Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

The Wittig Bioconjugation of Maleimide Derived, Water Soluble Phosphonium Ylides to Aldehyde-Tagged Proteins

Supporting Information

Rafael W. Hartmann,[a,b] Matthijs Pijnappel,[a] Johan Nilvebrant,[c] Hildur Run Helgudóttir,[d] Arni Asbjarnsson,[d] Gunnhildur Asta Traustadottir,[d] Thorarinn Gudjonsson,[d,e] Per-Åke Nygren,[c] Fredrik Lehmann,[a] Luke R. Odell*[c]

[a]Recipharm OT Chemistry, Virdings allé 16, 75450 Uppsala, Sweden.
[b]Department of Medicinal Chemistry, Uppsala University, Uppsala Biomediciniska Centrum, Husargatan 3, 75123 Uppsala, Sweden.
[c]Department of Protein Science, Division of Protein Engineering, KTH School of Engineering Sciences in Chemistry, Biology and Health, AlbaNova Universitetscentrum, Roslagsvägen 30B, 10961 Stockholm, Sweden.
[d]Stem Cell Research Unit, Biomedical Center, University of Iceland, 101 Reykjavik, Iceland.
[e]Department of Laboratory Hematology, Landspítali-University Hospital, Reykjavik, Iceland.

Contents

1. General remarks

All solvents used in the synthesis and purification of the compounds described herein were purchased from Chemtronica (Sweden) and used without further purification. Solvents used in non-aqueous reactions were stored over molecular sieves (3Å or 4Å, Sigma-Aldrich/Merck, Germany) prior to use. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP x HCl) was purchased from TCI Europe *via* Chemtronica (Sweden). *N*-benzyl maleimide was purchased from Lancaster Synthesis *via* Chemtronica (Sweden). All other chemicals used as starting materials, reagents or buffer components in the syntheses outlined herein, specifically 6-maleimidocaproic acid, (±)-3-amino-1,2-propanediol, *N*-succinimidyl 4-(maleimidomethyl) cyclohexanecarboxylate, glycine, N_{α}-(*tert*-butoxycarbonyl)-L-lysine, doxorubicin hydrochloride, pyridoxal-5'-phosphate, equine heart myoglobin and Coomassie Blue were purchased from Sigma-Aldrich/Merck (Germany). Cell lines and chromatographic media for the expression and purification of ADAPT6 were purchased from the following suppliers: *E. coli* BL21* (Thermo Fisher Scientific, USA) and TALON metal affinity resin (Takara Bio, Japan).

Aqueous buffers were prepared using an S47 SevenMulti pH-meter equipped with an InLab Routine Pro-ISM electrode (Mettler Toledo, Switzerland). Non-aqueous reactions were carried out in an atmosphere of nitrogen, which was pre-dried using a Drierite gas drying unit (W.A. Hammond Drierite Company, US). Unless otherwise noted, reactions containing only small molecule components were monitored using an Agilent (US) 1100 series LC/MS (single quadrupole) system equipped with an electrospray interface, a UV diode array detector and an ACE3 C8 (3.0 x 50 mm) column (ACE, UK) with a gradient of acetonitrile in 0.1% aqueous trifluoroacetic acid over 3 min and a flow of 1 mL/min.

Flash chromatography was performed automatically using a CombiFlash Rf^{+} Lumen flash machine (Teledyne Isco, US) equipped with a wide-range UV and evaporative light scattering (ELS) detector and prepacked silica columns (SiliCycle, Canada). Preparative HPLC was performed using a Gilson HPLC System (US) equipped with a UV diode array detector and an ACE3 C18-HL (250 x 21.2 mm) column (ACE, UK) with a gradient of acetonitrile in 0.1% aqueous trifluoroacetic acid over 10 min and a flow of 25 mL/min.

Unless otherwise noted, NMR spectra were recorded using an Ascend 400 spectrometer (Bruker, US) at 298K, corresponding to a Larmor frequency of 400MHz for ^{1}H (^{31}P, 162MHz; ^{13}C, 101MHz). In one instance, a ^{13}C-NMR spectrum was recorded using an Avance NEO 500 spectrometer (Bruker, US) at 298K, corresponding to a Larmor frequency of 126MHz for ^{13}C. All NMR-experiments were performed using commercially obtained, deuterated solvents with no further purification (Sigma-Aldrich/Merck, Germany). DMSO-d_6 was stored over molecular sieves (4Å, Sigma-Aldrich/Merck, Germany). NMR spectra were processed and interpreted using MestreNova (Mestrelab, Spain). High-resolution mass spectra of small molecules were acquired using a Premiere LCT mass spectrometer (Waters, US) operating in ES+ mode. The analytes were introduced into the mass spectrometer after chromatography on an Acquity UPLC system, running a gradient of 50% aqueous acetonitrile.

A NanoDrop One^{C} spectrophotometer (Thermo Fisher Scientific, US) was used for the UV-spectrophotometric quantification of proteins. Upon expression, purification and lyophilization, ADAPT6 was analyzed by means of mass spectrometry using 6520 accurate mass q-ToF mass spectrometer (Agilent, US). Upon derivatization, proteins and conjugates were analyzed using an Agilent (US) 1290 Infinity II series LC system equipped with a UV diode array detector and an Agilent (US) 6550 iFunnel q-ToF mass spectrometer. An Acquity UPLC Protein BEH C4 (2.1mm x 50 mm) column (Waters, US) and unless otherwise noted, a gradient of 0.1% formic acid in acetonitrile in 0.1% aqueous formic acid over 10 min and a flow of 1 mL/min were used. In one case, a gradient of acetonitrile in 0.1% aqueous ammonia was used instead. Prior to application to the gel, SDS-PAGE samples were incubated with premixed Laemmli sample buffer (Bio-Rad, US) at 90°C for 5min. Electrophoresis was then run on a NuPAGE™ (Thermo Fisher Scientific, US) 4-12% Bis-Tris Gel. One band was loaded with Amersham Low Molecular Weight calibration kit (GE Healthcare, US). The gel was stained with Coomassie Brilliant Blue G-250 Dye (Thermo Fisher Scientific, US).

Cells used for cytotoxicity experiments were cultured in flasks coated with collagen I (Advanced BioMatrix, US). PrestoBlue™ (Thermo Fisher Scientific, US) was used as a viability indicator.

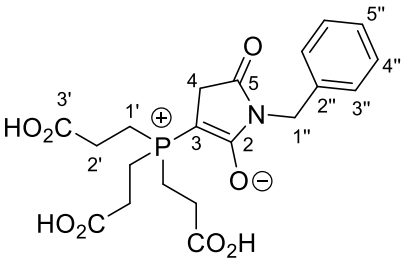
2. Characterization of the reaction between maleimides and TCEP

2.1 *N*-benzyl maleimide (2) derived TCE-phosphonium ylide

2.1.1 Synthesis *in situ*

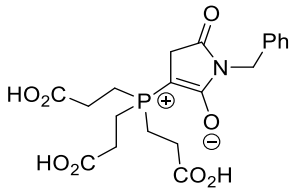
Tris(2-carboxyethyl)phosphine hydrochloride (7.2mg, 25µmol, 1.0eq) was dissolved in phosphate buffered saline[1] (50mM, pH=7.4, 60µL). The resulting mixture was diluted with a solution of *N*-benzyl maleimide (4.7mg, 25µmol, 1.0eq) in DMSO-d_6 (540µL), and triethylamine (0.4µL, 2.8µmol, 0.1eq) was added. The solution was then analyzed by NMR. Aside from confirming the tautomeric identity of the ylide, the analysis also confirmed the presence of trace amounts of both starting materials (N-benzyl maleimide: ^{1}H NMR (400 MHz, DMSO-d_6), δ 7.01 ppm (s, C*H*-C=O); TCEP: ^{1}H NMR (400 MHz, DMSO-d_6), δ 2.47-2.38 ppm and 1.99-1.90 ppm (m, 2xCH_2).), lending credence to the hypothesis that ylide formation reaches an equilibrium and is therefore reversible.

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.33 – 7.15 (m, 5H, H-3''/4''/5''), 4.57 (s, 2H, H-1''), 3.26 (d, *J* = 11.3 Hz, 2H, H-4), 2.73 – 2.55 (m, 12H, H-1'/2'). **^{31}P NMR** (162 MHz, DMSO-d_6, {^{1}H} decoupled) δ 38.4. **^{13}C NMR** (101 MHz, DMSO-d_6) δ 174.5, (d, *J* = 8.8Hz, C-2), 172.8 (d, *J* = 14.8Hz, C-3'), 172.5 (d, *J* = 3.4Hz, C-5), 135.6 (C-2''), 129.0 (C-4''), 128.1 (C-5''), 127.9 (C-3''), 42.9 (C-1''), 33.7 (d, *J* = 53.4Hz, C-3) 29.8 (s, *J* ≈ 0Hz, C-4), 26.2 (d, *J* = 3.3 Hz, C-2'), 14.4 (d, J = 49.7 Hz, C-1'). **HRMS** calculated for $C_{20}H_{25}NO_8P^+$ $[M+H]^+$, 438.1318; found, 438.1319; Calculated[1] for $C_{20}H_{24}DNO_8P^+$ $[M+H]^+$, 439.1380; found, 439.1392.



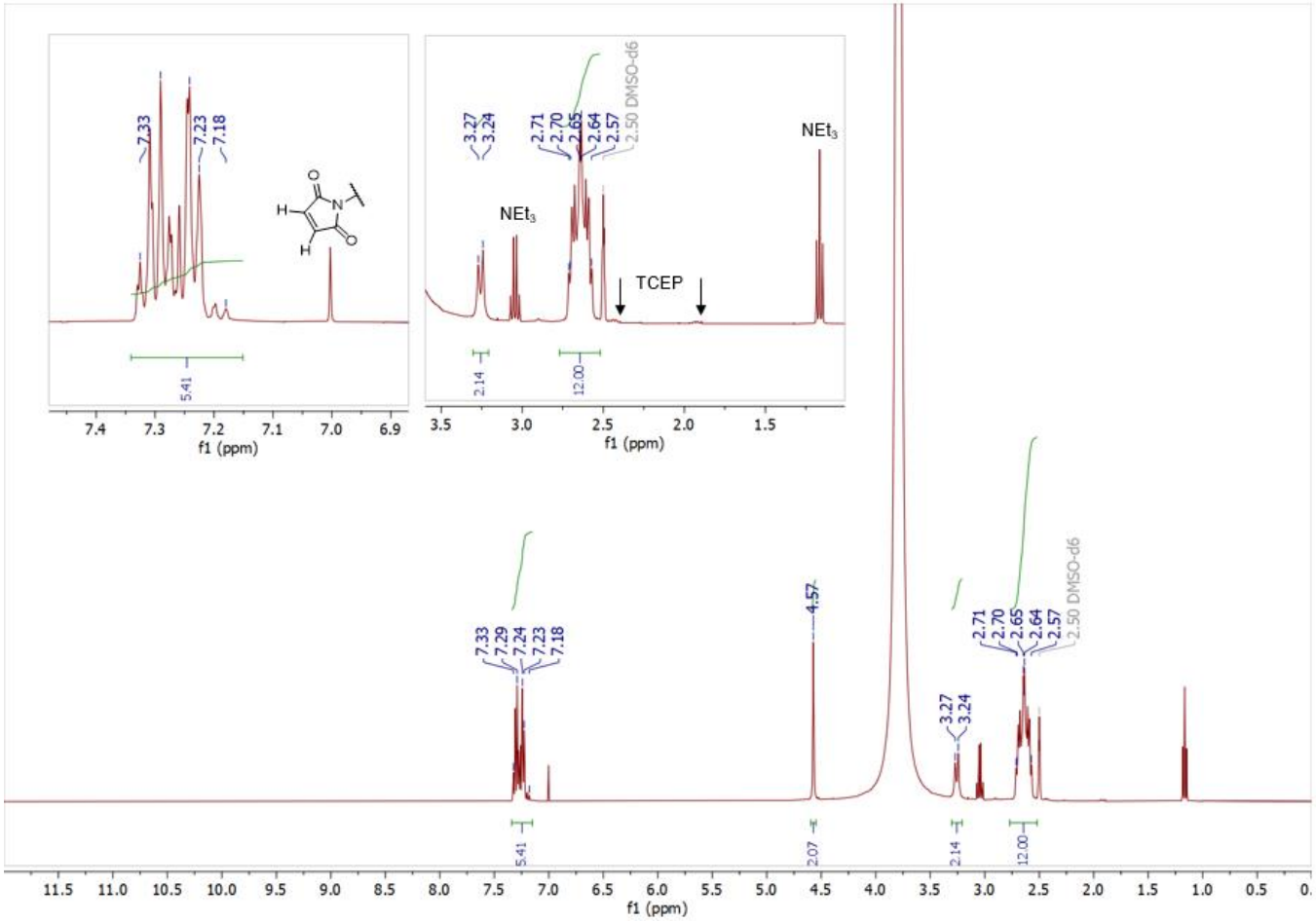
[1] To obtain the deuterated HRMS sample, the ylide was formed in PBS buffer made with heavy water.

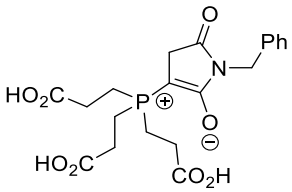
2.1.2 NMR spectra



^{1}H NMR (400MHz),
H_2O/DMSO-d_6 (1:9)

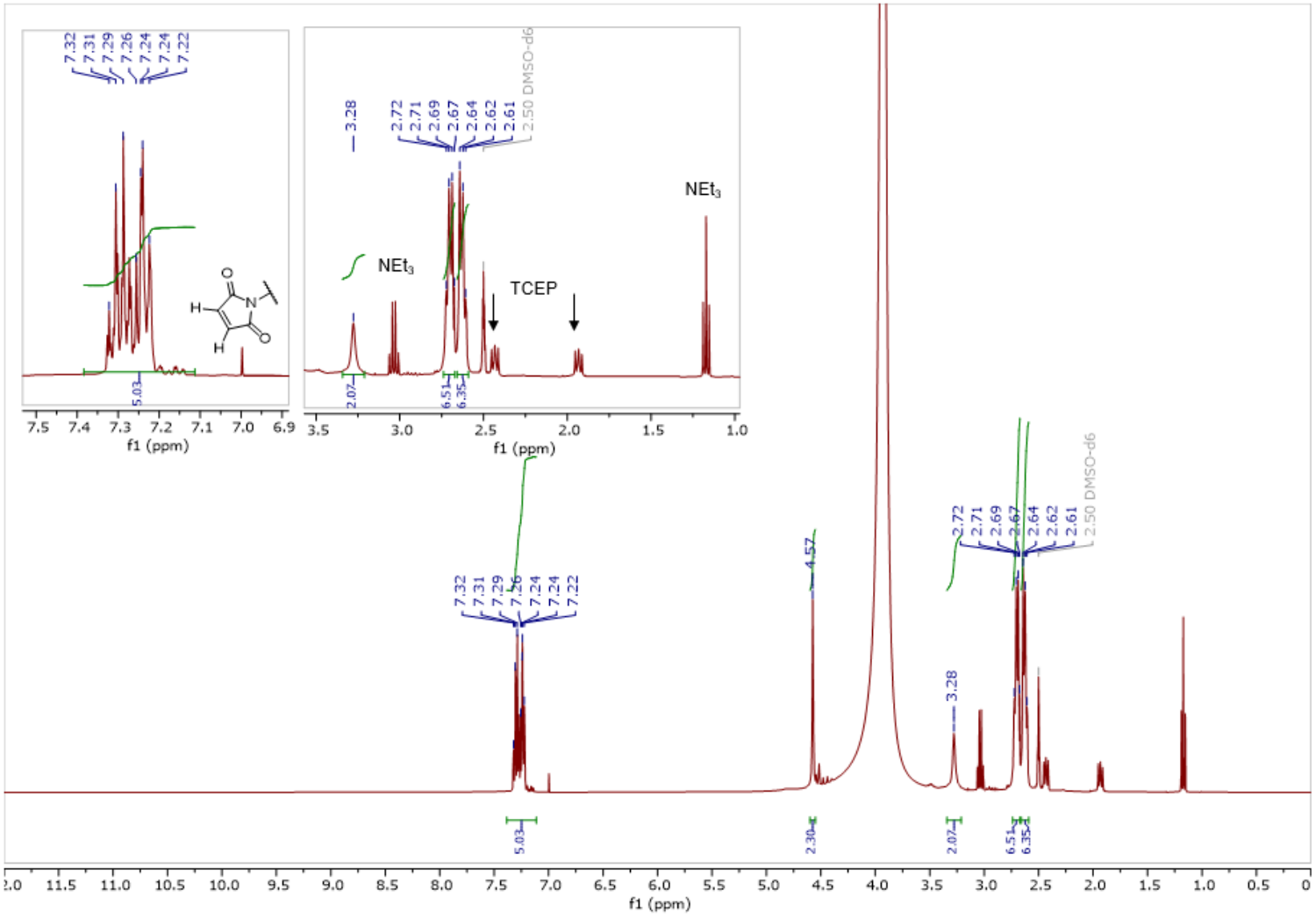
No {^{31}P} decoupling

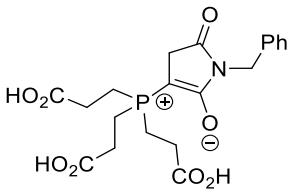




^{1}H NMR (400MHz),
H_2O/DMSO-d_6 (1:9)

{^{31}P} decoupled



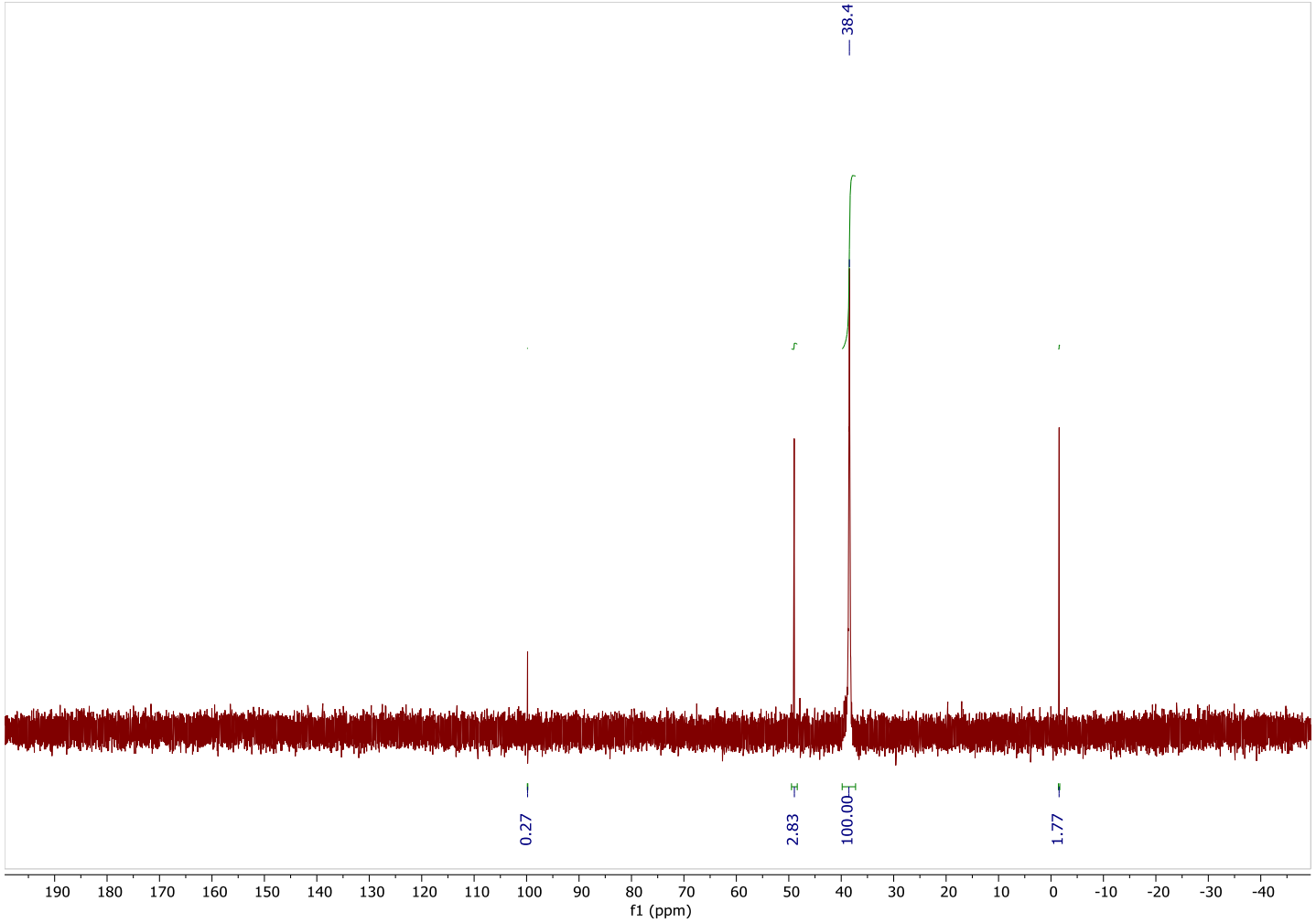


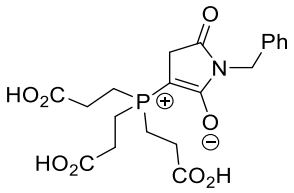
^{31}P NMR (162MHz), H_2O/DMSO-d_6 (1:9)

{^{1}H}- decoupled

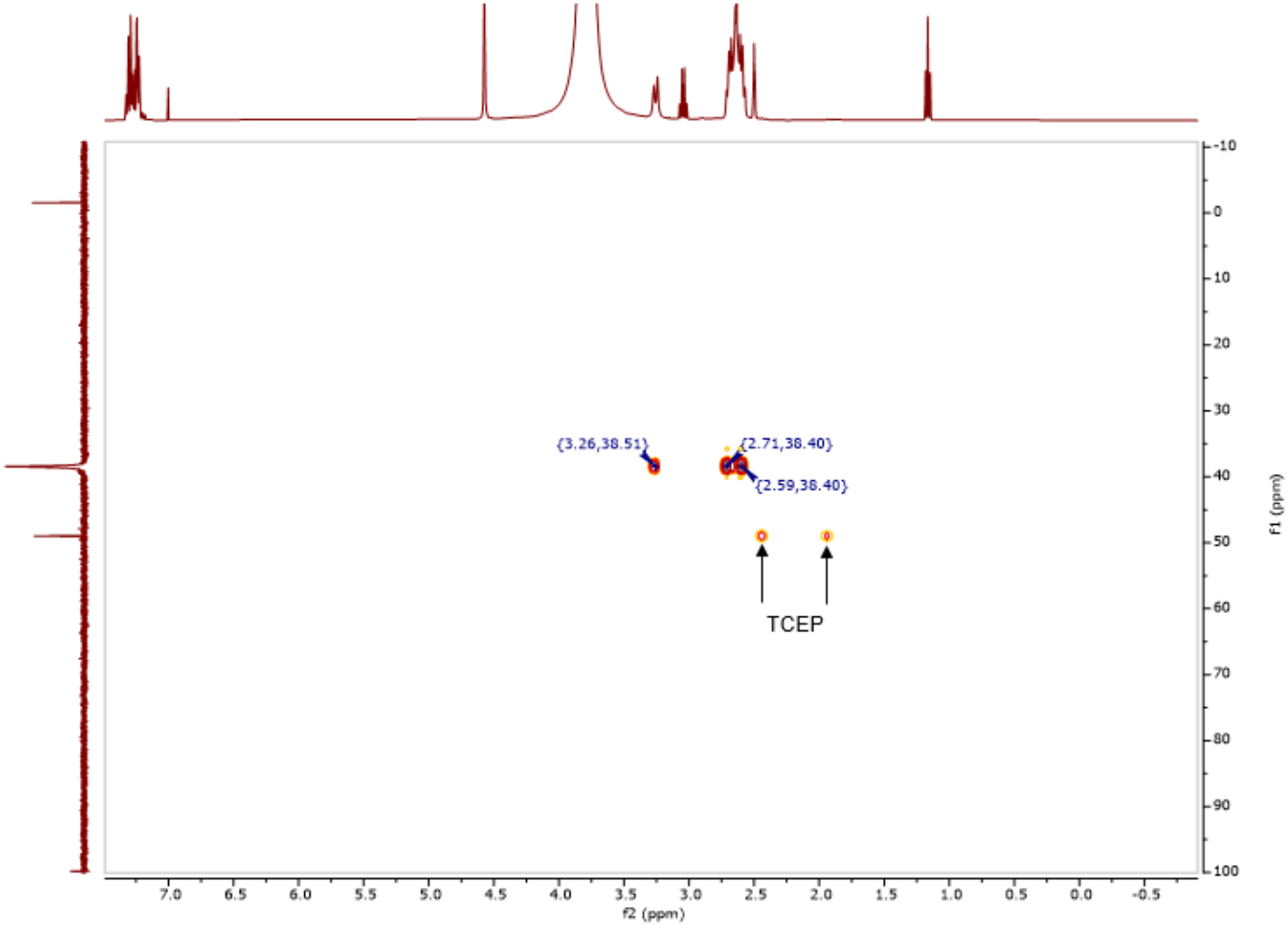
Note: Integrals are not accurate as a result of uneven NOE and an insufficiently long longitudinal relaxation delay but serve to indicate that phosphorous containing impurities/byproducts are only present in trace amounts.

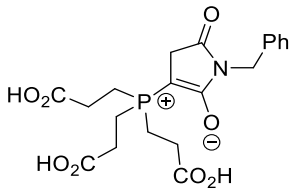
The signal at δ=-1.6ppm likely originates from the phosphate buffer and that at δ=49.0ppm corresponds to residual TCEP as confirmed by ^{1}H,^{31}P HMBC (see next page).



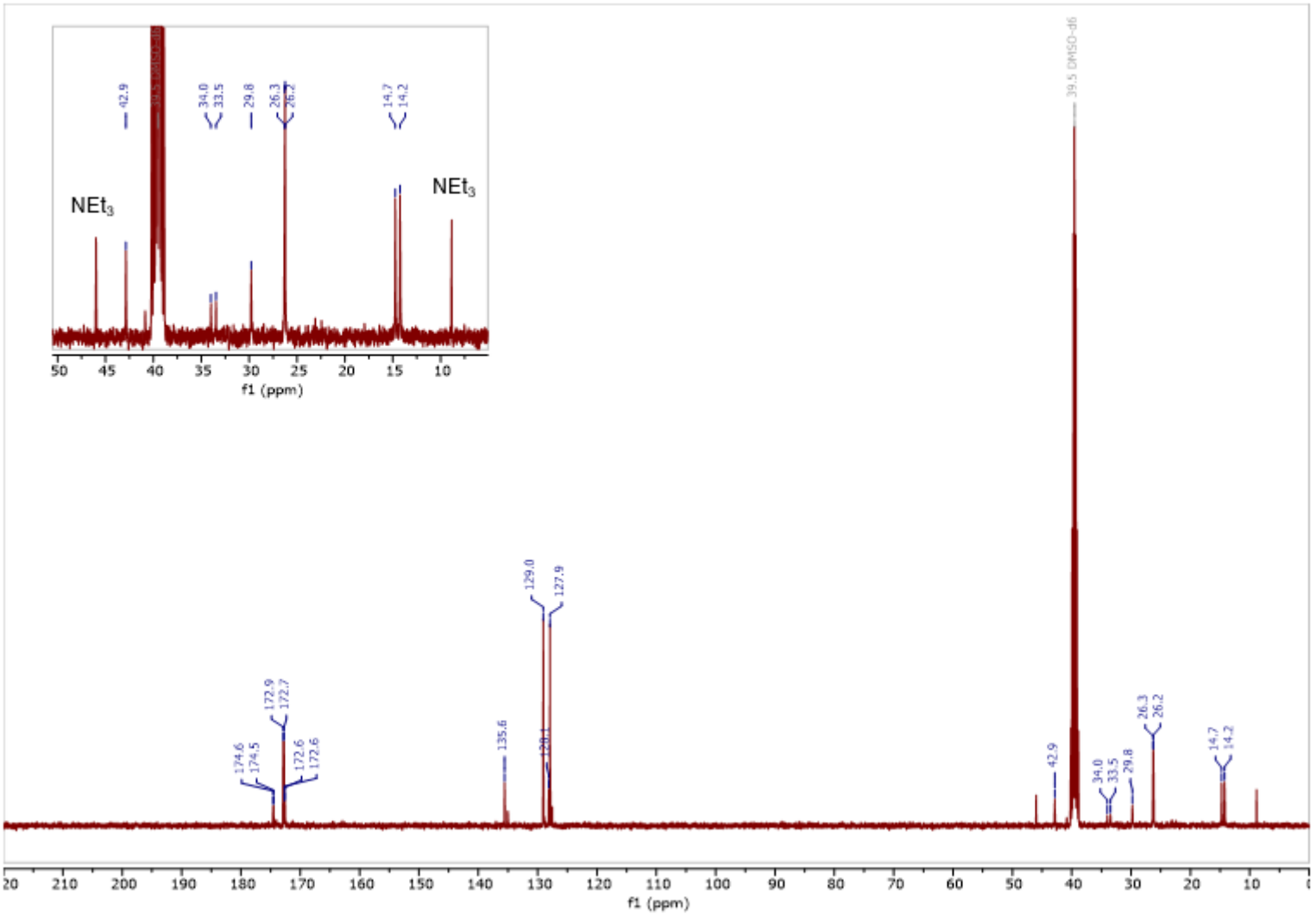


$^{1}H,^{31}P$ HMBC
H_2O/DMSO-d_6 (1:9)





^{13}C NMR (101MHz)
H_2O/DMSO-d_6 (1:9)



2.2 Solubility comparison: PPh_3 versus TCEP + maleimide 2

2.2.1 PPh_3 experiment

A 20mL vial was charged with triphenylphosphine (15.7mg, 59.8µmol, 1.0eq.) and *N*-benzylmaleimide (11.2mg, 59.8µM, 1.0eq.). DMSO (1mL) was added and the mixture was gently shaken until a solution has been obtained (see image **S1**). A drop of water containing triethylamine (0.8µL, 6µmol, 0.1eq.) was added to yield a slightly cloudy solution (see image **S2**). Finally, the mixture was diluted to the target concentration of 6mM by the addition of phosphate buffered saline (20mM, pH=7.4, 9mL) to yield a white suspension (see image **S3**). The mixture remained heterogeneous even after 24 hours (see image **S4**).

Image S1
Maleimide and PPh_3
in DMSO (60mM)

(clear solution)

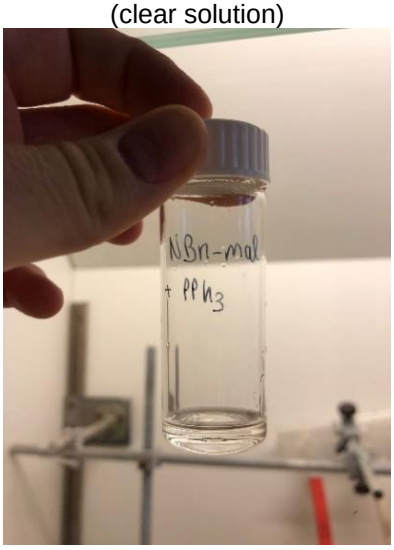


Image S2
Maleimide, PPh_3 and NEt_3
in wet DMSO (~60mM)

(cloudy solution)

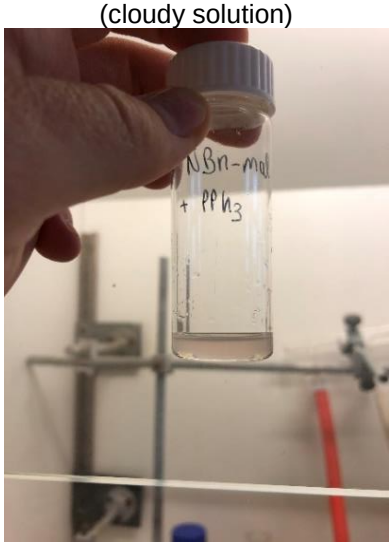


Image S3
Maleimide, PPh_3 and NEt_3
in PBS (20mM, pH7.4)/DMSO 9:1 (6mM)

(suspension)

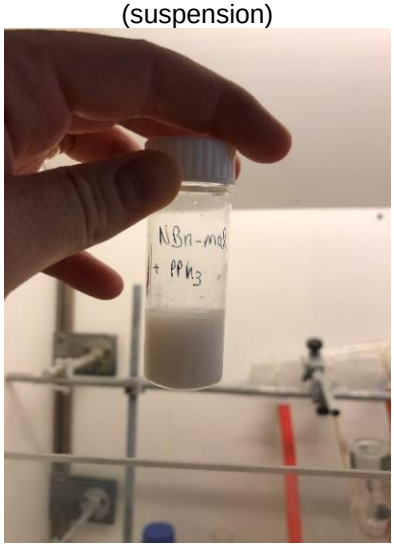
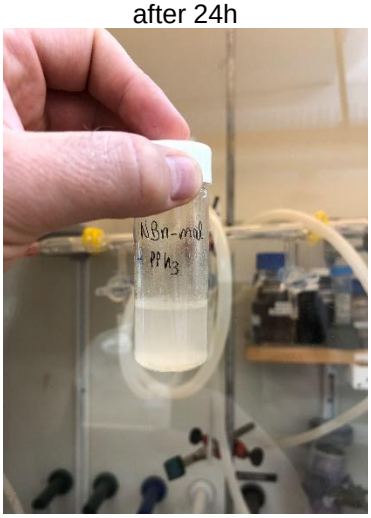


Image S4
Maleimide, PPh_3 and NEt_3
in PBS (20mM, pH7.4)/DMSO 9:1 (6mM)

after 24h



2.2.2 TCEP experiment

A 20mL vial was charged with tris(2-carboxyethyl)phosphine hydrochloride (17.2mg, 59.8µmol, 1.0eq.) and to another identical vial was added *N*-benzylmaleimide (11.2mg, 59.8µM, 1.0eq.). Phosphate buffered saline (20mM, pH7.4, 1mL) was added to the phosphine-containing vial, and DMSO (1mL) was added to the maleimide. Both vials yielded solutions almost immediately (see image **S5**). Triethylamine (0.8µL, 6µmol, 0.1eq.) was added to the TCEP-solution, which was thereupon added dropwise to the maleimide solution. No precipitation occurred at any point (see image **S6**). The reaction mixture was diluted with more phosphate buffered saline (8mL) and remained homogeneous (see image **S7**), even after 24 hours (see image **S8**).

Another image illustrating the pronounced aqueous solubility was taken from the preparative scale synthesis of ADAPT=**14** (see image S9, next page).

Image S5
Maleimide (60mM)
in DMSO (right)
TCEP x HCl (60mM)
in PBS (20mM, pH=7.4)
(left)
(clear solutions)

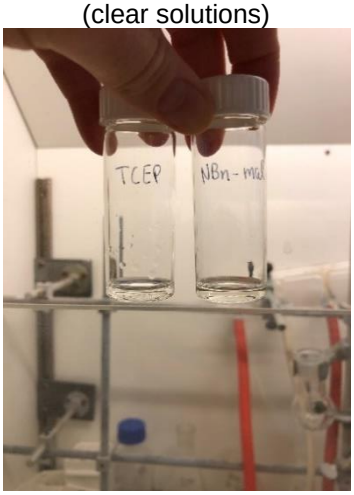


Image S6
Maleimide, TCEP x HCl
and NEt_3
in PBS (20mM, pH7.4)/
DMSO (30mM), 1:1
(right)
(clear solution)

Image S7
Maleimide, TCEP x HCl
and NEt_3
in PBS (20mM, pH7.4)/
DMSO (6mM), 9:1
(right)
(clear solution)

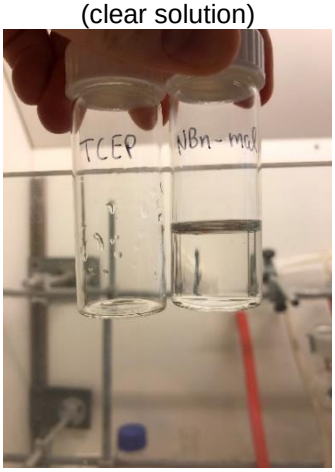
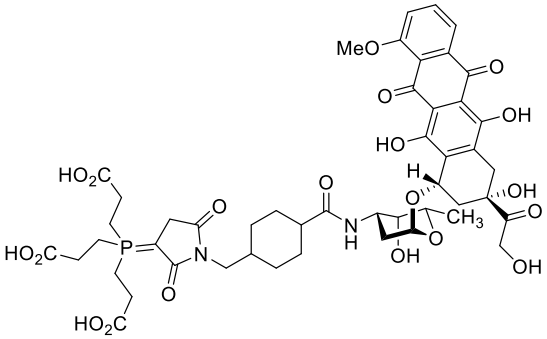


Image S8
Maleimide, TCEP x HCl
and NEt_3
in PBS (20mM, pH7.4)/
DMSO (6mM), 9:1

After 24h

Image S9



c=3mM
(phosphonium ylide derived from maleimide 14)
in PBS (10mM, pH7.4)/DMSO 9:1 + ADAPT-CHO (30µM)

(clear solution)

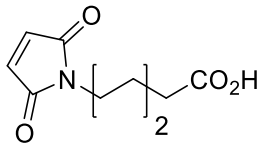
See synthesis on page 77

2.3 Acceleratory effect of NEt_3 on the formation of 6-maleimidocaproic acid (7) derived phosphonium ylide

2.3.1 Reaction setup

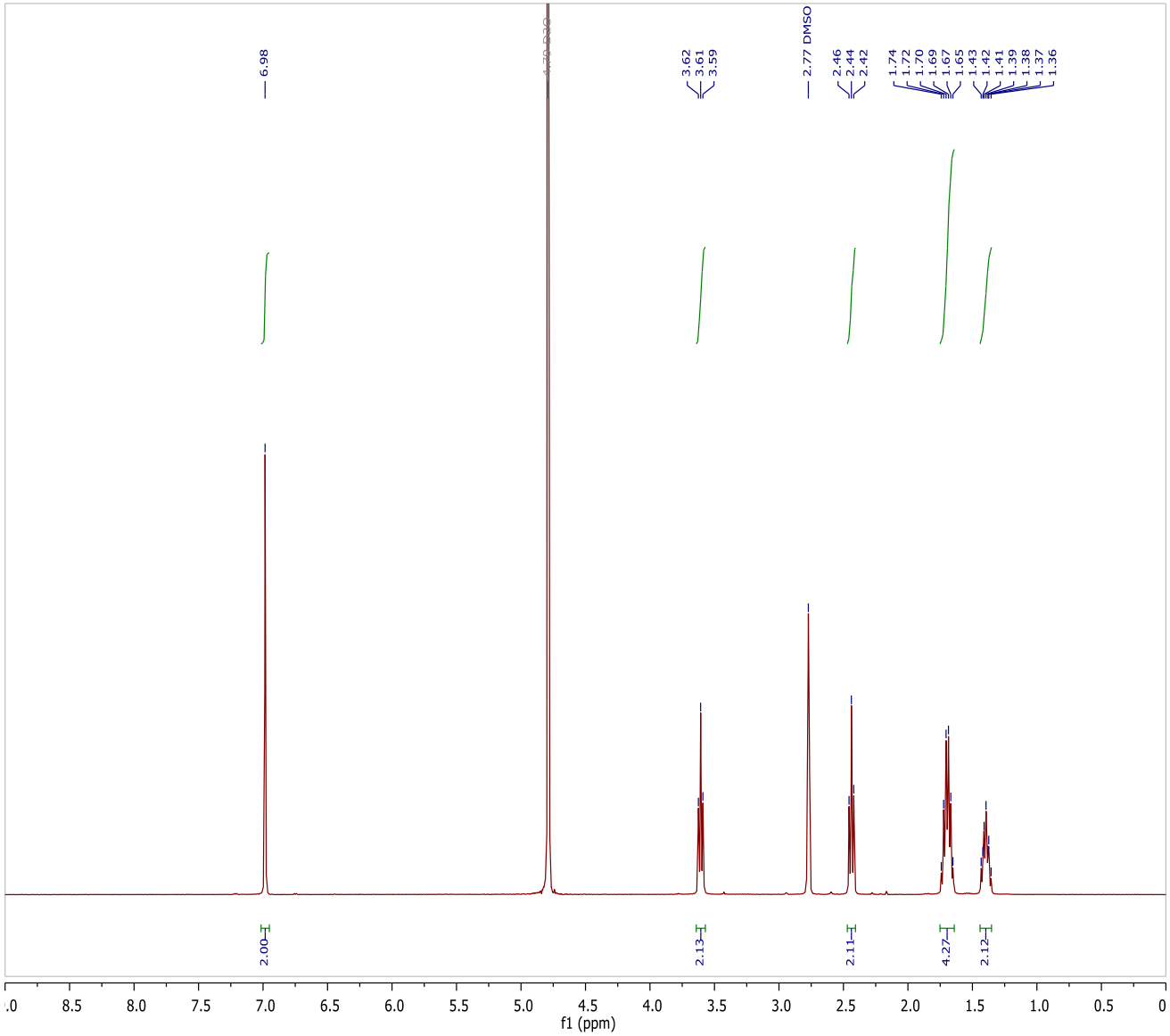
Two vials were each charged with tris(2-carboxyethyl)phosphine hydrochloride (27mg, 95µmol, 1.0eq) and 6-maleimidocaproic acid (20mg, 95µmol, 1.0eq). DMSO-d_6 (400µL) was added to each vial, followed by D_2O (2.6mL) and phosphate buffered saline in D_2O (200mM, pH=7.4, 1.0mL). Triethylamine (1.3µL, 9.5µmol, 0.1eq.) was added to one of the vials and they were both shaken briefly to cause complete dissolution of the analytes. ^1H NMR analysis revealed to complete consumption of the maleimide (δ=6.85ppm relative to H_2O at δ=4.79ppm) in the vial into which triethylamine had been added. In its absence, approximately 21% remained after 1h, 16% remained after 2h, and 8% remained after 6h. A spectrum of the maleimide in D_2O/DMSO-d_6 was recorded for reference (see all five spectra on the following pages).

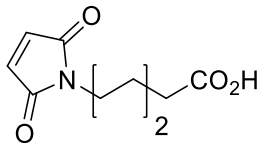
2.3.2 NMR spectra



^{1}H NMR (400MHz)
D_2O/DMSO-d_6 (9:1)
Starting material reference

The signal at δ=6.98ppm belongs to the two maleimide protons

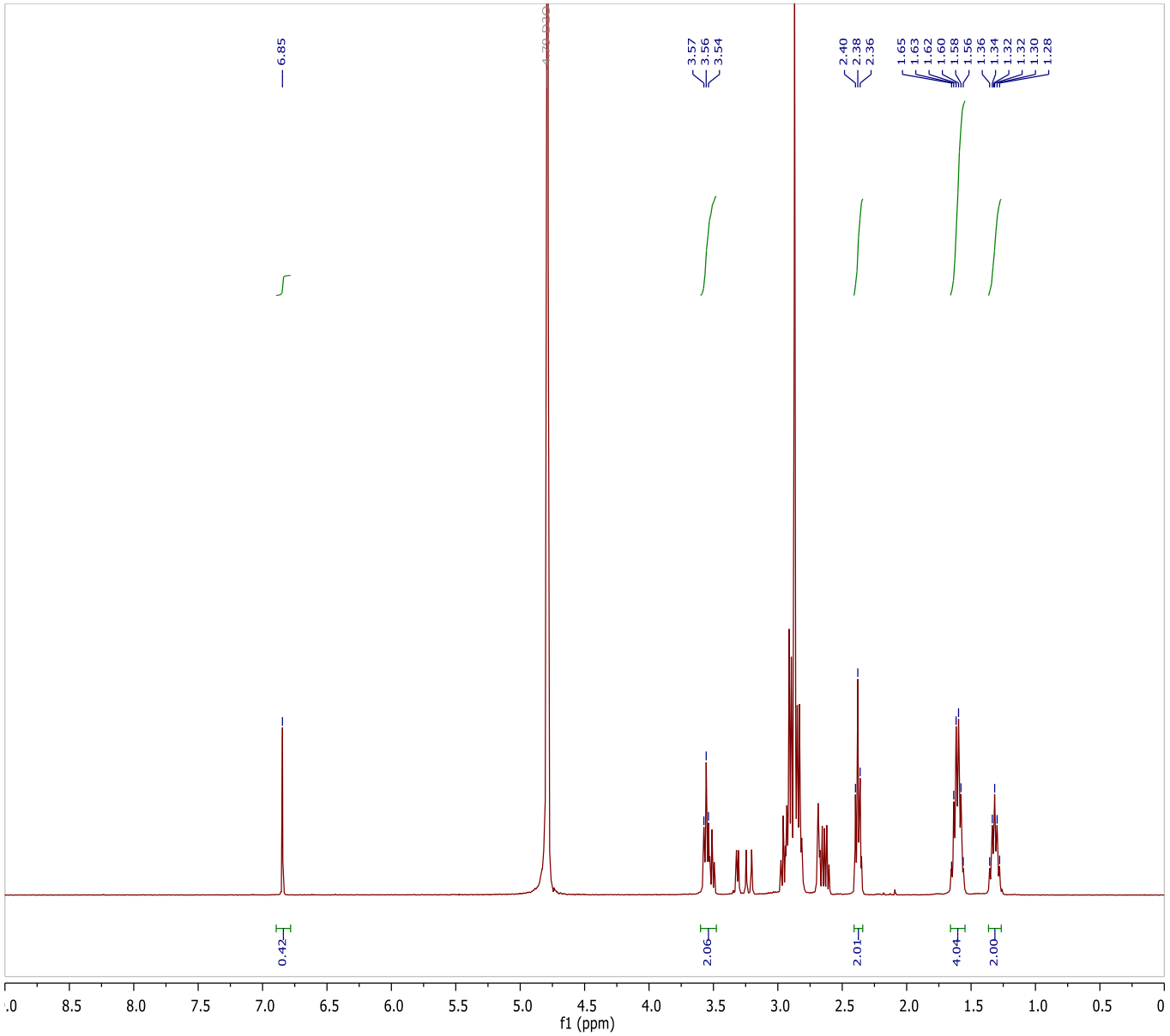


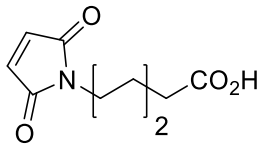


+TCEP (1eq), no NEt_3
after **1h** at room temperature

^1H NMR (400MHz)
PBS (50mM) in D_2O/
DMSO-d_6 (9:1)

The integral of the signal at δ=6.85ppm indicates that approximately 21% (=0.42/2) of maleimide remained

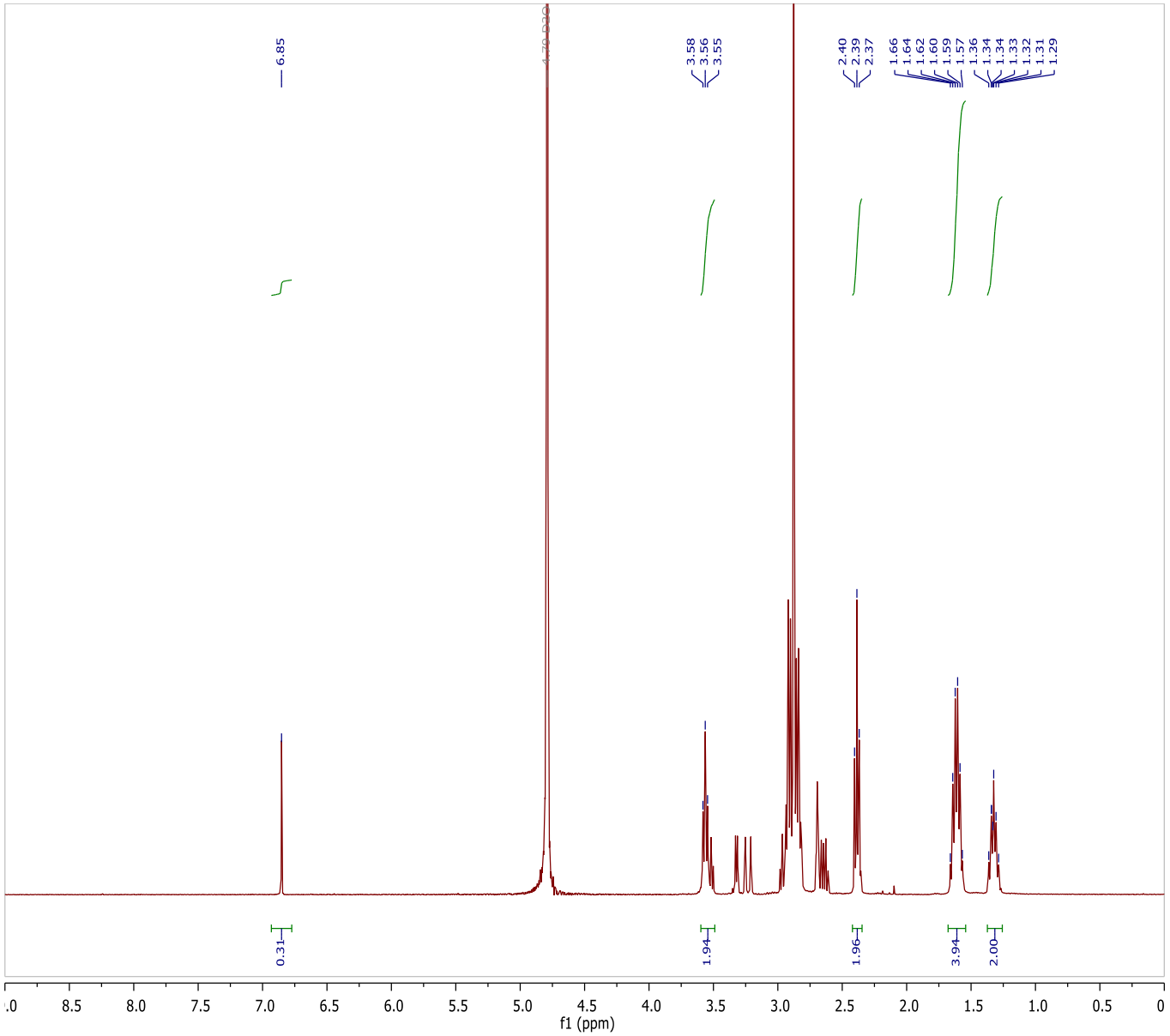


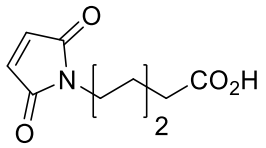


+TCEP (1eq), no NEt_3
after **2h** at room temperature

^1H NMR (400MHz)
PBS (50mM) in D_2O/
DMSO-d_6 (9:1)

The integral of the signal at δ=6.85ppm indicates that approximately 16% (=0.31/2) of maleimide remained

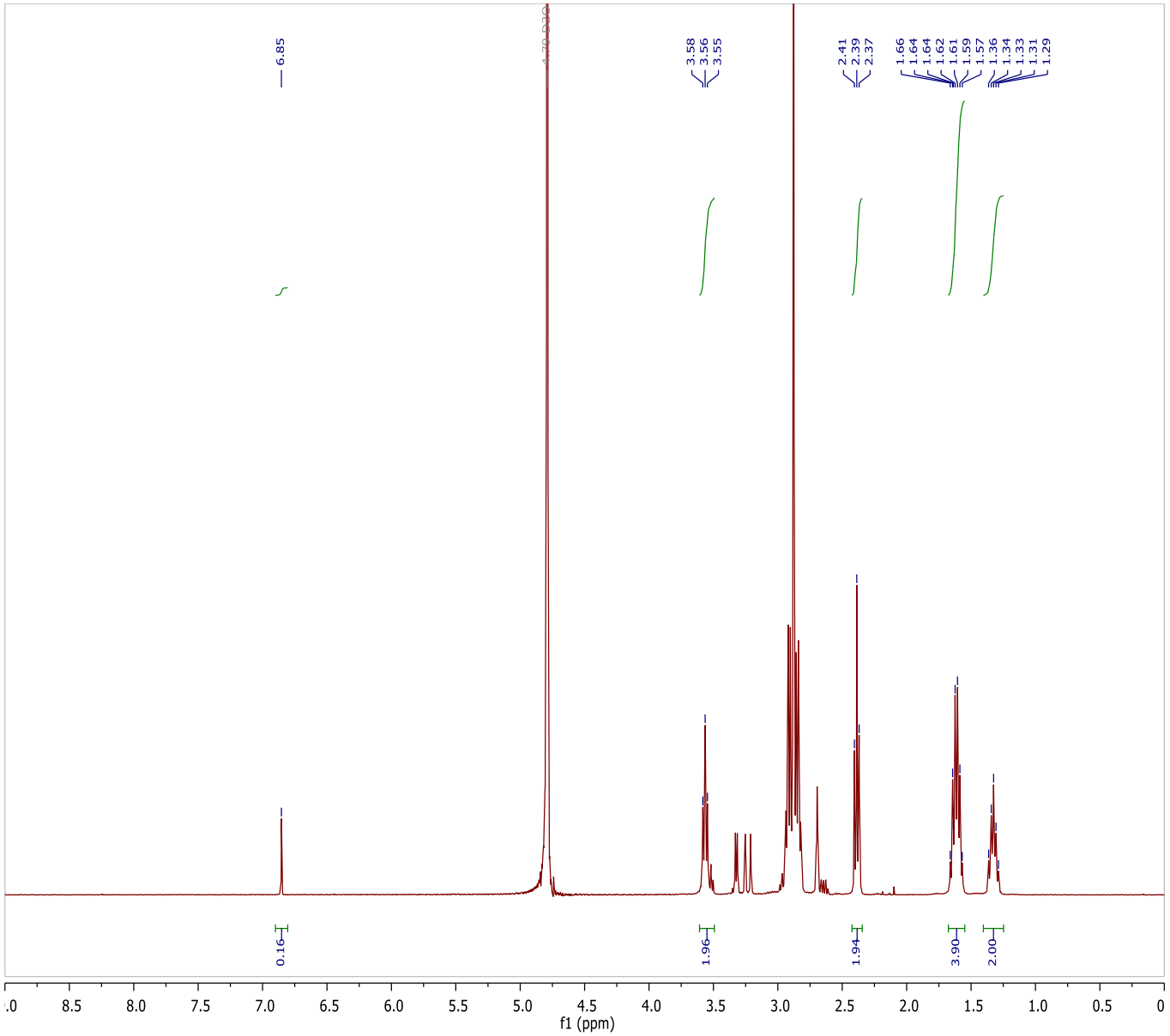


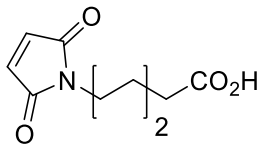


+TCEP (1eq), no NEt_3
after **6h** at room temperature

^1H NMR (400MHz)
PBS (50mM) in D_2O/
DMSO-d_6 (9:1)

The integral of the signal at δ=6.85ppm indicates that approximately 8% (0.16/2) of maleimide remained



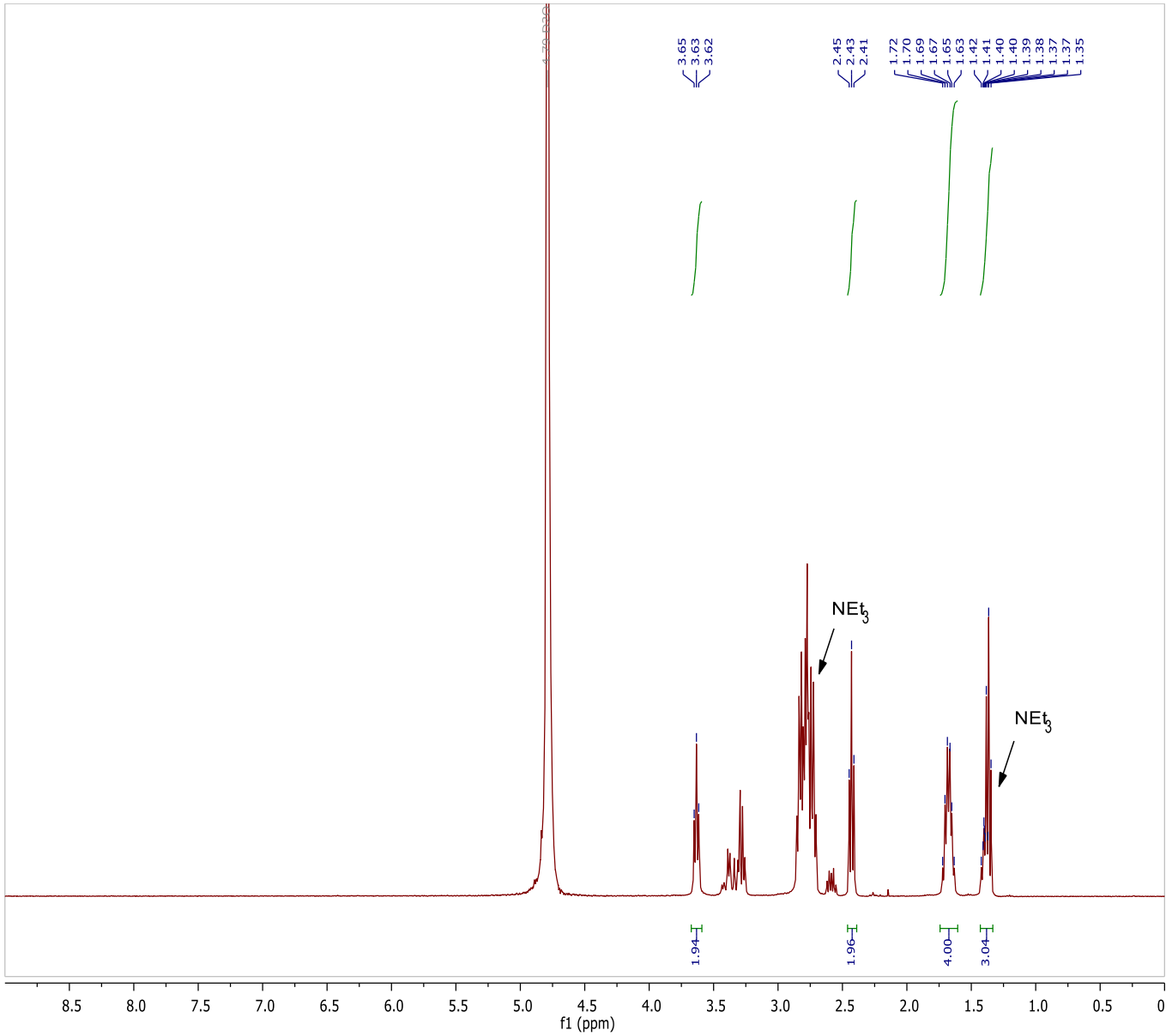


+TCEP (1eq)
+NEt_3 (0.1eq)
after **5min** at room temperature

^1H NMR (400MHz)
PBS (50mM) in D_2O/
DMSO-d_6 (9:1)

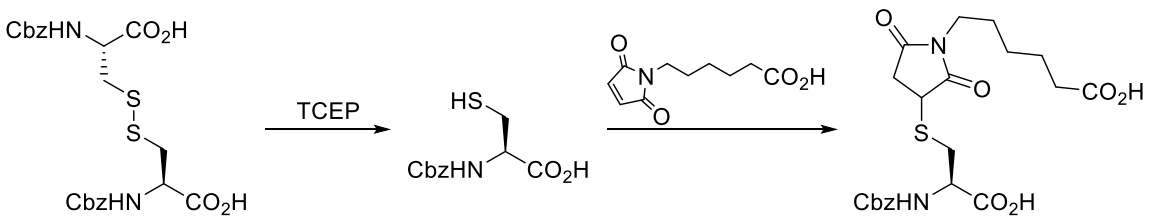
The absence of a signal at δ=6.85ppm indicates full conversion of the starting material

(Note that the multiplet at δ=1.42-1.34ppm overlaps with one of triethylamine's signals)



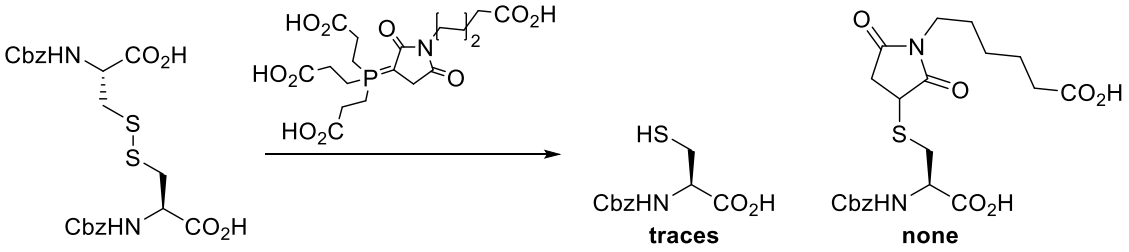
3. TCEP-phosphonium ylide tolerance of disulfides

3.1 Sequential reduction and alkylation of cysteine dimer (positive control experiment)[2]



A 100mL roundbottom flask was charged with (2*R*,2'*R*)-3,3'-disulfanediylbis(2-(((benzyloxy)carbonyl) amino)propanoic acid) (Cbz-Cys dimer, 25.4mg, 50.0µmol, 1.0eq) and a stir bar. DMSO (2.5mL) and phosphate buffered saline (50mM, pH=7.4, 22.5mL) were added and the mixture was stirred until a solution had been obtained, which was subsequently analyzed by LCMS (see image **S10**). Added to the solution was tris(2-carboxyethyl)phosphine hydrochloride (43.0mg, 150µmol, 3.0eq) and LCMS analysis after 15 minutes revealed complete disulfide reduction (see image **S11**). Lastly, 6-maleimidocaproic acid (31.6mg, 150µmol, 3.0eq) was added to the solution, which was stirred at room temperature for 24 hours. LCMS analysis after that time revealed two predominant compounds: The expected maleimido thioether along with the 6-maleimidocaproic acid derived phosphonium ylide (see image **S12**).

3.2 Treatment of cysteine dimer with TCE-phosphonium ylide[2]



A 100mL roundbottom flask was charged with tris(2-carboxyethyl)phosphine hydrochloride (42.9mg, 150µmol, 3.0eq), 6-maleimidocaproic acid (30.5mg, 150µmol, 3.0eq), and a stirbar. DMSO (2.5mL) and phosphate buffered saline (50mM, pH=7.4, 2.5mL) were added, the mixture was stirred until homogenous, and then triethylamine (2.0µL, 1.5µmol, 3.0eq) was added. LCMS analysis after 10 minutes revealed the complete formation of the 6-maleimidocaproic acid derived phosphonium ylide (see image **S13**). (2*R*,2'*R*)-3,3'-disulfanediylbis(2-(((benzyloxy)carbonyl) amino)propanoic acid) (Cbz-Cys dimer, 25.3mg, 150µmol, 1.0eq) was added to the reaction mixture, it was ultrasonified until homogenous, and heated at 37°C for 24 hours. LCMS analysis after that time revealed two predominant compounds: Unreacted Cbz-cysteine dimer along with the intact 6-maleimidocaproic acid derived phosphonium ylide. Small amounts (<5%) of reduced disulfide had also formed (see image **S14**).

These experiments demonstrate that while small amounts of disulfide reduction may occur in the presence of TCE-phosphonium ylides, no detectable amount of subsequent alkylation occurs. Please refer to the **next pages** for all **chromatograms** referenced in the paragraphs above.

[2] Note: In both experiments, the concentration of Cbz-cysteine dimer was chosen to be 2mM in order to mimic the conditions used for the Wittig couplings with small molecule substrates (see section 5).

3.3 Reaction monitoring by LC-MS

Sequential reduction, followed by alkylation:

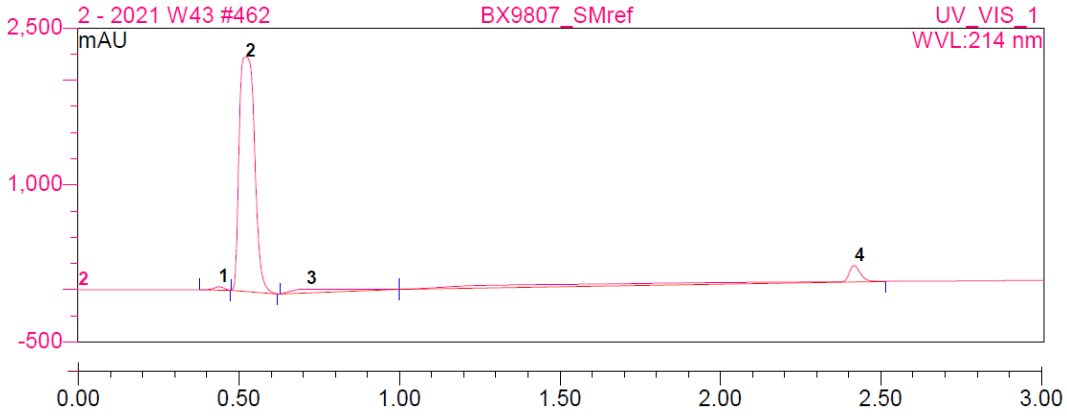


Image **S10**

Cbz-Cys dimer (m/z=531, $[M+Na]^+$, 509 $[M+H]^+$, R_t=2.41min) in PBS/DMSO (R_t=0.52min)

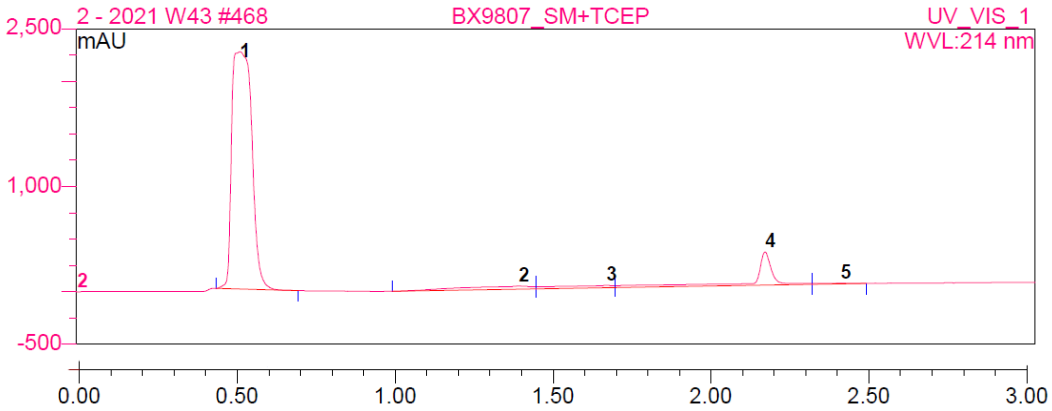


Image **S11**

Cbz-Cys monomer (R_t=2.16min) obtained by TCEP-mediated reduction of dimer

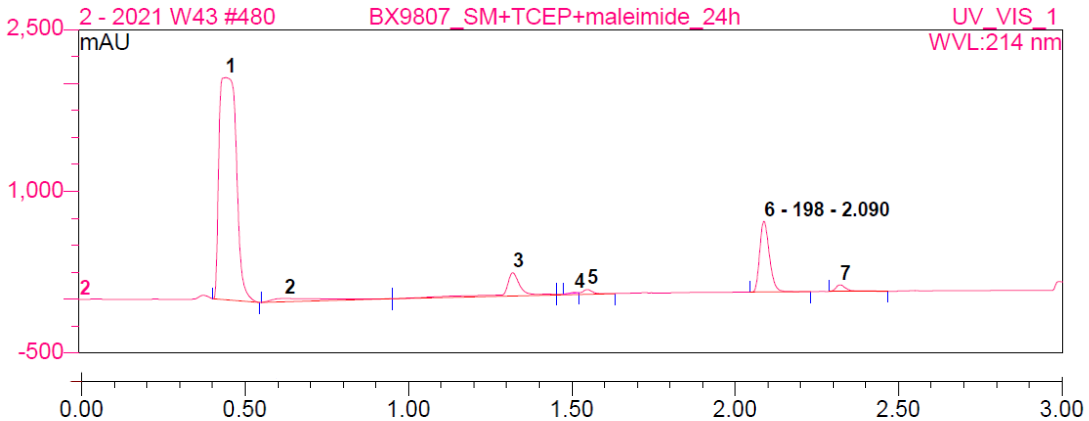
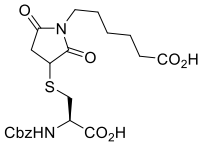
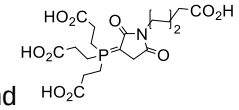


Image **S12**



(m/z=467 $[M+H]^+$, 489 $[M+Na]^+$, R_t=2.09min) and



(m/z=462 $[M+H]^+$, R_t=1.33min)

Treatment of disulfide with phosphonium ylide:

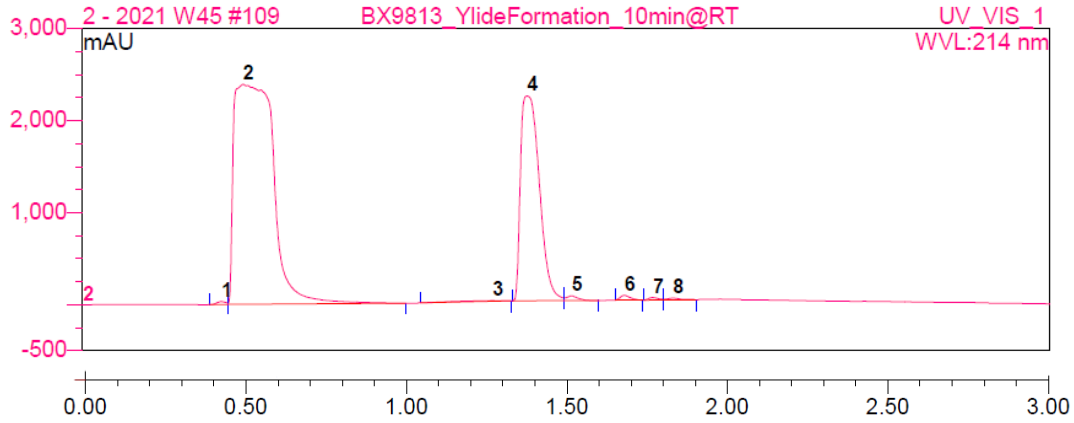
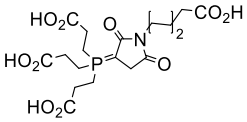


Image **S13**



(m/z=462 $[M+H]^+$, R_t=1.40min) in PBS/DMSO (R_t=0.50 min)

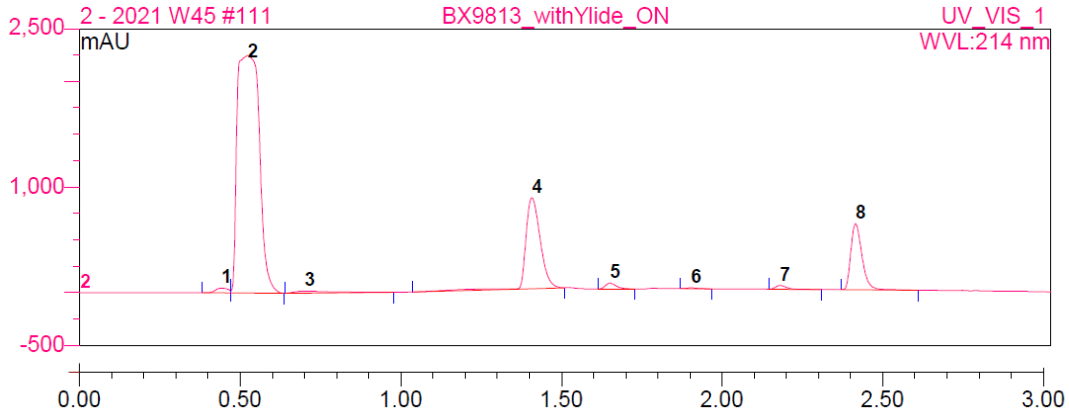
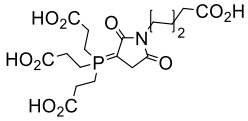


Image **S14**

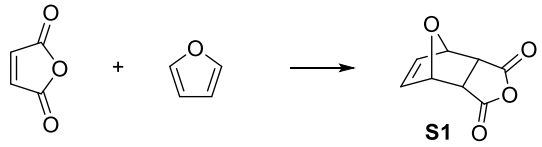


(m/z=462 $[M+H]^+$, R_t=1.40min), Cbz-Cys dimer (m/z=531, $[M+Na]^+$, 509 $[M+H]^+$, R_t=2.40min) and trace amounts of presumed Cbz-Cys monomer (R_t=2.19min)

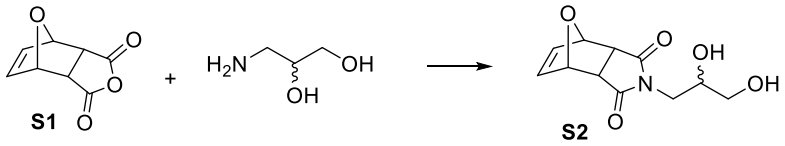
4. Synthesis of small molecule intermediates

4.1 1-(2,3-dihydroxypropyl)-1H-pyrrole-2,5-dione 8

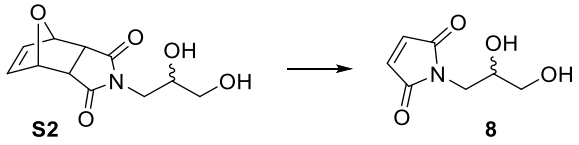
4.1.1 Synthesis



Exo-Diels-Alder adduct **S1**. A 100mL round bottom flask was charged with maleic anhydride (10.1g, 103mmol, 1.0eq.) and a stir bar. Furan (30mL) was added and the resulting suspension was stirred vigorously. The mixture solidified overnight. Diethyl ether was added, the mixture was ultrasonified, the volatiles were removed *in vacuo*. The solid was transferred to a beaker and dried under fine vacuum for an additional three days to obtain **S1** as a white solid (15.0g, 90.2mmol, 87% yield). The compound's spectroscopic data were in good agreement with literature.[3]



Imide **S2**. To a 100mL round bottom flask charged with **S1** (15.0g, 90.2mmol, 1.0eq.), ethanol (24mL), a solution of (±)-3-amino-1,2-propanediol (8.23g, 90.3mmol, 1.0eq.) in ethanol (6mL) and a stir bar were added and the resulting suspension was refluxed overnight. After 20 hours of stirring at reflux, the now homogenous, orange solution was allowed to attain room temperature while continuing to stir and was then cooled on ice. After one hour, the reaction mixture was filtered, and the filtrate was washed with ice-cold ethanol (2 x 40mL) and diethyl ether (40mL). The solid was transferred to a beaker and dried under fine vacuum for an additional two days to obtain **S2** as a white solid (9.40g, 39.3mmol, 43% yield). The compound's spectroscopic data were in good agreement with literature.[3]

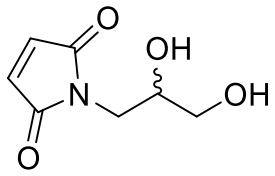


1-(2,3-dihydroxypropyl)-1H-pyrrole-2,5-dione (**8**). A 100mL round bottom flask was charged with **S2** (4.97g, 20.8mmol, 1.0eq.) and a stir bar. Toluene (50mL) was added, a reflux condenser was attached, and with no flow of cooling water, the flask was heated to reflux to allow for the furan to easily evaporate. After 18 hours of stirring, a two phasic mixture consisting of two liquid layers was obtained. While still heating the flask, the liquids were allowed to separate, the top layer was decanted off, and more toluene (50mL) was added. The contents of the reaction flask were stirred for ten minutes, and the top layer was decanted once again. The process of adding toluene (50mL) and extracting the denser of the two phases was repeated four more times. The toluene layers were then pooled, about half of the solvent was removed *in vacuo*, and the remaining solution was stored in the freezer overnight to cause the precipitation of the product. The crystals were crushed, transferred to a beaker, and dried under fine vacuum for four days to obtain **8** as a crystalline, white solid (1.46g, 8.53mmol, 41% yield). The compound's spectroscopic data were in good agreement with literature.[3]

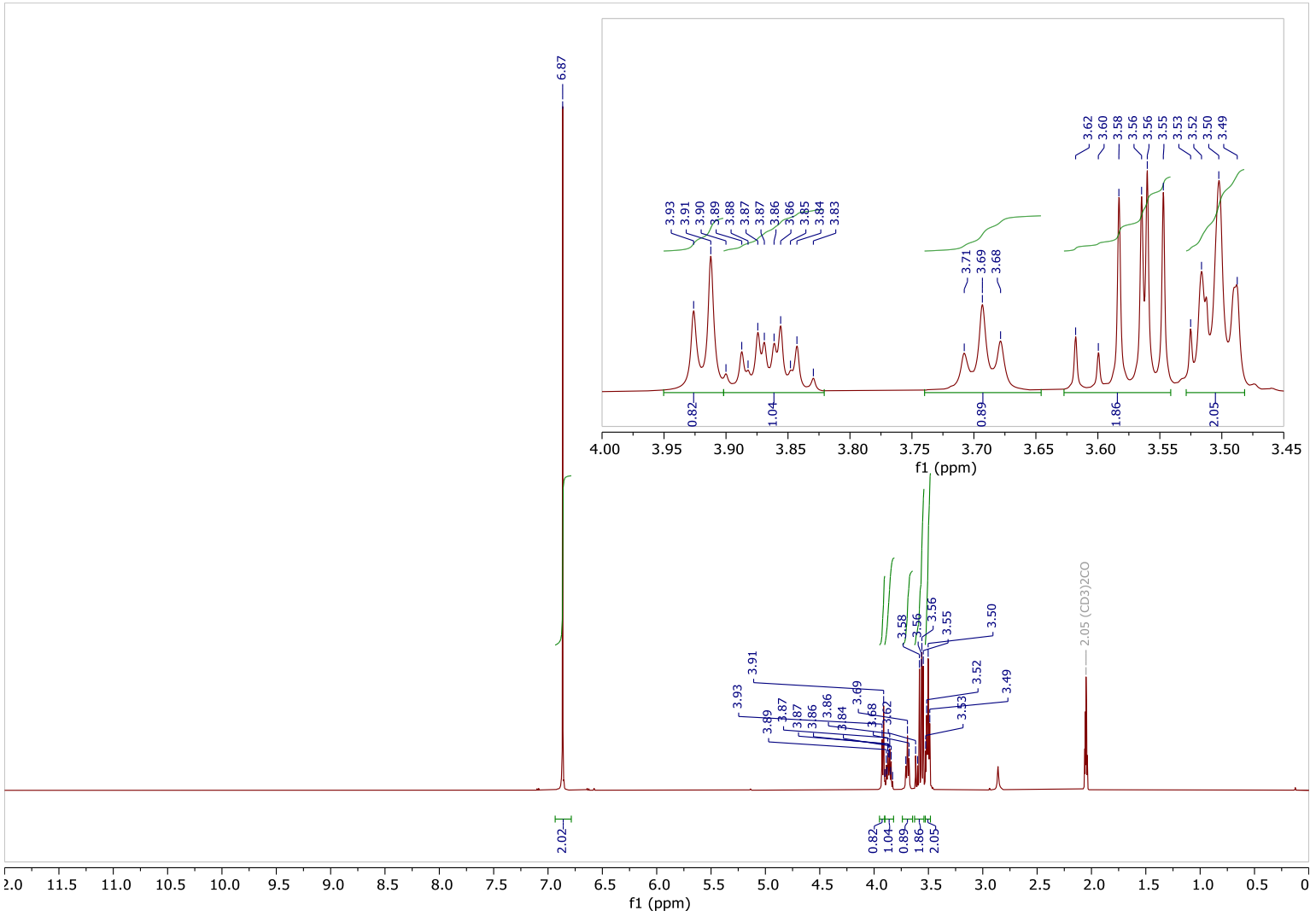
^{1}H NMR (400 MHz, acetone-d_6) δ 6.87 (s, 2H), 3.92 (d, J = 5.5 Hz, 1H), 3.87 (dp, J = 7.1, 5.5 Hz, 1H), 3.69 (t, J = 5.9 Hz, 1H), 3.62 – 3.54 (m, 2H), 3.53 – 3.48 (m, 2H). **^{13}C NMR** (101 MHz, acetone-d_6) δ 171.9, 135.2, 70.2, 65.1, 41.7. **HRMS** calculated for $C_7H_{10}NO_4$ $[M+H]^+$, 172.0610; found, 172.0607.

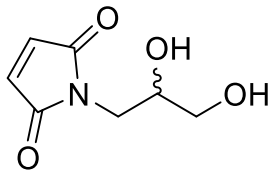
[3] Lu, X., et. al., *J. Mater. Chem. A*, **2014**, *2*, 16051-60.

4.1.2 NMR spectra

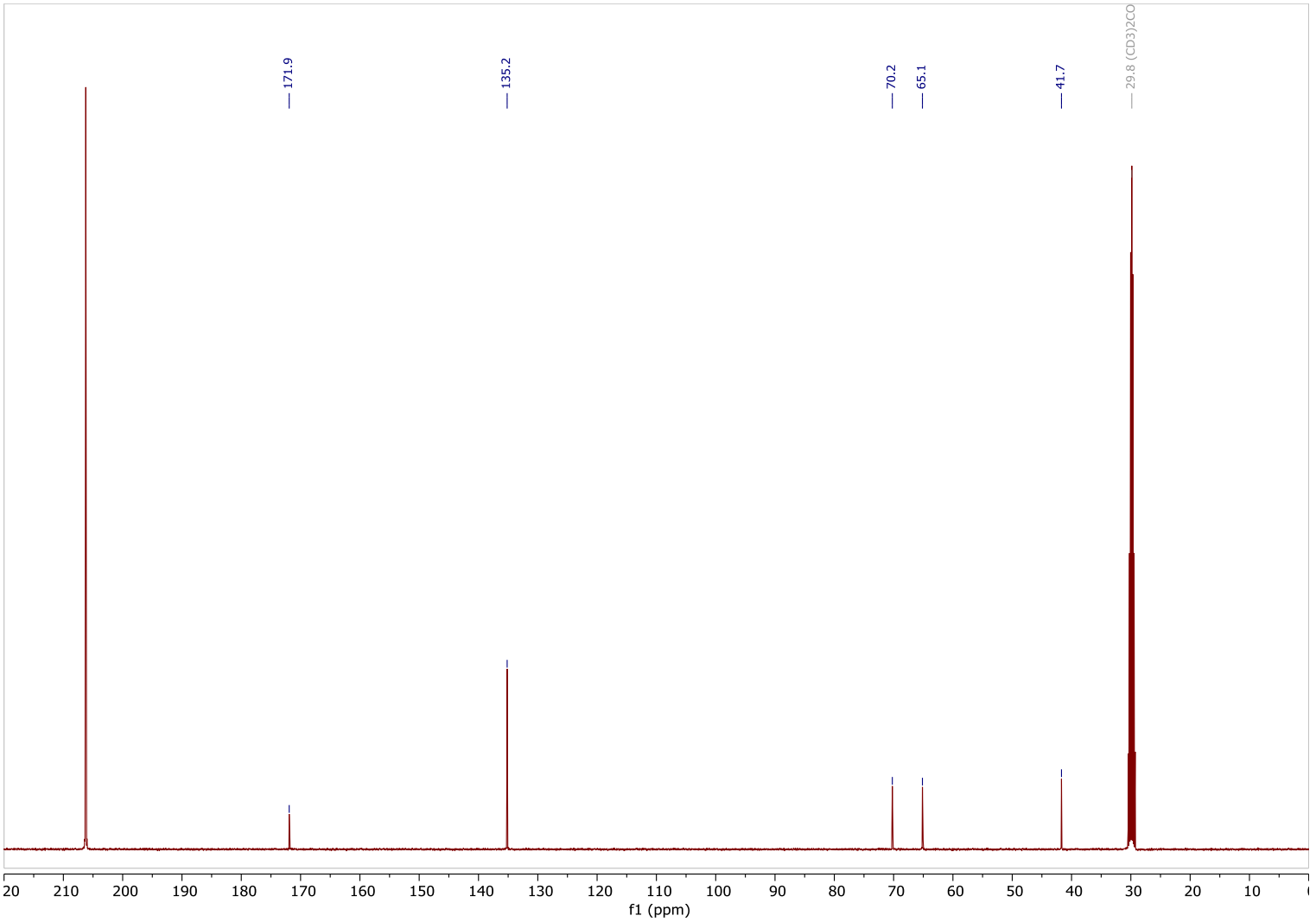


^{1}H NMR (400MHz), acetone-d_6



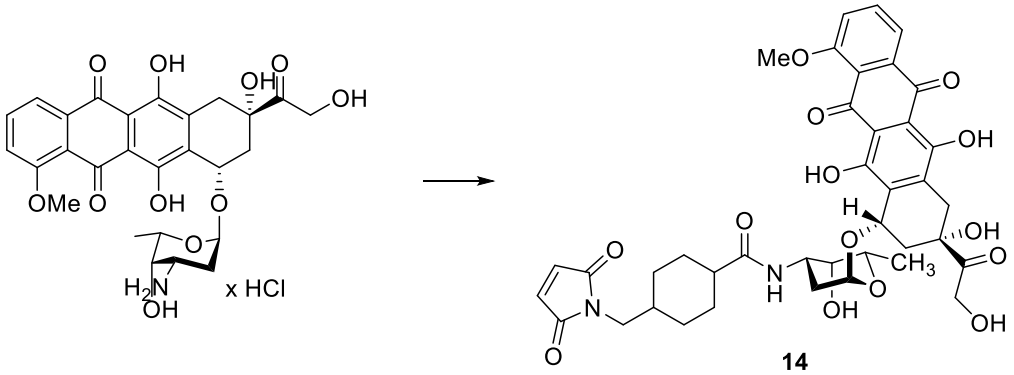


^{13}C NMR (101MHz), acetone-d_6



4.2 SMCC-derivatized doxorubicin 14

4.2.1 Synthesis

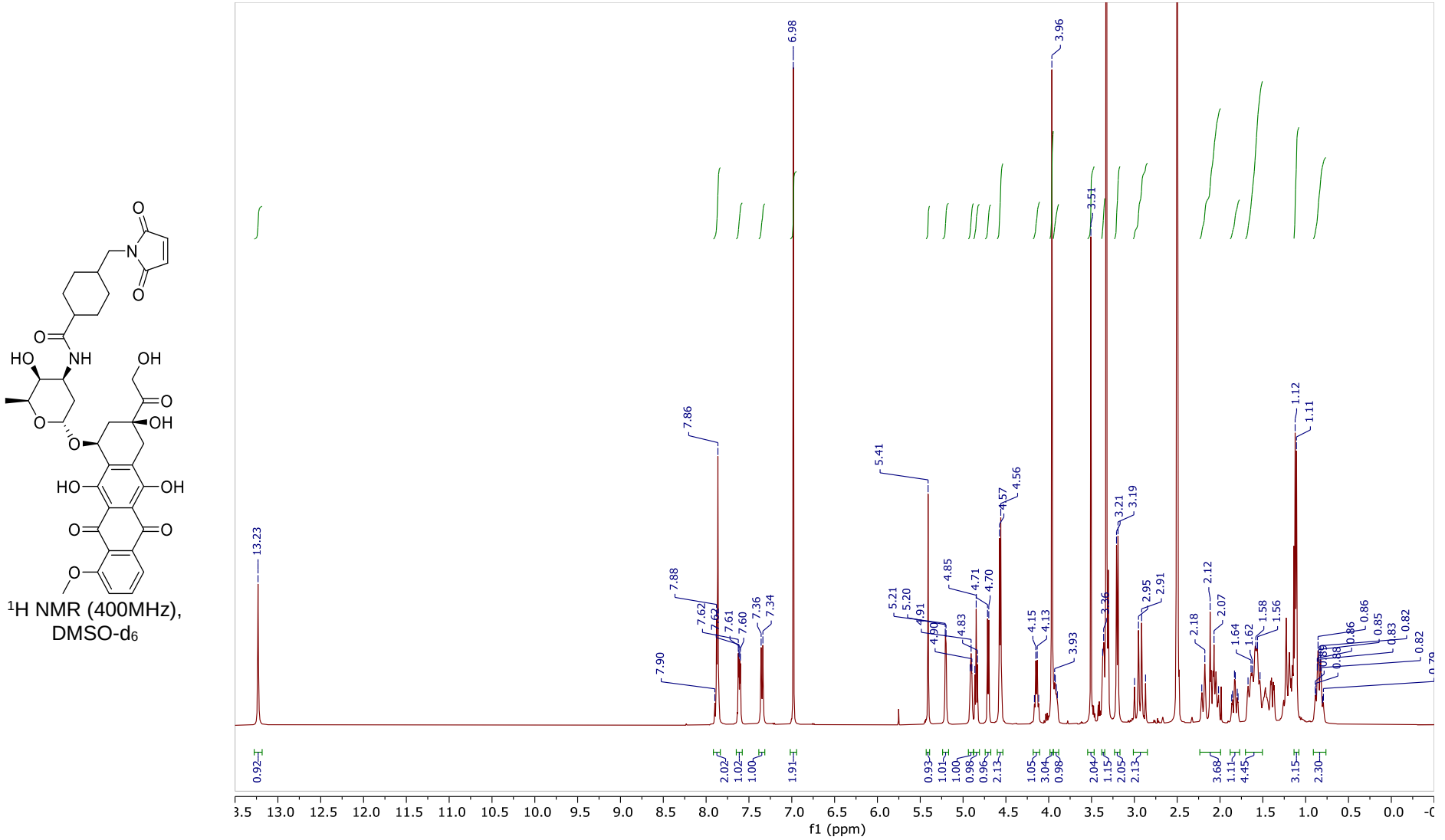


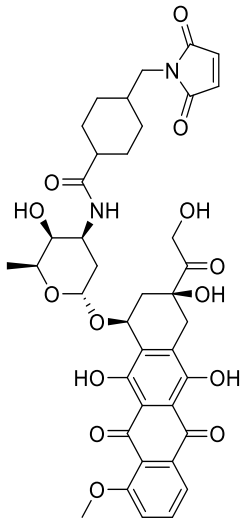
Doxorubicin SMCC-amide **14**. A 25mL round bottom flask was charged with doxorubicin hydrochloride (30mg, 52µmol, 1.0eq.), *N*-succinimidyl 4-(maleimidomethyl) cyclohexanecarboxylate (21mg, 62µmol, 1.2eq.), and a stir bar. The outside of the flask was covered with aluminum foil. Dimethylformamide (4mL) and diisopropylethyl amine (14µL, 78µmol, 1.5eq.) were added, and the resulting solution was stirred at room temperature overnight. The volatiles were removed *in vacuo*, the residue was taken up in dichloromethane (50mL) and washed with brine (2 x 25mL). The organic layer was then dried over Na_2SO_4. The solids were filtered off and washed with dichloromethane twice more (25mL each). The pooled organic layers were concentrated *in vacuo* and purified by flash chromatography on silica (MeOH/DCM 0:100 → MeOH/DCM 5:95%) to yield **14** as a red solid (24mg, 32µmol, 61%). The compound's spectroscopic data were in good agreement with literature.[4]

^{1}H NMR (400 MHz, DMSO-d_6) δ 13.99 (s, 1H), 13.23 (s, 1H), 7.91 – 7.84 (m, 2H), 7.61 (dd, *J* = 6.7, 3.0 Hz, 1H), 7.35 (d, *J* = 8.2 Hz, 1H), 6.98 (s, 2H), 5.41 (s, 1H), 5.20 (d, *J* = 3.6 Hz, 1H), 4.91 (m, w=8.3Hz, 1H), 4.85 (t, *J* = 5.9 Hz, 1H), 4.71 (d, *J* = 5.9 Hz, 1H), 4.57 (d, *J* = 6.0 Hz, 2H), 4.14 (q, *J* = 6.6 Hz, 1H), 3.98 – 3.88 (m, 4H), 3.51 (s, 2H), 3.36 (m, w=12.3Hz, 1H), 3.20 (d, *J* = 7.0 Hz, 2H), 3.01 – 2.85 (m, 2H), 2.24 – 2.00 (m, 3H), 1.89 - 1.77 (m, 1H), 1.72-1.51 (m, 4H), 1.50 – 1.35 (m, 2H), 1.11 (d, *J* = 6.4 Hz, 3H), 0.92 – 0.77 (m, 2H). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 213.8, 186.5, 186.4, 174.3, 171.3, 160.8, 156.1, 154.5, 136.2, 135.5, 134.6, 134.3, 134.0, 120.0, 119.7, 119.0, 110.8, 110.6, 100.5, 75.0, 70.0, 69.9, 68.1, 66.8, 63.7, 56.6, 44.7, 43.5, 43.1, 36.6, 36.1, 32.1, 29.7, 29.4, 28.6, 28.4, 17.0. **HRMS** calculated for $C_{39}H_{41}N_2O_{14}$ $[M-H]^-$, 761.2558; found, 761.2562.

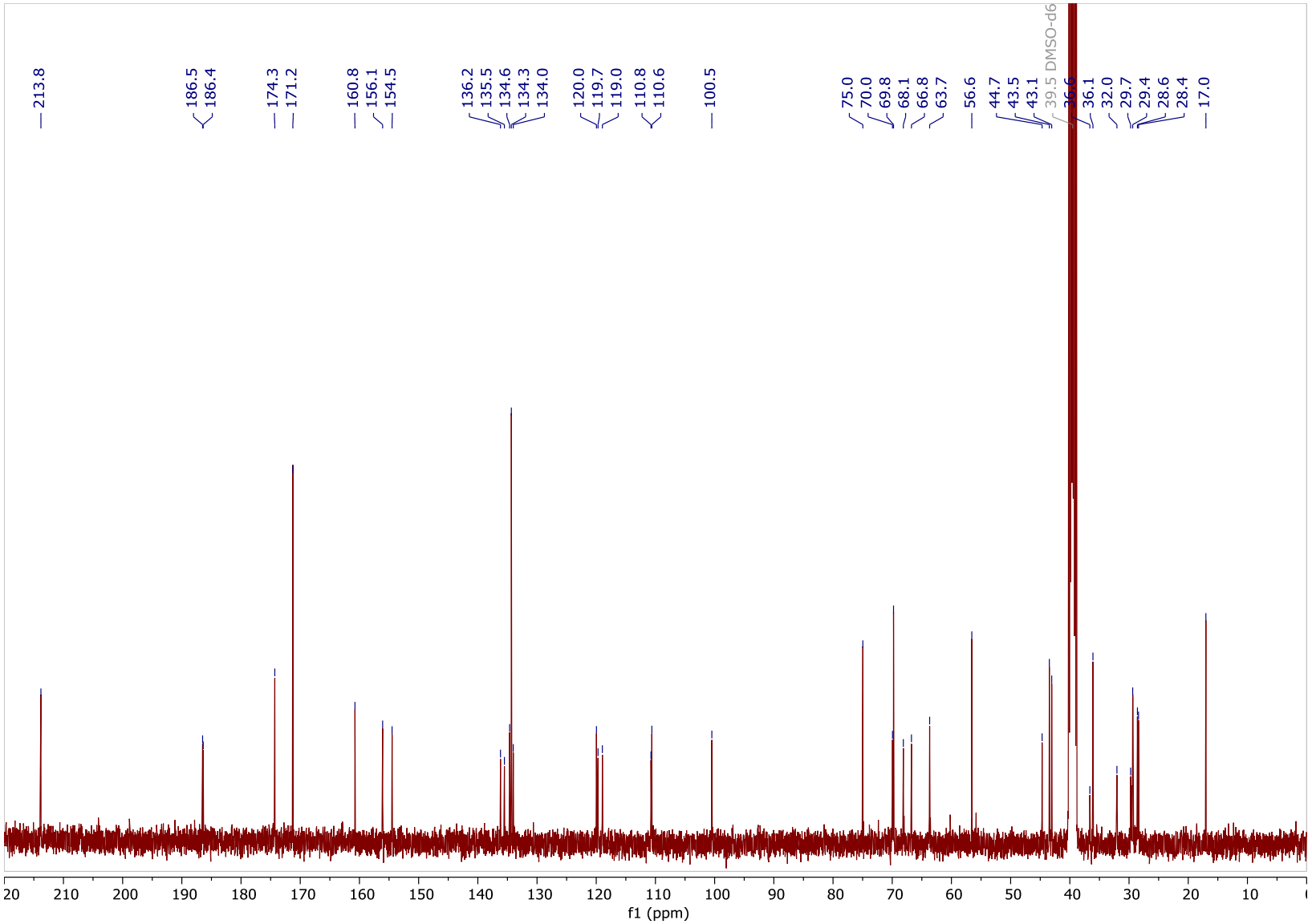
[4] Chen, Q., et al., *Synth. Comm.*, **2004**, *34*, 2407-14.

4.2.2 NMR spectra



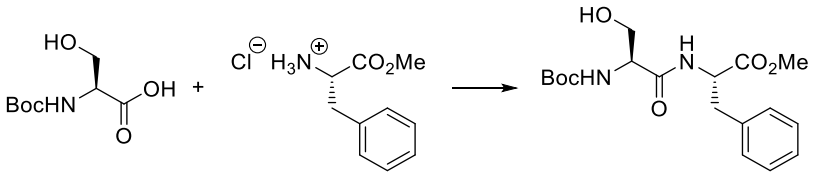


^{13}C NMR (101MHz), DMSO-d_6



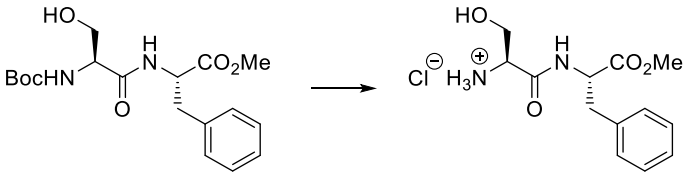
4.3 Dipeptide derived methyl-hemiacetal 5

4.3.1 Synthesis



S3

Methyl (*tert*-butoxycarbonyl)-L-seryl-L-phenylalaninate (**S3**). A 100mL round bottom flask was charged with *N*-(*tert*-butoxycarbonyl)-L-serine (2.0g, 9.8mmol, 1.0eq.), methyl L-phenylalaninate hydrochloride (2.7g, 13mmol, 1.3eq.) and a stir bar. Dichloromethane (50mL) was added to yield a suspension. Di*iso*propylethyl amine (5.1mL, 29mmol, 3.0eq.) was added, and the resulting solution was cooled on ice. 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU, 4.6g, 13mmol, 1.3eq.) was added, and the reaction mixture was stirred on ice for one hour and then at room temperature for 17 hours. The reaction mixture was transferred to a separatory funnel, washed with saturated aqueous ammonium chloride (50mL) and brine (2x25mL), and the organic layer was dried over Na_2SO_4. After removal of the volatiles *in vacuo*, the resulting oil was purified by flash chromatography on silica (ethyl acetate: petroleum ether 2:3→9:1) to yield **S1** as a colorless, viscous oil (2.5g, 6.9mmol, 71% yield). The compound's spectroscopic data were in good agreement with literature.[5]

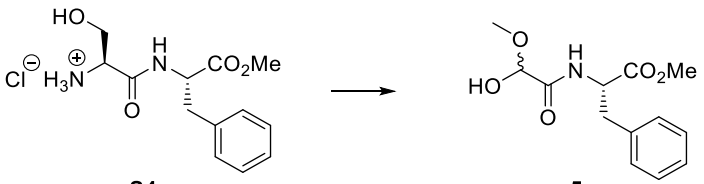


S3 **S4**

(*S*)-3-hydroxy-1-(((*S*)-1-methoxy-1-oxo-3-phenylpropan-2-yl)amino)-1-oxopropan-2-aminium chloride (**S4**). Methyl (*tert*-butoxycarbonyl)-L-seryl-L-phenylalaninate (**S1**, 150mg, 410µmol, 1.0eq.) was weighed in to a 25mL round bottom flask. 1,4-dioxane (1mL) and a stir bar were added and the resulting mixture was stirred until a solution had been obtained. The latter was cooled on ice and, immediately so as to avoid freezing, 4M HCl in dioxane was added (3.0mL, 12mmol, 29eq.). The resulting mixture was stirred on ice for one hour, then allowed to attain room temperature, and stirred for another two hours. The volatiles were then thoroughly removed *in vacuo* and the residue was recrystallized from *iso*-propanol to yield **S4** as a crystalline, white solid (340mg, 340µmol, 83% yield). ^{1}H NMR analysis showed trace contamination with *iso*-propanol.

^{1}H NMR (400 MHz, DMSO-d_6) δ 9.07 (d, *J* = 7.5 Hz, 1H), 8.27 (br s, 3H), 7.32-7.19 (m, 5H), 5.55 (br s, 1H), 4.51 (td, *J* = 8.5, 5.6 Hz, 1H), 3.87 – 3.76 (m, 2H), 3.66 (m, w=17.5 Hz, 1H), 3.60 (s, 3H), 3.06 (dd, *J* = 13.8, 5.6 Hz, 1H), 2.98 (dd, *J* = 13.8, 8.5 Hz, 1H). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 171.3, 167.0, 136.9, 129.2, 128.3, 126.7, 60.3, 54.2, 54.1, 52.0, 36.4. **HRMS** calculated for $C_{13}H_{19}N_2O_4$ $[M+H]^+$, 267.1345; found, 267.1340.

[5] *J. Am. Chem. Soc.* **2010**, *132*, 9546–9548

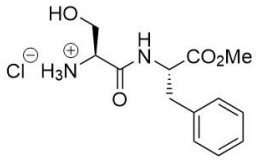


S4 **5**

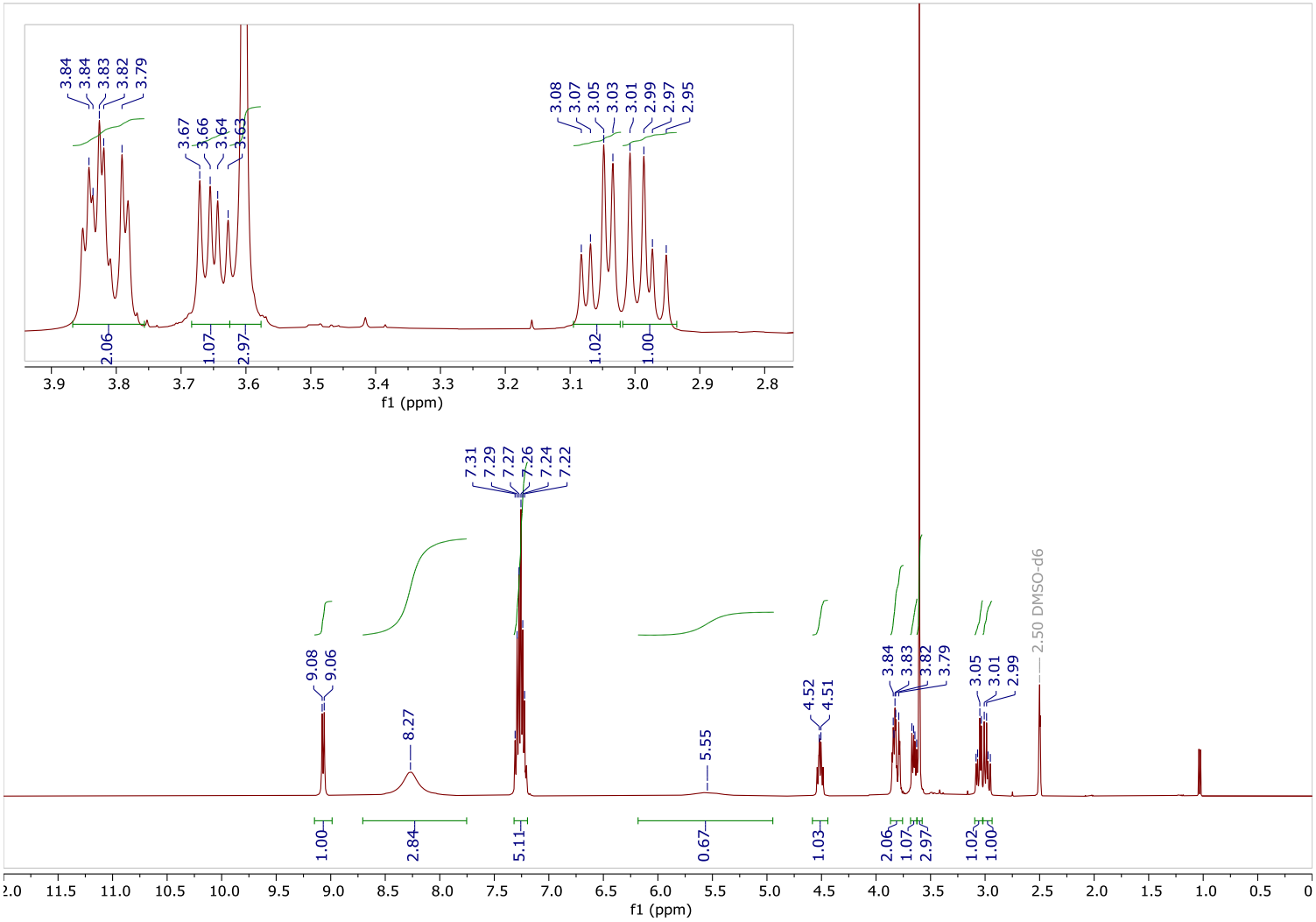
Methyl (2-hydroxy-2-methoxyacetyl)-L-phenylalaninate (**5**). A tainted vial was charged with (*S*)-3-hydroxy-1-(((*S*)-1-methoxy-1-oxo-3-phenylpropan-2-yl)amino)-1-oxopropan-2-aminium chloride (**S4**, 93mg, 310µmol, 1.0 eq.) and a suspension of sodium periodate (130mg, 630µmol, 2.1 eq.) in aqueous phosphate buffered saline (3mL, 50mM phosphate) was added. The resulting mixture was stirred at room temperature for one hour. Then, a solution of sodium sulfite (120mg, 920µmol, 3.0eq.) in aqueous phosphate buffered saline (1mL, 50mM phosphate) was added to quench the residual periodate, and the resulting solution was desalted into methanol (15mL) by means of solid-phase extraction (Oasis HLB 6cc cartridge, Waters, USA). The solvent was removed to yield **5** as a yellow gum (76mg, 290µmol, 93% yield).

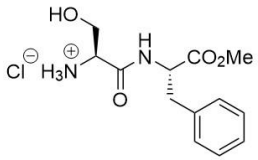
^{1}H NMR (400 MHz, CD_3OD), mixture of diastereomers δ 7.33 – 7.15 (m, 5H), 4.81 and 4.80 (s and s, 1H), 4.72 and 4.70 (t and t, J = 7.9 Hz, 1H), 3.71 and 3.70 (s and s, 3H), 3.35 (s, 3H), 3.22 – 3.15 (m, 1H), 3.10 – 3.02 (m, 1H). **^{13}C NMR** (101 MHz, CD_3OD), mixture of diastereomers δ 172.98 and 172.96, 171.38 and 171.32, 137.84 and 137.81, 130.32 and 130.30, 129.54 and 129.53, 128.00 and 127.98, 94.9, 54.8 and 54.7, 52.79 and 52.77, 49.8, 38.3 and 38.2. **HRMS** calculated for $C_{13}H_{18}NO_5$ $[M+H]^+$, 268.1185; found, 268.1192.

3.3.2 NMR spectra

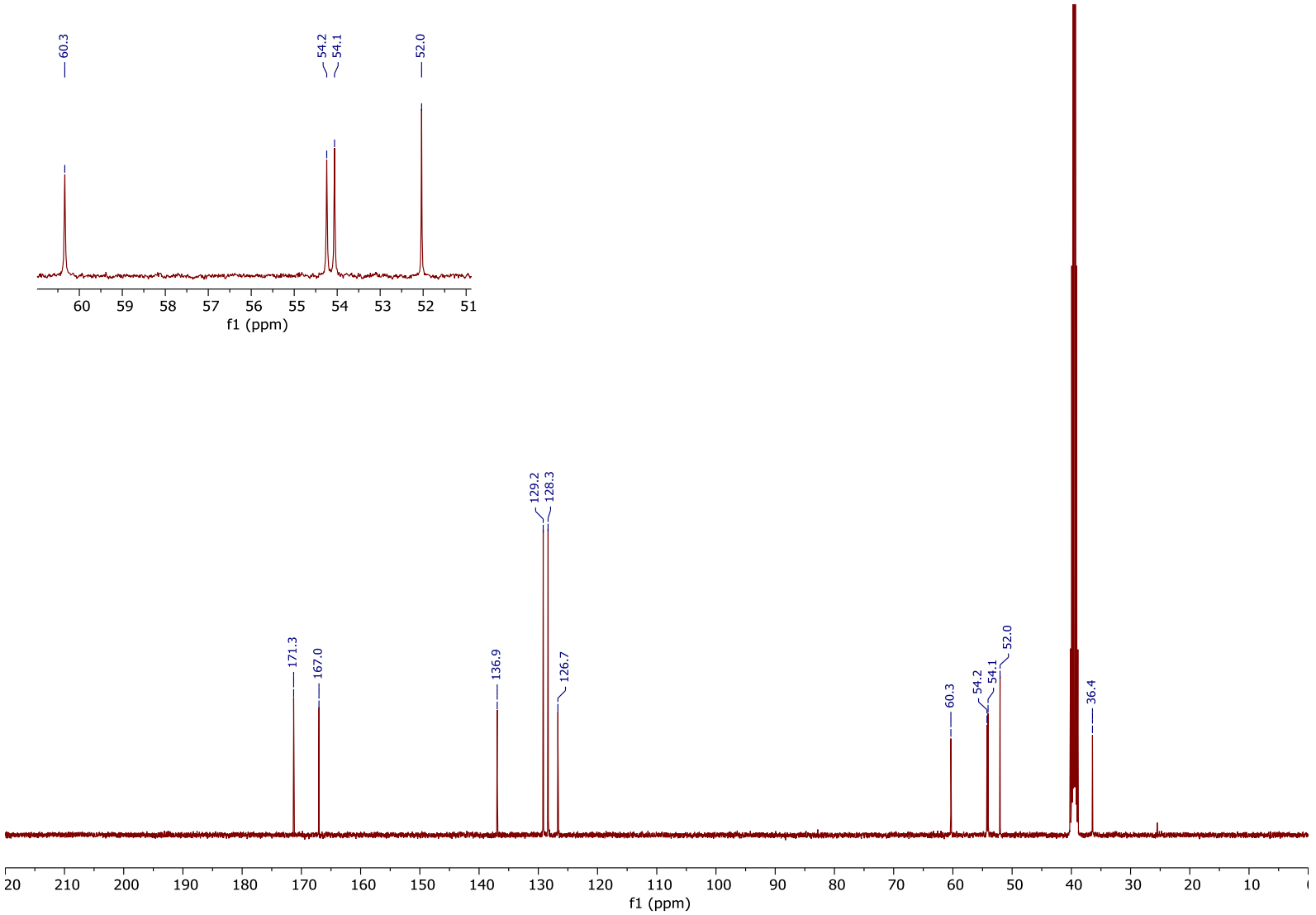


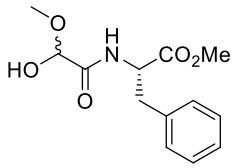
^{1}H NMR (400MHz), DMSO-d_6



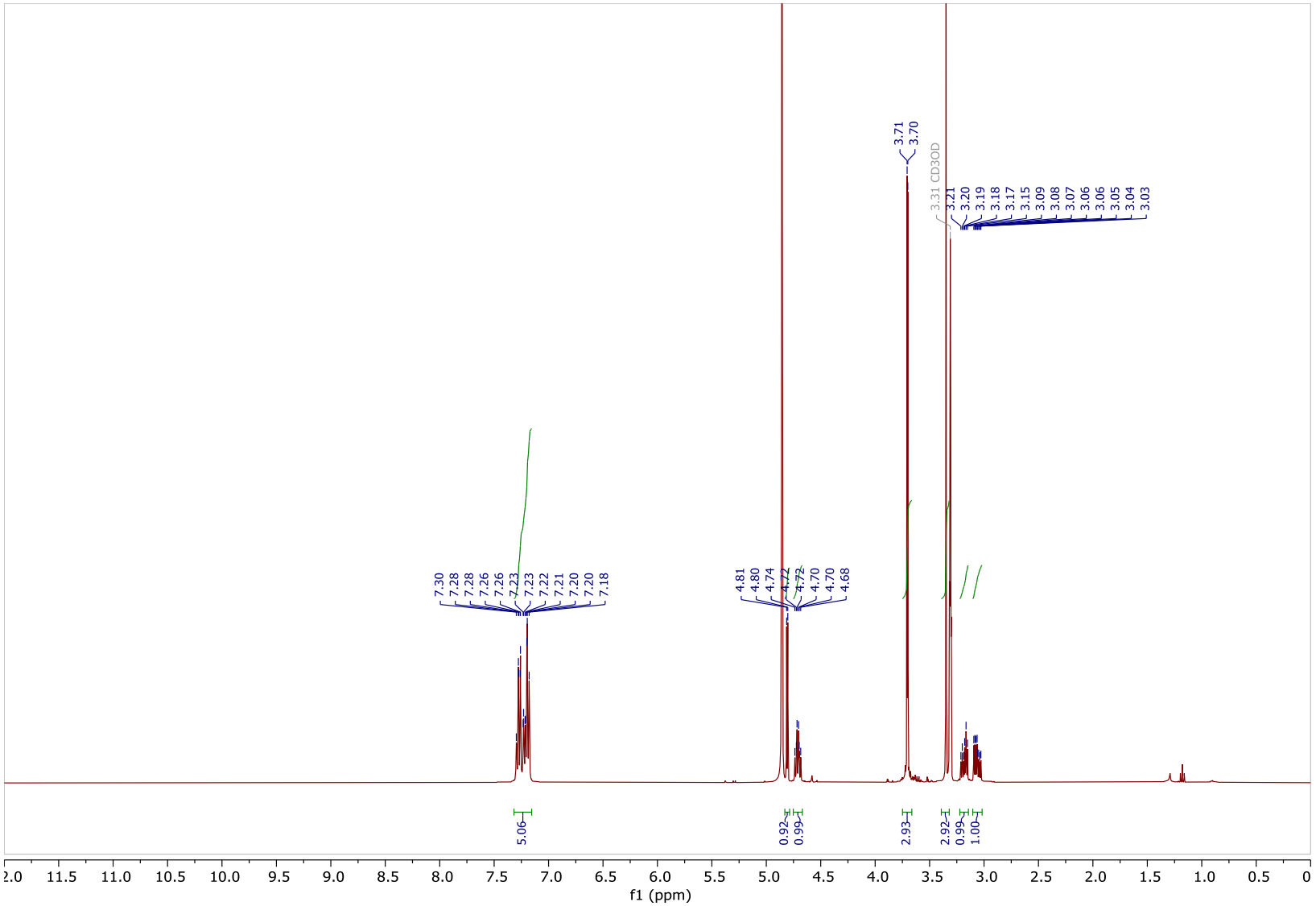


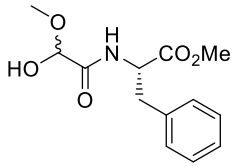
^{13}C NMR (101MHz), DMSO-d_6



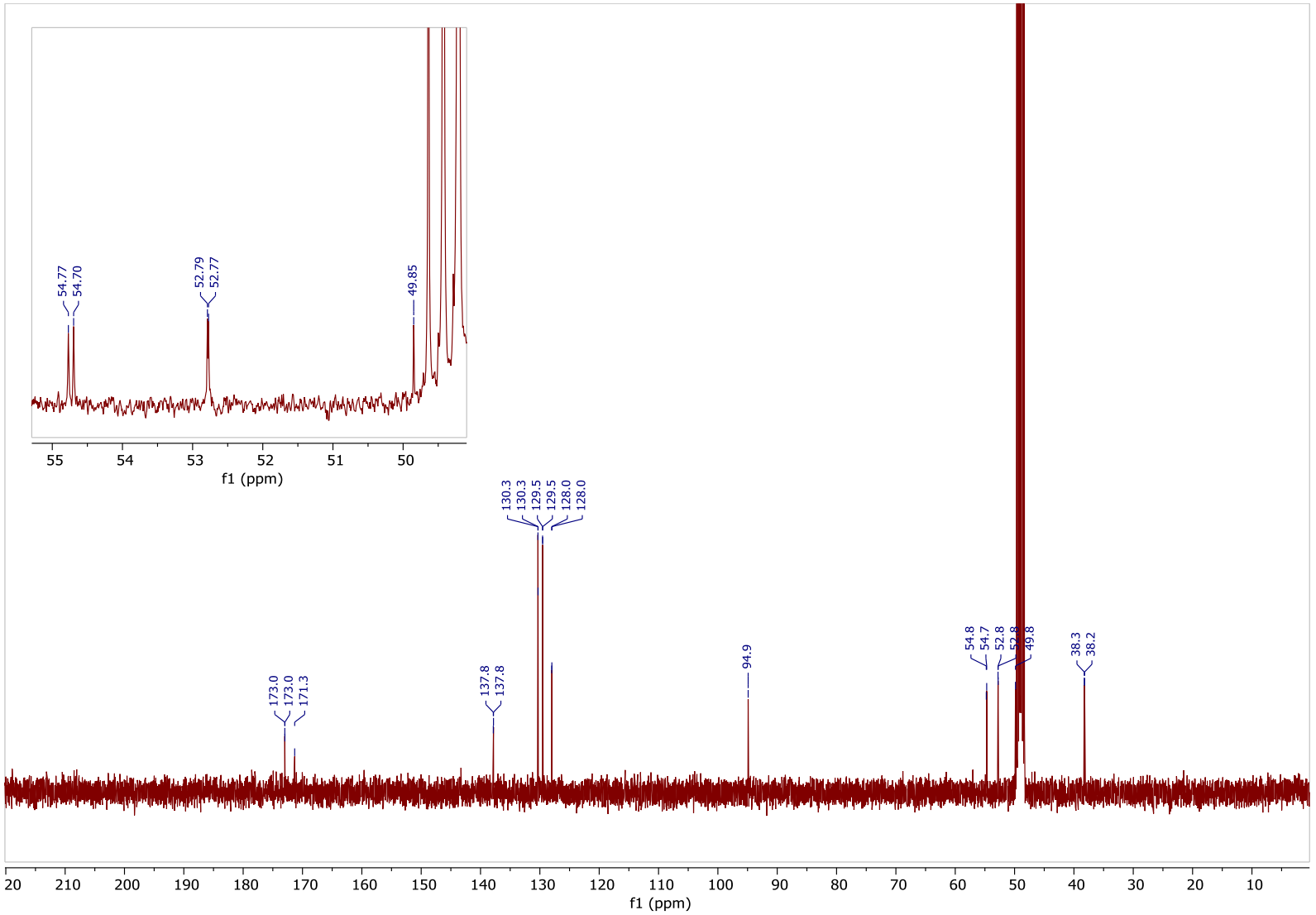


^{1}H NMR (400MHz), CD_3OD



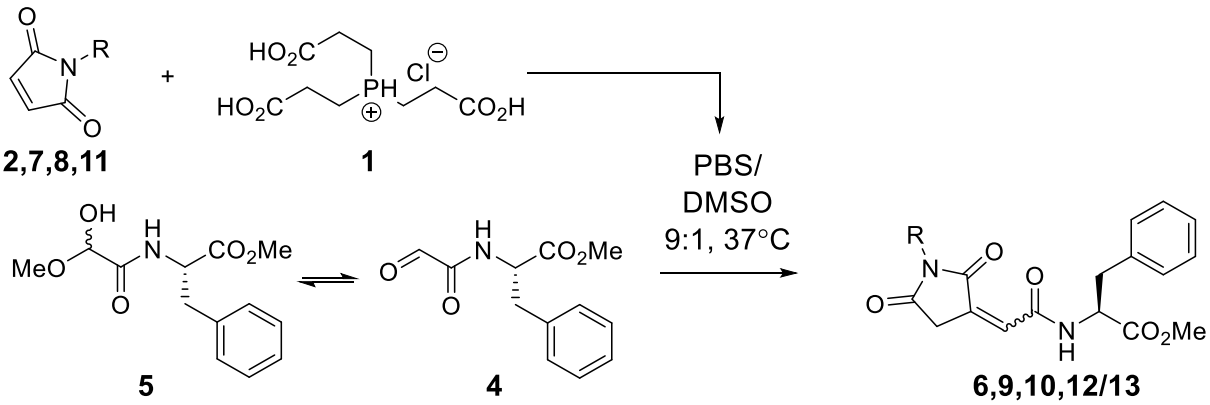


^{13}C NMR (101MHz), CD_3OD



5. Synthesis and characterization of small molecule alkenes 6, 9, 10, 12/13

5.1 General synthetic procedure



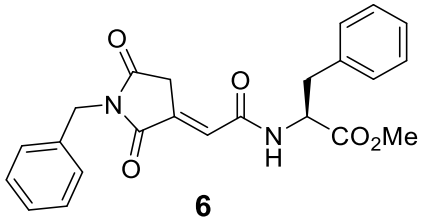
A 100mL RBF was charged with methylhemiacetal **5** (5.4mg, 20µmol, 1.0eq.). The respective maleimide (60µmol, 3.0eq.) and TCEP hydrochloride (**1**, 17 mg, 60µmol, 3.0eq.) were weighed into separate 4mL vials, the former was dissolved in DMSO (1mL) and the latter in aqueous PBS buffer (10mM phosphate, pH=7.4, 1mL). The TCEP solution was added to the DMSO solution in a single increment, which yielded a solution in all cases, and triethylamine (0.5µL, 6µmol, 0.3eq.) was added by micropipette. The formation of the phosphonium ylide was immediately assessed by means of HPLC and in each case was found to be complete. In the meantime, aqueous PBS buffer (10mM phosphate, pH=7.4, 8mL) and a stir bar were added to the RBF containing **5**, and the mixture was agitated until it appeared homogenous to the naked eye. The ylide solution was added, the resulting mixture was warmed to 37°C and stirred for 22-24h.

Quantification method A: In cases where the reaction mixture remained homogenous, i.e. the product did not precipitate or form an immiscible oil, the reaction mixture was analyzed by HPLC directly as follows: 10µL of reaction mixture were diluted twofold with water, and 5µL of the resulting sample were injected. The degree of conversion at the specified time-point was calculated retroactively after purification of the product, followed by calibration with the pure alkene.

Quantification method B: In cases where the reaction yielded insoluble products and became heterogeneous over time, the reaction mixture was extracted with $CHCl_3$ (3x3mL), the homogenous organic layers were pooled inside a volumetric flask, and diluted to exactly 10.0mL. 10µL of the organic layer were diluted twofold with methanol and analyzed in a fashion identical to that outlined above.

Isolation and HPLC-Calibration: Reaction products were purified by preparative HPLC, the volatiles were removed on a rotary evaporator, and the resulting aqueous solution was freeze-dried. The pure product was used for NMR-and HRMS-analyses as well as the retroactive calibration of the HPLC (λ=254nm) based on which conversions were determined.

5.2 *N*-benzylmaleimide derived alkene 6

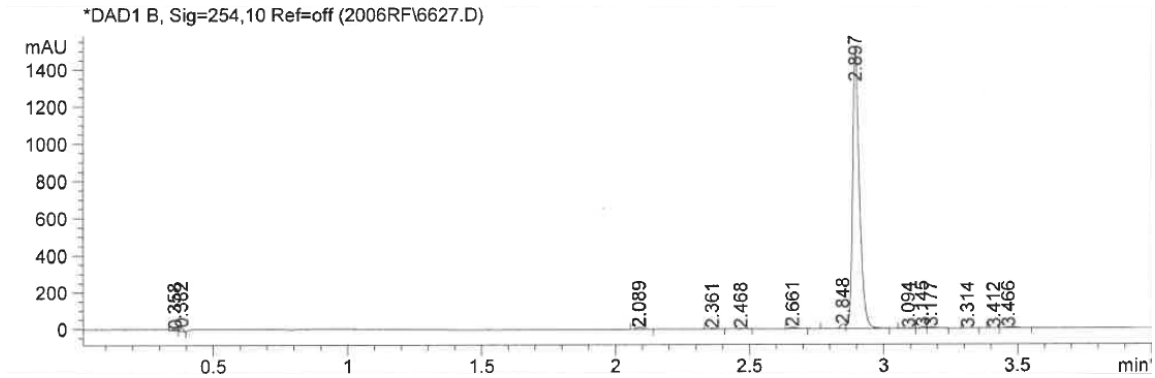


6

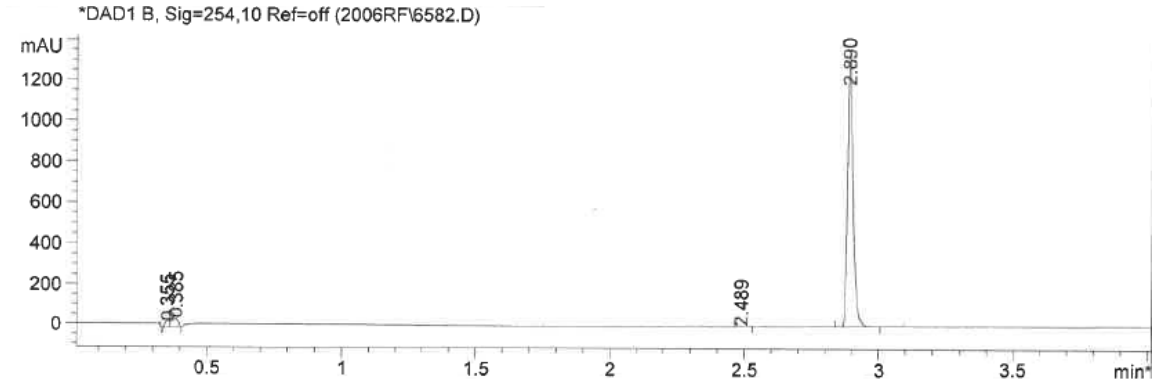
Double bond geometry not determined

methyl (2-(1-benzyl-2,5-dioxopyrrolidin-3-ylidene)acetyl)-L-phenylalaninate (**6**). Formed an immiscible oil as the reaction progressed. **Method B** was used to quantify conversion after 23h. A white solid was obtained after purification by preparative HPLC.

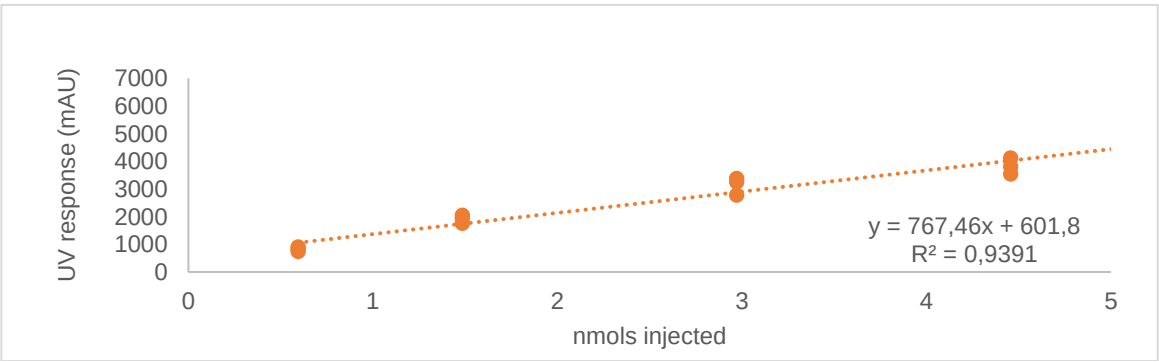
HPLC traces and calibration:



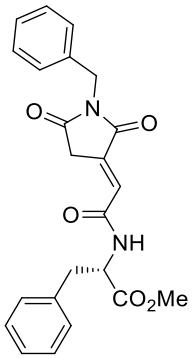
Reaction mixture after 23 hours and extraction into $CHCl_3$, in which the major isomer was exclusive



Purified product after preparative HPLC (99% purity)

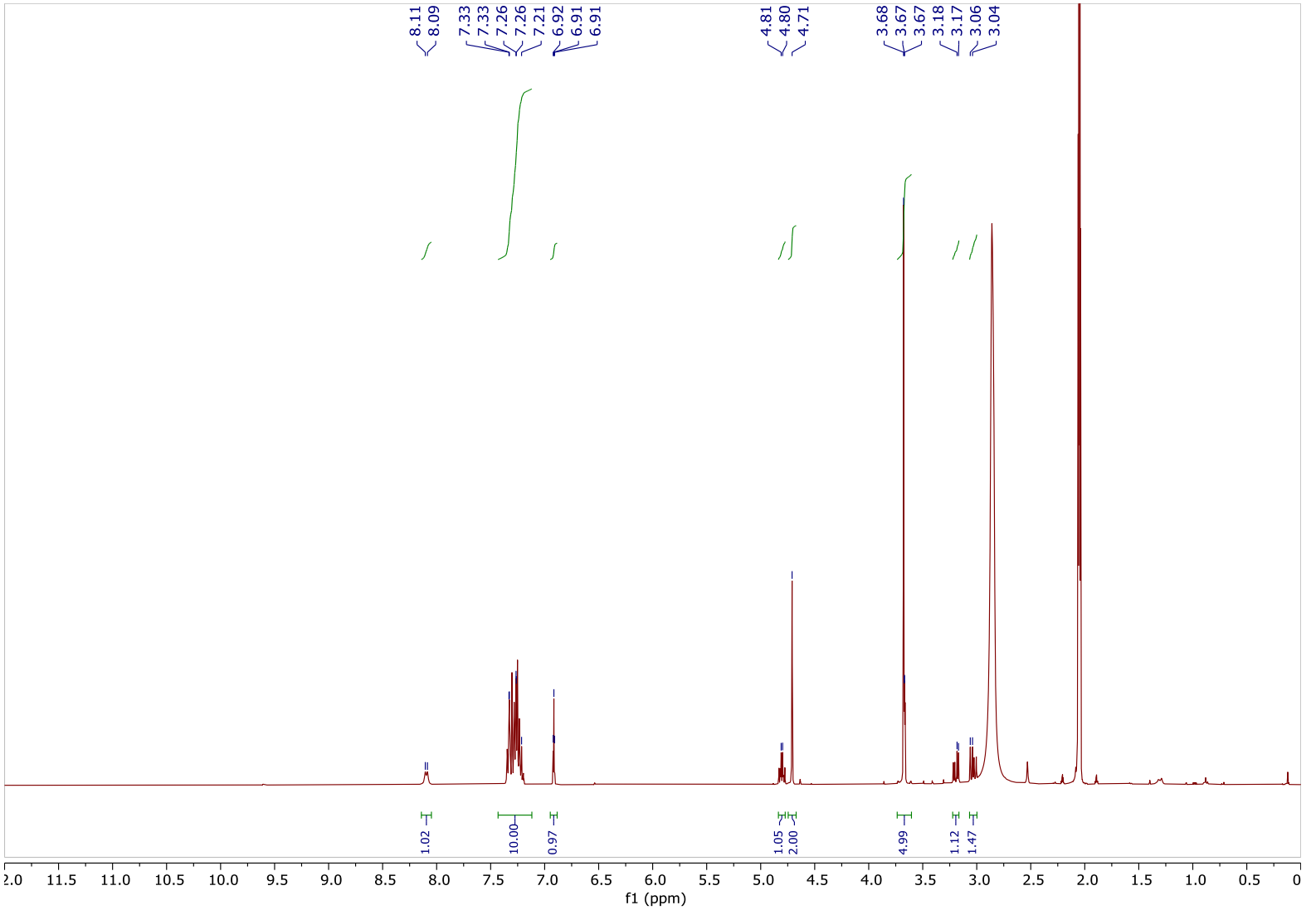


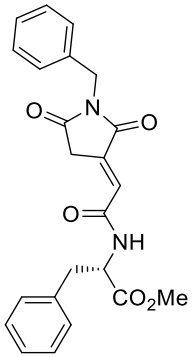
^{1}H NMR (400 MHz, acetone-d_6) δ 8.10 (d, *J* = 8.0 Hz, 1H), 7.37 – 7.17 (m, 10H), 6.91 (dd, *J* = 2.6, 2.5 Hz, 1H), 4.80 (td, *J* = 8.3, 5.5 Hz, 1H), 4.71 (s, 2H), 3.70 – 3.64 (m, 5H), 3.19 (dd, *J* = 13.9, 5.5 Hz, 1H), 3.03 (dd, *J* = 13.9, 8.5 Hz, 1H). **^{13}C NMR** (101 MHz, acetone-d_6) δ 174.7, 172.3, 170.2, 164.6, 139.1, 137.8, 137.3, 130.1, 129.31, 129.26, 129.0, 128.4, 127.6, 123.7, 54.8, 52.4, 42.7, 38.2, 35.0. **HRMS** calculated for $C_{23}H_{22}N_2O_5Na^+$ $[M+Na]^+$, 429.1426; found, 429.1430.



^{1}H NMR (400MHz), acetone-d_6

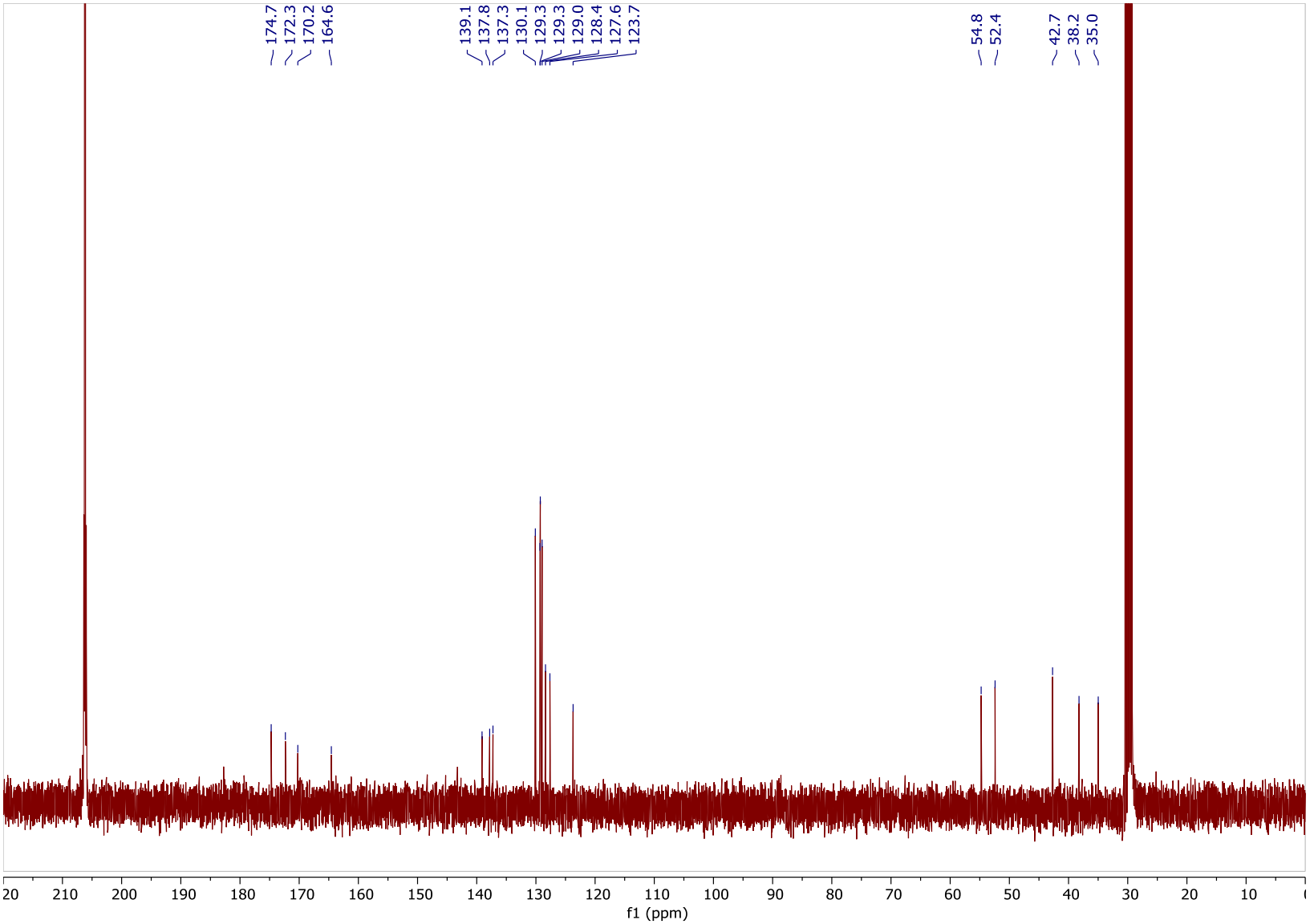
(Double bond geometry not determined)



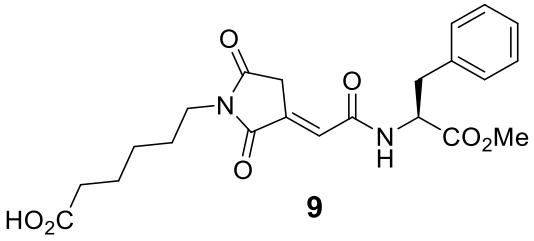


^{13}C NMR (101MHz), acetone-d_6

(Double bond geometry not determined)



5.3 6-Maleimidocaproic acid derived alkene 9

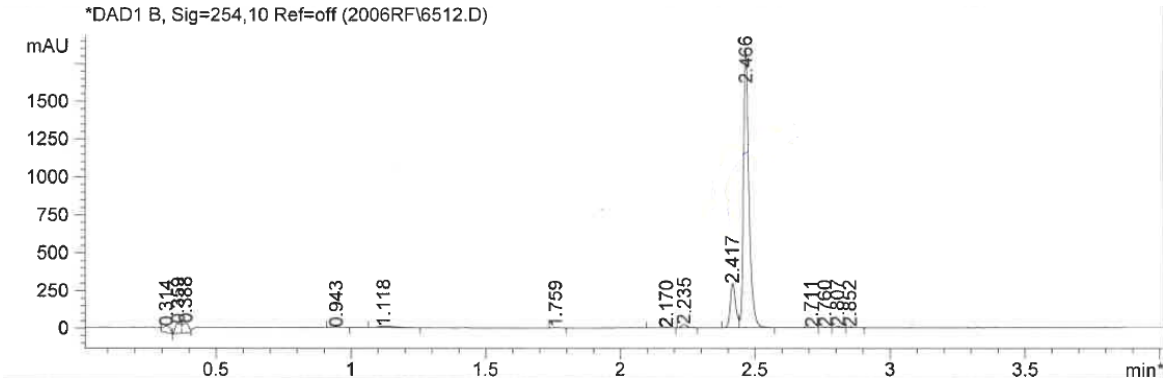


9

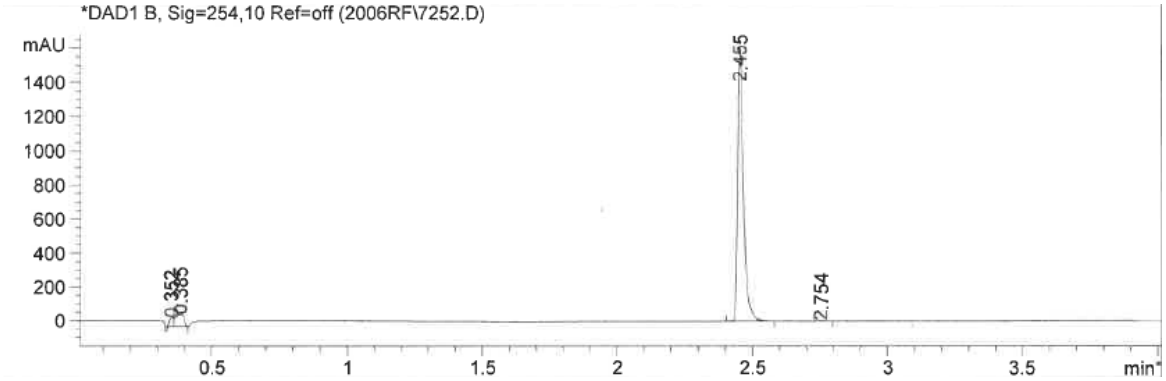
Double bond geometry not determined

(*S*)-6-(3-(2-((1-methoxy-1-oxo-3-phenylpropan-2-yl)amino)-2-oxoethylidene)-2,5-dioxopyrrolidin-1-yl)hexanoic acid (**9**). Yielded a homogenous reaction mixture. **Method A** was used to quantify conversion after 24h.

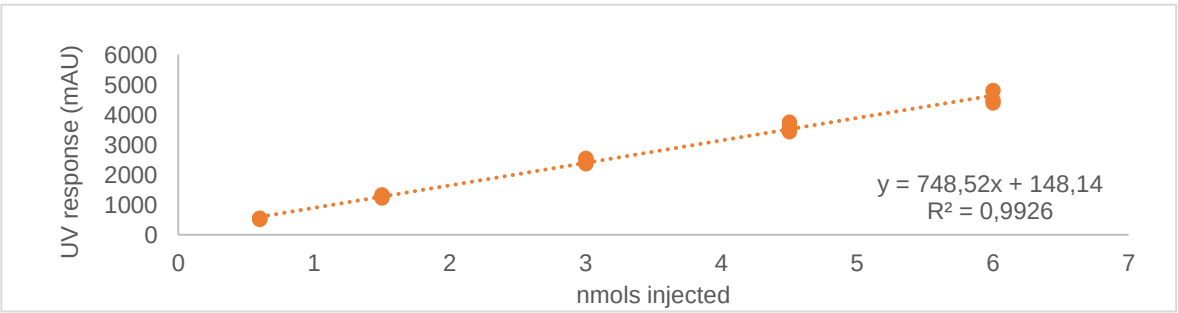
HPLC traces and calibration:



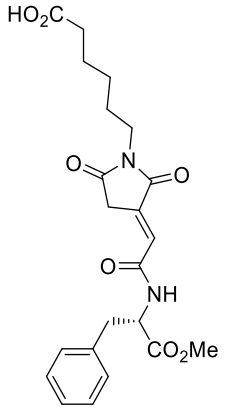
Crude reaction mixture after 24 hours. Major isomer: R_t = 2.466min, minor isomer: R_t = 2.417min



Purified product after preparative HPLC (99% purity)

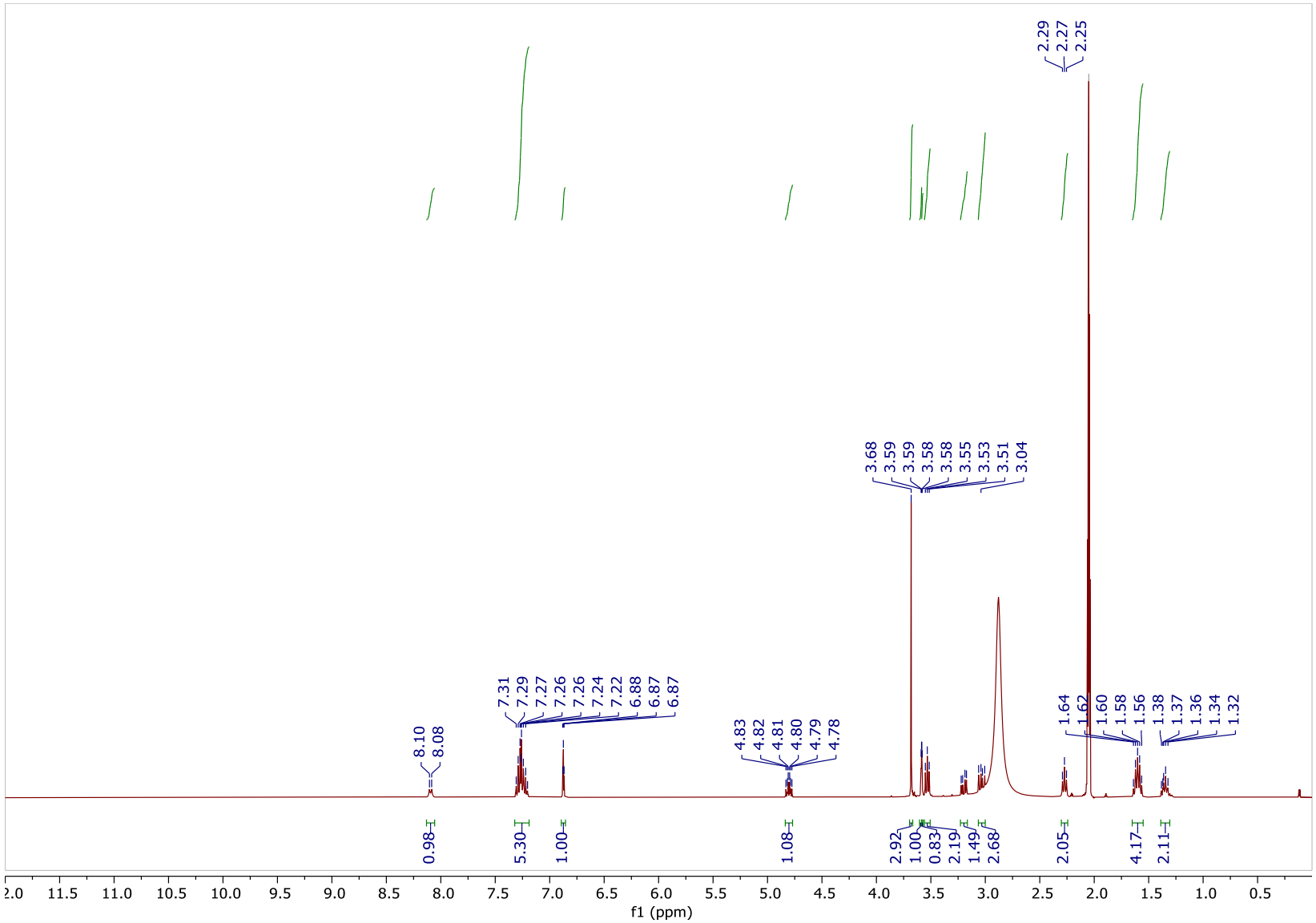


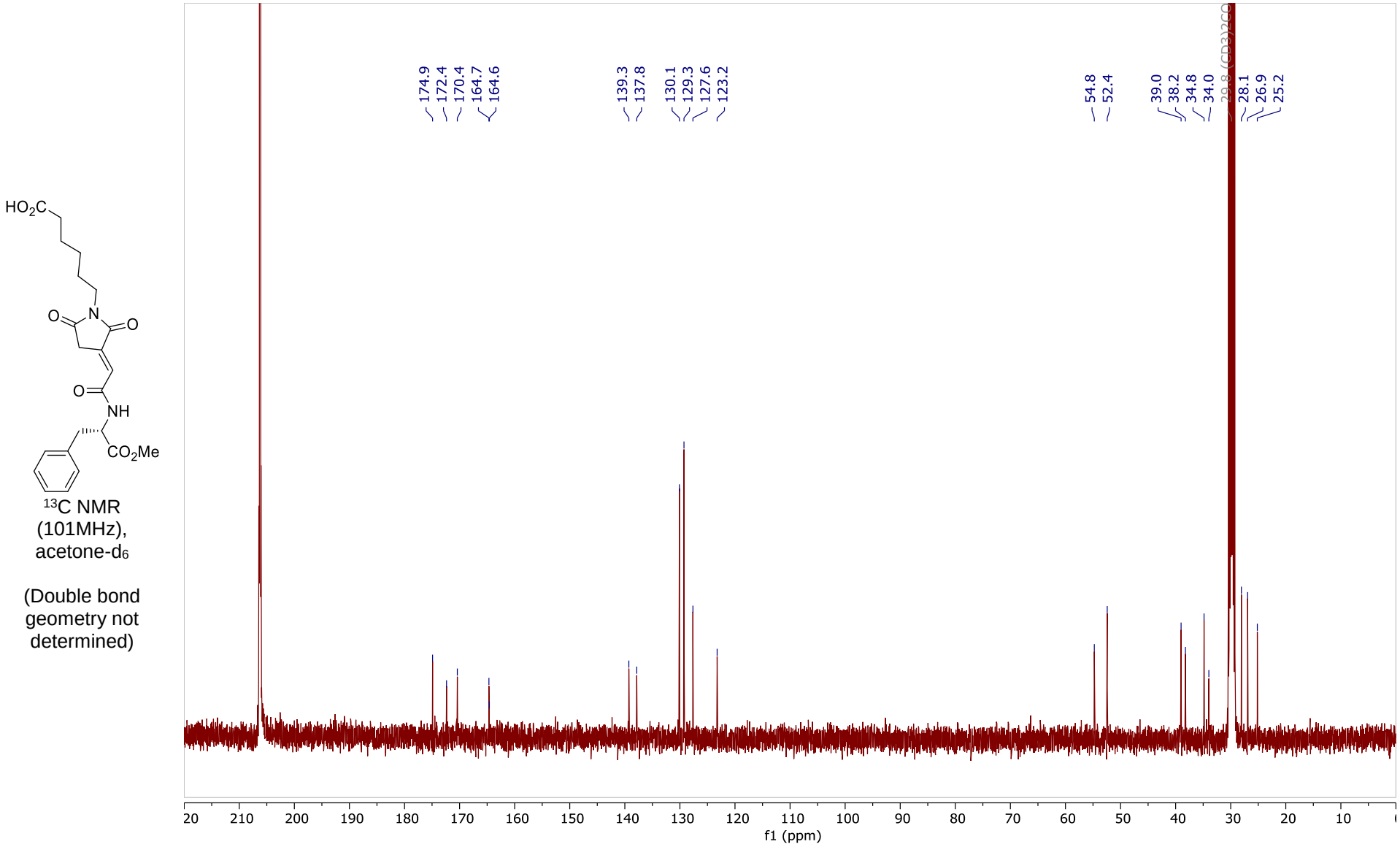
^{1}H NMR (400 MHz, acetone-d_6) δ 8.09 (d, *J* = 8.0 Hz, 1H), 7.33 – 7.18 (m, 5H), 6.87 (dd, *J* = 2.5, 2.4 Hz, 1H), 4.80 (ddd, *J* = 8.3, 8.0, 5.5 Hz, 1H), 3.68 (s, 3H), 3.59 (d, *J* = 2.5 Hz, 1H), 3.58 (d, *J* = 2.4 Hz, 1H) 3.53 (t, *J* = 7.2 Hz, 2H), 3.20 (dd, *J* = 13.9, 5.5 Hz, 1H), 2.03 (dd, *J* = 13.9, 8.5 Hz, 1H, obscured by H_2O), 2.27 (t, *J* = 7.4 Hz, 2H), 1.65 – 1.55 (m, 4H), 1.41 – 1.30 (m, 2H). ^{13}C NMR (101 MHz, acetone-d_6) δ 174.9, 172.4, 170.4, 164.7, 164.6, 139.3, 137.8, 130.1, 129.3, 127.6, 123.2, 54.8, 52.4, 39.0, 38.2, 34.8, 34.0, 28.1, 26.9, 25.2. **HRMS** calculated for $C_{22}H_{27}N_2O_7^+$ $[M+H]^+$, 431.1818; found, 431.1821.



^{1}H NMR (400MHz), acetone-d_6

(Double bond geometry not determined)

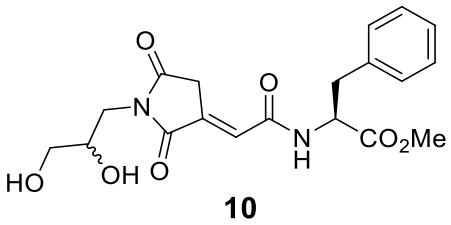




^{13}C NMR (101MHz), acetone-d_6

(Double bond geometry not determined)

5.4 *N*-(2,3-dihydroxypropyl)maleimide derived alkene 10

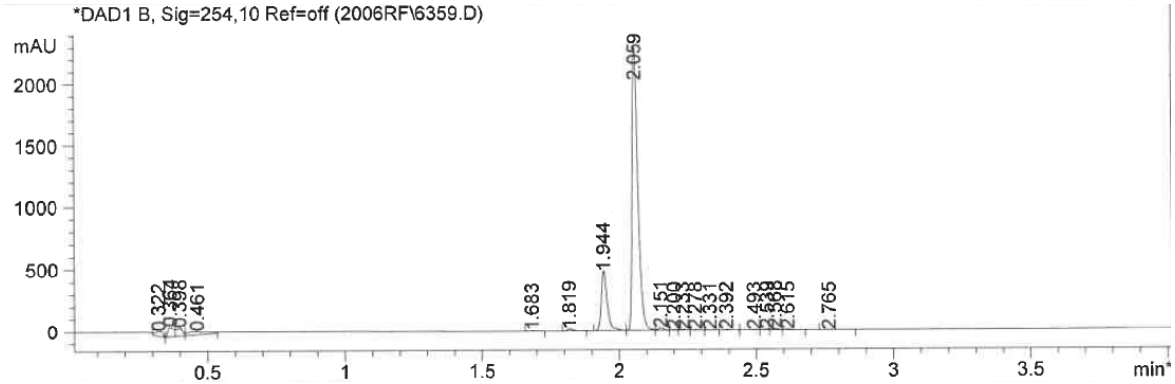


10

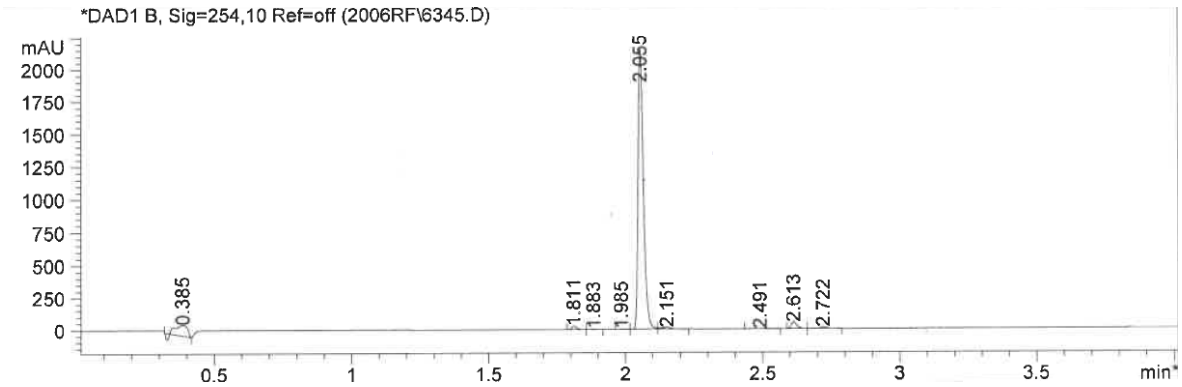
Double bond geometry not determined

methyl (2-(1-(2,3-dihydroxypropyl)-2,5-dioxopyrrolidin-3-ylidene)acetyl)-L-phenylalaninate (**10**). Yielded a homogenous reaction mixture. **Method A** was used to quantify conversion after 22h.

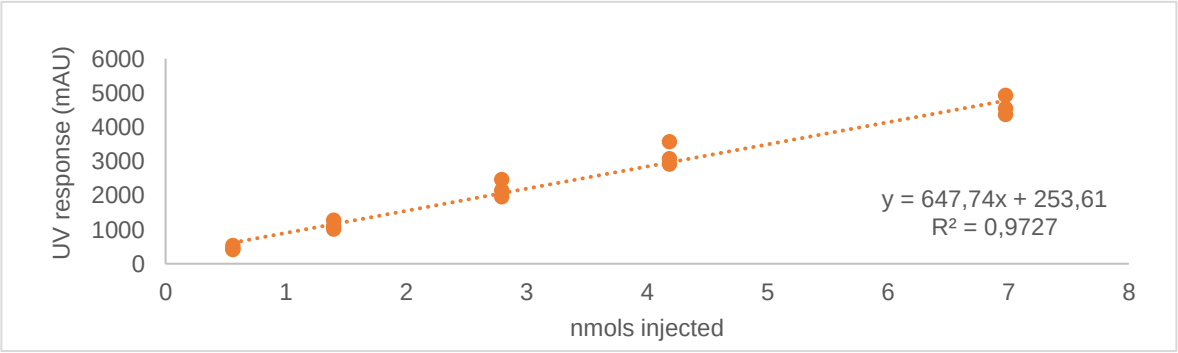
HPLC traces and calibration:



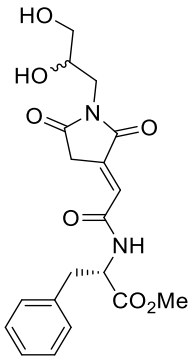
Crude reaction mixture after 22 hours. Major isomer: R_t = 2.059min, minor isomer: R_t = 1.944min



Purified product after preparative HPLC (94% purity)

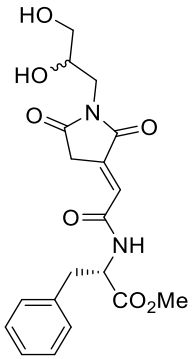


^{1}H NMR (400 MHz, acetone-d_6) δ 8.10 (d, *J* = 7.8 Hz, 1H), 7.33 – 7.18 (m, 5H), 6.88 (dd, *J* = 2.6, 2.5 Hz, 1H), 4.81 (td, *J* = 8.4, 5.4 Hz, 1H), 3.92 (m, w=23.2 Hz, 1H), 3.72 – 3.60 (m, 2H), 3.68 (s, 3H, obscured by previous multiplet), 3.61 – 3.57 (m, 2H), 3.51 (d, *J* = 5.3 Hz, 2H), 3.20 (dd, *J* = 13.9, 5.5 Hz, 1H), 3.03 (dd, *J* = 13.9, 8.6 Hz, 1H, obscured by H_2O-peak). **^{13}C NMR** (101 MHz, acetone-d_6) δ 175.3, 172.4, 170.8, 164.6, 139.2, 137.8, 130.1, 129.3, 127.6, 123.4, 69.8, 65.1, 54.8, 52.4, 42.9, 38.2, 34.8. **HRMS** calculated for $C_{19}H_{23}N_2O_7^+$ $[M+H]^+$, 391.1512; found, 391.1505.



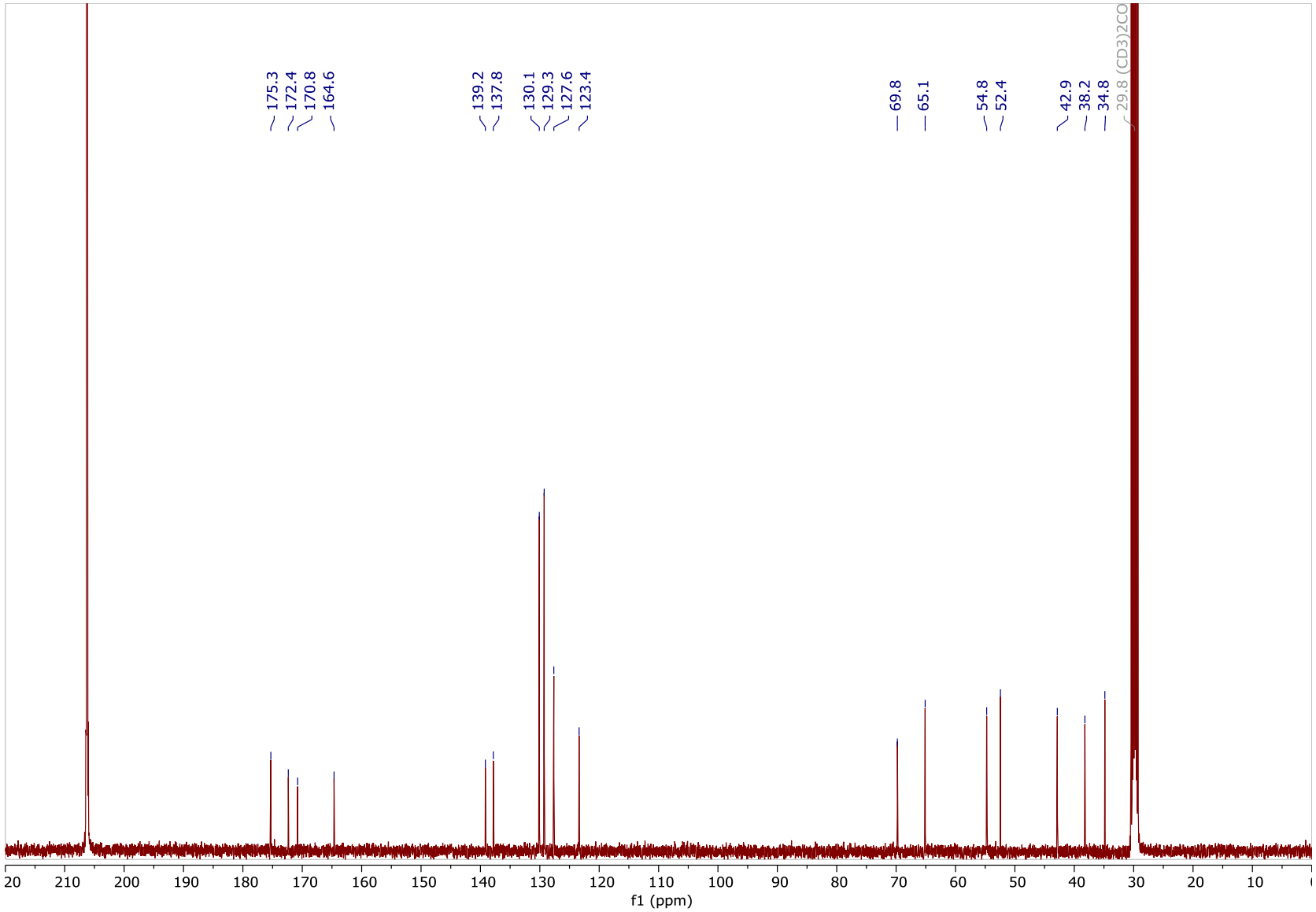
^{1}H NMR (400MHz), acetone-d_6

(Double bond geometry not determined)

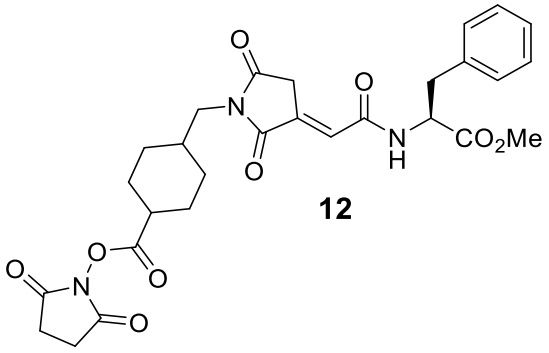


^{13}C NMR (101MHz), acetone-d_6

(Double bond geometry not determined)



5.5 SMCC derived alkene 12 (NHS ester intact)

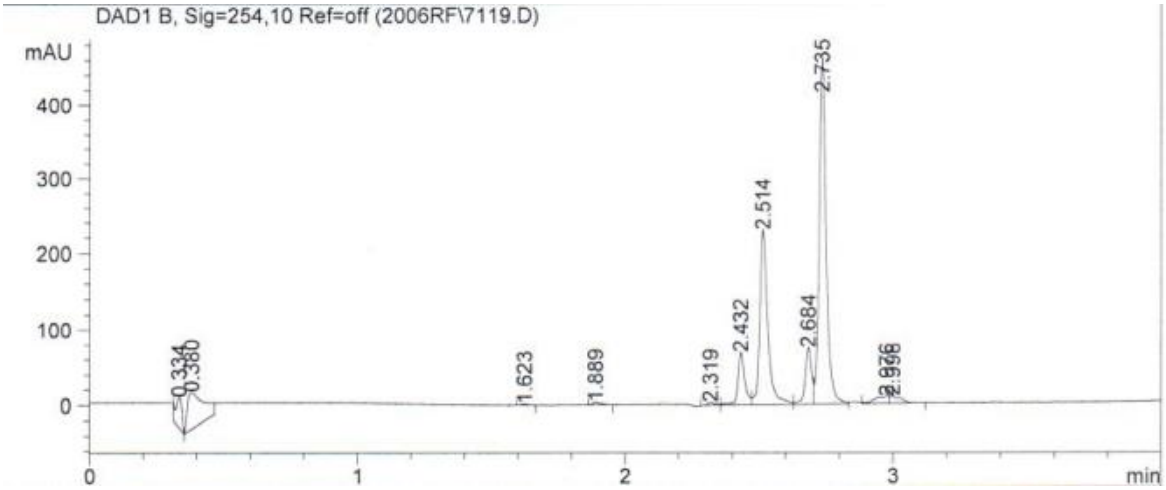


12

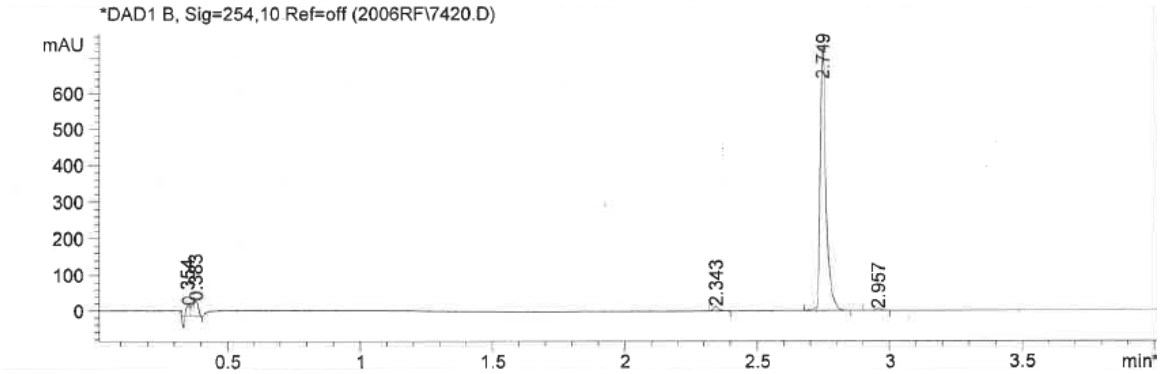
Double bond geometry not determined

2,5-dioxopyrrolidin-1-yl (*S*)-4-((3-(2-((1-methoxy-1-oxo-3-phenylpropan-2-yl)amino)-2-oxoethylidene) - 2,5-dioxopyrrolidin-1-yl)methyl)cyclohexane-1-carboxylate (**12**). Formed an immiscible solid as the reaction progressed. **Method B** was used to quantify conversion after 24h.

HPLC traces:

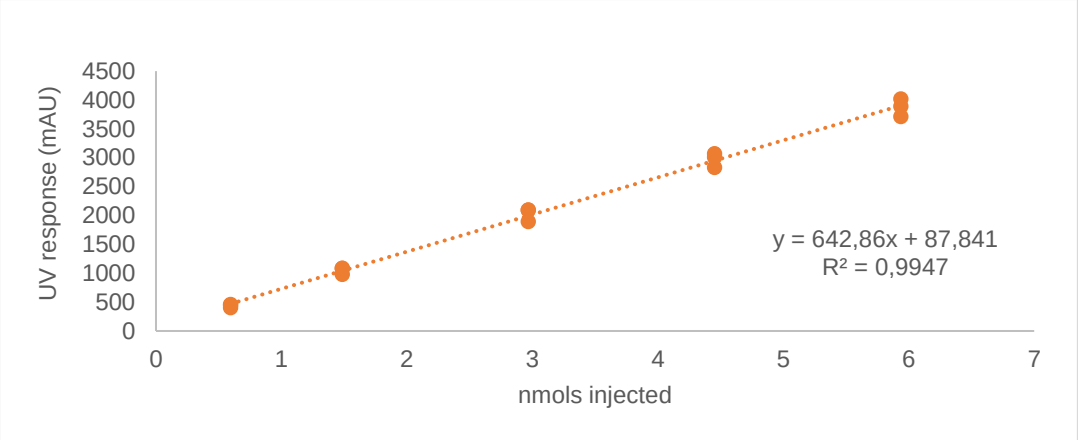


Reaction mixture after 24 hours and extraction into $CHCl_3$, NHS-ester: R_t=2.735min (major) and R_t=2.684min (minor isomer), carboxylic acid (see page **Error! Bookmark not defined.**): R_t=2.514min (major) and R_t=2.514min (minor isomer).

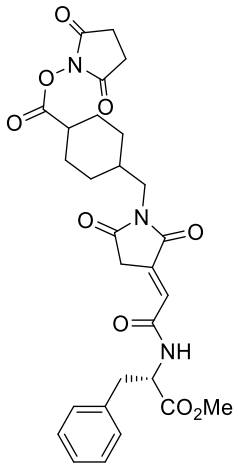


Purified product after preparative HPLC (96% purity)

HPLC calibration:

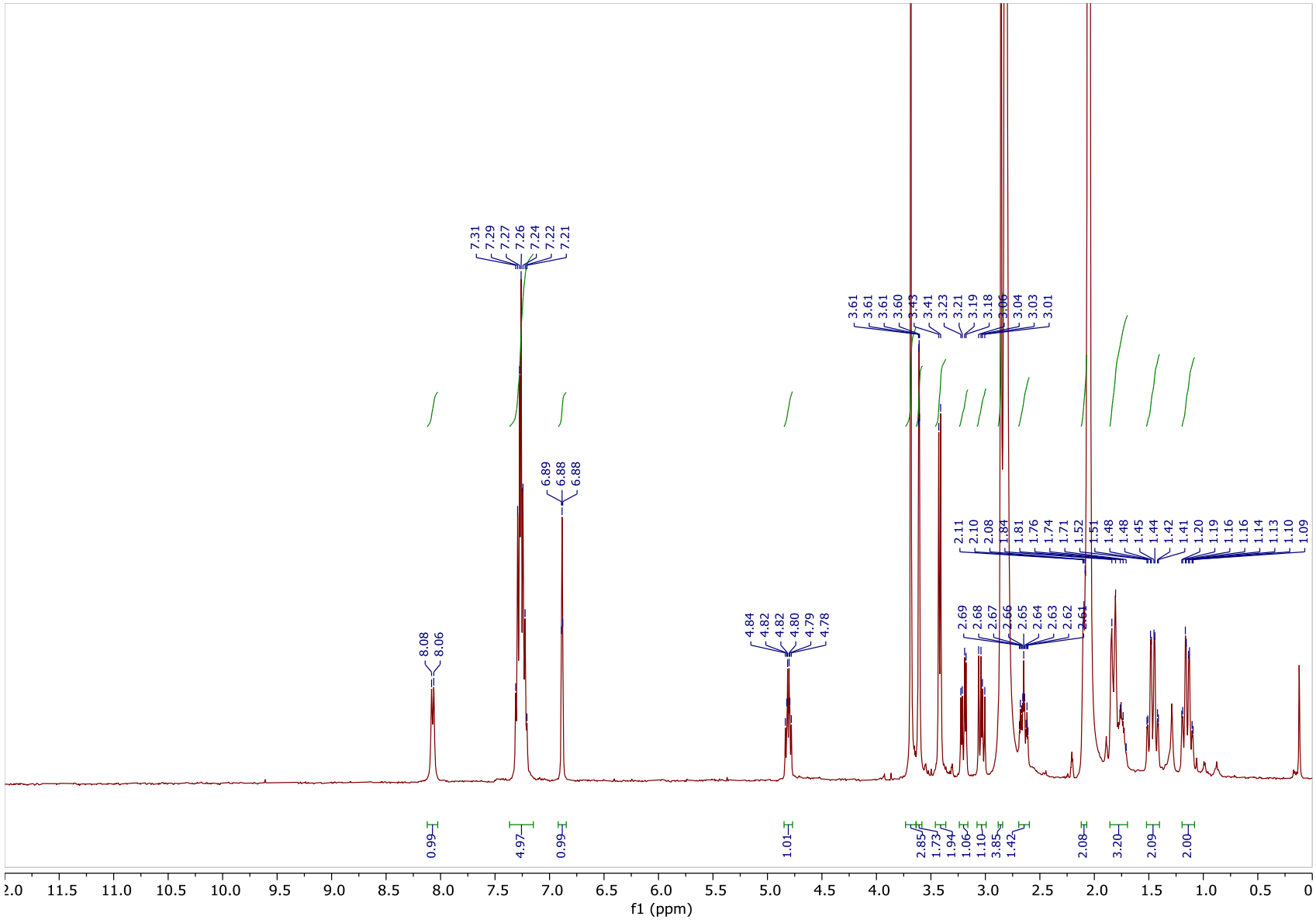


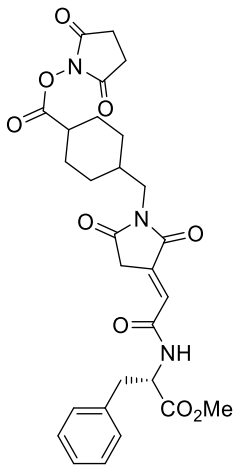
^{1}H NMR (400 MHz, acetone-d_6) δ 8.07 (d, *J* = 8.1 Hz, 1H), 7.33 – 7.18 (m, 5H), 6.88 (dd, *J* = 2.3, 2.1 Hz, 1H), 4.81 (ddd, *J* = 8.3, 8.1, 5.4 Hz, 1H), 3.68 (s, 3H), 3.61 (d, *J* = 2.3 Hz, 1H, overlaps with following doublet), 3.60 (d, *J* = 2.1 Hz, 1H, overlaps with preceding doublet), 3.42 (d, *J* = 6.9 Hz, 2H), 3.20 (dd, *J* = 13.9, 5.4 Hz, 1H), 3.03 (dd, *J* = 13.9, 8.3 Hz, 1H), 2.86 (s, 4H), 2.65 (tt, *J* = 12.1, 3.5 Hz, 1H), 2.13 – 2.06 (m, 2H, obscured by solvent peak), 1.88 – 1.67 (m, 3H), 1.54 – 1.39 (m, 2H), 1.22 – 1.07 (m, 2H). **^{13}C NMR** (126 MHz, acetone-d_6) δ 175.1, 172.3, 171.7, 170.63, 170.49, 164.64, 164.57, 139.2, 137.8, 130.1, 129.3, 127.6, 123.3, 54.8, 54.7, 52.4, 44.9, 41.0, 38.23, 38.20, 36.6, 34.8, 29.1, 26.3. **HRMS** calculated for $C_{28}H_{32}N_3O_9^+$ $[M+H]^+$, 554.2133; found, 554.2138.



^{1}H NMR (400MHz), acetone-d_6

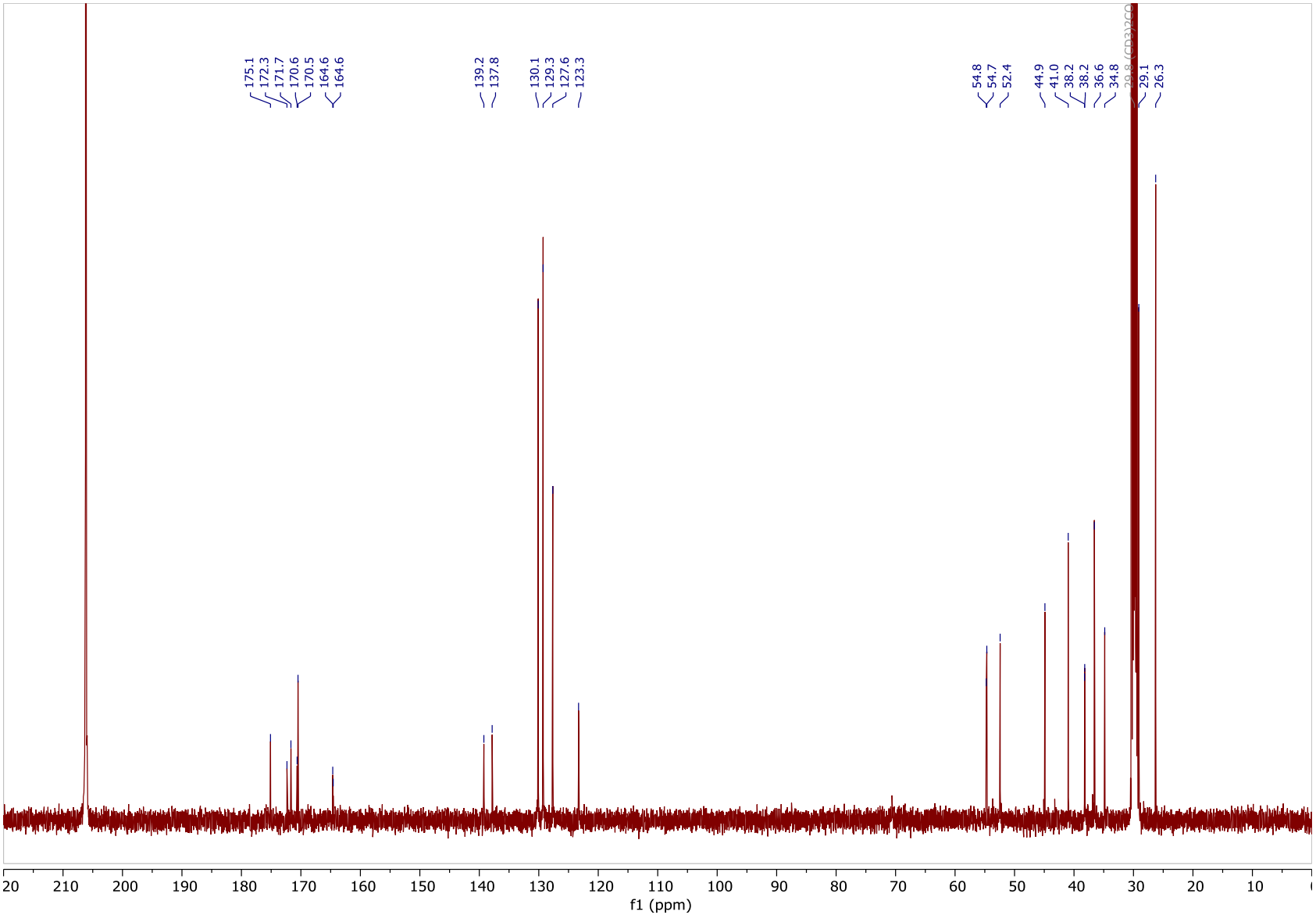
(Double bond geometry not determined)



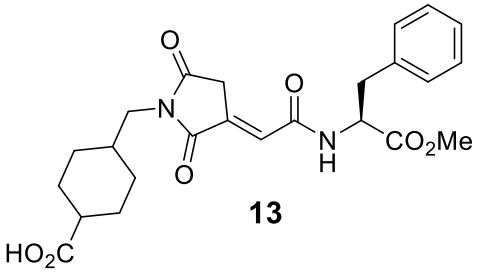


^{13}C NMR (126MHz), acetone-d_6

(Double bond geometry not determined)



5.6 SMCC derived alkene 13 (hydrolyzed NHS-ester)



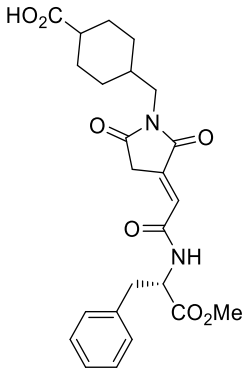
13

Double bond geometry not determined

(*S*)-4-((3-(2-((1-methoxy-1-oxo-3-phenylpropan-2-yl)amino)-2-oxoethylidene)-2,5-dioxopyrrolidin-1-yl)methyl)cyclohexane-1-carboxylic acid (**13**). **Method B** was used to quantify conversion after 24h.

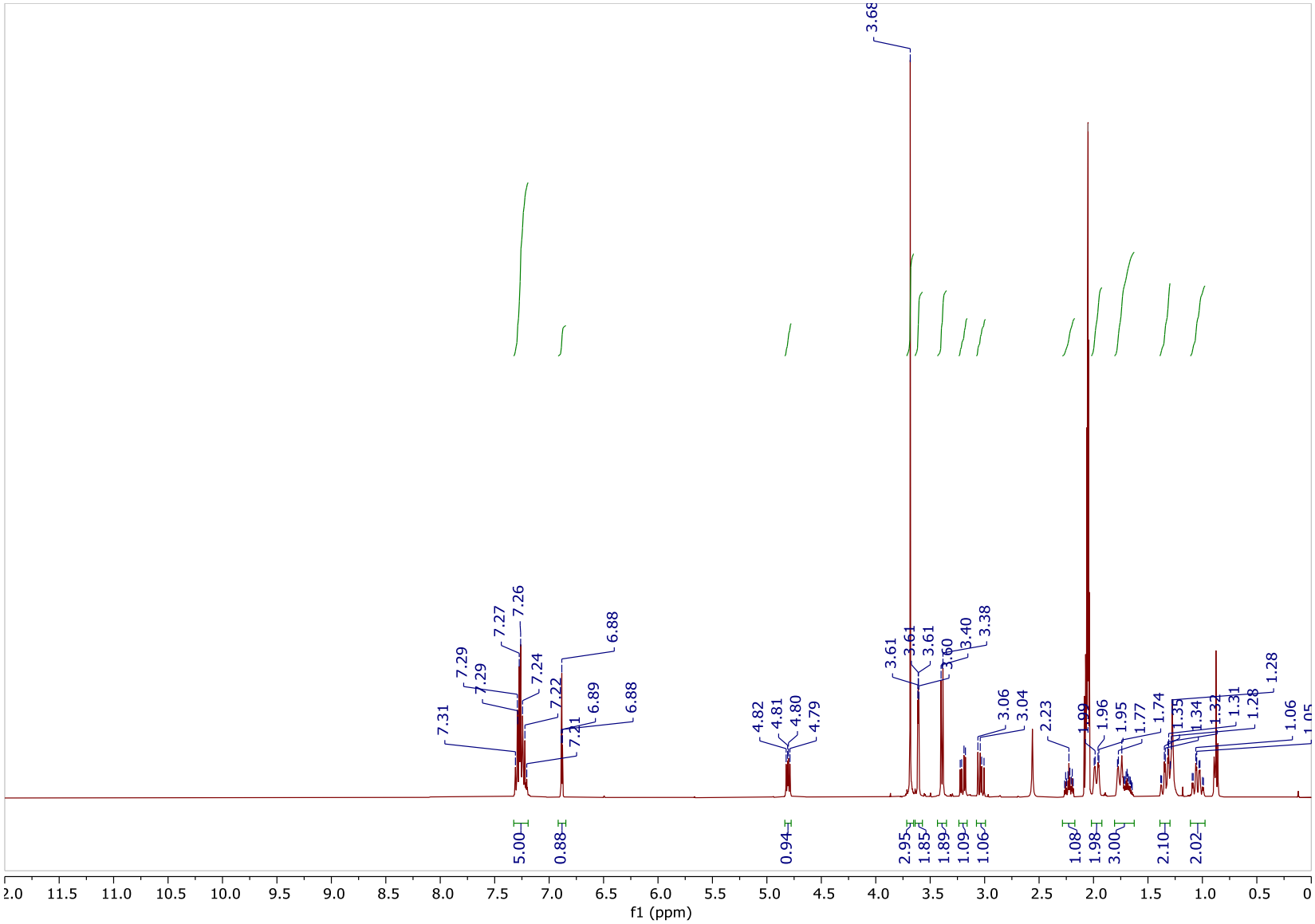
HPLC traces and calibration: See page 48 and the following one.

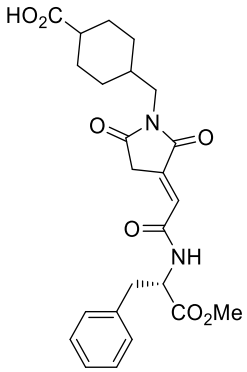
^{1}H NMR (400 MHz, acetone-d_6) δ 7.33 – 7.18 (m, 5H), 6.88 (dd, *J* = 2.6, 2.5 Hz, 1H), 4.81 (dd, *J* = 8.6, 5.4 Hz, 1H), 3.68 (s, 3H), 3.61 (d, *J* = 2.5 Hz, 1H, overlaps with following doublet), 3.60 (d, *J* = 2.6 Hz, 1H, overlaps with preceding doublet), 3.39 (d, *J* = 7.0 Hz, 2H), 3.20 (dd, *J* = 13.9, 5.4 Hz, 1H), 3.03 (dd, *J* = 13.9, 8.6 Hz, 1H), 2.23 (tt, *J* = 12.2, 3.6 Hz, 1H), 2.00 – 1.94 (m, 2H), 1.81 – 1.63 (m, 3H), 1.39 – 1.28 (m, 2H), 1.10 – 0.98 (m, 2H). **^{13}C NMR** (101 MHz, acetone-d_6) δ 175.1, 173.3, 172.4, 170.6, 164.6, 139.2, 137.8, 130.1, 129.3, 127.6, 123.2, 54.7, 52.4, 45.1, 43.3, 38.2, 36.9, 34.8, 32.6, 23.3. **HRMS** calculated for $C_{24}H_{29}N_2O_7^+$ $[M+H]^+$, 457.1975; found, 457.1981.



^{1}H NMR (400MHz), acetone-d_6

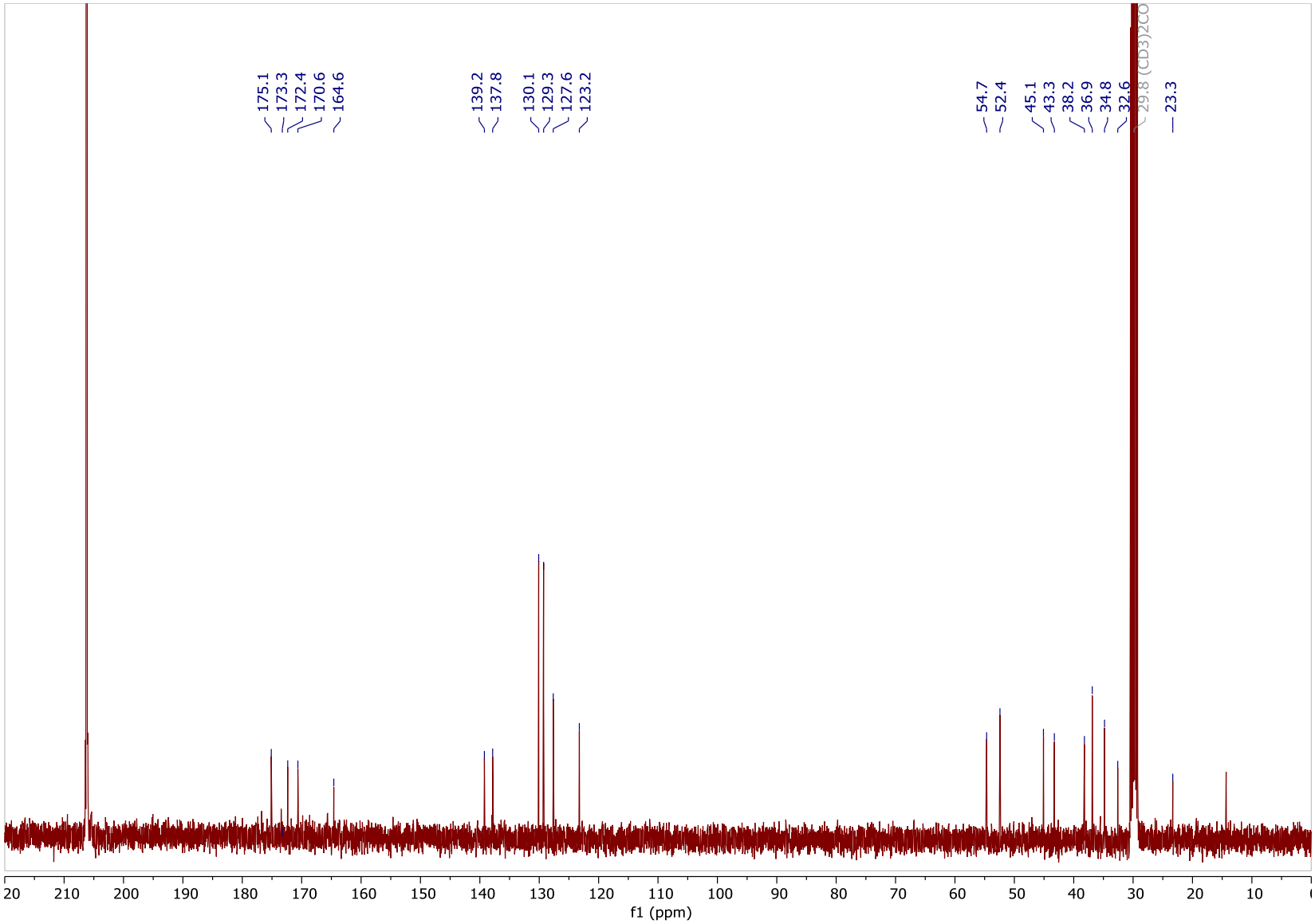
(Double bond geometry not determined)





^{13}C NMR (101MHz), acetone-d_6

(Double bond geometry not determined)



6. Oxidation of proteins, bioconjugation of small molecule phosphonium ylides and characterization of the resulting conjugates

6.1 ADAPT6 conjugates

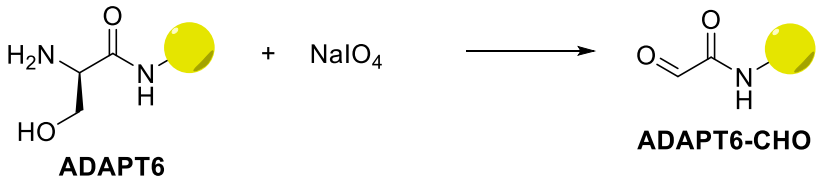
6.1.1 Protein expression and purification

An expression construct for **ADAPT6** with an N-terminal serine and a C-terminal hexahistidine tag (see table S1) was constructed using standard cloning procedures. Sequence-verified plasmids were transformed into E. coli BL21* and single transformants were grown at 37°C overnight (ON) in Terrific Soy Broth (TSB) supplemented with yeast extract and 50 µg/ml kanamycin (TSB+Y/kan). 200 ml TSB+Y/kan was inoculated with 2 ml ON-culture and grown at 37°C until an OD600 of 1 was reached. Cultures were induced with isopropyl-β-D-1-thiogalactopyranoside to a final concentration of 1 mM and grown at 25°C ON. Pelleted bacterial cells were resuspended in denaturing lysis buffer (7 M guanidium hydrochloride, 50 mM sodium phosphate, 10 mM Tris, 100 mM NaCl, pH 8.0) and incubated shaking for 2 h at 37°C. Cell lysates were subjected to immobilized metal ion affinity chromatography (IMAC) using TALON metal affinity resin and denaturing washing- (6 M guanidium hydrochloride, 50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 8.0) and elution (6 M urea, 50 mM sodium phosphate, 100 mM sodium chloride, 30 mM acetic acid, 70 mM sodium acetate, pH 5.0) buffers. Purified protein was buffer exchanged into 5 mM ammonium acetate pH 5.5 using PD10 desalting columns and lyophilized. Concentrations were determined from absorbance at 280 nm prior to lyophilization and molar extinction coefficients derived from the amino acid sequence. Purity was analyzed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and the molecular weight of the protein was confirmed by ESI-TOF MS (deconvoluted **MS** calculated for $C_{328}H_{496}N_{94}O_{97}$ $[M+H]^+$, 7308; found, 7308).

SGDEAVDANSLAAAKETALYHLDRLGVADAYKDLIDK AKTVEGVKARYFEILHALPYYGGHHHHHH

Table S1 - The amino acid sequence of H_2N-Ser-ADAPT6-His_6-CO_2H

6.1.2 Protein oxidation

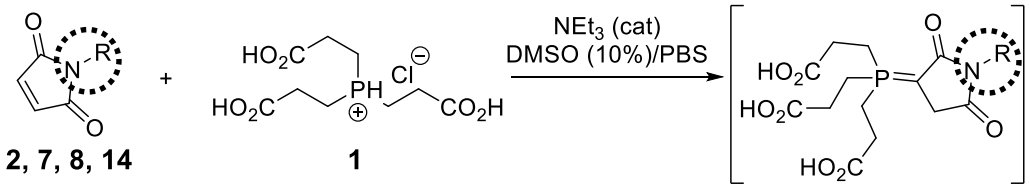


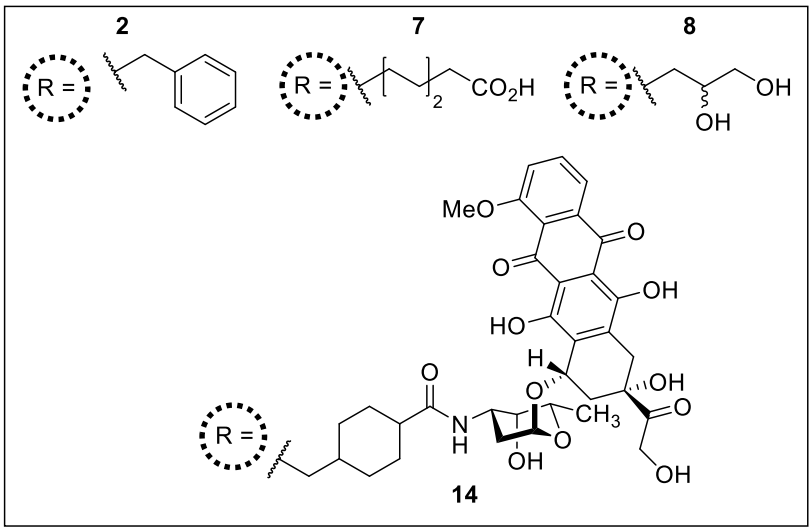
ADAPT6 **ADAPT6-CHO**

A solution of sodium periodate (6.02mg, 28.1µmol) in aqueous PBS buffer (50mM phosphate, pH=7.4, 10mL) was prepared and 1mL (corresponding to about 15eq of periodate) of it was used to dissolve a pellet of **ADAPT6** (1.31mg, 179nmol, 1.0eq.). After one hour of incubation at room temperature, the reaction mixture was desalted into PBS (10mM phosphate, pH=7.4) using PD Spintrap G25 desalting units and the samples of **ADAPT6-CHO** were pooled and frozen at -80°C.

Determination of protein concentration. The molar extinction coefficient of **ADAPT6-CHO** at λ=280nm was assumed to be approximately the same as that of the unoxidized protein, which in turn was calculated to be 7450$M^{-1}cm^{-1}$ based on its amino acid sequence. The absorption of five samples of protein at that wavelength were determined to reveal a protein concentration of 0.98±0.04mg/mL, indicating about 95% recovery.

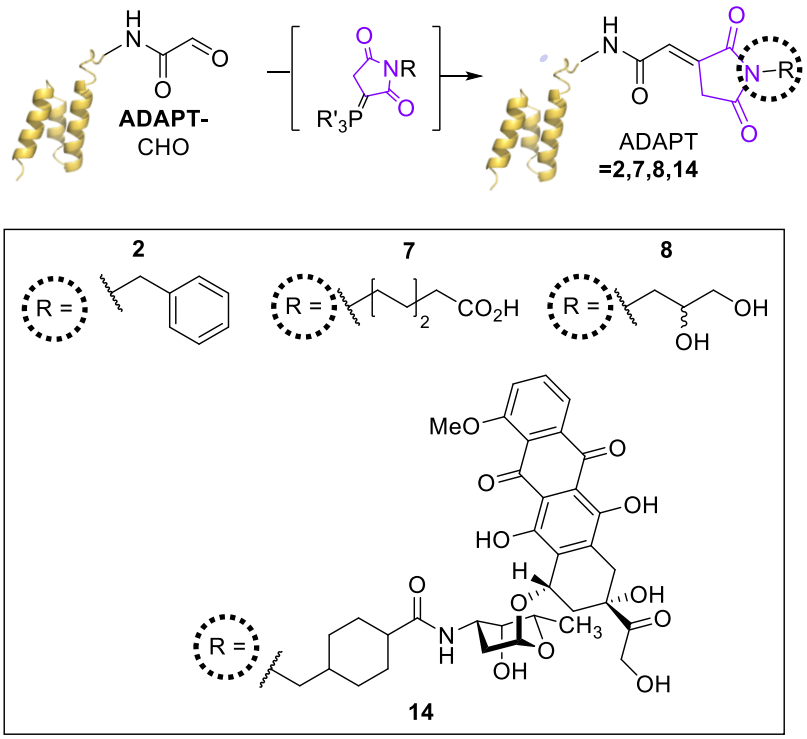
6.1.3 Ylide formation





A stock solution of maleimide (19.8µmol) in DMSO (268µL, c=73.9mM) was diluted with 2.41 mL of a solution of TCEP hydrochloride (**1**, 23.4mg, 81.6µmol) and triethylamine (1.2µL, 8.2µmol) in PBS (10mL, 10mM, pH=7.4), thus producing a solvent ratio of 9:1 (PBS/DMSO), an equimolar ratio between maleimide and TCEP hydrochloride and 10mol% triethylamine. The mixture was subsequently shaken for one hour, whereupon it was frozen at -20°C until immediately prior to its use in the subsequent Wittig reaction.

6.1.4 Wittig reactions



An aliquot of **ADAPT6-CHO** (102µL, c = 0.98mg/mL, 13.7nmol) in PBS (10mM, pH=7.4) was pipetted into a tainted glass vial, diluted with PBS (10mM, pH=7.4, 136µL), DMSO (26µL) and a solution of phosphonium ylide in aqueous DMSO (186µL, 1.37µmol, 100eq.) as prepared above was added. The reaction mixture was then gently agitated and incubated at 37°C without stirring for 24 hours. The reaction mixture was then frozen at -80°C and only thawed immediately before analysis.

Conversions were calculated based on AUCs in the deconvoluted mass spectra as shown in section 6.1.6.[6] In all but one case, the mass spectrometer was run in positive mode and a gradient involving aqueous formic acid was employed. ADAPT=**14** was found to ionize badly under these conditions, which is why in this case, the mass spectrometer was instead run in negative mode with a gradient involving aqueous ammonia. For further details on the characterization of the conjugates by LC/MS, please refer to section 1 (General remarks).

[6] K. Palla, et al., *J. Am. Chem. Soc.*, **2015**, *137*, 1123-1129.

6.1.5 SDS-PAGE

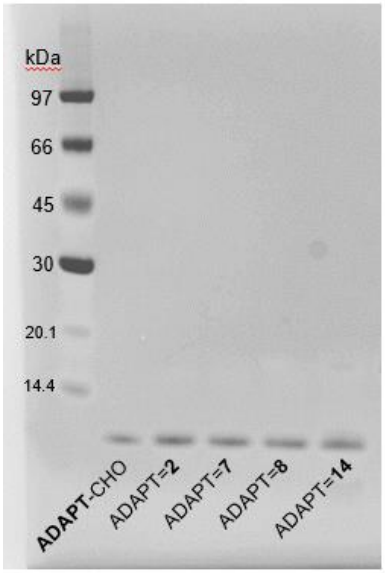
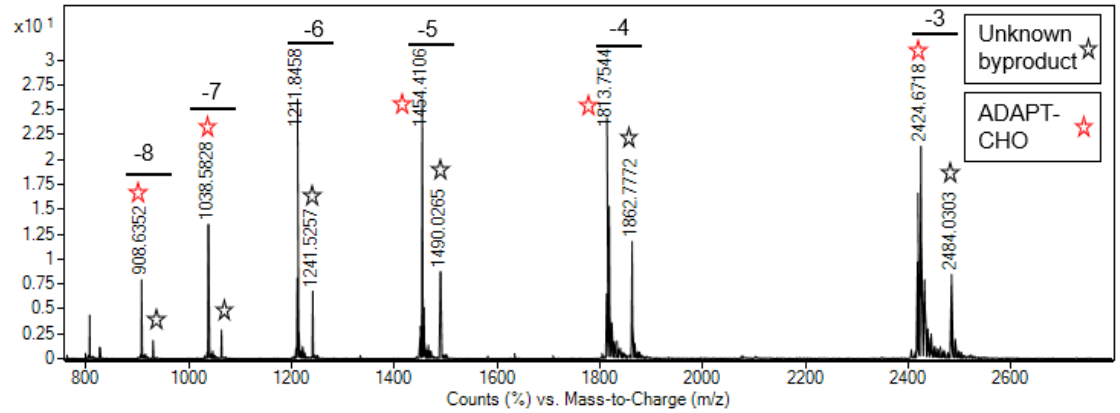
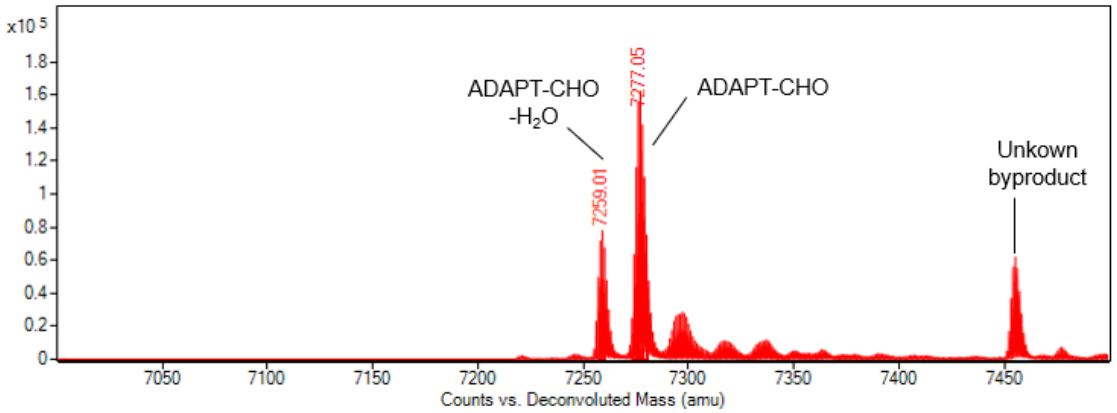


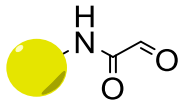
Image S10

SDS-PAGE analysis of
ADAPT-bioconjugates

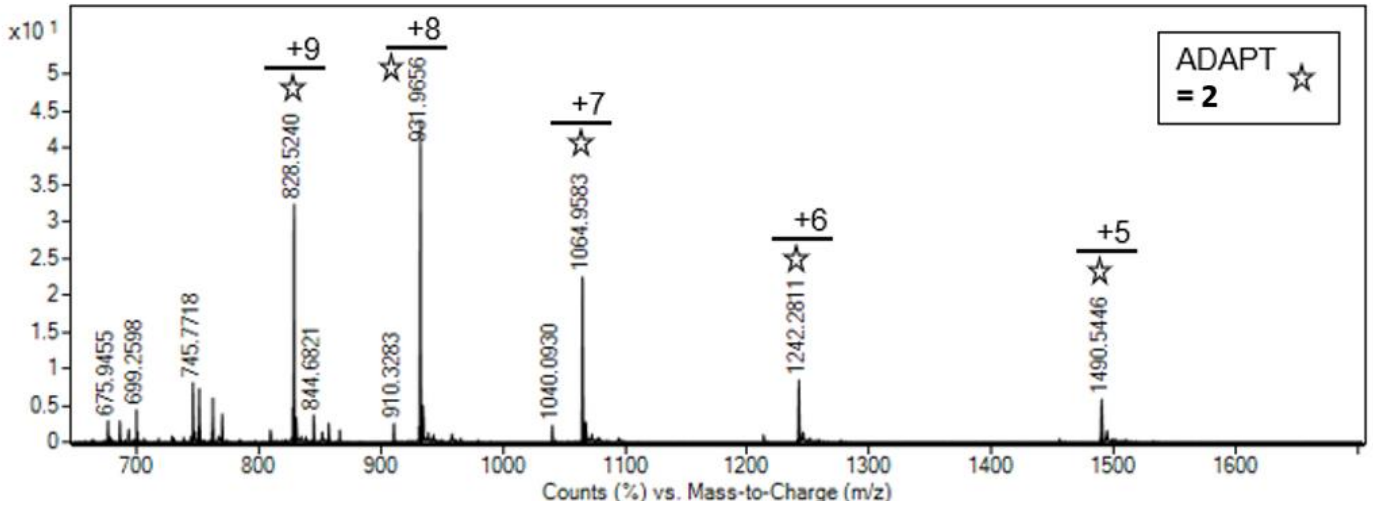
6.1.6 Mass spectra

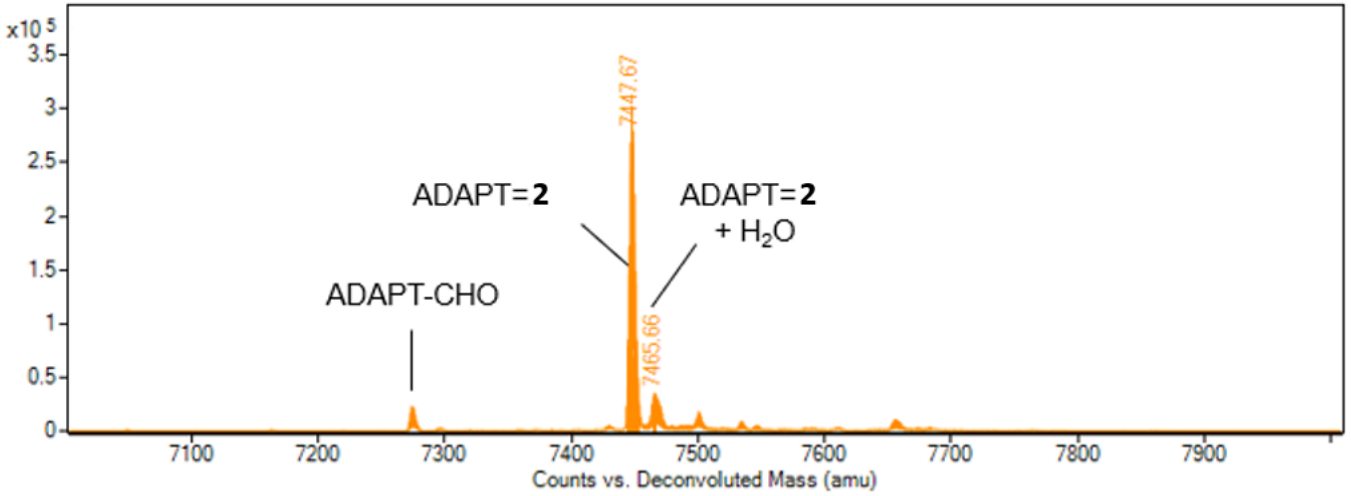


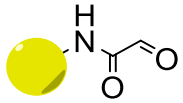




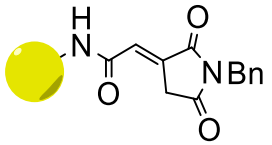
ADAPT-CHO
7277



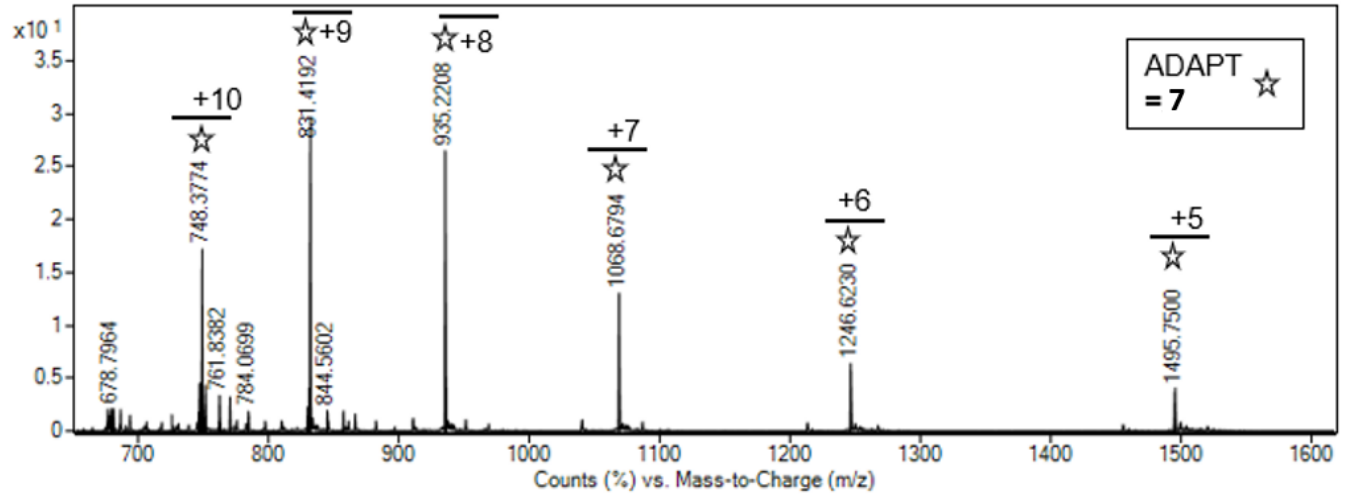


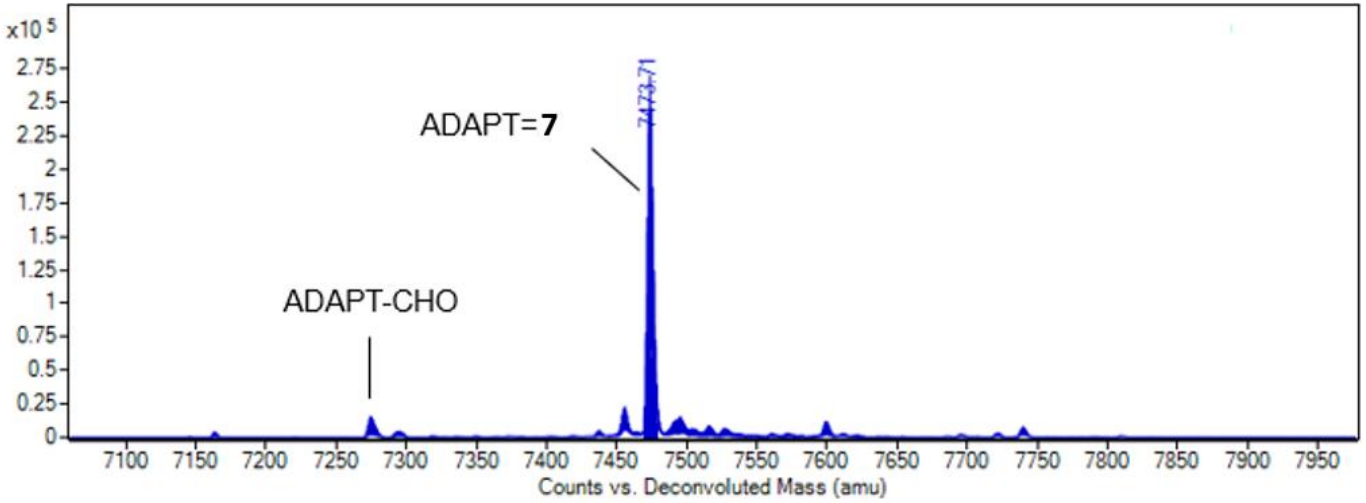


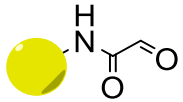
ADAPT-CHO
7277Da



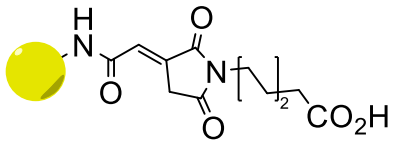
ADAPT=**2**
7448Da



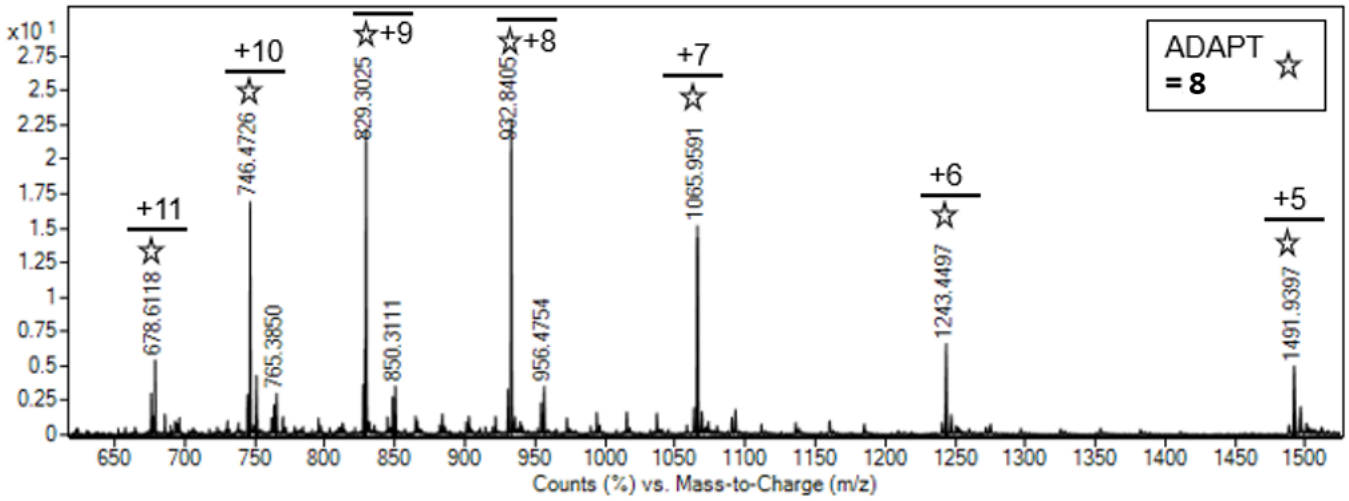


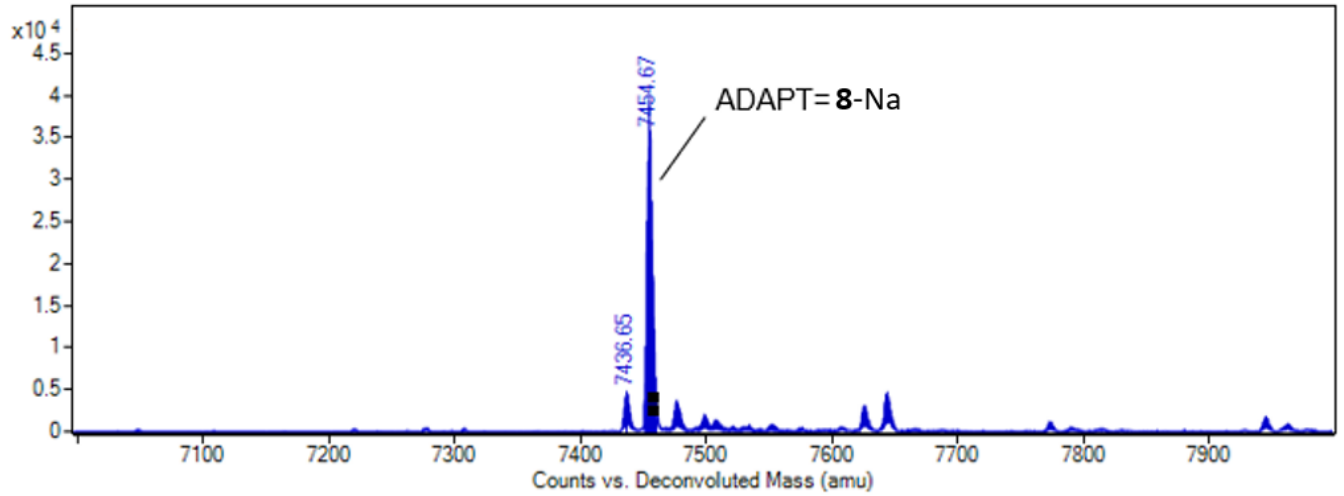


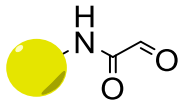
ADAPT-CHO
7277Da



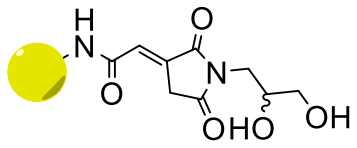
ADAPT=**7**
7472Da



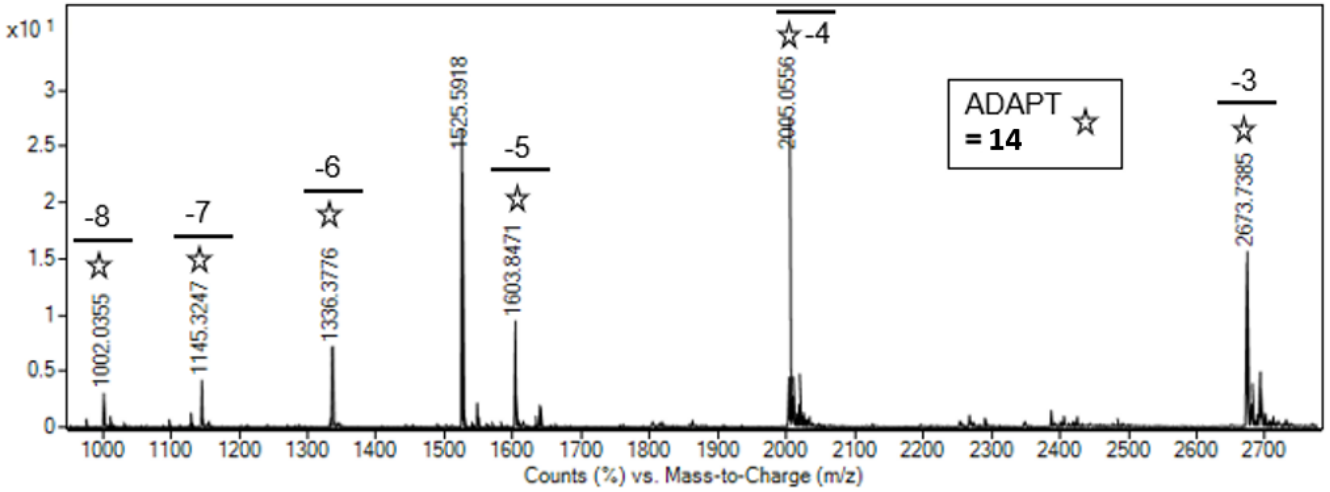


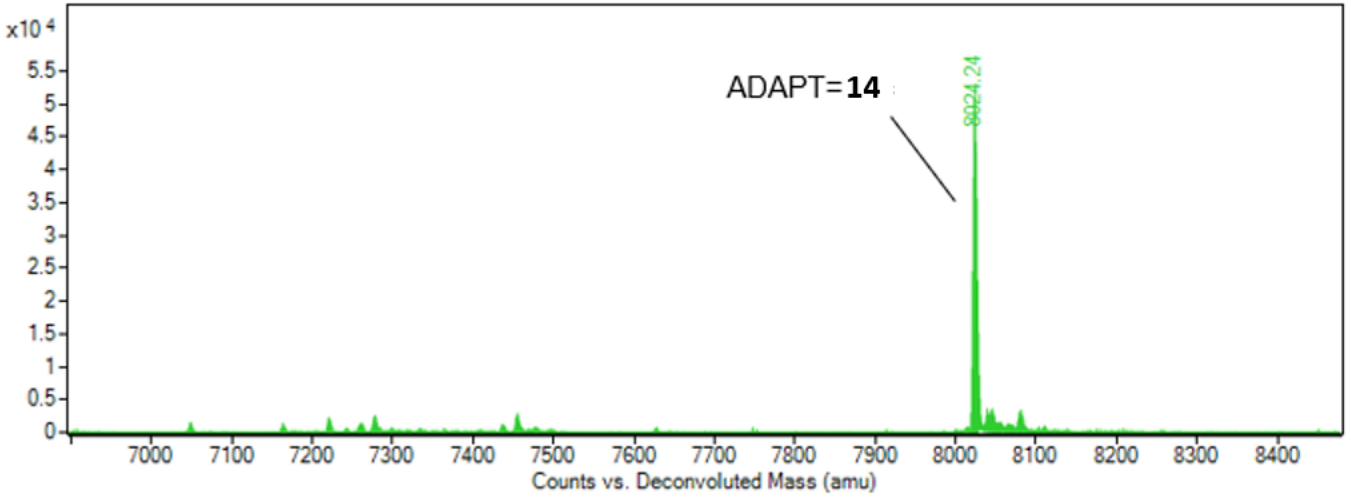


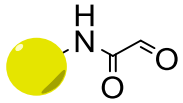
ADAPT-CHO
7277Da



ADAPT=**8**
7432/7454Da (Na-adduct)

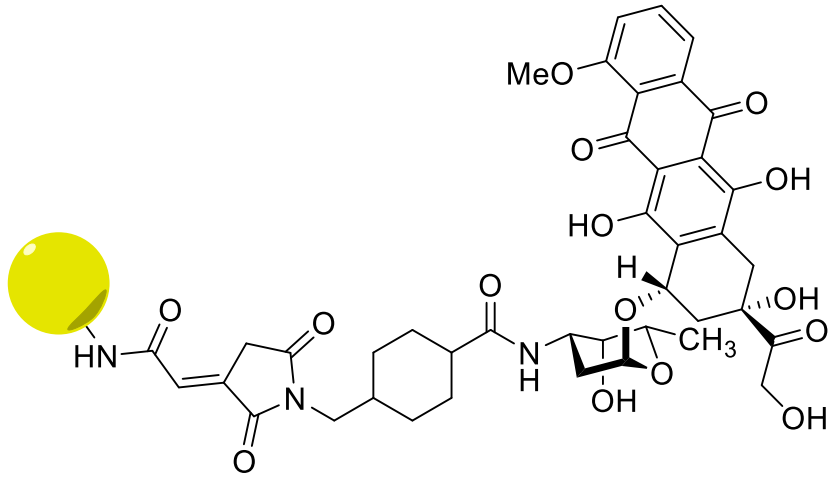






ADAPT-CHO
7277Da

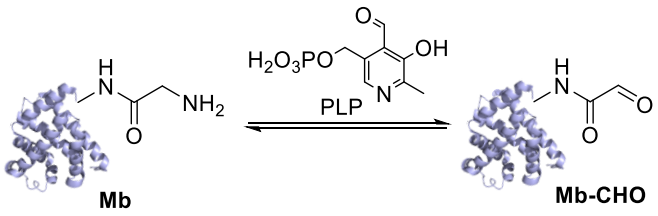
ADAPT=**14**
(see structure on next page)
8023Da



ADAPT=**14**

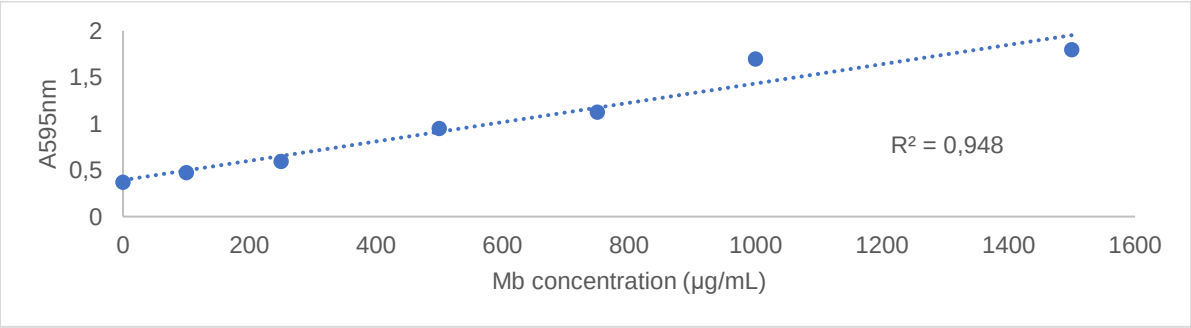
6.2 Myoglobin conjugates

6.2.1 Protein oxidation



To a pellet of equine heart myoglobin (**Mb**, 12.8mg, 755nmol) was added aqueous PBS buffer (25mM phosphate, pH=6.5, 6mL) to yield a cloudy solution (c=126µM). After 20 seconds of gentle agitation, the mixture was centrifuged and 4mL of the supernatant were transferred to a tainted 15mL vial. The now slightly more acidic solution was basified to pH=6.5 by the addition of aqueous NaOH (4M). In the meantime, a suspension of pyridoxal-5'-phosphate (**PLP**, 53.0mg, 200µmol) in aqueous PBS buffer (25mM phosphate, pH=6.5, 6mL) was prepared and the pH was adjusted back to 6.5, yielding a solution. The entirety of the PLP-solution (200µmol, 398eq.) was added to the **Mb**-solution (503nmol, 1.00eq.) to yield a combined volume of approximately 10mL and a protein concentration of 50µM. The vial was incubated in an oil bath at 37°C without stirring for six hours. The protein was thereupon desalted by means of ultracentrifugation (Amicon Ultra centrifugation tubes, MWCO=3kDa), fresh PBS-buffer was added, and the process was repeated three times. The protein solution was diluted back to 4mL and 6mL of a freshly made **PLP**-solution identical to the one above was added. The protein was oxidized for an additional 6 hours, whereupon it was desalted again in a manner identical to the one outlined above. The desalted protein was diluted to a volume of 8mL using aqueous PBS (10mM phosphate, pH=7.4), divided into aliquots of 500µL and frozen at -80°C.

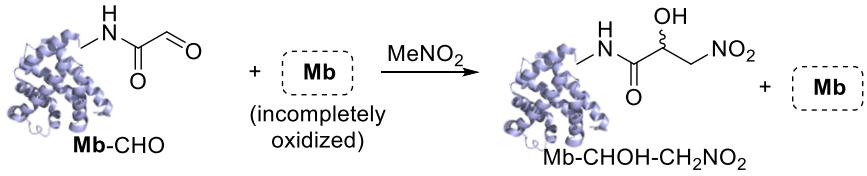
Determination of protein concentration (Bradford assay). A stock solution of unmodified **Mb** in PBS (10mM, pH=7.4, c = 1.50mg/mL) was prepared. A dilution series was prepared with the following concentrations: 1.50mg/mL (undiluted), 1.00mg/mL, 0.75mg/mL, 0.50mg/mL, 0.25mg/mL, 0.10mg/mL, 0mg/mL (plain PBS). To 50µL of each protein solution were added 1.5mL of Coomassie Brilliant Blue R-250 stain, the resulting solutions were agitated and incubated at room temperature for 15 minutes, and their UV-absorption at λ=595nm was recorded in triplicates using a cuvette (l=1cm). The following standard curve was obtained (average values per concentration shown):[7]



An aliquot of **Mb-CHO** was warmed to room temperature, gently agitated, and five 50µL samples were diluted with Coomassie Brilliant Blue R-250 stain (1.5mL each). After 15 minutes of incubation, their absorption at λ=595nm was recorded to reveal a protein concentration of 1.10±0.02mg/mL, indicating roughly quantitative recovery.

[7] Note: The concentrations on the horizontal axis of the graph correspond to those of the protein sample prior to dilution with Coomassie stain.

6.2.2 Henry reaction with $MeNO_2$



An aliquot of oxidized **Mb**-solution with an unknown aldehyde content (91µL, c = 1.10mg/mL, 5.9nmol) in PBS (10mM, pH=7.4) was diluted with a mixture of DMSO (18.3%) in PBS (10mM, pH=7.4, total volume 190µL) so as to yield a DMSO-concentration of 10% and a protein concentration of 0.50mg/mL (30µM). Nitromethane (0.6µL, 12µmol, 2000eq.) was added, the resulting solution was gently agitated and then incubated at 37°C without agitation. After 24 hours, the still homogenous reaction mixture was frozen at -80°C and only thawed immediately before analysis.

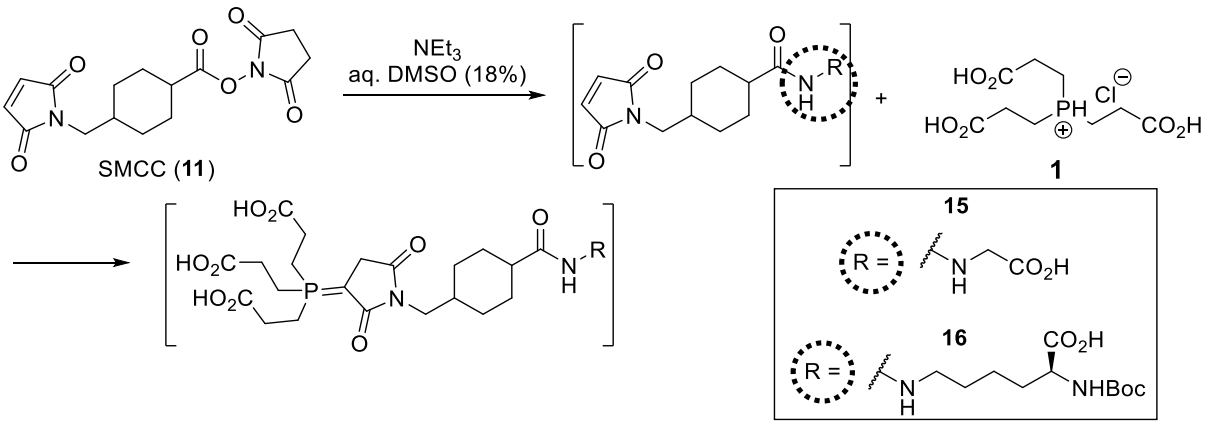
The content of unmodified **Mb** in the reaction mixture was found to be approximately 38%, which was determined based on relative AUCs in the deconvoluted mass spectrum (see below). In order to ensure that both **Mb** and **Mb-CHOH-CH_2NO_2** had fully eluted in the period of time for which the TIC (total ion chromatogram) was integrated, the EICs (extracted ion chromatograms) of each protein were consulted after deconvolution and the integration was adjusted where necessary.

The Henry reaction was generally performed simultaneously to the Wittig reactions (see below) and their analyses were executed sequentially. In order to determine the conversions in the case of the Wittig reactions, only the ratio of AUCs of starting material relative to product was used. No quantification *per se* was attempted, meaning that no calibration of the mass spectrometer was necessary and variations in its sensitivity between injections did not affect the outcome.

6.2.3 Ylide formation

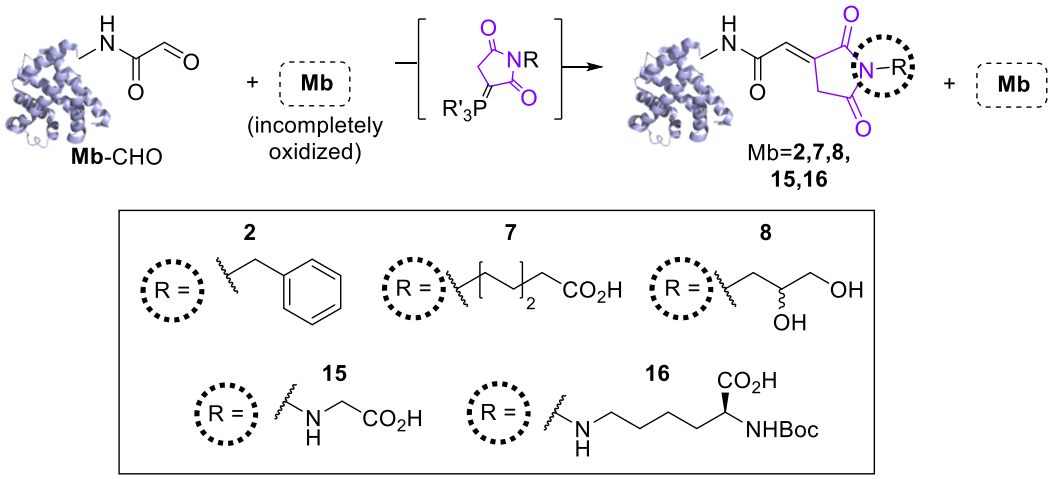
Derivatization of pure maleimides (compounds **2, 7, 8**) to phosphonium ylides: See section 6.1.3.

SMCC-derivatization of amino acids, followed by formation of phosphonium ylide *in situ*



A glass vial was charged with SMCC (**11**, 54.6mg, 12.5µmol) and the respective amino acid (12.5µmol). A solution of DMSO (1.8mL) and triethylamine (1.73µL, 12.5µmol) in aqueous $NaHCO_3$-buffer (100mM bicarbonate + 100mM NaCl + 150mM Na-caproate) with a total volume of 10mL was prepared. 0.5mL of the aqueous DMSO-solution was added to the vial containing SMCC and the amino acid and the resulting suspensions were shaken overnight. Another set of vials was then charged with TCEP hydrochloride (14.3mg, 5.00µmol) and a portion (200µL, corresponding to 5.00µmol of maleimide) of the SMCC-solution was added to each vial. The remainder of the SMCC-reaction mixtures were frozen at this point. The vials containing TCEP, SMCC and amino acid were shaken for 30 minutes until complete dissolution had been achieved. The ylide solutions were frozen at -20°C until immediately prior to its use in the subsequent Wittig reaction.

6.2.4 Wittig reactions

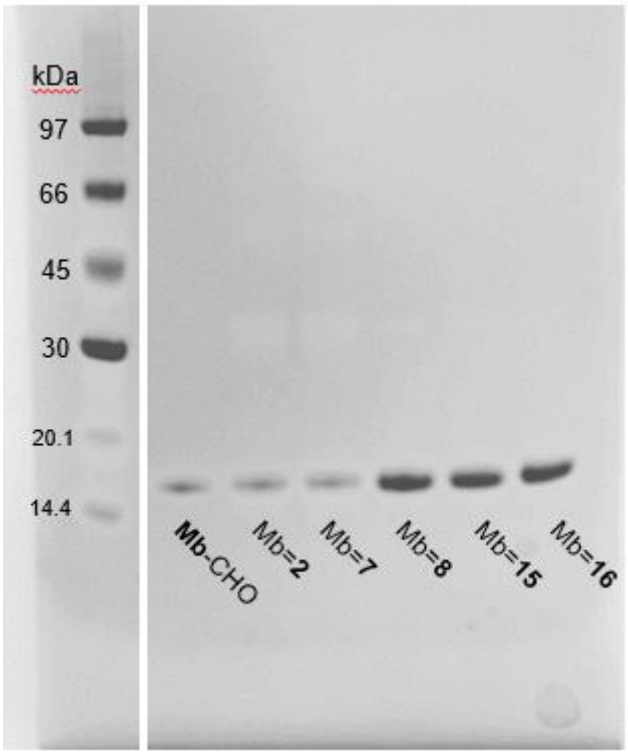


An aliquot of oxidized **Mb**-solution (91µL, c = 1.10mg/mL, 5.9nmol) in PBS (10mM, pH=7.4) was diluted with a mixture of DMSO (18%) in PBS (10mM, pH=7.4) and a solution of phosphonium ylide in DMSO (18%) as prepared above. In reactions with 20 molar equivalents of ylide, 104µL of the aqueous DMSO solution were added, followed by 4.7µL of ylide solution (25mM, 120nmol, 20eq.). In reactions with 100 molar equivalents of ylide, 85µL of the aqueous DMSO solution were added, followed by 24µL of ylide solution (25mM, 600nmol, 100eq.). In either case, the total volume of the reaction mixture was approximately 200µL. The reaction mixtures were then gently agitated and incubated without stirring for 24 hours. To reactions with 20eq. of ylide were added another 4.7µL of freshly thawed ylide solution (25mM, 120nmol, 20eq.) after six hours. All reaction mixtures were frozen at -80°C and only thawed immediately before analysis.

Conversions were calculated based on AUCs in the deconvoluted mass spectra (see section 6.2.6) relative to the AUCs obtained for the Henry reference (see section 6.2.2).[8] In all cases, the mass spectrometer was run in positive mode and a gradient involving aqueous formic acid was employed. For further details on the characterization of the conjugates by LC/MS, please refer to section 1.

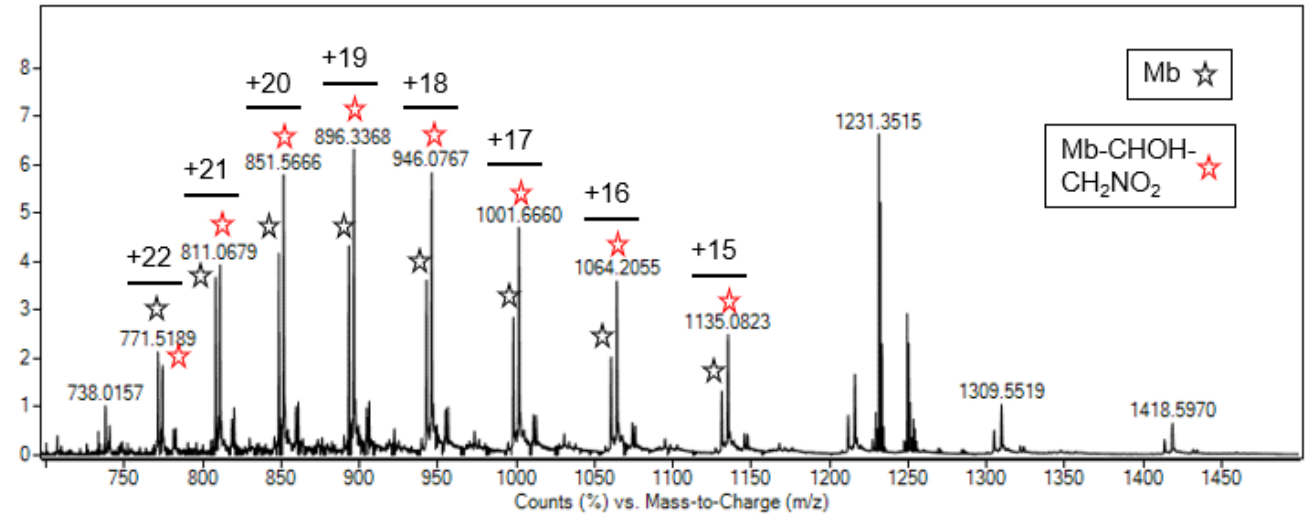
[8] Palla, K., et al., *J. Am. Chem. Soc.*, **2015**, *137*, 1123-1129.

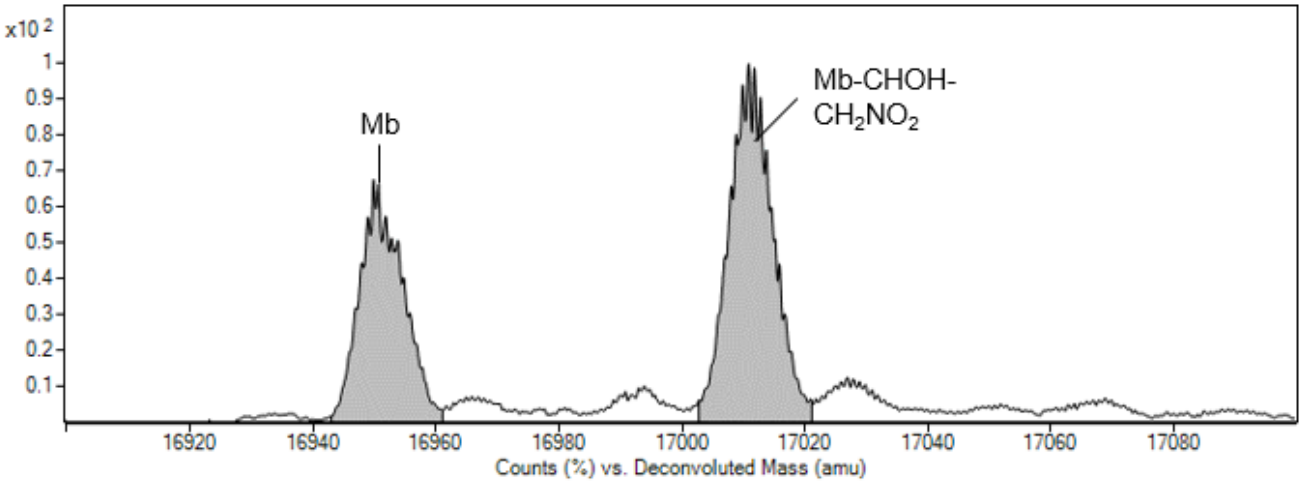
6.2.5 SDS-PAGE

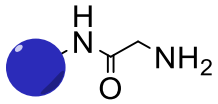


SDS-PAGE analysis of
Mb-bioconjugates

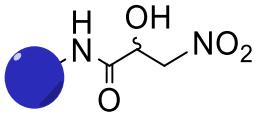
6.2.6 Mass spectra



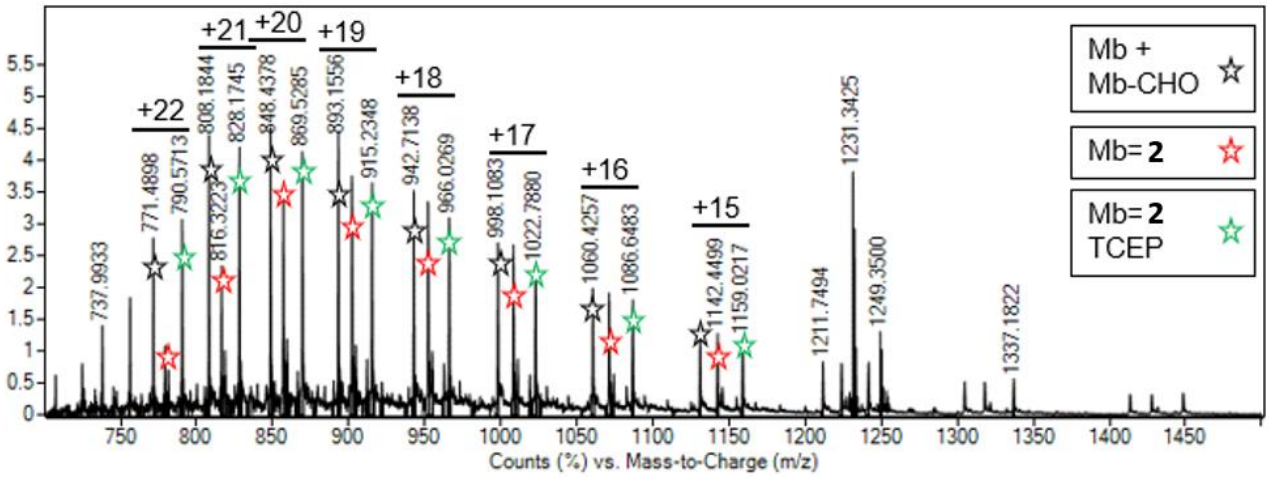


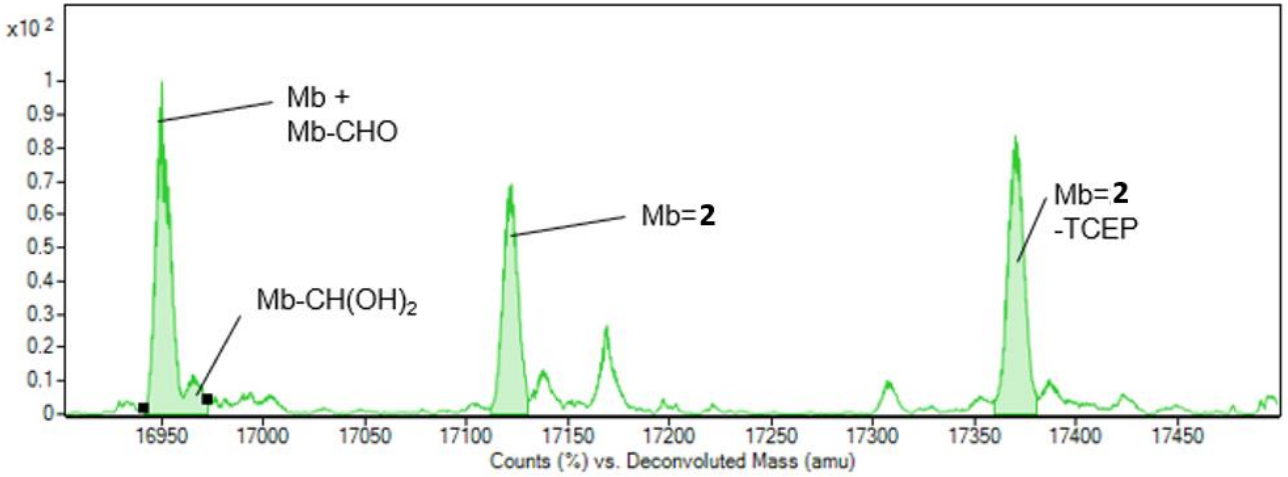


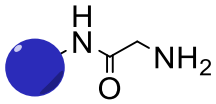
Mb
16950Da



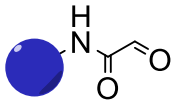
Mb-CHOH-**CH_2NO_2**
17010Da



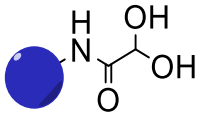




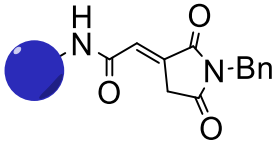
Mb
16950Da



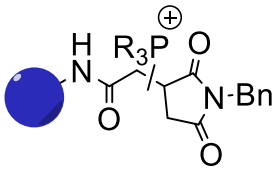
Mb-CHO
16949Da



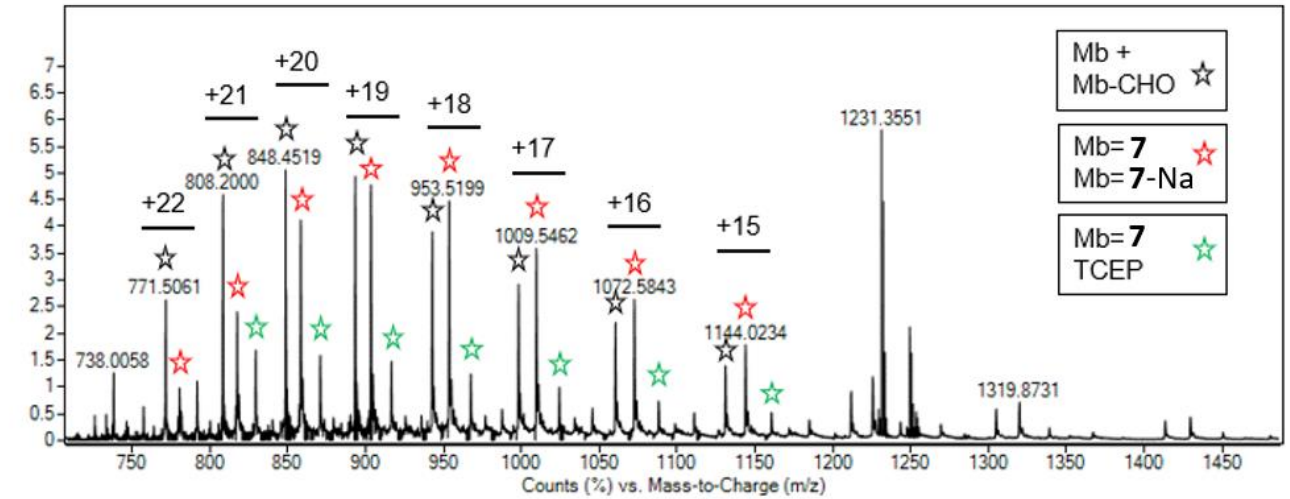
Mb-CH(OH)$_2$
16967Da

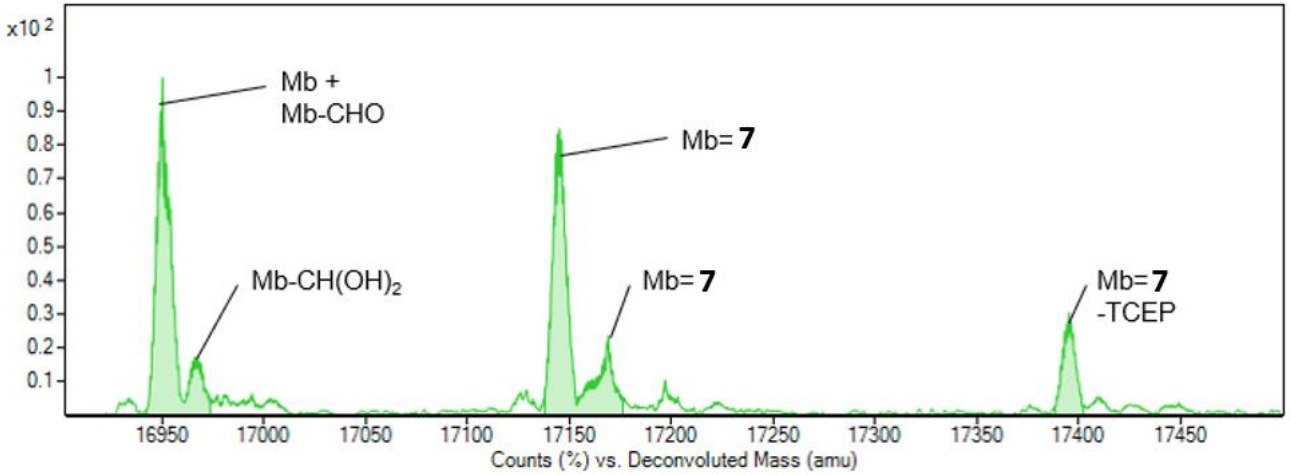


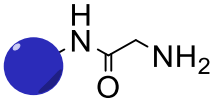
Mb=**2**
17121Da



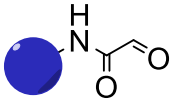
Mb=**2**-TCEP
17371Da



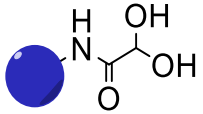




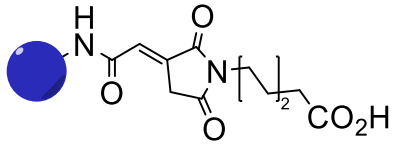
Mb
16950Da



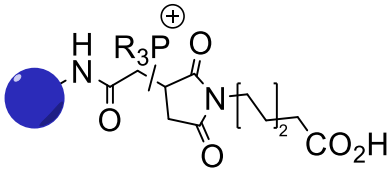
Mb-CHO
16949Da



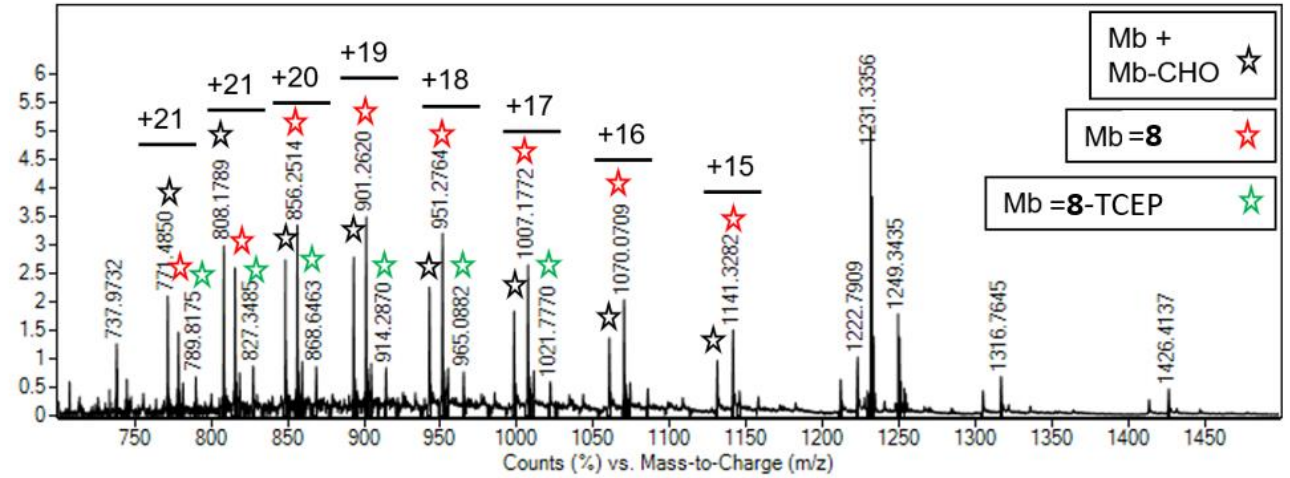
Mb-CH(OH)$_2$
16967Da

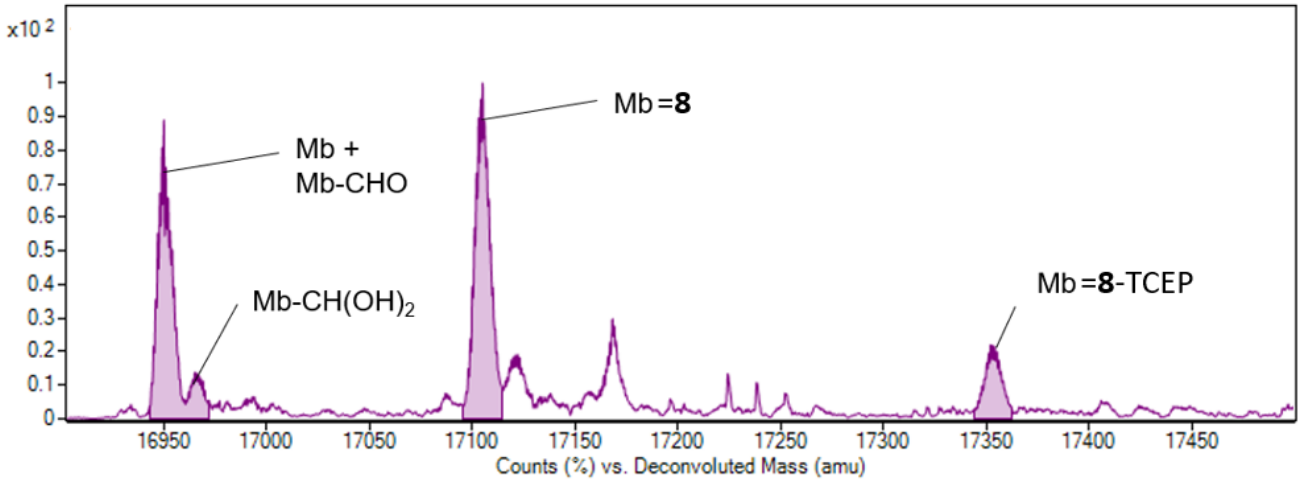


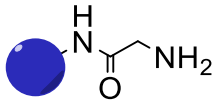
Mb=**7**
17145Da/ 17168Da (Na-adduct)



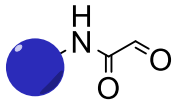
Mb=**7**-TCEP
17395Da



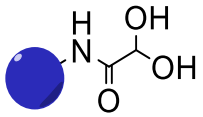




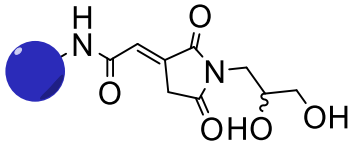
Mb
16950Da



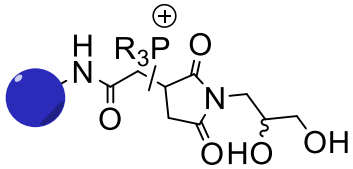
Mb-**CHO**
16949Da



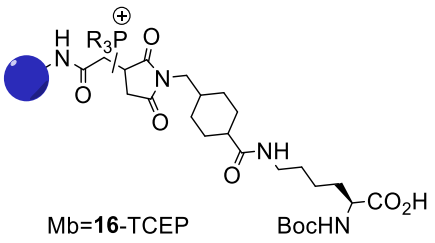
Mb-CH(OH)$_2$
16967Da



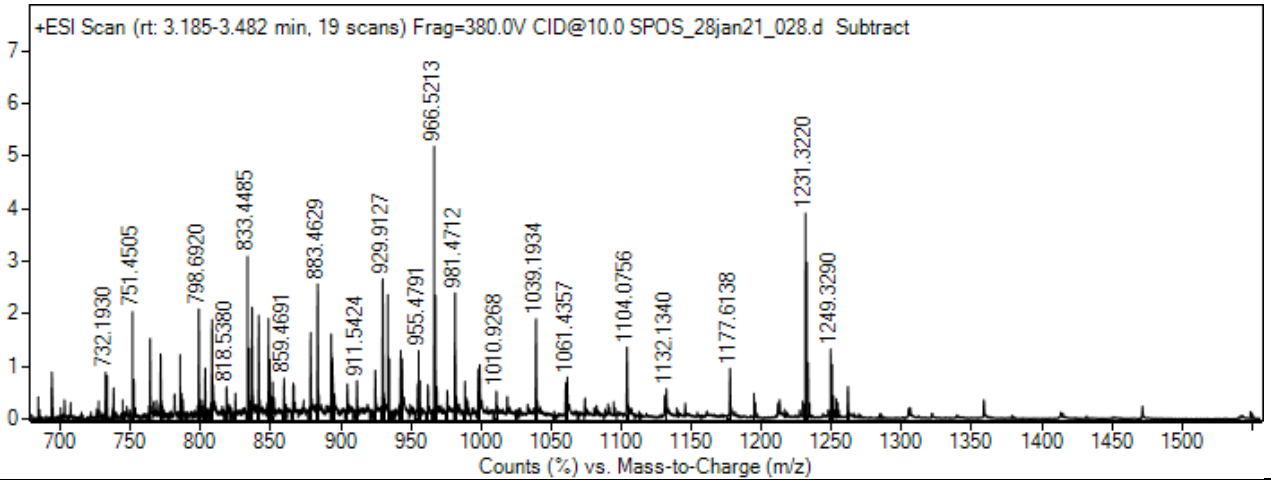
Mb=**8**
17105Da



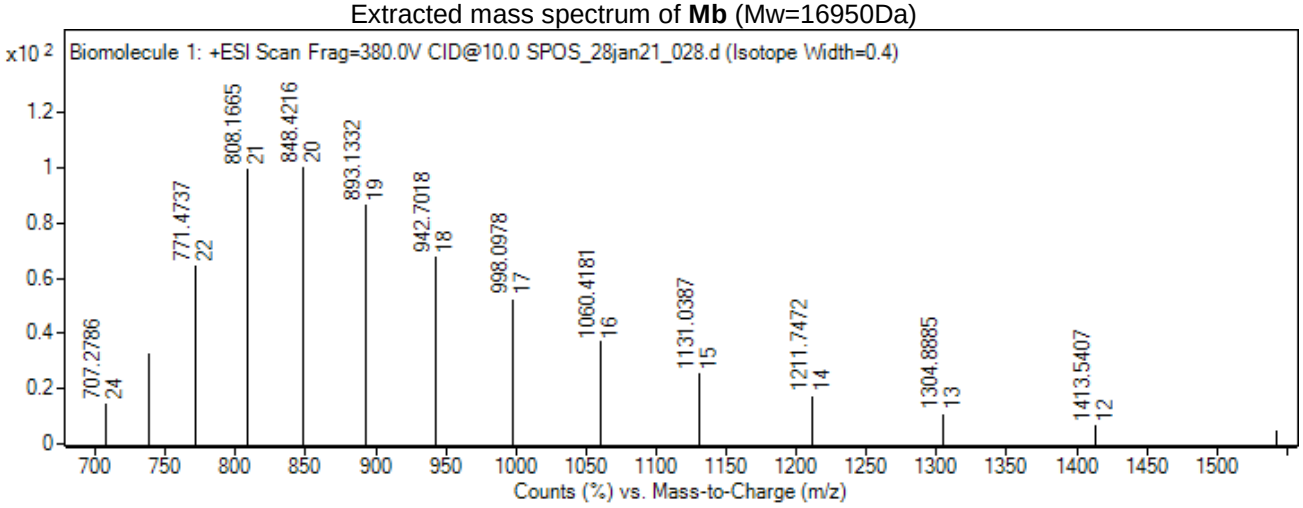
Mb=**8**-TCEP
17355Da



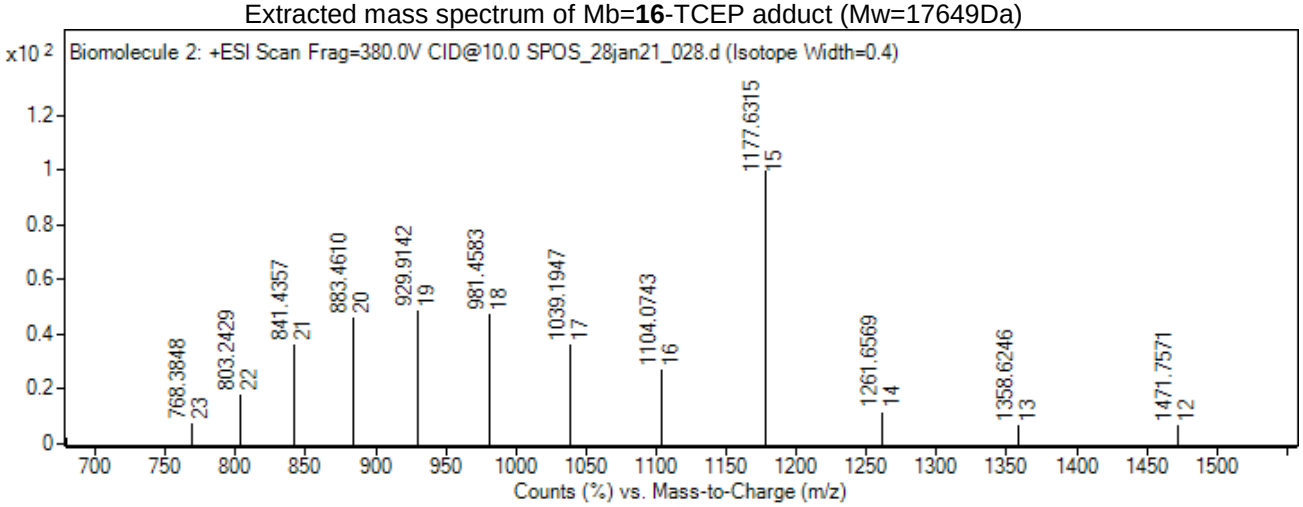
Raw data:[9]



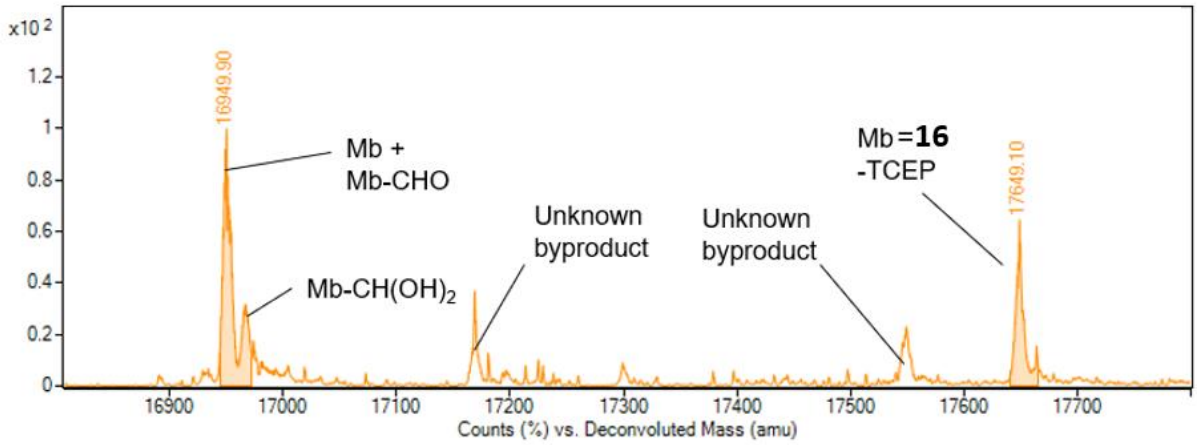
Extracted mass spectrum of **Mb** (Mw=16950Da)

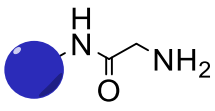


Extracted mass spectrum of Mb=**16**-TCEP adduct (Mw=17649Da)

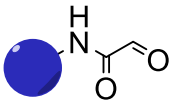


[9] For this particular compound, the Wittig reaction was relatively low yielding, resulting in the formation of two unidentified byproducts (see next page). As a consequence, the mass spectra are instead displayed in the format shown above (with the starting material's and product's extracted mass spectra stacked below the raw data file instead of picking the corresponding peaks within) for better visibility.

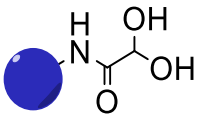




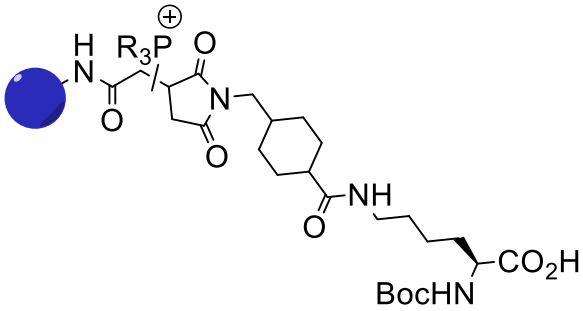
Mb
16950Da



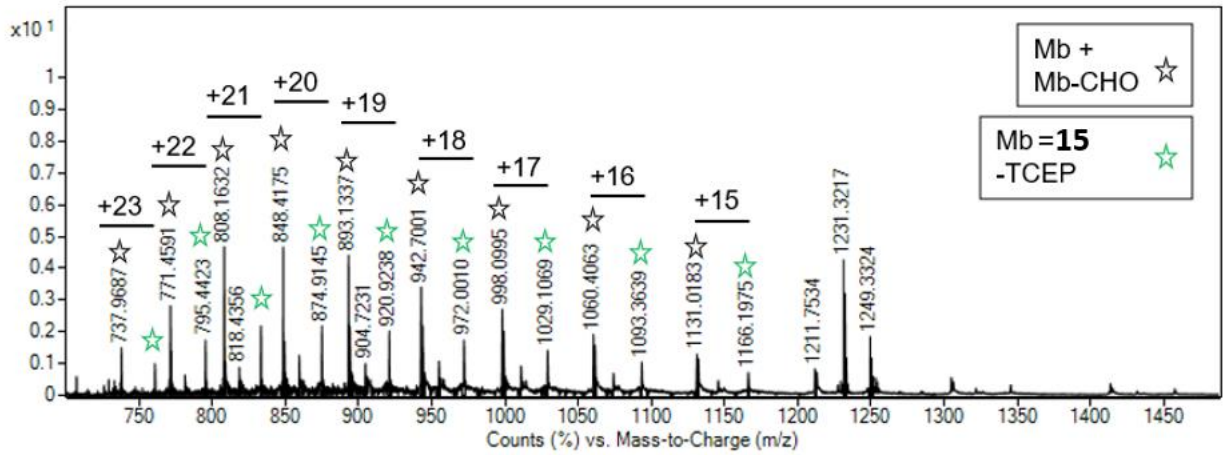
Mb-CHO
16949Da

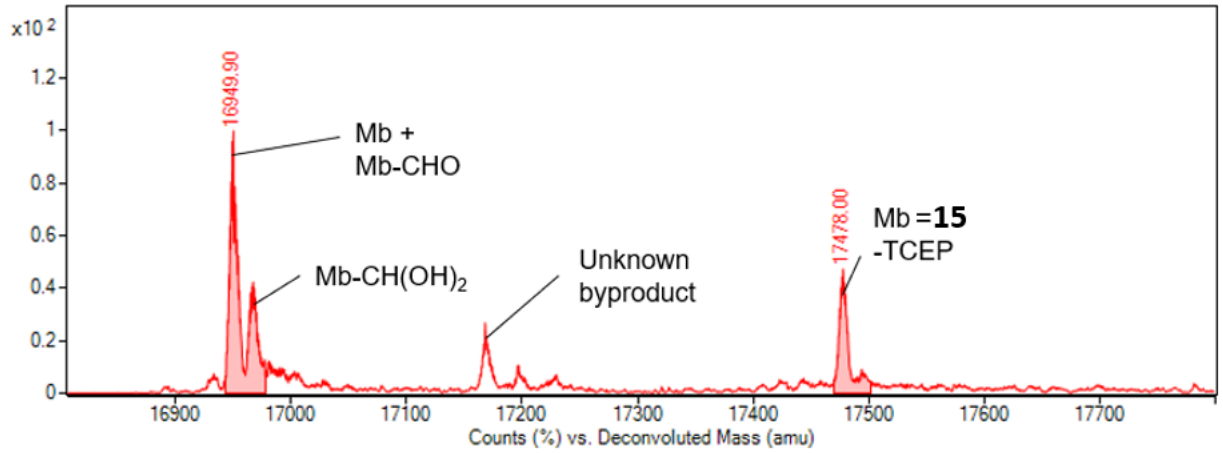


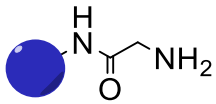
Mb-CH(OH)$_2$
16967Da

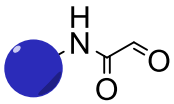


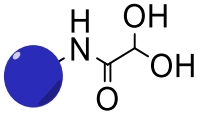
Mb=**16**-TCEP
17649Da

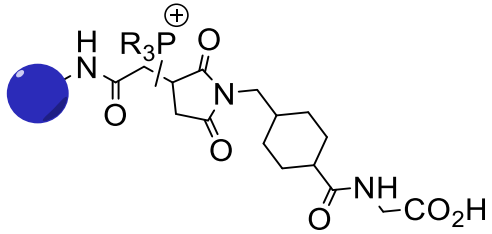





Mb
16950Da


Mb-CHO
16949Da


Mb-CH(OH)$_2$
16967Da


Mb=**15**-TCEP
17477Da

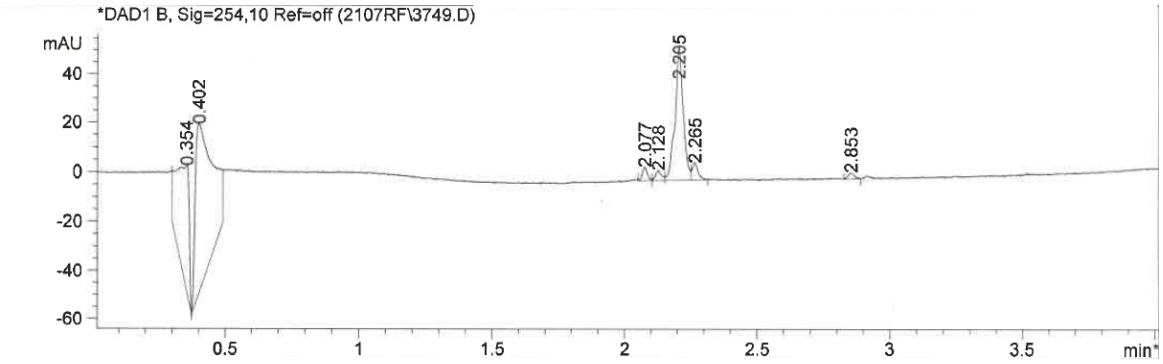
6.3 Biological evaluation of ADAPT=14

SMCC derivatized doxorubicin (**14**) was synthesized, isolated (see page 26) and transformed into the corresponding TCE-phosphonium ylide (see page 56) as described above. This preparative experiment was conducted on the following scale: **ADAPT**-CHO (2.05mg, 281nmol, 1.0eq.) was reacted with ylide derived from maleimide **14** (21.4mg, 28.1µmol, 100eq.) in a total of 9.4mL of solvent (10mM PBS/DMSO 9:1). The mixture remained homogenous throughout the entire 24 hour reaction time (see image S9).

Image S9

After 24 hours of reaction time, the homogenous reaction mixture was directly purified by preparative HPLC (25 to 50% gradient of acetonitrile in 0.1% aqueous TFA over 10 minutes). For more information about the HPLC set-up, please refer to section 1 (General remarks).

These conditions presumably denatured the protein, but other ADAPT-conjugates have been shown to refold upon chemical or thermal denaturation.[10] Product containing fractions were pooled and freeze-dried to yield ADAPT=**14** (1.91mg, 238nmol) as a faintly red solid (85% isolated yield) in 90% purity (see chromatogram below).



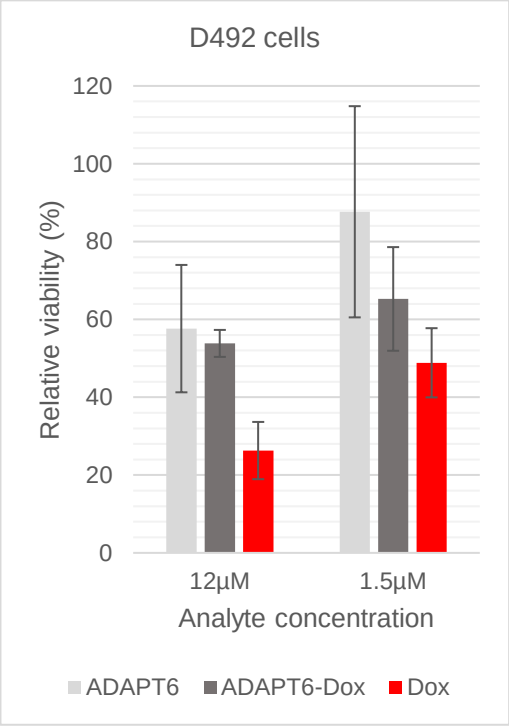
ADAPT=14 (Rt=2.205min) after purification by preparative HPLC (λ=254nm).

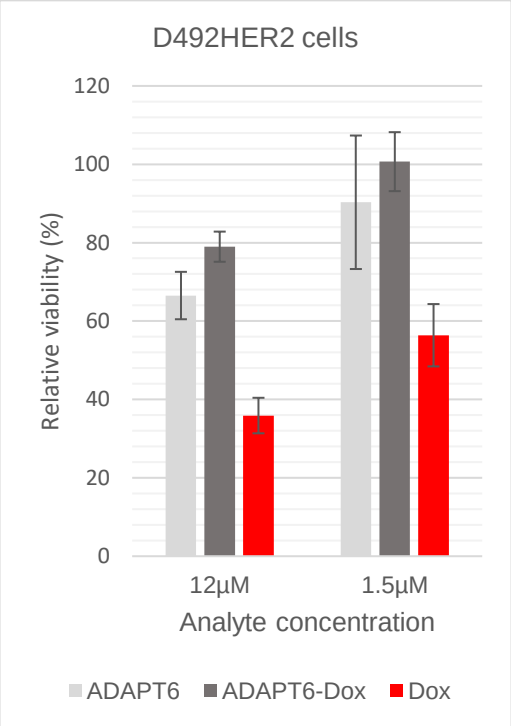
[10] Garousi, S., et. al., *Cancer Res.*, **2015**, *75*, 4364–4371.

The conjugate was taken up in 0.3M aqueous acetic acid (2mL) and the mixture was gently homogenized with the help of a micropipette. It was then diluted with aqueous phosphate buffered saline (10mL) and the pH was adjusted to 7.4 using 4M sodium hydroxide solution. The solution was then diluted with more phosphate buffered saline to yield a total volume of 20mL (c=12µM).

The conjugate's cytotoxicity in D492 and D492HER cells was then determined as follows: The cells were cultured in H14 media supplemented with penicillin (100U/mL) and streptomycin (100µg/mL) in culture flasks coated with collagen I as described previously.[11]

Cells were seeded in 96 well plates at a density of 10'000 cells/well and cultivated overnight prior to addition of the analyte. ADAPT=**14**, unoxidized **ADAPT** or doxorubicin were added at concentrations of 12µM or 1.5µM and the treated cells were incubated for 24 hours. PrestoBlue™ viability indicator was added to each well, and after two hours, the absorptivity at λ=570nm and 590nm was determined. Cell viability was then compared to untreated cells. The data (including 95% confidence intervals), which is based on four technical replicates, are depicted below.





[11] Blaschke, R. J., et. al., *Methods Enzymol.*, 1994, **245**, 535–556.