Total chemical synthesis of the site-selective azide-labeled [I66A]HIV-1 protease

Yun-Kun Qi,[‡ab] Hao-Nan Chang,[‡b] Kai-Mai Pan[b] and Ji-Shen Zheng*[ab]

a.High Magnetic Field Laboratory, Chinese Academy of Sciences Hefei 230031, China.

b. Department of Chemistry, Tsinghua University, Beijing 100084, China

‡ These authors contributed equally to this work.

E-mail: jszheng@hmfl.cas.cn.

Table of Contents

1. Reagents and materials

Rink amide AM resin and 2-Chlorotrityl resin were purchased from Hecheng Technology (Tianjing, China). Fmoc-amino acids were purchased from GL Biochem (Shanghai, China), C S Bio or Bo Mai Jie Technology (Beijing, China). Other reagents and materials were purchased as previously described.[1]

2. HPLC and mass spectrometry (MS)

Reversed phase HPLC was all performed on Shimadzu Prominence HPLC. For peptide analysis, Vydac C18 (4.6 × 150 mm) and C8 (4.6 × 150 mm) columns were used at a flow rate of 1.0 mL/min. For peptide purification, Vydac C18 (10 × 250 mm), C8 (10 × 250 mm) and C4 (10 × 250 mm) columns were used at a flow rate of 3-4 mL/min. The UV absorption at 214 nm and 254 nm were monitored for the injections.

ESI-MS spectra were recorded on an Agilent 1200 Series HPLC system with LC-MS (Agilent 6340 ion trap as the mass spectrometer) or an Esquire~LC iron trap mass spectrometer. MALDI-TOF/MS was performed on an Applied Biosystems 4800PLUS MALDI-TOF/TOF mass spectrometer in the Center of Biomedical Analysis, Tsinghua University.

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

3.1. General procedures for Fmoc-SPPS

Peptides were synthesized manually by Fmoc-SPPS.[1] The Rink amide AM resin was swelled in DCM/DMF (1:1). After 30 min, the Fmoc group of the resin was removed by treatment with 20% piperidine in DMF (twice: 5 min, 10 min). Then the coupling was carried out using a solution of the Fmoc-amino acid (5 equiv.), *O*-(6-Chlorobenzotriazol-1-yl)-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate (HCTU) (4.5 equiv.) and DIPEA (10 equiv.) in DMF (the Fmoc-amino acid was activated for about 15 s before addition to the resin). 1.0-1.3 h reaction at 30 °C is enough for most amino acids. Double coupling could be carried out when needed (for high sterically hindered amino acids and the amino acid right after proline). After the coupling reaction, the resin was thoroughly washed with DMF and DCM. The Fmoc group was removed by 20% piperidine in DMF

(twice: 5 min and then 10 min). Then the resin was thoroughly washed with DMF and DCM. Then the next coupling reaction could be conducted. After all the coupling reactions completed, the N-terminal Fmoc group was removed by treatment with 20% piperidine in DMF (twice: 5 min, 10 min). After washing the resin thoroughly with DMF and DCM, the cleavage and deprotection steps was carried out by cocktail K (TFA/phenol/water/thioanisole/EDT, 82.5:5:5:5:2.5) or B (TFA/phenol/water/TIPS, 88:5:5:2) for 2.5-3.0 h. Then the TFA solutions were concentrated by pure argon. The crude peptide was then precipitated with cold Et_2O and then centrifuged (three times). Finally, the crude peptide was dissolved in H_2O/CH_3CN and purified by semi-preparative RP-HPLC.

3.2. Synthesis of peptide hydrazides

Peptide hydrazides **1**, **2, 6, 11** and **12** were synthesized following the Fmoc-SPPS procedures described in Section 3.1.

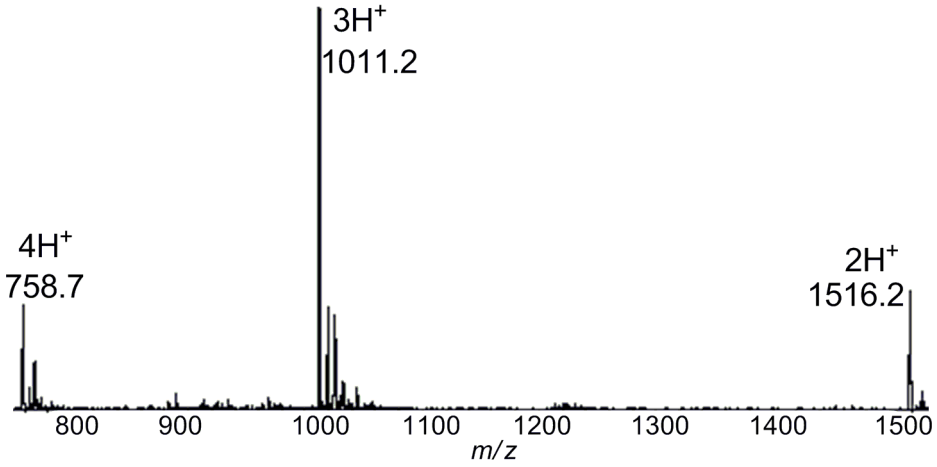


Figure S1. ESI-MS spectrum of purified Peptide **1.** The spectrum gave an observed mass of 3030.6 Da (calculated 3031.5 Da, average isotopes).

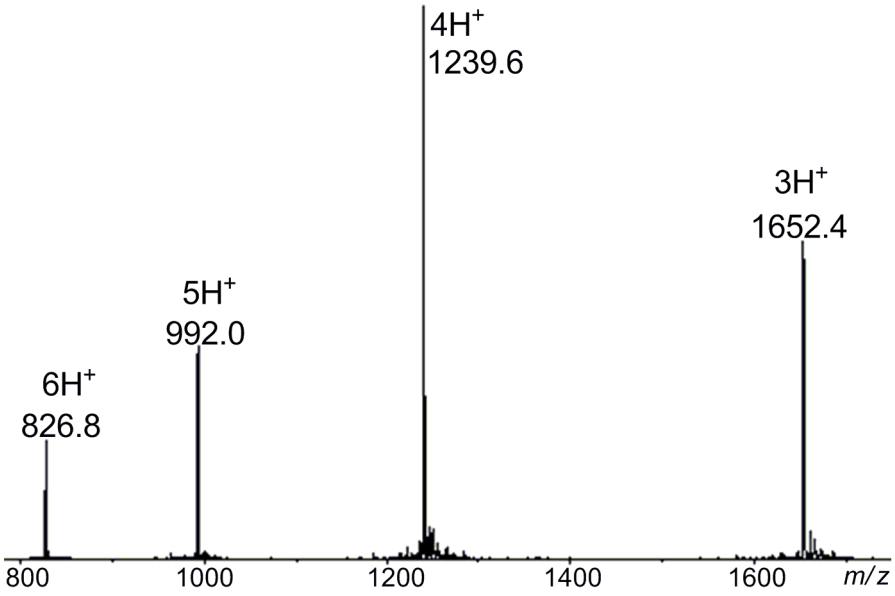


Figure S2. ESI-MS spectrum of purified Peptide **2.** The spectrum gave an observed mass of 4954.6 Da (calculated 4955.7 Da, average isotopes).

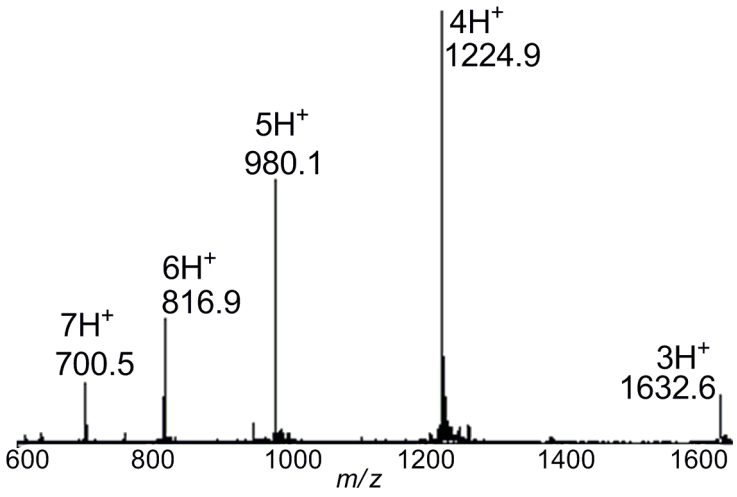


Figure S3. ESI-MS spectrum of purified Peptide **6.** The spectrum gave an observed mass of 4895.6 Da (calculated 4896.7 Da, average isotopes).

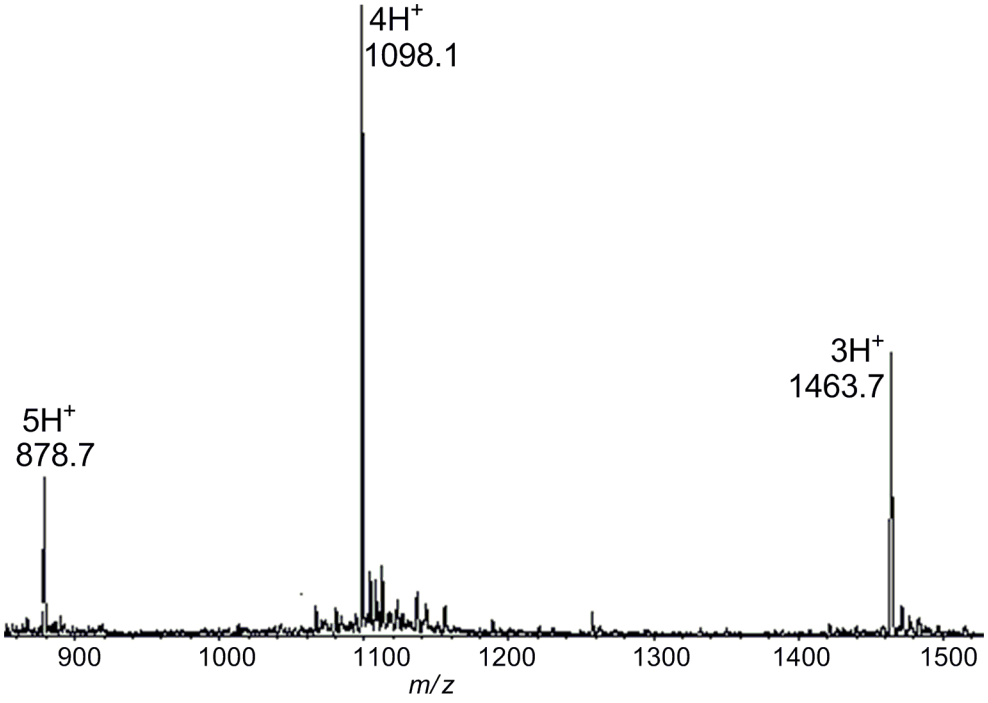


Figure S4. ESI-MS spectrum of purified Peptide **11.** The spectrum gave an observed mass of 4388.3 Da (calculated 4389.0 Da, average isotopes).

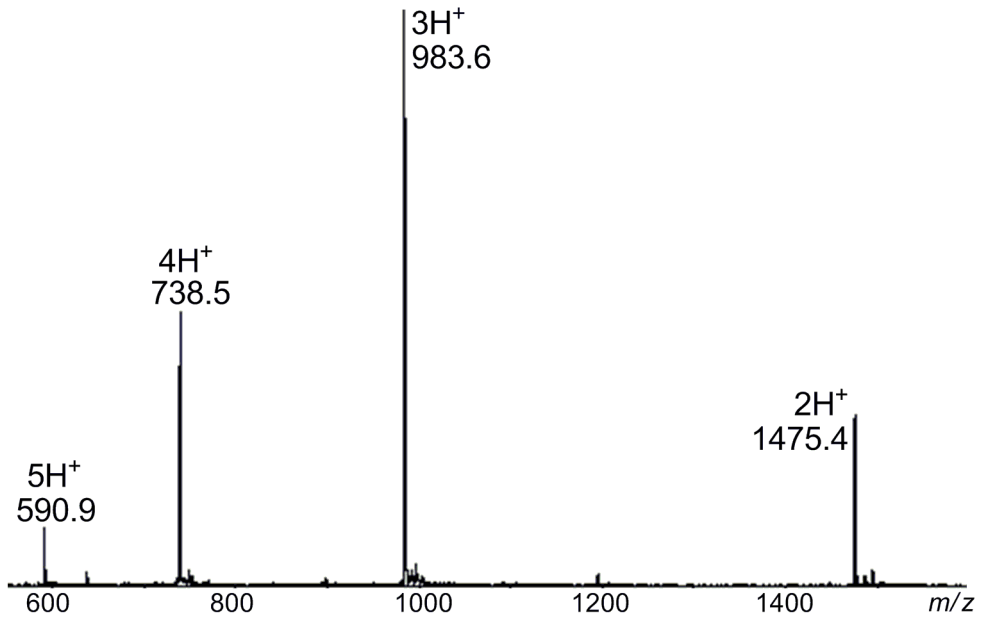


Figure S5. ESI-MS spectrum of purified Peptide **12.** The spectrum gave an observed mass of 2949.0 Da (calculated 2948.5 Da, average isotopes).

3.3. Synthesis of peptide amides

Peptide amides **3**, **7**, **13** and **19** were synthesized on Rink amide AM resin following the Fmoc-SPPS procedures described in Section 3.1.

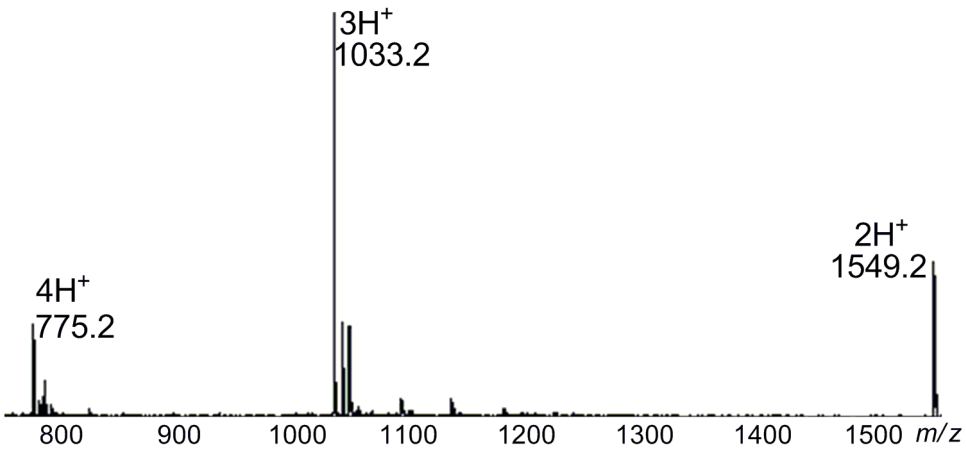


Figure S6. ESI-MS spectrum of purified Peptide **3.** The spectrum gave an observed mass of 3096.6 Da (calculated 3097.6 Da, average isotopes).

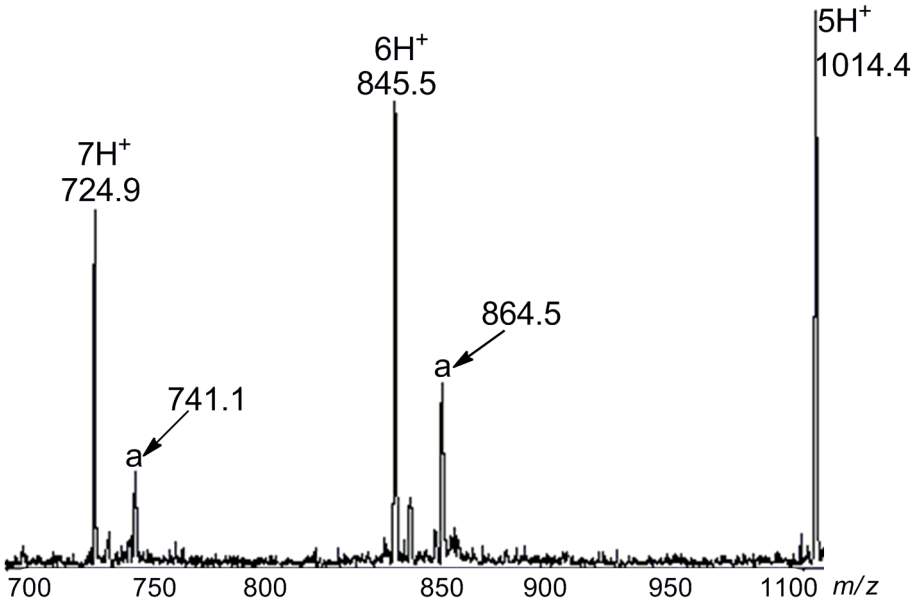


Figure S7. ESI-MS spectrum of purified Peptide **7.** The spectrum gave an observed mass of 5067.1 Da (calculated 5067.9 Da, average isotopes). Peaks marked by **a** corresponded to compound with the mass of approximately M+114. The ratio of these peaks to the main peaks varied from injections to

injections. These peaks were thought to be associated with the C-terminus Arg_6 tag because these peaks were observed only in ESI-MS spectra of peptides containing the Arg_6 tag.[2] Because these peaks may disappear after the detachment of the Arg_6 tag from the peptide, our synthesis process was not hampered by this problem.

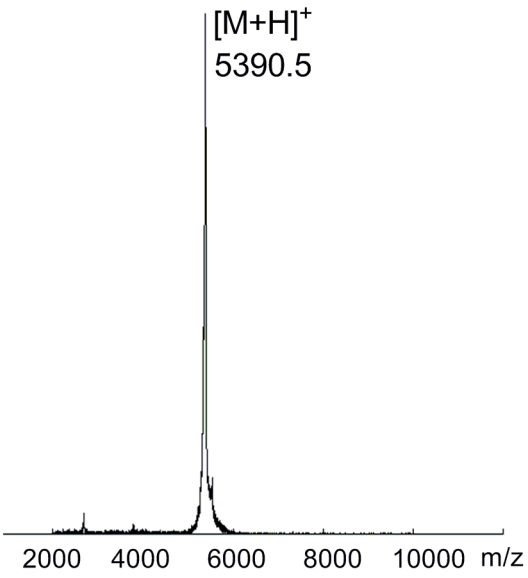


Figure S8. MALDI-TOF MS spectrum of purified Peptide **13.** The spectrum gave an observed mass of 5389.5 Da (calculated 5390.4 Da, average isotopes).

4. The attempts to synthesize azide-labled HIV-1 protease

Analytic RP-HPLC chromatogram (λ = 214 nm) of purified **5** (Figure 1E in main text). HPLC conditions: 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 C8 (4.6 × 150 mm) column.

Although the free-radical-based desulfurization reaction is often used in protein chemical synthesis, the understanding of the compatibility of azide groups with free-radical-based desulfurization reaction is limited. To test the compatibility of azide groups with free-radical-based desulfurization reaction, we treated PR[Pro^{1}−$Lys(N_3)^{41}$−Lys^{70}]-$NHNH_2$ **4** with aqueous TCEP (500 mM) containing 6 M Gn·HCl and 0.2 M Na_2HPO_4 (pH = 6.6), *t*BuSH and 2,2'-azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride (VA-044).[1] After stirred at room temperature for 2 min, PR[Pro^{1}−$Lys(N_3)^{41}$−Lys^{70}]-$NHNH_2$ **4** was completely converted to PR[Pro^{1}−Lys^{41}−Lys^{70}]-$NHNH_2$ **5**. ESI-MS analysis of **5** showed an exact 26 Da decrease, which could be attributed to the reduction of azide by high concentration of TCEP. Through the initial attempts, we found out that although azide group could survive the TCEP during NCL because of the low TCEP concentration and a quick operation, it can not bear such a high concentration of TCEP during free-radical-based desulfurization reaction.

5. The attempts to synthesize alkyne-labled HIV-1 protease

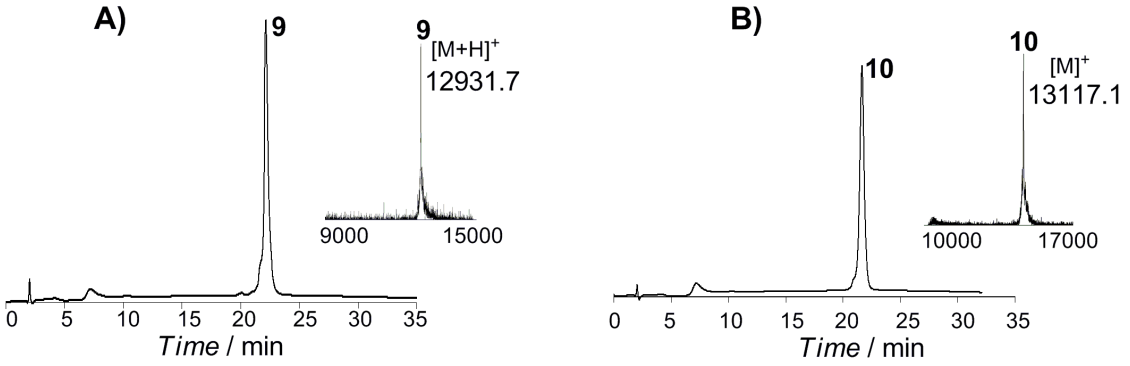


Figure S9. A) The analytical RP-HPLC trace (λ = 214 nm) and MALDI-TOF mass (inner) of purified peptide **9**. The mass spectrum gave an observed mass of 12930.7 Da (calcd 12932.2 Da, average isotopes). B) The analytical RP-HPLC trace (λ = 214 nm) and MALDI-TOF mass (inner) of purified peptide **10**. The mass spectrum gave an observed mass of 13117.1 Da. RP-HPLC conditions for **9** and **10**: 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 C8 (4.6 × 150 mm) column.

5.1. The study of the side reaction during free-radical-based desulfurization

5.1.1.

Peptide **22** PR[Asn^{37}–Pra^{41}–Lys^{43}]-NH_2 and Peptide **23** PR[Asn^{37}–Ala^{41}–Lys^{43}]-NH_2 were synthesized through standard Fmoc-SPPS. To study the site of side reaction and compatibility of the alkyne groups with the free-radical-based desulfurization reaction often involved in protein chemical synthesis, we treated Fmoc-Pra-COOH **25**, **22** and **23** with Tris(2-carboxyethyl)phosphine (TCEP), 2-methyl-2-propanethiol (*t*BuSH) and (2,2'-azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride) (VA-044).

Fmoc-Pra-COOH **25** (0.34 mg, 1 μmol) was dissolved in 0.3 mL DMF and 0.7 mL of aqueous solution (pH 6.6) containing 500 mM TCEP. Then 50 μl of *t*BuSH and 400 μl of VA-044 (0.1 M in water, PH7.0) were added. The reaction mixture was adjusted to pH 6.6 and then stirred on a magnetic stirrer at 37 ^{o}C. Under this reaction condition, **25** was mainly converted to **26** within 6 h. **26** showed an increase of 251.2 Da compared to **25**. The NMR spectra of **26** (Figure S11). ^{1}H NMR

(600 MHz, [D_4]CH_3OH): δ (ppm) = 7.80 (d, *J* = 6 Hz, 2H), 7.67 (d, *J* = 6 Hz, 2H), 7.40-7.31 (m, 4H), 6.95 (m, 1H), 6.08 (m, 1H), 4.43-4.23 (m, 4H), 2.81-2.53 (m, 14H). ^{31}P NMR (243 MHz, [D_6]DMSO): δ (ppm) = + 35.80. HRMS (Positive ESI) Calcd. For $C_{29}H_{33}NO_{10}P^+$: 586.1842, Found: 586.1847. It was speculated that TCEP (Calcd mass 250.2 Da, average isotopes) was added to the alkyne group of **26** under the free-radical-based desulfurization condition.

Peptide **22** (0.4 mg, 0.5 μmol) was dissolved in 1 mL aqueous solution (pH 6.8) containing 0.2 M Na_2HPO_4, 6 M Gn·HCl and 500 mM TCEP. Then 50 μl of *t*BuSH and 500 μl of 0.1 M VA-044 in 0.2 M Na_2HPO_4 solution (pH 7.0) were added. The reaction mixture was adjusted to pH 6.9 and then stirred on a magnetic stirrer at 37 ^oC. Peptide **23** was also treated with TCEP, *t*BuSH and VA-044. The major procedures of **23** were exactly the same as **22**.

Under the commonly used free-radical-based desulfurization condition, **22** was gradually converted to **24** within 18 h. ESI-MS analysis of product **24** showed an increase of 251.2 Da compared to peptide **22**. It could be concluded that peptide **22** was not compatible with the commonly used free-radical-based desulfurization condition. On the contrary, **23** was stable in the same reaction condition within 24 h. Considering that there was only one amino acid residue difference between **22** and **23**, we concluded that Pra^{41} may be the modification site during free-radical-based desulfurization reaction.

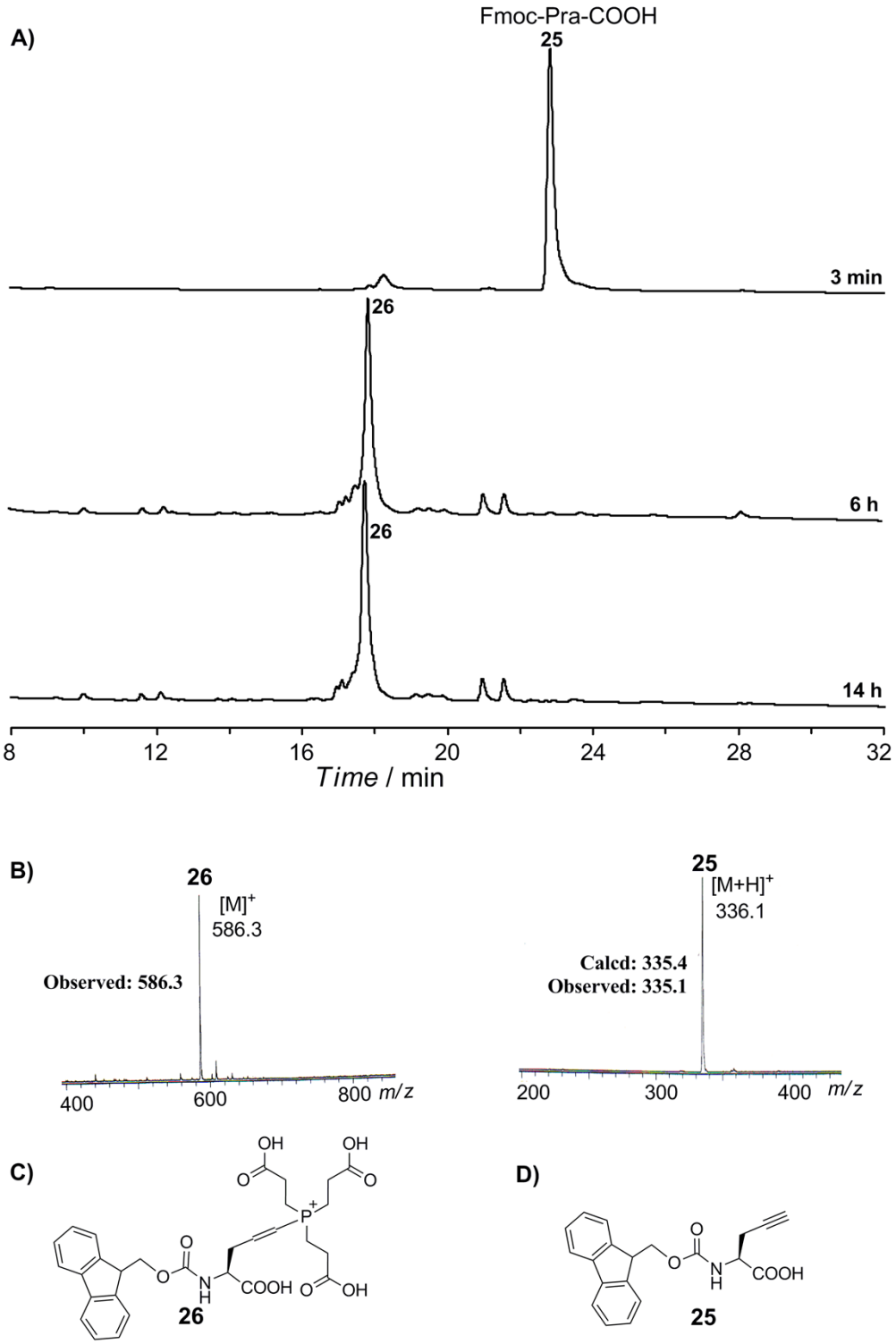


Figure S10. A) The reaction mixture of **25** was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). HPLC conditions: a linear gradient of 5-80% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 5%-80% for 30 min) on a Vydac C18 (4.6 × 150 mm) column. B) The ESI-MS spectra of purified **25** and **26**. C) The estimated molecular structure of **26**. D) The molecular structure of **25.**

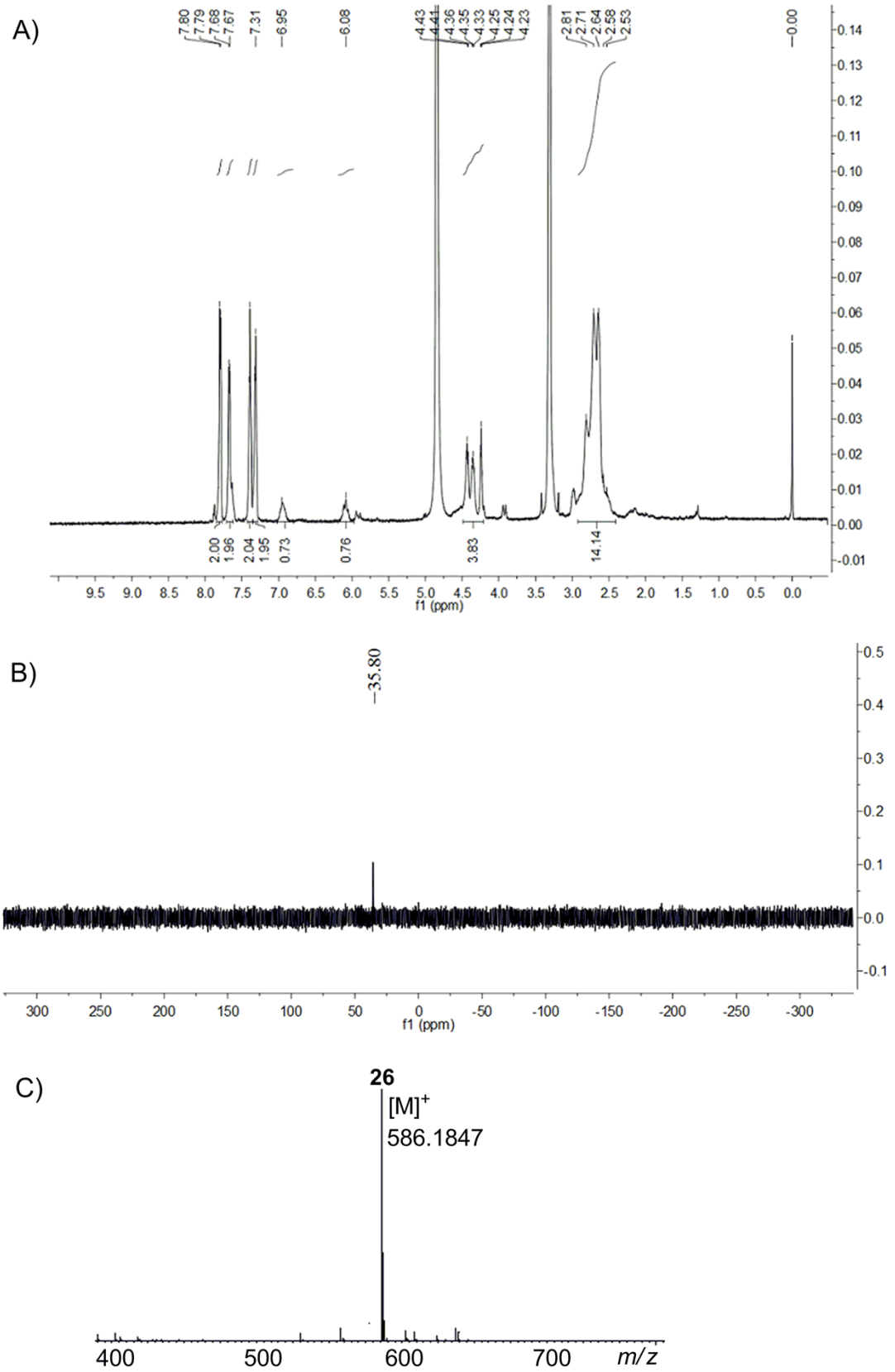


Figure S11. A) The ^{1}H NMR spectrum of **26**. B) The ^{31}P NMR spectrum of **26**. C) The HRMS spectrum of **26**.

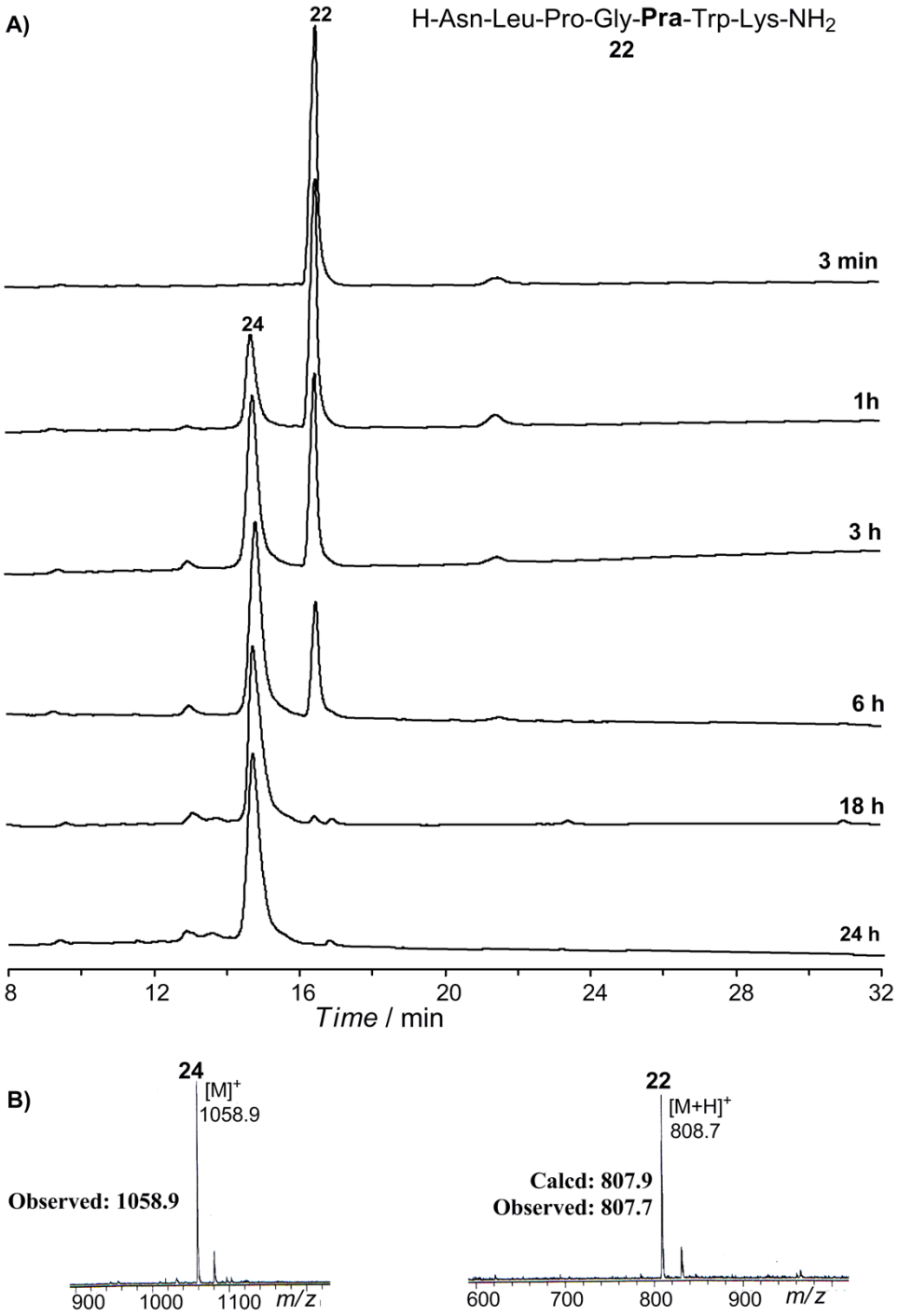


Figure S12. A) The reaction mixture of **22** was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). HPLC conditions: a linear gradient of 5-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 5%-50% for 30 min) on a Vydac C18 (4.6 × 150 mm) column. B) The ESI-MS spectra of purified **24** and **22**.

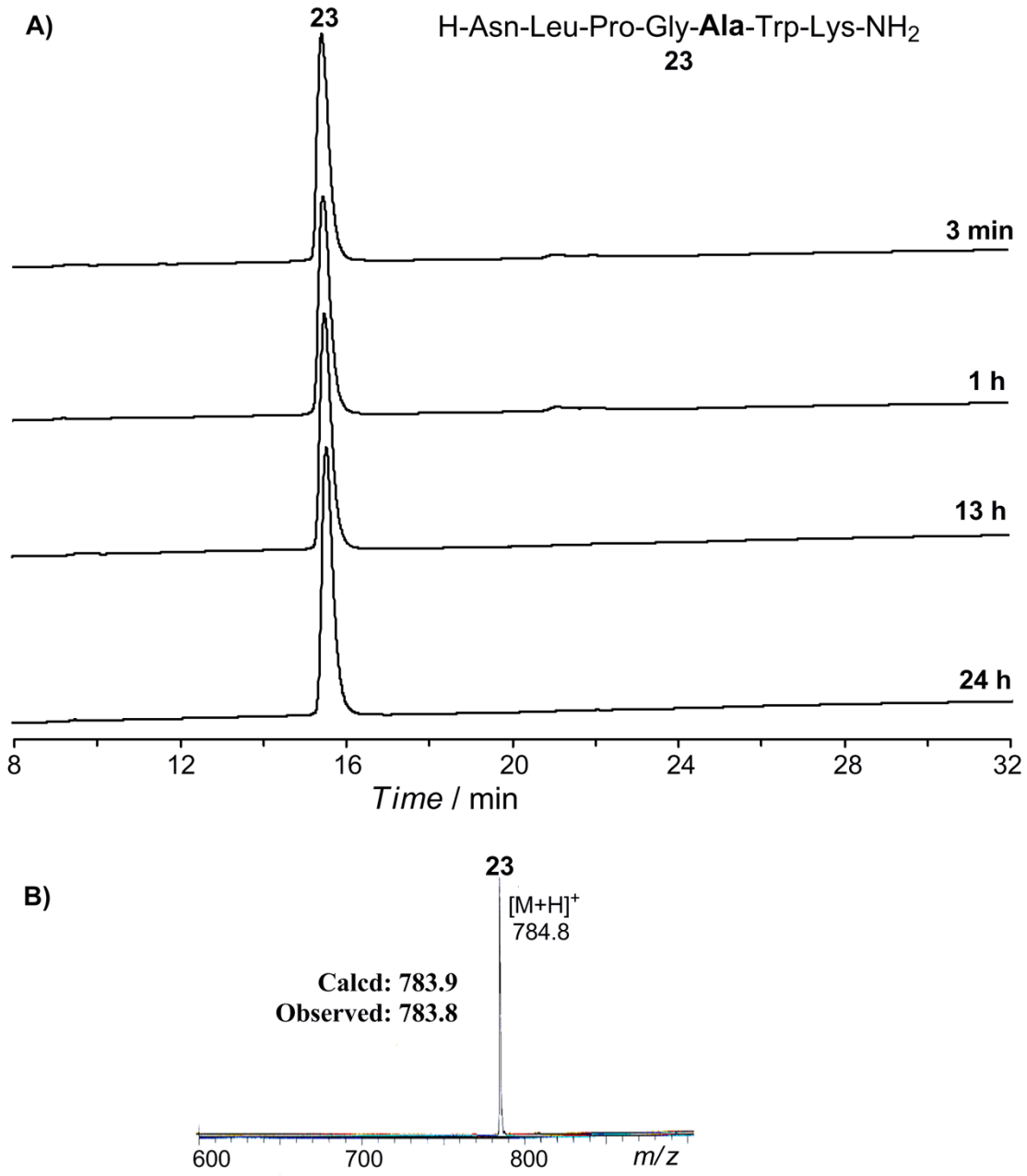


Figure S13. A) The reaction mixture of **23** was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). HPLC conditions: a linear gradient of 5-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 5%-50% for 30 min) on a Vydac C18 (4.6 × 150 mm) column. B) The ESI-MS spectrum of purified **23**.

5.1.2.

Note that VA-044 (Calcd mass 250.3 Da, average isotopes) and TCEP (Calcd mass 250.2 Da, average isotopes) have almost the same molecular weight. To further study the source of the

modification and the compatibility of alkyne groups with free-radical-based desulfurization reaction, we treated Fmoc-Pra-COOH **25**, **22** and **23** with TCEP, *t*BuSH and 2,2'-Azobis(2-methylpropionamidine)dihydrochloride (V-50). Both VA-044 and V-50 are water-soluble radical initiators.

Fmoc-Pra-COOH **25** (0.34 mg, 1 μmol) was dissolved in 0.3 mL DMF and 0.7 mL of aqueous solution (pH 6.6) containing 500 mM TCEP. Then 50 μl of *t*BuSH and 400 μl of V-50 (0.1 M in water, PH7.0) were added. The reaction mixture was adjusted to pH 6.6 and then stirred on a magnetic stirrer at 39 ^{o}C. Under this reaction condition, **25** was mainly converted to **26** within 8 h. **26** showed an increase of 251 Da compared to **25**.

Peptide **22** (0.4 mg, 0.5 μmol) was dissolved in 1 mL aqueous solution (pH 6.8) containing 0.2 M Na_2HPO_4, 6 M Gn·HCl and 500 mM TCEP. Then 50 μl of *t*BuSH and 500 μl of 0.1 M V-50 in 0.2 M Na_2HPO_4 solution (pH 7.0) were added. The reaction mixture was adjusted to pH 6.9 and then stirred on a magnetic stirrer at 39 ^{o}C. Peptide **23** was also treated with TCEP, *t*BuSH and V-50. The major procedures of **23** were exactly the same as **22**.

Under the desulfurization condition using V-50 as the radical initiator, **22** was gradually converted to **24** within 2 h. ESI-MS analysis of product **24** showed an increase of 251.2 Da compared to peptide **22**. It could be concluded that peptide **22** was not compatible with the desulfurization condition using V-50 as the radical initiator. On the contrary, **23** was stable in the same reaction condition within 12 h. Considering that there was only one amino acid residue difference between **22** and **23**, we further estimated that Pra^{41} may be the modification site during free-radical-based desulfurization reaction. Moreover, TCEP may be the source of the modification during the desulfurization reaction.

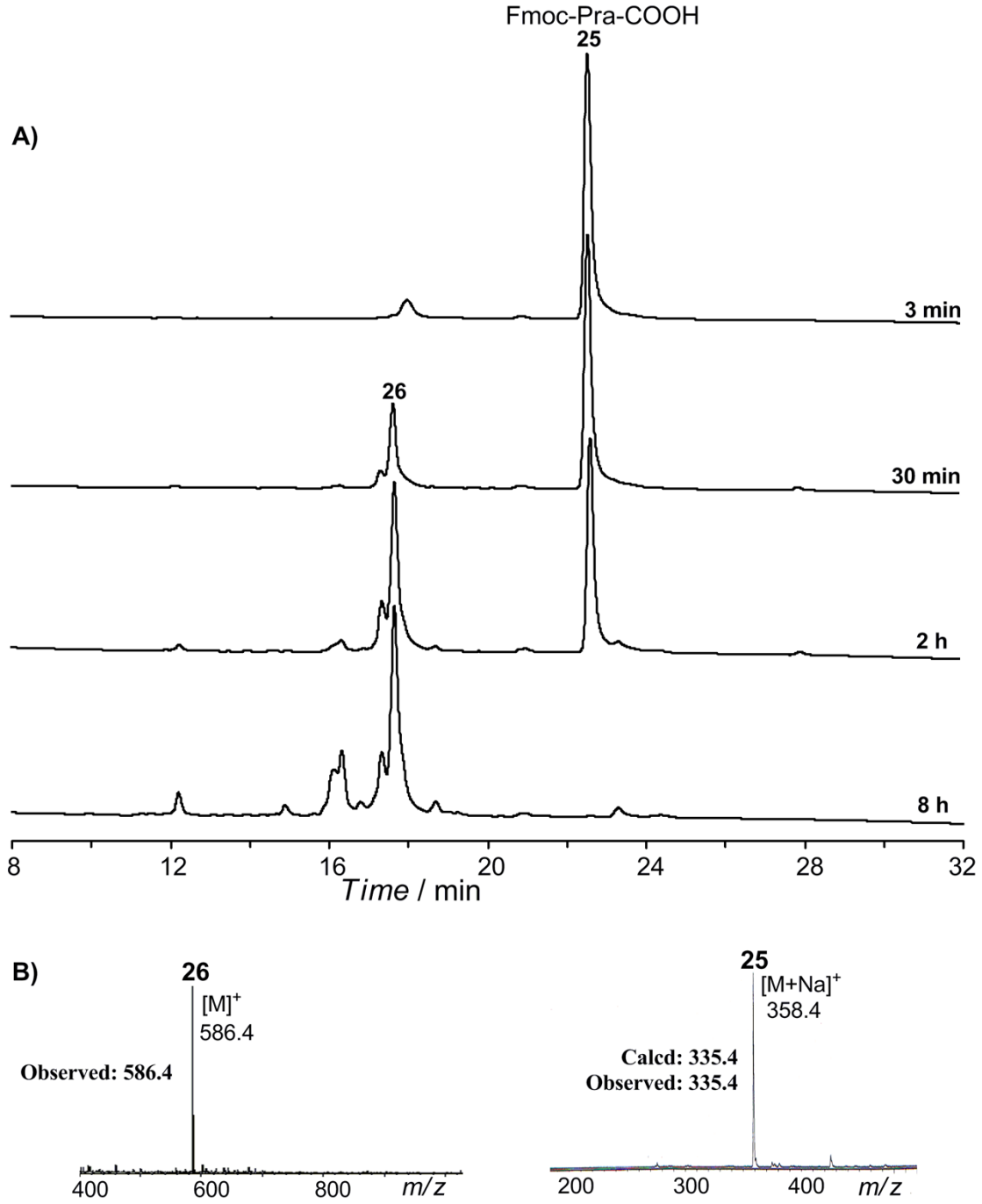


Figure S14. A) The reaction mixture of **25** were monitored by analytical RP-HPLC chromatograms (λ = 214 nm). HPLC conditions: a linear gradient of 5-80% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 5%-80% for 30 min) on a Vydac C18 (4.6 × 150 mm) column. B) The ESI-MS spectra of purified **25** and **26**.

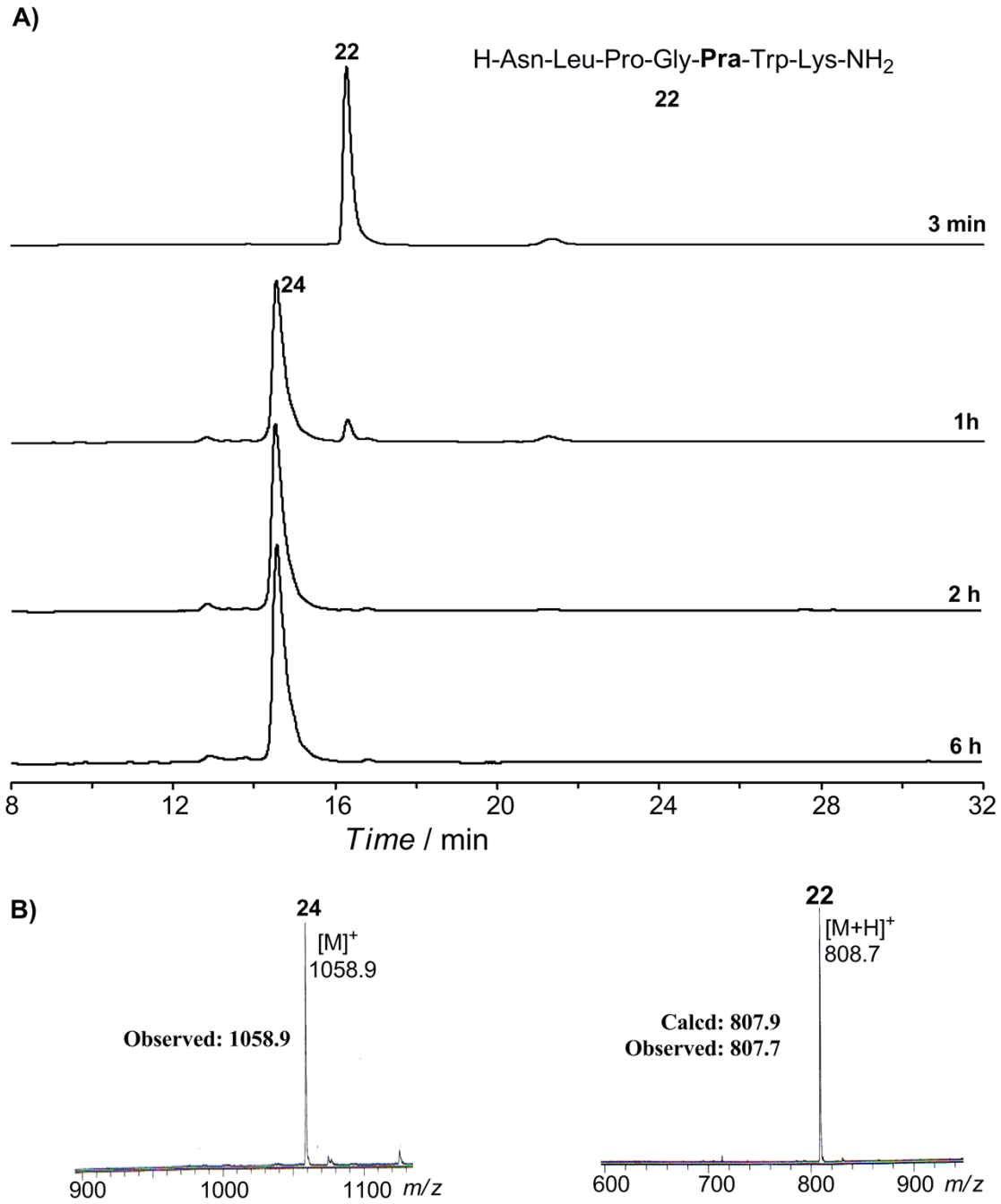


Figure S15. A) The reaction mixture of **22** was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). HPLC conditions: a linear gradient of 5-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 5%-50% for 30 min) on a Vydac C18 (4.6 × 150 mm) column. B) The ESI-MS spectra of purified **24** and **22**.

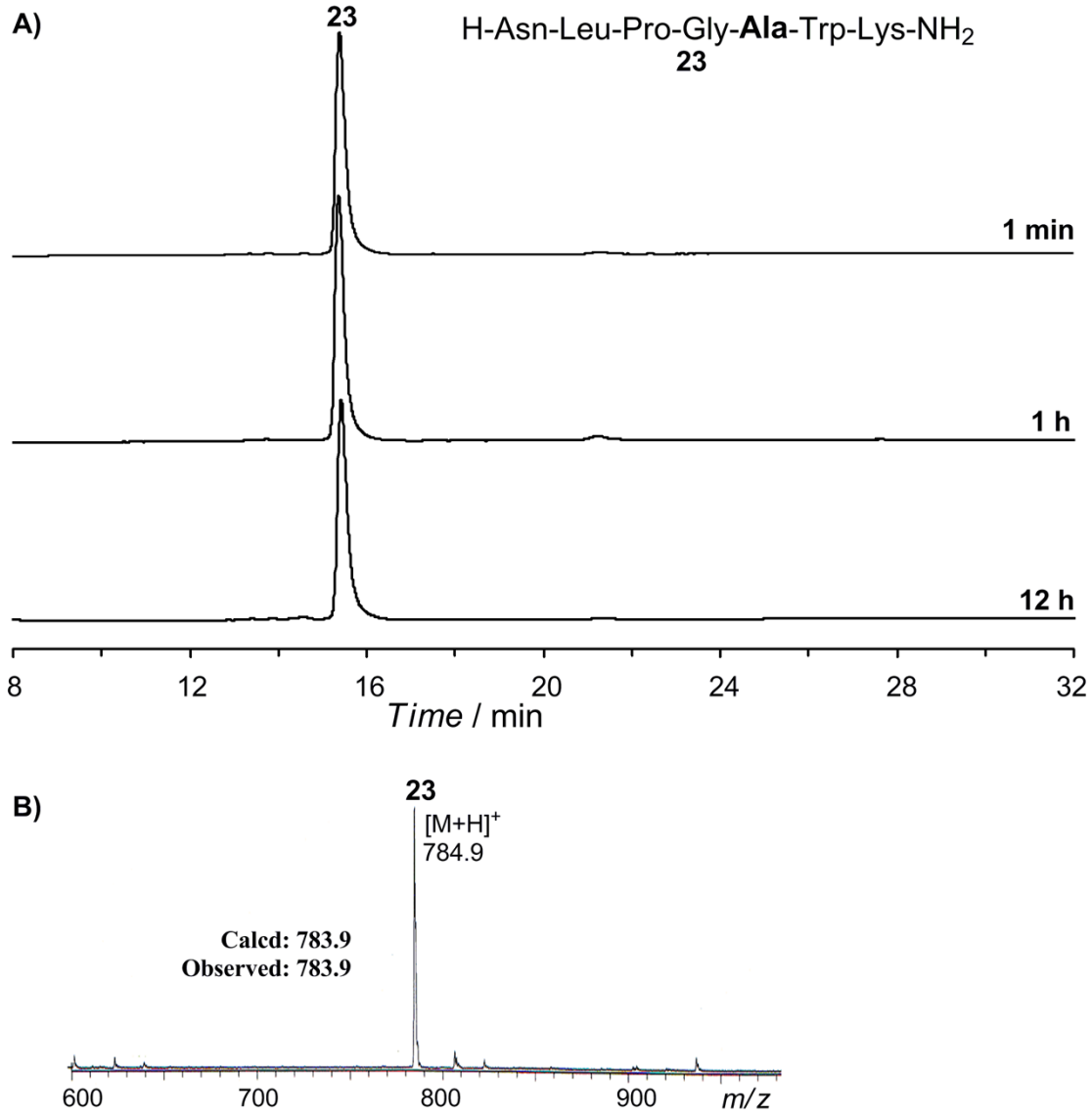


Figure S16. A) The reaction mixture of **23** was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). HPLC conditions: a linear gradient of 5-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 5%-50% for 30 min) on a Vydac C18 (4.6 × 150 mm) column. B) The ESI-MS spectrum of purified **23**.

5.2. The attempts to prevent the side reaction during free radical desulfurization

To prevent the side reaction during the free-radical-based desulfurization process, three different conditions of desulfurization reaction were studied. Petide **22** was used as the model to test the compatibility of alkyne groups with desulfurization reaction. Unfortunately, Petide **22** was not compatible with any of the three desulfurization reactions.

5.2.1.

According to Danishefsky' study,[3] ethanethiol could also be used as the hydrogen source during desulfurization reaction. We tried to replace the *t*BuSH with ethanethiol. Peptide **22** (0.8 mg, 1 μmol) was dissolved in 1 mL aqueous solution (pH 6.9) containing 0.2 M Na_2HPO_4, 6 M Gn·HCl under argon. To the solution, 1 mL aqueous solution (pH 6.9) containing 0.2 M Na_2HPO_4, 6 M Gn·HCl and 500 mM TCEP was added. Then, 100 μl of ethanethiol and 50 μl of VA-044 (0.1 M in water) were added. The reaction mixture was adjusted to pH 6.8 and then stirred on a magnetic stirrer at 37 °C. In this reaction condition, **22** was gradually converted to **24** within 20 h.

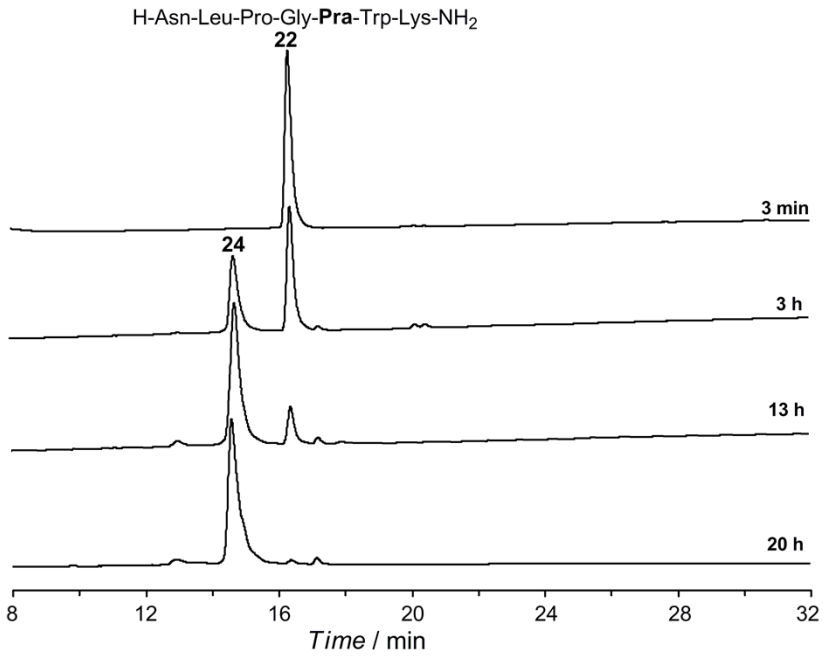


Figure S17. The reaction mixture of **22** was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). HPLC conditions: a linear gradient of 5-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 5%-50% for 30 min) on a Vydac C18 (4.6 × 150 mm) column.

5.2.2.

We tried to reduce the concentration of VA-044 during desulfurization reaction. Peptide **22** (0.4 mg, 0.5 μmol) was dissolved in 1 mL aqueous solution (pH 6.8) containing 0.2 M Na_2HPO_4, 6 M Gn·HCl and 500 mM TCEP. Then 50 μl of *t*BuSH and 500 μl of 0.02 M VA-044 in 0.2 M Na_2HPO_4 solution (pH 7.0) were added. The reaction mixture was adjusted to pH 6.9 and then stirred on a magnetic stirrer at 37 °C. In this reaction condition, **22** was gradually converted to **24** within 15 h.

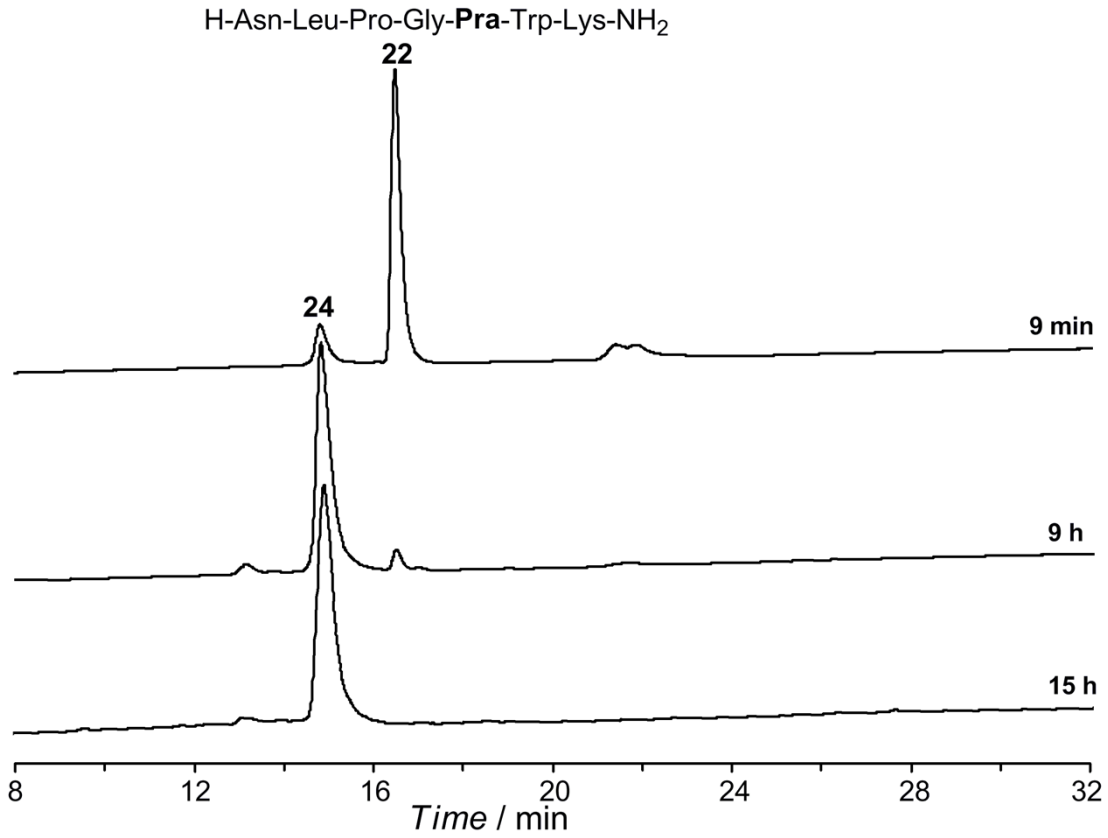


Figure S18. The reaction mixture of **22** was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). HPLC conditions: a linear gradient of 5-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 5%-50% for 30 min) on a Vydac C18 (4.6 × 150 mm) column.

5.2.3.

We tried to change the pH value of desulfurization reaction. Peptide **22** (0.4 mg, 0.5 μmol) was dissolved in 1 mL aqueous solution (pH 4.0) containing 0.2 M Na_2HPO_4, 6 M Gn·HCl and 500 mM TCEP. Then 50 μl of *t*BuSH and 500 μl of 0.1 M VA-044 in 0.2 M Na_2HPO_4 solution (pH 7.0) were

added. The reaction mixture was adjusted to pH 4.0 and then stirred on a magnetic stirrer at 37 °C. In this reaction condition, **22** was mainly converted to **24** within 12 h.

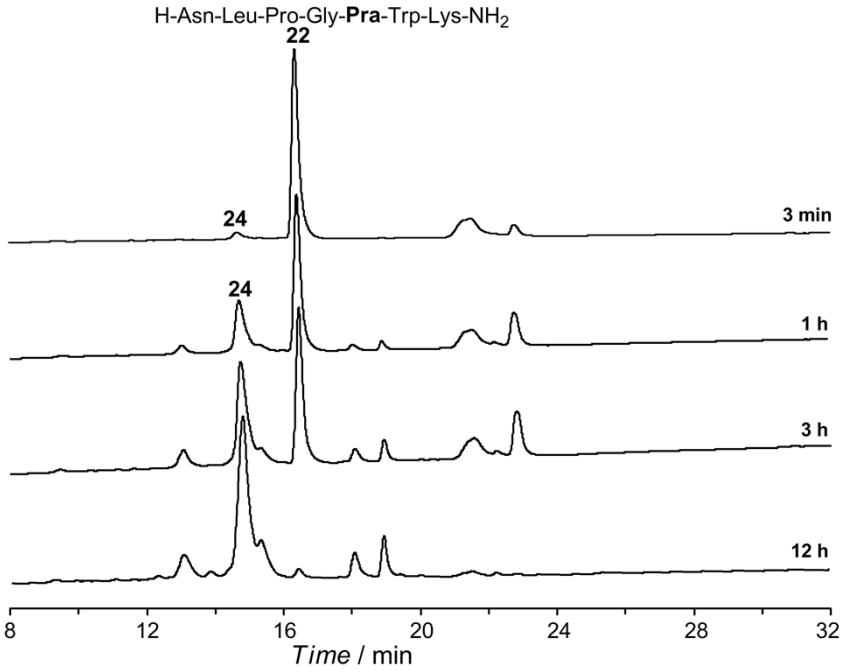


Figure S19. The reaction mixture of **22** was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). HPLC conditions: a linear gradient of 5-50% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 5%-50% for 30 min) on a Vydac C18 (4.6 × 150 mm) column.

6. The model of HIV-1 protease precursor with solubilizing tag

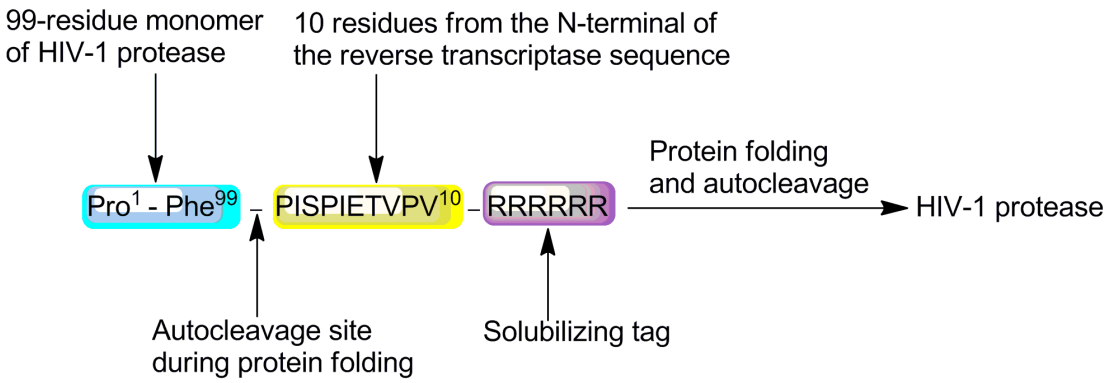
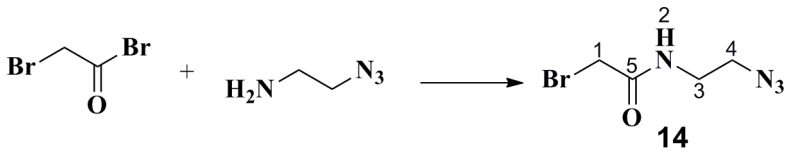


Figure S20. The HIV-1 protease precursor PR[Pro^{1}–$Cys(N_3)^{41}$–Arg^{115}]-NH_2 contains the 1-99 protease polypeptide in addition to 10 residues from the N-terminal of the reverse transcriptase polypeptide and 6 arginine residues. When successfully folded, the synthesized HIV-1 protease precursor was supposed to automatically cleave the C-terminal 16 amino-acid residues and then become the mature active HIV-1 protease, just as the HIV-1 protease did naturally.

7. Synthesis of compound 14



A solution of 2-bromoacetyl bromide (8.58 g, 42.5 mmol) in dry THF (25 mL) was added dropwise to a solution of 2-azidoethanamine (1.66 g, 19.3 mmol) and triethylamine (4.48 g, 44.3 mmol) in dry THF (50 mL) at 0ºC. After complete addition, the reaction mixture was allowed to stir for 1.3 h at room temperature. The excess of acid bromide was quenched by addition of methanol (8 mL). The formed triethylammonium bromide salt was filtered off and the solvent was removed in vacuo. The crude product was dissolved in dichloromethane and washed three times with a saturated ammonium chloride solution and then two times with distilled water. The organic layer was dried with magnesium sulfate. Evaporation of the solvents afforded a pale brown oil product. The pale brown oil was then purified by semi-preparative RP-HPLC, affording the desired compound **14** as a colorless oil (2.12 g, 53 %). ^{1}H NMR (300 MHz, [D_6]DMSO): δ (ppm) = 8.53 (s, 1H, H-2), 3.86 (s, 2H, H-1), 3.39 (t, *J* = 4.5 Hz, 2H, H-3), 3.28 (t, *J* = 4.7 Hz, 2H, H-4). ^{13}C NMR (300 MHz, [D_6]DMSO): δ (ppm) = 166.7 (C-5), 50.0 (C-4), 39.0 (C-3), 29.4 (C-1). ESI mass calcd for $C_4H_7BrN_4O$: 207.0 (average isotopes), observed: 207.0.

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

Peptide **1** (21.2 mg, 7 μmol, 1 equiv., final concentration 2 mM) was dissolved in an aqueous buffer containing 6 M guanidine hydrochloride (Gn·HCl) and 0.2 M NaH_2PO_4 (pH = 3.0), and cooled to approximately -13 °C in an ice-salt bath. A solution of $NaNO_2$ (7 equiv.) in the same buffer (pH = 3.0) was then added dropwise to activate peptide **1**. After stirred at -13°C for 25 min, 4-mercaptophenylacetic acid (MPAA, 50 equiv., 100 mM) dissolved in 0.2 M Na_2HPO_4 solution containing 6 M Gn·HCl (pH 7.0) was added into the mixture. The reaction was then taken out of the ice-salt bath and stirred at room temperature for 2 min. Subsequently, PR[Cys^{28}–$Lys(N_3)^{41}$–Lys^{70}]-$NHNH_2$ **2** (52.0 mg, 10.5 μmol, 1.5 equiv., final concentration 3 mM) was added into the reaction mixture. The pH value of reaction mixture was then adjusted to 6.6-6.8 slowly with aqueous NaOH solution (2 M) to initiate the NCL at room temperature. The reaction process was monitored by analytic RP-HPLC. The product was purified by semi-preparative RP-HPLC and characterized by ESI-MS. The reaction solution was reduced by aqueous Tris(2-carboxyethyl)phosphine (TCEP) (10 mM) containing 6 M Gn·HCl (pH = 4.0) with a quick operation before reaction analysis or product isolation. After 5 h, the ligation between **1** and **2** was completed, affording the desired product PR[Pro^1–$Lys(N_3)^{41}$–Lys^{70}]-$NHNH_2$ **4** (26.7 mg, 3.36 μmol) in 48% isolated yield.

8.1. Ligation of peptide hydrazide 1 with peptide hydrazide 2

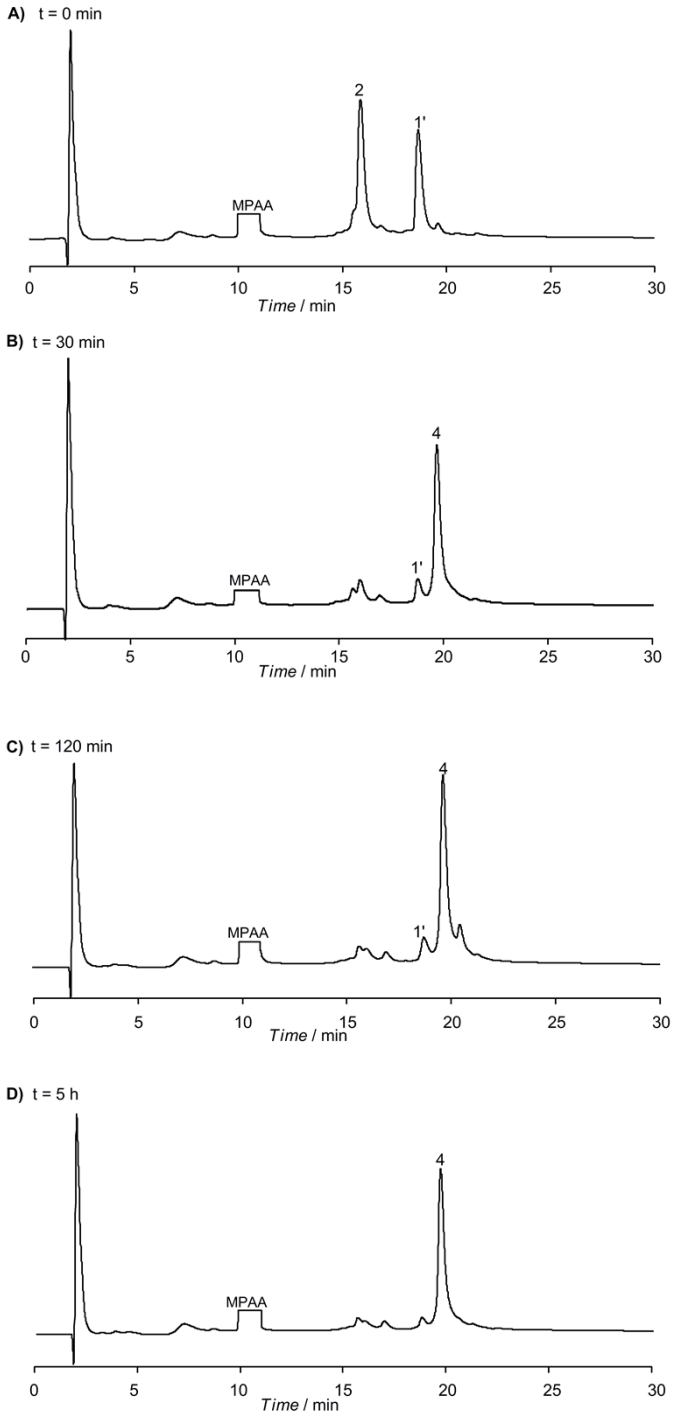


Figure S21. The ligation reaction between peptide hydrazide **1** and peptide hydrazide **2**. The ligation was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). Time points: 0 min, 30 min, 120 min and 5 h after the addition of peptide hydrazide **2**. The peak marked with **1'** corresponds to the peptide thioester generated from peptide hydrazide **1**. HPLC conditions: 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 C8 (4.6 × 150 mm) column.

8.2. The failed ligation of peptide hydrazide 4 with peptide amide 3

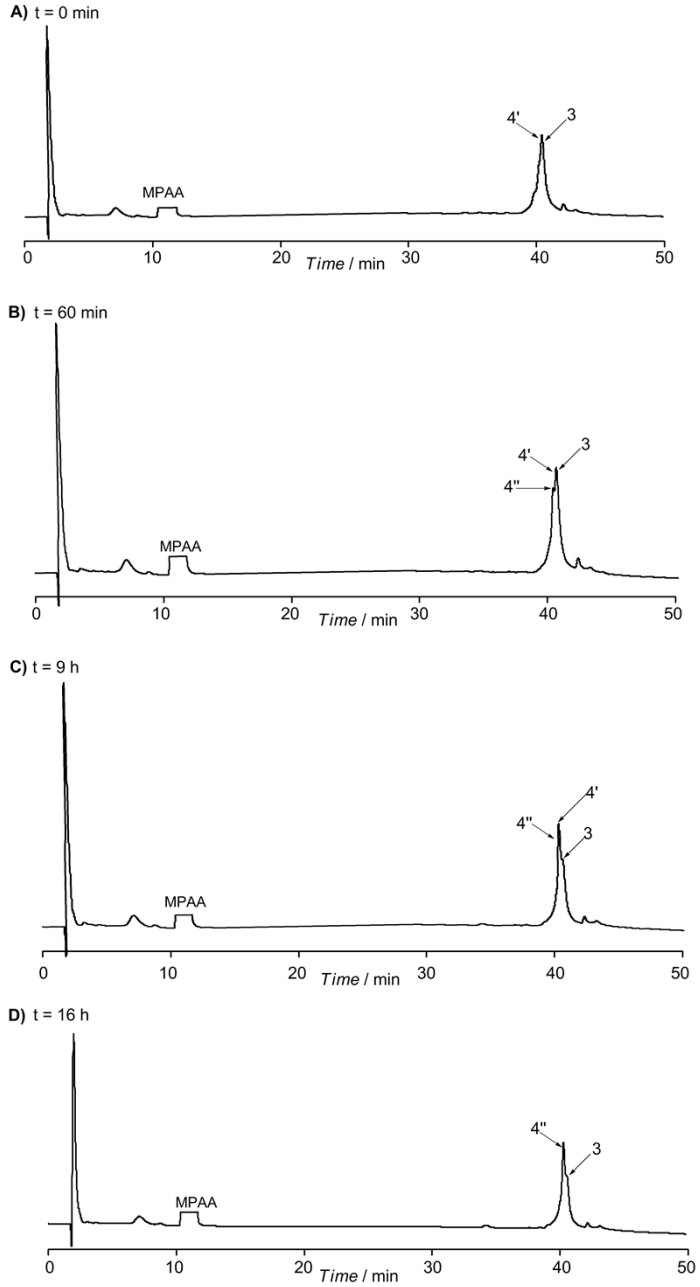


Figure S22. The failed ligation reaction between peptide hydrazide **4** (8.0 mg, 1 μmol, 1 equiv.) and peptide amide **3** (4.6 mg, 1.5 μmol, 1.5 equiv.). The major procedures were the same as described

above. The ligation was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). Time points: 0 min, 60 min, 9 h and 16 h after the addition of peptide amide **3**. The peak marked with **4'** corresponds to the peptide thioester generated from peptide hydrazide **4**. The peak marked with **4''** corresponds to the hydrolysis product of **4'**. Peptides **3**, **4'** and **4''** had almost the same retention time. HPLC conditions: a linear gradient of 20-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 50 min (5% for 2 min, then 20% for 20 min, then 20%-60% for 30 min) on a Vydac C8 (4.6 × 150 mm) column.

8.3. Ligation of peptide hydrazide 1 with peptide hydrazide 6

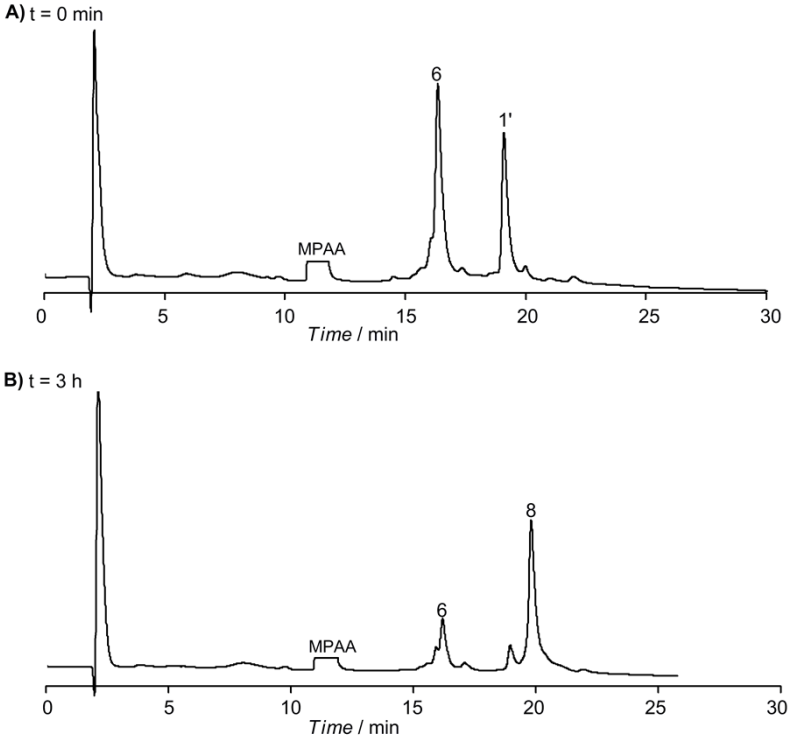


Figure S23. The ligation reaction between peptide hydrazide **1** (18.2 mg, 6 μmol, 1 equiv.) and peptide hydrazide **6** (44.1 mg, 9 μmol, 1.5 equiv.). The major procedures were the same as described above. The ligation was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). Time points: 0 min, 3 h after the addition of peptide hydrazide **6**. The peak marked with **1'** corresponds to the peptide thioester generated from peptide hydrazide **1**. The reaction was completed within 3 h. The isolated yield of **8** (24.6 mg, 3.12 μmol) was 52%. HPLC conditions: 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 C8 (4.6 × 150 mm) column.

8.4. Ligation of peptide hydrazide 8 with peptide amide 7

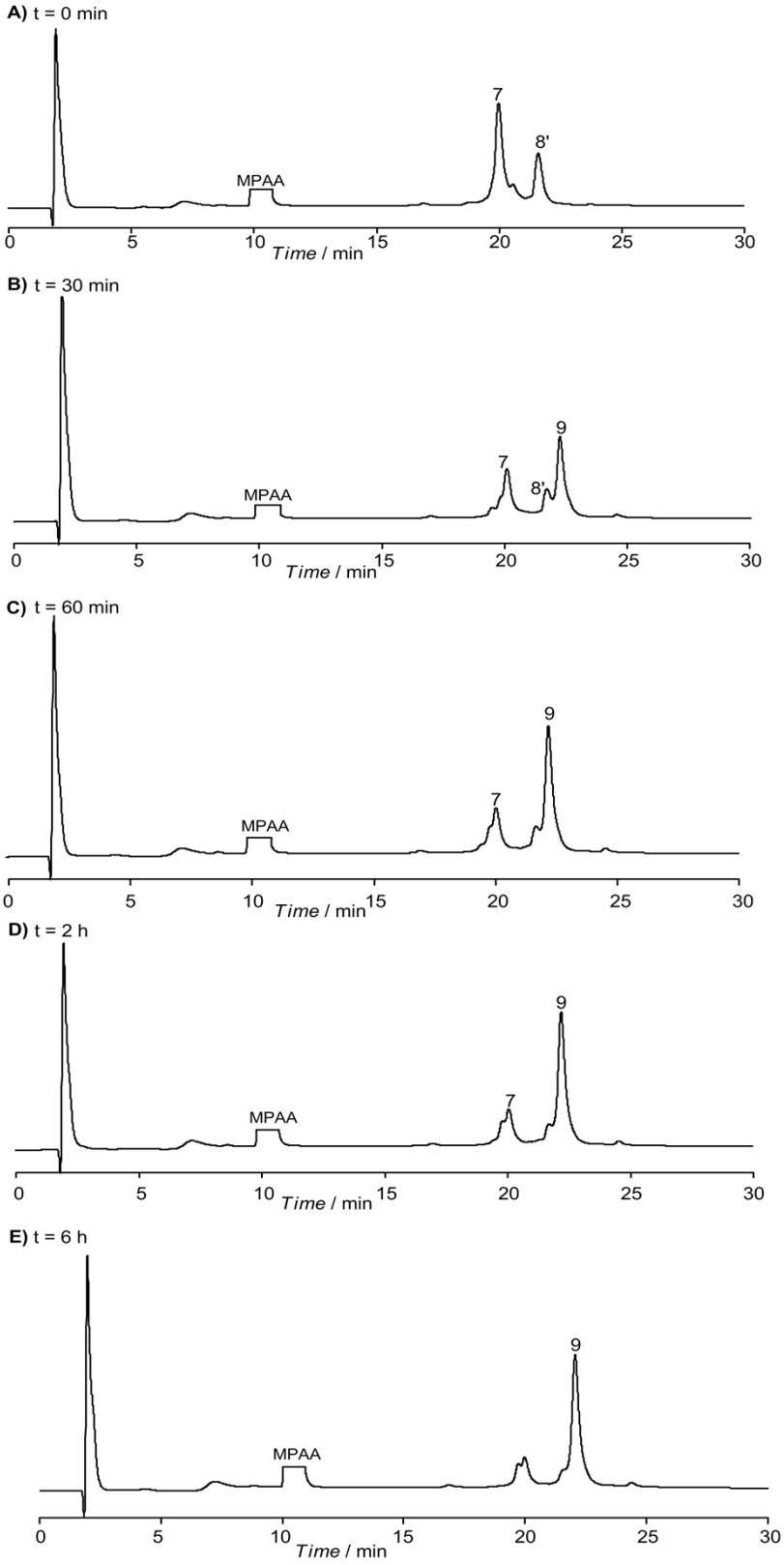


Figure S24. The ligation reaction between peptide hydrazide **8** (23.7 mg, 3 μmol, 1 equiv.) and peptide amide **7** (30.4 mg, 6 μmol, 2 equiv.). The major procedures were the same as described

above. The ligation was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). Time points: 0 min, 30 min, 60 min, 2 h and 6 h after the addition of peptide amide **7**. The peak marked with **8'** corresponds to the peptide thioester generated from peptide hydrazide **8**. The reaction was completed within 6 h. The isolated yield of **9** (16.7 mg, 1.29 μmol) was 43%. HPLC conditions: 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 C8 (4.6 × 150 mm) column.

8.5. Ligation of peptide hydrazide 11 with peptide hydrazide 12

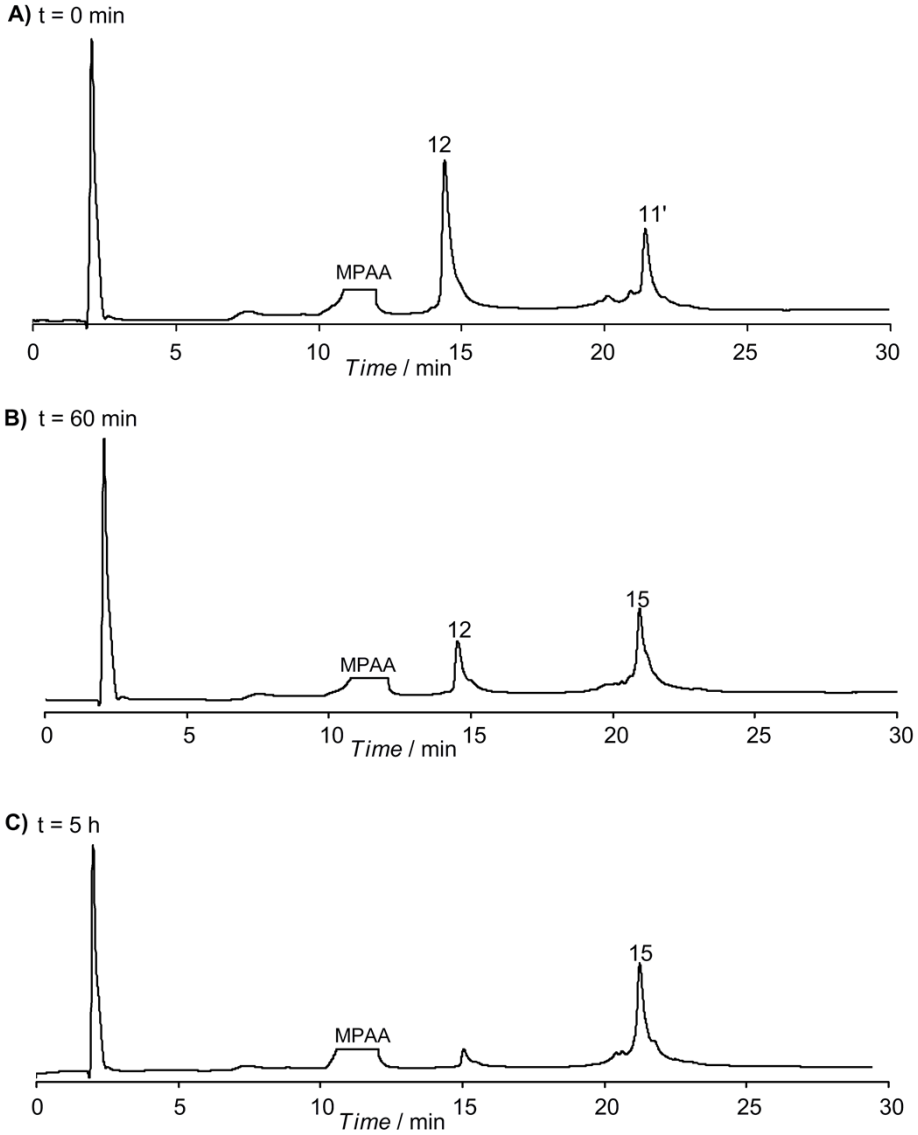


Figure S25. The ligation reaction between peptide hydrazide **11** (21.9 mg, 5 μmol, 1 equiv.) and peptide hydrazide **12** (29.5 mg, 10 μmol, 2 equiv.). The major procedures were the same as

described above. The ligation was monitored by analytical RP-HPLC chromatograms (λ = 214 nm). Time points: 0 min, 60 min and 5 h after the addition of peptide hydrazide **12**. The peak marked with **11'** corresponds to the peptide thioester generated from peptide hydrazide **11**. The reaction was completed within 5 h. The isolated yield of **15** (15.7 mg, 2.15 μmol) was 43%. HPLC conditions: 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 C8 (4.6 × 150 mm) column.

8.6. Ligation of peptide hydrazide 16 with peptide amide 13

The ligation reaction between peptide hydrazide **16** (11.1 mg, 1.5 μmol, 1 equiv.) and peptide amide **13** (16.2 mg, 3 μmol, 2 equiv.). The major procedures were the same as described above. The ligation was monitored by analytical RP-HPLC chromatograms (λ = 214 nm) (Figure 3 in main text). Time points: 0 min, 60 min and 6 h after the addition of peptide amide **13**. The peak marked with **16'** corresponds to the peptide thioester generated from peptide hydrazide **16**. The reaction was completed within 6 h. The isolated yield of **17** (7.9 mg, 0.62 μmol) was 41%. HPLC conditions: 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 C8 (4.6 × 150 mm) column.

9. Substitution reaction between peptide 15 and compound 14

The purified peptide PR[Pro^{1}−Ala^{66}]-$NHNH_2$ **15** (21.9 mg, 3 μmol, 1 equiv., final concentration 0.2 mM) containing only one Cys residue was dissolved in aqueous buffer containing 6 M Gn·HCl and 0.2 M Na_2HPO_4 (pH = 7.0). Then compound **14** (9.3 mg, 45 μmol, 15 equiv.) was added into the solution. The substitution reaction process was monitored by analytic RP-HPLC and ESI-MS. After stirred at room temperature for 60 min, the substitution reaction between the thiol group of Cys^{41} residue and compound **14** was completed. Subsequently, sodium 2-mercaptoethanesulfonate (MESNa) (7.9 mg, 48 μmol, 16 equiv.) was added to quench the excess of compound **14**. The product was purified by semi-preparative RP-HPLC, furnishing the site-selective azide-labeled peptide PR[Pro^{1}−Cys(N_3)^{41}−Ala^{66}]-$NHNH_2$ **16** (12.3 mg, 1.66 μmol) in 55% isolated yield. The reaction was monitored by analytical RP-HPLC chromatograms (λ = 214 nm) (Figure 3 in main text). Time points: 0 min and 60 min after the addition of compound **14**. HPLC conditions: 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 C8 (4.6 × 150 mm) column.

10. Protein folding

Purified PR[Pro^{1}−Cys(N_3)^{41}−Arg^{115}]-NH_2 **17** (6.4 mg, 0.5 μmol) was dissolved in aqueous 0.2 M Na_2HPO_4 solution containing 6 M Gn·HCl (pH 7.4) at a concentration of 0.9 mg/mL inside a dialysis bag (Molecular Weight Cut-Off = 3500). Then the protein was folded by step gradient dialysis against sodium acetate buffers at 4 °C. The first dialysis was against 1.6 L of 50 mM sodium acetate buffer (pH 5.6) for 3 h. The second dialysis was against 1.6 L of 25 mM sodium acetate buffer (pH 5.6) for 2 h. The last dialysis was against 2 L of 10 mM sodium acetate buffer (pH 5.6) overnight at 4 °C. The final protein concentration was 0.33 mg/mL (determined by absorbance at 280 nm using a calculated extinction coefficient of 25120 M^{-1} cm^{-1})

Figure 3E in main text. Analytic RP-HPLC chromatogram (λ = 214 nm) of protein folding solution (after the overnight dialysis, crude folding solution). HPLC conditions: 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 C8 (4.6 × 150 mm) column.

11. Substrate hydrolysis by synthetic azide-labeled HIV-1 protease

The synthetic substrate peptide **19** (0.2 mg, 0.2 μmol) was dissolved in 1 mL of 50 mM NaOAc, 0.2 M NaCl buffer with 1% (v/v) DMSO at pH 5.6. Then, 5 μL solution of folded HIV-1 protease (1.65 μg, 0.152 nmol) was added into the solution. The mixture was then incubated at 37 ºC. The hydrolysis process was monitored by analytic RP-HPLC (λ = 214 nm) and ESI-MS. RP-HPLC conditions: a linear gradient of 5-70% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min on a Vydac C8 (4.6 × 150 mm) column.

12. Substrate hydrolysis by the commercial HIV-1 protease

To examine that substrate peptide **19** could be hydrolyzed by natural HIV-1 protease through the aforementioned method, we conducted the substrate hydrolysis assay using a commercial HIV-1 protease. According to the RP-HPLC analysis, peptide **19** was mostly hydrolyzed within 30 minutes by the commercial HIV-1 protease.

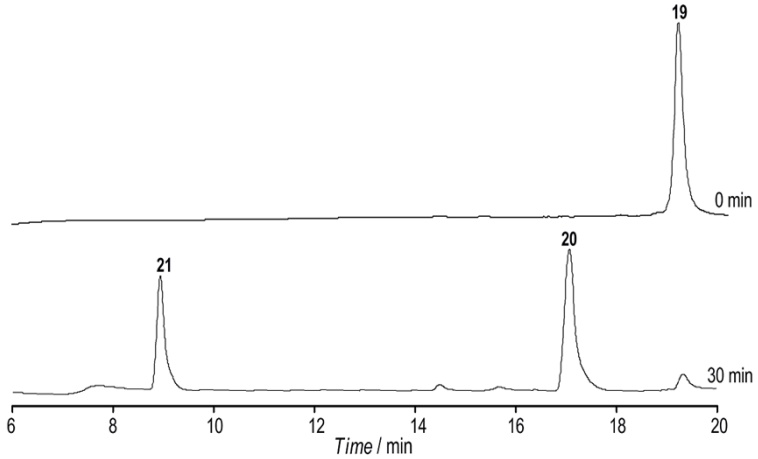


Figure S26. Substrate hydrolysis by a commercial HIV-1 protease. The hydrolysis process was monitored by analytic RP-HPLC (λ = 214 nm). RP-HPLC conditions: a linear gradient of 5-70% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min on a Vydac C8 (4.6 × 150 mm) column. Time points: 0 min and 30 min after the incubation.

13. Characterization of peptides 4, 15, 16, 17, 18

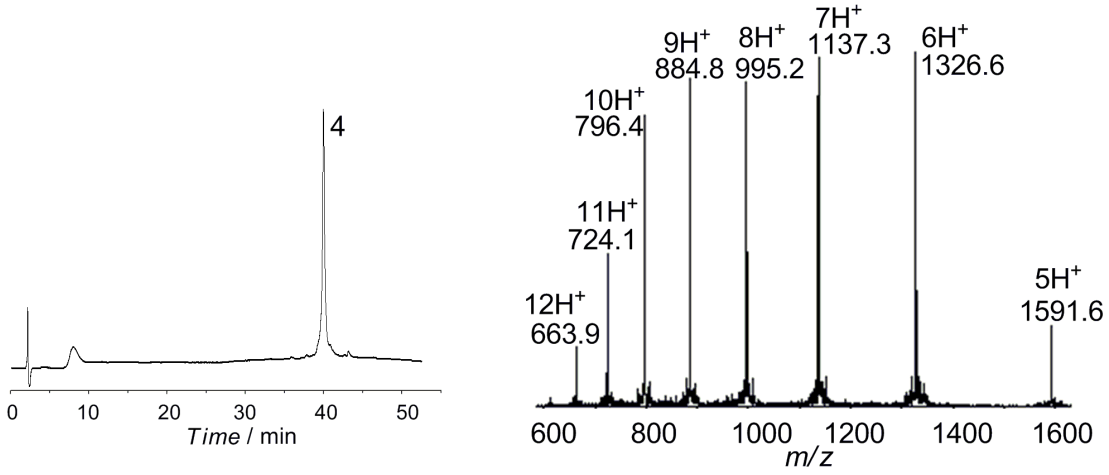


Figure S27. Characterization of **4,** the ligation product of **1** with **2**, by analytical RP-HPLC chromatogram (λ = 214 nm) and ESI-MS. HPLC conditions: a linear gradient of 20-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 50 min (5% for 2 min, then 20% for 20 min, then 20%-60% for 30 min) on a Vydac C8 (4.6 × 150 mm) column. The ESI-MS spectrum gave an observed mass of 7954.2 Da (calcd 7955.2 Da, average isotopes).

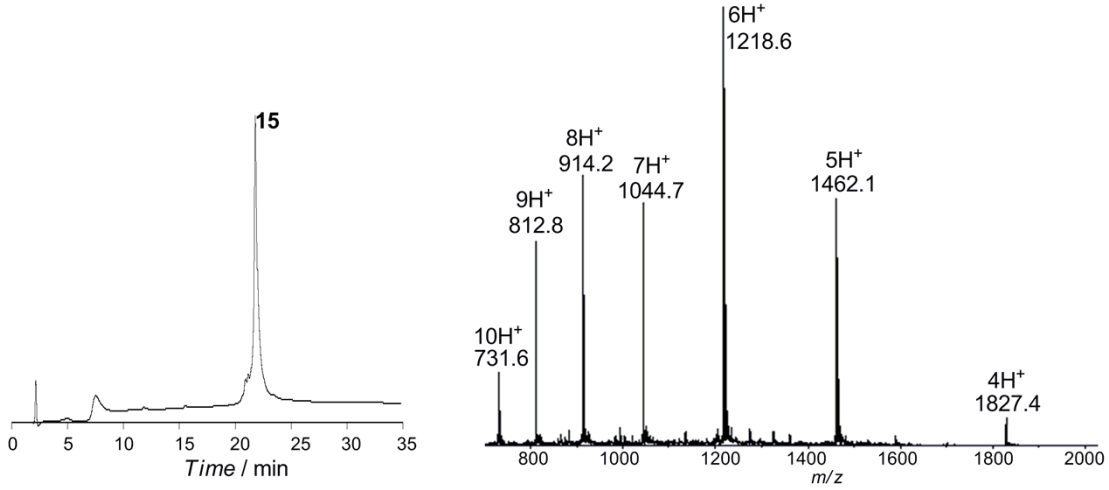


Figure S28. Characterization of **15,** the ligation product of **11** with **12**, by analytical RP-HPLC chromatograms (λ = 214 nm) and ESI-MS. HPLC conditions: a linear gradient of 20-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-

60% for 30 min) on a Vydac C8 (4.6 × 150 mm) column. The ESI-MS spectrum gave an observed mass of 7305.8 Da (calcd 7305.5 Da, average isotopes).

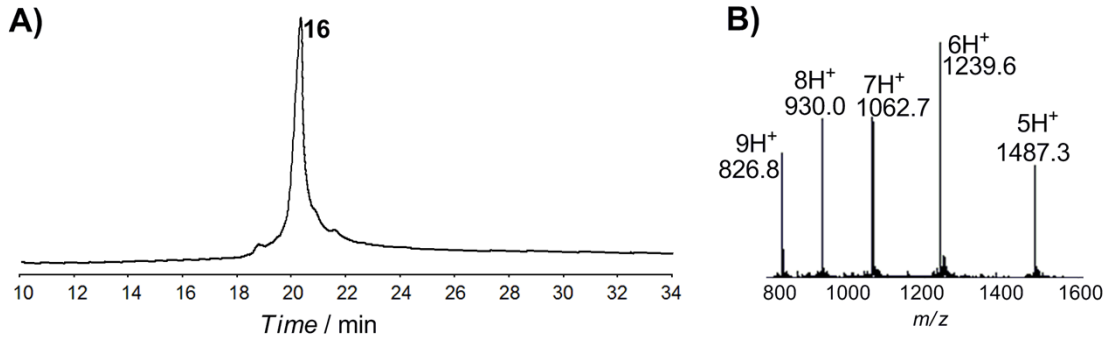


Figure S29. Characterization of **16**, the reaction product of **14** with **15**, by analytical RP-HPLC chromatograms (λ = 214 nm) and ESI-MS. HPLC conditions: a linear gradient of 20-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-60% for 30 min) on a Vydac C8 (4.6 × 150 mm) column. A) Analytical RP-HPLC chromatogram of purified **16**. B) The ESI-MS spectrum of **16** gave an observed mass of 7431.8 Da (calcd 7431.5 Da, average isotopes).

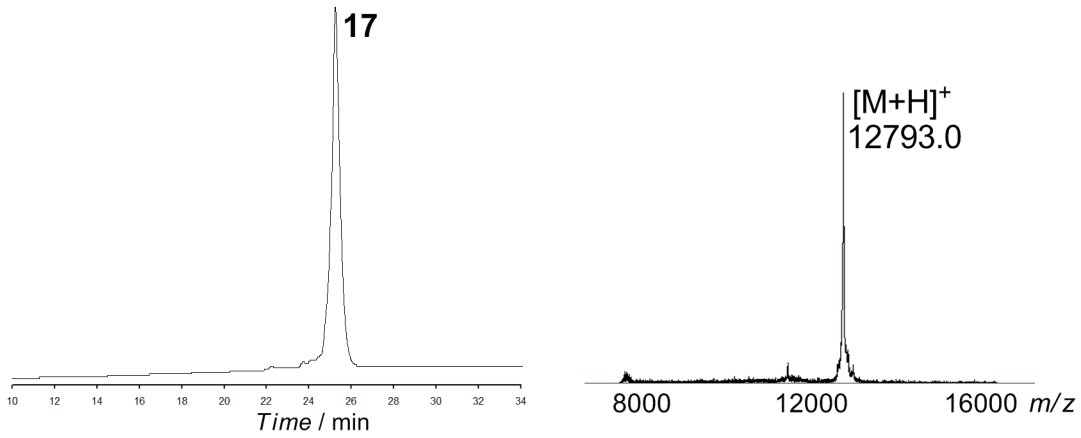


Figure S30. Characterization of **17,** the ligation product of **16** with **13**, by analytical HPLC chromatograms (λ = 214 nm) and MALDI-TOF/MS. HPLC conditions: a linear gradient of 20-60% acetonitrile (with 0.08-0.1% TFA) in water (with 0.1% TFA) over 30 min (5% for 2 min, then 20%-60% for 30 min) on a Vydac C8 (4.6 × 150 mm) column. MALDI-TOF/MS observed 12792.0 Da (calculated 12790.1 Da, average isotopes).

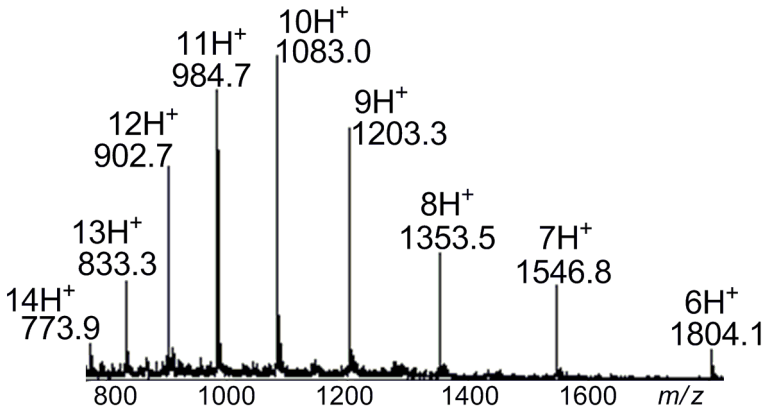


Figure S31. The ESI-MS spectrum of protein **18**. The spectrum gave an observed mass of 10820.2 Da (calcd 10820.7 Da, average isotopes).

14. ESI-MS spectra of peptides 19, 20 and 21

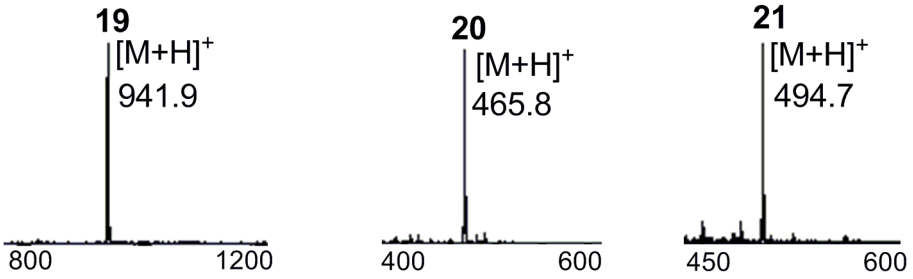


Figure S32. The ESI-MS spectra of purified **19**, **20** and **21**. For **19**, observed mass = 940.9 Da (calcd 940.1 Da, average isotopes). For **20,** observed mass = 464.8 Da (calcd 464.6 Da, average isotopes). For **21,** observed mass = 493.7 Da (calcd 493.5 Da, average isotopes).

15. References

[1] J. S. Zheng, S. Tang, Y. K. Qi, Z. P. Wang, L. Liu, *Nat. Protoc.* 2013, **8**, 2483.

[2] J. X. Wang, G. M. Fang, Y. He, D. L. Qu, M. Yu, Z. Y. Hong and L. Liu, *Angew. Chem., Int. Ed.*, 2015, **54**, 2194-2198.

[3] Q. Wan and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2007, **46**, 9248.