Supporting Information

Structure-based Discovery of Cell-Potent Peptidomimetic Inhibitors for Protein N-terminal Methyltransferase 1

Dongxing Chen,[†] Guangping Dong,[†] Youchao Deng,[†] Nicholas Noinaj,[§] and Rong Huang[*,†]

[†] Department of Medicinal Chemistry and Molecular Pharmacology, Purdue Institute for Drug Discovery, Purdue University Center for Cancer Research, Purdue University, West Lafayette, Indiana 47907, United States

[§] Department of Biological Sciences, Markey Center for Structural Biology, and the Purdue Institute of Inflammation, Immunology and Infectious Disease, Purdue University, West Lafayette, Indiana 47907, United States

Corresponding Author * huang-r@purdue.edu

Table of Contents

Experimental section

Chemistry general procedures. Starting materials, reagents, and solvents were obtained from commercial sources. Preparative high-pressure liquid chromatography (RP-HPLC) was performed on Agilent 1260 Series system. Systems were run with a 0-95% methanol/water gradient with a 0.1% TFA modifier. Matrix-assisted laser desorption ionization mass spectra (MALDI-MS) data were acquired in positive-ion mode using a Sciex 4800 MALDI TOF/TOF MS. Except for the known compounds, new compounds were also characterized and confirmed by TLC-MS or MALDI-MS. The peptides were synthesized using a Liberty Automated Microwave Peptide Synthesizer (CEM) with the manufacturer's standard coupling cycles at 0.1 mmol scales. The purity of final compounds was confirmed by Waters LC-MS system and/or Agilent 1260 Series system. Systems were run with a 0-40% methanol/water gradient with 0.1% TFA. All the purity of target compounds showed >95%.

(S)-N-((S)-1-(((1H-benzo[d]imidazol-2-yl)methyl)amino)-6-amino-1-oxohexan-2-yl)-1-(2-(naphthalen-1-yl)acetyl)pyrrolidine-2-carboxamide (**15**). To a solution of **20** (0.14 g, 0.27 mmol) in MeOH: H_2O = 3:1 (4 mL) was added $LiOH·H_2O$ (0.05 g, 1.08 mmol) at 0 ºC. Then the reaction was stirred overnight at r.t. 1 N HCl (aq.) was added to the reaction mixture in an ice bath to adjust the pH to around 4 and the mixture was extracted with DCM (5 mL X 3). The combined organic layer was dried over Na_2SO_4 and then was filtrated, concentrated to get crude product. To a solution of the crude product (0.05g, 0.10 mmol) in DMF (2 mL), HBTU (g, 0.13 mmol) and HOBt (g, 0.13 mmol) was added in portion at 0 ºC. The mixture was stirred at r.t. for 30 min, (1H-benzo[d]imidazol-2-yl)methanamine hydrochloride (0.02 g, 0.13 mmol) and DIPEA (44 μL, 0.25 mmol) were added into the solution. The mixture was stirred overnight at r.t. Then 20 ml brine was added into the mixture and the mixture was extract with DCM (3 mL X 3). The combined organic layer was dried over Na_2SO4 and then was filtrated, concentrated and the residue was purified by flash column chromatography (CH_2Cl_2 : MeOH = 20:1) and get 45 mg product (70%).Then the purified product (45 mg, 0.07mmol) was dissolved in CH_2Cl_2 (1mL), and 4N HCl in dioxane (1 mL) was added into the solution at 0 ºC. Then the mixture was stirred at r.t. for 30 min and the volatiles were removed under vacuum. The residue was purified through HPLC to obtain 26 mg compound (68%).^{1}H NMR (500 MHz, D_2O) δ 7.55 (d, *J* = 8.5 Hz, 1H), 7.47 (d, *J* = 8.2 Hz, 1H), 7.39 (d, *J* = 7.5 Hz, 1H), 7.29-7.25 (m, 5H), 7.23 – 7.16 (m, 2H), 7.14 – 7.05 (m, 2H), 4.34 – 4.18 (m, 3H), 4.19 – 4.08 (m, 1H), 4.01 – 3.90 (m, 1H), 3.85 – 3.76 (m, 1H), 3.74 – 3.55 (m, 2H), 2.78 – 2.68 (m, 2H), 2.30 – 2.15 (m, 1H), 2.03 – 1.72 (m, 4H), 1.73 – 1.58 (m, 1H), 1.58 – 1.38 (m, 2H), 1.34 – 1.07 (m, 2H). ^{13}C NMR (126 MHz, D_2O) δ 175.23, 174.41, 173.39, 148.78, 133.06, 131.54, 130.79, 129.92, 128.39, 127.96, 127.61, 126.29, 126.14, 125.82, 125.50, 123.46, 117.46, 115.14, 113.48, 61.31, 53.31, 48.15, 39.08, 38.20, 34.96, 29.49, 29.45, 25.97, 24.60, 22.16. ESI-MS (m/z): 541.3635 $[M + H]^+$.

(S)-N-((S)-6-amino-1-oxo-1-(phenylamino)hexan-2-yl)-1-(2-(naphthalen-1-yl)acetyl)pyrrolidine-2-carboxamide (**18**). Compound was prepared according to the procedure for **15**. ^{1}H NMR (500 MHz, D_2O) δ 7.87 – 7.71 (m, 3H), 7.52 – 7.36 (m, 3H), 7.35 – 7.22 (m, 2H), 7.22 – 7.14 (m, 1H), 7.06 – 6.97 (m, 2H), 6.81 – 6.74 (m, 1H),

4.39 – 4.33 (m, 1H), 4.28 – 4.23 (m, 1H), 4.16 – 4.01 (m, 1H), 3.85 – 3.74 (m, 2H), 3.58 – 3.50 (m, 1H), 2.90 – 2.66 (m, 2H), 2.34 – 2.22 (m, 1H), 2.13 – 1.82 (m, 4H), 1.76 – 1.60 (m, 1H), 1.61 – 1.46 (m, 2H), 1.42 – 1.26 (m, 2H). ESI-MS (m/z): 487.2479 $[M + H]^+$.

(S)-N-((S)-6-amino-1-(naphthalen-1-ylamino)-1-oxohexan-2-yl)-1-(2-(naphthalen-1-yl)acetyl)pyrrolidine-2-carboxamide (**19**). Compound was prepared according to the procedure for **15**. ^{1}H NMR (500 MHz, Methanol-*d*4) δ 7.99 – 7.68 (m, 6H), 7.57 – 7.32 (m, 6H), 7.28 – 7.21 (m, 1H), 7.20 – 7.14 (m, 1H), 4.75 – 4.69 (m, 1H), 4.62 – 4.55 (m, 1H), 4.56 – 4.43 (m, 1H), 4.27 – 4.19 (m, 1H), 4.18 – 4.09 (m, 1H), 3.94 – 3.77 (m, 2H), 2.92 – 2.76 (m, 1H), 2.74 – 2.52 (m, 1H), 2.36 – 2.26 (m, 1H), 2.26 – 2.13 (m, 1H), 2.14 – 2.01 (m, 3H), 1.94 – 1.81 (m, 1H), 1.70 – 1.35 (m, 6H). ESI-MS (m/z): 537.1959 $[M + H]^+$.

Methyl N^6-(tert-butoxycarbonyl)-N^2-((2-(naphthalen-1-yl)acetyl)-L-prolyl)-L-lysinate (**20**). To a solution of (2-(naphthalen-1-yl)acetyl)-L-proline (0.1 g, 0.35 mmol) in DMF (3 mL), HBTU (0.17 g, 0.46 mmol) and HOBt (0.06 g, 0.46 mmol) was added in portion at 0 °C. After stirring at r.t. for 1 h, methyl N^6-(tert-butoxycarbonyl)-L-lysinate hydrochloride (0.12 g, 0.46 mmol) and DIPEA (152 μL, 0.88 mmol) were added into the solution. The reaction was stirred overnight at r.t. and quenched by addition of 30 mL brine into the mixture. Extract with DCM (5 mL X 3), and the combined organic layer was dried over Na_2SO4. Filtration, concentrated and the residue was purified by flash column chromatography (CH_2Cl_2 : MeOH = 25:1) and get 0.14g product (76%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.01 (s, 1H), 7.97 – 7.92 (m, 1H), 7.86 (dd, J = 7.7, 1.7 Hz, 1H), 7.78 (d, J = 8.2 Hz, 1H), 7.55 – 7.45 (m, 1H), 7.45 – 7.38 (m, 1H), 7.36 (dd, J = 7.0, 1.2 Hz, 1H), 7.24 (d, J = 7.9 Hz, 1H), 4.65 – 4.59 (m, 1H), 4.50 – 4.42 (m, 1H), 4.23 – 4.07 (m, 2H), 3.70 (s, 3H), 3.68 – 3.64 (m, 1H), 3.57 – 3.48 (m, 1H), 3.11 – 3.03 (m, 1H), 2.39 – 2.29 (m, 1H), 2.20 – 2.10 (m, 1H), 2.01 – 1.87 (m, 2H), 1.80 – 1.68 (m, 1H), 1.59 – 1.49 (m, 1H), 1.39 (s, 9H), 1.34 – 1.28 (m, 2H), 1.27 – 1.13 (m, 2H).

Protein expression and purification. Expression and purification of NTMT1, G9a, SETD7, PRMT1, PRMT3, and *Tb*PRMT7 was performed as previously described [1,2,3–6]

NTMT1 biochemical assays. A fluorescence-based SAHH-coupled assay was applied to study the IC_{50} values for all the compounds. The assay was performed under the following conditions in a final well volume of 40 μL: 5 mM Tris (pH = 7.5), 50 mM KCl, 0.01% Triton X-100, 5 μM SAHH, 0.2 μM NTMT1, 100 μM SAM, and 15 μM ThioGlo1; or 25 mM Tris (pH 7.5), 50 mM KCl, 0.01% Triton X-100, 10μM SAHH, 0.1μM NTMT1, 100 μM SAM, and 10 μM ThioGlo4. The inhibitors were added at concentrations ranging from 0.15 nM to 10 μM. After 10 min incubation, reactions were initiated by the addition of 0.5 μM GPKRIA peptide. Fluorescence was monitored on a BMG ClariOtar microplate reader with excitation 370 nm and emission 500 nm (for Thioglo 1), or excitation 400 nm and emission 465 nm (for Thioglo 4). Data were processed by using GraphPad Prism software 7.0. For the inhibition mechanism studies was performed as previously described.[7]

Selectivity assays. The selectivity studies of G9a, SETD7, PRMT1, and *Tb*PRMT7

were performed as previously described.[8,9]

Cell permeability assay. Cell permeability assay was performed as previously described.[7] The cell lysate was then analyzed with TLC-MS (Advion) to identify the presence of the compound inside the cell.

Cellular N-methylation level. HCT116 cells were seeded 20,000 cells/well on 24-well plates in the presence of 1x PBS or DC541 at different concentrations for 3 days. Then cells were lysed in 1x RIPA cell lysis buffer (25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, protease inhibitors) and incubated for 30 min on ice. The cell lysates were centrifuged at 15,000 rpm for 10 min and the precipitates were removed. The concentration of total protein was quantified by bicinchoninic acid (BCA) protein assay kit (ThermoFisher, #23228). Equal amounts of total protein were mixed with 4x loading dye and loaded onto a 12.5% SDS–PAGE gel and separated. The gel was transferred to a polyvinylidene difluoride (PVDF) membrane using the BioRad Trans-Blot Turbo system. The membrane was then blocked for 1 h in 5% milk TBST solution and washed with 1x TBST solution three times. The membrane was incubated with the anti-me3-SPK antibody at 4 °C for 12 h, washed with 1x TBST solution three times, and then incubated with Rabbit IgG-HRP antibody (Cell Signaling, #7074S) for 1 h at room temperature. The membrane was again washed with 1x TBST solution three times and detected using a Protein Simple FluorChem imaging system. The data were analyzed in GraphPad Prism software 7.0.

Co-crystallization and structure determination. The co-crystallization and structure determination of DC113 with NTMT1 was performed as previously described.[7] Purified NTMT1 (30 mg/mL) was mixed with DC113 at a 1:2 molar ratio, and the mixture was incubated for 1 h at 4°C. Using the hanging-drop vapor diffusion method and incubation at 20°C, the crystals of **DC113** were grown in 2.5 M ammonium sulfate, 0.1 M MES (pH = 8.5). The crystals were then harvested directly from the drop and flash frozen in liquid nitrogen. Data were collected on a single crystal at 12.0 keV at the GM/CA ID-B beamline at the Advanced Photon Source, Argonne National Laboratory. The data were processed and the structure solved as previously described.[7]

Cell viability assay. HCT116 and HT29 cells were seeded as 5,000 cells/well on 96-well plates in the presence of 1x PBS or **DC541** at different concentrations for 24, 48 or 72 hours. Cell viability was assessed using 0.2 mg/ml resazurin solutions prepared from resazurin sodium salt (Acros Organics™, AC418900050) dissolved in sterile 1xPBS. Then, the cells were incubated with 10 µl resazurin solution (10% of cell culture volume) for four hours at 37 °C. The fluorescence was measured using a CLARIOstar microplate reader (Ex = 540 nm, Em = 620 nm) at 37 °C. Cell viability was calculated as 100% × (fluorescence of treated cells – fluorescence of background controls) / (fluorescence of 1x PBS controls – fluorescence of background controls).

MS characterization of compounds 1-22.

Compound **1** (**DC113**). MALDI-MS (positive) m/z: calcd for $C_{29}H_{43}N_8O_4$ $[M + H]^+$ m/z 567.3402, found m/z 567.2305.

Compound **2**. MALDI-MS (positive) m/z: calcd for $C_{28}H_{42}N_9O_4$ $[M + H]^+$ m/z 568.3354, found m/z 568.4898.

Compound **3**. MALDI-MS (positive) m/z: calcd for $C_{28}H_{42}N_9O_4$ $[M + H]^+$ m/z 568.3354, found m/z 568.5264.

Compound **4**. MALDI-MS (positive) m/z: calcd for $C_{29}H_{42}N_8O_4Br$ $[M + H]^+$ m/z 645.2507, found m/z 645.3284.

Compound **5**. MALDI-MS (positive) m/z: calcd for $C_{29}H_{42}N_8O_4F$ $[M + H]^+$ m/z 585.3308, found m/z 585.4217.

Compound **6**. MALDI-MS (positive) m/z: calcd for $C_{29}H_{43}N_6O_4$ $[M + H]^+$ m/z 539.3340, found m/z 539.2995.

Compound **7**. MALDI-MS (positive) m/z: calcd for $C_{29}H_{38}N_7O_4$ $[M + H]^+$ m/z 548.2980, found m/z 548.4167.

Compound **8**. MALDI-MS (positive) m/z: calcd for $C_{25}H_{34}N_5O_4$ $[M + H]^+$ m/z 468.2605, found m/z 468.2419.

Compound **9**. MALDI-MS (positive) m/z: calcd for $C_{26}H_{36}N_5O_4$ $[M + H]^+$ m/z 482.2762, found m/z 482.3716.

Compound **10**. MALDI-MS (positive) m/z: calcd for $C_{28}H_{38}N_5O_6$ $[M + H]^+$ m/z 540.2817, found m/z 540.3713.

Compound **11**. MALDI-MS (positive) m/z: calcd for $C_{34}H_{41}N_6O_4$ $[M + H]^+$ m/z 597.3184, found m/z 597.4349.

Compound **12**. MALDI-MS (positive) m/z: calcd for $C_{32}H_{40}N_5O_5$ $[M + H]^+$ m/z 574.3024, found m/z 574.3626.

Compound **13**. MALDI-MS (positive) m/z: calcd for $C_{32}H_{40}N_5O_4$ $[M + H]^+$ m/z 558.3075, found m/z 558.4128.

Compound **14**. MALDI-MS (positive) m/z: calcd for $C_{23}H_{31}N_4O_3$ $[M + H]^+$ m/z 411.2391, found m/z 411.2515.

Compound **15**. MALDI-MS (positive) m/z: calcd for $C_{30}H_{45}N_8O_4$ $[M + H]^+$ m/z 581.3558, found m/z 581.2948.

Compound **16**. MALDI-MS (positive) m/z: calcd for $C_{30}H_{44}N_6O_4Na$ $[M + Na]^+$ m/z 575.3316, found m/z 575.2672.

Compound **17**. MALDI-MS (positive) m/z: calcd for $C_{27}H_{37}N_5O_4Na$ $[M + Na]^+$ m/z 518.2738, found m/z 518.1962.

Compound **18**. ESI-MS (positive) m/z: calcd for $C_{31}H_{37}N_6O_3$ $[M + H]^+$ m/z 541.2922, found m/z 541.3635.

Compound **19**. MALDI-MS (positive) m/z: calcd for $C_{30}H_{36}N_5O_4$ $[M + H]^+$ m/z 530.2762, found m/z 530.3354.

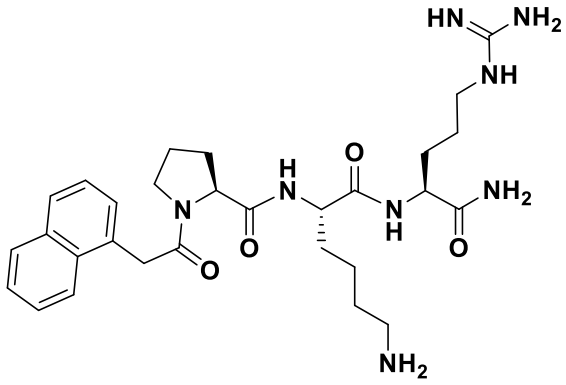
Compound **20** (**DC541**). MALDI-MS (positive) m/z: calcd for $C_{30}H_{35}N_5O_4Na$ $[M + Na]^+$ m/z 552.2581, found m/z 552.1934.

Compound **21**. ESI-MS (positive) m/z: calcd for $C_{29}H_{35}N_4O_3$ $[M + H]^+$ m/z 487.2704, found m/z 487.2479.

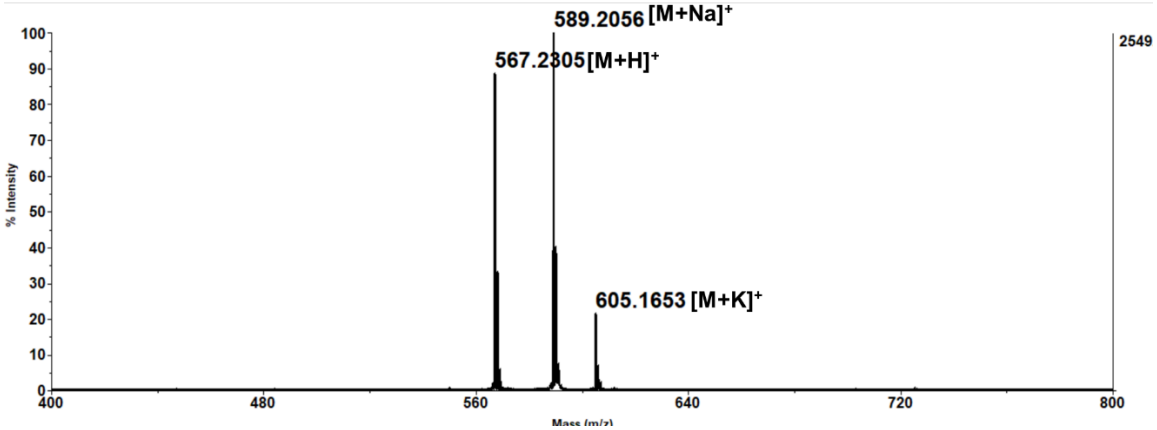
Compound **22**. ESI-MS (positive) m/z: calcd for $C_{33}H_{37}N_4O_3$ $[M + H]^+$ m/z 537.2860,

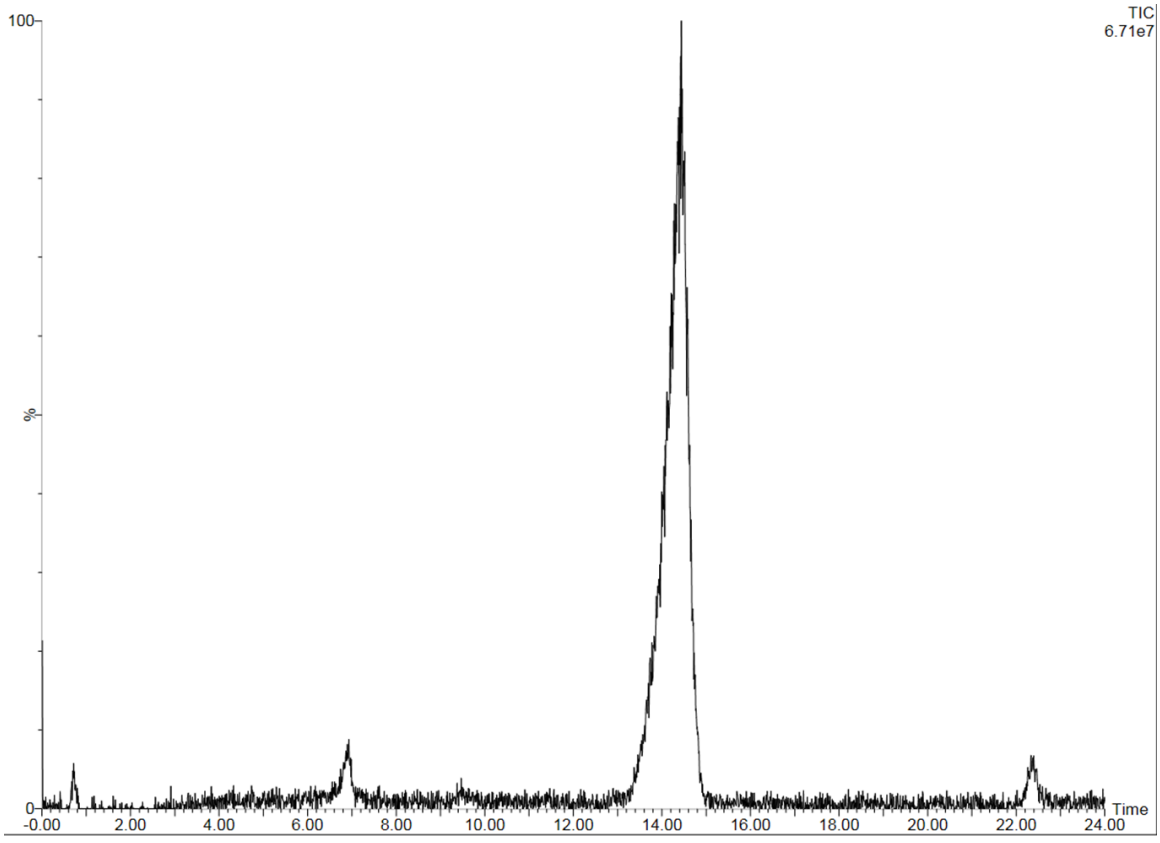
found m/z 537.1959.

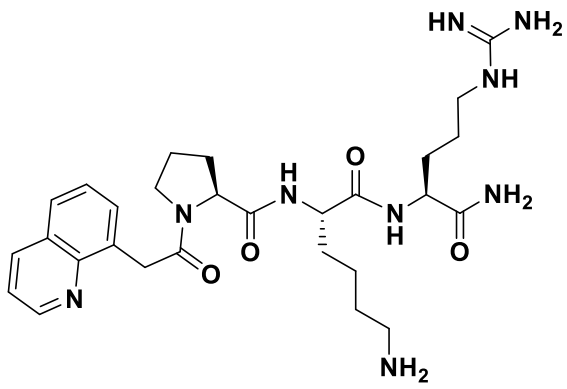
MS and HPLC analysis of compounds **1-22**



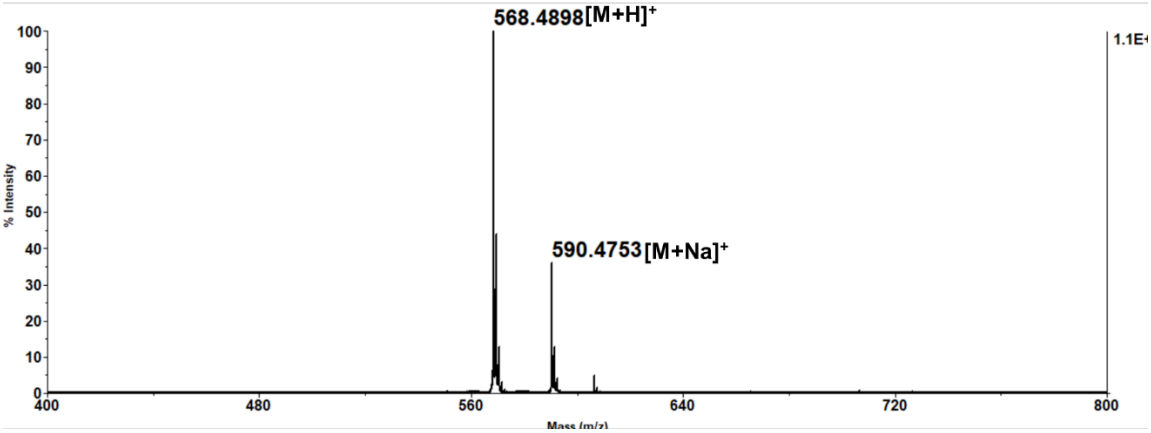
(1)

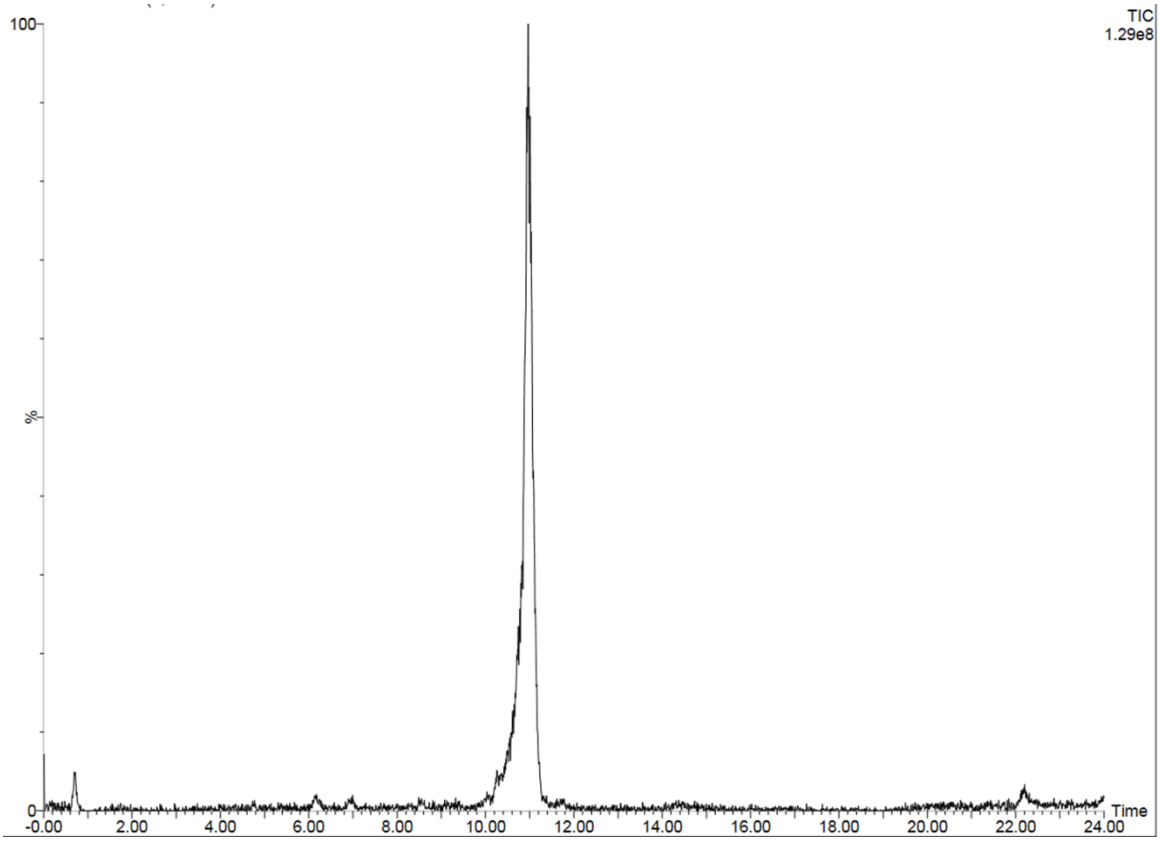


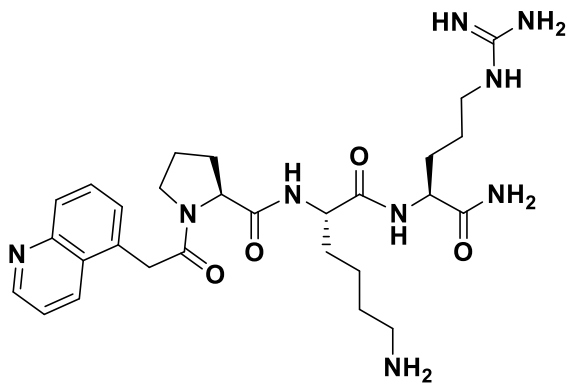



HN
NH2
NH
O
H
N
N
O
N
H
NH2
O
O
N
NH2

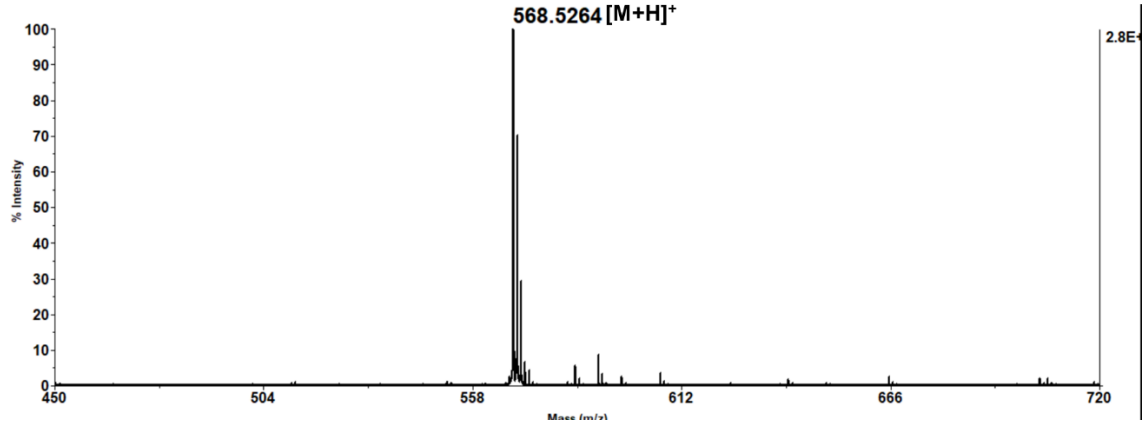
(2)

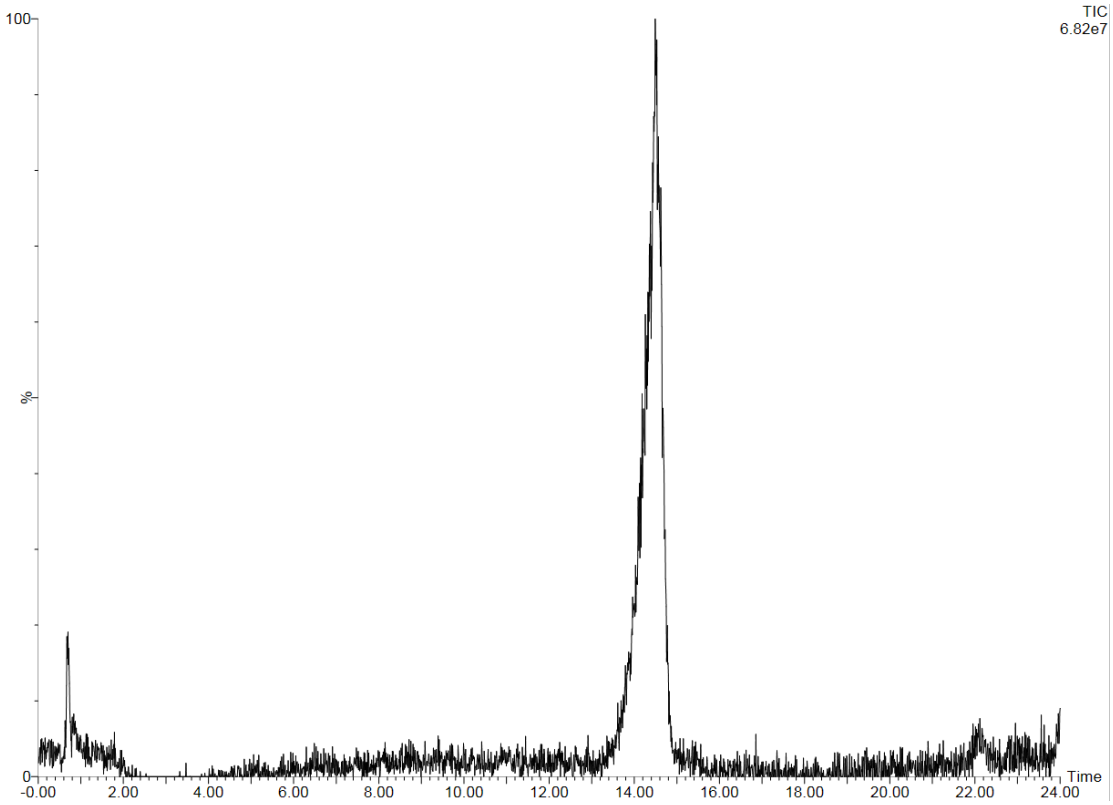

568.4898[M+H]+
590.4753[M+Na]+
% Intensity
100
90
80
70
60
50
40
30
20
10
0
400
480
560
640
720
800
1.1E+

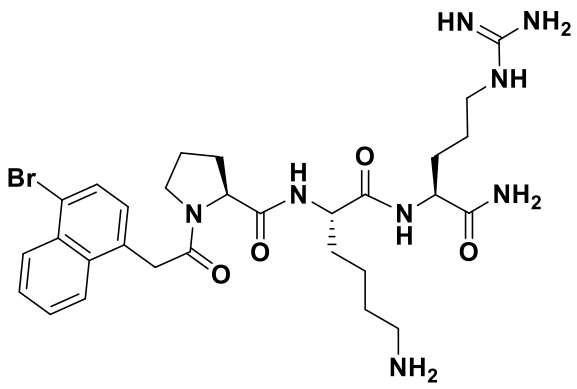

TIC
1.29e8
100
%
0
-0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
Time


HN
NH2
NH
O
H
N
N
O
O
N
H
NH2
O
N
O
NH2

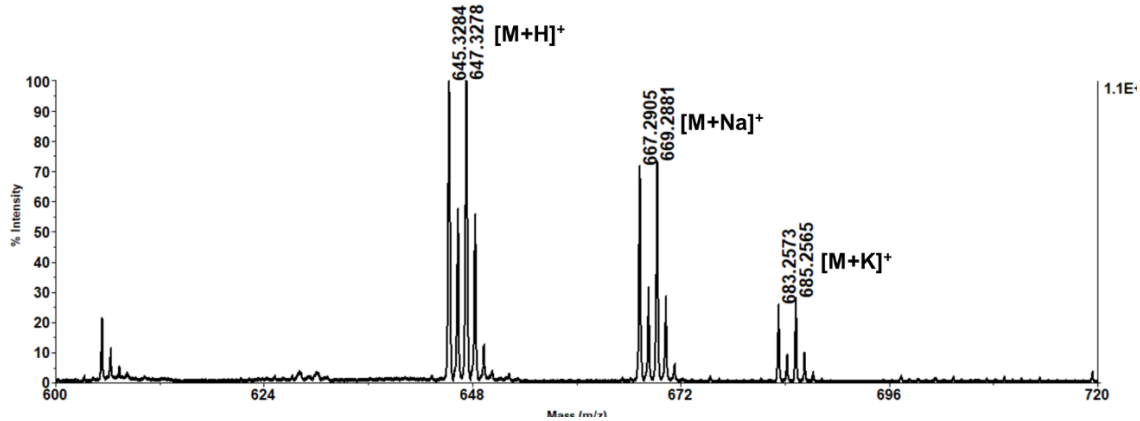
(3)

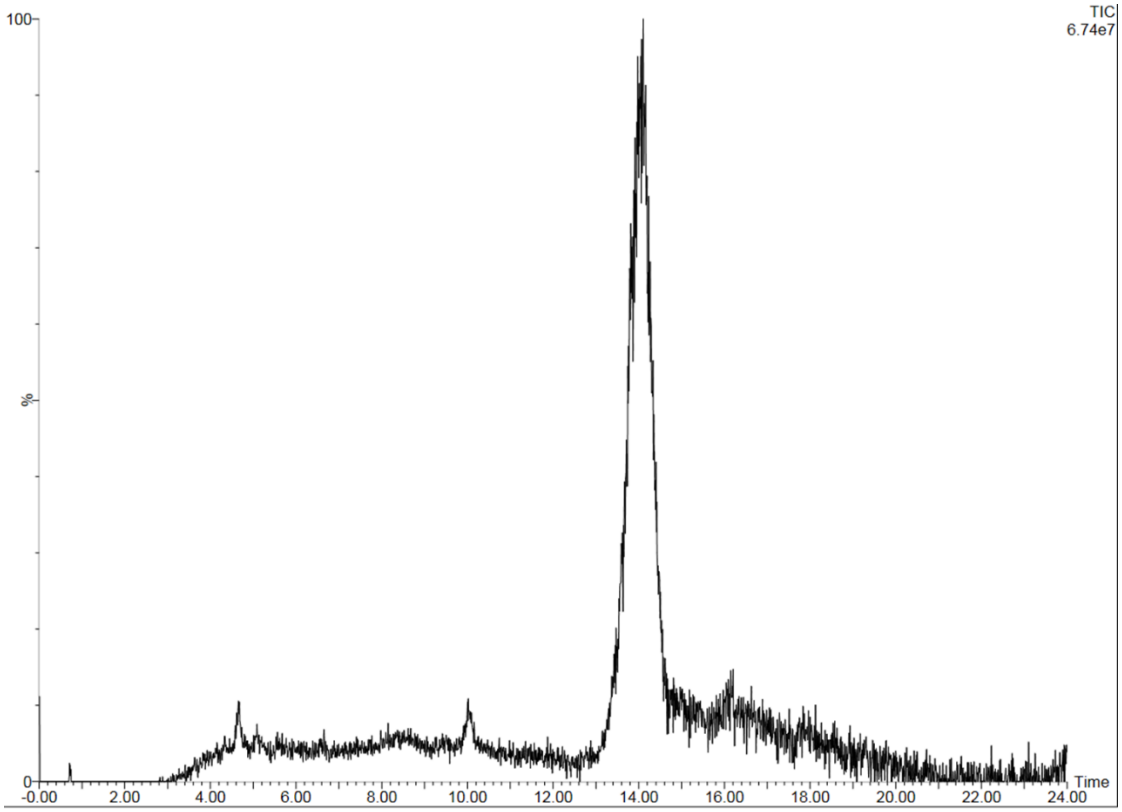

568.5264 [M+H]+
2.8E
% Intensity
100
90
80
70
60
50
40
30
20
10
0
450
504
558
612
666
720

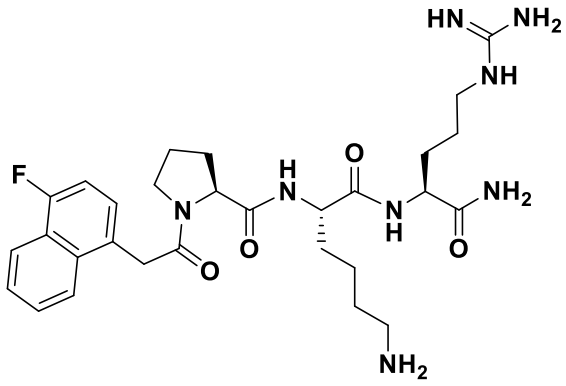

TIC
6.82e7
100
%
0
-0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
Time


HN
NH2
NH
Br
H
N
O
N
N
H
NH2
O
O
O
NH2

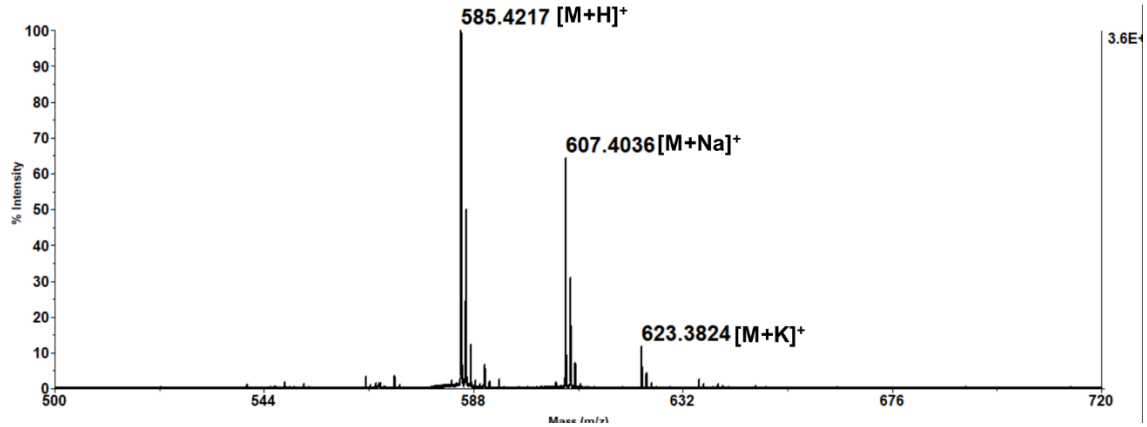
(4)

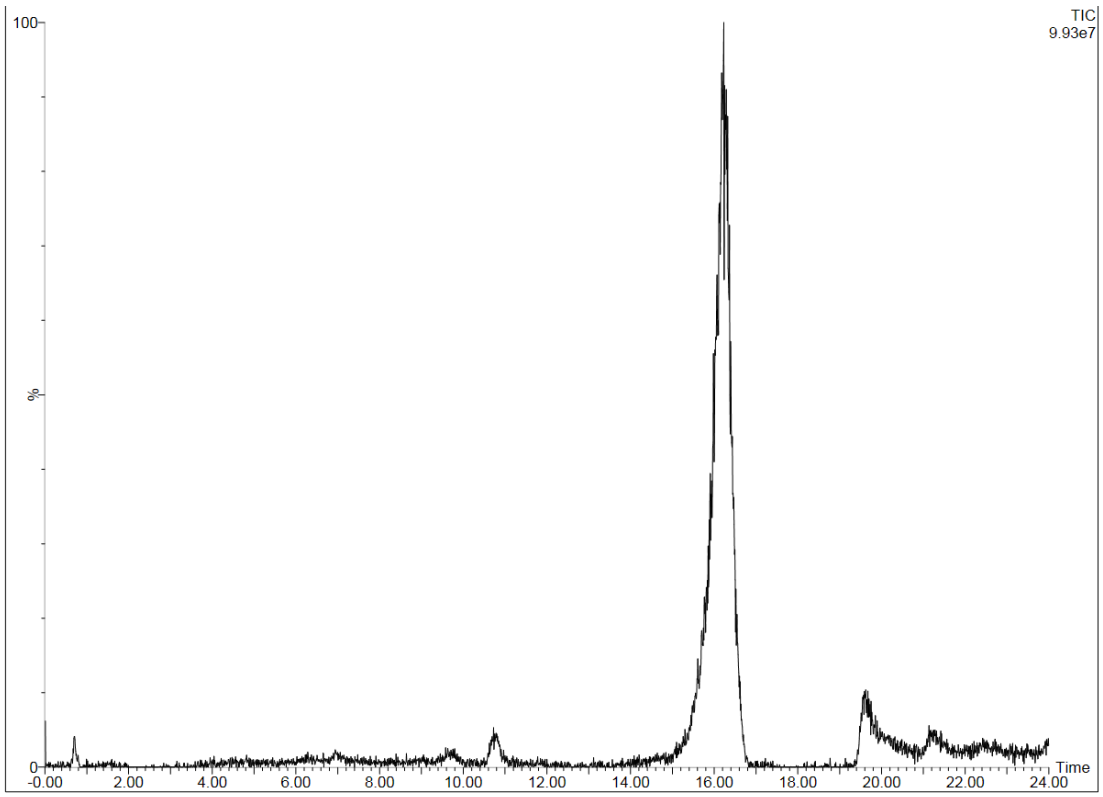

645.3284
647.3278
[M+H]+
667.2905
669.2881
[M+Na]+
683.2573
685.2565
[M+K]+
% Intensity
100
90
80
70
60
50
40
30
20
10
0
600
624
648
672
696
720
1.1E

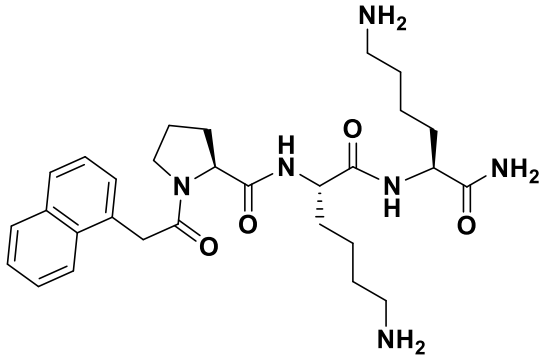

TIC
6.74e7
100
%
0
-0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
Time


HN
NH2
NH
F
N
H
N
O
N
H
NH2
O
O
O
NH2

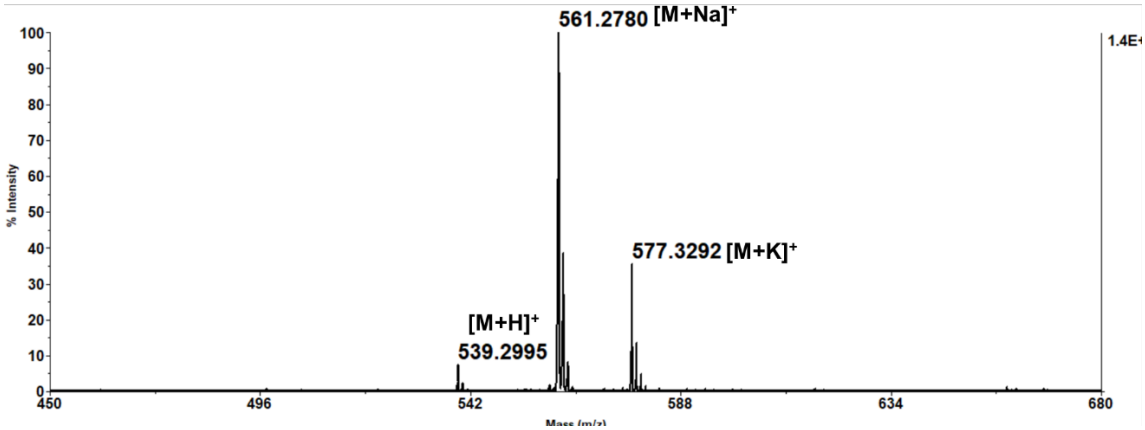
(5)

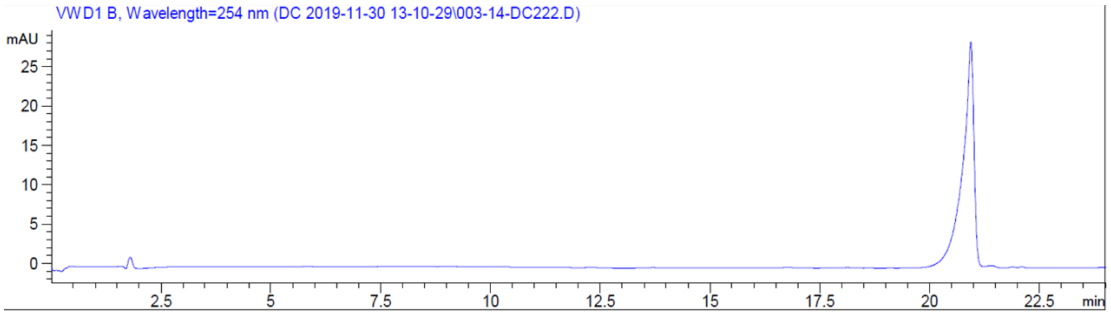

585.4217 [M+H]+
607.4036 [M+Na]+
623.3824 [M+K]+
% Intensity
100
90
80
70
60
50
40
30
20
10
0
500
544
588
632
676
720
3.6E+

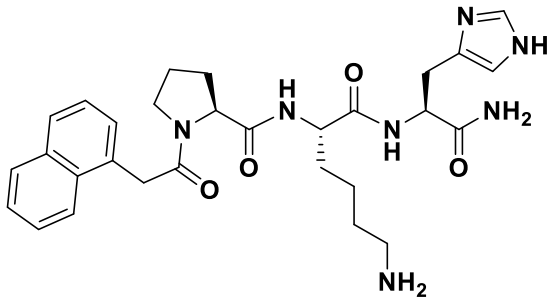

TIC
9.93e7
100
%
0
-0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
Time


NH2
O
H
N
N
N
H
NH2
O
O
O
NH2

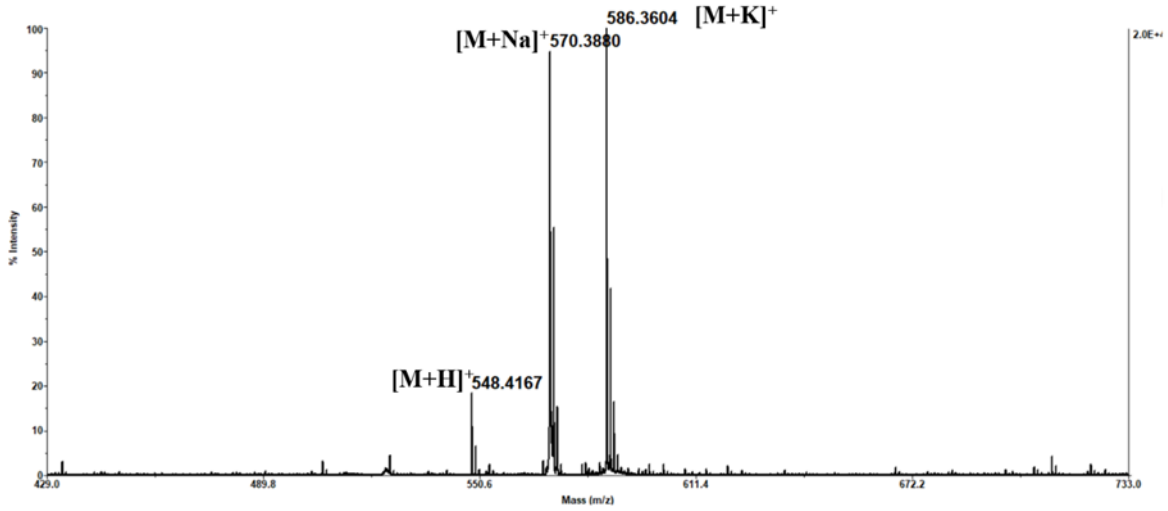
(6)

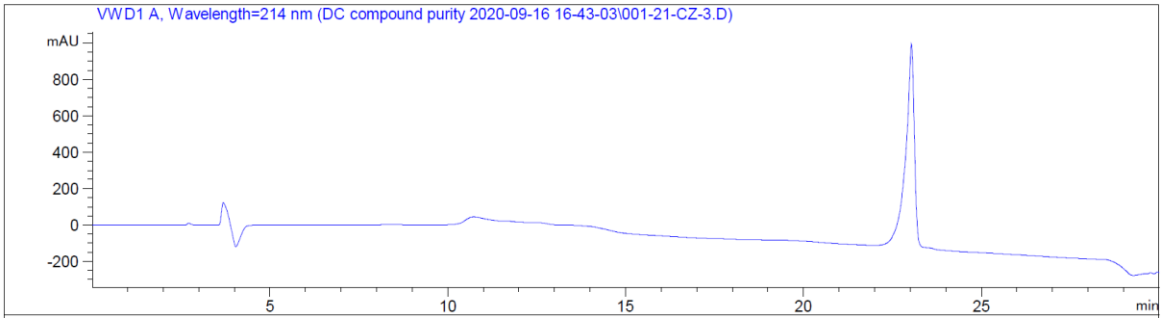

561.2780 [M+Na]+
577.3292 [M+K]+
[M+H]+
539.2995
% Intensity
100
90
80
70
60
50
40
30
20
10
0
450
496
542
588
634
680

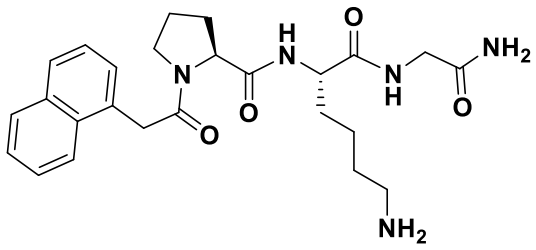

VWD1 B, Wavelength=254 nm (DC 2019-11-30 13-10-29\003-14-DC222.D)
mAU
25
20
15
10
5
0
2.5
5
7.5
10
12.5
15
17.5
20
22.5
min



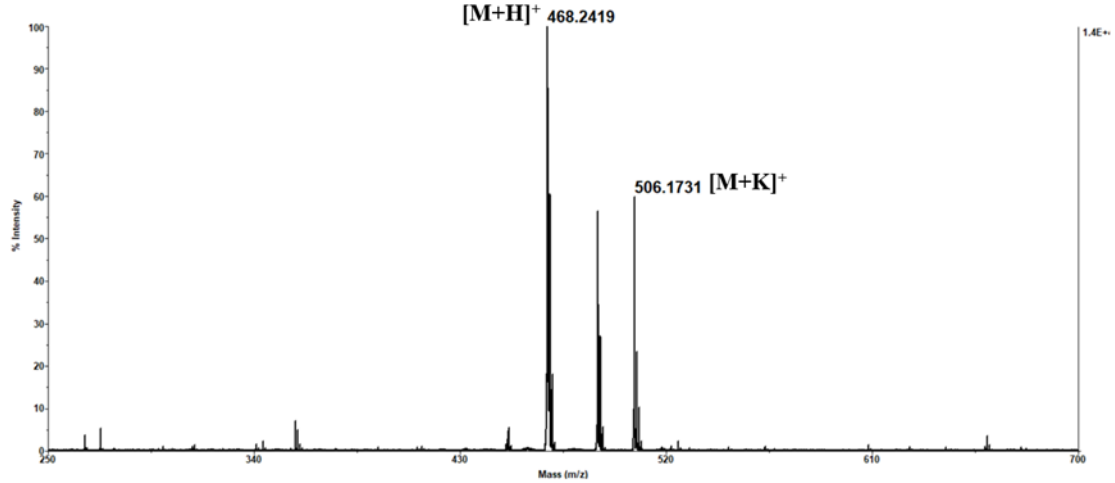
(7)

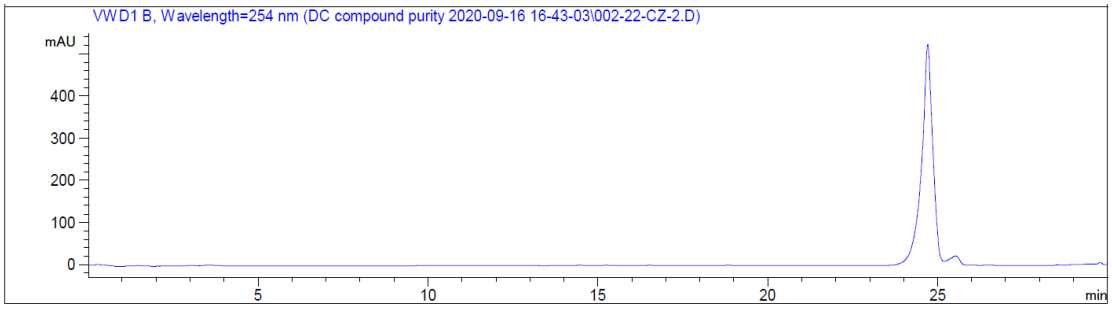

586.3604
[M+K]+
[M+Na]+570.3880
[M+H]+548.4167
% Intensity
Mass (m/z)

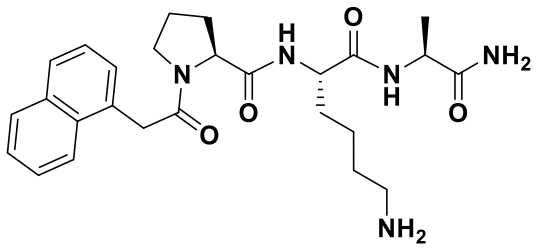

VWD1 A, Wavelength=214 nm (DC compound purity 2020-09-16 16-43-03\001-21-CZ-3.D)
mAU
min


O
H
N
NH2
N
N
H
O
O
O
NH2

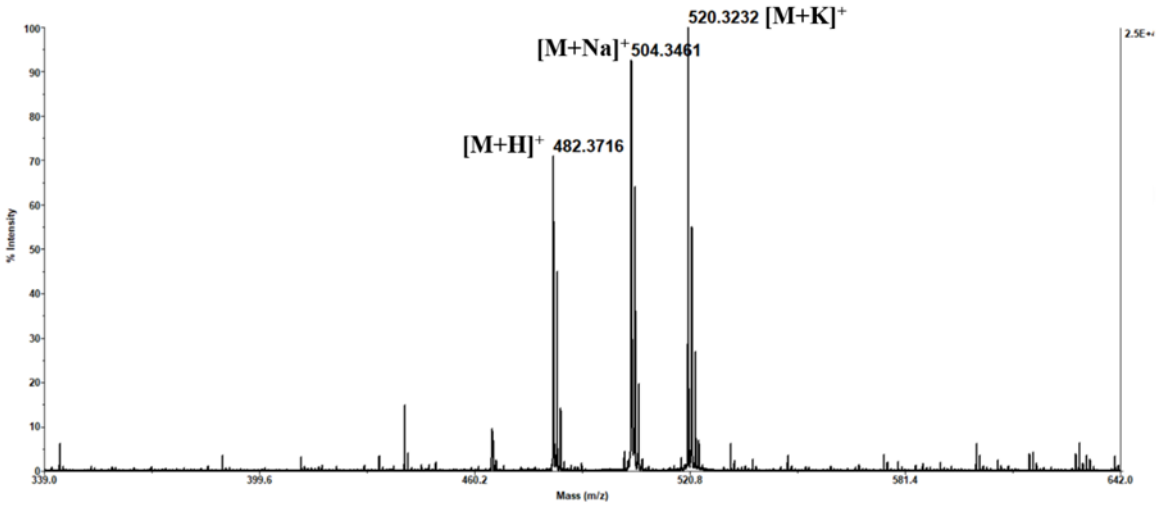
(8)

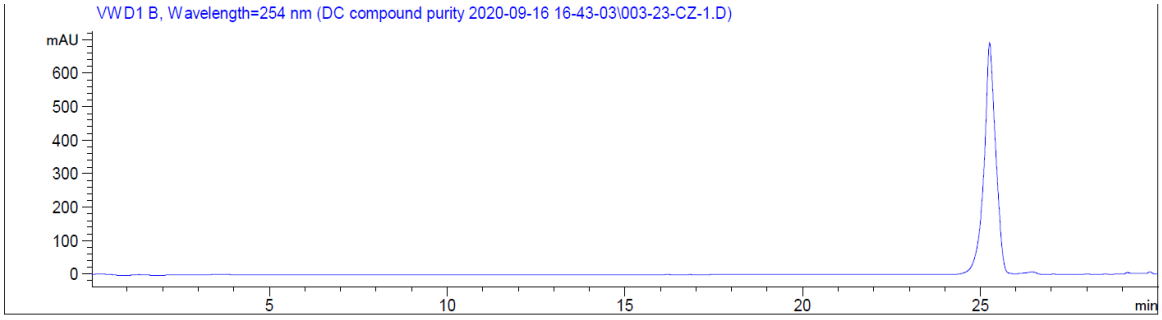

[M+H]+ 468.2419
506.1731 [M+K]+
% Intensity
Mass (m/z)
250
340
430
520
610
700

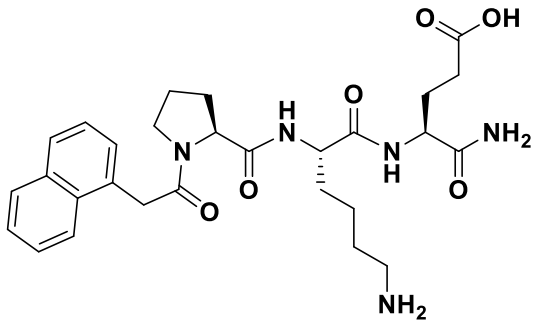

VWD1 B, Wavelength=254 nm (DC compound purity 2020-09-16 16-43-03\002-22-CZ-2.D)
mAU
400
300
200
100
0
5
10
15
20
25
min



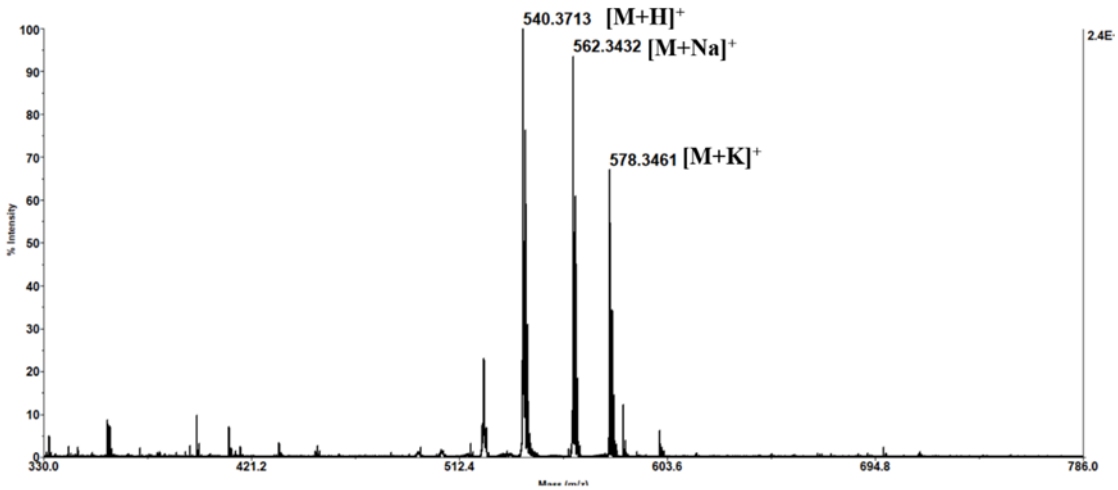
(9)

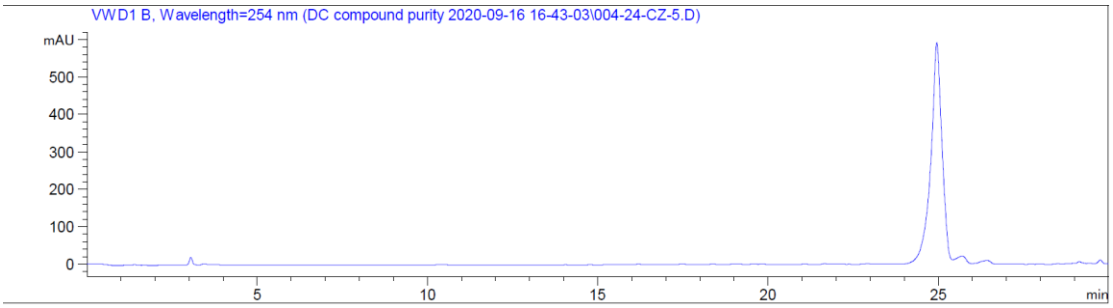

520.3232 [M+K]+
[M+Na]+504.3461
[M+H]+ 482.3716
% Intensity
Mass (m/z)

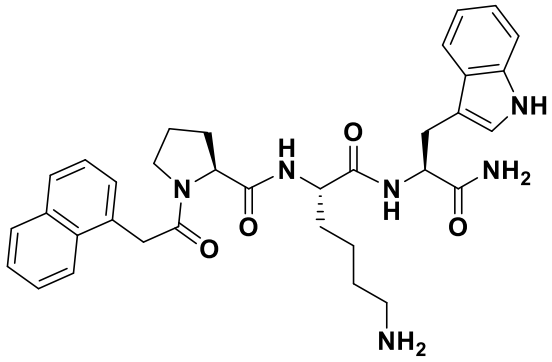

VWD1 B, Wavelength=254 nm (DC compound purity 2020-09-16 16-43-03\003-23-CZ-1.D)
mAU
min


O
OH
H
N
O
N
H
NH2
N
O
O
O
NH2

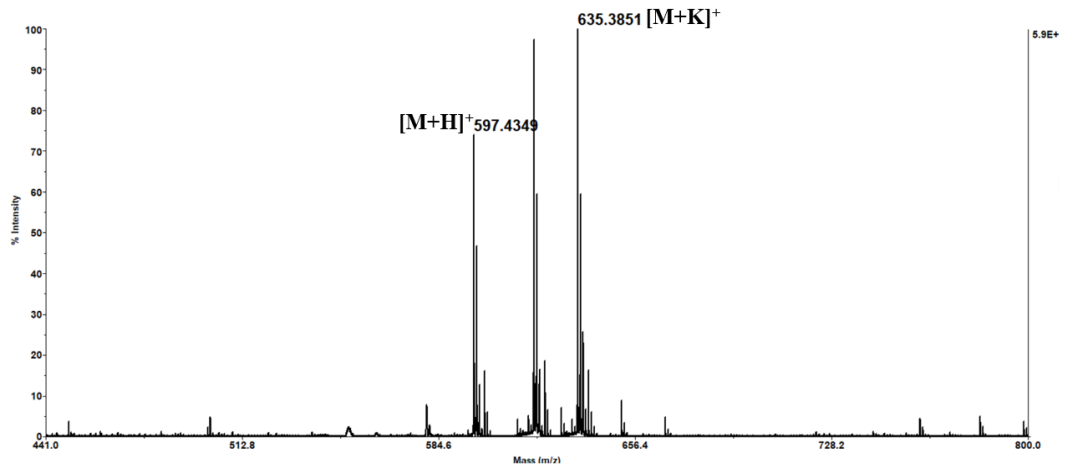
(10)

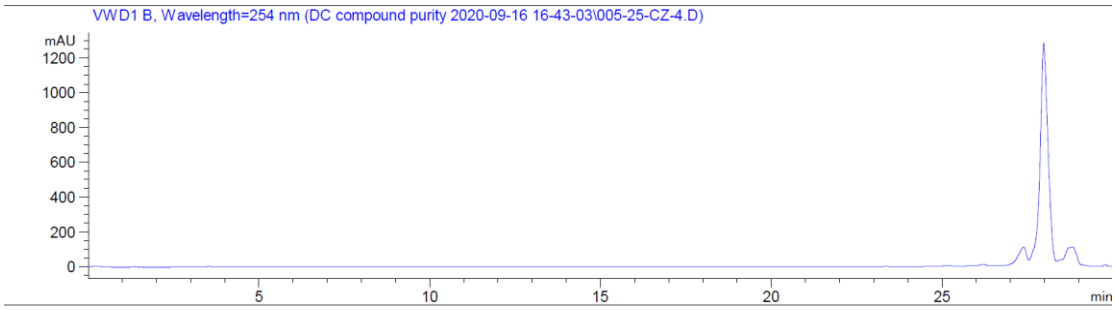

540.3713 [M+H]+
562.3432 [M+Na]+
578.3461 [M+K]+
% Intensity
100
90
80
70
60
50
40
30
20
10
0
330.0
421.2
512.4
603.6
694.8
786.0
2.4E

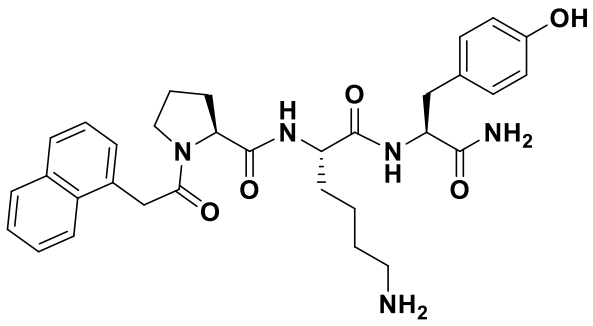

VWD1 B, Wavelength=254 nm (DC compound purity 2020-09-16 16-43-03\004-24-CZ-5.D)
mAU
500
400
300
200
100
0
5
10
15
20
25
min


NH
H
N
O
N
H
NH2
N
O
O
O
NH2

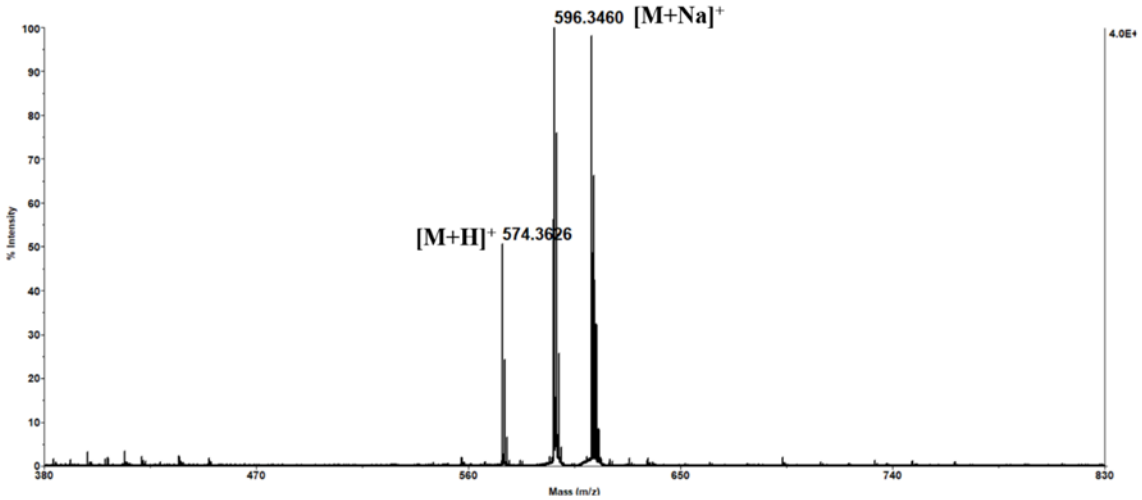
(11)

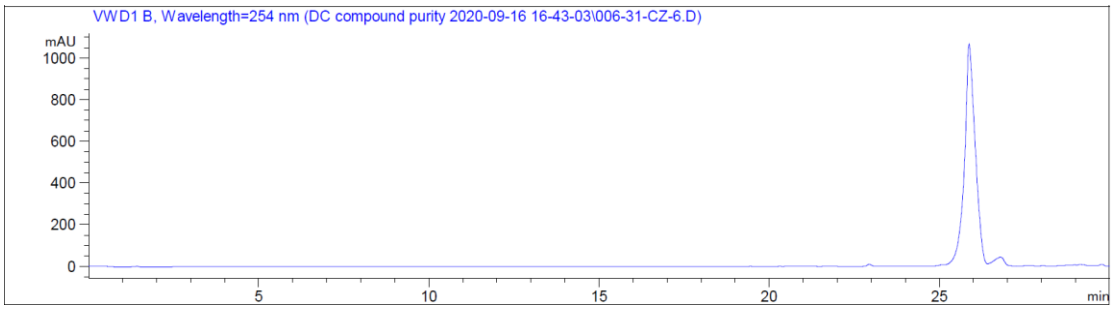

635.3851 [M+K]+
[M+H]+597.4349
5.9E+
% Intensity
Mass (m/z)
441.0
512.8
584.6
656.4
728.2
800.0

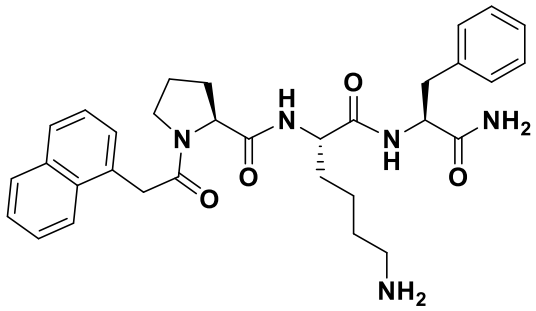

VWD1 B, Wavelength=254 nm (DC compound purity 2020-09-16 16-43-03\005-25-CZ-4.D)
mAU
min


OH
H
N
O
N
H
NH2
N
O
O
NH2

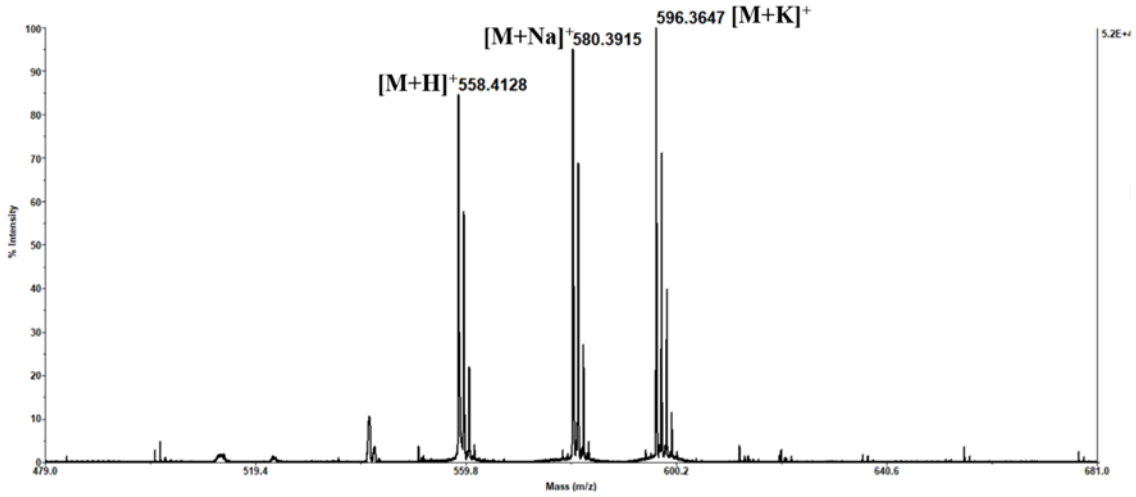
(12)

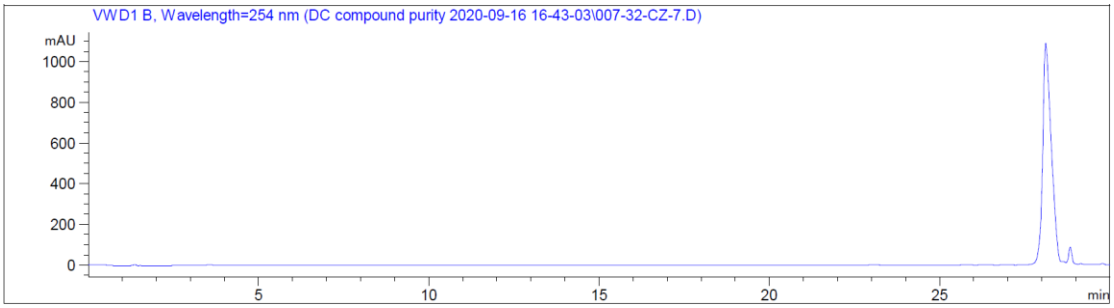

596.3460 [M+Na]+
[M+H]+ 574.3626
% Intensity
Mass (m/z)
4.0E+

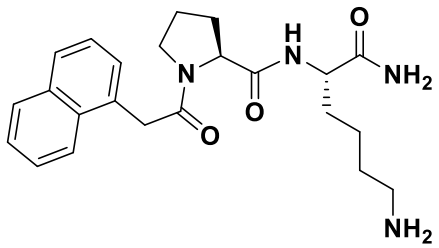

VWD1 B, Wavelength=254 nm (DC compound purity 2020-09-16 16-43-03\006-31-CZ-6.D)
mAU
min


H
N
O
N
H
NH2
O
N
O
NH2

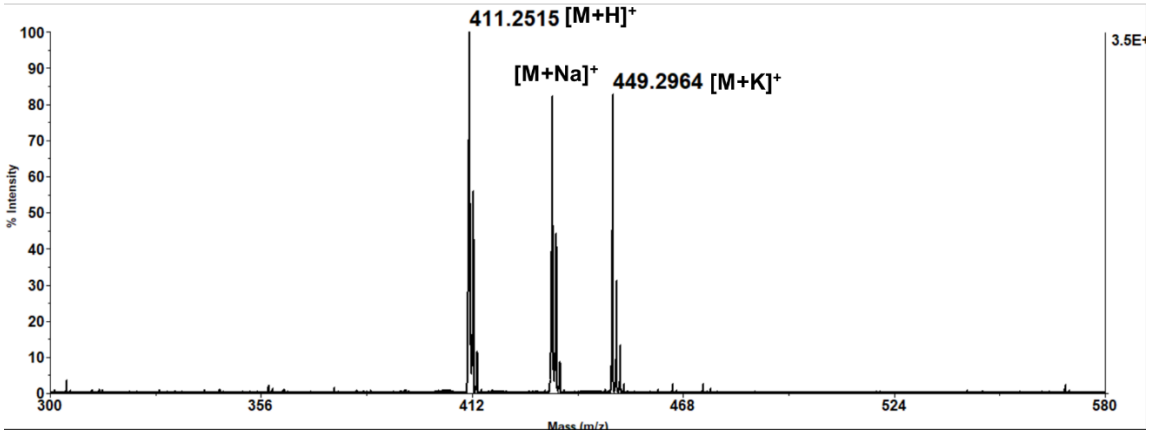
(13)

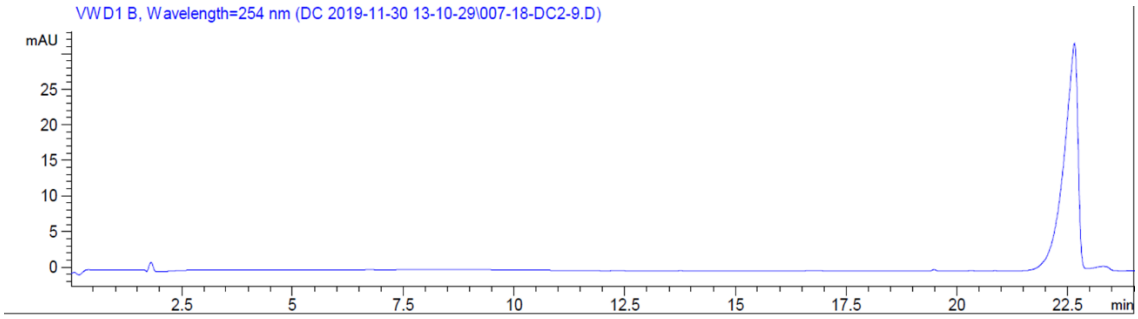

[M+H]+558.4128
[M+Na]+580.3915
596.3647 [M+K]+
% Intensity
Mass (m/z)
479.0
519.4
559.8
600.2
640.6
681.0

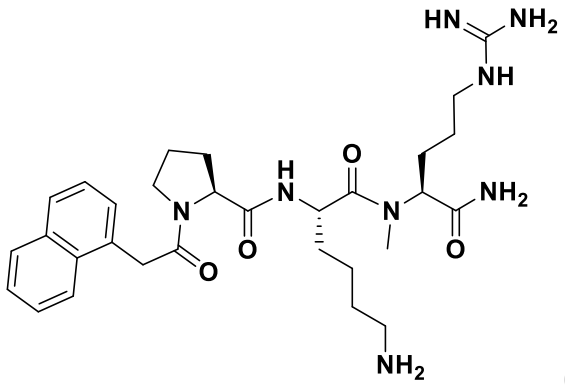

VWD1 B, Wavelength=254 nm (DC compound purity 2020-09-16 16-43-03\007-32-CZ-7.D)
mAU
1000
800
600
400
200
0
5
10
15
20
25
min


O
H
N
N
O
NH2
N
O
NH2

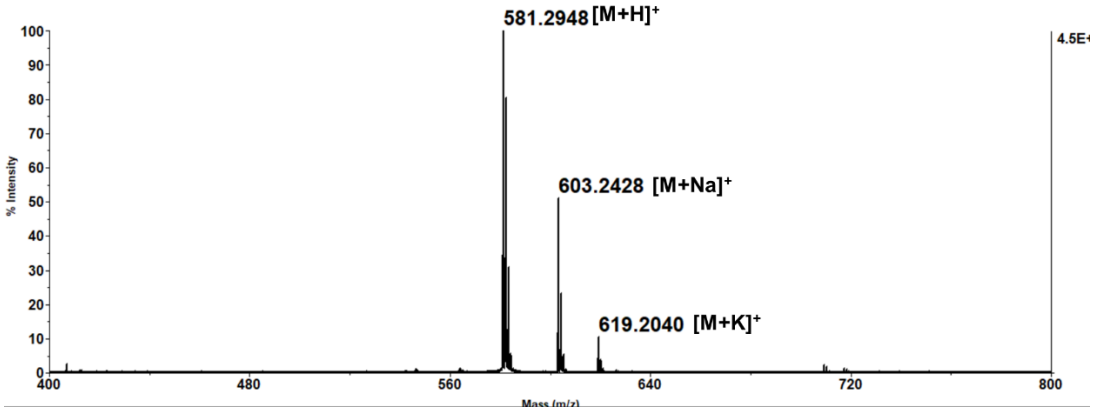
(14)

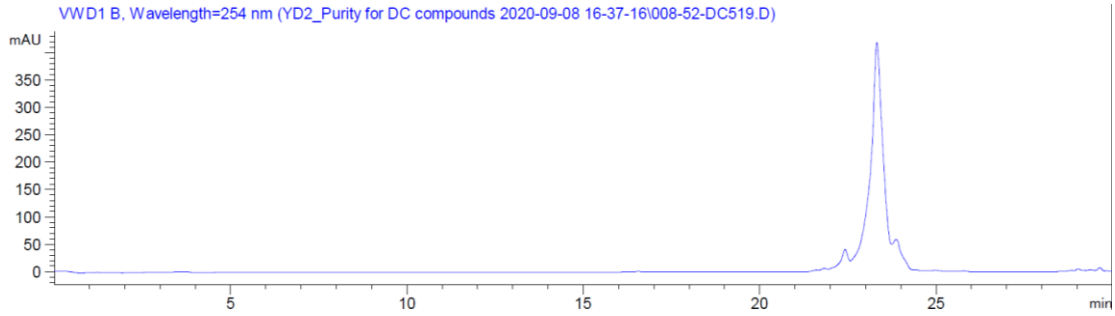

411.2515 [M+H]+
[M+Na]+
449.2964 [M+K]+
% Intensity
100
90
80
70
60
50
40
30
20
10
0
300
356
412
468
524
580
Mass (m/z)

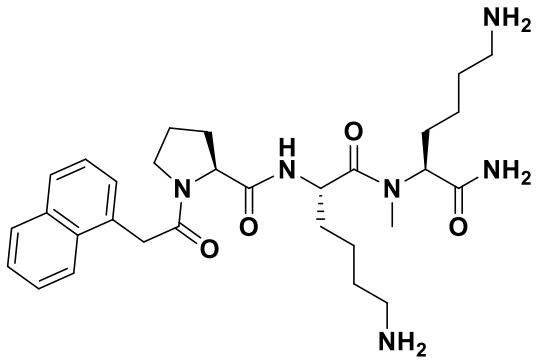

VWD1 B, Wavelength=254 nm (DC 2019-11-30 13-10-29\007-18-DC2-9.D)
mAU
25
20
15
10
5
0
2.5
5
7.5
10
12.5
15
17.5
20
22.5
min


HN
NH2
NH
H
N
O
N
NH2
O
N
O
O
NH2

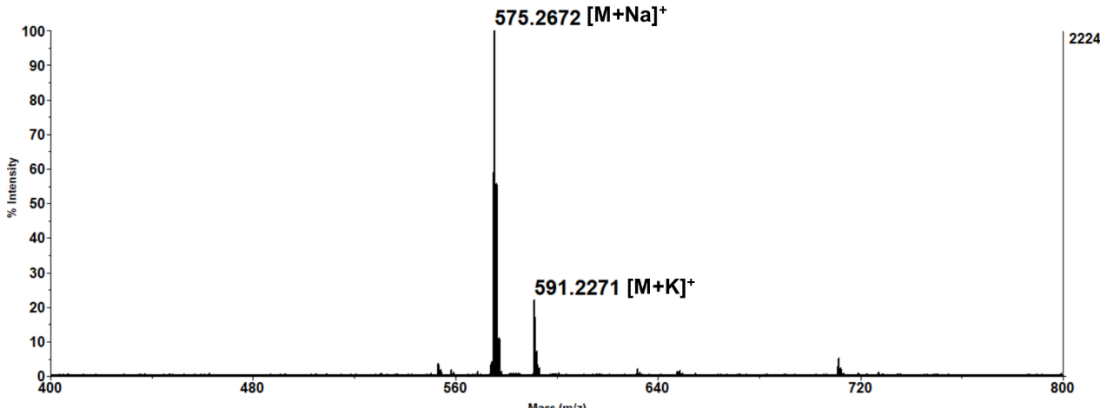
(15)


581.2948 [M+H]+
603.2428 [M+Na]+
619.2040 [M+K]+
% Intensity
100
90
80
70
60
50
40
30
20
10
0
400
480
560
640
720
800
Mass (m/z)

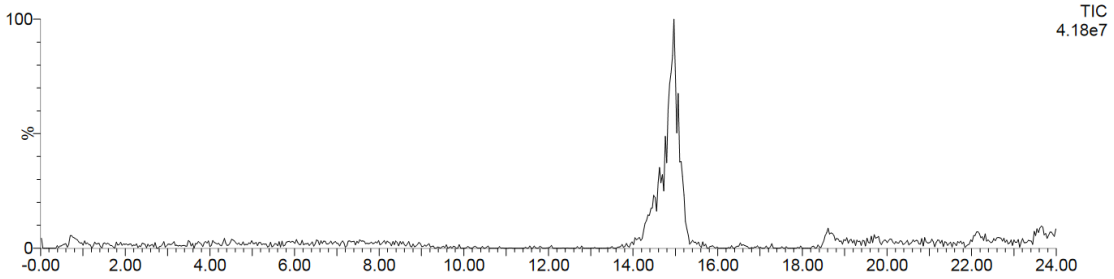

VWD1 B, Wavelength=254 nm (YD2_Purity for DC compounds 2020-09-08 16-37-16\008-52-DC519.D)
mAU
350
300
250
200
150
100
50
0
5
10
15
20
25
min

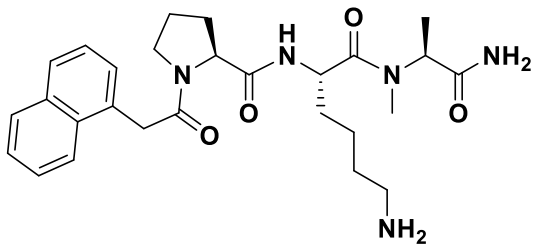

NH2
H
N
O
N
NH2
O
N
O
O
NH2

(16)

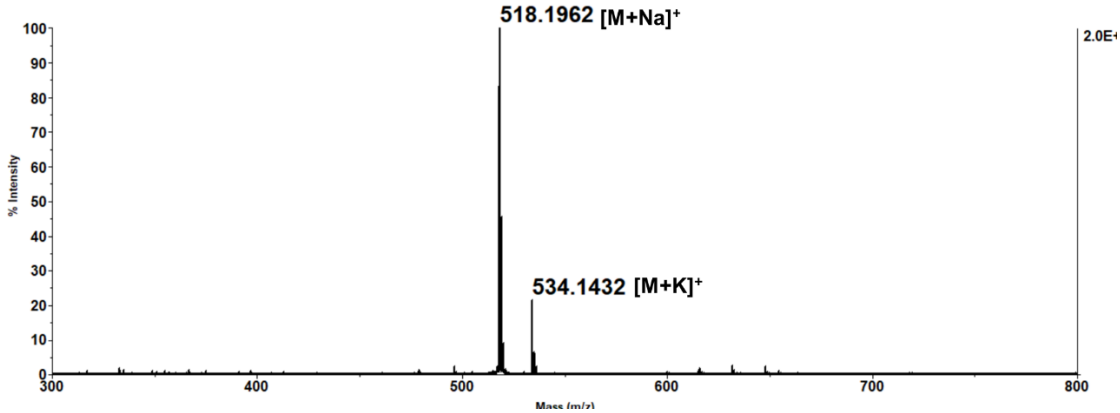

575.2672 [M+Na]+
591.2271 [M+K]+
2224
% Intensity
100
90
80
70
60
50
40
30
20
10
0
400
480
560
640
720
800

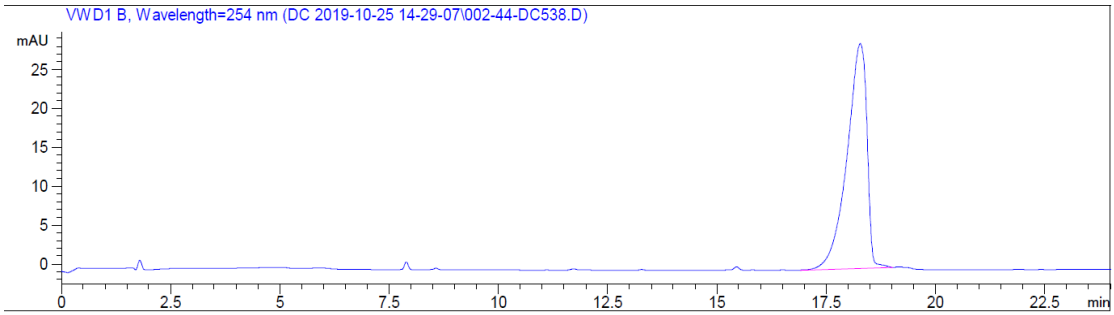
LC-MS

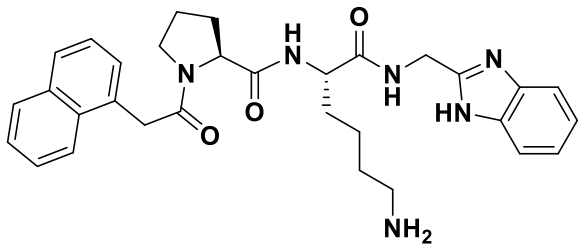

TIC
4.18e7
100
%
0
-0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00


H
N
O
N
NH2
O
N
O
O
NH2

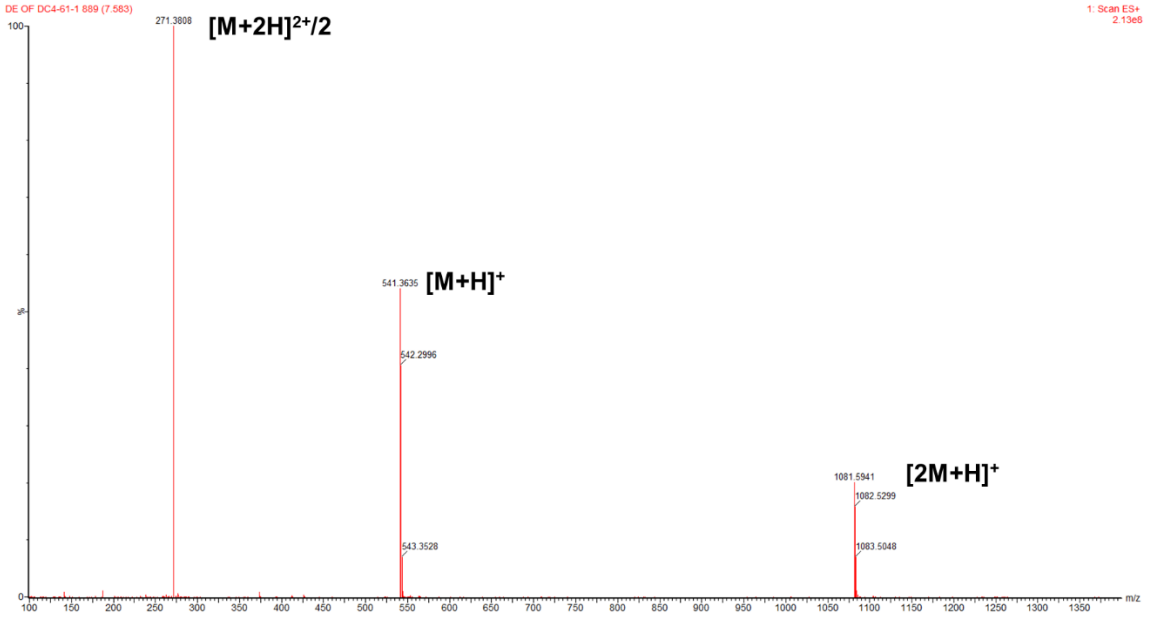
(17)

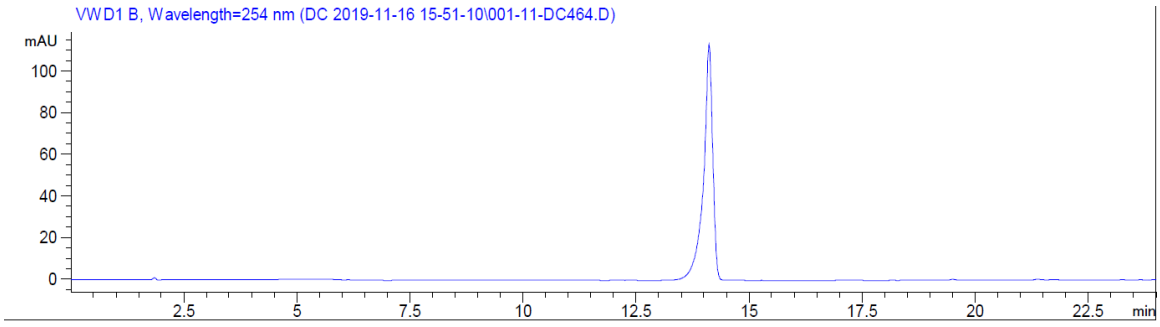

518.1962 [M+Na]+
534.1432 [M+K]+
% Intensity
100
90
80
70
60
50
40
30
20
10
0
300
400
500
600
700
800
Mass (m/z)

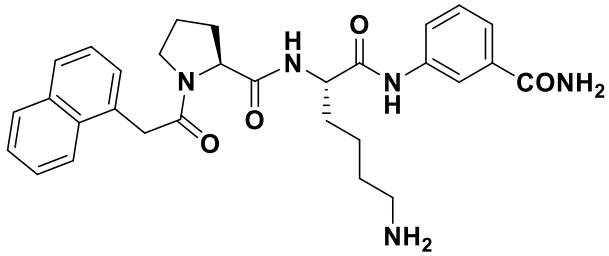

VWD1 B, Wavelength=254 nm (DC 2019-10-25 14-29-07\002-44-DC538.D)
mAU
25
20
15
10
5
0
0
2.5
5
7.5
10
12.5
15
17.5
20
22.5
min


O
H
N
N
O
N
H
N
HN
O
NH2

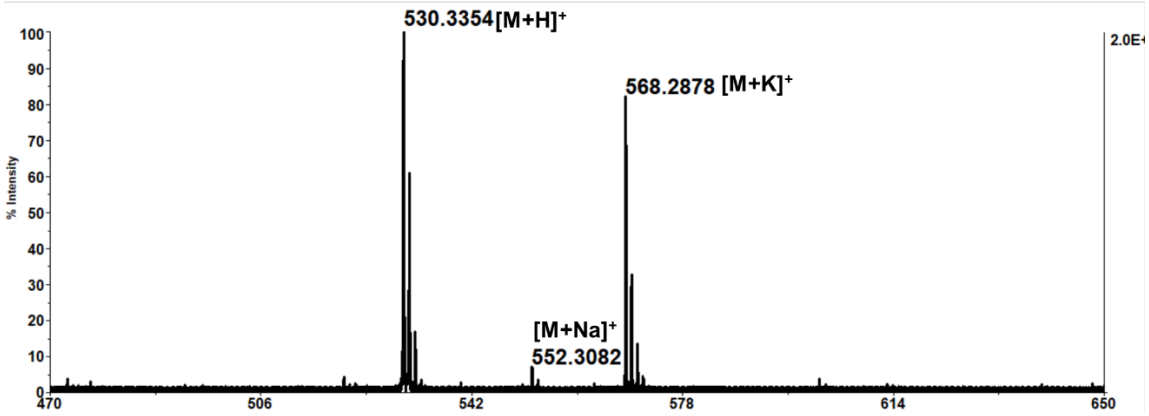
(18)

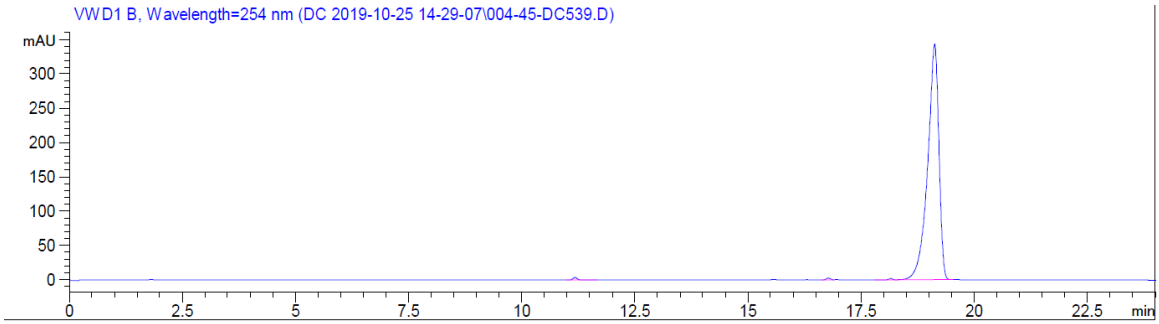

DE OF DC4-61-1 889 (7.583)
1: Scan ES+
2.13e8
271.3808
[M+2H]2+/2
541.3635
[M+H]+
542.2996
1081.5941
[2M+H]+
1082.5299
543.3528
1083.5048
m/z

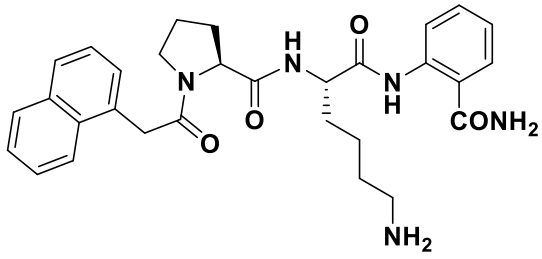

VWD1 B, Wavelength=254 nm (DC 2019-11-16 15-51-10\001-11-DC464.D)
mAU
min



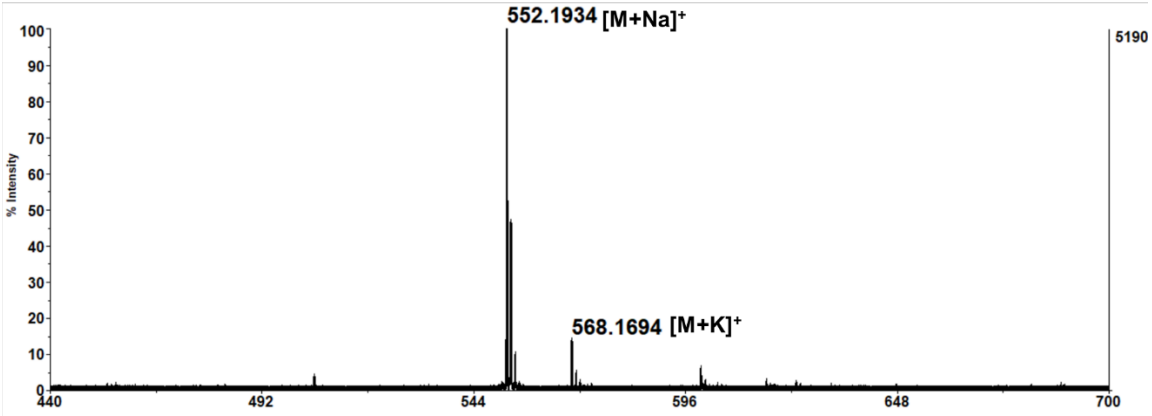
(19)


530.3354[M+H]+
568.2878 [M+K]+
[M+Na]+
552.3082
% Intensity
2.0E+
100
90
80
70
60
50
40
30
20
10
0
470
506
542
578
614
650

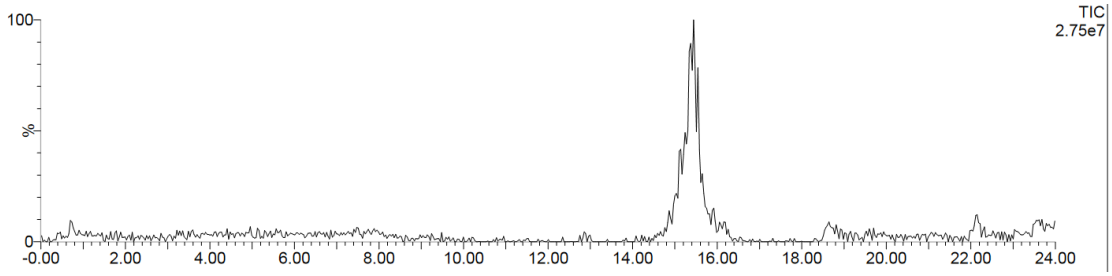

VWD1 B, Wavelength=254 nm (DC 2019-10-25 14-29-07\004-45-DC539.D)
mAU
300
250
200
150
100
50
0
0
2.5
5
7.5
10
12.5
15
17.5
20
22.5
min

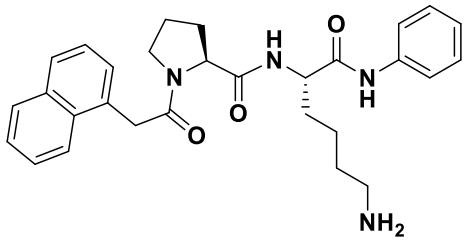

O
H
N
N
O
N
H
CONH2
O
NH2

(20)

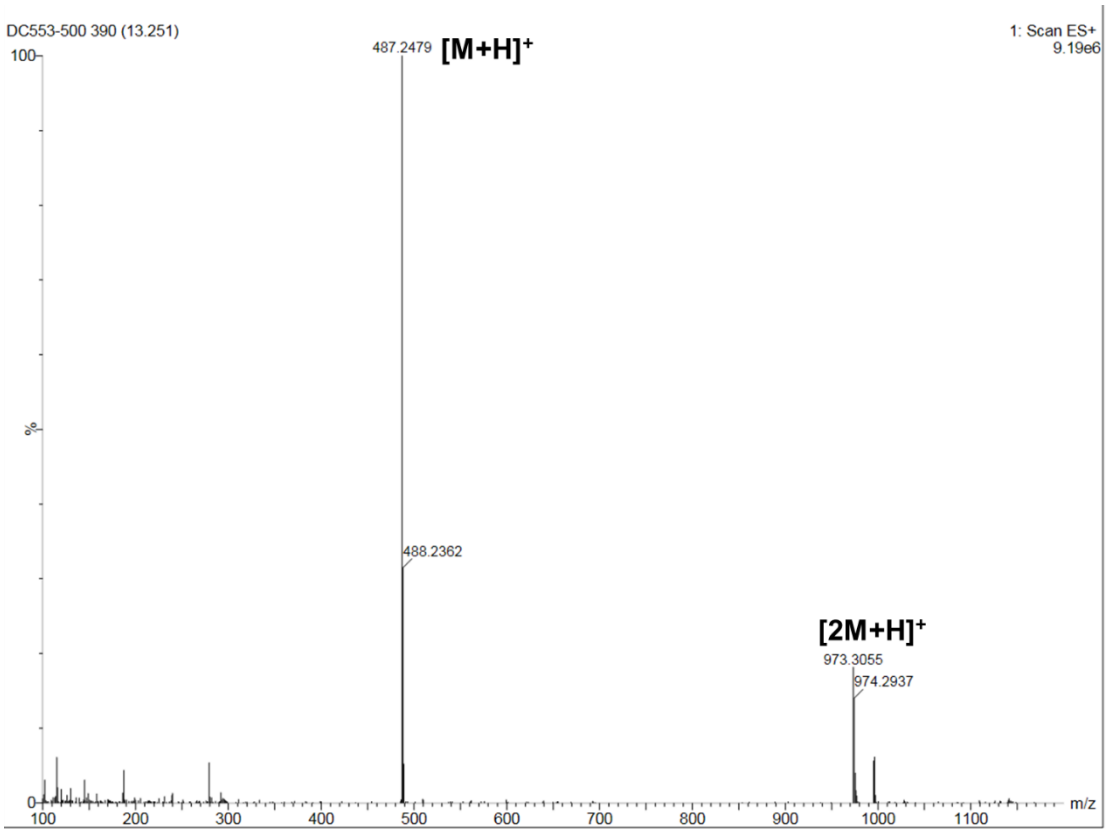

552.1934 [M+Na]+
568.1694 [M+K]+
5190
% Intensity
100
90
80
70
60
50
40
30
20
10
0
440
492
544
596
648
700

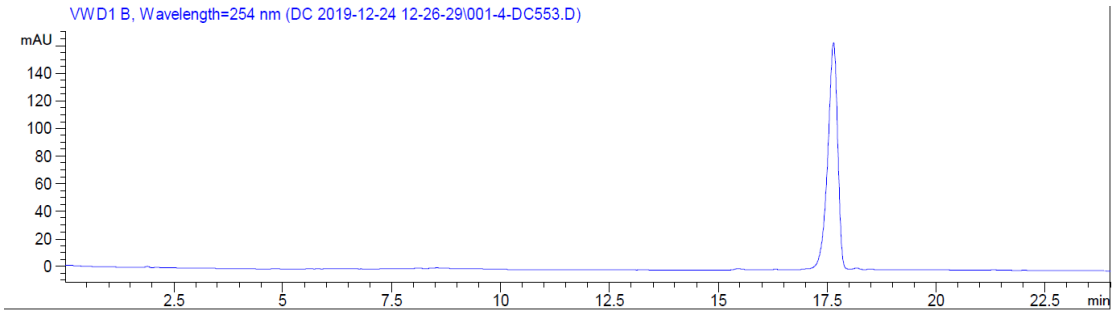
LC-MS

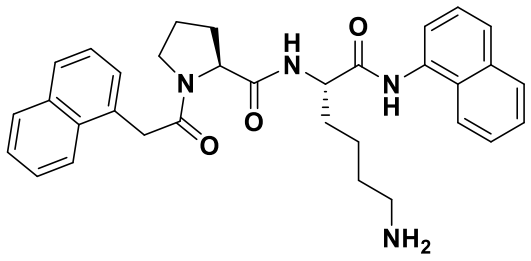

TIC
2.75e7
100
%
0
-0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00


O
H
N
N
O
N
H
O
NH2

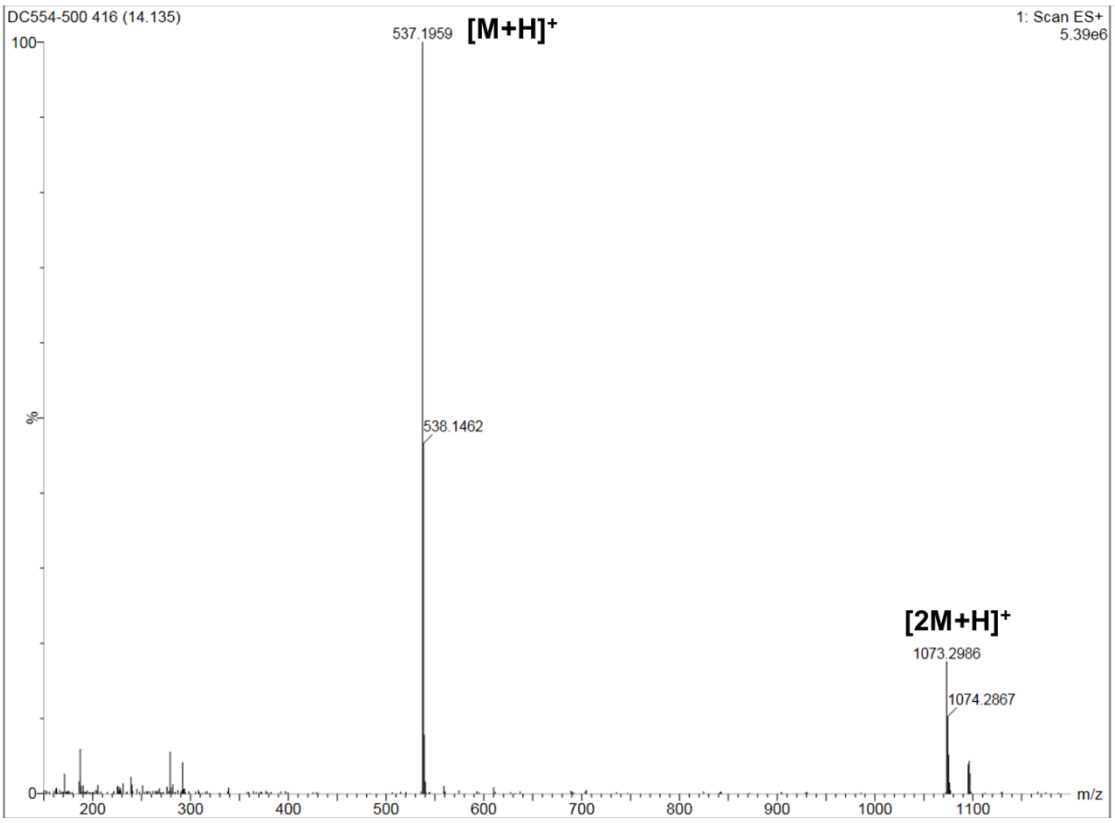
(21)

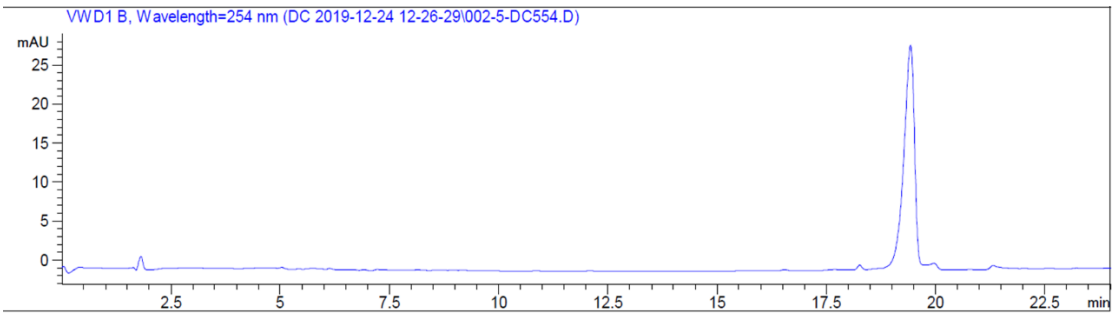

DC553-500 390 (13.251)
1: Scan ES+
9.19e6
100
%
0
487.2479
[M+H]+
488.2362
[2M+H]+
973.3055
974.2937
100
200
300
400
500
600
700
800
900
1000
1100
m/z


VWD1 B, Wavelength=254 nm (DC 2019-12-24 12-26-29\001-4-DC553.D)
mAU
140
120
100
80
60
40
20
0
2.5
5
7.5
10
12.5
15
17.5
20
22.5
min

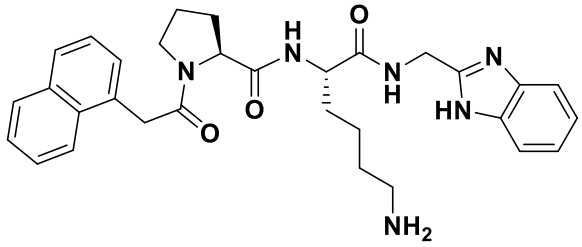

O
H
N
N
O
N
H
O
NH2

(22)

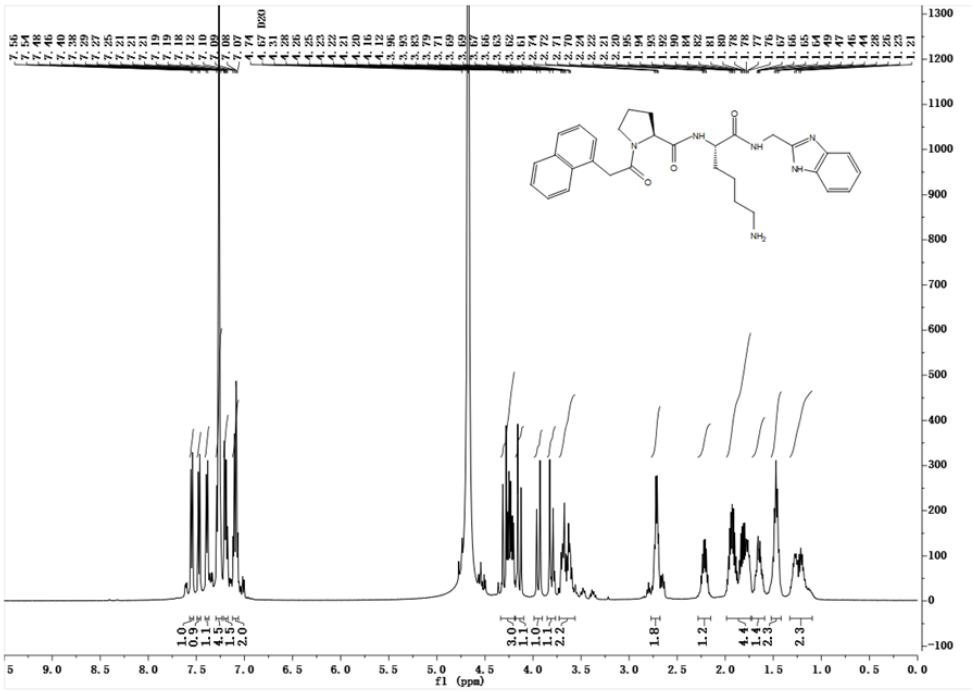

DC554-500 416 (14.135)
1: Scan ES+
5.39e6
100
%
0
537.1959
[M+H]+
538.1462
[2M+H]+
1073.2986
1074.2867
200
300
400
500
600
700
800
900
1000
1100
m/z

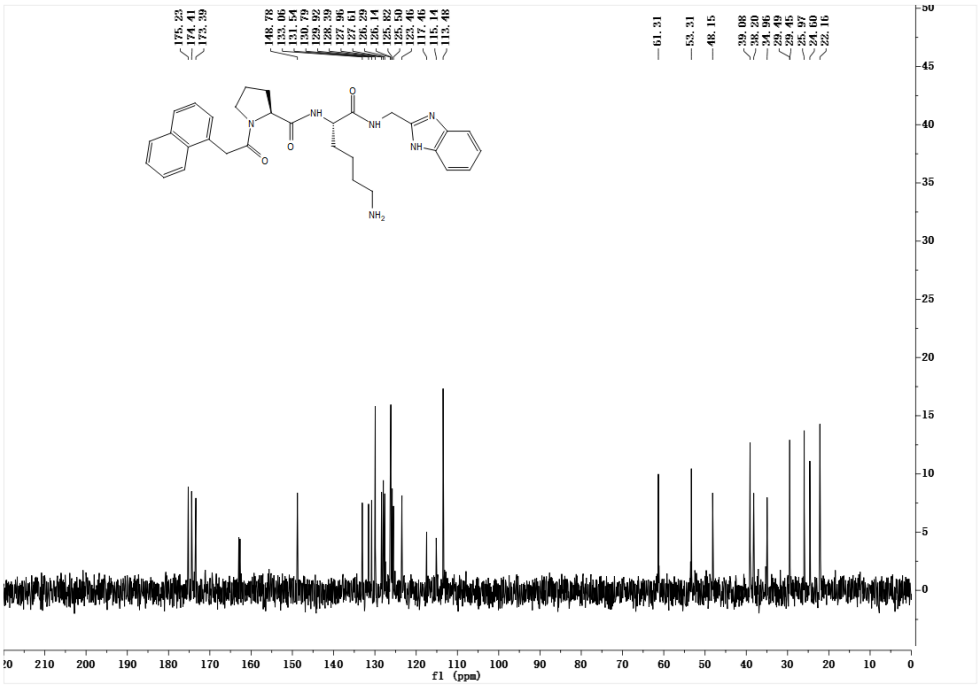

VWD1 B, Wavelength=254 nm (DC 2019-12-24 12-26-29\002-5-DC554.D)
mAU
25
20
15
10
5
0
2.5
5
7.5
10
12.5
15
17.5
20
22.5
min

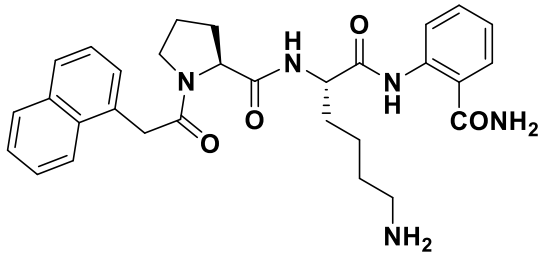
NMR spectrums of compounds **18**, **20**, **21**, **22** and **23**



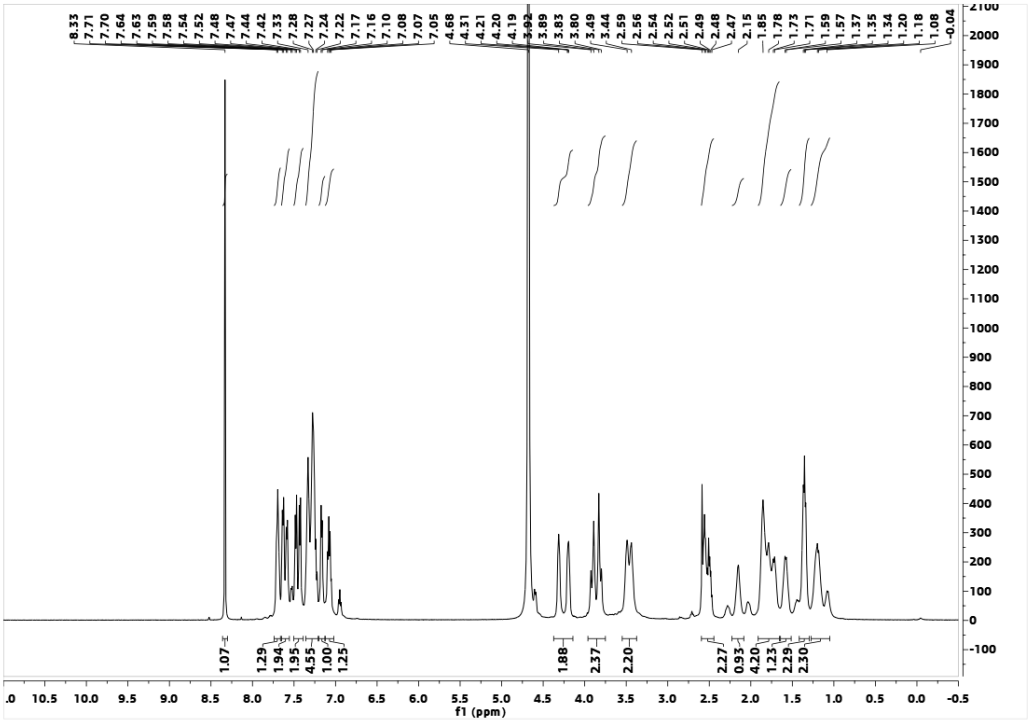
(18)

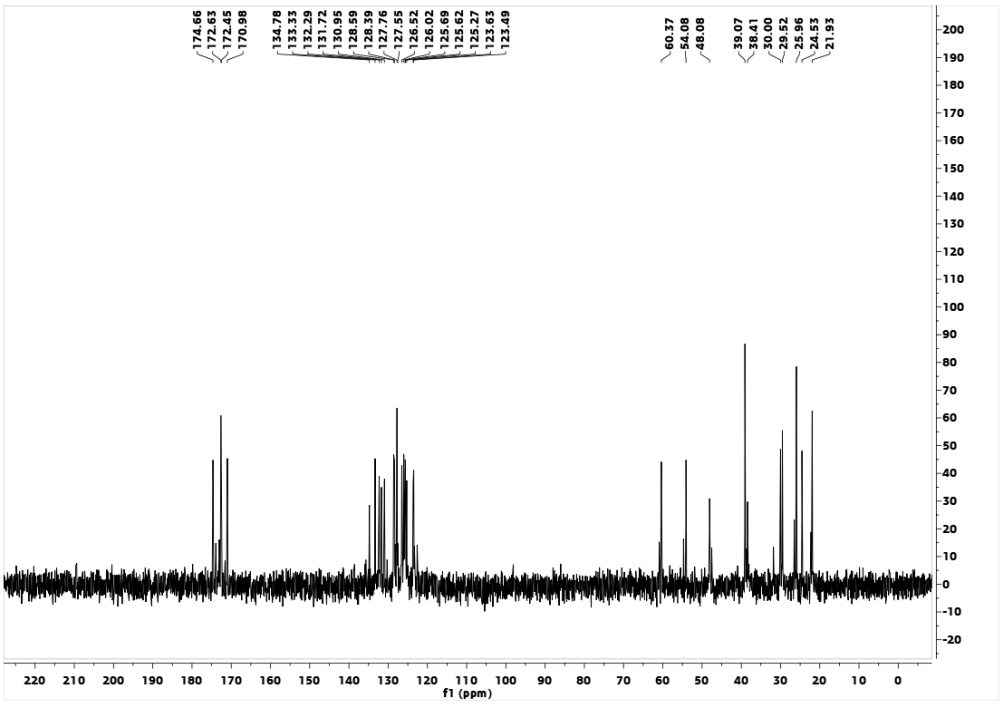


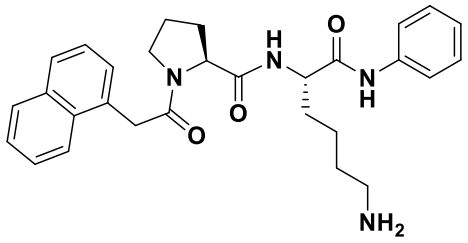



O
H
N
N
O
N
H
CONH2
O
O
NH2

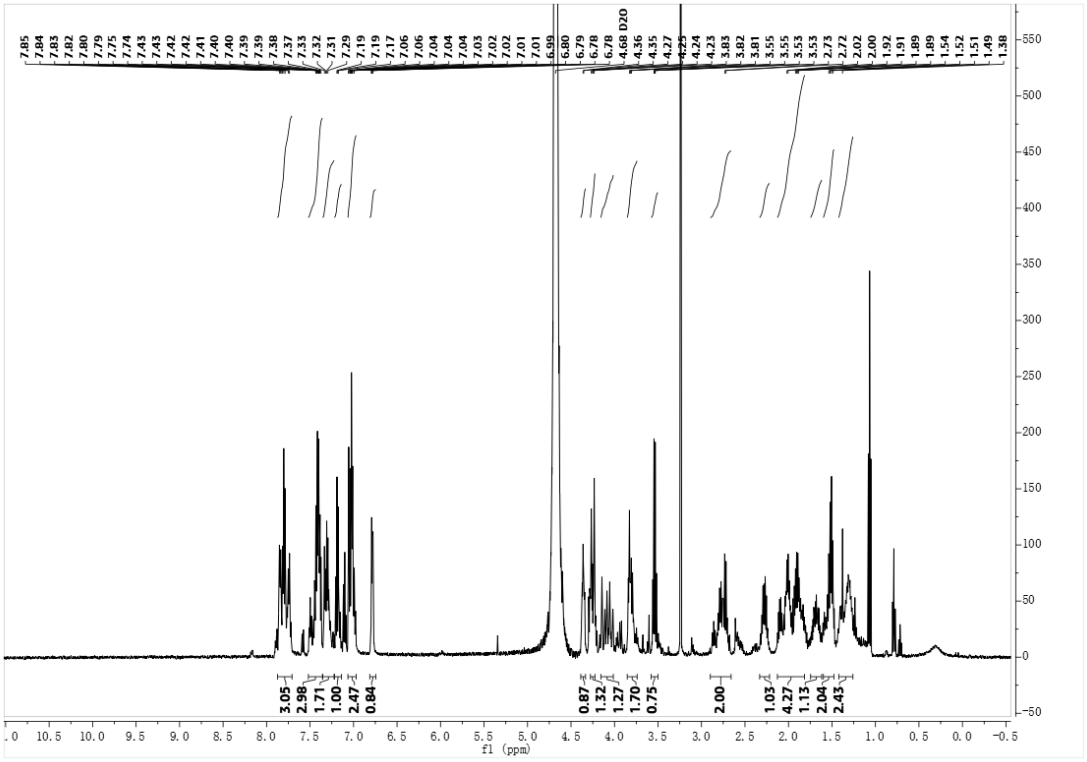
(20)

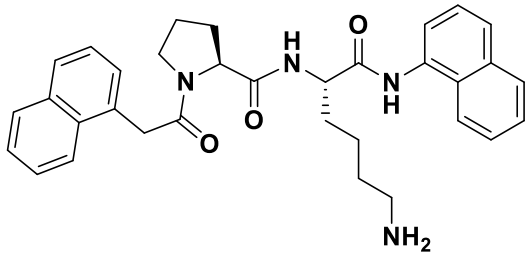

8.33
7.71
7.70
7.64
7.63
7.59
7.58
7.54
7.52
7.48
7.47
7.44
7.42
7.33
7.28
7.27
7.24
7.22
7.17
7.16
7.10
7.08
7.07
7.05
4.68
4.31
4.21
4.20
4.19
3.92
3.89
3.83
3.80
3.49
3.44
2.59
2.56
2.54
2.52
2.51
2.49
2.48
2.47
2.15
1.85
1.78
1.73
1.71
1.59
1.57
1.37
1.35
1.34
1.20
1.18
1.08
-0.04
1.07
1.29
1.94
1.95
4.55
1.00
1.25
1.88
2.37
2.20
2.27
0.93
4.20
1.23
2.29
2.30
f1 (ppm)


174.66
172.63
172.45
170.98
134.78
133.33
132.29
131.72
130.95
128.59
128.39
127.76
127.55
126.52
126.02
125.69
125.62
125.27
123.63
123.49
60.37
54.08
48.08
39.07
38.41
30.00
29.52
25.96
24.53
21.93
f1 (ppm)

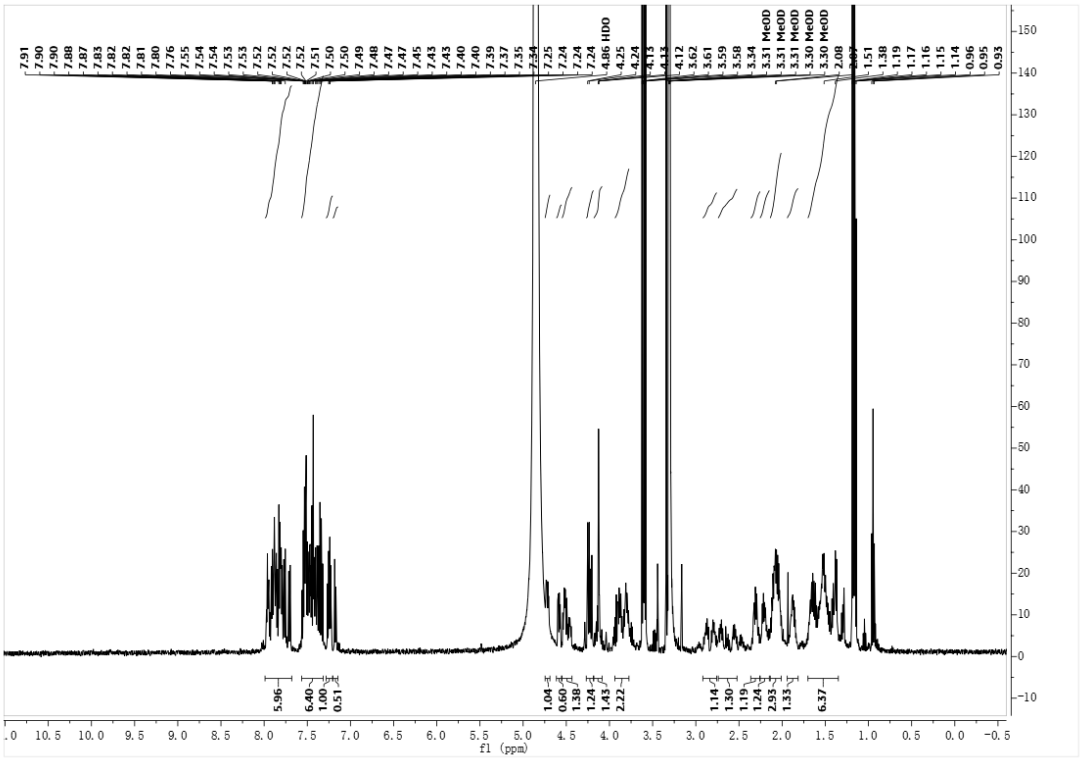

H
N
O
N
H
N
O
O
NH2

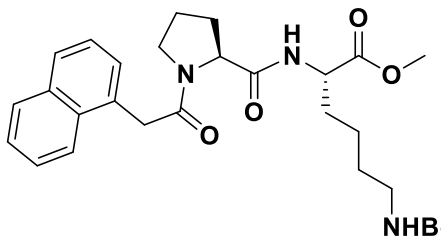
(21)


7.85 7.84 7.83 7.82 7.80 7.79 7.75 7.74 7.43 7.43 7.42 7.42 7.41 7.40 7.40 7.39 7.39 7.38 7.37 7.33 7.32 7.31 7.29 7.19 7.17 7.06 7.04 7.04 7.03 7.02 7.02 7.01 7.01 6.99 6.80 6.79 6.78 6.78 4.68 D20 4.36 4.35 4.27 4.25 4.24 4.23 3.83 3.82 3.81 3.55 3.55 3.53 3.53 2.73 2.72 2.02 2.00 1.92 1.91 1.89 1.89 1.54 1.52 1.51 1.49 1.38
3.05 2.98 1.71 1.00 2.47 0.84 0.87 1.32 1.27 1.70 0.75 2.00 1.03 4.27 1.13 2.04 2.43
f1 (ppm)

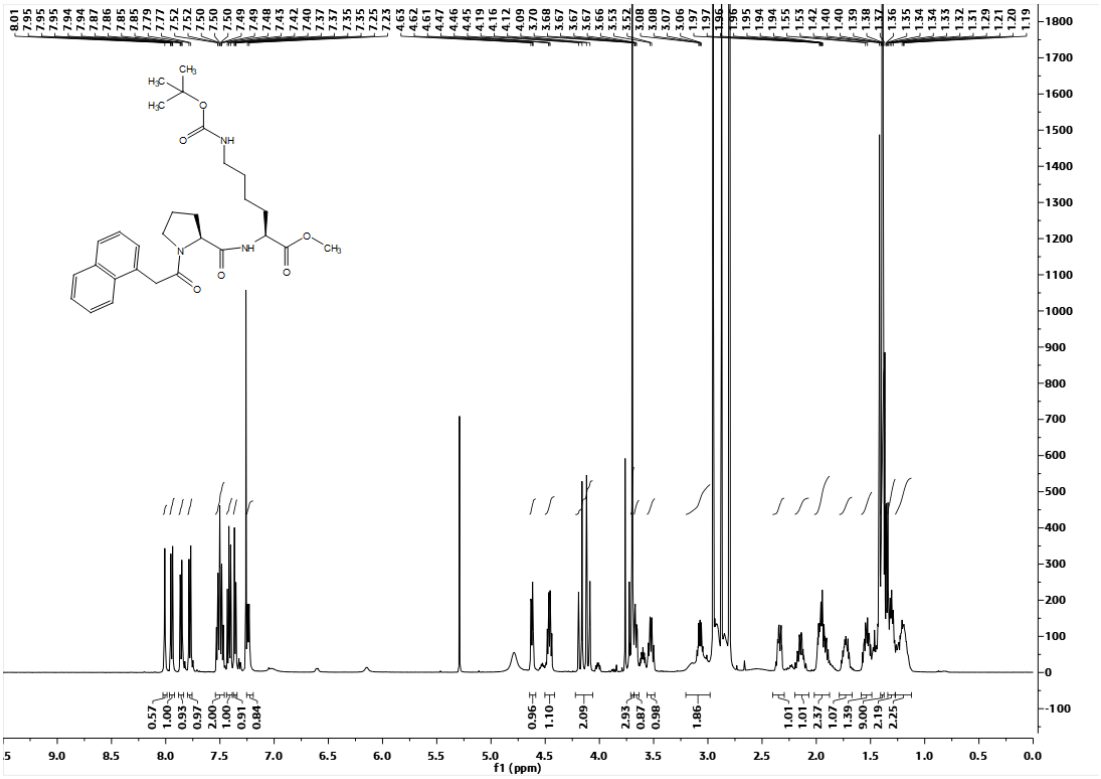


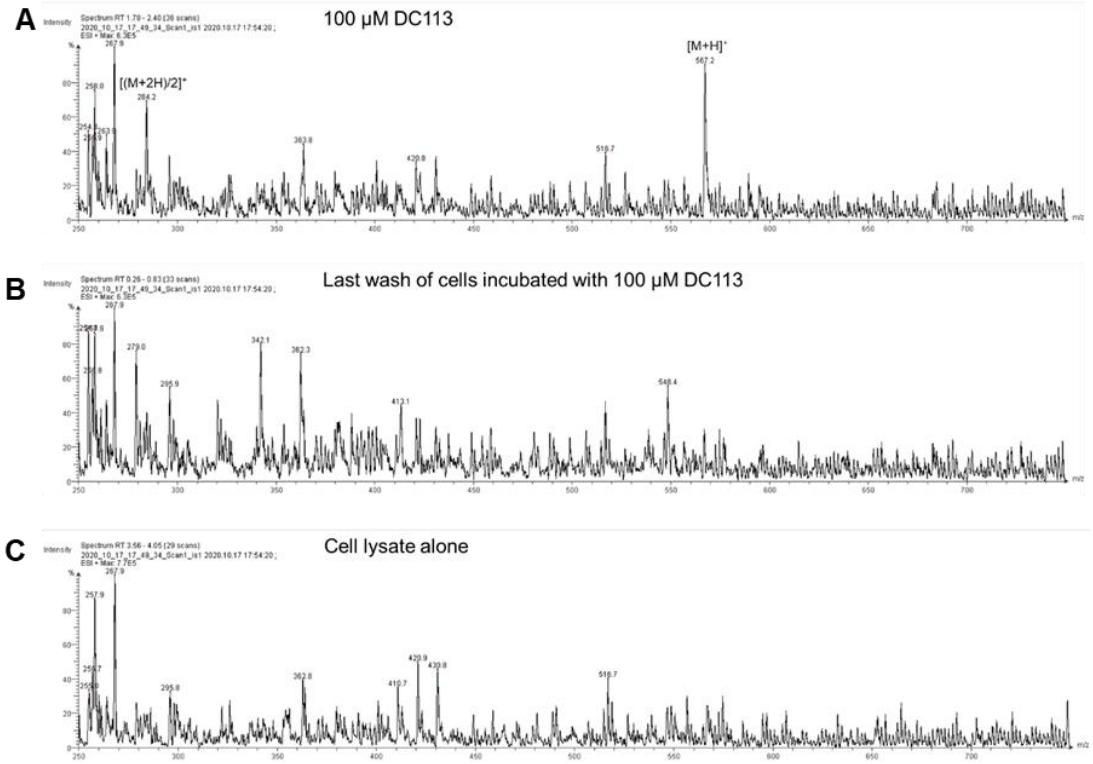
(22)



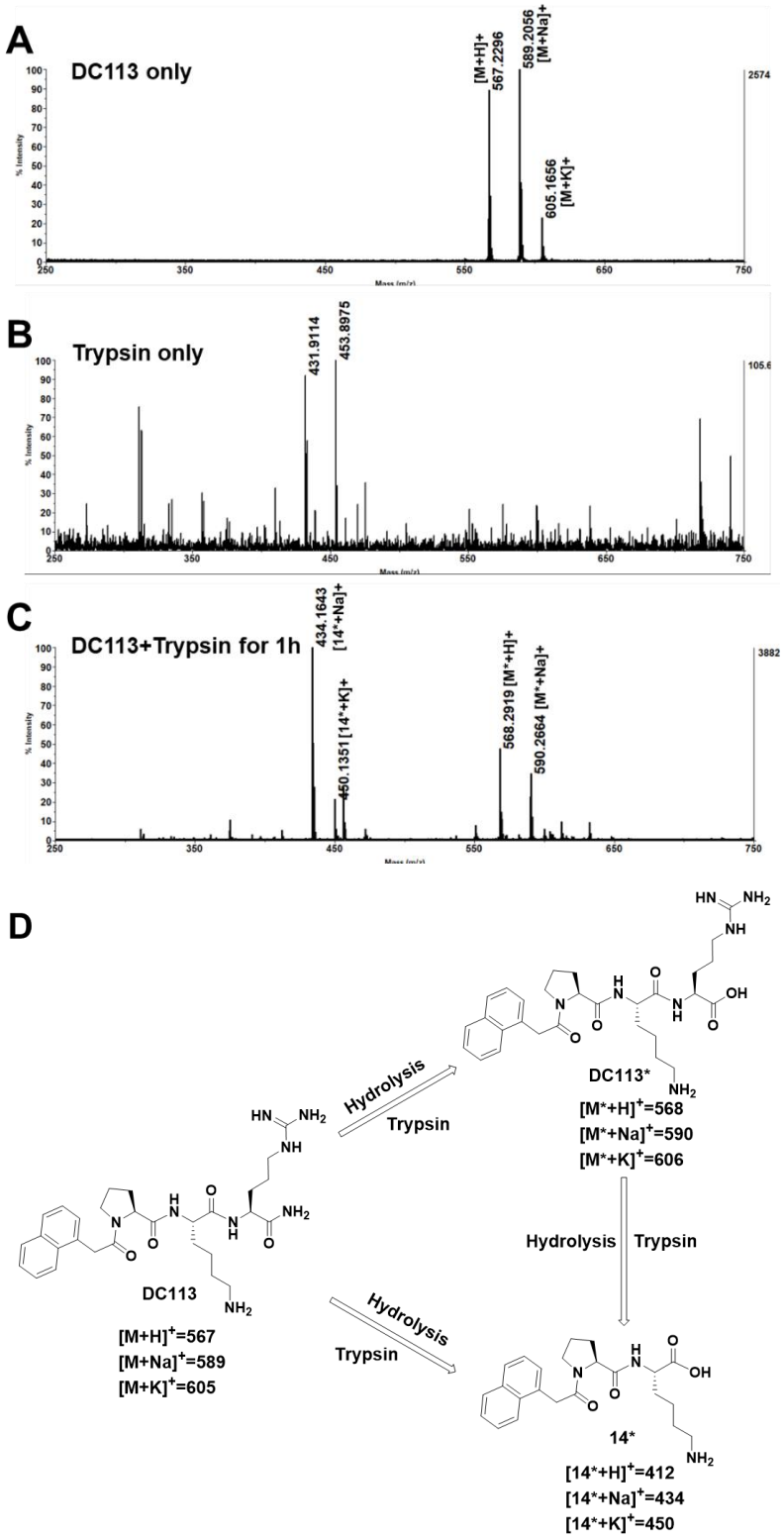


(23)

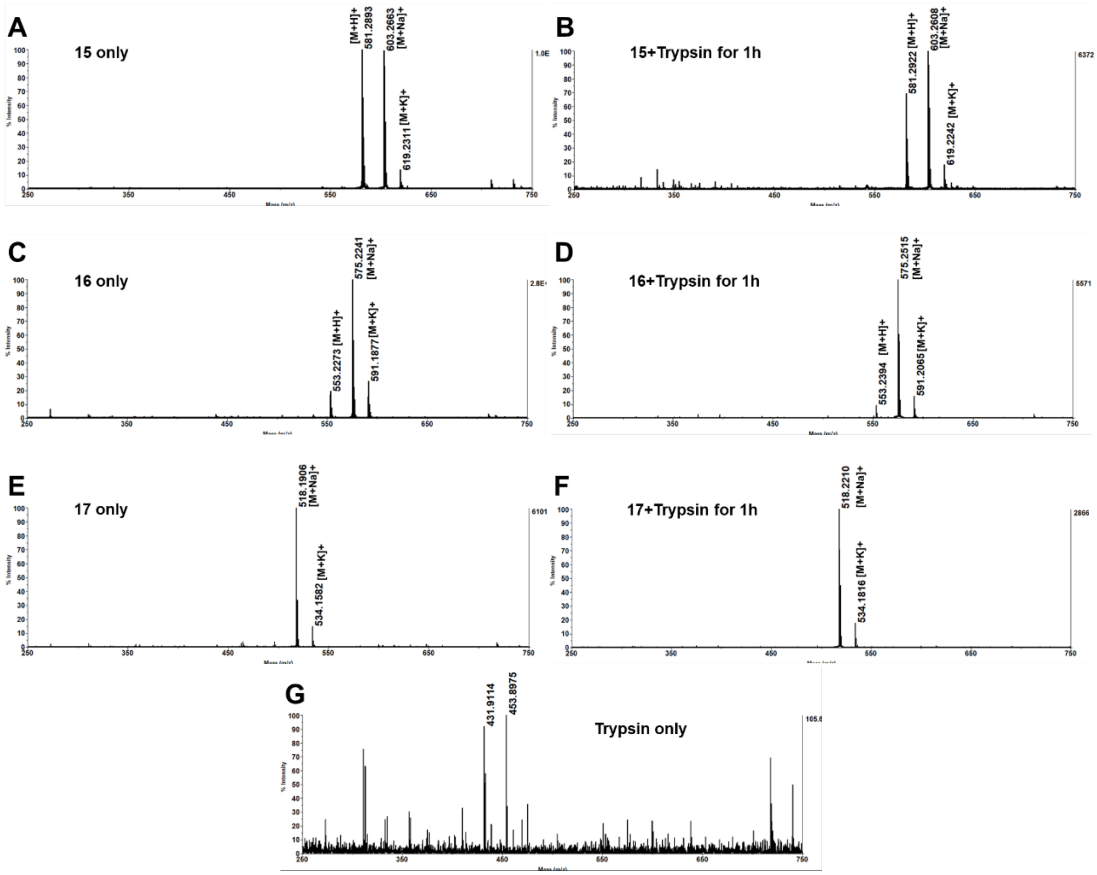




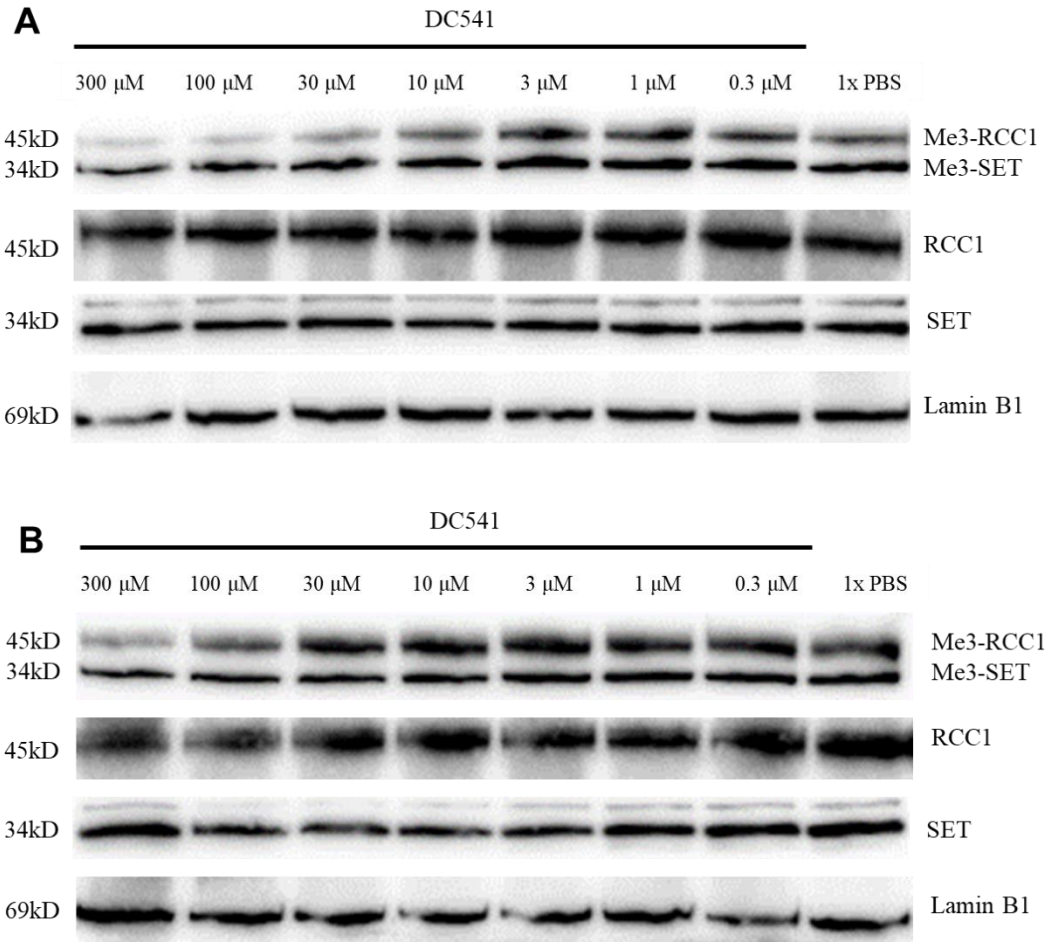
Supplementary Figure S1. Cell permeability evaluation of DC113. (A) MS for incubation of 100 μM DC113 with HCT116 cells; (B) MS for last wash of cells incubated with DC113; (C) MS for cell lysates only. [**DC113**+H]$^+$ = 567, [**DC113**+Na]$^+$ = 589, [**DC113**+K]$^+$ = 605, [(**DC113**+2H)/2]$^+$=284.



Supplementary Figure S2. Stability evaluation of compound **DC113**. (A) MALDI-MS for compound **DC113** only; (B) MALDI-MS for Trypsin only; (C) MALDI-MS for incubation of 10 μM **DC113** with Trypsin for 1h; (D) Proposed hydrolysis process for DC113 with Trypsin.



Supplementary Figure S3. Stability evaluation of compounds **15-17**. (A), (C) and (E) MALDI-MS for compounds **15-17** only, respectively; (B), (D) and (F) MALDI-MS for incubation of 10 µM compounds 15-17 with Trypsin for 1h, respectively; (G) MALDI-MS for Trypsin only. [**15**+H]$^+$ = 581, [**15**+Na]$^+$ = 603, [**15**+K]$^+$ = 619; [**16**+H]$^+$ = 553, [**16**+Na]$^+$ = 575, [**15**+K]$^+$ = 591; [**17**+H]$^+$ =481, [**15**+Na]$^+$ =518, [**15**+K]$^+$ = 534.



Supplementary Figure S4. Inhibition of cellular N-terminal methylation by **DC541** in HT29 and HCT116 cells. Western blot results of effects of **DC541** (0 – 300 μM) on the cellular methylation level of (A) HT29 and (B) HCT116 cells. RCC1, SET, and Lamin B1 blots are shown as loading controls.

Supplementary Table S1. Crystallography data and refinement statistics.

Data Collection	**NTMT1/DC113**
λ (Å)	1.0332
Space group	$P2_12_12_1$
a, b, c (Å)	74.86, 81.66, 84.34
α, β, γ (º)	90, 90, 90
Resolution (Å)*	50 – 2.35 (2.43 – 2.35)
Completeness (%)*	88.1 (92.0)
Redundancy*	3.6 (3.6)
R_{sym} †*	0.24 (0.90)
I / σ (I)*	9.7 (2.4)
$CC_{1/2}$	0.96 (0.52)
Refinement	
Resolution (Å)	42 – 2.34
No. reflections	19,700
$R^{§}/R_{free}$¶	0.21/0.26
r.m.s. deviations	
Bonds (Å)	0.004
Angles (º)	0.845
No. Protein atoms	3536
No. Ligand atoms	134
No. Waters	209
B-factors (Å^2)	
Wilson B	20.95
Protein	26.01
Ligands	26.81
Waters	29.76
Ramachandran Analysis¥	
Favored (%)	96.58
Allowed (%)	3.42
Outliers (%)	0
PDB code	7K3D

† $R_{sym} = \Sigma_{hkl,j} (|I_{hkl}-\langle I_{hkl}\rangle|) / \Sigma_{hkl,j} I_{hkl}$, where $\langle I_{hkl}\rangle$ is the average intensity for a set of j symmetry related reflections and I_{hkl} is the value of the intensity for a single reflection within a set of symmetry-related reflections.

§ *R* factor $= \Sigma_{hkl} (||F_o| - |F_c||) / \Sigma_{hkl}|F_o|$ where F_o is the observed structure factor amplitude and F_c is the calculated structure factor amplitude.

¶ $R_{free} = \Sigma_{hkl,T} (||F_o| - |F_c||) / \Sigma_{hkl,T}|F_o|$, where a test set, T (5% of the data), is omitted from the refinement.

¥ Performed using Molprobity within PHENIX.

* Indicates statistics for last resolution shell shown in parenthesis.

References

(1) Richardson, S. L.; Mao, Y.; Zhang, G.; Hanjra, P.; Peterson, D. L.; Huang, R. Kinetic Mechanism of Protein N-Terminal Methyltransferase 1. *J. Biol. Chem.* **2015**, *290*, 11601–11610.

(2) Dong, C.; Mao, Y.; Tempel, W.; Qin, S.; Li, L.; Loppnau, P.; Huang, R.; Min, J. Structural Basis for Substrate Recognition by the Human N-Terminal Methyltransferase 1. *Genes Dev.* **2015**, *29*, 2343-2348.

(3) Wu, H.; Min, J.; Lunin, V. V.; Antoshenko, T.; Dombrovski, L.; Zeng, H.; Allali-Hassani, A.; Campagna-Slater, V.; Vedadi, M.; Arrowsmith, C. H.; Plotnikov, A. N.; Schapira, M. Structural Biology of Human H3K9 Methyltransferases. *PLoS One* **2010**, *5*, e8570.

(4) Feng, Y.; Xie, N.; Jin, M.; Stahley, M. R.; Stivers, J. T.; Zheng, Y. G. A Transient Kinetic Analysis of PRMT1 Catalysis. *Biochemistry* **2011**, *50*, 7033–7044.

(5) Barsyte-Lovejoy, D.; Li, F.; Oudhoff, M. J.; Tatlock, J. H.; Dong, A.; Zeng, H.; Wu, H.; Freeman, S. A.; Schapira, M.; Senisterra, G. A.; Kuznetsova, E.; Marcellus, R.; Allali-Hassani, A.; Kennedy, S.; Lambert, J.-P.; Couzens, A. L.; Aman, A.; Gingras, A.-C.; Al-Awar, R.; Fish, P. V.; Gerstenberger, B. S.; Roberts, L.; Benn, C. L.; Grimley, R. L.; Braam, M. J. S.; Rossi, F. M. V.; Sudol, M.; Brown, P. J.; Bunnage, M. E.; Owen, D. R.; Zaph, C.; Vedadi, M.; Arrowsmith, C. H. (R)-PFI-2 Is a Potent and Selective Inhibitor of SETD7 Methyltransferase Activity in Cells. *Proc. Natl. Acad. Sci.* **2014**, *111*, 12853–12858.

(6) Debler, E. W.; Jain, K.; Warmack, R. A.; Feng, Y.; Clarke, S. G.; Blobel, G. A Glutamate / Aspartate Switch Controls Product Specificity in a Protein Arginine Methyltransferase. *Proc. Natl. Acad. Sci.* **2016**, *113*, 2068-2073.

(7) Mackie, B. D.; Chen, D.; Dong, G.; Dong, C.; Parker, H.; Schaner Tooley, C. E.; Noinaj, N.; Min, J.; Huang, R. Selective Peptidomimetic Inhibitors of NTMT1/2: Rational Design, Synthesis, Characterization, and Crystallographic Studies. *J. Med. Chem.* **2020**, *63*, 9512−9522.

(8) Chen, D.; Li, L.; Diaz, K.; Iyamu, I. D.; Yadav, R.; Noinaj, N.; Huang, R. Novel Propargyl-Linked Bisubstrate Analogues as Tight-Binding Inhibitors for Nicotinamide N-Methyltransferase. *J. Med. Chem.* **2019**, *62*, 10783–10797.

(9) Chen, D.; Dong, G.; Noinaj, N.; Huang, R. Discovery of Bisubstrate Inhibitors for Protein N-Terminal Methyltransferase 1. *J. Med. Chem.* **2019**, *62*, 3773–3779.