이학석사 학위논문

Characterization of a topologically distinct microviridin-like modified peptide with multiple bicyclic core motifs

새로운 구조의 이중 고리를 반복적으로 갖는 microviridin 유사 변형 펩타이드의 특성 규명

2019년 2월

서울대학교 대학원

화학부 생화학 전공

노희진

Abstract

Characterization of a topologically distinct microviridin-like modified peptide with multiple bicyclic core motifs

Hee Jin Roh

Department of Chemistry

College of Natural Science

Seoul National University

Microviridins are ribosomally synthesized and post-translationally modified peptides (RiPPs) in which two ω-ester bonds and one ω-amide bond are formed between the side chains of the peptide residues. This type of the side-to-side cross-links can make cyclic peptides with diverse topology. However, the majority of currently found RiPPs modified with this microviridin-like modification are microviridins showing only a single consensus sequence and the unique tricyclic architecture. Here, I characterized new RiPPs called mTgnA and its homologs, revealing their cross-linking connectivity. They contain a novel consensus sequence

of the core peptide, TxxTxxxExxDxD, and the leader peptide, KPYxxxYxE. By the cognate ATP-grasp enzyme, the core motif is modified into a distinct hairpin-like bicyclic structure with two ω-ester bonds. The systematic employment of the ester-specific reactions and the tandem mass spectrometry was a powerful approach to determine the cross-linking connectivity of the modified core motifs in mTgnA and its homologs. This distinct connectivity makes mTgnA and its homologous RiPPs as a new microviridin-like RiPPs group, named the thuringinin group. The cross-linking reaction proceeded from the inner to outer rings, and was selective for acidic amino acids. Therefore, in this study I expanded the natural diversity of the microviridn-like RiPPs and proposed to define this new RiPPs family as omega-ester containing peptides (OEPs).

Keywords: RiPPs, microviridin, plesiocin, ATP-grasp, omega-ester containing peptides, ester cross-link, macrocyclization

Student Number: 2017-25338

Contents

1. Introduction

The presented work in this paper has been in press in ChemBioChem.[1]

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are one of the groups of natural products having large structural diversity.[2] RiPPs are grouped into different families based on their structural motifs and biosynthetic pathways.[3] Many RiPPs have pharmaceutical and biotechnological applications such as antibacterial activity (lanthipeptide), enzyme inhibitors (microviridin) or receptor antagonists (lasso peptide).[3-10]

RiPPs are initially synthesized as precursor peptides from ribosomes, and chemically modified by the cognate enzymes. The modification enzymes are often encoded by genes in the same gene cluster with those of the precursor peptides.[2, 3, 11] In the precursor peptide, there are two parts, a leader and a core peptide. The leader peptide is bound to enzyme for the substrate recognition and the core peptide is modified by the enzyme.[3, 12, 13] After the proteolytic removal of the leader peptide, the RiPP is transformed into the mature form.[3, 14, 15] Discovering new RiPPs is important because it expands the chemical diversity of peptide modification. Furthermore, the leader-guided biosynthetic pathway is a simple and powerful method for generating peptides with diverse

structures, suggesting a great potential for finding biologically active molecules.[12, 16] By using various modification enzymes in many combinations, we may construct peptide libraries at low genetic cost.[3, 17]

The microviridin family of RiPPs presents a distinct macrocyclic structure with intramolecular ω-ester or ω-amide bonds.[6-8, 14, 15, 18-20] These lactones or lactams are formed by condensation of the carboxyl side chain of glutamate or aspartate with the hydroxyl group of threonine or serine, or with the ε-amino group of lysine, respectively. An ATP-grasp enzymes mediate this macrocyclization reaction by phosphorylating the carboxyl side chain with ATP and subsequently ligating it to the hydroxyl or amine side chain.[14, 15, 21, 22] After the proteolytic cleavage of the precursor peptide to remove the leader peptide and acetylation by a GNAT-type N-acetyltransferase, the mature form of microviridin is produced.[14, 15] In the previous study, more than 300 precursor peptides of microviridin were identified from 174 microviridin biosynthetic gene clusters by the comprehensive phylogenomic mining study.[23] The gene clusters were frequently found in cyanobacteria, bacteroidetes, and proteobacteria phyla. Although the precursor peptides were diverse in the number of core peptides and the presence of processing signals, all of them have highly conserved PFFARFL α-helix motif at the leader peptide and TxKxPSDx(E/D)(D/E) motifs at the core peptides.[23] The ATP-grasp ligases interact with the conserved motif of the leader peptide

to recognize their substrate and the cyclization occurs in the conserved residues of the core peptide.[21-25] Many of the precursor peptides also have double-glycine motif (LSxxELxxIxGG) between the leader and the core peptides which is often associated with an MdnE type of membrane transporter.[23] The microviridin is one of the rare families of RiPPs having multiple core peptides within a precursor peptide.[26, 27] The cross-linking reaction in multiple core peptides occurs in a distributive fashion and follows an unstrict N-to-C overall directionality.[26] However, there is a strict order in macrolactonizing each core peptide in which the larger macrolactone is formed before the smaller macrolactone.[22, 26] Microviridins often have inhibitory activity against various serine proteases including chymotrypsin, trypsin and elastase.[7, 8, 16, 18-20, 28]

Many RiPPs have the macrocycles which are formed in an end-to-end fashion (cross-links between the C-terminus and the N-terminus) or end-to-side fashion (cross-links between the N-/C-terminus and a side chain of residue). However, these types of cross-links show a limitation in the structural variation because the monocycles are different only in the number and the types of amino acids in the ring. In contrast, the microviridins contain side-to-side cross-links between side chains of peptide residues which may additionally introduce topologically diverse architectures by combining multiple macrocycles. However, all of the bioinformatically found microviridin precursor peptides contain a single conserved sequence.[23] In addition, all the biochemically

characterized microviridins have the core peptides with uniform tricyclic connectivity.[6–8, 18–20] Little is known about the topological diversity of the microviridin–like peptides that contain the side–to–side cross–links via intramolecular ester or amide linkages. Furthermore, MvdD, the microviridin biosynthetic enzyme that makes omega ester bonds, has a limited substrate scope.[22] The ring sizes of lactones that MvdD can produce are not flexible. This result suggests that it may be difficult to produce various microviridin variants using only microviridin biosynthetic enzymes. Nevertheless, there was a previous study in which microviridin variants libraries were constructed, using microviridin biosynthetic ATP–grasp ligases attached with the leader peptide.[16] However, constructing cyclic peptide libraries with high structural diversity was challenging because the efficiency of the cyclization was not the same for all the variants and there was a topological limitation that the microviridin biosynthetic enzyme could make. Therefore, to expand the structure of microviridin–like RiPPs, it is required to discover new enzymes that modify peptides in different topology and to understand their reaction properties.

According to the development of bioinformatics, finding RiPPs with novel structure has become possible through genome mining.[2, 29] Using these data, our group previously discovered plesiocin which is a microviridin–like RiPP whose cross–linking connectivity is different from that of microviridin.[30] Plesiocin has four TTxxxxEE motifs which form multiple hairpin–like bicyclic

structure by a single ATP-grasp enzyme.[30] Each of the repeat in the core peptide has two omega-ester bonds which are formed within the highly conserved residues (Thr and Glu). The full cross-linking connectivity of plesiocin was determined by using ester-specific reactions and tandem mass spectrometry (MS/MS). Plesiocin is the first example that presents a totally different core motif and ring topology with the same type of macrocycles as those of microviridins. In addition, plesiocin has been the only microviridin-like RiPP in which the full connectivity of multiple core repeats was determined.[26]

Here, following the microviridin and plesiocin groups, I present the third distinct group of the microviridin-like RiPPs with multiple core repeats, termed thuringinin group.[1] I characterized this new microviridin-like RiPPs group by determining their cross-linking connectivity and their biosynthetic properties.

2. Results and Discussion

2.1. A new RiPP with an unprecedented conserved sequence pattern was found.

Bioinformatically found new gene cluster from *Bacillus thuringiensis serovar huazhongensis* BGSC 4BD1 consists of genes encoding a putative ATP grasp ligase (*tgnB*), a transmembrane protein (*tgnC*) and a putative RiPP precursor peptide (*tgnA*) (Figure 1a).[29] By doing BLAST, additional fifteen gene clusters that encode homologous protein of TgnB were discovered from various bacteria strain both in Firmicutes and Proteobacteria (Figure 2). All of them include a gene encoding the putative precursor peptide near that of the ATP grasp enzymes. Those putative precursor peptides are composed of the N-terminal leader sequence and the C-terminal core sequence which shows 2 to 7 repeats of a consensus sequence, TxxTxxxExxDxD.[31] This highly conserved sequence pattern in every repeat of the core peptides is distinguished from that of microviridins or plesiocins (Figure 1b). In particular, TgnA has three core repeats (Figure 1c).

To verify that TgnB can modify TgnA, I initially tried to express TgnB and TgnA separately by screening multiple protein over-expression conditions, changing the culture temperature, the induction time, IPTG concentration and the type of competent cells. However, the protein expression level of TgnB was too low in

every condition I tested. Even the expressed TgnB had very low solubility (~4 μM) in various buffer condition, seriously hampering the *in vitro* reconstitution of the cross-linking reaction. In addition, in the process of preparing TgnA, TgnA was susceptible to degradation and most of it was cleaved during purification steps. The problems of over-expressing TgnA and TgnB have not been solved by tagging MBP with the proteins. However, when TgnA was co-expressed with TgnB, the modified TgnA (mTgnA) was directly expressed and it was much less degradable than the linear TgnA. This might because the cross-linking modification increases metabolic stability of the peptide.[3] Furthermore, when TgnB and TgnA were co-expressed at 23° C, the expression level of TgnB tremendously increased and the solubility of TgnA and TgnB mixture was very high (>100 μM; Figure 3). This might because the hydrophobic surface of TgnB is covered by the binding of TgnA. I could regulate the expression ratio of TgnB and TgnA by changing the culture temperature. At 37° C, mTgnA was mainly expressed compared to TgnB (Figure 3). After incubating the mixture of TgnA and TgnB with ATP (5 mM) and $MgCl_2$ (10 mM) for twelve hours to complete the reaction, I analyzed the molecular weight (MW) of mTgnA by MALDI-TOF-MS. The MW of mTgnA was less by 108.13Da than that of TgnA demonstrating the dehydration of six water molecules in TgnA due to condensation reactions (Figure 1d). I hypothesized that two cyclizations accompanied by loss of two water molecules occur in each repeat of the core peptide.

2.2. Determining the full cross-linking connectivity of the modified TgnA.

To verify the hypothesis, each repeat of the core peptide was isolated by fragmenting mTgnA with proteases. Chymotrypsin preferentially cleaves peptide bonds C-terminal to tryptophan, tyrosine and phenylalanine separating the leader peptide and each repeat of the core peptide (Figure 4a and 4b). Even though the third repeat could be obtained by the chymotrypsin proteolysis, its MALDI-TOF-MSMS data was difficult to interpret in the later process. Instead, I used GluC that selectively cleaves peptide bonds C-terminal to glutamic acid and another fragment of the third repeat core peptide (**R3**) was obtained to be used for further analysis (Figure 4a and 4c). The MW of the first (**R1**) or the third repeat (**R3**) of the core peptide showed the Ile32-Tyr62 peptide or the Thr89-Ser106 peptide missing two water molecules respectively (Figure 4a and 4d). In hydrolysis condition, both **R1** and **R3** returned to the linear peptide (Figure 5a, 5b and 5e). Thus, there might be two ester bonds in **R1** and **R3** respectively. However, unlike **R1** and **R3**, the second repeat (**R2**) of the core peptide has an additional chymotrypsin cleavage site inside the Phe63-Tyr84 peptide (Figure 4a and 4d). MW of **R2** shows one water molecule loss and in the hydrolysis condition, **R2** was separated into two small peptides, Phe63-Phe73 (**R2-1**) and

Thr74-Tyr84 (**R2-2**) (Figure 5a, 5c and 5d). This is because **R2** has a structure that two peptide fragments, **R2-1** and **R2-2** are cross-linked by two ester bonds (Figure 4d). These results suggest that there is no inter-motif cross-link and that each core motif in mTgnA forms two ester cross-links. Next, I determined the cross-linking connectivity of the ester bonds for each repeat.

Structures of new macrocyclized peptides can be usually determined by the extensive 2D NMR or X-ray crystallography. However, these methods require a lot of efforts and time. Our group has previously suggested a new strategy to determine the cross-linking connectivity of ester bonds in modified peptides using MALDI-TOF-MS/MS analysis.[30] Although we cannot know the complete 3D-structure of the modified peptide, the proposed method provides accurate information about its ring topology. This method makes it possible to analyze many peptides in a fairly short time. In this study, I have improved the previously reported method to expand the peptides that can be analyzed.

Depending on the type of repeats, different strategy was used. Figure 6a shows a strategy I used to determine ester bonds connectivity of **R1** and **R3**. First, to identify the acidic amino acids that participated in esterification, two ester bonds were $NaOCH_3$-mediated methanolyzed. Only the acidic residues that made ester bonds transformed into methyl esters which were subsequently identified by MALDI-TOF-MS/MS. For **R1**, Glu55 and Asp58 were methylated in **pep1** (Figure 6a and 6b, upper panel). Second, **R1**

was partially methanolyzed and the two intermediates, **pep2** and **pep3** with a different methanolyzed ester bond were isolated by HPLC respectively. It has been previously reported that upon MALDI-TOF-MS/MS fragmentation, the ester bond between a threonine and an acidic residue transforms into the dehydrated threonine and the normal form of acidic residue.[30] Therefore, through the MALDI-TOF-MS/MS analysis of the methanolysis intermediates, I could determine the ester bonds making threonines and their connectivity. For example, **pep2** which is the peptide intermediate having inner ester bond, transformed into **pep2'** (' designates the rearranged linear peptide in the MS analyzer) after the MALDI-TOF-MS/MS analysis and it shows that Thr51 is dehydrated and Asp58 is methylated (Figure 6a and 6b, middle panel). As I know that Glu55 and Asp58 make ester bonds in **R1**, non-methylated Glu55 is cross-linked to Thr51 (Thr51-Glu55). With the same logic, as Thr48 is dehydrated in the MALDI-TOF-MS/MS spectrum of **pep3'**, the outer ester bond is formed between Thr48 and Asp58 (Thr48-Asp58; Figure 6b, lower panel). Therefore, combining the three MALDI-TOF-MS/MS spectra, the cross-linking connectivity of two ester bonds in **R1** is Thr48-D58 and T51-E55.

For **R2**, similar but different strategy was used. After the complete methanolysis, **R2** was dissembled into **R2-1** and **R2-2-2Me** (**pep4**). All ester bond making acids were methylated and they were Glu76 and Asp79 (Figure 6c and 6d, upper panel). Two

esterified threonines were determined by the direct MALDI-TOF-MSMS analysis of **R2** (Figure 6c). In the MS/MS spectrum of **pep5'** , the peak of the dehydrated Thr69 suggests that one ester bond is made with Thr69, and the observation of the y13-y15 ions indicates that Thr72 makes the second ester bond (Figure 6d, middle panel). As **R2-1** and **R2-2** are separated in the C-terminus of Phe73, in order for y13-y15 to be observed, Thr72 must be covalently connected to **R2-2**. To determine the cross-linking connectivity of **R2**, partial methanolysis was conducted and only one product, **pep6** was detected in HPLC (Figure 6c). In the MS/MS spectrum of **pep6**, y13-y15 ions were observed and Asp79 was methylated (Figure 6d, lower panel). This result indicates that Thr72 is connected to unmethylated Glu76. Moreover, y-series-ion (y2*) that came from **R2-1** showed dehydrated Thr72 which additionally supported that the inner cross-links is made with Thr72. Therefore, the two ester bonds in **R2** are Thr69-Asp79 and Thr72-Glu76. Unlike methanolysis, two types of intermediates were detected after a partial hydrolysis. In the MS/MS spectrum of the hydrolysis intermediate with Thr69-Asp79 ester bond, y13-15 peaks were not observed which was consistent with the above results (Figure 7).

The cross-linking connectivity of **R3** was determined using the same strategy as the one used for **R1** (Figure 8). Because **R3** contains an arginine, the intensity of ions that are fragmented at the C-terminus of unmethylated aspartic acid predominated due to the

aspartic acid effect.[32-34] For example, the low peak intensity of b12 in Figure 8a and 8b, compared to that in Figure 8c, supports that Asp100 in Figure 8a and 8b is methylated. As a result, **R3** has two ester bonds having the same cross-linking connectivity pattern as **R1** (Thr90-Asp100 and Thr93-Glu97).

In conclusion, all of the three conserved TxxTxxxExxD motifs in mTgnA core peptide have the same connectivity of the two ester cross-links, resulting in three repeats of a hairpin-like bicyclic structure. mTgnA is the second example of microviridin-like RiPP whose full cross-linking connectivity was determined. Although both microviridin and plesiocin have lactones in multiple repeats of the core peptide, their cross-linking patterns are different from that of mTgnA.[26, 30] The core peptide of TgnA contains 13 threonines and 17 acidic amino acids (Glu or Asp), some of which are even adjacent to each other. The number of possible products with different esterification connections is more than 1.5×10^{10}. However, only one major product, mTgnA has been detected, demonstrating that the modification reaction is highly selective.

I do not know the bioactivity of mTgnA because it did not inhibit several proteases that I tested (chymotrypsin, GluC, trypsin, and AspN) and there was no obvious toxic effect while expressing mTgnA. I also do not know the structure of the naturally matured mTgnA because in the *tgnA* gene cluster, there was no gene for protease that cleaves the leader peptide. For convenience, the mature form of mTgnA after the leader cleavage was named as

thuringinin.

2.3. mTgnA and its homologs have the same cross-linking connectivity.

As TgnA was susceptible to degradation and its multiple repeat system is too complicated, I constructed a TgnA mutant having a single repeat (TgnA-R1). TgnA-R1 is composed of the leader peptide and the first repeat of the core peptide, and it was expressed without the degradation problem. In the presence of ATP and Mg^{2+}, TgnA-R1 was modified by TgnB in the *in vitro* reaction (Figure 9). Using the same strategy as the one used for **R1**, I found that the modified TgnA-R1, mTgnA-R1 is double dehydrated with the same cross-linking connectivity as **R1** (Figure 10, 11a and 11b). To identify that its unique ester bonds connectivity is common in mTgnA homologs, I additionally analyzed two TgnA homologs in a single repeat form. CwnA-R1 is from *Colwellia* sp. UCD-KL2 and DbnA-R1 is from *Dehalobacter sp*. TeCB2 which is proteobacteria and firmicutes respectively (Figure 2). CwnA-R1 was modified by the cognate enzyme, CwnB ATP grasp ligase and cleaved by trypsin. Although there are two trypsin-cleavage sites (peptide bonds C-terminal to Lys20 and Arg27) in CwnA-R1, I optimized the proteolysis condition to obtain the major peptide fragment only with the cleavage after Lys20. DbnA-R1 was modified by the cognate ATP-grasp enzyme, DbnB and cleaved by chymotrypsin after the C-terminus of Leu43 and Phe62. Both of

the core peptides of mCwnA-R1 and mDbnA-R1 were double dehydrated and their two ester bonds connectivities were determined using the same strategy as the one used for R1 (Figure 11c-f). As a result, both mCwnA-R1 and mDbnA-R1 have the same cross-linking connectivity as mTgnA-R1 which suggests that this distinct cross-links pattern is a unique property of mTgnA homologs (Figure 10). Based on this cross-linking connectivity, I defined mTgnA and its homologs as the third group of microviridin-like RiPPs, named the thuringinin group. Their topology is similar to that of prochlorosin 3.3 which is a lanthipeptide whose two thioethers topology and the number of amino acid spacing between reactive sites (TxxCxxxTxxC) are the same as thuringinin.[35-37]

In the cross reactions in which one ATP-grasp enzyme was mixed with a non-cognate precursor peptide within the thuringinin group, the modification reaction proceeded in different levels (Figure 12). This different reaction efficiency may be due to the variation in the non-conserved sequence of the precursor peptides or the different linker length between the leader and the core peptides.

All amino acids that participate in forming the ester bonds (TxxTxxxExxD) are included in the highly conserved sequence (TxxTxxxExxDxD) of the thuringinin group (Figure 1b). At first, I thought that the last conserved aspartic acid might be also important to the reaction. However, when Asp60 in TgnA-R1 was mutated into alanine (TgnA-R1-D60A), this mutation did not

significantly affect the reaction rate (Figure 13). In addition, the ester bonds connectivity of mTgnA-R1-D60A was the same as that of mTgnA-R1 (Figure 10, 11g and 11h). This result suggests that all of the conserved amino acids are not necessarily critical to the modification reaction.

2.4. The inner ester bond is formed before the outer ester bond.

If there is a reaction order in the formation of two ester bonds in mTgnA-R1, only one type of intermediate having one ester bond would be detected. As the intermediate was observed with MALDI, the enzyme reaction is distributive like that of microviridin or plesiocin (Figure 13a).[26, 30] First, I conducted HPLC analysis of the partially hydrolyzed mTgnA-R1 core peptide to identify the retention times of the possible intermediates (Figure 14a). Two hydrolysis intermediates with the inner or outer ester bond were detected in different retention times. Each intermediate structure was determined by MALDI-TOF-MS/MS analysis either directly or after the methanolysis (Figure 15a and 15b). I found that in HPLC analysis, the intermediate with the inner ester bond was eluted at around 30.5 min and that with the outer ester bond was eluted at around 36.5 min. Next, the *in vitro* reactions of TgnA-R1 and TgnB were quenched at different time points by adding EDTA (Figure 9). After the removal of the leader peptide with chymotrypsin, the core

peptides were subsequently analyzed by HPLC (Figure 14b). While the amount of intermediate with the inner ester bond increased and decreased sequentially over the reaction time, the peptide with the outer ring was not noticeably detected. Because I know the exact retention times of the peptide intermediates, it is certain that the peptide with an outer ring does not appear during *in vitro* reaction, excluding the possibility that I have missed the peptide during the HPLC analysis. Therefore, in TgnA-R1 modification the inner ester bond is formed before the outer ester bond (Figure 14C). Similarly, there are specific reaction order in the formations of two lactones in microviridins and two thioethers in prochlorosin 3.3, in which the bigger ring or the smaller ring is formed first, respectively.[22, 38, 39] One possible reason for the specific reaction order is that the first lactone helps create the other lactone by changing its structure. However, for TgnA-R1, the formation of the inner ring is not prerequisite for the formation of the outer ring. When the residues making the ester bonds were respectively mutated to alanine, all of these mutants showed singly dehydrated products after the reaction with TgnB (Figure 15c). These products have either the inner or outer ring in mTgnA-R1. This is different from microviridin in which the second lactone ring cannot be formed without the first lactone ring.[22] Therefore, our model for reaction order is the result of different kinetics.

2.5. Cross-linking reaction is selective with acidic

amino acids.

In the sequence of CwnA-R1, there are both $T^{26}xxT^{29}xxxE^{33}xxD^{36}$ and $T^{23}xxT^{26}xxxE^{30}xxE^{33}$ patterns (Figure 16a). If the CwnB reaction does not distinguish the two patterns, two products might be formed. However, only the native mCwnA-R1 was observed in the reaction with CwnB and it can be said that there is some selectivity in the reaction (Figure 10 and 16a). To further investigate this reaction selectivity property, I constructed peptide mutants, CwnA-R1-E33D and CwnA-R1-D36E.

In the MALDI-TOF-MS/MS spectra of CwnA-R1 mutants, the intensity of the b-ions with a C-terminal aspartate predominated and the other ions were difficult to observe. This effect arises because the high basicity of the arginine side chain (Arg27) sequesters the ionizing proton and suppresses random cleavage of backbone bonds.[32] Therefore, to obtain more informative MS/MS spectra, I chemically modified the arginine to a less basic N5-(4,6-dimethyl-2-pyrimidinyl) ornithine (Pyo) (Figure 17).[40]

If CwnB do not have a preference in the acidic residue, CwnA-R1-E33D would be modified into two products having two ester bonds in $T^{26}xxT^{29}xxxD^{33}xxD^{36}$ or $T^{23}xxT^{26}xxxE^{30}xxD^{33}$ (Figure 16a). However, after the twelve hours reaction, only the singly dehydrated product was observed where the ester bond was formed between Thr26 and Glu30 which is the inner ester bond of $T^{23}xxT^{26}xxxE^{30}xxD^{33}$ (Figure 16a and 18). It seems that the glutamic acid that makes the inner ester bond cannot be replaced

with the aspartic acid because no cross-link was formed in $T^{26}xxT^{29}xxxD^{33}xxD^{36}$ sequence. Furthermore, although the $T^{23}xxT^{26}xxxE^{30}xxD^{33}$ has the same sequence pattern with the conserved residues of thuringinin group, the outer Thr23-Asp33 ring was not formed. It suggests that amino acids other than those participate in esterification also affect the reaction.

In Cwn-R1-D36E, the same sequence pattern is observed in both $T^{23}xxT^{26}xxxE^{30}xxE^{33}$ and $T^{26}xxT^{29}xxxE^{33}xxE^{36}$. After the reaction with CwnB, CwnA-R1-D36E showed double dehydrated product and the major and the minor products were obtained at 10 to 1 ratio. As a result of MS/MS analysis, I found that the major product has two ester bonds with the same connectivity with mCwnA-R1 (Figure 16a and 19). Although it was difficult to precisely determine the connectivity of the minor product due to its small amount, it is likely that there are two ester bonds inside $T^{26}RLTE^{30}TIE^{33}$ where T26, E30, and E33 participate in the esterification (Figure 16a and 19). Despite the same sequence pattern of $T^{23}xxT^{26}xxxE^{30}xxE^{33}$ and $T^{26}xxT^{29}xxxE^{33}xxE^{36}$, most of the products are selectively made at the native reactive sites, $T^{26}xxT^{29}xxxE^{33}xxE^{36}$ which indicates that the surrounding residues influence the reaction selectivity. In addition, the reaction rate of CwnA-R1-D36E decreased significantly compared to that of CwnA-R1 (Figure 16b). Although the outer ester bonds can be formed with glutamic acid, this is not an optimal condition for the reaction. Collectively, I found that the cross-linking reaction is

selective in both of the acidic residues that make the cross-links and the residues surrounding the reactive sites.

The reason why esterification did not proceed completely in both mutants is not because of the short linker length between the leader and the core peptide. When a flexible linker with three glycines was inserted between the leader and the core peptide of both CwnA-R1 and its mutants, it did not significantly change the reaction rate (Figure 20). Therefore, the different reaction efficiency of the mutant peptides is certainly due to the peptide sequence.

Relaxed selectivity to acidic residues in reactive sites was observed in lasso peptides that make a macrolactam by the condensation of the N-terminal amino group with a side-chain carboxylate of glutamate or aspartate.[41, 42] Not only the type of acidic residue but also the location of the acidic residue that makes the macrolactam is flexible. The predicted microviridin precursors also present both glutamate and aspartate in positions where the lactam and the smaller lactone ring are formed (TxKxPSDx(E/D)(D/E)), and, in particular, the formation of the lactam rings with either acidic residues was biochemically characterized.[3, 6-8, 14, 15, 18-20, 23] However, total 73 core motifs in the sixteen precursor peptides of the thuringinin group exclusively showed glutamate for the smaller ring and aspartate for the larger ring (Figures 1b and 2). The different connectivity and the topology between microviridins and thuringinins might lead to the contrasting selectivity of acidic residues for macrocyclization. The effect of

surrounding residues was also showed in microviridins as the efficiency of the cyclization was not the same for all the core peptide variants.[16, 22, 28]

2.6. Defining the minimal leader of TgnA

Microviridins and plesiocins have conserved sequence motifs, 'PFFARFL' or 'LFIx(D/E)L' in their leader peptides respectively. These motifs are critical to the cyclization reactions (Figure 21).[21, 24, 30] I investigated the sequence of TgnA leader peptide and defined the minimal leader of TgnA based on its key residues. First, the leader peptide of TgnA-R1 was divided into six parts (L1 to L6) (Figure 22a). In the sequence logo of the thuringinin group' s leader peptide, highly conserved residues, KPYxxxYxE motif, is shown in L1 and L2 (Figure 22b). Even many leader peptides in thuringinin group are much shorter than that of TgnA, not having L4-L6 sequences (Figure 2). To identify which sequences are critical to the reaction, I constructed TgnA-R1 variants in which each part of the leader peptide (L1-L6) is truncated (Figure 22c). When L1 or L2 are deleted, the reaction with TgnB did not proceed at all, suggesting that L1 and L2 are critical to the reaction. In contrast, the deletion of one or two parts of L3-L6 did not influence the reaction. However, when there are only L1 and L2 (△L3L4L5L6), single dehydrated product was major which suggested that a certain length of linker between L1L2 and the core peptide is required for the full reaction. Therefore, I

constructed the TgnA-R1 mutant with the minimal leader (△L3L4L5L6_[Thm]_core) by inserting thrombin cleavage site (GKLVPRG) between L1L2 and the core peptide. This minimal leader system is useful because the peptide is short enough to be engineered. Furthermore, the leader and the core peptide can be easily dissembled by the thrombin cleavage (Figure 23).

3. Conclusion

In this work, I identified the full cross-linking connectivity of mTgnA, a new RiPP having multiple repeats of the core peptide which are macrocylized by two ω-ester bonds respectively. mTgnA and its homologs have the highly conserved sequence, TxxTxxxExxDxD motif which makes unique cross-links pattern. I named these microviridin-like RiPPs as thuringinin group. I also propose to define a new RiPP family that includes microviridin and microviridin-like RiPPs. Most of the RiPPs families are defined based on their modification chemistry.[3] As all of the microviridins and microviridn-like RiPPs have ω-ester bonds, I suggest naming this new RiPPs family as "omega-ester containing peptides (OEPs)". Microviridin, plesiocin and thuringinin can be OEPs and each of their homologs can be grouped into microviridin group, plesiocin group and thuringinin group respectively. As this RiPPs family is characterized for direct cross-links between side chains of peptide, it does not include some thiopeptides such as nosiheptide and thiostrepton where ester bonds are formed by additional attachment of indolic acid or quinaldic acid to the macrocyclic core.[43, 44] The topology of OEPs may be easily identified by the strategy that our group suggested, using methanolysis reaction and MALDI-TOF-MS/MS spectrometry. Furthermore, according to the bioinformatic studies, there are likely

to be additional OEPs with novel topology that have not yet been revealed.[29] For example, one potential precursor peptides have multiple repeats of the core peptide with unique conserved sequence motif, TKTxxxExDD.[29] Discovering new OEPs expands the natural diversity of microviridin-like RiPPs. Furthermore, it suggests the usefulness of enzymes as tools to make topologically diverse cyclic peptides which may show various bioactivities such as drug candidates or enzyme inhibitors.

4. Materials and Methods

4.1 General

Reagents for cloning were purchased from Enzynomics or Toyobo. Oligonucleotides were purchased from Macrogen. All PCR reactions were performed using KOD-plus-Neo-DNA polymerase. All cells carrying a plasmid were grown on LB-agar plates or LB media containing antibiotic that matches to the plasmid's antibiotic resistance gene. The antibiotic concentrations used in this study are 50 μg/mL kanamycin, 100 μg/mL ampicillin, and 100 μg/mL streptomycin.

Proteases (trypsin, chymotrypsin, and Glu-C) and ATP were purchased from Sigma-Aldrich. All protein concentrations were determined by UV absorbance at 280 nm. MALDI-TOF-MS and MALDI-TOF-MS/MS analysis were performed by Bruker Microflex MALDI-TOF mass spectrometer or Bruker Ultraflextreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics). All MALDI samples were desalted by C18 Zip-Tip (Millipore). HPLC analysis and HPLC purification were performed by Agilent 1260 Infinity (Agilent) that was connected to either ZORBAX SB-C18 StableBond Analytical HPLC Column (4.6 mm x 250 mm, 5 μm particle size; flow rate, 1.0 mL/min; Agilent) or ZORBAX SB-C18 Semi-Preparative HPLC Column (9.4 mm x 250 mm, 5 μm particle size; flow rate, 4.0 mL/min; Agilent). 0.05%

TFA in H_2O (Solvent A), 0.05% TFA in Acetonitrile (Solvent B), and 0.05% TFA in CH_3OH (Solvent C) were used as mobile phases. Proteins were monitored at 214 nm, and the eluted fractions were analyzed by MALDI.

4.2 Cloning

Genes optimized for heterologous expression in Escherichia Coli were synthesized by General Biosystems (Table 1). Plasmids and primers used in this study are listed in Table 2 and 3, respectively. Genes encoding TgnA, CwnA, DbnA, and TgnB were inserted into pET28b or pET22b vector with NdeI and NotI sites at 5' – and 3' –ends of each gene to make pRHJ1, pRHJ3, pRHJ5, and pRHJ2, respectively (Table 1 and 2). Genes encoding CwnB and DbnB were inserted into pCDFDuet-1 vector with BamHI and BglII sites at 5' – and 3' –ends of each gene, to make pRHJ4 and pRHJ6, respectively (Table 1 and 2).

To construct a plasmid co-expressing TgnA and TgnB (pRHJ7), I used the restriction-free (RF) cloning method.[45] In the first PCR, sequence of TgnB was amplified from pRHJ2 using TgnA_TgnB coexpression_FW and TgnA_TgnB coexpression_RV as primers. The PCR product was digested with DpnI, purified, and used as primers for the second PCR with pRHJ1 as a template. The PCR product was digested with DpnI, purified, and used for transformation. Plasmids expressing variants of TgnA, CwnA, and DbnA (pRHJ8-25) were obtained by the inverse-PCR method

using pRHJ1, pRHJ3, or pRHJ5 as a template.[46] To obtain a plasmid expressing ΔL3L4L5L6-[Thm]-core with a N-terminal His_6-tag (pRHJ26), the thrombin cleavage site was inserted between the leader peptide and the core peptide by the inverse-PCR method with pRHJ25 as a template. I also inserted E. coli maltose-binding protein (MBP) and a TEV-protease cleavage site before the leader peptide by the RF cloning method using pSK387 as a template to make pRHJ27.[47]

To verify the plasmid sequence, plasmids were electroporated into E. coli DH10β, and cells were grown on an LB agar plate containing antibiotics at 37° C for 14 hours. Single colonies were inoculated in 5 mL of LB media containing antibiotics and further incubated at 37° C for 12 hours before pelleting. Plasmids were isolated using the LaboPass Plasmid Miniprep Kit (Cosmogenetech). Mutations were verified by DNA sequencing (Macrogen).

4.3 Over-expression and purification of mTgnA and TgnB

ER2566 cells carrying pRHJ7 were grown on an LB agar plate with kanamycin at 37° C for 14 hours. For mTgnA preparation, a single colony was inoculated into 80 mL LB containing kanamycin and grown at 37° C for 12 hours. This culture was diluted with 8 L of LB containing kanamycin and grown at 37° C. 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added at an OD600 of 0.4 and the cell culture was incubated for additional 6 hours at

37° C. For TgnB preparation, 20 mL LB containing kanamycin was inoculated with a single colony and grown at 37° C for 6 hours. This culture was diluted with 2 L of LB containing kanamycin and grown at 37° C. 0.1 mM IPTG was added at an OD600 of 0.4 and the cell culture was incubated for additional 12 hours at 23° C. The different incubation temperature after IPTG induction is important as it controls the expression ratio of mTgnA and TgnB. Cells were harvested and resuspended with Tris-wash buffer (100 mM Tris (pH 8.0), 20 mM imidazole, and 300 mM NaCl). The resuspended cells were lysed by sonication and centrifuged at 20,000 x g for 30 min. The supernatant was passed through Ni Sepharose 6 FastFlow beads (GE Healthcare), and after washing the beads with Tris-wash buffer, the bound proteins having His_6-tag were eluted by Tris-elution buffer (100 mM Tris (pH 8.0), 500 mM imidazole, and 100 mM NaCl). Using an Amicon centrifugal filter (Millipore), the eluted solution was concentrated and its buffer was exchanged to 10 mM Tris (pH 8.0).

In SDS-PAGE gel, the eluted proteins expressed at 23° C showed the bands for both mTgnA and TgnB, but those expressed at 37° C showed only the band corresponding to mTgnA (Figure 3). I used the former as a source of the TgnB enzyme in the *in vitro* reaction. I could not separate TgnB and mTgnA by ion-exchange chromatography nor gel-filtration chromatography, but the presence of mTgnA was important for high TgnB expression level and solubility. After gel-filtration chromatography (Superose 6; GE

Healthcare), the TgnB concentration was estimated assuming that the molar ratio of TgnB:TgnA is 1:1. To complete the modification, mTgnA (40 μM) from the latter was mixed with TgnB (2 μM) from the former in the presence of ATP (5 mM) and $MgCl_2$ (10 mM). After incubation at room temperature for 12 hours, the reaction mixture was quenched by adding TFA (0.05% final) and separated by HPLC using the gradient of 25-55% Solvent B (in Solvent A) over 30 min. The collected fractions of mTgnA were freeze-dried.

4.4 Protease cleavage of mTgnA and purification of Rn (n = 1-3)

To obtain **R1** and **R2**, mTgnA was digested with chymotrypsin at 30°C for 3 hours (1:100 (w/w) of chymotrypsin:mTgnA; 100 mM Tris (pH 8.0) and 10 mM $CaCl_2$). To obtain **R3**, mTgnA was cleaved by GluC at 25°C for 12 hours (1:20 (w/w) of GluC:mTgnA; 50 mM Tris (pH8.0)). Reaction mixtures were quenched by TFA (0.05% final) and the peptides were separated at a gradient of 20-30% Solvent B (in Solvent A) over 100 min in HPLC. The collected fractions were freeze-dried.

4.5 Connectivity analysis of Rn (n = 1-3)

The **R1** and **R3** peptides were dissolved in 10 mM $NaOCH_3$ in CH_3OH for methanolysis. For partial methanolysis, the reaction

mixtures were quenched by adding large amounts of H_2O with 0.05% TFA. The peptides were separated at a gradient of 20−30% Solvent B (in Solvent A) over 100 min in HPLC. The fully methanolyzed and two partially methanolyzed peptides, which were eluted at different retention times, were analyzed by MALDI−TOF−MS/MS.

Half of the **R2** peptide was dissolved in 10 mM $NaOCH_3$ in CH_3OH for methanolysis. The reaction mixture was quenched by adding large amounts of H_2O with 0.05% TFA. The peptides were purified at a gradient of 10−80% Solvent B (in Solvent A) over 100 min in HPLC. **R2−1**, **R2−2**, and partially methanolyzed **R2**, which were eluted at different retention times, were analyzed by MALDI−TOF−MS/MS. The other half of **R2** was dissolved in 10 mM NaOH in H_2O for hydrolysis. The reaction mixture was quenched and the peptides was separated at a gradient of 42−52% Solvent C (in Solvent A) over 100 min in HPLC. The partially hydrolyzed **R2** with the T69−D79 ester bond was eluted at around 16.9 min and analyzed by MALDI−TOF−MS/MS (Figure 7).

4.6 Over−expression and purification of other peptides and enzymes (CwnB and DbnB)

BL21(DE3) cells carrying a plasmid expressing each protein were grown on an LB agar plate with antibiotics at 37° C for 14 hours. I obtained the precursor peptides (TgnA−R1, CwnA−R1, DbnA−R1, TgnA−R1−D60A, and the CwnA−R1 mutants) using a similar

protocol to the one used for TgnB. The purified precursor peptides were treated with thrombin (EMD Millipore) to remove His_6-tags after the buffer was exchanged to Tris 10 mM (pH 8.0). The leader truncation variants were obtained by the same method as that for TgnA-R1, except for △L3L4L5L6_[Thm]_core. △L3L4L5L6_[Thm]_core was initially purified as a fusion protein with MBP, cleaved by the TEV protease to cut off MBP, and separated from MBP with 30 kDa Amicon centrifugal filter (Millipore). To obtain ATP grasp enzymes (CwnB and DbnB), the same method as the one for TgnB was used except that the culture volume was reduced to 1 L with 100 μg/mL streptomycin.

4.7 Connectivity analysis of various single core peptides

To obtain mTgnA-R1 and its modified variants, TgnA-R1 and its variant precursor peptides were mixed with TgnB. To obtain mCwnA-R1, CwnA-R1 was mixed with CwnB. To obtain mDbnA-R1, DbnA-R1 was mixed with DbnB. The reaction solutions contained 5 μM precursor peptide, 0.5 μM enzyme, 5 mM ATP, and 10 mM $MgCl_2$. To quench the reaction at each time point, EDTA (100 mM final) was added to the reaction solution.

mTgnA-R1 and its modified variants were digested with chymotrypsin (1:100 (w/w) of chymotrypsin:peptide) at 30° C for 3 hours in the presence of 10 mM $CaCl_2$. The reaction mixtures were quenched with TFA (0.05% final) and injected into HPLC. The

peptides were separated at a gradient of 20–30% Solvent B (in Solvent A) over 100 min. mCwnA–R1 were digested with trypsin (1:100 (w/w) of trypsin:peptide) at 25° C for 12 hours. After quenching the reaction, the mixture was injected into HPLC and separated at a gradient of 25–40% Solvent B (in Solvent A) over 75 min. mDbnA–R1 was cleaved by chymotrypsin (1:20 (w/w) of chymotrypsin:peptide) at 30° C for 12 hours in the presence of 10 mM $CaCl_2$. After quenching the reaction, the mixture was injected into HPLC and separated at a gradient of 21–25% Solvent B (in Solvent A) over 30 min. The purified core peptides were freeze–dried. The connectivity of these peptides was analyzed by the same method used for **R1**.

4.8 Connectivity analysis of the CwnA–R1 mutants

The purified core peptides of CwnA–R1 mutants were obtained by the same method used for CwnA–R1. However, before performing MALDI–TOF–MS/MS analysis of CwnA–R1 mutants, the arginine residue in the core peptides of CwnA–R1 mutants was chemically modified to a less basic N5–(4,6–dimethyl–2–pyrimidinyl)ornithine (Pyo).[40] For the Arg–to–Pyo transformation, the core peptide was dissolved in 0.1 M Na_2CO_3/H_2O (pH 8.5) in the presence of 10% acetylacetone (Figure 17). After seven days, the reaction mixture was quenched by adding large amounts of H_2O with TFA (0.05% final), desalted, and analyzed by MALDI–TOF–MS/MS. Between the direct MALDI–TOF–MS/MS spectrum and

that of the arginine modified sample, I used one that gave more information. Therefore, the modified CwnA-R1 mutants were analyzed by MALDI-TOF-MS/MS in three different ways: (i) directly or after arginine modification to Pyo; (ii) after both full methanolysis (50 mM $NaOCH_3$ in CH_3OH) and the arginine modification; (iii) after trypsin digestion, or after both trypsin digestion and arginine modification. Trypsin cleaved the peptide bond that is C-terminal to R27 and transformed the core peptide into two fragments that are cross-linked together.

4.9 Determination of the reaction order of TgnA-R1

The mTgn-R1 core peptide was partially hydrolyzed in 10 mM NaOH. After quenching the reaction, peptides were separated at a gradient of 20-30% Solvent B (in Solvent A) over 100 min in HPLC. The retention times of hydrolysis intermediates were obtained by analyzing each peak with MALDI-TOF-MS/MS.

TgnA-R1 was mixed with TgnB in the same reaction condition as above. The same amount of reaction mixture was quenched at each time point by adding EDTA (100 mM final). They were digested with chymotrypsin (1:100 (w/w) of chymotrypsin:peptide) at 30° C for 3 hours in the presence of 10 mM $CaCl_2$ and separated at a gradient of 20-30% Solvent B (in Solvent A) over 100 min in HPLC.

5. References

[1] Roh, H., Han, Y., Lee, H., and Kim, S. (2018) A topologically distinct modified peptide with multiple bicyclic core motifs expands the diversity of microviridin-like peptides, *ChemBioChem* Epub Dec 21, 2018. DOI:10.1002/cbic.201800678.

[2] Ortega, Manuel A., and van der Donk, W. A. (2016) New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products, *Cell Chem. Biol. 23*, 31-44.

[3] Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K.-D., Fischbach, M. A., Garavelli, J. S., Goransson, U., Gruber, C. W., Haft, D. H., Hemscheidt, T. K., Hertweck, C., Hill, C., Horswill, A. R., Jaspars, M., Kelly, W. L., Klinman, J. P., Kuipers, O. P., Link, A. J., Liu, W., Marahiel, M. A., Mitchell, D. A., Moll, G. N., Moore, B. S., Muller, R., Nair, S. K., Nes, I. F., Norris, G. E., Olivera, B. M., Onaka, H., Patchett, M. L., Piel, J., Reaney, M. J. T., Rebuffat, S., Ross, R. P., Sahl, H.-G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Sussmuth, R. D., Tagg, J. R., Tang, G.-L., Truman, A. W., Vederas, J. C., Walsh, C. T., Walton, J. D., Wenzel, S. C., Willey, J. M., and van der Donk, W. A. (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature, *Nat. Prod. Rep. 30*, 108-160.

[4] White, C. J., and Yudin, A. K. (2011) Contemporary strategies for peptide macrocyclization, *Nature Chemistry 3*, 509.

[5] Schnell, N., Entian, K.-D., Schneider, U., Götz, F., Zähner, H., Kellner, R., and Jung, G. (1988) Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide-rings, *Nature 333*, 276.

[6] Ishitsuka, M. O., Kusumi, T., Kakisawa, H., Kaya, K., and Watanabe, M. M. (1990) Microviridin. A novel tricyclic depsipeptide from the toxic cyanobacterium Microcystis viridis, *J. Am. Chem. Soc. 112*, 8180-8182.

[7] Murakami, M., Sun, Q., Ishida, K., Matsuda, H., Okino, T., and Yamaguchi, K. (1997) Microviridins, elastase inhibitors from the cyanobacterium Nostoc minutum (NIES-26), *Phytochemistry 45*, 1197-1202.

[8] Okino, T., Matsuda, H., Murakami, M., and Yamaguchi, K. (1995) New microviridins, elastase inhibitors from the blue-green alga Microcystis aeruginosa, *Tetrahedron 51*, 10679-10686.

[9] Morishita, Y., Chiba, S., Tsukuda, E., Tanaka, T., Ogawa, T., Yamasaki, M., Yoshida, M., Kawamoto, I., and Matsuda, Y. (1994) RES-701-1, a novel and selective endothelin type B receptor antagonist produced by Streptomyces sp. RE-701. I. Characterization of producing strain, fermentation, isolation, physico-chemical and biological properties, *J Antibiot (Tokyo) 47*, 269-275.

[10] Weber, W., Fischli, W., Hochuli, E., Kupfer, E., and Weibel, E. K. (1991) Anantin--a peptide antagonist of the atrial natriuretic factor (ANF). I. Producing organism, fermentation, isolation and biological activity, *J Antibiot (Tokyo) 44*, 164-171.

[11] McIntosh, J. A., Donia, M. S., and Schmidt, E. W. (2009) Ribosomal peptide natural products: bridging the ribosomal and nonribosomal worlds,

Nat. Prod. Rep. 26, 537-559.

[12] Oman, T. J., and van der Donk, W. A. (2009) Follow the leader: the use of leader peptides to guide natural product biosynthesis, *Nat. Chem. Biol. 6*, 9.

[13] Yang, X., and van der Donk, W. A. (2013) Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: New Insights into the Role of Leader and Core Peptides during Biosynthesis, *Chemistry - A European Journal 19*, 7662-7677.

[14] Ziemert, N., Ishida, K., Liaimer, A., Hertweck, C., and Dittmann, E. (2008) Ribosomal Synthesis of Tricyclic Depsipeptides in Bloom-Forming Cyanobacteria, *Angew. Chem. Int. Ed. 47*, 7756-7759.

[15] Benjamin, P., Guntram, C., Y., Y. W., and K., H. T. (2008) Post-translational Modification in Microviridin Biosynthesis, *ChemBioChem 9*, 3066-3073.

[16] Reyna-González, E., Schmid, B., Petras, D., Süssmuth, R. D., and Dittmann, E. (2016) Leader Peptide-Free In Vitro Reconstitution of Microviridin Biosynthesis Enables Design of Synthetic Protease-Targeted Libraries, *Angew. Chem. 128*, 9544-9547.

[17] Sardar, D., Lin, Z., and Schmidt, Eric W. (2015) Modularity of RiPP Enzymes Enables Designed Synthesis of Decorated Peptides, *Chem. Biol. 22*, 907-916.

[18] Shin, H. J., Murakami, M., Matsuda, H., and Yamaguchi, K. (1996) Microviridins D-F, serine protease inhibitors from the cyanobacterium Oscillatoria agardhii (NIES-204), *Tetrahedron 52*, 8159-8168.

[19] Fujii, K., Sivonen, K., Naganawa, E., and Harada, K.-i. (2000) Non-Toxic Peptides from Toxic Cyanobacteria, Oscillatoria agardhii, *Tetrahedron 56*, 725-733.

[20] Rohrlack, T., Christoffersen, K., Hansen, P. E., Zhang, W., Czarnecki, O., Henning, M., Fastner, J., Erhard, M., Neilan, B. A., and Kaebernick, M. (2003) Isolation, Characterization, and Quantitative Analysis of Microviridin J, a New Microcystis Metabolite Toxic to Daphnia, *J. Chem. Ecol. 29*, 1757-1770.

[21] Li, K., Condurso, H. L., Li, G., Ding, Y., and Bruner, S. D. (2016) Structural basis for precursor protein-directed ribosomal peptide macrocyclization, *Nat. Chem. Biol. 12*, 973.

[22] Philmus, B., Guerrette, J. P., and Hemscheidt, T. K. (2009) Substrate Specificity and Scope of MvdD, a GRASP-like Ligase from the Microviridin Biosynthetic Gene Cluster, *ACS Chem. Biol. 4*, 429-434.

[23] Ahmed, M. N., Reyna-González, E., Schmid, B., Wiebach, V., Süssmuth, R. D., Dittmann, E., and Fewer, D. P. (2017) Phylogenomic Analysis of the Microviridin Biosynthetic Pathway Coupled with Targeted Chemo-Enzymatic Synthesis Yields Potent Protease Inhibitors, *ACS Chem. Biol. 12*, 1538-1546.

[24] Weiz, Annika R., Ishida, K., Makower, K., Ziemert, N., Hertweck, C., and Dittmann, E. (2011) Leader Peptide and a Membrane Protein Scaffold Guide the Biosynthesis of the Tricyclic Peptide Microviridin, *Chem. Biol. 18*, 1413-1421.

[25] Ziemert, N., Ishida, K., Weiz, A., Hertweck, C., and Dittmann, E. (2010) Exploiting the Natural Diversity of Microviridin Gene Clusters for Discovery of Novel Tricyclic Depsipeptides, *Appl. Environ. Microbiol. 76*, 3568-3574.

[26] Zhang, Y., Li, K., Yang, G., McBride, J. L., Bruner, S. D., and Ding, Y. (2018) A distributive peptide cyclase processes multiple microviridin core peptides within a single polypeptide substrate, *Nat. Commun. 9*, 1780.

[27] Gu, W., Sardar, D., Pierce, E., and Schmidt, E. W. (2018) Roads to Rome: Role of Multiple Cassettes in Cyanobactin RiPP Biosynthesis, *J. Am. Chem. Soc. 140*, 16213−16221.

[28] Weiz, A. R., Ishida, K., Quitterer, F., Meyer, S., Kehr, J.−C., Müller, K. M., Groll, M., Hertweck, C., and Dittmann, E. (2014) Harnessing the Evolvability of Tricyclic Microviridins To Dissect Protease−Inhibitor Interactions, *Angew. Chem. Int. Ed. 53*, 3735−3738.

[29] Iyer, L. M., Abhiman, S., Maxwell Burroughs, A., and Aravind, L. (2009) Amidoligases with ATP−grasp, glutamine synthetase−like and acetyltransferase−like domains: synthesis of novel metabolites and peptide modifications of proteins, *Mol. BioSyst. 5*, 1636−1660.

[30] Lee, H., Park, Y., and Kim, S. (2017) Enzymatic Cross−Linking of Side Chains Generates a Modified Peptide with Four Hairpin−like Bicyclic Repeats, *Biochemistry 56*, 4927−4930.

[31] Crooks, G. E., Hon, G., Chandonia, J.−M., and Brenner, S. E. (2004) WebLogo: A Sequence Logo Generator, *Genome Res. 14*, 1188−1190.

[32] Béla, P., and Sándor, S. (2005) Fragmentation pathways of protonated peptides, *Mass Spectrom. Rev. 24*, 508−548.

[33] Hjernø, K., and Højrup, P. (2015) Interpretation of Tandem Mass Spectrometry (MSMS) Spectra for Peptide Analysis, In *Peptide Antibodies: Methods and Protocols* (Houen, G., Ed.), pp 83−102, Springer New York, New York.

[34] Yu, W., Vath, J. E., Huberty, M. C., and Martin, S. A. (1993) Identification of the facile gas−phase cleavage of the Asp−Pro and Asp−Xxx peptide bonds in matrix−assisted laser desorption time−of−flight mass spectrometry, *Anal. Chem. 65*, 3015−3023.

[35] Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P. J., Joewono, I.,

Rusch, D., Chisholm, S. W., and van der Donk, W. A. (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria, *Proc. Natl. Acad. Sci. U. S. A. 107*, 10430−10435.

[36] Tang, W., and van der Donk, W. A. (2012) Structural Characterization of Four Prochlorosins: A Novel Class of Lantipeptides Produced by Planktonic Marine Cyanobacteria, *Biochemistry 51*, 4271−4279.

[37] Shi, Y., Yang, X., Garg, N., and van der Donk, W. A. (2011) Production of Lantipeptides in Escherichia coli, *J. Am. Chem. Soc. 133*, 2338−2341.

[38] Yu, Y., Mukherjee, S., and van der Donk, W. A. (2015) Product Formation by the Promiscuous Lanthipeptide Synthetase ProcM is under Kinetic Control, *J. Am. Chem. Soc. 137*, 5140−5148.

[39] Mukherjee, S., and van der Donk, W. A. (2014) Mechanistic Studies on the Substrate−Tolerant Lanthipeptide Synthetase ProcM, *J. Am. Chem. Soc. 136*, 10450−10459.

[40] Sergei, D., W., K. J., and H., R. D. (1997) Improving mass spectrometric sequencing of arginine-containing peptides by derivatization with acetylacetone, *J. Mass Spectrom. 32*, 1337−1349.

[41] Hegemann, J. D., Zimmermann, M., Xie, X., and Marahiel, M. A. (2015) Lasso Peptides: An Intriguing Class of Bacterial Natural Products, *Acc. Chem. Res. 48*, 1909−1919.

[42] Tietz, J. I., Schwalen, C. J., Patel, P. S., Maxson, T., Blair, P. M., Tai, H.−C., Zakai, U. I., and Mitchell, D. A. (2017) A new genome−mining tool redefines the lasso peptide biosynthetic landscape, *Nat. Chem. Biol. 13*, 470.

[43] Yu, Y., Duan, L., Zhang, Q., Liao, R., Ding, Y., Pan, H., Wendt−Pienkowski, E., Tang, G., Shen, B., and Liu, W. (2009) Nosiheptide

Biosynthesis Featuring a Unique Indole Side Ring Formation on the Characteristic Thiopeptide Framework, *ACS Chem. Biol. 4*, 855-864.

[44] Liao, R., Duan, L., Lei, C., Pan, H., Ding, Y., Zhang, Q., Chen, D., Shen, B., Yu, Y., and Liu, W. (2009) Thiopeptide Biosynthesis Featuring Ribosomally Synthesized Precursor Peptides and Conserved Posttranslational Modifications, *Chem. Biol. 16*, 141-147.

[45] van den Ent, F., and Löwe, J. (2006) RF cloning: A restriction-free method for inserting target genes into plasmids, *J. Biochem. Bioph. Methods 67*, 67-74.

[46] Ochman, H., Gerber, A. S., and Hartl, D. L. (1988) Genetic Applications of an Inverse Polymerase Chain Reaction, *Genetics 120*, 621-623.

[47] Kim, S., Grant, R. A., and Sauer, R. T. (2011) Covalent linkage of distinct substrate degrons controls assembly and disassembly of DegP proteolytic cages, *Cell 145*, 67-78.

6. Figures

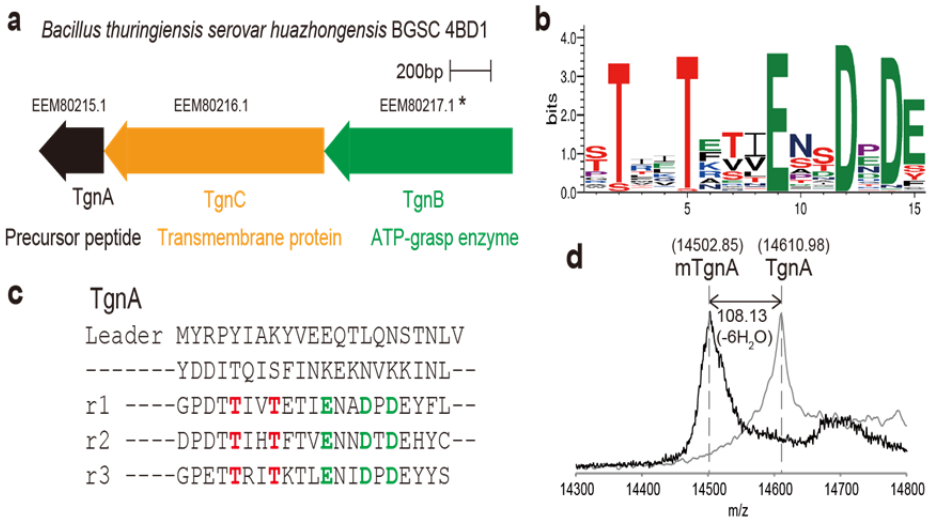
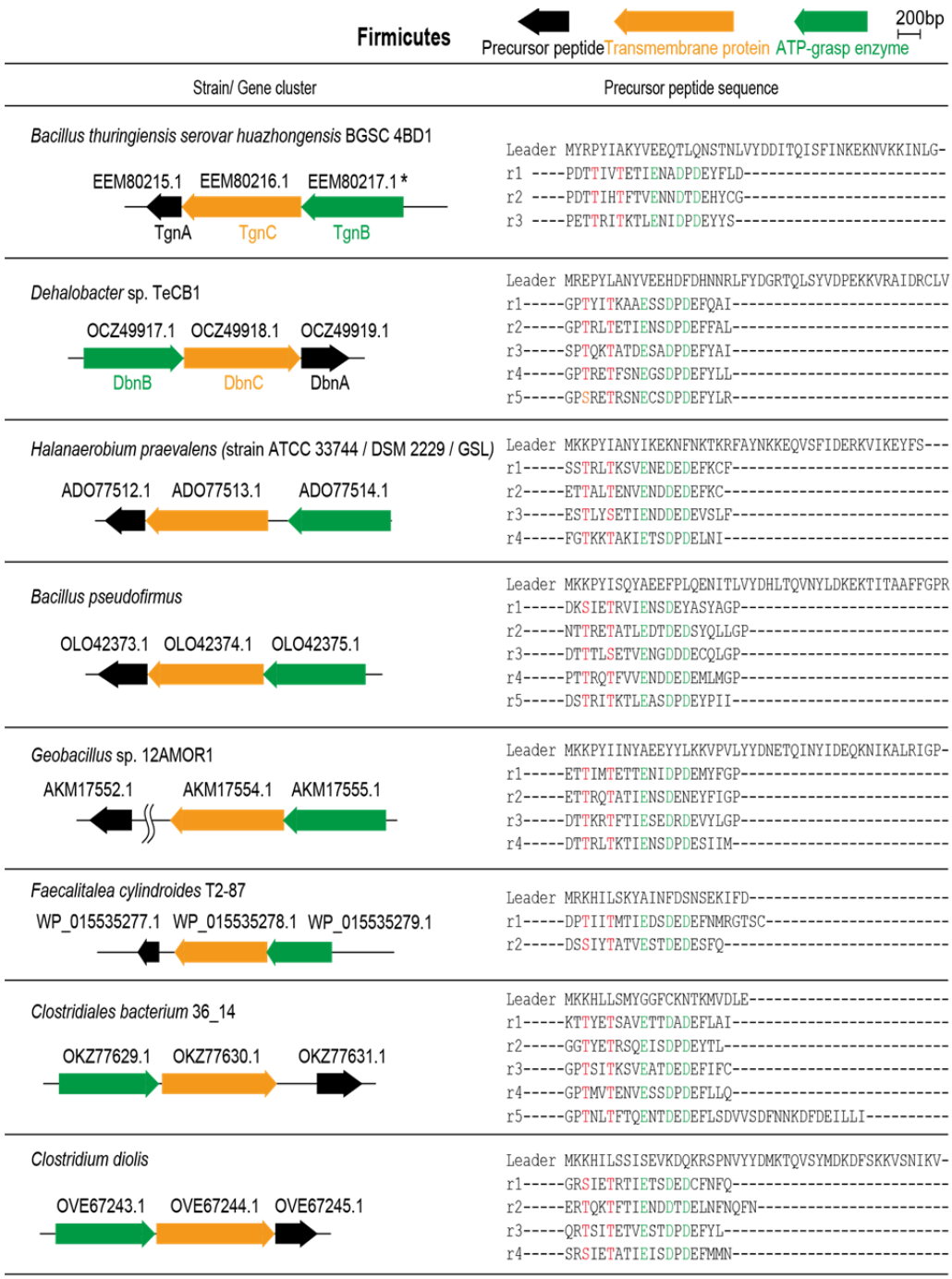
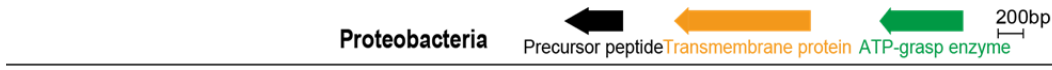


Figure 1. Identification of a new microviridin-like modified peptide, mTgnA. (a) A new putative RiPP gene cluster identified from *Bacillus thuringiensis serovar huazhongensis* BGSC 4BD1. Arrows indicate the size and orientation of the genes. GenBank accession numbers and new names for encoded proteins are shown above and below the arrows, respectively. The *tgnB* gene contains additional 261 bases upstream of the gene that encodes EEM80217.1 (EEM80217.1*). (b) Sequence logo of the core repeats of TgnA and its fifteen homologs. The TxxTxxxExxDxD motif is highly conserved. (c) Sequence alignment of three core repeats of TgnA. The highly conserved threonines and acidic amino acids are colored red and green, respectively. (d) MALDI-TOF-MS spectra of TgnA and mTgnA. To obtain mTgnA, TgnA was co-expressed and co-purified with TgnB, and the mixed solution was further incubated with ATP (5 mM) and $MgCl_2$ (10 mM) for twelve hours *in vitro*.


Firmicutes
Precursor peptide
Transmembrane protein
ATP-grasp enzyme
200bp
Strain/ Gene cluster
Precursor peptide sequence
Bacillus thuringiensis serovar huazhongensis BGSC 4BD1
EEM80215.1 EEM80216.1 EEM80217.1 *
TgnA TgnC TgnB
Leader MYRPYIAKYVEEQTLQNSTNLVYDDITQISFINKEKNVKKINLG-
r1 ----PDTTIVTETIENADPDEYFLD----------
r2 ----PDTTIHTFTVENNDTDEHYCG----------
r3 ----PETTRITKTLENIDPDEYYS----------
Dehalobacter sp. TeCB1
OCZ49917.1 OCZ49918.1 OCZ49919.1
DbnB DbnC DbnA
Leader MREPYLANYVEEHDFDHNNRLFYDGRTQLSYVDPEKKVRAIDRCLV
r1-----GPTYITKAAESSDPDEFQAI----------
r2-----GPTRLTETIENSDPDEFFAL----------
r3-----SPTQKTATDESADPDEFYAI----------
r4-----GPTRETFSNEGSDPDEFYLL----------
r5-----GPSRETRSNECSDPDEFYLR----------
Halanaerobium praevalens (strain ATCC 33744 / DSM 2229 / GSL)
ADO77512.1 ADO77513.1 ADO77514.1
Leader MKKPYIANYIKEKNFNKTKRFAYNKKEQVSFIDERKVIKEYFS---
r1-----SSTRLTKSVENEDEDEFKCF----------
r2-----ETTALTENVENDDEDEFKC----------
r3-----ESTLYSETIENDDEDEVSLF----------
r4-----FGTKKTAKIETSDPDELNI----------
Bacillus pseudofirmus
OLO42373.1 OLO42374.1 OLO42375.1
Leader MKKPYISQYAEEFPLQENITLVYDHLTQVNYLDKEKTITAAFFGPR
r1-----DKSIETRVIENSDEYASYAGP----------
r2-----NTTRETATLEDTDEDSYQLLGP----------
r3-----DTTTLSETVENGDDDECQLGP----------
r4-----PTTRQTFVVENDDEDEMLMGP----------
r5-----DSTRITKTLEASDPDEYPII----------
Geobacillus sp. 12AMOR1
AKM17552.1 AKM17554.1 AKM17555.1
Leader MKKPYIINYAEEYYLKKVPVLYYDNETQINYIDEQKNIKALRIGP-
r1-----ETTIMTETTENIDPDEMYFGP----------
r2-----ETTRQTATIENSDENEYFIGP----------
r3-----DTTKRTFTIESEDRDEVYLGP----------
r4-----DTTRLTKTIENSDPDESIIM----------
Faecalitalea cylindroides T2-87
WP_015535277.1 WP_015535278.1 WP_015535279.1
Leader MRKHILSKYAINFDSNSEKIFD----------
r1-----DPTIITMTIEDSDEDEFNMRGTSC----------
r2-----DSSIYTATVESTDEDESFQ----------
Clostridiales bacterium 36_14
OKZ77629.1 OKZ77630.1 OKZ77631.1
Leader MKKHLLSMYGGFCKNTKMVDLE----------
r1-----KTTYETSAVETTDADEFLAI----------
r2-----GGTYETRSQEISDPDEYTL----------
r3-----GPTSITKSVEATDEDEFIFC----------
r4-----GPTMVTENVESSDPDEFLLQ----------
r5-----GPTNLTFTQENTDEDEFLSDVVSDFNNKDFDEILLI----------
Clostridium diolis
OVE67243.1 OVE67244.1 OVE67245.1
Leader MKKHILSSISEVKDQKRSPNVYYDMKTQVSYMDKDFSKKVSNIKV-
r1-----GRSIETRTIETSDEDCFNFQ----------
r2-----ERTQKTFTIENDDTDELNFNQFN----------
r3-----QRTSITETVESTDPDEFYL----------
r4-----SRSIETATIEISDPDEFMMN----------



Strain/ Gene cluster	Precursor peptide sequence
Colwellia sp. UCD-KL20 WP_076419414.1 WP_076419415.1 WP_076419416.1 CwnB CwnC CwnA	Leader MKPYILNYSESIDIKHVTSKSIT---------- r1 ----DTTRLTETIEQIDQDELATL---------- r2 ----DYTSKTESIEPSDNDELFT---------- r3 ----NSTTQTFVIEDSDEDQVGSEIQ---------- r4 ----CSTLVTKTSEPSDNDEILT---------- r5 ----CSTWATESTEPSDNDEIFS---------- r6 ----FSTLQTRTIEPSDDDSINQ---------- r7 ----LSTLQTNTSENWDHDEIILN----------
Vibrio alginolyticus KOE80492.1 KOE80493.1 KOE80494.1	Leader MKPYIMNYSETIEISPVGSTFHTI---------- r1-----DTTTKTFTIESHDNDEIRC---------- r2-----DWTIQTNVTEPTDNDSILCL---------- r3-----DETAITKTLEPTDQDCMHL---------- r4-----DATYETRVVESSDNDSLYF---------- r5-----DTTVATESIEGLDPDEIVLAH----------
Serratia marcescens CVF09414.1 CVF09447.1 CVF09495.1	Leader MKPYVMLYSETYRKPTQSLEHLI---------- r1-----DSTVRTFTIESDDNDSIN---------- r2-----AGTMITKVNENGDDDFIH---------- r3-----MSTFETREIENSDQNFYY---------- r4-----QSTVVTESIENDDNDYLS---------- r5-----CATLVTNVVESNDNDFHSIYNDKIVSYDHLIN----------
Yersinia frederiksenii Y225 AJI88106.1 AJI87333.1 AJI87385.1	Leader MKPYIMNYSETITMKQSVAIQNTNIL---------- r1 ----DVTTRTFTIETDDHDEILG---------- r2 ----LATIITERIENSDRDEISA---------- r3 ----YSTFQTNIVENSDKDEISC---------- r4 ----AGTMVTAVIENSDTDEYLL---------- r5 ----FSTMETRELEAADSEEISYLMS---------- r6 ----DSTIMTKVFEATDNDE----------
Yersinia mollaretii CQQ19315.1 CQQ19343.1 CQQ19371.1	Leader MKPYIMNYSETITMKQSDAIQNKNIL---------- r1 ----DITTKTFTIETDDHDEVLG---------- r2 ----LATIITERIENSDRDEISA---------- r3 ----YSTFQTHTVENSDKDEISC---------- r4 ----AGTMVTAVIENSDTDECLL---------- r5 ----FSTVETRELEAADSDEISYLMADT---------- r6 ----DTTIITNELEATDDDE----------
Escherichia coli O127:H6 (strain E2348/70 / EPEC) CAS09659.1 CAS09660.1 CAS09661.1	Leader MKPYTMLYSESFEINASTISSKT---------- r1-----NSTVKTFTTEPNDEDYLN---------- r2-----TWTTHTRVIENQDEDYSLAPSNCERKYEISNEKQSA------- r3-----MDTFFTESIENNDDDFIV---------- r4-----CRTFVTNVVESGDDDFYFN----------
Shewanella oneidensis (strain MR-1) AFV73513.1 AAN53577.1	Leader MKPYILQYSETFEFIQQAHQMQLTK---------- r1-----GETWITRQMEPSDEDQVSFA---------- r2-----DTTTLTFVVENIDHDELLLVD---------- r3-----ETTRLTKTTEVSDPDEIVMFG---------- r4-----ETTLITNVVENVDPDEITLM----------
Enterobacter sp. GN02600 KLQ25722.1 KLQ25723.1 KLQ25724.1	Leader MKPYILQYSETFEFIQQAHQMQLTK---------- r1-----DSTRKTFSIESDDNDYLDTS---------- r2-----VWTTITKVNEATDQDSIFLVSNPPQKQSNFDYIA--------- r3-----TSTLLTERTENNDDDYIS---------- r4-----LSTIITNITESQDDDYSYS----------

Figure 2. Putative gene clusters for the mTgnA-like RiPPs. Most gene clusters sequentially encode an ATP grasp enzyme (green), a transmembrane protein (orange), and a potential precursor peptide (black). Arrows indicate the size and orientation of the genes. The accession numbers for encoded proteins are listed above the arrows. The *tgnB* gene contains additional 261 bases upstream of the gene that encodes EEM80217.1 (EEM80217.1*). The precursor peptides consist of the leader peptide and multiple repeats of the core peptide. The highly-conserved threonines and acidic residues (glutamate and aspartate) in the TxxTxxxExxDxD motif of the core peptide are colored red and green, respectively.

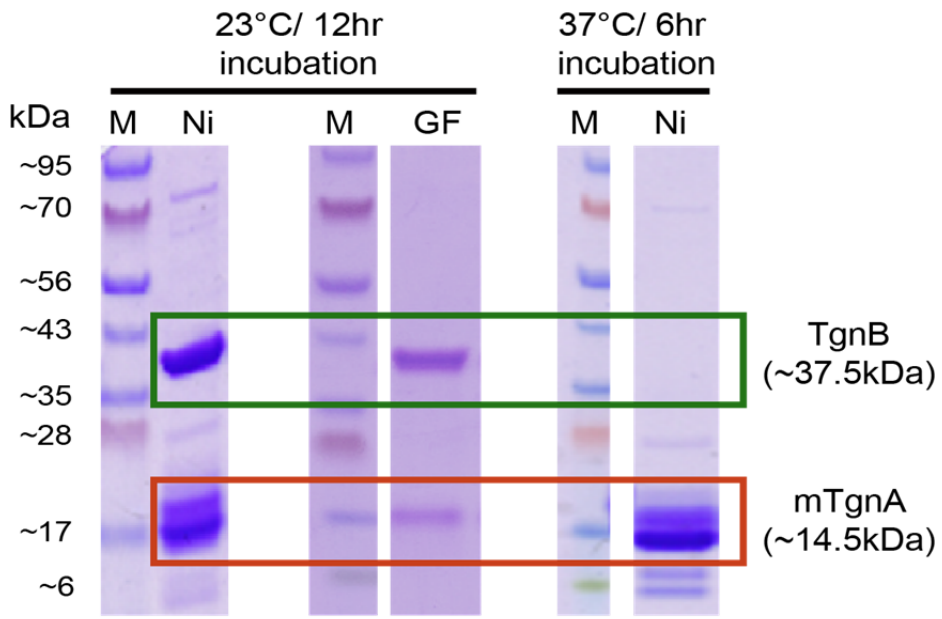
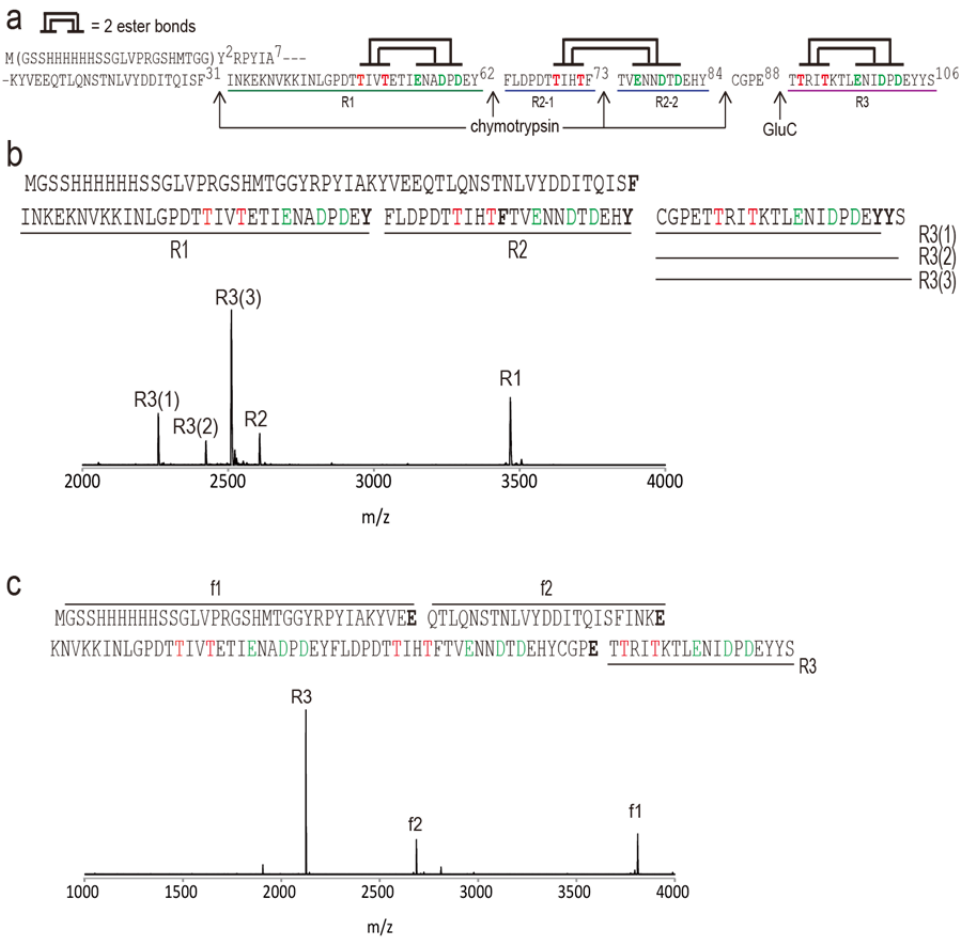


Figure 3. The SDS-PAGE gel of purified mTgnA and TgnB from TgnA-TgnB co-expression in different culture conditions. All panels show the purified proteins from cells incubated at 23° C or 37° C. M: GenDEPOT Protein marker (DAWINBIO), Ni: Eluate after immobilized metal ion affinity chromatography using Ni beads, GF: Eluate after gel filtration chromatography.



Repeat	R1	R2	R3
Cleavage pattern		R2-1 R2-2	
Precursor m/z (Calc.)	3501.796	2624.143	2159.051
Product m/z (Obs.)	3465.530	2606.251	2123.025
△ (product-precursor)	-36.266 (-2H_2O)	-17.892 (-H_2O)	-36.026 (-2H_2O)

Figure 4. Proteolysis of mTgnA to produce R1-R3 (a) The full sequence of TgnA with the conserved residues colored in red and green. Residues are numbered based on the original sequence of TgnA, and the his-tag and thrombin cleavage site shown in parentheses are inserted between Met^{1} and Tyr^{2}. Cleavage sites of chymotrypsin or GluC are indicated in arrows, and the positions of putative ester cross-links are shown above the sequence. (b and c) MALDI spectrum of mTgnA fragments after chymotrypsin cleavage (b) or GluC cleavage (c). Proteases cleave the peptide bonds that are located at C-termini of amino acids indicated in bold print. (d) Schemes and molecular weights for three peptide fragments (**R1-R3**) that are obtained from the protease cleavage of mTgnA and that contain each core repeat.

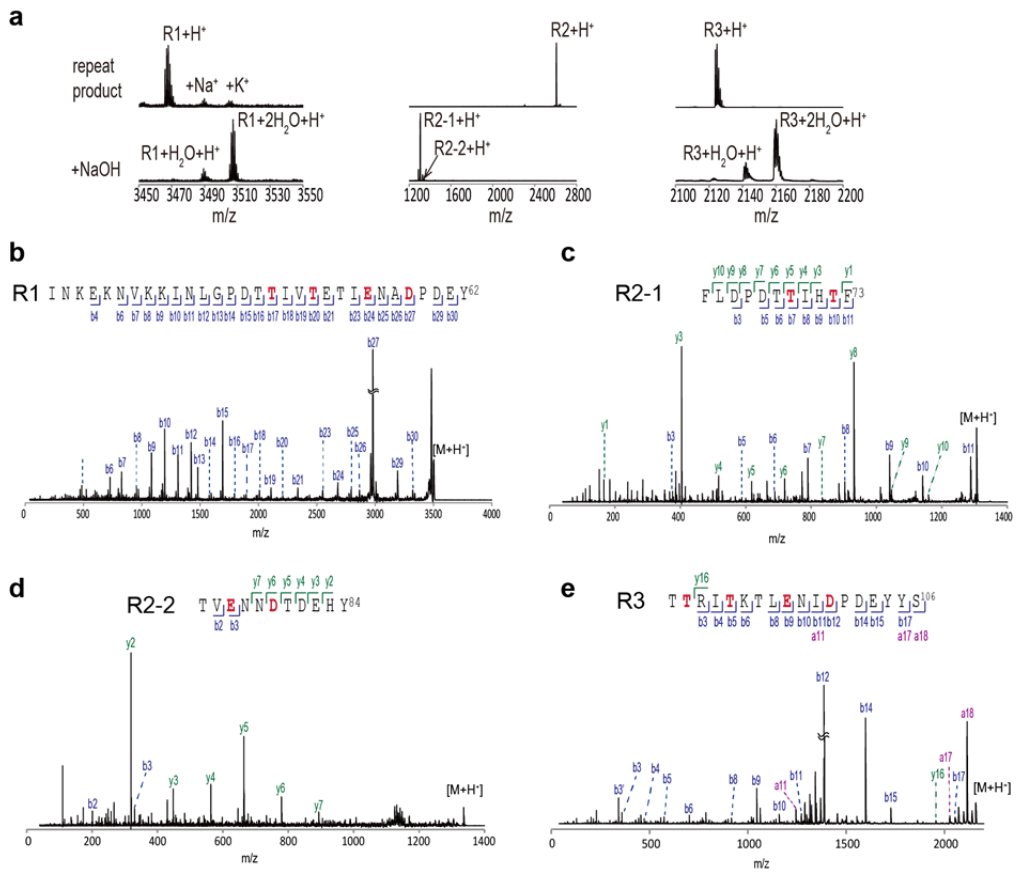


Figure 5. Hydrolysis of R1−R3 (a) MALDI spectra of **R1−R3** and their hydrolyzed peptides. **R1−R3** were hydrolyzed with NaOH (100 mM) for 12 hours. (b) MALDI−TOF−MS/MS spectrum of the fully hydrolyzed **R1**. (c and d) MALDI−TOF−MS/MS spectrum of **R2−1** (c) or **R2−2** (d) after full hydrolysis of **R2**. (e) MALDI−TOF−MS/MS spectrum of the fully hydrolyzed **R3**. The b3' ion is the b3 ion with loss of ammonia. The amino acids making ester bonds in **R1−R3** are colored red. Calculated and observed masses for all mass spectra are in Table 5.

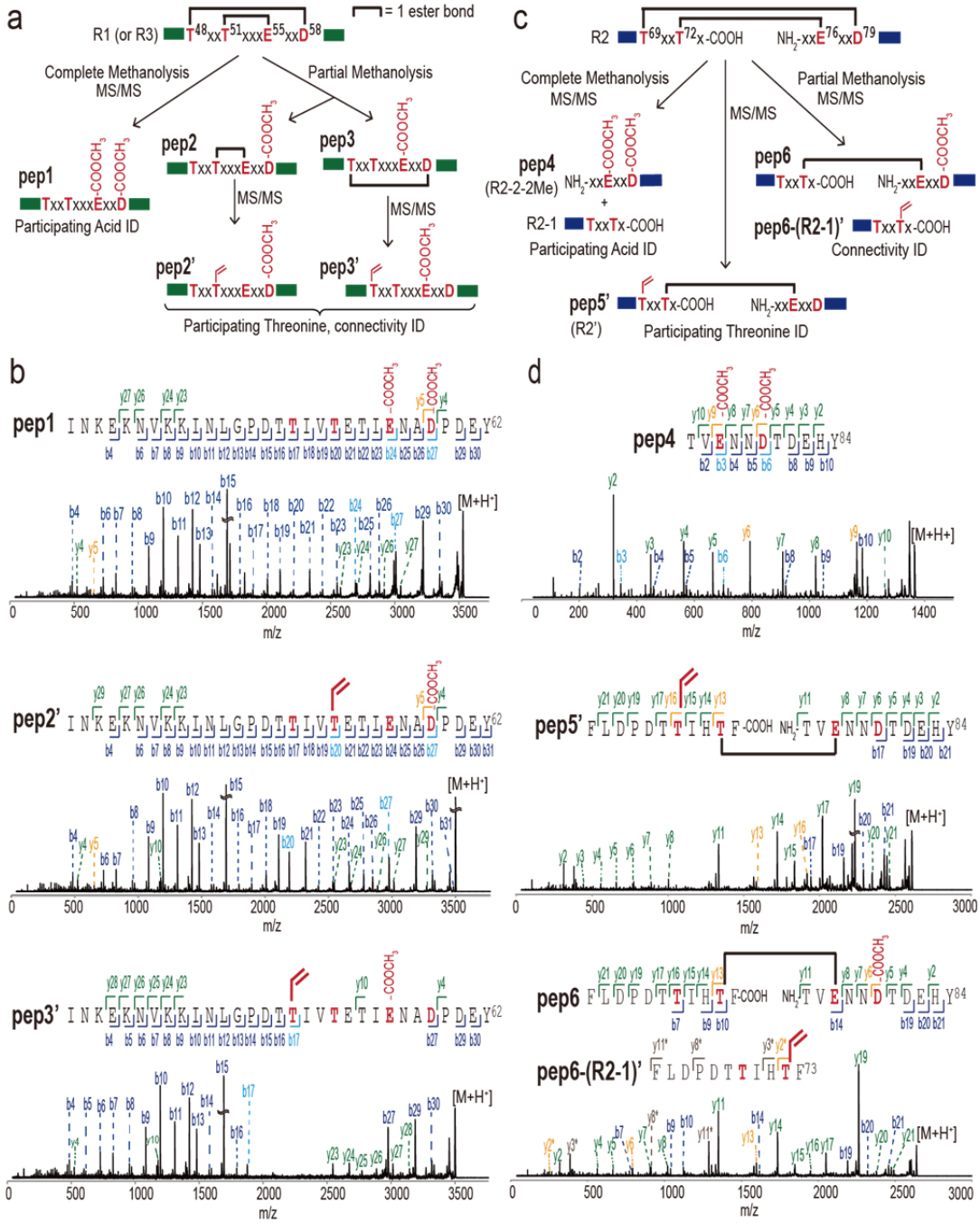

a
R1 (or R3)
= 1 ester bond
Complete Methanolysis MS/MS
Partial Methanolysis
pep1
pep2
pep3
MS/MS
pep2'
pep3'
Participating Acid ID
Participating Threonine, connectivity ID
c
R2
Complete Methanolysis MS/MS
Partial Methanolysis MS/MS
MS/MS
pep4
(R2-2-2Me)
R2-1
Participating Acid ID
pep6
pep6-(R2-1)'
Connectivity ID
pep5'
(R2')
Participating Threonine ID
b
pep1
pep2'
pep3'
[M+H+]
m/z
d
pep4
pep5'
pep6
pep6-(R2-1)'
[M+H+]
m/z

Figure 6. Determining the cross-linking connectivity of R1-R3 (a and c) Schematic representations of the work flow to determine the cross-linking connectivity of **R1** or **R3** (a) and **R2** (c). Residues participating in ester bond formation and their modified forms are shown in red. Complete or partial methanolysis of **R1** or **R2** results in **pep1-pep4** or **pep6** peptides. **PepN'** is a peptide where its ester bond is rearranged in the mass analyzer. (b and d) MALDI-TOF-MS/MS spectra of the indicated peptides that are originated from **R1** (b) and **R2** (d). The b-series ions and y-series ions are colored blue and green, of which ions with distinct modifications are colored light blue and orange, respectively. The y-series ions with asterisk (*) are fragments from **pep6-(R2-1)'**. Predicted and observed masses for all mass spectra are in Table 6.

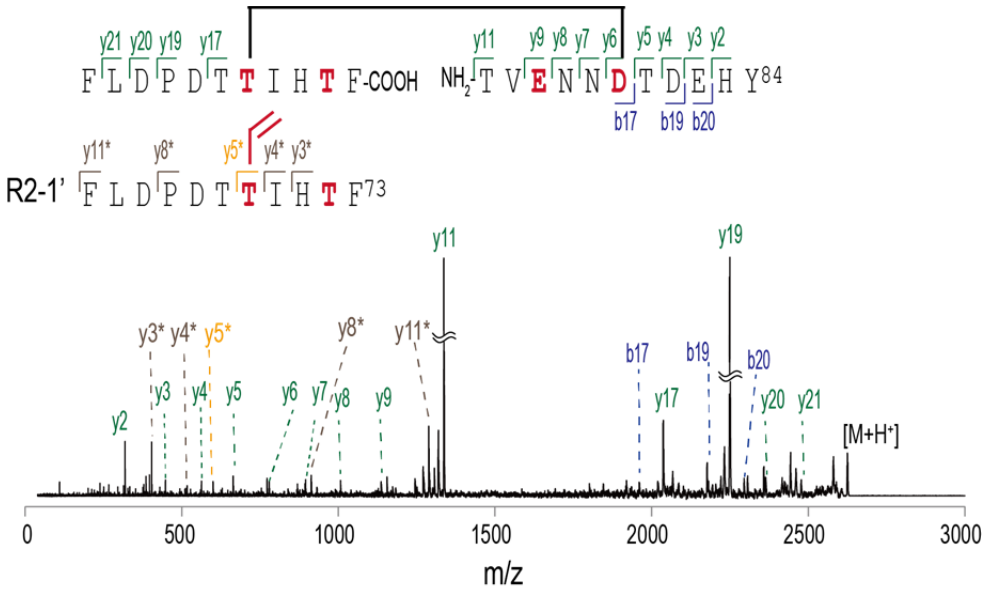


Figure 7. MALDI-TOF-MS/MS spectrum of the partially hydrolyzed R2 with the T69-D79 ester bond. The y-series ions with asterisk are fragments from R2-1' formed by the MS/MS-induced rearrangement. The red amino acids present residues that make ester bonds in R2 and the orange y-series ion (y5*) indicates the dehydration. Calculated and observed masses for all mass spectra are in Table 7.

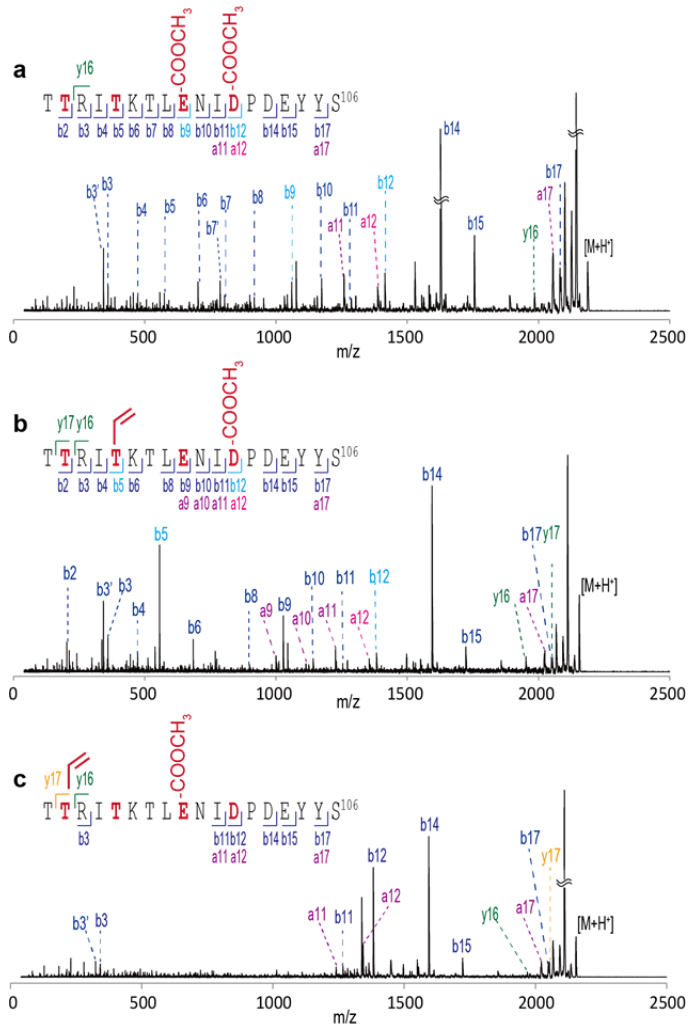


Figure 8. MALDI-TOF-MS/MS spectra of the methanolyzed R3. (a) MALDI-TOF-MS/MS spectrum of the fully methanolyzed **R3**. (b) MALDI-TOF-MS/MS spectrum of the partially methanolyzed **R3** where the outer ester bond (T90-D100) is methanolyzed. (c) MALDI-TOF-MS/MS spectra of the partially methanolyzed **R3** where the inner ester bond (T93-E97) is methanolyzed. The b*n*' ion is the b*n* ion with loss of ammonia. The amino acids making ester bonds in **R3** are colored red. Light blue b-series ions, pink a-series ions, and orange y-series ions indicate the modifications. Calculated and observed masses for all mass spectra are in Table 8.

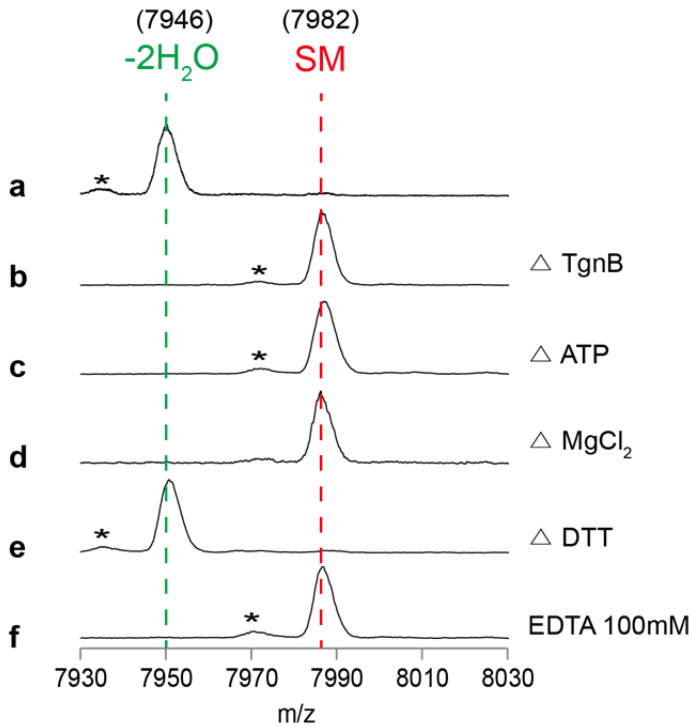


Figure 9. MALDI analyses of the *in vitro* reactions with TgnA-R1. All the reaction solutions were quenched after two hours. (a) The reaction solution contained 5 μM TgnA-R1, 0.5 μM TgnB, 5 mM ATP, 10 mM $MgCl_2$, and 1 mM DTT. (b-e) The reaction conditions were the same as (a) but each factor on the right is missing. (f) The reaction condition was the same as (a) but EDTA (100 mM final) was added. Red and green dotted lines indicate the peaks for the starting material (SM) and the doubly dehydrated peptide (-$2H_2O$), respectively. The numbers in parentheses are monoisotopic masses. The asterisk indicates a peak resulting from the laser-induced deamination which is not related to the reaction.

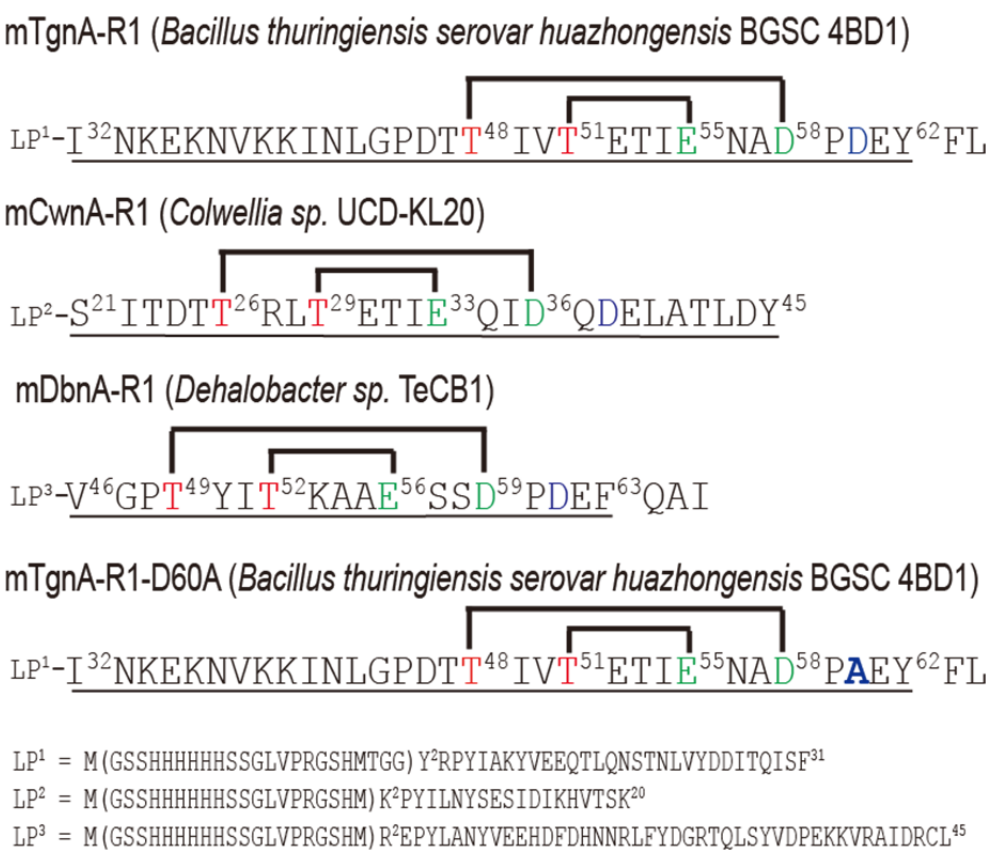
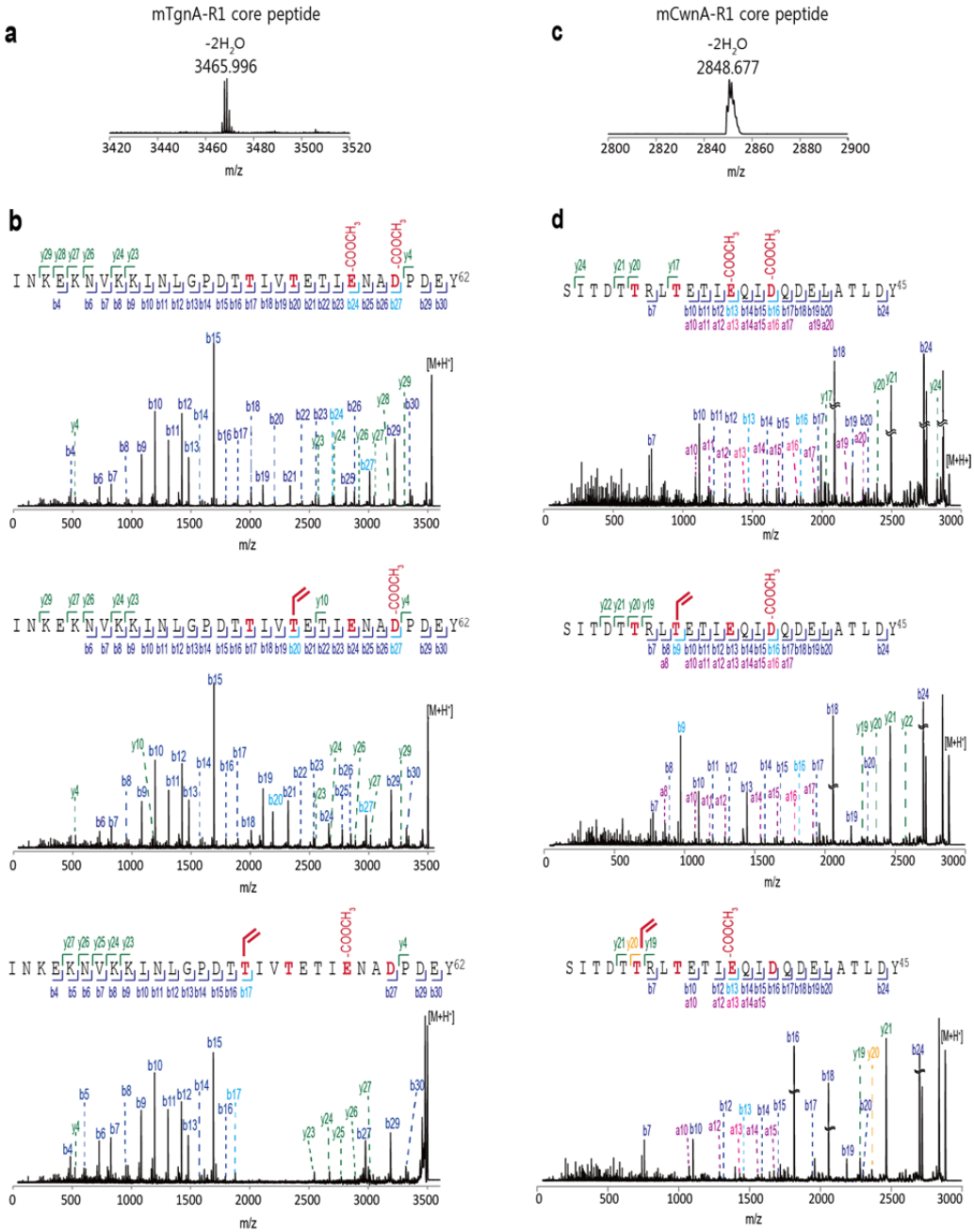


Figure 10. The cross-linking connectivity of mTgnA-R1, mCwnA-R1, mDbn-R1 and mTgnA-R1-D60A, all of which contain only the first core repeat with modification as well as the leader peptide. CwnA and DbnA are homologs of TgnA. Genes encoding homologous precursors are found in bacterial strains shown in parentheses. The underlined peptides are fragments used in MALDI-TOF-MS/MS analysis after proteolysis. All modified peptides show the same cross-linking connectivity.


a
mTgnA-R1 core peptide
-2H2O
3465.996
3420 3440 3460 3480 3500 3520
m/z
c
mCwnA-R1 core peptide
-2H2O
2848.677
2800 2820 2840 2860 2880 2900
m/z
b
d
[M+H+]
m/z

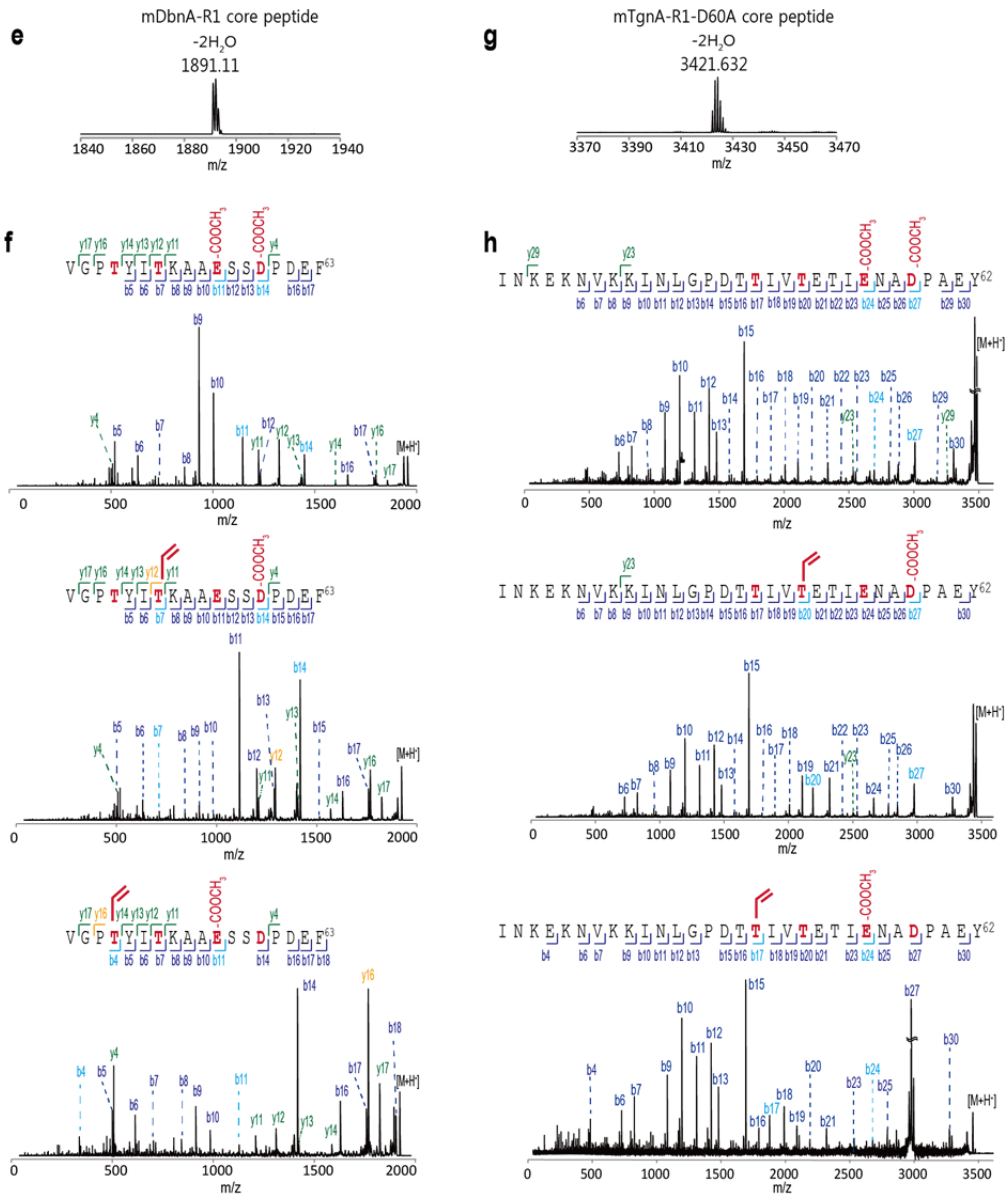

e
mDbnA-R1 core peptide
-2H2O
1891.11
m/z
g
mTgnA-R1-D60A core peptide
-2H2O
3421.632
m/z
f
h
[M+H+]
m/z

Figure 11. MALDI-TOF-MS and MALDI-TOF-MS/MS spectra of various single core peptides. MALDI-TOF-MS and MALDI-TOF-MS/MS spectra of the core peptides of (a and b) mTgnA-R1, (c and d) mCwnA-R1, (e and f) mDbn-R1, and (g and h) mTgnA-R1-D60A. All core peptides showed two dehydrations and the number above each peak is the observed monoisotopic mass of the core peptide. The three MALDI-TOF-MS/MS spectra were obtained with the fully methanolyzed core peptides (upper), the partially methanolyzed core peptides in which the outer ester bond is methanolyzed (middle), and the partially methanolyzed core peptides in which the inner ester bond is methanolyzed (lower). The amino acids making ester bonds in modified peptides are colored red. Light blue b-series ions, pink a-series ions, and orange y-series ions indicate the modifications. Calculated and observed masses for all mass spectra are in Table 9.

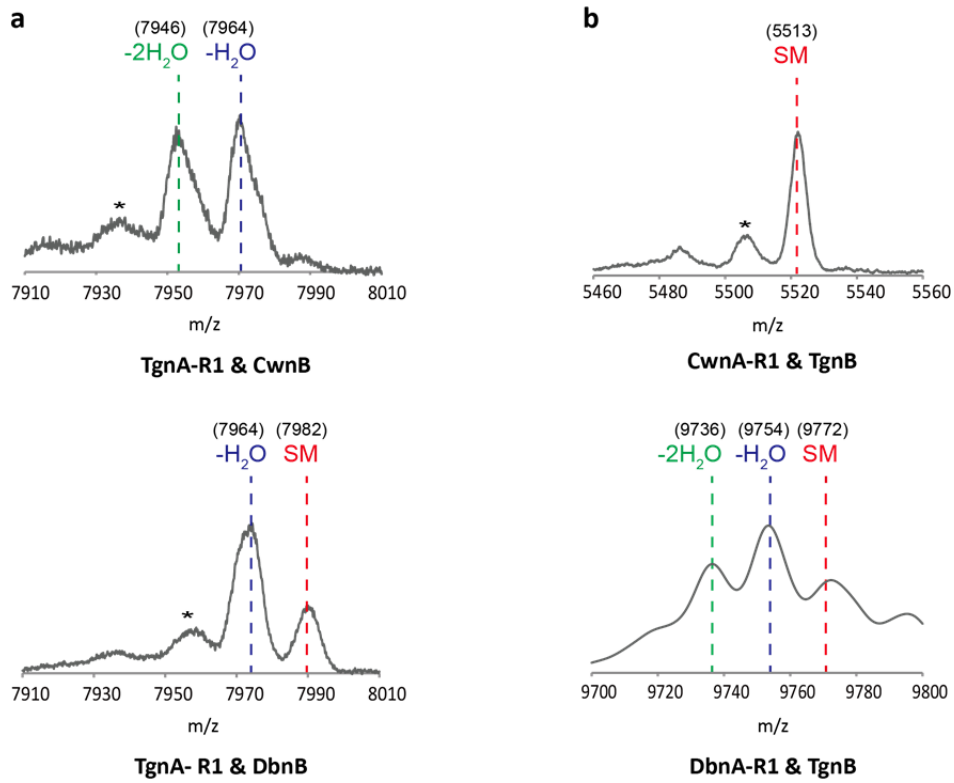


Figure 12. MALDI analysis of the cross reaction within the thuringinin group. (A) MALDI spectra of TgnA-R1 after the twelve hours *in vitro* reaction with CwnB (upper) or DbnB (lower). (B) MALDI spectrum of CwnA-R1 after the twelve hours *in vitro* reaction with TgnB (upper), and that of DbnA-R1 after the co-expression with TgnB (lower). DbnA-R1 was detected by MALDI-TOF in linear mode. The *in vitro* reaction condition was the same as that in Figure 9e. Red, blue, and green dotted lines indicate the peaks for SM, the singly-dehydrated peptides, and the doubly-dehydrated peptides, respectively. The numbers in parentheses are monoisotopic masses except those of DbnA-R1 which are average masses. The asterisk indicates a peak resulting from the laser-induced deamination which is not related to the reaction.

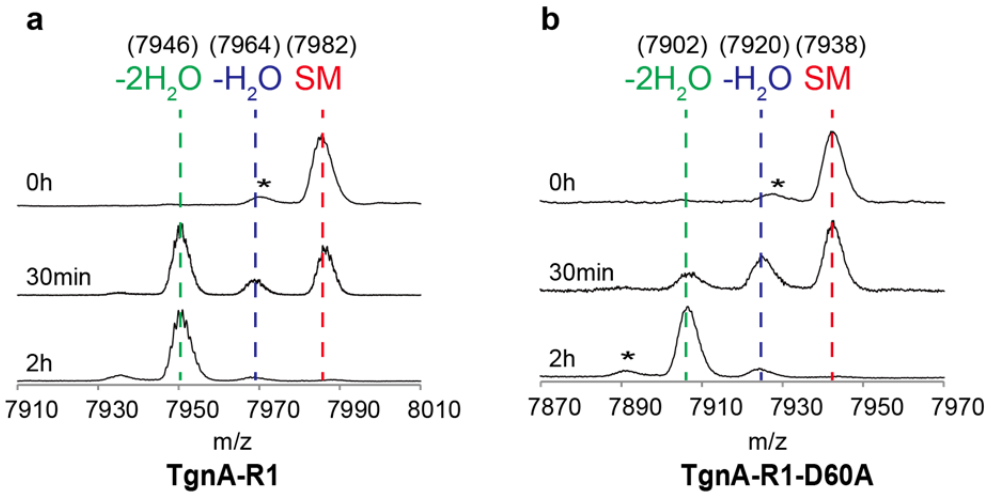


Figure 13. MALDI analysis of the *in vitro* reactions of TgnA-R1 and TgnA-R1-D60A. MALDI spectra of the reaction of (a) TgnA-R1 and (b) TgnA-R1-D60A with TgnB. The reaction condition was the same as that in Figure 9e. The reaction was quenched with EDTA (100 mM final) at each time point. Red, blue, and green dotted lines indicate the peaks for SM, the singly-dehydrated peptides, and the doubly-dehydrated peptides, respectively. The numbers in parentheses are monoisotopic masses. The asterisk indicates a peak resulting from the laser-induced deamination which is not related to the reaction.

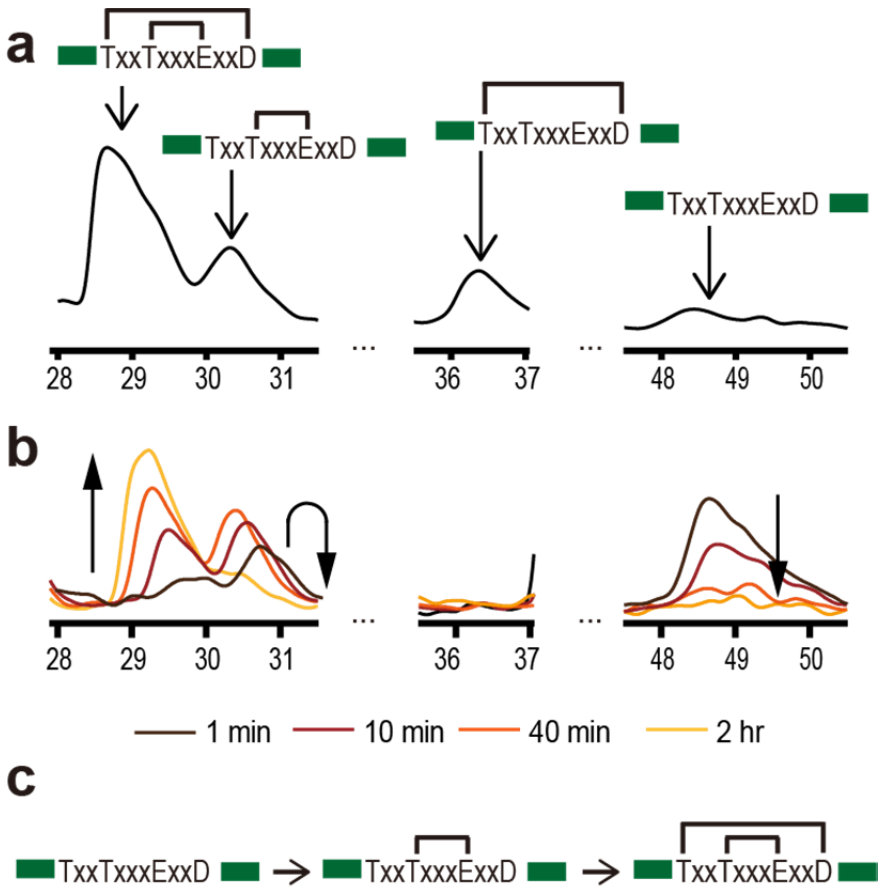
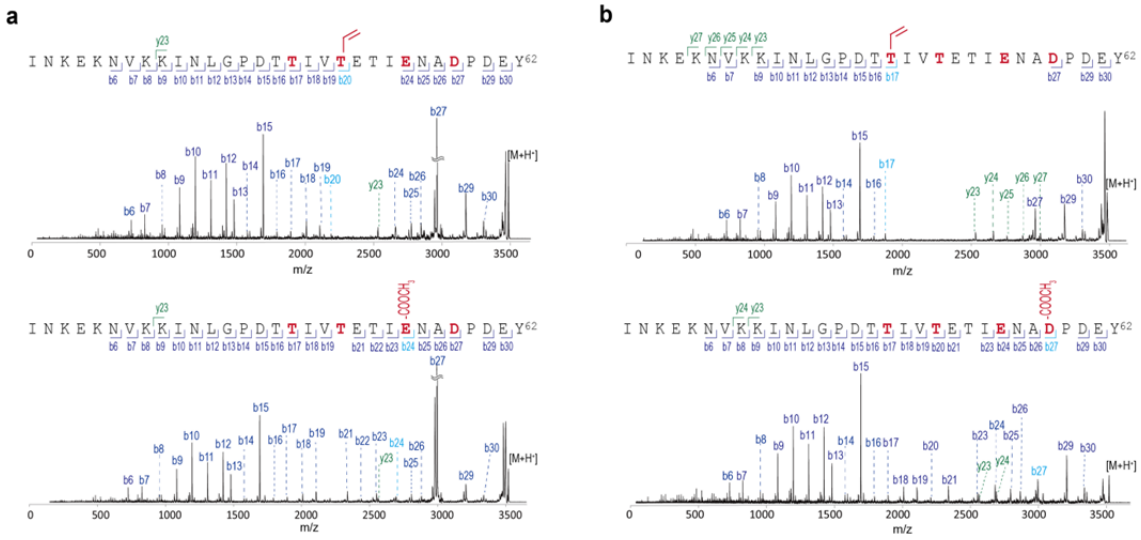


Figure 14. The order of the two macrocyclizations. (a) HPLC chromatogram of the core peptide of mTgnA-R1 after partial hydrolysis. Four different peptides were identified in separate peaks. (b) The reaction of TgnA-R1 and TgnB was quenched by EDTA (100 mM) at different time points and the core peptides were analyzed by HPLC after chymotrypsin cleavage. Arrows indicate the change in amounts of each core peptide over time. (c) A model of the reaction order. The inner ester bond is formed before the outer ester bond.



c

Peptide	AxxTxxxExxD	TxxxAxxExxD	TxxTxxxAxxD	TxxTxxxExxA
Reaction	(7934) (7952) -H2O SM	(7934) (7952) -H2O SM	(7906) -H2O	(7920) -H2O
Single-dehydrated product	INKEKNVKKINLGPDTAIVTETIENADPDEY	INKEKNVKKINLGPDTTIVAETIENADPDEY	INKEKNVKKINLGPDTTIVTETIANADPDEY	INKEKNVKKINLGPDTTIVTETIENAAPDEY
Single-dehydrated product MS/MS				
Methanolysis product MS/MS				N/A

Figure 15. MALDI-TOF-MS/MS spectra of the mTgnA-R1 hydrolysis products (a,b) and the conserved residues mutational study (c). (A) MALDI-TOF-MS/MS spectra of the partially hydrolyzed core peptide of mTgnA-R1 that was eluted at 30.5 min in HPLC analysis (top) and its methanolysis product (bottom). This partial hydrolysis product has the inner ester bond (T51-E55). (B) MALDI-TOF-MS/MS spectrum of the partially hydrolyzed mTgnA-R1 core peptide that was eluted at 36.5 min in HPLC analysis (top) and its methanolysis product (bottom). This partial hydrolysis product has the outer ester bond (T48-D58). The amino acids making ester bonds in mTgnA-R1 are colored red. Light blue b-series ions indicate the modifications. (C) Mutated TgnA peptides, in which each of the highly-conserved residues (Thr, Glu, or Asp) in TgnA-R1 is mutated to alanine, were modified by TgnB and showed single-dehydrated products. We determined each single-dehydrated product by MALDI-TOF-MS/MS. Calculated and observed masses for all mass spectra are in Table 10.

a

peptide	expected	observed
mCwnA-R1	LP-S^{21}ITDTT^{26}RLT^{29}ETIE^{33}QID^{36}QDELATLDY^{45} LP-S^{21}IT^{23}DTT^{26}RLTE^{30}TIE^{33}QIDQDELATLDY^{45}	LP-S^{21}ITDTT^{26}RLT^{29}ETIE^{33}QID^{36}QDELATLDY^{45}
mCwnA-R1-E33D	LP-S^{21}ITDTT^{26}RLT^{29}ETID^{33}QID^{36}QDELATLDY^{45} LP-S^{21}IT^{23}DTT^{26}RLTE^{30}TID^{33}QIDQDELATLDY^{45}	LP-S^{21}IT^{23}DTT^{26}RLTE^{30}TID^{33}QIDQDELATLDY^{45}
mCwnA-R1-D36E	LP-S^{21}ITDTT^{26}RLT^{29}ETIE^{33}QIE^{36}QDELATLDY^{45} LP-S^{21}IT^{23}DTT^{26}RLTE^{30}TIE^{33}QIEQDELATLDY^{45}	LP-S^{21}ITDTT^{26}RLT^{29}ETIE^{33}QIE^{36}QDELATLDY^{45} (major) LP-S^{21}IT^{23}DTT^{26}RLTE^{30}TIE^{33}QIEQDELATLDY^{45} (minor)

LP = M(GSSHHHHHHSSGLVPRGSHM)K^{2}PYILNYSESIDIKHVTSK^{20}

b

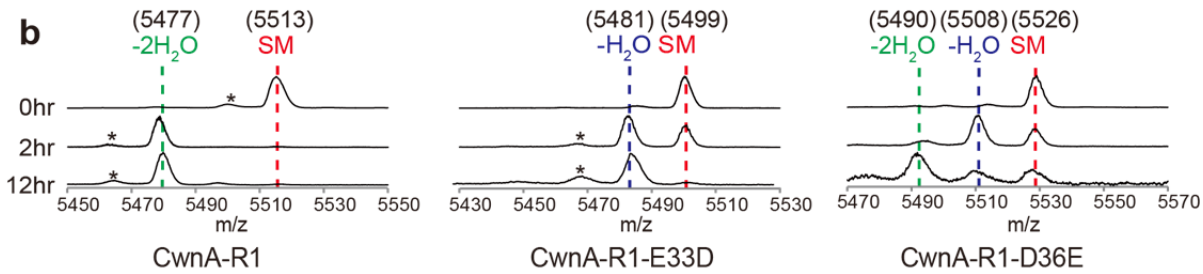


Figure 16. Cross-linking reaction is selective with acidic residues at reactive sites. (a) Expected and observed cross-linking patterns for mCwnA-R1 and its two mutant variations, E33D and D36E. Expected products may be obtained assuming that the reaction does not have the preference for acidic residues at reactive site or residues that surround it. Residues that make ester bonds are bolded and the mutated residues are colored in red. In the minor product of mCwnA-R1-D36E, the line above T^{26}RLTE^{30}TIE^{33} indicates that there are two ester bonds inside the sequence but the exact connectivity is unknown. (b) MALDI spectra of CwnA-R1 and its mutants (5 μM) before or after the *in vitro* reaction with CwnB (0.5 μM) in the presence of ATP (5 mM) and $MgCl_2$ (10 mM). Peaks for the starting material (SM), singly and doubly dehydrated peptides ($-H_2O$ and $-2H_2O$) are presented with red, blue, and green dotted lines, respectively, and the numbers in parentheses are monoisotopic masses. The asterisk (*) indicates a peak resulting from laser-induce deamination which is not related to the reaction.

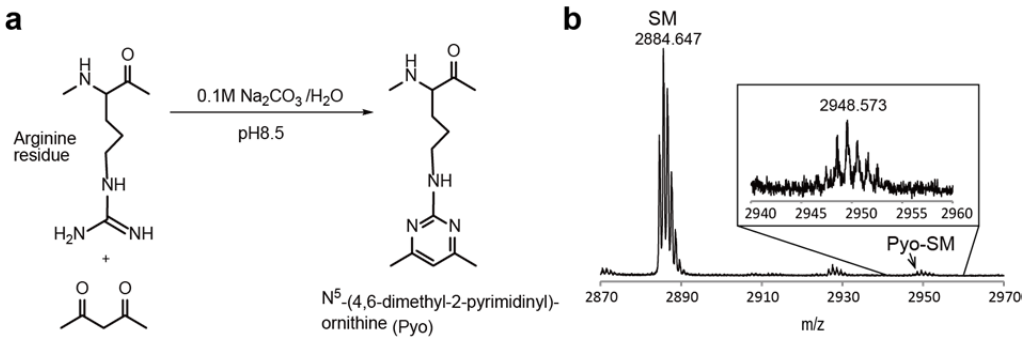


Figure 17. Modification of the arginine residue. (a) Scheme of modifying the guanidino group of arginine into N5-(4,6-dimethyl-2-pyrimidinyl)ornithine (Pyo). (b) MALDI spectrum of mCwnA-R1-E33D core peptide after arginine modification reaction for seven days. Observed monoisotopic masses of the core peptide (SM) and its modified form (Pyo-SM) are 2884.647 and 2948.573, respectively.

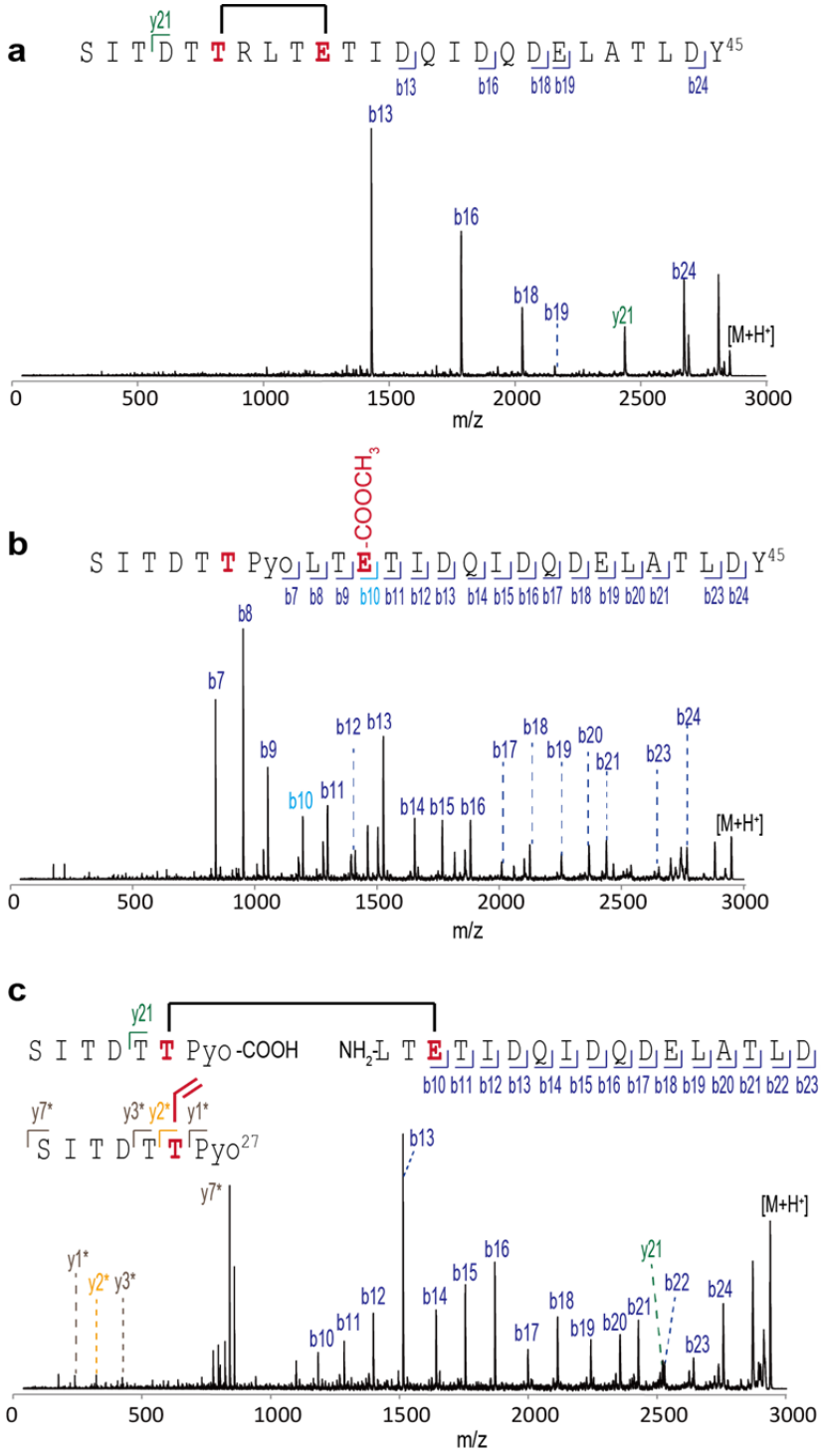

a
S I T D T T R L T E T I D Q I D Q D E L A T L D Y 45
y21
b13 b16 b18 b19 b24
b13
b16
b18
b19
y21
b24
[M+H+]
0 500 1000 1500 2000 2500 3000
m/z
b
-COOCH3
S I T D T T Pyo L T E T I D Q I D Q D E L A T L D Y 45
b7 b8 b9 b10 b11 b12 b13 b14 b15 b16 b17 b18 b19 b20 b21 b23 b24
b8
b7
b9
b12 b13
b10 b11
b14 b15 b16
b17
b18
b19
b20
b21
b23
b24
[M+H+]
0 500 1000 1500 2000 2500 3000
m/z
c
y21
S I T D T T Pyo-COOH NH2- T E T I D Q I D Q D E L A T L D Y 45
b10 b11 b12 b13 b14 b15 b16 b17 b18 b19 b20 b21 b22 b23 b24
y7* y3* y2* y1*
S I T D T T Pyo 27
y7*
b13
[M+H+]
y1*
y2* y3*
b16
b15
y21
b12
b14
b18
b22
b24
b11
b21
b10
b17
b19
b20
b23
0 500 1000 1500 2000 2500 3000
m/z

Figure 18. MALDI-TOF-MS/MS spectra of the core peptide of mCwnA-R1-E33D. (a) MALDI-TOF-MS/MS spectrum of the core peptide of mCwnA-R1-E33D. (b) MALDI-TOF-MS/MS spectrum of the core peptide of mCwnA-R1-E33D after methanolysis and arginine modification. Pyo is the modified arginine by acetylacetone. (c) MALDI-TOF-MS/MS spectrum of the mCwnA-R1-E33D core peptide after trypsin cleavage and arginine modification. The y-series ions with asterisk are fragments from SITDTTPyo[27] formed by the MS/MS-induced rearrangement. The amino acids making ester bonds in mCwnA-R1-E33D are colored red. Light blue b-series ions and orange y-series ions indicate the modifications. Calculated and observed masses for all mass spectra are in Table 11.

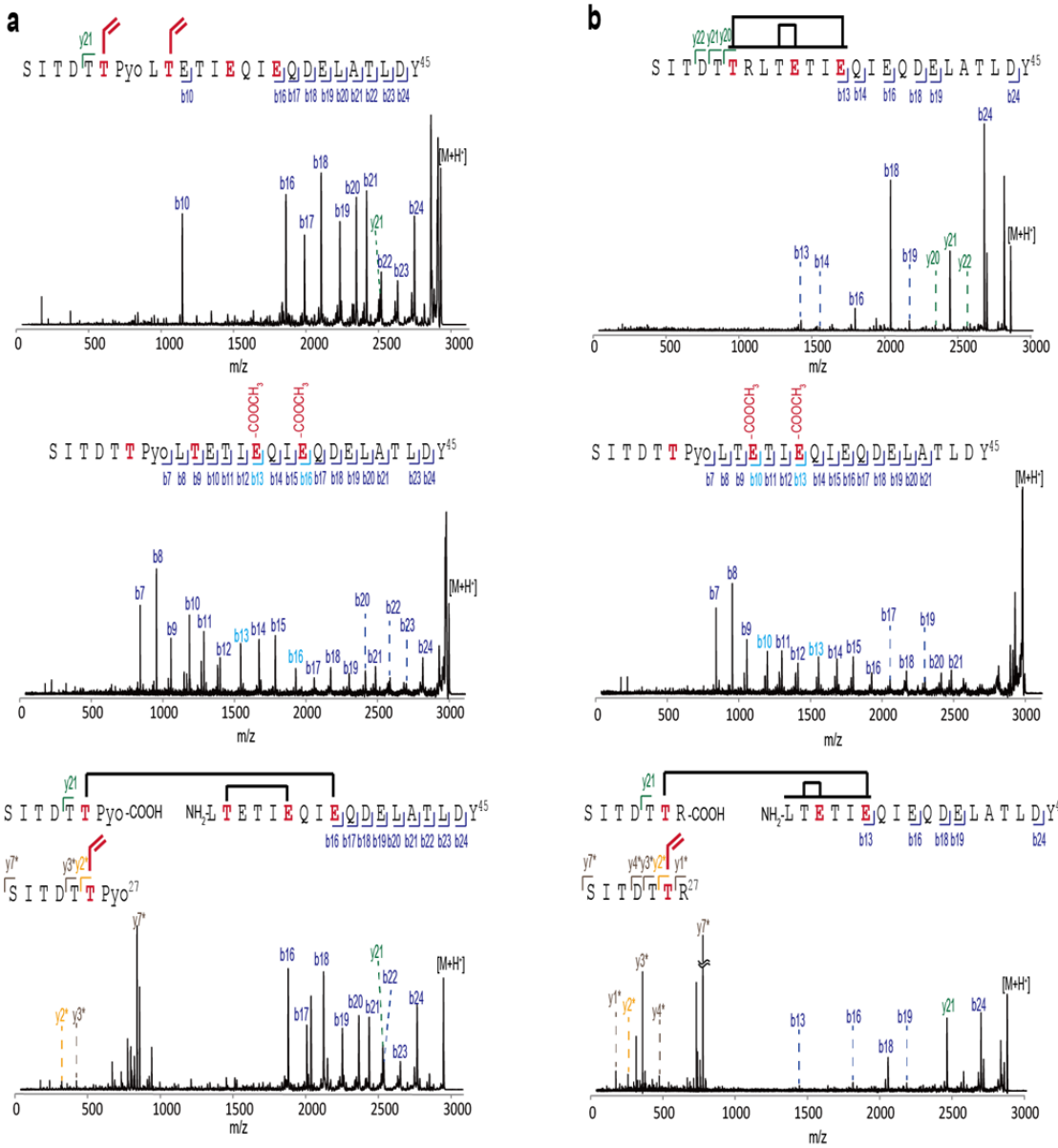

a
b
m/z
[M+H+]
-COOCH3
-COOCH3
NH2-L T E T I E Q I E Q D E L A T L D Y45
S I T D T T Pyo-COOH
S I T D T T R-COOH

Figure 19. MALDI-TOF-MS/MS spectra of the core peptides of the modified CwnA-R1-D36E. (a) The major and (b) the minor products of the modification reaction with CwnA-R1-D36E were analyzed by MALDI-TOF-MS/MS. The three panels present the spectra of the modified core peptides (upper), the modified core peptides with complete methanolysis (middle), and the modified core peptides after trypsin cleavage (lower). Pyo is the modified arginine by acetylacetone. The y-series ions with asterisk are fragments from SITDTT(Pyo/R)[27] formed by the MS/MS-induced rearrangement. The amino acids identified to make ester bonds are colored red. Light blue b-series ions and orange y-series ions indicate the modifications. Calculated and observed masses for all mass spectra are in Table 12.

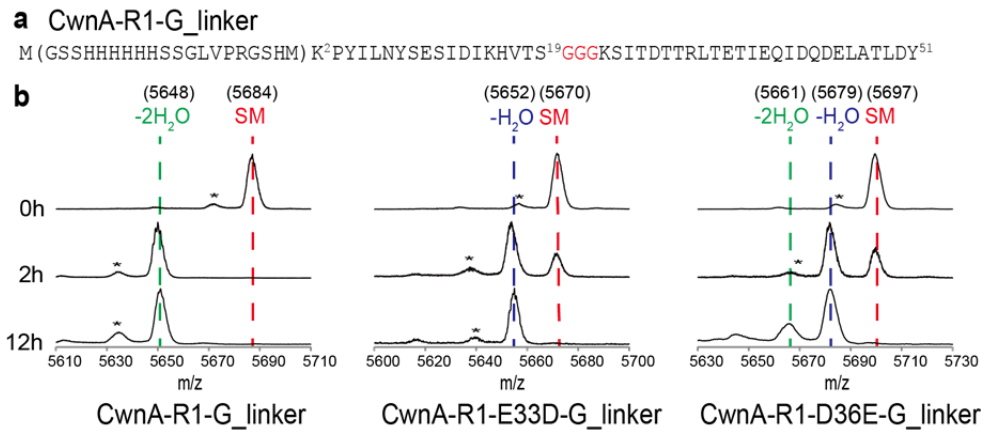


Figure 20. MALDI analysis of the enzyme reactions of CwnA-R1-G_linker and its mutants. (a) The peptide sequence of CwnA-R1-G_linker in which the GGG linker is inserted between S19 and K20 of CwnA-R1. (b) MALDI spectra of CwnA-R1-G_linker and its mutants after enzyme reactions. The reaction condition is the same as the one in Figure 9e. The reactions were quenched by EDTA (100 mM final) at three time points. Red, blue, and green dotted lines indicate the peaks for SM, the singly-dehydrated peptide, and the doubly-dehydrated peptide. The numbers in parentheses are monoisotopic masses. The asterisk indicates a peak resulting from the laser-induced deamination which is not related to the reaction. Compared to the CwnA-R1-G_linker, CwnA-R1-E33D-G_linker and CwnA-R1-D36E-G_linker showed incomplete reactions and these results are the same as those of precursor peptides without the G_linker.

a

Bacteria strain	ATP-grasp enzyme Accession number	Precursor peptide Accession number	Precursor peptide sequence
Stigmatella aurantiaca	SEM66409.1	SEM66380.1	MKRENQEQKAQRTQQLFIQDLARIEGKKSAASKVSASGPFTTYALAEESGA
Myxococcus fulvus 124B02	AKF84605.1	AKF84606.1	MTRKHMHPRRELFITELGATPKPPGTGKPGGPKPMTTMAVGEESKPPDSGSRD
Myxococcus macrosporus DSM 14697	ATB44709.1	ATB44710.1	MPENAPPRPPPALFIQDWGPTPPEPSREETGQVTTMQVGEETGDGGELQR
Plesiocystis pacifica SIR-1	EDM79249.1	EDM79248.1	MSNDKKNNELKELFIEDLGQVTGGCRRPHPTPVGEGPATTLALGEEEPTTLSLQAEEPTTLSLQAEEPTTLSLSAED
Aeribacillus pallidus	ASS89204.1	ASS89203.1	MKKKLFIEELEVRQSSSPLFVSPIVDPGYATTLALGEEGGYTTQAIGEEGGPIVTTYALGEEGGFIK
Stigmatella aurantiaca (strain DW4/3-1)	ADO75084.1	ADO75085.1	MEKSADPYLRGMGPRDSAFREVVMKRENQEHTVKRIKALFVQDLARIEGKQGVASTESASGPFTTFALAEESGA
Stigmatella erecta	SET73852.1	SET73831.1	MKRENQEQKTQRTQQLFIQDLARIEGKKSAASTVSASGPFTTYALAEESGA
Myxococcus fulvus	SES77654.1	SES77688.1	MTRKDIHRRRELFITELGERPKPPGTGKPGGPKPITTMAVGEESKPPDSGTRD
Myxococcus hansupus	AKQ69827.1	AKQ69826.1	MLFIRDWAPTSAGPRPQQTDPVTTMKVGEESDSGDDLQR
Nannocystis exedens	SFF22126.1	SFF22140.1	MKKNDKKSKKQLKDLFIEDLGQVTGGKGPIYTTQAIGEECPEPTTLAIGEEGPDVTTLALGEEGPEATTKALGEE
Geobacillus sp. 8	KZM56108.1	KZM56107.1	MKKKLFIEELEVRQSSSPLFVSPIVDPGYATTLALGEEGGYTTQAIGEEGGPIVTTYALGEEGGFIK
Archangium gephyra	AKJ02156.1	AKJ02155.1	MEPKKEKKVKKIKELFVQDLARIEGGAGPVQDKKCPDITTYACGEEPPGCLAC
Kyrpidia tusciae (strain DSM 2912 / NBRC 15312 / T2) (Bacillus tusciae)	ADG06746.1	ADG06747.1	MAKKRLFIEQLEERQSTSPLFVTPVDPSLIMTTMMVGEEGGPIHPPITTQAIGEEGGITMAFGEEGMHFTSMLYGEEGGHFIPLG
Paenibacillaceae bacterium JTherm	PDO09934.1	PDO09933.1	MVSKRLFIEELEEVPSSSPLFAAGAGTDAVATTLAVGEECGPVYTTLAIGEESGPIYTTMAIGEEGCWPKR

b

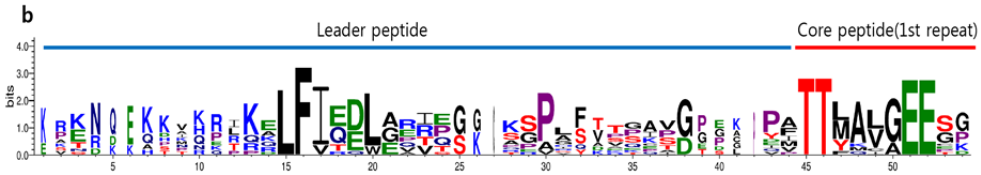


Figure 21. Sequence analysis of the precursor peptides in the plesiocin group. (a) A list of bacterial strains, the accession numbers for the ATP-grasp enzyme and the precursor peptide, and the sequence of the precursor peptide in the plesiocin group. The highly-conserved sequence motifs in the leader peptide and the first repeat of the core peptide are colored blue and red, respectively. (b) Sequence logo of the precursor peptides in the plesiocin group. Blue and red lines above the sequence logo indicate the leader peptide and the first repeat of the core peptide, respectively, in which the LFIx(D/E)L and TTxxxxEE motifs are highly conserved.

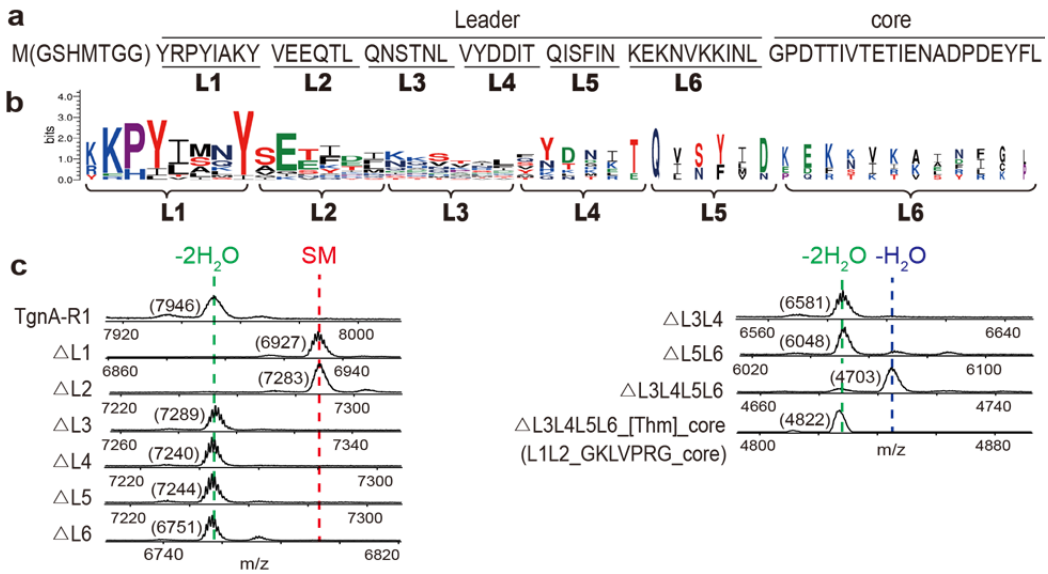


Figure 22. The highly-conserved sequence of the leader peptide is critical to the reaction. (a) The peptide sequence of TgnA-R1. The leader peptide is divided into six parts, L1-L6. (b) The sequence logo of the leader peptides of TgnA and its fifteen homologs. Highly conserved sequences are in L1 and L2. Many TgnA homologs do not have L4-L6, and thus, the sequence logo at L4-L6 is sparse. (c) MALDI spectra of TgnA-R1 and its truncated variants (5 μM) after the *in vitro* reaction with TgnB (0.5 μM), ATP (5 mM), and $MgCl_2$ (10 mM) for 12 hours. Peaks for the SM, singly and doubly dehydrated peptides ($-H_2O$ and $-2H_2O$) are presented with red, blue, and green dotted lines, respectively, and the numbers in parentheses are monoisotopic masses. ⊿Ln (n = 1-6) indicates a peptide where Ln is truncated. ⊿L3L4L5L6_[Thm]_core (L1L2_GKLVPRG_core) indicates a peptide where L3-L6 is truncated and the thrombin cleavage site ([Thm]), GKLVPRG, is inserted between L1L2 and the core peptide.

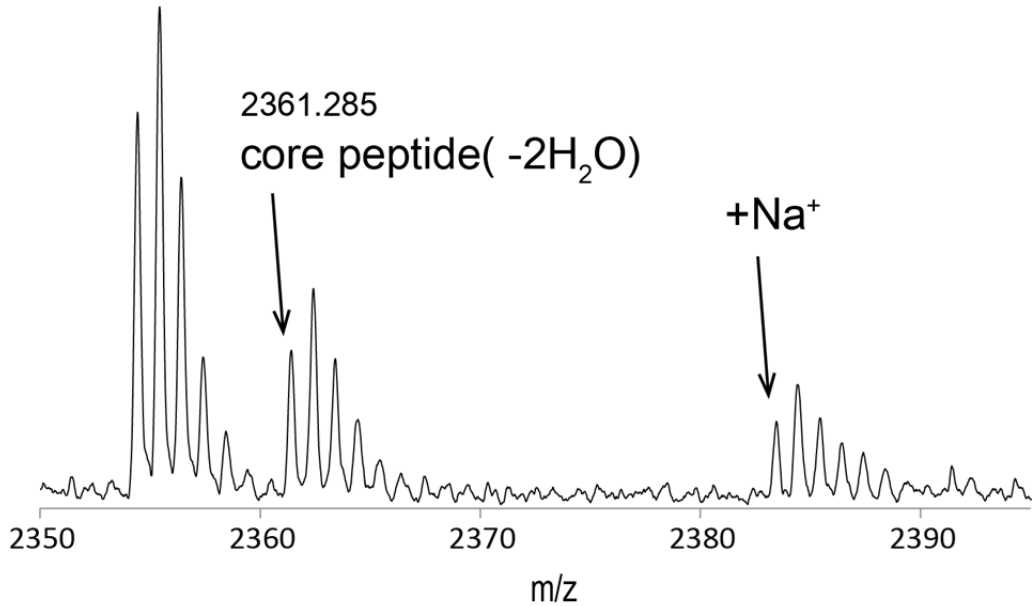


Figure 23. MALDI spectrum of the core peptide of the modified △L3L4L5L6_[Thm]_core after thrombin cleavage. The observed monoisotopic mass of the core peptide is 2361.285 that indicates loss of two water molecules.

7. Tables

Table 1. Optimized sequences of the synthesized genes for heterologous protein expression. Restriction sites are underlined.

strain	Protein name	Accession number	Restriction sites	Gene sequence (5' to 3')
Bacillus thuringiensis serovar huazhongensis BGSC 4BD1	TgnA	EEM80215.1	5' *Nde*I 3' *Not*I	CATATGACCGGTGGCTATCGCCCGTATATTGCGAAATATGTGG AAGAACAGACCCTGCAGAATTCTACCAATCTGGTGTATGATGA TATTACCCAGATTTCATTTATTAATAAAGAAAAAAATGTTAAAAA AATTAATCTGGGTCCGGATACCACCATTGTGACCGAAACCATT GAAAATGCCGATCCGGATGAATATTTTCTGGATCCGGATACCA CCATTCATACCTTTACCGTTGAAAATAATGATACCGATGAACAT TATTGCGGCCCGGAAACCACCCGTATTACCAAAACCCTGGAA AATATTGATCCGGATGAATATTATAGCTAAGCGGCCGC
	TgnB	EEM80217.1*	5' *Nde*I 3' *Not*I	CATATGAAAACCATTCTGATTATTACCAATACCCTGGATCTGAC CGTGGATTATATTATTAATCGCTATAATCATACCGCTAAATTTTT TCGTCTGAATACCGATCGTTTTTTTGATTATGATATTAATATTAC CAATAGCGGTACCAGCATTCGTAATCGTAAATCTAATCTGATTA TTAATATTCAGGAAATTCATAGCCTGTATTATCGCAAAATTACC CTGCCGAATCTGGATGGCTATGAAAGTAAATATTGGACCCTGA TGCAGCGCGAAATGATGAGTATTGTTGAAGGCATTGCAGAAA CCGCTGGCAATTTTGCACTGACCCGTCCGTCTGTGCTGCGC AAAGCTGATAATAAAATTGTGCAGATGAAACTGGCAGAAGAAA TTGGTTTTATTCTGCCGCAGAGTCTGATTACCAATTCAAATCA GGCGGCAGCCTCATTTTGCAATAAAAATAATACCAGCATTGTG AAACCGCTGAGTACCGGCCGCATTCTGGGTAAAAATAAAATT GGCATTATTCAGACCAATCTGGTTGAAACCCATGAAAATATTC AGGGCCTGGAACTGTCTCCGGCTTATTTTCAGGATTATATTCC GAAAGATACCGAAATTCGTCTGACCATTGTTGGTAATAAACTG TTTGGCGCCAATATTAAATCAACCAATCAGGTTGATTGGCGCA AAAATGATGCACTGCTGGAATATAAACCGGCCAATATTCCGGA TAAAATTGCCAAAATGTGTCTGGAAATGATGGAAAAACTGGAA ATTAATTTTGCGGCGTTTGATTTTATTATTCGTAATGGTGATTAT ATTTTTCTGGAACTGAATGCCAATGGTCAGTGGCTGTGGCTG GAAGATATTCTGAAATTTGATATTTCAAATACCATTATTAATTATC TGCTGGGTGAACCGATTTAAGCGGCCGC
Colwellia sp. UCD-KL20	CwnA	WP_076419416.1	5' *Nde*I 3' *Not*I	CATATGAAGCCGTATATTCTGAATTACAGTGAAAGTATTGACAT CAAACACGTTACCAGTAAAAGCATTACCGATACCACCCGCCT GACCGAAACCATTGAACAGATTGATCAGGATGAACTGGCCAC CCTGGATTATACCAGTAAAACCGAAAGTATTGAACCGAGTGAT AATGATGAACTGTTTACCAATAGTACCACCCAGACCTTTGTTA TTGAAGATAGTGATGAAGATCAGGTTGGCAGTGAAATTCAGT GCAGTACCCTGGTTACCAAAACCAGTGAACCGAGTGACAATG ATGAAATTCTGACCTGCAGCACCTGGGCCACCGAAAGCACC GAACCGAGTGATAACGATGAAATTTTTAGCTTTAGCACCCTGC AGACCCGTACCATTGAACCGTCAGATGATGATAGTATTAATCA GCTGAGTACCCTGCAGACAAATACCAGTGAAAATTGGGATCA TGATGAAATTATCCTGAATTAAGCGGCCGC
	CwnB	WP_076419414.1	5' *Bam*HI 3' *Bgl*II	GGATCCGCTGGTGCCTCGTGGCAGTGATAAAATTCTGATGAT TACCAGCAGTTATGATAAAACCTGCGATTATCTGATTAAGAAAT ACAATCACGTGAACTTTTTCCGCCTGGATGTGGATTGTTTTAG TAAATATAGTATCAGCTTCGCAGATAATCAGTTTGTGATTAGCG ATGAAGAAAGCAGTCTGCTGGAAGAAAATTGCTGTAGTATTTA TTACCGTAAGCCGAGTCCGGAAAATCTGAATAATATTATTGAC AACAAGTACCAGAGCTTTGCACATAAAGAAAGCTATAGTCTGA TTGAAGGTATTGTGGAAGCCTATAGCGGTATTTGTCTGAGTAA

				ACCGAGCATTATGCGCCCGGCAGGTAATAAAGTTCTGCAGGC ACGCATTGCCAATGTGGTTGGCTTTGATCTGCCGAAATATCT GATTACCAATGATGCAGAAAAGATTCATGAACTGGCACCGGA AAAAGCCATTGTGAAACCGCTGAGCGTTGGTGTTATTCAGGA TGAAAAAAGCAAAGAATTCGTTCAGACCAATCTGCTGAATAAA AATATTGATCTGAAGGCACTGAAGTATAGTCCGGCCTATTTTC AGAAATATCAGAAAAAAGACTACGAGTGCCGTGTGACCTTTG TGGGCAAAAAAGCCTTTCCGGTTAAAATTGAAAGCGAAAATG ATATTGACTGGCGCAAACAGGATAATAATATTATCTATAGCGTG TGCGAAATCCCGCATGATGTTTATTGTAAATGCCTGCTGTTTA TGGATAAATGCAAAATGAATTTCGGCTGCTTTGATTTTATCATT CATGATAATTACTGGTACTTCCTGGAAATGAATGTTAATGGCC AGTGGGCCTGGCTGGAATTTGAAGTGGGTATTGATATTAGCG GTGAAATTATGCGCTATCTGTGTGATAAATAAAGATCT
Dehalobacter sp. TeCB1	DbnA	OCZ49919.1	5' *Nde*I 3' *Not*I	CATATGCGTGAACCGTATCTGGCCAATTATGTTGAAGAACATG ATTTTGATCACAACAATCGCCTGTTTTATGATGGCCGCACCCA GCTGAGTTATGTTGATCCGGAAAAAAAAGTTCGTGCAATTGAT CGTTGCCTGGTGGGCCCGACCTATATTACCAAAGCCGCAGAA AGCAGCGATCCGGATGAATTTCAGGCCATTGGCCCGACCCG TCTGACCGAAACCATTGAAAATAGTGATCCGGATGAGTTTTTT GCCCTGAGCCCGACCCAGAAAACCGCAACCGATGAAAGCG CCGATCCGGATGAATTCTATGCAATTGGTCCGACCCGCGAAA CCTTTAGCAATGAAGGTAGCGATCCGGACGAATTTTATCTGCT GGGCCCGAGCCGCGAAACCCGCTCAAATGAATGTAGCGATC CGGATGAATTTTATTTACGCTAAGCGGCCGC
	DbnB	OCZ49917.1	5' *Bam*HI 3' *Bgl*II	GGATCCGCTGGTGCCTCGTGGCAGTAAACAGATTCTGGTTAT TACCAGTAGCTATGATCTGACCGTGGATTATATTATTAACAAAT TCCAGGCACTGAACTTTTTTCGCCTGAATACCGATCGTCTGG ATGAATATCAGATTAGCTTTTATAAGAGCAAGTGGGAAATTACC GATAGCTTTAAAAGCATTACCAATAATGACATCAAGAGTATTTA CTACCGTAAACCGATTCTGCCGCATCTGGATCAGTTTAATCCG CAGTATCATAATTATATGCACAAAGAAATCCTGACCCTGATTGA AGGCATTATTGAACAGTTTGATGGCAATTGCCTGAGTAAACC GAGTATTCTGAAACGTGCCGATAATAAAATTGTGCAGCTGCAG CTGGCAGAAAAAATTGGCCTGCAGATGCCGCTGAGTCTGATT ACCAATAGTAATCATAGTGCAAAAATCTTCTGCAGTGATAAAAA AACCATCGTTAAACCGATTAGTGTTGGCAAACTGCTGTATGAA AATAAAGTGGGTATTATTCAGACCAATTACGTGGACCCTGCCT TTGGTTTTGAAGGCCTGGAACTGGCACCGAGTTATTTTCAGG AATATATTGATAAGGACTTCGAACTGCGTGTGACCGTTGTTCA TGATGAATTTTTTCCGGTGAAAATTGAAACCACCGAAAAAGTT GATTGGCGTAAAGCCACCAATGGTAATATTCTGTATAGCAAAG ATATCCTGCCGGATGATATTAAAAGCCTGAGTCTGCAGATGAT GCGTCTGCTGAATCTGAAATATGCAGCCTTTGATTTTATCGTG AGCAAAGGTCGCTATGTTTTTCTGGAACTGAATGCAAATGGT CAGTGGGCCTGGCTGGAACAGGAACTGAATCTGAATATTAGC GATAGCGTTATTCGCTATCTGGAACGTTAAAGATCT

Table 2. Plasmids used in this study.

Plasmid	vector	Expressed Protein
pRHJ1	pET28b	6His_[Thm]_TgnA
pRHJ2	pET28b	6His_[Thm]_TgnB
pRHJ3	pET22b	6His_[Thm]_CwnA
pRHJ4	pCDF Duet-1	6His_[Thm]_CwnB
pRHJ5	pET22b	6His_[Thm]_DbnA
pRHJ6	pCDF Duet-1	6His_[Thm]_DbnB
pRHJ7	pET28b	6His_[Thm]_TgnA & 6His_[Thm]_TgnB
pRHJ8	pET28b	6His_[Thm]_TgnA-R1
pRHJ9	pET22b	6His_[Thm]_CwnA-R1
pRHJ10	pET22b	6His_[Thm]_DbnA-R1
pRHJ11	pET28b	6His_[Thm]_TgnA-R1-D60A
pRHJ12	pET22b	6His_[Thm]_CwnA-R1-E33D
pRHJ13	pET22b	6His_[Thm]_CwnA-R1-D36E
pRHJ14	pET22b	6His_[Thm]_CwnA-R1-G_linker
pRHJ15	pET22b	6His_[Thm]_CwnA-R1-E33D-G_linker
pRHJ16	pET22b	6His_[Thm]_CwnA-R1-D36E-G_linker
pRHJ17	pET28b	6His_[Thm]_ΔL1
pRHJ18	pET28b	6His_[Thm]_ΔL2
pRHJ19	pET28b	6His_[Thm]_ΔL3
pRHJ20	pET28b	6His_[Thm]_ΔL4
pRHJ21	pET28b	6His_[Thm]_ΔL5
pRHJ22	pET28b	6His_[Thm]_ΔL6
pRHJ23	pET28b	6His_[Thm]_ΔL3L4
pRHJ24	pET28b	6His_[Thm]_ΔL5L6
pRHJ25	pET28b	6His_[Thm]_ΔL3L4L5L6
pRHJ26	pET28b	6His_[Thm]_ΔL3L4L5L6_Thr cleavage site_core
pRHJ27	pET28b	6His_[Thm]_MBP_[Tev]_ΔL3L4L5L6_Thr cleavage site_core
pSK387	pET28b	6His_[Thm]_MBP_[Tev]_*lysC_CS*(D18-I58)

[Thm]: Thrombin cleavage site, [Tev]: Tev protease cleavage site

Table 3. Oligonucleotide primers used for producing each plasmid.

Each forward (FW) and reverse (RV) primers were used for constructing each plasmid.

Plasmid	Primer names	Sequence(5' to 3')
pRHJ7	TgnA_TgnB coexpression_FW	GAAAATATTGATCCGGATGAATATTATAGCTAATCTAGAAATA ATTTTGTTTAACTTTAAGAA
	TgnA_TgnB coexpression_RV	GCTAGTTATTGCTCAGCGG (T7 terminator)
pRHJ8	TgnA-R1-FW	TAAGCGGCCGCACTCGAGCACC
	TgnA-R1-RV	CAGAAAATATTCATCCGGATCGGC
pRHJ9	CwnA-R1-FW	Same as TgnA-R1-FW
	CwnA-R1-RV	ATAATCCAGGGTGGCCAGTTCATC
pRHJ10	DbnA-R1-FW	Same as TgnA-R1-FW
	DbnA-R1-RV	AATGGCCTGAAATTCATCCGGATC
pRHJ11	TgnA-R1-D60A-FW	GAATATTTTCTGTAAGCGGCCGCA
	TgnA-R1-D60A-RV	GGCCGGATCGGCATTTTCAATGGTTTCGGTCAC
pRHJ12	CwnA-R1-E33D-FW	GATCAGATTGATCAGGATGAACTGGCC
	CwnA-R1-E33D-RV	AATGGTTTCGGTCAGGCGGGT
pRHJ13	CwnA-R1-D36E-FW	GAACAGGATGAACTGGCCACCCTG
	CwnA-R1-D36E-RV	AATCTGTTCAATGGTTTCGGTCAG
pRHJ14-16	CwnA-R1-G_linker-FW	GGCGGTGGCAAAAGCATTACCGATACCACCCGC
	CwnA-R1-G_linker-RV	ACTGGTAACGTGTTTGATGTCAAT
pRHJ17	ΔL1-FW	GTGGAAGAACAGACCCTGCAGAAT
	ΔL1-RV	GCCACCGGTCATATGGCTGCC
pRHJ18	ΔL2-FW	CAGAATTCTACCAATCTGGTGTATGAT
	ΔL2-RV	ATATTTCGCAATATACGGGCGATAGCC
pRHJ19	ΔL3-FW	GTGTATGATGATATTACCCAGATTTCATTT
	ΔL3-RV	CAGGGTCTGTTCTTCCACATATTT
pRHJ20	ΔL4-FW	CAGATTTCATTTATTAATAAAGAAAAAAATGTT
	ΔL4-RV	CAGATTGGTAGAATTCTGCAGGGT
pRHJ21	ΔL5-FW	AAAGAAAAAAATGTTAAAAAAATTAATCTGGGT
	ΔL5-RV	GGTAATATCATCATACACCAGATTGGTAGA
pRHJ22	ΔL6-FW	GGTCCGGATACCACCATTGTGACC
	ΔL6-RV	ATTAATAAATGAAATCTGGGTAATATCATC
pRHJ23	ΔL3L4-FW	Same as ΔL4-FW
	ΔL3L4-RV	Same as ΔL3-RV
pRHJ24	ΔL5L6-FW	Same as ΔL6-FW
	ΔL5L6-RV	Same as ΔL5-RV
pRHJ25	ΔL3L4L5L6-FW	Same as ΔL6-FW
	ΔL3L4L5L6-RV	Same as ΔL3-RV
pRHJ26	ΔL3L4L5L6_[Thm]_core-FW1	Same as ΔL6-FW
	ΔL3L4L5L6_[Thm]_core-RV1	CTGGAAATACAGGTTTTCCAGATTAATTTTTTTAACATTTTTT TCTTTATT
pRHJ27	MBP fusion-FW	TCCGGAGAAAACCTGTATTTCCAGGGCTATCGCCCGTATAT TGCG
	MBP fusion-RV	Same as TgnA_TgnB coexpression_RV

Table 4. Predicted and observed masses for various ions from MALDI-TOF-MS spectra.△is observed mass (Obs.) - calculated mass (Calc.).

Figure	construct	Calc.	Obs.	△
1d	TgnA	14610.83	14610.98	0.15
1d	mTgnA	14502.76	14502.85	0.09
4d	R1	3465.775	3465.53	-0.245
4d	R2	2606.132	2606.251	0.119
4d	R3	2123.03	2123.025	-0.005
5a	R1+H_2O+H^+	3483.786	3483.747	-0.039
5a	R1+2H_2O+H^+	3501.796	3501.752	-0.044
5a	R2-1+H^+	1306.632	1306.568	-0.064
5a	R2-2+H^+	1336.529	1336.621	0.092
5a	R3+H_2O+H^+	2141.04	2140.968	-0.072
5a	R3+2H_2O+H^+	2159.051	2159.008	-0.043
11a	mTgnA-R1 core peptide(-2H_2O)	3465.775	3465.996	0.221
11c	mCwnA-R1 core peptide(-2H_2O)	2848.374	2848.677	0.303
11e	mDbnA-R1 core peptide(-2H_2O)	1890.876	1891.11	0.234
11g	mTgnA-R1-D60A core peptide(-2H_2O)	3421.785	3421.632	-0.153
17b	SM	2884.395	2884.647	0.252
17b	Pyo-SM	2948.426	2948.573	0.147
23	core peptide(-2H_2O)	2361.077	2361.285	0.208

Table 5. Calculated and observed masses for various ions in MALDI-TOF-MS/MS spectra of Figure 5. △is observed mass (Obs.) - calculated mass (Calc.).

Figure 5b R1				Figure 5c R2-1			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b4	485.272	485.232	-0.04	b3	376.187	376.122	-0.065
b6	727.41	727.317	-0.093	b5	588.266	588.174	-0.092
b7	826.478	826.405	-0.073	b6	689.314	689.219	-0.095
b8	954.573	954.483	-0.09	b7	790.362	790.242	-0.12
b9	1082.668	1082.618	-0.05	b8	903.446	903.289	-0.157
b10	1195.752	1195.739	-0.013	b9	1040.505	1040.408	-0.097
b11	1309.795	1309.772	-0.023	b10	1141.552	1141.46	-0.092
b12	1422.879	1422.856	-0.023	b11	1288.621	1288.48	-0.141
b13	1479.901	1479.801	-0.1	y1	166.086	166.178	0.092
b14	1576.953	1576.946	-0.007	y3	404.193	404.115	-0.078
b15	1691.98	1691.808	-0.172	y4	517.277	517.171	-0.106
b16	1793.028	1792.872	-0.156	y5	618.325	618.204	-0.121
b17	1894.076	1893.913	-0.163	y6	719.372	719.254	-0.118
b18	2007.16	2007	-0.16	y7	834.399	834.293	-0.106
b19	2106.228	2106.2	-0.028	y8	931.452	931.331	-0.121
b20	2207.276	2207.091	-0.185	y9	1046.479	1046.4	-0.079
b21	2336.318	2336.284	-0.034	y10	1159.563	1159.521	-0.042
b23	2550.45	2550.352	-0.098	[M+H]+	1306.631	1306.585	-0.046
b24	2679.493	2679.38	-0.113				
b25	2793.536	2793.375	-0.161				
b26	2864.573	2864.471	-0.102				
b27	2979.6	2979.486	-0.114				
b29	3191.679	3191.585	-0.094				
b30	3320.722	3320.428	-0.294				
[M+H]+	3501.796	3502.633	0.837				

Figure 5d R2-2				Figure 5e R3			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b2	201.123	201.182	0.059	b3	359.204	359.205	0.001
b3	330.166	330.164	-0.002	b4	472.288	472.309	0.021
y2	319.14	319.224	0.084	b5	573.336	573.374	0.038
y3	448.183	448.25	0.067	b6	701.43	701.357	-0.073
y4	563.21	563.275	0.065	b8	915.562	915.448	-0.114
y5	664.257	664.33	0.073	b9	1044.605	1044.353	-0.252

y6	779.284	779.377	0.093	b10	1158.648	1158.2	-0.448
y7	893.327	893.405	0.078	b11	1271.732	1271.767	0.035
[M+H]+	1336.529	1336.529	0	b12	1386.759	1386.692	-0.067
				b14	1598.838	1598.821	-0.017
				b15	1727.881	1727.794	-0.087
				b17	2054.008	2054.033	0.025
				a11	1243.737	1243.727	-0.01
				a17	2026.013	2026.021	0.008
				a18	2113.045	2113.036	-0.009
				y16	1956.955	1956.958	0.003
				[M+H]+	2159.05	2159.049	-0.001

Table 6. Calculated and observed masses for various ions in MALDI-TOF-MS/MS spectra of Figure 6. △is observed mass (Obs.) − calculated mass (Calc.).

Figure 6b pep1				Figure 6b pep2'				Figure 6b pep3'			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b4	485.272	485.232	-0.04	b4	485.272	485.226	-0.046	b4	485.272	485.159	-0.113
b6	727.41	727.391	-0.019	b6	727.41	727.356	-0.054	b5	613.367	613.265	-0.102
b7	826.478	826.421	-0.057	b7	826.478	826.452	-0.026	b6	727.41	727.325	-0.085
b8	954.573	954.383	-0.19	b8	954.573	955.071	0.498	b7	826.478	826.375	-0.103
b9	1082.668	1082.667	-0.001	b9	1082.668	1082.674	0.006	b8	954.573	954.481	-0.092
b10	1195.752	1195.771	0.019	b10	1195.752	1195.768	0.016	b9	1082.668	1082.628	-0.04
b11	1309.795	1309.826	0.031	b11	1309.795	1309.837	0.042	b10	1195.752	1195.703	-0.049
b12	1422.879	1422.916	0.037	b12	1422.879	1422.933	0.054	b11	1309.795	1309.777	-0.018
b13	1479.901	1479.933	0.032	b13	1479.901	1479.937	0.036	b12	1422.879	1422.867	-0.012
b14	1576.953	1576.955	0.002	b14	1576.953	1576.986	0.033	b13	1479.901	1479.894	-0.007
b15	1691.98	1691.995	0.015	b15	1691.98	1692.03	0.05	b14	1576.953	1576.991	0.038
b16	1793.028	1793.028	0	b16	1793.028	1793.06	0.032	b15	1691.98	1691.95	-0.03
b17	1894.076	1894.066	-0.01	b17	1894.076	1894.057	-0.019	b16	1793.028	1792.924	-0.104
b18	2007.16	2007.12	-0.04	b18	2007.16	2007.149	-0.011	b17	1876.065	1876.002	-0.063
b19	2106.228	2106.207	-0.021	b19	2106.228	2106.206	-0.022	b27	2975.605	2975.46	-0.145
b20	2207.276	2207.215	-0.061	b20	2189.265	2189.209	-0.056	b29	3187.684	3187.603	-0.081
b21	2336.318	2336.235	-0.083	b21	2318.307	2318.288	-0.019	b30	3316.727	3316.701	-0.026
b22	2437.366	2437.243	-0.123	b22	2419.355	2419.229	-0.126	y4	523.204	523.142	-0.062
b23	2550.45	2550.336	-0.114	b23	2532.439	2532.369	-0.07	y10	1180.501	1180.57	0.069
b24	2693.509	2693.452	-0.057	b24	2661.482	2661.403	-0.079	y23	2544.235	2543.948	-0.287
b25	2807.552	2807.441	-0.111	b25	2775.525	2775.422	-0.103	y24	2672.33	2672.173	-0.157
b26	2878.589	2878.531	-0.058	b26	2846.562	2846.496	-0.066	y25	2771.399	2771.293	-0.106
b27	3007.632	3007.657	0.025	b27	2975.605	2975.547	-0.058	y26	2885.441	2885.153	-0.288
b28	3104.685	3104.738	0.053	b29	3187.684	3187.796	0.112	y27	3013.536	3013.465	-0.071
b29	3219.711	3219.736	0.025	b30	3316.727	3316.79	0.063	y28	3142.579	3142.434	-0.145
b30	3348.754	3348.852	0.098	b31	3479.79	3479.749	-0.041	$[M+H]^+$	3497.801	3497.801	0

y4	523.204	523.173	-0.031	y4	523.204	523.163	-0.041
y5	652.246	652.203	-0.043	y5	652.246	652.219	-0.027
y23	2576.262	2576.11	-0.152	y10	1180.501	1180.705	0.204
y24	2704.357	2704.254	-0.103	y23	2544.235	2544.052	-0.183
y26	2917.468	2917.411	-0.057	y24	2672.33	2672.227	-0.103
y27	3045.563	3045.581	0.018	y26	2885.441	2885.362	-0.079
$[M+H]^+$	3529.828	3529.811	-0.017	y27	3013.536	3013.562	0.026
				y29	3270.674	3270.728	0.054
				$[M+H]^+$	3497.801	3497.796	-0.005

Figure 6d pep4				Figure 6d pep5'				Figure 6d pep6, pep6-(R2-1)'			
ion	Calc.	Obs.	△	ion	Calc.	Obs.	△	ion	Calc.	Obs.	△
b2	201.123	201.208	0.085	b17	1942.881	1942.88	-0.001	b7	790.362	790.352	-0.01
b3	344.182	344.285	0.103	b19	2158.956	2158.985	0.029	b9	1040.505	1040.456	-0.049
b4	458.225	458.267	0.042	b20	2287.999	2288.043	0.044	b10	1141.552	1141.645	0.093
b5	572.268	572.276	0.008	b21	2425.057	2425.13	0.073	b14	1617.78	1617.562	-0.218
b6	701.311	701.292	-0.019	y2	319.14	319.098	-0.042	b19	2190.983	2190.943	-0.04
b8	917.385	917.329	-0.056	y3	448.183	448.133	-0.05	b20	2320.026	2319.949	-0.077
b9	1046.428	1046.336	-0.092	y4	563.21	563.195	-0.015	b21	2457.084	2457.084	0
b10	1183.487	1183.379	-0.108	y5	664.257	664.299	0.042	y2	319.14	319.069	-0.071
b11	1346.55	1346.897	0.347	y6	779.284	779.387	0.103	y4	563.21	563.151	-0.059
y2	319.14	319.24	0.1	y7	893.327	893.382	0.055	y5	664.257	665.275	1.018
y3	448.183	448.224	0.041	y8	1007.37	1007.424	0.054	y6	793.3	793.271	-0.029
y4	563.21	563.23	0.02	y11	1336.529	1336.551	0.022	y7	907.343	907.287	-0.056
y5	664.257	664.252	-0.005	y13	1584.645	1584.638	-0.007	y8	1021.386	1021.356	-0.03
y6	793.3	793.288	-0.012	y14	1721.704	1721.693	-0.011	y11	1350.545	1350.435	-0.11
y7	907.343	907.31	-0.033	y15	1834.788	1834.765	-0.023	y13	1598.661	1598.497	-0.164
y8	1021.386	1021.349	-0.037	y16	1917.825	1917.791	-0.034	y14	1735.72	1735.542	-0.178
y9	1164.445	1164.377	-0.068	y17	2018.872	2018.879	0.007	y15	1848.804	1848.672	-0.132
y10	1263.513	1263.437	-0.076	y19	2230.952	2231.042	0.09	y16	1949.852	1949.681	-0.171
$[M+H]+$	1364.561	1364.529	-0.032	y20	2345.979	2346.054	0.075	y17	2050.899	2050.814	-0.085
				y21	2459.063	2459.126	0.063	y19	2262.979	2262.957	-0.022
				$[M+H]+$	2606.131	2606.131	0	y20	2378.006	2378.076	0.07
								y21	2491.09	2491.076	-0.014
								y2*	249.123	249.127	0.004
								y3*	386.182	386.193	0.011
								y8*	913.441	913.568	0.127
								y11*	1288.62	1288.671	0.051
								$[M+H]+$	2638.158	2638.114	-0.044

Table 7. Calculated and observed masses for various ions in MALDI-TOF-MS/MS spectra of Figure 7. △is observed mass (Obs.) - calculated mass (Calc.).

Figure 7							
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b7	790.362	790.261	-0.101	y11	1336.529	1336.502	-0.027
b17	1960.892	1960.355	-0.537	y17	2036.883	2036.929	0.046
b19	2176.967	2177.01	0.043	y19	2248.963	2249.098	0.135
b20	2306.01	2306.069	0.059	y20	2363.99	2364.061	0.071
y2	319.14	319.127	-0.013	y21	2477.074	2477.316	0.242
y3	448.183	448.193	0.01	y3*	404.193	404.169	-0.024
y4	563.21	563.214	0.004	y4*	517.277	517.287	0.01
y5	664.257	664.301	0.044	y5*	600.314	600.324	0.01
y6	779.284	779.317	0.033	y8*	913.441	913.474	0.033
y7	893.327	893.389	0.062	y11*	1288.62	1288.577	-0.043
y8	1007.37	1007.383	0.013	$[M+H]^+$	2624.142	2624.122	-0.02
y9	1136.413	1136.493	0.08				

Table 8. Calculated and observed masses for various ions in MALDI-TOF-MS/MS spectra of Figure 8. △is observed mass (Obs.) − calculated mass (Calc.).

Figure 8a				Figure 8b				Figure 8c			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b2	203.103	203.121	0.018	b2	203.103	203.144	0.041	b3	341.193	341.173	-0.02
b3	359.204	359.183	-0.021	b3	359.204	359.218	0.014	b3'	324.166	324.177	0.011
b3'	342.177	342.16	-0.017	b3'	342.177	342.192	0.015	b11	1267.737	1267.718	-0.019
b4	472.288	472.227	-0.061	b4	472.288	472.259	-0.029	b12	1382.764	1382.694	-0.07
b5	573.336	573.289	-0.047	b5	555.325	555.309	-0.016	b14	1594.843	1594.765	-0.078
b6	701.43	701.367	-0.063	b6	683.419	683.38	-0.039	b15	1723.886	1723.737	-0.149
b7	802.478	802.417	-0.061	b8	897.551	897.497	-0.054	b17	2050.013	2050.08	0.067
b7'	785.451	785.412	-0.039	b9	1026.594	1026.544	-0.05	a11	1239.742	1239.721	-0.021
b8	915.562	915.483	-0.079	b10	1140.637	1140.587	-0.05	a12	1354.769	1354.7	-0.069
b9	1058.621	1058.537	-0.084	b11	1253.721	1253.648	-0.073	a17	2022.018	2022.012	-0.006
b10	1172.664	1172.571	-0.093	b12	1382.764	1382.677	-0.087	y16	1970.971	1970.736	-0.235
b11	1285.748	1285.655	-0.093	b14	1594.843	1594.774	-0.069	y17	2054.008	2053.965	-0.043
b12	1414.791	1414.685	-0.106	b15	1723.886	1723.788	-0.098	$[M+H]^+$	2155.055	2155.055	0
b14	1626.87	1626.741	-0.129	b17	2050.013	2049.962	-0.051				
b15	1755.913	1755.764	-0.149	a9	998.599	998.588	-0.011				
b17	2082.04	2081.848	-0.192	a10	1112.642	1112.619	-0.023				

a11	1257.753	1257.695	-0.058	a11	1225.726	1225.707	-0.019
a12	1386.796	1386.721	-0.075	a12	1354.769	1354.744	-0.025
a17	2054.045	2053.96	-0.085	a17	2022.018	2022.027	0.009
y16	1984.987	1984.852	-0.135	y16	1952.96	1952.862	-0.098
$[M+H]^+$	2187.082	2187.05	-0.032	y17	2054.008	2053.925	-0.083
				$[M+H]^+$	2155.055	2155.05	-0.005

Table 9. Calculated and observed masses for various ions in MALDI-TOF-MS/MS spectra of Figure 11. △is observed mass (Obs.) − calculated mass (Calc.).

Figure 11b (upper)				Figure 11b (middle)				Figure 11b (lower)			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ	ion	Calc.	Obgis.	Δ
b4	485.272	485.236	-0.036	b6	727.41	727.327	-0.083	b4	485.272	485.193	-0.079
b6	727.41	727.359	-0.051	b7	826.478	826.433	-0.045	b5	613.367	613.235	-0.132
b7	826.478	826.43	-0.048	b8	954.573	954.547	-0.026	b6	727.41	727.307	-0.103
b8	954.573	954.539	-0.034	b9	1082.668	1082.654	-0.014	b7	826.478	826.372	-0.106
b9	1082.668	1082.669	0.001	b10	1195.752	1195.787	0.035	b8	954.573	954.413	-0.16
b10	1195.752	1195.766	0.014	b11	1309.795	1309.823	0.028	b9	1082.668	1082.594	-0.074
b11	1309.795	1309.821	0.026	b12	1422.879	1422.92	0.041	b10	1195.752	1195.703	-0.049
b12	1422.879	1422.895	0.016	b13	1479.901	1479.938	0.037	b11	1309.795	1309.773	-0.022
b13	1479.901	1479.932	0.031	b14	1576.953	1576.996	0.043	b12	1422.879	1422.851	-0.028
b14	1576.953	1576.976	0.023	b15	1691.98	1692.018	0.038	b13	1479.901	1479.893	-0.008
b15	1691.98	1691.989	0.009	b16	1793.028	1793.088	0.06	b14	1576.953	1576.926	-0.027
b16	1793.028	1793.028	0	b17	1894.076	1894.078	0.002	b15	1691.98	1691.959	-0.021
b17	1894.076	1894.03	-0.046	b18	2007.16	2007.163	0.003	b16	1793.028	1792.959	-0.069
b18	2007.16	2007.1	-0.06	b19	2106.228	2106.227	-0.001	b17	1876.065	1876.057	-0.008
b19	2106.228	2106.206	-0.022	b20	2189.265	2189.21	-0.055	b27	2975.605	2975.456	-0.149
b20	2207.276	2207.184	-0.092	b21	2318.307	2318.242	-0.065	b29	3187.684	3187.647	-0.037
b21	2336.318	2336.229	-0.089	b22	2419.355	2419.131	-0.224	b30	3316.727	3316.657	-0.07
b22	2437.366	2437.229	-0.137	b23	2532.439	2532.659	0.22	y4	523.204	523.135	-0.069
b23	2550.45	2550.379	-0.071	b24	2661.482	2661.204	-0.278	y23	2544.235	2543.992	-0.243
b24	2693.509	2693.367	-0.142	b25	2775.525	2775.294	-0.231	y24	2672.33	2672.245	-0.085
b25	2807.552	2807.353	-0.199	b26	2846.562	2846.43	-0.132	y25	2771.399	2771.212	-0.187
b26	2878.589	2878.462	-0.127	b27	2975.605	2975.432	-0.173	y26	2885.441	2885.344	-0.097
b27	3007.632	3007.612	-0.02	b29	3187.684	3187.504	-0.18	y27	3013.536	3013.461	-0.075

b29	3219.711	3219.793	0.082	b30	3316.727	3316.766	0.039	$[M+H]^+$	3497.801	3497.796	-0.0051
b30	3348.754	3348.817	0.063	y4	523.204	523.128	-0.076				
y4	523.204	523.165	-0.039	y10	1180.501	1180.719	0.218				
y5	652.246	652.205	-0.041	y23	2544.235	2543.995	-0.24				
y23	2576.262	2576.159	-0.103	y24	2672.33	2672.155	-0.175				
y24	2704.357	2704.299	-0.058	y26	2885.441	2885.385	-0.056				
y26	2917.468	2917.427	-0.041	y27	3013.536	3013.381	-0.155				
y27	3045.563	3045.539	-0.024	y29	3270.674	3270.157	-0.517				
$[M+H]^+$	3529.828	3529.804	-0.024	$[M+H]^+$	3497.801	3497.801	0				

Figure 11d (upper)				Figure 11d (middle)				Figure 11d (lower)			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b7	775.395	775.45	0.055	b7	775.395	775.419	0.024	b7	757.384	757.388	0.004
b10	1118.569	1118.609	0.04	b8	888.479	888.519	0.04	b10	1100.558	1100.596	0.038
b11	1219.617	1219.625	0.008	b9	971.515	971.58	0.065	b12	1314.69	1314.752	0.062
b12	1332.701	1332.662	-0.039	b10	1100.558	1100.603	0.045	b13	1457.748	1457.811	0.063
b13	1475.759	1475.704	-0.055	b11	1201.606	1201.6	-0.006	b14	1585.807	1585.754	-0.053
b14	1603.818	1603.693	-0.125	b12	1314.69	1314.658	-0.032	b15	1698.891	1698.87	-0.021
b15	1716.902	1715.777	-1.125	b13	1443.732	1443.683	-0.049	b16	1813.918	1813.839	-0.079
b16	1845.945	1845.8	-0.145	b14	1571.791	1571.665	-0.126	b17	1941.976	1941.904	-0.072
b17	1974.003	1973.901	-0.102	b15	1684.875	1684.72	-0.155	b18	2057.003	2056.956	-0.047
b18	2089.03	2088.955	-0.075	b16	1813.918	1813.757	-0.161	b19	2186.046	2186.011	-0.035
b19	2218.073	2218.021	-0.052	b17	1941.976	1941.844	-0.132	b20	2299.13	2300.181	1.051
b20	2331.157	2330.355	-0.802	b18	2057.003	2056.934	-0.069	b24	2699.326	2699.404	0.078
b24	2731.353	2731.562	0.209	b19	2186.046	2185.97	-0.076	a10	1072.563	1072.612	0.049
a10	1090.574	1090.667	0.093	b20	2299.13	2299.82	0.69	a12	1286.695	1286.683	-0.012
a11	1191.622	1191.691	0.069	b24	2699.326	2699.425	0.099	a13	1429.753	1429.75	-0.003
a12	1304.706	1304.739	0.033	a8	860.484	860.537	0.053	a14	1557.812	1557.818	0.006
a13	1447.764	1447.777	0.013	a10	1072.563	1072.624	0.061	a15	1670.896	1670.866	-0.03
a14	1575.823	1575.778	-0.045	a11	1173.611	1173.672	0.061	y19	2280.125	2280.189	0.064
a15	1688.907	1688.836	-0.071	a12	1286.695	1286.681	-0.014	y20	2363.161	2363.242	0.081
a16	1817.95	1817.85	-0.1	a13	1415.737	1415.763	0.026	y21	2464.209	2464.267	0.058
a17	1946.008	1945.968	-0.04	a14	1543.796	1543.765	-0.031	$[M+H]^+$	2880.4	2880.4	0
a19	2190.078	2190.333	0.255	a15	1656.88	1656.848	-0.032				
a20	2303.162	2303.341	0.179	a16	1785.923	1785.804	-0.119				
y17	2024.955	2024.738	-0.217	a17	1913.981	1913.958	-0.023				
y20	2395.188	2395.208	0.02	y19	2262.114	2262.083	-0.031				
y21	2496.236	2496.352	0.116	y20	2363.161	2363.126	-0.035				
y24	2825.395	2825.694	0.299	y21	2464.209	2464.285	0.076				

$[M+H]^+$	2912.427	2912.389	-0.038	y22	2579.236	2579.282	0.046
				$[M+H]^+$	2880.4	2880.335	-0.065

Figure 11f (upper)				Figure 11f (middle)				Figure 11f (lower)			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b5	518.261	518.258	-0.003	b5	518.261	518.274	0.013	b4	337.187	337.223	0.036
b6	631.345	631.332	-0.013	b6	631.345	631.367	0.022	b5	500.25	500.285	0.035
b7	732.393	732.376	-0.017	b7	714.382	714.396	0.014	b6	613.334	613.351	0.017
b8	860.488	860.458	-0.03	b8	842.477	842.498	0.021	b7	714.382	714.434	0.052
b9	931.525	931.519	-0.006	b9	913.514	913.548	0.034	b8	842.477	842.473	-0.004
b10	1002.562	1002.553	-0.009	b10	984.551	984.576	0.025	b9	913.514	913.545	0.031
b11	1145.62	1145.599	-0.021	b11	1113.593	1113.621	0.028	b10	984.551	984.597	0.046
b12	1232.653	1232.619	-0.034	b12	1200.626	1200.648	0.022	b11	1127.609	1127.659	0.05
b13	1319.685	1319.648	-0.037	b13	1287.658	1287.675	0.017	b14	1416.7	1416.723	0.023
b14	1448.727	1448.681	-0.046	b14	1416.7	1416.711	0.011	b16	1628.78	1628.759	-0.021
b16	1660.807	1660.76	-0.047	b15	1513.753	1513.771	0.018	b17	1757.823	1757.777	-0.046
b17	1789.85	1789.814	-0.036	b16	1628.78	1628.764	-0.016	y4	507.209	507.238	0.029
y4	507.209	507.208	-0.001	b17	1757.823	1757.764	-0.059	y11	1209.527	1209.563	0.036
y11	1223.543	1223.527	-0.016	y4	507.209	507.208	-0.001	y12	1310.575	1310.613	0.038
y12	1324.591	1324.569	-0.022	y11	1209.527	1209.562	0.035	y13	1423.659	1423.696	0.037
y13	1437.675	1437.642	-0.033	y12	1292.564	1292.585	0.021	y14	1586.722	1586.713	-0.009
y14	1600.738	1600.695	-0.043	y13	1405.648	1405.651	0.003	y16	1766.812	1766.79	-0.022
y15	1701.786	1701.742	-0.044	y14	1568.711	1568.699	-0.012	y17	1823.833	1823.806	-0.027
y16	1798.839	1798.798	-0.041	y16	1766.812	1766.794	-0.018	$[M+H]^+$	1922.902	1922.902	0
y17	1855.86	1855.823	-0.037	y17	1823.833	1823.802	-0.031				
$[M+H]^+$	1954.929	1954.958	0.029	$[M+H]^+$	1922.902	1922.902	0				

Figure 11h (upper)				Figure 11h (middle)				Figure 11h (lower)			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b6	727.41	727.36	-0.05	b6	727.41	727.396	-0.014	b4	485.272	485.228	-0.044
b7	826.478	826.456	-0.022	b7	826.478	826.472	-0.006	b6	727.41	727.283	-0.127
b8	954.573	954.559	-0.014	b8	954.573	954.59	0.017	b7	826.478	826.297	-0.181
b9	1082.668	1082.678	0.01	b9	1082.668	1082.704	0.036	b9	1082.668	1082.511	-0.157
b10	1195.752	1195.772	0.02	b10	1195.752	1195.824	0.072	b10	1195.752	1195.813	0.061
b11	1309.795	1309.844	0.049	b11	1309.795	1309.879	0.084	b11	1309.795	1309.738	-0.057
b12	1422.879	1422.919	0.04	b12	1422.879	1422.956	0.077	b12	1422.879	1422.861	-0.018

b13	1479.901	1479.933	0.032	b13	1479.901	1479.983	0.082	b13	1479.901	1479.937	0.036
b14	1576.953	1577.05	0.097	b14	1576.953	1576.993	0.04	b15	1691.98	1692.125	0.145
b15	1691.98	1692.002	0.022	b15	1691.98	1692.049	0.069	b16	1793.028	1793.032	0.004
b16	1793.028	1793.066	0.038	b16	1793.028	1793.046	0.018	b17	1876.065	1876.209	0.144
b17	1894.076	1894.053	-0.023	b17	1894.076	1894.143	0.067	b18	1989.149	1989.017	-0.132
b18	2007.16	2007.111	-0.049	b18	2007.16	2007.219	0.059	b19	2088.217	2088.087	-0.13
b19	2106.228	2106.18	-0.048	b19	2106.228	2106.254	0.026	b20	2189.265	2189.211	-0.054
b20	2207.276	2207.146	-0.13	b20	2189.265	2189.318	0.053	b21	2318.307	2318.129	-0.178
b21	2336.318	2336.246	-0.072	b21	2318.307	2318.333	0.026	b23	2532.439	2532.785	0.346
b22	2437.366	2437.243	-0.123	b22	2419.355	2419.317	-0.038	b24	2675.498	2675.276	-0.222
b23	2550.45	2550.149	-0.301	b23	2532.439	2532.344	-0.095	b25	2789.541	2790.107	0.566
b24	2693.509	2693.457	-0.052	b24	2661.482	2661.491	0.009	b27	2975.605	2974.796	-0.8093
b25	2807.552	2807.437	-0.115	b25	2775.525	2775.573	0.048	b30	3272.737	3272.813	0.076
b26	2878.589	2878.39	-0.199	b26	2846.562	2846.541	-0.021	$[M+H]^+$	3453.811	3453.865	0.054
b27	3007.632	3007.54	-0.092	b27	2975.605	2975.614	0.009				
b29	3175.722	3175.635	-0.087	b29	3143.695	3143.6	-0.095				
b30	3304.764	3304.9	0.136	b30	3272.737	3272.791	0.054				
y23	2532.272	2532.099	-0.173	y23	2500.245	2500.16	-0.085				
y29	3258.711	3258.695	-0.016	$[M+H]^+$	3453.811	3453.8061	-0.0049				
$[M+H]^+$	3485.838	3485.838	0								

Table 10. Calculated and observed masses for various ions in MALDI-TOF-MS/MS spectra of Figure 15. △is observed mass (Obs.) − calculated mass (Calc.).

Figure 15a (top)				Figure 15a (bottom)			
ion	Calc.	Obs.	△	ion	Calc.	Obs.	△
b6	727.41	727.363	-0.047	b6	727.4098	727.388	-0.02176
b7	826.478	826.435	-0.043	b7	826.4782	826.444	-0.03417
b8	954.573	954.532	-0.041	b8	954.5731	954.555	-0.01813
b9	1082.668	1082.657	-0.011	b9	1082.668	1082.665	-0.0031
b10	1195.752	1195.754	0.002	b10	1195.752	1195.742	-0.01016
b11	1309.795	1309.824	0.029	b11	1309.795	1309.781	-0.01409
b12	1422.879	1422.908	0.029	b12	1422.879	1422.872	-0.00715

b13	1479.901	1479.935	0.034	b13	1479.901	1479.895	-0.00561
b14	1576.953	1576.81	-0.143	b14	1576.953	1576.957	0.00362
b15	1691.98	1691.977	-0.003	b15	1691.98	1691.957	-0.02332
b16	1793.028	1792.964	-0.064	b16	1793.028	1792.927	-0.101
b17	1894.076	1894.045	-0.031	b17	1894.076	1894.046	-0.02968
b18	2007.16	2007.054	-0.106	b18	2007.16	2007.074	-0.08574
b19	2106.228	2106.087	-0.141	b19	2106.228	2106.137	-0.09115
b20	2189.265	2189.178	-0.087	b21	2336.318	2336.235	-0.08343
b24	2661.482	2661.994	0.512	b22	2437.366	2437.204	-0.1621
b25	2775.525	2775.236	-0.289	b23	2550.45	2550.281	-0.16917
b26	2846.562	2846.493	-0.069	b24	2693.509	2693.323	-0.18576
b27	2961.589	2961.489	-0.1	b25	2807.552	2807.364	-0.18769
b29	3173.668	3173.531	-0.137	b26	2878.589	2878.617	0.0282
b30	3302.711	3302.583	-0.128	b27	2993.616	2993.586	-0.02974
y23	2530.219	2529.984	-0.235	b29	3205.695	3205.717	0.02155
$[M+H]^+$	3483.785	3483.785	0	b30	3334.738	3334.811	0.07296
				y23	2562.2461	2561.968	-0.27812
				$[M+H]^+$	3515.812	3515.812	0

Figure 15b (top)				Figure 15b (bottom)			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b6	727.41	727.317	-0.093	b6	727.40976	727.344	-0.06576
b7	826.478	826.396	-0.082	b7	826.47817	826.424	-0.05417
b9	1082.668	1082.635	-0.033	b8	954.57313	954.588	0.01487
b10	1195.752	1195.71	-0.042	b9	1082.6681	1082.659	-0.0091
b11	1309.795	1309.775	-0.02	b10	1195.7522	1195.77	0.01784
b12	1422.879	1422.854	-0.025	b11	1309.7951	1309.834	0.03891
b13	1479.901	1479.874	-0.027	b12	1422.8792	1422.946	0.06685
b14	1576.953	1576.96	0.007	b13	1479.9006	1479.971	0.07039
b15	1691.98	1691.931	-0.049	b14	1576.9534	1576.993	0.03962
b16	1793.028	1792.91	-0.118	b15	1691.9803	1692.045	0.06468
b17	1876.065	1875.947	-0.118	b16	1793.028	1792.856	-0.172
b27	2961.589	2961.454	-0.135	b17	1894.0757	1894.127	0.05132
b29	3173.668	3173.705	0.037	b18	2007.1597	2007.105	-0.05474
b30	3302.711	3302.753	0.042	b19	2106.2282	2106.252	0.02385
y23	2530.219	2530.058	-0.161	b20	2207.2758	2207.131	-0.14483
y24	2658.314	2658.205	-0.109	b21	2336.3184	2336.316	-0.00243
y25	2757.383	2756.941	-0.442	b23	2550.4502	2550.245	-0.20517
y26	2871.425	2871.21	-0.215	b24	2679.4928	2679.426	-0.06676
y27	2999.52	2999.286	-0.234	b25	2793.5357	2793.276	-0.25969
$[M+H]^+$	3483.785	3483.785	0	b26	2864.5728	2864.731	0.1582

	b27	2993.616	2993.655	0.03926
	b29	3205.695	3205.794	0.09855
	b30	3334.738	3334.626	-0.11204
	y23	2562.2461	2562.122	-0.12412
	y24	2690.3411	2690.479	0.13791
	$[M+H]^+$	3515.812	3515.812	0

Figure 15C AxxTxxxExxD				Figure 15C AxxTxxxExxD methanolysis			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b6	727.41	727.677	0.267	b6	727.41	727.553	0.143
b7	826.478	826.734	0.256	b7	826.478	826.68	0.202
b8	954.573	954.803	0.23	b8	954.573	954.609	0.036
b9	1082.668	1083.075	0.407	b9	1082.668	1082.949	0.281
b10	1195.752	1196.192	0.44	b10	1195.752	1196.044	0.292
b11	1309.795	1310.242	0.447	b11	1309.795	1310.112	0.317
b12	1422.879	1423.345	0.466	b12	1422.879	1423.207	0.328
b13	1479.901	1480.355	0.454	b13	1479.901	1480.221	0.32
b14	1576.953	1577.437	0.484	b14	1576.953	1577.239	0.286
b15	1691.98	1692.438	0.458	b15	1691.98	1692.304	0.324
b16	1793.028	1793.465	0.437	b16	1793.028	1793.333	0.305
b17	1864.065	1864.493	0.428	b17	1864.065	1864.359	0.294
b18	1977.149	1977.569	0.42	b18	1977.149	1977.376	0.227
b19	2076.218	2076.635	0.417	b19	2076.218	2076.452	0.234
b20	2159.254	2159.523	0.269	b21	2306.308	2306.569	0.261
b24	2631.471	2631.378	-0.093	b22	2407.356	2407.356	0
b25	2745.514	2745.91	0.396	b23	2520.44	2520.494	0.054
b26	2816.551	2816.704	0.153	b24	2663.482	2663.597	0.115
b27	2931.578	2931.78	0.202	b25	2777.525	2777.709	0.184
b29	3143.658	3143.728	0.07	b26	2848.562	2848.815	0.253
b30	3272.7	3272.594	-0.106	b27	2963.589	2963.803	0.214
y23	2500.209	2500.431	0.222	b29	3175.669	3175.836	0.167
[M+H]+	3453.785	3453.774	-0.01063	b30	3304.711	3304.821	0.11
				y23	2532.22	2532.293	0.073
				[M+H]+	3485.785	3485.798	0.013

Figure 15C TxxAxxxExxD				Figure 15C TxxAxxxExxD methanolysis			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b6	727.41	727.377	-0.033	b6	727.41	727.359	-0.051
b7	826.478	826.409	-0.069	b7	826.478	826.386	-0.092
b8	954.573	954.499	-0.074	b8	954.573	954.591	0.018
b9	1082.668	1082.637	-0.031	b9	1082.668	1082.669	0.001
b10	1195.752	1195.713	-0.039	b10	1195.752	1195.77	0.018

b11	1309.795	1309.783	-0.012	b11	1309.795	1309.812	0.017
b12	1422.879	1422.84	-0.039	b12	1422.879	1422.911	0.032
b13	1479.901	1479.854	-0.047	b13	1479.901	1479.929	0.028
b14	1576.953	1576.899	-0.054	b14	1576.953	1576.921	-0.032
b15	1691.98	1691.926	-0.054	b15	1691.98	1691.97	-0.01
b16	1793.028	1792.951	-0.077	b16	1793.028	1792.969	-0.059
b17	1876.076	1875.965	-0.111	b17	1894.076	1894.036	-0.04
b27	2931.589	2931.49	-0.099	b18	2007.16	2007.054	-0.106
b29	3143.669	3143.693	0.024	b19	2106.228	2106.072	-0.156
b30	3272.711	3272.926	0.215	b20	2177.265	2177.013	-0.252
y23	2500.22	2500.058	-0.162	b21	2306.308	2306.088	-0.22
y24	2628.315	2628.16	-0.155	b23	2520.44	2520.171	-0.269
y25	2727.383	2727.219	-0.164	b24	2649.482	2649.042	-0.44
y26	2841.426	2841.304	-0.122	b25	2763.525	2763.379	-0.146
[M+H]+	3453.785	3453.785	0	b26	2834.562	2834.284	-0.278
				b27	2963.589	2963.129	-0.46
				b29	3175.669	3175.527	-0.142
				[M+H]+	3485.785	3485.785	0

Figure 15C TxxTxxxAxxD				Figure 15C TxxTxxxAxxD methanolysis			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b6	727.41	727.372	-0.038	b6	727.41	727.643	0.233
b7	826.478	826.434	-0.044	b7	826.478	826.665	0.187
b8	954.573	954.534	-0.039	b8	954.573	954.738	0.165
b9	1082.668	1082.65	-0.018	b9	1082.668	1082.836	0.168
b10	1195.752	1195.751	-0.001	b10	1195.752	1195.902	0.15
b11	1309.795	1309.785	-0.01	b11	1309.795	1309.916	0.121
b12	1422.879	1422.866	-0.013	b12	1422.879	1422.966	0.087
b13	1479.901	1479.879	-0.022	b13	1479.901	1479.967	0.066
b14	1576.953	1576.922	-0.031	b14	1576.953	1577.009	0.056
b15	1691.98	1691.934	-0.046	b15	1691.98	1692.017	0.037
b16	1793.028	1792.965	-0.063	b16	1793.028	1792.912	-0.116
b17	1876.076	1876.024	-0.052	b17	1894.076	1894.433	0.357
b27	2903.594	2903.592	-0.002	b18	2007.16	2007.026	-0.134
b29	3115.674	3115.622	-0.052	b19	2106.228	2106.059	-0.169
b30	3244.717	3244.566	-0.151	b20	2207.276	2207.205	-0.071
y23	2472.225	2472.037	-0.188	b21	2336.318	2336.083	-0.235
y24	2600.32	2600.164	-0.156	b23	2550.45	2550.183	-0.267
y25	2699.388	2699.26	-0.128	b24	2621.487	2622.17	0.683
y26	2813.431	2813.288	-0.143	b25	2735.53	2735.607	0.077
[M+H]+	3425.79	3425.79	0	b26	2806.567	2806.13	-0.437
				b27	2935.594	2935.39	-0.204
				b29	3147.674	3147.558	-0.116

				b30	3276.717	3276.734	0.017
				[M+H]+	3457.79	3456.817	-0.973
Figure 15C TxxTxxxExxA							
ion	Calc.	Obs.	Δ				
b6	727.41	727.376	-0.034				
b7	826.478	826.442	-0.036				
b8	954.573	954.541	-0.032				
b9	1082.668	1082.671	0.003				
b10	1195.752	1195.764	0.012				
b11	1309.795	1309.81	0.015				
b12	1422.879	1422.893	0.014				
b13	1479.901	1479.894	-0.007				
b14	1576.953	1576.958	0.005				
b15	1691.98	1691.971	-0.009				
b16	1793.028	1792.995	-0.033				
b17	1894.076	1893.98	-0.096				
b18	2007.16	2007.102	-0.058				
b19	2106.228	2106.141	-0.087				
b20	2189.276	2189.232	-0.044				
b24	2661.493	2661.311	-0.182				
b25	2775.536	2775.477	-0.059				
b26	2846.573	2846.58	0.007				
b27	2917.61	2917.583	-0.027				
b29	3129.69	3129.752	0.062				
b30	3258.732	3258.725	-0.007				
y23	2486.24	2486.175	-0.065				
y24	2614.335	2614.302	-0.033				
y25	2713.404	2713.286	-0.118				
[M+H]+	3439.806	3439.806	0				

Table 11. Calculated and observed masses for various ions in MALDI-TOF-MS/MS spectra of Figure 18. △is observed mass (Obs.) - calculated mass (Calc.).

Figure 18a				Figure 18b				Figure 18c			
ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ	ion	Calc.	Obs.	Δ
b13	1429.717	1429.705	-0.012	b7	839.426	839.489	0.063	b10	1182.6	1182.573	-0.027
b16	1785.886	1785.759	-0.127	b8	952.51	952.594	0.084	b11	1283.648	1283.589	-0.059

b18	2028.972	2028.876	-0.096	b9	1053.557	1053.633	0.076	b12	1396.732	1396.686	-0.046
b19	2158.014	2157.872	-0.142	b10	1196.616	1196.699	0.083	b13	1511.759	1511.71	-0.049
b24	2671.294	2671.507	0.213	b11	1297.664	1297.687	0.023	b14	1639.817	1639.729	-0.088
y21	2436.177	2436.279	0.102	b12	1410.748	1410.696	-0.052	b15	1752.901	1752.735	-0.166
$[M+H]^+$	2852.368	2852.368	0	b13	1525.775	1525.679	-0.096	b16	1867.928	1867.802	-0.126
				b14	1653.833	1653.723	-0.11	b17	1995.987	1995.889	-0.098
				b15	1766.917	1766.813	-0.104	b18	2111.014	2110.887	-0.127
				b16	1881.944	1881.729	-0.215	b19	2240.056	2239.987	-0.069
				b17	2010.003	2009.82	-0.183	b20	2353.14	2353.051	-0.089
				b18	2125.03	2124.817	-0.213	b21	2424.177	2424.249	0.072
				b19	2254.072	2254.012	-0.06	b22	2525.225	2525.29	0.065
				b20	2367.156	2367.16	0.004	b23	2638.309	2638.46	0.151
				b21	2438.193	2438.248	0.055	b24	2753.336	2753.536	0.2
				b23	2652.325	2652.403	0.078	y21	2518.219	2518.29	0.071
				b24	2767.352	2767.503	0.151	y1*	239.15	239.101	-0.049
				$[M+H]^+$	2948.426	2948.573	0.147	y2*	322.187	322.108	-0.079
								y3*	423.234	423.168	-0.066
								y7*	839.425	839.459	0.034
								$[M+H]^+$	2934.41	2934.42	0.01

Table 12. Calculated and observed masses for various ions in MALDI-TOF-MS/MS spectra of Figure 19. △is observed mass (Obs.) − calculated mass (Calc.).

Figure 19a (upper)				Figure 19a (middle)				Figure 19a (lower)			
ion	Calc.	Obs.	△	ion	Calc.	Obs.	△	ion	Calc.	Obs.	△
b10	1146.578	1145.592	-0.986	b7	839.426	839.482	0.056	b16	1877.948	1877.725	-0.223
b16	1859.937	1859.936	-0.001	b8	952.51	952.535	0.025	b17	2006.007	2005.802	-0.205
b17	1987.996	1987.974	-0.022	b9	1053.557	1053.589	0.032	b18	2121.034	2120.831	-0.203
b18	2103.023	2103.044	0.021	b10	1182.6	1182.621	0.021	b19	2250.077	2249.941	-0.136
b19	2232.066	2232.108	0.042	b11	1283.648	1283.645	-0.003	b20	2363.161	2363.129	-0.032
b20	2345.15	2345.272	0.122	b12	1396.732	1396.785	0.053	b21	2434.198	2434.162	-0.036
b21	2416.187	2416.335	0.148	b13	1539.79	1539.764	-0.026	b22	2535.245	2535.223	-0.022
b22	2517.234	2517.485	0.251	b14	1667.849	1667.756	-0.093	b23	2648.329	2648.382	0.053
b23	2630.318	2630.539	0.221	b15	1780.933	1780.772	-0.161	b24	2763.356	2763.552	0.196

b24	2745.345	2745.639	0.294	b16	1923.991	1923.828	-0.163	y21	2528.24	2528.242	0.002
y21	2510.229	2510.466	0.237	b17	2052.05	2051.932	-0.118	y2*	258.156	258.056	-0.1
$[M+H]^+$	2926.419	2926.419	0	b18	2167.077	2166.945	-0.132	y3*	359.203	359.105	-0.098
				b19	2296.12	2296.001	-0.119	y7*	775.394	775.471	0.077
				b20	2409.204	2409.074	-0.13	$[M+H]^+$	2944.43	2944.436	0.006
				b21	2480.241	2480.263	0.022				
				b22	2581.288	2581.39	0.102				
				b23	2694.372	2694.252	-0.12				
				b24	2809.399	2809.657	0.258				
				$[M+H]^+$	2990.473	2990.473	0				

Figure 19b (upper)

ion	Calc.	Obs.	Δ
b13	1425.721	1425.751	0.03
b14	1553.78	1553.777	-0.003
b16	1795.906	1795.856	-0.05
b18	2038.992	2039.001	0.009
b19	2168.035	2168.051	0.016
b24	2681.314	2681.503	0.189
y20	2345.15	2345.284	0.134
y21	2446.198	2446.342	0.144
y22	2561.225	2561.438	0.213
$[M+H]^+$	2862.388	2862.416	0.028

Figure 19b (middle)

ion	Calc.	Obs.	Δ
b7	839.426	839.347	-0.079
b8	952.51	952.449	-0.061
b9	1053.557	1053.446	-0.111
b10	1196.616	1196.762	0.146
b11	1297.664	1297.79	0.126
b12	1410.748	1410.799	0.051
b13	1553.806	1553.78	-0.026
b14	1681.865	1681.894	0.029
b15	1794.949	1794.842	-0.107
b16	1923.991	1923.983	-0.008
b17	2052.05	2051.994	-0.056
b18	2167.077	2166.909	-0.168
b19	2296.12	2295.958	-0.162
b20	2409.204	2409.198	-0.006
b21	2480.241	2480.075	-0.166
$[M+H]^+$	2990.473	2990.585	0.112

Figure 19b (lower)

ion	Calc.	Obs.	Δ
b13	1443.732	1443.947	0.215
b16	1813.917	1814.085	0.168
b18	2057.003	2056.797	-0.206
b19	2186.046	2186.196	0.15
b24	2699.325	2699.216	-0.109
y21	2464.209	2464.115	-0.094
y1*	175.119	175.114	-0.005
y2*	258.156	258.111	-0.045
y3*	359.203	359.161	-0.042
y4*	474.23	474.237	0.007
y7*	775.394	775.423	0.029
$[M+H]^+$	2880.399	2880.833	0.434

8. 요약(국문초록)

Microviridin은 리보솜에서 합성된 후 번역 후 변형이 일어나는 펩타이드(RiPP)이며, 두 아미노산 곁사슬 사이에 오메가 에스터 결합이나 오메가 아마이드 결합이 있다는 점에서 특징적이다. 이러한 아미노산 곁사슬 사이의 결합으로 고리가 생성되는 펩타이드들은 다양한 구조를 가질 수 있다. 하지만 현재까지 밝혀진 대부분의 RiPPs 중에 microviridin만이 이러한 방식으로 변형이 일어나며, microviridin은 오직 하나의 보존된 펩타이드 서열과 다중 고리 구조를 가진다. 이번 연구에서는 mTgnA라고 하는 새로운 RiPP와 두 개의 mTgnA 상동 펩타이드들의 연결 구조를 밝힘으로써 microviridin와 유사하게 변형되는 펩타이드의 다양성을 넓혔다. 이 펩타이드들은 코어와 리더에 각각 TxxTxxxExxDxD 와 KPYxxxYxE 의 보존서열을 가지며, 코어 펩타이드는 같은 gene cluster에 속한 ATP-grasp enzyme에 의해서 독특한 이중 고리 구조로 바뀐다. 따라서 이러한 특징적인 고리 패턴에 기반하여, mTgnA와 mTgnA 상동 펩타이드들을 묶어서 thuringinin group이라고 정의하였다. Thuringinin group의 core motif의 고리 패턴은 에스터 결합을 풀 수 있는 화학 반응과 텐덤 질량분석기를 사용하여 쉽게 밝힐 수 있었다. 이러한 연결 구조를 형성하는 효소 반응은 순차적이며, 안쪽 고리가 먼저 생기고 그 다음에 바깥쪽 고리가 생긴다. 또한

이 반응은 반응에 참여하는 산성 아미노산의 종류에 선택적이다. 따라서, 이번 연구에서는 microviridin 유사 변형 펩타이드가 기존에 알고 있던 것보다 자연계에 훨씬 더 다양하게 존재한다는 것을 제시하였다. 또한 이 새로운 RiPP 의 집합을 오페가 에스터 결합 포함 펩타이드, omega-ester containing peptides (OEPs) 라고 명명할 것을 제시하였다.

주요어: RiPPs, microviridin, plesiocin, ATP-grasp, 오메가 에스터 결합 포함 펩타이드, 에스터 결합, 대고리화

학번: 2017-25338