Electronic Supplementary Material (ESI) for Chemical Communications.

Fast and Catalyst-Free Hydrazone Ligation via Ortho-halo-substituted Benzaldehydes for Protein C-terminal Labeling at Neutral pH

Yang Xu,[a,b‡] Lin Xu,[a‡] Yuan Xia,[c] Chao-Jian Guan,[b] Yao Fu,[a] Qing-Xiang Guo,[a] Chen Wang,[c,*] Yi-Ming Li,[a,b*]

[a] *Department of Chemistry, University of Science and Technology of China, Hefei 230026, China. E-mail: lym2007@mail.ustc.edu.cn,*

[b] *School of Medical Engineering, Hefei University of Technology, Hefei, Anhui 230009, China, E-mail: ymli@hfut.edu.cn*

[c] *Zhejiang Key Laboratory of Alternative Technologies for Fine Chemicals Process, Shaoxing University, Shaoxing 312000, China, Email: wangchen@usx.edu.cn*

‡These authors contributed equally to this work

1. General Information

a. Materials

All chemical reagents and solvents were purchased from Sinopharm Chemical Reagent Co. Ltd., Alfa Aesar China Co. Ltd., CS Bio Co. (shanghai), GL Biochem (shanghai), Aladdin-reagent Co. (shanghai), J&K Chemical Co. Ltd. and were purified when necessary. TLC was executed on plates pre-coated with silica gel 60 F254 (250 layer thickness). Visualization was achieved using UV light, iodine vapors, permanganate solution. Flash column chromatographic purification of products was achieved using forced-flow chromatography on Silica Gel (200-300 mesh on small-scale or 300-400 mesh on large-scale). Manual peptide-synthesis apparatus was using the peptide synthesis vessel and in a constant-temperature shaker at 30℃. Automated peptide-synthesis apparatus was using a CS Bio 136XT automated synthesizer conducting with a 0.25 mmol resin scale.

b. HPLC

Analytical HPLC was run on a SHIMADZU (Prominence LC-20AT) instrument using analytical column (Grace Vydac "Protein & Peptide C18", 250 × 4.6 mm, 5 μm particle size, flow rate1.5 mL/min, R.T.). Analytical samples were monitored at 214 nm and 254 nm. Semi-preparative HPLC was run on a SHIMADZU (Prominence LC-20AT) instrument using a semi preparative column (Grace Vydac "Peptide C18", 250 × 10 mm, 10 μm particle size, flow rate 5 mL/min, rt). Solution A was 0.08 % trifluoroacetic acid in acetonitrile, and solution B was 0.1 % trifluoroacetic acid in ddH_2O.

c. Mass spectrometry and NMR

ESI-MS (/MS) spectra were recorded on a Finnigan LCQ Advantage MAX ion trap mass spectrometer (Thermo Fisher Scientific. USA) equipped with a standard ESI ion source. Data acquisition and analysis were done with the Xcalibur (version 2.0, Thermo quest Finnigan) software package.

^{1}H NMR spectra were recorded on a Bruker 400 MHz spectrometer in deuteriochloroform ($CDCl_3$) with the solvent residual peak ($CDCl_3$: 7.26 ppm (^{1}H)as internal reference unless otherwise stated. ^{13}C-NMR spectra were recorded with^{1}H-decoupling on a Brucker 101 MHz spectrometer. Data are reported in the following order: chemical shifts are given (δ); multiplicities are indicated as br (broadened), s(singlet), d (doublet), t (triplet), q (quartet), m (multiplet), app (apparent); coupling constants, J, are reported (Hz); integration is provided.

d. General procedures for SPPS of peptides following the Fmoc strategy

The peptide synthesis reaction vessels were attained from commercial sources. Hydrazine 2CTC resin was initially swelled with DCM/DMF (1/1, V/V) about 0.5 h. For pre-activation of the first protected amino acid, 3.6 eq. of HCTU, 8 eq. of DIEA were added to a solution of 4 eq. protected amino acid in DMF. After pre-activation for 1 min, the mixture was added to the resin. After 1 h the resin was washed with DMF (3 times), DCM (3 times) and DMF (3 times). Then treatment with 20% piperidine/DMF (2 min, 10min), the resin was washed again with DMF (3

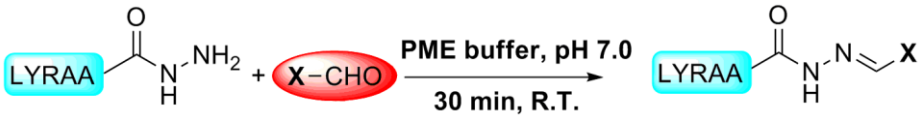
times), DCM (3 times) and DMF (3 times). Amino acid residues were coupled by a pre-activated solution of 4 eq. protected amino acid using 3.6 eq. HBTU, 8 eq. DIEA to the resin. After 1 h, the resin was washed with DMF (3 times), DCM (3 times) and DMF (3 times). Removal of Fmoc roup: piperidine (20% in DMF) was added to the resin for 12 min (twice: 2 min and 10 min). Couplings were checked by ninhydrin test. Cleavage: A mixture of 88%TFA, 5% water, 5% phenol and 2% TIPS was added. After 2.5 h, the resin was washed with TFA. The combined solutions were concentrated by blowing with N_2. The crude peptides were acquired by precipitation with cold ether and centrifugation. The residue was dissolved in water, purified by preparative HPLC and analyzed by High-resolution ESI mass spectra.

2. Experimental Section

a. Synthesis of hydrazine 2CTC resin

2-Chlorotrityl chloride resin (1 g, loading = 0.55 mmol/g) was swelled in 10 mL DCM/DMF (1/1, V/V) at 30℃. Then 10 mL $NH_2NH_2{\cdot}H_2O$/DMF (1/20, V/V) were added. The reaction was conducted overnight. 10 mL Methanol/DMF (1/20, V/V) was added to quench the remaining 2-Chlorotrityl chloride resin. After 30 min, the resin was washed with DMF, H_2O, Methanol, Ethylether and kept under high vacuum for 2 h.

b. Hydrazone ligation by using H-LYRAA-$NHNH_2$ and X-CHO



The peptide hydrazide (30 μM in final concentration) were added together in the aqueous PME buffer (100 mM PIPES, 1 mM $MgSO_4$, 2 mM EGTA at pH 7.0). The compounds of aryl aldehyde was firstly dissolved in ethanol (10 mM) for stock solution and then added to reaction buffer in 200 μM final concentration. The concentration of ethanol in the buffer is about 2% (V/V). Then, the mixture was stirred at neutral pH and 25℃ for about 0.5 h before the yield was determined by RP-HPLC.

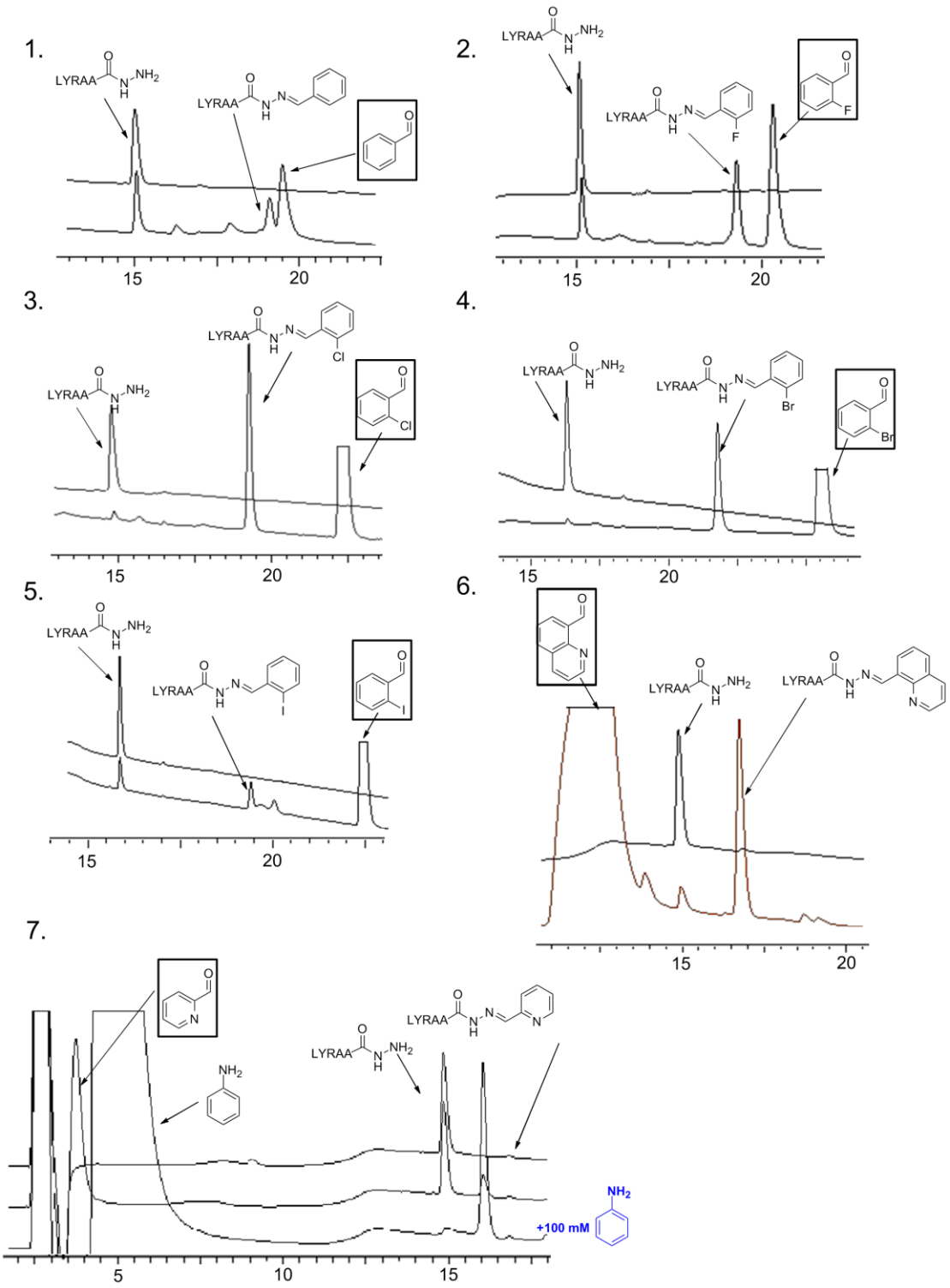
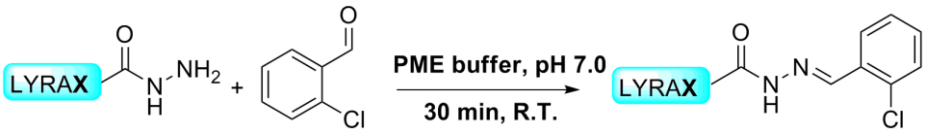


Figure S1: The chromatogram of the reactions between model peptide hydrazides Leu-Tyr-Arg-Ala-Ala-$NHNH_2$ and different aldehydes. a) 1,2,3,4,5,6,7 were monitored by RP-HPLC at 30 min. b) 7 was added 100 mM aniline to promote the reaction after 30 min.

c. Hydrazone ligation by using H-LYRAX-$NHNH_2$ and 2-Chlorobenzaldehyde



The peptide hydrazide (30 μM in final concentration) were added together in the aqueous PME buffer (100 mM PIPES, 1 mM $MgSO_4$, 2 mM EGTA at pH 7.0). 2-Chlorobenzaldehyde was firstly dissolved in ethanol (10 mM) for stock solution and then added to reaction buffer in 200 μM final concentration. The concentration of ethanol in the buffer is about 2% (V/V). Then, the mixture was stirred at neutral pH and 25℃ for about 0.5 h before the yield was determined by RP-HPLC. We found that the reaction system is clear during the experimental course (Fig. S2).

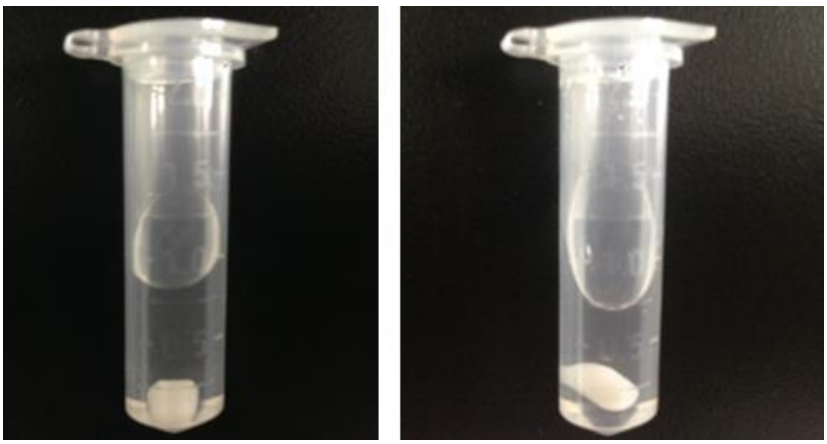
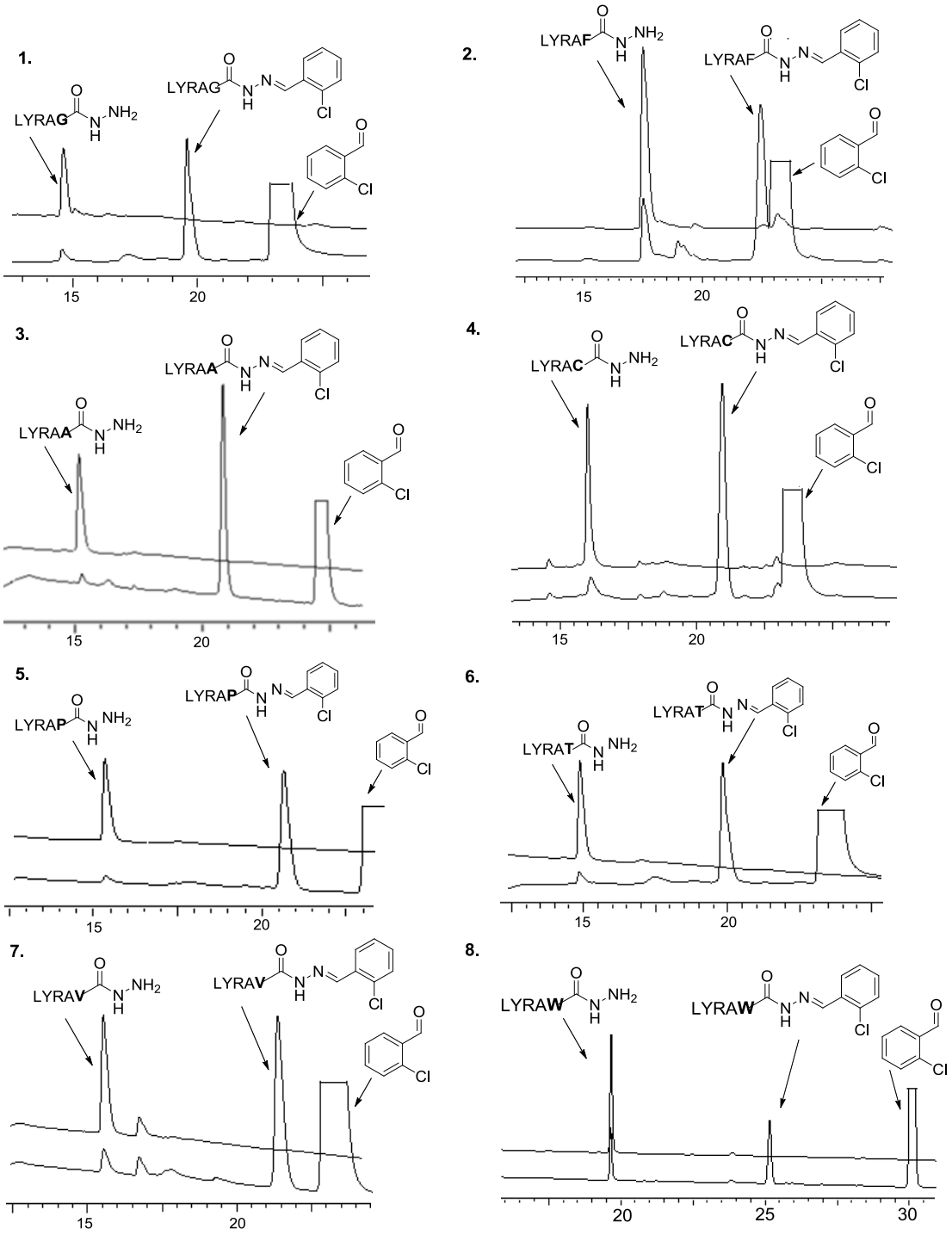


Figure S2: The photograph of the reaction between H-LYRAX-$NHNH_2$ and 2-Chlorobenzaldehyde. It showed that the reaction system is clear.


1.
LYRAG
LYRAG
2.
LYRAF
LYRAF
3.
LYRAA
LYRAA
4.
LYRAC
LYRAC
5.
LYRAP
LYRAP
6.
LYRAT
LYRAT
7.
LYRAV
LYRAV
8.
LYRAW
LYRAW
Cl
O
N
NH2
H
15
20
25
30

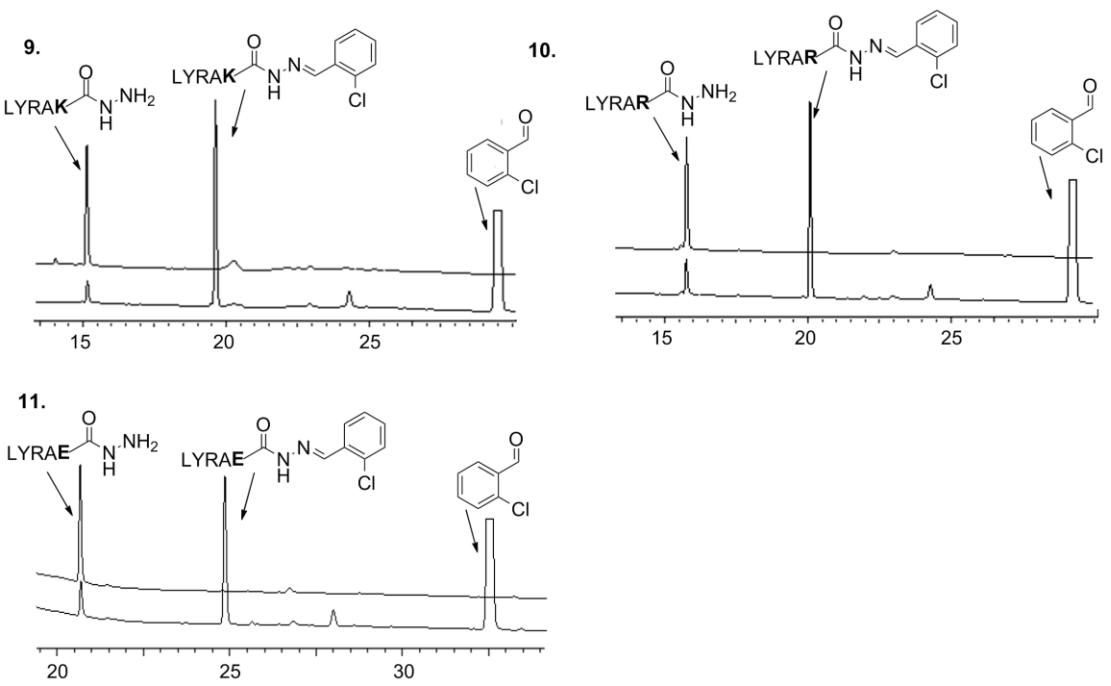
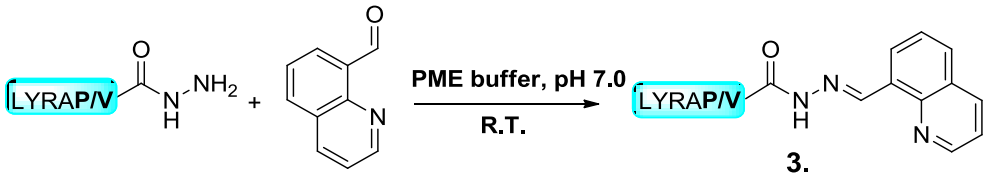
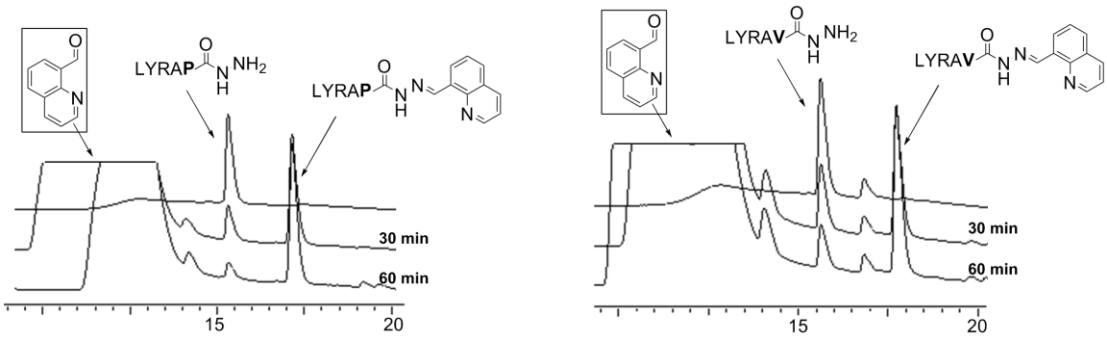


Figure S3: The chromatogram of the reactions between different peptide hydrazide and 2-Chlorobenzaldehyde. 1,2,3,4,5,6,7,8,9,10,11 were monitored by RP-HPLC after 30 min.

d. Hydrazone ligation by using H-LYRAP/V-NHNH$_2$ and 8-Quinolinecarbaldehyde



The peptide hydrazide and 8-Quinolinecarbaldehyde (30 μM and 200 μM in final concentration) were added together in the aqueous PME buffer (100 mM PIPES, 1 mM $MgSO_4$, 2 mM EGTA at pH 7.0). Then, the mixture was stirred at neutral pH and 25℃.



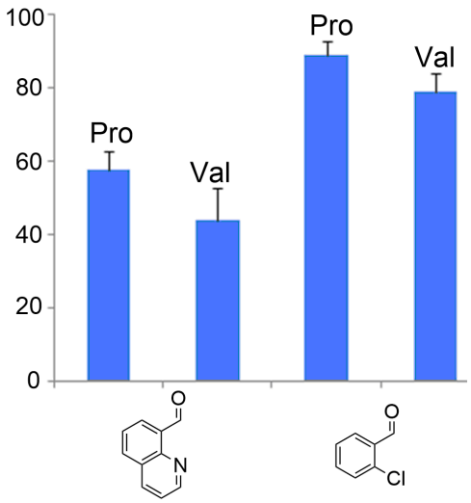
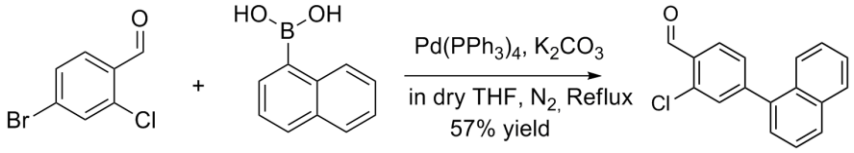


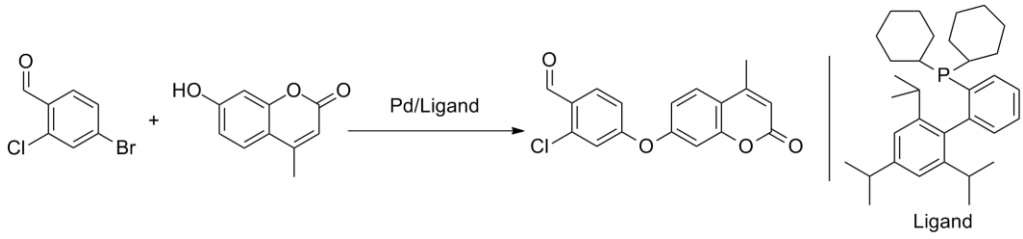
Figure S4: The comparison of ligation efficiency between 8-Quinolinecarbaldehyde and 2-chlorobenzaldehydewithpeptide that containing large steric hindrance residues. The reaction was monitored by RP-HPLC.

e. The synthesis of fluorophore 5, 7



2-chloro-4-bromobenzaldehyde (219 mg, 1mmol) and 1-naphthylboronic acid 189 mg, 1.1 mmol) were dissolved in dry THF (2 mL) under nitrogen. Then $Pd(PPh_3)_4$ (115 mg, 0.1 mmol) and potassium carbonate (276 mg, 2 mmol) added to the reaction. The reaction was stirred vigorously for 20 h under reflux condition. The reaction was diluted with EtOAc, washed with water and brine, dried over Na_2SO_4, filtrated and concentrated. The crude product was purified by chromatography (SiO_2, 10:1 Petroleum ether /EtOAc) to yield fluorophore **5** (excitation wavelength 286 nm, emission wavelength 321 nm, 150 mg, 57%).

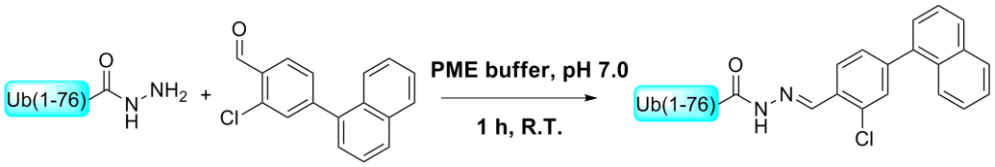
fluorophore 7



2-chloro-4-bromobenzaldehyde (87.6 mg, 0.4 mmol) and 4-Methylumbelliferone (84.5 mg, 0.48 mmol) were dissolved in dry toluene (1.5 mL) under nitrogen. Then $Pd(OAc)_2$ (1.8 mg, 0.004 mmol), X-Phos (11.4 mg, 0.024 mmol) and potassium phosphate (169.6 mg, 0.8 mmol) added to the reaction. The reaction was stirred vigorously for 10 h under 100℃. Then the reaction was diluted with EtOAc, washed with water and brine, dried over Na_2SO_4, filtrated and concentrated. The crude product was purified by chromatography (SiO_2, 4:1 Petroleum ether /EtOAc) to yield

fluorophore 7 (excitation wavelength 372 nm, emission wavelength 445 nm, 38.9 mg, 31%).

f. The expression of Ub-(1-76)-$NHNH_2$ and its C terminal labeling

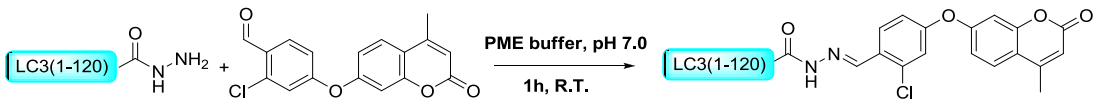


Ub sequence:
MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQ
KESTLHLVLRLRGG

The DNA sequence for the first 76 amino residues of ubiquitin was isolated by PCR using the 5' primer (5'-GGTGGTCATATGCCGTCCGAGAAGACCT-3') and 3' primer (5'-GGTGGTTGCTCTTCCGCACGTCTCCTGGGAGGC-3'), encoding NdeI and BspQI restriction site, respectively. The PCR amplified DNA was purified by agarose gel extraction (QIA quick kit), digested with NdeI and BspQI, then ligated into a NdeI, BspQI digested pTXB1 bacterial expression vector (New England Biolab, Beverly, MA). Then, *E. coli* ER2566 (NEB) cells transformed with the pTXB1-Ub(1-76) plasmid were grown to mid-log phase (OD600 = 0.6) in 1 liter LB medium and induced with 0.2 mM IPTG at 16°C over night. Cells were harvested by using ultrasonification and resuspended in lysis buffer containing 20 mM Hepes (pH 7.4), 500 mM NaCl, 0.1% Triton X-100 (v/v) and 1 mM EDTA. After cell lysis, the insoluable material was removed by centrifugation the soluble fraction was applied to chitin resin pre-equilibrated with lysis buffer and incubated 2 h at room temperature. The beads were washed with 10 column volumes of lysis buffer. The protein was cleaved from the resin by incubation of the beads with 1 column volume of cleavage buffer (20 mM Tris-HCl, 300 mM NaCl, 4% $NH_2NH_2 \cdot H_2O$, pH 7.5) for 3-4 hours to obtain Ub(1-76)-$NHNH_2$. The column was eluted and the resin was washed with 2 column volume cleavage buffer. Then the cleavage protein was concentrated by using centrifugal filters (3000 MW cutoff) and further purified by process RP-HPLC using a 30-45% B gradient over 60 min, yielding 5 mg of lyophilized protein. The protein acylhydrazine was characterized by ESI-MS.

The compounds **5** was firstly dissolved in ethanol (10 mM) for stock solution and then added to reaction buffer in 300-500 μM final concentration. Ub-$NHNH_2$ was dissolved in 6 M Gn·HCl (1 mg in 100 μL) for stock solution and then added to reaction buffer in 30-50 μM final concentration. 10%-20% (v/v) ethanol was added for dissolving the insoluble compounds. The reaction was incubated for 60 min and detected by analytical RP-HPLC.

g. The C terminal labeling of LC3(1-120)-$NHNH_2$



LC3 sequence:
PSEKTFKQRRSFEQRVEDVRLIREQHPTKIPVIIERYKGEKQLPVLDKTKFLVPDHVNMSEL
IKIIRRRLQLNANQAFFLLVNGHSMVSVSTPISEVYESERDEDGFLYMVYASQETFG
LC3(1-120)-$NHNH_2$ was obtained by using established method.[1]

Fluorophore **8** was firstly dissolved in ethanol (10 mM) for stock solution and then added to reaction buffer in 300 μM final concentration. LC3-$NHNH_2$ was dissolved in 6 M Gn·HCl (1 mg in 100 μL) for stock solution and then added to reaction buffer in 30 μM final concentration. 10%-20% (v/v) ethanol was added for dissolving the insoluble compounds. The reaction was incubated for 60 min and detected by analytical RP-HPLC.

h. Fluorescence SDS-PAGE analysis of labeled protein

For SDS-PAGE, samples were loaded onto 12% SDS-PAGE gelsand ran for 30 min at 80 V and 50 min at150 V. All the protein samples were prepared in Tris buffer containing bromophenol blue. When samples were ran near the edge of gel, it was removed from electrophoresis chamber and first irradiated under 365 nm UV for fluorescence analysis. Then, the gel was stained in Coomassie brilliant blue for 1 h for SDS-PAGE analysis.

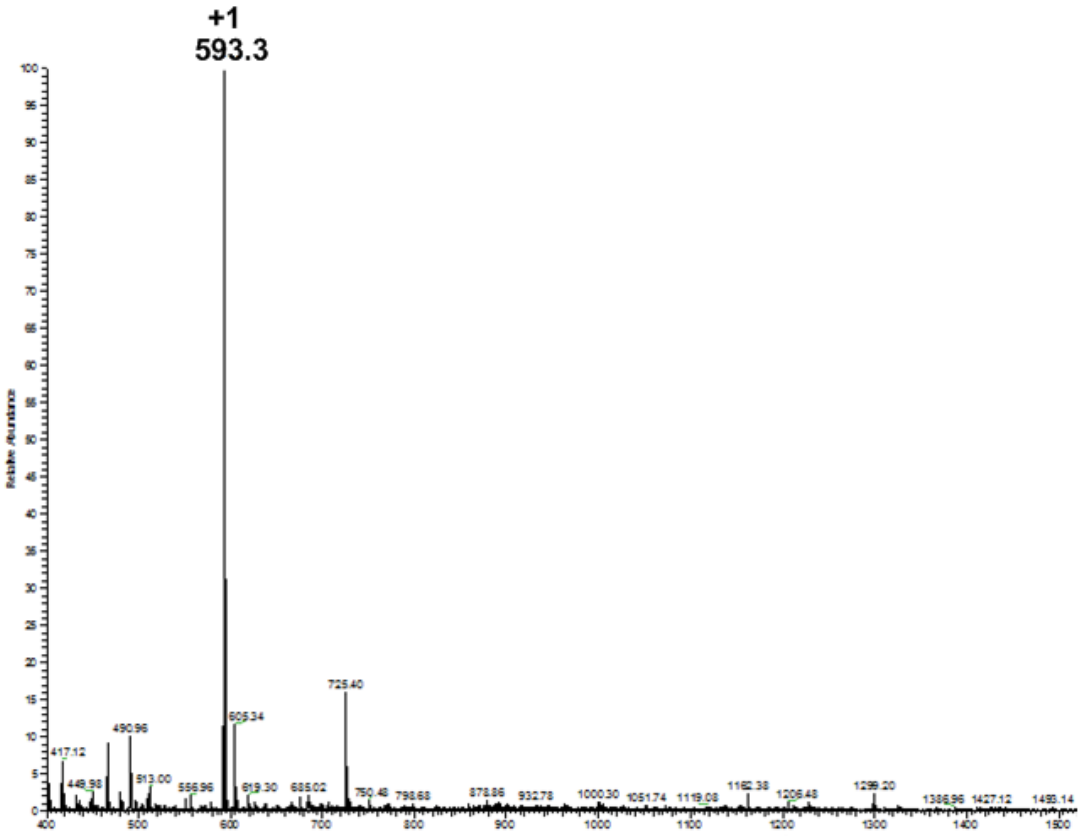
Reference:

1. G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li, Y. Lin, H. C. Cui and L. Liu, *Angew. Chem., Int. Ed.*, 2011, **50**, 7645–7649.

3. Spectra

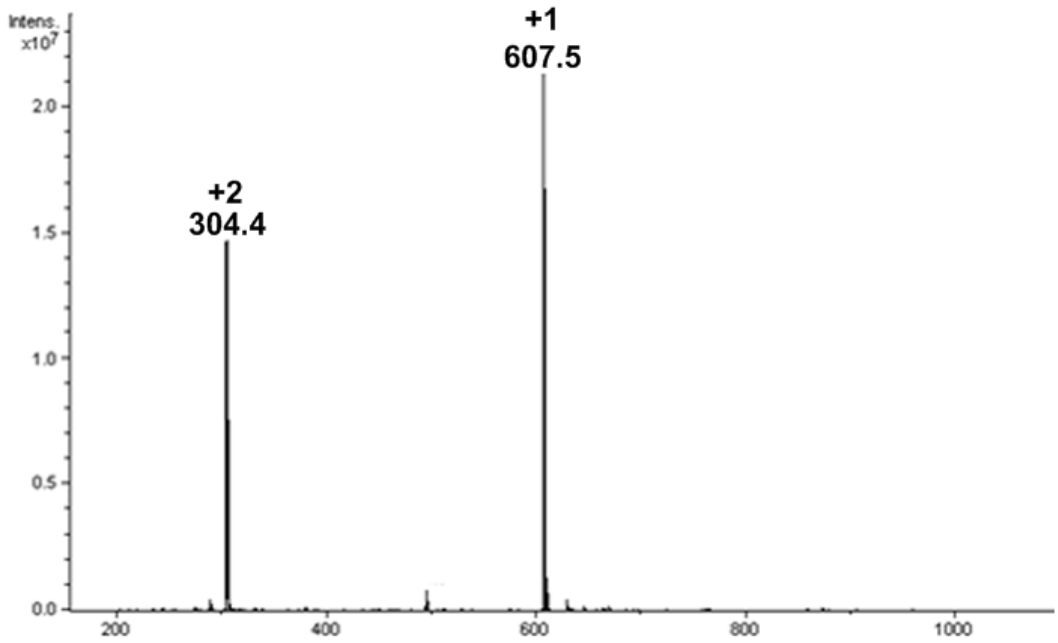
a. MS

1) LYRAG-$NHNH_2$



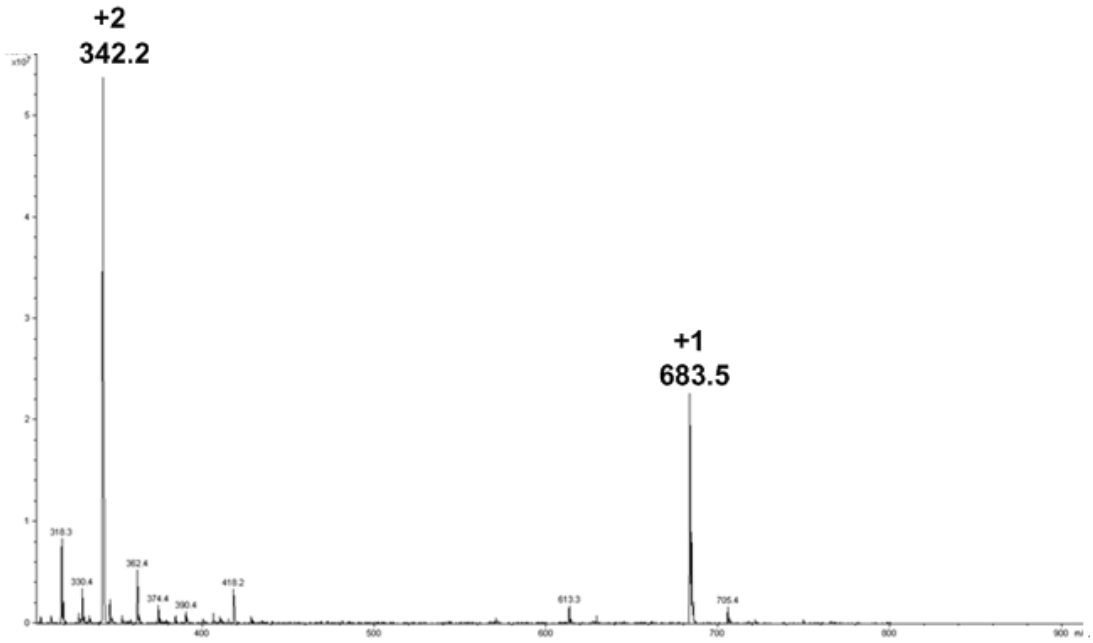
Mass observed: 593.3 Da, mass calculated: 592.6 Da

2) LYRAA-$NHNH_2$



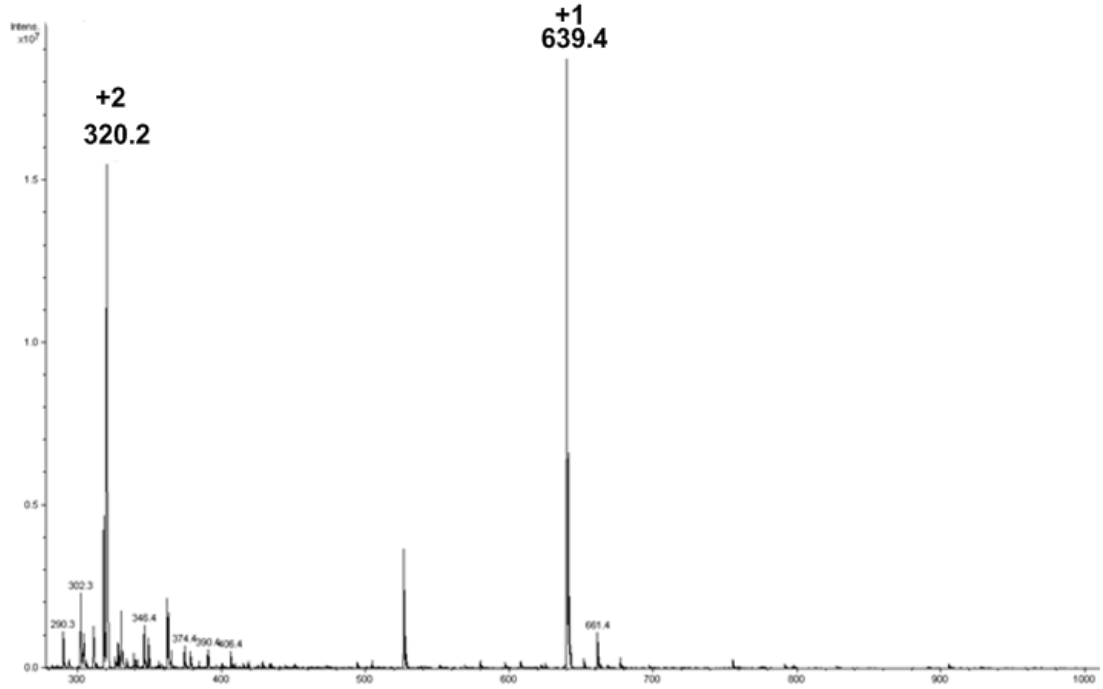
Mass observed: 606.5 Da, mass calculated: 606.6 Da

3) LYRAF-$NHNH_2$



Mass observed: 682.5 Da, mass calculated: 682.7 Da

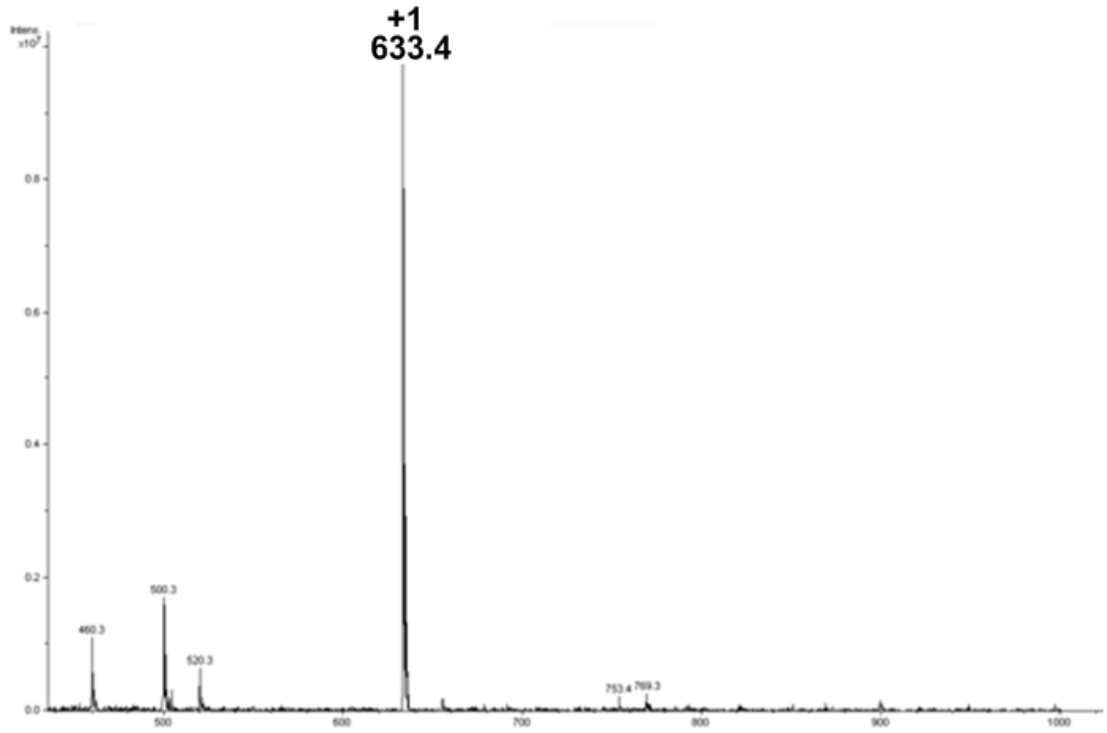
4) LYRAC-$NHNH_2$



Mass observed: 638.4 Da, mass calculated: 638.7 Da

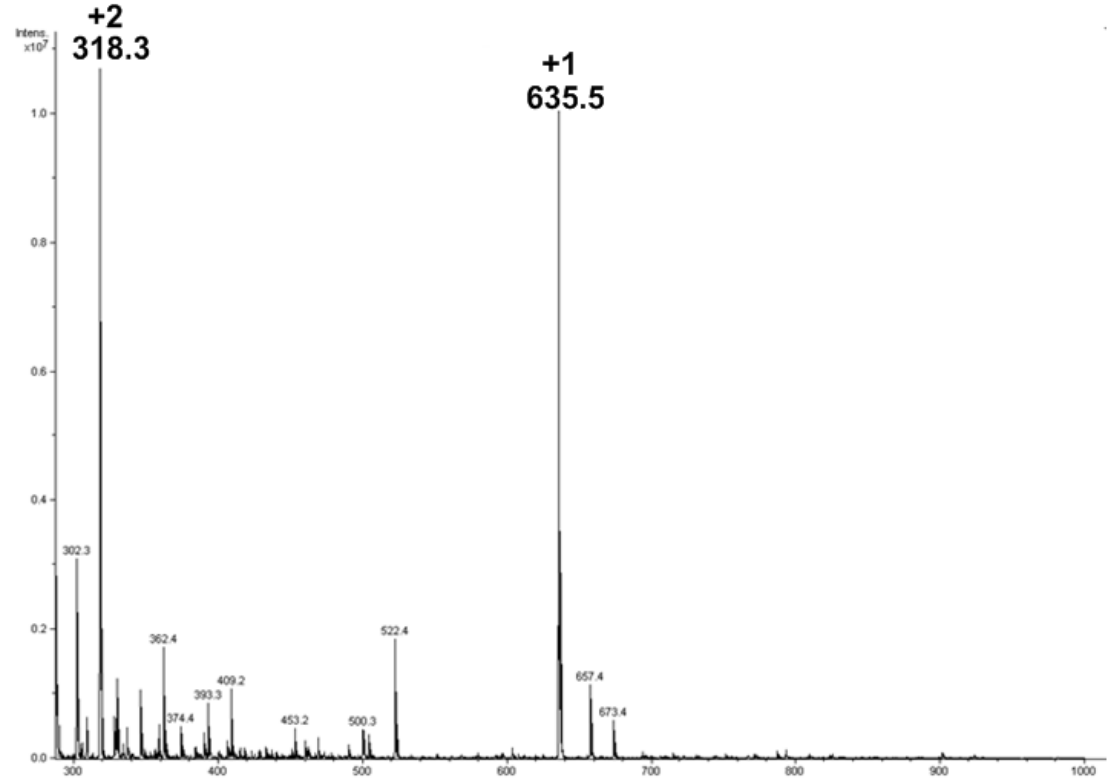
\

5) LYRAP-$NHNH_2$



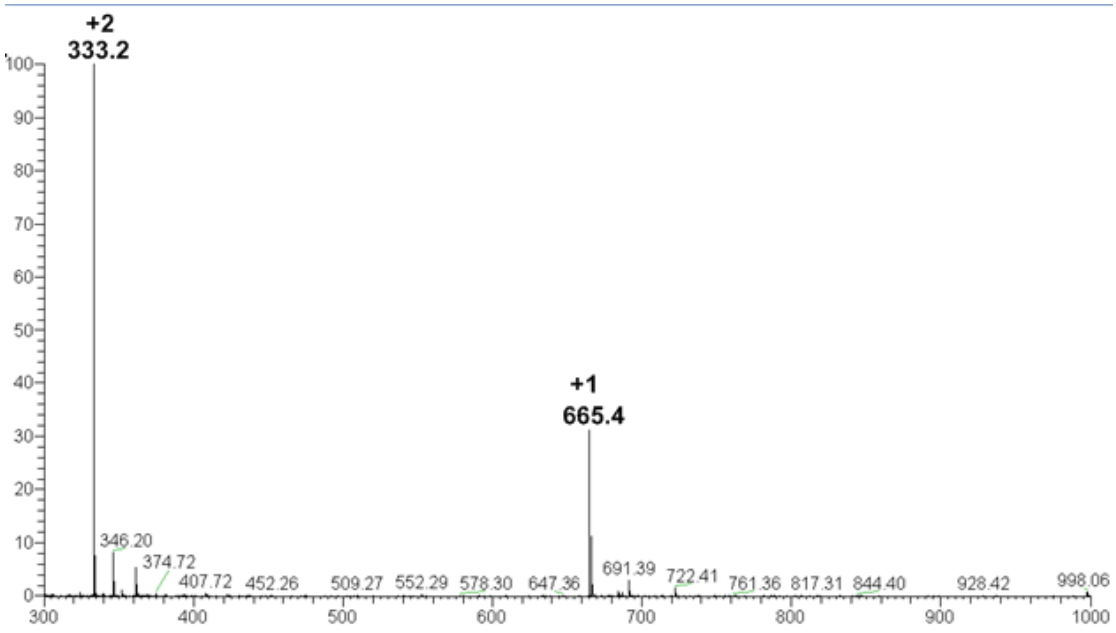
Mass observed: 632.4 Da, mass calculated: 632.7 Da

6) LYRAV-$NHNH_2$



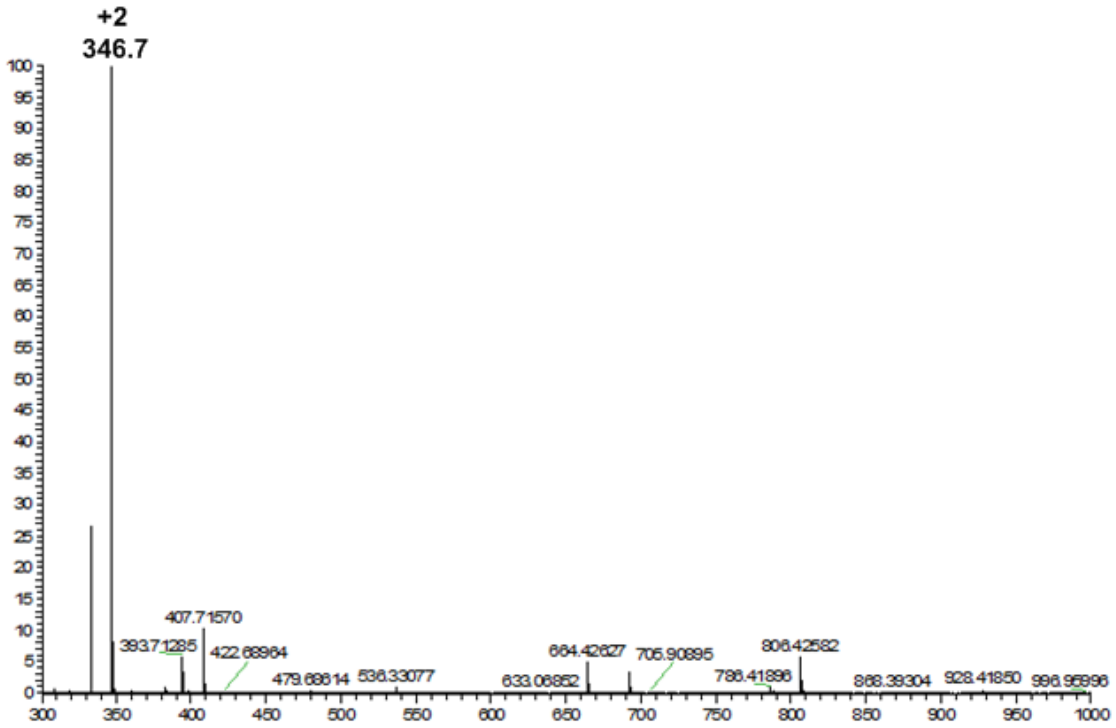
Mass observed: 634.5 Da, mass calculated: 634.7 Da

7) LYRAE-$NHNH_2$



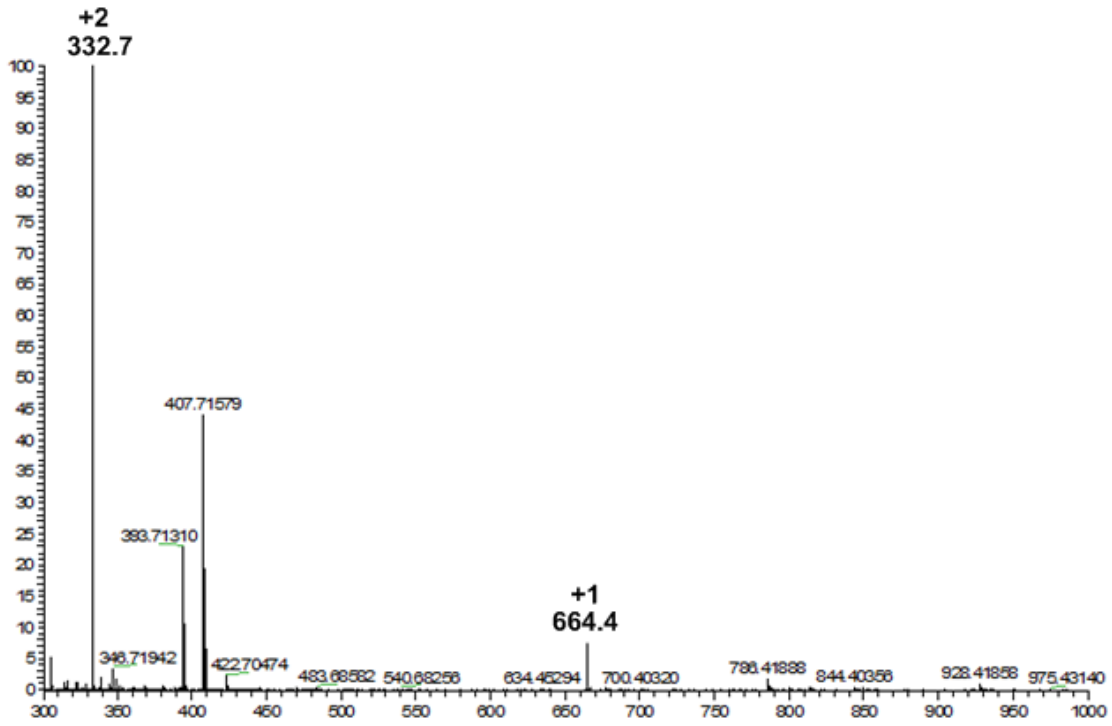
Mass observed: 664.4 Da, mass calculated: 664.7 Da

8) LYRAR-$NHNH_2$



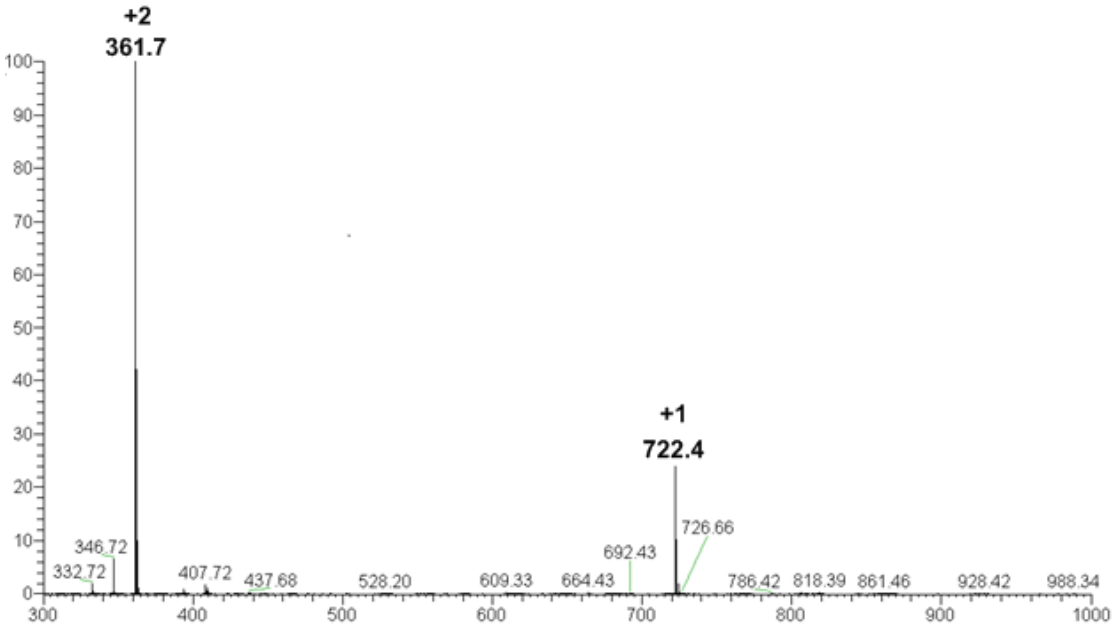
Mass observed: 691.4 Da, mass calculated: 691.8 Da

9) LYRAK-$NHNH_2$

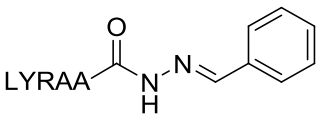


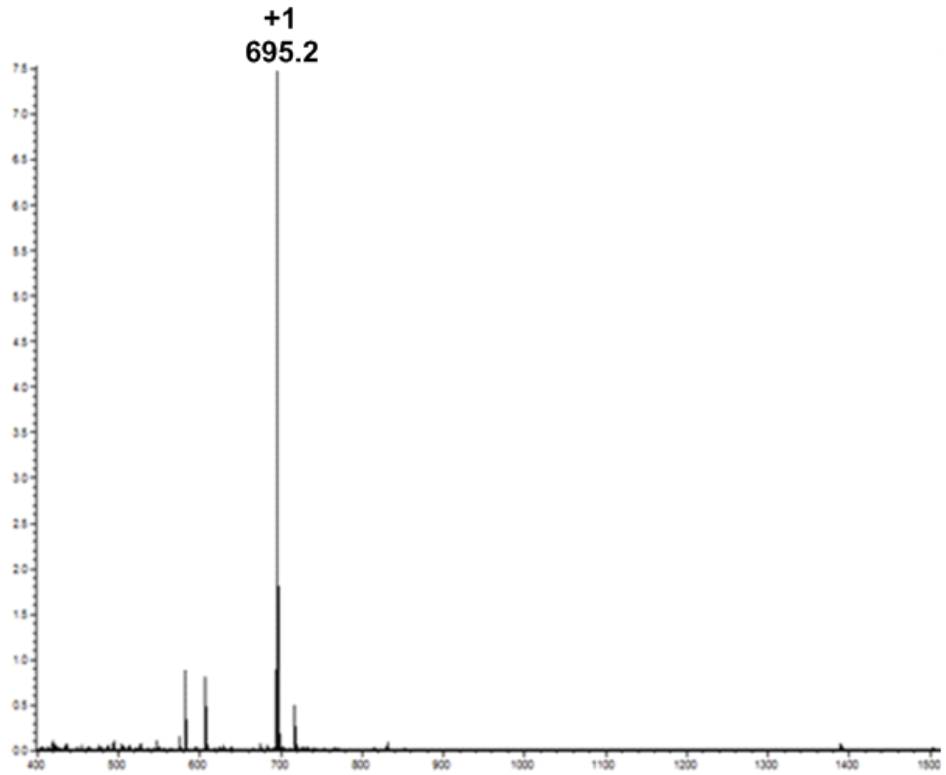
Mass observed: 663.4 Da, mass calculated: 663.7 Da

10) LYRAW-NHNH$_2$



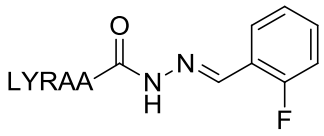
Mass observed: 721.4 Da, mass calculated: 721.8 Da

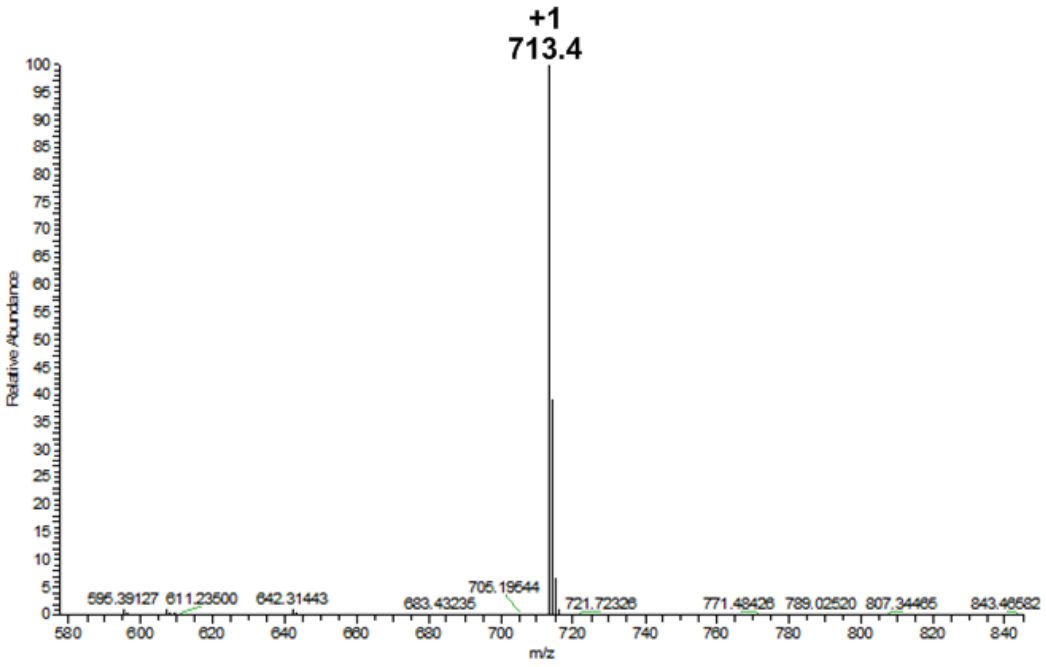
11) 



Mass observed: 694.2 Da, mass calculated: 694.6 Da

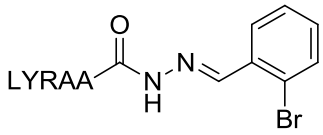
12)

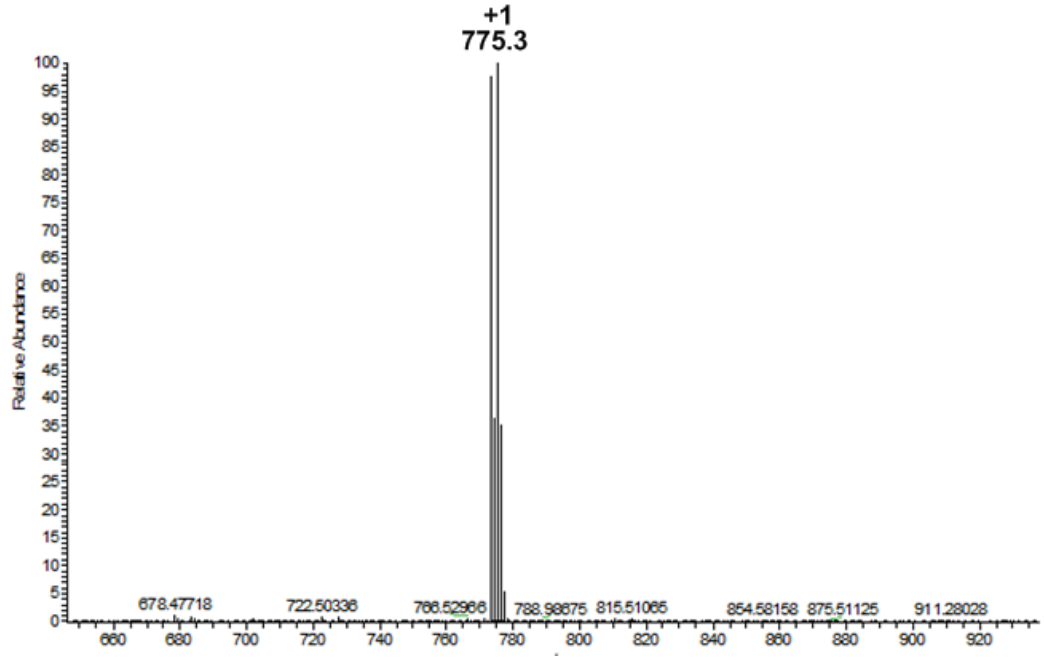




Mass observed: 712.4 Da, mass calculated: 712.6 Da

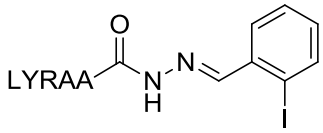
13)

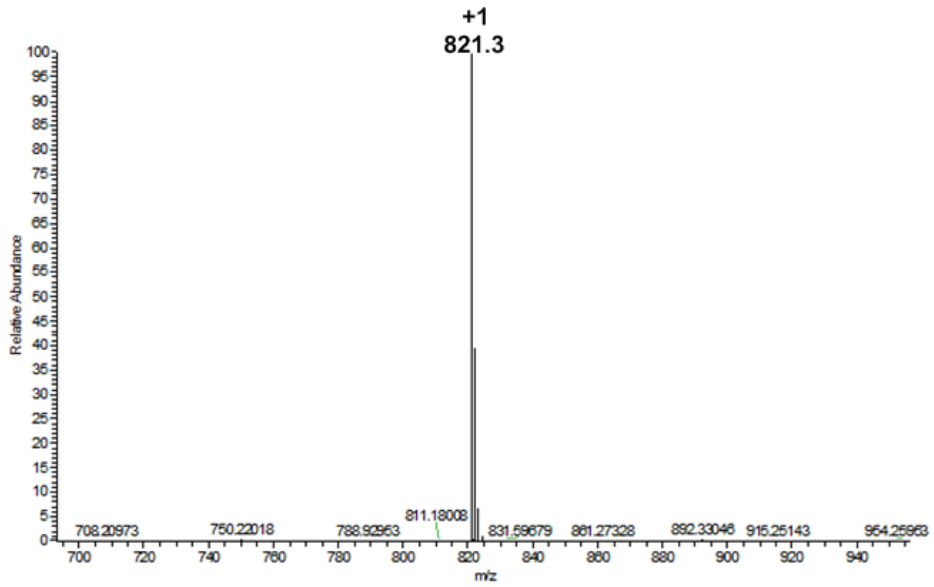




Mass observed: 774.3 Da, mass calculated: 773.5 Da

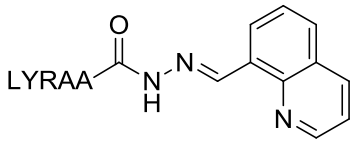
14)

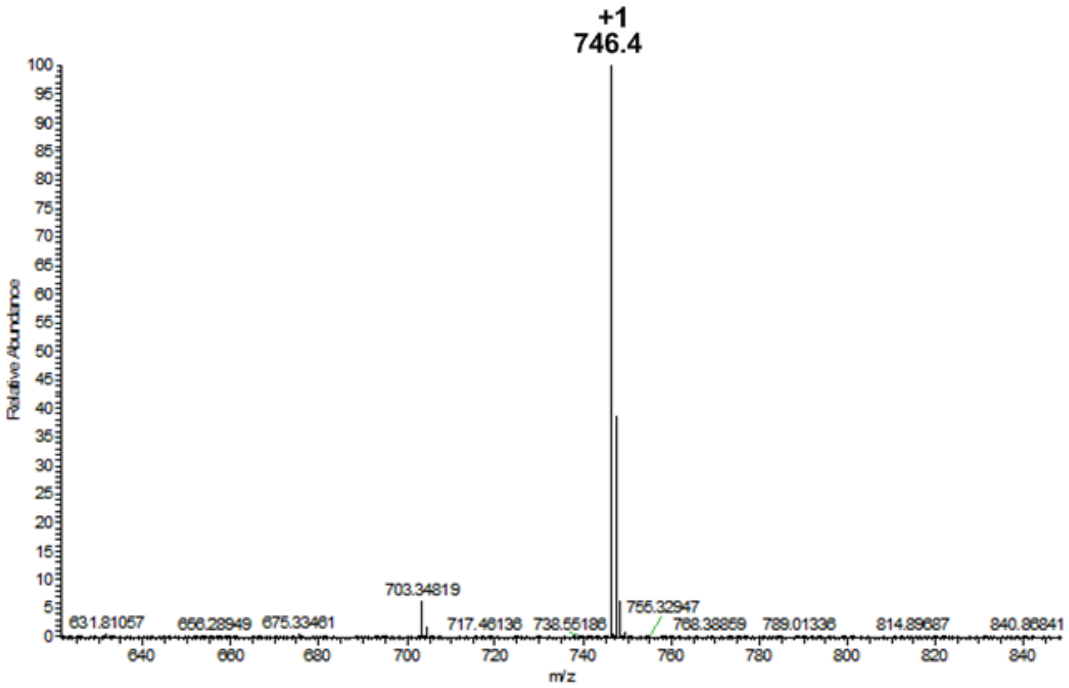




Mass observed: 820.3 Da, mass calculated: 820.5 Da

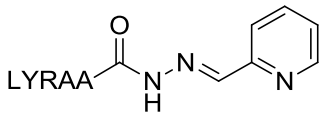
15)

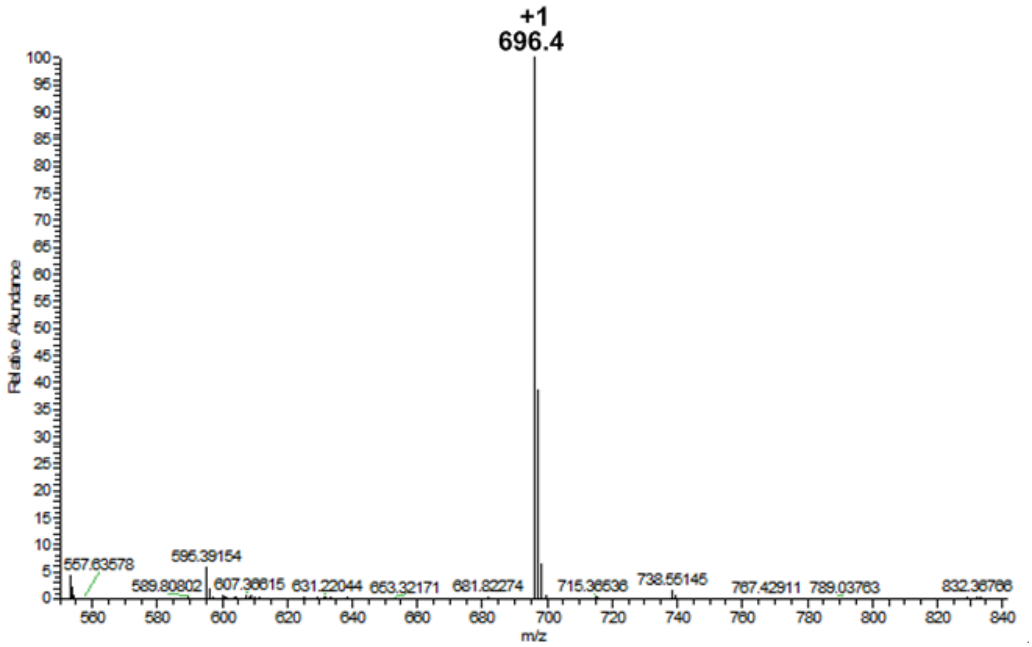




Mass observed: 745.4 Da, mass calculated: 745.7 Da

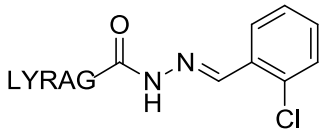
16)

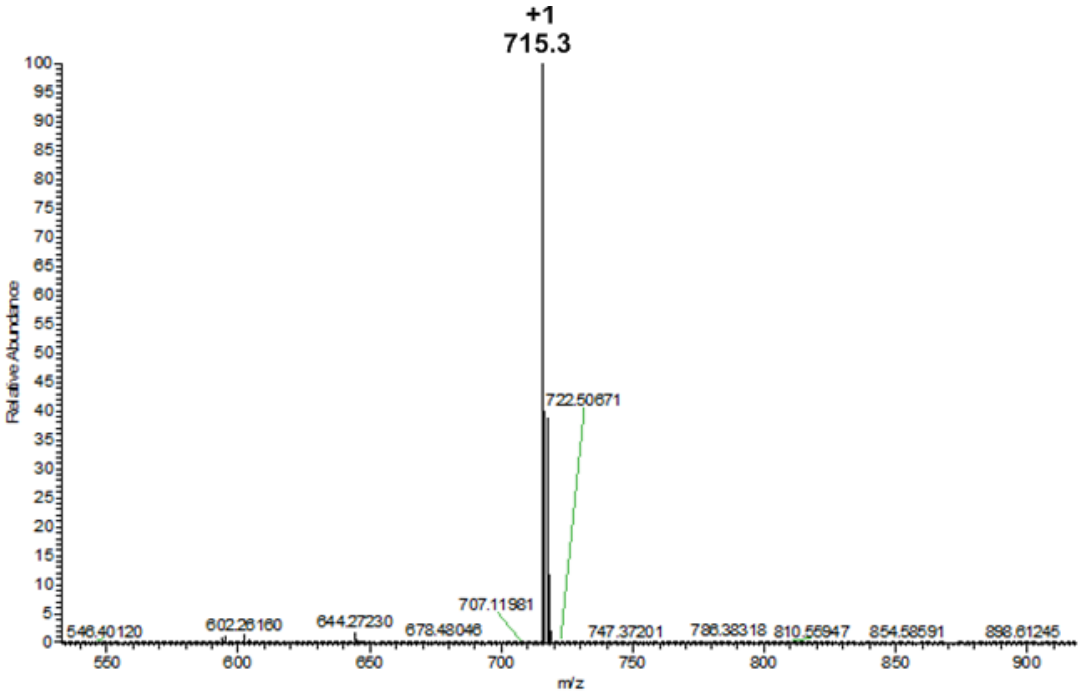




Mass observed: 695.4 Da, mass calculated: 695.6 Da

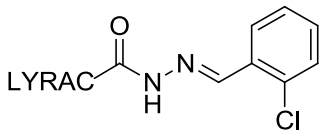
17)

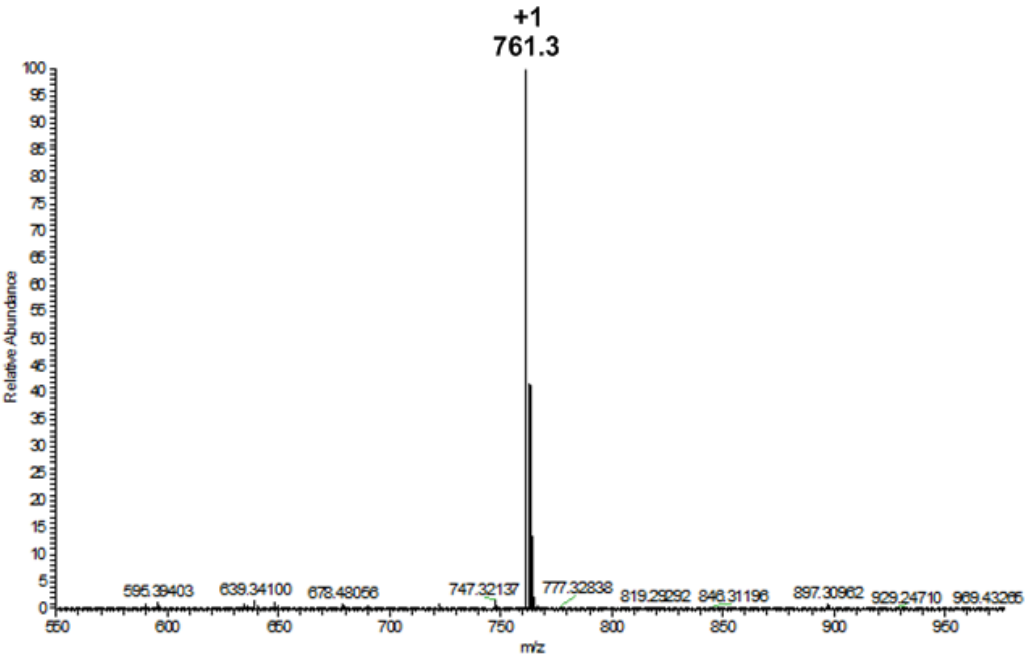




Mass observed: 714.3 Da, mass calculated: 715.9 Da

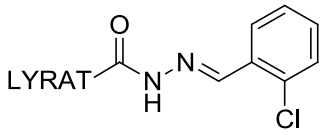
18)

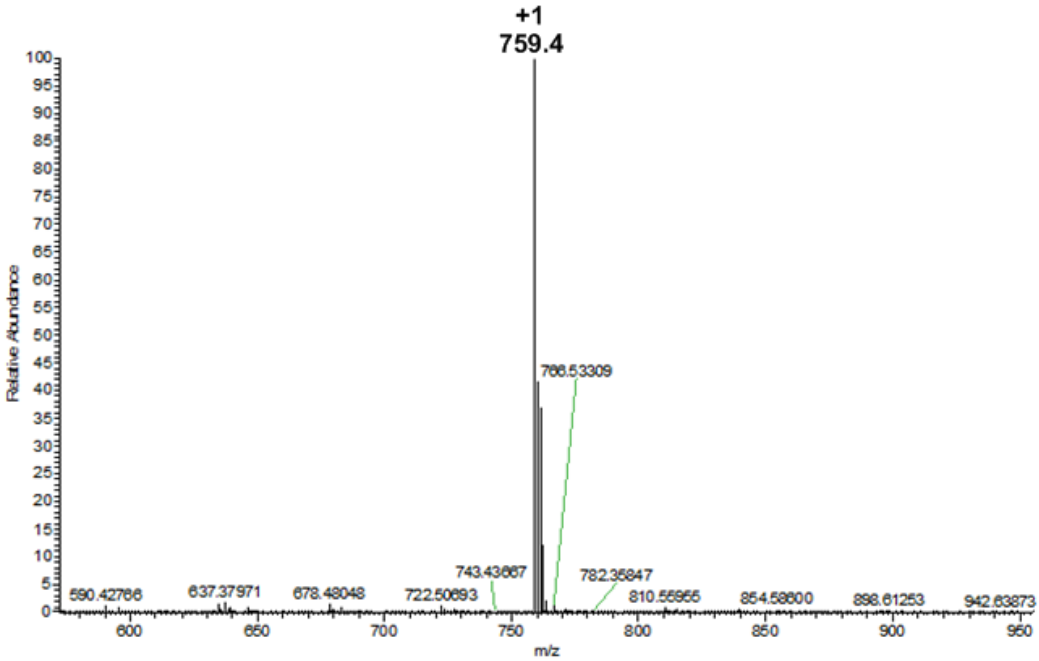




Mass observed: 762.3 Da, mass calculated: 761.3 Da

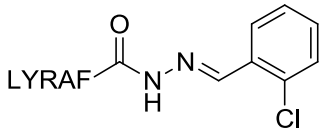
19)

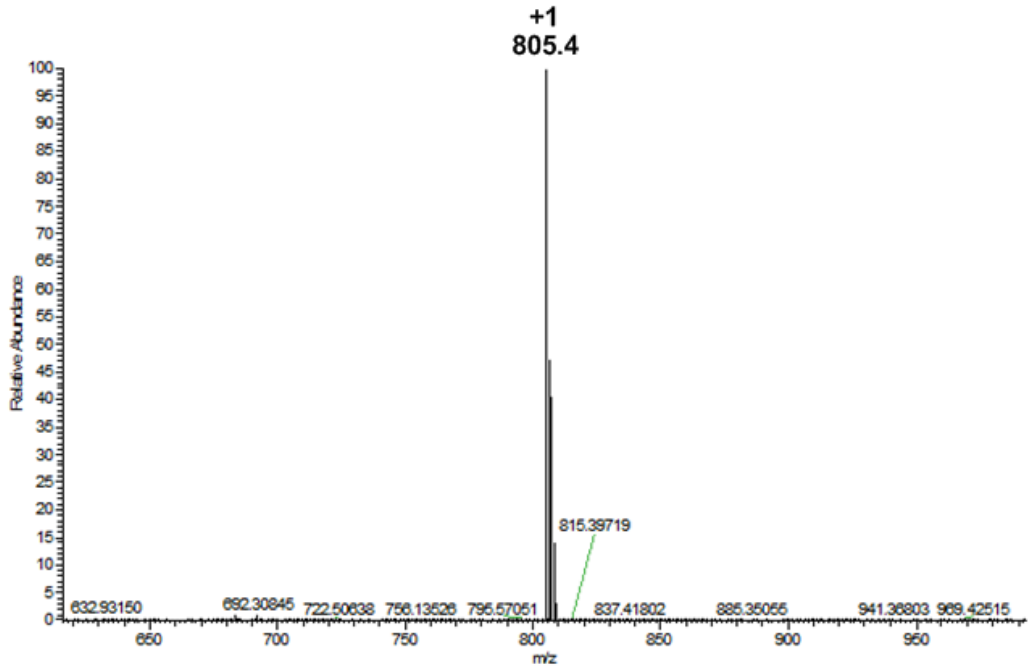




Mass observed: 758.4 Da, mass calculated: 759.3 Da

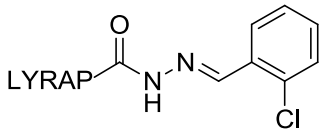
20)

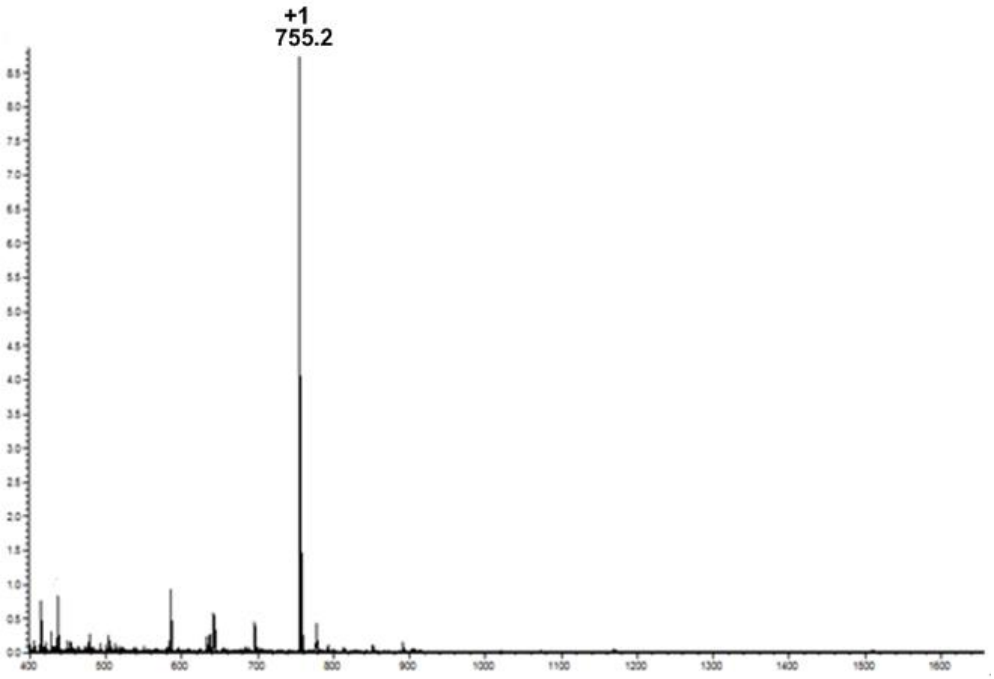




Mass observed: 804.4 Da, mass calculated:805.3 Da

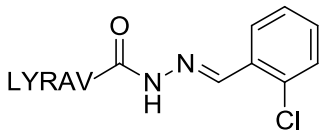
21)

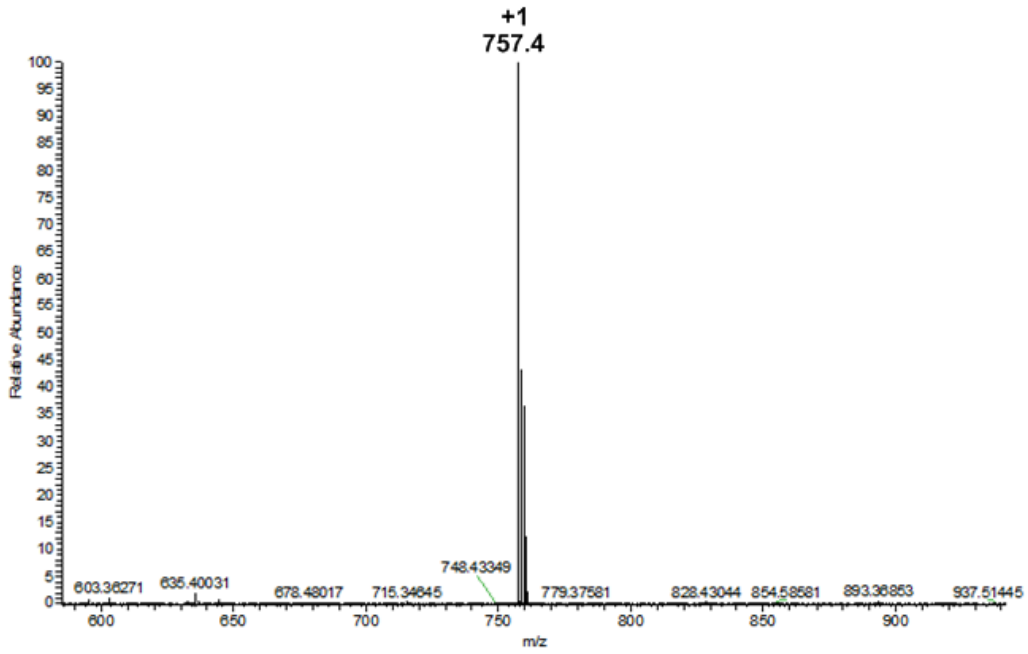




Mass observed: 754.2 Da, mass calculated: 755.3 Da

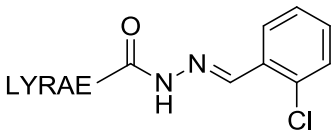
22)

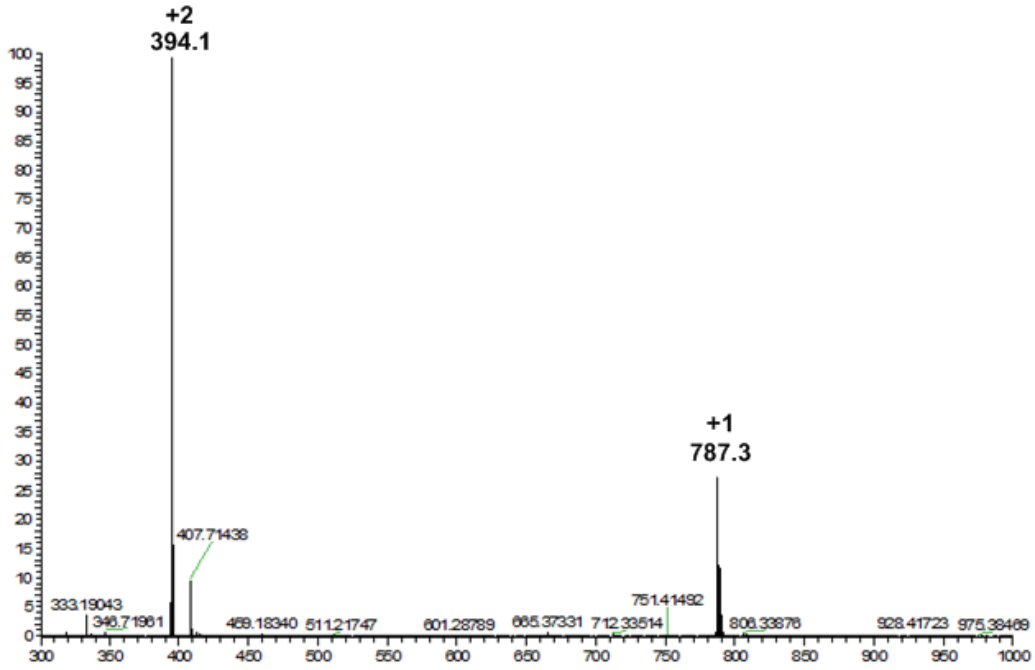




Mass observed: 756.4 Da, mass calculated: 757.1 Da

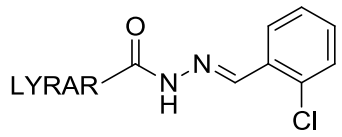
23)

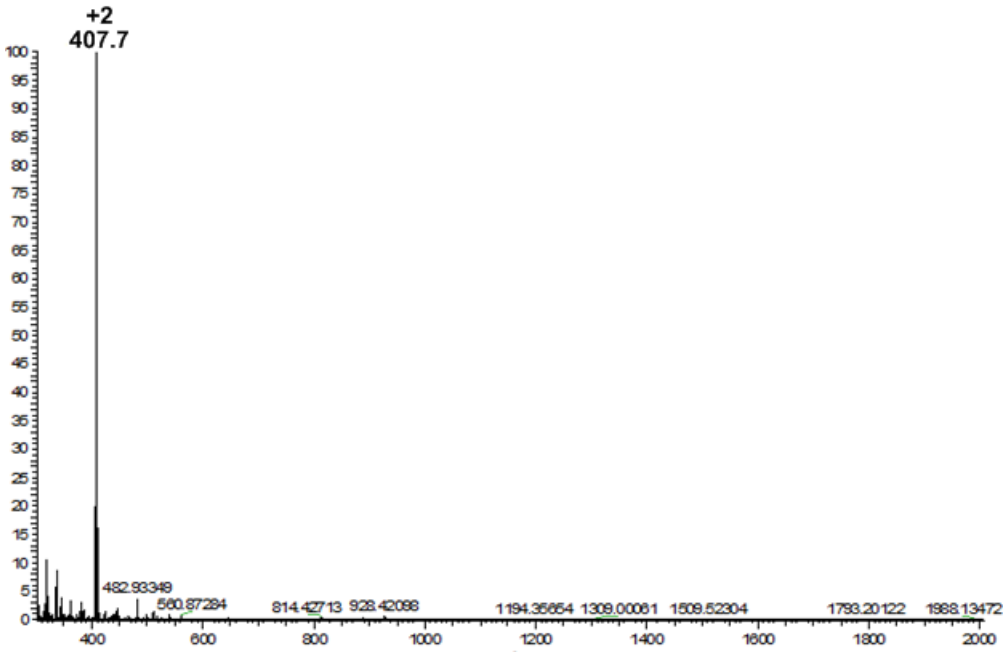




Mass observed: 786.3 Da, mass calculated: 787.3 Da

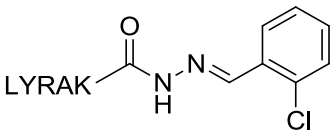
24)

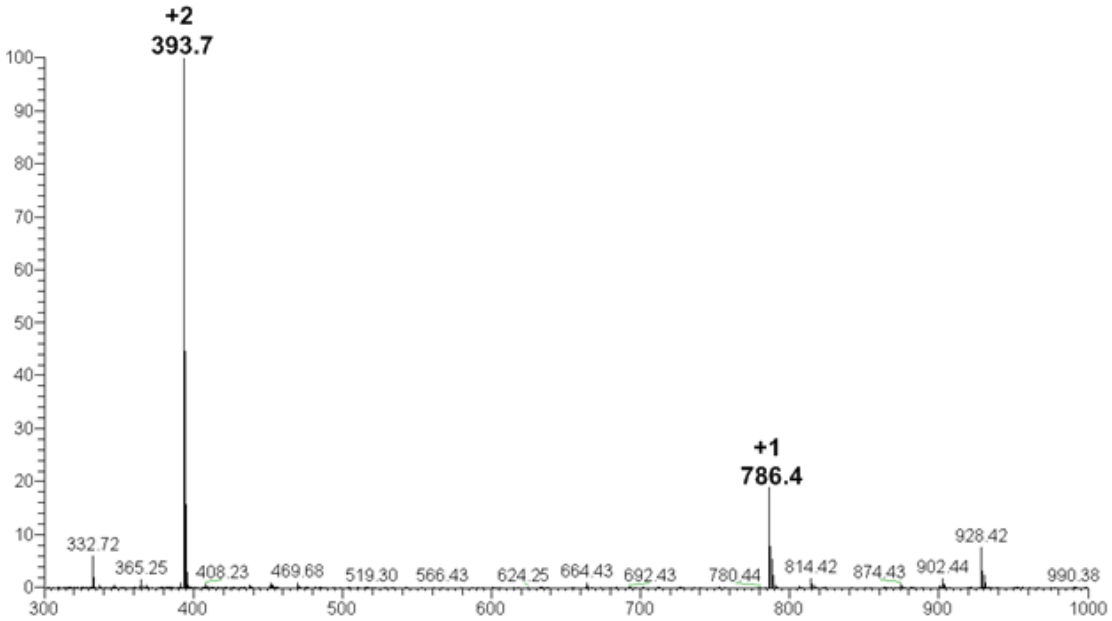




Mass observed: 813.4 Da, mass calculated: 814.4 Da

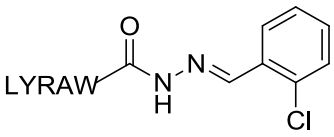
25)

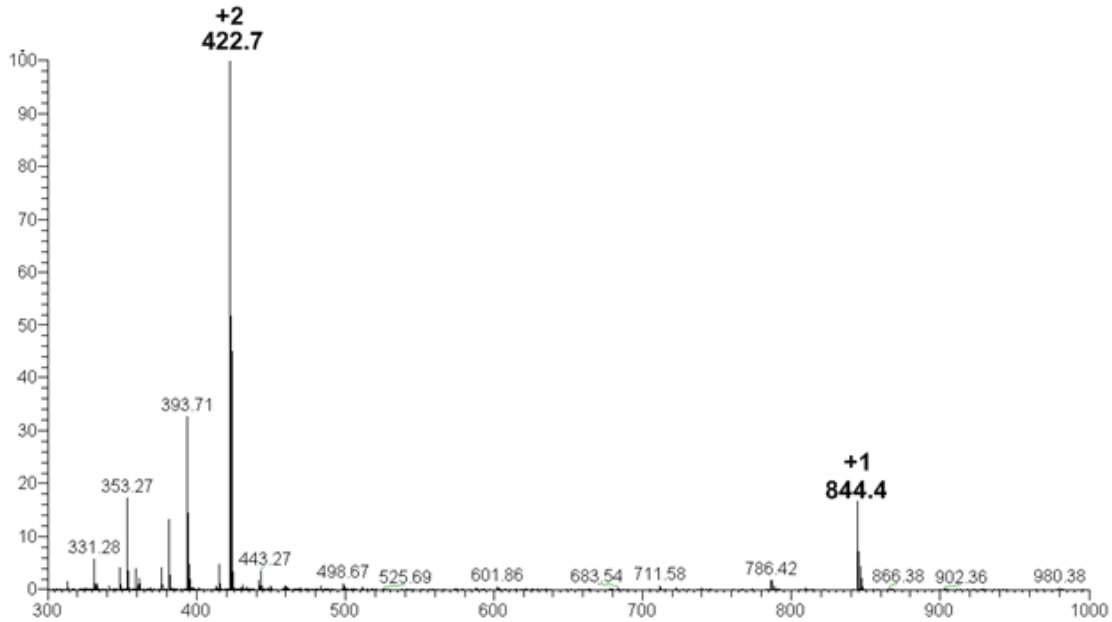




Mass observed: 785.4 Da, mass calculated: 786.3 Da

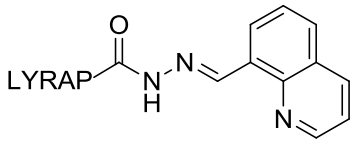
26)

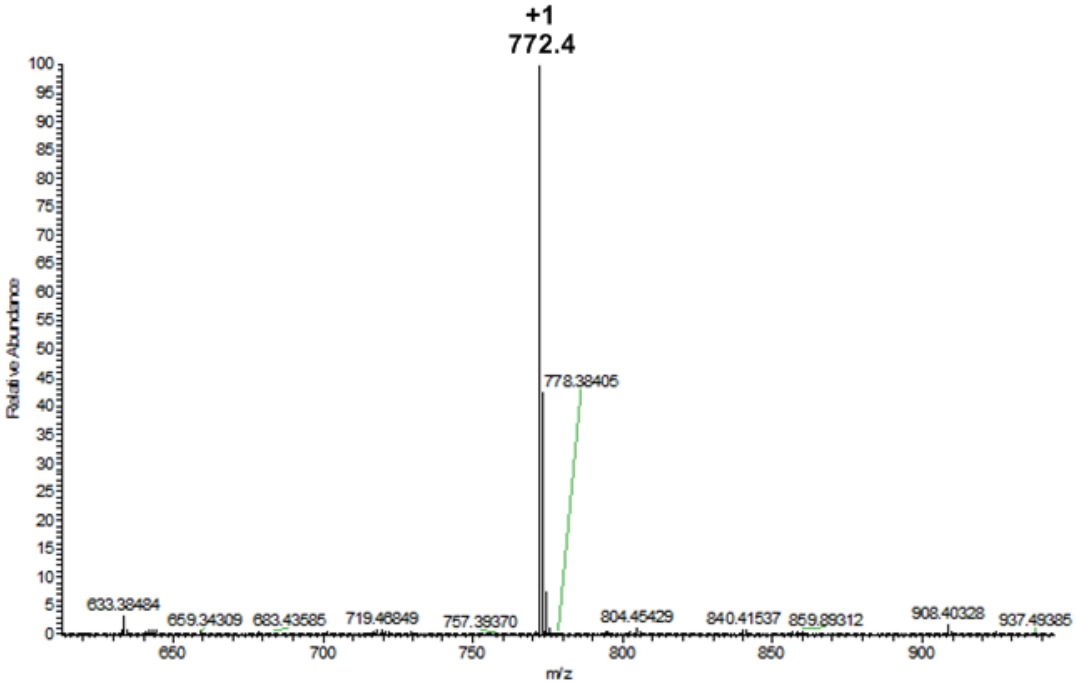




Mass observed: 843.4 Da, mass calculated: 844.4 Da

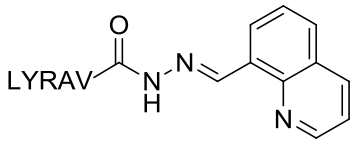
27)

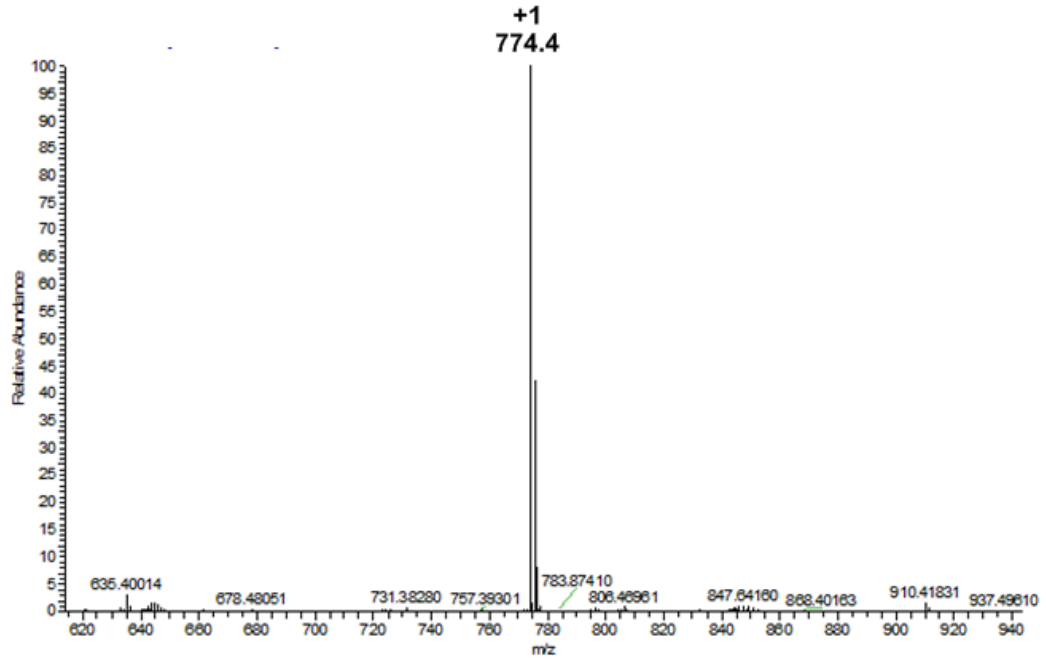




Mass observed: 771.4 Da, mass calculated: 771.9 Da

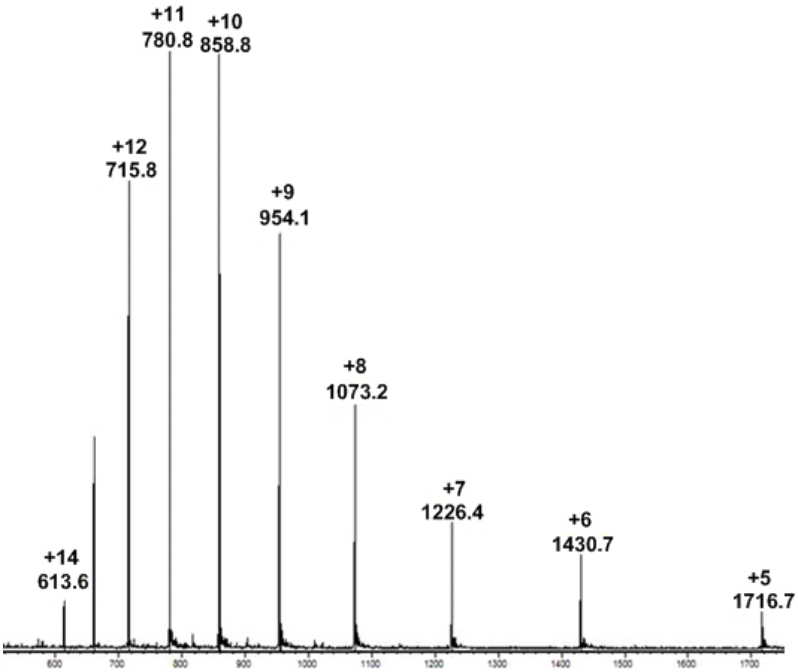
28)





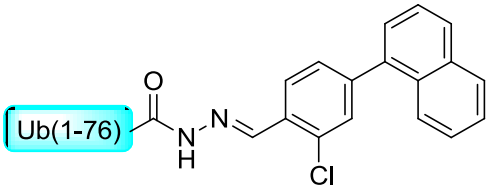
Mass observed: 773.4 Da, mass calculated: 773.9 Da

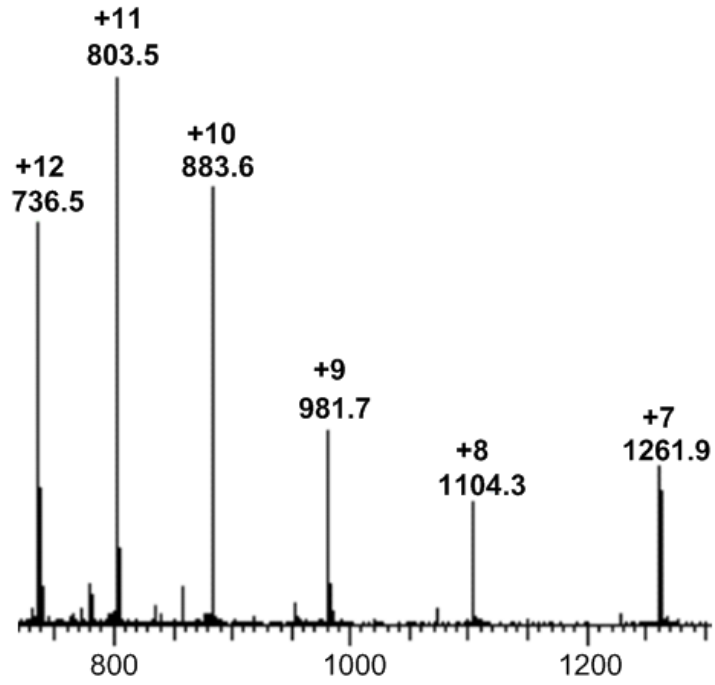
29) Ub(1-76)—$NHNH_2$



Mass observed: 8577.9 Da, mass calculated: 8578.8 Da

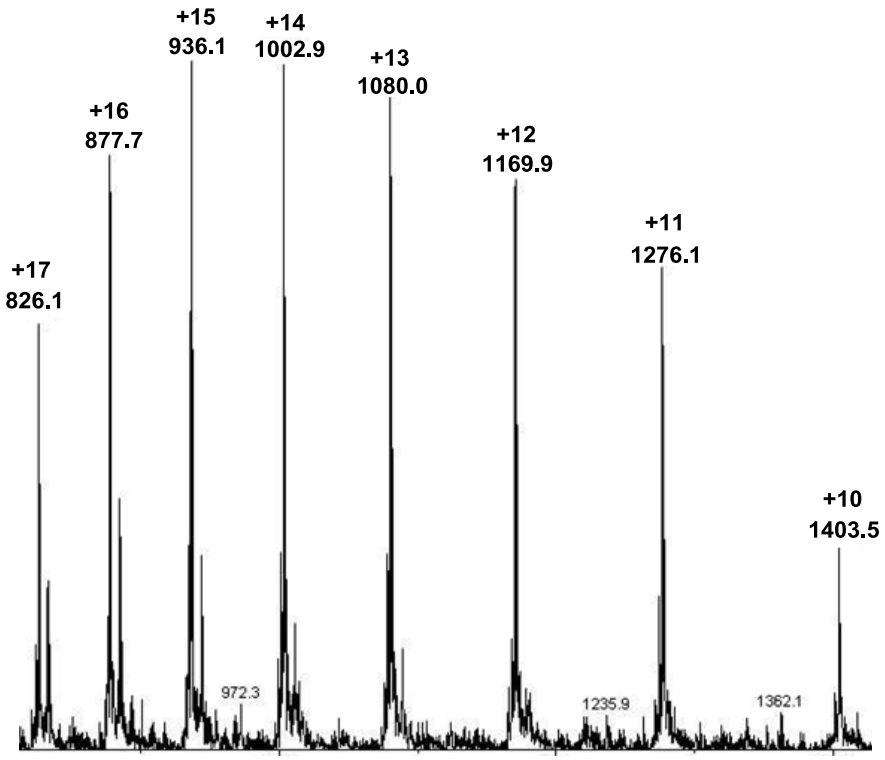
30)





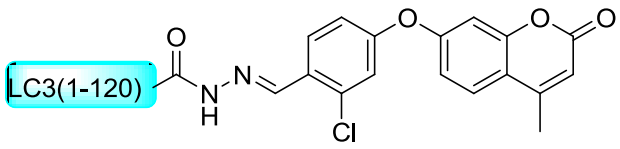
Mass observed: 8826.7 Da, mass calculated: 8827.6 Da

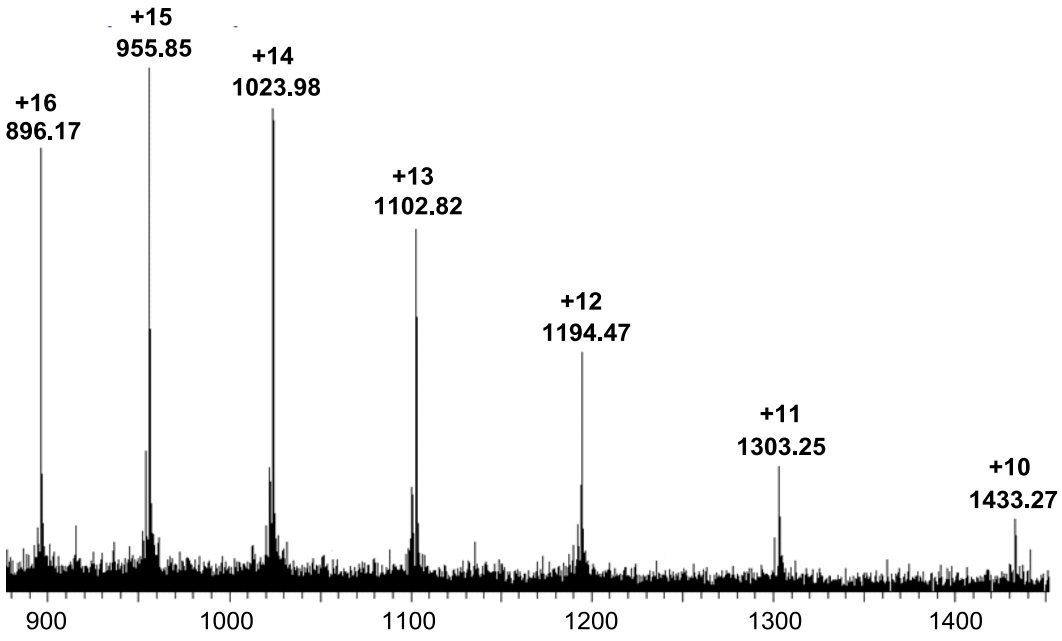
31) LC3(1-120)—$NHNH_2$



Mass observed: 14026.7 Da, mass calculated: 14026.1 Da

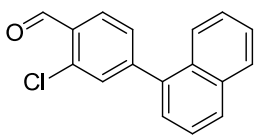
32) LC3(1-120)

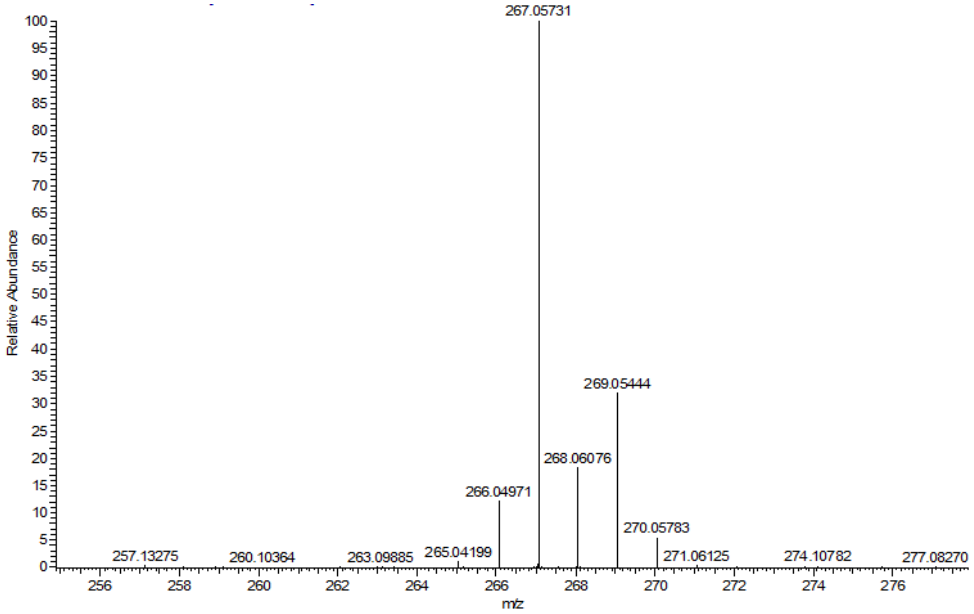




Mass observed: 14322.75 Da, mass calculated: 14322.84 Da

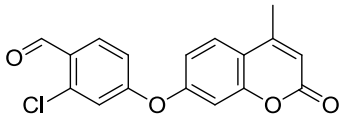
33)

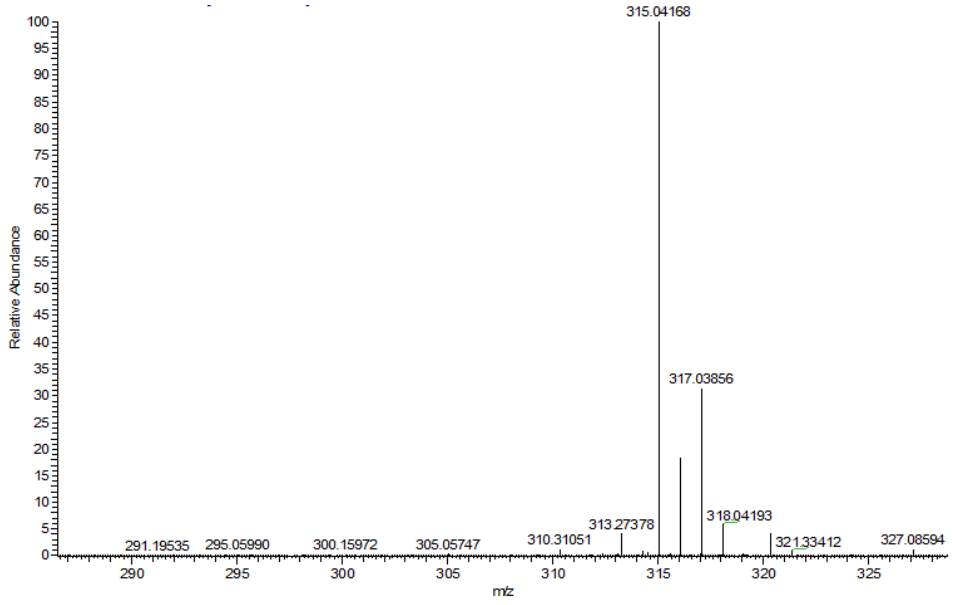




m/z	Intensity	Relative	Theo. Mas	Delta (mm	Composition
267.0573	3.37E+08	100	267.0571	0.19	C17 H12 O Cl
269.0544	1.07E+08	31.76	269.0542	0.27	C17 H12 O [37]Cl

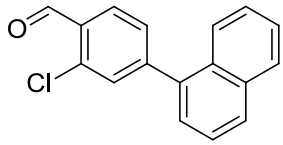
34)





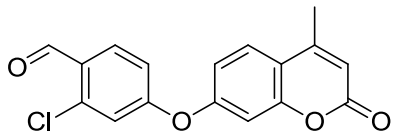
m/z	Intensity	Relative	Theo. Mas	Delta (mm	Composition
315.0417	3.14E+08	100	315.0419	-0.18	C17 H12 O4 Cl
317.0386	98274168	31.33	317.0389	-0.36	C17 H12 O4 [37]Cl

b. NMR spectra data



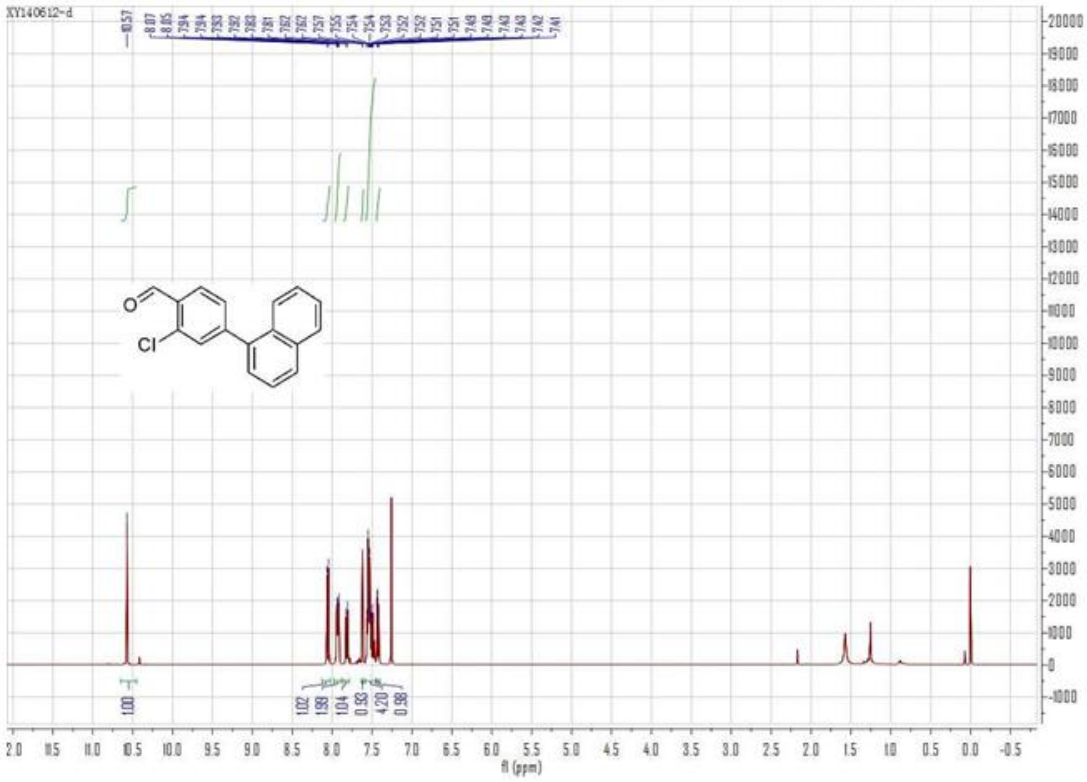
Fluorophore 5: (2-chloro-4-(naphthalen-1-yl)benzaldehyde)

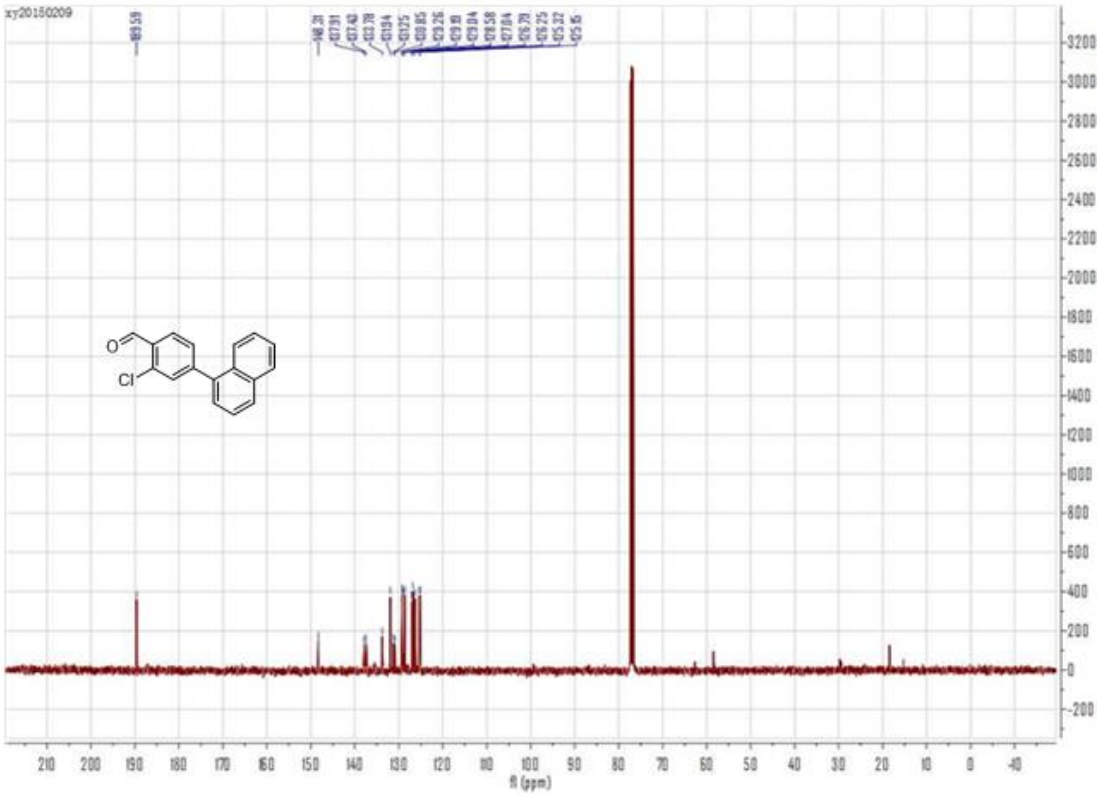
^{1}H NMR (400 MHz, $CDCl_3$) δ 10.57 (s, 1H), 8.05 (d, *J* = 7.9 Hz, 1H), 7.93 (dd, *J* = 7.9, 3.4 Hz, 2H), 7.82 (d, *J* = 8.4 Hz, 1H), 7.62 (d, *J* = 1.5 Hz, 1H), 7.58 – 7.45 (m, 4H), 7.42 (dd, *J* = 7.0, 1.0 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 189.59, 148.31, 137.91, 137.43, 133.78, 131.94, 131.25, 130.85, 129.26, 129.19, 129.04, 128.58, 127.04, 126.79, 126.25, 125.32, 125.15.

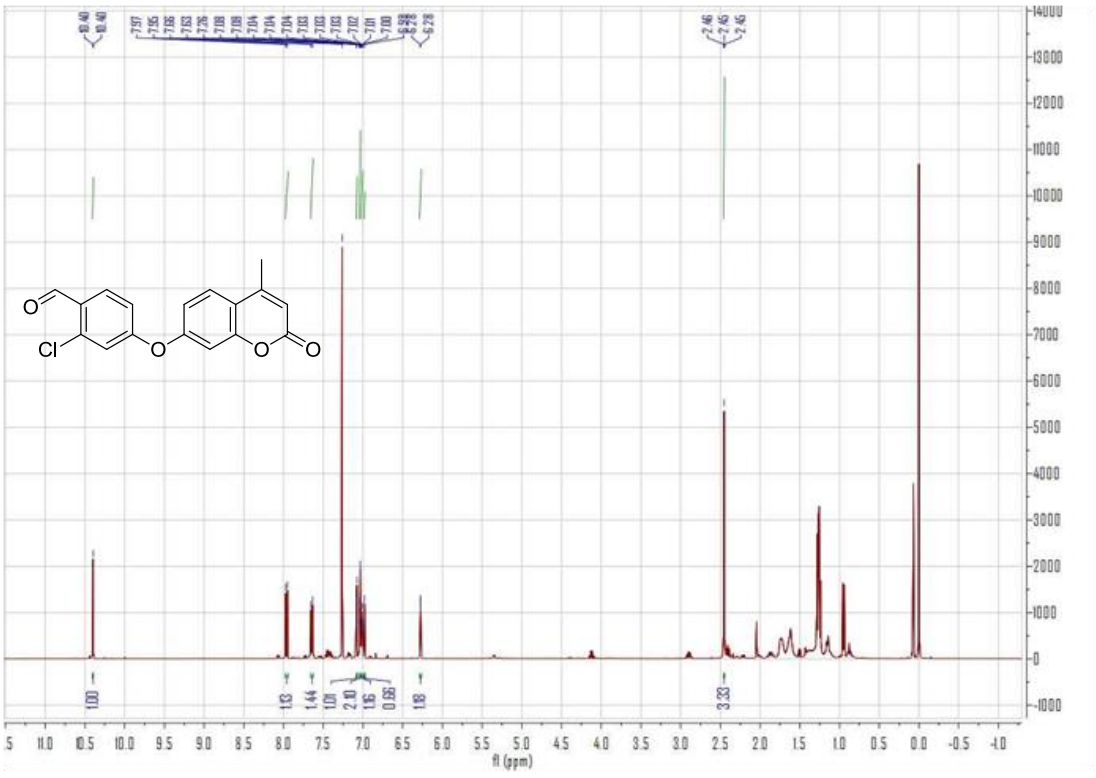


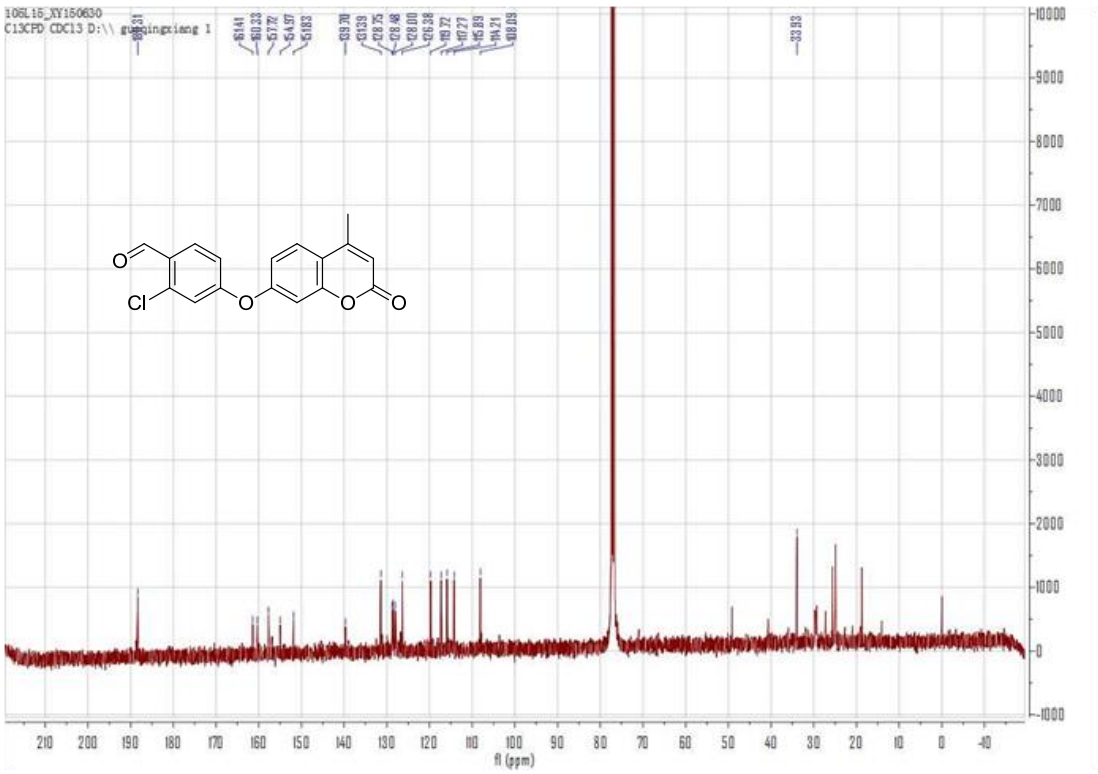
Fluorophore 7: (2-chloro-4-((4-methyl-2-oxo-2H-chromen-7-yl)oxy)benzaldehyde)

^{1}H NMR (400 MHz, $CDCl_3$) δ 10.40 (d, *J* = 0.8 Hz, 1H), 7.96 (d, *J* = 8.6 Hz, 1H), 7.65 (d, *J* = 8.3 Hz, 1H), 7.08 (d, *J* = 2.3 Hz, 1H), 7.05 – 7.03 (m, 2H), 7.03 – 6.99 (m, 1H), 6.28 (d, *J* = 1.2 Hz, 1H), 2.45 (t, *J* = 1.9 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 188.31, 161.41, 160.33, 157.72, 154.97, 151.83, 139.70, 131.39, 128.75, 128.48, 128.00, 119.72, 117.27, 115.89, 114.21, 108.09, 33.93.




xy20150209
O
Cl
f1 (ppm)


O
Cl
O
O
O
f1 (ppm)


106L15_XY150630
C13CPD CDC13 D:\\ gugingxiang 1
O
Cl
O
O
O
f1 (ppm)