Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

Controlled assembly of macromolecular β-sheet fibrils

Jurgen M. Smeenk, Matthijs B. J. Otten, Jens Thies, David A. Tirrell, Hendrik G. Stunnenberg, and Jan C. M. van Hest*

Contents

Materials and methods

DNA constructs. The double-stranded oligonucleotide 1 (Figure S1a) encoding two copies of the octapeptide repeating unit $[(AG)_3EG]_2$ was designed with the following considerations: (i) avoidance of codons with a low frequency of occurrence in *E. coli*, (ii) minimization of sequence repetitiveness and (iii) non-palindromic overhangs for unidirectional multimerization. The plasmid pSK-JS2 was prepared by ligation of the 66-bp oligonucleotide 2 into the *Eco*RI and *Bam*HI site of pBluescript® II SK (-) (Stratagene). Digestion of pSK-JS2 with *Bsp*MI allowed directional insertion of multimerized oligonucleotide 1 (Figure S1b). The desired multimeric gene can subsequently be cloned into the *Bam*HI site of the pET-3b expression vector.

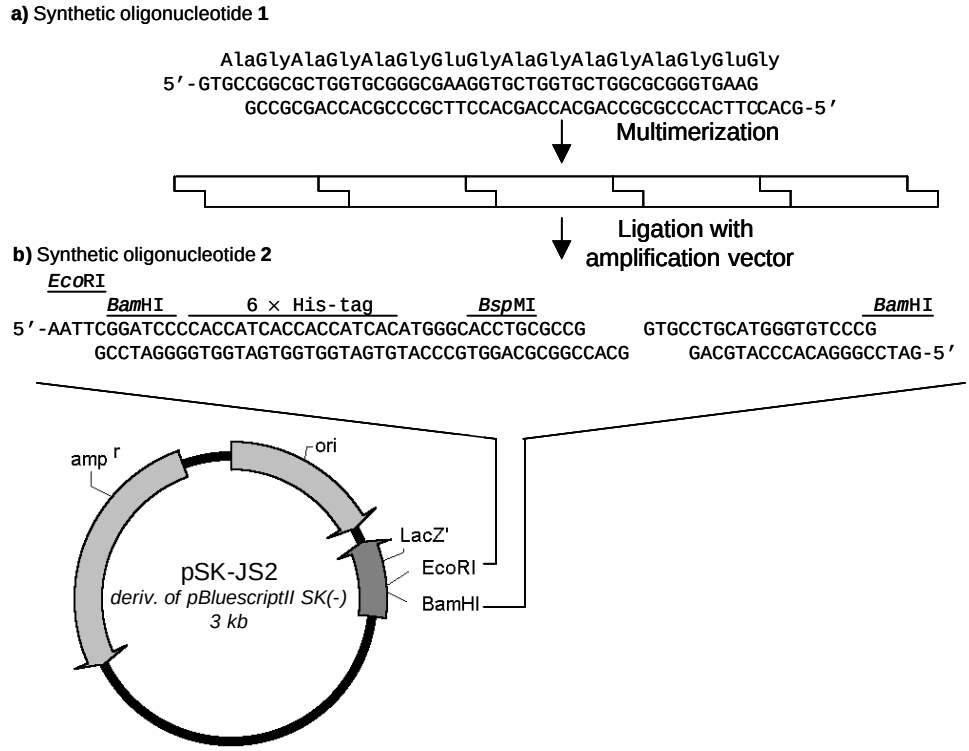


Figure S1. Cloning strategy for $[(AG)_3EG]_n$ polypeptides. A. Sequence and multimerisation of synthetic oligonucleotide 1 coding for 2 octapeptide repeats. B. The amplification plasmid pSK-JS2 was made by cloning of synthetic oligonucleotide 2 into the *Eco*RI and *Bam*HI site of pBluescript® II SK (-). After digestion of this plasmid with *Bsp*MI the multimeric gene was introduced. Finally the sequence was transferred to the pET-3b expression vector using *Bam*HI (not shown).

Expression and purification protocol. The pET-3b expression vectors carrying the artificial genes coding for 10 or 20 repeats of the octapeptide sequence $(AG)_3EG$ were transformed to BL21(DE3)pLysS cells (Novagen Inc., Madison, USA) and grown overnight at 30 °C. A single colony was used to inoculate 250 ml 2 × YT medium containing 100 µg/ml ampicillin, 34 µg/ml chloramphenicol and 1% glucose. After growth overnight at 30 °C this preculture was

used to inoculate 4.5 L of 2 × YT medium to OD = 0.1 and cells were grown at 37°C. Protein expression was induced during logarithmic growth by addition of IPTG to a final concentration of 1 mM. Cells were harvested after 4 hours of expression by centrifugation at 6,000 × g for 15 minutes at 4 °C. Cells were resuspended in 50 ml lysis buffer (50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol and 1 mM PMSF). Cells were disrupted by sonication on ice for 5 minutes using a 250 W Branson sonicator (50% duty cycle, 5 units power). RNase A (10 μg/ml) and DNase I (5 μg/ml) were added followed by incubation on ice for 15 minutes. The lysate was centrifuged at 10,000 × g for 30 minutes to pellet the cellular debris. The supernatant was incubated with 5 ml Ni-NTA agarose beads (Qiagen) for 1 hour at 4 °C. The suspension was then loaded onto the column followed by washing with 40 ml wash buffer (as lysis buffer, except 20 mM imidazole). The protein was eluted with 10 ml elution buffer containing 200 mM imidazole. For $[(AG)_3EG]_{20}$ the eluate was heated for 10 minutes at 70 °C followed by centrifugation at 6,000 × g for 15 minutes. Gel filtration chromatography with a Superdex-75 Hi-Load™ 26/60 column (Amersham Biosciences) resulted in the final purified product (eluent: 50 mM NaH_2PO_4 (pH 8.0), 150 mM NaCl, 10 mM β-mercaptoethanol; flow 2 ml/min; room temperature). Finally, the product was dialysed against demi-water using a dialysis membrane with a molecular weight cut-off of 3,500 Da (Spectrapor) for 2 days and lyophilized.

CNBr cleavage. Protein was dissolved in 70% formic acid (2 mg/ml) and an equal volume of CNBr in 70% formic acid (100 mg/ml) was added followed by incubation on a rotary arm for 2 days at room temperature in the dark.[1] The samples were dried in a centrifugal dryer at room temperature. The pellets were redissolved in demi-water and dialyzed for 2 days against demi-water using a dialysis membrane with a molecular weight cut-off of 3,500 Da (Spectrapor).

Crystallization experiments. Crystallization was induced by vapor diffusion of methanol into a 70% formic acid 10 mg/ml protein solution. This was achieved by placing an eppendorf tube with protein solution (100 μl) in a bigger container filled with 20 ml methanol. Gelation was observed after 2 days.

NMR spectroscopy. ^1H NMR spectra were recorded on a Varian Inova-400 instrument at 298 K. Chemical shifts are reported in ppm relative to the H_2O signal (4.79 ppm). ^{13}C NMR spectra were recorded on a Bruker AC-300 instrument at 298 K.

Maldi-TOF analysis of conjugates. MALDI-TOF-MS spectra were measured on a Bruker Biflex III spectrometer. Samples were dissolved in water/acetonitrile (1:1 v/v) containing 1% trifluoroacetic acid and mixed in a 1:1 ratio with a solution of 20 mg/ml of sinapinic acid in the same solvent and spotted on a MALDI-plate.

Infrared and circular dichroism spectroscopy. IR spectra were recorded on a Thermo Mattson Genesis IR-300 spectrometer equipped with an ATR cell. Circular dichroism spectroscopy was performed on a Jasco J-810 spectropolarimeter (band width: 1nm, response: 1 sec., sensitivity: standard, l range: 260-180 nm, data pitch: 0.5 nm, scanning speed: 100 nm/min., accumulation: 5).

Protein expression and purification

The expression and purification of cysteine-flanked $[(AG)_3EG]_{10}$ and $[(AG)_3EG]_{20}$ is depicted in Figure S2a and S2b, respectively. Lanes 1 and 2 show the total protein content before and after induction with IPTG. The expression levels of both proteins were low and therefore T7 tag Western blot analysis was used to follow the purification of both proteins (not shown). Both proteins were present in the soluble fraction of the lysate (lane3). Lane 5 shows the purity after native Ni-NTA purification. For $[(AG)_3EG]_{20}$ the purity could be increased substantially by heating the sample for 10 minutes at 70 °C which resulted in the precipitation of endogenous *E. coli* protein whereas the protein of interest remained soluble (Figure S2b, lane 6). The molecular weight, as determined by electrophoresis, were higher than theoretically predicted, due to the highly acidic nature of these polypeptides. The anomalous behaviour of highly acidic proteins has been described before by Mc Grath et al.[2] The expected molecular weights of 11.1 kDa for $[(AG)_3EG]_{10}$ and 16.8 kDa for $[(AG)_3EG]_{20}$ were confirmed my MALDI-TOF mass spectrometry. In addition to the full-length protein we observed for both $[(AG)_3EG]_{10}$ and $[(AG)_3EG]_{20}$ the presence of truncated product at the height of the 21 kDa molecular weight marker. Western blot analysis showed that this product contained the N-terminus of the desired protein. Maldi-TOF analysis indicated that this band is not a discrete protein product, but a mixture of truncated products from 5.9 kDa to 6.9 kDa. These truncated products were removed using gel filtration chromatography (Figure S2a, lane 6 and Figure S2b, lane 7). Isolated yields for $[(AG)_3EG]_{10}$ and $[(AG)_3EG]_{20}$ were 6 mg/L culture and 3 mg/L, respectively.

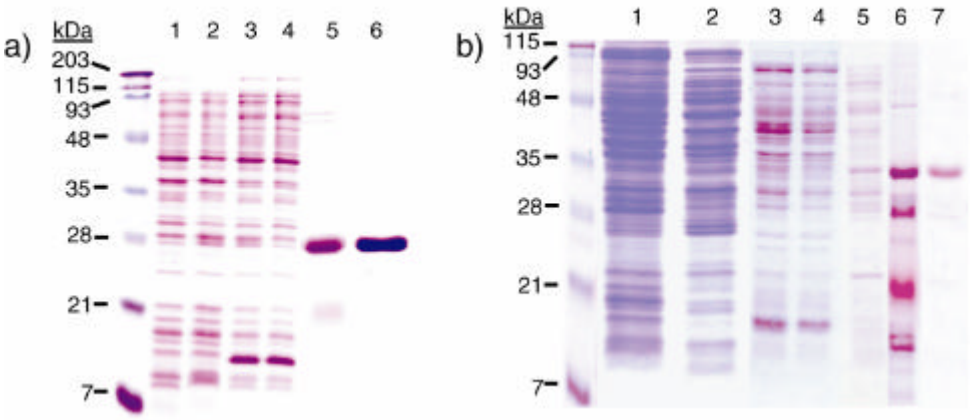


Figure S2. SDS-PAGE analysis of expression and purification of $[(AG)_3EG]_n$ polypeptides (15% acrylamide, staining with Coomassie Brilliant Blue R-250). a) $[(AG)_3EG]_{10}$. Lane 1 and 2: whole cell lysate of *E. coli* BL21(DE3)pLysS uninduced and 3 hours after induction with IPTG; lane 3: clarified soluble fraction of lysate; lane 4: flow-through of Ni-NTA column; lane 5 elution fraction of Ni-NTA column; lane 6: protein after purification with Superdex-75 column. b) $[(AG)_3EG]_{20}$. Lane 1-5: as a); lane 6: soluble fraction after heating for 10 minutes at 70 °C; lane 7: protein after purification with Superdex-75 column.

Synthesis of maleimide-functionalized poly(ethylene glycol)-750

Poly(ethylene glycol)-750 monofunctionalized with an amine group was purchased from Rapp Polymere GmbH (Tübingen, Germany). 391 mg PEG-750-NH_2 (0.52 mmol) was dried by azeotropic evaporation with benzene. The dried product was dissolved in 5 ml DMF together with 110 mg ε-maleimidocaproic acid (0.52 mmol; Sigma), 230 mg BOP coupling reagent (0.52 mmol ; benzotriazole-1-yl-oxy-tris-(dimethylamino)-phosphonium hexafluorophosphate; Advanced Chemtech) and 192 mg diisopropylethylamine (1.60 mmol; Fluka). After 24 hours stirring at room temperature DMF was evaporated and the resulting solid was redissolved in dichloromethane. This solution was subsequently extracted with 1N HCl (2 ×), water, 5% $NaHCO_3$ (2 ×), water and saturated NaCl. After evaporation the product was further purified with a Sephadex LH-20 gel filtration column (resin from Amersham Biosciences) using methanol/dichloromethane 1:1 v/v as eluent. The final yield was 217 mg pure product (0.35 mmol, 68%). R_f = 0.60 – 0.75 (methanol/chloroform = 1 : 4 v/v, maleimide detection); ^{1}H NMR (400 MHz, $CDCl_3$): ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.69 (s, 2H,C*H*=C*H*), 6.27 (br s, 1H, N*H*CO), 3.61-3.68 (br m, 60H, O(CH_2)$_2$O), 3.55 (t, 2H, CH_2CH$_2$NHCO), 3.51 (t, 2H, NCH_2), 3.44 (m, 2H, CH_2NHCO), 3.38 (s, 3H, CH_3O), 2.17 (t, 2H, CH_2CONH), 1.66 (m, 2H, NHCOCH$_2$CH_2), 1.60 (m, 2H, CH_2CH$_2$N), 1.31 (m, 2H, CH$_2$CH_2CH$_2$CH$_2$N); ^{13}C NMR (75 MHz, $CDCl_3$): δ = 172.1 (1C, *C*ONH), 163.8 (2C, N*C*O), 133.5 (2C, C=C), 71.6 (1C, *C*H$_2$CH$_2$NH), 70.2 (30C, O(*C*H$_2$)$_2$O), 58.7 (1C, *C*H$_3$O), 38.9 (1C, *C*H$_2$NH), 37.5 (1C, *C*H$_2$N), 36.1 (1C, *C*H$_2$CONH), 28.1 (1C, *C*H$_2$CH$_2$N), 26.2 (1C, CH$_2$*C*H$_2$CH$_2$), 24.9 (1C, *C*H$_2$CH$_2$CONH); MS (MALDI-TOF): Calcd. for n = 15, $C_{43}H_{80}N_2O_{19}$ $[M + Na]^+$: 952.05, found m/z = 951.70, M_n = 1045, M_w/M_n = 1.01.

Conjugation of PEG-750-maleimide to $[(AG)_3EG]_n$ polypeptides

10 mg of freeze-dried protein was dissolved in 5 ml 20 mM NaH_2PO_4 buffer (pH 8.0) containing 150 mM NaCl, 1 mM EDTA and 200 mM dithiotreitol (DTT) and incubated for 1 hour at room temperature. The protein was precipitated by addition of 0.25 volumes of ice-cold 100% trichloroacetic acid (TCA), followed by incubation at -20 °C for 30 minutes. After centrifugation at 13000 rpm for 10 minutes at 4 °C the pellet was washed with 2.5 ml ice-cold 20% TCA followed by a second wash of 2.5 ml 1% TCA for $[(AG)_3EG]_{10}$ and 2.5 ml milli-Q for $[(AG)_3EG]_{20}$. After each wash the solution was centrifuged for 5 minutes at 4 °C. The pellet was redissolved in 5 ml 100 mM NaH_2PO_4 buffer (pH 6.8) containing 150 mM NaCl. The thiol

content of this solution was determined using the Ellman's essay[3]: The thiol content for $[(AG)_3EG]_{10}$ and $[(AG)_3EG]_{20}$ was determined to be respectively a factor 1.4 and 1.8 times higher than the theoretical thiol content, indicating the presence of residual DTT. For complete PEGylation of the cysteine residues an excess of maleimide functionalised poly(ethylene glycol) (M_n =750 g/mol) was immediately added as a 5 ml solution in the same buffer. For $[(AG)_3EG]_{10}$ a 5-fold excess was used whereas a 20-fold excess was used for $[(AG)_3EG]_{20}$. The reaction mixture was incubated overnight on a rotating arm.

Ni-NTA chromatography was used to remove the excess poly(ethylene glycol). The Ni-NTA beads were pre-equilibrated in 100 mM NaH_2PO_4 buffer (pH 8.0) containing 150 mM NaCl. The PEG-conjugation reaction was added to 10 ml equilibrated 50% Ni-NTA suspension and incubated for 1 hour at room temperature. The suspension was centrifuged for 5 minutes at 1500 rpm and the supernatant was removed. 10 ml wash buffer (100 mM NaH_2PO_4 (pH 8.0), 150 mM NaCl) was added and a column was loaded. The beads were washed until all the poly(ethylene glycol) was removed (monitored by measuring absorption profile at 214 nm). Subsequently the protein-PEG conjugate was eluted by increasing the imidazole concentration to 200 mM. Finally, the product was dialysed for 2 days against demi-water using a dialysis membrane (MWCO = 3,500 Da; Spectrapor) and lyophilized. SDS-polyacrylamide gel electrophoresis showed that the attachment of PEG-chains resulted in a different electrophoretic mobility for both proteins (Figure S3). Noteworthy is the different effect of the attachment of poly(ethylene glycol)-750 on the electrophoretic mobility of the conjugates. For $[(AG)_3EG]_{10}$ the band shifted to a lower apparent molecular weight, whereas the opposite was the case for $[(AG)_3EG]_{20}$.

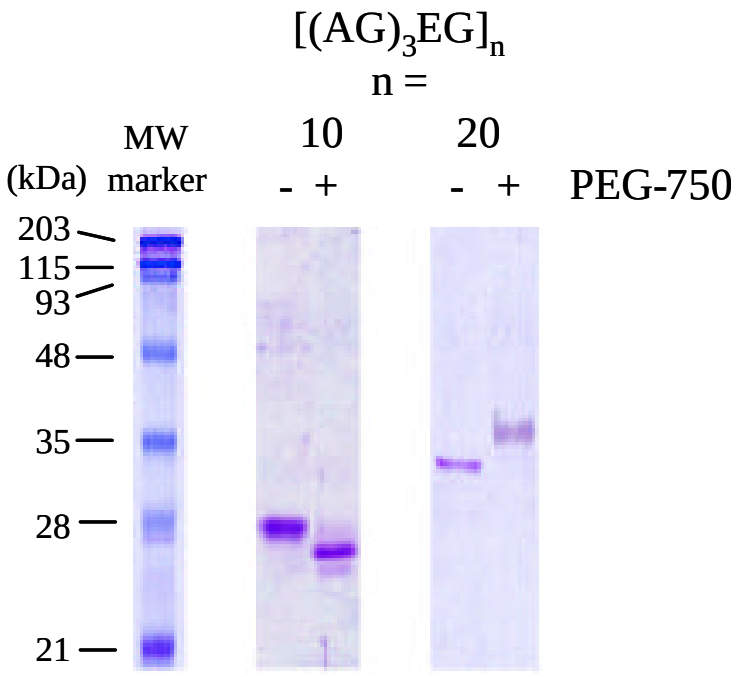


Figure S3. SDS-PAGE analysis of conjugation of PEG-750-maleimide with $[(AG)_3EG]_{10}$ and $[(AG)_3EG]_{20}$.

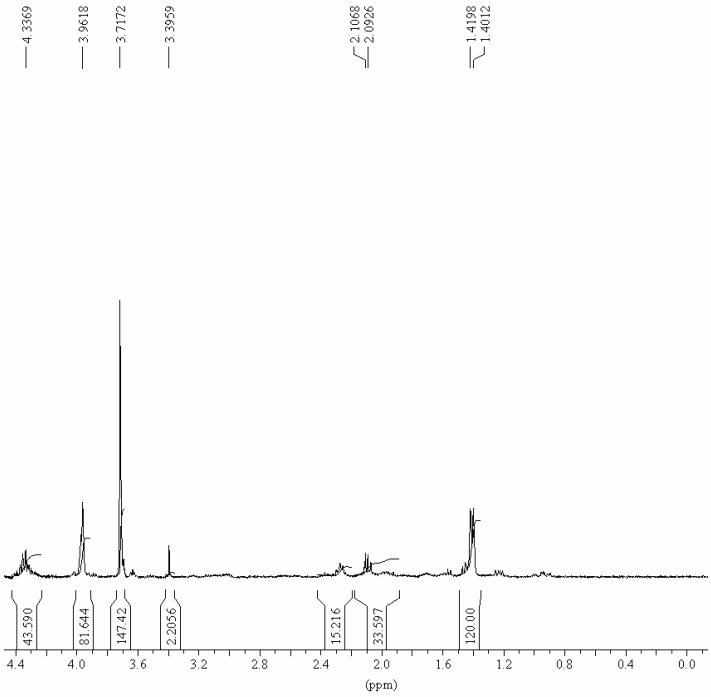
Chemical characterization of conjugates

^{1}H NMR spectroscopy. The conjugates were dissolved in D_2O. The ^{1}H NMR spectrum of conjugates of $[(AG)_3EG]_{10}$ and $[(AG)_3EG]_{20}$ with PEG-750 are depicted in Figure S4a and S4b, respectively.

For $[(AG)_3EG]_{10}$ reacted with PEG-750-maleimide the following (major) signals could be assigned: δ = 1.40 (d, 120H, Ala-Hβ), 2.10 (m, 26H, Glu-Hβ), 2.26 (m, 26H, Glu-Hγ), 3.40 (s, 6H, CH_3O), 3.72 (s, 128H, O(CH_2)O), 3.96 (s, 88H, Gly-Hα), 4.33 (m, 53H, Ala/Glu-Hα). To determine the efficiency of conjugation with poly(ethylene glycol) and the subsequent removal of excess non-reacted poly(ethylene glycol), the signal intensity of the methyl group of alanine (Ala-Hβ) was compared to the signal intensity of the repeating ethyleneoxide units (O(CH$_2$)$_2$O) of poly(ethylene glycol). The ethylene oxide signal was 1.2 times the expected value and therefore in reasonable agreement with the attachment of two poly(ethylene glycol) chains per protein molecule (Figure S4a).

For $[(AG)_3EG]_{20}$ reacted with PEG-750-maleimide the following signals could be assigned: δ = 1.40 (d, 210H, Ala-Hβ), 2.10 (m, 46H, Glu-Hβ), 2.26 (m, 46H, Glu-Hγ), 3.39 (s, 6H, CH_3O), 3.71 (s, 120H, O(CH_2)O), 3.96 (s, 168H, Gly-Hα), 4.34 (m, 93H, Ala/Glu-Hα). Comparison of the ethylene oxide signal with the methyl group of alanine showed that the ethylene oxide signal was 1.4 times the expected value (Figure S4B).

a)



b)

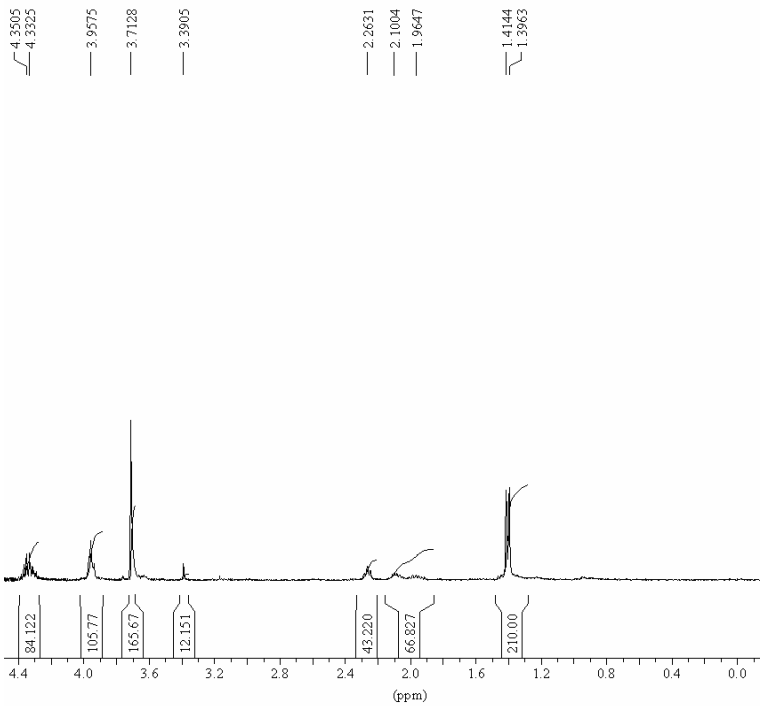


Figure S4. 400 MHz ^{1}H NMR spectrum of PEG-$[(AG)_3EG]_n$-PEG: a) n = 10 and b) n = 20. Solvent: D_2O.

Maldi-TOF mass spectrometry. The main peak after conjugation of maleimide-functionalized PEG-750 was observed at 13280 Da which is in agreement with the attachment of two PEG-chains (calculated mass 13202). After cyanogen bromide cleavage the main peak was observed at 8529 Da in good agreement with the theoretical mass of 8508 Da. For $[(AG)_3EG]_{20}$ a molecular ion was observed at m/z 16811 (1+) (expected value = 16817). Upon conjugation with maleimide-functionalized poly(ethylene glycol) a shift in mass was observed in agreement with the conjugation of two PEG-chains (observed m/z 18679). In addition a smaller peak (m/z 19807) was observed which could indicate the attachment of three PEG-chains, possibly caused by attachment to a lysine residue on the C-terminal side of the protein.

Secondary structure characterization

Infrared spectroscopy. Infrared spectra were recorded by pipetting the gelated sample in methanol (1mg/ml) on the surface of the ATR crystal followed by drying. The IR spectra of PEG-$[(AG)_3EG]_{10}$-PEG conjugates exhibited strong amide I and II vibrational modes at ~1623 cm^{-1} and ~1522 cm^{-1}, characteristic of the β-sheet conformation[4], whereas the weak amide I component observed at ~1697 cm^{-1} indicated its antiparallel nature (Figure S5a)[5]. Similar spectra were recorded for PEG-$[(AG)_3EG]_{20}$-PEG (Figure S5b). We can conclude that the antiparallel β-sheet structure is retained upon conjugation of poly(ethylene glycol). The amide I band around 1654 cm^{-1} (and also amide II bands around 1550, not indicated) indicates that some fraction of the polypeptide chain has adopted a secondary structure other than that of the antiparallel β-sheet. The amide I component has been assigned previously to reverse turns of the β or γ type[6], but may also partly result from a random coil/α-helical conformation.

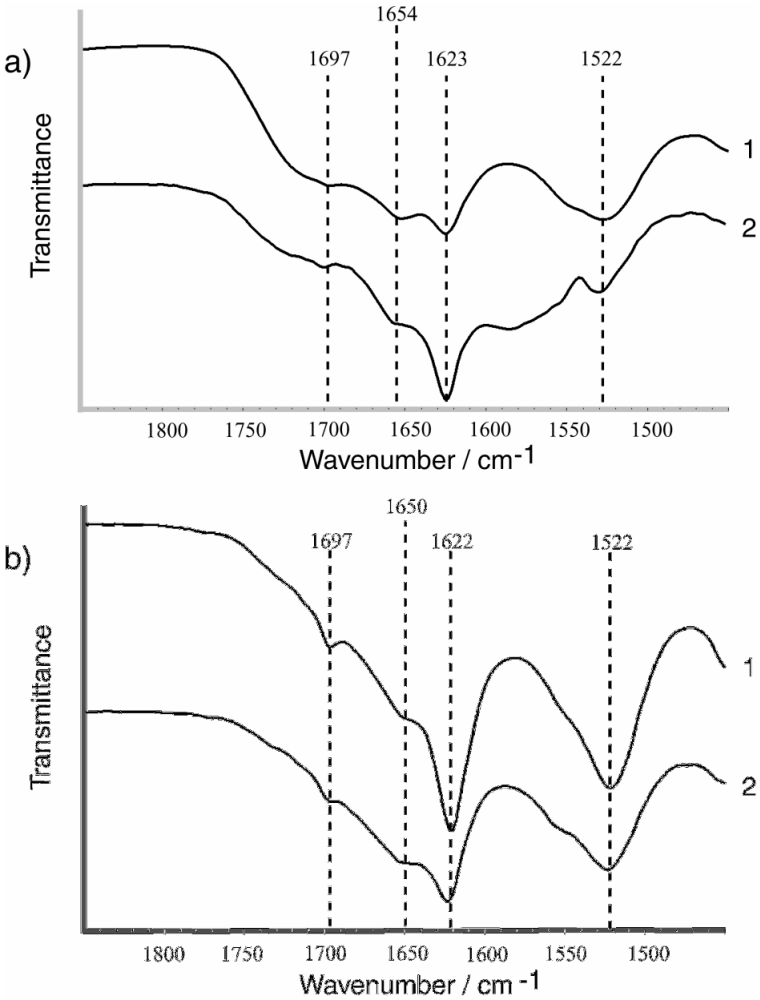


Figure S5. ATR-infrared spectra of PEG-$[(AG)_3EG]_n$-PEG conjugates in amide I and amide II regions for a) n = 10: 1) uncleaved 2) CNBr cleaved b) n = 20: 1) uncleaved 2) CNBr cleaved.

Circular dichroism spectroscopy. CD spectroscopy was performed on PEG-$[(AG)_3EG]_{20}$-PEG conjugates. Only the uncleaved variant gave a strong CD-effect and confirmed the β-sheet conformation with a negative ellipticity at 215 nm and a positive ellipticity at 195 nm (Figure S6).[7]

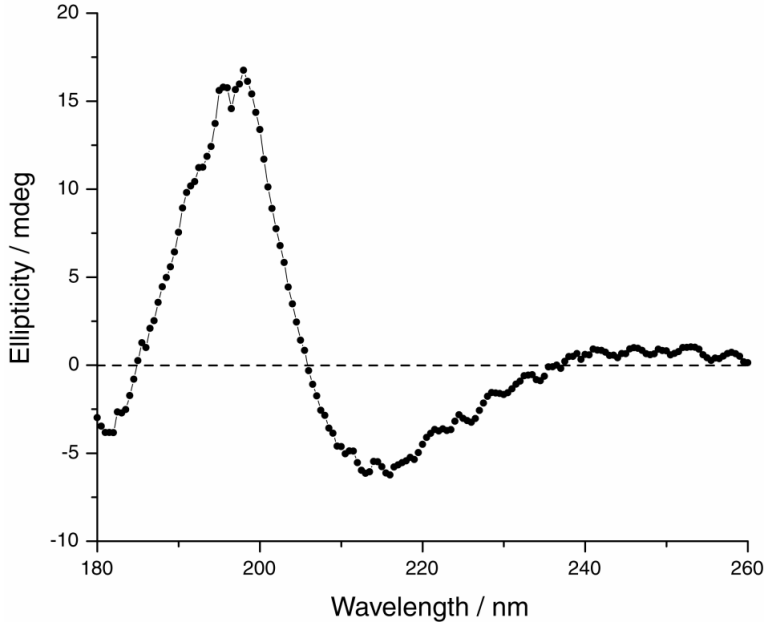


Figure S6. Circular dichroism spectrum of uncleaved PEG-$[(AG)_3EG]_{20}$-PEG conjugate (T = 20 °C). 50 µl crystallized sample (1 mg/ml in methanol) was applied to a quartz surface and allowed to dry.

Transmission electron microscopy analysis

For the proteins without conjugated poly(ethylene glycol) only few individual fibres were found, which is most likely the result of the formation of bigger aggregated structures, due to increased interaction between the β-sheet polypeptides. TEM images for uncleaved and CNBr cleaved samples of $[(AG)_3EG]_{20}$ are depicted in Figure S7a and b, respectively. TEM analysis of $[(AG)_3EG]_{36}$ by Krejchi et al.[6] showed crystalline platelets of ~2000 Å (assigned as a) by 150 Å (assigned as b) by about 40 to 100 Å (from shadow lengths; assigned as c). No indication of similar platelets were found in our samples.

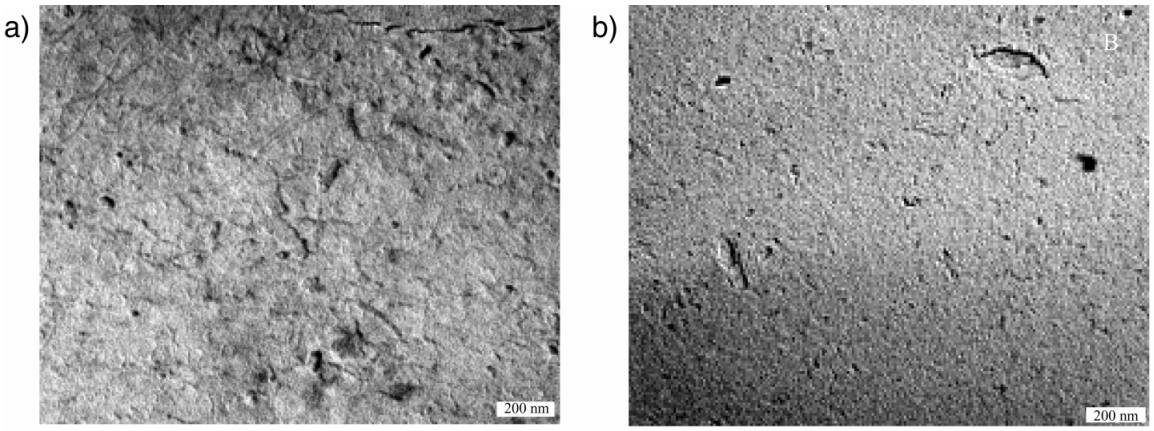


Figure S7. TEM micrographs of crystallized $[(AG)_3EG]_n$ polypeptides without conjugated poly(ethylene glycol)-750. Platinum shadowed samples (1 mg/ml) of a) uncleaved $[(AG)_3EG]_{20}$; b) CNBr cleaved $[(AG)_3EG]_{20}$.

References

[1] F. Ausubel, R. Brent, R. Kingston, D. Moore, J. Seidman, J. Smith, K. Struhl, *Current protocols in molecular biology, volume 3, sec. 16.4.13, Vol. 16.4.13*, John Wiley & sons, New York, **1999**.

[2] K. P. McGrath, M. J. Fournier, T. L. Mason, D. A. Tirrell, *J. Am. Chem. Soc.* **1992**, *114*, 727.

[3] G. L. Ellman, *Arch. Bioch. Biophys.* **1959**, *82*, 70.

[4] S. Krimm, J. Bandekar, *Adv. Protein Chem.* **1986**, *38*, 181.

[5] M. Miyazawa, E. R. Blout, *J. Am. Chem. Soc.* **1960**, *83*, 712.

[6] M. T. Krejchi, E. D. T. Atkins, A. J. Waddon, M. J. Fournier, T. J. Mason, D. A. Tirrell, *Science* **1994**, *265*, 1427.

[7] N. Greenfield, G. D. Fasman, *Biochemistry* **1969**, *8*, 4108.