Convergent Chemical Synthesis of Histone H3 Protein for Sites-Specific Acetylation at Lys56 and Ubiquitination at Lys122

Yun-Kun Qi,[ab] Qiao-Qiao He,[a] Hua-Song Ai,[a] Jing Guo,[a] and Jia-Bin Li*[ab]

a. Tsinghua-Peking Center for Life Sciences, Department of Chemistry, Tsinghua University, Beijing 100084, China.

b. School of Life Sciences, University of Science and Technology of China, Hefei 230026, China.

Table of Contents

1. Reagents and materials

2-Chlorotrityl resin was purchased from Hecheng Technology (Tianjing, China). Nova-PEG Wang Resin was purchased from Novabiochem. Fmoc-amino acids were purchased from C S Bio, or GL Biochem (Shanghai, China). Tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl) was purchased from Adamas-beta Co., Ltd. (Shanghai, China). Glycyl 1-(2,4-dimethoxyphenyl)-2-mercaptoethyl auxiliary was purchased from Nantong PPTIDE Biotech ltd (Jiangsu, China). Other reagents and materials were purchased as previously described.[1] All reagents and solvents were purified when necessary.

2. HPLC, FPLC, mass spectrometry (MS)

The reversed phase HPLC (RP-HPLC) was performed on Shimadzu Prominence HPLC at room temperature. For peptide analysis, analytical Vydac C18 (4.6 × 150 mm, 5 μm particle size) and C4 (4.6 × 250 mm, 5 μm particle size) columns were used at a flow rate of 1.0 mL/min. For peptide purification, semi-preparative Vydac C18 (10 × 250 mm, 10 μm particle size) and C4 (10 × 250 mm, 10 μm particle size) columns were used at a flow rate of 3-4 mL/min. The UV absorption at 214 nm and 254 nm were monitored throughout the injections. Buffers for RP-HPLC: buffer A (0.1% TFA in CH_3CN) and buffer B (0.1% TFA in water). Both solvents were sonicated for 30 min before use.

FPLC was performed on an AKTA (GE Healthcare Life Science) with Size-exclusion chromatography Superdex 200 column (10/300 GL). All the buffers were filtered through 0.22 μm filter paper and sonicated for 10 min before use. Every injection was monitored at 280 nm and 214 nm.

The crude peptides and reaction products were characterized by normal ESI mass on LC/MS 2020 (SHIMADZU). The ultimate products were characterized by high-resolution ESI mass spectra and deconvoluted electrospray ionization mass spectra on Agilent 6210 Time of Flight Mass Spectrometer.

3. SDS-PAGE and Western blotting

For SDS-PAGE, samples were loaded onto 15-18% SDS-PAGE gels. Then the samples were electrophoresed for 10 min at 80 V and then 40 min at 200 V. The SDS–PAGE was stained with coomassie brilliant blue (CBB).

For western blotting, the samples were separated by SDS–PAGE gels and then transferred to PVDF membranes (Bio-Rad). The blocking and anibody incubations were conducted in Tris-buffer containing 5% (w/v) skimmed milk powder. For western blotting, two commercial primary antibodies were used: rabbit anti-H3 (ab1791, Abcam) and rabbit anti-ubiquitin (ab19247, Abcam). HRP-conjugated goat anti-rabbit IgG (Beyotime, Shanghai, China) were used as secondary antibody in immunoblotting. The chemiluminescence solution (Clarity™ ECL, Bio-Rad) was used for protein detection. The pictures of SDS–PAGE and immunoblotting were taken on ChemiDoc™ XRS+ system (Bio-Rad).

In transfer buffer (25 mM Tris-base, 192 mM Glycine, 20% methanol), the samples on the SDS–PAGE gel were transferred to the PVDF membrane for 80 mins under 300 mA at 4 °C. Then,

the PVDF membrane was blocked by Tris-buffer (20 mM Tris-base, 137 mM NaCl, 5% (w/v) skimmed milk powder, pH 7.6) under rotation for 2 h at room temperature. Rabbit anti-H3 or rabbit anti-ubiquitin primary antibody was diluted by 2000 fold using Tris-buffer (20 mM Tris-base, 137 mM NaCl, 5% (w/v) skimmed milk powder, pH 7.6). Then, the diluted primary antibody solution was added onto the PVDF membrane and incubated under rotation for 2 h at room temperature. Subsequently, the PVDF membrane was washed with wash buffer TBST (20 mM Tris-base, 137 mM NaCl, 0.1% (v/v) Tween-20, pH 7.6) three times (under rotation for about 10 mins per time) at room temperature. The anti-rabbit IgG secondary antibody was diluted by 2500 fold using Tris-buffer (20 mM Tris-base, 137 mM NaCl, 5% (w/v) skimmed milk powder, pH 7.6). Then, the diluted secondary antibody solution was added onto the above PVDF membrane and incubated under rotation for 1.5 h at room temperature. Subsequently, the PVDF membrane was washed with TBST buffer four times (under rotation for at least 10 mins per time) at room temperature. Then, the freshly prepared HRP chemiluminescent reagents were added onto the PVDF membrane. At last, the PVDF membrane was exposed and analyzed by ChemiDoc™ XRS + system (Bio-Rad).

4. Fmoc-based solid-phase peptide synthesis (Fmoc-SPPS)

4.1. General procedures for Fmoc-SPPS

The synthesis of peptide segments was performed on automated peptide synthesizer (CS136XT, C.S. Bio, Co.) as previously described.[1C,2]

Generally, 0.2 mmol scale of desired resin was swelled in 10 mL DMF/DCM (1:1, v/v) for about 3 h. Then, the first amino acid residue was coupled for twice (10 min and then 60 min), using a DMF solution of the Fmoc-amino acid (4 equiv.), HCTU (3.8 equiv.) and DIPEA (8 equiv.) (The Fmoc-amino acid was activated for about 10-15 s before adding to the resin). Generally, double coupling (10 min and then 60 min) is enough for most amino acids.

The Fmoc group was removed by the treatment of 20% (v/v) piperidine in DMF (twice: 5 min and then 10 min). After the coupling of the last amino acid residue, the resin was thoroughly dried and then treated with cocktail (TFA/phenol/water/thioanisole/EDT, 85:2.5:5:5:2.5) for about 2.5-3 h. Subsequently, the resin was completely removed by filtration and then washed twice with above cocktail. The pooled filtrate was then concentrated by pure argon. After the majority of the TFA was removed, the crude peptide was obtained through precipitating and washing with cold Et_2O (three to four times). The obtained crude peptides were further purified by semi-preparative RP-HPLC and then lyophilized. Each purified peptide segments could be obtained on at least tens-of-milligrams scale through a synthesis process (using 0.2 mmol resin within 3 days).

Peptide hydrazides **1**, **2**, **8**, **9**, and **17** were synthesized on hydrazine 2-chlorotrityl resin following the Fmoc-SPPS procedures. Especially, peptide 2 with site specific K56 acetylation could be readily assembled combining the use of commercially available Fmoc-Lys(Ac)-OH and standard Fmoc SPPS.

The synthesis of peptide **3** was critical, which should be conducted on Nova-PEG Wang resin with standard Fmoc SPPS.

Peptide **7** which contains glycyl auxiliary group at Lys122 site could be readily assembled through standard Fmoc SPPS on Nova-PEG Wang resin.

Peptide acid **18** was synthesized on 2-chlorotrityl resin following the Fmoc-SPPS.

4.2. The synthetic scheme for peptide 7

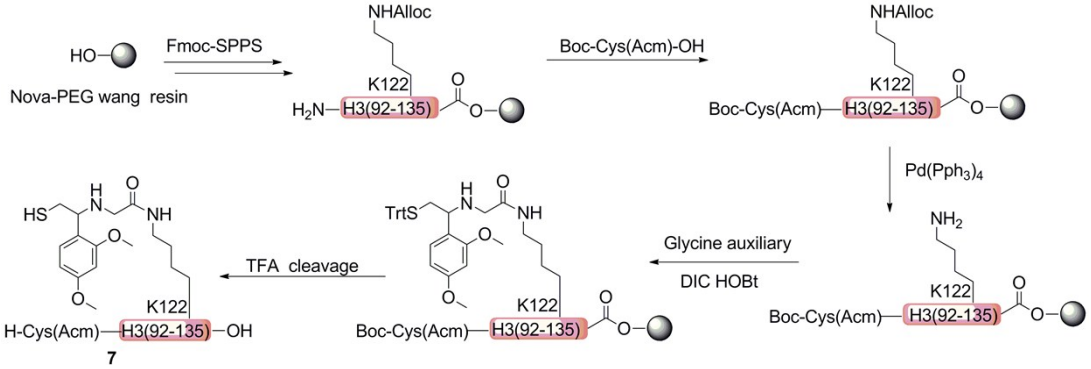


Figure S1. Synthetic scheme for peptide **7**. Glycyl 1-(2,4-dimethoxyphenyl)-2- mercaptoethyl auxiliary was commercially available and compatible with standard Fmoc SPPS. During the synthesis of peptide **7**, residue Lys122 was introduced through Fmoc-Lys(Alloc)-OH instead of Fmoc-Lys(Boc)-OH. After the introduction of Boc-Cys(Acm)-OH, the resin was treated with $Pd(PPh_3)_4/PhSiH_3$ to remove the Alloc group. Then the resin was coupled with glycine auxiliary using DIC/HOBt reaction system.

5. The preparation of peptide 12 from peptide 11

5.1 Removal of Auxiliary group

Peptide **11** (37.8mg, 2.7 μmol) was treated with 5 mL of TFA cocktail (TFA/water/EDT/ thioanisole, 95:2.5:1.25:1.25) for 3 h. Then, the reaction solution was concentrated by pure N_2. The reaction mixture was then precipitated with cold Et_2O and then centrifuged (3-4 times). Subsequently, the crude peptide **11**' was dissolved in H_2O/CH_3CN then lyophilized.

5.2 Removal of Acm group

Peptide **11**' (about 32 mg, 2.3 μmol, 1 equiv., final concentration 0.8 mM) was dissolved in 2.9 mL of 50% AcOH solution (AcOH:H_2O, 1:1, v/v). To the above solution, silver acetate (23 mg, 138 μmol, 60 equiv.) was added and the reaction mixture was stirred for about 11 h at room temperature. Then, the mixture was treated with 5.5 mL of DTT solution (1.0 M in 6.0 M Gn·HCl) and stirred for another 20 min at room temperature. After centrifugation, the crude product was purified by semi-preparative RP-HPLC and confirmed by ESI-MS.

After the removal of auxiliary and Acm groups, pure peptide **12** (20.6 mg, 1.5 μmol) was obtained with an overall isolated yield of 56% over two deprotection steps.

6. Desulfurization

The desulfurization was conducted as previously reported.[3] Peptide **5** (62.9 mg, 4.1 μmol, final concentration 0.28 mM) was dissolved in the 8.3 mL aqueous solution (pH 6.9) containing of 6.0 M Gn·HCl, 0.2 M Na_2HPO_4 and 500 mM TCEP. Then 579 μl of *t*BuSH and 5790 μl of 0.1 M VA-044 in aqueous solution (pH 6.9) containing of 6.0 M Gn·HCl and 0.2 M Na_2HPO_4 were added into the reaction. The reaction mixture was adjusted to pH 6.9 and then stirred for about 12 h at 37 °C. The desulfurization product was then purified by semi-preparative RP-HPLC, affording pure peptide **6** (42 mg, 2.8 μmol) with an isolated yield of 69%.

7. Protein expression and purification

The wild type yeast H3 was obtained as the methods described previously.[4] The histone H3 plasmid was transformed into *E. coli* BL21(DE3) cells. The *E. coli* BL21 (DE3) cells were then grown in 1 litre of Luria-Bertani media with Ampicillin (100 ug/ml) at 37 °C. When the OD_{600} reached about 0.8, 0.4 mM (final concentration) IPTG was added to induce the expression, at 37 °C for about 8-10 h. The cells were harvested by centrifugation (6000 rpm, 20 min, 4 °C) to get the cell precipitation. The cells were resuspended by lysis buffer (20 mM Tris, 150 mM NaCl, 1 mM DTT, 1mM PMSF, pH 8.0) and then lysed by sonication for 30 min. Then, the inclusion body was obtained by centrifugation (12000 rpm, 30 min, 4 °C). The inclusion body was washed by TW buffer (50 mM Tris-HCl , 100 mM NaCl, 1 mM Na-EDTA, 1 mM Benzamidine, 1% (v/v) Triton X-100, 5 mM DTT, pH 7.5) three times and then washed by wash buffer (TW buffer without Triton X-100) twice. After thorough wash, the precipitation was obtained through centrifugation, and then resolved by unfolding buffer (6 M Gn·HCl, 20 mM Tris-HCl, 10 mM DTT, pH 7.5). The insoluble impurities were removed through centrifugation (12000 rpm, 30 min, 4 °C; twice if necessary). The supernatant was filtered through 0.22 μm filter paper and purified by semi-preparative RP-HPLC. After lyophilization, desired histone H3 (20 mg, 1.3 μmol) was obtained.

8. Histones tetramer reconstitution

Generally, the tetramer $(H3\text{-}H4)_2$ was reconstituted as the methods described previously with minor modifications.[4]

Histones H3 and H4 were dissolved respectively in unfolding buffer (6 M Gn·HCl, 20 mM sodium acetate, 5 mM dithiothreitol, pH 4.5) at a concentration of 2 mg/mL. After 1 hour incubation at room temperature, equimolar ratios of H3 and H4 solution were mixed together in a tube. Then, the combined solution was diluted to a total final protein concentration of 1 mg/mL with unfolding buffer. Then, the solution was transferred into a dialysis bags and dialyzed against 600 mL of refolding buffer (2 M sodium chloride, 10 mM sodium acetate, 1 mM Na-EDTA, 5 mM 2-mercaptoethanol, pH 4.5) at 4 °C. The refolding buffer was changed every 12 hours. After 36 hours (three changes of refolding buffer, at 4 °C), the protein solution was concentrated to a final volume of 500-600 μL. The precipitated proteins were removed by centrifugation. Subsequently, the supernatant solution was further purified by size-exclusion chromatography (Superdex 200 10/300 GL column), using the refolding buffer as the elution buffer. Check the purity and stoichiometry of the fractions on an 18% SDS-PAGE. The fractions should always be kept at 4°C.

Through gel filtration chromatography, the refolded heterotetramer $(H3\text{-}H4)_2$ was isolated as a single mono-dispersed peak eluted at 14.5 mL (Figure 3A in the main text), which was considered as the characteristic retention volume for histone $(H3\text{-}H4)_2$. Meanwhile, through the same tetramer assembly strategy described above, $(H3K56Ac\text{-}H4)_2$ heterotetramer was also successfully reconstituted.

9. LC-MS/MS identification

The MS/MS identifications of H3K56Ac and H3K56AcK122Ub were conducted as previously described.[5]

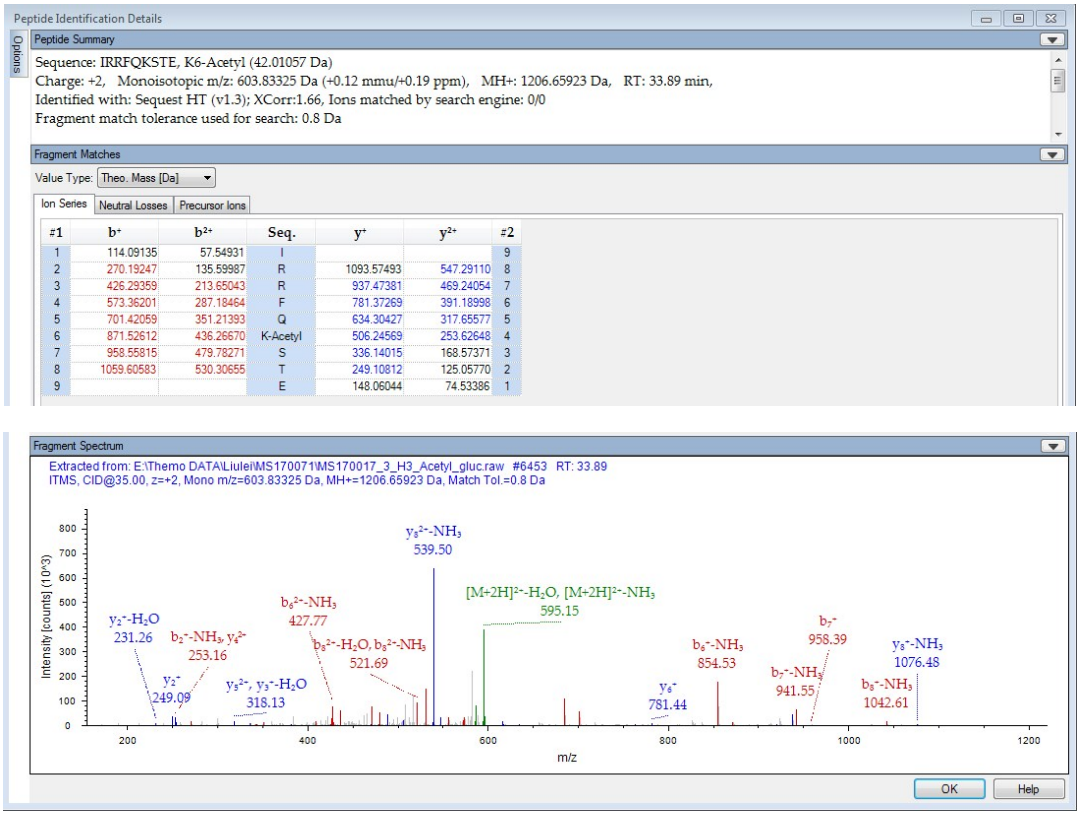
Purified H3K56Ac or H3K56AcK122Ub was reduced with 5 mM of dithiothreitol (DTT) and alkylated with 10 mM of iodoacetamide. Then, in-gel digestion was performed using sequencing Glu-C protease in 50 mM ammonium bicarbonate at 37 °C for about 12 h. Then resulted peptide fragments were extracted twice with aqueous solution (containing 50% acetonitrile (v/v) and 1% TFA (v/v)) for 30 min. The extraction solution was combined and then concentrated in a Speedvac. Then, the concentrated mixtures were separated by a 120 min gradient elution at a flow rate of 0.300 μL/min. The Thermo-Dionex Ultimate 3000 HPLC system, which was directly coupled with a Thermo LTQ-Orbitrap VELOS mass spectrometer (Thermo Fisher Scientific), was used for the LC-MS/MS analysis. The analytical C18 chromatography column was used for the HPLC. The buffers for LC-MS/MS analysis consisted of 0.1% formic acid. The LTQ-Orbitrap VELOS mass spectrometer was operated in the data-dependent acquisition mode and the Xcalibur 2.0.7 software was used. The experiment consisted of a single full-scan mass spectrum in the Orbitrap (300–1500 m/z, 60,000 resolution) followed by 20 data-dependent MS/MS scans in the ion trap at 35% normalized collision energy (CID). Then, MS/MS spectra from each LC-MS/MS experiment were searched against the histone and ubiquitin database, using Proteome Discoverer searching algorithm (Version 1.4).

The search criteria were as follows. Full Glu-C (cleavage at the C-terminal of residues Asp and Glu) specificity was needed. Carbamidomethylation was set as the fixed modification; oxidation (M), Acetylation (K), Ub (K) (Glu-C cleavage of Ub, STLHLVLRLRGG (-H_2O) = 1302.78846 Da) were set as variable modifications. Two missed cleavages were allowed. Precursor ion mass tolerance was 5 ppm for all MS data acquired in the Orbitrap mass analyzer. Fragment ion mass tolerance was 0.8 Da for all MS2 spectra. High confidence score filter (FDR <1%) was applied to select the “hit” peptide segments.

We thank the protein chemistry facility at the Center for Biomedical Analysis of Tsinghua University for sample analysis.

9.1. LC-MS/MS identification of H3K56Ac

The resulting modified polypeptide from H3K56Ac was digested by Glu-C protease to the specific piece: IRRFQK*STE (**fragment 1**，K* means acetylation at Lysine). The (LC-MS)/MS results of **fragment 1** were shown in Figure **S2**. The mass data were identical to that expected from the database.



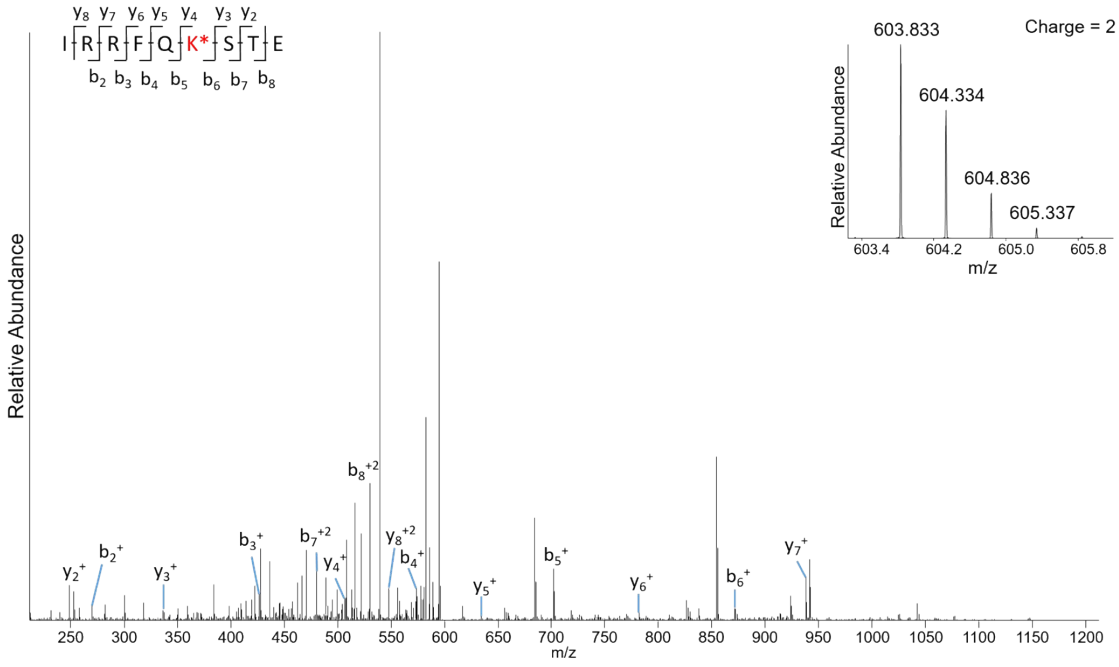


Figure S2. (LC-MS)/MS of **fragment 1** of H3K56Ac.

9.2. LC-MS/MS identification of H3K56AcK122Ub

The resulting modified polypeptide from H3K56AcK122Ub was digested by Glu-C protease to the two specific piece: IRRFQK*STE (**fragment 2**，K* means acetylation at Lysine) and DTNLAAIHAKRVTIQKK*D (**fragment 3**, K* means Ub segment at Lysine). The (LC-MS)/MS results of **fragment 2** were shown in Figure **S3**. The (LC-MS)/MS results of **fragment 3** were shown in Figure **S4**. The mass data were identical to that expected from the database.

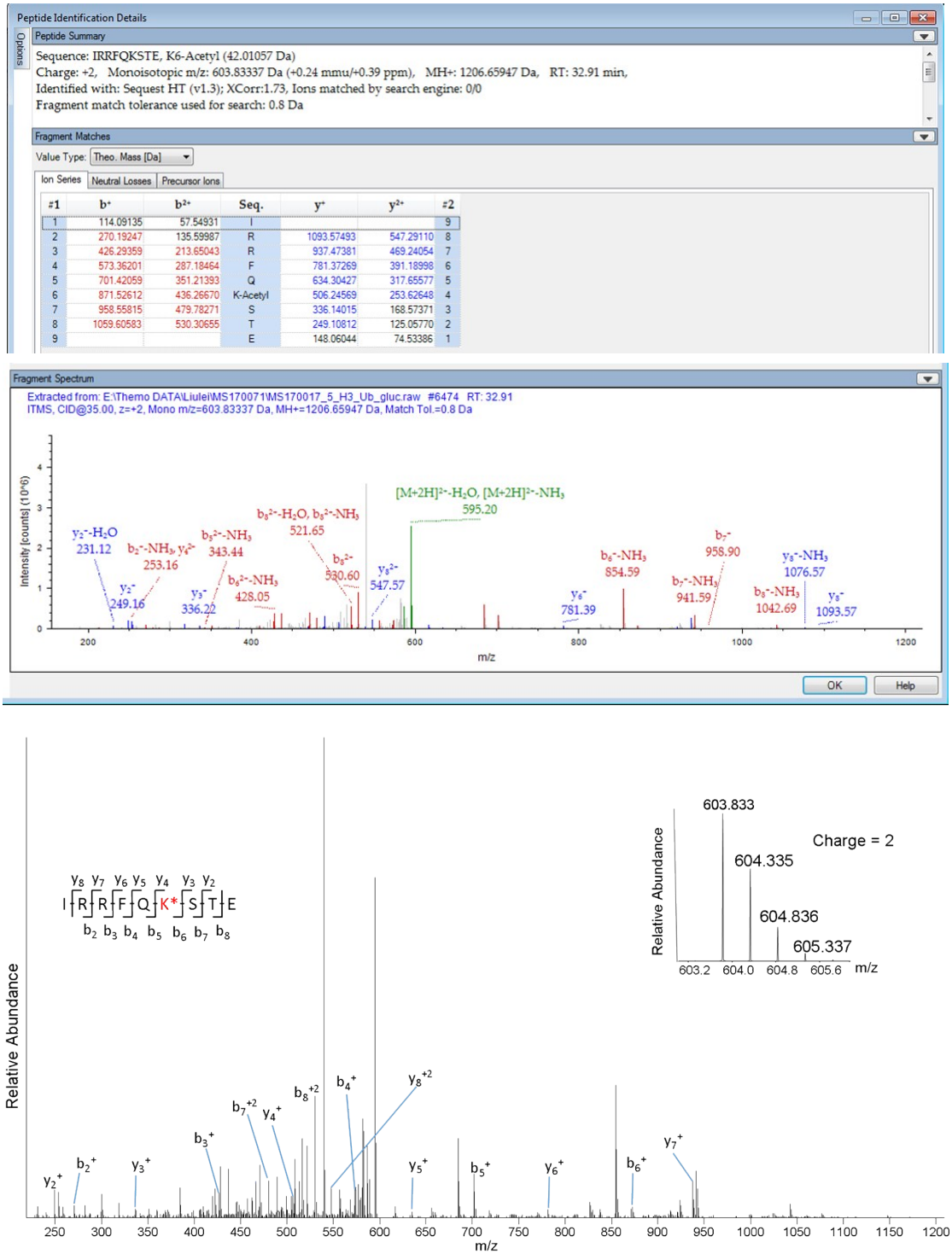
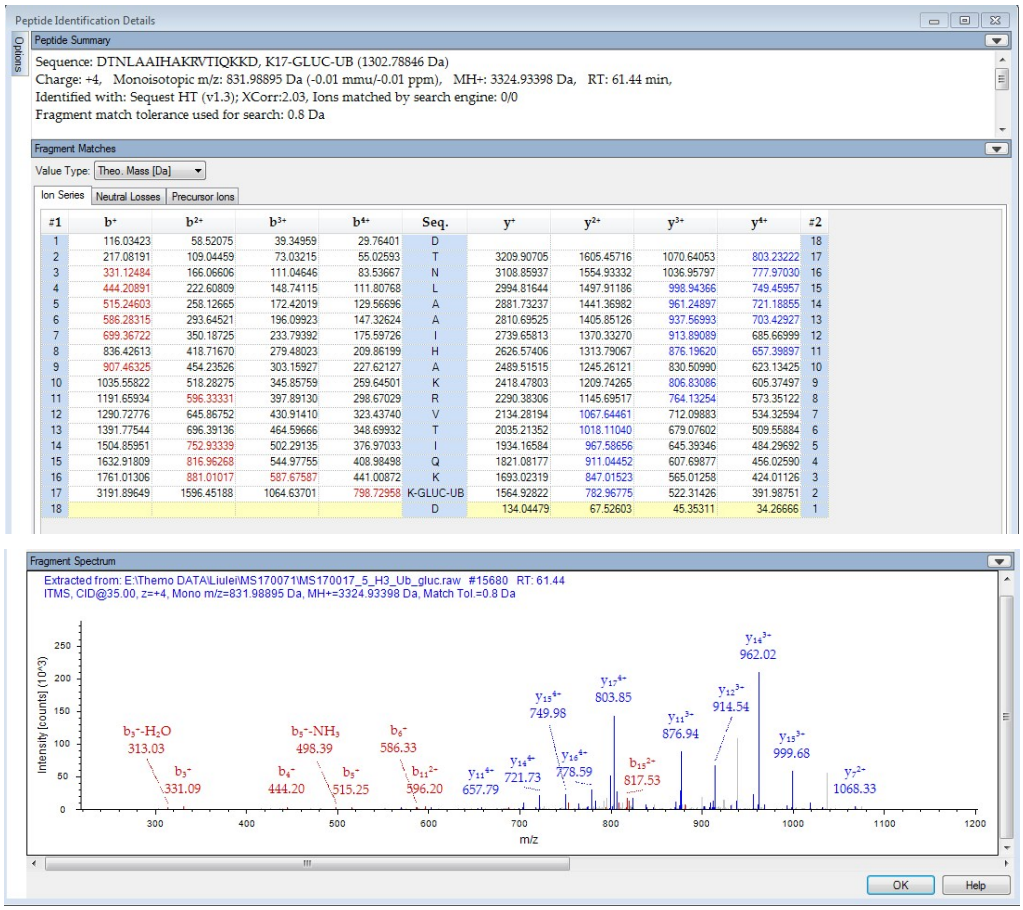


Figure S3. (LC-MS)/MS of **fragment 2** of H3K56AcK122Ub.



#1	b+	b2+	b3+	b4+	Seq.	y+	y2+	y3+	y4+	#2
1	116.03423	58.52075	39.34959	29.76401	D					18
2	217.08191	109.04459	73.03215	55.02593	T	3209.90705	1605.45716	1070.64053	803.23222	17
3	331.12484	166.06606	111.04646	83.53667	N	3108.85937	1554.93332	1036.95797	777.97030	16
4	444.20891	222.60809	148.74115	111.80768	L	2994.81644	1497.91186	998.94366	749.45957	15
5	515.24603	258.12665	172.42019	129.56696	A	2881.73237	1441.36982	961.24897	721.18855	14
6	586.28315	293.64521	196.09923	147.32624	A	2810.69525	1405.85126	937.56993	703.42927	13
7	699.36722	350.18725	233.79392	175.59726	I	2739.65813	1370.33270	913.89089	685.66999	12
8	836.42613	418.71670	279.48023	209.86199	H	2626.57406	1313.79067	876.19620	657.39897	11
9	907.46325	454.23526	303.15927	227.62127	A	2489.51515	1245.26121	830.50990	623.13425	10
10	1035.55822	518.28275	345.85759	259.64501	K	2418.47803	1209.74265	806.83086	605.37497	9
11	1191.65934	596.33331	397.89130	298.67029	R	2290.38306	1145.69517	764.13254	573.35122	8
12	1290.72776	645.86752	430.91410	323.43740	V	2134.28194	1067.64461	712.09883	534.32594	7
13	1391.77544	696.39136	464.59666	348.69932	T	2035.21352	1018.11040	679.07602	509.55884	6
14	1504.85951	752.93339	502.29135	376.97033	I	1934.16584	967.58656	645.39346	484.29692	5
15	1632.91809	816.96268	544.97755	408.98498	Q	1821.08177	911.04452	607.69877	456.02590	4
16	1761.01306	881.01017	587.67587	441.00872	K	1693.02319	847.01523	565.01258	424.01126	3
17	3191.89649	1596.45188	1064.63701	798.72958	K-GLUC-UB	1564.92822	782.96775	522.31426	391.98751	2
18					D	134.04479	67.52603	45.35311	34.26666	1

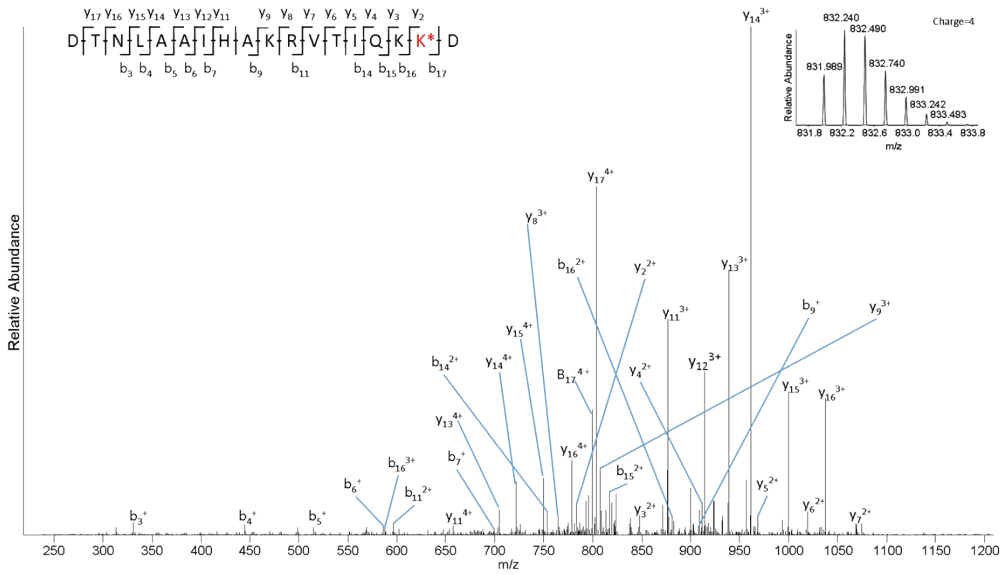


Figure S4. (LC-MS)/MS of **fragment 3** of H3K56AcK122Ub.

10. The amino acid sequence of yeast Ub

MQIFVKTLTG KTITLEVESS
DTIDNVKSKI QDKEGIPPDQ
QRLIFAGKQL EDGRTLSDYN
IQKESTLHLV LRLRGG
(76 amino acids)

Figure S5. The amino acid sequence of yeast Ub.

11. The synthetic scheme for native H4

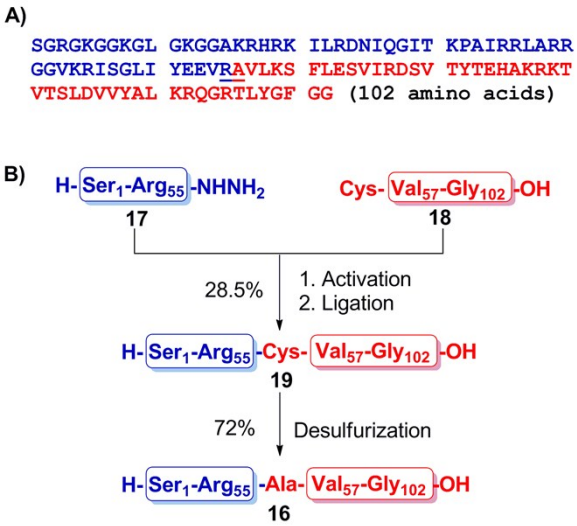


Figure S6. The synthetic scheme for native yeast H4. (A) The amino acid sequence of 102-residue H3. (B) Synthetic route.

12. Characterizations of peptides 1-19

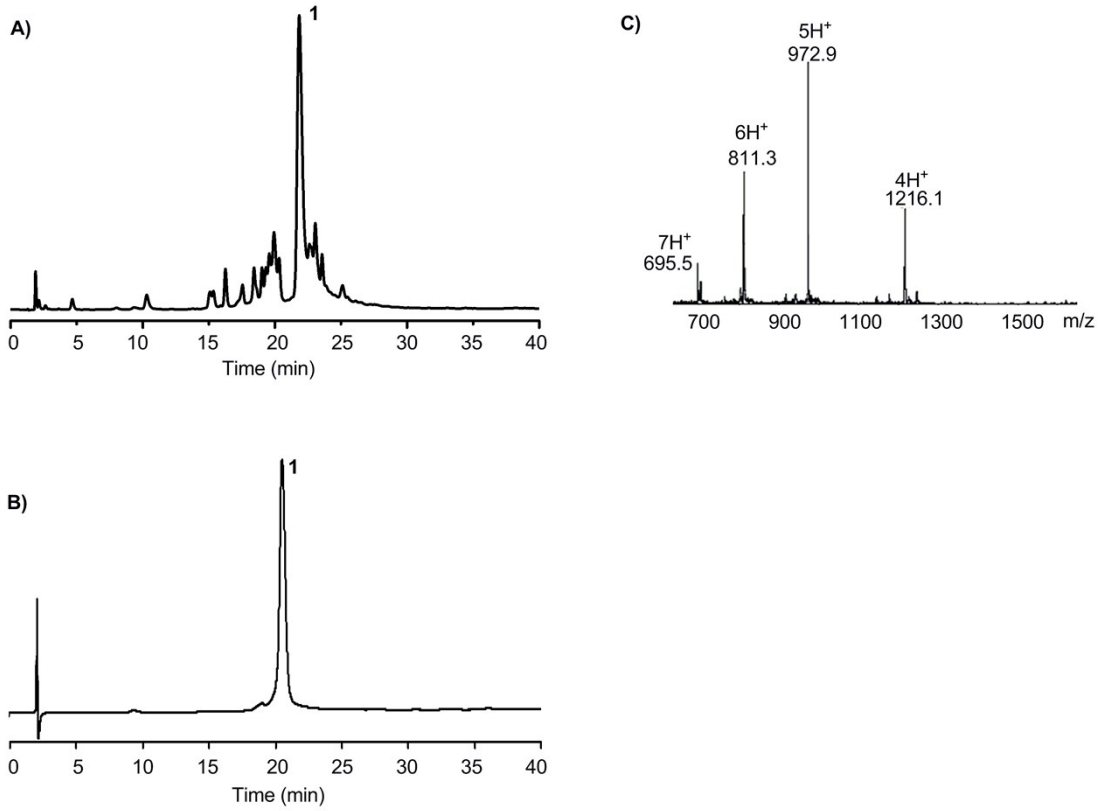


Figure S7. Characterization of synthetic peptide segment **1**. The peptide **1** was purified by HPLC with an isolated yield of 29.1%. A) Analytical HPLC chromatogram (λ = 214 nm) of crude peptide **1**. B) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **1**. HPLC condition: a linear gradient of 5%-30% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 38 min (5% for 2 min, then 5%-30% for 38 min) on a Vydac C18 column. C) ESI-MS spectrum of purified peptide **1.** The spectrum gave an observed mass of 4860.8 Da (calculated 4859.7 Da, average isotopes).

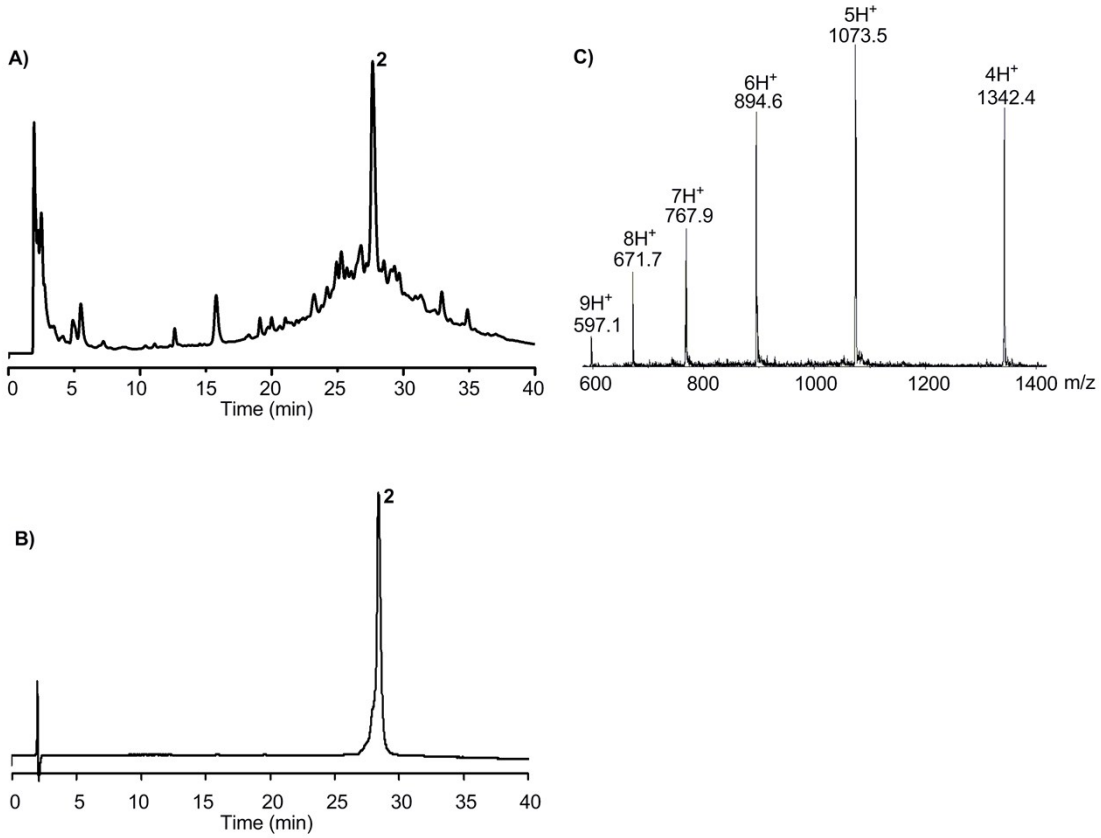


Figure S8. Characterization of synthetic peptide segment **2**. The peptide **2** was purified by HPLC with an isolated yield of 9.2%. A) Analytical HPLC chromatogram (λ = 214 nm) of crude peptide **2**. B) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **2**. HPLC condition: a linear gradient of 25%-55% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 38 min (5% for 2 min, then 25%-55% for 38 min) on a Vydac C18 column. C) ESI-MS spectrum of purified peptide **2.** The spectrum gave an observed mass of 5364.8 Da (calculated 5364.3 Da, average isotopes).

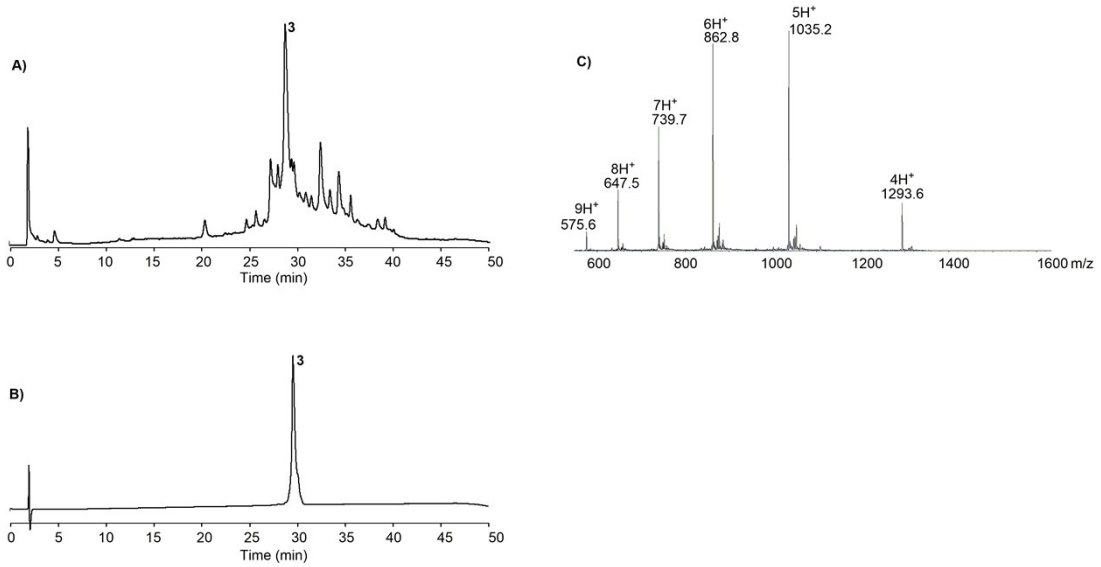


Figure S9. Characterization of peptide **3.** The peptide **3** was purified by HPLC with an isolated

yield of 16.3%. A) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **3**. HPLC condition: a linear gradient of 20%-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 38 min (5% for 2 min, then 20%-50% for 38 min) on a Vydac C18 column. B) ESI-MS spectrum of purified peptide **3**. The spectrum gave an observed mass of 5171.1 Da (calculated 5171.0 Da, average isotopes).

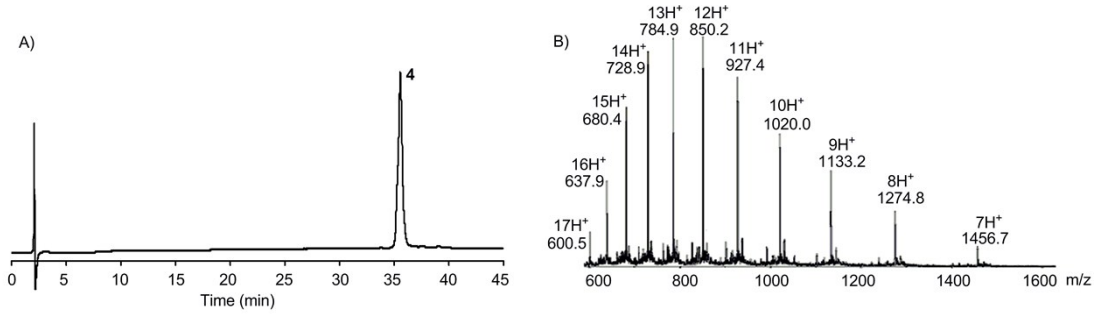


Figure S10. Characterization of peptide **4**. The peptide **4** was purified by HPLC with an isolated yield of 42.3%. A) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **4**. HPLC condition: a linear gradient of 5%-55% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 5%-55% for 40 min) on a Vydac C18 column. B) ESI-MS spectrum of purified peptide **4**. The spectrum gave an observed mass of 10190.5 Da (calculated 10192.0 Da, average isotopes).

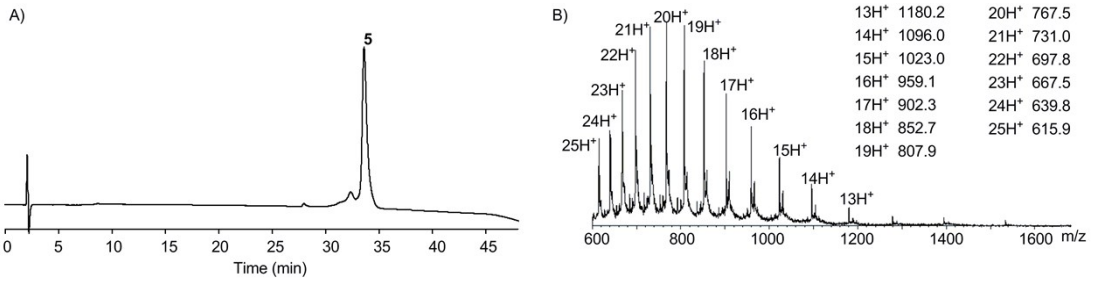


Figure S11. Characterization of peptide **5**. The peptide **5** was purified by HPLC with an isolated yield of 52.1%. A) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **5**. HPLC condition: a linear gradient of 5%-55% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-60% for 38 min) on a Vydac C18 column. B) ESI-MS spectrum of purified peptide **5**. The spectrum gave an observed mass of 15329.6 Da (calculated 15331.0 Da, average isotopes).

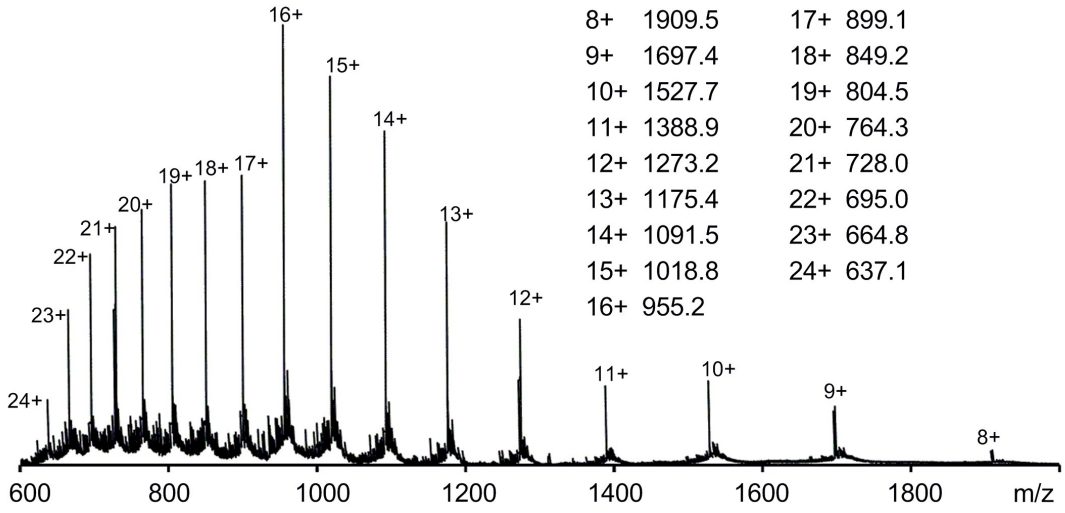


Figure S12. Characterization of peptide **6**. HPLC condition of Figure 1D in the main text: a linear gradient of 20%-75% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-75% for 30 min) on a Vydac C4 (4.6 × 250 mm) column. The ESI-MS spectrum of purified peptide **6**. The spectrum gave an observed mass of 15267.0 Da (calculated 15266.8.0 Da, average isotopes).

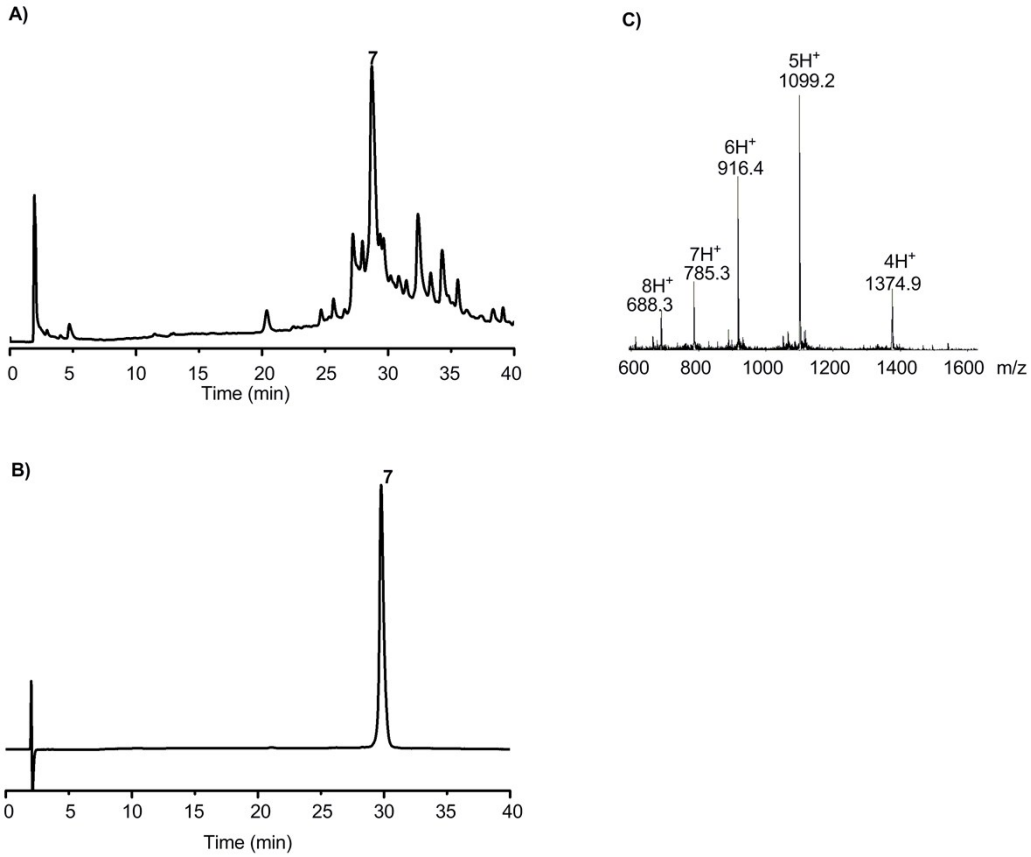


Figure S13. Characterization of synthetic peptide segment **7**. The peptide **7** was purified by HPLC with an isolated yield of 9%. A) Analytical HPLC chromatogram (λ = 214 nm) of crude peptide **7**. B) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **7**. HPLC condition: a linear

gradient of 20%-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 38 min (5% for 2 min, then 20%-50% for 38 min) on a Vydac C18 column. C) ESI-MS spectrum of purified peptide **7.** The spectrum gave an observed mass of 5495.0 Da (calculated 5495.4 Da, average isotopes).

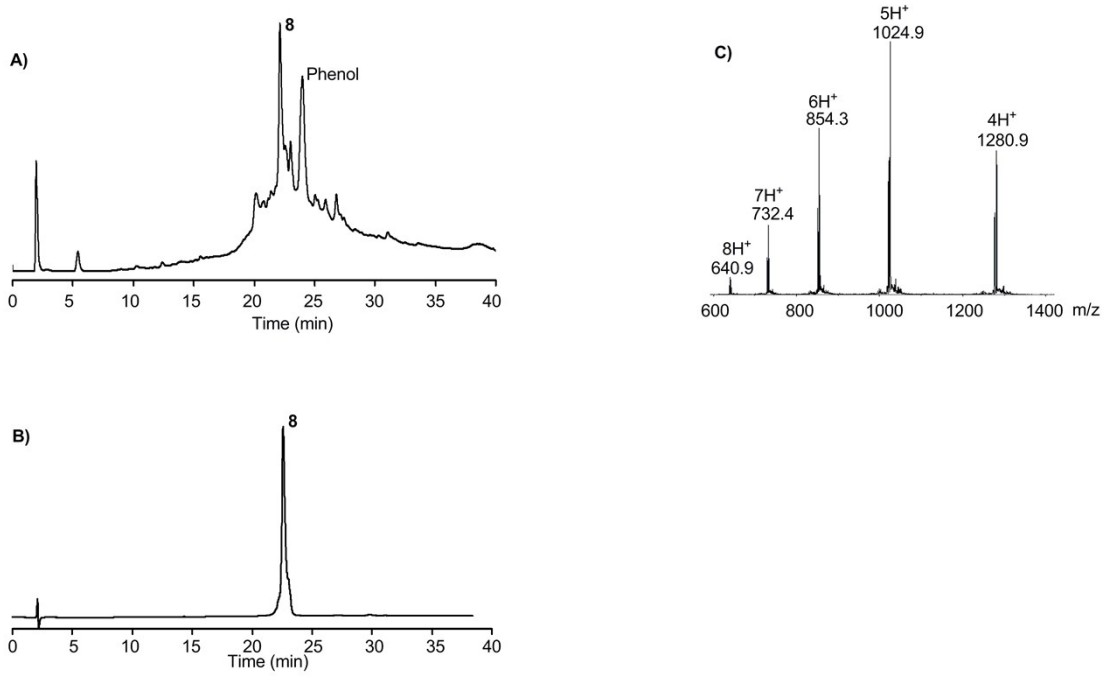


Figure S14. Characterization of synthetic peptide segment **8**. The peptide **8** was purified by HPLC with an isolated yield of 25%. A) Analytical HPLC chromatogram (λ = 214 nm) of crude peptide **8**. B) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **8**. HPLC condition: a linear gradient of 15%-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (15% for 2 min, then 15%-50% for 30 min) on a Vydac C18 column. C) ESI-MS spectrum of purified peptide **8.** The spectrum gave an observed mass of 5119.6 Da (calculated 5119.9 Da, average isotopes).

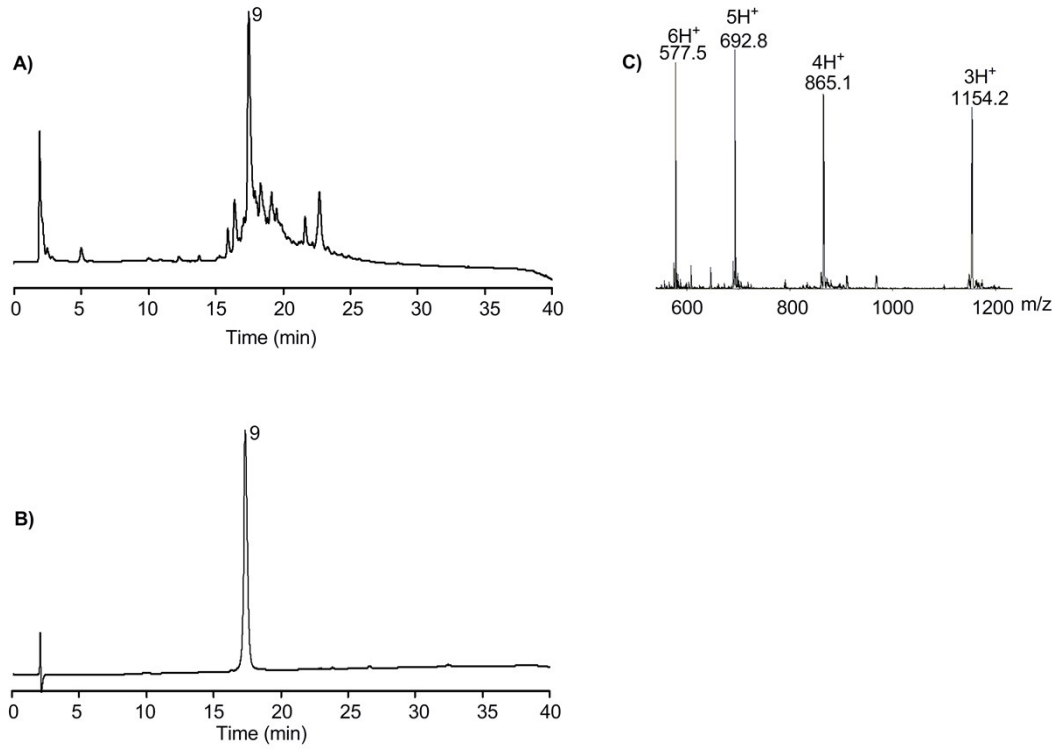


Figure S15. Characterization of synthetic peptide segment **9**. The peptide **9** was purified by HPLC with an isolated yield of 33%. A) Analytical HPLC chromatogram (λ = 214 nm) of crude peptide **9**. B) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **9**. HPLC condition: a linear gradient of 15%-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (15% for 2 min, then 15%-50% for 30 min) on a Vydac C18 column. C) ESI-MS spectrum of purified peptide **9.** The spectrum gave an observed mass of 3457.8 Da (calculated 3457.9 Da, average isotopes).

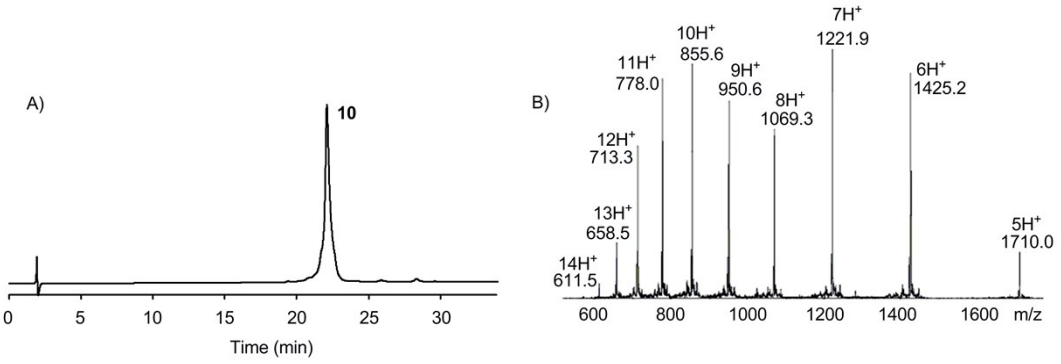


Figure S16. Characterization of synthetic peptide **10**. The peptide **10** was purified by HPLC with an isolated yield of 66%. A) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **10**. HPLC condition: a linear gradient of 20%-45% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-45% for 30 min) on a Vydac C18 column. B) ESI-MS spectrum of purified peptide **10.** The spectrum gave an observed mass of 8546.4 Da (calculated 8545.8 Da, average isotopes).

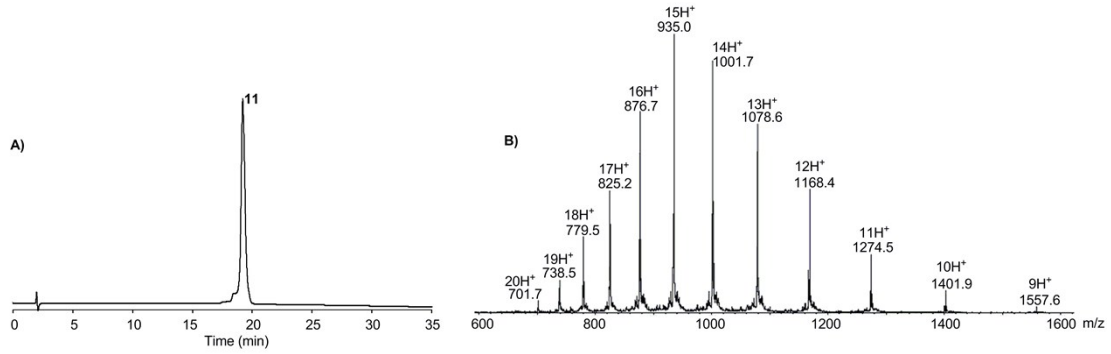


Figure S17. Characterization of synthetic peptide segment **11**. Peptide **11** is the key intermediate, which bears one glycyl auxiliary at Lys122 and the N terminal Acm group at Cys91. The peptide **11** was purified by HPLC with an isolated yield of 39%. A) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **11**. HPLC condition: a linear gradient of 25%-55% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 25%-55% for 30 min) on a Vydac C18 column. B) ESI-MS spectrum of purified peptide **11.** The spectrum gave an observed mass of 14010.3 Da (calculated 14009.2 Da, average isotopes).

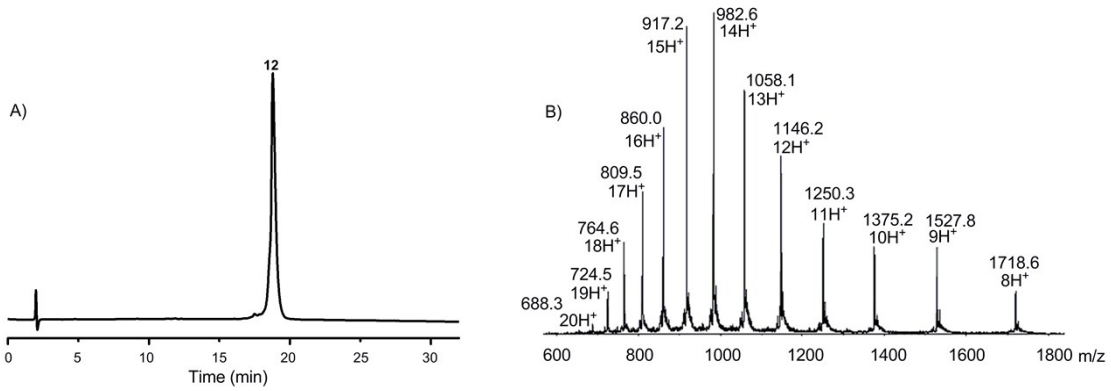


Figure S18. Characterization of synthetic peptide **12**. The peptide **12** was purified by HPLC with an isolated yield of 56% over two deprotection steps (auxiliary removal and Acm deprotection.). A) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **12**. HPLC condition: a linear gradient of 25%-55% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 25%-55% for 30 min) on a Vydac C18 column. B) ESI-MS spectrum of purified peptide **12.** The spectrum gave an observed mass of 13743.0 Da (calculated 13741.8 Da, average isotopes).

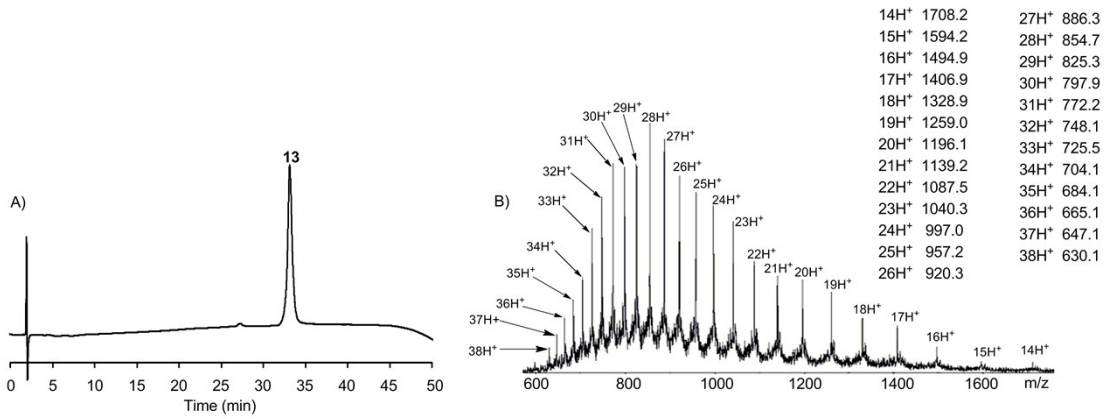


Figure S19. Characterization of synthetic peptide **13**. The peptide **13** was purified by HPLC with an isolated yield of 46%. A) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **13**. HPLC condition: a linear gradient of 20%-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 38 min (5% for 2 min, then 20%-60% for 38 min) on a Vydac C18 column. B) ESI-MS spectrum of purified peptide **13.** The spectrum gave an observed mass of 23902.6 Da (calculated 23901.8 Da, average isotopes).

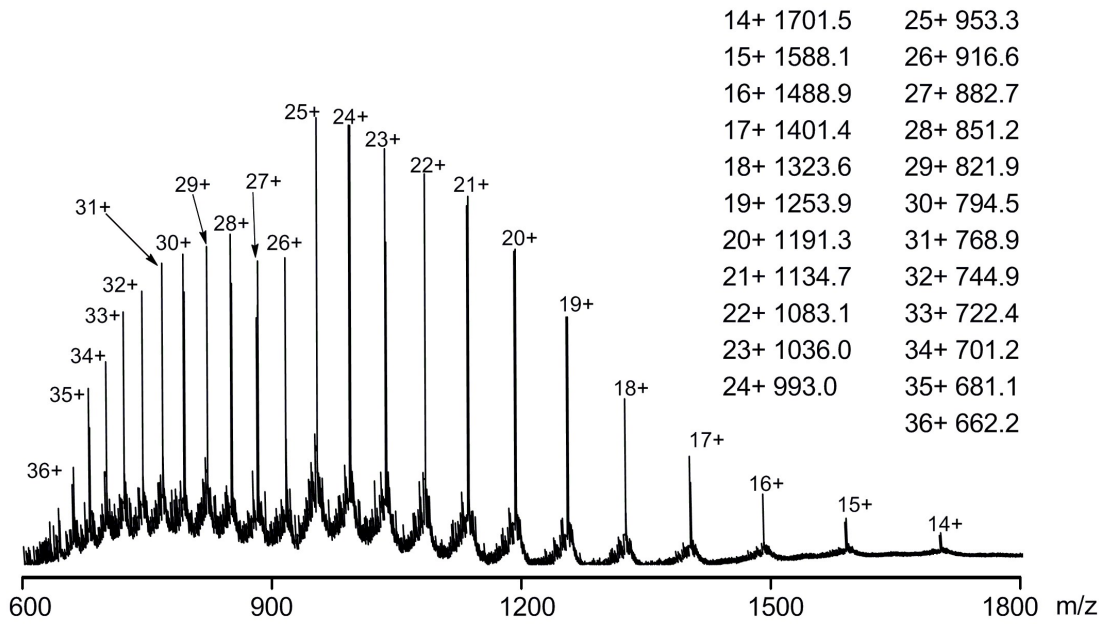


Figure S20. ESI-MS spectrum of purified peptide **14.** The spectrum gave an observed mass of 23806.0 Da (calculated 23805.6 Da, average isotopes).

HPLC condition of Figure 2C in the main text: a linear gradient of 20%-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 40 min (5% for 2 min, then 20%-60% for 38 min) on a Vydac C18 column.

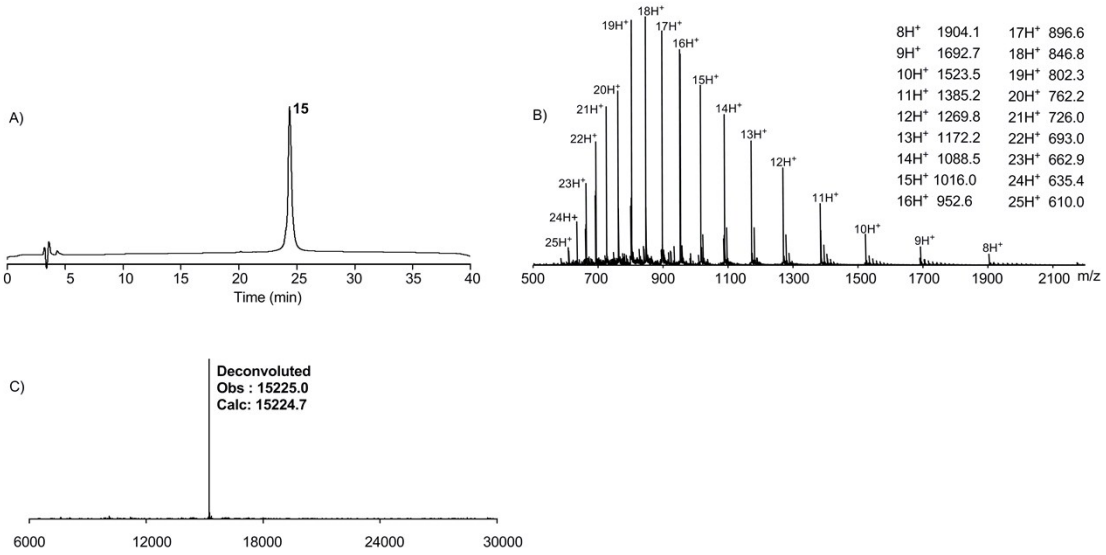


Figure S21. Characterization of protein **15** (native histone H3). A) Analytical HPLC chromatogram (λ = 214 nm) of purified protein **15**. HPLC condition: a linear gradient of 20%-75% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-75% for 30 min) on a Vydac C4 (4.6 × 250 mm) column. B) The ESI-MS spectrum of purified **15**. C) The deconvoluted electrospray ionization mass spectrum of **15**. The spectra gave an observed mass of 15225.0 Da (calculated 15224.7 Da, average isotopes).

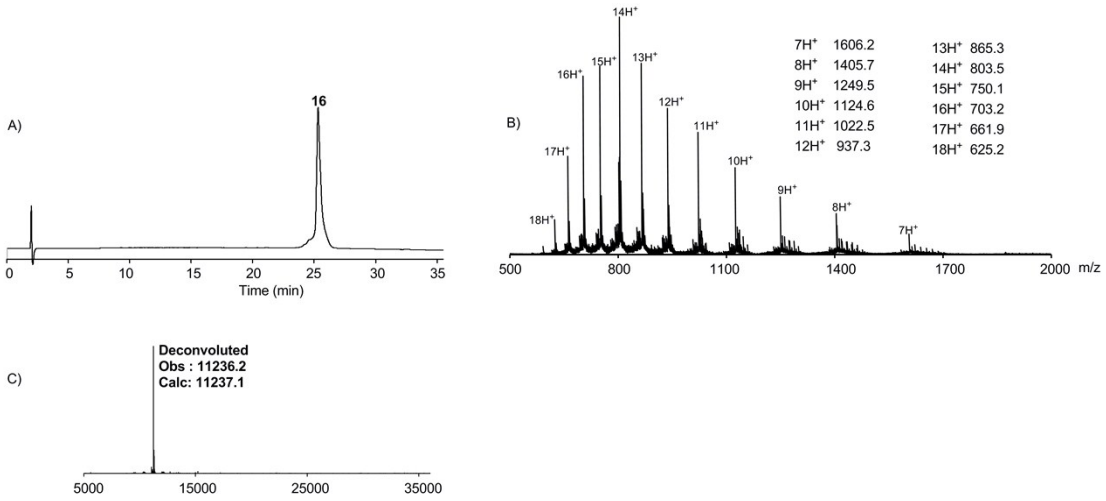


Figure S22. Characterization of protein **16**. The protein **16** was purified by HPLC with an isolated yield of 72%. A) Analytical HPLC chromatogram (λ = 214 nm) of purified **16**. HPLC condition: a linear gradient of 20%-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-60% for 30 min) on a Vydac C18 column. B) The ESI-MS spectrum of purified **16**. C) The deconvoluted electrospray ionization mass spectrum of **16**. The spectra gave an observed mass of 11236.2 Da (calculated 11237.1 Da, average isotopes).

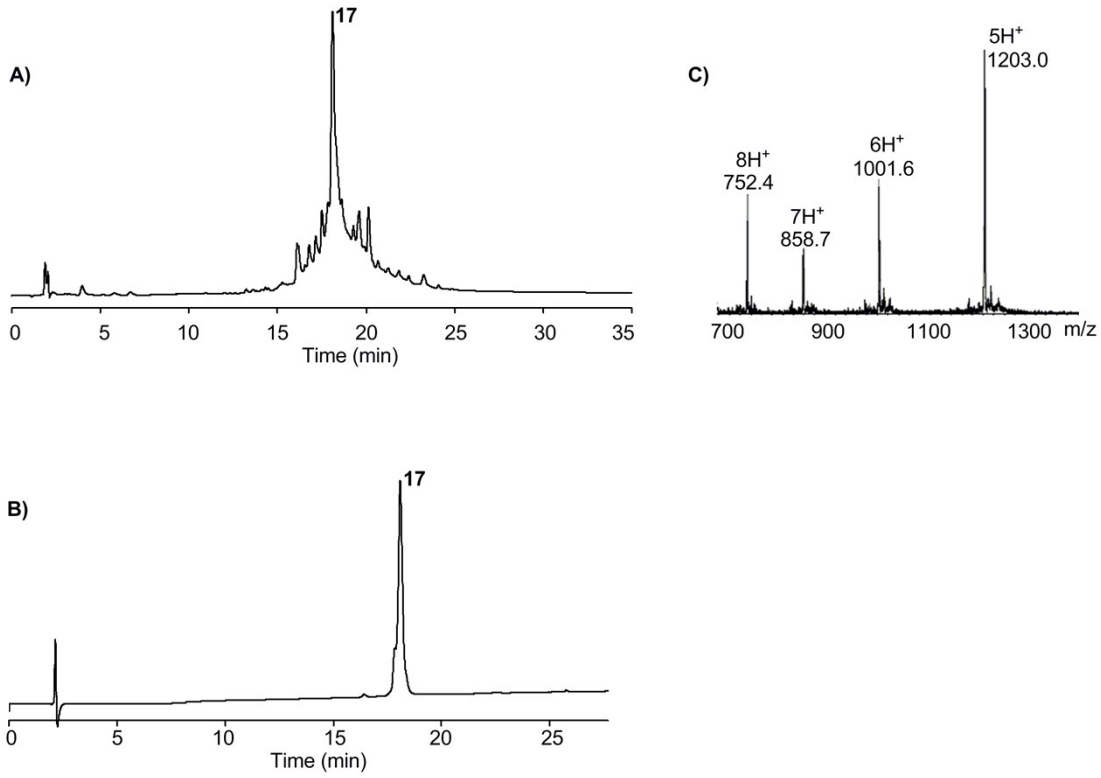


Figure S23. Characterization of synthetic peptide **17**. The peptide **17** was purified by HPLC with an isolated yield of 36%. A) Analytical HPLC chromatogram (λ = 214 nm) of crude peptide **17**. B) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **17**. HPLC condition: a linear gradient of 10%-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 10%-60% for 30 min) on a Vydac C18 column. C) ESI-MS spectrum of purified peptide **17.** The spectrum gave an observed mass of 6007.2 Da (calculated 6007.1 Da, average isotopes).

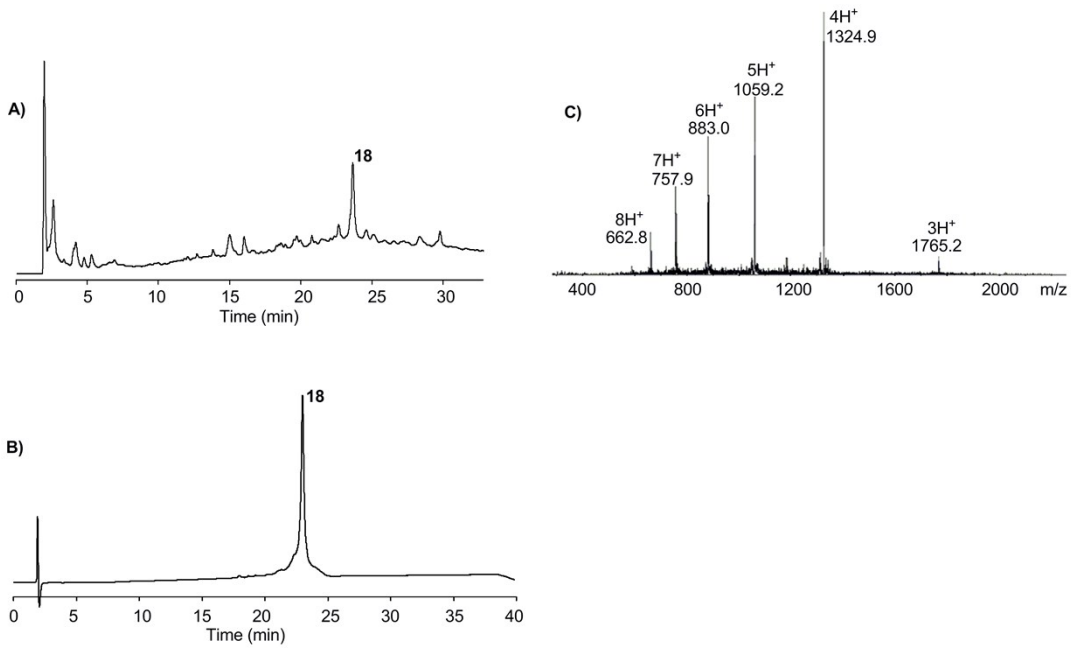


Figure S24. Characterization of synthetic peptide **18**. The peptide **18** was purified by HPLC with an isolated yield of 19%. A) Analytical HPLC chromatogram (λ = 214 nm) of crude peptide **18**. B) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **18**. HPLC condition: a linear gradient of 25%-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 25%-50% for 30 min) on a Vydac C18 column. C) ESI-MS spectrum of purified peptide **18.** The spectrum gave an observed mass of 5294.0 Da (calculated 5294.1 Da, average isotopes).

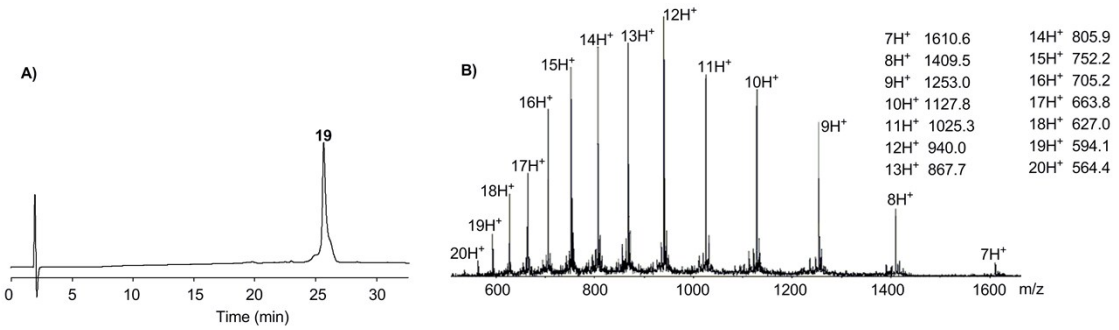


Figure S25. Characterization of synthetic peptide **19**. The peptide **19** was purified by HPLC with an isolated yield of 29%. A) Analytical HPLC chromatogram (λ = 214 nm) of purified peptide **19**. HPLC condition: a linear gradient of 20%-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-60% for 30 min) on a Vydac C18 column. B) ESI-MS spectrum of purified peptide **19.** The spectrum gave an observed mass of 11268.0 Da (calculated 11269.2 Da, average isotopes).

13. General procedure for the native chemical ligation of peptide hydrazids

Native chemical ligation of peptide hydrazides was conducted as previously described.[1,2]

The hydrazide peptide (1 equiv., final concentration 1-3 mM) was dissolved in the aqueous buffer containing 6 M guanidine hydrochloride (Gn·HCl) and 200 mM NaH_2PO_4 (pH = 3.0). Then the solution was cooled to approximately -10 °C to -13 °C in an ice-salt bath. Then, a solution of 200 mM $NaNO_2$ (7 equiv., dissolved in the same aqueous buffer as above) was added dropwise to activate hydrazide peptide. Then, the reaction mixture was incubated for 30 min (at -10 °C to -13 °C) to fully convent the acyl hydrazide to the acyl azide. Then, a solution of 200 mM MPAA (70 equiv.) dissolved in the aqueous buffer containing 6 M guanidine hydrochloride (Gn·HCl) and 200 mM Na_2HPO_4 (pH = 7.0) was added into the reaction mixture. The reaction mixture was taken out of the ice-salt bath and stirred at room temperature for 3 min. Then, the N-terminal Cys peptide (1-1.2 equiv.) was added into the reaction mixture. Then, the pH value of reaction was slowly adjusted to 6.5-6.9 with aqueous NaOH solution (2 M). Then, the reaction was monitored by analytic RP-HPLC. After the completion of the ligation, the reaction solution was reduced by a solution of 200 mM TCEP (equal volumes to the reaction mixture). The ligation product was then purified by semi-preparative RP-HPLC and characterized by ESI-MS.

Auxiliary-mediated ligation of peptide hydrazide between peptide **7** and peptide **10** was conducted as the normal native chemical ligation of peptide hydrazides.

13.1. Ligation of peptide 1 with peptide 2

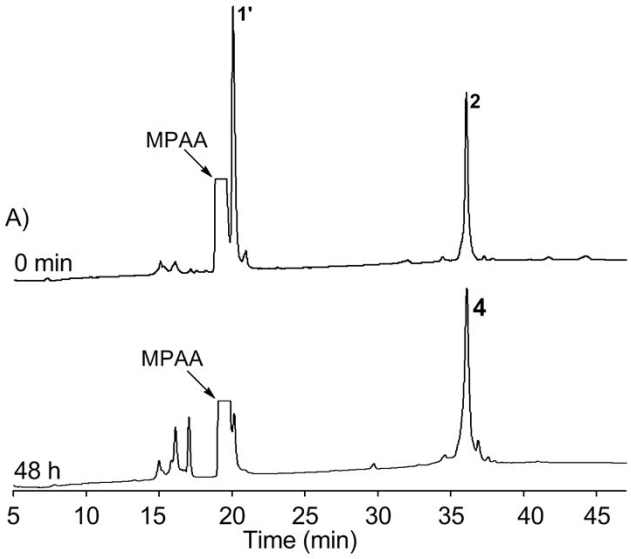


Figure S26. The analytical RP-HPLC traces (λ = 214 nm) for the ligation between peptide **1** and peptide **2**. 0 min and 48 h after the addition of peptide **2**. The peak marked with **1′** corresponds to the peptide thioester generated from peptide hydrazide **1**. HPLC condition: a linear gradient of 5%-55% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 38 min (5% for 2 min, then 5%-55% for 38 min) on a Vydac C18 column.

First, peptide **1** (12 μmol, 1.5 equiv.) was converted to the thioester by in situ $NaNO_2$ activation and subsequential thiolysis. Then, peptide **2** (8 μmol, 1.0 equiv.) was added into the reaction solution. Under continuous stirring, the pH value of the ligation solution was adjusted to 6.9 to initiate the NCL reaction. Analytical RP-HPLC was used to monitor the reaction process.

It should be mentioned that during ligation of peptides **1** and **2**, the product **4** has almost the same retention time with **2**. Moreover, this ligation was conducted at Val46-Cys47 that was considered as sterically hindered ligation site. Hence, excess peptide **1** (1.5 equivalents of **1** compared to peptide **2**), higher pH value (about pH 6.9), higher reaction temperature (37 °C), and more ligation time (about 40-50 hours) with gentle stirring were necessary for the ligation of peptides **1** and **2**.

13.2. Ligation of peptide 3 with peptide 4

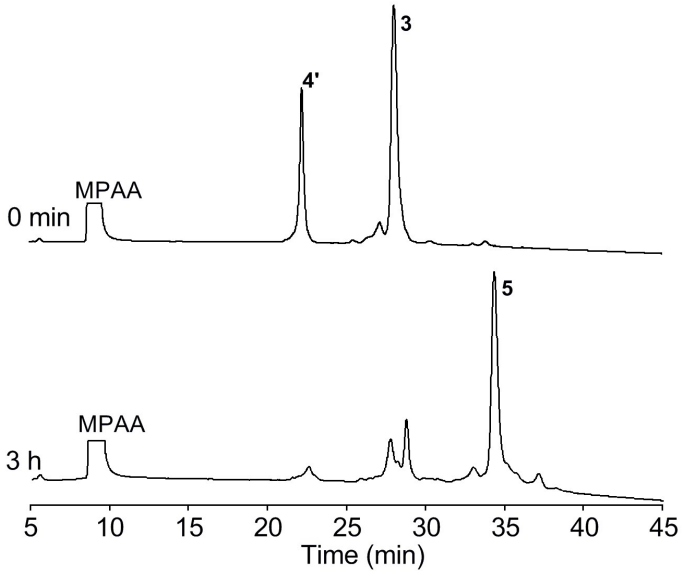


Figure S27. The analytical RP-HPLC traces (λ = 214 nm) for the ligation between peptide **4** and peptide **3**. 0 min and 3 h after the addition of peptide **3**. The peak marked with **4′** corresponds to the peptide thioester generated from peptide hydrazide **4**. HPLC condition: a linear gradient of 20%-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 38 min (5% for 2 min, then 20%-60% for 38 min) on a Vydac C18 column.

13.3. Ligation of peptide 8 with peptide 9

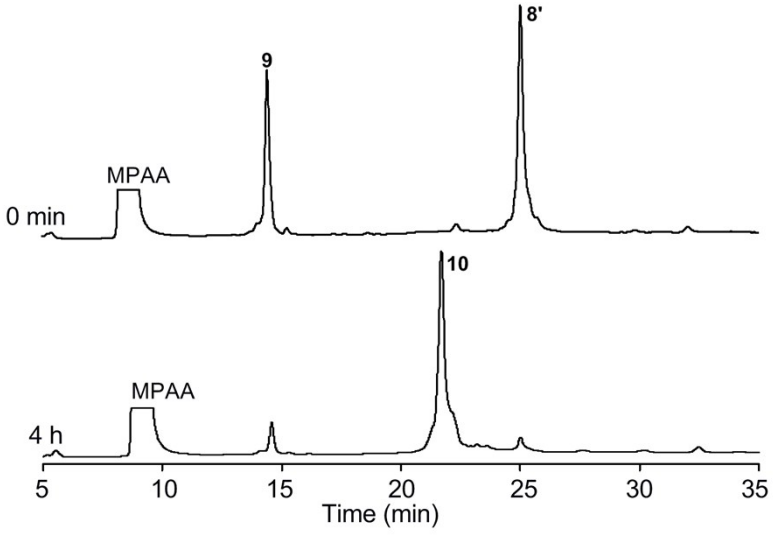


Figure S28. The analytical RP-HPLC traces (λ = 214 nm) for the ligation between peptide **8** and peptide **9**. 0 min and 4 h after the addition of peptide **9**. The peak marked with **8′** corresponds to the peptide thioester generated from peptide hydrazide **8**. HPLC condition: a linear gradient of 20%-45% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-45% for 30 min) on a Vydac C18 column.

13.4. Ligation of peptide 10 with peptide 7

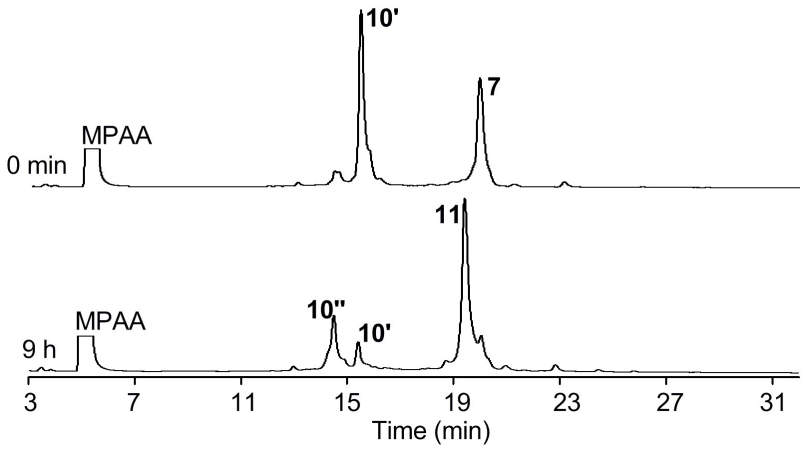


Figure S29. The analytical RP-HPLC traces (λ = 214 nm) for the ligation between peptide **7** and peptide **10**. 0 min and 9 h after the addition of peptide **7**. The peak marked with **10′** corresponds to the peptide thioester generated from peptide hydrazide **10**. The peak marked with **10″** corresponds to the hydrolysis product of **10′**. HPLC condition: a linear gradient of 25%-55% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 25%-55% for 30 min) on a Vydac C18 column. It should be mentioned that auxiliary-mediated ligation proceeds more slowly than that of standard Cysteine-mediated NCL. Moreover, the product **11** has similar retention time with **7**. Hence, excess peptide **10** should be used to facilitate the ligation reaction as well as to reduce the difficulty of the following separating operation.

13.5. Ligation of peptide 4 with peptide 12

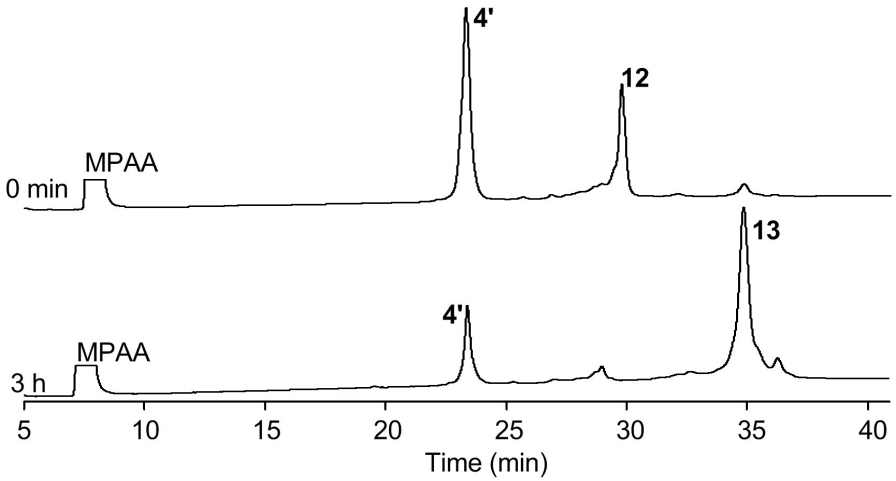


Figure S30. The analytical RP-HPLC traces (λ = 214 nm) for the ligation between peptide **4** and peptide **12**. 0 min and 3 h after the addition of peptide **12**. The peak marked with **4′** corresponds to the peptide thioester generated from peptide **4**. HPLC condition: a linear gradient of 20%-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 38 min (5% for 2 min, then 20%-60% for 38 min) on a Vydac C18 column. Especially, the presence of ubiquitin in peptide **12** and two internal Cysteine residues (Cys47 in peptide **4** and Cys46 in ubiquitin moiety) did not affect the condensation between peptide **4** and **12**.

13.6. Ligation of peptide 17 with peptide 18

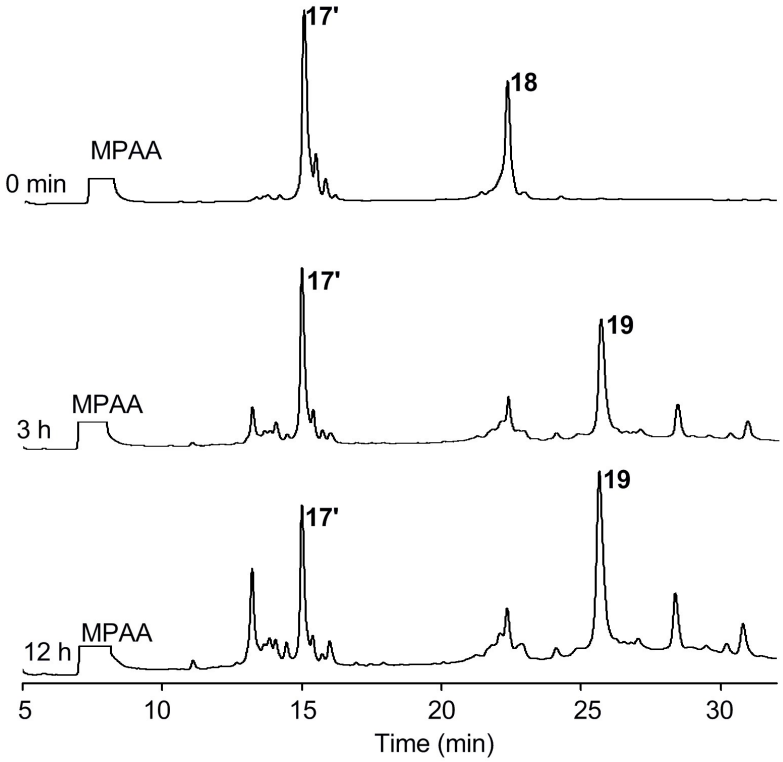


Figure S31. The analytical RP-HPLC traces (λ = 214 nm) for the ligation between peptide **17** (7.8 μmol, 1.3 equiv.) and peptide **18** (6 μmol, 1.0 equiv.). 0 min, 3 h and 12 h after the addition of peptide **18**. The peak marked with **17′** corresponds to the peptide thioester generated from peptide **17**. HPLC condition: a linear gradient of 20%-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-60% for 30 min) on a Vydac C18 column.

14. References

[1] (a) J. S. Zheng, S. Tang, Y. K. Qi, Z. P. Wang, L. Liu, *Nat. Protoc.* 2013, **8**, 2483; (b) Y. K. Qi, S. Tang, Y. C. Huang, M. Pan, J. S. Zheng, L. Liu, *Org. Biomol. Chem.*, 2016, **14**, 4194; (c) J. B. Li, Y. Y. Li, Q. Q. He, Y. M. Li, H. T. Li, L. Liu, *Org. Biomol. Chem.*, 2014, 12, 5435.

[2] J. B. Li, Q. Q. He, Y. T. Liu, S. L. Liu, S. Tang, C. M. Li, D. M. Sun, X. R. Li, M. Zhou, P. Zhu, G. Q. Bi, Z. H. Zhou, J. S. Zheng, C. L. Tian, *ChemBioChem.*, 2017, **18**, 176.

[3] Q. Wan, S. J. Danishefsky, *Angew. Chem. Int. Ed. Engl.*, 2007, **46**, 9248; *Angew. Chem.*, 2007, **119**, 9408.

[4] P. N. Dyer, R. S. Edayathumangalam, C. L. White, Y. Bao, S. Chakravarthy, U. M. Muthurajan, K. Luger, *Methods. Enzymol.*, 2004, **375**, 23.

[5] M. Pan, S. Gao, Y. Zheng, X. D. Tan, H. Lan, X. L. Tan, D. M. Sun, L. N. Lu, T. Wang, Q. Y. Zheng, Y. C. Huang, J. W. Wang, L. Liu. *J. Am. Chem. Soc.*, 2016, **138**, 7429.