Supporting Information

for

Bulky Toroidal and Vesicular Self-Assembled Nanostructures from Fullerene End-Capped Rod-Like Polymers

Daniela Mazzier, Miriam Mba, Mirco Zerbetto, and Alessandro Moretto**

CONTENTS

GENERALMETHODS

NMR: ^{1}H and ^{13}C NMR spectra were recorded at room temperature on a Bruker AC-200 (200 MHz) and Bruker AC-300 (300 MHz) instrument using TMS as internal reference. The multiplicity of a signal is indicated as: s-singlet, d-doublet, m-multiplet. Chemical shifts (δ) are expressed in ppm and coupling constants (J) in Hertz.

FT-IR absorption: FT-IR absorption spectra were recorded with a Perkin-Elmer 1720X spectrophotometer; ν_{max} is given for the main absorption bands.

UV-Vis Absorption: The UV-Vis absorption spectra were recorded using a Shimadzu model UV-2501PC spectrophotometer. A 1-cm pathlength quartz cell was used.

CD: Circular dichroism measurements were carried out at room temperature using a Jasco J-715 spectropolarimeter. A fused quartz cell of 0.2-mm path length (Hellma) was used.

TEM: Samples were analyzed on a Jeol 300PX instrument. Samples were prepared immediately before used, by dilution of the dialyzed solutions with MilliQ water. A small drop of solutions was floated on a glow discharged carbon coated grid and excess was removed by #50 hardened Whatman filter paper. For the samples with negative staining, the grid was then floated on 2% uranyl acetate solution for 10 seconds, and the excess was removed by #50 hardened Whatman filter paper.

SEM: A Carl Zeiss Merlin field emission scanning electron microscope operating at 5kV accelerating voltage was used. A small drop of the milk-like aqueous suspension was placed on a microscope glass cover slip and allowed to dry overnight. The dry material was coated with platinum.

AFM: AFM experiments were performed on Agilent Technologies 5500 scanning probe microscope, operating in acoustic AC AFM mode (tapping mode) with a silicon Asylum Reasearch high frequency cantilever displaying a resonance frequency of 305 kHz. A small drop of sample was placed on a mica or silicon surface and allowed to dry in air for at least 60 min before measurement.

SEC analysis: measurements were done on a Agilent 1260 Infinity system equipped with 1260 isopump, 1260 TCC, 1260 VWD VL, 1260 RID, Phenogel 5u linear/mixed guard column (30 x 4.6 mm), followed by Phenomenex Phenogel 5u 10^4 Å (300 x 4.6 mm) column working at 60 °C. DMF was used as eluent at a flow rate of 1 ml/min. Before SEC analysis is performed, the samples were filtered through a 0.2 μm PTFE filter (15 mm, Phenomenex).The molecular weights were calculated using polystyrene standards.

SYNTHESISAND CHARACTERIZATION

General

γ-benzyl L-glutamic acid was obtained from Fluka. 1-hydroxy-7-aza-1,2,3-benzotriazole (HOAt) and 1-ethyl-3-(3-Dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl) were purchased from GL Biochem (Shanghai) Ltd. Triphosgene, cysteamine hydrochloride, triisopropylsilane (TIS) and trifluoroacetic acid (TFA), α-pinene and triphenylmethanol were obtained from Sigma-Aldrich. The deuterated solvent $CDCl_3$ and DMSO-d_6 was purchased from Euriso-Top (France). All other chemicals and solvents were Sigma-Aldrich, Fluka or Acros products and used as provided without further purifications.

Synthesis of γ-benzyl L-glutamate *N*-carboxyanhydride (BLG-NCA)

γ-benzyl L-glutamic acid (5.05 g, 21.3 mmol) and α-pinene (6.64 g, 48.7 mmol) were dissolved with 70 mL of ethyl acetate in a three-neck flask and heated under reflux. Triphosgene (4.24 g, 14.2 mmol) was dissolved in 25 ml ethyl acetate and added slowly with a dropping funnel once the reflux started. After 4h of reaction, the heating was interrupted and 3/4 of the solvent was evaporated under reduced pressure. The compound was precipitated by addition of petroleum ether and the recovered solid was recrystallized twice, subsequently filtered and washed with petroleum ether. The NCA was recovered as a white solid (4.6 g, yield 84%).

^{1}H-NMR (200 MHz, $CDCl_3$): δ 7.35 (s, 5H, ArH), 6.61 (s, 1H, NH), 5.14 (s, 2H, Ar-CH_2), 4.37 (t, 1H, αCH), 2.59 (t, 2H, γCH_2), 2.36-2.02 (m, 2H, β CH_2).

Reference

G. J. M. Habraken, M. Peeters, C. H. J. T. Dietz, C. E. Koning, A. Heise, *Polym. Chem.*, **2010**, *1*, 514-524.

Synthesis of 2-(tritylsulfanyl)ethanamine

A solution of cysteamine hydrochloride (1.04 g, 9.17 mmol) and triphenylmethanol (2.18 g, 8.37 mmol) in TFA (5 mL) was stirred at room temperature for 1 h. After co-evaporation with acetonitrile the residue was dissolved in ethyl acetate and washed with $NaOH_{(aq)}$ 1 M, water and brine. The organic layer was dried over Na_2SO_4, filtered and the solvent was evaporated under reduced pressure. The compound was recovered as white solid (2.4 g, yield 89 %).

^{1}H NMR (200 MHz, $CDCl_3$) δ 7.48-7.43 (m, 7H, ArH), 7.35-7.23 (m, 8H, ArH), 2.61 (t, 2H, CH_2), 2.35 (t, 2H, CH_2), 1.42 (s, 2H, NH_2).

Reference

K. Görmer, M. Bürger, J. A. W. Kruijtzer, I. Vetter, N. Vartak, L. Brunsveld, P. I. H. Bastiaens, R. M. J. Liskamp, G. Triola, and H. Waldmann, *Chem Bio Chem* **2012**, *13*, 1017-1023.

Synthesis of 3-(tritylsulfanyl)propanoic acid

To a suspension of triphenylmethanol (2.6 g, 10 mmol) in 10 mL of TFA was added 3-mercaptopropionic acid (1.4 g, 13 mmol). The mixture was stirred for 1 h at room temperature. After that the TFA was removed under vacuum and the residue was dissolved in ethyl acetate and washed with $NaHCO_{3(aq)}$ 5%, $KHSO_{4(aq)}$ 10%,water and brine. The organic layer was dried over Na_2SO_4, filtered and the solvent was evaporated under reduced pressure. The compound was recovered as white solid (3.05 g, yield88 %).

^{1}H NMR (200 MHz, $CDCl_3$) δ 7.47-7.40 (m, 6H, ArH), 7.35-7.22 (m, 9H, ArH), 2.49 (t, 2H, CH_2), 2.25 (t, 2H, CH_2).

Synthesis of PBLG derivatives

NH_2-PBLG-S-Trt (1)

The BLG-NCA (1 g, 3.8 mmol) was dissolved in 3 mL of dry dioxane under N_2 atmosphere and the flask was cooled with a MeOH/ice bath. Then (tritylsulfanyl)ethanamine (12 mg, 0.038 mmol) dissolved in 500 µL of dry dioxane was added to the solution and the cool bath was removed. The reaction was stirred at 25 °C for one day, until the mixture became very viscous. The product was precipitated by adding MeOH to the solution, the polymer was filtered and dried in vacuo. The polymer was recovered as a white fibrous solid (810 mg, recovered yield 81 %).

NH_2-PBLG-C_{60} (3)

The removal of Trt protective group was performed by treatment of polymer **1** (200 mg, 10.5 µmol) with 4 mL of a solution 50% TFA in CH_2Cl_2 with TIS for 20 min. The solvent was evaporated under reduced pressure and the residue was suspended in Et_2O, filtered and washed with Et_2O. The resulting polymer (**2**) was dissolved in 10 mL of a solution of DMF/toluene 2:8 v/v. Then fullerene (37 mg, 0.052mmol) and AIBN (24 mg, 0.16mmol) were added. The reaction mixture was heated at 90 °C for 5 h, after that the solvent was removed under reduced pressure. The residue was suspended in DMF and filtered through a 20 µm filter to eliminate the non-reacted fullerene. The polymer was precipitated from the filtrate solution by adding MeOH. The solid was filtered and washed several times with MeOH and acetone to remove the fullerene by-products. The polymer **3** was recovered as brown solid (171 mg, yield 85 %).

Trt-S-PBLG-C_{60} (6)

Trt-*mercapto*propionic acid (60 mg, 0.17 mmol) and HOAt (24 mg, 0.17 mmol) were dissolved in 3 mL of anhydrous DMF, cooled at 0 °C and then EDC·HCl (33 mg, 0.17 mmol) was added. The mixture was stirred for 15 min at 0 °C. Polymer **3** (90 mg, 4.6 µmol) was dissolved in 2 mL of DMF and added with TEA (50 µL, 0.36 mmol) to the solution of active ester. After stirring the solution at r. t. for 48 h, the polymer was precipitated by adding MeOH, then filtered and washed several times with MeOH. The polymer was dry to obtain 92 mg, yield 98 %.

C_{60}-PBLG-C_{60} (6)

Polymer **4** (75 mg, 3.8 μmol) was treated with 4 mL of a solution 50% TFA in CH_2Cl_2 with TIS for 20 min. The solvent was evaporated under reduced pressure and the residue was suspended in Et_2O, filtered and washed with Et_2O. The obtained compound (**5**) was dissolved in 5 mL of a solution of DMF/toluene 2:8 v/v. Successively fullerene (15 mg, 0.021 mmol) and AIBN (10 mg, 0.061 mmol) were added. The mixture was heated at 90 °C for 5 h, after that the solvent was removed under reduced pressure. The residue was suspended in DMF and filtered through a 20 μm filter. To the filtrated solution MeOH was added to precipitate the polymer. The compound was filtered and washed several times with MeOH and acetone. The compound was obtained as a brown solid (60 mg, yield 78 %).

Microstructure Preparation

Typically, 20 mg of polymer were dissolved in 10 mL of a solvent mixture DMF/THF 3:7 v/v. This solution was put into a cellulose dialysis tubing with a molecular weight cutoff of 12 KD (flat with 35 mm, Sigma). The dialysis process was carried out against deionized water for 48 h.

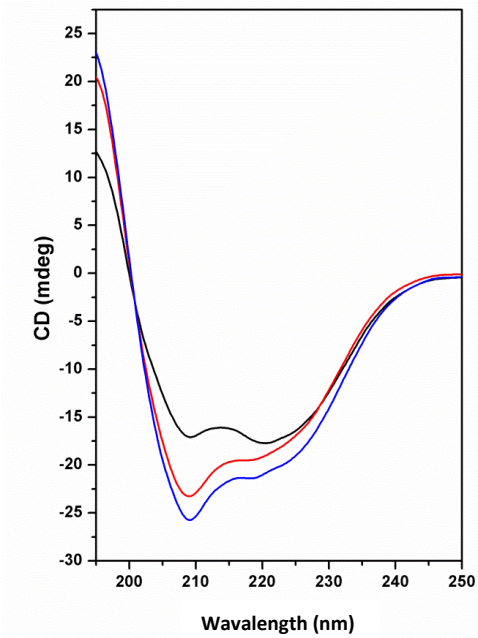


Figure S1. Comparison of ECD spectra of polymers **1** (black), **3** (red) and **6** (blue).

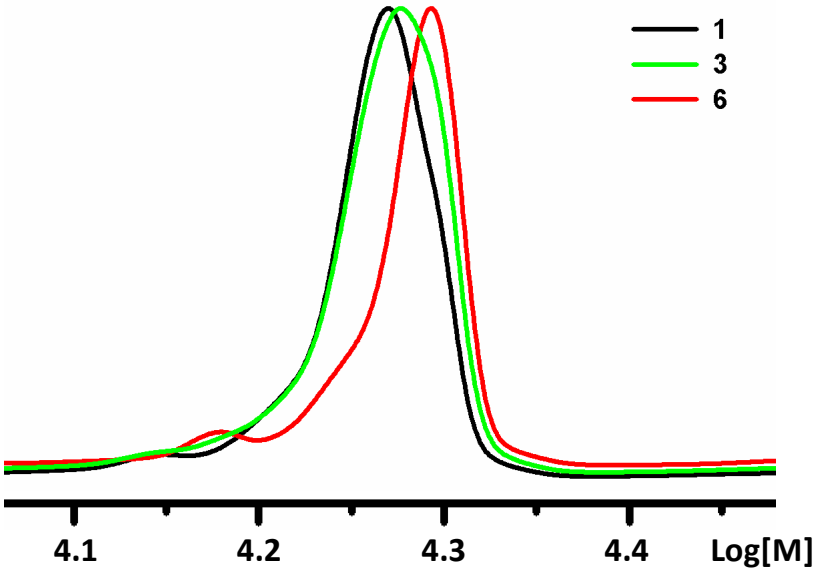


Figure S2. Polymers **1**, **3** and **6** SEC traces. Low molecular weight polymers were also detected in less than 2%, according with other reported N-carboxyanhydride ring opening polymerizations they come from a secondary initiation mechanism.

Reference

G. J. M. Habraken, M. Peeters, C. H. J. T. Dietz, C. E. Koning, A. Heise, *Polym. Chem.*, **2010**, *1*, 514-524.

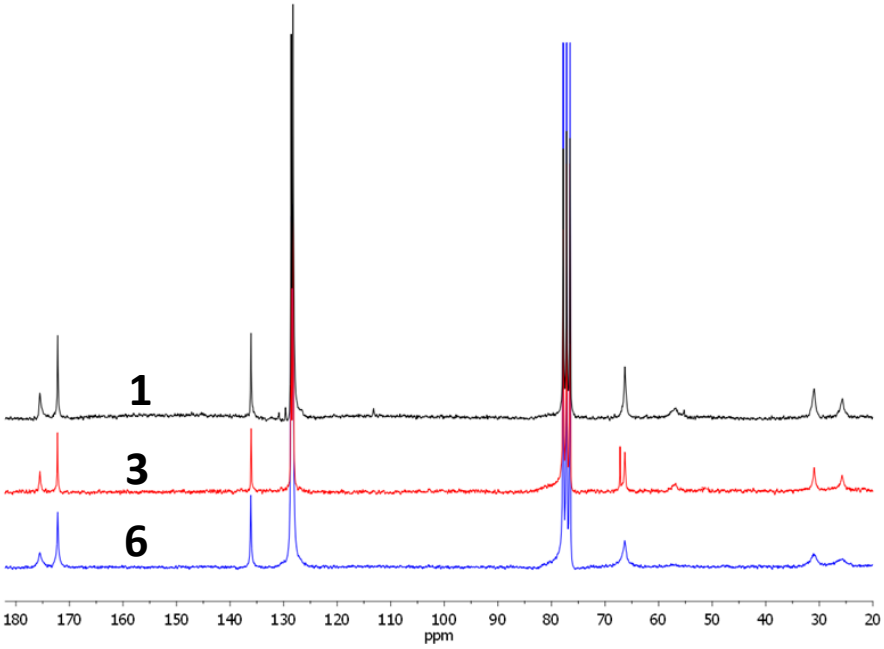


Figure S3. Comparison of ^{13}C-NMR ($CDCl_3$) spectra of polymers **1**, **3** and **6**.

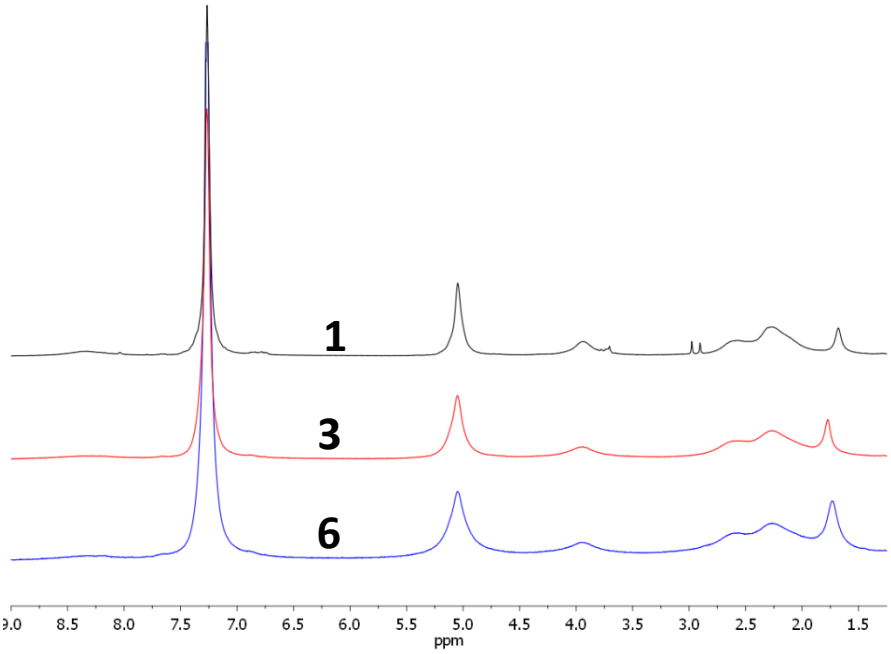


Figure S4. Comparison of ^{1}H-NMR ($CDCl_3$) spectra of polymers **1**, **3** and **6**.

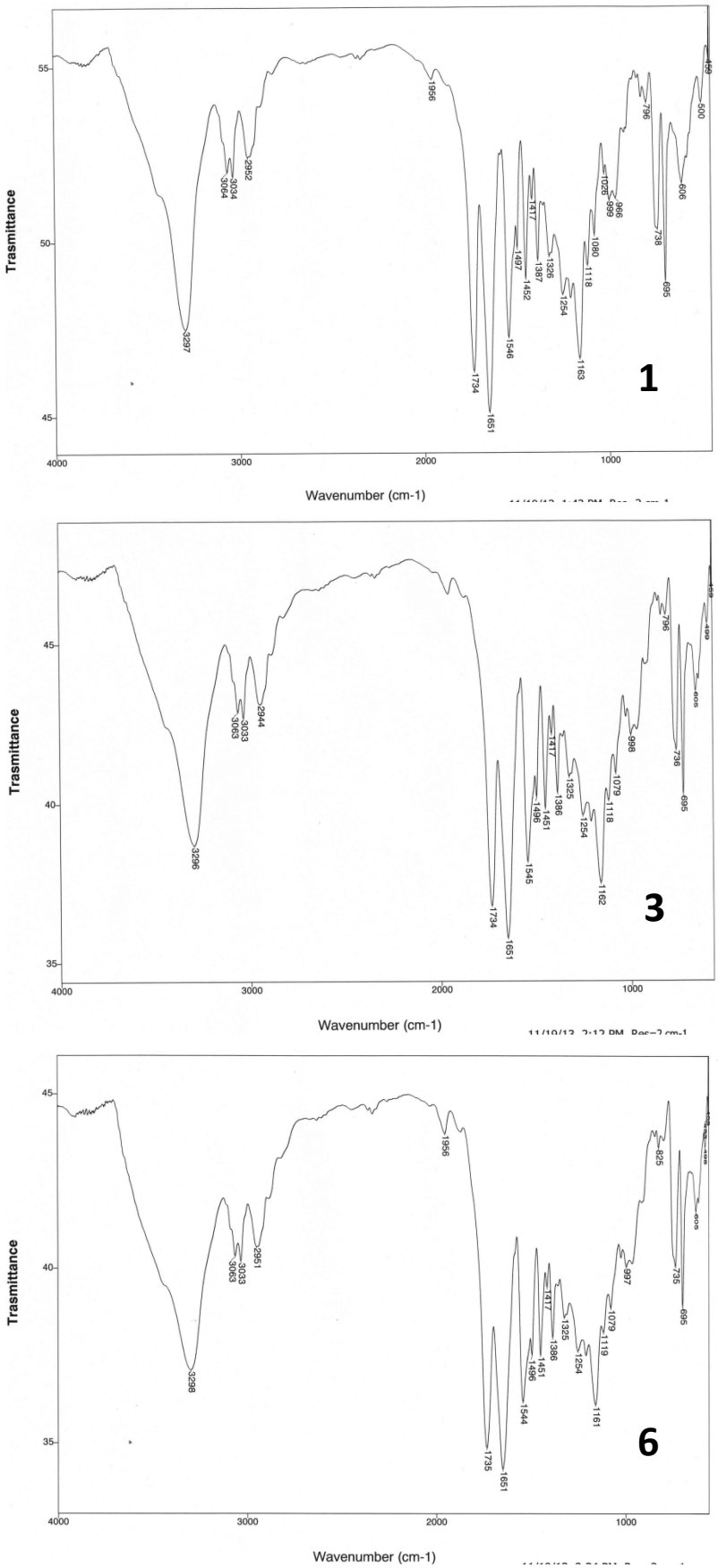


Figure S5. Comparison of solid state (KBr) IR spectra of polymers **1**, **3** and **6**.

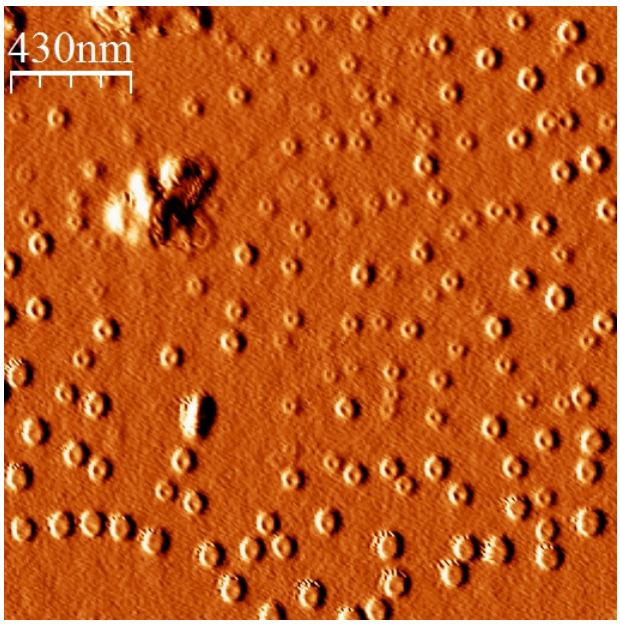


Figure S6. AFM image of toroid, obtained from **3** diluted solution.

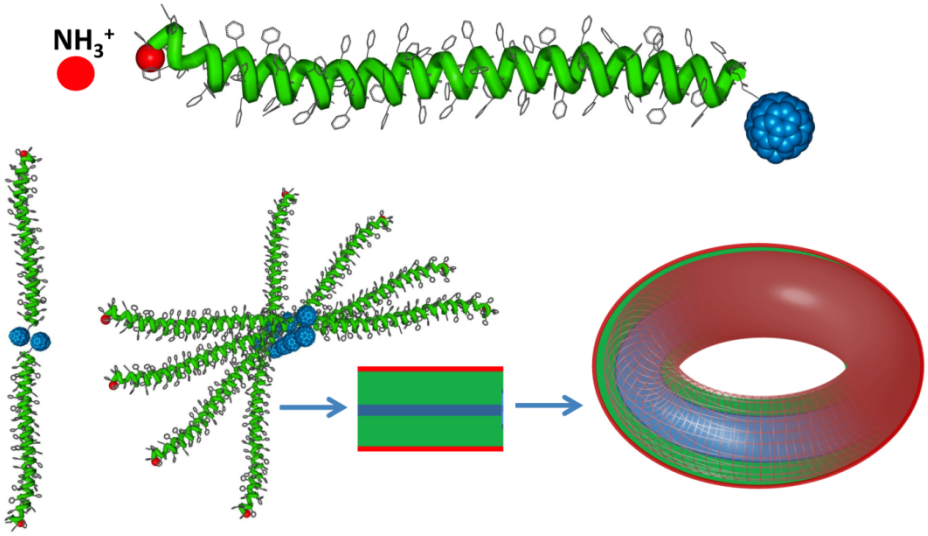


Figure S7. Top: representation of polymer **3** single chain (benzyl glutammate/fullerene 84/1, MW= 19359). The red balls represent the charged NH_3^+ groups, the blue parts represent the fullerene moieties and the green parts the main polymers helical backbone. Bottom: representation of the hypothetic mechanism for the self-assembly in ordered toroidal-like microstructures occurring for polymer **3**.

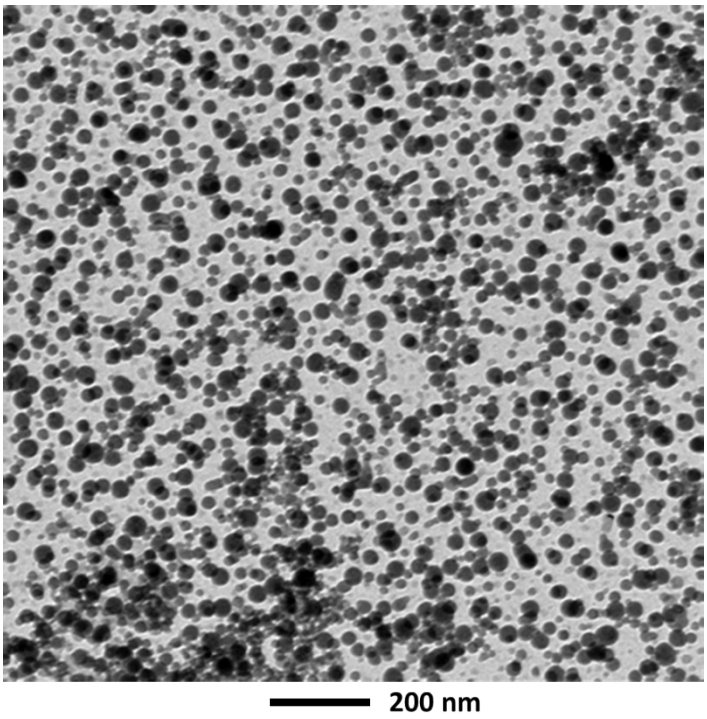


Figure S8. TEM image (unstained) of vesicles, obtained from **6**, shorter polymer.

Coarse-grained model of short-polymer 2

As underlined in the main text, for the simulations we considered as model molecule the short 25 BLG units polymer **6**, mainly to keep the required computational effort to reasonable levels. The atomistic structure of the polymer is shown in the left part of Figures S8a,c. Our modeling approach is based on a bottom-up protocol to build and parameterize the coarse-grained model of the polymer starting from full-atom molecular dynamics (FA-MD) simulations.

Full-atom simulations. Two FA-MD calculations in explicit water have been conducted. In the former a single polymer molecule in water has been simulated. In the latter, the calculation has been repeated for two polymer molecules in explicit water, to study their approaching and condensing process. Both trajectories have been carried out with the NAMD package.[1] The CHARMM27 force field with CMAP correction[2] was employed for the backbone and the first part of the side chain of the benzyl glutamic amino acid, while the standard general force field for organic moieties was employed for the benzyl group and the terminal moieties (C_{60}-S-CH_2-CH_2-NH- and C_{60}-S-CH_2-CH_2-CO-). Finally, the TIP3P force field was used for water.[3] In both simulations a cubic box of 8 nm of side length was built with 1 or 2 polymer molecules (placed in the middle) and 15733 or 15445 water molecules. Periodic boundary conditions were used. After energy minimization, slow heating from 0 to 293.15 K and 5 ns of equilibration, production runs of 50 ns in canonical NPT conditions (constant number of particles, pressure and temperature) were generated. Other simulation parameters were: 0.002 ps of integration time step, coordinates saved every 2 ps, 1.2 nm cutoff for non-bonded interactions (with function switching at 1.0 nm); pair list distance of 1.35 nm (pair list update every 0.02 ps); scaling of 1-4 interactions set to 1.0; particle mesh Ewald (PME) treatment of electrostatics. A Langevin thermostat at 293.15 K and a Langevin piston (target pressure 1 atm, piston period 0.2 ps, piston decay 0.1 ps) were coupled to the system to ensure maintenance of the canonical ensemble properties. Also, the SHAKE algorithm is used to fix all stretches of bonds with hydrogen atoms. AF-MD simulations were run over 1 hybrid compute node, allocating 8 CPU (AMD Opteron 6128 1.6 GHz) cores and 2 GPU (nVidia Tesla M2050) boards. All simulations have been run on the C3P ("Centro di Chimica Computazionale di Padova") high performance computing facility of our department (http://www.chimica.unipd.it/licc).

Parameterization of the coarse-grained model. The FA-MD calculations have been employed to parameterize the coarse-grained potential. We recall that our coarse-graining approach is based on representing each of the BLG units and the two terminal moieties (namely the C-terminal C_{60}-S-CH_2-CH_2-NH- and the N-terminal C_{60}-S-CH_2-CH_2-CO-) with a single bead each. Beads are located in the center of mass of the groups of atoms that they represent and the bead-bead interaction potential is assumed *a priori* to be partitioned into the canonical three bonded terms (bond, bond angle and torsion angle energies) plus a short-range dispersion interaction contribution modeled with a standard 12-6 Lennard-Jones potential (neither charges nor polarizability are included in the model), i.e.

$$E = \sum_{b \in bonds} k_b \left(b - b_0\right)^2 + \sum_{\theta \in angles} k_\theta \left(\theta - \theta_0\right)^2 + \sum_{\phi \in torsions} k_\phi \left[1 + \cos\left(n\phi - \phi_0\right)\right] + \sum_{i=1}^{N_{beads}} \sum_{j=i+1}^{N_{beads}} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \qquad (1)$$

where the first three (bonded) energy terms run respectively over all the bonds, bond angles and torsion angles, while the non-bonded interactions are calculated among all the beads but those connected with less then 4 bonds. The energy constants (k_b, k_θ, k_ϕ, ε) and the minimum energy parameters (b_0, θ_0, ϕ_0, σ) are the unknown parameters.

Let us firstly underline that our coarse-grained representation of the polymer is a non-branched chain of 27 adjacent beads of two types: *i*) the head and tail beads (A-type) represent the terminal groups and *ii*) the body beads (B-type) represent the BLG amino acids. Thus the set of parameters to be determined is limited to: the A-B and B-B bond, the A-B-B and B-B-B bond angle, the A-B-B-B and B-B-B-B torsion angle and the A and B Lennard-Jones parameters, i.e. a total of 16 quantities.

The force constants and minimum energy values of the bonded interactions have been parameterized using the single-polymer FA-MD simulation under the assumption that, since the structure shows high rigidity, variations of the internal non-bonded energy are negligible and thus the intra-molecular partition function is approximately factorized into contributions from the single internal coordinates. As a consequence, the Boltzmann equilibrium distribution calculated from the FA-MD trajectory for any of the internal coordinates (bond, bond angle or torsion angle) connecting the beads can be compared with the Boltzmann probability density defined analytically given the shape of the energy in eq. 1. From this comparison, the force constants (k_b, k_θ, k_ϕ) and the minimum energy values (b_0, θ_0, ϕ_0) can be obtained by means of a simple non-linear fitting procedure. Thus, as an example, the bond energy parameters for two beads *i* and *j* are obtained by the comparison

$$\exp\left[-\beta k_{b,ij}\left(b_{ij} - b_{0,ij}\right)^2\right] = Z_{FA-MD}\left(b_{ij}\right) \propto \int d\mathbf{x} \exp\left[-\beta E_{FA}\left(\mathbf{x}\right)\right] \delta\left(\hat{b}_{ij}\left(\mathbf{x}\right) - b_{ij}\right) \qquad (2)$$

where $1/\beta = k_B T$, with k_B the Boltzmann constant and T the absolute temperature, b_{ij} is the distance (bond length) between the two beads, Z_{AF-MD} is the reduced partition function extracted from the FA-MD simulation which, under the assumption of ergodicity, corresponds to the definition given in the rightmost equality of eq. 2, in which E_{FA} is the full-atom energy landscape, $\mathbf{x}$ is the array of the coordinates of all the atoms and $\hat{b}_{ij}(\mathbf{x})$ is the map that projects the atoms coordinates into the distance between beads *i* and *j*. The reduced partition function Z_{AF-MD} is obtained from the histogram analysis of the bead-bead distance time series extracted from the full-atom trajectory and the parameters $k_{b,ij}$ and $b_{0,ij}$ are obtained by fitting the partition function with the analytical expression in the leftmost part of eq. 2.

This procedure ensures, within the range of validity of the approximations made, that the statistical properties of the coarse-grained model resemble those of the full-atom model.[4] The fitting process briefly described above has been repeated for all the bonded interactions among the beads and average values were taken for similar couples (bonds), triplets (bond angles) or quartets (torsion angles) of beads. The full set of bonded parameters is reported in Table S1. We must underline here that the interactions with water molecules are also averaged

out when calculating the Z_{AF-MD} functions. Thus, for consistency, the water should not enter into the coarse-grained molecular dynamics (CG-MD) calculations. However, it proofed to be important to consider explicitly the presence of water in the simulation of the self-assembly, since the process is driven by the low affinity of water with the polymer. As we consider that the water scarcely interacts with the polymer, we shall assume that adding explicit water molecules in the CG simulation will have a negligible effect on the reduced partition functions of the internal degrees of freedom of the polymer.

The polymer-polymer non-bonded interactions (i.e. the parameters of the Lennard-Jones potential) have been evaluated using the two-polymer FA-MD simulation. In particular, taking the last snapshot of the trajectory, in which the two polymer molecules are condensed, as the equilibrium (minimum energy) configuration of the system, we searched for the set of the Lennard-Jones parameters such that the derivative of the energy in that configuration is (as close as possible to) zero. We tested the parameters of the non-bonded energy term obtained from the fitting procedure by producing a CG-MD simulation of the two polymer molecules starting from the same initial configuration of the FA-MD calculation. We checked that the condensing time and the final configuration of the CG system resembled those observed in the full-atom simulation. A comparison between the full-atom and coarse-grained dimers is shown in Figures S8b,d. Parameters are reported in Table S1. The phenyl-phenyl potential well was found to be 12 kJ/mol deep. This value is compatible with the stacking energy of toluene dimers calculated at quantum mechanical level.[5] On the other hand, for the fullerene-fullerene interaction while the minimum energy distance is similar to that found from the a simplified coarse-graining introduced by Girifalco[6] the potential depth is much smaller then that found in ref. 6 probably due to the fact that the averaging over an hypothetical sphere circumscribing the fullerene is not reasonable due to the steric hindrance of the polymer and the reduced mobility of the fullerene with respect to the polymer body. Since, in our modelling, BLG-C60 and C60-C60 interactions were found to be weaker than BLG-BLG interactions, it is expected that in the assembly process the C60 moieties will be localized mostly on the boundaries of the layer. This also explains, in part, why a limit on the thickness of the layer has been observed in the simulations.

To complete the modeling, as underlined in the main text, water should be explicitly taken into consideration. Since water is quite dense, a large number of molecules is required to fill the simulation box and this number may soon become prohibitive as the box side length approaches the hundred of nanometers. Thus, we used the 4-water CG model in which water is represented by beads interacting by means of a Lennard-Jones potential and each bead represents a cluster of 4 water molecules.[7] Thus, the required number of beads is four times smaller than the number of full-atom, single-molecule water models. Water-water and water-polymer interactions have been parameterized taking as reference the interaction table in ref. 7, in which the 4-water model is applied to the simulation of phospholipids vesicles. Since, differently to the application to phospholipids, we have only polar (water) and non-polar (polymer) beads, we consider the two limiting cases in the interaction table. The values of the parameters are reported in Table S1.

A comment is in order. The coarse-grained model here adopted is able to well reproduce the experimental observations for the short polymer with 25 BLG amino acids (see the results presented in the main text). Since the parametrization has been conducted without using the experiments to bias the modeling, we can conjecture

that our simulations could be predictive also for longer polymers Simulations of the 85-units polymer, however, showedcomplications to be faced mainly due tothe fact that: (i) larger simulation boxes are required implying the need of a prohibitive number of water molecules to be included; (ii) since the rigidity of the backbone is not explicitly modeled, it could happen that long polymers can show an unphysical bending. A modelling targeted to face these issues is in order.

Coarse-grained simulations. Once the CG force filed has been fully parameterized the CG-MD simulation was carried out using the Gromacs 4.6.3 software package.[8] A box of 62 nm of side length containing 1600 molecules of the coarse-grained 27-bead polymer and 1600608 4-water beads was created using the PACKMOL package.[9] The polymer molecules are placed and oriented randomly in the sub-box of 52 nm of side length to let a 5 nm of water layer from the box boundaries. The minimum distance between two beads of two different polymers was set to 1.5 nm, while the minimum distance between a water bead and any other bead was set to 0.5 nm. After energy minimization a 5 ns simulation (1 ns of equilibration + 4 ns of production) was calculated employing periodic boundary conditions and under constant particles, volume and temperature (NVT) canonical conditions. Parameters of the simulation were: 0.01 ps of integration time step, coordinates saved every 20 ps, 2.8 nm cutoff for non-bonded interactions (without function); pair list distance of 3.0 nm (pair list update every 0.2 ps); scaling of 1-4 interactions set to 1.0. A Langevin thermostat at 293.15 K with correlation time of 0.1 was coupled to the system to ensure maintenance of the canonical ensemble properties. The CG-MD calculation was run over 3 hybrid nodes, allocating a total of 12 CPU (AMD Opteron 6128 1.6 GHz) cores and 6 GPU (nVidia Tesla M2050) boards.

References

1.J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781-1802.

2. A. D. Jr. MacKerell, D. Bashford, M. Bellott, R. L. Jr. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, S.; T. Ngo, D. T. Nguyen, B. Prodhom, W. E. III, Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586-3616.

3. W. L. Jorgensen, J. D. Madura, *J. Am. Chem. Soc.* **1983**, *105*, 1407-1413.

4. W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, H. C. Andersen, *J. Chem. Phys* .**2008**, *128*, 244114.

5. D. M. Rogers, J. D. Hirst, E. P. F. Lee, T. G. Wright, *Chem. Phys. Lett.* **2006**, *31*, 410-413.

6. L. A. Girifalco, *J. Phys. Chem.* **1992**, *96*, 858-861.

7. S. J. Marrink, A. E. Mark, *J. Am. Chem. Soc.* **2003**, *125*, 15233-15242.

8. B. Hess, C. Kutzner, D. van der Spoel, E. Lindhal, *J. Chem. Theory Comput.* **2008**, *4*, 435-447.

9. L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, *J. Comp. Chem.* **2009**, *30*, 2157-2164.

Bond interaction $k_b(b-b_0)^2$		
	b_0 / nm	k_b / kJ mol^{-1} nm^{-2}
AB	1.05	569
B-B	0.754	658
Bond angle interaction $k_\theta(\theta-\theta_0)^2$		
	θ_0 / deg	k_θ / kJ mol^{-1} rad^{-2}
A-B-B	94.8	163
B-B-B	82.3	216
Torsion angle interaction $k_\phi[1+\cos(\phi-\phi_0)]$		
	ϕ_0 / deg	k_ϕ / kJ mol^{-1}
A-B-B-B	-11.3	290
B-B-B-B	-25.0	430
Non-bonded interaction $4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^{6}\right]$		
	σ / nm	ε / kJ mol^{-1}
A, A	0.900	1.07
B, B	0.650	12.0
A, B	0.775	3.59
W, W	0.470	5.00
A, W	0.685	1.80
B, W	0.560	1.80

Table S1. Parameterization of the coarse-grained force field. A and B bead types represent, respectively, the terminal parts of the polymer and the BLG unit, while the W bead is the coarse-grained 4-water model. The quantities $\sigma_{W,W}$, $\varepsilon_{W,W}$, $\varepsilon_{A,W}$, and $\varepsilon_{B,W}$ are taken from ref. SI_MD_R7, while $\sigma_{A,B}$, $\varepsilon_{A,B}$, $\sigma_{A,W}$, and $\sigma_{B,W}$ have been calculated using the rules $\sigma_{i,j}=(\sigma_{i,i}+\sigma_{j,j})/2$ and $\varepsilon_{i,j}=\sqrt{\varepsilon_{i,i}\varepsilon_{j,j}}$.

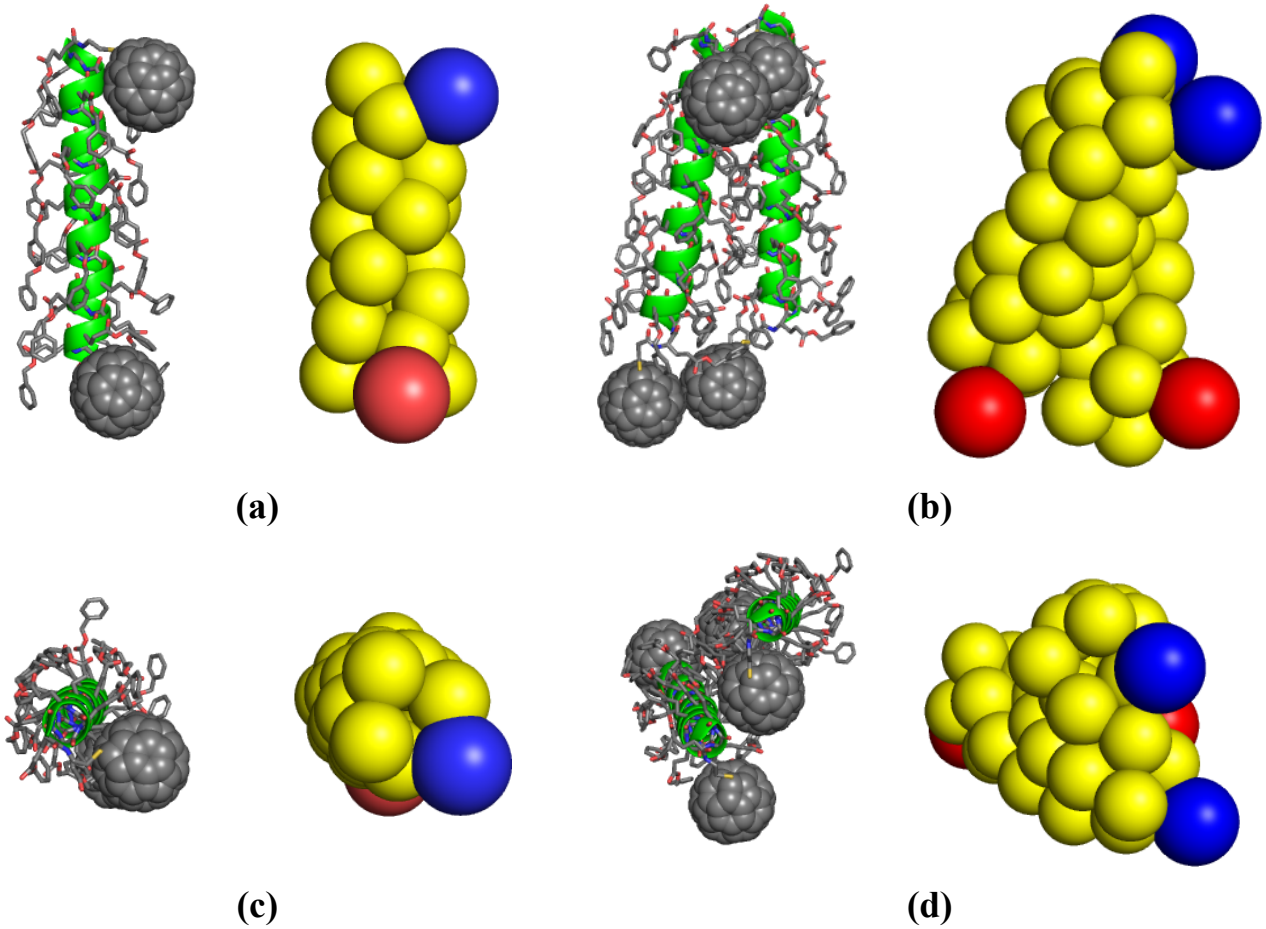
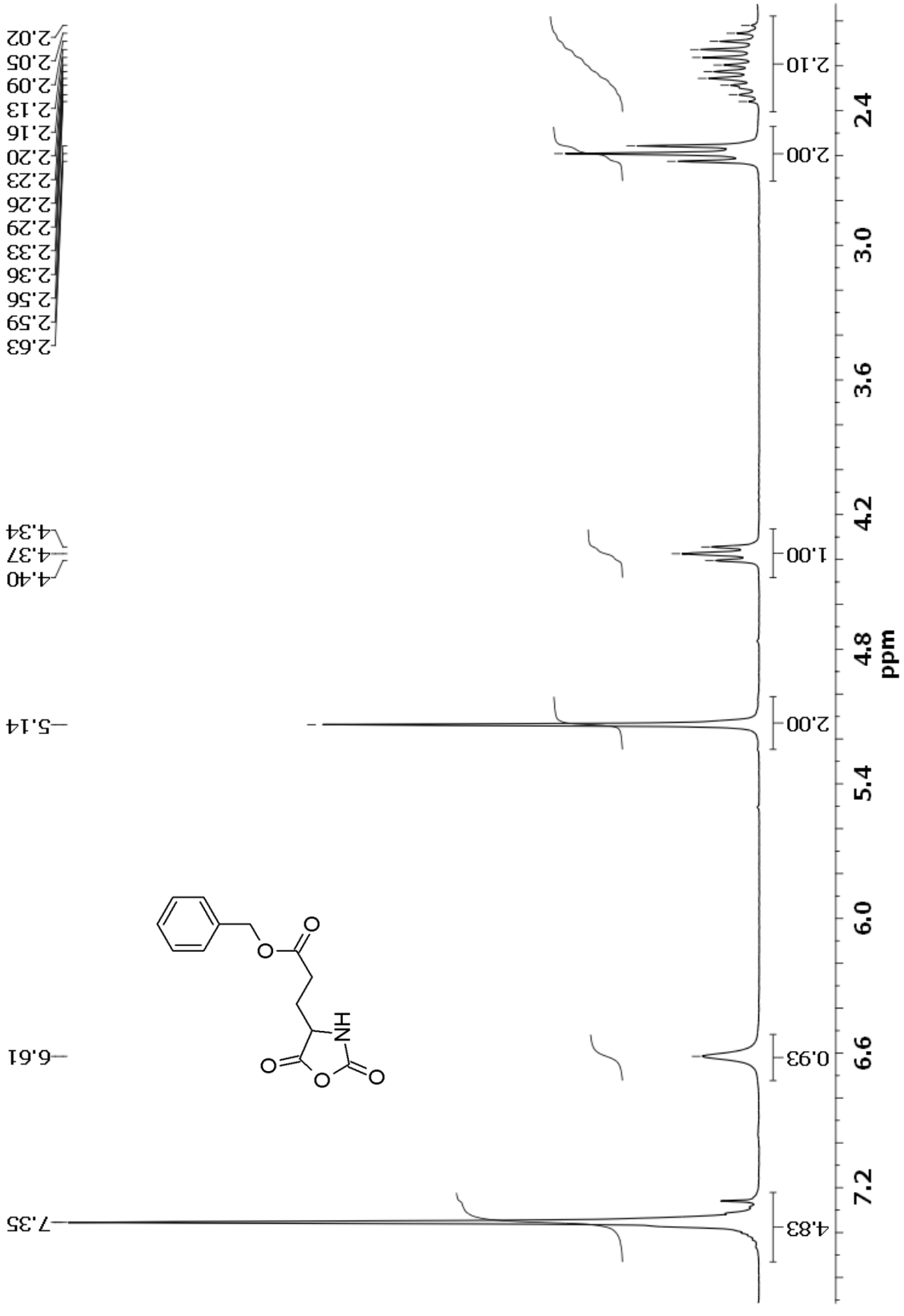
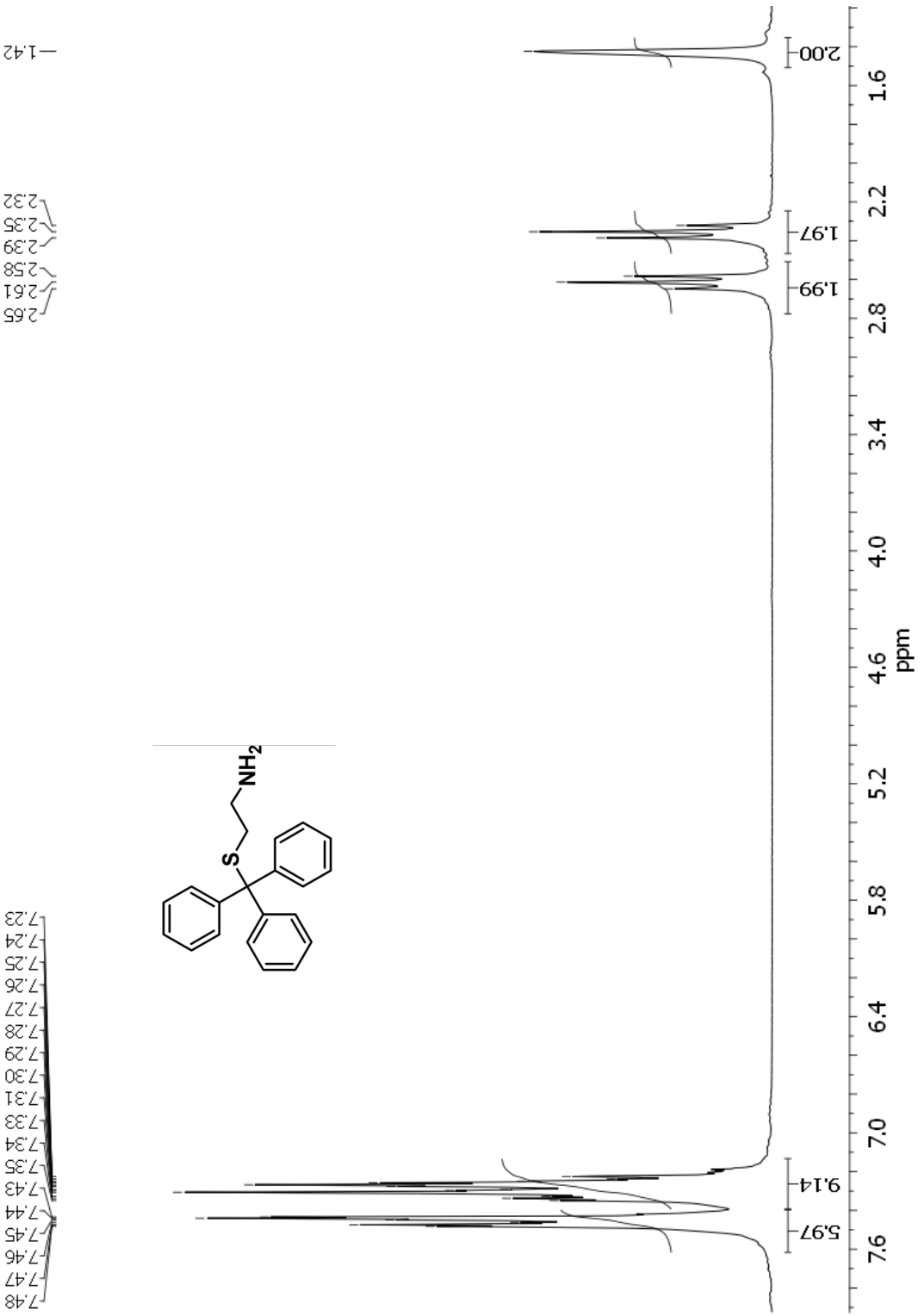


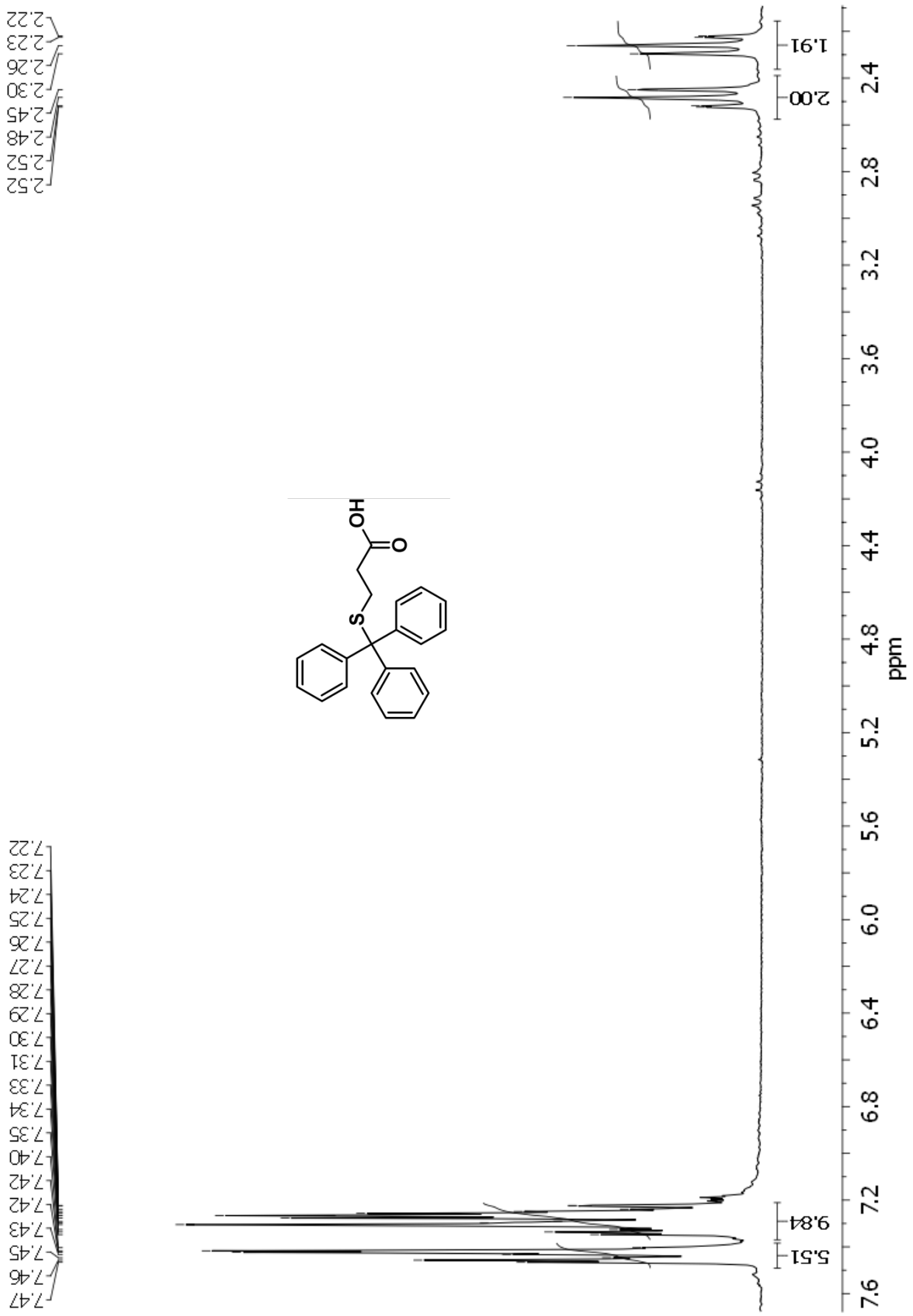
Figure S9. Comparison of full-atom and coarse-grained representations of polymer **6** with 25 BLG amino acids. The (a) front and (c) top views of a single molecule are shown on the left. The (b) front and (d) top views of a dimer are shown on the right. The molecular configurations are taken from the last snapshot of the full-atom and coarse-grained molecular simulations.



^{1}H-NMR spectrum of BLG-NCA in $CDCl_3$.



^{1}H-NMR spectrum of 2-(tritylsulfanyl)ethanamine in $CDCl_3$.



^{1}H-NMR spectrum of 3-(tritylsulfanyl)propanoic acid in $CDCl_3$.