

## Supplementary Information

### Engineered unnatural ubiquitin for optimal detection of deubiquitinating enzymes

Wioletta Rut<sup>1\*</sup>, Mikolaj Zmudzinski<sup>1</sup>, Scott J. Snipas<sup>2</sup>, Miklos Bekes<sup>3,4</sup>, Tony T. Huang<sup>3</sup>, Marcin Drag<sup>1,2\*</sup>

<sup>1</sup>Department of Chemical Biology and Bioimaging, Wrocław University of Science and Technology, Wyb. Wyspińskiego 27, 50-370 Wrocław, Poland

<sup>2</sup>Sanford Burnham Prebys Medical Discovery Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA

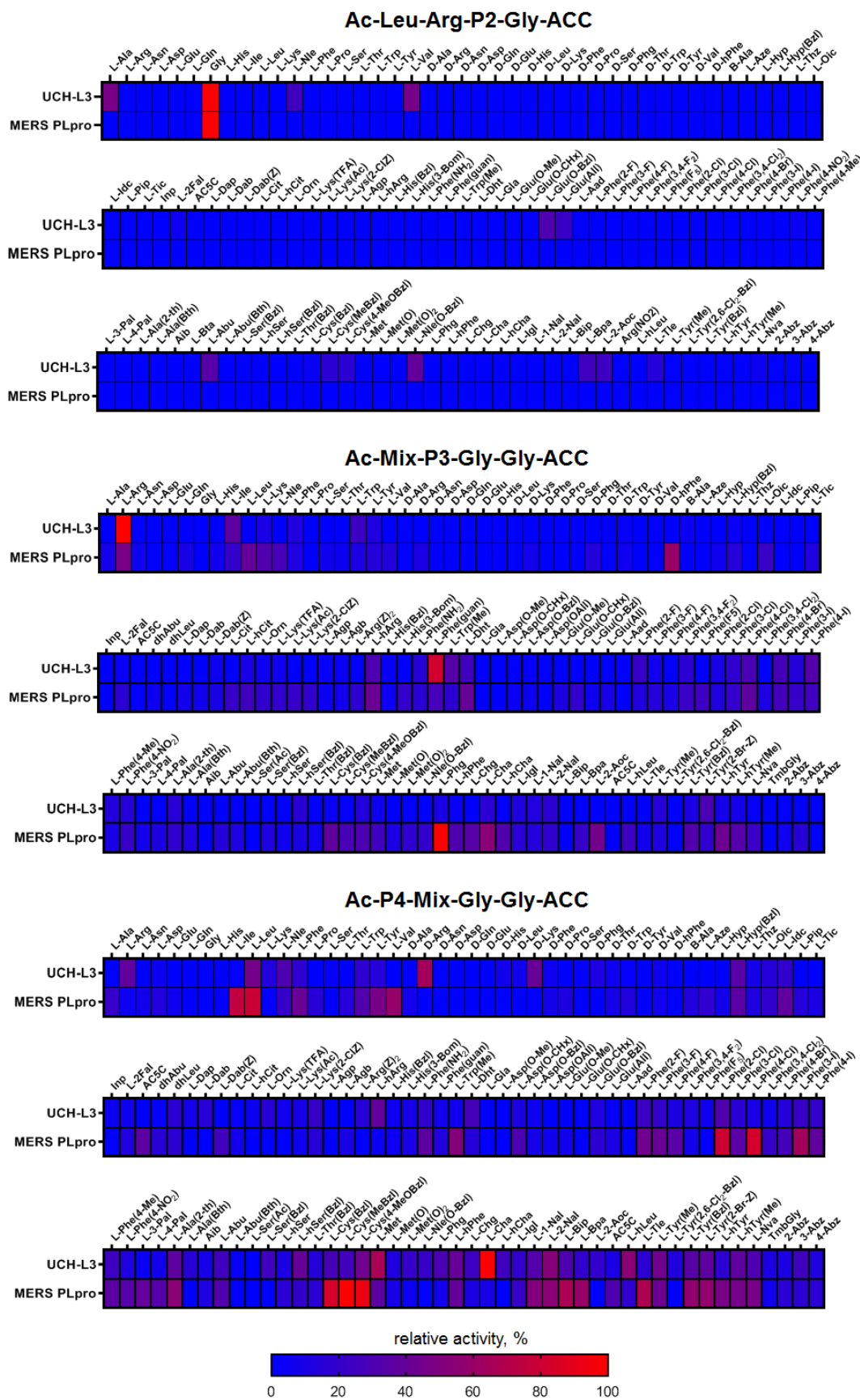
<sup>3</sup>Department of Biochemistry & Molecular Pharmacology, New York University School of Medicine, New York, NY 10016, USA

<sup>4</sup>present address: Arvinas, Inc., 5 Science Park, New Haven, CT, 06511, USA

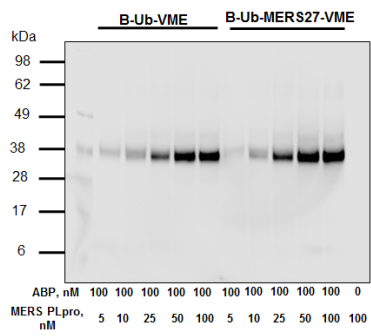
\*corresponding authors: wioletta.rut@pwr.edu.pl, marcin.drag@pwr.edu.pl

#### Content of Supporting Information

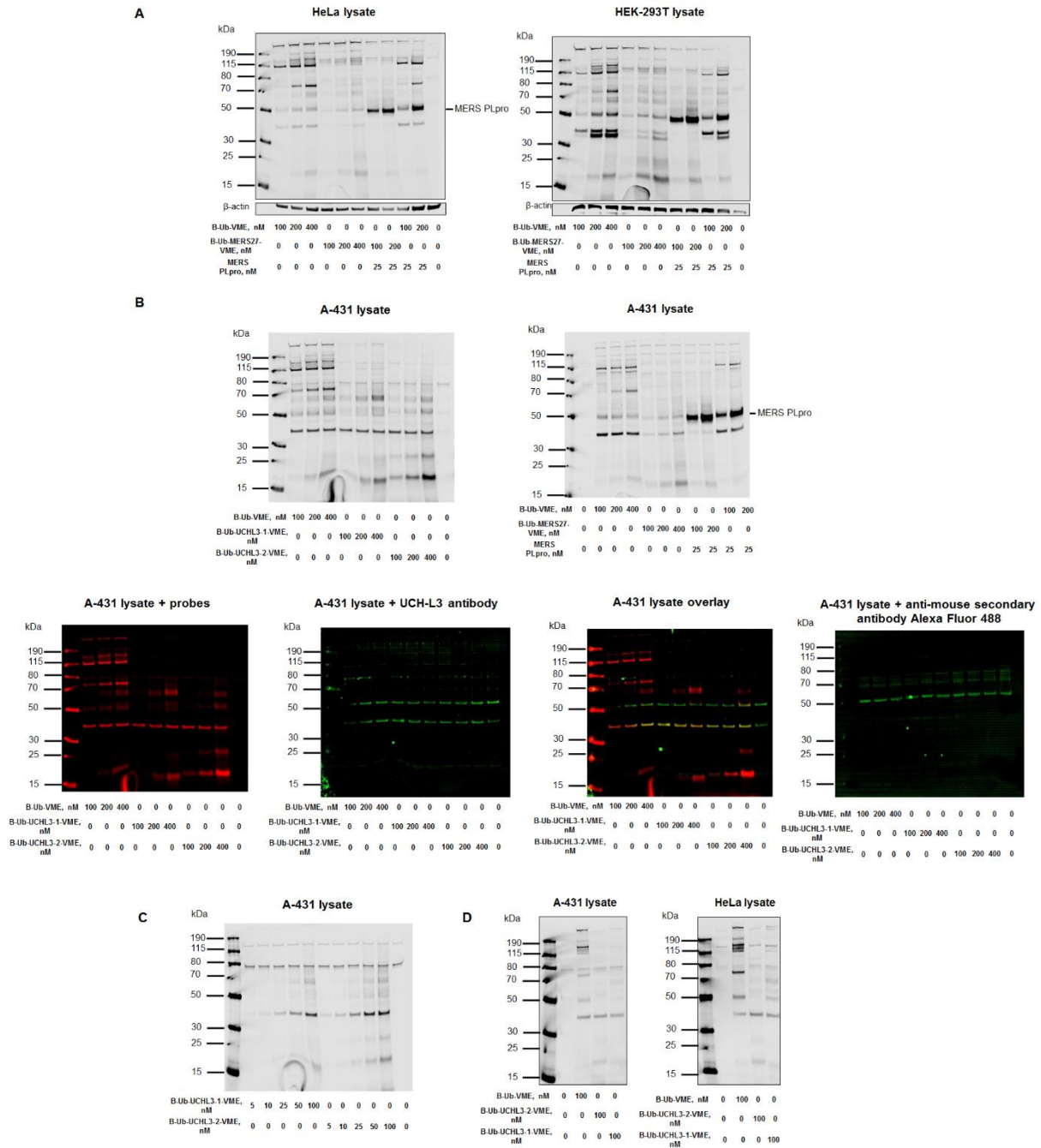
1. Supplementary Figure.....	S2-S4
2. Supplementary Tables.....	S5-S15
3. Analysis data of synthesized substrates and activity-based probes.....	S16-S40



**Figure S1.** DUBs substrate specificity profiles presented as heat maps.



**Figure S2.** MERS PLpro labelling by Ub-based probes. Ub-based probes (100 nM) were incubated with different enzyme concentrations for 30 min at 37°C. Then SDS-PAGE analysis was performed.



**Figure S3. Ub-based probe selectivity.**

(A) DUB labeling in cell lysates by B-Ub-VME and B-Ub-MERS27-VME (B) DUB labeling in cell lysates by B-Ub-UCHL3-1-VME, B-Ub-UCHL3-2-VME and anti-UCH-L3 antibody. The second band at ca. 50 kDa which is presented on the membrane “A431-lysate + UCH-L3 antibody” is unspecific labelling of protein from lysate by secondary antibody (goat anti-mouse IgG secondary antibody Alexa Fluor 488). This unspecific labelling was confirmed by re-running the gel and incubating membrane only with anti-mouse secondary antibody Alexa Fluor 488 (without anti-UCH-L3 antibody). (C) Dose-dependent labelling of UCH-L3 in A-431 lysate by Ub-probes. (D) Pull-down experiment using B-Ub-VME, B-Ub-UCHL3-1-VME and B-Ub-UCHL3-2-VME. A-431 and HeLa lysates were incubated with or without probes for 30 min at 37°C. Then streptavidin beads were added and samples were incubated for 2.5 h at 4°C. Samples were analyzed by gel electrophoresis, followed by western blotting.

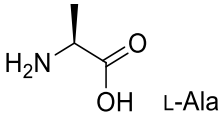
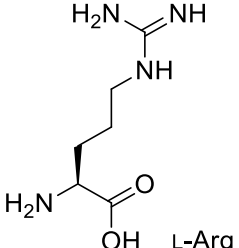
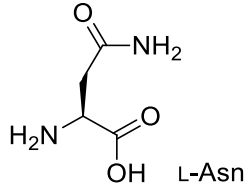
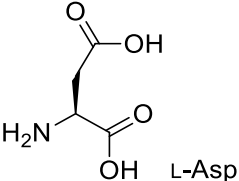
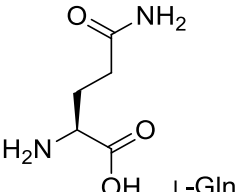
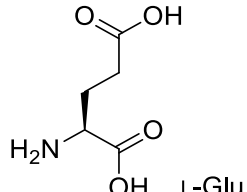
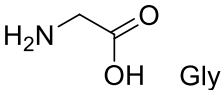
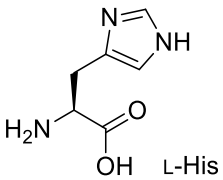
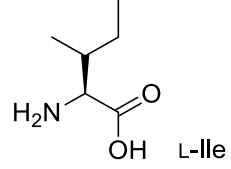
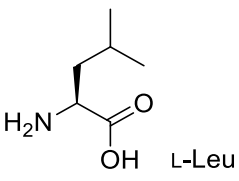
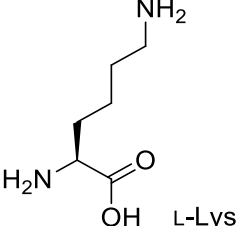
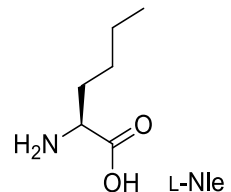
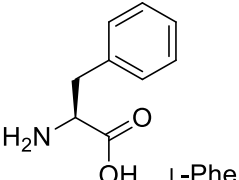
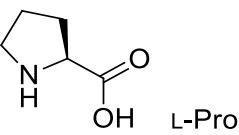
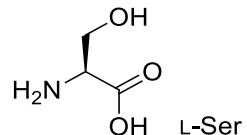
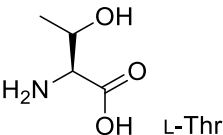
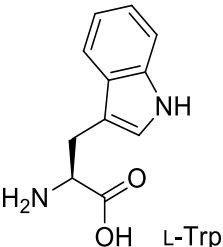
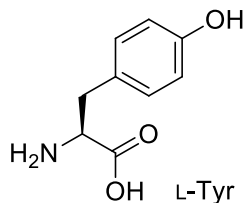
**Table S1.** Calculated and found m/z for Ac-Leu-Arg-P2-Gly-ACC substrate library.

<b>Substrate</b>	<b>Ac-Leu-Arg-P2-Gly-ACC</b>	<b>m/z<sub>calcd</sub></b>	<b>m/z<sub>found</sub></b>
P2/1	L-Ala	658.3308	658.28
P2/2	L-Arg	743.3948	743.37
P2/3	L-Asn	701.3366	701.30
P2/4	L-Asp	702.3206	702.28
P2/5	L-Glu	716.3363	716.30
P2/6	L-Gln	715.3522	715.15
P2/7	Gly	644.3151	644.15
P2/8	L-His	724.3526	724.25
P2/9	L-Ile	700.3777	700.28
P2/10	L-Leu	700.3777	700.29
P2/11	L-Lys	715.3886	715.30
P2/12	L-Nle	700.3777	700.20
P2/13	L-Phe	734.3621	734.23
P2/14	L-Pro	684.3464	684.19
P2/15	L-Ser	674.3257	674.19
P2/16	L-Thr	688.3413	688.21
P2/17	L-Trp	773.3730	773.29
P2/18	L-Tyr	750.3570	750.27
P2/19	L-Val	686.3621	686.26
P2/20	D-Ala	658.3308	658.20
P2/21	D-Arg	743.3948	743.29
P2/22	D-Asn	701.3366	701.24
P2/23	D-Asp	702.3206	702.22
P2/24	D-Gln	715.3522	715.25
P2/25	D-Glu	716.3363	716.23
P2/26	D-His	724.3526	724.27
P2/27	D-Leu	700.3777	700.28
P2/28	D-Lys	715.3886	715.29
P2/29	D-Phe	734.3621	734.28
P2/30	D-Pro	684.3464	684.22
P2/31	D-Ser	674.3257	674.21
P2/32	D-Phg	720.3464	720.25
P2/33	D-Thr	688.3413	688.23
P2/34	D-Trp	773.3730	773.31
P2/35	D-Tyr	750.3570	750.30
P2/36	D-Val	686.3621	686.25
P2/37	D-hPhe	748.3777	748.29
P2/38	β-Ala	658.3308	658.21
P2/39	L-Aze	670.3308	670.23
P2/40	L-Hyp	700.3413	700.22
P2/41	L-Hyp(Bzl)	790.3883	790.23
P2/42	L-Thz	702.3029	702.12
P2/43	L-Oic	738.3934	738.23
P2/44	L-Idc	732.3464	732.18
P2/45	L-Pip	698.3621	698.19
P2/46	L-Tic	746.3621	746.20

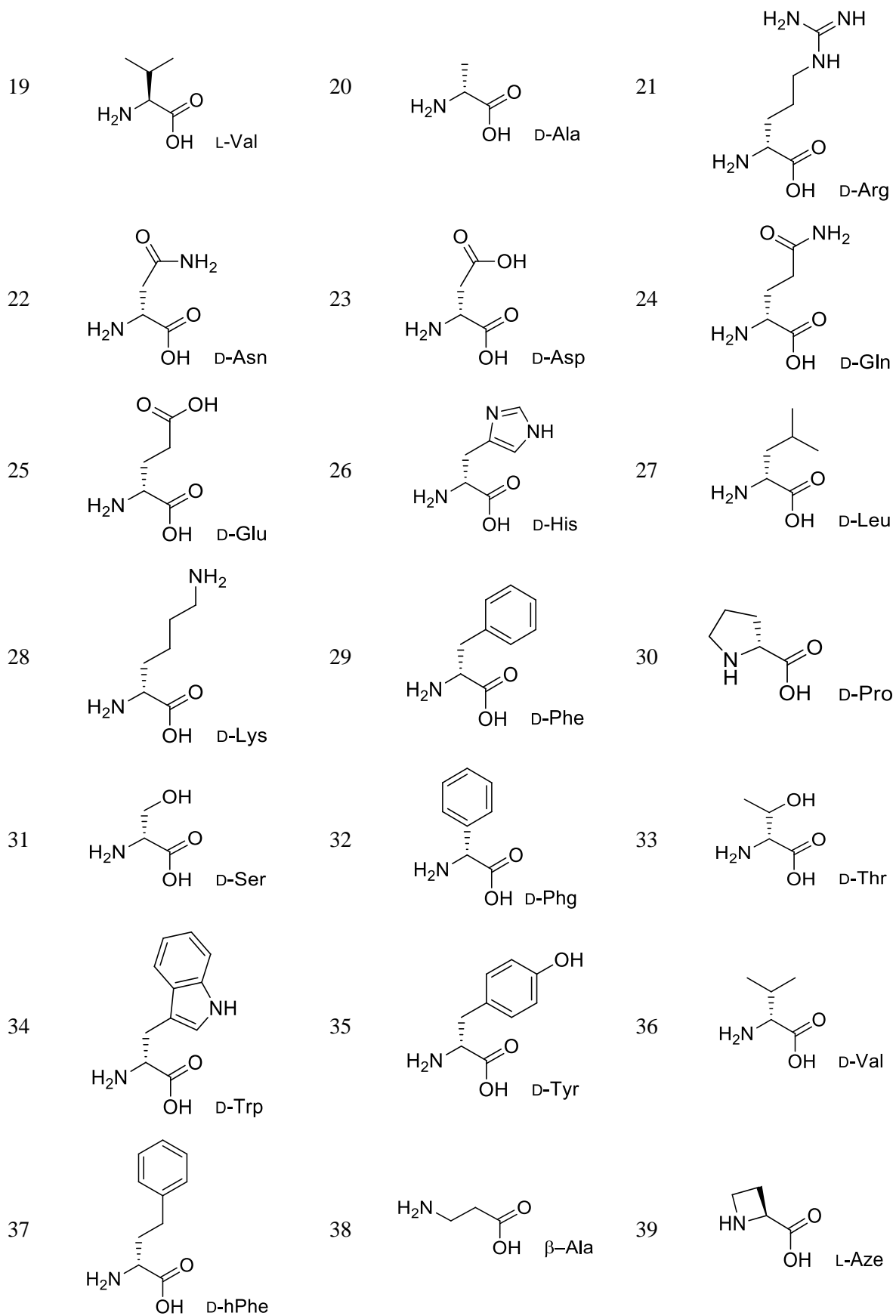
P2/47	Inp	698.3621	698.20
P2/48	L-2Fal	724.3413	724.17
P2/49	AC5C	698.3621	698.19
P2/50	L-Dap	673.3417	673.26
P2/51	L-Dab	687.3573	687.26
P2/52	L-Dab(Z)	821.3941	821.36
P2/53	L-Cit	744.3788	744.32
P2/54	L-hCit	758.3944	758.36
P2/55	L-Orn	701.3730	701.29
P2/56	L-Lys(TFA)	811.3709	811.32
P2/57	L-Lys(Ac)	757.3992	757.28
P2/58	L-Lys(2-ClZ)	883.3864	883.28
P2/59	L-Agp	715.3635	715.24
P2/60	L-hArg	757.4104	757.31
P2/61	L-His(Bzl)	814.3995	814.31
P2/62	L-His(3-Bom)	844.4101	844.34
P2/63	L-Phe(NH <sub>2</sub> )	749.3730	749.29
P2/64	L-Phe(guan)	791.3948	791.30
P2/65	L-Trp(Me)	787.3886	787.29
P2/66	L-Dht	775.3886	775.30
P2/67	L-Gla	760.3261	760.23
P2/68	L-Glu(O-Me)	730.3519	730.25
P2/69	L-Glu(O-cHx)	798.4145	798.34
P2/70	L-Glu(O-Bzl)	806.3832	806.32
P2/71	L-Glu(All)	756.3676	756.31
P2/72	L-Aad	730.3519	730.28
P2/73	L-Phe(2-F)	752.3527	752.29
P2/74	L-Phe(3-F)	752.3527	752.28
P2/75	L-Phe(4-F)	752.3527	752.29
P2/76	L-Phe(3,4-F <sub>2</sub> )	770.3432	770.27
P2/77	L-Phe(F <sub>5</sub> )	824.3150	824.26
P2/78	L-Phe(2-Cl)	768.3231	768.24
P2/79	L-Phe(3-Cl)	768.3231	768.25
P2/80	L-Phe(4-Cl)	768.3231	768.25
P2/81	L-Phe(3,4-Cl <sub>2</sub> )	802.2841	802.31
P2/82	L-Phe(4-Br)	812.2726	812.30
P2/83	L-Phe(3-I)	860.2587	860.31
P2/84	L-Phe(4-I)	860.2587	860.31
P2/85	L-Phe(4-Me)	748.3777	748.40
P2/86	L-Phe(4-NO <sub>2</sub> )	779.3472	779.37
P2/87	L-3-Pal	735.3573	735.35
P2/88	L-4-Pal	735.3573	735.34
P2/89	L-Ala(2-th)	740.3185	740.30
P2/90	L-Ala(Bth)	791.3294	791.35
P2/91	Aib	672.3464	672.30
P2/92	L-Bta	790.3342	790.38
P2/93	L-Abu	672.3464	672.32
P2/94	L-Abu(Bth)	805.3451	805.41

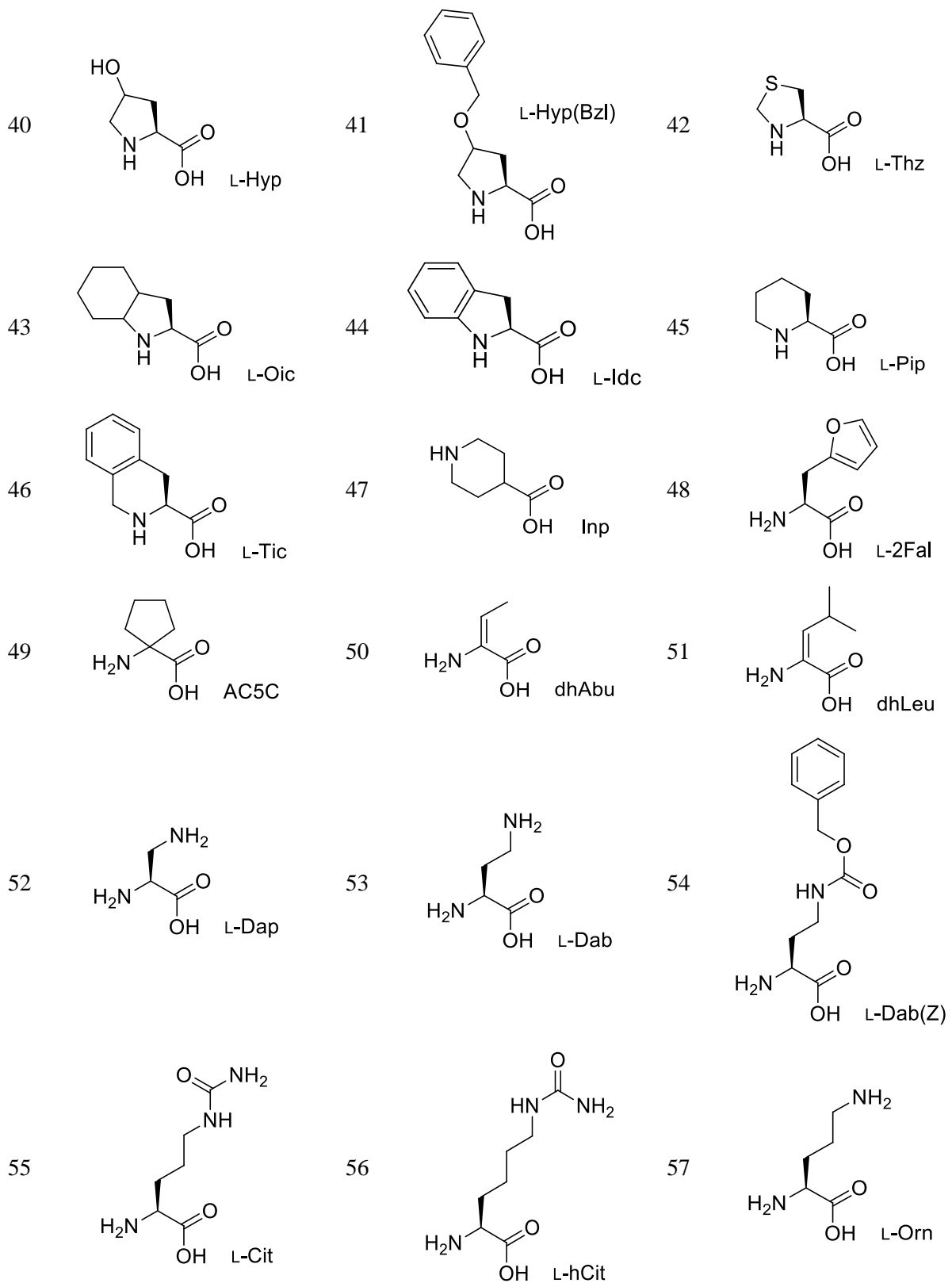
P2/95	L-Ser(Bzl)	764.3276	764.40
P2/96	L-hSer	688.3413	688.31
P2/97	L-hSer(Bzl)	778.3883	778.42
P2/98	L-Thr(Bzl)	778.3883	778.43
P2/99	L-Cys(Bzl)	780.3498	780.38
P2/100	L-Cys(MeBzl)	794.3655	794.38
P2/101	L-Cys(4-MeOBzl)	810.3604	810.33
P2/102	L-Met	718.3342	718.27
P2/103	L-Met(O)	734.3291	734.29
P2/104	L-Met(O) <sub>2</sub>	750.3240	750.29
P2/105	L-Nle(O-Bzl)	806.4196	806.40
P2/106	L-Phg	720.3464	720.34
P2/107	L-hPhe	748.3777	748.39
P2/108	L-Chg	726.3934	726.39
P2/109	L-Cha	740.4090	740.41
P2/110	L-hCha	754.4247	754.44
P2/111	L-Igl	760.3777	760.39
P2/112	L-1-Nal	784.3777	784.42
P2/113	L-2-Nal	784.3777	784.42
P2/114	L-Bip	810.3934	810.44
P2/115	L-Bpa	838.3883	838.44
P2/116	L-2-Aoc	728.4090	728.40
P2/117	L-Arg(NO <sub>2</sub> )	788.3799	788.40
P2/118	L-hLeu	714.3934	714.37
P2/119	L-Tle	700.3777	700.36
P2/120	L-Tyr(Me)	764.3726	764.43
P2/121	L-Tyr(2,6-Cl <sub>2</sub> -Bzl)	908.3260	908.41
P2/122	L-Tyr(Bzl)	840.4039	840.48
P2/123	L-hTyr	764.3726	764.41
P2/124	L-hTyr(Me)	778.3883	778.43
P2/125	L-Nva	686.3621	686.33
P2/126	2-Abz	706.3308	706.33
P2/127	3-Abz	706.3308	706.31
P2/128	4-Abz	706.3308	706.30

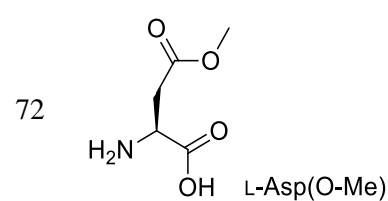
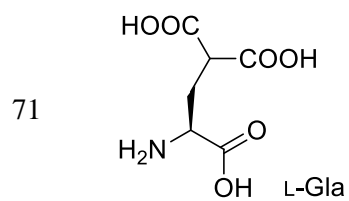
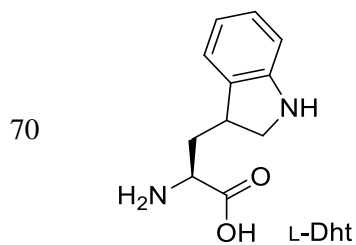
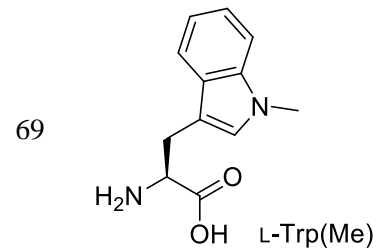
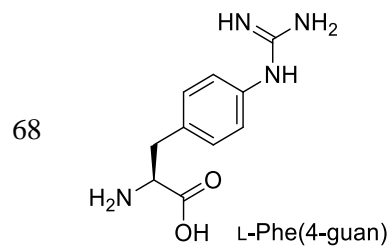
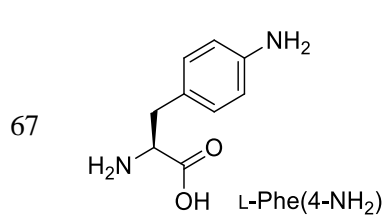
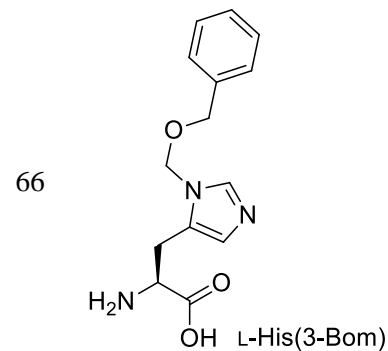
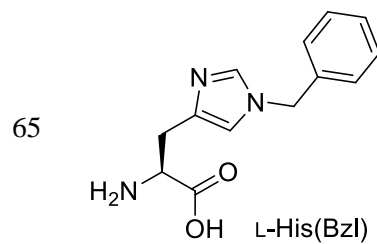
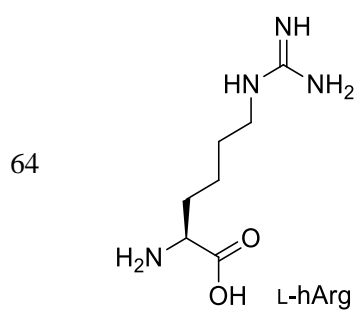
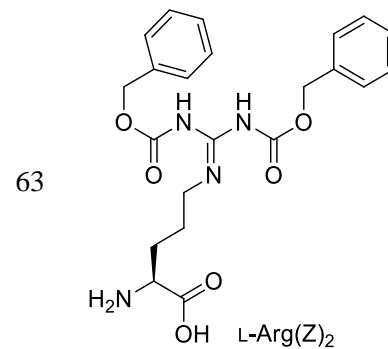
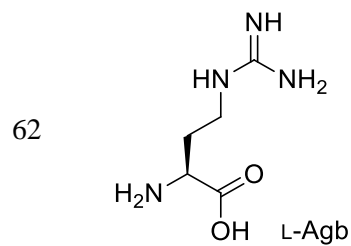
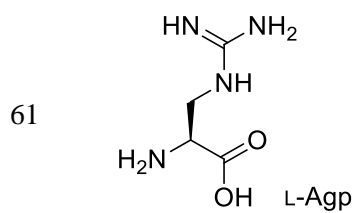
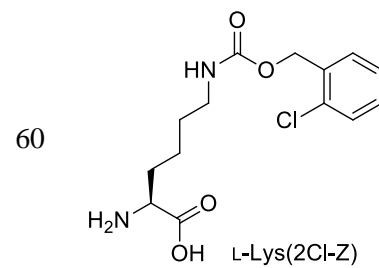
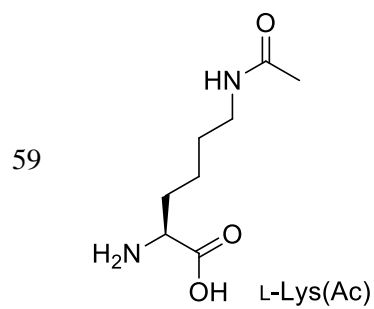
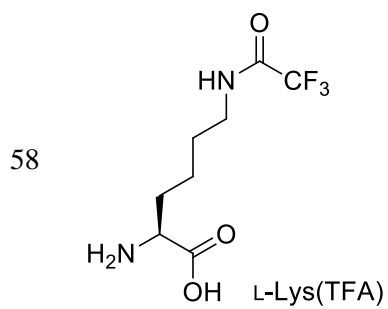
**Table S2.** Structures of fixed natural and unnatural amino acids used in combinatorial library.

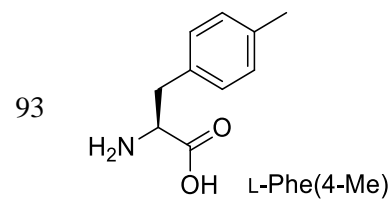
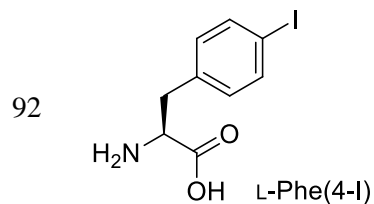
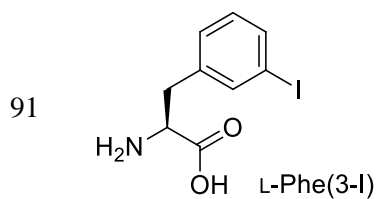
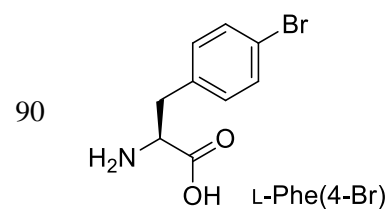
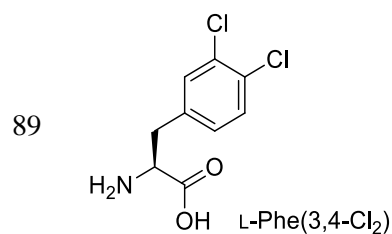
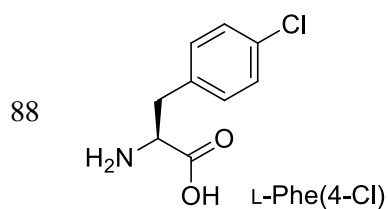
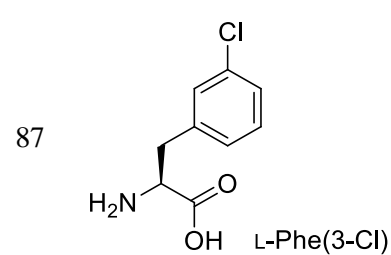
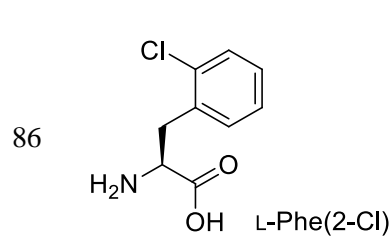
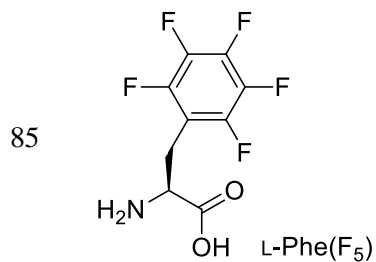
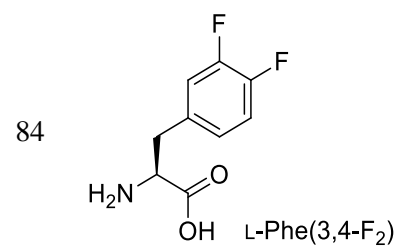
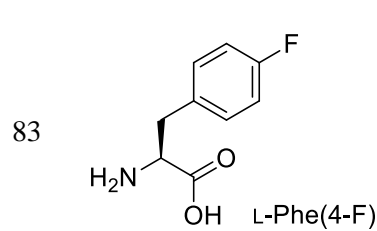
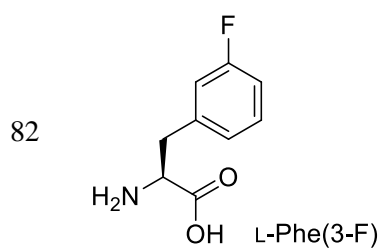
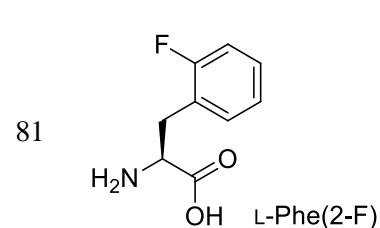
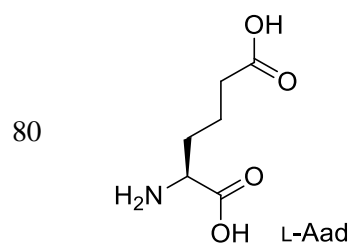
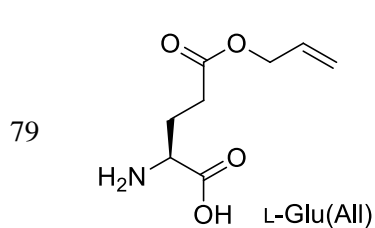
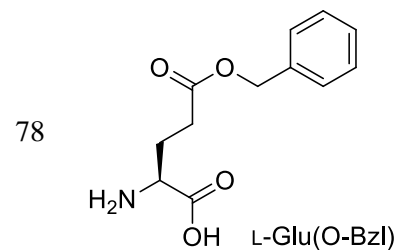
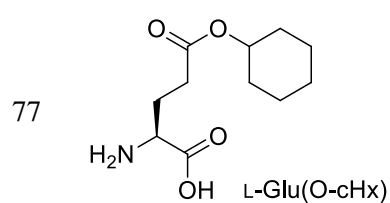
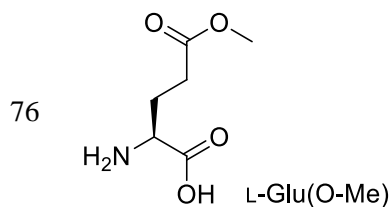
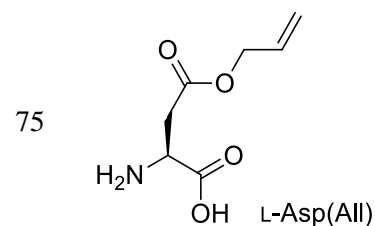
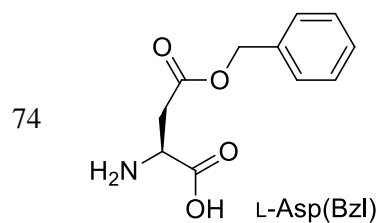
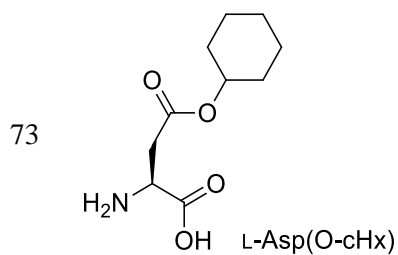
No	Structure + code	No	Structure + code	No	Structure + code
1	 L-Ala	2	 L-Arg	3	 L-Asn
4	 L-Asp	5	 L-Gln	6	 L-Glu
7	 Gly	8	 L-His	9	 L-Ile
10	 L-Leu	11	 L-Lys	12	 L-Nle
13	 L-Phe	14	 L-Pro	15	 L-Ser
16	 L-Thr	17	 L-Trp	18	 L-Tyr

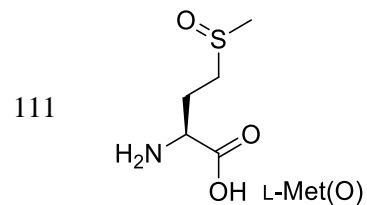
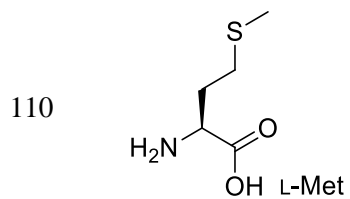
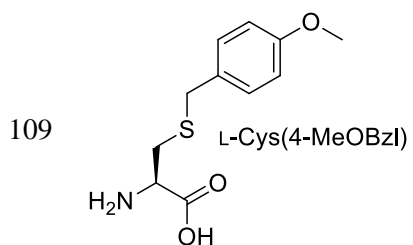
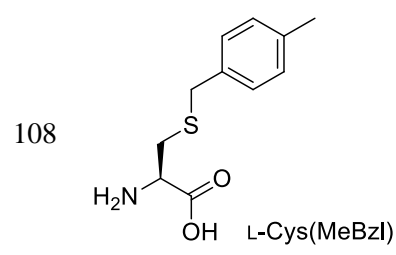
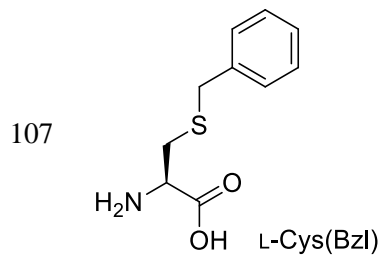
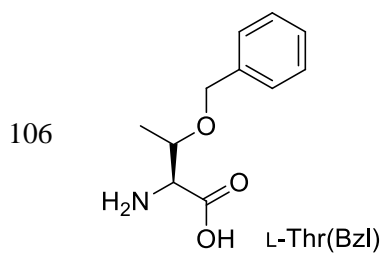
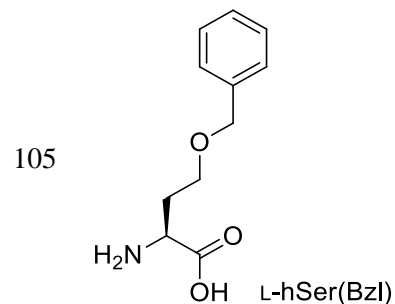
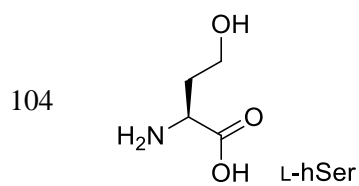
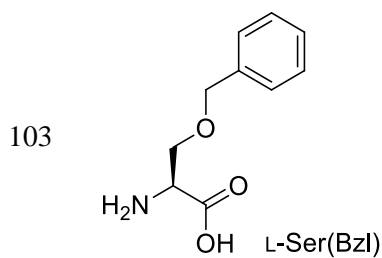
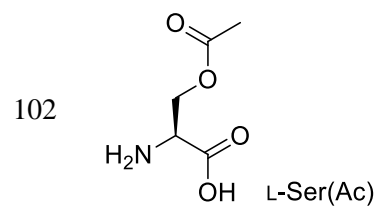
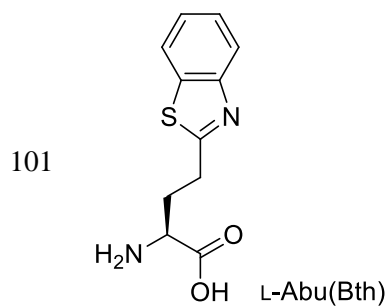
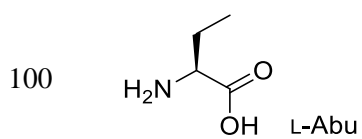
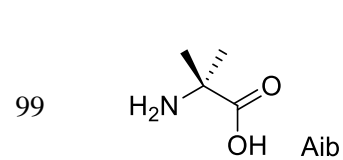
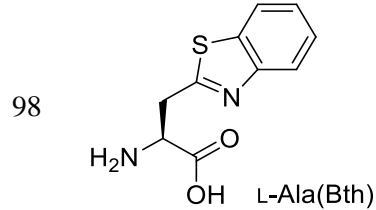
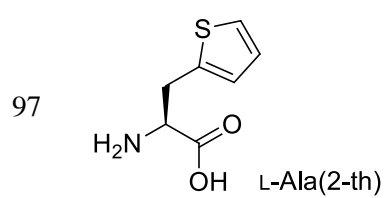
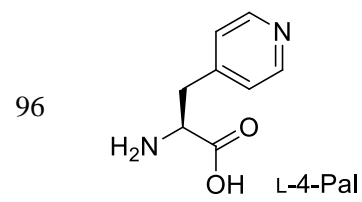
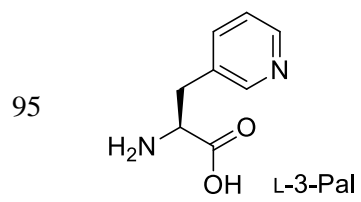
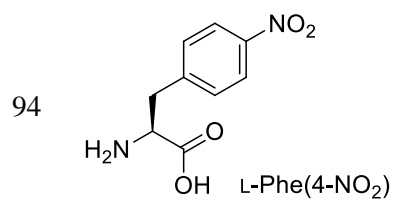


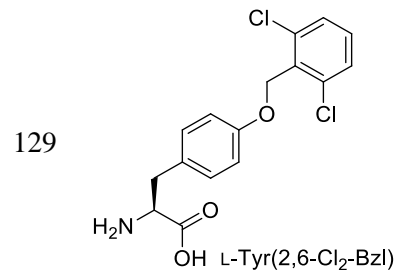
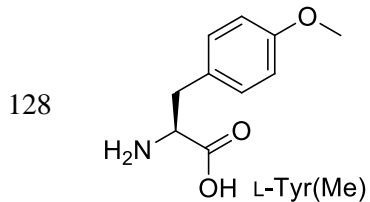
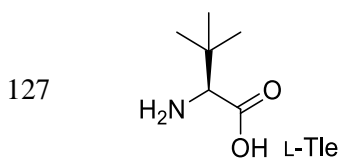
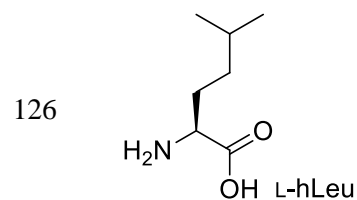
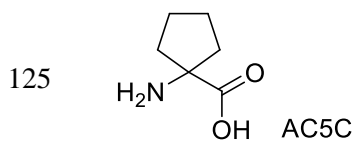
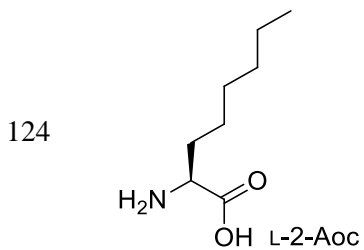
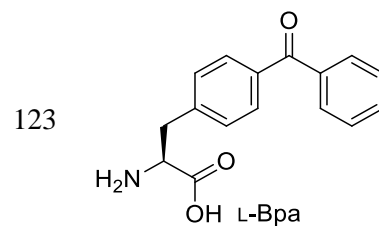
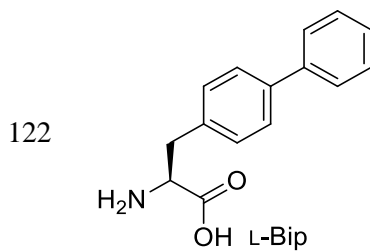
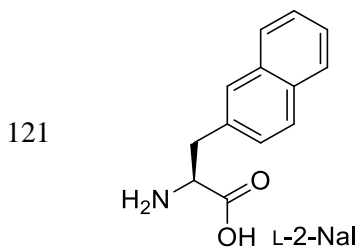
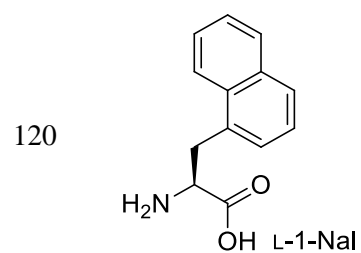
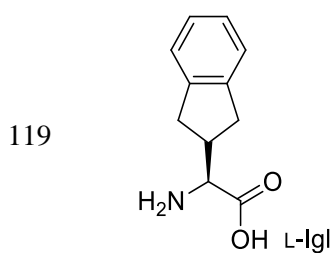
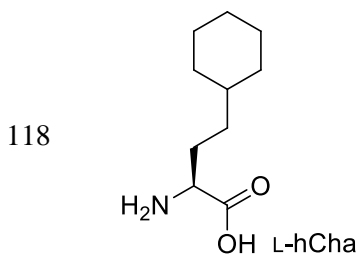
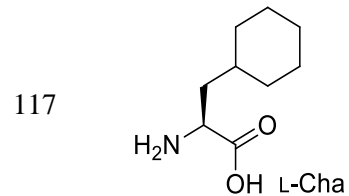
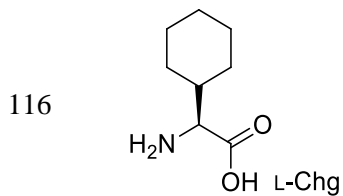
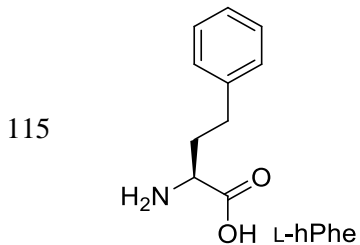
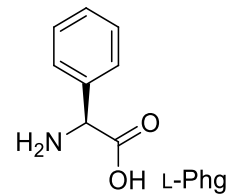
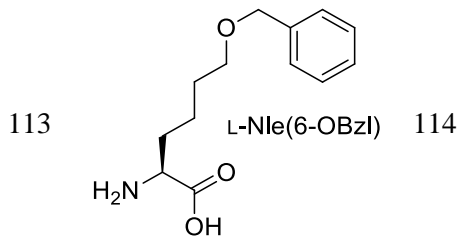
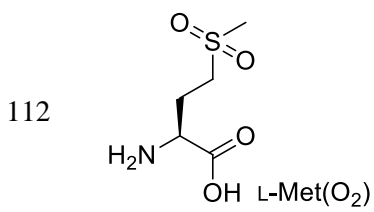




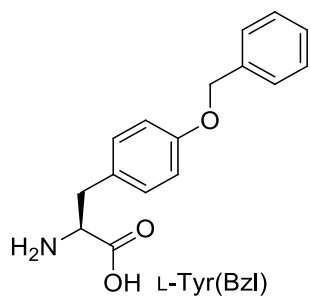




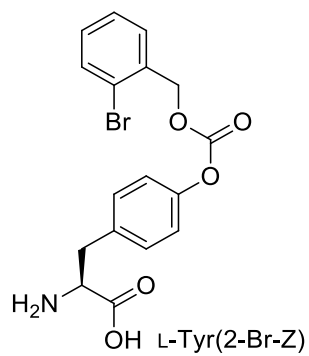




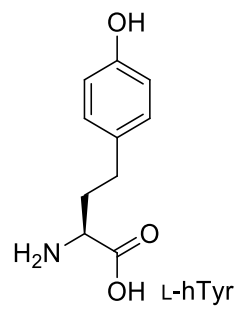
130



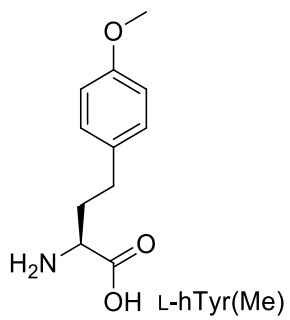
131



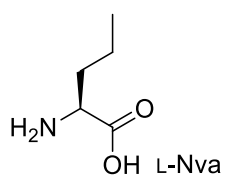
132



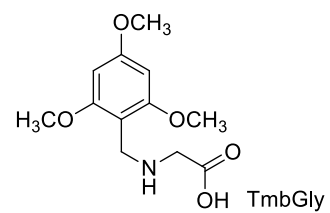
133



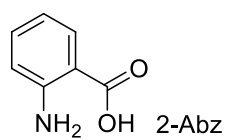
134



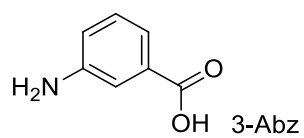
135



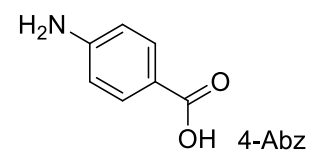
136



137



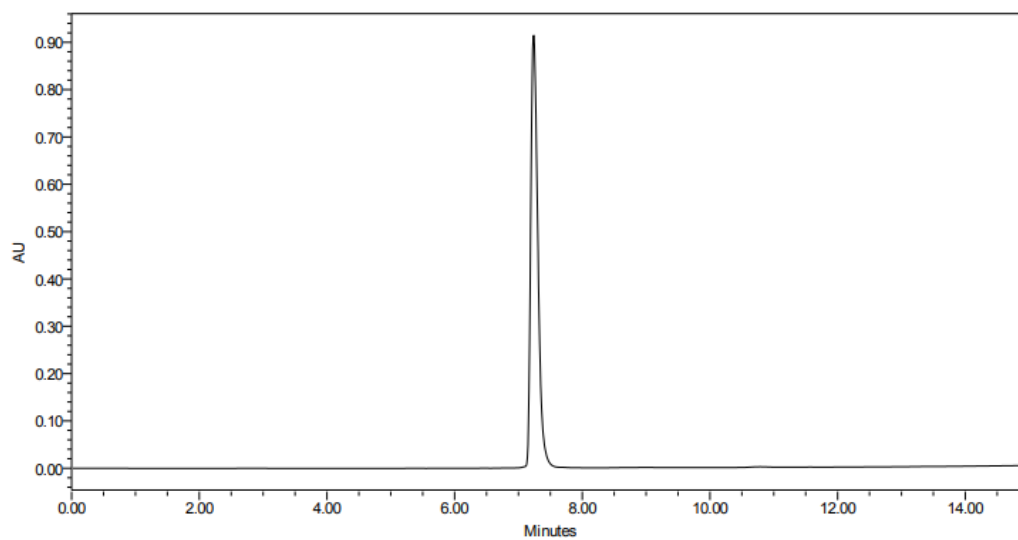
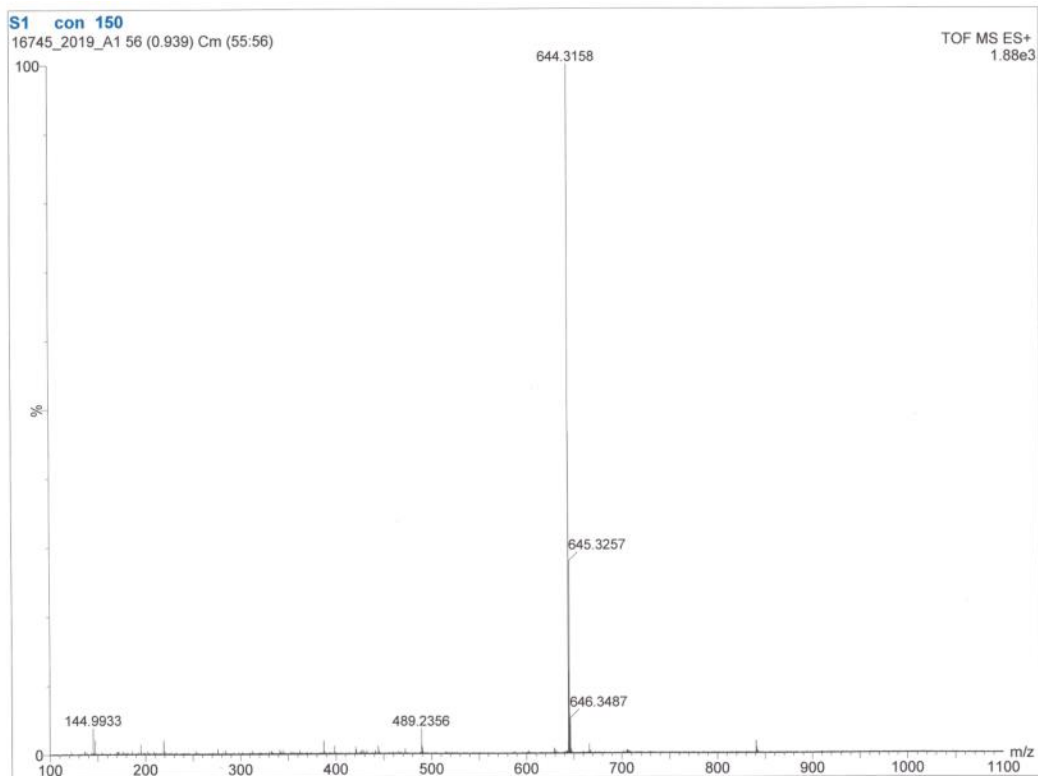
138



## HRMS and analytical chromatograms of synthesized substrates and activity-based probes

### Ac-Leu-Arg-Gly-Gly-ACC

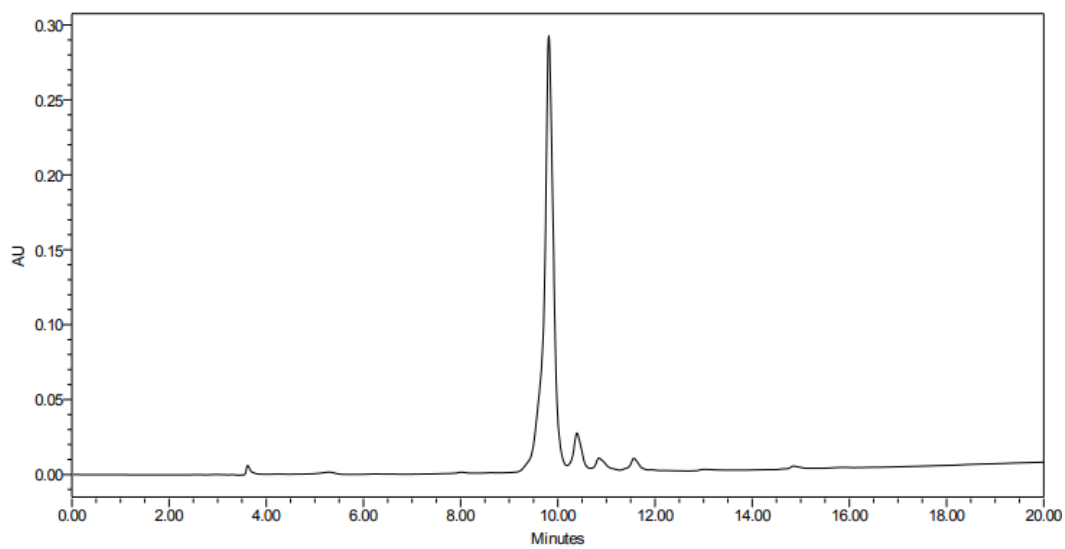
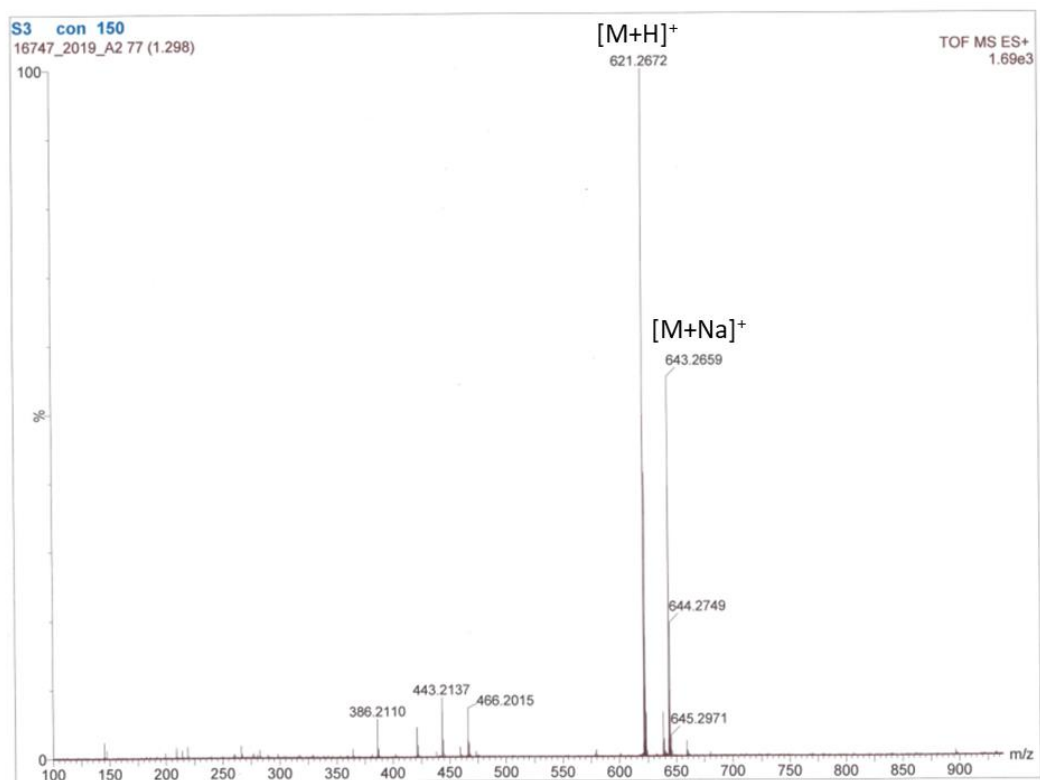
HRMS for  $C_{29}H_{41}N_9O_8$  ( $m/z$  calcd = 644.3151;  $m/z$  found = 644.3158)





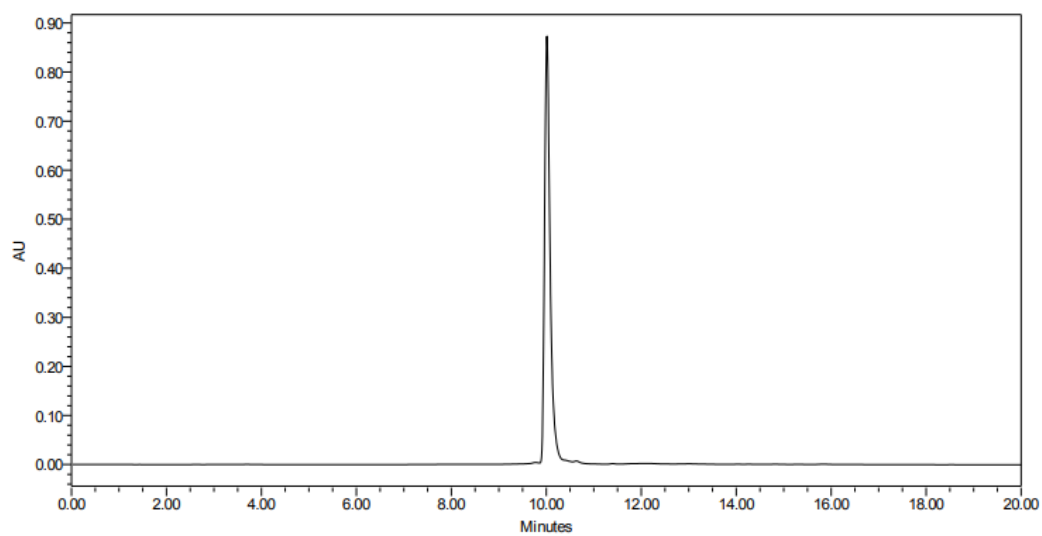
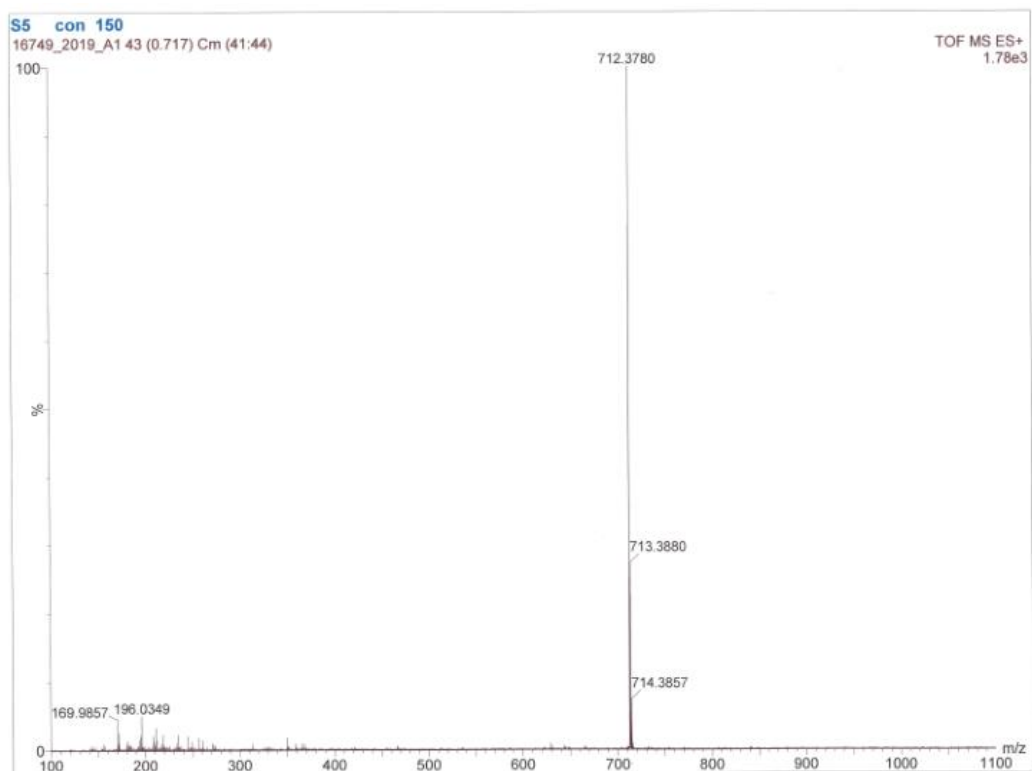
# Ac-Tle-Phg-Gly-Gly-ACC

HRMS for  $C_{31}H_{36}N_6O_8$  ( $m/z$   $_{\text{calcd}} = 621.2595$ ;  $m/z$   $_{\text{found}} = 621.2672$ )



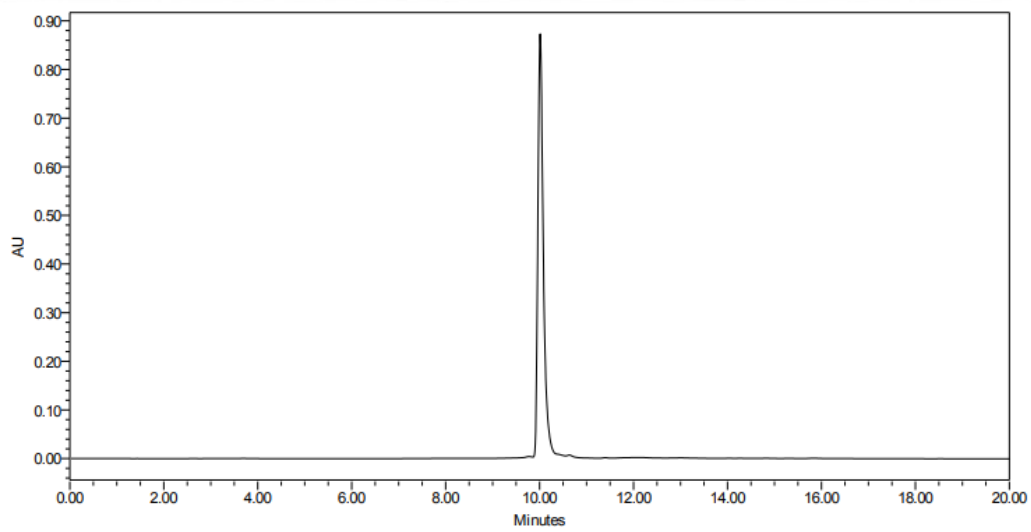
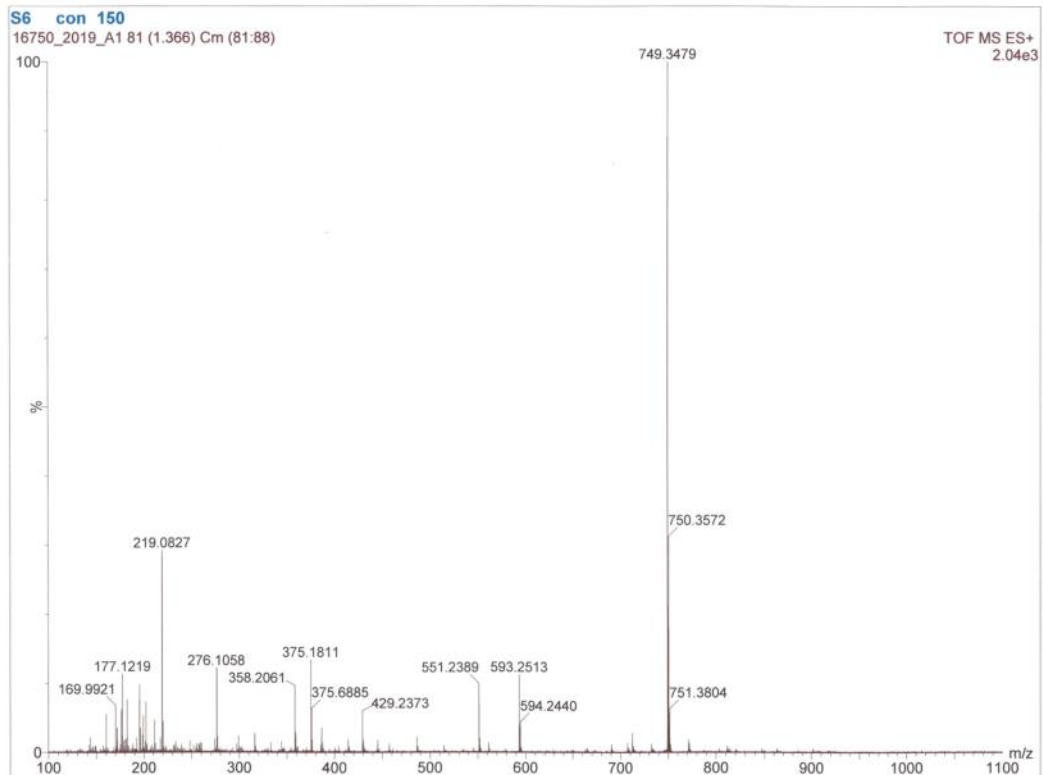
### Ac-Cha-Arg-Abu-Gly-ACC

HRMS for  $C_{34}H_{49}N_9O_8$  ( $m/z$   $_{\text{calcd}} = 712.3777$ ;  $m/z$   $_{\text{found}} = 712.3780$ )



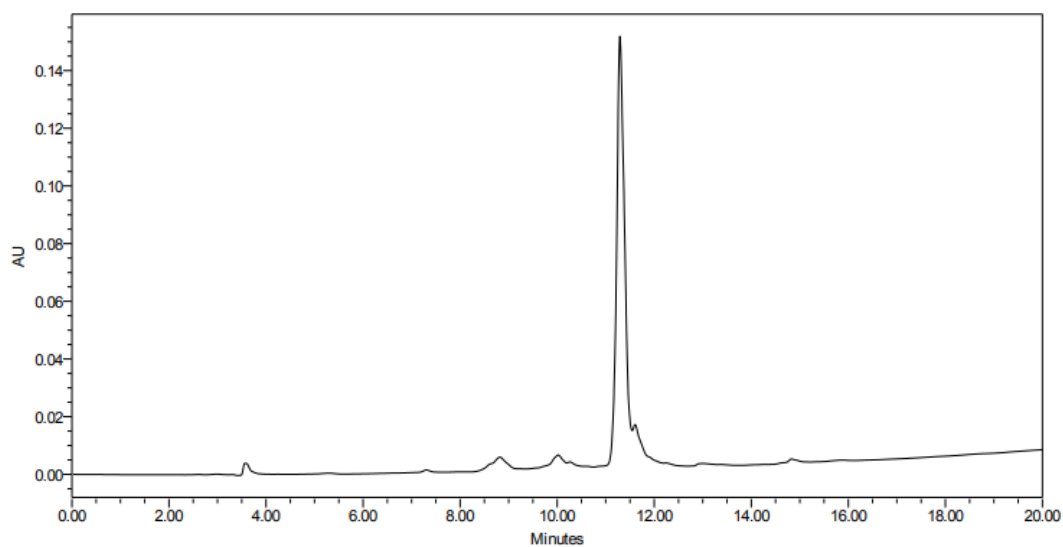
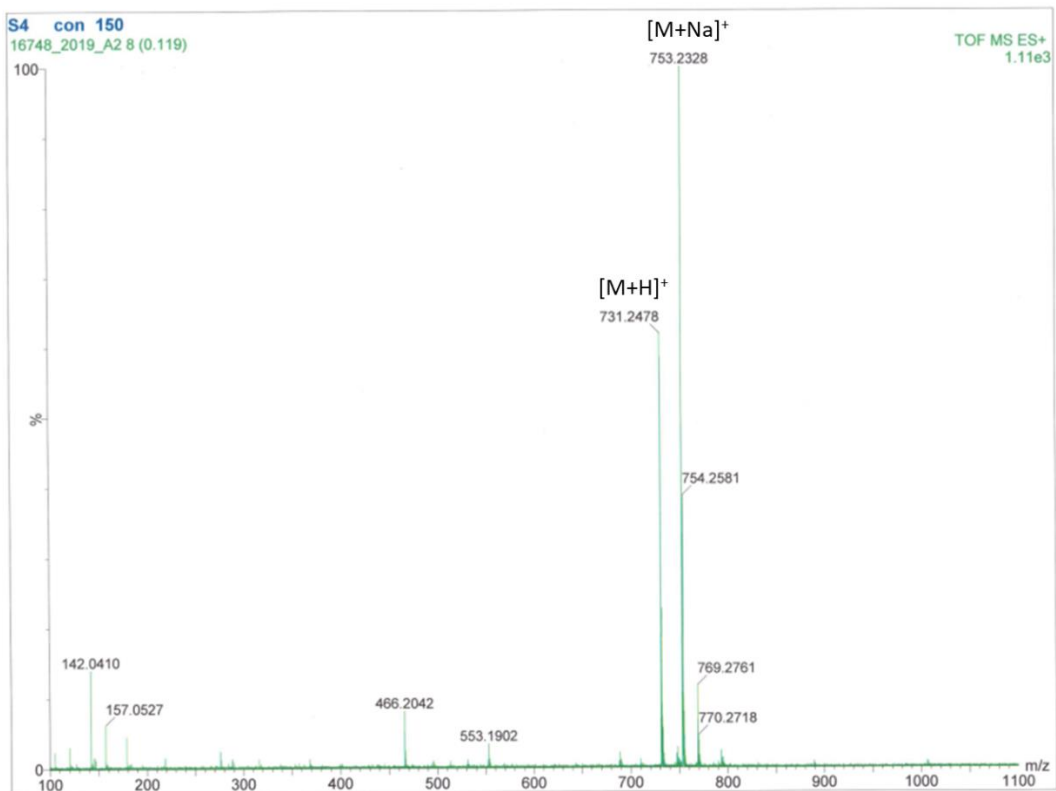
# Ac-DArg-Phe(guan)-Ala-Gly-ACC

HRMS for  $C_{34}H_{44}N_{12}O_8$  ( $m/z_{\text{calcd}} = 749.3478$ ;  $m/z_{\text{found}} = 749.3479$ )



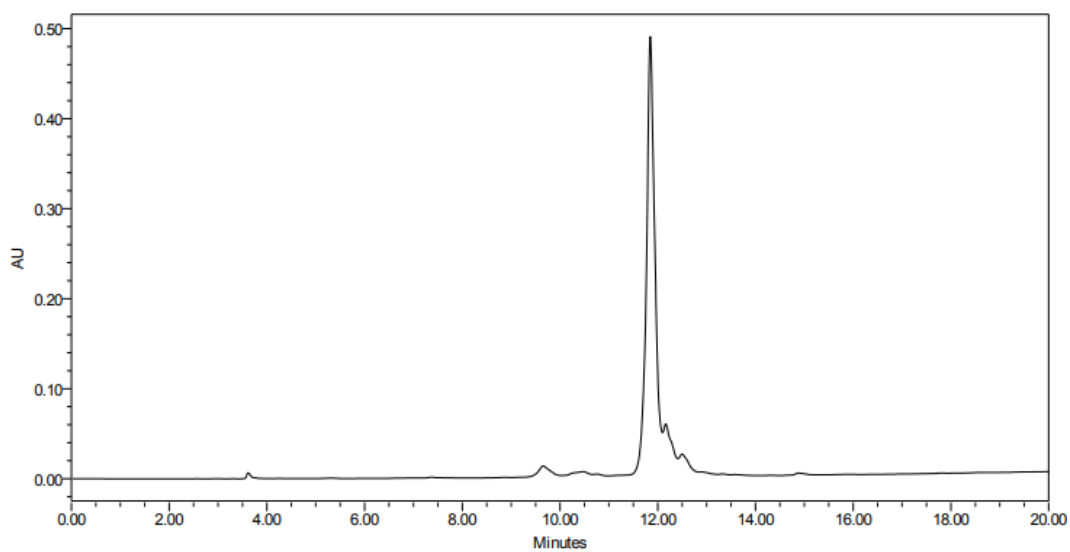
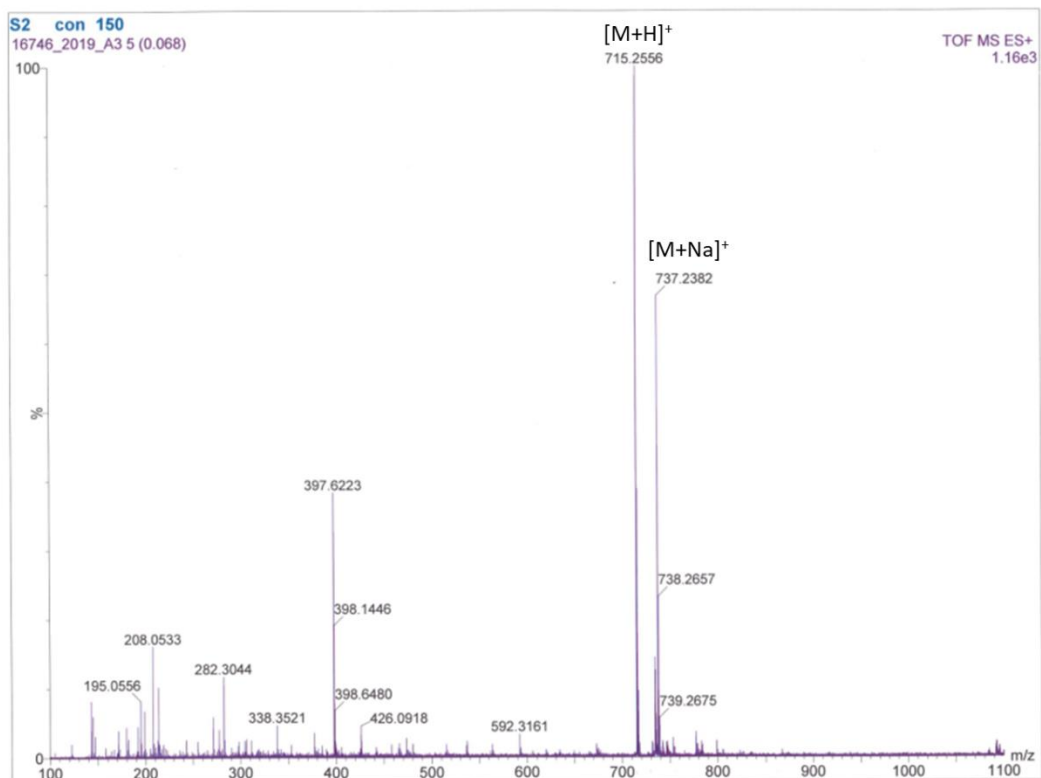
# Ac-Cys(4-MeOBzl)-Phg-Gly-Gly-ACC

HRMS for  $C_{36}H_{38}N_6O_9S$  ( $m/z$  calcd = 731.2494;  $m/z$  found = 731.2478)



# Ac-Cys(MeBzl)-Phg-Gly-Gly-ACC

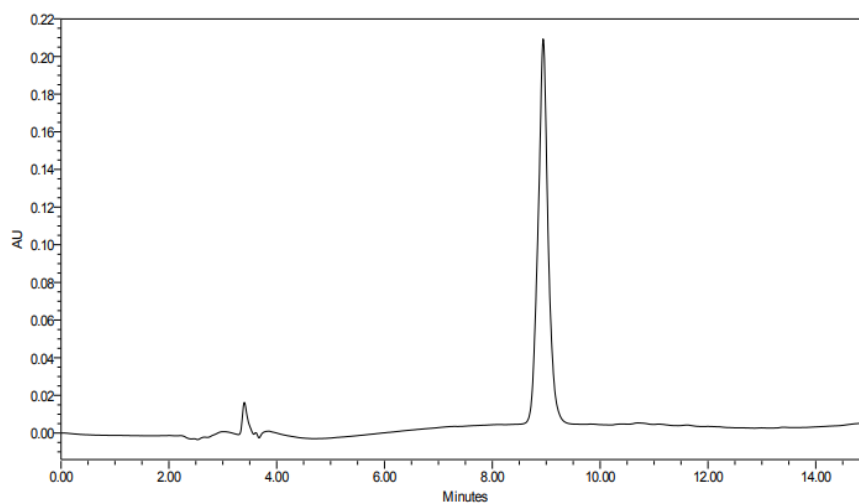
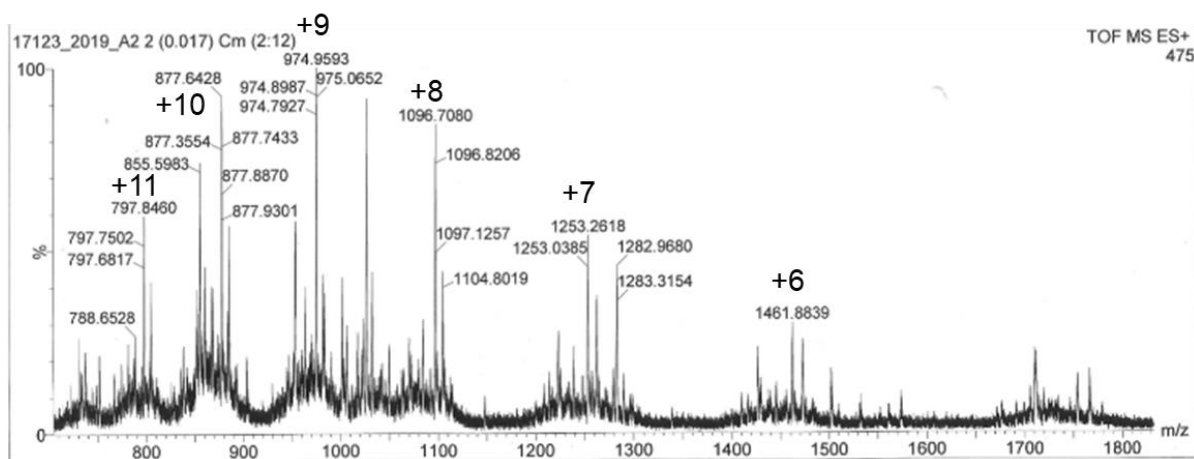
HRMS for  $C_{36}H_{38}N_6O_8S$  ( $m/z$  calcd = 715.2545;  $m/z$  found = 715.2556)



# Ub-ACC

HRMS for C<sub>389</sub>H<sub>637</sub>N<sub>107</sub>O<sub>120</sub>S

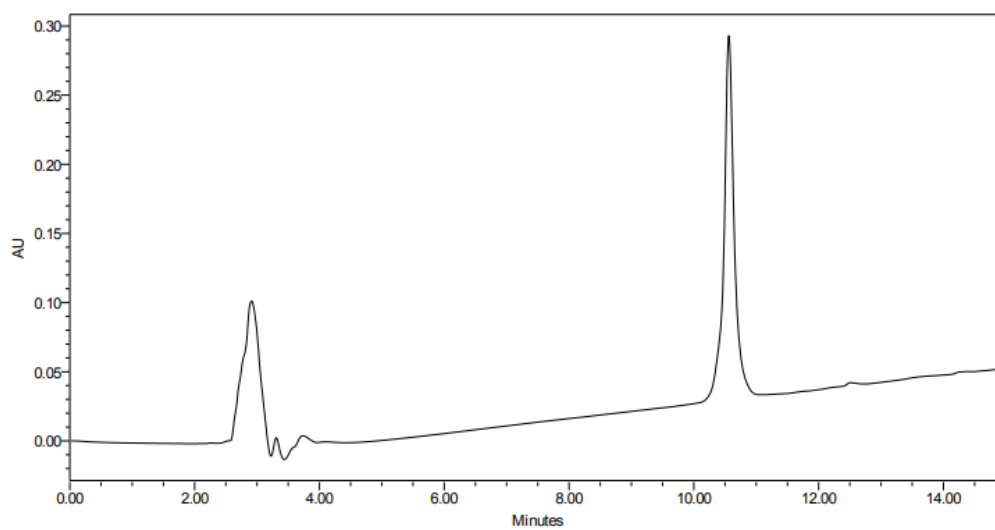
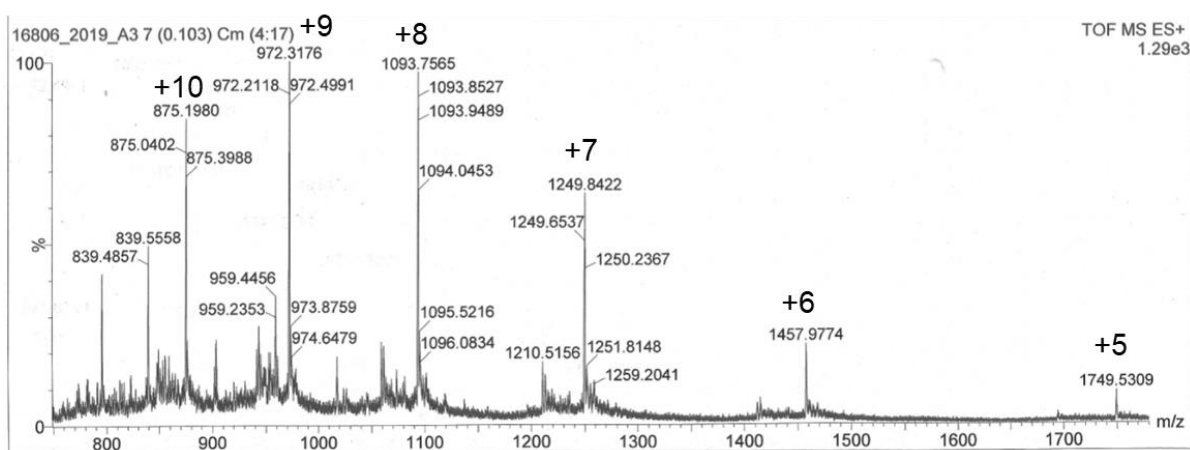
charge	m/z <sub>calcd</sub>	m/z <sub>found</sub>
+6	1461.6221	1461.8839
+7	1252.9628	1253.2618
+8	1096.4684	1096.7080
+9	974.7505	974.9593
+10	877.3762	877.6428
+11	797.7063	797.8460



## Ub-Tle-Phg-Gly-Gly-ACC (Ub-MERS27-ACC)

HRMS for C<sub>391</sub>H<sub>632</sub>N<sub>104</sub>O<sub>120</sub>S

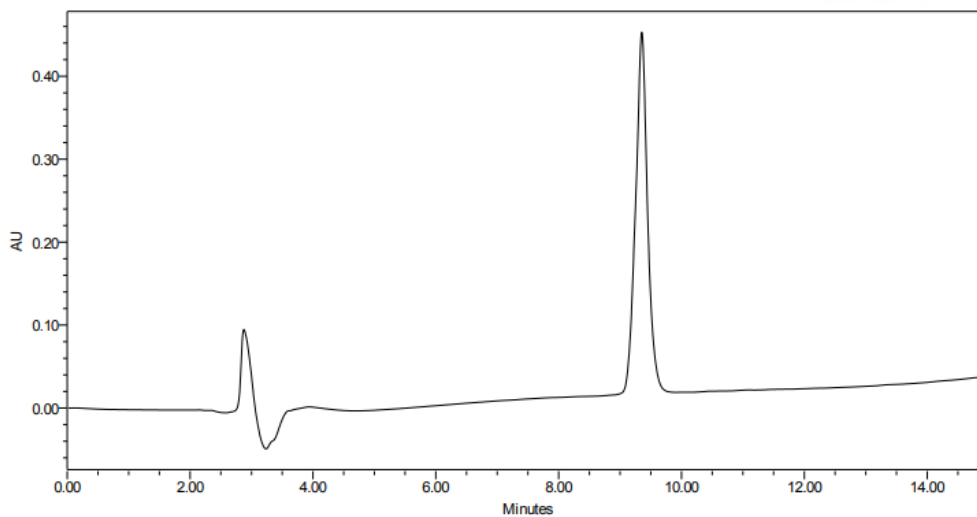
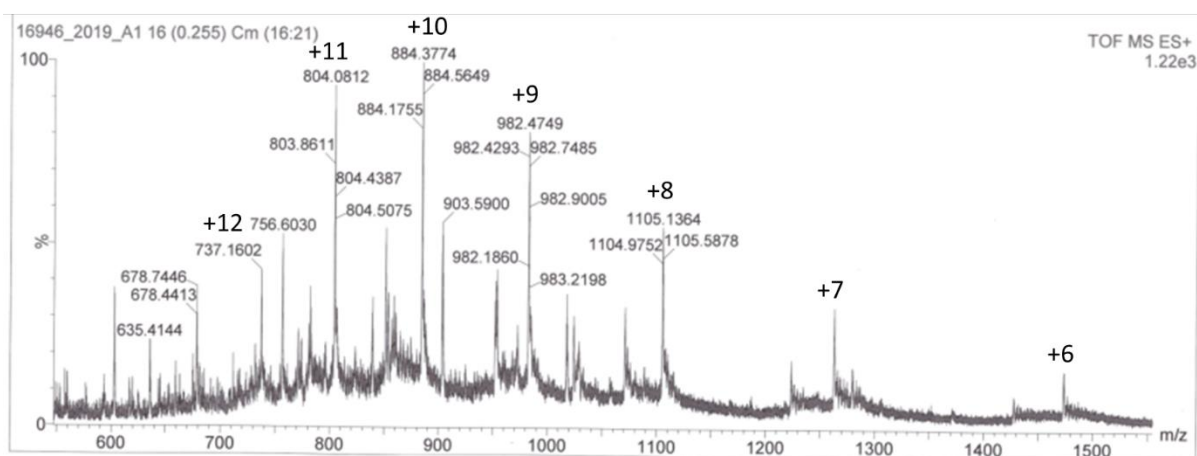
charge	m/z <sub>calcd</sub>	m/z <sub>found</sub>
+5	1749.1354	1749.5309
+6	1457.7807	1457.9774
+7	1249.6702	1249.8422
+8	1093.5874	1093.7565
+9	972.1896	972.3176
+10	875.0713	875.1980



# Ub-Cha-Arg-Abu-Gly-ACC (Ub-UCHL3-1-ACC)

HRMS for C<sub>394</sub>H<sub>645</sub>N<sub>107</sub>O<sub>120</sub>S

charge	m/z <sub>calcd</sub>	m/z <sub>found</sub>
+5	1767.3576	1767.8867
+6	1472.9659	1473.3879
+7	1262.6861	1263.0165
+8	1104.9762	1105.1364
+9	982.3130	982.4749
+10	884.1824	884.3774
+11	803.8938	804.0812
+12	736.9866	737.1602

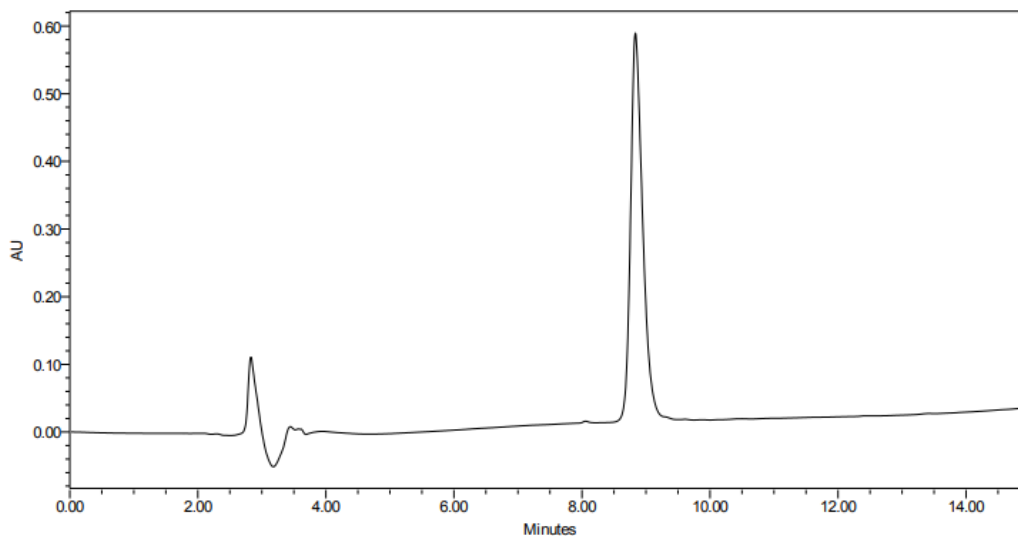
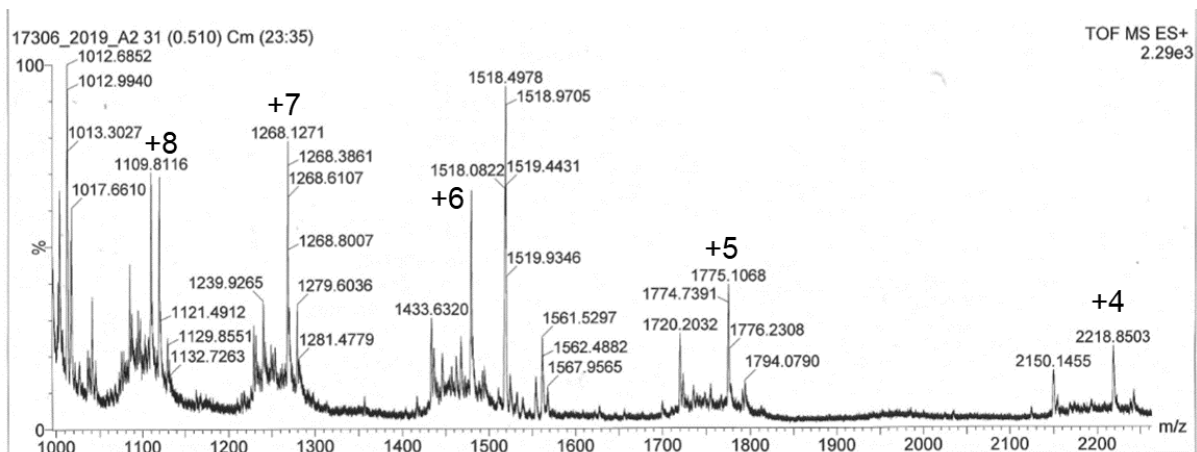




# Ub-DArg-Phe(guan)-Ala-Gly-ACC (Ub-UCHL3-2-ACC)

HRMS for C<sub>394</sub>H<sub>640</sub>N<sub>110</sub>O<sub>120</sub>S

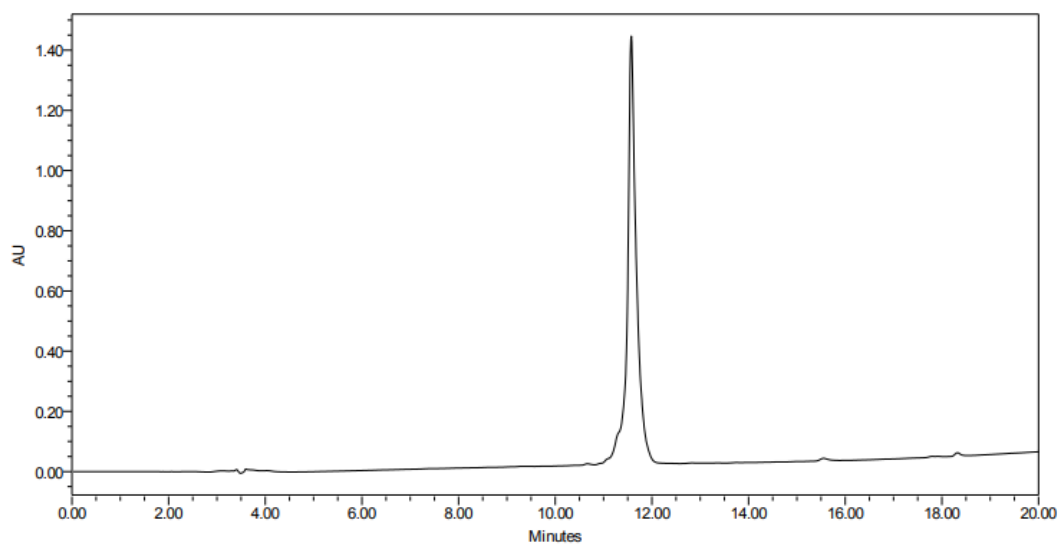
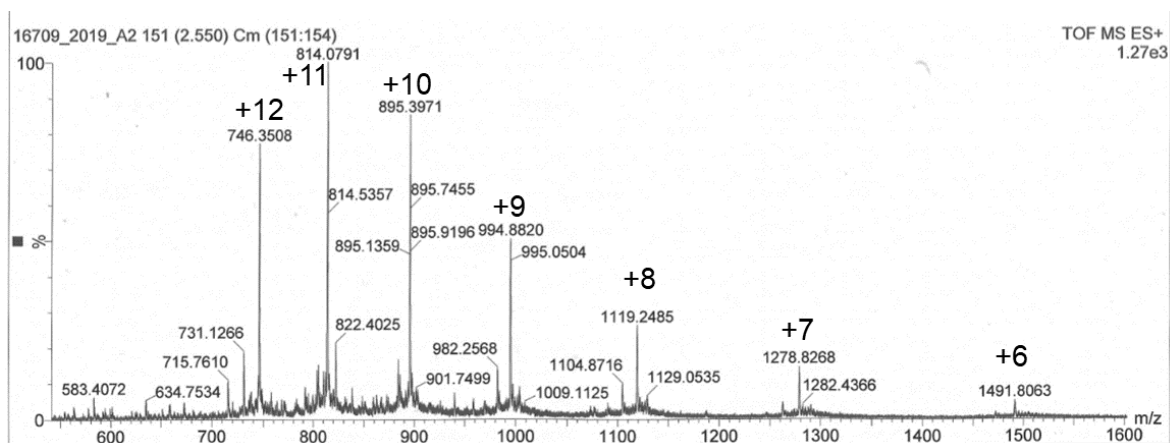
charge	m/z calcd	m/z found
+4	2218.1876	2218.8503
+5	1774.7516	1775.1068
+6	1479.1275	1479.4045
+7	1267.9675	1268.1271
+8	1109.5975	1109.8116



### Biot-6-Ahx-Ub-VME (B-Ub-VME)

HRMS for  $C_{397}H_{658}N_{108}O_{121}S_2$

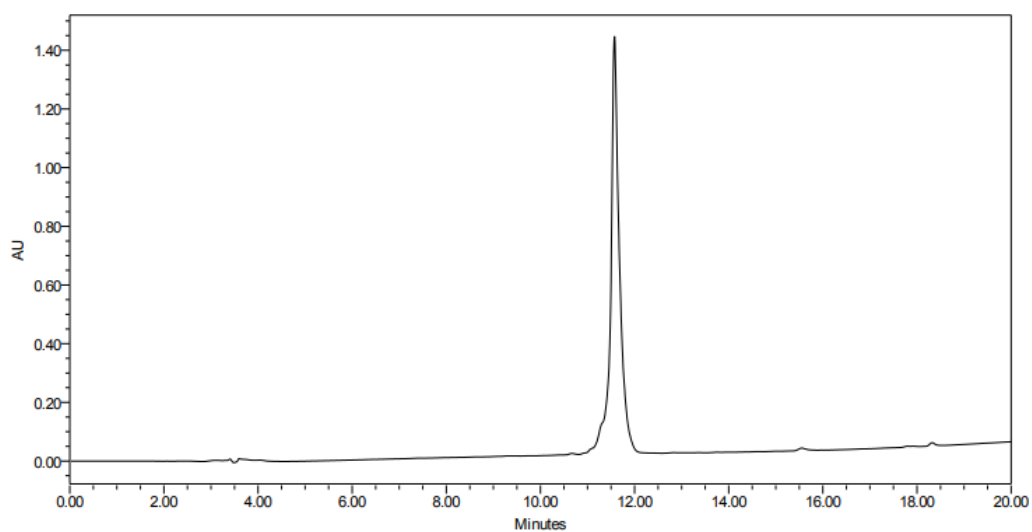
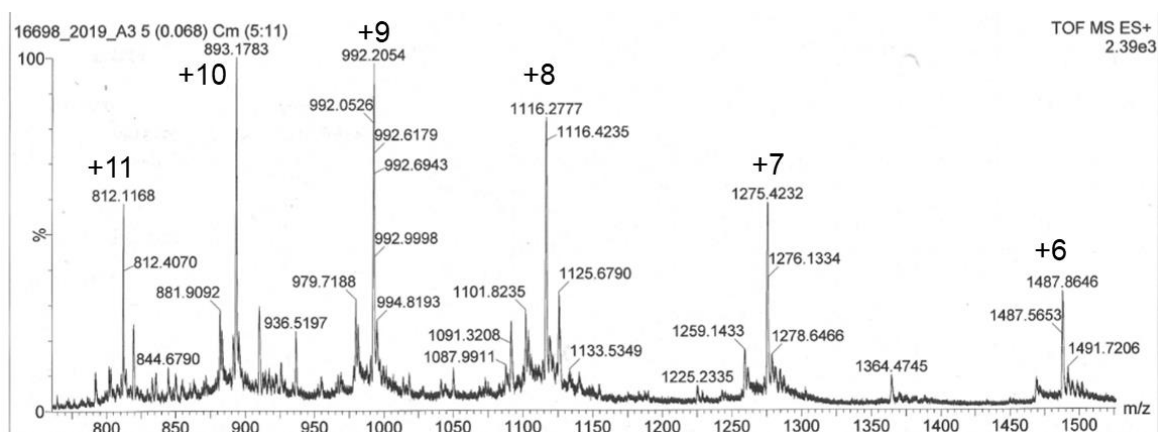
charge	m/z calcd	m/z found
+6	1491.4778	1491.8063
+7	1278.5535	1278.8268
+8	1118.8602	1119.2485
+9	994.6543	994.8820
+10	895.2896	895.3971
+11	813.9912	814.0791
+12	746.2426	746.3508



### Biot-6-Ahx-Ub-Tle-Phg-Gly-Gly-VME (B-Ub-MERS27-VME)

HRMS for C<sub>399</sub>H<sub>653</sub>N<sub>105</sub>O<sub>121</sub>S<sub>2</sub>

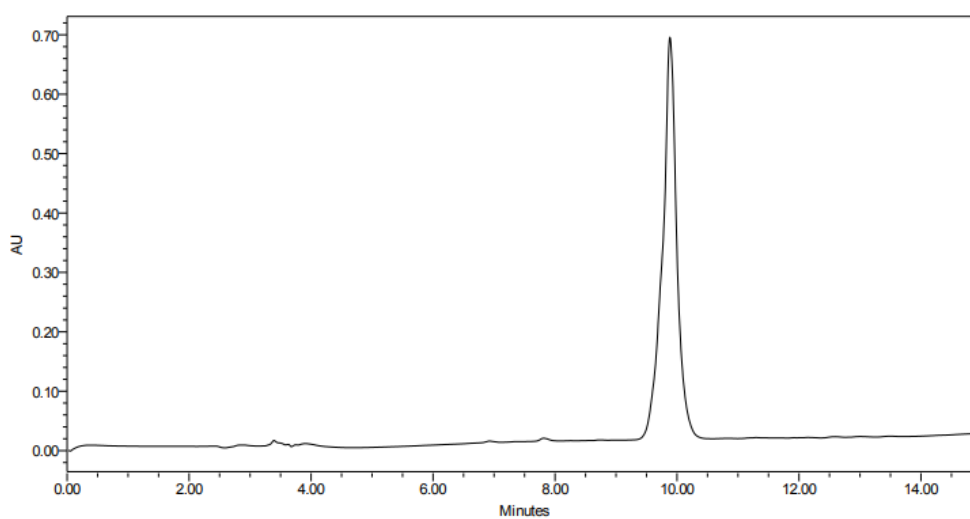
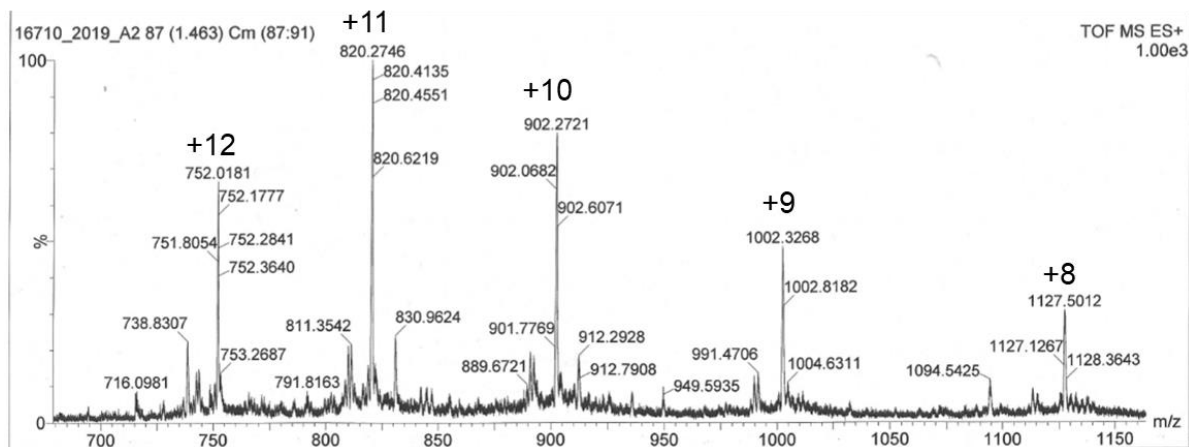
charge	m/z calcd	m/z found
+6	1487.6364	1487.8646
+7	1275.2608	1275.4232
+8	1115.9791	1116.2777
+9	992.0934	992.2054
+10	892.9848	893.1783
+11	811.8959	812.1168



# Biot-6-Ahx-Ub-Cha-Arg-Abu-Gly-VME (B-Ub-UCHL3-1-VME)

HRMS for C<sub>402</sub>H<sub>666</sub>N<sub>108</sub>O<sub>121</sub>S<sub>2</sub>

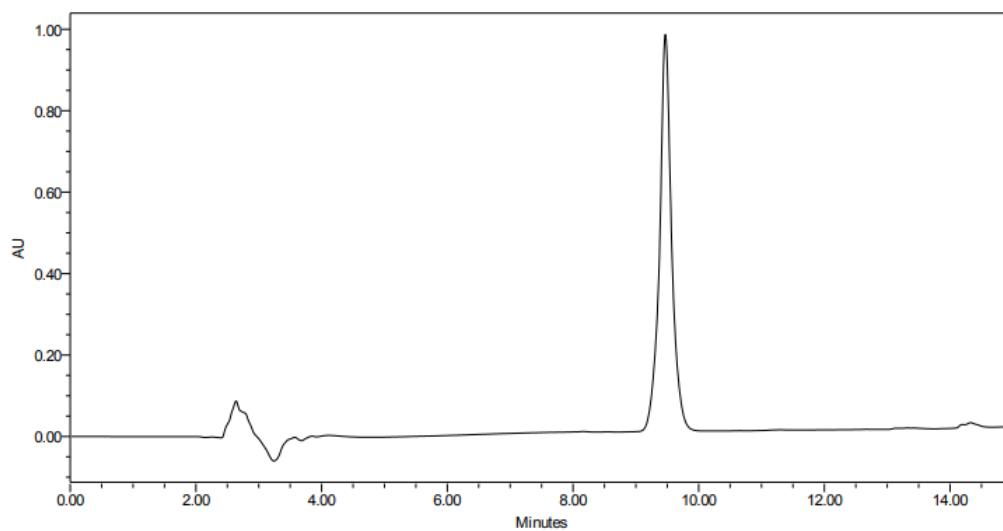
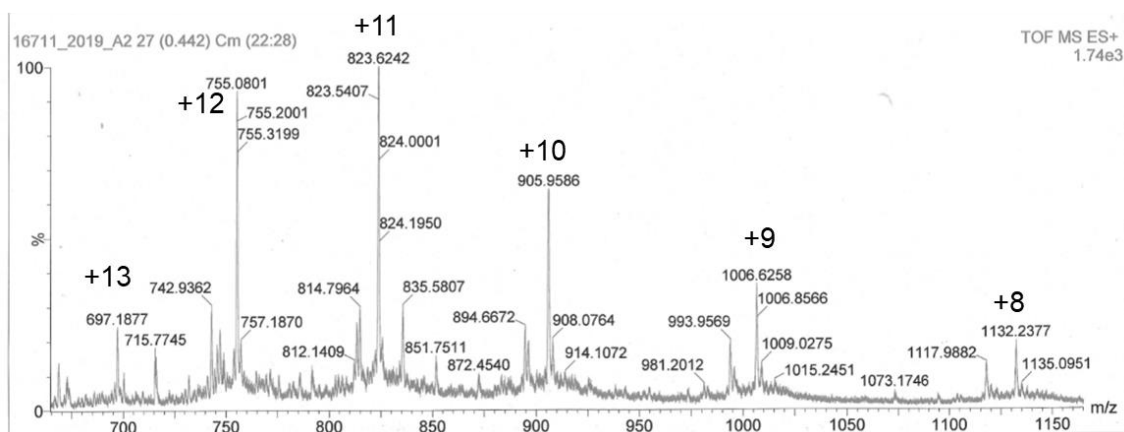
charge	m/z calcd	m/z found
+8	1127.3680	1127.5012
+9	1002.2168	1002.3268
+10	902.0959	902.2721
+11	820.1787	820.2746
+12	751.9144	752.0181



# Biot-6-Ahx-Ub-DArg-Phe(guan)-Ala-Gly-VME (B-Ub-UCHL3-2-VME)

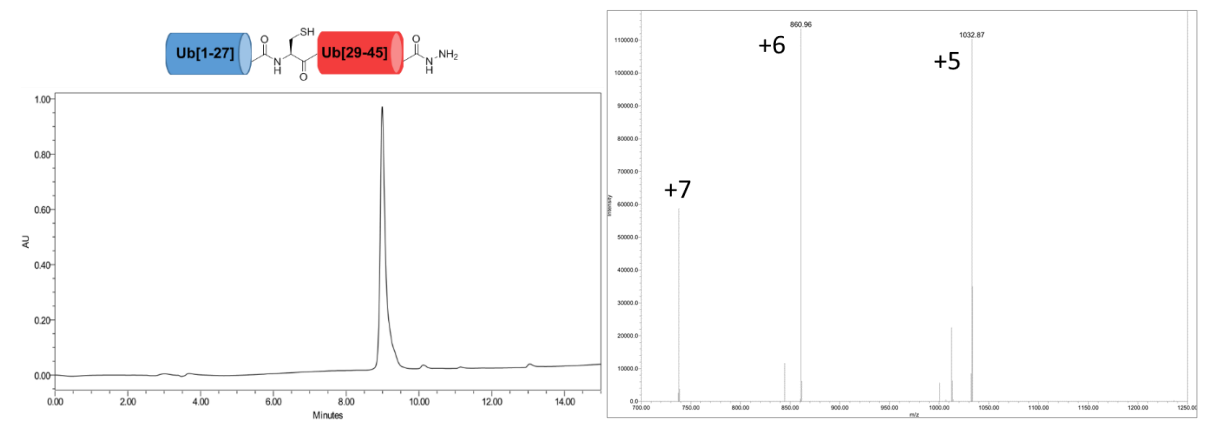
HRMS for  $C_{402}H_{661}N_{111}O_{121}S_2$

charge	m/z calcd	m/z found
+8	1131.9893	1132.2377
+9	1006.3246	1006.6258
+10	905.7929	905.9586
+11	823.5396	823.6242
+12	754.9953	755.0801
+13	696.9962	697.1877

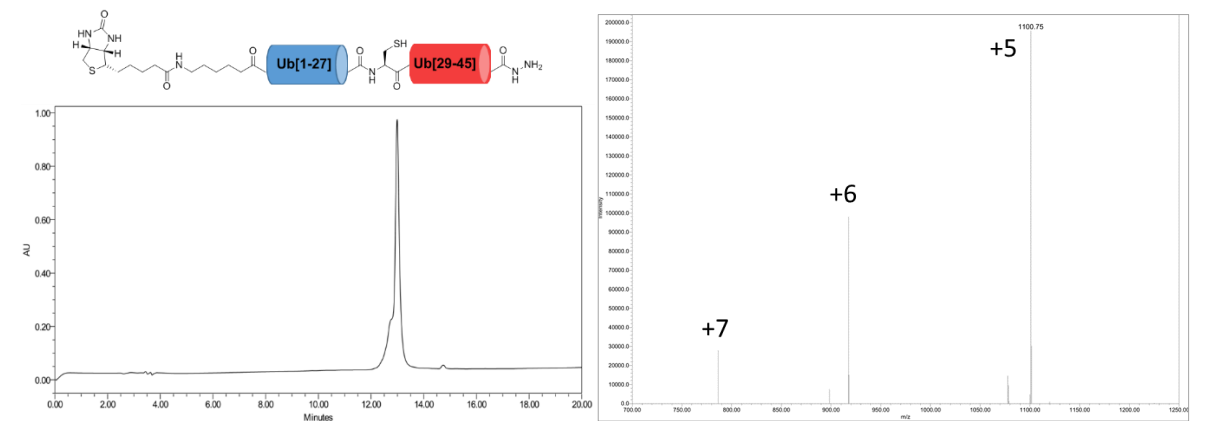




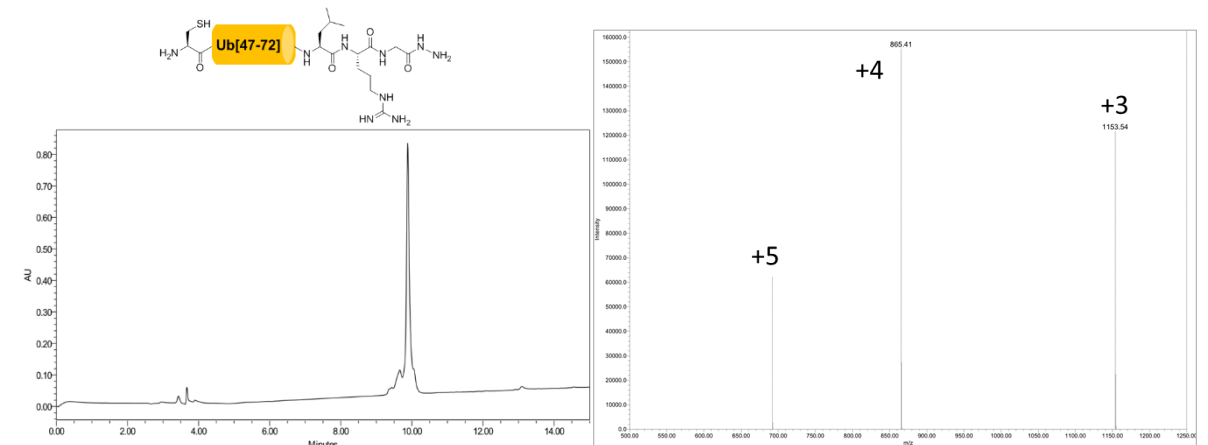
### Ub[1-45]-(A<sup>28</sup>C)-NH-NH<sub>2</sub>



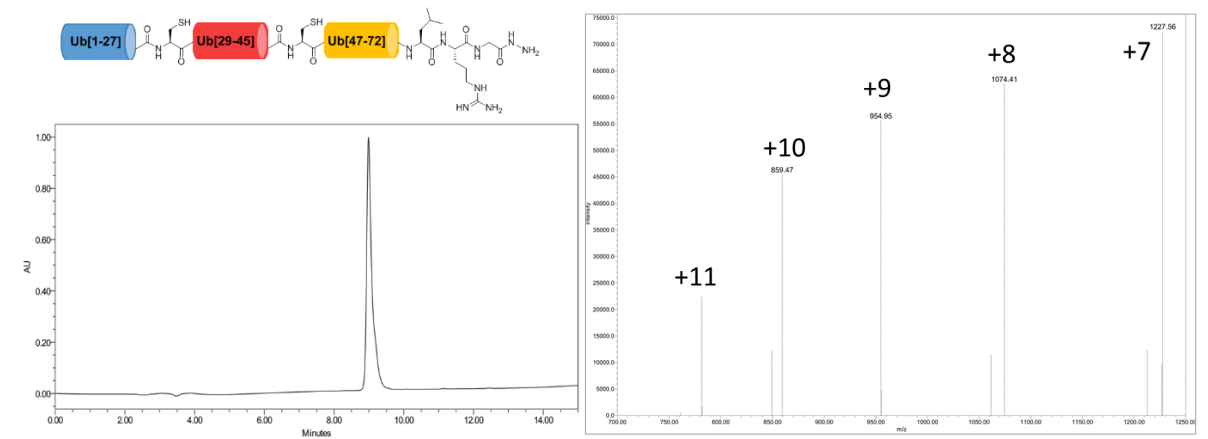
### Biot-6-Ahx-Ub[1-45]-(A<sup>28</sup>C)-NH-NH<sub>2</sub>



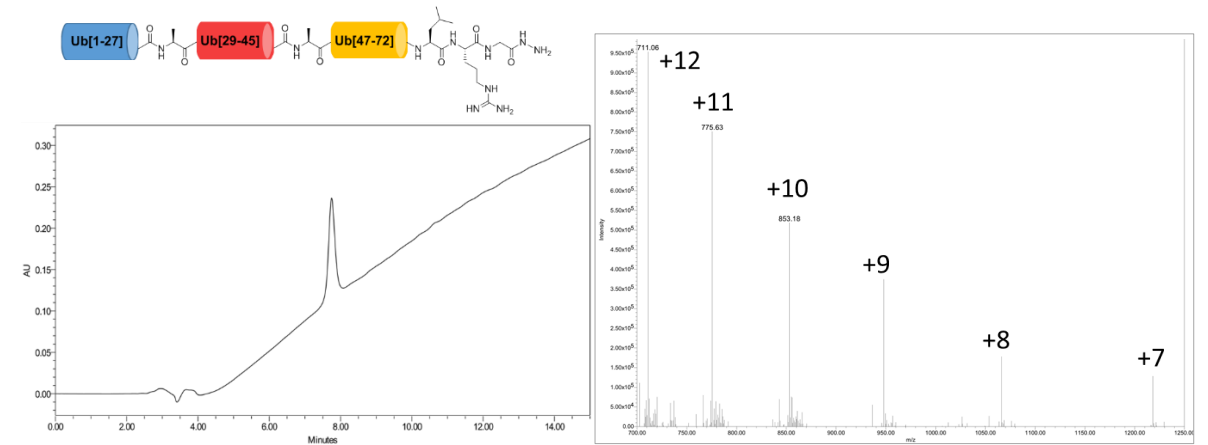
### Ub[46-72]-Leu-Arg-Gly-(A<sup>46</sup>C)-NH-NH<sub>2</sub>



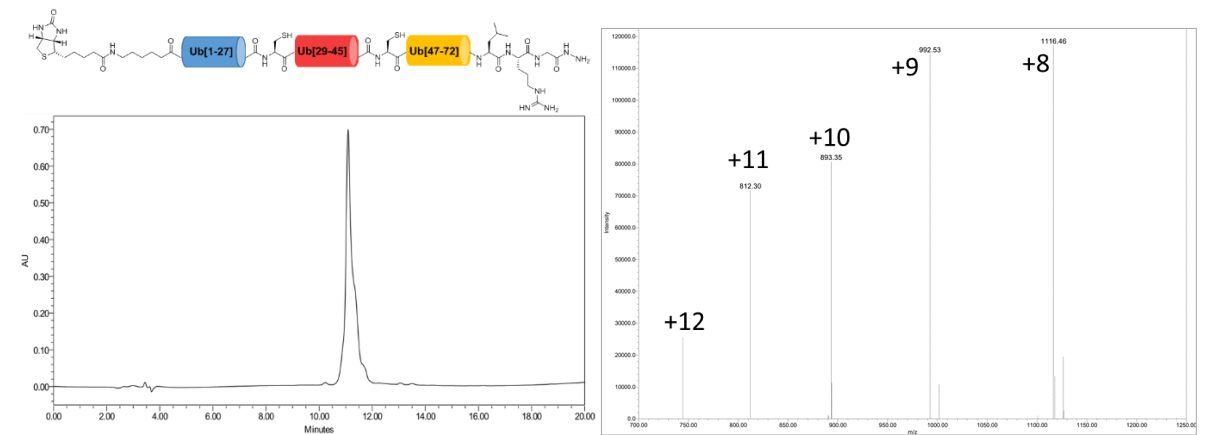
### Ub[1-72]-Leu-Arg-Gly-(A<sup>28</sup>C,A<sup>46</sup>C)-NH-NH<sub>2</sub>



### Ub[1-72]-Leu-Arg-Gly-NH-NH<sub>2</sub>



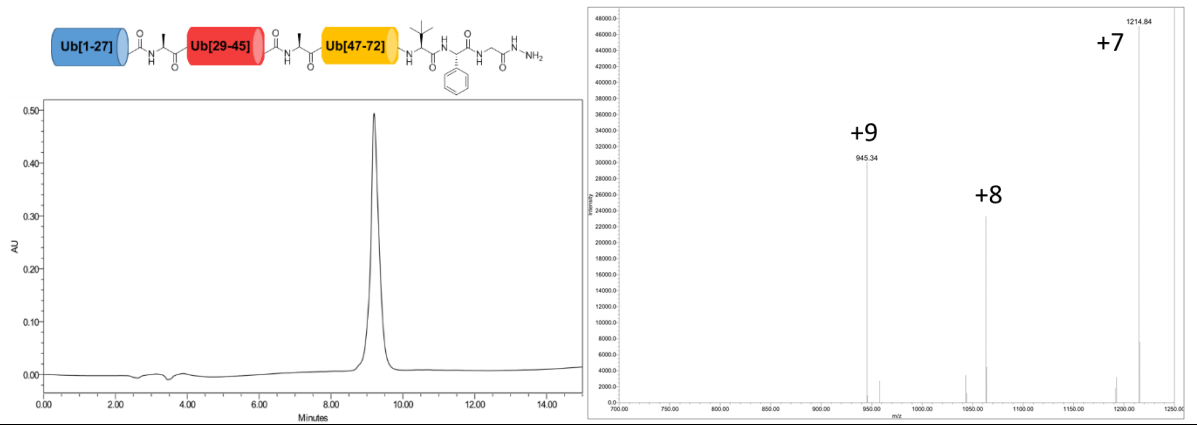
### Biot-6-Ahx-Ub[1-72]-Leu-Arg-Gly-(A<sup>28</sup>C,A<sup>46</sup>C)-NH-NH<sub>2</sub>



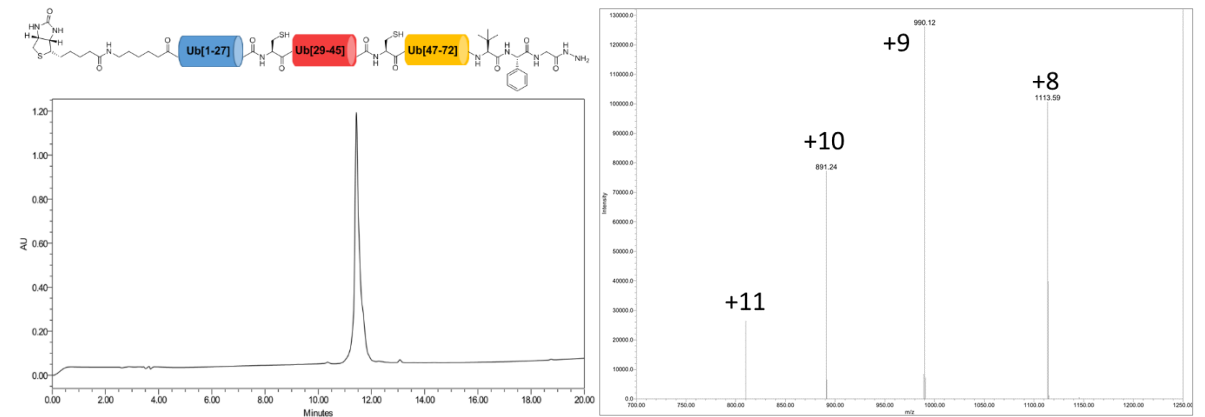




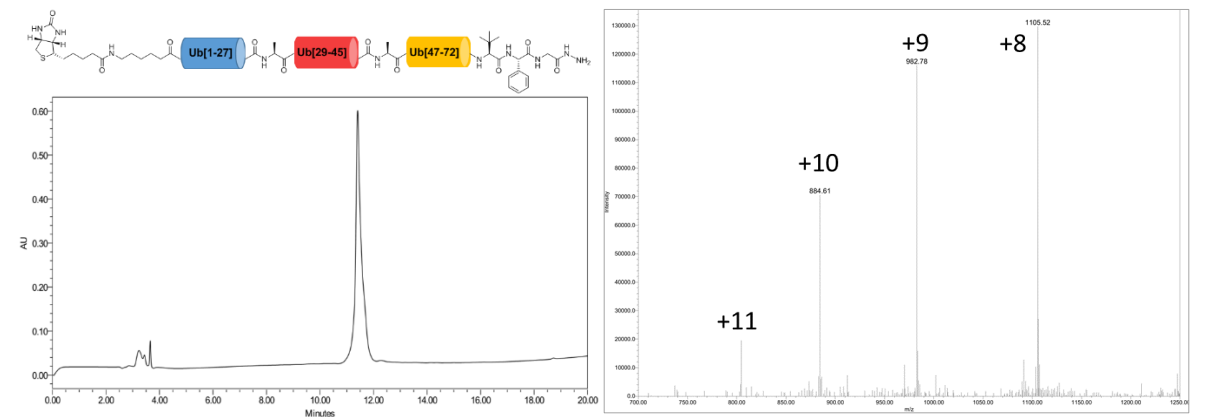
### Ub[1-72]-Tle-Phg-Gly-NH-NH<sub>2</sub>



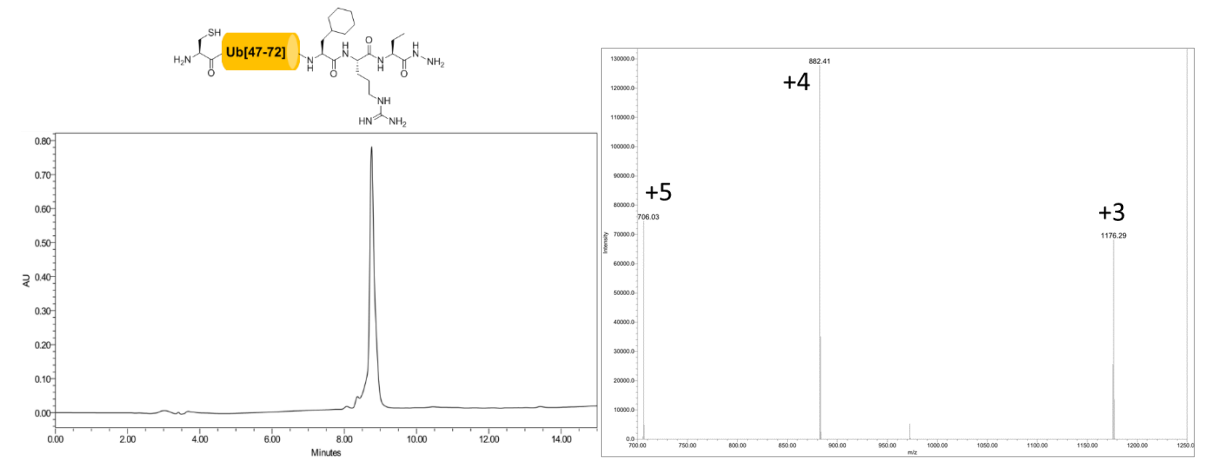
### Biot-6-Ahx-Ub[1-72]-Tle-Phg-Gly-(A<sup>28</sup>C,A<sup>46</sup>C)-NH-NH<sub>2</sub>



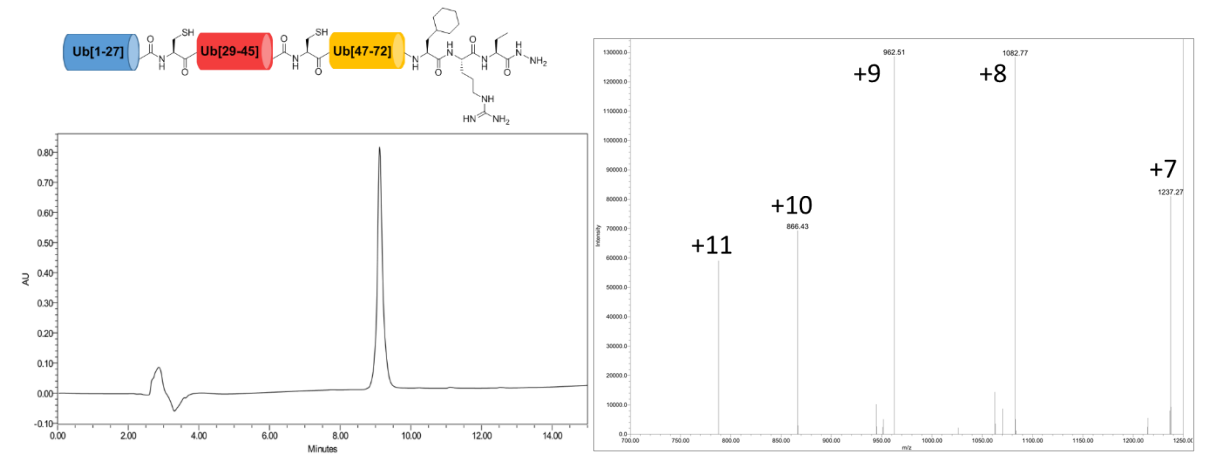
### Biot-6-Ahx-Ub[1-72]-Tle-Phg-Gly-NH-NH<sub>2</sub>



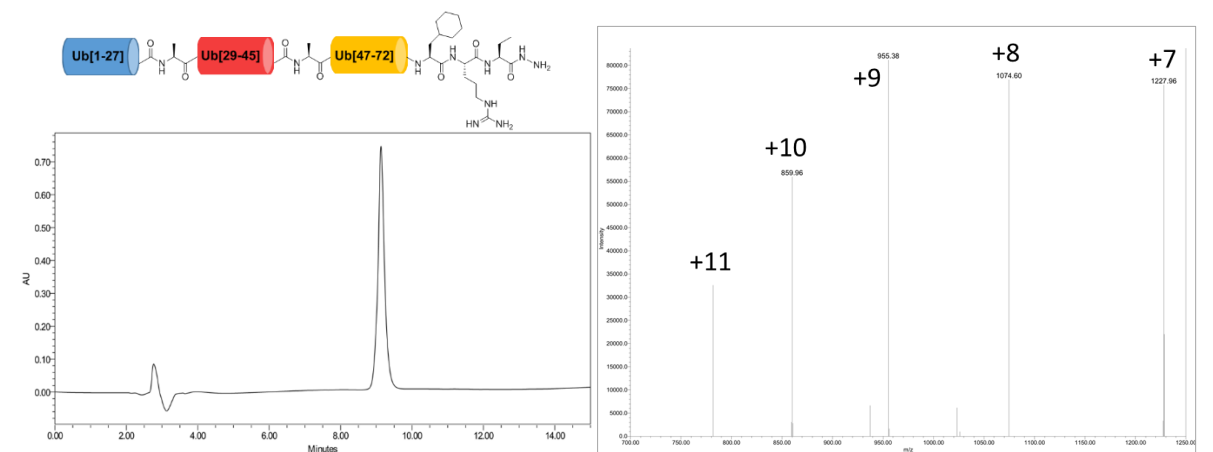
**Ub[46-72]-Cha-Arg-Abu-(A<sup>46</sup>C)-NH-NH<sub>2</sub>**



**Ub[1-72]-Cha-Arg-Abu-(A<sup>28</sup>C,A<sup>46</sup>C)-NH-NH<sub>2</sub>**

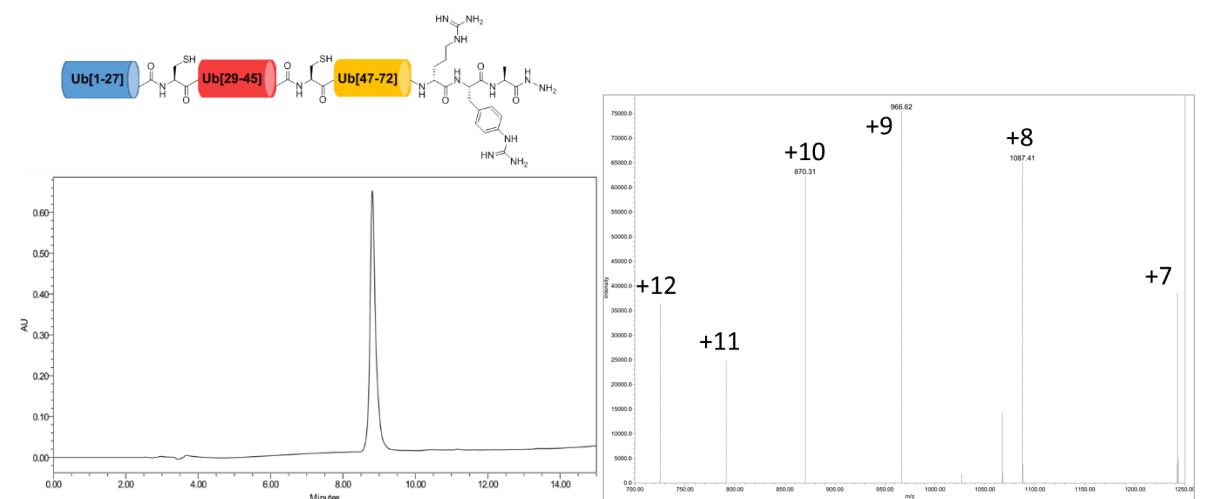


**Ub[1-72]-Cha-Arg-Abu-NH-NH<sub>2</sub>**

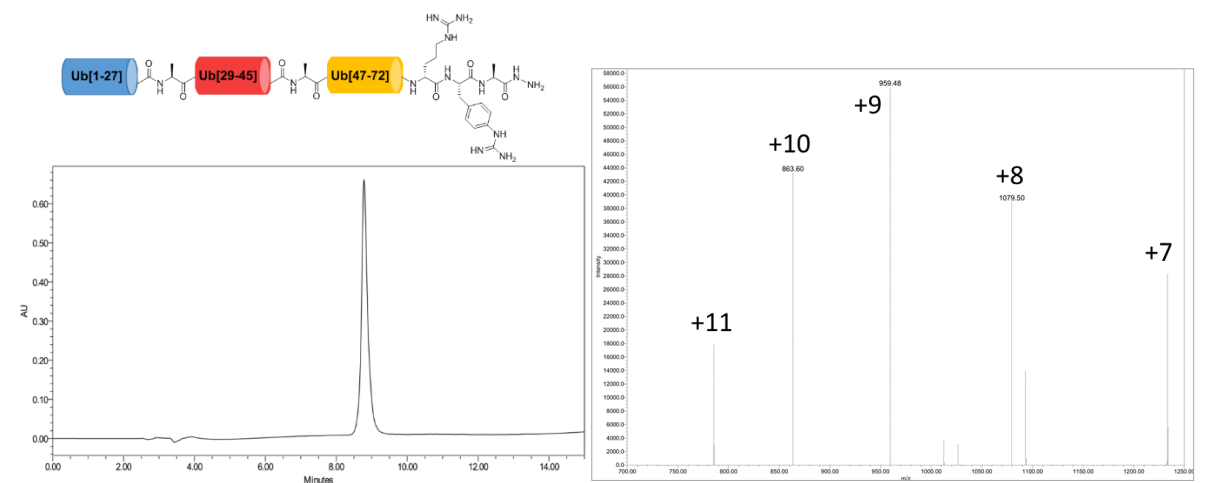




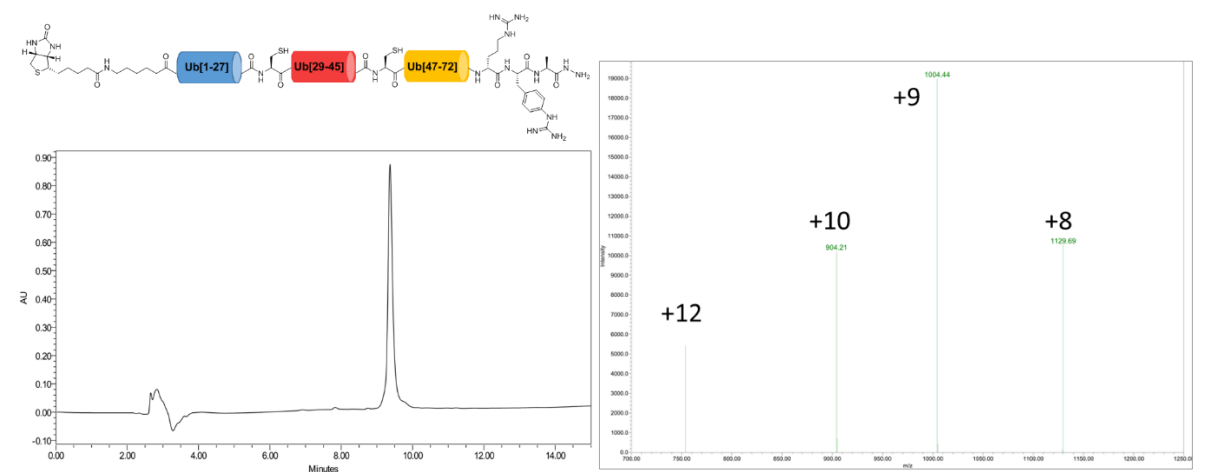
**Ub[1-72]-DArg-Phe(guan)-Ala-(A<sup>28</sup>C,A<sup>46</sup>C)-NH-NH<sub>2</sub>**



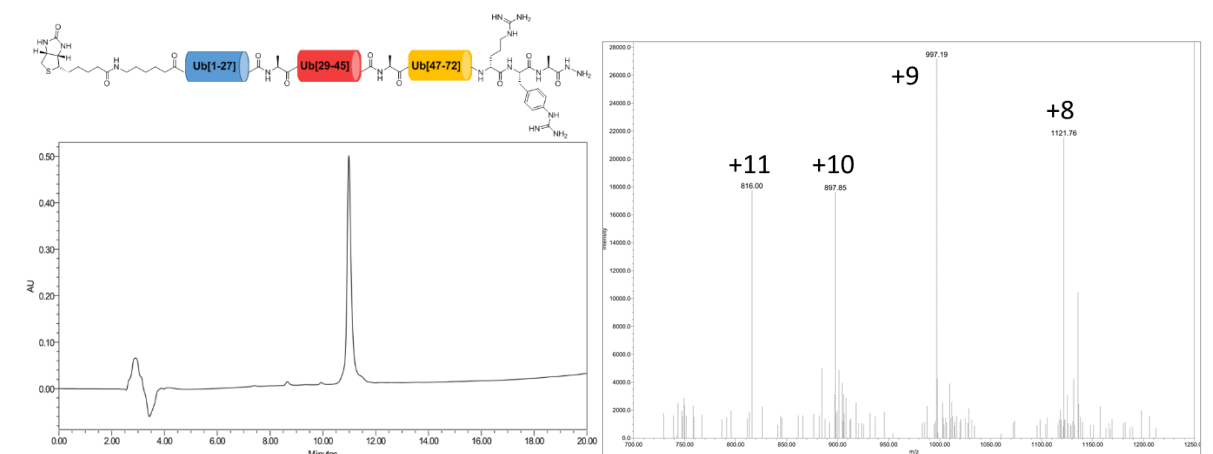
**Ub[1-72]-DArg-Phe(guan)-Ala-NH-NH<sub>2</sub>**



**Biot-6-Ahx-Ub[1-72]-DArg-Phe(guan)-Ala-(A<sup>28</sup>C,A<sup>46</sup>C)-NH-NH<sub>2</sub>**



# Biot-6-Ahx-Ub[1-72]-DArg-Phe(guan)-Ala-NH-NH<sub>2</sub>

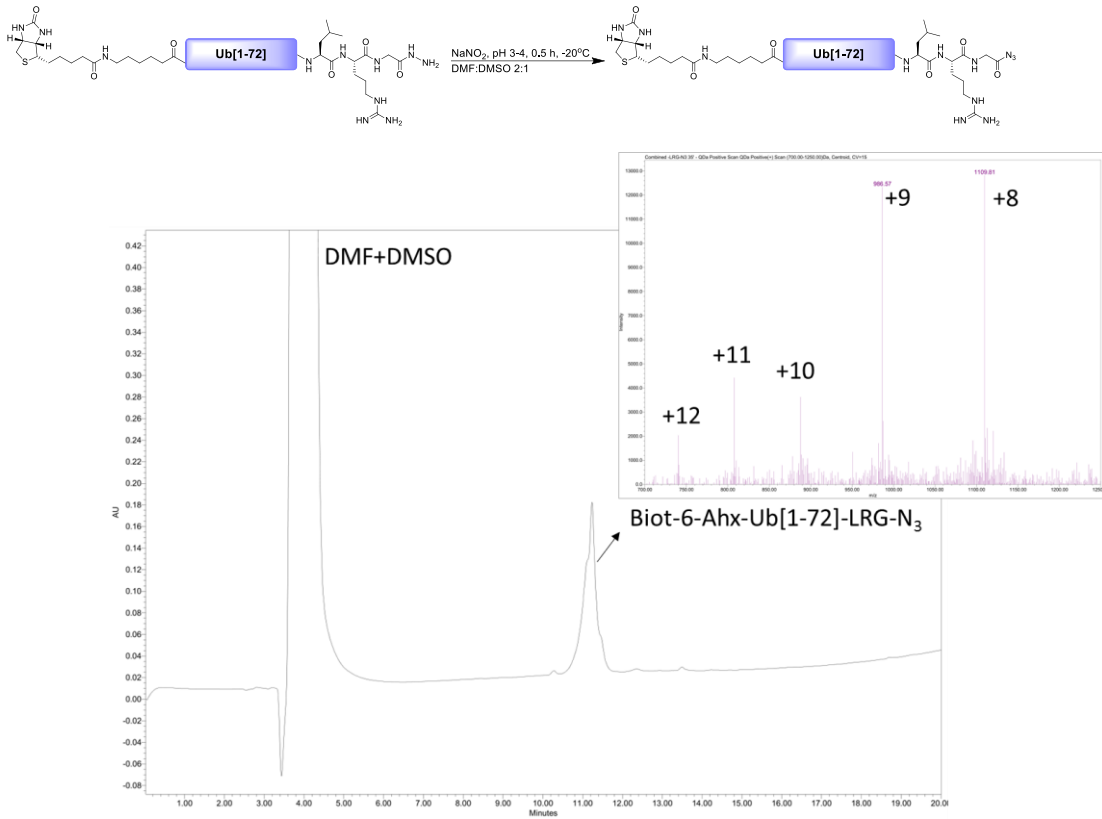




## C-terminal derivatization with the H<sub>2</sub>N-Gly-VME warhead

Shown on the example of Biot-6-Ahx-Ub-VME synthesis.

**Step 1:** acyl hydrazide to azide conversion using NaNO<sub>2</sub>



**Step 2:** H<sub>2</sub>N-Gly-VME coupling

