Supporting Information to "Artificial lantipeptides from *in vitro* translations"

Florian P. Seebeck, Alonso Ricardo and Jack W. Szostak

Synthesis of 4-Selenoisoleucine

General Methods.

Thin layer chromatography (TLC) was performed on Merck Silica Gel 60 matrix plates (analytical) and Analtech silica gel plates (preparative) containing a fluorescent indicator. Compounds were characterized by proton (^{1}H), and carbon (^{13}C) nuclear magnetic resonance spectroscopy on a Varian NMR spectrometer (Oxford AS-400). Chemical shift values are reported in ppm and referenced to the chemical shift of the solvent. Electrospray mass spectrometry were recorded on a Bruker Daltonics Esquire 3000 and 6000 ESI-MS. Solvents and other chemicals were purchased from Sigma Aldrich. Deuterated solvents for NMR were purchased from Cambridge Isotope laboratories.

Scheme S1. Synthesis of Z-protected 4-selenoisoleucine

(2*R*,3*R*)-2-(benzyloxycarbonylamino)-3-(methylselanyl)butanoic acid (2)

To a round bottom flask (100 mL) under an argon atmosphere containing dry ethanol (6 mL) was added via syringe dimethyl diselenide* (1.38g, 7.34 mmol). The resulting yellow colored biphasic solution was treated with sodium borohydride (0.15g, added in 30 mg portions) and stirred at room temperature until no more hydrogen gas evolution was observed and the solution turned uniform and colorless. At this point, a solution of (2*R*,3*S*)-2-(benzyloxycarbonylamino)-3-chlorobutanoic acid (**1**) (1.0g; 3.69 mmol) in dry tetrahydrofuran (30 mL) was added and the resulting mixture was heated at

reflux temperature under an argon atmosphere for 16 h. The solution was allowed to cool to room temperature and the volatiles were evaporated under reduced pressure. The residue was dissolved in ethyl acetate (20 mL) and extracted with a saturated solution of $NaHCO_3$ (4x15 mL). The aqueous extracts were combined, acidified using HCl (6N) until a pH value of 2 and then extracted with ethyl acetate (4x 30 mL). The organic extract was washed with water (50 mL), dried with Na_2SO_4 and the solvent evaporated under reduced pressure to obtain compound **2** as a clear colorless oil (1.08 g, 89% yield).

^{1}H-NMR (400 MHz, $CDCl_3$): δ 7.38-7.31(m, 5H, Ar), 5.613 (d, 1H; *J*= 8.8 Hz, NH), 5.122 (s, 2H, Ar-CH_2), 4.618 (dd, 1H; *J*= 8.9, 3.2 Hz, C2-H), 3.53 (m, 1H, C3-H), 2.029 (s, 3H, Se-CH_3), 1.47 (d, 3H, *J*=7.1 Hz, C4-3H)

^{13}C-NMR (100 MHz, $CDCl_3$): δ 175.85, 156.8, 136.19, 128.82, 128.55, 128.38, 67.62, 58.75, 37.36, 20.20, 4.37

ESI-MS (-): m/z calcd. 330.02, found 330

*It is recommended to use an appropriate full-face respirator while manipulating dimethyldiselenide (or methylselenide) due to its toxicity and stench, perform all manipulations, including evaporations, under a fume hood. Frequent exchange of gloves in between procedures it is also recommended to prevent contamination of surfaces

After the reactions were completed, waste solutions containing dimethyldiselenide (methylselenide) were treated with sodium hydroxide (1/10 volume; 10 M) and household bleach (1 volume) before being transferred into the chemical waste container. The glassware that has been exposed to the selenium reagent was submerged overnight in a container of household bleach before cleaning it.

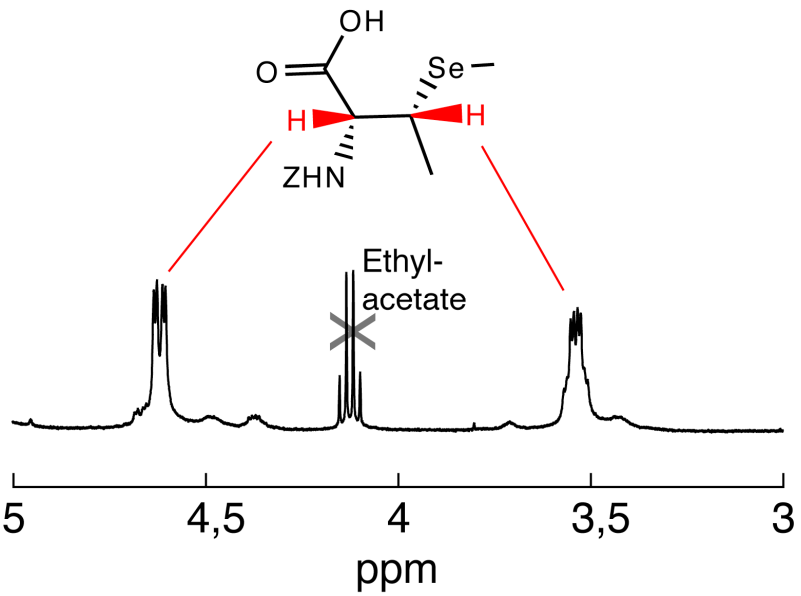


Figure S1. Stereochemical homogeneity of (2*R*,3*R*)-2-(benzyloxycarbonylamino)-3-(methylselanyl)butanoic acid (**2**). Inspection of the ^{1}H-NMR of unpurified **2** in $CDCl_3$ reveals only one set of C2-H and C3-H signals, suggesting that nucleophilic substitution (from **1** to **2**) produced only one diastereomer via a S_N2 reaction rather than by an elimination-addition mechanism. Similar stereoselectivity has been found in related reactions.[1]

Scheme S2. Synthesis of 4-selenoisoleucine hydrobromide

(2*R*,3*R*)-2-amino-3-(methylselanyl)butanoic acid hydrobromide (3)

Compound **2** (50 mg, 0.15 mmol) was dissolved in diethyl ether 1 mL and cooled on an ice bath, HBr (4 M solution in HOAc, 10 eq) was added and the solution stirred for 24 hours. The volatiles were then evaporated under vacuum and the residue was partitioned between water (2 mL) and ethyl acetate (2 mL). The aqueous fraction was washed with ethyl acetate (2x 2mL) and evaporated to dryness to yield product **3** (21 mg, 50%) as white paste.

^{1}H-NMR (400 MHz, D_2O): δ 3.90 (d, 1H; J=4.4 Hz, C2-H), 3.35 (m, 1H, C3-H), 1.86(s, 3H, -$SeCH_3$), 1.33 (d, 3H; J= 7.2 Hz, C4-3H)

^{13}C-NMR (100 MHz, D_2O): δ 171.31, 58.35, 34.02, 18.94, 3.58

ESI(-): m/z calcd. 195.99, found 196

Oxidative-elimination studies using a model peptide

Scheme S3. Synthesis of 4-selenoisolucine contaning dipeptide

(*S*)-methyl 2-((2*R*,3*R*)-2-(benzyloxycarbonylamino)-3-(methylselanyl) butanamido)-3-hydroxypropanoate (4)

Compound **2** (0.1g; 0.3 mmol) and L-serine methyl ester hydrochloride (0.056 g; 0.36 mmol) were dissolved in DMF (0.6 mL) followed by addition of di-isopropylethyl amine (0.313 mL, 6 eq), HOBT (0.049g; 0.36 mmol) and N,N-diisopropylcarbodiimide (0.056 mL; 1.2 eq) and the resulting solution was stirred overnight at room temperature. The volatiles were evaporated under vacuum and the residue dissolved in ethyl acetate (10 mL), washed with water (2x 5 mL) and brine (5 mL). The organic phase was dried with sodium sulfate and the solvent evaporated to dryness. The crude dipeptide was purified by column chromatography on silica (toluene-acetone 1:1) to yield compound **4** (97 mg, 75%) as a off-white solid.

^{1}H-NMR (400 MHz, CD_3CN): δ 7.41-7.30 (m, 5H), 6.11 (d, 1H, J= 6 Hz), 5.12 (s, 2H), 4.49 (m, 1H), 4.33(m, 1H), 3.86 (m, 1H), 3.74 (m, 1H), 3.71 (s, 3H), 3.30 (m, 1H), 2.01(s, 3H), 1.41 (d, 3H, J=7.13 Hz)

^{13}C-NMR (100 MHz, CD_3CN): δ 170.67, 169.95,128.55, 128.03,127.77, 127.01, 125.46, 116.55, 116.50, 66.4, 61.65, 58.99, 56.46, 54.73, 54.64, 51.97, 37.71, 37.32, 31.95, 31.14, 29.70, 28.77, 18.18, 2.57.

ESI-MS (+): m/z calcd. 433.09, found 433

Scheme S4. Formation of E-Dhb containing dipeptide

(*S*,*E*)-methyl 2-(2-(benzyloxycarbonylamino)but-2-enamido)-3-hydroxy propanoate (5)

Dipeptide **4** (20 mg) was dissolved in deuterated acetonitrile (0.7 mL), treated with hydrogen peroxide (4x 2 uL of a 30 % hydrogen peroxide solution in water, every 15 min) and the reaction progress monitored by ^{1}H-NMR (Figure S2). After one hour, the solution was transferred to a round bottom flask (5 mL) and evaporated to dryness. The oxidation product **5** was purified by preparative TLC on silica (Toluene:acetone 1:1) (11mg, 65% recovery).

^{1}H-NMR (400 MHz, CD_3CN): δ 7.41-7.33 (5H, m), 7.1 (1H, bs), 5.84 (1H, q, 7.2 Hz), 5.08 (2H, s), 4.53 (1H, m), 3.82 (2H, m), 3.70 (3H, s), 3.40(1H, m), 1.90 (3H, d, 7.2 Hz)

^{13}C-NMR (100 MHz, CD_3CN): δ 171.03, 164.63, 154.90, 136.99, 130.31, 128.79, 128.34, 128.24, 128.13, 66.78, 61.84, 54.98, 52.21, 25.52, 22.77, 12.88

ESI-MS (+): m/z calcd. 337.14, found 337

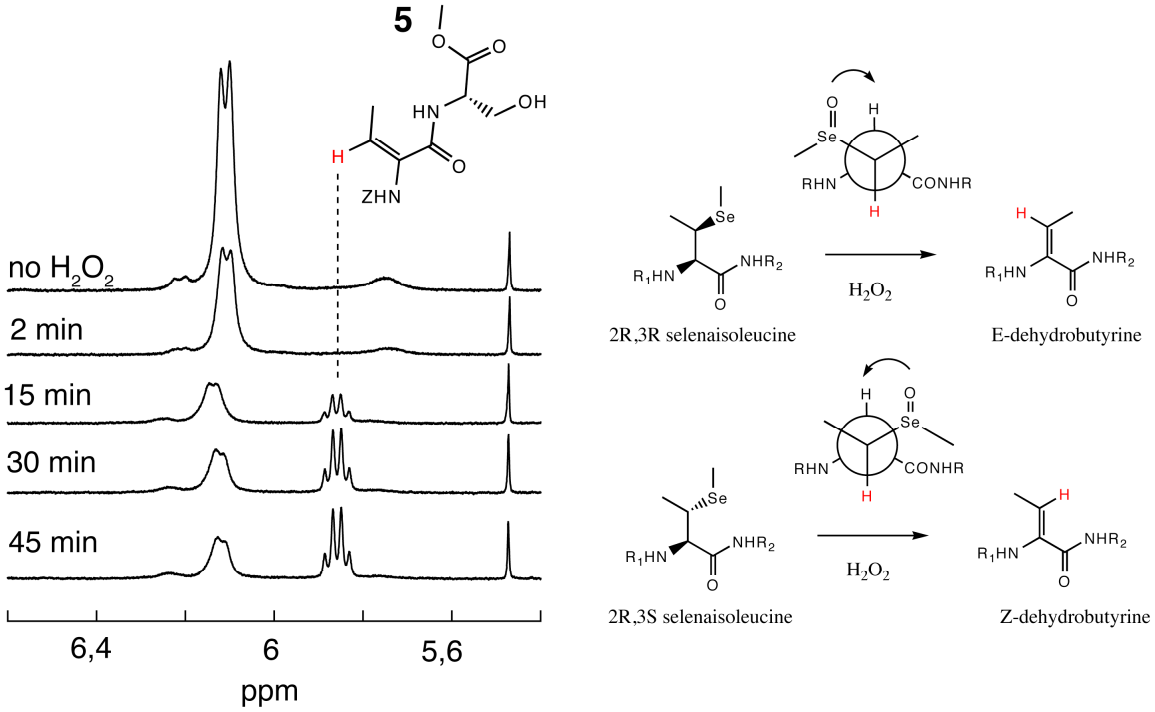
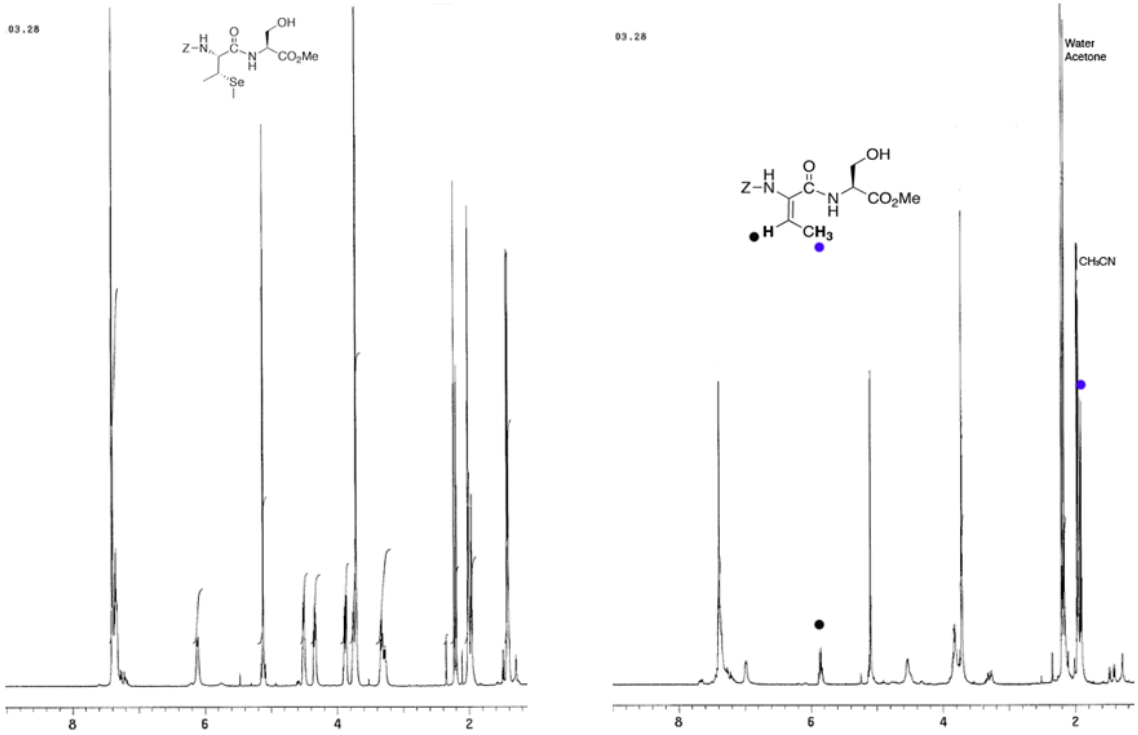


Figure S2. Right: ^{1}H-NMR monitoring of conversion of compound **4** to dipeptide **5**. Treatment of **4** with H_2O_2 induced the appearance of a quartet signal in the olefinic region consistent with the formation of the unsaturated Dhb side chain. A previous comparison of peptides containing either Z- or E-Dhb reported that the chemical shifts for the olefinic protons differ by 0.3 ppm.[1] Conversion of **4** to **5** produces only one detectable olefinic proton signal suggesting that only one Dhb-isomer is formed (the minor isomer amounts to less than 5 %). **Left:** Assignment of Dhb stereochemistry. Selenoxides are known to undergo syn-elimination which predicts that 2R,3R-selenaisoleucine eliminates to E-Dhb.[2] In principle a two-step reaction including selenoxide γ-elimination to form vinylglycine followed by isomerization could lead to the more stable Z-Dhb. Isomerization of vinylglycine does proceed at 95°C within 30 min but is very slow at room temperature.[3] We would therefore expect vinyl glycine to accumulate during the conversion of **4** to **5** under our reaction conditions (room temperature), but we could not detect the corresponding ^{1}H NMR signals. We therefore conclude that **5** contains E-Dhb.

Figure S3. ^{1}H-NMR spectra of compound **4** and E-Dhb containing dipeptide **5.**



***In vitro* peptide translation**

mRNA Templates

DNA templates for *in vitro* transcription of peptide coding mRNAs were prepared from synthetic oligonucleotides (Table S1). Templates for peptides discussed in Figs 1 and 3 were constructed by ligating complementary oligonucleotides (ILESE1S and ILESE1A; ILESE4S and ILESE4A, Table S1) to Nde1/Xho1 digested pMG209.[4] The resulting plasmids were transformed into the *E. coli* strain TOP10. Monoclonal plasmids were isolated, sequenced and used as PCR templates. DNA templates for peptides discussed in Fig. 2 were generated by PCR amplification of the synthetic oligonucleotides (ILESE2S and ILESE2,3A; ILESE3S + ILESE2,3A, Table S1). The DNA template for the C-terminal domain of T4 fibritin was amplified from the oligonucleotides FIB1-4. The amplified DNA-fragment was Nde1/Xho1 digested and ligated into pMG209.

For mRNA production, PCR amplified DNA was used as template for standard T7 RNA polymerase *in vitro* transcription reactions.[5] The full lengths transcripts were purified by preparative polyacrylamide gel electrophoresis.

Peptide translation and modification

The composition of the reconstituted *E. coli* translation system has been described previously.[6, 7] Translation reactions (100 µl) were typically incubated for 1 h at 37°C. The reactions were diluted with 200 µl wash buffer (50 mM Tris-HCl, 300 mM NaCl, 0.2 mM TCEP, pH 8) and supplemented with 100 µl suspended NTA-agarose beads (Quiagen). After incubation (10 min) the agarose beads were treated with wash buffer and with the following modification buffers at room temperature. A) Cysteine protection: 10 mM oxidized glutathione in wash buffer without TCEP for 20 min; B) Oxidative elimination: 50 mM H_2O_2 in wash buffer at pH 7.2 without TCEP for 1 – 2 h; C) Lanthionine formation: wash buffer for 20 min. The peptides were eluted with a 0.2 % TFA solution. For MALDI-TOF analysis peptides were desalted and concentrated by reversed phase micro-chromatography (C18 Zip Tips, Millipore) and eluted with a 70 %

acetonitrile, 0.1 % TFA solution saturated with α-cyano-4-hydroxycinnamic acid. Mass measurements were performed using an Applied Biosystems Voyager MALDI-TOF.

Table S1. Synthetic oligonucleotides

FIB1	5'- AAACATATGTGTTATATTCCTGAAGCTCCAA
FIB2	5'- TTACGAACGTAAGCTTGCCCATCTCTTGGAGCTTCAGGAATA
FIB3	5'- CAAGCTTACGTTCGTAAAGATGGCGAATGGGTATTGCTTTCTA
FIB4	5'- TTTCTCGAGAAAGGTAGAAAGCAATACCCAT
ILESE1S	5'- TATGCGTAAGATTC
ILESE1A	5'- TCGAGAATCTTACGCA
ILESE2S	5'-TAATACGACTCACTATAGGGTTAACTTTAGTAAGGAGGACAGCTAAATGAA GAGTCGTTTCGGTTGCAGTCGTTTCGGTATTGGTCATCATCATCATCATCAT
ILESE3S	5'-TAATACGACTCACTATAGGGTTAACTTTAGTAAGGAGGACAGCTAAATGAT TAGTCGTTTCGGTTGCAGTCGTTTCGGTAAGGGTCATCATCATCATCATCAT
ILESE2,3A	5'-CTAGCTACCTAATGATGATGATGATGATGACC
ILESE4S	5'- TATGATTAAGGTTATTCTCTGCAAACCTGGTTGCCGTC
ILESE4A	5'- TCGAGACGGCAACCAGGTTTGCAGAGAATAACCTTAATCA

Tryptic digest

Peptides immobilized on 100 ul NTA agarose were incubated with 10 mM Tris-HCl pH 7.6, 20 mM CaCl2 and 0.5 ug trypsin in 200 ul for 1 h at 37°C. The resin was washed with 0.5 ml 10 mM Tris-HCl pH 7.6, 300 mM NaCl and with 0.5 ml 10 mM Tris-HCl pH 7.6. The peptide fragments were eluted in 0.2 % TFA.

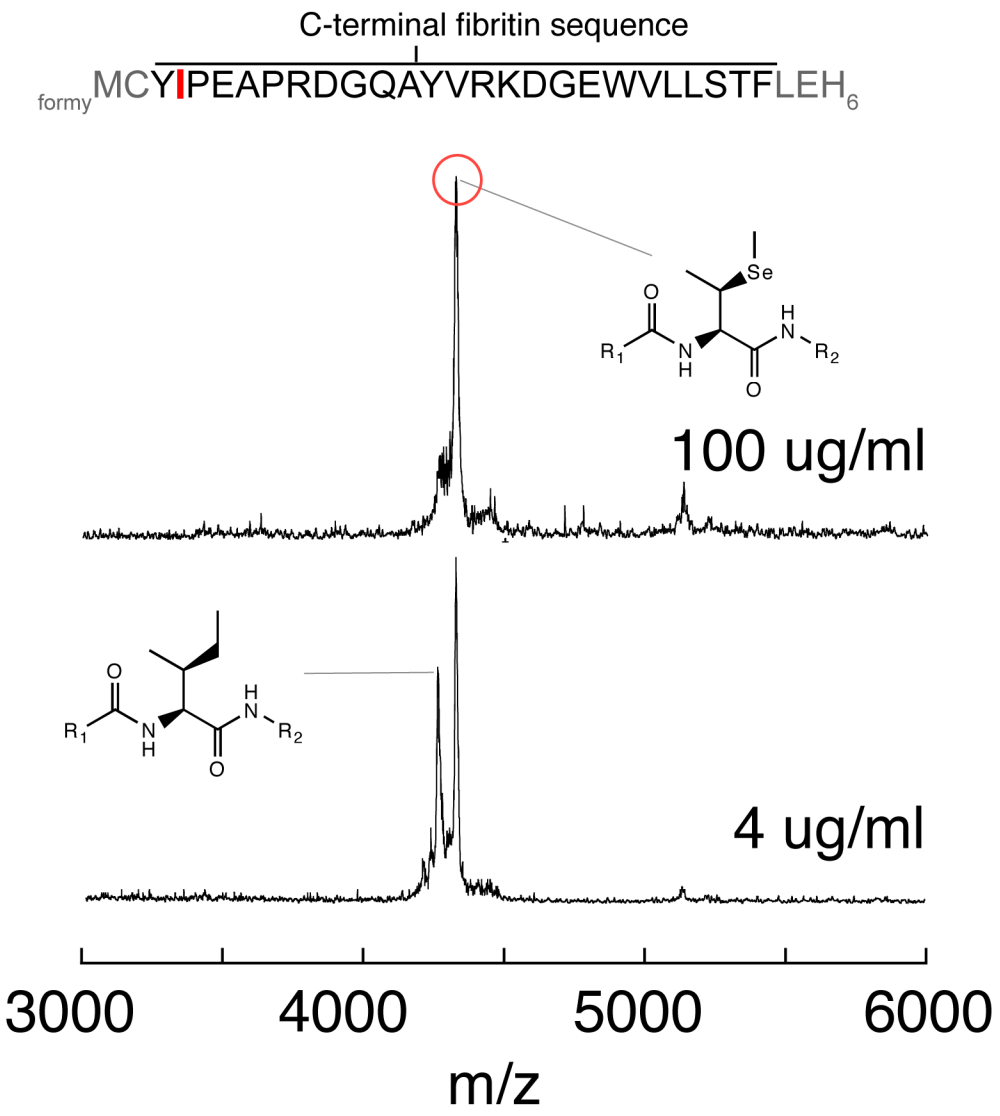


Figure S4. 4-Selenoisoleucine is a specific isoleucine substitute. The C-terminal domain of T4 fibritin was translated in reaction mixture with 4-selenoisoleucine as a substitute for isoleucine (fibritin : m/z calcd. 3237.0, found. 4238; seleno-fibiritin: m/z calcd. 4303.0, found.4304). Traces of endogenous isoleucine compete with 4-selenoisoleucine present at a concentration of 4 ug/ml in the translation reaction. In a reaction containing 100 ug/ml 4-selenoisoleucine this competition is suppressed.

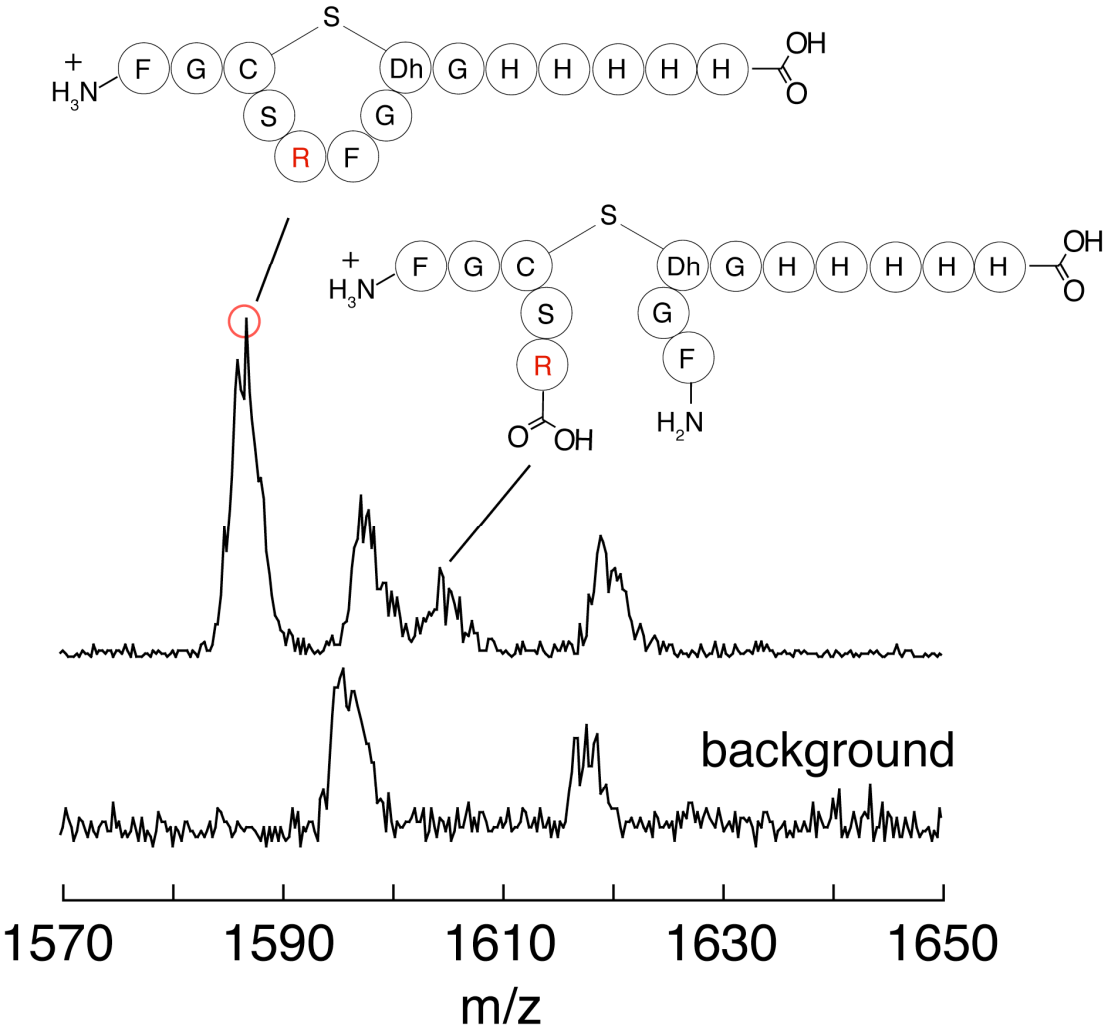


Figure S5. Tryptic digest of peptide 2 (see Fig. 2) produces two C-terminal fragments resulting from cleavage after arginine 4 and partial cleavage after arginine 12.

1. H. Zhou and W. A. van der Donk, *Org. Lett.*, 2002, **4**, 1335-1338.
2. K. B. Sharpless, M. W. Young and R. F. Lauer, *Tet. Lett.*, 1973, 1979-1982.
3. Y. Goto, K. Iwasaki, K. Torikai, H. Murakami and H. Suga, *Chem. Commun.*, 2009, 3419-3421.
4. M. Gamper, D. Hilvert and P. Kast, *Biochemistry*, 2000, **39**, 14087 - 14094.
5. J. F. Milligan and O. C. Uhlenbeck, *Methods Enzymol.*, 1989, **180**, 51 - 62.
6. F. P. Seebeck and J. W. Szostak, *J. Am. Chem. Soc.*, 2006, **128**, 7150-7151.
7. K. Josephson, M. C. Hartman and J. W. Szostak, *J. Am. Chem. Soc.*, 2005, **127**, 11727-11735.