Accepted Manuscript

Peptide macrocyclization through amide-to-amide transpeptidation

Xinya Hemu , Yibo Qiu , James P. Tam

PII: S0040-4020(14)00842-4

DOI: 10.1016/j.tet.2014.05.112

Reference: TET 25661

To appear in: *Tetrahedron*

Received Date: 21 April 2014

Revised Date: 28 May 2014

Accepted Date: 30 May 2014

Please cite this article as: Hemu X, Qiu Y, Tam JP, Peptide macrocyclization through amide-to-amide transpeptidation, *Tetrahedron* (2014), doi: 10.1016/j.tet.2014.05.112.

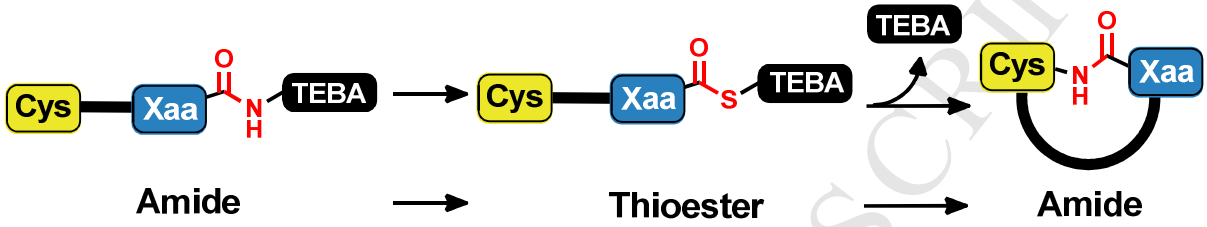
Graphical Abstract

To create your abstract, type over the instructions in the template box below.
Fonts or abstract dimensions should not be changed or altered.

Peptide macrocyclization through amide-to-amide transpeptidation

Leave this area blank for abstract info.

Xinya Hemu, Yibo Qiu, James P. Tam
School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, 03s-71, Singapore 637551



Tetrahedron
journal homepage: www.elsevier.com

Peptide macrocyclization through amide-to-amide transpeptidation

Xinya Hemu[a] , Yibo Qiu[a], and James P. Tam[a]*

[a] *School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, 03s-71, Singapore 63755*

ARTICLE INFO

Article history:
Received
Received in revised form
Accepted
Available online

Keywords:
Macrocyclization
Amide-to-amide synthesis
Transpeptidation
Amido thioester surrogate
Cys-thioester ligation

ABSTRACT

Cyclic cysteine peptides are peptide macrocycles endowed with enhanced metabolic stability and potentially, with membrane permeability. They have attracted attention in drug design and interest in their synthesis. The chemical approach for macrocyclization through transpeptidation bears striking similarity to the biological approach using an intein. Both use a similar design of thioester precursors and an amide-to-amide transpeptidation scheme, employing a series of acyl shifts to break and make amide bonds. Here we describe the synthesis of two cyclic cysteine peptides, hedyotide B1 and sunflower trypsin inhibitor-1, highlighting the similarities between the intein-based and chemical amide-to-amide schemes. In our intein-based and chemical schemes, we employed an intein *Mxe* or a thioethylbutylamido linkage at the C-terminus of their linear precursors, respectively. Our results demonstrated that the chemical approach provides a useful alternative to the intein approach with high efficiency. (136 words, 997 characters)

* Corresponding author. Tel.: +65-63162863; fax: +65-65151632; e-mail: jptam@ntu.edu.sg

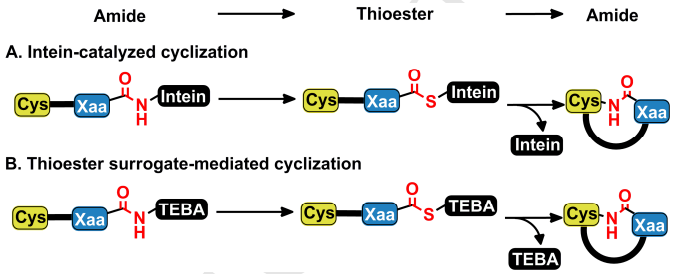
1. Introduction

Ribosome-derived, end-to-end macrocyclic peptides commonly occur in nature. They include pilins[1] and circular bacteriocins[2] from bacteria, *θ*-defensins[3] from animals, cyclotides[4,5] and sunflower trypsin inhibitors[6] from plants. The cyclized backbone protects these peptides from degradation by exopeptidases.[7] In addition, cyclic cysteine-rich peptides possessing one or more disulfide bonds further enhance their resistance to degradation by heat, chemical or enzymatic treatment.[8,9] Certain cyclic cysteine-rich peptides such as cyclotides and SFTI-1 are cell membrane permeable and orally active.[10,11] Consequently, naturally-occurring cyclic cysteine-containing peptides provide useful scaffolds as templates for peptidyl drug development.[12,13] Our laboratory has a strong interest in the discovery of cysteine-rich peptides, both linear and cyclic forms, from medicinal plants[14,15] in addition to methods for their synthesis.[16-22]

In the bioprocessing of a ribosome-derived peptide macrocycle, a cyclase is required for transpeptidation reactions, catalyzing the cleavage of a mature domain from its precursor and the formation of an end-to-end peptide bond by joining their terminal residues. However, only three cyclases, TraF, PATG and PCY1, have been isolated and characterized,[23-25] but not been exploited for chemical synthesis of peptide macrocyclization. Cyclases mediating the cleavage and cyclization reactions for *θ*-defensins and cyclotides remain to be discovered. Asparaginyl endopeptidase (AEP) has been implicated in the cleavage and cyclization of cyclotides, but its isolation and characterization remain elusive.[26] However, intein and Sortase A, two transpeptidases which are not naturally-occurring cyclases, have been exploited successfully for peptide macrocyclization.[27-30]

Intein, embedded as an endopeptidase in a protein, is a protein-splicing enzyme that mediates self-excision and ligation of the two flanking segments, N- and C-extein, to produce a newly joined N-C extein as a mature sequence.[31-33] Such splicing process catalyzed by intein involves four-step acyl shift reactions.[34,35] The first step that initiates the chain of acyl shifts is the N-S/O acyl shift, which breaks the amide bond between N-extein and N-terminal Cys/Ser of intein and forms an ester or a thioester in the catalytic pocket. This step is facilitated by a *cis*-conformation of the scissile amide bond. The second acyl shift reaction is an S/O-S/O transesterification reaction by the nucleophilic attack from N-terminal Cys/Ser of C-extein joining N- and C-exteins as a branched peptide via an ester or a thioester bond. The third step involves intein self-excision through an N-N acyl shift reaction that occurs at C-terminal Asn of intein via succinimide formation. The final step is not enzyme catalyzed and occurs spontaneously via an S/O-N acyl shift reaction that joins the N- and C-exteins as a new peptide bond, a mechanism identical to Native Chemical Ligation.[36,37] Muir's group has successfully exploited the intein biosynthesis for protein engineering in expressed protein ligation (EPL).[38] Since then intein-splicing techniques have been applied for ligation, cyclization and protein labeling. For preparing cyclic peptides using a recombinant system, a useful construct is a linear precursor design of Cys-peptide-intein with an N-terminal Cys, as demonstrated by the synthesis of kalata B1.[28]

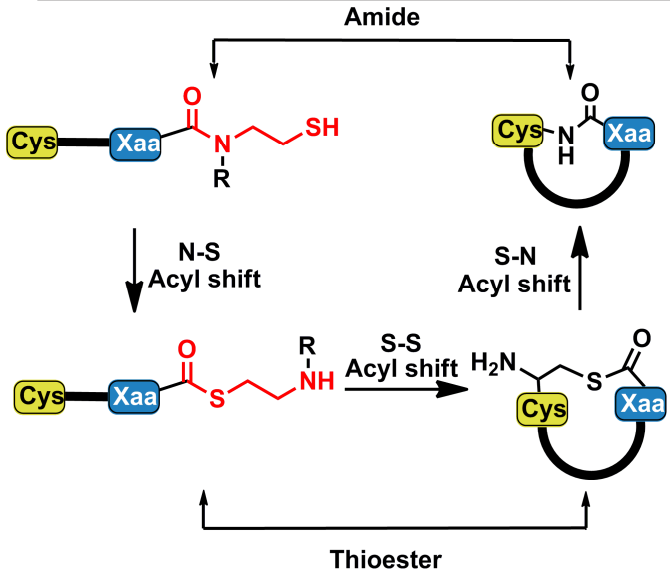
The chemical scheme of an amide-to-amide approach to prepare peptide macrocycles employs an amide bond as a latent thioester to form a new amide bond, mimicking an intein-based peptide cyclization scheme (**Scheme 1**). The use of amides as latent thioesters is best exemplified by N-alkyl cysteine, secondary cysteinyl amides that have been successfully exploited by Hojo and his coworkers in solid-phase peptide synthesis using Fmoc chemistry.[39] Under acidic conditions, N-alkyl cysteinyl amide readily undergoes an N-S acyl shift from an amide to a thioester at a much faster rate than the cysteinyl amide bond.[40] Various simplified versions of N-alkyl cysteine have also been developed[41,42] and an example is the C-terminal thioethylbutylamido (TEBA) linker recently introduced by our laboratory.[22] An advantage of the TEBA linker compared to a MeCys moiety is that it does not contain an α-carbonyl group. And eliminating the carbonyl group in TEBA also eliminates the β-elimination side reaction of the MeCys in peptide synthesis using Fmoc chemistry.[21]



Scheme 1. Schematic illustration of intein-catalyzed synthesis and chemical approaches using thioester surrogates for preparing macrocyclic peptides.

For its application in macrocyclization, the TEBA linker is placed at the C-terminus of a linear construct Cys-peptide-TEBA with Cys at the N-terminus (**Scheme 2**). Such a construct is similar in arrangement to the precursor construct of Cys-peptide-intein in which intein mediates the cyclization reaction. The TEBA linker also mediates macrocyclization of a Cys-peptide-TEBA construct. Under acidic conditions, the TEBA moiety isomerizes via an N-S acyl shift reaction from an amide to a thioester, which is trapped by an N-terminal Cys-thiol via a thiol-thioester exchange, an S-S acyl shift reaction, to form an end-to-end thiolactone. Interestingly, such thiolactone is relatively stable under acidic conditions as shown by our recent synthesis of cyclic conotoxin[21] and kalata B1.[22] Under basic conditions, the end-to-end thiolactone undergoes a proximity-driven S-N acyl shift reaction, resulting in the formation of a new amide bond to yield a macrocycle. Advantages of the TEBA linkage as a thioester surrogate include its ease of synthesis and compatibility with Fmoc chemistry.

Here we describe two amide-to-amide schemes in the synthesis of cyclic cysteine peptides, a cyclotide hedyotide B1 (hB1) and a sunflower seed trypsin inhibitor SFTI-1. In the biological scheme using an intein, we investigated the intein splicing site specificity to optimize its biosynthesis of the cyclotide hB1. In the chemical scheme, we applied the thioethylbutylamido (TEBA) as a C-terminal linker for preparing SFTI-1 and showed that the amide-to-amide cyclization can be performed in a one-pot reaction using the crude peptide after its cleavage from the resin support.



Scheme 2. A scheme demonstrating the biomimetic peptide cyclization mediated by thiolethylamido (TEA) group. Three acyl shift reactions (N-S, S-S, S-N) mediate the breaking and making of the peptide bond. Through a series of acyl shifts, the thioester surrogate was excluded from the target peptide with the N and C termini joined by an amide bond.

2. Results and Discussion

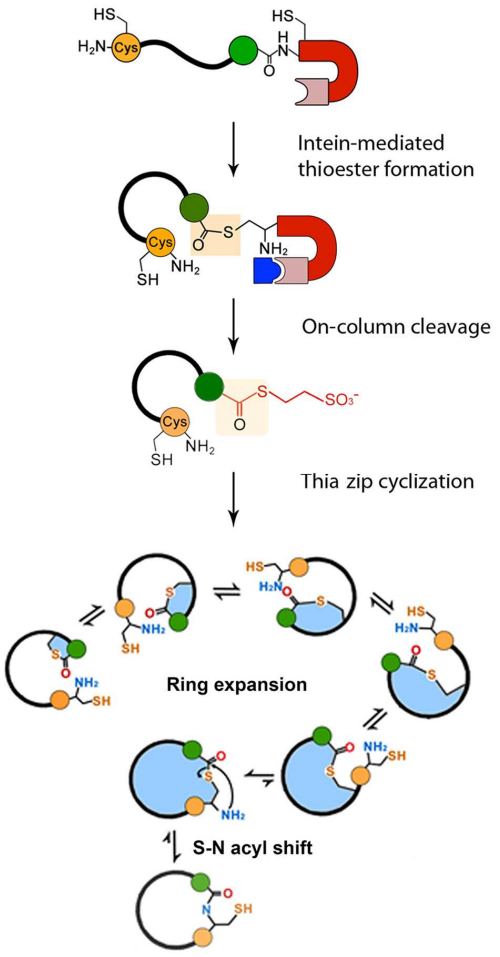
2.1. Intein-mediated synthesis of hedyotide B1

Hedyotide B1 (hB1), the first bracelet cyclotide identified from herbal plant *Hedyotis biflora*, is a cyclic cysteine-knot peptide of 30 residues.[43] For an intein-based synthesis of hB1, each of the six Cys residues can serve as a ligation site for macrocyclization using a Cys-peptide-intein linear precursor. To determine the optimal ligation site, we prepared all six linear precursors of hB1, named as B1X (X = R, T, P, G, Y, and F), based on the six Xaa-Cys dipeptide sequences in hB1 (Xaa = Arg, Thr, Pro, Gly, Tyr and Phe), each with an N-terminal Cys residues and a different Xaa-residue at the C-terminus (**Table 1**). These precursors were co-expressed with intein *Mxe* that linked to a chitin-binding domain (CBD) in an *E. coli* system (**Scheme 3**). The fusion protein B1X-*Mxe*-CBD was isolated in a chitin column, where the enzyme-catalyzed thioesterification and thiolysis took place. The reducing thiol sodium 2-mercaptoethanesulfonate (MESNa) was added to facilitate B1X release by thiolysis, a thiol-thioester reaction. It also protected multiple Cys residues from disulfide formation. Previously we have shown that cysteine-rich peptides, both the N-terminal and internal Cys side-chain thiol groups can undergo an intramolecular thiol-thioester reaction, an S-S acyl shift reaction with the C-terminal thioester, leading to the formation of multiple peptide thiolactones.[21] These thiolactones facilitate the cyclization reaction through a thia zip mechanism of thiolactone ring expansion via a series of entropy-driven thiol-thiolactone exchange reactions.[19] The thiolactone ring expansion leads to an end-to-end thiolactone with the N^{α}-amine and C^{α}-carboxyl brought into close proximity to allow an S-N acyl shift reaction to occur resulting in a peptide macrocycle.

Table 1. Six hedyotide B1 precursors B1X (X for R, T, P, G, Y, F) with different ligation site residues

Precursor	Amino acid sequences
hB1	GTR C GET C FVLP C WSAKFG C Y C QKGF C YRN (I II III IV V VI)
B1R	C GET C FVLP C WSAKFG C Y C QKGF C YRNGTR
B1T	C FVLP C WSAKFG C Y C QKGF C YRNGTR C GET
B1P	C WSAKFG C Y C QKGF C YRNGTR C GET C FVLP
B1G	C Y C QKGF C YRNGTR C GET C FVLP C WSAKFG
B1Y	C QKGF C YRNGTR C GET C FVLP C WSAKFG C Y
B1F	C YRNGTR C GET C FVLP C WSAKFG C Y C QKGF

Cysteine residues (in yellow) are aligned with the native sequence.
Different C-terminal residues are highlighted in red.



Scheme 3. Schematic illustration of Intein-mediated thioester formation and enzyme-free thia zip cyclization. Cysteine-rich peptides (N-terminal cysteine in orange and C-terminal residues in green) were co-expressed with intein (red) and chitin-binding domain (brown) in *E. coli*. The fusion proteins were extracted and purified on the affinity column containing chitin beads (blue). Intein-mediated N-S acyl shift reaction resulted in formation of peptide-thioesters that were cleaved from the fusion proteins through S-S acyl shift reactions either by MESNa-mediated thiolysis that produced MES-peptide thioester or by internal thiol-mediated thiolactone formation. The on-column cyclization reaction proceeded through thia zip mechanism and resulted in end-to-end cyclic cysteine-rich peptides.

For the release of B1X from the intein, we found that both inter- and intramolecular thiolysis occurred under neutral to basic pH. For our experiments, the release of B1X was performed by an intermolecular thiolysis using a large excess of MESNa and the amount released at 12 h and 24 h intervals was analyzed by SDS-PAGE based on the intact fusion protein and intein-CBD which remained on the chitin-beads (**Fig. 1A**). Our results showed that the cleavage of B1G (X = Gly) from fusion protein was complete within 12 h whereas cleavage of B1R, B1Y or B1F (X = Arg, Tyr and Phe, respectively) was incomplete in 12 h, with yields about 50%, but increased to >80% after 24 h (**Fig. 1B**). The thiolytic release from the fusion protein of both B1T (X = Thr) and B1P (X = Pro) was slow, with yields about 60% for B1T and 50% for B1P after 24 h.

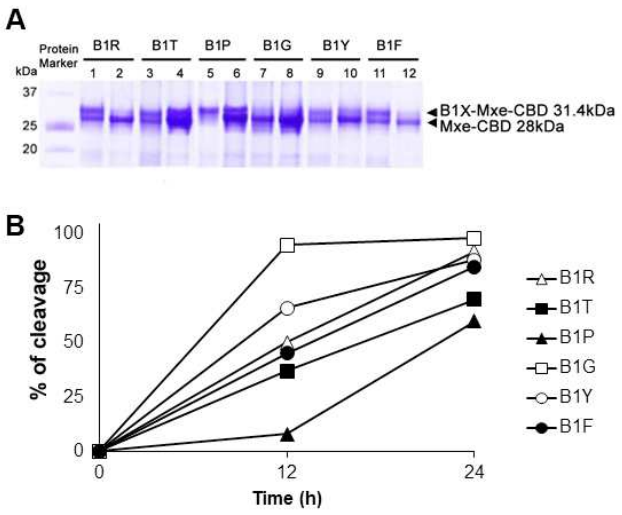


Fig. 1. Comparison of cleavage efficiency of six B1X precursors. (A) SDS-PAGE analysis of proteins remained on chitin column. Protein samples were prepared by dissolving 30 µl chitin beads in 15 µl 2x loading buffer and heating at 100 °C for 1 min. The odd-number lane of each clone shows cleavage ratio after 12 h incubation in column buffer containing MESNa and the even-number lane shows the cleavage ration after 24 h incubation. (B) Comparison of cleavage efficiency of six B1X-*Mxe*-CBD fusion proteins.

Next, we identified the products released from the thiolytic reactions by HPLC and MS **(Fig. 2A).** We anticipated that the B1X-thioesters (MES thioesters) and/or B1X-thiolactones would undergo the thia zip mediated macrocyclization to form the end-to-end lactam, the cyclic but S-reduced form of hB1. Indeed, HPLC analysis showed that the cyclic hB1 was obtained as the major product for B1G, B1Y, B1F, B1R (>90%) and B1T (> 60%). In contrast, B1P afforded <5% yield of cyclic B1P, with B1P-OH, the hydrolyzed product (60% yield), as the major side product together with about 30% yield of the unreacted thioester B1P-MES (**Fig. 2B**). By calculating the peptide release from the intein and cyclization yield, we obtained the ranking of cyclization efficiency as Gly>Arg, Tyr, Phe>Thr>>>Pro with a difference of 20-fold between B1G and B1P. Our results suggest that in designing a Cys-peptide-intein construct, placing a Gly at the C-terminus as the C-terminal ligation site is highly desirable while a hindered amino acid such as Thr or Pro undesirable.

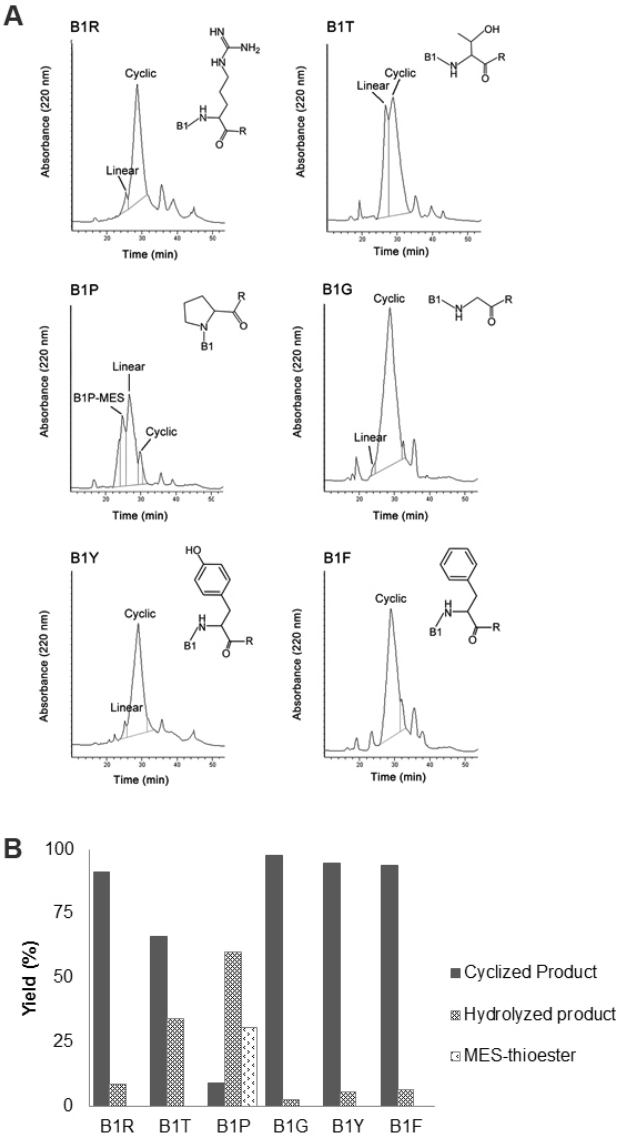


Fig. 2. Comparison of cyclization efficiency of six B1X precursors. (A) HPLC profiles of product mixtures eluted from chitin column. (B) Cyclization products analysis and yield.

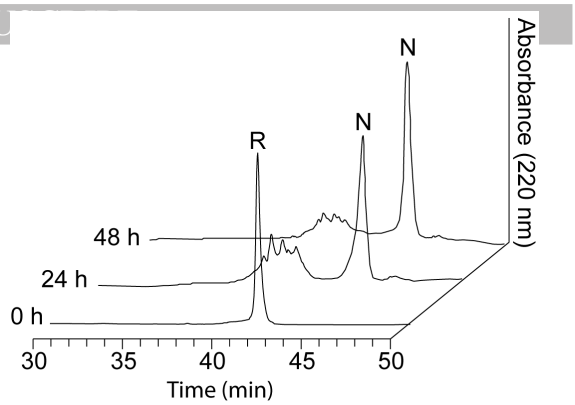


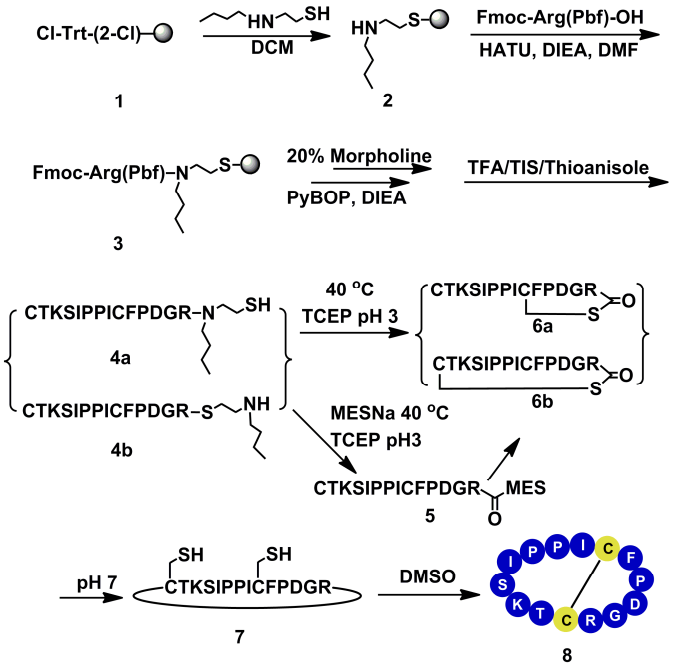
Fig. 3. Global oxidative folding of hedyotide B1. Reduced hedyotide B1 (peak R) was eluted at 42 min and native hedyotide B1 (peak N) at 43.8 min.

The fully reduced cyclic hB1 was subjected to global oxidative folding with 10 μM cyclized hB1 in the presence of reduced/oxidized glutathione (molar ratio peptide:GSH:GSSG = 1:10:100) and 70% isopropanol at pH 8.5, an optimized condition which we have previously reported.[12] The reaction was completed after 48 h at room temperature to give 48% yield of the native-folded hB1 (**Fig. 3**, peak N). Cyclotides contain an inner sulfur core consisting of a cysteine knot, which displaces the bulky hydrophobic side chains to its surface. Consequently, a cyclotide such as hB1 contains hydrophobic patches and tends to aggregate in the folding process. Addition of co-solvent isopropanol improved the folding yield in two aspects, increasing solubility of peptide by reducing non-covalent interactions, and facilitated thiol-disulfide exchange reactions by promoting the native-like conformation of hB1. It should be noted that hB1 belongs to the bracelet family of cyclotides. Comparing with Möbius family of cyclotides, oxidative folding of bracelet cyclotides has been known to be difficult to obtain high yields.[17,44]

2.2. TEBA-mediated synthesis of SFTI-1

Sunflower trypsin inhibitor-1 (SFTI-1) is a 14-residue cyclic peptide with one disulfide linking Cys3 and Cys11 to stabilize two β-sheets. It is one of the most potent trypsin inhibitor (K_i=0.1 nM) belonging to the Bowman-Birk family.[6] It is relatively stable against enzyme degradation and could be a useful template for peptide engineering for therapeutics.

In our synthesis of SFTI-1 utilizing a precursor design of Cys-peptide-TEBA, we started with coupling 2-(butylamino)ethanethiol (TEBA amine form) to the Cl-Trt(2-Cl) resin to afford the TEBA resin (**Scheme 4**). 2-(butylamino)ethanethiol is commercially available and readily reacted with Cl-Trt(2-Cl) resin **1** to give TEBA-Trt(2-Cl) resin **2** with a thioether linkage in 1 h. An advantage of this approach over an N-alkyl Cys as a linker is that the resin support serves as a protecting group (Trt) for the thiol moiety of TEBA, which is deprotected during the TFA cleavage step concurrent with other protecting groups of the peptide. In the Cys-peptide-TEBA scheme, there are two logical ligation sites, Cys-Arg and Cys-Ile. Based on the lessons learned from the cyclotide hB1 synthesis using intein, Arg was chosen over the β-branched amino acid Ile as the ligation site to minimize steric hindrance in the cyclization step.



Scheme 4. A scheme of the TEBA-mediated synthesis of cyclic peptide SFTI-1.

The synthesis of Cys-SFTI-TEBA commenced with the coupling of Fmoc-Arg(Pbf)-OH as the first amino acid to the TEBA resin using HATU as the coupling reagent. The reaction completed in 1 h as determined by the chloranil test, which detects the presence of free secondary amine. The Fmoc-Arg(Pbf)-TEBA resin **3** was then used as the starting material for the synthesis of SFTI-1 in a microwave peptide synthesizer by a Fmoc-chemistry protocol with a modified deprotection step using 20% morpholine instead of 20% piperidine (see experimental section 4.8). SFTI-1 contains a base-sensitive Asp-Gly sequence, which can lead to aspartimide formation during synthesis. To suppress the aspartimide formation, we employed a milder base in the deprotection steps, using morpholine instead of piperidine. In this case, less than 1% aspartimide formation was observed. After the completion of the SFTI-1 synthesis and its release from the resin support by TFA, HPLC analysis showed that two products, the amide form (N form) SFTI-TEBA **4a** and thioester form (S form) **4b**, were obtained.

The cleaved peptides were subjected to cyclization in two approaches (**Scheme 4**). The first approach was a two-step process, a process similar to intein approach with the thiolytic release of peptide as a thioester by a large excess of an external thiol, followed by cyclization. Indeed, the Cys-SFTI-TEBA was released from the TEBA linker as Cys-SFTI-MES thioester using a large excess of MESNa. The thiolytic release was achieved at 40 °C, pH 3.[22] 10 mM TCEP was added to prevent disulfide bond formation (**Fig. 4**). The thioester SFTI-MES **5** was observed as the major product after 15 h. Subsequently, two thiolactones (**6a** and **6b**) were also generated through thiol-thioester exchange between the thiol side chain of Cys residues and the C-terminal thioester bond, a 19-atom thiolactone **6a** generated from internal Cys and a 43-atom thiolactone **6b** generated from the N-terminal Cys. The thiolactones and the end-to-end lactam SFTI-1 **7** (the reduced cyclic form) are isomers with the same MW 1515 Da, but thiolactones can be distinguished from lactam by the stronger UV-absorbance at 260 nm, a characteristic absorbance of a thioester bond. In HPLC trace, product **6b** eluted close to **7** whereas **6a** eluted earlier than **6b**. The product mixtures containing all three products **5, 6a** and **6b** were subjected to the thia zip cyclization at pH 7 at room temperature to afford **7** after 4 h. After HPLC purification and oxidative folding with 20% DMSO, native SFTI-1 **8** was obtained in 33% yield.

With two cysteine residues in the SFTI-1 sequence, we anticipated that they could undergo a direct thiolactone formation to release the TEBA linker, rendering the use of an external thiol for the extra step of thioesterification unnecessary. Thus, the second approach in the macrocyclization of Cys-SFTI-TEBA was performed without MESNa or any external thiol under the experimental conditions similar as the previous approach. After 15 h, except for the thioester **5**, all expected products, SFTI-TEBA thioester **4b**, thiolactones **6a** and **6b,** were found. A small amount of hydrolysis product (<5%) eluted earlier than thiolactones **6a** in the HPLC trace was also observed. After cyclization at pH 7, the reaction mixture containing product **7** was directly subjected to oxidative folding with 20% DMSO without purification. Both approach with or without the use of an external thiol such as MESNa gave comparable cyclization yield, which was 89% and 84%, respectively. However, under our experimental conditions, the approach without an external thiol required only one purification step from the linear precursor to native product whereas the other approach required two purification steps. Thus the overall yield was 10% higher (43% vs 33% isolated yield) when no external thiol was included in the reaction.

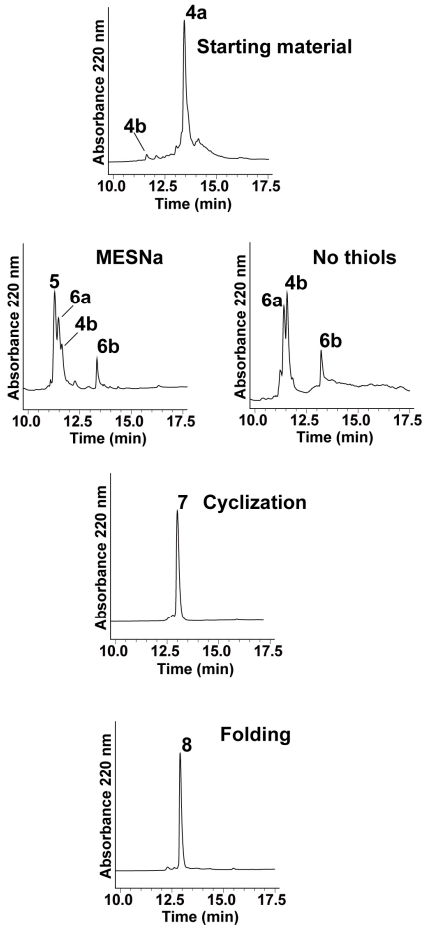


Fig. 4. RP-HPLC profiles of synthetic intermediates in the synthesis of SFTI-1. Starting material was subjected to two conditions with or without MESNa at 40 °C pH 3 for 18 h. After the pH adjusted to 7, the cyclization was conducted at room temperature for 4 h. Then 20% DMSO was used to perform the oxidative folding reaction. Each intermediates were monitored by RP-HPLC.

3. Conclusions

Our chemical scheme of an amide-to-amide approach in peptide macrocyclization has two major features in common with the biological approach using an intein. First, the design of the linear precursor construct of Cys-peptide-intein/TEBA is similar, although the size of TEBA is several hundred fold smaller than an intein. Second, they share similar chemical principle in the amide-to-amide splicing process via a series of acyl shifts, resulting in the excision of the C-terminal amide moiety as an intein or the TEBA linker, followed by forming a new amide cysteinyl bond to generate a peptide macrocycle.

There are also significant differences between these two methods. Whereas an intein catalyzes the amide to thioester at the C-terminal junction at neutral pH, the TEBA linker requires fairly acidic conditions for transforming an amide to a thioester via an N-S acyl shift. The N-terminal cysteinyl thiol, and in cysteine-rich peptides, an internally placed cysteinyl thiol, play an important role to release or excise either an intein or TEBA linker to complete the excision process through one or more thiol-thioester exchange reactions. Thus, the amide-to-amide chemical approach mediated by the TEBA is particular suitable for preparing cyclic cysteine peptides since the presence of multiple internal thiols greatly facilitates the release the TEBA linker as thiolactones which can then directly undergo thia zip cyclization. Again, there is a large pH difference between the chemical means to release the TEBA and biological means to release the intein, with the former at an acidic pH and the later in neutral pH. Thus, the TEBA-mediated scheme depends on the catalysis by general acids and bases and differs from catalysis by specific acid and base of intein to break and make peptide bonds in peptide macrocyclization.

4. Experimental section

4.1. General

All Fmoc-amino acids reagents and solvents were used without purification. Amino acids were purchased from Chem-Impex. HOBt, PyBOP and HATU were purchased from GL Biochem. Chlorotrityl chloride resin (1.2 mmol/g) and solvents DMF, DCM, piperidine, morpholine, and diethylether were purchased from Merck. All HPLC was performed on Shimadzu liquid chromatography using an analytical column (4.6 mm x 250 mm) with a linear gradient of 10 to 60% acetonitrile (ACN)/0.1% trifluoroacetic acid (TFA) for 25 min at a flow rate of 1 ml/min, and a preparative column C18 column (22 mm x 250 mm) at a flow rate of 6 ml/min. Molecular weights were measured with 4800 plus MALDI-TOF/TOFTM Analyzer.

4.2. Cloning of Hedyotide B1 precursors

The sequence of B1 cDNA (cB1) has been confirmed by 3'-RACE and 5'-RACE reactions from previous study. The cDNA sequence was optimized according to *E. coli* codon bias index (**Table S1**) by modifications on G1, R3, P12 and R29. The six hB1 precursors B1X (X=R, T, P, G, Y, and F) were synthesized by tandem-extension PCR without adding templates in the PCR mix (**Fig. S1**). For each precursor, four primers were designed to fully cover the complete nucleotide sequence with overlapping oligos. Primers 1f and 4r carries the NdeI and SapI restriction sites respectively for integrating into vector pTXB1 (NEB, IMPACT kit) (**Fig. S2**). PCR products were purified after 25 cycles of amplification by agarose-gel (2% *w/v*) DNA extraction (Invitrogen, PureLink™ Quick Gel Extraction Kit). The PCR products obtained from each tandem-extension reaction were integrated into pTXB1 by NdeI/SapI double digestion and T4 DNA ligation. Ligation products were transformed into competent DH5α *E. coli*. Correct clones were screened by EcoRI/EcoRV digestion (1 h, 37 °C) and agarose-electrophoresis (**Fig. S3**). The purified plasmids were transformed into high-performance expression strain ER2566 (BL21).

4.3. IPTG-induced expression and fusion protein extraction

One fresh-growing colony was innoculated into 3 ml LB medium in a 15 ml culture tube for 3 h at 37 °C. One ml culture was then transferred to 500 ml LB medium and shaking-incubated in a 2 L flask. Until A_{600} reached 0.7, decreased incubator temperature to 16 °C and added 0.3 mM IPTG, shaked for 24 h at 150 rpm.

Bacterial cells were collected by 10 min centrifugation at 10,000 rcf. The pellets were resuspended in cell lysis buffer (20 mM Tris-Cl, pH 8.5, 500 mM NaCl, 1 mM EDTA, 0.1% Triton X-100) and lysed by ultrasonication for 1 h on ice (4s-operation/10s-pause). The clarified cell extract was purified by 40 min centrifugation of the cell lysate at 10,000 rcf at 4 °C and filtration (0.22 μm). Expressed proteins in the cell lysates and clarified extracts were examined by SDS-PAGE (**Fig. S4**).

4.4. On-column purification and cleavage

Chitin beads (10 ml) were loaded into chitin-binding column and then calibrated by 100 ml column buffer (20 mM Tris-Cl, 500 mM NaCl, 1 mM EDTA, pH 8.5). Pre-chilled 50ml clarified cell extract from 1 L cell culture was slowly loaded on to the chitin column (flow rate <0.5 ml/min). Chitin beads were washed with at least 200 ml column buffer (flow rate = 2 ml/min) to stabilize the affinity binding between chitin and CBD. After wash, beads were flushed with 50 ml column buffer to remove unspecific binding. This purification step was conducted under 4 °C in the cold room. Chitin beads were then flushed with 20 ml cleavage buffer containing mercaptoethanesulfonate (20 mM Tris-Cl, 500 mM NaCl, 1 mM MESNa, pH 8.5) and incubated at room temperature for 24 h, or at 4 °C for not more than 48 h. The cleavage efficiency was checked by analyzing the column-bound proteins using SDS-PAGE. Splicing products were eluted with 10 ml cleavage buffer. The eluents were desalted using SPE ODS-C18 Cartridges (1 ml) and eluted with 1 ml 60% ACN, 0.05% TFA, pH 2.0. TCEP (30 mM) was added to the eluents and incubated for 1 h at 37 °C to reduce any unexpected disulfide bonds.

4.5. Tricine SDS-PAGE

Samples for SDS-PAGE included cell pellet from 1 ml cell culture before induction, cell pellet from 0.5 ml cell culture after induction, cell debris from 50 μl crude extract, clarified extract from 50 μl crude extract, 20 μl chitin beads collected after step sample loading, and 20 μl chitin beads after cleavage. Six samples were resuspended/mixed with 40 μl 2x sample loading buffer (Biorad, tricine-SDS sample buffer) and heated at 100 °C for 10 min. Tricine-SDS separation gel contained 10% acrylamide/bis-acrylamide mix (29:1 stock solution), 1M Tris-Cl/0.1% SDS (*w/v*, 3x stock solution, pH 8.5), 20% glycerol, 0.05% ammonium persulfate (APS) (*w/v*, 10% stock solution), and 0.005% TEMED (*w/v*). Tricine-SDS stacking gel contained 4.05% acrylamide/bis-acrylamide mix (29:1 stock solution), 7.5% Tris-Cl/0.1% SDS (*w/v*, 3x stock slution, pH 8.5), 0.05% ammonium persulfate (APS, *w/v*, 10% stock solution), and 0.01% TEMED. The 5x Tricine-SDS PAGE running buffer stock was prepared by 0.5M tricine, 0.5M Tris-base, 0.5% SDS (*w/v*).

4.6. RP-HPLC analysis

Analytical RP-HPLC was performed using C18 column (5 μm, 300 Å, 4.6x250 mm) with a flow rate of 1 ml/min. The mobile phase included buffer A (0.05% TFA, milli-Q water) and buffer B (80% ACN, 0.035% TFA, milli-Q water). The gradient for cyclic products was 10-70% ACN for 60 min. The gradient for folding products was 0-50% ACN for 50 min.

4.7. Disulfide folding and characterization

Lyophilized B1 precursors were oxidized by GSH and GSSG in the molar ratio of peptide:GSSG:GSH = 1:10:100. The powder mixture was dissolved in 70% iPrOH with 0.1 mM Tris-Cl (pH 8.5) and mixed by vortex. The reaction solution was degassed by nitrogen to eliminate oxygen interference to the thiol-thioester exhange reaction. Folding reactions were conducted for 48 h at room temperature.

4.8. Synthesis of SFTI-1 on TEBA-loaded resin

Cl-Trt(2-Cl) resin (417 mg, 0.5 mmol) was swollen for 30 min in CH_2Cl_2 (15 ml) and washed with CH_2Cl_2 (10 ml x 2). The resin was re-suspended in 10 ml CH_2Cl_2. To the suspension, 2-(butylamino)ethanethiol (37.6 μl, 0.25 mmol) was added, and the suspension was shaken for 1 h at room temperature. Then DIEA (133 μl, 0.75 mmol) in MeOH (10 ml) was added to the suspension and the suspension was shaken for 10 min. After filtration, the resin was washed with CH_2Cl_2 (10 ml x 3) and DMF (10 ml x 3). To the resin, a pre-mixed solution containing 4 equiv Fmoc-Arg(Pbf)-OH (M.W. 648.8, 1 mmol 649 mg), 4 equiv HATU (M. W. 380.2, 1 mmol, 380 mg) and 6 equiv DIEA (265 μl, 0.75 mmol) was added. The coupling reaction was performed at room temperature for 1 h and monitored by chloranil test (2% acetaldehyde, 2% chloranil in DMF, 5 min at room temperature). Peptide elongation was carried out in an automatic microwave peptide synthesizer using PyBOP/DIEA. Fmoc deprotection was performed by the use of 20% morpholine in DMF. 940 mg SFTI-TEBA-loaded resin was produced and cleaved with 10 ml cleavage cocktail containing TFA/TIS/H_2O (95/2.5/2.5, *v/v*) to remove the solid support and side chain protection groups. The cleavage reaction was allowed to proceed for 2 h and cold diethyl ether was used to precipitate the peptide. Then it was dried *in vacuo* and analyzed by RP-HPLC. 330 mg precipitation was generated and used directly for the following cyclization reaction.

4.9. Cyclization and folding of SFTI-1

4.9 mg crude peptide precipitation containing SFTI-TEBA **4a** and **4b** was dissolved in 0.1M sodium phosphate buffer (pH 3, 1 ml) at a concentration of 3 mM. The reaction mixture was incubated at 40 °C with or without MESNa (150 μmol, 24.7 mg) for 18 h. TCEP (10 μmol, 2.9 mg) was added to prevent the disulfide formation through inter- or intra-molecular linkage. The formation of synthetic intermediates such as thioesters and thiolactones was monitored by RP-HPLC at both 220 and 260 nm. Then the reaction mixture was diluted with 0.1M sodium phosphate (pH 3, 2 ml) and adjusted with 1 N NaOH. After the pH was adjusted to 7, the thia zip cyclization was allowed to proceed for 4 h at room temperature. The cyclization process was monitored by RP-HPLC. After the cyclization completed, the reaction mixture was purified by RP-HPLC when MESNa was included as an external thiol. Reduced cyclic SFTI-1 was then dissolved in pH 7.5 phosphate buffer at concentration of 1 mM and 20% DMSO (*v/v*) were added. The oxidative folding reaction was allowed to proceed at room temperature and monitored by RP-HPLC. For conditions without MESNa, 20% DMSO was added directly to the cyclization mixture to perform the oxidative folding reaction. For both conditions, the oxidative folding reaction was quenched with 1 N HCl to pH 2 and purified by using RP-HPLC to give native SFTI-1.

Acknowledgments

This research was supported in part by Biomedical Research Council (BMRC 09/1/22/19/612) of A*STAR and National Research Foundation (CRP8-2011-05) of Singapore.

References and notes

1. Eisenbrandt, R.; Kalkum, M.; Lai, E.-M.; Lurz, R.; Kado, C. I.; Lanka, E. *J. Biol. Chem.* 1999, *274*, 22548-22555.
2. González, C.; Langdon, G. M.; Bruix, M.; Gálvez, A.; Valdivia, E.; Maqueda, M.; Rico, M. *Proc. Natl. Acad. Sci. U. S. A.* 2000, *97*, 11221-11226.
3. Tran, D.; Tran, P. A.; Tang, Y. Q.; Yuan, J.; Cole, T.; Selsted, M. E. *J. Biol. Chem.* 2002, *277*, 3079-3084.
4. Trabi, M.; Craik, D. J. *Trends Biochem. Sci.* 2002, *27*, 132-138.
5. Pelegrini, P.; Quirino, B.; Franco, O. *Peptides* 2007, *28*, 1475-1481.
6. Luckett, S.; Santiago Garcia, R.; Barker, J. J.; Konarev, A. V.; Shewry, P. R.; Clarke, A. R.; Brady, R. L. *J. Mol. Biol.* 1999, *290*, 525-533.
7. Hruby, V. J. *Life Sci.* 1982, *31*, 189-199.
8. Colgrave, M. L.; Craik, D. J. *Biochemistry* 2004, *43*, 5965-5975.
9. Thorstholm, L.; Craik, D. J. *Drug Discov. Today Tech.* 2012, *9*, e13-e21.
10. Gran, L. *Acta Pharmacol. Toxicol.* 1973, *33*, 400-408.
11. Cascales, L.; Henriques, S. T.; Kerr, M. C.; Huang, Y. H.; Sweet, M. J.; Daly, N. L.; Craik, D. J. *J. Biol. Chem.* 2011, *286*, 36932-36943.
12. Wong, C. T. T.; Rowlands, D. K.; Wong, C.-H.; Lo, T. W. C.; Nguyen, G. K. T.; Li, H.-Y.; Tam, J. P. *Angew. Chem. Int. Ed.* 2012, *51*, 5620-5624.
13. Clark, R. J.; Jensen, J.; Nevin, S. T.; Callaghan, B. P.; Adams, D. J.; Craik, D. J. *Angew. Chem. Int. Ed.* 2010, *49*, 6545-6548.
14. Nguyen, G. K. T.; Lim, W. H.; Nguyen, P. Q. T.; Tam, J. P. *J. Biol. Chem.* 2012, *287*, 17598-17607.
15. Nguyen, G. K. T.; Lian, Y.; Pang, E. W.; Nguyen, P. Q. T.; Tran, T. D.; Tam, J. P. *J. Biol. Chem.* 2013, *288*, 3370-3380.
16. Tam, J. P.; Lu, Y. A. *Tetrahedron Lett.* 1997, *38*, 5599-5602.
17. Tam, J. P.; Lu, Y. A. *Protein Sci.* 1998, *7*, 1583-1592.
18. Tam, J. P.; Lu, Y. A.; Yang, J. L.; Chiu, K. W. *Proc. Natl. Acad. Sci. U. S. A.* 1999, *96*, 8913-8918.
19. Tam, J. P.; Lu, Y. A.; Yu, Q. *J. Am. Chem. Soc.* 1999, *121*, 4316-4324.
20. Tam, J. P.; Wong, C. T. T. *J. Biol. Chem.* 2012, *287*, 27020-27025.
21. Hemu, X.; Taichi, M.; Qiu, Y.; Liu, D. X.; Tam, J. P. *Pept. Sci.* 2013, *100*, 492-501.
22. Taichi, M.; Hemu, X.; Qiu, Y.; Tam, J. P. *Org. Lett.* 2013, *15*, 2620-2623.
23. Haase, J.; Lanka, E. *J. Bacteriol.* 1997, *179*, 5728-5735.
24. Lee, J.; McIntosh, J.; Hathaway, B. J.; Schmidt, E. W. *J. Am. Chem. Soc.* 2009, *131*, 2122-2124.
25. Barber, C. J. S.; Pujara, P. T.; Reed, D. W.; Chiwocha, S.; Zhang, H.; Covello, P. S. *J. Biol. Chem.* 2013, *288*, 12500-12510.
26. Saska, I.; Gillon, A. D.; Hatsugai, N.; Dietzgen, R. G.; Hara-Nishimura, I.; Anderson, M. A.; Craik, D. J. *J. Biol. Chem.* 2007, *282*, 29721-29728.
27. Camarero, J. A.; Muir, T. W. *J. Am. Chem. Soc.* 1999, *121*, 5597-5598.
28. Kimura, R. H.; Tran, A.-T.; Camarero, J. A. *Angew. Chem. Int. Ed.* 2006, *45*, 973-976.
29. Wu, Z.; Guo, X.; Guo, Z. *Chem. Commun.* 2011, *47*, 9218-9220.
30. Jia, X.; Kwon, S.; Wang, C. I.; Huang, Y. H.; Chan, L. Y.; Tan, C. C.; Rosengren, K. J.; Mulvenna, J. P.; Schroeder, C. I.; Craik, D. J. *J. Biol. Chem.* 2014, *289*, 6627-6638.
31. Hirata, R.; Ohsumk, Y.; Nakano, A.; Kawasaki, H.; Suzuki, K.; Anraku, Y. *J. Biol. Chem.* 1990, *265*, 6726-6733.
32. Kane, P.; Yamashiro, C.; Wolczyk, D.; Neff, N.; Goebl, M.; Stevens, T. *Science* 1990, *250*, 651-657.
33. Perler, F. B.; Davis, E. O.; Dean, G. E.; Gimble, F. S.; Jack, W. E.; Neff, N.; Noren, C. J.; Thorner, J.; Belfort, M. *Nucleic Acids Res.* 1994, *22*, 1125-1127.
34. Xu, M. Q.; Southworth, M. W.; Mersha, F. B.; Hornstra, L. J. *Cell* 1993, *75*, 1371-1377.
35. Xu, M. Q.; Perler, F. B. *EMBO J.* 1996, *15*, 5146-5153.
36. Dawson, P. E.; Muir, T. W.; Clark-lewis, I.; Kent, S. B. H. *Science* 1994, *266*, 776-779.
37. Tam, J. P.; Lu, Y. A.; Liu, C.-F.; Shao, J. *Proc. Natl. Acad. Sci. U. S. A.* 1995, *92*, 12485-12489.
38. Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U. S. A.* 1998, *95*, 6705-6710.
39. Hojo, H.; Onuma, Y.; Akimoto, Y.; Nakahara, Y.; Nakahara, Y. *Tetrahedron Lett.* 2007, *48*, 25-28.
40. Nakamura, K.; Sumida, M.; Kawakami, T.; Vorherr, T.; Aimoto, S. *Bull. Chem. Soc. Jpn.* 2006, *79*, 1773-1780.
41. Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. *Org. Lett.* 2010, *12*, 5238-5241.

42. **Hou, W.; Zhang, X.; Li, F.; Liu, C.-F.** ***Org. Lett.*** **2011,** ***13*****, 386-389.**

43. **Wong, C. T. T.; Taichi, M.; Nishio, H.; Nishiuchi, Y.; Tam, J. P.** ***Biochemistry*** **2011,** ***50*****, 7275-7283.**

44. **Gunasekera, S.; Daly, N. L.; Clark, R. J.; Craik, D. J.** ***Antioxid. Redox Signal*** **2009,** ***11*****, 971-980.**

Supplementary Material

These data include experimental procedures, characterization of PCR products and plasmids by agarose gel, SDS-PAGE gel analysis of expressed fusion proteins, all MALDI-TOF mass spectra of reduced-cyclic/hydrolyzed/native hB1 and B1-MES thioester, and chemically synthesized compounds **4-8**.

Supplementary material

Peptide macrocyclization through transpeptidation

Xinya Hemu[a] , Yibo Qiu[a,] and James P. Tam[a] *

[a]*School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, 03s-71, Singapore 637551*

Table of Content

Table S1. Primer design for tandem-extension PCR to amplify CRP sequences flanked by NdeI/SapI restriction sites. The longer nucleotide lengths and higher GC content of the complementary region on 2f/3r conferred higher melting temperature (Tm) for the first extension reaction to prepare double-strand template for the second extension. The annealing temperatures based on Tm were set as 54˚C for 1st reaction and 47˚C for the 2nd reaction.

Primer	Sequence	Tm (˚C)
B1R-1f	5′-GGTGGT**CATATG**TGCGGTGAGACTTGCTTCGTTTTACCGTG	51.3
B1R-2f	5′-TGCTTCGTTTTACCGTGCTGGTCCGCCAAGTTCGGCTGCTACTGCC	59.1
B1R-3r	5′-CCGTTGCGGTAACAAAAACCCTTTTGGCAGTAGCAGCCGAACTT	53.1
B1R-4r	5′-GGTGGTT**GCTCTTC**CGCAGCGAGTGCCGTTGCGGTAACAAA	
B1T-1f	5′-GGTGGT**CATATG**TGCTTCGTTTTACCGTGCTGGTCCGCCAAGTTCGG	53.8
B1T-2f	5′-TCCGCCAAGTTCGGCTGCTACTGCCAAAAGGGTTTTTGTTAC	57.6
B1T-3r	5′-ACCGCAGCGAGTGCCGTTGCGGTAACAAAAACCCTTTTGGCA	54.1
B1T-4r	5′-GGTGGTT**GCTCTTC**CGCAAGTCTCACCGCAGCGAGTGC	
B1P-1f	5′-GGTGGT**CATATG**TGCTGGTCCGCCAAGTTCGGCTGCTACTGCC	53.2
B1P-2f	5′-TCGGCTGCTACTGCCAAAAGGGTTTTTGTTACCGCAACGGCAC	58.6
B1P-3r	5′-AACGAAGCAAGTCTCACCGCAGCGAGTGCCGTTGCGGTAACA	51.9
B1P-4r	5′-GGTGGTT**GCTCTTC**CGCACGGTAAAACGAAGCAAGTCTCACC	
B1G-1f	5′-GGTGGT**CATATG**TGCTACTGCCAAAAGGGTTTTTGTTACCGCAACGG	53.1
B1G-2f	5′-TTTTGTTACCGCAACGGCACTCGCTGCGGTGAGACTTGCTTCG	59.1
B1G-3r	5′-GAACTTGGCGGACCAGCACGGTAAAACGAAGCAAGTCTCACCGC	52.5
B1G-4r	5′-GGTGGTT**GCTCTTC**CGCAGCCGAACTTGGCGGACCA	
B1Y-1f	5′-GGTGGT**CATATG**TGCCAAAAGGGTTTTTGTTACCGCAACGG	53.1
B1Y-2f	5′-TTTGTTACCGCAACGGCACTCGCTGCGGTGAGACTTGCTTCG	59.1
B1Y-3r	5′-GAACTTGGCGGACCAGCACGGTAAAACGAAGCAAGTCTCACCGC	52.5
B1Y-4r	5′-GGTGGTT**GCTCTTC**CGCAGTAGCAGCCGAACTTGGCGGACCA	
B1F-1f	5′-GGTGGT**CATATG**TGTTACCGCAACGGCACTCGCTGCGGTGAG	53.0
B1F-2f	5′-ACTCGCTGCGGTGAGACTTGCTTCGTTTTACCGTGCTGGTCCGC	58.6
B1F-3r	5′-TTTGGCAGTAGCAGCCGAACTTGGCGGACCAGCACGGT	54.3
B1F-4r	5′-GGTGGTT**GCTCTTC**CGCAAAAACCCTTTTGGCAGTAGCAGCC	

NdeI/SapI restriction sequences on primers 1f/4r were in bold. Complementary sequences on 2f/3r pair for the first annealing were highlighted in blue. The annealing sequences for the second extension reaction on 1f/2f and 3r/4r were highlighted with red and orange, respectively.

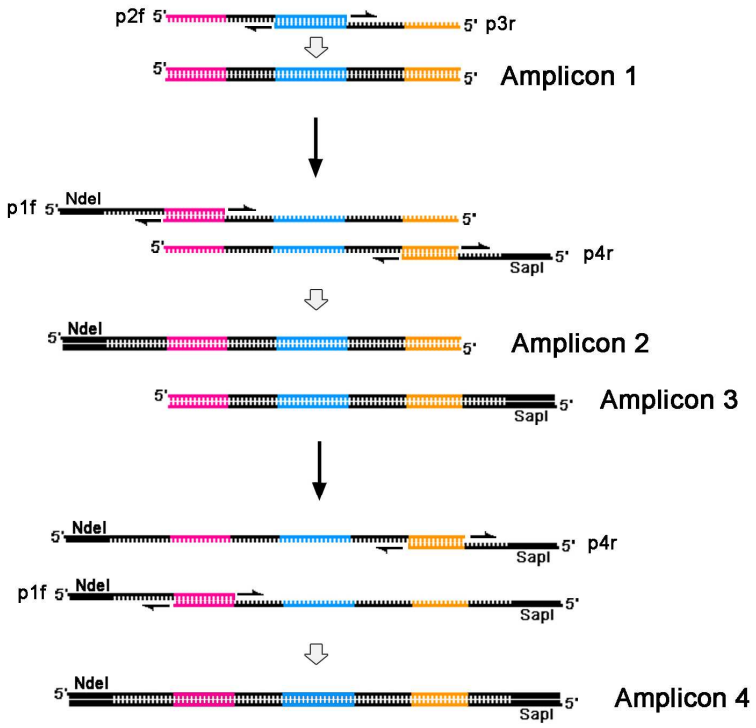


Figure S1. **Schema illustration of tandem-extension PCR**. Forward primer p1f carried NdeI restriction site and 5'-end of CRP cDNA. Forward primer p2f contained 14-17bp overlapping sequence that complementary to p1f (red) and 15-21bp overlapping sequence that complementary to p3r (blue). Reversed primer p3r overlapped with p2f (blue) and p4r (orange). Reversed primer p4r overlapped with p3r (orgnge) and carried SapI restriction site on its 3'-end. In the first chain reaction with higher annealing temperature (56-58°C), only p2f and p3r participated in the polymerization to generate amplicon 1 for 5 runs. In the second chain reaction, p1f and p4r annealed with amplicon 1 at lower annealing temperature (51-53°C) to produce amplicon 2 and 3. In the following runs, all four primers participated in extension reactions to produce amplicon 4, the expected product that possessed the full length of CRP cDNA sequences flanked by NdeI/SapI restriction sites.

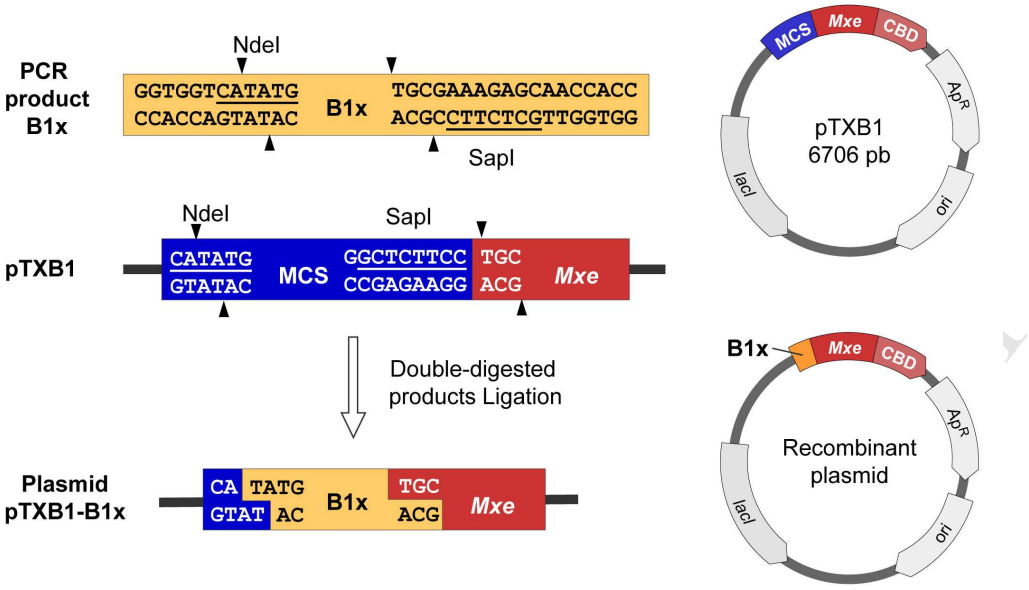


Figure S2. **Preparation of recombinant plasmid pTXB1-B1x**. The PCR products B1x that flanked with enzyme binding sites (orange) and pTXB1 vector (MCS in blue, *Mxe* intein in red) were subjected to NdeI/SapI double digestion. Ligation of the digested products gave a recombinant plasmid with the in-frame insertion of B1x sequence upstream of *Mxe* intein.

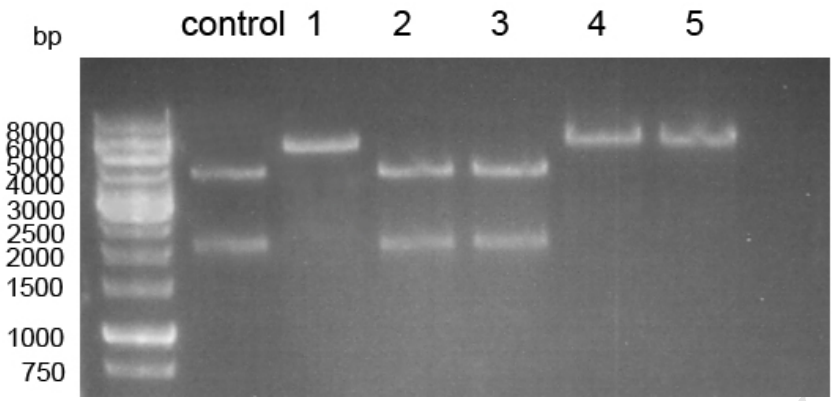


Figure S3. EcoRI/EcoRV double digestion screening of plasmids with PCR product insertion. Vector pTXB1 was used as negative control, which was digested into two fragments. Five colonies were picked for overnight culture. Plasmids extracted by mini-prep were subject to double digestion and agarose gel electrophoresis. Plasmids 1, 4 and 5 were suggested to contain PCR product insertion, which replaced MCS region containing EcoRI restriction site. Plasmids 2 and 3 were empty vectors.

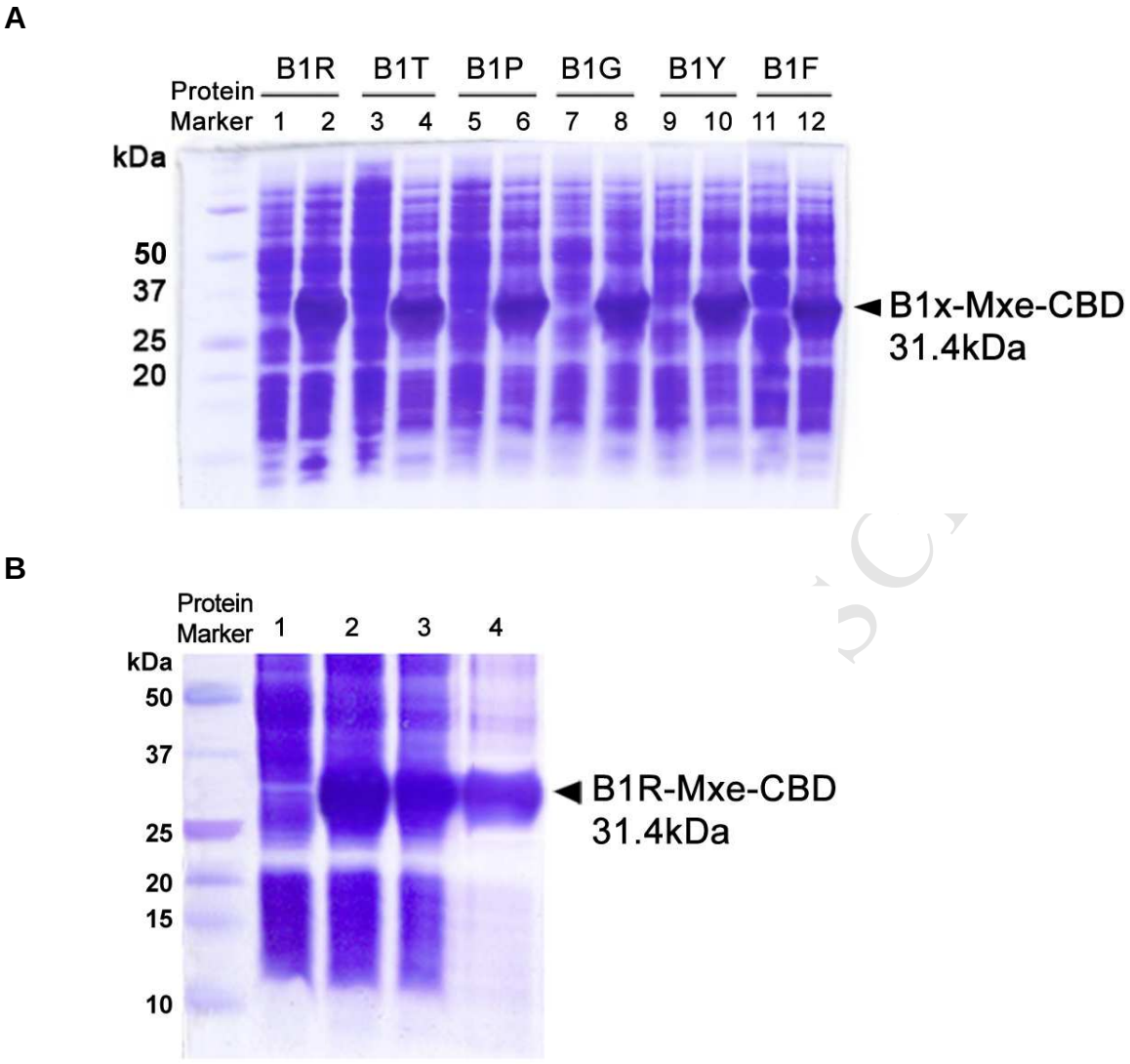
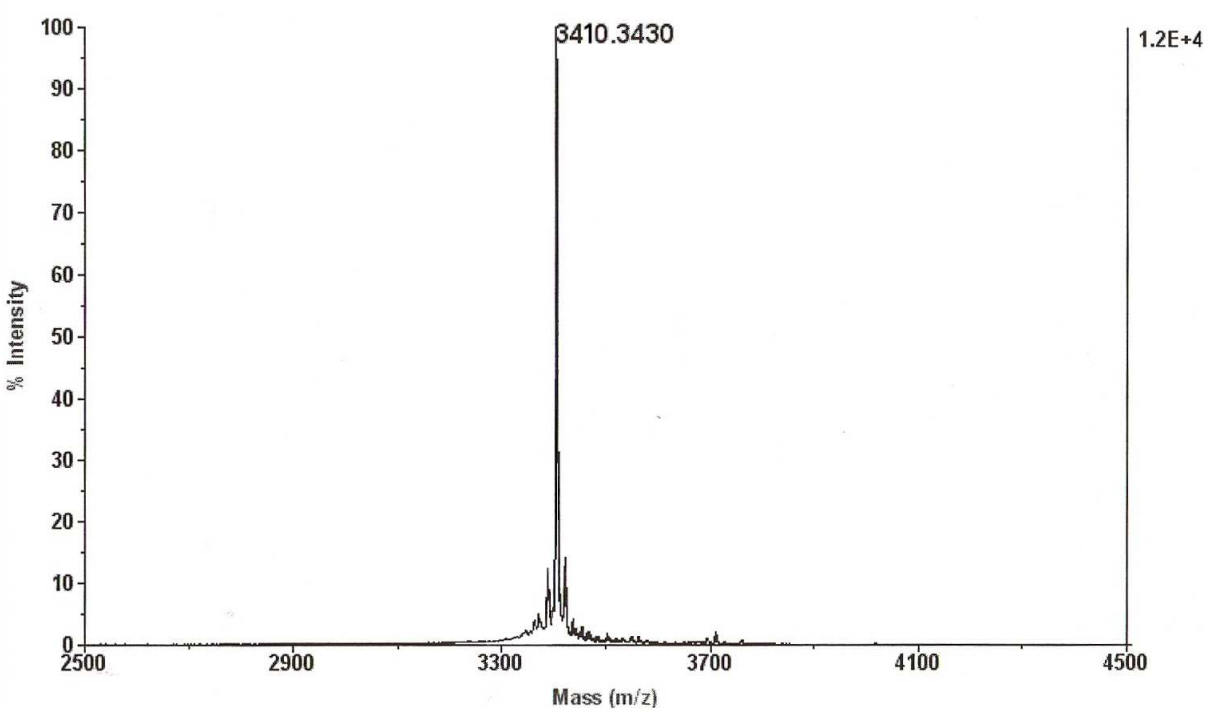
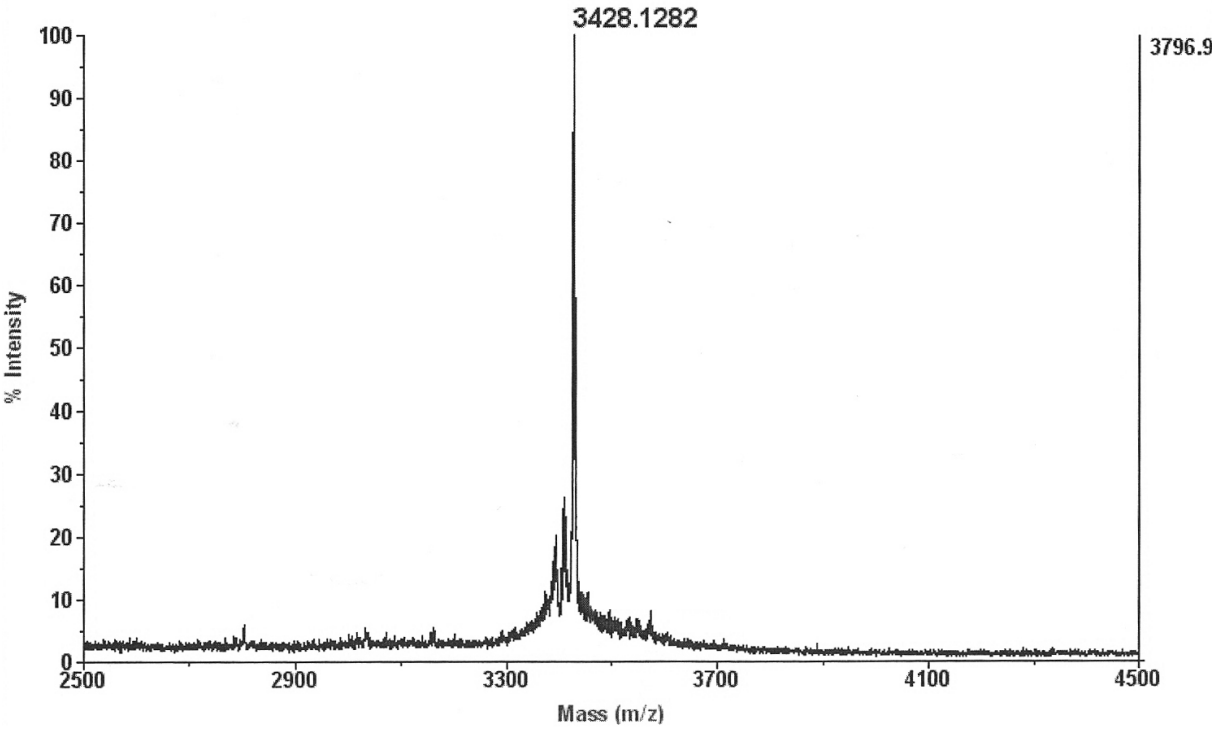


Figure S4. **Expression and extraction of fusion protein B1X-*Mxe*-CBD**. (A) The overall expression level of fusion proteins (X for R, T, P, G, Y or F). Lanes in odd number: cell lysate before IPTG induction. Lanes in even number: cell lysate after IPTG induction (B) The solubility of single-intein fusion protein B1R-Mxe CBD. Lane 1: cell lysate before induction. Lane 2: cell lysate after induction. Lane 3: insoluble cell pellet containing inclusion bodies. Lane 4: clarified cell extraction.

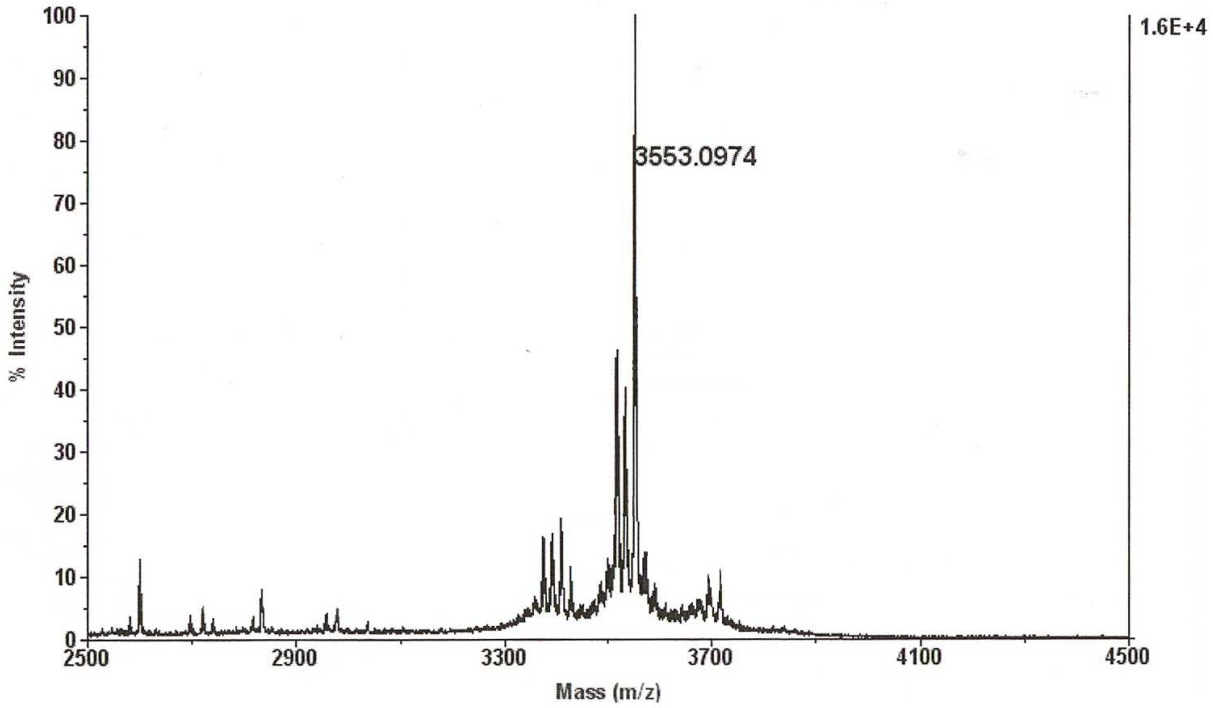
Cyclic reduced hB1 (Calc. 3410.8, Fd. 3410.3)



Linear hydrolyzed hB1 (Calc. 3428.9, fd 3428.1)



B1P-MES thioester (Calc. 3551.2, Fd. 3552.1)



Native-folded cyclic hB1 (Calc. 3403.2, Fd. 3403.5)

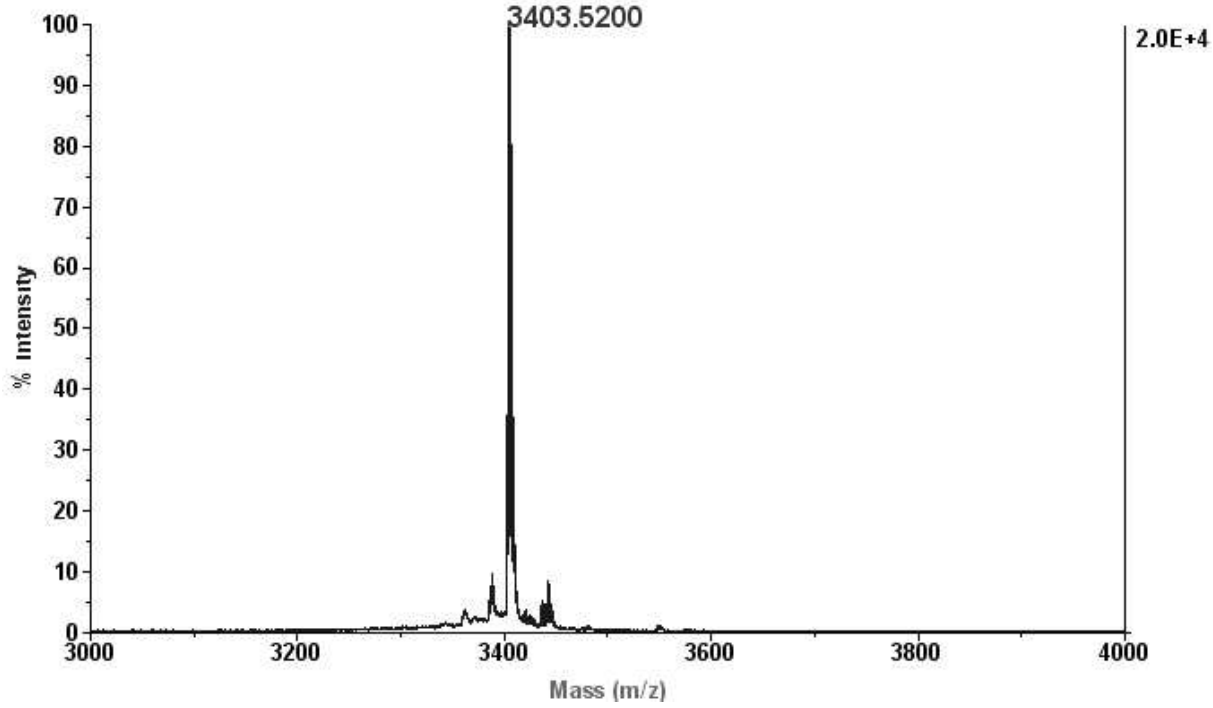
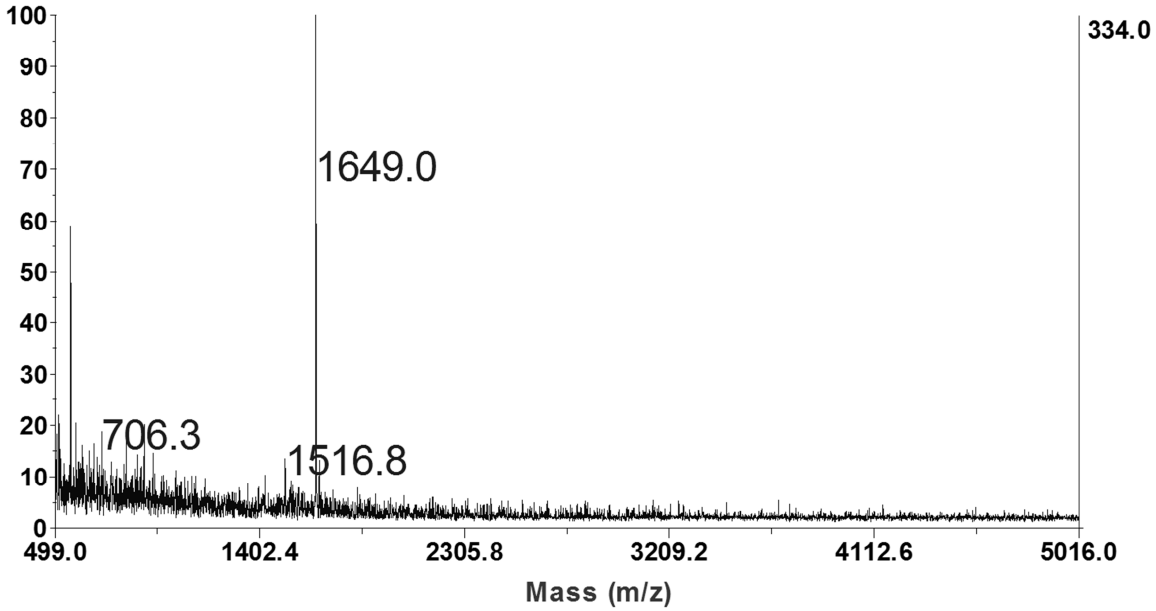
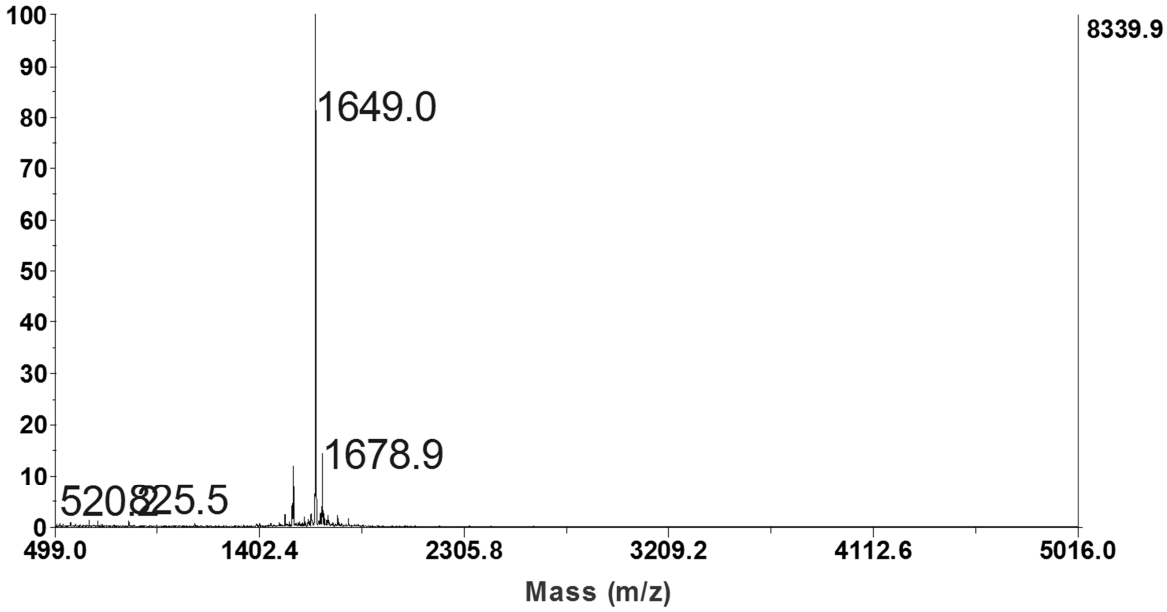


Figure S5. MALDI-TOF spectra of cyclic reduced hB1, hydrolyzed reduced hB1x, B1P-MES thioester and native-folded cyclic hB1.

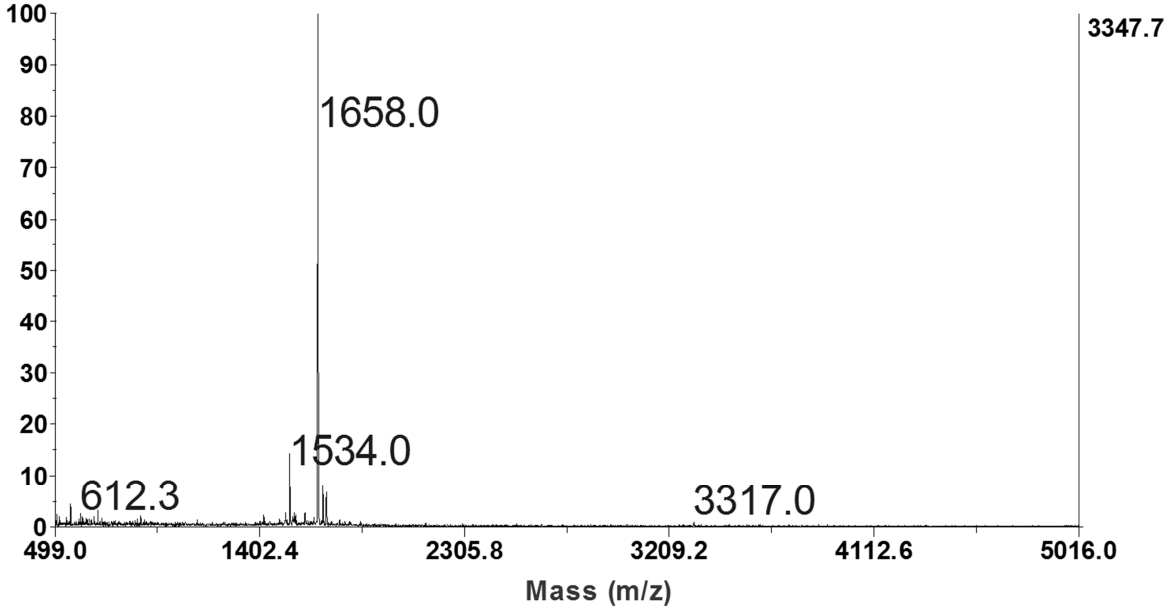
4a (calc. 1647.8, Fd. 1649.0)



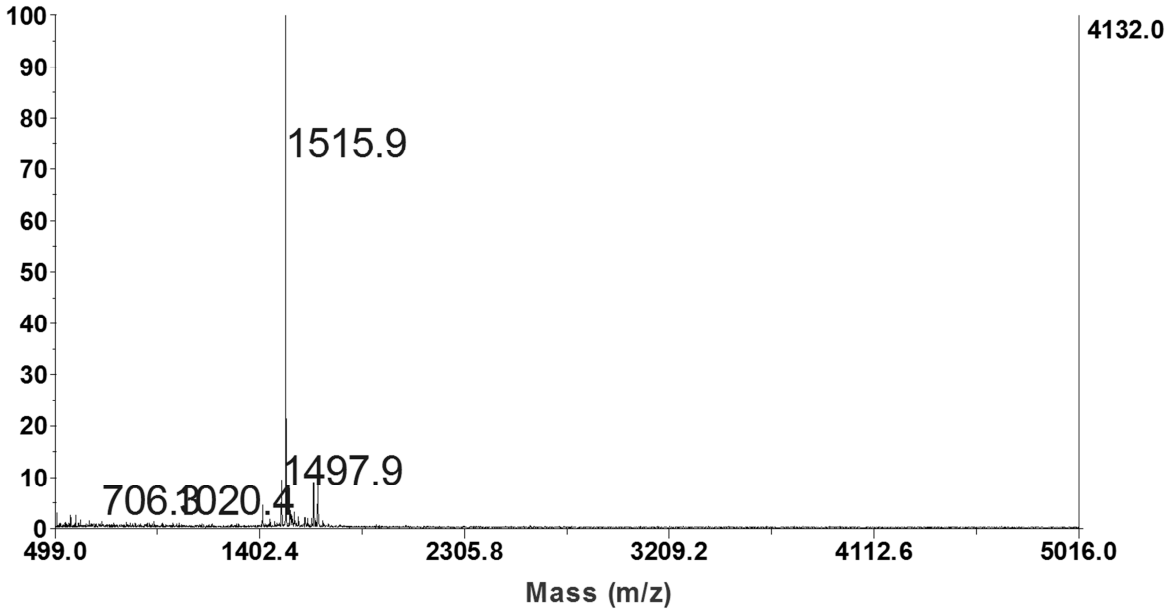
4b (calc. 1647.8, fd. 1649.0)



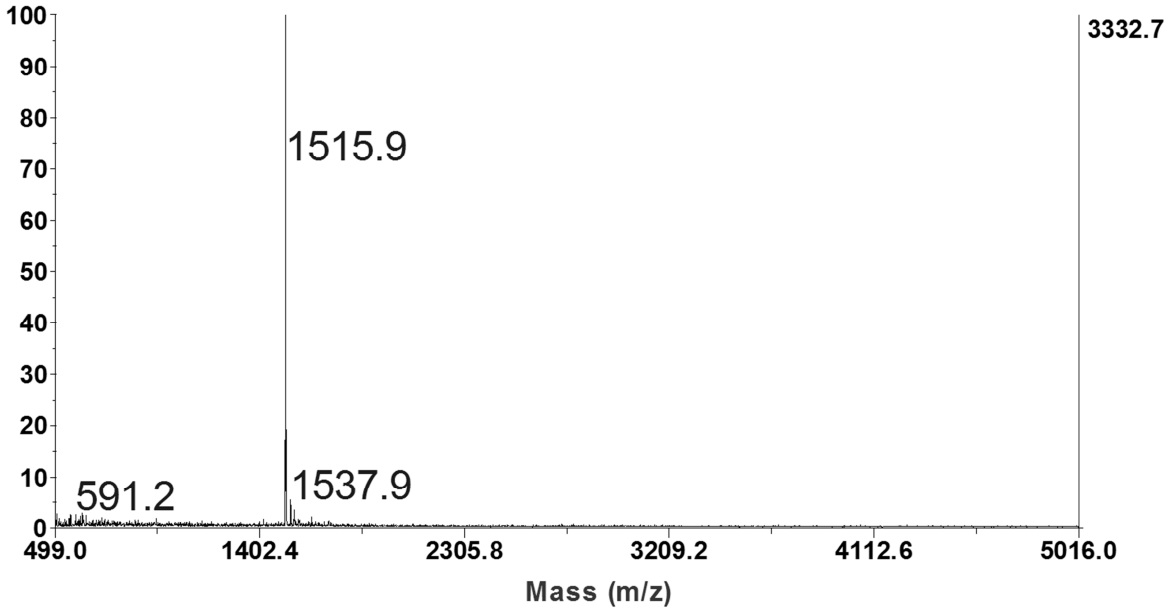
5 (calc. 1656.8, fd. 1658.0)



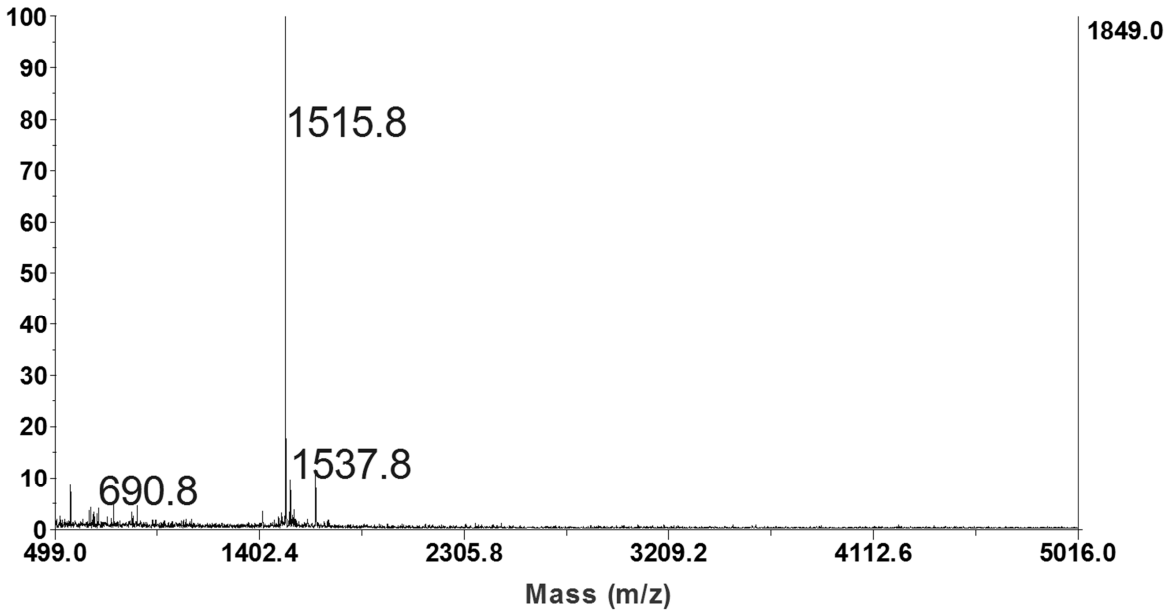
6a (calc. 1514.8, fd. 1515.9)



6b (calc. 1514.8, fd. 1515.9)



7 (calc. 1514.8, fd. 1515.8)



8 (calc. 1512.8, fd. 1513.7)

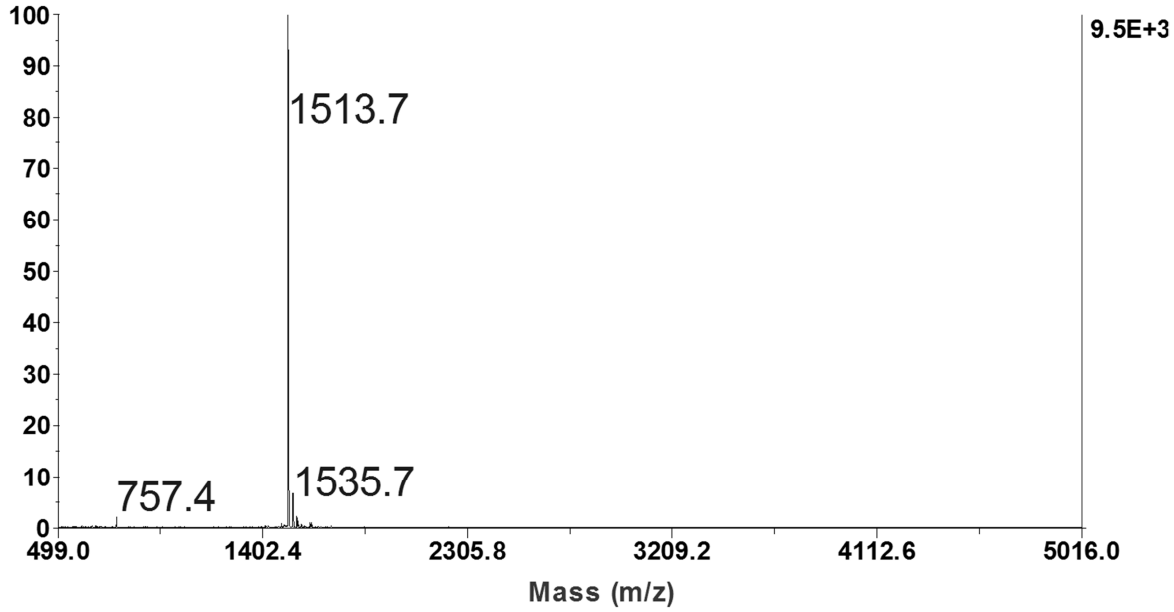


Figure S6. MALDI-TOF spectra of SFTI-TEBA (N-form **4a**), SFTI-TEBA (S-form **4b**), SFTI-MES thioester (**5**), 19-atom SFTI-thiolactone (**6a**), 43-atom SFTI-thiolactone (**6b**), cyclic reduced SFTI (**7**) and cyclic folded SFTI (**8**).