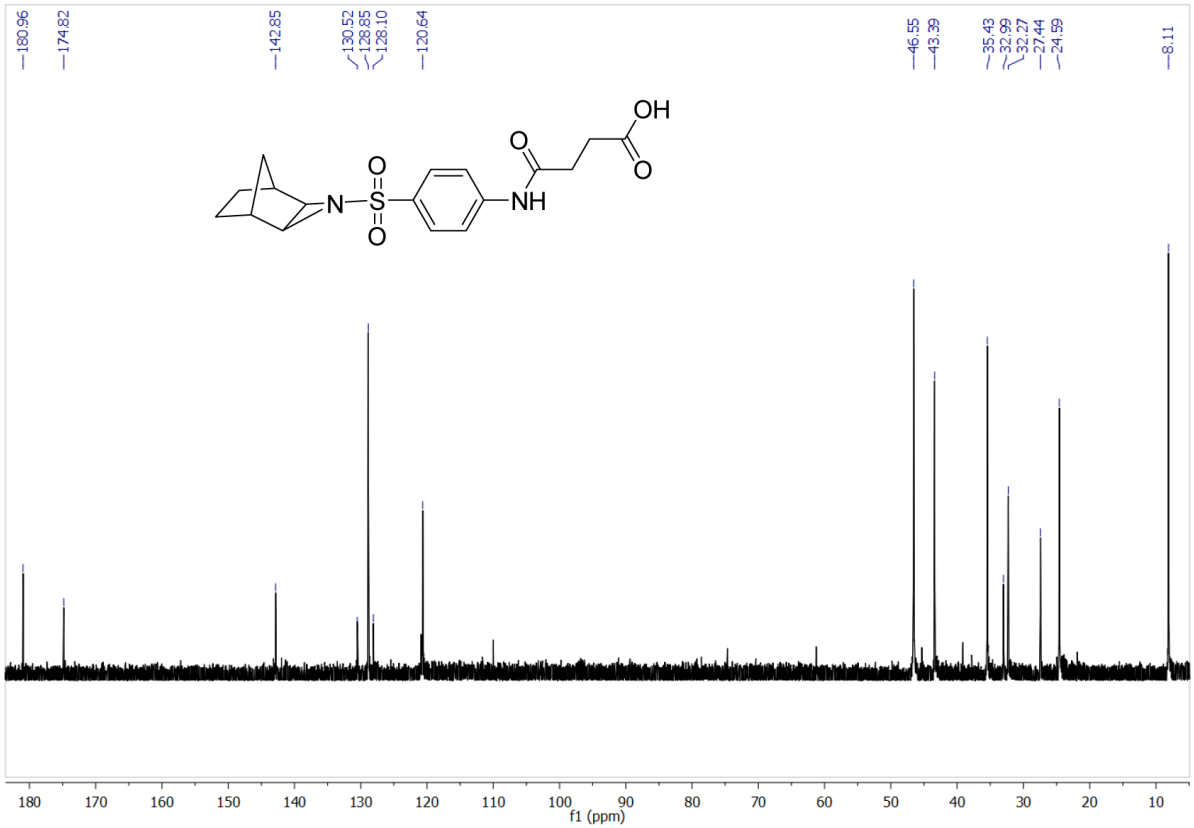

8.56
8.35
8.33
8.19
8.18
7.88
7.58
7.57
7.27
7.26
3.55
3.48
3.33
3.25
3.24
3.23
2.93
2.87
2.72
2.55
1.73
1.58
1.31
1.02
1.00
1.00
4.06
1.97
1.00
3.99
3.95
3.82
2.08
2.11
5.60
2.00
2.10
2.14
2.19
0.83
f1 (ppm)

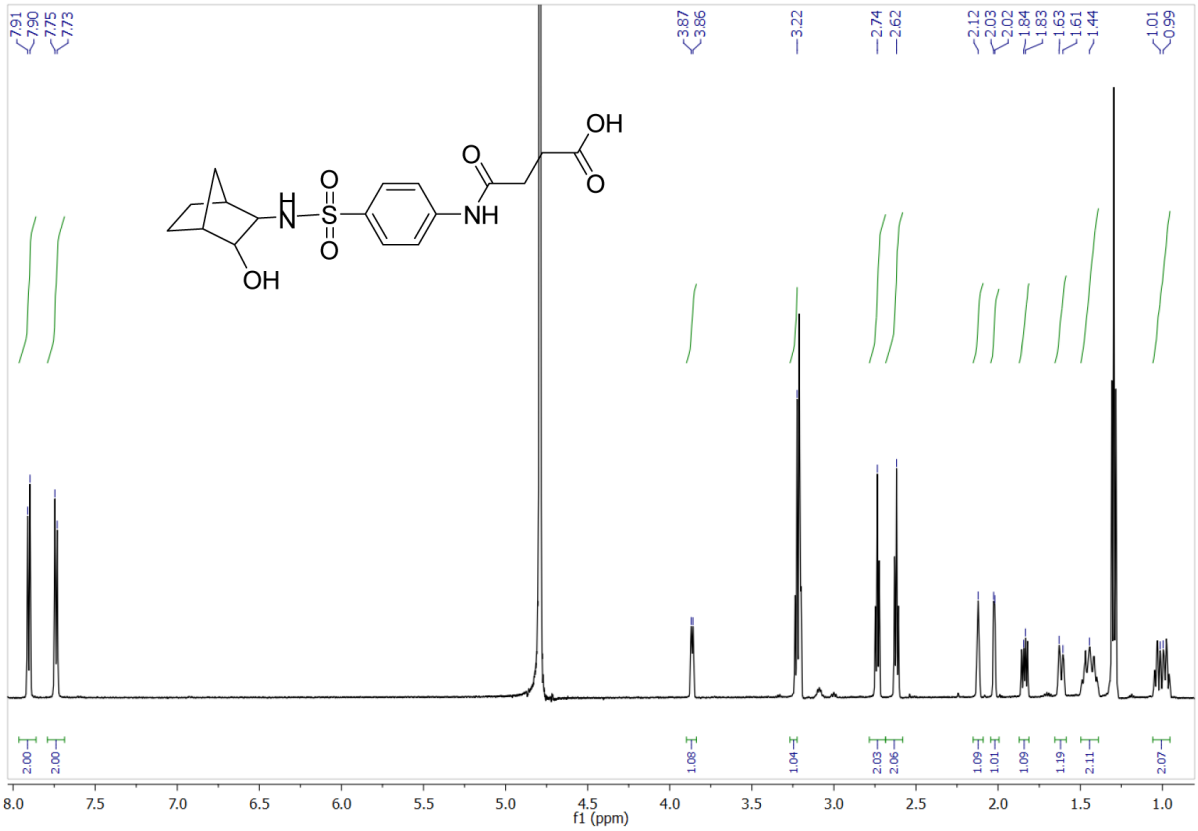

174.41
173.45
153.20
146.44
136.86
133.24
131.19
131.15
130.94
130.30
129.93
129.11
124.32
120.61
116.41
71.44
71.37
71.14
70.98
69.81
69.36
45.81
41.26
37.84
33.04
31.50
30.54
30.30
f1 (ppm)

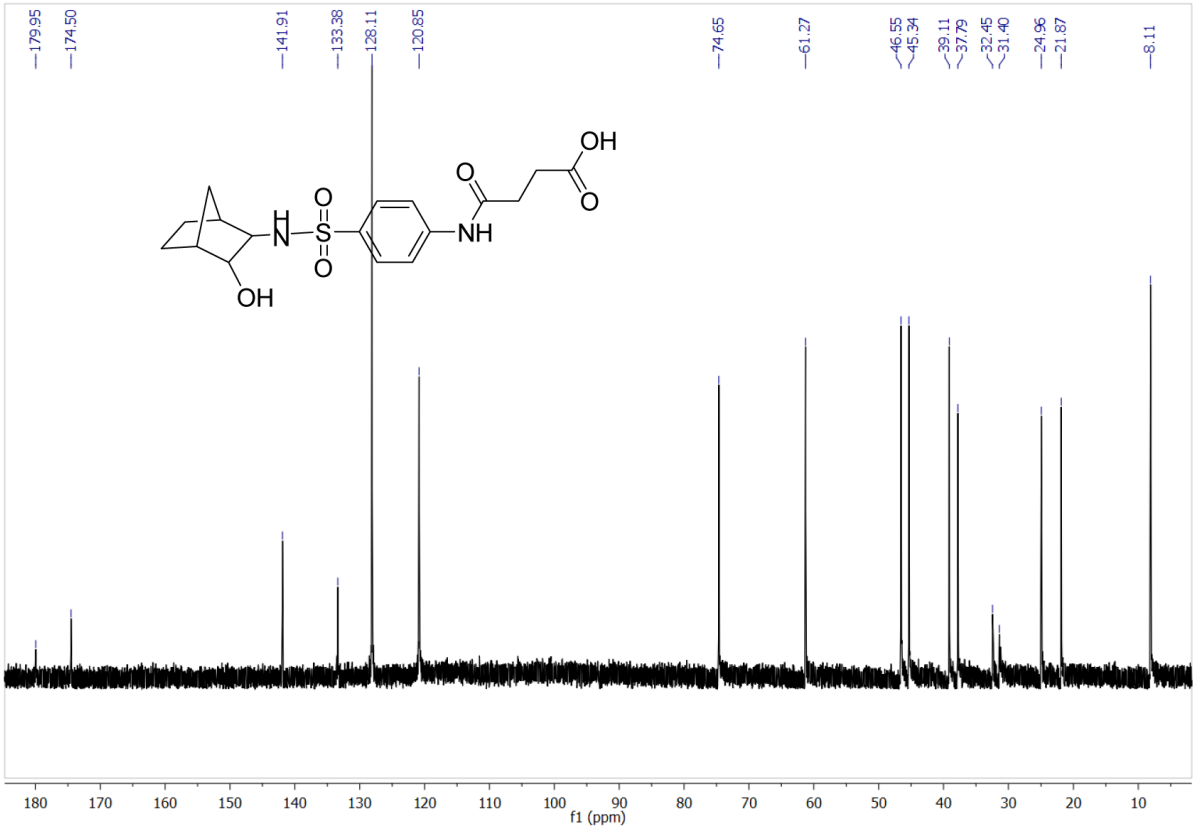

7.91
4.49
4.31
3.62
3.54
3.37
3.31
3.20
2.95
2.92
2.72
2.60
2.22
1.65
1.46
1.37
4.00
0.99
0.98
33.04
4.26
6.78
1.61
0.99
2.99
2.01
1.96
3.80
1.97
3.04
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
f1 (ppm)


176.07
174.56
173.44
166.05
146.51
133.21
129.97
120.64
71.56
71.29
70.54
63.35
61.60
57.00
55.68
47.71
41.06
40.37
36.73
33.07
31.43
29.77
29.51
26.85
23.54
13.21
9.30
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
f1 (ppm)


7.88
7.86
7.73
7.71
3.03
2.67
2.52
2.45
1.47
1.45
1.25
1.20
1.17
0.83
0.80
OH
O
O
N
S
O
O
NH
1.94
2.01
1.96
1.95
2.00
1.84
2.05
1.92
1.01
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)


180.96
174.82
142.85
130.52
128.85
128.10
120.64
46.55
43.39
35.43
32.99
32.27
27.44
24.59
8.11
OH
O
O
N
S
O
O
NH
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
f1 (ppm)


7.91
7.90
7.75
7.73
3.87
3.86
3.22
2.74
2.62
2.12
2.03
2.02
1.84
1.83
1.63
1.61
1.44
1.01
0.99
OH
O
O
NH
O
S
O
H
N
OH
2.00
2.00
1.08
1.04
2.03
2.06
1.09
1.01
1.09
1.19
2.11
2.07
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)


179.95
174.50
141.91
133.38
128.11
120.85
74.65
61.27
46.55
45.34
39.11
37.79
32.45
31.40
24.96
21.87
8.11
OH
O
O
NH
O
S
O
H
N
OH
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
f1 (ppm)

11. Literature

(1) Kaya, E.; Vrabel, M.; Deiml, C.; Prill, S.; Fluxa, V. S.; Carell, T. *Angew. Chem. Int. Ed.***2012**, *51*, 4466.
(2) Teimouri, A.; Chermahini, A. N.; Emami, M. *Tetrahedron* **2008**, *64*, 11776.
(3) Guo, H. M.; Minakawa, M.; Ueno, L.; Tanaka, F. *Bioorg. Med. Chem. Lett.* **2009**, *19*,1210.
(4) Staros, J. V.; Bayley, H.; Standring, D. N.; Knowles, J. R. *Biochem. Biophys. Res. Commun.* **1978**, *80*, 568.
(5) Handlon, A. L.; Oppenheimer, N. J. *Pharm. Res.* **1988**, *5*, 297.
(6) Ueki, M.; Maruyama, H.; Mukaiyama, T. *Bull. Chem. Soc. Jpn* **1971**, 1108.
(7) Warren, B. K.; Knaus, E. E. *J. Heterocyclic Chem.* **1987**, *24*, 1413.
(8) Cartwright, I. L.; Hutchinson, D. W.; Armstrong, V. W. *Nucleic Acids Res.* **1976**, *3*.
(9) Bailey, A. S.; Churn, M. C.; Wedgwood, J. J. *Tetrahedron Lett.* **1968**, *57*, 5953.
(10) Harmon, R. E.; Wellman, G.; Gupta, S. K. *J. Heterocyclic Chem.* **1972**, *9*, 1191.
(11) Nair, S. K.; Calderone, T. L.; Christianson, D. W.; Fierke, C. A. *J. Biol. Chem.* **1991**, *266*, 17320.