# Supporting Information

# Site-Specific Conjugation of RAFT Polymers to Proteins via Expressed Protein Ligation

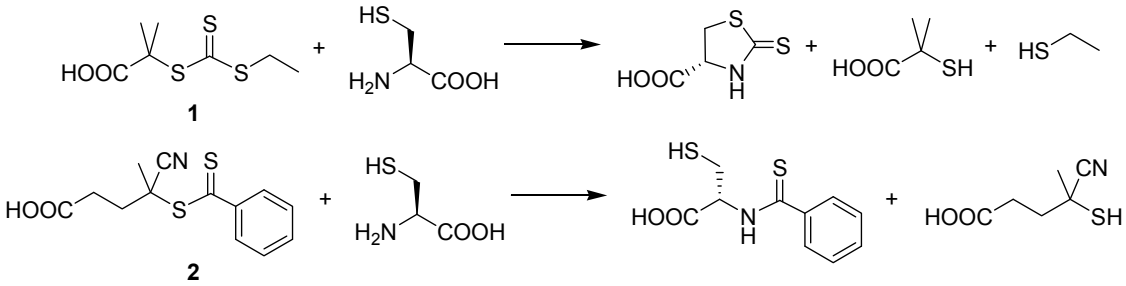
**Yan Xia, Shengchang Tang, and Bradley D. Olsen**

**Materials.** 2-Ethylsulfanylthiocarbonylsulfanyl-2-methylpropionic acid[1] and *N*-Boc-*S*-trityl-*N*-2-aminoethyl-L-cysteinamide[2] were synthesized as previously reported. *N*,*N*'-dimethylacrylamide (DMA) and poly(ethylene glycol) methyl ether methacrylate (PEGMA, $M_n$ = 475 and DP = 8~9) were passed through an alumina column to remove the inhibitor. *N*-isopropylacrylamide (NIPAM) was purified by sublimation. Azobisisobutylonitrile (AIBN) was recrystallized twice from ethanol. All other materials were obtained from commercial sources and used as received. The intein gene (pTWIN2) was purchased from New England Biolabs.

**Instrumentation.** $^1$H and $^{13}$C NMR spectra were recorded in $CDCl_3$ using a Varian 300 spectrometer. GPC was carried out on an Agilent 1260 GPC system with three DMF ResiPore columns, a Wyatt Mini-DAWN TREOS 3-angle static light scattering detector, and a Wyatt Optilab T-rEX refractive index detector. DMF with 0.02 M LiBr was used as eluent at 1.0 mL/min at 70 °C. LC-MS was performed on Waters ZQ 2000 mass spectrometer. UV/Vis spectra were measured using a Varian Cary 50 Bio UV-Vis spectrophotometer. CD spectra of protein and bioconjugate solutions (4.5 μM in 20 mM Tris buffer, pH = 8.0) were measured using an Aviv model 202 circular dichroism spectrometer at 25 °C. DLS of bioconjugate solutions (150 μM in 20 mM Tris buffer, pH

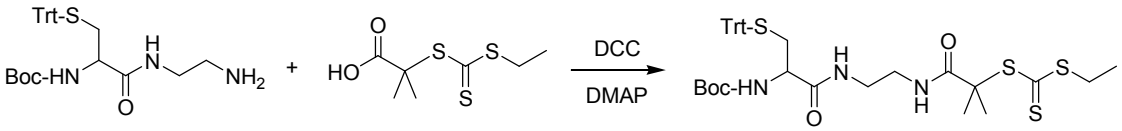
= 8.0) was performed on a Wyatt DynaPro Nanostar using 658 nm laser at 25 and 50 °C. A JEOL 200CX TEM was used to obtain bright field images of micelles resulting from bioconjugates using an accelerating voltage of 120 kV and a tungsten filament. Images were captured using a digital camera in a fixed bottom mount configuration. Dilute solutions of bioconjugates (0.1-0.2 wt % in 20 mM Tris buffer, pH = 8.0, at 50 °C) was drop cast on Formvar-coated TEM grids and stained with 1% uranyl acetate.

**Putative decomposition pathways of common RAFT agents with cysteine.**




**1** and **2** were used as model compounds to study the compatibility of RAFT agents with cysteine under common NCL conditions. **1** or **2** was dissolved in $H_2O$/MeOH (80/20 v/v) at 30 mM, and pH was adjusted to 8 at 25 °C. Upon addition of 3 eq. of cysteine dissolved in $H_2O$, rapid fading of color was observed, indicating reaction of **1** or **2** with cysteine. After stirring for 1 h, the reaction mixture was withdrawn for LC-MS analysis, which suggested the formation of 5-carboxythiazolidine-2-thione or *N*-thiobenzoylcysteine resulting from the decomposition of **1** and **2**, respectively (Fig S1).

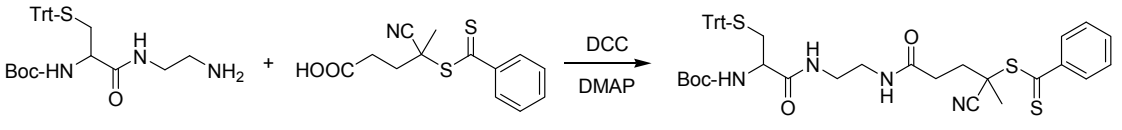
**Synthesis of RAFT agent *N*-Boc-*S*-Trityl-*N*-[2-[(2-ethylsulfanylthiocarbonylsulfanyl-2-methylpropionyl)amino]ethyl]-L-cysteinamide 3.**




**3**

To a round-bottom flask was added 2-Ethylsulfanylthiocarbonylsulfanyl-2-methylpropionic acid (0.66 g, 2.9 mmol), *N*,*N*'-Dicyclohexylcarbodiimide (DCC) (0.61g, 2.9 mmol) and 4-dimethylaminopyridine (DMAP) (72 mg, 0.58 mmol), followed by 20 mL DCM. N-Boc-S-Trityl-N-2-aminoethyl-L-cysteinamide (1.5 g, 2.9 mmol) was then added to the stirring solution. The reaction mixture was allowed to stir at room temperature overnight. The reaction mixture was purified by silica gel chromatography (ethyl acetate/hexanes, 4:6 v/v) to give 1.1 g **3** as a yellow solid (52% yield). $^{1}$H NMR (300 MHz, $CDCl_3$): δ 7.42-7.19 (15H, m), 6.94 (1H, br), 6.39 (1H, br), 4.84 (1H, d, J=6 Hz), 3.86 (1H, ddd, J = 6, 6, 6 Hz), 3.35-3.23 (6H, m), 2.72 (1H, dd, J = 12, 6 Hz ), 2.50 (1H, dd, J = 13.5, 4.5 Hz), 1.64 (6H, d, J = 6 Hz), 1.41 (9H, s), 1.30 (3H, t, J = 6 Hz). $^{13}$C NMR (75 MHz, $CDCl_3$): δ221.25, 177.94, 173.37, 171.38, 144.53, 129.78, 128.31, 127.15, 80.56, 77.73, 77.31, 76.89, 67.43, 57.21, 53.88, 40.51, 39.54, 34.32, 31.65, 28.54, 28.52, 26.07, 25.84, 13.04. HRMS (FAB+) m/z calcd for $C_{36}H_{46}N_3O_4S_4$ $[M+H]^+$ : 712.2366, found 712.2368.

**Synthesis of RAFT agent *N*-Boc-*S*-Trityl-*N*-[2-[(4-cyano-4-(phenylcarbonothioylthio)pentanoyl) amino]ethyl]-L-cysteinamide 4.**




**4**

**4** was synthesized in a similar manner to **3** and isolated as a deep pink solid (65% yield). $^{1}$H NMR (300 MHz, $CDCl_3$): δ 7.89 (2H, d, J = 9 Hz), 7.55 (1H, t, J = 6 Hz), 7.42-7.18 (17H, m), 6.55 (1H, br), 6.48 (1H, br), 4.88 (1H, d, J = 9 Hz), 3.77 (1H, ddd, J = 6, 6, 6Hz), 3.45-3.22 (4H, m), 2.77-2.70 (1H, dd, J = 12, 6 Hz), 2.64-2.32 (4H, m), 1.89 (3H, s), 1.40 (9H, s). $^{13}$C NMR (75 MHz, $CDCl_3$) δ 222.81, 177.94, 171.75, 171.39, 144.73, 144.51, 133.25, 129.73, 128.80, 128.36, 127.22, 126.91, 126.91, 118.95, 80.61, 77.85, 77.43, 77.00, 67.43, 54.03, 46.28, 39.83, 39.56, 34.21, 31.74, 31.70, 28.58, 24.29. HRMS (FAB+) *m/z* calcd for $C_{42}H_{47}N_4O_4S_3$ $[M+H]^+$ : 767.2754, found 767.2753.

**General procedure for RAFT.** RAFT polymerization of DMA and NIPAM was performed in acetonitrile at 60 $^{o}$C using **3** as the RAFT agent and AIBN as the initiator. The ratio between the RAFT agent 3 and AIBN was fixed at 5:1, and the monomer/CTA ratio was varied to obtain polymers with desired molecular weights. The reaction mixture was degassed by three freeze-pump-thaw cycles. After reaching 80 – 95% conversion, the reaction flask was immersed into liquid nitrogen and opened to air. Polymer was obtained by precipitating in diethylether, filtered and dried *in vacuo*. RAFT polymerization of PEGMA was performed in a similar manner in dioxane at 60 $^{o}$C with **4** used as the RAFT agent.

**Procedures for polymer end group removal.**

*Removal of trithiocarbonate via aminolysis.* 800 mg polymer was dissolved in 10 mL of anhydrous THF. A ten-fold molar excess of hexylamine was added to the polymer solution. After stirring overnight at room temperature, the reaction mixture became colorless. Polymer was recovered by precipitating in diethylether.

*Removal of trithiocarbonate via radical-induced reduction.* Using reaction conditions slightly modified from the literature,[3] a 25 mL Schlenk tube was charged with 800 mg polymer, 50 eq of hypophosphate 1-ethylpiperidine hypophosphite, 0.5 eq of AIBN, and 5 mL of acetonitrile. The reaction mixture was degassed by three cycles of freeze-pump-thaw. The removal of trithiocarbonate was performed at 85 $^{o}$C for 2 hours. The solvent was removed by evaporation. The residual solid was redissolved in 10 mL DCM and washed with 10 mL of brine three times. The organic phase was separated, dried over $Na_2SO_4$ and concentrated under vacuum. Polymer was obtained by precipitating in diethylether.

*Deprotection of t-Boc and Trt groups.* 600 mg polymer was dissolved in 8 mL DCM in a 20 mL scintillation vial. 1 mL of TFA was added, and the reaction mixture turned yellow. Triethylsilane was gradually added to the reaction until the yellow color disappeared. Solvent and excess reagents were then removed by evaporation. The residual solid was redissolved in a small amount of acetone and precipitated in diethylether.

**Cloning and Protein Expression.**

A gene encoding for GFP with an *N*-terminal His tag was subcloned into the pTWIN2 vector (NEB) via NcoI and SpeI sites to yield intein-fused GFP construct with two Gly residues at the cleavage site. The protein was expressed in BL21(DE3), grown in Luria Bertani (LB) medium supplemented with 100 μg/mL of ampicillin at 37 °C in shaker incubator to an $OD_{600}$ of 1, then induced using 1 mM isopropyl thiogalactosidase (IPTG) and temperature was lowered to 28 °C to continue incubation for 5 h. Upon harvest (6000 g, 10 min, 4 °C), cells were resuspended in lysis buffer (20 mM Tris-HCl, pH 7.0, 0.5 M NaCl, 1 mM EDTA, and 1 mM PMSF) and lysed by sonication. The clear lysate was

collected by centrifugation (10000 g, 30 min, 4 °C) and loaded onto a chitin column pre-equilibrated with lysis buffer. To purify the fusion protein, the clear cell lysate was incubated on the column for 30 min at 4 °C with gentle agitation to ensure maximum protein binding. Unbound impurities were then washed away with lysis buffer. To induce intein cleavage, cleavage buffer (20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 1 mM EDTA, 50 mM MESNa) was quickly passed through the column, and the flow was stopped. The column was incubated at 4 °C overnight. The cleaved GFP was then collected in the cleavage buffer and directly used for conjugation (Fig S2).

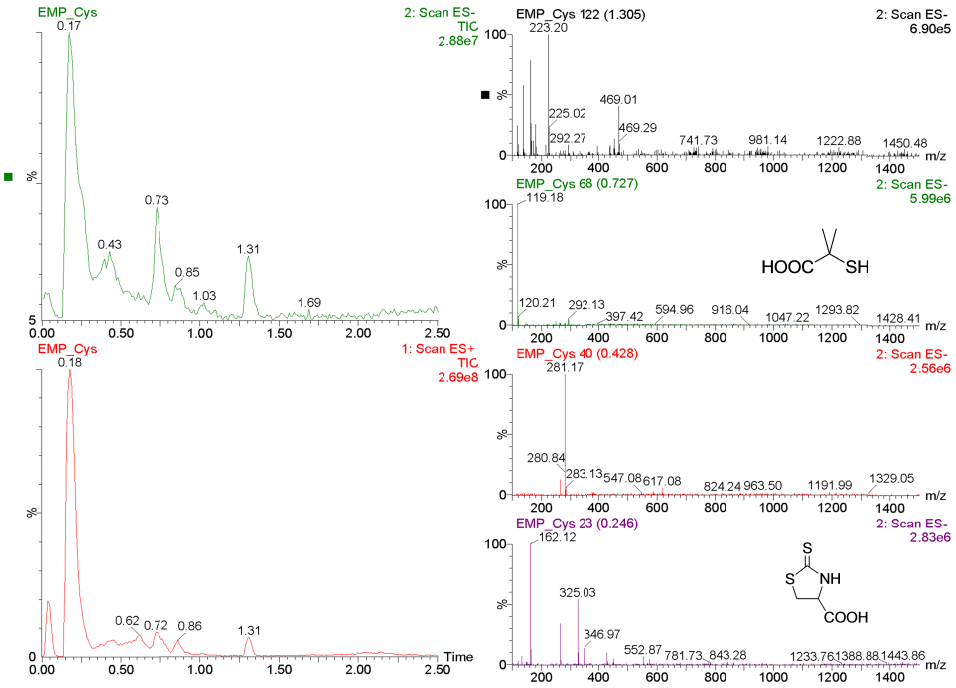
The cleaved GFP has an amino acid sequence:

MRGSHHHHHHGSASELMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGD
ATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPE
GYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNY
NSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
YLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKKLGG

**Synthesis and purification of bioconjugates.** To the cleaved GFP (1-2 mg/mL) in 20 mM Tris-HCl (pH 8.5), 0.5 M NaCl, 1 mM EDTA, and 50 mM MESNa was added. 10 eq. polymer with a cysteine end group was dissolved in the same buffer. The reaction was allowed to proceed at room temperature on a rocking shaker for 12 h. The linear PDMA-*b*-PNIPAM conjugate was purified by precipitation in 1.0 M ammonium sulfate solution followed by centrifugation at room temperature. The pellet containing conjugate and excess polymer was then resuspended in 20 mM Tris buffer (pH 8.0) and precipitated a second time. The excess polymer was removed from the second pellet using Ni-NTA affinity chromatography to yield purified conjugate. For the synthesis of Y-shaped

conjugate, cysteine-functionalized PDMA was first conjugated to GFP via native chemical ligation, and both unreacted GFP and GFP-PDMA conjugate were isolated from the excess PDMA via Ni-NTA affinity chromatography. This mixture was dialyzed into 20 mM Tris buffer (pH 8.0), and was added with 10 eq. of maleimide-functionalized PNIPAM and 10 eq. TCEP. The reaction was allowed to proceed at 4 °C for 24 h. Ammonium sulfate precipitation was less effective to purify the Y-shaped conjugate; therefore, the conjugate was purified using anion exchange chromatography with a NaCl gradient from 0 to 250 mM.

(a)



(b)

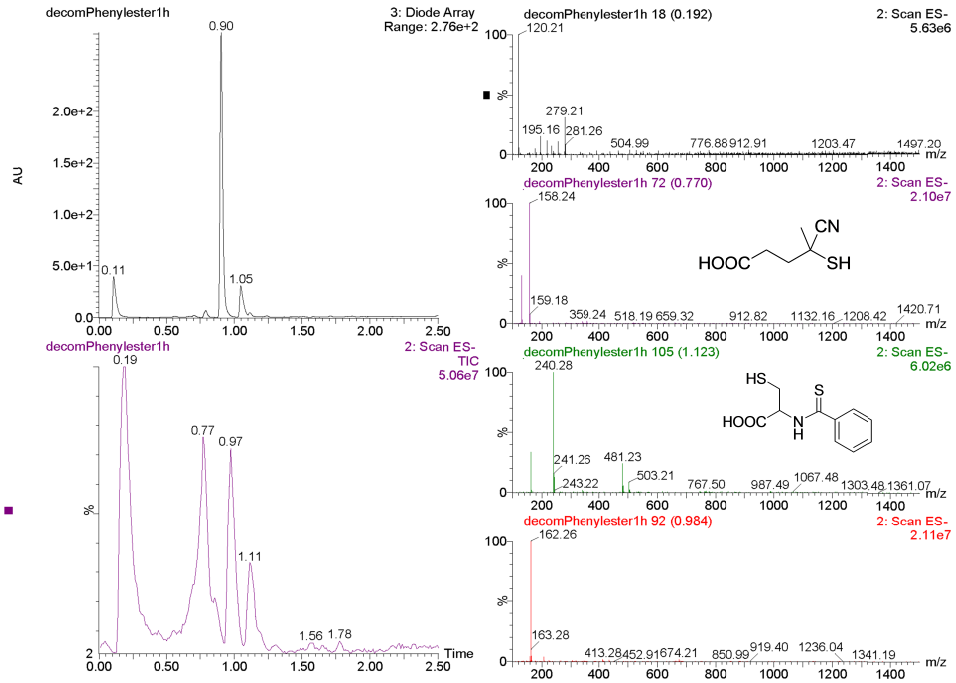



**Figure S1.** LC-MS traces of reaction mixture of RAFT agent (a) **1** and (b) **2** with cysteine.

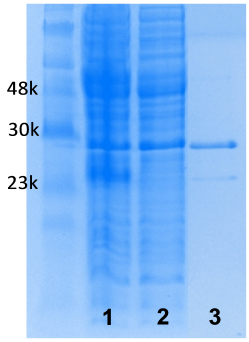



**Figure S2.** SDS-PAGE of (1) cell lysate of GFP-intein expression, (2) cell lysate after passing chitin column, and (3) eluted GFP after intein cleavage.

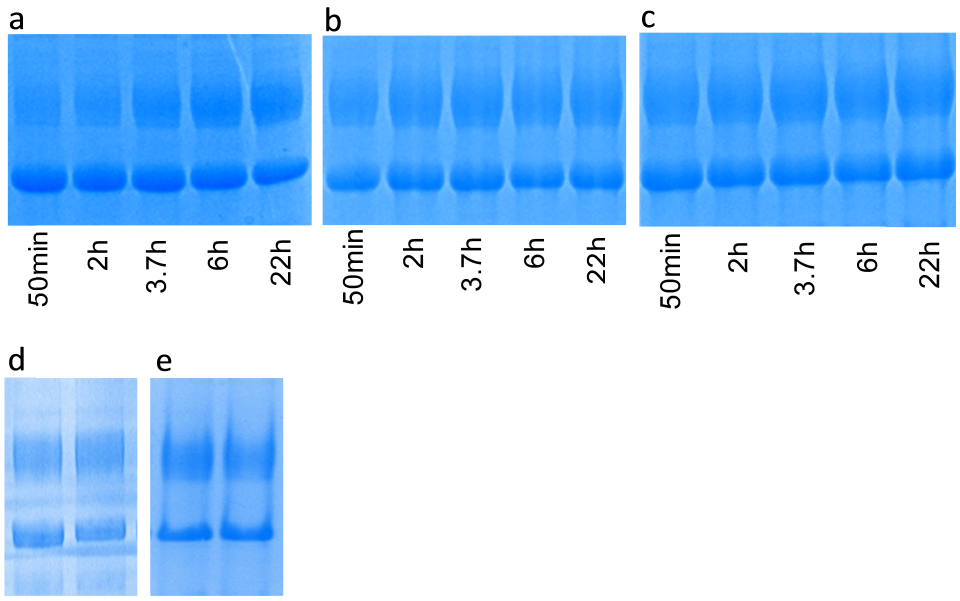



**Figure S3.** Investigation of conjugation conditions using N-Cys PDMA and GFP. Conjugation was performed at (a) 4 °C, (b) room temperature, and (c) 37 °C at different time intervals; using (d) 10 eq.(left) and 20 eq.(right) PDMA or (e) no TCEP (left) and 5 eq. TCEP (right) at room temperature.

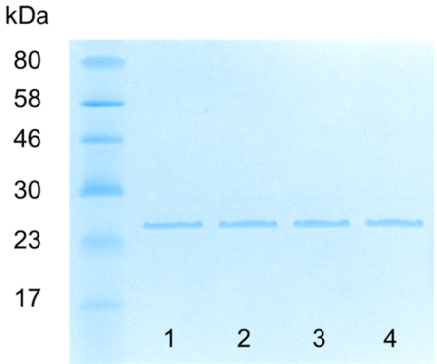
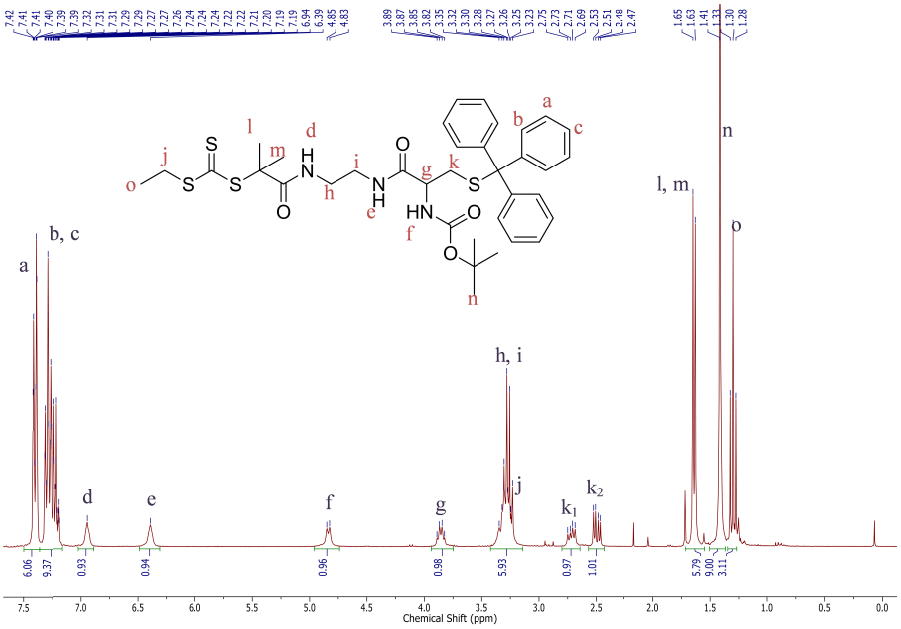
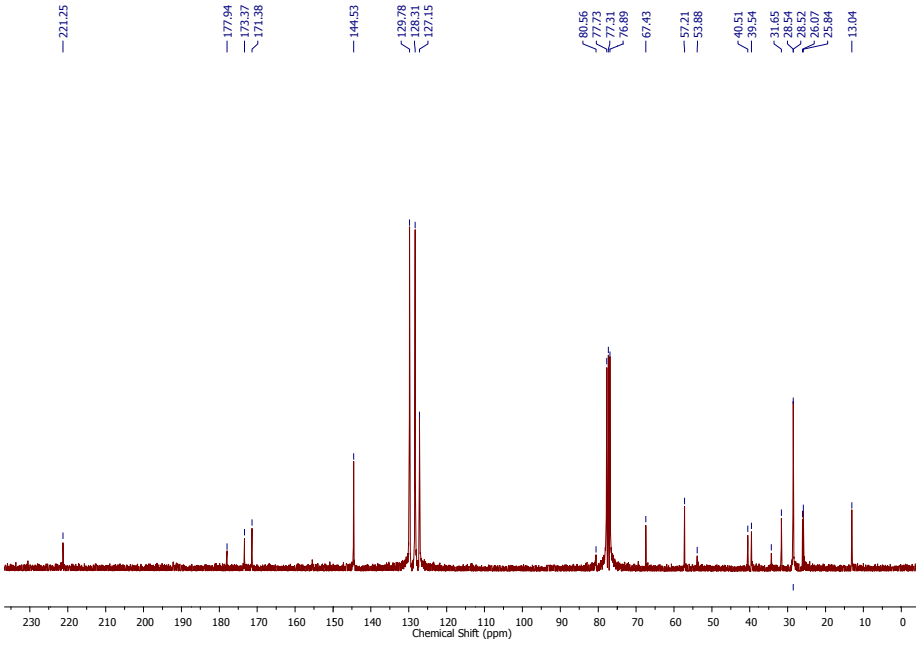



**Figure S4.** SDS-PAGE of (1) GFP, (2) equimolar physical mixture of GFP and PDMA, (3) equimolar physical mixture of GFP and PNIPAM, and (4) equimolar physical mixture of GFP, PDMA and PNIPAM. No extra band is observed in SDS-PAGE analysis.

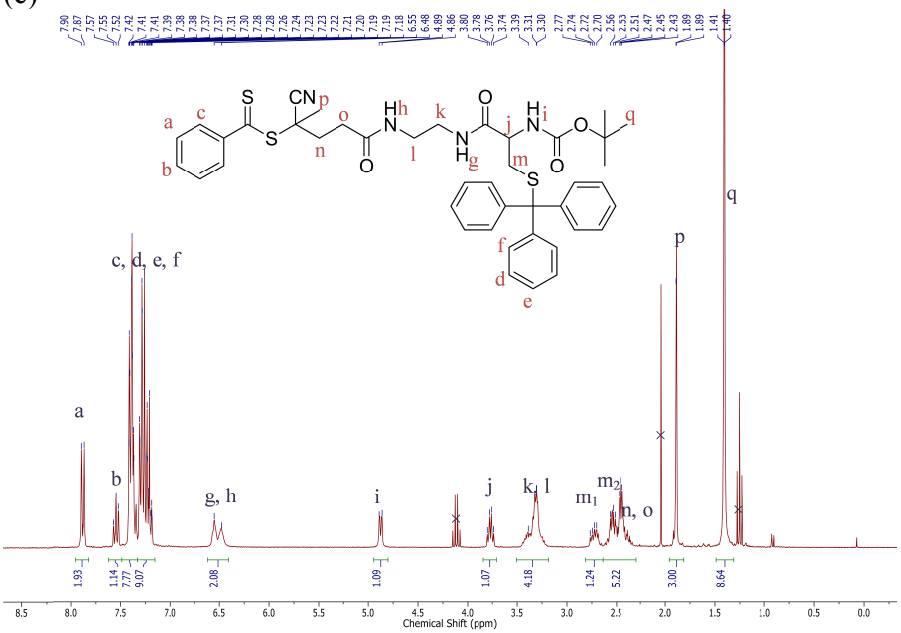
(a)


a
b, c
d
e
f
g
h, i
j
k1
k2
l, m
n
o
Chemical Shift (ppm)

(b)


221.25
177.94
173.37
171.38
144.53
129.78
128.31
127.15
80.56
77.73
77.31
76.89
67.43
57.21
53.88
40.51
39.54
31.65
28.54
28.52
26.07
25.84
13.04
Chemical Shift (ppm)

(c)




(d)

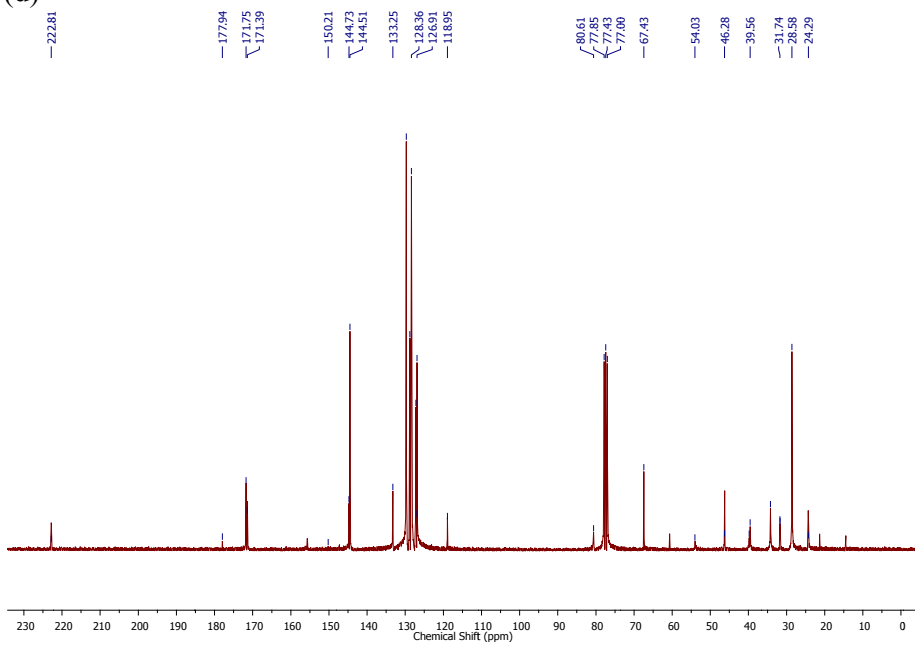



**Figure S5.** (a) $^{1}$H-NMR spectrum of RAFT agent **3**, (b) $^{13}$C-NMR spectrum of RAFT agent **3**, (c) $^{1}$H-NMR spectrum of RAFT agent **4**, (d) $^{13}$C-NMR spectrum of RAFT agent **4**.

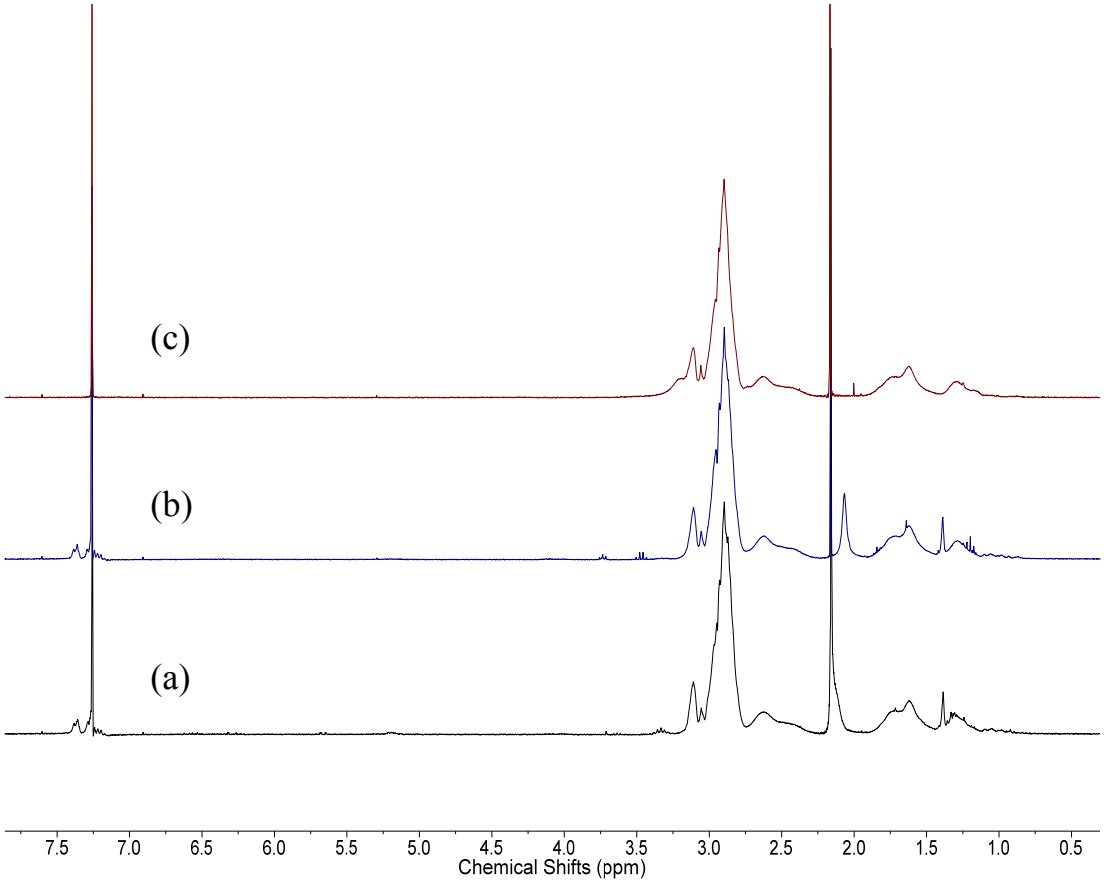



**Figure S6.** Representative $^{1}$H NMR spectra of PDMA (a) after RAFT polymerization, (b) after removal of trithiocarbonate, and (c) after deprotection of cysteine.

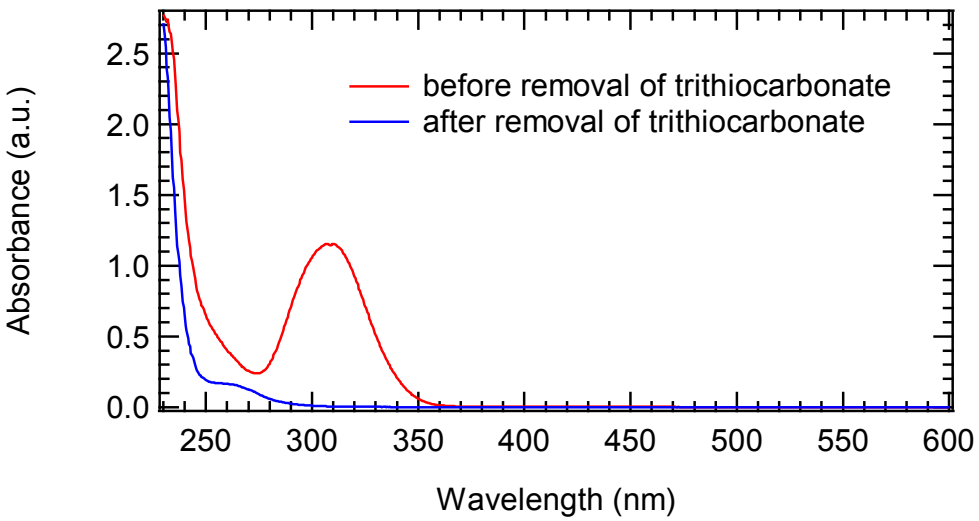



**Figure S7.** UV-Vis spectra of 0.5 mg/mL PDMA ($M_n$ = 6.7 k, PDI = 1.02) in mill-Q water at room temperature. Trithiocarbonate is removed by aminolysis. The complete removal of trithiocarbonate is evidenced by the disappearance of absorbance at 310 nm.

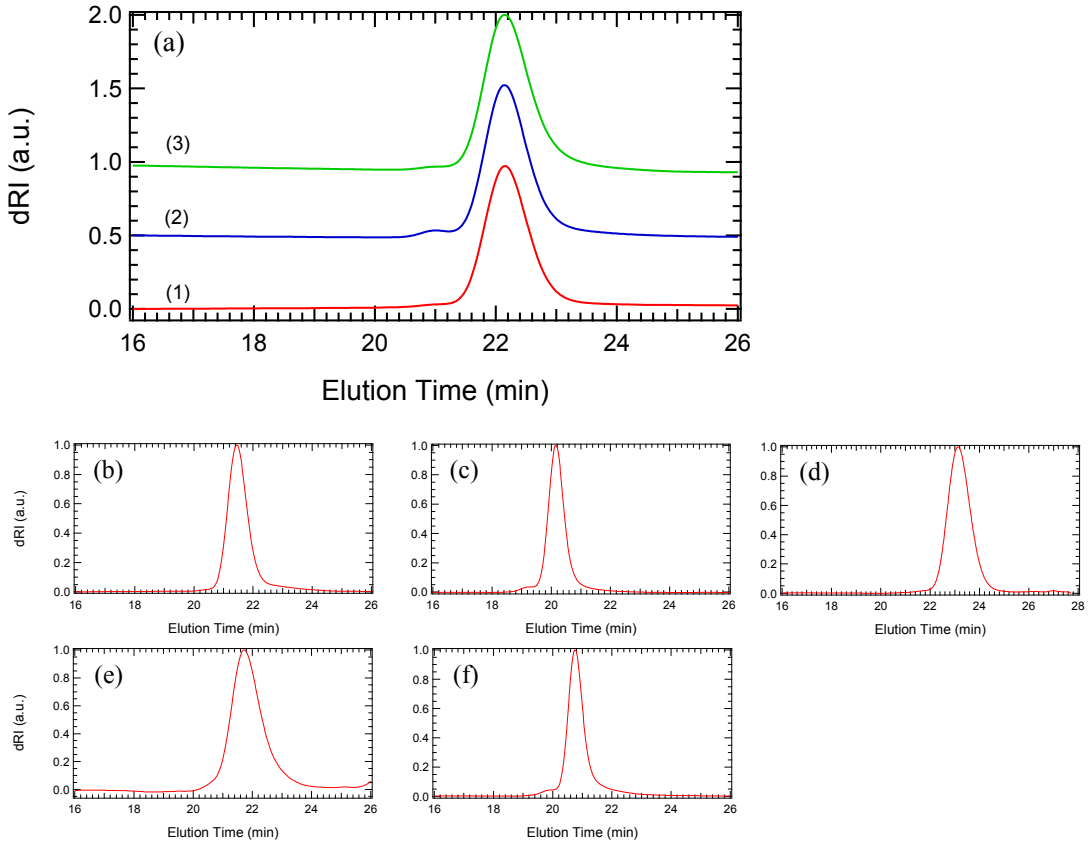



**Figure S8.** GPC traces of RAFT polymers. (a) PDMA, $M_n$ = 6.7 k, PDI = 1.02; (1) after RAFT polymerization, (2) after removal of trithiocarbonate by aminolysis, and (3) after deprotection of cysteine. A slight shoulder is observed at high molecular weight in trace 2 due to oxidative thiol-thiol coupling; this coupling peak disappears upon further deprotection of the cysteine in trace 3. (b) PDMA, $M_n$ = 10.9 k, PDI = 1.04; (c) PDMA, $M_n$ = 21.0 k, PDI = 1.06; (d) PDMA, $M_n$ = 6.4 k, PDI = 1.04, trithiocarbonate removed by radical-induced reduction; (e) PPEGMA, $M_n$ = 9.2 k, PDI = 1.04; (f) PDMA(6.9 k)-b-PNIPAM(7.9 k), PDI = 1.08.

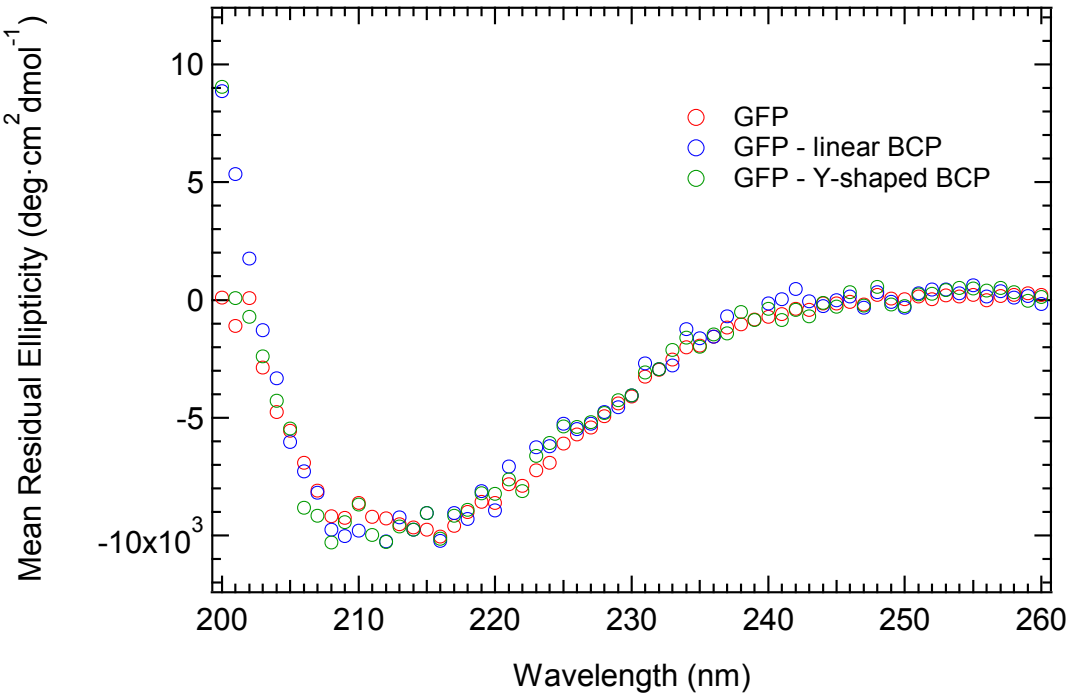



**Figure S9.** CD spectra of unconjugated GFP, linear BCP conjugate, and Y-shaped conjugate (4.5 μM in 20 mM Tris buffer, pH = 8.0) at 25 ºC.

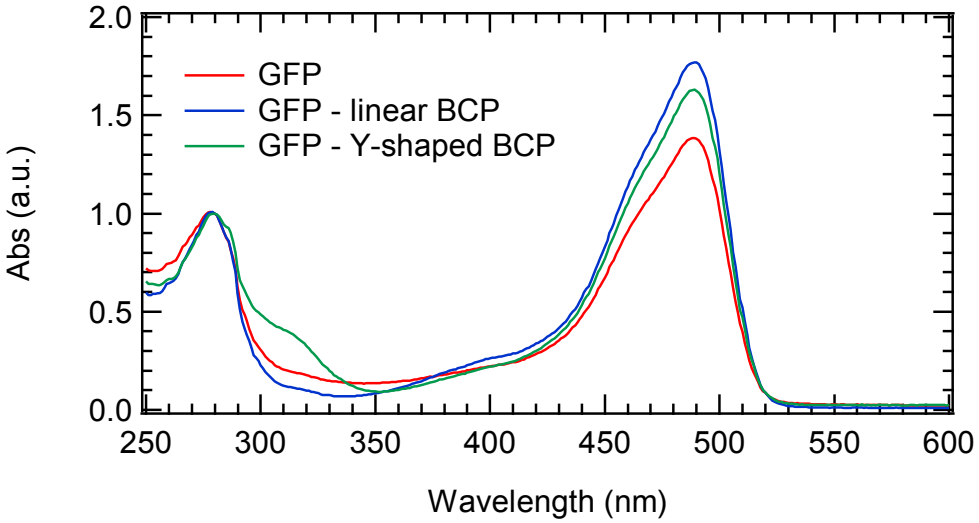



**Figure S10.** UV-Vis spectra of unconjugated GFP, linear BCP conjugate, and Y-shaped conjugate (10 μM in 20 mM Tris buffer, pH = 8.0) at room temperature. The absorption at 320 nm for GFP Y-shaped conjugate is due to the trithiocarbonate end group present on PNIPAM polymer. Spectra are normalized to the signal at 280 nm.

## References

(1) Thomas, C. S.; Glassman, M. J.; Olsen, B. D. *ACS Nano* **2011**, *5*, 5697-5707.
(2) Wood, R. J.; Pascoe, D. D.; Brown, Z. K.; Medlicott, E. M.; Kriek, M.; Neylon, C.; Roach, P. L. *Bioconjugate Chem.* **2004**, *15*, 366-372.
(3) Chong, Y. K.; Moad, G.; Rizzardo, E.; Thang, S. H. *Macromolecules* **2007**, *40*, 4446-4455.