Supporting Information

Enzyme-Catalyzed Macrocyclization of Long Unprotected Peptides

Chi Zhang[1], *Peng Dai*[1], *Alexander M. Spokoyny*[1], *Bradley L. Pentelute*[1*]

[1] Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, United States

*Corresponding author
Email: blp@mit.edu
Phone: +1-617-324-0180

Table of Contents

1. General Considerations **3**
Chemicals 3
Glutathione S-transferase (GST) 3
Solid phase peptide synthesis 4
Peptide purification 5
LC-MS analysis 5
Determination of reaction yields 6
2. General Protocol for Preparation of S-Perfluoroarylated Peptides **6**
3. General Protocol for GST-catalyzed Peptide Macrocyclization **7**
4. Synthesis and Purification of Peptide 7 ..... **7**
5. Expression and Purification of GST-M5 ..... **7**
6. References **9**
7. Supplementary Tables and Figures **9**
Table S1 9
Figure S1 10
Figure S2 11
Figure S3 12
Figure S4 13
Figure S5 14
Figure S6 15
Figure S7 16
Figure S8 17
Figure S9 18
Figure S10 19
Figure S11 20
Figure S12 21
Figure S13 22
Figure S14 23
Figure S15 24
Figure S16 25
Figure S17 26
Figure S18 27

1. General Considerations

Chemicals

Hexafluorobenzene and decafluorobiphenyl were purchased from Oakwood Chemicals (West Columbia, SC). Decafluorobiphenyl sulfide was purchased from SynQuest Laboratories (Alachua, FL). *Tris*(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl) was purchased from Hampton Research (Aliso Viejo, CA). 2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), Fmoc-Rink amide linker, Boc-α-Glu(OtBu)-OH, Fmoc-Gly-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Ala-OH, Fmoc-Cys(Trt)-OH, Fmoc-Cys(S-tBu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Phe-OH, Fmoc-Ile-OH, Fmoc-Ser(tBu)-OH, Fmoc-Tyr(tBu)-OH and Fmoc-Pro-OH were purchased from Chem-Impex International (Wood Dale, IL). 4-methylbenzhydrylamine (MBHA) resin was obtained from Anaspec (Fremont, CA). *N,N*-Dimethylformamide (DMF), dichloromethane (DCM), diethyl ether, HPLC-grade acetonitrile, and guanidine hydrochloride were obtained from VWR International (Philadelphia, PA). All other reagents were purchased from Sigma-Aldrich and used as received unless otherwise noted.

Glutathione *S*-transferase (GST)

GST from equine liver (Sigma-Aldrich catalog number: G6511) was received as lyophilized powder containing Tris and phosphate buffer salts, reduced glutathione, and EDTA. Water was added to dissolve the GST sample to a final concentration of 10 mg/mL followed by dialysis against 20 mM Tris, 150 mM NaCl, pH 7.5 buffer to remove reduced glutathione, and aliquots of GSTs were stored at -80^{o}C.

Human GST-M5 was cloned from pcDNA3-HA-GSTM5 (Addgene number: 11986) using Champion^{TM} pET SUMO expression system (Life Technology) following manufacturer's recommended protocol. Aliquots of expressed GST-M5 were stored in 20 mM Tris, 150 mM NaCl, pH 7.5 buffer at 6 mg/mL concentration in -80 ^{o}C refrigerator.

Solid phase peptide synthesis

All peptides were synthesized on a 0.2 mmol scale using manual Fmoc-SPPS chemistry under flow[1] using 3 minutes cycle for each amino acid. Specifically, all reagents and solvents are delivered to a stainless steel reactor containing resins at a constant flow rate using HPLC pump; temperature of the reactor was maintained at 60 °C during the synthesis using water bath. Procedure for amino-acid residue coupling cycle contained 30 second coupling with 1 mmol Fmoc protected amino acids, 1.2 mmol HBTU, and 500 μL of diisopropyl ethyl amine (DIEA) in 2.5 mL of DMF, flow rate was 6 mL/min, note that for coupling of cysteine, 190 μL of DIEA was used to prevent racemization; 1 minute wash with DMF, flow rate was 20 mL/min; 20 seconds deprotection with 50% (v/v) piperidine in DMF, flow rate was 20 mL/min; and 1 minute wash with DMF, flow rate was 20 mL/min. The resin was washed thoroughly with DCM and air dried after completion of the stepwise SPPS. The peptide is then simultaneously cleaved from the resin and side-chain deprotected by treatment with 2.5% (v/v) water, 2.5% (v/v) 1,2-ethanedithiol (EDT), and 1% (v/v) triisoproprylsilane in neat trifluoroacetic acid (TFA) for 2 hours at room temperature. Resulting solution containing peptide was evaporated by blowing a stream of nitrogen gas over its surface for 15 minutes, then triturated and washed with cold diethyl ether three times. Obtained gummy-like solid was dissolved in 50% H_2O: 50% acetonitrile containing 0.1% TFA and lyophilized. ***These same solvent compositions were used in most experiments and will be referred to as A: 0.1% TFA in H_2O and B: 0.1% TFA in acetonitrile.*** All peptides were synthesized using MBHA resin, unless noted otherwise. Peptide **5** was synthesized using hydrazide resin[2]. All peptides containing glutathione tag at *N*-terminus were synthesized with Boc-α-Glu(OtBu)-OH as the last amino acid residue.

Peptide purification

The crude peptide was dissolved in 95% A: 5% B with 6 M guanidinium hydrochloride and purified by semi-preparative RP-HPLC (Agilent Zorbax SB C_{18} column: 21.2 x 250 mm, 7 μm, linear gradient: 5-50% B over 90 min, flow rate: 5 mL/min). 1 μL of each HPLC fraction was mixed with 1 μL of alpha-cyano-4-hydroxycinnamic acid (CHCA) matrix in 75% A: 25% B, spotted with MALDI, and checked for fractions with desired molecular mass. The purity of fractions was confirmed by analytical RP-HPLC (Agilent Zorbax SB C_3 column: 2.1 x 150 mm, 5 μm, gradient: 0-2 minutes 5% B, 2-11 minutes 5-65% B, 11-12 minutes 65% B, flow rate: 0.8 mL/min). HPLC fractions containing only product material were confirmed by LC-MS analysis, combined, and then lyophilized. Peptides synthesized using fast flow-based SPPS and purified by RP-HPLC are listed in Table S1.

LC-MS analysis

LC-MS chromatograms and associated mass spectra were acquired using Agilent 6520 ESI-Q-TOF mass spectrometer. Following LC methods were used:

Method A*:* Zorbax SB C_3 column: 2.1 x 150 mm, 5 μm, gradient: 0-2 minutes 5% B, 2-11 minutes 5-65% B, 11-12 minutes 65% B, flow rate: 0.8 mL/min.

Method B*:* Zorbax SB C_{18} column: 2.1 x 250 mm, 5 μm, gradient: 0-2 minutes 1% B, 2-15 minutes 1-70% B, 15-17 minutes 70% B, flow rate: 0.4 mL/min.

All samples were analyzed using ***Method A*** unless otherwise noted. Data was processed using Agilent Mass Hunter software package.

Y-axis in all chromatograms shown in supplementary figures represents total ion current (TIC); mass spectrum insets correspond to the maxima point of the TIC peak.

Determination of reaction yields

All yields reported were determined by measuring total ion currents using LC-MS data. First, using Agilent Mass Hunter software package, the peak areas for all relevant peptidic species on the chromatogram were integrated. Then yield was calculated as following: %yield = S_{pro}/S_{all} where S_{pro} is the peak area of the desired product and S_{all} is the sum of all TIC peaks.

2. General Protocol for Preparation of *S*-Perfluoroarylated Peptides

Peptides **1a−1c:** To a solid sample of peptide **1** (5 μmoles) dissolved in 20 mM Tris base in 1mL of DMF in a plastic Eppendorf tube was added 500 μmoles of perfluoroaromatic reagent (decafluorobiphenyl for product **1a**, hexafluorobenzene for **1b**, decafluorobiphenyl sulfide for **1c**). The tube was vortexed to ensure complete reagent mixing and left at room temperature for 30 minutes. 1 μL of each reaction mixture was quenched by addition of 20 μL of 50% A: 50% B and subjected to LC-MS analysis. Resulting reaction mixtures were quenched by addition of 20 mL of 95% A: 5% B, filtered through 0.22 μm nylon syringe filter, and purified by RP-HPLC.

Peptides **2a−4a** were prepared using the same protocol as **1a−1c** by reacting peptide **2−4** with decafluorobiphenyl. Note that for preparation of peptides **3a** and **4a**, 20 mL of 90% A: 10% B was used as a quenching solution. Peptide ***6a*** was prepared using the same protocol by reacting peptide **6** with decafluorobiphenyl.

LC-MS data for all reactions obtained *in situ* and RP-HPLC-purified products of peptides **1a** to **1c** and **2a−4a** are summarized in Figure S1 and Figure S2.

3. General Protocol for GST-catalyzed Peptide Macrocyclization

Macrocyclization reactions described in Figure 3 and Figure 4 were performed on a 100 μL scale. 10 μL of the peptide stock solution (1 mM in water), 2 μL of GST stock solution (10 mg/mL), 10 μL of 10X reaction buffer (1M phosphate, 200 mM TCEP•HCl, pH 8.0), and 78 μL of water were combined in a vial (total volume of 100 μL). The reaction mixture was pipetted up and down for 20 times to ensure thorough mixing and left at room temperature (25 ^{o}C). 10 μL of the reaction mixture was quenched by addition of 100 μL of 50% A: 50% B, and subjected to LC-MS analysis.

4. Synthesis and Purification of Peptide 7

Peptide **7** was synthesized *via* hydrazide-based native chemical ligation protocol.[2] To a solution of peptide **5** (2.17 mg, 0.92 μmol) dissolved in 210 μL of 200 mM phosphate buffer, pH 3.0 at -10 ^{o}C was slowly added 20 μL of 100 mM $NaNO_2$ in water and reacted for 20 min at the same temperature. 230 μL of 200 mM sodium 2-sulfanylenthanesulfonate in 200 mM phosphate buffer, pH 7.0 was then added to the reaction mixture. The pH of the solution was adjusted to 7.0 by addition of 5M sodium hydroxide solution. The reaction mixture was left at room temperature for 10 minutes followed by the addition of 460 μL of 2 mM peptide **6a** (*S*-perfluoroarylated peptide **6**) solution and 100 μL of buffer (2M phosphate, 200 mM TCEP•HCl, pH 7.0). The reaction was left at room temperature and monitored by LC-MS analysis. After ligation for two hours, the product peptide **7** was purified by RP-HPLC (Figure S15).

5. Expression and Purification of GST-M5

GST-M5 gene was amplified from pcDNA3-HA-GSTM5 (Addgene number: 11986) using primers: 5'-ATGCCCATGACTCTGGG and 5'-CTATTTGCTGTTCCATGTAGCTGAC, after gel

purification, the gene was ligated into pET SUMO vector using protocol provided in Champion™ pET SUMO protein expression system manual. The *N*-terminal oligo-glycine sequence was introduced by site-directed mutagenesis using QuickChange Lightning Multi Site-Directed mutagenesis kit (Agilent) with primer: 5'-GGCTCACAGAGAACAGATTGGTGGTggtggtggtggcggcATGCCCATGACTCT-GGGG. The generated pET-SUMO-G5-GST-M5 construct encoded for the following protein sequence (G5-GST-M5 sequence is shown in bold):

SUMO-GST-M5
MSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLR
FLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGG**GGGGGMPMTLGYWDIRGLAHAIRLLLEY**
TDSSYVEKKYTLGDAPDYDRSQWLNEKFKLGLDFPNLPYLIDGAHKITQSNAILRYIARKH
NLCGETEEEKIRVDILENQVMDNHMELVRLCYDPDFEKLKPKYLEELPEKLKLYSEFLGK
RPWFAGDKITFVDFLAYDVLDMKRIFEPKCLDAFLNLKDFISRFEGLKKISAYMKSSQFLR
GLLFGKSATWNSK

E. coli BL21(DE3) cells transformed with pET-SUMO-G5-GST-M5 plasmid were grown in 1 L of LB medium containing kanamycin (30 μg/mL) at 37 ℃ until OD_{600} = 0.6. Then, expression was induced by addition of 0.5 mM IPTG overnight at 30 ℃. After harvesting the cells by centrifugation (6,000 rpm, 10 min), the cell pellet was lysed by sonication in 25 mL of 50 mM Tris, 150 mM NaCl, pH 7.5 buffer containing 15 mg lysozyme (Calbiochem), 1 mg DNase I (Sigma-Aldrich), and 0.5 tablet of protease inhibitor cocktail (Roche Diagnostics, Germany). The suspension was centrifuged at 17,000 rpm for 30 min to remove cell debris. The supernatant was loaded onto a 5 mL HisTrap FF crude Ni-NTA column (GE Healthcare, UK), first washed with 40 mL of 20 mM Tris, 150 mM NaCl, pH 8.5, and then washed with 40 mL of 40 mM imidazole in 20 mM Tris, 150 mM NaCl, pH 8.5. The protein was eluted from the column with buffer containing 500 mM imidazole in 20 mM Tris, 150 mM NaCl, pH 8.5. Imidazole was removed from protein using HiPrep 26/10 Desalting column (GE Healthcare, UK), protein was eluted into 20 mM Tris, 150 mM NaCl, pH 7.5 buffer. Purified protein was analyzed using an Any kD Mini-

PROTEAN TGX Precast Gel (Bio-Rad, CA). In addition, the protein was analyzed by LC-MS to confirm its purity and molecular weight.

SUMO group on SUMO-G5-GST-M5 was cleaved by incubating 1 μg of SUMO protease per mg protein at room temperature for 60 minutes. The crude reaction mixture was loaded onto a 5 mL HisTrap FF crude Ni-NTA column (GE Healthcare, UK) and the flow through containing G5-GST-M5 was collected. The protein was analyzed by LC-MS confirming sample purity and molecular weight (analyzed via high-resolution ESI-Q-TOF MS, observed mass: 25960.2, calculated mass: 25960.5).

6. References

1. Simon, M. D.; Heider, P. L.; Adamo, A.; Vinagradov, A. A.; Mong, S. K.; Li, X.; Berger, T.; Policarpo, R. L.; Zhang, C.; Zou, Y.; Liao, X.; Spokoyny, A. M.; Jensen, K. F.; Pentelute, B. L. *ChemBioChem* **2014**, *15*, 713-720.

2. Fang, G.-M.; Li, Y.-M.; Shen, F.; Huang, Y.-C.; Li, J.-B.; Lin, Y.; Cui, H.-K.; Liu, L. *Angew. Chem. Int. Ed.* **2011**, *50*, 7645-7649.

7. Supplementary Table and Figures

Table S1. Sequences and masses of peptides synthesized by fast-flow solid-phase peptide synthesizer.

Peptide	Sequence*	Calc. mass	Obs. mass
1	NH_2-γ-EC(S-tBu)G-GLKAG-C-$CONH_2$	923.40	923.40
2	NH_2-γ-EC(S-tBu)G-$(GLKAG)_2$-C-$CONH_2$	1349.66	1349.66
3	NH_2-γ-EC(S-tBu)G-$(GLKAG)_3$-C-$CONH_2$	1775.92	1775.92
4	NH_2-γ-EC(S-tBu)G-$(GLKAG)_4$-C-$CONH_2$	2202.18	2202.18
2b	NH_2-γ-ECG-$(GLKAG)_2$-C-$CONH_2$	1261.63	1261.63
3b	NH_2-γ-ECG-$(GLKAG)_3$-C-$CONH_2$	1687.89	1687.89
4b	NH_2-γ-ECG-$(GLKAG)_4$-C-$CONH_2$	2114.15	2114.15
5	NH_2-γ-ECGGNQQKRAFIRSLYDDPSG-$CONHNH_2$	2354.12	2354.12
6	NH_2-C(S-^{t}Bu)ANLLAEAKKLNDAQAPKC-$CONH_2$	2087.07	2087.07

* Amino acids are shown in one-letter code. γ-E stands for γ-glutamic acid; C(S-^{t}Bu) denotes *tert*-butylthio-protected cysteine.

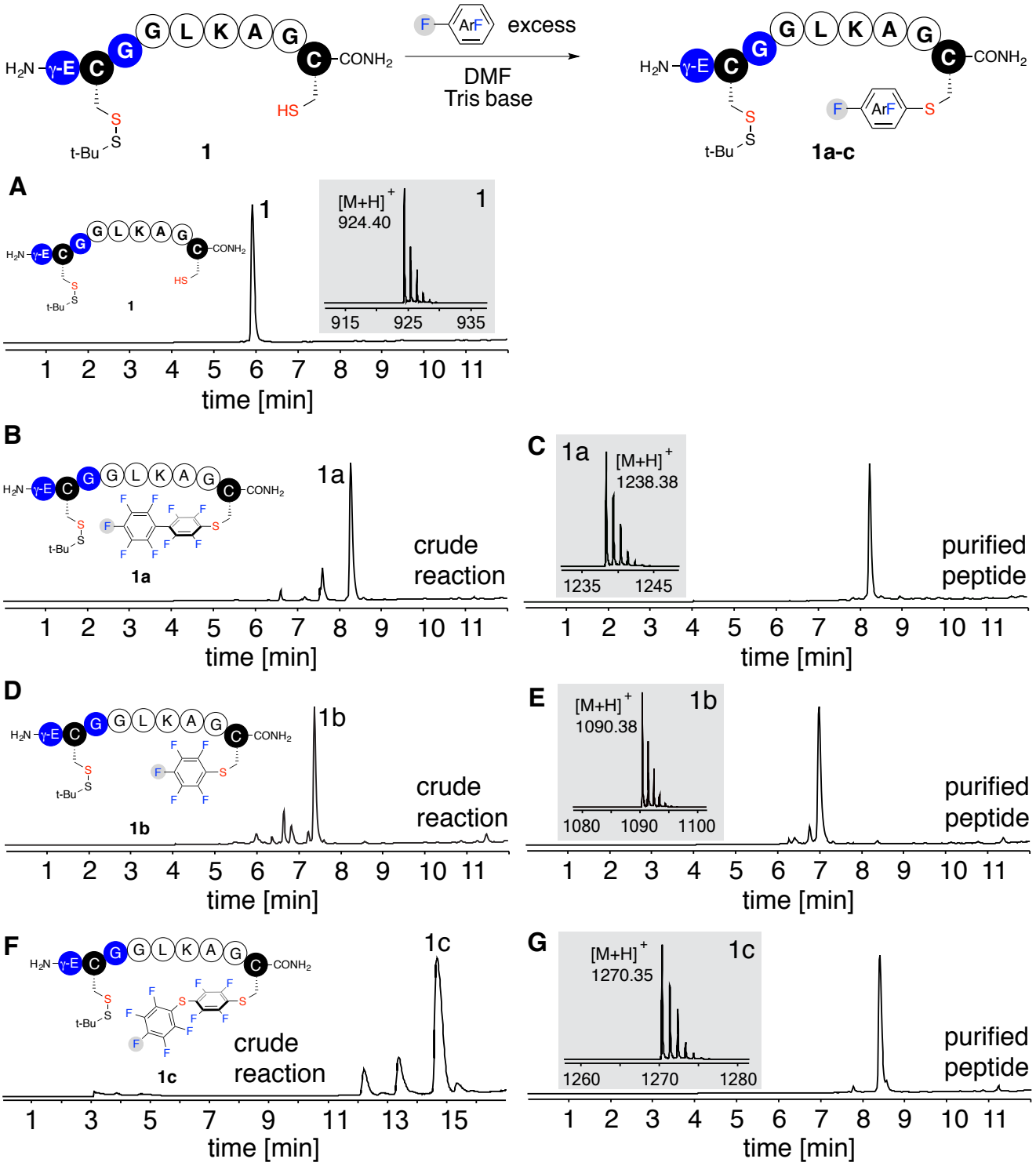


Figure S1. Synthesis and purification of peptides **1a-c**. Reaction conditions are described in the general protocol section. Chromatograms B, D, and F correspond to crude reactions used to prepare peptides **1a**, **1b**, and **1c**, respectively; chromatograms C, E and G represent LC-MS analysis data for RP-HPLC-purified peptides **1a**, **1b**, and **1c**, respectively. Chromatogram F was acquired using LC-MS method B.

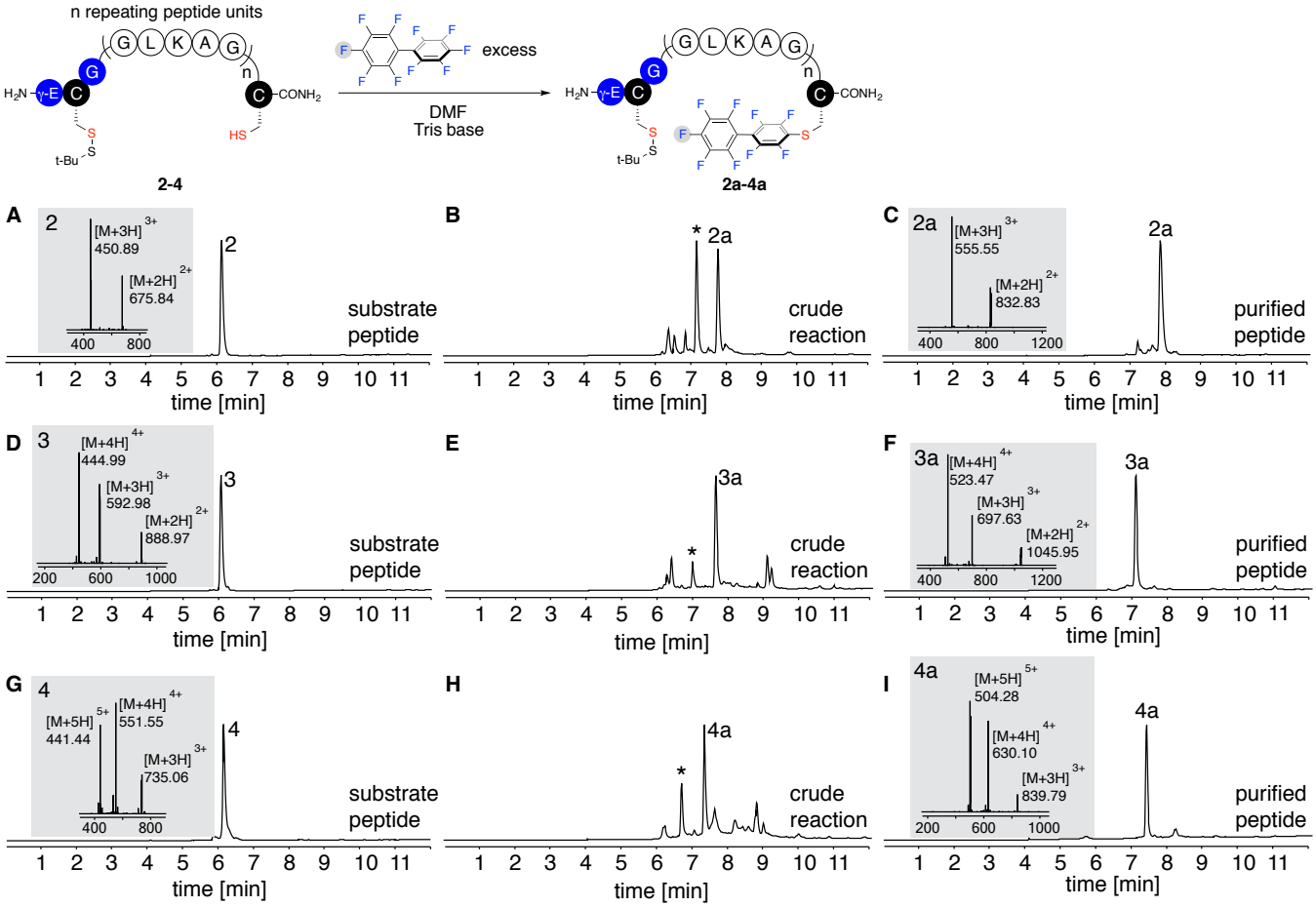


Figure S2. Synthesis and purification of peptides **2a-4a**. Shown are three series of chromatograms, each series consists of TIC chromatograms for LC-MS analysis of the pure substrate peptide, crude reaction mixture, and purified perfluoroaryl-linked peptide. A, B, and C are for peptide **2a**; D, E, and F are for peptide **3a**; G, H, and I are for peptide **4a**. Peaks labeled with "*" denote side products with two peptides crosslinked by one decafluorobiphenyl molecule.

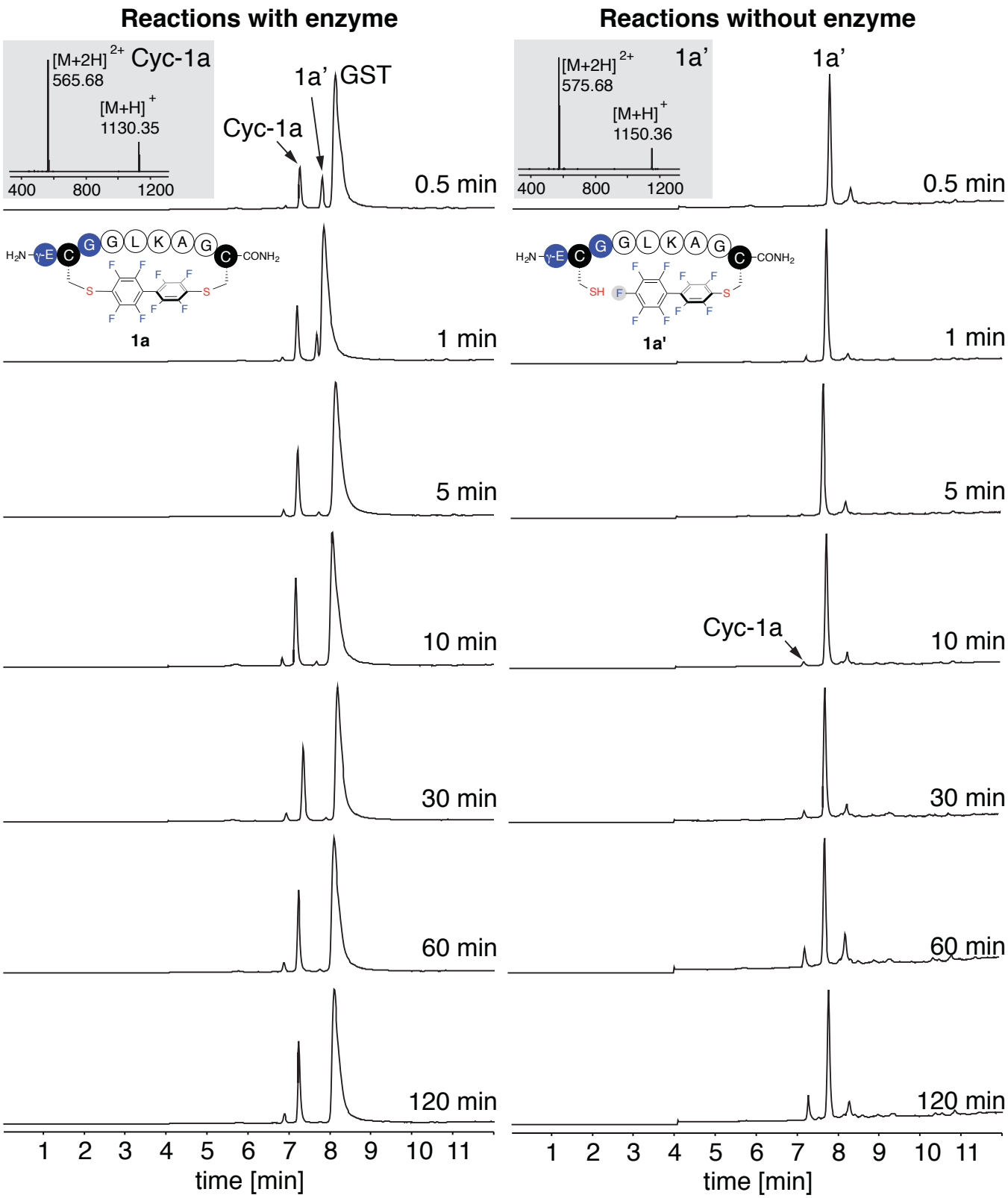


Figure S3. Macrocyclization of peptide **1a** with and without GST analyzed by LC-MS at different time points. Reaction conditions are described in Figure 3 of the manuscript.

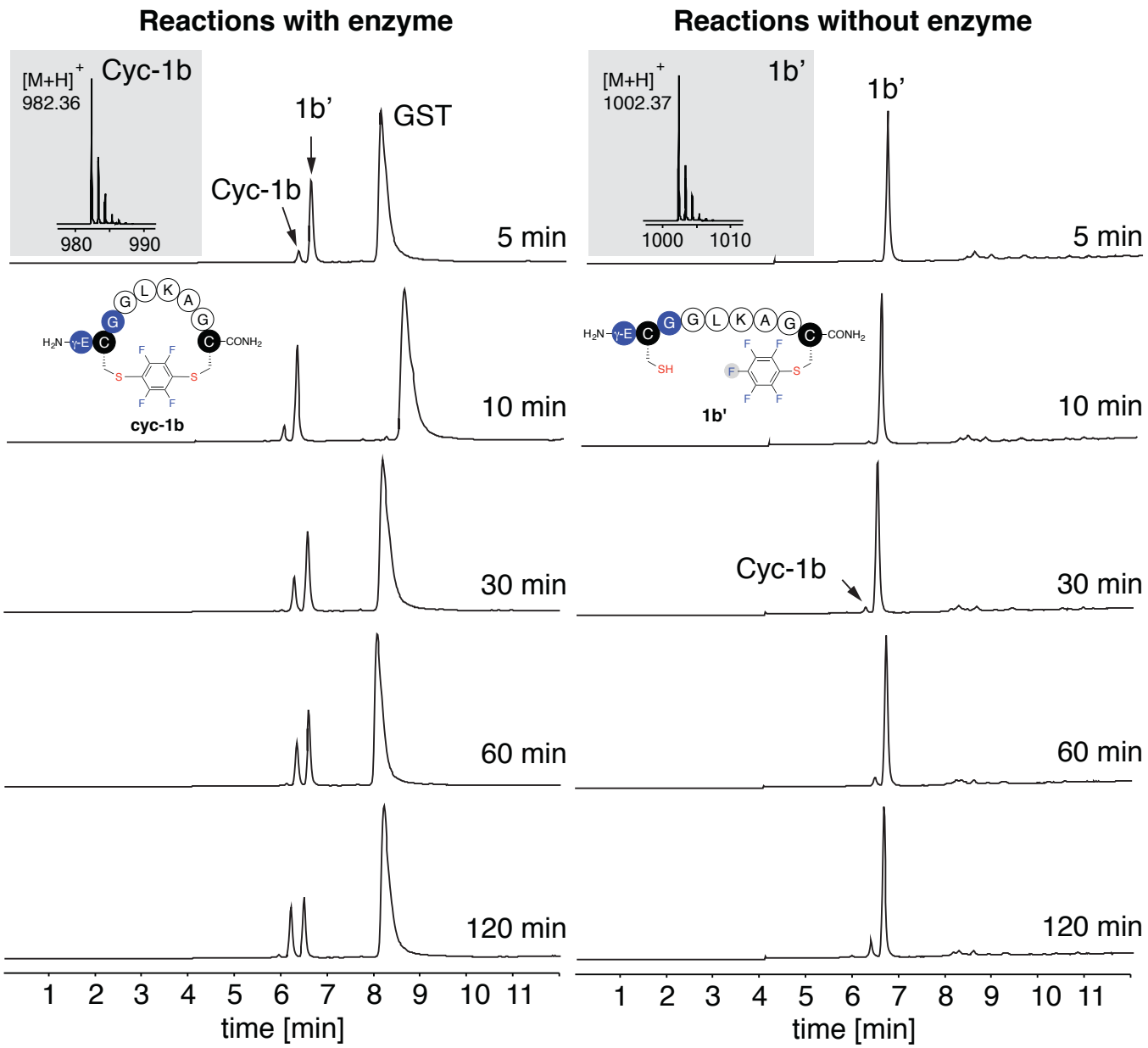


Figure S4. Macrocyclization of peptide **1b** with and without GST analyzed by LC-MS at different time points. Reaction conditions are described in Figure 3 of the manuscript.

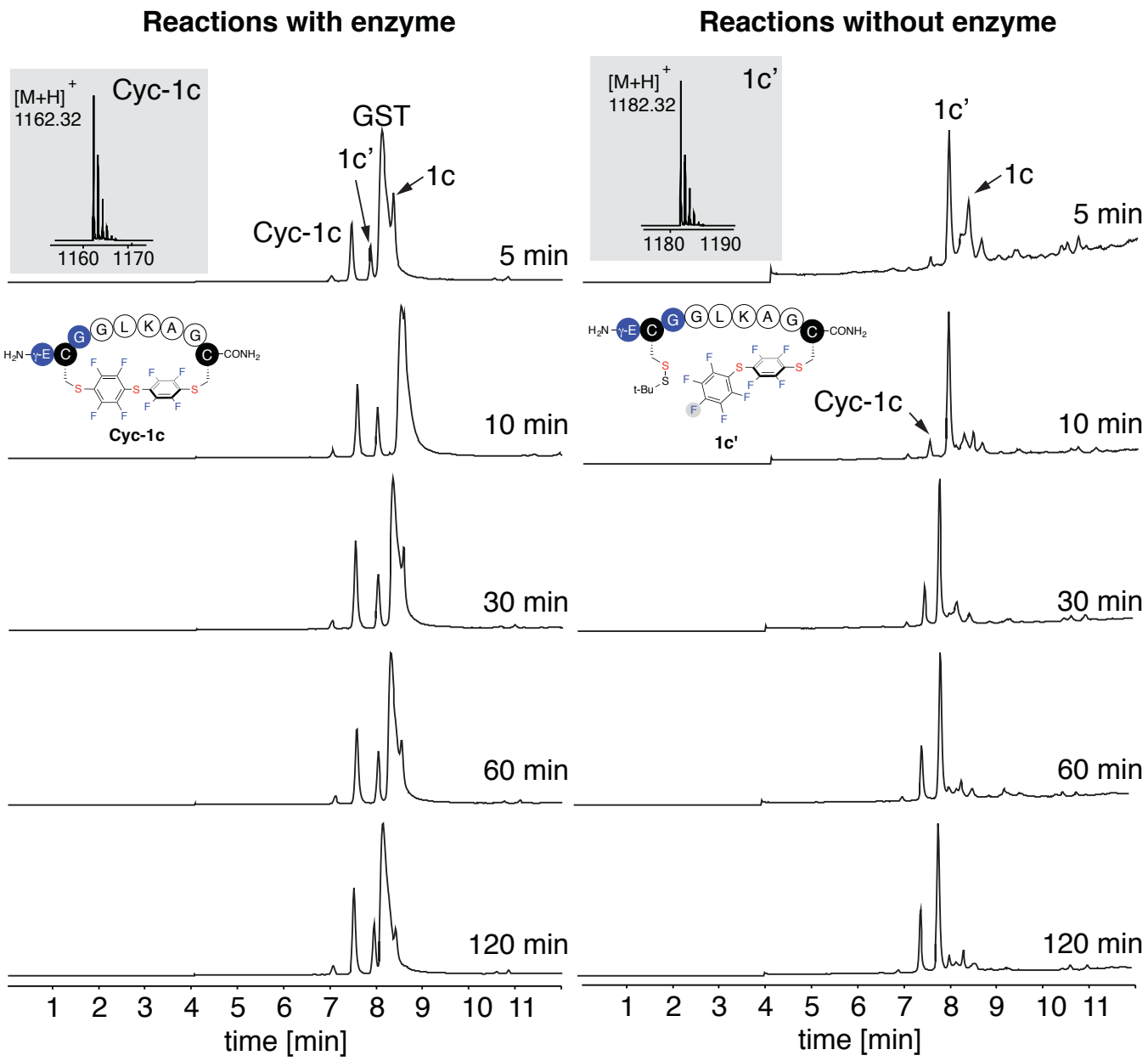


Figure S5. Macrocyclization of peptide **1c** with and without GST analyzed by LC-MS at different time points. Reaction conditions are described in Figure 3 of the manuscript.

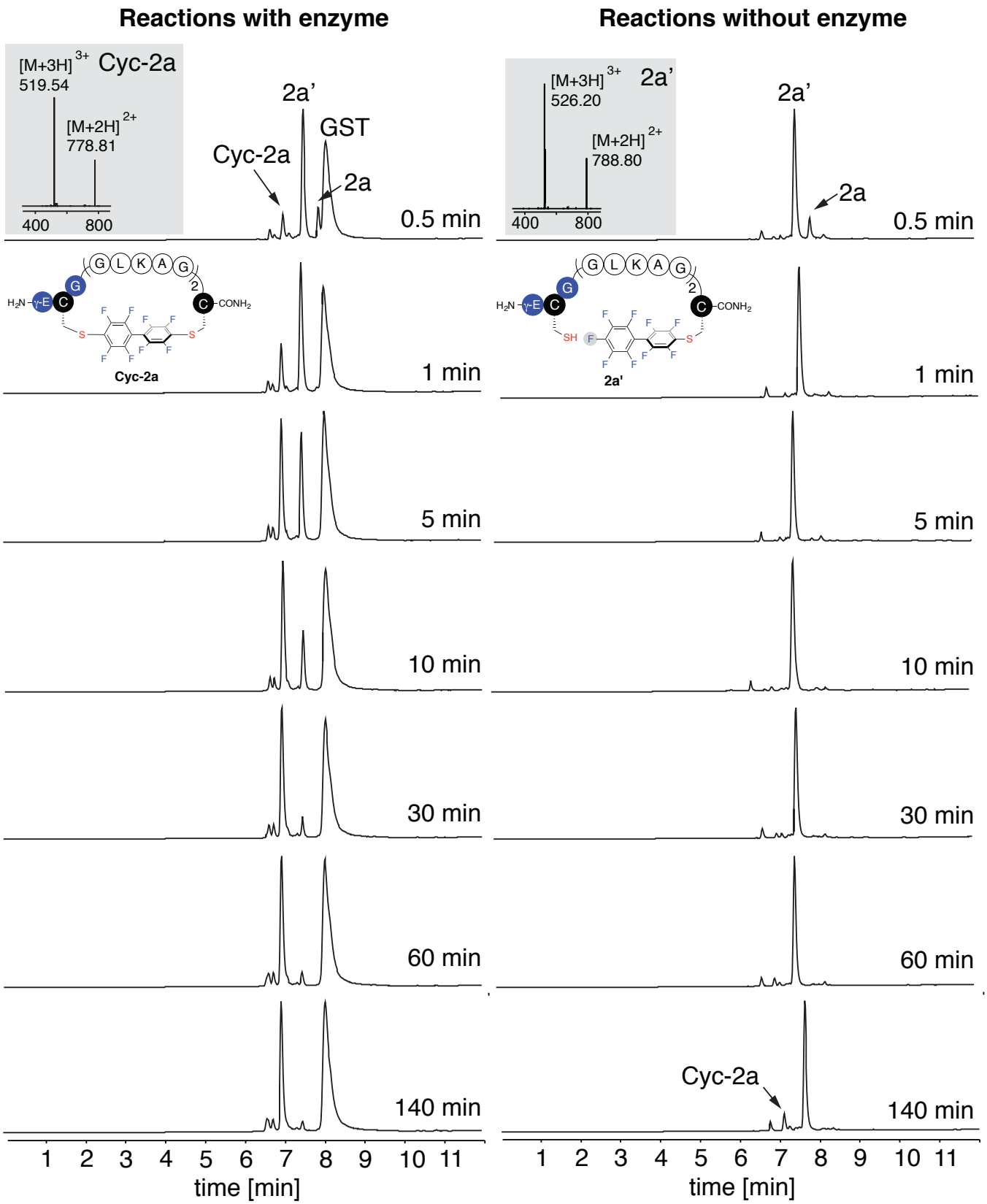


Figure S6. Macrocyclization of peptide **2a** with and without GST analyzed by LC-MS at different time points. Reaction conditions are described in Figure 4 of the manuscript.

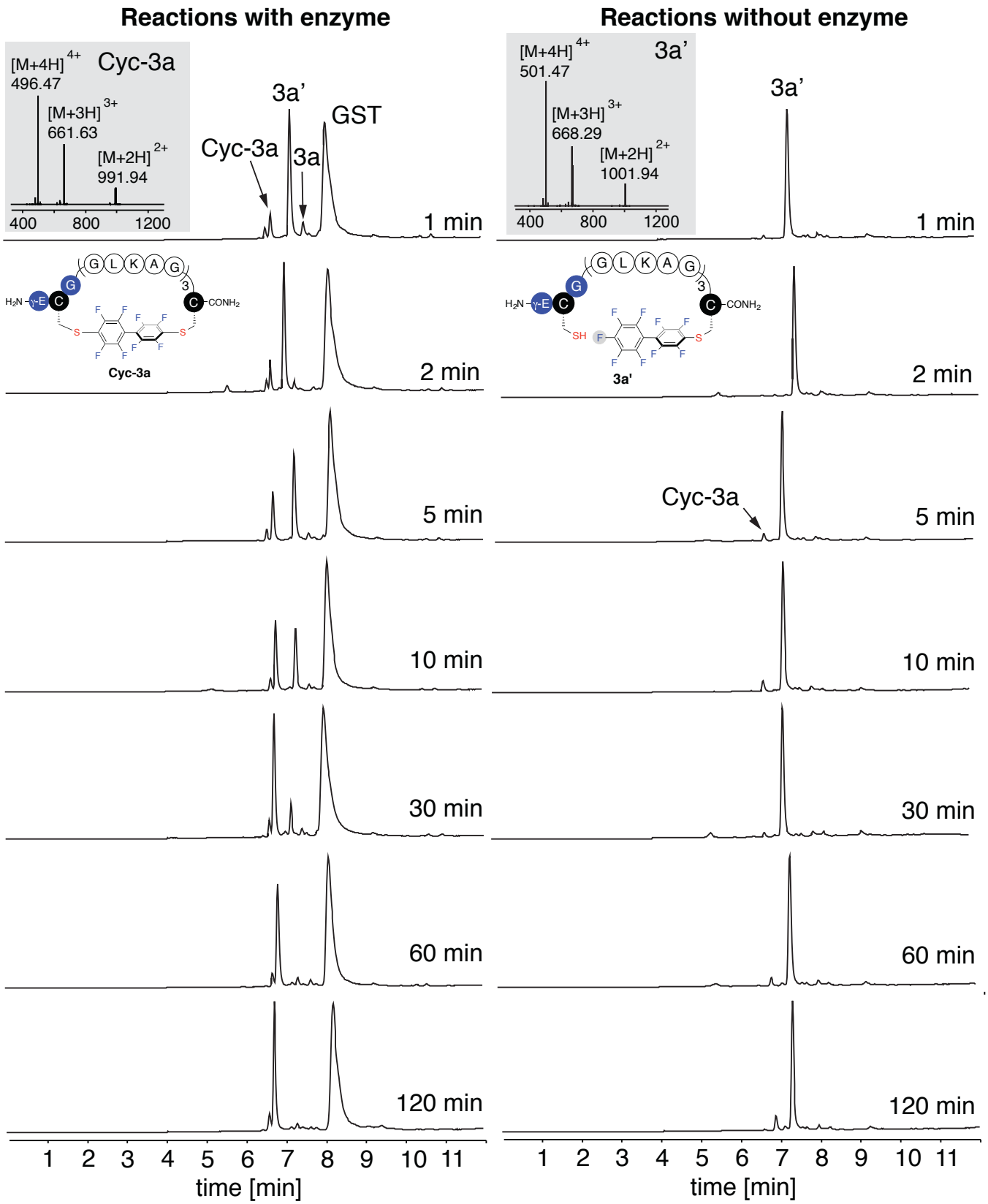


Figure S7. Macrocyclization of peptide **3a** with and without GST analyzed by LC-MS at different time points. Reaction conditions are described in Figure 4 of the manuscript.

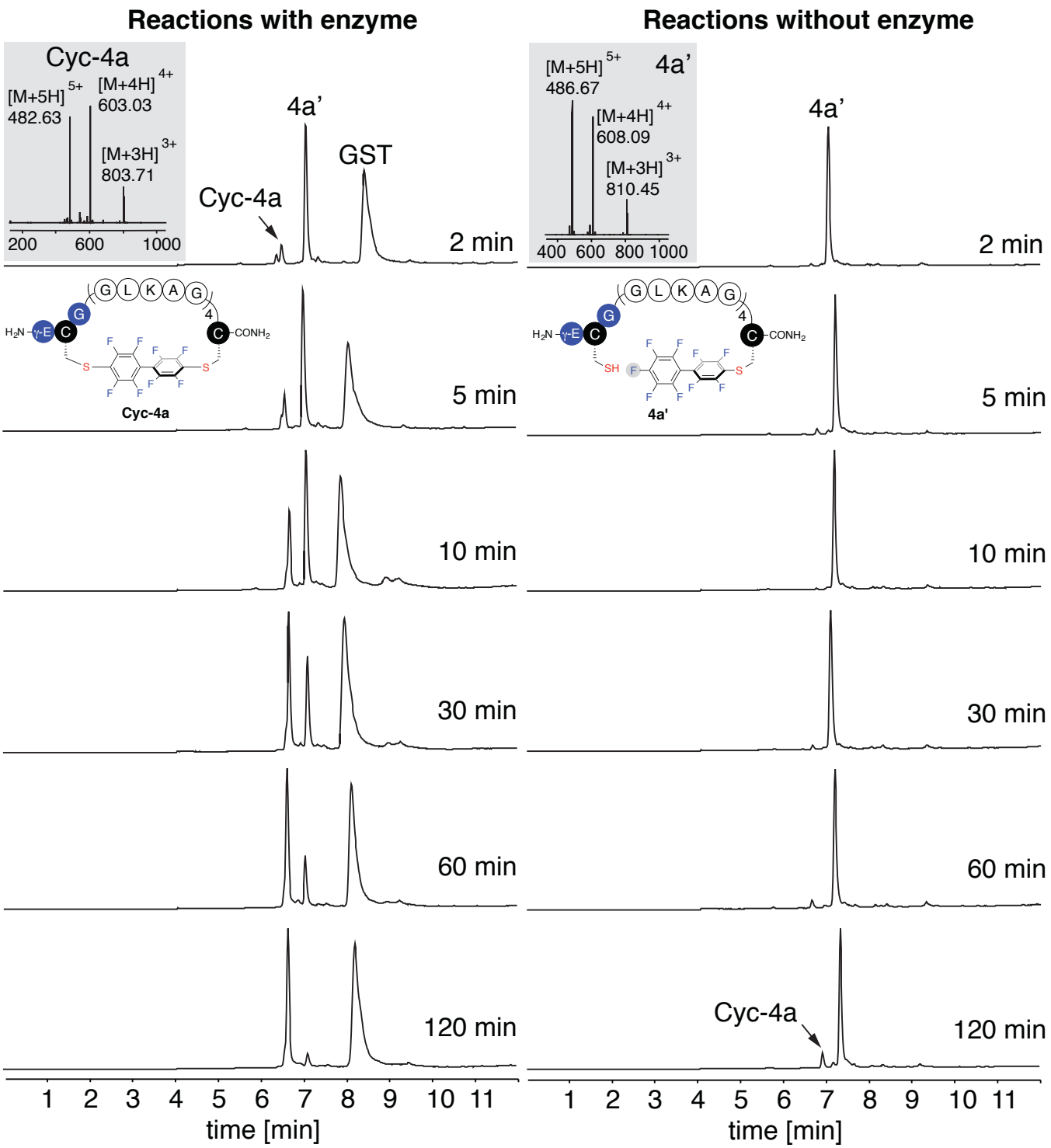


Figure S8. Macrocyclization of peptide **4a** with and without GST analyzed by LC-MS at different time points. Reaction conditions are described in Figure 4 of the manuscript.

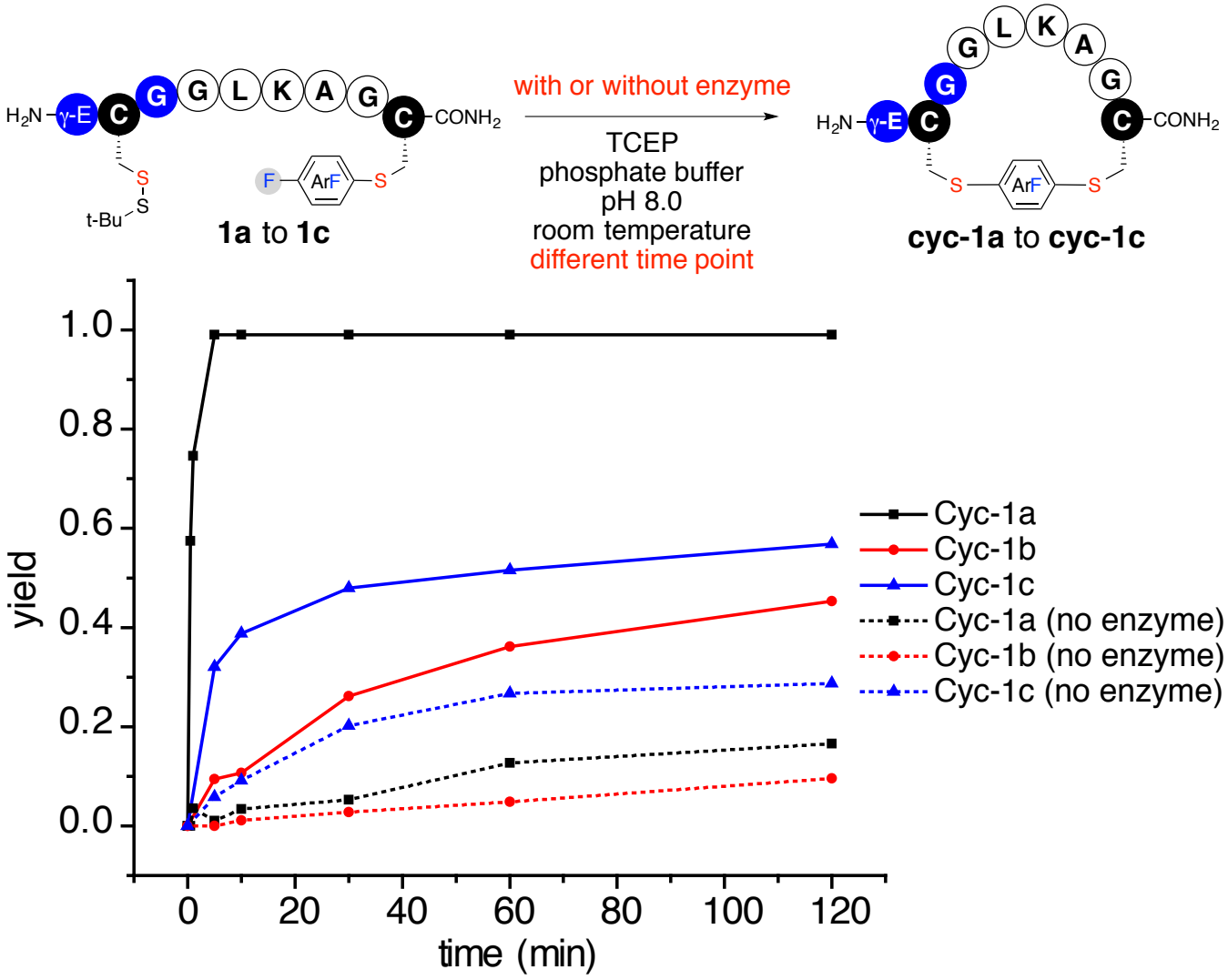
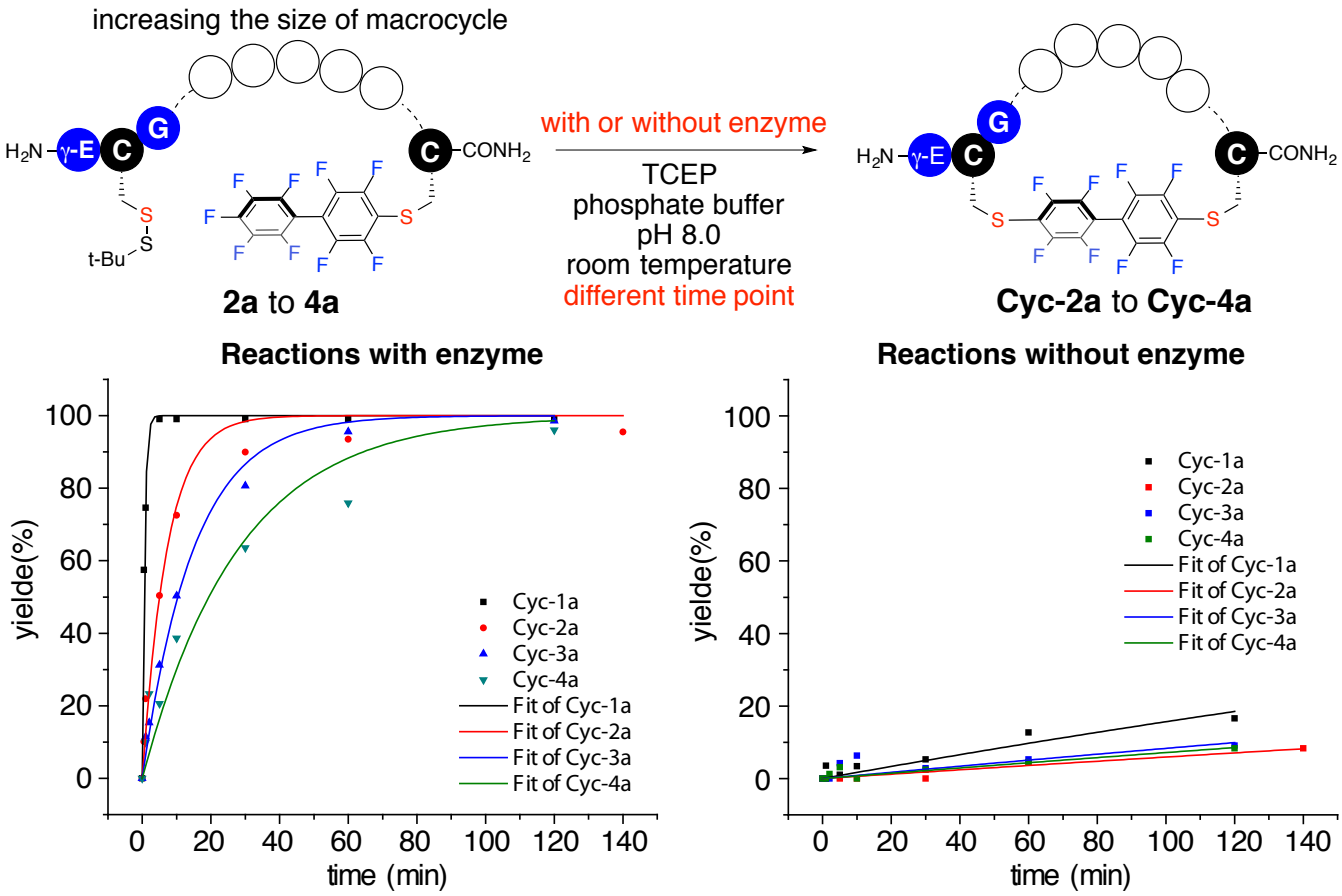


Figure S9. Kinetic profiles for macrocyclization of peptides **1a-c**. Yields were calculated from LC-MS analysis of the crude reaction mixture at different time points. All reactions with enzyme showed higher reaction rate than reactions without enzyme. Peptide **1a** showed highest reaction rate under GST catalysis. Reaction conditions are same as those describe in Figure 3 of the manuscript. At each time point, 10 μL of the reaction mixture was quenched with 100 μL of 50% A: 50% B and was then subjected to LC-MS analysis. LC-MS chromatograms are summarized in Figure S3-S5.



Peptide	$K_{enzyme(+)}$/min^{-1}	$K_{enzyme(-)}$ /min^{-1}	$K_{enzyme(+)}/K_{enzyme(-)}$
Cyc-1a	1.54	0.0017	906
Cyc-2a	0.13	0.00061	213
Cyc-3a	0.067	0.0009	74
Cyc-4a	0.036	0.00074	49

Figure S10. Kinetic profiles for macrocyclization of peptides **2a-4a**. Yields are obtained from LC-MS analysis of the crude reaction mixture at different time points. Using Origin software, apparent rate constants of reactions are obtained by fitting the data with first-order kinetics using function: yield = $1-e^{-Kt}$ where K is the apparent rate constant (min^{-1}) and t is the time (min). Reaction conditions are same as those describe in Figure 4 of the manuscript. At each time point, 10 μL of the reaction mixture was quenched with 100 μL of 50% A: 50% B and was then subjected to LC-MS analysis. LC-MS chromatograms are summarized in Figure S6-S8.

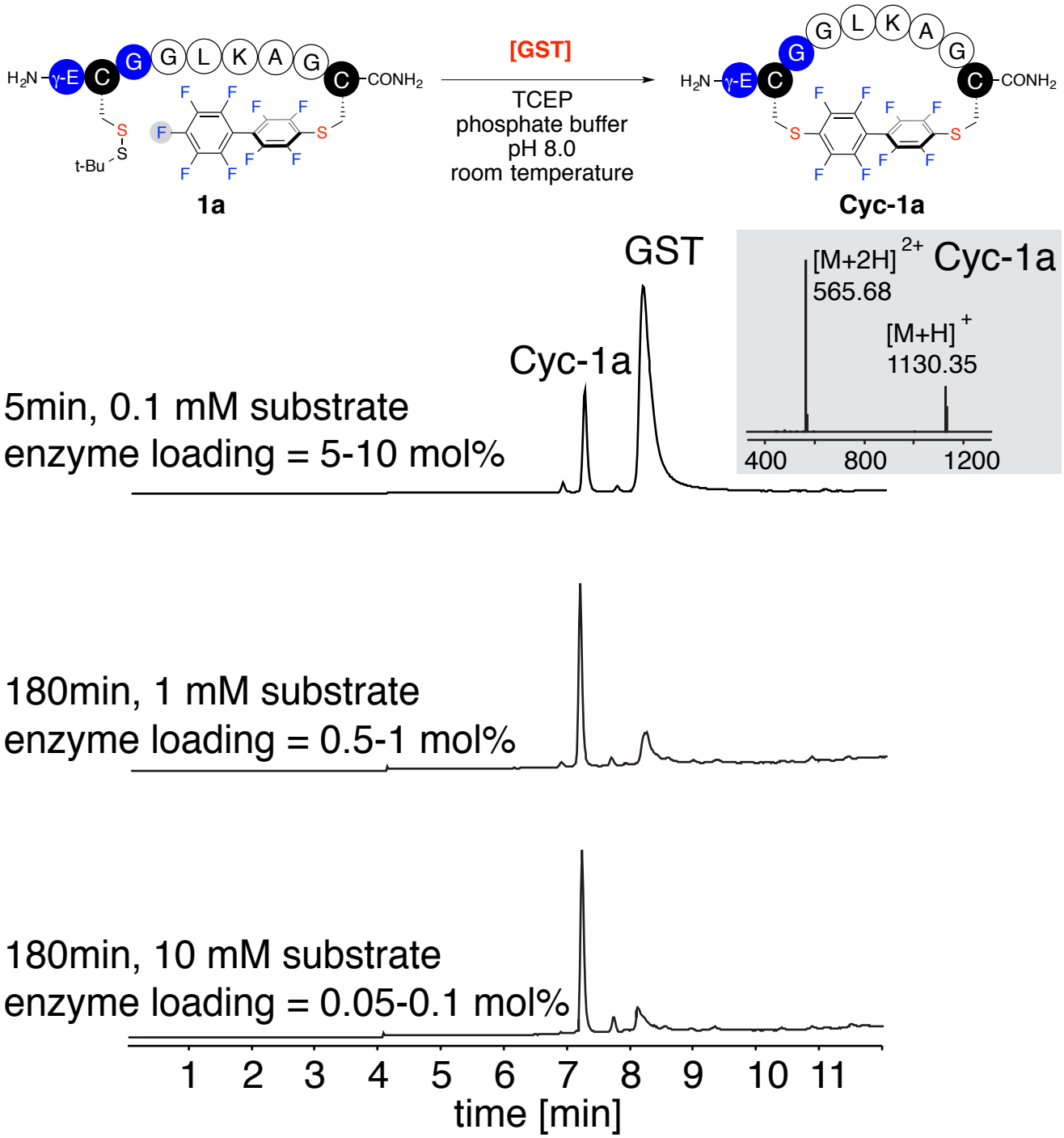


Figure S11. Macrocyclization of peptide **1a** at increased concentrations. Concentration of GST was kept constant at 0.2 mg/mL; LC-MS analysis of reaction where concentrations of substrate **1a** is 0.1 mM (top chromatogram), 1 mM (middle chromatogram), and 10 mM (bottom chromatogram). Reactions were performed in 0.1 mM phosphate buffer with 20 mM TCEP•HCl, pH 8.0 at room temperature.

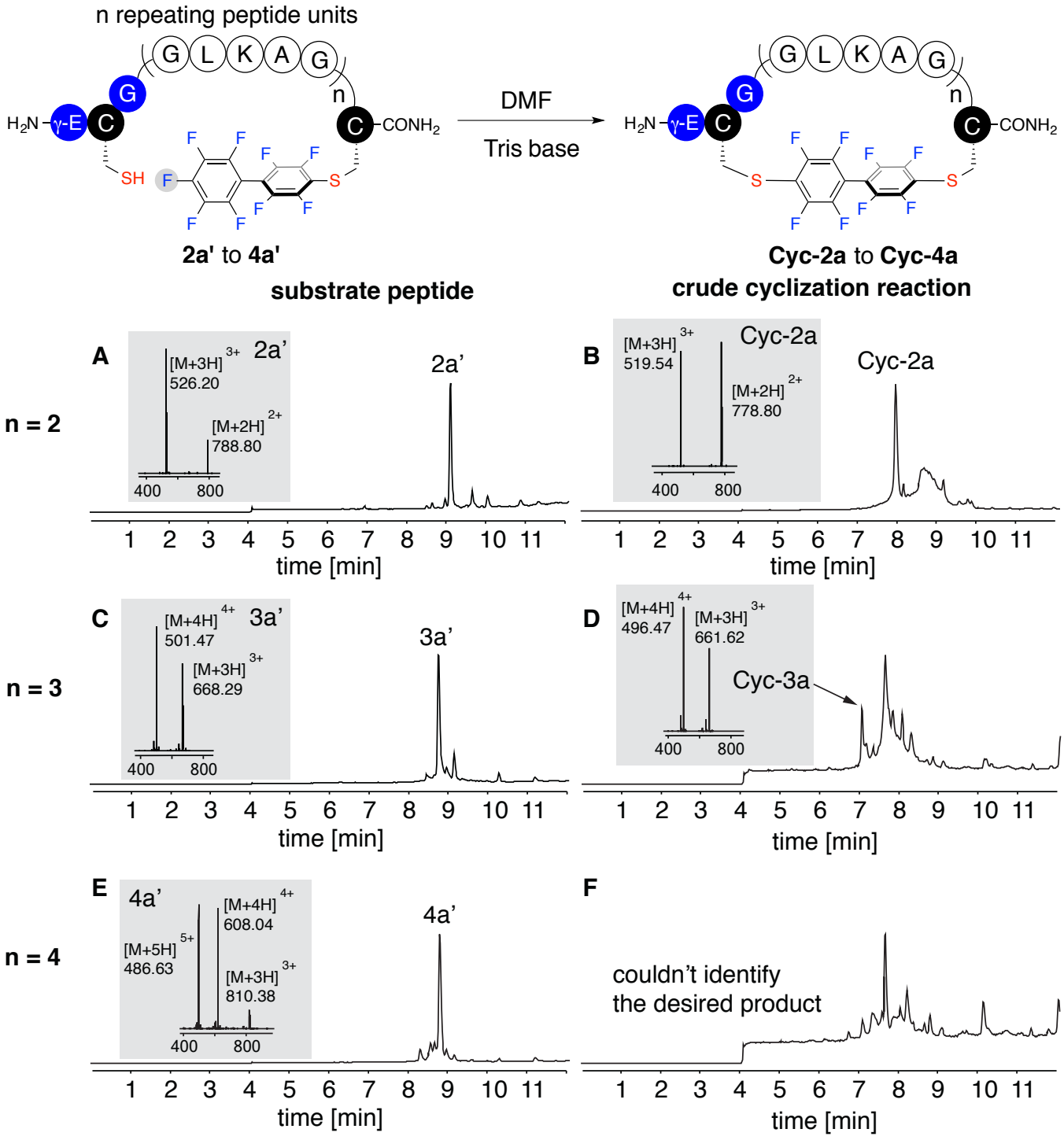


Figure S12. Macrocyclization of peptide **Lin-2a** to **Lin-4a** in DMF. Shown are two series of chromatograms, each series consist of TIC chromatograms for LC-MS analysis of the pure substrate peptide and crude reaction mixture. A and B are for peptide **Lin-2a**; C and D are for peptide **Lin-3a**; E and F are for peptide **Lin-4a**. Reaction conditions: 0.1 mM peptide **Lin-2a** to **Lin-4a**, 20 mM tris(hydroxymethyl)aminomethane, DMF, 2 hours at room tempterature. 20 μL of the each reaction mixture was quenched by addition of 200 μL of 50% A: 50% B, and subjected to LC-MS analysis.

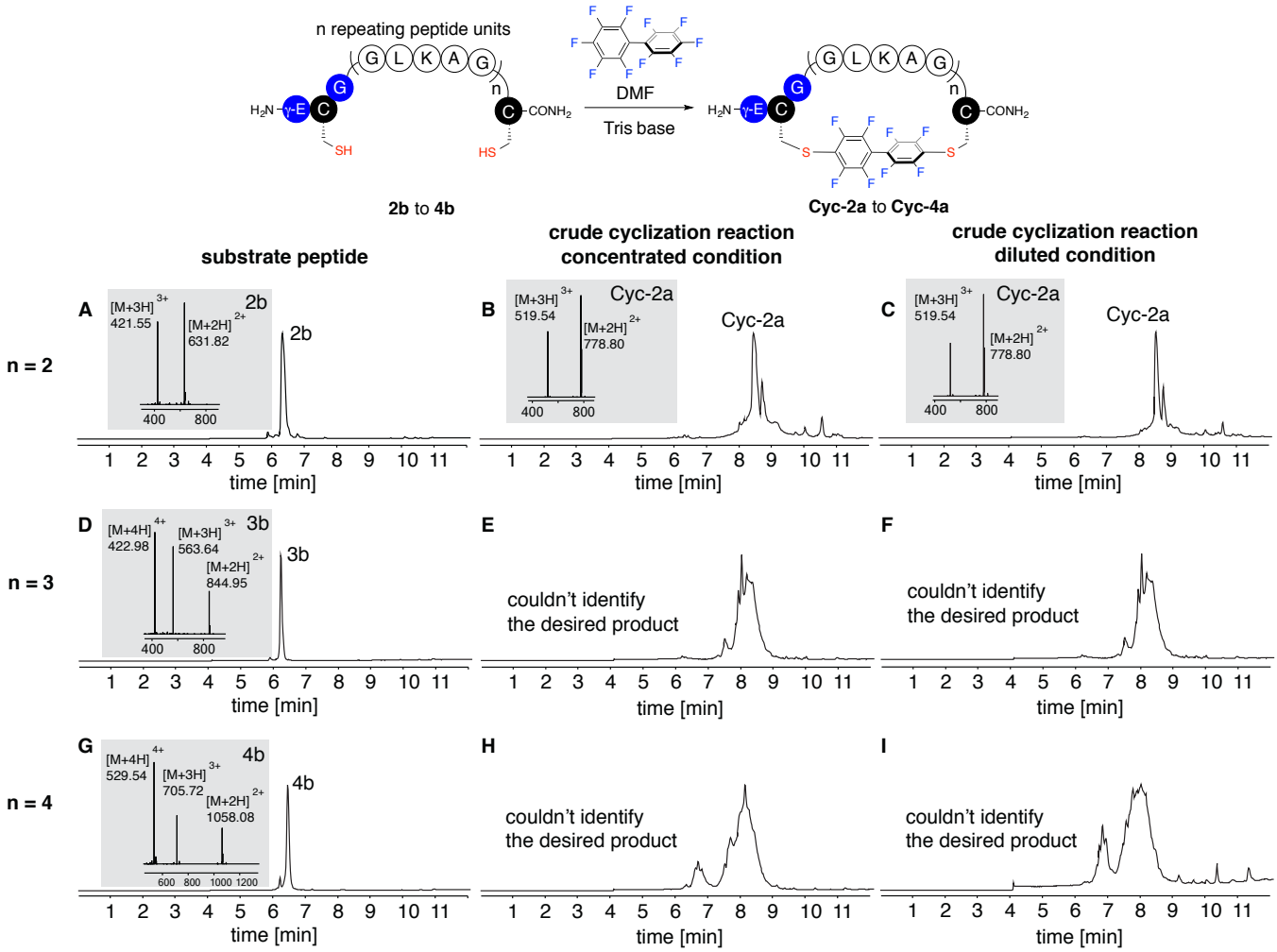
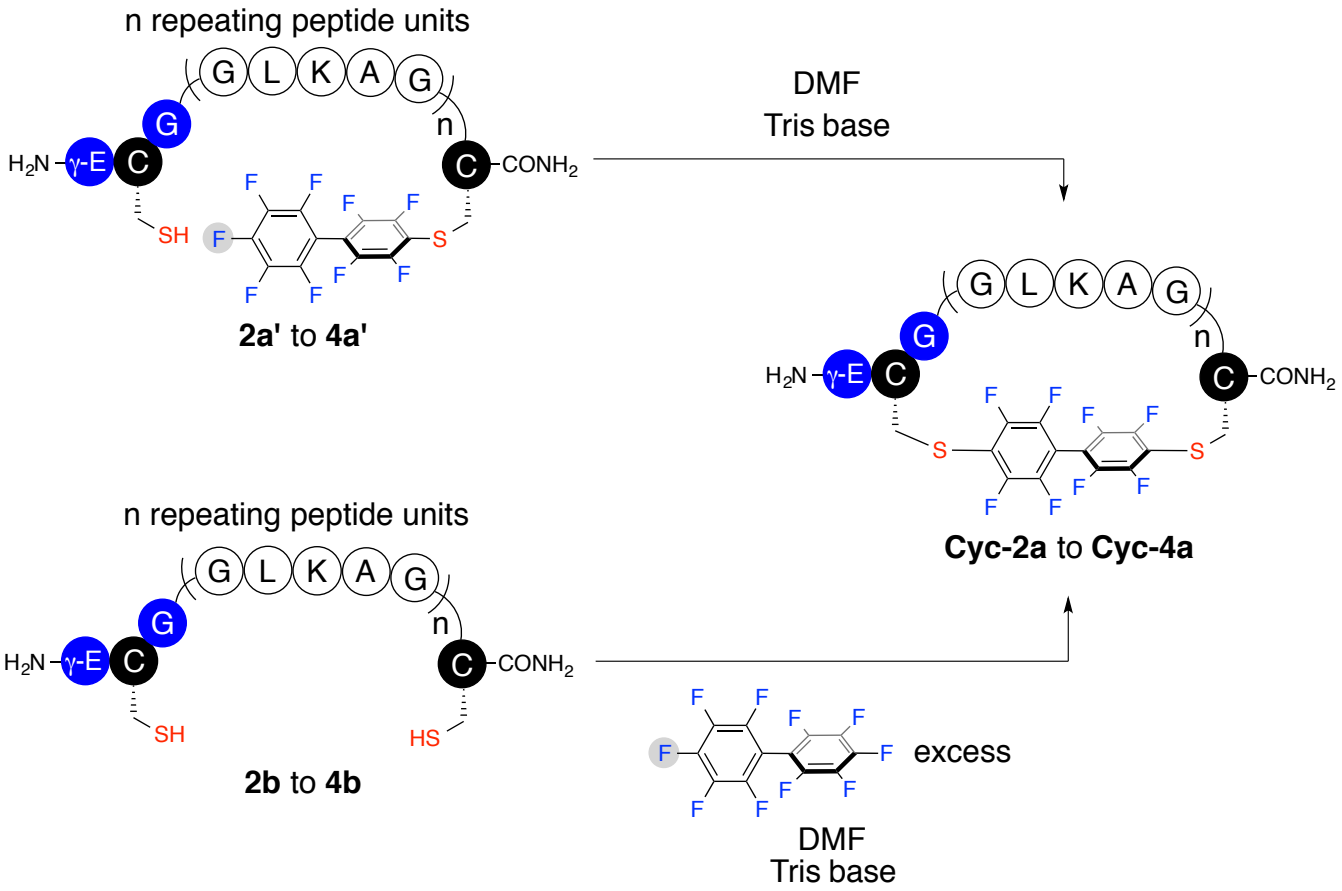


Figure S13. Macrocyclization of peptide **2b**−**4b** with decafluorobiphenyl in DMF. Shown are three series of chromatograms, each series consists of TIC chromatograms for LC-MS analysis of the pure substrate peptide, crude reaction mixture under concentrated condition, and crude reaction mixture under diluted condition. A, B, and C correspond to reaction with **2b**; D, E, and F - **3b**; G, H, and I - **4b**. Concentrated conditions: 2 mM substrate peptide, 2.5 mM decafluorobiphenyl, 20 mM tris(hydroxymethyl)aminomethane, 2 hours at room temperature. Diluted conditions: 0.1 mM substrate peptide, 0.125 mM decafluorobiphenyl, 10 mM tris(hydroxymethyl)aminomethane, 2 hours at room temperature. 20 μL of the reaction mixture was quenched by addition of 200 μL of 50% A: 50% B, and subjected to LC-MS analysis.



Peptide	Length (residues)	Yield (%)	Yield at diluted condition (%)
2a'	14	39	-
3a'	19	4	-
4a'	24	0	-
2b	14	37	42
3b	19	N/A	N/A
4b	24	N/A	N/A

Figure S14. Summary of macrocyclization yields of peptide **2a'** to **4a'** and peptide **2b** to **4b** in organic solvent.

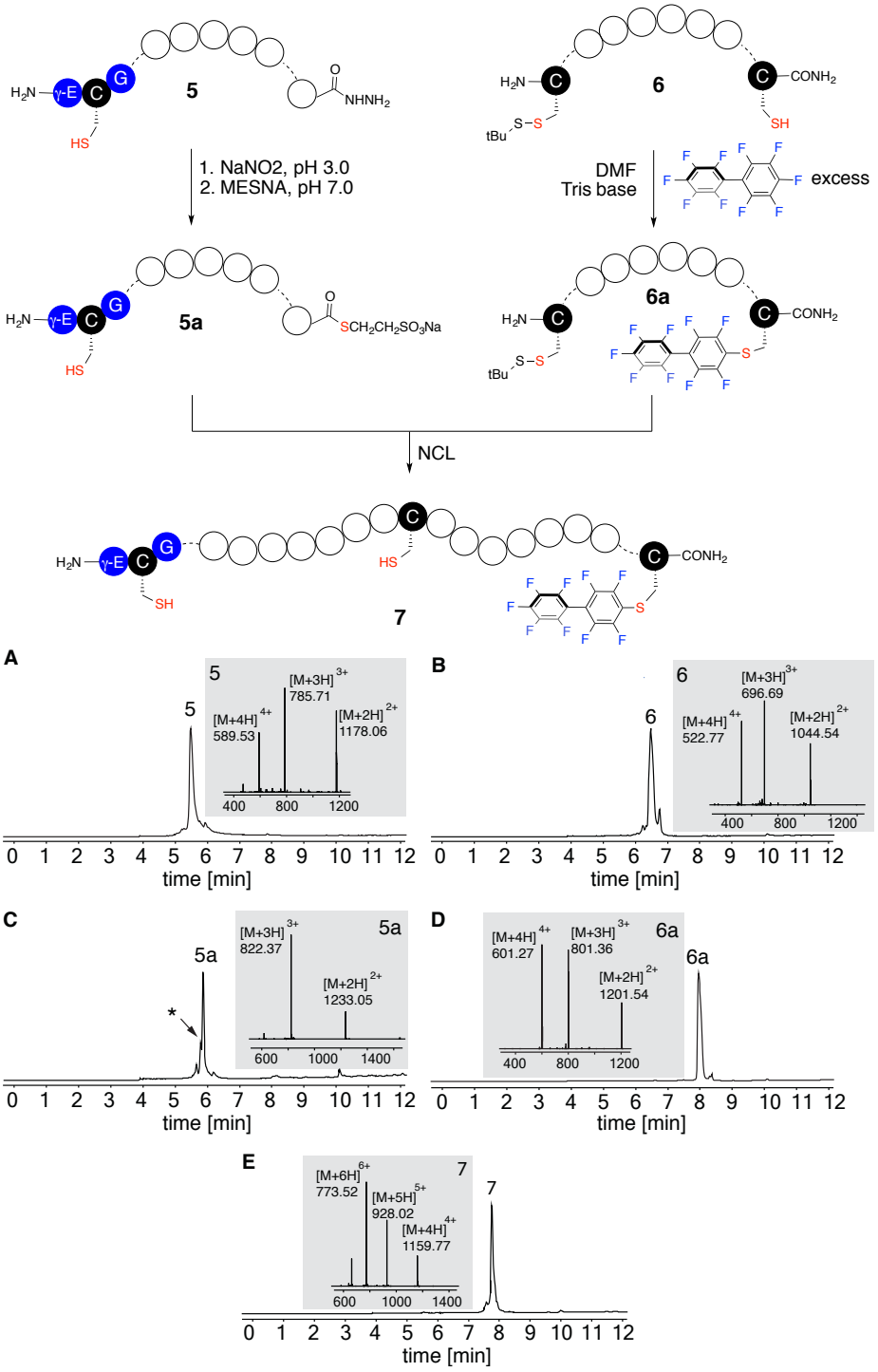


Figure S15. Synthesis and purification of peptide **7**. TIC chromatograms for LC-MS analysis of crude peptide thioester **5a** and RP-HPLC purified peptides **5**, **6**, **6a** and **7**. Peak labeled with "*" denotes thiolactone formed at the GSH cysteine from peptide **5a**. Reaction conditions are described in section 4 of supporting information.

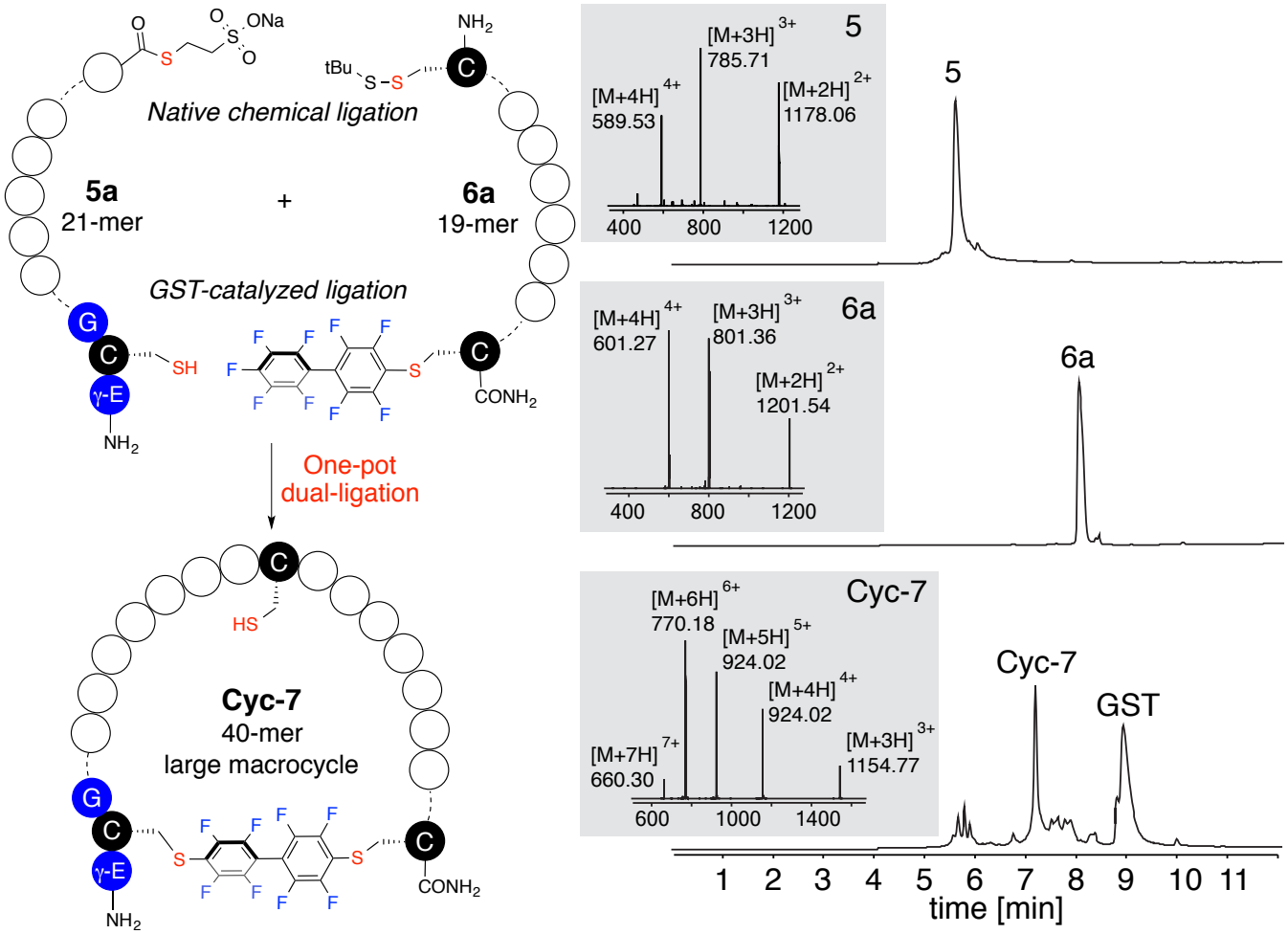


Figure S16. One-pot dual-ligation combining NCL and GST-catalyzed ligation for the synthesis of macrocyclic peptide **Cyc-7**. LC-MS analysis of crude thioester **5a** from oxidation of peptide hydrazide **5** (top chromatogram), purified peptide **6a** (middle chromatogram), and crude dual-ligation reaction mixture (bottom chromatogram). Reaction conditions for one-pot dual ligation: 1 mM peptide **5a**, 1 mM peptide **6a**, 1 mg/mL GST M5, 0.2 M phosphate, 20 mM TCEP•HCl, pH 8.0.

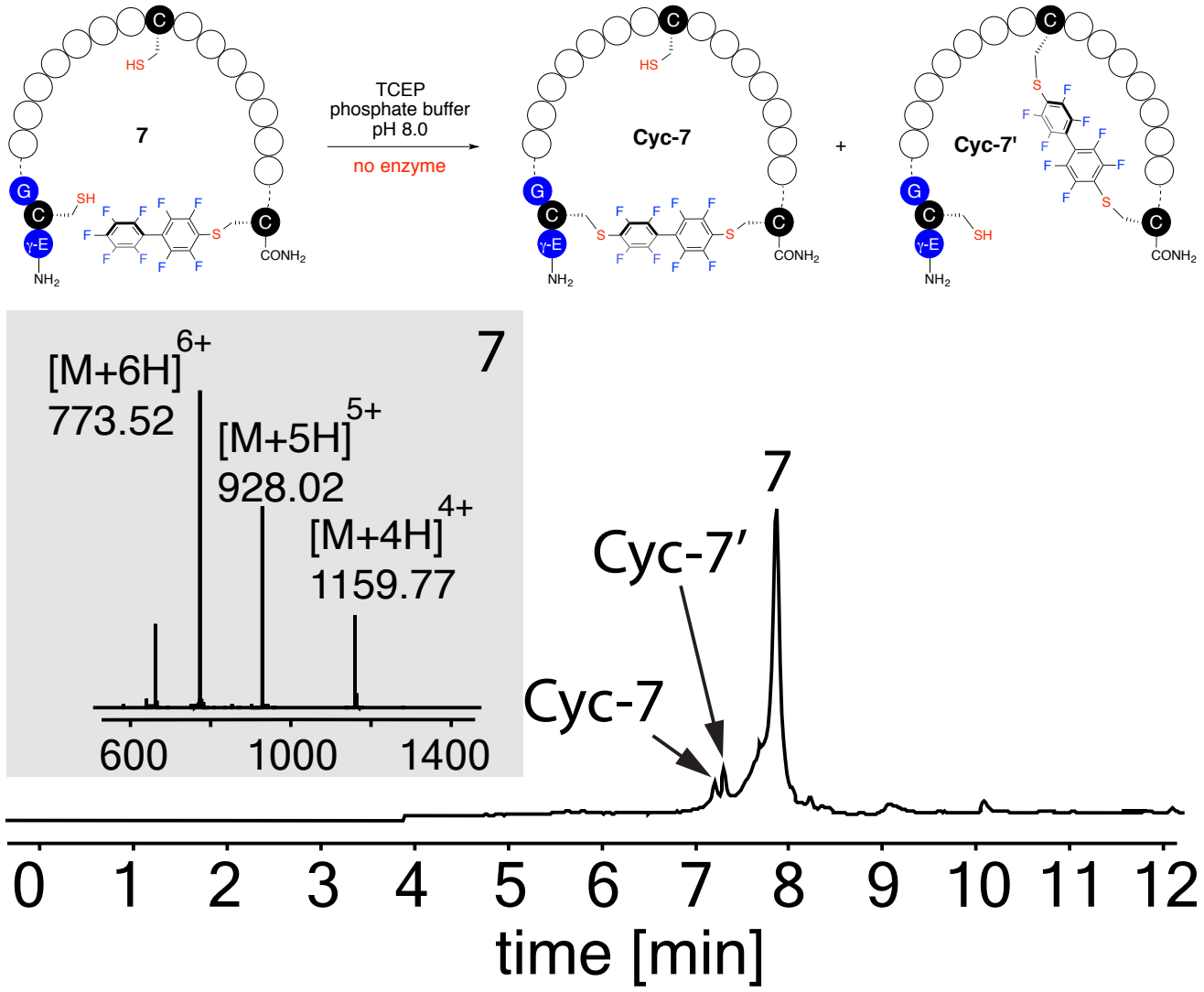
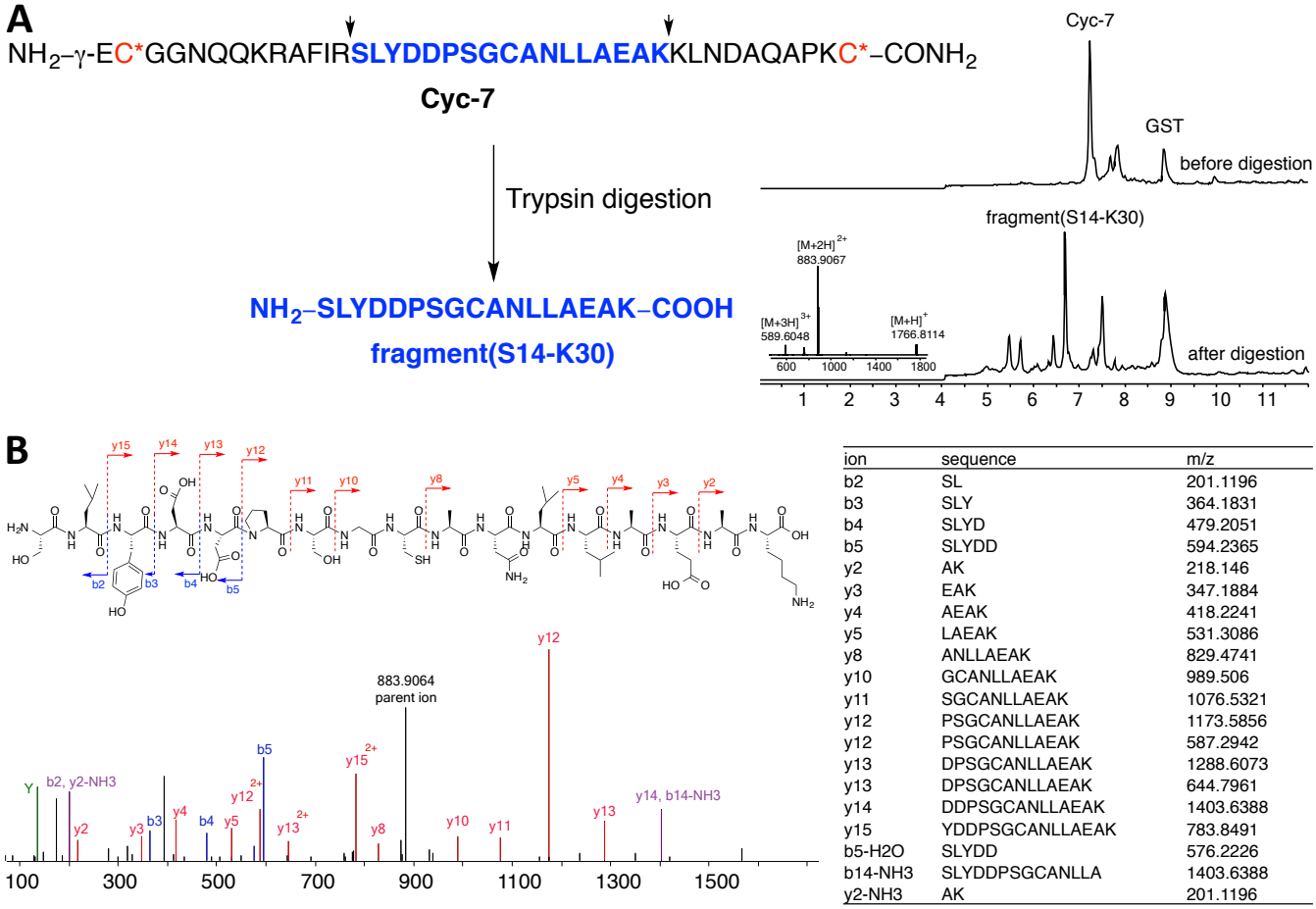


Figure S17. Cyclization of peptide **7** without GST catalyst. Reaction conditions: 0.1 mM peptide **7**, 0.2 M phosphate, 20 mM TCEP•HCl, pH 8.0, room temperature, 2 hours. The reaction lacks regioselectivity producing two isomers **Cyc-7** and **Cys-7'** with arylation at both GSH cysteine and the NCL linkage cysteine, respectively. Peak corresponding to **Cyc-7** was assigned based on retention time identity of the product obtained in the GST-catalyzed cyclization reaction shown in Figure 6.



ion	sequence	m/z
b2	SL	201.1196
b3	SLY	364.1831
b4	SLYD	479.2051
b5	SLYDD	594.2365
y2	AK	218.146
y3	EAK	347.1884
y4	AEAK	418.2241
y5	LAEAK	531.3086
y8	ANLLAEAK	829.4741
y10	GCANLLAEAK	989.506
y11	SGCANLLAEAK	1076.5321
y12	PSGCANLLAEAK	1173.5856
y12	PSGCANLLAEAK	587.2942
y13	DPSGCANLLAEAK	1288.6073
y13	DPSGCANLLAEAK	644.7961
y14	DDPSGCANLLAEAK	1403.6388
y15	YDDPSGCANLLAEAK	783.8491
b5-H2O	SLYDD	576.2226
b14-NH3	SLYDDPSGCANLLA	1403.6388
y2-NH3	AK	201.1196

Figure S18. Trypsin digestion and MS/MS analysis of **Cyc-7**. Results show that GST regio-selectively catalyzed the cyclization at the GSH cysteine, and the central cysteine remained unmodified after GST-catalyzed cyclization reaction. (**A**) LC-MS analysis of the crude GST-catalyzed ligation product before trypsin digestion (top chromatogram) and after trypsin digestion (bottom chromatogram). Reaction conditions for cyclization reaction were same as those described in Figure 5. For trypsin digestion, 20 μL of the crude cyclization product was mixed with 20 μL of trypsin stock (100 μg/mL) and 1 μL of 1.5 M NH_4HCO_3. The mixture was incubated at 37 °C for 3 hours and was then subjected to LC-MS/MS analysis. The major product after trypsin digestion was identified as fragment (S14-K30), its mass spectrum is shown as the inset in bottom chromatogram. Cross-linked cysteines are labeled as C* and highlighted in red; arrows indicate the sites of trypsin digestion; sequence of the digested product fragment (S14-K30) is highlighted in blue. (**B**) MS/MS analysis of fragment (S14-K30), masses and sequences of identified ion species are listed in the table. MS/MS result confirmed that the central cysteine C22 was not modified.