Electronic Supporting Information

Sequence-encoded Macromolecules with Increased Data Storage Capacity through a Thiol-epoxy Reaction

Matthieu Soete, Chiel Mertens, Resat Aksakal, Nezha Badi, Filip Du Prez*

Polymer Chemistry Research group, Centre of Macromolecular Chemistry (CMaC), Department of organic and Macromolecular Chemistry, Faculty of Sciences, Ghent University, Krijgslaan 281 S4-bis, B-9000 Ghent, Belgium

* Corresponding author: E-mail: Filip.DuPrez@UGent.be

Table of contents

Materials

All chemicals were used as supplied, unless otherwise stated. Deionised water was used in the procedures. Tetrahydrofuran was purified by filtration over activated basic aluminium oxide to remove organic peroxides. DMSO-d_6 and $CDCl_3$ (≥99.8%) were purchased from Eurisotop. Homocysteine thiolactone hydrochloride (99%) was purchased from Haihang Industry (Jinan City, China). Magnesium sulphate (dried ≥99%), sodium hydrogen carbonate (99%) and sodium chloride (99%) were purchased from Carl Roth. Hydrochloric acid (36 wt%) was purchased from Chem Lab NV. Trifluoroacetic acid (peptide grade) and Fmoc-protected rink-amide resin (100–200 mesh, 1% DVB, 0.4 mmol g^{-1}) were purchased from Iris Biotech GmbH. Dichloromethane (HPLC grade, 99.5%), anhydrous chloroform (99.8%), N,N′-dimethylformamide (HPLC grade, 99.8%), piperidine (anhydrous, 99%), 1,4-dioxane (HPLC grade, 99%), pyridine (anhydrous, 99%), triethylamine (99%) and methanol (HPLC grade, 99.5%) were purchased from Acros Organics. Dimethylphenylphosphine (99%), N,N′-dicyclohexylcarbodiimide (99%), glutaric anhydride (95%), zirconium(IV) acetylacetonate (97%), 4-dimethylaminopyridine (99%), Phenethylamine (99%), Benzylamine (99%), Allylamine (99%), Pyrrolidine (99%), 2-methoxyethylamine (99%), octylamine (99%), copper sulfate pentahydrate (>98%), (+)-sodium *L*-ascorbate (>98%), *N*,*N*-Dimethylethylenediamine (95%), Phenyl glycidyl ether (99%), (±)-Propylene oxide (99%), ethyl acetate (HPLC grade, 99.7%) and chloroform (HPLC grade, 99.9%) were also purchased from Sigma Aldrich. Triphosgene (98.0%), *n*-butylamine (99%)Acetone (HPLC grade, 99.5%), ethanol (absolute HPLC grade), acetonitrile (HPLC grade, 99.9%), diethyl ether (HPLC grade, 99%), 4-fluorobenzylamine (98%) and *N,N′*-dimethylformamide (HPLC grade, 99.7%) were purchased from Fischer Scientific. Propargylamine (97%), 3-morpholinopropylamine (98%), 2-morpholinoethylamine (98%), 4-methoxyphenethylamine (98%), 1,2-epoxyhexane (96%), benzyl glycidyl ether (97%), 1,2-epoxydecane (96%) and 1,2-epoxyoctane (96%) were purchased from TCI Europe**.**

Instrumentation

Nuclear magnetic resonance (NMR) spectroscopy

All ^{1}H and ^{13}C spectra were recorded in DMSO-d_6 or $CDCl_3$, with a Bruker Avance 300 (300 MHz) or a Bruker Avance 400 (400 MHz) device. The assignment of the ^{1}H NMR spectra of **M1-M3** was aided by COSY analysis.

Electrospray Ionization Mass Spectrometry (ESI-MS) and Liquid Chromatography Mass Spectrometry (LC-MS)

An Agilent technologies 1100 series LC/MSD system equipped with a diode array detector and single quad MS detector (VL) with an electrospray source (ESI-MS) was used for classic reversed phase LC-MS and MS analysis. Analytic reversed phase HPLC (high-performance liquid chromatography) was performed with a Phenomenex Kmetex C_{18} (2) column with a solid core at 35°C and a flow rate of 1.5 mL min^{-1} (5 µm, 250 × 4.6 mm) using a solvent gradient (0 → 100 %) acetonitrile in H_2O over 6 min, unless otherwise stated. The eluting compounds were detected via UV-detection (λ = 214 nm).

Size Exclusion Chromatography (SEC)

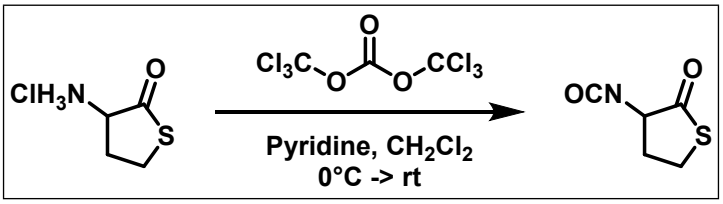
Oligomers were characterized on a Waters SEC system equipped with a Waters 1515 isocratic pump, Waters 2410 refractive index detector (24 °C), Waters 717plus autosampler and a Waters 2487 dual λ absorbance UV detector and column oven. For separation, a three-column setup was used with one SDV 3 µm, 8 × 50 mm precolumn and two SDV 3 µm, 1000 Å, 8×300 mm columns supplied by PSS, Germany. Tetrahydrofuran (THF) stabilized with butylated hydroxytoluene (BHT, HPLC-SEC grade) supplied by Biosolve was used at a flow rate 1.0 mL min^{-1}. Calibration was carried out by three injections of a mixture of narrow polystyrene standards ranging from 162 to 38640 Da. To each sample, 2 µL of toluene was added as internal standard to check the accuracy of the calibrations.

Matrix-Assisted Laser Desorption/Ionization Tandem Mass Spectrometry (MALDI-MS/MS)

For the MALDI measurements a stock solution of the matrix, *trans*-2-[3-(4-*tert*-butylphenyl)-2-methyl-2-propenylidene]malonitrile (DCTB, 30 mg/mL) and sodium trifluoroacetate (NaTFA, 10 mg/mL) were prepared in tetrahydrofuran. Samples were solubilized in tetrahydrofuran (10 mg/mL). 45 µL of the matrix solution, 15 µL of the salt and 15 µL of the sample solution were mixed and subsequently spotted on the MALDI plate. The spots were dried on the plate at room temperature and loaded into an Applied Biosystems Sciex 4800+ MALDI-TOF/TOF analyzer, controlled by 4000 Series Explorer software (Applied Biosystems, Germany). The instrument was operated in reflective positive ion mode with delayed extraction and an acceleration voltage of 20 kV with a grid of 15.6 kV. Fragmentation (MS/MS) was performed in positive ion mode at 1 kV using the no gas option.

Synthetic Procedures

Synthesis of *α*-isocyanato-*γ*-thiolactone



Triphosgene (50 g, 168 mmol, 0.33 eq.) was dissolved in ice-cooled CH_2Cl_2 (900 mL) in a two-neck flask, placed under argon atmosphere and stirred for 15 minutes. Subsequently, *DL*-homocysteine thiolactone hydrochloride (74 g, 482 mmol, 1 eq.) was added. Next, dry pyridine (120 mL, 1.5 mol, 3.11 eq.) was added dropwise over a time span of 20 minutes using an addition funnel. After 1 hour, the ice-bath was removed and the mixture was stirred for an additional 4 hours. The aqueous work-up of the crude mixture was performed fast to avoid degradation. The crude mixture was directly filtered into a separation funnel to remove the pyridinium hydrochloride salt and was rinsed with cold CH_2Cl_2. The organic phase was washed with an ice-cooled 2 M HCl solution (500 mL), ice water (500 mL) and brine (500 mL). The water phase was extracted an additional time with 300 mL CH_2Cl_2. All the organic phases were collected and dried with magnesium sulphate and subsequently filtered over a large Büchner filter and concentrated *in vacuo*. A brown liquid was obtained that was further purified by vacuum distillation, the fraction between 83–95 °C (0.08 mbar) was collected, yielding a transparent liquid (61.04 g, 88.5%).

^{1}H-NMR (300 MHz, $CDCl_3$, δ): 4.23 (dd, 1H, J=12.6, 6.8 Hz, H_1), 3.30 (m, 2H, $H_{4,5}$), 2.64 (m, 1H, H_2), 2.11 (m, 1H, H_3). **^{13}C-NMR** (125 MHz, $CDCl_3$, δ): 203.1 (C), 127.6 (C), 62.6 (CH), 32.1 (CH_2), 27.0 (CH_2).

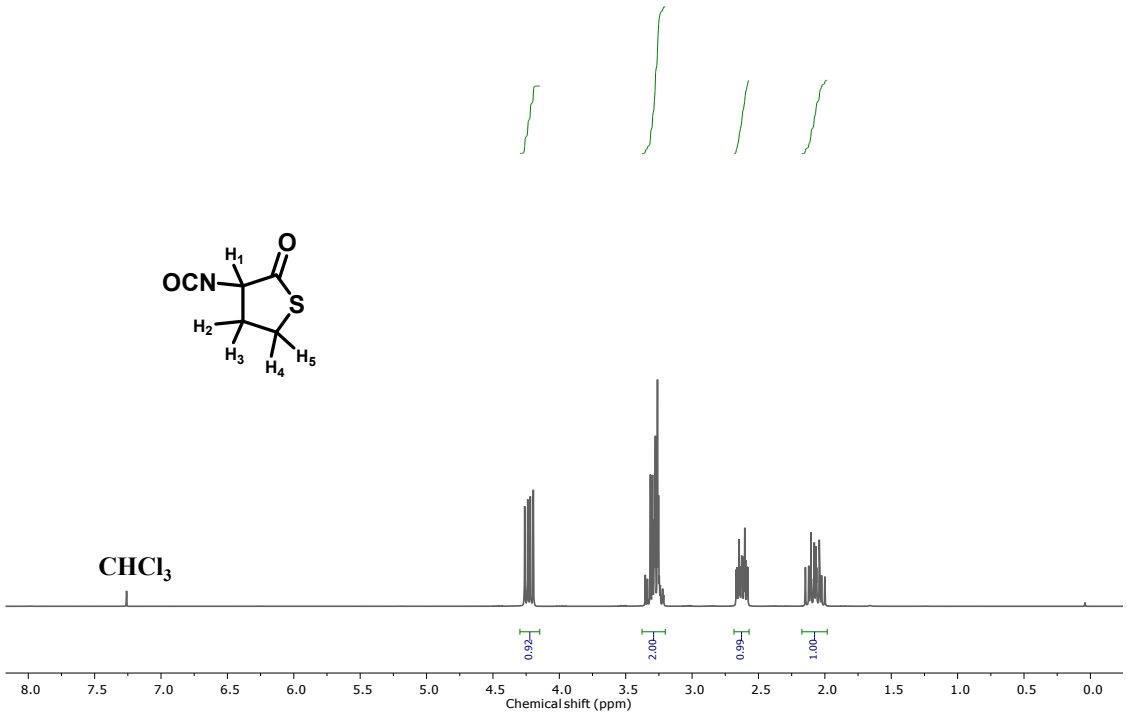
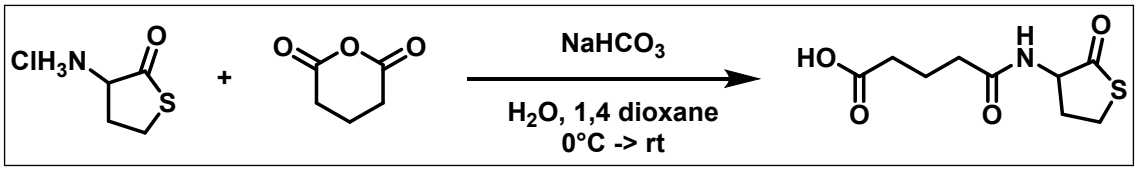


Figure S1: ^{1}H NMR analysis ($CDCl_3$) of *α*-isocyanato-*γ*-thiolactone

Synthesis of thiolactone-carboxylic acid (TLa-COOH)



DL-homocysteine thiolactone hydrochloride (30 g, 195 mmol, 1eq.) was dissolved in an ice-cooled solution of H_2O/1,4-dioxane (1:1 400 mL). $NaHCO_3$ (82.05 g, 819 mmol, 4.2 eq.) was added and the reaction was stirred for 30 minutes at 0 °C. Glutaric anhydride (44,55 g, 390 mmol, 2 eq.) was added in several portions to the reaction mixture and the reaction was allowed to reach room temperature and left to stir for 14 hours. The reaction mixture was acidified to pH 1 through the dropwise addition of a 12 M aqueous HCl solution. Next, the reaction mixture was extracted with ethyl acetate (3 x 300 mL). The combined organic phases were washed with brine (200 mL), dried over $MgSO_4$, filtered and dried *in vacuo*. The white solid was recrystallized from a minimal amount of acetone, yielding the thiolactone-carboxylic acid (TLa-COOH) as a white solid (32.92 g, 73.1%).

^{1}H NMR (300 MHz, DMSO-d_6, δ): 12.04 (s, 1H, H_1), 8.16 (d, J = 8.3 Hz, 1H, H_8), 4.59 (ddd, J = 12.6, 8.4, 7.0 Hz, 1H, H_9), 3.55-3.13 (m, 2H, $H_{12,13}$), 2.46-2.33 (m, 1H, H_{10}) 2.27-1.97 (m, 5H, $H_{2,3,6,7,11}$), 1.71 (p, J = 7.5 Hz, 2H, $H_{4,5}$); **^{13}C NMR** (75 MHz, DMSO-d_6, δ): 206.51 (C), 174.18 (C), 171.85, (C), 58.11 (CH), 34.29 (CH_2), 32.82 (CH_2), 30.13 (CH_2), 26.73 (CH_2), 20.57(CH_2).

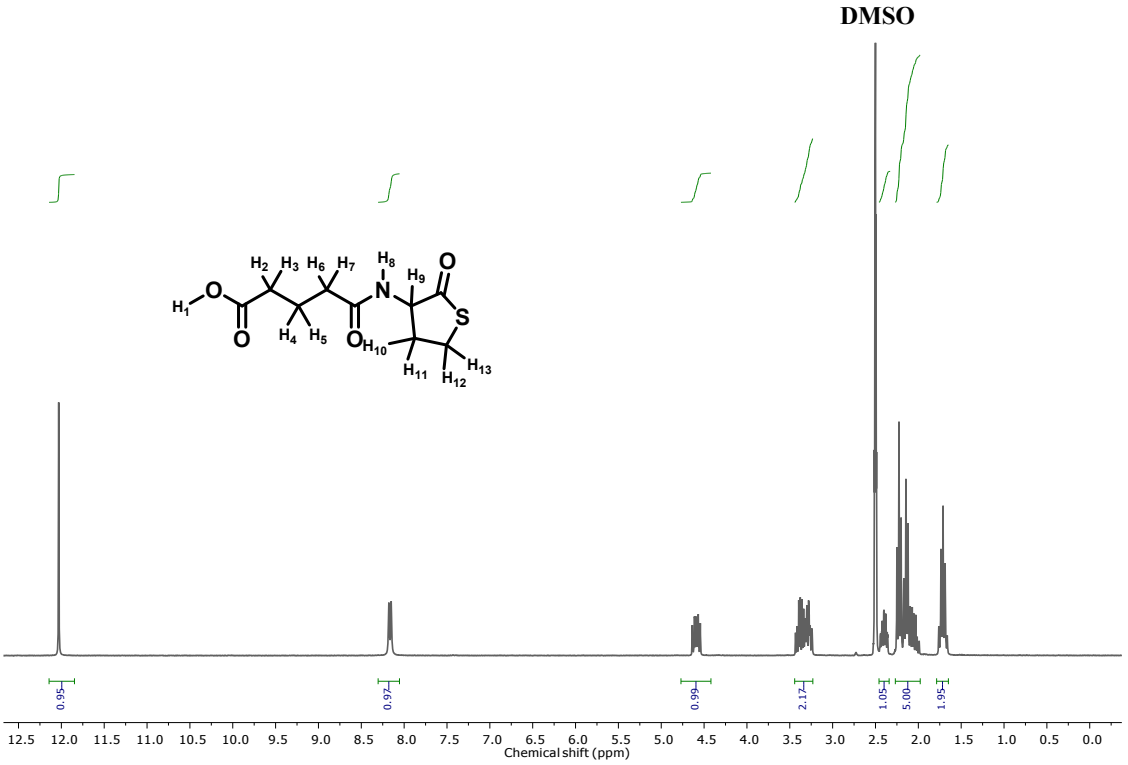
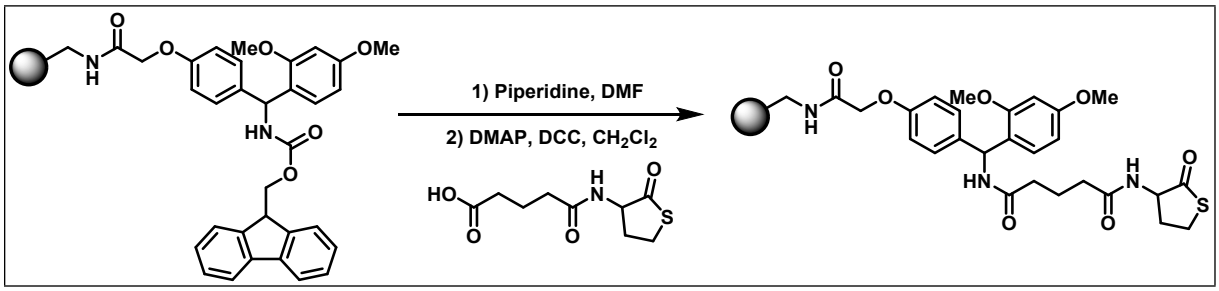


Figure S2: ^{1}H NMR analysis (DMSO-d_6) of TLa-COOH

General protocol for the synthesis of uniform macromolecules

Loading of Fmoc-protected rink-amide resin with TLa-COOH



The Fmoc-protected Rink-amide resin (5 g, 2 mmol, 1 eq.) with a commercial loading of 0.4 mmol g^{-1} was swollen in dry DMF (40 mL) and shaken for 10 min. Following removal of the solvent, the solid support was washed with DMF (3 x 40 mL). Next, the resin was treated with a premade 20 *v/v*% piperidine solution in dry DMF (40 mL) and the reactor was shaken vigorously at room temperature for 5 hours to ensure full fluorenylmethoxycarbonyl (Fmoc) removal. The solvent was filtered and the solid support was washed with DMF (2 x 40 mL) and CH_2Cl_2 (2 x 40 mL).

Next, dry CH_2Cl_2 (32 mL) was added to the solid support, which was placed under Argon atmosphere and shaken for 10 minutes. Subsequently, TLa-COOH (2.31 g, 10 mmol, 5 eq., see **Figure S2**), 4-dimethylaminopyridine (DMAP, 0.023 g, 0.2 mmol, 0.1 eq.) and *N*,*N*'-dicyclohexylcarbodiimide (DCC, 2.06 g, 10 mmol, 5 eq.) were added while preserving the inert conditions. At this time, dry DMF (8 mL) was added in order to facilitate the dissolution of the acid. The reaction mixture was shaken vigorously for 36 hours under argon atmosphere, before the resin was washed with DMF (2 x 40 mL), CH_2Cl_2 (2 x 40 mL) and Et_2O (2 x 40 mL). Finally, the resin was dried under reduced pressure at room temperature for 24 hours. A 100% loading efficiency was assumed as a result of the excess of the added reagents in combination with the extended reaction times.

Aminolysis and thiol-epoxy coupling

THF was added to the solid support (1 mL solvent per 100 mg of resin) and the beads were swollen for 10 minutes. Next, the desired amine (10 eq. relative to the number of thiolactone units on the solid support) and epoxide (25 eq. relative to the number of thiolactone units on the solid support) were added, together with 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU, 0.5 eq. relative to the number of thiolactone units on the solid support), dimethylphenylphosphine (DMPP, 1 eq. relative to the number of thiolactone units on the solid support) and a drop of water before the mixture was shaken for 1 hour. The reaction mixture was removed through filtration, and the resin was washed with DMF (4 x 2 mL), methanol (4 x 2 mL), $CHCl_3$ (4 x 2 mL) and Et_2O (4 x 2 mL). This procedure was performed twice in order to ensure near-quantitative conversions.

Chain extension

Dry $CHCl_3$ was added to the solid support (1 mL per 100 mg of resin) and the beads were swollen for 10 minutes. Next, α-isocyanato-γ-thiolactone (10 eq. relative to the number of alcohol moieties on the solid support) and zirconium(IV) acetylacetonate (0.025 eq. relative to the number of alcohol moieties on the solid support) were added to the reaction mixture, which was shaken for 1 hour. The reaction mixture was removed through filtration, and the resin was washed with DMF (4 x 2 mL), methanol (4 x 2 mL), $CHCl_3$ (4 x 2 mL) and Et_2O (4 x 2 mL).

Cleavage and LC-MS/SEC analysis

Cleavage of the product prior to LC-MS/SEC analysis was achieved through the treatment of the resin (2 mg for LC-MS, 5 mg for SEC) with a 10% (v/v) trifluoroacetic acid (TFA) solution in CH_2Cl_2 (1 mL) for 15 minutes. The immediate formation of a red color is related to the formation of on-resin cations, and their intensity allows a qualitative determination of the cleavage process. Next, the solid support was removed through filtration, and the sample was concentrated *in vacuo*. Finally, the obtained sample was dissolved in acetonitrile (1 mL) for LC-MS analysis or THF (1 mL) for SEC analysis.

LC-MS analysis

Optimization of the synthetic protocol

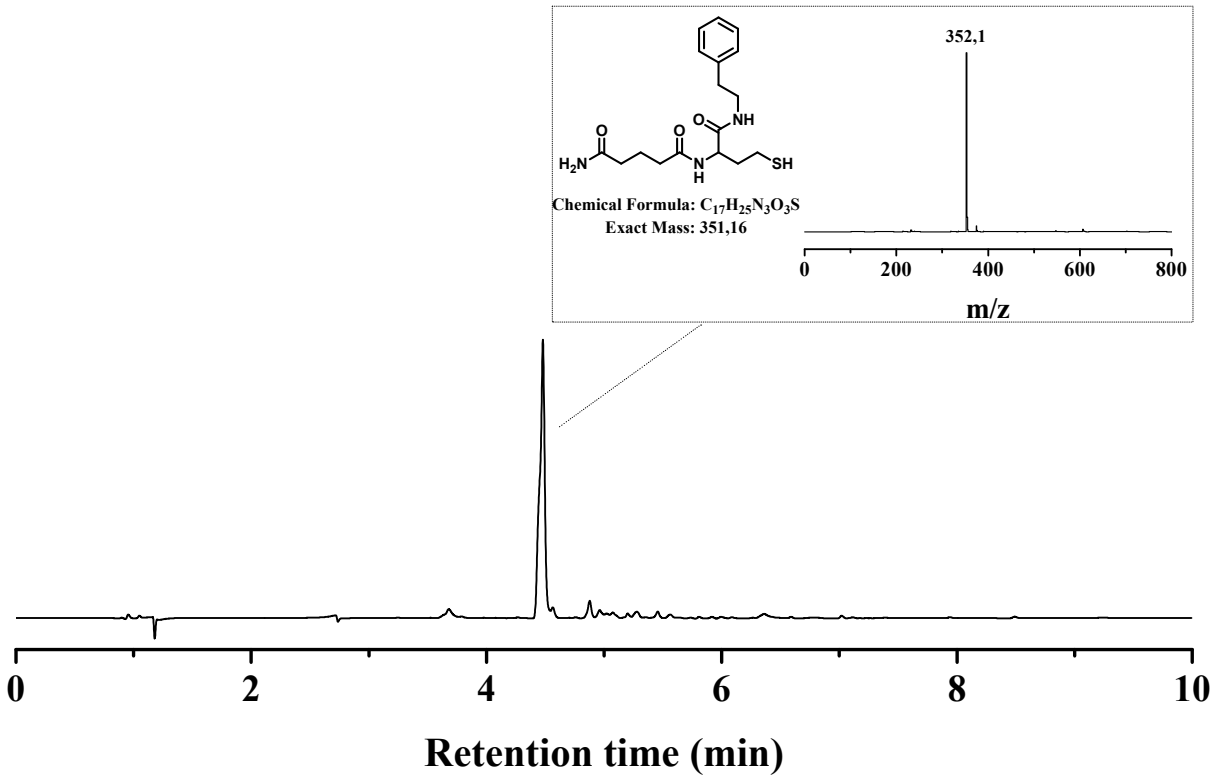


Figure S3: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, phenylethylamine (10 eq.) and 1,2-epoxybutane (25 eq.), catalyzed by triethylamine (0.5 eq.) in the presence of dimethylphenylphosphine (DMPP, 1 eq.) and a drop of water. The intermediate free thiol was the only product detected (mass detected as $[M+H]^+$, t = 4.46 min).

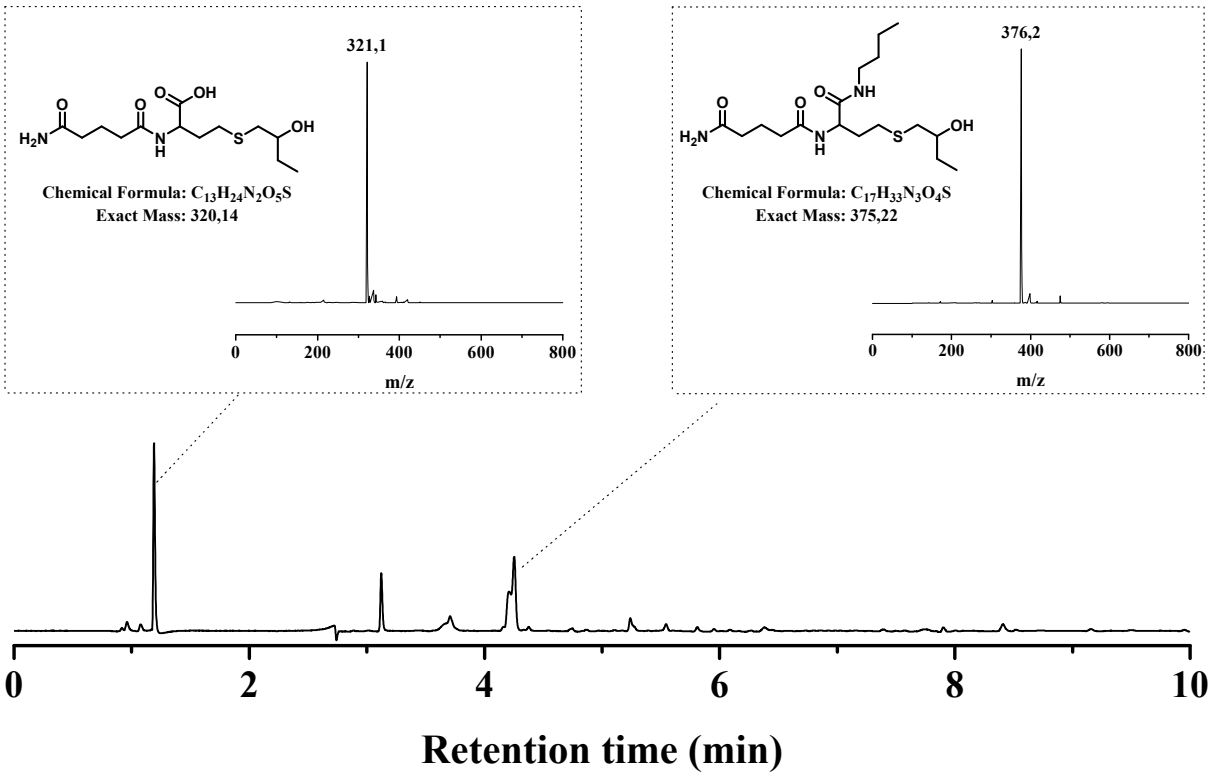


Figure S4: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, butylamine (10 eq.) and 1,2-epoxybutane (25 eq.), catalyzed by LiOH (1 eq.) in a THF/H_2O (9:1) mixture. The following products can be identified: hydrolysis of the thiolactone unit followed by the thiol-epoxy reaction (mass detected as $[M+H]^+$, t = 1.19 min) and the desired product (mass detected as $[M+H]^+$, t = 4.25 min).

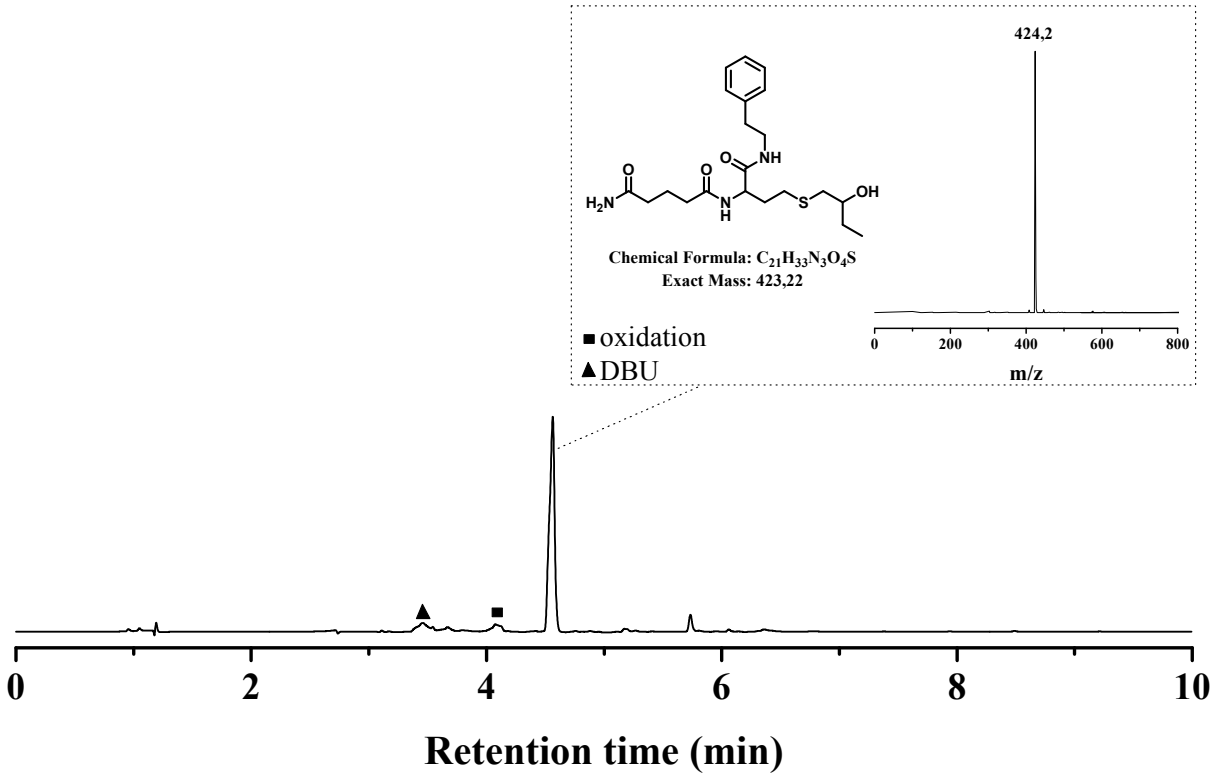


Figure S5: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, phenylethylamine (10 eq.) and 1,2-epoxybutane (25 eq.), catalyzed by DBU (0.5 eq.) in the presence of DMPP (1 eq.) and a drop of water. The following products can be identified: DBU (t = 3.47 min), oxidation of the product (t = 4.04 min), the product (mass detected as $[M+H]^+$, t = 4.54 min) and the TFA-ester of the product (t = 5.72 min). The latter is a result of the cleavage cocktail and is not observed with a thiolactone end-group, while DBU can be readily removed by performing extra washing steps after the reaction.

A selection of the different monomers tested within the protocol

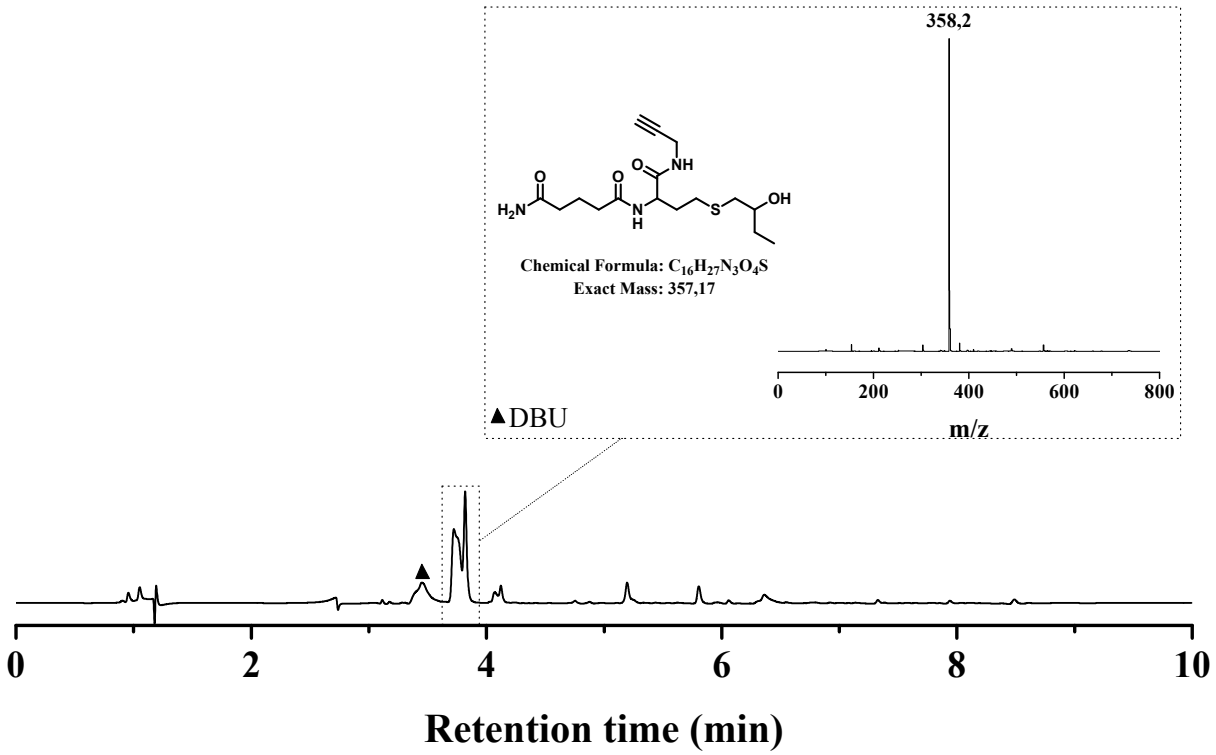


Figure S6: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, propargylamine (10 eq.) and 1,2-epoxybutane (25 eq.), catalyzed by DBU (0.5 eq.) in the presence of DMPP (1 eq.) and a drop of water. The following products can be identified: DBU (t = 3.47 min), the product (mass detected as $[M+H]^+$, t = 3.77 min) and the TFA-ester of the product (t = 5.19 min). The latter is a result of the cleavage cocktail and is not observed with a thiolactone end-group, while DBU can be readily removed by performing extra washing steps after the reaction.

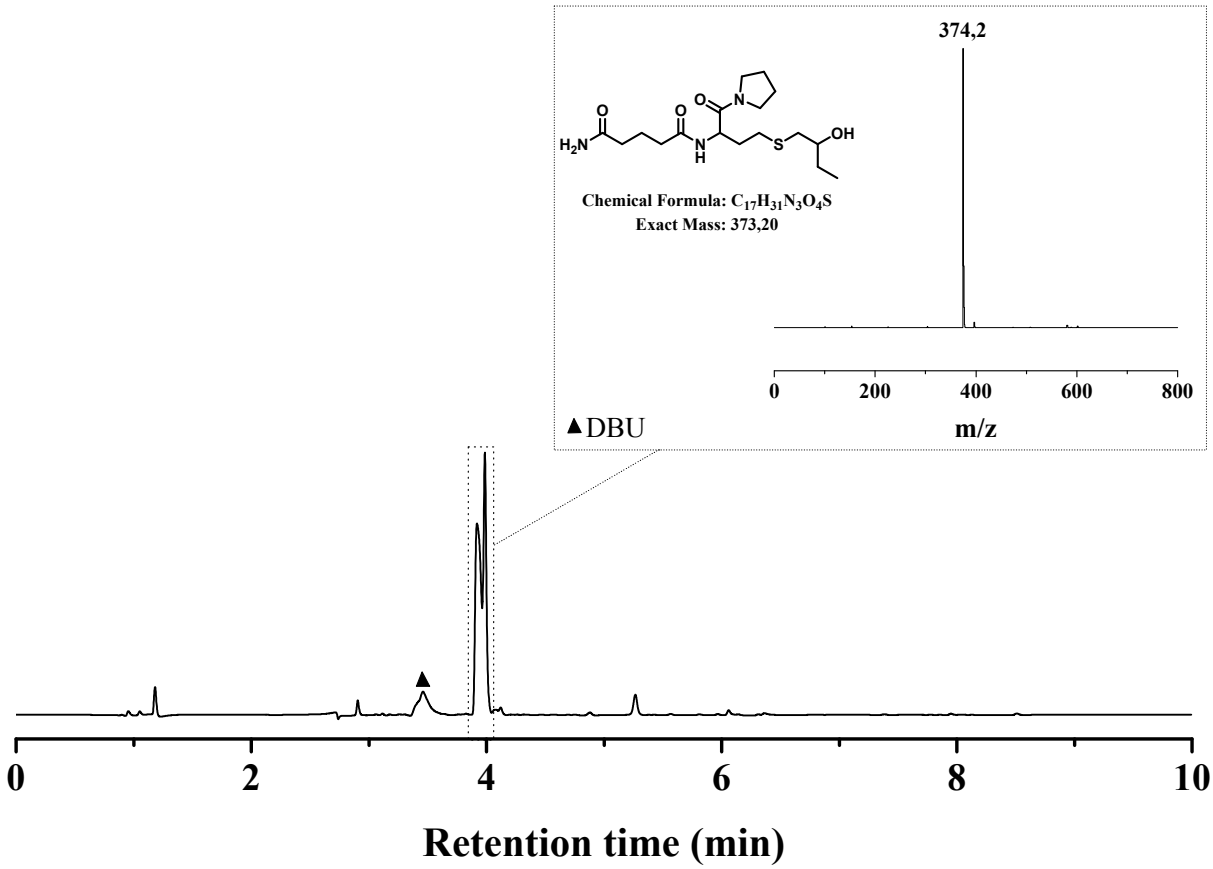


Figure S7: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, pyrrolidine (10 eq.) and 1,2-epoxybutane (25 eq.), catalyzed by DBU (0.5 eq.) in the presence of DMPP (1 eq.) and a drop of water. The following products can be identified: DBU (t = 3.47 min), the product (mass detected as $[M+H]^+$, t = 3.96 min) and the TFA-ester of the product (t = 5.27 min). The latter is a result of the cleavage cocktail and is not observed with a thiolactone end-group, while DBU can be readily removed by performing extra washing steps after the reaction.

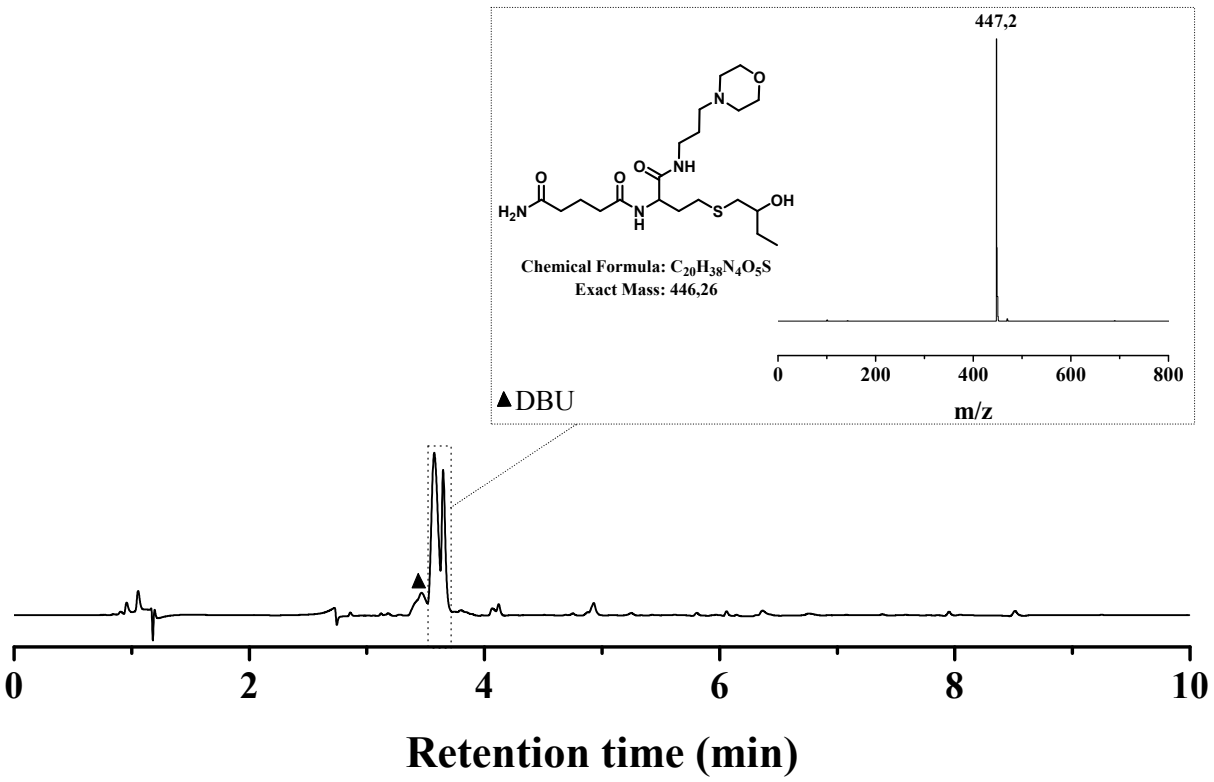


Figure S8: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, 3-morpholinopropylamine (10 eq.) and 1,2-epoxybutane (25 eq.), catalyzed by DBU (0.5 eq.) in the presence of DMPP (1 eq.) and a drop of water. The following products can be identified: DBU (t = 3.47 min), the product (mass detected as $[M+H]^+$, t = 3.61 min) and the TFA-ester of the product (t = 4.92 min). The latter is a result of the cleavage cocktail and is not observed with a thiolactone end-group, while DBU can be readily removed by performing extra washing steps after the reaction.

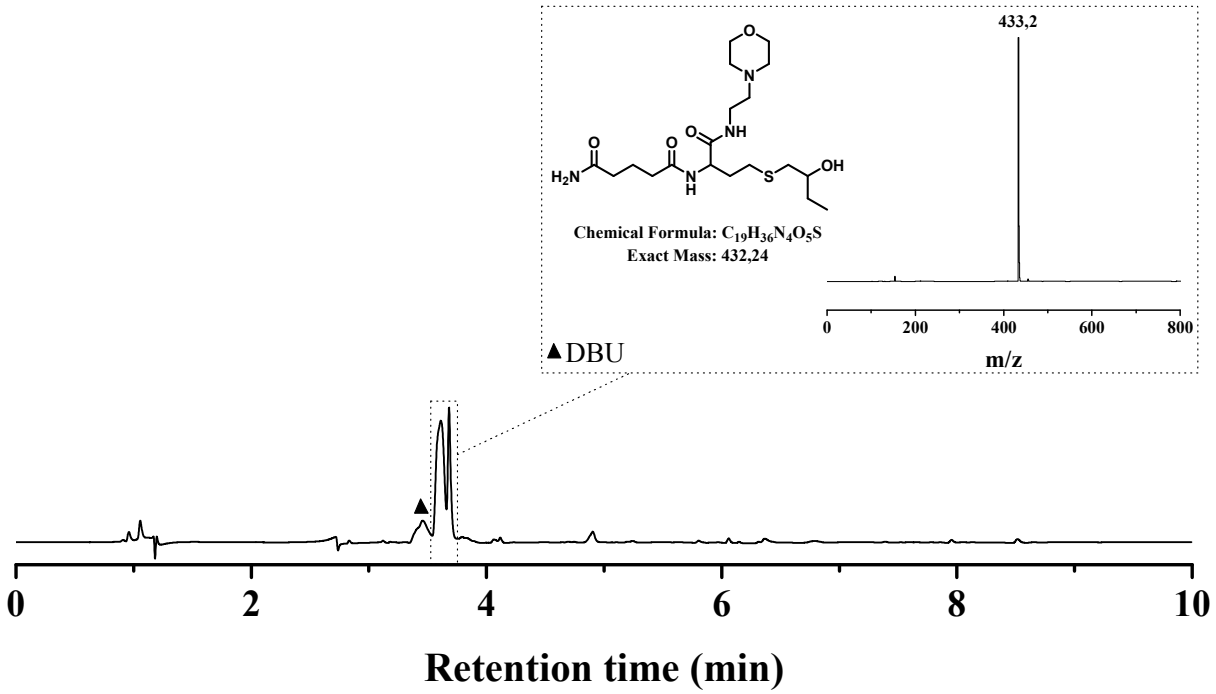


Figure S9: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, 2-morpholinoethylamine (10 eq.) and 1,2-epoxybutane (25 eq.), catalyzed by DBU (0.5 eq.) in the presence of DMPP (1 eq.) and a drop of water. The following products can be identified: DBU (t = 3.47 min), the product (mass detected as $[M+H]^+$, t = 3.65 min) and the TFA-ester of the product (t = 4.90 min). The latter is a result of the cleavage cocktail and is not observed with a thiolactone end-group, while DBU can be readily removed by performing extra washing steps after the reaction.

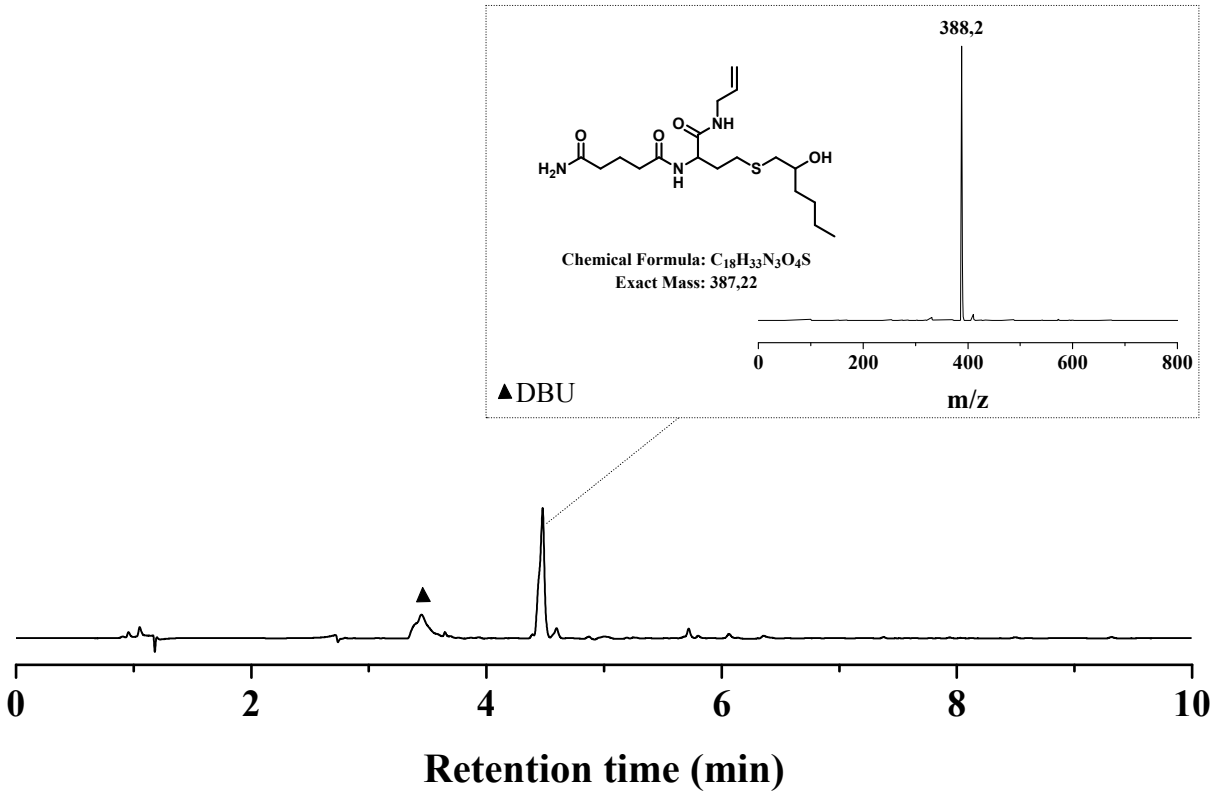


Figure S10: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, allylamine (10 eq.) and 1,2-epoxyhexane (25 eq.), catalyzed by DBU (0.5 eq.) in the presence of DMPP (1 eq.) and a drop of water. The following products can be identified: DBU (t = 3.47 min), the product (mass detected as $[M+H]^+$, t = 4.48 min) and the TFA-ester of the product (t = 5.71 min). The latter is a result of the cleavage cocktail and is not observed with a thiolactone end-group, while DBU can be readily removed by performing extra washing steps after the reaction.

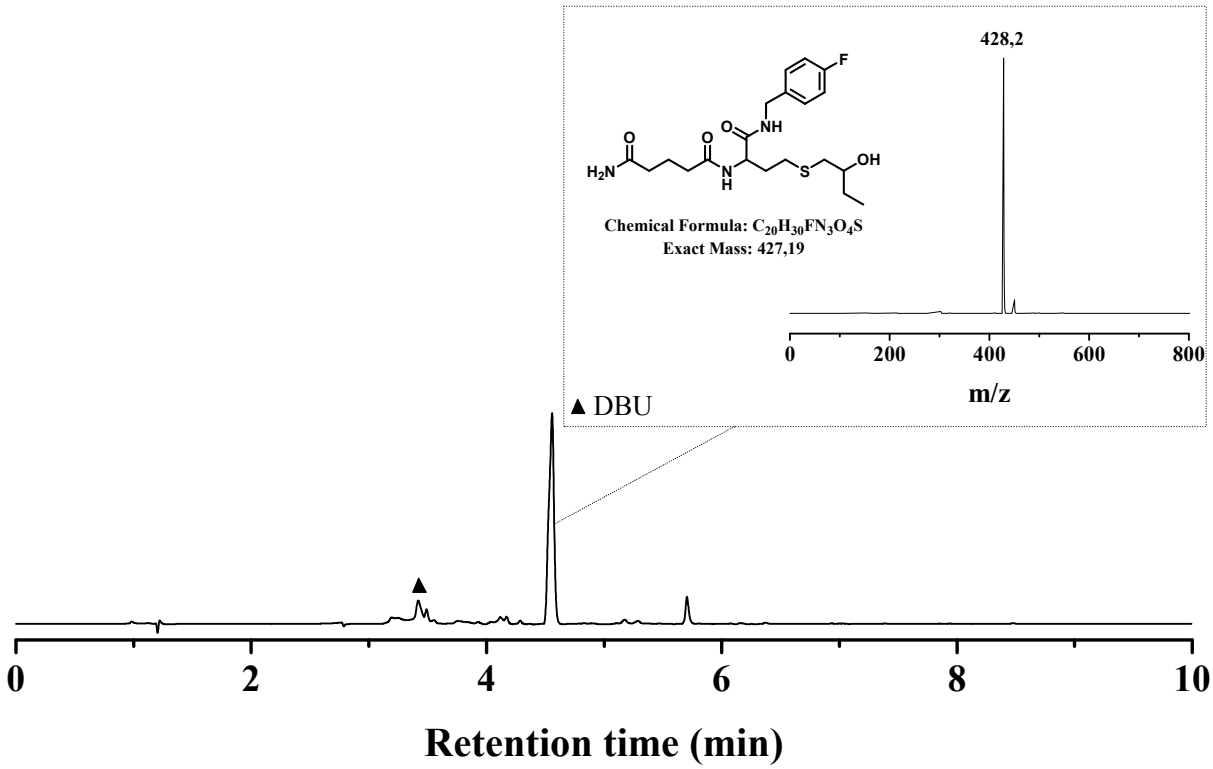


Figure S11: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, 4-fluorobenzylamine (10 eq.) and 1,2-epoxybutane (25 eq.), catalyzed by DBU (0.5 eq.) in the presence of DMPP (1 eq.) and a drop of water. The following products can be identified: DBU (t = 3.47 min), the product (mass detected as $[M+H]^+$, t = 4.56 min) and the TFA-ester of the product (t = 5.83 min). The latter is a result of the cleavage cocktail and is not observed with a thiolactone end-group, while DBU can be readily removed by performing extra washing steps after the reaction.

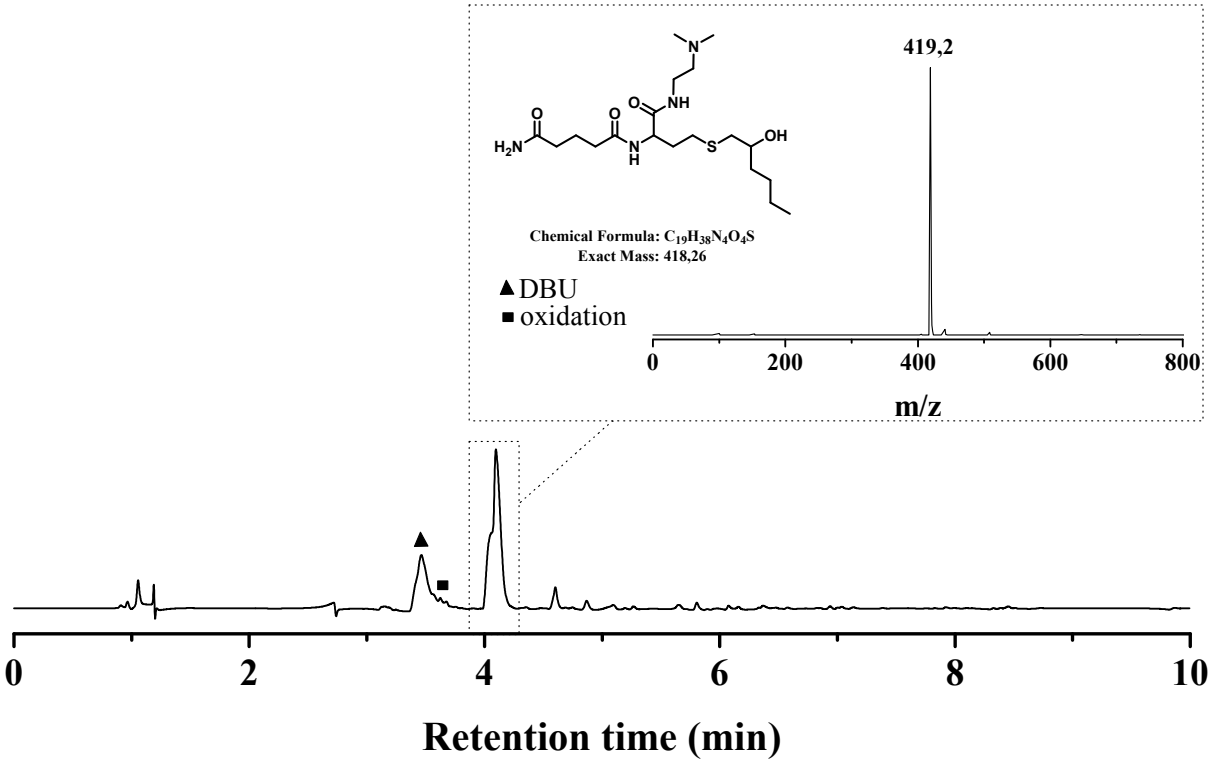


Figure S12: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, *N*,*N*-Dimethylethylenediamine (10 eq.) and 1,2-epoxyhexane (25 eq.), catalyzed by DBU (0.5 eq.) in the presence of DMPP (1 eq.) and a drop of water. The following products can be identified: DBU (t = 3.47 min), the product (mass detected as $[M+H]^+$, t = 4.10 min) and the TFA-ester of the product (t = 4.58 min). The latter is a result of the cleavage cocktail and is not observed with a thiolactone end-group, while DBU can be readily removed by performing extra washing steps after the reaction.

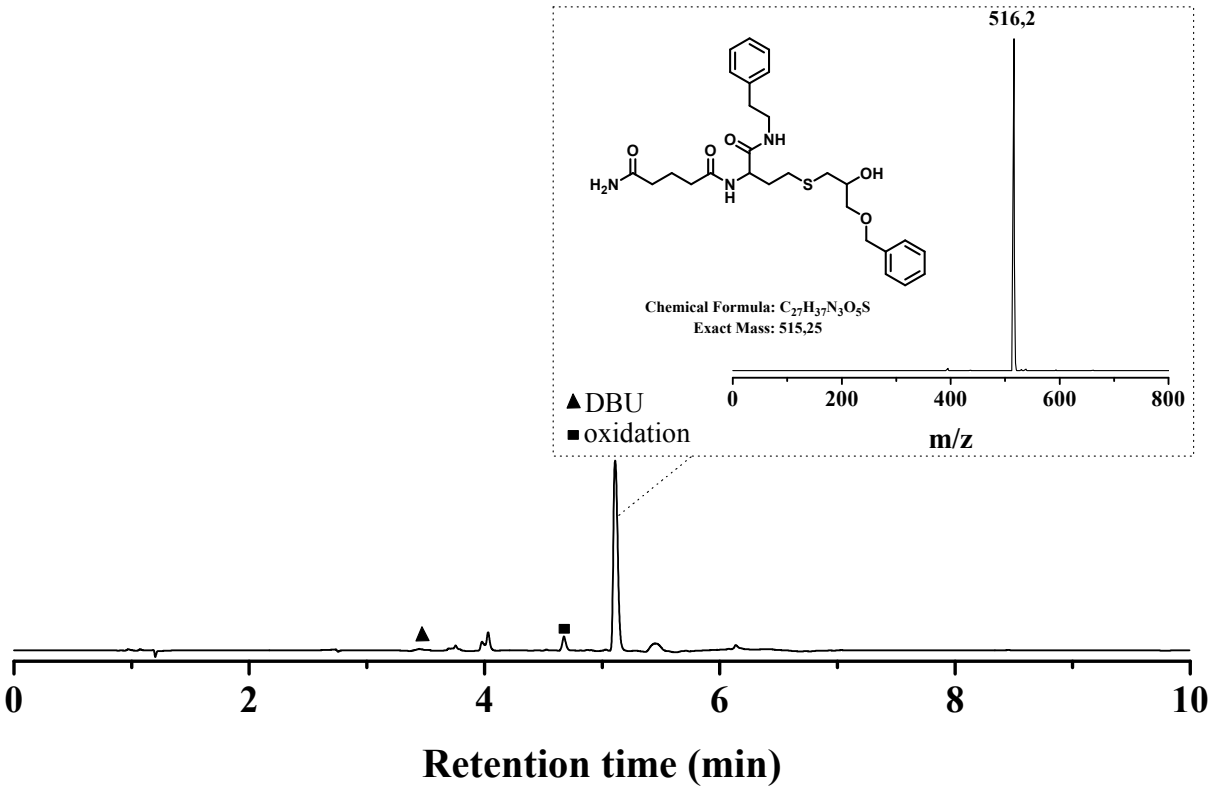


Figure S13: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound thiolactone, phenethylamine (10 eq.) and benzyl glycidyl ether (25 eq.), catalyzed by DBU (0.5 eq.) in the presence of DMPP (1 eq.) and a drop of water. The following products can be identified: DBU (t = 3.47 min), oxidation of the product (t = 4.73 min), the product (mass detected as $[M+H]^+$, t = 5.10 min) and the TFA-ester of the product (t = 6.15 min). The latter is a result of the cleavage cocktail and is not observed with a thiolactone end-group, while DBU can be readily removed by performing extra washing steps after the reaction.

Post-modification using CuAAC on resin

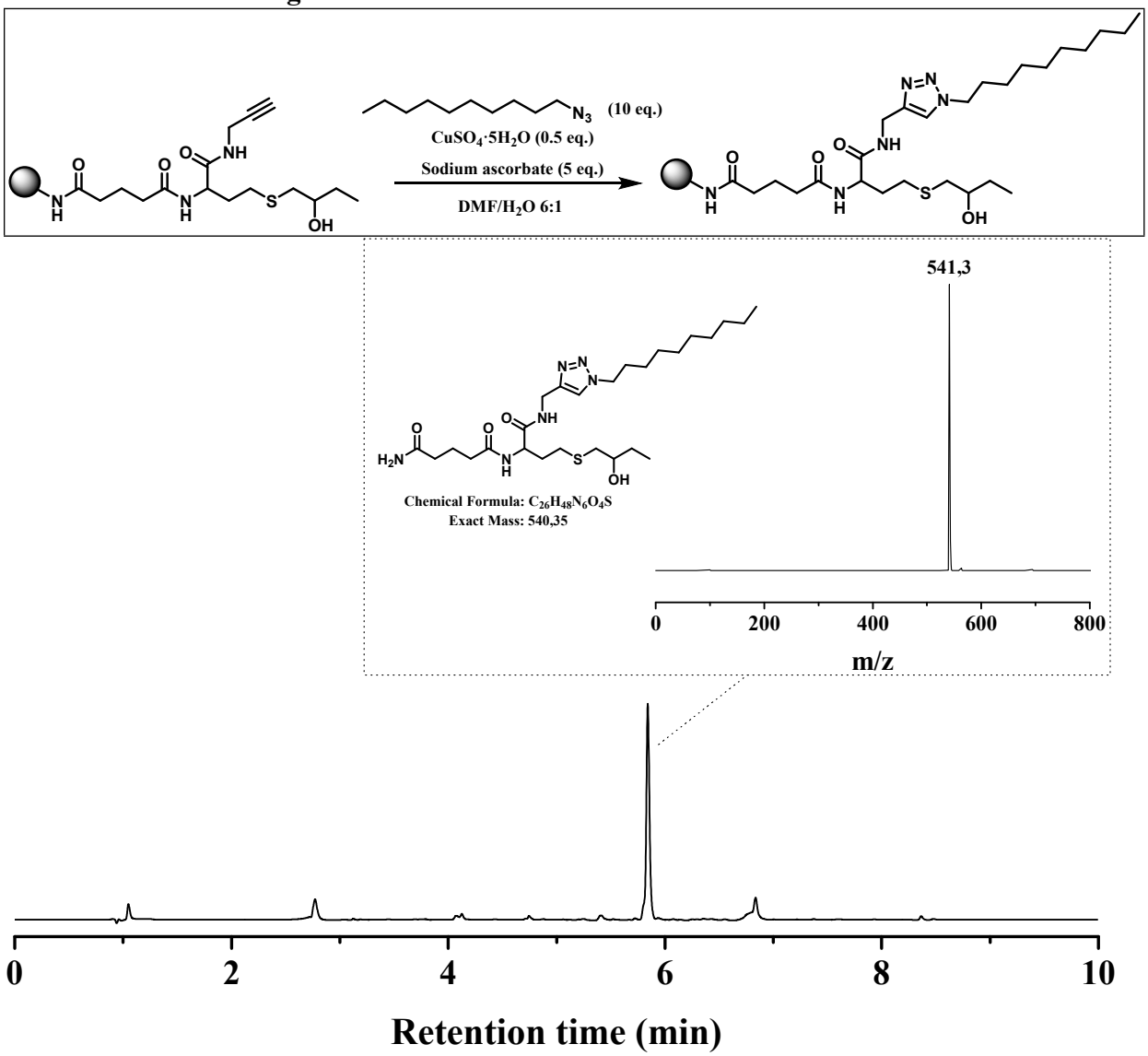


Figure S14: LC-MS trace (λ = 214 nm) of the reaction with a solid phase bound alkyne moiety, 1-azidodecane (10 eq.), $CuSO_4 \cdot 5H_2O$ (0.5 eq.) and *L*-sodium ascorbate (5 eq.) in a DMF/H_2O 6:1 mixture. The following products can be identified: oxidation of the product (t = 5.41 min), the product (mass detected as $[M+H]^+$, t = 5.85 min), the TFA-ester of the product (t = 6.83 min). The latter is a result of the cleavage cocktail and is not observed with a thiolactone end-group.

Characterization of M2

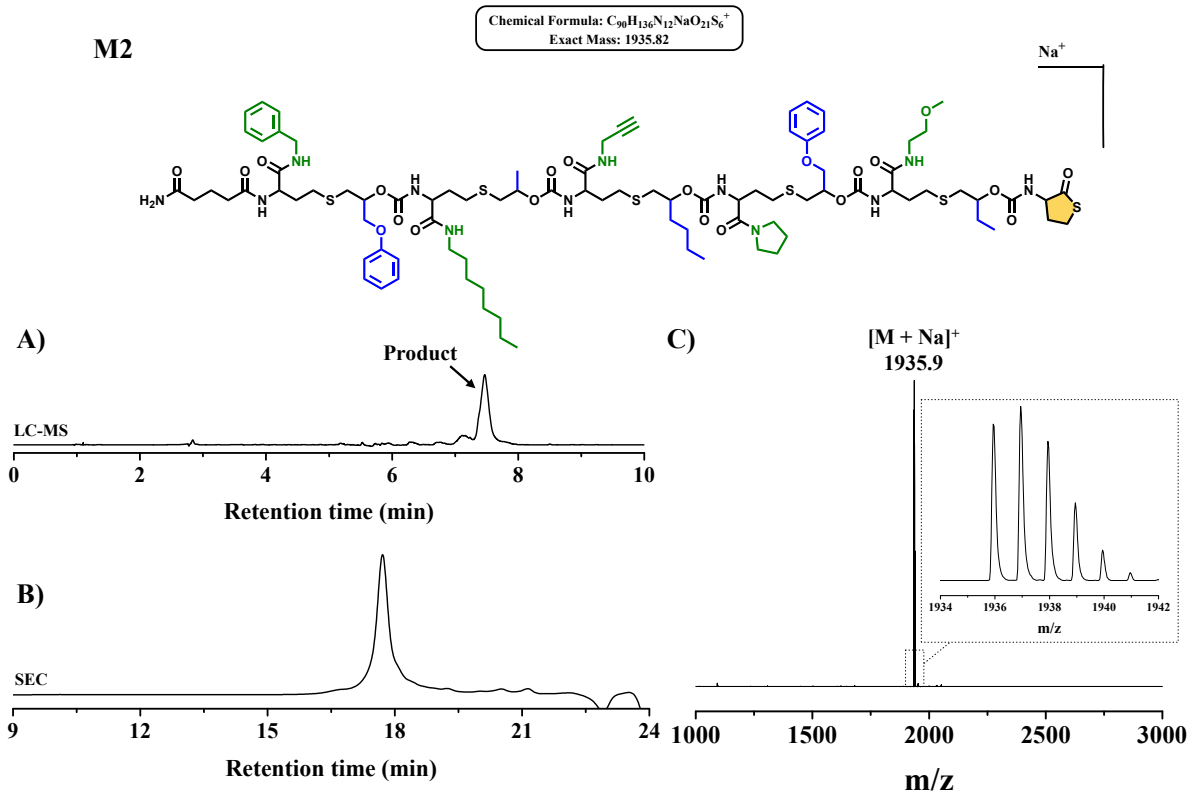


Figure S15: Structure of the synthesized sequence-defined macromolecule **M2** based on the amine-thiolactone-epoxy conjugation. The final product was characterized via a) LC-MS analysis ($\lambda = 214$ nm), showing the desired product, b) SEC chromatogram in THF and c) MALDI-TOF, displaying the sodium adduct of the desired macromolecule **M2**.

Characterization of M3

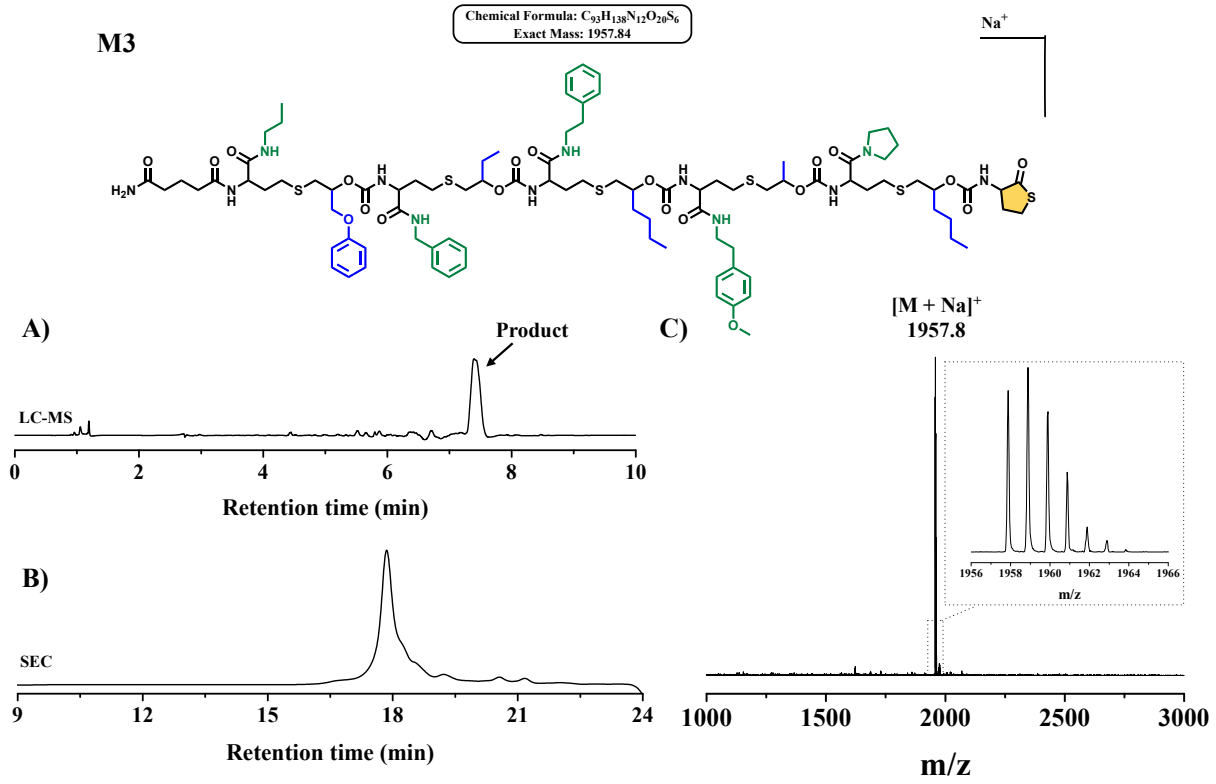


Figure S16: Structure of the synthesized sequence-defined macromolecule **M3** based on the amine-thiolactone-epoxy conjugation. The final product was characterized via a) LC-MS analysis ($\lambda = 214$ nm), showing the desired product b) SEC chromatogram in THF and c) MALDI-TOF, displaying the sodium adduct of the desired macromolecule **M3**.

MALDI-TOF/TOF analysis

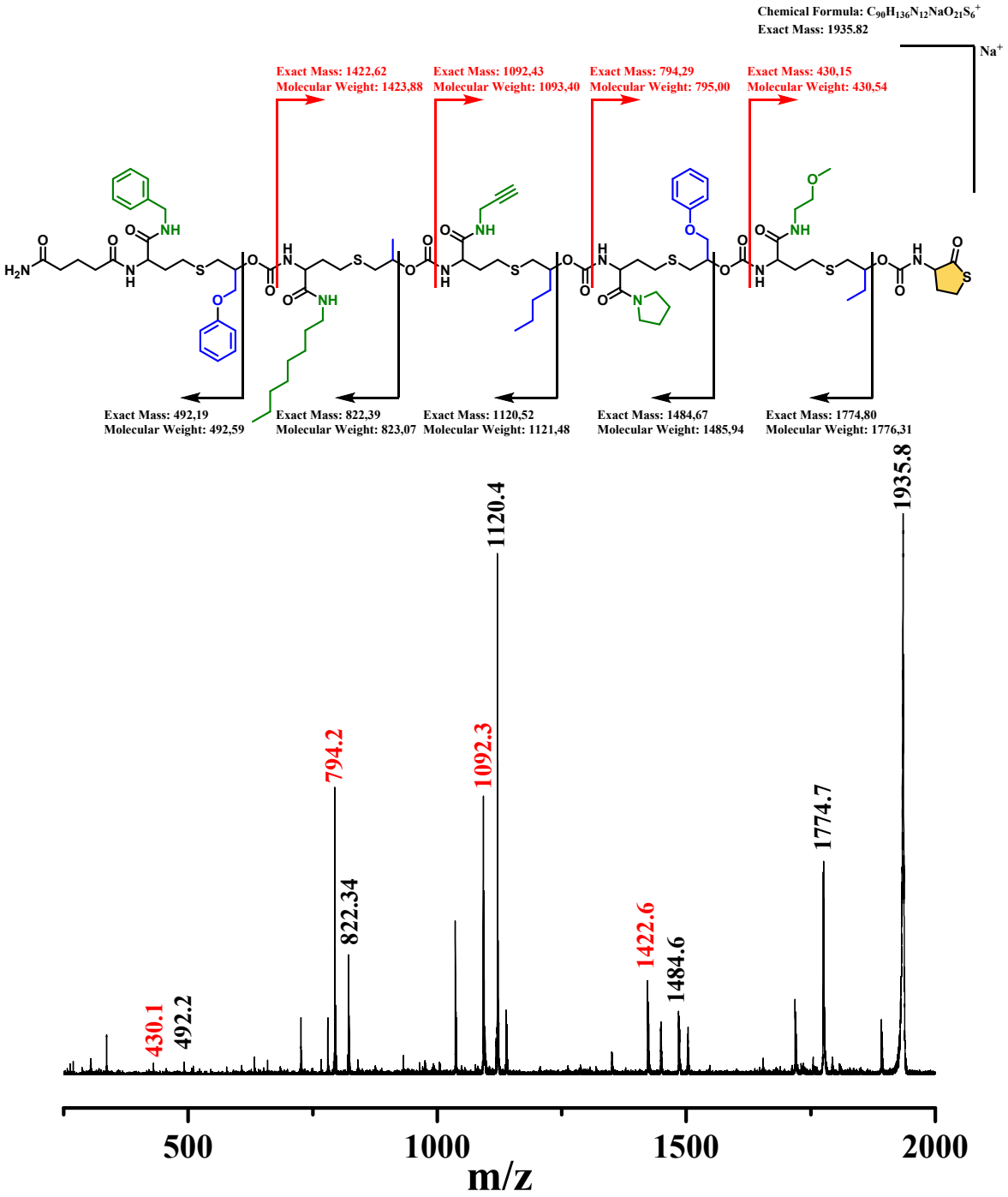


Figure S17: MALDI-TOF/TOF sequencing of the uniform macromolecule **M2**.

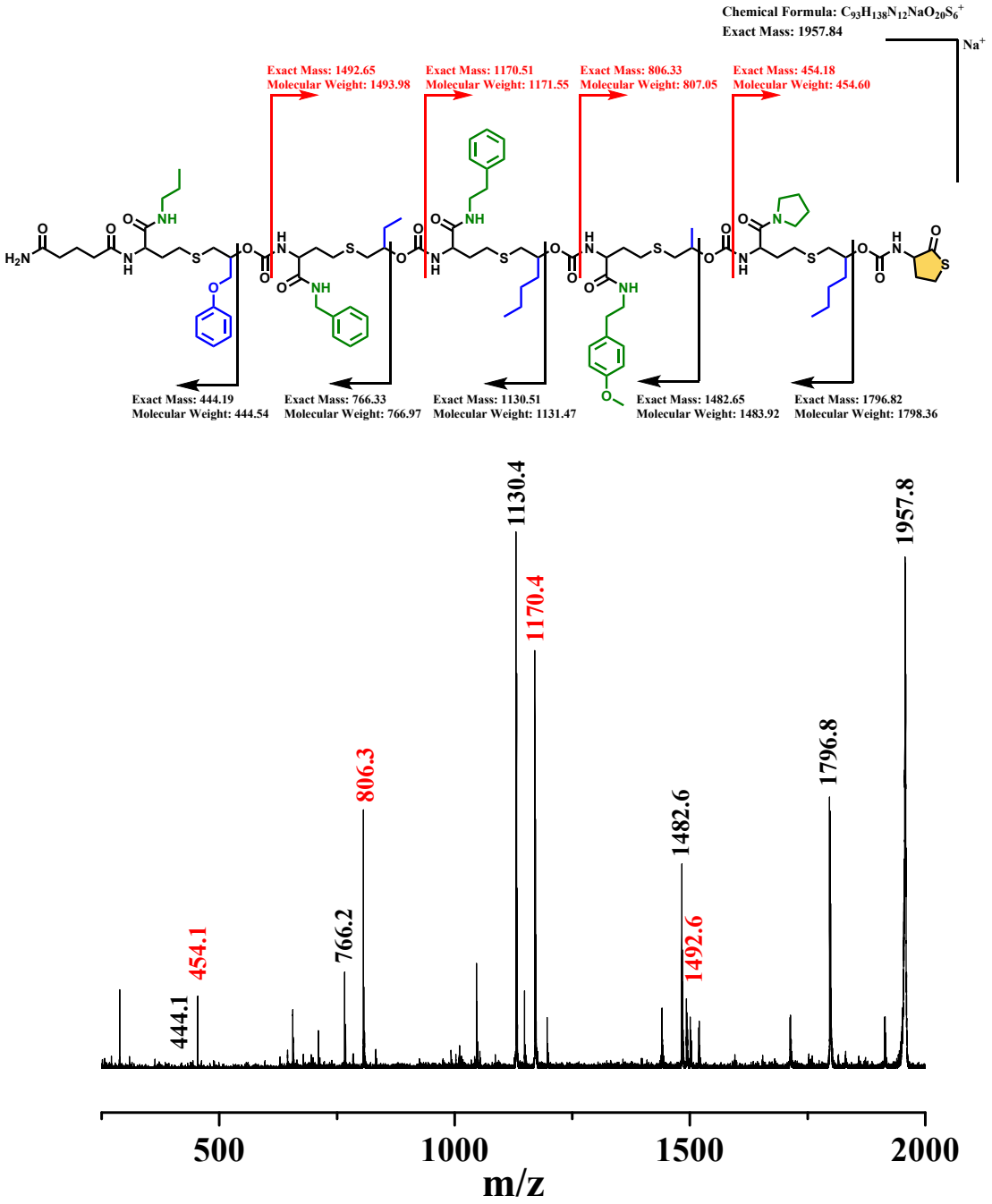


Figure S18: MALDI-TOF/TOF sequencing of the uniform macromolecule **M3**.

NMR analysis of M1-M3

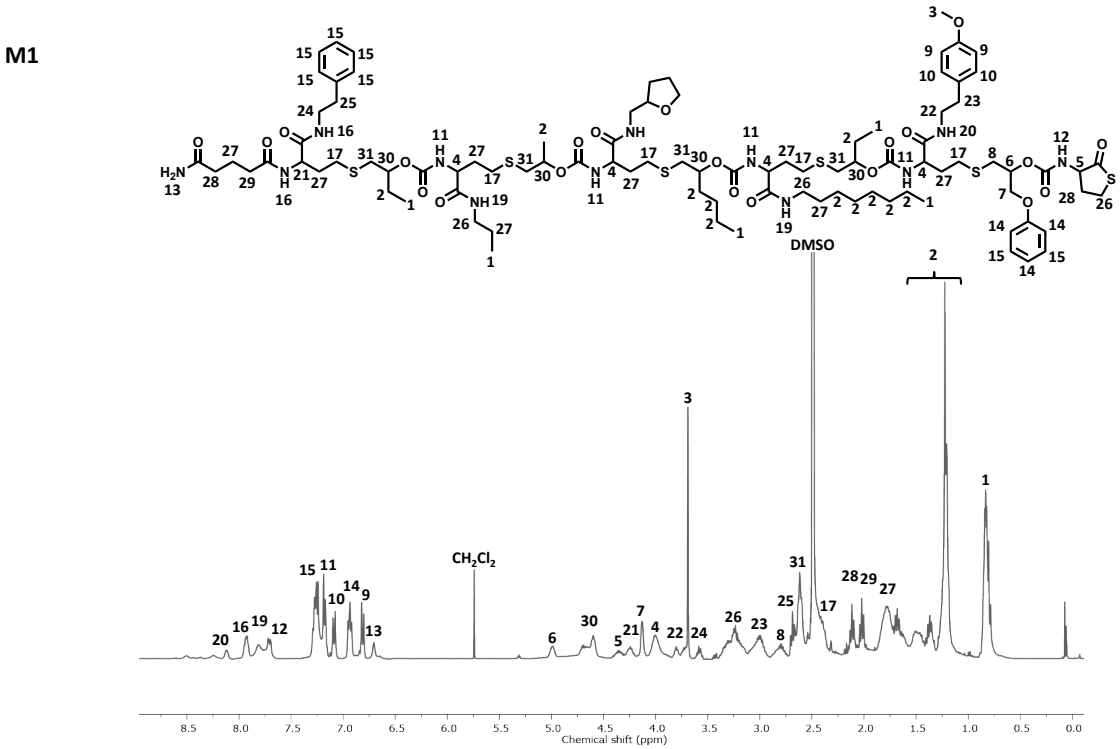


Figure S19: ^{1}H NMR analysis (DMSO-d_6) of uniform macromolecule **M1.**

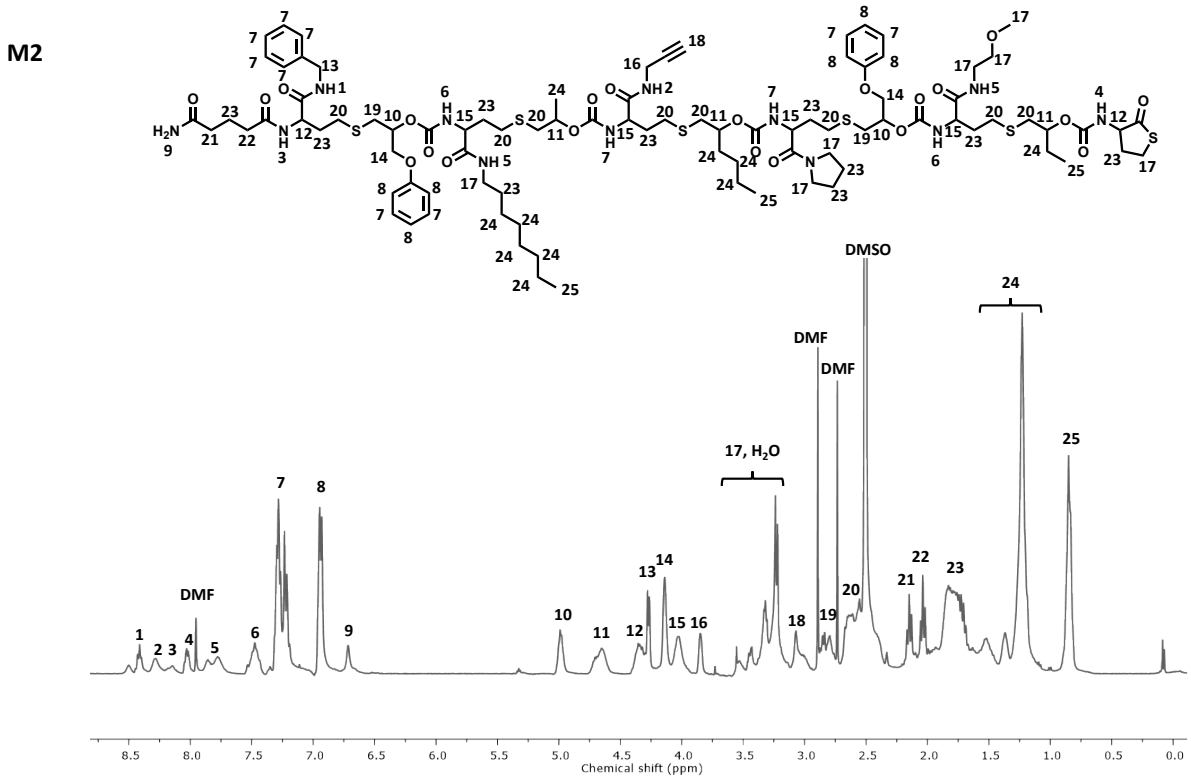


Figure S20: ^{1}H NMR analysis (DMSO-d_6) of uniform macromolecule **M2.**

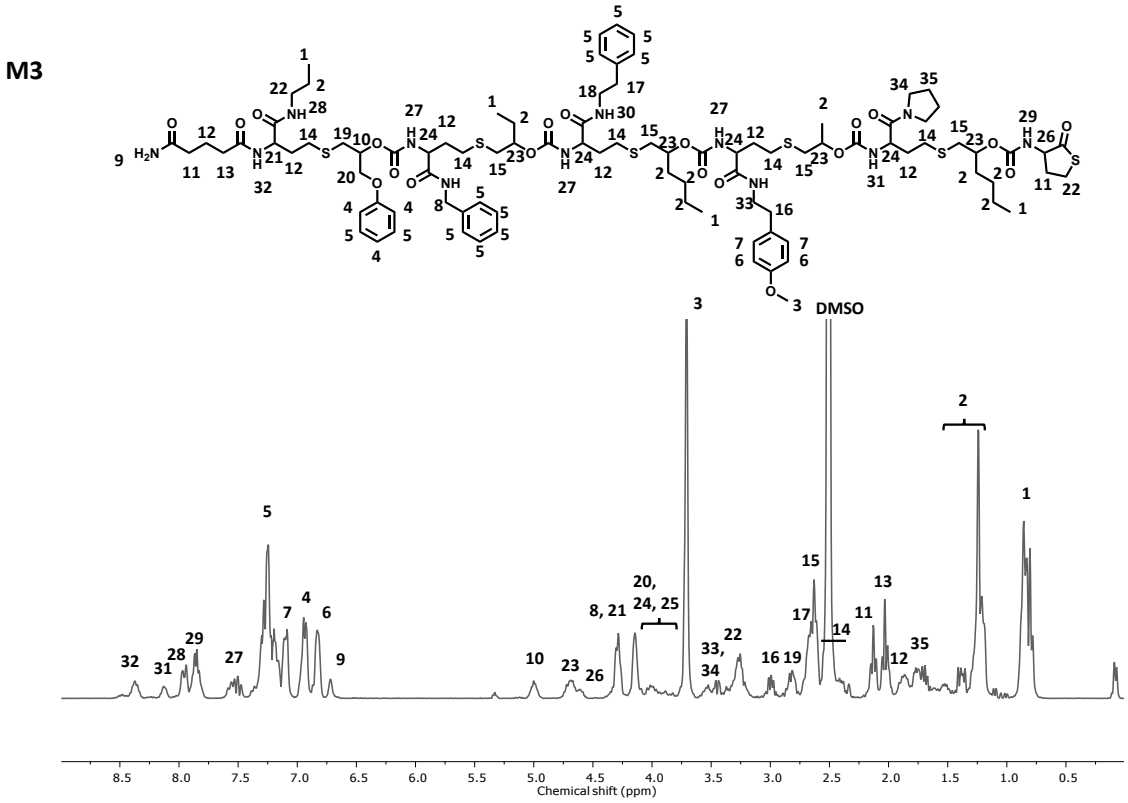


Figure 21: ^{1}H NMR analysis (DMSO-d_6) of uniform macromolecule **M3.**

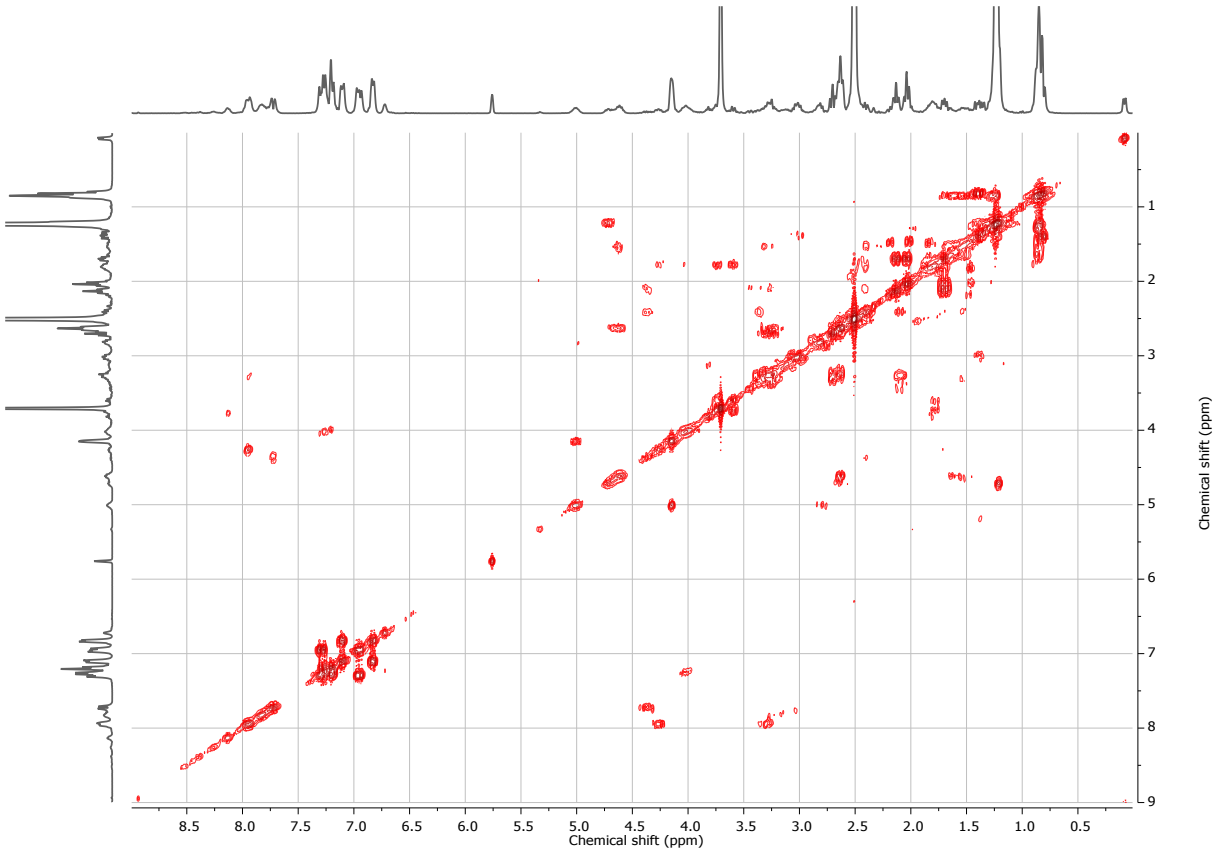


Figure S22: COSY NMR analysis (DMSO-d_6) of uniform macromolecule **M1.**

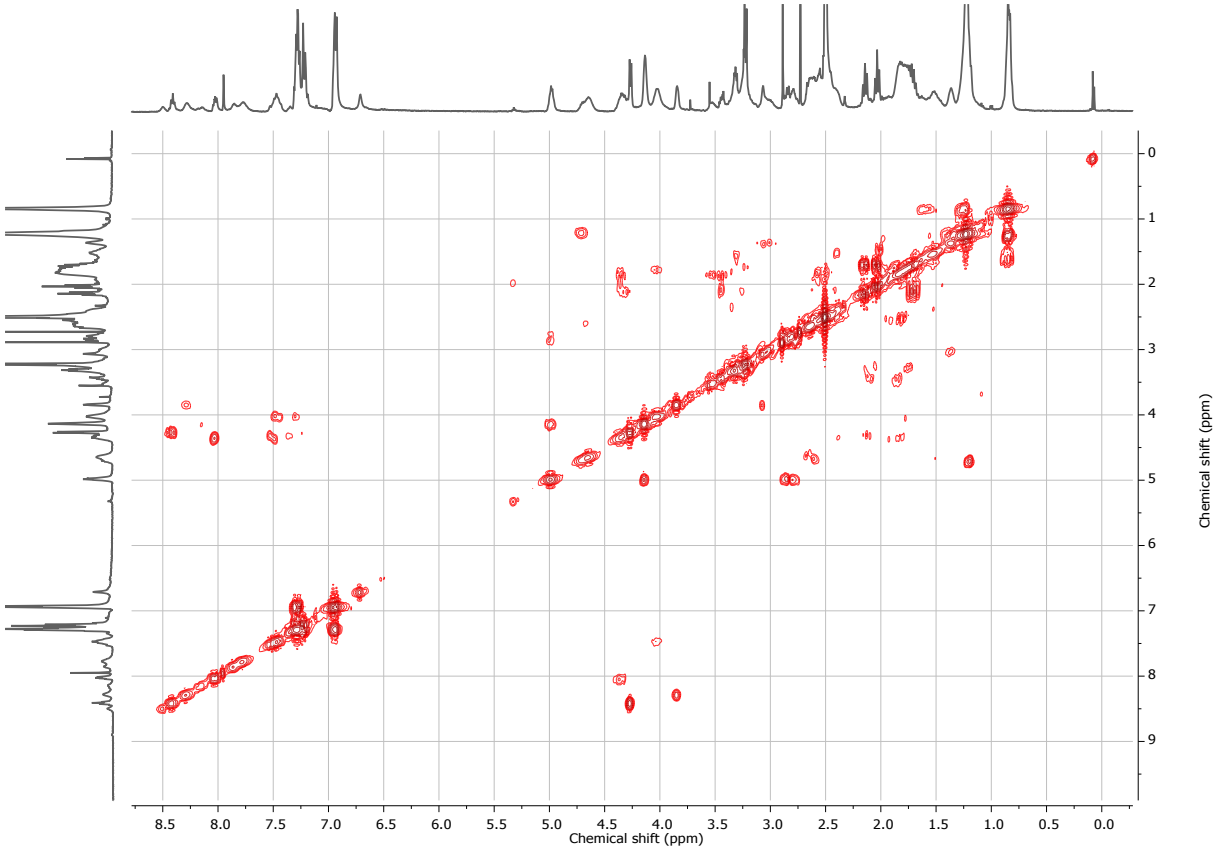


Figure S23: COSY NMR analysis (DMSO-d_6) of uniform macromolecule **M2.**

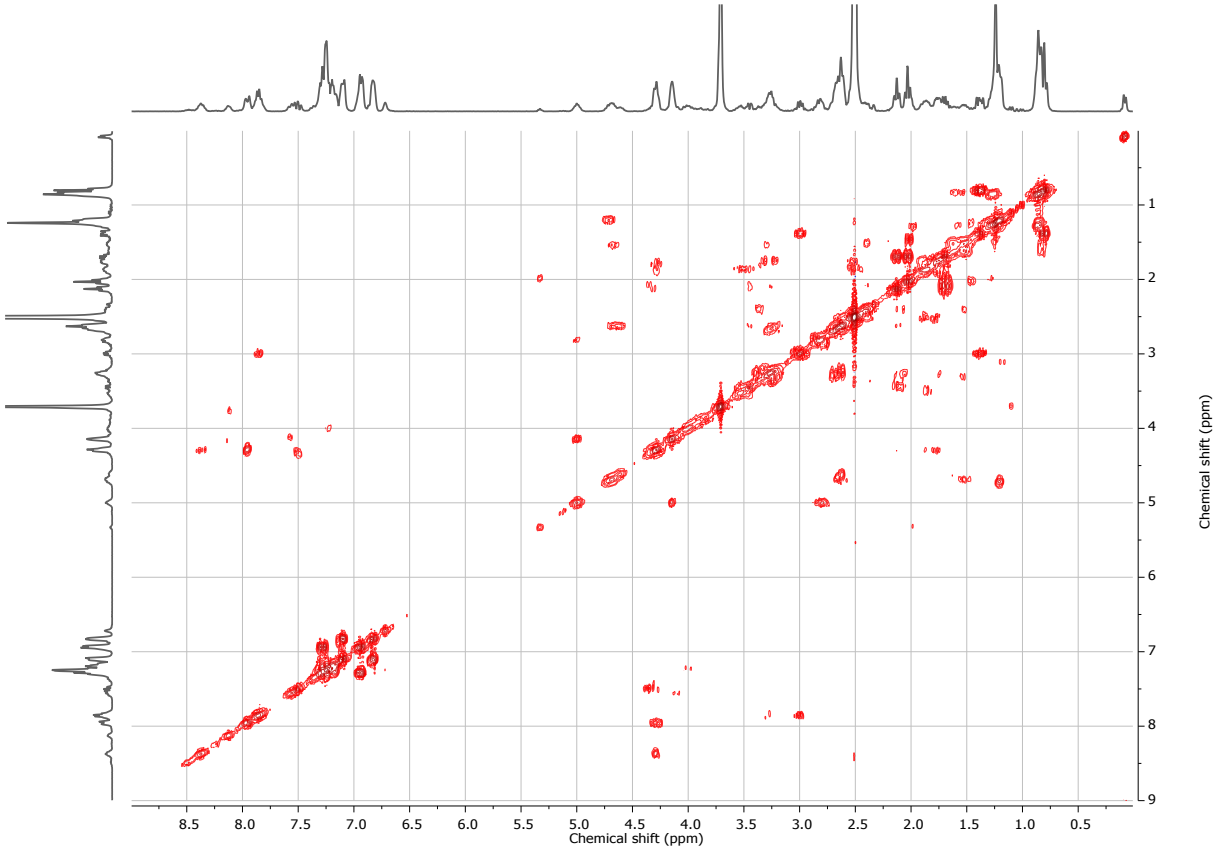


Figure S24: COSY NMR analysis (DMSO-d_6) of uniform macromolecule **M3.**