A Novel Epimerase Catalyzing Multiple Isomerization of Amino Acid Residues of Ribosomal Peptide

リボソームペプチドの複数アミノ酸残基を異性化する新規エピメラーゼ

Wanlu Xiao

Graduate School of Chemical Sciences and Engineering
Hokkaido University

Abstract

Salinipeptins, grisemycin, and cypemycin are ribosomally synthesized and post-translationally modified peptides (RiPPs). Among these, salinipeptins were reported to comprise 22 amino acid residues with multiple D-amino acids, and its biosynthetic gene cluster was identified. However, no genes homologous to known isomerases such as epimerases and racemases existed in the cluster, but a gene, *sinL*, which showed no similarities to function known enzymes, located in the cluster, suggesting that SinL might be a novel epimerase. Actually, biosynthetic gene clusters of grisemycin and cypemycin also possess orthologs of *sinL*, although the chirality of amino acids composing grisemycin and cypemycin has not been reported. Therefore, I first examined grisemycin also contains D-amino acid residues. By heterologous expression of grisemycin biosynthetic gene cluster (*grm*) in *Streptomyces lividans*, grisemycin was confirmed to contain multiple D-amino acids, in a similar manner to salinipeptins. The heterologous expression experiments also confirmed the involvement of a novel peptide epimerase in grisemycin biosynthesis. Gene-deletion experiments indicated that *grmL*, an ortholog of *sinL*, was indispensable for grisemycin production and that the epimerization preceded decarboxylation and methylation, which are other modifications installed into the precursor peptides of grisemycin (GrmA).

To obtain further evidence that GrmL encodes the novel epimerase,

recombinant precursor peptide (GrmA) and GrmL were prepared and used for *in vitro* analysis. However, no isomerase activity was observed under various conditions. Considering that grisemycin contains dehydroamino acids and its putative biosynthetic gene, *grmH*, exists in the gene cluster, the dehydration reaction might occur before isomerization. To examine the possibility, recombinant GrmH was prepared and incubated with GrmA, but no dehydration activity was detected. Because GrmL is the novel enzyme found in the *Streptomyces* strain, I considered the possibility that GrmL might require a cofactor specifically utilized in the *Streptomyces* strain. I therefore employed the above-mentioned heterologous expression system again. When *grmA*, *grmH*, and *grmL* were co-expressed, a dehydrated and isomerized GrmA was produced, but *grmA* and either of the two resulted in the production of no modified GrmA. The results suggested that both GrmH and GrmL, which were shown to constitute a protein complex by a pulldown assay, were required to catalyze the dehydration, epimerization, and proteolytic cleavage of a precursor peptide GrmA by *in vivo* experiments.

Table of Contents

Chapter 1

General Introduction

1.1 Peptide natural products

Natural secondary metabolites possess highly complex structures with a wide range of biological activities. In recent years, more attention has been paid to numerous peptide and peptide-derived natural products, which have been isolated from bacteria, fungi, and plants, to discover drug-lead compounds in the pharmaceutical industry[1–5].

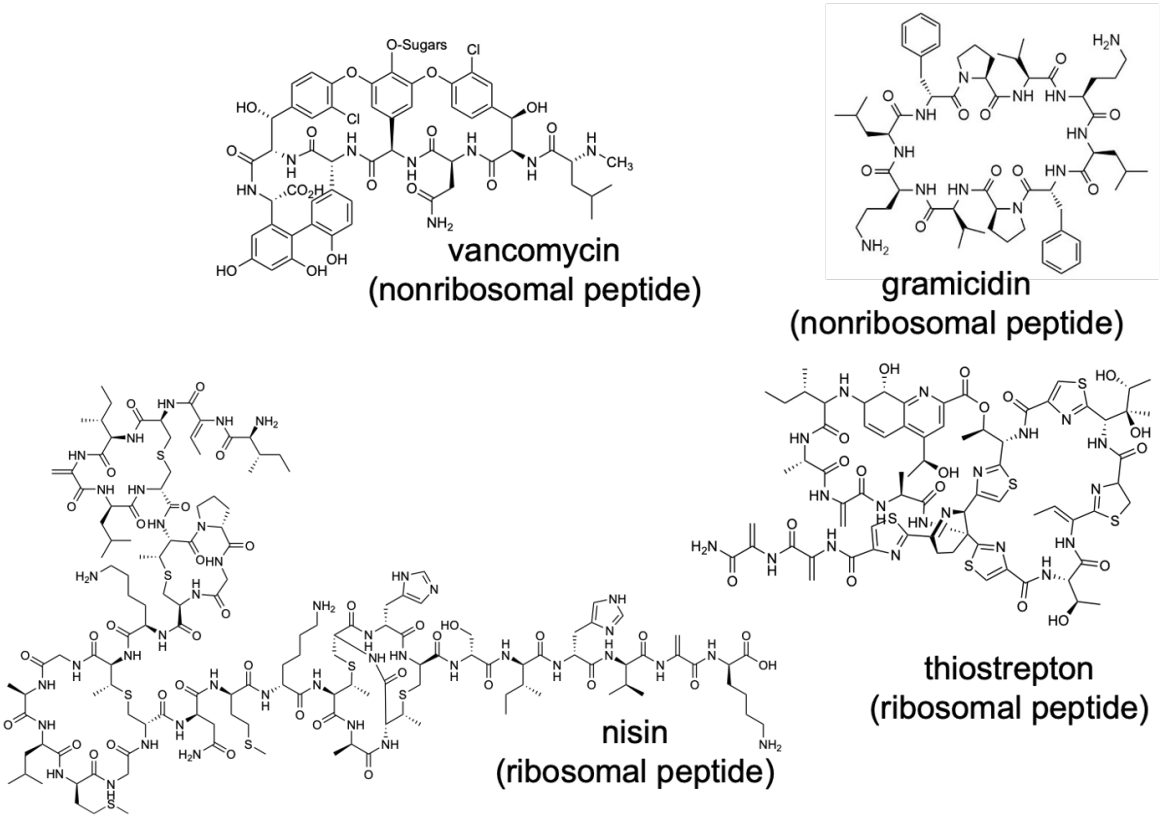


Figure 1-1 Examples of peptide natural products.

Many natural peptides possess middle molecular weight[6–9] (Figure 1-1), which would increase the affinity between the peptides and the target molecule in cells and promote the formation of distinct secondary structures to show the specific biological activities[10–12]. Therefore, the peptides are expected to fill the gap between small molecule pharmaceuticals and antibody drugs.

Peptide natural products are synthesized by ribosomal or nonribosomal pathways. Most of all nonribosomal peptides (NRPs) are biosynthesized by nonribosomal peptide synthetase (NRPS) composed of several modules[13], which contain adenylation (A) domain, peptidyl carrier protein (PCP), and condensation (C) domain as a minimum set. Growing peptidyl chains are covalently tethered as thioesters on PCP and the amine group of amino-acyl substrate on the next PCP attacks to carbonyl group of the growing peptidyl chain on PCP to make the next amido bond. The final product tethered on PCP is released by thioesterase (TE) domain. NPRSs can utilize many nonproteinogenic amino acids as building blocks and therefore create the structural and functional diversities[4,14,15] (Figure 1-2).

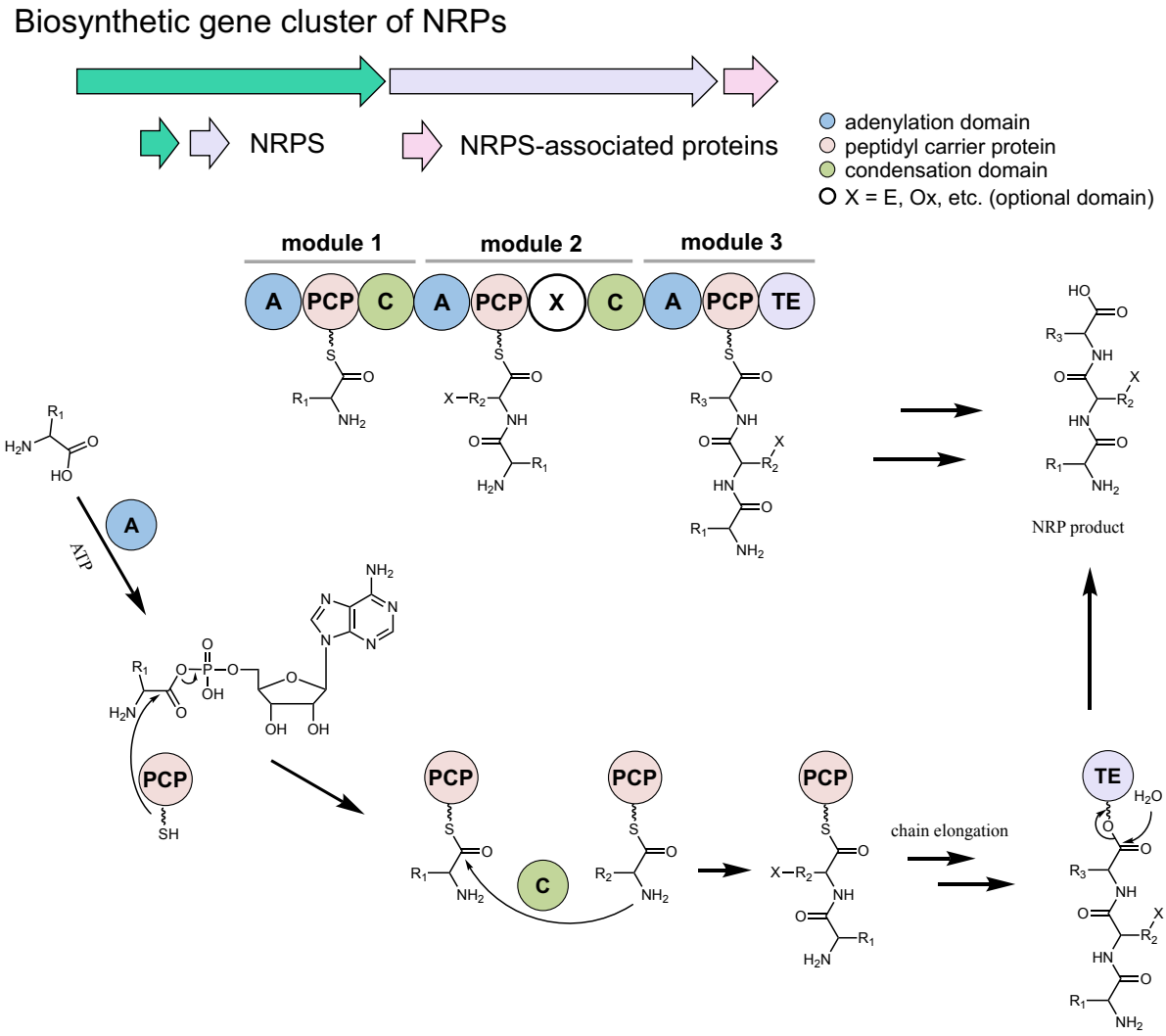


Figure 1-2 General biosynthesis of NRP.

Another class of peptide natural products is ribosomally synthesized and post-translationally modified peptides (RiPPs)[16,17] (Figure 1-3). In ribosomal pathways, mRNA is used as a template for precursor peptide biosynthesis. Genes encoding the precursor peptide and posttranslational maturation enzymes usually form a biosynthetic gene cluster (BGC). The precursor peptide usually consists of a core peptide region and an additional *N*-terminal leader region and/or a C-terminal follower region. The leader- and follower-peptide are usually important for the recognition of many post-translational modification (PTM) enzymes. The maturation enzymes install modifications in the core peptide region prior to the proteolytic release of the leader- and/or follower-peptides. Although 20 encoded amino acids are used for the biosynthesis of precursor peptides, extensive posttranslational modifications provide mature peptides with enormous chemical and biological diversity.

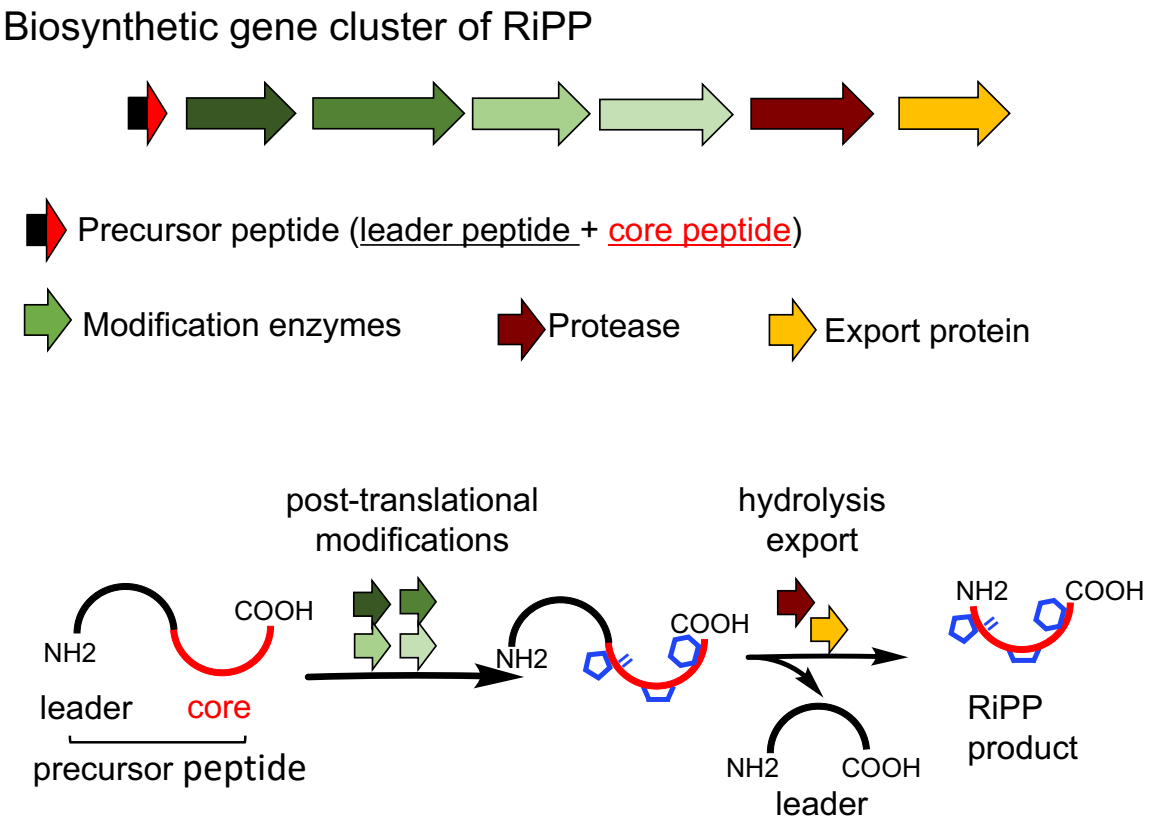


Figure 1-3 General biosynthesis of RiPP

NRPS is usually composed of a huge enzyme complex and therefore gene manipulation of the NRPS genes is difficult[5,18]. In comparison, RiPP pathways are genetically simple because each biosynthetic enzyme is encoded by a separate gene, and therefore extensive posttranslational modifications can be introduced into the products with enormous chemical and biological diversities[17].

Therefore, I focus on modification enzymes found in ribosomal pathways.

1.2 D-Amino acids containing peptides

Generally, the peptide bonds forming the backbone of proteins and peptides are susceptible to proteases[19]. Therefore, many bioactive natural peptides are protected from enzymatic degradations by modifications, such as cyclization[14,20], glycosylation[6,21], and incorporation of nonproteinogenic amino acids[7,22] (Fig. 1-4).

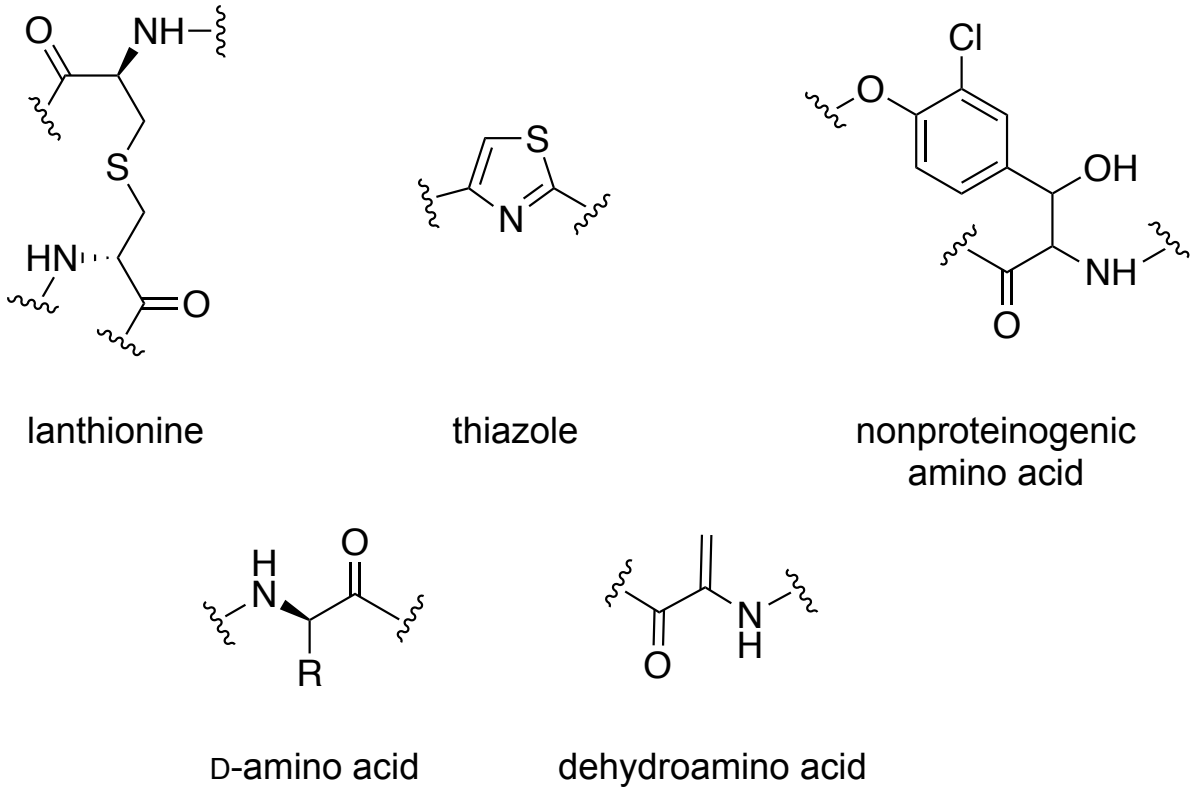


Figure 1-4 Examples of modification in peptide natural products.

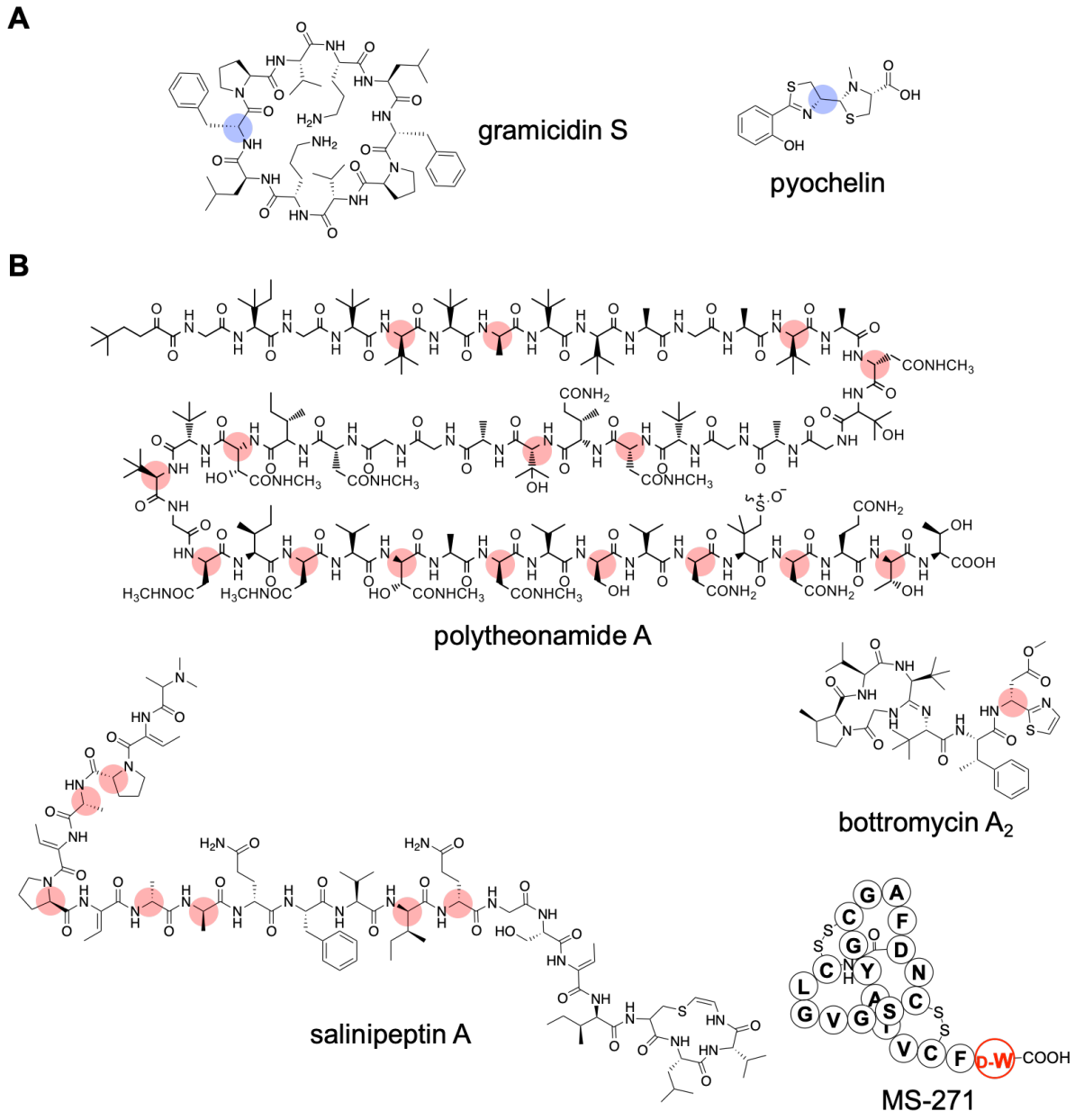


Figure 1-5 D-Amino acids are common structural features in both NRP (A), and RiPP (B) natural products.

As an alternative strategy, NRPS and RiPP peptides possess D-amino acids (Figure 1-5). In NRPSs, D-amino acids are formed by *in situ* epimerization of the Cα center of the PCP-tethered intermediate[7,13]. Commonly, the reaction is catalyzed by an epimerization (E) domain located between the PCP and C domains and generates an equilibrium of D- and L-configurated products. Then, the downstream C domain recognizing the strict D-stereoisomer catalyzes the next condensation. As an example, gramicidin S biosynthesis[23] is schematically shown in Figure 1-6.

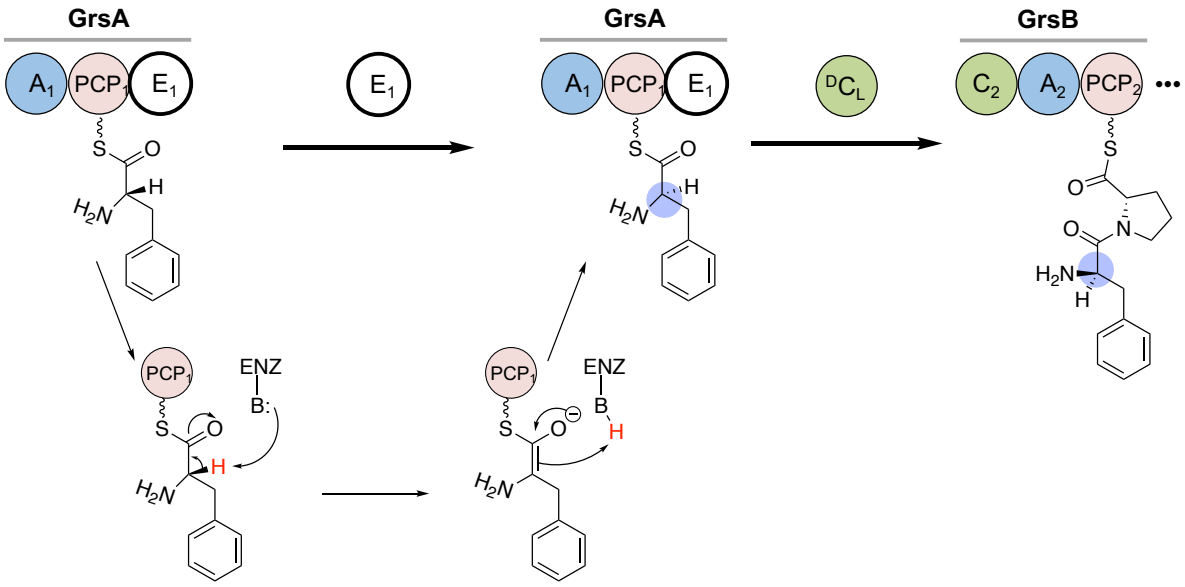


Figure 1-6 D-Phe formation catalyzed by epimerase domains in NRPS assembly lines.

In RiPPs, since the precursor peptide is ribosomal origin, post-translation epimerization is required to convert L-amino acid residues into D-configuration. The reaction is chemically challenging, but some RiPPs are reported to contain D-amino acids[10,11,24] (Figure 1-5 B).

To date, five enzymatic reactions catalyzing the epimerization of ribosomal peptides have been characterized, involving radical *S*-adenosylmethionine (rSAM) enzymes[25–28] (Figure 1-7), α/β-hydrolase family enzymes like BotH[29,30] (Figure 1-8), a metallo-dependent enzyme MslH[31,32] (Figure 1-9), and a two-step dehydration–hydrogenation process[16,33] (Figure 1-10).

Polytheonamide A, isolated from *Theonella swinhoei*, came to light because of its high potency as cytotoxic agent against leukemia cell lines. The most remarkable structural feature of polytheonamide A is the presence of 18 D-residues (Figure 1-5).

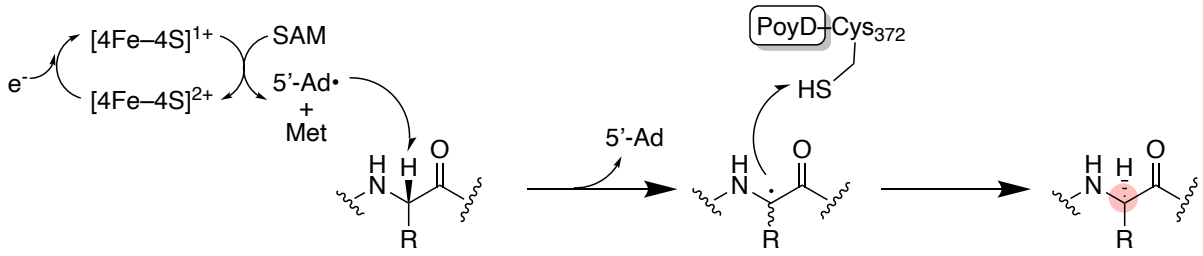


Figure 1-7 The mechanism of the the radiacal SAM epimerase, PoyD.

Through extensive studies, PoyD, a radical SAM (rSAM) enzyme, was shown to catalyze all the epimerization reactions. To date, two subfamilies of rSAM enzymes, PoyD and YydG, were reported to utilize the oxygen-sensitive [4Fe–4S] cluster as a cofactor. Based on the general properties of rSAM enzymes, the [4Fe–4S] cluster is reduced to generate a 5′-deoxyadenosyl radical (5′-dA•) and abstracts a Cα hydrogen atom of an amino acid residue. Then, the generated carbon-centered radical is quenched by the thiolate H atom of Cys-372 of PoyD leading to the formation of a D-amino acid residue[26] (Figure 1-7).

Bottromycins, isolated from *S. bottropensis*, are active against human pathogens. They are featured to have a macrocyclic amidine and a D-Asp residue. An unusual α/β-hydrolase (ABH) fold enzyme BotH, was responsible for the epimerization of the Asp residue. Based on the biochemical and structural analysis, the mechanism of epimerization in bottromycin biosynthesis was proposed as shown in Fig. 1-8.

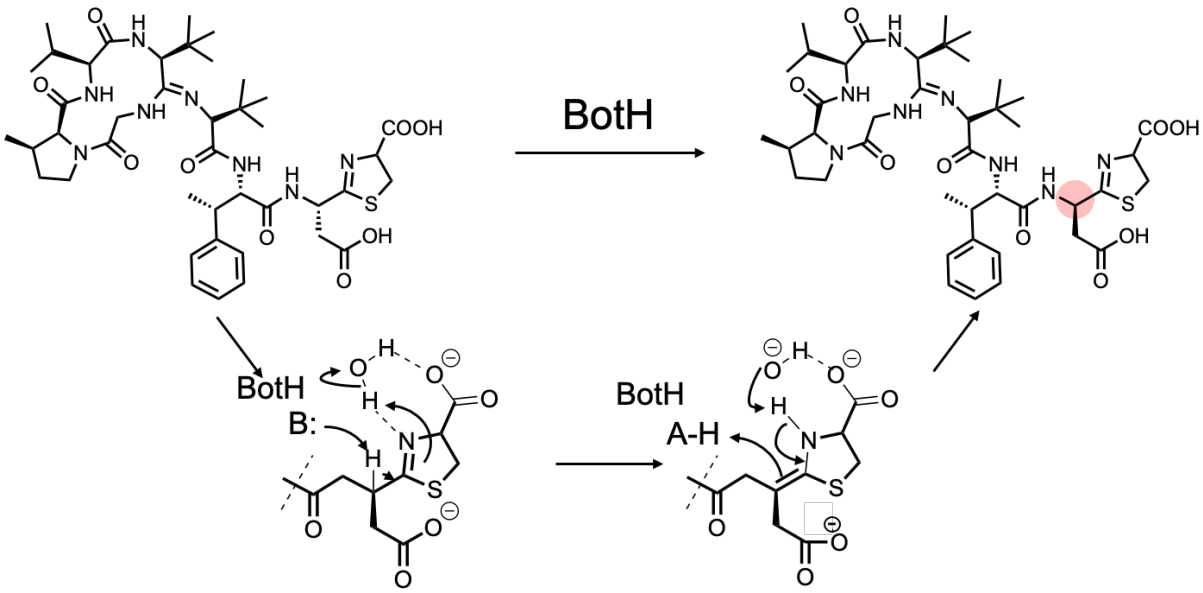


Figure 1-8 Epimerization by BotH with atypical α/β-hydrolase fold.

MS-271, originally isolated from *Streptomyces* sp. M-271, comprises 21 residues including two disulfides, as well as a C-terminal D-Trp. The biosynthetic gene cluster of MS-271 identified by genome sequence analysis revealed that the precursor peptide gene (*mslA*) encoded 42-residue peptide with a C-terminal Trp, and that a gene encoding protein of unknown function (MslH) existed in the cluster. Heterologous expression of the BGC and functional analysis of MslH suggested that MslH is a novel peptide epimerase catalyzing the epimerization at the Cα center adjacent to a carboxylic acid of C-terminal Trp. In addition, the recent X-ray crystallographic studies in association with MslA core peptide analogs, demonstrated that MslH is a metallo-dependent peptide epimerase with a calcineurin-like fold and that MslH utilizes two pairs of His/Asp catalytic residues to facilitate the reversible epimerization of the C-terminal Trp21 of MslA (Figure 1-9).

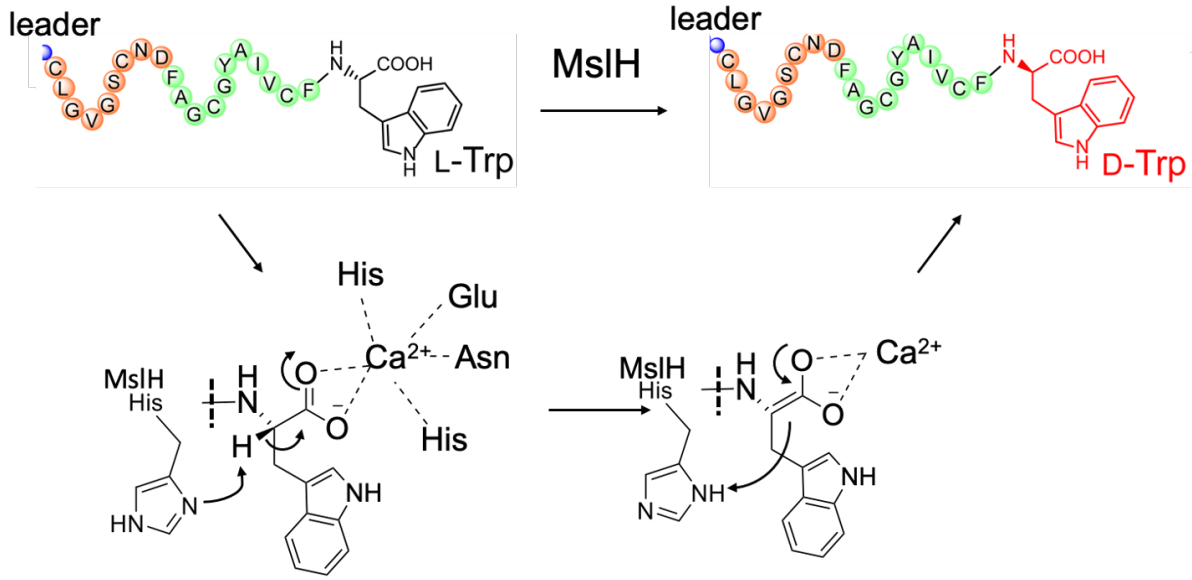


Figure 1-9 A metallo-dependent enzyme MslH responsible for the epimerization of C-terminal D-Trp in MS-271.

Some lanthipeptides contain D-Ala or D-aminobutyric acid residue, which arise from a two-step dehydration–hydrogenation process. The key enzymes are LanB and LanJ. The former dehydrates L-Ser to install Dha, and then, the latter catalyzes a stereospecific enzymatic hydrogenation event to form D-alanine (Figure 1-10).

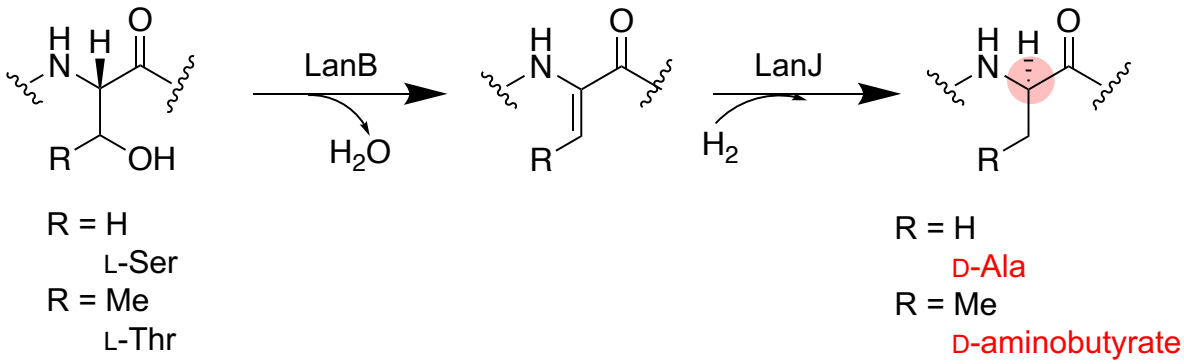


Figure 1-10 D-Amino acids formation in lanthipeptide biosynthesis.

Although salinipeptin A, a member of linaridin natural peptides, was also reported to contain nine D-amino acid residues[34,35] (Figure 1-5), the source of D-amino acids remains unclear, because the gene cluster lacks any homologs of known epimerases described above[34,36,37].

1.3 Biosynthesis of linaridins

Linaridins[37,38] (linear arid peptides) are the family of RiPPs characterized by the presence of dehydrobutyrine (Dhb). To date, there are 7 linaridins isolated, namely cypemycin[38,39], grisemycin[40], salinipeptins[34], legonaridin[41], pegvadins[35], mononaridin[42], and corynaridin[43] (Figure 1-11), however, *in silico* analysis has shown that this family of RiPPs is widespread in nature[35,37]. Linaridins are featured to contain Dhb, *N,N*-dimethylalanine, and C-terminal aminovinyl cysteine (AviCys) moiety. Interestingly, multiple D-amino acids are also contained in linaridins, nevertheless, an enzyme responsible for D-amino acids introduction remains unclear.

Type A linaridins

salinipeptin

Streptomyces sp. GSL-6C

cypemycin

Streptomyces sp. OH-4156

grisemycin

Streptomyces griseus

Type B linaridins

legonaridin

Streptomyces sp. CT34

monoaridin

Streptomyces monomycini

pegvadins

Streptomyces noursei

corynaridin

Corynebacterium lactis

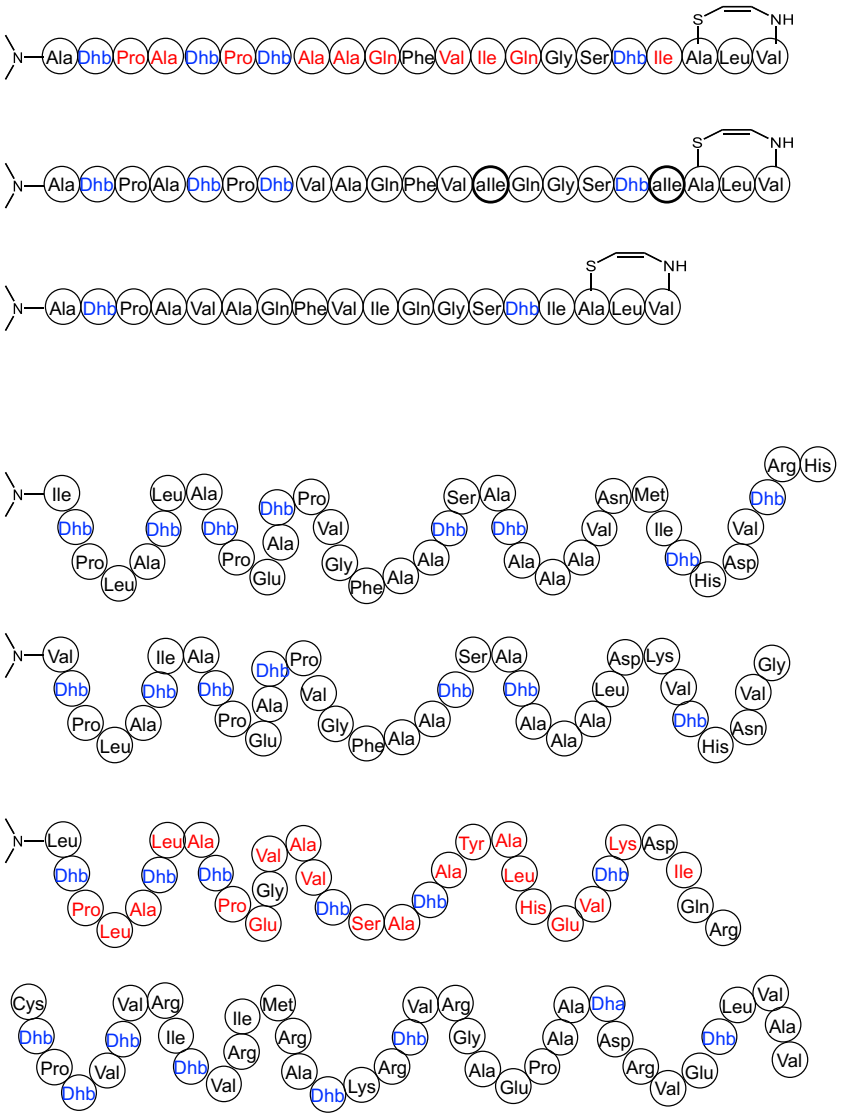


Figure 1-11 Characterized linaridins. D-amino acid in red, and dehydroamino acid in blue.

In this study, I identified a novel epimerase catalyzing the introduction of multiple D-amino acid residues in the type-A linaridins. In Chapter 2, I examined whether a novel peptide epimerase is involved in the biosynthesis of type-A linaridins by heterologous expression experiments and gene-deletion experiments. In Chapter 3, I performed *in vivo* experiments for functional analysis of the putative epimerase.

1.4 Reference:

1. Hancock, R. E. W. & Chapple, D. S. Peptide antibiotics. *Antimicrob. Agents Chemother.* **43**, 1317–1323 (1999).
2. Walsh, C. T. Polyketide and Nonribosomal Peptide Antibiotics: Modularity and Versatility. *Science (80-.).* **303**, 1805–1810 (2004).
3. Ongpipattanakul, C. *et al.* Mechanism of Action of Ribosomally Synthesized and Post-Translationally Modified Peptides. *Chem. Rev.* (2022).
4. Walsh, C. T. *et al.* Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr. Opin. Chem. Biol.* **5**, 525–534 (2001).
5. Dang, T. & Süssmuth, R. D. Bioactive Peptide Natural Products as Lead Structures for Medicinal Use. *Acc. Chem. Res.* **50**, 1566–1576 (2017).
6. Hadatsch, B. *et al.* The Biosynthesis of Teicoplanin-Type Glycopeptide Antibiotics: Assignment of P450 Mono-Oxygenases to Side Chain Cyclizations of Glycopeptide A47934. *Chem. Biol.* **14**, 1078–1089 (2007).
7. Woithe, K. *et al.* Oxidative phenol coupling reactions catalyzed by OxyB: A cytochrome P450 from the vancomycin producing organism. Implications for vancomycin biosynthesis. *J. Am. Chem. Soc.* **129**, 6887–6895 (2007).
8. Mavaro, A. *et al.* Substrate recognition and specificity of the NisB protein, the lantibiotic dehydratase involved in nisin biosynthesis. *J. Biol. Chem.* **286**, 30552–30560 (2011).
9. Ortega, M. A. *et al.* Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature* **517**, 509–512 (2015).
10. Walsh, C. T. Blurring the lines between ribosomal and nonribosomal

peptide scaffolds. *ACS Chem. Biol.* **9**, 1653–1661 (2014).

11. Mordhorst, S., Ruijne, F., Vagstad, A. L., Kuipers, O. P. & Piel, J. Emulating nonribosomal peptides with ribosomal biosynthetic strategies. *RSC Chem. Biol.* **4**, 7–36 (2022).

12. Sikandar, A. & Koehnke, J. The role of protein-protein interactions in the biosynthesis of ribosomally synthesized and post-translationally modified peptides. *Natural Product Reports* **36**, 1576–1588 (2019).

13. Süssmuth, R. D. & Mainz, A. Nonribosomal Peptide Synthesis—Principles and Prospects. *Angew. Chemie - Int. Ed.* **56**, 3770–3821 (2017).

14. Dekimpe, S. & Masschelein, J. Beyond peptide bond formation: The versatile role of condensation domains in natural product biosynthesis. *Nat. Prod. Rep.* **38**, 1910–1937 (2021).

15. Hur, G. H., Vickery, C. R. & Burkart, M. D. Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Natural Product Reports* **29**, 1074–1098 (2012).

16. Arnison, P. G. *et al.* Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **30**, 108–160 (2013).

17. Montalbán-López, M. *et al.* New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rep.* **38**, 130–239 (2021).

18. Winter, J. M. & Tang, Y. Natural products: Getting a handle on peptides. *Nat. Chem.* **6**, 1037–1038 (2014).

19. Busche, T. *et al.* Multi-omics and targeted approaches to determine the role of cellular proteases in streptomyces protein secretion. *Front. Microbiol.* **9**, 1–12 (2018).

20. Qiu, Y., Liu, J., Li, Y., Xue, Y. & Liu, W. Formation of an aminovinyl-

cysteine residue in thioviridamides occurs through a path independent of known lanthionine synthetase activity. *Cell Chem. Biol.* **28**, 675-685.e5 (2021).

21. Zerbe, K. *et al.* An oxidative phenol coupling reaction catalyzed by OxyB, a cytochrome P450 from the vancomycin-producing microorganism. *Angew. Chemie - Int. Ed.* **43**, 6709–6713 (2004).

22. Strieker, M. & Marahiel, M. A. The structural diversity of acidic lipopeptide antibiotics. *ChemBioChem* **10**, 607–616 (2009).

23. Stachelhaus, T. & Walsh, C. T. Mutational analysis of the epimerization domain in the initiation module PheATE of gramicidin S synthetase. *Biochemistry* **39**, 5775–5787 (2000).

24. Ogasawara, Y. & Dairi, T. Peptide epimerization machineries found in microorganisms. *Front. Microbiol.* **9**, 1–9 (2018).

25. Bhushan, A., Egli, P. J., Peters, E. E., Freeman, M. F. & Piel, J. Genome mining- and synthetic biology-enabled production of hypermodified peptides. *Nat. Chem.* **11**, 931–939 (2019).

26. Morinaka, B. I. *et al.* Radical S-adenosyl methionine epimerases: Regioselective introduction of diverse D -amino acid patterns into peptide natural products. *Angew. Chemie - Int. Ed.* **53**, 8503–8507 (2014).

27. Parent, A. *et al.* Mechanistic Investigations of PoyD, a Radical S-Adenosyl-l -methionine Enzyme Catalyzing Iterative and Directional Epimerizations in Polytheonamide A Biosynthesis. *J. Am. Chem. Soc.* **140**, 2469–2477 (2018).

28. Zhu, W. & Klinman, J. P. Biogenesis of the peptide-derived redox cofactor pyrroloquinoline quinone. *Curr. Opin. Chem. Biol.* **59**, 93–103 (2020).

29. Franz, L., Kazmaier, U., Truman, A. W. & Koehnke, J. Bottromycins -

Biosynthesis, synthesis and activity. *Nat. Prod. Rep.* **38**, 1659–1683 (2021).

30. Sikandar, A. *et al.* The bottromycin epimerase BotH defines a group of atypical α/β-hydrolase-fold enzymes. *Nat. Chem. Biol.* **16**, 1013–1018 (2020).

31. Feng, Z., Ogasawara, Y. & Dairi, T. Identification of the peptide epimerase MslH responsible ford-amino acid introduction at the C-terminus of ribosomal peptides. *Chem. Sci.* **12**, 2567–2574 (2021).

32. Nakashima, Y., Kawakami, A., Ogasawara, Y. *et al*. Structure of lasso peptide epimerase MslH reveals metal-dependent acid/base catalytic mechanism. *Nat Commun* **14**, 4752 (2023).

33. Ayikpoe, R. S. & van der Donk, W. A. *Peptide backbone modifications in lanthipeptides*. *Methods in Enzymology* **656**, (Elsevier Inc., 2021).

34. Shang, Z., Winter, J. M., Kauffman, C. A., Yang, I. & Fenical, W. Salinipeptins: Integrated Genomic and Chemical Approaches Reveal Unusual d -Amino Acid-Containing Ribosomally Synthesized and Post-Translationally Modified Peptides (RiPPs) from a Great Salt Lake Streptomyces sp. *ACS Chem. Biol.* **14**, 415–425 (2019).

35. Georgiou, M. A., Dommaraju, S. R., Guo, X., Mast, D. H. & Mitchell, D. A. Bioinformatic and reactivity-based discovery of linaridins. *ACS Chem. Biol.* **15**, 2976–2985 (2020).

36. Xiao, W., Satoh, Y., Ogasawara, Y. & Dairi, T. Biosynthetic Gene Cluster of Linaridin Peptides Contains Epimerase Gene. *ChemBioChem* **23**, (2022).

37. Ma, S. & Zhang, Q. Linaridin natural products. *Nat. Prod. Rep.* **37**, 1152–1163 (2020).

38. Claesen, J. & Bibb, M. Genome mining and genetic analysis of cypemycin biosynthesis reveal an unusual class of posttranslationally modified

peptides. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 16297–16302 (2010).

39. Kanki, K. *et al.* New Antibiotic, Cypemycin Taxonomy, Fermentation, Isolation and Biological Characteristics. *J. Antibiot. (Tokyo).* **46**, 1666–1671 (1993).

40. Claesen, J. & Bibb, M. J. Biosynthesis and regulation of grisemycin, a new member of the linaridin family of ribosomally synthesized peptides produced by Streptomyces griseus IFO 13350. *J. Bacteriol.* **193**, 2510–2516 (2011).

41. Rateb, M. E. *et al.* Legonaridin, a new member of linaridin RiPP from a Ghanaian Streptomyces isolate. *Org. Biomol. Chem.* **13**, 9585–9592 (2015).

42. Wang, F. *et al.* Genome Mining and Biosynthesis Study of a Type B Linaridin Reveals a Highly Versatile α-N-Methyltransferase. *CCS Chem.* **3**, 1049–1057 (2021).

43. Pashou, E. *et al.* Identification and Characterization of Corynaridin, a Novel Linaridin from Corynebacterium lactis. *Microbiol. Spectr.* **11**, (2023).

Chapter 2

Involvement of A New Type of Peptide Epimerase

2.1 Introduction

As mentioned in Chapter 1, salinipeptins, produced by halotolerant *Streptomyces* sp. GSL-6C, are members of the linaridin class of natural products[1]. Although salinipeptins possess multiple D-amino acid residues that were introduced via epimerization at the α-carbon of amino acid residues of the precursor peptide, the *sin* cluster does not contain an obvious candidate gene for the epimerization. Considering that SinH, SinD, and SinM are required for the formation of the dehydrobutyrine (Dhb), AviCys, and *N*,*N*-dimethylalanine, respectively[1,2], the sole function of an unknown protein, SinL, may be involved in the epimerization during salinipeptin biosynthesis (Figure 2-1).

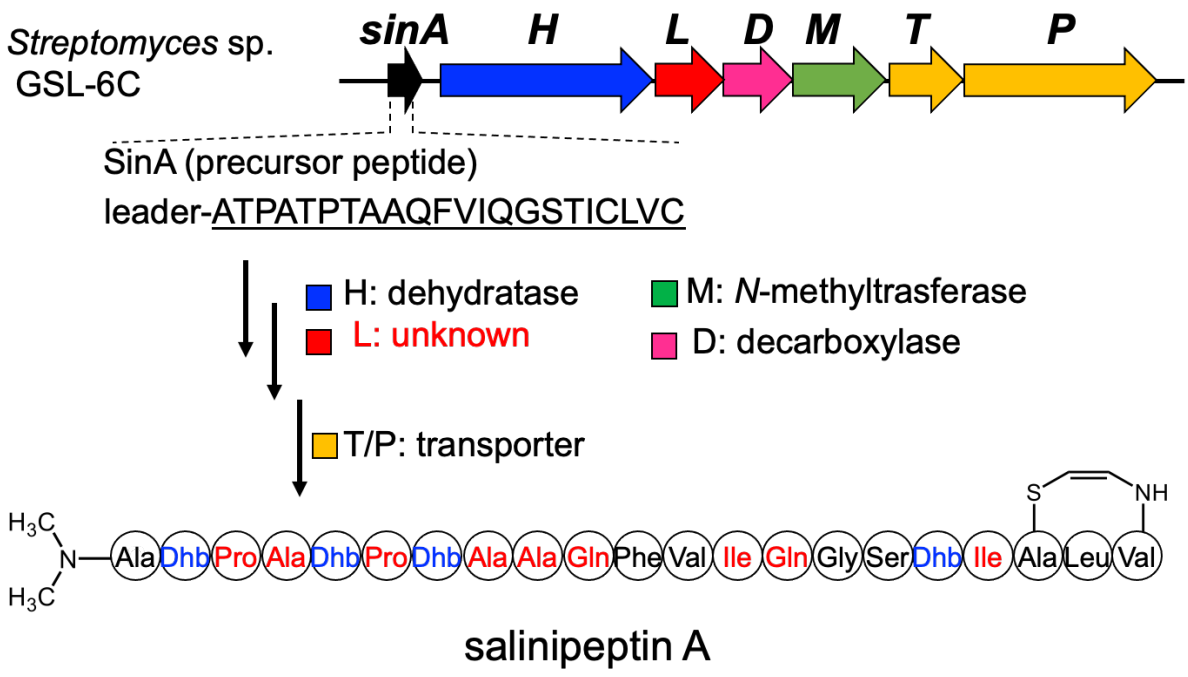


Figure 2-1 Proposed biosynthesis of salinipeptin A. D-amino acids in red, dehydroamino acids in blue. Dhb, dehydrobutyrine.

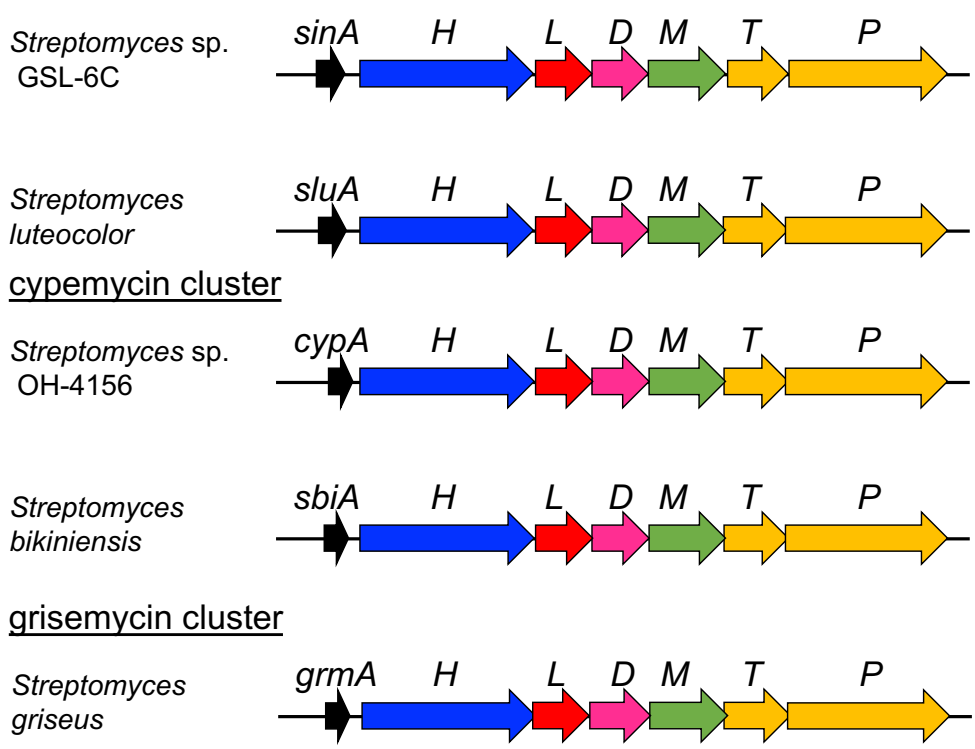
Interestingly, bioinformatics analysis revealed that gene clusters with similar gene organization were widely distributed in many different bacterial strains and that biosynthetic gene clusters of type-A linaridins of grisemycin[3] (*grm*) and cypemycin[4] (*cyp*) had orthologs organized similar to those of the *sin* cluster (Figure 2-2). Although the absolute stereochemistry of amino acid residues in grisemycin and cypemycin has not been reported, cypemycin contained *allo*-Ile, an epimer of Ile.

In this chapter, I examined the plausibility that a novel peptide epimerase is involved in type-A linaridins biosynthesis, and that the presence of D-amino acid residues is a common feature of type-A linaridin natural products.

2.2 Results and Discussion

To test the hypothesis, I performed heterologous expression experiments of linaridin clusters and examined whether the metabolites contained D-amino acid residues. Because I could not obtain producing strains of salinipeptins (*Streptomyces* sp. GSL-6C) and cypemycin (*Streptomyces* sp. OH4156), I first examined the metabolites produced by grisemycin (*grm*) cluster with producing strain *S. griseus* NBRC 13350 in hand. Furthermore, based on the deduced amino acid sequence of the precursor peptide, I selected the *sbi* cluster (*S. bikiniensis* NBRC 14598) and *slu* cluster (*S. luteocolor* NBRC 13826), whose metabolites would be cypemycin- and salinipeptin-like linaridins, respectively (Figure 2-2, Table 2-1).

A



B

SinA	MRSELTSTRPDSSATHPESLAAQEFANTALGAATPGFHADCETPAMATPATPTAAQFVIQGSTICLVC
SluA	MQSESTVTRPDS-AIRPEALAAQEFANTVLSGAAPGFHADCETPAMATPATPTAAQFVIQGSTICLVC
CypA	----MRSEMTLTSTNSAEALAAQDFANTVLSAAAPGFHADCETPAMATPATPTVAQFVIQGSTICLVC
SbiA	------MTRQDSTAVHPEVLAAQDFANTVLEGAAPGFHSNCETPAMATPATPTVAQFVIQGSTICLVC
GrmA	MRLDS-IATQETATALPESMATQDFANSVLAGAVPGFHSDAETPAMATP---AVAQFVIQGSTICLVC

Figure 2-2 (A) Organization of the type-A linaridin BGC, and (B) deduced amino acid sequences of precursor peptides, core peptides are underlined.

Table 2-1 Amino acid identities (%) of biosynthetic enzymes found in the linaridin clusters.

	CypH	**SbiH**	**SinH**	**SluH**
GrmH	74.1	74.2	73.7	74.2
CypH	–	82.9	82.1	86.3
SbiH	–	–	81.5	83.8
SinH	–	–	–	84.1
	CypL	**SbiL**	**SinL**	**SluL**
GrmL	73.0	74.6	77.8	77.2
CypL	–	86.1	85.6	90.4
SbiL	–	–	86.1	89.8
SinL	–	–	–	90.9
	CypD	**SbiD**	**SinD**	**SluD**
GrmD	75.4	75.4	73.8	76.4
CypD	–	83.8	82.2	92.1
SbiD	–	–	82.2	85.9
SinD	–	–	–	86.4
	CypM	**SbiM**	**SinM**	**SluM**
GrmM	76.0	74.0	73.6	76.0
CypM	–	87.0	84.6	87.8
SbiM	–	–	82.9	86.2
SinM	–	–	–	85.4
	CypT	**SbiT**	**SinT**	**SluT**
GrmT	74.4	74.9	73.5	77.7
CypT	–	86.7	83.9	88.6
SbiT	–	–	83.4	86.7
SinT	–	–	–	88.6
	CypP	**SbiP**	**SinP**	**SluP**
GrmP	61.3	61.5	59.8	62.9
CypP	–	80.3	75.8	83.0
SbiP	–	–	74.5	78.9
SinP	–	–	–	79.9

2.2.1 Heterologous expression of grisemycin biosynthetic gene cluster

The *grm* cluster (7 kbp) of *S. griseus* NBRC 13350 was cloned into pWHM3[5] to generate plasmid pWHM3-grm and then introduced into *S. lividans* TK23. After cultivation of the resulting transformants, metabolites were extracted from 50 mL of culture broth with chloroform (50 mL × 3), concentrated in vacuo, and re-dissolved in 0.1% aqueous formic acid (500 μL). LC-MS analysis of the chloroform extract revealed the production of specific metabolite (**1**) at a retention time of 22.2 min (Figure 2-3 trace b).

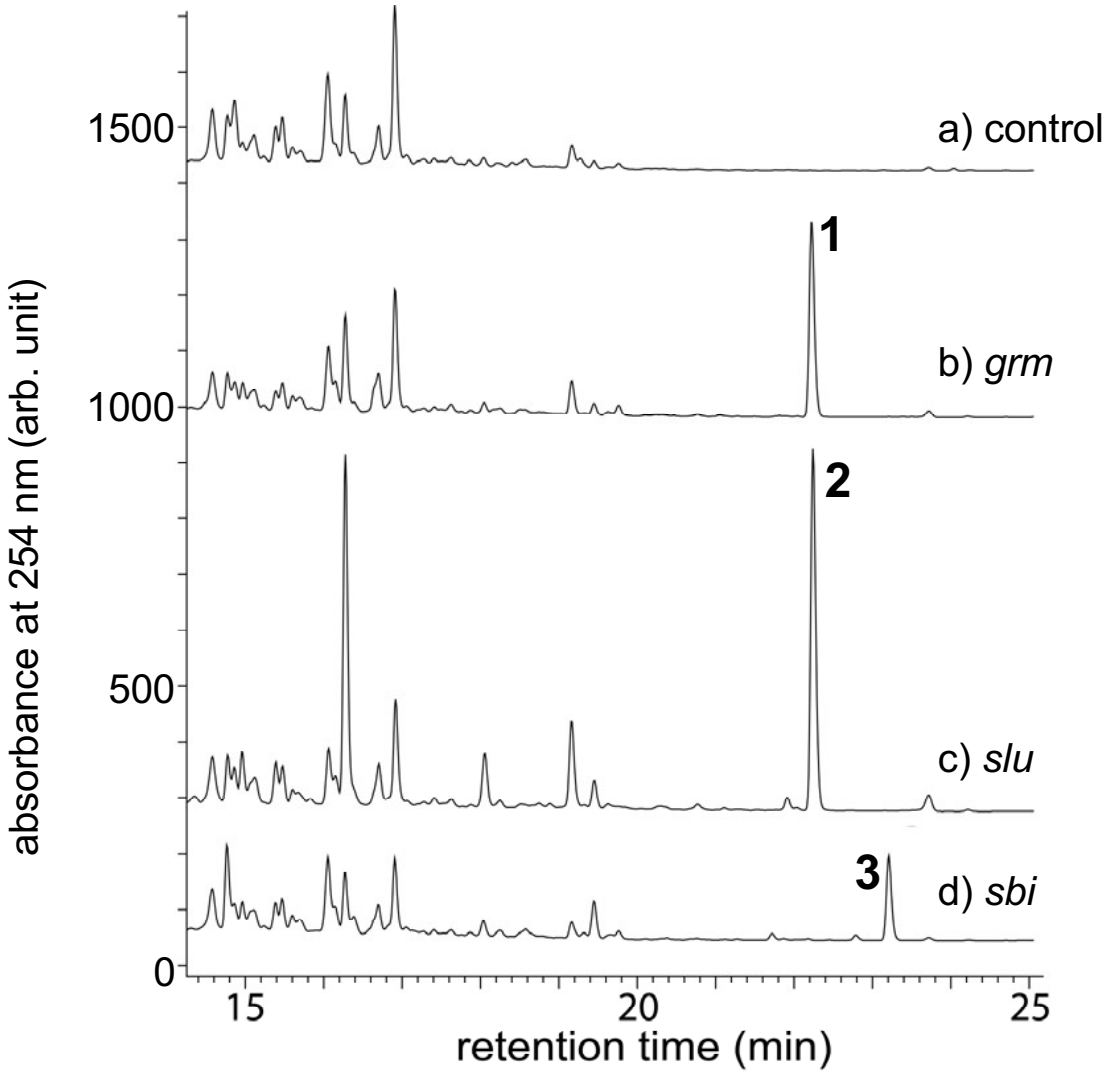
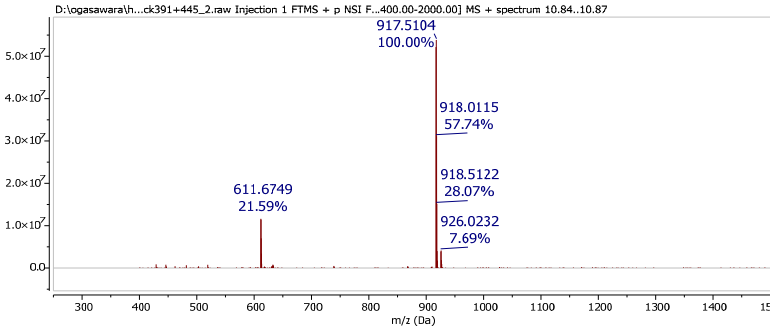


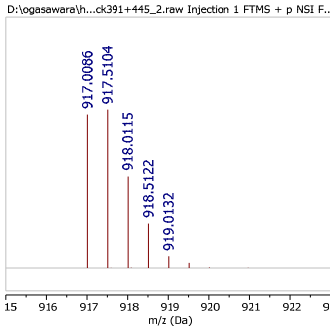
Figure 2-3 LC-MS analysis of the chloroform extracts of the culture broths. a) *S. lividans*/pWHM3 (empty vector control), b) *S. lividans*/pWHM3-grm, c) *S. lividans*/pWHM3-slu, d) *S. lividans*/pWHM3-sbi.

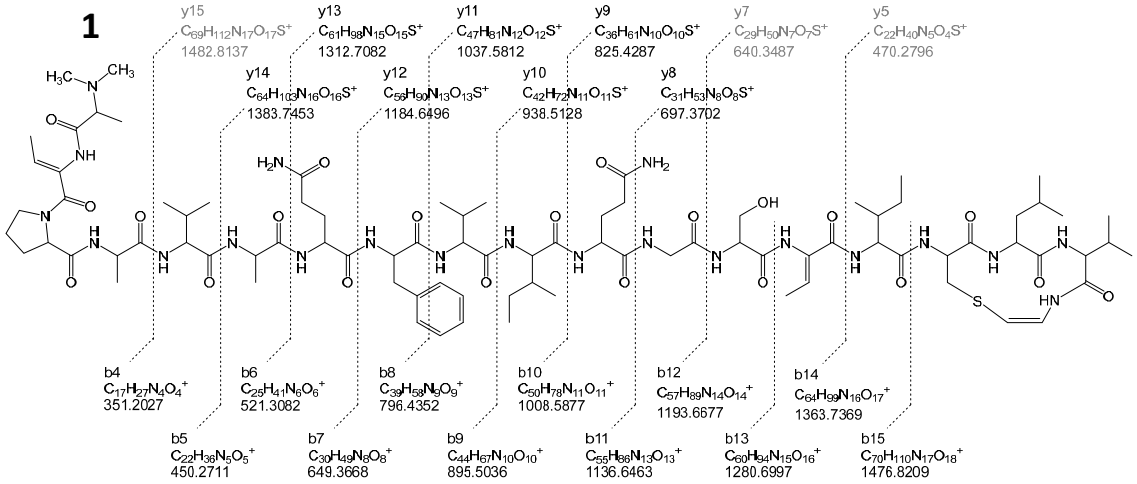
High-resolution electrospray ionization (HR-ESI) MS and MS/MS analysis of the compound showed that a doubly protonated molecular ion at *m*/*z* 917.0086 ($C_{86}H_{137}N_{21}O_{21}S$: calcd. 917.0082 $[M+2H]^{2+}$) and the fragment ions (*b* and *y* ions) were fully consistent with the theoretical values for grisemycin (Figure 2-4).

HR-MS spectrum



HR-MS spectrum (enlarged)





HR-MS/MS spectrum

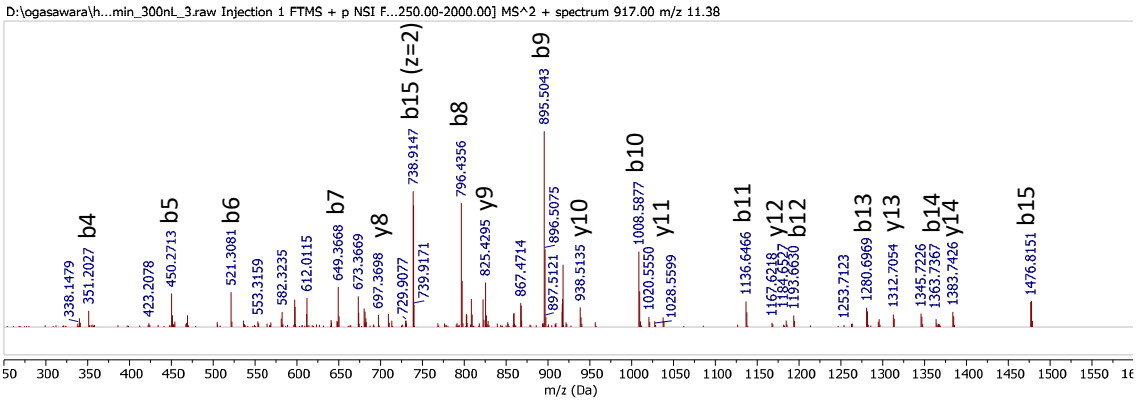


Figure 2-4 HR-MS and MS/MS spectra of compound **1**.

The planar structure of **1** was further confirmed by NMR analysis using 0.5 mg of **1** purified from large-scale cultures (50 mL × 100 flasks). The ^{1}H and ^{13}C-NMR data, NMR spectra (^{1}H, ^{13}C, COSY, TOCSY, HSQC, HSQC-TOCSY, HMBC, and ROESY), and 2D NMR correlation are summarized in Table 2-2 and Figures 2-5–2-14. The presence of 18 amino acid residues (Ala ×3, Dhb ×2, Pro ×1, Val ×3, Gln ×2, Phe ×1, Ile/*allo*-Ile ×2, Gly ×1, Ser ×1, Leu ×1, and AviCys ×1) was confirmed mainly by COSY, TOCSY, and HSQC-TOCSY analysis, and the sequence of these amino acids was established with HMBC and ROESY analysis. For nonproteinogenic amino acid residues, the *N*-terminal dimethyl group was confirmed with HMBC correlation from the *N*-CH 3 (δ_H 2.86, 6H) to C_α of Ala-1 (δ_C 64.05). In addition, the geometry of the double bond in AviCys-16 was deduced to be *Z*, based on the small coupling constant of 7.0 Hz between H-4 (δ_H 5.48) and H-5 (δ_H 7.11), as well as Nuclear Overhauser Effect (NOE) correlations between them. The geometry of the double bonds in both Dhb residues was also deduced to be *Z* because NOE correlations between H-4 (δ_H 1.81) and NH (δ_H 9.90) in Dhb-2, as well as H-4 (δ_H 1.78) and NH (δ_H 10.00) in Dhb-14, were observed. It is worth noting that both cypemycin and salinipeptins contained (*Z*)-Dhb residues[1,4].

Table 2-2 NMR data of **1**.

position		δH	δC
Ala-1	NCH_3	2.86 (s, 6H)	41.32
	1	—	169.00
	2	4.31 (m, 1H)	64.05
	3	1.65 (d, *J* = 7.0 Hz, 3H)	13.67
Dhb-2	NH	9.90 (s, 1H)	—
	1	-	167.75
	2	-	130.85
	3	6.08 (q, *J* = 7.5 Hz, 1H)	128.48
	4	1.81 (d, *J* = 7.2 Hz, 3H)	12.72
Pro-3	1	—	175.26
	2	4.22 (m, 1H)	64.30
	3	2.33 (m, 1H) 1.89 (m, 1H)	30.05
	4	2.04 (m, 1H) 1.87 (m, 1H)	26.60
	5	3.58 (m, 1H) 3.47 (m, 1H)	50.04
Ala-4	NH	7.17 (m, 1H)	—
	1	—	176.87
	2	3.88 (m, 1H)	52.88
	3	1.37 (d, *J* = 7.4 Hz, 3H)	16.78
Val-5	NH	7.15 (m, 1H)	—
	1	—	175.30
	2	3.44 (m, 1H)	64.30
	3	2.07 (m, 1H)	29.71
	4	0.86 (m, 3H)	19.27
	5	0.86 (m, 3H)	20.33
Ala-6	NH	7.86 (m, 1H)	—
	1	—	177.01
	2	3.94 (m, 1H)	53.48
	3	1.39 (d, *J* = 7.3 Hz, 3H)	16.48
Gln-7	NH	7.86 (m, 1H)	—
	1	—	175.67
	2	3.80 (m, 1H)	57.51
	3	1.96 (m, 1H) 1.87 (m, 1H)	26.33
	4	2.31 (m, 2H)	31.53
	5	—	175.72

position		δH	δC
Phe-8	NH	8.26 (d, *J* = 7.1 Hz, 1H)	—
	1	—	173.95
	2	4.00 (m, 1H)	56.23
	3	3.18 (m, 2H)	35.08
	4	—	139.55
	5	7.17 (m, 2H)	130.54
	6	7.21 (m, 2H)	129.07
	7	7.15 (m, 1H)	127.05
Val-9	NH	8.18 (d, *J* = 4.4 Hz, 1H)	—
	1	—	176.71
	2	3.48 (m, 1H)	65.54
	3	2.17 (m, 1H)	30.10
	4	0.71 (d, *J* = 6.7 Hz, 3H)	19.49
	5	1.00 (m, 3H)	21.03
allo-Ile-10	NH	8.30 (d, *J* = 5.0 Hz, 1H)	—
	1	—	178.47
	2	3.68 (m, 1H)	64.81
	3	2.00 (m, 1H)	36.32
	4	1.51 (m, 1H) 1.17 (m, 1H)	26.33
	5	0.91 (t, *J* = 7.4 Hz, 3H)	10.92
	6	1.03 (d, *J* = 6.5 Hz, 3H)	17.22
Gln-11	NH	9.14 (d, *J* = 3.2 Hz, 1H)	—
	1	—	174.02
	2	4.11 (m, 1H)	57.83
	3	2.22 (m, 2H)	27.06
	4	2.58 (m, 1H) 2.47 (m, 1H)	33.25
	5	—	176.21
Gly-12	NH	7.96 (t, *J* = 6.2 Hz, 1H)	—
	1	—	170.51
	2	4.21 (m, 1H) 3.66 (m, 1H)	43.45
Ser-13	NH	7.63 (d, *J* = 6.6 Hz, 1H)	—
	1	—	176.60
	2	4.25 (m, 1H)	59.02
	3	4.29 (m, 1H) 3.94 (m, 1H)	63.53

position		δH	δC
Dhb-14	NH	10.00 (s, 1H)	—
	1	—	167.50
	2	—	130.85
	3	6.72 (m, 1H)	133.82
	4	1.78 (d, *J* = 7.1 Hz, 3H)	13.40
allo-Ile-15	NH	7.54 (d, *J* = 5.3 Hz, 1H)	—
	1	—	172.10
	2	4.09 (m, 1H)	60.09
	3	2.00 (m, 1H)	36.16
	4	1.34 (m, 1H) 1.28 (m, 1H)	26.85
	5	0.84 (m, 3H)	11.95
	6	0.98 (m, 3H)	15.59
aviCys-16	NαH	7.28 (d, *J* = 10.1 Hz, 1H)	—
	1	—	170.84
	2	4.98 (ddd, *J* = 10.1, 4.9, 1.8 Hz, 1H)	52.50
	3	3.16 (m, 1H) 3.04 (dd, *J* = 13.8, 4.9 Hz, 1H)	38.32
	4	5.48 (d, *J* = 7.0 Hz, 1H)	101.19
	5	7.11 (dd, *J* = 11.2, 7.0 Hz, 1H)	132.92
	NεH	8.82 (d, *J* = 11.2 Hz, 1H)	—
Leu-17	NH	7.01 (d, *J* = 4.0 Hz, 1H)	—
	1	—	174.81
	2	3.90 (m, 1H)	56.97
	3	1.68 (m, 1H) 1.49 (m, 1H)	41.10
	4	1.71 (m, 1H)	25.59
	5	0.88 (d, *J* = 6.4 Hz, 3H)	22.20
	6	0.99 (m, 3H)	22.80
Val-18	NH	7.36 (d, *J* = 9.3 Hz, 1H)	—
	1	—	171.45
	2	3.86 (m, 1H)	62.50
	3	2.07 (m, 1H)	30.10
	4	0.96 (m, 3H)	19.66
	5	0.86 (m, 3H)	20.63

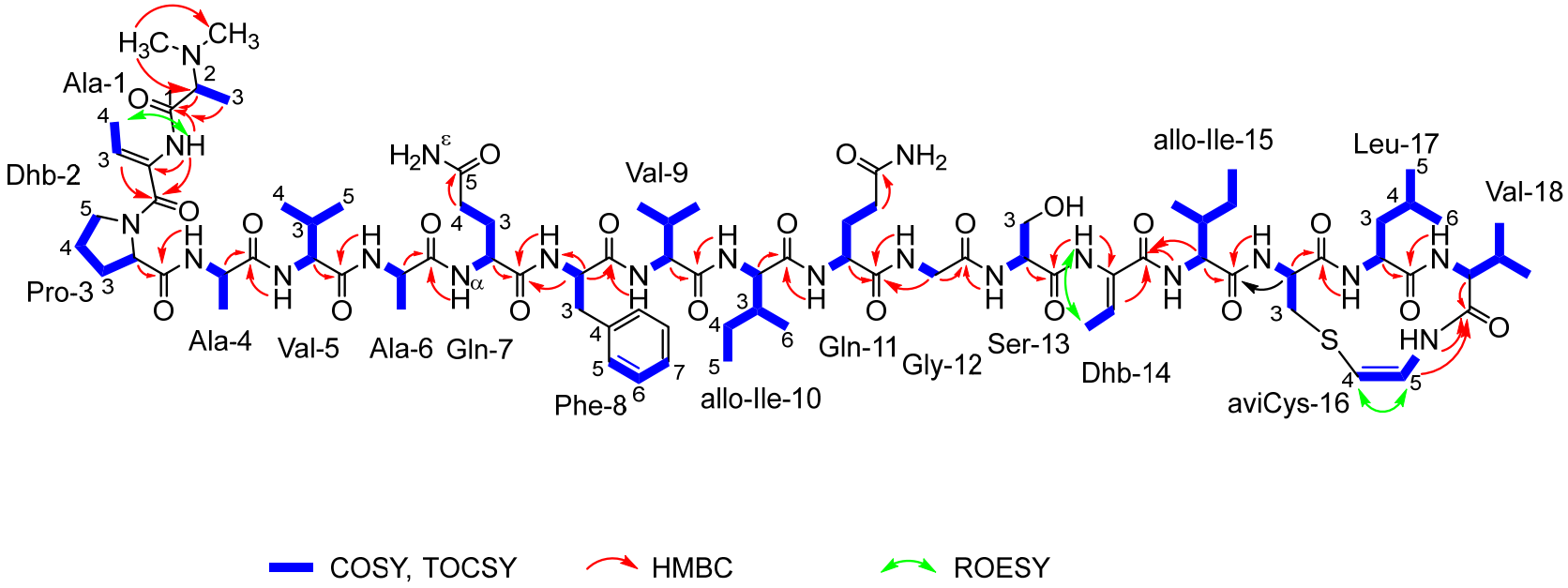


Figure 2-5 Selected 2D-NMR correlations of **1**.

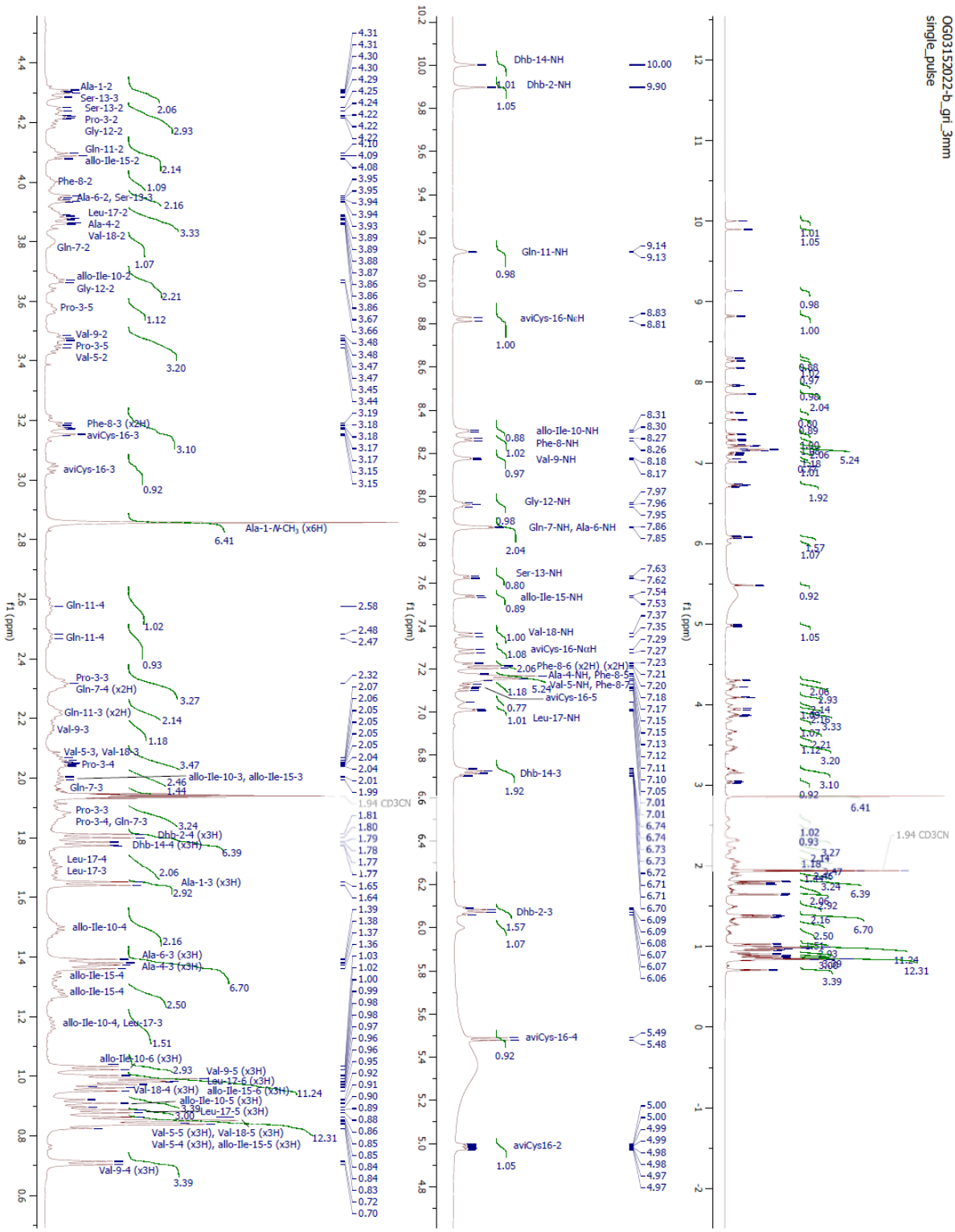


Figure 2-6 ^{1}H NMR spectrum of **1** (CD_3CN, 600 MHz).

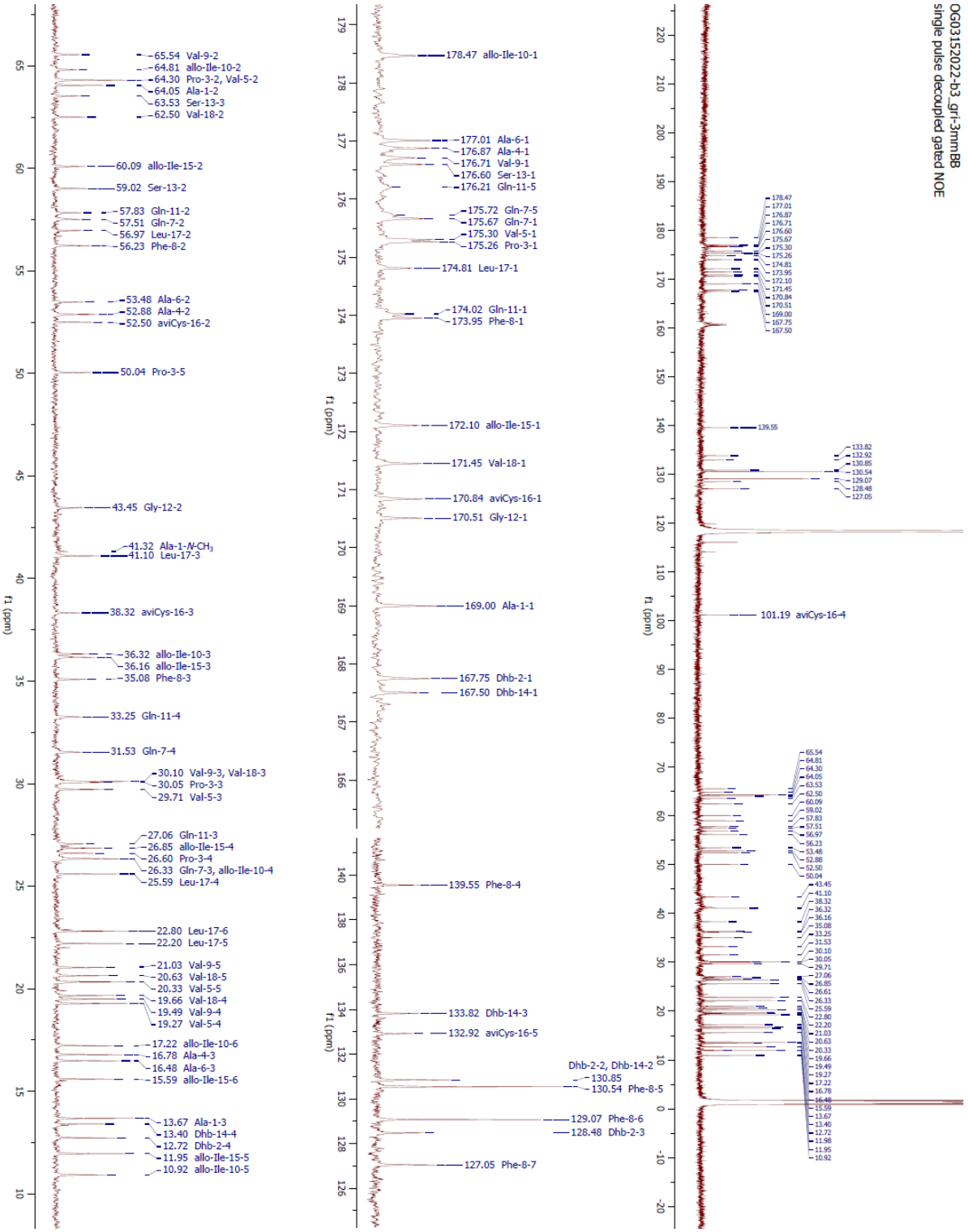


Figure 2-7 ^{13}C NMR spectrum of **1** (CD_3CN, 150 MHz).

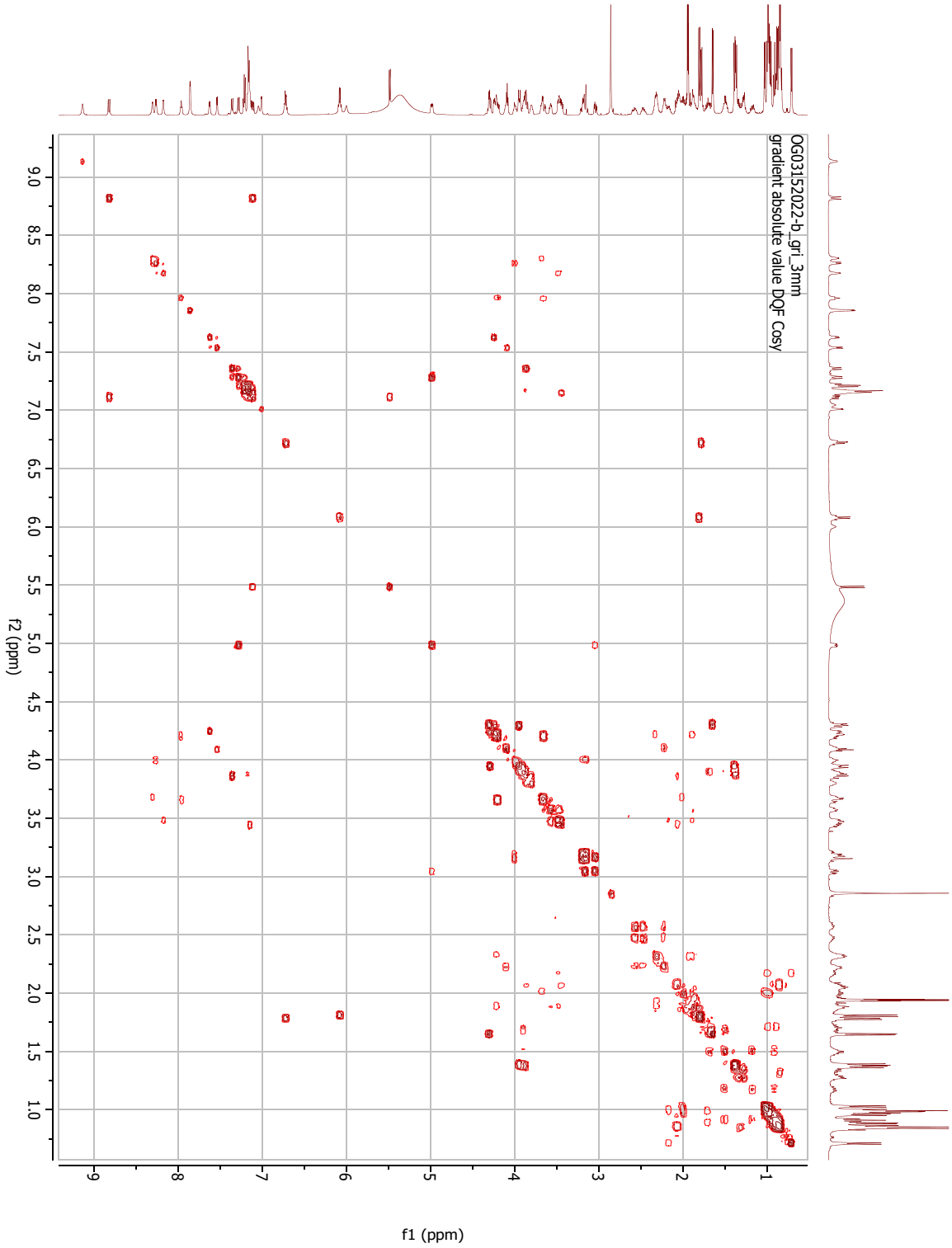


Figure 2-8 COSY spectrum of **1**.

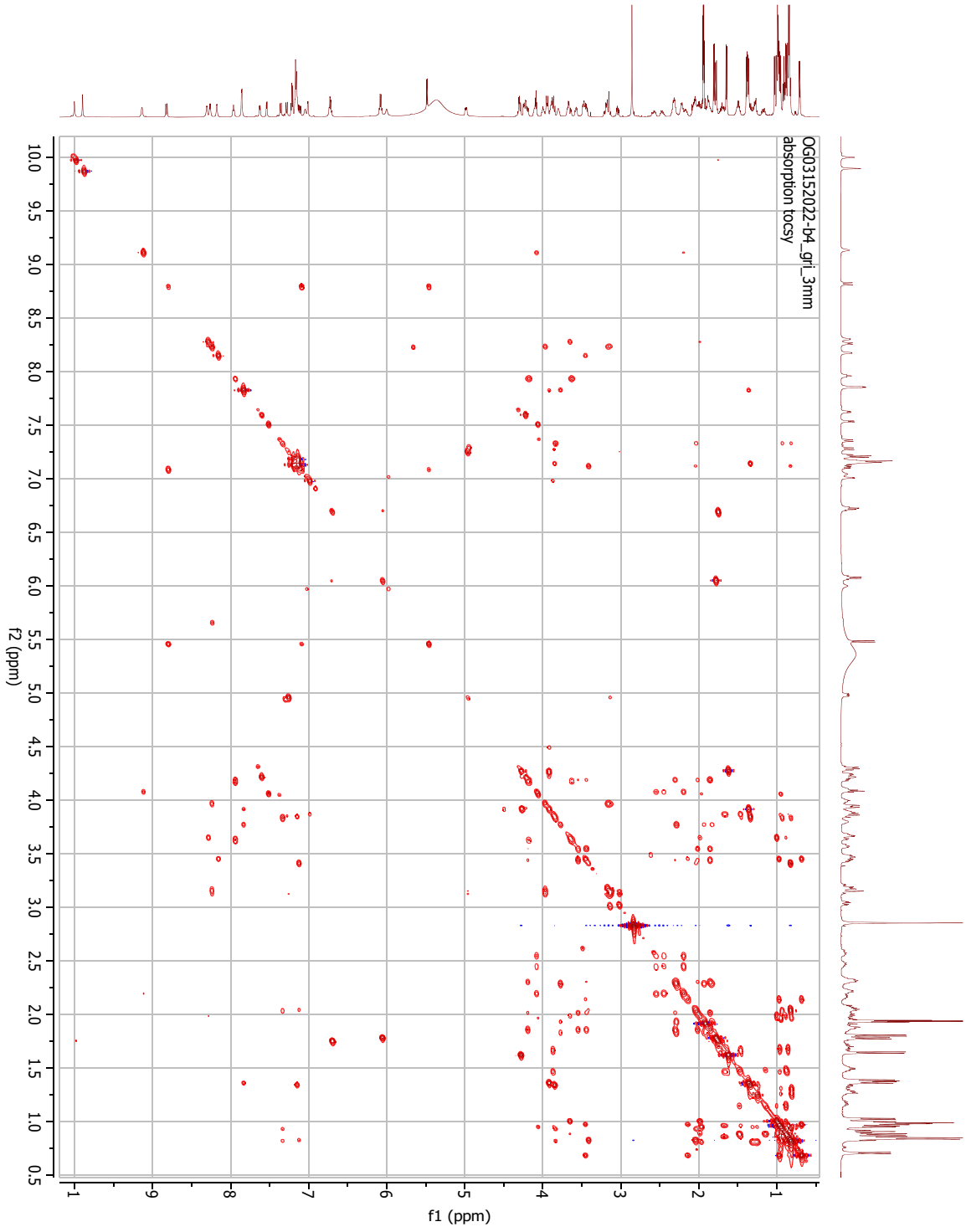


Figure 2-9 TOCSY spectrum of **1**.

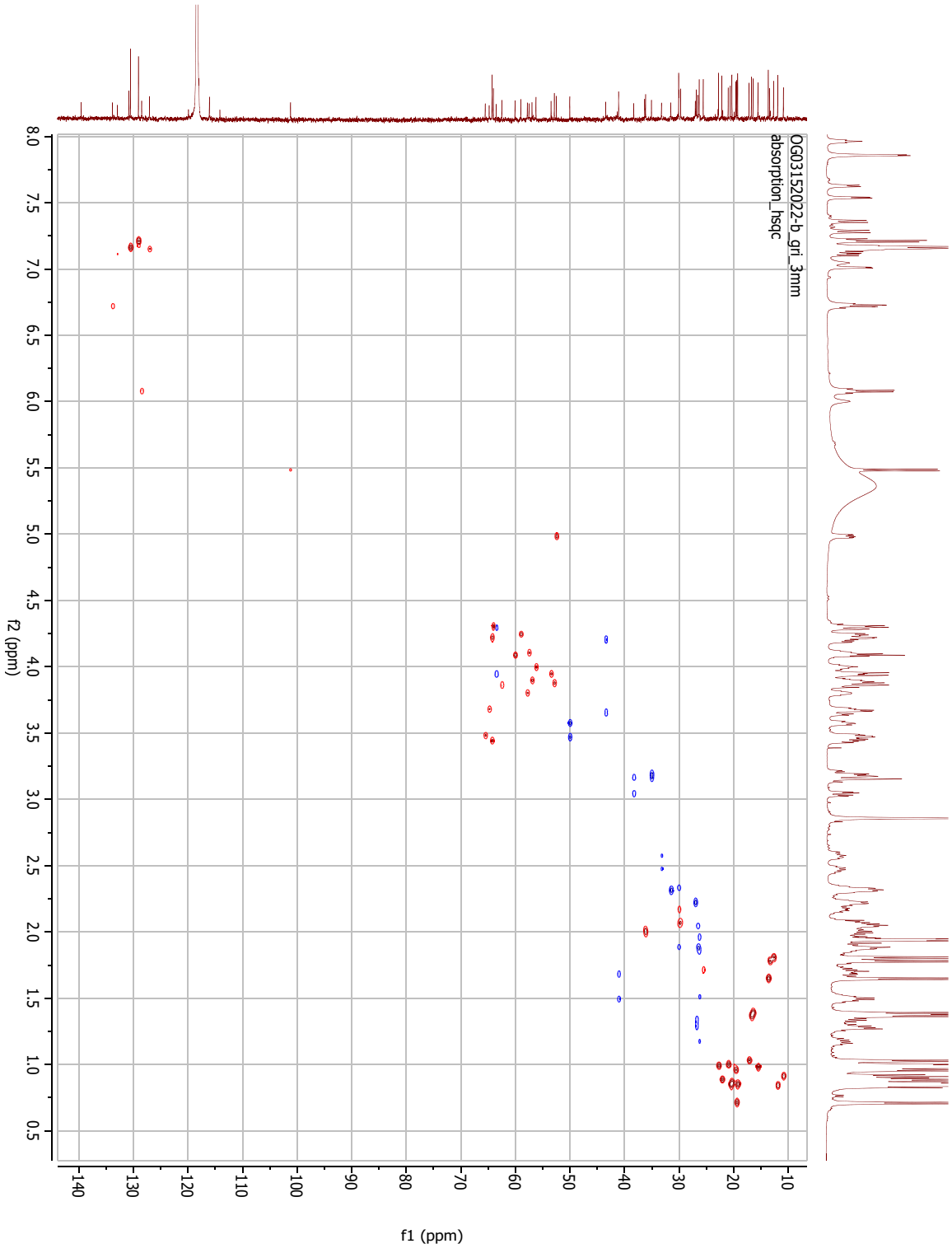


Figure 2-10 HSQC spectrum of **1**.

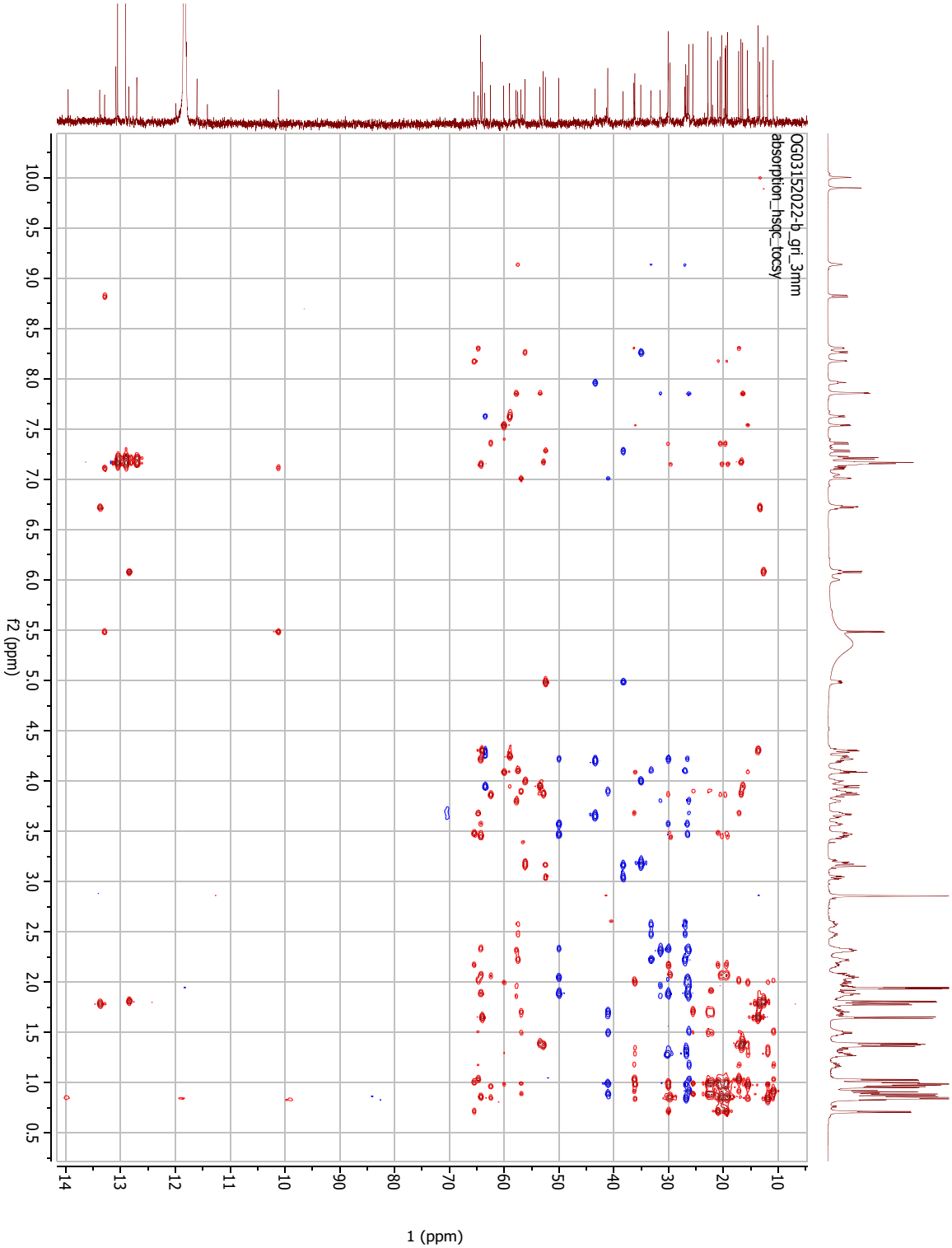


Figure 2-11 HSQC-TOCSY spectrum of **1**.

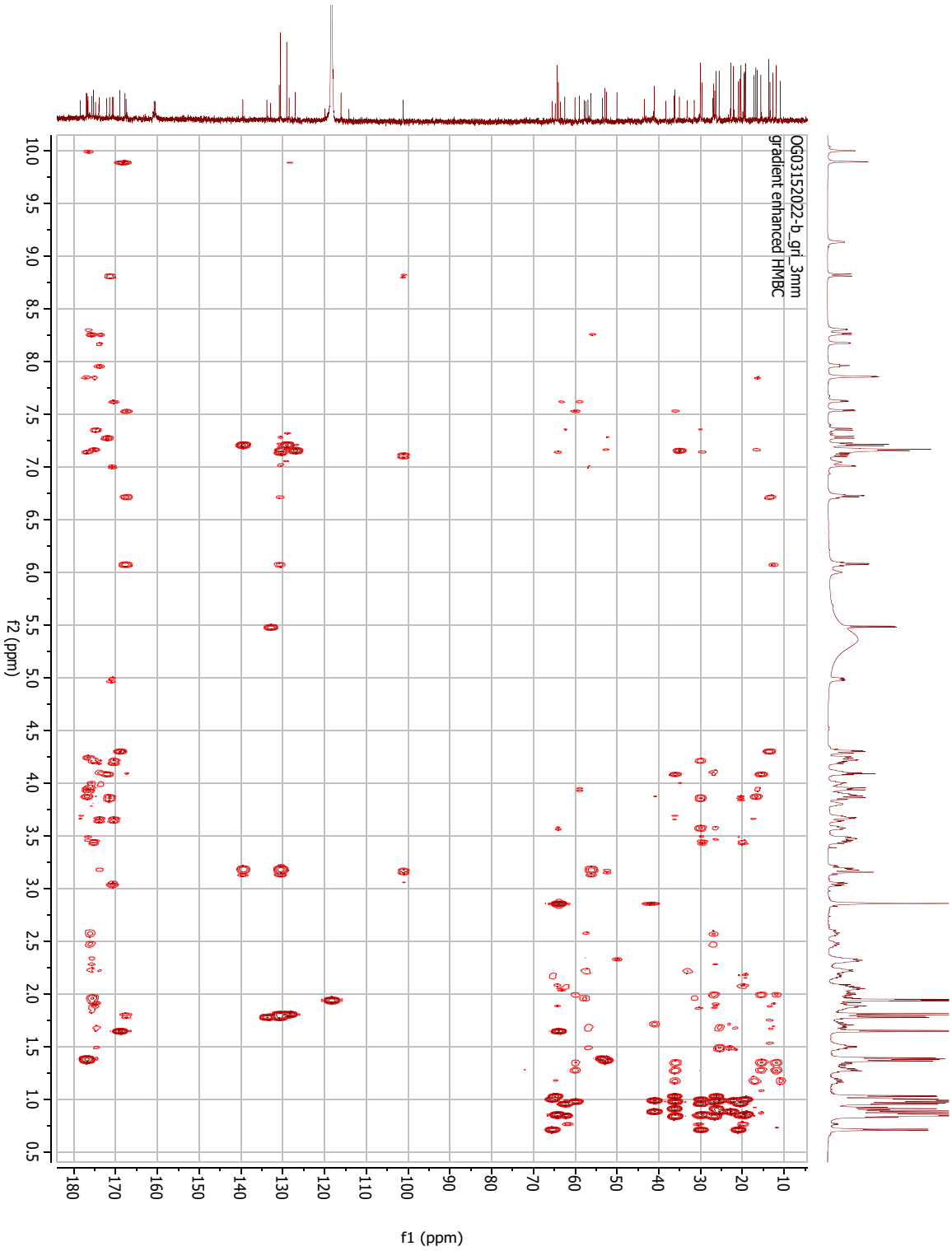


Figure 2-12 HMBC spectrum of **1**.

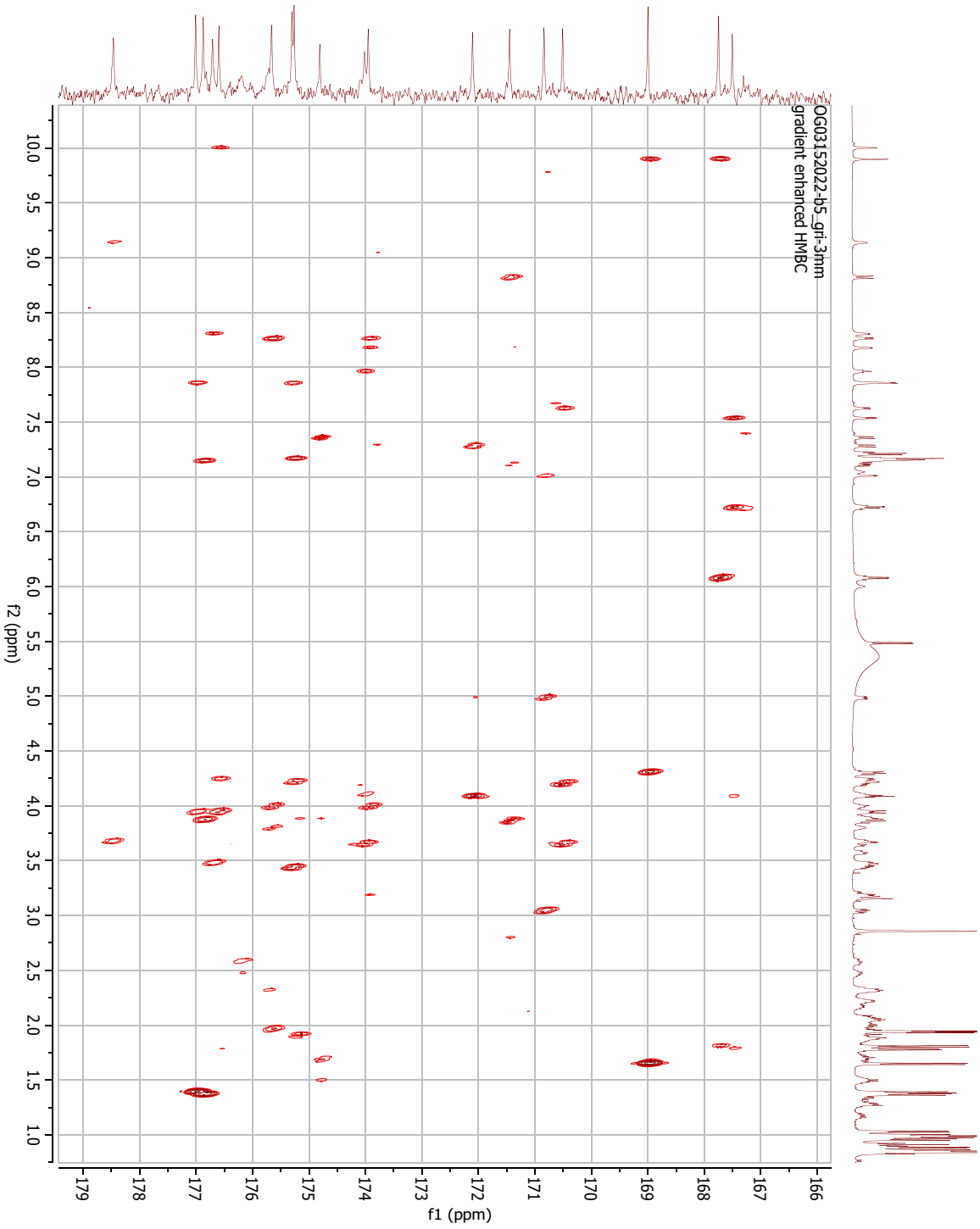


Figure 2-13 Band selective HMBC spectrum of **1**.

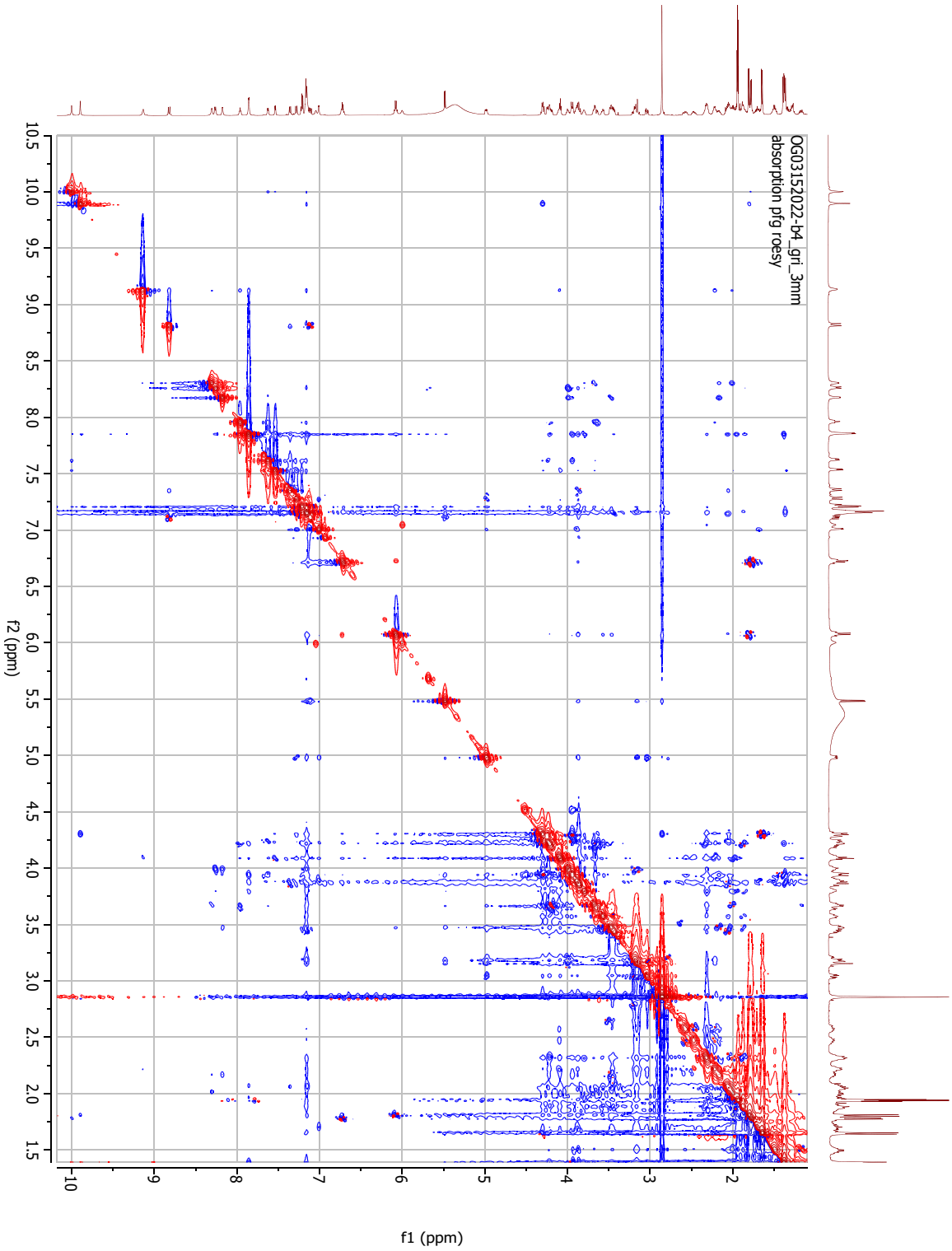


Figure 2-14 ROESY spectrum of **1**.

I then examined whether compound **1** contained D-amino acid residues. The purified compound **1** was hydrolyzed in 6 M deuterium chloride at 110 °C for 6 h, and the released amino acids were analyzed by LC-MS after derivatization with D- and L-1-fluoro-2,4-dinitrophenyl-5-leucine amide (FDLA). In this method, the influence of unwanted racemization during acid hydrolysis can be excluded[8]. As shown in Figures 2-15 and 2-16, comparison with authentic standards of D-FDLA-derivatized amino acids revealed that all Ala, Gln, Pro, and *allo*-Ile (from Ile) residues and two Val residues are D-enantiomers. The presence of D-*allo*-Ile was confirmed by C_3 Marfey's method[7] because Ile and *allo*-Ile could not be differentiated using the original Marfey's method. In contrast, all Phe, Ser, and Leu residues and one Val residue were determined to be L-enantiomers.

These results clearly suggested that compound **1** contains multiple D-amino acids in a similar manner to salinipeptins, and that a novel peptide epimerase is involved in grisemycin biosynthesis (Figure 2-17).

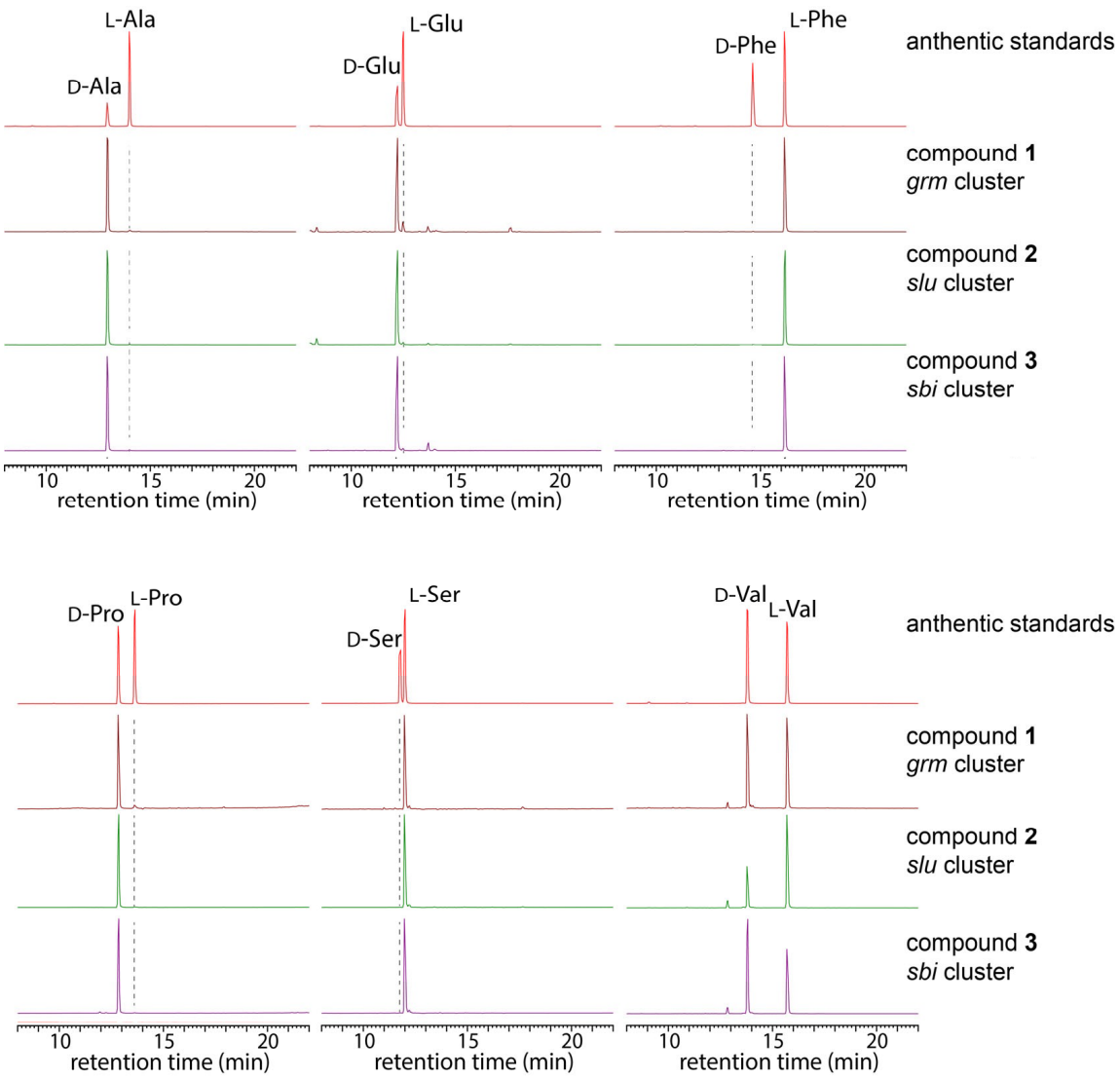


Figure 2-15 LC-MS analysis of D-FDLA derivatives of amino acids (Ala, Glu, Phe, Pro, Ser, and Val, ESI negative ion mode) observed in compounds **1–3**. Chromatograms were monitored by m/z = 382 for Ala, 398 for Ser, 408 for Pro, 410 for Val, 440 for Glu (hydrolyzed product from Gln), and 458 for Phe.

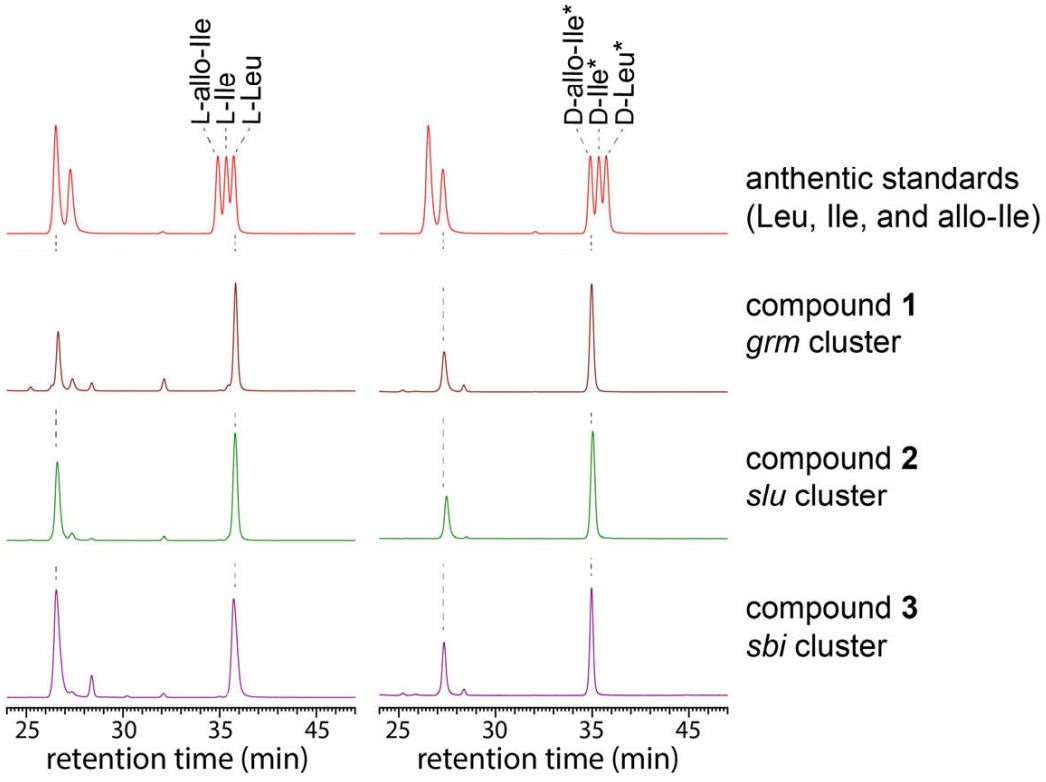


Figure 2-16 LC-MS analysis of L- and D-FDLA derivatives of amino acids (Leu, Ile, and *allo*-Ile, ESI negative ion mode). Chromatograms were monitored by m/z = 424 for Leu/Ile/*allo*-Ile. (Left) D-FDLA and (Right) L-FDLA derivatives of authentic standards and hydrolyzed product of **1**−**3**.

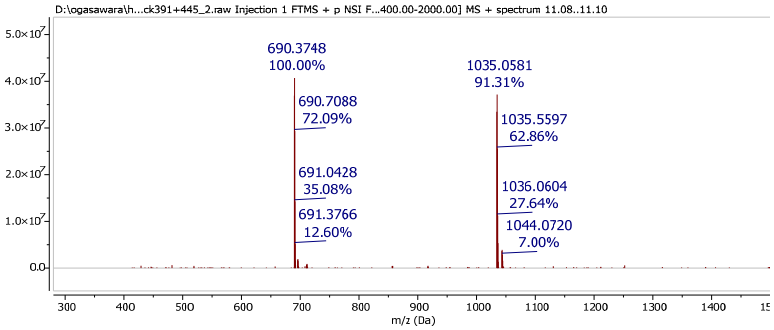
Figure 2-17 Structures of **1**–**3**. D-amino acids in red.

2.2.2 Heterologous expression of other type-A linaridin biosynthetic gene clusters

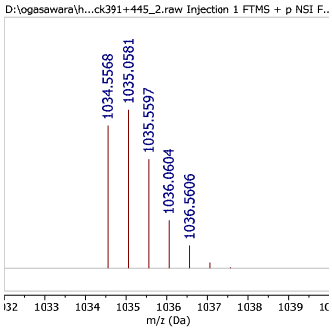
To examine whether the presence of D-amino acid residues is the common feature of type-A linaridins, I heterologously expressed the BGCs of the *slu* (*sluA–sluP*) and *sbi* (*sbiA–sbiP*) clusters using a similar method to that described above, the specific metabolites were analyzed by LC-MS (Figure 2-3). As shown in Figures 2-18 and 2-19, HR-MS and MS/MS analyses revealed the production of salinipeptin-like linaridin **2** (obs.: *m*/*z* 1034.5568, calcd. for $C_{97}H_{150}N_{24}O_{24}S$, calcd.: 1034.5561 $[M+2H]^{2+}$) and cypemycin-like linaridin **3** (*m*/*z* obs.: 1048.5722, calcd. for $C_{99}H_{154}N_{24}O_{24}S$: 1048.5717 $[M+2H]^{2+}$) from the *slu* and *sbi* clusters, respectively. Furthermore, chiral analysis of the hydrolytic products of these compounds revealed that both contained multiple D-amino acid residues in a similar manner to grisemycin (Figure 2-15–2-17).

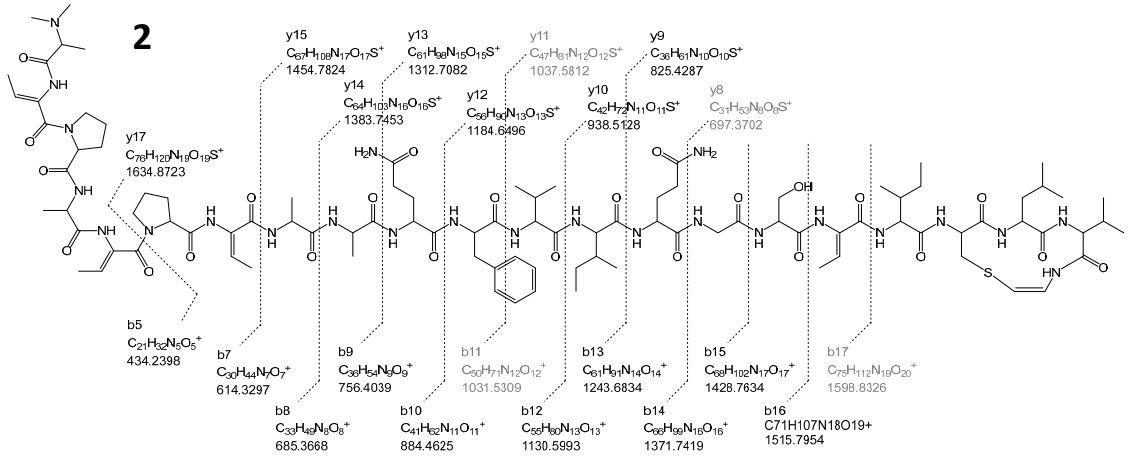
These results indicated that the presence of D-amino acids is likely a common feature of type-A linaridin natural products.

HR-MS spectrum



HR-MS spectrum (enlarged)





HR-MS/MS spectrum

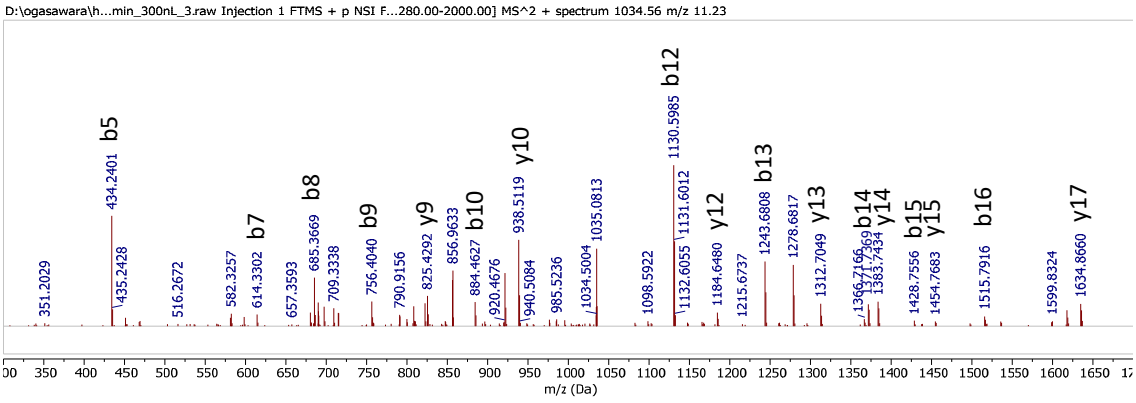
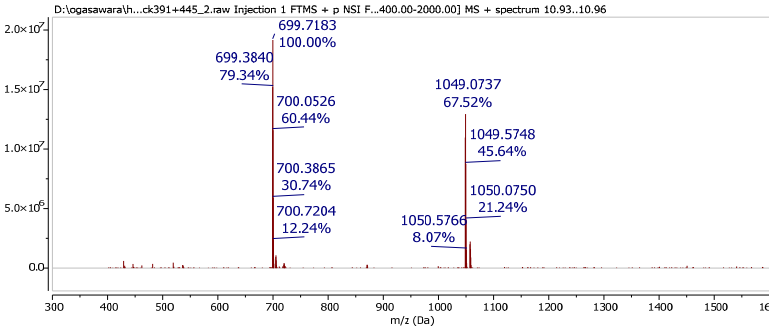
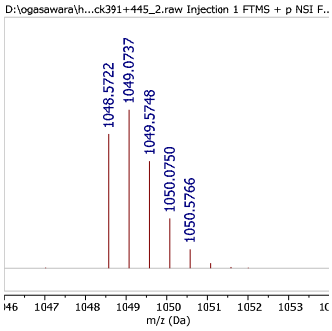


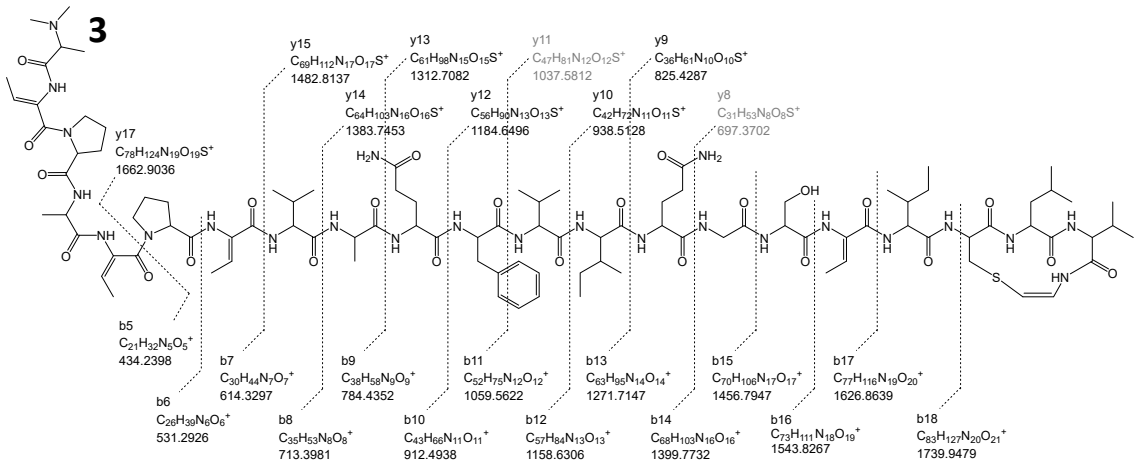
Figure 2-18 HR-MS and MS/MS spectra of compound **2**.

HR-MS spectrum



HR-MS spectrum (enlarged)





HR-MS/MS spectrum

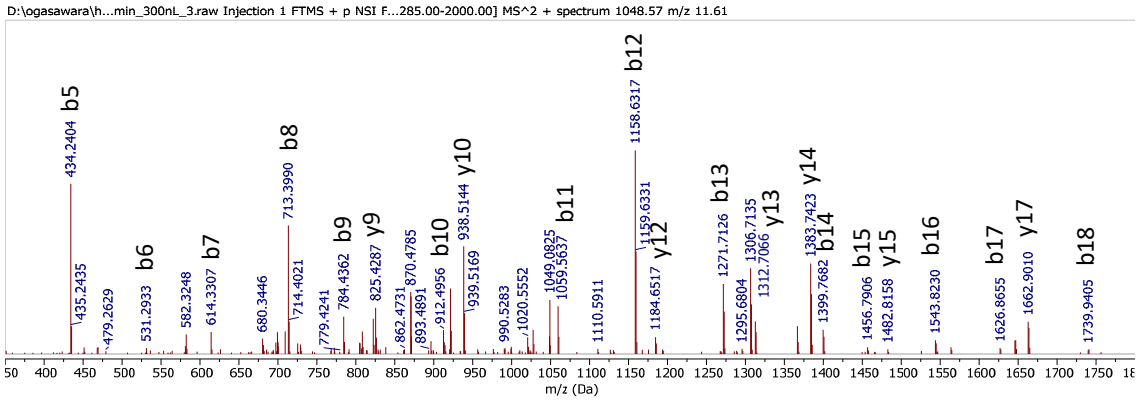


Figure 2-19 HR-MS and MS/MS spectra of compound **3**.

2.2.3. Gene-deletion experiments

I next investigated the function of each gene by gene-deletion experiments of the *grm* cluster. After construction of pWHM3-grm-derived in-frame deletion plasmids lacking the *grmH*, *grmL*, *grmM*, or *grmD* gene, each plasmid was introduced into *S. lividans,* and the metabolites were analyzed by LC-MS. When any one of the genes was removed, grisemycin production was completely abolished (Figure 2-20). The result clearly indicated that all four genes are indispensable for grisemycin biosynthesis.

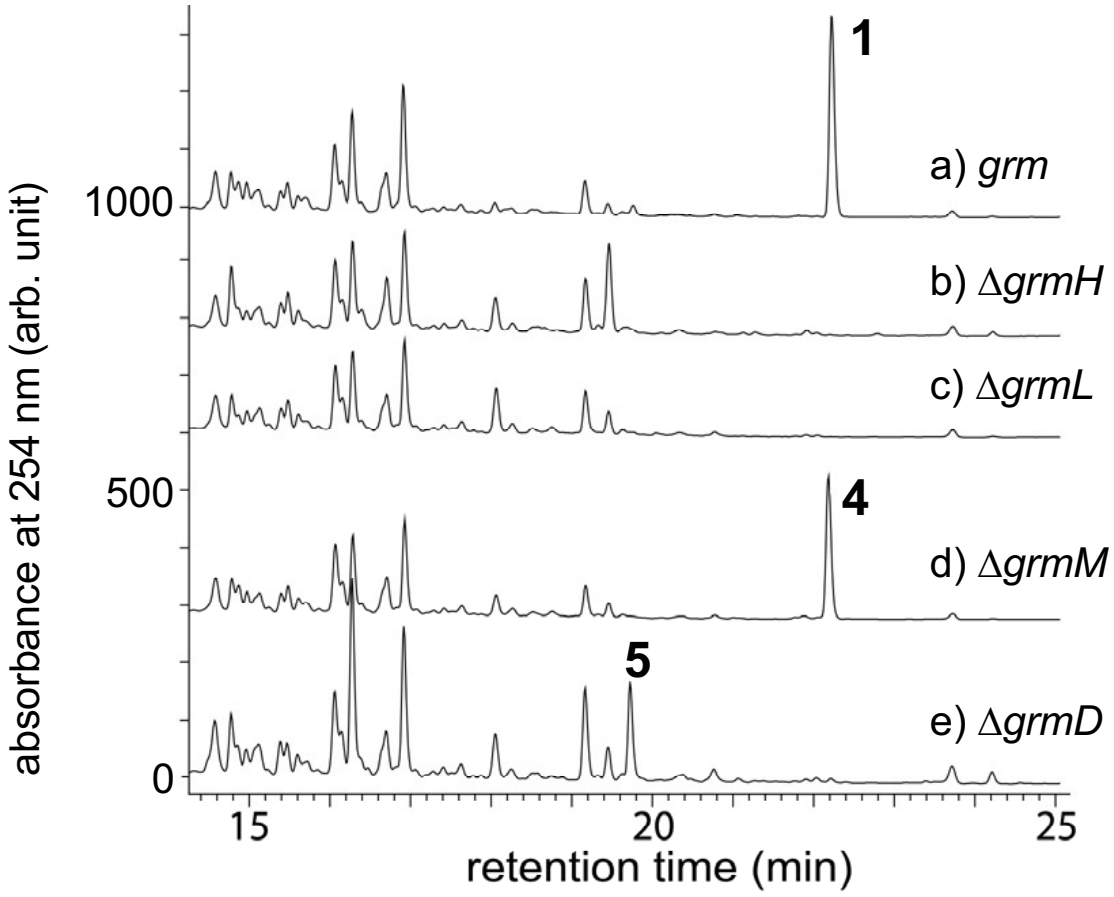
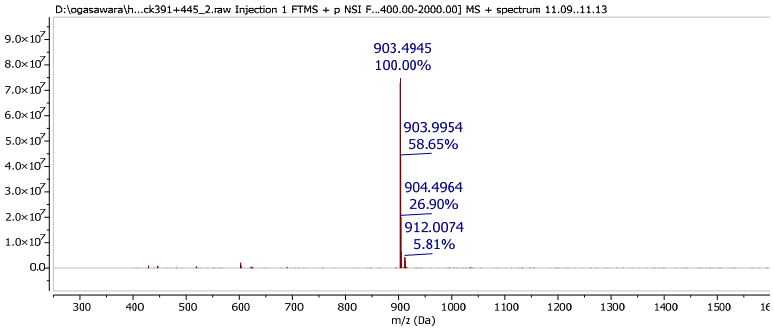


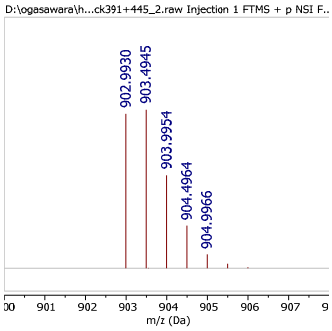
Figure 2-20 LC-MS analysis of the chloroform extracts of the culture broths. a) *S. lividans*/pWHM3-grm, b) *S. lividans*/pWHM3-grm-ΔgrmH, c) *S. lividans*/pWHM3-grm-ΔgrmL, d) *S. lividans*/pWHM3-grm-ΔgrmM, e) *S. lividans*/pWHM3-grm-ΔgrmD.

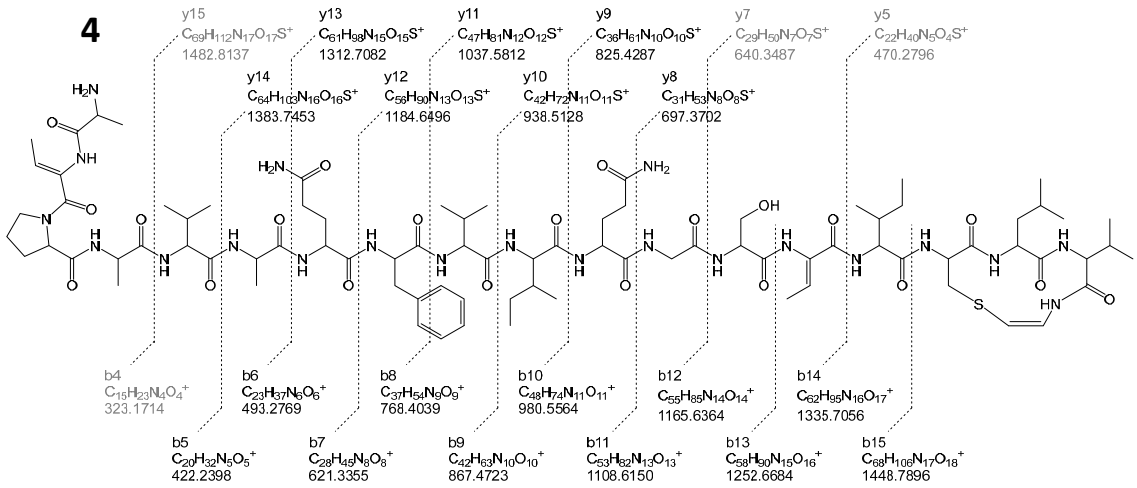
Furthermore, chromatograms revealed that the *grmM* and *grmD* mutants produced new compounds **4** and **5**, while no intermediates were detected in the *grmH* and *grmL* mutants. Then, the structures of these compounds were analyzed by LC-MS and MS/MS. The compound **4** produced by the *grmM*-disruptant had a similar structure to grisemycin except for the absence of the *N*-terminal *N*,*N*-dimethyl groups (*m*/*z* obs.: 902.9930, calcd. for $C_{84}H_{133}N_{21}O_{21}S$: 902.9926 $[M+2H]^{2+}$), as expected from the predicted function of GrmM (Figure 2-21). In contrast, the molecular formula of the compound **5** produced by the *grmD*-disruptant deduced from HR-MS was $C_{79}H_{126}N_{20}O_{20}$ (*m*/*z* obs.: 838.4801, calcd.: 838.4802 $[M+2H]^{2+}$), which was 203 mass units smaller than that of the expected non-decarboxylated peptide of the *ΔgrmD* mutant (Figure 2-22). Considering that the *N*-terminal 16 amino acid residues of the compound were suggested to be the same as those of grisemycin by observation of a series of fragment ions (*b5–b16* ions), I proposed the structure as shown in Figures 2-22. Removal of two amino acid residues (Val-18 and Cys-19) from the C-terminus, presumably catalyzed by endogenous enzymes in the heterologous host, might occur to generate the compound from the non-decarboxylated intermediate.

HR-MS spectrum



HR-MS spectrum (enlarged)





HR-MS/MS spectrum

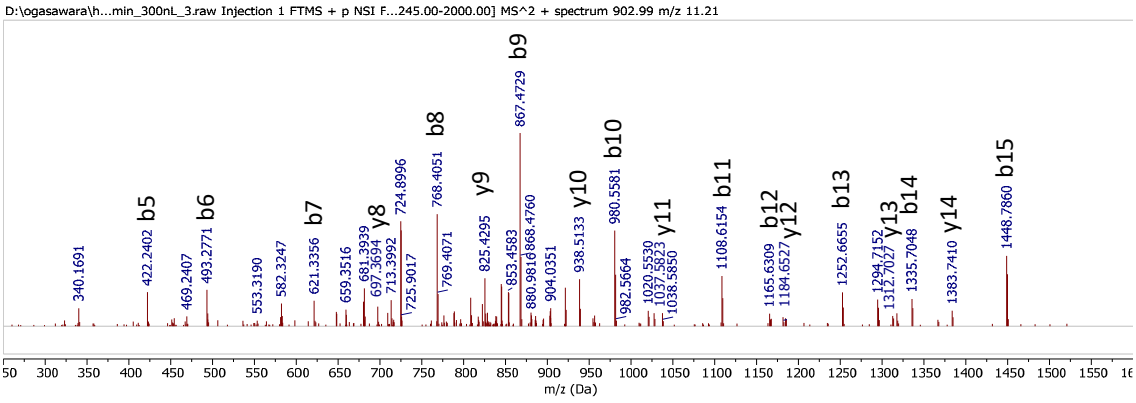
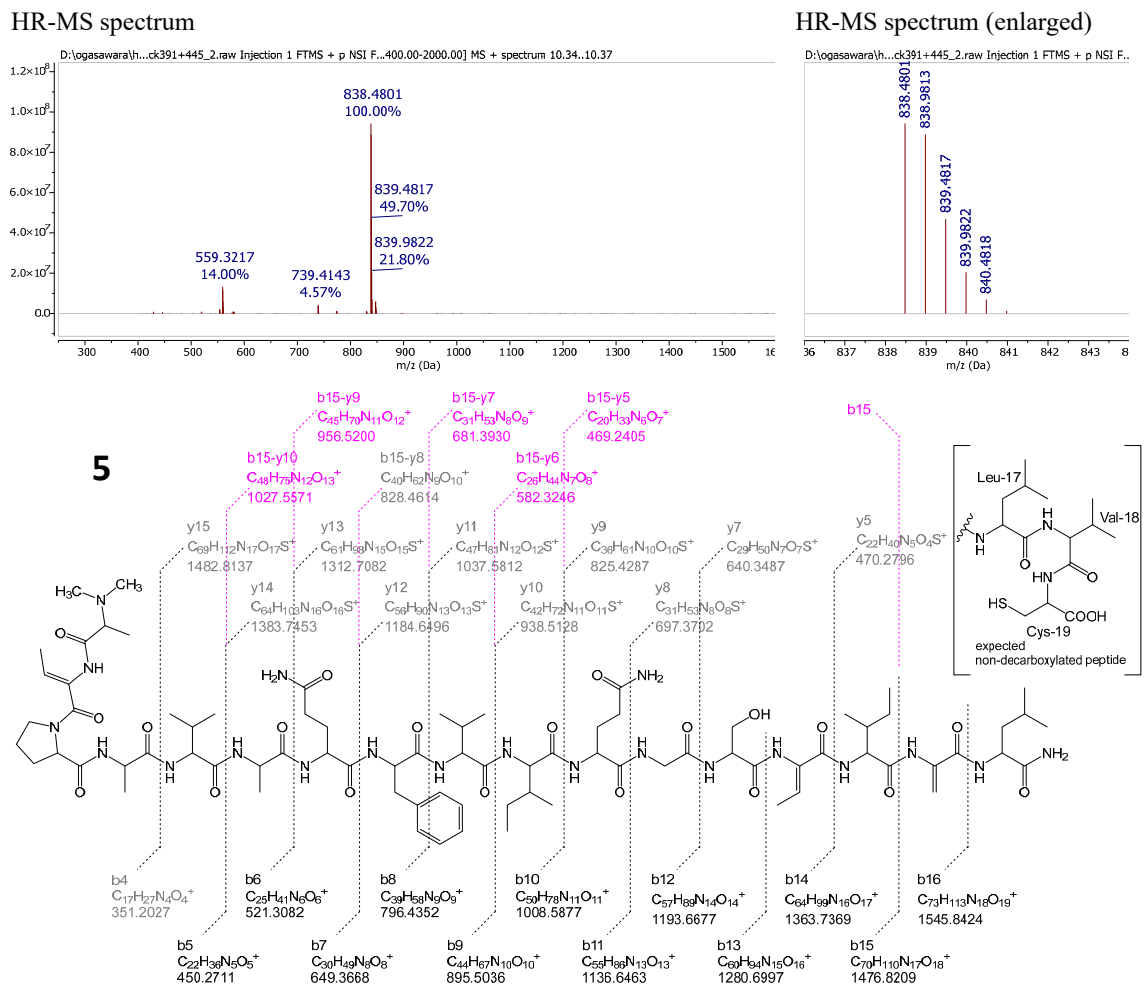


Figure 2-21 HR-MS and MS/MS spectra of compound **4**.



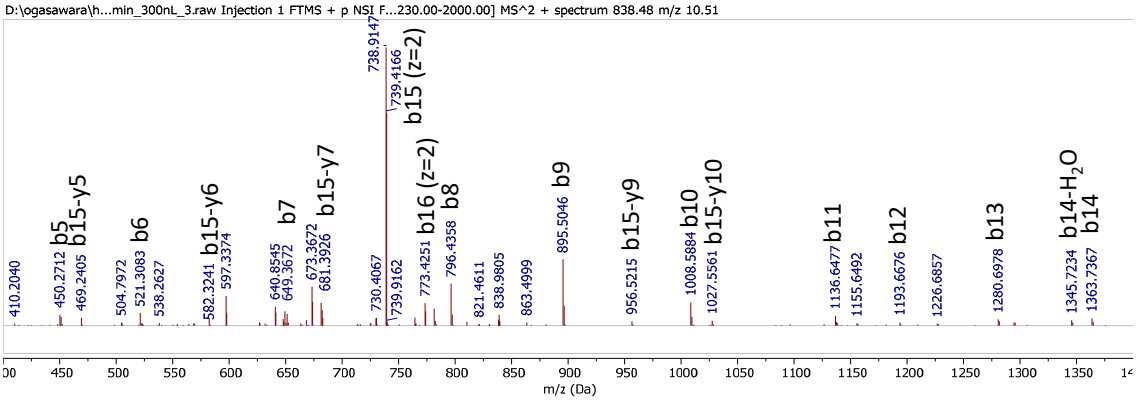


Figure 2-22 HR-MS and MS/MS spectra of compound **5** and the structure of the expected non-decarboxylated peptide.

Importantly, compounds **4** and **5** were further analyzed by the modified Marfey's method described above. As shown in Figures 2-23–2-24, both compounds **4** and **5** produced by the *grmM*-disruptant and the *grmD*-disruptant contained multiple D-amino acids that correspond to grisemycin.

These results suggested that the introduction of D-amino acids by the novel epimerase occurred prior to methylation and decarboxylation. Considering that GrmL was the only enzyme of unknown function, GrmL may be responsible for the introduction of multiple D-amino acid residues in the early stage of grisemycin biosynthesis.

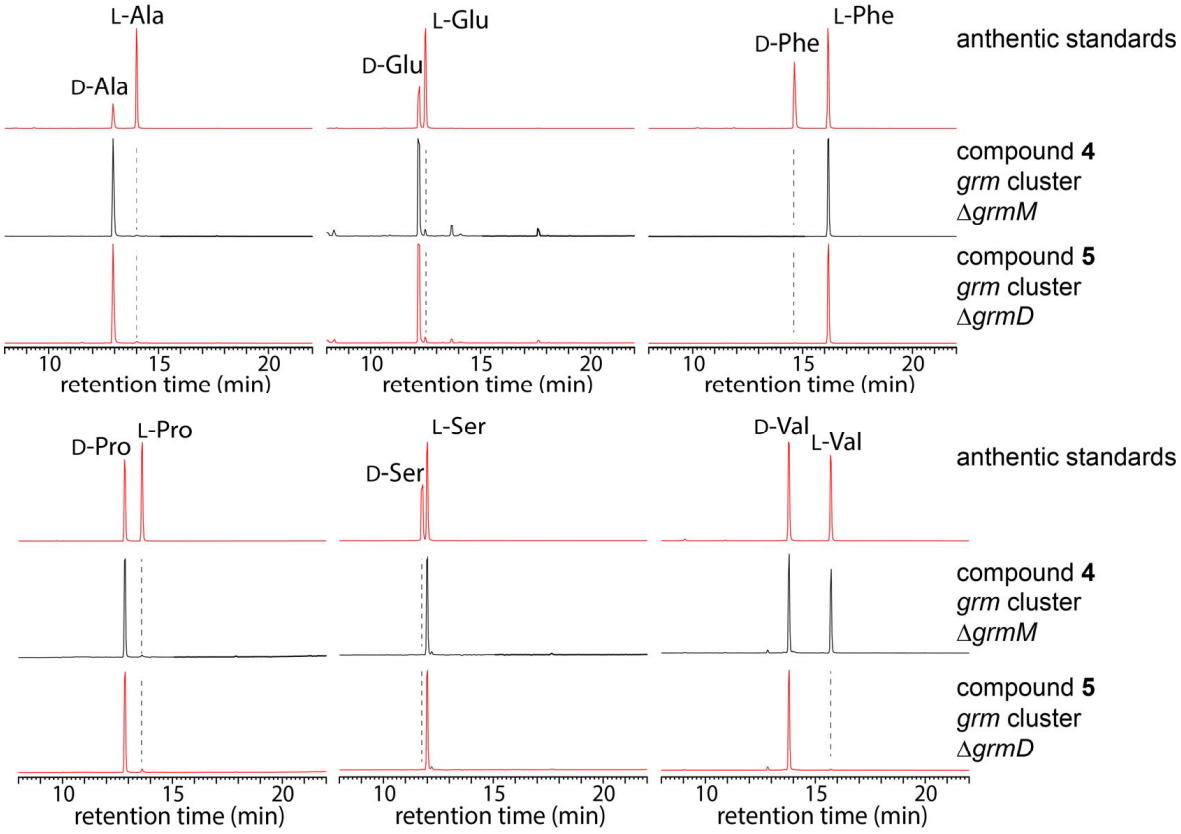


Figure 2-23 LC-MS analysis of D-FDLA derivatives of amino acids (Ala, Glu, Phe, Pro, Ser, and Val, ESI negative ion mode) observed in compounds **4** and **5**. Chromatograms were monitored by *m*/*z* = 382 for Ala, 398 for Ser, 408 for Pro, 410 for Val, 440 for Glu (hydrolyzed product from Gln), and 458 for Phe. For Val, both enantiomers were observed in compounds **1**–**4**, but not in compound **5**, and thus, penultimate Val residue in peptide was proposed to have L-configuration and the others had D-configuration.

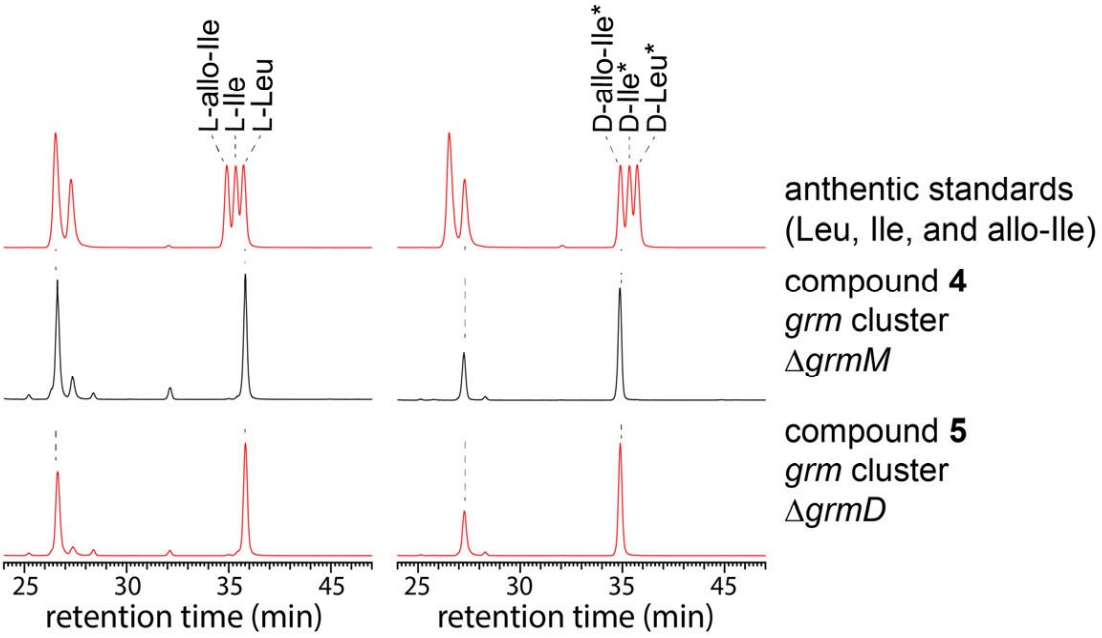


Figure 2-24 LC-MS analysis of L- and D-FDLA derivatives of amino acids (Leu, Ile, and *allo*-Ile, ESI negative ion mode). Chromatograms were monitored by *m*/*z* = 424 for Leu/Ile/*allo*-Ile. (Left) D-FDLA and (Right) L-FDLA derivatives of authentic standards and hydrolyzed products of **4** and **5**.

2.3 Conclusion

I performed heterologous expression of three type-A linaridin biosynthetic gene clusters and demonstrated that all of the metabolites contained multiple D-amino acids in a similar manner to salinipeptin A. Importantly, these observations strongly indicated that linaridin clusters contained a gene encoding a novel peptide epimerase.

Furthermore, gene-deletion experiments of the *grm* cluster showed that the *grmL*, which was the sole gene of unknown function, and *grmH* in the cluster, were indispensable for grisemycin biosynthesis and that the reactions of epimerization and dehydration preceded decarboxylation and methylation.

2.4 Experimental Section

2.4.1 General

All chemicals were purchased from Fujifilm Wako Pure Chemical Corporation (Osaka, Japan), Sigma-Aldrich Japan (Tokyo, Japan), or Tokyo Chemical Industry Co. Ltd. (Tokyo, Japan) unless specified otherwise. Enzymes and kits for DNA manipulations were purchased from Takara Bio (Shiga, Japan), Nippon Gene Co. Ltd. (Tokyo, Japan), or New England Biolabs Japan Inc. (Tokyo, Japan). Polymerase chain reaction (PCR) was carried out using a SimpliAmp thermal cycler (Life Technologies Japan Ltd., Tokyo, Japan) and Tks Gflex DNA polymerase (Takara Bio). Oligonucleotide synthesis and DNA sequencing were performed in Fasmac (Kanagawa, Japan). High-resolution MS and MS/MS data were recorded using an LTQ-orbitrap XL mass spectrometer (Thermo Scientific) at the Instrument Analysis Division, Global Facility Center, Creative Research Institution, Hokkaido University. NMR spectra were recorded using a JEOL JNM-ECZ600R spectrometer in Frontier Chemistry Center in Hokkaido University. General genetic manipulations of *Escherichia coli* and *S. lividans* TK23 were performed according to standard protocols. *S. griseus* NBRC 13350, *S. luteocolor* NBRC 13826, and *S. bikiniensis* NBRC 14598 were obtained from NITE Biological Resource Center, the National Institute of Technology and Evaluation (Tokyo, Japan).

2.4.2 Heterologous expression of biosynthetic gene clusters

DNA fragments carrying the *slu* (*S. luteocolor* NBRC 13826), *sbi* (*S. bikiniensis* NBRC 14598), and *grm* (*S. griseus* NBRC 13350) clusters were amplified by PCR from corresponding genomic DNA with primer pairs, slu-cluster_F (XbaI)/slu-cluster_R (HindIII), sbi-cluster_F (XbaI)/sbi-cluster_R (HindIII), and grm-cluster_F (XbaI)/grm-cluster_R (HindIII), respectively. The amplified fragments were digested with *Xba*I and *Hin*dIII and inserted into the same sites of pWHM3. The resulting plasmids were introduced into *S. lividans* TK23, and the transformants were grown in TSB medium (10 mL) supplied with thiostrepton (5 μg/ml). A portion of the seed culture (1 mL) was transferred into a 250-ml Erlenmeyer flask containing 50 ml of SK#2 medium (2% soluble starch, 0.5% glucose, 0.5% yeast extract, 0.3% peptone, 0.3% fish extract, 0.02% KH_2PO_4, and 0.06% $MgSO_4 \cdot 7H_2O$, pH 7.6, supplemented with 5 μg/mL thiostrepton). After 3-day cultivation at 30°C, 200 rpm, 50 mL of culture was centrifuged (15,000 ×rpm, 10 min) to remove the cells. The resulting supernatant was extracted with the same volume of chloroform three times. The combined organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure. The extract was re-dissolved in water with 0.1% formic acid (300 μL) and analyzed by HPLC and LC-MS. HPLC analysis and purification was performed with a Shimadzu Prominence HPLC system equipped with a photo diode array (PDA) detector using the following conditions: column, Zorbax 300SB-C8 column (250 × 4.6 mm, 5 μm, Agilent); column temperature, 40 °C;

detection, PDA (190–350 nm); mobile phase, A: water with 0.1% trifluoroacetic acid (TFA), B: acetonitrile with 0.1% TFA, 5% solvent B for 0–5 min, a linear gradient to 95% solvent B for 5–30 min, and 95% solvent B for 30–40 min; flow rate, 1.0 mL/min. For LC-MS analysis, Shimadzu Prominence HPLC system equipped with amaZon SL DB-1 (Bruker, MA, USA) was used under the following conditions: column, Zorbax 300SB-C8 column (150 × 2.1 mm, 3 μm, Agilent); column temperature, 40 °C; detection, PDA (190–350 nm) and ESI-MS positive mode (*m*/*z* 50–2000); capillary voltage, 4500 V; end plate offset voltage, 500 V; nebulizer gas, 1.5 bar; dry gas flow, 7.0 L/min; dry temp., 200 °C; mobile phase, A: water with 0.1% trifluoroacetic acid (TFA), B: acetonitrile with 0.1% TFA, 5% solvent B for 0–5 min, a linear gradient to 95% solvent B for 5–30 min, and 95% solvent B for 30–40 min; flow rate, 0.2 mL/min.

2.4.3 Chiral analysis of amino acid residues

The chirality of the amino acid residues in the peptide products was analyzed using the modified Marfey's method. Each product obtained by heterologous expression was first hydrolyzed at 110 °C for 6 h with 200 μL of 6 M deuterium chloride (DCl) in D_2O. After evaporation to dryness, the sample was re-dissolved in 100 μL of water. To the sample solution 20 μL was added 5 μL of 1 M sodium bicarbonate and 25 μL of 1% 1-fluoro-2,4-dinitrophenyl-5-D-leucinamide (D-FDLA) or 1-fluoro-2,4-dinitrophenyl-5-L-leucinamide (L-FDLA) in acetone. After incubation for 2 h at 37 °C, the reaction was quenched by the addition of 5

μL of 1 M HCl, and then diluted with 200 μL of methanol and 250 μL of water. The sample containing D-FDLA derivatives was analyzed by LC-MS using Waters ACQUITY UPLC system equipped with an SQ Detector2 and Acquity PDA detector under the following conditions: column, Mightysil RP-18GP Aqua column (150 × 2.1 mm, Kanto Chemical Co. Inc., Tokyo, Japan); column temperature, 40 °C; detection, ESI negative mode with selected ion monitoring (± 0.5 Da) at *m*/*z* = 382 for Ala, 398 for Ser, 408 for Pro, 410 for Val, 424 for Leu/Ile/*allo*-Ile, 440 for Glu (hydrolyzed product from Gln), and 458 for Phe and PDA (200–500 nm); cone voltage, 40 V; mobile phase, solvent A (water with 0.05% TFA) and solvent B (acetonitrile with 0.05% TFA), 15% solvent B for 0–2 min and a linear gradient to 95% solvent B for an additional 18 min; flow rate, 0.2 mL/min. Because the above conditions cannot distinguish Ile from *allo*-Ile, the chirality of Leu, Ile, and *allo*-Ile was analyzed with the C3 Marfey's method under the following procedure. The sample containing D- or L-FDLA derivatives was analyzed by LC-MS under the following conditions: column, Zorbax SB-C3 (150 × 2.1 mm, 5 μm, Agilent); column temperature, 50 °C; detection, ESI-negative mode with selected ion monitoring (± 0.5 Da) at *m*/*z* = 424 (Leu/Ile/*allo*-Ile) and PDA (200–500 nm); cone voltage, 40 V; mobile phase, solvent A (water), solvent B (methanol), and solvent C (acetonitrile with 0.1% formic acid), a linear gradient from solvents A:B:C = 70:25:5 to 5:90:5 for 50 min; flow rate, 0.2 mL/min.

2.4.4 Gene deletion of the biosynthetic genes

To construct pWHM3-grm derived plasmid lacking *grmH* (pWHM3-grm-ΔgrmH), upstream and downstream-regions of *grmH* gene were amplified using the primers pairs grm-cluster_up_BsmI_F/grm-cluster_del-grmH_up_R and grm-cluster_del-grmH_down_F/grm-cluster_down_AfeI_R, respectively, and two DNA fragments were simultaneously inserted into *Bsm*I–*Afe*I sites of pWHM3-grm by three fragments in-fusion cloning. Likewise, pWHM3-grm-ΔgrmL and pWHM3-grm-ΔgrmD were constructed using appropriate primers as listed in Table 2-3. For the construction of pWHM3-grm-ΔgrmM, a PCR fragment was prepared using a primer pair grm-cluster_up_BsmI_F/grm-cluster_del-grmM_up_R and was cloned into *Bsm*I–*Afe*I sites of pWHM3-grm by in-fusion cloning. Nucleotide sequences of the PCR-derived region (*Bsm*I–*Afe*I sites) in plasmids were confirmed by DNA sequencing.

Table 2-3 Oligonucleotides used in this study. The introduced restriction sites are underlined. Homology arm sequences for in-fusion cloning are shown in bold letters.

primer name	sequences (5′–3′)
slu-cluster_F(XbaI)	ATATTCTAGATGGGTTCCCCCTCCGAGAGCAACAGTTG
slu-cluster_R(HindIII)	ATATAAGCTTCGACTAGGCCCGGCGTAGGCGCGTGAAC
sbi-cluster_F(XbaI)	TATATCTAGACTTCACCTGCCCACCGAGCTGCTGAC
sbi-cluster_R(HindIII)	ATATAAGCTTACTCGGCGAAACCGGCCTGAGAGAACGC
grm-cluster_F(XbaI)	TATATCTAGAGCCCACCGGTGGAACAGATACAGATCC
grm-cluster_R(HindIII)	ATATAAGCTTCATCGACCTGGACCTCGGGCTGATCC
grm-cluster_up_BsmI_F	**GACGAGCTGCATCGGACCC**AGCATTC
grm-cluster_del-grmH_up_R	**GTATCCAATGACCGCCGAC**GCCGCGATC
grm-cluster_del-grmH_down_F	**GTCGGCGGTCATTGGATAC**AGACATCAC
grm-cluster_down_AfeI_R	**GAACGAGGTGAGGGTCGAAGC**
grm-cluster_del-grmL_up_R	**TCGGATGTCAGAACGGTC**GAGACCAGGTG
grm-cluster_del-grmL_down_F	**GACCGTTCTGACATCCGA**GTGGATCCCGCTC
grm-cluster_del-grmD_up_R	**TCCCACCTCGGCACCCTC**GAACTGCTCCACGTTC
grm-cluster_del-grmD_down_F	**GAGGGTGCCGAGGTGGGA**TTCAATCTC
grm-cluster_del-grmM_up_R	**CGAGGTGAGGGTCGAAGC**CGCCGCGTCGTAAACG

2.5 References

1. Shang, Z., Winter, J. M., Kauffman, C. A., Yang, I. & Fenical, W. Salinipeptins: Integrated Genomic and Chemical Approaches Reveal Unusual d -Amino Acid-Containing Ribosomally Synthesized and Post-Translationally Modified Peptides (RiPPs) from a Great Salt Lake Streptomyces sp. *ACS Chem. Biol.* **14**, 415–425 (2019).

2. Ma, S. & Zhang, Q. Linaridin natural products. *Nat. Prod. Rep.* **37**, 1152–1163 (2020).

3. Claesen, J. & Bibb, M. J. Biosynthesis and regulation of grisemycin, a new member of the linaridin family of ribosomally synthesized peptides produced by Streptomyces griseus IFO 13350. *J. Bacteriol.* **193**, 2510–2516 (2011).

4. Claesen, J. & Bibb, M. Genome mining and genetic analysis of cypemycin biosynthesis reveal an unusual class of posttranslationally modified peptides. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 16297–16302 (2010).

5. Vara, J., Lewandowska-Skarbek, M., Wang, Y. G., Donadio, S. & Hutchinson, C. R. Cloning of genes governing the deoxysugar portion of the erythromycin biosynthesis pathway in Saccharopolyspora erythraea (Streptomyces erythreus). *J. Bacteriol.* **171**, 5872–5881 (1989).

6. Chu, L. *et al.* Hijacking a Linaridin Biosynthetic Intermediate for Lanthipeptide Production. *ACS Chem. Biol.* **17**, 3198–3206 (2022).

7. Vijayasarathy, S. *et al.* C3 and 2D C3 Marfey's Methods for Amino Acid Analysis in Natural Products. *J. Nat. Prod.* **79**, 421–427 (2016).

8. Goodlett, D. R. *et al.* Peptide chiral purity determination: hydrolysis in deuterated acid, derivatization with Marfey's reagent and analysis using high-performance liquid chromatography-electrospray ionization-mass spectrometry. *J. Chromatogr. A* **707**, 233–244 (1995).

Chapter 3

Functional Analysis of the Putative Epimerase

3.1 Introduction

In Chapter 2, I revealed that the type-A linaridin natural products such as salinipeptins, grisemycin, and cypemycin contain multiple D-amino acid residues, suggesting that *linL* locating in the biosynthetic gene clusters will be responsible for D-amino acid introduction. Moreover, the results of the gene-deletion experiments showed that dehydration and epimerization precede decarboxylation and methylation, which are catalyzed by GrmD and GrmM, respectively. Considering that GrmH, composing two domains (an *N*-terminal horizontally transferred transmembrane domain, and a C-terminal α/β hydrolase domain), is a putative dehydratase to catalyze dehydration of GrmA. Therefore, I hypothesized that epimerization of nascent precursor peptide GrmA occurs before dehydration catalyzed by GrmH and *vice versa* (Figure 3-1).

In this chapter, I examined the plausibility by *in vitro* and *in vivo* experiments.

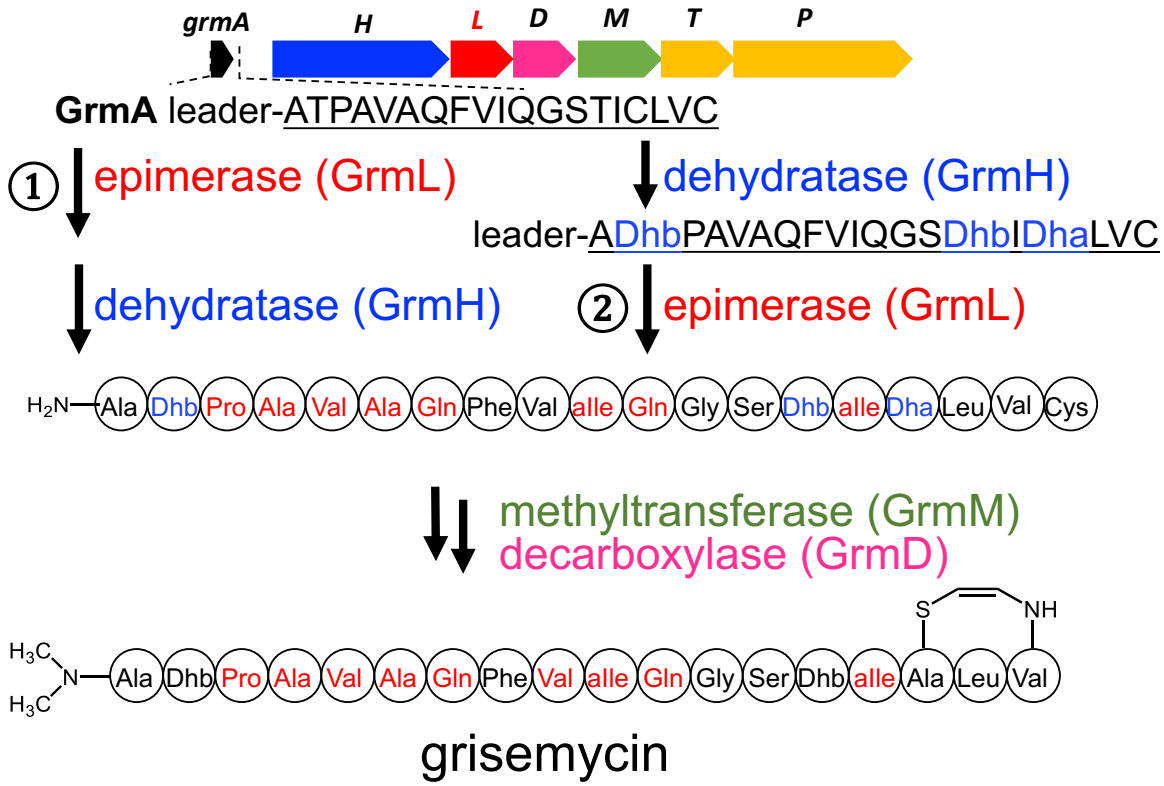


Figure 3-1 Proposed biosynthetic pathway of grisemycin. D-amino acids in red, and dehydroamino acids in blue. Dhb, dehydrobutyrine.

3.2 Results and Discussion

3.2.1. *In vitro* characterization of GrmH and GrmL

To confirm that *grmL* encodes a new type of peptide epimerase and *grmH* encodes a dehydratase, I first employed an *in vitro* experiment with recombinant precursor peptide GrmA, GrmL, and GrmH. However, at first, no expression of GrmA was observed by analysis of SDS-PAGE. Because the SKIK tag, which is the additional *N*-terminal Ser-Lys-Ile-Lys sequence, was reported to improve the productivity of peptides in *E. coli*[1], the productivity of the precursor peptide (such as MslA) with SKIK-tag could be increased[2–4]. Using the same strategy, I successfully expressed GrmA fused with SKIK and His tags at its *N*-terminus (Figure 3-2 A). I next tried to express the recombinant His-tagged GrmL and His-tagged GrmH in *E. coli* but did not obtain enough amounts of the soluble recombinants for *in vitro* assays even by many trials with various expression vectors and culture conditions (Figure 3-2 B, C).

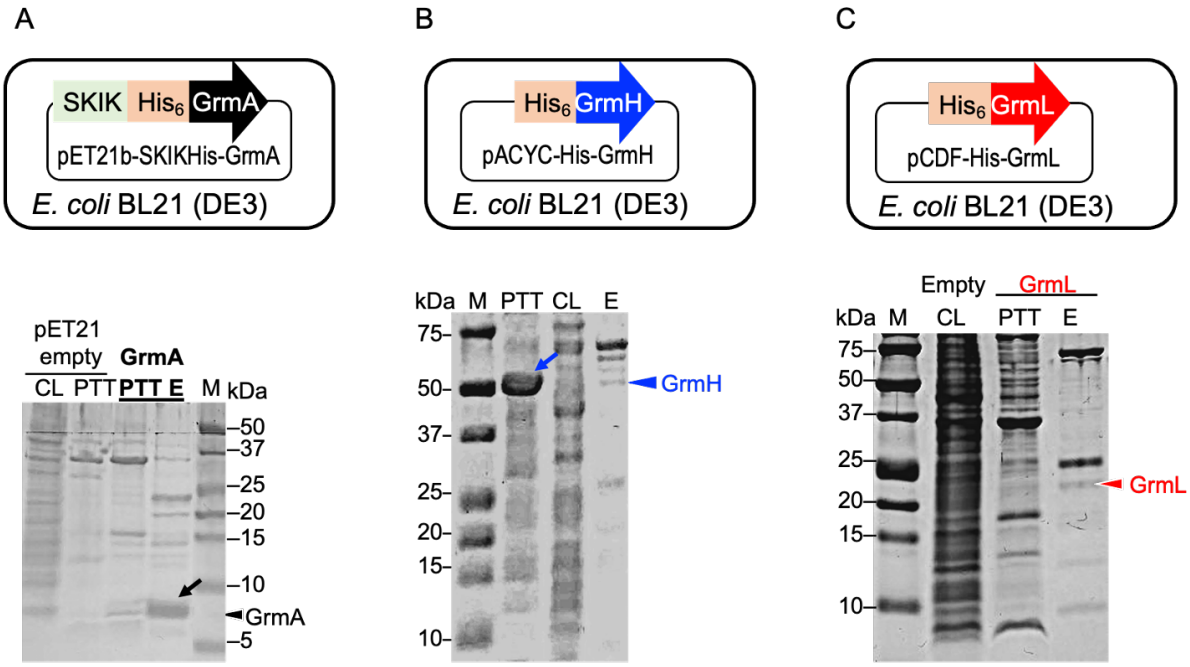


Figure 3-2 SDS-PAGE analysis of GrmA, GrmH, and GrmL expression in *E. coli*. M, marker; PTT, insoluble proteins; CL, soluble proteins; E, purified proteins by Ni-NTA affinity column.

3.2.2. *In vivo* characterization of GrmH and GrmL in *E. coli*

I therefore employed an *in vivo* heterologous expression strategy in *E. coli*, by which expression of recombinant enzymes could be detected with comparative ease. Then, I tried to coexpress recombinant GrmH and/or GrmL together with GrmA. In these cases, recombinant GrmL and GrmH were weakly expressed in contrast to the above-mentioned single GrmH and GrmL expression, which were extremely lower than that of GrmA. I then examined various plasmid constructs and succeeded in coexpression of GrmA with GrmL, with GrmH, and with both GrmL and GrmH (Figure 3-3).

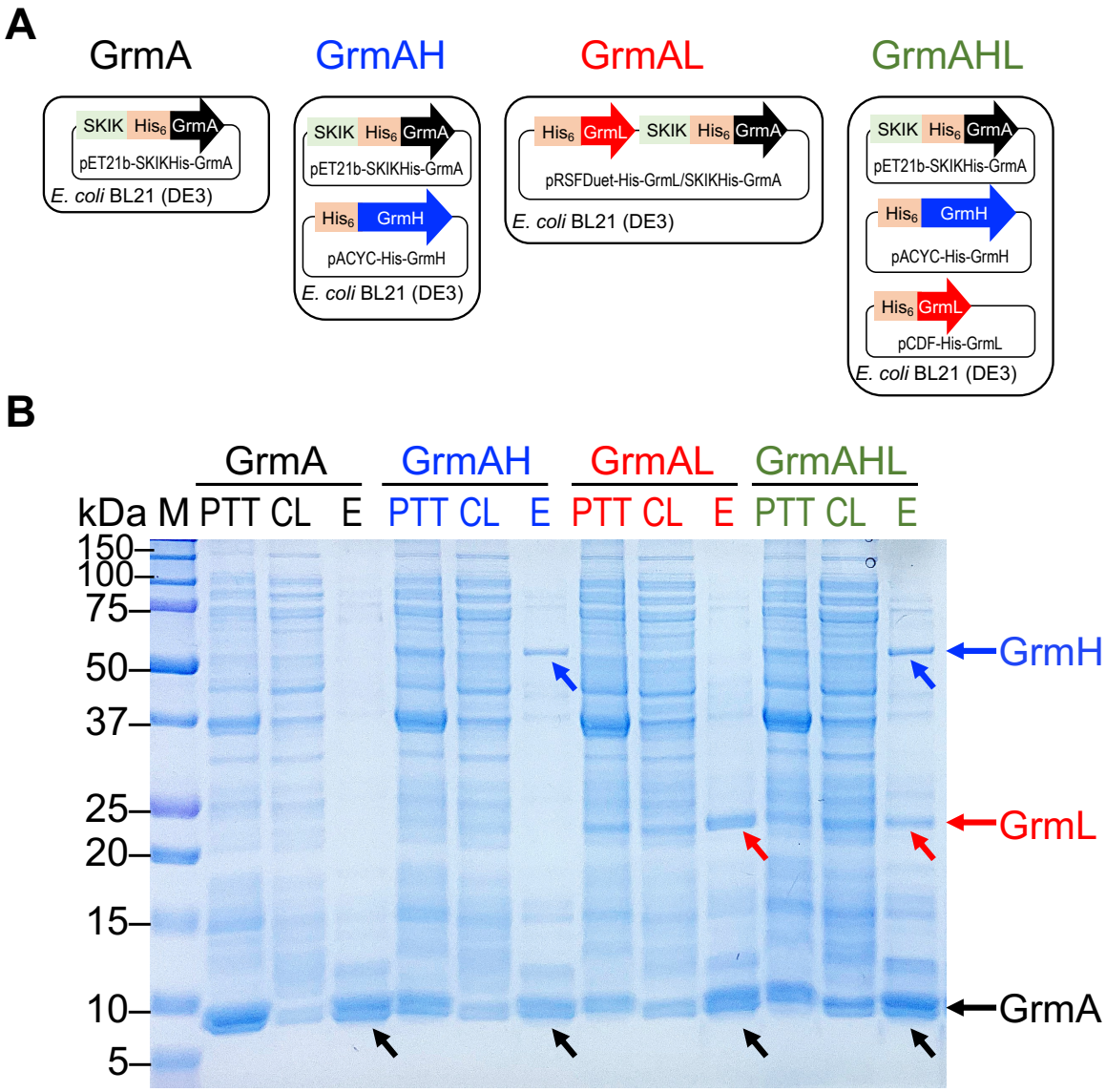


Figure 3-3 The plasmids to express *grmA*, *grmH*, and/or *grmL* in *E. coli* a), and SDS-PAGE analysis of GrmA, GrmH, and GrmL co-expression in *E. coli* b). GrmA (~8 kDa), GrmH (~56 kDa), and GrmL (~22 kDa). PTT, insoluble proteins; CL, soluble proteins; E, purified proteins by Ni-NTA affinity column.

To monitor *in vivo* epimerization reactions, the GrmA product was purified by Ni-NTA affinity chromatography and SDS-PAGE, and the chirality of amino acid residues in GrmA was analyzed by the modified Marfey's method after acid hydrolysis. However, no D-amino acids were observed in GrmA even though both GrmL and GrmH were coexpressed, and the same was true when GrmL or GrmH alone was coexpressed with GrmA (Figure 3-4). Furthermore, LC-MS analysis of purified GrmA revealed that *in vivo* dehydration reactions did not occur upon coexpressions (Figure 3-5). Considering that grisemycin was isolated from the *Streptomyces* strain, GrmL and GrmH may require cofactors absent in *E. coli*.

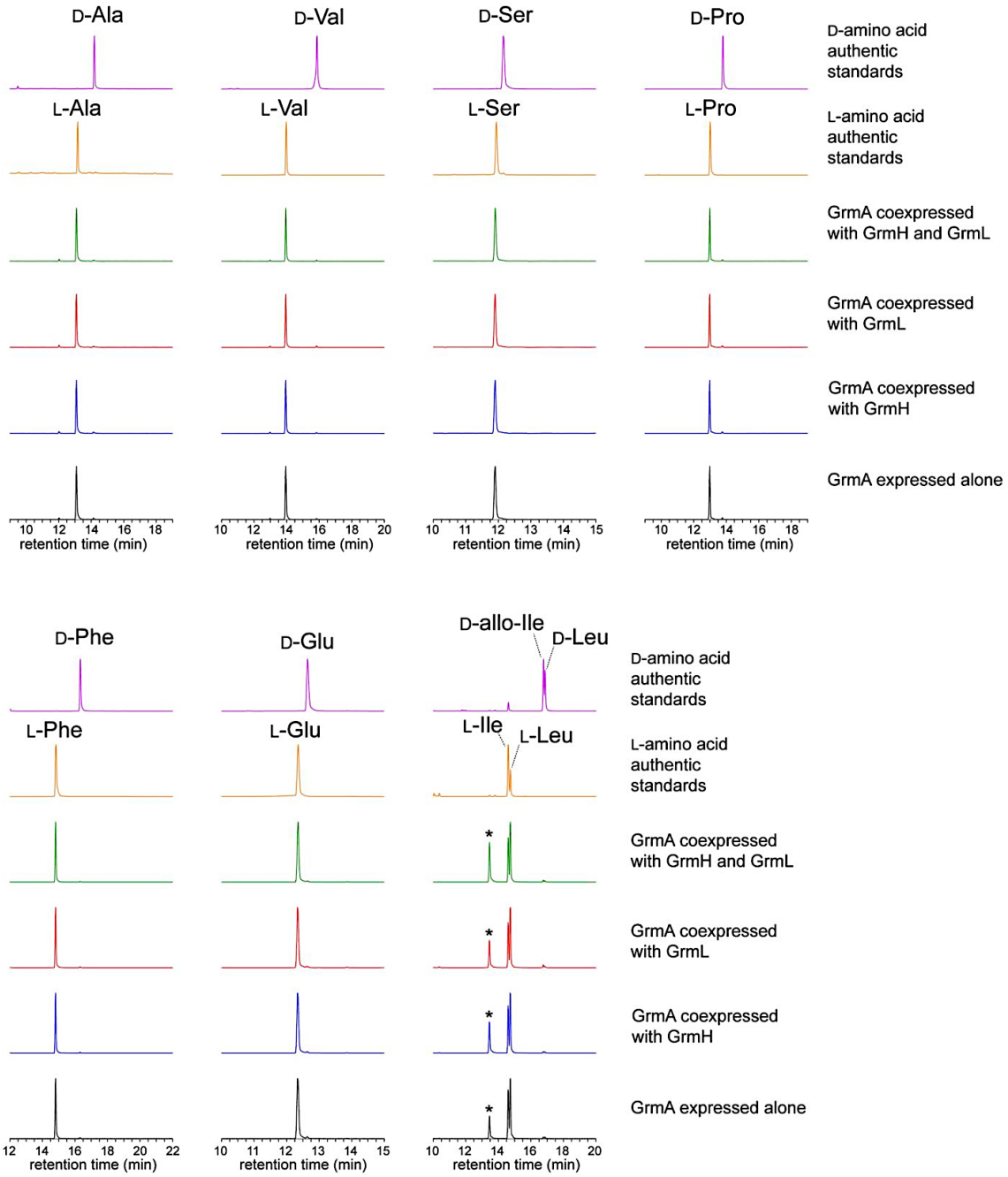


Figure 3-4 LC-MS analysis of L-FDLA derivatives of amino acids in GrmA prepared from *E. coli*. Chromatograms were monitored by m/z 382 for Ala, 410 for Val, 398 for Ser, 408 for Pro, 458 for Phe, 440 for Glu (hydrolyzed product from Gln), 424 for Leu/Ile. Peaks marked with an asterisk are not related to Leu, Ile, or *allo*-Ile.

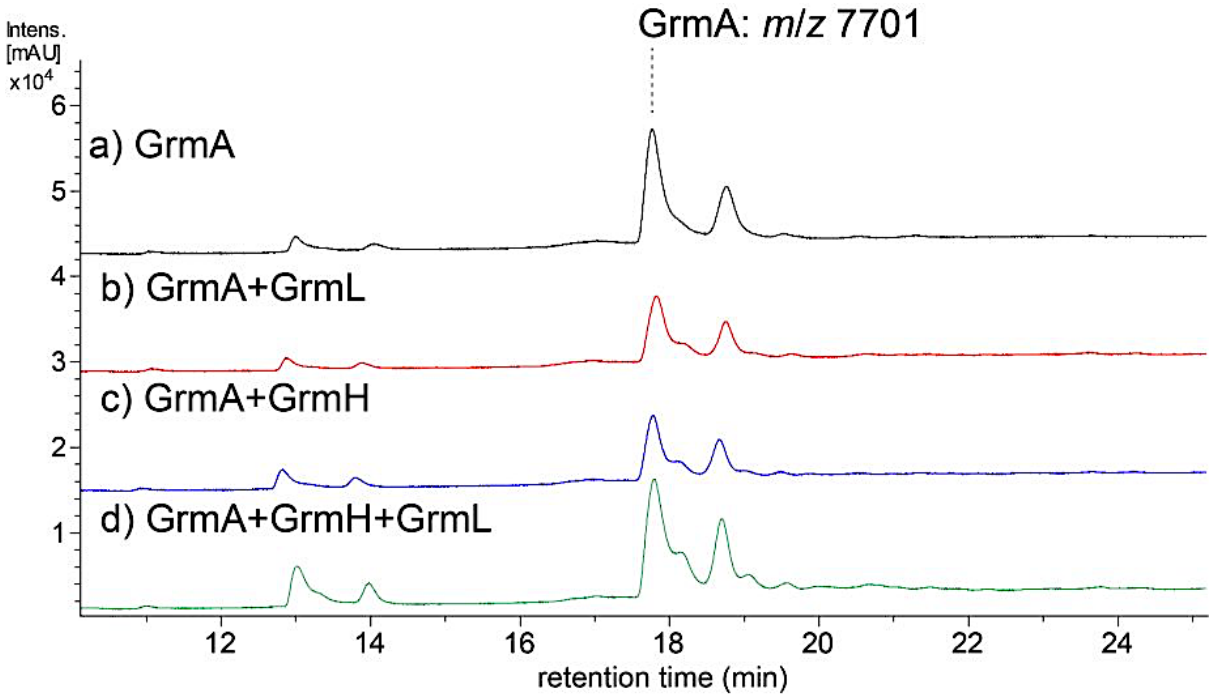


Figure 3-5 LC-MS analysis of purified GrmA from *E. coli* cells expressing a) *grmA* alone, b) *grmA* and *grmL*, c) *grmA* and *grmH*, and d) *grmA*, *grmH*, and *grmL*. Chromatograms were monitored with UV absorption at 254 nm.

3.2.3. *In vivo* characterization of GrmH and GrmL in *S. lividans*

I therefore used *S. lividans* as the heterologous host. DNA regions containing *grmA*, *grmA–grmH*, and *grmA–grmL* were respectively cloned into pSE101-N8H to construct pSE101-N8H-grmA, pSE101-N8H-grmAH, and pSE101-N8H-grmAHL, so that *grmA* product is fused to *N*-terminal 8×His-tag. Similarly, pSE101-N8H-grmAL, a plasmid containing His-tagged-*grmA* and *grmL*, was constructed by in-frame deletion of *grmH*. After *S. lividans* TK23 was transformed by the plasmids, the transformants were cultivated in YEME medium for 3 days. The resulting cultures were then centrifuged and both supernatant and cell extracts were analyzed by LC-MS. As shown in Figure 3-6, a specific metabolite (**6**) was observed in 20% acetone extract of the cells expressing *grmA*, *grmH*, and *grmL*. The monoisotopic mass of **6** was 925.9988 ($[M+2H]^{2+}$, calcd. for $C_{85}H_{135}O_{23}N_{21}S$: 925.9953), which corresponds to the predicted dehydrated product of GrmA core peptide. More importantly, chiral analysis by the modified Marfey's method revealed that compound **6** contained multiple D-amino acid residues identical to those in grisemycin (Figure 3-7, Figure 3-8).

These results clearly indicated that GrmH and GrmL were responsible for the dehydration and epimerization reactions. By contrast, no specific metabolites were observed in 20% acetone extract of the other transformants containing pSE101N8H-grmA, pSE101N8HgrmAH, or pSE101N8H-grmAL (Figure 3-6).

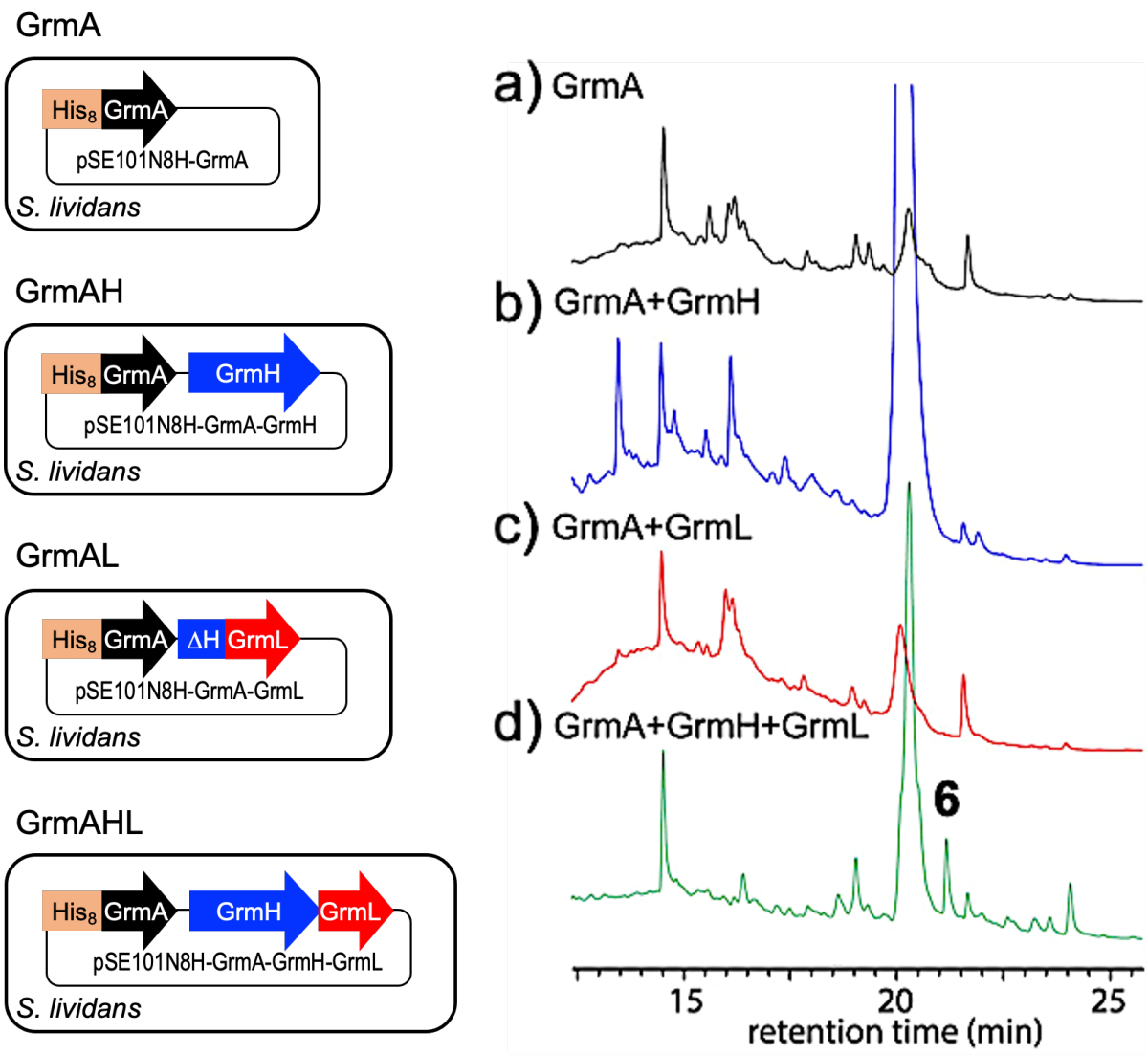


Figure 3-6 LC-MS analysis of the extract of cells expressing a) GrmA, b) GrmA + GrmH, c) GrmA + GrmL, d) GrmA + GrmH + GrmL. The chromatograms were monitored by UV absorption at 254 nm.

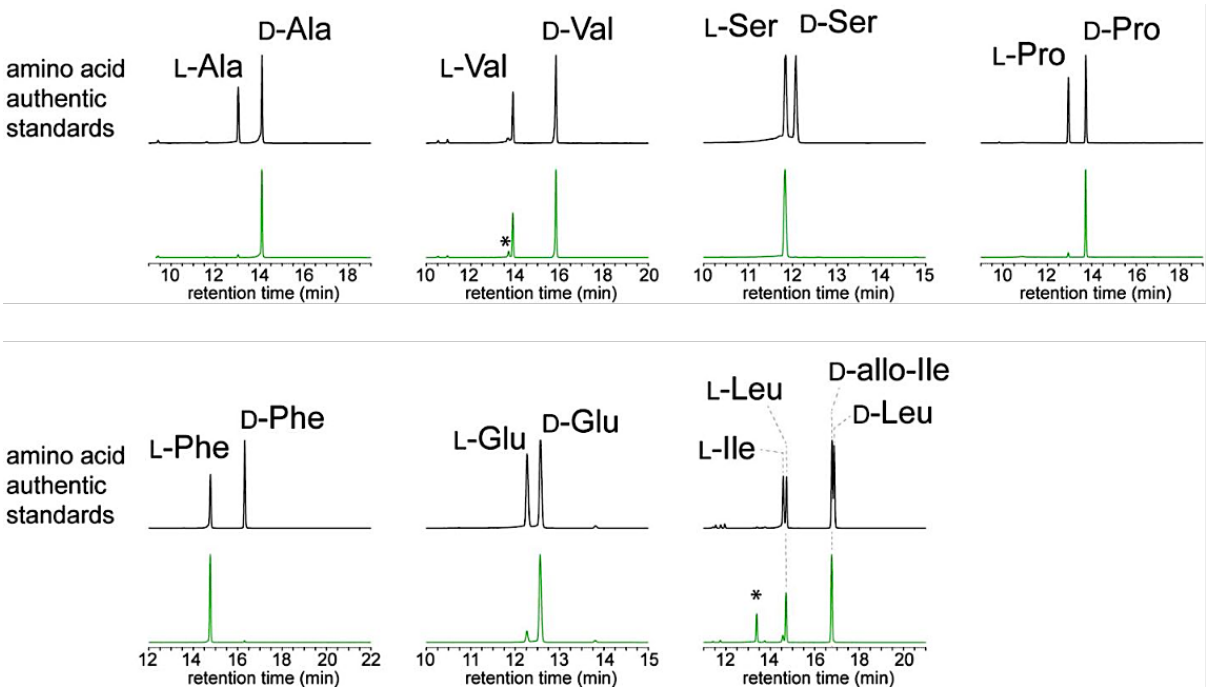


Figure 3-7 Chiral analysis of amino acids in metabolite **6** (Ala, Val, Ser, Pro, Phe, Glu, Leu, and Ile, ESI negative ion mode). Chromatograms were monitored by m/z 382 for Ala, 410 for Val, 398 for Ser, 408 for Pro, 458 for Phe, 440 for Glu (hydrolyzed product from Gln), 424 for Leu/Ile. Peaks marked with an asterisk are not related to amino acids.

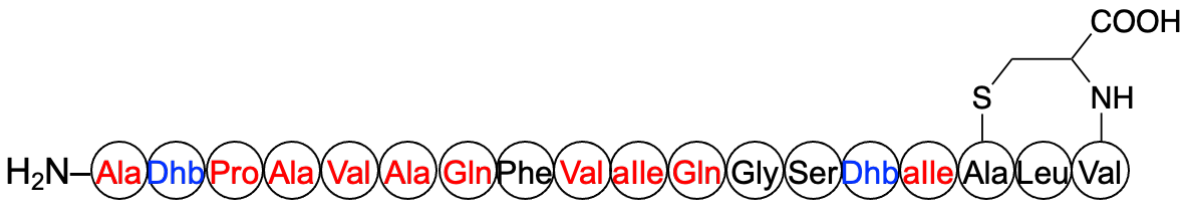


Figure 3-8 Proposed structure of compound **6**. D-amino acids are shown in red, and dehydroamino acids are shown in blue. A C-terminal lanthionine residue likely arose from Michael addition of the Cys19 residue to Dha16 in a proposed intermediate[1].

To further investigate the fate of the precursor peptide in these transformants, I purified GrmA by Ni-NTA affinity chromatography from cell extracts and analyzed by LC-MS and SDS-PAGE. In all the cases, expression of GrmA was observed on SDS-PAGE (Figure 3-9). However, no dehydration product was observed by LC-MS (Figure 3-10). Notably, the results revealed that GrmA was observed when grmA was expressed alone and when *grmA* was coexpressed with either *grmH* or *grmL*, but not when *grmA* was coexpressed with both *grmH* and *grmL*.

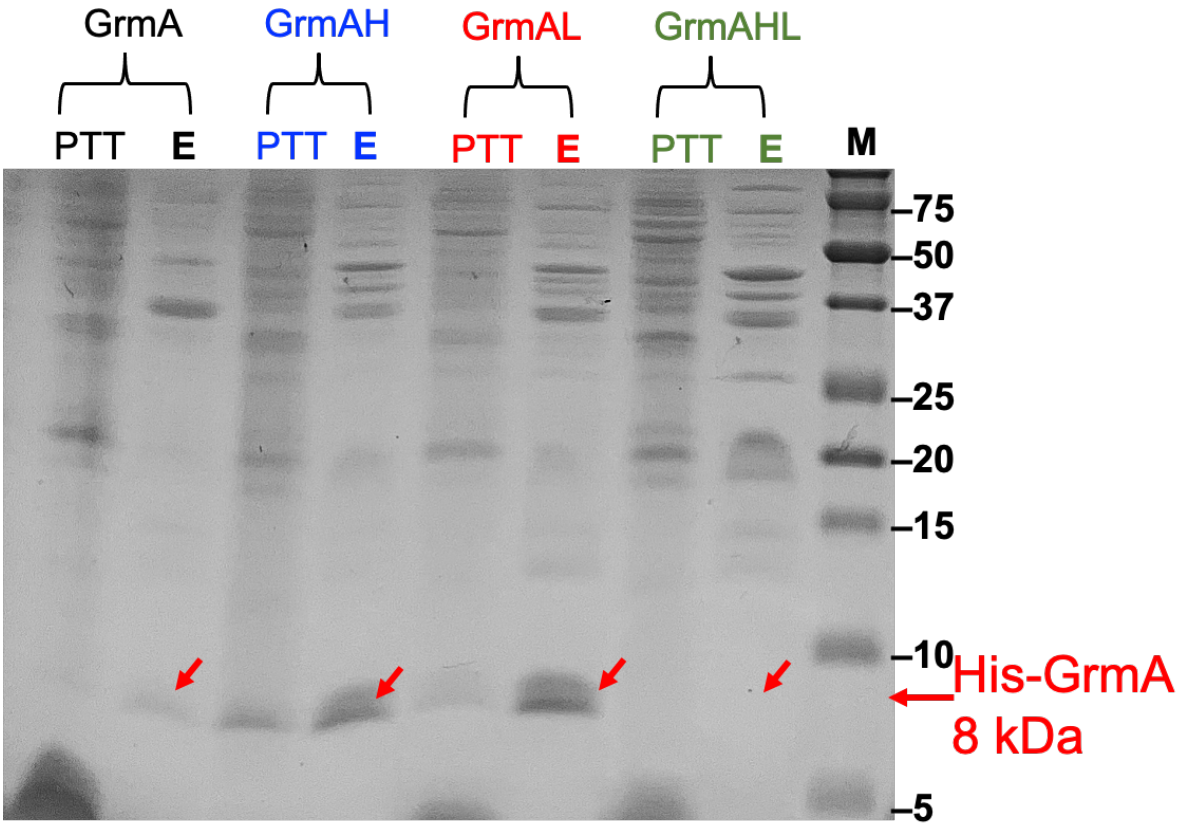


Figure 3-9 SDS-PAGE of *in vivo* coexpression in *S. lividans*; PTT, insoluble proteins; E, purified proteins by Ni-NTA affinity column.

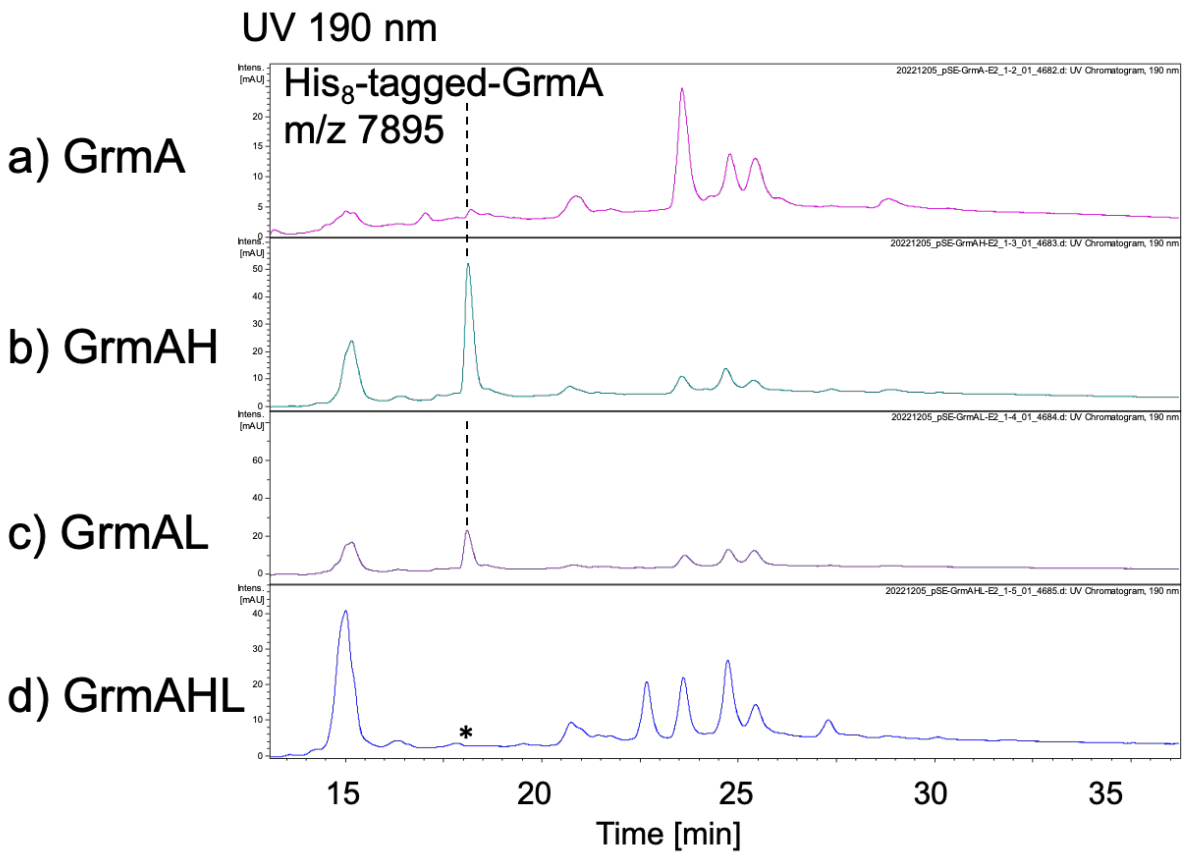


Figure 3-10 LC-MS analysis of purified GrmA from *S. lividans* cells expressing a) *grmA* alone, b) *grmA* and *grmH*, c) *grmA* and *grmL*, and d) *grmA*, *grmH*, and *grmL*. Chromatograms were monitored with UV absorption at 190 nm. A peak marked with an asterisk could not be identified as a compound.

The observed GrmA was further purified by SDS-PAGE to a single band (Figure 3-11) and the chirality of GrmA was analyzed by the Marfey's method. However, no epimerization product was observed as shown in Figure 3-12.

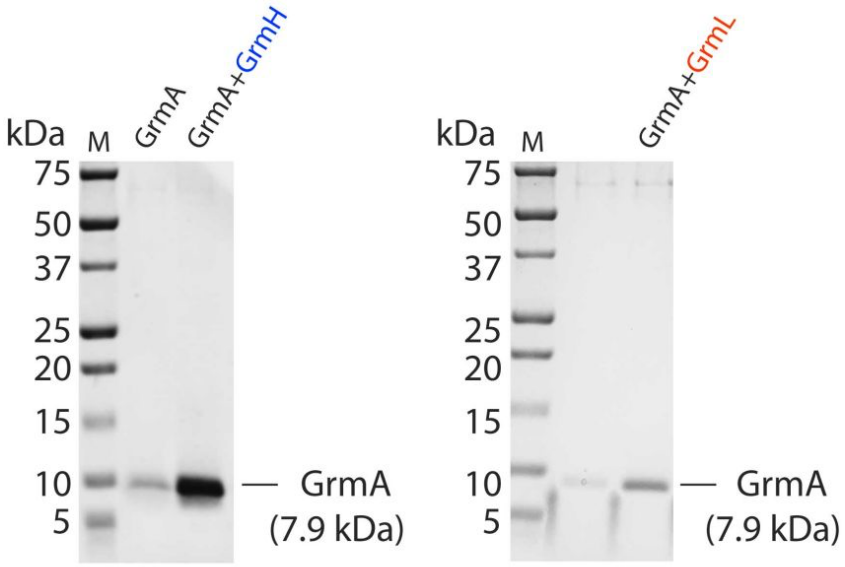


Figure 3-11 SDS-PAGE analysis of purified GrmA. GrmA was purified from *S. lividans* transformants harboring either pSE101N8H-grmA, pSE101N8H-grmAH, or pSE101N8H-grmAL.

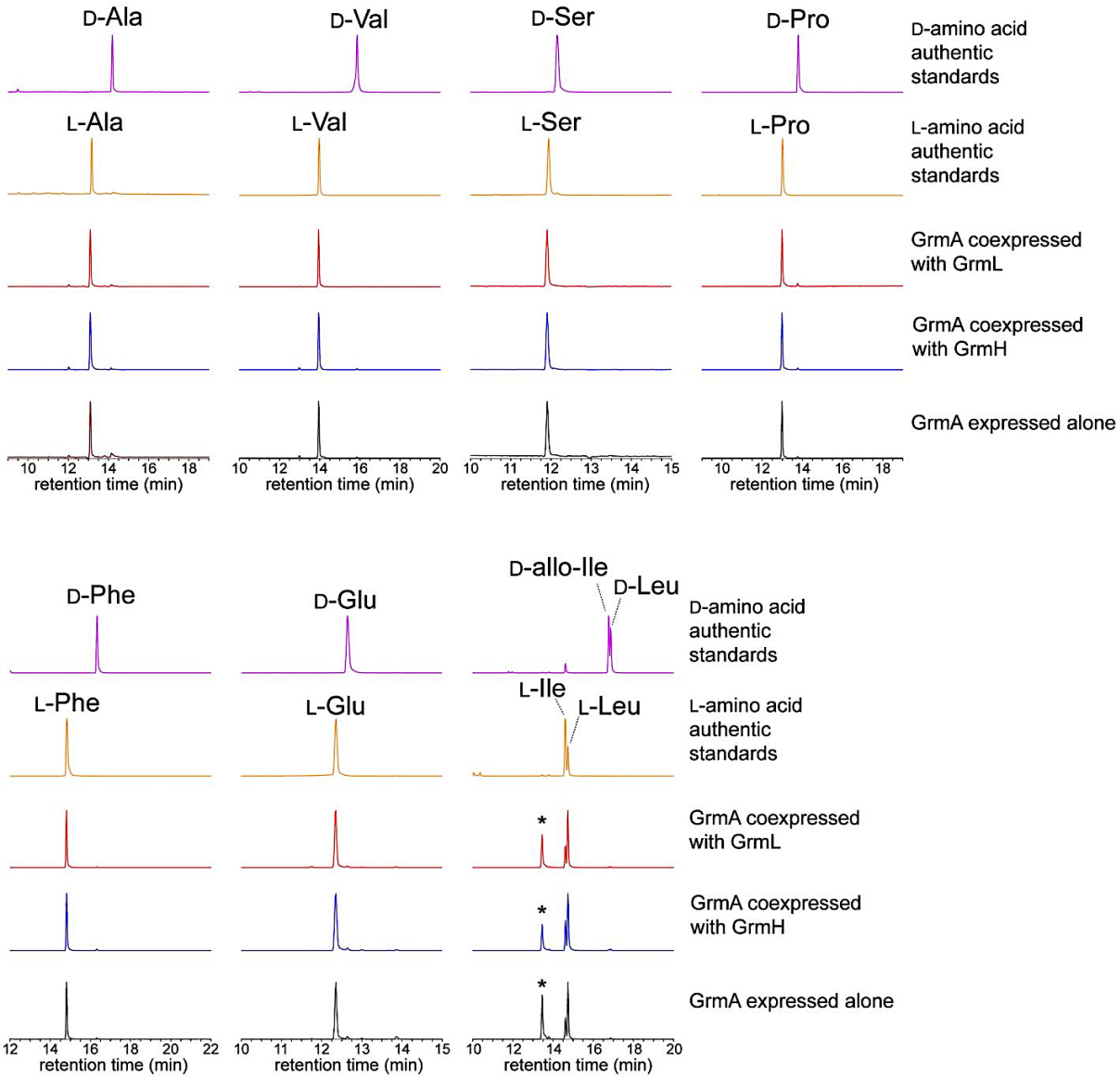


Figure 3-12 LC-MS analysis of L-FDLA derivatives of amino acids in GrmA prepared from *S. lividans* (Ala, Val, Ser, Pro, Phe, Glu, Leu, and Ile, ESI negative ion mode). Chromatograms are monitored by m/z 382 for Ala, 410 for Val, 398 for Ser, 408 for Pro, 458 for Phe, 440 for Glu (hydrolyzed product from Gln), 424 for Leu/Ile. Peaks marked with an asterisk are not related to Leu, Ile, or *allo*-Ile.

These observations suggested that neither GrmH nor GrmL alone could process GrmA. Because both LC-MS and SDS-PAGE analysis of cell extracts indicated that proteolytic cleavage between the leader and core peptides occurred only when *grmA* was coexpressed with both *grmH* and *grmL*, removal of the leader peptide may not be due to endogenous protease in the heterologous host.

Taken together, these findings revealed that GrmH and GrmL catalyze the dehydration, epimerization, and removal of the leader peptide on GrmA .

3.2.4. Analysis of the substrate tolerance of GrmH and GrmL

I next examined the substrate tolerance of GrmH and GrmL by replacing *grmA* gene in pSE101N8H-grmAHL with *sluA* and *sbiA,* which are the precursor peptide genes from the salinipeptin cluster of *S. luteocolor* and the cypemycin cluster of *S. bikiniensis*, respectively.

Although the amino acid sequences of GrmA, SluA, and SbiA are similar to each other, the core peptide in grisemycin is shorter by three amino acids compared to those in salinipeptin and cypemycin (Fig. 3-13).

```
GrmA -------MRLDS-IATQETATALPESMATQDFANSVLAGAVPGFHSDAETPAMATP---AVAQFVIQGSTICLVC* 64
SbiA -------------MTRQDSTAVHPEVLAAQDFANTVLEGAAPGFHSNCETPAMATPATPTVAQFVIQGSTICLVC* 62
SluA -------MQSESTVTRPDS-AIRPEALAAQEFANTVLSGAAPGFHADCETPAMATPATPTAAQFVIQGSTICLVC* 67
```

Figure 3-13 The sequence of nascent precursor peptides GrmA, SbiA, and SluA. Core peptides are underlined.

After cultivation of the *S. lividans* expressing *sluA* with *grmH* and *grmL* (pSE101N8H-sluA-grmH/L) and *sbiA* with *grmH* and *grmL* (pSE101N8H-sbiA-grmH/L), metabolites were analyzed by LC-MS (Figure 3-14). Specific metabolites **7** (m/z 1043.5439, $[M+2H]^{2+}$, calcd. for $C_{96}H_{148}N_{24}O_{26}S$: 1043.5431) and **8** (m/z 1057.5584, $[M+2H]^{2+}$, calcd. for $C_{98}H_{152}N_{24}O_{26}S$: 1057.5588) were detected in the *S. lividans* transformants harboring pSE101N8H-sluA-grmH/L and pSE101N8H-sbiA-grmH/L, respectively, and the observed molecular mass was consistent with expected dehydrated products of the core peptides in both cases.

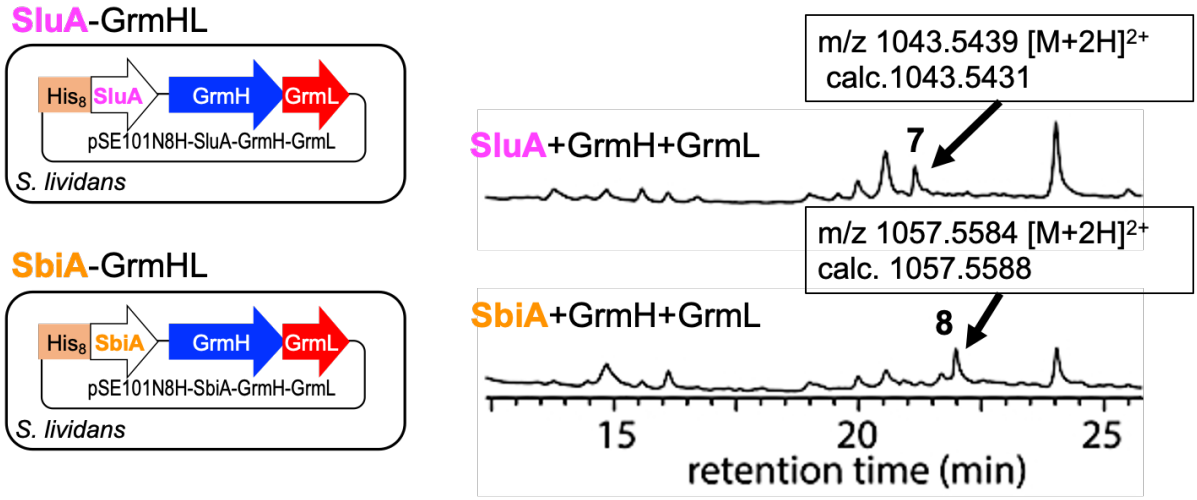


Figure 3-14 LC-MS analysis of the extract of cells expressing a) SluA + GrmH + GrmL, and b) SbiA + GrmH + GrmL. The chromatograms were monitored by UV absorption at 254 nm.

Furthermore, the chiral analysis revealed that both **7** and **8** contained D-amino acid residues almost identical to those in salinipeptin and cypemycin, respectively (Figure 3-15, Figure 3-16), except that **8** contained both D- and L-Gln while all Gln residues in cypemycin have D-configuration.

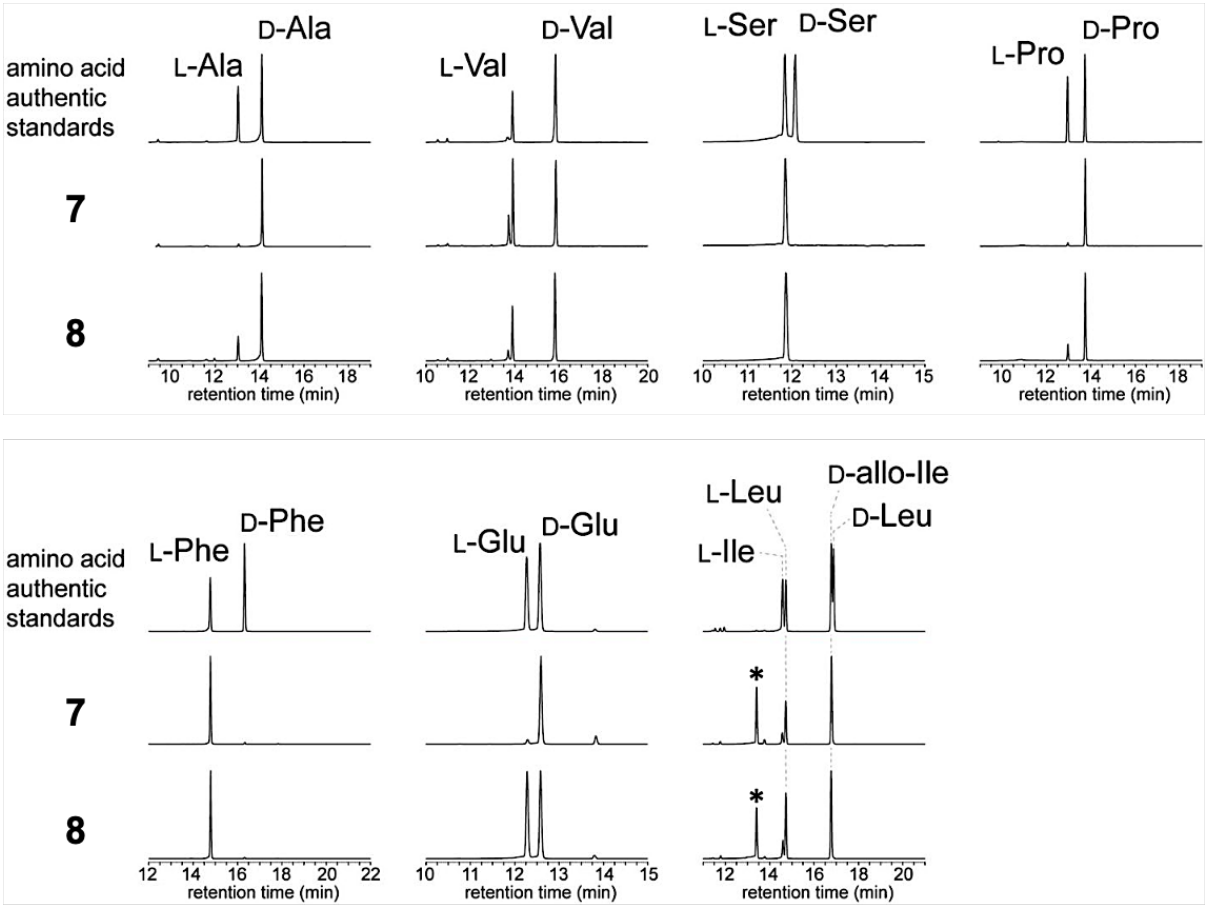


Figure 3-15 LC-MS analysis of L-FDLA derivatives of amino acids in compounds **7** and **8** (Ala, Val, Ser, Pro, Phe, Glu, Leu, and Ile, ESI negative ion mode). Chromatograms are monitored by m/z 382 for Ala, 410 for Val, 398 for Ser, 408 for Pro, 458 for Phe, 440 for Glu (hydrolyzed product from Gln), 424 for Leu/Ile. Peaks marked with an asterisk are not related to amino acid monitored.

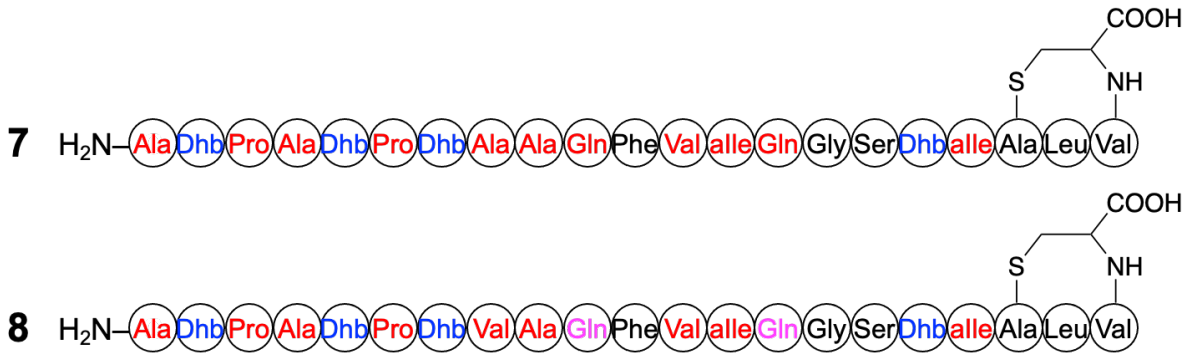


Figure 3-16 Proposed structure of compounds **7** and **8**. D-amino acids are shown in red, and dehydroamino acids are shown in blue. Gln is shown in pink, because Gln residue was unable to be mapped with L- or D-configuration. A C-terminal lanthionine residue likely arose from Michael addition of the Cys22 residue to Dha19 in a proposed intermediate[1].

These results indicated that GrmH and GrmL accept the different length of precursor peptides. It expanded insight into the substrate tolerance of modification enzymes (LinH and LinL) and would be useful to generate unnatural linaridins by precursor peptide bioengineering.

3.2.5. Analysis of the interaction between GrmH and GrmL

To further investigate the function of GrmH and GrmL, the structures were analyzed by AlphaFold 2 using the LocalColabFold program[7], which suggested a strong interaction between GrmH and GrmL (Figure 3-17).

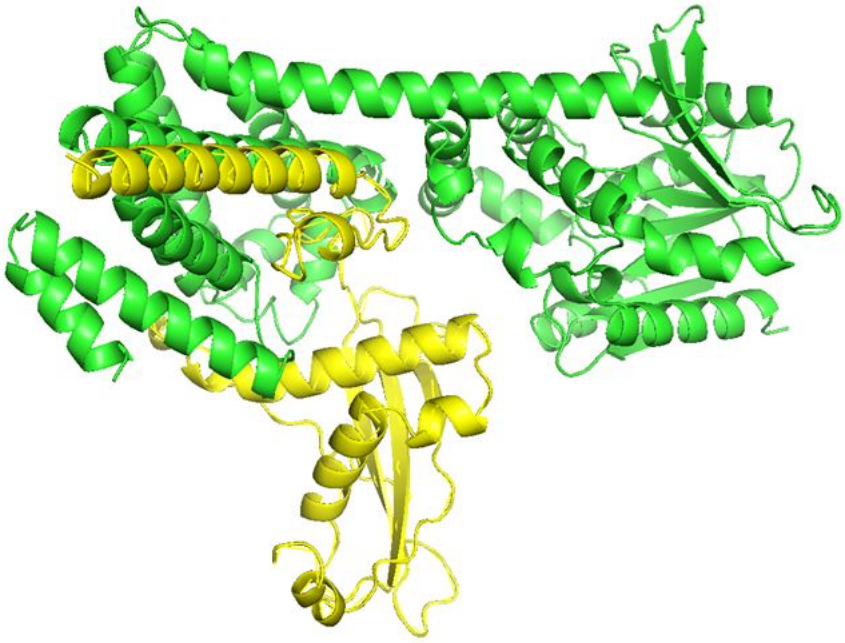


Figure 3-17 AlphaFold predicted GrmH/GrmL complex structure generated with the LocalColabFold program. The structure was analyzed and visualized with PyMol Software. GrmH and GrmL are shown in lime color and yellow color, respectively.

I next confirmed whether GrmH and GrmL form a complex by carrying out a pulldown assay using His-tagged GrmL (or GrmH) and its untagged counterpart. Although I did not obtain GrmL and GrmH as soluble proteins when I expressed the individual genes in *S. lividans*, I succeeded in coexpression of His-tagged GrmL with untagged GrmH as soluble proteins upon additional coexpression of the untagged GrmA (Figure 3-18 a). After affinity purification with Ni-NTA resin, I observed two bands (1:1 molar ratio based on the band intensity) with the expected molecular weight of His-tagged GrmL (~21 kDa) and GrmH (~63 kDa) on SDS-PAGE (Figure 3-18 b).

LC-MS analysis further confirmed that the copurified sample contained GrmH (observed: 63242.6 Da and calculated for $C_{2820}H_{4480}N_{830}O_{813}S_7$: 63242.4 Da) and His-tagged GrmL (observed: 21163.3 Da and calculated for $C_{952}H_{1448}N_{274}O_{269}S_4$: 21162.7 Da) (Figure 3-19).

Although the formation of the CypH/CypL complex was previously proposed by AlphaFold2 analysis[1], this is the first experimental evidence that LinH and LinL form a stable complex.

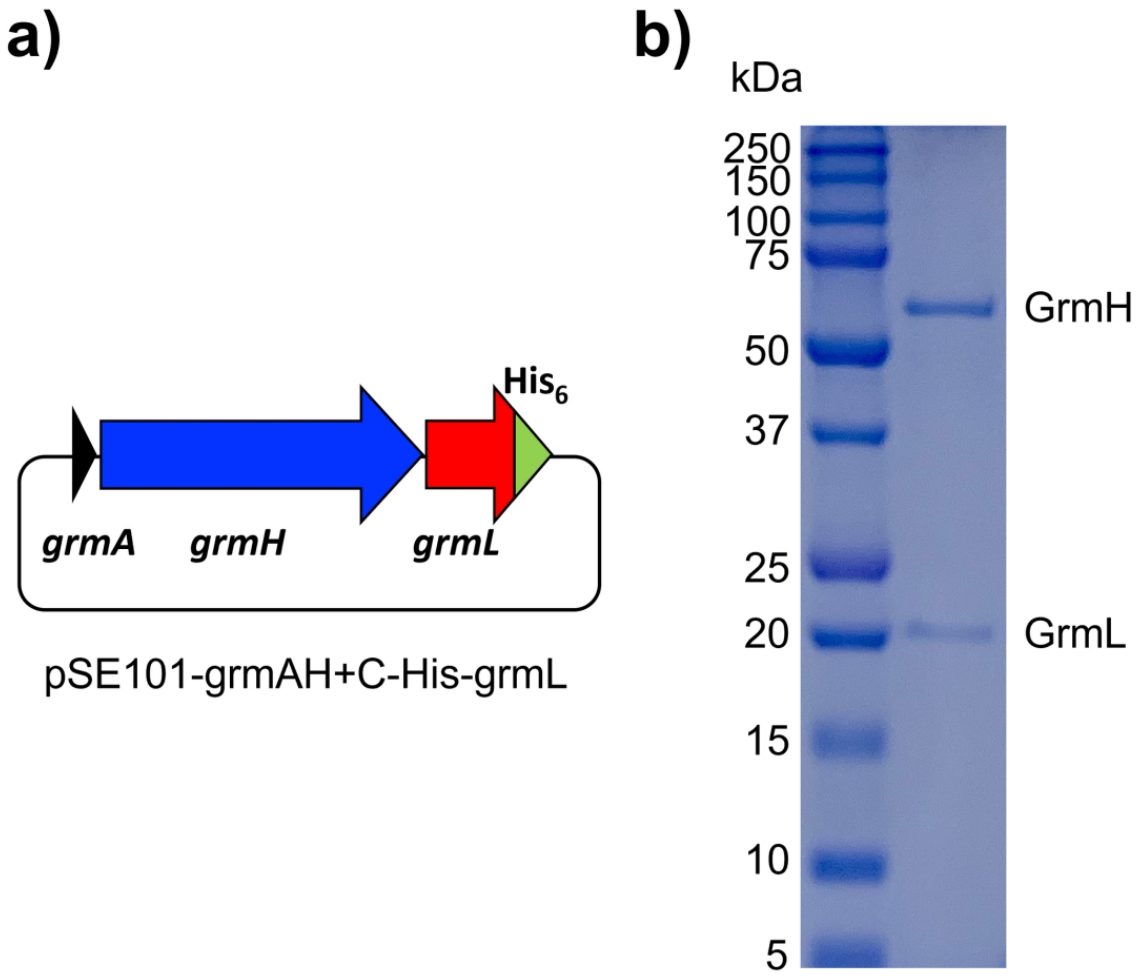


Figure 3-18 a) The plasmid to express *grmA*, *grmH*, and *grmL* in *S. lividans*, and b) SDS-PAGE analysis of copurified GrmH/GrmL complex. A stoichiometric amount of untagged GrmH (63.2 kDa) was coeluted with His-tagged GrmL (21.3 kDa).

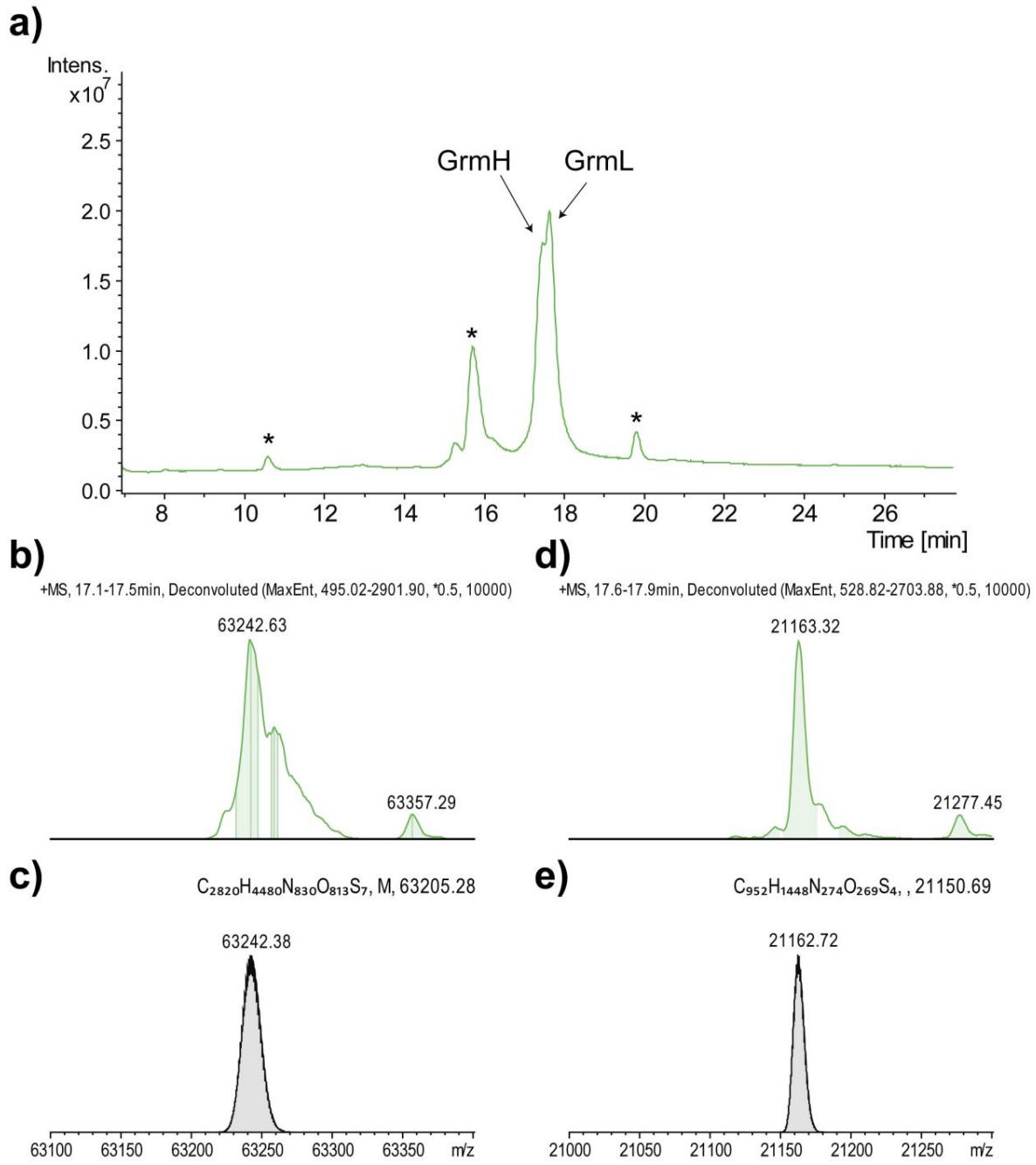


Figure 3-19 LC-MS analysis of copurified GrmH/GrmL complex. a) A total ion chromatogram, b) deconvoluted mass spectrum of GrmH eluted at retention time (17.1–17.5 min), c) simulated mass spectrum of GrmH ($C_{2820}H_{4480}N_{830}O_{813}S_7$), d) deconvoluted mass spectrum of GrmL eluted at retention time (17.6–17.9 min), and e) simulated mass spectrum of GrmL ($C_{952}H_{1448}N_{274}O_{269}S_4$). Asterisks denote unrelated small molecules.

Because I obtained the copurified GrmH/GrmL complex in soluble form, I next performed *in vitro* analysis of the GrmH/GrmL reaction. To prepare full-length GrmA as the substrate in large amounts, I used *E. coli* as the heterologous host as mentioned in previous sections. Considering the substrate tolerance of GrmH and GrmL as described in the previous section, I also prepared the precursor peptides of SluA and SbiA as substrates of GrmH/L complex in a similar manner to GrmA. Because *in vivo* analysis showed that compound **6** was observed only when *S. lividans* was used as the heterologous host, the GrmH/GrmL reaction may require cofactors present in *S. lividans*. Thus, I incubated GrmA with the GrmH/GrmL complex in the presence of a cell-free extract prepared from *S. lividans,* and the reaction was analyzed by LC-MS.

However, the formation of **6** (m/z 926, $[M+2H]^{2+}$) or another product was not observed in the reaction mixture, despite many different conditions being tried. In addition, incubation of SluA or SbiA with the GrmH/GrmL complex and the cell-free extract did not generate a product corresponding to **7** (m/z 1043.5, $[M+2H]^{2+}$) for SluA or **8** (m/z 1057.5, $[M+2H]^{2+}$) for SbiA.

Considering that epimerization and dehydration activities were detected only by the simultaneous coexpression of *grmA*, *grmH*, and *grmL*, simultaneous enzyme expression and the successive formation of a complex structure of the three enzymes might be necessary for the dehydration and epimerization reactions. Otherwise, other conditions such as an anaerobic environment or the addition of a macromolecule fraction from *S. lividans* may be required for the

enzymatic activity of the GrmH/GrmL complex.

3.3 Conclusion

In this chapter, I demonstrated that neither GrmH nor GrmL alone could process GrmA but that a GrmH/GrmL complex can catalyze dehydration, epimerization, and proteolytic cleavage of a precursor peptide GrmA. Moreover, I experimentally revealed that GrmH (LinH) and GrmL (LinL) constitute a protein complex by a pulldown assay with His-tagged GrmL and untagged GrmH. Although the reactions catalyzed by the GrmH/GrmL complex have yet to be carefully investigated by additional *in vitro* analysis, these results provide new functionality to catalyze both dehydration and epimerization by the GrmH/GrmL complex involved in the biosynthesis of RiPP natural products.

3.4 Experimental Section

3.4.1 General

Chemicals, enzymes, and kits for DNA manipulations, and primers for PCR were obtained from the same supplier as those described in Chapter 2. MS and MS/MS data were recorded by the same methods as described in Chapter 2. Heterologous expression and chiral analysis of the products were carried out by the same methods as described in Chapter 2. A vector pSE101N8H was kindly provided by Prof. Tomohisa Kuzuyama at The University of Tokyo.

3.4.2 *In vitro* and *in vivo* analysis in *E. coli*.

To construct the plasmids of pCDFD-His_6-grmL, pACYCD-His_6-GrmH, pET21a-SKIK-His_6-grmA: A primer pair grmL-F(BamHI)/grmL-R(HindIII) was used to amplify the grmL gene and the PCR products, digested with *Bam*HI and *Hind*III, were cloned into the same sites of pCDF-Duet or pRSF-Duet to obtain pCDF-His_6-grmL or pRSF-His_6-grmL. A plasmid pACYC-His_6-GrmH was constructed with a primer pair, grmH-F(BamHI)/grmH-R(HindIII), and pACYC-Duet vector similarly. For the expression of SKIK- and His-tagged GrmA, a DNA fragment containing the grmA gene was amplified with the primer pair, grmA-F(NdeI)/grmA-R(HindIII), and cloned into pET21a vector using *Nde*I and *Hind*III restriction enzymes to obtain pET21a-SKIK-His_6-grmA. Oligonucleotides used for plasmid construction are summarized in Table 3-1. All plasmids were confirmed by DNA sequencing.

The plasmids, pCDF-His_6-grmL, pACYC-His_6-GrmH, pET21a-SKIKHis_6-GrmA, pET21a-SKIKHis_6-GrmA/pACYC-His_6-GrmH, pET21a-SKIKHis_6-GrmA/ pRFS-His_6-GrmL, or pET21a-SKIKHis_6-GrmA/ pACYC-His_6-GrmH/ pRFS-His_6-GrmL, were transferred into *E. coli* BL21 (DE3) (Merck) for expression. *E. coli* transformants were grown in LB medium (2 mL) supplemented with appropriate antibiotics. After aerobic cultivation for 12–14 h at 37 °C, the cultures were transferred to fresh LB medium (50 mL) containing appropriate antibiotics and further cultivated (37 °C, 200 rpm) until OD_{600} reached 0.6–0.8. The culture was cooled down to 4 °C in iced water, to which isopropyl β-D-1-thiogalactopyranoside (0.5 mM) was added, followed by cultivation with shaking at 200 rpm, for 16 h at 16°C. The cell pellets were obtained by centrifugation (15,000 ×rpm, 4°C, 5 min), and washed with buffer I (50 mM sodium phosphate, 300 mM NaCl, and 20 mM imidazole, pH 8.0). The pellets were frozen with liquid nitrogen and stored at –80 °C for use. When using cell pellets were resuspended in 5 mL of buffer A and disrupted by sonication using an ultrasonic disruptor (Branson SFX 250). After centrifugation at 15,000 ×rpm for 30 min, the His-tagged recombinant proteins were adsorbed onto Ni-NTA agarose resin (Qiagen), washed with wash buffer II (50 mM sodium phosphate, 300 mM NaCl and 50 mM imidazole, pH 8.0), and eluted with elution buffer (50 mM sodium phosphate, 300 mM NaCl, and 250 mM imidazole, pH 8.0). Purified protein was rebuffered with 50 mM Tris-HCl (pH 8.0). The purified proteins were analyzed by SDS-PAGE, LC-MS, and Marfey's method as

described in Chapter 2.

3.4.3 In vivo analysis in S. lividans.

The constructed plasmids were designated as pSE101-N8H-grmA, pSE101-N8H-grmAH, and pSE101-N8H-grmAHL. A plasmid, pSE101-N8H-grmAL, which lacked *grmH,* was also constructed by in-frame deletion of an internal region of *grmH* as described in Chapter 2. The DNA fragments containing grmA alone, grmA–grmH, and grmA–grmL were amplified by PCR from the plasmid pWHM3-grm with the primer pairs of grmA-F(HindIII)/grmA-R(XbaI), grmA-F(HindIII)/grmH-R(XbaI), and grmA-F(HindIII)/grmL-R(XbaI), respectively. Each PCR product was cloned into the *Hind*III–*Xba*I site of the pSE101N8H vector so that the *N*-terminal 8×His-tag and GrmA sequences were expressed in a single open reading frame. The deduced amino acid sequence of the precursor peptide was:

“MHHHHHHHHKLRLDSIATQETATALPESMATQDFANSVLAGAVPG FHSDAETPAMATPAVAQFVIQGSTICLVC”.

Each plasmid was introduced into *S. lividans* and the transformant was grown at 30°C, 200 rpm in Tryptic Soy Broth medium (15 mL) supplemented with 8 μg/mL of thiostrepton. After 2 days, 1 mL of the culture was transferred into 50 mL YEME medium supplemented with 8 μg/mL of thiostrepton in 250 mL Erlenmeyer flask, and cells were grown at 28°C, 200 rpm for 3 days. The cells were harvested from the culture by centrifugation (15,000 ×rpm, 10 min)

and washed with water. To search for metabolites produced by the transformants, the pellets were resuspended with 30 mL of 20% (v/v) acetone or 100% methanol and agitated for 1 hour. The supernatant obtained by centrifugation (15,000 rpm, 30 min) was dried in vacuo and solubilized with 600 µL of milliQ. The obtained samples were analyzed by HPLC and LC-MS.

To construct pSE101N8H-sluA-grmH/L, the DNA fragments containing *sluA* were amplified by PCR with the primer pair of sluA-F(HindIII)/sluA-R. Independently, the DNA fragment containing *grmH* and *grmL* was prepared with the primer pair grmH-F/grmL-R(XbaI). Two DNA fragments were then fused with overlap extension PCR with the primer pair sluA-F(HindIII)/grmL-R(XbaI) and the resulting PCR product was cloned into *Hind*III–*Xba*I sites of the pSE101N8H vector. Likewise, pSE101N8H-sbiA-grmH/L was prepared using primers of sbiA-F(HindIII)/sbiA-R. The deduced amino acid sequences of SluA and SbiA were

“MHHHHHHHHKLQSESTVTRPDSAIRPEALAAQEFANTVLSGAAPG FHADCETPAMATPATPTAAQFVIQGSTICLVC” and

“MHHHHHHHHKLTRQDSTAVHPEVLAAQDFANTVLEGAAPGFHSN CETPAMATPATPT VAQFVIQGSTICLVC”, respectively.

To construct plasmid, pSE101-grmAH/C-His_6-grmL: The DNA fragment containing genes *grmAHL* was amplified with the primer pair of grmA-F(HindIII)/grmAHL_C-His-R(XbaI) to introduce a 6×His-tag sequence at the C-terminus of the *grmL* gene. The fragment was then cloned into pSE101 using

*Hind*III and *Xba*I restriction enzymes.

3.4.4 Analysis of the interaction between GrmH and GrmL by a pulldown assay.

The interaction between GrmH and GrmL was evaluated by a pulldown assay using His-tagged GrmL and untagged GrmH. An *S. lividans* transformant harboring pSE101-grmA/H-C-His_6-grmL was cultivated in YEME medium (300 ml) for 3 days. The cells were collected by centrifugation, lysed by sonication, and His-tagged proteins were purified using Ni-NTA resin as described above. The elution fraction (~300 μl) containing GrmL and the proteins copurified with GrmL was analyzed by SDS-PAGE and LC-MS. LC-MS analysis was performed with an HPLC instrument (Agilent Technologies Inc.) equipped with a Maxis Plus (Bruker). The analytical conditions were as follows: column, Sunshell C8-30HT (2.1 × 150 mm, 3.4 μm, ChromaNik Technologies Inc., Osaka, Japan); flow ratio, 0.3 ml/min; temperature, 70 °C; mobile phase, (A) 0.1% TFA in water and (B) 0.1% TFA in acetonitrile; gradient conditions, 30%–60% B (30 min); detection, 280 nm, and positive ion mode; injection volume, 10 μl. The mass spectra of multiply charged ions were deconvoluted using the DataAnalysis ver. 4.3 software.

3.4.5 *In vitro* reactions of the GrmH/GrmL complex with GrmA, SluA, or SbiA

The plasmids of pET21a-SKIK-His_6-sbiA, and pET21a-SKIK-His_6-sluA were prepared as follows: The DNA fragment containing the *sbi* or *slu* gene was amplified with primer pairs of sluA-F(NdeI)/sluA-R(HindIII) and sbiA-F(NdeI)/sbiA-R(HindIII) were used to construct pET21a-SKIK-His-sluA and pET21a-SKIK-His-sbiA, respectively, and cloned into pET21a vector using *Nde*I and *Hind*III restriction enzymes to obtain the plasmids. The recombinants were purified as described in Chapter 2.

Reaction mixtures (50 μL) containing the precursor peptides (100 μM), the GrmH/GrmL complex (4 μM), and the cell-free extract in reaction buffer (50 mM Tris-HCl, pH 8.0) of *S. lividans* were incubated at 30 °C overnight. The reactions were quenched by the addition of 50 μL of methanol. After centrifugation (16 000 ×g, 10 min), the supernatants were analyzed by LC-MS as described above. The cell-free extract of *S. lividans* was prepared as follows: The cell pellet (1 g wet weight) of *S. lividans* TK23 was sonicated in 4 mL of reaction buffer. After centrifugation, the supernatant was passed through an Amicon Ultra (MWCO; 3K) to remove macromolecules.

Table 3-1 Oligonucleotides used to clone *grmA*, *grmH*, and *grmL* for *in vitro* and *in vivo* analysis in *E. coli* and *S. lividans* as the heterologous hosts. The introduced restriction sites are underlined. SKIK and His tag sequence are shown in bold letters.

primer name	sequences (5′–3′)
grmL-F(BamHI)	CACCACAGCCAGGATCCAAGTAGCCGGATTCGTGGTG
grmL-R(HindIII)	TTATGCGGCCGCAAGCTTATCCAGCGGTTTCCTTGTGTTC
grmH-F(BamHI)	CACCACAGCCAGGATCCGACACTCGCGGCGGCGACCG
grmH-R(HindIII)	TTATGCGGCCGCAAGCTTTACTTGGCATGTATGAACCTTCGGATGG
grmA-F(NdeI)	TCTCGAACATATG**TCTAAAATAAAACACCACCACCACCAC**CACCGACTCGATTCGATTGCAACG
grmA-R(HindIII)	ACAGAAGCTTTCAGCAGACGAGACAGATGGTGC
sluA-F(NdeI)	GAGATATACATATG**TCTAAAATAAAACACCACCACCACCA**CCACCAATCTGAATCGACTGTGACGCGACCGG
sluA-R(HindIII)	TATATAAGCTTCAGCAGACGAGGCATATCGTGCTGCCCT
sbiA-F(NdeI)	GAGATATACATATG**TCTAAAATAAAACACCACCACCACCA**CCACACGCGACAGGATTCCACTGC
sbiA-R(HindIII)	TATATAAGCTTCAGCAGACCAGGCAAATGGTGCTGCCCTGGATGAC
grmA-F(HindIII)	ATATAAGCTTCGACTCGATTCGATTGCAACGCAGGAGA
grmA-R(XbaI)	ATATTCTAGAGTCAGCAGACGAGACAGATGGTGCTGCCC
grmH-R(XbaI)	ATATTCTAGAGCTACTTGGCATGTATGAACCTTCGGATG
grmL-R(XbaI)	ATATTCTAGAGTCATCCAGCGGTTTCCTTGTGTTCGGACGCG
grmH-F	TGGTGGGCTGACTCGCGTGTCGGAGGGG

sluA-F(HindIII)	CACCACCACAAGCTTCAATCTGAATCGACTGTGACGCGACCGG
sluA-R	CGAGTCAGCCCACCATCAGCAGACGAGGCATATCGTGCTGCCC
sbiA-F(HindIII)	CACCACCACAAGCTTACGCGACAGGATTCCACTGCCGTCCACC
sbiA-R	CGAGTCAGCCCACCATCAGCAGACCAGGCATATGGTGCTGCCC
grmAHL_C-His-R(XbaI)	ATATTCTAGAGTCAGTGGTGGTGGTGGTGGTGTCCAGCGGTTTCCTTGTGTTCGGACGC

3.5 References

1. Xue, Y., Wang, X. & Liu, W. Reconstitution of the Linaridin Pathway Provides Access to the Family-Determining Activity of Two Membrane-Associated Proteins in the Formation of Structurally Underestimated Cypemycin. *J. Am. Chem. Soc.* **145**, 7040–7047 (2023).

2. Feng, Z., Ogasawara, Y. & Dairi, T. Identification of the peptide epimerase MslH responsible ford-amino acid introduction at the C-terminus of ribosomal peptides. *Chem. Sci.* **12**, 2567–2574 (2021).

3. Feng, Z., Ogasawara, Y., Nomura, S. & Dairi, T. Biosynthetic Gene Cluster of a d-Tryptophan-Containing Lasso Peptide, MS-271. *ChemBioChem* **19**, 2045–2048 (2018).

4. Shi, Y., Yang, X., Garg, N. & Van Der Donk, W. A. Production of lantipeptides in Escherichia coli. *J. Am. Chem. Soc.* **133**, 2338–2341 (2011).

5. Xue, Y., Wang, X. & Liu, W. Reconstitution of the Linaridin Pathway Provides Access to the Family-Determining Activity of Two Membrane-Associated Proteins in the Formation of Structurally Underestimated Cypemycin. *J. Am. Chem. Soc.* **145**, 7040–7047 (2023).

6. Mo, T. *et al.* Biosynthetic Insights into Linaridin Natural Products from Genome Mining and Precursor Peptide Mutagenesis. *ACS Chem. Biol.* **12**, 1484–1488 (2017).

7. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).

Chapter 4

Conclusion

Salinipeptin A, which is a member of type-A linaridins, was reported to comprise 22 amino acid residues with multiple D-amino acid residues, but no gene homologous to known epimerases existed in the biosynthetic gene cluster. Bioinformatic analysis suggested that an unknown functional gene *SinL* might encode a novel epimerase. In this study, I examined this plausibility.

In Chapter 2, I performed heterologous expression experiments of three type-A linaridin biosynthetic gene clusters, including *grm* (grisemycin), *cyp* (cypemcyin), and *sin* (salinipeptins), and revealed that the presence of multiple D-amino acids is a common feature of type-A linaridins. Importantly, these observations strongly indicated that type-A linaridin clusters contain a gene, *linL*, encoding a novel peptide epimerase. Furthermore, gene-deletion experiments of *grm* cluster indicated that *grmL* and *grmH* are indispensable for grisemycin production and that epimerization and dehydration precede decarboxylation and methylation.

In Chapter 3, I performed *in vivo* experiments in *S. lividans*, and the results suggested that neither GrmH nor GrmL alone can process GrmA but that a GrmH/GrmL complex can catalyze dehydration, epimerization, and proteolytic cleavage of the precursor peptide GrmA. In addition, I experimentally revealed that GrmH and GrmL constitute a protein complex.

Acknowledgment

At the beginning, I would like to express my deepest appreciation and gratitude to my supervisor, Prof. Tohru Dairi, for his guidance and support for five years during both my master's and doctoral studies. His invaluable guidance and kind suggestions have given me enormous help towards my research. Besides, I have learned from him not only technical knowledge but also his attitude as a researcher.

Secondly, I would like to express my greatest thanks to Associate Professor Dr. Yasushi Ogasawara, Assistant Professor Dr. Yasuharu Satoh, and Research Assistant Professor Dr. Takeshi Tsunoda, for their help and encouragement on my research. Appreciations are also extended to all my lab mates for their support and enjoyable days with them.

Last but not least, I am extremely grateful to my parents for their endless encouragement and unconditional love.