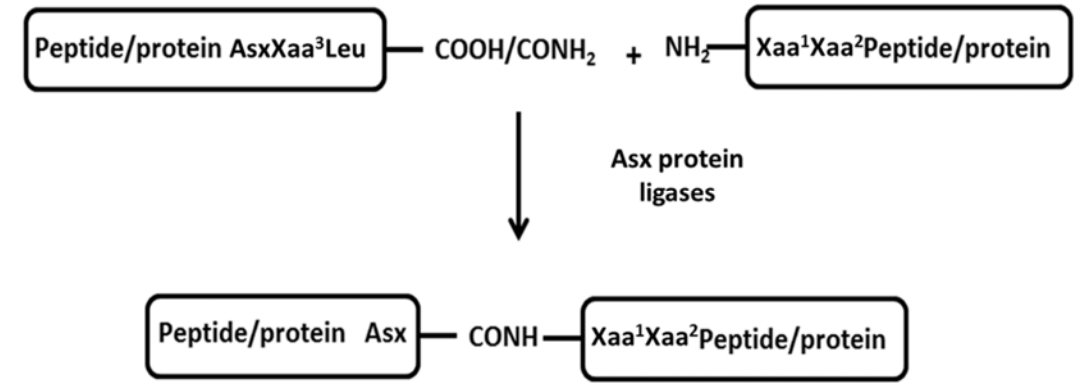
Figure 1

a


Peptide/protein AsxXaa³Leu
COOH/CONH₂
+
NH₂
Xaa¹Xaa²Peptide/protein
Asx protein ligases
Peptide/protein Asx
CONH
Xaa¹Xaa²Peptide/protein

b

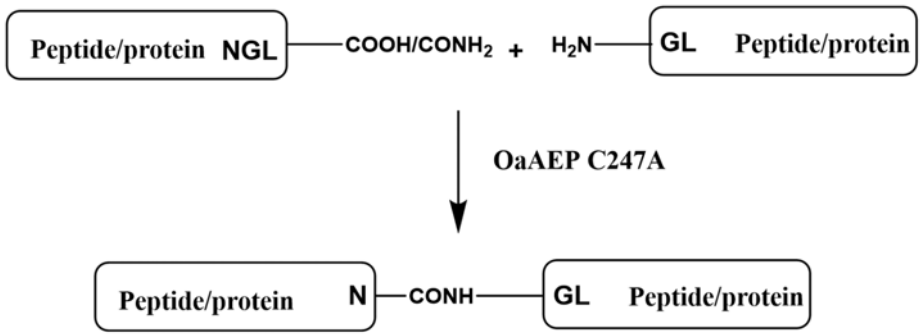

Peptide/protein NGL
COOH/CONH₂
+
H₂N
GL Peptide/protein
OaAEP C247A
Peptide/protein N
CONH
GL Peptide/protein

Figure 2

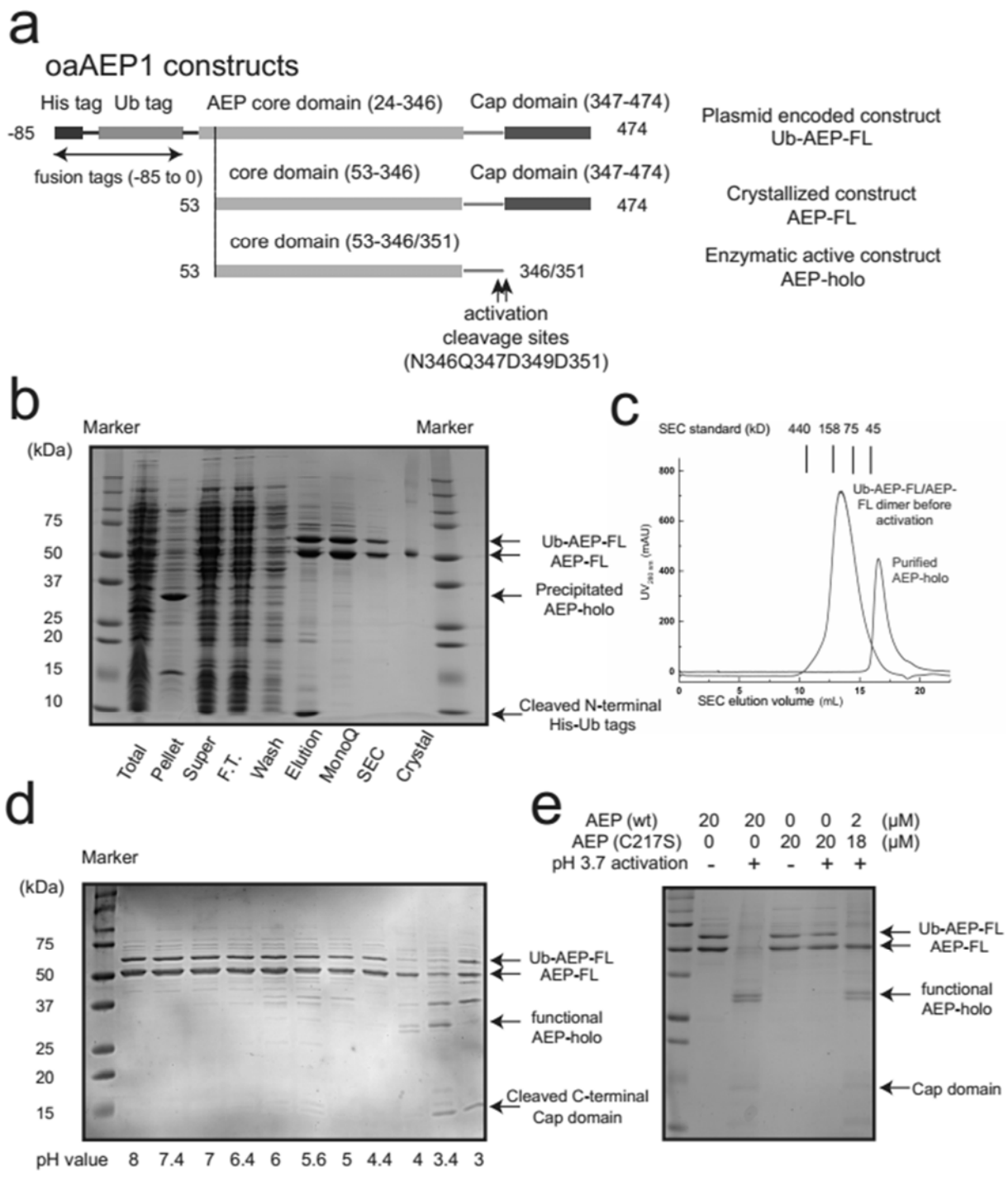

a
oaAEP1 constructs
His tag
Ub tag
AEP core domain (24-346)
Cap domain (347-474)
-85
474
fusion tags (-85 to 0)
core domain (53-346)
53
core domain (53-346/351)
346/351
activation
cleavage sites
(N346Q347D349D351)
Plasmid encoded construct
Ub-AEP-FL
Crystallized construct
AEP-FL
Enzymatic active construct
AEP-holo
b
Marker
Marker
(kDa)
75
50
37
25
20
15
10
Ub-AEP-FL
AEP-FL
Precipitated
AEP-holo
Cleaved N-terminal
His-Ub tags
Total
Pellet
Super
F.T.
Wash
Elution
MonoQ
SEC
Crystal
c
SEC standard (kD)
440
158
75
45
Ub-AEP-FL/AEP-
FL dimer before
activation
Purified
AEP-holo
SEC elution volume (mL)
d
Marker
(kDa)
75
50
37
25
20
15
Ub-AEP-FL
AEP-FL
functional
AEP-holo
Cleaved C-terminal
Cap domain
pH value 8 7.4 7 6.4 6 5.6 5 4.4 4 3.4 3
e
AEP (wt) 20 20 0 0 2 (μM)
AEP (C217S) 0 0 20 20 18 (μM)
pH 3.7 activation - + - + +
Ub-AEP-FL
AEP-FL
functional
AEP-holo
Cap domain

Figure 3

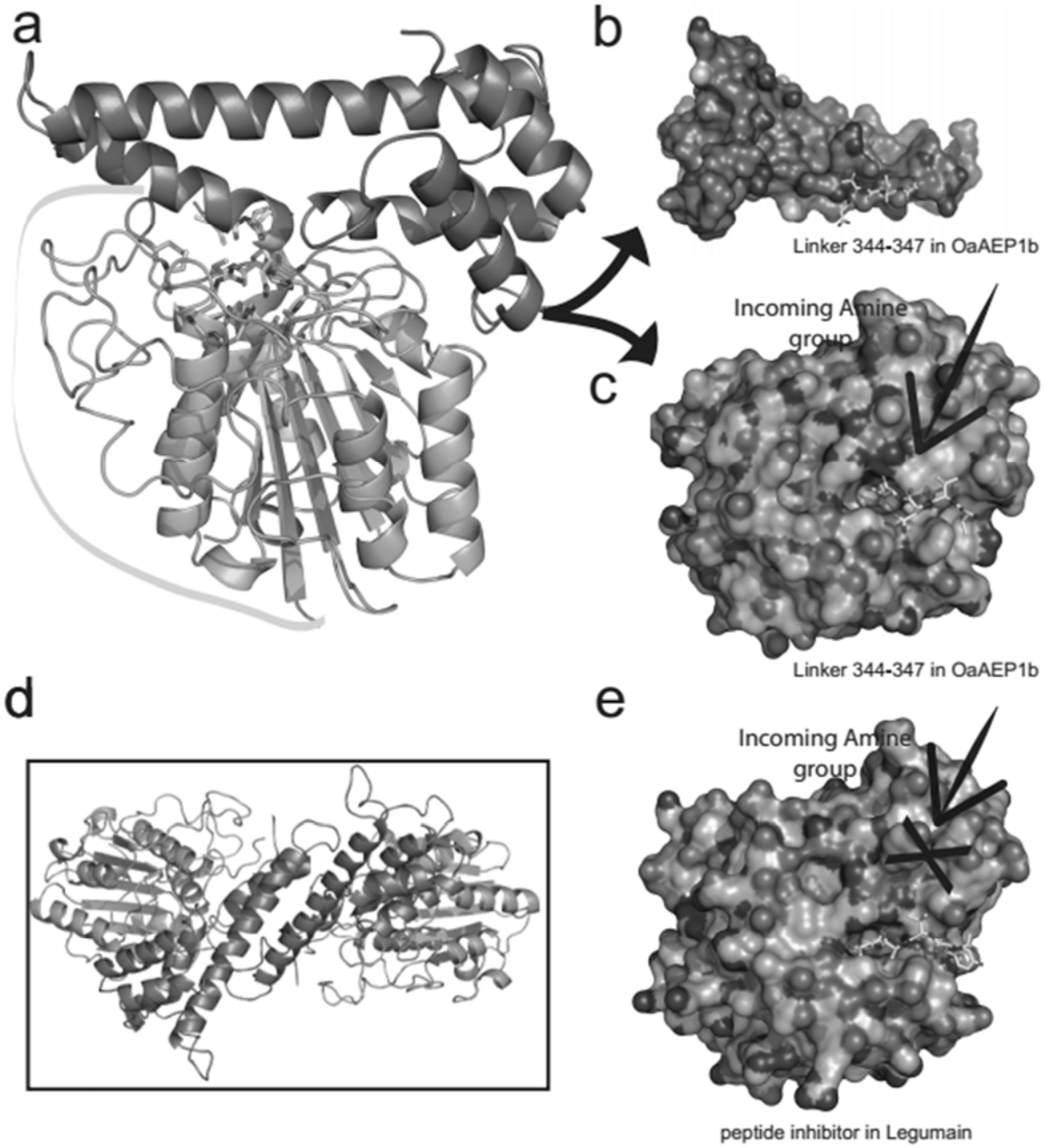

a
b
Linker 344-347 in OaAEP1b
Incoming Amine group
c
Linker 344-347 in OaAEP1b
d
e
Incoming Amine group
peptide inhibitor in Legumain

Figure 3f

f

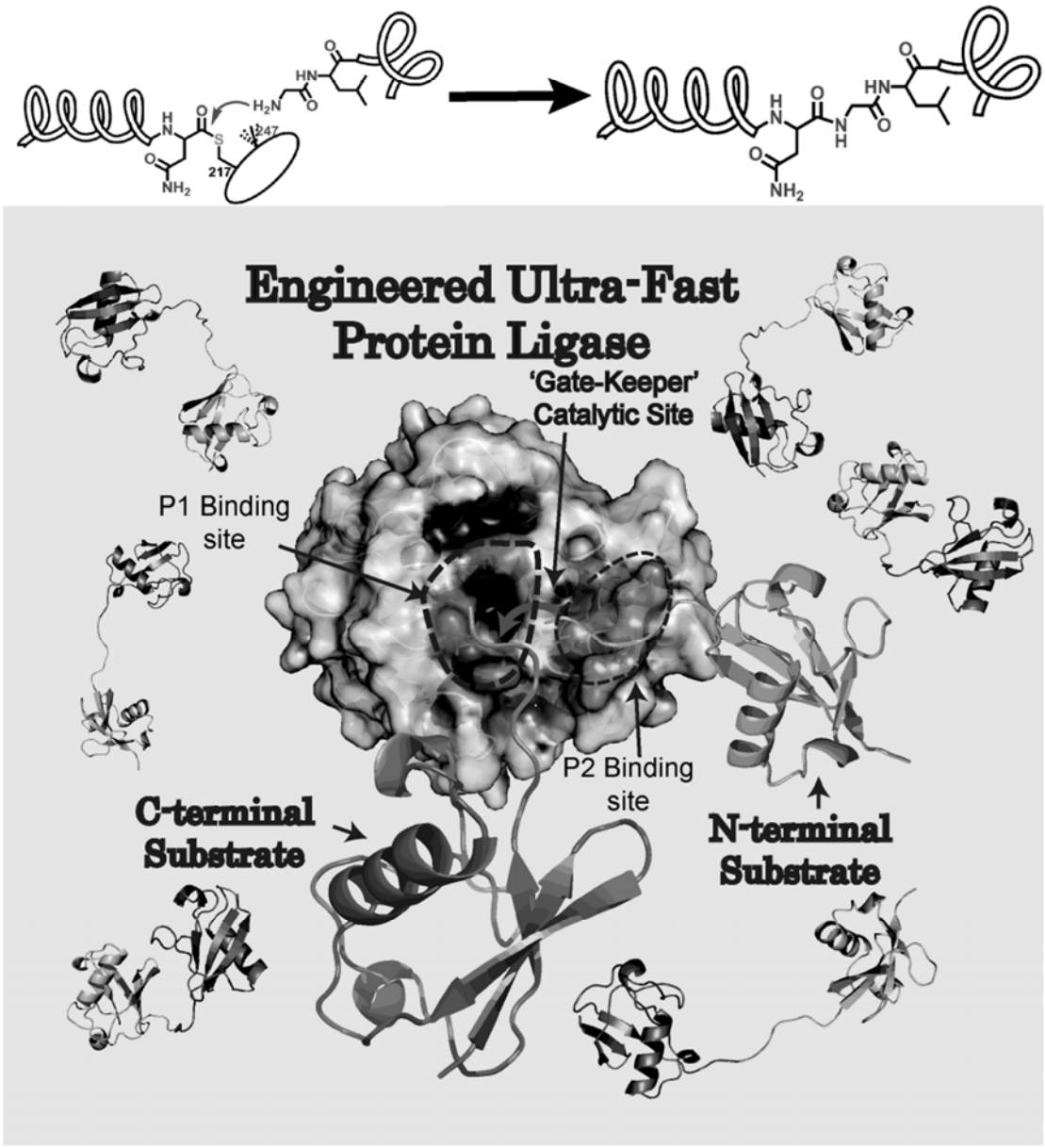

Engineered Ultra-Fast Protein Ligase
'Gate-Keeper' Catalytic Site
P1 Binding site
P2 Binding site
C-terminal Substrate
N-terminal Substrate

Figure 4

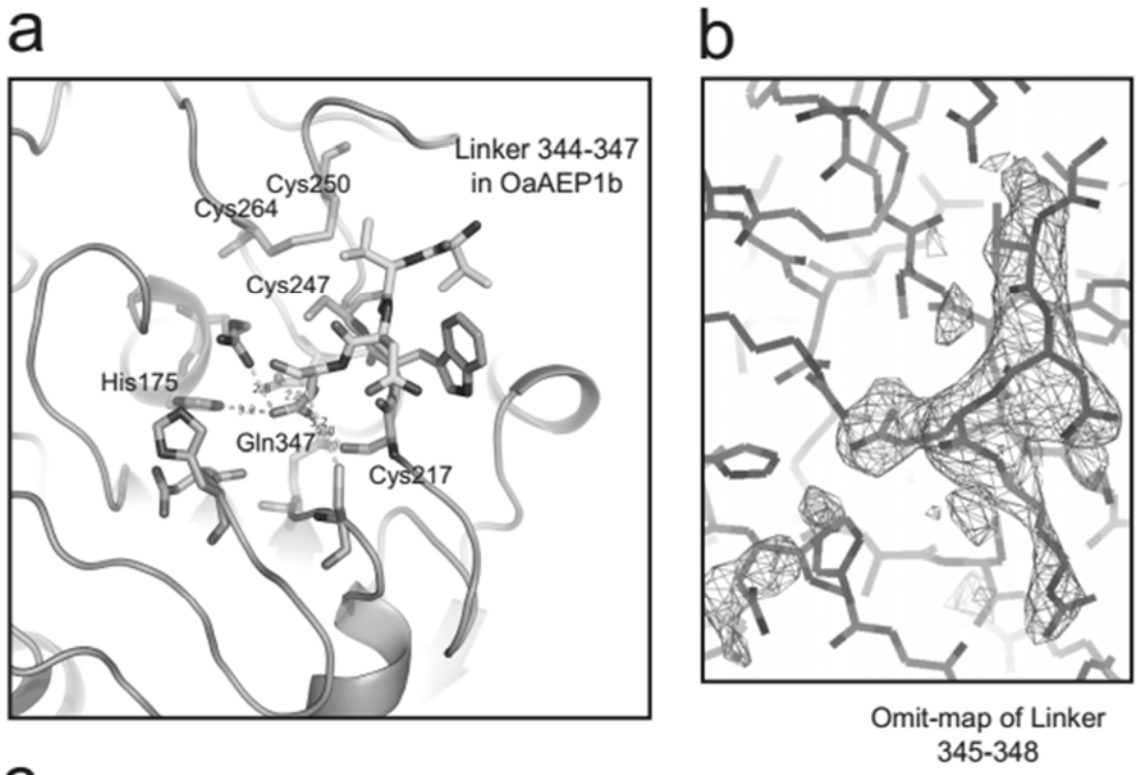
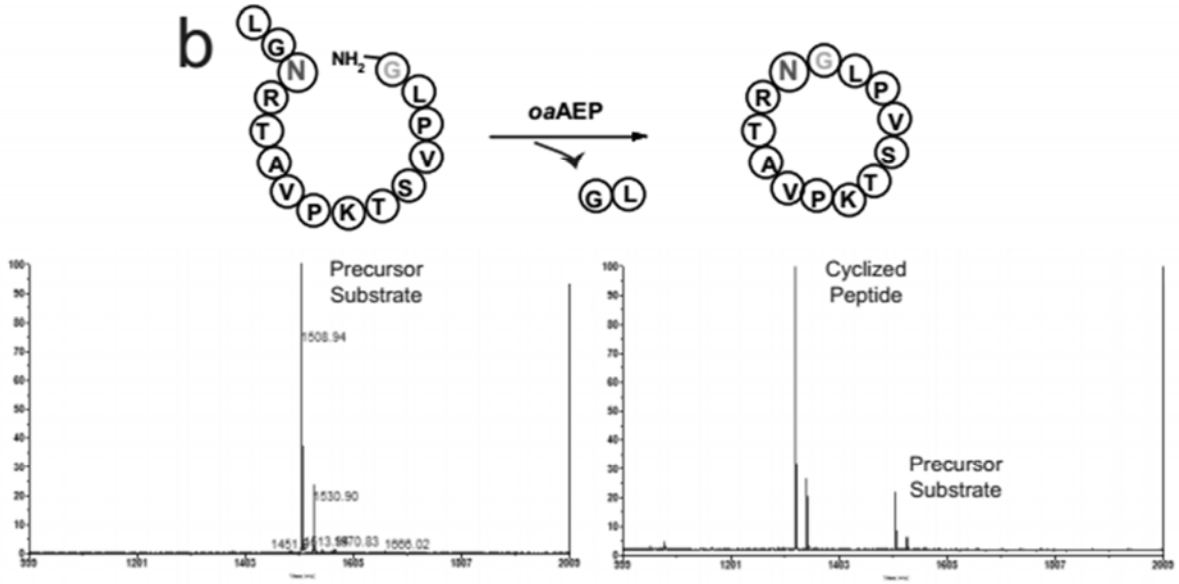

a
Linker 344-347
in OaAEP1b
Cys250
Cys264
Cys247
His175
Gln347
Cys217
b
Omit-map of Linker
345-348
c
WT
N346A
Q347A
NQ/AA
DD/AA
A3
A4
pH 3.7 activation
- + - + - + - + - + - + - + Marker
D
Activation efficiency
after 6 hours
100
80
60
40
20
0
WT
N346A
Q347A
NQ/AA
DD/AA
A3
A4

Figure 5

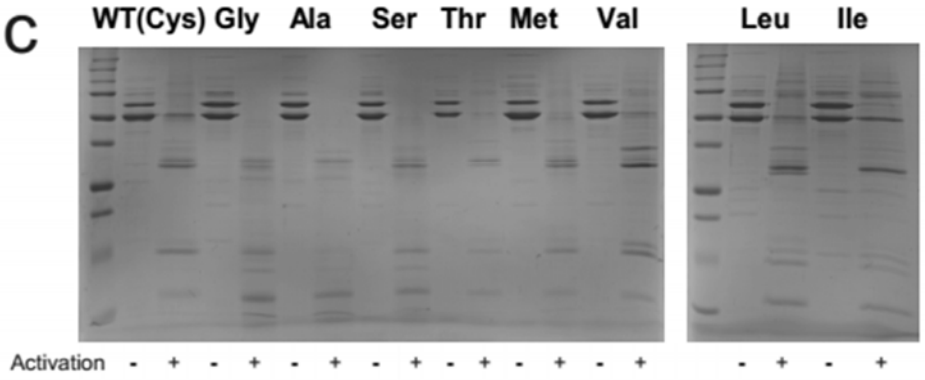
a

Ligase	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1} s^{-1}$)
WT AEP	0.052 ± 0.008	245 ± 60	215
AEP (Cys247Ala)	13.9 ± 1.1	407 ± 49	34,209
Butelase 1	4.83 ± 0.62	48.4 ± 14	99,793

b



c



d

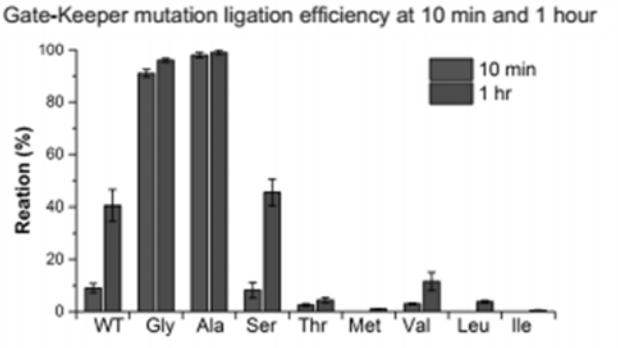


Figure 6

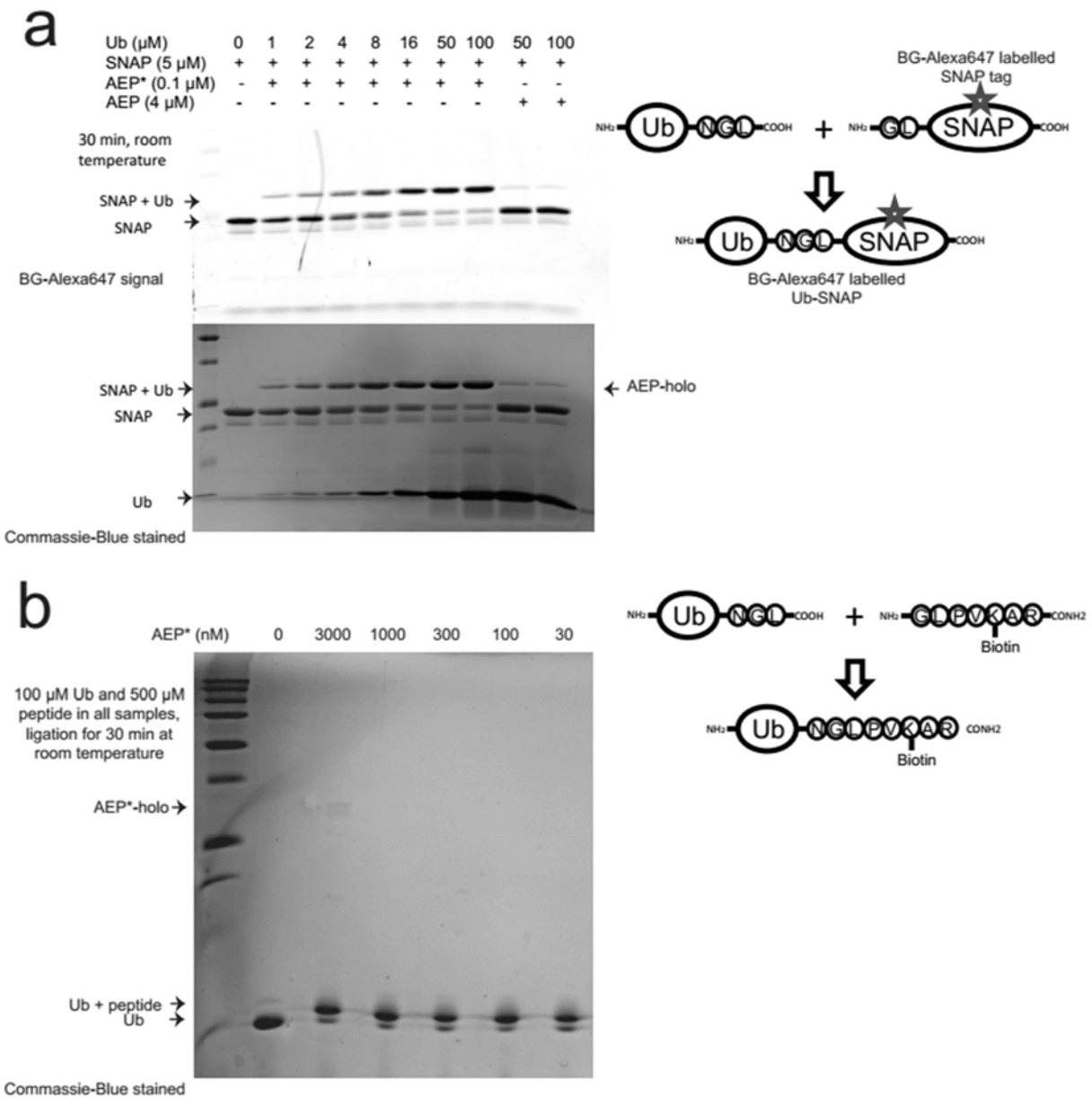
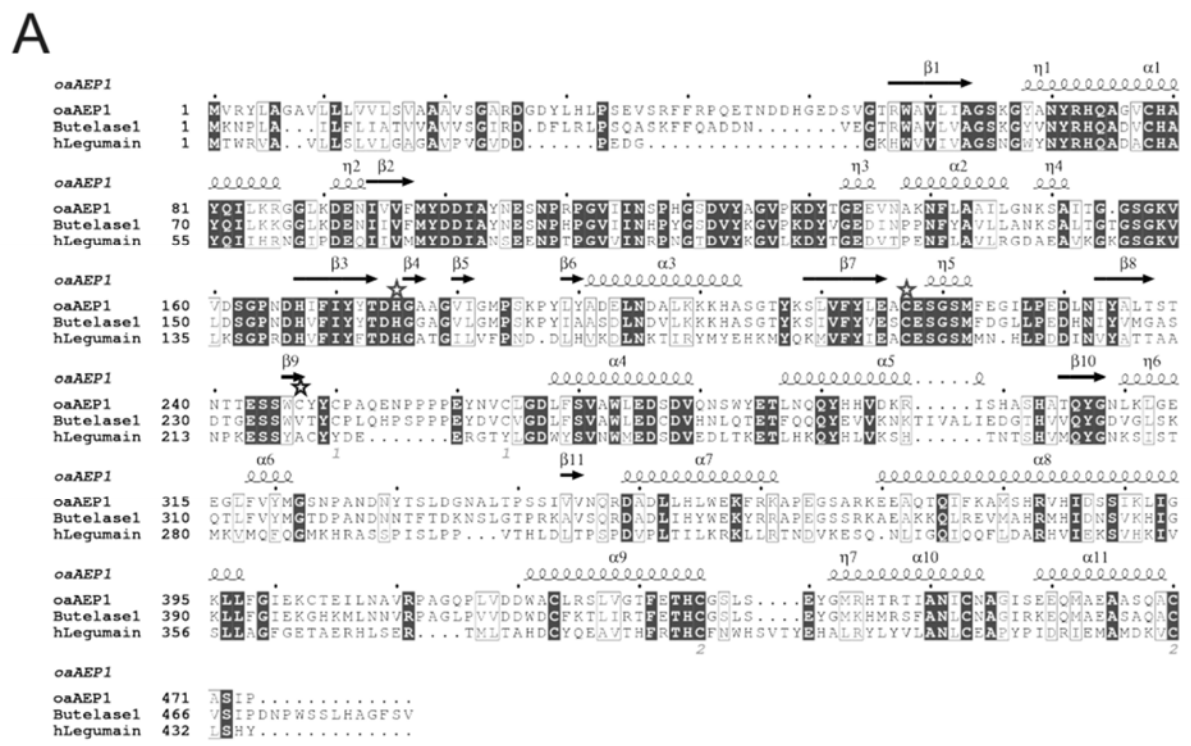

a
Ub (μM) 0 1 2 4 8 16 50 100 50 100
SNAP (5 μM) + + + + + + + + + +
AEP* (0.1 μM) - + + + + + + + - -
AEP (4 μM) - - - - - - - - + +
30 min, room temperature
SNAP + Ub
SNAP
BG-Alexa647 signal
BG-Alexa647 labelled SNAP tag
Ub
NGL
GL
SNAP
BG-Alexa647 labelled Ub-SNAP
SNAP + Ub
SNAP
Ub
AEP-holo
Commassie-Blue stained
b
AEP* (nM) 0 3000 1000 300 100 30
100 μM Ub and 500 μM peptide in all samples, ligation for 30 min at room temperature
AEP*-holo
Ub + peptide
Ub
GLPVKAR
CONH2
Biotin
Commassie-Blue stained

Figure 7


A
oaAEP1
oaAEP1 1
Butelase1 1
hLegumain 1
oaAEP1 81
Butelase1 70
hLegumain 55
oaAEP1 160
Butelase1 150
hLegumain 135
oaAEP1 240
Butelase1 230
hLegumain 213
oaAEP1 315
Butelase1 310
hLegumain 280
oaAEP1 395
Butelase1 390
hLegumain 356
oaAEP1 471
Butelase1 466
hLegumain 432


B
-85 MGMAHHHHHHMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQ
-36 RLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGARDGDYLHLPSEVSRF
40 FRPQETNDDHGEDSVGTRWAVLIAGSKGYANYRHQAGVCHAYQILKRGG
89 LKDENIVVFMYDDIAYNESNPRPGVIINSPHGSDVYAGVPKDYTGEEVNAK
140 NFLAAILGNKSAITGGSGKVVDSGPNDHIFIYYTDHGAAGVIGMPSKPYLY
191 ADELNDALKKKHASGTYKSLVFYLEACESGSMFEGILPEDLNIYALTSTNTT
243 ESSWCYYCPAQENPPPPEYNVCLGDLFSVAWLEDSDVQNSWYETLNQQY
292 HHVDKRISHASHATQYGNLKLGEEGLFVYMGSNPANDNYTSLDGNALTPS
342 SIVVNQRDADLLLHLWEKFRKAPEGSARKEEAQTQIFKAMSHRVHIDSSIKL
393 IGKLLFGIEKCTEILNAVRPAGQPLVDDWACLRSLVGTFETHCGSLSEYGM
443 RHTRTIANICNAGISEEQMAEAASQACASIP
Expression Fusion tags (-85 to 0)
Core Domain (24 to 325)
Linker (326 to 347)
Cap Domain (348 to 474)

Figure 8

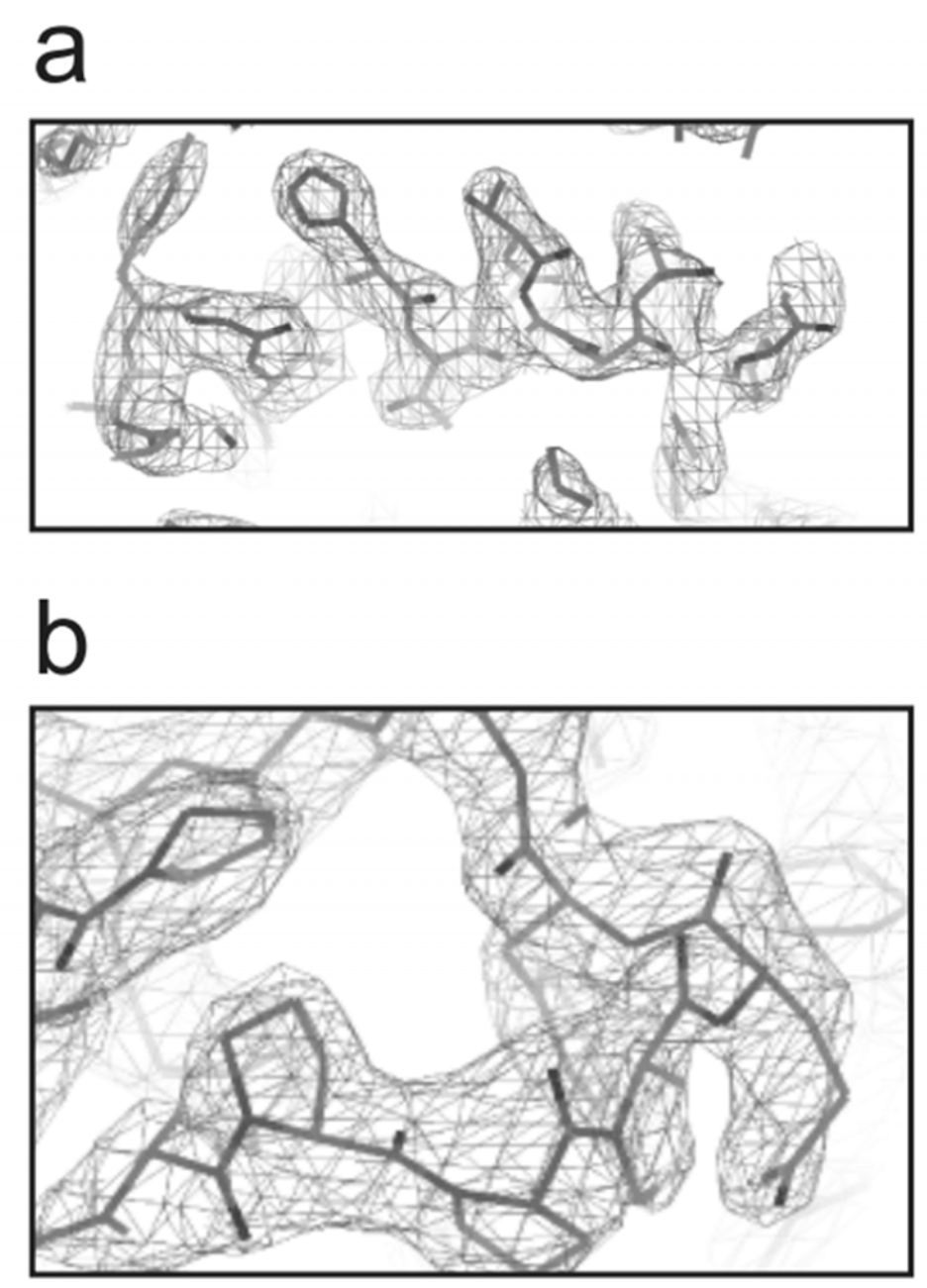
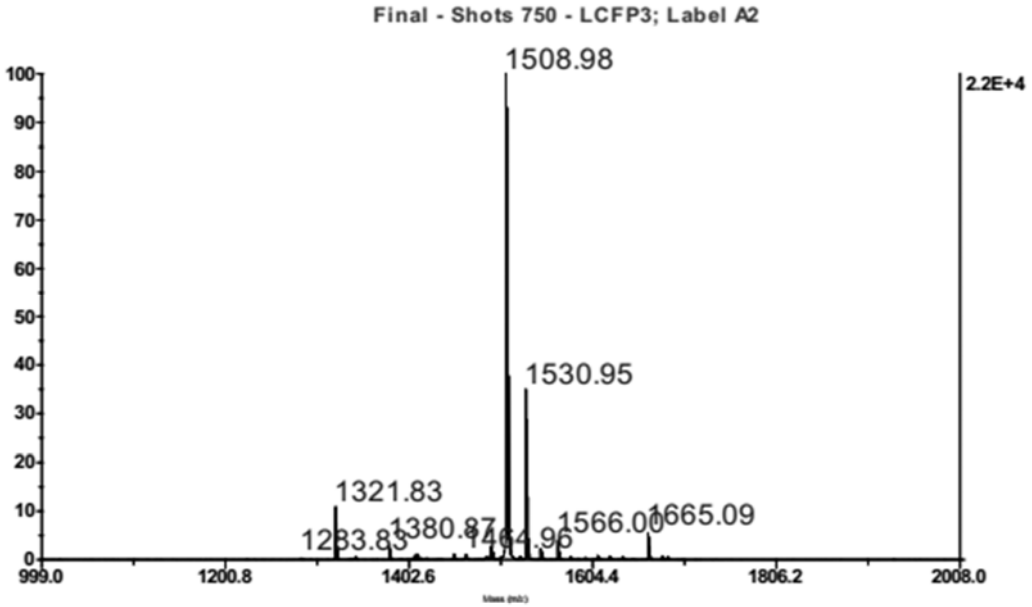
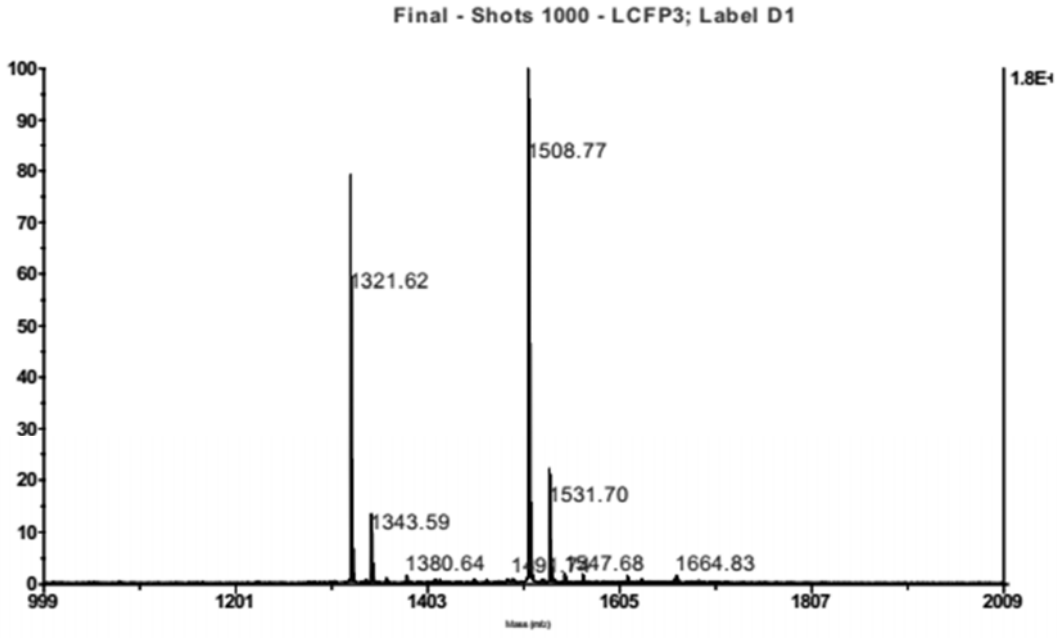


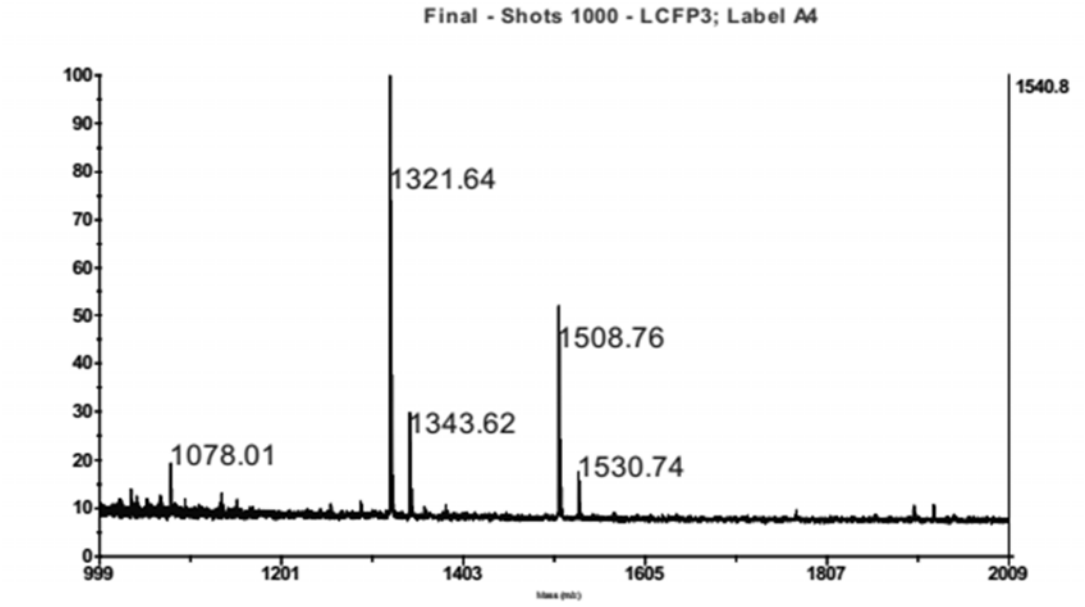
Figure 9



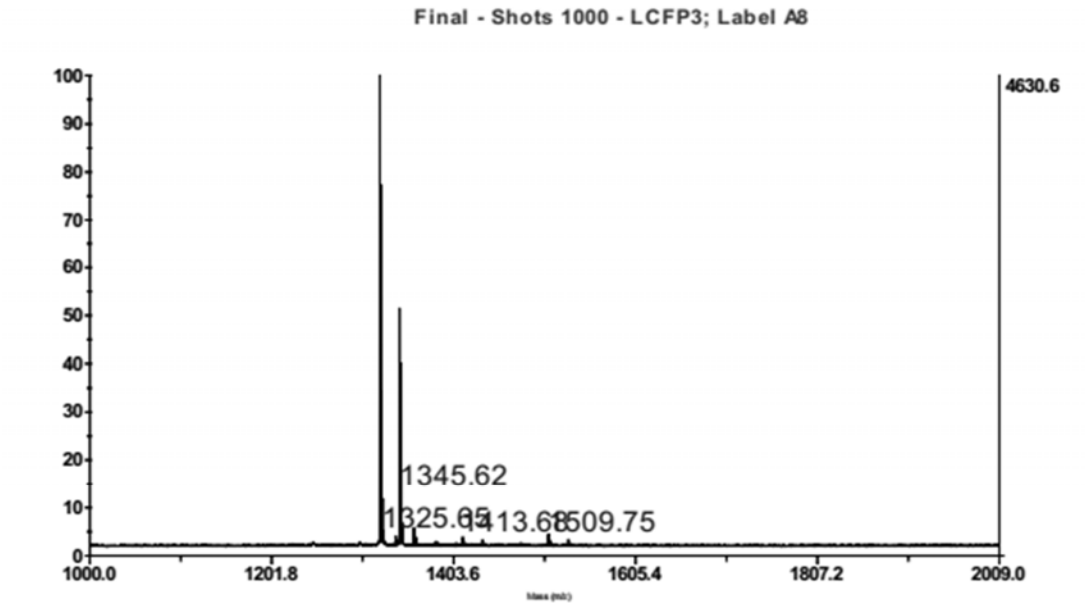
WT, 10MIN



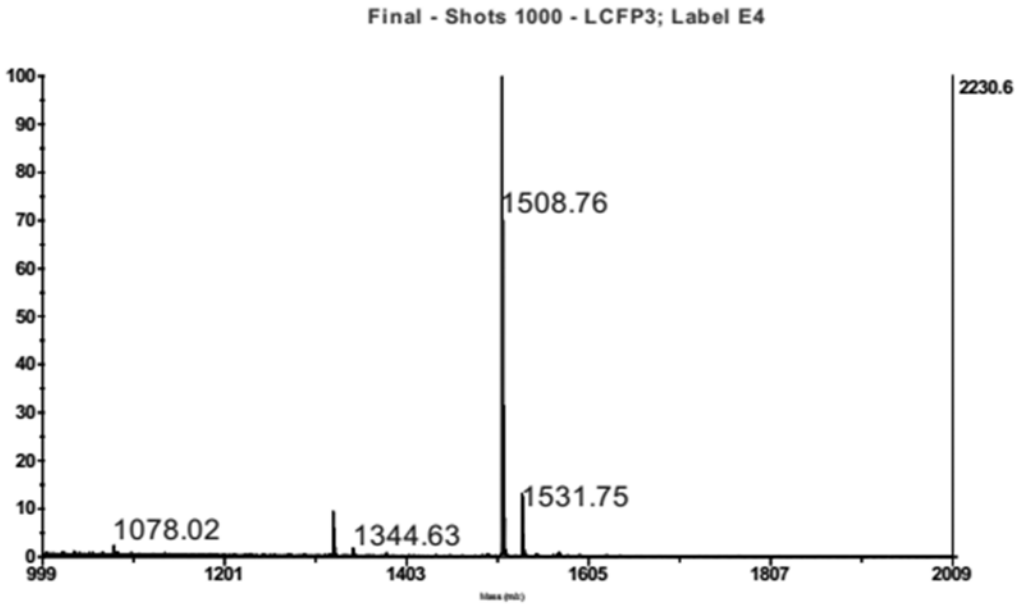
AEP WT, 1H



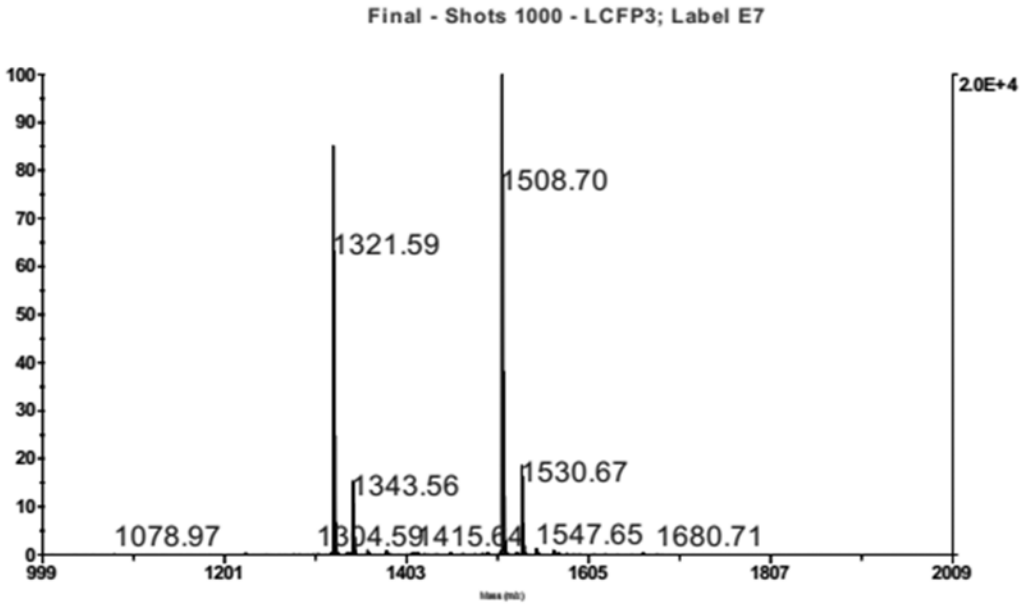
Gly, 10 min



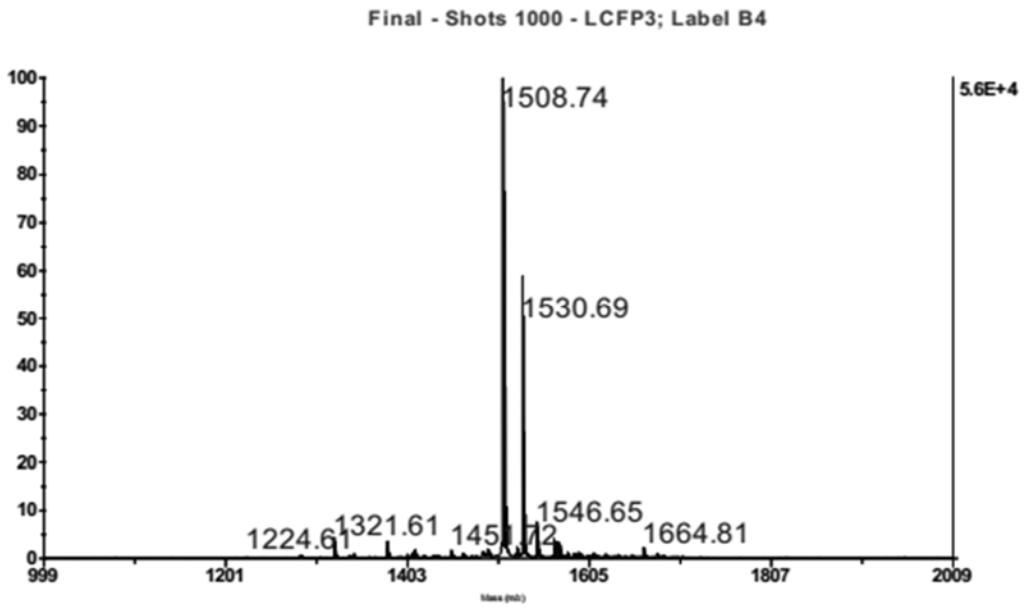
Gly, 1 h



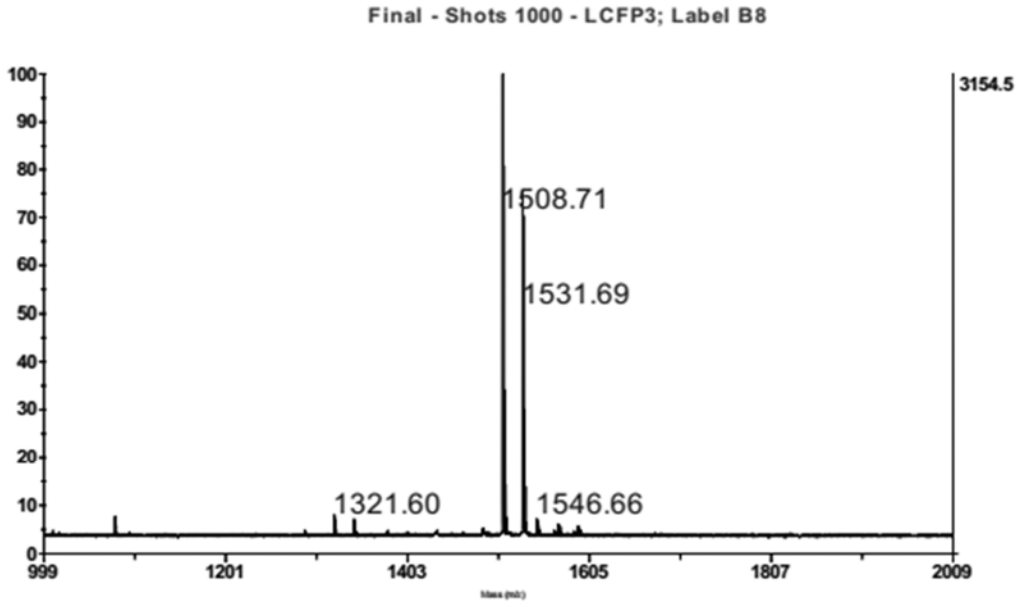
Ser, 10 min



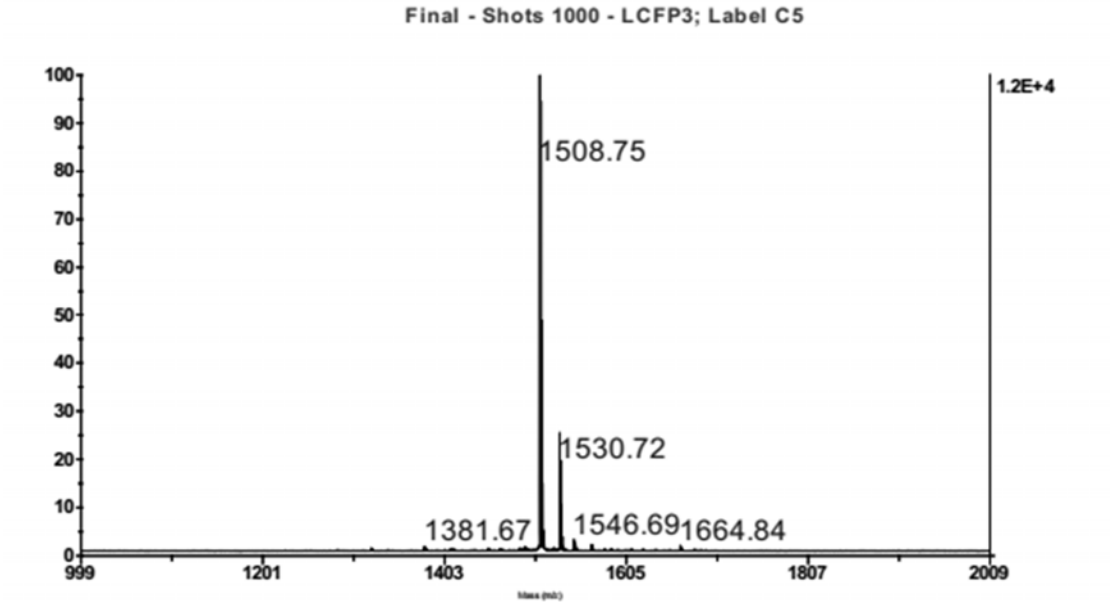
Ser, 1 h



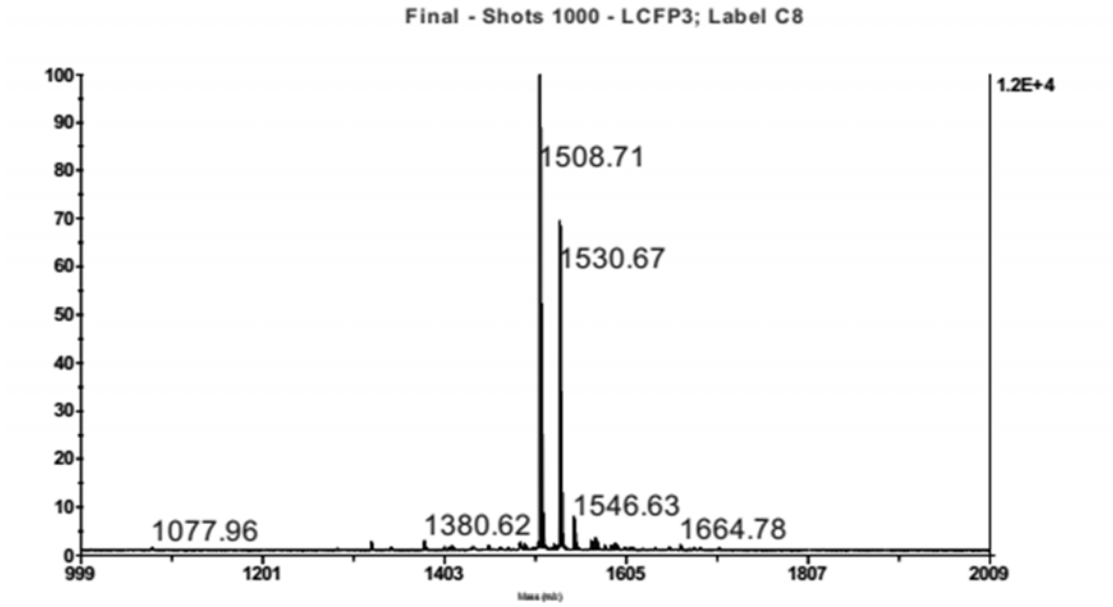
Thr, 10 min



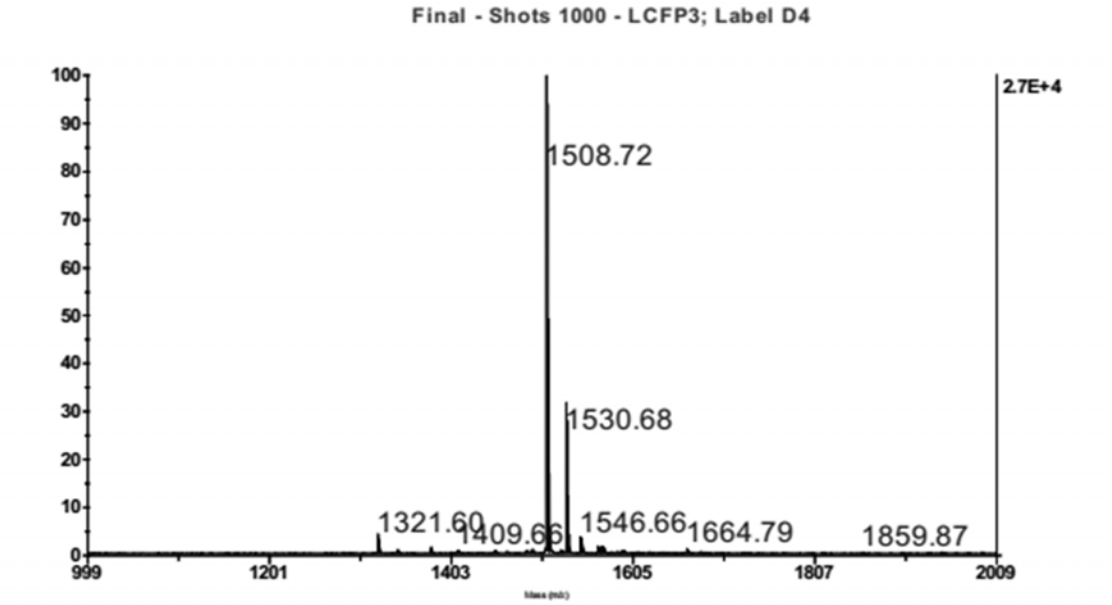
Thr, 1 h



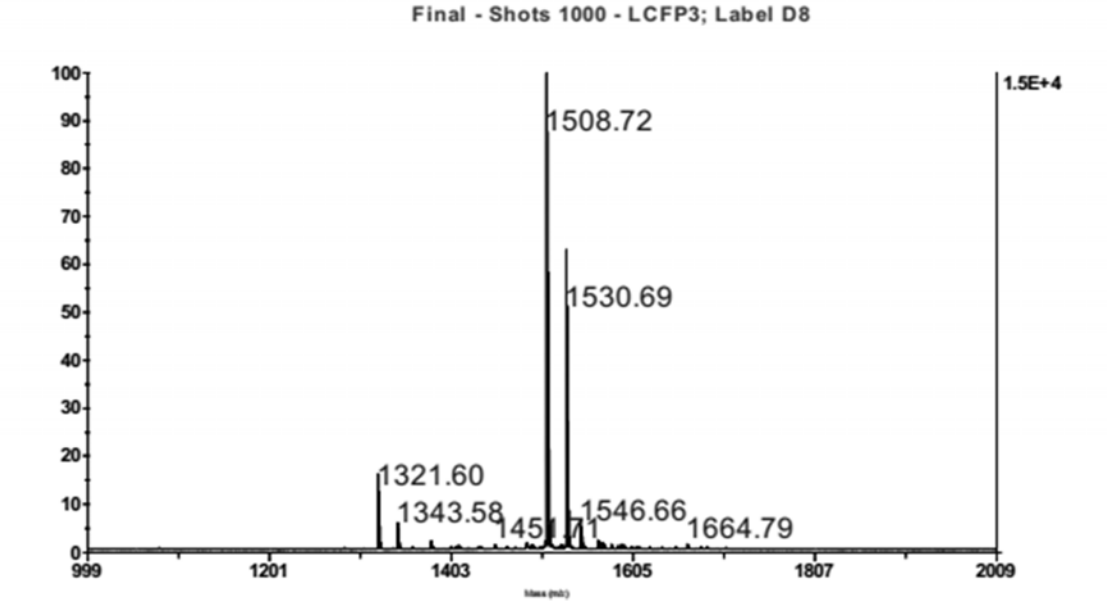
Met, 10 min



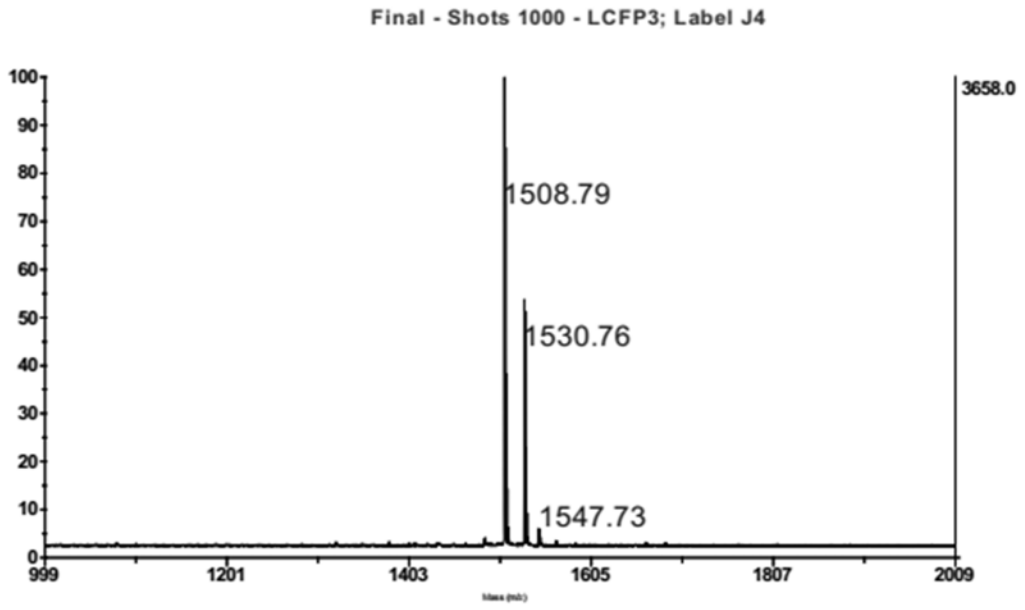
Met, 1 h



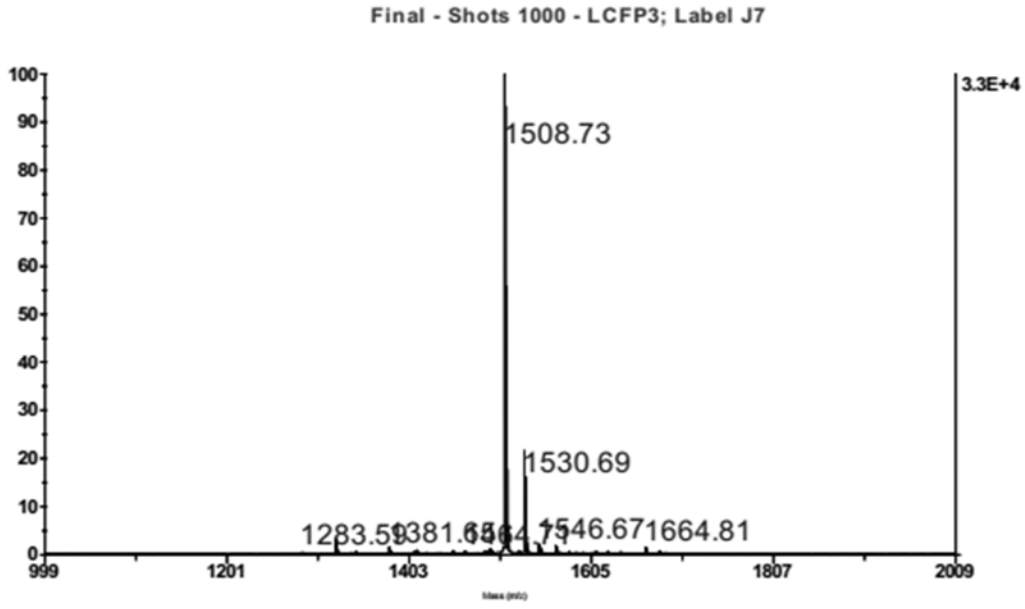
Val, 10 min



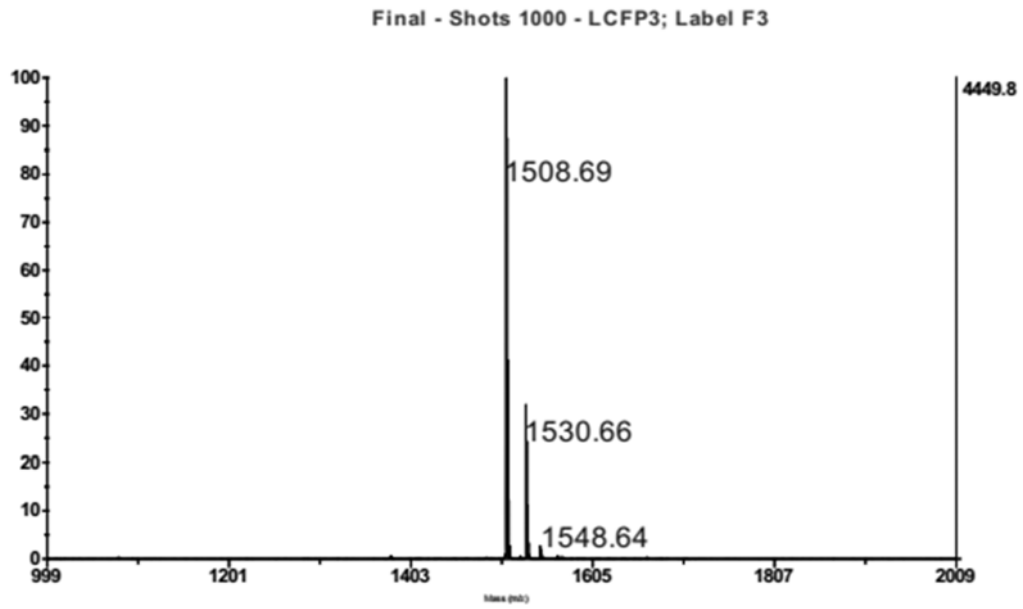
Val, 1 h



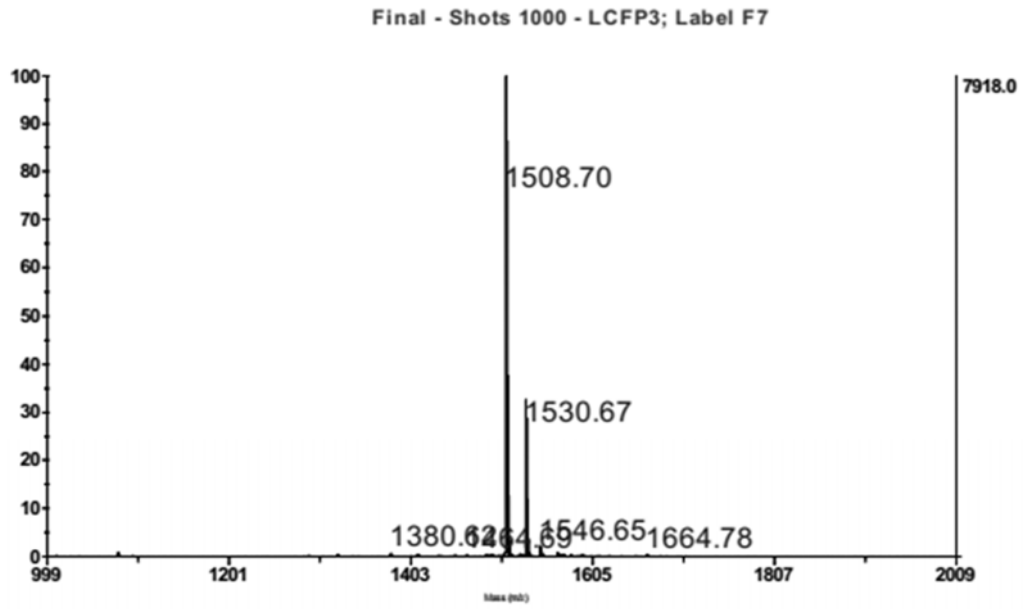
Leu, 10 min



Leu, 1 h



Ile, 10 min



Ile, 1 h

Figure 10

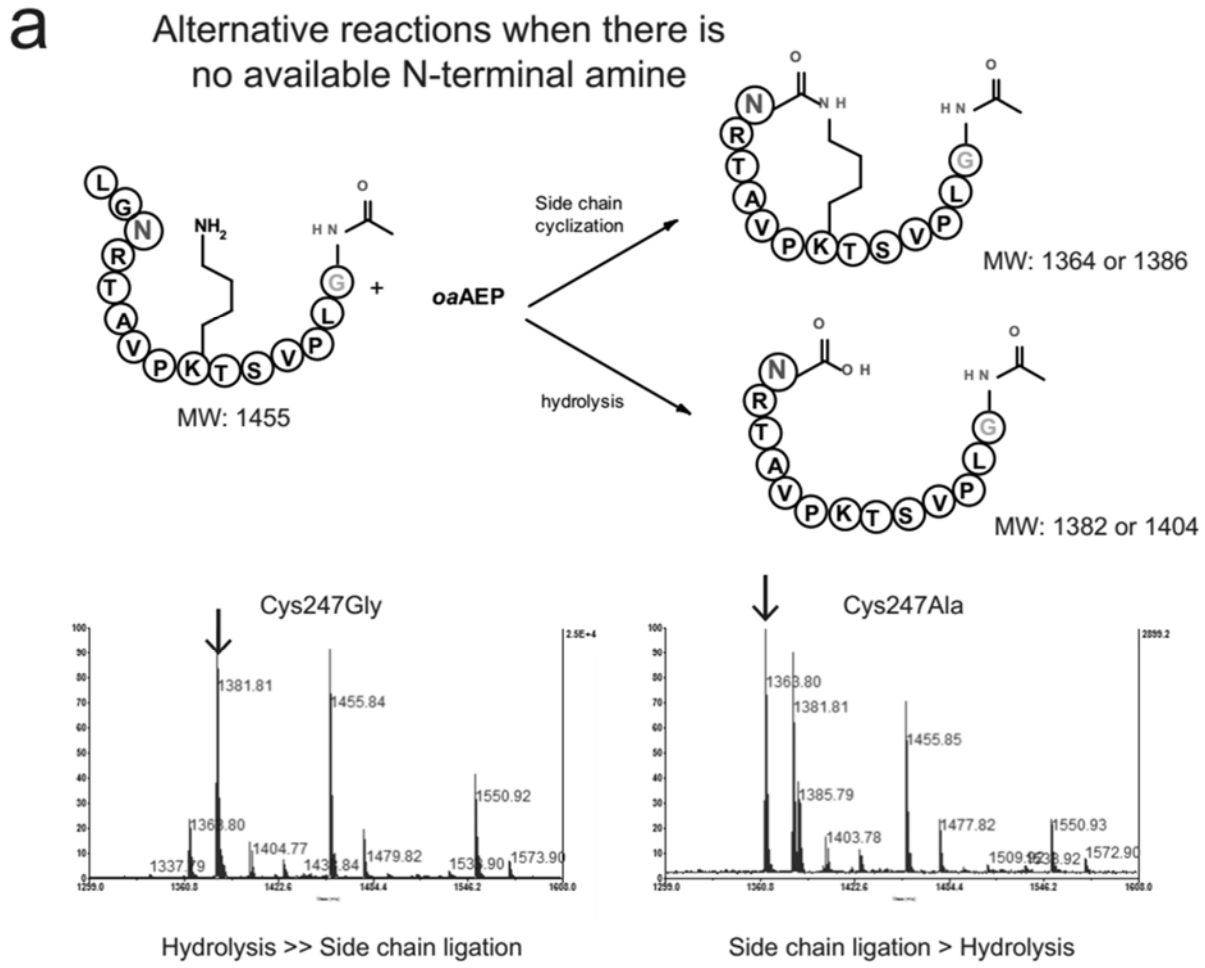
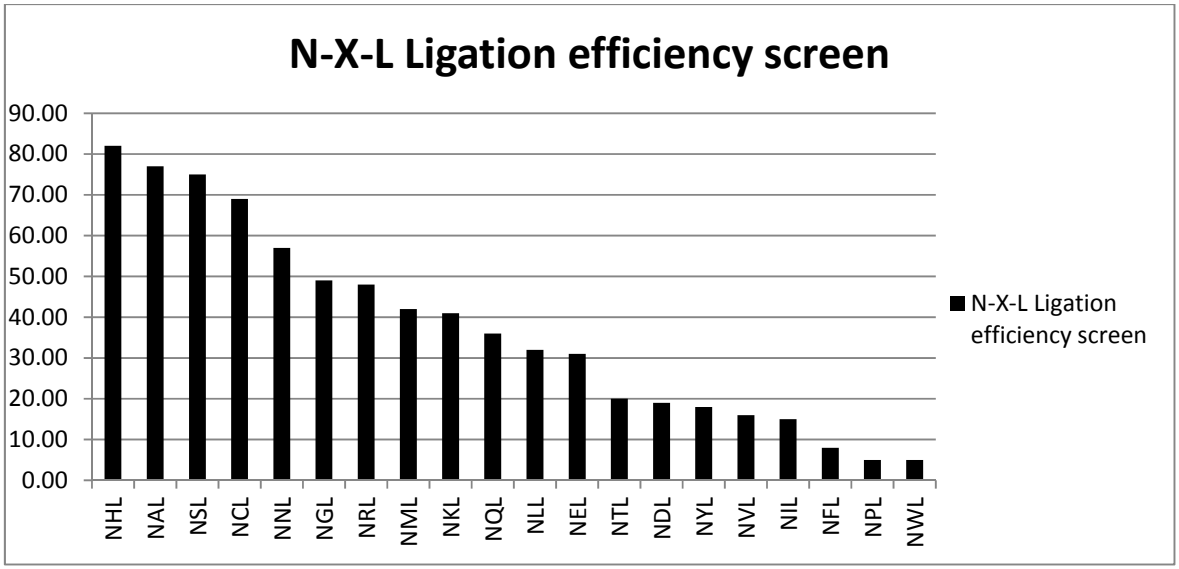
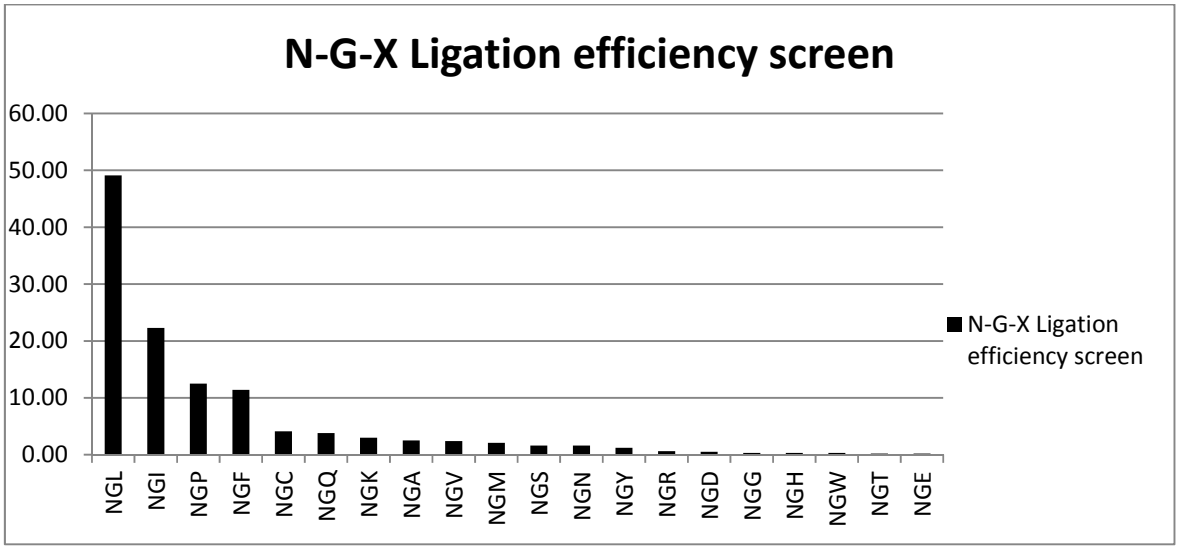

a
Alternative reactions when there is no available N-terminal amine
MW: 1455
oaAEP
Side chain cyclization
hydrolysis
MW: 1364 or 1386
MW: 1382 or 1404
Cys247Gly
1381.81
1455.84
1550.92
1368.80
1404.77
1479.82
1573.90
Cys247Ala
1368.80
1381.81
1455.85
1385.79
1403.78
1477.82
1550.93
1572.90
Hydrolysis >> Side chain ligation
Side chain ligation > Hydrolysis

Figure 11

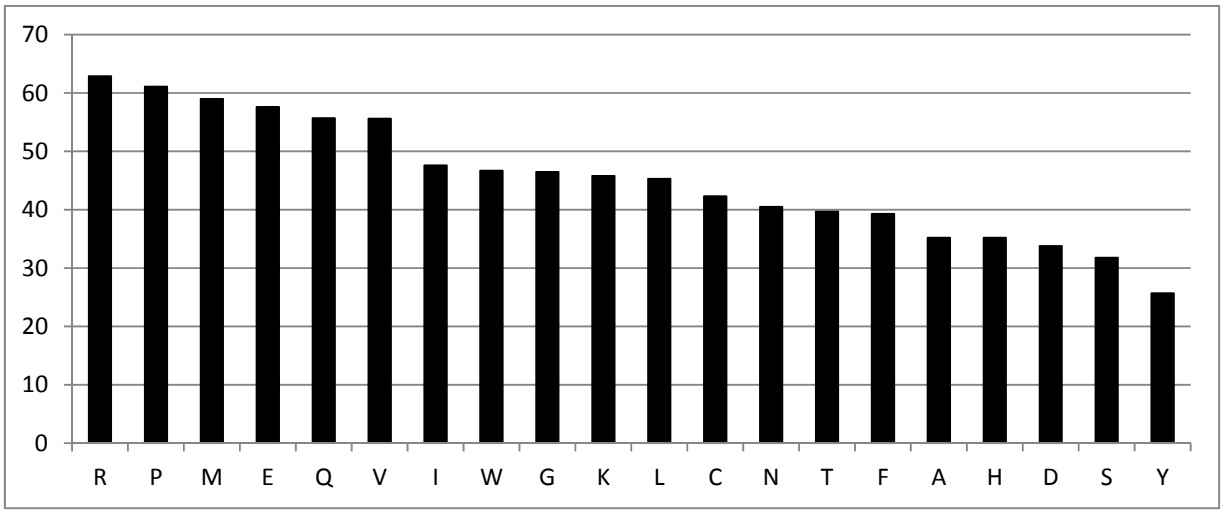
a


N-X-L Ligation efficiency screen
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
NHL
NAL
NSL
NCL
NNL
NGL
NRL
NML
NKL
NQL
NLL
NEL
NTL
NDL
NYL
NVL
NIL
NFL
NPL
NWL
N-X-L Ligation efficiency screen

b

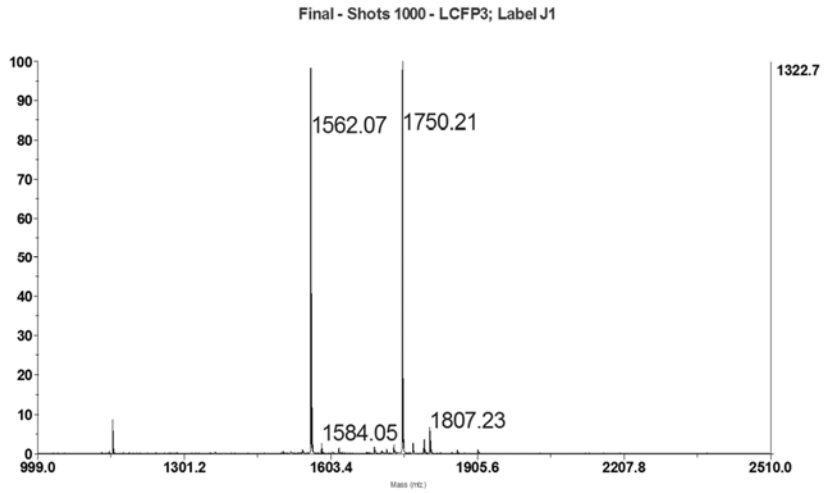

N-G-X Ligation efficiency screen
60.00
50.00
40.00
30.00
20.00
10.00
0.00
NGL
NGI
NGP
NGF
NGC
NGQ
NGK
NGA
NGV
NGM
NGS
NGN
NGY
NGR
NGD
NGG
NGH
NGW
NGT
NGE
N-G-X Ligation efficiency screen

c

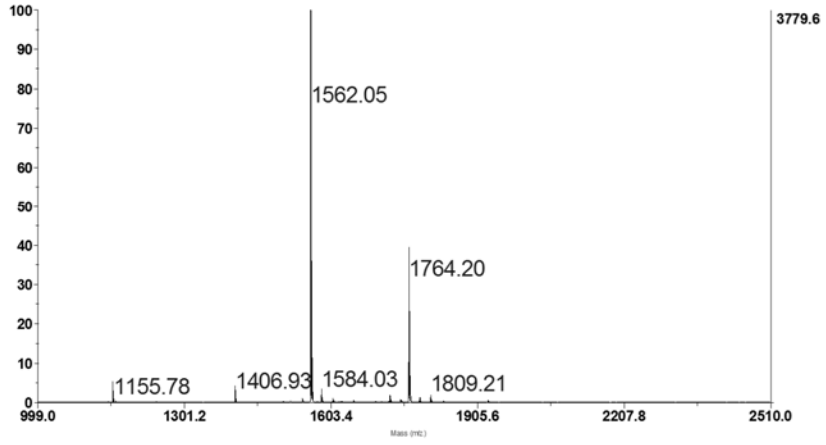


d

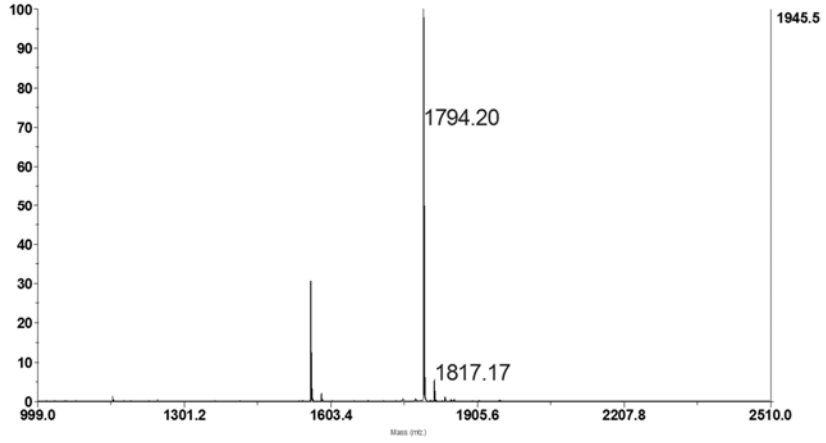
GLYRRGRLYRRNXL (J1-J20: NGL, NAL, NTL, NSL, NYL, NCL, NEL, NPL, NLL, NIL, NNL, NWL, NRL, NHL, NDL, NVL, NML, NKL, NQL, NFL, total 20 peptides)



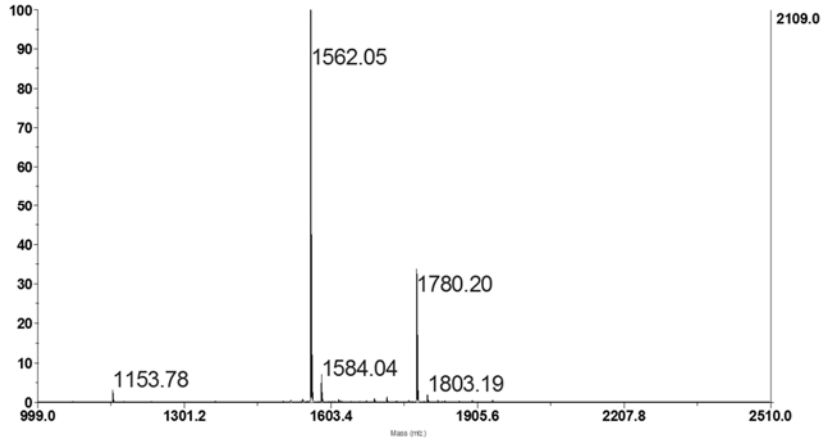
Final - Shots 1000 - LCFP3; Label J2



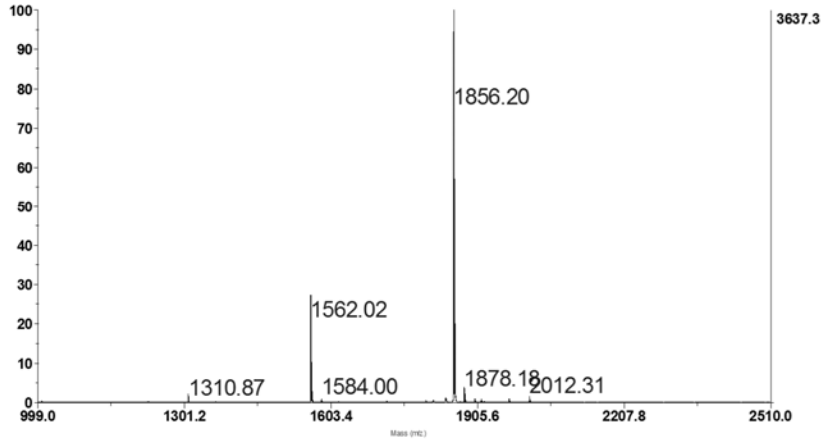
Final - Shots 1000 - LCFP3; Label J3



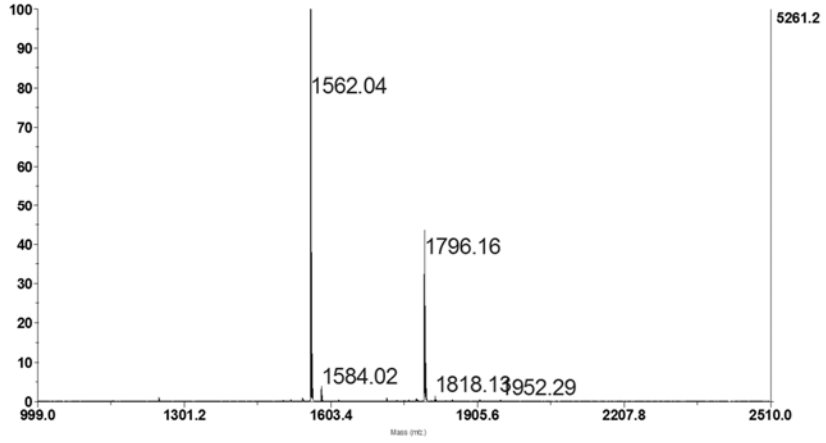
Final - Shots 1000 - LCFP3; Label J4



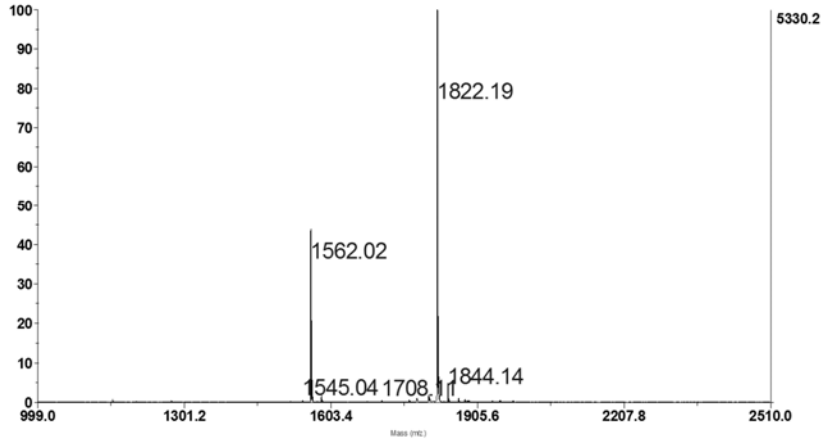
Final - Shots 1000 - LCFP3; Label J5



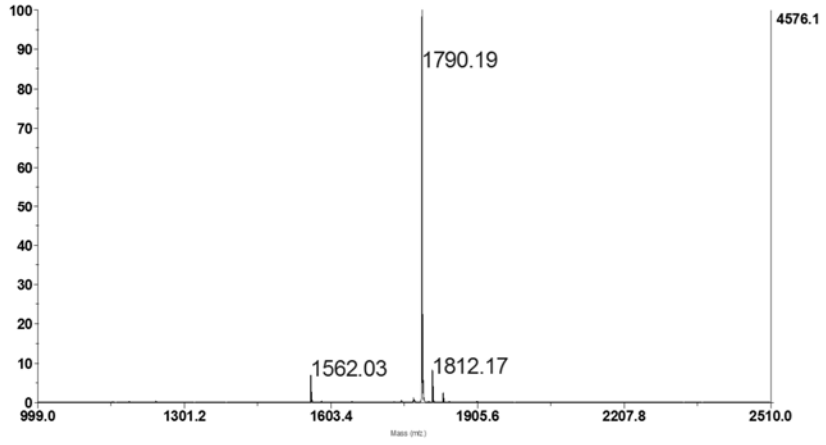
Final - Shots 1000 - LCFP3; Label J6



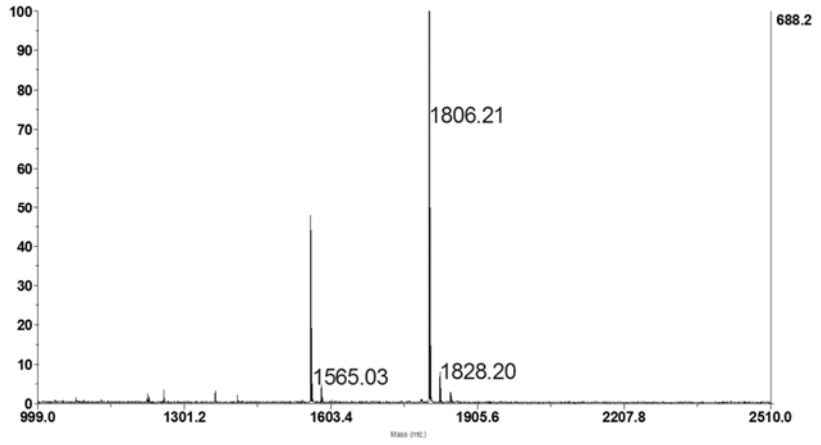
Final - Shots 1000 - LCFP3; Label J7



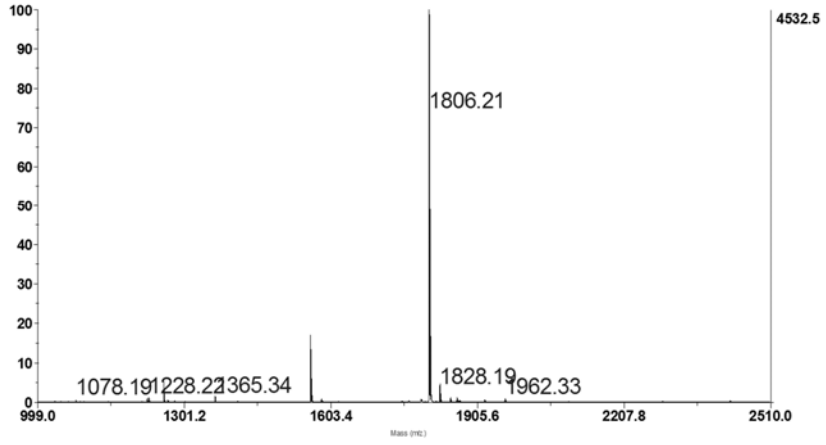
Final - Shots 1000 - LCFP3; Label J8



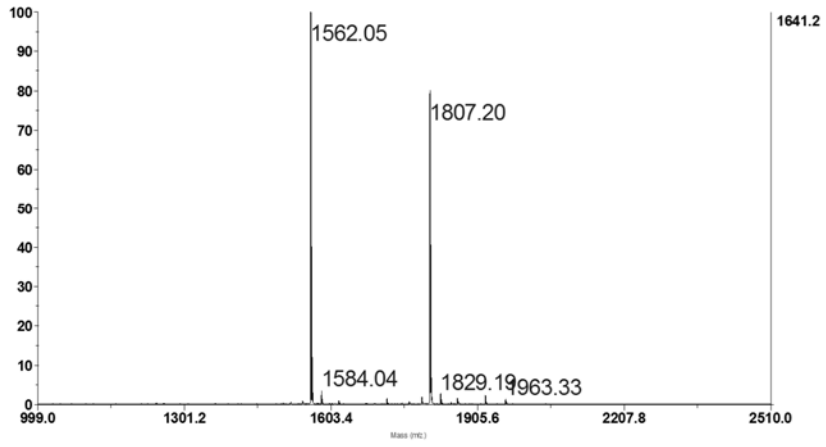
Final - Shots 1000 - LCFP3; Label J9



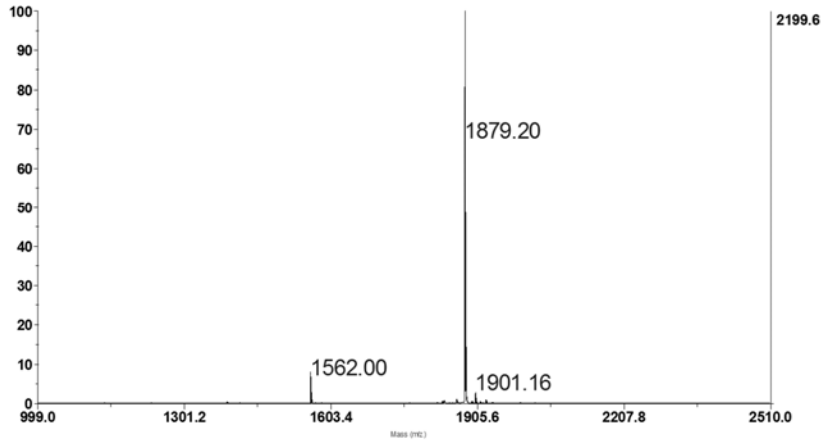
Final - Shots 1000 - LCFP3; Label J10



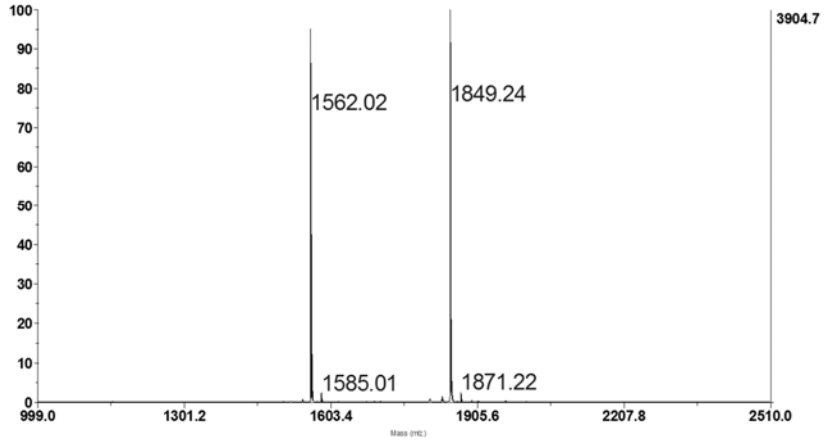
Final - Shots 1000 - LCFP3; Label J11



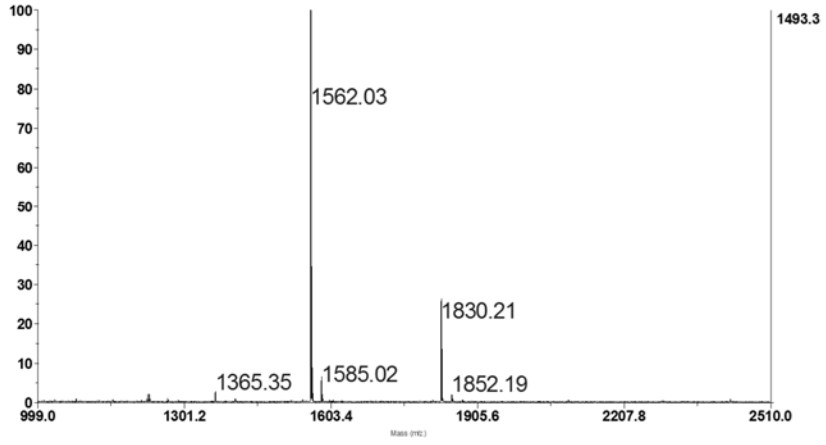
Final - Shots 1000 - LCFP3; Label J12



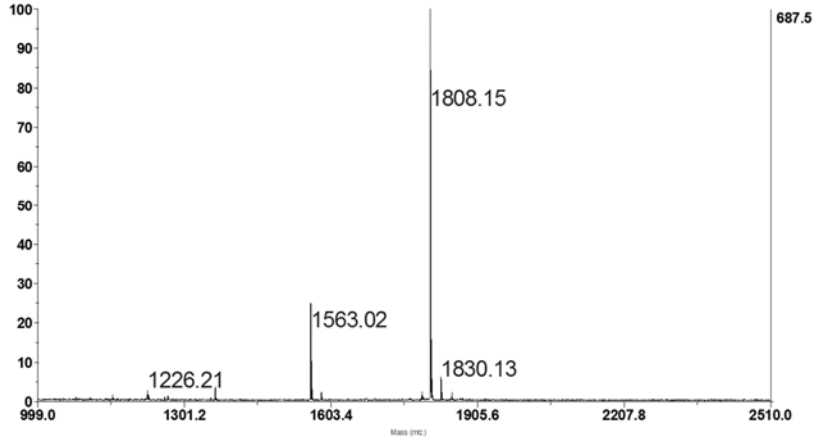
Final - Shots 1000 - LCFP3; Label J13



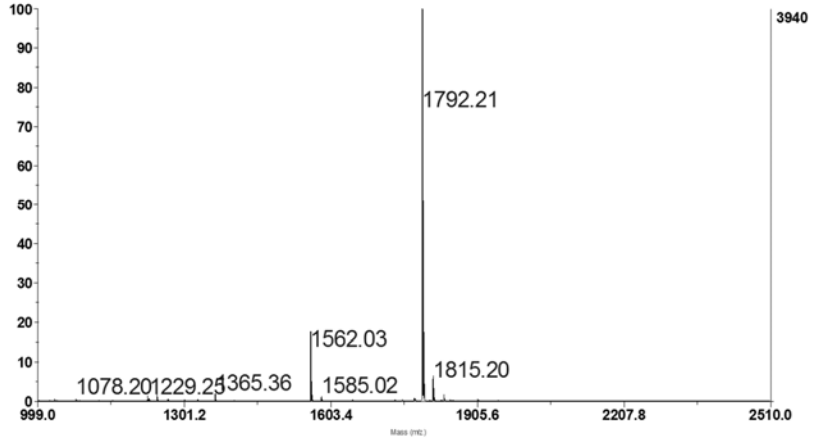
Final - Shots 1000 - LCFP3; Label J14



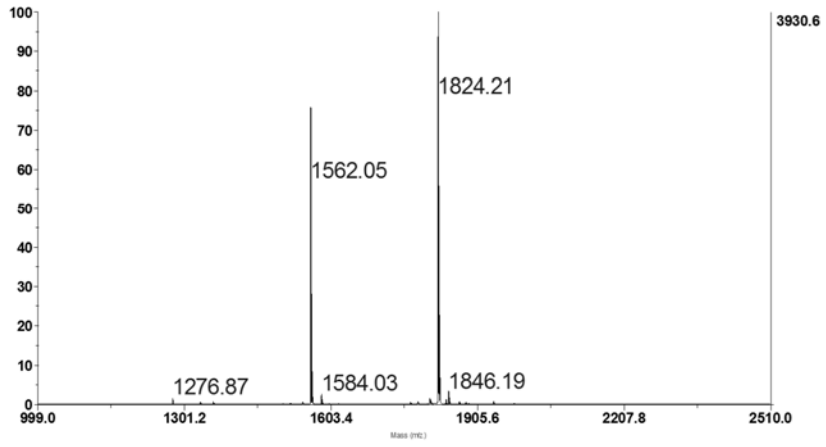
Final - Shots 1000 - LCFP3; Label J15



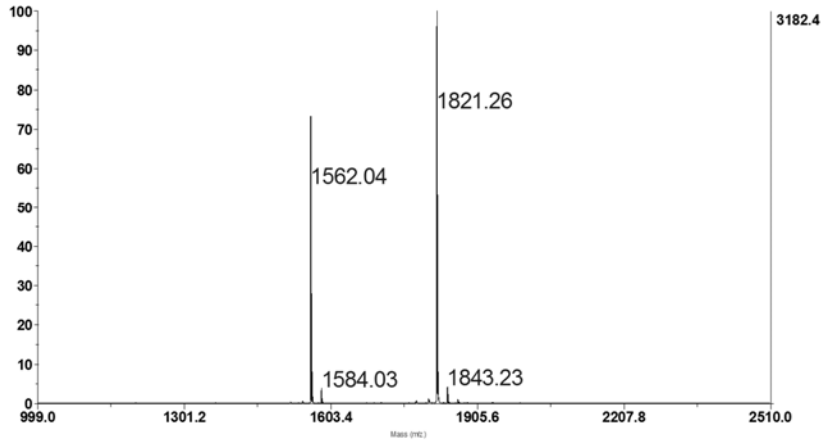
Final - Shots 1000 - LCFP3; Label J16



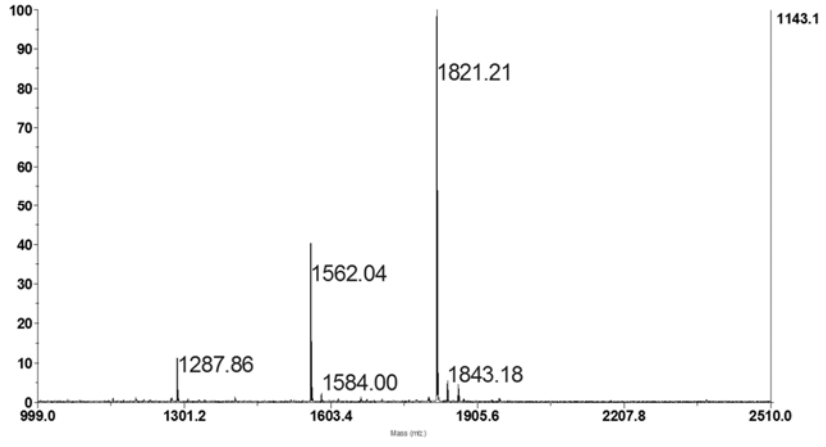
Final - Shots 1000 - LCFP3; Label J17

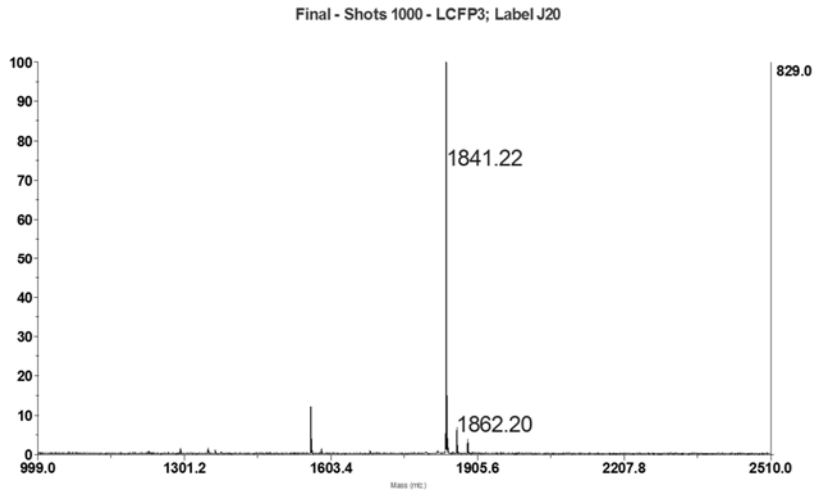


Final - Shots 1000 - LCFP3; Label J18

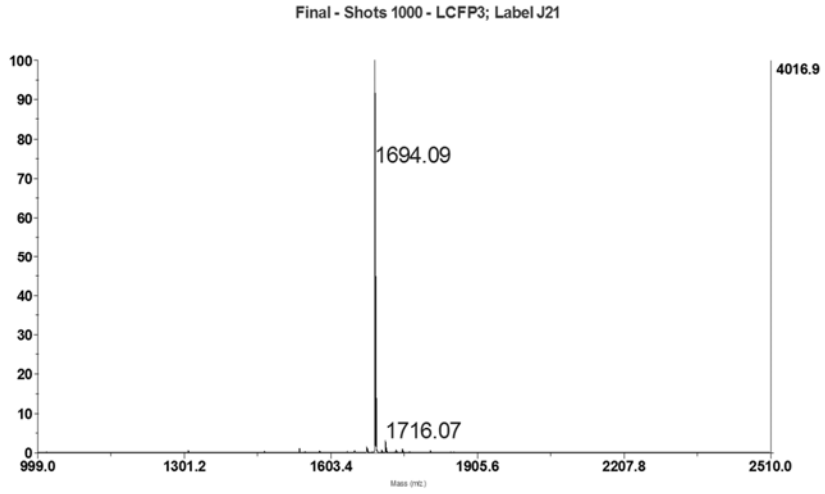


Final - Shots 1000 - LCFP3; Label J19

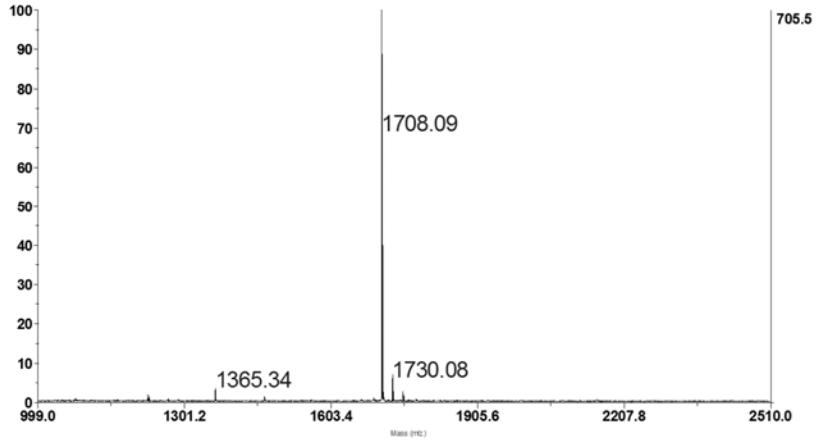




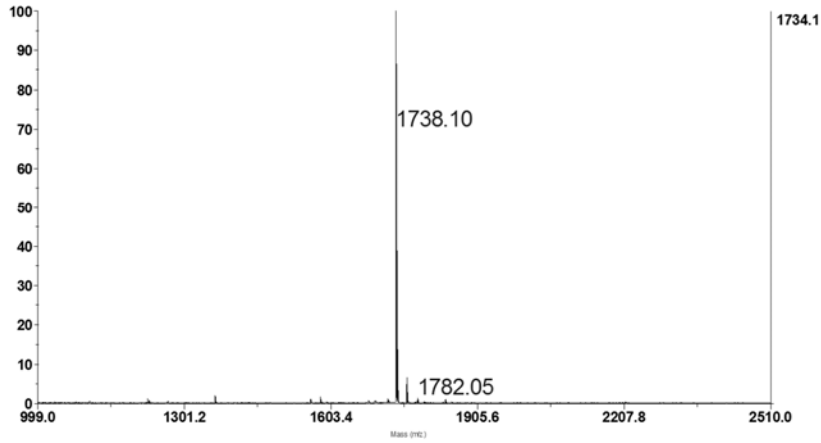
GLYRRGRLYRRNGX (J21-L16: NGG, NGA, NGT, NGS, NGY, NGC, NGE, NGP, NGI, NGN, NGW, NGR, NGH, NGD, NGV, NGM, NGK, NGQ, NGF total 19 peptides, NGL is the same peptide as the first plot-J1)



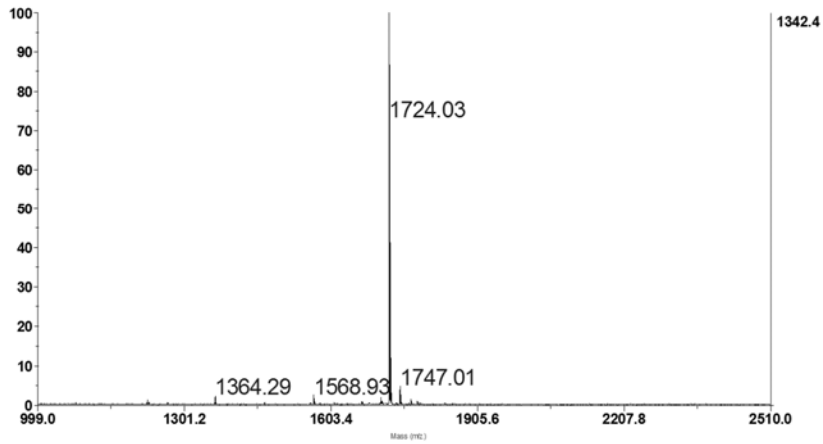
Final - Shots 1000 - LCFP3; Label J22



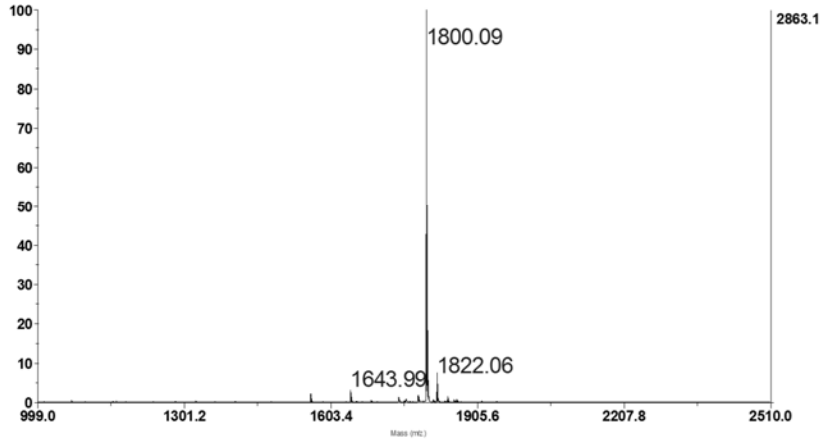
Final - Shots 1000 - LCFP3; Label J23



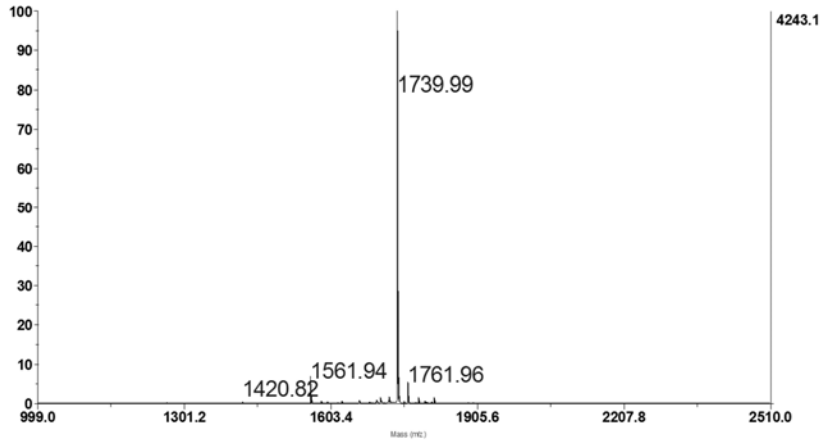
Final - Shots 1000 - LCFP3; Label K1



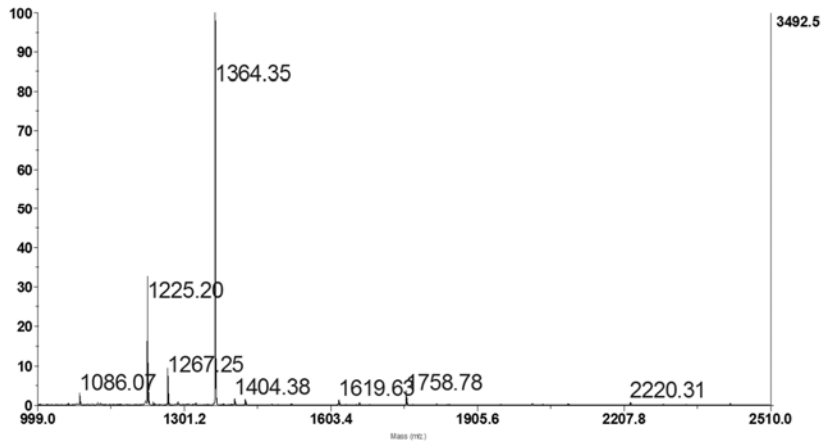
Final - Shots 1000 - LCFP3; Label L1



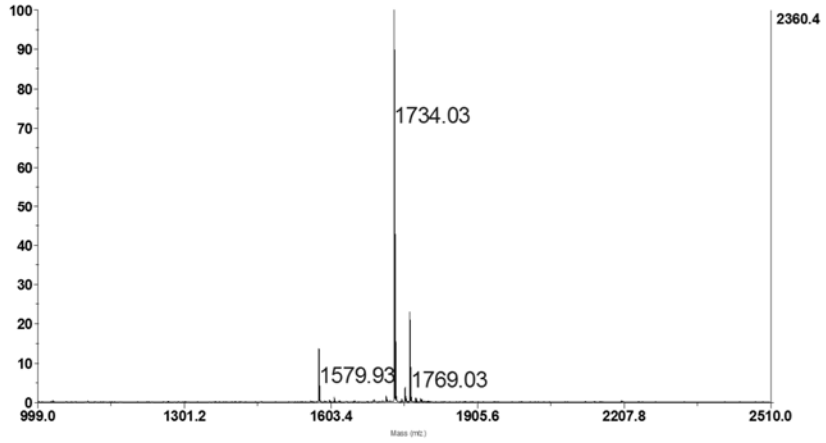
Final - Shots 1000 - LCFP3; Label L2



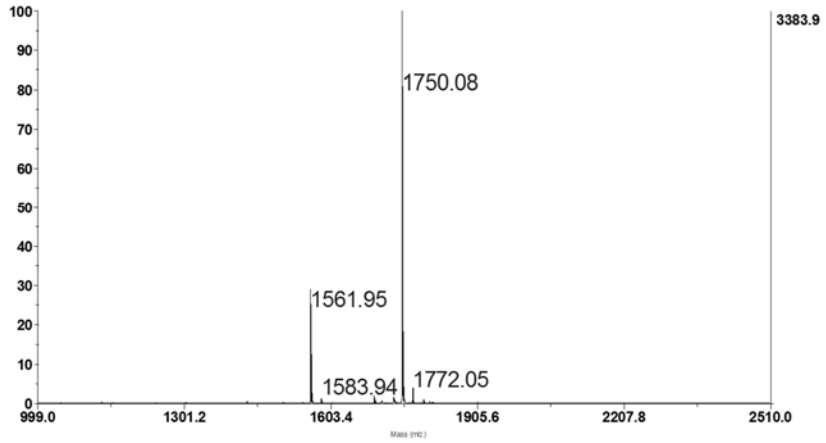
Final - Shots 1000 - LCFP3; Label F1



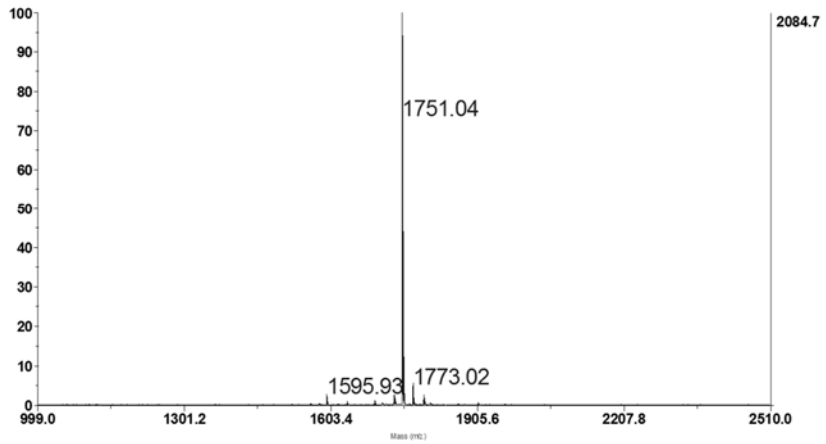
Final - Shots 1000 - LCFP3; Label L4



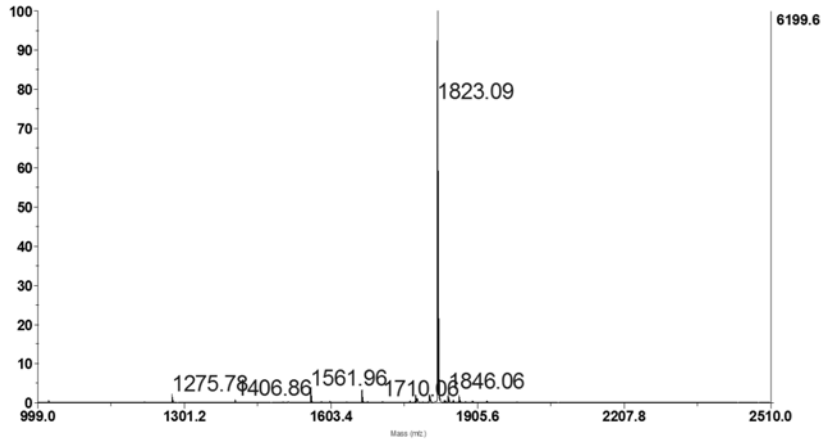
Final - Shots 1000 - LCFP3; Label L6



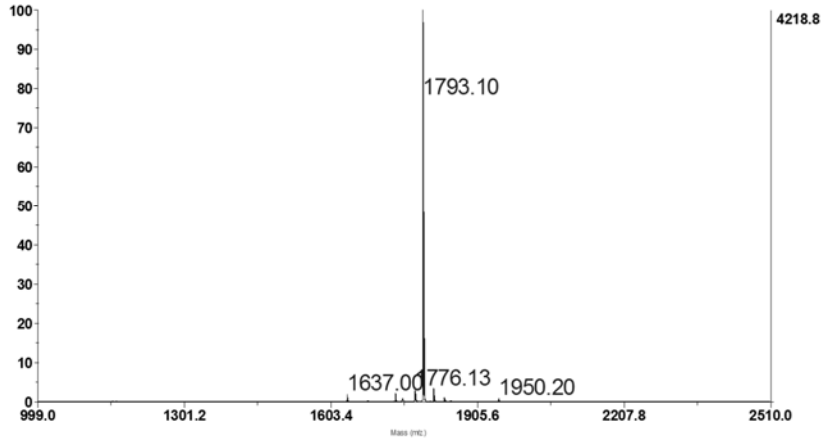
Final - Shots 1000 - LCFP3; Label L7



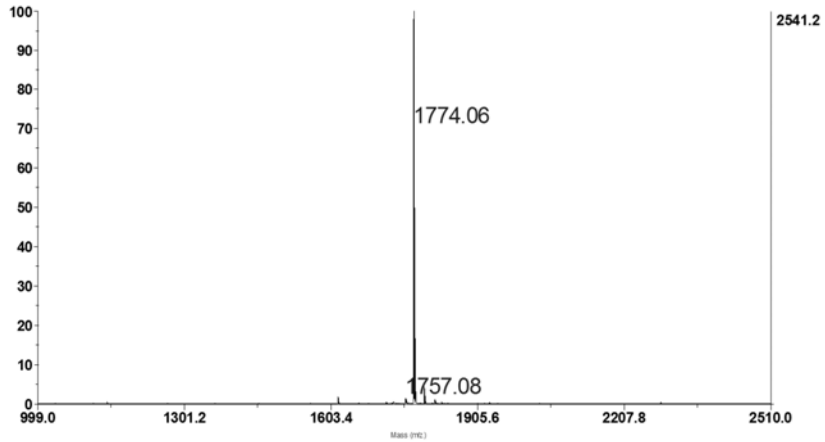
Final - Shots 1000 - LCFP3; Label L8



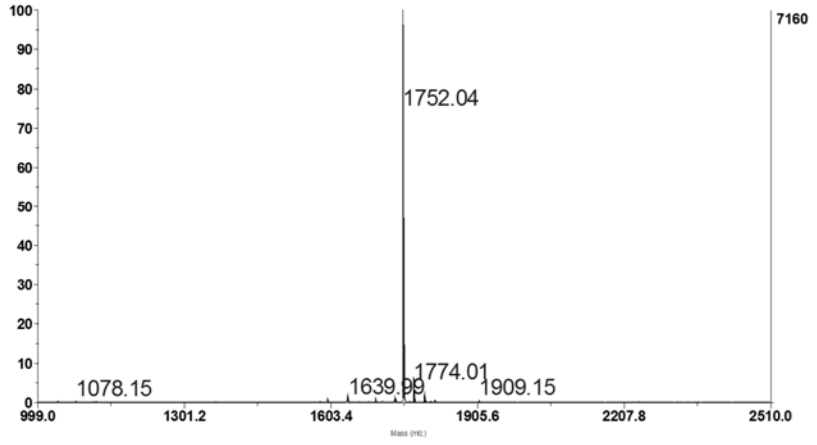
Final - Shots 1000 - LCFP3; Label L9



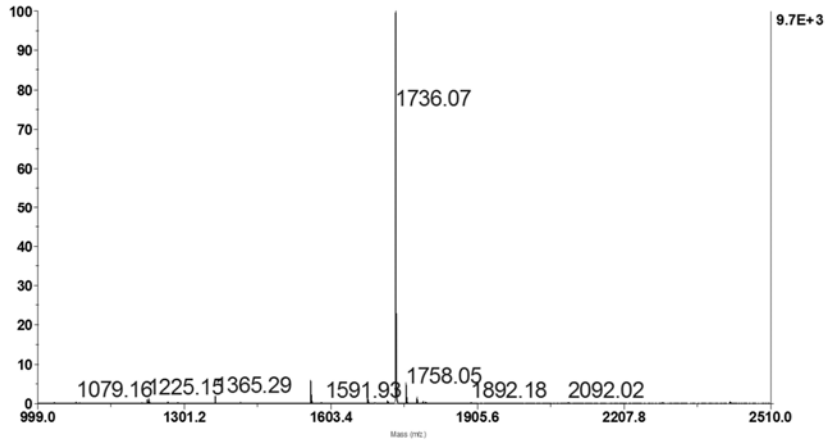
Final - Shots 1000 - LCFP3; Label L10



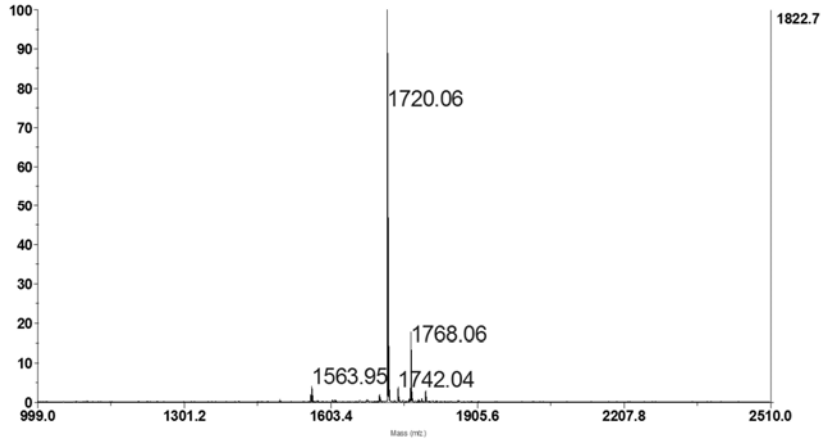
Final - Shots 1000 - LCFP3; Label L11



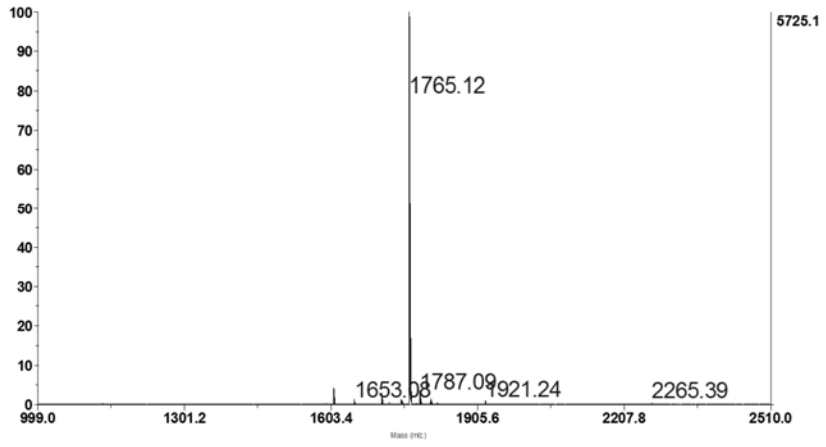
Final - Shots 1000 - LCFP3; Label L12



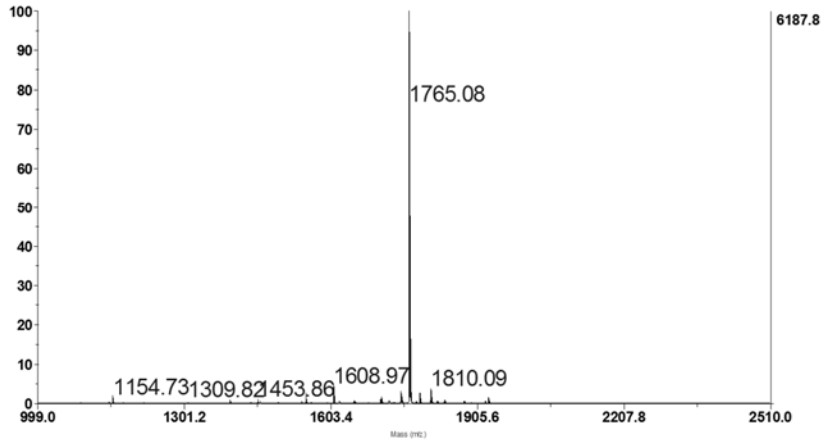
Final - Shots 1000 - LCFP3; Label L13



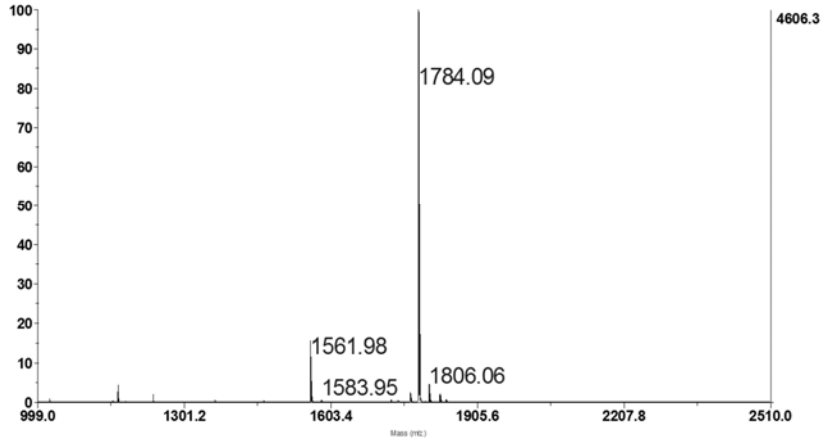
Final - Shots 1000 - LCFP3; Label L14



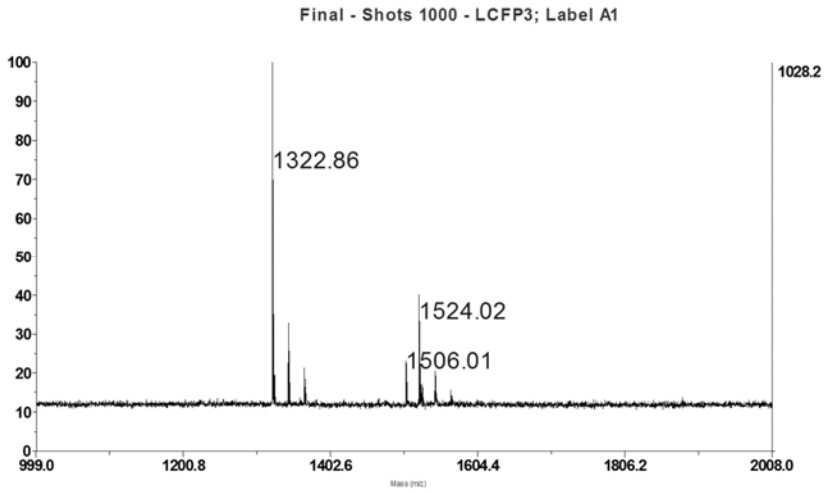
Final - Shots 1000 - LCFP3; Label L15



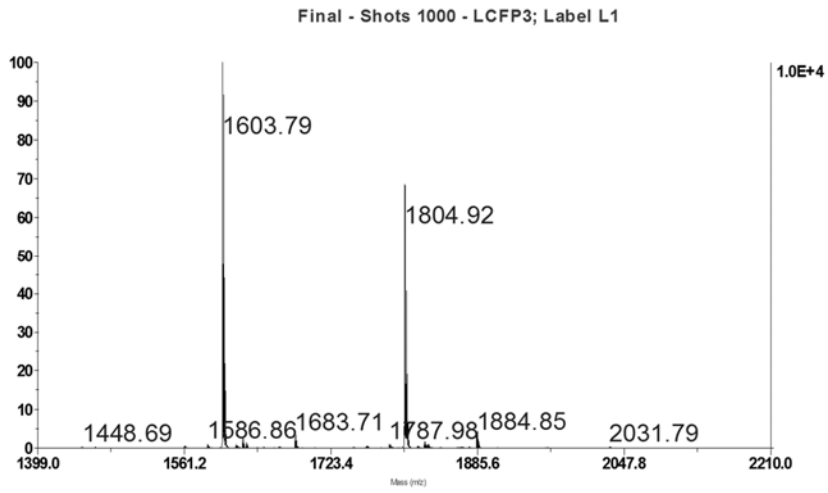
Final - Shots 1000 - LCFP3; Label L16



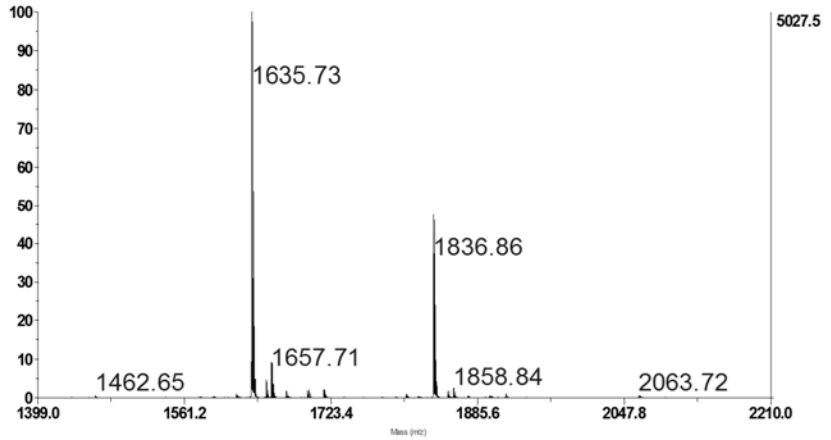
GL PVSTKPVATRDAL cyclization for 5 min, when replacing the Asn in NAL sequence with Asp in DAL sequence. The ligation is less efficient but still robust enough.



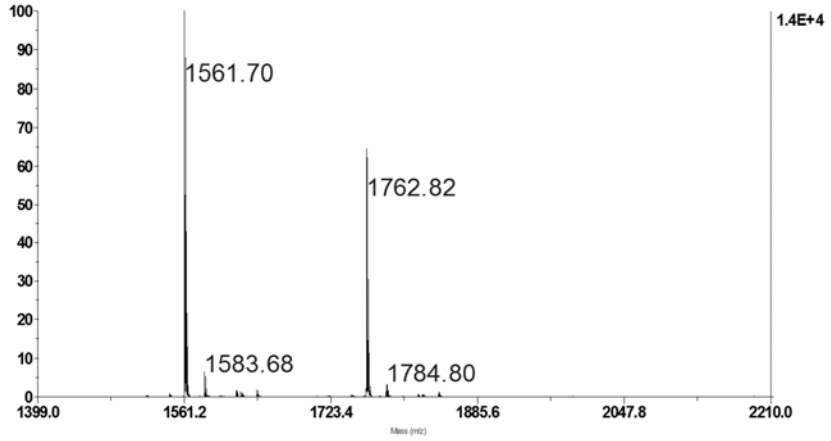
XLYRRGRLYRRNAL (L1-L20: VL, ML, GL, RL, QL, WL, EL, PL, DL, TL, SL, AL, IL, NL, KL, CL, YL, HL, FL, LL)



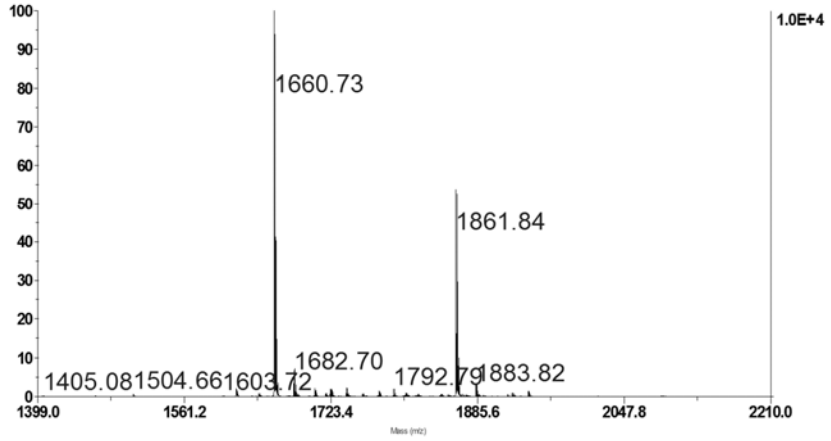
Final - Shots 1000 - LCFP3; Label L2



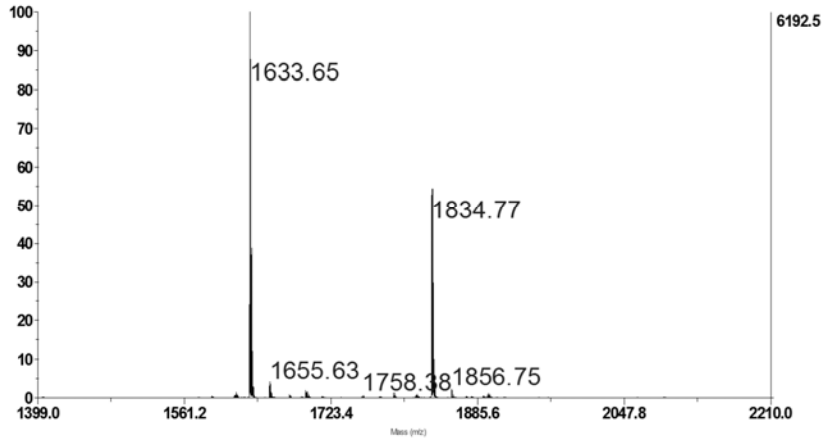
Final - Shots 1000 - LCFP3; Label L3



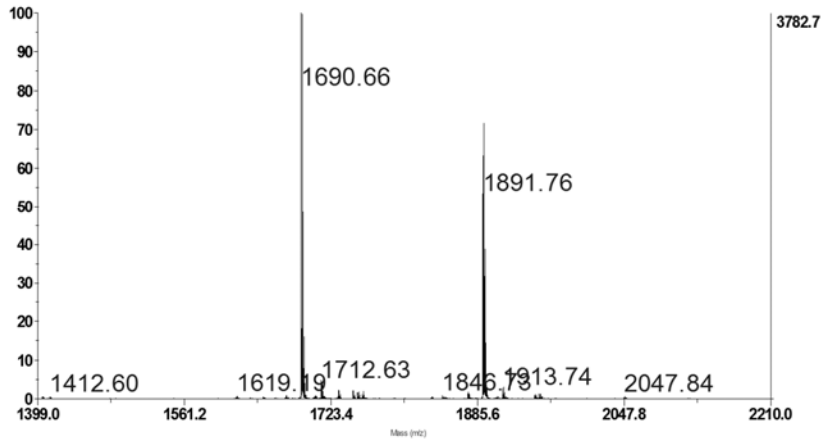
Final - Shots 1000 - LCFP3; Label L4



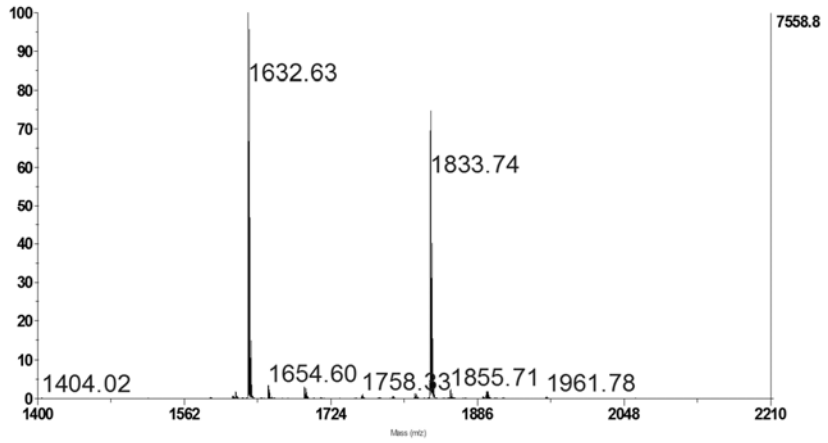
Final - Shots 1000 - LCFP3; Label L5



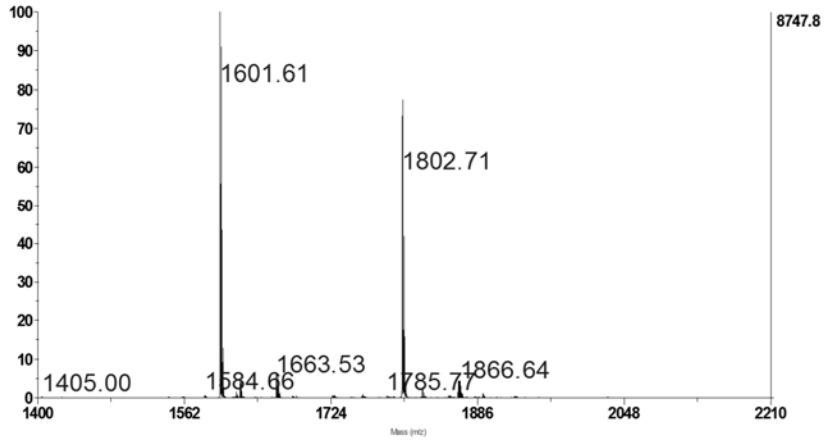
Final - Shots 1000 - LCFP3; Label L6



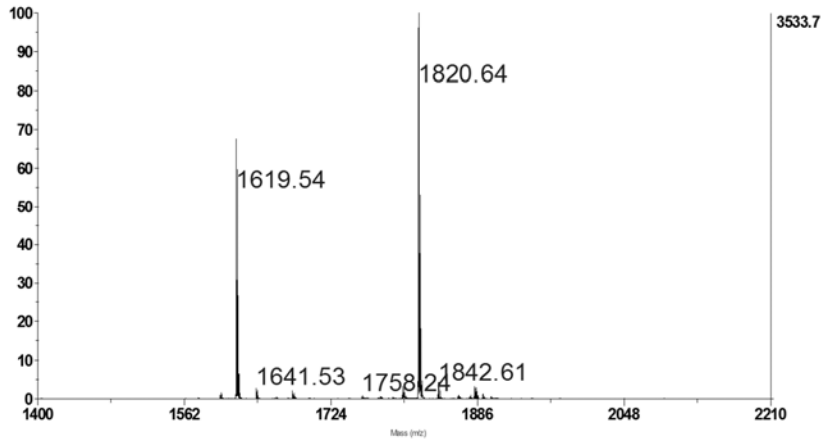
Final - Shots 1000 - LCFP3; Label L7



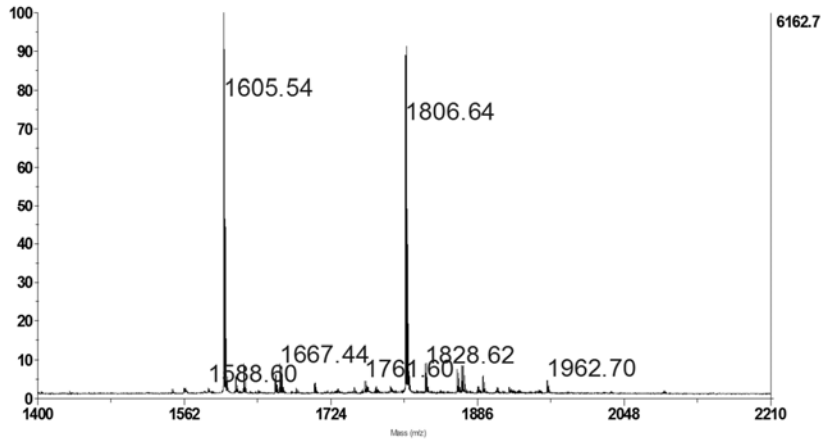
Final - Shots 1000 - LCFP3; Label L8



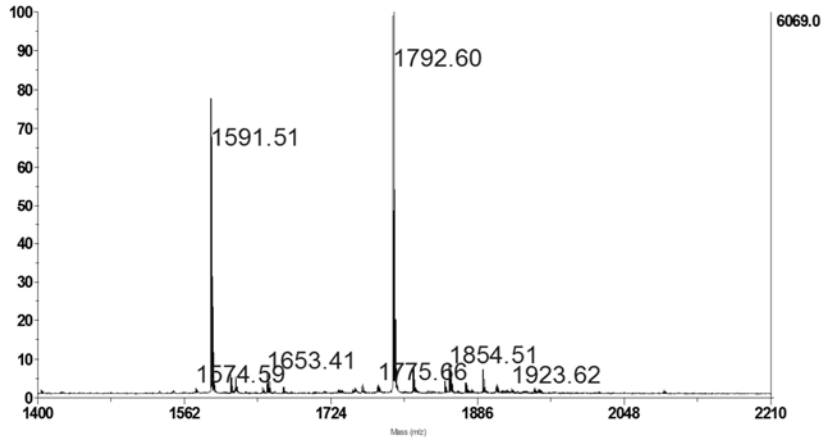
Final - Shots 1000 - LCFP3; Label L9



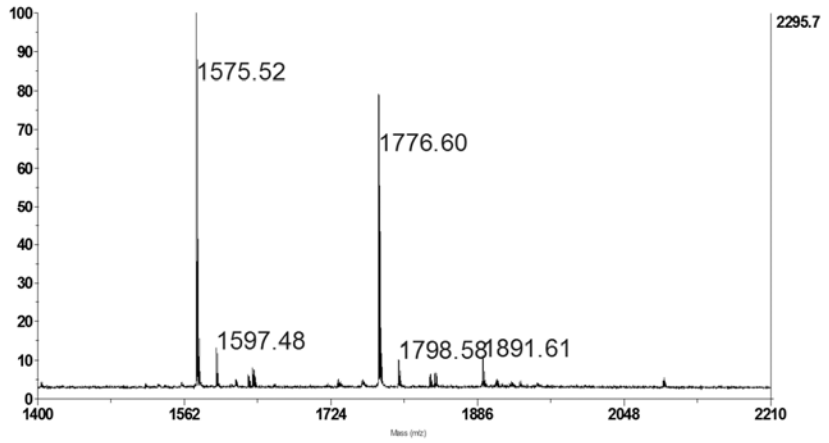
Final - Shots 1000 - LCFP3; Label L10



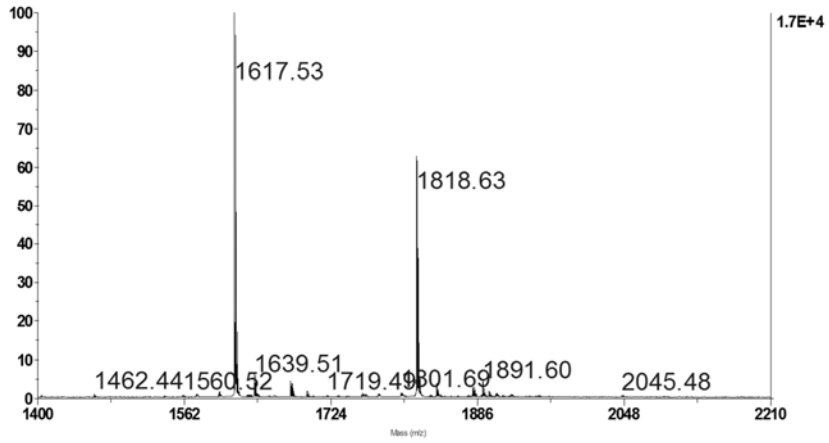
Final - Shots 1000 - LCFP3; Label L11



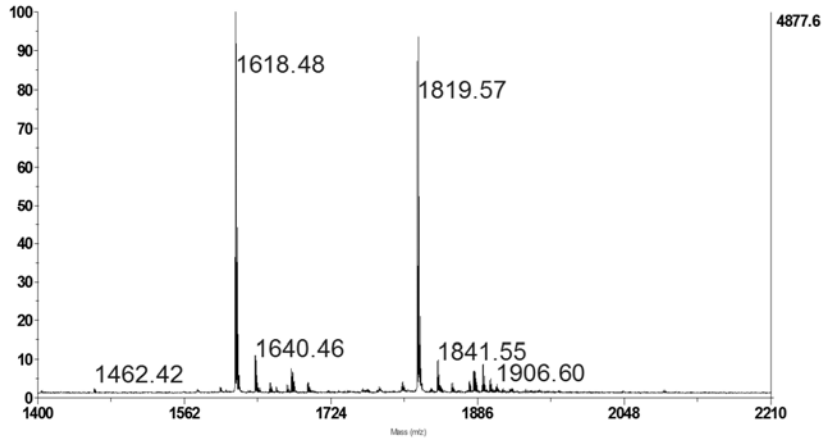
Final - Shots 1000 - LCFP3; Label L12



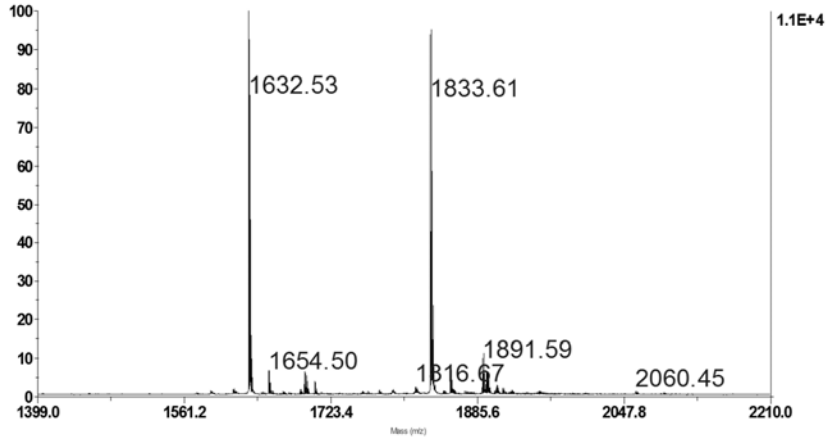
Final - Shots 1000 - LCFP3; Label L13



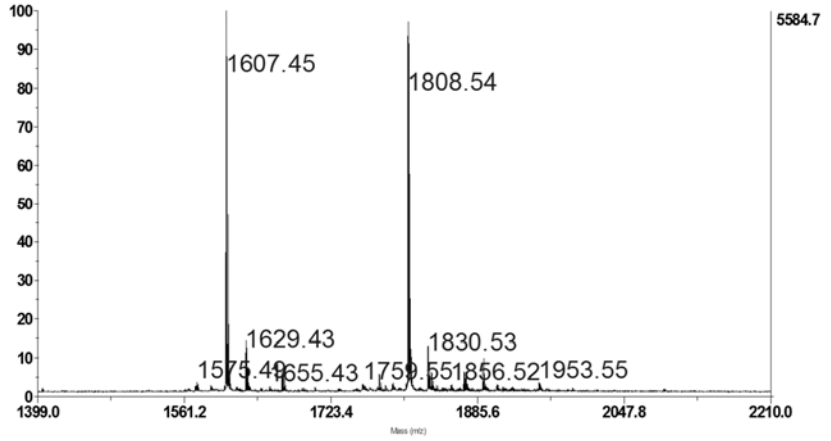
Final - Shots 1000 - LCFP3; Label L14



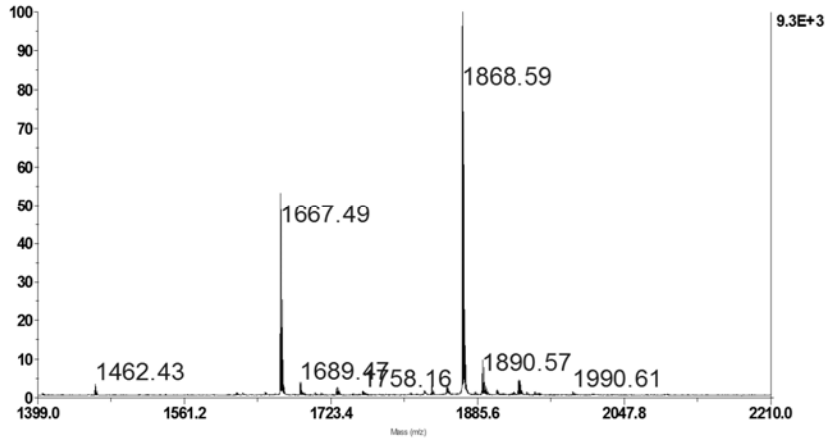
Final - Shots 1000 - LCFP3; Label L15



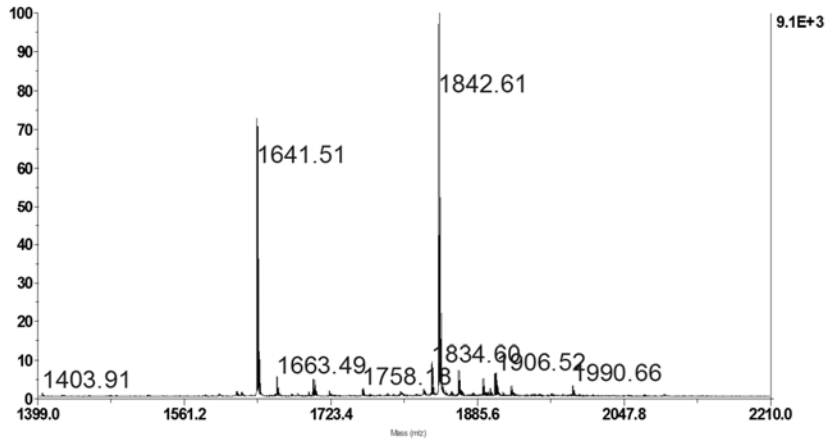
Final - Shots 1000 - LCFP3; Label L16



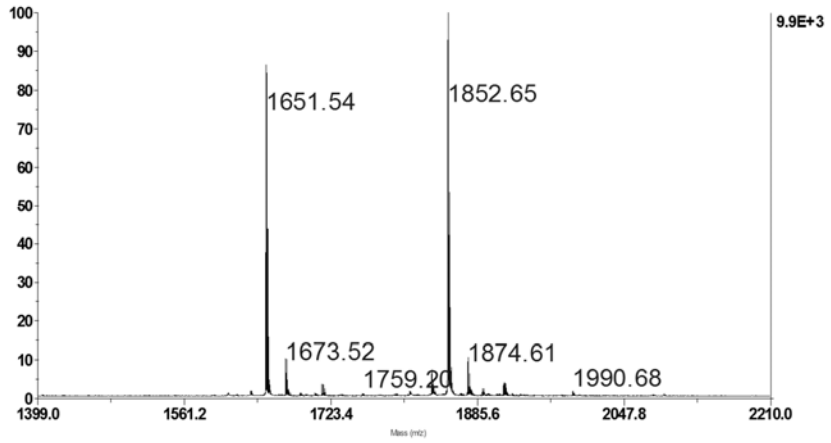
Final - Shots 1000 - LCFP3; Label L17



Final - Shots 1000 - LCFP3; Label L18



Final - Shots 1000 - LCFP3; Label L19



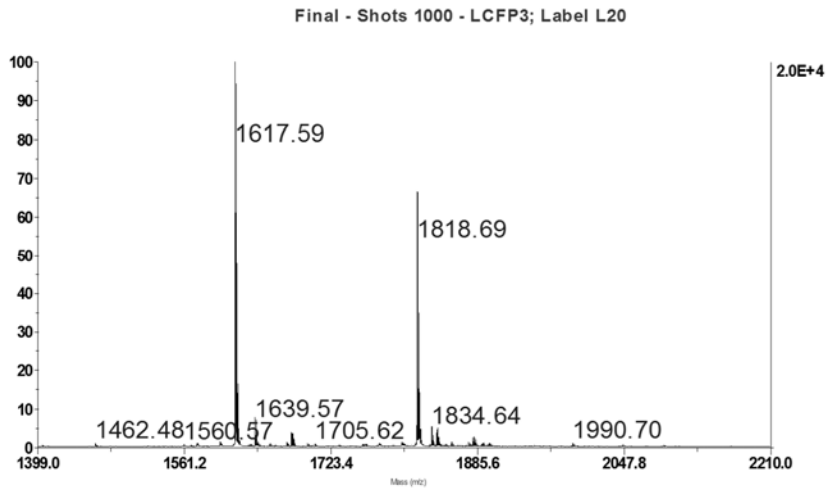

Final - Shots 1000 - LCFP3; Label L20
2.0E+4
1617.59
1818.69
1639.57
1462.48
1560.57
1705.62
1834.64
1990.70
100
90
80
70
60
50
40
30
20
10
0
1399.0
1561.2
1723.4
1885.6
2047.8
2210.0

Figure 12a

a

BLAST Results

Blast 2 sequences

Job title: Protein Sequence (474 letters)

RID HUCBXSS011N (Expires on 05-19 13:30 pm)

Query ID	lcl|Query_162141	**Subject ID**	7 subjects
Description	None	**Description**	See details
Molecule type	amino acid	**Molecule type**	amino acid
Query Length	474	**Subject Length**	3432
		Program	BLASTP 2.6.1+

New Analyze your query with SmartBLAST

Graphic Summary

Distribution of the top 7 Blast Hits on 7 subject sequences

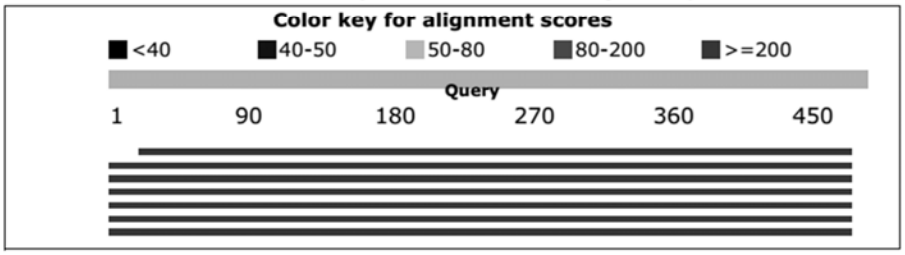


Figure 12b

b

Sequences producing significant alignments:

Description	Max score	Total score	Query cover	E value	Ident	Accession
unnamed protein product [Coffea canephora]	730	730	95%	0.0	76%	CDP08231.1
PREDICTED: vacuolar-processing enzyme-like [Daucus carota subsp. sativus]	708	708	100%	0.0	70%	XP_017229093.1
Peptidase C13, legumain [Corchorus capsularis]	707	707	100%	0.0	70%	OMO66906.1
Peptidase C13, legumain [Corchorus olitorius]	702	702	100%	0.0	69%	OMO86616.1
vacuolar-processing enzyme [Jatropha curcas]	701	701	100%	0.0	68%	XP_012077326.1
vacuolar processing enzyme [Malus hupehensis]	699	699	100%	0.0	69%	AGC94758.1
PREDICTED: vacuolar-processing enzyme-like [Pyrus x bretschneideri]	698	698	100%	0.0	69%	XP_009361606.1

Figure 13

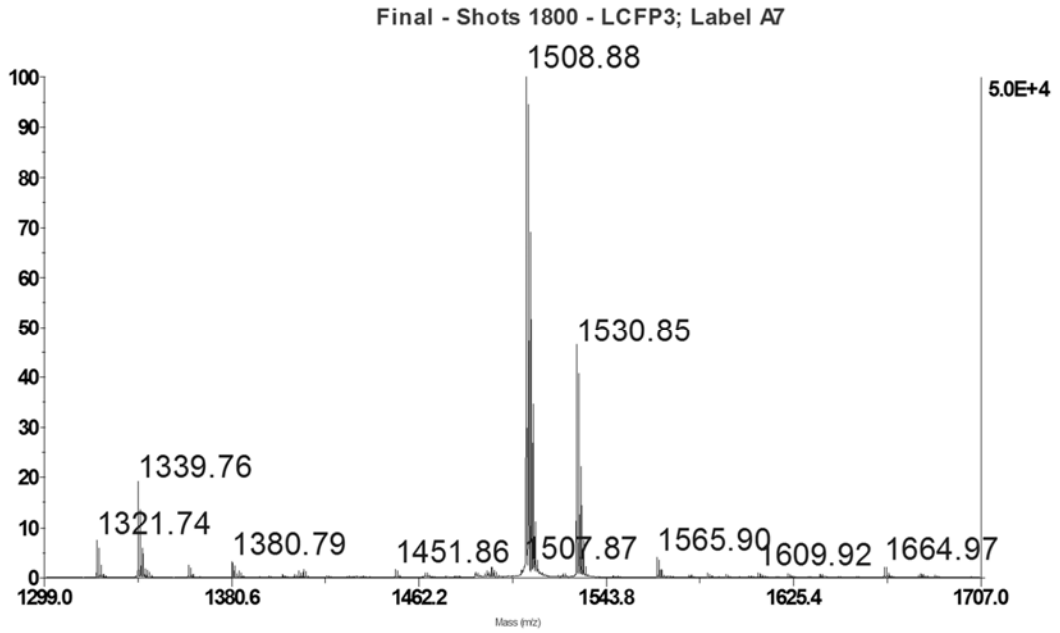
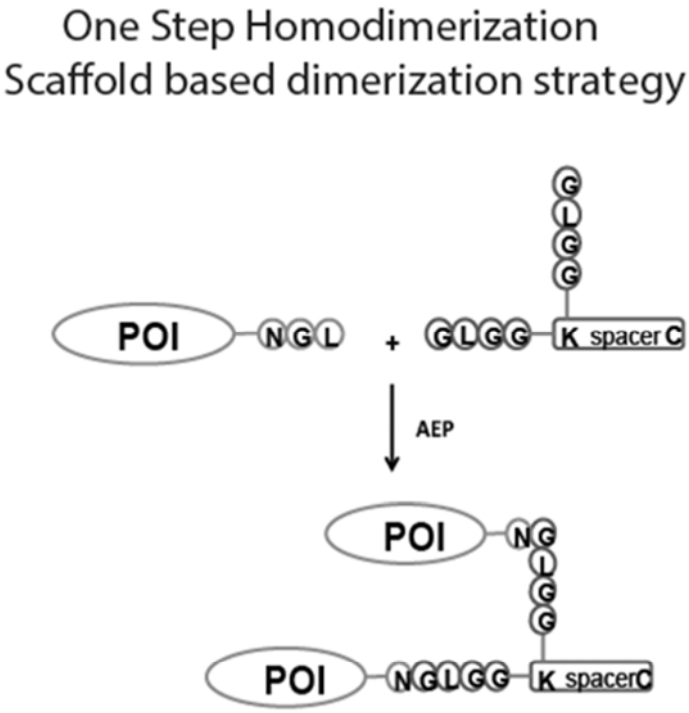
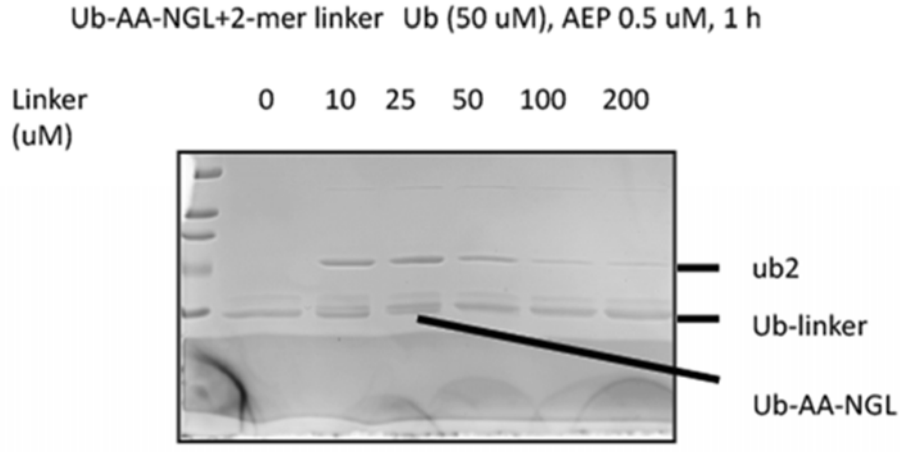

Final - Shots 1800 - LCFP3; Label A7
1508.88
5.0E+4
1530.85
1339.76
1321.74
1380.79
1451.86
1507.87
1565.90
1609.92
1664.97
100
90
80
70
60
50
40
30
20
10
0
1299.0
1380.6
1462.2
1543.8
1625.4
1707.0
Mass (m/z)

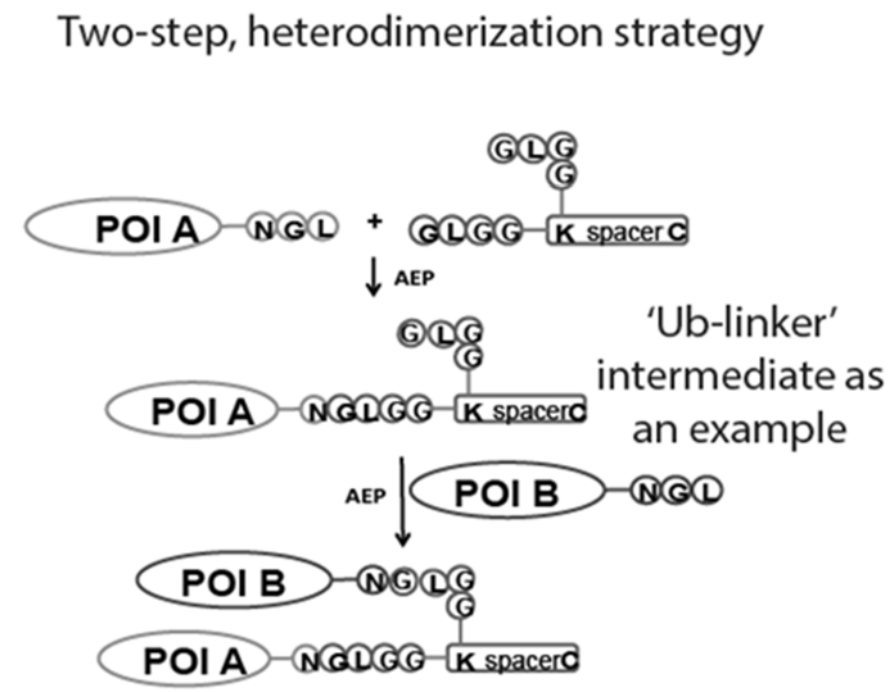
Figure 14

a

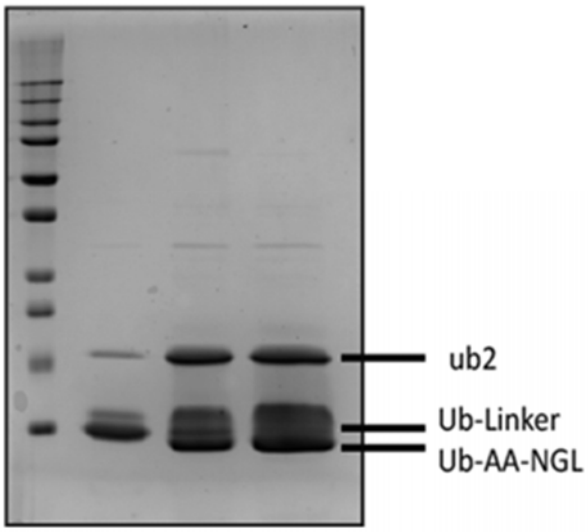

One Step Homodimerization
Scaffold based dimerization strategy
POI
N G L
+
G L G G
K spacer C
AEP
POI
N G
L
G
G
POI
N G L G G
K spacer C


Ub-AA-NGL+2-mer linker Ub (50 uM), AEP 0.5 uM, 1 h
Linker (uM)
0 10 25 50 100 200
ub2
Ub-linker
Ub-AA-NGL

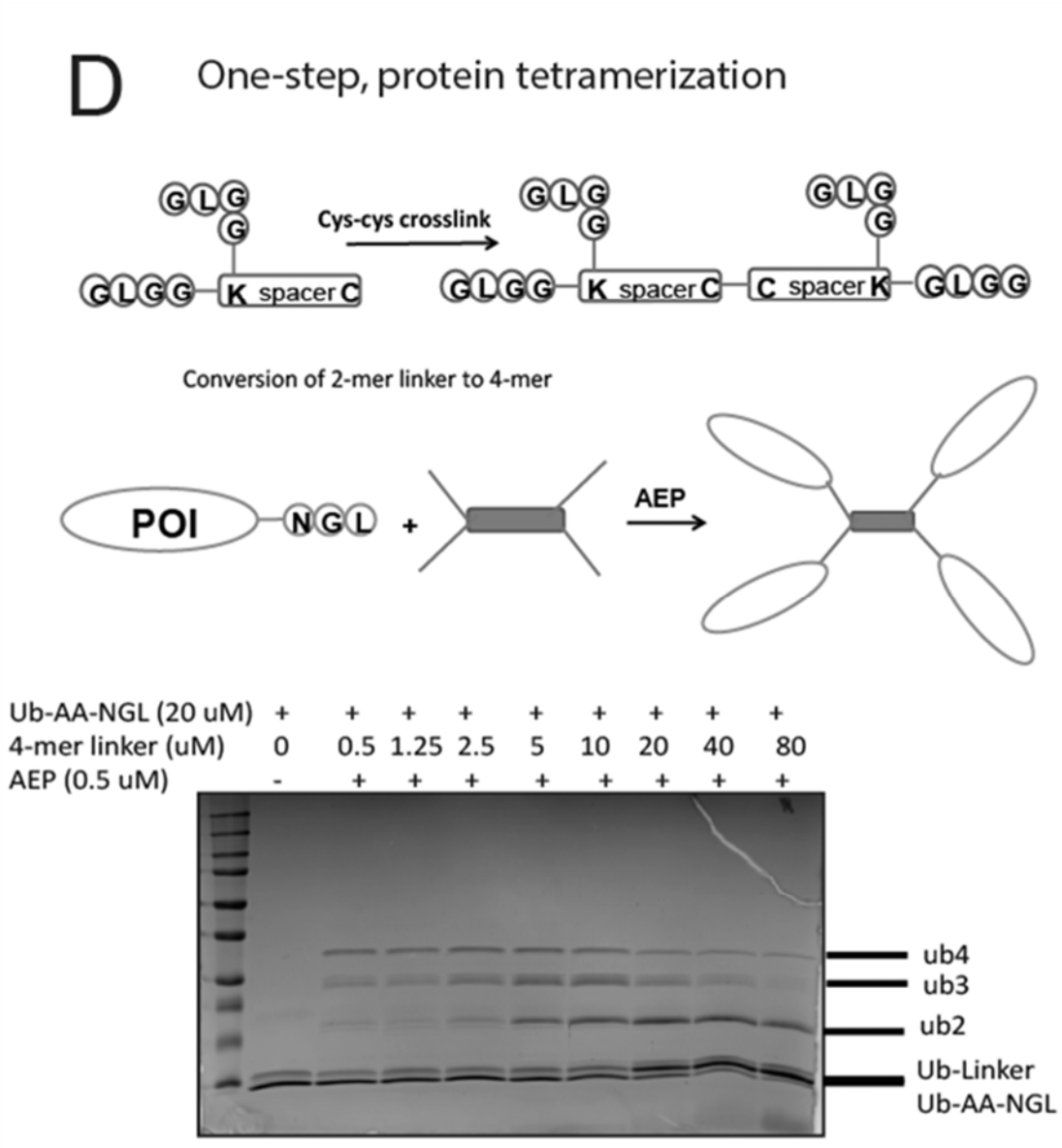
b



Ub-linker (uM)	50	50	50
Ub-AA-NGL (uM)	0	100	200
AEP (uM)	0	0.5	0.5

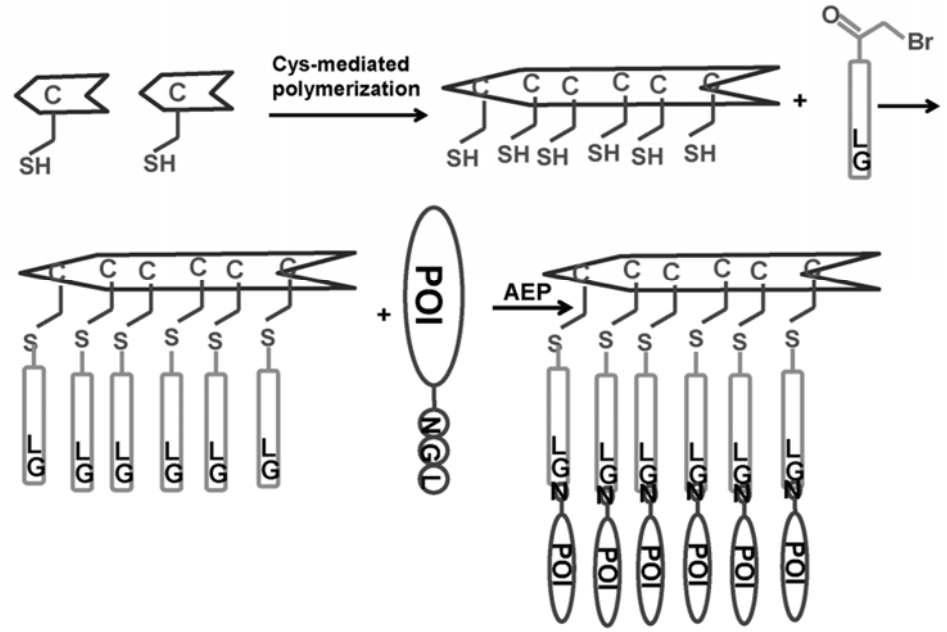


C

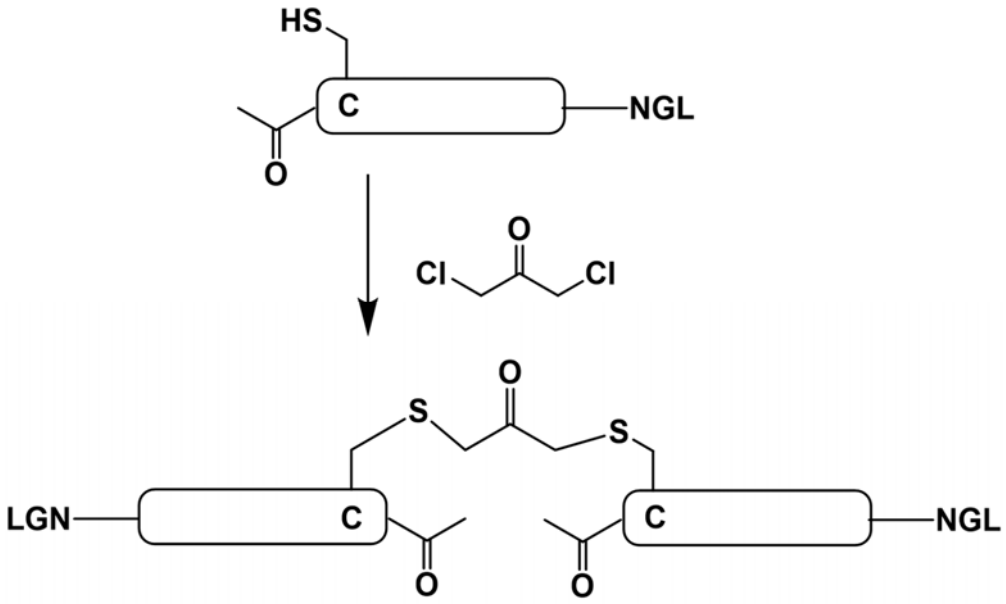

D
One-step, protein tetramerization
GLGG
K spacer C
Cys-cys crosslink
C spacer K
Conversion of 2-mer linker to 4-mer
POI
NGL
AEP
Ub-AA-NGL (20 uM) + + + + + + + + +
4-mer linker (uM) 0 0.5 1.25 2.5 5 10 20 40 80
AEP (0.5 uM) - + + + + + + + +
ub4
ub3
ub2
Ub-Linker
Ub-AA-NGL

d

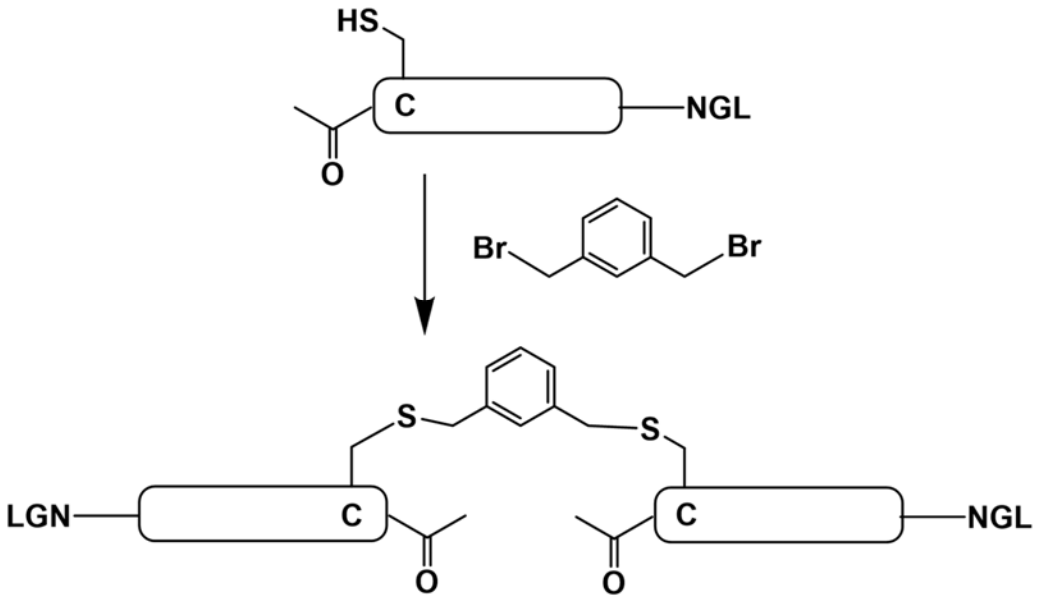
Schematic Ilustration of protein ultra-oligomer strategy



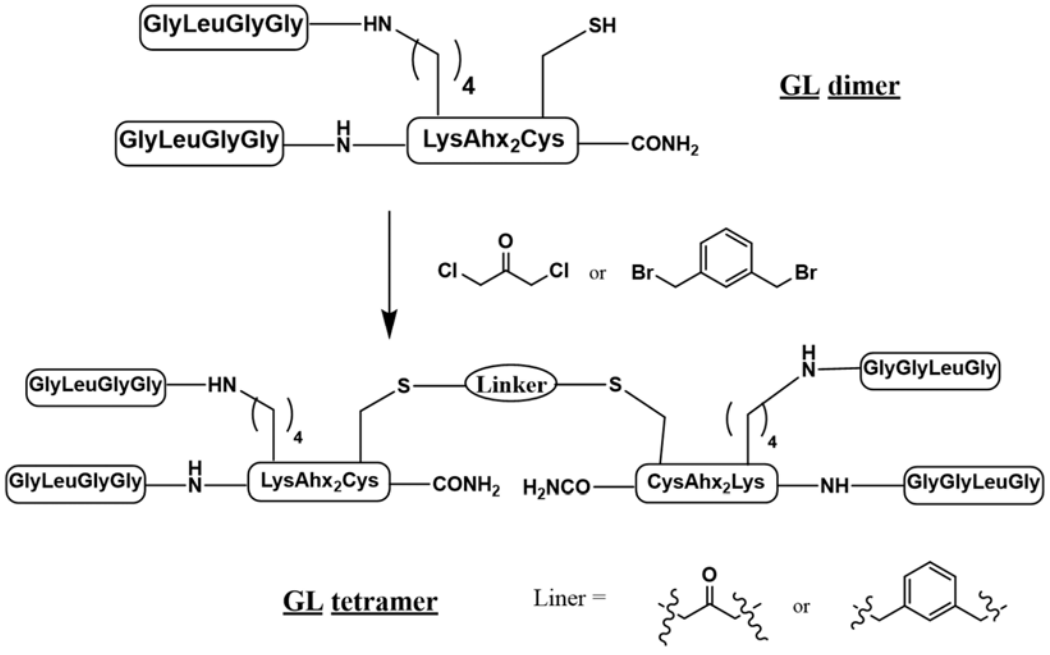
e



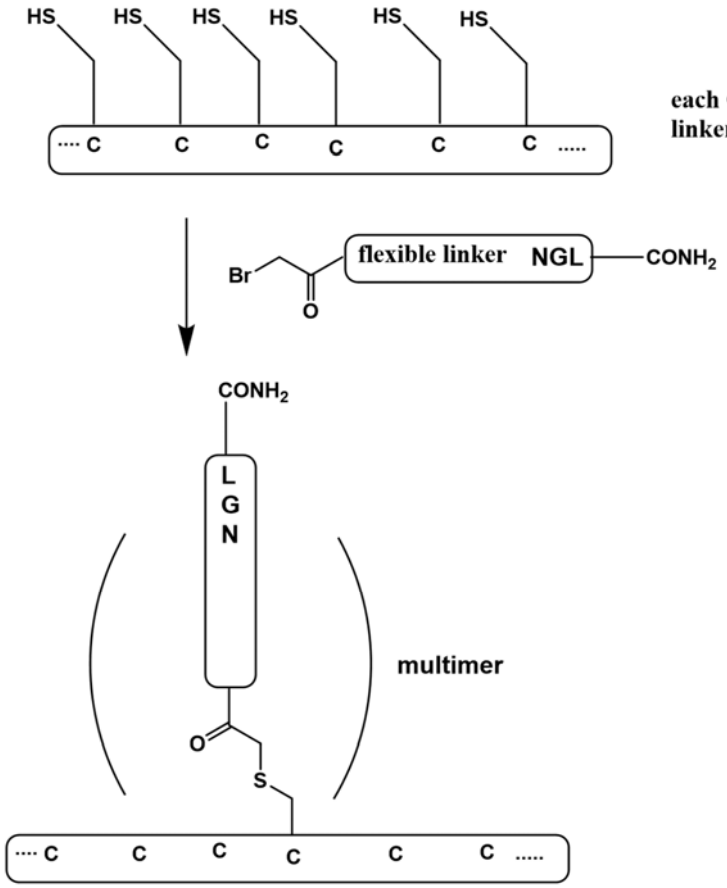
f


HS
C
NGL
O
Br
Br
S
S
LGN
C
C
NGL
O
O

g


GlyLeuGlyGly
HN
4
SH
GlyLeuGlyGly
H N
LysAhx2Cys
CONH2
GL dimer
O
Cl
Cl
or
Br
Br
GlyLeuGlyGly
HN
4
S
Linker
S
H N
GlyGlyLeuGly
4
GlyLeuGlyGly
H N
LysAhx2Cys
CONH2
H2NCO
CysAhx2Lys
NH
GlyGlyLeuGly
GL tetramer
Liner =
O
or

h


HS
HS
HS
HS
HS
HS
C
C
C
C
C
C
each Cys is separated by flexible linker made of 7-8 amio acids
Br
O
flexible linker
NGL
CONH2
CONH2
L
G
N
multimer
O
S
C
C
C
C
C
C

i

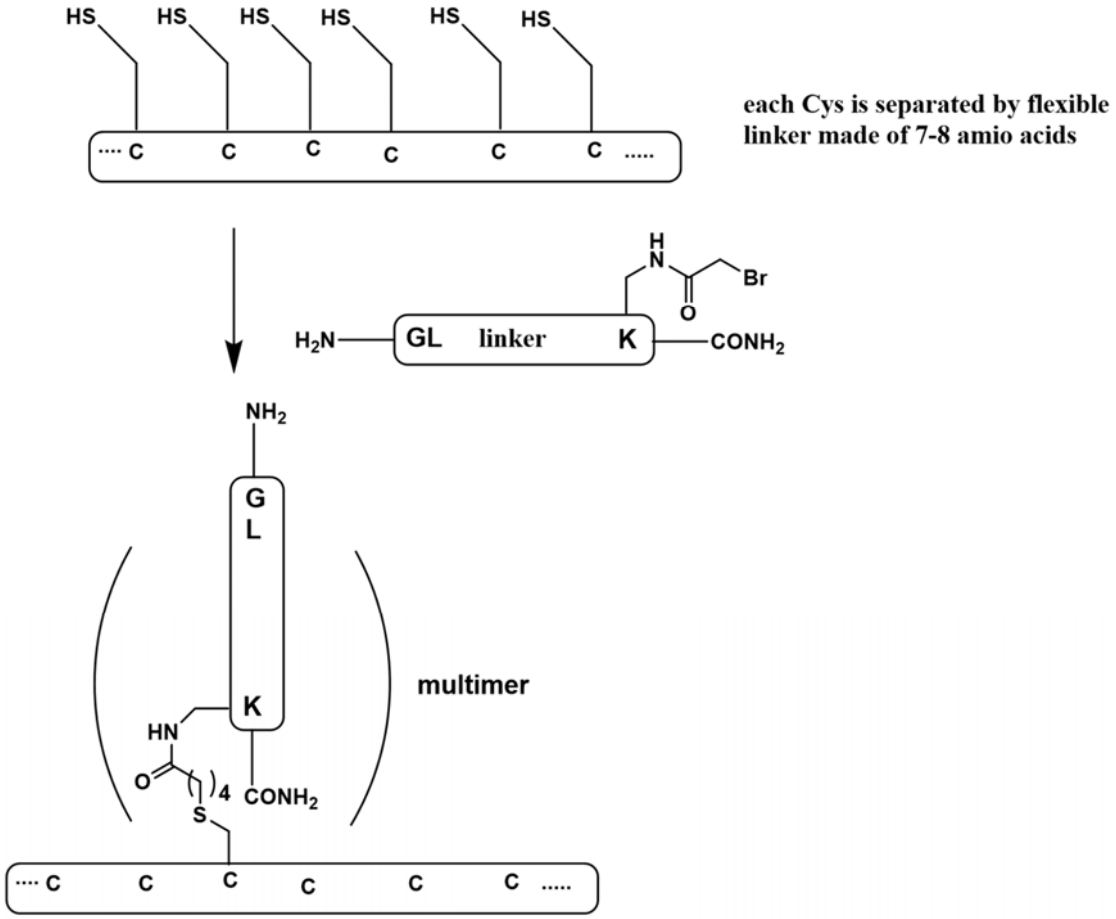
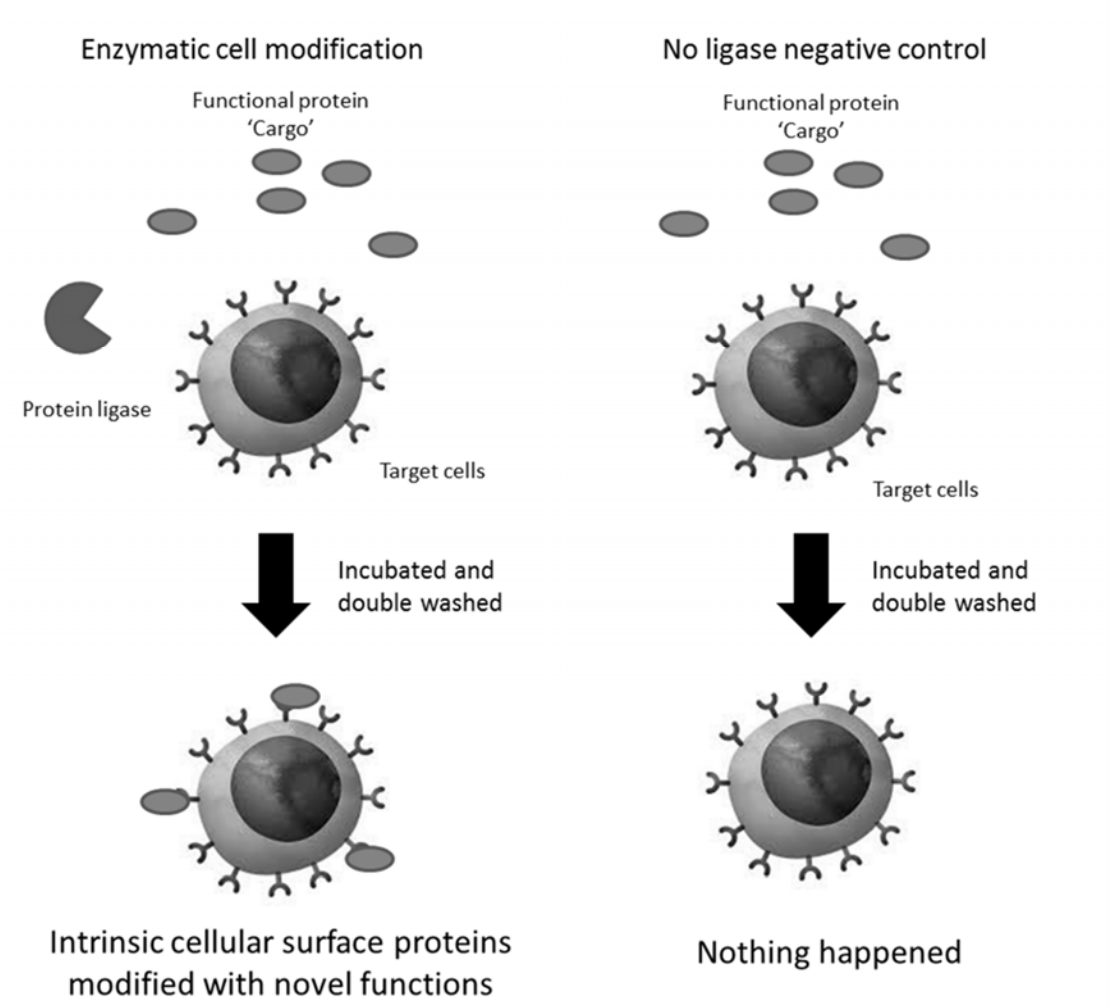
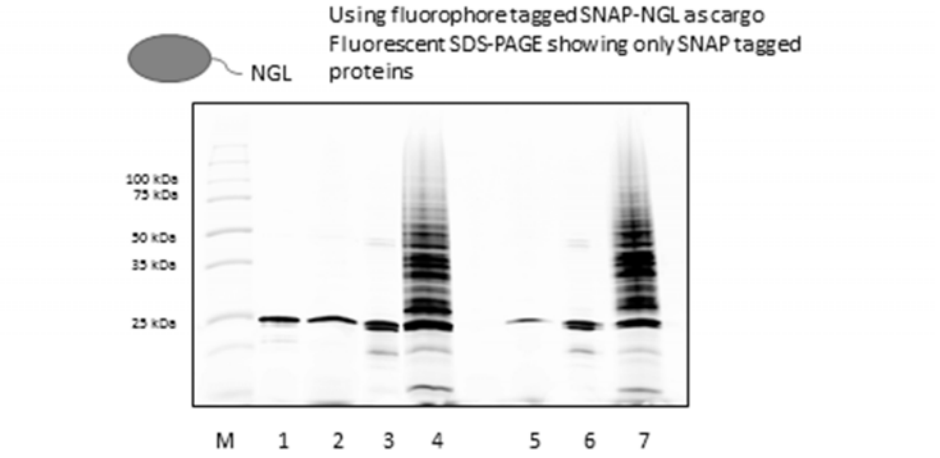

HS
HS
HS
HS
HS
HS
C
C
C
C
C
C
each Cys is separated by flexible linker made of 7-8 amio acids
H
N
Br
O
H2N
GL
linker
K
CONH2
NH2
G
L
K
HN
O
4
CONH2
S
multimer
C
C
C
C
C
C

Figure 15

a


Enzymatic cell modification
Functional protein
'Cargo'
Protein ligase
Target cells
Incubated and
double washed
Intrinsic cellular surface proteins
modified with novel functions
No ligase negative control
Functional protein
'Cargo'
Target cells
Incubated and
double washed
Nothing happened

b



M = Molecular weight
1 = SNAP-NGL protein control
2 = SNAP-NGL added to RAW cells, no ligase, cells after wash
3 = SNAP-NGL added to RAW cells, with ligase, wash solution
4 = SNAP-NGL added to RAW cells, with ligase, cells after wash
5 = same as 2, using HEK cells
6 = same as 3, using HEK cells
7 = same as 4, using HEK cells

c

