Electronic Supplementary Material (ESI) for ChemComm.
This journal is © The Royal Society of Chemistry 2021

Supporting Information

PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction

Peter Pasch[a†], Alexander Höing[b†], Serap Ueclue[a], Matthias Killa[c], Jens Voskuhl[c], Shirley K. Knauer[b*] and Laura Hartmann[a*]

a. Department for Organic Chemistry and Macromolecular Chemistry, Heinrich Heine University Düsseldorf, Universitätsstraße 1, Düsseldorf 40225, Germany.
b. Department for Molecular Biology II, Center of Medical Biotechnology (ZMB), University Duisburg-Essen, Universitätsstrasse 5, Essen, 45117, Germany.
c. Faculty of Chemistry (Organic Chemistry), University of Duisburg-Essen, Universitätsstraße 7, 45117 Essen, Germany.

†These authors contributed equally.
* Corresponding authors.

Materials:

Diethyl ether (with BHT as inhibitor, ≥ 99.8%), triisopropylsilane (TIPS) (98%), concentrated hydrochloric acid (pa), acetic anhydride (pa) and formic acid (pa) were purchased from Sigma Aldrich. N,N-Diisopropylethylamine (DIPEA) (≥ 99%) was purchased from Carl Roth. N,N-Dimethylformamide (DMF) (99.8%, for peptide synthesis), piperidine (99%) were obtained from Acros Organics. Dichloromethane (DCM) (99.99%), ethyl acetate (analytical reagent grade) and 1,4-dioxane (analytic reagent grade) were purchased from Fisher Scientific. Acetonitrile was purchased from AppliChem. Trifluoroacetic acid (TFA) (99%), (benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), and triethylsilane (analytical reagent grade) were purchased from Fluorochem. TentaGel® S RAM (Rink Amide) and TentaGel® PAP resins (loading: 0.23 mmol / g) were purchased from RAPP Polymer. N_α-Fmoc-N_ε-Boc-L-lysine (≥98.0%) was purchased from Iris Biotech. N_α-Fmoc-N_β-Boc-L-2,3-diaminopropionic acid (≥98.0%) was purchased from TCI. Polyethylene glycol 3000 was purchased from Merck.

Analytical Methods:

Preparative Reversed Phase- High Pressure Liquid Chromatography (prep RP-HPLC)

An Agilent 1260 Infinity device was used to purify the oligo(amidoamines), which is coupled to a variable wavelength detector (VWD) (set to 214 nm) and an automated fraction collector. The RP HPLC column, CAPCELL PAK C18 (20 x 250 mm, 5 µm), was used. The mobile phases A and B were H_2O and acetonitrile, each containing 0.1 vol% formic acid. The flow rate was set at 15 ml/min.

Reversed Phase- High Pressure Liquid Chromatography- Mass Spectrometry (RP-HPLC-MS)/Electron Spray Ionization- Mass Spectrometry (ESI-MS)

RP-HPLC-MS was carried out on an Agilent 1260 Infinity instrument coupled to a variable wavelength detector (VWD) (set to 214 nm) and a 6120 Quadrupole LC/MS containing an Electrospray Ionization (ESI) source (operation mode positive, m/z range from 200 to 2000). A MZ-AquaPerfect C18 (3.0 × 50 mm, 3 µm) RP column from Mz-Analysentechnik was used. As eluent system water/acetonitrile containing 0.1 vol% formic acid was applied. The mobile phases A and B were: System A) H_2O/acetonitrile (95/5, *v*/*v*); System B) H_2O / acetonitrile (5/95, v/v). The samples were analyzed at a flow rate of 0.4 ml/min using a linear gradient,

starting with 100% of system A) and reaching 100% system B) within 30 min. The temperature of the column room was set to 40 °C. All purities were determined using the OpenLab ChemStation software for LC/MS from Agilent Technologies.

Electron Spray Ionization- Mass Spectrometry (ESI-MS) measurements were performed with the above mentioned ESI source and quadrupole detector.

Ultra High Resolution - Mass Spectrometry (UHR-MS)

UHR-MS measurements were performed with a Bruker UHR-QTOF maXis 4G instrument with a direct inlet via syringe pump, an ESI source and a quadrupole followed by a Time of Flight (QTOF) mass analyzer.

Matrix-Assisted Laser Desorption Ionization-Time of Flight–Mass Spectrometry

(MALDI-TOF-MS) Compounds were detected using a Bruker MALDI-TOF Ultraflex I system with 2,5-dihydroxybenzoic acid (DHB) as matrix. The matrix to compound ratio of was 10:1. Spectra were acquired for reflector mode for a m/z range 2000-20000. The reflector mode was calibrated using a protein mixture.

Nuclear Magnetic Resonance Spectroscopy (NMR)

The ^{1}H-NMR spectra were recorded on a Bruker Avanace III 600 (600 MHz). These spectra were evaluated according to the following scheme: (frequency in MHz, deuterated solvent), chemical shift in ppm (multiplicity, coupling constant, integral, signal assignment). The chemical shift is given in relation to the ^{1}H signals of the deuterated solvents used (D_2O: 4.79 ppm). The multiplicities of the signals were abbreviated as follows: s (singlet), d (doublet), t (triplet), m (multiplet).

Freeze dryer

The final oligomers were lyophilized with an Alpha 1-4 LD plus instrument from Martin Christ Freeze Dryers GmbH. The drying method was set to -40 °C and 0.1 mbar.

Docking

Maestro 11.5 Schroedinger was used for the images.

DAPGCP, DAPLysGCP and PEG700 were used for the docking. The molecules were prepared with LigPrep. A model of the Taspase1 crystal structure extended by a NMR based structure of the loop (Taspase1_40-420_van_den_Boom [1][2]) was used for the grids.

A grid around the amino acids Arg190/201 Lys 225/218 with a size of 36 A (Loop), Asp233 with a size of 15 A, Asp337 with a size of 15 A and Glu207 with a size of 10 A were generated with glide grid generator.

The prepared molecules and grids were used for Docking. The method was XP (extra precise) and the sampling was flexible. The following conditions were chosen: sample nitrogen inversions, bias sampling of torsions for amides.

Visualisation of Ligand and Loop

The three DapLysGCPs were dragged to the Asp233, Asp337 and Glu207. The PEG3000 was coiled by hand and put at the loop. After that a minimization was performed.

Cloning

The plasmid for the inactive Taspase1$_{D233/T234A}$ mutant was generated as previously described [3].

The gene for Importinα was amplified from a "pc3DNA-Importinα-HA" plasmid and the ends modified via PCR (Forward primer: CAGGGGCCCTCCACCAACGAGAATGCTAAT, Reverse primer: TTCGGATCCTTAGAGAAAGTTAAAGGTCCC). The gene, now with overhangs for ApaI/BamHI digestion, was cloned in a blunt pJET1.2 vector (Thermo Fisher) according to the CloneJET PCR cloning kit (Thermo Fisher). After transfection of *E. coli* NEB-10ß (New England BioLabs), the plasmid was amplified using the NucleoBond Xtra Midi kit (Macherey-Nagel). The sequence for Importinα was then ApaI/BamHI cloned into a modified pET-41b vector containing an N-terminal GST tag and a PreScission protease cleavage site (GeneArt). The plasmid was then again amplified using *E. coli* NEB10-ß and isolated with the NucleoBond Xtra Midi kit. The sequence was verified by sequencing.

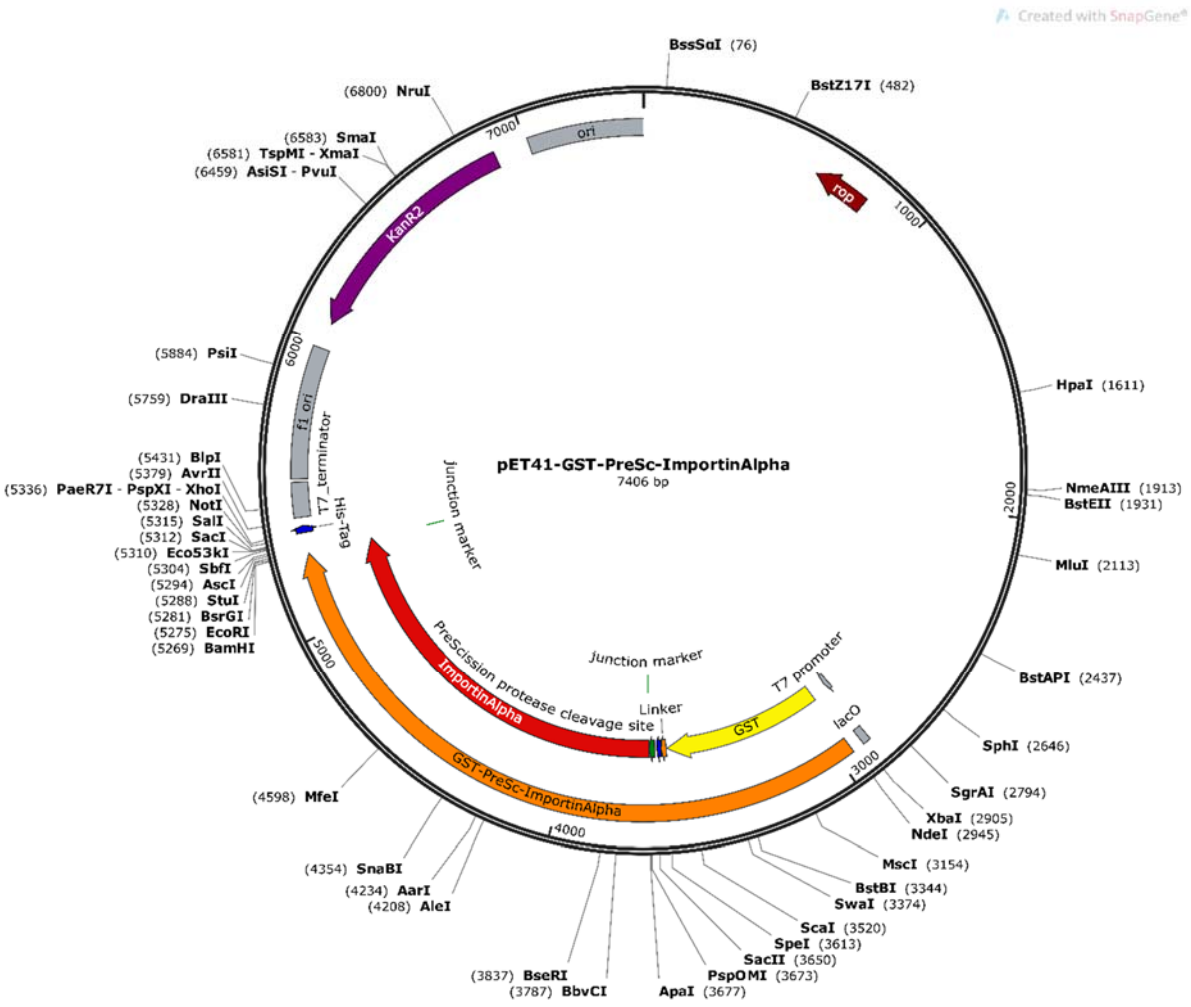

Figure S 1: Map of the plasmide pET41-GST-PreScn-Importinα generated by cloning. The map was created using "Gene Construction Kit" (Texco Biosoftware) and visualized with "Snap Gene Viewer" (GSL Biotech).

Purification of recombinant proteins

pET22-Taspase1$_{D233A/T234}$-His was expressed in *E. coli* BL21 (DE3). The cells were lysed using ultrasonic sheering and enzymatic lysis by lysozyme. The protein was purified using the His tag for affinity chromatography with a HisTrap FF (GE Healthcare). After imidazole elution the Tasapse1-His containing fractions were pooled and loaded onto a Superdex 200 HiLoad16/600 column (GE Healthcare) for size exclusion chromatography.

pET41-GST-PreScn-Importinα was expressed in *E. coli* BL21 (DE3), the cells were lysed using sonication and enzymatic lysis by lysozyme and the soluble fraction obtained with centrifugation and filtration. The protein was purified using the GST tag for affinity chromatography with a GSTrap 4B (GE Healthcare). After glutathion elution, the GST-Importinα containing fractions were pooled and loaded onto a Superdex 200 HiLoad16/600 column (GE Healthcare) for size exclusion chromatography. GST-Importinα containing fractions were pooled, frozen in liquid nitrogen and stored at – 20°C.

Pull-down assay

For this assay, all solutions were prepared with Dulbecco's Phosphate Buffered saline (Sigma-Aldrich) containing 0,1 % (v/v) Triton X-100 (Carl Roth) and 1 mM DTT (Carl Roth) (PBST), all incubation steps were carried out at 4 °C to preserve the proteins, all centrifugation steps were carried out at 400 xG and samples taken for later analysis were mixed with 5x sample buffer and heated to 95 °C for 5 Min. 50 µM Glutathione Sepharose 4 B (Merck) were transferred to a Spin Column (IBA Lifescience), equilibrated with 500 µL PBST followed by centrifugation. 500 µL 2,5 µM GST-Importinα were added to the column, a sample for the "Input" fraction was taken and the column then incubated for 2 h on a rotator. Unbound protein was then removed by three washing steps with PBST followed by centrifugation. 500 µL 2,2 µM inactive Taspase1-His with the respective concentration of compound were pre-incubated for 1 h on a rotator and a sample for the "Input" fraction was taken. The free binding sites on the column were blocked with 1 % (w/v) BSA (Carl Roth) in PBST for 30 Min on a rotator. The blocking solution was removed from the column by centrifugation for 1 Min. After that, the inactive Taspase1-His pre-incubated with the compound was added to the column and allowed to bind for 1 h on a rotator. A sample from the "Unbound" fraction was taken and unbound protein was then removed by three washing steps with PBST followed by centrifugation for 1 Min. 500 µL 1x sample Buffer were added to the column and heated to 95 °C for 10 Min. The proteins were eluted by centrifugation for 2 Min.

SDS-PAGE and Western Blotting

For these assays, we used the standard recipes for SDS-PAGE according to Laemmli [4] and for Western Blotting according to Towbin [5]. For SDS-PAGE, Tris-glycine gels with 10 % (v/v) acrylamide in the stacking gel and 4 %(v/v) acrylamide in the separating gel were cast according. For the electrophoresis, we used the TetraCell system (BioRad) set to 200 V for 45 Min. The proteins were then transferred to a nitrocellulose membrane using a wet blot tank (Peqlab) set to 360 mA for 90 Min at 4 °C. To detect the different proteins, the membrane was first reversibly stained with Ponceau S (AppliChem) and then cut between the protein bands according to the was cut according to the Spectra Multicolor Broad Range Protein Ladder (Thermo Fisher). Free binding sites were blocked with 5 % (w/v) powdered milk (Carl Roth) in Tris buffered saline with Tween-20 (TBST) (Carl Roth) for 30 Min at room temperature. After that, the membranes were incubated with the respective primary antibodies rabbit anti-Taspase1 1:2000 (sc-85945,Santa Cruz) and mouse anti-Karyopherinα2 1:1000 (sc-55538, Santa Cruz) in 5 % (w/v) powdered milk in TBST for 1 h at room temperature. Unbound antibodies were removed by three washing steps with TBST. The membranes were incubated with the respective secondary antibodies donkey anti-rabbit HPR-coupled 1:10000 (NA934,

GE Healthcare) and sheep anti-mouse HPR-coupled 1:10000 (NXA931, GE Healthcare) in 5 % w/v) powdered milk with TBST for 1 h. Unbound antibodies were removed by four washing steps in TBST. For the detection of chemiluminescence, we used Pierce ECL Plus Western Blotting Substrate (Thermo Fisher) and the ChemidocImaging System (BioRad).

The signal was quantified with “Fiji” [6]. If necessary, the signal of Taspase1 in the eluted fraction was corrected for the Taspase1 stuck to the column without Importinα. To correct possible loading differences the signal of Taspase1 in the eluted fraction was then normalized for the signal of Importinα in the eluted fractions. The data was evaluated using “Origin2019” (OriginLab).

Toxicity Assay

1 x 10^4 cells were cultured in Corning 96 Well microplates (Sigma-Aldrich) in 100 µl Dulbecco’s modified eagle medium (DMEM) (Thermo Fisher Scientific) supplied with 10 % (v/v) fetal calf serum (FCS) (Life Technologies GmbH), Antibiotic-Antimycotic (Life Technologies GmbH) and the respective compound concentration. The cells were then cultivated at 37 °C and 5 % CO_2 for 24 h. After that, the compound-containing medium was removed, and cells were washed with PBS once. 100 µL fresh DMEM with 10 % FCS and Antibiotic-Antimycotic were added to each well. After the Following addition of 20 µL Cell Titer Aqueous One (Promega), absorption at 490 nm was recorded with the plate reader Promega Glow Max (Promega) after 30 min of incubation. Since the compounds were dissolved in water, the results were then normalized to a water treated control to correct for the dilution of the media. The data are the mean of at least three replicates ± standard deviation.

Isothermal titration calorimetry (ITC)

ITC was performed with MicroCal ITC2000 (Malvern Pananalytical). The proteins were rebuffered five times into a tenfold volume of assay buffer and then concentrated using Vivaspin 6, 10000 MWCO (Sartorius) at 4900 xg and 4 °C. The rebuffered samples were degassed immediately before ITC with MicroCal ThermoVac (Malvern Pananalytical). The experimental setup was 18 injections of 2.0 µL with 1-1.5 mM ligand to 200 µL 70-80 µM inactive Taspase1-His with an injection time of 4 s and 180 s spacing between each injection at 25 °C and with constant stirring at 750 rpm. The first injection set to 0.4 µL to remove air and mixed reactants from the tip. For each experiment, we performed a ligand-to-buffer (LtB) titration as well as a buffer-to-protein (BtP) titration with the same experimental setup to correct for possible heat of dilution introduced by the ligand or protein. The data was analysed using MicroCal Analysis (OriginLab). The first injection peak was discarded and the isotherms of the LtB and BtP controls were subtracted from the experimental isotherm.

At this time, ITC experiments were not successful (see titration curves below). At this point we attribute this to the challenge in stabilizing Taspase1 at high concentrations which requires high ionic strength, which in turn is expected to affect interaction of the GCP units.

Surface Plasmon Resonance (SPR)

The measurements were performed on a Biacore X100 from GE Healthcare Life Sciences, Uppsala, Sweden. The sensograms were recorded with the Biacore X100 Control Software and evaluated with the Biacore X100 Evaluation Software.

For the measurements a C1 sensor chip from GE Healthcare Life Science was used. Before immobilization the sensor chip surface was activated by twofold injection of 0.1 M glycin-NaOH + 0.3 % Triton X 100, pH 12 and followed by washing with HBS-P^+ buffer. Taspase1 was immobilized on the sensor chip surface on flow cell 2 via carbodiimide chemistry by the wizard

template for immobilization. Therefore, a 126 µM stock solution of Taspase1 in PBS buffer was diluted in 10 mM acetate buffer (pH 5.5, GE Healthcare) to get a final protein concentration of 1.26 µM. For flow cell 2 an immobilization level of 1391,6 RU was reached. The flow cell 1 was blocked by a solution of ethanolamine (1 M, pH 8.5, GE Healthcare) and an immobilization level of 1.6 RU was obtained. As running buffer HBS-P$^+$ buffer (pH 7.4, GE Healthcare) at a flow rate of 5 µl min^{-1} was used.

After immobilization the system was primed with the running buffer and two startup cycles were performed. The PEGylated and non-PEGylated GCP macromolecules were injected in concentrations of 0.10 – 400 µM in HBS-P$^+$ buffer with a dilution factor of 2. A flow rate at 5 µL min^{-1} and the contact and dissociation time were 120 s, respectively 180 s were used. After each measurement, the sensor chip was regenerated by injection of 0.1 M arginine in HBS-P$^+$ buffer at a flow rate of 5 µL min^{-1} with a contact time of 90 s to ensure that all sample was washed out and to achieve a stable baseline for the following measurements. For each sample the measurements were repeated three times.

To test for reproducibility, the assay was performed on a second chip prepared as described above giving an immobilization level of 1282,6 RU for flow cell 2 and an immobilization level of 144,6 RU for flow cell 1. Due to different immobilization levels, different absolute values are determined that are, however, in the same range [µM] and show the same trend for the different ligands as observed for the first chip. These measurements were repeated two times.

Measurements for monovalent ligands as well as only PEG were performed on the second chip and repeated two times.

Macromolecule Synthesis:

Synthesis EDS and GCP

(4-((2-(2-(2-aminoethoxy)ethoxy)ethyl)- amino)-4-oxobutanoic)[7] as well as N-Boc-protected 5-(guanidinocarbonyl)-1H-pyrrole-2-carboxylic acid (GCP) as triethylamine salt [8] were synthesized according to literature procedures. The free acid of the GCP was obtained by crystallization from methanol.

General

Oligomer synthesis were carried out manually in 10 ml polypropylene syringe reactors with a polyethylene frit and a Luer stopper from Multisyntech GmbH. All oligomers were synthesized on the TentaGel® S RAM (Rink Amide) or TentaGel® PAP resin with a loading of 0.23 mmol/g. Batch size of all oligomers was 0.15 mM.

Fmoc cleavage

The resin was swollen in DCM for 30 min and subsequently washed three times with DMF. Secondly the Fmoc protecting group of the resin as well as from the coupled building blocks or amino acids was cleaved by means of a 25% solution of piperidine in DMF achieving an amine end group. The deprotection was carried out twice for 20 min. Afterwards the resin was washed 10 times with DMF.

Coupling protocol

First the resin was swollen in DCM for 30 min and then washed three times with DMF. The Fmoc protecting group has to be removed before further couplings!

For the building block, amino acid or GCP (5 eq.), 5 eq. PyBOP and 10 eq. DIPEA were dissolved in 5 ml DMF, drawn into the reactor syringe and shaken for 90 min, followed by washing ten times with DMF. A double coupling, adding fresh building block and coupling reagents, was performed each time the GCP motif was coupled.

Capping of N-terminal primary amine

With scaffold completion of the oligomer the N-terminal amine group was acetylated with 8 ml of acetic anhydride, shaking for 20 min. After that, the resin was washed 5 times with DMF.

Boc-Cleavage

For Boc-deprotection, 6 ml of a 4 M HCl in dioxane solution (2 ml HCL conc. and 4 ml dioxane) was drawn into the reactor syringe and shaken for 10 min. Afterwards the reaction mixture was washed 3 times with dioxane and again 6 ml fresh 4 M HCl dioxane solution was drawn into the syringe and shaken for 20 min. Subsequently the solution was removed and the resin washed three times alternately with dioxane and DCM. To neutralize the resin, a 10 volume percent ice-cold DIPEA DCM solution was drawn up twice and shaken for 10 min. Last the resin was washed alternately three times with dioxane and DCM and finally 10 times with DMF.

Cleavage from solid phase

The oligomers were cleaved from the TentaGel® S RAM resin by drawing up a solution of 5 vol% triisopropylsilane (TIPS) and 95 vol% TFA into the syringe and shaking for 1.5 hours. The TentaGel® PAP cleavage was achieved with TFA/thioanisole (95/5, *v*/*v*) for 24 hours at room temperature.

Afterwards the solution was placed in ice cooled diethyl ether. The resulting precipitate was centrifuged off and the supernatant was decanted off. The pellet was washed 3 times with diethyl ether.

The product was dried and dissolved in MilliQ water. The entire solution was collected in a falcon tube and freeze-dried to isolate the product. Subsequently, the products were purified by means of preparative HPLC. Due to the purification by preparative HPLC and the added 0.1 vol% formic acid in the mobile phases, the structures are present as formates. The number of formates was quantified by ^{1}H NMR spectra. Further information can be found in the respective ^{1}H NMR data.

Analytical data for macromolecules:

*Macromolecule **3G***

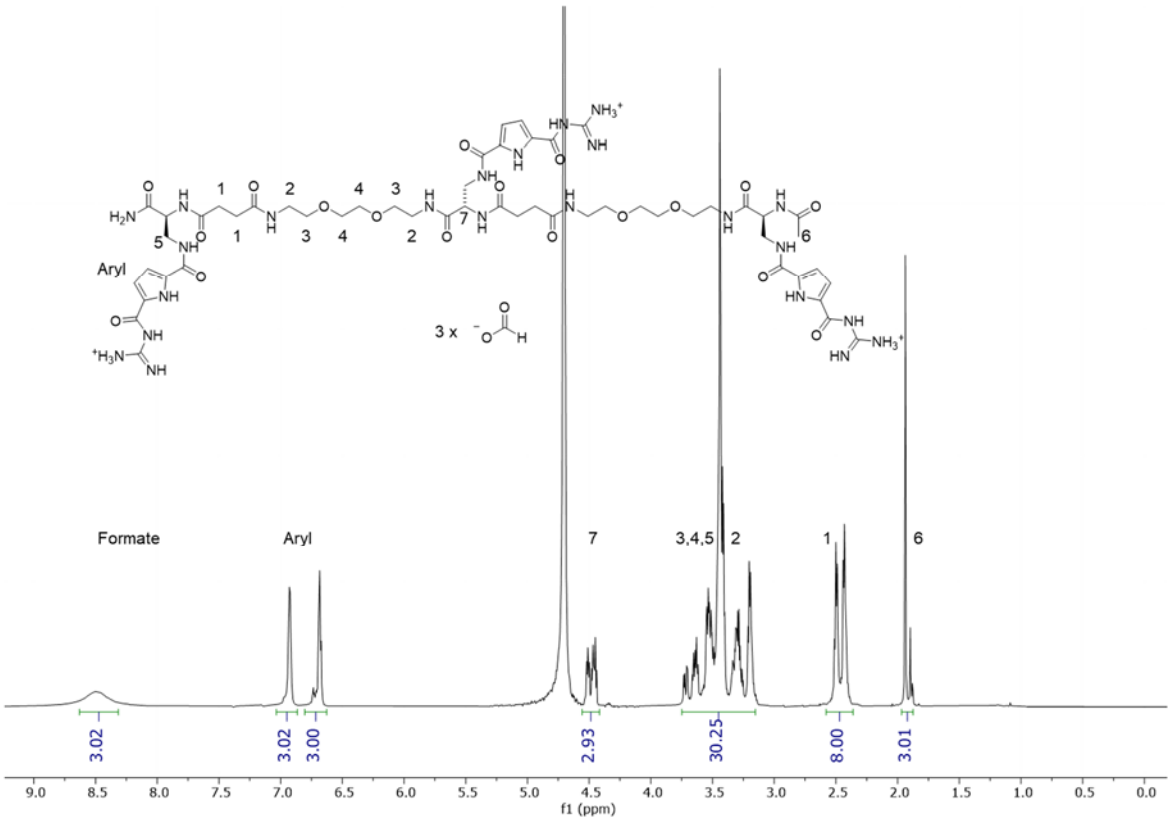

*Figure S 2: 600 MHz ^{1}H NMR spectrum of **3G** as formate salt in D_2O at 25°C.*

1H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.25 to 8.75 signal of the formate (s, 3H, **3G** is present with three formate anions), 7.09-6.85 (m, 3H, Ar-H), 6.79-6.24 (m, 3H, Ar-H), 4.50-4.40 (m, 3H, H7), 3.51-3.16 (m, 30H, H2-H5), 2.63-2.37 (m, 8H, H1), 1.94 (s, 3H, H6, second small signal cannot be assigned).

The effective molar mass for **3G** with three formates is 1450.4 g/mol.

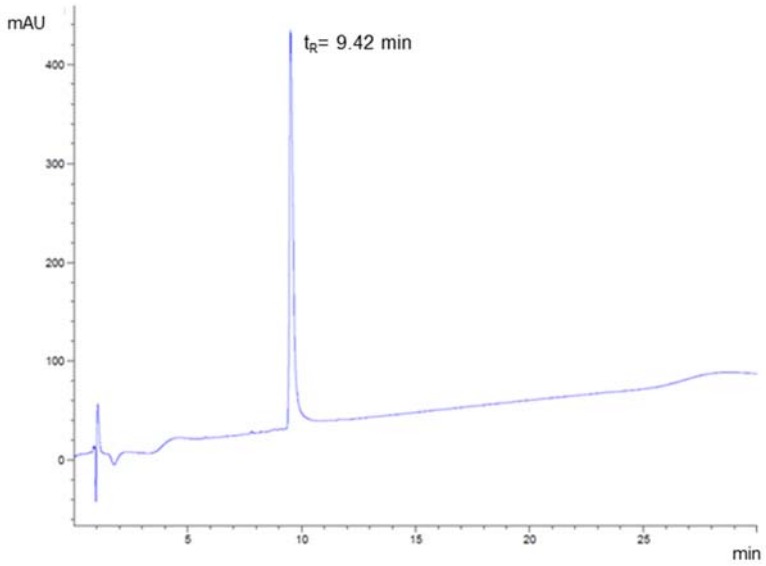

*Figure S 3: **3G** detected with relative purities >95% by RP-HPLC analysis (linear gradient from 5 – 50 vol% eluent H_20/acetonitrile) in 30 min at 40 °C.*

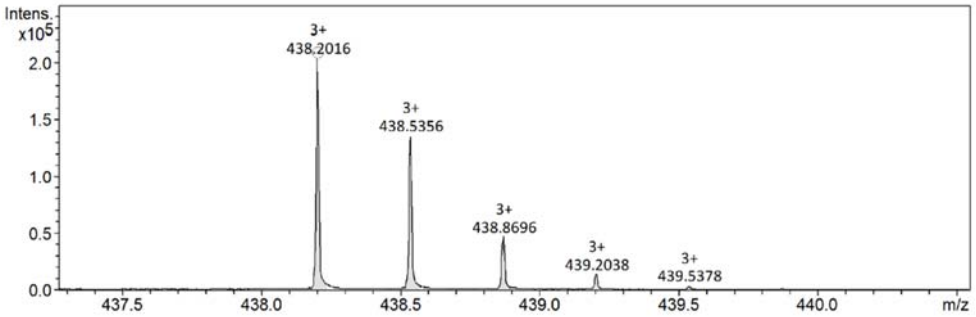

*Figure S 4: HR-ESI-MS of **3G**.*

HR-ESI-MS: for $C_{52}H_{80}N_{23}O_{18}$ m/z $[M+3H]^{3+}$ calcd.: 438.2012, found: 438.2016, mass accuracy -0.9 ppm.

*Macromolecule **3GL***

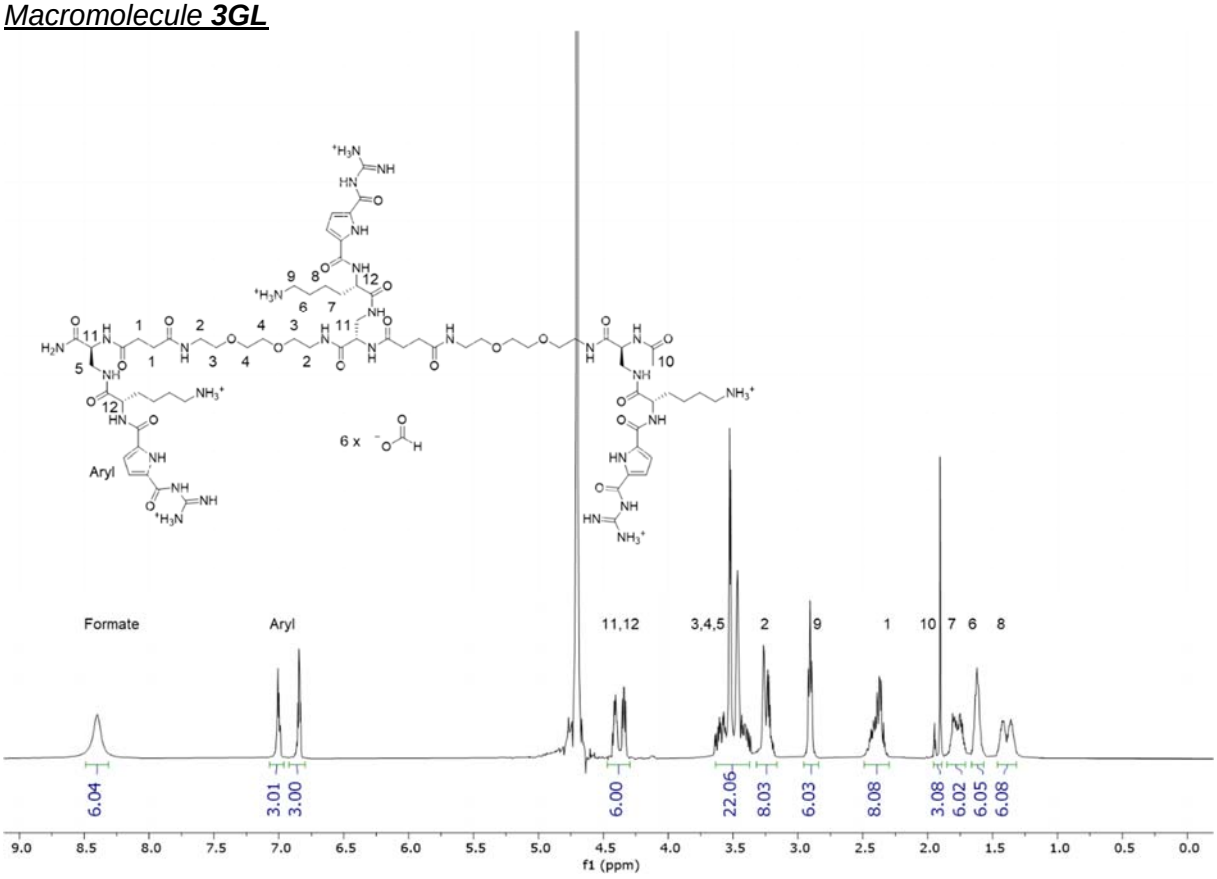

*Figure S 5: 600 MHz ^{1}H NMR spectrum of **3GL** as formate salt in D_2O at 25°C*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.30 to 8.50 signal of the formate (s, 6H, **3GL** is present with six formate anions), 7.12-6.92 (m, 3H, Ar-H), 6.87-6.79 (m, 3H, Ar-H), 4.48-4.27 (m, 6H, H11, H12), 3.69-3.38 (m, 22H, H3-H5), 3.33-3.19 (m, 8H, H2), 2.91-2.82 (m, 6H, H9), 2.49-2.27 (m, 8H, H1), 1.91 (s, 3H, H10, second small signal cannot be assigned), 1.82-1.71 (m, 6H, H7), 1.68-1.59 (m, 6H, H6), 1.45-1.31 (m, 6H, H8).

The effective molar mass for **3GL** with six formates is 1971.7 g/mol.

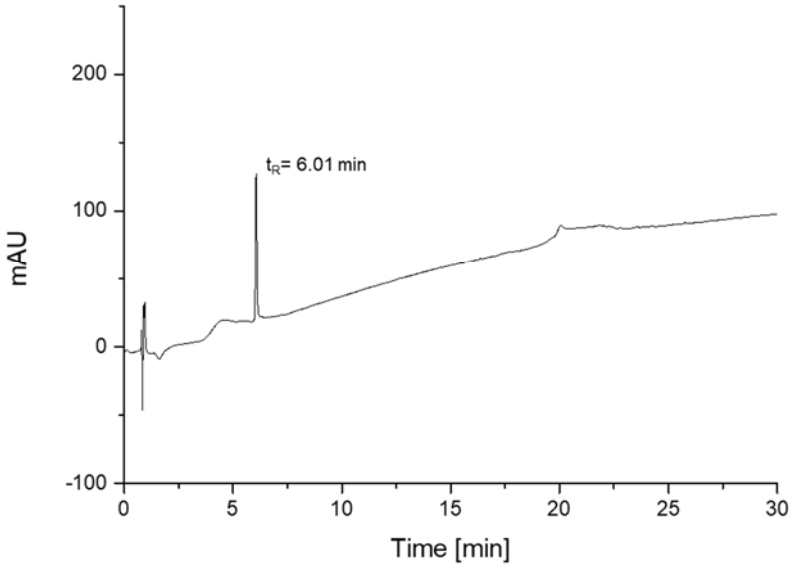

*Figure S 6: **3GL** detected with relative purities >95% by RP-HPLC analysis (linear gradient from 5 – 50 vol% eluent H_20/acetonitrile) in 30 min at 40 °C.*

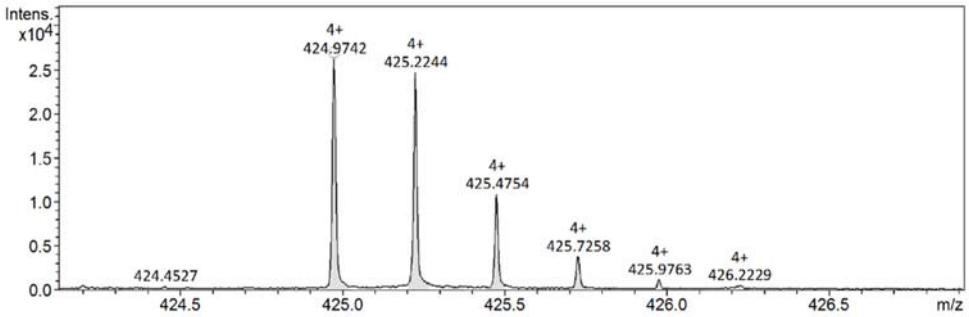

*Figure S 7: HR-ESI-MS of **3GL**.*

HR-ESI-MS: for $C_{70}H_{117}N_{29}O_{21}$ m/z $[M+4H]^{4+}$ calcd.: 424.9739, found: 424.9742, mass accuracy -0.6 ppm.

*Macromolecule **3GP***

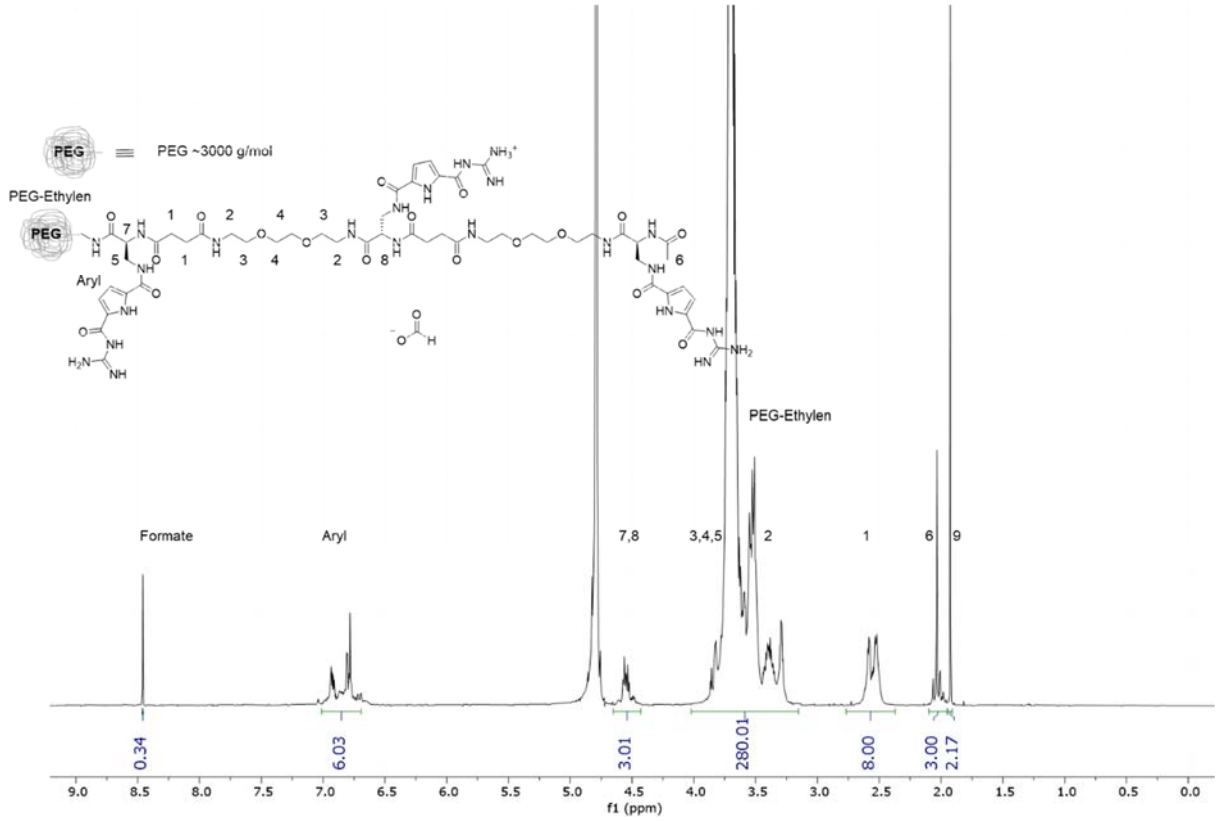

*Figure S 8: 600 MHz ^{1}H NMR spectrum of **3GP** as formate salt in D_2O at 25°C.*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.40 to 8.50 signal of the formate (s, 0.34H, **3GP** is present at an average of 0.34 formate anion per molecule), 7.08-6.65 (m, 6H, Ar-H), 4.70-4.39 (m, 3H, H7, H8), 4.06-3.19 (m, 280H, HPEG, H2-H5), 2.78-2.36 (m, 8H, H1), 2.11-1.97 (m, 3H, H6), 1.94 (s, 2H, H9, signal of the end group of the PEG chain).

The effective molar mass for **3GP** with formate is 4413.4 g/mol.

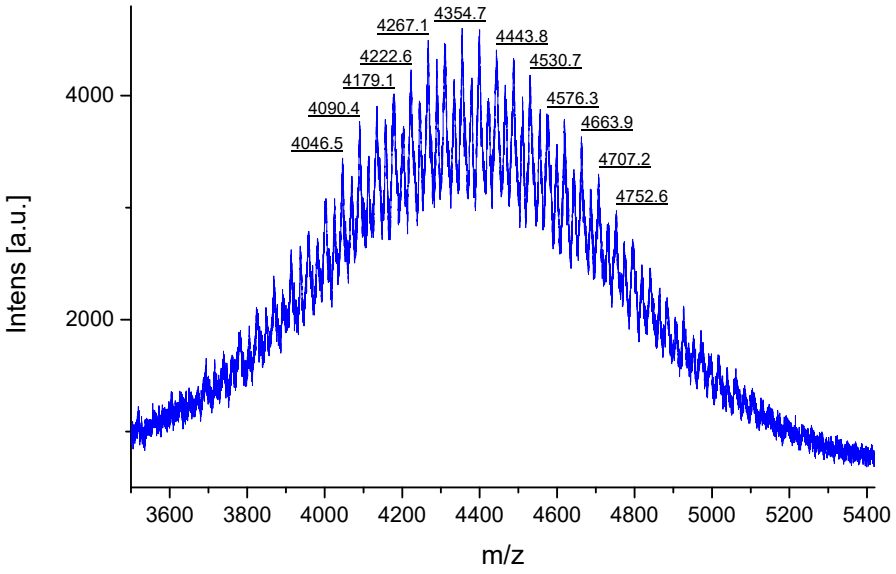

*Figure S 9: MALDI-TOF-MS of **3GP** in a m/z range using DHB as matrix in a compound to matrix ratio of 1:5.*

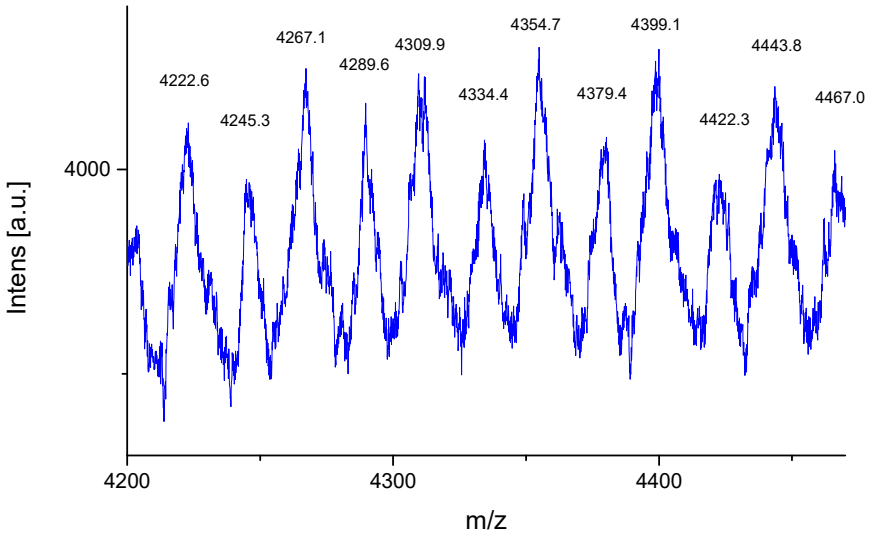

*Figure S 10: Detailed view on MALDI-TOF-MS of **3GP**, focusing on PEG-repeating units (every second signal corresponds to one PEG unit more, intermediate signal corresponds to one additional sodium ion).*

Mass analysis MALDI-TOF-MS of **3GP**: m/z found 4399.1- 3106.8 (PEG-Part) = 1292.3

(m/z calcd. Oligomer Part: 1311.6).

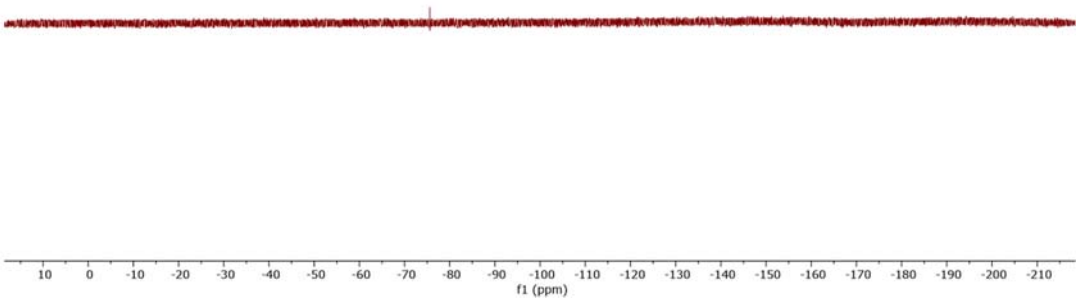

*Figure S 11: Exemplary 600 MHz ^{19}F NMR spectrum of **3GP** in D_2O at 25°C showing that no TFA counterions are present.*

*Macromolecule **3GLP***

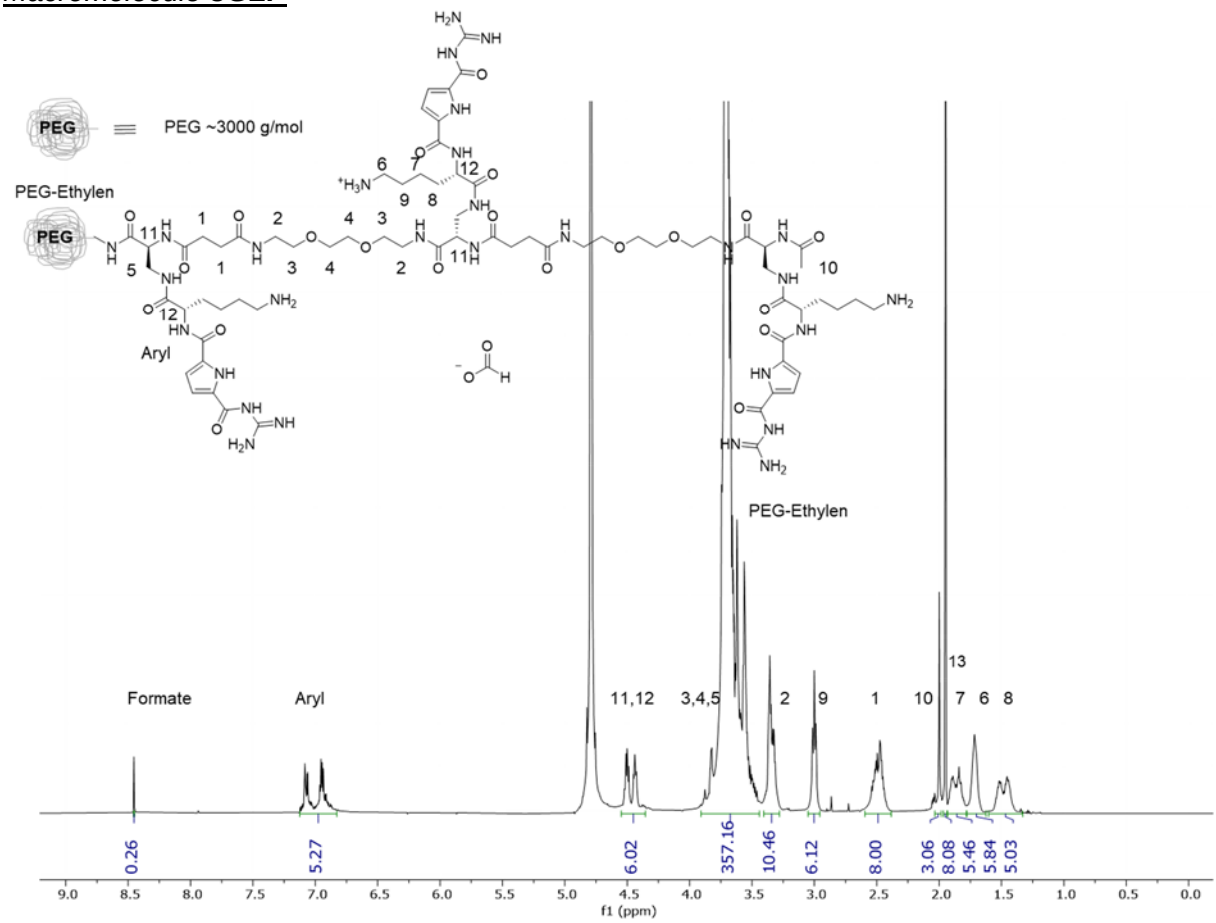

*Figure S 12: 600 MHz ^{1}H NMR spectrum of **3GLP** as formate salt in D_2O at 25°C.*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.40 to 8.50 signal of the formate (s, 0.26H, **3GLP** is present at an average of 0.26 formate anion per molecule), 7.18-6.79 (m, 6H, Ar-H), 4.57-4.32 (m, 6H, H11, H12), 3.82-3.46 (m, 357H, HPEG, H3-H5), 3.41-3.27 (m, 10H, H2), 3.08-2.90 (m, 6H, H9), 2.61-2.31 (m, 8H, H1), 2.04 (m, 3H, H10). 2.00-1.93 (m, H13, signal of the end group of the PEG chain) 1.90-1.78 (m, 6H, H7), 1.76-1.62 (m, 6H, H6), 1.59-1.30 (m, 6H, H8).

The effective molar mass for **3GLP** with formate is 4818.8 g/mol.

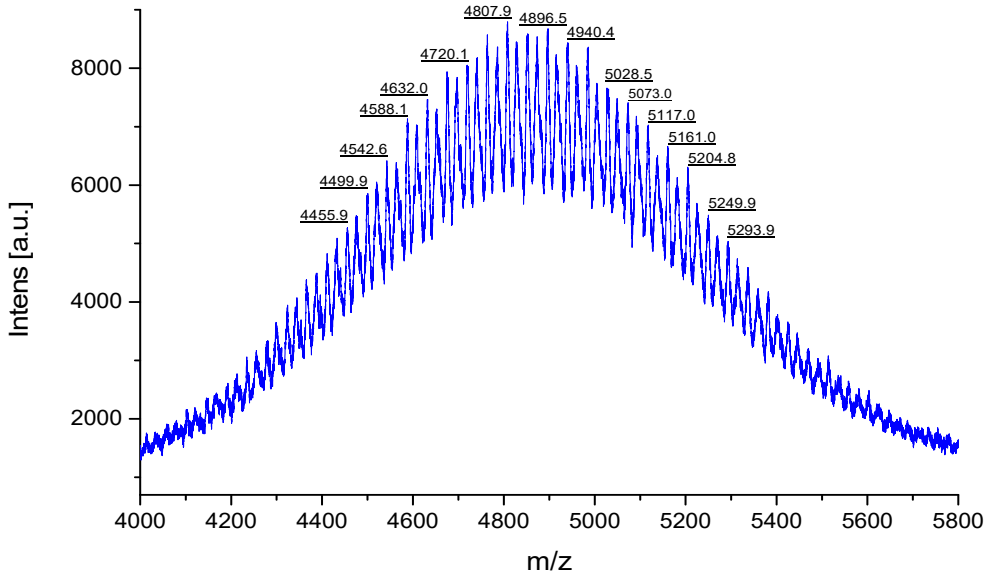

*Figure S 13: MALDI-TOF-MS of **3GLP** in a m/z range using DHB as matrix in a compound to matrix ratio of 1:10.*

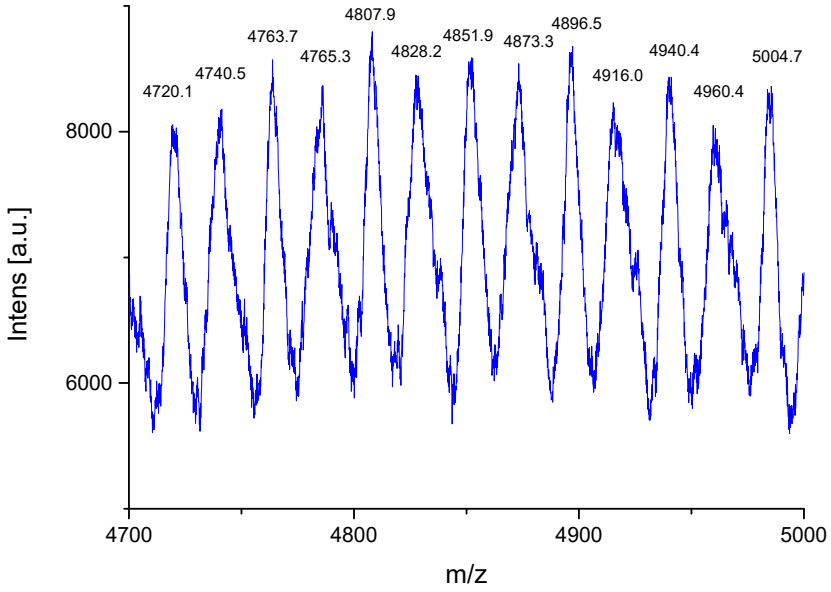

*Figure S 14: Detailed view on MALDI-TOF-MS of **3GLP**, focusing on PEG-repeating units (every second signal corresponds to one PEG unit more, intermediate signal corresponds to one additional sodium ion).*

Mass analysis MALDI-TOF-MS of **3GLP**: m/z found 4807.9- 3106.8 (PEG-Part) = 1701.2

(m/z calcd. Oligomer Part: 1695.7).

*Macromolecule **G***

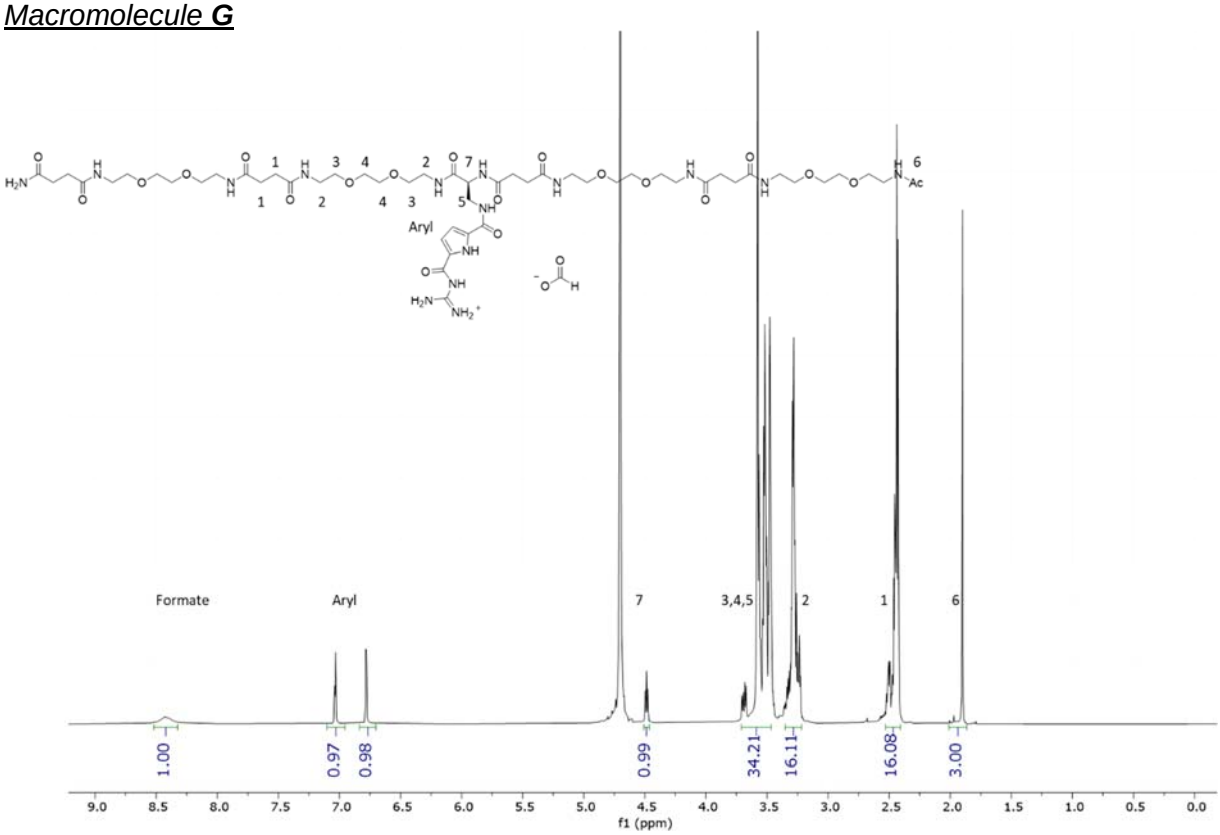

*Figure S 15: 600 MHz ^{1}H NMR spectrum of **G** as formate salt in D_2O at 25°C.*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.30 to 8.50 signal of the formate (s, 1H, **G** is present with one formate anion), 7.08 (s, 1H, Ar-H), 6.78 (s, 1H, Ar-H), 4.52-4.44 (m, 1H, H7), 3.69-3.41 (m, 34H, H3-H5), 3.35-3.23 (m,16H, H2), 2.56-2.39 (m, 16H, H1), 1.92 (s, 3H, H6).

The effective molar mass for **G** with one formate is 1291.42 g/mol.

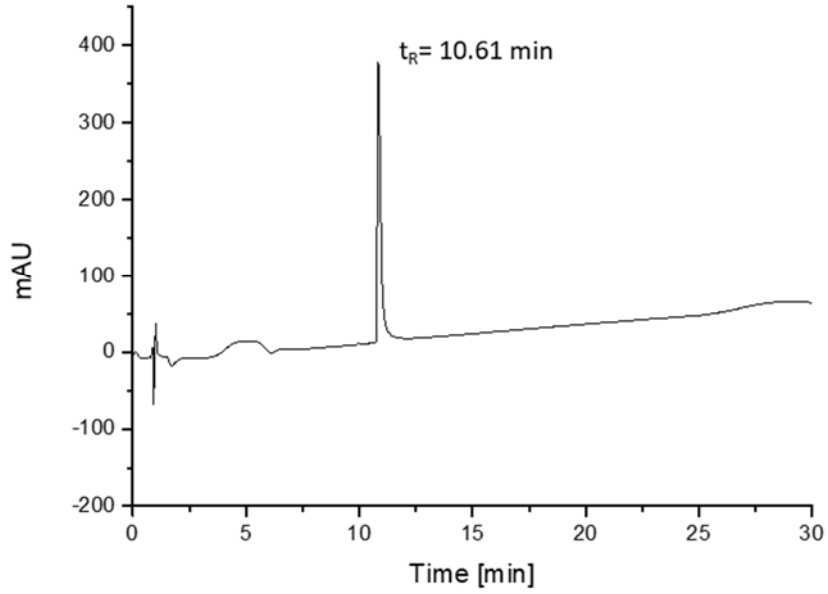

*Figure S 16: **G** detected with relative purities >95% by RP-HPLC analysis (linear gradient from 5 – 50 vol% eluent H_20/acetonitrile) in 30 min at 40 °C.*

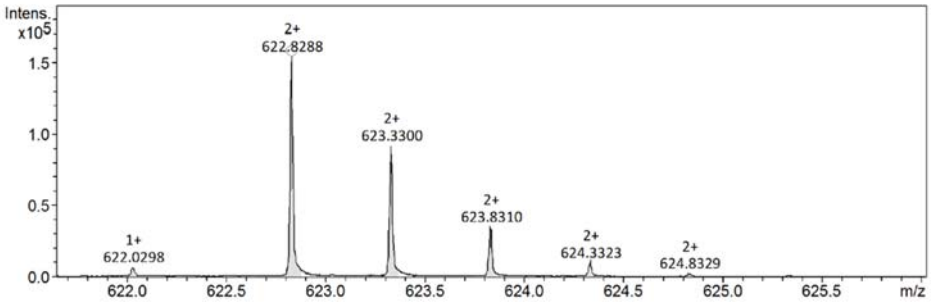

*Figure S 17: HR-ESI-MS of **G**.*

HR-ESI-MS: for $C_{52}H_{99}N_{15}O_{20}$ m/z $[M+2H]^{2+}$ calcd.: 622.8277, found: 622.8288, mass accuracy -1.7 ppm.

*Macromolecule **GL***

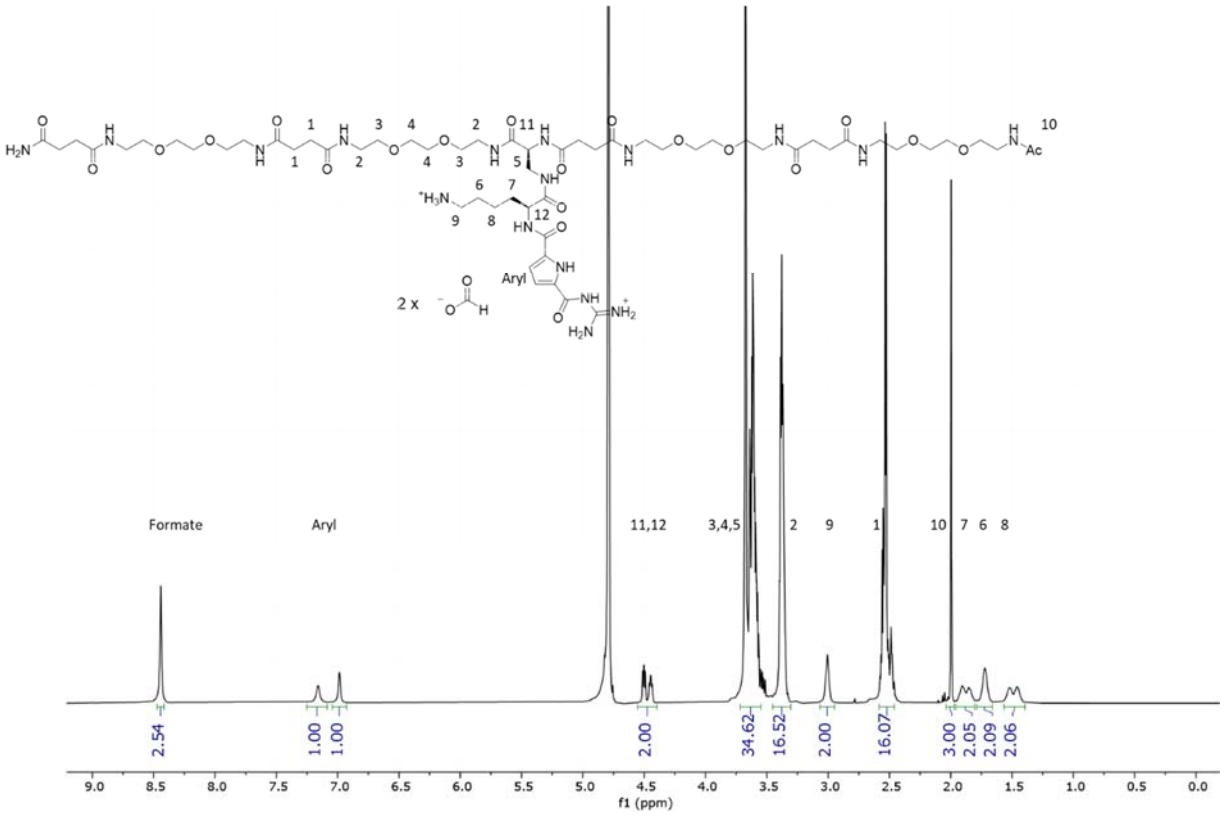

*Figure S 18: 600 MHz ^{1}H NMR spectrum of **GL** as formate salt in D_2O at 25°C.*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.40 to 8.50 signal of the formate (s, 2H, **GL** is present with two formate anions) amide, guanidino functionalities occur, 7.17 (s, 1H, Ar-H), 7.01 (s, 1H, Ar-H), 4.56-4.40 (m, 2H, H11, H12), 3.73-3.54 (m, 34H, H3-H5), 3.36-2.27 (m,16H, H2), 3.07-2.95 (m, 2H), H9), 2.59-2.44 (m, 16H, H1), 2.01 (s, 3H, H10), 1.93-1.80 (m, 2H, H7), 1.76 (s, 2H, H6), 1.57-1.40 (m, 2H, H8).

The effective molar mass for **GL** with two formates is 1457.7 g/mol.

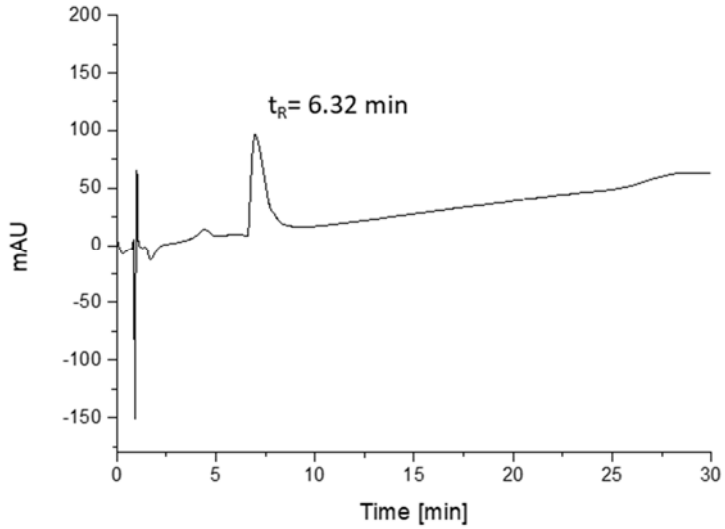

*Figure S 19: **GL** detected with relative purities >95% by RP-HPLC analysis (linear gradient from 5 – 50 vol% eluent H_20/acetonitrile) in 30 min at 40 °C.*

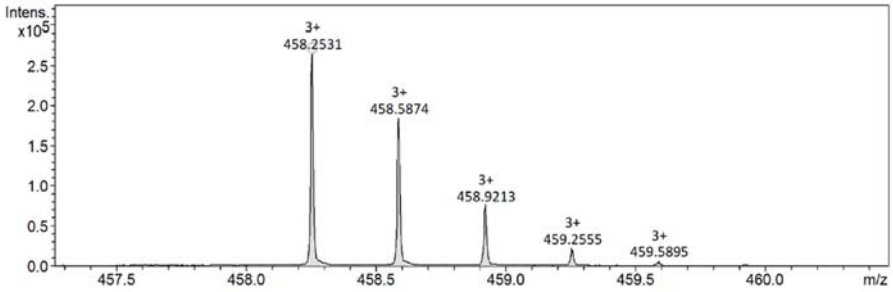

*Figure S 20: HR-ESI-MS of **1GL**.*

HR-ESI-MS: for $C_{58}H_{101}N_{17}O_{21}$ m/z $[M+3H]^{2+}$ calcd.: 458.2525, found: 458.2531, mass accuracy -1.2 ppm.

PEG

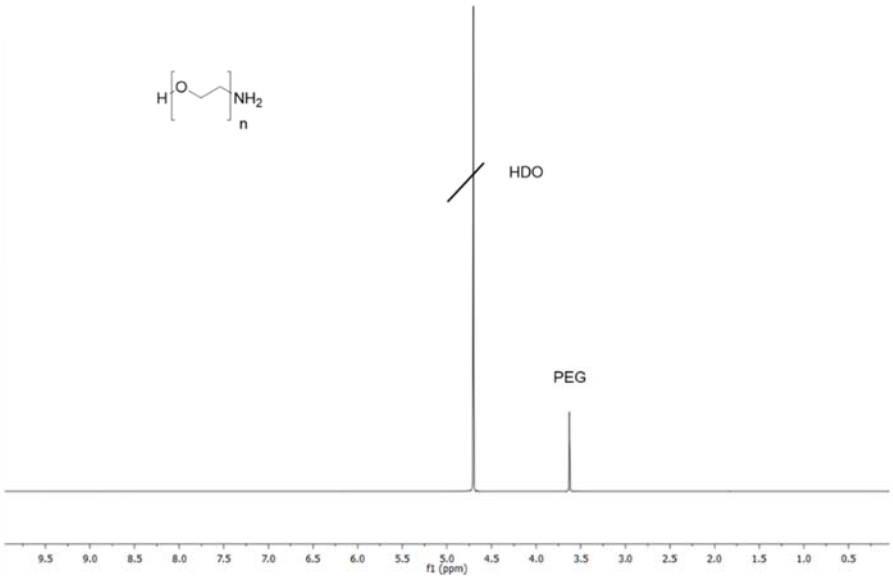

*Figure S 21: 600 MHz ^{1}H NMR spectrum of **PEG** in D_2O at 25°C.*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = 3.64 (s, HPEG).

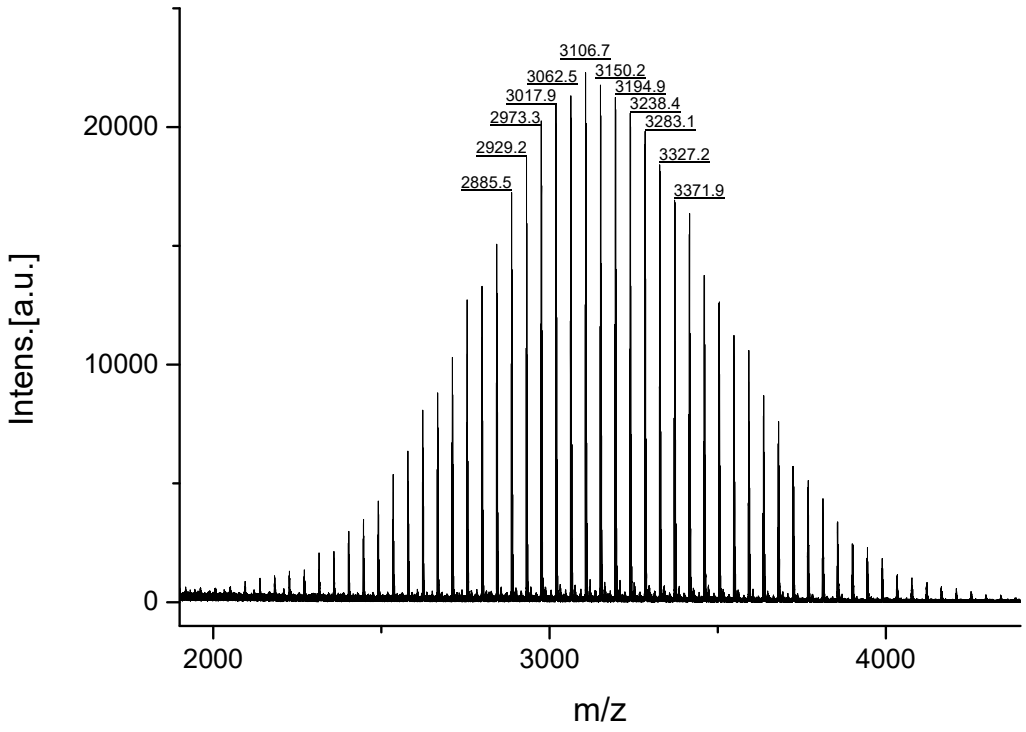

*Figure S 22: MALDI-TOF-MS of **PEG** in a m/z range using DHB as matrix in a compound to matrix ratio of 1:10. The value 3106.7 g/mol corresponds to 70 polyethylene glycol repeating units.*

Binding to Taspase1: SPR Assay

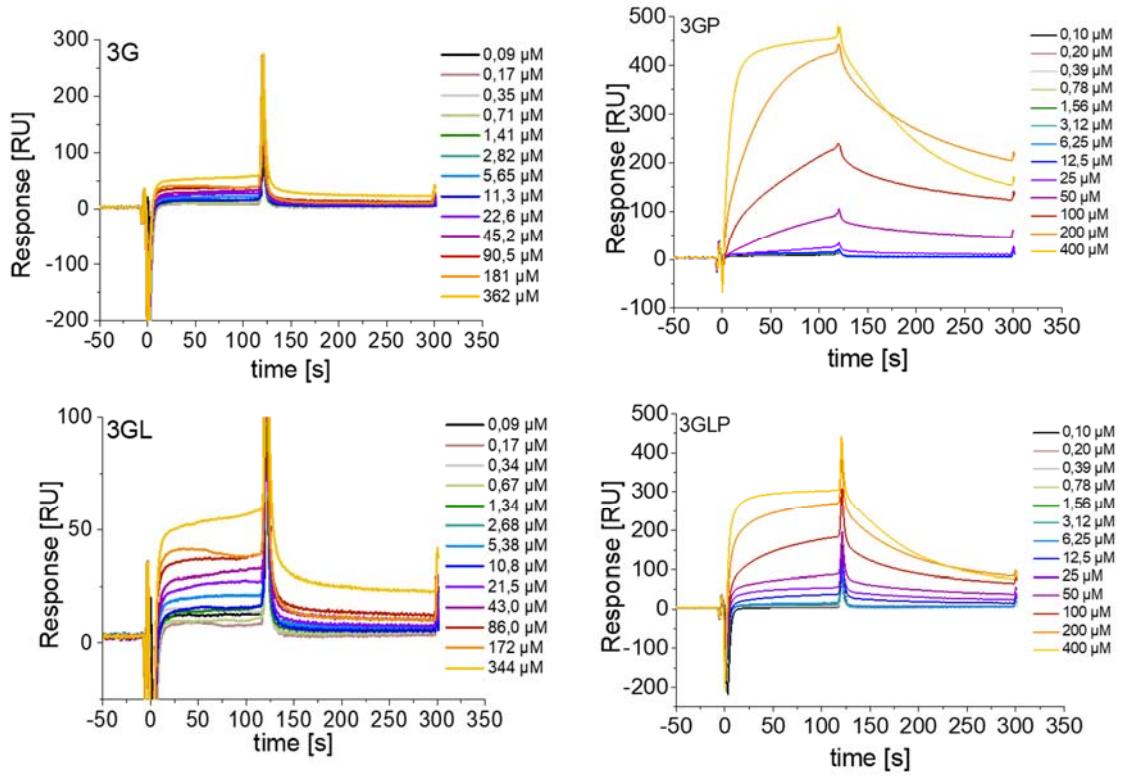

Figure S 23: SPR-Sensograms for PEGylated und non-PEGylated GCP macromolecules for the second C1 sensor chip. Each measurement was repeated two times.

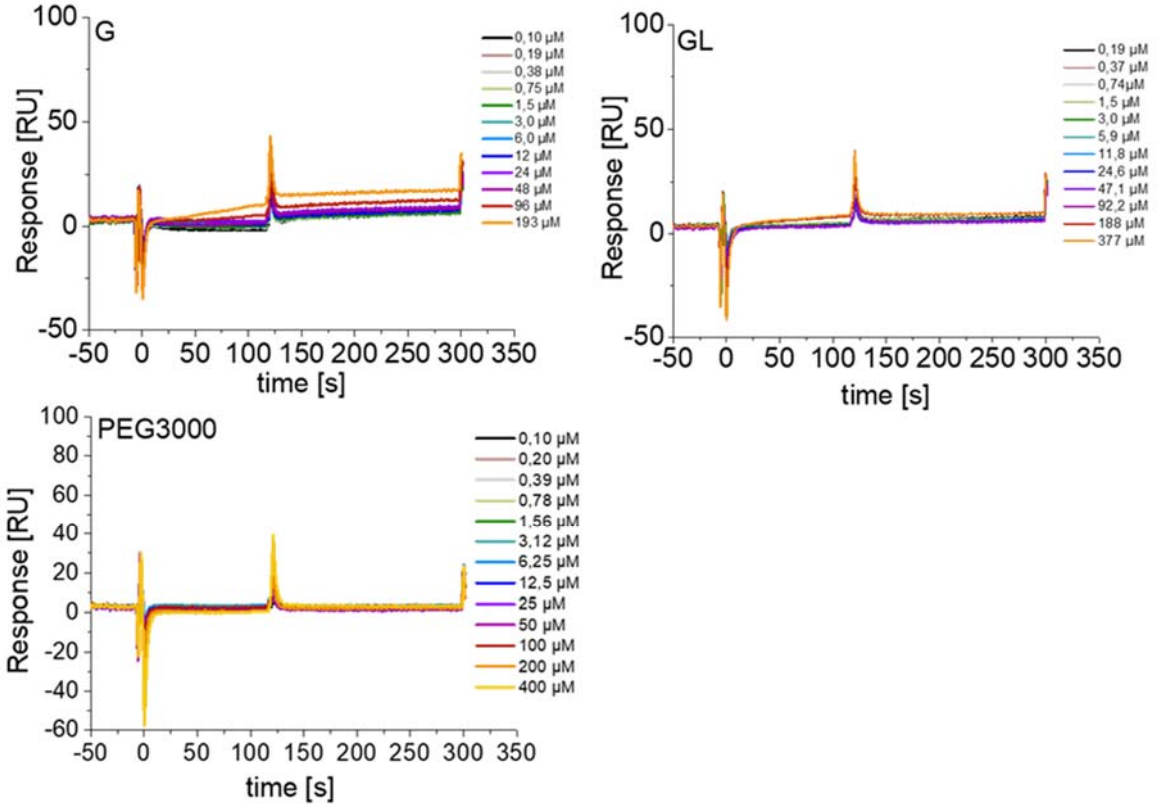

Figure S 24: SPR-Sensograms for monovalent non-PEGylated GCP macromolecules (top) and PEG3000 (bottom). The measurements showed no binding to Taspase1. Each measurement was repeated two times on the second C1 sensor chip.

Binding to Taspase1: Pull-down assay

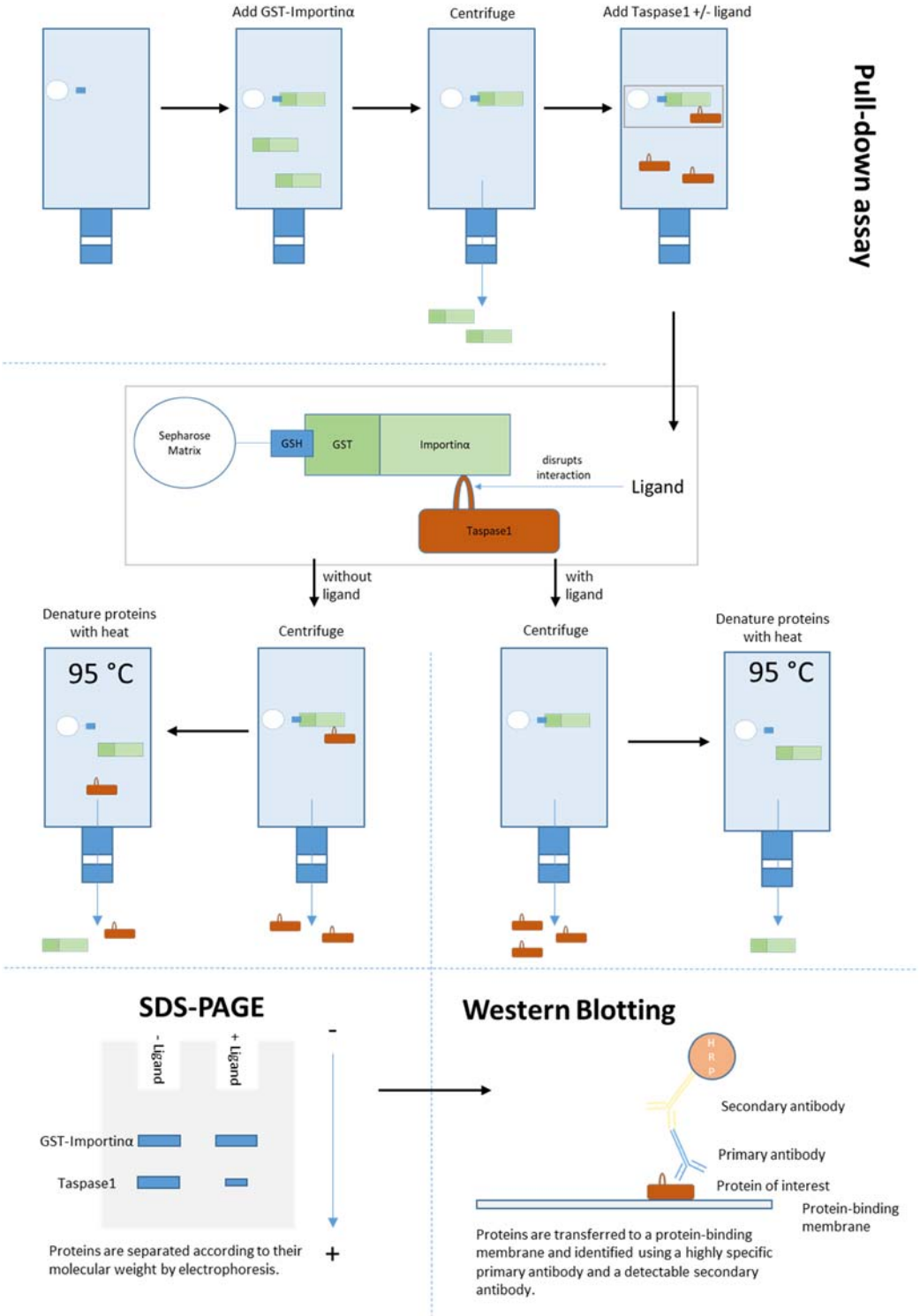

Figure S 25: Workflow of the modified pull-down assay. A spin column was used to fix GST-Importin α on a Sepharose matrix coated with glutathione. First, GST was allowed to bind to glutathione with high affinity, and unbound protein was removed by centrifugation. Then, Taspase1-His was pre-incubated with ligand or left untreated as indicated, subsequently added to the column, and unbound protein was again removed by centrifugation. Next, a buffer containing ionic detergents as well as reducing agents was applied to the column and heated to 95 °C to denature and thus dissociate all protein from the matrix. Finally, the proteins were separated according to their molecular weight by SDS page and analyzed by Western Blot analysis for quantification.

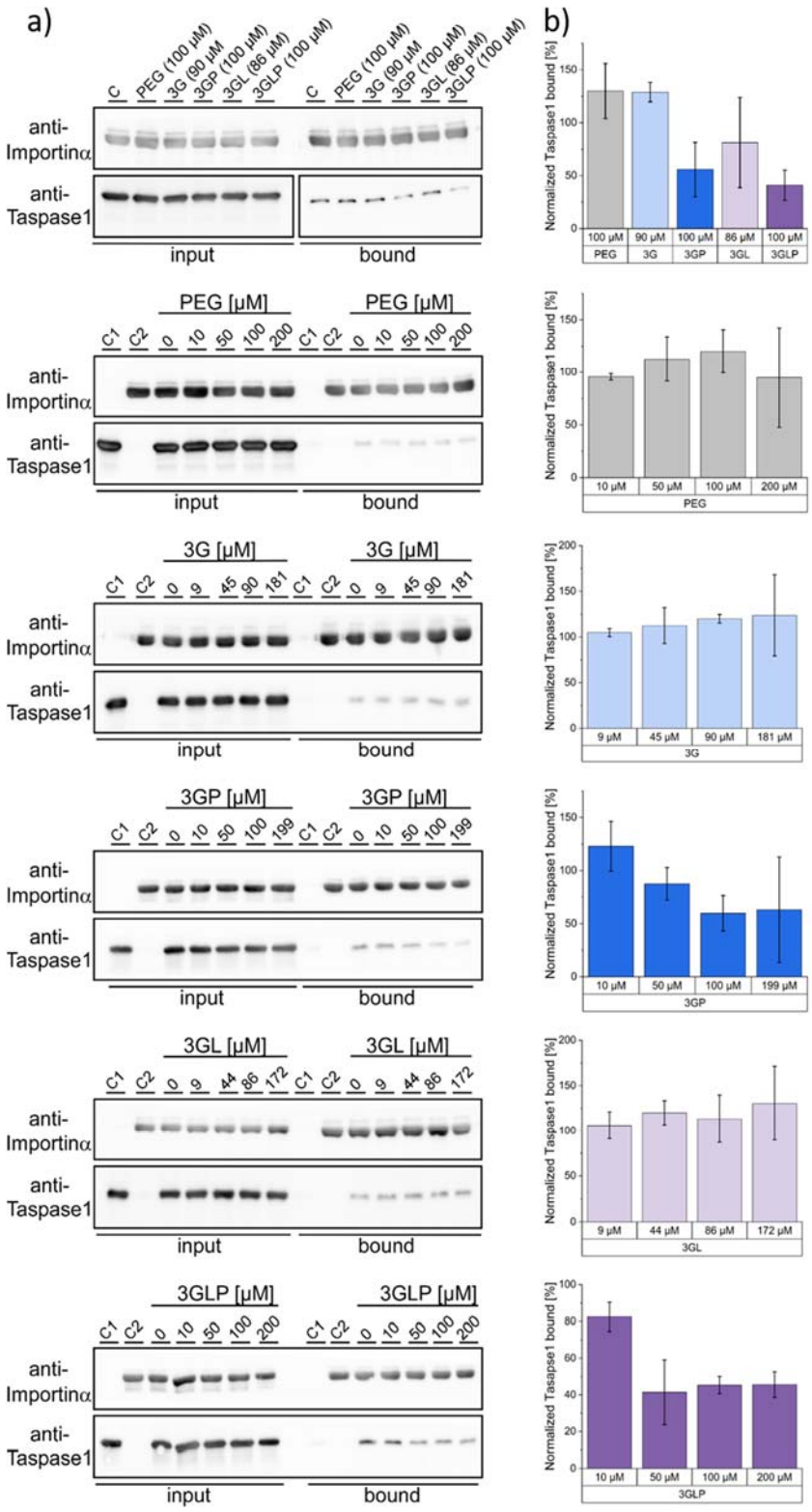

Figure S 26: a) Representative Western Blot analyses from the pull-down assays performed with the different compounds. The input fraction contains samples of the complete protein preparation added to the column, the bound fraction contains the respective portion bound to the column. The latter comprises GST-Importin α directly associated with the column and Taspase1 indirectly bound via its interaction with Importinα. C = Untreated control, C1 = Control with only Taspase1, C2 = Control with only Importin α. b) Densiometric quantification of the respective pull-down assays, comprising three replicates ± standard deviation. ***Please note:*** *originally concentrations were calculated not considering the counterions present in the structures. This was corrected leading to the here shown concentrations.*

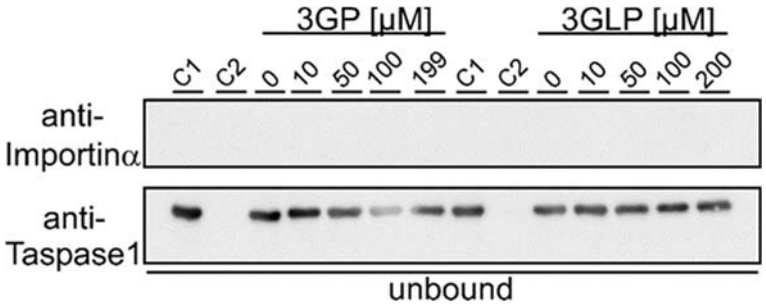

Figure S 27: Binding of Importin α to the column was not affected by the PEGylated ligands during the assay. Western Blot of the unbound fraction after incubation of the Importin α-loaded column with Taspase1 in the presence of the indicated ligands. C1 = Control with only Taspase1, C2 = Control with only Importin α.

Binding to Taspase1: ITC measurements

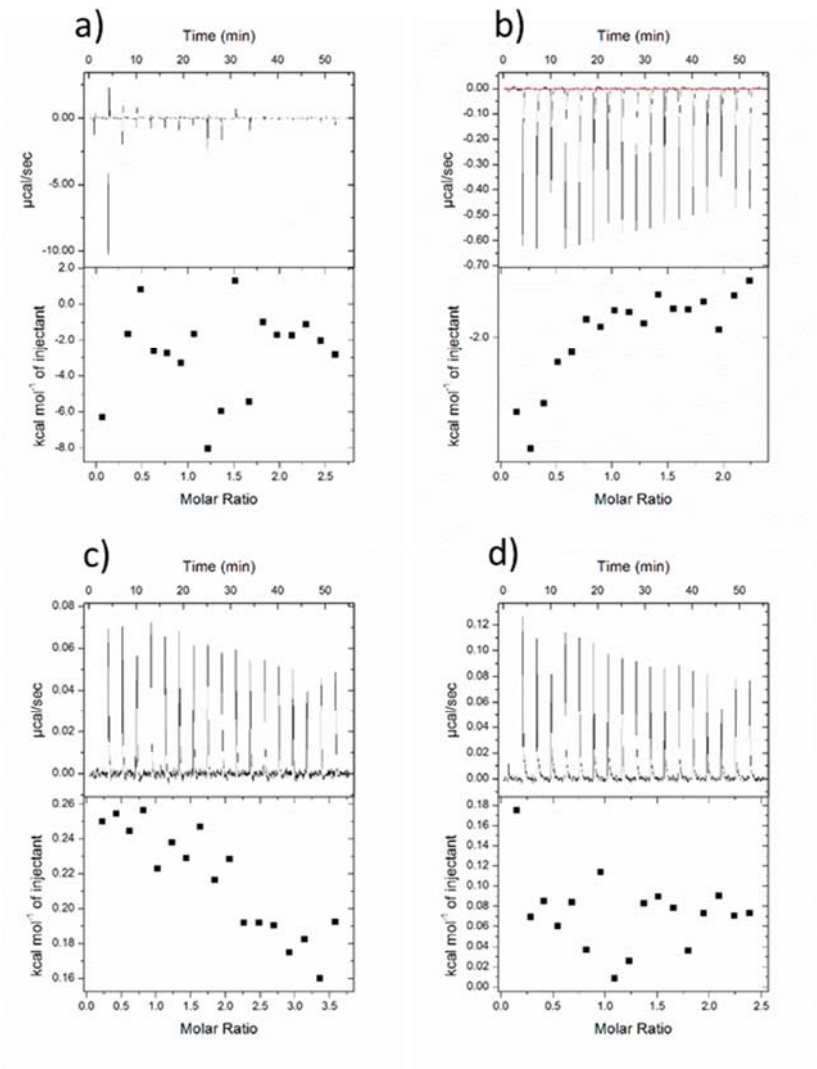

Figure S 28: Test of different ITC conditions with inactive Taspase1 and the respective ligands. a) 1 mM GLP to 70 µM inactive Taspase1, buffer: 150 mM NaCl, 50 mM NaH_2PO_4, pH 7,4 . b) 1 mM GLP to 80 µM inactive Taspase1, buffer: 10 % (w/v) Sucrose, 50 mM NaH_2PO_4, pH 7,4. c) 1,5 mM GLP to 75 µM inactive Taspase1, buffer: 10 % (w/v) Sucrose, 50 mM NaH_2PO_4, pH 7,4. d) 1 mM GP to 75 µM inactive Taspase1, buffer: 10 % (w/v) Sucrose, 50 mM NaH_2PO_4, pH 7,4. e) 1 mM GL to 75 µM inactive Taspase1, buffer: 10 % (w/v) Sucrose, 50 mM NaH_2PO_4, pH 7,4.

Toxicity study

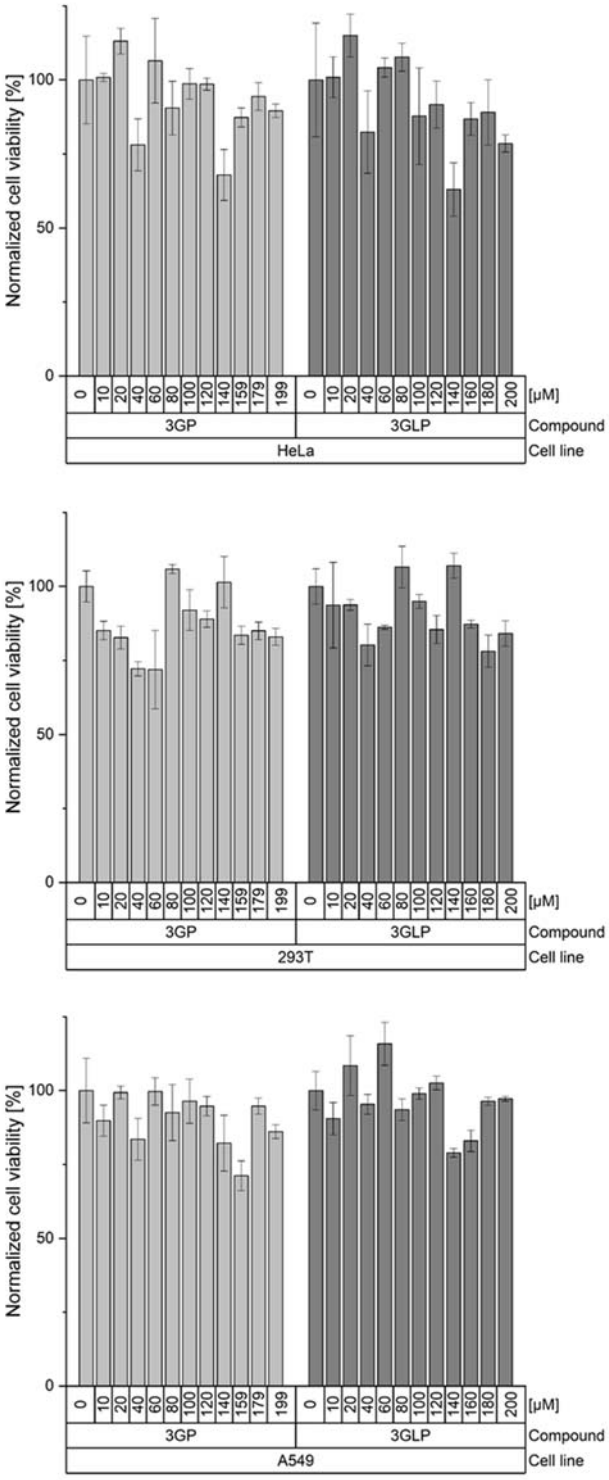

Figure S 29: The compounds do not affect the cell viability of various tumor cell lines. 293T (a), A549 (b) and HeLa (c) cells were cultivated in cell culture medium supplied with the respective concentrations of compound for 24 h. After that, we performed an MTS assay to determine the cell viability. The data points are the mean of triplicates ± standard deviation. ***Please note:*** *originally concentrations were calculated not considering the counterions present in the structures. This was corrected leading to the here shown concentrations.*

References:

[1] J. van den Boom, F. Trusch, L. Hoppstock, C. Beuck, P. Bayer, *PloS one* **2016**, *11* (3), 1-13.

[2] J. A. Khan, B. M. Dunn, L. Tong. *Structure*. **2005**, 13(10), 1443-1452.

[3] J. van den Boom, A. Hensel, F. Trusch, A. Matena, S. Siemer, D. Guel, D. Docter, A. Höing, P. Bayer, R. H. Stauber, S. K. Knauer, *Nanoscale* **2020**, *12*, 19093-19103.

[4] S. Ramakrishnan, K. N. Sulochana, R. Punitham, K. Arunagiri, *Glycoconj. J*. **1996**, *13*, 519-523.

[5] H. Towbin, T. Staehelin, J. Gordon, *PNAS* **1979,** *76* (9), 4350-4354.

[6] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, A. Cardona, *Nat. Methods* **2012**, *9* (7), 676-682.

[7] M. F. Ebbesen, C. Gerke, P. Hartwig, L. Hartmann, *Polym. Chem.* **2016***, 7* (46)*,* 7086-7093.

[8] C. Schmuck, V. Bickert, M. Merschky, L.Geiger, D. Rupprecht, J. Dudaczek, P. Wich, T. Rehm, U. Machon, *Eur. J. Org. Chem*. **2008**, *2*, 324-329.