Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Side-chain-to-tail cyclization of ribosomally derived peptides promoted by aryl and alkyl amino-functionalized unnatural amino acids

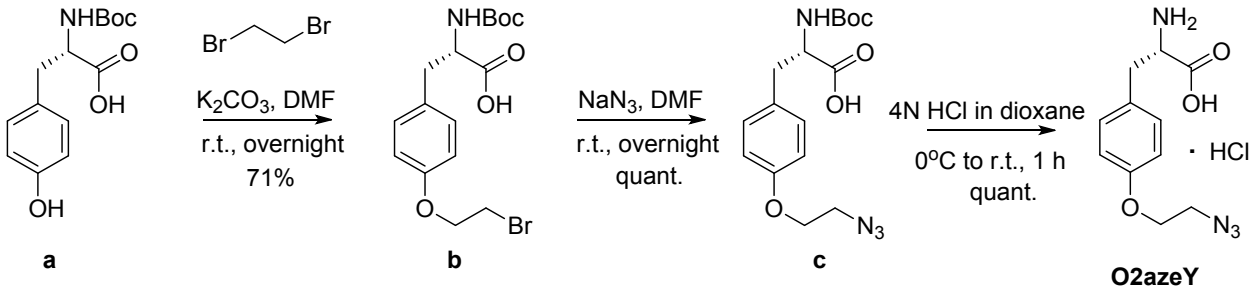
John R. Frost[§], Zhijie Wu[§], Yick Chong Lam[§], Andrew E. Owens and Rudi Fasan*

Department of Chemistry, University of Rochester, Hutchinson Hall, Rochester NY, 14627

[§]These authors contributed equally to this work

* Corresponding author: rfasan@ur.rochester.edu

Table of Contents



Scheme S1. Synthesis of O-2-azidoethyl-tyrosine (O2azeY).

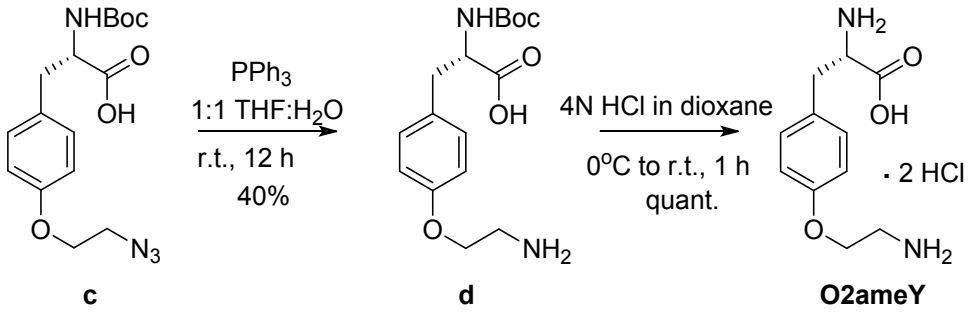
Synthesis of O-2-azidoethyl-tyrosine, O2azeY

Commercially available (tert-butoxycarbonyl)-L-tyrosine, (2 grams, 7.109 mmol, 1 eq) was added to a dry 100 mL round bottom flask under argon. The compound was dissolved in 20 mL anhydrous DMF and 1.23 mL of 1,2 dibromoethane (2.671 grams, 14.218 mmol, 2 eq) was added, followed by addition of K_2CO_3 (2.947 grams, 21.327 mmol, 3 eq). The reaction was stirred at room temperature for 12 hours. Following completion the solution was filtered by vacuum filtration to remove the solid, and the filtrate was transferred to a separatory funnel and washed 1x 40 mL of H_2O. The aqueous layer was acidified to pH 4 and was extracted with 100 mL of EtOAc. The organic layer was dried by $MgSO_4$ and the volatiles were removed *in vacuo*. The resulting mixture was purified by flash column chromatography using a gradient of 15:4:1 to 10:4:1 hexane: ethyl acetate: acetic acid to afford N-Boc-O-(2-bromoethyl)-tyrosine (**b**) as a yellow oil (1.95 grams, 5.022 mmol, 70.6% yield). ^{1}H NMR (400 MHz, 25 °C, $CDCl_3$): δ 6.99 (d, J = 8.4 Hz, 2H), 6.76 (d, J = 8.8 Hz, 2H), 5.15 (d, J = 8.4 Hz, 1H), 4.53 (d, J = 7.2 Hz, 1H), 4.36 (q, J = 4.8, 2H), 3.43 (t, J = 5.6 Hz, 2H), 3.02 (t, J = 6 Hz, 2H), 1.41 (s, 9H).

N-Boc-O-(2-bromoethyl)-tyrosine (0.8 grams, 2.06 mmol, 1 equiv.) was dissolved in 17 mL of anhydrous DMF in a 100 mL round bottom flask under argon. Sodium azide (0.147g, 2.26 mmol, 1.1 equiv.) was added. The reaction was heated at 50 °C, and stirred for 12 hours. The reaction mixture was then transferred to the separatory funnel and combined with 60 mL of deionized water. The resulting mixture was acidified to pH 4 and extracted with 2 x 50 mL of EtOAc. The organic layer was dried over $MgSO_4$ and the volatiles were removed *in vacuo* to yield N-Boc-O-(2-azidoethyl)-tyrosine (**c**) as a yellow oil (720 mg, 2.06 mmol, >99% yield) ^{1}H NMR (400 MHz, 25 °C, $CDCl_3$): δ 6.99 (d, J = 8.4 Hz, 2H), 6.78 (d, J = 8.4 Hz, 2H), 4.97 (d, J =

7.6 Hz, 1H), 4.55 (d, J = 6.8 Hz, 1H), 4.24 (t, J = 4.8, 2H), 3.45 (q, J = 5.2 Hz, 2H), 3.01 (t, J = 6 Hz, 2H), 1.42 (s, 9H).

N-Boc-O-(2-azidoethyl)-tyrosine (720 mg, 2.06 mmol, 1 eq) was dissolved in a solution of 4N HCl in dioxane (20 mL). The reaction was stirred at room temperature for 4 hours until completion, as monitored by TLC. The completed reaction was dried under reduced pressure to yield the hydrochloride salt of O-2-azidoethyl-tyrosine as a white solid (590 mg, quantitative). ^{1}H NMR (500 MHz, 25°C, CD_3OD): δ 7.10 (d, J=8.7 Hz, 2H), 6.80 (d, J=8.6 Hz, 2H), 4.35 (t, J=5.28 Hz, 2H), 4.28 (dd, J=7.3, 6.2 Hz, 1H), 3.6-3.5 (m, 2H), 3.19 (dd, J=14.50, 6.26 Hz, 1H), 3.11 (dd, 14.5, 7.5 Hz, 1H).^{13}C NMR (126 MHz, 25°C, CD_3OD) δ 170.1, 158.5, 131.5, 125.5, 116.9, 65.9, 55.4, 50.6, 36.7. MS (ESI) for $C_{11}H_{14}N_4O_3$ $[M+H]^+$ calc: 251.25; $[M+H]^+$ obs: 251.1.



c **d** **O2ameY**

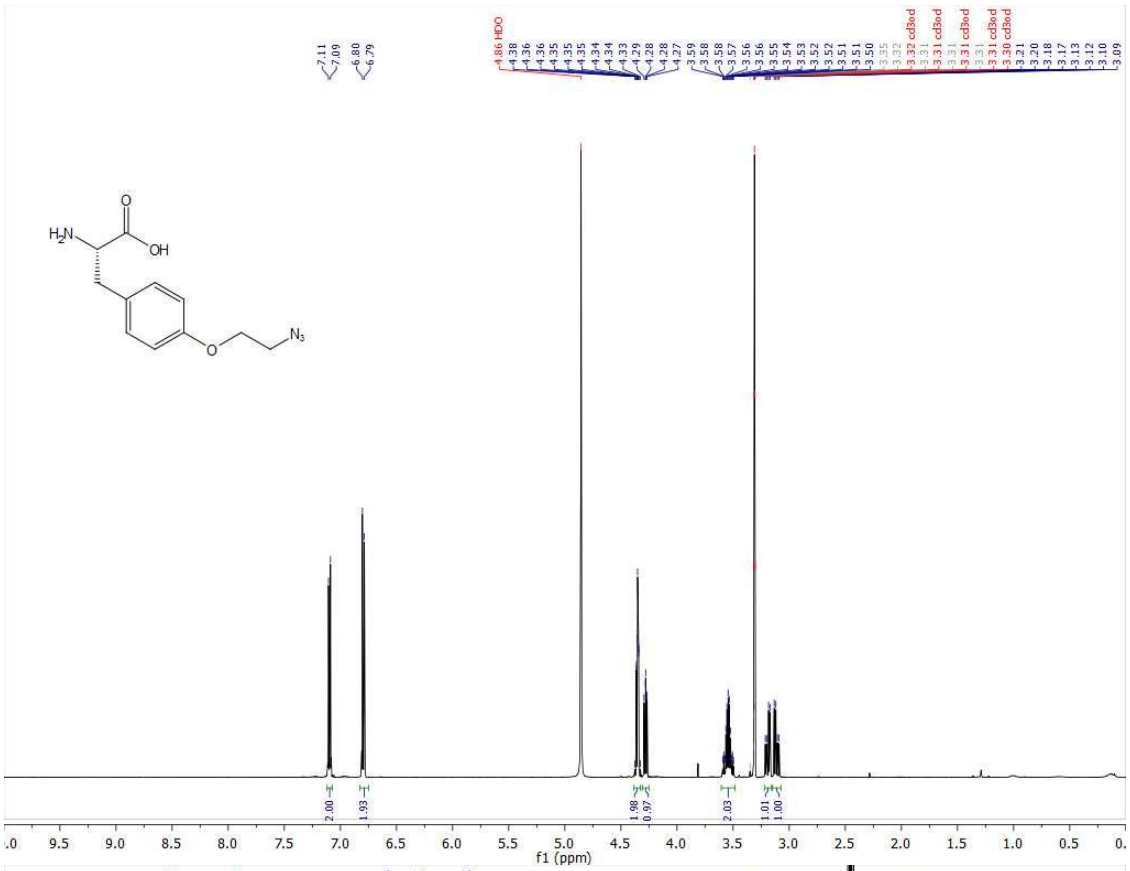
Scheme S2. Synthesis of O-2-aminoethyl-tyrosine (O2ameY).

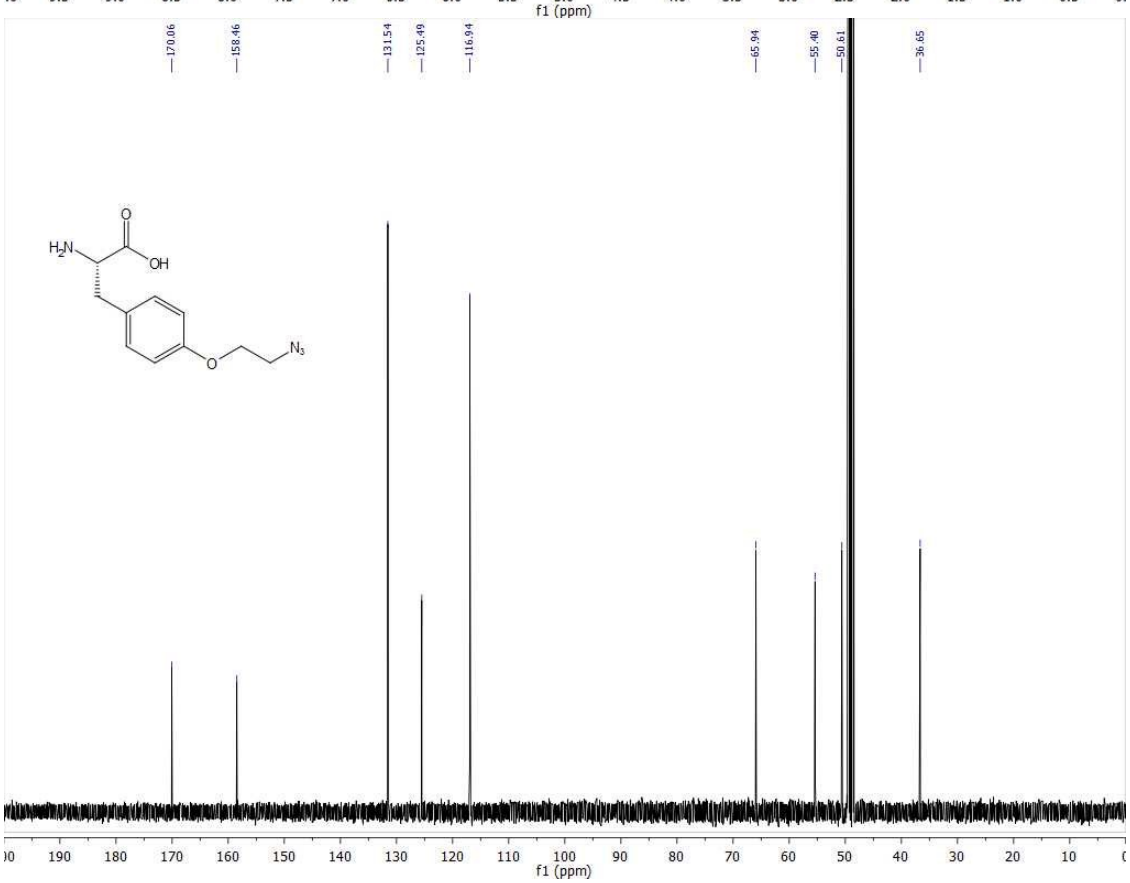
Synthesis of O-2-aminoethyl-tyrosine, O2ameY

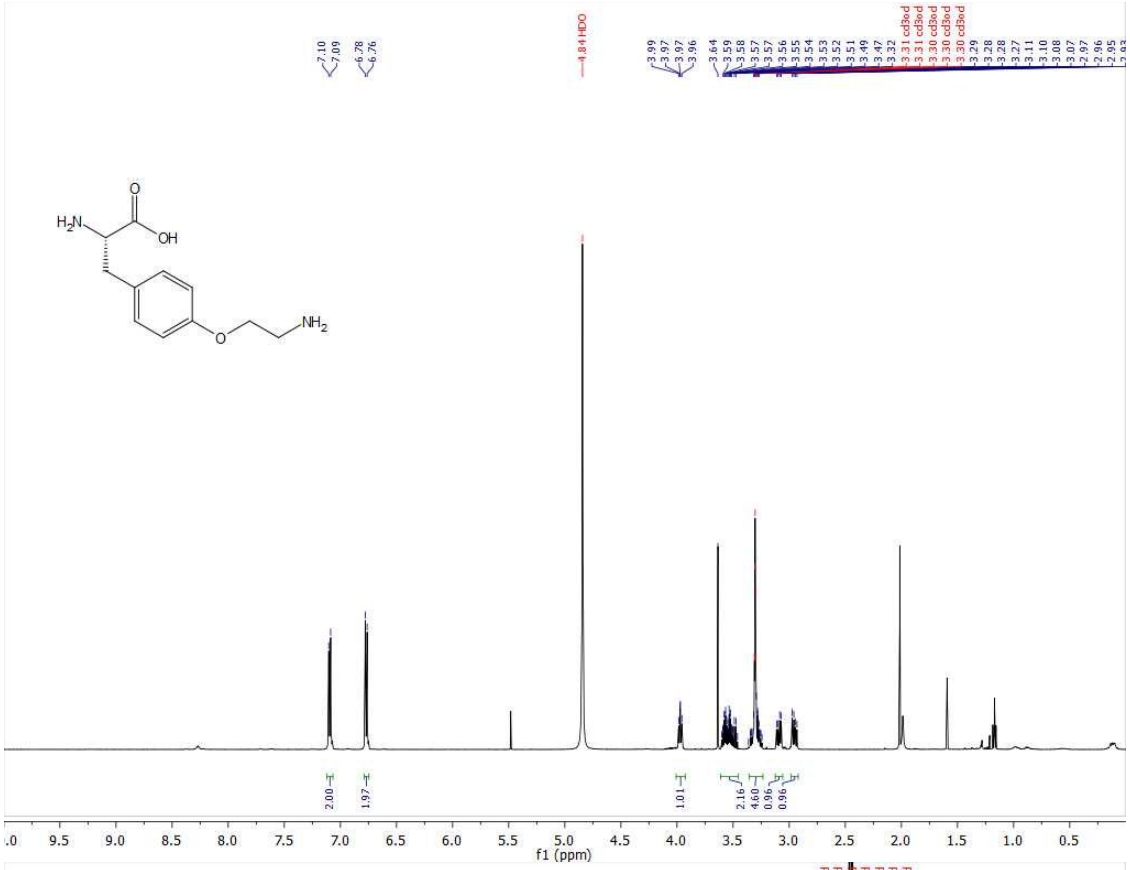
N-Boc-O-(2-azidoethyl)-tyrosine (**c**), (0.9 grams, 2.565 mmol, 1 eq) was dissolved in 25mL of THF in a 100 mL round-bottom flask and the solution was cooled to 0°C under argon. Triphenylphosphine (0.8 grams, 3.079 mmol, 1.2 eq) was added and the reaction was stirred at 0°C for 5 hours. 25 mL of de-ionized H_2O was then added and the reaction was stirred for 16 hours at ambient temperature and then heated at 50 °C for an additional 3 hours. The solvent was removed *in vacuo* and the reaction mixture was extracted with EtOAc 3 x 50 mL and dried with $MgSO_4$. The volatiles were removed under reduced pressure to afford the crude product, which was purified on silica gel using 10:9:1 hexane: ethyl acetate: acetic acid as the solvent. The purified N-Boc-O-(2-aminoethyl)-tyrosine (**4**) was isolated as a yellow oil (0.367grams, 1.131

mmol, 44% yield). ^{1}H NMR (500 MHz, 25 °C, $CDCl_3$): δ 7.07 (d, J = 8 Hz, 2H), 6.78 (d, J = 8.5 Hz, 2H), 5.11 (s, 1H), 4.21 (d, J = 6.5 Hz, 1H), 3.85 (s, 2H), 3.32 (s, 2H), 2.93 (dd, J = 8 Hz, 2H), 1.42 (s, 9H). MS (ESI) for $C_{16}H_{24}N_2O_5Na$ $[M+Na]^+$ calc: 347.38; $[M+Na]^+$ obs: 347.51.

N-Boc-O-(2-aminoethyl)-tyrosine (**d**) (160 mg, 0.493 mmol, 1 eq) was dissolved in a solution of 4 N HCl in dioxane (3 mL) at 0 °C. The reaction was stirred at room temperature for 1 hour. Following completion, the solvent was removed *in vacuo* and afforded the dihydrochloride salt of O-2-aminoethyl-tyrosine (O2ameY) as a yellow oil (146 mg, quantitative). ^{1}H NMR (500 MHz, 25°C, CD_3OD): δ 7.10 (d, J=8.7 Hz, 2H), 6.79 (d, J= 8.3Hz, 2H), 3.97 (t, J=7.2 Hz, 1H), 3.60-3.45 (m, 2H), 3.34-3.24 (m, 2H), 3.09 (dd, J=13.6, 6.6 Hz, 1H), 2.95 (dd, J=13.9, 7.6 Hz, 1H). ^{13}C NMR (126 MHz, 25°C, CD_3OD): δ 169.9, 158.3, 131.6, 126.1, 116.8, 61.2, 56.1, 43.0, 37.9. MS (ESI) for $C_{11}H_{17}N_2O_3$ $[M+H]^+$ calc: 225.15; $[M+H]^+$ obs: 225.47.


7.11
7.09
6.80
6.79
4.86 HDO
4.38
4.36
4.35
4.34
4.33
4.29
4.28
4.27
3.59
3.58
3.57
3.56
3.55
3.54
3.53
3.52
3.51
3.50
3.35
3.32
3.32 cd3od
3.31
3.31 cd3od
3.31
3.31 cd3od
3.31
3.31 cd3od
3.30 cd3od
3.21
3.20
3.18
3.17
3.13
3.12
3.10
3.09
2.00
1.93
1.98
0.97
2.03
1.01
1.00
f1 (ppm)


170.06
158.46
131.54
125.49
116.94
65.94
55.40
50.61
36.65
f1 (ppm)


7.10
7.09
6.78
6.76
4.84 HDO
3.99
3.97
3.97
3.96
3.64
3.59
3.58
3.57
3.57
3.56
3.55
3.54
3.53
3.52
3.51
3.49
3.47
3.32
3.31 cd3od
3.31 cd3od
3.30 cd3od
3.30 cd3od
3.30 cd3od
3.29
3.28
3.28
3.27
3.11
3.10
3.08
3.07
2.97
2.96
2.95
2.93
O
H2N
OH
NH2
O
2.00
1.97
1.01
2.16
4.60
0.96
0.96
f1 (ppm)

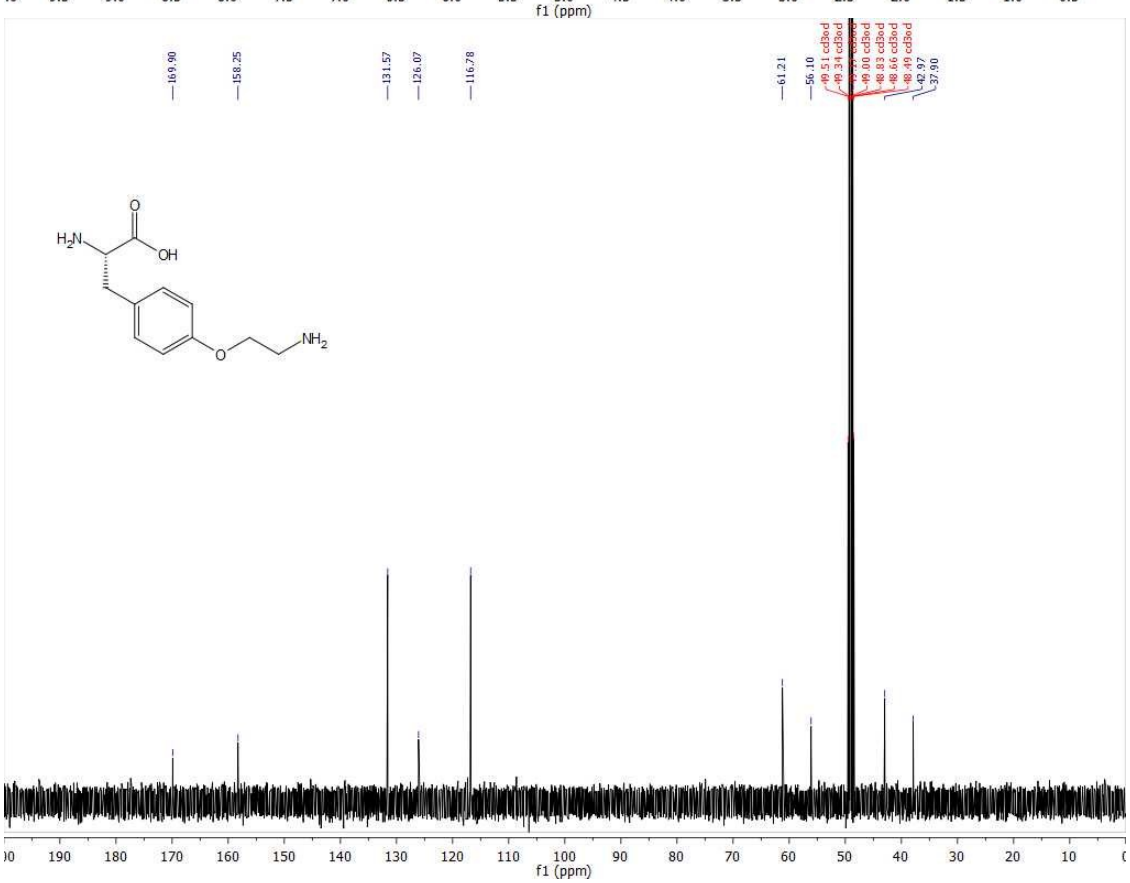

169.90
158.25
131.57
126.07
116.78
61.21
56.10
49.51 cd3od
49.34 cd3od
49.00 cd3od
48.83 cd3od
48.66 cd3od
48.49 cd3od
42.97
37.90
O
H2N
OH
NH2
O
f1 (ppm)

Figure S1. MALDI-TOF MS spectrum of precursor protein CBD-10mer(OpgY) after incubation with thiophenol under standard conditions. Only the acyclic product ('a') is observed indicating the lack of reactivity of the lysine comprised within the target peptide sequence toward inducing side-chain-to-tail cyclization under the applied conditions.

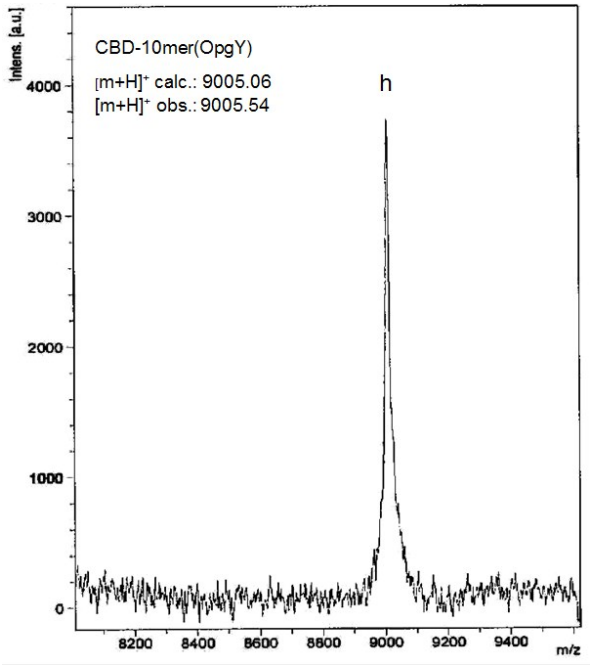


Figure S2. SDS-PAGE protein gels (left column) and densitometric analysis (graphs on right column) for benzyl mercaptan-induced cyclization reactions at pH 7.5, 8.2, and 9.0 at various time points. FL = full-length protein, GyrA = cleaved GyrA intein, CBD = CBD = CBD-fused macrocycle and acyclic product.

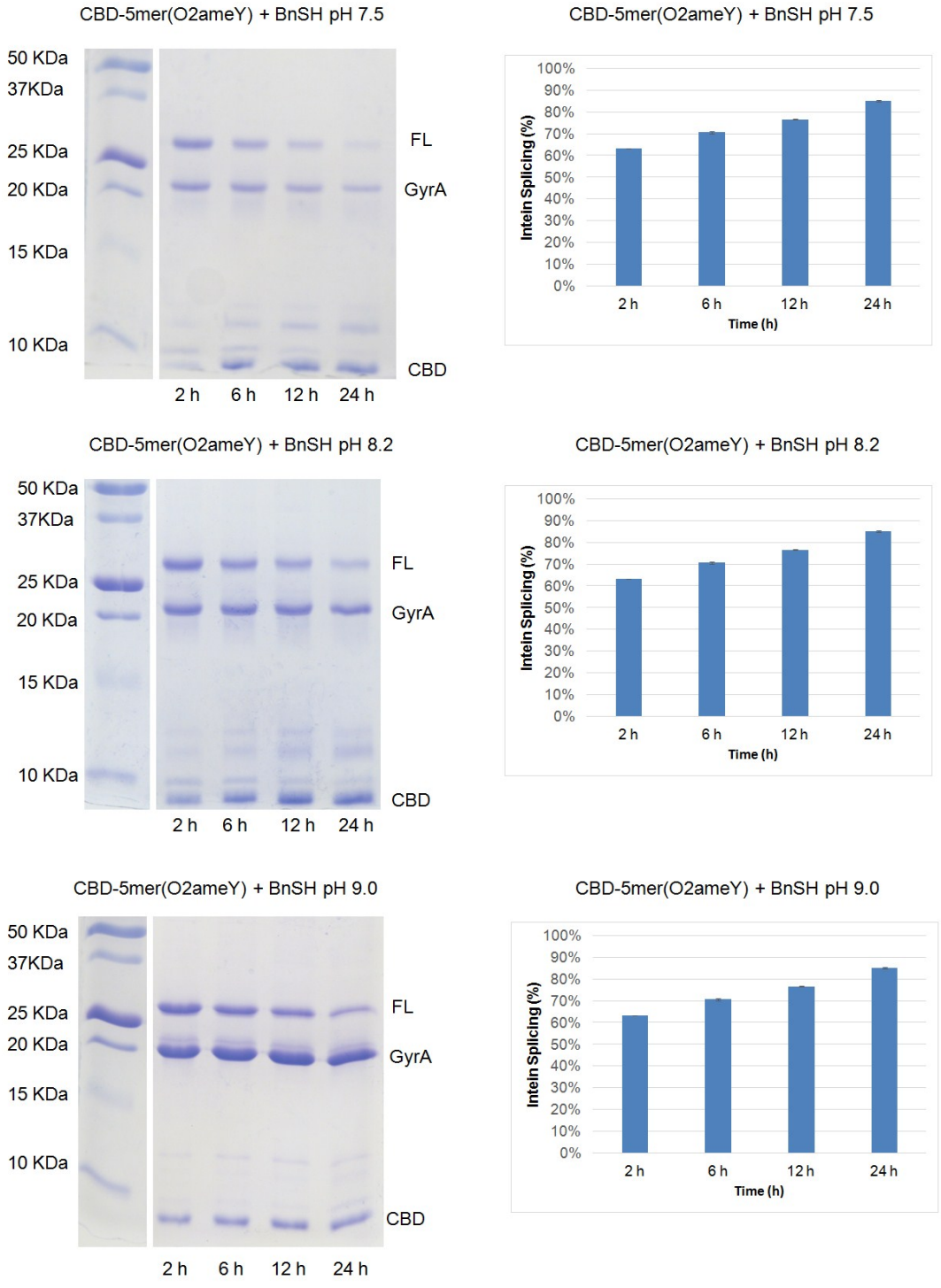


Figure S3. SDS-PAGE protein gels (left column) and densitometric analysis (graphs on right column) for thiophenol-induced cyclization reactions at pH 7.5, 8.2, and 9.0 at various time points. FL = full-length protein, GyrA = cleaved GyrA intein, CBD = CBD-fused macrocycle and acyclic product.

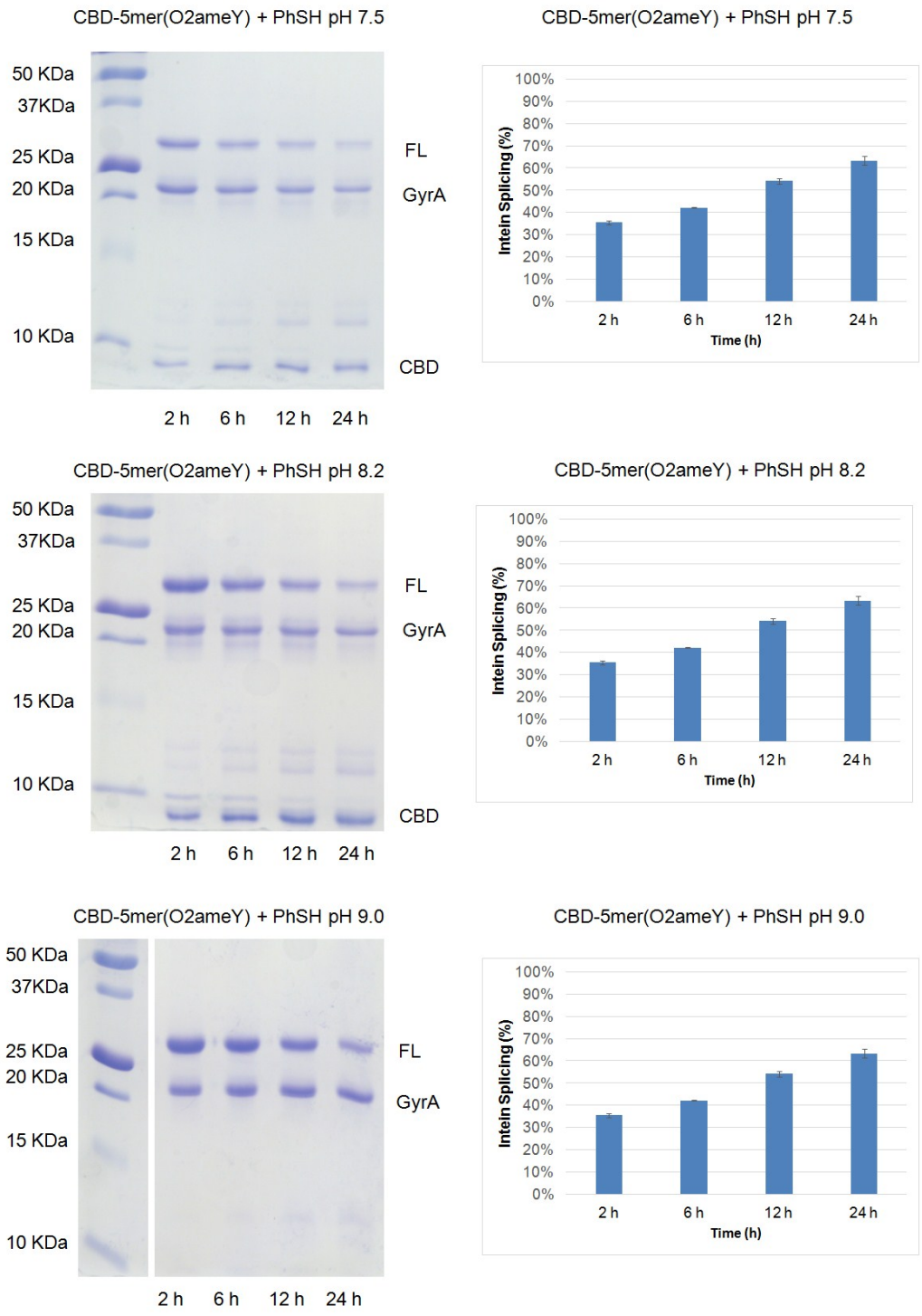


Figure S4. Representative LC-MS deconvolution spectra for large MW products from the cyclization reaction with 5mer(O2ameY) at different time points (2, 6, 12, 18, and 24 hours). FL = full-length protein, GyrA = cleaved GyrA intein.

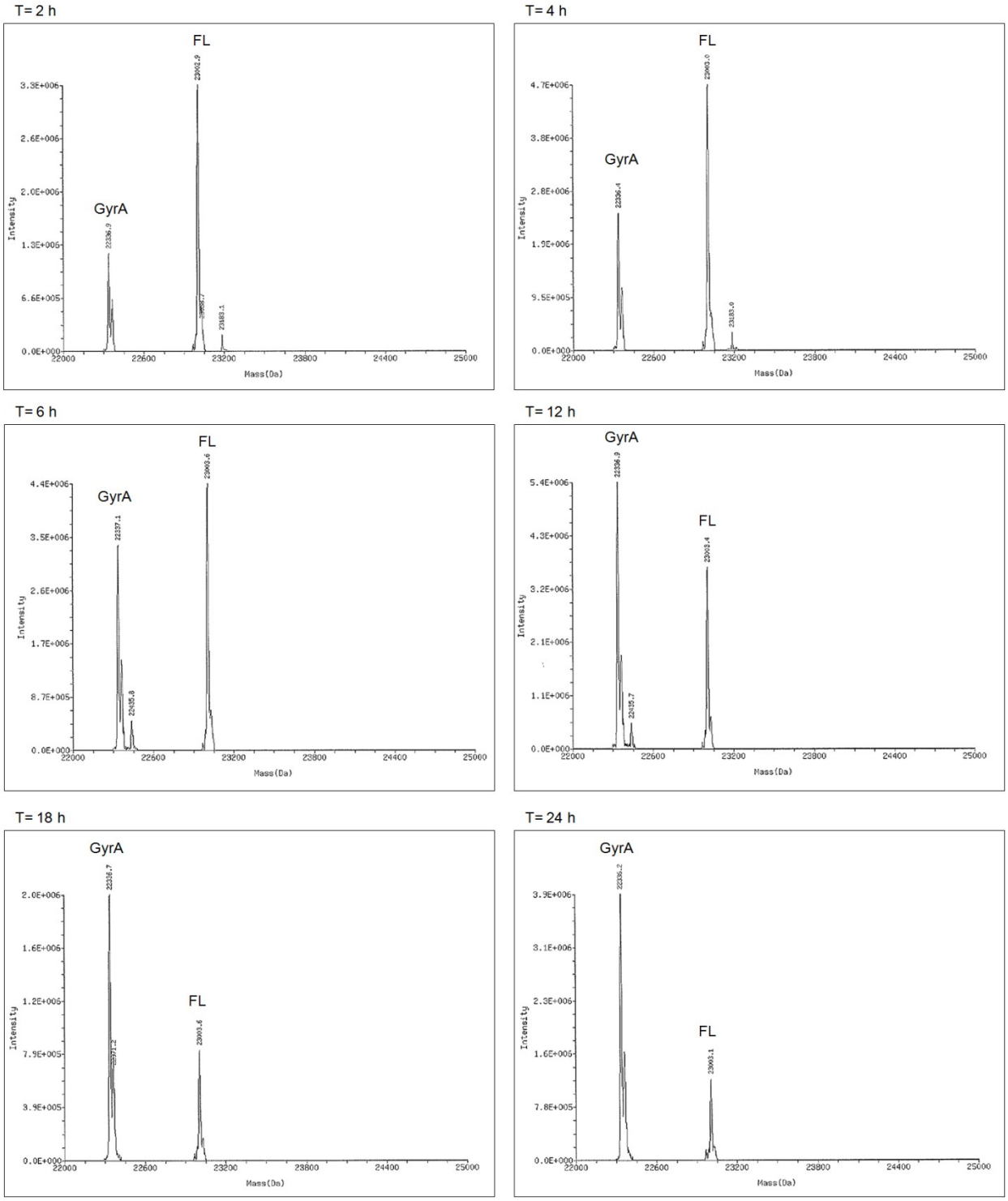


Figure S5. LC-MS ion extract chromatograms corresponding to O2ameY-containing macrocycles obtained via benzylmercaptan-induced cyclization of precursor proteins 5mer(OameY), 6mer(OameY), and 8mer(OameY), and Strep11mer(OameY).