Macrocyclization of Enzyme-based Supramolecular Polymers

Maartje M. C. Bastings, Tom F. A. de Greef, Joost L. J. van Dongen, Maarten Merkx and E. W. Meijer

Electronic Supporting Information

Synthesis and Characterization of AB1 and AB2. Before oxidation of the serine residue at the N-terminus of the S-protein, the commercially obtained RNase S first was separated into the S-peptide and S-protein using RP-HPLC with a gradient of acetonitrile in water both containing 0.1% TFA on a C4 semi-prep column. Products were characterized by analytical LC-MS (Figure S1).

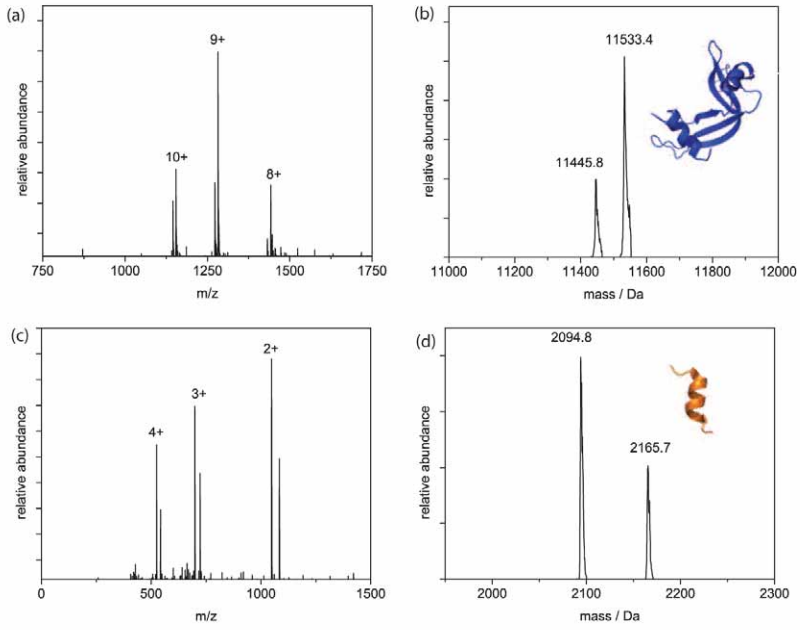
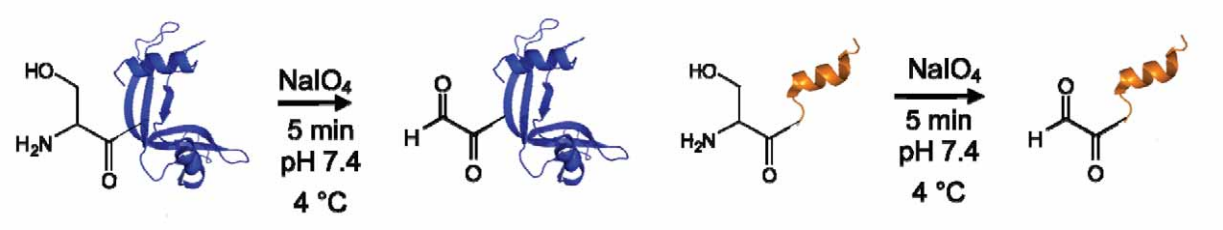


Figure S1 (a) m/z and deconvoluted mass spectra **(b)** of the S-protein (MW_{calc} = 11534 Da and $MW_{calc\ -ser}$= 11446.9 Da) and m/z **(c)** and deconvoluted mass spectra **(d)** of the S-peptide (MW_{calc} = 2095 Da and $MW_{calc\ +\ ser}$ = 2166 Da) after separation by RP-HPLC. The heterogeneous digestion by subtilisin is clearly detectable.

The ketone functionality was introduced in the protein and peptide by oxidation of the N-terminal serine residue using 1.2 equivalents $NaIO_4$ in a pH 7.4 PBS buffer and a reaction time of 5 minutes at 4 °C (Scheme S1). $NaIO_4$ was successfully separated from the peptide and protein by RP-HPLC and the oxidized protein and peptide were obtained as the quantitative product. Products were characterized by analytical LC-MS (Figure S2).



Scheme S1 Oxidation of the N-terminal serine residue by $NaIO_4$.

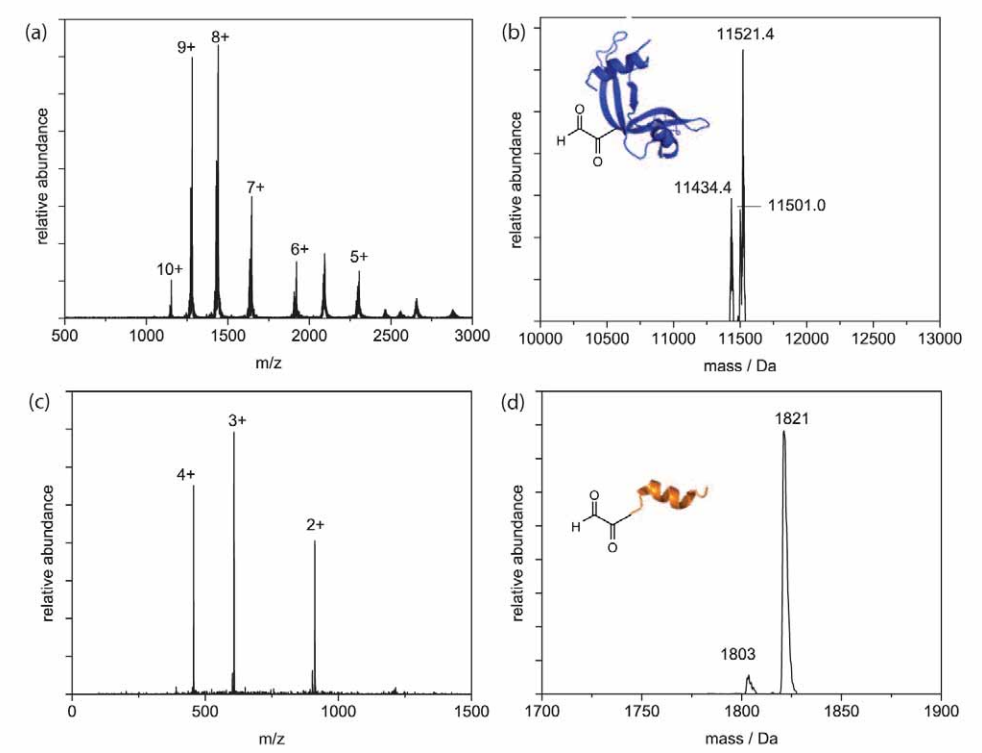
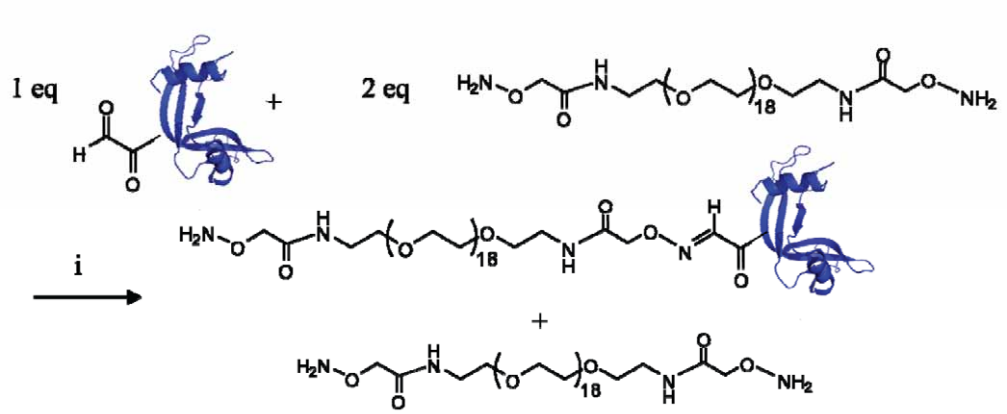


Figure S2 A Q-TOF MS of oxidized S-protein, with the hydrated product as main species ($MW_{calc+OH}$ = 11521 Da). **B** Q-TOF MS of the oxidized S-peptide, mainly present as hydrated species (MWcalc + OH = 1821 Da, MWcalc = 1803 Da)

The protein and peptide were then ligated to amino-oxy functionalized PEG-linkers of 18 and 5 PEG repeats in two steps. First the S-protein was ligated to the linker to form the mono-substituted compound (overnight at 37 °C, Scheme S2). Due to the excess of aminooxy groups and steric hindrance when one protein is coupled, only the monoligated product will be present, together with unreacted aminooxy linker (Figure S3). Subsequently, 4 equivalents of the oxidized S-peptide were added to react with the remaining aminooxy end groups (4 hours at 37 °C, Scheme S3). Ligation reactions were performed in 0.1 M anilinium acetate buffer at pH 4.5. Both the desired AB-monomer product as well as the AA di-peptide monomer are formed (Figure S5), these reaction products are easily separated using RP-HPLC (Figure S4) and yielded the expected AB-monomers in 40 % yield relative to the S-protein (Figure S6).



Scheme S2 Reaction scheme of the first step in the AB-monomer synthesis: oxime ligation of oxidized S-protein with an excess of diaminooxy linker. (i = 0.1 M anilinium acetate buffer pH 4.5, 37 °C, o/n)

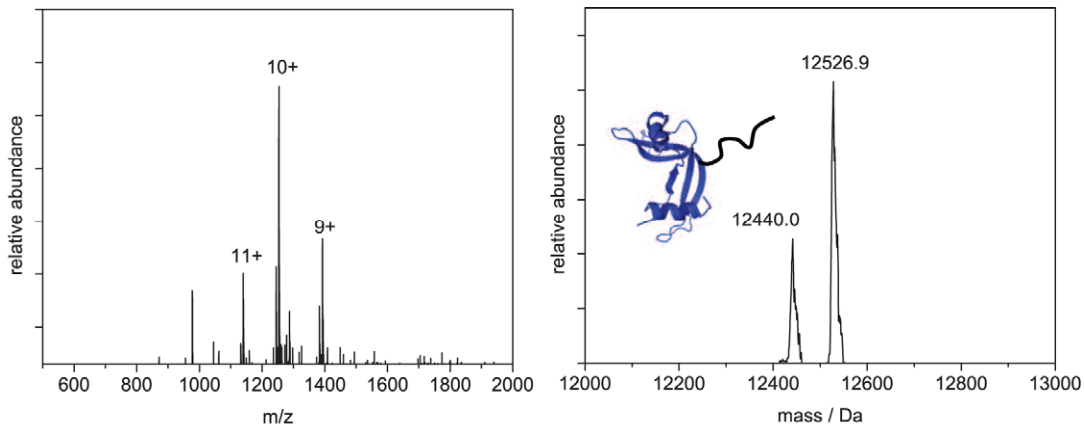
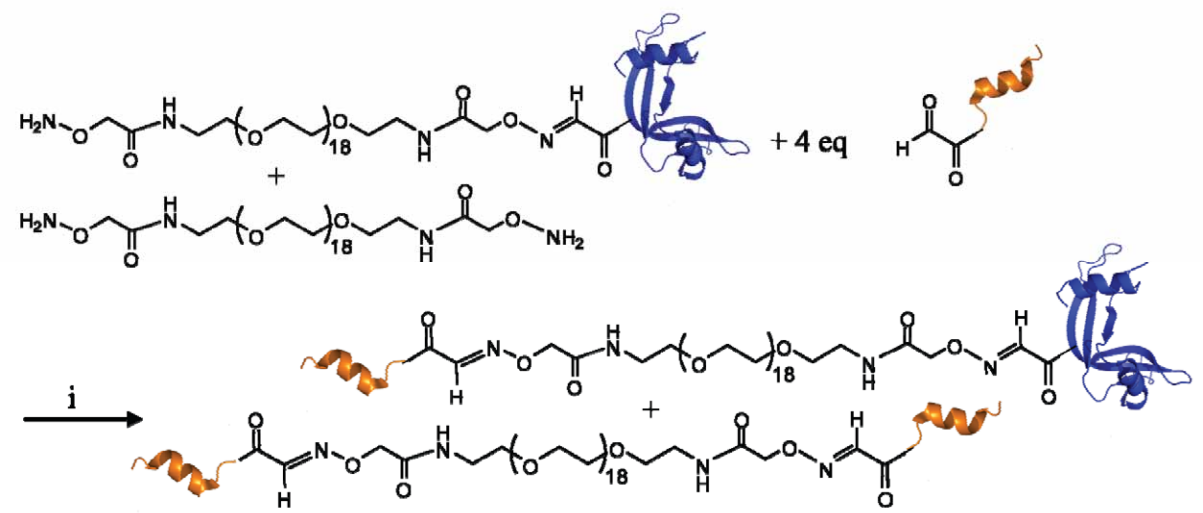


Figure S3 m/z and deconvoluted mass spectra of the mono ligated S-protein, (MW_{calc} = 12528 Da, $MW_{calc\ -\ ser}$ = 12441 Da)



Scheme S3 Reaction scheme of the second step in the AB-monomer synthesis: addition of 4 eq of oxidized S-peptide and oxime ligation with the unreactied aminooxy groups after the first step to form the AB-monomer as well as the AA S-peptide monomer. (i = 0.1 M anilinium acetate buffer pH 4.5, 37 °C, 4 hours)

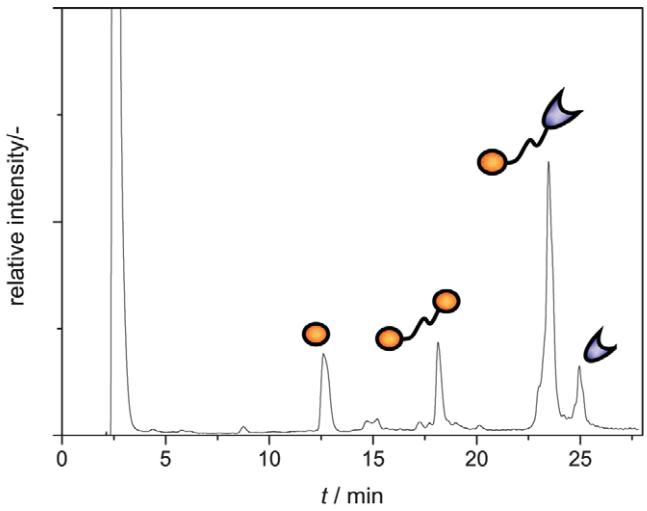


Figure S4 TIC trace of RP-HPLC purification of the AB-monomer, the first peak is from GuHCl

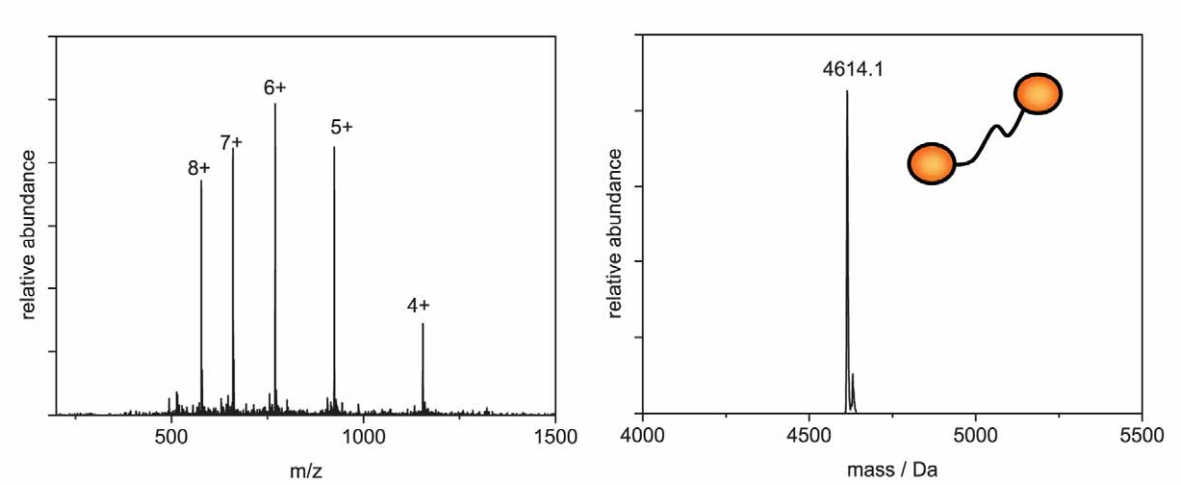


Figure S5 m/z and deconvoluted mass spectra of the S-peptide dimer (MW_{calc} = 4615 Da)

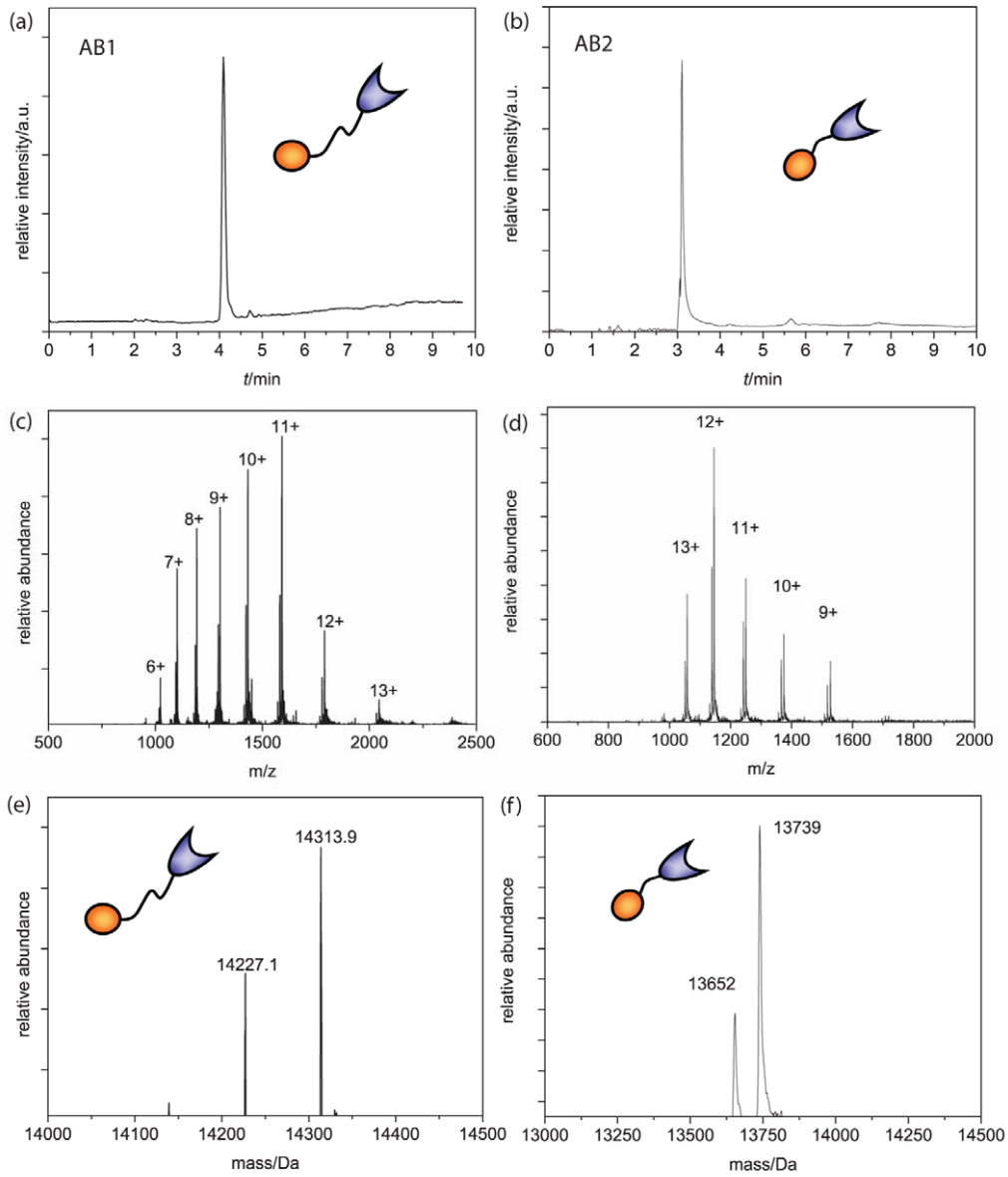


Figure S6 TIC trace of the RP-HPLC purification of **AB1 (a)** and **AB2 (b)**. Q-TOF-MS m/z spectrum of **AB1 (c)** and **AB2 (d)**, and deconvoluted mass of **AB1 (e)** and **AB2 (f)**.

Activity Assay The fluorescent 6-FAM-dArUdAdA-6TAMRA substrate was used to monitor enzymatic activity. Activity measurements at different enzyme concentrations (~nM range), verified with UV-vis, and using a fixed substrate concentration of 400 nM were recorded. Fresh dilutions from the protein stock (20 μM) were made for each measurement to prevent deviations from expected protein concentrations by sticking to the tube walls.

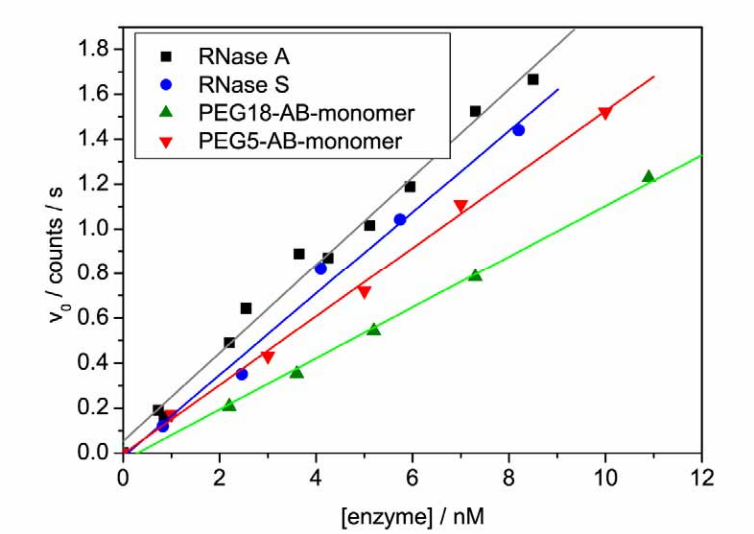


Figure S7 Plot of the activity of RNase A (black squares), RNase S (blue circles) and the synthesized AB-PEG18-monomer (green triangles) and AB-PEG5-monomer (red inversed triangles) versus the enzyme concentration. All lines are linear regression fits through the data points.

For commercial RNase A and RNase S, comparable activity results are obtained with slopes 0.1962 and 0.1816 respectively, determined by linear regression (Figure S7). Linear regression analysis of **AB2** shows a slightly decreased slope of 0.1533 and **AB1** activity shows a slope of even less magnitude, 0.1135. From these results we can conclude that **AB2** contains an activity which is ~80 and **AB1** an activity of ~60% of that of commercially available RNase A and S.

Dynamic Light Scattering The scattering intensity scales linearly with the concentration but is dependent on the size with r^6. Larger rings therefore cause significantly more scattering than monomeric rings. Disadvanteous about DLS is a large sample volume of 800 µL in the cuvette, therefore DLS measurements of **AB1** were performed only at 50 µM concentrations. From the number weighted graph we clearly detect the monomeric species to be the main product present (Figure S8). The radius of 0.8 nm thus a diameter of 1.6 nm is consistent with the size of a monomeric ring, confirming that no larger assemblies or aggregates are present.

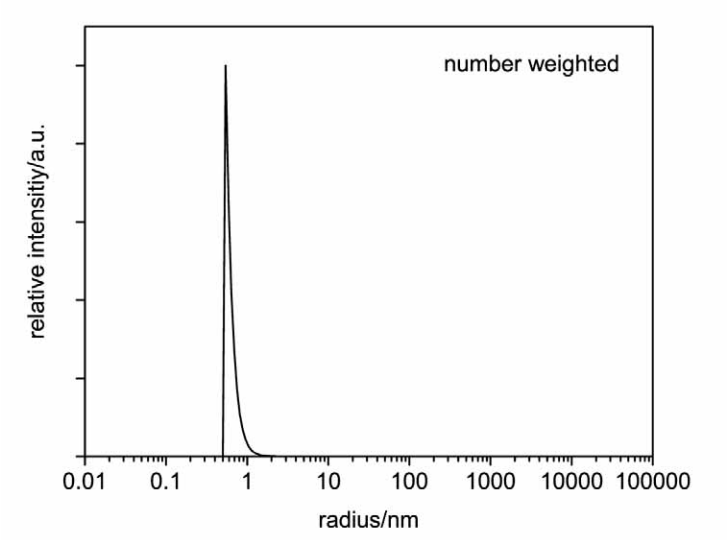


Figure S8 Number weighted DLS graph of 50 µM **AB1** in water.

Detailed modeling of ring-chain equilibrium for various EM_1.

When all cycles are strainless, the yield of cyclic monomer is always higher than any other cyclic oligomer for concentrations up to and slightly above EM_1 . This is demonstrated by simulations with the ring-chain model using various EM1 values as input, eg. 0.1 mM, 1 mM, 10 mM and 100 mM. These values are chosen to be situated around the calculated EM values for our measured systems. Figure S9 shows the modeled results for the monomeric to pentameric cyclic assemblies and the fraction of linear polymeric species, related to the analyte concentration.

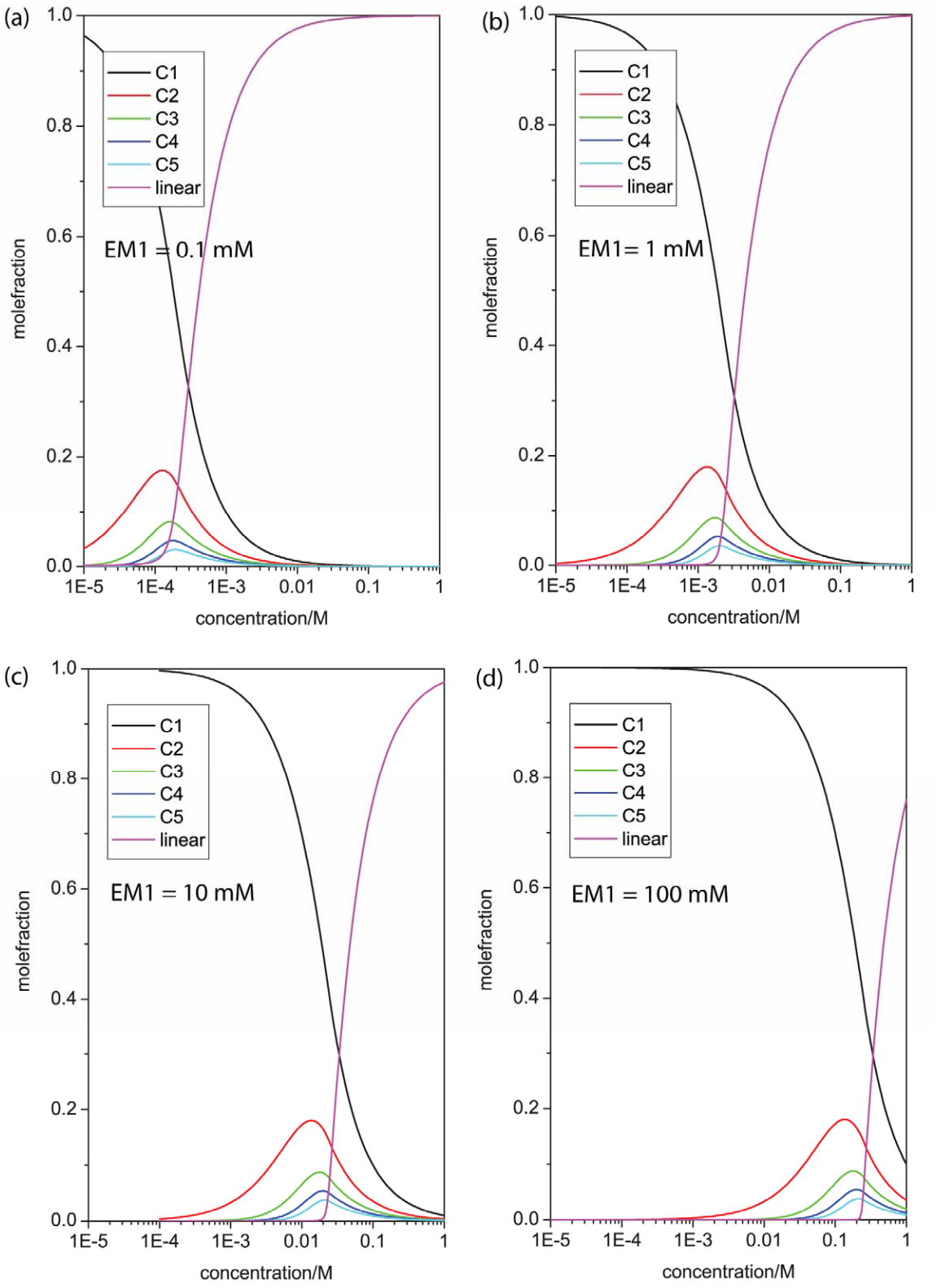


Figure S9 Modeled results for the monomeric to pentameric cyclic assemblies and the fraction of linear polymeric species, related to the analyte concentration. Used EM1 values as input, are **(a)** 0.1 mM, **(b)** 1 mM, **(c)** 10 mM and **(d)** 100 mM.

SEC-flow speed analysis

The presence of architectures assembled of up to three monomeric **AB1** moieties and up to four monomeric **AB2** moieties demonstrates the slow dissociation between the RNase S-peptide and S-protein on the time scale of the measurement. We performed control experiments using lower flow rates, 0.075 mL/min and 0.05 mL/min to confirm that k_{off} rates do not influence the distribution of architectures on the column. Figure S10 clearly shows that the composition of supramolecular architectures stays comparable when the flowspeed is changed from 0.1 mL/min, to 0.075 mL/min. Monomers are present for 85 ± 5 %, dimers for 15 ± 5 %.

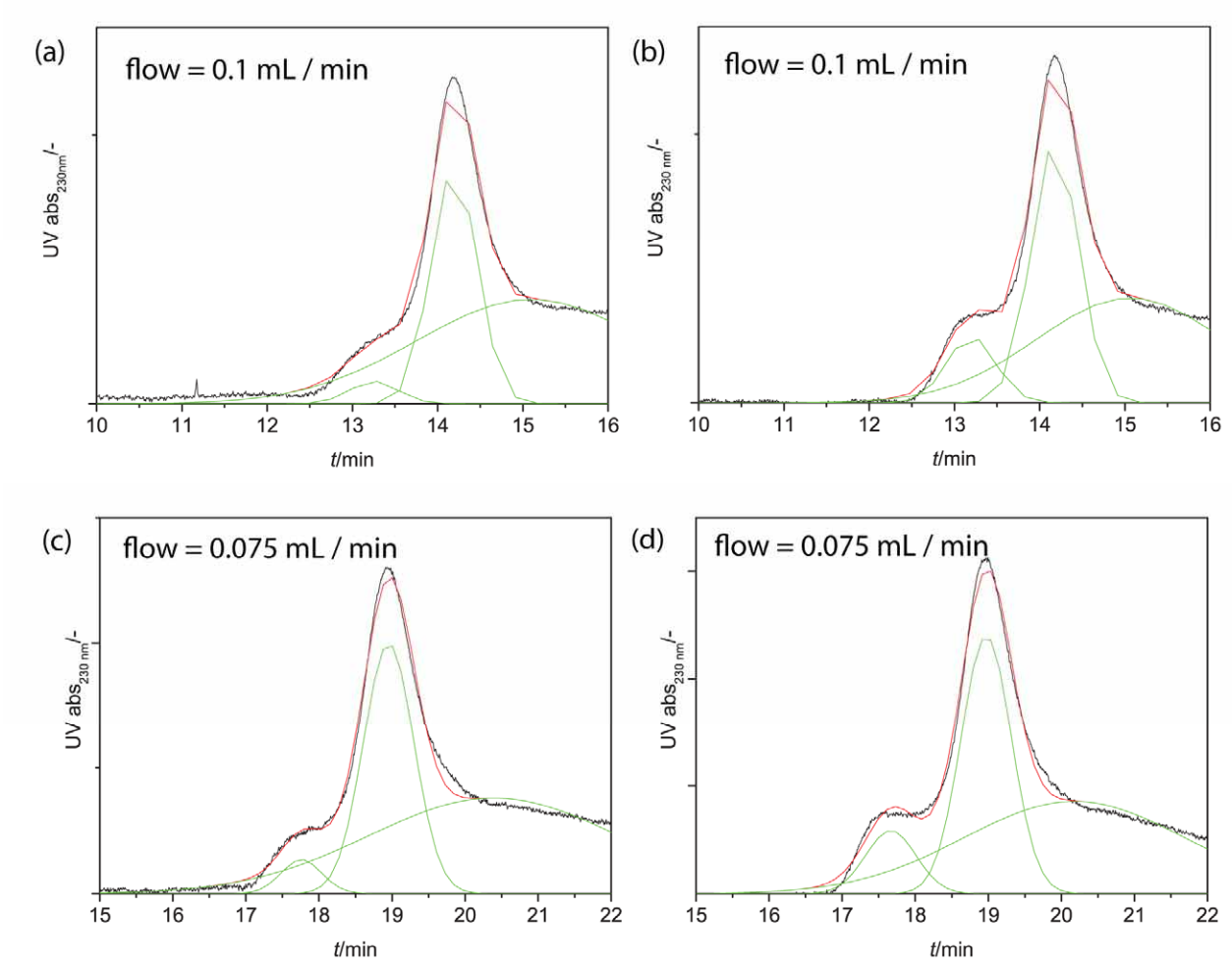


Figure S10 Composition of supramolecular architectures stays comparable at various flow-speeds **(a)** + **(b)** 0.1 mL/min, **(c)** + **(d)** 0.075 mL/min.