

Butelase-mediated Cyclization and Ligation of Peptides and Proteins

Giang K. T. Nguyen, Yibo Qiu, Yuan Cao, Xinya Hemu, Chuan-Fa Liu & James P. Tam*

School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore

Correspondence should be addressed to J.P.T. (jptam@ntu.edu.sg)

Editor Summary

This protocol from Nguyen *et al.* describes the use of the plant cyclase butelase 1 for the efficient cyclization and ligation of peptides and proteins. After extraction from *Clitoria ternatea*, the protocol describes reactions for cyclization, ligation, and synthesis of protein thioesters.

Keywords: Butelase, cyclization, Asx-specific ligase, site-specific conjugation

Abstract

Enzymes catalyzing efficient macrocyclization or site-specific ligation of peptides and proteins are enabling tools for drug design and protein engineering. Here we describe a protocol to use butelase 1, a recently discovered peptide ligase, for cyclization and ligation of peptides and proteins ranging in sizes from 10 to >200 residues with high efficiency. Butelase 1 is the fastest known ligase and is found in pods of the common medicinal plant *Clitoria ternatea* (also known as butterfly pea). It has a very simple C-terminal-specific recognition motif that requires Asn/Asp (Asx) at the P1 position and a C-terminal dipeptide His-Val at P1' and P2' positions. Substrates for butelase-mediated ligation can be prepared by standard Fmoc chemistry or recombinant expression with the minimal addition of this tripeptide Asn-His-Val motif at the C-terminus. Butelase 1 achieves cyclizations that are 20,000 times faster than sortase A, a commonly used enzyme for backbone cyclization. Unlike sortase A, butelase is traceless and can be used for the total synthesis of naturally occurring peptides and proteins. Furthermore, butelase 1 is also useful for intermolecular ligations and synthesis of peptide or protein thioesters which are versatile activated intermediates necessary and compatible with many chemical ligation methods. The protocol describes steps for isolation and purification of butelase 1 from plant extract using a four-step chromatography procedure which takes about three days. We then describe steps for intramolecular cyclization, intermolecular ligation, and butelase-mediated synthesis of protein thioesters. Butelase reactions are generally complete within minutes and often achieve excellent yields.

INTRODUCTION

Peptide ligases, enzymes that form peptide bonds with site specificity and substrate selectivity, are highly desirable tools for engineering, chemoselective labeling and macrocyclization of proteins¹⁻⁴. However, they are exceedingly rare with only few known ligases such as PATG, PCY1, POBP and sortase A. The first three ligases are specific cyclases involved in the biosynthesis of small cyclic peptides ranging from 5-11 residues in cyanobacteria, plants, and mushrooms, respectively⁵⁻⁸, whereas sortase A is a transpeptidase known for anchoring surface proteins to bacterial cell walls. These ligases, however, exhibit slow kinetics or narrow substrate specificity, disadvantages that limit their broad applications. A ligase with broad substrate specificity and efficient kinetic would be highly desirable.

Discovery of butelase 1 and diversity of applications

Recently, we discovered a novel ligase, an Asn/Asp (Asx)-specific ligase called butelase 1 from butterfly pea (*Clitoria ternatea*) which is known locally as bunga telang from which the name butelase is derived^{9, 10}. *C. ternatea* is a common medicinal plant found in Southeast Asia, which is also widely used for food dye, shampoo, fodder, and ornamental. Butelase 1 can be purified from pods of *Clitoria ternatea* by a four-step chromatographic procedure to give about 5 mg enzyme per kg of fresh plant material. Butelase 1 is a cyclase involved in the biosynthesis of cyclotides, the largest family of plant cyclic peptides¹⁰⁻¹⁴. The first cyclotide kalata B1 was reported in the early 1970s, its structure determined in 1995 and total synthesis in 1997¹⁵⁻¹⁷. Thus, our work on butelase 1 provides a satisfactory closure on the biosynthetic processing of how cyclotides form a head-to-tail macrocycle, some 40 years after its discovery⁹.

Butelase 1 exhibits kinetics unmatched by the other known ligases, with catalytic efficiency attaining $1,340,000 \text{ M}^{-1} \text{ s}^{-1}$ for

medium-size peptides, and >10,000 times faster than PATG, PCY1, and sortase A¹⁸. It is useful for both intra- and intermolecular ligation, efficiently cyclizing or ligating various peptides and proteins ranging in size from 10 to >200 amino acids¹⁸⁻²⁰. Importantly, butelase 1 is C-terminus-specific for Asx but accepts most N-terminal amino acids to mediate intermolecular peptide and protein ligation under mild conditions. As such, butelase 1 allows selective ligation of novel functional groups and biochemical probes into proteins without affecting their functions. Thus, the high catalytic efficiency coupled with a broad substrate specificity of butelase 1 could augment new applications, both in vitro and in vivo systems for basic and translational research (**Fig. 1**).

Recently a homolog of butelase 1, OaAEP1b, has been recombinantly expressed in *E. coli*²¹. However, the expression yield remains low and the purification requires a cumbersome western-blot-guided chromatography to identify fractions containing the target enzyme. Furthermore, the catalytic activity of OaAEP1b is 500 times slower than butelase 1, which would require up to 22 h for peptide cyclization as compared to only 5 min of butelase 1.

Advantages of the method

Butelase 1 is the fastest known ligase with a very high catalytic efficiency. A typical butelase-mediated reaction requires 100- to 1000-fold less enzyme than with sortase A or PATG^{6,22}. Butelase 1 is C-terminus-specific for Asx with a sorting sequence of two amino acids or less after Asx. Importantly, no part of the recognition sequence except for Asx is left behind in the cyclized or ligated products, a feature making butelase 1 an attractive candidate for the total synthesis of natural occurring linear or circular proteins. Butelase 1 exhibits a very broad specificity for the acceptor nucleophilic amino acids (all natural amino acids except Pro) to form a new Asx-Xaa peptide bond.

Compared to chemical methods such as native chemical ligation or Click chemistry²³⁻²⁶, butelase requires no chemical modifications of the peptide substrates. There is no requirement for C-terminal thioester, peptide azide or N-terminal cysteine to allow ligation. The substrates for butelase 1 can be readily prepared by standard Fmoc chemistry or recombinant expression. These advantages make butelase 1 a powerful new tool for peptide and protein engineering.

Experimental Design

Butelase 1 can be used for both intramolecular ligation (cyclization) and intermolecular ligation (peptide or protein synthesis, conversion to a thioester, site-specific conjugation and labeling). In this protocol, we define substrate specificity of butelase 1 and the design concept to direct butelase 1 for cyclization or ligation (**Fig. 2**).

Overview of the protocol In the procedure section, we first describe steps to purify butelase 1 from *Clitoria ternatea* plant extracts. We then describe detail protocols for butelase-mediated intramolecular cyclization, intermolecular ligation, and synthesis of protein thioesters.

Butelase-mediated macrocyclization

Macrocyclization reduces conformational entropy and often enhances proteolytic stability, thermostability, pharmacological potency, and in certain cases, oral bioavailability²⁷⁻²⁹. Butelase 1 can cyclize a broad range of peptides and proteins ranging in sizes from 10 to >200 residues with up to 95% yields. We estimate that, under our proposed experimental conditions, butelase 1 could be up to 200,000-fold faster than sortase A. Butelase 1 is also the only known naturally-occurring cyclase with the ability to cyclize large proteins.

To efficiently apply butelase 1 for peptide macrocyclization, an understanding of substrate scope at the N- and C-termini is required. For butelase-mediated cyclization, a tripeptide recognition sequence Asx-His-Val must be present at the C-terminus. For the N-terminus, butelase 1 display a loose specificity for the residue at the P1" position but a more stringent requirement at the P2" position. Probing the substrate specificity by combinatorial libraries showed that butelase 1 accepts 20 natural amino acids at the P1" position except for proline, and highly favors bulky hydrophobic amino acids such as Ile, Leu, Val and to some extent Cys at the P2" position. Butelase 1 has no substrate requirement beyond P2" residue.

Butelase-mediated intermolecular peptide ligation

Although the native function of butelase 1 is for macrocyclization, we can direct it for intermolecular peptide ligation by manipulating residues at the N-terminus. The components of an intermolecular ligation reaction are: butelase 1, an Asx-His-Val containing substrate (peptide **1**) and a suitable acceptor nucleophile (peptide **2**). For an intramolecular cyclization, both the acceptor nucleophile and the Asx-His-Val recognition sequence are found in the same peptide. To utilize butelase 1 for intermolecular ligation, these two components have to be separated into two different peptides. Peptide **1** carrying the C-terminal Asx-His-Val recognition sequence must not contain the Ile/Leu/Val/Cys residue at the P2" position at the N-terminus. On the other hand, peptide **2** must contain Ile/Leu/Val/Cys residue at the P2" position.

It should be noted that unlike a cyclization reaction which often proceed in high yield due to the intramolecular location of an acceptor nucleophile, the intermolecular ligation is reversible as the released dipeptide His-Val acts as a competing nucleophile with the incoming peptide **2**. To improve the ligation yield, either peptide **1** or peptide **2** have to be used in excess (>5 folds).

Butelase-mediated synthesis of protein thioesters

A peptide/protein thioester is a versatile intermediate required for various chemical ligation methods³⁰. Currently, the intein-based approach is the most popular method for both protein cyclization and thioesterification^{3, 31, 32}. However, the requirement for genetic fusion of the target protein with the intein domain may limit its applications in many circumstances due to the unpredictability of the folding and solubility of the fusion protein. The minimal substrate modification of butelase 1, which only requires an addition of a C-terminal tripeptide Asn-His-Val, would provide an alternative and complimentary approach to the intein method. We found that butelase 1 efficiently ligates various glycine thioesters with proteins to generate the desired protein thioesters (Fig. 3)¹⁹. This strategy has been applied to prepare protein thioesters with good yields (up to 90%).

MATERIALS

REAGENTS

CAUTION! Many of the reagents are potentially harmful. Wear appropriate protective clothing and safety glasses, and work in a fume hood.

- Pods of *Clitoria ternatea* (~1 kg) were purchased from a local nursery
- Ammonium sulfate (Regent Chemicals, cat. no. 011566)
- *N,N*-Diisopropylethylamine (DIEA; Sigma-Aldrich, cat. no. D125806, 500 ml)
- (Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP; Chem-Impex International, cat. no. 02276, 1 kg)
- Dichloromethane (DCM; Fisher, cat. no. ZZLFK/D/1852/17, 2.5 liters)
- Trifluoroacetic acid (TFA; Alfa Aesar, cat. no. L06374, 500 ml)
- Acetonitrile (ACN; Fisher, cat. no. ZZLFS/A998-4, 4 liters)
- Ethylenediaminetetraacetic acid (EDTA; Bio-Rad, cat. no. 1610729, 500 g)
- β -mercaptoethanol (β -ME; Sigma-Aldrich, cat. no. M6250, 250 ml)
- Sodium dihydrogen phosphate monohydrate (Merck, cat. no. 1.06346.1000, 1 kg)
- Disodium hydrogen phosphate dihydrate (Merck, cat. no. 1.06580.1000, 1 kg)
- Phenylmethylsulfonyl fluoride (PMSF; Sigma-Aldrich, cat. no. P7626, 25 g)
- Q-Sepharose Fast Flow (GE Healthcare, cat. no. 17-0510-01, 300 ml)
- Boc-Gly-OH (GL biochem, cat. no. 30701, 25 g)
- 3-Methyl-1-butanethiol (MBT; Sigma-Aldrich, cat. no. 115924, 5 g)
- Hexane (Fischer chemical, cat. no. 142641, 4 liters)
- Ethyl acetate (Fischer chemical, cat. no. 1545128, 2.5 liters)
- Ethanol (Sigma-Aldrich, cat. no. 277649, 1 liter)

Limitations of the method

For intramolecular cyclization of large peptides (>50 residues) and proteins in general, the N- and C-termini have to be in close proximity to allow efficient reaction. Under our experimental conditions, butelase 1 cannot cyclize denatured proteins and requires a refolding step to achieve a properly folded protein for N-to-C terminal cyclization. Furthermore, we have also found that butelase 1 cannot efficiently cyclize peptide substrates smaller than 9 residues.

For intermolecular ligation, a limitation is that butelase 1 requires an excess amount of a substrate (>5 fold) to compete with the cleaved dipeptide His-Val which acts as a competitive nucleophile to reverse the ligation reaction. Recently, we have shown that this limitation can be minimized by using a thiopeptide as a substrate to render the ligation reaction irreversible²⁰.

- Ninhydrin (Alfa Aesar, cat. no. 10170490, 50 g)
- Sodium chloride (Merck, cat. no. 1.06404.5000, 5 kg)
- Target peptides or proteins: Target peptides for intramolecular cyclization or intermolecular ligation reactions can be produced by solid-phase synthesis or purchased from commercial vendors (e.g. GL Biochem, Shanghai, China). Target proteins for cyclization can be produced by recombinant expression in *E. coli* or a suitable expression system³³
- Peptides to investigate ligase activity e.g. sunflower trypsin inhibitor (SFTI) peptide (GRCTKISPPICFPNHV) (GL Biochem, Shanghai, China)

EQUIPMENT

- HPLC (Shimadzu, CBM-20A)
- MALDI-TOF MS (AB SCIEX 4800 Plus MALDI TOF/TOFTM Analyzer)
- Thermo Finnigan LCQ DECA XP MAX (Thermo Fisher Scientific)
- Q Exactive hybrid quadrupole-Orbitrap (Thermo Fisher Scientific)
- Freeze dryer (Virtis sentry 2.0)
- Vacuum pump (Knf Lab N 838.1.2 K.18)
- Milli-Q[®] Integral Water Purification System
- NanoDrop 2000c Spectrophotometer
- Blender (Philips, 650W) with a 2-liter glass jar
- Nalgene Filter Units, 1000 ml volume, 0.2 μ m (597-4520) or 0.45 μ m (167-0045)
- Filter paper (Whatman, Grade 54)
- SnakeSkin dialysis tubing with a molecular weight cut-off (MWCO) of 10 kDa (ThermoFisher, cat. no. 88245)
- Amicon Ultra Centrifugal Filter (Millipore, cat. no. UFC901024)
- Biosuite 250, 13 μ m SEC column (Waters, 21.5 x 300 mm, cat. no. 186002171)
- PolyWax analytical anion-exchange column (Poly LC, 4.6 x 200 mm, cat. no. 204WX05)
- PolyWAX preparative anion-exchange column (PolyLC, 21 x 250 mm, cat. no. 2521WX12)

- C18 analytical column (Phenomenex, Aeris PEPTIDE XB-C18, 3.6 μm , 4.6 mm x 250 mm, cat. no. 00G-4507-E0)
- C18 preparative column (Phenomenex, Jupiter, 5 μm , 300 \AA , 21.2 mm x 250 mm, cat. no. 00G-4053-PO)
- Buchner funnels, 550 ml (VWR; cat. no. 89038-126) and 1.86 liter (VWR; cat. no. 89038-132)
- Buchner flask, 4 liter (VWR; 89000-391)
- 400 MHz NMR spectrometer
- Thin-layer chromatography (TLC) plates
- Single-neck round-bottomed flasks, 25-500 ml
- Magnetic stirring plate
- Magnetic stir bars
- Speedvac
- Heat gun
- MagTran 1.03 (<http://www.ionsource.com/links/programs.htm>)
- ESIProt 1.0 (<http://www.bioprocess.org/esiprot/ESIProt-1.0.exe>)

REAGENT SETUP

Ninhydrin solution Add 0.15 g ninhydrin into 50 ml ethanol. Solution can be stored at 25 $^{\circ}\text{C}$ for 3 months.

Ligation and cyclization buffer The ligation and cyclization buffer contains 20 mM phosphate buffer, and 1 mM EDTA, pH 6. Prepare 1 liter of 1X buffer by adding 200 ml of 0.1 M phosphate buffer pH 6, 292.2 mg EDTA and 800 ml of water. Buffer can be stored at 25 $^{\circ}\text{C}$ for 6 months.

0.1 M Phosphate Buffer Prepare 0.1 M NaH_2PO_4 buffer by dissolving 13.8 g of $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (monobasic, m.w. = 138) in water to make a final volume of 1 liter. Prepare 0.1 M Na_2HPO_4 buffer by dissolving 17.8 g $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ (dibasic, m.w. = 178) in water to make a final volume of 1 liter. Mix 877 ml of 0.1 M NaH_2PO_4 and 123 ml of 0.1 M Na_2HPO_4 stock solutions to make a 0.1 M phosphate buffer, pH 6 (5X stock). Buffer can be stored at 25 $^{\circ}\text{C}$ for 6 months.

1 M Sodium Phosphate Prepare 1 M NaH_2PO_4 buffer by dissolving 138 g of $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (monobasic, m.w. = 138) in water to make a final volume of 1 liter. Prepare 1 M Na_2HPO_4 buffer by dissolving 178 g $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ (dibasic, m.w. = 178) in water to make a final volume of 1 liter. Mix 877 ml of 1 M NaH_2PO_4 and 123 ml of 1 M Na_2HPO_4 stock solutions to make a 1 M phosphate buffer, pH 6. Buffer can be stored at 25 $^{\circ}\text{C}$ for 6 months.

0.5 M EDTA Prepare 0.5 M EDTA solution by dissolving 146 g of EDTA (m.w. = 292) in 800 mL of water. Adjust the pH to 8.0 with NaOH and bring the volume to 1 L with water.

Extraction buffer EB1 Prepare 1 liter of EB1 buffer (20 mM sodium phosphate, 1 mM EDTA, 1 mM PMSF, 5 mM β -ME, pH 6.0) by combining 20 ml of 1 M sodium phosphate, 2 ml of 0.5 M EDTA, 174

mg of PMSF, and 0.35 ml of β -ME. Top up with water to 1 liter and adjust the pH to 6.0 with NaOH and HCl. Due to instability of PMSF in aqueous solution, prepare fresh buffer before use.

Extraction buffer EB2 Prepare 1 liter of EB2 buffer (EB1 without PMSF) by combining 20 ml of 1 M sodium phosphate, 2 ml of 0.5 M EDTA, and 0.35 ml of β -ME. Top up with water to 1 liter and adjust the pH to 6.0 with NaOH and HCl. Buffer can be stored at 25 $^{\circ}\text{C}$ for 6 months.

Extraction buffer EB3 Prepare 1 liter of EB3 buffer (EB2 + 0.2 M NaCl) by combining 20 ml of 1 M sodium phosphate, 2 ml of 0.5 M EDTA, 0.35 ml of β -ME, and 11.7 g NaCl. Top up with water to 1 liter and adjust the pH to 6.0 with NaOH and HCl. Buffer can be stored at 25 $^{\circ}\text{C}$ for 6 months.

Extraction buffer EB4 Prepare 1 liter of EB4 buffer (EB2 + 1 M NaCl) by combining 20 ml of 1 M sodium phosphate, 2 ml of 0.5 M EDTA, 0.35 ml of β -ME, and 58.5 g NaCl. Top up with water to 1 liter and adjust the pH to 6.0 with NaOH and HCl. Buffer can be stored at 25 $^{\circ}\text{C}$ for 6 months.

EQUIPMENT SETUP

HPLC Analyze peptides using a Shimadzu CBM-20A HPLC system (Tokyo, Japan) equipped with a Phenomenex analytical C18 column (Aeris PEPTIDE XB-C18, 00G-4507-E0, 3.6 μm , 100 \AA , 4.6 mm x 250 mm) with a linear gradient 1: 10-60% ACN/0.1% TFA for 25 min at a flow rate of 1 ml/min or gradient 2: 10-40% ACN/0.1% TFA for 35 min at a flow rate of 1 ml/min. For peptide purification, we use a Phenomenex preparative C18 column (Jupiter, 00G-4053-PO, 5 μm , 300 \AA , 21.2 mm x 250 mm) at a flow rate of 6 ml/min with a gradient of 10% to 60% ACN/0.1% TFA for 70 min.

MALDI-TOF MS Mass spectrometry was performed on an ABI 4800 MALDI-TOF/TOF system (Applied Biosystems, Framingham, MA, USA). The instrument was equipped with a solid-state laser (diode pumped Nd:YAG laser) pulsing at a repetition rate of 200 Hz. A 10 mg/mL solution of α -cyano-4-hydroxycinnamic acid in 60% acetonitrile, 0.05% TFA was used as a MALDI matrix. Samples were mixed thoroughly with the matrix solution at the ratio of 1:1 (v/v) and 0.5 μl of the mixture was spotted onto a target plate. The instrument was calibrated externally using a mixture of peptide standards obtained from Sigma-Aldrich (MSCal1). All mass spectra can be acquired using dual-stage reflectron mirror. The laser intensity was set between 3500-4500 with an accelerating voltage of 20 kV.

ESI-MS ESI-MS was performed on a Thermo Finnigan LCQ DECA XP MAX or a Q Exactive hybrid quadrupole-Orbitrap (Thermo Fisher). The deconvoluted data were obtained using the software of MegTran 1.03 and ESIProt 1.0.

PROCEDURE

Isolation and purification of butelase 1 • TIMING 3 d

CRITICAL We use a four-step chromatographic procedure involving flash anion exchange, preparative anion-exchange HPLC, size exclusion, and analytical anion-exchange HPLC to purify butelase 1. The entire purification process takes about three days.

1| *Preparation of plant extract (day 1, 3h):* Homogenize 1 kg of freshly collected pods of *Clitoria ternatea* with 2 liters of extraction buffer EB1 using a kitchen blender (Philips, 650W).

CRITICAL STEP To minimize protein degradation, perform the purification of butelase 1 in a 4 °C cold room.

PAUSE POINT Fresh plant material can be stored at -80 °C for up to 12 months.

2| Centrifuge at 9000 g for 10 min at 4 °C, and filtrate the plant extract through a Buchner funnel covered with a filter paper (Whatman Grade 54 or normal tissue paper worked well for us) to remove the plant debris.

3| Add solid ammonium sulfate salt to the plant extract to reach 15% saturation and incubate for 10 min. Centrifuge at 9000 g for 10 min at 4 °C to remove precipitated proteins. Transfer the aqueous plant extract to a clean beaker.

4| Add solid ammonium sulfate salt to the plant extract to reach 85% saturation and incubate for 20 min. Centrifuge at 9000 g for 30 min at 4 °C and discard the supernatant. Resuspend the precipitated proteins with 1 liter of buffer EB1 and dialyze overnight with 20 liters of buffer EB2 using a SnakeSkin dialysis tubing with a MWCO of 10 kDa (ThermoFisher).

5| After overnight dialysis, centrifuge the extract at 9000 g for 20 min and discard the precipitated proteins. Keep the supernatant for subsequent chromatographic fractionations. Filter the extract with Nalgene Filter Units (1000 ml receiver volume), with a pore size of either 0.2 or 0.45 µm.

6| *Fractionation by flash anion exchange and preparative anion exchange (day 2, 8 h):* Prepare the flash chromatography column by adding 300 ml slurry of Q-Sepharose Fast Flow anion-exchange resin (GE Healthcare) to a 600 ml Buchner funnel covered with a filter paper (Whatman Grade 54). Connect the funnel to a 4-liter Buchner flask with a tube leading to a vacuum pump.

7| Load the plant extract to the flash column and turn on the vacuum pump to accelerate the flow. The flow-through extract can be reloaded into the column for maximum yield. Discard the flow through.

8| Wash the column with 2 liters of buffer EB2 and elute with 800 ml of buffer EB3

9| Dialyze the eluent with 10 liters of buffer EB2 for 4 h.

CRITICAL STEP Ensure that no solution leak from the dialysis tubing.

10| Filter the dialyzed eluent using Nalgene Filter Units (1000 ml receiver volume) with a pore size of 0.2 µm and load the sample into a preparative PolyWAX anion-exchange column (PolyLC, 21 x 250 mm). The column must be balanced with buffer EB2 before loading.

CRITICAL STEP We use the solvent line C of a Shimadzu HPLC system, at a flow rate of 8 ml/min to load the sample into the HPLC anion-exchange column. Ideally, the HPLC system should be set up in a cold room or a 4 °C fridge to prevent protein degradation. Alternatively, the column and HPLC buffers can be submerged in ice boxes. Thoroughly purge the solvent line C with HPLC grade water after loading sample.

11| Prepare the HPLC buffers EB2 and EB4 for anion exchange separation. Fractionate the plant extract by running a linear gradient from 0% to 70% buffer EB4 in 70 min, at a flow rate of 8 ml/min (**Supplementary Fig. 1**). Collect 8 to 10 ml per fraction.

12| Screen for fractions containing ligase activity using a linear derivative of sunflower trypsin inhibitor (SFTI) peptide (GRCTKISPPICFPNHV or any substrates with a suitable recognition motif). We usually add 0.5 µl of each fraction to a 5-µl solution containing 20 µM peptide substrate (dissolved in buffer EB2) and incubate for 5 min (**Supplementary Fig. 2**). Immediately spot the samples on MALDI plate for MALDI-MS analysis. SFTI has a molecular mass of 1766 Da which is converted to 1512 Da in the presence of butelase 1. Butelase 1 generally elutes at about 25 to 40% buffer EB4.

PAUSE POINT Store the fractions at 4 °C overnight

13| *Size exclusion and analytical anion exchange (day 3, 8h)* Pool all the fractions containing ligase activity and concentrate to a final volume of 3 ml using an Amicon Ultra Centrifugal Filter (Millipore, Ireland Ltd) with a 10 kDa MWCO.

CRITICAL STEP Ensure that the membrane of the centrifugal filter does not leak and no ligase activity in the flow-through solution.

14| Fractionate the concentrated sample by size-exclusion chromatography using a Biosuite HPLC column (Waters, 21.5 × 300 mm) and buffer EB2 at a flow rate of 1.5 ml/min (**Supplementary Fig. 3**). Collect 2 ml per fraction.

15| Screen for fractions containing ligase activity as performed in step 12.

16| Further purify by anion-exchange chromatography using an analytical PolyWax HPLC column (Poly LC, 4.6 × 200 mm). Use a linear gradient from 0% to 70% buffer EB4 in 70 min, at a flow rate of 1 ml/min (**Supplementary Fig. 4**). Collect 1 ml per fraction.

17| Determine the purity of isolated butelase 1 by SDS-PAGE and Coomassie-blue staining. Butelase 1 migrates as a 37-kDa band in SDS gel.

? TROUBLESHOOTING

18| Determine the enzyme concentration by OD280 base on the predicted mature domain of butelase 1. The predicted ϵ_{280} for butelase 1 is 54780.

? TROUBLESHOOTING

19| Confirm the identity of butelase 1 by tryptic in gel digestion. The following fragments should be observed: TIVALIEDGTHVVQYGDVGLSK (MW = 2314), HQADVCHAYQLIK (1582 Da), WAVLVAGSK (930 Da), GYVNYR (771 Da).

20| For quality control of butelase 1 activity, we incubate the purified enzyme (0.1 μM) with SFTI peptide substrate (50 μM) for 10 min at 42 °C in the cyclization buffer. Approximately 35% of SFTI should be converted into the circular form based on the known kinetic parameter of butelase 1 for SFTI ($k_{cat} = 0.6 \text{ s}^{-1}$, $K_m = 51 \text{ }\mu\text{M}$). The conversion of SFTI from linear to circular form can be monitored by MS or HPLC.

PAUSE POINT Butelase 1 can be stored for 3-4 weeks at 4 °C. For long term storage, we freeze the butelase solution at -80 °C for up to 2 years at a concentration of 0.2-0.5 mg/ml. The enzyme should be aliquoted into small tubes with a volume of 25-50 μl per tube to avoid repeated freeze-thaw cycles.

CRITICAL The optimal pH for butelase 1 activity is in the range of 5.0 to 6.5. Exposure to pH above 7.5 will denature the enzyme.

Butelase-mediated cyclization, intermolecular ligation, or synthesis of protein thioesters

21| For butelase-mediated ligation, there are four options. Option A describes the protocol for intramolecular cyclization of peptides. Option B describes intramolecular cyclization of proteins. Option C describes the protocol for intermolecular ligation between two peptides. Option D describes the protocol for the preparation of peptide and protein thioesters.

(A) Intramolecular cyclization of peptides • TIMING 1 d

CRITICAL The following protocols are for a laboratory scale reaction, which can be scaled up depending on experimental needs. The general peptide/protein sequence for a butelase-mediated cyclization can be defined as $X_1X_2\dots X_n\text{Asx-His-Val}$ (X_1 = all amino acids except Pro, X_2 = Ile/Leu/Val/Cys). Target peptides can be prepared by standard solid phase synthesis using Fmoc chemistry³⁴. Alternatively, they can be purchased from vendors inexpensively. We usually order our peptides from GL Biochem (Shanghai, China).

(i) *Preparation of stock solutions (20 min)*: Prepare 500 μM peptide substrates in water and store at -20 °C.

(ii) Prepare a 10- μ M solution of butelase 1 in buffer EB2 and store at 4 °C.

(iii) *Typical peptide cyclization procedure (15 min)*: Conduct the cyclization reaction in a 50- μ l reaction mixture, at an enzyme-to-peptide ratio of 1:500 (0.1 μ M butelase 1, 50 μ M peptide) in the cyclization buffer (20 mM phosphate, 1 mM EDTA, pH 6).

Table 1. Components for a typical peptide cyclization

Component	Amount (μ l)	Final concentration (μ M)
500 μ M target peptide	5	50
10 μ M butelase 1	0.5	0.1
Cyclization buffer	44.5	
Total	50	

CRITICAL STEP For cysteine-containing peptides, we usually add 10 mM DTT in the reaction buffer to prevent disulfide bond formation. Butelase 1 has been shown to cyclize the reduced and alkylated peptide at a faster rate than the oxidized peptide. Reducing reagents are not necessary for non-cysteine containing peptides.

(iv) Incubate the reaction mixture at 42°C for 5 to 10 min.

(v) Monitor the progress of the reaction by HPLC and MALDI-TOF MS (**Fig. 4**). We should observed a mass shift of 254 Da upon conversion of a target peptide from the linear to circular form.

? TROUBLESHOOTING

(vi) Once the cyclization is completed, isolate the cyclic product by RP-HPLC and lyophilize the sample for further characterization.

PAUSE POINT The reaction can be quenched by adding HCl to a final concentration of 50 mM.

(B) Intramolecular cyclization of proteins • TIMING 1 d

CRITICAL Target proteins can be prepared by recombinant expression in *E. coli* or a suitable expression system³³

(i) *Preparation of stock solutions (20 min)*: Prepare 250 to 500 μ M solutions of recombinant proteins in buffers that they are most stable. Store at -80 °C.

(ii) Prepare a 10- μ M solution of butelase 1 in buffer EB2 and store at 4 °C.

(iii) *Typical protein cyclization procedure (30 min)*: Conduct the cyclization in a 50- μ l reaction mixture, at an enzyme-to-protein ratio of 1:250 (0.1 μ M butelase 1, 25 μ M protein substrate) in the cyclization buffer (20 mM phosphate, 1 mM EDTA, pH 6).

Table 2. Components for a typical protein cyclization

Component	Amount (μ l)	Final concentration (μ M)
250 μ M target protein	5	25
10 μ M butelase 1	0.5	0.1
Cyclization buffer	44.5	
Total	50	

CRITICAL STEP Butelase 1 activity will reduce if salt concentration is higher than 200 mM. Ideally, the total salt concentration in the final reaction buffer should be less than 100 mM.

(iv) Incubate the reaction mixture at between 25 °C and 42 °C for 15 to 30 min.

(v) Monitor the reactions by SDS-PAGE and ESI-LC-MS (**Fig. 5**). The cyclizations generally complete within 15 min. The cyclized protein should migrate faster than the linear protein precursor. A mass shift of 254 Da should be observed upon cyclization.

? TROUBLESHOOTING

(vi) *Purification of the cyclized protein (1 d)*. Dialyze the protein solutions overnight at 4 °C against suitable buffers at pH 7.5 to 8.5 to deactivate butelase 1. Exposure to buffers with pH above 7.5 will irreversibly deactivate butelase 1. The dialyzed proteins can be used directly without further purification. If a higher purity is required, ion-exchange or hydrophobic interaction chromatography can be used to purify the target circular proteins.

(C) Intermolecular peptide ligation • TIMING 4 h

CRITICAL: Here we use KALVINHV and GIGGIR as representatives for peptides 1 and 2, respectively, to demonstrate the intermolecular ligation. The general sequence for peptide 1 which contains the Asx-His-Val recognition sequence can be defined as $X_1X_2\dots X_n$ Asx-His-Val (X_2 should not be Ile/Leu/Val/Cys). Requirement beyond the Asx-His-Val sequence appears to be minimal. The general sequence for peptide 2 (acceptor peptide) can be defined as $X_1X_2\dots X_n$ (X_2 =Ile/Leu/Val/Cys). Requirement for acceptor peptide beyond X_2 residue appears to be minimal. The peptides 1 and 2 may contain unusual amino acids or labeling probes such as biotin and fluorophores depending on experimental needs.

(i) Prepare stock solutions of GIGGIR and KALVINHV at concentrations of 10 mM and 1 mM in water, respectively.

(ii) Prepare a 10- μ M solution of butelase 1 and store at 4 °C.

(iii) Conduct ligation reactions at the ratio of 1:500:10000 for butelase:KALVINHV:GIGGIR at the final concentrations of 0.1:50:1000 (μ M).

Table 3. Components for a typical intermolecular peptide ligation

Component	Amount (μ l)	Final concentration (μ M)
10 mM GIGGIR	5	1000
1 mM KALVINHV	2.5	50
10 μ M butelase 1	0.5	0.1
1X Ligation buffer	42	
Total	50	

CRITICAL STEP Make sure butelase 1 is the last component to add.

(iv) Incubate the reaction mixture at 42 °C for 20 min to 2 h.

(v) Monitor the ligation by injecting 25 μ l of the reaction solution at different time points into HPLC. The reaction can be quenched by adding HCl to a final concentration of 50 mM. Calculate the ligation yield based on the area under the curve in the HPLC profile.

? TROUBLESHOOTING

(vi) Purify the ligated product KALVINGIGGIR by RP-HPLC and lyophilize the sample for further characterization

(D) Butelase-mediated synthesis of protein thioesters • TIMING 3 d

(i) *Preparation of Boc-Gly-MBT (4 h)*. Add 20 ml of DCM into a 100-ml round-bottomed flask equipped with a stirring bar. Add the reagents in order: 88.57 mg (0.5 mmol) of Boc-Gly-OH, 312 mg (0.6 mmol) of PyBOP, 53.6 mg (0.5 mmol) of 3-methyl-1-butanethiol (MBT). Add 127.5 μ l (0.75 mmol) of DIEA at the final step.

(ii) Stir the reaction for 3 h at room temperature.

(iii) Monitor the conversion of Boc-Gly-OH to Boc-Gly-MBT by TLC using hexane:EA (3:1) mixture as the eluent. The product is visualized by ninhydrin solution after heating. We should observe a quantitative conversion indicated by the disappearance of the Boc-Gly-OH band and the appearance of a new band which is Boc-Gly-MBT.

? TROUBLESHOOTING

(iv) Remove the solvent by rotary evaporator at 35 °C.

(v) *Purification of Boc-Gly-MBT by chromatography (2 h)*. Prepare a silica gel slurry by adding 100 ml of a 5:1 mixture of hexane:ethyl acetate to 70 g of silica gel. The column used should be 2.5 to 3 cm inner diameter and should be packed until the compressed silica gel layer reaches 25 to 30 cm in height.

(vi) Dissolve the crude product in a minimal amount of DCM. Apply the product to the top of the column and cover the top with a thin layer of sand.

(vii) Elute with a mixture of hexane and ethyl acetate. Collect fractions of 40 ml and analyze with TLC. The desired product could be visualized in ninhydrin solution and should have a R_f of 0.8 in a mixture of hexane:ethyl acetate (3:1).

(viii) Combine all the fractions containing products and dry them by rotary evaporator at 35 °C. The product appears as yellow sticky oil and should be confirmed by NMR.

PAUSE POINT The Boc-Gly-MBT can be stored at 4 °C for at least 3 weeks.

(ix) *Preparation of H-Gly-MBT (2 h)*. Add 130 mg of Boc-Gly-MBT to a 50-ml round-bottomed flask with a magnetic stir bar in it. To this flask, add a 5-ml mixture of TFA:DCM (3:7).

(x) Stir the solution at room temperature for 45 min. After the reaction, remove DCM and TFA by rotary evaporation.

(xi) Dissolve the residual in a minimum amount of water and lyophilize the sample overnight. The lyophilized product is obtained in the form of transparent sticky oil and can be used in the following reactions without any purification.

PAUSE POINT The product could be stored in -20 °C for 6 months.

(xii) *Typical procedure for protein thioester preparation (2 h)*. Prepare 250 μ M stock solutions of recombinant proteins containing the C-terminal recognition sequence Asn-His-Val in buffers that they are most stable.

(xiii) Prepare a solution of 50 mM glycine thioester (H-Gly-MBT) in water and store at -20 °C.

(xiv) In a clean PCR tube, add 10 μ l of protein stock, 1 μ l of glycine thioester, 38.5 μ l of ligation buffer and 0.5 μ l of 10 μ M butelase 1 solution. The final concentrations of protein, glycine thioester, and butelase 1 are 50 μ M, 1 mM, and 0.1 μ M, respectively.

Table 4. Components for a typical protein thioester preparation

Component	Amount (μ l)	Final concentration (μ M)
250 μ M target protein	10	50

50 mM glycine thioester	1	1000
10 μ M butelase 1	0.5	0.1
1X Ligation buffer	38.5	
Total	50	

(xv) Thoroughly mix and incubate the reaction at 42 °C for 30 min to 2 h.

(xvi) Monitor the reaction progress by HPLC and identify the peaks by MS. Reaction is normally completed within one hour.

? TROUBLESHOOTING

(xvii) *Purification of the protein thioester (2 h)*. Dialyze the protein solutions overnight against 50 mM Tris buffer, pH 7.5 to 8.5 to remove excess glycine thioester and to deactivate butelase 1.

(xviii) Ion exchange or hydrophobic interaction chromatography can be used to separate or partially separate the desired protein thioester from unreacted and hydrolyzed proteins.

● TIMING

Steps 1-20, preparation of butelase 1: 3 d

Steps 21A i-ii, preparation of stock solutions: variable

Steps 21A iii-v, typical peptide cyclization procedure: 15 min

Step 21A vi, purification of the cyclized product: 1 h

Steps 21B i-ii, preparation of stock solutions: variable

Steps 21B iii-v, typical protein cyclization procedure: 30 min

Step 21B vi, purification of the cyclized protein product: 1 d

Steps 21C i-ii, preparation of stock solutions: variable

Steps 21C iii-iv, typical intermolecular peptide ligation procedure: 20 min – 2 h

Step 21C v, HPLC analysis: 30 min

Step 21C vi, purification of the ligated peptide: 1 h

Steps 21D i-viii, preparation of Boc-Gly-MBT: 8 h

Steps 21D ix-xi, preparation of H-Gly-MBT: 2 h and overnight lyophilization

Steps 21D xii-xv, typical procedure for protein thioester preparation: 1-2 h

Step 21D xvi, HPLC and MS analysis: 1 h

Step 21D xvii-xviii, purification of the protein thioester: 1d

ANTICIPATED RESULTS

Isolation and purification of butelase 1

The expected yield of butelase 1 is about 5 mg per kg of fresh plant materials. This amount of enzyme is sufficient for over 25,000 cyclization experiments described in this study using a 50- μ l reaction mixture and 0.1 μ M butelase 1. Purified butelase 1 should migrate as a single band about 37 kDa in SDS-PAGE gel (**Supplementary Fig. 5**).

Intramolecular cyclization

The cyclization reaction is generally completed within 10 min with >95% yield using the condition described. A mass shift of 254 Da should be observed upon cyclization (**Fig. 4**). If the cyclization yield is low, then the N-terminus may be hindered or masked because of the peptide conformation. Refolding may be required to expose both termini for efficient cyclization.

Intermolecular ligation of KALVINHV and GIGGIR

KALVINHV and GIGGIR have a m/z values of 892.2 and 572.3, respectively. You will observe a mass of 1209.5 for the ligated product KALVINGIGGIR, which elutes at 7.6 min in the RP-HPLC (**Fig. 6**). The intermolecular ligation reaction is completed within 15 min with about 90% yield.

Butelase-mediated synthesis of protein thioesters

The conversion from a peptide/protein containing the C-terminal Asn-His-Val recognition motif to the corresponding thioesters takes about 30 to 60 min. The conversion rate is highly dependent on the C-terminal accessibility of the target proteins. For peptides less than 50 residues, we usually achieve >95% yields in less than 30 min. For proteins, the average conversion yields are about 80-90% or up to 95% yields under optimized conditions (**Fig. 7**).

Synthesis of H-Gly-MBT

H-Gly- MBT.Transparent sticky oil.

TLC (Hexane:ethyl acetate 3:1 (vol/vol)) Rf = 0.8

¹H NMR (400 MHz, CDCl₃): δ in ppm 4.104 (s, 2H, NH₂-CH₂), 3.101-3.063 (t, 2H, S-CH₂), 1.729-1.662 (m, 1H, CH), 1.570-1.515 (m, 2H, CH₂), 0.973 (s, 3H, CH₃), 0.957 (s, 3H, CH₃)

¹³C NMR (400 MHz, CDCl₃): δ in ppm 192.804 (C=O), 46.839 (NH₂-CH₂), 38.045 (S-CH₂), 27.208 (CH₂), 26.542 (CH), 21.112 (CH₃)

Table 1 | Troubleshooting table.

Step	Problem	Possible reason	Solution
17	Low purity of butelase 1	Chromatographic columns may be overloaded	Reduce the amount of sample loaded onto the column
18	Low yield of butelase 1	Leaking of sample during dialysis or centrifugal ultrafiltration	Ensure no leaking of sample
21A v	Slow cyclization	The sequence of the target peptide is not optimal for butelase 1 The target peptide may have conformation unfavorable for cyclization	Increase the enzyme-to- peptide ratio from 1:500 to 1:100 Refold the peptide before proceeding to cyclization
21B v	No or slow cyclization	The N- and C-termini of proteins may be buried inside the protein core Some proteins are not stable at pH 6 High salt concentration in the protein buffer	Introduce a short linker sequence to the protein termini. Length of the linker may vary depending on the distance between the N- and C-termini of the proteins Raise pH of the cyclization buffer to 6.5-7.0. Reduce the salt concentration in the final buffer
21C v	Low ligation efficiency Hydrolysis of the Asx bond of peptide 1	When concentration of the acceptor peptide 2 is not high, water may act as a competing nucleophile	Increase the amount of the peptide 2 of up to 30 times in excess
21D iii	Inefficient coupling	MBT is oxidized	Use new bottle of MBT and store the reagent in dark bottle at 4 °C
21D xvi	Low yield of the protein thioesters	The C-termini of proteins may be buried inside the protein core	Add a short linker sequence to the C-terminus of the proteins

Some proteins are not stable at 42 °C

Lower the reaction temperature to 25-37 °C

FIGURE LEGENDS

Figure 1 | Schematic concepts and proposed applications of butelase 1 in protein engineering and biological research. (a) A molecular tag can be chemoselectively ligated to a target protein with a C-terminal NHV motif. (b) Cyclization of therapeutic peptides or proteins carrying a C-terminal NHV and a suitable acceptor residue at the N-terminus. (c) Production of edible biologics in plants by transgenic expression of both butelase 1 and a grafted therapeutic peptide into a cyclotide scaffold. (d) Generation of protein-protein fusion. (e) Immobilization of target peptides and proteins in to a solids support such as polystyrene beads, hydrogel or biosensor chip. (f) Tandem ligations using combination of ligases with different specificities such as butelase 1 and sortase A.

Figure 2 | Schematic presentation of butelase-mediated ligation. (a) For intramolecular cyclization, the substrate contains both of the Asn-His-Val recognition motif at the C-terminus and a suitable X₂ residue (I/L/V/C) at the N-terminus. (b) For intermolecular ligation, peptide 1 contains the C-terminal recognition sequence Asn-His-Val and peptide 2 contains a suitable X₂ residue (I/L/V/C).

Figure 3 | Protein thioester preparation using butelase-mediated ligation. The thioacyl intermediate is resolved by an amino group of the Gly-MBT to form the desired protein thioester. It should be noted that butelase 1 does not accept a thiol group for the incoming nucleophile and highly favor an amino group.

Figure 4 | Butelase-mediated cyclization of human hormonal peptides. (a, b) MS and HPLC profiles of neuromedin cyclization. (c, d) MS and HPLC profiles of galanin cyclization. The cyclization reactions were performed at 42 °C for 5 min in the presence of 0.1 μM butelase 1 and 50 μM peptide substrates. This figure is adapted from reference 18.

Figure 5 | An example of butelase-mediated cyclization of green fluorescent protein. The reaction was performed at 42 °C in the presence of 0.1 μM butelase 1 and 25 μM protein substrate. Analysis by ESI-MS and SDS-PAGE showed that the cyclization yield reaches over 90% in 15 min. The SDS gel was visualized by silver staining. This figure is adapted from reference 18.

Figure 6 | Butelase mediated intermolecular ligation of KALVINHV and GIGGIR. The reaction was performed in the presence of 0.1 μM butelase 1, 50 μM KALVINHV, and 1 mM GIGGIR, at 42 °C for 15 min. KALVINHV and GIGGIR have retention times of 5.9 and 5.2 min, respectively. The ligation product KALVINGIGGIR elutes at 7.6 min.

Figure 7 | An example of butelase-mediated thioesterification of ubiquitin. Observed molecular masses for the starting material (Compound 1) and product (Compound 2) are 9982.5 Da (calc. 9982.3 Da) and 9066 Da (calc. 9066.2 Da), respectively. Proteins are separated by RP-HPLC using an analytical C4 column (250 x 4.6 mm, Phenomenex) with a linear gradient of 0 to 60% acetonitrile in 30 min. This figure is adapted from reference 19.

Supplementary data

Supplementary Figure 1: Chromatogram of the *C. ternatea* extract by preparative anion exchange. Fractions containing ligase activity are indicated by the double arrow. The right axis indicate the percentage of the elution buffer. The cyan line indicates the gradient used for HPLC separation.

Supplementary Figure 2: A typical MS-guided screening of ligase-containing fractions. The linear precursor of SFTI has a m/z value of 1766.9 which is converted to 1512.7 in the presence of butelase 1.

Supplementary Figure 3: Chromatogram of the *C. ternatea* extract by size exclusion chromatography. Fractions containing ligase activity are indicated by the double arrow.

Supplementary Figure 4: Chromatogram of butelase 1 separation by analytical anion exchange. Fractions containing ligase activity are indicated by the double arrow. The cyan line indicates the gradient used for HPLC separation.

Supplementary Figure 5: SDS-PAGE analysis of purified butelase 1. The gel was stained with Coomassie blue. Purified butelase 1 should migrate as a single band with a molecular weight of 37 kDa.

ACKNOWLEDGEMENTS

This work was supported by the NTU iFood Research Grant M4081467.080 and Singapore National Research Foundation NRF-CRP8-2011-05.

AUTHOR CONTRIBUTIONS

G.K.T.N. conceived of the idea and drafted the manuscript. Y.Q. and X.H. developed the protocol for intermolecular ligation. Y.C. and C.F.L. developed the protocol for butelase-mediated protein thioester synthesis. J.P.T. supervised the project and revised the manuscript.

COMPETING FINANCIAL INTERESTS

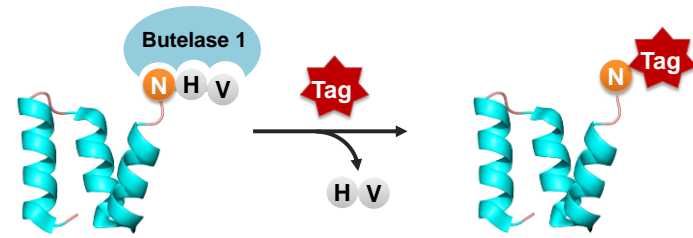
The authors declare no competing financial interests.

1. Mao, H., Hart, S.A., Schink, A. & Pollok, B.A. Sortase-mediated protein ligation: a new method for protein engineering. *J. Am. Chem. Soc.* **126**, 2670-2671 (2004).
2. Chang, T.K., Jackson, D.Y., Burnier, J.P. & Wells, J.A. Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12544-12548 (1994).
3. Wood, D.W. & Camarero, J.A. Intein applications: from protein purification and labeling to metabolic control methods. *J. Biol. Chem.* **289**, 14512-14519 (2014).
4. Rashidian, M., Dozier, J.K. & Distefano, M.D. Enzymatic labeling of proteins: techniques and approaches. *Bioconjug. Chem.* **24**, 1277-1294 (2013).
5. Barber, C.J. et al. The two-step biosynthesis of cyclic peptides from linear precursors in a member of the plant family Caryophyllaceae involves cyclization by a serine protease-like enzyme. *J. Biol. Chem.* **288**, 12500-12510 (2013).
6. Lee, J., McIntosh, J., Hathaway, B.J. & Schmidt, E.W. Using Marine Natural Products to Discover a Protease that Catalyzes Peptide Macrocyclization of Diverse Substrates. *J. Am. Chem. Soc.* **131**, 2122-2124 (2009).
7. Luo, H. et al. Peptide macrocyclization catalyzed by a prolyl oligopeptidase involved in alpha-amanitin biosynthesis. *Chem. Biol.* **21**, 1610-1617 (2014).
8. Mazmanian, S.K., Liu, G., Ton-That, H. & Schneewind, O. Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285**, 760-763 (1999).
9. Nguyen, G.K. et al. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. Chem. Biol.* **10**, 732-738 (2014).
10. Serra, A. et al. A high-throughput peptidomic strategy to decipher the molecular diversity of cyclic cysteine-rich peptides. *Scientific reports* **6**, 23005 (2016).
11. Craik, D.J., Daly, N.L., Bond, T. & Waite, C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol.* **294**, 1327-1336 (1999).

12. Nguyen, G.K. et al. Discovery and Characterization of Novel Cyclotides Originated from Chimeric Precursors Consisting of Albumin-1 Chain a and Cyclotide Domains in the Fabaceae Family. *J. Biol. Chem.* **286**, 24275-24287 (2011).
13. Nguyen, G.K., Lim, W.H., Nguyen, P.Q. & Tam, J.P. Novel Cyclotides and Uncyclotides with Highly Shortened Precursors from *Chassalia chartacea* and Effects of Methionine Oxidation on Bioactivities. *J. Biol. Chem.* **287**, 17598-17607 (2012).
14. Nguyen, K.N. et al. Immunostimulating and Gram-negative-specific Antibacterial Cyclotides from the Butterfly Pea *Clitoria ternatea*. *FEBS J.* (2016).
15. Gran, L. On the effect of a polypeptide isolated from "Kalata-Kalata" (*Oldenlandia affinis* DC) on the oestrogen dominated uterus. *Acta Pharmacol. Toxicol. (Copenh.)* **33**, 400-408 (1973).
16. Saether, O. et al. Elucidation of the primary and three-dimensional structure of the uterotonic polypeptide kalata B1. *Biochemistry (Mosc.)* **34**, 4147-4158 (1995).
17. Tam, J.P. & Lu, Y.A. Synthesis of large cyclic cystine-knot peptide by orthogonal coupling strategy using unprotected peptide precursor. *Tetrahedron Lett.* **38**, 5599-5602 (1997).
18. Nguyen, G.K. et al. Butelase 1: A Versatile Ligase for Peptide and Protein Macrocyclization. *J. Am. Chem. Soc.* **137**, 15398-15401 (2015).
19. Cao, Y., Nguyen, G.K., Tam, J.P. & Liu, C.F. Butelase-mediated synthesis of protein thioesters and its application for tandem chemoenzymatic ligation. *Chem. Commun. (Camb.)* **51**, 17289-17292 (2015).
20. Nguyen, G.K., Cao, Y., Wang, W., Liu, C.F. & Tam, J.P. Site-Specific N-Terminal Labeling of Peptides and Proteins using Butelase 1 and Thiodepsipeptide. *Angew. Chem. Int. Ed. Engl.* **54**, 15694-15698 (2015).
21. Harris, K.S. et al. Efficient backbone cyclization of linear peptides by a recombinant asparaginyl endopeptidase. *Nat. Commun.* **6**, 10199 (2015).
22. Bolscher, J.G. et al. Sortase A as a tool for high-yield histatin cyclization. *FASEB J.* **25**, 2650-2658 (2011).
23. Dawson, P.E., Muir, T.W., Clark-Lewis, I. & Kent, S.B. Synthesis of proteins by native chemical ligation. *Science* **266**, 776-779 (1994).
24. Liu, C.-F. & Tam, J.P. Peptide segment ligation strategy without use of protecting groups. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6584-6588 (1994).
25. Kolb, H.C., Finn, M.G. & Sharpless, K.B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem. Int. Ed.* **40**, 2004-2021 (2001).
26. Tam, J.P., Lu, Y.A., Liu, C.F. & Shao, J. Peptide synthesis using unprotected peptides through orthogonal coupling methods. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 12485-12489 (1995).
27. Wong, C.T. et al. Orally active peptidic bradykinin B1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment. *Angew. Chem. Int. Ed. Engl.* **51**, 5620-5624 (2012).
28. Clark, R.J., Akcan, M., Kaas, Q., Daly, N.L. & Craik, D.J. Cyclization of conotoxins to improve their biopharmaceutical properties. *Toxicon* **59**, 446-455 (2012).
29. Driggers, E.M., Hale, S.P., Lee, J. & Terrett, N.K. The exploration of macrocycles for drug discovery--an underexploited structural class. *Nat. Rev. Drug Discov.* **7**, 608-624 (2008).
30. Tam, J.P. & Wong, C.T.T. Chemical synthesis of circular proteins. *J. Biol. Chem.* **287**, 27020-27025 (2012).
31. Xu, M.Q. & Evans, T.C. Intein-mediated ligation and cyclization of expressed proteins. *Methods* **24**, 257-277 (2001).
32. Muir, T.W., Sondhi, D. & Cole, P.A. Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 6705-6710 (1998).
33. Baneyx, F. Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* **10**, 411-421 (1999).
34. Wellings, D.A. & Atherton, E. in *Methods in Enzymology*, Vol. Volume 289 44-67 (Academic Press, 1997).

Protein immobilization to solid support

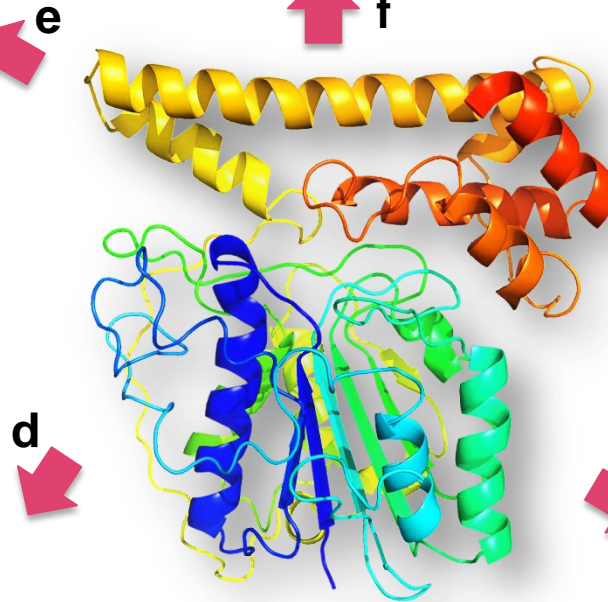
Tandem enzymatic ligation



Chemoselective labeling of proteins

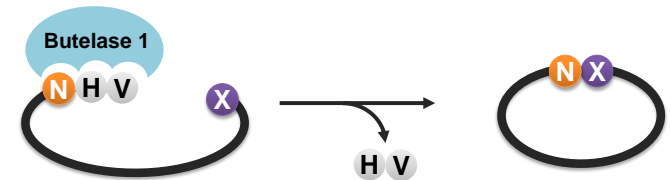
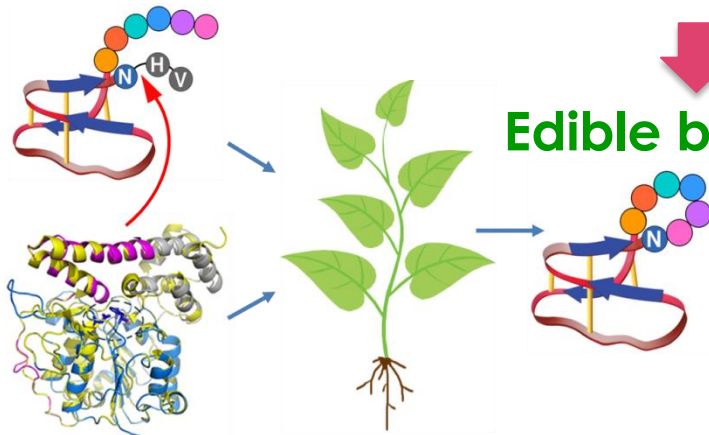
Protein-protein fusion

Cyclization of therapeutic proteins

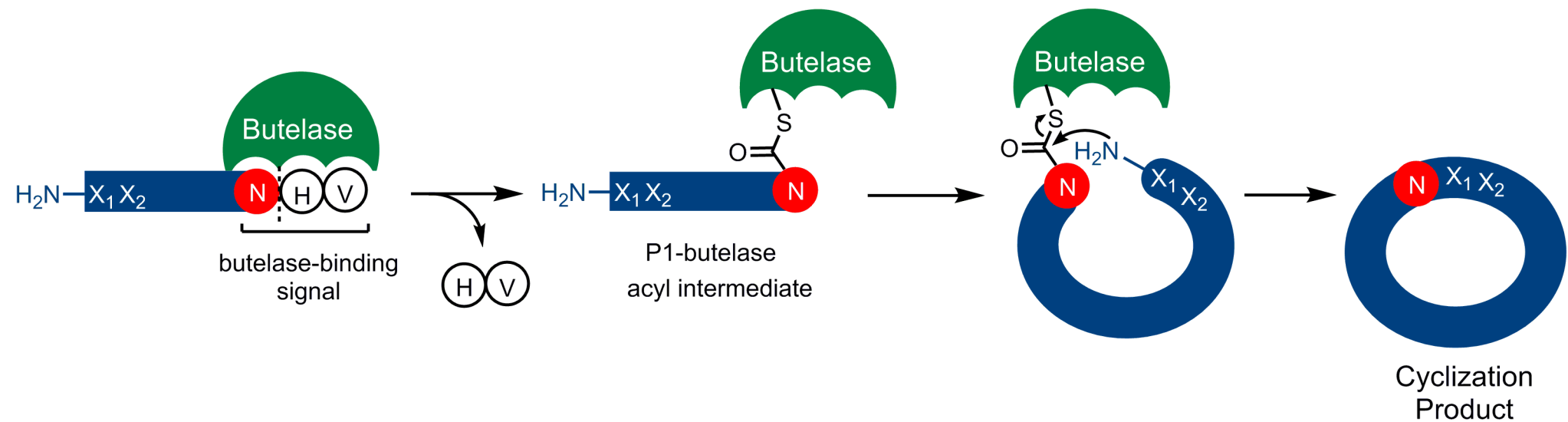


Butelase 1

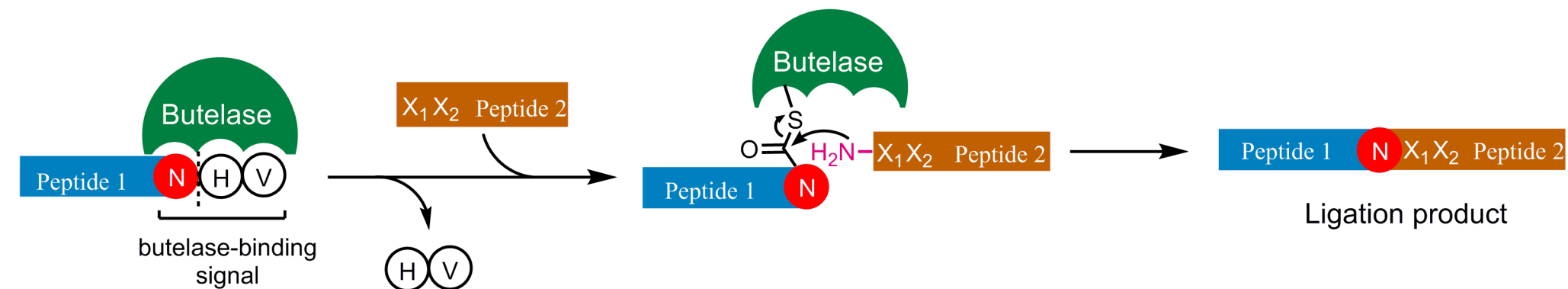
Edible biologics



a Intramolecular ligation

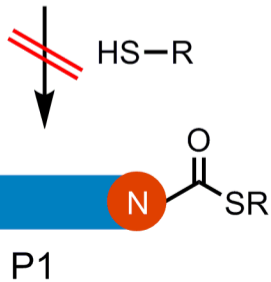
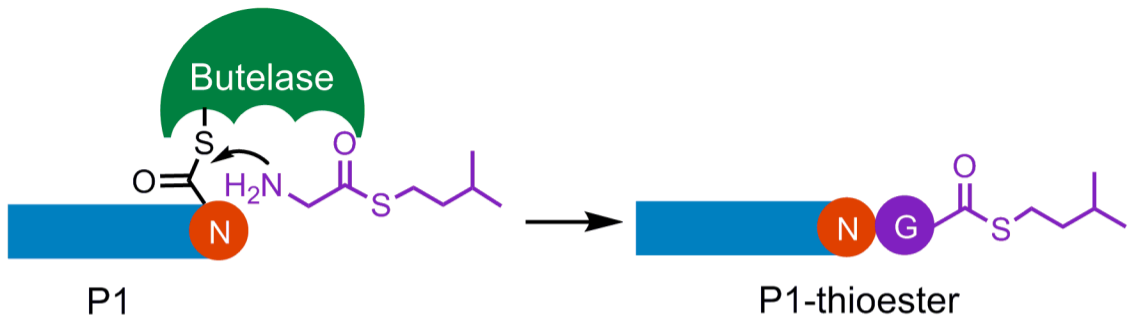


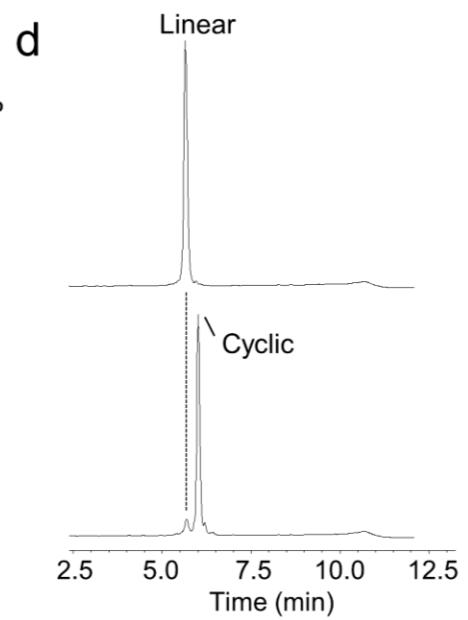
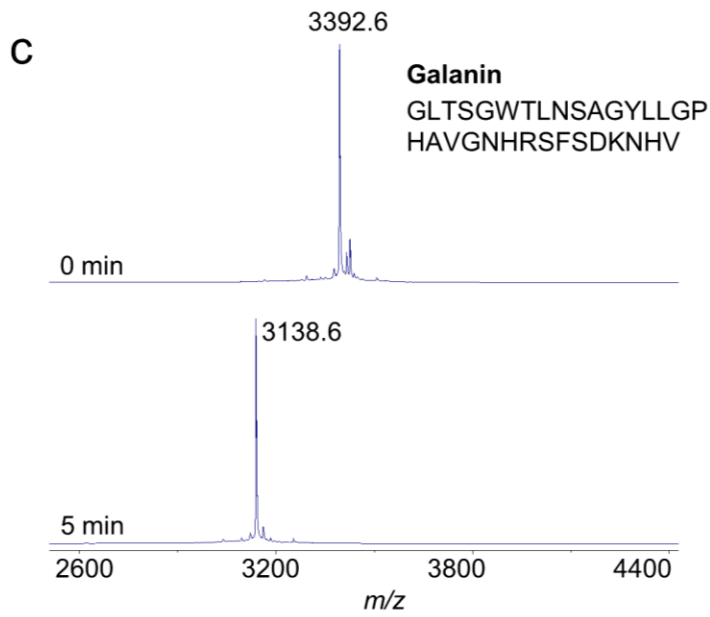
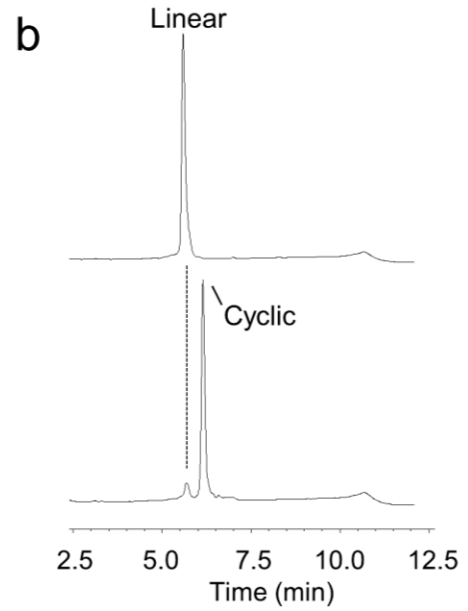
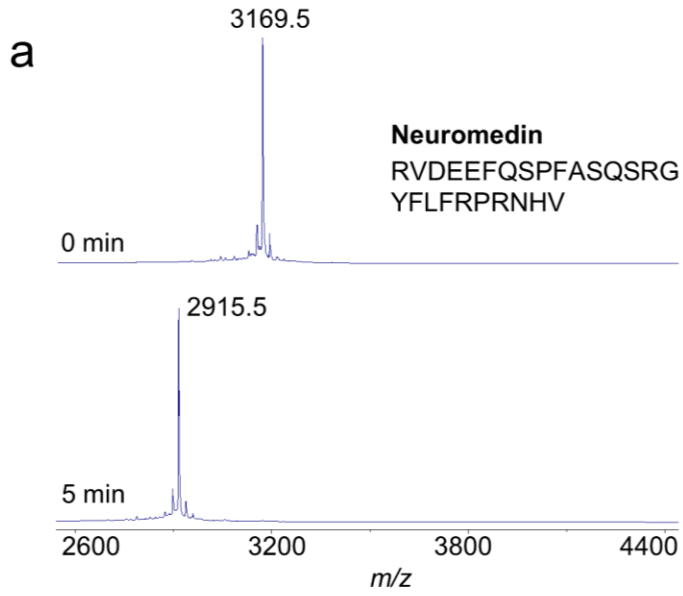
b Intermolecular ligation

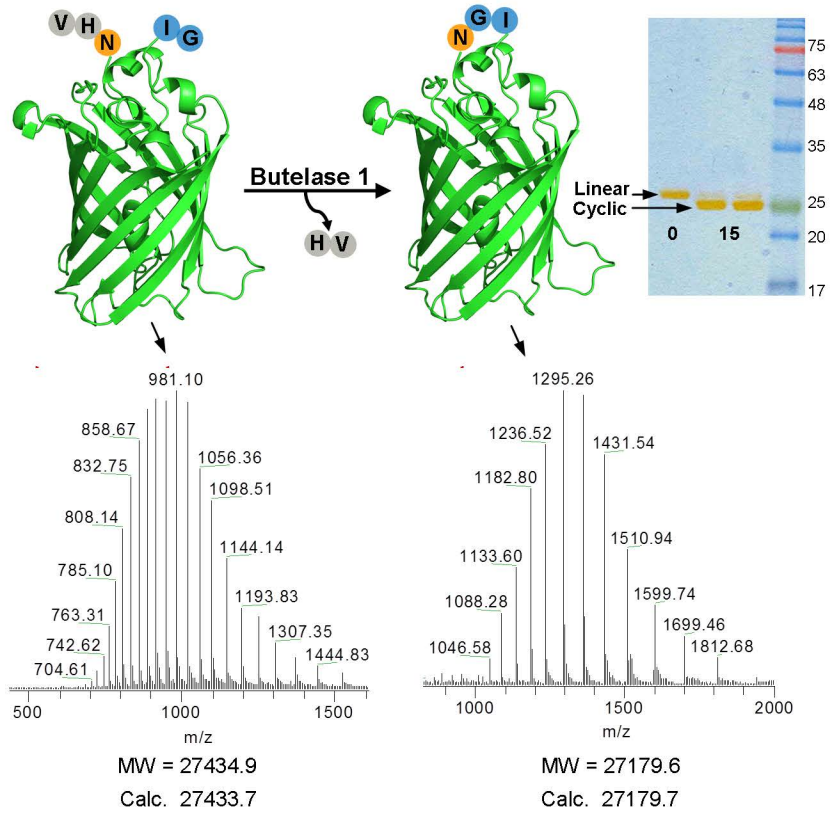


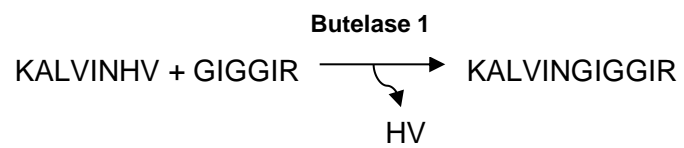
X_1 = any amino acid except P

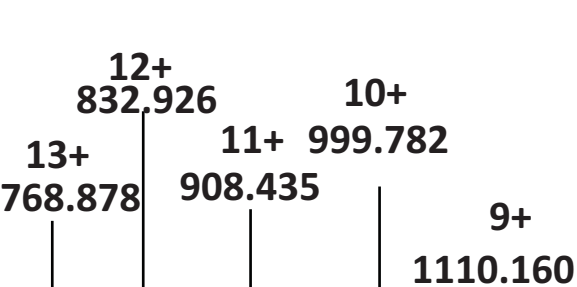
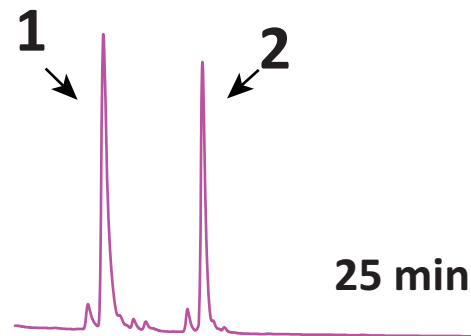
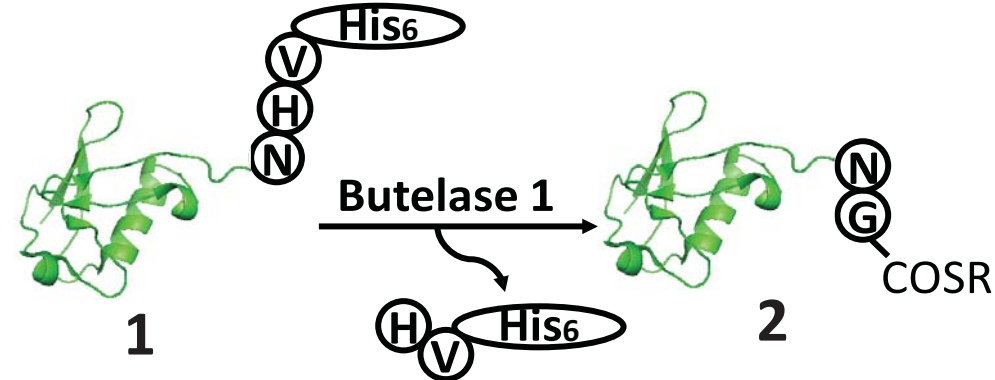
X_2 = L / V / I / C



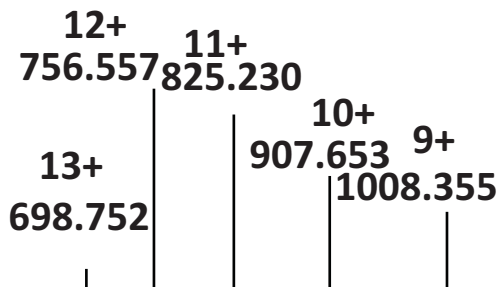








MW = 9982.5



MW = 9066

