Electronic Supplementary Material (ESI) for Chemical Communications
This journal is © The Royal Society of Chemistry 2013

Electronic Supplementary Information for:

One-pot Squaric Acid Diester Mediated Aqueous Protein Conjugation

*Frederik Wurm, Tobias Steinbach, and Harm-Anton Klok**

Ecole Polytechnique Fédérale de Lausanne (EPFL), Institut des Matériaux and Institut des Sciences et Ingénierie Chimiques, Laboratoire des Polymères, Bâtiment MXD, Station 12, CH-1015 Lausanne, Switzerland.

Materials. All chemicals and solvents were purchased from commercial suppliers and used without further purification. 4'-(Aminomethyl)fluorescein was purchased from Invitrogen. Deuterated solvents for NMR spectroscopy were purchased from Armar Chemicals (Döttigen, Switzerland) and Sigma-Aldrich. Dry solvents were obtained from a solvent purification system (PureSolv™, Innovative Technology, Inc., Newburyport, Massachusetts, USA). CelluSep H1 regenerated cellulose tubular was used as dialysis membrane. Carboxylic acid terminated poly(ethylene glycol) was synthesized according to literature.[1]

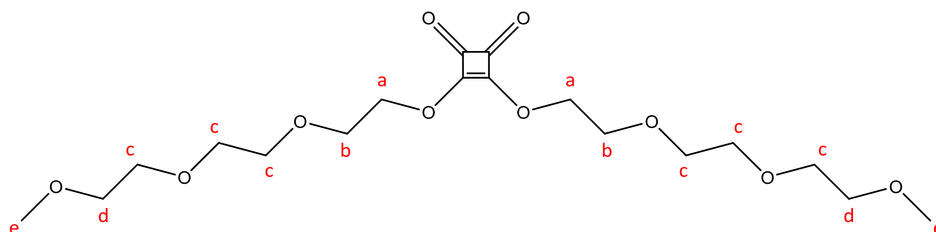
Instrumentation & Methods. ESI-MS analysis was performed on a TSQ 7000 (Thermo Fischer) single quadrupole mass spectrometer equipped with an electrospray (ES) ionization interface. Data were acquired using the ICIS software running on a Digital Unix workstation. SDS-PAGE was carried out with 4-8% Tris-HCl gels (Biorad, 0.75 mm, 10 well). Gel permeation chromatography (GPC) measurements were performed in a 9 : 1 mixture of phosphate buffer (0.1 M, pH = 6.5): methanol at a flow rate of 0.5 $mL \cdot min^{-1}$ on PSS protema columns on a Viscotek TDA 300 with triple detection at 25 °C equipped with a MetaChem

degasser, a Visotek VE 1121 GPC solvent pump and a VE 5200 GPC autosampler. Matrix-Assisted-Laser-Desorption/Ionization-Time-of-Flight mass spectrometry was conducted with a Shimadzu Axima CFR plus™ MALDI-ToF mass spectrometer equipped with a nitrogen laser at a wavelength of 337 nm and a pulse length of 3 ns. External calibration was carried out with a mixture of seven peptides (CHCA, CHCA dimer, reserpine, angiotensin 2, substance P, Glu fib, ACTH frag 18 & 39, melittin, chain B Ins). Data processing was performed using the Kompact v2.4.3 software. Samples were dissolved in a H_2O/acetonitrile/TFA (50 : 50 : 0.005) solution by mixing with sinapinic acid as the matrix (10 mg/mL), protein (1 mg/mL). A volume of ca. 1 μL of the sample solution was deposited on the MALDI target and allowed to dry at room temperature prior to the measurement. Grazing angle attenuated total reflectance (ATR) Fourier transform infrared spectroscopy was performed on a nitrogen purged Nicolet 6700 FT-IR spectrometer equipped with a VariGATR grazing angle ATR accessory (Harrick Scientific Products Inc., NY) fixing the incident angle at 65°. ^{1}H NMR and ^{13}C NMR spectra were recorded on a 400 MHz (^{1}H) or 100 MHz (^{13}C) Bruker Avance-400™ spectrometer. All spectra were recorded at room temperature. The chemical shift (δ) is given in ppm and referenced to the residual proton or carbon-13 signal of the appropriate deuterated solvent. Elemental analyses were carried out using an Elementar Vario EL 2 instrument.

Experimental

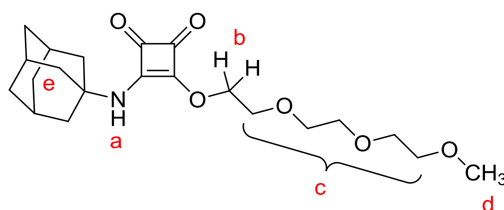
Di(triethylene glycol monomethyl ether) squaric acid ester (**1**). Triethylene glycol monomethyl ether (13.8 g, 84 mmol) was dissolved in 60 mL of dry benzene followed by addition of squaric acid (4.6 g, 40 mmol). Concentrated sulfuric acid (10 drops) was added as the catalyst. The vigorously stirred suspension was heated to reflux (oil bath temperature: 90 °C) in a Soxhlet apparatus. The extraction thimble of the Soxhlet apparatus was filled with 3

cm of oven-dry molecular sieve (4 Å) covered by 1 cm of calcium hydride. The reaction mixture was kept under nitrogen and heated for 12 hours. The solvent was removed in vacuo to yield the final product as a viscous yellow oil (15.3 g, 38 mmol, 95%). The product was used without further purification (R_f: 0.55 (1 : 2 : 2 ethyl acetate : hexane: acetone; R_f: 0.9 (8 : 2 DCM : MeOH).



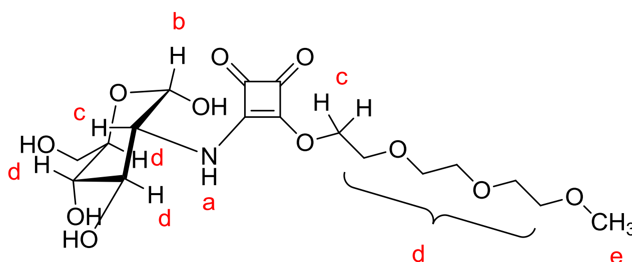
^{1}H-NMR (400 MHz, $CDCl_3$): δ [ppm] = 4.82 (m, 4 H, **a**), 3.84 (m, 4 H, **b**), 3.66 (m, 12 H, **c**), 3.54 (m, 4 H, **d**), 3.38 (s, 6 H, **e**). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ [ppm] = 189.04 (s, $\underline{C}$=O), 183.93 (s, $\underline{C}$=$\underline{C}$), ~ 71 (m, O-$\underline{C}H_2$-$\underline{C}H_2$-), ~ 59 (m, O-$\underline{C}H_3$). **FTIR**: $\tilde{\nu}$ [cm^{-1}] = 2873 (CH_{Alk}, b), 1812 ($\nu_{s(C=O)}$, s), 1733 ($\nu_{as(C=O)}$, s), 1601 ($\nu_{(C=C)}$, s), 1464, 1415, 1333, 1295, 1248, 1197, 1097 ($\nu_{(C\text{-}O\text{-}C)}$, s), 1026, 930, 850, 809, 674. **ESI-mass:** 407 (MH^+), 429 (MNa^+)

*3-Adamantylamino-4-(triethylene glycol methyl ether)cyclobut-3-ene-1,2-dione (**2**).* **1** (150 mg, 0.37 mmol) was dissolved in 10 mL phosphate buffer (pH 7.4, 50 mM) and adamantylamine (56 mg, 0.37 mmol) was added. The reaction can be followed with the naked eye: the turbid solution was vigorously stirred at room temperature over a period of 2 h until it became clear. The formation of **2** can also be monitored by thin layer chromatography (TLC) (R_f of **1**: 0.55, R_f of **2**: 0.75 (1 : 2 : 2 ethyl acetate : hexane : acetone)). The final reaction solution can be directly used for the second amidation step. For characterization of the intermediate, the product was extracted with chloroform to isolate (desalt) **2** in 99% yield (144 mg) as a colorless, viscous material.



^{1}H-NMR (400 MHz, 335K, $CDCl_3$): δ [ppm] = 5.65 (s, 1H, a), 4.91 (1, 2H, b), 3.87 (m, 2H, **c**), 3.68 (m, 6 H, **c**), 3.56 (m, 2H, **c**), 3.40 (s, 3H, d), 2.19-1.48 (m, 15H, e). **^{13}C-NMR** (100 MHz, 298K, $CDCl_3$): δ [ppm] = 190, 184 ($\underline{C}$=O), 177, 171 ($\underline{C}$=$\underline{C}$), 72-69 (m, O-$\underline{C}H_2$-$\underline{C}H_2$-), 59 (NH-$\underline{C}$-, and O-$\underline{C}H_3$), 43, 36, 30 (m, adamantyl). **ESI-mass:** 394 (MH^+), 416 (MNa^+), 787 ($2MH^+$), 809 ($2MNa^+$). **Elemental analysis:** $C_{21}H_{31}O_6N$ (theor: 64.1%C, 7.94%H, 3.56%N, 24.40%O; found: 63.8%C, 8.0%H, 3.7%N, 25.0%O).

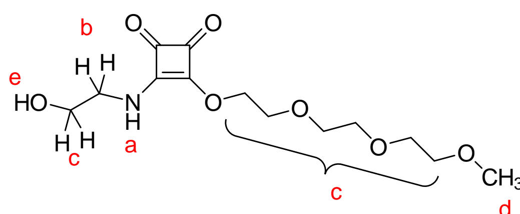
*3-Glucosamino-4-(triethylene glycol methyl ether)cyclobut-3-ene-1,2-dione (**3**).* **1** (100 mg, 0.25 mmol) was dissolved in 10 mL phosphate buffer (pH7.4, 50 mM) and glucosamine (0.25 mmol (neutralized from 54 mg (0.25 mmol) glucosamine hydrochloride with 14 mg (0.026 mmol) sodium methoxide) was added. The solution was vigorously stirred at room temperature over a period of 90 min to generate **3** in full conversion (TLC control, R_f of **3**: 0.37 (8 : 2 DCM: MeOH)). This solution can be directly used for the second amidation step. For characterization of the intermediate, the solution was freeze-dried and extracted with acetonitrile. Afterwards the solvent was removed *in vacuo* and the crude product was washed with diethyl ether to remove the released triethylene glycol monomethyl ether and isolate **3** in 85% yield (87 mg) as a slightly yellow viscous material.



Electronic Supplementary Material (ESI) for Chemical Communications
This journal is © The Royal Society of Chemistry 2013

^{1}H-NMR (400 MHz, 327K, CD_3CN): δ [ppm] = 7.38 (0.5H, a), 6.98 (0.5H, a), 5.23 (s, 1H, b), 4.81-4.66 (3H, c), 4.72 (s, 2H, d), 3.86-3.34 (m, 14H, d), 3.31 (s, 3H, e), 3.1-2.5 (broad, OH). **^{13}C-NMR** (100 MHz, 298K, CD_3CN): δ [ppm] = 189.6, 184 (C=O), 177.5, 173.9 (C=C), 92.3 (acetal carbon), 72.2-69.3 (m, O-CH_2-CH_2- and -CH-OH), 66.6 (-CH_2-OH), 61.1 (-CH-NH-), 58.4 (O-CH_3). **ESI-mass:** 423 (MH^+), 837 ($2MH^+$). **Elemental analysis:** $C_{17}H_{27}O_{11}N$ (theor: 48.45%C, 6.46%H, 4.32%N, 41.76%O; found: 49.2%C, 7.1%H, 3.9%N, 42.1%O).

*3-Ethanolamino-4-(triethylene glycol methyl ether)cyclobut-3-ene-1,2-dione (**4**).* **1** (3 g, 7.4 mmol) was dissolved in 100 mL acetonitrile and freshly distilled ethanolamine (452 mg, 7.4 mmol) in 50 mL acetonitril was added dropwise over a period of 1 h. The mixture was stirred over a period of 2h until TLC control proved the complete consumption of the starting material (R_f of **4** 0.6 (dichloromethane: methanol 8:1). The solvent was removed *in vacuo*, the crude mixture was dissolved in dichloromethane and washed with brine to remove most of the triethylene glycol monomethyl ether. For most following reactions the product is pure enough after simple washing (yields typically 90-95%). A sample was purified via column chromatography (dichloromethane: methanol 8:1) to yield the pure compound as a slightly yellow viscous oil (yield: 1.9 g, 85%).



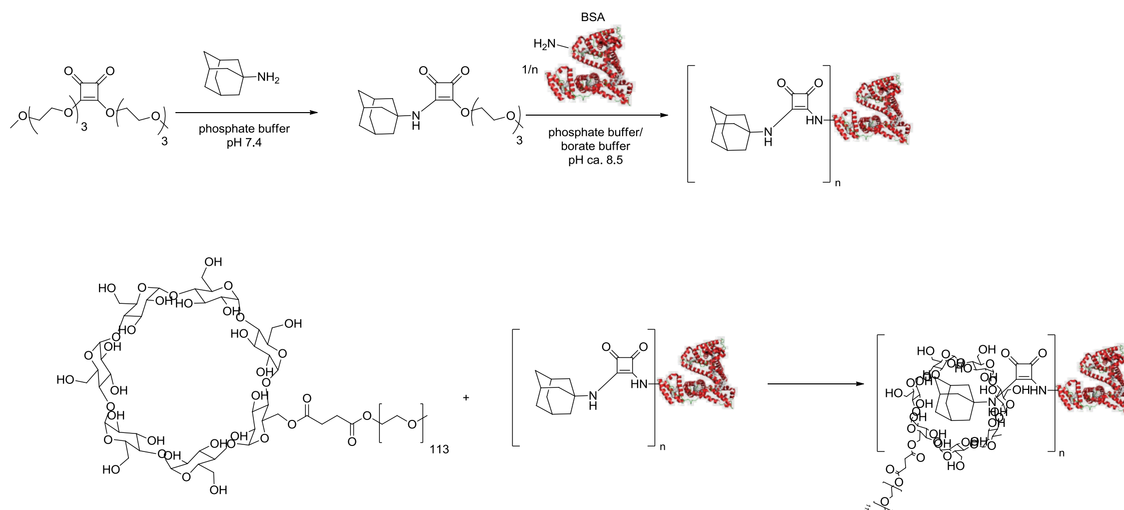
^{1}H-NMR (300 MHz, 298K, $CDCl_3$): δ [ppm] = 7.45 (0.5H, a), 7.14 (0.5H, a), 4.85-4.75 (2H, b), 3.87-3.5 (m, H, c), 3.35 (s, 3H, d), 3.29 (2H, e). **ESI-mass:** 304.4 (MH^+). **Elemental analysis:** $C_{13}H_{21}O_7N$ (theor: 51.46%C, 6.98%H, 4.62%N, 36.94%O; found 51.0%C, 7.1%H, 4.4%N, 37.0%O).

Modification of BSA with **2**. The protein-reactive ester amide was synthesized as described above. From the reaction mixture 0.5 mL were removed and added to a solution of 30 mg BSA in borate buffer (equivalent to ca 27.5 equivalents of **2**). The mixture was allowed to stir over night at room temperature, dialyzed against water, and subsequently freeze-dried to yield the adamantyl-BSA conjugate (31 mg) as a colorless powder. The conjugate was analyzed via MALDI-ToF mass: 72400 g/mol (ca. 25 adamantyl-units attached).

Synthesis of polyethylene glycol-modified β-cyclodextrin (PEG-β-CD). Carboxylic acid terminated polyethylene glycol (500 mg, 0.1 mmol, M_n = 5100 g/mol, M_w/M_n = 1.05) and β-cyclodextrin (113.5 mg, 0.1 mmol) were dissolved in 250 mL dry dimethylformamide. The solution was cooled with an ice bath and dicyclohexylcarbodiimide (23 mg, 0.11 mmol) and dimethylaminopyridine (7 mg, 0.06 mmol) were added. The mixture was stirred for a period of 16 h, then the solvent was removed in vacuo and the crude reaction mixture was dissolved in 5 mL water, filtered, and dialyzed extensively against water. The final product was isolated after freeze-drying as a colorless powder (500 mg, 79%). GPC (DMF): M_n = 6,300 g/mol, M_w/M_n= 1.15 (*note:* also a small fraction of oligomers was detected with 2 or 3 PEG-chains attached to CD, see below). **^{1}H-NMR** (400 MHz, 327K, CD_3CN): δ [ppm] = 5.5 (m, 7H, CD-acetals), 4.86 (m, OH), 4.2-4.1 (5-6 H, -CH_2-OCO-), 3.74-3.30 (CD & PEG backbone), 3.27 (3H, O-CH_3) (See also Figure S11).

Non-covalent PEGylation of BSA. Adamantyl-modified BSA (1 mg, $1.38x10^{-5}$ mmol) and BSA were both separately dissolved in phosphate-buffer/ methanol (9 : 1, solvent for GPC) and to both solutions PEG-*β*-CD (2.6 mg, $4.13x10^{-4}$ mmol) was added. Both solutions were injected to the GPC and the elugrams were compared. Only the adamantyl-modified BSA proved successful conjugation (peak at ca. 26.2 mL).

This journal is © The Royal Society of Chemistry 2013



Scheme S1: Protocol for the synthesis of adamantyl-modified BSA and subsequent conjugation of PEG-β-CD in a non-covalent manner.

One-pot glycosylation of BSA. The ester-amide **3** was prepared as described above. From the reaction mixture 0.5 mL were removed and added to a solution of 55, 41 and 33 mg BSA (equivalent to 15, 20, and 25-fold excess of **3**, respectively) in borate buffer (0.1 M, pH 9.1). The final pH value was in the range between 8.5-9.0. The mixture was allowed to stir overnight. Then it was dialyzed against water and freeze dried to yield the glycoconjugates in 58, 44 and 38 mg yield as colorless powders.

Hydrolysis experiments. All experiments were conducted in NMR tubes (diameter 5 mm). For each spectrum 4 scans were run. Locking, matching and tuning of the NMR spectrometer was done before the actual experiment with a NMR tube containing the final reaction mixture which was prepared several hours in advance and consisting of approximately the same concentration of reactants as the actual experiment.

This journal is © The Royal Society of Chemistry 2013

*Hydrolysis of **1** in D_2O.* A NMR tube was charged with **1** (20 mg, 50 µmol). Immediately after addition of deuterium oxide (D_2O, 0.7 mL) the NMR experiment was started.

*Hydrolysis of **1** in borate buffer.* A NMR tube was charged with **1** (20 mg, 50 µmol). Immediately after addition of borate buffer (0.01 M in D_2O, 0.7 mL, pD 9.5) the NMR experiment was started.

*Hydrolysis of **3** in borate buffer.* A NMR tube was charged with purified **5** (15 mg, 43 µmol) and borate buffer (0.01 M in D_2O, 0.7 mL, pD 9.5) was added. The experiment was started immediately and experiments were run over a period of 14 days.

Additional Figures:

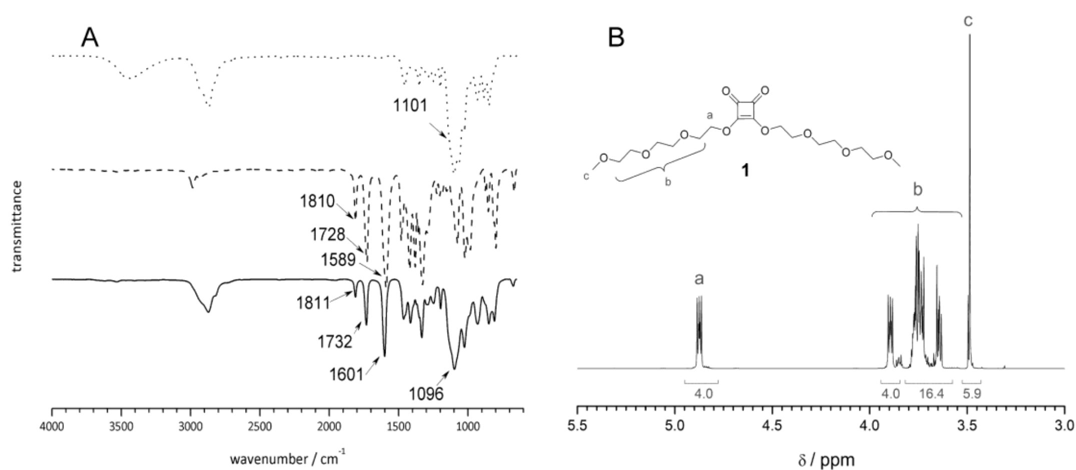


Figure S1 : (A) FTIR spectra of triethylene glycol monomethylether (…), squaric acid diethyl ester (- - -), squaric acid di(triethylene glycol monomethyl ether)ester (**1**, **-**); (B): ^{1}H NMR spectrum of **1** in $CDCl_3$

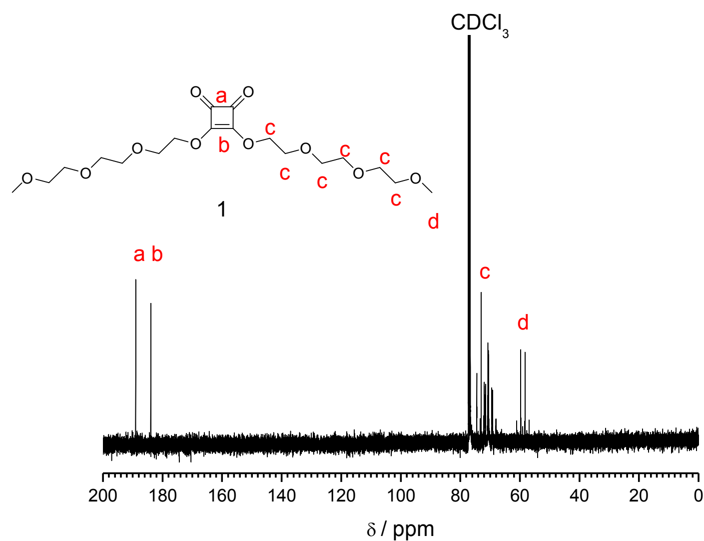


Figure S2: ^{13}C NMR spectrum of di(triethylene glycol monomethyl ether) squaric acid ester (**1**) ($CDCl_3$, 100 MHz).

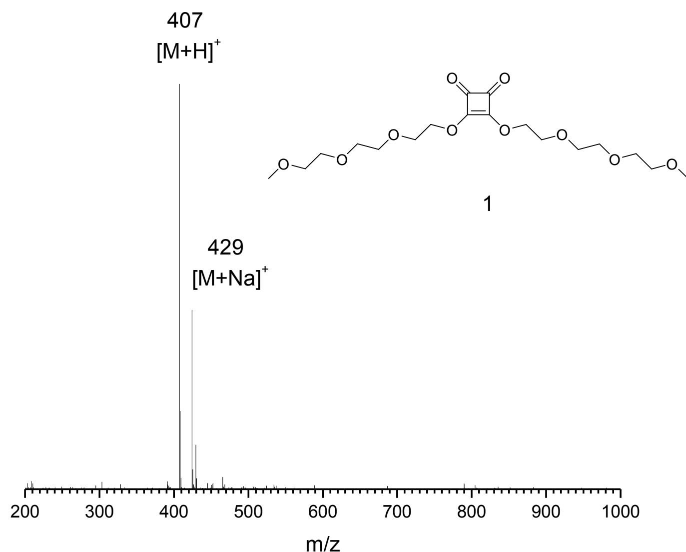


Figure S3: ESI mass spectrum of di(triethylene glycol monomethyl ether) squaric acid ester (**1**).

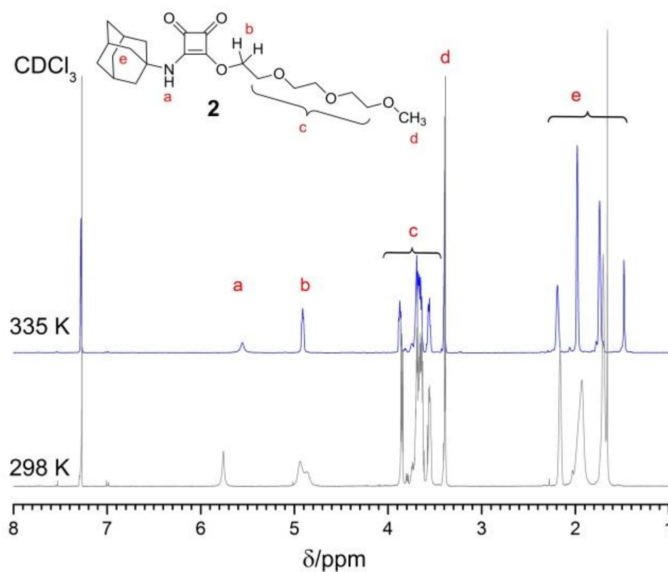


Figure S4: ^{1}H NMR spectra of **2** in $CDCl_3$ at 298 K (bottom) and 335 K (top). The signal due to the methylene groups of **2** that is labeled as *b* splits in two signals due to the presence of anti- and syn-squaramide rotamers at room temperature.

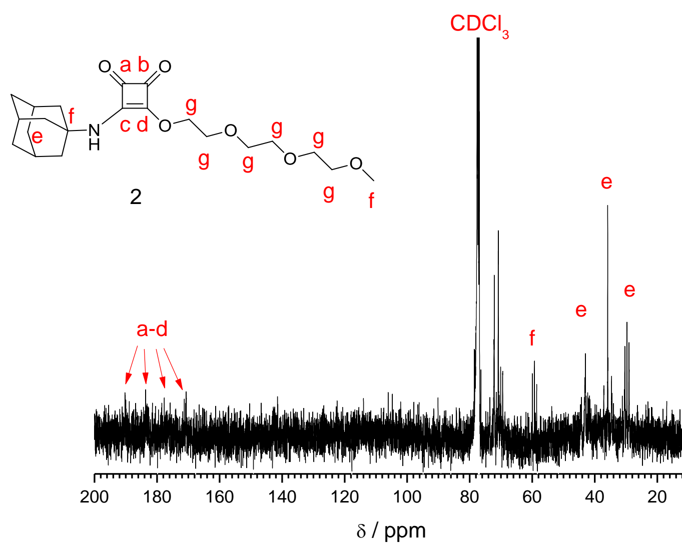


Figure S5: ^{13}C NMR of 3-adamantylamino-4-(triethylene glycol methyl ether)cyclobut-3-ene-1,2-dione (**2**) ($CDCl_3$, 100 MHz).

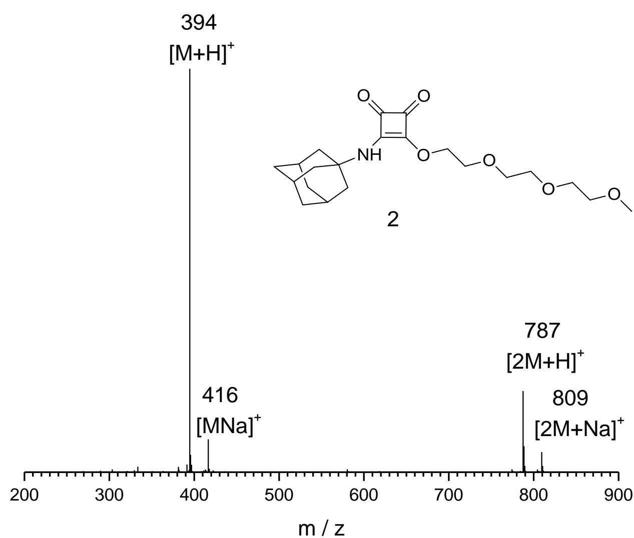


Figure S6: ESI mass spectrum of 3-adamantylamino-4-(triethylene glycol methyl ether)cyclobut-3-ene-1,2-dione (**2**).

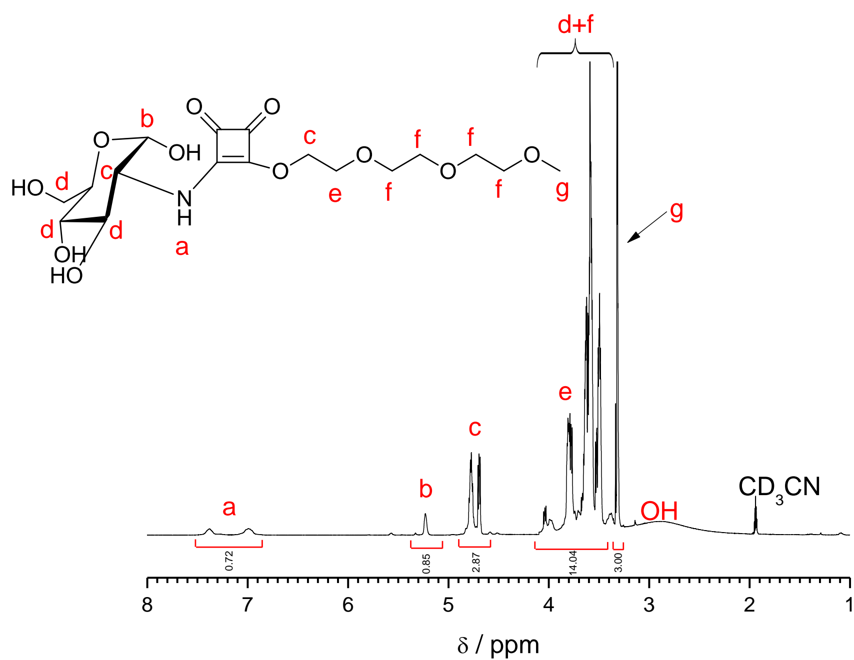


Figure S7: ^{1}H NMR spectrum (CD_3CN, 400 MHz) of 3-glucosamino-4-(triethylene glycol methyl ether)cyclobut-3-ene-1,2-dione (**3**).

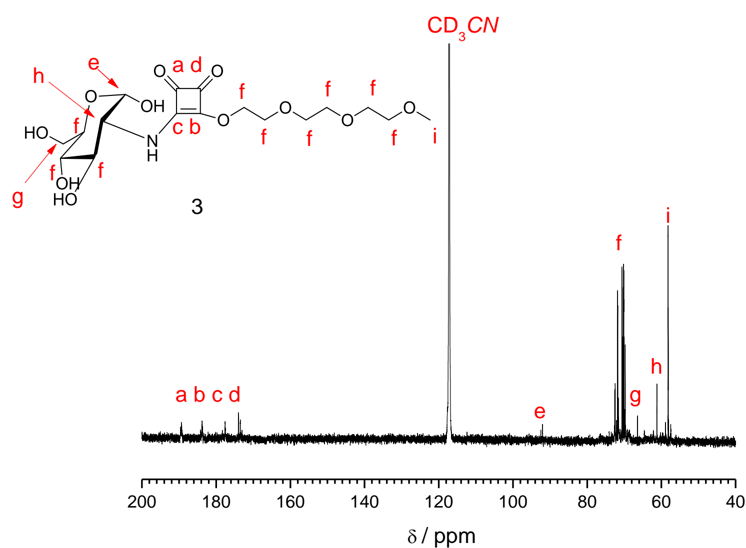


Figure S8: ^{13}C NMR spectrum of 3-glucosamino-4-(triethylene glycol methyl ether)cyclobut-3-ene-1,2-dione (**3**) (CD_3CN, 100 MHz).

This journal is © The Royal Society of Chemistry 2013

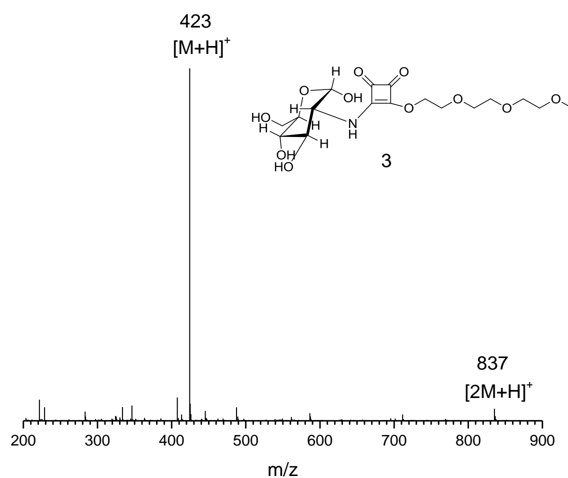


Figure S9: ESI mass spectrum of 3-glucosamino-4-(triethylene glycol methyl ether)cyclobut-3-ene-1,2-dione (**3**).

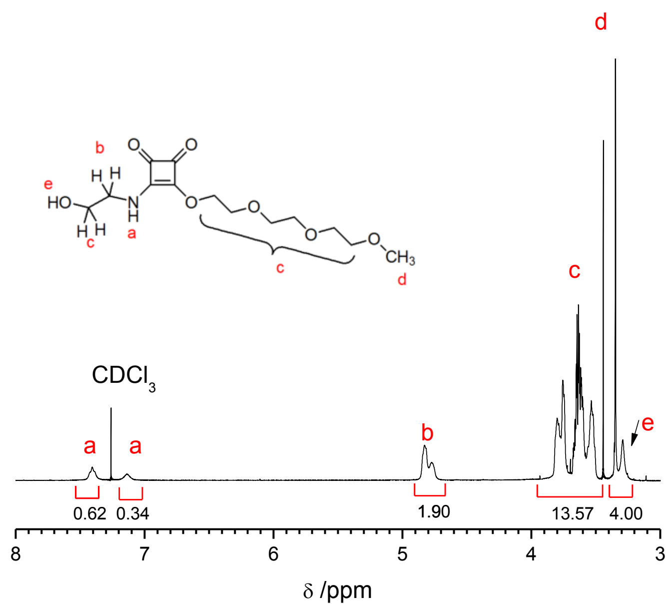


Figure S10: ^{1}H NMR spectrum (CD_3Cl, 400 MHz) of 3-ethanolamino-4-(triethylene glycol methyl ether)cyclobut-3-ene-1,2-dione (**4**).

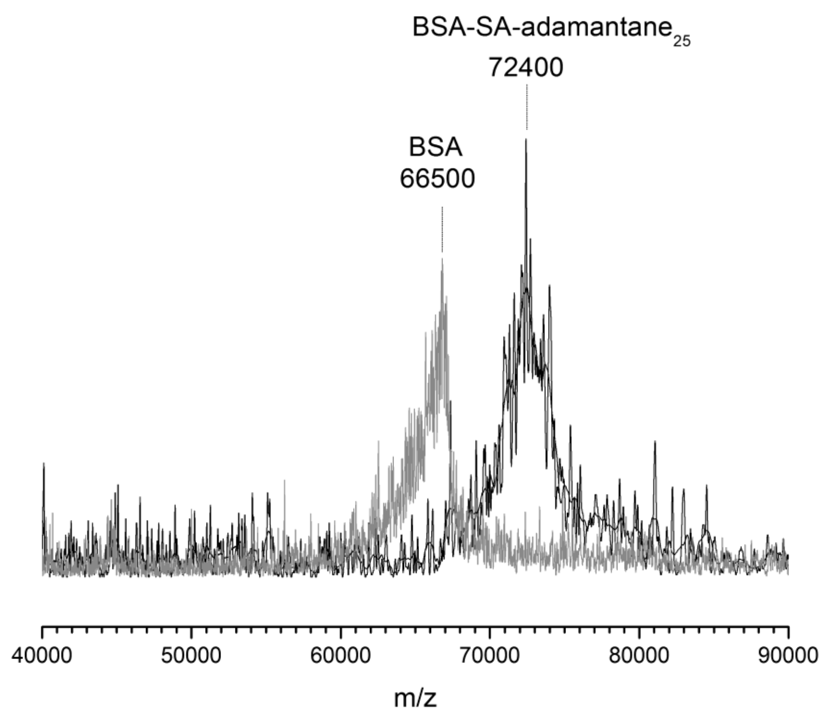


Figure S11: MALDI-ToF mass spectra of unmodified as well as adamantyl-modified BSA containing on average 25 units adamantyl groups as estimated from the difference in mass between the peak maxima in the two spectra (measured in linear mode with sinapic acid as matrix).

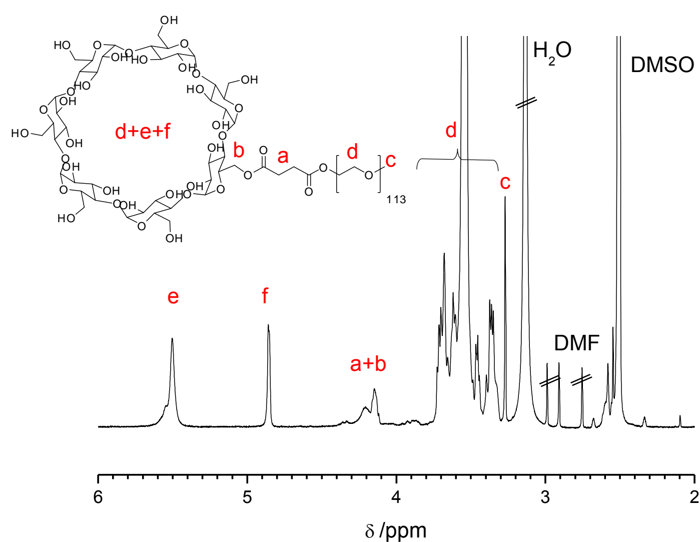


Figure S12: ^{1}H NMR spectrum of PEG-β-CD (400 MHz, DMSO-d_6, 335K).

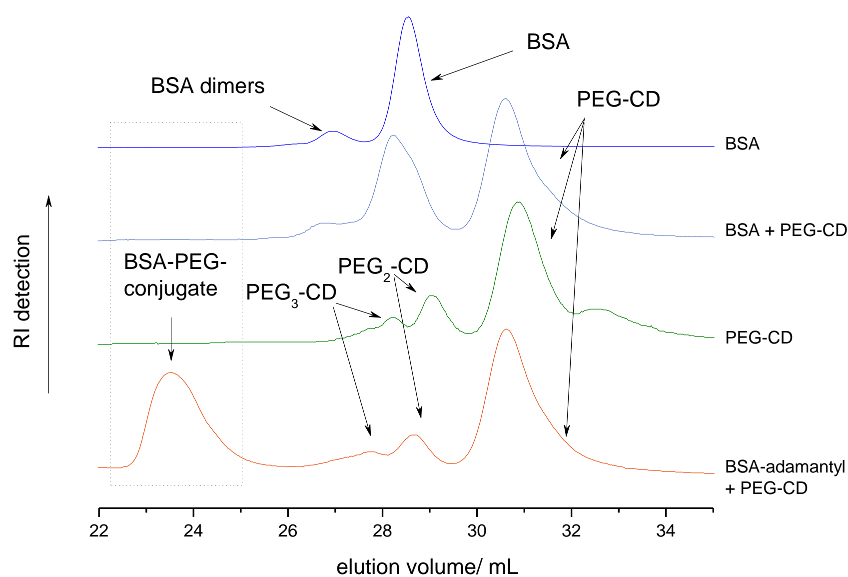


Figure S13: GPC elugrams of BSA (top), BSA-PEG-conjugate (middle), mixture of BSA and PEG-β-CD (bottom).

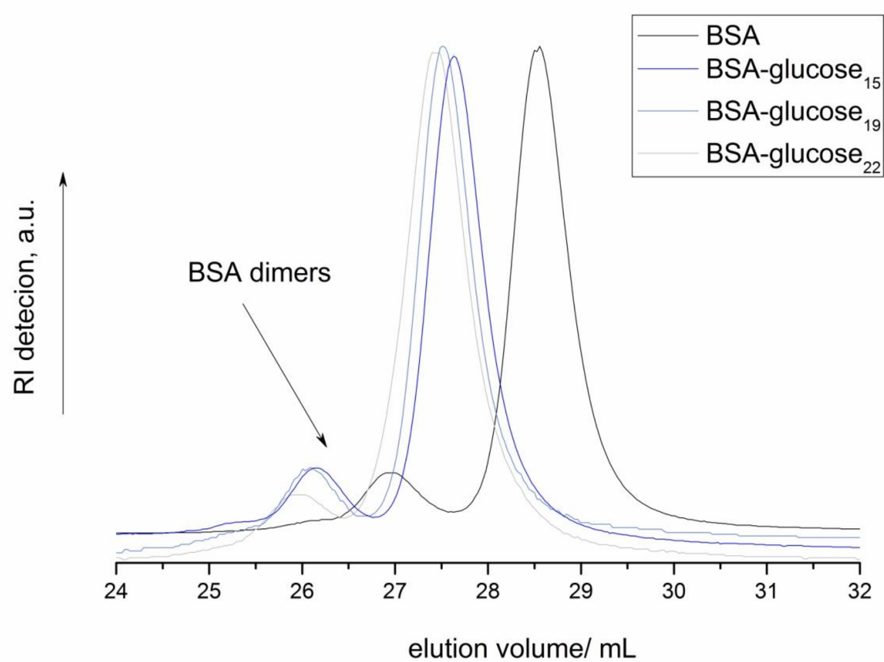


Figure S14: GPC elugrams (RI detection) of BSA and glycosylated BSA.

This journal is © The Royal Society of Chemistry 2013

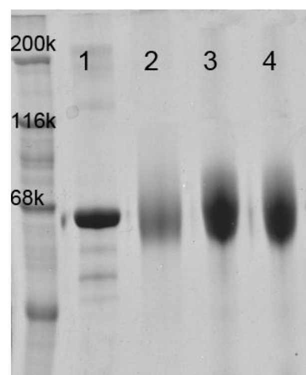


Figure S15: SDS-PAGE of glycosylated BSA; Lane 1: BSA; lane 2: BSA-glucose$_{15}$; lane 3: BSA-glucose$_{19}$; lane 4: BSA-glucose$_{22}$ (subscript = number of glucose-units determined via MALDI ToF mass spectrometry).

References

1. Kynclova, E.; Elsner, E.; Kopf, A.; Hawa, G.; Schalkhammer, T.; Pittner, F. *J. Mol. Recog.* **1996,** 9, 644-651.