

## Supplementary Information for

Synthesis of polypeptides via bioinspired polymerization of in situ purified *N*-carboxyanhydrides

Ziyuan Song, Hailin Fu, Jiang Wang, Jingshu Hui, Tianrui Xue, Lazaro A. Pacheco, Haoyuan Yan, Ryan Baumgartner, Zhiyu Wang, Yingchun Xia, Xuefang Wang, Lichen Yin, Chongyi Chen, Joaquín Rodríguez-López, Andrew L. Ferguson, Yao Lin\*, and Jianjun Cheng\*

Email: jianjunc@illinois.edu (J.C.); yao.lin@uconn.edu (Y.L.)

### **This PDF file includes:**

Supplementary text: methods for experiments, simulation, and kinetic modeling  
Figs. S1 to S21  
Captions for movies S1 to S2  
References for SI reference citations

### **Other supplementary materials for this manuscript include the following:**

Movies S1 to S2

## 1. General Experimental Section

**Materials.** All chemicals were purchased from MilliporeSigma (St. Louis, MO, USA) unless otherwise specified. Amino acids were purchased from Chem-Impex International Inc. (Wood Dale, IL, USA). Methoxy poly(ethylene glycol) amine (PEG, 5 kDa) was purchased from Laysan Bio, Inc. (Arab, AL, USA). The end-group fidelity of PEG (*i.e.*, primary amino group at chain terminus) was determined to be 58% using a previously reported method (1). Deuterated solvents were purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA, USA). Anhydrous tetrahydrofuran (THF) and hexane were dried by a column packed with alumina. Anhydrous *N,N*-dimethylformamide (DMF) were treated with polymer-bound isocyanates (MilliporeSigma, St. Louis, MO, USA) to remove any amino residues. Anhydrous dichloromethane (DCM) were stored over 3 Å molecular sieves in a freezer. Poly(norbornene)s with pendent *N*-trimethylsilyl amine groups (PNB, degree of polymerization (DP) = 100) were prepared according to previous literature procedures (2). Aqueous buffer with wide pH ranges were prepared according to previous literature procedures (3), which contains H<sub>3</sub>BO<sub>3</sub>, citric acid, and H<sub>3</sub>PO<sub>4</sub> as active components.

**Instrumentation.** Proton and carbon nuclear magnetic resonance (NMR) spectra were recorded on a Varian U500 or VNS750NB spectrometer in the NMR laboratory, University of Illinois. Chemical shifts were reported in ppm and referenced to the residual protons in the deuterated solvents. MestReNova software (version 8.1.1, Mestrelab Research, Escondido, CA, USA) was used for all NMR analysis. Gel permeation chromatography (GPC) experiments were performed on a system equipped with an isocratic pump (1260 Infinity II, Agilent, Santa Clara, CA, USA), a multi-angle static light scattering (MALS) detector (DAWN HELEOS-II, Wyatt Technology, Santa Barbara, CA, USA), and a differential refractometer (DRI) detector (Optilab T-rEX, Wyatt Technology, Santa Barbara, CA, USA). The detection wavelength of HELEOS was set at 658 nm. Separations were performed using serially connected size exclusion columns (three PLgel MIXED-B columns, 10 µm, 7.5 × 300 mm, Agilent, Santa Clara, CA, USA) at 40 °C using DMF containing 0.1 mol/L LiBr as the mobile phase. The MALS detector was calibrated using pure toluene and can be used for the determination of the absolute molecular weights (MWs). The MWs of polymers were determined based on the *dn/dc* value of each polymer sample calculated offline by using the internal calibration system processed by the ASTRA 7 software (version 7.1.3.15, Wyatt Technology, Santa Barbara, CA, USA). Fourier transform infrared (FTIR) spectra were recorded on a Perkin Elmer 100 serial FTIR spectrophotometer (PerkinElmer, Santa Clara, CA, USA) calibrated with polystyrene film. Circular dichroism (CD) measurements were carried out on a JASCO J-815 CD spectrometer (JASCO, Easton, MD, USA). The mean residue molar ellipticity of each polypeptide was calculated on the basis of the measured apparent ellipticity following the literature-reported formulas: Ellipticity ( $[\theta]$  in deg cm<sup>2</sup> dmol<sup>-1</sup>) = (millidegrees × mean residue weight)/(path length in millimetres × concentration of polypeptide in mg mL<sup>-1</sup>) (4, 5). High performance liquid chromatography (HPLC) was performed on a Shimadzu LC system (Shimadzu, Columbia, MD, USA) with a pump (LC-20AT), a photodiode array (PDA) detector (SPD-M20A), and an LC column (Eclipse plus C18, 3.5 µm, 4.6 × 100 mm, Agilent, Santa Clara, CA, USA). Gradient method was adopted using 0.1 % TFA-H<sub>2</sub>O and acetonitrile as mobile phase. Electron ionization (EI) and matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectra were collected in the mass spectrometry laboratory, University of Illinois. MALDI-TOF spectra were obtained on a Bruker ultrafleXtreme with  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) as the matrix. CHN analysis and chlorine analysis were performed in the microanalysis laboratory, University of Illinois. Emulsification was done with a probe sonicator (Fisherbrand, Model 705, Thermo Fisher Scientific, Waltham, MA, USA).

## 2. Accelerated Polymerization with Interfacially Anchored Initiators

**Water-sensitivity of NCA molecules.** The water-sensitivity of NCA molecules were demonstrated in both a water/DCM biphasic system and a water/THF solution.

Purified BLG NCAs were dissolved in DCM containing aqueous buffer (50 mM, 1 wt%, pH = 7.0), and the resulting biphasic mixture was vigorously stirred at room temperature. At different time intervals, an aliquot of mixture (~ 10  $\mu$ L) was taken out, dried on a KBr salt plate, and characterized by FTIR. The decrease in anhydride peaks (1857 and 1786  $\text{cm}^{-1}$ ) from NCAs and the increase in amide peaks (1652  $\text{cm}^{-1}$ ) from oligopeptides/polypeptides evidenced the oligomerization/polymerization of NCAs (Fig. S1B).

In a separate experiment, purified BLG NCAs were dissolved in THF:water mixture (50  $\mu$ g/mL, 1:1, vol/vol, pH = 7.0). At different time intervals, an aliquot of mixture was taken out for HPLC-UV analysis ( $\lambda = 210$  nm). The increase in amino acid peaks (2.1 min) and oligopeptide peaks (3.3 min) demonstrated the degradation of NCAs (Fig. S1C).

**Preparation of macroinitiators.** Methoxy poly(ethylene glycol)-*b*-poly( $\gamma$ -benzyl-L-glutamate) amine (PEG-PBLG) macroinitiators were prepared by polymerizing purified BLG NCAs with a PEG initiator (Fig. S2A). In a glovebox, PEG (5 kDa, 38.0 mg, 0.0076 mmol) and purified BLG NCA (60 mg, 0.23 mmol) were dissolved in anhydrous DMF (800  $\mu$ L). The mixture was incubated at 4  $^{\circ}$ C for 60 h. FTIR characterization confirmed the complete conversion of BLG NCA. After purification by precipitation in cold hexane/ether (1:1, vol/vol), the obtained macroinitiators were re-dissolved in DCM (100 mg/mL) and stored at -30  $^{\circ}$ C. The MW of PEG-PBLG was determined with GPC (Fig. S2B). The stock solution of PEG-PBLG macroinitiators was used within one week to avoid the degradation of terminal amino groups (*i.e.*, terminal amino groups react with the side-chain esters to form amide groups (6), which no longer initiate NCA polymerization). PBLG macroinitiators were prepared in a similar way of PEG-PBLG, but using hexylamine as the initiators (Fig. S2A). In addition to GPC characterization, the end-group fidelity of PBLG macroinitiator was confirmed by MALDI-TOF (Fig. S2C).

The DCM solution of both PEG-PBLG and PBLG macroinitiators were characterized by FTIR and CD, which indicated an  $\alpha$ -helical conformation for the polypeptide segments of both macroinitiators (Fig. S2D-E) (7).

**Polymerization of purified NCAs in a w/o emulsion.** For polymerizations in a w/o emulsion, macroinitiators also behave as emulsion stabilizers (*e.g.*, PEG-PBLG, PEG, and PBLG). When no macroinitiators were added (*i.e.*, to evaluate the water-induced side reactions, Fig. S5), *N*-acetyl methoxy poly(ethylene glycol) amine (mPEG-NHAc) was used as an inert emulsion stabilizer to ensure comparable size of water droplets (See Fig. S3), which were found to significantly alter the kinetics of water-induced side reactions. mPEG-NHAc was prepared by reacting PEG (5 kDa) with acetic anhydrides (20 equiv.) in DCM overnight and purified through precipitation in ether. All polymerizations were conducted in regular DCM without anhydrous setups.

Typically, the stock solution of PEG-PBLG macroinitiators (100 mg/mL, 59.5  $\mu$ L) was diluted with DCM (340  $\mu$ L), into which the aqueous buffer (pH = 7.0, 10.6  $\mu$ L) was added. The resulting mixture was emulsified with a probe sonicator using a pulse sequence (13 W, 1 s pulse on and 1 s

pulse off, total pulse on time = 10 s). An aliquot of the emulsion (257  $\mu\text{L}$ ) was then mixed with DCM solution of purified NCAs (250  $\mu\text{L}$ , 0.1 M) to start the polymerization. Final condition:  $[\text{M}]_0 = 50 \text{ mM}$ ,  $[\text{I}]_0 = 0.5 \text{ mM}$ , water:DCM = 1:100, wt/wt.

Deuterated DCM was used instead of DCM in the polymerization kinetic studies with  $^1\text{H}$  NMR (Fig. S4). The conversion of NCA was calculated through the integrals of the ring N-H peak of NCA on  $^1\text{H}$  NMR spectra ( $\delta$  in the range of 6.4-7.1 ppm, depending on the side-chain structure and the concentration of NCAs). In order to obtain the integral of ring N-H peak at  $t = 0$  (100% remaining NCA), the NCA conversion in the first  $^1\text{H}$ -NMR spectrum (typically  $t = 2 \text{ min}$ ) was calculated based on the integral ratios of side-chain benzyl peaks between NCA and polypeptide ( $\delta$  in the range of 4.9-5.2 ppm, for BLG NCA and ZLL NCA), or the integral ratios of backbone  $\alpha$ -H peaks between NCA and polypeptide ( $\delta$  in the range of 3.8-4.5 ppm, for ELG NCA), which was then used to normalize the integral of the ring N-H peak in the first  $^1\text{H}$ -NMR spectrum. When PEG-PBLG and PBLG were used as the macroinitiators, the contribution of macroinitiators to the integral of side-chain benzyl peaks was subtracted before the calculation of conversion.

For the MW analysis of polymers, the emulsion was dried under vacuum and the residues were dissolved in DMF containing 0.1 M LiBr. The resulting solution was then filtered through a 0.45  $\mu\text{m}$  PTFE membrane (Thermo Fisher Scientific, Waltham, MA, USA) before injection into DMF GPC.

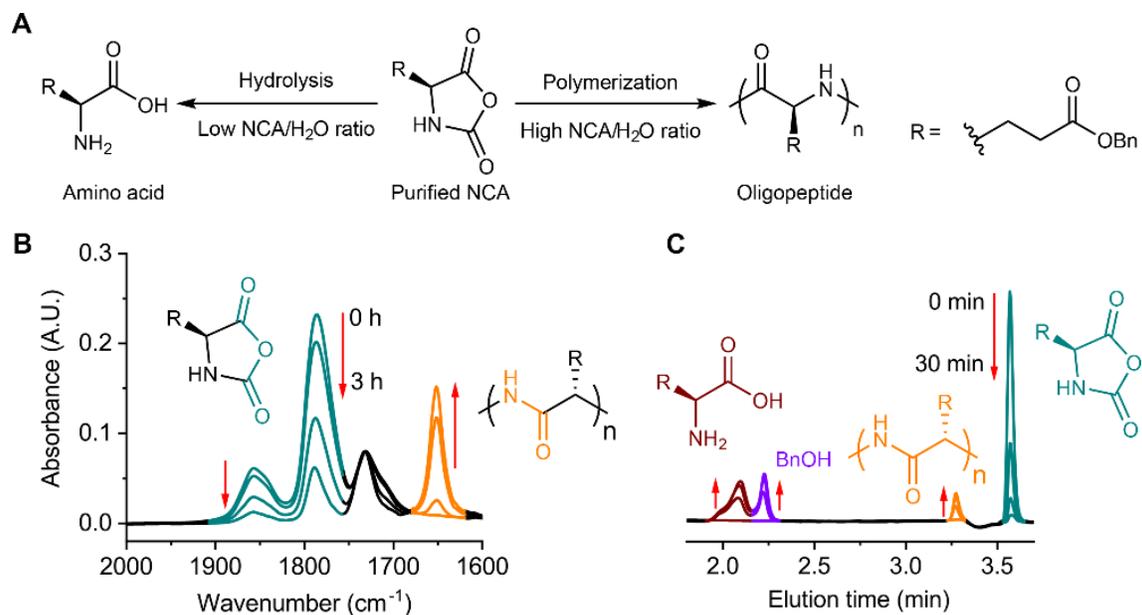
**Characterization of nonpurified NCAs.** All three nonpurified NCAs are characterized by NMR, HR EI-MS, FTIR, and elemental analysis. However, these routine characterization methods are difficult to analyze the impurities. Therefore, we used chlorine analysis to indicate the amount of impurities (results are represented as mean  $\pm$  s.d. from three independent analyses).

**BLG NCA:**  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ,  $\delta$ , 500 MHz): 7.42-7.32 (m, 5H, ArH), 6.27 (s, 1H, ring NH), 5.14 (s, 2H,  $-\text{CH}_2\text{Ar}$ ), 4.37 (t,  $J = 6.10 \text{ Hz}$ , 1H,  $\alpha$ -H), 2.61 (t,  $J = 6.41 \text{ Hz}$ , 2H,  $-\text{CH}_2\text{CH}_2\text{COO}-$ ), 2.34-2.07 (m, 2H,  $-\text{CH}_2\text{CH}_2\text{COO}-$ ).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ,  $\delta$ , 125 MHz): 172.5, 169.6, 152.3, 135.3, 128.8, 128.7, 128.4, 67.2, 57.0, 29.8, 26.9. **HRMS** (EI,  $m/z$ ):  $[\text{M}+\text{H}]^+$  Calcd. for  $\text{C}_{13}\text{H}_{14}\text{NO}_5$  264.0872. Found: 264.0877. **Anal.** (%): Calcd. for  $\text{C}_{13}\text{H}_{13}\text{NO}_5$ : C, 59.31; H, 4.98; N, 5.32. Found: C, 58.95; H, 4.82; N, 5.20. **Chlorine analysis** (%):  $0.45 \pm 0.04$ . **FTIR** ( $\text{cm}^{-1}$ ): 3310 (br), 1856, 1786, 1731.

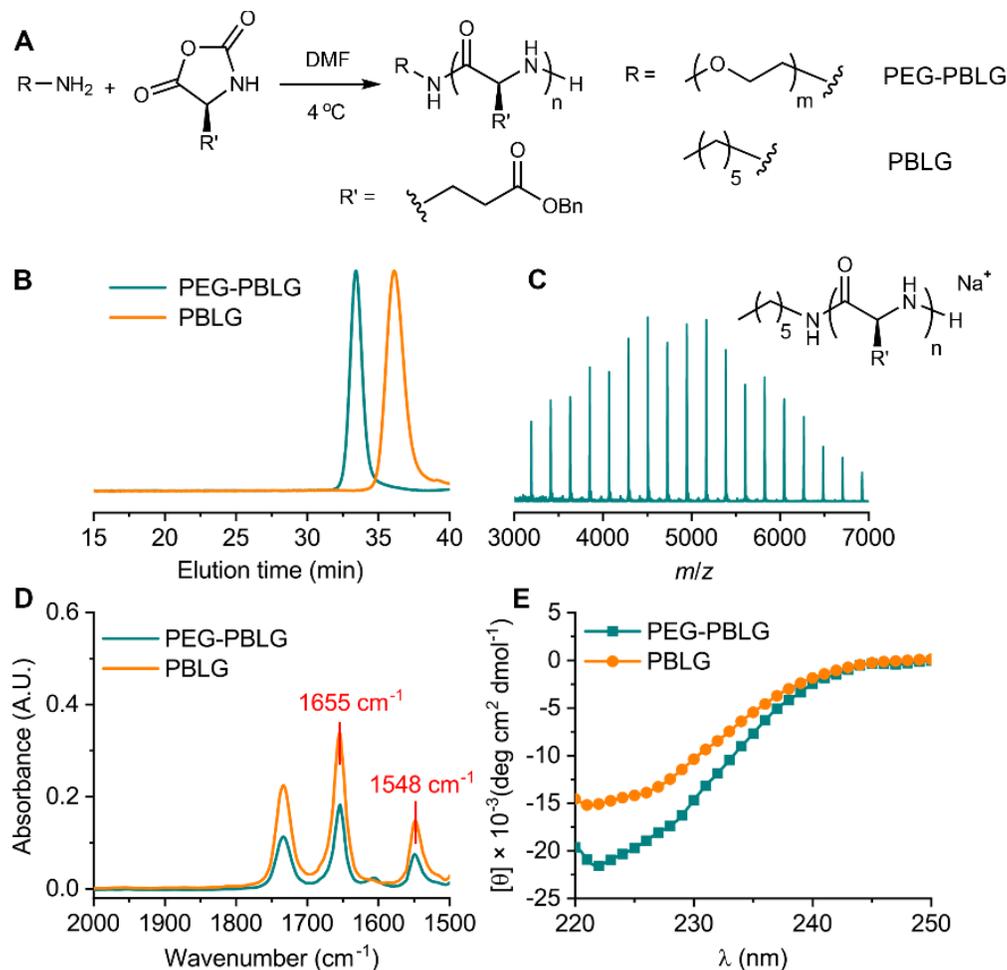
**ELG NCA:**  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ,  $\delta$ , 500 MHz): 6.59 (s, 1H, ring NH), 4.41 (t,  $J = 6.00 \text{ Hz}$ , 1H,  $\alpha$ -H), 4.16 (q,  $J = 7.14 \text{ Hz}$ , 2H,  $-\text{COOCH}_2\text{CH}_3$ ), 2.55 (t,  $J = 6.80 \text{ Hz}$ , 2H,  $-\text{CH}_2\text{CH}_2\text{COO}-$ ), 2.34-2.07 (m, 2H,  $-\text{CH}_2\text{CH}_2\text{COO}-$ ), 1.27 (t,  $J = 7.14 \text{ Hz}$ , 3H,  $-\text{COOCH}_2\text{CH}_3$ ).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ,  $\delta$ , 125 MHz): 172.7, 169.8, 152.5, 61.4, 57.0, 29.6, 26.9, 14.2. **HRMS** (EI,  $m/z$ ):  $[\text{M}+\text{H}]^+$  Calcd. for  $\text{C}_8\text{H}_{12}\text{NO}_5$  202.0715. Found: 202.0719. **Anal.** (%): Calcd. for  $\text{C}_8\text{H}_{11}\text{NO}_5$ : C, 47.76; H, 5.51; N, 6.96. Found: C, 47.69; H, 5.53; N, 6.71. **Chlorine analysis** (%):  $0.63 \pm 0.01$ . **FTIR** ( $\text{cm}^{-1}$ ): 3303 (br), 1857, 1786, 1729.

**ZLL NCA:**  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ,  $\delta$ , 500 MHz): 7.45-7.30 (m, 5H, ArH), 6.67 (s, 1H, ring NH), 5.11 (s, 2H,  $-\text{CH}_2\text{Ar}$ ), 4.89 (s, 1H,  $-(\text{CH}_2)_3\text{CH}_2\text{NH}-$ ), 4.28 (t,  $J = 5.81 \text{ Hz}$ , 1H,  $\alpha$ -H), 3.20 (m, 2H,  $-(\text{CH}_2)_3\text{CH}_2\text{NH}-$ ), 2.05-1.35 (m, 6H,  $-(\text{CH}_2)_3\text{CH}_2\text{NH}-$ ).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ,  $\delta$ , 125 MHz): 170.1, 157.0, 152.8, 136.5, 128.7, 128.3, 128.1, 67.0, 57.6, 40.3, 31.0, 29.2, 21.5. **HRMS** (EI,  $m/z$ ):

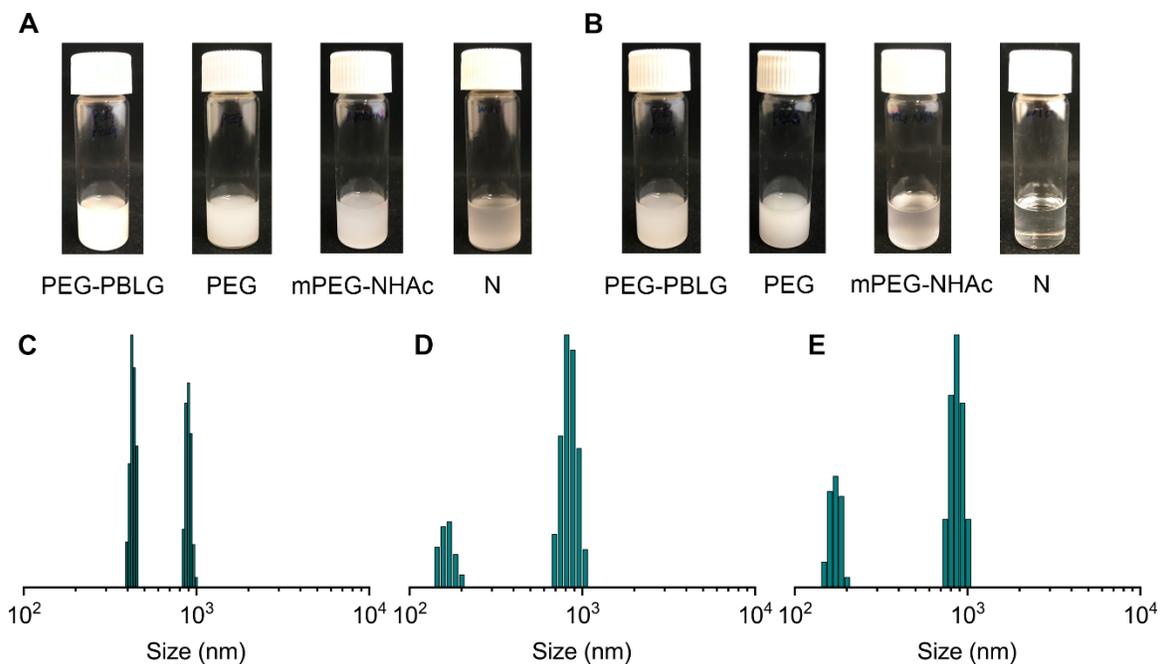
$[M+H]^+$  Calcd. for  $C_{15}H_{19}N_2O_5$  307.1294. Found: 307.1288. **Anal.** (%): Calcd. for  $C_{15}H_{18}N_2O_5$ : C, 58.82; H, 5.92; N, 9.15. Found: C, 58.37; H, 5.89; N, 9.06. **Chlorine analysis** (%):  $0.80 \pm 0.04$ . **FTIR** ( $cm^{-1}$ ): 3331 (br), 1855, 1786, 1695, 1531.



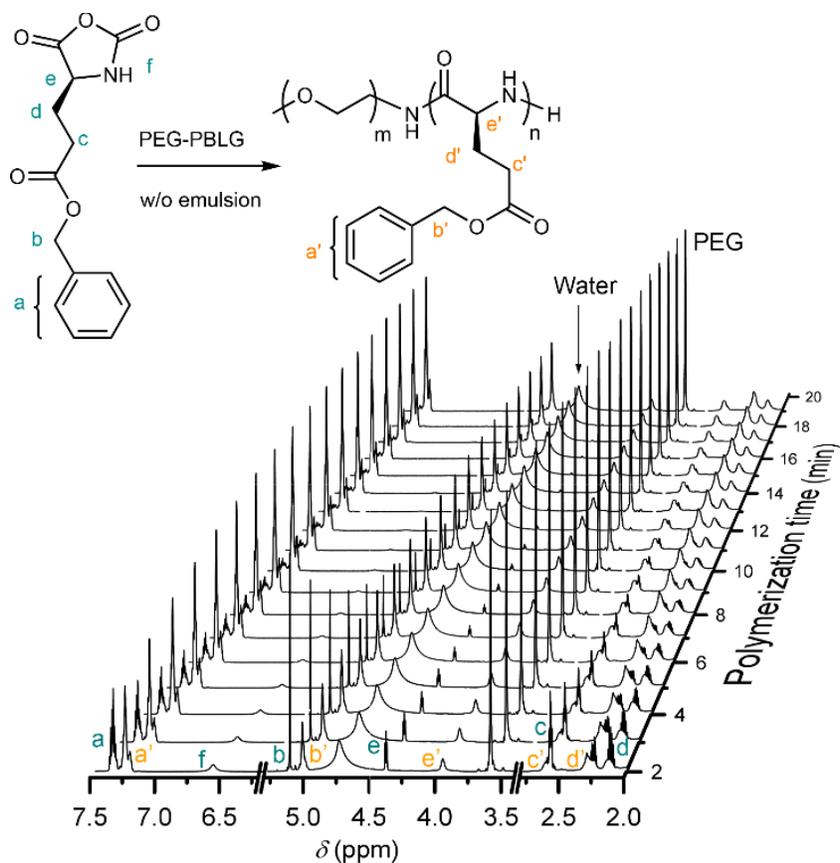
**Fig. S1.** Water-induced degradation of NCA molecules. (A) Scheme illustrating the degradation of NCAs in the presence of water. (B) FTIR spectra showing water-induced oligomerization or polymerization of purified BLG NCAs in water/DCM mixture (1 wt% water). [NCA]<sub>0</sub> = 50 mM. (C) HPLC-UV traces showing water-induced degradation of purified BLG NCAs in THF/water co-solvent (1:1, vol/vol). [NCA]<sub>0</sub> = 50 μg/mL.



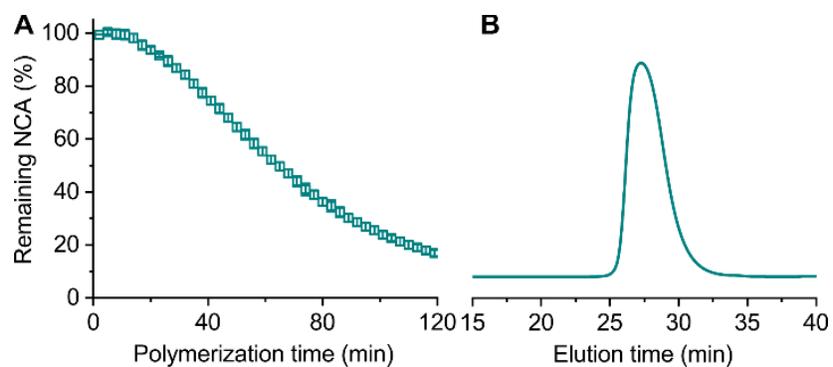
**Fig. S2.** Characterization of macroinitiators. (A) Synthetic routes to PEG–PBLG and PBLG macroinitiators. (B) Normalized GPC-LS traces of macroinitiators. PEG–PBLG:  $M_n = 14.8$  kDa,  $D = 1.06$ ; PBLG:  $M_n = 7.5$  kDa,  $D = 1.06$ . (C) MALDI-TOF spectrum of PBLG macroinitiator. The peaks agree well with the calculated value ( $124.11 + 219.09n$ ), which corresponds to  $[M+Na]^+$  form of PBLG bearing amino terminus with negligible degradation. (D) FTIR spectra of macroinitiators in DCM (10 mg/mL). The characteristic amide I ( $1655\text{ cm}^{-1}$ ) and amide II ( $1548\text{ cm}^{-1}$ ) peaks indicate an  $\alpha$ -helical conformation for both macroinitiators. (E) CD spectra of macroinitiators in DCM (0.5 mg/mL). The characteristic CD curve confirmed the  $\alpha$ -helical structure of both macroinitiators.



**Fig. S3.** Characterization of w/o emulsions containing macroinitiators. (*A* and *B*) Photographs of w/o emulsions 1 min (*A*) and 4 h (*B*) after emulsification. N represents w/o emulsion with no added macroinitiators.  $[I]_0 = 0.5$  mM, water:DCM = 1:100 (wt/wt). (*C-E*) Normalized DLS results (size vs intensity) of w/o emulsions containing PEG-PBLG (*C*), PEG (*D*), and mPEG-NHAc (*E*).  $[I]_0 = 0.17$  mM, water:DCM = 1:100 (wt/wt). The polymer mPEG-NHAc was used as an inert stabilizer to study the water-induced polymerization. The resulting emulsion has comparable property with those containing PEG-PBLG and PEG.



**Fig. S4.** Representative  $^1\text{H}$  NMR spectra showing the polymerization kinetics. Polymerization was conducted in  $\text{CD}_2\text{Cl}_2$  (750 MHz). The rapid decrease in peak intensity of NCA protons (labelled in dark cyan) and the increase in peak intensity of PBLG protons (labelled in orange) suggest fast polymerization kinetics.  $[\text{M}]_0 = 50 \text{ mM}$ ,  $[\text{I}]_0 = 0.5 \text{ mM}$ , water: $\text{CD}_2\text{Cl}_2 = 1:100$  (wt/wt).



**Fig. S5.** Water-induced NCA polymerization in a w/o emulsion. (A) Conversion of purified BLG NCA in a w/o emulsion stabilized by inert mPEG–NHAc emulsifier.  $[M]_0 = 50$  mM, water:DCM = 1:100 (wt/wt). Results represent means  $\pm$  s.d. of three replicates. (B) GPC-LS traces of resulting polymers from water-induced polymerization.  $M_n = 149.7$  kDa,  $D = 1.47$ .

### 3. Molecular Dynamics Simulation

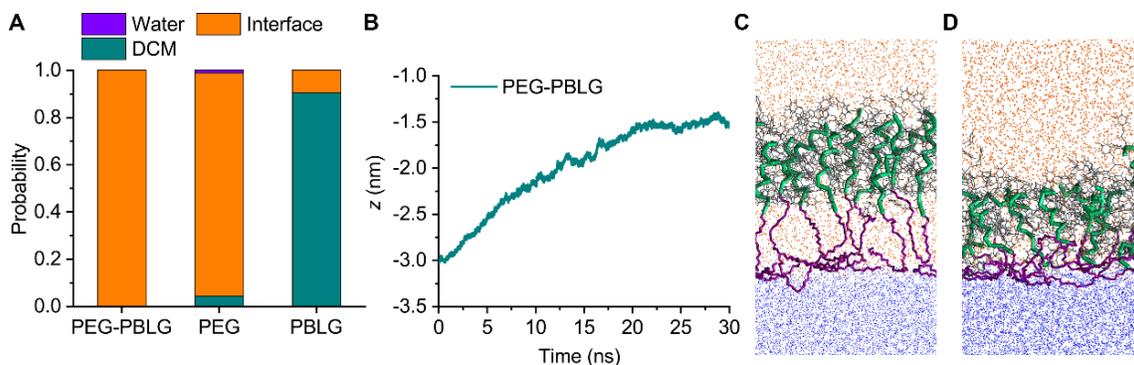
The interfacial behaviors of PEG–PBLG (3.9 kDa, DP = 11 and 15 for PEG and PBLG block, respectively), PEG (0.55 kDa, DP = 11), and PBLG (3.4 kDa, DP = 15) were studied. Due to the limitation of computational power, shorter lengths of PEG and PBLG segments were used compared with the experimental setup (which has a DP of 113 and 45 for PEG and PBLG segment, respectively).

All-atom molecular models of PEG–PBLG, PEG, PBLG, and DCM solvent were generated using the Automated Topology Builder (ATB) server (<http://atb.uq.edu.au>) (8) and modeled using the GROMOS 54A7 force field (9). Atomic partial charges were assigned using semiempirical quantum mechanical calculations conducted using the MOPAC method (10) and all molecules carried zero net charge.

Molecular dynamics simulations were conducted using the GROMACS 4.6 simulation suite (11). Lennard-Jones interactions were shifted smoothly to zero at 1.4 nm, and interactions between unlike atoms specified by Lorentz-Berthelot combining rules (12). Coulomb interactions were treated by Particle Mesh Ewald (PME) with a real-space cutoff of 1.4 nm and a 0.12 nm reciprocal-space grid spacing (13). Bond lengths were fixed to their equilibrium values using the LINCS algorithm (14). Temperature was maintained at 300 K using a Nosé-Hoover thermostat (15) and pressure at 1.0 bar using an isotropic Parrinello–Rahman barostat (16). Newton's equations of motion were integrated using the leap-frog algorithm with time step of 2 fs (17). System configurations were saved for analysis every 2 ps. Calculations were conducted on NVIDIA Quadro K1200 GPU cards achieving execution speeds of about 30 ns/day. Simulation trajectories were visualized using Visual Molecular Dynamics (VMD) (18).

Individual PBLG and PEG–PBLG molecules were placed in a  $5 \times 5 \times 5$  nm<sup>3</sup> box of DCM solvent and subjected to steepest descent energy minimization until the maximum force on any given atom was less than a threshold of 10 kJ/mol·nm. Atomic velocities were initialized from a Maxwell distribution at 300 K, and the systems then simulated over the course of 10 ns until equilibrated structures were obtained: an  $\alpha$ -helix for PBLG and a coil for the PEG. These structures served as the initial configurations for the umbrella sampling calculations.

Umbrella sampling of passage of the molecules across the water/DCM interface were conducted in a box of size of  $5 \times 5 \times 20$  nm<sup>3</sup>, where the length of the box in  $z$  direction is 20 nm. DCM solvent fills up the portion with  $z < 10$ nm, and H<sub>2</sub>O the portion where  $z > 10$ nm. Fifty individual umbrella sampling runs on the  $z$ -coordinate of the molecular center of mass (COM) were conducted over the range  $z = (5 \text{ nm}, 15 \text{ nm})$  with umbrella windows placed uniformly at 0.2 nm intervals and harmonic restraining potentials with force constants of 1000 kJ/mol·nm<sup>2</sup> placed at the center of each window (19). Each umbrella simulation was conducted for 10 ns, and the unbiased potential of mean force (PMF) curve  $W(z)$  reconstructed from the biased umbrella simulation trajectories using the Weighted Histogram Analysis Method (WHAM) (20, 21) implemented using `g_wham` in GROMACS 4.6 (22).



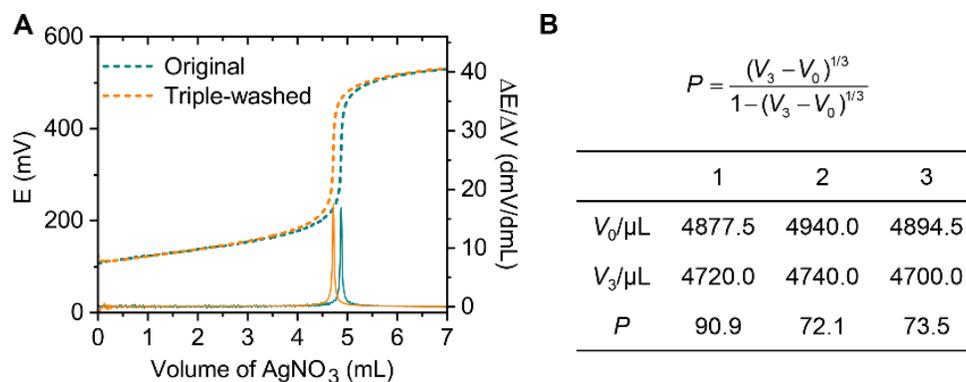
**Fig. S6.** Molecular dynamic simulation revealing the anchoring effect of PEG segments. (A) Probability to locate PEG–PBLG, PEG, and PBLG in a DCM/water biphasic system. The “interface region” was defined as the region  $\pm 2$  nm away from the interfacial plane. Assuming the volume ratio of water to DCM is 1.33:100, with  $V(\text{H}_2\text{O}) = 6.65 \mu\text{L}$ ,  $V(\text{DCM}) = 500 \mu\text{L}$ , and water droplets are non-interacting spheres with a diameter of 600 nm. (B) The  $z$ -coordinate of the center of mass of 9 PEG–PBLGs over the course of 30 ns of molecular dynamics simulation. The interface was set as  $z = 0$  nm, 9 PEG–PBLGs were placed at  $z = -3$  nm initially at  $t = 0$  ns, with the PEG segments touching the interface. The system was then simulated for 30 ns. The PEG chains initially touching the interface rapidly dragged the PEG–PBLGs from the DCM phase to the interface corresponding to the minimum of the PMF and most favorable location of the molecule (Fig. 1G). (C and D) Snapshots of the simulation trajectories at  $t = 0$  ns (C) and 30 ns (D). The structures are color coded as follows: PBLG backbone, green; PBLG side chains, grey; PEG, purple; DCM, orange; water, blue.

#### 4. SIMPLE Polymerization

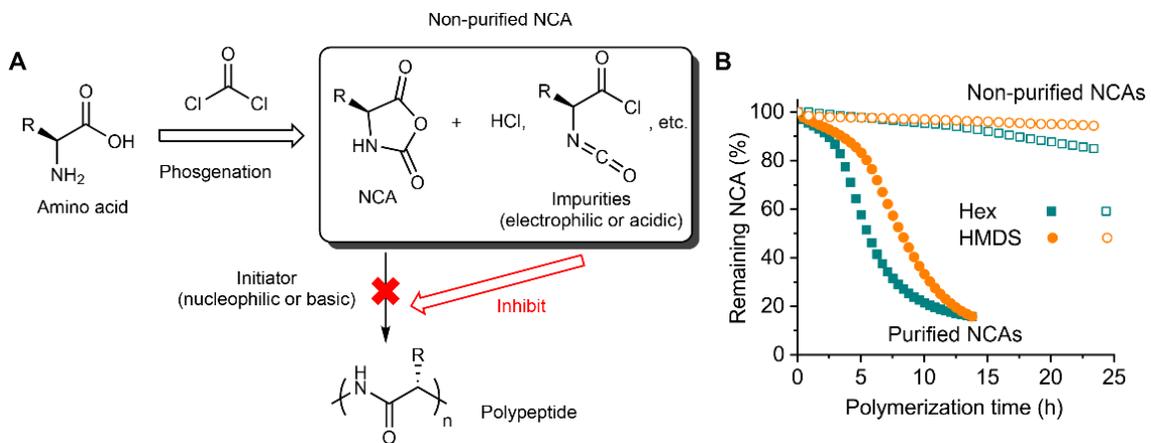
**Determination of Partition Coefficient of HCl.** In order to obtain the partition coefficient of HCl in water/DCM, the aqueous solution of HCl (0.1 M, 20 mL) was washed three times with DCM (20 mL  $\times$  3) using a separation funnel. The concentration of HCl in the original, non-washed solution and the triple-washed solution was then determined through potentiometric titration (23). The experiment was repeated three times and the results were shown as mean  $\pm$  s.d. Titration experiments were performed using CH Instrument 760E bipotentiostat. The aqueous solution of AgNO<sub>3</sub> (0.1 M) was used as titrant. A saturated Ag/AgCl, Ag wire, and Pt wire were used as reference, working, and counter electrode, respectively. During the titration, a syringe pump (KD Scientific Legato 100, KD Scientific Inc., Holliston, MA, USA) was used to inject AgNO<sub>3</sub> titrant into test solutions (4 mL) at a flow rate of 5  $\mu$ L/s under constant stirring. The potential changes of Ag working electrode were simultaneously monitored at different titrant volumes. The titration ends were indicated by the inflection points of the titration curves, which were determined via first order derivation of the curves. The detailed calculation of partition coefficient (*P*) of HCl in water/DCM biphasic mixture was shown in Fig. S7B.

**Inhibitory effect of impurities.** In a glovebox, purified or nonpurified BLG NCA were dissolved in anhydrous DCM (100 mM, 300  $\mu$ L) and mixed with DCM solution of initiators (1 mM, 300  $\mu$ L). The resulting solution was transferred into a KBr sealed liquid cell with circular FTIR aperture (Model SL-3, pathlength = 0.1 mm, International Crystal Laboratories, Garfield, NJ, USA). The liquid cell was then sealed and transferred out of glovebox, and the solution was scanned at pre-determined time points with FTIR. The concentration of NCA was calculated with a standard curve based on the peak area at 1792  $\text{cm}^{-1}$  (*i.e.*, the peak of NCA anhydrides).

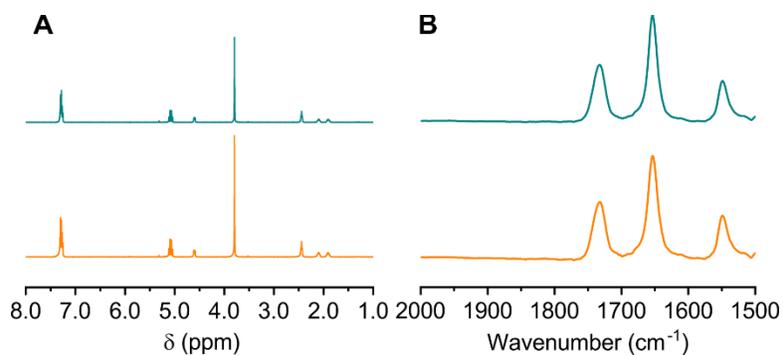
**SIMPLE polymerization of nonpurified NCAs.** Detailed procedures conducting SIMPLE polymerization of nonpurified NCAs were shown in Materials and Methods section. Polymerization kinetics were monitored in deuterated DCM via <sup>1</sup>H NMR in a similar way with purified NCA. After complete conversion of NCA, the top aqueous phase was first removed from the mixture, and the emulsion phase was then dried under vacuum for MW analysis by GPC.



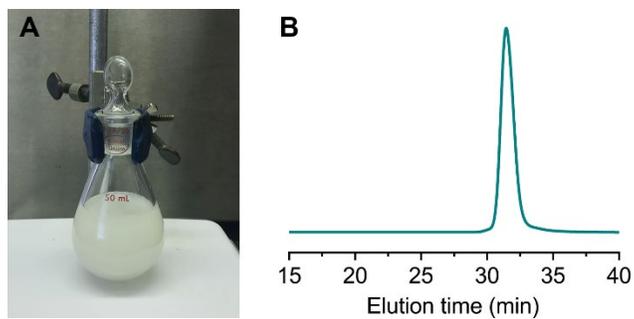
**Fig. S7.** Determination of partition coefficient of HCl in water/DCM. (A) Representative potentiometric titration curve (dashed lines) of HCl in water before and after triple-washing with DCM. The first derivative was shown in solid lines. (B) Summary of titration results and calculation of partition coefficient.  $V_0$ , titrated volumes of the original solution;  $V_3$ , titrated volumes of triple-washed solutions;  $P$ , partition coefficient.



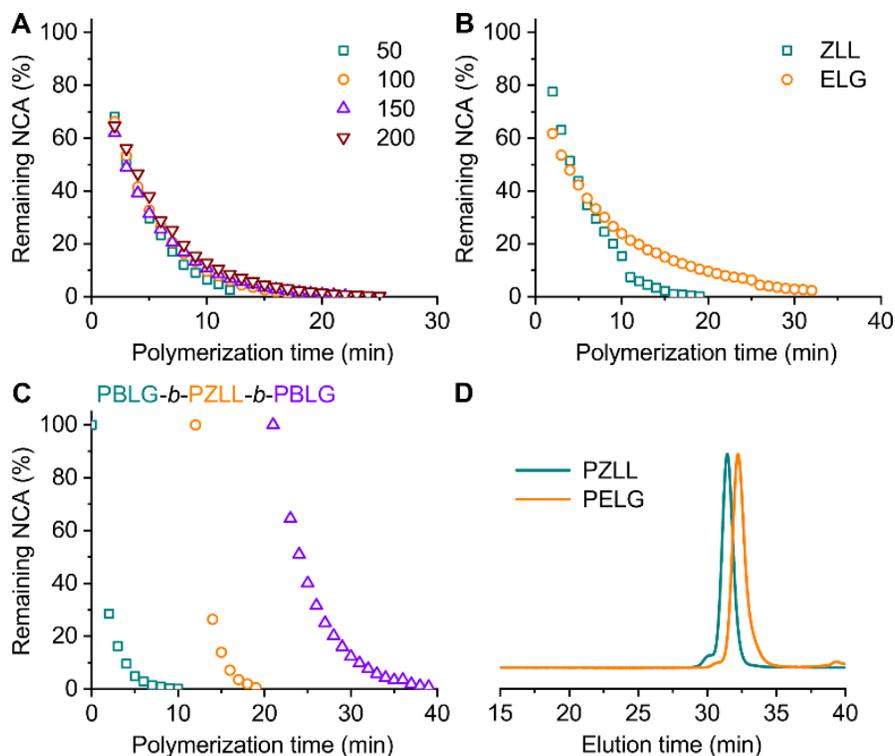
**Fig. S8.** Inhibitory effect of impurities on NCA polymerization. (A) Scheme illustrating the generation of impurities during NCA synthesis and their inhibitory effect on NCA polymerization. (B) Conversion of purified (solid symbols) and nonpurified (open symbols) BLG NCA in a DCM solution under moisture-free conditions. Hex, hexylamine; HMDS, hexamethyldisilazane.  $[M]_0 = 50 \text{ mM}$ ,  $[I]_0 = 0.5 \text{ mM}$ .



**Fig. S9.** Comparison of resulting polypeptide materials synthesized from SIMPLE polymerization and conventional polymerization.  $^1\text{H}$  NMR (*A*) and FTIR (*B*) spectra of resulting polymers synthesized from SIMPLE polymerization of nonpurified BLG NCA (top) and conventional polymerization of purified BLG NCA in anhydrous DCM (bottom). Polymerization condition:  $[\text{M}]_0 = 50 \text{ mM}$ ,  $[\text{I}]_0 = 0.5 \text{ mM}$ . For SIMPLE polymerization, water:DCM = 1:10 (wt/wt), pH = 7.0. The comparable spectra suggested the similar molecular structures and secondary structures of polypeptide materials prepared from SIMPLE polymerization and conventional polymerization.



**Fig. S10.** Gram-scale synthesis of polypeptide materials with SIMPLE polymerization. (A) Photograph showing SIMPLE polymerization of nonpurified BLG NCA (1.0 g). Polymerization condition:  $[M]_0 = 75$  mM,  $[I]_0 = 0.5$  mM, water:DCM = 1:10 (wt/wt), pH = 7.0. (B) Normalized GPC-LS traces of resulting polymers from gram-scale synthesis.  $M_n = 51.7$  kDa ( $M_n^* = 47.8$  kDa),  $D = 1.05$ .

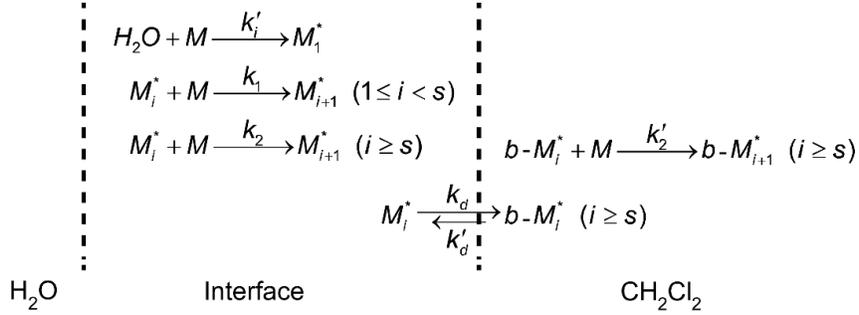


**Fig. S11.** SIMPLE polymerization kinetics and additional GPC analysis. (*A* and *B*) Conversion of nonpurified BLG NCA at various feeding  $[M]_0/[I]_0$  ratios (*A*) and nonpurified ZLL NCA and ELG NCA at  $[M]_0/[I]_0 = 100$  (*B*) in a water/DCM biphasic system.  $[I]_0 = 0.5$  mM in all cases, water:DCM = 1:10 (wt/wt). (*C*) Conversion of nonpurified NCA for block copolymer synthesis. Feeding  $[M]_0/[I]_0 = 50$  for all three blocks,  $[M]_0 = 100, 50,$  and  $25$  mM for the synthesis of first, second, and third block, respectively. (*D*) Normalized GPC-LS traces of resulting polymers from SIMPLE polymerization of nonpurified ZLL NCAs and ELG NCAs.

## 5. Kinetic Modeling

### Differential equations for kinetic modeling

Water/DCM emulsion stabilized with inert PEG (mPEG–NHAc)



The back diffusion is slow compared to the reaction rate and the diffusion equilibrium can't be reached until the reaction is almost done. Therefore, we ignore the back diffusion and only consider the net diffusion from interface to the bulk DCM (*i.e.*,  $k'_d = 0$ ).

Differential equations, with  $[X]$  representing the concentrations of various species:

$$\frac{\partial [M_1^*]}{\partial t} = k'_1[H_2O][M] - k_1[M_1^*][M] \quad i = 1 \quad (\text{S1})$$

$$\frac{\partial [M_i^*]}{\partial t} = k_1[M]([M_{i-1}^*] - [M_i^*]) \quad 1 < i < s \quad (\text{S2})$$

$$\frac{\partial [M_s^*]}{\partial t} = k_1[M][M_{s-1}^*] - k_2[M][M_s^*] - k_d[M_s^*] \quad (\text{S3})$$

$$\frac{\partial [M_i^*]}{\partial t} = k_2[M]([M_{i-1}^*] - [M_i^*]) - k_d[M_i^*] \quad i > s \quad (\text{S4})$$

$$\frac{\partial [b\text{-M}_s^*]}{\partial t} = k_d[M_s^*] - k'_2[b\text{-M}_s^*][M] \quad (\text{S5})$$

$$\frac{\partial [b\text{-M}_i^*]}{\partial t} = k'_2[M]([b\text{-M}_{i-1}^*] - [b\text{-M}_i^*]) + k_d[M_i^*] \quad i > s \quad (\text{S6})$$

$$\frac{\partial [M]}{\partial t} = -k'_1[H_2O][M] - k_1[M] \sum_{i=1}^{s-1} [M_i^*] - k_2[M] \sum_{i=s}^{\infty} [M_i^*] - k'_2[M] \sum_{i=s}^{\infty} [b\text{-M}_i^*] \quad (\text{S7})$$

$$[F1] = \sum_{i=s}^{\infty} [M_i^*] \quad (\text{S8})$$

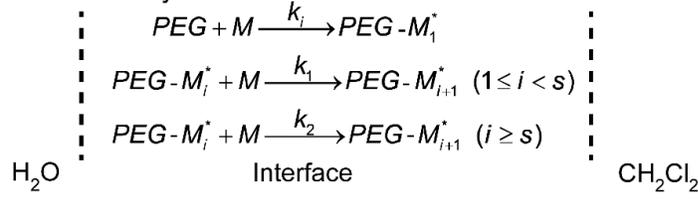
$$\frac{\partial [F1]}{\partial t} = k_1[M_{s-1}^*][M] - k_d[F1] \quad (\text{S9})$$

$$[F2] = \sum_{i=s}^{\infty} [b\text{-M}_i^*] \quad (\text{S10})$$

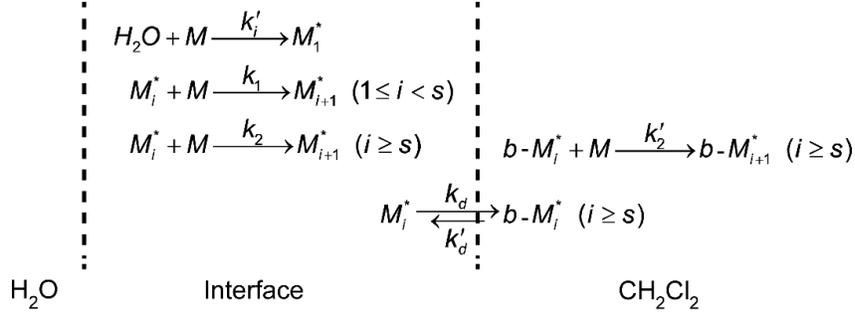
$$\frac{\partial [F2]}{\partial t} = k_d[F1] \quad (\text{S11})$$

Water/DCM emulsion stabilized with PEG macroinitiators

Main Pathway:



Competing Pathway:



Differential equations:

$$\begin{aligned}
 \frac{\partial [M_1^*]}{\partial t} &= k'_i [\text{H}_2\text{O}] [M] - k_1 [M_1^*] [M] & i = 1 \\
 \frac{\partial [M_i^*]}{\partial t} &= k_1 [M] ([M_{i-1}^*] - [M_i^*]) & 1 < i < s \\
 \frac{\partial [M_s^*]}{\partial t} &= k_1 [M] [M_{s-1}^*] - k_2 [M] [M_s^*] - k_d [M_s^*] \\
 \frac{\partial [M_i^*]}{\partial t} &= k_2 [M] ([M_{i-1}^*] - [M_i^*]) - k_d [M_i^*] & i > s \\
 \frac{\partial [b - M_s^*]}{\partial t} &= k_d [M_s^*] - k'_2 [b - M_s^*] [M] \\
 \frac{\partial [b - M_i^*]}{\partial t} &= k'_2 [M] ([b - M_{i-1}^*] - [b - M_i^*]) + k_d [M_i^*] & i > s \\
 \frac{\partial [\text{PEG}]}{\partial t} &= -k_1 [M] [\text{PEG}] & \text{(S12)}
 \end{aligned}$$

$$\frac{\partial [\text{PEG} - M_1^*]}{\partial t} = k_1 [M] ([\text{PEG}] - [\text{PEG} - M_1^*]) \quad i = 1 \quad \text{(S13)}$$

$$\frac{\partial [\text{PEG} - M_i^*]}{\partial t} = k_1 [M] ([\text{PEG} - M_{i-1}^*] - [\text{PEG} - M_i^*]) \quad 1 < i < s \quad \text{(S14)}$$

$$\frac{\partial [\text{PEG} - M_s^*]}{\partial t} = k_1 [M] [\text{PEG} - M_{s-1}^*] - k_2 [M] [\text{PEG} - M_s^*] \quad \text{(S15)}$$

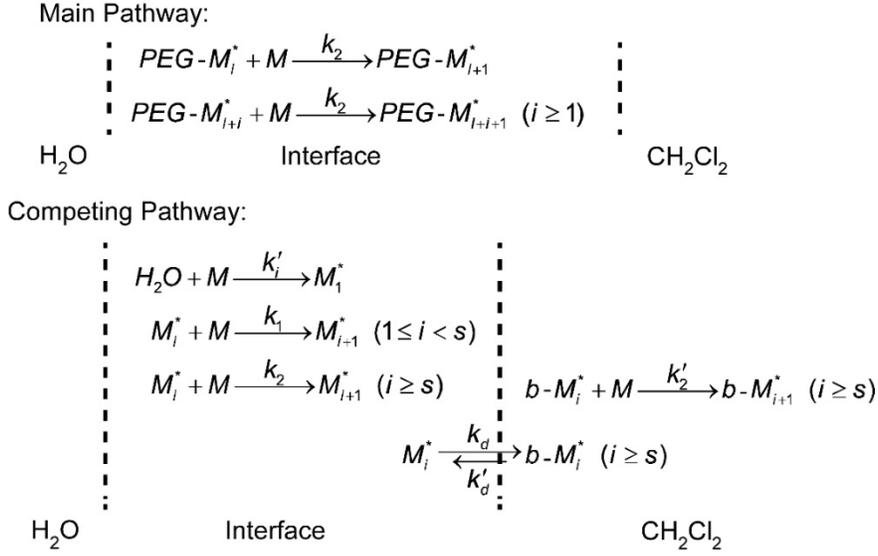
$$\frac{\partial [\text{PEG} - M_i^*]}{\partial t} = k_2 [M] ([\text{PEG} - M_{i-1}^*] - [\text{PEG} - M_i^*]) \quad i > s \quad \text{(S16)}$$

$$\frac{\partial[M]}{\partial t} = -k'_1[H_2O][M] - k_1[M]\sum_{i=1}^{s-1}([M_i^*] + [PEG-M_i^*]) - k_1[M][PEG] - k_2[M]\sum_{i=s}^{\infty}([M_i^*] + [PEG-M_i^*]) - k'_2[M]\sum_{i=s}^{\infty}[b-M_i^*] \quad (S17)$$

$$[F3] = \sum_{i=s}^{\infty}[PEG-M_i^*] \quad (S18)$$

$$\frac{\partial[F3]}{\partial t} = k_1[PEG-M_{s-1}^*][M] \quad (S19)$$

Water/DCM emulsion stabilized with PEG-PBLG macroinitiators



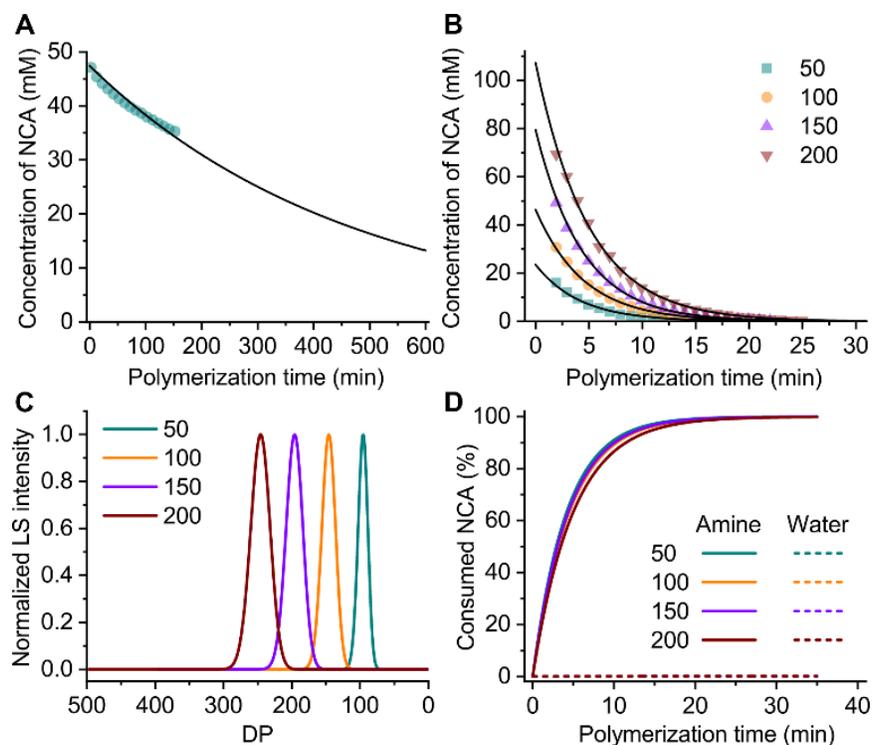
Differential equations:

$$\begin{aligned} \frac{\partial[M_1^*]}{\partial t} &= k'_1[H_2O][M] - k_1[M_1^*][M] & i=1 \\ \frac{\partial[M_i^*]}{\partial t} &= k_1[M]([M_{i-1}^*] - [M_i^*]) & 1 < i < s \\ \frac{\partial[M_s^*]}{\partial t} &= k_1[M][M_{s-1}^*] - k_2[M][M_s^*] - k_d[M_s^*] \\ \frac{\partial[M_i^*]}{\partial t} &= k_2[M]([M_{i-1}^*] - [M_i^*]) - k_d[M_i^*] & i > s \\ \frac{\partial[b-M_s^*]}{\partial t} &= k_d[M_s^*] - k'_2[b-M_s^*][M] \\ \frac{\partial[b-M_i^*]}{\partial t} &= k'_2[M]([b-M_{i-1}^*] - [b-M_i^*]) + k_d[M_i^*] & i > s \\ \frac{\partial[PEG-M_i^*]}{\partial t} &= -k_2[M][PEG-M_i^*] & (S20) \end{aligned}$$

$$\frac{\partial[PEG-M_{i+i}^*]}{\partial t} = k_2[M]([PEG-M_{i+i-1}^*] - [PEG-M_{i+i}^*]) \quad i \geq 1 \quad (S21)$$

$$\begin{aligned} \frac{\partial[M]}{\partial t} = & -k_1'[H_2O][M] - k_1[M] \sum_{i=1}^{s-1} [M_i^*] - k_2[M] \sum_{i=s}^{\infty} [M_i^*] \\ & - k_2[M][PEG-M_i^*]_0 - k_2'[M] \sum_{i=s}^{\infty} [b \cdot M_i^*] \end{aligned} \quad (S22)$$

In all the cases, the water concentration at the interface is assumed to be constant. At  $t = 0$ ,  $[PEG] = [PEG]_0 \times IE$ , where  $IE$  is the initiation efficiency of PEG.  $IE$  was first calculated to be 26% from the GPC of PEG-initiated polymers (Fig. 1F), which was then used as a reference in the modeling, the calculated  $IE$  was determined to be 22% by fitting.

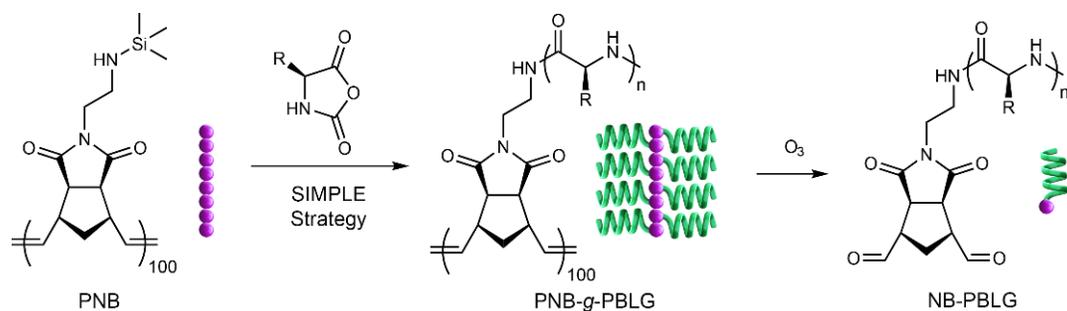


**Fig. S12.** Kinetic modeling of SIMPLE polymerization. (A) Kinetic data from PBLG-initiated polymerization of purified BLG NCAs in DCM solution (dark cyan circles) was fit with the cooperative kinetic model (black line). Rate constant  $k'_2$  was determined as  $0.075 \text{ M}^{-1} \text{ s}^{-1}$ . (B) Kinetic data from SIMPLE polymerization of nonpurified BLG NCAs (colored symbols) in a water/DCM biphasic system was fit with the cooperative kinetic model (black lines). Rate constant  $k_2$  was determined as 8.5, 8.0, 7.2, and  $6.2 \text{ M}^{-1} \text{ s}^{-1}$  for  $[M]_0/[I]_0$  ratio 50, 100, 150, and 200, respectively. (C) Simulated GPC-LS traces from SIMPLE polymerization with various feeding  $[M]_0/[I]_0$  ratios predicted by the kinetic model. (D) Prediction of water competition in SIMPLE polymerization of nonpurified BLG NCAs.

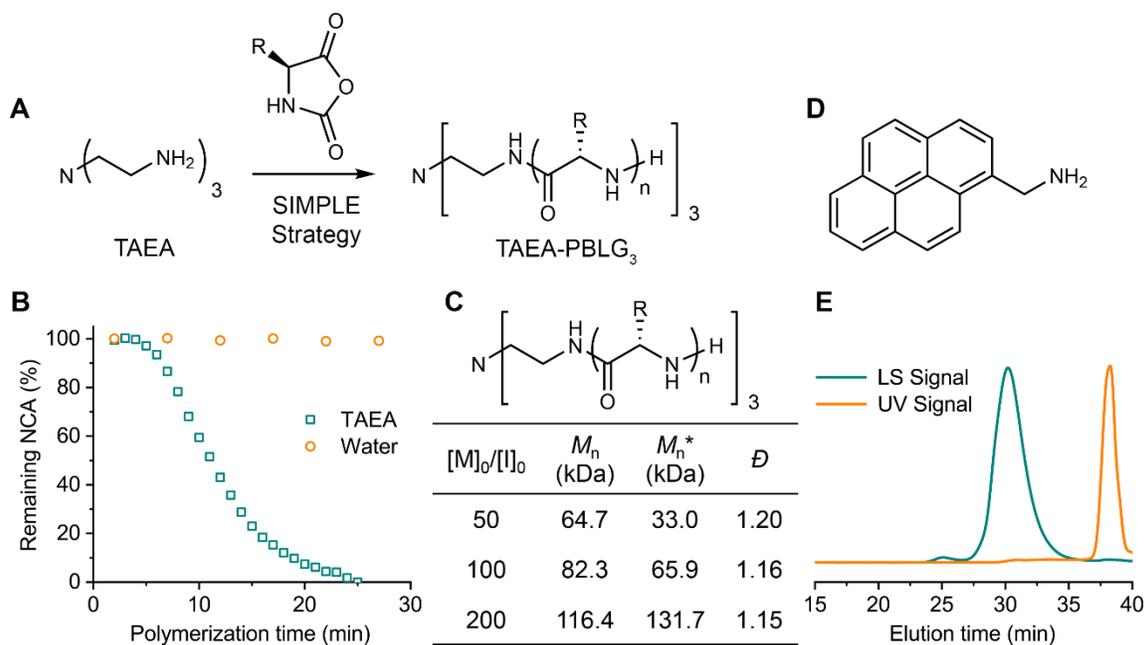
## 6. Extension of SIMPLE concept

**SIMPLE polymerization in DCM phase.** SIMPLE polymerization in DCM phase was carried out in a similar way with SIMPLE polymerization at water/DCM interface. Typically, nonpurified NCAs were dissolved in regular DCM (0.1 M), mixed with aqueous buffer (pH = 7.0, 18 wt%), and vortexed for 10 s. DCM solution of various initiators was immediately added into the biphasic mixture containing NCA to start the polymerization. No stable emulsion was formed throughout the polymerization process.

**Preparation of NB-PBLG homopolypeptide from PNB-g-PBLG.** Backbone cleavage of PNB-g-PBLG using ozone was conducted following a literature procedure (2). DCM solution of PNB-g-PBLG (5 mg/mL) was cooled to -78 °C using a dry ice/acetone bath. The solution was bubbled with oxygen for 2 min through ozone generator (Model: VMUS-4S, AZCO Industries Limited, Langley, B.C., Canada), followed by the bubbling of ozone for ~ 5 min. The solution turned blue during ozone bubbling, indicating successful reaction. After the oxidation, the solution was bubbled with oxygen until blue colors fade away (~ 2-5 min). Dimethyl sulfide (100 µL) was added into the resulting solution, which was allowed to sit overnight at room temperature. The solvent was then removed under vacuum, and the unfractionated solid residues were dissolved in DMF containing LiBr (0.1M) and analyzed by GPC.

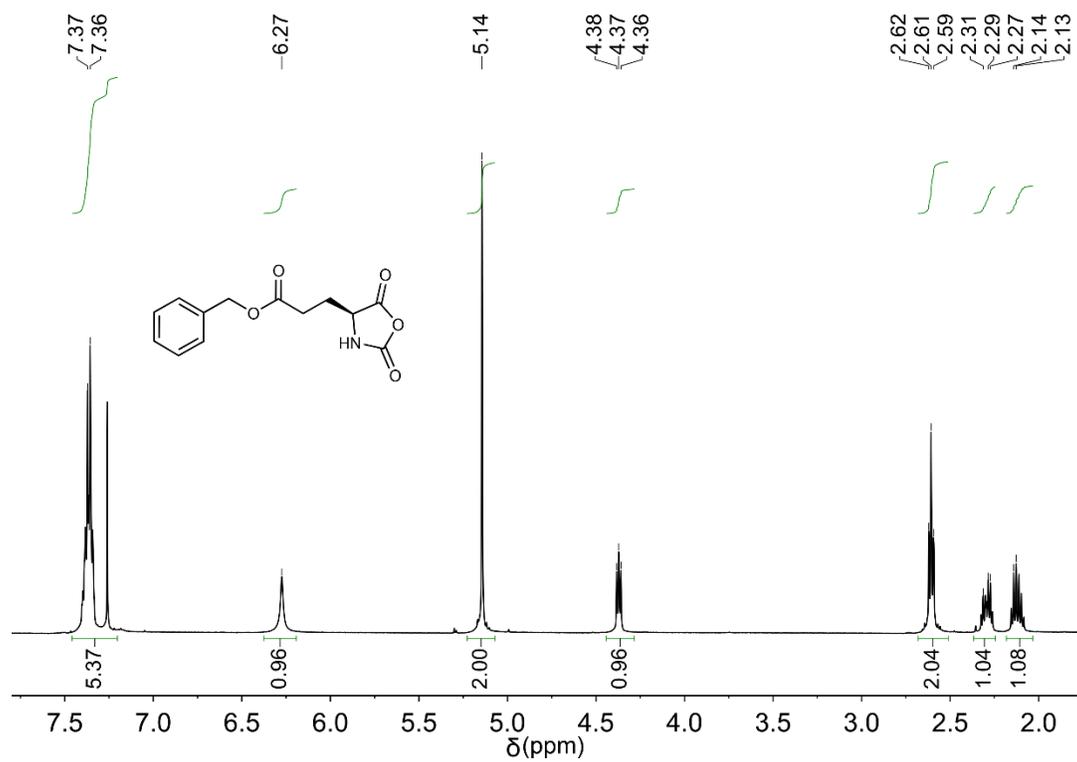


**Fig. S13.** Synthetic route to homopolypeptides from nonpurified NCAs in the presence of water. A brush-like initiator, PNB, was first used as the initiators for SIMPLE polymerization of nonpurified BLG NCAs in the presence of water. The obtained brush-like polymers were then treated with ozone to cleave the backbone, yielding PBLG homopolypeptides.

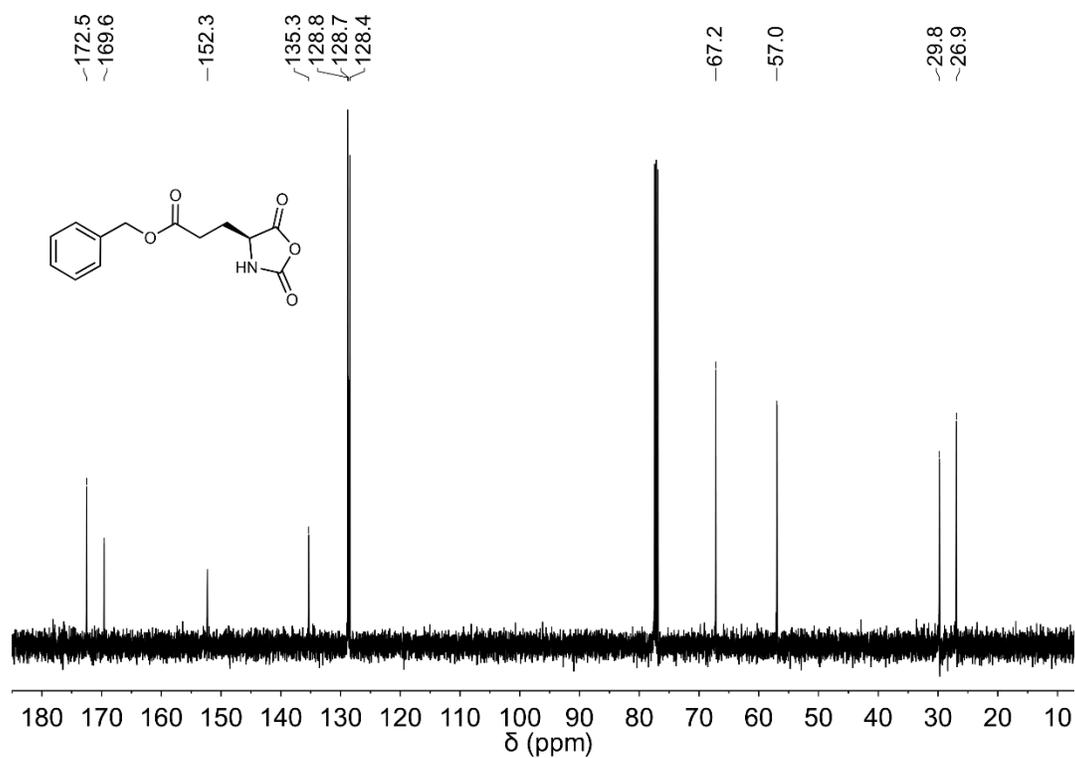


**Fig. S14.** SIMPLE polymerization of nonpurified BLG NCA in the presence of TAEA or 1-pyrenemethyl amine initiators. (A) Scheme illustrating the preparation of polypeptides from tris(2-aminoethyl)amine (TAEA), an initiator bearing three primary amine groups and a tertiary amine group. (B) Conversion of nonpurified BLG NCA in a water/DCM biphasic system using TAEA as initiators.  $[M]_0 = 50$  mM,  $[M]_0/[I]_0 = 100$ , water:DCM = 1:10 (wt/wt). (C) GPC characterization of resulting polypeptides from TAEA-initiated polymerization of nonpurified BLG NCAs. (D) Chemical structure of 1-pyrenemethyl amine. (E) Overlaid GPC-LS and GPC-UV traces of resulting polymers from SIMPLE polymerization of nonpurified BLG NCAs in the presence of 1-pyrenemethyl amine. For GPC-UV,  $\lambda = 345$  nm is used to detect species containing pyrene.

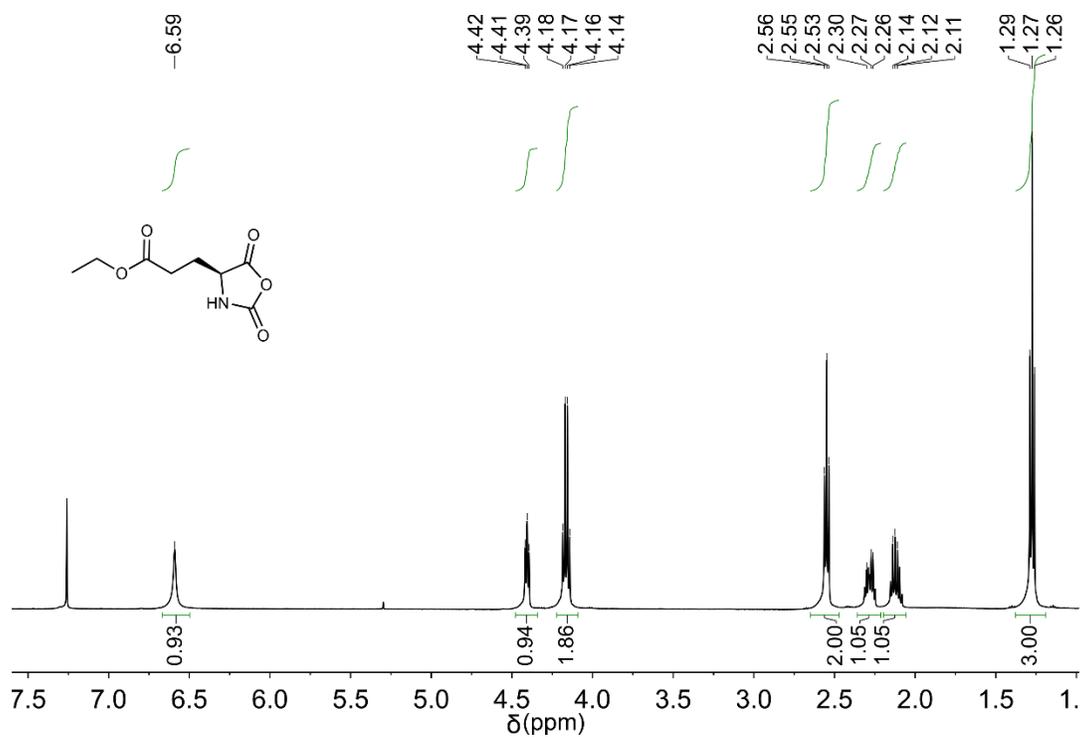
## 7. NMR Spectra



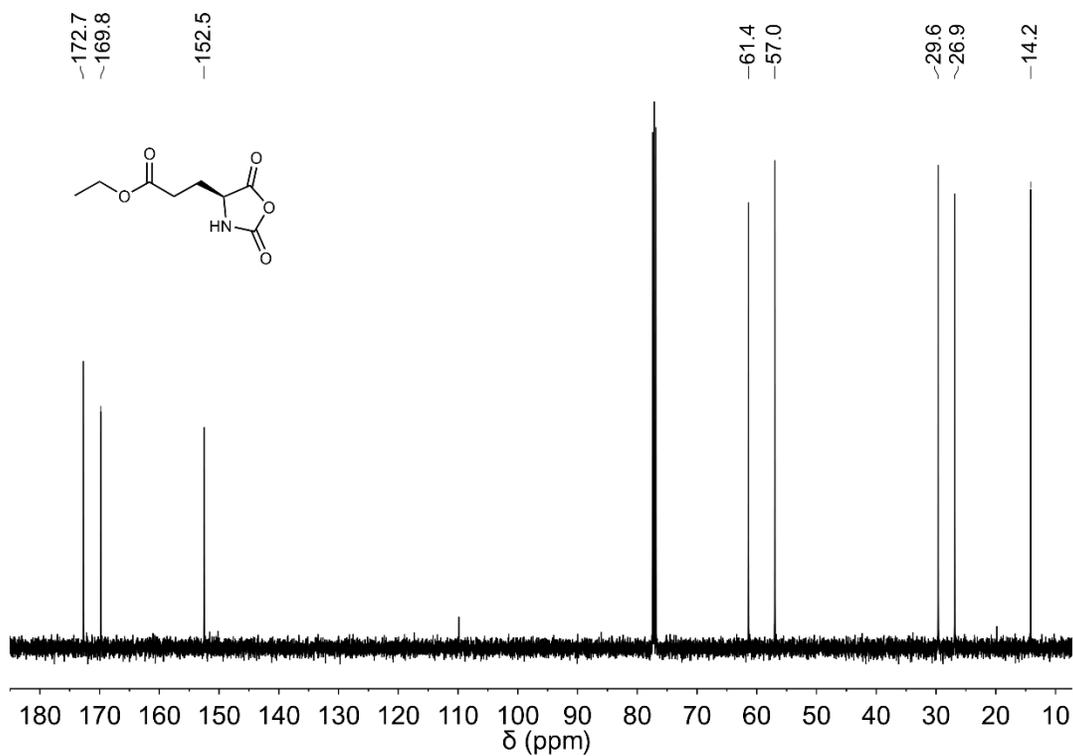
**Fig. S15.**  $^1\text{H}$  NMR spectrum (500 MHz) of nonpurified  $\gamma$ -benzyl-L-glutamate *N*-carboxyanhydride (BLG NCA) in  $\text{CDCl}_3$ .



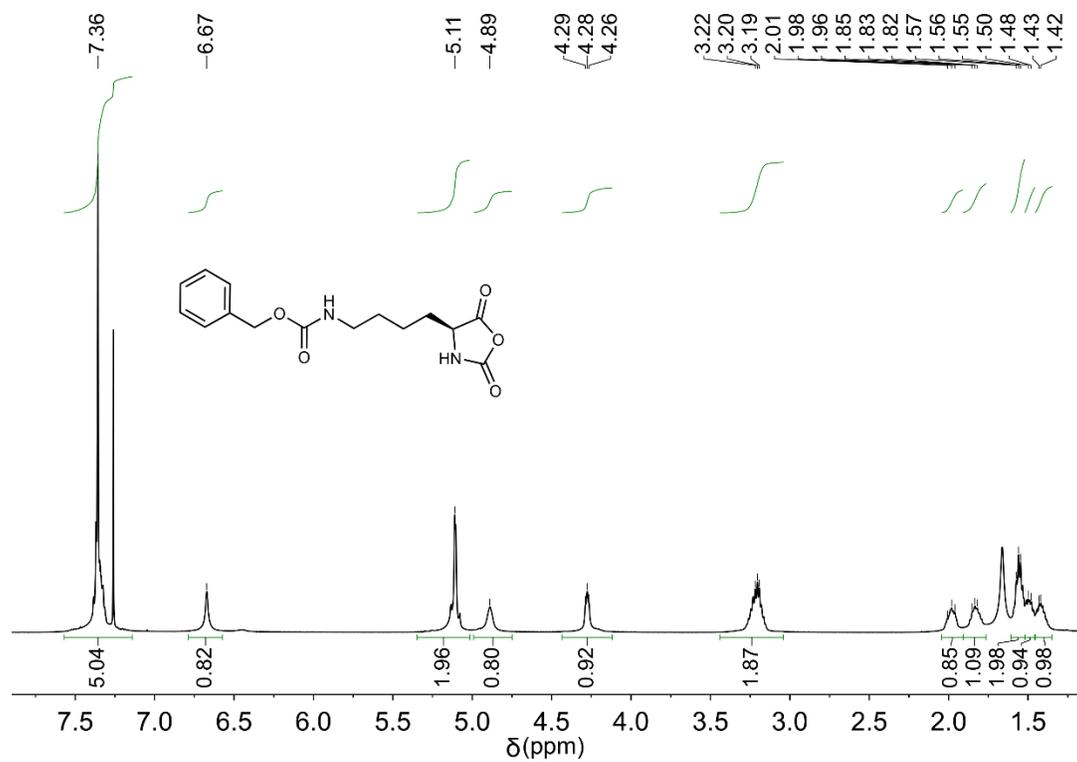
**Fig. S16.** <sup>13</sup>C NMR spectrum (125 MHz) of nonpurified  $\gamma$ -benzyl-L-glutamate *N*-carboxyanhydride (BLG NCA) in CDCl<sub>3</sub>.



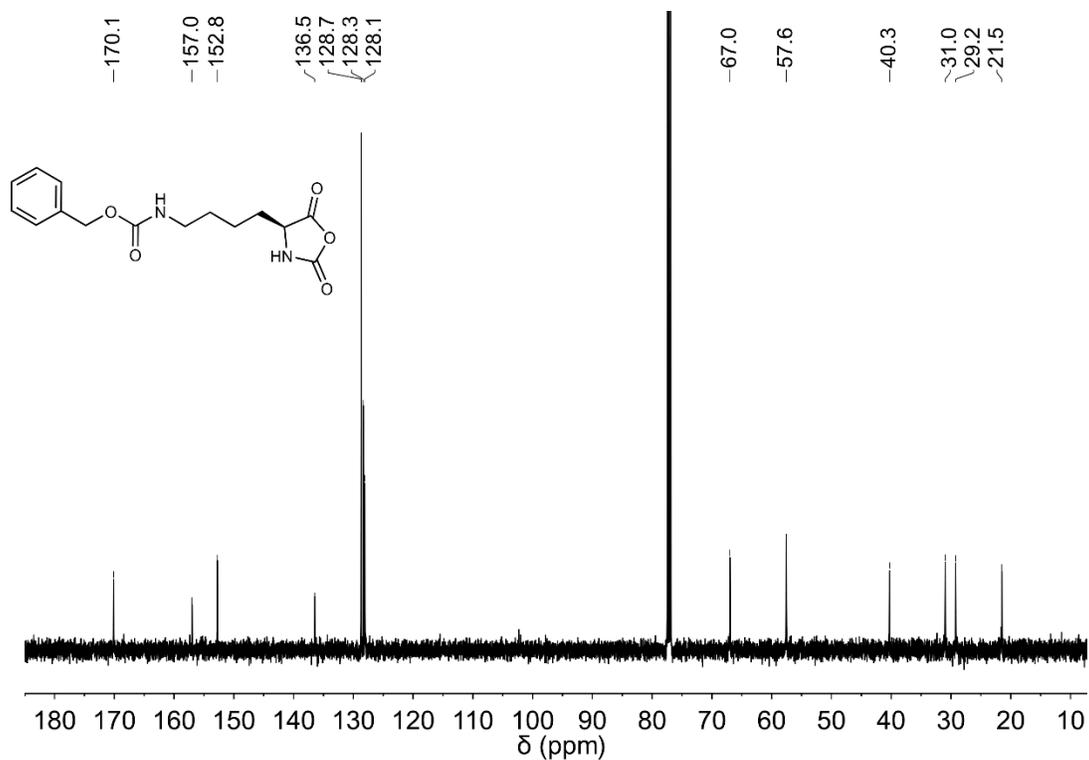
**Fig. S17.** <sup>1</sup>H NMR spectrum (500 MHz) of nonpurified  $\gamma$ -ethyl-L-glutamate *N*-carboxyanhydride (ELG NCA) in CDCl<sub>3</sub>.



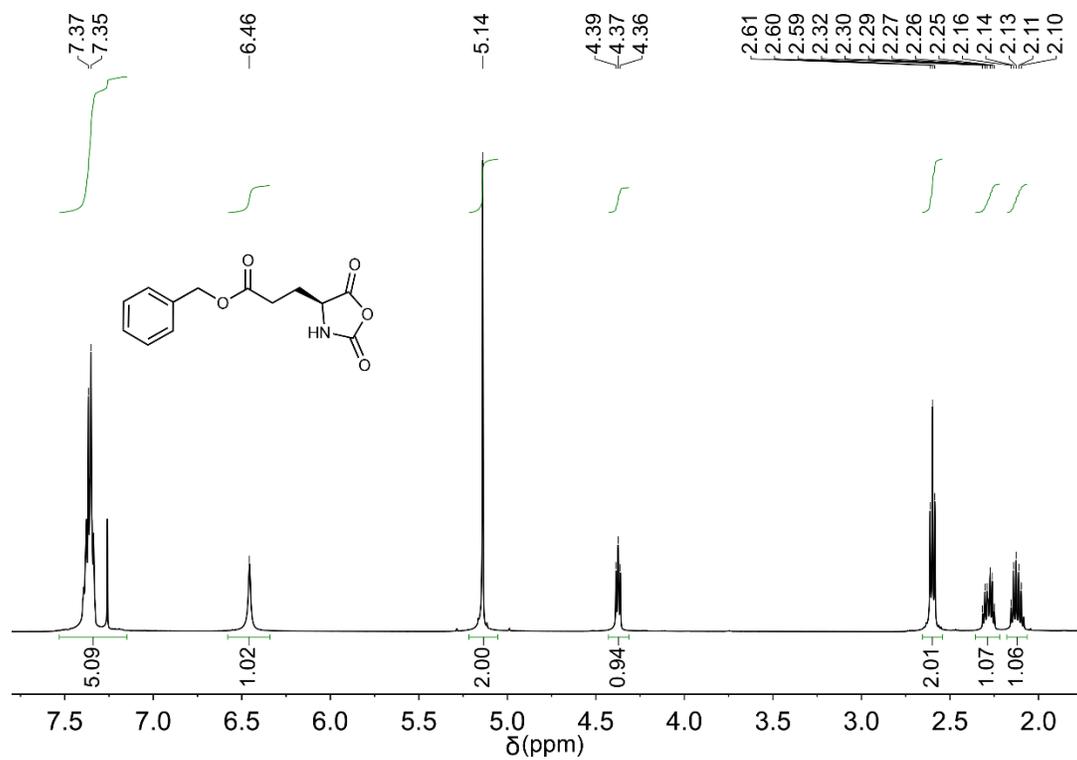
**Fig. S18.** <sup>13</sup>C NMR spectrum (125 MHz) of nonpurified γ-ethyl-L-glutamate N-carboxyanhydride (ELG NCA) in CDCl<sub>3</sub>.



**Fig. S19.** <sup>1</sup>H NMR spectrum (500 MHz) of nonpurified ε-carboxybenzyl-L-lysine *N*-carboxyanhydride (ZLL NCA) in CDCl<sub>3</sub>.



**Fig. S20.** <sup>13</sup>C NMR spectrum (125 MHz) of nonpurified ε-carboxybenzyl-L-lysine N-carboxyanhydride (ZLL NCA) in CDCl<sub>3</sub>.



**Fig. S21.** <sup>1</sup>H NMR spectrum (500 MHz) of purified  $\gamma$ -benzyl-L-glutamate *N*-carboxyanhydride (BLG NCA) in CDCl<sub>3</sub>.

## 8. Supplementary Movies

**Movie S1.** Movie illustrating the spontaneous partitioning of PEG–PBLG macroinitiators towards the interface due to PEG segments touching the interface in molecular dynamics simulation. The interface was set as  $z = 0$  nm, 9 PEG–PBLGs were placed at  $z = -3$  nm initially at  $t = 0$  ns, with the PEG segments touching the interface. The molecules were then simulated for 30 ns. The PEG chains initially touching the interface rapidly dragged the PEG–PBLGs from the DCM phase to the interface corresponding to the minimum of the PMF and most favorable location of the molecule (Fig. 1G). Two replicas of the fundamental simulation cell are shown connected through the periodic boundary.

**Movie S2.** Movie illustrating SIMPLE polymerization of nonpurified NCAs in the presence of water. Nonpurified BLG NCA at a feeding  $[M]_0/[I]_0$  ratio of 100 was used as an example. (a) The DCM solution of PEG–PBLG (1 mM) was mixed with aqueous buffer (pH = 7, water:DCM = 1:50, wt/wt), which was emulsified with a probe sonicator. (b) Nonpurified BLG NCA was dissolved in DCM (100 mM), which was treated with aqueous buffer to extract impurities (pH = 7, water:DCM = 9:50, wt/wt). Vortexing of the mixture (10 s) promoted the segregation process. (c) The w/o emulsion of PEG–PBLG was added into the mixture of NCA at 1:1 ratio to initiated the polymerization. Final condition:  $[M]_0 = 50$  mM,  $[I]_0 = 0.5$  mM, water:DCM = 1:10 (wt/wt).

## References

1. Ji S, Hoye TR, Macosko CW (2005) Primary Amine ( $-NH_2$ ) Quantification in Polymers: Functionality by  $^{19}F$  NMR Spectroscopy. *Macromolecules* 38(11):4679-4686.
2. Baumgartner R, Fu H, Song Z, Lin Y, Cheng J (2017) Cooperative polymerization of  $\alpha$ -helices induced by macromolecular architecture. *Nat Chem* 9:614-622.
3. Carmody WR (1961) Easily prepared wide range buffer series. *J Chem Educ* 38(11):559.
4. Adler AJ, Greenfield NJ, Fasman GD (1973) Circular dichroism and optical rotatory dispersion of proteins and polypeptides. *Methods Enzymol*, eds Hirs CHW, Timasheff SN (Academic Press), Vol 27, pp 675-735.
5. Greenfield NJ (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* 1(6):2876-2890.
6. Habraken GJM, Peeters M, Dietz CHJT, Koning CE, Heise A (2010) How controlled and versatile is N-carboxy anhydride (NCA) polymerization at 0 °C? Effect of temperature on homo-, block- and graft (co)polymerization. *Polym Chem* 1(4):514-524.
7. Miyazawa T, Blout ER (1961) The infrared spectra of polypeptides in various conformations: Amide I and II bands. *J Am Chem Soc* 83(3):712-719.
8. Malde AK, et al. (2011) An automated force field topology builder (ATB) and repository: version 1.0. *J Chem Theory Comput* 7(12):4026-4037.
9. Schmid N, et al. (2011) Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur Biophys J* 40(7):843-856.
10. Stewart JJP (1990) MOPAC: a semiempirical molecular orbital program. *J Comput-Aided Mol Des* 4(1):1-103.
11. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4(3):435-447.
12. Allen MP, Tildesley DJ (1989) *Computer Simulation of Liquids* (Clarendon Press, New York).
13. Essmann U, et al. (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103(19):8577-8593.
14. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: A linear constraint solver for molecular simulations. *J Comput Chem* 18(12):1463-1472.
15. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. *J Chem Phys* 81(1):511-519.
16. Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys* 52(12):7182-7190.
17. Hockney RW, Eastwood JW (1988) *Computer Simulation Using Particles* (CRC Press, Bristol).
18. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graphics* 14(1):33-38.
19. Torrie GM, Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J Comput Phys* 23(2):187-199.

20. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA (1992) THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13(8):1011-1021.
21. Roux B (1995) The calculation of the potential of mean force using computer simulations. *Comput Phys Commun* 91(1):275-282.
22. Hub JS, de Groot BL, van der Spoel D (2010) g\_wham—A free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J Chem Theory Comput* 6(12):3713-3720.
23. Tang T-F, Gordon G (1980) Quantitative determination of chloride, chlorite, and chlorate ions in a mixture by successive potentiometric titrations. *Anal Chem* 52(9):1430-1433.