Harnessing selenocysteine reactivity for oxidative protein folding

Norman Metanis and Donald Hilvert

Supporting Information

Materials and Methods

Protein Sequence. The amino acid sequence of wt-BPTI is: $R^{1}PDFC^{5}LEPPYTGP$-$C^{14}KARIIRYFYNAKAGL^{29}C^{30}QTFVYGG^{37}C^{38}RAKRNNFKSAEDC^{51}MRTC^{55}GGA^{58}$. The synthetic wt BPTI protein and selenocysteine-containing analogs used in this study contain norleucine in place of Met52 to prevent undesired oxidative side reactions.

Reagents. Buffers for kinetic measurements were prepared using de-ionized water. KH_2PO4 and K_2HPO4 and 2-amino-2-hydroxymethyl-propane-1,3-diol (Tris) were purchased from Fisher Biotech. Bovine pancreatic trypsin inhibitor (BPTI) was a gift from Bayer AG and Bachem AG. All Boc-amino acids were obtained from Peptides International (Louisville, Kentucky), with the following side-chain protecting groups: Arg(Tos), Asp(OcHxl), Asn(Xan), Cys(MeBzl), Glu(OcHxl), Ser(Bzl), Thr(Bzl), Tyr(Br-Z), Lys(2ClZ), (OcHxl = cyclohexyl; Bzl = benzyl; MeBzl = 4-methylbenzyl; 2ClZ = 2-chlorobenzyloxycarbonyl; Br-Z = 2-bromobenzyloxycarbonyl; Xan = N-xanthyl). All Fmoc-amino acids were obtained from Novabiochem (Merck, Darmstadt, Germany) or Bachem (Switzerland) with the following side-chain protecting groups: Cys(StBu), Glu(OtBu), Lys(Boc), Thr(tBu), Ser(tBu), Asn(Trt), Gln(Trt), Arg(Pbf), Tyr(tBu) (Boc = tert-butyloxycarbonyl; Trt = trityl; tBu = tert-butyl; Pbf = 2,2,4,6,7- pentamethyldihydrobenzofuran-5-sulfonyl). H-Gly-2-chlorotrityl-resin and Fmoc-Ala-Wang resin were obtained from Novabiochem. 2-(6-Chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU), 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), hydroxybenzotriazole (HOBt), and S-trityl-β-mercaptopropionic acid-Leu-OCH_2-Pam (TMPAL-Pam) resin were obtained from Peptides International (Louisville, KY). All solvents were HPLC-grade. *N*-Methyl-2-pyrrolidone (NMP),

N,N-dimethylformamide (DMF), dichloromethane (DCM), acetonitrile (ACN), *N,N*-diisopropylethyl amine (DIEA), and anisole were purchased from Sigma or Applied Biosystems Europe. Trifluoroacetic acid (TFA) was obtained from Halocabon Products (River Edge, NJ). Anhydrous HF was purchased from Pangas (Linde, Germany). All other chemicals were obtained from Fisher or Sigma-Aldrich, Inc.

Peptides synthesis. Synthetic BPTI analogs were prepared as previously described with minor modifications.[1]

The peptide thioesters C5U BPTI(R1–L29)-COSR and C14U BPTI(R1–L29)-COSR were prepared by manual solid-phase peptide synthesis (SPPS), typically on a 0.3 mmol scale using the *in situ* neutralization/HCTU activation procedure for Boc-SPPS.[1, 2] The trityl-protecting group of the Trt-SCH_2CH_2CO-Leu-OCH_2-PAM resin was first removed with a mixture of TFA:triisopropylsilane:H_2O–95:2.5:2.5. Peptide couplings were carried out with 2.2 mmol activated amino acid for 20 min. Coupling of Boc-Sec(4-MeBzl)-OH was carried out manually using a DIC/HOBt activation method as previously described.[1, 3] Upon completion of assembly, the polypeptides was globally deprotected and cleaved from the solid supports by treatment of the dry peptide resins (~1 g) with 12-15 mL of anhydrous HF containing 10% anisole and 2.0 equiv DTNP[4, 5] for 1 h at 0 °C. The crude peptides were precipitated and washed with cold anhydrous ether, dissolved in aqueous acetonitrile, lyophilized (~700 mg crude peptides), and purified by preparative reversed-phase HPLC (220 nm, Waters XBridge C18 column, 5 µm, 19x150 mm, 10:90 to 50:50 (acetonitrile-0.08% TFA:H_2O-0.1% TFA) over 60 min) to give the desired product with a single oxidized Se–S bond. High-resolution electrospray ionization (HR-ESI) mass spectrometry: C5U BPTI(R1–L29)-COSR calc. 3606.7502 Da, obs. 3607.7312 Da, and C14U BPTI(R1–L29)-COSR calc. 3606.7502 Da, obs. 3606.7264 Da.

Peptide BPTI(Thz30–37)-COSR was prepared previously.[1]

Peptide BPTI(Sec38–A58) was synthesized by machine-assisted Fmoc-SPPS (ABI 433A Applied Biosystems) on Fmoc-Ala-Wang resin, on 0.21 mmol scale, using 1 mmol amino

acid and the HOBt/HBTU activation procedure. Coupling of Boc-Sec(Mob)-OH was carried out manually using a DIC/HOBt activation method as in the case with Boc-Sec(4-MeBzl)-OH. Upon completion, the peptide was deprotected and cleaved from the support by treatment of the dry resin with 20 mL cleavage cocktail (TFA:H_2O:1,2-ethanedithiol:triisopropylsilane–95:2.5:2.5:1 with 2.0 equiv DTNP)[4, 5] for 2.5 h at room temperature in a syringe with a frit. The resin was removed and cold Et_2O was added, the precipitate peptide was removed by centrifugation, dissolved in 50% ACN and lyophilized (450 mg crude peptide), followed by purification with preparative RP-HPLC (220 nm, Waters XBridge C18 column, 5 µm, 19x150 mm, 10:90 to 50:50 (acetonitrile-0.08% TFA:H_2O-0.1% TFA) over 40 min) to give the desired product containing a single oxidized Se–S bond (mixture of isomers), mass calc. 2345.48 Da, obs. 2346.19 Da.

Native chemical ligations (NCL). The C38U BPTI(C30–A58) was prepared as in the case of BPTI(C30–A58)[1] by NCL reaction between BPTI(Thz30–G37)-COSR and BPTI(U38–A58) at room temperature for 20 h with periodic vortexing and monitoring by analytical RP-HPLC and LC-MS. Upon completion, thiols were extracted with Et_2O, followed by addition of NH_2OMe (2 M in water) at pH 4. The conversion of Thz to Cys was completed within 4 h. The desired peptide, containing a single oxidized Se–S bond (mixture of isomers), was purified by HPLC. ESI-MS: calc. 3201.44 Da, obs. 3202.00 Da.

For the second NCL reactions between C5U BPTI(R1–L29)-COSR and BPTI(C30–A58) and between C14U BPTI(R1–L29)-COSR and C38U BPTI(C30–A58) were performed as described for the preparation of C5U/C14U BPTI[1]. The ligations, which gave C5U BPTI and C14U/C38U BPTI, respectively, were monitored by analytical RP-HPLC and LC-MS.

Upon completion of the ligation reaction yielding **C14U/C38U BPTI**, the thiol additives were extracted with Et_2O, followed by purification by analytical RP-HPLC (multiple injections). C14U/C38U BPTI eluted as a single peak with a retention time of 24.6 min (220 nm, Waters XBridge C18 column, 3.5 µm, 4.6x150 mm, 10:90 to 50:50 (acetonitrile-0.08% TFA:water-0.1% TFA) over 40 min). The deconvoluted high-resolution dual MALDI/ESI-MS (calc.

6591.0148 Da, obs. 6591.0589 Da) and isotopic pattern corresponded to a protein containing a single oxidized Se–Se bond and four reduced thiols (Figure S1).

C5U BPTI gave a mixture of Se–S containing isomers (broad peak in the HPLC, retention time 24-30 min), which complicated its isolation in homogeneous form. For this reason, a different approach was taken involving aerobic oxidative folding of the crude ligated mixture to give a single folded protein, which was purified by HPLC, reduced with DTT, and re-isolated. To that end, upon completion of the ligation reaction and extraction of thiol additives with Et_2O, the sample was lyophilized. The crude product was first dissolved in 10 mM HCl, then diluted to a final concentration of ~30 μM in Tris-HCl buffer (100 mM Tris-HCl, 200 mM KCl, 1 mM EDTA, pH 8.7), and folded aerobically for 3 h. After exhaustive dialysis (Fisher Scientific MWCO 1000 Da) against 10 mM HCl, the oxidized protein was lyophilized and then purified by HPLC. The deconvoluted high-resolution mass spectrum of folded C5U BPTI gave the exact mass (calc. 6540.0212 Da, obs. 6540.0282 Da, Figure S2) and isotopic pattern expected for a protein containing a single oxidized Se–S bond and two S–S bonds. After lyophilization, the protein was dissolved in phosphate buffer (200 mM sodium phosphate, 6 M GdmCl, pH 8.0) containing 50 mM DTT. After 2 h, the partially reduced protein (mixture of Se-S isomers) was desalted on a PD10 column and lyophilized. The deconvoluted high-resolution mass spectrum of C5U BPTI gave the exact mass (calc. 6544.0695 Da, obs. 6544.0678 Da, Figure S3) and isotopic pattern expected for a protein containing a single Se–S bond and four reduced thiols.

Oxidative folding of BPTI, C5U BPTI and C14U/C38U BPTI. Oxidative renaturation of BPTI analogs was monitored by analytical HPLC as previously described.[1] Unless otherwise stated, the folding reactions were carried out in degassed buffer in a glove-bag under a mixture of argon and nitrogen. Oxidized glutathione (GSSG) or selenoglutathione (GSeSeG) (final concentration = 150 μM) was added to the BPTI analog (30 μM) in Tris-HCl buffer (100 mM Tris-HCl, 200 mM KCl, 1 mM EDTA, pH 8.7). At various time intervals, 80 μL aliquots were removed, quenched with 30 μL 1 M HCl to pH <2, and stored at -20 °C prior to analysis.

The samples were injected onto an Atlantis T3 column (3 μm 4.6 x 150 mm), heated to 40 °C, and eluted with a 10:90 to 25:75 gradient (acetonitrile-0.08% TFA:water-0.1% TFA) over 8 minutes, increasing to 40:60 (acetonitrile-0.08% TFA:water-0.1% TFA) over 50 min. The chromatograms were monitored at 220 and 280 nm. For kinetic analysis, all peak areas were summed and the total was assigned a value of 100%. The peak assigned to native BPTI was confirmed by co-injection of the oxidized folding product with an authentic standard of native BPTI. The peaks of the intermediates that accumulate during folding were assigned according to the literature.

Characterization of the C5U folding intermediate. Because of the facile oxidation of selenols and the susceptibility of selenosulfides to rearrangement, the previously described acid quench method[6] could not be used to identify the intermediate that formed during folding of C5U BPTI. Instead, the partially renatured sample was treated directly with excess iodoacetamide (IAM; >1000 fold) and purified by HPLC. The major fraction containing incompletely folded protein (retention time ~22 min) had the mass expected for doubly alkylated C5U BPTI (calc. 6656.5 Da, obs. 6656.2 Da, Figure S4). It was lyophilized, reduced with DTT for 1 h, and then treated with 4-vinylpyridine (4-VP) to alkylate the previously oxidized thiols/selenol. This treatment increased the mass of the adduct as expected for addition of four pyridylethyl (PE) groups (calc. 7082.1 Da, obs. 7084.1 Da, Figure S5). After desalting on a PD-10 column and lyophilization, the modified protein was digested with trypsin (at a ratio of 100:1) for 2 h at 55 °C. The fragments were filtered through a C_{18}-ZipTip and spotted on a MALDI-MS plate for analysis. Ions for two unique peptides—(C5U BPTI(1–15) and BPTI(27–39)) containing four of the six possible Cys or Sec sites—were detected (Figure S6). Peptides BPTI(47–53) and BPTI(54–58) were never observed, probably because they were further degraded under the reaction conditions.[7]

The first mass, 1979.8 Da, corresponds to BPTI(1–15) modified with two PE groups. The selenium isotopic pattern is clearly evident in the parent ion, but not in the 1792.6 Da secondary ion, which corresponds to loss of $HSeCH_2CH_2Py$ (-187 Da) upon cleavage of the

relatively weak C–Se bond by the MALDI laser beam. The finding that Sec5 and Cys14 were both alkylated by 4-VP in the second step rather than by IAM in the first indicates that these two residues were already oxidized in the folding intermediate. The second mass, 1536.6 Da, corresponds to the BPTI(27–39) peptide containing one carbamidomethyl (CAM) and one PE group (Figure S7 and S8). Further MS/MS analysis of the 1536.6 Da ion showed that the CAM group was appended to Cys30, whereas the PE group was attached to Cys38 (Figure S9).

Five two-disulfide intermediates have been previously identified in partially folded BPTI samples (Figure S7).[8] Of these, only N* [5–55, 14–38] and the nonnative [5–14, 30–51] species would yield trypsin fragments with masses matching those that we observed experimentally for the intermediate in the C5U BPTI folding reaction. The finding that Cys30 was alkylated with iodoacetamide and Cys38 with 4-VP provides support for the N* assignment and rules out [5–14, 30–51] as the intermediate (Figure S7 and S8). Although we cannot completely exclude other nonnative two-disulfide intermediates, such as [14–38, 5–51], such species have never been observed before in BPTI folding reactions. Given that the retention time of the transient species formed during folding of C5U BPTI (Figure 3a) closely matches that of N* in the wt folding pathway (Figure 3b), we conclude that the intermediate is likely to be an N* analog.

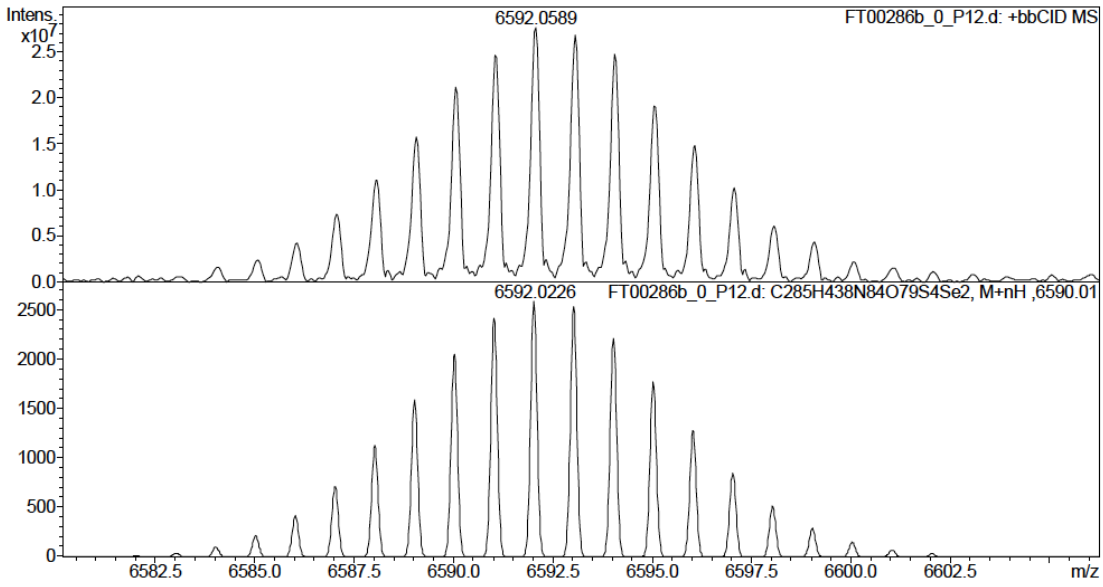


Figure S1. High-resolution dual MALDI/ESI-MS of purified C14U/C38U BPTI. The HR-MS indicates that the protein contains a single Se-Se bond between Sec14 and Sec38 plus four reduced thiols.

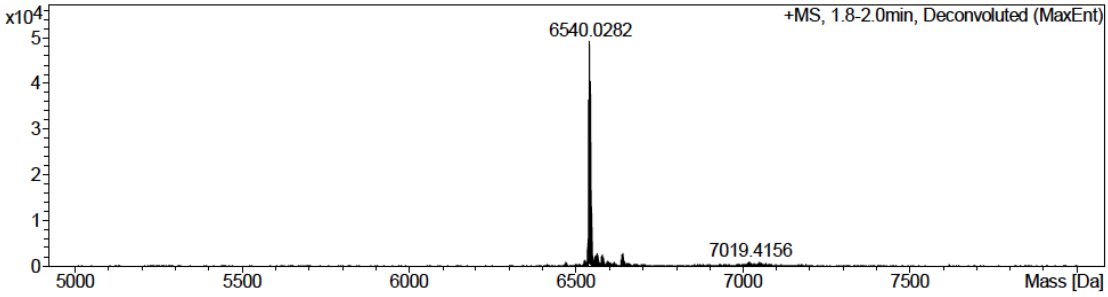


Figure S2. Deconvoluted HR ESI-MS of folded C5U BPTI. The HR-MS indicates that there are three oxidized bonds, one Se-S and two S-S crosslinks.

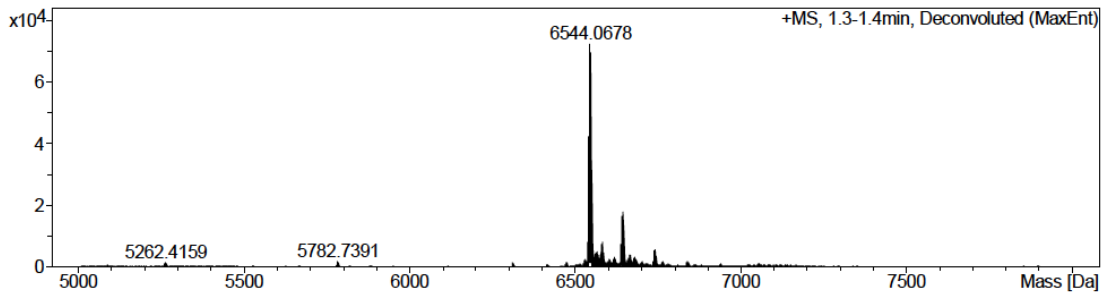


Figure S3. Deconvoluted HR ESI-MS of partly reduced C5U BPTI. The HR-MS indicates that the protein contains one Se-S bond and four reduced thiols.

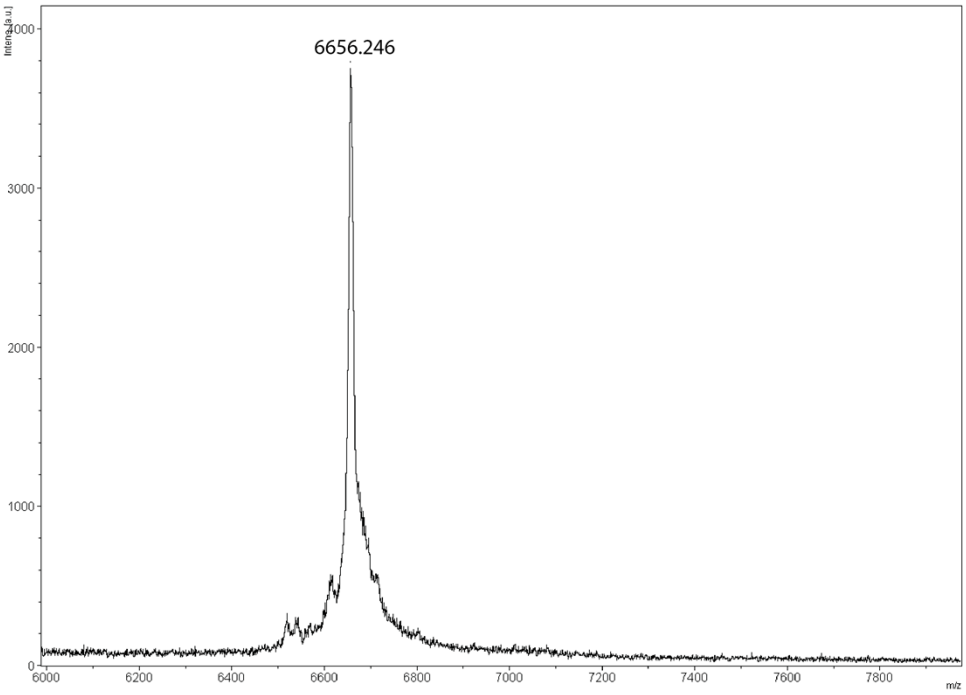


Figure S4. MALDI-MS spectrum of doubly alkylated C5U BPTI intermediate after reaction with IAM.

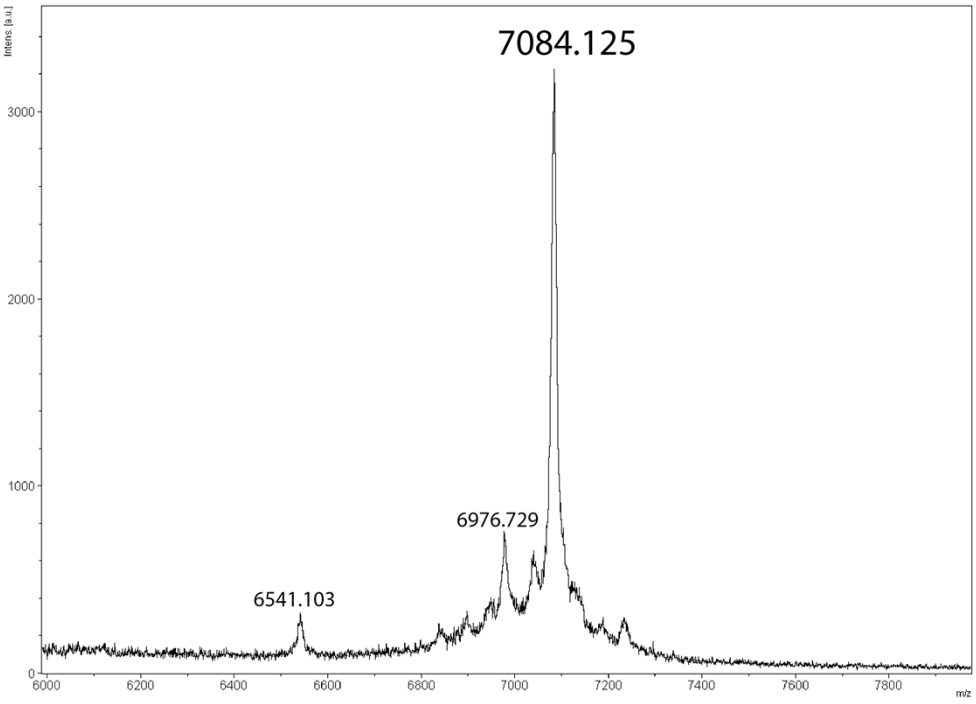


Figure S5. MALDI-MS for the C5U BPTI intermediate after reaction with IAM, lyophilization, reduction with DTT, and alkylation with 4-VP. The mass indicates the intermediate contains two CAM groups (from the reaction with IAM) and four PE groups (from the reaction with 4-VP).

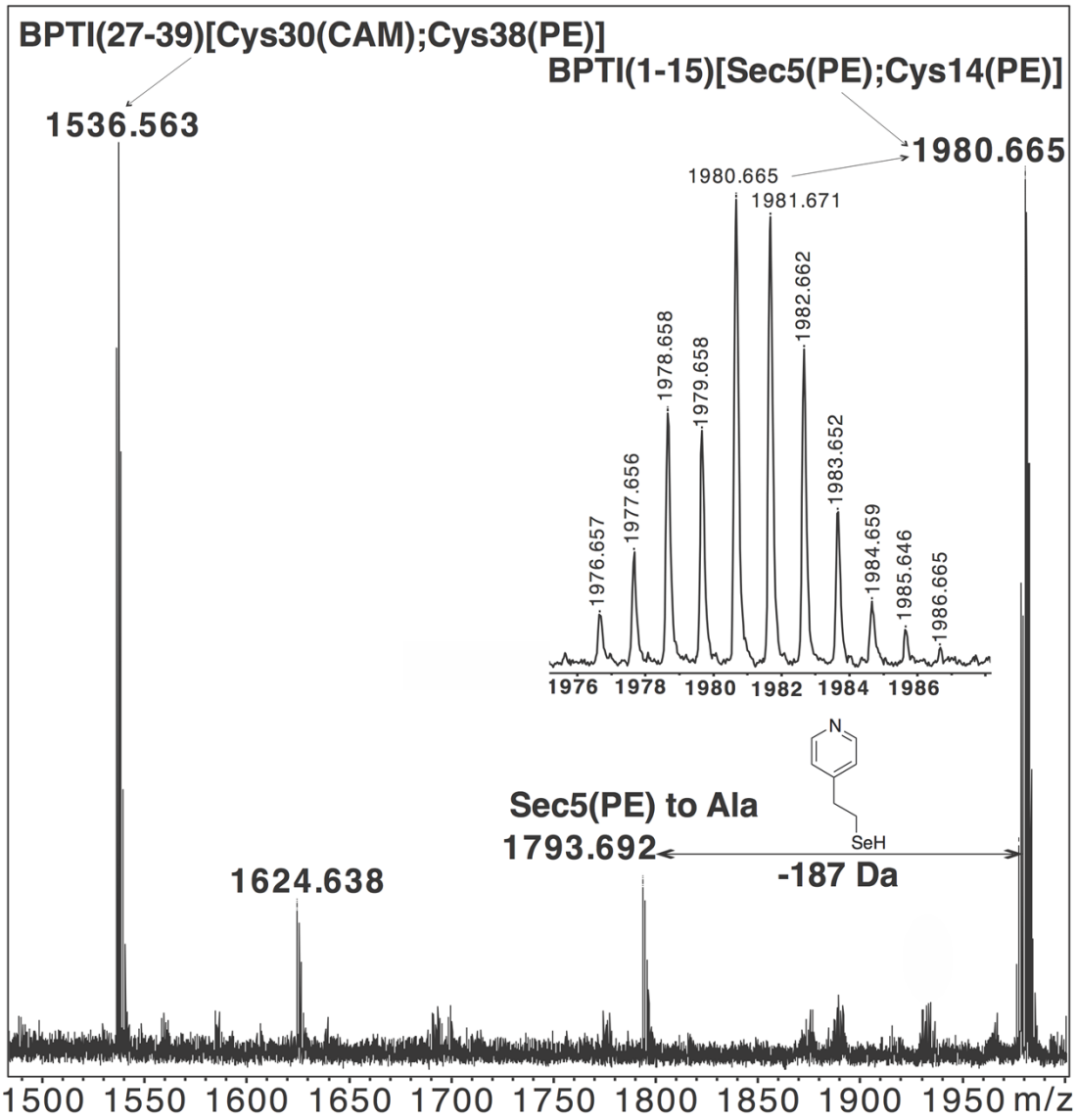
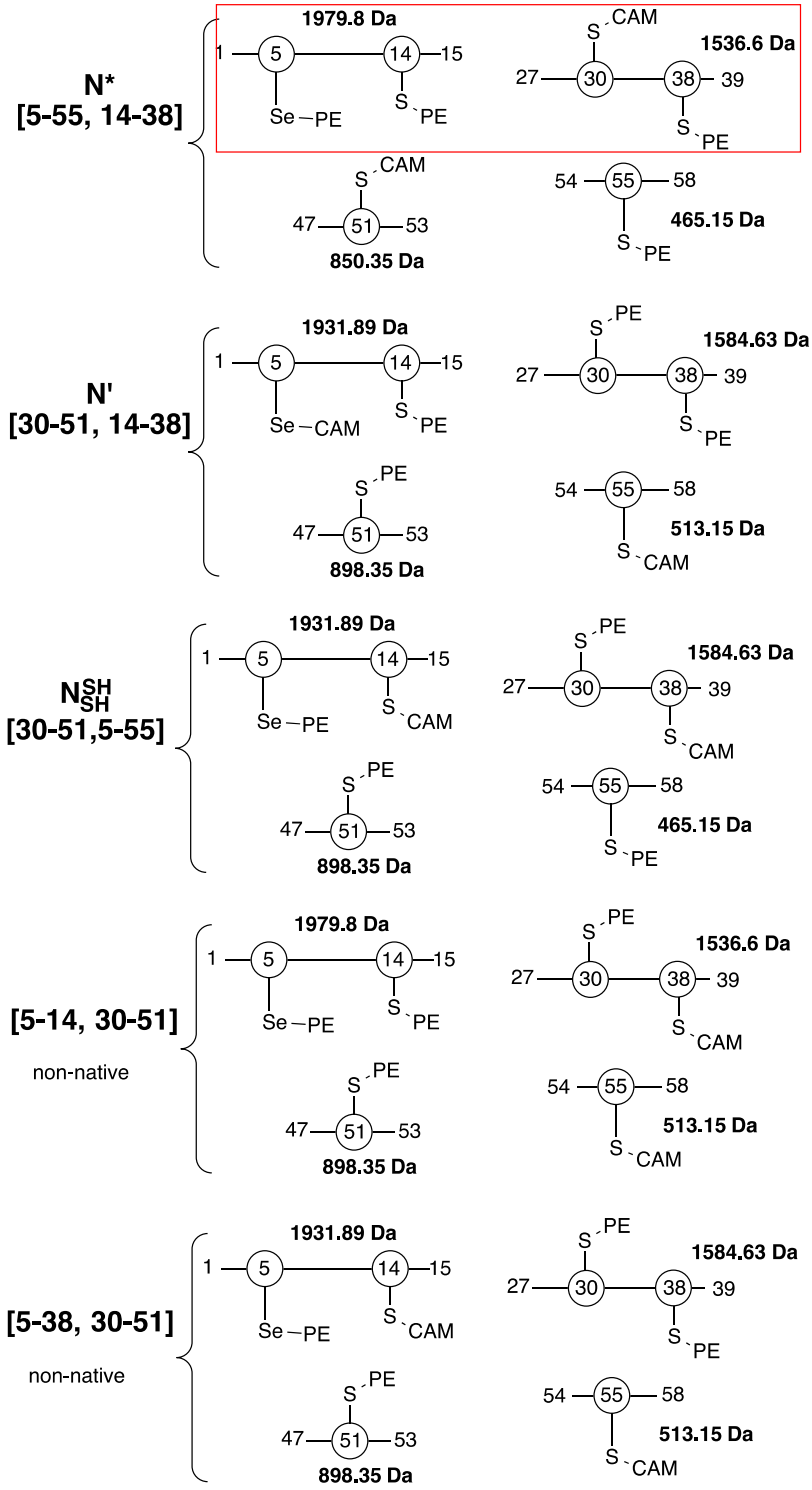
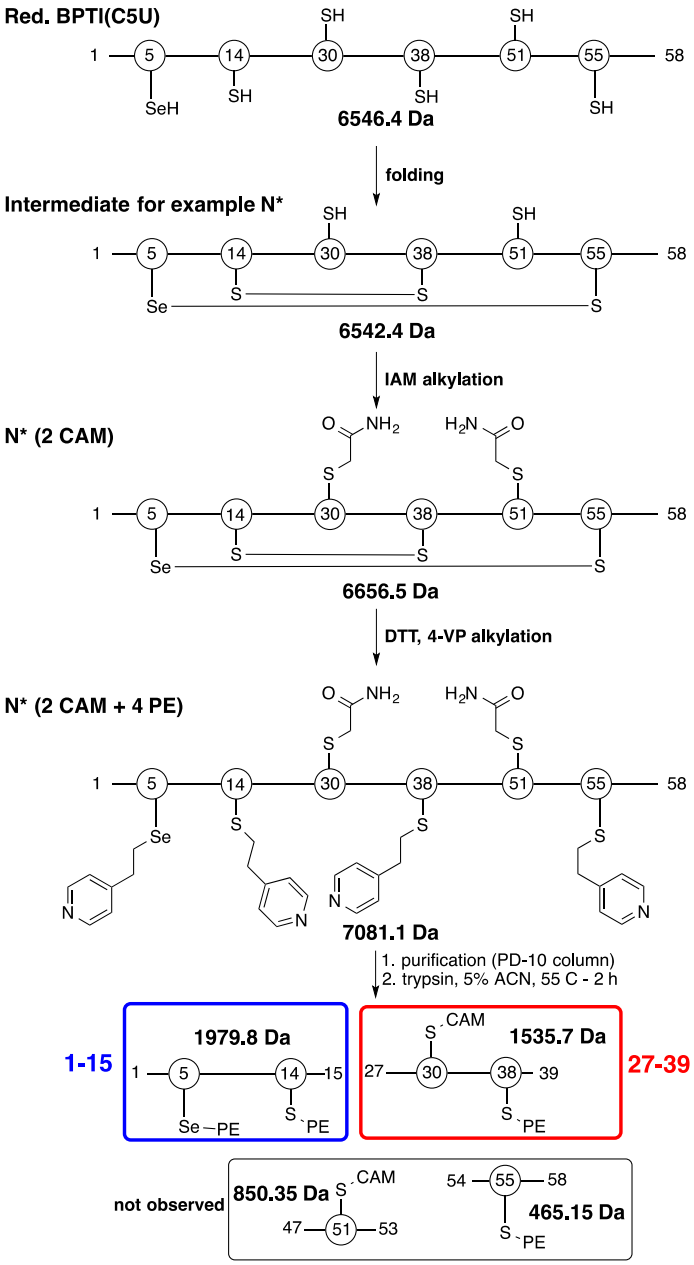


Figure S6. MALDI-MS of the C5U BPTI folding intermediate following alkylation and trypsin digestion. The peptide with m/z 1980.664 is C5U BPTI(1–15)[Sec5(PE);Cys14(PE)], whereas m/z 1536.563 corresponds to BPTI(27–39)[Cys30(CAM); Cys38(PE)]. The m/z 1793.692 peak originated from the m/z 1980.664 ion as a result of HSe-$(CH_2)_2$-Py elimination (-187 Da) caused by the MALDI laser beam. The m/z 1980.664 peak is enlarged in the inset to show the selenium isotope pattern.



Scheme S7. Expected trypsin fragmentation patterns for possible C5U BPTI folding intermediates containing two crosslinks. The crosslinked residues in each case are indicated in the brackets. Only bona fide intermediates that were previously detected in folding reactions with wild-type BPTI are shown.[8] The different candidates can be distinguished by MALDI-MS and MS/MS based on the unique masses of the individual peptides and their fragments.



Scheme S8. Analysis of the C5U BPTI folding mechanism, assuming formation of an N*-like intermediate. After folding for 2 min, the reaction was quenched with iodoacetamide (IAM). The mass of the modified intermediate (6656 Da) reflects addition of two carbamidomethyl (CAM) groups, indicating that it originally contained two free thiols, one Se-S bond, and one S-S bond (or one thiol, one selenol, and two S-S bonds). After purification, lyophilization, and reduction with DTT, subsequent alkylation with 4-vinylpyridine (4-VP) gave a new adduct (7081 Da) containing two CAM groups and four pyridylethyl (PE) groups. After purification and trypsin digestion, two peptides were obtained that correspond to BPTI(1–15) [Sec5(PE);Cys14(PE)] (1980 Da) and BPTI(27–39)[Cys30(CAM); Cys38(PE)] (1536 Da). MS/MS analysis of the 1536 Da fragment (Figure 9) proved that Cys30 is alkylated with CAM and Cys38 with PE. Peptides BPTI(47–53) and BPTI(54–58) were never detected under our conditions.

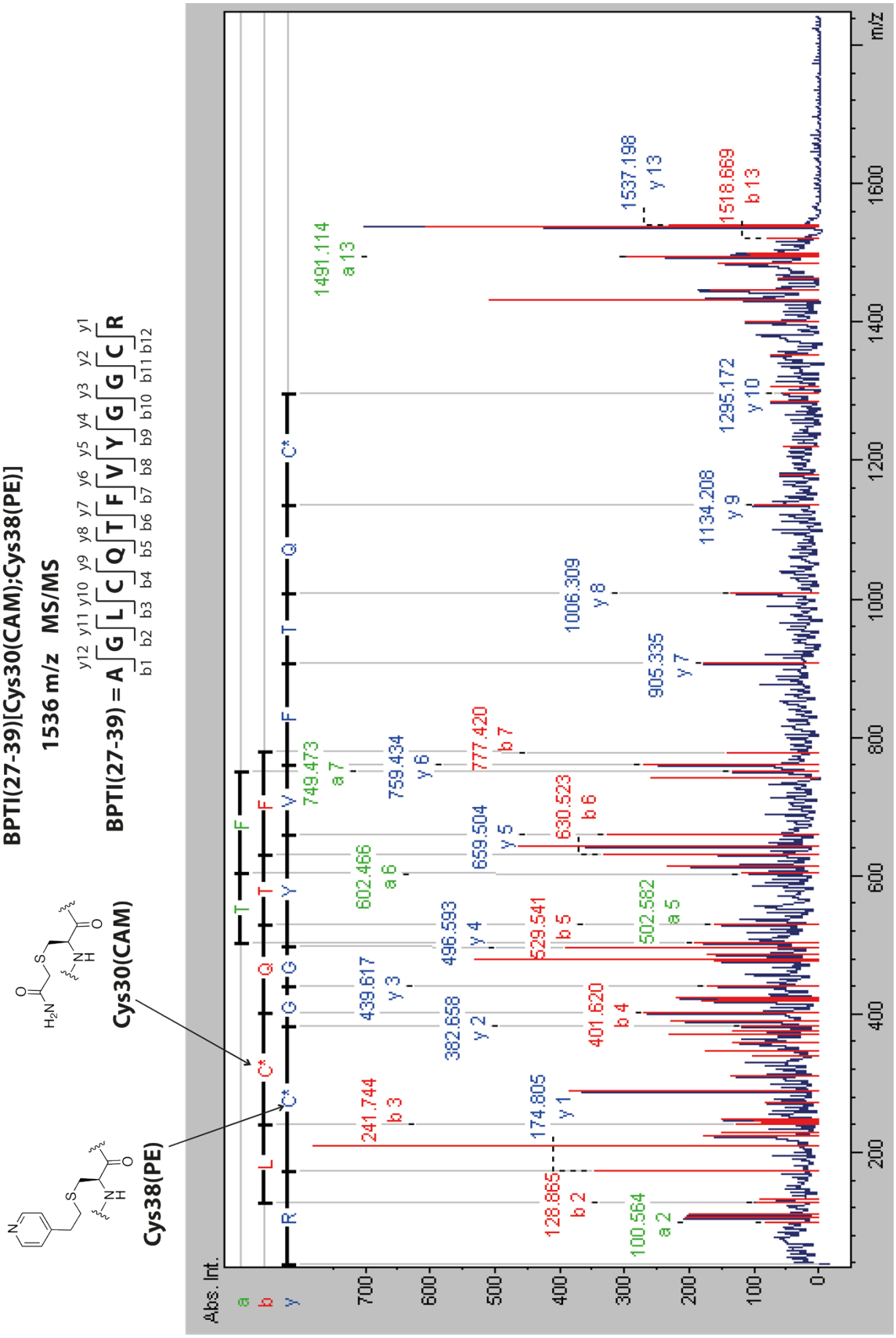


Figure 9. MS/MS analysis of the m/z 1536.563 peak from the MALDI-MS spectrum shown in Figure S6. The (*b4-b3*) and (*y10-y9*) ions show that the CAM group is attached to Cys30, whereas the (*y2-y1*) ion shows that the PE group is attached to Cys38.

References

1. N. Metanis and D. Hilvert, *Angewandte Chemie*, 2012, 51, 5585-5588.
2. M. Schnolzer, P. Alewood, A. Jones, D. Alewood and S. B. H. Kent, *International Journal of Peptide Research and Therapeutics*, 2007, 13, 31-44.
3. C. J. Armishaw, N. L. Daly, S. T. Nevin, D. J. Adams, D. J. Craik and P. F. Alewood, *J. Biol. Chem.*, 2006, 281, 14136-14143.
4. S. Flemer and R. J. Hondal, *Biopolymers*, 2011, 96, 439-439.
5. A. L. Schroll, R. J. Hondal and S. Flemer, *J. Pept. Sci.*, 2012, 18, 1-9.
6. J. S. Weissman and P. S. Kim, *Science*, 1991, 253, 1386-1393.
7. D. F. Dyckes, Creighto.T and R. C. Sheppard, *Nature*, 1974, 247, 202-204.
8. T. E. Creighton, *J. Mol. Biol.*, 1977, 113, 275-293.