Supplemental Information

Efficient Synthesis of Amino Acid Polymers for Protein Stabilization

Bing Li,[a,b,c] Yueming Wu,[a, b, c] Wenjing Zhang,[a, b, c] Si Zhang,[a, b, c] Ning Shao,[a, b, c] Weiwei Zhang,[a, b, c] Lixin Zhang,[a] Jian Fei,[d] Yidong Dai[e] and Runhui Liu[a, b, c]*

[a] State Key Laboratory of Bioreactor Engineering, [b] Key Laboratory for Ultrafine Materials of Ministry of Education, [c] Research Center for Biomedical Materials of Ministry of Education, East China University of Science and Technology, Shanghai 200237, China

[d] Department of General Surgery, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

[e] Shanghai Ruijin Rehabilitation Hospital, Shanghai 200023, China

Correspondence should be addressed to R.L. (rliu@ecust.edu.cn)

Synthesis of γ-tert-butyl-L-glutamate NCA (*t*Bu-L-Glu NCA)

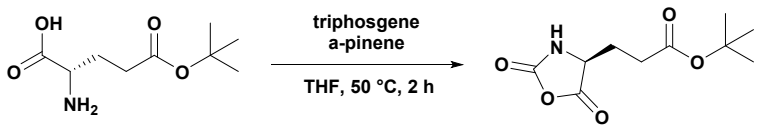

Fig. S1 Synthesis of *t*Bu-L-Glu NCA.

*t*Bu-L-Glu NCA was prepared by following the precedent literature with slight modification.[1] L-glutamic acid 5-tert-butyl ester (5.00 g, 24.60 mmol) and α-pinene (9.35 mL, 59.04 mmol) were suspended in 80 mL anhydrous tetrahydrofuran (THF) in a 250 mL round bottom flask, followed by dropwise addition of triphosgene (2.92 g, 9.84 mmol) dissolved in anhydrous THF under an ice-water bath in nitrogen environment. Final concentration of L-glutamic acid 5-tert-butyl ester was 0.3 M. The reaction was stirred at 50 °C under nitrogen environment for 2 hours, and then the solvent was removed by rotovap. The product was quickly dissolved in ethyl acetate (EtOAc) (100 mL) and then was washed sequentially with cold water (1 × 100 mL) and cold brine (1 × 100 mL). The combined organic phase was dried over anhydrous $MgSO_4$ and then concentrated to give the crude NCA. The crude product was recrystallized three times in a glovebox in nitrogen environment using anhydrous EtOAc and hexane to afford the pure NCA as a white powder (3.10 g, 55% yield). ^{1}H NMR (400 MHz, $CDCl_3$, Fig. S16): δ 6.66 (s, 1H), 4.41-4.34 (m, 1H), 2.46 (t, *J* = 6.8 Hz, 2H), 2.29-2.19 (m, 1H), 2.12-2.00 (m, 1H), 1.45 (s, 9H).

Synthesis of Nε-tert-butyloxycarbonyl-L-lysine NCA (Boc-L-Lys NCA)

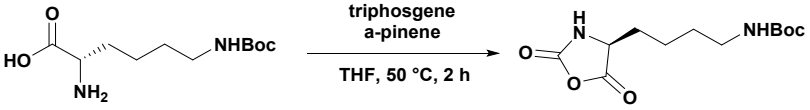

Fig. S2 Synthesis of Boc-L-Lys NCA.

Boc-L-Lys NCA was prepared by following a method that was reported previously.[2, 3] Nε-tert-butyloxycarbonyl-L-lysine (10.00 g, 40.60 mmol) and α-pinene (16.98 mL, 107.18 mmol) were suspended in 150 mL anhydrous THF in a 250 mL round bottom flask followed by dropwise addition of triphosgene (5.30 g, 17.86 mmol) dissolved in anhydrous THF under an ice-water bath in nitrogen environment. Final concentration of Nε-tert-butyloxycarbonyl-L-lysine was 0.3 M. The reaction was stirred at 50 °C under nitrogen environment for 2 hours, and then the solvent was removed by rotovap. The product was quickly dissolved in EtOAc (100 mL) and then was washed sequentially with cold water (1 × 100 mL) and cold brine (1 × 100 mL). The combined organic

phase was dried over anhydrous $MgSO_4$ and then concentrated to give the crude NCA. The crude product was recrystallized three times in a glovebox in nitrogen environment using anhydrous EtOAc and hexane to afford the pure NCA as a white powder (5.30 g, 48% yield). ^{1}H NMR (400 MHz, $CDCl_3$, Fig. S17): δ 6.84 (s, 1H), 4.67 (s, 1H), 4.33 (dd, J = 6.5, 5.0 Hz, 1H), 3.13 (s, 2H), 2.08-1.95 (m, 1H), 1.90-1.79 (m, 1H), 1.62-1.48 (m, 4H), 1.45 (s, 9H).

Polymer synthesis and characterization

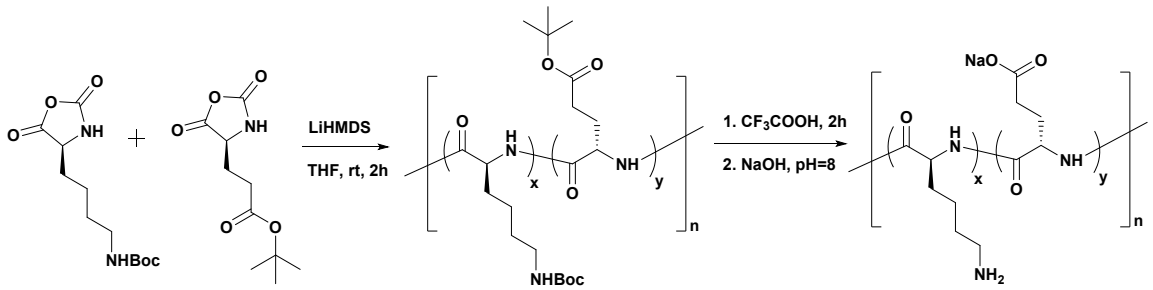

Fig. S3 Synthesis of PLG-*r*-PLL. x+ y = 100%, x = 20%, 40%, 50%, 60% or 80%.

For the synthesis of binary copolymers, a representative synthesis of PLG50-*r*-PLL50 is described below. In a nitrogen purged glovebox, Boc-L-Lys NCA (0.25 mL, 0.3 M) and *t*Bu-L-Glu NCA (0.25 mL, 0.3 M) were added to a dried reaction vial with a magnetic stir bar, followed by addition of 0.03 mL LiHMDS solution of 0.1 M. The reaction was reacted for 2 hours and then removed out of the glovebox, followed by pouring into cold petroleum ether (40 mL) to give a white precipitation. The precipitate was collected by centrifugation and dried under air flow. The collected product was dissolved in THF (1.5 mL) and then was precipitated out again by adding petroleum ether (40 mL) to the solution. This dissolution-precipitation process was repeated three times to get PLG50-*r*-PLL50 in the sidechain protected stage as a white solid. These amino acid polymers at the sidechain protected stage were characterized by GPC. To remove *N*-Boc and tert-butyl protecting group of PLG50-*r*-PLL50, 1.5 mL trifluoroacetic acid (TFA) was added to the solid. After gentle shaking 2 hours at room temperature, the TFA was removed by air flow to afford a viscous liquid. After that, MeOH (0.5 mL) was added to the viscous liquid followed by addition of Et_2O (40 mL) to precipitate out the deprotected polymer. The solid was collected after centrifugation and dried under nitrogen flow. After three cycles of dissolution-precipitation process, the collected solid was dissolved in milli-Q, followed by the addition of NaOH to milli-Q and adjusted the pH to 8. The final amino acid polymers were obtained as a fluffy white solid by lyophilization and used for further protein activity studies. PLG-*r*-PLL ^{1}H NMR (400 MHz, D_2O, Fig. S18-S22).

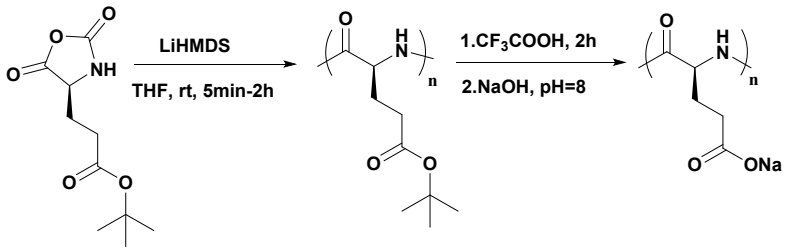

Fig. S4 Synthesis of PLG with different chain lengths.

PLG were prepared by following the method that was reported previously with slight modification.[3] A representative synthesis of PLG1 is described below. In a nitrogen purged glovebox, *t*Bu-L-Glu NCA (103.15mg, 0.45 mmol) was weighed out, and added to a dried reaction vial with a magnetic stir bar, then it was resolved using THF to reach the concentration of 0.1 M. A 0.9 mL solution of lithium bis(trimethylsilyl)amide (LiHMDS) (0.1 M) was added to the reaction vial and the mixture was stirred at room temperature for 5 min. The other polymers with different chain lengths were synthesized similarly. The reaction time was increased depending on target polymer chain length, with PLG2 for 10 min, PLG3 and PLG4 for 2 hours respectively. After the reaction completion, cold petroleum ether (40 mL) was poured into the reaction mixture to precipitate out a white fluffy solid. The precipitate was collected by centrifugation of 4000 rpm and dried by air flow. The collected product was dissolved in THF (1.5 mL) and then precipitated out again by adding petroleum ether (40 mL) to the solution. This dissolution-precipitation process was repeated three times to obtain the side-chain protected polymer. To remove the tert-butyl protecting group of PLG, 1.5 mL TFA was added to the solid. After gentle shaking 2 hours at room temperature, the TFA was removed by air flow to afford a viscous liquid. Then MeOH (0.5 mL) was added to the viscous liquid followed by the addition of Et_2O (40 mL) to get the white precipitate. The precipitate was obtained after centrifugation and dried under nitrogen flow. After three cycles of dissolution-precipitation process, the collected solid was dissolved in milli-Q followed by addition of NaOH to adjusted the pH to 8. The final amino acid polymers were obtained after lyophilization as a fluffy white solid. These amino acid polymers at the deprotected stage were characterized by GPC and used for further protein activity studies.

The feed ratio between tBu-L-Glu and Boc-L-Lys and comonomer composition observed by ^{1}H NMR

Table S1 Comonomer composition observed by ^{1}H NMR

PLG-r-PLL	Feed ratio between tBu-L-Glu and Boc-L-Lys	Comonomer composition observed by 1H NMR
PLG80-r-PLL20	**8:2**	**3.67:1**
PLG60-r-PLL40	**6:4**	**1.47:1**
PLG50-r-PLL50	**5:5**	**1.04:1**
PLG40-r-PLL60	**4:6**	**0.67:1**
PLG20-r-PLL80	**2:8**	**0.25:1**

Zeta Potential of PLG-*r*-PLL

The zeta potential of PLG-*r*-PLL were determined with a Zetasizer Nano-ZS from Malvern Instruments using 1.3 mg/mL sample solution in water. Each reported measurement was conducted for three times.

Table S2 Zeta Potential of PLG-*r*-PLL

PLG-*r*-PLL	Zeta Potential (mV)
PLG80-*r*-PLL20	**-46.5±1.8**
PLG60-*r*-PLL40	**-25.7±2.6**
PLG50-*r*-PLL50	**-5.5±0.3**
PLG40-*r*-PLL60	**30.5±0.1**
PLG20-*r*-PLL80	**48.1±1.1**

Activity of amino acid polymers only (no protein) on β-Gal activity indicator, 2-Nitrophenyl-beta-D-galactopyranoside (ONPG).

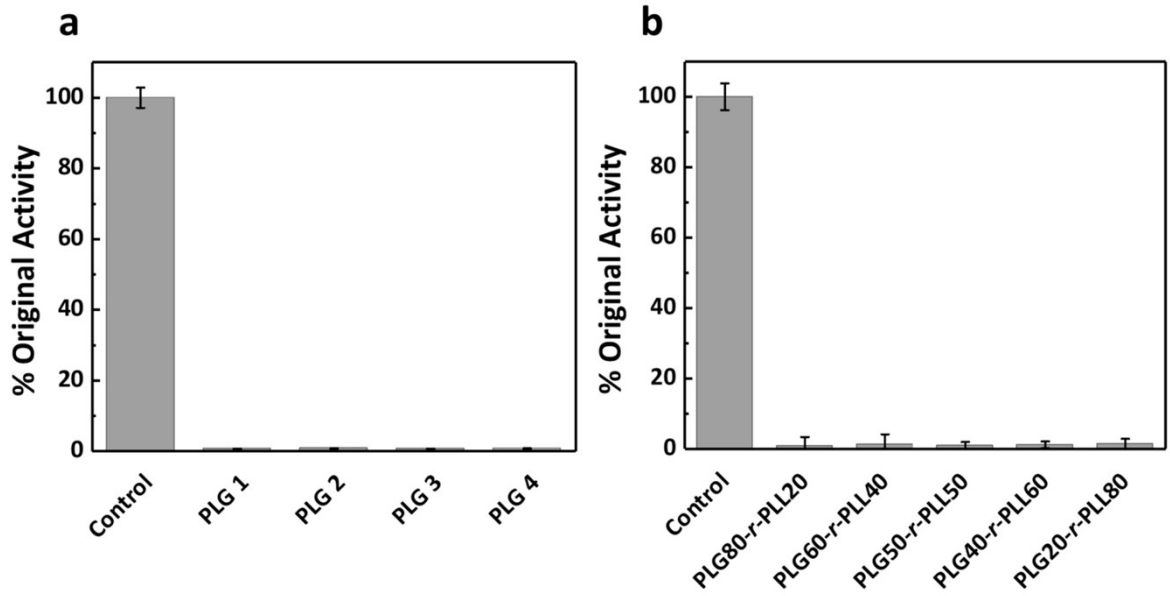

Fig. S5 Activity of amino acid polymers only (no protein) on β-Gal activity indicator, 2-Nitrophenyl-beta-D-galactopyranoside (ONPG). (a) Activity of PLG on ONPG. (b) Activity of PLG-*r*-PLL on ONPG. Data shown are based on 3 repeats. These data indicated that all amino acid polymers themselves in this study have no effect on β-Gal activity analysis.

β-Gal activity after lyophilization using poly-L-lysine (PLL), L-lysine (L-Lys), L-Glutamic acid (L-Glu) as the control samples.

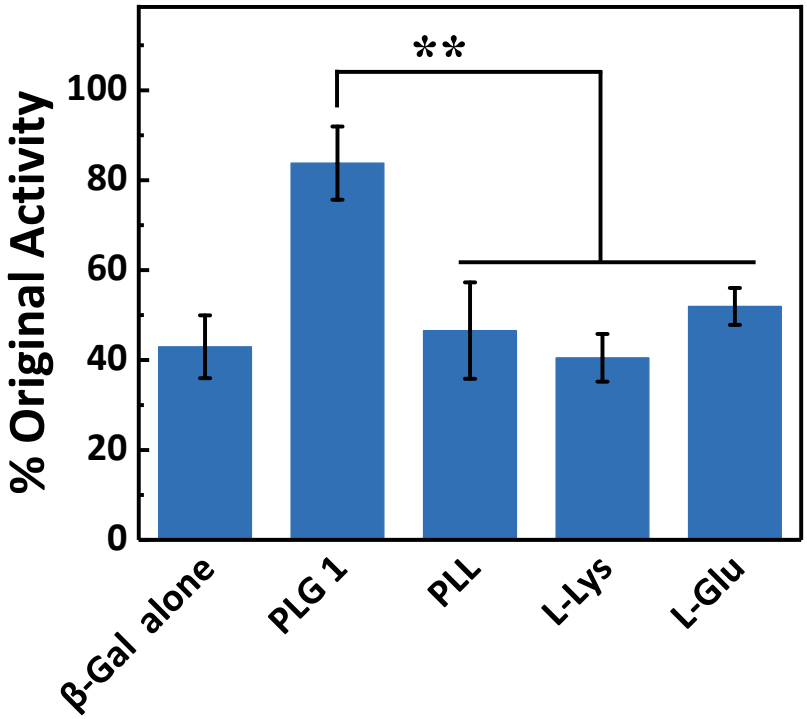

Fig. S6 β-Gal activity after lyophilization, with or without protection, compared to fresh β-Gal. 10 wt equiv of PLG1 (Mn=7700), poly-L-lysine (PLL) (Mn= 4700), L-lysine (L-Lys) (Mn=146.19), L-Glutamic acid (L-Glu) (Mn=147.13) as the control samples was used. (**) $p < 0.01$, n=3 for all samples.

HPLC analysis on NCA monomer conversion

Table S3 Monomer conversion and polymer yield of PLG and PLG-*r*-PLL

	monomer conversion	polymer yield (before deprotection)	polymer yield (deprotection)
PLG 1	**99.9%**	**92.8%**	**75.9%**
PLG 2	**99.8%**	**88.5%**	**77%**
PLG 3	**99.6%**	**93.9%**	**73.7%**
PLG 4	**98.7%**	**90.6%**	**77.1%**
PLG80-*r*-PLL20	**99.8%**	**86.2%**	**74.2%**
PLG60-*r*-PLL40	**99.9%**	**87.6%**	**75.5%**
PLG50-*r*-PLL50	**99.9%**	**87.7%**	**76.4%**
PLG40-*r*-PLL60	**99.9%**	**89.7%**	**72.6%**
PLG20-*r*-PLL80	**99.7%**	**88.6%**	**71.8%**

Reverse phase HPLC analysis on NCA monomer conversion used the combination of water (eluent A) and acetonitrile (eluent B) as the mobile phase. NCA monomer and polymers were analyzed by integration of the absorbance peak at 210 nm. Then the NCA solution was injected into a HPLC using 100% acetonitrile as the mobile phase to calculate the monomer conversion using an internal standard triphenylmethane (TPM). The monomer conversion were calculated by the equation(% conversion = $\frac{\frac{A_{monomer}}{A_{TPM}} - \frac{A_{polymer}}{A_{TPM}}}{\frac{A_{polymer}}{A_{TPM}}} \times 100$). where $A_{monomer}/A_{TPM}$ is the integrated areas ratios of HPLC peaks for NCA monomer and TPM, $A_{polymer}/A_{TPM}$ is the integrated areas ratios of HPLC peaks for polymer and TPM. The monomer conversion were calculated in table S3 and the detailed HPLC spectra were in Fig.S7- Fig.S15.

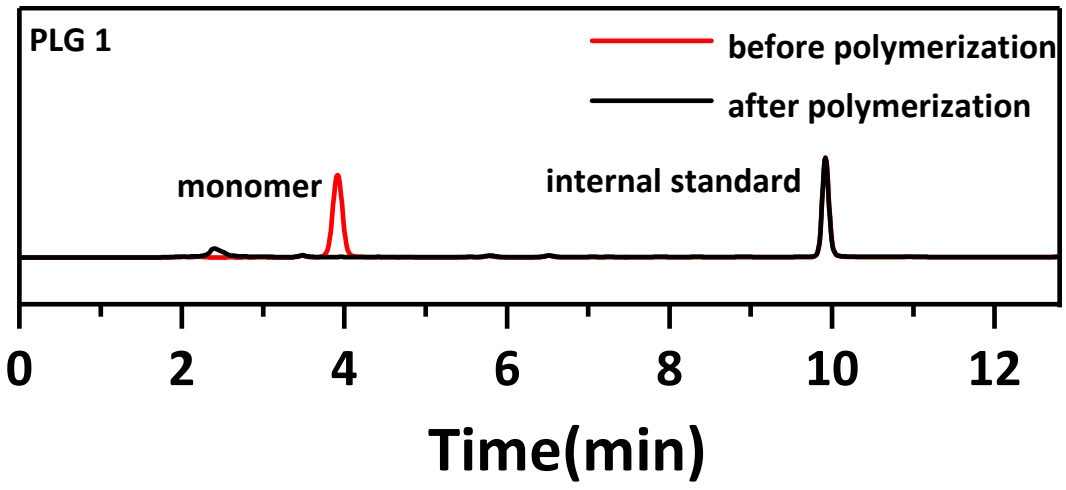

Fig. S7 Analysis on PLG 1 monomer conversion.

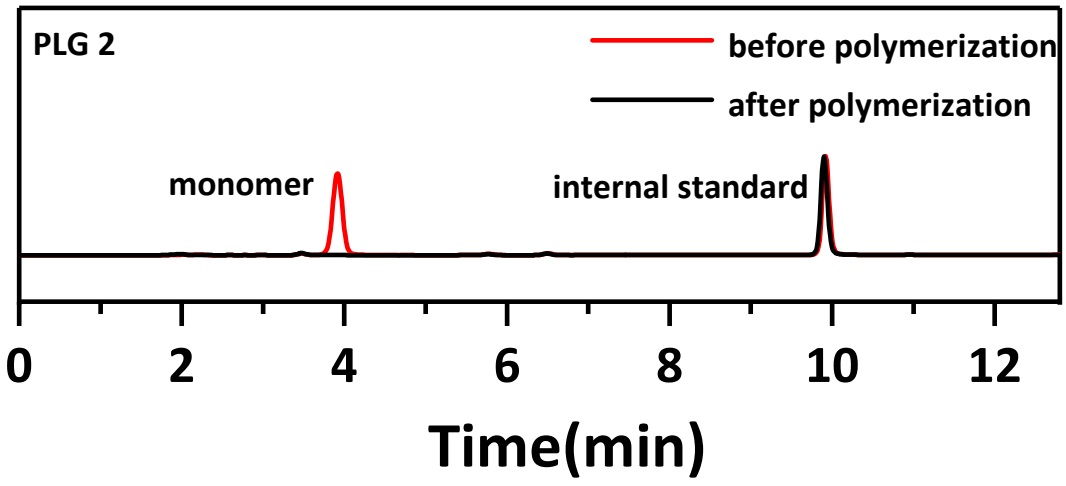

Fig. S8 Analysis on PLG 2 monomer conversion.

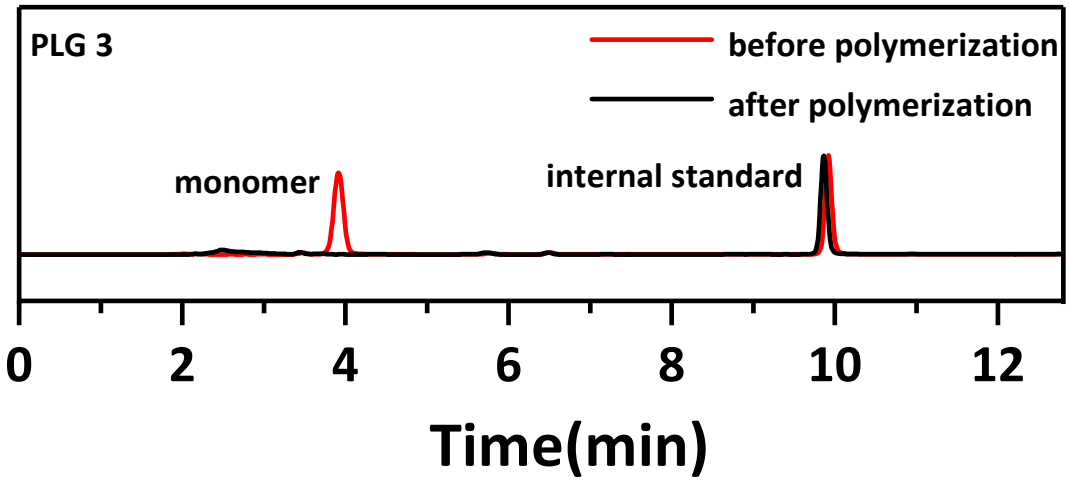

Fig. S9 Analysis on PLG 3 monomer conversion.

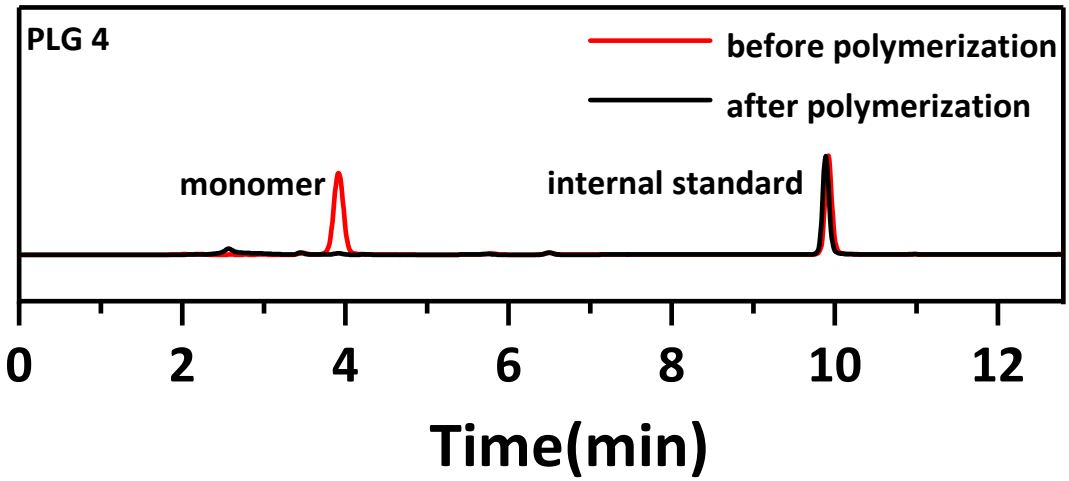

Fig. S10 Analysis on PLG 4 monomer conversion.

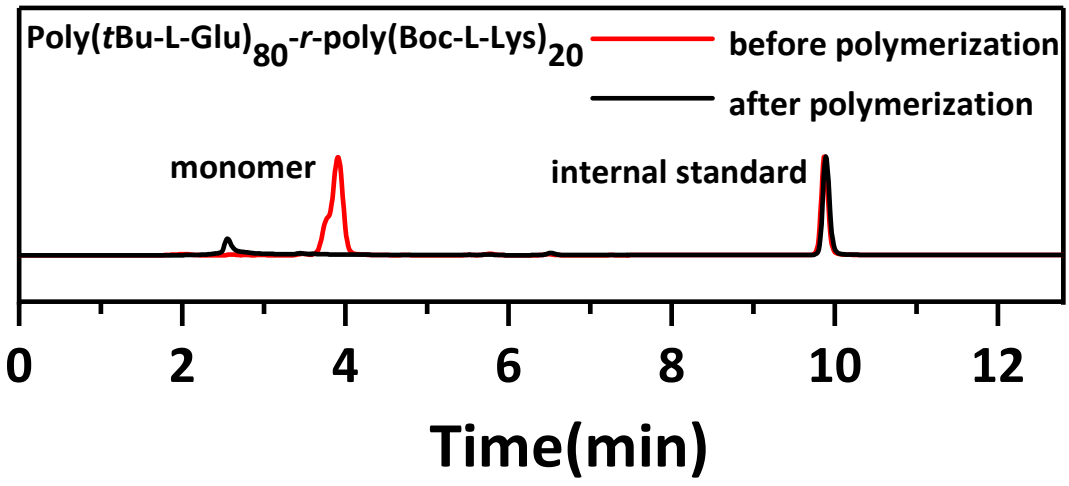

Fig. S11 Analysis on poly(*t*Bu-L-Glu)$_{80}$-*r*-poly(Boc-L-Lys)$_{20}$ monomer conversion.

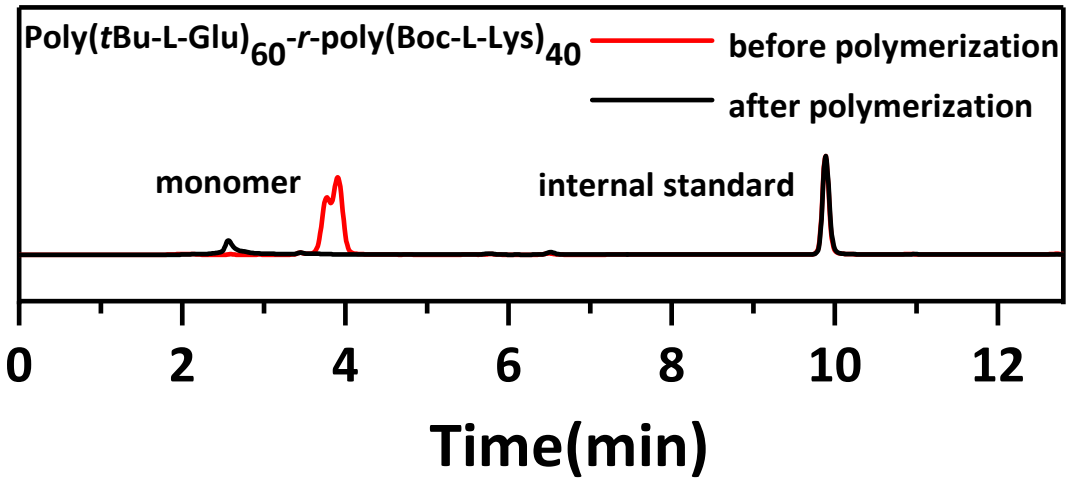

Fig. S12 Analysis on poly(*t*Bu-L-Glu)$_{60}$-*r*-poly(Boc-L-Lys)$_{40}$ monomer conversion.

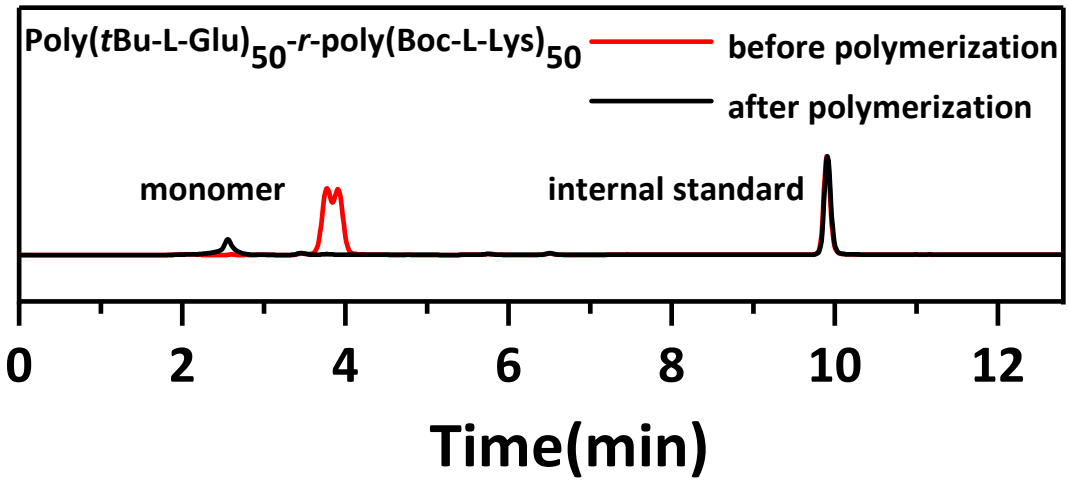

Fig. S13 Analysis on poly(*t*Bu-L-Glu)$_{50}$-*r*-poly(Boc-L-Lys)$_{50}$ monomer conversion.

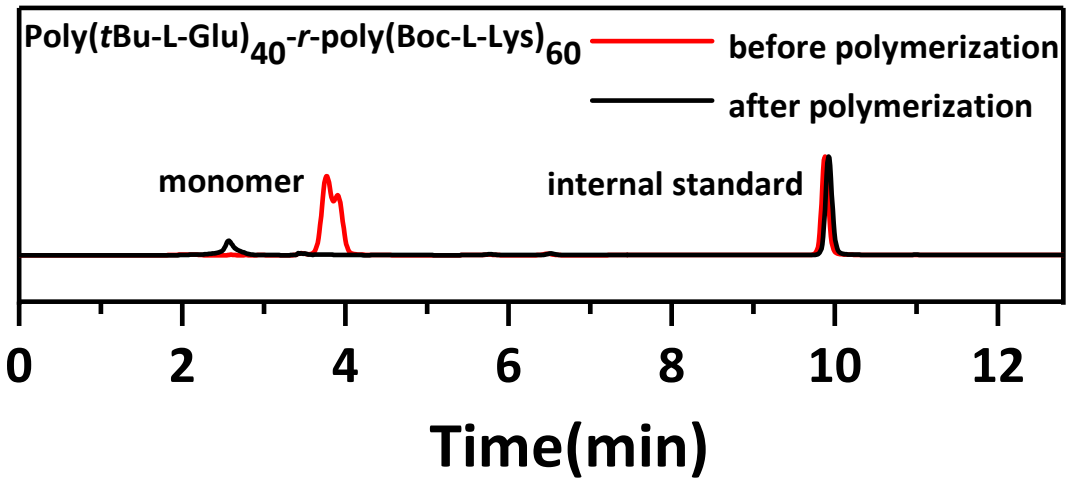

Fig. S14 Analysis on poly(*t*Bu-L-Glu)$_{40}$-*r*-poly(Boc-L-Lys)$_{60}$ monomer conversion.

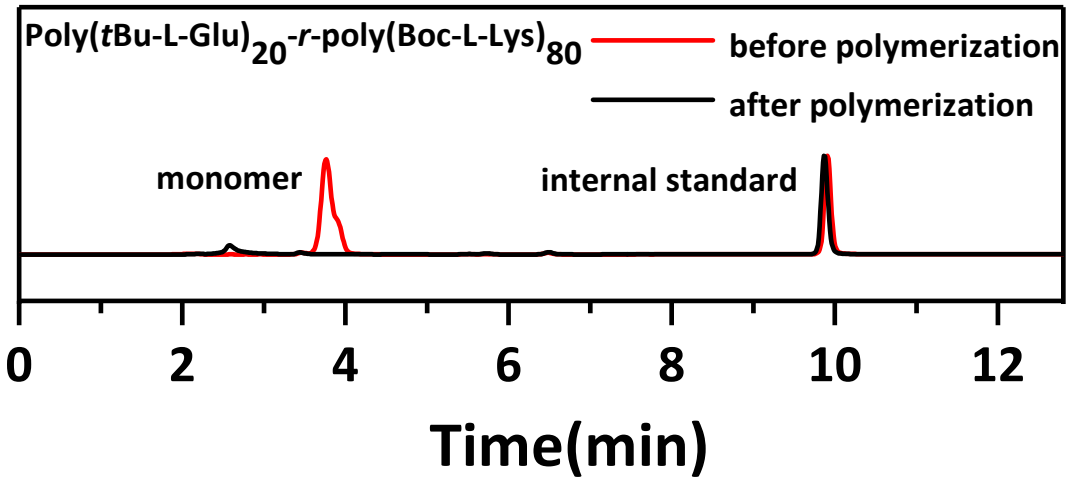

Fig. S15 Analysis on poly(*t*Bu-L-Glu)$_{20}$-*r*-poly(Boc-L-Lys)$_{80}$ monomer conversion.

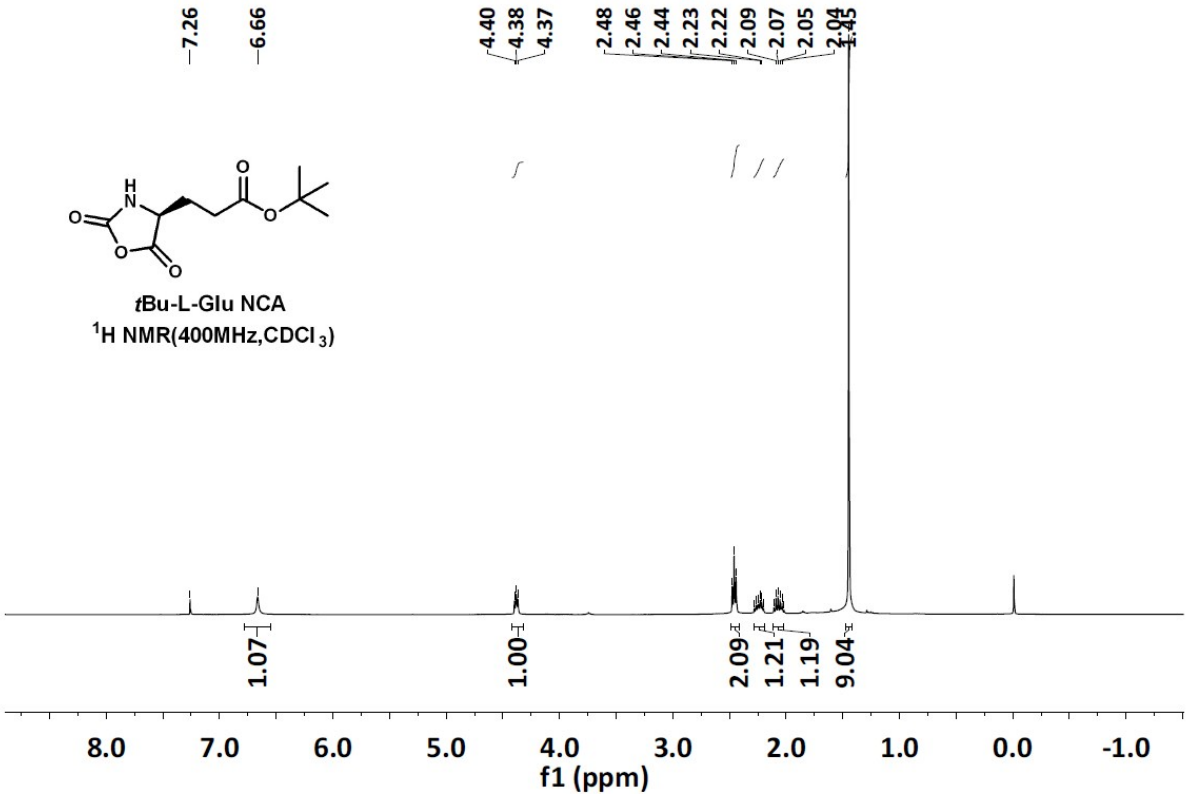

Fig. S16 ^{1}H NMR of *t*Bu-L-Glu NCA in $CDCl_3$, 400MHz.

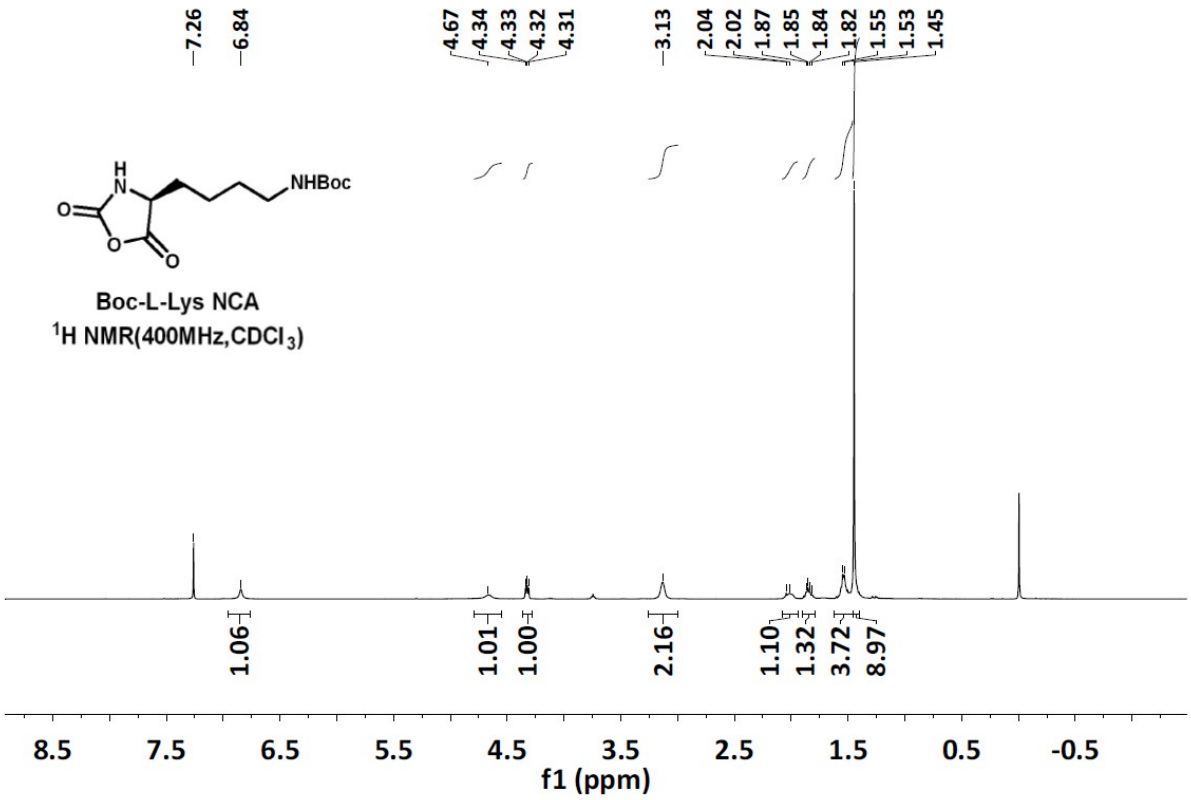

Fig. S17 ^{1}H NMR of Boc-L-Lys NCA in $CDCl_3$, 400MHz.

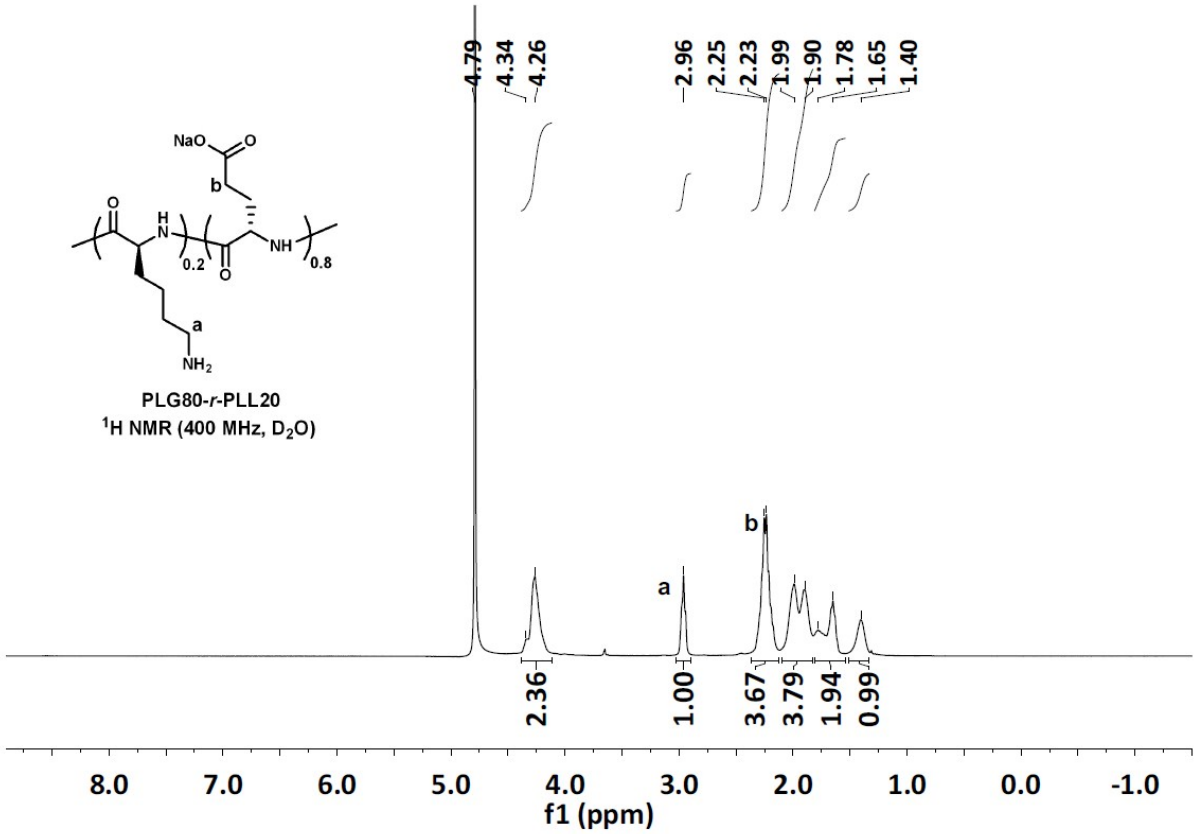

Fig. S18 ^{1}H NMR of PLG80-*r*-PLL20 in D_2O, 400MHz.

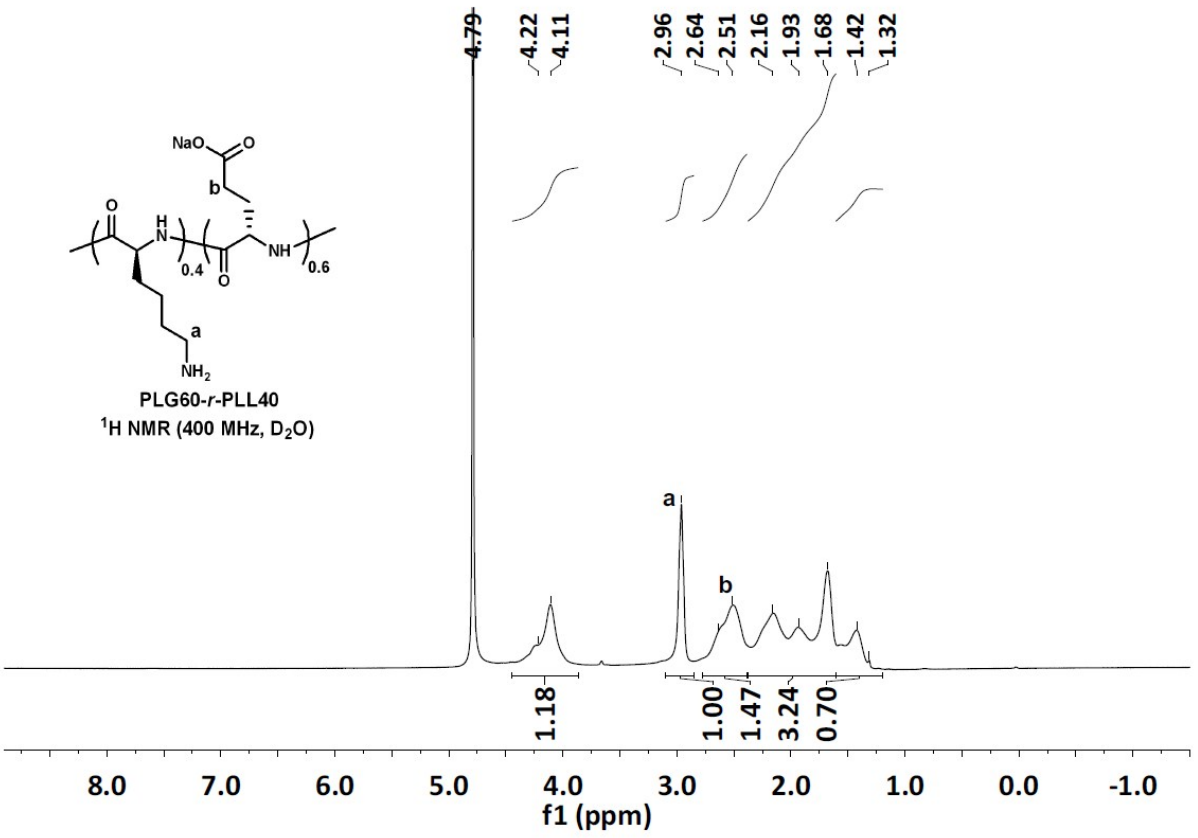

Fig. S19 ^{1}H NMR of PLG60-*r*-PLL40 in D_2O, 400MHz.

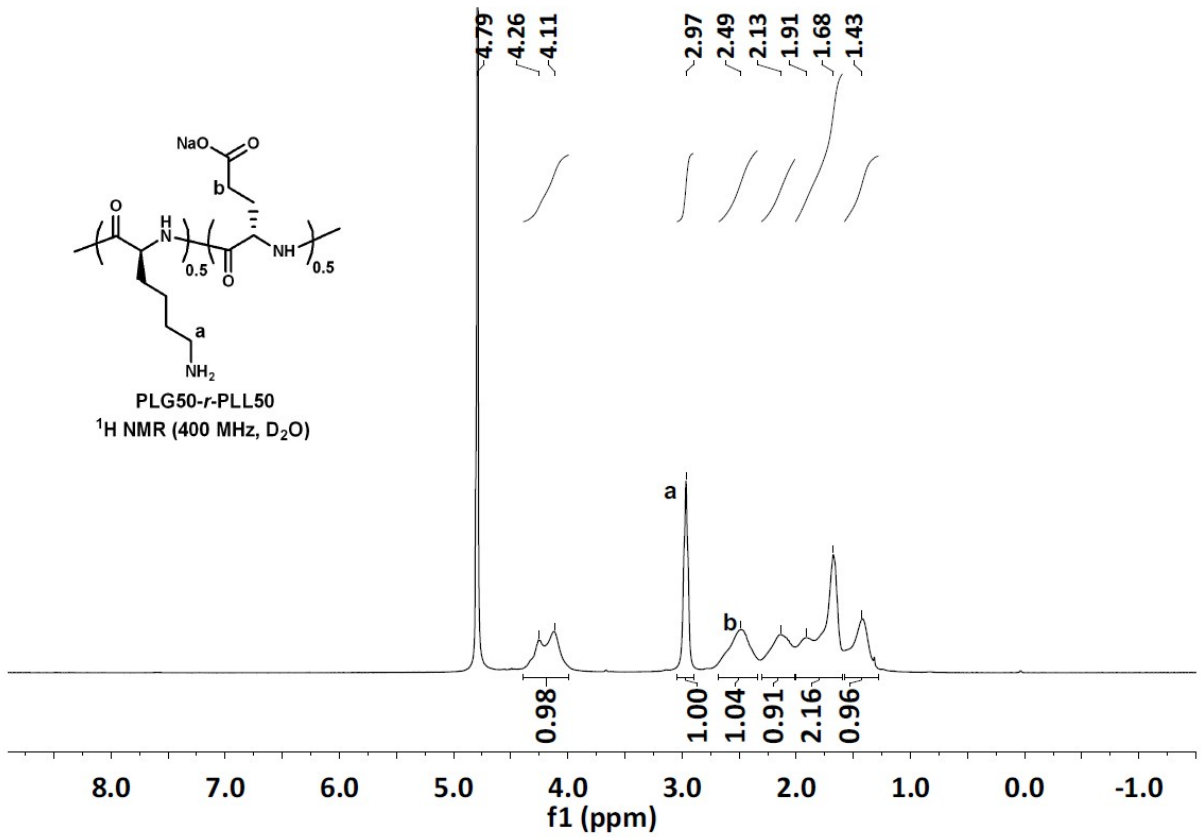

Fig. S20 ^{1}H NMR of PLG50-*r*-PLL50 in D_2O, 400MHz.

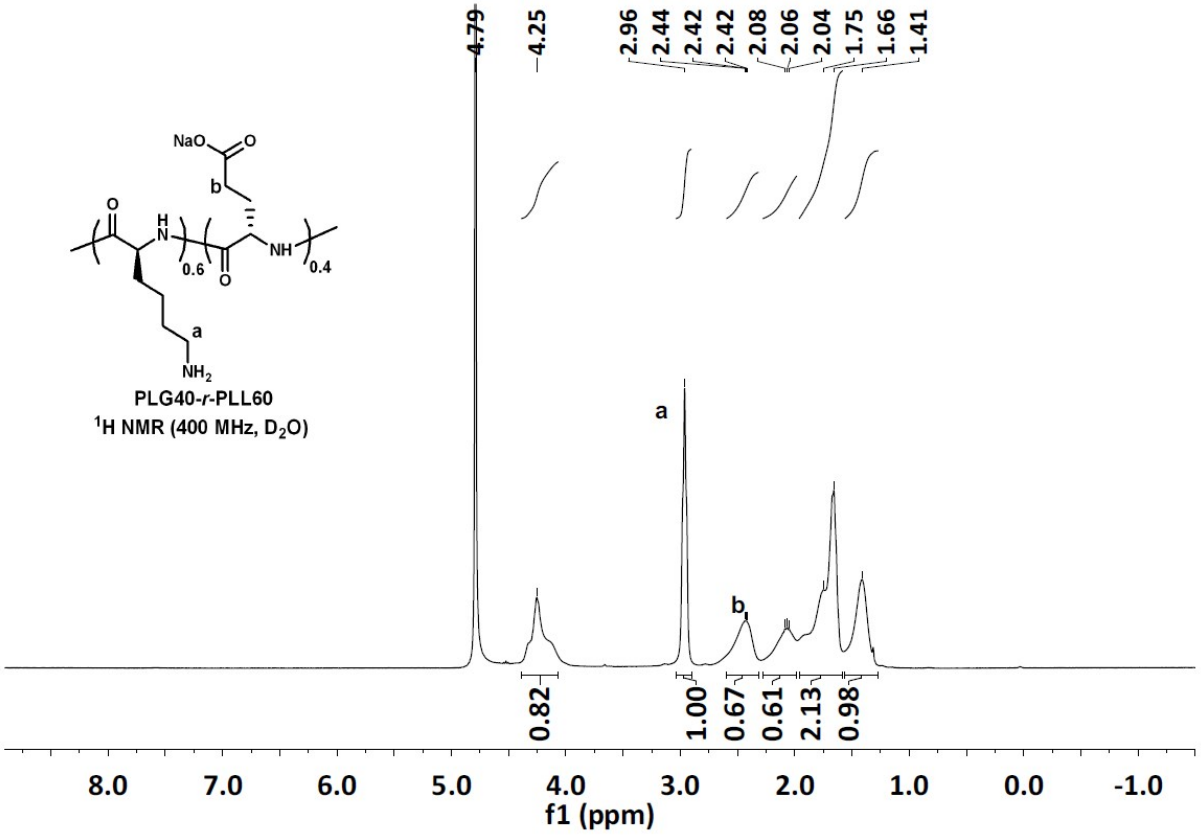

Fig. S21 ^{1}H NMR of PLG40-*r*-PLL60 in D_2O, 400MHz.

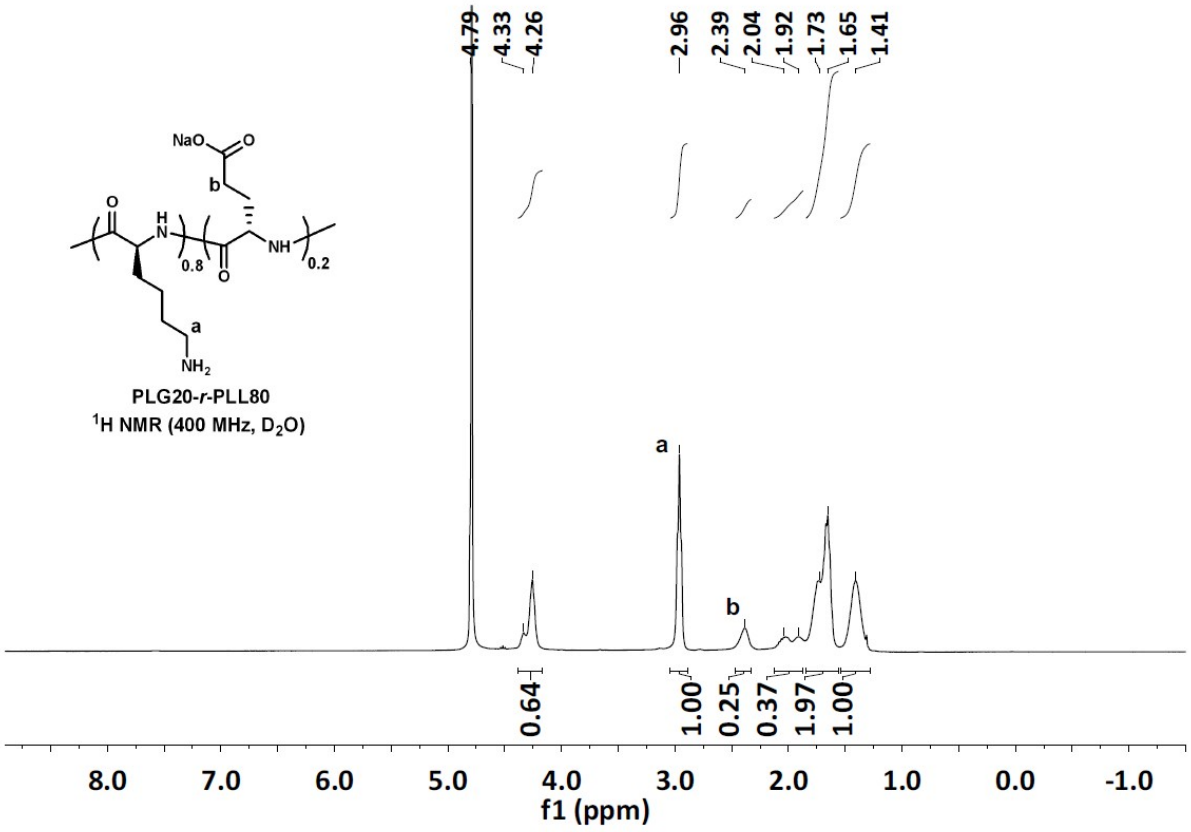

Fig. S22 ^{1}H NMR of PLG20-*r*-PLL80 in D_2O, 400MHz.

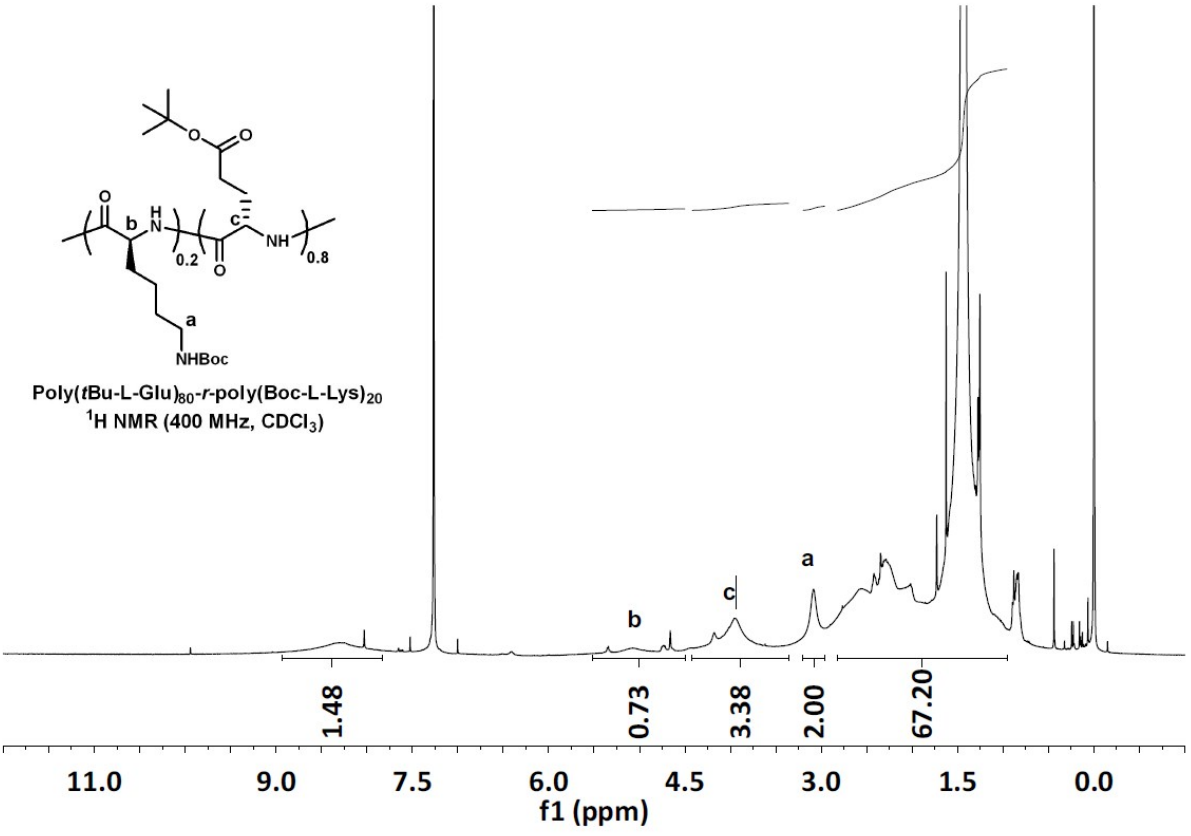

Fig. S23 ^{1}H NMR of Poly(*t*Bu-L-Glu)$_{80}$-*r*-poly(Boc-L-Lys)$_{20}$ in CDCl3, 400MHz.

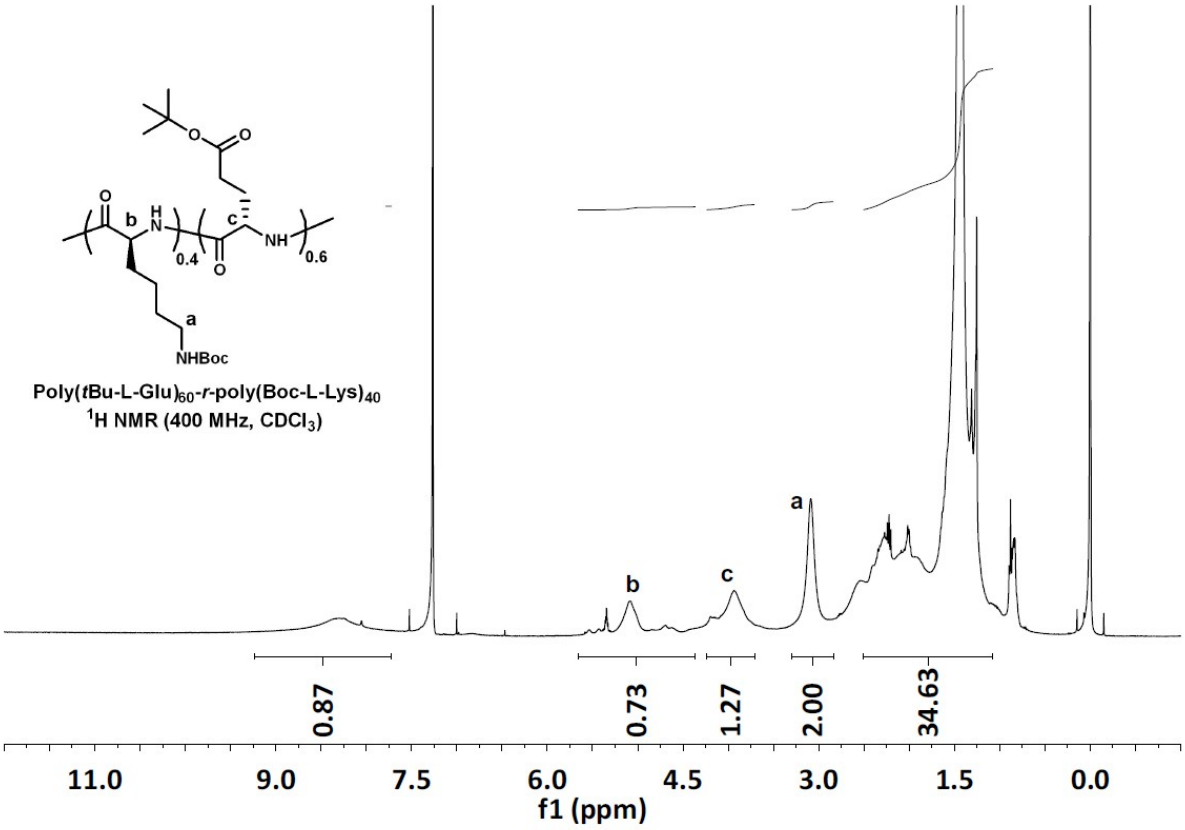

Fig. S24 ^{1}H NMR of Poly(*t*Bu-L-Glu)$_{60}$-*r*-poly(Boc-L-Lys)$_{40}$ in CDCl3, 400MHz.

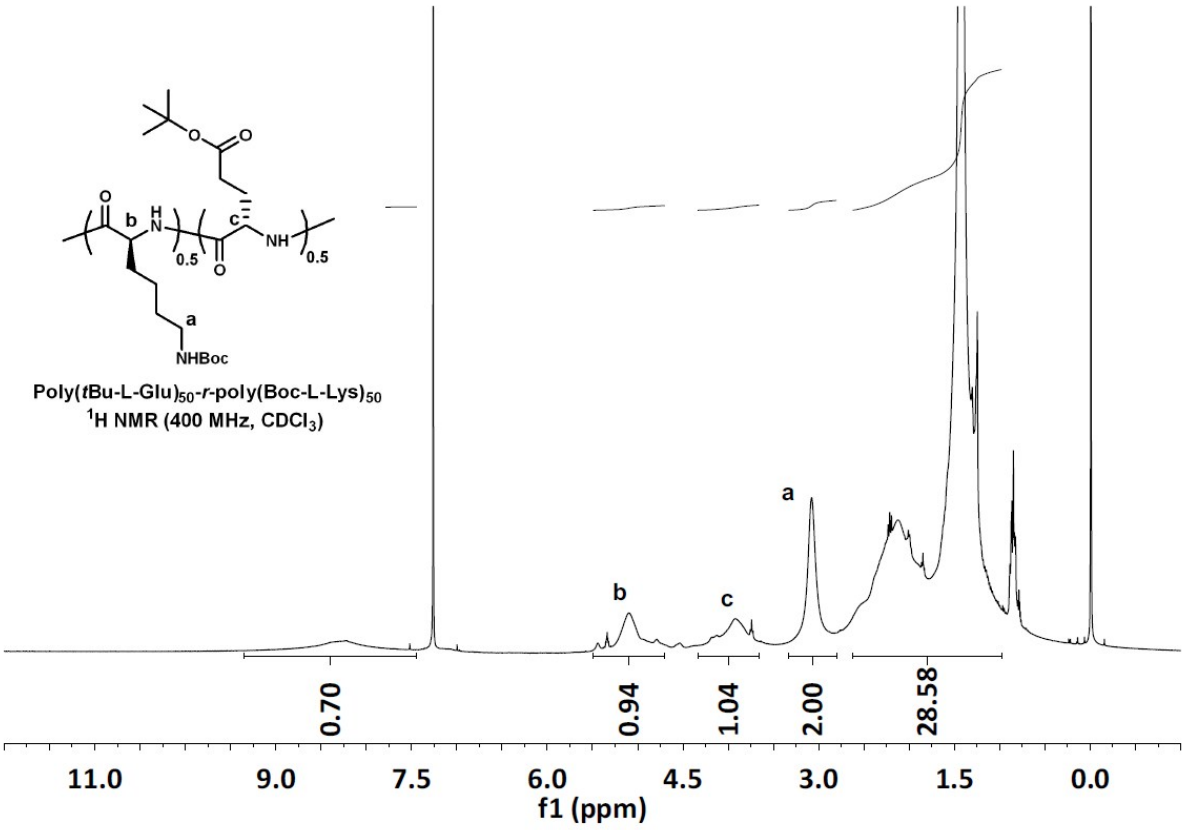

Fig. S25 ^{1}H NMR of Poly(*t*Bu-L-Glu)$_{50}$-*r*-poly(Boc-L-Lys)$_{50}$ in CDCl3, 400MHz.

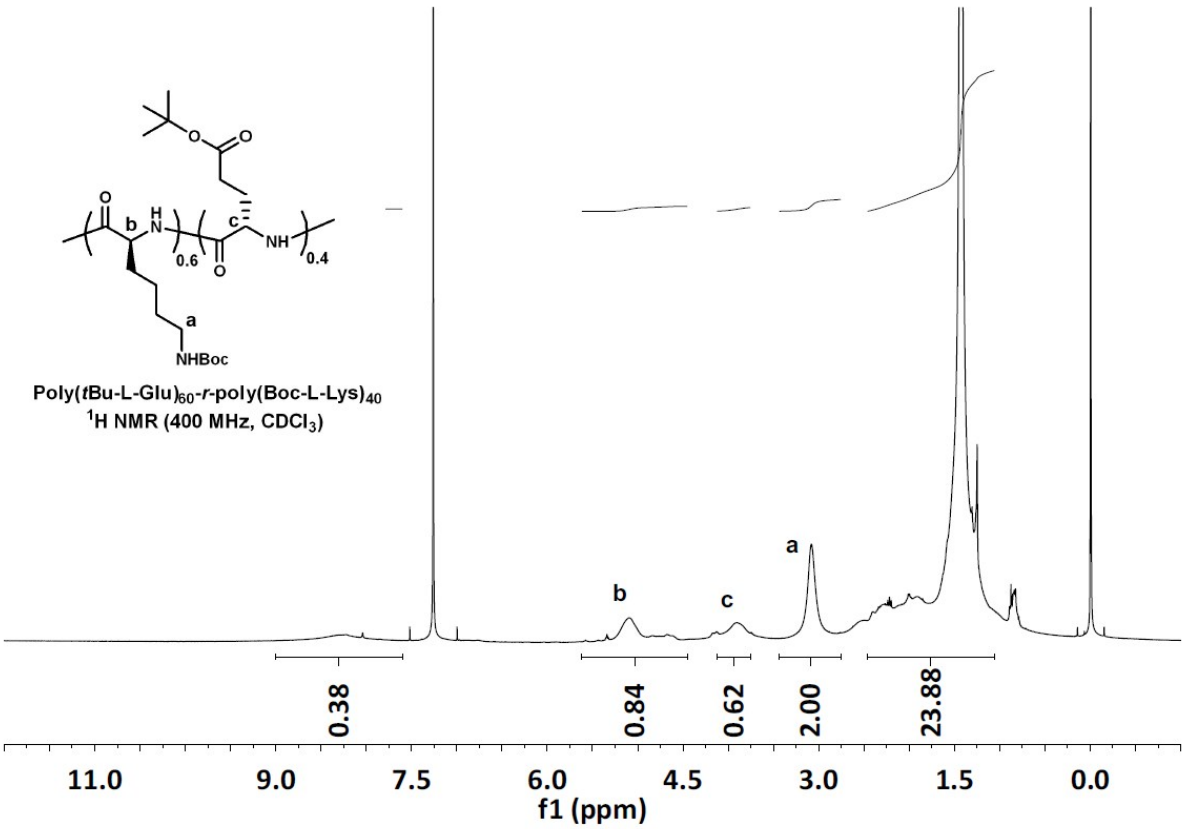

Fig. S26 ^{1}H NMR of Poly(*t*Bu-L-Glu)$_{40}$-*r*-poly(Boc-L-Lys)$_{60}$ in CDCl3, 400MHz.

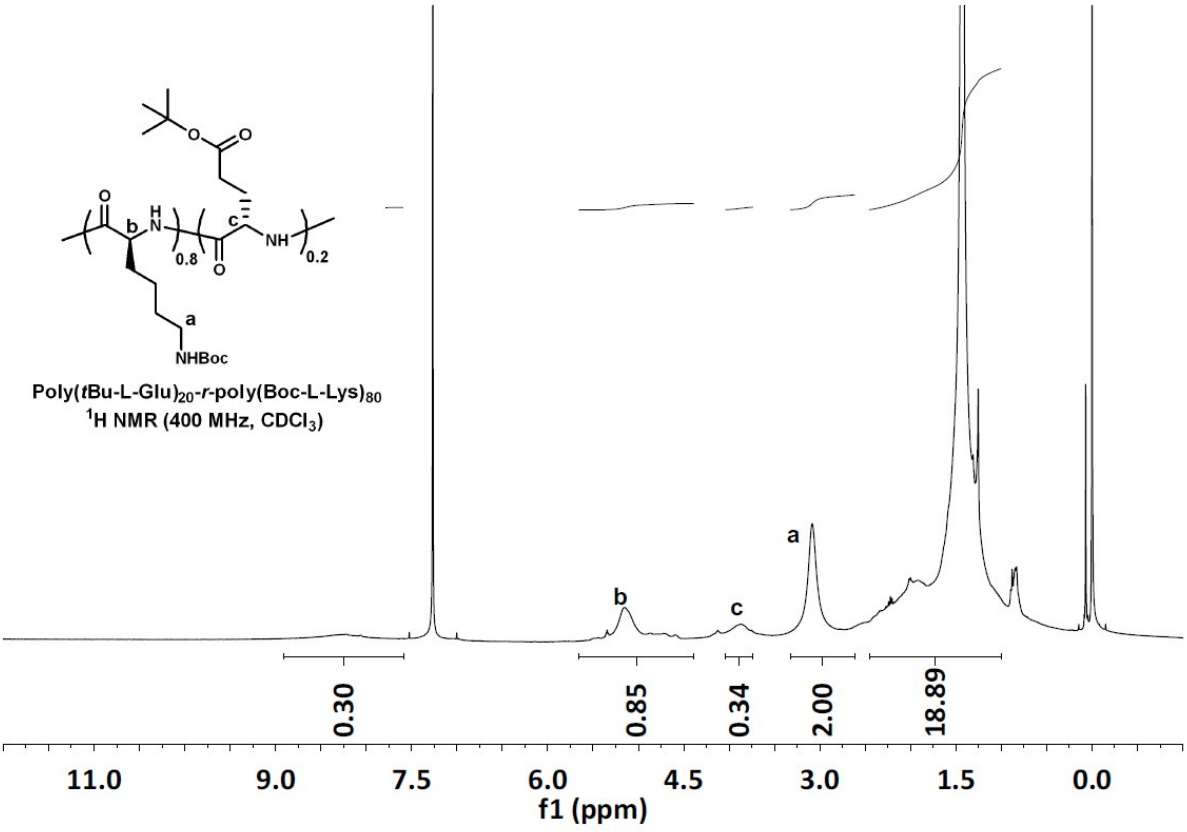

Fig. S27 ^{1}H NMR of Poly(*t*Bu-L-Glu)$_{20}$-*r*-poly(Boc-L-Lys)$_{80}$ in CDCl3, 400MHz.

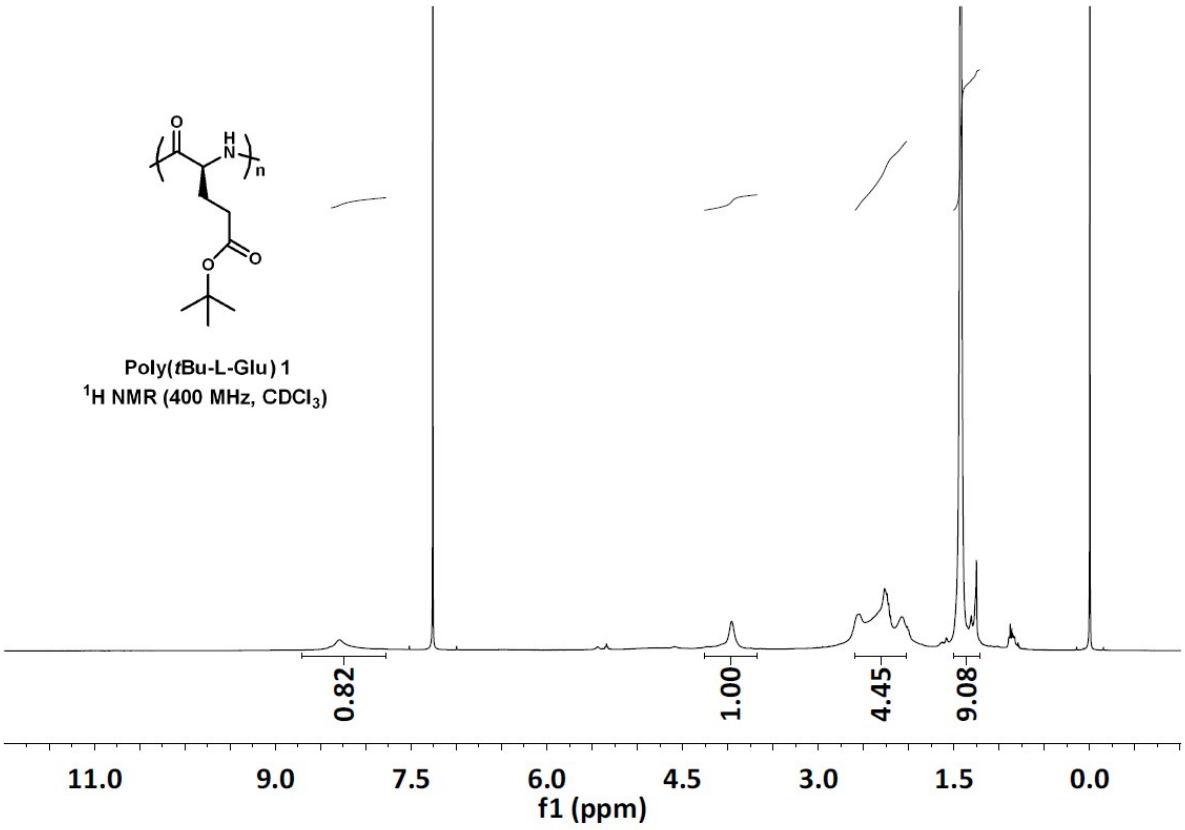

Fig. S28 ^{1}H NMR of Poly(*t*Bu-L-Glu) 1 in CDCl3, 400MHz.

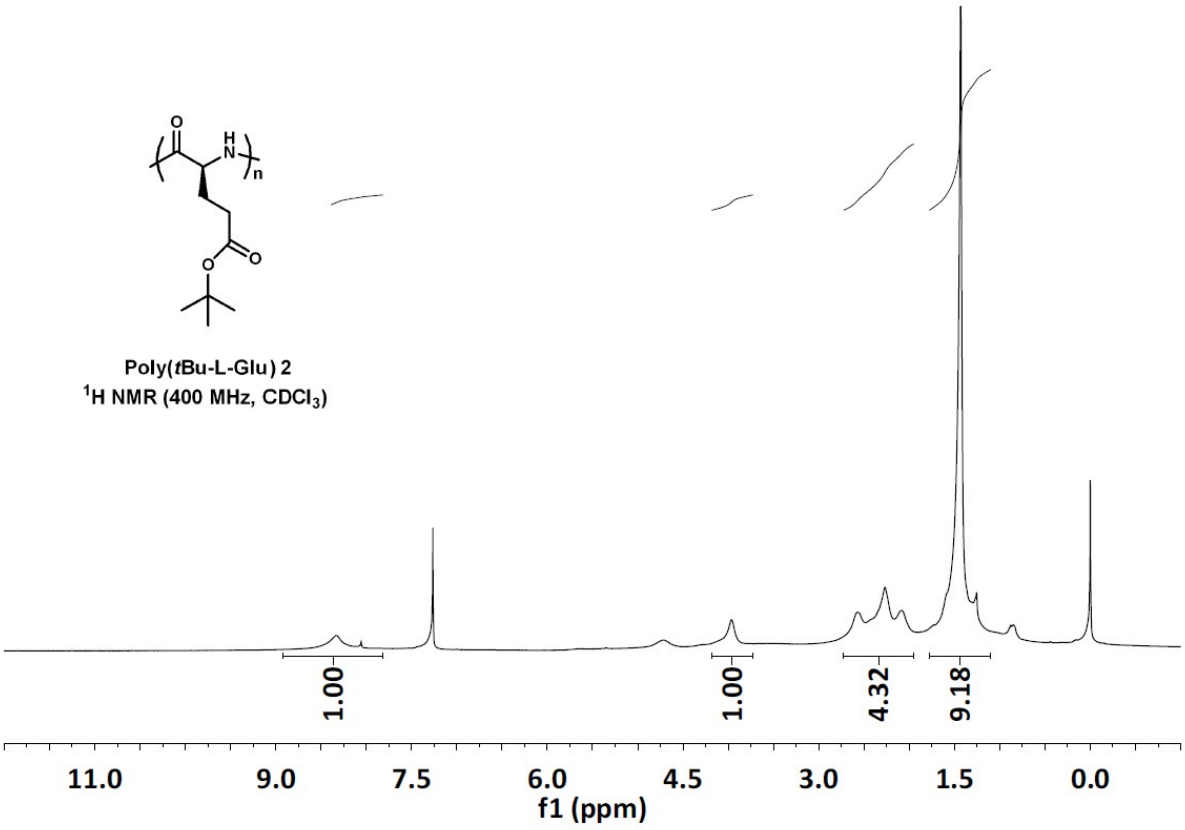

Fig. S29 ^{1}H NMR of Poly(*t*Bu-L-Glu) 2 in CDCl3, 400MHz.

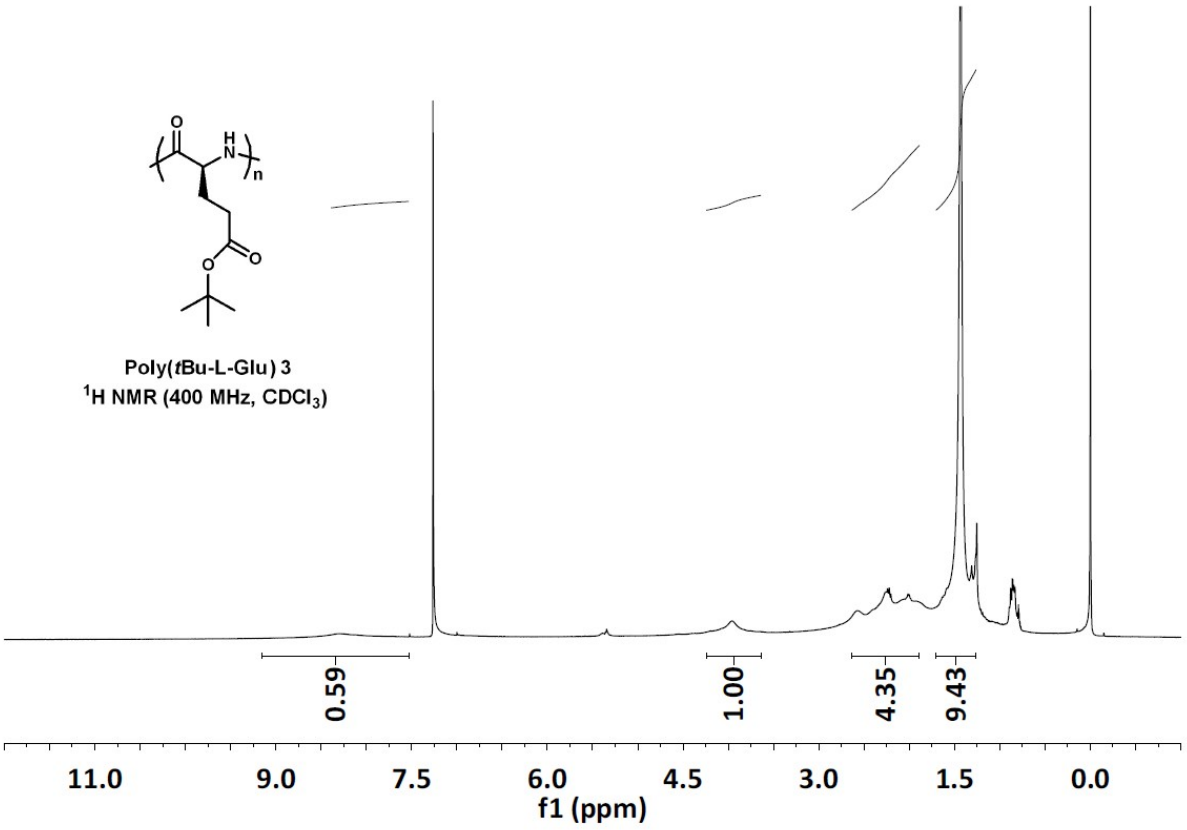

Fig. S30 ^{1}H NMR of Poly(*t*Bu-L-Glu) 3 in CDCl3, 400MHz.

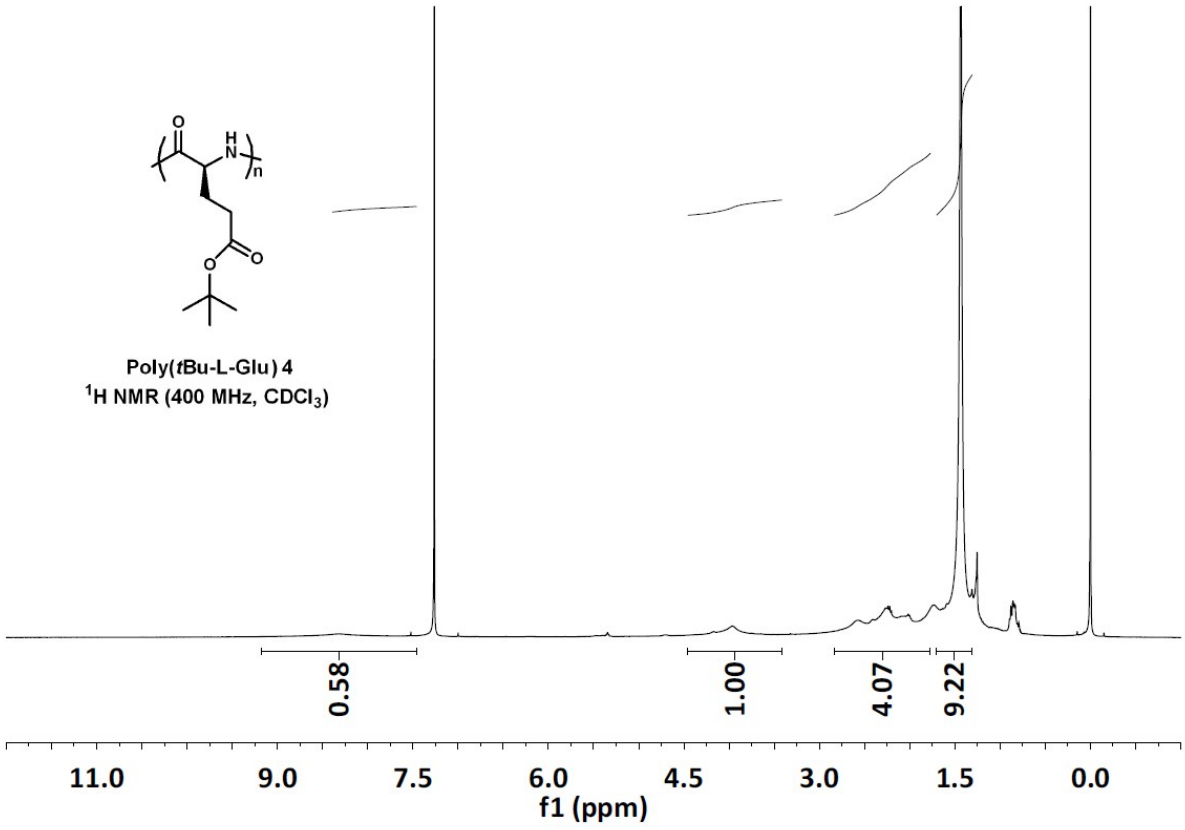

Fig. S31 ^{1}H NMR of Poly(*t*Bu-L-Glu) 4 in CDCl3, 400MHz.

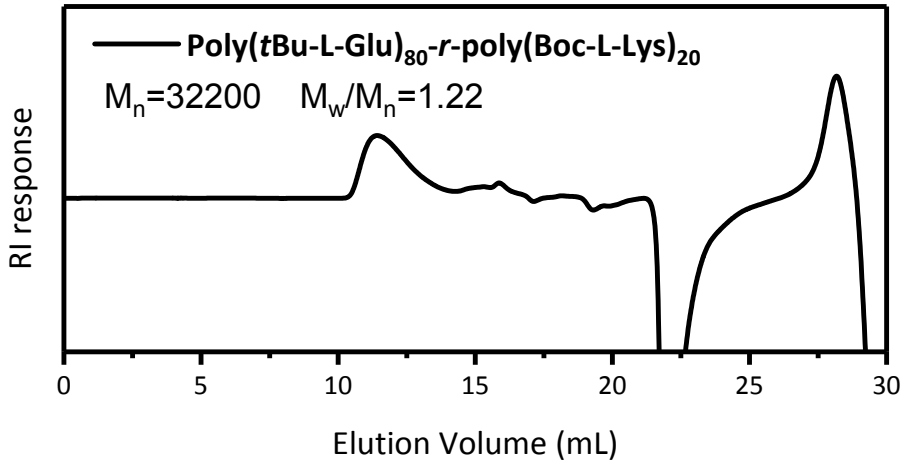

Fig. S32 GPC trace of PLG80-*r*-PLL20 at the sidechain protected stage using DMF as the mobile phase.

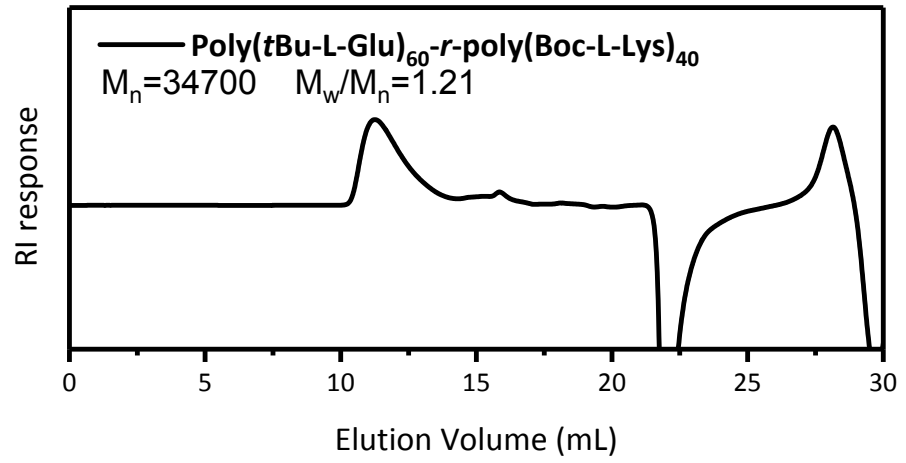

Fig. S33 GPC trace of PLG60-*r*-PLL40 at the sidechain protected stage using DMF as the mobile phase.

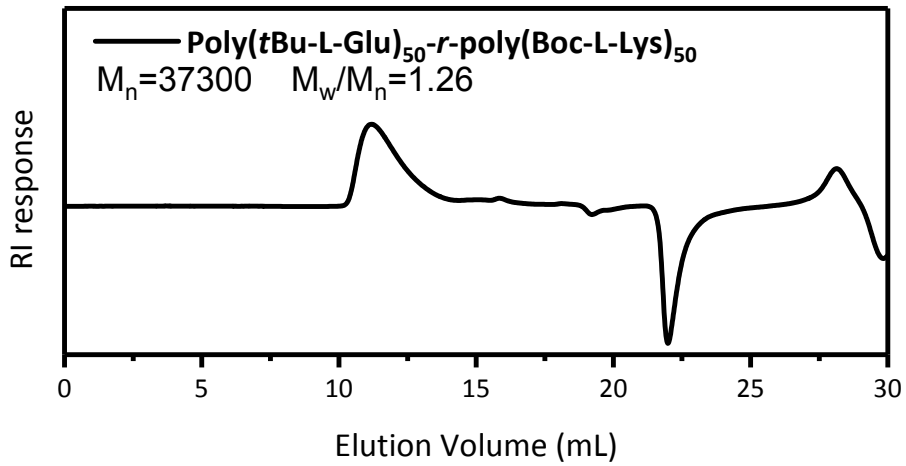

Fig. S34 GPC trace of PLG50-*r*-PLL50 at the sidechain protected stage using DMF as the mobile phase.

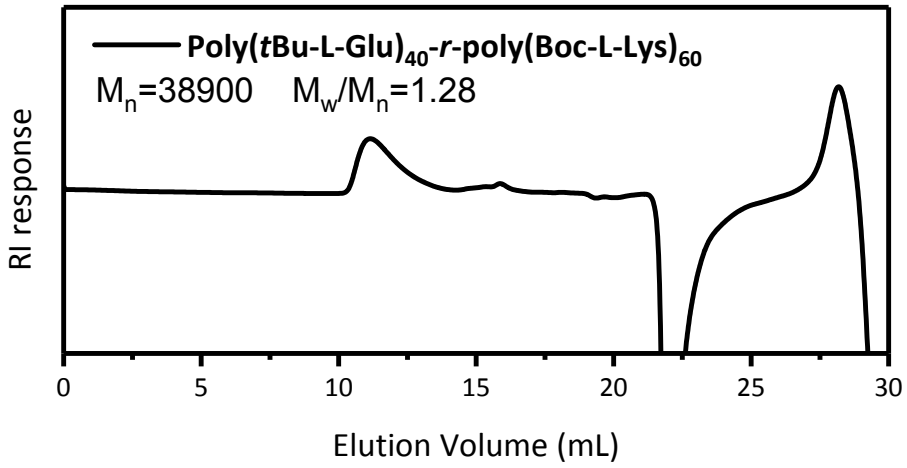

Fig. S35 GPC trace of PLG40-*r*-PLL60 at the sidechain protected stage using DMF as the mobile phase.

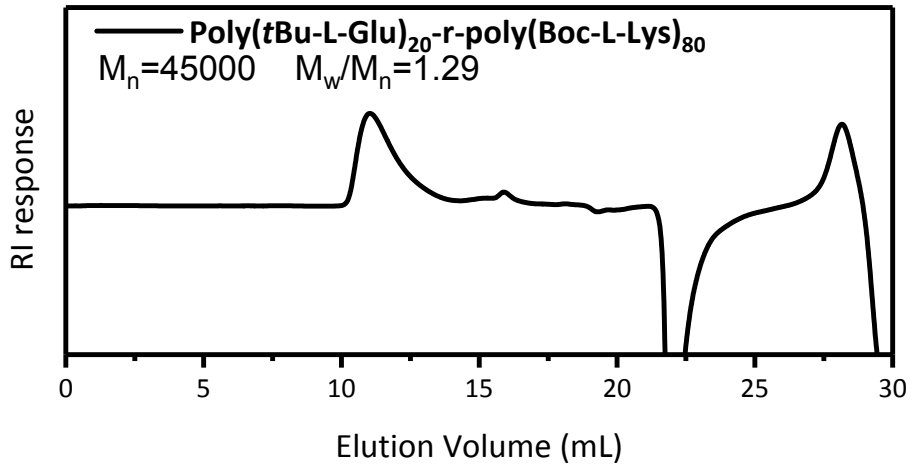

Fig. S36 GPC trace of PLG20-*r*-PLL80 at the sidechain protected stage using DMF as the mobile phase.

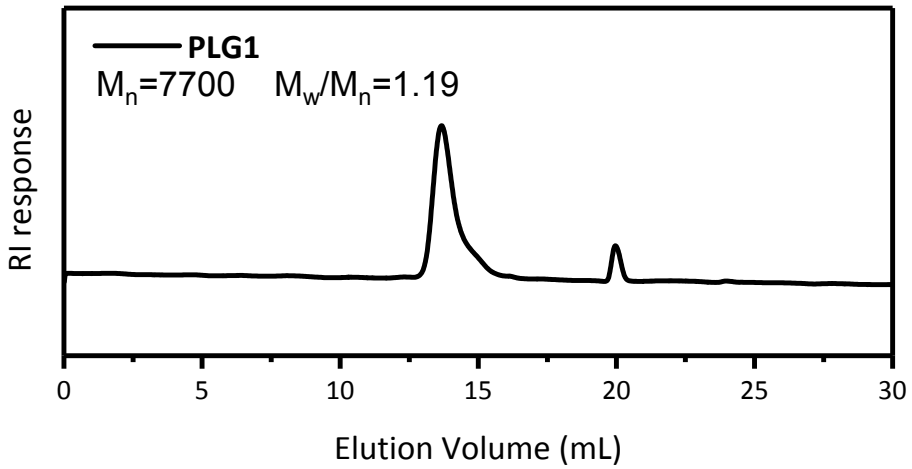

Fig. S37 GPC trace of PLG1 at the deprotected stage using water as the mobile phase.

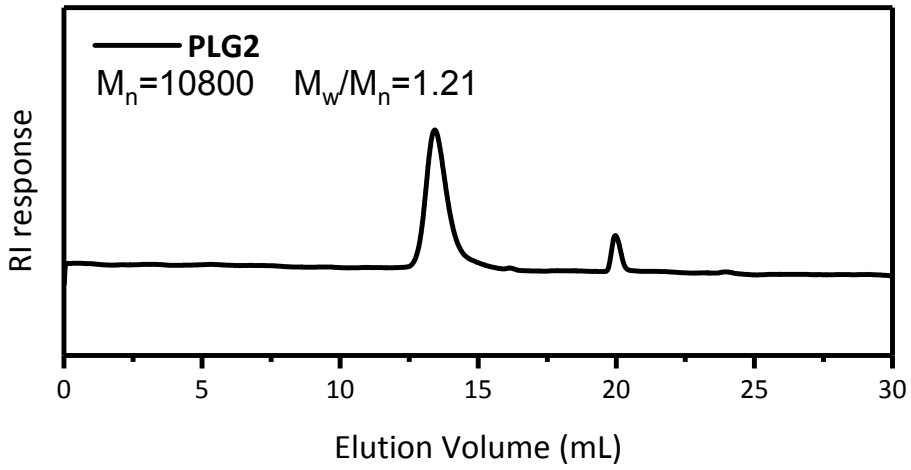

Fig. S38 GPC trace of PLG2 at the deprotected stage using water as the mobile phase.

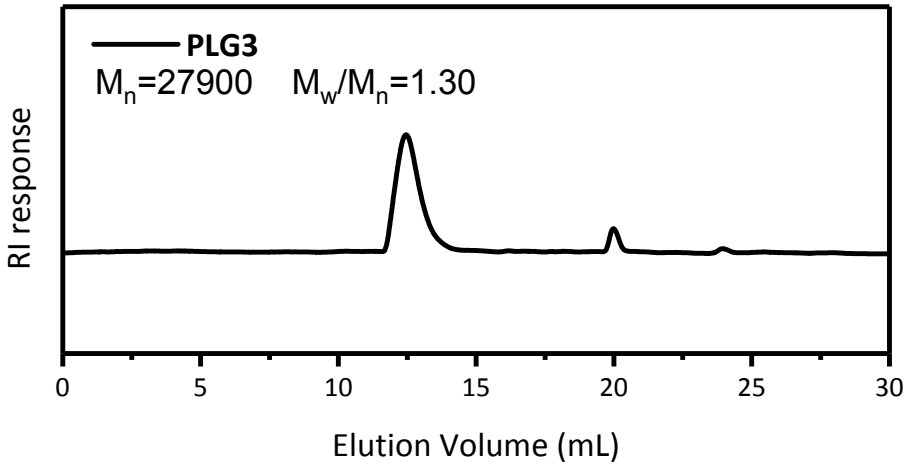

Fig. S39 GPC trace of PLG3 at the deprotected stage using water as the mobile phase.

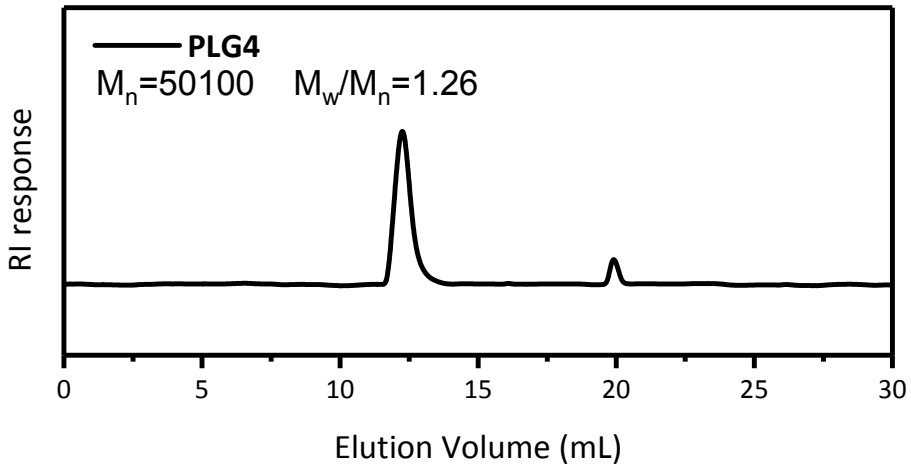

Fig. S40 GPC trace of PLG4 at the deprotected stage using water as the mobile phase.

References

1. C. Lavilla, M. Byrne and A. Heise, *Macromolecules*, 2016, **49**, 2942-2947.
2. D. Zhang, Y. Qian, S. Zhang, P. Ma, Q. Zhang, N. Shao, F. Qi, J. Xie, C. Dai, R. Zhou, Z. Qiao, W. Zhang, S. Chen and R. Liu, *Sci. China-Mater.*, 2018, **62**, 604-610.
3. Y. Wu, D. Zhang, P. Ma, R. Zhou, L. Hua and R. Liu, *Nat Commun*, 2018, **9**, 5297-5307.