nature portfolio

https://doi.org/10.1038/s41557-021-00733-y

Supplementary information

Site-selective tyrosine bioconjugation via photoredox catalysis for native-to-bioorthogonal protein transformation

In the format provided by the authors and unedited

Site-selective tyrosine bioconjugation via photoredox catalysis for native-to-bioorthogonal protein transformation

Daniel K. Kim[1,5], Beryl X. Li[1,5], Steven Bloom[1], Richard Y.-C. Huang[2], Jennifer X. Qiao[3], William R. Ewing[3], Daniel G. Oblinsky[4], Gregory D. Scholes[4], and David W. C. MacMillan[1]✉

[1]Merck Center for Catalysis at Princeton University, Washington Road, Princeton, New Jersey 08544, USA.

[2]Pharmaceutical Candidate Optimization, Research and Development, Bristol-Myers Squibb Company, Route 206 and Province Line Road, Princeton, New Jersey 08543, USA.

[3]Discovery Chemistry, Research and Development, Bristol-Myers Squibb Company, Route 206 and Province Line Road, Princeton, New Jersey, USA.

[4]Department of Chemistry, Princeton University, Washington Road, Princeton, New Jersey 08544, USA.

[5]These authors contributed equally to this work.

✉e-mail: dmacmill@princeton.edu

TABLE OF CONTENTS

I. General Information

Reagents. All commercial reagents were used as received without further purification. Lumiflavin was purchased from Santa Cruz Biotechnology. Bivalirudin was purchased from Apex Bio. ZVar affibody was provided by the Muir Laboratory at Princeton University. All remaining proteins investigated (human insulin, human lysozyme, ribonuclease A, cytochrome C, α-lactalbumin, myoglobin, chymotrypsinogen A, serotransferrin, carbonic anhydrase, bovine serum albumin) were purchased from Sigma Aldrich in lyophilized form. Protein structures were obtained from the Protein Data Bank (rcsb.org[1]). Milli-Q ultrapure water purified through Millipore Milli-Q Advantage A1 purification system was used for all bioconjugation reactions. Small molecule reagents and buffer salts were purchased from Sigma Aldrich and used without purification.

Photoredox light set-up. Photoredox reactions on bivalirudin were performed with Kessil PR160 LED Photoredox Lights, 440 nm blue light. Reactions on all protein substrates were performed with the Integrated photoreactor.

Protein purification. Protein reactions were purified via one of the following: Amicon Ultra 0.5 mL filters (3 kDa or 10 kDa), Thermo Slide-A-Lyzer mini dialysis units (3500 MWCO), or semi-preparatory HPLC method. Protein samples were lyophilized with a Millrock Technology MD 85 lyophilizer.

HPLC. HPLC analyses and purifications of bioconjugates were performed on an Agilent 1260 Infinity LC system using an Agilent Eclipse Plus C18 column (3.5 µm, 4.6 mm x 100 mm, part number 959961-902). Purification of small molecules beyond SiO_2 flash chromatography were performed on an ACCQPrep HP150 preparative HPLC system by Teledyne Isco (20 mm x 100 mm C18 column).

Mass spectrometry. High Resolution Mass Spectra were obtained from the Princeton University Mass Spectral Facility. For intact protein LC-MS, samples were desalted and eluted from a reversed-phase C18 Scaling column (ACQUITY UPLC BEH C18 column 300A 1.7um 2.1x150mm). For LC-MS/MS peptide mapping analysis, digested peptides were separated on a reversed-phase C18 column (ACQUITY UPLC BEH C18 column 130A 1.7um 2.1x150mm). Data was acquired on a QEPlus mass spectrometer (ThermoFisher Scientific, Waltham, MA) connected to Aquity UPLC (Waters, Milford, MA), and analyzed using Thermo BioPharma Finder software (ThermoFisher Scientific, Waltham, MA).

NMR spectroscopy. ^{1}H NMR spectra were recorded on Bruker UltraShield Plus Avance III 500 MHz. Data for ^{1}H NMR are reported as follows: chemical shift (δ ppm), multiplicity (s = singlet, d = doublet, t = triplet, m = multiplet, dd = doublet of doublets, ddd = doublet of doublet of doublets, br = broad), coupling constant (Hz), and integration.

Cyclic Voltammetry. All CV experiments were performed on CH Instruments Electrochemical Workstation (Austin, TX). Data acquisition was performed on CHI software (Ver14.05).

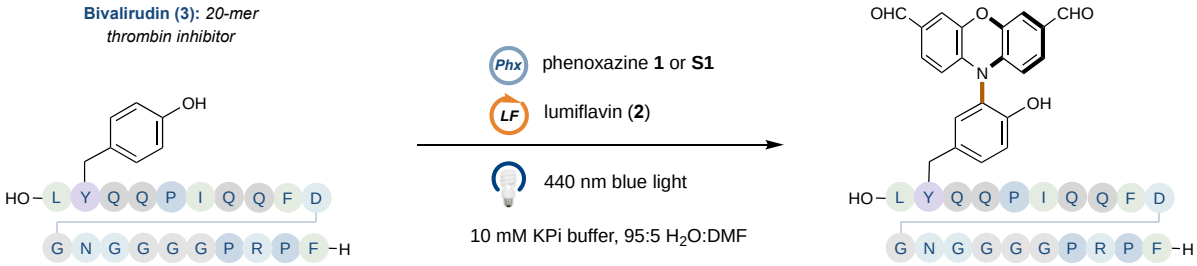
Circular Dichroism. All CD experiments were performed on Chirascan CD spectrometer from Applied Photophysics.

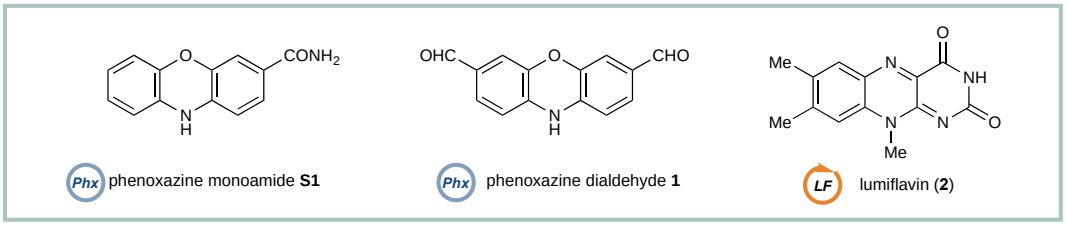
---

[1] The Protein Data Bank. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P.E. *Nucleic Acids Research*, **28**, 235–242 (2000).

UV-Vis and fluorescence spectroscopy. UV-Vis spectra were recorded on NanoDrop One Microvolume UV-Vis Spectrophotometer. UV-Vis spectra and fluorescence emission spectra were recorded on SpectraMax iD5 Multi-Mode Microplate Reader, at room temperature, using Corning 96-well Flat Clear Bottom Polystyrene TC-treated Microplates.

II. Reaction optimization with bivalirudin (3)



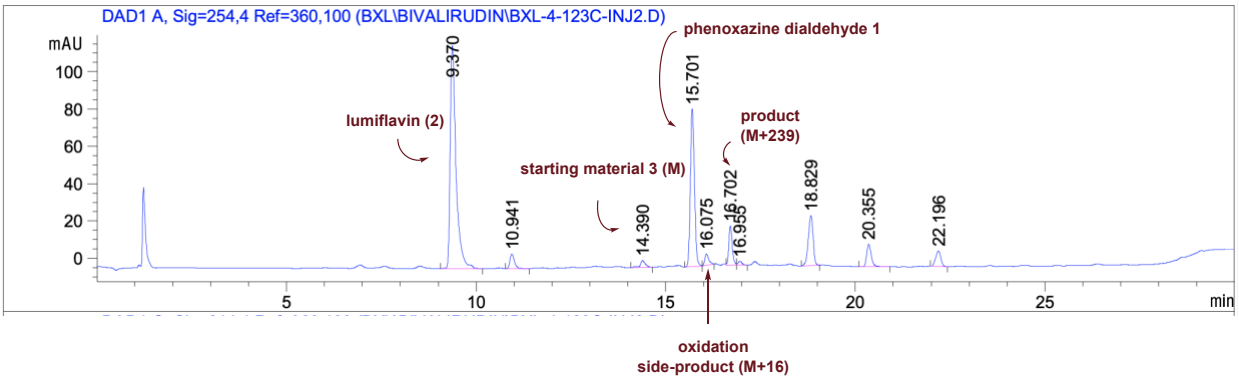


entry	phenoxazine (equiv.)	lumiflavin	reaction time	conversion
1	phenoxazine monoamide **S1** (10 equiv.)	1 equiv.	3 hours	65%
2	phenoxazine dialdehyde **1** (10 equiv.)	3 equiv.	3 hours	46%
3	phenoxazine dialdehyde **1** (10 equiv.)	3 equiv.	5 hours	74%
4	phenoxazine dialdehyde **1** (100 equiv.)	3 equiv.	5 hours	95%
5	phenoxazine dialdehyde **1** (100 equiv.)	0 equiv.	5 hours	0% (sm recovered)
6	phenoxazine dialdehyde **1** (100 equiv.)	3 equiv.	5 hours (dark)	0% (sm recovered)

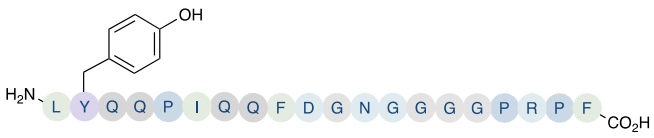
a. General reaction conditions and characterization methods

A 2-mL septum-topped screwcap HPLC via was equipped with a stir bar. Bivalirudin (**3**, 10 nmol, 10 µL of 2 mM stock solution in water, 0.1 mM final concentration), phenoxazine **1** or **S1** (0.1 or 1 µmol, 0.5 or 5 µL of 200 mM stock solution in DMF, 1 or 10 mM final concentration), lumiflavin (**2**, 30 nmol, 3 µL of 10 mM stock solution in water, 0.3 mM final concentration), potassium phosphate buffer (10 µL of 100 mM stock solution, 10 mM final concentration, pH 7) were added using micropipettes. Then, H_2O was added to adjust the total reaction volume to 100 µL and obtain the correct aforementioned final concentrations of each reagent, catalyst, and buffer. The reaction was sparged with N_2 for 15 min, then parafilmed and irradiated with a 440 nm Kessil lamp for 3 or 5 hours. The crude reaction mixture was centrifuged to remove any particulates and the supernatant was analyzed by HPLC (254 nm) to determine conversion.

b. Sample HPLC trace (entry 4)



c. LC-MS/MS results



3, Bivalirudin (2.2 kDa)
1 tyrosine
Y2 single product

LC-MS/MS analysis showed good peptide coverage (100%) and indicated the reaction was selective exclusively for Tyr 2. Predicted ***b*** and ***y*** ions are tabulated below each substrate and detected ions are bolded in blue. Raw spectra for each table can be found in section XV of this document.

N-terminus					*C-terminus*
b^+	**b^{2+}**		**Sequence**		**y^+**
–	–	1	**F**	20	–
245.1285	–	2	**P**	19	2269.9667
401.2296	201.1184	3	**R**	18	2172.9139
498.2823	249.6448	4	**P**	17	2016.8128
555.3038	278.1555	5	**G**	16	**1919.76**
612.3253	306.6663	6	**G**	15	1862.7386
669.3467	335.177	7	**G**	14	1805.7171
726.3682	363.6877	8	**G**	13	1748.6956
840.4111	420.7092	9	**N**	12	1691.6742
897.4326	449.2199	10	**G**	11	1577.6313
1012.4595	506.7334	11	**D**	10	1520.6098
1159.5279	**580.2676**	12	**F**	9	1405.5828
1288.5705	**644.7889**	13	**E**	8	1258.5144
1417.6131	709.3102	14	**E**	7	1129.4718
1530.6972	765.8522	15	**I**	6	1000.4292
1627.7499	814.3786	16	**P**	5	**887.3452**
1756.7925	878.8999	17	**E**	4	790.2924
1885.8351	943.4212	18	**E**	3	661.2498
2285.9405	1143.4739	19	Phx-**Y**	2	**532.2072**
–	–	20	**L**	1	132.1019

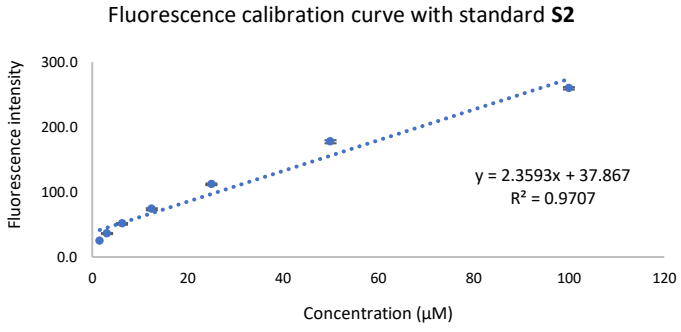
III. Protein scope and characterization

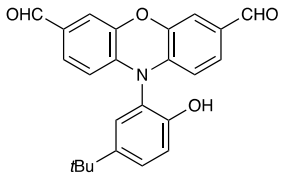
a. General protein reaction conditions

A 1.5 mL Lo-Bind Eppendorf tube was equipped with a triangular stir bar. Under an oxygen-free atmosphere, protein (20 nmol, 10 µL of 2 mM stock solution in water, 0.1 mM final concentration), phenoxazine **1** (1 µmol, 5 µL of 200 mM stock solution in DMF, 5 mM final concentration), lumiflavin (**2**, 60 nmol, 6 µL of 10 mM stock solution in water, 0.3 mM final concentration), DMF (5 µL, to achieve 95:5 water/DMF ratio in reaction), potassium phosphate buffer (20 µL of 100 mM stock solution, 10 mM final concentration, pH 7) were added using micropipettes. Then, H_2O was added to adjust the total reaction volume to 200 µL and obtain the correct aforementioned final concentrations of each reagent, catalyst, and buffer. The reaction tube was parafilmed and irradiated with a 440 nm photoreactor at 4 ºC and 5% light intensity for 30 min. The crude reaction mixture was centrifuged through a 0.22 µM spin filter to remove any particulates, then purified by a combination of Amicon Ultra spin concentrators, dialysis devices, and/or semi-preparatory HPLC method. HPLC analysis verified that no lumiflavin or phenoxazine is present in the product samples prior to determining conversion via fluorescence. Purified products were digested via trypsin and subjected to LC-MS/MS analysis to determine tyrosine selectivity.

b. Fluorescence assay calibration curve

Considering that the bioconjugation product is fluorescent, a small molecule proxy **S2** was synthesized and used to prepare a fluorescence calibration curve. Each data point in the standard curve was an average of eight runs (see error bar on graph below). Excitation wavelength is set to 430 nm, and emission is collected at 525 nm. Using this standard curve, effective concentrations of phenoxy-phenol fluorophores in each bioconjugation product were determined. Protein concentrations in the same product samples were determined by bicinchoninic acid (BCA) assay, using the commercially available Pierce™ BCA Protein Assay Kit (Thermo Scientific) and following its protocols. With the fluorophore concentration and the protein concentration, reaction conversion was able to be calculated from the ratio of the two.





Fluorescence standard **S2**

Shown below is a sample calculation for bivalirudin (**3**) reaction conversion. Note that the conversion calculated from fluorescence is in good agreement with that from HPLC analysis. For all BCA assay calibration curves, see Section XI of this document.

$$\frac{\text{Fluorescence (μM)}}{\text{Protein concentration by BCA assay (μM)}} = \frac{38.14\ \mu M}{42.12\ \mu M} = 0.91 = 91\%\ \text{conversion}$$

c. LC-MS/MS (proteomics analysis)

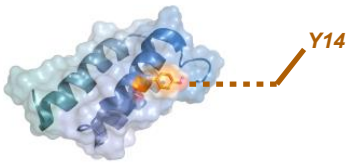
Conjugation was characterized at the intact protein level and the conjugation site was identified via peptide mapping strategy. For intact protein LC-MS, samples were prepared at 1 mg/mL in PBS and analyzed by a Waters Synapt G2 QTOF mass spectrometer connected to Aquity UPLC (Waters, Milford, MA). Samples were desalted and eluted from a reversed-phase C4 column (ACQUITY UPLC BEH C18 column 300A 1.7um 2.1x150mm).

For LC-MS/MS peptide mapping analysis, samples were denatured in the presence of 0.5 % Rapigest surfactant, reduced, alkylated and digested by trypsin and/or Glu-C. Digested peptides were separated on a reversed-phase C18 column (ACQUITY UPLC BEH C18 column 130A 1.7um 2.1x150mm). Data was acquired on a QEPlus mass spectrometer (ThermoFisher Scientific, Waltham, MA) connected to Aquity UPLC (Waters, Milford, MA), and analyzed using Thermo BioPharma Finder software (ThermoFisher Scientific, Waltham, MA). MS results were manually verified.

Predicted ***b*** and ***y*** ions are tabulated below each substrate. Observed ***b*** and ***y*** ions are highlighted in blue. Phenoxazine dialdehyde modification on tyrosine (Phx-**Y**) results in +237.0425 in mass. Carbamidomethylation of cysteine (CAM-**C**) is an artifact due to trypsin digestion during proteomics sample preparation. Raw spectra for each table can be found in section XV of this document.

IV. Protein substrates

*a. ZVar affibody (**4**)*



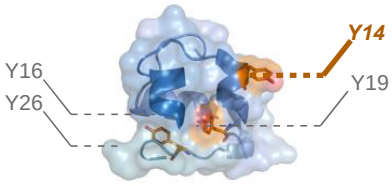
4, ZVar affibody (8.1 kDa)
1 tyrosine

ZVar affibody (**4**, PDB ID 2spz) reacted with phenoxazine dialdehyde **1** under the aforementioned general reaction conditions. The reaction mixture was irradiated for 3 hours, then purified via dialysis (MWCO 3500) to remove any small molecule reaction components. Conversion was determined by fluorescence assay to be 76%. LC-MS/MS analysis showed good peptide coverage (100%) and indicated the reaction was selective exclusively for Tyr 14. The full protein sequence is shown below, with the relevant peptide highlighted.

[1]V D A K F D K E Q Q N A F **Y** E I L H L P N L T E E Q R N A F
I Q S L K D D P S Q S A N L L A E A K K L N D A Q A P K[58]

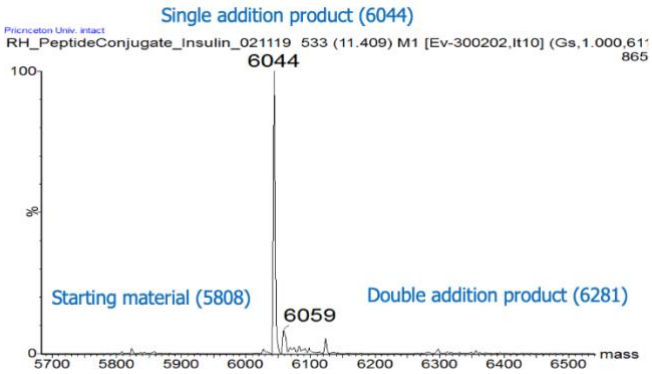
N-terminus				*C-terminus*
b$^+$		**Sequence**		**y**$^+$
–	1	**Q**	7	–
257.1244	2	**Q**	6	1008.373
371.1674	3	**N**	5	880.3142
442.2045	4	**A**	4	766.2713
589.2729	5	**F**	3	**695.2342**
989.3782	6	Phx-**Y**	2	**548.1658**
–	7	**E**	1	**148.0604**

*b. Human insulin (**5**)*



5, Human insulin (5.8 kDa)
4 tyrosines

Human insulin (**5**, PDB ID 3i40) reacted with phenoxazine dialdehyde **1** under the aforementioned general reaction conditions. The reaction mixture was irradiated for 1 hour, then purified via semi-preparatory HPLC to remove any small molecule reaction components. Conversion was determined by fluorescence assay to be 100%. It's worth to note that we observe double addition product, where both Tyr 14 and Tyr 19 were modified. We observe the double addition mass by intact mass analysis (below).



LC-MS/MS analysis showed good peptide coverage (100%) and indicated the reaction was predominantly selective for Tyr 14 (Y14/Y19 6:1). The full protein sequence is shown below, with the relevant peptide and modified tyrosine residues highlighted.

A chain

^{1}G I V E Q C C T S I C S L **Y** Q L E N **Y** C N^{21}

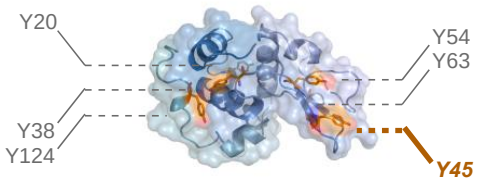
B chain

^{1}F V N Q H L C G S H L V E A L **Y** L V C G E R G F F **Y** T P K T^{30}

N-terminus					*C-terminus*
b^{+}	**b^{2+}**		**Sequence**		**y^{+}**
–	–	1	**Q**	13	–
–	**289.0965**	2	CAM-**C**	12	1770.6802
–	**449.1272**	3	CAM-**C**	11	1610.6496
532.1643	550.1748	4	**T**	10	1450.6189
619.1963	637.2069	5	**S**	9	1349.5712
732.2804	**750.2909**	6	**I**	8	1262.5392
892.311	**910.3216**	7	CAM-**C**	7	1149.4551
979.343	997.3536	8	**S**	6	989.4245
1092.4271	1110.4377	9	**L**	5	902.3925
1492.5324	1510.543	10	Phx-**Y**	4	789.3084
1620.591	1638.6016	11	**Q**	3	389.2031
1733.6751	1751.6856	12	**L**	2	**261.1445**
–	–	13	**E**	1	**148.0604**

N-terminus			*C-terminus*
b^{+}		**Sequence**	**y^{+}**
–	1	**N**	4	–
515.1555	2	Phx-**Y**	3	693.1967
675.1862	3	CAM-**C**	2	**293.0914**
–	4	**N**	1	**133.0608**

*c. Human lysozyme (**6**)*



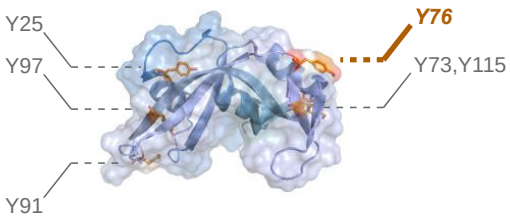
6, Human lysozyme (14.4 kDa)
6 tyrosines

Human lysozyme (**6**, PDB ID 1rex) reacted with phenoxazine dialdehyde **1** under the aforementioned general reaction conditions. The reaction mixture was irradiated for 3 hours, then purified via semi-preparatory HPLC to remove any small molecule reaction components. Conversion was determined by fluorescence assay to be 65%. LC-MS/MS analysis showed good peptide coverage (85%) and indicated the reaction was selective exclusively for Tyr 45. The full protein sequence is shown below, with the relevant peptide highlighted.

[1]K V F E R	C E L A R	T L K R L	G M D G **Y**	R G I S L	A N W M C
L A K W E	S G **Y** N T	R A T N **Y**	N A G D R	S T D **Y** G	I F Q I N
S R **Y** W C	N D G K T	P G A V N	A C H L S	C S A L L	Q D N I A
D A V A C	A K R V V	R D P Q G	I R A W V	A W R N R	C Q N R D
V R Q **Y** V	Q G C G V[130]				

N-terminus				*C-terminus*
b^+–H_2O		**Sequence**		**y^+**
–	1	**A**	9	–
155.0815	2	**T**	8	**1147.443**
269.1244	3	**N**	7	**1046.396**
669.2298	4	Phx-**Y**	6	**932.3527**
783.2727	5	**N**	5	**532.2474**
854.3098	6	**A**	4	**418.2045**
911.3313	7	**G**	3	**347.1674**
1026.3582	8	**D**	2	**290.1459**
–	9	**R**	1	**175.119**

d. Ribonuclease A (7)



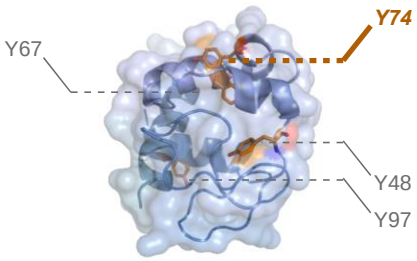
7, Ribonuclease A (13.7 kDa)
6 tyrosines

Ribonuclease A (**7**, PDB ID 1kf5) reacted with phenoxazine dialdehyde **1** under the aforementioned general reaction conditions. The reaction mixture was irradiated for 3 hours, then purified via an Amicon Ultra 0.5 mL filter (3 kDa) to remove any small molecule reaction components. Conversion was determined by fluorescence assay to be 49%. LC-MS/MS analysis showed good peptide coverage (100%) and indicated the reaction was selective exclusively for Tyr 76. The full protein sequence is shown below, with the relevant peptide highlighted.

^{1}K E T A A A K F E R Q H M D S S T S A A S S S N **Y** C N Q M M
K S R N L T K D R C K P V N T F V H E S L A D V Q A V C S Q
K N V A C K N G Q T N C **Y** Q S **Y** S T M S I T D C R E T G S S
K **Y** P N C A **Y** K T T Q A N K H I I V A C E G N P **Y** V P V H F
D A S V^{124}

N-terminus **b^{+}**		**Sequence**		*C-terminus* **y^{+}**
---	1	**N**	19	---
172.0717	2	**G**	18	2408.9211
300.1302	3	**Q**	17	2351.8996
401.1779	4	**T**	16	2223.841
515.2209	5	**N**	15	2122.7933
675.2515	6	CAM-**C**	14	2008.7504
838.3148	7	**Y**	13	1848.7198
966.3734	8	**Q**	12	1685.6564
1053.4054	9	**S**	11	**1557.5979**
1453.5108	10	Phx-**Y**	10	**1470.5658**
1540.5428	11	**S**	9	**1070.4605**
1641.5905	12	**T**	8	**983.4285**
1772.631	13	**M**	7	**882.3808**
1859.663	14	**S**	6	**751.3403**
1972.747	15	**I**	5	**664.3083**
2073.7947	16	**T**	4	**551.2242**
2188.8217	17	**D**	3	**450.1765**
2348.8523	18	CAM-**C**	2	**335.1496**
---	19	**R**	1	175.119

*e. Cytochrome C (**8**)*



8, Cytochrome C (12.4 kDa)
4 tyrosines

Cytochrome C (**8**, PDB ID 2b4z) reacted with phenoxazine dialdehyde **1** under the aforementioned general reaction conditions. The reaction mixture was irradiated for 3 hours, then purified via dialysis (MWCO 3500) to remove any small molecule reaction components. Conversion was determined by fluorescence assay to be 64%. LC-MS/MS analysis showed good peptide coverage (94%) and indicated the reaction was predominantly selective for Tyr 74 (Y74/[Y48,Y97] 20:1). The full protein sequence is shown below, with the relevant peptide highlighted.

[1]G D V E K	G K K I F	V Q K C A	Q C H T V	E K G G K	H K T G P
N L H G L	F G R K T	G Q A P G	F S **Y** T D	A N K N K	G I T W G
E E T L M	E **Y** L E N	P K K **Y** I	P G T K M	I F A G I	K K K G E
R E D L I	A **Y** L K K	A T N E[104]			

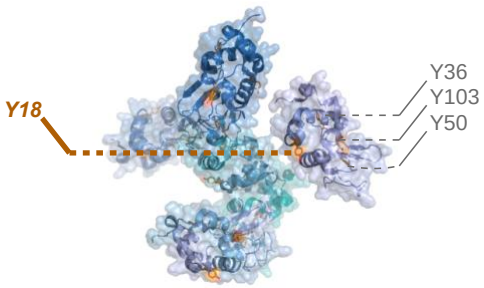
N-terminus				*C-terminus*	
b^+		**Sequence**		**y^+**	**$y^+–H_2O$**
–	1	**T**	14		
159.0764	2	**G**	13	1606.6803	1588.6697
287.135	3	**Q**	12	1549.6588	1531.6482
358.1721	4	**A**	11	**1421.6002**	1403.5897
455.2249	5	**P**	10	**1350.5631**	1332.5526
512.2463	6	**G**	9	**1253.5104**	1235.4998
659.3148	7	**F**	8	**1196.4889**	1178.4783
760.3624	8	**T**	7	**1049.4205**	1031.4099
1160.4678	9	Phx-**Y**	6	**948.3728**	930.3622
1261.5154	10	T	5	**548.2675**	530.2569
1376.5424	11	**D**	4	**447.2198**	429.2092
1447.5795	12	**A**	3	**332.1928**	
1561.6224	13	**N**	2	**261.1557**	
–	14	**K**	1	147.1128	

N-terminus				*C-terminus*
b^+		**Sequence**		**y^+**
	1	**K**	7	
529.2076	2	Phx-**Y**	6	**915.4241**
642.2916	3	**I**	5	**515.3188**
739.3444	4	**P**	4	**402.2347**

796.3659	5	**G**	3	**305.1819**
897.4135	6	**T**	2	**248.1605**
	7	**K**	1	**147.1128**

N-terminus				*C-terminus*
b⁺		**Sequence**		**y⁺**
231.0975	1	**E**	10	1486.7207
387.1987	2	**R**	9	1357.6781
516.2413	3	**E**	8	1201.577
631.2682	4	**D**	7	1072.5344
744.3523	5	**L**	6	957.5074
857.4363	6	**I**	5	844.4234
928.4734	7	**A**	4	**731.3393**
1328.5788	8	Phx-**Y**	3	**660.3022**
1441.6628	9	**L**	2	**260.1969**
	10	**K**	1	147.1128

*f. α-Lactalbumin (**9**)*



9, α-Lactalbumin (14.2 kDa, homo 6-mer)
4 tyrosines

α-Lactalbumin (**9**, PDB ID 1f6s) reacted with phenoxazine dialdehyde **1** under the aforementioned general reaction conditions. The reaction mixture was irradiated for 1 hour, then purified via semi-preparatory HPLC to remove any small molecule reaction components. Conversion was determined by fluorescence assay to be 89%. LC-MS/MS analysis showed good peptide coverage (100%) and indicated the reaction was predominantly selective for Tyr 18 (Y18/Y103 30:1). The full protein sequence is shown below, with the relevant peptide highlighted.

A chain/ B chain/ C chain/ D chain/ E chain/ F chain

[1]E Q L T K C E V F R E L K D L K G **Y** G G V S L P E W V C T T
F H T S G **Y** D T Q A I V Q N N D S T E **Y** G L F Q I N N K I W
C K D D Q N P H S S N I C N I S C D K F L D D D L T D D I M
C V K K I L D K V G I N **Y** W L A H K A L C S E K L D Q W L C
E K L[123]

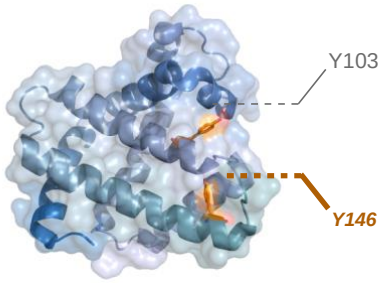
N-terminus					*C-terminus*			
b⁺	**b²⁺**	**b³⁺**	**b⁴⁺**	**Sequence**	**y⁺**	**y²⁺**	**y³⁺**	**y⁴⁺**

58.0287	29.5180	20.0144	15.2626	1	**G**	42				
458.1346	229.5709	**153.3830**	**115.2891**	2	Phx-**Y**	41	4891.1892	2446.0982	1631.0679	**1223.5528**
515.1560	**258.0817**	**172.3902**	**129.5445**	3	**G**	40	4491.0834	2246.0453	1497.6993	**1123.5263**
572.1775	286.5924	191.3974	143.7998	4	**G**	39	4434.0619	2217.5346	1478.6922	1109.2709
671.2459	**336.1266**	224.4202	168.5669	5	**V**	38	4377.0405	2189.0239	1459.6850	**1095.0156**
758.2779	379.6426	**253.4308**	**190.3249**	6	**S**	37	4277.9720	2139.4897	**1426.6622**	**1070.2485**
871.3620	436.1846	**291.1255**	218.5960	7	**L**	36	4190.9400	2095.9736	**1397.6515**	**1048.4905**
968.4148	484.7110	**323.4764**	**242.8592**	8	**P**	35	4077.8560	2039.4316	1359.9568	1020.2194
1097.4574	**549.2323**	**366.4906**	**275.1198**	9	**E**	34	3980.8032	1990.9052	**1327.6059**	**995.9563**
1283.5367	**642.2720**	**428.5171**	**321.6396**	10	**W**	33	3851.7606	1926.3839	**1284.5917**	963.6956
1382.6051	691.8062	**461.5399**	**346.4067**	11	**V**	32	3665.6813	1833.3443	**1222.5653**	**917.1758**
1542.6357	771.8215	**514.8834**	**386.4144**	12	CAM-**C**	31	3566.6129	1783.8101	1189.5425	**892.4087**
1643.6834	**822.3453**	**548.5660**	411.6763	13	**T**	30	3406.5822	1703.7948	1136.1989	**852.4010**
1744.7311	872.8692	**582.2486**	**436.9382**	14	**T**	29	3305.5345	1653.2709	**1102.5164**	**827.1391**
1891.7995	**946.4034**	**631.2714**	473.7053	15	**F**	28	3204.4869	1602.7471	**1068.8338**	**801.8772**
2028.8584	**1014.9329**	676.9577	**507.9701**	16	**H**	27	3057.4185	1529.2129	**1019.8110**	**765.1101**
2129.9061	**1065.4567**	**710.6402**	**533.2320**	17	**T**	26	2920.3595	**1460.6834**	**974.1247**	730.8453
2216.9381	**1108.9727**	**739.6509**	**554.9900**	18	**S**	25	2819.3119	**1410.1596**	**940.4421**	**705.5834**
2273.9596	1137.4834	**758.6581**	**569.2454**	19	**G**	24	2732.2798	**1366.6436**	**911.4315**	683.8254
2437.0229	1219.0151	**813.0125**	**610.0112**	20	**Y**	23	2675.2584	1338.1328	**892.4243**	669.5701
2552.0499	1276.5286	**851.3548**	638.7679	21	**D**	22	2512.1950	**1256.6012**	838.0699	628.8042
2653.0975	**1327.0524**	**885.0374**	**664.0298**	22	**T**	21	2397.1681	**1199.0877**	**799.7276**	**600.0475**
2781.1561	1391.0817	**927.7236**	696.0445	23	**Q**	20	2296.1204	**1148.5639**	**766.0450**	**574.7856**
2852.1932	**1426.6003**	**951.4026**	**713.8038**	24	**A**	19	2168.0618	1084.5346	**723.3588**	542.7709
2965.2773	1483.1423	**989.0973**	**742.0748**	25	**I**	18	2097.0247	**1049.0160**	**699.6798**	**525.0116**
3064.3457	1532.6765	**1022.1201**	766.8419	26	**V**	17	1983.9407	**992.4740**	661.9851	**496.7406**
3192.4043	1596.7058	**1064.8063**	**798.8565**	27	**Q**	16	1884.8723	**942.9398**	**628.9623**	**471.9735**
3306.4472	1653.7272	**1102.8206**	**827.3673**	28	**N**	15	1756.8137	**878.9105**	**586.2761**	**439.9589**
3420.4901	1710.7487	**1140.8349**	**855.8780**	29	**N**	14	1642.7707	**821.8890**	**548.2618**	**411.4481**
3535.5171	1768.2622	1179.1772	884.6347	30	**D**	13	**1528.7278**	764.8675	**510.2475**	**382.9374**
3622.5491	**1811.7782**	1208.1879	906.3927	31	**S**	12	**1413.7009**	**707.3541**	**471.9051**	**354.1807**
3723.5968	1862.3020	**1241.8705**	**931.6547**	32	**T**	11	**1326.6688**	**663.8381**	**442.8945**	**332.4227**
3852.6394	1926.8233	**1284.8846**	**963.9153**	33	**E**	10	**1225.6212**	**613.3142**	**409.2119**	**307.1608**
4015.7027	2008.3550	1339.2391	**1004.6811**	34	**Y**	9	**1096.5786**	548.7929	**366.1977**	**274.9001**
4072.7242	2036.8657	**1358.2462**	1018.9365	35	**G**	8	**933.5152**	467.2613	**311.8433**	**234.1343**
4185.8082	2093.4078	**1395.9409**	**1047.2075**	36	**L**	7	**876.4938**	438.7505	**292.8361**	**219.8789**
4332.8767	2166.9420	1444.9637	**1083.9746**	37	**F**	6	**763.4097**	**382.2085**	**255.1414**	191.6079
4460.9352	2230.9713	**1487.6499**	**1115.9893**	38	**Q**	5	**616.3413**	308.6743	**206.1186**	**154.8408**
4574.0193	2287.5133	**1525.3446**	**1144.2603**	39	**I**	4	**488.2827**	244.6450	163.4324	122.8261
4688.0622	2344.5348	1563.3589	**1172.7710**	40	**N**	3	**375.1987**	**188.1030**	**125.7377**	94.5551
4802.1052	2401.5562	1601.3732	1201.2817	41	**N**	2	**261.1557**	**131.0815**	87.7234	66.0444
				42	**K**	1	**147.1128**	74.0600	49.7091	37.5337

N-terminus						*C-terminus*		
b^+	**b^{2+}**	**b^{3+}**		**Sequence**		**y^+**	**y^{2+}**	**y^{3+}**
114.0913	57.5493	38.7020	1	**I**	14			
227.1754	114.0913	76.3967	2	**L**	13	1793.9009	**897.4541**	**598.6385**
342.2024	171.6048	114.7390	3	**D**	12	1680.8168	**840.9120**	**560.9438**
470.2973	235.6523	**157.4373**	4	**K**	11	1565.7899	**783.3986**	**522.6015**
569.3657	**285.1865**	190.4601	5	**V**	10	**1437.6949**	**719.3511**	479.9032
626.3872	313.6972	**209.4673**	6	**G**	9	**1338.6265**	669.8169	446.8804
739.4713	**370.2393**	247.1619	7	**I**	8	**1281.6050**	**641.3062**	427.8732
853.5142	**427.2607**	**285.1762**	8	**N**	7	**1168.5210**	584.7641	390.1785
1253.6200	**627.3136**	418.5449	9	Phx-**Y**	6	**1054.4780**	527.7427	352.1642
1439.6993	720.3533	**480.5713**	10	**W**	5	**654.3722**	327.6897	218.7956
1552.7834	776.8953	518.2660	11	**L**	4	**468.2929**	234.6501	**156.7692**
1623.8205	**812.4139**	**541.9450**	12	**A**	3	**355.2088**	178.1081	119.0745
1760.8794	880.9433	587.6313	13	**H**	2	**284.1717**	142.5895	95.3954

	14	**K**	1	**147.1128**	74.0600	49.7091

*g. Myoglobin (**10**)*



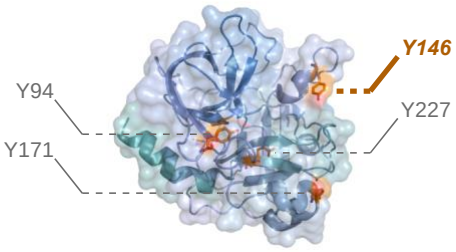
10, Myoglobin (13.7 kDa)
2 tyrosines

Myoglobin (**10**, PDB ID 1wla) reacted with phenoxazine dialdehyde **1** under the aforementioned general reaction conditions. The reaction mixture was irradiated for 1 hour, then purified via an Amicon Ultra 0.5 mL filter (3 kDa) to remove any small molecule reaction components. Conversion was determined by fluorescence assay to be 72%. LC-MS/MS analysis showed good peptide coverage (96%) and indicated the reaction was selective exclusively for Tyr 146. The full protein sequence is shown below, with the relevant peptide highlighted.

[1]G L S D G	E W Q Q V	L N V W G	K V E A D	I A G H G	Q E V L I
R L F T G	H P E T L	E K F D K	F K H L K	T E A E M	K A S E D
L K K H G	T V V L T	A L G G I	L K K K G	H H E A E	L K P L A
Q S H A T	K H K I P	I K **Y** L E	F I S D A	I I H V L	H S K H P
G D F G A	D A Q G A	M T K A L	E L F R N	D I A A K	**Y** K E L G
F Q G[153]					

N-terminus					*C-terminus*	
b^+	**b^{2+}**		**Sequence**		**y^+**	**y^{2+}**
115.05020	58.02874	1	**N**	9		
230.07715	**115.54221**	2	**D**	8	**1045.49883**	**523.25305**
343.16121	172.08424	3	**I**	7	**930.47189**	**465.73958**
414.19832	207.60280	4	**A**	6	**817.38782**	**409.19755**
485.23544	**243.12136**	5	**A**	5	**746.35071**	373.67899
613.33040	307.16884	6	**K**	4	**675.31360**	338.16044
1013.43623	507.22175	7	Phx-**Y**	2	**547.21863**	274.11295
		8	**K**	1	**147.11280**	74.06004

*h. Chymotrypsinogen A (**11**)*



11, Chymotrypsinogen A (25.7 kDa)
4 tyrosines

Chymotrypsinogen A (**11**, PDB ID 1ex3) reacted with phenoxazine dialdehyde **1** under the aforementioned general reaction conditions. The reaction mixture was irradiated for 1 hour, then purified via semi-preparatory HPLC to remove any small molecule reaction components. Conversion was determined by fluorescence assay to be 76%. LC-MS/MS analysis showed good peptide coverage (99%) and indicated the reaction was predominantly selective for Y146 (Y146/Y171 10:1). The full protein sequence is shown below, with the relevant peptide highlighted.

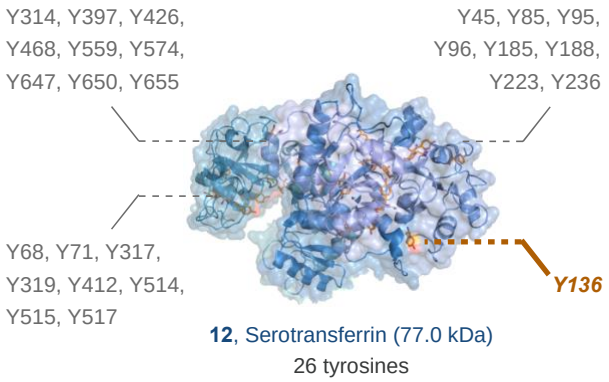
^{1}C G V P A	I Q P V L	S G L S R	I V N G E	E A V P G	S W P W Q
V S L Q D	K T G F H	F C G G S	L I N E N	W V V T A	A H C G V
T T S D V	V V A G E	F D Q G S	S S E K I	Q K L K I	A K V F K
N S K **Y** N	S L T I N	N D I T L	L K L S T	A A S F S	Q T V S A
V C L P S	A S D D F	A A G T T	C V T T G	W G L T R	**Y** T N A N
T P D R L	Q Q A S L	P L L S N	T N C K K	**Y** W G T K	I K D A M
I C A G A	S G V S S	C M G D S	G P L V C	K K N G A	W T L V G
I V S W G	S S T C S	T S T P G	V **Y** A R V	T A L V N	W V Q Q T
L A A N^{244}					

N-terminus					*C-terminus*	
b^{+}	**b^{2+}**		**Sequence**		**y^{+}**	**y^{2+}**
401.11311	**201.06019**	1	Phx-**Y**	9		
502.16078	**251.58403**	2	**T**	8	**888.41698**	**444.71213**
616.20371	**308.60549**	3	**N**	7	**787.36931**	**394.18829**
687.24082	**344.12405**	4	**A**	6	**673.32638**	**337.16683**
801.28375	**401.14551**	5	**N**	5	**602.28926**	**301.64827**
902.33143	**451.66935**	6	**T**	4	**488.24634**	**244.62681**
999.38419	**500.19574**	7	**P**	3	**387.19866**	194.10297
1114.41114	**557.70921**	8	**D**	2	**290.14590**	**145.57659**
–	–	9	**R**	1	175.11895	88.06311

N-terminus					*C-terminus*	
b^{+}	**b^{2+}**		**Sequence**		**y^{+}**	**y^{2+}**
129.10224	65.05476	1	**K**	6	–	–
529.20807	265.10767	2	Phx-**Y**	5	891.36709	446.18718
715.28738	358.14733	3	**W**	4	491.26126	246.13427

772.30884	386.65806	4	**G**	3	305.18195	153.09461
873.35652	437.18190	5	**T**	2	248.16048	124.58388
–	–	6	**K**	1	147.11280	74.06004

*i. Serotransferrin (**12**)*



12, Serotransferrin (77.0 kDa)
26 tyrosines

Serotransferrin (**10**, PDB ID 3v8x) reacted with phenoxazine dialdehyde **1** under the aforementioned general reaction conditions. The reaction mixture was irradiated for 1 hour, then purified via semi-preparatory HPLC to remove any small molecule reaction components. Conversion was determined by fluorescence assay to be 55%. LC-MS/MS analysis showed good peptide coverage (70%) and indicated the reaction was selective exclusively for Tyr 136. The full protein sequence is shown below, with the relevant peptide highlighted.

[1]MRLAV	GALLV	CAVLG	LCLAV	PDKTV	RWCAV
SEHEA	TKCQS	FRDHM	KSVIP	SDGPS	VACVK
KASYL	DCIRA	IAANE	ADAVT	LDAGL	VYDAY
LAPNN	LKPVV	AEFYG	SKEDP	QTFYY	AVAVV
KKDSG	FQMNQ	LRGKK	SCHTG	LGRSA	GWNIP
IGLLY	CDLPE	PRKPL	EKAVA	NFFSG	SCAPC
ADGTD	FPQLC	QLCPG	CGCST	LNQYF	GYSGA
FKCLK	DGAGD	VAFVK	HSTIF	ENLAN	KADRD
QYELL	CLDNT	RKPVD	EYKDC	HLAQV	PSHTV
VARSM	GGKED	LIWEL	LNQAQ	EHFGK	DKSKE[300]
[301]FQLFS	SPHGK	DLLFK	DSAHG	FLKVP	PRMDA
KMYLG	YEYVT	AIRNL	REGTC	PEAPT	DECKP
VKWCA	LSHHE	RLKCD	EWSVN	SVGKI	ECVSA
ETTED	CIAKI	MNGEA	DAMSL	DGGFV	YIAGK
CGLVP	VLAEN	YNKSD	NCEDT	PEAGY	FAIAV
VKKSA	SDLTW	DNLKG	KKSCH	TAVGR	TAGWN
IPMGL	LYNKI	NHCRF	DEFFS	EGCAP	GSKKD
SSLCK	LCMGS	GLNLC	EPNNK	EGYYG	YTGAF

RCLVE	KGDVA	FVKHQ	TVPQN	TGGKN	PDPWA
KNLNE	KDYEL	LCLDG	TRKPV	EEYAN	CHLAR
APNHA	VVTRK	DKEAC	VHKIL	RQQQH	LFGSN[600]
[601]VTDCS	GNFCL	FRSET	KDLLF	RDDTV	CLAKL
HDRNT	YEKYL	GEEYV	KAVGN	LRKCS	TSSLL
EACTF	RRP[668]				

N-terminus						*C-terminus*		
b^{+}	**b^{2+}**	**b^{3+}**		**Sequence**		**y^{+}**	**y^{2+}**	**y^{3+}**
88.03930	44.52329	30.01795	1	**S**	19			
159.07642	80.04185	53.69699	2	**A**	18	2321.10586	**1161.05657**	774.37347
216.09788	108.55258	72.70415	3	**G**	17	2250.06874	**1125.53801**	750.69443
402.17720	**201.59224**	134.73058	4	**W**	16	2193.04728	**1097.02728**	**731.68728**
516.22012	258.61370	**172.74489**	5	**N**	15	2006.96797	**1003.98762**	**669.66084**
629.30419	**315.15573**	**210.43958**	6	**I**	14	1892.92504	946.96616	**631.64653**
726.35695	**363.68211**	**242.79050**	7	**P**	13	1779.84098	**890.42413**	**593.95184**
839.44101	**420.22415**	**280.48519**	8	**I**	12	1682.78821	**841.89774**	**561.60092**
896.46248	**448.73488**	**299.49234**	9	**G**	11	1569.70415	**785.35571**	**523.90623**
1009.54654	**505.27691**	**337.18703**	10	**L**	10	1512.68268	756.84498	504.89908
1122.63061	**561.81894**	**374.88172**	11	**L**	9	**1399.59862**	700.30295	**467.20439**
1522.73643	761.87186	508.25033	12	Phx-**Y**	8	**1286.51456**	643.76092	**429.50970**
1682.76708	**841.88718**	**561.59388**	13	CAM-**C**	7	**886.40873**	443.70800	**296.14109**
1797.79403	899.40065	599.93619	14	**D**	6	**726.37808**	**363.69268**	**242.79754**
1910.87809	955.94268	**637.63088**	15	**L**	5	**611.35114**	**306.17921**	204.45523
2007.93085	**1004.46907**	**669.98180**	16	**P**	4	**498.26707**	**249.63717**	**166.76054**
2136.97345	1068.99036	**712.99600**	17	**E**	3	**401.21431**	**201.11079**	134.40962
2234.02621	1117.51674	**745.34692**	18	**P**	2	**272.17172**	136.58950	91.39542
			19	**R**	1	**175.11895**	88.06311	59.04450

V. Human lysozyme (6) derivatives via aminooxy- or hydrazide- condensation

a. General Procedure for human lysozyme derivatization through aldehyde-click condensations

A 1.5 mL Lo-Bind Eppendorf tube was equipped with a triangular stir bar. Under an oxygen-free atmosphere, lysozyme (20 nmol, 10 µL of 2 mM stock solution in water, 0.1 mM final concentration), phenoxazine **1** (1 µmol, 5 µL of 200 mM stock solution in DMF, 5 mM final concentration), lumiflavin (**2**, 60 nmol, 6 µL of 10 mM stock solution in water, 0.3 mM final concentration), DMF (5 µL, to achieve 95:5 water/DMF ratio in reaction), potassium phosphate buffer (20 µL of 100 mM stock solution, 10 mM final concentration, pH 7) were added using micropipettes. Then, H_2O was added to adjust the total reaction volume to 200 µL and obtain the correct aforementioned final concentrations of each reagent, catalyst, and buffer. The reaction tube was parafilmed and irradiated with a 440 nm photoreactor at 4 ºC and 5% light intensity for 30 min. The crude reaction mixture was centrifuged through a 0.22 µM spin filter to remove any particulates. Aminooxy- or hydrazide-linked biofunctionalized handles were prepared as 200 mM stock solutions in DMSO. 5-Methoxyanthranilic acid was also prepared as a 200 mM stock solution in DMSO. Both stock solutions were added to the lysozyme crude product mixture to reach 10 mM final concentration, respectively (10% DMSO in total). The reaction was incubated at 4 °C for overnight.

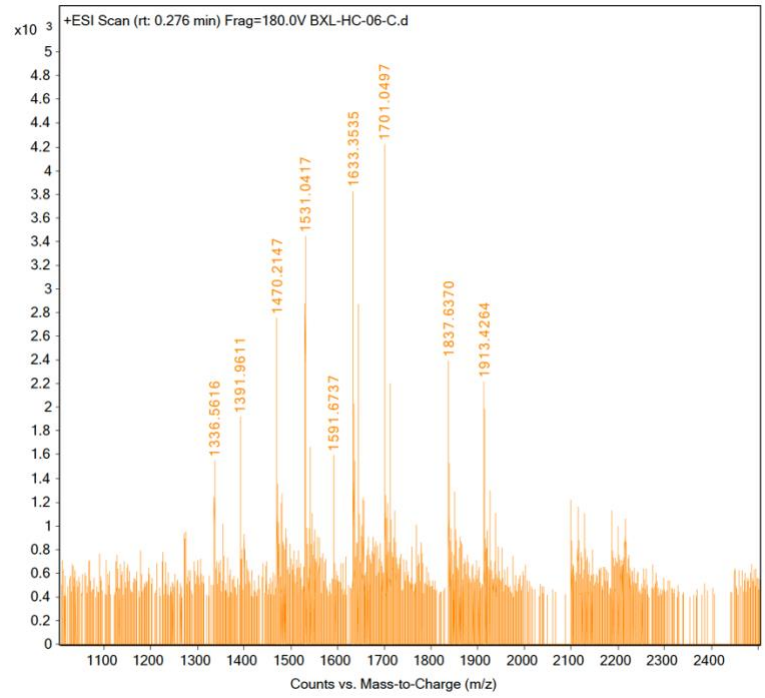
All samples were concentrated and prepared directly from protein bioconjugation reactions to 1 mg/mL prior to intact mass analysis using ThermoScientific Pierce™ Concentrator, PES, 10K MWCO, 0.5mL concentrator device. Proteins were washed and concentrated with Milli-Q water containing 0.1% formic acid (3 × 400 µL) to remove excess salt and buffer from the protein solution.

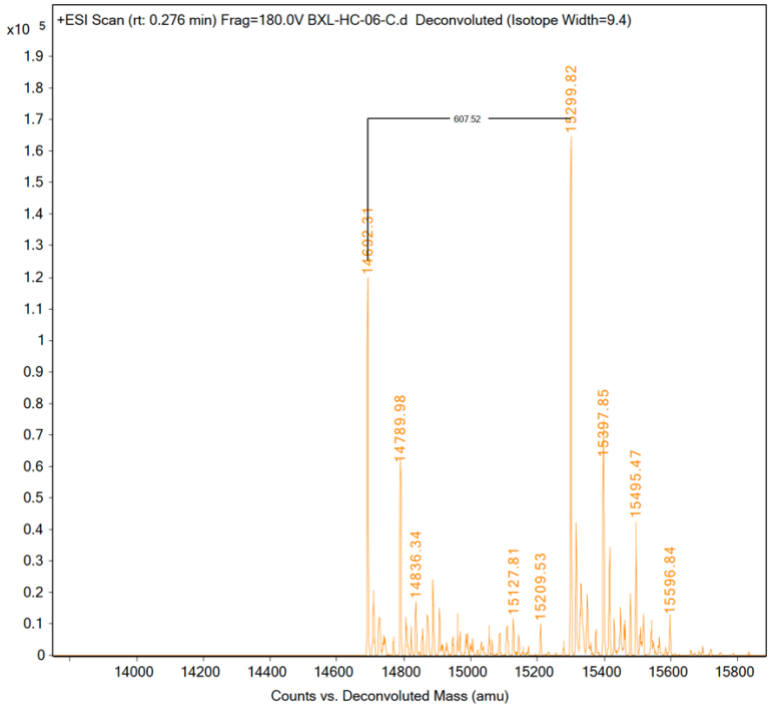
b. Intact mass analysis of alkyne, azide, biotin, and Cy 7.5 derivatives

Lysozyme + dCHOphenoxazine + aminooxy-PEG3-alkyne

native lysozyme starting material mass: 14692 Da

bioconjugated lysozyme + modification mass (+2 mod): 15299 Da (14692 + 237 + 185 + 185)

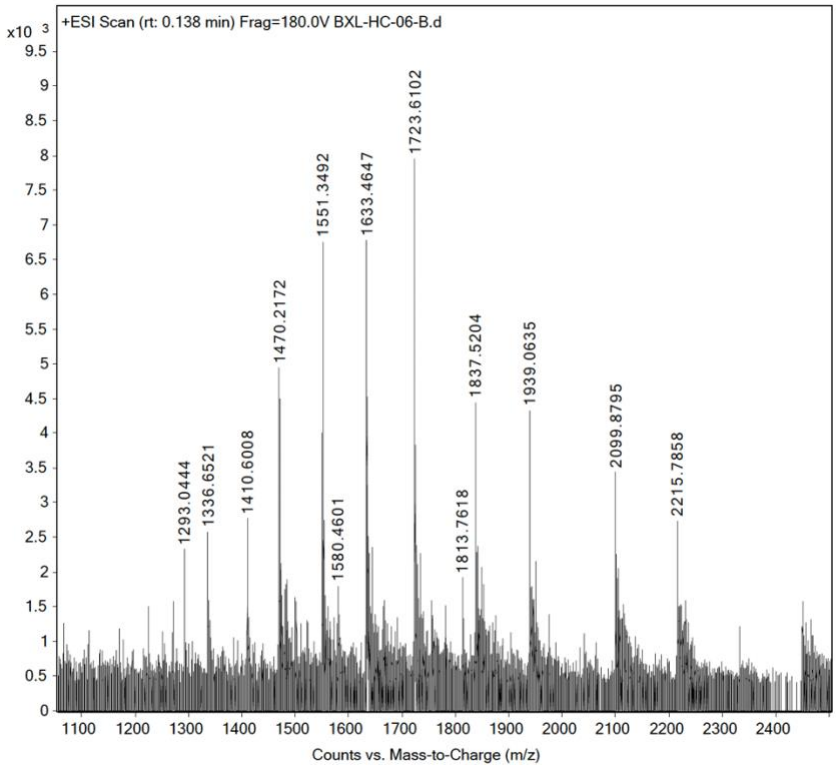


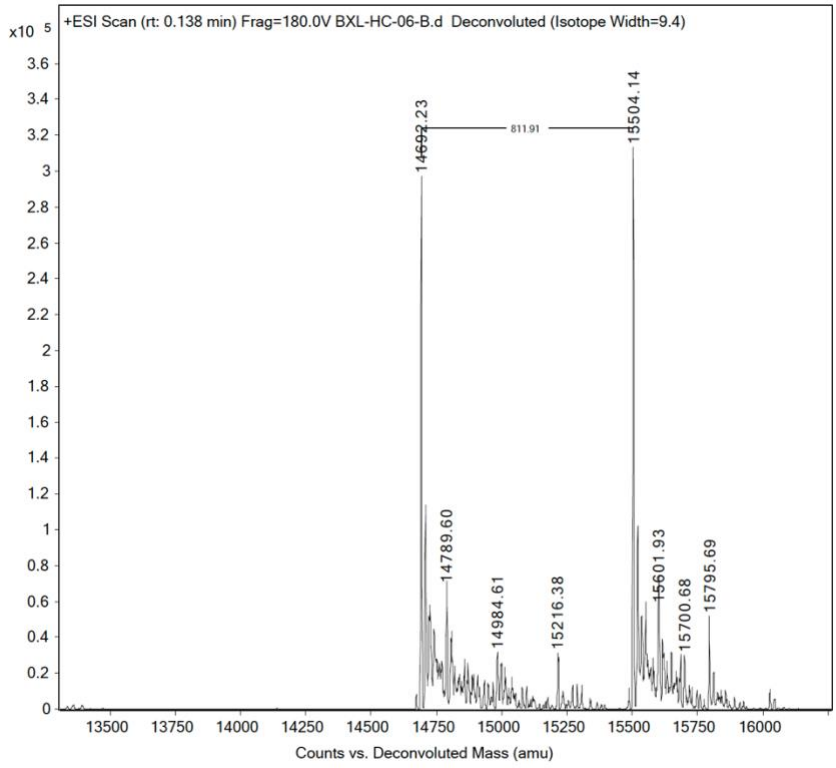


Lysozyme + dCHOphenoxazine + hydrazide-PEG4-azide

native lysozyme starting material mass: 14692 Da

bioconjugated lysozyme + modification mass (+2 mod): 15503 Da (14692 + 237 + 287 + 287)

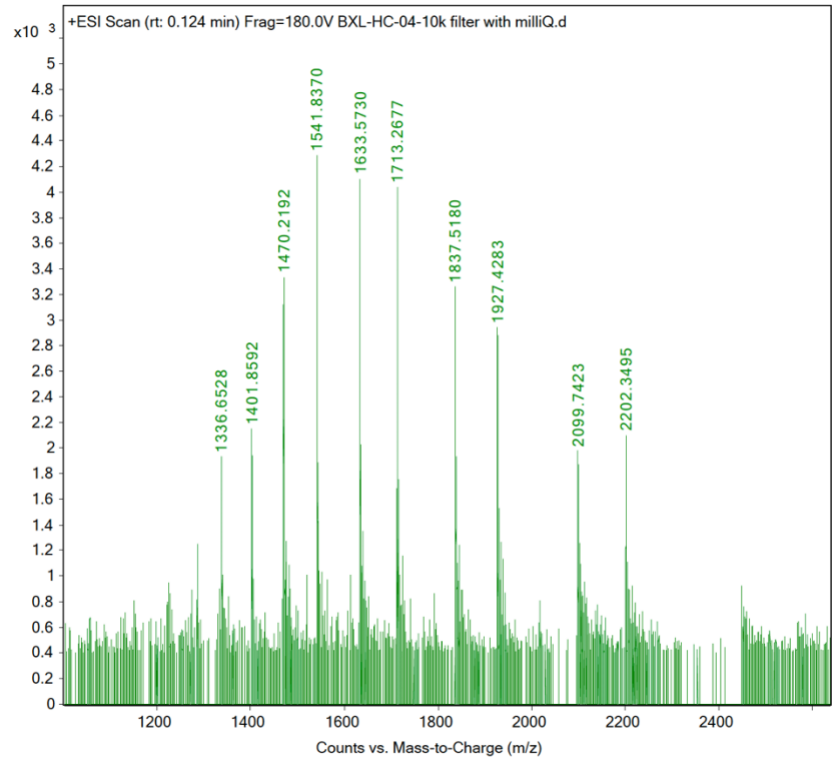


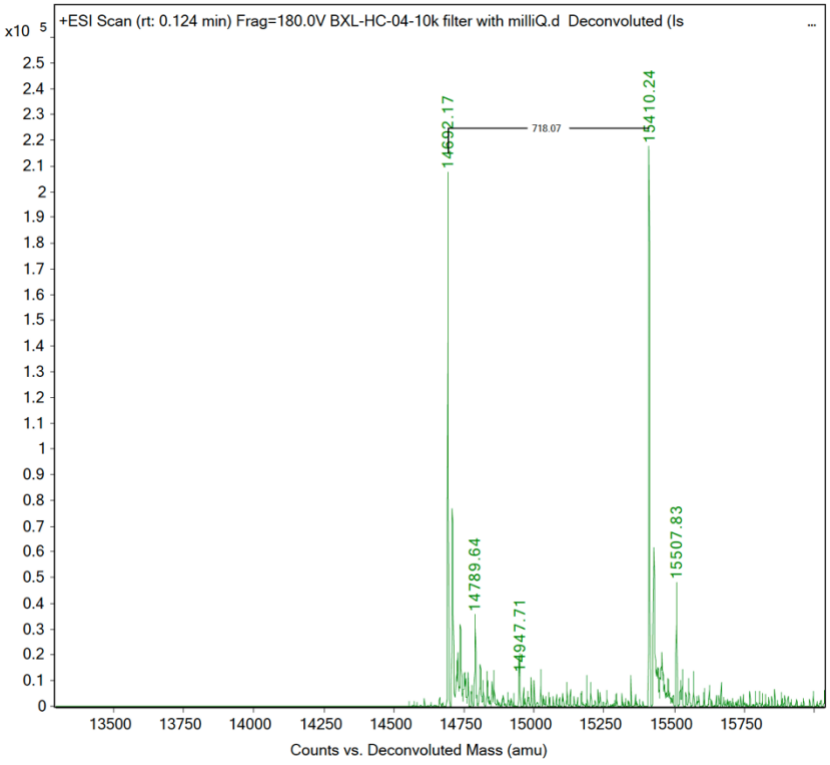


Lysozyme + dCHOphenoxazine + hydrazide-biotin

native lysozyme starting material mass: 14692 Da

bioconjugated lysozyme + modification mass (+2 mod): 15411 Da (14692 + 237 + 241 + 241)

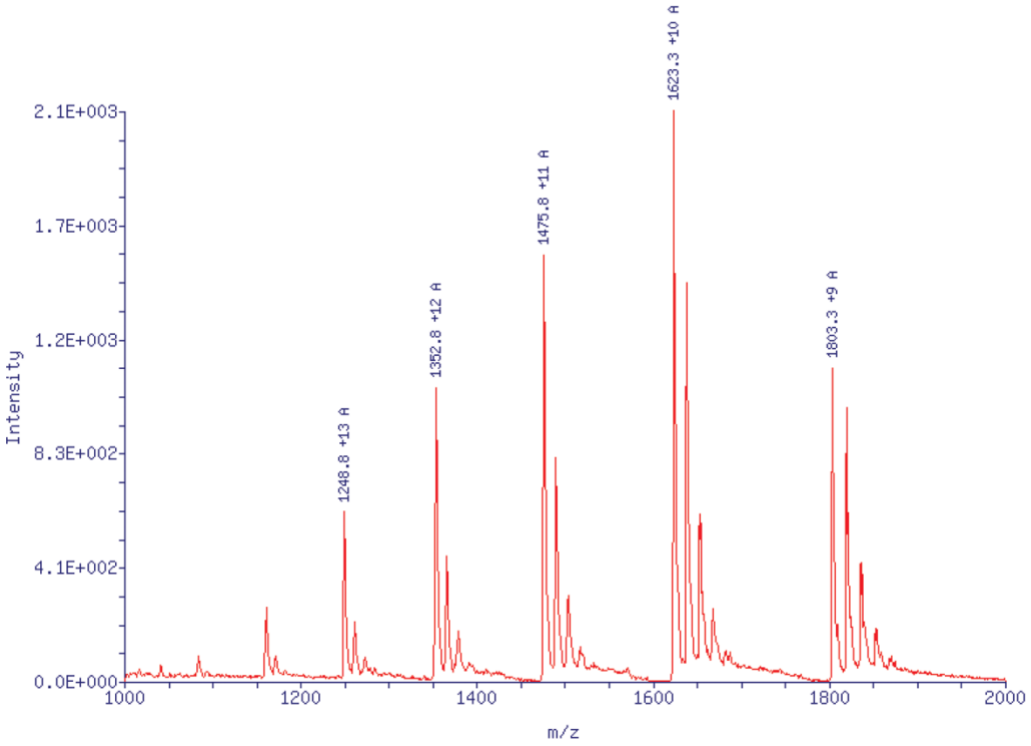


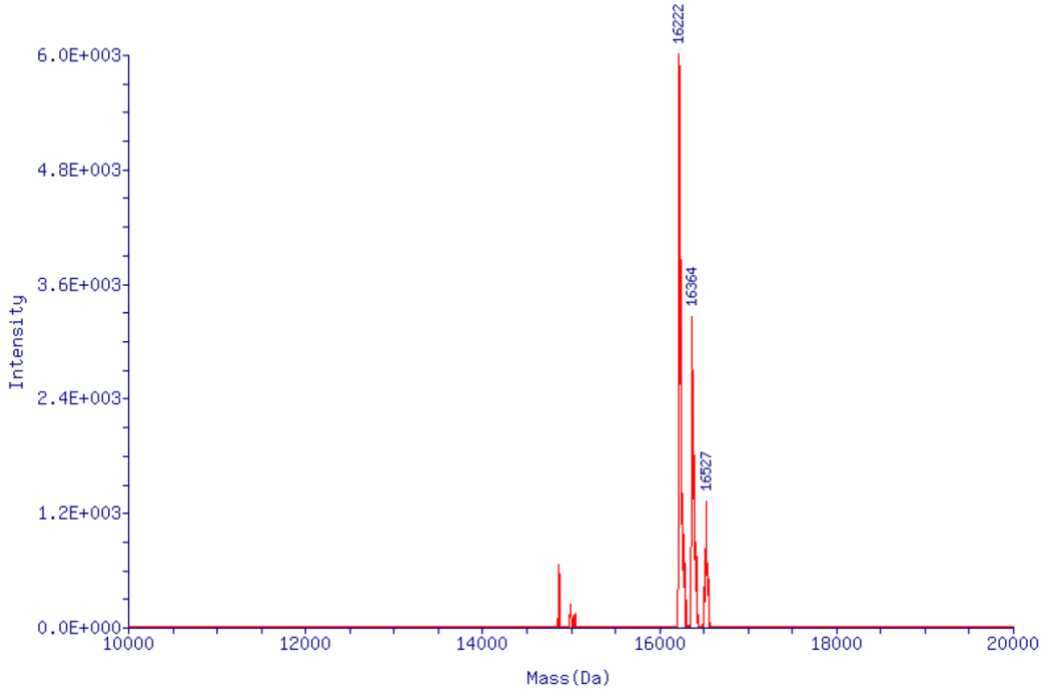


Lysozyme + dCHOphenoxazine + hydrazide Cy 7.5 fluorophore

native lysozyme starting material mass: 14692 Da

bioconjugated lysozyme + modification mass (+2 mod): 16221 Da (14692 + 237 + 646 + 646)





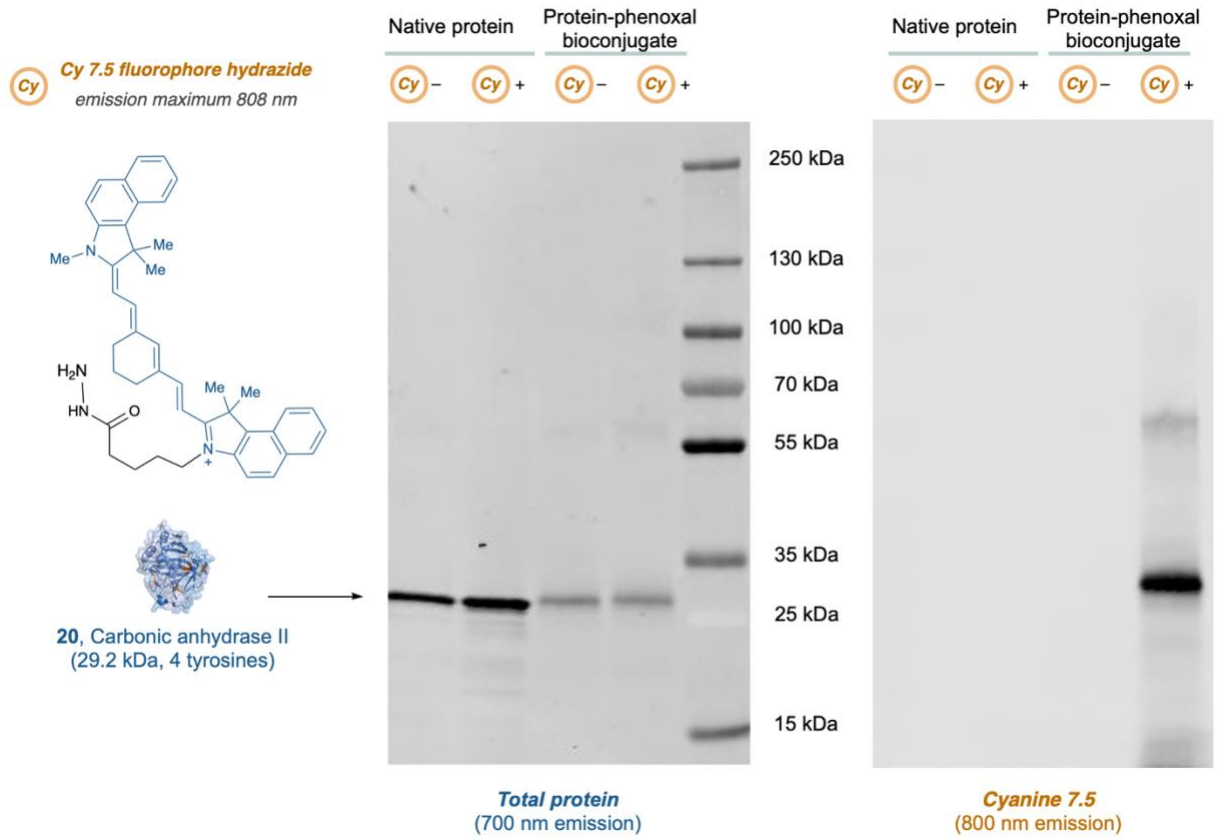
VI. SDS-PAGE of carbonic anhydrase (20) and bovine serum albumin (BSA) (21)

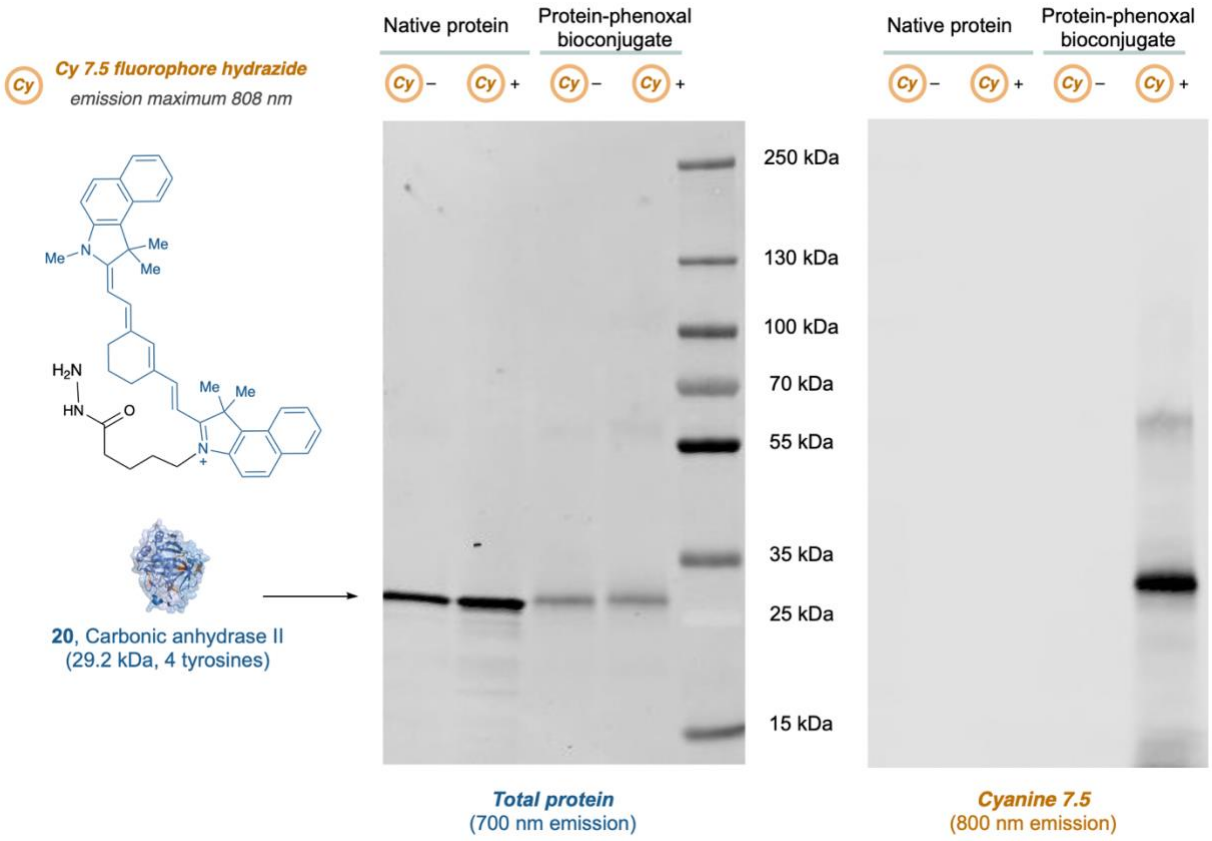
a. Reaction conditions

Carbonic anhydrase II (**20**) was bioconjugated with phenoxazine dCHO **1** under general reaction conditions reported in Section III. To the protein mixture was added an excess of Cy 7.5 hydrazide (Lumiprobe, Catalog # 16070) of 50–100 equivalence. The resulting mixture was incubated in the dark at 37 ºC for 1 hour. The excess small molecule reagents was removed from the protein bioconjugate with Amicon Ultra spin filters (3 kDa size, 3 x 10 min centrifugation at 10 kDa) before loaded for gel electrophoresis analysis. BSA (**21**) was subjected to the same reaction conditions.

For protein analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was accomplished on a Criterion™ Cell midi electrophoresis apparatus (Bio-Rad, USA). Commercially available markers PageRuler™ Prestained Protein Ladder (Bio-Rad, USA) were applied to at least one lane of each gel for calculation of apparent molecular weights. Fluorescence visualization of gels was obtained by UV transillumination at 800 nm. Visualization of protein bands was accomplished by staining with REVERT™ Total Protein Stain (Li-COR, USA). Gel imaging was performed on an Li-COR Odyssey CLx system (USA). Protein recovery was estimated from standard optical density measurements of the observed gel bands with Image Studio software (version 5.2, Li-COR).

b. SDS-PAGE

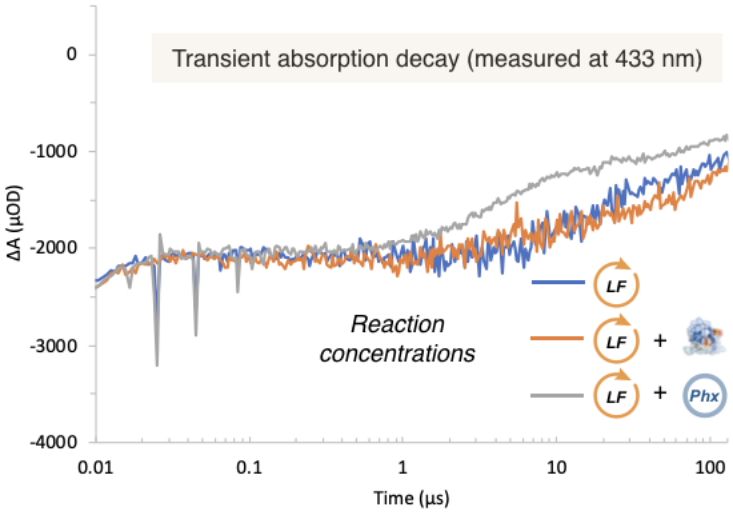



Cy 7.5 fluorophore hydrazide
emission maximum 808 nm
Me
Me
Me
N
H2N
HN
O
Me
Me
N+
20, Carbonic anhydrase II
(29.2 kDa, 4 tyrosines)
Native protein
Protein-phenoxal bioconjugate
Cy –
Cy +
Cy –
Cy +
250 kDa
130 kDa
100 kDa
70 kDa
55 kDa
35 kDa
25 kDa
15 kDa
Total protein
(700 nm emission)
Native protein
Protein-phenoxal bioconjugate
Cy –
Cy +
Cy –
Cy +
Cyanine 7.5
(800 nm emission)

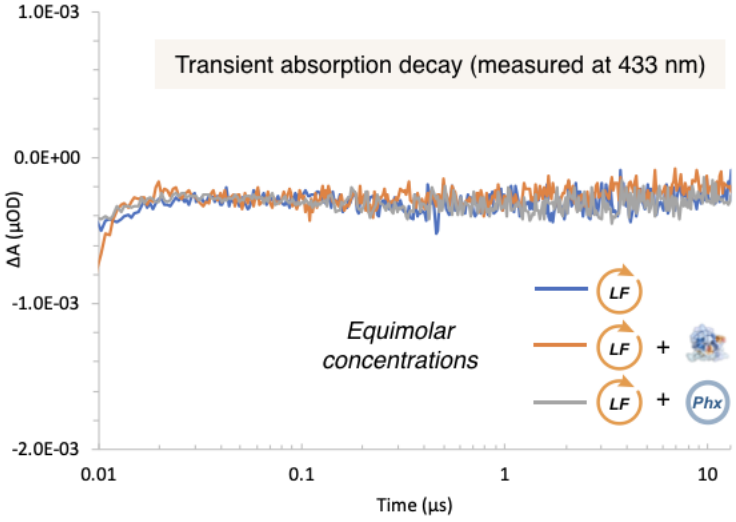
VII. Transient absorption spectroscopy

The long-time transient absorption spectrometer (EOS) is a commercial system built by Ultrafast Systems. The setup consists a Coherent Libra Ti:Sapphire (800 nm:45 fs 1 kHz) to seed a Coherent optical parametric amplifier to generate pulses from 300 nm to 1600 nm (~20 nm full-width half maximum) which are subsequently used to excite the sample. The probe pulse is from a Leukos Disco laser. This is accomplished by pumping a photonic crystal fiber with 1064 nm photons which generates a 350–1600 nm broad band pulse (800 ps, 2 kHz). The two lasers are synchronized via a TTL pulse and a delay is introduced through the electronic triggering system.

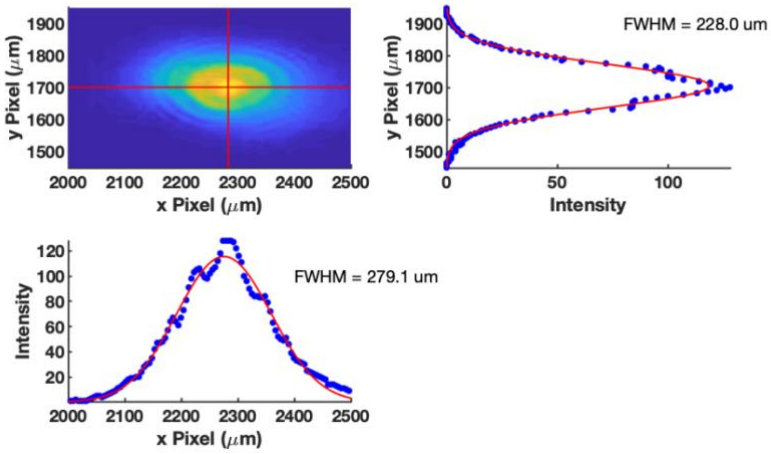
All samples were excited with a 370 nm pulse with an energy of 300 nJ yielding a total photon flux of 2.77E14 photons/cm^2. Samples were continuously stirred in a 2 mm cuvette for the duration of the experiment. Molecular oxygen from each solution was removed by undergoing three freeze-pump-thaw cycles. Sample preparation was performed in an inert atmosphere (MBraun Labmaster Pro; O2 < 0.1 ppm). The concentration for each species was kept constant for each transient absorption experiment. They were as follows: 1 mM lumiflavin (**2**), 50 mM phenoxazine dialdehyde **1**, 5 mg/mL (~1 mM) insulin (**5**).



To elucidate the effects of concentration dependent quenching on the dynamics we modified all concentrations of lumiflavin, phenoxazine, and insulin to be equimolar (30 µM). Solutions of equimolar lumiflavin, lumiflavin + insulin, and lumiflavin + phenoxazine where studied under the same spectroscopic conditions as above. It can be noted that at these concentrations no observable quenching of the flavin bleach is observed when overlaying the kinetics at 433nm. Therefore, an increase in the concentration of the phenoxazine over that of insulin allows for quenching of the species in higher concentration.



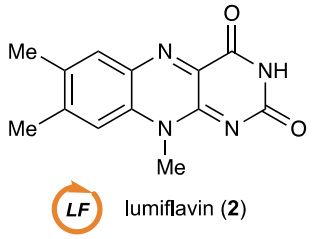
Pump spot size as visualized on CCD camera (Thorlabs DCU224) and fit to gaussian in x and y coordinate (see figure below).

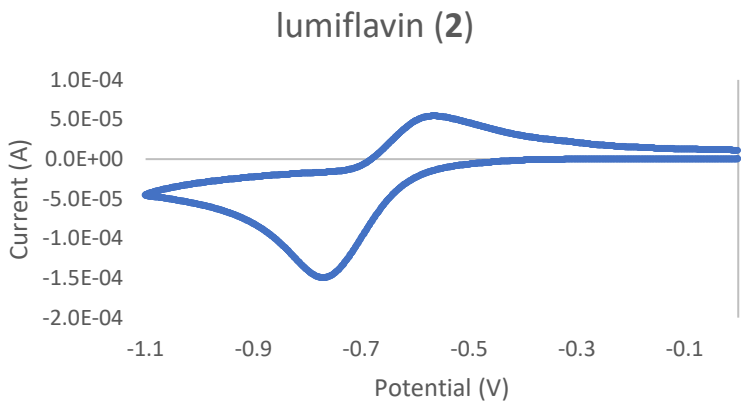


VIII. Cyclic voltammetry data

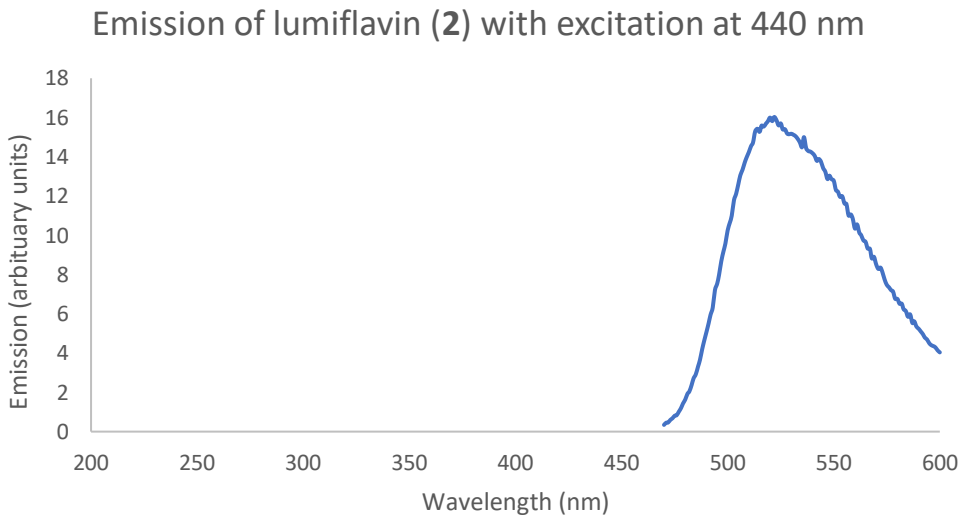
*a. Lumiflavin (**2**)*

In a three-neck round-bottom flask a solution of **2** (2.6 mg, 0.01 mmol, 1 mM) and KCl (electrolyte, 74.5 mg, 1.00 mmol, 100 mM) and diluted up with 9:1 Milli-Q water:DMF (10 mL total volume). Voltammograms recorded at υ = 200 mV/s with a 3 mm glassy carbon working electrode, a 3 mm SCE reference electrode, and a platinum wire counter electrode under nitrogen. $E_{1/2}^{red}$ [**2**/**2**$^{\bullet-}$] = – 0.66 V.



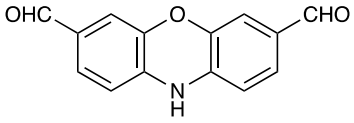


Excited state oxidation potential of lumiflavin (**2**) was calculated using the Rehm-Weller formalism. Em_{max} of lumiflavin was measured to be 525 nm. $E_{1/2}^{red}$ [^{3}LF/LF$^{\bullet-}$] = + 1.68 V

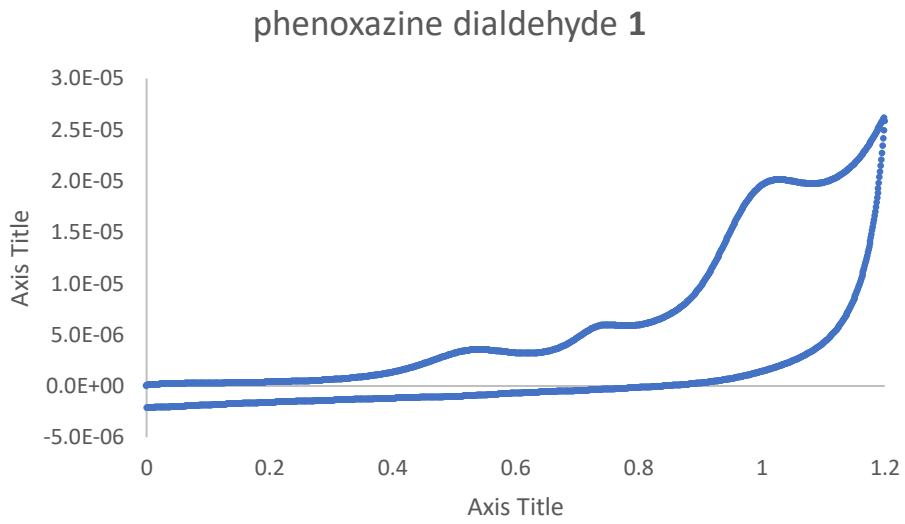


*b. Phenoxazine dialdehyde **1***

In a three-neck round-bottom flask a solution of **1** (4.8 mg, 0.02 mmol, 2 mM) and KCl (electrolyte, 74.5 mg, 1.00 mmol, 100 mM) and diluted up with milli-Q water (10 mL total volume). Voltammograms recorded at υ = 200 mV/s with a 3 mm glassy carbon working electrode, a 3 mm SCE reference electrode, and a platinum wire counter electrode under air. E_{pa} [**1**/**1**$^{\bullet+}$] = + 0.53 V.

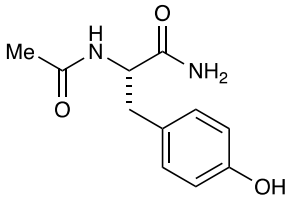


phenoxazine bisaldehyde **1**

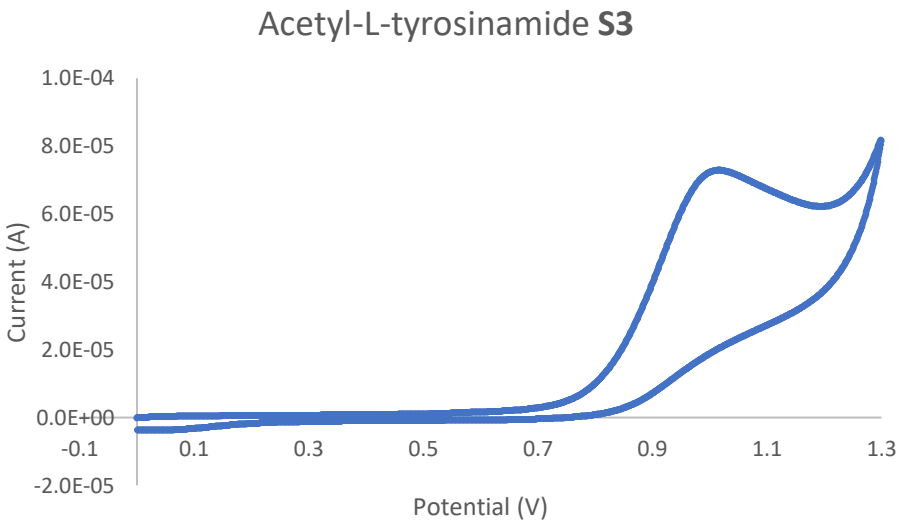


*c. Acetyl-L-tyrosinamide **S3***

In a three-neck round-bottom flask a solution of **S3** (6.6 mg, 0.03 mmol, 3 mM) and KCl (electrolyte, 74.5 mg, 1.00 mmol, 100 mM) and diluted up with milli-Q water (10 mL total volume). Voltammograms recorded at υ = 200 mV/s with a 3 mm glassy carbon working electrode, a 3 mm SCE reference electrode, and a platinum wire counter electrode under air. E_{pa} [**S3**/**S3**$^{\bullet+}$] = + 1.01 V.

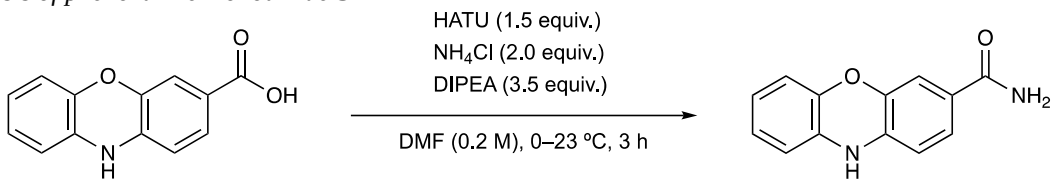


Acetyl-L-tyrosinamide **S3**



IX. Synthesis of phenoxazine derivatives

*a. Synthesis of phenoxazine monoamide **S1***

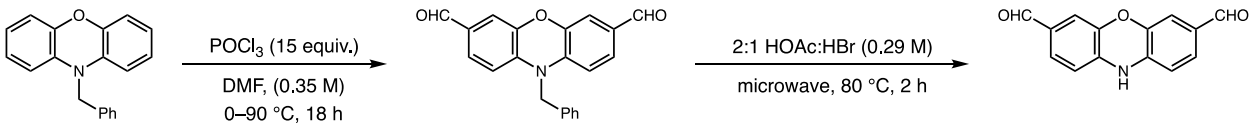


S1

10*H*-phenoxazine-3-carboxamide (S1). Starting material 10*H*-phenoxazine-3-carboxylic acid was synthesized following literature procedure[2]. To a 2-dram (8 mL) vial was added 10*H*-phenoxazine-3-carboxylic acid (100 mg, 0.44 mmol, 1 equiv.) in DMF (2.2 mL, 0.2 M) and was cooled at 0 °C in an ice-water bath. To the solution was added HATU (155 mg, 0.66 mmol, 1.5 equiv.) followedL by NH_4Cl (47.1 mg, 0.88 mmol, 2 equiv.). Finally, DIPEA was added dropwise (268 µL, 1.54 mmol, 3.5 equiv.). The reaction was allowed to warm to room temperature (23 °C) and was stirred for 3 hours. Upon completion, the reaction was quenched with water and was extracted with EtOAc (3 × 20 mL). The combined organic layer was washed with a saturated aqueous LiCl solution (1 × 25 mL). The organic layer was dried over magnesium sulfate, filtered, and concentrated under reduced pressure. The crude reaction mixture was purified by flash column chromatography using 15–40% ethyl acetate in hexanes to yield a beige powder in 98% yield (98 mg, 0.431 mmol).

^{1}H NMR (500 MHz; DMSO-d_6) δ 8.57 (s, 1H), 7.65 (s, 1H), 7.29 (dd, J = 8.1, 1.9 Hz, 1H), 7.11 (d, J = 1.9 Hz, 1H), 7.08 (s, 1H), 6.74 (ddd, J = 7.9, 6.2, 2.6 Hz, 1H), 6.66–6.55 (m, 2H), 6.47 (dd, J = 8.1, 1.9 Hz, 1H), 6.43 (d, J = 8.1 Hz, 1H). ^{13}C NMR (126 MHz; DMSO-d_6) δ 166.8, 142.6, 142.1, 135.3, 131.4, 126.1, 124.3, 124.1, 121.1, 115.2, 114.1, 113.6, 112.3. HRMS (ESI-TOF): *m/z* calculated for $C_{13}H_{11}N_2O_2^+$ $[M+H]^+$ 227.0815, found 227.0767.

*b. Synthesis of phenoxazine dialdehyde **2***



S4 **1**

10-benzyl-10*H*-phenoxazine-3,7-dicarbaldehyde (S4). Starting material 10-benzyl-10*H*-phenoxazine was synthesized following literature procedure[3]. 10-benzyl-10*H*-phenoxazine (2.73 g, 10 mmol) was dissolved in DMF

[2] Moritomo, A.; Yamada, H.; Watanabe, T.; Itahana, H.; Akuzawa, S.; Okada, M. & Ohta, M. Synthesis and structure–activity relationships of new carbonyl guanidine derivatives as novel dual 5-HT2B and 5-HT7 receptor antagonists. *Bioorg. Med. Chem.* **21**, 7841–7852 (2013).

[3] Odyniec, M. L.; Sedgwick, A. C.; Swan, A. H.; Weber, M.; Tang, T. M. S.; Gardiner, J. E.; Zhang, M.; Jiang, Y.-B.; Kociok-Kohn, G.; Elmes, R. B. P.; Bull, S. D.; He, X.-P. & James, T. D. 'AND'-based fluorescence scaffold for the detection of ROS/RNS and a second analyte. *Chem. Commun.* **54**, 8466–8469, (2018).

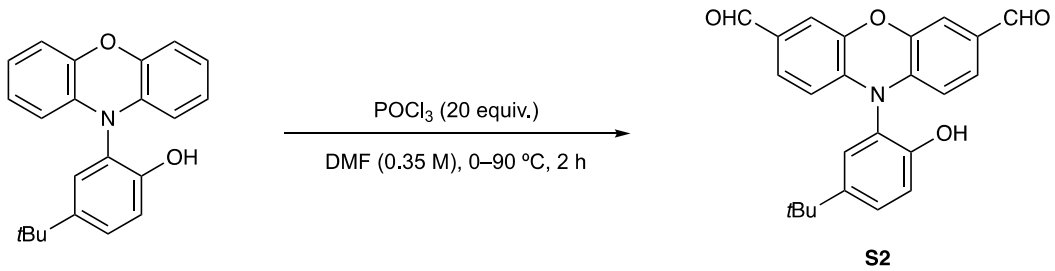
(28 mL, 0.36 M) and the solution was cooled down to 0 °C ice-water bath. Phosphorus oxychloride ($POCl_3$, 13.94 mL, 150 mmol, 15 equiv.) was then added dropwise to the reaction mixture at 0 °C. When all the required phosphorus oxychloride was added, the temperature of the reaction mixture was increased to 90 °C and the reaction was stirred for 18 h at this temperature. The reaction mixture was then cooled down to room temperature and was diluted with DCM (50 mL) and poured into a large beaker containing an ice. The quenched reaction mixture was extracted with dichloromethane (3 × 50 mL) and washed with brine (1 × 50 mL). The organic layer was separated and dried with anhydrous magnesium sulfate, filtered, and concentrated under reduced vacuum. The crude product was purified by column chromatography using 10–30% ethyl acetate in hexanes to yield a yellow powder (**S4**) in 43% yield (1.44 g, 4.3 mmol).

^{1}H NMR (500 MHz; CD_3CN-d_3) δ 9.69 (s, 2H), 7.41–7.34 (m, 5H), 7.32 (dd, J = 8.3, 1.9 Hz, 2H), 7.14 (d, J = 1.8 Hz, 2H), 6.64 (d, J = 8.3 Hz, 2H), 4.97 (s, 2H). ^{13}C NMR (126 MHz; CD_3CN-d_3) δ 190.9, 145.9, 139.0, 135.8, 132.6, 130.0, 129.4, 128.5, 127.0, 114.7, 113.9, 49.5.

10*H*-phenoxazine-3,7-dicarbaldehyde (1). To a microwave vial (2.0–5.0 mL, Biotage) fitted with a stirbar was added 10-benzyl-10*H*-phenoxazine-3,7-dicarbaldehyde (200 mg, 0.607 mmol, 1 equiv.) and was sonicated in a 2:1 solvent mixture of acetic acid (glacial):HBr (aq, 48%) (1.4:0.7 mL. total reaction volume 2.1 mL). The red-orange mixture was microwaved at 80 °C for 2 hours. After the reaction, the reaction solution becomes homogenous and turns deep red in color. The reaction was diluted with EtOAc (50 mL) and quenched with $NaHCO_3$ (saturated aqueous solution) until bubbling ceases. The mixture was extracted with EtOAc (3 × 50 mL) and washed with brine (1 × 50 mL). The organic layer was separated and dried with anhydrous magnesium sulfate, filtered, and concentrated under reduced vacuum. The crude product was purified by column chromatography using 10–30% ethyl acetate in hexanes to yield a neon yellow solid. The solid was further purified using reverse phase column using 0–90% acetonitrile in water. The fractions were lyophilized under reduced pressure to yield an orange solid in 24% yield (35 mg, 0.146 mmol).

^{1}H NMR (500 MHz; CD_3CN-d_3) δ 9.45 (s, 2H), 7.36 (s, 1H), 7.12 (dd, J = 8.2, 1.8 Hz, 2H), 6.85 (d, J = 1.8 Hz, 2H), 6.41 (d, J = 8.3 Hz, 2H). ^{13}C NMR (126 MHz; CD_3CN-d_3) δ 190.8, 144.4, 137.4, 132.3, 129.3, 115.2, 114.5. HRMS (ESI-TOF): *m/z* calculated for $C_{27}H_{22}N_5O_5^+$ $[M+H]^+$ 240.0655, found 240.0605.

c. Synthesis of fluorescence standard ***S2***



S2

10-(5-(*tert*-butyl)-2-hydroxyphenyl)-10*H*-phenoxazine-3,7-dicarbaldehyde (S2). Starting material 4-(*tert*-butyl)-2-(10*H*-phenoxazin-10-yl)phenol was prepared according to the following literature procedure[4]. 4-(*tert*-butyl)-2-(10*H*-phenoxazin-10-yl)phenol (63 mg, 0.191 mmol, 1 equiv.) was dissolved in DMF (350 µL, 0.5 M) in a 2-dram, 8-mL vial under nitrogen. The mixture was cooled down to 0 °C in an ice-water bath, and phosphorus oxychloride (353 µL, 3.82 mmol, 20 equiv.) was added dropwise via a microsyringe. The reaction vessel was then warmed to 90 °C and stirred for 2 hours. The cooled crude reaction mixture was analyzed by TLC, then directly purified by flash column chromatography using 10–40% EtOAc in hexanes. Fluorescent standard **S2** was isolated as a brown powder in 39% yield (29 mg, 0.075 mmol).

^{1}H NMR (500 MHz; DMSO-d_6) δ 9.93 (s, 1H), 9.67 (s, 2H), 7.45 (dd, J = 9.8, 2.0 Hz, 1H), overlapping 7.32 (apparent dd, J = 9.8, 2.0 Hz, 2H) and 7.32 (apparent d, J = 1.8 Hz, 1H), 7.17 (d, J = 1.8 Hz, 2H), 7.09 (d, J = 8.6 Hz, 1H), 5.99 (d, J = 8.2 Hz, 2H), 1.27 (s, 9H). ^{13}C NMR (126 MHz; DMSO-d_6) δ 190.3, 151.8, 144.0, 143.6, 137.9, 131.0, 128.2, 127.9, 127.1, 121.4, 117.5, 114.1, 113.4, 34.0, 31.3. (ESI-TOF): *m/z* calculated for $C_{24}H_{22}NO_4^+$ $[M+H]^+$ 388.1544, found 388.1467.

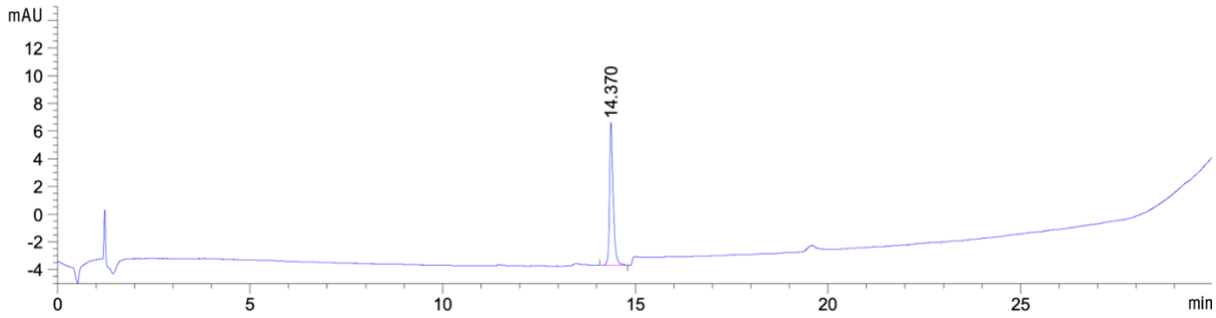
[4] Louillat-Habermeyer, M.-L., Jin, R. & Patureau, F. W. *Angew. Chem. Int. Ed.* **54**, 4102–4104 (2015).

X. HPLC spectra of substrates and relevant reaction components

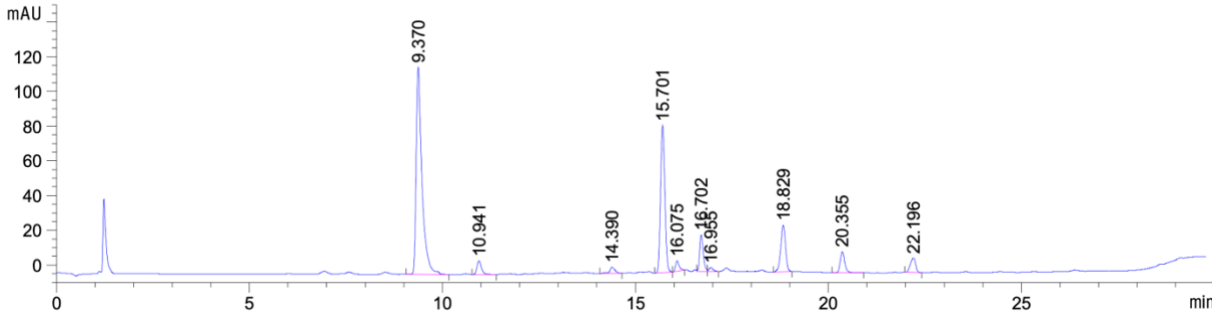
HPLC eluent conditions: 0–73% MeCN/H_2O over 25 min, 0.1% trifluoroacetic acid, flow rate 1 mL/min.

*a. Bivalirudin (**3**) (254 nm)*

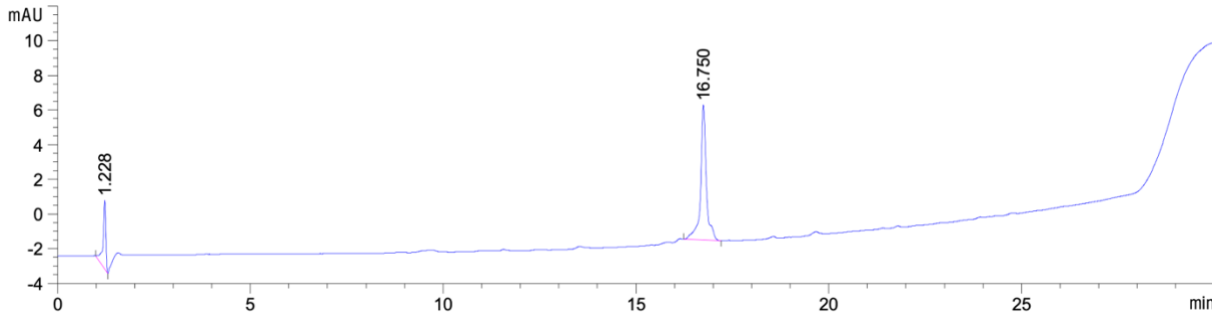
Starting material:



Crude reaction mixture:

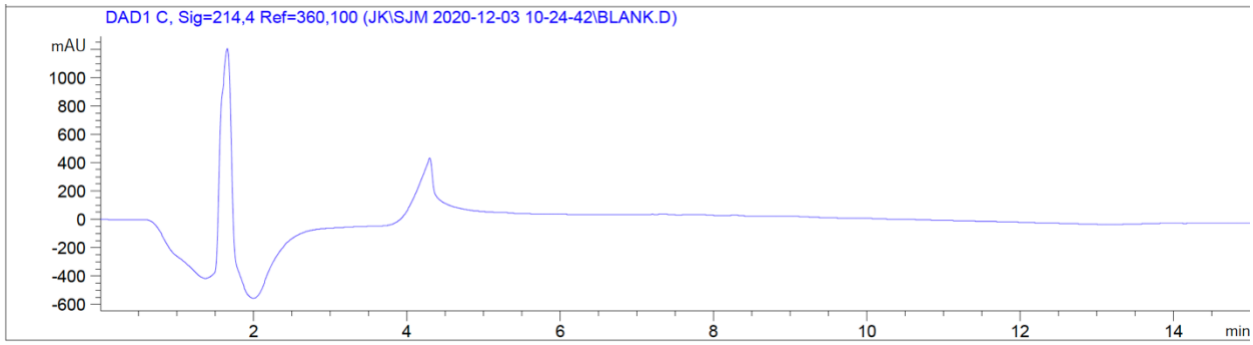


After removal of small molecule reaction components:



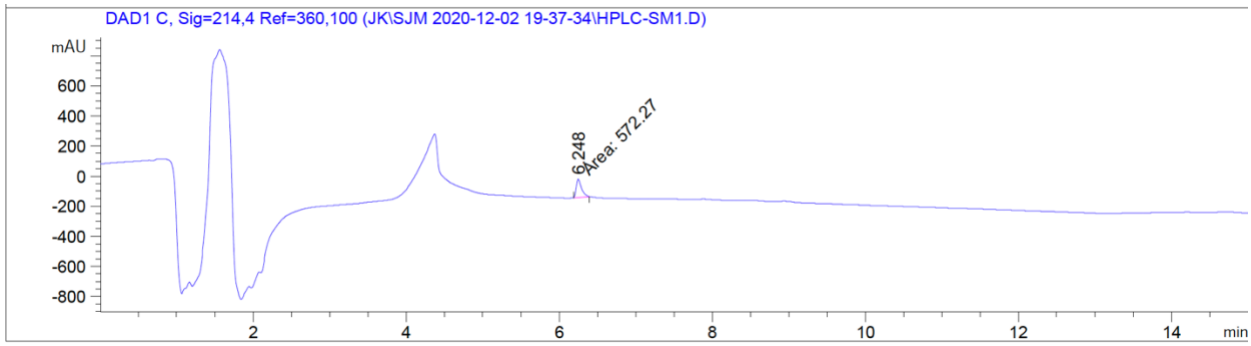
HPLC eluent conditions: 0–73% MeCN/H_2O over 10 min, 0.1% formic acid, flow rate 1 mL/min.

Blank injection:

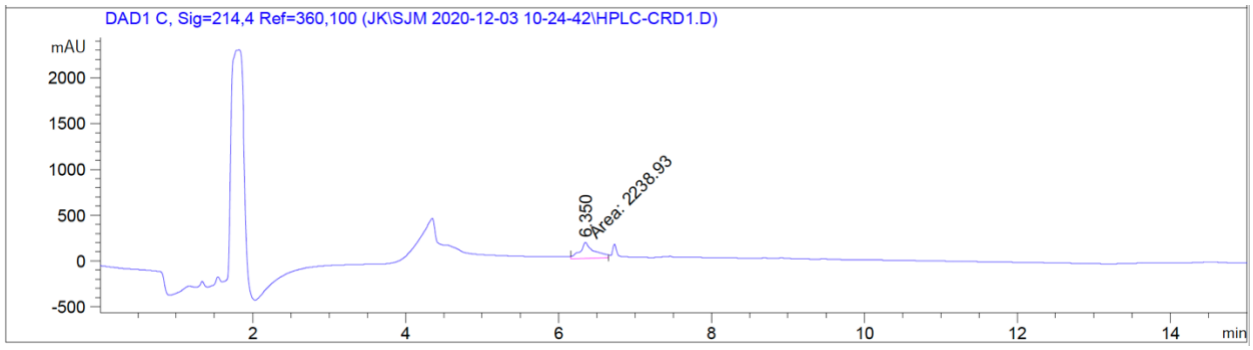


*b. ZVar affibody (**4**)*

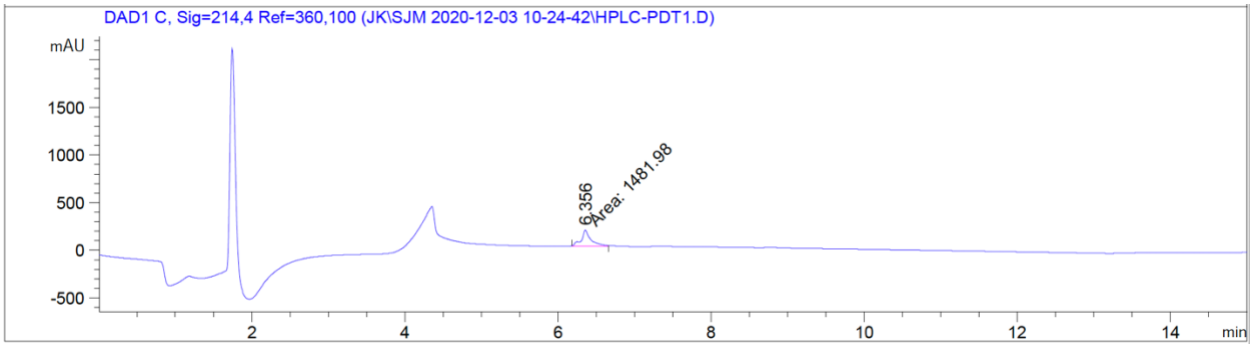
Starting material:



Crude reaction mixture:

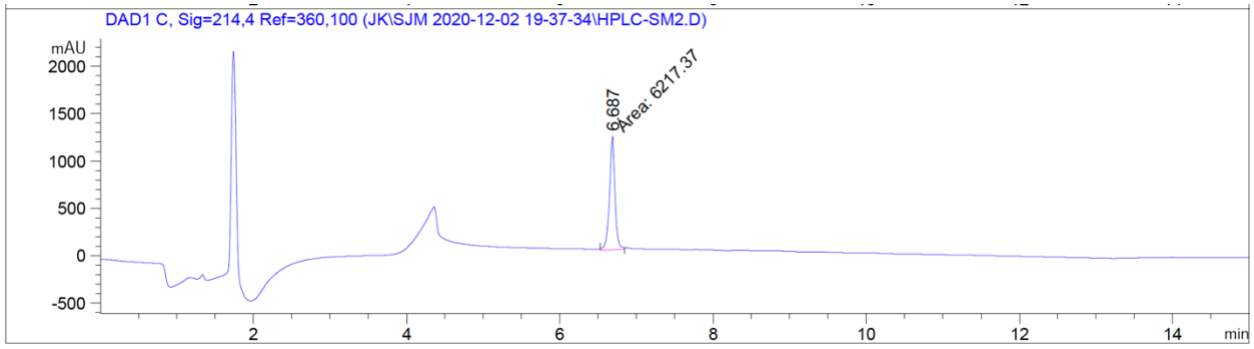


After removal of small molecule reaction components:

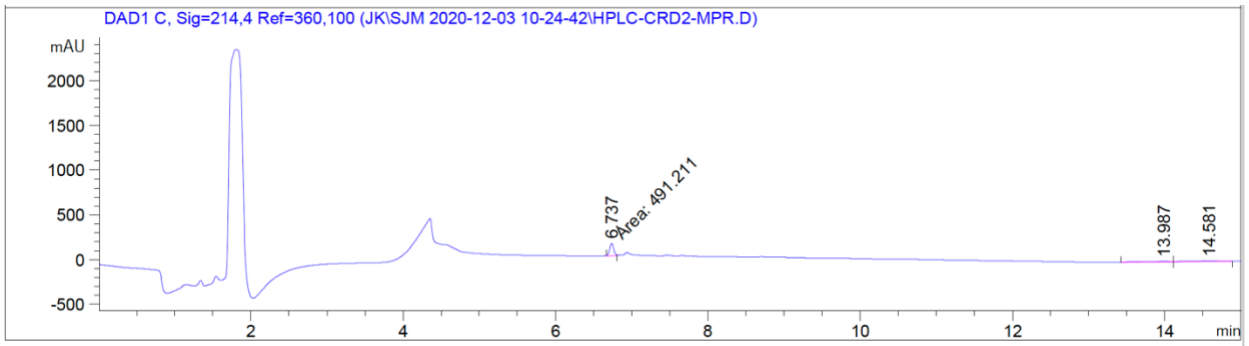


*b. Human insulin (**5**)*

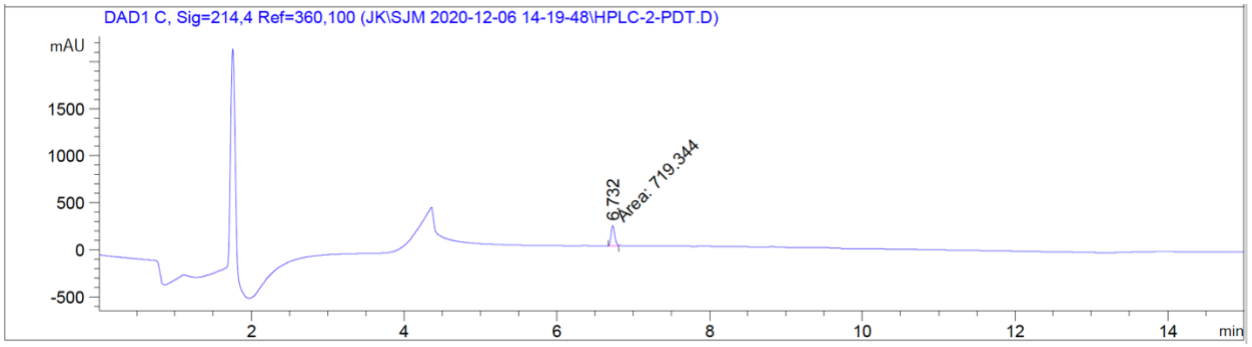
Starting material:



Crude reaction mixture:

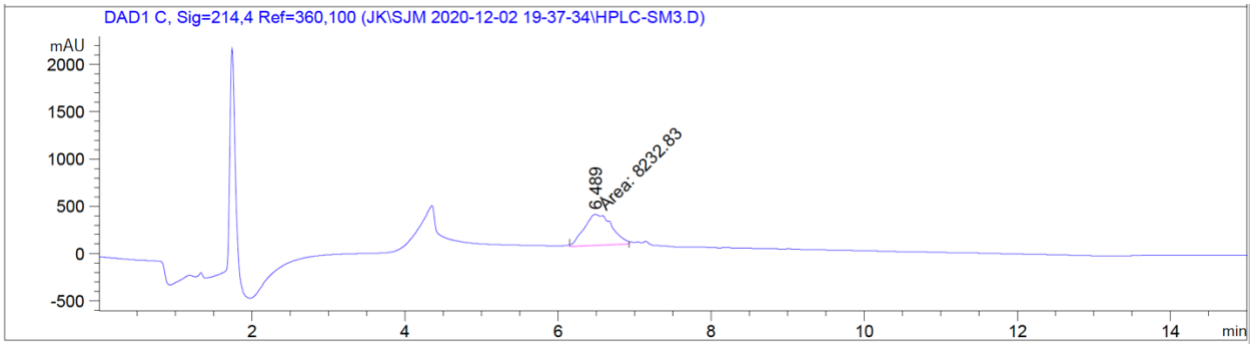


After removal of small molecule reaction components:

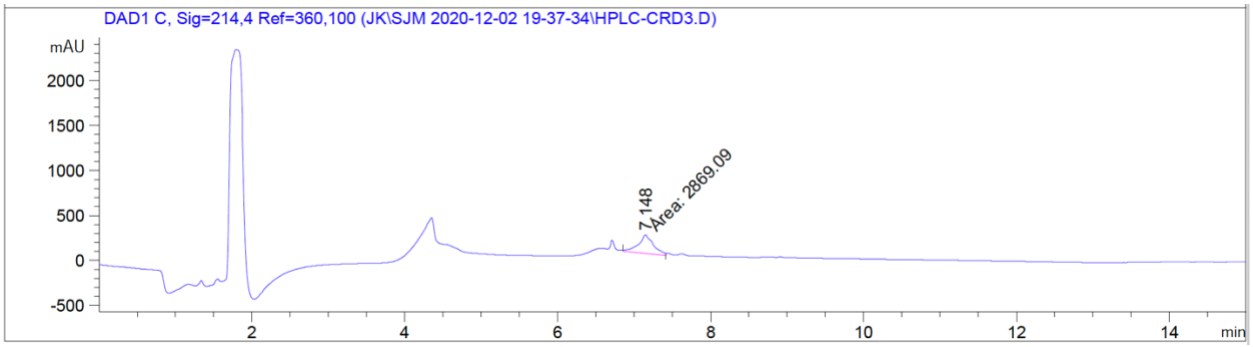


c. Human lysozyme ***(6)***

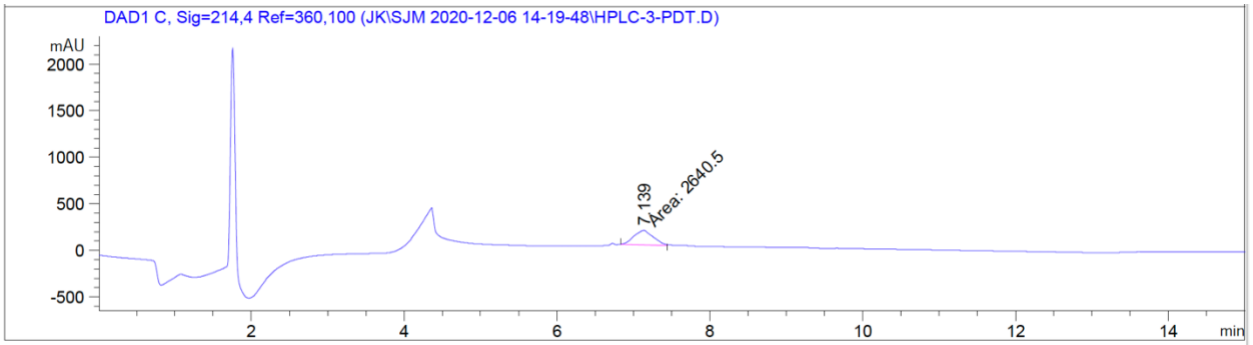
Starting material:



Crude reaction mixture:

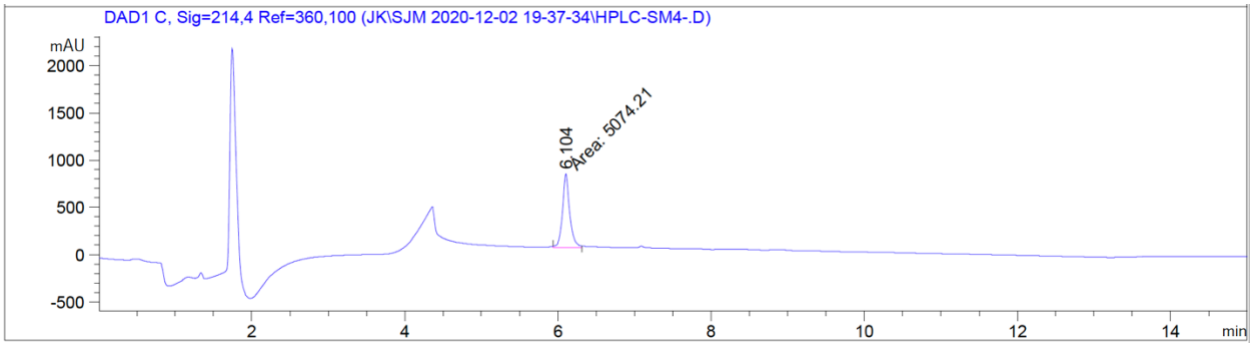


After removal of small molecule reaction components:

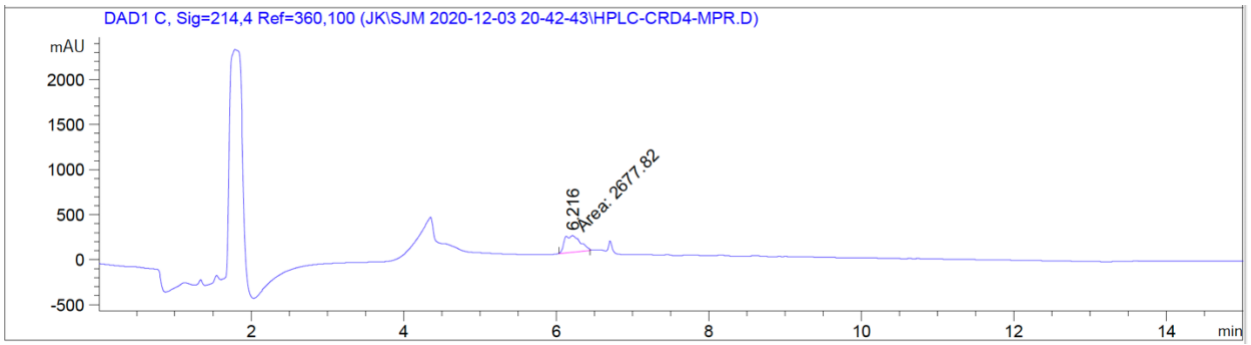


d. Ribonuclease A (7)

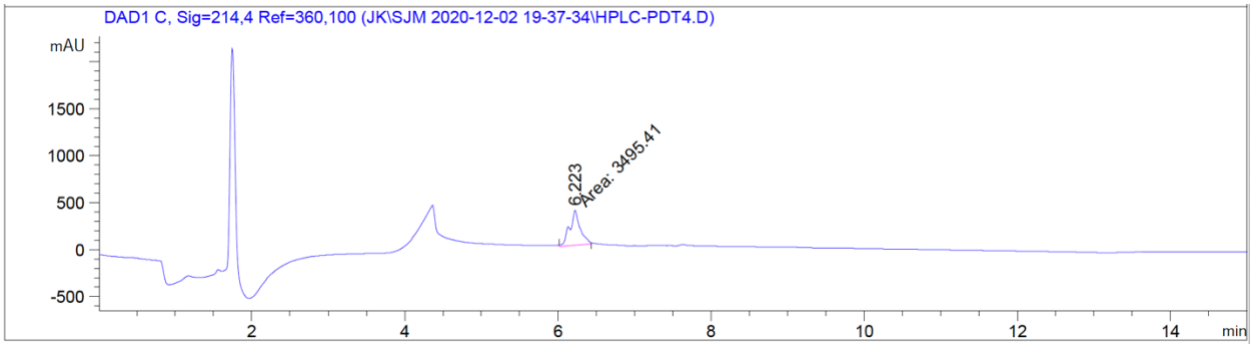
Starting material:



Crude reaction mixture:

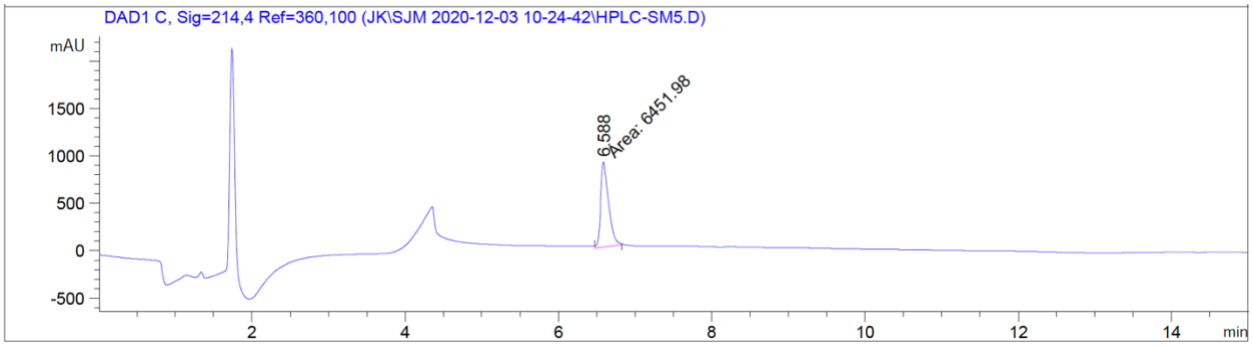


After removal of small molecule reaction components:

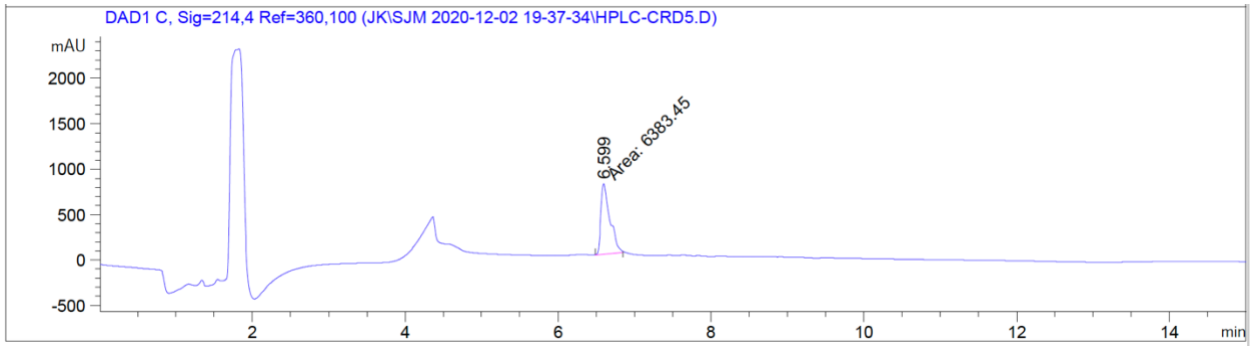


e. Cytochrome C ***(8)***

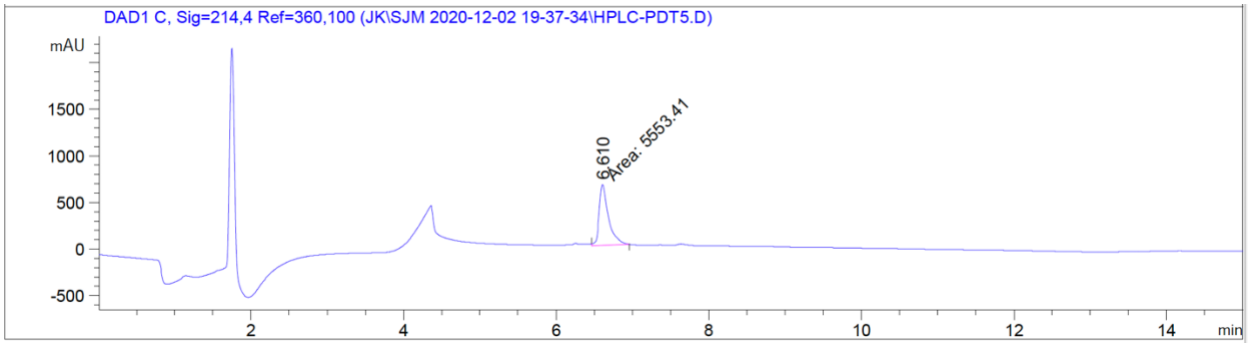
Starting material:



Crude reaction mixture:

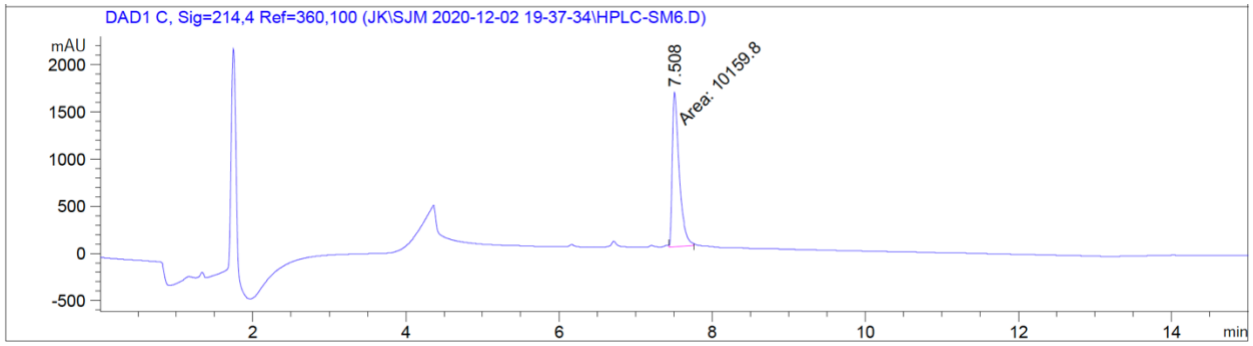


After removal of small molecule reaction components:

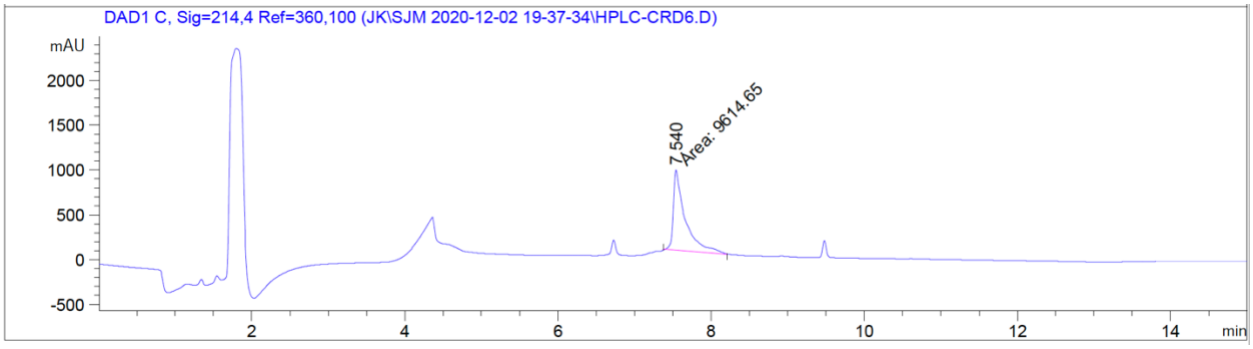


*f. α-Lactalbumin (**9**)*

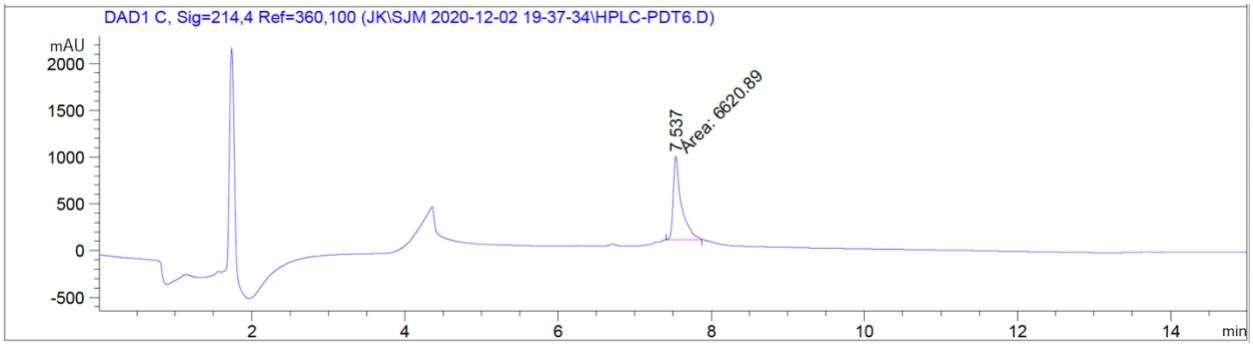
Starting material:



Crude reaction mixture:

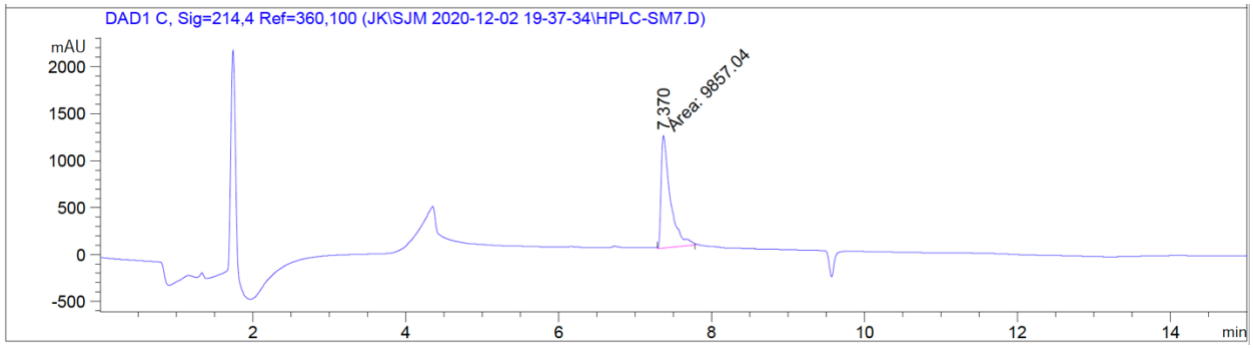


After removal of small molecule reaction components:

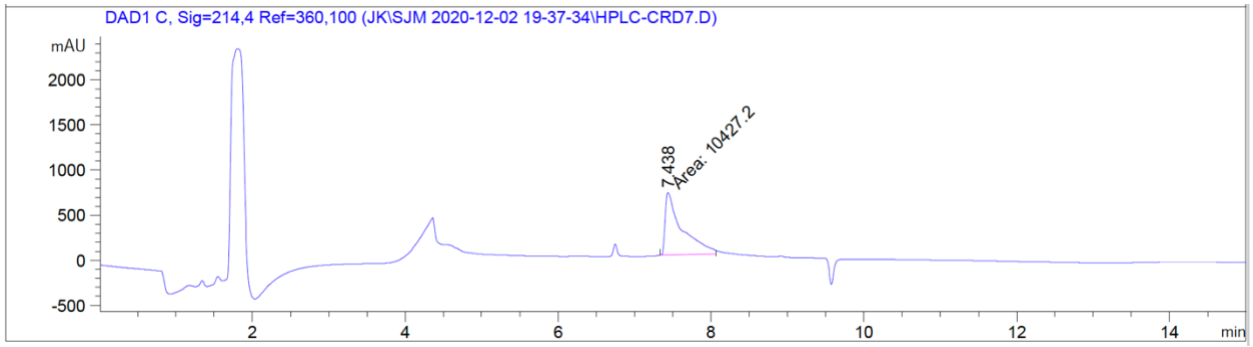


*g. Myoglobin (**10**)*

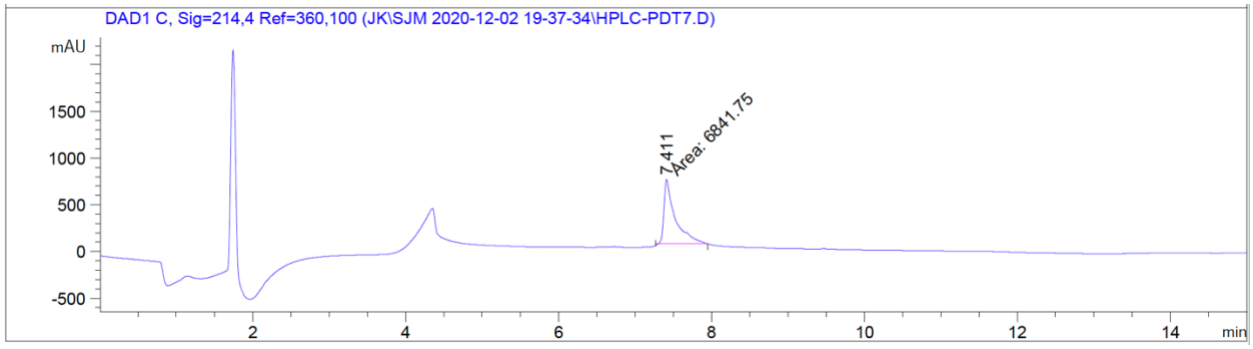
Starting material:



Crude reaction mixture:

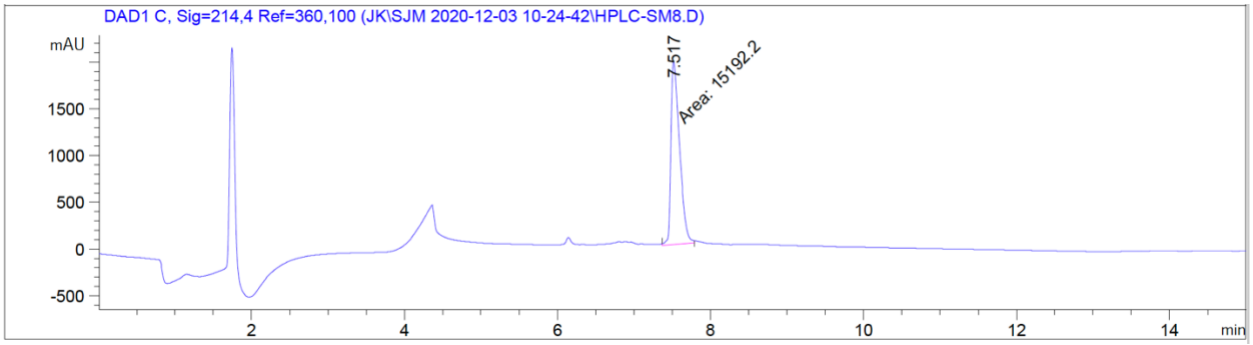


After removal of small molecule reaction components:

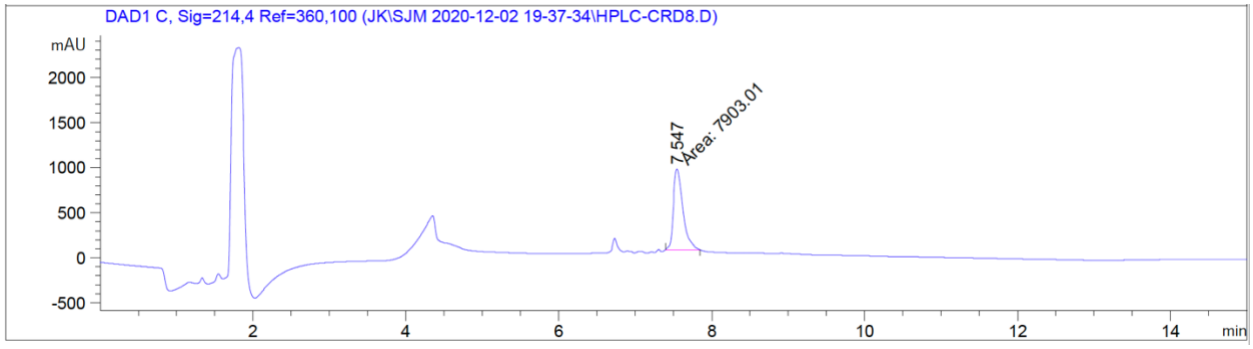


*h. Chymotrypsinogen A (**11**)*

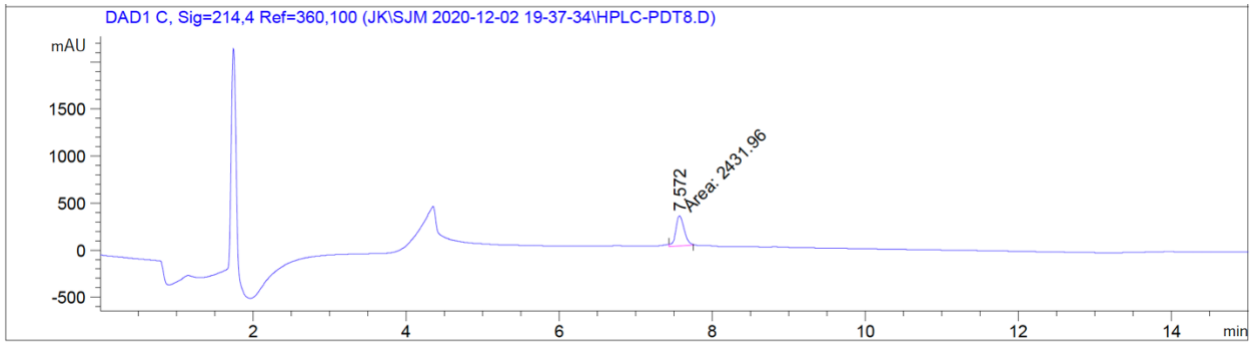
Starting material:



Crude reaction mixture:

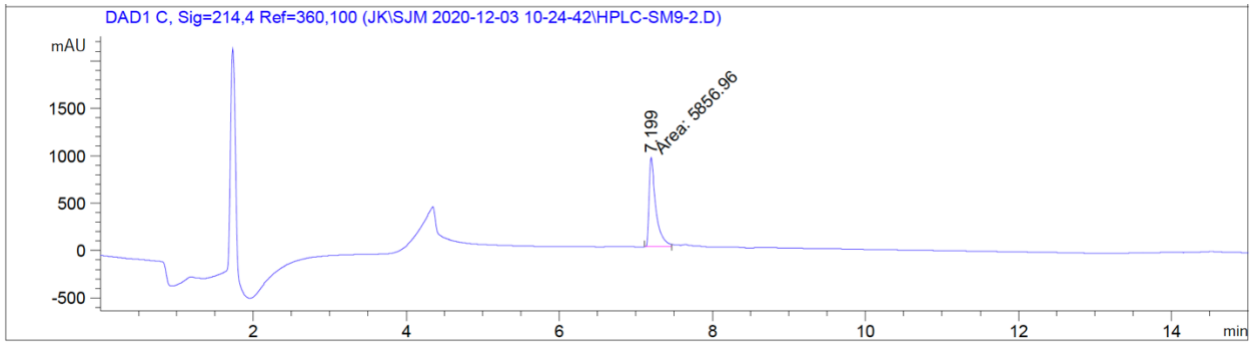


After removal of small molecule reaction components:

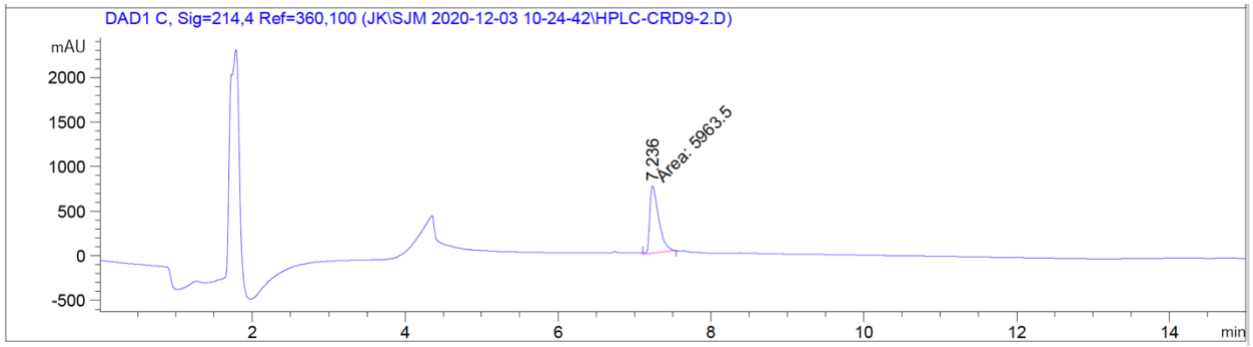


*i. Serotransferrin (**12**)*

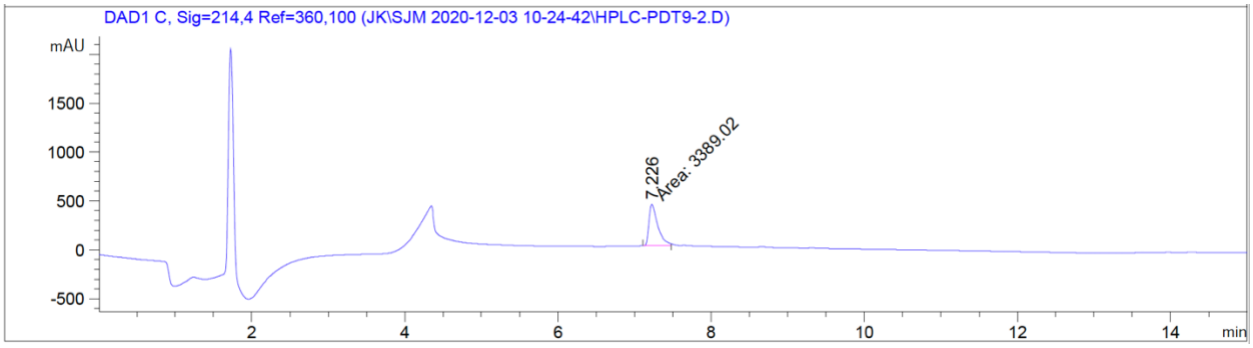
Starting material:



Crude reaction mixture:



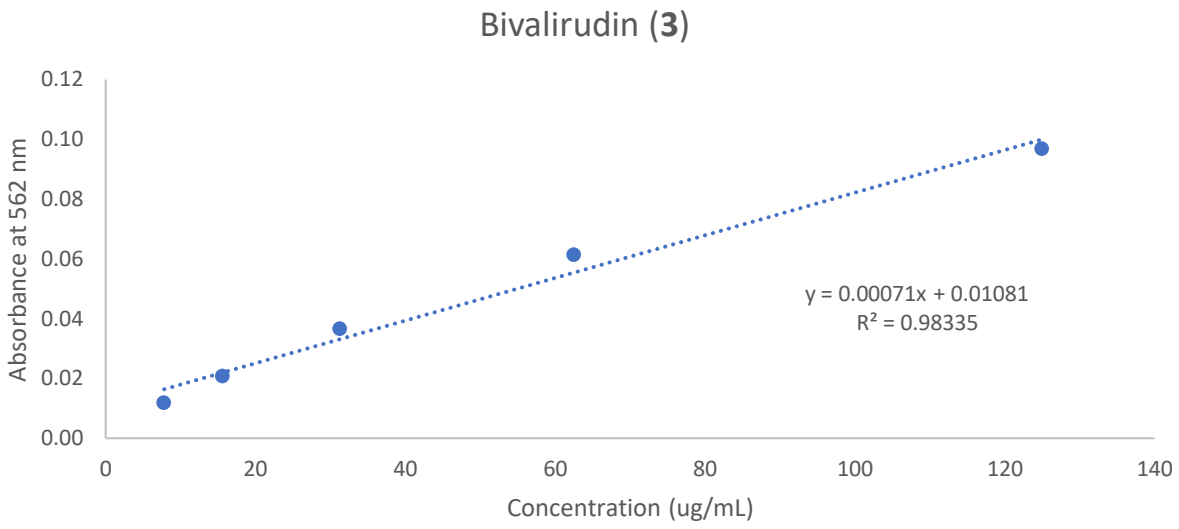
After removal of small molecule reaction components:



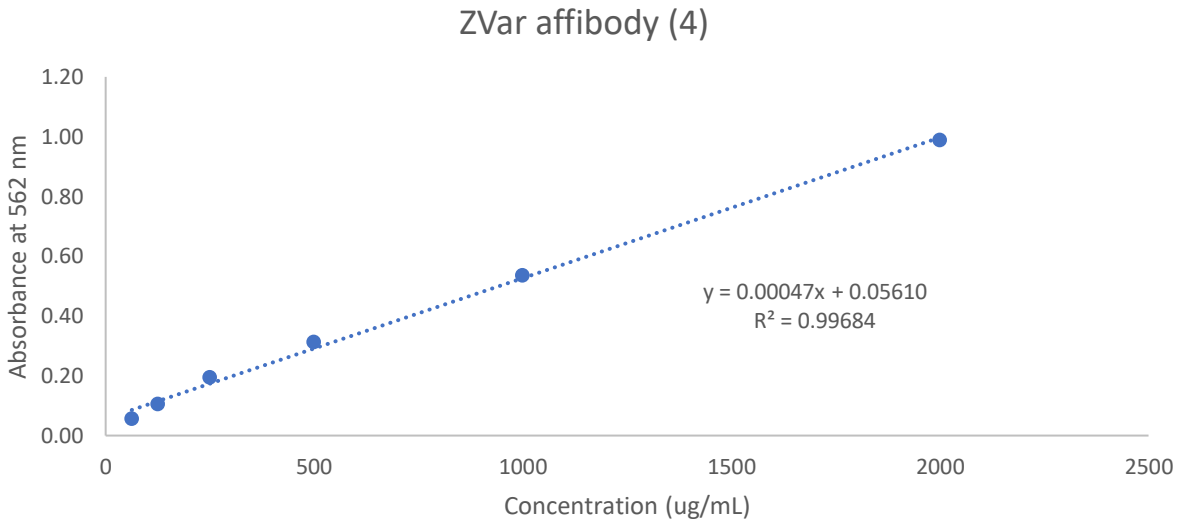
XI. Protein concentration calibration curves via BCA assay

A separate BCA assay calibration curve is produced for each protein substrate. Combined with the fluorescence calibration curve, they are used to determine conversion for each of the protein substrates.

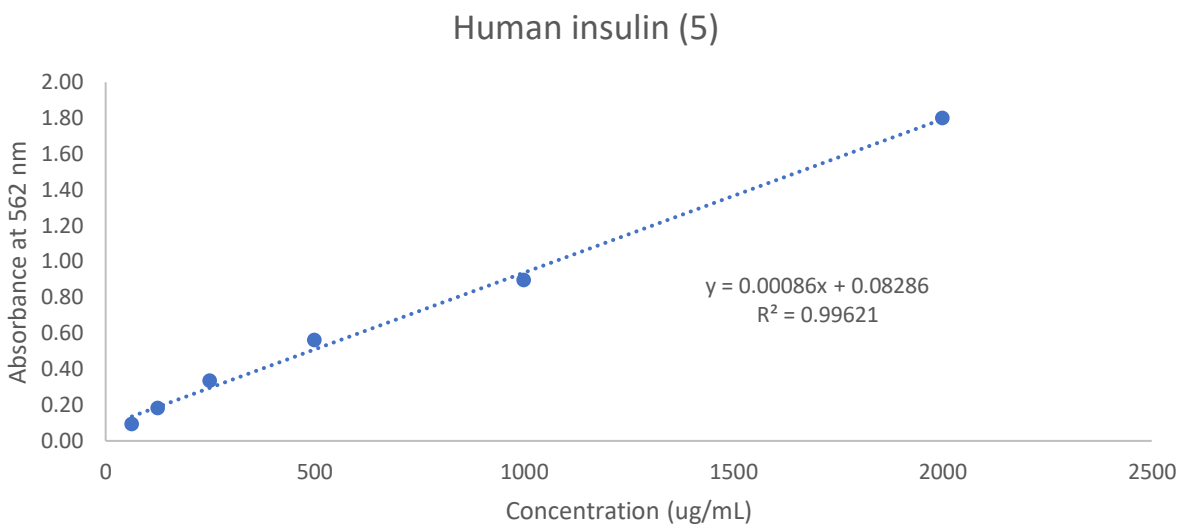
*a. Bivalirudin (**3**)*



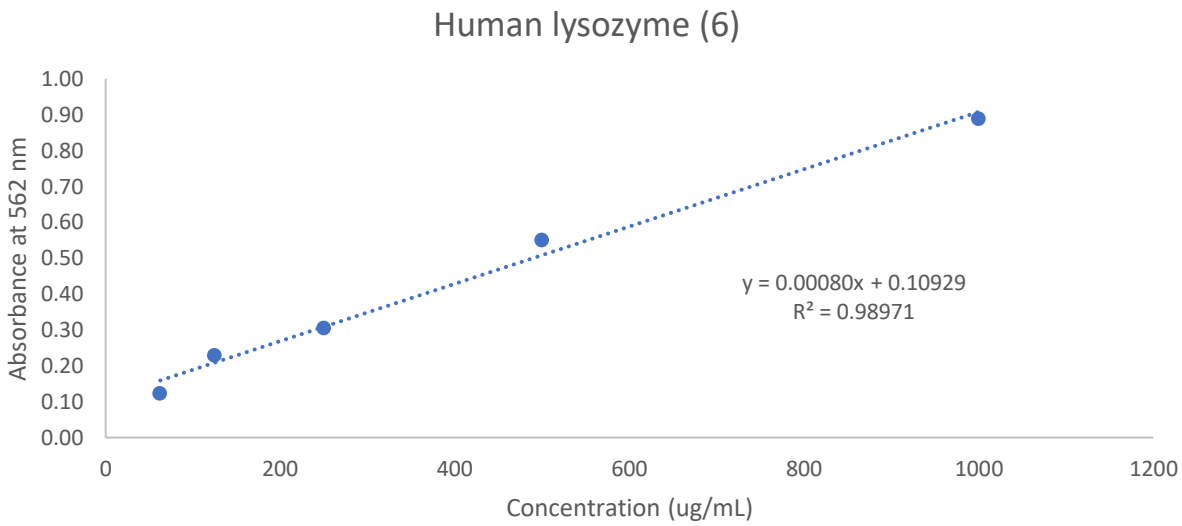
*b. ZVar affibody (**4**)*



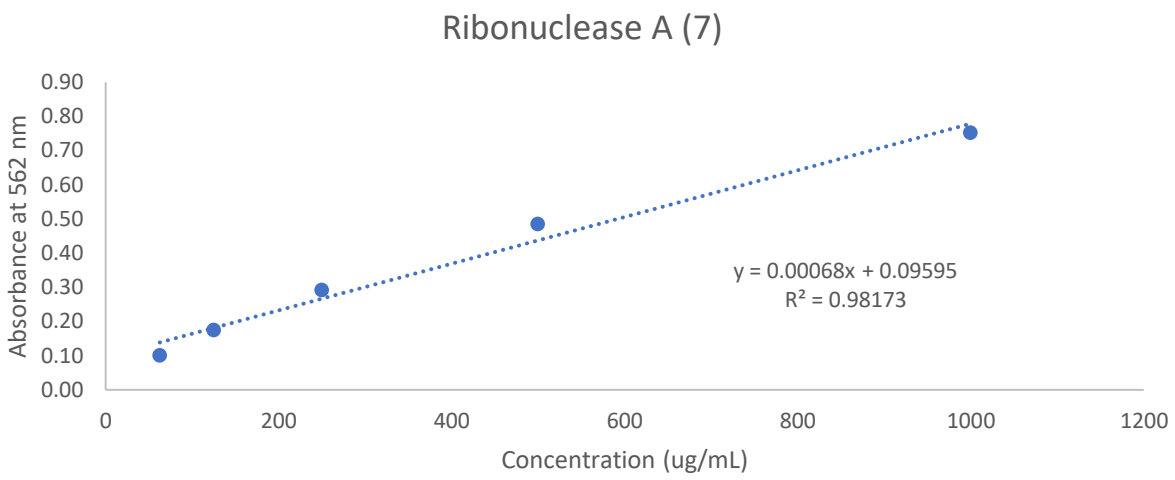
*c. Human insulin (**5**)*



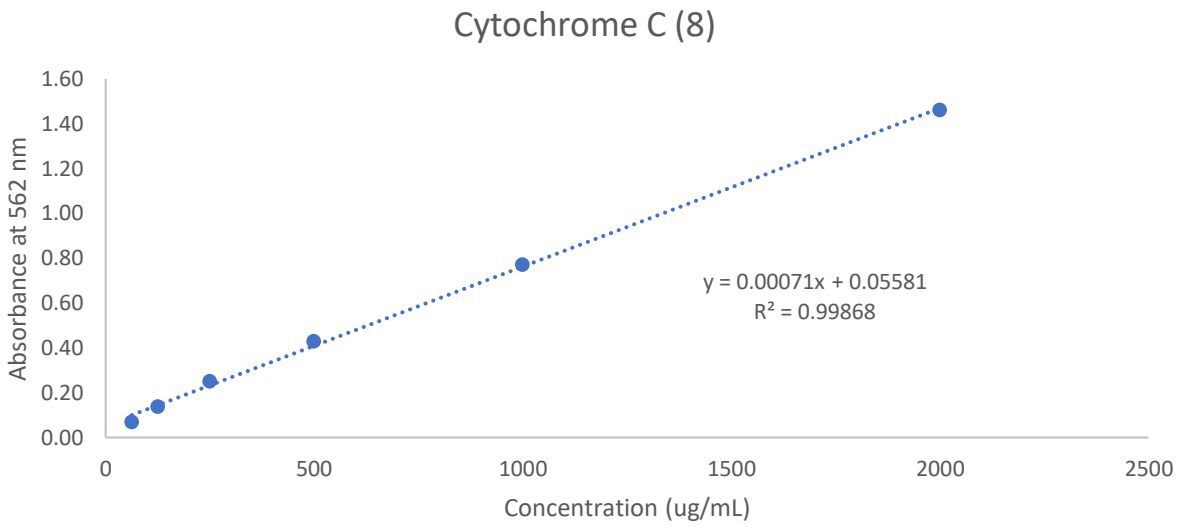
*d. Human lysozyme (**6**)*



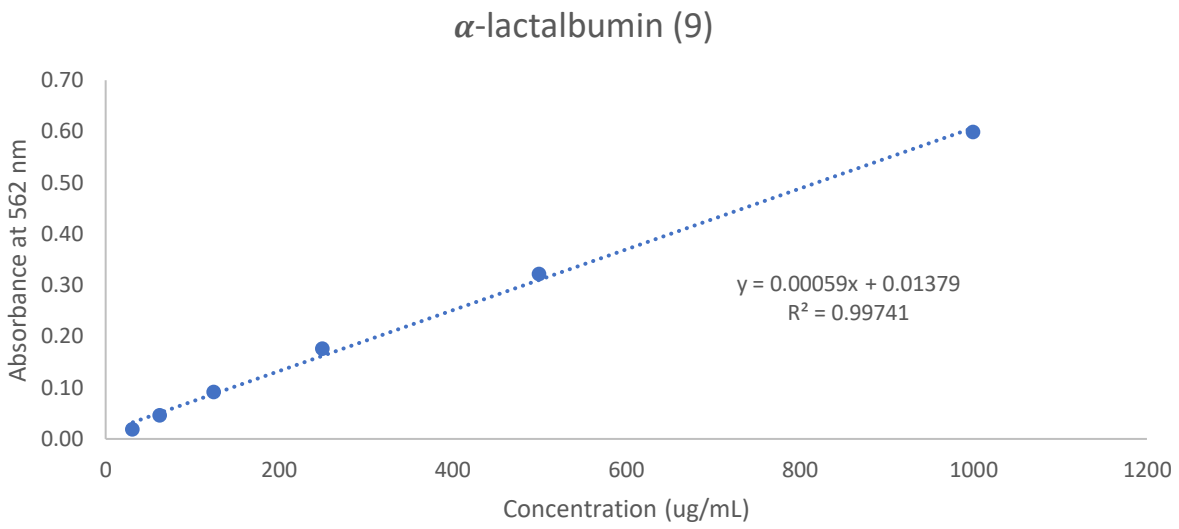
e. Ribonuclease A (7)



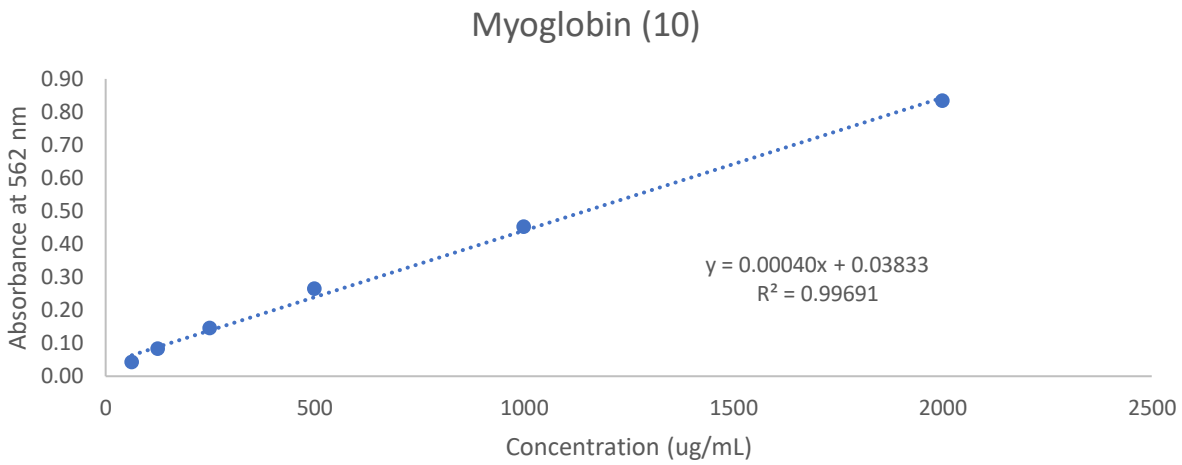
*f. Cytochrome C (**8**)*



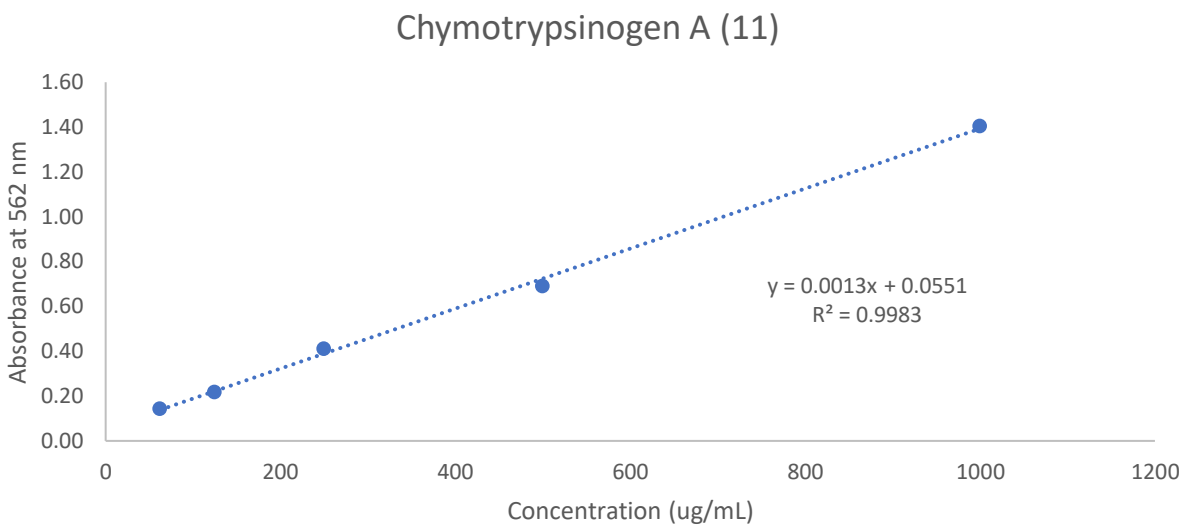
*g. α-Lactalbumin (**9**)*



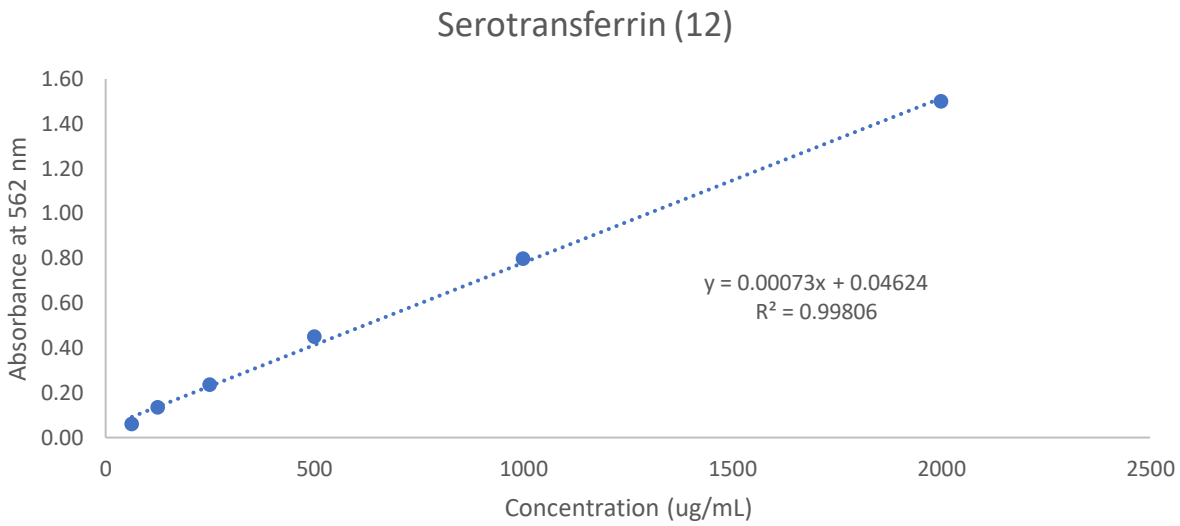
*h. Myoglobin (**10**)*



*i. Chymotrypsinogen A **(11)***



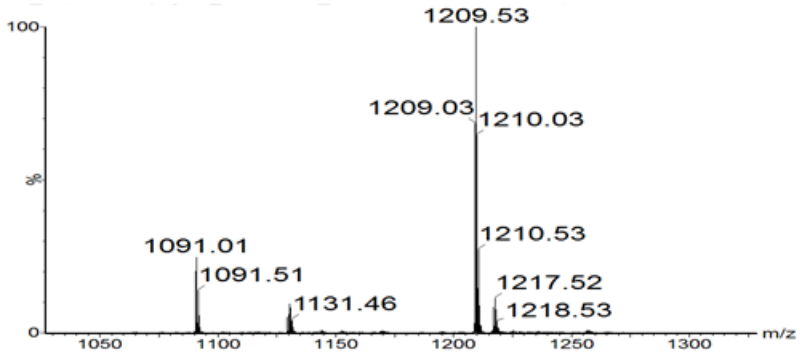
*j. Serotransferrin **(12)***



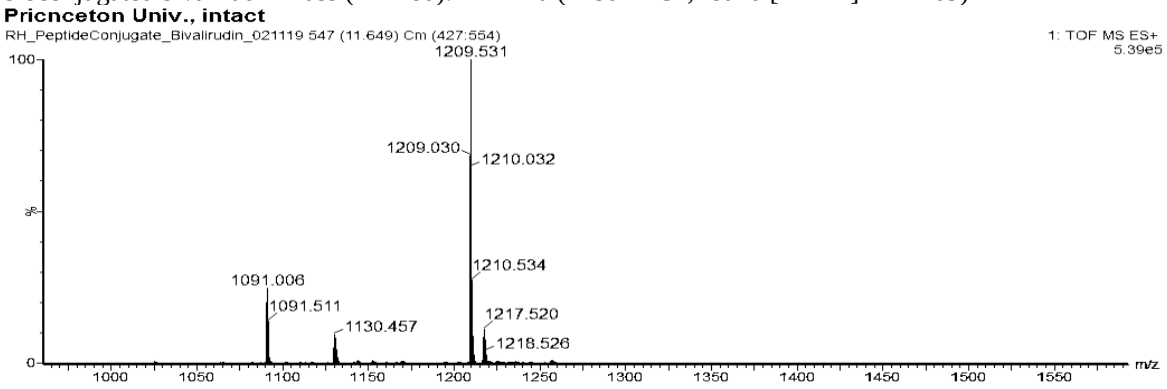
XII. Intact mass spectra

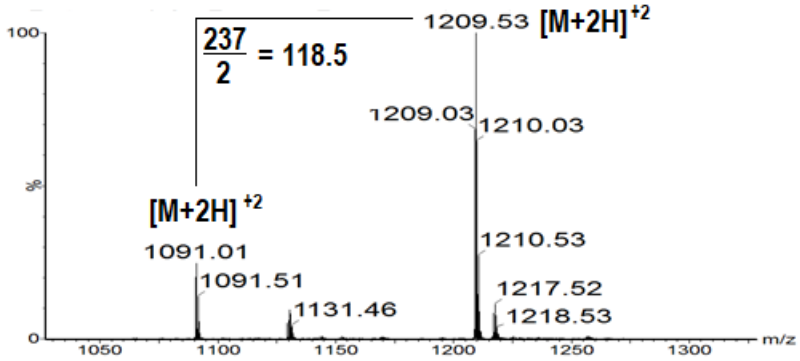
Mass spectrometry. High Resolution Intact-Mass Spectra & Analysis were obtained from Princeton University Mass Spectral Facility using an Agilent 6546 LC/Q-TOF & MassHunter BioConfirm Software or from Bristol-Myers Squibb (BMS) using Waters Synapt G2 QTOF mass spectrometer connected to Aquity UPLC (Waters, Milford, MA). Samples at BMS were desalted and eluted from a reversed-phase C4 column (ACQUITY UPLC BEH C18 column 300A 1.7um 2.1x150mm).

*a. Bivalirudin (**3**)*
native bivalirudin starting material mass: 2180 Da (found $[M+2H]^{+2}$ = 1091)



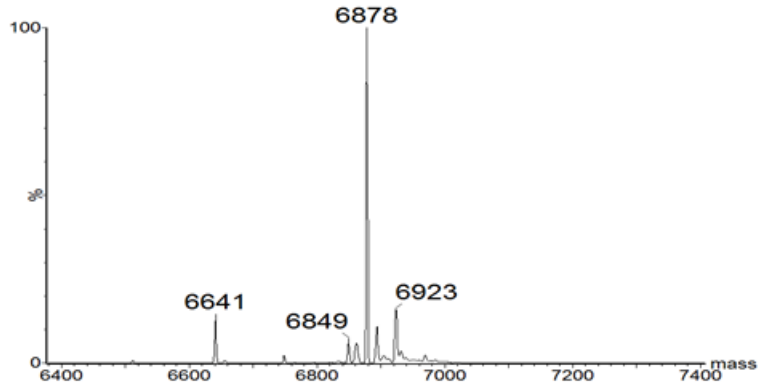
bioconjugated bivalirudin mass (+1 mod): 2417 Da (2180 + 237; found $[M+2H]^{+2}$ = 1209)



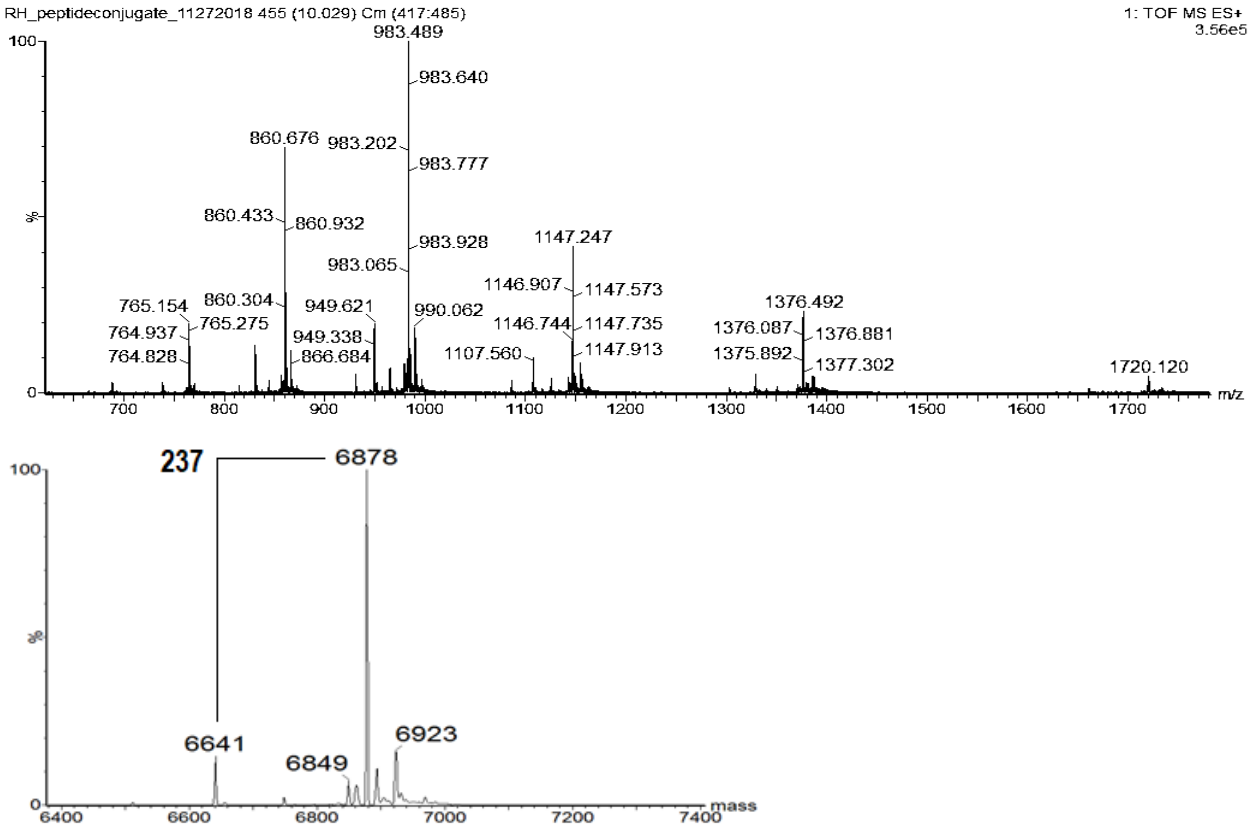


*b. ZVar affibody (**4**)*

native ZVar affibody starting material mass: 6641 Da

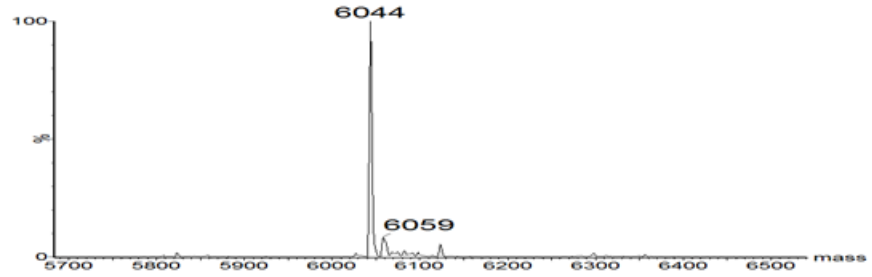


bioconjugated ZVar affibody mass (+1 modification): 6878 Da (6641 + 237)

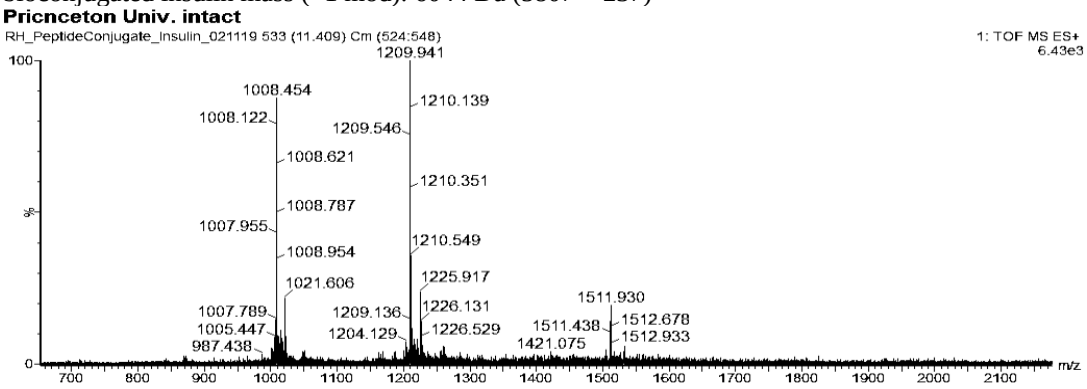


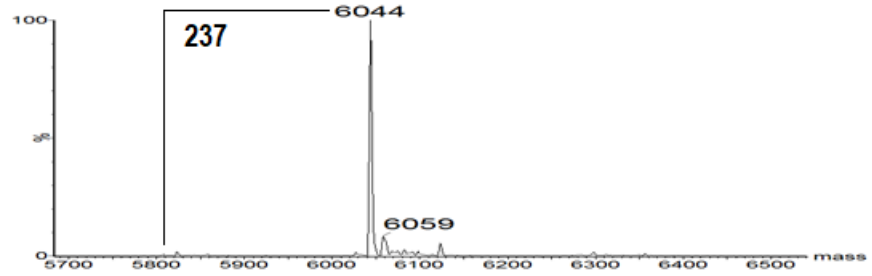
*c. Human insulin (**5**)*

native insulin starting material mass: 5807 Da



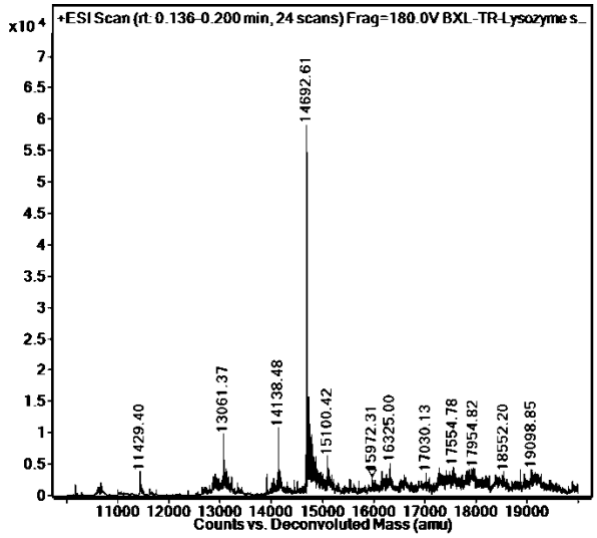
bioconjugated insulin mass (+1 mod): 6044 Da (5807 + 237)



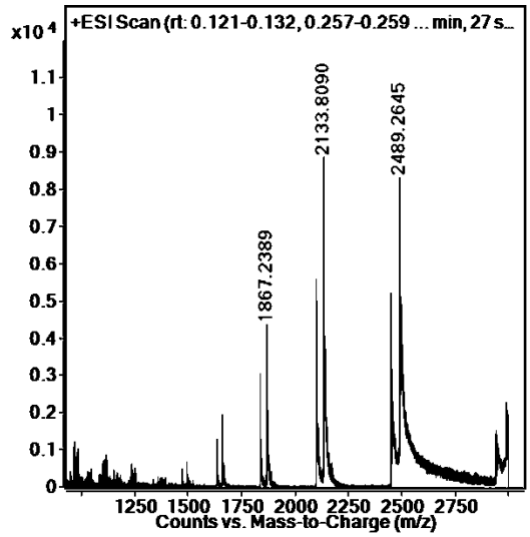


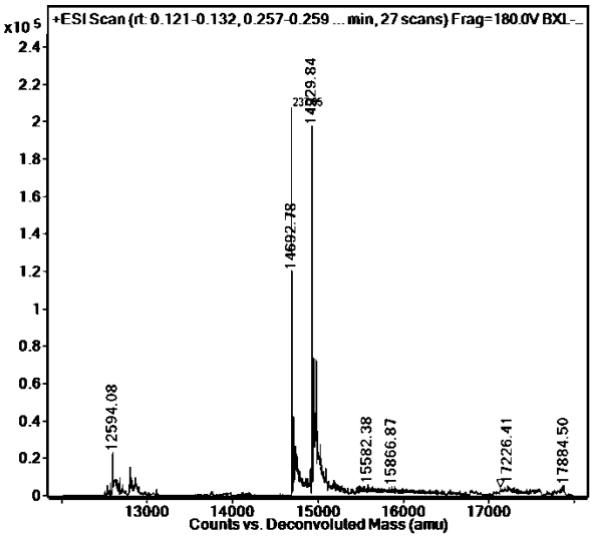
*d. Human lysozyme (**6**)*

native lysozyme starting material mass: 14693 Da



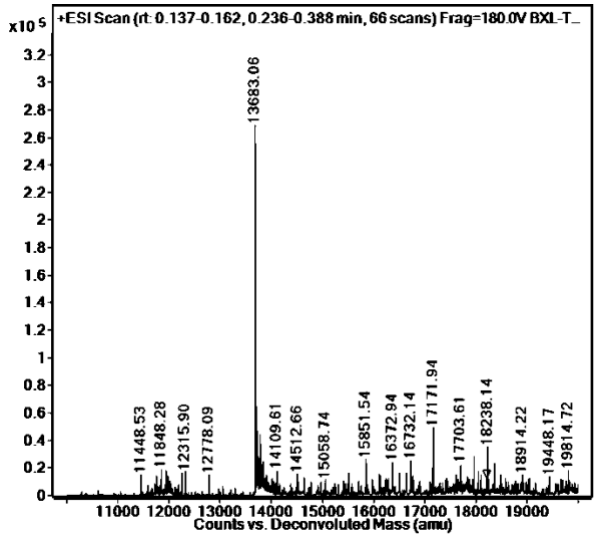
bioconjugated lysozyme mass (+1 modification): 14930 Da (14693 + 237)



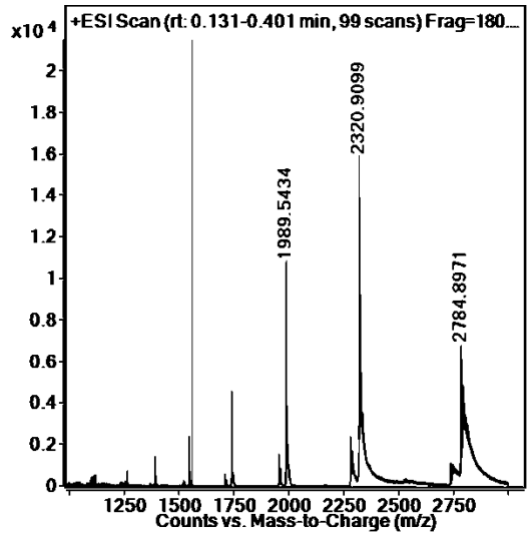


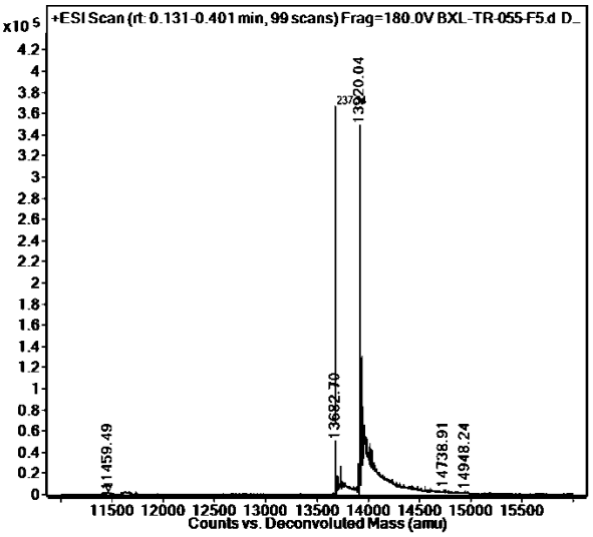
e. Ribonuclease A (7)

native ribonuclease A starting material mass: 13683 Da



bioconjugated ribonuclease A mass (+1 modification): 13920 Da (13683 + 237)

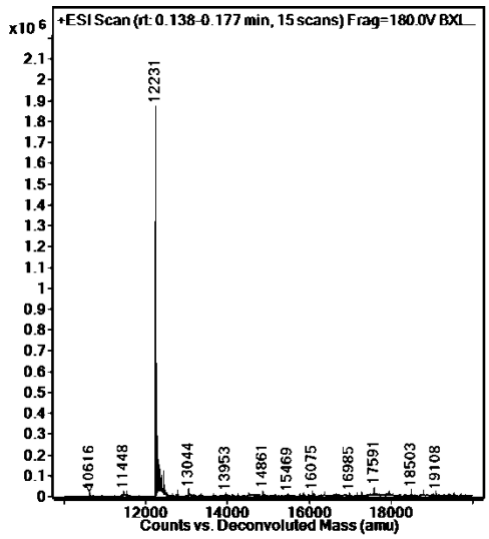




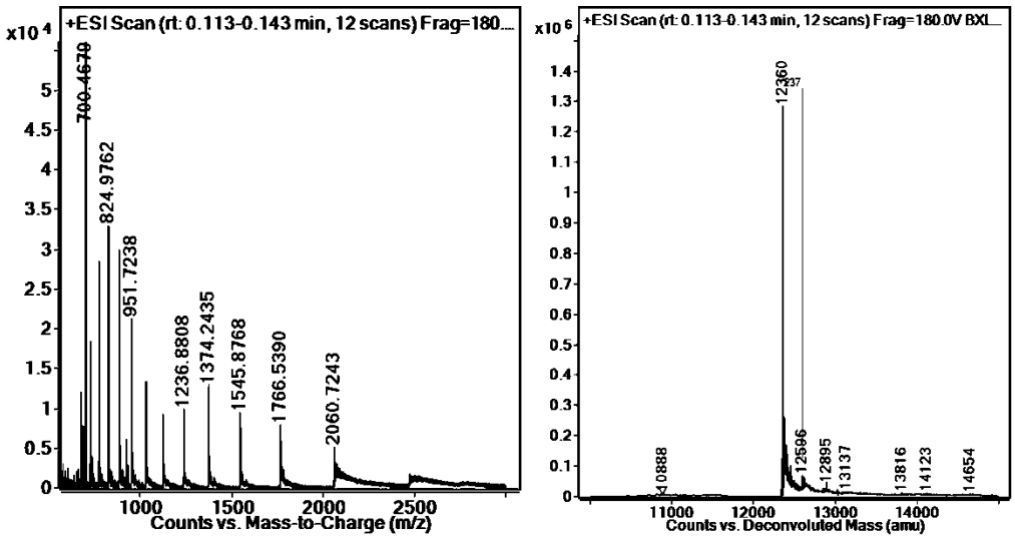
*f. Cytochrome C (**8**)*

Intact spectrum collected at under incomplete conversion conditions.

native cytochrome C starting material mass: 12358 Da



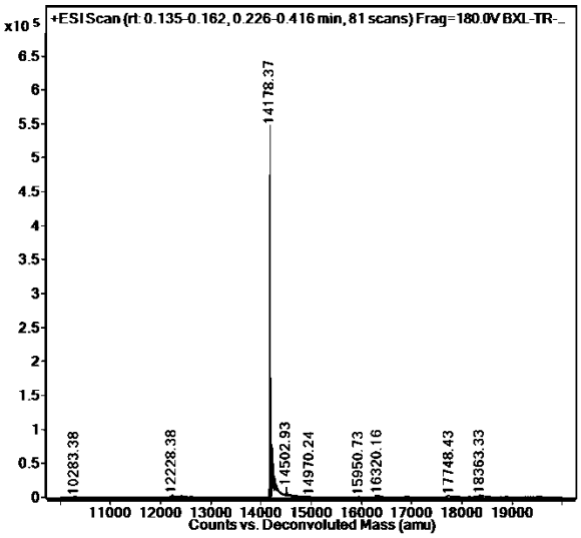
bioconjugated cytochrome C mass (+1 modification): 12595 Da (12358 + 237)



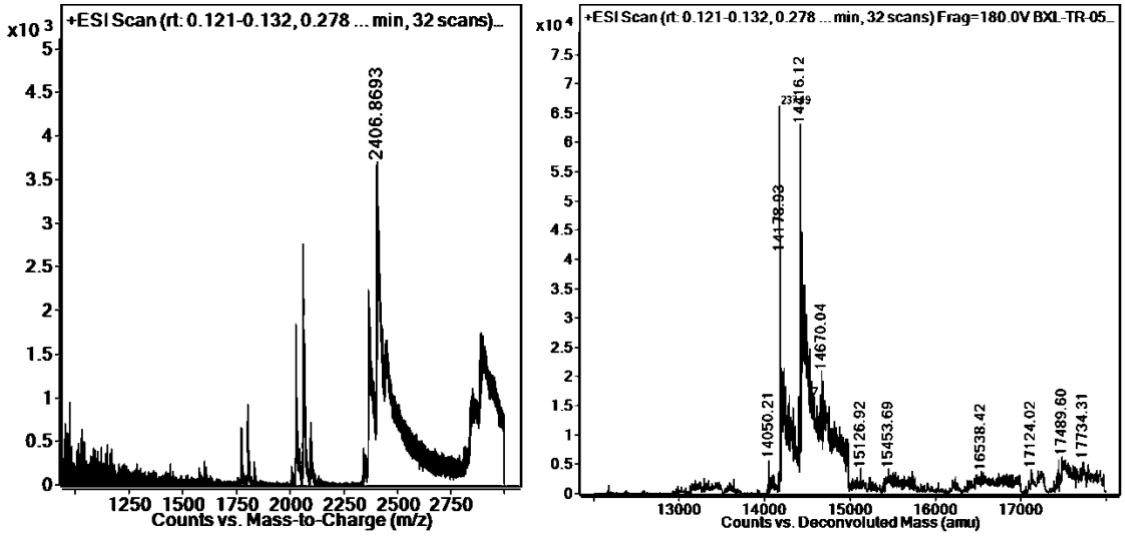
*g. α-Lactalbumin (**9**)*

Using General Procedure B

native *α*-lactalbumin starting material mass: 14178 Da



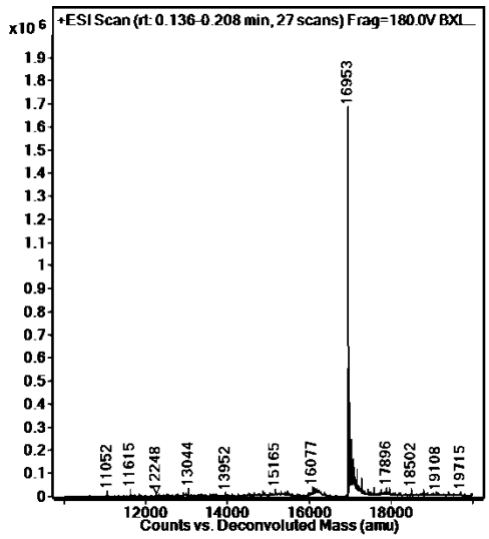
bioconjugated *α*-lactalbumin mass (+1 modification): 14415 Da (14178 + 237)



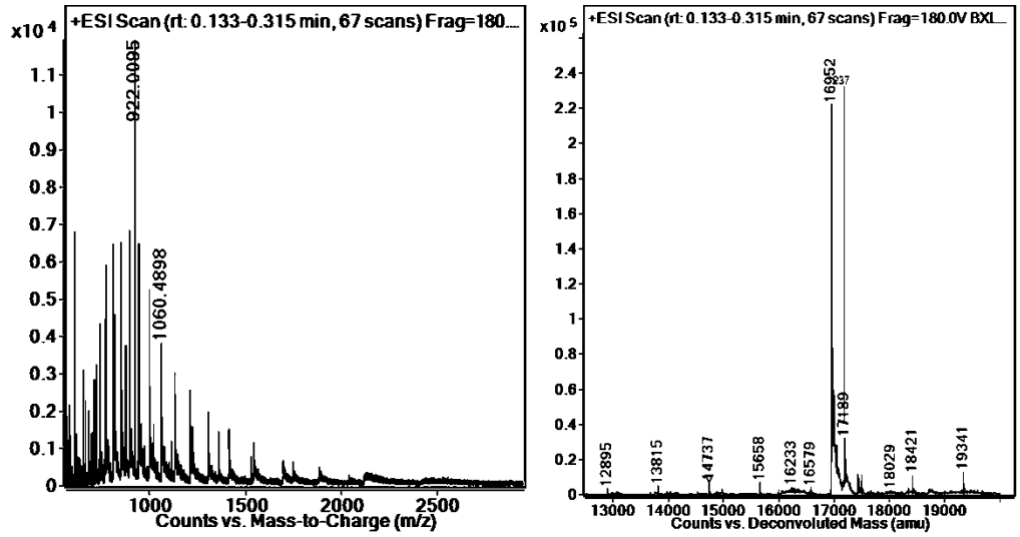
*h. Myoglobin (**10**)*

Intact spectrum collected at under incomplete conversion conditions.

native myoglobin starting material mass: 16952 Da

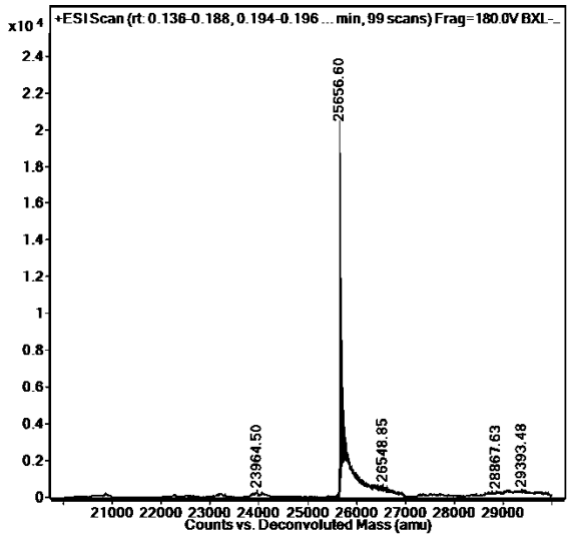


bioconjugated myoglobin mass (+1 modification): 17189 Da (16952 + 237)

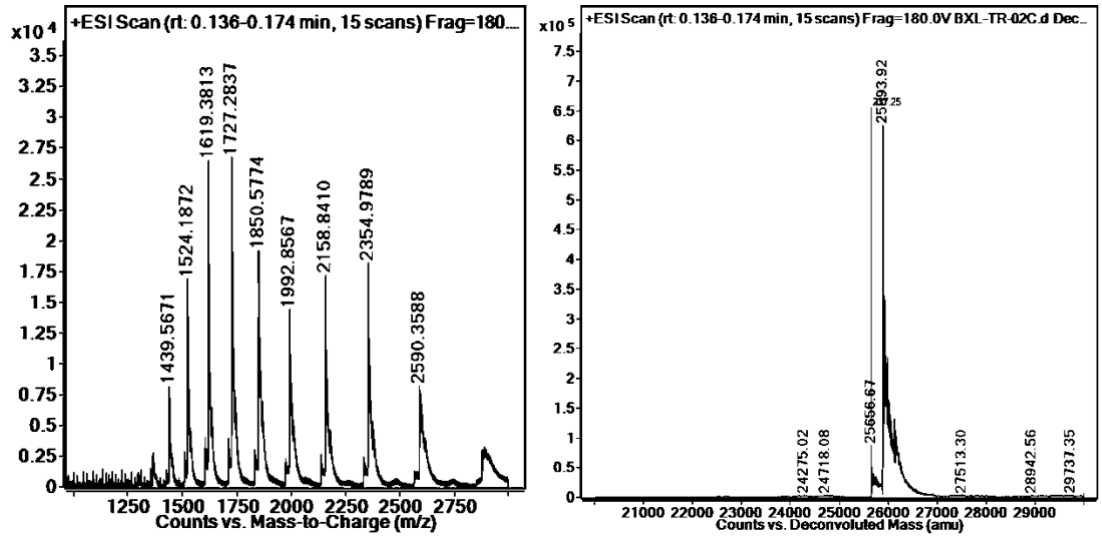


i. Chymotrypsinogen A ***(11)***

native chymotrypsinogen starting material mass: 25657 Da



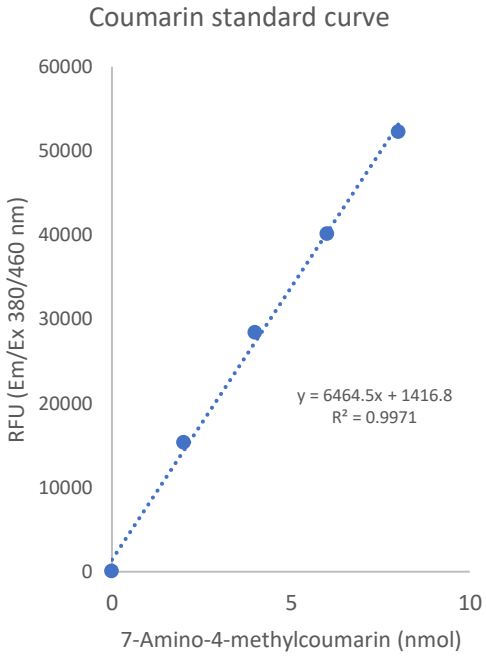
bioconjugated chymotrypsinogen mass (+1 modification): 25894 Da (25657 + 237)

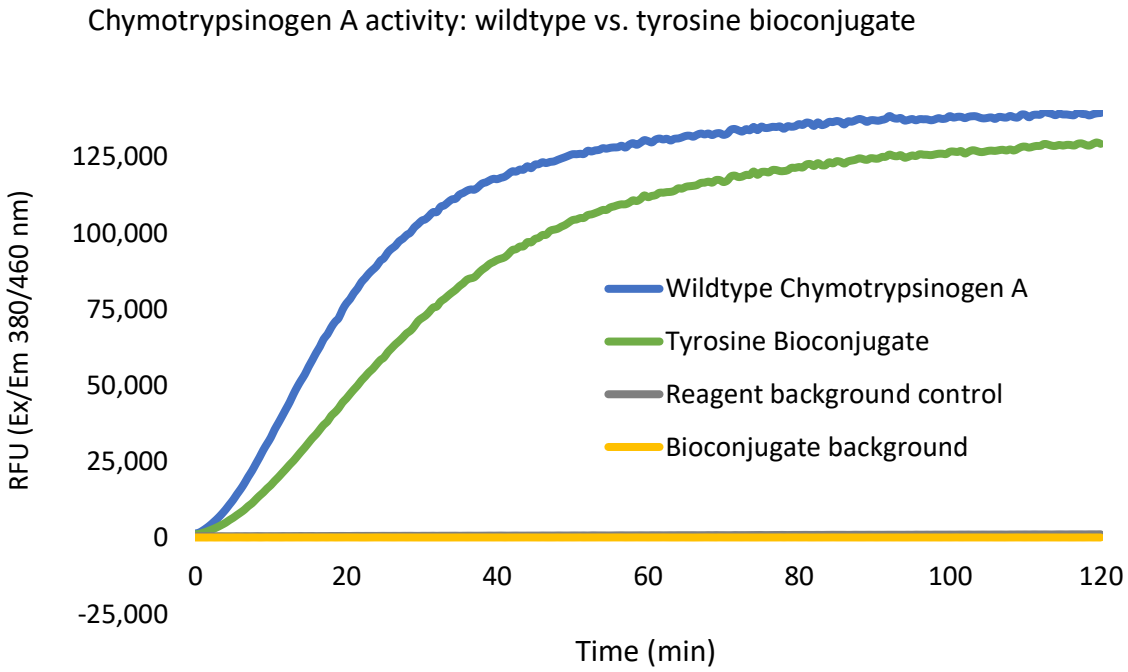


XIII. Chymotrypsinogen A assay

We assessed the integrity of one of our bioconjugation products, Chymotrypsinogen A, by assessing the peptide cleavage efficiencies of this protein before and after modification. We were able to do so with the commercially available Chymotrypsin Assay Kit (Abcam plc., ab234051). A chymotrypsin activator cleaves chymotrypsinogen to form active chymotrypsin, which then hydrolyzes the non-fluorescent substrate to release a stable fluorophore (Ex/Em 380 nm/4460nm). The fluorescent signal is compared to a coumarin standard curve in order to obtain specific activities of the enzyme samples. Based on calculations, the bioconjugate sample subjected to reaction conditions retained 78% of the wildtype specific activity.

Data	**Wildtype**	**Bioconjugate**
t1	15 min	15 min
t2	25 min	25 min
RFU t1	56073.633	31147.707
RFU t2	91789.289	59173.098
RFU background t1	148.451	73.183
RFU background t2	140.488	75.755
∂F = RFU1 – RFU2	35715.656	28025.391
∂F background	0	2.572
∂∂F	35715.656	28022.819
Coumarin curve (m)	6464.5	6464.5
Coumarin curve (b)	1416.8	1416.8
∂B (pmol)	5305.724495	4115.711811
∂t (min)	10	10
concentration (mmol)	0.1	0.1
MW	25657	25657
concentration (µg/mL)	2565.7	2565.7
amount/well (mg)	25.657	25.657
Chymo specific acitivity (pmol/min/mg)	20.7	16.0
% activity of WT	100%	78%

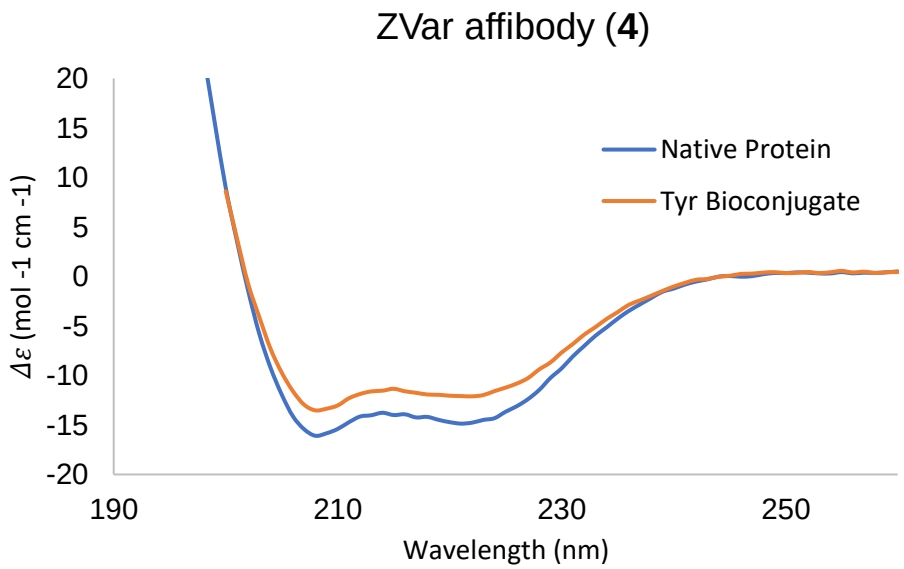



Chymotrypsinogen A activity: wildtype vs. tyrosine bioconjugate
RFU (Ex/Em 380/460 nm)
125,000
100,000
75,000
50,000
25,000
0
-25,000
0
20
40
60
80
100
120
Time (min)
Wildtype Chymotrypsinogen A
Tyrosine Bioconjugate
Reagent background control
Bioconjugate background

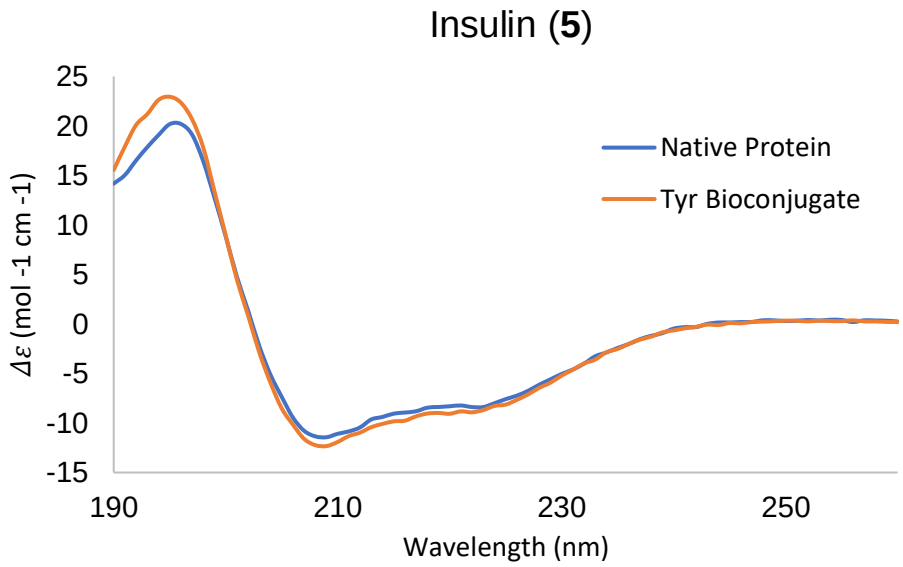
XIV. CD spectra

Circular dichroism (CD) spectroscopy was performed on all protein substrates (**4**–**12**) using a Chirascan CD spectrometer (Applied Photophysics). Wildtype or purified bioconjugated proteins are dissolved and diluted in 10 mM potassium phosphate buffer (final concentration 0.05 or 0.1 mg/mL) and were placed into a 0.1 mm pathlength quartz cuvette (Hellma Analytics, Germany). 10 scans were made from 260 to 190 nm at a scan speed of 65 nm/min and 0.5 nm data pitch. The slit width was 1 nm. The dynode voltage never exceeded 0.80 kV. Buffer spectra were subtracted to compare the secondary structure of bioconjugated samples to that of the wildtype. Deconvolution of the CD spectrum to gain the fractional content of secondary structure elements was done using DichroWeb (Whitmore, L. & Wallace, B.A. *Biopolymers* **89**, 392–400 (2008)). This tool gives the fractional secondary structure content from 3 algorithms, CONTINLL, SELCON, and CDSSTR and calculates the mean output (MeanCDPro).

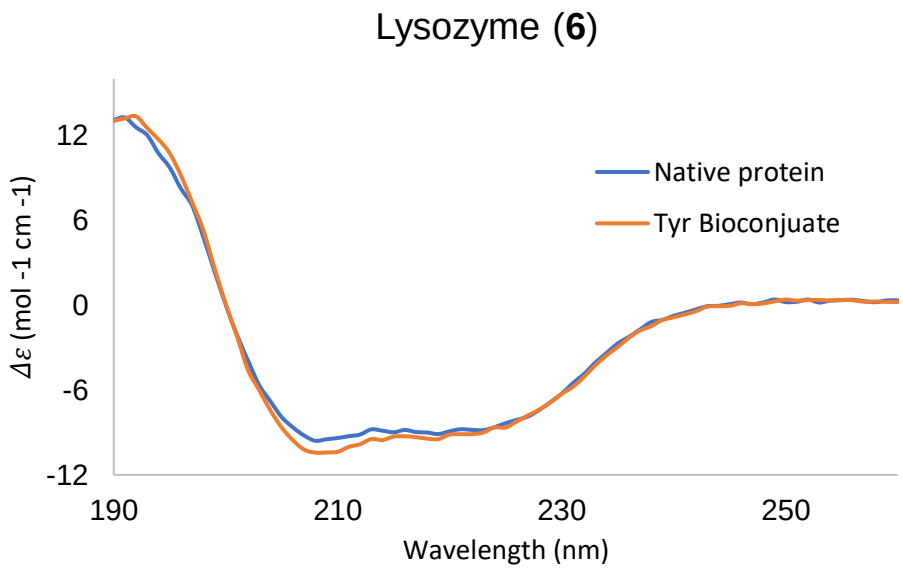
*a. ZVar affibody (**4**)*



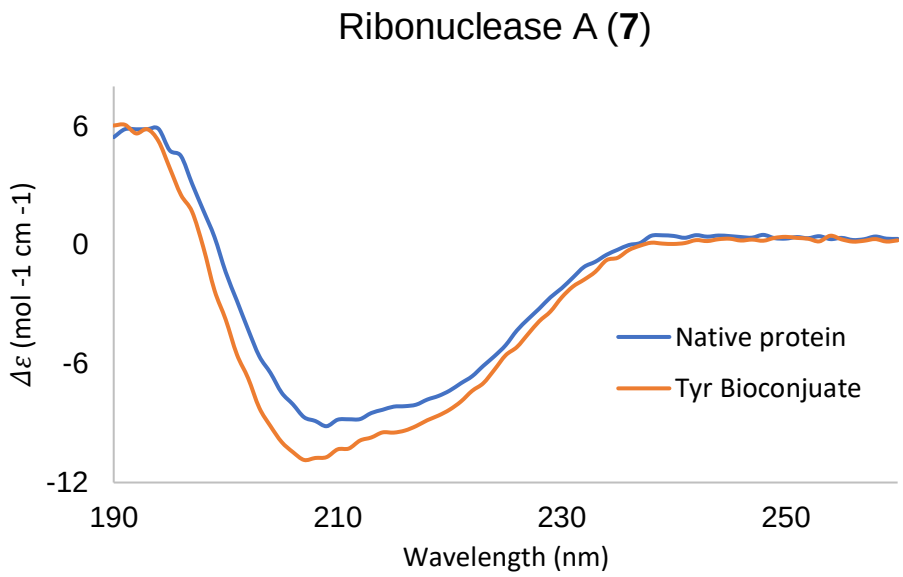
*b. Human insulin (**5**)*



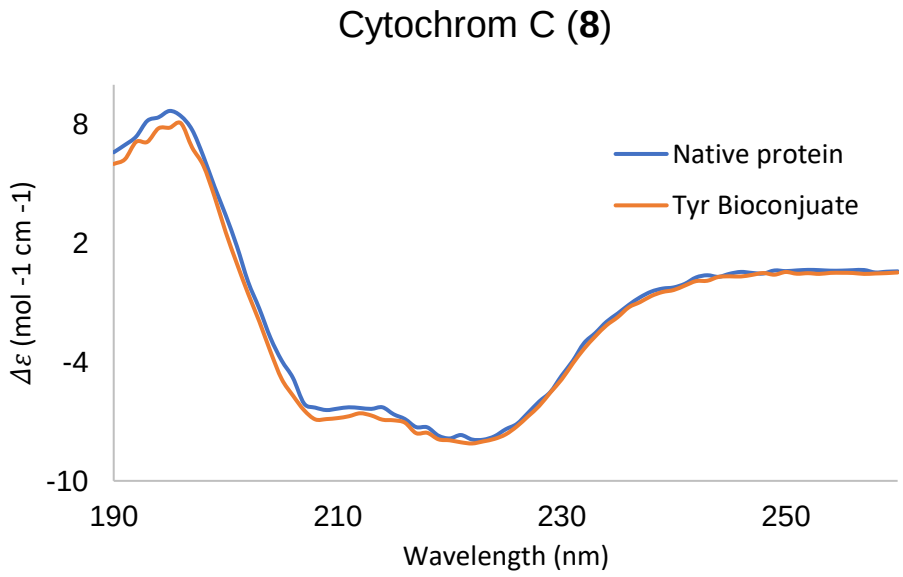
*c. Human lysozyme (**6**)*



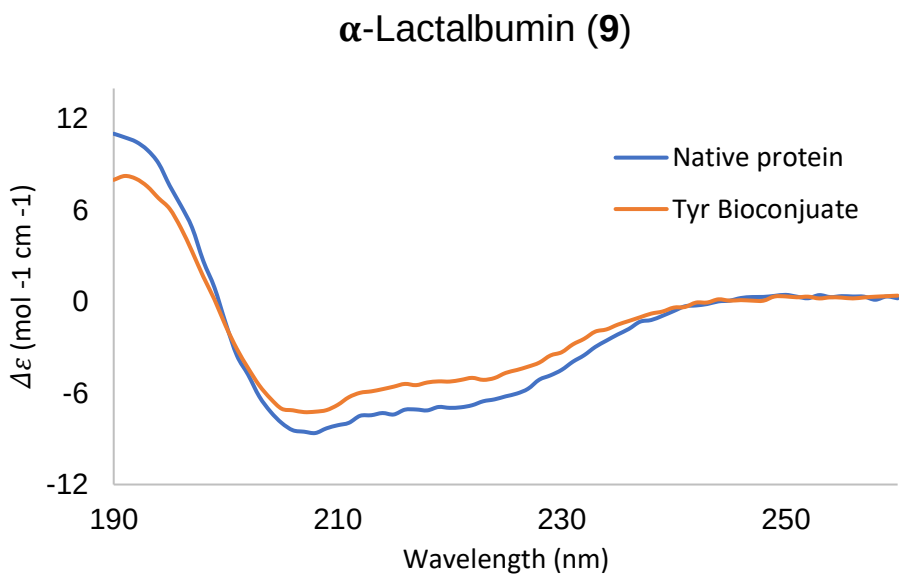
*d. Ribonuclease A (**7**)*



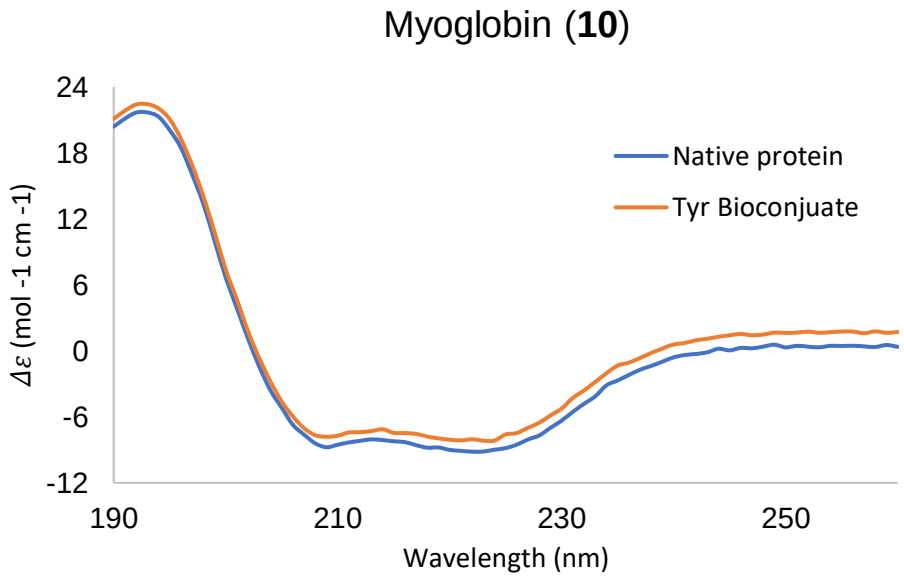
*e. Cytochrome C (**8**)*



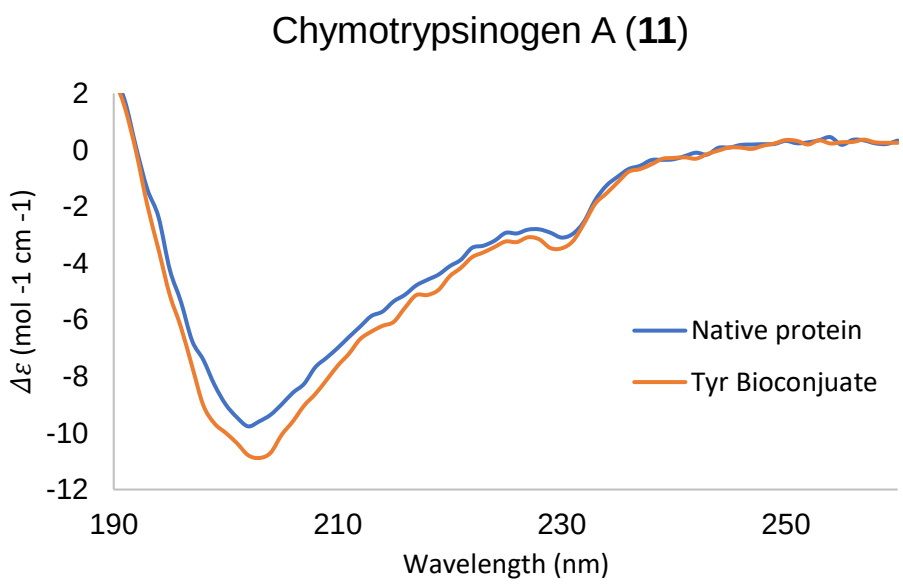
*f. α-Lactalbumin (**9**)*



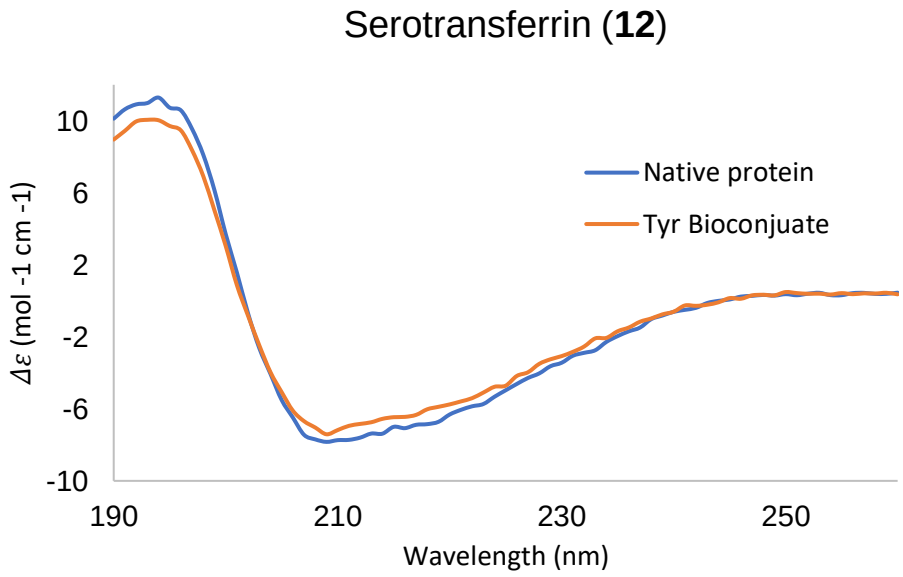
*g. Myoglobin (**10**)*



*h. Chymotrypsinogen A (**11**)*



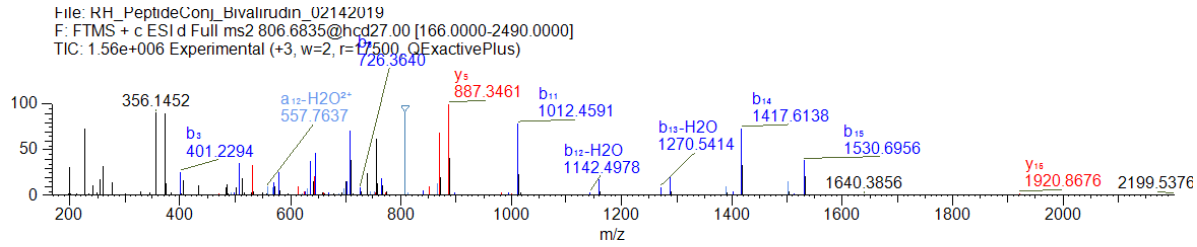
*i. Serotransferrin (**12**)*



XV. LC-MS/MS spectra

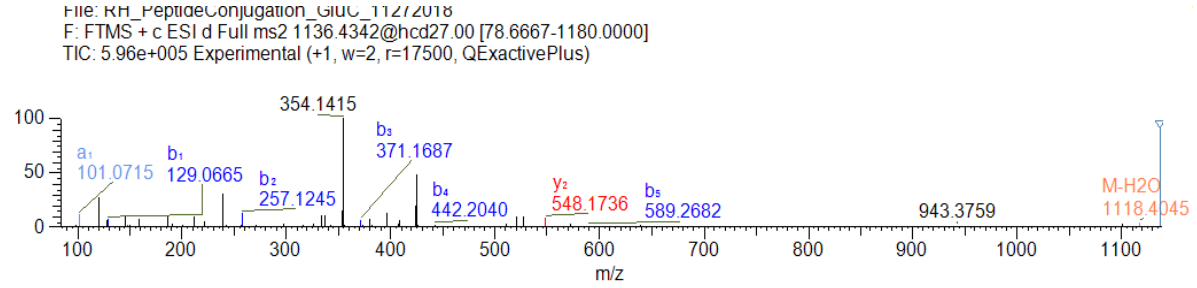
*a. Bivalirudin (**3**)*

F P R P G G G G N G D F Q Q I P Q Q **Y[19]** L



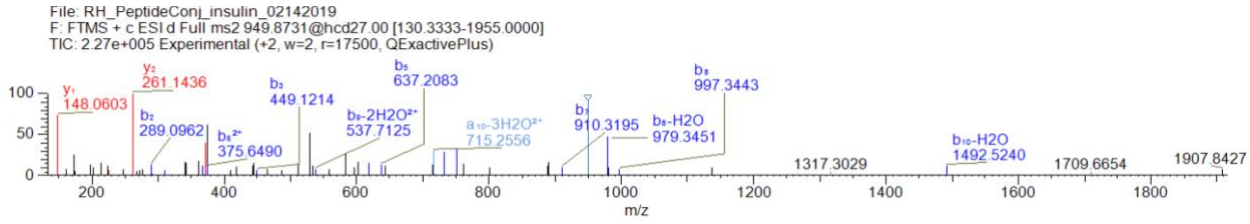
*b. ZVar affibody (**4**)*

Q Q N A F **Y[14]** E

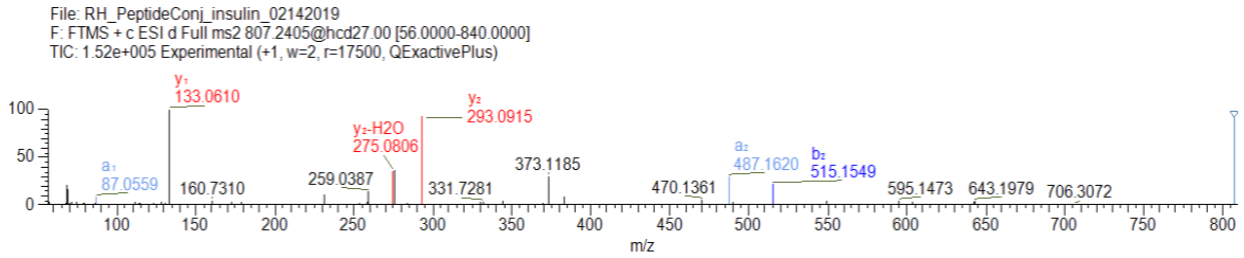


*c. Human insulin (**5**)*

Q C C T S I C S L **Y[14]** Q L E

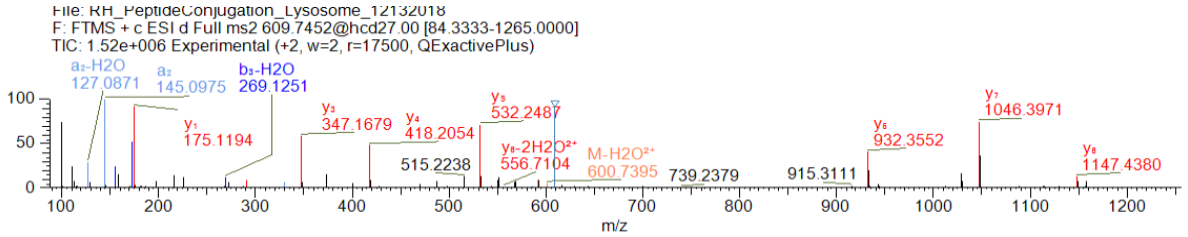


N **Y[19]** C N



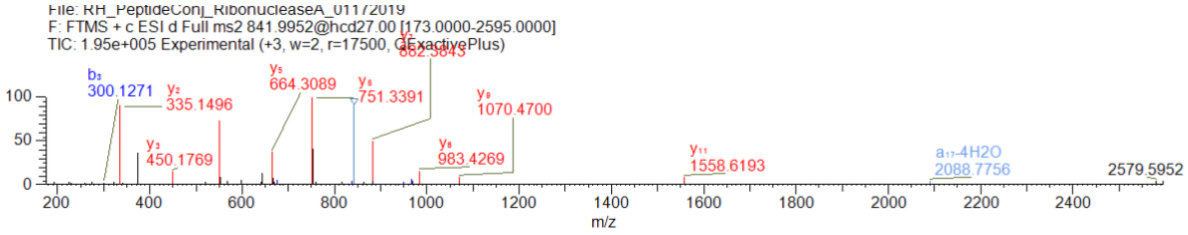
*d. Human lysozyme (**6**) and derivatives*

A T N **Y[45]** N A G D R



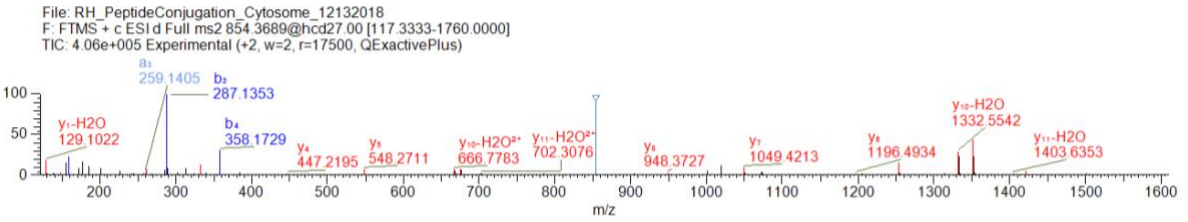
e. Ribonuclease A (7)

N G Q T N C **Y** Q S **Y**[76] S T M S I T D C R

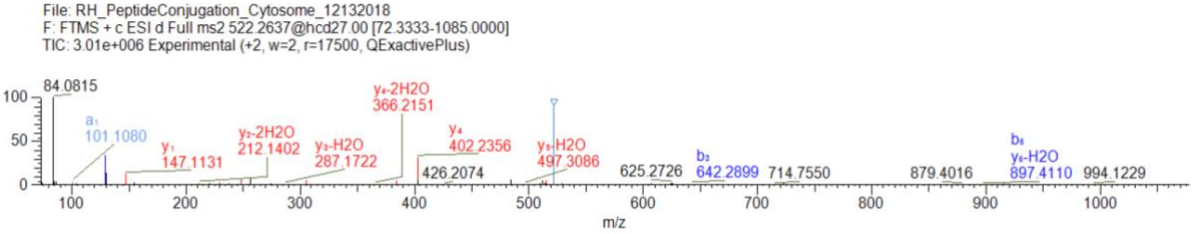


*f. Cytochrome C (**8**)*

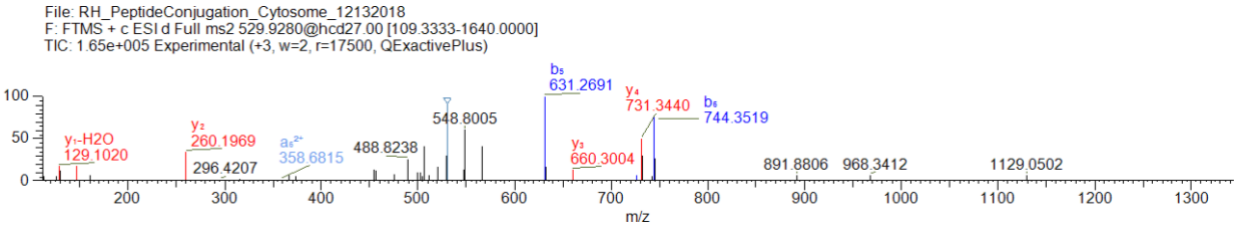
G Q A P G F S **Y**[48] T D A N K



K **Y**[74] I P G T K

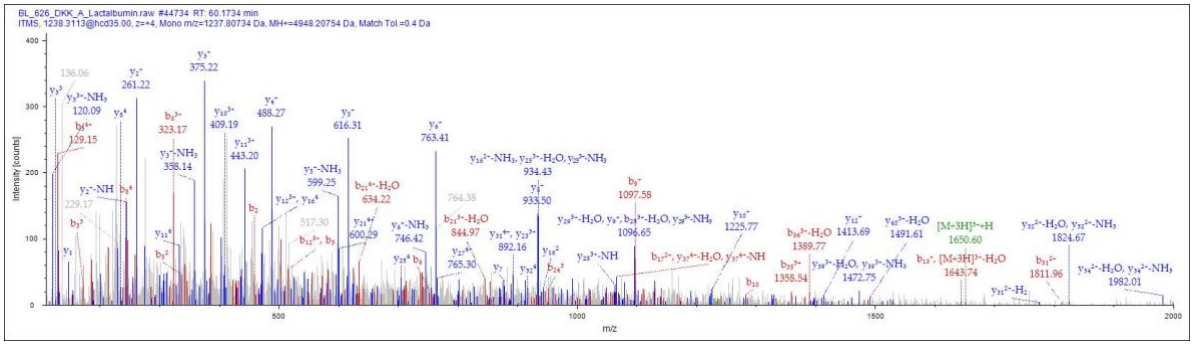


T E R E D L I A **Y**[97] L K

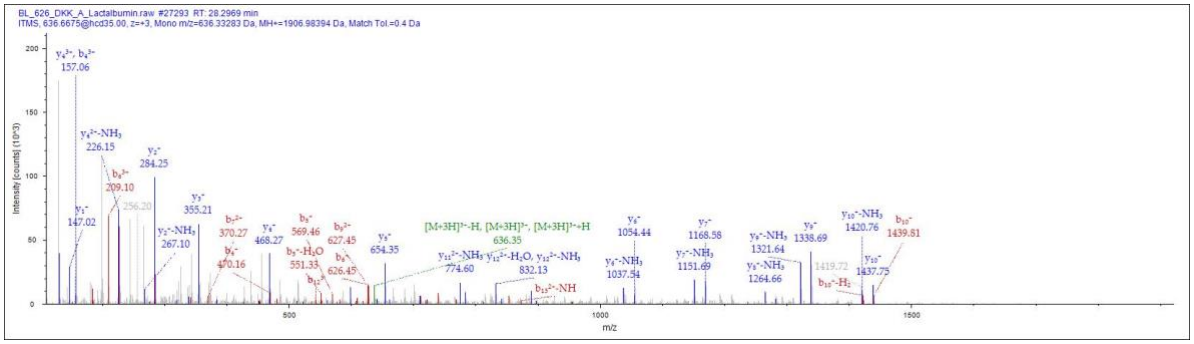


*g. α-Lactalbumin (**9**)*

G **Y**[18] G G V S L P E W V C T T F H T S G **Y** D T Q A I V Q N N D S T E **Y** G L F Q I N N K

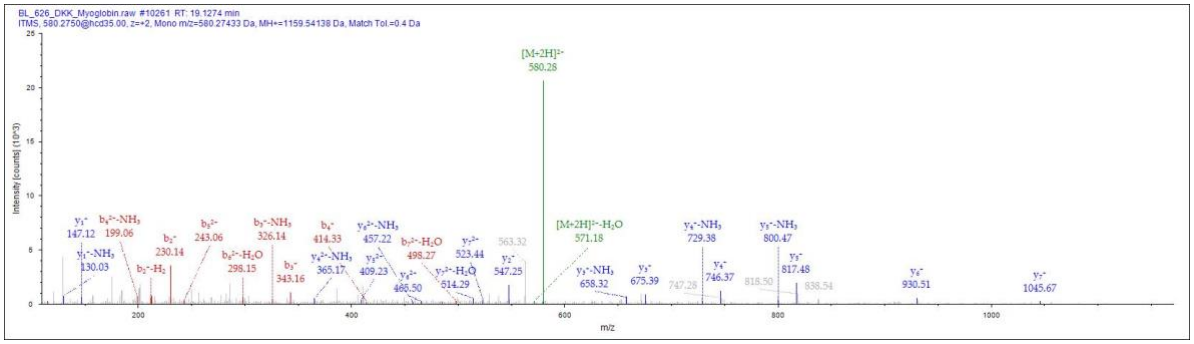


I L D K V G I N **Y[103]** W L A H K



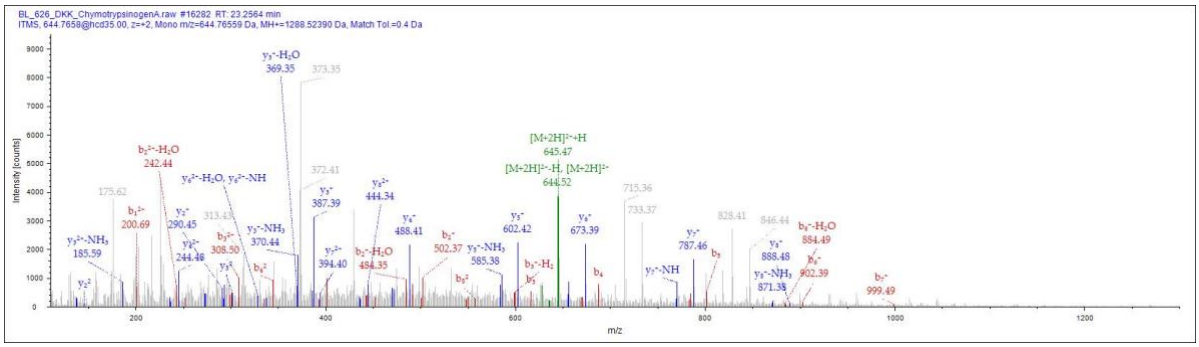
*h. Myoglobin (**10**)*

N D I A A K **Y[146]** K

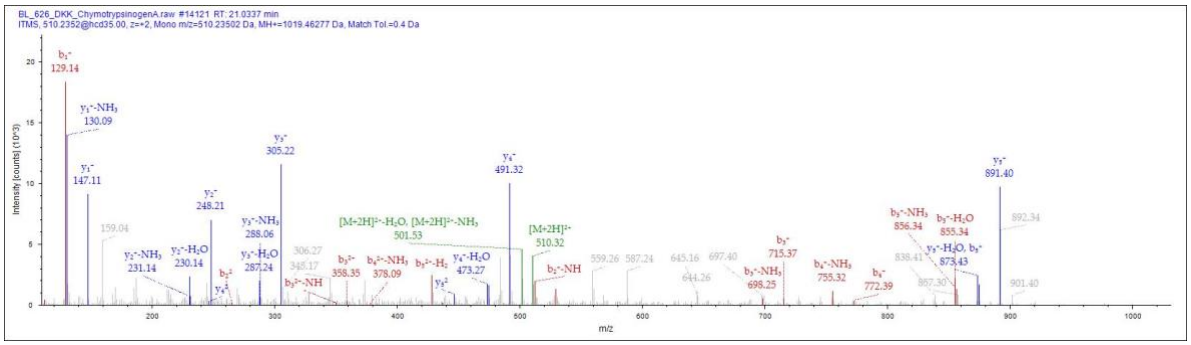


*i. Chymotrypsinogen A (**11**)*

Y[146] T N A N T P D R

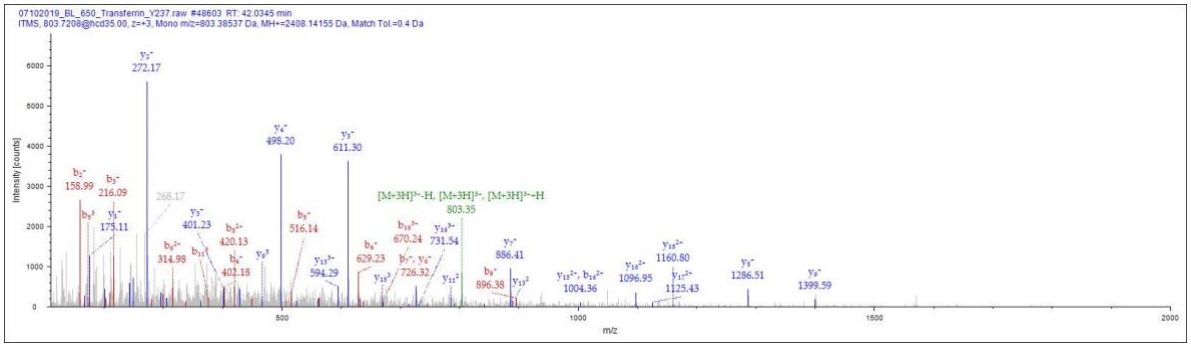


K K **Y[171]** W G T K



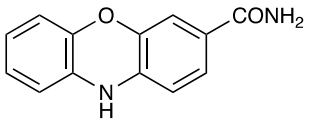
*j. Serotransferrin (**12**)*

S A G W N I P I G L L **Y[136]** C D L P E P R

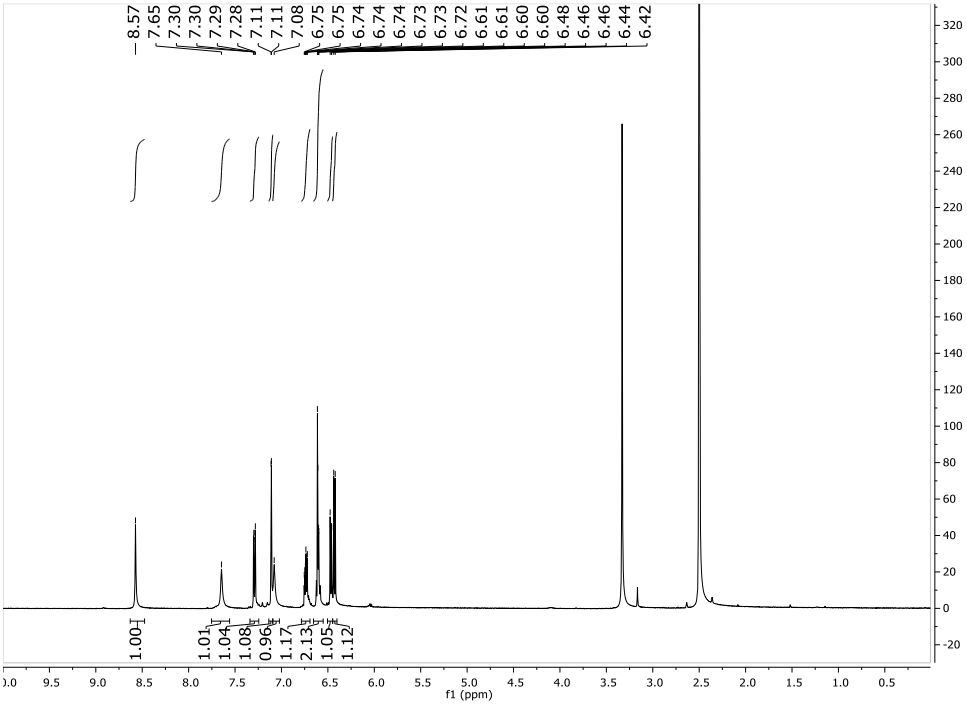


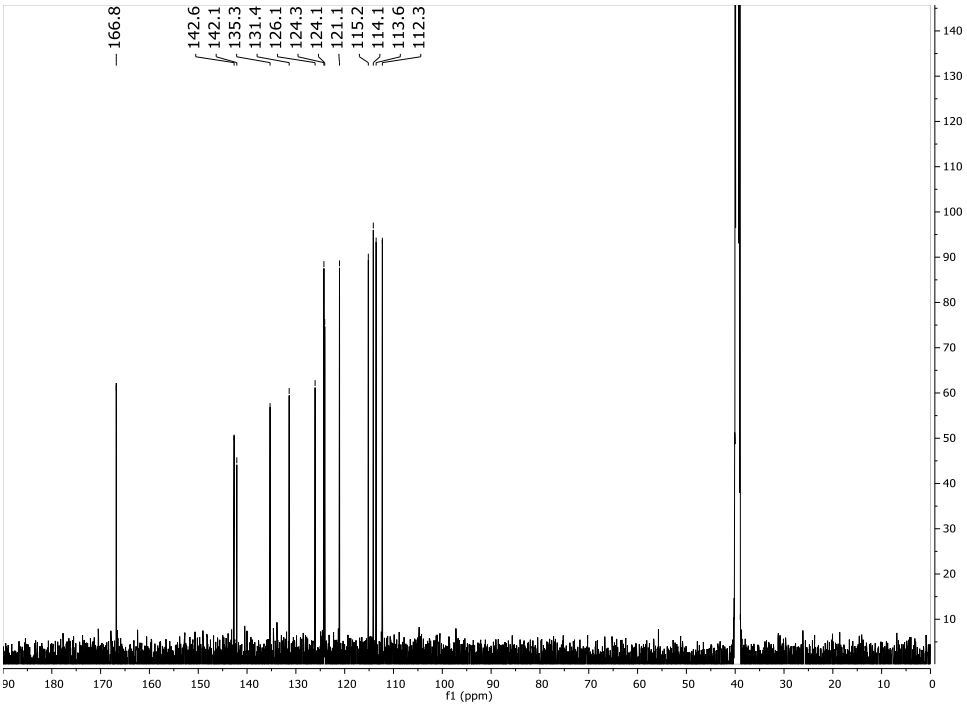
XVI. NMR Spectra

*a. Phenoxazine monoamide **S1***

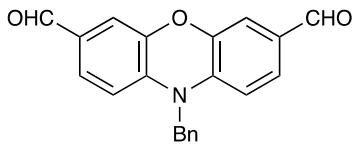


phenoxazine monoamide **S1**

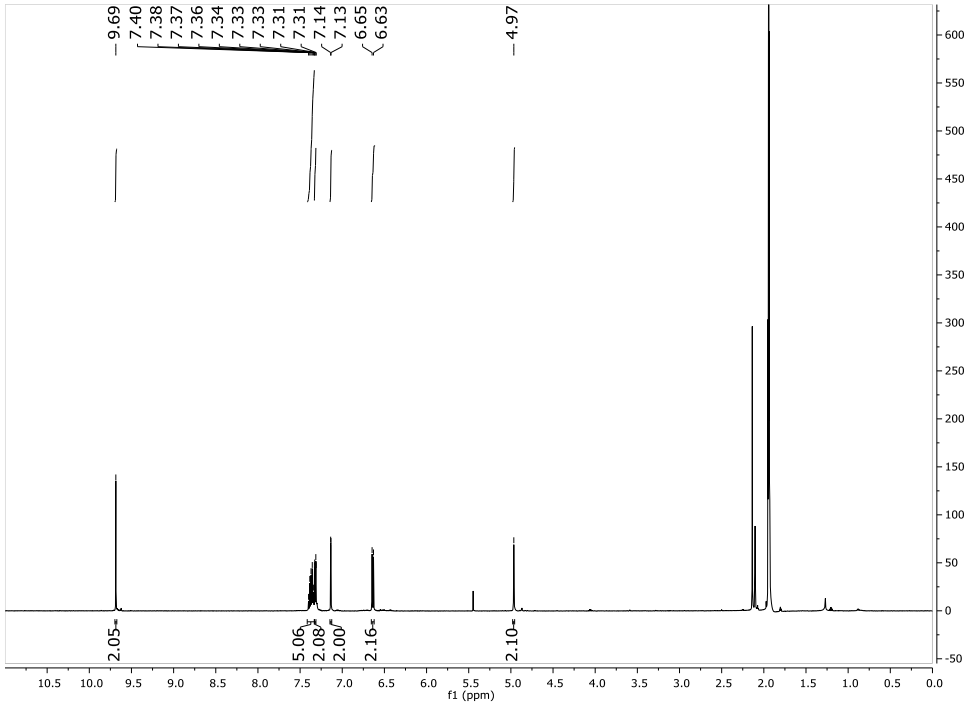


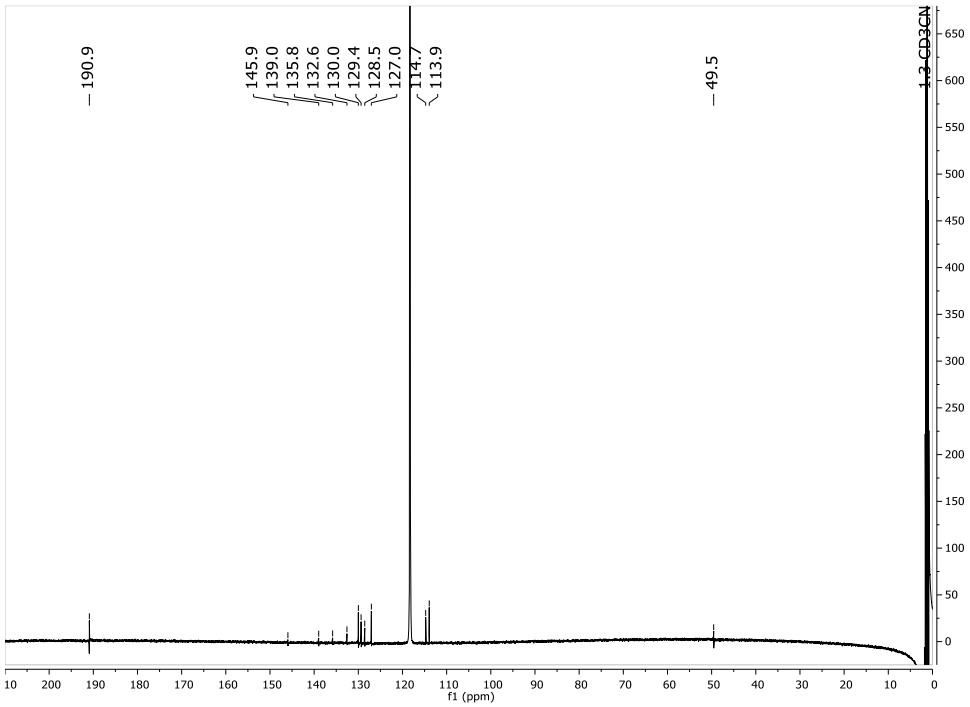


*b. Benzylated phenoxazine dialdehyde **S4***

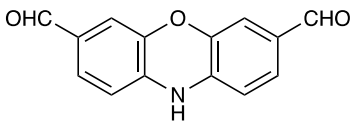


S4

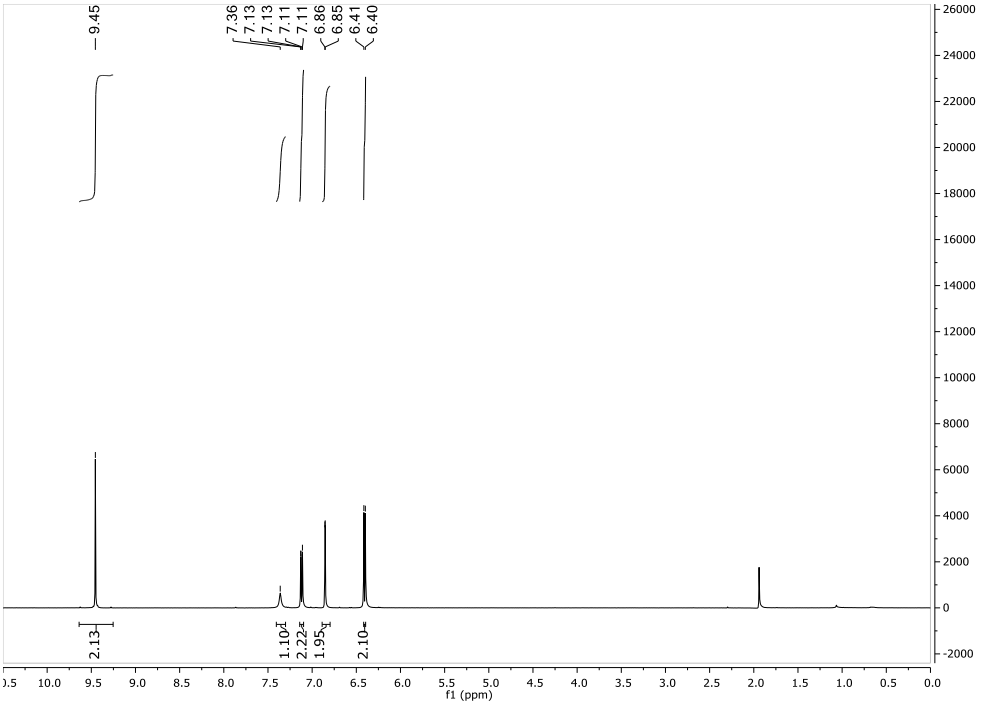


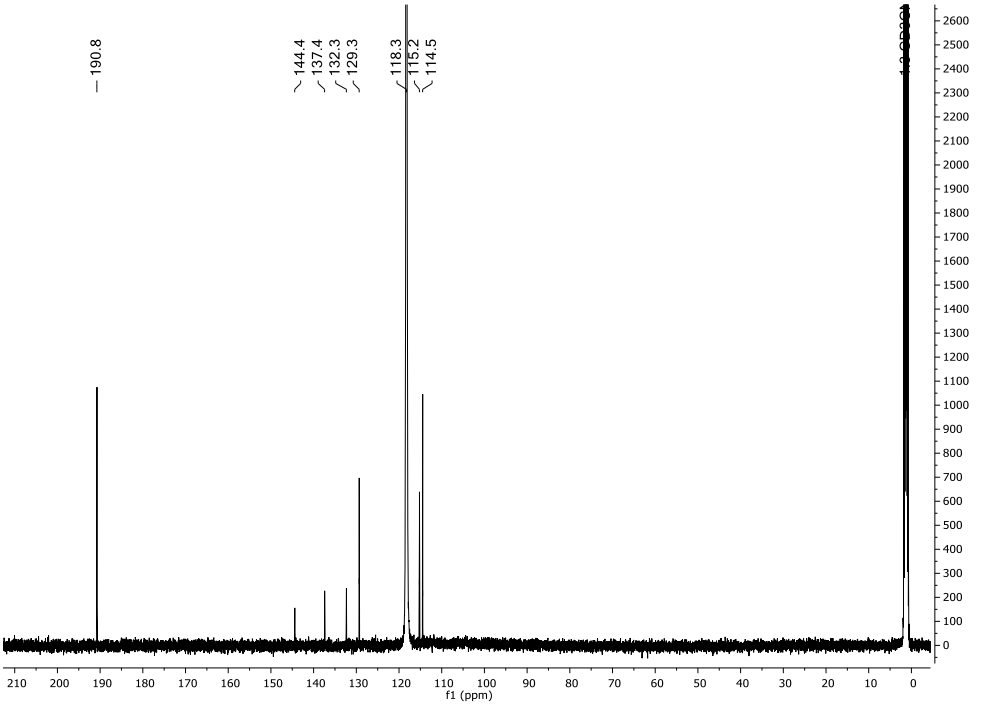


*c. Phenoxazine dialdehyde **1***

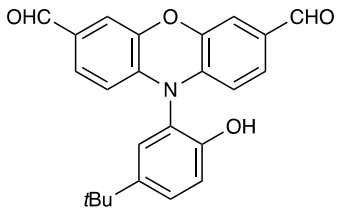


phenoxazine dialdehyde **1**





*d. Fluorescence standard **S2***



Fluorescence standard **S2**