Electronic Supporting Information

Synthesis of selenocysteine-containing cyclic peptides via tandem *N*-to-*S* acyl migration and intramolecular selenocysteine-mediated native chemical ligation

Shingo Shimodaira,[a] Toshiki Takei,[b], Hironobu Hojo*[b] and Michio Iwaoka*[a]

[a] School of Science, Department of Chemistry, Tokai University, Kitakaname, Hiratsuka-shi, Kanagawa 259-1292, Japan

[b] Institute for Protein Research, Osaka University, Yamadaoka, Suita-shi, Osaka 565-0871, Japan

Contents

1. General experimental protocols for SPPS

Solid-phase peptide synthesis (SPPS) of precursor selenopeptides was carried out with the Fmoc strategy using the HBTU-DIEA or DIC-HOBt activation method. Completion of the amino acid coupling was assessed by the Kaiser test. When selenocysteine (Sec, U) was being loaded to the peptide, solvent was changed to dichloromethane (DCM) from *N*-methyl-2-pyrrolidone (NMP). The resulting resin was treated with 20 % piperidine/DCM and then a TFA cocktail (TFA:TIS:H_2O:TA:Py_2S_2=90:2.5:2.5:2.5:2.5). After agitation, a precursor selenopeptide was obtained as a mixture of a linear selenopeptide with SPy protection on the Se atom and a cyclic selenopeptide with an intramolecular Se—S linkage (for example, **9** and **10** for the synthesis of **1**). The crude mixture was subsequently cyclized using MESNa under intramolecular Sec-NCL conditions

and then treated with DTT (or 2-mercaptoethanol for the case of **2**). The obtained selenopeptide was purified by preparative PR-HPLC and was characterized by MALDI-TOF-MS and amino acid analysis (AAA). The AAA sample solution was prepared by hydrolyzing the cyclic selenopeptides with 6 M HCl at 150°C for 2 h in an evacuated sealed tube. Details of the synthesis of selenopeptides **1-5** are described below. Linear selenopeptide **6** is a known compound.[19a] Linear selenopeptides **7** and **8** were prepared according to the method standardized previously for the synthesis of **5**. *N*-(9-Fluorenylmethoxycarbonyl)-*Se*-(*p*-methoxyphenylmethyl)-selenocysteine (Fmoc-Sec(MPM)-OH) was synthesized by the literature method.[S1] Fmoc-Gly-(Et)Cys(Trt)-OH was synthesized by condensation of *N*-ethyl-*S*-triphenylmethyl-L-cysteine (Et-Cys(Trt)-OH), which was prepared by reductive amination of acetaldehyde with H-Cys(Trt)-OH using $NaBH_3CN$, with Fmoc-Gly-OPfp, which was prepared by reacting Fmoc-Gly-OH with pentafluorophenol (PfpOH), in the presence of HOOBt according to literature.[17,18]

Analytical PR-HPLC was carried out using a Mightysil RP-18 GP II reverse-phase column (4.6 x 150 mm, Kanto, Tokyo) equilibrated at 50°C. A volume of the sample solution loop was 1 mL. An injection volume was 100 μL. A solvent gradient was applied by linearly increasing a ratio of eluent B (0.1 % TFA in acetonitrile) from 5 to 30 % in 25 min against eluent A (0.1 % TFA in H_2O) at a flow rate of 1.0 mL/min after sample injection. For cyclic selenopeptide **2** and linear selenopeptides **7** and **8**, the ratio of eluent B was linearly increased from 1 to 11 % in 20 min and 10 to 35 % in 25 min, respectively. Peptides were detected by the absorbance at 220 nm.

2. Synthesis of [c-KQGUGWGN-]$_2$ (1)

Fmoc-Rink amide MBHA resin (52.6 mg, 20 μmol) was weighed into a plastic reaction vessel, where the resin was swelled with NMP for 30 min at r.t. The resin was treated with 20 % piperidine/NMP for 1 min with vortex mixing. The deprotection reaction was repeated in fleshly prepared 20 % piperidine/NMP for 5 min. The resin was then washed with NMP (x5). Fmoc-Lys(Boc)-OBt, which was prepared by mixing Fmoc-Lys(Boc)-OH (49.2 mg, 100 μmol), 0.45M HBTU/DMF (211 μL, 95 μmol) and DIEA (34.8 μL, 200 μmol) for 5 min, was added to the resin. The mixture was vortexed for 12 min at r.t. After the coupling, the resin was washed with NMP (x3). The deprotection of the Fmoc group was repeated with a similar procedure as described above. Fmoc-Gly-(Et)Cys(Trt)-OBt prepared by mixing Fmoc-Gly-(Et)Cys(Trt)-OH (27.7 mg, 40 μmol), HOBt (10.8 mg, 80 μmol) and DIC (12.4 μL, 80 μmol) was coupled to the resin in DCE for overnight at 37°C. Fmoc-Gln(Trt)-OH (57.6 mg, 100 μmol), Fmoc-Lys(Boc)-OH (51.2 mg, 100 μmol), Fmoc-Asn(Trt)-OH (62.2 mg, 100 μmol), Fmoc-Gly-OH (32.9 mg, 100 μmol), Fmoc-Trp(Boc)-OH (54.9 mg, 100 μmol) and Fmoc-Gly-OH (32.3 mg, 100 μmol) were then introduced sequentially to the resin by a similar procedure to that for Fmoc-Lys(Boc)-OH to yield Fmoc-Gly-Trp(Boc)-Gly-Asn(Trt)-Lys(Boc)-Gln(Trt)-Gly-(Et)Cys(Trt)-Lys(Boc)-resin. After deprotection of the Fmoc group, the solvent was changed to DCM from NMP. Fmoc-Sec(MPM)-OBt, which was prepared by mixing Fmoc-Sec(MPM)-OH (22.4 mg, 40 μmol), HOBt (51.4 mg, 80 μmol) and DIC (12.3 μL, 80 μmol) in DCM for 30 min at r.t., was added to the resin, and the mixture was vortexed for 1 h at r.t. After the coupling, the resin was washed with DCM (x3). The resin was treated with 20 % piperidine/DCM for 5 min, and it was repeated with the fresh 20 % piperidine/DCM for 15 min. The

resulting resin was washed with DCM (x5) and diethyl ether (x3) and was dried in vacuo to yield H-Sec(MPM)-Gly-Trp(Boc)-Gly-Asn(Trt)-Lys(Boc)-Gln(Trt)-Gly-(Et)Cys(Trt)-Lys(Boc)-resin.

The obtained resin was treated with the TFA cocktail (2 mL) for 2 h at r.t. After the removal of TFA by N_2 stream, the deprotected peptide was precipitated with diethyl ether, washed with diethyl ether (x3), and dried in vacuo. The crude peptide was treated with a 0.1 M phosphate buffer solution at pH 5.5 (5 mL) containing 6 M GdmCl and 10 % (v/v) MESNa at 37 °C. After stirring overnight, DTT (250 mg) was added to the mixture solution and stirred for 3 h at the same temperature. The resulting mixture was purified by preparative RP-HPLC to yield **1** (385.5 nmol, 3.1 % based on the resin employed). MALDI-TOF-MS (*m*/*z*) found: 1755.52, calcd for $[M+H]^+$: 1755.58. Amino acid analysis: $Asp_{1.00}Glu_{0.99}Gly_{3.00}Lys_{0.93}$. The isolated yield of **11** from **10** was 62.0 %.

3. Synthesis of [c-KQGUGWGA-]$_2$ (3)

A similar procedure was applied for the synthesis of **3** (1377.8 nmol, 13.8 % based on the resin employed). The isolated yield of **11b** from **10b** was 59.0 %. MALDI-TOF-MS (*m*/*z*) found:1669.37, calcd for $[M+H]^+$: 1669.57. Amino acid analysis: $Glu_{0.86}Gly_{3.00}$ $Ala_{1.02}Lys_{0.95}$.

4. Synthesis of [c-KAGUGWGN-]$_2$ (4)

A similar procedure was applied for the synthesis of **4** (1017.3 nmol, 10.2 % based on the resin employed). The isolated yield of **11c** from **10c** was 60.4 %. MALDI-TOF-MS (*m*/*z*) found: 1641.37, calcd for $[M+H]^+$: 1641.54. Amino acid analysis: $Asp_{1.02}Gly_{3.00}Ala_{0.99}Lys_{1.00}$.

5. Synthesis of [c-KQGUGAGN-]$_2$ (2)

A similar procedure was applied for the synthesis of **2** until cyclization. After cyclization, the resulting mixture was purified by preparative RP-HPLC to yield the MESNa adduct corresponding to cyclic selenopeptide **11**. After lyophilization, the obtained peptide was treated with a 0.1 M phosphate buffer solution at pH 7.0 (5 mL) containing 6 M GdmCl and 1 % (v/v) 2-mercaptoethanol for 2 h at 37 °C. After removal of excess 2-mercaptoethanol by extraction with Et_2O, the resulting mixture was purified by preparative RP-HPLC to yield **2** (710.9 nmol, 7.1 % based on the resin employed). The isolated yield of **11a** from **10a** was 68.8 %. MALDI-TOF-MS (*m*/*z*) found: 1525.35, calcd for $[M+H]^+$: 1525.50. Amino acid analysis: $Asp_{1.34}Glu_{1.18}Gly_{3.00}Ala_{1.27}Lys_{1.27}$.

6. Synthesis of [KQGUGWGN]$_2$ (5)

Fmoc-Asn(Trt)-CREAR Acid resin (50.0 mg, 20 μmol) was weighed into a plastic reaction vessel, where the resin was swelled with NMP for 30 min at r.t. The resin was treated with 20 % piperidine/NMP for 1 min with vortex mixing. The deprotection reaction was repeated in fleshly prepared 20 % piperidine for 5 min. The resin was washed with NMP (x5). Fmoc-Gly-OBt prepared by mixing Fmoc-Gly-OH (33.2 mg, 100 μmol), 0.45M HBTU/DMF (211 μL, 95 μmol) and DIEA (34.8 μL, 200 μmol) for 5 min was added to the resin. The mixture was vortexed for 12 min at r.t. After the coupling, the resin was washed with NMP (x3). Fmoc-

Trp(Boc)-OH (55.5 mg, 100 μmol) and Fmoc-Gly-OH (31.0 mg, 100 μmol) were then introduced to the resin by a similar procedure to yield Fmoc-Gly-Trp(Boc)-Gly-Asn(Trt)-resin. After deprotection of the Fmoc group, the solvent was changed to DCM from NMP. Fmoc-Sec(MPM)-OBt, which was prepared by mixing Fmoc-Sec(MPM)-OH (24.2 mg, 40 μmol) with HOBt (10.8 mg, 80 μmol) and DIC (12.3 μL, 80 μmol) in DCM for 30 min at r.t., was added to the resin, and the mixture was vortexed for 1 h at r.t. After the coupling, the resin was washed with DCM (x3) and treated with 20 % piperidine/DCM for 5 min. The deprotection reaction was repeated with the fleshly prepared 20 % piperidine/DCM for 15 min. the resin was washed with DCM (x5). Fmoc-Gly-OH (30.8 mg, 100 μmol), Fmoc-Gln(Trt)-OH (67.5 mg, 100 μmol) and Fmoc-Lys(Boc)-OH (47.4 mg, 100 μmol) were sequentially introduced to the resin by a similar procedure. The Fmoc group of the *N* terminal was finally deprotected. The resulting resin was washed with DCM (x5) and diethyl ether (x3) and was dried in vacuo to yield H-Lys(Boc)-Gln(Trt)-Gly-Sec(MPM)-Gly-Trp(Boc)-Gly-Asn(Trt)-resin.

The obtained resin was treated with the TFA cocktail (2 mL) for 2 h at r.t. After removal of TFA by N_2 stream, the deprotected peptide was precipitated with diethyl ether, washed with diethyl ether (x3), and dried in vacuo. The obtained crude peptide was reacted with DTT (12.3 mg, 80 μmol) in 10 % aqueous acetonitrile (5 mL) containing 0.1 % TFA. The resulting mixture was purified to yield **5** (593.3 nmol, 5.9 %). MALDI-TOF-MS (*m*/*z*) found: 1791.38, calcd for $[M+H]^+$: 1791.60. Amino acid analysis: $Asp_{1.03}Glu_{0.72}Gly_{3.00}Lys_{0.90}$.

7. Synthesis of [RQAUAWNG]$_2$ (7)

A similar procedure to the synthesis of **5** was applied in a 50 μmol scale. Yield of **7**: 2575 nmol (10.3 %). MALDI-TOF-MS (*m*/*z*) found: 1903.59, calcd for $[M+H]^+$: 1903.68. Amino acid analysis: $Arg_{1.06}Glu_{0.95}Ala_{2.11}$ $Gly_{1.00}$ $Asp_{1.05}$.

8. Synthesis of [RQPUPWNG]$_2$ (8)

A similar procedure to the synthesis of **5** was applied in a 50 μmol scale. Yield of **8**: 5300 nmol (21.2 %). MALDI-TOF-MS (*m*/*z*) found: 2007.70, calcd for $[M+H]^+$: 2007.74. Amino acid analysis: $Arg_{0.94}Glu_{0.95}Pro_{2.18}$ $Gly_{1.00}$ $Asp_{0.95}$.

9. Characterization of the products and intermediates

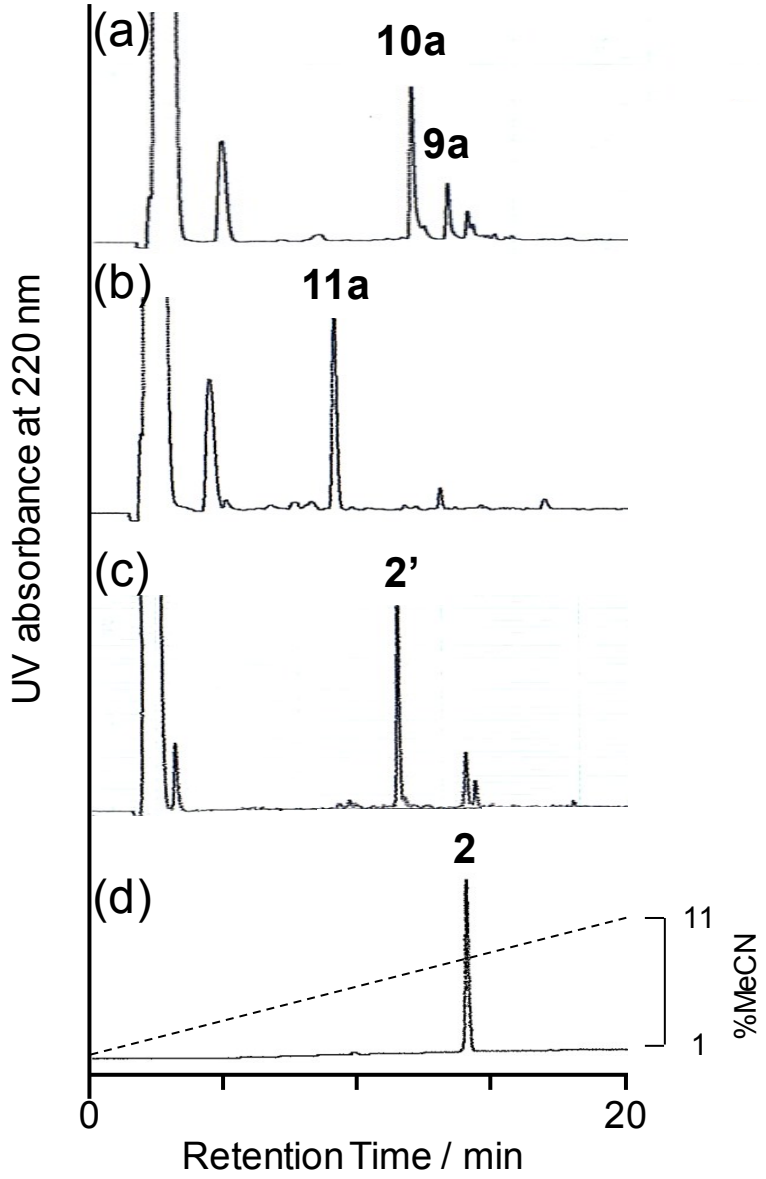

Fig. S1 Changes of RP-HPLC chromatograms during transformation of a crude peptide liberated from the resin to **2**. (a) Before the reaction. (b) After the treatment with MESNa at pH 5.5 and 37˚C overnight. in the presence of 6 M GdmCl. (c) After addition of 2-mercaptoethanol to the purified product (**11a**). (d) After purification by preparative RP-HPLC.

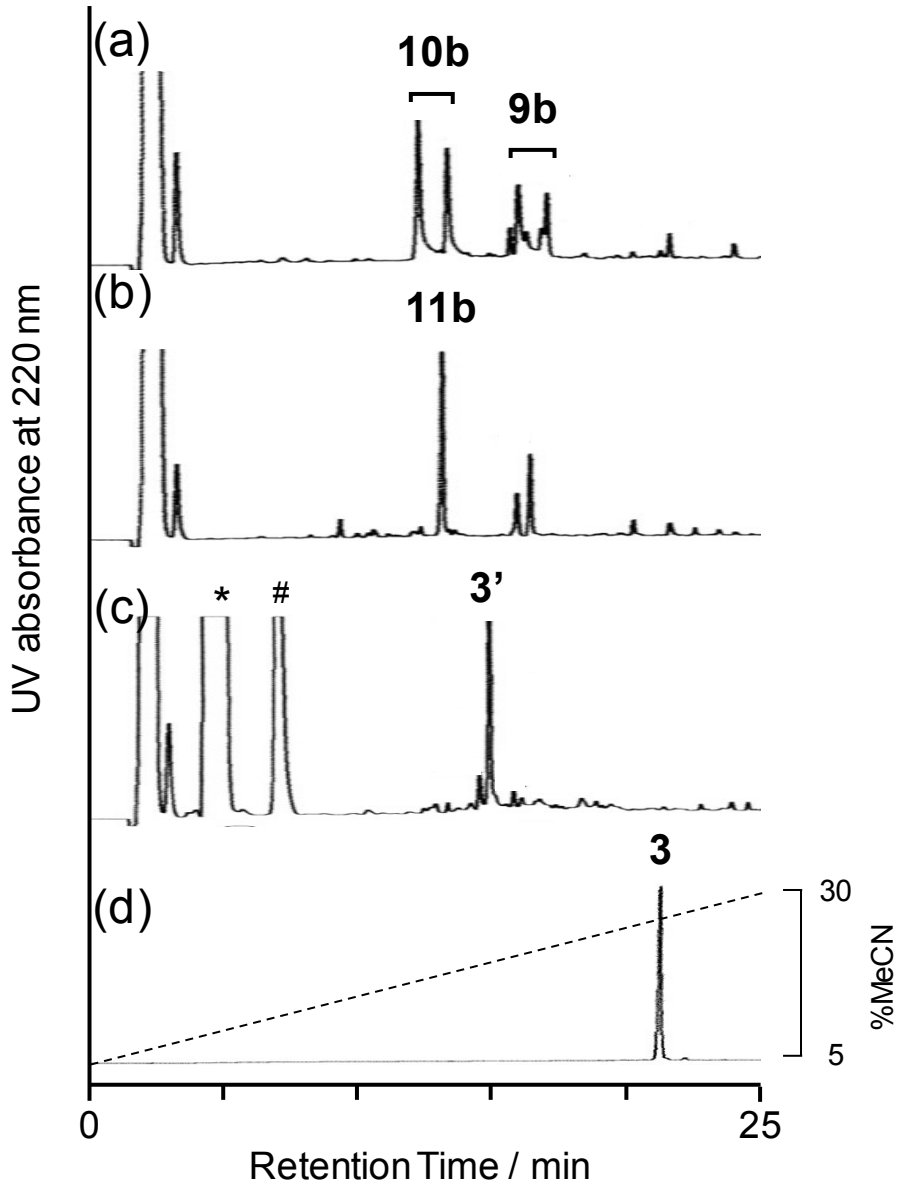

Fig. S2 Changes of RP-HPLC chromatograms during transformation of a crude peptide liberated from the resin to **3**. (a) Before the reaction. (b) After the treatment with MESNa at pH 5.5 and 37˚C overnight. in the presence of 6 M GdmCl. (c) After addition of DTT to the reaction mixture. *=DTT and #=oxidized DTT. (d) After purification by preparative RP-HPLC.

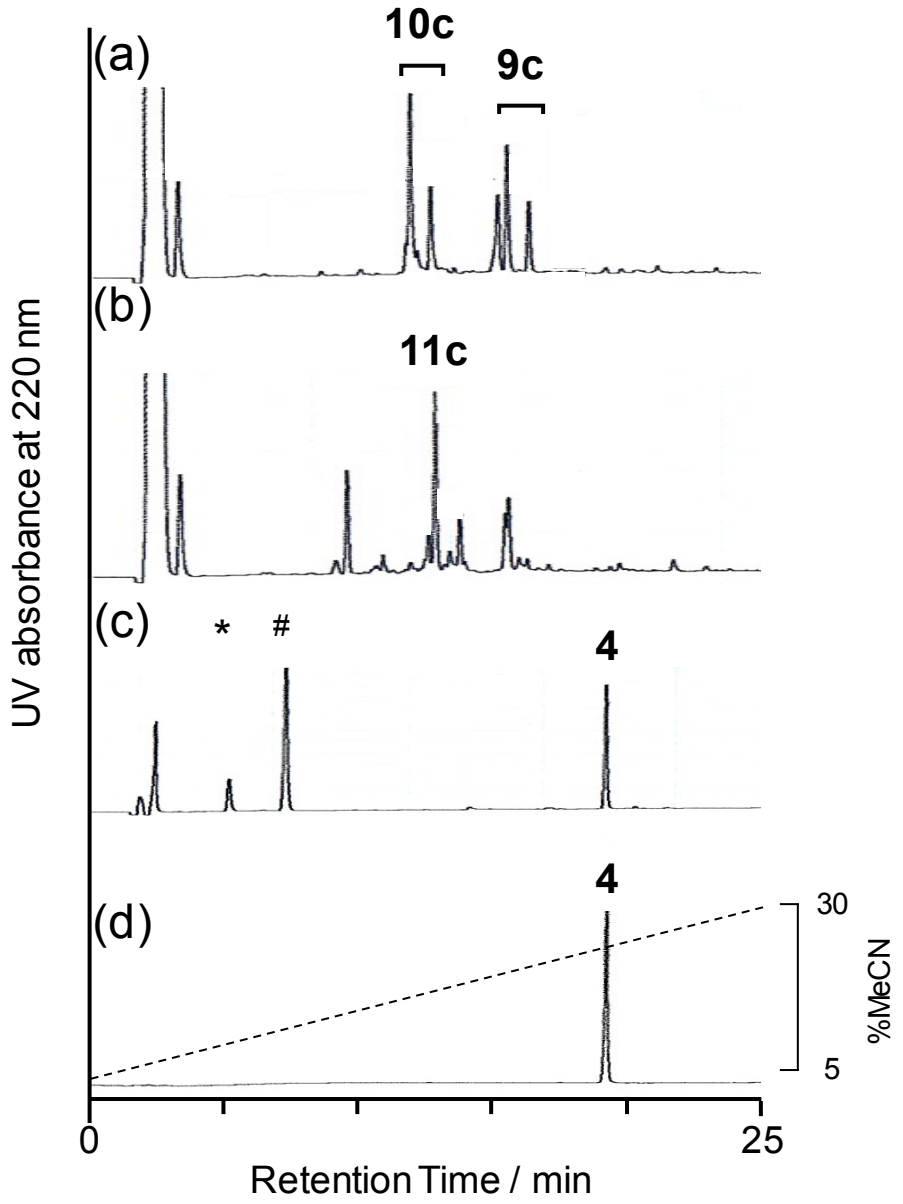

Fig. S3 Changes of RP-HPLC chromatograms during transformation of a crude peptide liberated from the resin to **4**. (a) Before the reaction. (b) After the treatment with MESNa at pH 5.5 and 37˚C overnight. in the presence of 6 M GdmCl. (c) After addition of DTT to the purified product (**11**). *=DTT and #=oxidized DTT. (d) After purification by preparative RP-HPLC.

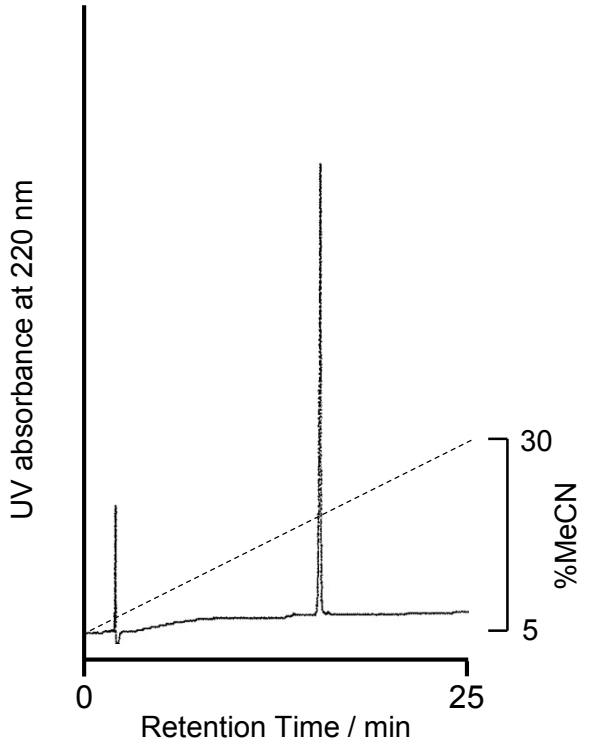

Fig. S4 RP-HPLC chromatogram for purified **5**.

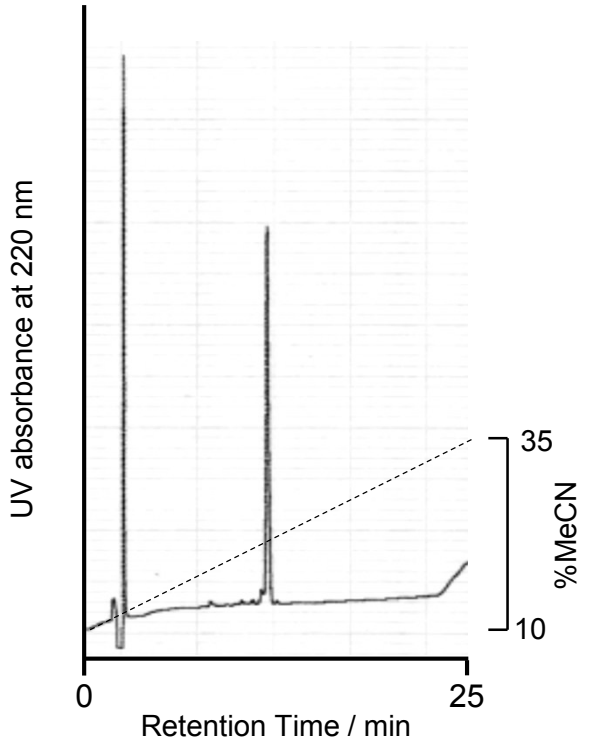

Fig. S5 RP-HPLC chromatogram for purified **7**.

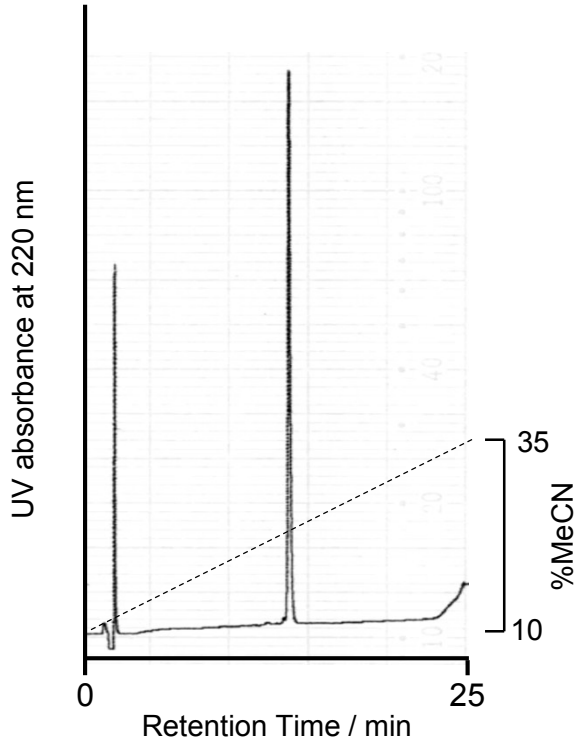

Fig. S6 RP-HPLC chromatogram for purified **8**.

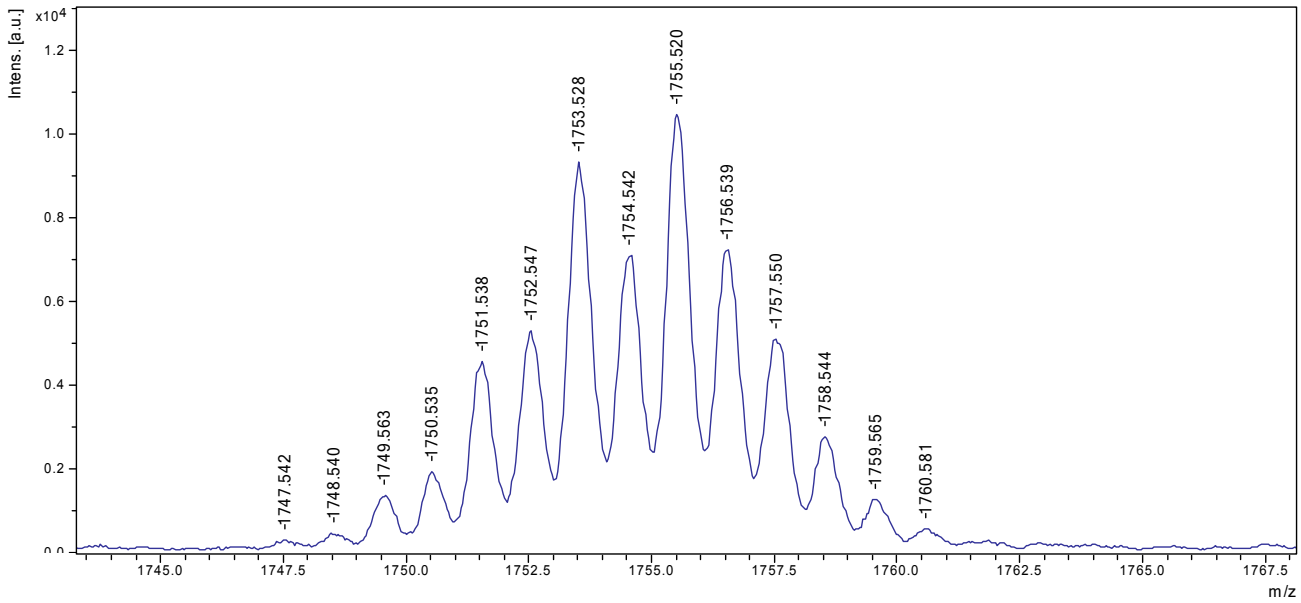

Fig. S7 MALDI-TOF-MS spectrum for purified **1**. MALDI-MS (*m*/*z*) found: 1755.52, calcd for $[M+H]^+$: 1755.58.

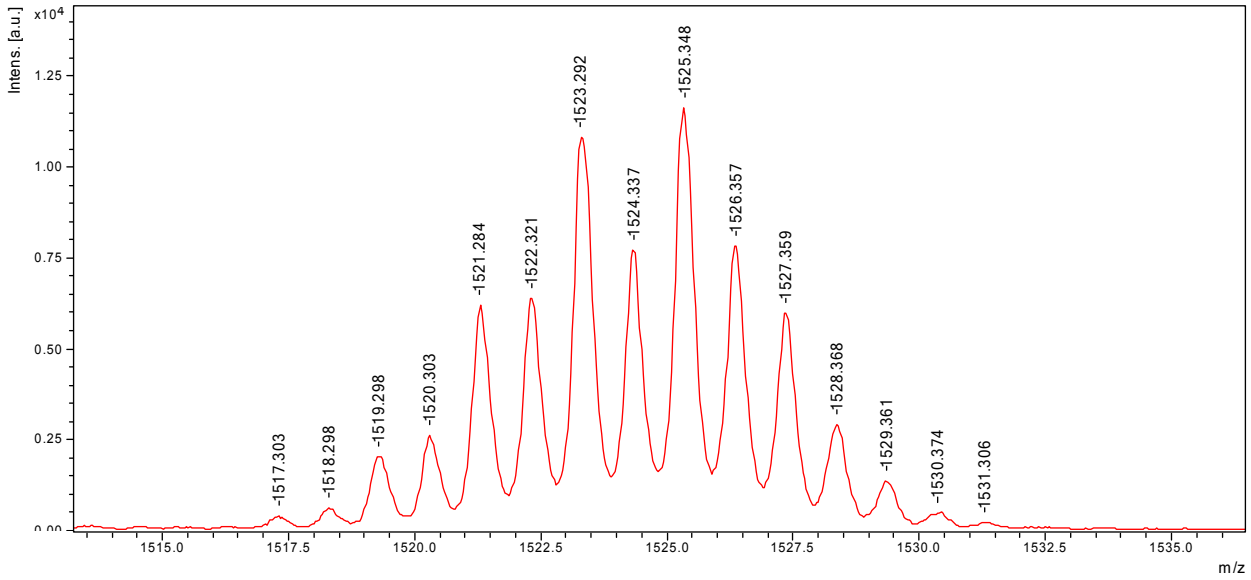

Fig. S8 MALDI-TOF-MS spectrum for purified **2**. MALDI-MS (*m*/*z*) found: 1525.35, calcd for $[M+H]^+$: 1525.50.

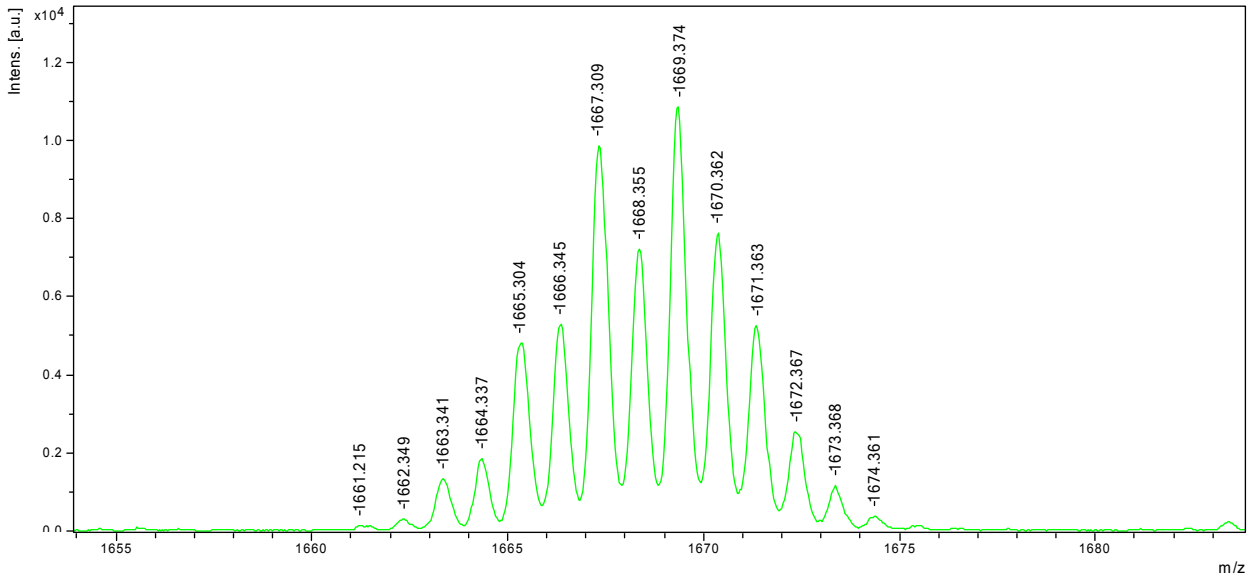

Fig. S9 MALDI-TOF-MS spectrum for purified **3**. MALDI-MS (*m*/*z*) found: 1669.37, calcd for $[M+H]^+$: 1669.57.

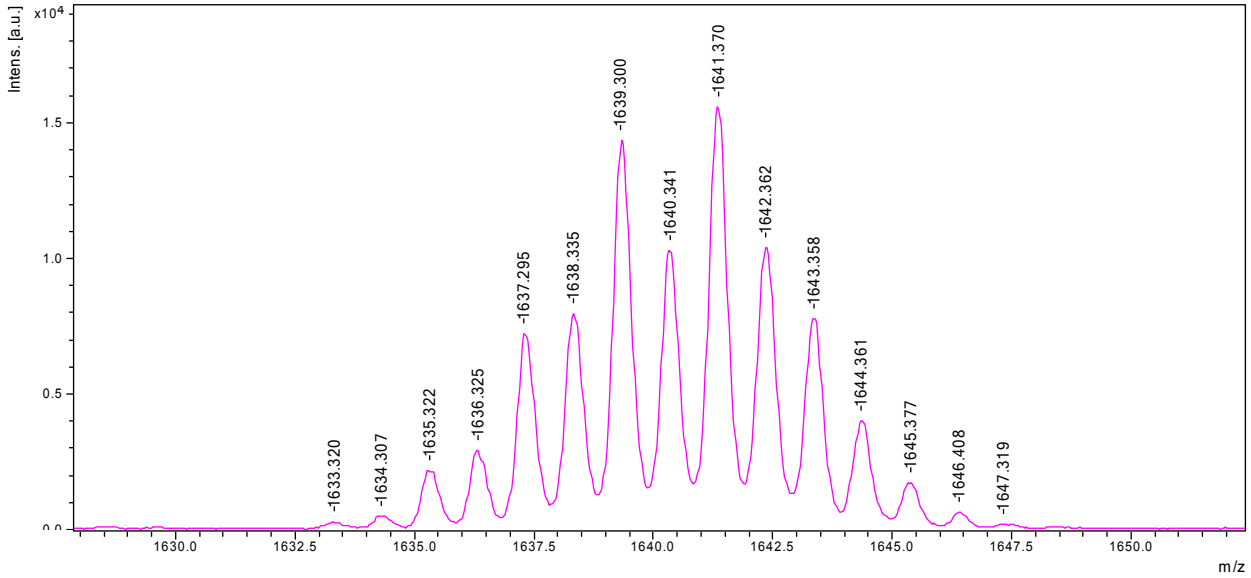

Fig. S10 MALDI-TOF-MS spectrum for purified **4**. MALDI-MS (*m*/*z*) found: 1641.37, calcd for $[M+H]^+$: 1641.54.

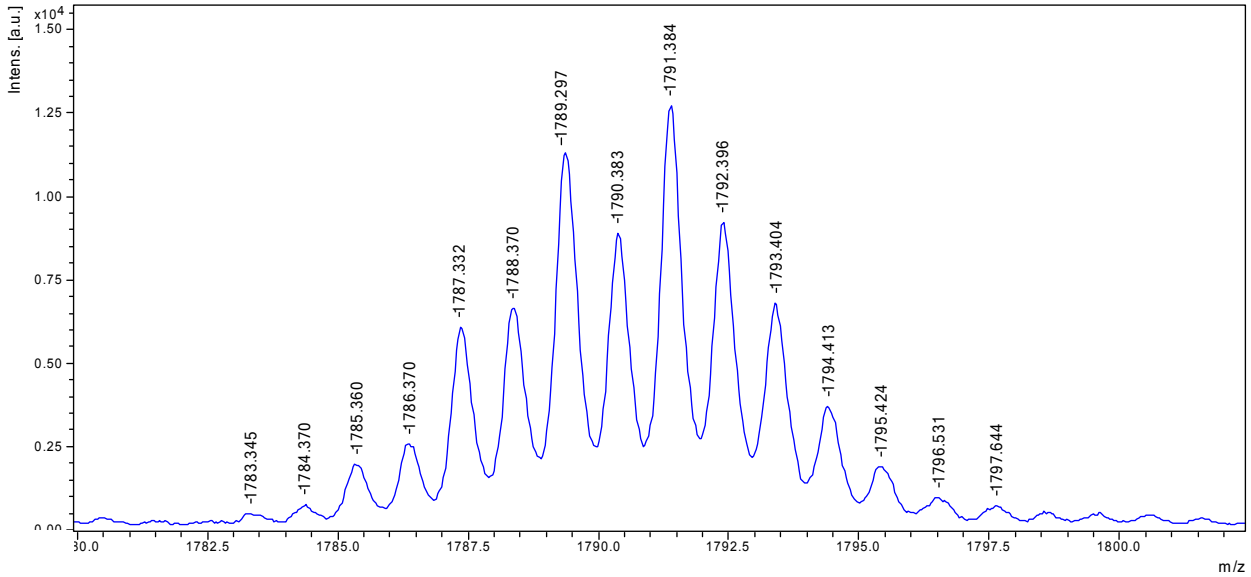

Fig. S11 MALDI-TOF-MS spectrum for purified **5**. MALDI-MS (*m*/*z*) found: 1791.38, calcd for $[M+H]^+$: 1791.60.

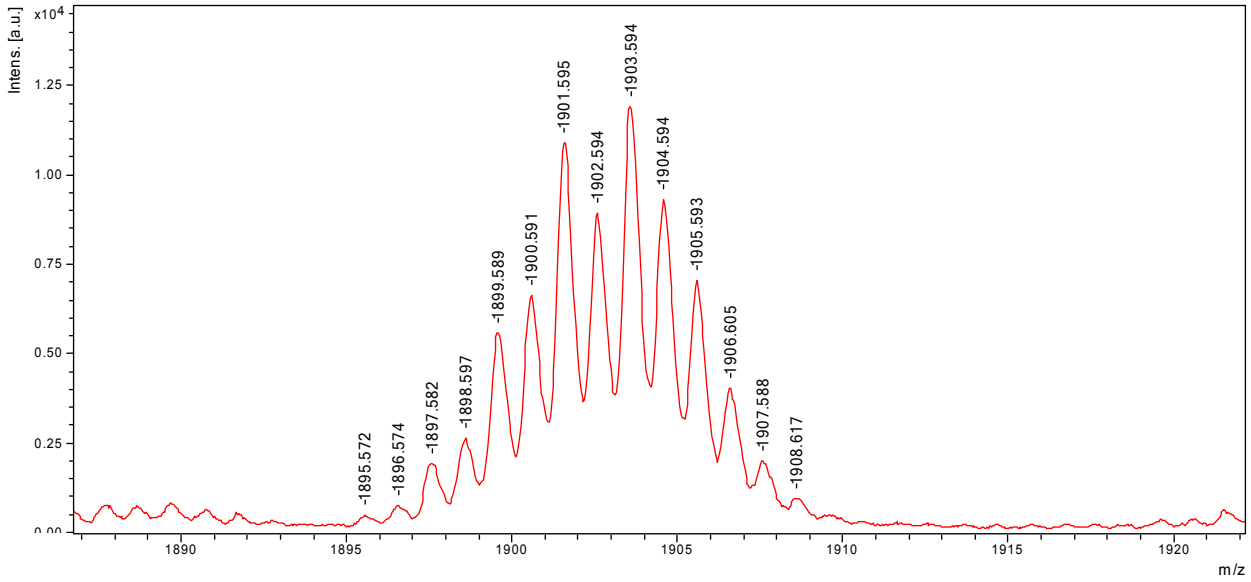

Fig. S12 MALDI-TOF-MS spectrum for purified **7**. MALDI-MS (*m*/*z*) found: 1903.59, calcd for $[M+H]^+$: 1903.68.

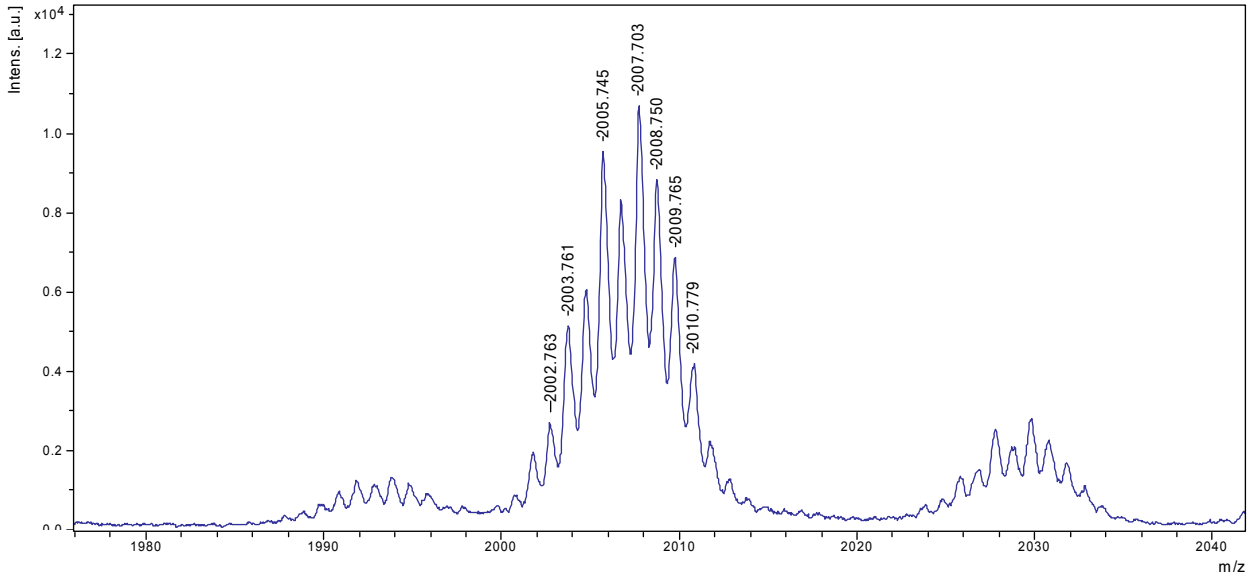

Fig. S13 MALDI-TOF-MS spectrum for **8**. MALDI-TOF-MS (*m*/*z*) found: 2007.70, calcd for $[M+H]^+$: 2007.74.

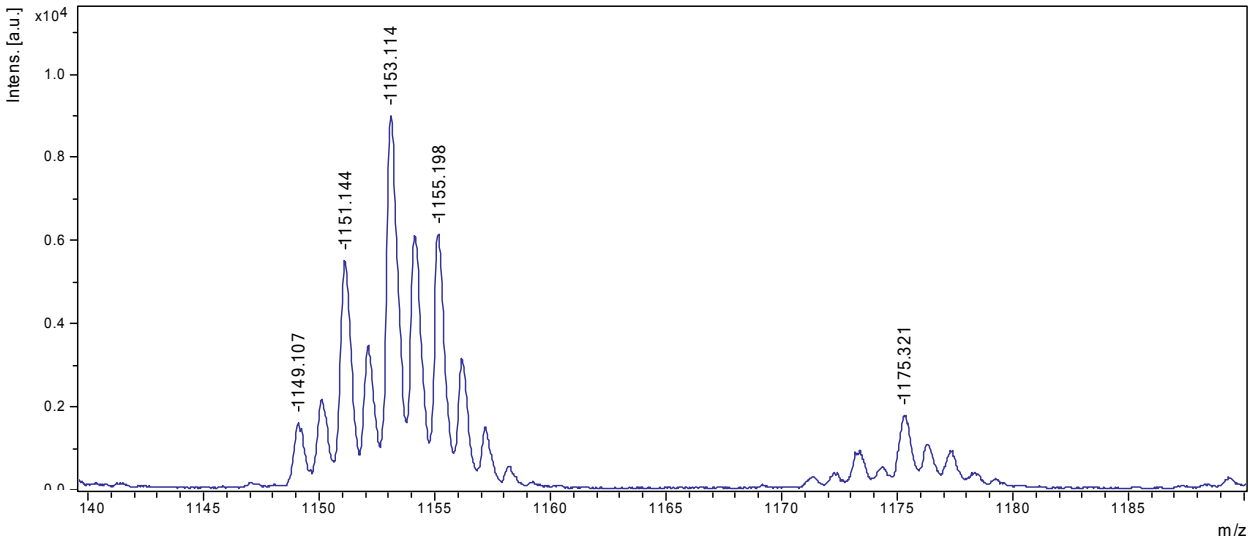

Fig. S14 MALDI-TOF-MS spectrum for **9**. MALDI-TOF-MS (*m*/*z*) found: 1153.11, calcd for $[M+H]^+$: 1153.45.

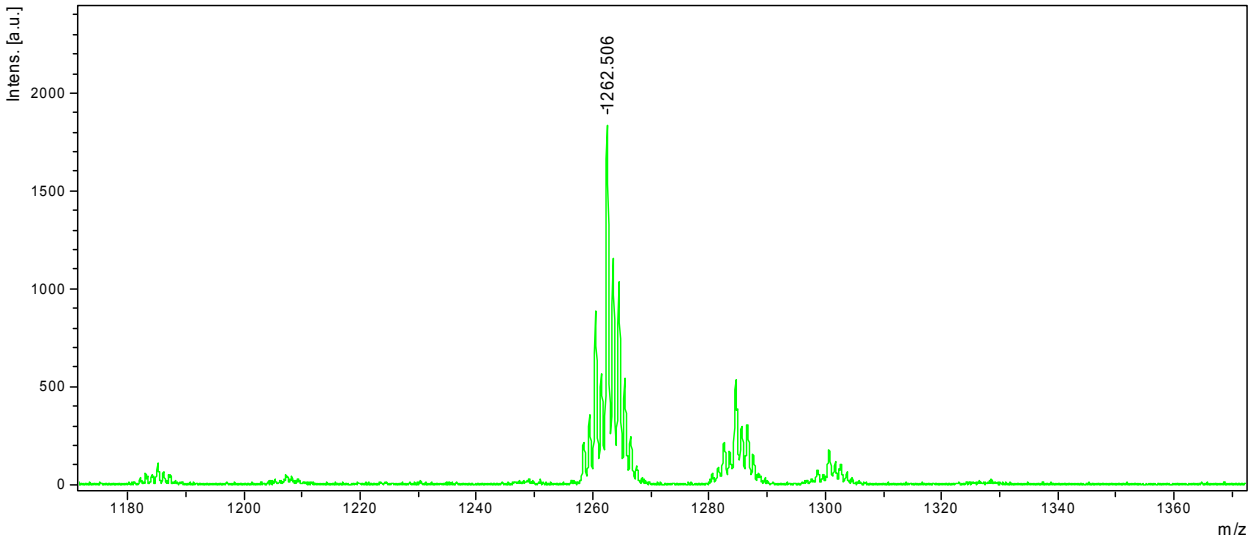

Fig. S15 MALDI-TOF-MS spectrum for **10**. MALDI-TOF-MS (*m*/*z*) found: 1262.51, calcd for $[M+H]^+$: 1264.46.

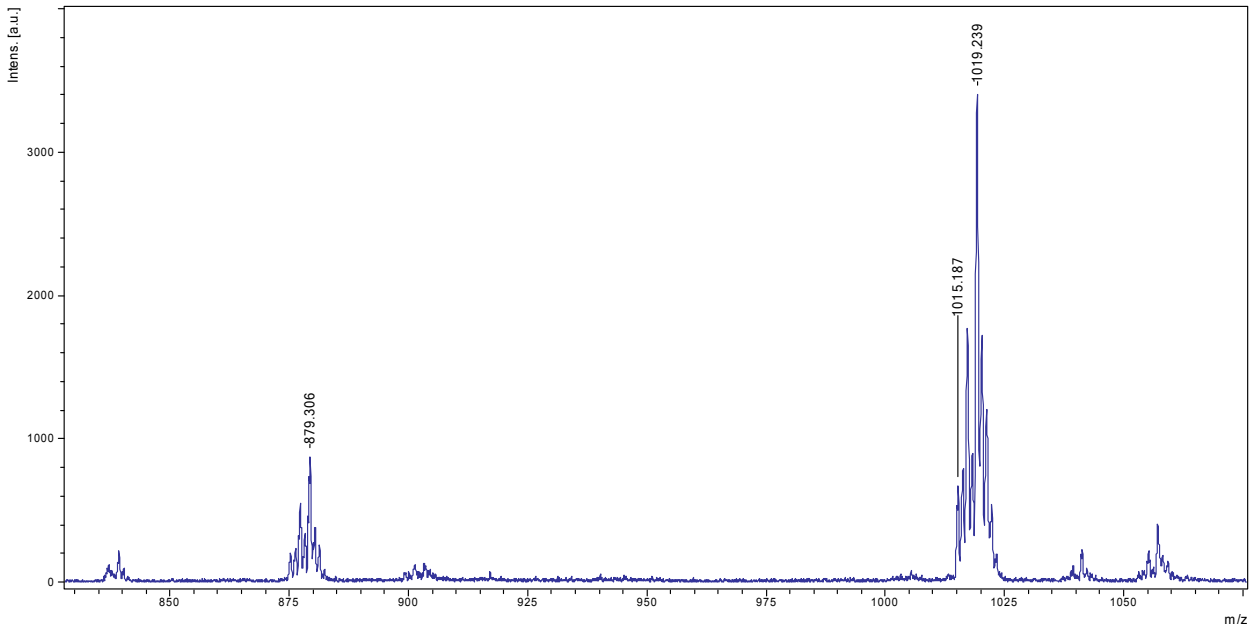

Fig. S16 MALDI-TOF-MS spectrum for **11**. MALDI-TOF-MS (*m*/*z*) found: 1019.24, calcd for $[M+H]^+$: 1019.26.

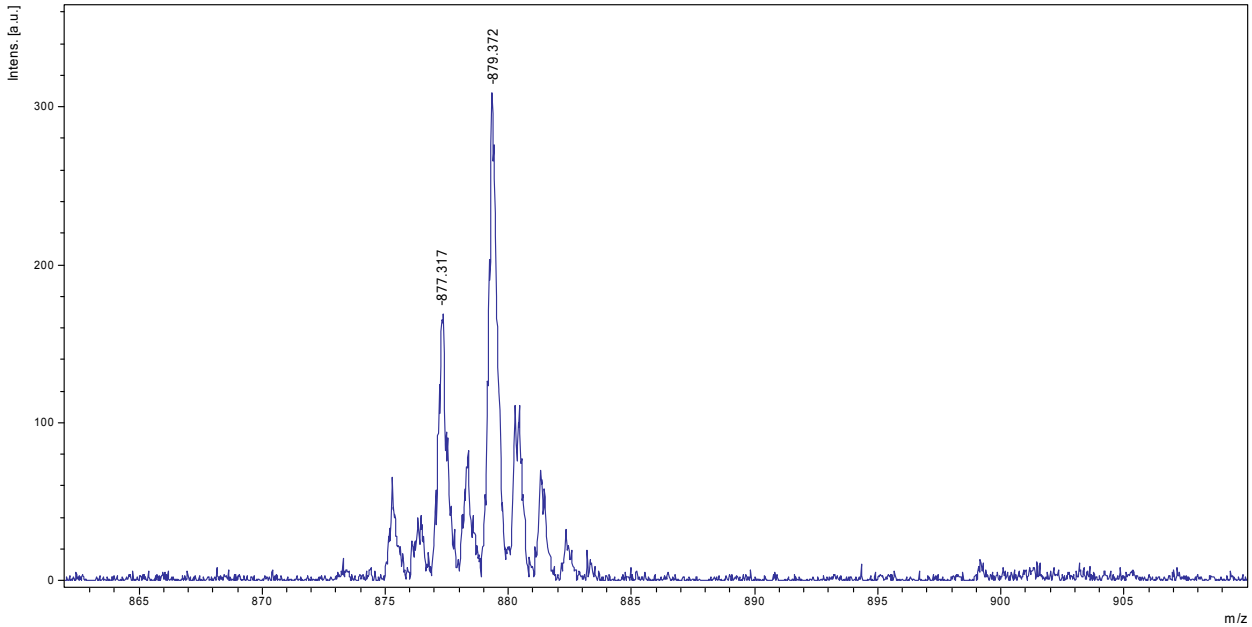

Fig. S17 MALDI-TOF-MS spectrum for **1'**. MALDI-TOF-MS (*m*/*z*) found: 879.37, calcd for $[M+H]^+$: 879.30.

10. GPx-like activity assay

The GPx-like catalytic activity of selenopeptides was assayed according to the literature method.[19a] A test solution (300 μL), containing 1 mM NADPH, 13.3 mM GSH and 13.3 U/mL glutathione reductase (GR) in a 100 mM phosphate/6 mM EDTA buffer solution at pH 7.35 was added with a phosphate buffer solution (570 μL) and a 830 μM selenopeptide solution (60 μL) in a 100 mM phosphate/6 mM EDTA buffer solution at pH

7.35, and the mixture was vortexed. The reaction was initiated by addition of a 3.6 mM H_2O_2 solution (70 μL) to the mixture solution. The reduction rate of H_2O_2 was monitored at 25 °C by absorption change at 340 nm due to consumption of NADPH, which was added in the assay solution to reduce GSSG (a counterproduct of the H_2O_2 reduction) to GSH in the presence of GR. The initial concentrations were $[GSH]_0$ = 4 mM, $[H_2O_2]_0$ = 0.25 mM, $[NADPH]_0$ = 0.3 mM, [GR] = 4 units/mL and [selenopeptide] = 50 μM. PhSeSePh was used as a control of the GPx activity in each set of measurements. The assay was repeated more than three times.

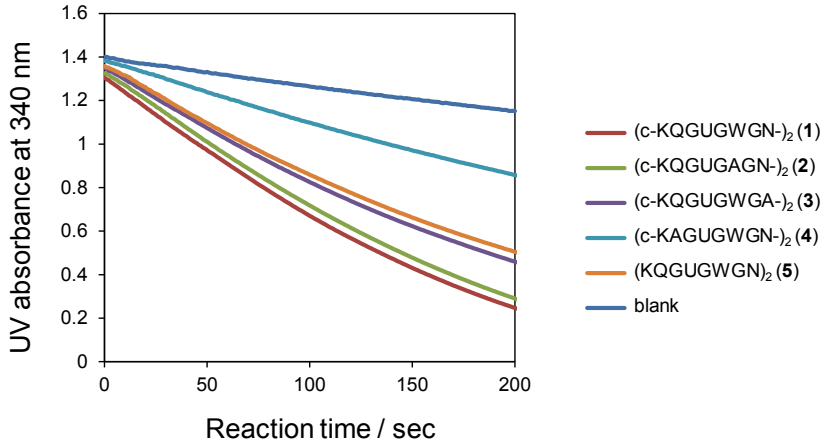

Fig. S18 GPx-like activity of cyclic and acyclic selenopeptides **1-5**.

Table S1. Summary of GPx-like activities.

Compounds	Initial velocity of H_2O_2 reduction (mM/s)	Relative activity
Blank	17 ± 1	1.0
1	89 ± 8	5.2 ± 0.5
2	84 ± 4	4.9 ± 0.2
3	72 ± 3	4.2 ± 0.2
4	38 ± 4	2.2 ± 0.2
5	70 ± 5	4.1 ± 0.3
6[a]	37 ± 3	2.2 ± 0.2
7	47 ± 3	2.7 ± 0.2
8	50 ± 3	2.9 ± 0.2
PhSeSePh	95 ± 5	5.5 ± 0.3

[a] Data are quoted from ref. 19a.

11. REMC/SAAP simulation

The replica-exchange Monte Carlo simulation using the SAAP3D force field (REMC/SAAP3D)[16] was carried out for linear selenopeptides, KQGUGWGN (a reduced selenol form of **5**), KQAUAWAN and KQPUPWPN, using four replicas at different temperatures at 300, 370, 440 or 510 K. At each MC step, one dihedral angle of the peptide molecule was changed randomly with a maximum displacement angle of ± 32 degrees. The total Monte Carlo step was set to 400 million, and 20,000 structures were output. In each 20,000 MC steps, the temperatures of replicas were attempted to exchange using Metropolis–Hastings algorithm. An extended structure was selected as the initial structure. After the REMC/SAAP simulation, the structure ensembles obtained at 300 K were statistically analyzed by measuring atomic distances between Sec and Gln, Trp or Asn.

Structure clustering was also performed based on the main-chain RMSD values using the k-means method.

Table S2. Number of the structures with specific atomic interactions obtained at 300 K by the REMC/SAAP simulation.[a]

Peptides	O(Q)···Se(U)	N(Q)···Se(U)	N(W)···Se(U)	O(N)···Se(U)	N(N)···Se(U)	Triad (U,Q,W)	Tetrad (U,Q,W,N)
KQGUGWGN	3394 (17)	2702 (14)	6305 (32)	1594 (8)	1531 (8)	2622 (13)	106 (1)
KQAUAWAN	6042 (30)	5558 (28)	1542 (8)	496 (2)	591 (3)	463 (2)	1 (0)
KQPUPWPN	6487 (32)	4649 (23)	6698 (33)	3998 (20)	4330 (22)	4888 (24)	1149 (6)

[a] The criterion of the distance was less than 5 Å. The number in parentheses is the probability (%) in the 20,000 output structures.

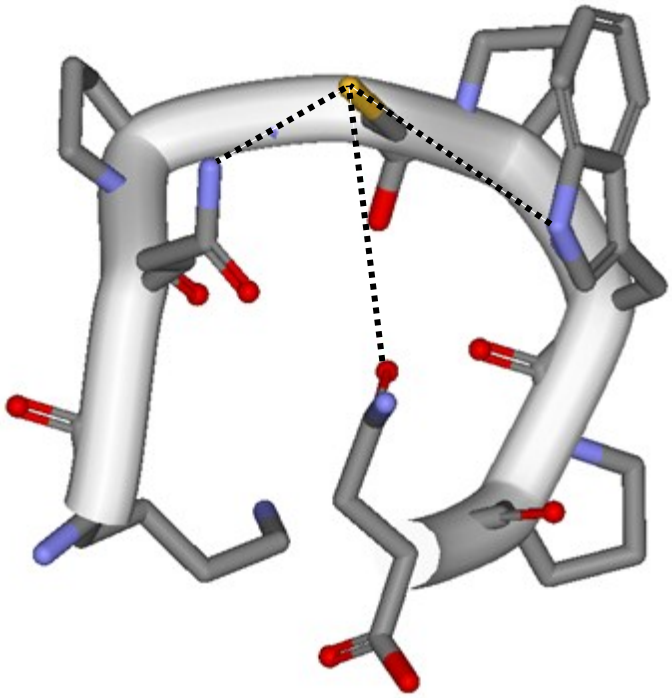

Fig. S19 The representative structure of the cluster (8%) obtained for KQPUPWPN at 300 K by the REMC/SAAP simulation. N(Q)···Se(U) = 4.16 Å, N(W)···Se(U) = 4.47 Å and O(N)···Se(U) = 4.99 Å.

12. References

S1 S. Shimodaira and M. Iwaoka, Arkivoc, 2017, **ii**, 260.