© 2011 JUAN ESTEBAN VELÁSQUEZ

BIOSYNTHESIS OF THE ANTIMICROBIAL PEPTIDE EPILANCIN 15X

BY

JUAN ESTEBAN VELÁSQUEZ

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor Wilfred A. van der Donk, Chair
Professor William W. Metcalf
Professor Eric Oldfield
Professor Scott K. Silverman

ABSTRACT

Lantibiotics are polycyclic antimicrobial peptides that are ribosomally synthesized and posttranslationally modified to their biologically active forms. The recently discovered lantibiotic epilancin 15X produced by *Staphylococcus epidermidis* 15X154 is active against several pathogenic bacteria, including methicillin-resistant *S. aureus* and vancomycin-resistant Enterococci. Epilancin 15X contains an unusual N-terminal lactate group (Lac) that could be important for its bioactivity. To understand its biosynthesis, the epilancin 15X gene cluster was identified. The Lac group is introduced by dehydration of a Ser residue in the first position of the core peptide by a lantibiotic dehydratase (ElxB), followed by proteolytic removal of the leader peptide by a serine-type protease (ElxP), and hydrolysis of the resulting N-terminal dehydroalanine residue. A pyruvyl group thus formed is finally reduced to Lac by a dehydrogenase (ElxO). Using substrate analogs synthesized by solid-phase peptide synthesis or heterologously expressed in *Escherichia coli*, the enzymatic activities of ElxO and ElxP and their substrate specificities were investigated *in vitro*. Furthermore, the dehydratase activity of ElxB was demonstrated *in vivo*. ElxO tolerates several structural modifications along the substrate sequence and can catalyze the NADPH-dependent reduction of diverse pyruvyl- and oxobutyryl-containing peptides, such as the lantibiotic lactocin S, to the corresponding alcohols. The reaction mechanism of ElxO was proposed based on its x-ray crystal structure and a mutagenesis study. The characterization of ElxP demonstrated that the protease can cleave the unmodified precursor peptide and relies on the C-terminal residues of the leader region for substrate recognition. Finally, the

stereochemical configuration and the importance of the N-terminal Lac for peptide stability against bacterial aminopeptidases were determined. Understanding the biosynthetic route to epilancin 15X opens the door for engineering of novel lantibiotics with more potent biological activities or different spectra of action.

To my parents Óscar and Oliva
To my brother Óscar
To my aunts
To Consuelo, we miss you so much
To my Champaign family

ACKNOWLEDGMENTS

The current work was possible due to the invaluable contribution of many people. First, I would like to express my gratitude to my adviser Professor Wilfred A. van der Donk. His guidance and assistance were essential for the completion of this process. In difficult times, he always had smart and serene suggestions that allowed me to advance on my research. From him, I have learnt not only vast knowledge, but also that honesty, integrity, and respect must define a truly successful career in science. I would also like to thank my dissertation committee Professors William Metcalf, Scott Silverman, Eric Oldfield, and Neil Kelleher. They always had the availability to discuss my research, but most importantly, their advice and encouragement were especially motivating. Special thanks to Professors Alan Hill and Kenneth Suslick for invaluable support during my transition to graduate school.

I am particularly grateful to former postdoctoral researchers in the van der Donk laboratory Robert Cichillo and Svetlana Borisova. They had the willingness, availability, and patience to teach me uncountable laboratories techniques and to give me valuable advice for my research and for my life as a graduate student. Special thanks to Jun Kai Zhang, Neha Garg, Laura Guest, Isabel Neacato, Xingang Zhang, and Jiangtao Gao that have made important contributions to this project.

My deepest gratitude to my friend Trent Oman. It was an honor to take classes, joint a research group, teach a laboratory course, and work together for five years. Special thanks to Joel Cioni, Despina Bougioukou, Heather Cooke, Jin Hee Lee, Spencer Peck, Kou-San Ju, Annette Erb, Tobias Erb, John Whitteck, Benjamin Griffin,

Benjamin Circello, Xioamin Yu, James Doroghazi, Bradley Evans, Stephanie Bumpus, and Ioanna Ntai for help, discussions, advice, and specially for laughs. It was great working with you side by side.

I am also indebted to previous and current labmates in the van der Donk lab. Thank you to Gregory Patton, Lisa Cooper, Bo Li, Leigh Anne Furgerson Ihnken, Ian Gut, Lindsey Johnstone, Yanxiang Shi, Ayşe Ökesli, John Hung, and Gabrielle Thibodeaux for immeasurable help with biology and meaningful discussions. Thank you to Kevin Clark, Remco Merkx, Noah Bindman, and Patrick Knerr for incalculable support with organic synthesis.

Special thanks to Courtney Evans, Martha Freeland, Debra Piper, and Nancy Holda for managing the administrative details of the laboratories and for their infinite willingness to support my work in the laboratory. I am also grateful to the friendly staff from the Chemical Biology, Organic Chemistry, Graduate Student Services, Career Services, and Graduate Admission Offices from the Department of Chemistry. Thanks to Casondra Anastasiadis, Stacy Olson, Susan Lighty, Becky Duffield, Connie Knight, Dorothy Gordon, Andrea Wynn, Krista Smith, Patricia Simpson, and Karen York. Thank you to Furong Sung, Haijun Yao, Feng Lin, and Vera Mainz from the Mass Spectrometry and NMR Spectroscopy laboratories.

My research projects were funded by the National Institute of Health and the Howard Hughes Medical Institute. Thanks to SACNAS for showing me that it is possible to combine science with my cultural background. Thanks to the NIH Training Program at the Chemistry-Biology Interface (CBI) for exceptional opportunities to meet great professors and be continuously fascinated by science.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION*

1.1. LANTIBIOTICS AND THE PROBLEM OF BACTERIAL RESISTANCE

Over the last several decades, bacterial resistance to known classes of antibacterial agents has been rising at an alarming rate, mainly due to the widespread overuse of antibiotics. The exclusive dependence on broad-spectrum antibiotics has intensified the problem by inducing the development of multidrug-resistant pathogens. For instance, 60% to 70% of all *Staphylococcus aureus* strains isolated in hospitals are currently multidrug-resistant (*1*). In recent years, the number of fatalities due to infections caused by this microorganism in the United States was larger than the number of deaths attributed to AIDS (*1*). Furthermore, the rising prevalence of methicillin-resistant *S. aureus* (MRSA) increases the likelihood that vancomycin-resistant *S. aureus* (VRSA), which is more challenging to treat, will become a new scourge in hospitals (*2*). Despite the rise of resistant pathogens, the rate of discovery and approval of new antibiotics has been dramatically decreasing. Indeed, during the last five decades, only four new classes of antibiotics have been introduced, in part due to high rediscovery rates of old scaffolds and to the unfavorable economics of antibiotic development compared to other drugs (*2, 3*). In order to prevent potential epidemic outbreaks, new antibacterial drug classes not affected by existing resistance mechanisms are therefore much needed.

Lantibiotics are ribosomally synthesized and posttranslationally modified polycyclic peptides that contain thioether cross-links and that demonstrate promising

* Reproduced in part with permission from: "Velásquez, J. E., and van der Donk, W. A. (2011) Genome mining for ribosomally synthesized natural products, *Curr. Opin. Chem. Biol.* 15, 11-21." Copyright © 2010 Elsevier Ltd.

activity against pathogenic bacteria due to their high potency, low frequency of antimicrobial resistance, and lack of significant cytotoxicity (*4*). In addition to the characteristic lanthionine and methyllanthionine rings, members of this family are commonly characterized by the presence of the dehydrated amino acids dehydroalanine (Dha) and dehydrobutyrine (Dhb) and several other unnatural amino acids, such as lysinoalanine and hydroxy-aspartate in cinnamycin; S-aminovinyl-D-cysteine, hydroxyproline, and chlorinated tryptophan in microbisporicin; and D-alanine and N-terminal pyruvyl (Pyr) in lactocin S (Figure 1.1). Nisin, the most studied lantibiotic (Figure 1.2), has been used in food preservation and treatment of bovine mastitis during the last 40 years in more than 80 countries without the development of natural bacterial resistance (*5*), possibly as a consequence of its multiple modes of action (*6*). The N-terminal portion of nisin, as well as several other lantibiotics, recognizes the membrane-bound cell wall precursor lipid II, inhibiting peptidoglycan biosynthesis (Figure 1.3) (*7*-*9*). Once the molecule is docked to the membrane, the C-terminal portion generates stable pores that result in membrane damage and depolarization (*10*).

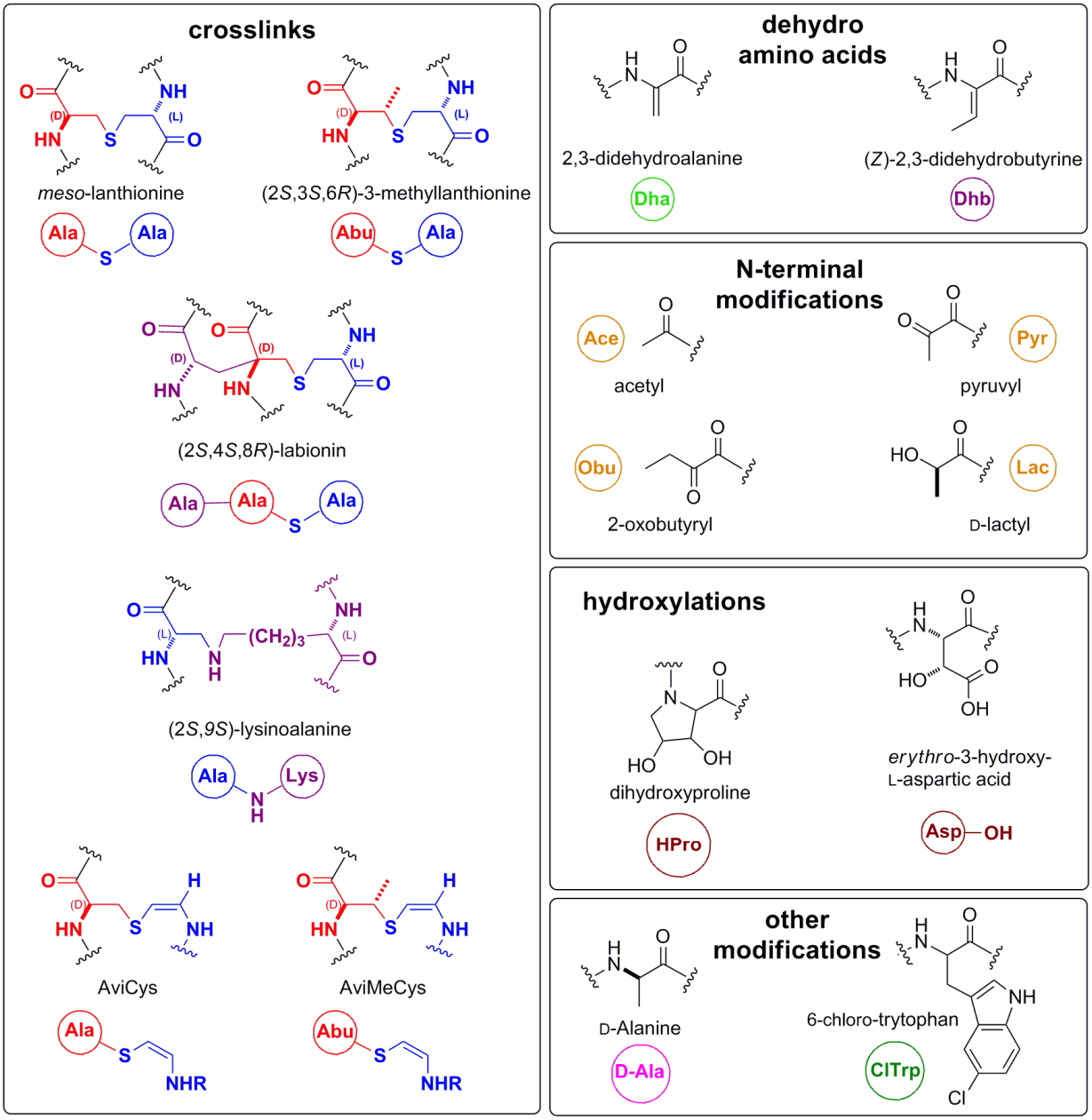


Figure 1.1. Posttranslational modifications in lantibiotics.

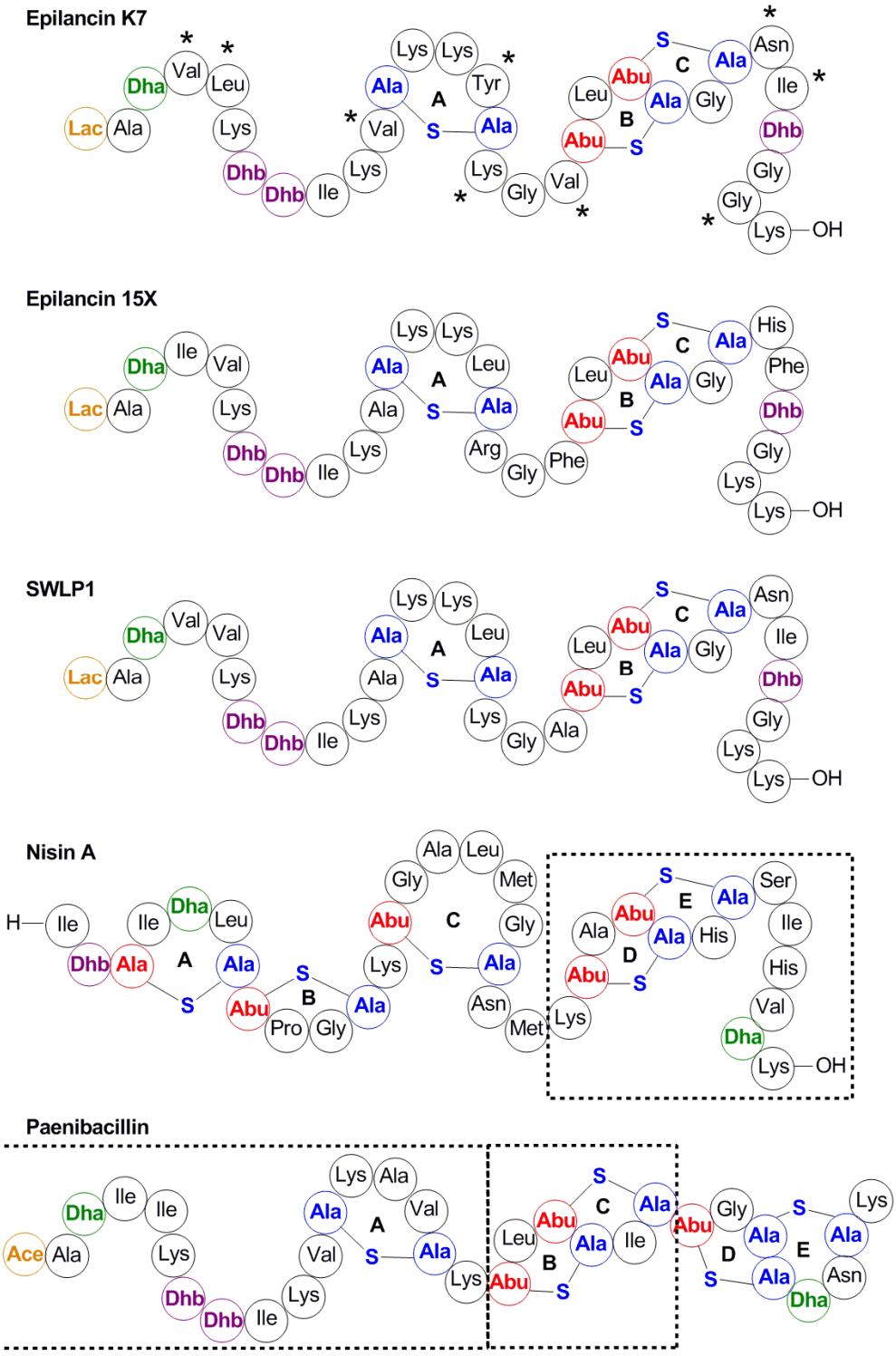


Figure 1.2. Structures of the epilancin-group of lantibiotics (epilancin K7, epilancin 15X, and SWLP1) using the shorthand notation from Figure 1.1. The non-conserved residues are highlighted with an asterisk in the structure of epilancin K7. The structures of nisin A and paenibacillin with considerably homology with the C- and N-terminus of the epilancin-group of peptides, respectively, are also shown.

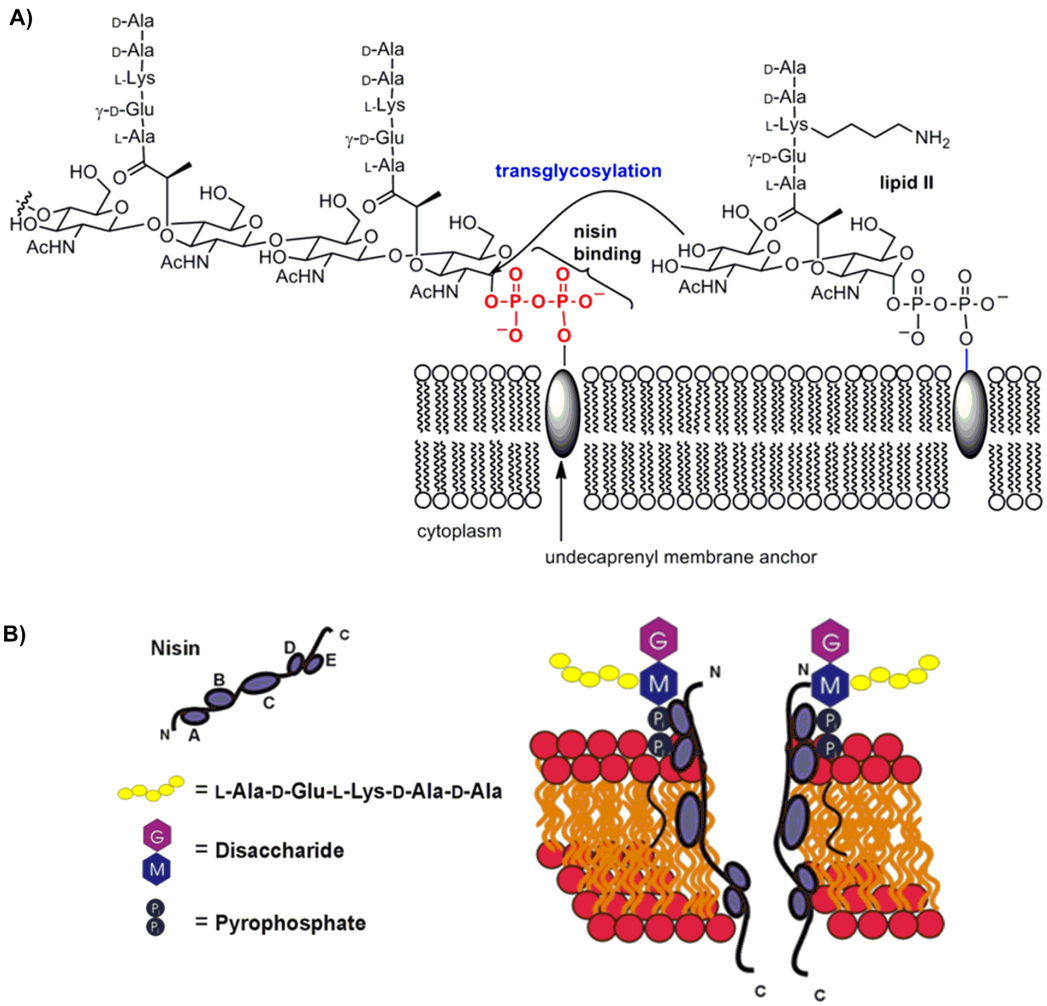


Figure 1.3. A dual mode of action of the lantibiotic nisin A. By using the pyrophosphate moiety of lipid II as a docking molecule, the N-terminal rings of nisin inhibit polymerization of murein subunits in bacterial cell wall biosynthesis (A), while the C-terminal portion forms pores in the plasma membrane (B) leading to cell death. Reproduced in part with permission from (*5*).

Lantibiotics have been known for decades and many of them are active against pathogenic bacteria at nanomolar levels (*11-13*). For instance, as early as 1952, nisin A (Figure 1.2) was shown to be as effective as penicillin in the treatment of mice infected with *Mycobacterium tuberculosis*, *Streptococcus pyogenes*, and *S. aureus* (*14*). Moreover, despite the short half-life and rapid clearance of nisin A from the blood, the peptide was shown to be more effective than vancomycin in an *S. pneumoniae* mouse

infection model (*15*). However, lantibiotics have not been extensively tested for their potential utility in the treatment of infectious diseases and no lantibiotics are currently approved in this area. Previous attempts have failed mainly due to the difficulty of obtaining these molecules in sufficient, cost effective amounts to enable their testing and commercialization (*16*). Standard fermentations methods, such as those used to make a variety of other antibiotics, have been rarely successful (*16*), with only a few cases, such as for the staphylococcal lantibiotics epidermin and gallidermin (*17-20*), being produced in significant amounts (more than 400 mg/L). Furthermore, in cases in which large amounts of lantibiotics are produced, such as for nisin A, the peptides have limitations due to chemical/biological stability under physiological conditions or the purity standards required for clinical testing have not been achieved (*21*). Thus, current challenges in the field include the identification of lantibiotics with increased stability that are effective against pathogenic bacteria and the development of alternative approaches for their production in a cost effective manner. The *in vitro* reconstitution of the biosynthetic machinery using purified *Escherichia coli* recombinant expression products, as demonstrated for lacticin 481 (*22, 23*), or the *in vivo* heterologous expression in a manageable host, as recently shown for three different peptides (*24*), have opened a door for more efficient production and purification strategies of lantibiotics. Furthermore, the recent discovery of several lantibiotics, such as the epilancin-group of peptides (Figure 1.2), with remarkable antimicrobial activity against pathogenic bacteria, including MRSA and vancomycin-resistant *Enterococci* (VRE), may aid in the development of novel peptide-based drug therapies.

1.2. THE EPILANCIN-GROUP OF ANTIMICROBIAL COMPOUNDS

At present more than 60 different lantibiotics have been structurally characterized and they have been classified in two groups based on the topology of their rings (*5, 13*). The type A lantibiotics, such as the well studied nisin A and Pep5, have a propensity to form elongated amphipathic screw-shaped membrane depolarizing structures and have a net positive charge, whereas type B lantibiotics, as in the case of cinnamycin or mersacidin, are more globular and compact in structure and are less charged (*5*). Within the type A lantibiotics, the epilancin-group is characterized by an N-terminal linear tail composed of 11 residues including several dehydrated amino acids and capped by the unusual N-terminal 2-hydroxypropionyl group (lactate, Lac) (Figure 1.2). The C-terminus of the polypeptides contains two intertwined MeLan rings, with a spacing of two residues between the bridging residues. Interestingly, two of the dehydrated amino acids in the N-terminus are neighbors, which may impose certain conformational restrictions (*25*).

Epilancin K7 was the first member of this group of peptides to be characterized (*25, 26*). It was isolated from *S. epidermidis* K7 and it is a large (31 residues, 3032 Da), elongated, and basic peptide due to the presence of six positively charged amino acids. The structure was determined by extensive two-dimensional and three-dimensional NMR spectroscopy and contains one lanthionine (Lan) and two 3-methyllanthionine (MeLan) bridges, one 2,3-didehydroalanine (Dha), three (*Z*)-2,3-didehydrobutyrine (Dhb), and an unprecedented N-terminal Lac. Epilancin K7 adopts an extended, flexible structure in water, although the rings likely have defined conformations (*25*). However, after interaction with membranes, or under conditions that mimic such situations, the

peptide may adopt a more ordered structure than in water, as evidenced by NMR spectroscopy experiments with dodecylphosphocholine micelles (*25*).

The lantibiotic epilancin 15X was isolated from *S. epidermidis* 15X154 and highly resembles epilancin K7 with 68% sequence identity (*27*). It also contains 31 residues (3175 Da), including ten amino acids that are posttranslationally modified and seven positively charged residues. The primary and tertiary structures were determined by mass spectrometry and NMR spectroscopy confirming the presence of the unusual N-terminal Lac group (Figure 1.2). Finally, the lantibiotic SWLP1 was isolated from *S. warneri* DSM16081, also known to produce the class II lantibiotic nukacin ISK1 (*28*), and its structure was determined by mass spectrometry (2999 Da) (*29*). As epilancin K7 and 15X, SWLP1 also contains an N-terminal Lac (Figure 1.2).

The epilancin-like lantibiotics are active against several strains of pathogenic bacteria, including MRSA, VRE, and *S. epidermidis*. Indeed, epilancin 15X has minimal inhibitory concentration (MIC) values against MRSA and VRE among the lowest reported for lantibiotics (*27, 30*) and epilancin K7 is approximately tenfold more active against *S. simulans* 22 than the staphylococcal lantibiotics Pep5, epidermin, and gallidermin (*25*). In addition, this group of peptides may possess considerable stability against acid/alkali treatment as evidenced for epilancin 15X in the current work. Moreover, the multiple positively charged residues make these peptides quite soluble in water. In contrast, the solubility and efficacy of nisin A are highly pH dependent and it loses activity sharply in the neutral pH region limiting its application to acidic environments (*31, 32*).

Currently, the mode of action of the epilancin-group of lantibiotics is unknown. The bactericidal activity of the cationic type-A lantibiotics, such as nisin A, subtilin, epidermin, and Pep5, is thought to be a consequence of their ability to induce pore formation in cell membranes. Indeed, the C-terminal B and C rings of the epilancin-like peptides are structurally similar to the D and E rings of nisin A (Figure 1.2) that are believed to be involved in pore formation (*10*). Furthermore, experiments with phosphatidylcholine liposomes suggested that epilancin K7 can penetrate into the hydrophobic carbon region of the lipid bilayer permeabilizing the membranes of susceptible cells (*33*). Interestingly, members of the epilancin family of peptides, with six to seven charged residues at neutral pH, together with Pep5, are the most positively charged lantibiotics reported so far (*34*). The electrostatic interactions between the peptides and the negatively charged bacterial membrane are likely important for the initial binding. Indeed, experiments with nisin A suggested that the sensitivity of the target bacteria is dependent on the charged state of the cell wall and membrane (*35, 36*). Moreover, all the conserved positively charged residues in the epilancin-group of peptides (Lys6, Lys10, Lys13, Lys14, Arg17, and Lys31 in epilancin 15X) are expected to be located on one side of the molecule and the hydrophobic residues on the other side, as suggested by helical wheel projections of the unmodified core peptides. The hydrophobic side may face the lumen of the pore while the hydrophilic part could interact with the membrane leading to the leakage of charged ions, amino acids, small metabolites, and ATP with a simultaneous drop in membrane potential as suggested for Pep5 (*37, 38*). Thus, the efficacy of the epilancin-like peptides may be correlated with their cationic charge.

However, formation of pores by itself would not explain the low MIC values of epilancin 15X, epilancin K7, and SWLP1. While a number of cationic peptides, such as those produced by mammals or insects, have antimicrobial activities in the micromolar (μM) range, some lantibiotics, including nisin A and the epilancin-like peptides, have been shown to be active at nanomolar (nM) concentrations. In the case of nisin A, the potent antibacterial activity is a consequence of docking onto lipid II by the A and B rings (*39*), which greatly enhances the pore forming ability (*8*). However, the epilancin-family of peptides does not contain the A and B-rings found in nisin, and lipid II does not serve as a docking molecule for epilancin K7 (*7*). Thus, the N-terminal portion of this group of peptides and the unusual lactate group may be involved in a currently unknown alternative mechanism. Indeed, additional posttranslational modifications beyond the Lan or MeLan rings are often important for biological activity of lantibiotics. In the case of cinnamycin, a β-hydroxylated aspartate residue and a lysinoalanine ring are important for recognition of its target (*40, 41*). In another example, the lantibiotic microbisporicin, which has a very similar ring topology as epidermin, is two orders of magnitude more potent than the latter compound against several strains of *S. aureus*, presumably because of the posttranslationally modified hydroxylated proline and chlorinated tryptophan residues that are absent in epidermin (*42*). Thus, the N-terminal lactate group in the epilancin family of peptides might also be important for the antimicrobial activity. Interestingly, the lantibiotic paenibacillin produced by the food-borne bacterial strain *Paenibacillus polymyxa* OSY-DF that contains an N-terminal acetyl group, instead of a Lac group, (*43, 44*) shares considerable N-terminal sequence similarity with the epilancin-group of peptides, suggesting convergent evolution for

interaction with a similar bacterial target (Figure 1.2). Contrary to nisin, paenibacillin has shown considerable thermal and acid/alkali stability even after long incubation periods (*43*). Despite the potential of the epilancin-group of lantibiotics for treatment of infectious diseases, their biosynthetic routes and the role of the Lac group in bioactivity have not been determined. The elucidation of their biosynthesis may support the development of more efficient production strategies and facilitate the introduction of modified amino acids to generate compounds that are more potent or target specific bacteria.

1.3. THE BIOSYNTHESIS OF LANTIPEPTIDES

Lantipeptides are defined by their characteristic lanthionine and/or methyllanthionine thioether crosslinks (*45*); lantipeptides with confirmed antimicrobial activity are called lantibiotics. A classification scheme for this family of posttranslational modified peptides has been previously proposed based on the enzymes responsible for ring formation (Figure 1.4) and the biosynthetic genes have been designated with the generic locus symbol *lan* (or Lan for the encoded proteins), with a more specific name for each lantibiotic (such as *nis* in nisin A biosynthesis) (*4*). In all cases, lantipeptides are synthesized as precursor peptides (LanA) that contain an N-terminal leader region that is not modified and a C-terminal core peptide encoding the backbone of the final product. The core peptides in class I lantipeptides are modified by two distinct enzymes designated LanB and LanC that catalyze the dehydration of serine and threonine residues and the subsequent intramolecular addition of cysteine thiols onto the dehydrated amino acids, respectively, resulting in the formation of the thioether crosslinks (Figure 1.4). For class II, III, and IV precursor peptides, bifunctional enzymes

designated LanM, RamC-like (or LabKC in labyrinthopeptin biosynthesis, a group of peptides with additional carbacyclic side-chain linkages) (*46-48*), and LanL (*45*), respectively, are responsible for both dehydration and cyclization steps of the core peptide (Figure 1.4) (*4, 45, 47*). All the lanthionine forming enzymes (LanC, LanM, RamC-like/LabKC, and LanL) share significant homology at the C-terminus and are members of a protein superfamily designated LanC-like, while the LanB dehydratases bear no homology with other known proteins. After ring formation, and in some cases further modifications, the lantipeptides are secreted and the leader peptides are removed by the protease domain of transporters (LanT) or dedicated proteases (LanP), generating the biologically active compound. The biosynthetic routes to several lantipeptides have been characterized with nisin A and lacticin 481 the most studied class I and class II members, respectively (Figure 1.5). These pathways possess several advantages for bioengineering, including the direct link between precursor peptide gene and natural product structure and the relatively short biosynthetic pathways. Furthermore, many of the biosynthetic enzymes have demonstrated substrate promiscuity. In the following sections, the different families of enzymes are described in more detailed.

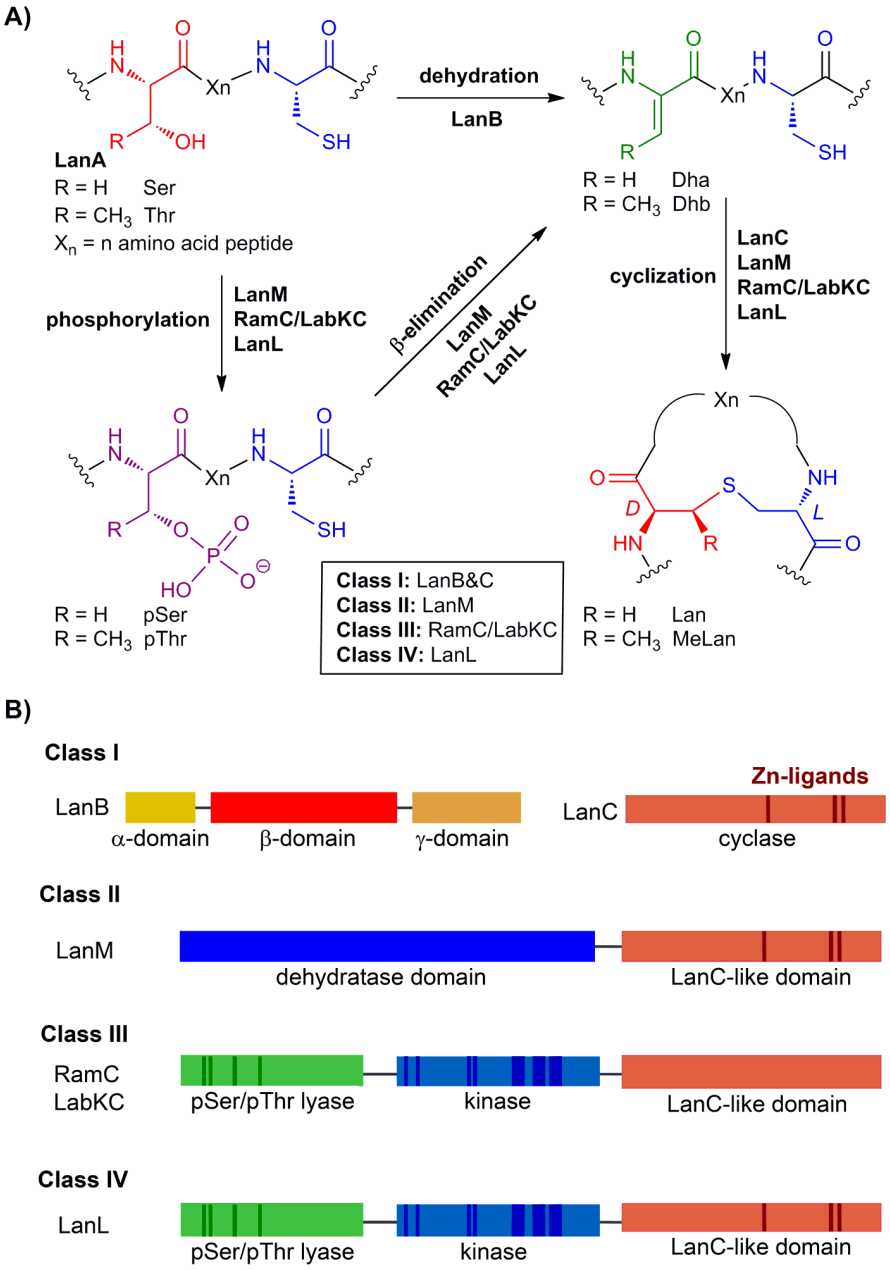


Figure 1.4. Biosynthesis of lantipeptides by four classes of lanthionine-forming enzymes. (A) After ribosomal synthesis of the precursor peptides (LanA), different enzymes denotated LanB, LanM, RamC-like (LabKC), or LanL can catalyze dehydration of the Ser or Thr residues to generate the dehydrated amino acids Dha or Dhb, respectively. For class II, III, and IV lantipeptides, this modification proceeds through a two-step process, phosphorylation of Ser/Thr and subsequent β-elimination of the phosphate to generate Dha/Dhb. Following dehydration, LanC, LanM, LanL, or RamC-like/LabKC enzymes can catalyze intramolecular addition of Cys thiols onto the dehydro amino acids generating the thioether-containing residues Lan and MeLan. The conserved zinc-ligands are missing in RamC-like/LabKC proteins and the mechanism of cyclization is currently unknown. (B) Schematic representation of the domain architecture of class I−IV (methyl)lanthionine-forming enzymes. The positions of the expected active site residues are shown in darker colors (*49*).

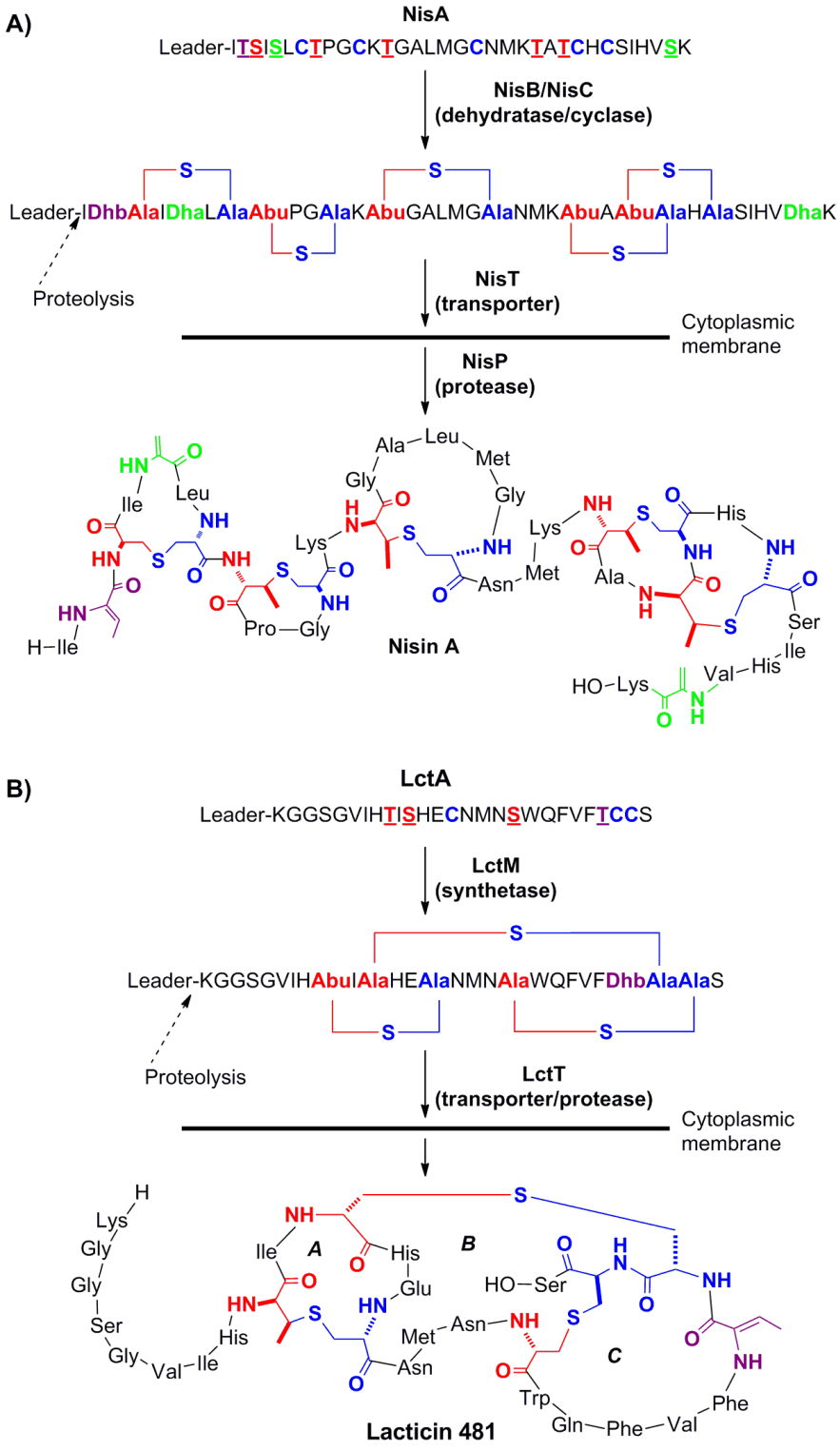


Figure 1.5. Biosynthetic route to the lantibiotics nisin A (A) and lacticin 481 (B).

Dehydration

The dehydration of the Ser and Thr residues in the core peptide leading to the formation of Dha or Dhb is the first key reaction in lantipeptide biosynthesis and can be performed by different enzymes. The LanB proteins from class I lantipeptides possess about 1000 residues and seem to be associated with the membrane, possibly as part of a biosynthetic membrane-associated complex, which may explain the failure of previous attempts for the reconstitution of their enzymatic activity *in vitro*. LanB proteins do not show any significant similarity to the N-terminal domains of the lanthionine synthetases (LanM, RamC-like/LabKC, and LanL) that catalyze the dehydration reaction in other classes of lantipeptides and therefore both groups of proteins are unlikely to be evolutionary related. Dehydration by LanB seems to be an organized directional process that starts at the N-terminus of the core peptide and continues towards the C-terminus, as suggested for NisB (*50*). Further details regarding the reaction mechanism, cosubstrates, and the structure of the LanB proteins remain unknown. These dehydratases contain nine short segments with significant homology that may have functional significance (*51*). However, prediction of the residues involved in catalysis or other functions remains challenging because more than 40 residues are highly conserved in this family of proteins, with only one residue (Pro612 in NisB) being fully conserved based on the alignment of 66 putative LanBs (*51*). The mechanism for substrate recognition is also unknown, but the N-terminal part of the leader peptide, including a partially conserved FDLD motif, was shown to be important for the interaction of NisA and NisB (*52-54*). Furthermore, NisB possesses low substrate

specificity, as demonstrated by the *in vivo* dehydration of nonlantibiotic peptides fused to the NisA leader peptide (*55*).

Three independent domains have been identified in LanB proteins based on sequence alignments: an N-terminal Lant_dehyd_N domain (PF04737), a central Lant_dehyd_C domain (PF04738), and a C-terminal thiopep_ocin domain (TIGR03891) that will be designated as α, β, and γ domains for simplicity (Figure 1.4). Interestingly, genes encoding for LanB dehydratase homologs are also present in the biosynthetic gene clusters of thiazolyl peptides. In this case, the dehydratase is split in two proteins, one consisting of the α and β domains (TsrC for thiostrepton using the nomenclature in (*56*)) and the other possessing only the γ domain (TsrD) (*56-59*). A split dehydratase (GodF and GodG) is also encoded by the goadsporin biosynthetic gene cluster (*60*). Although dehydrated residues derived from Ser and Thr are also found in numerous NRPS-derived products, such as syringomycin (*61*), syringotoxin B (*62*), or microcystin (*63-65*), the enzymatic mechanism of dehydration remains elusive. However, single or multiple putative dehydratases containing the α and β domains, but not the γ domain, are present in several gene clusters encoding for putative non-ribosomal peptides identified by genome mining (Figure 1.6), suggesting that this family of proteins may also be involved in the dehydration of Ser/Thr residues of NRPS-derived metabolites. Importantly, the α and/or β dehydratase domains can be fused to adenylation (A) or thioesterase (TE) domains, as evidenced for gene clusters from *Micromonospora aurantiaca* and *Salinispora arenicola* (Figure 1.6). This NRPS-domain architecture has not been reported in the literature to the best of my knowledge. The conservation of the γ domain in biosynthetic proteins from different families of ribosomal peptide

(lantipeptides, thiazolyl peptides, and goadsporin) and the absence of this domain in NRPS-related biosynthetic proteins suggests that the γ domain may function as a docking protein that recognizes the leader peptide allowing dehydration by the α and β domains. A similar role was proposed for McbD, SagD, and the C-terminal domain of PatD in the biosynthesis of the thiazole/oxazole-containing ribosomal peptides microcin B17, streptolysin S, and patellamides, respectively (*66-68*).

The N-terminal catalytic domain of LanM proteins is responsible for dehydration in class II lantipeptide biosynthesis as demonstrated by the *in vitro* reconstitution of the enzymatic activity of LctM (the synthetase involved in lacticin 481 biosynthesis) and several mutants (*22, 69*). LctM dehydration involves initial phosphorylation of the hydroxyl-bearing amino acids in the presence of ATP and Mg^{2+}, followed by phosphate elimination (Figure 1.4) (*70*). In addition, LctM is distributive with respect to the dehydration reaction, releasing the intermediates after each dehydration (*71*). Similarly, NisB and NisC are alternating enzymes (*50*). Sequence alignment of the N-terminal domain of LctM and site-directed mutagenesis experiments on the conserved N-terminal residues of LctM provided some insights into the reaction mechanism of dehydration (Figure 1.7). A general base, such as Asp242 or Asp259, has been proposed to deprotonate the Ser/Thr residues targeted for dehydration generating a transient alkoxide group that could attack the γ-phosphate of ATP producing the phosphorylated peptide and ADP. Then, another base, possibly Arg299 or Asp259, could deprotonate the α-carbon of pSer/pThr to generate an enolate, which subsequently eliminates the phosphate group, possibly assisted by a general acid, such as Lys159, and/or Mg^{2+} (*69*). Additional mutagenesis studies on the precursor peptide

suggested that the secondary structure of the leader peptide is important for peptide processing (*72*).

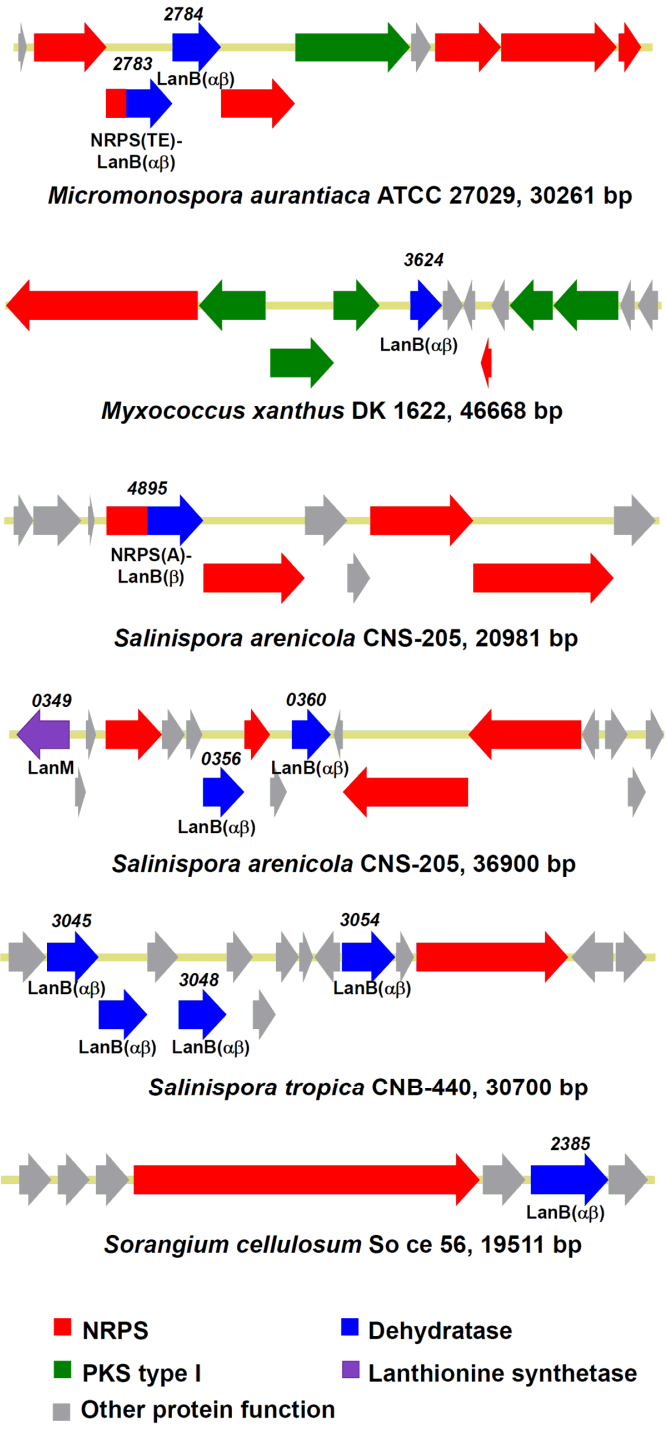


Figure 1.6. Representative examples of non-ribosomal peptide biosynthetic gene clusters identified by genome mining for LanB proteins.

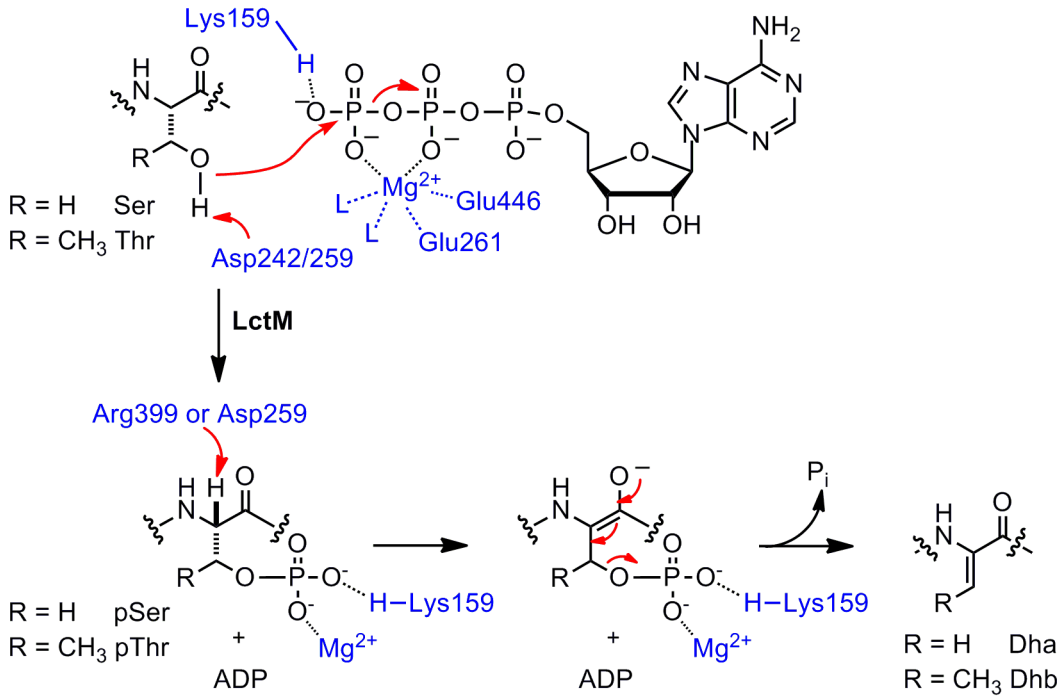


Figure 1.7. Proposed general mechanism of dehydration of Ser/Thr by LctM (*69*).

LanL-like lanthionine synthetases contain a central Ser/Thr kinase domain and an N-terminal pSer/pThr lyase domain that catalyze phosphorylation and elimination of phosphorylated hydroxyl-containing amino acids, respectively, as demonstrated for VenL (*45, 49*). These two domains have no significant sequence homology with either the LanB or the LanM proteins. Based on mutagenesis studies on VenL lacking the C-terminal cyclase domain, a reaction mechanism for the elimination reaction was proposed (Figure 1.8) (*49*). Lys80 could deprotonate the α-proton of pSer/pThr residues to initiate the reaction, while His53 may protonate the phosphate leaving group, leading to elimination and formation of the dehydrated amino acids (Figure 1.8) (*49*). Finally, the N-terminal domains of LabKC (a class III biosynthetic protein), which share significant homology to the corresponding domains of LanL proteins, were proposed to catalyze

the dehydration of Ser residues in the core peptide by phosphorylation/elimination using GTP as cosubstrate rather than ATP (*48*). Furthermore, a conserved N-terminal hydrophobic patch in the precursor peptide and predicted to possess a helical structure was recently shown to be important for proper peptide processing (*73*).

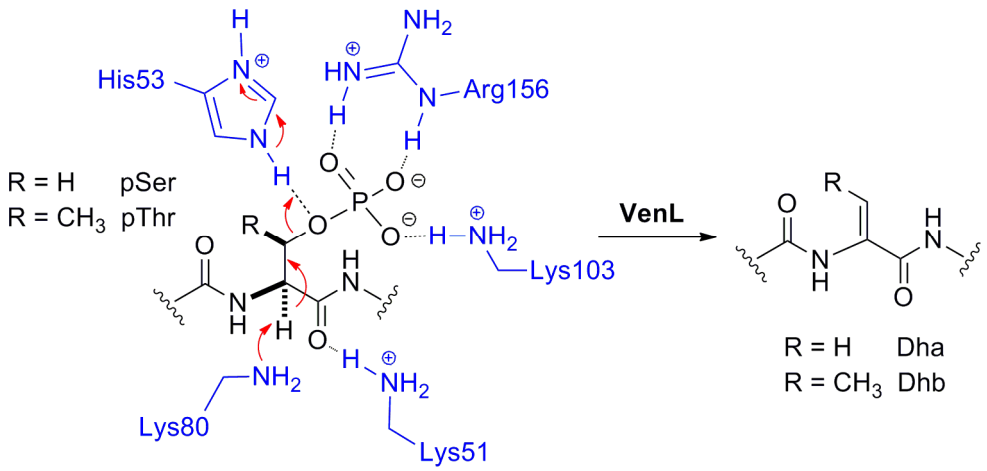


Figure 1.8. Proposed mechanism for β-elimination of the phosphate of pSer/Thr catalyzed by the lyase domain of VenL (*49*).

Cyclization and the LanC-like superfamily

In lantipeptide biosynthesis, the formation of the thioether bonds is catalyzed by enzymes from the LanC-like domain superfamily (PF05147), a highly divergent group of peptide-modifying enzymes in bacterial and eukaryotic cells (Figure 1.9). Although spontaneous nucleophilic addition of thiol groups into dehydrated amino acids is well known and relatively fast, enzymatic control over the chemoselectivity and/or processivity is required for formation of the correct ring topologies. The LanC cyclases are single domain proteins of about 400 residues long, whereas LanM, LanL, and RamC-like modification enzymes are larger proteins (900-1000 aa) that bear the C-terminal cyclization domain (LanC-like) as well as the N-terminal domains involved in dehydrations as discussed above (Figure 1.4). The LanC-like domain is characterized

by seven conserved motifs that are present in all the lanthionine-forming enzymes. In addition, Cys and His residues (Cys284, Cys330, and His331 in NisC or Cys781, Cys836, and His837 in LctM) that function as zinc ligands and that are important for catalysis (Figure 1.10) (*74-76*) are conserved in the LanC, LanM, and LanL proteins, although the His residue is sometimes replaced by a Cys (Figure 1.9, diamonds). In contrast, these metal ligands are absent from the RamC-like/LabKC proteins and the mechanism of cyclization may be different (*46*). In addition to the metal ligands, the coordination sphere is completed by a water molecule believed to be displaced by the Cys residues in the substrate (Figure 1.10).

The enzymatic activity of several members of the LanC-like superfamily, including proteins from all the lantipeptide classes, has been reconstituted *in vitro* and the catalytic mechanism has been investigated (*22, 45, 48, 74-77*). The thiol groups in the lantipeptide core peptides are likely activated by the metal that lowers their pKa allowing the nucleophilic addition onto the β-carbon of the dehydro amino acids. The enolate intermediate thus generated could be protonated by an active site base (Figure 1.10). A conserved His (His212 in NisC or His725 in LctM) was suggested to be the base that deprotonates the thiol of the substrate or the acid that protonates the enolate intermediate (*75*). The mechanism controlling the regioselectivity of cyclization remains unclear. However, the crystal structure of NisC evidenced that the active site consists of a shallow groove that accommodates the formation of rings of different sizes and that the zinc is located on top of an α,α-barrel that resembles the β-subunit of farnesyl transferase (*74*). Interestingly, the directionality of ring formation is not always uniform. The LanC cyclases from class I lantipeptides incorporate lanthionine rings solely in the

N-to-C terminal direction, whereas for some LanM enzymes some of the lanthionine rings in the final structure (such as in mersacidin (*78*) and cinnamycin (*79*)) result from cyclization in the opposite direction (*5*). Like the dehydratases, LanC-like proteins, including NisC, LctM, and ProcM (in prochlorosins biosynthesis (*80*)), have low substrate specificity, modifying nonlantibiotic peptides and cysteine derivatives (*55, 81*).

Proteolysis and transport

After ring formation, and frequently additional posttranslational modifications, the leader peptide is removed and the peptide is secreted. However, the order of these steps is not conserved. In some cases, the mature bioactive compound is generated inside the cytoplasm followed by secretion, as in Pep5 or lactocin S biosynthesis (*82, 83*). In other cases, leader removal and peptide secretion are coupled and are performed by membrane proteins, such as in lacticin 481 or nukacin ISK-1 biosynthesis (*23, 84*). Finally, the leader peptide can be removed after secretion as in nisin A or epidermin biosynthesis (*85, 86*).

ATP-binding cassette (ABC) transporters designated as LanT are commonly encoded in lantipeptide gene clusters and are involved in the secretion of fully modified precursor peptides or mature lantipeptides, as demonstrated for NisT, CylT, and LasT (in nisin A, cytolysin, and lactocin S biosynthesis, respectively (*82, 87, 88*)). These proteins contain an intracellular domain with a highly conserved ATP binding motif, indicating that ATP hydrolysis is utilized as a source of energy for secretion. Furthermore, LanTs contain a membrane-spanning domain that usually possesses six transmembrane helices.

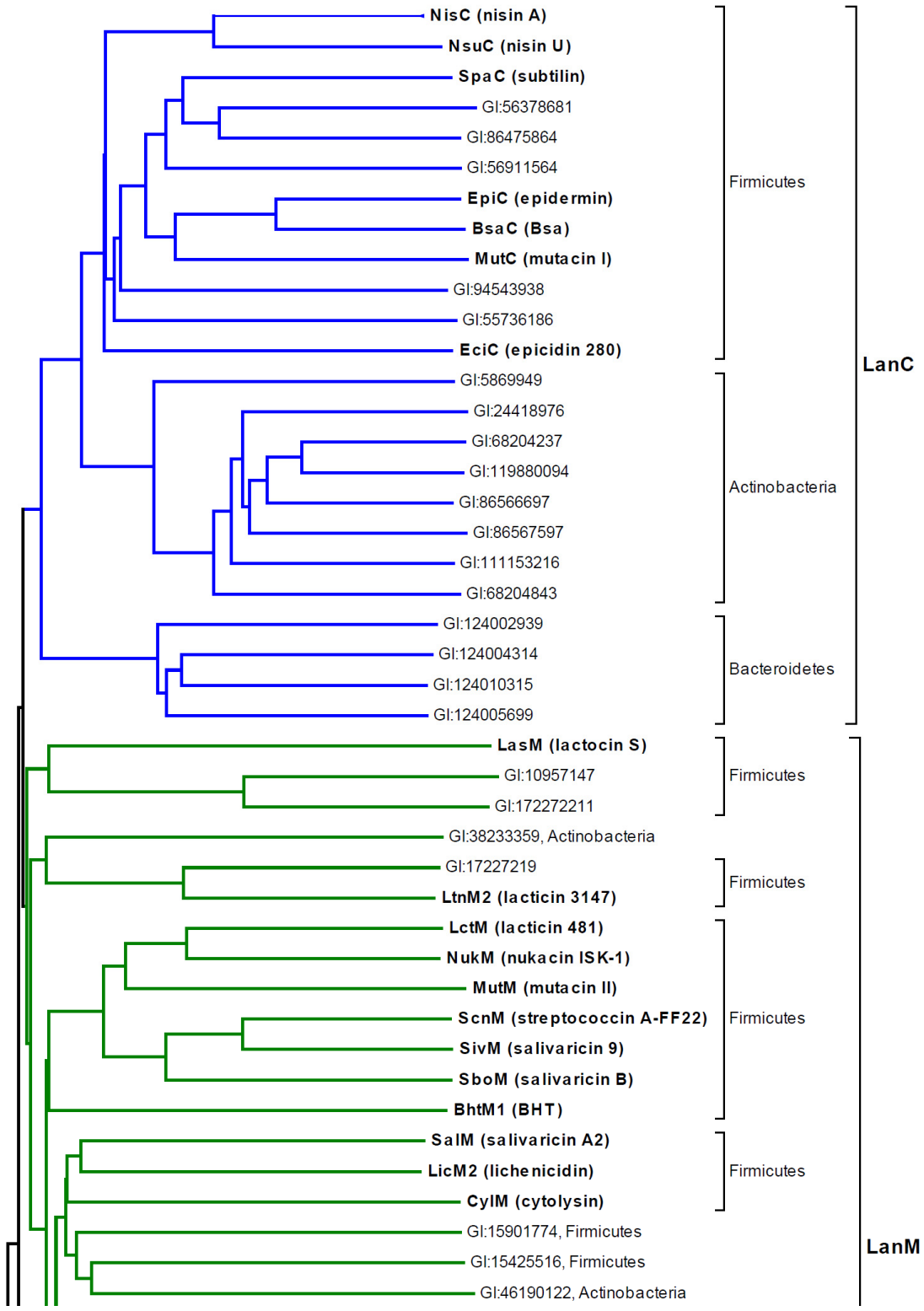


Figure 1.9. Cladograph of the LanC-like domains from different lanthionine-forming enzymes.

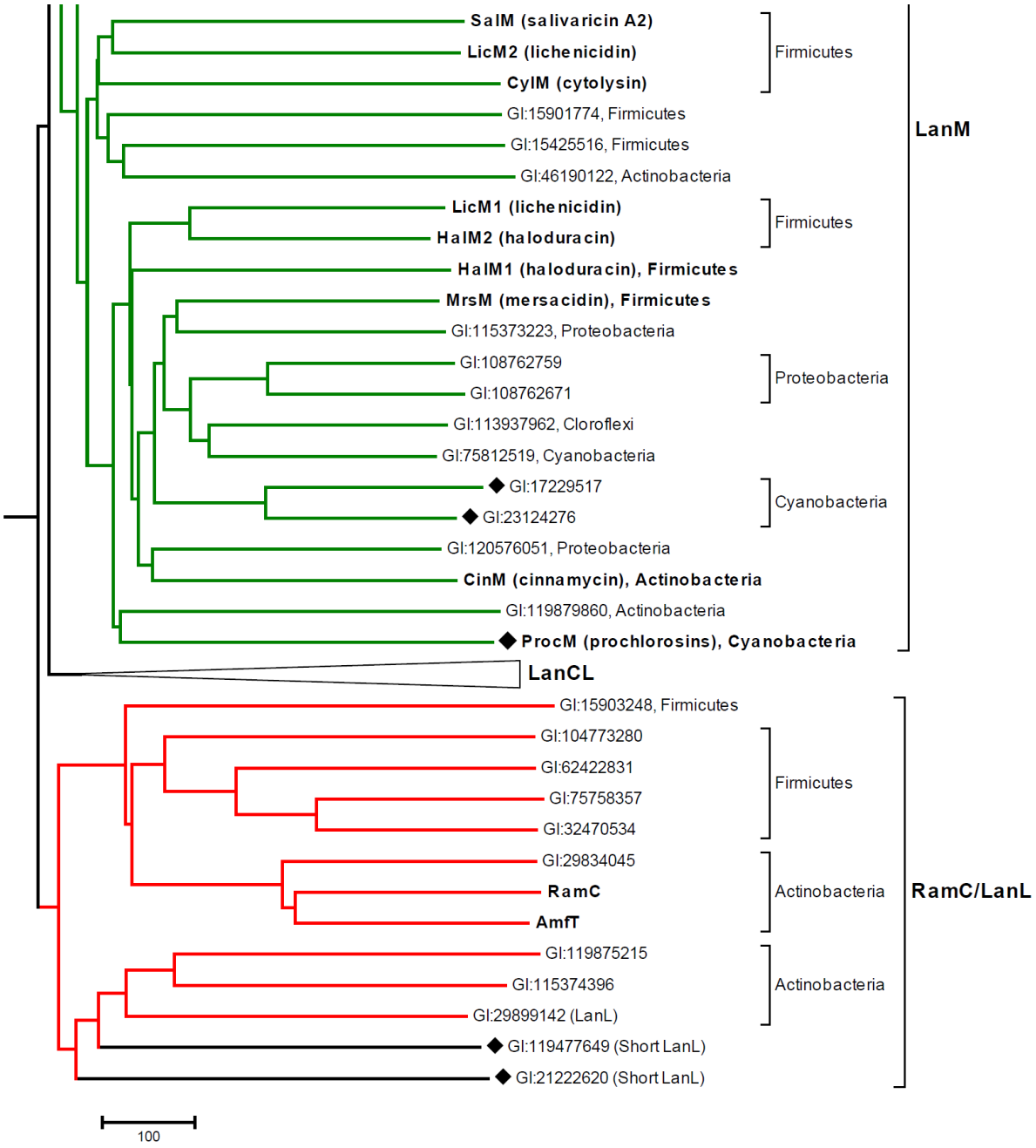


Figure 1.9. *Continuation*. Cladograph of the LanC-like domains from different lanthionine-forming enzymes. Four distinctive clades corresponding to different classes of LanC-like families are shown. LanC (blue), LanM (green), and the group of RamC & LanL (red) proteins are involved in lantipeptide biosynthesis in bacteria. The clade designated as LanCL is mostly composed of eukaryotic lanthionine synthetases and is shown compressed. Branches with diamonds correspond to proteins in which the conserved His residue (His331 in NisC) is replaced by Cys. The evolutionary history was inferred using the Neighbor-Joining method (13) and considering only the C-terminal LanC-like domain from LanM, RamC-like, and LanL proteins. The tree was adapted from the National Center for Bioinformatic Information (NCBI) curated domain hierarchy for the LanC-like domain superfamily (*89*) and modified by using CDTree3.1 (*90*) and MEGA5 (*91*).

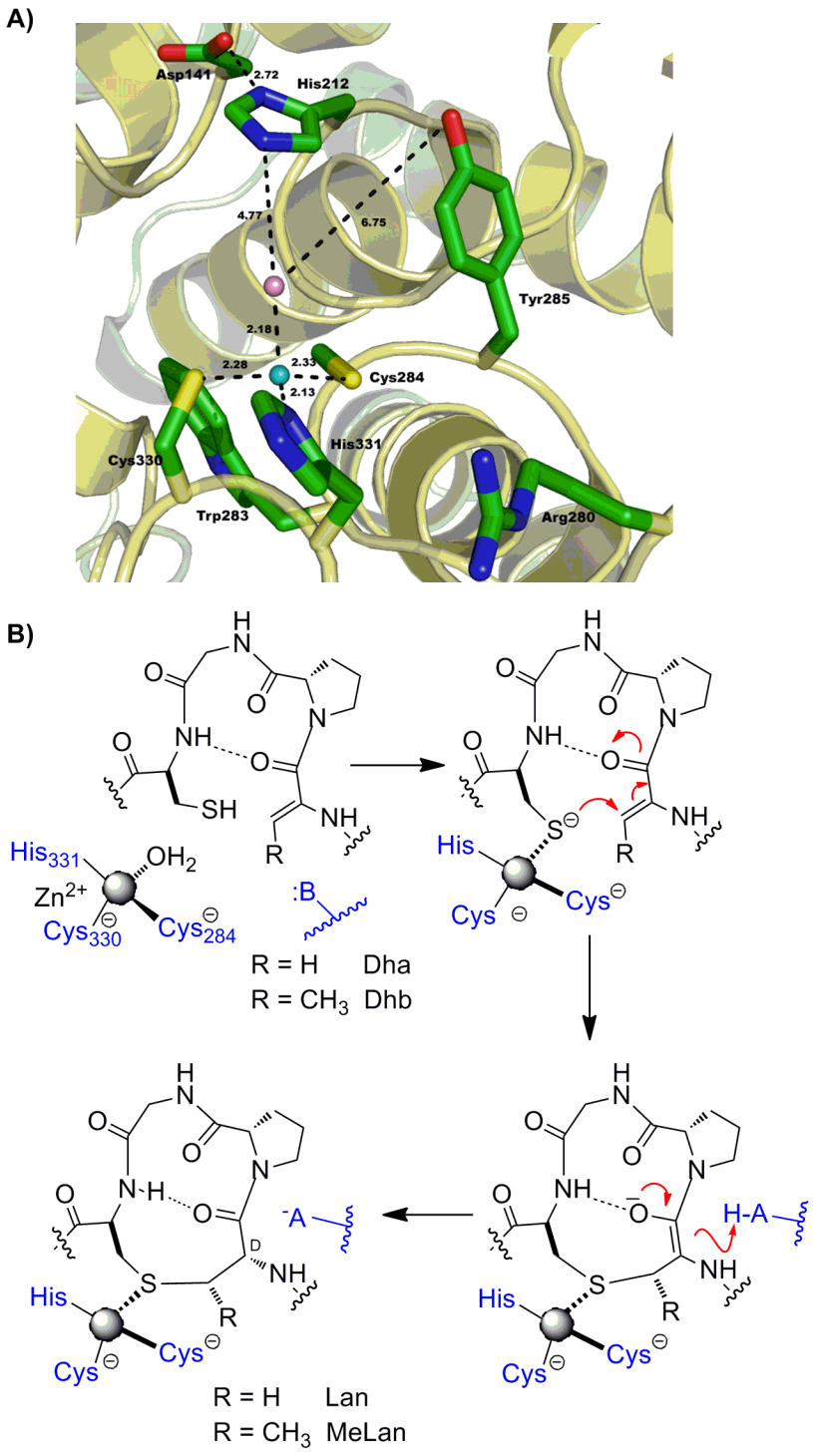


Figure 1.10. Active site of NisC (A) and proposed mechanism for the cyclization reaction (B). Reproduced in part with permission from (*75*).

One group of LanT proteins, frequently found in class II lantipeptide biosynthesis, contains an extra N-terminal, intracellular, cysteine-type protease domain of about 150 residues. A similar domain architecture is also found in several other ABC transporter, maturation, and secretion proteins (AMS) (*92*). The protease domain cleaves the leader peptide after a highly conserved double-glycine motif with consensus sequence (I/L/V)(S/T)X_2E(L/M)X_2(I/L/V)X**G**(**G**/A/S), where X can be any amino acid (*92*). The *in vitro* enzymatic activity, mechanism, and substrate specificity of the protease domain of LctT (*23*) and NukT (*84*) (in lacticin 481 and nukacin ISK-1 biosynthesis) have been investigated. The conserved Cys12 residue in LctT seems to be essential for catalysis, while His90 could play a role in the regiospecificity of amide bond hydrolysis (*23*). In addition, the GA motif C-terminal to the cleavage site, Glu−8, and the secondary structure of LctA are all important for substrate processing (*23*). In the case of NukT, the peptidase activity depends on ATP hydrolysis, suggesting that the N-terminal protease domain and the C-terminal ATP-binding domain may function cooperatively maybe through conformational changes (*84*).

For most class I lantipeptides, and also for some class II members, removal of the leader region is performed by a subtilisin-like protease commonly designated as LanP. These proteins vary greatly in size ranging from about 250 to 680 amino acids, due to the presence or absence of preprosequences, which direct the enzyme to the *sec*-dependent secretory pathway. In some cases, an additional C-terminal extension, consisting of a spacer of about 100 residues and a C-terminal transmembrane sequence with the cell wall anchor consensus motif (LPXTG), is also found (*85*). The most conserved regions of these serine proteases are around the active site and

include the catalytic triad residues (Asp, His, and Ser) as well as the oxyanion-hole residue (Asn). The substrate-scope of this family of proteases will be discussed in more detail in chapter 4.

1.4. GENOME MINING FOR LANTIPEPTIDES

The ever-increasing amount of DNA sequence data and the knowledge about the biosynthetic routes to lantipeptides, described above, have lead to the discovery of novel gene clusters encoding previously unknown peptides. Upon identification of the clusters, prediction of the chemical structures of the mature lantipeptides based on the genomic context of the precursor genes, in combination with several chemical and biological strategies, has aided bioassay-independent discovery and isolation of novel lantipeptides. In the first such example, a bioinformatic analysis of the genome of *Bacillus halodurans* C-125 revealed the presence of two genes encoding for precursor peptides clustered with two additional open reading frames (ORF) encoding for LanM-type synthetases. Analysis of cell-free supernatants by mass spectrometry (MS), *in vitro* reconstitution of the LanM enzymes, antimicrobial assays, and mutagenesis experiments, allowed the identification and structural characterization of the two-component lantibiotic haloduracin (Figure 1.11) (*77, 93, 94*).

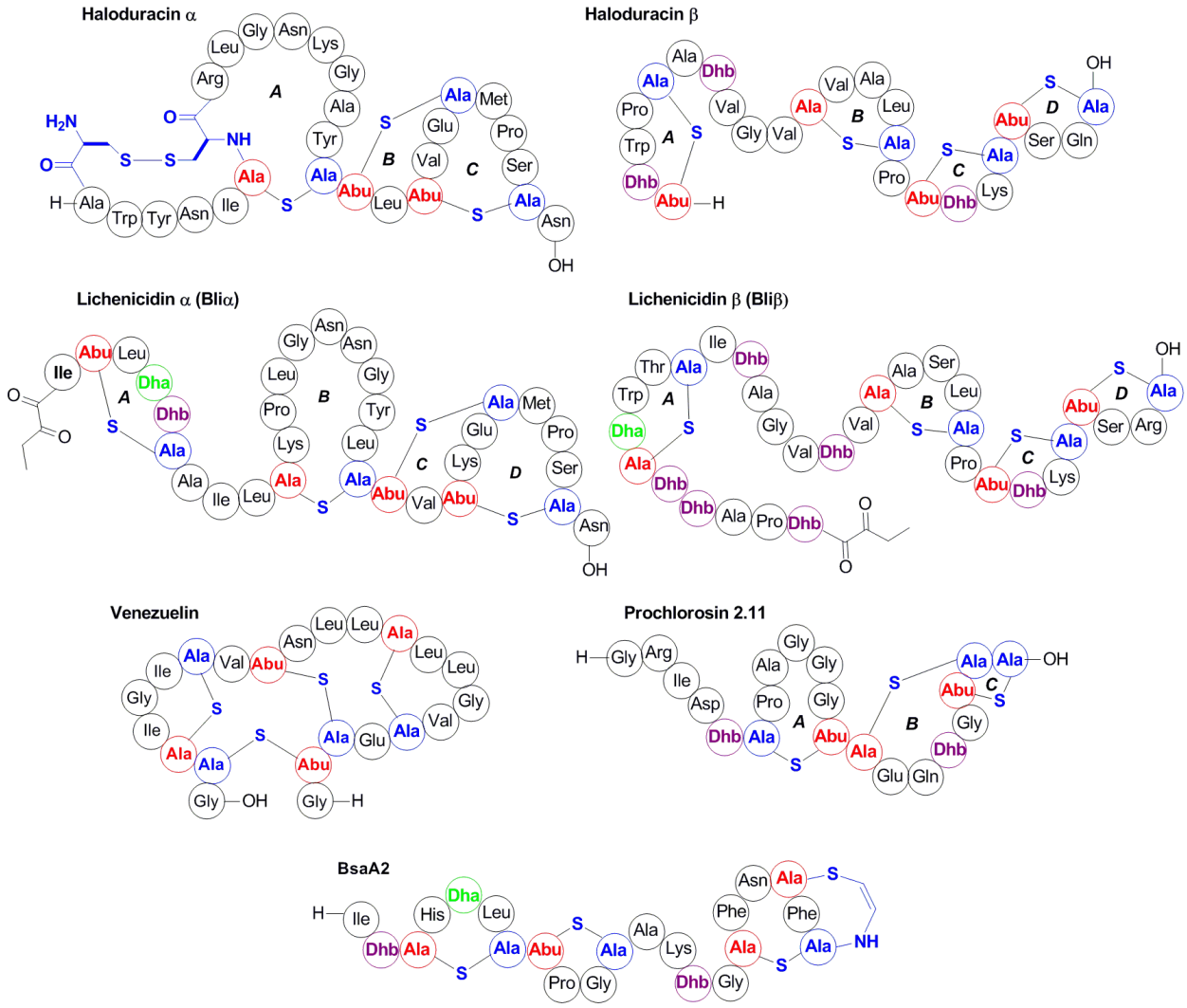


Figure 1.11. Lantipeptides recently discovered by genome mining using the notation described in Figure 1.1. The two-component lantibiotics haloduracin from *B. halodurans* C-125 and lichenicidin from *B. licheniformis* ATCC14580, DSM13, and VK21 were discovered after genome mining for LanM lanthionine synthetases. Venezuelin, a lantipeptide predicted from the genome sequence of *S. venezuelae* ATCC10712, was produced *in vitro* after reconstitution of a novel LanL lanthionine synthetase. Prochlorosin 2.11 is one of 29 lantipeptides produced by *Prochlorococcus* MIT9313. BsaA2 was discovered after genome sequence analysis of several *S. aureus* strains. Reproduced in part with permission from (*95*).

Using similar bioinformatic approaches or PCR amplification of conserved DNA sequences, several *lanM*-containing biosynthetic gene clusters were discovered subsequently, including the cluster encoding for lichenicidin in the genome of *Bacillus licheniformis* ATCC14580 (or DSM13) and VK21 (*77, 96-98*). In follow-up studies, this

two-component lantibiotic was detected in bacterial cultures by antimicrobial assays and MS analysis (Figure 1.11) (*96-98*). In a similar study, several variants of the lantibiotic epidermin, designated Bsa, and produced by MRSA strains were identified (Figure 1.11), suggesting that these bacteriocins may confer a competitive ecological advantage on community acquired infections (*99*).

A new route to lantipeptides was discovered by analysis of the draft genome sequence of *S. venezuelae* ATCC10712 revealing a lantibiotic-like gene cluster with an ORF encoding for a novel class of putative lanthionine synthetases (LanL), as described above (*45*). *In vitro* reconstitution of the enzyme activity with the predicted precursor peptide resulted in the production of venezuelin, the first class IV lantipeptide (Figure 1.11). LanL homologs were also identified in other species of Actinobacteria and Firmicutes (*45*). Another recent genome database analysis revealed that several strains of marine *Prochlorococcus* and *Synechococcus* contain multiple ORFs encoding a wide diversity of lantipeptide precursor peptides but only a single gene encoding a LanM-like synthetase (*80*). The precursor peptides have highly homologous leader peptides but display great diversity in the core peptide. The enzymatic activity of the predicted LanM from *Prochlorococcus* MIT9313 was reconstituted *in vitro* and 17 selected precursor peptides (out of the 29 encoded in the genome) were converted to the corresponding lantipeptides (prochlorosins), providing an example of natural combinatorial biosynthesis. Analysis of the spent media of *Prochlorococcus* MIT9313 by MS showed *in vivo* production of at least three of these compounds (Figure 1.11), confirming that lantipeptide production is not restricted to Gram-positive or soil bacteria as long believed (*80*). A search of sequences in the Global Oceanic Survey uncovered more

than 150 prochlorosin precursor genes from many different locations suggesting that prochlorosin production is widespread. Interestingly, several LanA peptides with long leader sequences (as in prochlorosins biosynthesis) were identified in bacterial genomes and classified in two families (NHLP and N11P) based on homology to regions of the alpha subnunit of nitrile hydratase and the cyanobacterial Nif11 nitrogen-fixing protein families (*100*). Importantly, some of the ORFs encoding for the precursor peptides are distant to the genomic regions encoding for LanM proteins and their identification would be challenging by using only BLAST searches for lanthionine synthetases. More recently, comparative analysis and BLAST searches of 58 cyanobacterial genomes, by using cyanobacterial LanM sequences as a query, led to the identification of sixteen gene clusters encoding for 61 lantipeptide precursor peptides, among other bacteriocin precursors, including several members of the NHLP and N11P families (*101*).

In two additional studies, PSI-BLAST searches of the NCBI database for LanB, LanC, and LanM homologs uncovered 49 class I and 61 class II lantipeptide gene clusters, including examples in phyla (Chloroflexi, Bacteroidetes, and Proteobacteria) not previously identified as lantipeptide producers (*51, 96*). Similarly, a computerized discovery strategy using the genome-mining software BAGEL2 revealed approximately 150 putative lantibiotic gene clusters based on conserved biosynthetic, transport, and immunity machinery (*102*). By using BAGEL, the biosynthetic gene cluster of pneumococcin A1 and A2 was identified in the genome sequence of *S. pneumoniae* R6 (*103*). Interestingly, when the predicted pneumococcin core peptides (belonging to class II lantipeptides) were modified *in vivo* using the nisin A synthetase machinery (a class I

system), lanthionine containing-peptides with antimicrobial activity were produced (*103*). Thus, genome sequencing has contributed to the discovery of novel lantibiotics and provided a better understanding of their biosynthetic and regulatory mechanisms, which may facilitate production and generation of more potent or stable variants.

1.5. BIOENGINEERING OF LANTIBIOTICS

An alternative approach to access novel/improved lantibiotics involves the derivatization of known peptides. Since lantibiotics are gene-encoded as precursor peptides, they possess a key advantage over classical antibiotics, as they are more amenable to engineering strategies and their antimicrobial activity and stability can be potentially enhanced. Mutant precursor peptides can be modified *in vitro* or *in vivo* using the lantibiotic-associated enzymes to generate new molecules. The simplest systems involve *in vivo* expression of mutant LanA precursor peptides in a variant of the original producer strain, either *in cis* by replacing the original LanA gene or *in trans* by complementing a copy of the mutant gene (*104*). Indeed, coupling these strategies with random and site-directed mutagenesis have allowed the generation of large libraries of bioengineered peptides and the identification of enhanced forms of the lantibiotics nisin A (*105*), mersacidin (*106*), and nukacin ISK-1 (*107*). For instance, the mutants nisin A M21V (renamed nisin V) or K22T were shown to display enhanced antimicrobial activity against medically significant pathogens, including MRSA, VRE, *Clostridium difficile*, *S. agalactiae*, and *Listeria monocytogenes (105, 108)*. However, this approach is limited to mutant peptides that are compatible with the native biosynthetic, transport, and immunity machineries.

An alternative strategy relies on the *in vitro* reconstitution of biosynthetic enzymes. This approach is not affected by reduced or aborted *in vivo* production of novel variants caused by incompatibility with the biosynthetic machinery or self-destruction of the host. Furthermore, the strategy can be easily coupled with the chemical incorporation of non-proteogenic amino acids to generate novel variants not easily accessible by using traditional *in vivo* expression strategies. Using this method, two lacticin 481 variants with increased antibacterial activity against a target strain were discovered, extending the chemical and structural diversity of lantibiotics (*109*). Thus, the recent advancements in the understanding of lantibiotics regarding biosynthetic machinery and mode of action can lead to the design of tailor-made peptides with enhanced activity and stability.

The following chapters focus on the biosynthesis of the antimicrobial compound epilancin 15X. In chapter 2, the molecular cloning and sequencing of the biosynthetic gene cluster are described. Furthermore, the reconstitution of the enzymatic activities of the dehydratase ElxB, the protease ElxP, and the dehydrogenase ElxO involved in formation of the N-terminal Lac group is demonstrated, and a biosynthetic route to epilancin 15X is proposed. Chapter 3 focuses on the biochemical and structural characterization of the dehydrogenase ElxO, displaying relaxed substrate specificity that can be used to modify or introduce other chemical functionalities into different peptides. Finally, chapter 4 describes the biochemical characterization and bioinformatic analysis of the protease ElxP giving some insights into the substrate scope of this family of proteins. Additional experiments towards the discovery of novel phosphonate antibiotics are described in supplemental chapter A.

1.6. REFERENCES

1. Taubes, G. (2008) The bacteria fight back, *Science 321*, 356-361.

2. Fischbach, M. A., and Walsh, C. T. (2009) Antibiotics for emerging pathogens, *Science 325*, 1089-1093.

3. Baltz, R. H. (2006) Marcel Faber Roundtable: is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration?, *J. Ind. Microbiol. Biotechnol. 33*, 507-513.

4. Willey, J. M., and van der Donk, W. A. (2007) Lantibiotics: Peptides of diverse structure and function, *Annu. Rev. Microbiol. 61*, 477-501.

5. Chatterjee, C., Paul, M., Xie, L., and van der Donk, W. A. (2005) Biosynthesis and mode of action of lantibiotics, *Chem. Rev. 105*, 633-684.

6. Breukink, E., and de Kruijff, B. (2006) Lipid II as a target for antibiotics, *Nat. Rev. Drug. Discov. 5*, 321-332.

7. Brötz, H., Josten, M., Wiedemann, I., Schneider, U., Götz, F., Bierbaum, G., and Sahl, H. G. (1998) Role of lipid-bound peptidoglycan precursors in the formation of pores by nisin, epidermin and other lantibiotics, *Mol. Microbiol. 30*, 317-327.

8. Breukink, E., Wiedemann, I., van Kraaij, C., Kuipers, O. P., Sahl, H., and de Kruijff, B. (1999) Use of the cell wall precursor lipid II by a pore-forming peptide antibiotic, *Science 286*, 2361-2364.

9. Hasper, H. E., Kramer, N. E., Smith, J. L., Hillman, J. D., Zachariah, C., Kuipers, O. P., de Kruijff, B., and Breukink, E. (2006) An alternative bactericidal mechanism of action for lantibiotic peptides that target lipid II, *Science 313*, 1636-1637.

10. Wiedemann, I., Breukink, E., van Kraaij, C., Kuipers, O. P., Bierbaum, G., de Kruijff, B., and Sahl, H. G. (2001) Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity, *J. Biol. Chem. 276*, 1772-1779.

11. Brumfitt, W., Salton, M. R., and Hamilton-Miller, J. M. (2002) Nisin, alone and combined with peptidoglycan-modulating antibiotics: activity against methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant *Enterococci*, *J. Antimicrob. Chemother. 50*, 731-734.

12. Mota-Meira, M., LaPointe, G., Lacroix, C., and Lavoie, M. C. (2000) MICs of mutacin B-Ny266, nisin A, vancomycin, and oxacillin against bacterial pathogens, *Antimicrob. Agents Chemother. 44*, 24-29.

13. Piper, C., Cotter, P. D., Ross, R. P., and Hill, C. (2009) Discovery of medically significant lantibiotics, *Curr. Drug. Discov. Technol. 6*, 1-18.

14. Bavin, E. M., Beach, A. S., Falconer, R., and Friedmann, R. (1952) Nisin in experimental tuberculosis, *Lancet 1*, 127-129.

15. Goldstein, B. P., Wei, J., Greenberg, K., and Novick, R. (1998) Activity of nisin against *Streptococcus pneumoniae*, *in vitro*, and in a mouse infection model, *J. Antimicrob. Chemother. 42*, 277-278.

16. Smith, L., and Hillman, J. (2008) Therapeutic potential of type A (I) lantibiotics, a group of cationic peptide antibiotics, *Curr. Opin. Microbiol. 11*, 401-408.

17. Kempf, M., Theobald, U., and Fiedler, H. P. (1999) Economic improvement of the fermentative production of gallidermin by *Staphylococcus gallinarum*, *Biotechnol. Lett. 21*, 663-667.

18. Medaglia, G., and Panke, S. (2010) Development of a fermentation process based on a defined medium for the production of pregallidermin, a nontoxic precursor of the lantibiotic gallidermin, *Appl. Microbiol. Biotechnol. 87*, 145-157.

19. Horner, T., Ungermann, V., Zahner, H., Fiedler, H. P., Utz, R., Kellner, R., and Jung, G. (1990) Comparative-studies on the fermentative production of lantibiotics by *Staphylococci*, *Appl. Microbiol. Biotechnol. 32*, 511-517.

20. Horner, T., Zahner, H., Kellner, R., and Jung, G. (1989) Fermentation and isolation of epidermin, a lanthionine containing polypeptide antibiotic from *Staphylococcus epidermidis*, *Appl. Microbiol. Biotechnol. 30*, 219-225.

21. Anonymous, Orgenics, Inc. - MU 1140 and the DPOLT™ platform. http://www.oragenics.com/?q=lantibiotics. Accessed on Sept. 18, 2011.

22. Xie, L., Miller, L. M., Chatterjee, C., Averin, O., Kelleher, N. L., and van der Donk, W. A. (2004) Lacticin 481: *In vitro* reconstitution of lantibiotic synthetase activity, *Science 303*, 679-681.

23. Furgerson Ihnken, L. A., Chatterjee, C., and van der Donk, W. A. (2008) *In vitro* reconstitution and substrate specificity of a lantibiotic protease, *Biochemistry 47*, 7352-7363.

24. Shi, Y., Yang, X., Garg, N., and van der Donk, W. A. (2011) Production of lantipeptides in *Escherichia coli*, *J. Am. Chem. Soc. 133*, 2338-2341.

25. van de Kamp, M., Horstink, L. M., van den Hooven, H. W., Konings, R. N. H., Hilbers, C. W., Frey, A., Sahl, H.-G., Metzger, J. W., and van de Ven, F. J. M. (1995) Sequence analysis by NMR spectroscopy of the peptide lantibiotic epilancin K7 from *Staphylococcus epidermidis* K7, *Eur. J. Biochem. 227*, 757-771.

26. van de Kamp, M., van den Hooven, H. W., Konings, R. N. H., Bierbaum, G., Sahl, H.-G., Kuipers, O. P., Siezen, R. J., de Vos, W., Hilbers, C. W., and van de Ven, F. J. (1995) Elucidation of the primary structure of the lantibiotic epilancin K7 from *Staphylococcus epidermis* K7. Cloning and characterization of the epilancin K7-encoding gene and NMR analysis of mature epilancin K7, *Eur. J. Biochem. 230*, 587-600.

27. Ekkelenkamp, M. B., Hanssen, M., Danny Hsu, S. T., de Jong, A., Milatovic, D., Verhoef, J., and van Nuland, N. A. (2005) Isolation and structural characterization of epilancin 15X, a novel lantibiotic from a clinical strain of *Staphylococcus epidermidis*, *FEBS Lett. 579*, 1917-1922.

28. Sashihara, T., Kimura, H., Higuchi, T., Adachi, A., Matsusaki, H., Sonomoto, K., and Ishizaki, A. (2000) A novel lantibiotic, nukacin ISK-1, of *Staphylococcus warneri* ISK-1: cloning of the structural gene and identification of the structure, *Biosci. Biotechnol. Biochem. 64*, 2420-2428.

29. Petersen, J., Boysen, A., Fogh, L., Tabermann, K., Kofoed, T., King, A., Schrotz-King, P., and Hansen, M. C. (2009) Identification and characterization of a bioactive lantibiotic produced by *Staphylococcus warneri*, *Biol. Chem. 390*, 437-444.

30. Verhoef, J., Milatovic, D., and Ekkelenkamp, M. B. (2005) Antimicrobial compounds, WO/2005/023852.

31. Davies, E. A., Bevis, H. E., Potter, R., Harris, J., Williams, G. C., and Delves-Broughton, J. (1998) Research note: The effect of pH on the stability of nisin solution during autoclaving, *Lett. Appl. Microbiol. 27*, 186–187.

32. Rollema, H. S., Kuipers, O. P., Both, P., de Vos, W. M., and Siezen, R. J. (1995) Improvement of solubility and stability of the antimicrobial peptide nisin by protein engineering, *Appl. Environ. Microbiol. 61*, 2873-2878.

33. Driessen, A. J., van den Hooven, H. W., Kuiper, W., van de Kamp, M., Sahl, H. G., Konings, R. N., and Konings, W. N. (1995) Mechanistic studies of lantibiotic-induced permeabilization of phospholipid vesicles, *Biochemistry 34*, 1606-1614.

34. Suda, S., Hill, C., Cotter, P. D., and Ross, R. P. (2010) Investigating the importance of charged residues in lantibiotics, *Bioeng. Bugs 1*, 345-351.

35. Breukink, E., van Kraaij, C., Demel, R. A., Siezen, R. J., Kuipers, O. P., and de Kruijff, B. (1997) The C-terminal region of nisin is responsible for the initial interaction of nisin with the target membrane, *Biochemistry 36*, 6968-6976.

36. Breukink, E., van Kraaij, C., van Dalen, A., Demel, R. A., Siezen, R. J., de Kruijff, B., and Kuipers, O. P. (1998) The orientation of nisin in membranes, *Biochemistry 37*, 8153-8162.

37. Sahl, H. G. (1991) Pore formation in bacterial membranes by cationic lantibiotics, In *Nisin and novel lantibiotics: Proceedings of the first International workshop on lantibiotics* (Jung, G., and Sahl, H. G., Eds.), p 347, Escom, Leiden.

38. Sahl, H. G. (1985) Influence of the staphylococcinlike peptide Pep5 on membrane potential of bacterial cells and cytoplasmic membrane vesicles, *J. Bacteriol. 162*, 833-836.

39. Hsu, S. T., Breukink, E., Tischenko, E., Lutters, M. A., De Kruijff, B., Kaptein, R., Bonvin, A. M., and Van Nuland, N. A. (2004) The nisin-lipid II complex reveals a pyrophosphate cage that provides a blueprint for novel antibiotics, *Nat. Struct. Mol. Biol. 11*, 963-967.

40. Hosoda, K., Ohya, M., Kohno, T., Maeda, T., Endo, S., and Wakamatsu, K. (1996) Structure determination of an immunopotentiator peptide, cinnamycin, complexed with lysophosphatidylethanolamine by ^{1}H-NMR, *J. Biochem. 119*, 226-230.

41. Ökesli, A., Cooper, L. E., Fogle, E. J., and van der Donk, W. A. (2011) Nine post-translational modifications during the biosynthesis of cinnamycin, *J. Am. Chem. Soc. 133*, 13753-13760.

42. Castiglione, F., Lazzarini, A., Carrano, L., Corti, E., Ciciliato, I., Gastaldo, L., Candiani, P., Losi, D., Marinelli, F., Selva, E., and Parenti, F. (2008) Determining the structure and mode of action of microbisporicin, a potent lantibiotic active against multiresistant pathogens, *Chem. Biol. 15*, 22-31.

43. He, Z., Kisla, D., Zhang, L., Yuan, C., Green-Church, K. B., and Yousef, A. E. (2007) Isolation and identification of a *Paenibacillus polymyxa* strain that coproduces a novel lantibiotic and polymyxin, *Appl. Environ. Microbiol. 73*, 168-178.

44. He, Z., Yuan, C., Zhang, L., and Yousef, A. E. (2008) N-terminal acetylation in paenibacillin, a novel lantibiotic, *FEBS Lett. 582*, 2787-2792.

45. Goto, Y., Li, B., Claesen, J., Shi, Y., Bibb, M. J., and van der Donk, W. A. (2010) Discovery of unique lanthionine synthetases reveals new mechanistic and evolutionary insights, *Plos Biology 8*, e1000339.

46. Kodani, S., Hudson, M. E., Durrant, M. C., Buttner, M. J., Nodwell, J. R., and Willey, J. M. (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene *ramS* in *Streptomyces coelicolor*, *Proc. Natl. Acad. Sci. U. S. A. 101*, 11448-11453.

47. Meindl, K., Schmiederer, T., Schneider, K., Reicke, A., Butz, D., Keller, S., Guhring, H., Vertesy, L., Wink, J., Hoffmann, H., Bronstrup, M., Sheldrick, G. M., and Süssmuth, R. D. (2010) Labyrinthopeptins: A new class of carbacyclic lantibiotics, *Angew. Chem. Int. Ed. Engl. 49*, 1151-1154.

48. Müller, W. M., Schmiederer, T., Ensle, P., and Süssmuth, R. D. (2010) *In vitro* biosynthesis of the prepeptide of type-III lantibiotic labyrinthopeptin A2 including formation of a C-C bond as a post-translational modification, *Angew. Chem. Int. Ed. Engl. 49*, 2436-2440.

49. Goto, Y., Ökesli, A., and van der Donk, W. A. (2011) Mechanistic studies of Ser/Thr dehydration catalyzed by a member of the LanL lanthionine synthetase family, *Biochemistry 50*, 891-898.

50. Lubelski, J., Khusainov, R., and Kuipers, O. P. (2009) Directionality and coordination of dehydration and ring formation during biosynthesis of the lantibiotic nisin, *J. Biol. Chem. 284*, 25962-25972.

51. Marsh, A. J., O'Sullivan, O., Ross, R. P., Cotter, P. D., and Hill, C. (2010) *In silico* analysis highlights the frequency and diversity of type 1 lantibiotic gene clusters in genome sequenced bacteria, *BMC Genomics 11*, 679.

52. van der Meer, J. R., Rollema, H. S., Siezen, R. J., Beerthuyzen, M. M., Kuipers, O. P., and de Vos, W. M. (1994) Influence of amino acid substitutions in the nisin leader peptide on biosynthesis and secretion of nisin by *Lactococcus lactis*, *J. Biol. Chem. 269*, 3555-3562.

53. Plat, A., Kluskens, L. D., Kuipers, A., Rink, R., and Moll, G. N. (2011) Requirements of the engineered leader peptide of nisin for inducing modification, export, and cleavage, *Appl. Environ. Microbiol. 77*, 604-611.

54. Mavaro, A., Abts, A., Bakkes, P. J., Moll, G. N., Driessen, A. J., Smits, S. H., and Schmitt, L. (2011) Substrate recognition and specificity of the NisB protein, the lantibiotic dehydratase involved in nisin biosynthesis, *J. Biol. Chem. 286*, 30552-30560.

55. Kluskens, L. D., Kuipers, A., Rink, R., de Boef, E., Fekken, S., Driessen, A. J., Kuipers, O. P., and Moll, G. N. (2005) Post-translational modification of therapeutic peptides by NisB, the dehydratase of the lantibiotic nisin, *Biochemistry 44*, 12827-12834.

56. Kelly, W. L., Pan, L., and Li, C. (2009) Thiostrepton biosynthesis: prototype for a new family of bacteriocins, *J. Am. Chem. Soc. 131*, 4327-4334.

57. Wieland-Brown, L. C., Acker, M. G., Clardy, J., Walsh, C. T., and Fischbach, M. A. (2009) Thirteen posttranslational modifications convert a 14-residue peptide into the antibiotic thiocillin, *Proc. Natl. Acad. Sci. U. S. A. 106*, 2549-2553.

58. Liao, R., Duan, L., Lei, C., Pan, H., Ding, Y., Zhang, Q., Chen, D., Shen, B., Yu, Y., and Liu, W. (2009) Thiopeptide biosynthesis featuring ribosomally synthesized precursor peptides and conserved posttranslational modifications, *Chem. Biol. 16*, 141-147.

59. Morris, R. P., Leeds, J. A., Naegeli, H. U., Oberer, L., Memmert, K., Weber, E., LaMarche, M. J., Parker, C. N., Burrer, N., Esterow, S., Hein, A. E., Schmitt, E. K., and Krastel, P. (2009) Ribosomally synthesized thiopeptide antibiotics targeting elongation factor Tu, *J. Am. Chem. Soc. 131*, 5946-5955.

60. Onaka, H., Nakaho, M., Hayashi, K., Igarashi, Y., and Furumai, T. (2005) Cloning and characterization of the goadsporin biosynthetic gene cluster from *Streptomyces* sp. TP-A0584, *Microbiology 151*, 3923-3933.

61. Tillett, D., Dittmann, E., Erhard, M., von Dohren, H., Borner, T., and Neilan, B. A. (2000) Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system, *Chem. Biol. 7*, 753-764.

62. Fukuchi, N., Isogai, A., Nakayama, J., and Suzuki, A. (1990) Structure of syringotoxin B, a phytotoxin produced by citrus isolates of *Pseudomonas syringae* pv. *syringae*, *Agric. Biol. Chem. 54*, 3377-3379.

63. Nishizawa, T., Asayama, M., Fujii, K., Harada, K., and Shirai, M. (1999) Genetic analysis of the peptide synthetase genes for a cyclic heptapeptide microcystin in *Microcystis* spp., *J. Biochem. 126*, 520-529.

64. Christiansen, G., Fastner, J., Erhard, M., Borner, T., and Dittmann, E. (2003) Microcystin biosynthesis in *Planktothrix*: genes, evolution, and manipulation, *J. Bacteriol. 185*, 564-572.

65. Nishizawa, T., Ueda, A., Asayama, M., Fujii, K., Harada, K., Ochi, K., and Shirai, M. (2000) Polyketide synthase gene coupled to the peptide synthetase module involved in the biosynthesis of the cyclic heptapeptide microcystin, *J. Biochem. 127*, 779-789.

66. Milne, J. C., Roy, R. S., Eliot, A. C., Kelleher, N. L., Wokhlu, A., Nickels, B., and Walsh, C. T. (1999) Cofactor requirements and reconstitution of microcin B17 synthetase: a multienzyme complex that catalyzes the formation of oxazoles and thiazoles in the antibiotic microcin B17, *Biochemistry 38*, 4768-4781.

67. Lee, S. W., Mitchell, D. A., Markley, A. L., Hensler, M. E., Gonzalez, D., Wohlrab, A., Dorrestein, P. C., Nizet, V., and Dixon, J. E. (2008) Discovery of a widely distributed toxin biosynthetic gene cluster, *Proc. Natl. Acad. Sci. U. S. A. 105*, 5879-5884.

68. Schmidt, E. W., McIntosh, J. A., and Donia, M. S. (2010) Insights into heterocyclization from two highly similar enzymes, *J. Am. Chem. Soc. 132*, 4089-4091.

69. You, Y. O., and van der Donk, W. A. (2007) Mechanistic investigations of the dehydration reaction of lacticin 481 synthetase using site-directed mutagenesis, *Biochemistry 46*, 5991-6000.

70. Chatterjee, C., Miller, L. M., Leung, Y. L., Xie, L., Yi, M., Kelleher, N. L., and van der Donk, W. A. (2005) Lacticin 481 synthetase phosphorylates its substrate during lantibiotic production, *J. Am. Chem. Soc. 127*, 15332-15333.

71. Lee, M. V., Ihnken, L. A., You, Y. O., McClerren, A. L., van der Donk, W. A., and Kelleher, N. L. (2009) Distributive and directional behavior of lantibiotic synthetases revealed by high-resolution tandem mass spectrometry, *J. Am. Chem. Soc. 131*, 12258-12264.

72. Patton, G. C., Paul, M., Cooper, L. E., Chatterjee, C., and van der Donk, W. A. (2008) The importance of the leader sequence for directing lanthionine formation in lacticin 481, *Biochemistry 47*, 7342-7351.

73. Müller, W. M., Ensle, P., Krawczyk, B., and Süssmuth, R. D. (2011) Leader peptide-directed processing of labyrinthopeptin A2 precursor peptide by the modifying enzyme LabKC, *Biochemistry*, [Epub ahead of print]. dx.doi.org/10.1021/bi200526q.

74. Li, B., Yu, J. P., Brunzelle, J. S., Moll, G. N., van der Donk, W. A., and Nair, S. K. (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis, *Science 311*, 1464-1467.

75. Li, B., and van der Donk, W. A. (2007) Identification of essential catalytic residues of the cyclase NisC involved in the biosynthesis of nisin, *J. Biol. Chem. 282*, 21169-21175.

76. Paul, M., Patton, G. C., and van der Donk, W. A. (2007) Mutants of the zinc ligands of lacticin 481 synthetase retain dehydration activity but have impaired cyclization activity, *Biochemistry 46*, 6268-6276.

77. McClerren, A. L., Cooper, L. E., Quan, C., Thomas, P. M., Kelleher, N. L., and van der Donk, W. A. (2006) Discovery and *in vitro* biosynthesis of haloduracin, a two-component lantibiotic, *Proc. Natl. Acad. Sci. U. S. A. 103*, 17243-17248.

78. Prasch, T., Naumann, T., Markert, R. L., Sattler, M., Schubert, W., Schaal, S., Bauch, M., Kogler, H., and Griesinger, C. (1997) Constitution and solution conformation of the antibiotic mersacidin determined by NMR and molecular dynamics, *Eur. J. Biochem. 244*, 501-512.

79. Kaletta, C., Entian, K. D., and Jung, G. (1991) Prepeptide sequence of cinnamycin (Ro 09-0198): the first structural gene of a duramycin-type lantibiotic, *Eur. J. Biochem. 199*, 411-415.

80. Li, B., Sher, D., Kelly, L., Shi, Y. X., Huang, K., Knerr, P. J., Joewono, I., Rusch, D., Chisholm, S. W., and van der Donk, W. A. (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria, *Proc. Natl. Acad. Sci. U. S. A. 107*, 10430-10435.

81. Chatterjee, C., Patton, G. C., Cooper, L., Paul, M., and van der Donk, W. A. (2006) Engineering dehydro amino acids and thioethers into peptides using lacticin 481 synthetase, *Chem. Biol. 13*, 1109-1117.

82. Skaugen, M., Abildgaard, C. I., and Nes, I. F. (1997) Organization and expression of a gene cluster involved in the biosynthesis of the lantibiotic lactocin S, *Mol. Gen. Genet. 253*, 674-686.

83. Meyer, C., Bierbaum, G., Heidrich, C., Reis, M., Süling, J., Iglesias-Wind, M. I., Kempter, C., Molitor, E., and Sahl, H.-G. (1995) Nucleotide sequence of the lantibiotic Pep5 biosynthetic gene cluster and functional analysis of PepP and PepC, *Eur. J. Biochem. 232*, 478-489.

84. Nishie, M., Shioya, K., Nagao, J., Jikuya, H., and Sonomoto, K. (2009) ATP-dependent leader peptide cleavage by NukT, a bifunctional ABC transporter, during lantibiotic biosynthesis, *J. Biosci. Bioeng. 108*, 460-464.

85. van der Meer, J. R., Polman, J., Beerthuyzen, M. M., Siezen, R. J., Kuipers, O. P., and de Vos, W. M. (1993) Characterization of the *Lactococcus lactis* nisin A operon genes *nisP*, encoding a subtilisin-like serine protease involved in precursor processing, and *nisR*, encoding a regulatory protein involved in nisin biosynthesis, *J. Bacteriol. 175*, 2578-2588.

86. Schnell, N., Engelke, G., Augustin, J., Rosenstein, R., Ungermann, V., Götz, F., and Entian, K.-D. (1992) Analysis of genes involved in the biosynthesis of the lantibiotic epidermin, *Eur. J. Biochem. 204*, 57-68.

87. Gilmore, M. S., Segarra, R. A., and Booth, M. C. (1990) An HlyB-type function is required for expression of the *Enterococcus faecalis* hemolysin/bacteriocin, *Infect. Immun. 58*, 3914-3923.

88. Kuipers, A., de Boef, E., Rink, R., Fekken, S., Kluskens, L. D., Driessen, A. J., Leenhouts, K., Kuipers, O. P., and Moll, G. N. (2004) NisT, the transporter of the lantibiotic nisin, can transport fully modified, dehydrated, and unmodified prenisin and fusions of the leader peptide with non-lantibiotic peptides, *J. Biol. Chem. 279*, 22176-22182.

89. NCBI, Conserved domains: LanC_like (cd004434). http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=cd04434. Accessed on September 24, 2011.

90. Marchler-Bauer, A., Anderson, J. B., Derbyshire, M. K., DeWeese-Scott, C., Gonzales, N. R., Gwadz, M., Hao, L., He, S., Hurwitz, D. I., Jackson, J. D., Ke, Z., Krylov, D., Lanczycki, C. J., Liebert, C. A., Liu, C., Lu, F., Lu, S., Marchler, G. H., Mullokandov, M., Song, J. S., Thanki, N., Yamashita, R. A., Yin, J. J., Zhang, D., and Bryant, S. H. (2007) CDD: a conserved domain database for interactive domain family analysis, *Nucleic Acids Res. 35*, D237-240.

91. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol. Biol. Evol. 28*, 2731-2739.

92. Dirix, G., Monsieurs, P., Dombrecht, B., Daniels, R., Marchal, K., Vanderleyden, J., and Michiels, J. (2004) Peptide signal molecules and bacteriocins in Gram-negative bacteria: a genome-wide *in silico* screening for peptides containing a double-glycine leader sequence and their cognate transporters, *Peptides 25*, 1425-1440.

93. Lawton, E. M., Cotter, P. D., Hill, C., and Ross, R. P. (2007) Identification of a novel two-peptide lantibiotic, haloduracin, produced by the alkaliphile *Bacillus halodurans* C-125, *FEMS Microbiol. Lett. 267*, 64-71.

94. Cooper, L. E., McClerren, A. L., Chary, A., and van der Donk, W. A. (2008) Structure-activity relationship studies of the two-component lantibiotic haloduracin, *Chem. Biol. 15*, 1035-1045.

95. Velásquez, J. E., and van der Donk, W. A. (2011) Genome mining for ribosomally synthesized natural products, *Curr. Opin. Chem. Biol. 15*, 11-21.

96. Begley, M., Cotter, P. D., Hill, C., and Ross, R. P. (2009) Identification of a novel two-peptide lantibiotic, lichenicidin, following rational genome mining for LanM proteins, *Appl. Environ. Microbiol. 75*, 5451-5460.

97. Dischinger, J., Josten, M., Szekat, C., Sahl, H. G., and Bierbaum, G. (2009) Production of the novel two-peptide lantibiotic lichenicidin by *Bacillus licheniformis* DSM 13, *Plos One 4*, e6788.

98. Shenkarev, Z. O., Finkina, E. I., Nurmukhamedova, E. K., Balandin, S. V., Mineev, K. S., Nadezhdin, K. D., Yakimenko, Z. A., Tagaev, A. A., Temirov, Y. V., Arseniev, A. S., and Ovchinnikova, T. V. (2010) Isolation, structure elucidation, and synergistic antibacterial activity of a novel two-component lantibiotic lichenicidin from *Bacillus licheniformis* VK21, *Biochemistry 49*, 6462-6472.

99. Daly, K. M., Upton, M., Sandiford, S. K., Draper, L. A., Wescombe, P. A., Jack, R. W., O'Connor, P. M., Rossney, A., Götz, F., Hill, C., Cotter, P. D., Ross, R. P., and Tagg, J. R. (2010) Production of the Bsa lantibiotic by community-acquired *Staphylococcus aureus* strains, *J. Bacteriol. 192*, 1131-1142.

100. Haft, D. H., Basu, M. K., and Mitchell, D. A. (2010) Expansion of ribosomally produced natural products: a nitrile hydratase- and Nif11-related precursor family, *BMC Biol 8*, 70.

101. Wang, H., Fewer, D. P., and Sivonen, K. (2011) Genome mining demonstrates the widespread occurrence of gene clusters encoding bacteriocins in cyanobacteria, *Plos One 6*, e22384.

102. de Jong, A., van Heel, A. J., Kok, J., and Kuipers, O. P. (2010) BAGEL2: mining for bacteriocins in genomic data, *Nucleic Acids Res. 38 Suppl*, W647-651.

103. Majchrzykiewicz, J. A., Lubelski, J., Moll, G. N., Kuipers, A., Bijlsma, J. J., Kuipers, O. P., and Rink, R. (2010) Production of a class II two-component lantibiotic of *Streptococcus pneumoniae* using the class I nisin synthetic machinery and leader sequence, *Antimicrob. Agents Chemother. 54*, 1498-1505.

104. Field, D., Hill, C., Cotter, P. D., and Ross, R. P. (2010) The dawning of a 'Golden era' in lantibiotic bioengineering, *Mol. Microbiol. 78*, 1077-1087.

105. Field, D., Connor, P. M., Cotter, P. D., Hill, C., and Ross, R. P. (2008) The generation of nisin variants with enhanced activity against specific gram-positive pathogens, *Mol. Microbiol. 69*, 218-230.

106. Appleyard, A. N., Choi, S., Read, D. M., Lightfoot, A., Boakes, S., Hoffmann, A., Chopra, I., Bierbaum, G., Rudd, B. A., Dawson, M. J., and Cortes, J. (2009) Dissecting structural and functional diversity of the lantibiotic mersacidin, *Chem. Biol. 16*, 490-498.

107. Islam, M. R., Shioya, K., Nagao, J., Nishie, M., Jikuya, H., Zendo, T., Nakayama, J., and Sonomoto, K. (2009) Evaluation of essential and variable residues of nukacin ISK-1 by NNK scanning, *Mol. Microbiol. 72*, 1438-1447.

108. Field, D., Quigley, L., O'Connor, P. M., Rea, M. C., Daly, K., Cotter, P. D., Hill, C., and Ross, R. P. (2010) Studies with bioengineered nisin peptides highlight the broad-spectrum potency of nisin V, *Microb. Biotechnol. 3*, 473-486.

109. Levengood, M. R., Knerr, P. J., Oman, T. J., and van der Donk, W. A. (2009) *In vitro* mutasynthesis of lantibiotic analogues containing nonproteinogenic amino acids, *J. Am. Chem. Soc. 131*, 12024-12025.

CHAPTER 2. CHARACTERIZATION OF THE GENE CLUSTER AND THE BIOSYNTHETIC PATHWAY OF THE LANTIBIOTIC EPILANCIN 15X*

2.1. INTRODUCTION

The lantibiotic epilancin 15X produced by *Staphylococcus epidermidis* 15X154 is a member of the epilancin-group of peptides that have shown potent antimicrobial activity against pathogenic bacteria, including methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant *Enterococci* (VRE). Like the other members of the family, epilancin 15X contains one lanthionine (Lan) and two 3-methyllanthionine (MeLan) bridges, one 2,3-didehydroalanine (Dha), three (*Z*)-2,3-didehydrobutyrine (Dhb) residues, and an unusual N-terminal 2-hydroxypropionyl group (lactate, Lac) that might be important for antimicrobial activity. Despite the promising biological properties of epilancin 15X, as well as other members of the family, its biosynthetic route has not been characterized and the stereochemical configuration and role of the N-terminal modification in the biological activity have not been determined.

The gene clusters and biosynthetic pathways of some other staphylococcal lantibiotics, such as Pep5, epidermin, and epicidin 280, have been previously characterized (*1*-*3*). Furthermore, an N-terminal Lac is also present in epicidin 280 and a putative oxidoreductase designated EciO was speculated to be involved in its biosynthesis, but this hypothesis has not been confirmed experimentally (*2*). Understanding how the N-terminal Lac group is introduced into epilancin 15X and other

* Reproduced in part with permission from: "Velásquez, J. E., Zhang, X., and van der Donk, W. A. (2011) Biosynthesis of the antimicrobial peptide epilancin 15X and its N-terminal lactate, *Chem. Biol.* 18, 857-867." Copyright © 2011 Elsevier Ltd.

lantibiotics may facilitate the elucidation of its biological role and the design of novel peptides with enhanced antimicrobial properties. This chapter focuses on the characterization of the biosynthetic gene cluster of epilancin 15X, obtained after the construction of a genome library of *S. epidermidis* 15X154 and sequencing of fosmids encoding for lantibiotic protein homologs. Additionally, an NADP(H)-dependent short-chain alcohol dehydrogenase/reductase (SDR) designated as ElxO is demonstrated to catalyze the conversion of an N-terminal pyruvyl (Pyr) group in a epilancin 15X precursor to Lac. The expression, purification, and *in vitro* reconstitution of the enzymatic activity of a peptidase designated ElxP that is involved in leader peptide cleavage is also reported, to the best of my knowledge the first such example for a member of the serine-type lantibiotic proteases. Finally, the *in vivo* reconstitution of the enzymatic activity of a Ser/Thr dehydratase designated ElxB is described, demonstrating its role in biosynthesis. The proposed biosynthetic route of epilancin 15X constitutes the first example for a member of the epilancin-group of lantibiotics and opens the door for engineering of novel peptides with biological activity against pathogenic bacteria.

2.2. EXPERIMENTAL METHODS

Materials, organisms, media, and growth conditions

Chemical reagents and media components used in this study were purchased from Sigma-Aldrich or Thermo Fisher Scientific, unless otherwise specified, and were used without further purification. The strains and plasmids are listed in Table 2.1 and the primers are shown in Table 2.2. *S. epidermidis* 15X154 and *S. carnosus* TM300

were grown on Luria – Bertani (LB) or Mueller Hinton (Oxoid) solid agar or liquid broth at 37 °C. *Escherichia coli* strains were routinely grown in LB solid agar or broth supplemented with antibiotics at 37 °C. Strain *E. coli* WM4489 (*4, 5*), used as the host for genomic library construction, was supplemented with maltose (10 mM) before phage transfection, while strains containing fosmid pJK050 (*4, 5*) or its derivatives were supplemented with rhamnose (10 mM) preceding fosmid isolation. For agar diffusion bioactivity assays, 25 mL of agar medium inoculated with culture (1/100 dilution) was poured into a sterile plate. Aliquots of antibacterial compounds were placed into wells made on the solidified agar and the plates were incubated at 37 or 30 °C overnight. The bioactivity of the sample was confirmed if a clear growth inhibition zone surrounding the well was formed.

Table 2.1. Microorganisms and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source or reference
Escherichia coli		
BL21(DE)	*fhuA2* [*lon*] *ompT gal* (λ DE3) [*dcm*] Δ*hsdS* λ DE3 = λ *sBamHIo* Δ*EcoRI-B int*::(*lacI*::*PlacUV5*::*T7 gene1*) *i21* Δ*nin5*	Novagen
DH5*α*	*λpir/ϕ80dlacZΔM15* Δ(*lacZYA-argF*)*U169 recA1 hsdR17 deoR thi-1 supE44 gyrA96 relA1*	*(6)*
Rosetta2(DE3)	F^- *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) λ(DE3 [*lacI lacUV5-T7 gene 1 ind1 sam7 nin5*]) pRARE2	Novagen
T7 Express	*fhuA2 lacZ*::*T7 gene1* [*lon*] *ompT gal sulA11* R(*mcr-73*::*miniTn10*--Tet^S)*2* [*dcm*] R(*zgb-210*::*Tn10*--Tet^S) *endA1* Δ(*mcrC-mrr*)*114*::*IS10*	New England Biolabs
WM4489	*E. coli* DH10B derivative: *mcrA* Δ(*mrr hsdRMS mcrBC*) *ϕ80*(Δ*lacM15*) Δ*lacX74 endA1 recA1 deoR* Δ(*ara-leu*)*7697 araD139 galU galK nupG rpsL λattB*::pAE12(*PrhaB*::*trfA33* Δ*oriR6K-cat*::*frt5*)	(*5*)
Staphylococcus		
epidermidis 15X154	Epilancin 15X producer strain	(*7*)
carnosus TM300	Epilancin 15X sensitive strain	G. Bierbaum, U. of Bonn
Plasmids		
pACYCDuet-1	Cm^R *E. coli* T7 based vector for coexpression vector of two target genes	Novagen
pACYC.ElxC	*elxC* cloned in pACYCDuet-1 vector	This study
pACYC.NisBC	*nisB and nisC* cloned in pACYCDuet-1	N. Bindman, van der Donk laboratory
pACYC.NisC	*nisC* cloned in pACYCDuet-1	(*8*)
pAE5	Source of mini-Mu transposon	(*5*)
pCDF.NisB	*nisB* cloned in pCDFDuet-1	N. Bindman, van der Donk laboratory
pET28b	Kan^R *E. coli* T7 based histidine-tag fusion expression vector	Novagen
pET.His_6-ElxA	*elxA* cloned in pET28b vector	This study
pET.His_6-ElxO	*elxO* cloned in pET28b vector	This study
pET.His_6-MBP-ElxP	*elxP* cloned in pET28b vector	This study
pJK050	*oriV*, *oriS*, copy-control cosvector, Cm^R	(*5*)
pRSFDuet-1	Kan^R *E. coli* T7 based histidine-tag fusion vector for coexpression of two target genes	Novagen
pRSF.His_6-ElxAB	*elxA* and *elxB* cloned in pRSFDuet-1 vector	This study
pRSF.His_6-NisElxA.NisB	Chimeric *nisA*(leader)-*elxA*(core) gene and *nisB* cloned in pRSFDuet-1 vector	(*8*)
pRSF.His_6-NisAB	*nisA* and *nisB* cloned in pRSFDuet-1 vector	This study

[a]ATCC, American Type Culture Collection, Manassas, VA.

Table 2.2. Primers used in this study

Name	Sequence
elxA.BamHI.F	5'- GCG AGC CAG GAT CCG ATG AAA AAA GAA TTA TTT GAT TTA AAT CTT AAT AAA G -3'
elxA.EcoRI.R	5'- GGC GCG CCG AAT TCT TAT TTT TTA CCA GTA AAG TG -3'
elxA.NdeI.F	5'- GGC GCG CCC ATA TGA AAA AAG AAT TAT TTG ATT TAA ATC TTA AT -3'
elxA.NotI.R	5'- GCA TTA TGC GGC CGC TTA TTT TTT ACC AGT AAA GTG ACA TCC ACA AGT TAG -3'
elxA-F1	5'-ATG AAT AAC GAA TTA TT(C/T) (A/G)AT TTG GAT C-3'
elxA-R2	5'-CTT TGT AGA GGA TTT ACA CTA ACT TG-3'
elxB.BglII.F	5'- GAA GGA GAT ATA CAT ATG GCA GAT CTC ATG AAC ATC TTC AAA AAA T -3'
elxB.XhoI.R	5'- CGG TTT CTT TAC CAG ACT CGA GTT AGT TGA TTT TTT TGT AG -3'
elxBgapF	5'- GTT ATT TCT TGA ATT ATT CTG CTA CAG AGC -3'
elxC.NdeI.F	5'- GGA TAT ACA TAT GGA AAA TAG TAT CCA AAA ATC CTT ATC ATA CCT TTC AG -3'
elxC.XhoI.R	5'- ACC AGA CTC GAG TTA CGC AAA ACA AAA CAA TTT ATA CC -3'
elxC-596F	5'-AT(C/T) TAG G(A/T)T A(T/C)G C(A/T/G)C ATG G(A/T)A T-3'
elxC-781R	5'- CC(A/G) TA(A/G) CAC CAA (C/G)C(A/G) T(C/T)T CT-3'
elxCgapR	5'- AAC CCA TAA CAC CAG TAA ATA ATG AAG TAG -3'
elxO.NheI.F	5'- GGT TGG TTG CTA GCA TGA AAA AAA ATG TTC TTA TTA CAG -3'
elxO.XhoI.R	5'- AAG TCG ACC TCG AGT TAT TGG GAT AAA TAT CCT C -3'
elxP.28MBP.F	5'- GAA CCT GTA CTT CCA ATC CGG ATC CAT GGA TAA TTT TCT TAG TTG GCC TAA TAA A -3'
elxP.28MBP.R	5'- GGT GGT GGT GGT GGT GCT CGA GTT AAA TAT GTT CAT TAG TAA TAA TCT CCT TTG TG -3'
nis-elxA.R-1E.F	5'- GAA AGA TTC AGG TGC ATC ACC AGA GTC AGC TAG TAT TGT TAA AAC AAC T -3'
nis-elxA.R-1E.R	5'- CTT AAG CAT TAT GCG GCC GCA AGC TTT TAT TTT TTA CCA GTA AAG TGA C -3'
seqaetf	5'- TCG CCT TCT TGA CGA GTT CT -3'
seqaetr	5'- TAG GAA CTT CGG GAT CCG TT- 3'

S. epidermidis 15X154 genomic DNA isolation

An aliquot of *S. epidermidis* 15X154 culture was pelleted by centrifugation and cells were washed with TE25S buffer (25 mM Tris-HCl, 25 mM EDTA, 0.3 M sucrose, pH 8). Cells were then lysed with 2 mg/mL of lysozyme (Sigma) and 50 μg/mL of lysostaphin (Sigma) in TE25S buffer at 37 °C for 60 min. The washed protoplasted cells

were incubated with proteinase K (0.15 mg/mL, Sigma) in TE25S buffer at 50 °C for 30 min. Sodium dodecyl sulfate (SDS) was added to 0.5% (w/v) final concentration and gently mixed for 10 min at room temperature. The cell lysate was mixed with an equal volume of saturated buffer of phenol, chloroform, and isoamyl alcohol (25:24:1). After centrifugation, the aqueous layer was removed and mixed gently with an equal volume of a mixture of chloroform and isoamyl alcohol (24:1). The top layer was removed again and 0.1 volumes of 5 M NaCl and 0.7 volumes of isopropanol were added to precipitate the DNA. The DNA was washed three times with 70% ethanol and once with 100% ethanol, air-dried, and resuspended in TE buffer (pH 8.0). See *Notebook III, page 05*.

Genomic library construction, screening, and DNA sequencing

Genomic DNA was partially digested with *Sau*3A1 (New England Biolabs) in the presence of RNAse A (Sigma) using serial dilutions of the enzyme. The DNA solutions were analyzed by field inversion gel electrophoresis on a 1% agarose/TBE gel and the fraction containing ~20-60 kb DNA fragments was treated with shrimp alkaline phosphatase (Roche Diagnostics). Cosmid pJK050 was digested with *Nhe*I (New England Biolabs), treated with shrimp alkaline phosphatase (Roche Diagnostics), and purified by sequential extraction with phenol:chloroform:isoamyl alcohol (25:24:1) and precipitation with ammonium sulfate and ethanol. The cosmid was further treated with *Bam*HI (Invitrogen). Digested pJK050 and genomic DNA were ligated with T4 DNA ligase (New England Biolabs) and cosmid constructs were packaged into lambda phage with a MaxPlax Lambda Packaging Extract Kit (Epicentre) according to the manufacturer's protocol, followed by transfection of *E. coli* WM4489. Library clones

were screened by PCR for the *elxA* gene using forward primer elxA-F1 and reverse primer elxA-R2 (Table 2.2), or for the *elxC* gene with forward primer elxC-596F and reverse primer elxC-781R, *Taq* polymerase (Invitrogen), and 1× PCR premix A (Epicentre).

Plasmid pAE5 was digested with *Bgl*II and the 1084 bp fragment to be used as a transposon was gel purified. For each positive fosmid, the fragment and the positive fosmid were mixed and treated with MuA transposase (Finnzymes) and the dialyzed DNA was used to transform *E. coli* WM4489. Sets of 192 clones were sequenced from the ends of the transposon insert using primers seqaetf and seqaetr at the Keck Center for Comparative and Functional Genomics at the University of Illinois at Urbana-Champaign. Using *S. epidermidis* 15X154 as template, the specific primers elxBgapF and elxCgapR were used to amplify and sequence a DNA fragment of 1.2 kb, closing the sequence gap between the fosmids. See also *Notebook III, page 25*.

Construction of plasmids pET.His$_6$-ElxO, pET.His$_6$-ElxA, and pET.His$_6$-MBP-ElxP

The gene *elxO* was amplified by PCR from *S. epidermidis* 15X154 genomic DNA using a forward primer containing an *Nhe*I restriction site (elxO.NheI.F) and a reverse primer containing a *Xho*I site (elxO.XhoI.R). The PCR product and the vector pET28b (Novagen) were digested with restriction endonucleases *Nhe*I and *Xho*I (Invitrogen) and ligated using T4 DNA ligase (New England Biolabs) to produce the plasmid pET.His$_6$-ElxO. See also *Notebook III, page 60*.

The gene *elxA* was amplified by PCR from *S. epidermidis* 15X154 genomic DNA using a forward primer containing an *Nde*I restriction site (elxA.NdeI.F) and a reverse

primer containing a *EcoR*I site (elxA.EcoRI.R). The PCR product and the vector pET28b (Novagen) were digested with restriction endonucleases *Nde*I and *EcoR*I (Invitrogen) and ligated using T4 DNA ligase (New England Biolabs) to produce the plasmid pET.His_6-ElxA. See also *Notebook III, page 71*.

The gene *elxP* was amplified by PCR from *S. epidermidis* 15X154 genomic DNA using a forward primer (elxP.28MBP.F) and a reverse primer (elxP.28MBP.R). The PCR product contained annealing regions to a modified pET28b vector (Novagen) that encodes for a fusion of pelB to hexahistidine tagged MBP. The purified PCR product was used as a primer to amplify the vector using Phusion® Hot Start DNA polymerase (New England Biolabs), followed by treatment with *Dpn*I (New England Biolabs) before transformation of *E. coli* DH5α cells to generate the plasmid pET.His_6-MBP-ElxP that encodes for ElxP fused at its N-terminus to a TEV protease cleavage site, an MBP tag, a hexahistidine tag, and a pelB signal peptide. The correct sequences of the inserts were confirmed by sequencing at the Keck Center at the University of Illinois at Urbana-Champaign. See also *Notebook VII, page 05*.

Overproduction and purification of His_6-ElxO

Electrocompetent *E. coli* Rosetta2(DE3) cells were transformed with pET.His_6-ElxO and a single colony was grown in LB medium supplemented with antibiotics. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM when OD_{600} = 0.7 and the culture was shaken for additional 16 h at 18 °C. The cells were harvested by centrifugation and the pellet was resuspended in lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 5 mM imidazole, pH 7.5) and stored at −80 °C until used.

The cell pellet was thawed, lysozyme was added to a final concentration of 1 mg/mL, and the suspension was passed through a French-press after incubation for 30 min at 4 °C. The protein was purified by immobilized metal ion affinity chromatography (IMAC) using an ÄKTApurifier (Amersham Biosciences, GE Healthcare) equipped with a HisTrap HP 5 mL column prepacked with Ni Sepharose™ (GE Healthcare). A gradient of 0-100% buffer B (20 mM Na_2HPO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4) in buffer A (20 mM Na_2HPO_4, 500 mM NaCl, 20 mM imidazole, pH 7.4) over 20 column volumes (CV) was used to elute the proteins. The fractions containing His_6-ElxO were concentrated using an Amicon Ultracel 10k filter (Millipore). Imidazole was removed by using a PD-10 desalting column (GE Healthcare) and the protein was eluted and stored in buffer (50 mM Tris-HCl, 300 mM NaCl, 20% glycerol, pH 7.5). The native molecular weight of His_6-ElxO was determined by size exclusion chromatography using an ÄKTApurifier equipped with a Superdex 200 HR 10/30 GL column (GE Healthcare) and utilizing cytrochrome c (12.4 kDa), carbonic anydrase (29 kDa), bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), β-amylase (200 kDa), and blue dextran (2000 kDa) as standards (MWGF200, Sigma). See also *Notebook V, page 82*.

Overexpression and purification of His_6-ElxA

Electrocompetent *E. coli* Rosetta2(DE3) cells were transformed with pET.His_6-ElxA and a single colony was inoculated in 5 mL of LB medium containing 50 µg/mL kanamycin and 12.5 µg/mL chloramphenicol and grown for 12 h with shaking. An aliquot of 1 mL was used to inoculate 100 mL of LB medium containing the same antibiotics followed by overnight incubation at 37 °C. Finally, an aliquot of 20 mL of the overnight

culture was used to inoculate 2 L cultures. The cells were grown at 37 °C until OD_{600} = 1.0. IPTG was added to a final concentration of 1.0 mM and the cultures were shaken for 16 h at 18 °C. The cells were harvested by centrifugation and resuspended in lysis buffer (20 mM NaH_2PO_4, 500 mM NaCl, 0.5 mM imidazole, 20% glycerol, pH 7.5). After sonication and centrifugation, the pellets were resuspended in denaturing buffer (6 M guanidine hydrochloride, 20 mM NaH_2PO_4, 0.5 mM imidazole, 500 mM NaCl, pH 7.5) and sonicated and centrifuged again. The supernatants were loaded onto a HisTrap HP 5 mL column prepacked with Ni Sepharose™ (GE Healthcare) and washed with buffer (4 M guanidine hydrochloride, 20 mM NaH_2PO_4, 30 mM imidazole, 300 mM NaCl, pH 7.5). The peptides were then eluted with elution buffer (20 mM Tris-HCl, 100 mM NaCl, 1 M imidazole, 4 M guanidine hydrochloride, pH 7.5). Finally, the peptides were purified by C_4 semi preparative reverse phase high performance liquid chromatography (HPLC) using an Agilent 1200 instrument equipped with a Delta-Pak C4 column (25 mm i.d. x 100 mm L, Waters) and a variable wavelength detector set at 220 nm. The mobile phase was 0.1% trifluoroacetic acid (TFA) in water (A) and 0.086% TFA in 80% acetonitrile / 20% water (B). A gradient of 2-100% B in A over 60 min and a flow rate of 8 mL/min were used. The masses of the purified peptides were determined by matrix-assisted laser desorption/ionization – time-of-flight mass spectrometry (MALDI-TOF MS) on a Voyager DE-STR Biospectrometry Workstation using α-cyano-4-hydroxycinnamic acid as matrix at the Mass Spectrometry Laboratory of the University of Illinois at Urbana Champaign. See also *Notebook V, page 19*.

Overexpression and purification of His_6-MBP-ElxP

Electrocompetent *E. coli* Rosetta2(DE3) cells were transformed with pET.His_6-MBP-ElxP and a single colony was resuspended in 100 µL of LB and plated on an LB agar plate containing 50 µg/mL kanamycin and 12.5 µg/mL chloramphenicol and grown overnight at 37 °C. The colonies were scraped-off and resuspended in 2 mL of LB medium and used to inoculate a 2 L culture. The culture was incubated at 37 °C with shaking until OD_{600} = 1.0. IPTG was added to a final concentration of 0.1 mM and the culture was shaken for an additional 16 h at 18 °C. The cells were harvested by centrifugation and the cell pellet was resuspended in 20 mL of lysis buffer (50 mM Na_2HPO4, 500 mM NaCl, 20 mM imidazole, 5 mg/mL lysozyme, pH 7.5). After incubation at 4 °C for 45 min, the cells were lysed using French-press. The suspension was centrifuged for 45 min at 16,000 × g and 4 °C and the supernatant was filtered through a 0.45 µm syringe-tip filter (Millipore). The protein was purified by IMAC using an ÄKTApurifier (Amersham Biosciences, GE Healthcare) equipped with a HisTrap HP 5 mL column prepacked with Ni Sepharose^{TM} (GE Healthcare). After loading and washing the column with 15 CV of buffer A (20 mM Na_2HPO_4, 500 mM NaCl, 20 mM imidazole, pH 7.4), a gradient of 0-100% buffer B (20 mM Na_2HPO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4) in buffer A over 20 CV was used to elute the protein. The fractions containing His_6-MBP-ElxP, as confirmed by SDS-PAGE analysis, were concentrated using an Amicon Ultracel 50k filter (Millipore). Imidazole was removed by using a PD-10 desalting column (GE Healthcare) and the protein was eluted from the column using an elution buffer containing glycerol (50 mM HEPES, 300 mM NaCl, 20% glycerol, pH 7.5). Aliquots were frozen in liquid N_2 and stored at −80 °C. An aliquot

containing His_6-MBP-ElxP was incubated with TEV protease (0.1 mg TEV/mg protein) in the presence of TCEP (1 mM) at 4 °C overnight. The suspension was incubated with Ni-NTA agarose beads (Qiagen) for 30 min with shaking at 4 °C to remove the hexahistidine tagged MBP-containing peptides. The supernatant was recovered and aliquots were frozen in liquid N_2 and stored at −80 °C. See also *Notebook VIII, page 19*.

Synthesis of substrate analogues and products

The peptides Pyr-AAIVK, D-Lac-AAIKV, L-Lac-AAIVK, ASIVK, Pyr-AAIVKBBIKA, and AAIVKBBIKA were synthesized by Fmoc-based solid phase peptide synthesis (SPPS) using a PS3 peptide synthesizer (Protein Technologies). Fmoc groups were removed during the deprotection steps with 20% piperidine in dimethylformamide (DMF). Coupling of the amino acids was performed using DMF as solvent and 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU) and 0.4 M N-methylmorpholine (NMM) as activating reagents. Coupling reactions of pyruvic acid or L/D-lactic acid were performed using hydroxybenzotriazole (HOBt) and diisopropylcarbodiimide (DIC) as activating reagents. Coupling of unprotected Ser in ASIVK was performed using 3-(diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(3H)-one (DEPBT) and N,N'-diisopropylethylamine (DIPEA) as activating reagents. ADhaIVK was produced after solid phase dehydration of Ser with large excess of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and CuCl as described elsewhere (*9*). D-Lac-ADhaIKV and L-Lac-ADhaIVK were produced after coupling of L- or D-lactic acid to ADhaIVK as described above. Peptides were cleaved from the resins using a mixture of TFA/water/phenol (90:5:5) or TFA/triisopropylsilane (95:5) in the case of D/L-Lac-

ADhaIKV. The cleavage solutions were evaporated using a rotary evaporator and the peptides were precipitated from the solution with cold diethyl ether. The peptide was purified by C_{18} semi preparative reverse phase HPLC using an Agilent 1200 instrument equipped with an Eclipse XDB-C18 column (9.4 mm i.d. x 250 mm L, Agilent). A gradient of 2-30% B (acetonitrile) in A (0.1% formic acid) in 30 min was used. The masses of the purified peptides were determined by electrospray ionization – mass spectrometry (ESI-MS) using a Waters ZMD quadrupole instrument at the Mass Spectrometry Laboratory of the University of Illinois at Urbana-Champaign. See also *Notebook VI, page 41*.

His_6-ElxO and ElxP activity assays

His_6-ElxO (10 μM) and the peptide Pyr-AAIVK (1 mM) generated by Fmoc-based SPPS were incubated with NADPH (1 mM) in assay buffer (100 mM HEPES, 500 mM NaCl, pH 7.5) at room temperature for 6 h. Reaction progress was monitored by UV-vis spectrophotometry, measuring the disappearance of the NADPH peak at 340 nm. Formation of reduced peptide was confirmed by liquid chromatography mass spectrometry (LC-MS) using an Agilent 1200 instrument equipped with a single quadrupole multimode ESI/APCI ion source mass spectrometry detector (Agilent) and a Synergi Fusion-RP column (4.6 mm i.d. × 150 mm L, Phenomenex). See also *Notebook V, page 93*.

ElxP or His_6-MBP-ElxP (5 μM) and purified His_6-ElxA (*10*) (50 μM) were incubated in the presence of assay buffer (50 mM HEPES, pH 7.5) at room temperature for 2–3 h. A control sample lacking enzyme was incubated under identical conditions.

Cleavage of the peptide in the reaction sample was confirmed by MALDI-TOF MS on a Voyager DE-STR Biospectrometry Workstation as described above. See also *Notebook VIII, page 21*.

Production and purification of epilancin 15X

A medium composed of Lab-Lemco meat extract (10%, Oxoid), malt extract (3%, BD), ammonium chloride (20 mM), $Ca(OH)_2$ (0.4%), and NaCl (2%) was inoculated with an overnight pre-culture of *S. epidermidis* 15X154 in LB broth (1/100 dilution). Cells were incubated at 37 °C with shaking for 12 h and harvested by centrifugation. The supernatant was filtered through a 0.22 μm pore size filter and heated at 80 °C for 1 h to deactivate proteases. Solid $(NH_4)_2SO_4$ was added to the culture supernatant to reach 80% saturation at 4 °C and stirred for 4 h followed by centrifugation. The pellet was resuspended in water and loaded onto a Vydac® C_4 reverse phase solid phase extraction column (SPE) (214SPE1000, Discovery Sciences). The column was washed with 25% acetonitrile in ammonium acetate buffer (20 mM, pH 5.0) to remove impurities and the peptide was eluted with 0.1% TFA in 80% methanol / 20% water. The lantibiotic was further purified by HPLC using an Agilent 1200 instrument equipped with a Vydac® 214TP54 C_4 reverse phase column (4.6 mm i.d. × 250 mm L, Discovery Sciences). A gradient of 50-60% B (0.1% TFA in methanol) in A (0.1% TFA in water) over 50 min was used. The fractions corresponding to the major peak were collected and analyzed by MALDI-TOF MS on a Voyager DE-STR Biospectrometry Workstation as described above. See also *Notebook VI, page 68*.

Determination of stereochemical configuration of N-terminal Lac

After incubation of Pyr-AAIVK (1 mM) with ElxO (10 μM) and NADPH (1 mM) in reaction buffer (100 mM HEPES, 500 mM NaCl, pH 7.5), the reaction mixture and the synthetic peptides D-Lac-AAIKV and L-Lac-AAIVK were analyzed individually or combined by HPLC using an Agilent 1200 instrument equipped with a Synergi Fusion-RP column (4.6 mm i.d. × 150 mm L, Phenomenex). A gradient of 0-70% B (methanol) in A (0.1% formic acid in water) over 30 min was used and absorbance at 210 nm was monitored. See also *Notebook VI, page 51.*

Epilancin 15X (30 μM) was incubated with Trypsin (5 μM, Sigma) in buffer (50 mM HEPES, pH 7.5) at room temperature for 3 h. The resulting proteolyzed peptide and the synthetic peptides D-Lac-ADhaIKV and L-Lac-ADhaIVK were analyzed individually or combined by LC-MS using a Waters SYNAPT™ mass spectrometry system equipped with an ACQUITY UPLC®, an ESI ion source, a quadrupole time-of-flight detector (Waters) scanning a range 570 - 573 Da, and an ACQUITY Bridged Ethyl Hybrid (BEH) C18 column (2.1 mm i.d. × 50 mm L, 1.7 μm, Waters). A gradient of 9-12% B (0.1% formic acid in acetonitrile) in A (0.1% formic acid in water) over 15 min was used. See also *Notebook VIII, page 43*.

Cloning and coexpression of *elxA*, *elxB*, and *elxC*

The gene *elxA* was amplified by PCR from *S. epidermidis* 15X154 genomic DNA using a forward primer containing a *BamH*I restriction site (elxA.BamHI.F) and a reverse primer containing a *Not*I site (elxA.NotI.R). The PCR product and the vector pRSFDuet-1 (Novagen) were digested with restriction endonucleases *BamH*I and *Not*I (New

England Biolabs) and ligated using T4 DNA ligase (New England Biolabs) to produce the plasmid pRSF.His_6-ElxA encoding ElxA fused at its N-terminus to a His_6-tag. A sequence-optimized synthetic *elxB* gene (Geneart) was amplified by PCR using a forward primer containing an *Bgl*II restriction site (elxB.BglII.F) and a reverse primer containing an *Xho*I restriction site (elxB.XhoI.R). The PCR product and the vector pRSF.His_6-ElxA were digested with restriction endonucleases *Bgl*II and *Xho*I (Invitrogen) and ligated using T4 DNA ligase (New England Biolabs) to produce the plasmid pRSF.His_6-ElxAB encoding for ElxA fused at its N-terminus to a hexahistidine tag and for ElxB. See also *Notebook VIII, page 23*.

The gene *elxC* was amplified by PCR from *S. epidermidis* 15X154 genomic DNA using a forward primer containing an *Nde*I restriction site (elxC.NdeI.F) and a reverse primer containing a *Xho*I site (elxC.XhoI.R). The PCR product and the vector pACYCDuet-1 (Novagen) were digested with restriction endonucleases *Nde*I and *Xho*I (Invitrogen) and ligated using T4 DNA ligase (New England Biolabs) to produce the plasmid pACYC.ElxC encoding for ElxC.

Electrocompetent *E. coli* BL21(DE3) cells were transformed with pRSF.His_6-ElxAB or cotransformed with pRSF.His_6-ElxAB and pACYC.ElxC. Single colonies were inoculated in 5 mL of LB medium containing the appropriate antibiotics (50 µg/mL kanamycin or 12.5 µg/mL chloramphenicol) and grown for 12 h with shaking. Aliquots of 2 mL were used to inoculate 200 mL of LB medium containing the same antibiotics followed by shaken at 37 °C until OD_{600} = 0.7. IPTG was added to a final concentration of 0.5 mM and the cultures were shaken for 16 h at 18 °C. The cells were harvested by centrifugation and resuspended in lysis buffer (20 mM NaH_2PO_4, 500 mM NaCl, 20 mM

imidazole, 2 mg/mL lysozyme, pH 7.4). After sonication and centrifugation, the supernatants were loaded onto HisTrap HP 5 mL columns prepacked with Ni Sepharose™ (GE Healthcare) and washed with buffer (20 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole, pH 7.4). The peptides were then eluted with elution buffer (20 mM NaH_2PO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4) and loaded onto Vydac® C_4 reverse phase SPE columns (214SPE1000, Discovery Sciences). The columns were washed with 0.1% TFA in 5% methanol / 95% water to remove impurities and the peptides were eluted with 0.1% TFA in 80% methanol / 20% water and 0.1% TFA in 80% acetonitrile / 20% water. After lyophilization, the peptides were resuspended in buffer (50 mM HEPES, pH 7.5) and incubated with His_6-MBP-ElxP at room temperature overnight to cleave the leader peptide. Ni-NTA agarose beads (Qiagen) were added and the supernatant was recovered and analyzed by MALDI-TOF MS on a Voyager DE-STR Biospectrometry Workstation as described above. See also *Notebook VIII, page 25*.

Cloning of *nisA*(leader)-*elxA*(core) and coexpression with *nisB* and *nisC*

The fragment of the gene *elxA* encoding for the core peptide was amplified by PCR from *S. epidermidis* 15X154 genomic DNA using a forward (nis-elxA.R-1E.F) and a reverse (nis-elxA.R-1E.R) primer and iProof™ high-fidelity DNA polymerase (Bio-Rad). A mutation for cleavage with GluC at the N-terminus of the encoded core peptide was also introduced in the forward primer. The PCR product contained annealing regions to the pRSF.His_6-NisAB plasmid (*8*), which encodes for hexahistidine tagged NisA and for untagged NisB, allowing replacement of the *nisA* core region for *elxA* by

PCR amplification of the entire plasmid. After treatment with *Dpn*I (New England Biolabs) and transformation of *E. coli* DH5α cells, the plasmid pRSF.His_6-NisElxA.NisB was generated. This plasmid encodes for the hexahistidine tagged chimera NisA(leader)-ElxA(core) R-1E precursor peptide and for NisB. See also *Notebook IX, page 05*.

Chemically competent *E. coli* T7 Express or electrocompetent BL21(DE3) cells were transformed with pRSF.His_6-NisElxA.NisB, or cotransformed with pRSF.His_6-NisElxA.NisB and pACYC.NisC (*8*), pACYC.NisBC, or pCDF.NisB in different combinations. Single colonies were inoculated in 5 mL of LB medium containing the appropriated antibiotics (50 µg/mL kanamycin, 12.5 µg/mL chloramphenicol, or 25 µg/mL streptomycin) and grown for 12 h at 37 °C with shaking. Aliquots of 2.5 mL were used to inoculate 250 mL of LB medium containing the same antibiotics followed by incubation at 37 °C until OD_{600} = 0.6. IPTG was added to a final concentration of 0.5 mM and the cultures were shaken for 16 h at 18 °C. The cells were harvested by centrifugation and resuspended in lysis buffer (20 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole, 2 mg/mL lysozyme, pH 7.4). After sonication and centrifugation, the supernatants were loaded onto HisTrap HP 5 mL columns prepacked with Ni Sepharose™ (GE Healthcare) and washed with buffer (20 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole, pH 7.4). The peptides were then eluted with elution buffer (20 mM NaH_2PO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4) and loaded onto Vydac® C_4 reverse phase SPE columns (214SPE1000, Discovery Sciences). The columns were washed with 0.1% TFA in 5% methanol / 95% water to remove impurities and the peptides were eluted with 0.1% TFA in 80% methanol / 20% water and 0.1% TFA in

80% acetonitrile / 20% water. After lyophilization, the peptides were resuspended in buffer (50 mM HEPES, pH 7.5) and incubated with the endopeptidase GluC (2 ng/μL, New England Biolabs) at room temperature overnight to cleave the leader peptide. The products were analyzed by MALDI-TOF MS using a Bruker Daltonics UltrafleXtreme MALDI TOF/TOF mass spectrometer at the Mass Spectrometry Laboratory of the University of Illinois at Urbana-Champaign. See also *Notebook IX, page 06*.

2.3. RESULTS

Cloning and sequencing of the epilancin 15X biosynthetic gene cluster

To identify and sequence the epilancin 15X biosynthetic gene cluster, a fosmid library of *S. epidermidis* 15X154 genomic DNA was constructed in *E. coli.* The fosmid library was screened by PCR using degenerate primers to amplify a fragment of *elxA* (the gene encoding the precursor peptide) and *elxC* (the gene encoding a lanthionine cyclase). The primers for *elxA* were designed based on the amino acid sequence of epilancin 15X by Dr. Xingang Zhang, a previous postdoctoral researcher at the van der Donk laboratory, while primers for *elxC* were designed based on the conserved amino acid sequences in the cyclase enzymes. Two positive clones containing non-overlapping DNA fragments were isolated and the fosmids were sequenced using transposon insertions. Specific primers annealing with regions of *S. epidermidis* 15X154 genomic DNA were then used to amplify by PCR a bridging 1.2 kb DNA fragment that was sequenced to obtain the biosynthetic gene cluster (Figure 2.1). The open reading frames (ORFs) were analyzed with the Basic Local Alignment Search Tool (BLAST) (*11*) and six ORFs encoding putative proteins with high sequence identity to enzymes

involved in the production and transport of lantibiotics were identified: *elxA*, *elxB*, *elxC*, *elxP*, *elxT,* and *elxO* (Table 2.3 and Figure 2.1). Additionally, three genes with no homology to characterized lantibiotic genes and presumably involved in immunity were identified and designated as *elxI1*, *elxI2,* and *elxI3*. With the exception of *elxO*, the role of the other genes in epilancin 15X biosynthesis can be predicted based on homology to previously characterized lantibiotic genes. The gene encoding the precursor peptide, *elxA*, encodes a serine residue at the first position of the core peptide, which corresponds to lactate in epilancin 15X. Thus, Ser1 must be posttranslationally modified by an undetermined enzyme.

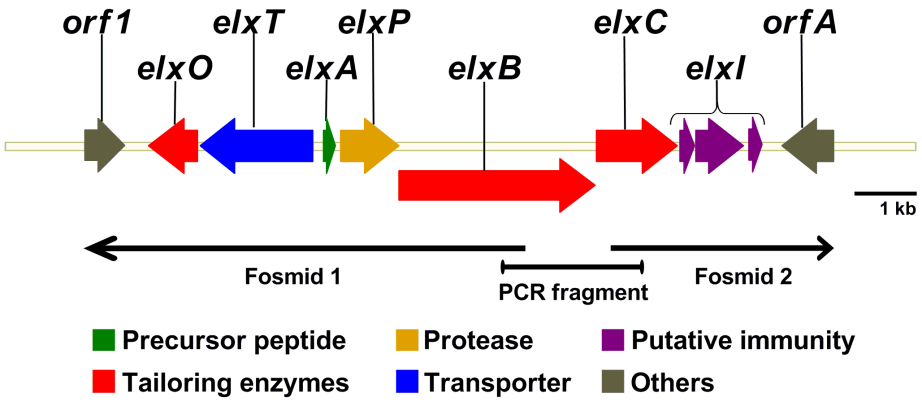


Figure 2.1. Epilancin 15X biosynthetic gene cluster.

Table 2.3. Open reading frames analysis of the epilancin 15X gene cluster using BLASTp at the NCBI website.

Predicted ORF	Number of AA[1]	Protein homology[2]	Identity[3]	Expect[4]
elxA	55	ElkA, epilancin K7 precursor peptide, *S. epidermidis* K17, (AAA79236) (55 aa)	38/55 (69%)	7e-15
elxB	986	PepB, Pep5 dehydratase, *S. epidermidis* 5, (CAA90025) (967 aa)	327/994 (32%)	7e-104
elxC	402	PepC, Pep5 cyclase, *S. epidermidis* 5, (CAA90026) (398 aa)	142/379 (37%)	8e-56
elxO	248	EciO, oxidoreductase, *S. epidermidis* BN280, (CAA74346) (247 aa)	(126/248) 50%	8e-64
elxP	297	EciP, epicidin 280 protease, *S. epidermidis* 5, (CAA74349) (300 aa)	123/286 (43%)	3e-48
elxT	573	PepT, Pep5 ABC transporter, *S. epidermidis* 5, (CAA90021) (571 aa)	354/571 (61%)	0.0
elxI1	72	Hypothetical protein SE2390, *S. epidermidis* ATCC12228, (NP_765945) (76 aa)	43/65 (66%)	7e-18
elxI2	241	CAAX amino protease, *S. epidermidis* M23864:W1, (ZP_04817536) (248 aa)	70/178 (39%)	8e-24
elxI3	71	Hypothetical protein, *S. aureus* subsp. *aureus* USA300_TCH959, (ZP_04865952) (75 aa)	43/71 (60%)	6e-17
orf1	177	Recombinase, *S. aureus* subsp. *aureus* TCH70, (ACZ58811) (182 aa)	154/176 (87%)	4e-85
orfA	261	Membrane spanning protein, *S. hominis* SK119, (ZP_04060547) (257 aa)	249/257 (96%)	1e-110

[1]AA: amino acid. [2]Results are from a BLASTp search of the GenBank protein database on January 2010. [3]Identities in the aligned region. [4]Expectation value.

Cloning and overexpression of *elxO*

To study the role of ElxO, the corresponding gene was cloned into a pET28b vector to generate pET.His_6-ElxO that encodes an N-terminal hexahistidine fusion protein (His_6-ElxO). His_6-ElxO was heterologously produced in *E. coli* Rosetta2(DE3) cells and the enzyme was purified by immobilized metal ion affinity chromatography (IMAC) with Ni^{2+}, resulting in 60 to 90 mg of purified protein per liter of cell culture. The enzyme migrated as a protein of approximately 30 kDa by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis, close to the predicted monomeric molecular weight of His_6-ElxO (29.7 kDa). Native molecular weight analysis

using gel filtration chromatography suggests that His_6-ElxO exists as a dimer (59 kDa, observed).

In vitro reconstitution of the enzymatic activity of His_6-ElxO and determination of the stereochemistry of the reaction

On the basis of its amino acid sequence, ElxO is a member of the short-chain dehydrogenase/reductase (SDR) superfamily that catalyzes the interconversion of alcohols to aldehydes or ketones using NAD(P)(H) as a cofactor. Dehydroepilancin 15X, a peptide containing an N-terminal Pyr group, could therefore be the substrate for ElxO. Since dehydroepilancin 15X was not available, the small peptide AAIVK was synthesized by Fmoc-based SPPS followed by coupling of pyruvic acid to produce the ketone containing substrate Pyr-AAIVK. This peptide resembles the N-terminal portion of dehydroepilancin 15X, with Dha at position 3 replaced with Ala for simplicity. Incubation of Pyr-AAIVK with ElxO resulted in a decrease in absorbance at 340 nm over time with NADPH but not NADH. The reaction sample was also analyzed by LC-MS and a new peak with slightly shorter retention time and with a mass of 572.4 Da, corresponding to Lac-AAIVK, was observed (Figure 2.2).

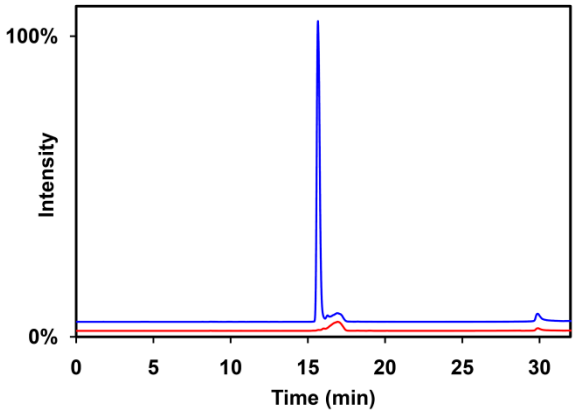


Figure 2.2. Enzymatic assay of His_6-ElxO with Pyr-AAIVK. Single ion chromatograms at *m/z* = 573.4, corresponding to the expected $[M+H]^+$ ion for Lac-AAIVK, of the peptide incubated with (blue) or without (red) His_6-ElxO are shown. The peak at 16 min observed in the reaction sample but not in the control sample confirms the enzymatic reduction of the peptide. The peak at 17 min in the control sample corresponds to higher molecular weight isotopes of the substrate (monoisotopic *m/z* = 571.4).

To determine the stereochemical configuration of the N-terminal Lac, the two possible reaction products (D-Lac-AAIVK and L-Lac-AAIVK) were synthesized by Fmoc-based SPPS using D- or L-lactic acid during the last coupling step. The enzymatic product of Pyr-AAIVK, after incubation with His_6-ElxO and NADPH, was combined with D-Lac-AAIVK or L-Lac-AAIVK and analyzed by HPLC (Figure 2.3). The enzymatic product of His_6-ElxO co-eluted with D-Lac-AAIVK, but not with L-Lac-AAIVK, demonstrating that ElxO catalyzes the formation of an N-terminal D-lactate ((*R*)-2-hydroxypropionate).

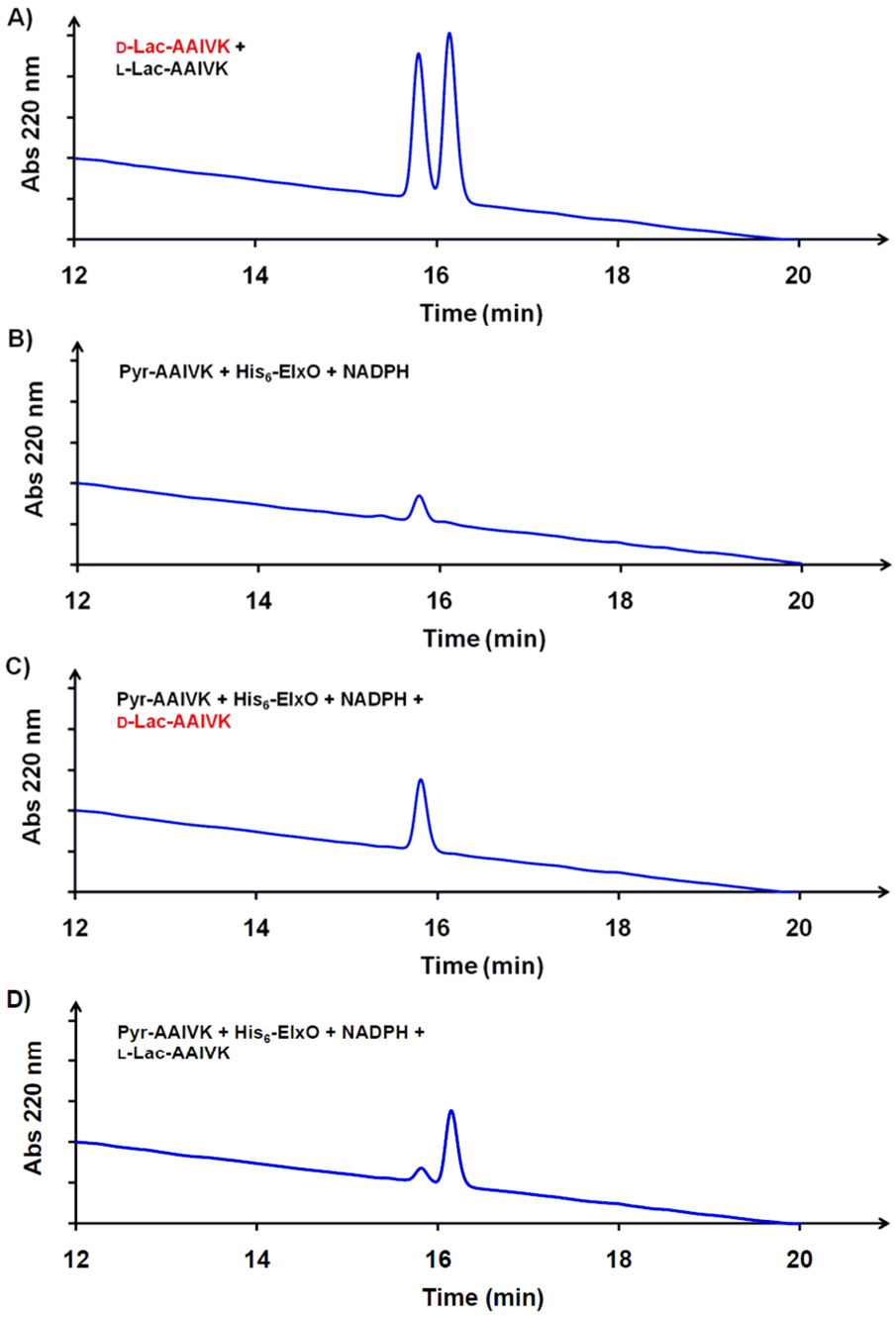
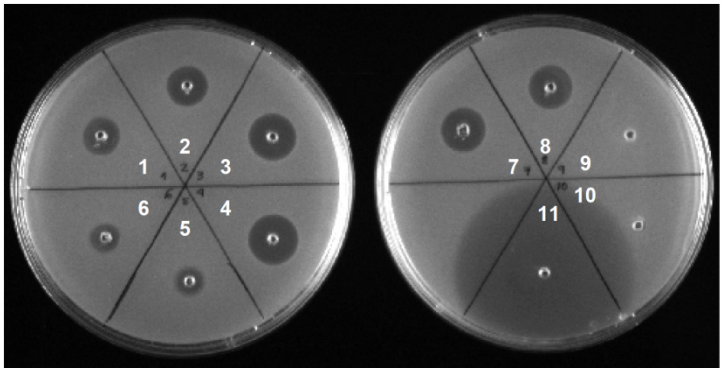


Figure 2.3. Determination of the stereochemical configuration of the enzymatically synthesized N-terminal lactate (Lac). The two possible diastereoisomers of Lac-AAIKV have different retention times when analyzed by reverse-phase HPLC. A) A mixture of chemically synthesized D-Lac-AAIVK and L-Lac-AAIVK separated by HPLC. B) The reaction product of Pyr-AAIVK after incubation with His_6-ElxO and NADPH analyzed under the same conditions as A). C) A mixture of the enzymatic reaction product of Pyr-AAIVK and D-Lac-AAIVK produced a single peak. D) The enzymatic product of Pyr-AAIVK combined with L-Lac-AAIVK produced two well separated peaks.

Production of epilancin 15X and stereochemical characterization of its N-terminal Lac group

The production of lanthionine-containing polypeptide antibiotics by staphylococci is highly dependent on the composition of the media (*12, 13*). Media containing Lab-Lemco meat extract, malt extract, ammonia, $Ca(OH)_2$, and NaCl have been shown to support the production of Pep5, gallidermin, and epidermin, whereas small concentrations of phosphate and glucose can repress bacteriocin production (*12, 13*). To identify the optimal conditions for the production of epilancin 15X, a set of cultures, in which the concentrations of meat extract, NaCl, and NH_4Cl were systematically varied, were grown and epilancin production was compared with the original conditions (*7*). Analysis of culture supernatants for bioactivity using an agar diffusion assay with *Staphylococcus carnosus* TM300 as indicator strain (Figure 2.4) demonstrated that a medium containing 10% Lab-Lemco meat extract, 2% NaCl, 20 mM NH_4Cl, 3% malt extract, and 0.4% $Ca(OH)_2$ produced the highest concentration of epilancin 15X under the conditions tested. Purification yielded about 3.0 mg of bacteriocin per liter of culture, compared with a yield of 0.5 mg per liter reported previously (*7*). The identity of epilancin 15X was determined by MALDI-TOF MS, confirming the presence of only the lantibiotic and one possible oxidation product or hydrated dehydroepilancin 15X (Figure 2.4).

A)



	Sample							
Component	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Meat extract, %	10	10	10	**10**	2	2	2	2
NaCl, %	10	10	2	**2**	10	10	2	2
NH_4Cl, mM	200	20	200	**20**	200	20	200	20

B)

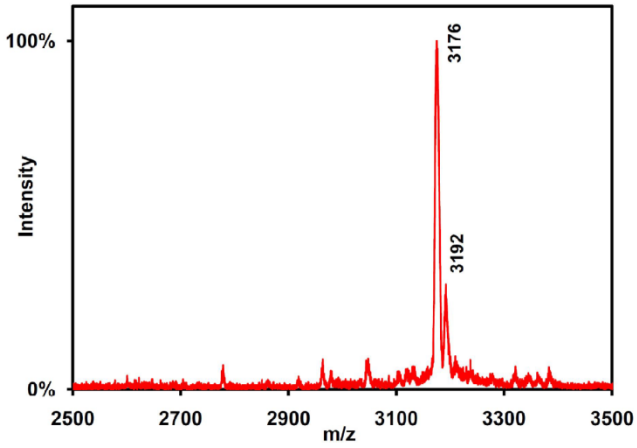


Figure 2.4. Production of epilancin 15X. A) *S. epidermidis* 15X was grown in media containing different concentrations of meat extract, NaCl, and NH_4Cl. The supernatants were concentrated (5×) and tested against *S. carnosus* TM300 for bioactivity. A medium containing 10% meat extract, 2% NaCl, 20 mM NH_4Cl, 3% malt extract, and 0.4% $Ca(OH)_2$ produced the higher concentration of epilancin 15X (sample 4). No significant production was observed when Mueller Hinton broth II or nutrient broth were used (samples 9 and 10, respectively). Ampicillin: sample 11 (positive control). B) MALDI-TOF MS spectrum of epilancin 15X after HPLC purification. The peaks at *m/z* = 3176 corresponds to the $[M+H]^+$ ion of epilancin 15X (expected average mass 3175 Da). The peak at *m/z* = 3192 may correspond to an oxidation product of epilancin 15X or hydrated epilancin 15X.

To determine the stereochemical configuration of the N-terminal Lac in epilancin 15X, a sample of purified lantibiotic was treated with trypsin generating Lac-ADhaIVK,

among other peptide fragments. The resulting peptide mixture and synthetic samples of D-Lac-ADhaIVK and L-Lac-ADhaIVK, produced by SPPS, were analyzed by LC-MS (Figure 2.5). The N-terminal proteolytic fragment of epilancin 15X co-eluted with D-Lac-ADhaIVK, confirming the stereochemical configuration of the N-terminal Lac.

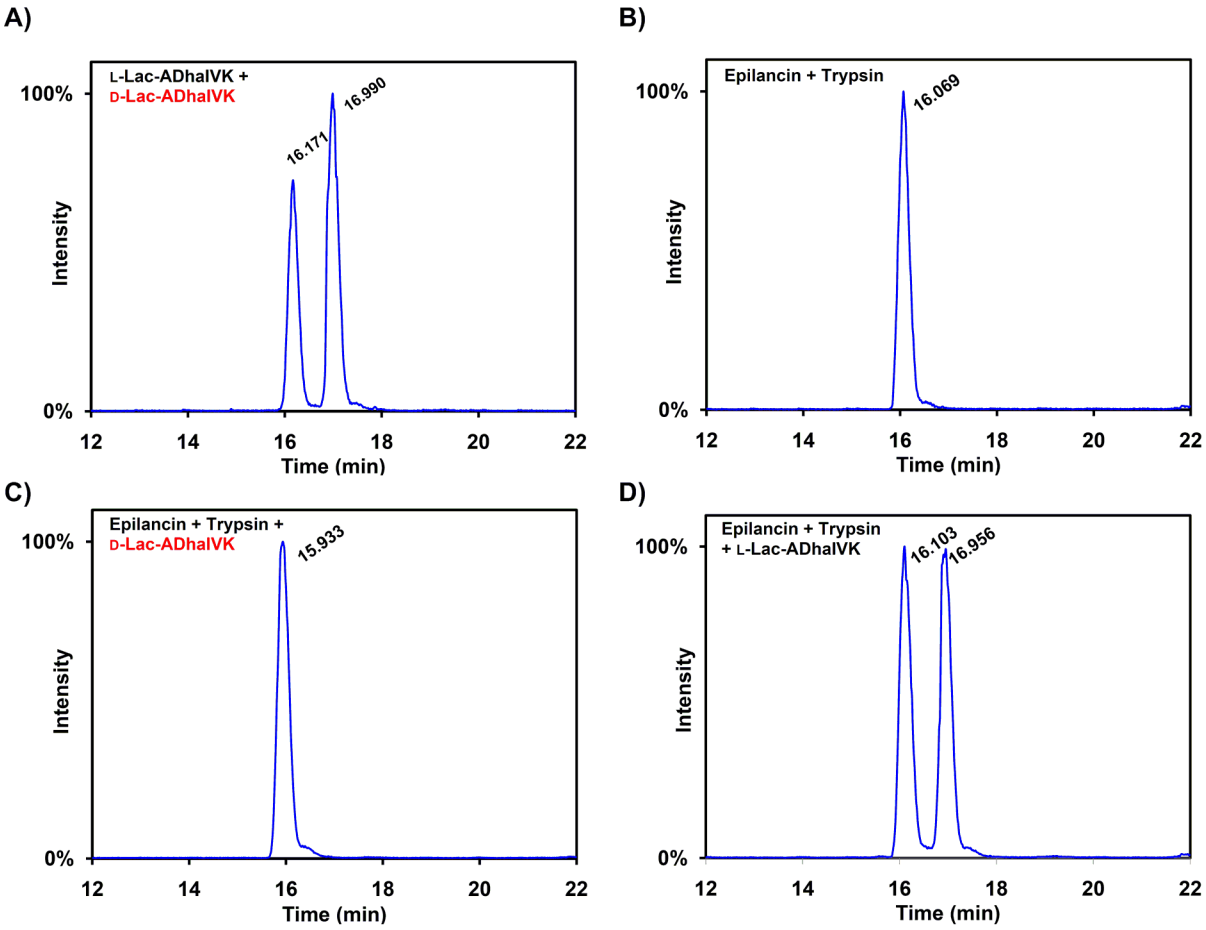


Figure 2.5. Determination of the stereochemical configuration of the N-terminal lactate (Lac) in epilancin 15X. The peptides were detected using a quadrupole time-of-flight MS detector scanning the range 570 - 573 Da. A) A mixture of chemically synthesized D-Lac-ADhaIVK and L-Lac-ADhaIVK separated by UPLC. B) The reaction product of epilancin 15X after digestion with trypsin analyzed under the same conditions as A). C) A mixture of the enzymatic digestion product of epilancin 15X and D-Lac-AAIVK produced a single peak. D) The enzymatic digestion product of epilancin 15X combined with L-Lac-AAIVK produced two well-separated peaks.

Cloning and overexpression of *elxP*

The gene *elxP* was cloned initially into a pET28b vector to generate an N-terminal hexahistidine fusion of ElxP (His_6-ElxP). However, attempts to overexpress the protein in *E. coli* were unsuccessful, since the rate of growth of the host was greatly reduced after induction with IPTG, suggesting that ElxP is toxic to the heterologous host. To overcome toxicity problems and improve solubility, *elxP* was cloned into pET.His_6-MBP-ElxP that encodes for a fusion protein containing an N-terminal pelB signal, followed by a hexahistidine tag for purification, and a maltose binding protein (MBP) tag for solubility, separated from ElxP by a Tobacco Etch Virus (TEV) protease cleavage site (Figure 2.6). Heterologous expression trials in *E. coli* Rosetta2(DE3) by using the plating method (*14*) afforded successful production of His_6-MBP-ElxP. The enzyme was purified by IMAC with Ni^{2+}, resulting in about 9 mg of purified protein per liter of cell culture. After treatment of His_6-MBP-ElxP with TEV protease and SDS-PAGE analysis, protein bands at about 34 kDa and 45 kDa, corresponding to ElxP (predicted mass 34.3 kDa) and His_6-MBP (predicted mass of 45.5 kDa), respectively, were observed.

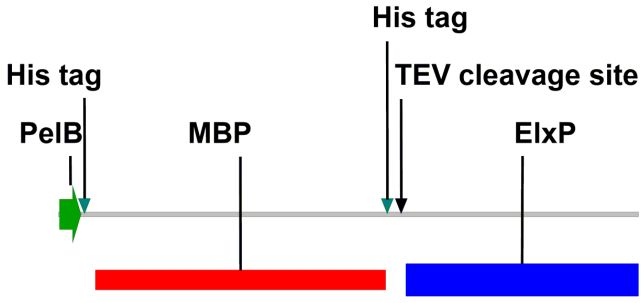


Figure 2.6. Scheme of ElxP fused to a TEV protease cleavage site, a MBP tag, a hexahistidine tag, and a pelB signal peptide encoded by pET.His_6-MBP-ElxP.

In vitro reconstitution of the enzymatic activity of ElxP

Based on its amino acid sequence, ElxP is a serine protease that may cleave the leader peptide from the fully cyclized precursor peptide. Since modified ElxA was not available, linear His_6-ElxA was tested as substrate. The peptide was obtained by expression from a pET28b vector in *E. coli* Rosetta2(DE3) cells, followed by purification by IMAC with Ni^{2+} and subsequent HPLC. Linear His_6-ElxA was incubated with ElxP or His_6-MBP-ElxP at pH 7.5 and the reaction products were analyzed by MALDI-TOF MS, confirming the formation of two proteolytic products in both cases (Figure 2.7). The peak at *m/z* = 4862 corresponds to the N-terminal hexahistidine tagged leader peptide (residues −1 to −43) resulting from amide bond hydrolysis at the predicted Gln−1/ Ser1 cleavage site, whereas the peak at *m/z* = 3313 corresponds to the C-terminal unmodified core peptide.

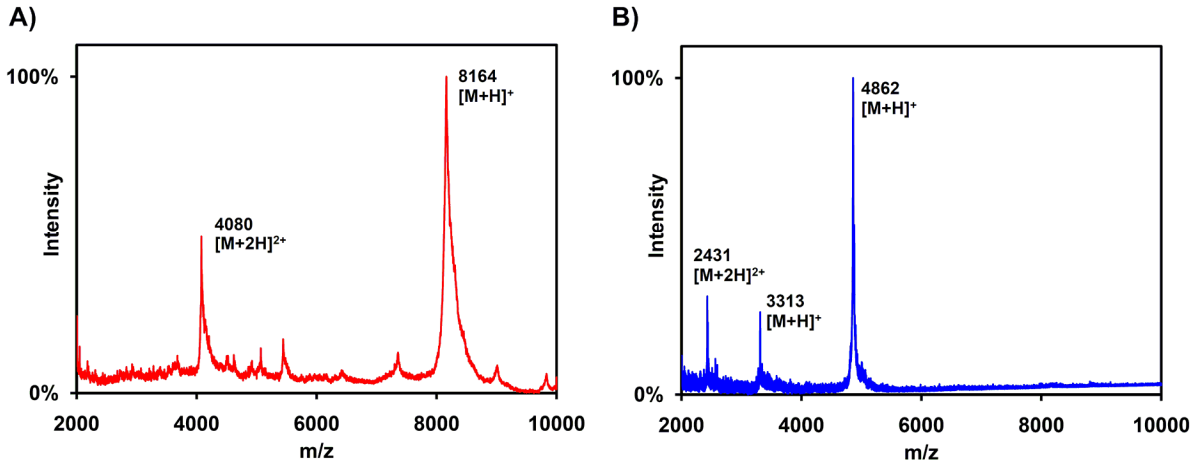


Figure 2.7. MALDI-TOF MS analysis of the products after digestion of His_6-ElxA by His_6-MBP-ElxP. Similar results were obtained when ElxP was used. A) Analysis of His_6-ElxA (calculated average mass = 8162 Da) after incubation without enzyme in reaction buffer. B) Analysis of peptide mixture after incubation with enzyme under the same conditions. Peaks corresponding to the leader peptide (calculated average mass = 4864 Da) and to the core peptide (calculated average mass = 3316 Da) were observed.

Cloning and coexpression of *elxA, elxB, and elxC* in a heterologous host

Based on bioinformatic analysis (Table 1), ElxB likely catalyzes the dehydration of the Ser and Thr residues in the precursor peptide ElxA, while ElxC catalyzes the regioselective addition of the thiol groups from the Cys residues into the dehydrated amino acids, generating the Lan/MeLan rings. To confirm the role of ElxB in epilancin 15X biosynthesis, the genes *elxA* and a synthetic codon optimized version of *elxB* were cloned into a pRSFDuet-1 vector, as a biscistronic construct as previously described for nisin (*8*). The generated plasmid pRSF.His_6-ElxAB encodes for an N-terminal hexahistidine fusion of ElxA (His_6-ElxA) allowing rapid purification and for untagged ElxB (Figure 2.8A). Additionally, Ms. Isabel Neacato, a graduate student at the van der Donk laboratory, designed and prepared a pACYCDuet-1 derivative plasmid encoding for ElxC (Figure 2.8A). Upon coexpression of *elxA* and *elxB* in *E. coli* BL21(DE3), followed by peptide purification and cleavage of the leader region with His_6-MBP-ElxP, a mixture of peptides potentially containing up to three out of eight expected dehydrations was observed by MALDI-TOF MS (Figure 2.8B). Coexpression of the same genes and *elxC*, followed by a similar treatment, allowed the production of peptides with up to five dehydrations (Figure 2.8C). Attempts to obtain fully dehydrated peptide or to reconstitute the enzymatic activity of ElxB *in vitro* were not successful.

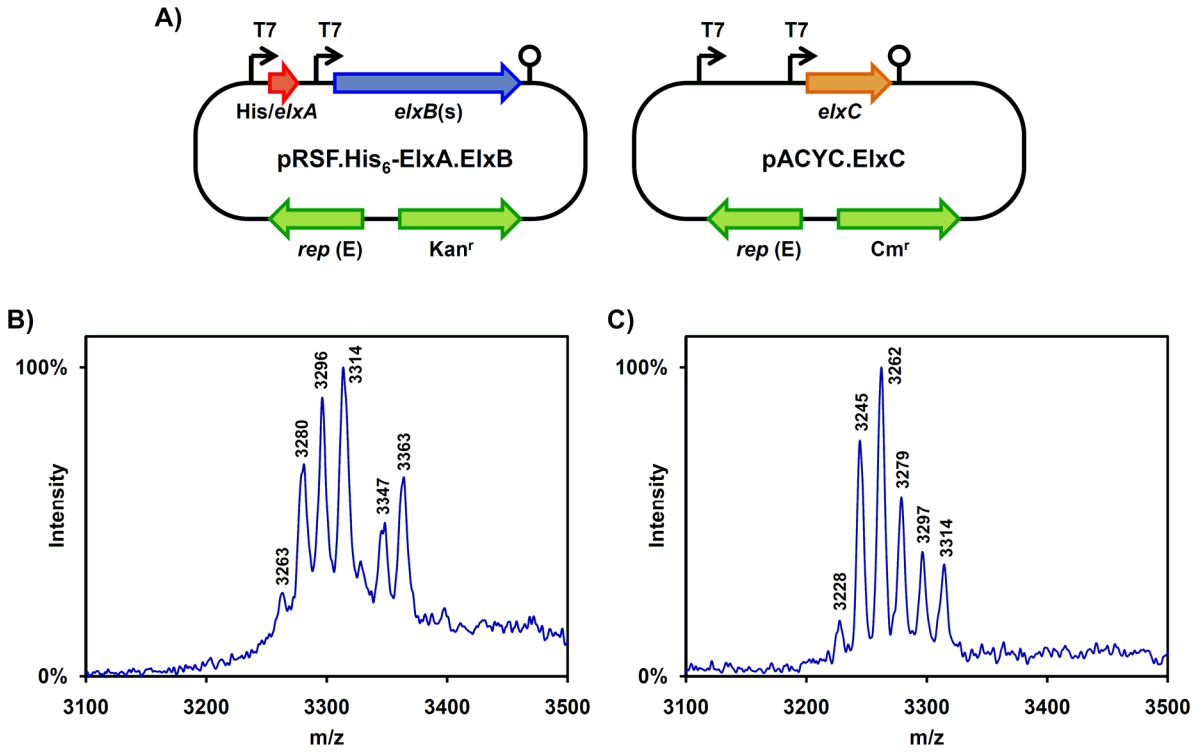


Figure 2.8. Coexpression of *elxA*, *elxB*, and *elxC*. The genes *elxA*, *elxB*, and *elxC* were cloned into compatible vectors for coexpression in *E. coli* (A). The MALDI-TOF MS spectra of the mixtures of peptides obtained after coexpression of *elxA* and *elxB* (B), or *elxA*, *elxB*, and *elxC* (C) in *E. coli* BL21(DE3) and cleavage of leader peptide with His_6-MBP-ElxP are shown. Peaks likely corresponding to the unmodified core peptide (calculated average *m/z* = 3317) or to oxidized single dehydrated peptide (calculated average *m/z* = 3315) and to 1× to 5× dehydrated peptides (calculated average *m/z* = 3299, 3281, 3263, 3245, and 3227) were observed. The peaks at *m/z* = 3347 and 3363 might correspond to oxidation products.

Cloning and coexpression of the chimera *nisA*(leader)-*elxA(core), nisB, and nisC* in a heterologous host

The nisin biosynthetic machinery has been successfully used to modify peptides other than nisin that are fused to the leader region of NisA using a *Lactococcus lactis* expression system (*15-19*). To evaluate if a similar strategy can be used to produce epilancin 15X, using the recently developed nisin A production system with *E. coli* as heterologous host (*8*), a chimeric gene encoding for the NisA leader peptide fused to the ElxA core peptide was cloned into a pRSFDuet-1 derivative plasmid that also

encodes for NisB (*8*), generating pRSF.His_6-NisElxA.NisB (Figure 2.9A). Furthermore, an N-terminal hexahistidine tag was encoded at the N-terminus of the chimeric NisA-ElxA peptide for rapid purification and an artificial GluC protease cleavage site was introduced between the leader and the core region to allow *in vitro* cleavage of the leader sequence after peptide purification. Upon coexpression of *nisA-elxA* and *nisB* in *E. coli* T7 express, a mixture of peptides, potentially containing up to six out of eight expected dehydrations, was observed by MALDI-TOF MS (Figure 2.9B). Additional assays cotransforming pRSF.His_6-NisElxA.NisB with the plasmids pACYC.NisC (encoding for NisC, Figures 2.9A and C), pACYC.NisBC (encoding for NisB and NisC, Figures 2.9A and D), or pACYC.NisBC and pCDF.NisB (containing an additional copy of *nisB*, Figures 2.9A and E), followed by coexpression in *E. coli* T7 express, were also performed generating peptides with up to seven dehydrations. Interestingly, an increase in the extent of dehydration was observed with the introduction of multiple copies of *nisB* based on the relative intensity of the peaks in the MALDI-TOF MS spectra. Thus, the nisin A biosynthetic machinery can be used to modify peptides different than NisA *in vivo* using *E. coli* as a host. Further optimization is required to produce fully dehydrated ElxA and dehydroepilancin 15X.

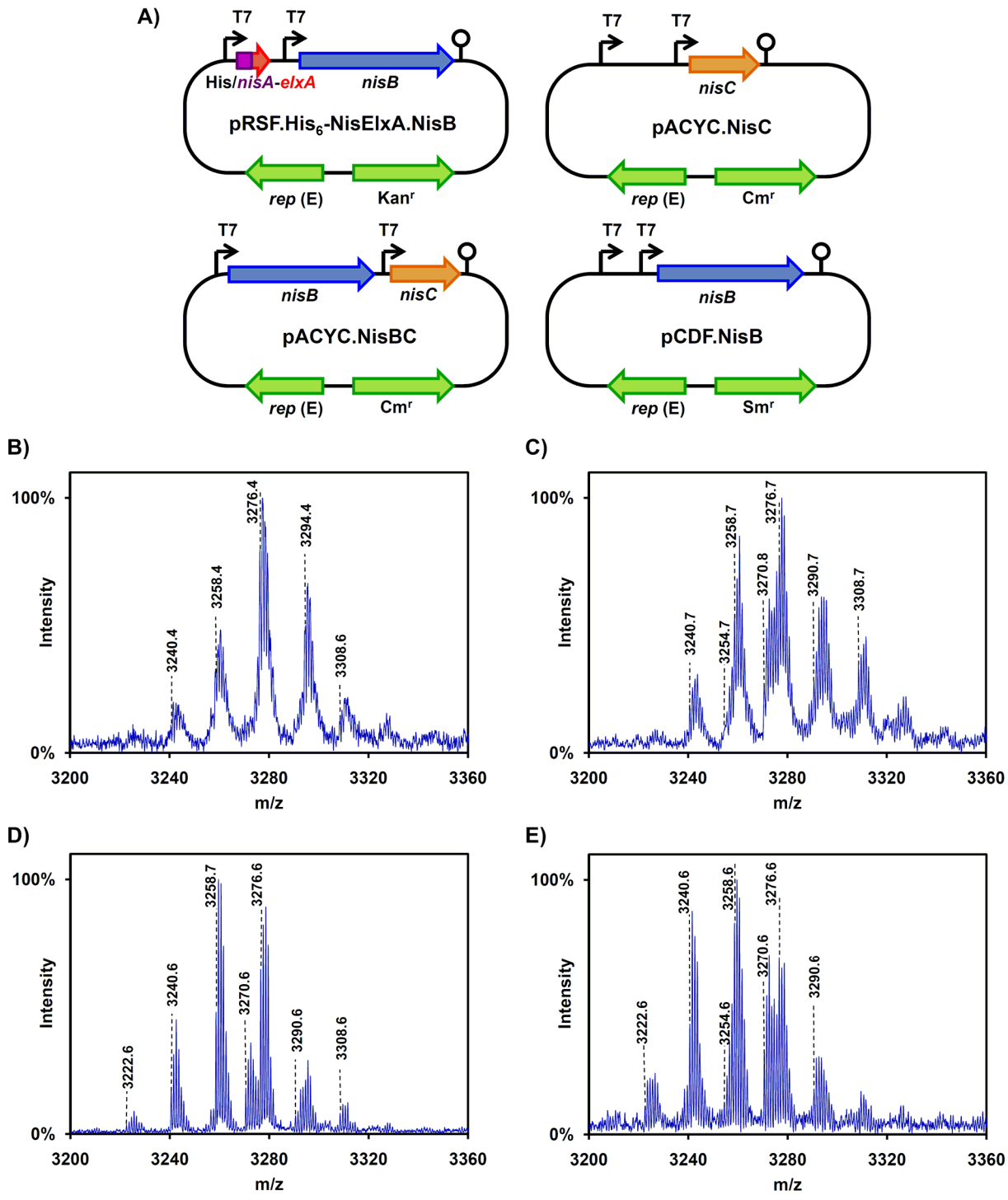


Figure 2.9. Coexpression of *nisA*(leader)-*elxA*(core), *nisB*, and *nisC*.

Figure 2.9. *Continuation*. A chimeric gene encoding for the NisA leader peptide fused to ElxA core sequence was cloned together with *nisB* to generate a biscistronic construct (A). The genes *nisB and nisC* were also cloned individually or together into compatible vectors for coexpression in *E. coli* (A). The MALDI-TOF MS spectra of the mixtures of peptides obtained after coexpression of *nisA*-*elxA* and *nisB* (B); *nisA*-*elxA*, *nisB*, and *nisC* (C); *nisA*-*elxA*, *nisB* (two copies), and *nisC* (D); and *nisA*-*elxA*, *nisB* (three copies), and *nisC* (E) in *E. coli* T7 express and cleavage of leader peptide with GluC are shown. Peaks corresponding to 3× to 7× dehydrated peptides (calculated monoisotopic *m/z* = 3293.8, 3275.8, 3257.8, 3239.8, and 3221.8) were observed. The observed masses correspond to the peptides with addition of methanol to N-terminal Pyr (M+MeOH) or double oxidized peptides (M+2O). Additional peaks might correspond to different oxidation products. A copy of *nisC* and multiple copies of *nisB* seem to favor dehydration.

2.4. DISCUSSION

As discussed in chapter 1, lantibiotics are produced from ribosomally synthesized linear precursor peptides that consist of an N-terminal leader region and a C-terminal core peptide. The mature lantibiotic is generated from the core peptide after several posttranslational modifications. The reactions involved in the formation of the characteristic Lan/MeLan rings formation have been investigated (*20*-*22*), but only a few enzymes introducing other posttranslational modifications have been studied (*23, 24*). This chapter focuses on the mechanism of formation of N-terminal lactate groups and on the enzymatic cleavage of the leader peptide.

The epilancin 15X gene cluster contains five genes involved in biosynthesis (*elxABCOP*), one gene involved in the export of the mature peptide (*elxT*), and three genes potentially involved in immunity (*elxI1, elxI2,* and *elxI3*) (Table 1 and Figure 2). The cluster organization resembles that of the lantibiotics Pep5 (*3*) and epicidin 280 (*2*) produced by different strains of *S. epidermidis*, suggesting that these clusters have evolved from a common ancestor. The predicted peptide ElxA has high amino acid sequence similarity to the epilancin K7 precursor peptide ElkA (*25*) (Figure 2.10). ElxA contains an N-terminal leader sequence of 24 amino acids and a C-terminal core

peptide region comprised of 31 amino acids, as predicted from the chemical structure of epilancin 15X. The leader region also contains the conserved motif F-(D/N)-L-(D/E/N) and a Pro at position −2 that are characteristic of class I lantibiotics (Figure 2.10) (*26*).

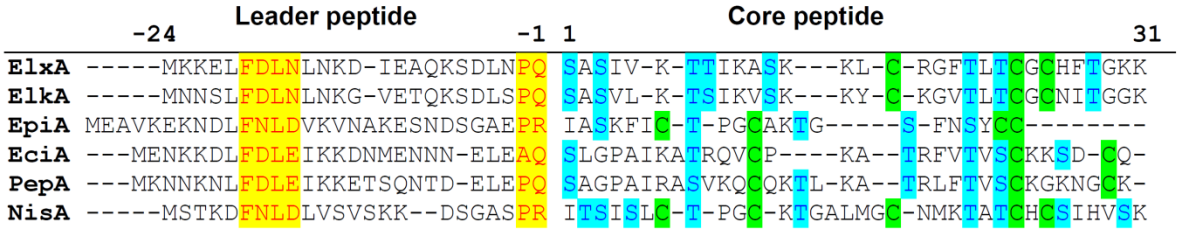


Figure 2.10. Sequence alignment of ElxA with related lantibiotic precursor peptides. ElxA: epilancin 15X precursor peptide, ElkA: epilancin K7 precursor peptide, EpiA: epidermin precursor peptide, EciA: epicidin 280 precursor peptide, PepA: Pep5 precursor peptide, NisA: nisin A precursor peptide.

Downstream of *elxA*, an ORF designated as *elxP* was identified. The encoded protein ElxP possesses high amino acid sequence similarity to EciP, the protease involved in the biosynthesis of epicidin 280 (*2*), and to other subtilisin-like serine proteases. Importantly, the residues of the predicted catalytic triad and oxyanion hole (Asp27, His62, Ser240, and Asn154) are conserved in ElxP. The lack of an N-terminal *sec*-signal sequence, a pro- sequence, and a C-terminal cell wall anchor sequence (LPXTG) suggests that ElxP is localized inside the cytoplasm, possibly as part of a membrane-bound biosynthetic complex. Thus, ElxP likely removes the leader peptide before the mature peptide is transported outside the cell, in contrast to other serine-type lantibiotic proteases, such as NisP or EpiP, that are located extracellularly and remove the leader region once the peptide has been secreted (see chapter 4) (*27, 28*).

In a previous attempt to study lantibiotic proteases, an *E. coli* host carrying a plasmid encoding for NisP was able to express the protease at low concentrations

based on SDS-PAGE analysis using ^{35}S-Met (*27*). Although NisP was not purified, the *E. coli* cell extracts cleaved the nisin A cyclized precursor peptide, producing a biologically active compound (*27*). Additional *in vivo* studies indicated that NisP is able to cleave the leader region only from fully processed precursor peptide, but not from uncyclized dehydrated or unmodified NisA (*29*). Similar results were obtained from *in vivo* studies of the lantibiotic Pep5 (*3*). In contrast, culture supernatants of *S. carnosus* TM300 expressing EpiP processed unmodified EpiA to the expected proteolytic products (*28*). In the present study, ElxP was successfully expressed in *E. coli* and its enzymatic activity was reconstituted *in vitro* after enzyme purification (Figure 2.7). The protease was able to process unmodified His_6-ElxA, indicating that neither the Lan/MeLan ring nor the dehydrated residues, including Dha at position 1, are strictly required for enzyme recognition and proteolytic processing.

Downstream of *elxP*, an ORF designated as *elxB* encodes a protein with homology to PepB, the enzyme that catalyzes the dehydration of Ser or Thr residues in the Pep5 precursor peptide (*3*). Analysis of ElxB with SignalP 3.0 using neuronal networks trained with Gram-positive bacteria (*30*) suggests that this protein may contain an N-terminal cell membrane anchor signal (Figure 2.11). The activity of ElxB homologs of class I lantibiotics has never been reconstituted *in vitro* (*31*) and the cofactors or metals involved in catalysis are currently unknown. In this work, I was also unable to reconstitute ElxB activity *in vitro*. However, coexpression of His_6-ElxA and ElxB with or without ElxC in *E. coli* produced a partially dehydrated peptide (Figure 2.8) confirming the role of the protein in the biosynthesis of epilancin 15X. Closer inspection of the ElxB sequence indicates that it contains an almost conserved Walker A motif (GXXXXGKT/S:

GLLENWKT) and a conserved Walker B motif (hhhhD: IIFPD, where h stands for hydrophobic residue) (Figure 2.12). In addition, the three potential binding sites characteristic of GTP binding proteins are also present (DXXG: DFLG, NKXG: NTID/NDID/NLND/NRND, SAX: SAT) (*32*) suggesting that GTP or another nucleotide may be required for dehydration by LanB proteins, similar to class II-IV lanthionine synthetases (*20, 22, 33, 34*). The ORF *elxC* encodes a protein with high amino acid sequence similarity to PepC, the cyclase responsible for Lan and MeLan ring formation in Pep5 (*3*). Additionally, ElxC contains the conserved residues comprising a zinc ion binding site (Cys269, Cys318, and His319) and the residues involved in acid-base catalysis (His205 and Asp142) that are characteristic of this family of proteins (*21, 35*).

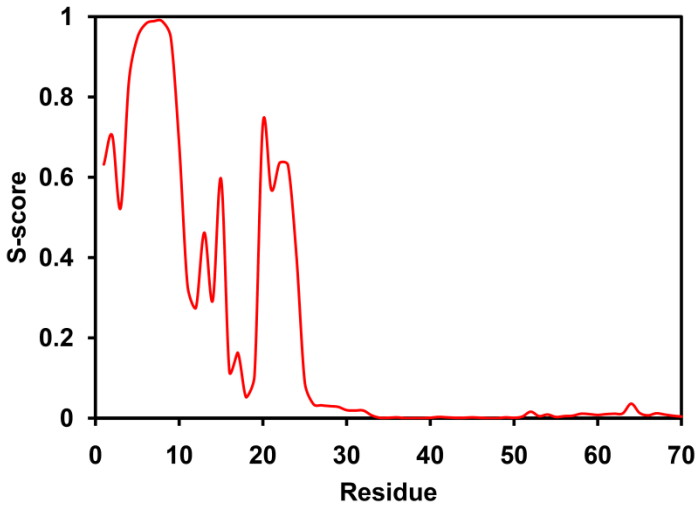


Figure 2.11. Analysis of ElxB with SignalP 3.0. High values of the S-score indicate that the corresponding amino acid is part of a signal peptide (*30*).

Figure 2.12. Potential binding motifs and protein domain analysis of ElxB. Lantibiotic dehydratase - N terminus (pfam04737, pink) and C terminus (pfam04738, blue) domains and an N-terminal signal peptide (red) are present in ElxB. Potential Walker A and B motifs (yellow) and GTP binding sites (green) are almost perfectly conserved. h, stands for hydrophilic residue.

Based on the experimental results and on bioinformatic analysis, the first steps of epilancin 15X biosynthesis can be postulated (Figure 2.13). The precursor peptide ElxA is modified by the dehydratase ElxB and the cyclase ElxC to produce the crosslinked

peptide. Then, the leader peptide is removed by the protease ElxP producing an N-terminal enamine (dehydroalanine, Dha) present in equilibrium with the corresponding imine that can be hydrolyzed to produce dehydroepilancin 15X. Although enamine hydrolysis is fast (*36*), enzymatic assistance (*e.g.* by ElxP) cannot be ruled out at present. The reduction of the N-terminal ketone to the respective alcohol will complete the synthesis of mature epilancin 15X.

The ORF *elxO* encodes a protein with homology to EciO, an oxidoreductase hypothesized to be involved in the reduction of pyruvate to lactate in the biosynthesis of epicidin 280 (*2*). ElxO contains a predicted N-terminal NAD(P)(H) binding site (GXXGXG: GGFKGIGK) and the catalytic triad residues (Ser139, Tyr152, and Lys156) of the short-chain dehydrogenase/reductase (SDR) protein superfamily (Figure 2.14) (*37, 38*). The absence of the 'proximal Asp residue' (Ser33 in ElxO) and the presence of the 'proximal basic residues' (Lys12 and Arg34), responsible for cofactor specificity, correctly predicted that ElxO is an NADPH-dependent enzyme and that it belongs to the cP3 subfamily (*38, 39*). The stereochemical course of the reaction established here allows assignment of the configuration of the N-terminal lactate group of epilancin 15X as (*R*). Because His_6-ElxO was able to reduce a hexamer peptide, the thioether rings or other structural motifs in dehydroepilancin 15X are not essential for enzyme recognition.

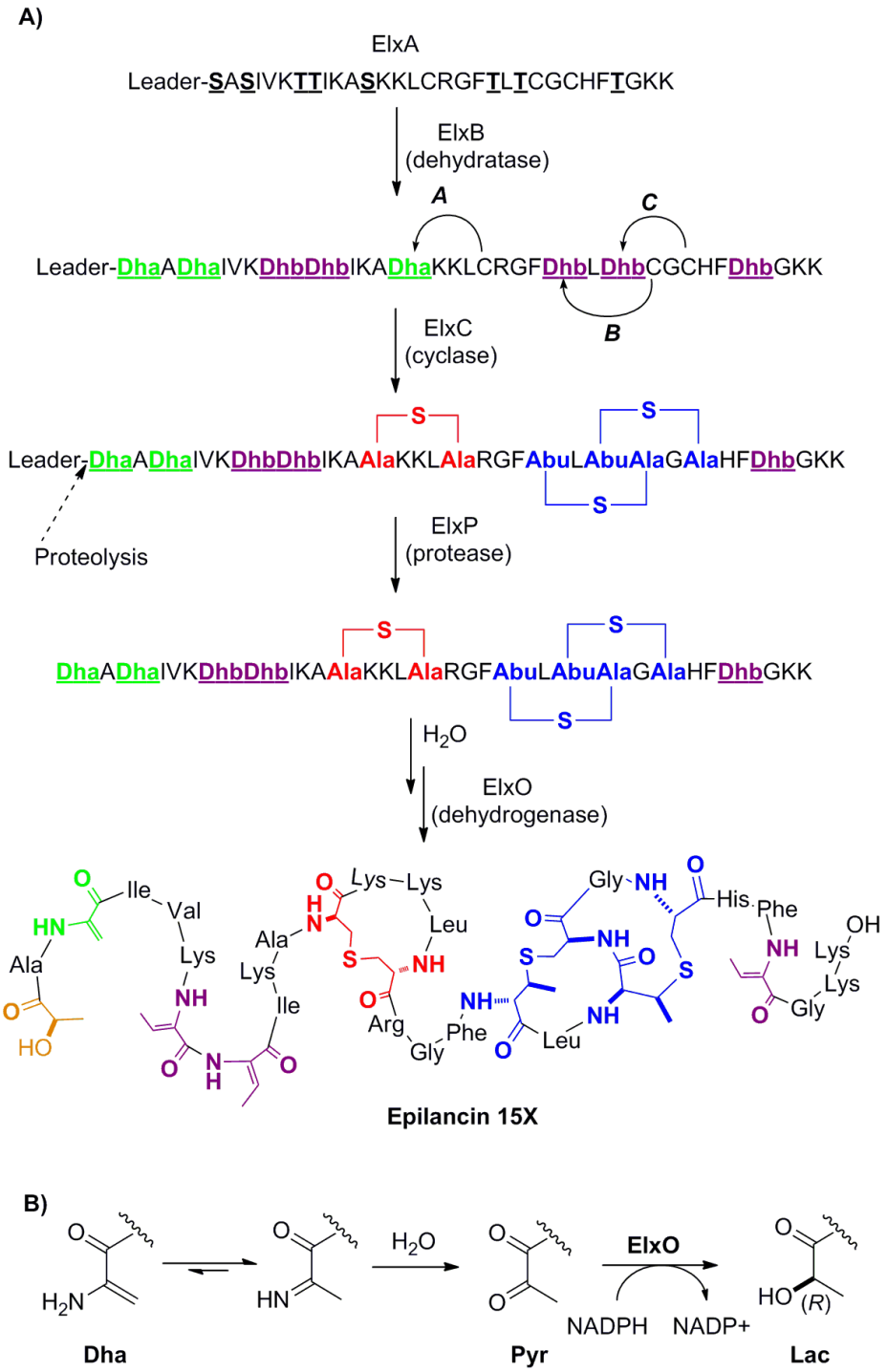


Figure 2.13. Proposed biosynthetic steps involved in the production of epilancin 15X (A) and its N-terminal D-lactate group (B).

The N-terminal Lac is not a common modification in lantibiotics. Indeed, besides epilancin 15X, only three other lantibiotics, epilancin K7 (*25, 40*), SWLP1 (*41*), and

epicidin 280 (*2*), have been reported to contain this modification. Interestingly, genome mining for LanC homologs allowed the identification of a lantipeptide gene cluster containing an open reading frame encoding for a putative SDR protein in the genome of *Streptomyces griseoflavus* Tu4000 (SgriT_010100000930) (*42*). Although the predicted lantipeptide has not been isolated and the function of the protein is unknown, this finding suggests that N-terminal Lac groups may not be restricted to staphylococcal lantibiotics.

The ORF *elxT* encodes a putative protein with homology to PepT (*3*). The C-terminal domain of ElxT contains an ATP-binding site characterized by the conserved Walker A motif (GXXXXGKT/S: GPSGAGKT) or P-loop, the Walker B motif (hhhhD: ILLLD), and the C motif or "signature" motif (LSGGQ) specific to ABC transporters (Figure 2.15). Additionally, analysis of ElxT with TMHMM 2.0 (*30*) indicates that its N-terminal portion contains six transmembrane helices (Figure 2.15). Thus, ElxT is likely involved in lantibiotic secretion.

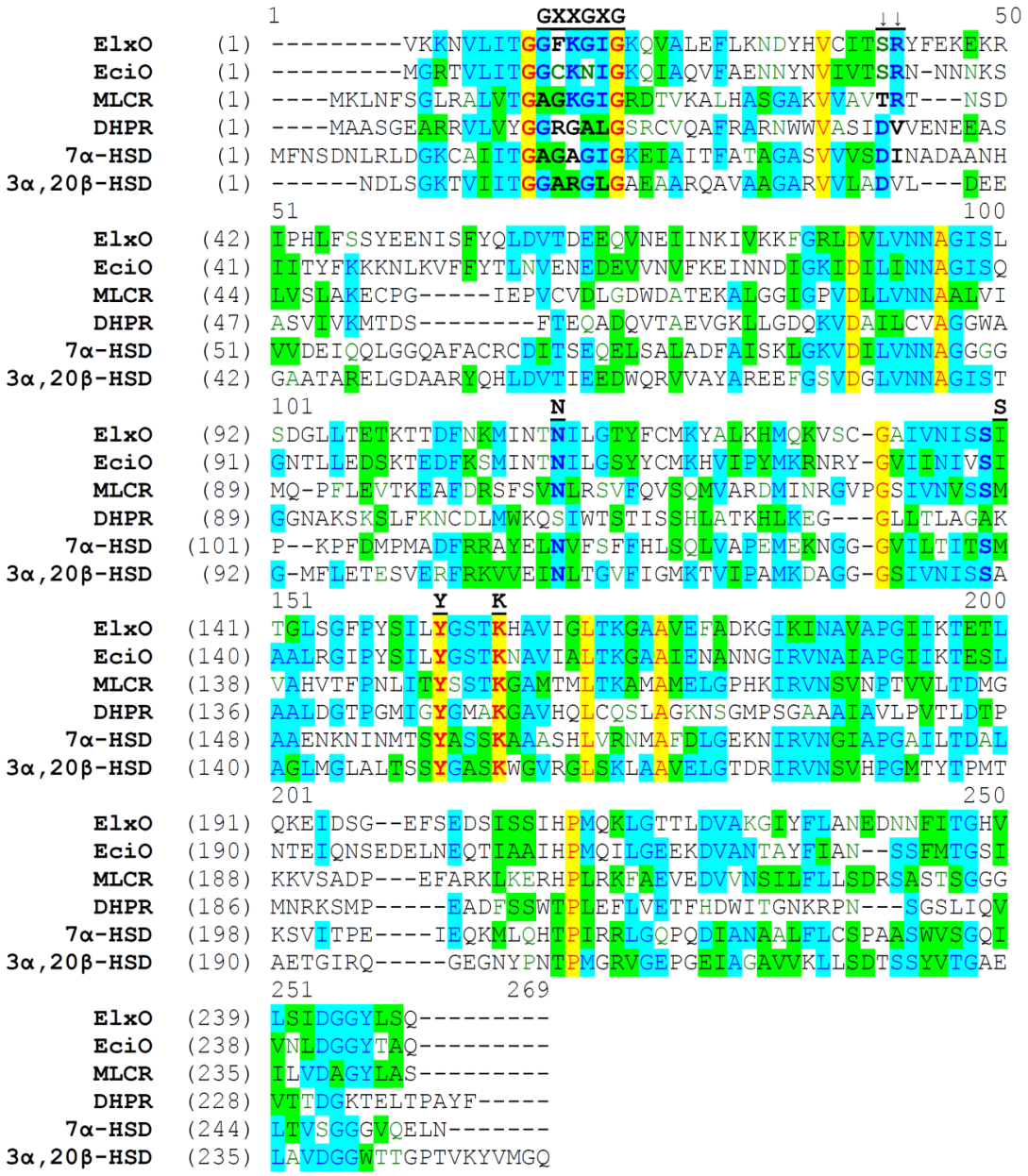


Figure 2.14. Sequence alignment of ElxO and several other short-chain dehydrogenase reductase proteins. The highly conserved residues Asn111, Ser139, Tyr152, and Lys156, the GXXGXG motif, and the proximal Asp residue (Asp42 in 7α-hydroxysteroid dehydrogenase, which is substituted by Ser33 in ElxO) and the 'proximal basic residues' (Arg34 in ElxO) are in bold. EciO: epicidin 280 oxidoreductase, MLCR: Mouse lung carbonyl reductase, DHPR: Dihydropteridine reductase, 7α-HSD: 7α-hydroxysteroid dehydrogenase, 3α,20β-HSD: 3α,20β-hydroxysteroid dehydrogenase.

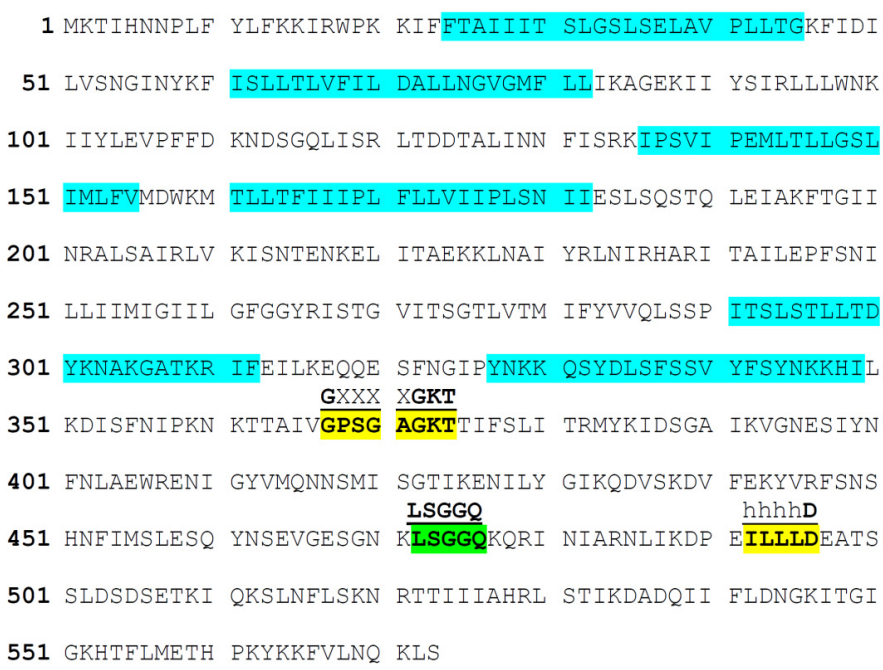


Figure 2.15. Domain analysis of ElxT. The N-terminal transmembrane domains predicted with TMHMM 2.0 (*30*) are highlighted in blue, the Walker A and B motifs in yellow and the C-motif characteristic of ABC transporters in green.

Epilancin 15X has potent activity against staphylococci, including strains of *S. epidermidis* (*43*), suggesting that the producer strain must have an effective self-resistance mechanism. Such immunity is particularly important for epilancin 15X because it is activated by leader peptide removal within the cytoplasm unlike most lantibiotics for which leader peptide cleavage occurs after or concomitant with secretion. Downstream of *elxC*, three ORFs *elxI1*, *elxI2*, and *elxI3* were identified. The genes *elxI1* and *elxI3* potentially encode for 72 aa and 71 aa paralog peptides (47% identity) with no sequence homology to previously characterized proteins. Analysis by TMHMM 2.0 (*30*) indicates that ElxI1 and ElxI3 contain two highly hydrophobic α-helical domains followed by strongly hydrophilic C-terminal segments (Figure 2.16). Thus, although ElxI1 and ElxI3 contain no signal peptides, these small proteins may be localized at the

cytoplasmic membrane. A similar pattern of helical domains followed by a hydrophilic region is found in small proteins encoded by the gene clusters of the closely related lantibiotics Pep5 (*44*) and epicidin 280 (*2*), the structurally unrelated lactosin S (*45*), the nonlantibiotic bacteriocin divergicin A (*46*), and the circular bacteriocins AS-48, acidocin B, butyrivibriocin AR10, and circularin A (*47*). In the case of Pep5, the protein was designated PepI and was shown to be a determinant for self-immunity of the producer strain (*48*). Interestingly, in all of the lantibiotics mentioned above, intracellular peptidases remove the leader peptide and the mature bacteriocin is produced inside the cell (see chapter 4). PepI has been suggested to bind to a (currently unknown) target molecule, avoiding docking of the lantibiotic Pep5 onto the target (*44*). Despite the absence of significant sequence homology between PepI and ElxI1/3, the topological similarity suggests that these peptides may protect the host organism in a similar fashion (Figure 2.16).

Interestingly, epilancin 15X and Pep5 are two of the most highly charged lantibiotics with a net positive charge of +7.1 and +8.0 at pH 7.4, respectively (compared, for instance, to a charge of +3.2 for nisin), suggesting that their targets might be negatively charged. The hydrophilic and positively charged tail of ElxI1/3 may protect this target by electrostatic repulsion of epilancin 15X. In another scenario, mature epilancin 15X may recognize the highly hydrophobic, membrane-like helical domains of ElxI1/3 to form a complex that protects the host against the lethal bioactivity. The entire complex, instead of the mature lantibiotic alone, may be recognized and secreted by ElxT. This potential role for ElxI1/3 is equivalent to that of substrate-binding proteins commonly found in ABC transporter systems (*49*).

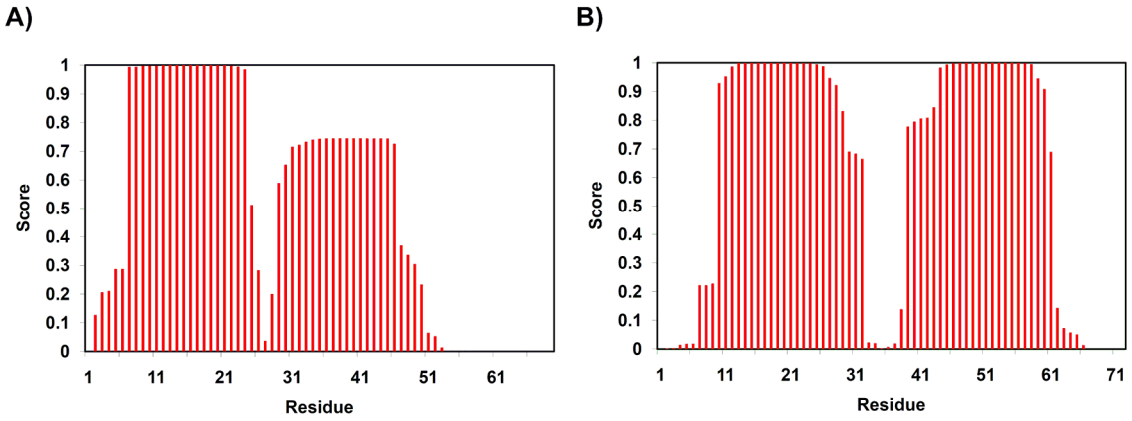


Figure 2.16. Domain analysis of PepI (A) and ElxI1 (B). Both proteins are predicted to contain two highly hydrophobic α-helical domains followed by strongly hydrophilic C-terminal segments. The helical regions were predicted using TMHMM 2.0 (*30*).

Finally, the putative immunity protein ElxI2 has high sequence similarity to Abi proteins, membrane-bound metalloproteases that are involved in self-immunity to plantaricin EF and JK, sakacin 23K, or streptolysin S (*50, 51*). ElxI2 is predicted to contain seven transmembrane domains, including the final four α-helices that form the Abi domain (Figure 2.17). Three highly conserved motifs (EEXXXR: EEILYR , FXXXH: FSLIH, and His226) likely constitute the active site of the protease (*52*). Thus, ElxI2 may protect the host against the bacteriocin by direct degradation of the peptide as in the case of the extracellular metalloprotease gelatinase in *Enterococcus faecalis* (*53*). An ortholog neighborhood analysis revealed that several putative ElxI2 orthologs exist in other staphylococci that are sensitive to epilancin 15X. The encoding ORFs are surrounded by genes likely involved in carbohydrate-phosphotransferase systems (PTS) responsible for the concomitant import and phosphorylation of carbohydrates. The mannose-PTS has previously been shown to serve as target receptor for several class IIa and IIc bacteriocins, including pediocin PA-1, enterocin P, sakacin P, and lactococcin A (*54, 55*). Thus, epilancin 15X might also interact with a PTS membrane

protein triggering pore formation and ElxI2 could be a resistant copy of the target instead of a bacteriocin peptidase. Experimental investigation of these hypotheses may help to elucidate the mode of action of epilancin 15X and the self-protection mechanism.

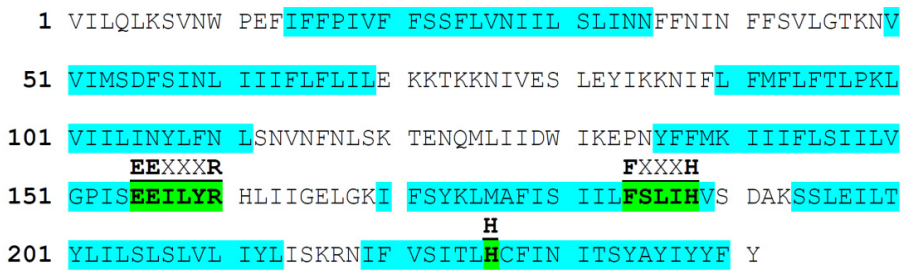


Figure 2.17. Domain analysis of ElxI2. The transmembrane domains are highlighted in blue and the conserved Abi motifs in green.

Whether *elxI1*, *elxI2*, and *elxI3* are part of the cluster and are involved in the immunity mechanism is at present not certain. Closer analysis of the non-coding sequences upstream of *elxC*, *elxI1*, and *elxI2* suggests the presence of only one relevant inverted repeat between *elxC* and *elxI1*, partially overlapping *elxC,* and with a calculated free energy of -10.2 kcal/mol. This repeat may work as a weak rho-independent transcriptional terminator that allows partial read-through, indicating that *elxC* and *elxI1-3* may be part of a single operon and the same gene cluster. Downstream of *elxI3*, a non-coding region of 329 bp is followed by the ORF *orfA* in a different operon that encodes a putative ABC transporter with no sequence homology to any known lantibiotic proteins, but with significant homology to transporters from staphylococci. Thus, OrfA is not likely to be related to epilancin 15X biosynthesis. Flanking the putative epilancin 15X gene cluster on the other side, an ORF designated *orf1* was identified. Orf1 is a putative recombinase, not likely to be involved in epilancin

15X biosynthesis, transport, or immunity. Thus, the epilancin 15X cluster likely spans a 9.2 kb region in the *S. epidermidis* 15X154 genome and includes the genes from *elxO* to *elxI3* (Figure 2.1).

Previously, the nisin biosynthetic machinery was successfully used to produce linear, dehydrated, or cyclized peptides fused to the leader region of NisA using a *L. lactis* expression system (*15-17*). Furthermore, production of fully modified nisinA(leader)/subtilin(core) or vice versa, with either nisin A or subtilin modification machineries was also demonstrated (*18, 56*). In the present study, a chimeric peptide containing the nisin A leader peptide and the epilancin 15X core peptide was dehydrated several fold when coexpressed with NisB, demonstrating that peptides different to NisA can be modified using the nisin tailoring enzymes in *E. coli* as heterologous host. Moreover, inclusion of multiple copies of *nisB* or simultaneous expression of NisC seem to favor the dehydration efficiency, in contrast with previous results in the *L. lactis* system (*16*).

The recently discovered lantibiotic epilancin 15X produced by *S. epidermidis* 15X154 has potent antimicrobial activity against drug-resistant strains of *S. aureus*. Epilancin 15X is structurally simple compared with other lantibiotics and yet is very active. The compound contains an unusual N-terminal D-lactate group that could be essential for biological activity. In this investigation, the gene cluster for epilancin 15X and the biosynthetic route of Lac were determined. Furthermore, the enzymatic activity of the dehydratase, protease, and oxidoreductase involved in the biosynthesis were demonstrated *in vitro* or *in vivo*, opening the door for engineering studies as well as for elucidation of the epilancin 15X mode of action.

2.5. REFERENCES

1. Augustin, J., Rosenstein, R., Wieland, B., Schneider, U., Schnell, N., Engelke, G., Entian, K. D., and Götz, F. (1992) Genetic analysis of epidermin biosynthetic genes and epidermin-negative mutants of *Staphylococcus epidermidis*, *Eur. J. Biochem. 204*, 1149-1154.

2. Heidrich, C., Pag, U., Josten, M., Metzger, J., Jack, R. W., Bierbaum, G., Jung, G., and Sahl, H. G. (1998) Isolation, characterization, and heterologous expression of the novel lantibiotic epicidin 280 and analysis of its biosynthetic gene cluster, *Appl. Environ. Microbiol. 64*, 3140-3146.

3. Meyer, C., Bierbaum, G., Heidrich, C., Reis, M., Suling, J., Iglesias-Wind, M. I., Kempter, C., Molitor, E., and Sahl, H. G. (1995) Nucleotide sequence of the lantibiotic Pep5 biosynthetic gene cluster and functional analysis of PepP and PepC. Evidence for a role of PepC in thioether formation, *Eur. J. Biochem. 232*, 478-489.

4. Blodgett, J. A., Zhang, J. K., and Metcalf, W. W. (2005) Molecular cloning, sequence analysis, and heterologous expression of the phosphinothricin tripeptide biosynthetic gene cluster from *Streptomyces viridochromogenes* DSM 40736, *Antimicrob. Agents Chemother. 49*, 230-240.

5. Eliot, A. C., Griffin, B. M., Thomas, P. M., Johannes, T. W., Kelleher, N. L., Zhao, H., and Metcalf, W. W. (2008) Cloning, expression, and biochemical characterization of *Streptomyces rubellomurinus* genes required for biosynthesis of antimalarial compound FR900098, *Chem. Biol. 15*, 765-770.

6. Grant, S. G., Jessee, J., Bloom, F. R., and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants, *Proc. Natl. Acad. Sci. U. S. A. 87*, 4645-4649.

7. Ekkelenkamp, M. B., Hanssen, M., Danny Hsu, S. T., de Jong, A., Milatovic, D., Verhoef, J., and van Nuland, N. A. (2005) Isolation and structural characterization of epilancin 15X, a novel lantibiotic from a clinical strain of *Staphylococcus epidermidis*, *FEBS Lett. 579*, 1917-1922.

8. Shi, Y., Yang, X., Garg, N., and van der Donk, W. A. (2011) Production of lantipeptides in *Escherichia coli*, *J. Am. Chem. Soc. 133*, 2338-2341.

9. Jiménez, J. C., Bayó, N., Chavarría, B., López-Macià, A., Royo, M., Nicolás, E., Giralt, E., and Albericio, F. (2002) Synthesis of peptides containing alpha, beta-didehydroamino acids. Scope and limitations, *Lett. Pept. Sci. 9*, 135-141.

10. Li, B., Cooper, L. E., and van der Donk, W. A. (2009) Chapter 21. In vitro studies of lantibiotic biosynthesis, *Methods Enzymol. 458*, 533-558.

11. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool, *J. Mol. Biol. 215*, 403-410.

12. Horner, T., Ungermann, V., Zahner, H., Fiedler, H. P., Utz, R., Kellner, R., and Jung, G. (1990) Comparative studies on the fermentative production of lantibiotics by staphylococci, *Appl. Microbiol. Biotechnol. 32*, 511-517.

13. Horner, T., Zahner, H., Kellner, R., and Jung, G. (1989) Fermentation and isolation of epidermin, a lanthionine containing polypeptide antibiotic from *Staphylococcus epidermidis*, *Appl. Microbiol. Biot. 30*, 219-225.

14. Suter-Crazzolara, C., and Unsicker, K. (1995) Improved expression of toxic proteins in *E. coli*, *Biotechniques 19*, 202-204.

15. Kluskens, L. D., Kuipers, A., Rink, R., de Boef, E., Fekken, S., Driessen, A. J., Kuipers, O. P., and Moll, G. N. (2005) Post-translational modification of therapeutic peptides by NisB, the dehydratase of the lantibiotic nisin, *Biochemistry 44*, 12827-12834.

16. Kuipers, A., Meijer-Wierenga, J., Rink, R., Kluskens, L. D., and Moll, G. N. (2008) Mechanistic dissection of the enzyme complexes involved in biosynthesis of lacticin 3147 and nisin, *Appl. Environ. Microbiol. 74*, 6591-6597.

17. Rink, R., Kluskens, L. D., Kuipers, A., Driessen, A. J., Kuipers, O. P., and Moll, G. N. (2007) NisC, the cyclase of the lantibiotic nisin, can catalyze cyclization of designed nonlantibiotic peptides, *Biochemistry 46*, 13179-13189.

18. Kuipers, O. P., Rollema, H. S., de Vos, W. M., and Siezen, R. J. (1993) Biosynthesis and secretion of a precursor of nisin Z by Lactococcus lactis, directed by the leader peptide of the homologous lantibiotic subtilin from Bacillus subtilis, *FEBS Lett. 330*, 23-27.

19. Majchrzykiewicz, J. A., Lubelski, J., Moll, G. N., Kuipers, A., Bijlsma, J. J., Kuipers, O. P., and Rink, R. (2010) Production of a class II two-component lantibiotic of *Streptococcus pneumoniae* using the class I nisin synthetic machinery and leader sequence, *Antimicrob. Agents Chemother. 54*, 1498-1505.

20. Xie, L., Miller, L. M., Chatterjee, C., Averin, O., Kelleher, N. L., and van der Donk, W. A. (2004) Lacticin 481: in vitro reconstitution of lantibiotic synthetase activity, *Science 303*, 679-681.

21. Li, B., Yu, J. P., Brunzelle, J. S., Moll, G. N., van der Donk, W. A., and Nair, S. K. (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis, *Science 311*, 1464-1467.

22. Goto, Y., Li, B., Claesen, J., Shi, Y., Bibb, M. J., and van der Donk, W. A. (2010) Discovery of unique lanthionine synthetases reveals new mechanistic and evolutionary insights, *PLoS Biol. 8*, e1000339.

23. Kupke, T., Kempter, C., Gnau, V., Jung, G., and Götz, F. (1994) Mass spectroscopic analysis of a novel enzymatic reaction. Oxidative decarboxylation of the lantibiotic precursor peptide EpiA catalyzed by the flavoprotein EpiD, *J. Biol. Chem. 269*, 5653-5659.

24. Majer, F., Schmid, D. G., Altena, K., Bierbaum, G., and Kupke, T. (2002) The flavoprotein MrsD catalyzes the oxidative decarboxylation reaction involved in formation of the peptidoglycan biosynthesis inhibitor mersacidin, *J. Bacteriol. 184*, 1234-1243.

25. van de Kamp, M., van den Hooven, H. W., Konings, R. N., Bierbaum, G., Sahl, H. G., Kuipers, O. P., Siezen, R. J., de Vos, W. M., Hilbers, C. W., and van de Ven, F. J. (1995) Elucidation of the primary structure of the lantibiotic epilancin K7 from *Staphylococcus epidermidis* K7. Cloning and characterisation of the epilancin-K7-encoding gene and NMR analysis of mature epilancin K7, *Eur. J. Biochem. 230*, 587-600.

26. van der Meer, J. R., Rollema, H. S., Siezen, R. J., Beerthuyzen, M. M., Kuipers, O. P., and de Vos, W. M. (1994) Influence of amino acid substitutions in the nisin leader peptide on biosynthesis and secretion of nisin by *Lactococcus lactis*, *J. Biol. Chem. 269*, 3555-3562.

27. van der Meer, J. R., Polman, J., Beerthuyzen, M. M., Siezen, R. J., Kuipers, O. P., and de Vos, W. M. (1993) Characterization of the *Lactococcus lactis* nisin A operon genes nisP, encoding a subtilisin-like serine protease involved in precursor processing, and nisR, encoding a regulatory protein involved in nisin biosynthesis, *J. Bacteriol. 175*, 2578-2588.

28. Geißler, S., Götz, F., and Kupke, T. (1996) Serine protease EpiP from *Staphylococcus epidermidis* catalyzes the processing of the epidermin precursor peptide, *J. Bacteriol. 178*, 284-288.

29. Kuipers, A., De Boef, E., Rink, R., Fekken, S., Kluskens, L. D., Driessen, A. J., Leenhouts, K., Kuipers, O. P., and Moll, G. N. (2004) NisT, the transporter of the lantibiotic nisin, can transport fully modified, dehydrated and unmodified prenisin and fusions of the leader peptide with non-lantibiotic peptides, *J. Biol. Chem. 279*, 22176-22182.

30. Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools, *Nat. Protoc. 2*, 953-971.

31. Xie, L., Chatterjee, C., Balsara, R., Okeley, N. M., and van der Donk, W. A. (2002) Heterologous expression and purification of SpaB involved in subtilin biosynthesis, *Biochem. Biophys. Res. Commun. 295*, 952-957.

32. Kjeldgaard, M., Nyborg, J., and Clark, B. F. C. (1996) Protein motifs .10. The GTP binding motif: Variations on a theme, *Faseb J. 10*, 1347-1368.

33. Chatterjee, C., Miller, L. M., Leung, Y. L., Xie, L., Yi, M., Kelleher, N. L., and van der Donk, W. A. (2005) Lacticin 481 synthetase phosphorylates its substrate during lantibiotic production, *J. Am. Chem. Soc. 127*, 15332-15333.

34. Müller, W. M., Schmiederer, T., Ensle, P., and Süssmuth, R. D. (2010) *In vitro* biosynthesis of the prepeptide of type-III lantibiotic labyrinthopeptin A2 including formation of a C-C bond as a post-translational modification, *Angew. Chem. Int. Ed. Engl. 49*, 2436-2440.

35. Li, B., and van der Donk, W. A. (2007) Identification of essential catalytic residues of the cyclase NisC involved in the biosynthesis of nisin, *J. Biol. Chem. 282*, 21169-21175.

36. Sollenberger, P. Y., and Martin, R. B. (1970) Mechanism of enamine hydrolysis, *J. Am. Chem. Soc. 92*, 4261-4270.

37. Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzàlez-Duarte, R., Jeffery, J., and Ghosh, D. (1995) Short-chain dehydrogenases/reductases (SDR), *Biochemistry 34*, 6003-6013.

38. Tanaka, N., Nonaka, T., Nakamura, K. T., and Hara, A. (2001) SDR: Structure, mechanism of action, and substrate recognition, *Curr. Org. Chem. 5*, 89-111.

39. Kallberg, Y., Oppermann, U., Jörnvall, H., and Persson, B. (2002) Short-chain dehydrogenases/reductases (SDRs) - Coenzyme-based functional assignments in completed genomes, *Eur. J. Biochem. 269*, 4409-4417.

40. van de Kamp, M., Horstink, L. M., van den Hooven, H. W., Konings, R. N., Hilbers, C. W., Frey, A., Sahl, H. G., Metzger, J. W., and van de Ven, F. J. (1995) Sequence analysis by NMR spectroscopy of the peptide lantibiotic epilancin K7 from *Staphylococcus epidermidis* K7, *Eur. J. Biochem. 227*, 757-771.

41. Petersen, J., Boysen, A., Fogh, L., Tabermann, K., Kofoed, T., King, A., Schrotz-King, P., and Hansen, M. C. (2009) Identification and characterization of a bioactive lantibiotic produced by *Staphylococcus warneri*, *Biol. Chem. 390*, 437-444.

42. Marsh, A. J., O'Sullivan, O., Ross, R. P., Cotter, P. D., and Hill, C. (2010) In silico analysis highlights the frequency and diversity of type 1 lantibiotic gene clusters in genome sequenced bacteria, *BMC Genomics 11*, 679.

43. Verhoef, J., Milatovic, D., and Ekkelenkamp, M. B. (2005) Antimicrobial compounds, WO/2005/023852.

44. Hoffmann, A., Schneider, T., Pag, U., and Sahl, H. G. (2004) Localization and functional analysis of PepI, the immunity peptide of Pep5-producing *Staphylococcus epidermidis* strain 5, *Appl. Environ. Microb. 70*, 3263-3271.

45. Skaugen, M., Abildgaard, C. I., and Nes, I. F. (1997) Organization and expression of a gene cluster involved in the biosynthesis of the lantibiotic lactocin S, *Mol. Gen. Genet. 253*, 674-686.

46. Worobo, R. W., Van Belkum, M. J., Sailer, M., Roy, K. L., Vederas, J. C., and Stiles, M. E. (1995) A signal peptide secretion-dependent bacteriocin from *Carnobacterium divergens*, *J. Bacteriol. 177*, 3143-3149.

47. Maqueda, M., Sanchez-Hidalgo, M., Fernandez, M., Montalban-Lopez, M., Valdivia, E., and Martinez-Bueno, M. (2008) Genetic features of circular bacteriocins produced by Gram-positive bacteria, *FEMS Microbiol. Rev. 32*, 2-22.

48. Reis, M., Eschbach-Bludau, M., Iglesias-Wind, M. I., Kupke, T., and Sahl, H. G. (1994) Producer immunity towards the lantibiotic Pep5: identification of the immunity gene pepI and localization and functional analysis of its gene product, *Appl. Environ. Microbiol. 60*, 2876-2883.

49. van der Heide, T., and Poolman, B. (2002) ABC transporters: one, two or four extracytoplasmic substrate-binding sites?, *EMBO Rep. 3*, 938-943.

50. Kjos, M., Snipen, L., Salehian, Z., Nes, I. F., and Diep, D. B. (2010) The Abi proteins and their involvement in bacteriocin self-immunity, *J. Bacteriol. 192*, 2068-2076.

51. Datta, V., Myskowski, S. M., Kwinn, L. A., Chiem, D. N., Varki, N., Kansal, R. G., Kotb, M., and Nizet, V. (2005) Mutational analysis of the group A streptococcal operon encoding streptolysin S and its virulence role in invasive infection, *Mol. Microbiol. 56*, 681-695.

52. Pei, J., and Grishin, N. V. (2001) Type II CAAX prenyl endopeptidases belong to a novel superfamily of putative membrane-bound metalloproteases, *Trends Biochem. Sci. 26*, 275-277.

53. Sedgley, C. M., Clewell, D. B., and Flannagan, S. E. (2009) Plasmid pAMS1-encoded, bacteriocin-related "Siblicide" in *Enterococcus faecalis*, *J. Bacteriol. 191*, 3183-3188.

54. Diep, D. B., Skaugen, M., Salehian, Z., Holo, H., and Nes, I. F. (2007) Common mechanisms of target cell recognition and immunity for class II bacteriocins, *Proc. Natl. Acad. Sci. U. S. A. 104*, 2384-2389.

55. Kjos, M., Nes, I. F., and Diep, D. B. (2009) Class II one-peptide bacteriocins target a phylogenetically defined subgroup of mannose phosphotransferase systems on sensitive cells, *Microbiology 155*, 2949-2961.

56. Chakicherla, A., and Hansen, J. N. (1995) Role of the leader and structural regions of prelantibiotic peptides as assessed by expressing nisin-subtilin chimeras in Bacillus subtilis 168, and characterization of their physical, chemical, and antimicrobial properties, *J. Biol. Chem. 270*, 23533-23539.

CHAPTER 3. BIOCHEMICAL AND STRUCTURAL CHARACTERIZATION OF A LANTIBIOTIC DEHYDROGENASE*

3.1. INTRODUCTION

Lantipeptides are characterized by the presence of the thioether-cross-linked amino acids lanthionine and methyllanthionine and commonly by the dehydrated amino acids dehydroalanine (Dha) and dehydrobutyrine (Dhb) (Figure 1.1) (*1*). In addition to these usual modifications, lantipeptides may contain several other unnatural amino acids (Figure 1.1), such as the lysinoalanine and hydroxy-aspartate residues in cinnamycin (*2*); S-aminovinyl-D-cysteine, hydroxyproline, and chlorinated tryptophan in microbisporicin (*3*); and D-alanine and N-terminal pyruvyl (Pyr) in lactocin S (*4*). These less common residues can facilitate the interaction with the cellular target or protect the peptides from proteolytic degradation (*5*). The recently discovered lantibiotic epilancin 15X produced by *Staphylococcus epidermidis* 15X154 is remarkably potent against antibiotic-resistant strains of *S. aureus* and *Enterococcus faecalis* and contains an unusual N-terminal D-lactate (Lac) group (*5, 6*). As described in chapter 2, the Lac group is produced by dehydration of the N-terminal Ser residue in the core peptide by the lantibiotic dehydratase ElxB, followed by cleavage of the leader peptide by the protease ElxP, and spontaneous or enzymatic hydrolysis of the resulting N-terminal Dha residue to generate a Pyr group. The resulting ketone-containing peptide, designated dehydroepilancin 15X, is finally reduced to the alcohol by a short chain

* Reproduced in part with permission from: "Velásquez, J. E., Zhang, X., and van der Donk, W. A. (2011) Biosynthesis of the antimicrobial peptide epilancin 15X and its N-terminal lactate, *Chem. Biol.* 18, 857-867." Copyright © 2011 Elsevier Ltd.

dehydrogenase/reductase (SDR) designated ElxO, generating the mature Lac-containing lantibiotic (Figure 2.13) (*5*).

Despite the large diversity of topologies and unusual amino acids in the structures of lantipeptides (*1*), only a few posttranslational modification enzymes have been reconstituted and characterized *in vitro*, including the lanthionine synthetases LanM (*7*-*15*), LanL (*16*), and RamC-like (or LabKC) (*17*), and the cyclase (LanC) (*18*-*20*) involved in (methyl)lanthionine ring formation. Chapter 2 describes the *in vitro* reconstitution of the enzymatic activity of the lantibiotic dehydrogenase ElxO (*5*), an enzyme that may alter the biological activity and the stability of epilancin 15X. In this chapter, additional studies regarding the structural and biochemical characterization of ElxO are presented, demonstrating that the protein is highly tolerant with respect to the structure of its substrates. Indeed, ElxO was used to modify lactocin S, a lantibiotic that is structurally unrelated to epilancin 15X and that contains an N-terminal Pyr, obtaining the Lac-containing analog peptide. Additionally, the reverse reaction catalyzed by ElxO was used to convert epilancin 15X to dehydroepilancin 15X, which was further modified to a fluorescent analog of the lantibiotic. Finally, the role of the N-terminal Lac group in the protection of epilancin 15X against degradation by aminopeptidases is also demonstrated (*5*), suggesting that the introduction of this modification at the N-terminus of different lantibiotics may be advantageous. Thus, ElxO could be used as a tool to generate peptides that are more stable against proteolytic degradation, possess improved biological activity, or contain additional functional groups.

3.2. EXPERIMENTAL METHODS

Materials, organisms, media, and growth conditions

Chemical reagents and media components used in this study were purchased from Sigma-Aldrich or Thermo Fisher Scientific, unless otherwise specified, and were used without further purification. All microorganisms and primer sequences are summarized in Tables 3.1 and 3.2, respectively. *Escherichia coli* strains were routinely grown in LB solid agar or broth supplemented appropriately at 37 °C. Kanamycin (50 µg/mL) and chloramphenicol (12.5 µg/mL) were used to select for plasmid propagation. The strains *Lactobacillus sake* L45 and *Pediococcus acidilactici* Pac1.0 were grown in de Man-Rogosa-Sharpe (MRS) solid agar or broth (Difco), a medium that favors the growth of *Lactobacilli* (*21*). The strains *Lactococcus lactis* subsp. *lactis* CNRZ481 and *L. lactis* subsp. *cremoris* HP were grown in Elliker broth medium without gelatin and supplemented with sodium β-glycerophosphate (15 g/L) as buffering agent (EG'P medium) or in M17 medium (Difco) containing 0.5% glucose (GM17 medium), respectively.

Table 3.1. Microorganisms and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source or reference
Escherichia coli		
DH5*α*	*λpir/ϕ80dlacZΔM15* Δ(*lacZYA-argF*)*U169 recA1 hsdR17 deoR thi-1 supE44 gyrA96 relA1*	(*22*)
Rosetta2(DE)	F^- *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) λ(DE3 [*lacI lacUV5-T7 gene 1 ind1 sam7 nin5*]) pRARE2	Novagen
Lactococcus		
lactis subsp. *lactis* CNRZ 481	Lacticin 481 producer strain	(*23*)
lactis subsp. *cremoris* HP ATCC12602	Lacticin 481 sensitive strain	ATCC[a]
Lactobacillus sake L45	Lactocin S producer strain	(*24*)
Pediococcus acidilactici Pac1.0	Lactocin S sensitive strain	(*24*)
Plasmids		
pET28b	Kan^R *E. coli* T7 based histidine-tag fusion expression vector	Novagen
pET.His_6-ElxO	*elxO* cloned into pET28b vector	(*5*)
pET.His_6-ElxO(S139A)	Mutant pET.His_6-ElxO encoding for His_6-ElxO S139A	This study
pET.His_6-ElxO(Y152F)	Mutant pET.His_6-ElxO encoding for His_6-ElxO Y152F	This study
pET.His_6-ElxO(K156A)	Mutant pET.His_6-ElxO encoding for His_6-ElxO K156A	This study
pET.His_6-ElxO(K156M)	Mutant pET.His_6-ElxO encoding for His_6-ElxO K152A	This study

[a]ATCC, American Type Culture Collection, Manassas, VA.

Table 3.2. Primers used in this study

Name	Sequence
ElxO.S139A.FP	5'- GGA AAT CCG CTT AAT CCT GTA ATA GCA GAA ATA TTT ACT ATT GCT CC -3'
ElxO.S139.RP	5'- GGA GCA ATA GTA AAT ATT TCT GCT ATT ACA GGA TTA AGC GGA TTT CC -3'
ElxO.Y152F.FP	5'- GCG GAT TTC CTT ACT CTA TAT TAT TCG GTA GCA CAA AAC ATG CTG -3'
ElxO.Y152F.RP	5'- CAG CAT GTT TTG TGC TAC CGA ATA ATA TAG AGT AAG GAA ATC CGC -3'
ElxO.K156A.FP	5'- CCT TTA GTT AAA CCA ATA ACA GCA TGT GCT GTG CTA CCG TAT AAT ATA GAG TAA GG -3'
ElxO.K156A.RP	5'- CCT TAC TCT ATA TTA TAC GGT AGC ACA GCA CAT GCT GTT ATT GGT TTA ACT AAA GG -3'
ElxO.K156M.FP	5'- CCT TTA GTT AAA CCA ATA ACA GCA TGC ATT GTG CTA CCG TAT AAT ATA GAG TAA GG -3'
ElxO.K156M.RP	5'- CCT TAC TCT ATA TTA TAC GGT AGC ACA ATG CAT GCT GTT ATT GGT TTA ACT AAA GG -3'

Synthesis of substrate analogues

Peptides were synthesized on a 0.1 mmol or 0.15 mmol scale by Fmoc-based SPPS using a PS3 peptide synthesizer (Protein Technologies) or an Apex SC-396 Peptide Synthesizer (Advanced Chemtech). Fmoc groups were removed during the deprotection steps with 20% piperidine in *N*,*N*-dimethylformamide (DMF) for 3 × 3 min. Coupling of the amino acids was performed for 1-1.5 h with 0.5 or 0.75 mmol (five equiv) of each Fmoc-AA (ChemImpex) and using DMF as solvent and 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU) with 0.4 M *N*-methylmorpholine (NMM) or hydroxybenzotriazole (HOBt) with diisopropylcarbodiimide (DIC) as activating reagents. Coupling reactions of pyruvic acid or α-ketobutyric acid were performed twice for 1-1.5 h using five equiv of acid in the presence of HOBt and DIC. After completion of synthesis, peptidyl-resins were washed with DMF and ethanol, and dried in vacuo. Peptides were cleaved from the resins using a mixture of TFA/water/phenol (90:5:5) for 2 h. For Pyr-MAIVK a cleavage cocktail

containing TFA/water/phenol/thioanisole/mercaptoethanol (82.5:5:5:5:2.5) was used, and for Pyr-YAIVK, Pyr-TAIVK, Pyr-RAIVK, and Obu-RAIVK a cleavage mixture of TFA/water/phenol/thioanisole (85:5:5:5) was used. The cleavage solutions were evaporated using a rotary evaporator to remove TFA and the peptides were precipitated from the solution with cold diethyl ether. The precipitated materials were resuspended in 0.1% TFA in acetonitrile/water (50:50) and lyophilized to dryness. The synthesis of Glx-AAIVK, where Glx stands for glyoxylic acid, was performed by oxidation of SAAIVK (5 mM) with sodium periodate (10 mM) in sodium phosphate buffer (40 mM, pH 7.5) for 5 min, followed by quenching with sodium sulfite (40 mM) (*25*). The peptides were purified by semi preparative reverse phase high performance liquid chromatography (HPLC) using an Agilent 1200 instrument equipped with an Eclipse XDB-C18 column (9.4 mm x 250 mm, Agilent) or a Synergi Fusion-RP column (9.4 mm × 150 mm, Phenomenex) and a variable wavelength detector set at 220 nm. The mobile phase was 0.1% formic acid in water (A) and acetonitrile (B). A gradient of 2-30% B in A and a flow rate of 4 mL/min were used. The masses of the purified peptides were determined by electrospray ionization mass spectrometry (ESI-MS) using a Waters ZMD quadrupole instrument at the Mass Spectrometry Laboratory of the University of Illinois at Urbana-Champaign. See also *Notebook VI, page 41*.

Construction of plasmids pET.His$_6$-ElxO(S139A), pET.His$_6$-ElxO(Y152F), pET.His$_6$-ElxO(K156A), and pET.His$_6$-ElxO(K156M) and expression and purification of proteins

To generate pET.His$_6$-ElxO(S139A), pET.His$_6$-ElxO(Y152F), pET.His$_6$-ElxO(K156A), and pET.His$_6$-ElxO(K156M), the entire pET.His$_6$-ElxO (*5*) was amplified by PCR using *Pfu*Turbo hotstart DNA polymerase (Stratagene) or iProof^{TM} high-fidelity DNA polymerase (BioRad) and the appropriate mutagenesis primers (Table 3.2) followed by treatment with *Dpn*I (New England Biolabs) and transformation of *E. coli* DH5α cells. The correct sequence of the insert was confirmed by sequencing at the W. M. Keck Center for Comparative and Functional Genomics at the University of Illinois at Urbana-Champaign. The proteins were expressed in *E. coli* Rosetta2(DE3) cells and purified using a HisTrap HP column (GE Healthcare) as described in chapter 2 for His$_6$-ElxO (*5*), followed by further purification by size exclusion chromatography using an ÄKTApurifier equipped with a HiLoad 16/60 Superdex 200 column (GE Healthcare) and a flow of 1.5 mL/min of running buffer (50 mM HEPES, 300 mM NaCl, 10% glycerol, pH 7.4). See also *Notebook VI, page 81*.

Wild-type and mutant His$_6$-ElxO activity assays

Wild-type or mutant His$_6$-ElxO (2 or 10 μM) and purified peptide (0.1 to 5 mM) were incubated with NADPH (2.5 mM) in assay buffer (100 mM HEPES, 500 mM NaCl, pH 7.5) at 25 °C. The reaction was initiated by addition of the enzyme and the progress was monitored by UV spectrophotometry, measuring the disappearance of NADPH at 340 nm by using a molar extinction coefficient for NADPH of 6.22 mM^{-1} cm^{-1}. Formation

of reduced peptides was confirmed by liquid chromatography coupled to mass spectrometry (LC-MS) using an Agilent 1200 instrument equipped with a single quadruple multimode ESI/APCI ion source mass spectrometry detector and a Synergi Fusion-RP column (4.6 mm × 150 mm, Phenomenex). The mobile phase was 0.1% formic acid in water (A) and methanol (B). A gradient of 0-70% B in A over 30 min and a flow rate of 0.5 mL/min were used. See also *Notebook VII, page 14*.

Production of dihydrolactocin S and bioactivity assays

Synthetic lactocin S (50 μM), obtained from Prof. J. Vederas (University of Alberta) (*26*), was incubated with His_6-ElxO (50 μM) and NADPH (10 mM) in assay buffer (100 mM HEPES, 500 mM NaCl, pH 7.5) at room temperature for 12 h. The formation of reduced peptide was confirmed by LC-MS using a Waters SYNAPT™ mass spectrometry system equipped with an ACQUITY UPLC®, an ESI ion source, a quadrupole time-of-flight detector, and an ACQUITY Bridged Ethyl Hybrid (BEH) C8 column (2.1 mm × 50 mm, 1.7 μm, Waters). A gradient of 3-97% B (0.1% formic acid in methanol) in A (0.1% formic acid in water) over 12 min was used.

Agar diffusion bioactivity assays were performed using MRS agar media (Difco) (*21*). For each assay, aliquots of agar medium inoculated with overnight cultures of indicator strain (1/100 dilution) were poured into sterile plates. Aliquots of 20 μL of sample were placed into wells made on the solidified agar and the plates were incubated at 37 °C overnight. For determination of critical concentration, the diameter of the inhibition zones were determined and fitted to the equation $D = a + b \times Ln(C)$, where *D* is the diameter of the inhibition zone, *C* is the concentration of bacteriocin, and

a and *b* are constant parameters (*27-29*). For MIC determinations, serial dilutions of peptides were prepared in MRS broth. Then, aliquots of 50 µL of peptide solutions were dissolved in 150 µL of a 1 to 50 dilution of an overnight culture of indicator strain in fresh MRS broth and placed on 96-well plates. The plates were incubated at 37 °C overnight and the wells with no bacterial growth (OD_{600} < 0.3 AU) were determined. See also *Notebook VII, page 41*.

Production of dehydroepilancin 15X and fluorescently labeled epilancin 15X

Epilancin 15X (30 µM) purified from the producer strain as described in chapter 2 (*5*) was incubated with His_6-ElxO (30 µM) and $NADP^+$ (80 mM) in assay buffer (575 mM boric acid, 500 mM NaCl, pH 9.5) at room temperature for 12 h. Control samples lacking enzyme, cofactor, or both were incubated under identical conditions. Formation of oxidized peptide was confirmed by LC-MS analysis using a Waters SYNAPT^{TM} mass spectrometry system as described above. Additionally, reaction and control samples containing (dehydro)epilancin 15X were reacted with phenylene-1,2-diamine (40 mM) in assay buffer (4.0 M sodium acetate, pH 4.8) at 37 °C for 48 h as described elsewhere (*30*). The reaction products were analyzed by matrix-assisted laser desorption/ionization – time-of-flight mass spectrometry (MALDI-TOF MS) on a Voyager DE-STR Biospectrometry Workstation using α-cyano-4-hydroxycinnamic acid as matrix at the Mass Spectrometry Laboratory of the University of Illinois at Urbana Champaign. See also *Notebook VII, page 48*.

For lantibiotic labeling, epilancin 15X (30 µM) was incubated with His_6-ElxO (30 µM), $NADP^+$ (70 mM), and cascade blue hydrazine trisodium salt (10 mM, Invitrogen) in

the presence of aniline (50 mM) and in assay buffer (100 mM HEPES, 500 mM NaCl, pH 7.5) at room temperature for 72 h (*31*). The formation of the hydrazone was confirmed by LC-MS using a Waters SYNAPT^{TM} mass spectrometry system as above. See also *Notebook VII, page 53*.

Production and purification of lacticin 481

Production of lacticin 481 was performed as suggested before (*23, 32*). In brief, *L. lactis* subsp. *lactis* CNRZ481 was grown for 9 h in EG'P media (Elliker broth medium without gelatin and supplemented with sodium β-glycerophosphate), maintaining the culture pH at 5.5 by adding aliquots of 3 M ammonium hydroxide. In previous studies, the purification of lacticin 481 and other lantibiotics has been accomplished using a complex combination of chromatographic techniques and/or organic solvent extractions with low recovery yields and high amounts of oxidized products that are difficult to separate (*32-35*). To purify lacticin 481, the culture supernatant was heat-treated to deactivate proteases and the lantibiotic was concentrated with an ammonium sulfate precipitation step saturating the solution at 60% and 4 °C. After resuspension of the pellet in a Sorensen buffer (NaH_2PO_4 176 mM, Na_2HPO_4 24 mM, pH 6.0), lacticin 481 was further purified by solid phase extraction (SPE) using Vydac C4 reverse phase columns (214SPE1000, Discovery Sciences) and 20 mM ammonium acetate and acetonitrile as eluting solvents. The bacteriocin was eluted with 32% acetonitrile, while most of the impurities were eluted at lower concentrations of organic solvent. The fraction containing lacticin 481 was concentrated by lyophilization and further purified by HPLC using an Agilent 1200 instrument equipped with a Vydac® 214TP54 C_4 reverse

phase column (4.6 mm i.d. × 250 mm L, Discovery Sciences). A gradient of 30-40% B (acetonitrile) in A (20 mM, ammonium acetate, pH 5.5) over 30 min was used. A yield of 1.1 mg of lacticin 481 per liter of culture was obtained, compared with a yield of 0.24 mg per liter reported previously (*32*). The identity of lacticin 481 was confirmed by MALDI-TOF MS. See also *Notebook II, page 64*.

Incubation of peptides with *Aeromonas proteolytica* aminopeptidase

After incubation of epilancin 15X and $NADP^+$ with (or without) His_6-ElxO, aliquots of the peptide solutions were incubated with *A. proteolytica* aminopeptidase (10 U/mL, Sigma-Aldrich) in assay buffer (40 mM Tris-HCl, pH 8.0) at room temperature for 24 h. Samples containing purified lacticin 481 with or without protease were incubated under the same conditions for 12 h. The reaction products were analyzed by MALDI-TOF MS. The lacticin 481 containing samples were tested by agar diffusion bioactivity assays using M17 agar medium supplemented with 0.5% glucose and *L. lactis* subsp. *cremoris* HP as indicator strain. See also *Notebook VII, page 78*.

A solution containing Lac-AAIVKBBIKA, generated enzymatically with His_6-ElxO as described above, and control samples containing Pyr-AAIVKBBIKA or AAIVKBBIKA were also incubated with the peptidase under the conditions described above for 6 h. The peptides were then purified by SPE using Discovery® DSC-18 columns (1 mL, 50 mg, Supelco, Bellefonte, PA) and 0.1% formic acid in water or 0.1% formic acid in 60% acetonitrile as washing and eluting solvents, respectively. The reaction products were finally analyzed by ESI-MS using a Waters ZMD quadrupole instrument at the Mass Spectrometry Laboratory of the University of Illinois at Urbana-Champaign.

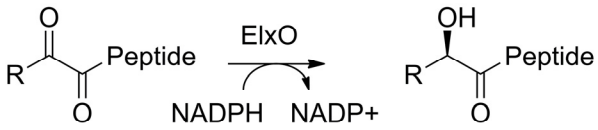
3.3. RESULTS AND DISCUSSION

The data presented in chapter 2 showed that His_6-ElxO catalyzes the conversion of the synthetic peptide Pyr-AAIVK, that resembles the N-terminal region of dehydroepilancin 15X, to D-Lac-AAIVK, demonstrating that the (methyl)lanthionine rings and full length ElxA peptide are not strictly required for substrate recognition by ElxO (*5*). Thus, ElxO could be potentially used to introduce N-terminal alcohols to other peptides or proteins that contain N-terminal Pyr or 2-oxobutyryl (Obu) groups, thereby enhancing their biological activity or chemical stability. The purpose of this study was to evaluate the substrate requirements of ElxO, to characterize its reaction mechanism, and to explore the application of the enzyme for the modification of peptides and other lantibiotics. A combination of site-directed mutagenesis, crystallographic, and kinetic studies using several synthetic substrates was conducted.

Synthesis of substrate analogs and substrate specificity analysis of ElxO

To explore the substrate specificity of the enzyme, a series of small potential substrates was synthesized by Fmoc-based solid phase peptide synthesis (SPPS) followed by coupling of pyruvic acid using hydroxybenzotriazole (HOBt) and diisopropylcarbodiimide (DIC) as activating reagents to produce the Pyr-containing peptides. Single residues of the originally tested substrate, Pyr-AAIVK, were replaced systematically with Ala, and Ala2 was changed to a wide variety of amino acids, including polar, nonpolar, acidic, and basic residues obtaining a diverse set of alternative substrates (Table 3.3).

Table 3.3. Substrates tested with His_6-ElxO. The closest substrate analog Pyr-AAIVK is shown in entry 1 and residues different from Pyr-AAIVK in the subsequent entries are highlighted in red.



R = Me: Pyr
R = Et: Obu
R = H: Glx

R = Me: Lac
R = Et: Hob

Entry	Substrate	k_{cat}/K_M ($M^{-1} \cdot s^{-1}$)	Relative k_{cat}/K_M
	1 23456		
1	Pyr-AAIVK	2.43 ± 0.06	1.00
2	Pyr-AAIV_	1.13 ± 0.00	0.47
3	Pyr-AAI__	0.06 ± 0.00	0.02
4	Pyr-AA___	< 0.03	< 0.01
5	Pyr-A____	< 0.03	< 0.01
6	Pyr-AAIVKBBIKAAKK[a]	14.17 ± 0.36	5.83
7	Pyr-AAIVA	1.33 ± 0.01	0.55
8	Pyr-AAIAK	1.59 ± 0.02	0.65
9	Pyr-AAAVK	0.29 ± 0.01	0.12
10	Pyr-RAIVK	5.50 ± 0.05	2.26
11	Pyr-KAIVK	4.22 ± 0.04	1.74
12	Pyr-DAIVK	0.29 ± 0.02	0.12
13	Pyr-NAIVK	7.60 ± 0.04	3.13
14	Pyr-PAIVK	0.13 ± 0.01	0.05
15	Pyr-MAIVK	15.54 ± 0.11	6.40
16	Obu-AAIVK	0.92 ± 0.03	0.38
17	Obu-RAIVK	1.51 ± 0.02	0.62
18	Glx-AAIVK	< 0.03	< 0.01
Other substrates confirmed by LC-MS:			
19	**Non-polar**	Pyr-GAIVK Pyr-IAIVK Pyr-VAIVK Pyr-FAIVK	
20	**Polar**	Pyr-TAIVK Pyr-YAIVK	
21	**Basic**	Pyr-HAIVK	
22	**Pyr-containing analogs**	Pyr-ASIVK	
23	**Obu-containing analogs**	Obu-AAAVK Obu-AAIAK	
24	**Lantibiotic analogs**	Pyr-AAVLK Pyr-LGPAIK Pyr-APVLA Obu-AGPAIR Pyr-BPVLAAVAVAKKK[a]	

Note: a: B stands for L-2-aminobutyric acid.

If Thr instead of Ser were located at the position 1 of the epilancin 15X core peptide (Figure 2.13), an N-terminal Obu group, instead of a Pyr group, would be

generated upon dehydration and hydrolysis of the N-terminal Dhb residue as in the lantipeptides Pep5 (*36, 37*), lacticin 3147 A2 (*38*), lichenicidin (*39-41*), and prochlorosin 1.7 (*14*). Being able to reduce enzymatically the Obu moiety to the corresponding alcohol would provide a means for generating a number of potentially useful analogs as the more reactive carbonyl group is removed. To explore the specificity of ElxO for the N-terminal ketone group, a series of analogs containing the N-terminal Obu group was also generated by coupling 2-oxobutanoic acid (α-ketobutyric acid) to the N-terminus of the synthetic peptides. To release the peptides from the solid supports, the peptidyl-resins were treated with TFA cleavage cocktails not containing triisopropylsilane (TIS) since presence of this reagent resulted in the chemical reduction of the ketone-containing peptides, as suggested before (*42*). Finally, the peptides were purified by revere phase high performance liquid chromatography (HPLC) and the identities of the compounds were confirmed by electrospray ionization mass spectrometry (ESI-MS).

Purified peptides at different concentrations were incubated with or without His_6-ElxO in the presence of NADPH and the changes in absorbance at 340 nm over time were monitored by UV spectrophotometry establishing the reaction rates. Attempts to determine the steady state kinetic parameters k_{cat} and K_M using a subset of peptides were not successful, since it was not possible to saturate the enzyme with the substrates before reaching the solubility limits of the peptides (Figure 3.1). However, the catalytic efficiencies (k_{cat}/K_M) for some of the substrates were determined after measuring the reaction rates at different peptide concentrations (Table 3.3). For all tested substrates, the k_{cat}/K_M constants were small, presumably because the peptides are lacking structural features compared to the expected physiological substrate, such

as the thioether rings or additional amino acids that may provide essential binding interactions with the enzyme. Alternatively, ElxO could be part of a biosynthetic complex in the native producer and the interaction with other proteins might be required to enhance the catalytic efficiency. Aliquots of the peptide solutions, incubated with NADPH and with or without enzyme, were also analyzed by LC-MS to confirm that the reduced peptides are only formed in the presence of enzyme.

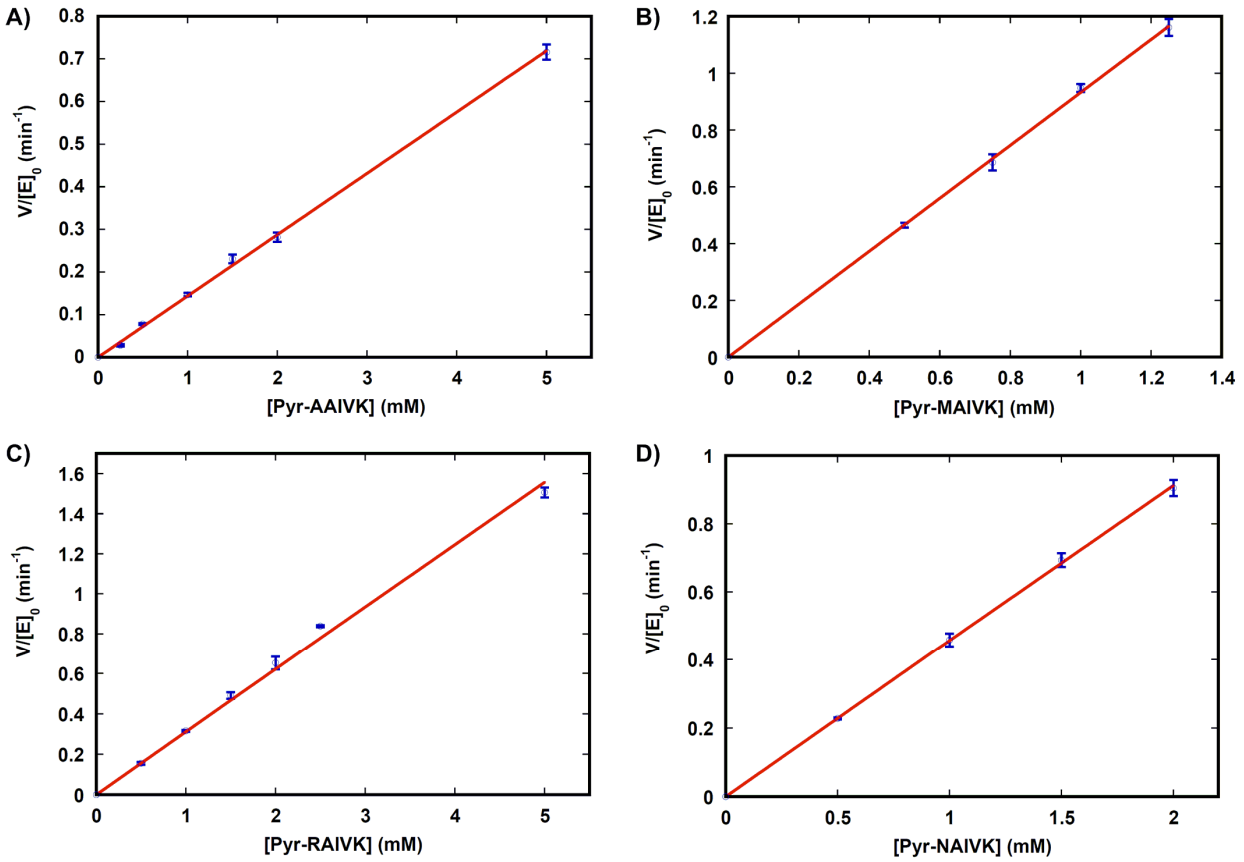


Figure 3.1. Dependence of the reaction rate on Pyr-AAIVK (A), Pyr-MAIVK (B), Pyr-RAIVK (C), and Pyr-NAIVK (D) concentrations by His_6-ElxO and using NADPH as co-substrate. Saturation of the enzyme with the peptide substrates was not achieved even at high peptide concentrations (5 mM) as exemplified with the trials with Pyr-AAIVK and Pyr-RAIVK.

The smaller peptides Pyr-AAIV and Pyr-AAI (entries 2 and 3), but not Pyr-AA and Pyr-A (entries 4 and 5), were reduced by His_6-ElxO in the presence of NADPH based on the spectrophotometric analysis, although with considerably lower reaction rates compared with Pyr-AAIVK. In contrast, the peptide Pyr-AAIVKBBIKAAKK, where B stands for L-2-aminobutyric acid, was converted at a higher rate (Table 3.3, entry 6), suggesting that the length of the peptide is important for substrate recognition. The Ala scanning analysis performed along the sequence of Pyr-AAIVK (Table 3.3, entries 7-9) indicated that the enzyme was able to reduce all the peptides tested; however, k_{cat}/K_M was lower for Pyr-AAAVK (Table 3.3, entry 9). Additionally, peptides containing amino acids with different polarity properties, such as nonpolar (Pro, Met, Gly, Ile, Val, Phe; Table 3.3, entries 14, 15, and 19), polar (Asn, Thr, Tyr; Table 3.3, entries 13 and 20), acidic (Asp; Table 3.3, entry 12), or basic (Arg, Lys, His; Table 3.3, entries 10, 11, and 21) residues at position 2 were transformed to the reduced products. These results suggest that no substrate residue is absolutely required for enzymatic activity and that the minimal length of the peptide to be accepted as substrate is four residues. Interestingly, Pyr-RAIVK, Pyr-KAIVK, Pyr-NAIVK, and Pyr-MAIVK are better substrates than Pyr-AAIVK (Table 3.3, entries 10, 11, 13, 15 vs. 1), while Pyr-DAIVK and Pyr-PAIVK are converted considerably less efficiently (Table 3.3, entries 12 and 14), suggesting that negatively charged residues next to the Pyr group are not well tolerated and that the conformation of the peptide is important for catalysis or substrate recognition.

Obu-AAIVK and Obu-RAIVK were also accepted as substrates leading to the formation of an N-terminal 2-hydroxybutyryl group (Hob), although at lower rates (Table

3.3, entries 16 and 17). Similarly, the peptides Obu-AAAVK and Obu-AAIAK were substrates for the enzyme (Table 3.3, entry 23). However, when Pyr was substituted for a glyoxylyl (Glx) group, such as in the peptide Glx-AAIVK (Table 3.3, entry 18), no significant formation of the reduced peptide was observed based on LC-MS analysis.

In addition to epilancin 15X, two other known lantibiotics, epilancin K7 and epicidin 280, contain an N-terminal Lac moiety. To explore the potential of using His_6-ElxO for the synthesis of other lantibiotics, peptides mimicking the N-terminal portion of dehydroepilancin K7 (Pyr-AAVLK) and dehydroepicidin 280 (Pyr-LGPAIK) were synthesized and tested as substrates (Table 3.3, entry 24). His_6-ElxO reduced both peptides, even though their sequences are quite different from the N-terminus of epilancin 15X. Several other lantibiotics, *e.g.* lactocin S and Pep5, contain N-terminal Pyr and Obu groups produced after dehydration of Ser or Thr at the N-terminus of the core peptide, followed by spontaneous or enzymatic enamine hydrolysis upon leader peptide removal (Figure 2.13). These lantibiotics might be modified to produce the corresponding unnatural dihydropeptides, provided they are substrates for His_6-ElxO. To test this possibility, peptides resembling the N-terminus of lactocin S (Pyr-APVLA and Pyr-BPVLAAVAVAKKK) and Pep5 (Obu-AGPAIR) were incubated with His_6-ElxO in the presence of NADPH (Table 3.3, entry 24). The peptides were reduced as confirmed by LC-MS analysis. These observations suggest that ElxO is quite promiscuous and may be used to synthesize a variety of peptides that contain N-terminal 2-hydroxypropionyl (lactate, Lac) or 2-hydroxybutyryl (Hob) groups.

Production of dihydrolactocin S

Lactocin S is a 37-residue lantibiotic produced by *L. sake* L45 that contains an N-terminal Pyr and that inhibits the growth of species from the genera *Lactobacillus*, *Pediococcus*, and *Leuconostoc* (Figure 3.2A) (*24*). To evaluate if lactocin S is a substrate for ElxO, a synthetic sample of the lantibiotic, obtained from Prof. J. Vederas laboratory, was incubated with NADPH in the presence of His_6-ElxO and the reduction of the peptide was monitored by high-resolution LC-MS (Figure 3.2B), confirming the formation of dihydrolactocin S.

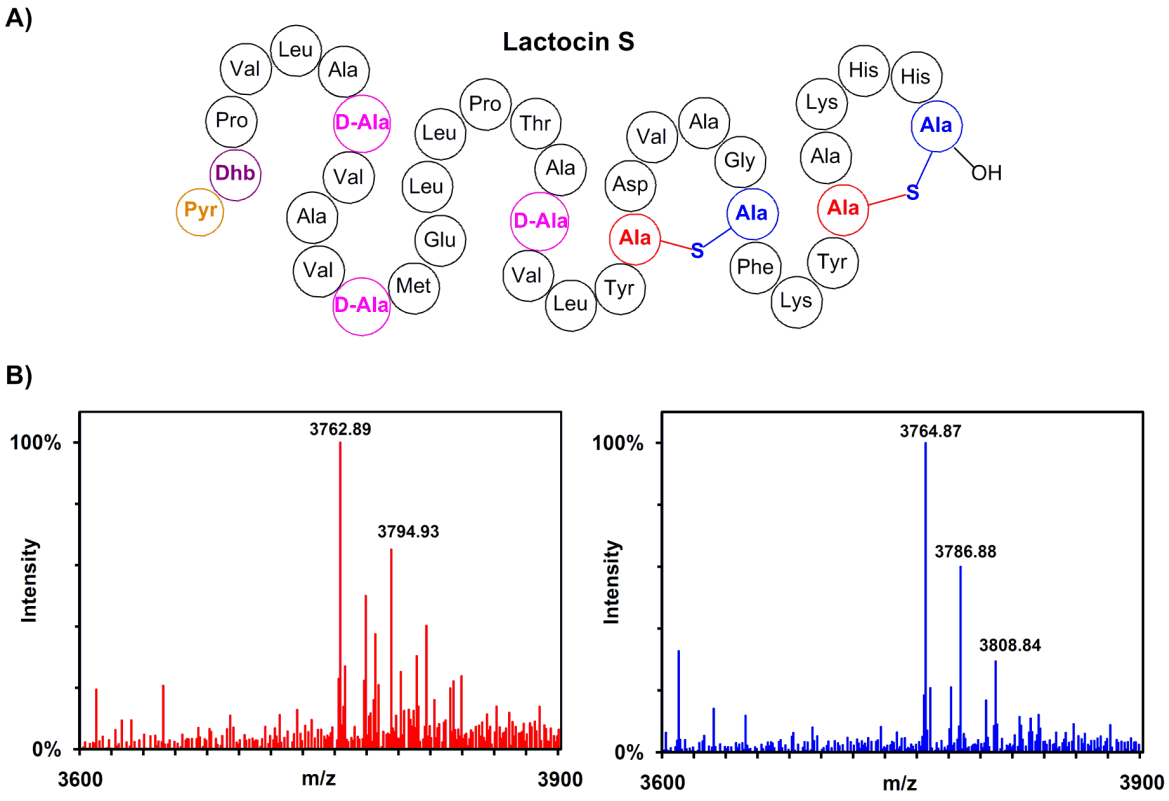
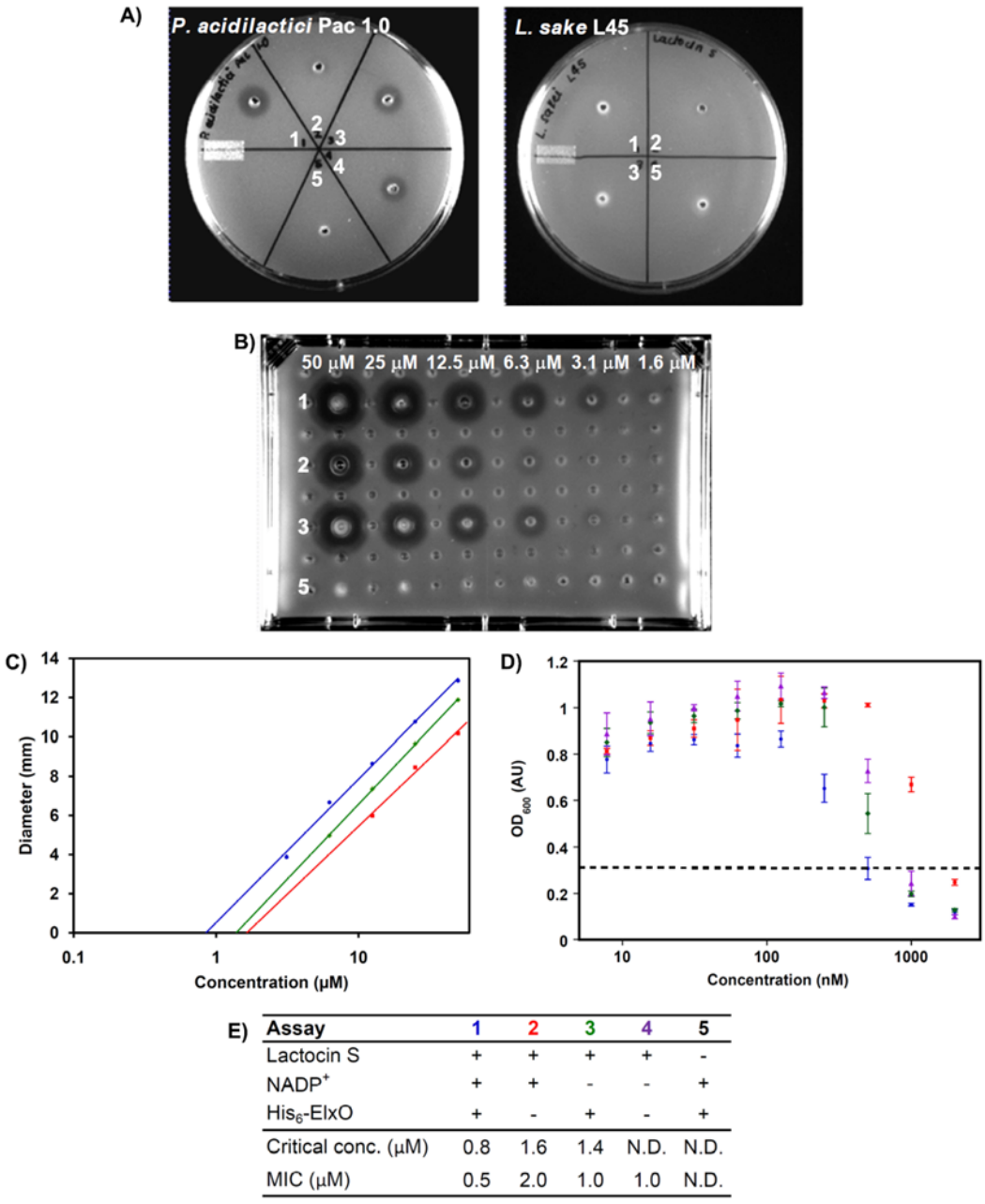


Figure 3.2. Schematic structure of lactocin S and formation of dihydrolactocin S. The lantibiotic lactocin S (A) that contains an N-terminal Pyr group was converted to dihydrolactocin S as evidenced by LC-MS analysis (B). B, left: MS analysis of lactocin S (calculated monoisotopic *m/z* = 3762.89) incubated with NADPH in the absence of His_6-ElxO. The peak at *m/z* = 3794.93 corresponds to oxidized lactocin S (M+O). B, right: MS analysis of dihydrolactocin S (calculated monoisotopic *m/z* = 3764.89) generated after incubation of lactocin S with both NADPH and His_6-ElxO. The peaks at *m/z* = 3786.88 and 3808.84 correspond to the sodium and disodium adducts of dihydrolactocin S.

Samples containing the reduced peptide and control samples containing unmodified lactocin S were tested by single concentration (Figure 3.3A) and serial dilution (Figure 3.3B) agar diffusion bioactivity assays using either *P. acidilactici* Pac1.0 as indicator strain or the lactocin S producer strain. As expected, all the peptides were active against *P. acidilactici* Pac1.0 but not against *L. sake* L45 suggesting that the N-terminal Pyr is not involved in self-immunity. From the serial dilution assay (Figure 3.3B), the sizes of the inhibition zones were determined and the critical concentrations, i.e. the concentrations at which no inhibition zones are observed (*28*), were calculated (Figure 3.3C). Interestingly, the sample containing dihydrolactocin S produced slightly larger inhibition zones and smaller critical concentrations than control samples containing the unreduced peptide (Figure 3.3E), suggesting that the Pyr group of lactocin S is not essential for bioactivity and that Lac may favor the biological activity. Similar results were obtained upon determination of apparent minimal inhibitory concentrations (MIC) by a serial dilution bioactivity assay in liquid media (Figure 3.3D and E).



E)

Assay	1	2	3	4	5
Lactocin S	+	+	+	+	-
$NADP^+$	+	+	-	-	+
His_6-ElxO	+	-	+	-	+
Critical conc. (µM)	0.8	1.6	1.4	N.D.	N.D.
MIC (µM)	0.5	2.0	1.0	1.0	N.D.

Figure 3.3. Antimicrobial activity assays with lactocin S and its reduced analog. Single concentration (A) and serial dilution (B) agar diffusion bioactivity assays of enzymatically synthesized dihydrolactocin S (sample 1, blue) and control samples lacking enzyme (sample 2, red), cofactor (sample 3, green), or both (sample 4, purple) and incubated under identical conditions are shown. Sample 5 was a control assay lacking lactocin S. The sizes of the inhibition zones of the serial dilution agar diffusion bioactivity assays (B) were determined and fitted to a linear model to establish the critical concentration (inhibition zone diameter, D = 0) (C) (*28*). Apparent minimal inhibitory concentrations (MIC) were also determined by a serial dilution bioactivity assay in liquid media (D). The MIC values were established based on the smallest concentration of antibiotic that caused OD_{600} to be less than 0.3 AU. A table summarizing the results is also presented (E).

Preparation of dehydroepilancin 15X and fluorescently labeled epilancin 15X

In addition to formation of alcohols in peptides other than epilancin 15X, as shown above for lactocin S, ElxO could be potentially used to catalyze the reverse reaction and oxidize existing alcohol moieties. The resulting ketone groups may then be conjugated with orthogonal functional groups such as hydrazines or aminooxy containing compounds to generate potentially useful lantibiotic analogs. Indeed, when epilancin 15X was incubated with His_6-ElxO, in the presence of an excess of $NADP^+$ and at high pH, a new peptide with a difference in mass of −2 Da and corresponding to dehydroepilancin 15X was observed by LC-MS (Figure 3.4A). Additionally, when the dehydrated peptide was incubated with phenylen-1,2-diamine (*30*), the N-terminal Pyr group was removed as determined by MALDI-TOF MS analysis (Figure 3.4B), confirming the formation of the oxidized peptide. Furthermore, when the dehydrogenation reaction was performed in the presence of cascade blue labeled hydrazine and aniline (*43*), the corresponding fluorophore-containing peptide was produced (Figure 3.5). Such labeled peptide could be highly valuable, for example, for studies of mode of action. Similarly, ElxO could be used to modify other Lac-containing lantibiotics such as epilancin K7 or epicidin 280 due to its relaxed substrate specificity.

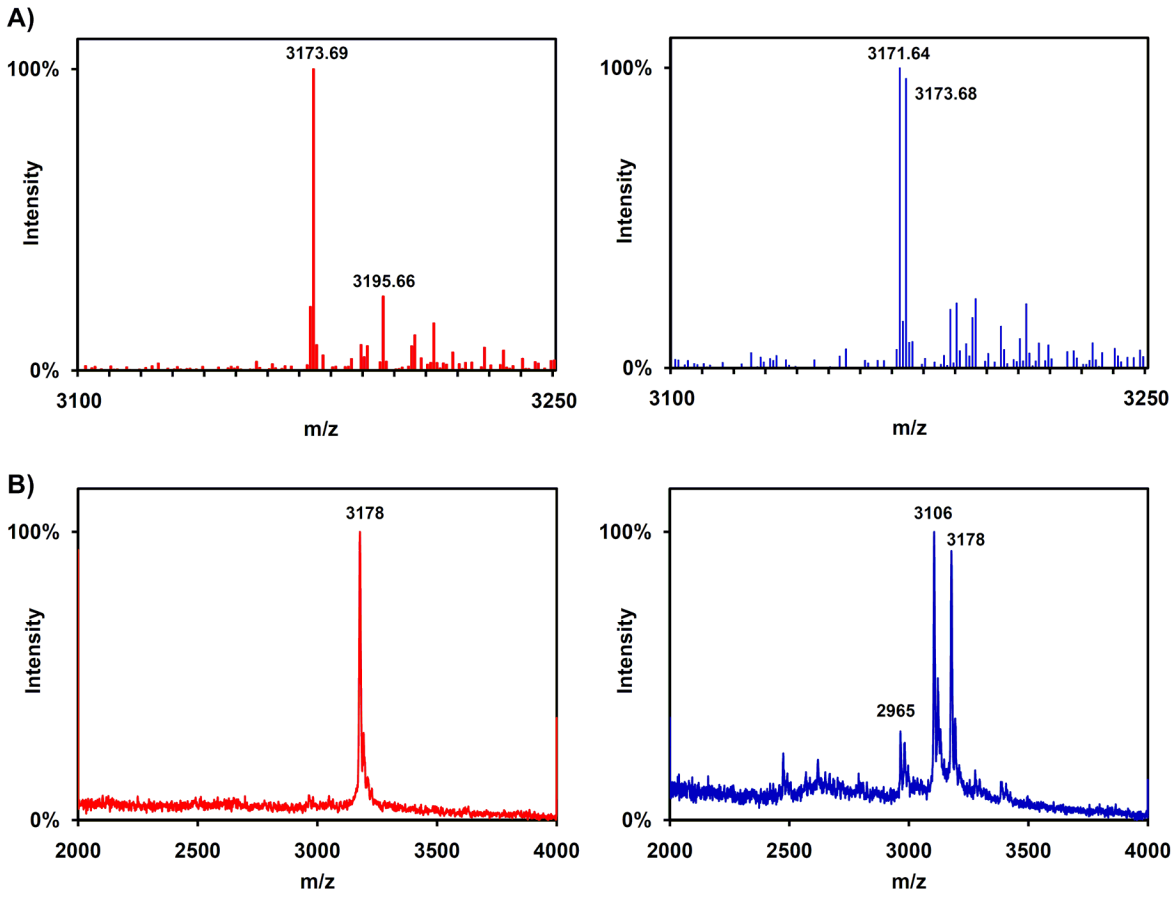


Figure 3.4. Production of dehydroepilancin 15X. A) LC-MS analysis of epilancin 15X after incubation with (blue) or without (red) His_6-ElxO in the presence of $NADP^+$. A new peak corresponding to dehydroepilancin 15X (calculated monoisotopic *m/z* = 3171.68) was observed only in the sample containing enzyme. B) MALDI-TOF MS analysis of samples from A), after treatment with phenylene-1,2-diamine. In addition to unreacted epilancin 15X (calculated average *m/z* = 3176), a new compound corresponding to epilancin(2-31) was observed (calculated average *m/z* = 3106).

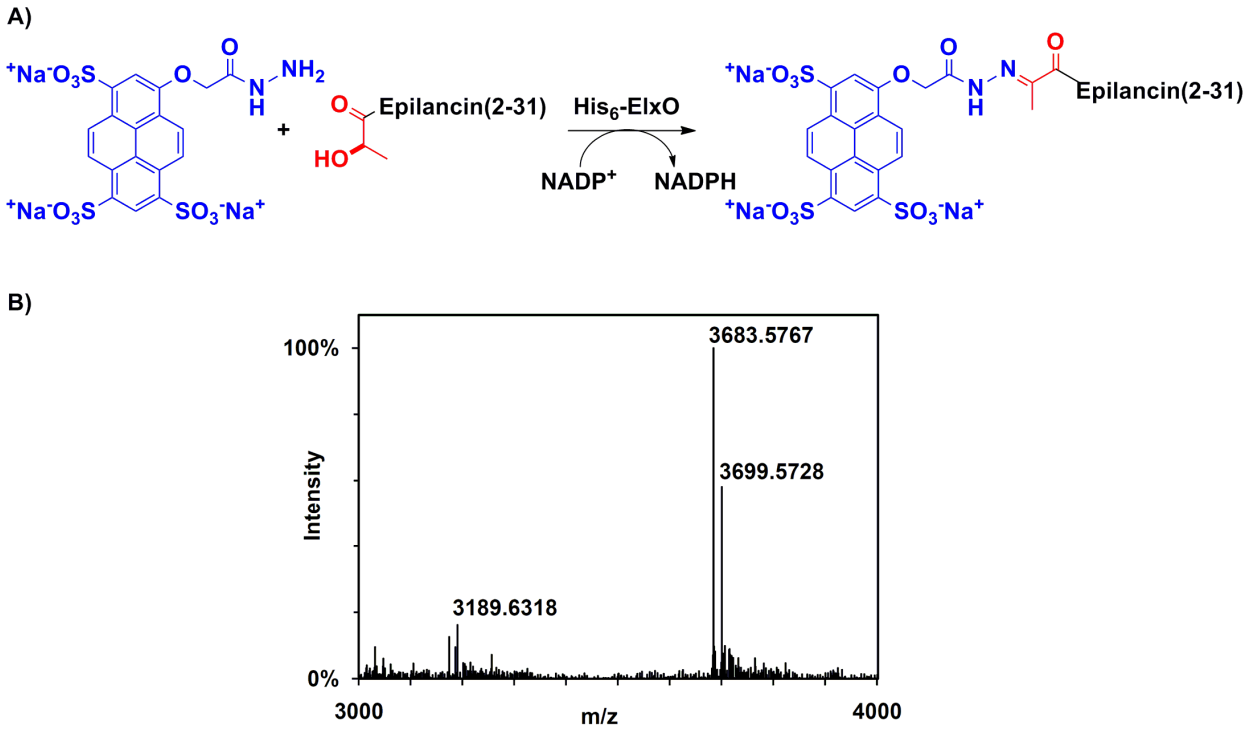


Figure 3.5. Chemo-enzymatic labeling of epilancin 15X with cascade blue. A) Scheme of the labeling reaction of epilancin 15X with cascade blue in the presence of His_6-ElxO. B) After the labeling reaction, the predicted product was observed by LC-MS analysis (calculated monoisotopic *m/z* = 3683.65). The peaks at *m/z* = 3189.63 and 3699.57 correspond to oxidation products (M+O) of epilancin 15X and the labeled peptide, respectively.

X-ray crystal structure of ElxO

At present, very little is known about the structural biology of lantibiotic biosynthetic enzymes, as structures are available only for LanC (*18*), LanD (*44, 45*), and LanP (*46*) enzymes. To obtain some insights into the substrate scope of ElxO and to characterize structurally this lantibiotic dehydrogenase, Ms. Nega Garg, a joint graduate student between the Nair and the van der Donk laboratories, and Professor Satish Nair determined the x-ray crystal structure of the complex of ElxO with NADPH at 1.8 Å resolution (Figure 3.6). Attempts to obtain a co-crystal structure with bound peptide

(Pyr-AAIVK, Pyr-DAIVK, Pyr-VAIVK, Obu-AAIVK, or Obu-RAIVK) were unsuccessful, presumably because of the expected high K_M values for these substrates.

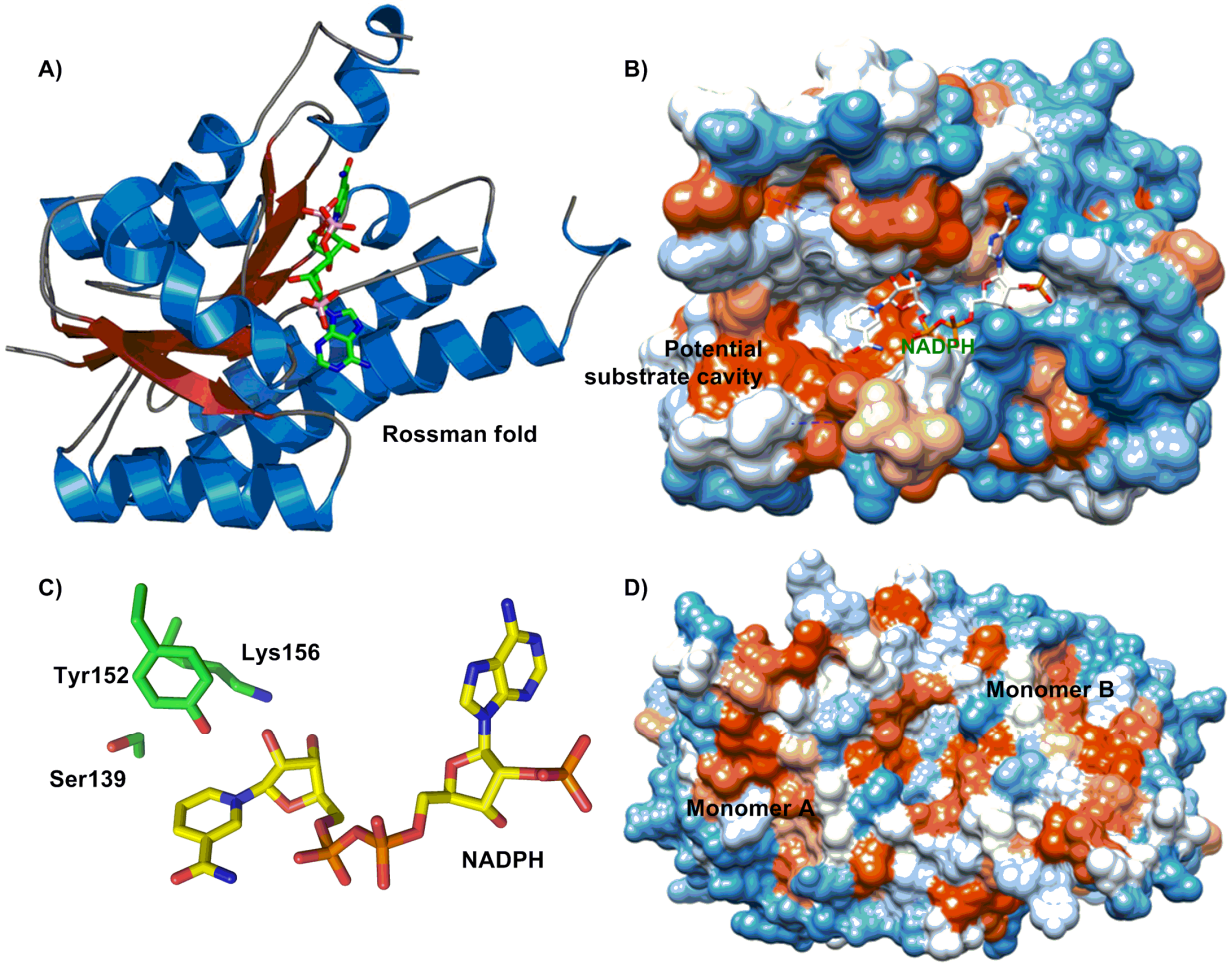


Figure 3.6. X-ray crystal structure of ElxO bound to NADPH. ElxO displays the characteristic Rossman fold of dinucleotide-binding proteins (A). A hydrophobic pocket surrounding the catalytic side may accommodate the pyruvyl-containing substrate (B). The three highly conserved catalytic triad residues (Ser139, Tyr152, and Lys156) (C) and an extended hydrophobic patch covering both monomers and in the opposite side of the protein (D) are shown. The hydrophobicity surface structures are shown in colors from blue (hydrophilic) to red (hydrophobic) based on the Kyte and Doolittle scale (*47*).

The crystallographic data revealed that ElxO displays a single domain architecture and the characteristic Rossman fold (*48*) common to dinucleotide-binding proteins (Figure 3.6A). The core of the enzyme is formed by seven strands creating a

twisted parallel β-sheet flanked by three α-helices from each side and a small one-turn α-helix on the top at the beginning of the active site loop (Figure 3.6A). The region between Glu188 and Ile204 was not observed in the electron density map and it is expected to fold in an additional helix on the top edge of the β-sheet forming a protective lid over the active site that may undergo an 'open to close' conformational change upon substrate-binding (*49, 50*). The structural characterization confirmed that the protein exists as a dimer in solution as suggested by size exclusion chromatography (*5*), with the C-terminal β-strand and an α-helix in each monomer at the edge of the interface.

The structural analysis also revealed that three highly conserved residues (Ser139, Tyr152, and Lys156) located in the active site of the protein (Figure 3.6C) likely play a catalytic role as suggested by amino acid sequence alignment with other proteins from the short-chain dehydrogenase/reductase (SDR) family (*5, 49, 51*). Furthermore, the highly conserved Asn111 presumably involved in a proton relay system that connects bulk solvent to Tyr152 was also identified (*49, 52*). In addition to these residues, other motifs that define classical SDR proteins are apparent. The backbone nitrogen atoms of Gly13, Ile14, and Gly15 and the ε-amino group of Lys12 in the conserved TGX_3GXG motif may stabilize the pyrophosphate moiety by hydrogen bonding interactions with the phosphate oxygens. Furthermore, the adenine group of the cofactor fits into a hydrophobic pocket bounded by Leu58, Val60, Ile70, Ala87, and Ile89 and is isolated from the external solvent by Arg34. A hydrophobic floor in the nicotinamide ring-binding pocket is defined by residues Ile14, Pro182, Ile223, and Val219. Lys156 forms hydrogen bonds to both 2' and 3' hydroxyl groups of the

nicotinamide ribose, positioning the cofactor, and possibly also lowering the pK_a of the catalytic base Tyr152 through an electrostatic influence (*50, 52*). Overall, NADP(H) binds in an extended conformation with the adenine and nicotinamide rings oriented roughly perpendicular to the planes of the respective ribose groups, resulting in an *anti* conformation for the adenine ring and a *syn* conformation for the nicotinamide group and suggesting a B-face 4-*pro*-*S* hydride transfer reaction.

Although a co-crystal structure with a pyruvyl-containing substrate is not available, the structural analysis revealed a hydrophobic pocket surrounding the catalytic site and bounded at the sides by branched aliphatic residues (Ile14, Ile184, Ile185, Ile207, Val219, and Ile223) (Figure 3.6B). This cavity may provide dominant interactions with the substrate. Indeed, when the charged residue Asp was located at position 2 of the substrate (Table 3.3, entry 12), the peptide was processed considerably less efficiently than the substrates containing the non-polar Met, the isosteric Asn, or even basic residues with long side chains such as Arg or Lys at the same position (Table 3.3, entries 10, 11, 13, and 15). Furthermore, when Ile at position 4 was substituted for Ala (Table 3.3, entry 9) the catalytic efficiency was also reduced. Thus, peptides containing long side chains at positions 2 or 4, and likely at position 3, may have better affinities for the enzyme due to hydrophobic interactions at the active site.

Interestingly, an extended hydrophobic patch with several branched aliphatic amino acids (Ile, Val, or Leu), exposed to the solvent, and covering both monomers was identified in the opposite face of the protein (Figure 3.6D). This region could anchor the enzyme by hydrophobic interactions with membrane lipids, or more specifically with

membrane proteins such as the ABC transporter ElxT or the putative immunity protein ElxI2. Thus, ElxO could be part of a complex of proteins involved in biosynthesis, activation, and secretion of epilancin 15X. Alternatively, this patch might interact with the hydrophobic B and C rings of dehydroepilancin 15X (Figure 1.2) providing additional binding energy and justifying the low catalytic efficiencies with the short peptides.

SDR proteins usually act on small substrates, like steroids, prostaglandins, sugars, and small xenobiotics or aliphatic alcohols (*50, 52*). Importantly, although more than 200 structures have been deposited in the Protein Data Bank and several members have been biochemically characterized (*52*), no SDR enzyme acting on larger substrates, such as lantibiotics, has been characterized so far to the best of my knowledge. Interestingly, ElxO structure still resembles the single domain structures of the enzymes recognizing small molecules and additional domains are not required to interact with the long cyclized peptide.

ElxO catalytic mechanism

To determine the importance of Tyr152, Ser139, and Lys156 for catalysis, the single mutant proteins His_6-ElxO Y152F, S139A, K156A, and K156M were prepared and purified. Enzymatic assays with wild-type and mutant proteins were performed using Pyr-AAIVK as substrate and the consumption of NADPH over time was determined by UV spectrophotometry (Figure 3.7). Single mutations of Tyr152 to Phe, Ser139 to Ala, or Lys156 to Ala or Met resulted in a considerable reduction of reaction rate (k_{cat}/K_M < 0.03 $M^{-1}\cdot s^{-1}$), supporting the role of these residues in catalysis and in agreement with similar results for other SDR proteins (*50*). Based on these findings, the

structural characterization of ElxO, and previous studies on SDR enzymes (*49, 50, 52*), an ordered 'bi-bi' catalytic mechanism, with NADP(H) binding first and leaving last, can be proposed for ElxO (Figure 3.8). Initially, the hydroxyl group of protonated Tyr152 and possibly the hydroxyl group of Ser139 are hydrogen bonded to the N-terminal carbonyl oxygen of dehydroepilancin 15X, promoting nucleophilic attack. Then, the 4-*pro*-*S* hydride anion is transferred from C-4 of the nicotinamide ring to the *Si*-face of the carbonyl. The resulting positive charge on the oxidized nicotinamide ring and the adjacent ε-amino group of Lys156 may lower the Tyr152 hydroxyl pK_a, allowing proton transfer to the formed alkoxide in the substrate, generating the *(R)*-Lac moiety of epilancin 15X and $NADP^+$ (*5*).

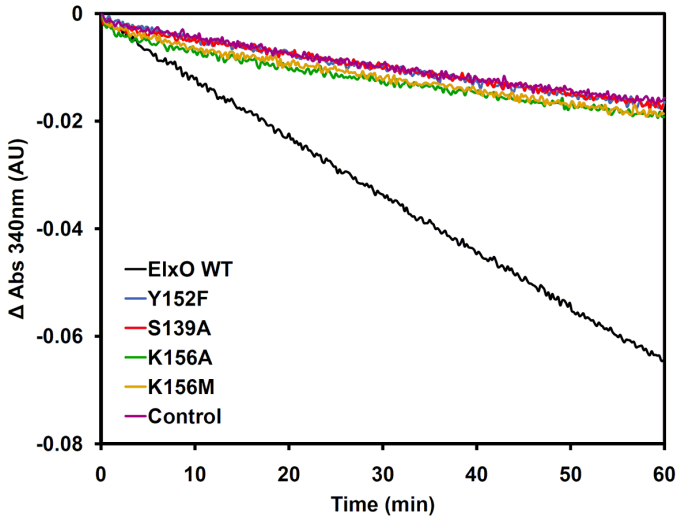


Figure 3.7. Consumption of NADPH (2.5 mM) over time upon reduction of Pyr-AAIVK (5.0 mM) by wild-type His_6-ElxO (black) or the mutants Y152F (blue), S139A (red), K156A (green), or K156M (orange) compared with a control sample lacking peptide (purple).

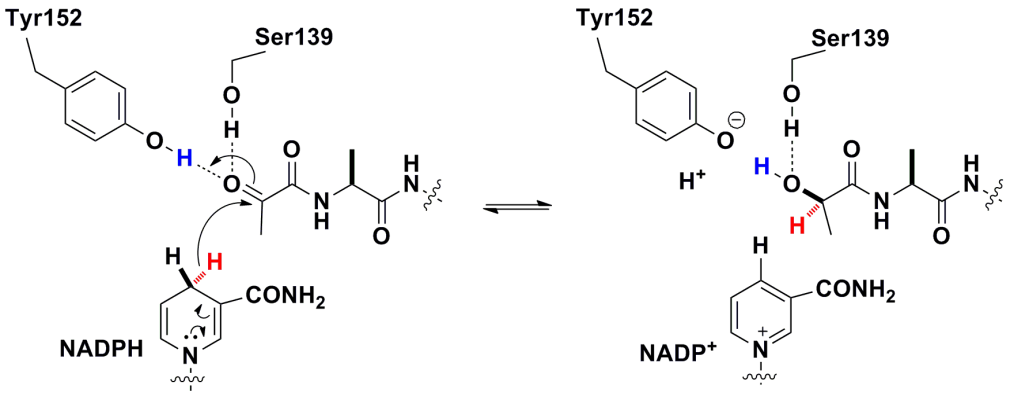


Figure 3.8. Proposed reaction mechanism for NADPH-dependent reduction of dehydroepilancin 15X catalyzed by ElxO.

Role of Lac in proteolytic stability

N-Terminal modification is fairly common in lantibiotics and includes (methyl)lanthionines, disulfides, pyruvate and lactate groups, 2-oxobutyrate groups, and acylations. The importance of these modifications is largely unknown, but the N-terminal disulfide in Halα (one of the two peptides in haloduracin) was shown not to be important for antimicrobial activity, but to protect the peptide from exoproteases (*15, 53*). In addition, the N-terminal lanthionine in lacticin 3147 A1 is not strictly required for antimicrobial activity (*54*), but may protect lacticin 3147 A1 from proteolysis. A similar role can be proposed for the N-terminal Lac group in epilancin 15X. To test this hypothesis, samples of dehydroepilancin 15X, enzymatically generated as described above, and epilancin 15X were incubated with a commercially available aminopeptidase from *A. proteolytica*. Interestingly, only the lantibiotics but not proteolyzed peptides were observed by MS, suggesting that both lactate and pyruvate groups protect the lantibiotics against degradation by the aminopeptidase (Figure 3.9). Although, the N-terminal pyruvate in dehydroepilancin 15X is sufficient to protect the bacteriocin from N-terminal proteolytic degradation, the reduction of this group to lactate may have evolved

because of the high reactivity of ketones. Without its reduction to an alcohol, the pyruvate group in dehydroepilancin 15X may react intramolecularly with the ε-amino groups in the side chains of lysine residues or with other external amines affecting its mode of action.

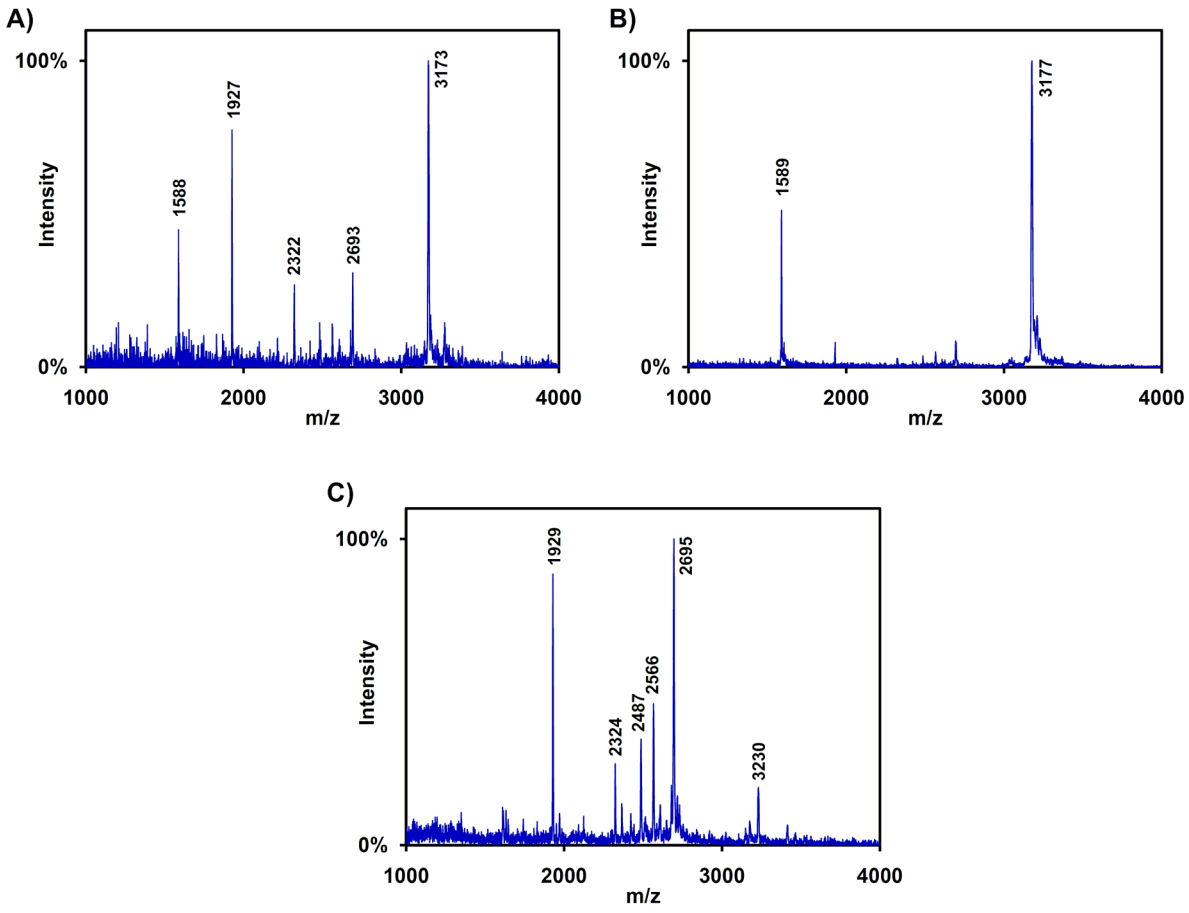


Figure 3.9. MALDI-TOF MS analysis of dehydroepilancin 15X and epilancin 15X after treatment with *A. proteolytica* aminopeptidase. Dehydroepilancin 15X (A) and epilancin 15X (B) are resistant to the protease. The peaks at *m/z* = 3173 and 1588 or m/z = 3177 and 1589 correspond to the $[M+H]^+$ and $[M+2H]^{2+}$ ions of dehydroepilancin 15X (calculated average mass 3173 Da) and epilancin 15X (calculated average mass 3175 Da), respectively. Peaks at m/z= 1927, 2322, and 2693 are also present in the control sample containing aminopeptidase but lacking lantibiotic (C).

To further evaluate the role of the lactate and pyruvate groups in the stability against aminopeptidases, the synthetic peptide Pyr-AAIVKBBIKA (where B stands for L-2-aminobutyric acid) mimicking the N-terminal portion of dehydroepilancin 15X and the corresponding peptide without the Pyr group were synthesized by SPPS. The reduced peptide containing the N-terminal Lac group was also generated enzymatically after incubation of the ketone-containing peptide with His_6-ElxO and NADPH. The three peptides were incubated with the aminopeptidase from *A. proteolytica*, followed by analysis by ESI-MS (Figure 3.10). Whereas the peptide lacking the lactate or pyruvate group was completely degraded, the peptides containing these N-terminal moieties were detected intact, confirming that lactate or pyruvate can prevent the proteolytic degradation by aminopeptidases.

In contrast, when lacticin 481 (Figure 3.11A), a lantibiotic lacking N-terminal modifications, was incubated with the aminopeptidase under similar conditions, a peptide lacking the first seven N-terminal residues was obtained as the main product (Figure 3.11B). Thus, the aminopeptidase was able to remove N-terminal unmodified amino acids with the exception of the His preceding the first MeLan ring that protects the lantibiotic from additional proteolysis. Lacticin 481(2-27) and lacticin 481(3-27) were also observed, suggesting that the GG motif at position 2 is processed at a lower rate and may partially protect the lantibiotic against the action of aminopeptidases. The proteolyzed lacticin 481 and control samples were tested by agar diffusion assays using the indicator strain *L. lactis* subsp. *cremoris* HP (Figure 3.11C), demonstrating a decrease in antibacterial activity of the proteolyzed peptides. These results are

consistent with previous reports that have shown that the linear N-terminal portion of lacticin 481 is important for bioactivity (*55, 56*).

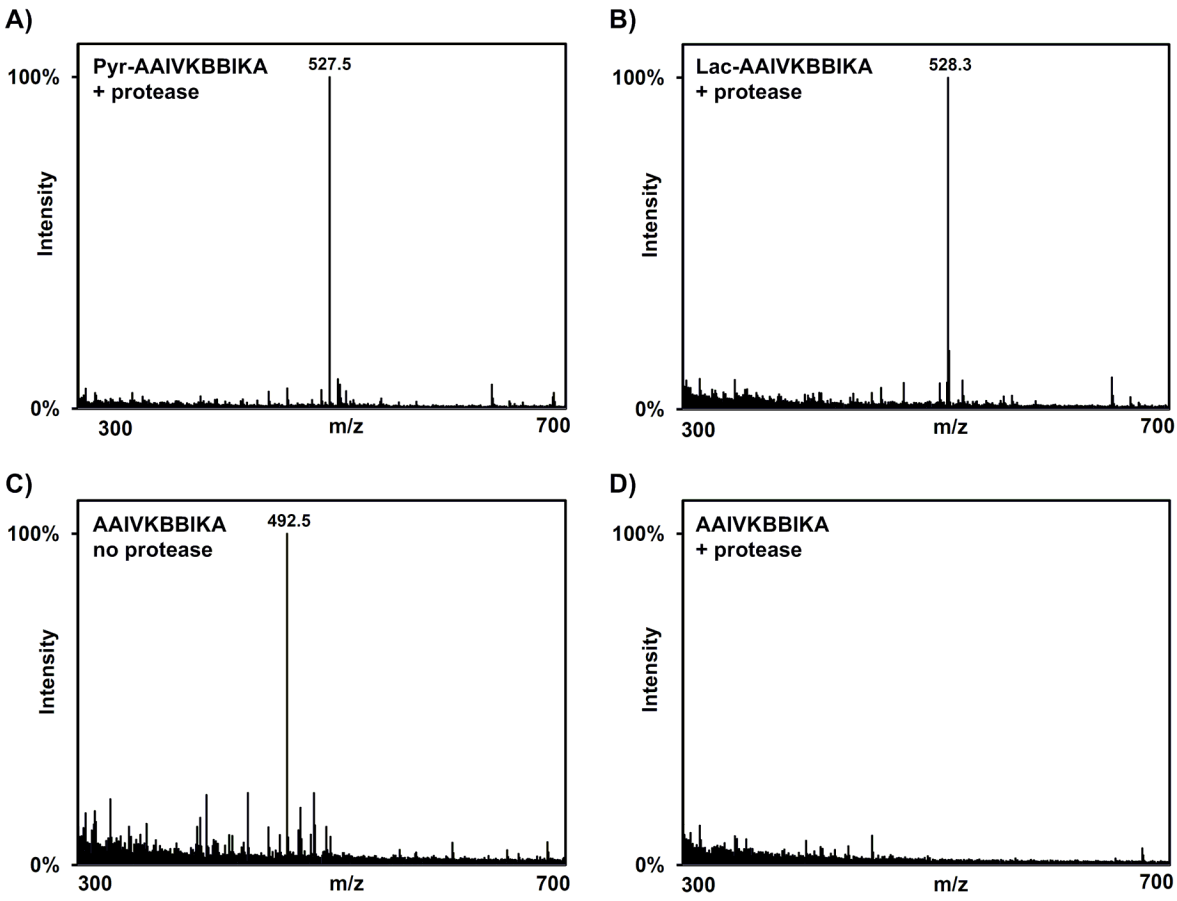


Figure 3.10. Degradation of peptides by *A. proteolytica* aminopeptidase. Samples containing the peptides Pyr-AAIVKBBIKA (A), Lac-AAIVKBBIKA (B), and AAIVKBBIKA (D) were incubated with *A. proteolytica* aminopeptidase and analyzed by ESI-MS. Only the peptide lacking the Lac or Pyr groups was completely degraded by the aminopeptidase (D) compared with a control sample lacking enzyme and incubated under identifcal conditions (C). The observed m/z peaks correspond to the $[M+2H]^{2+}$ ions.

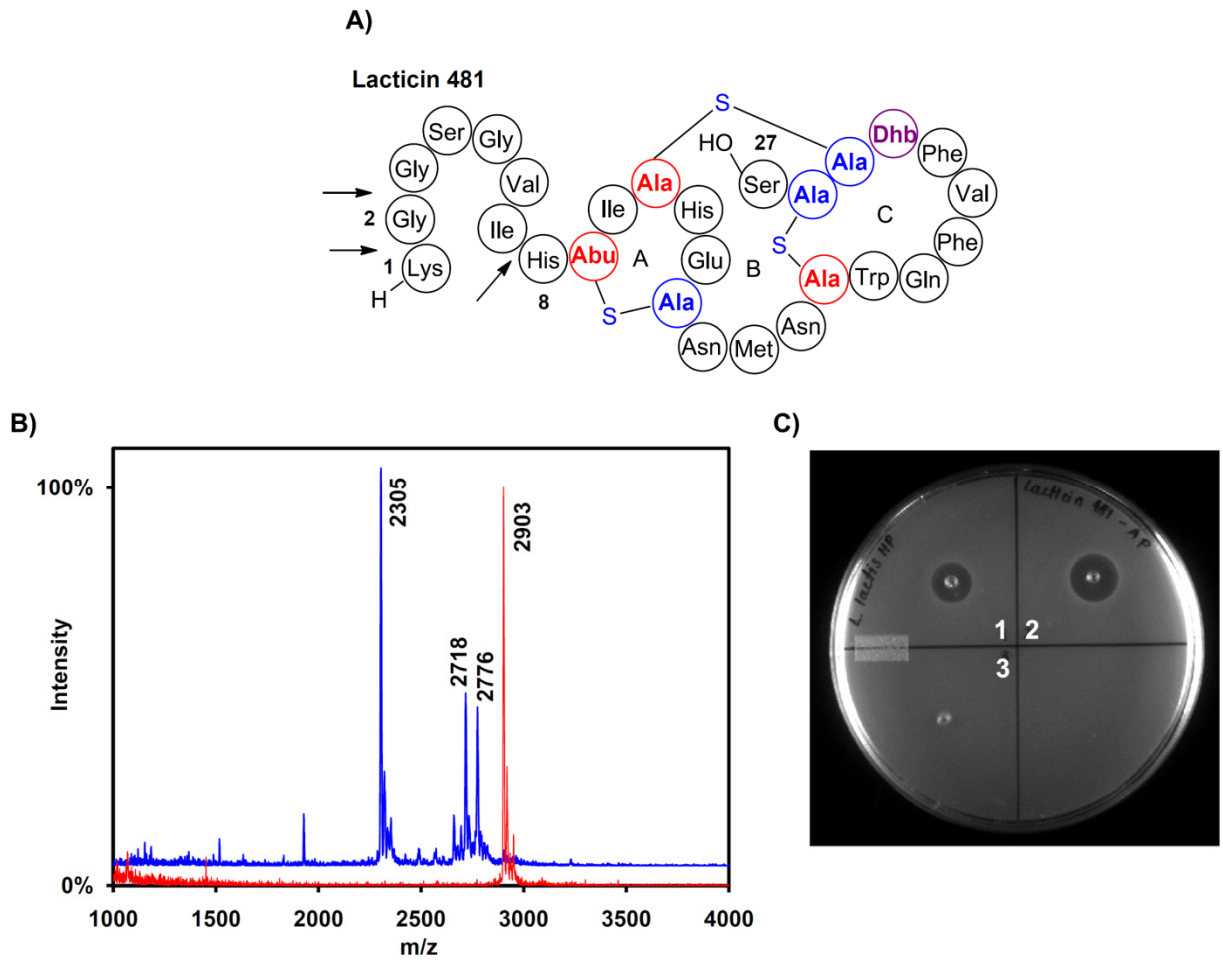


Figure 3.11. Degradation of lacticin 481 by *A. proteolytica* aminopeptidase. A) Structure of the lantibiotic lacticin 481, using the shorthand notation described in Figure 1.1. B) MALDI-TOF MS analysis of lacticin 481 (calculated average m/z = 2902) after incubation with (blue) or without (red) *A. proteolytica* aminopetidase. A new peak corresponding to lacticin 481(8-27) was observed (calculated average m/z = 2304). Peaks corresponding to the lantibiotic after removal of the first Lys (calculated average m/z = 2774) or LysGly (calculated average m/z = 2717) were also observed. C) Agar diffusion bioactivity assay of lacticin 481 after treatment with the peptidase (spot 1) and control sample with no peptidase (spot 2) using *L. lactis* subsp. *cremoris* HP as indicator strain. A control sample containing the peptidase but no lantibiotic was also tested (spot 3). The proteolyzed sample has a considerably lower antibacterial activity.

In summary, ElxO is a lantibiotic dehydrogenase that tolerates diverse N-terminal ketone-containing peptides as substrates and that can be used to generate novel lantibiotics, such as dihydrolactocin S. It can also be used to introduce reactive ketones into Lac-containing peptides that can be subsequently modified by reaction with

hydrazines or aminooxy groups. The x-ray crystal structure of ElxO and the mutagenesis studies suggest that Ser139, Tyr152, and Lys156 play an essential role in catalysis and that substrates with long side chains at position 2 are better tolerated. Furthermore, the N-terminal Lac group introduced by ElxO into epilancin 15X confers stability against proteolytic degradation by aminopeptidases, a feature that may be applied for the engineering of novel lantibiotics with enhanced antibacterial activities or different spectra of action.

3.4. REFERENCES

1. Willey, J. M., and van der Donk, W. A. (2007) Lantibiotics: Peptides of diverse structure and function, *Annu. Rev. Microbiol. 61*, 477-501.

2. Fredenhagen, A., Fendrich, G., Marki, F., Marki, W., Gruner, J., Raschdorf, F., and Peter, H. H. (1990) Duramycins B and C, two new lanthionine containing antibiotics as inhibitors of phospholipase A2. Structural revision of duramycin and cinnamycin, *J. Antibiot. 43*, 1403-1412.

3. Castiglione, F., Lazzarini, A., Carrano, L., Corti, E., Ciciliato, I., Gastaldo, L., Candiani, P., Losi, D., Marinelli, F., Selva, E., and Parenti, F. (2008) Determining the structure and mode of action of microbisporicin, a potent lantibiotic active against multiresistant pathogens, *Chem. Biol. 15*, 22-31.

4. Skaugen, M., Nissen-Meyer, J., Jung, G., Stevanovic, S., Sletten, K., Inger, C., Abildgaard, M., and Nes, I. F. (1994) In vivo conversion of L-serine to D-alanine in a ribosomally synthesized polypeptide, *J. Biol. Chem. 269*, 27183-27185.

5. Velásquez, J. E., Zhang, X., and van der Donk, W. A. (2011) Biosynthesis of the antimicrobial peptide epilancin 15X and its N-terminal lactate, *Chem. Biol. 18*, 857-867.

6. Ekkelenkamp, M. B., Hanssen, M., Danny Hsu, S. T., de Jong, A., Milatovic, D., Verhoef, J., and van Nuland, N. A. (2005) Isolation and structural characterization of epilancin 15X, a novel lantibiotic from a clinical strain of *Staphylococcus epidermidis*, *FEBS Lett. 579*, 1917-1922.

7. Xie, L., Miller, L. M., Chatterjee, C., Averin, O., Kelleher, N. L., and van der Donk, W. A. (2004) Lacticin 481: *In vitro* reconstitution of lantibiotic synthetase activity, *Science 303*, 679-681.

8. Chatterjee, C., Miller, L. M., Leung, Y. L., Xie, L., Yi, M., Kelleher, N. L., and van der Donk, W. A. (2005) Lacticin 481 synthetase phosphorylates its substrate during lantibiotic production, *J. Am. Chem. Soc. 127*, 15332-15333.

9. Miller, L. M., Chatterjee, C., van der Donk, W. A., and Kelleher, N. L. (2006) The Dehydration activity of lacticin 481 synthetase is highly processive, *J. Am. Chem. Soc. 128*, 1420-1421.

10. Patton, G. C., Paul, M., Cooper, L. E., Chatterjee, C., and van der Donk, W. A. (2008) The importance of the leader sequence for directing lanthionine formation in lacticin 481, *Biochemistry 47*, 7342-7351.

11. Paul, M., Patton, G. C., and van der Donk, W. A. (2007) Mutants of the zinc ligands of lacticin 481 synthetase retain dehydration activity but have impaired cyclization activity, *Biochemistry 46*, 6268-6276.

12. You, Y. O., and van der Donk, W. A. (2007) Mechanistic investigations of the dehydration reaction of lacticin 481 synthetase using site-directed mutagenesis, *Biochemistry 46*, 5991-6000.

13. Ökesli, A., Cooper, L. E., Fogle, E. J., and van der Donk, W. A. (2011) Nine post-translational modifications during the biosynthesis of cinnamycin, *J. Am. Chem. Soc. 133*, 13753-13760.

14. Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P. J., Joewono, I., Rusch, D., Chisholm, S. W., and van der Donk, W. A. (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria, *Proc. Natl. Acad. Sci. U. S. A. 107*, 10430-10435.

15. McClerren, A. L., Cooper, L. E., Quan, C., Thomas, P. M., Kelleher, N. L., and van der Donk, W. A. (2006) Discovery and *in vitro* biosynthesis of haloduracin, a two-component lantibiotic, *Proc. Natl. Acad. Sci. U. S. A. 103*, 17243-17248.

16. Goto, Y., Li, B., Claesen, J., Shi, Y., Bibb, M. J., and van der Donk, W. A. (2010) Discovery of unique lanthionine synthetases reveals new mechanistic and evolutionary insights, *Plos Biol. 8*, e1000339.

17. Müller, W. M., Schmiederer, T., Ensle, P., and Süssmuth, R. D. (2010) *In vitro* biosynthesis of the prepeptide of type-III lantibiotic labyrinthopeptin A2 including formation of a C-C bond as a post-translational modification, *Angew. Chem. Int. Ed. Engl. 49*, 2436-2440.

18. Li, B., Yu, J.-P. J., Brunzelle, J. S., Moll, G. N., van der Donk, W. A., and Nair, S. K. (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis, *Science 311*, 1464-1467.

19. Li, B., and van der Donk, W. A. (2007) Identification of essential catalytic residues of the cyclase NisC involved in the biosynthesis of nisin, *J. Biol. Chem. 282*, 21169-21175.

20. Helfrich, M., Entian, K. D., and Stein, T. (2007) Structure-function relationships of the lanthionine cyclase SpaC involved in biosynthesis of the *Bacillus subtilis* peptide antibiotic subtilin, *Biochemistry 46*, 3224-3233.

21. de Man, J. D., Rogosa, M., and Sharpe, M. E. (1960) A medium for the cultivation of *Lactobacilli*, *J. Appl. Bact. 23*, 130-135.

22. Grant, S. G., Jessee, J., Bloom, F. R., and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants, *Proc. Natl. Acad. Sci. U. S. A. 87*, 4645-4649.

23. Piard, J. C., Delorme, F., Giraffa, G., Commissaire, J., and Desmazeaud, M. (1990) Evidence for a bacteriocin produced by *Lactococcus lactis* CNRZ 481, *Neth. Milk Diary J. 44*, 143-158.

24. Mortvedt, C. I., Nissen-Meyer, J., Sletten, K., and Nes, I. F. (1991) Purification and amino acid sequence of lactocin S, a bacteriocin produced by *Lactobacillus sake* L45, *Appl. Environ. Microbiol. 57*, 1829-1834.

25. Geoghegan, K. F., and Stroh, J. G. (1992) Site-directed conjugation of non-peptide groups to peptides and proteins via periodate oxidation of a 2-amino alcohol. Application to modification at N-terminal serine, *Bioconjugate Chem. 3*, 138-146.

26. Ross, A. C., Liu, H., Pattabiraman, V. R., and Vederas, J. C. (2010) Synthesis of the lantibiotic lactocin S using peptide cyclizations on solid phase, *J. Am. Chem. Soc. 132*, 462-463.

27. Wolf, C. E., and Gibbons, W. R. (1996) Improved method for quantification of the bacteriocin nisin, *J. Appl. Bacteriol. 80*, 453-457.

28. Delgado, A., Brito, D., Fevereiro, P., Tenreiro, R., and Peres, C. (2005) Bioactivity quantification of crude bacteriocin solutions, *J. Microbiol. Meth. 62*, 121-124.

29. Parente, E., Brienza, C., Moles, M., and Ricciardi, A. (1995) A comparison of methods for the measurement of bacteriocin activity, *J. Microbiol. Meth. 22*, 95-108.

30. Stevens, J., and Dixon, H. B. (1995) The removal of 2-oxoacyl residues from the N-terminus of peptides and cystatin in non-denaturing conditions, *Biochim. Biophys. Acta 1252*, 195-202.

31. Dirksen, A., Dirksen, S., Hackeng, T. M., and Dawson, P. E. (2006) Nucleophilic catalysis of hydrazone formation and transimination: implications for dynamic covalent chemistry, *J. Am. Chem. Soc. 128*, 15602-15603.

32. Piard, J. C., Muriana, P. M., Desmazeaud, M. J., and Klaenhammer, T. R. (1992) Purification and partial characterization of lacticin 481, a lanthionine-containing bacteriocin produced by *Lactococcus lactis* subsp. *lactis* CNRZ 481, *Appl. Environ. Microbiol. 58*, 279-284.

33. Allgaier, H., Jung, G., Werner, R. G., Schneider, U., and Zahner, H. (1986) Epidermin: sequencing of a heterodetic tetracyclic 21-peptide amide antibiotic, *Eur. J. Biochem. 160*, 9-22.

34. Bailey, F. J., and Hurst, A. (1971) Preparation of a highly active form of nisin from *Streptococcus lactis*, *Can. J. Microbiol. 17*, 61-67.

35. Cheeseman, G. C., and Berridge, N. J. (1957) An improved method of preparing nisin, *Biochem. J. 65*, 603-608.

36. Kaletta, C., Entian, K. D., Kellner, R., Jung, G., Reis, M., and Sahl, H. G. (1989) Pep5, a new lantibiotic: structural gene isolation and prepeptide sequence, *Arch. Microbiol. 152*, 16-19.

37. Meyer, C., Bierbaum, G., Heidrich, C., Reis, M., Süling, J., Iglesias-Wind, M. I., Kempter, C., Molitor, E., and Sahl, H.-G. (1995) Nucleotide sequence of the lantibiotic Pep5 biosynthetic gene cluster and functional analysis of PepP and PepC, *Eur. J. Biochem. 232*, 478-489.

38. Ryan, M. P., Jack, R. W., Josten, M., Sahl, H. G., Jung, G., Ross, R. P., and Hill, C. (1999) Extensive post-translational modification, including serine to D-alanine conversion, in the two-component lantibiotic, lacticin 3147, *J. Biol. Chem. 274*, 37544-37550.

39. Begley, M., Cotter, P. D., Hill, C., and Ross, R. P. (2009) Identification of a novel two-peptide lantibiotic, lichenicidin, following rational genome mining for LanM proteins, *Appl. Environ. Microbiol. 75*, 5451-5460.

40. Dischinger, J., Josten, M., Szekat, C., Sahl, H. G., and Bierbaum, G. (2009) Production of the novel two-peptide lantibiotic lichenicidin by *Bacillus licheniformis* DSM 13, *Plos One 4*, e6788.

41. Shenkarev, Z. O., Finkina, E. I., Nurmukhamedova, E. K., Balandin, S. V., Mineev, K. S., Nadezhdin, K. D., Yakimenko, Z. A., Tagaev, A. A., Temirov, Y. V., Arseniev, A. S., and Ovchinnikova, T. V. (2010) Isolation, structure elucidation, and synergistic antibacterial activity of a novel two-component lantibiotic lichenicidin from *Bacillus licheniformis* VK21, *Biochemistry 49*, 6462-6472.

42. Marceau, P., Bure, C., and Delmas, A. F. (2005) Efficient synthesis of C-terminal modified peptide ketones for chemical ligations, *Bioorg. Med. Chem. Lett. 15*, 5442-5445.

43. Dirksen, A., and Dawson, P. E. (2008) Rapid oxime and hydrazone ligations with aromatic aldehydes for biomolecular labeling, *Bioconjugate Chem. 19*, 2543-2548.

44. Blaesse, M., Kupke, T., Huber, R., and Steinbacher, S. (2000) Crystal structure of the peptidyl-cysteine decarboxylase EpiD complexed with a pentapeptide substrate, *Embo J. 19*, 6299-6310.

45. Blaesse, M., Kupke, T., Huber, R., and Steinbacher, S. (2003) Structure of MrsD, an FAD-binding protein of the HFCD family, *Acta Cryst. Section D Biol. Cryst. D59*, 1414-1421.

46. Minasov, G., Kuhn, M., Ruan, J., Halavaty, A., Shuvalova, L., Dubrovska, I., Winsor, J., Bagnoli, F., Falugi, F., Bottomley, M., Grandi, G., and Anderson, W. F. (2011) 1.95 Angstrom resolution crystal structure of epidermin leader peptide processing serine protease (EpiP) S393A mutant from *Staphylococcus aureus*. PDB ID: 3T41, Center for Structural Genomics of Infectious Diseases (CSGID).

47. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol. 157*, 105-132.

48. Rossmann, M. G., Moras, D., and Olsen, K. W. (1974) Chemical and biological evolution of nucleotide-binding protein, *Nature 250*, 194-199.

49. Ladenstein, R., Winberg, J. O., and Benach, J. (2008) Structure-function relationships in short-chain alcohol dehydrogenases, *Cell. Mol. Life Sci. 65*, 3918-3935.

50. Tanaka, N., Nonaka, T., Nakamura, K. T., and Hara, A. (2001) SDR: Structure, mechanism of action, and substrate recognition, *Curr. Org. Chem. 5*, 89-111.

51. Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzàlez-Duarte, R., Jeffery, J., and Ghosh, D. (1995) Short-chain dehydrogenases/reductases (SDR), *Biochemistry 34*, 6003-6013.

52. Kavanagh, K. L., Jörnvall, H., Persson, B., and Oppermann, U. (2008) The SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes, *Cell. Mol. Life Sci. 65*, 3895-3906.

53. Cooper, L. E., McClerren, A. L., Chary, A., and van der Donk, W. A. (2008) Structure-activity relationship studies of the two-component lantibiotic haloduracin, *Chem. Biol. 15*, 1035-1045.

54. Cotter, P. D., Deegan, L. H., Lawton, E. M., Draper, L. A., O'Connor, P. M., Hill, C., and Ross, R. P. (2006) Complete alanine scanning of the two-component lantibiotic lacticin 3147: generating a blueprint for rational drug design, *Mol. Microbiol. 62*, 735-747.

55. Uguen, P., Hindré, T., Didelot, S., Marty, C., Haras, D., Le Pennec, J. P., Vallee-Rehel, K., and Dufour, A. (2005) Maturation by LctT is required for biosynthesis of full-length lantibiotic lacticin 481, *Appl. Environ. Microbiol. 71*, 562-565.

56. Levengood, M. R., Knerr, P. J., Oman, T. J., and van der Donk, W. A. (2009) *In vitro* mutasynthesis of lantibiotic analogues containing nonproteinogenic amino acids, *J. Am. Chem. Soc. 131*, 12024-12025.

CHAPTER 4. STUDIES ON THE SUBSTRATE SCOPE OF THE SERINE-TYPE LANTIBIOTIC PROTEASES

4.1. INTRODUCTION

Lantipeptides are ribosomally synthesized as precursor peptides, followed by posttranslational modifications that generate the biologically active compounds. The precursor molecules consist of a leader region that is not modified and a core peptide that comprises the backbone of the final metabolite. Based on the enzymes that introduce the thioether cross-links into the core region, lantipeptides can be categorized in four classes as discussed in chapter 1. In class I members (exemplified by the well studied nisin, Pep5, epidermin, and the recently discovered epilancin 15X), the serine and threonine residues in the core peptide are dehydrated by a dehydratase (ElxB in epilancin 15X) and the intramolecular additions of the cysteine residues into the dehydrated amino acids are catalyzed by a cyclase (ElxC in epilancin 15X) (Figure 4.1). In the other three classes, a bifunctional lanthionine synthetase (LanM in class II, RamC-like in class III, or LanL in class IV) catalyzes both dehydration and (methyl)lanthionine formation (*1-3*). Subsequent to ring formation, the leader peptide is cleaved by a protease releasing the bioactive molecule.

Two different classes of proteases are known to remove the leader peptide from the lantipeptide precursors. In most of the class II lantibiotics, as well as in linear bacteriocins and class II microcins (*4, 5*), proteolysis of the leader peptide is performed by a membrane located ATP-binding cassette (ABC) transporter, maturation, and secretion protein (AMS), designated as LanT, that contains an N-terminal cysteine-type

protease domain (Figure 4.2, purple diamonds). These LanT proteins cleave the leader peptide after a highly conserved double-glycine motif with consensus sequence (I/L/V)(S/T)X_2E(L/M)X_2(I/L/V)X**G**(**G**/A/S), where X can be any amino acid (*4, 5*). The *in vitro* enzymatic activity, mechanism, and substrate specificity of some AMS proteins, including LctT (*6*) and NukT (*7*) (the LanT enzymes that cleave off the leaders from lacticin 481 and nukacin ISK-1), among others (*8-12*), have been investigated.

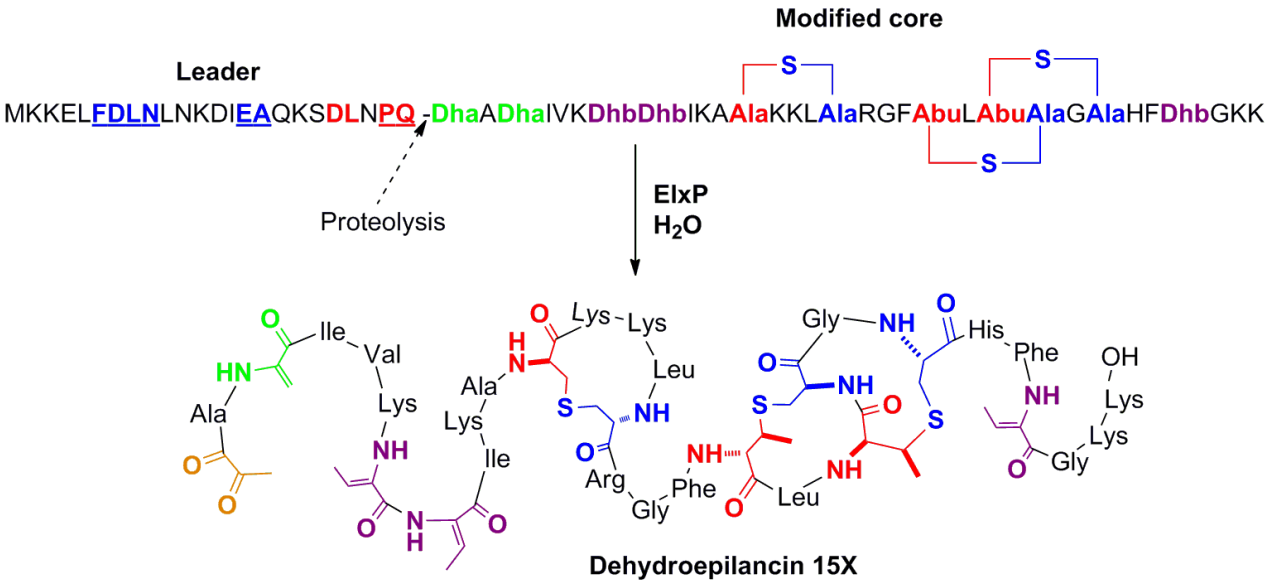


Figure 4.1. Cleavage of the leader peptide by ElxP and formation of the intermediate dehydroepilancin 15X in epilancin 15X biosynthesis. The positions in the leader peptide that were investigated in the current work are underlined and residues in the same region that are likely important for substrate recognition are shown in red (see results and discussion).

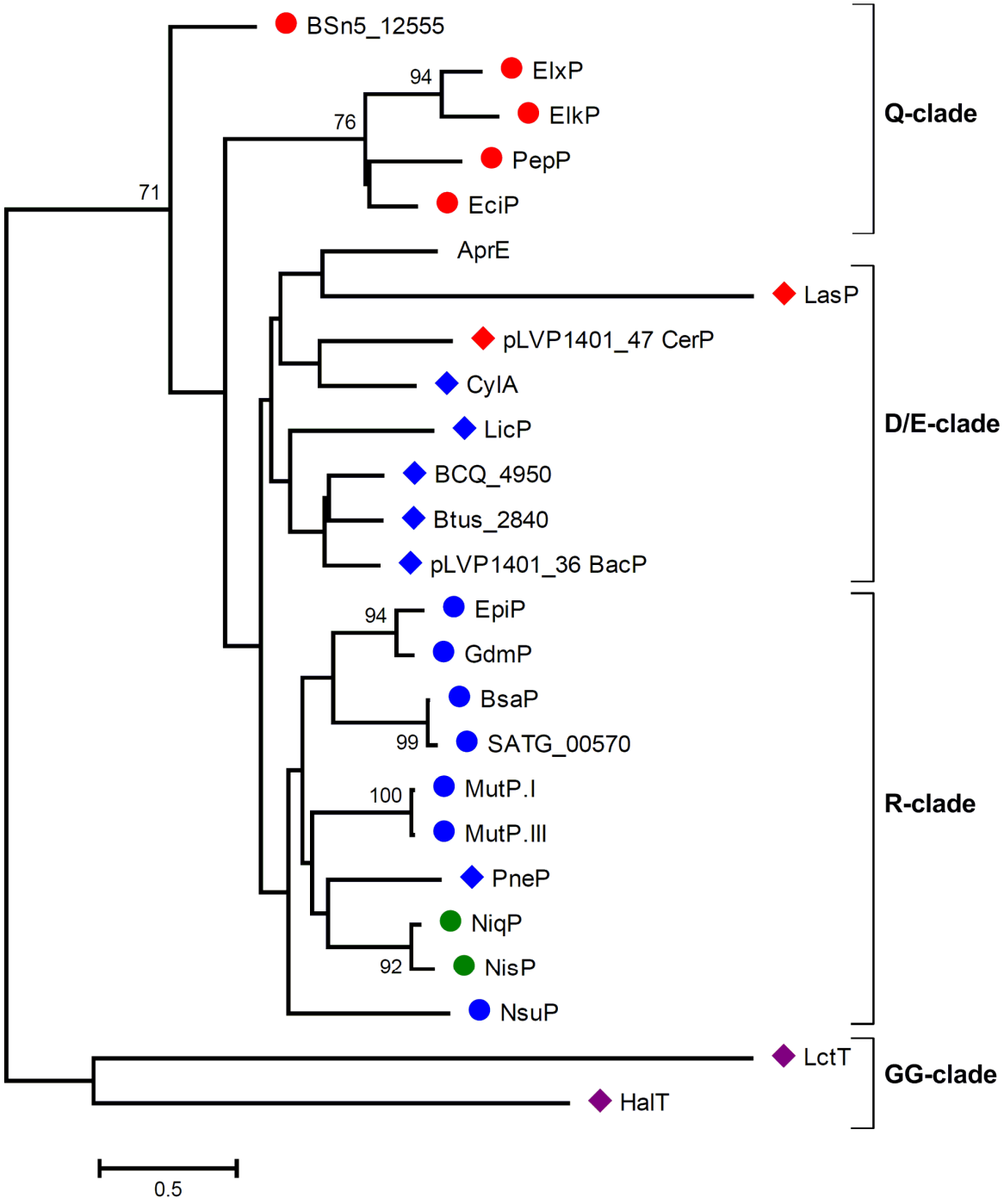


Figure 4.2. Cladograph of LanP proteases.

Figure 4.2. *Continuation*. Cladograph of LanP proteases. The clades are named based on the C-terminal residue of the corresponding leader peptides (see discussion and Figure 4.4). The tree was constructed using the Neighbor-Joining method (*13*) and the distances were computed using the JTT matrix-based method (*14*) in units of number of amino acid substitutions per site. The optimal tree with the sum of branch length = 14.72311067 is shown. The percentage of replicate trees in which the associated proteins clustered together in the bootstrap test are shown next to the branches, only for values over 70% (*15*). Evolutionary analyses were conducted in MEGA5 (*16*). The analysis involved the amino acid sequences from the predicted proteases for the biosynthesis of epilancin 15X (ElxP), epilancin K7 (ElkP), Pep5 (PepP), epicidin 280 (EciP), lactocin S (LasP), cytolysin L and S (CylA), lichenicidin (LicP), epidermin (EpiP), gallidermin (GdmP), Bsa (BsaP), mutacin I (MutP.I), mutacin III/1140 (MutP.III), pneumococcin (PneP), nisin Q (NiqP), nisin A (NisP), nisin U (NsuP), and additional members identified by genome mining. The AMS proteins LctT and HalT (in lacticin 481 and haloduracin biosynthesis) were included as an out-group. The subtilisin AprE is one of the proteases that can cleave the leader peptide from subtilin (*17*). Notation: Circles, class I lantipeptides; diamonds, class II lantipeptides; red, intracellular proteases; blue, extracellular proteases; green, cell wall-anchored proteases; purple, AMS proteins.

In contrast, for most class I lantipeptides (Figure 4.2, circles), such as epilancin 15X, and also for some class II members (Figure 4.2, diamonds), removal of the leader region is performed by a serine-type protease from the subtilisin group and commonly designated as LanP (Figure 4.1). However, the cellular localization and domain structure of LanP proteases vary considerably (Figures 4.2 and 4.3). In the biosynthesis of epidermin and several other lantipeptides (Figure 4.2, blue and Figure 4.3), the peptidases (EpiP in epidermin) contain an N-terminal signal and a pro-peptide sequence and are located extracellularly (*18, 19*). Similarly, the proteases NisP and NiqP (Figure 4.2, green and Figure 4.3) involved in nisin A and nisin Q biosynthesis, respectively, are also extracellular, but are attached to the cell wall through a C-terminal anchor peptide (*20*). For these two classes of proteases, the inactive and fully modified precursor peptide is secreted before leader removal. In contrast, in epilancin 15X biosynthesis, the peptidase ElxP is predicted to be located intracellularly (Figure 4.2, red and Figure 4.3), possibly as part of a membrane-bound biosynthetic complex, and the mature peptide is produced entirely in the cytoplasm (*21, 22*).

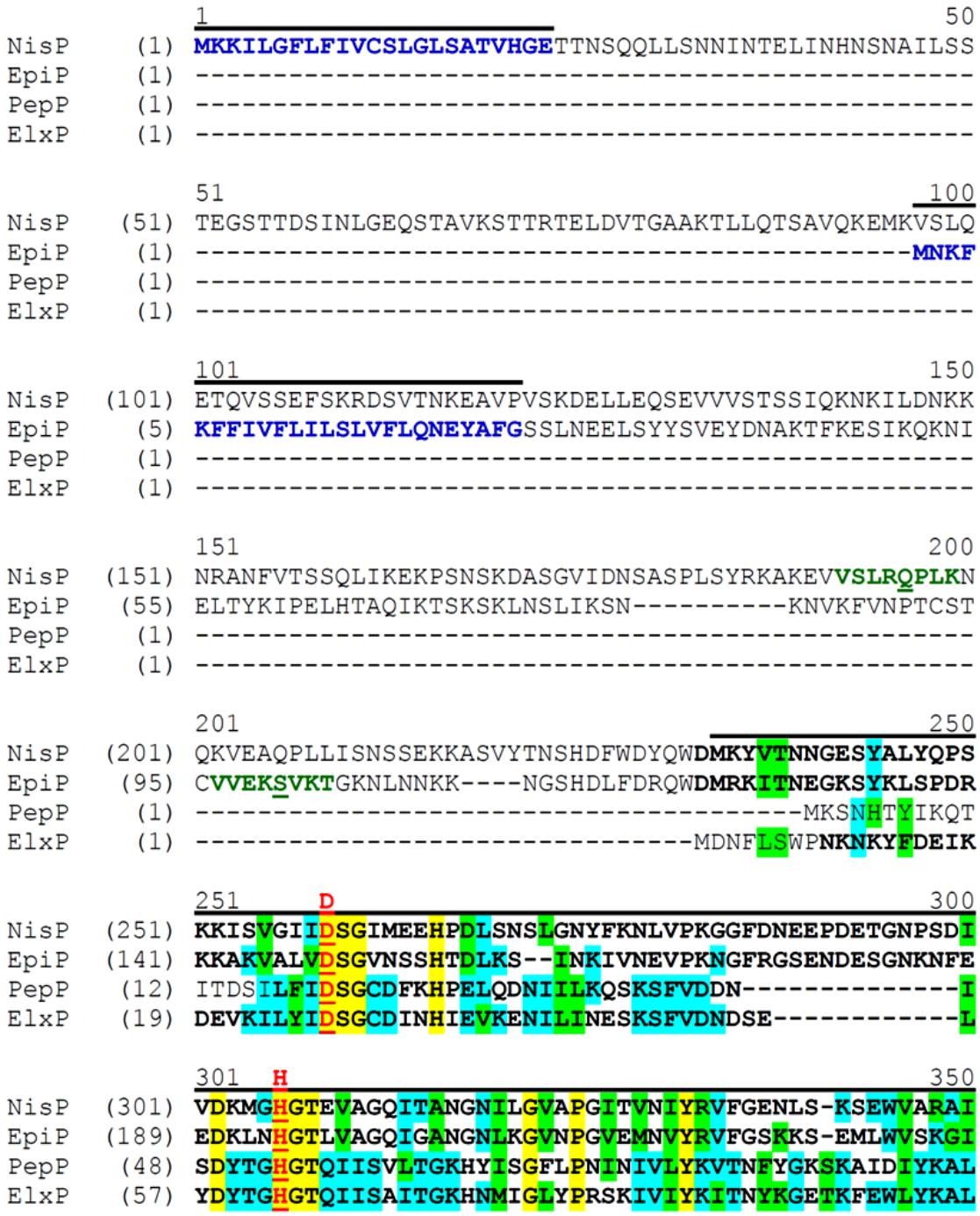


Figure 4.3. Sequence alignment of representative proteases from the different classes of LanP proteins. The N-terminal signal peptide of NisP and EpiP as well as the C-terminal anchor signal in NisP are shown in blue bold letters. The putative NisP and EpiP autocleavage sites after the pro-sequences are shown in green bold letters and the starting amino acids in the mature proteins are underlined. The central subtilisin-like serine protease domains are shown in black bold letters and the residues of the catalytic triad are shown in red bold letters. NisP: protease for nisin A, EpiP: protease for epidermin, PepP: protease for Pep5, and ElxP: protease for epilancin 15X.

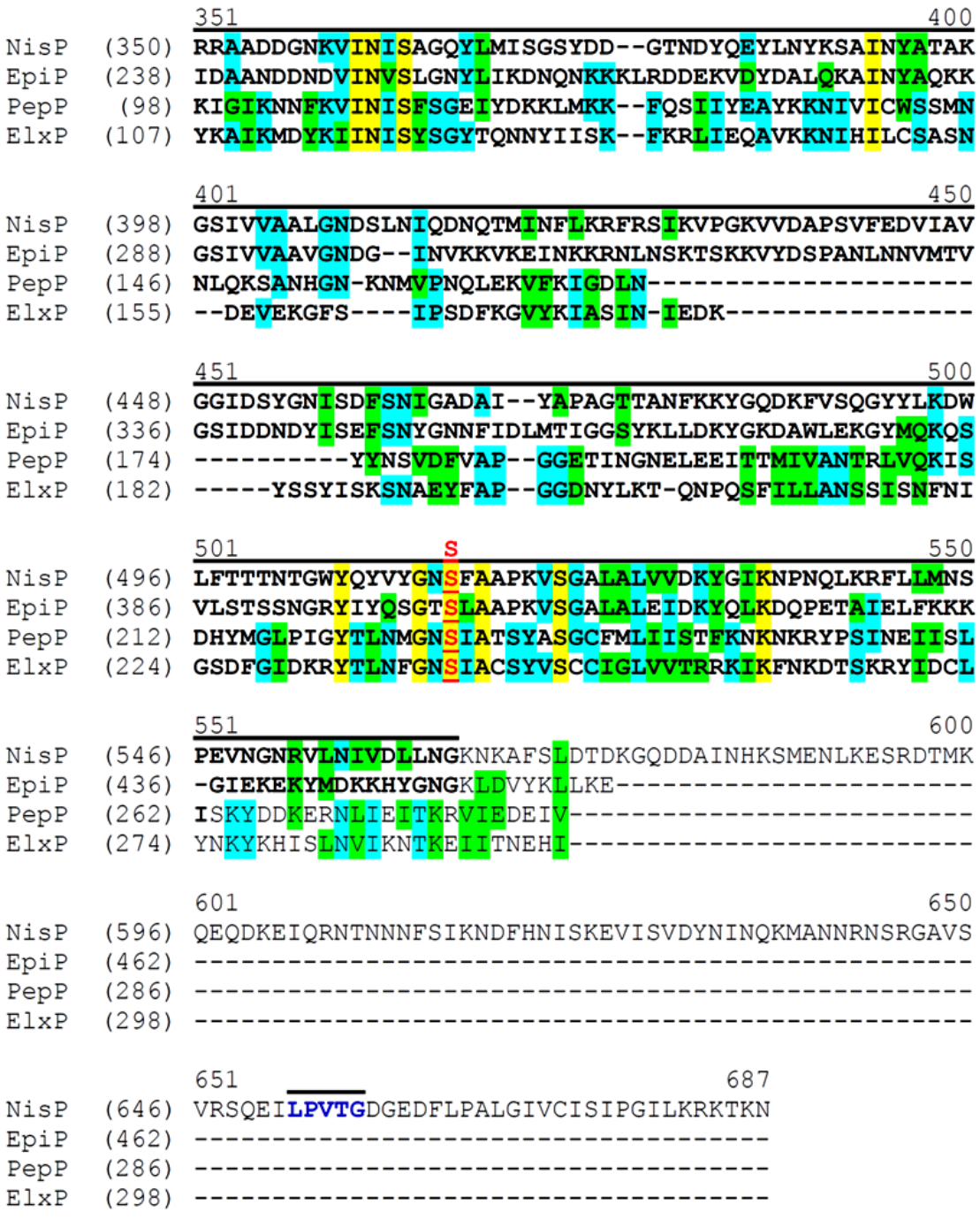


Figure 4.3. *Continuation*.

Despite considerable efforts in the study of lantipeptide biosynthesis, the mechanism and substrate specificity of LanP proteases have remained elusive, mainly due to the difficulties in the overexpression of these enzymes that are commonly toxic to

the heterologous hosts and the reconstitution of their enzymatic activities *in vitro* under cell extract- or culture supernatant-free conditions. In a previous attempt, an *Escherichia coli* host carrying a plasmid encoding for NisP was able to express the protease, albeit at low concentrations, based on a SDS-PAGE analysis after selective labeling of plasmid-encoded proteins with ^{35}S-Met (*20*). Although NisP was never purified, the *E. coli* cell extracts were able to cleave the nisin A cyclized precursor peptide, producing a biologically active compound (Table 4.1, entry 1) (*20*). Even though a large variety of fully modified NisA with mutations across the core sequence, including the ring A, are recognized by NisP (*23*), several studies have evidenced that mutations across the leader region are not as well tolerated (Table 4.1). For instance, when a gene encoding for the chimera subtilin(leader)/nisinZ(core) was transformed into a nisin A producer strain, only the cyclized peptide with the leader attached was detected, presumably because of the alteration at the cleavage site R−1Q (Table 4.1, entry 2) (*24*). Furthermore, mutations at Arg−1 and Ala−4 in nisin Z leader peptide, but not at the widely conserved Pro−2 and other amino acids, greatly affected amide bond hydrolysis *in vivo* (Table 4.1, entries 3 to 7) (*25*). In contrast, a more recent study evidenced that NisP-expressing *Lactococcus lactis* cells cleave the leader from fully modified NisA with double mutations at Ala−4 and Pro−2 or from unmodified NisA with the putative proNisP autocleavage site, VSLR-QP (Table1, entries 13 to 17) (*26*). However, a different study suggested that NisP is able to cleave the leader region only from fully processed precursor peptide, but not from dehydrated or unmodified NisA (Table 4.1, entries 8 to 12) (*27*).

Regarding other lantibiotic proteases, *in vivo* studies with Pep5 indicated that its protease PepP does not recognize the unmodified precursor peptide, but can process the dehydrated peptide, suggesting that dehydration of Thr1 may be important for substrate recognition (Table 4.1, entries 18 and 19) (*21*). In contrast, culture supernatants of *S. carnosus* TM300 expressing EpiP processed unmodified EpiA to the expected proteolytic products, but not the mutant peptide R−1Q (Table 4.1, entries 20 and 21) (*19*). Thus, the sequence of the amino acids near the cleavage site in the leader peptide and the structure of the cyclized core region are likely determinants for the recognition of the substrates by LanP proteases, including ElxP. Although these results regarding substrate specificity have provided valuable insights into the importance of certain leader peptide residues for proteolytic processing by LanP proteins, a comprehensive analysis on the importance of leader peptide residues in proteolysis is not available.

The epilancin 15X precursor peptide (ElxA), like other lantibiotic precursors (*28, 29*), is expected to fold in an α-helical configuration between residues Lys−7 and Asp−12 upon binding to biosynthetic proteins based on computational predictions (GOR V (*30*) and PredictProtein (*31*)). Disruption of a similar helical region in LctA, the precursor peptide of the class II lantibiotic lacticin 481, greatly affects the activity of the protease domain of LctT (*6*). Although LctT is not related to ElxP by sequence (Figure 4.2), the secondary structure of ElxA may also be important for proteolysis, in addition to the primary amino acid sequence as described above. Herein, we report the first *in vitro* analysis of the substrate scope of a serine-type lantibiotic protease. The importance of the partially conserved motifs F(D/N)L(D/N) at position −19 and Pro at

position −2, common to class I lantipeptide precursor peptides (Figure 4.4), and the leader peptide secondary structure were evaluated. Furthermore, homology modeling of ElxP, based on the crystal structures of other subtilisin-like proteases, provided additional insights into the substrate scope of the peptidase. Based on these results, protein engineering could be used to design peptidases with a desired specificity.

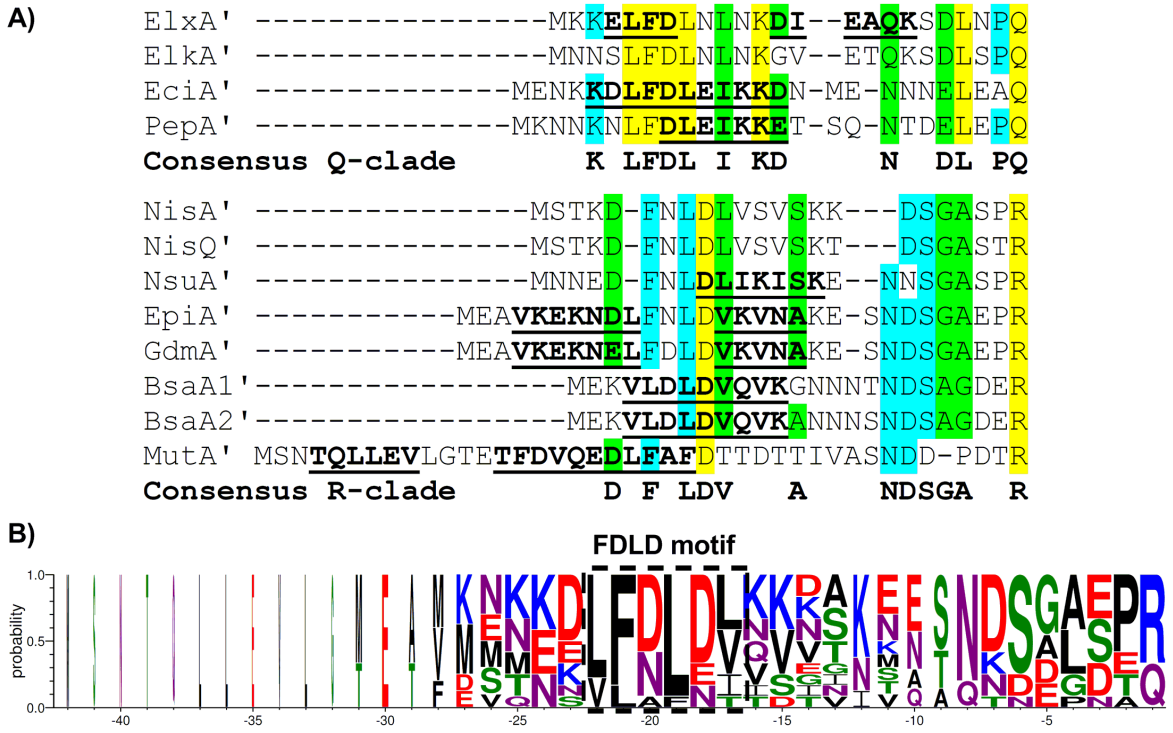


Figure 4.4. Leader peptides of selected class I lantibiotics. A) Sequence alignment of the leader regions of class I lantibiotic precursor peptides processed by proteases from the Q- and R-clades (Figure 4.2). The predicted helical regions are shown in bold and underlined. B) Sequence logo generated using all the leader peptide sequences from A). The partially conserved FDLD motif is highlighted. The logo shows the probability of each amino acid (height of the letter) and it is scaled (width of the letter) according to the number of sequences considered for a specific position (i.e. narrow letters were generated from a smaller number of sequences than wider letters). The alignment was performed using ClustalW (*32*), the sequence logo with WebLogo 3.1 (*33*), and the helical regions were predicted with GOR V (*30*). ElxA': Leader peptide of epilancin 15X; ElkA': Leader peptide of epilancin K7; EciA': Leader peptide of epicidin 280; PepA': Leader peptide of Pep5; NisA': Leader peptide of nisin A, nisin F, nisin Z; NisQ': Leader peptide of nisin Q; NsuA': Leader peptide of nisin U; EpiA': Leader peptide of epidermin; GdmA': Leader peptide of gallidermin; MutA': Leader peptide of mutacin I and mutacin III/1140; BsaA1' and BsaA2': Leader peptide of the Bsa group of lantibiotics.

Table 4.1. Cleavage of mutated lantibiotic precursor peptides by LanP proteases. Mutations in the precursor peptide sequences in comparison with the wild-type peptides are highlighted in red and the cleavage sites are shown in bold letters. Dha: 2,3-dehydroalanine, Dhb: 2,3-dehydrobutyrine, Lan: lanthionine.

Entry	LanP	Precursor Peptide	Cleavage site (Leader-Core)	Core peptide modifications	Assay description	Cleavage	Reference
1	NisP	NisA	SGASP**R-I**(Dhb)(Lan)	Fully modified	*In vitro* with cell extract	Yes	(*20*)
2	NisP	SpaS(leader)-NisZ(core)	SKITP**Q-I**(Dhb)(Lan)	Fully modified	*In vivo*	No	(*24*)
3	NisP	NisZ	SGASP**R-I**(Dhb)(Lan)	Fully modified	*In vivo*	Yes	(*25*)
4	NisP	NisZ	SGASP**Q-I**(Dhb)(Lan)	Fully modified	*In vivo*	No	(*25*)
5	NisP	NisZ	SGASG**R-I**(Dhb)(Lan)	Fully modified	*In vivo*	Yes	(*25*)
6	NisP	NisZ	SGASA**R-I**(Dhb)(Lan)	Fully modified	*In vivo*	Yes	(*25*)
7	NisP	NisZ	SGDSP**R-I**(Dhb)(Lan)	Fully modified	*In vivo*	No	(*25*)
8	NisP	NisA	SGASP**R-I**(Dhb)(Lan)	Fully modified	*In vivo*	Yes	(*27*)
9	NisP	NisA	SGASP**R-I**TS	Unmodified	*In vivo*	No	(*27*)
10	NisP	NisA	PGASL**R-I**(Dhb)(Dha)	Dehydrated	*In vivo*	No	(*27*)
11	NisP	NisA	PGASL**R-I**(Dhb)(Lan)[a]	Dehydrated	*In vivo*	Yes	(*27*)
12	NisP	NisA(leader)-Angiotensin (1-7)	SGASP**R-N**R(Dha)	Dehydrated	*In vivo*	No	(*27*)
13	NisP	NisA[b]	SGASP**R-I**(Dhb)(Lan)	Fully modified	*In vitro* with cell extract	Yes	(*26*)
14	NisP	NisA[b]	SGVSL**R-I**(Dhb)(Lan)	Fully modified	*In vitro* with cell extract	Yes	(*26*)
15	NisP	NisA[b]	SGVSL**R-I**TS	Unmodified	*In vitro* with cell extract	No	(*26*)
16	NisP	NisA[b]	SGVSL**R-Q**P(Lan)	Partially modified	*In vitro* with cell extract	Yes	(*26*)
17	NisP	NisA[b]	SGVSL**R-Q**PS	Unmodified	*In vitro* with cell extract	Yes	(*26*)
18	PepP	PepA	DELEP**Q-(Dhb)**AG	Fully modified	*In vivo*	Yes	(*21*)
19	PepP	PepA	DELEP**Q-T**AG	Unmodified	*In vivo*	No	(*21*)
20	EpiP	EpiA	SGAEP**R-I**AS	Unmodified	*In vitro* with culture supernatant	Yes	(*19*)
21	EpiP	EpiA	SGAEP**Q-I**AS	Unmodified	*In vitro* with culture supernatant	No	(*19*)

Notes: *a*: Ring topology was not confirmed.
b: Culture supernatant containing modified or unmodified NisA was used

4.2. EXPERIMENTAL METHODS

Materials, organisms, media, and growth conditions

Chemical reagents and media components used in this study were purchased from Sigma-Aldrich or Thermo Fisher Scientific and were used without further purification. All *E. coli* strains were grown at 37 °C in Luria-Bertani (LB) liquid or solid media with the addition of the appropriate antibiotics, unless otherwise specified. The list of strains and plasmids used in this study are shown in Table 4.2.

Construction of mutant pET.His_6-ElxA plasmids and overexpression and purification of wild-type and mutant His_6-ElxA peptides

Site directed mutagenesis was used to generate the mutant pET.His_6-ElxA plasmids (Table 4.2). Briefly, the entire pET.His_6-ElxA was amplified by PCR using iProof™ high-fidelity DNA polymerase (BioRad) and the appropriate mutagenesis primers (Table 4.3), followed by treatment with *Dpn*I (New England Biolabs) before transformation of *E. coli* DH5α cells. The correct sequences of the inserts were confirmed by sequencing at the W. M. Keck Center for Comparative and Functional Genomics at the University of Illinois at Urbana-Champaign. See also *Notebook IX, page 03*.

Table 4.2. Microorganisms and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source or reference
Strains		
E. coli DH5*α*	*λpir/ϕ80dlacZΔM15* Δ(*lacZYA-argF*)*U169 recA1 hsdR17 deoR thi-1 supE44 gyrA96 relA1*	(*34*)
E. coli Rosetta2(DE3)	F^- *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) λ(DE3 [*lacI lacUV5-T7 gene 1 ind1 sam7 nin5*]) pRARE2	Novagen
Plasmids		
pET.His_6-MBP-ElxP	*elxP* cloned into pET28b vector	(*22*)
pET.His_6-ElxA	*elxA* cloned into a pET28b plasmid (a Kan^R *E. coli* T7 based hexahistidine-tag fusion expression vector)	(*22*)
pET.His_6-ElxA(Q−1A)	Mutant pET.His_6-ElxA encoding for His_6-ElxA Q−1A	This study
pET.His_6-ElxA(Q−1E)	Mutant pET.His_6-ElxA encoding for His_6-ElxA Q−1E	This study
pET.His_6-ElxA(Q−1R)	Mutant pET.His_6-ElxA encoding for His_6-ElxA Q−1R	This study
pET.His_6-ElxA(P−2A)	Mutant pET.His_6-ElxA encoding for His_6-ElxA P−2A	This study
pET.His_6-ElxA(P−2D)	Mutant pET.His_6-ElxA encoding for His_6-ElxA P−2D	This study
pET.His_6-ElxA(P−2K)	Mutant pET.His_6-ElxA encoding for His_6-ElxA P−2K	This study
pET.His_6-ElxA(A−9P)	Mutant pET.His_6-ElxA encoding for His_6-ElxA A−9P	This study
pET.His_6-ElxA(E−10A)	Mutant pET.His_6-ElxA encoding for His_6-ElxA E−10A	This study
pET.His_6-ElxA(E−10P)	Mutant pET.His_6-ElxA encoding for His_6-ElxA E−10P	This study
pET.His_6-ElxA(N−16A)	Mutant pET.His_6-ElxA encoding for His_6-ElxA N−16A	This study
pET.His_6-ElxA(N−16D)	Mutant pET.His_6-ElxA encoding for His_6-ElxA N−16D	This study
pET.His_6-ElxA(L−17A)	Mutant pET.His_6-ElxA encoding for His_6-ElxA L−17A	This study
pET.His_6-ElxA(D−18A)	Mutant pET.His_6-ElxA encoding for His_6-ElxA D−18A	This study
pET.His_6-ElxA(D−18N)	Mutant pET.His_6-ElxA encoding for His_6-ElxA D−18N	This study
pET.His_6-ElxA(F−19A)	Mutant pET.His_6-ElxA encoding for His_6-ElxA F−19A	This study

Electrocompetent *E. coli* Rosetta2(DE3) cells were transformed with pET.His_6-ElxA or mutant plasmids (Table 4.2) and single colonies were inoculated in 5 mL of LB medium containing 50 μg/mL kanamycin and 12.5 μg/mL chloramphenicol and grown for 12 h with shaking. An aliquot of 3 mL was used to inoculate 300 mL LB medium with the appropriate antibiotics and the cells were grown at 37 °C until OD_{600} = 1.0 AU. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5

mM and the cultures were incubated for 16 h at 18 °C. Attempts to overexpress the peptides at 37 °C, as suggested before for other lantibiotic precursor peptides (*35*), were not successful, since His_6-ElxA is highly toxic to the host, most likely related to peptide glutathionylation by *E. coli* glutathione S transferases (Dr. Bradly Evans, personal communication). After centrifugation, the cell pellets were resuspended in denaturing buffer (6 M guanidine hydrochloride, 50 mM NaH_2PO_4, 20 mM imidazole, 500 mM NaCl, pH 7.5), followed by incubation at 4 °C with shaking for 1 h, sonication and centrifugation. The supernatants were mixed with 5 mL of 50% Ni-NTA (nitrilotriacetic acid) slurry (Qiagen) for 60 min. Unbound material was discarded and the resin was washed with 15 column volumes (CV) of washing buffer (6 M guanidine hydrochloride, 50 mM NaH_2PO_4, 20 mM imidazole, 500 mM NaCl, pH 7.5). The peptides were then eluted with 4 CV of elution buffer (4 M guanidine hydrochloride, 50 mM NaH_2PO_4, 500 mM imidazole, 500 mM NaCl, pH 7.5). Finally, the peptides were purified by solid phase extraction (SPE) using Varian Bond Elut C18EWP columns (Agilent). After loading the peptide-containing solutions, the resin was washed with 0.1% TFA in 20% acetonitrile / 80% water and the peptides were eluted with 0.1% TFA in 80% acetonitrile / 20% water. The masses of the purified peptides were determined by matrix-assisted laser desorption/ionization – time-of-flight mass spectrometry (MALDI-TOF) on a Voyager DE-STR Biospectrometry Workstation using α-cyano-4-hydroxycinnamic acid as matrix at the Mass Spectrometry Laboratory of the University of Illinois at Urbana Champaign. A sample or His_6-NisA was donated by Ms. Nega Garg, a current graduate student at the van der Donk lab at the University of Illinois. ElxP was

overexpressed and purified following a protocol described in chapter 2 (*22*). See also *Notebook IX, page 13*.

Cleavage of His_6-ElxA substrates by ElxP

ElxP (2 μM) and wild-type or mutant His_6-ElxA peptides (50 μM) were incubated in the presence of assay buffer (50 mM HEPES, pH 7.5) at room temperature for 2 h, unless otherwise specified. Control samples lacking protease were incubated under the same conditions. The reactions were quenched by the addition of 1% TFA to a final concentration of 0.5%, and the samples were prepared by C18 Zip-tip (Millipore) before analysis by MALDI-TOF MS as described above. See also *Notebook IX, page 15*.

Table 4.3. Primers used in this study

Name	Sequence
ElxA.Q−1A.FP	5'- GCA CAA AAA AGT GAC CTA AAT CCG GCA TCA GCT AGT ATT GTT AAA ACA AC -3'
ElxA.Q−1A.RP	5'- GTT GTT TTA ACA ATA CTA GCT GAT GCC GGA TTT AGG TCA CTT TTT TGT GC -3'
ElxA.Q−1E.FP	5'- CAA AAA AGT GAC CTA AAT CCG GAA TCA GCT AGT ATT GTT AAA A -3'
ElxA.Q−1E.RP	5'- TTT TAA CAA TAC TAG CTG ATT CCG GAT TTA GGT CAC TTT TTT G -3'
ElxA.Q−1R.FP	5'- GCA CAA AAA AGT GAC CTA AAT CCG CGA TCA GCT AGT ATT GTT AAA ACA AC -3'
ElxA.Q−1R.RP	5'- GTT GTT TTA ACA ATA CTA GCT GAT CGC GGA TTT AGG TCA CTT TTT TGT GC -3'
ElxA.P−2A.FP	5'- GCA CAA AAA AGT GAC CTA AAT GCG CAA TCA GCT AGT ATT GTT AAA ACA AC -3'
ElxA.P−2A.RP	5'- GTT GTT TTA ACA ATA CTA GCT GAT TGC GCA TTT AGG TCA CTT TTT TGT GC -3'
ElxA.P−2D.FP	5'- GGC ACA AAA AAG TGA CCT AAA TGA CCA ATC AGC TAG TAT TGT TAA AAC AAC -3'
ElxA.P−2D.RP	5'- GTT GTT TTA ACA ATA CTA GCT GAT TGG TCA TTT AGG TCA CTT TTT TGT GCC -3'
ElxA.P−2K.FP	5'- CGA GGC ACA AAA AAG TGA CCT AAA TAA GCA ATC AGC TAG TAT TGT TAA AAC AAC -3'
ElxA.P−2K .RP	5'- GTT GTT TTA ACA ATA CTA GCT GAT TGC TTA TTT AGG TCA CTT TTT TGT GCC TCG -3'
ElxA.A−9P.FP	5'- GAT TTA AAT CTT AAT AAA GAT ATC GAG CCA CAA AAA AGT GAC CTA AAT CCG C -3'
ElxA.A−9P.RP	5'- GCG GAT TTA GGT CAC TTT TTT GTG GCT CGA TAT CTT TAT TAA GAT TTA AAT C -3'
ElxA.E−10A.FP	5'- GAT TTA AAT CTT AAT AAA GAT ATC GCG GCA CAA AAA AGT GAC CTA AAT CCG -3'
ElxA.E−10A.RP	5'- CGG ATT TAG GTC ACT TTT TTG TGC CGC GAT ATC TTT ATT AAG ATT TAA ATC -3'
ElxA.E−10P.FP	5'- GAT TTA AAT CTT AAT AAA GAT ATC CCG GCA CAA AAA AGT GAC CTA AAT CCG -3'
ElxA.E−10P.RP	5'- CGG ATT TAG GTC ACT TTT TTG TGC CGG GAT ATC TTT ATT AAG ATT TAA ATC -3'
ElxA.N−16A.FP	5'- GCC ATA TGA AAA AAG AAT TAT TTG ATT TAG CTC TTA ATA AAG ATA TCG AGG CAC -3'
ElxA.N−16A.RP	5'- GTG CCT CGA TAT CTT TAT TAA GAG CTA AAT CAA ATA ATT CTT TTT TCA TAT GGC -3'
ElxA.N−16D.FP	5'- GCC ATA TGA AAA AAG AAT TAT TTG ATT TAG ATC TTA ATA AAG ATA TCG AGG CAC -3'
ElxA.N−16D.RP	5'- GTG CCT CGA TAT CTT TAT TAA GAT CTA AAT CAA ATA ATT CTT TTT TCA TAT GGC -3'

Table 4.3. *Continuation*.

Name	Sequence
ElxA.L−17A.FP	5'- GCC ATA TGA AAA AAG AAT TAT TTG ATG CAA ATC TTA ATA AAG ATA TCG AGG C -3'
ElxA.L−17A.RP	5'- GCC TCG ATA TCT TTA TTA AGA TTT GCA TCA AAT AAT TCT TTT TTC ATA TGG C -3'
ElxA.D−18A.FP	5'- CAG CCA TAT GAA AAA AGA ATT ATT TGC TTT AAA TCT TAA TAA AGA TAT CGA GGC -3'
ElxA.D−18A.RP	5'- GCC TCG ATA TCT TTA TTA AGA TTT AAA GCA AAT AAT TCT TTT TTC ATA TGG CTG -3'
ElxA.D−18N.FP	5'- GGC AGC CAT ATG AAA AAA GAA TTA TTT AAT TTA AAT CTT AAT AAA GAT ATC GAG GC -3'
ElxA.D−18N.RP	5'- GCC TCG ATA TCT TTA TTA AGA TTT AAA TTA AAT AAT TCT TTT TTC ATA TGG CTG CC -3'
ElxA.F−19A.FP	5'- GGC AGC CAT ATG AAA AAA GAA TTA GCT GAT TTA AAT CTT AAT AAA GAT ATC G -3'
ElxA.F−19A.RP	5'- CGA TAT CTT TAT TAA GAT TTA AAT CAG CTA ATT CTT TTT TCA TAT GGC TGC C -3'

4.3. RESULTS

Enzymatic activity of ElxP with His_6-ElxA mutant peptides

ElxP is a serine protease that catalyzes the cleavage of the leader peptide from fully cyclized precursor peptide ElxA (Figure 4.1) (*22*). To determine the role of the leader peptide sequence and the FDLD and PQ motifs (Figure 4.4) in proteolysis, a series of mutant substrates was expressed and purified from *E. coli* and tested as potential substrates under identical conditions (Table 4.4 and Figure 4.5). Since fully modified mutant ElxA peptides were not available, linear unmodified peptides were used instead, as previous experimental results have evidenced that ElxP is able to process linear wild-type His_6-ElxA (chapter 2) (*22*). The reaction products were analyzed by MALDI-TOF MS and peaks corresponding to the N-terminal hexahistidine tagged leader peptides and to the C-terminal unmodified core peptide (calculated mass 3316 Da) were only detected when ElxP processed the substrates. Smaller peaks with a difference in

mass of +178 Da with respect to the calculated masses for the precursor and leader peptides were also observed. These peaks correspond to the gluconoyl adducts generated by modification of the N-terminal α-amino group of the GSS-His-tag during peptide expression in *E. coli* (*36*).

Table 4.4. MALDI-TOF MS analysis of the cleavage products after incubation of wild-type or mutant His_6-ElxA peptides with or without ElxP for 2 h under identical conditions. The calculated masses are shown in parenthesis.

Entry	Peptide	Precursor peptide (calcd)	Leader peptide (calcd)	Core peptide (calcd)	Result[a]
1	His_6-ElxA wt	8161 (8162)	4866 (4864)	3316 (3316)	+++
2	Q−1A	8107 (8105)	4809 (4807)	3315 (3316)	+
3	Q−1E	8162 (8163)	4864 (4865)	3314 (3316)	+++
4	Q−1R	8190 (8190)	4890 (4892)	3314 (3316)	+
5	P−2A	8134 (8136)	4839 (4838)	3315 (3316)	+
6	P−2D	8179 (8180)	– (4882)	– (3316)	–
7	P−2K	8193 (8193)	– (4895)	– (3316)	–
8	N−16A	8119 (8119)	4819 (4821)	3314 (3316)	+++
9	N−16D	8161 (8163)	4865 (4865)	3314 (3316)	+++
10	L−17A	8118 (8120)	4823 (4822)	3315 (3316)	+++
11	D−18A	8118 (8118)	4818 (4820)	3314 (3316)	+++
12	D−18N	8159 (8161)	4863 (4863)	3316 (3316)	+++
13	F−19A	8086 (8086)	4789 (4788)	3315 (3316)	++
14	A−9P	8189 (8188)	4889 (4890)	3317 (3316)	++
15	E−10A	8105 (8104)	4805 (4806)	3316 (3316)	+++
16	E−10P	8132 (8130)	4833 (4832)	3317 (3316)	++
17	His_6-NisA wt	7444 (7444)	– (3964)	– (3498)	–

Note: a: The efficiency of the enzymatic activity was determined qualitatively based on the intensity of the leader and core peptide ions relative to the precursor peptide ion and it is reported in a scale from +++ for highly efficiently processed substrates compared to the wild-type to – when no cleavage products were detected.

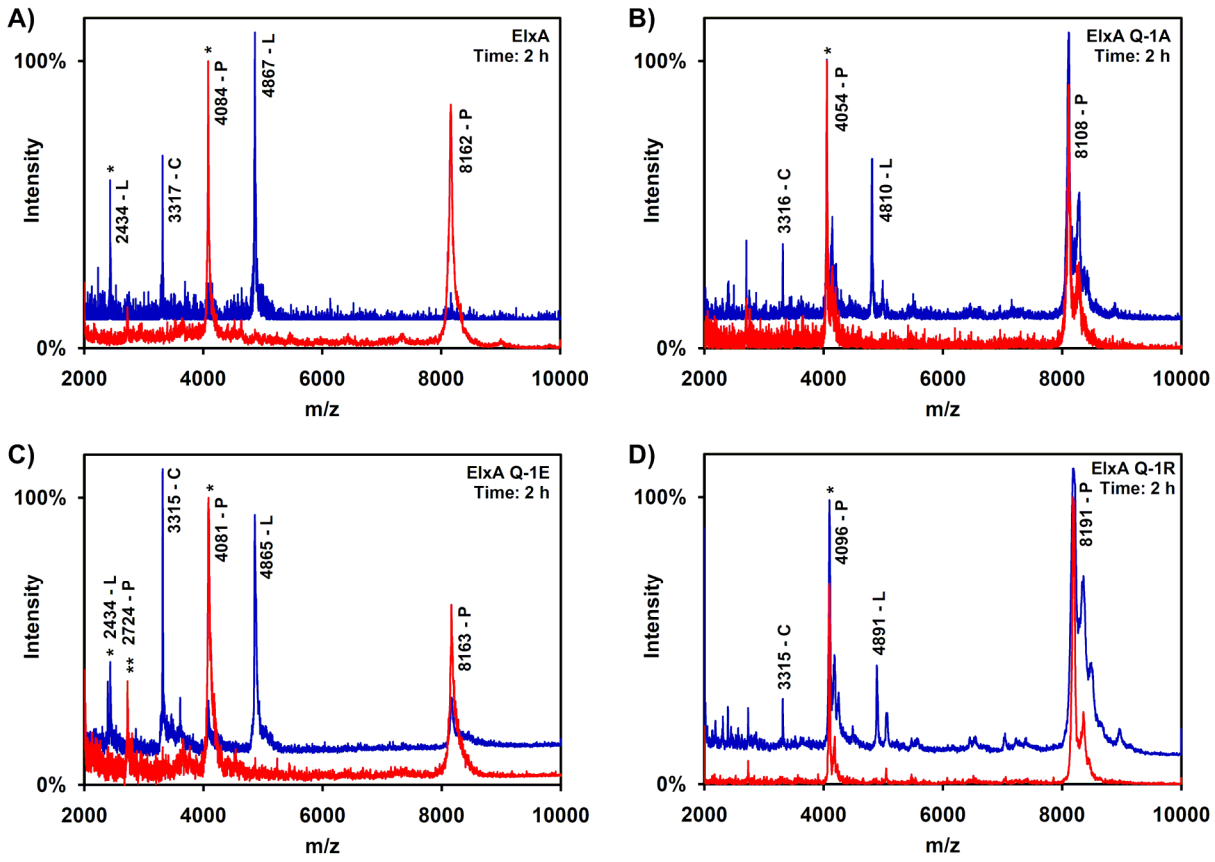


Figure 4.5. MALDI-TOF MS analysis of proteolytic processing of wild-type His_6-ElxA (A), His_6-ElxA Q−1A (B), His_6-ElxA Q−1E (C), and His_6-ElxA Q−1R (D) after incubation with (blue) or without (red) ElxP for 2 h. P, L, and C denotate precursor, leader, and core peptide, respectively. The single (*) and double (**) asterisks indicate peaks corresponding to double- or triple-charged ions, respectively.

Since mutations at the cleavage site of NisA and EpiA were shown to affect amide bond hydrolysis by NisP or EpiP (*19, 24, 25*), Gln−1 in wild-type ElxA was mutated to Ala or to charged residues (Glu, Arg) and the resulting peptides were tested as potential substrates (Table 4.4, entries 2 to 4 and Figure 4.5). Although all the mutants were accepted as substrates by the protease, the peptides Q−1A and Q−1R were processed to a much lesser extent than the wild-type or the mutant peptide Q−1E when incubated under identical conditions based on the intensity of the peaks in the MALDI-TOF MS spectra. Similarly, the highly conserved Pro−2 of ElxA was mutated to

Ala, Asp, or Lys (Table 4.4, entries 5 to 7 and Figure 4.6). Interestingly, with the substrates P−2D and P−2K a considerably decreased proteolytic activity was observed and the cleavage products were only detected after long reaction periods (30 h). Similarly, the mutant P−2A did not prevent proteolysis, but did reduce the enzyme efficiency, as only small amounts of cleavage products were detected under conditions that led to full conversion of His_6-ElxA.

To determine the importance of other conserved residues in the leader sequence, the single mutants F−19A, D−18A, D−18N, L−17A, N−16A, and N−16D were investigated (Table 4.4, entries 8 to 13 and Figure 4.7). Although we have not obtained quantitative kinetic parameters, all the mutant substrates were qualitatively processed with very similar efficiencies as wild-type His_6-ElxA based on relative intensities of the substrate and product ions observed by MALDI-TOF MS. The mutant F−19A was well processed, but the cleavage efficiency was reduced to some extent presumably due to experimental deviations.

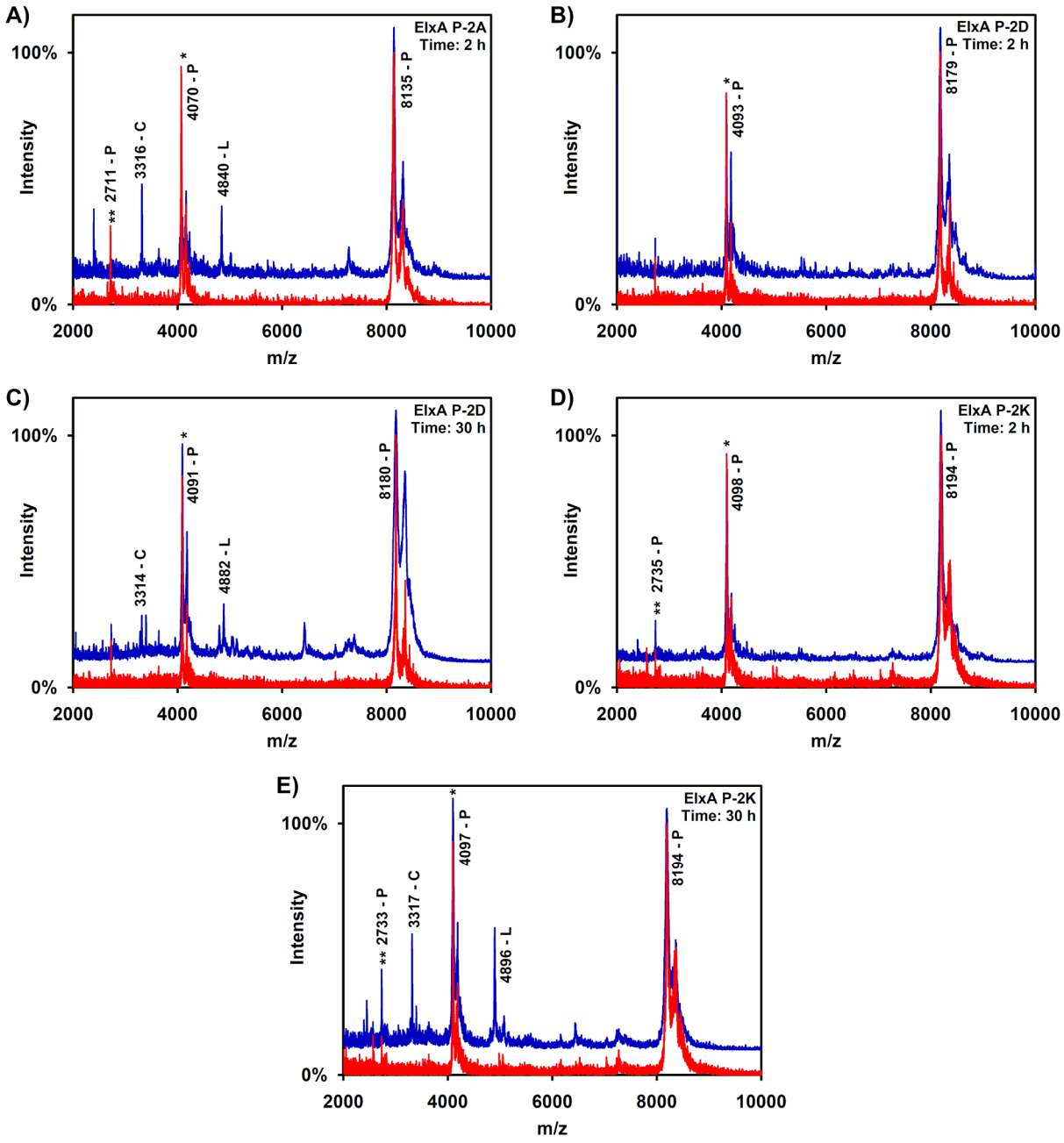


Figure 4.6. MALDI-TOF MS analysis of proteolytic processing of His_6-ElxA P−2A (A), His_6-ElxA P−2D (B and C), and His_6-ElxA P−2K (D and E) after incubation with (blue) or without (red) ElxP for 2 h (A, B and D) or 30 h (C and E). P, L, and C denotate precursor, leader, and core peptide, respectively. The single (*) and double (**) asterisks indicate peaks corresponding to double- or triple-charged ions, respectively.

The leader peptides of several class I lantibiotics, including gallidermin, epidermin, Pep5, and nisin, were previously suggested to adopt a helical conformation based on circular dichroism experiments (*28*). These studies were performed dissolving synthetic peptides in solutions containing trifluoroethanol that is known to promote helicity. However, biosynthetic proteins, including proteases, may also induce and recognize similar helical structures as demonstrated for the precursor peptide ComC upon binding to the peptidase domain of ComA (*9, 10*) and as suggested for LctA and the corresponding LctT protease domain (*6*). Secondary structure predicting programs anticipate α-helical character for the ElxA leader peptide (GOR V (*30*), residues −7 to −12 and −18 to −21; PredictProtein (*31*), residues −7 to −11 and −19 to −22), as well as for several other class I leader peptides (Figures 4.4 and 4.8). To assess if the predicted α-helical structure of ElxA is required for substrate recognition, the helix-breaking residue Pro was introduced at positions −9 or −10 (Figure 4.8). Interestingly, the mutants A−9P and E−10P did not prohibit the peptidase activity of ElxP, although they were processed to a lesser degree than the wild-type and the mutant E−10A under identical assay conditions, indicating that the predicted helical conformation is not strictly required for substrate-enzyme interaction (Table 4.4, entries 14 to 16, Figure 4.9).

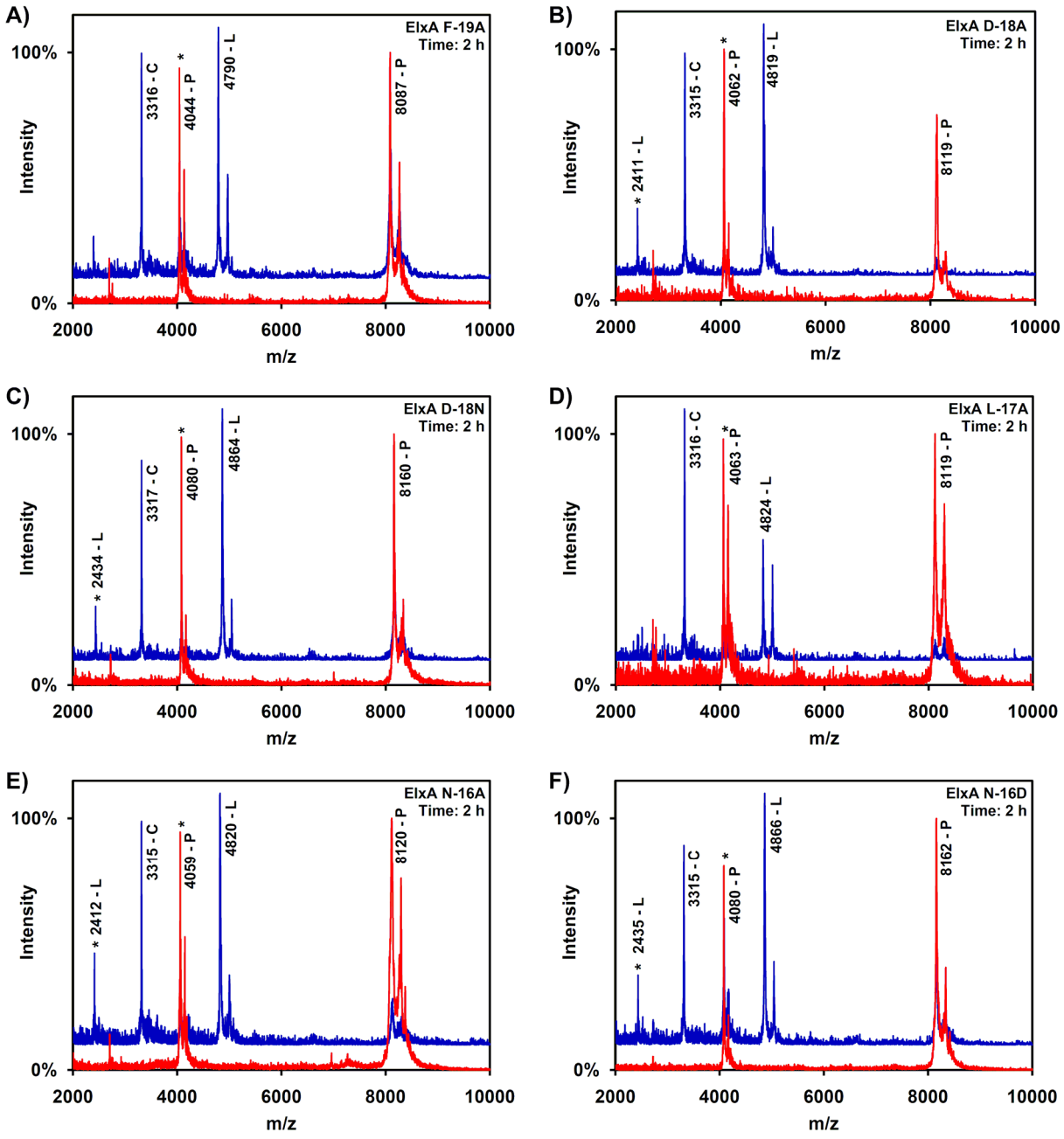


Figure 4.7. MALDI-TOF MS analysis of proteolytic processing of His_6-ElxA F−19A (A), His_6-ElxA D−18A (B), His_6-ElxA D−18N (C), His_6-ElxA L−17A (D), His_6-ElxA N−16A (E), and His_6-ElxA N−16D (F) after incubation with (blue) or without (red) ElxP for 2 h. P, L, and C denotate precursor, leader, and core peptide, respectively. The asterisk (*) indicates a peak corresponding to a double-charged ion.

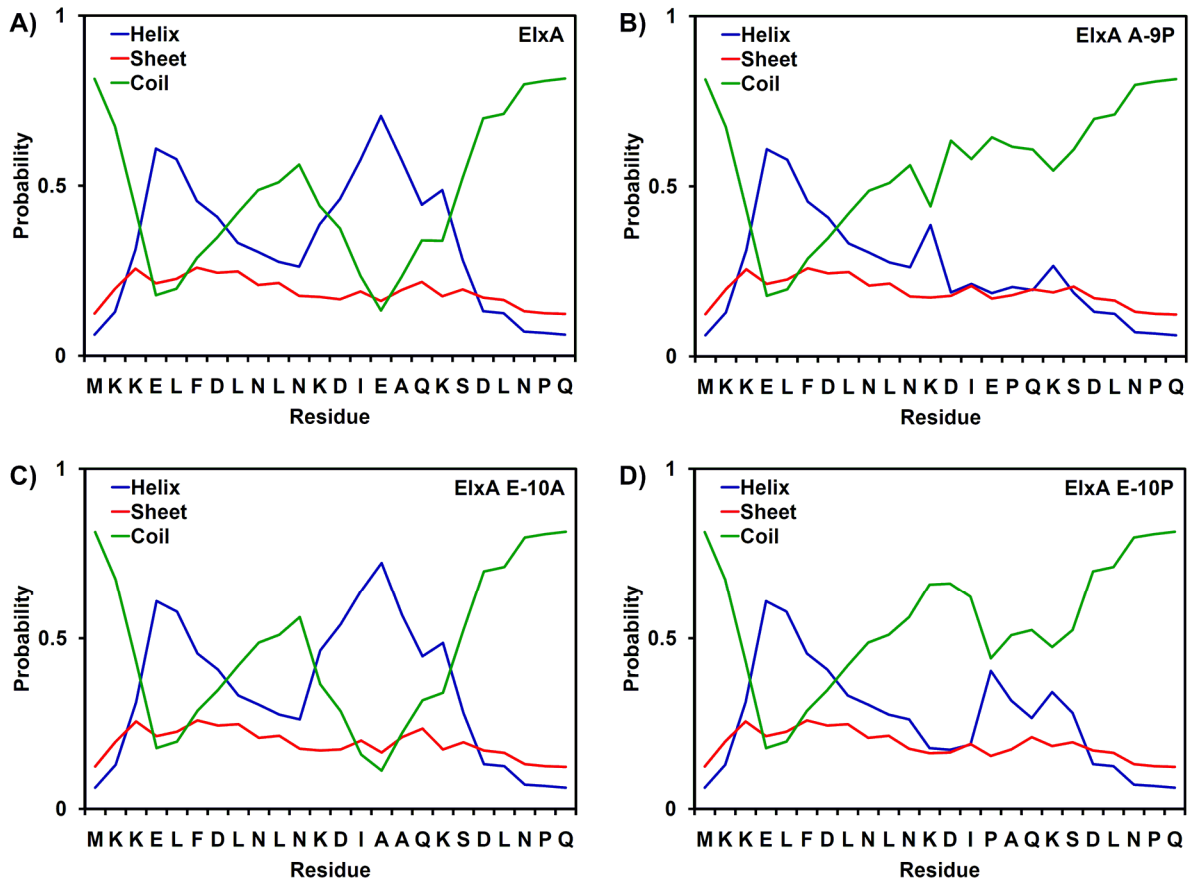


Figure 4.8. Predicted secondary structure of the leader region of ElxA (A) and the mutants A−9P (B), E−10A (C), and E−10P (D). The probability of α-helix (blue), β-sheet (red) or random coil (green) conformations are shown for every residue. Upon mutation of Ala−9 or Glu−10 to Pro, the helical structure at the C-terminus of the leader peptide is predicted to disappear. The predictions were performed using GOR V (*30*).

The leader peptides of NisA and ElxA share some sequence homology, with 7 out 23 residues being identical. Although these two leader peptides differ in the last residue (Gln−1 for ElxA as opposed to Arg−1 for NisA) and in the sequence of the FDLD motif (FDLN for ElxA and FNLD for NisA), the peptide His_6-ElxA Q−1R was still processed by ElxP to some extent and the mutants D−18N and N−16D were processed with very similar efficiencies as wild-type ElxA. Thus, it was anticipated that ElxP could cleave the linear unmodified NisA peptide. However, after incubation of His_6-NisA with the peptidase, the predicted cleavage products were not detected by MALDI-TOF MS

even after long incubation periods (24 h) (Table 4.4, entry 17 and Figure 4.10), suggesting that additional residues in the precursor peptide may be important for substrate recognition.

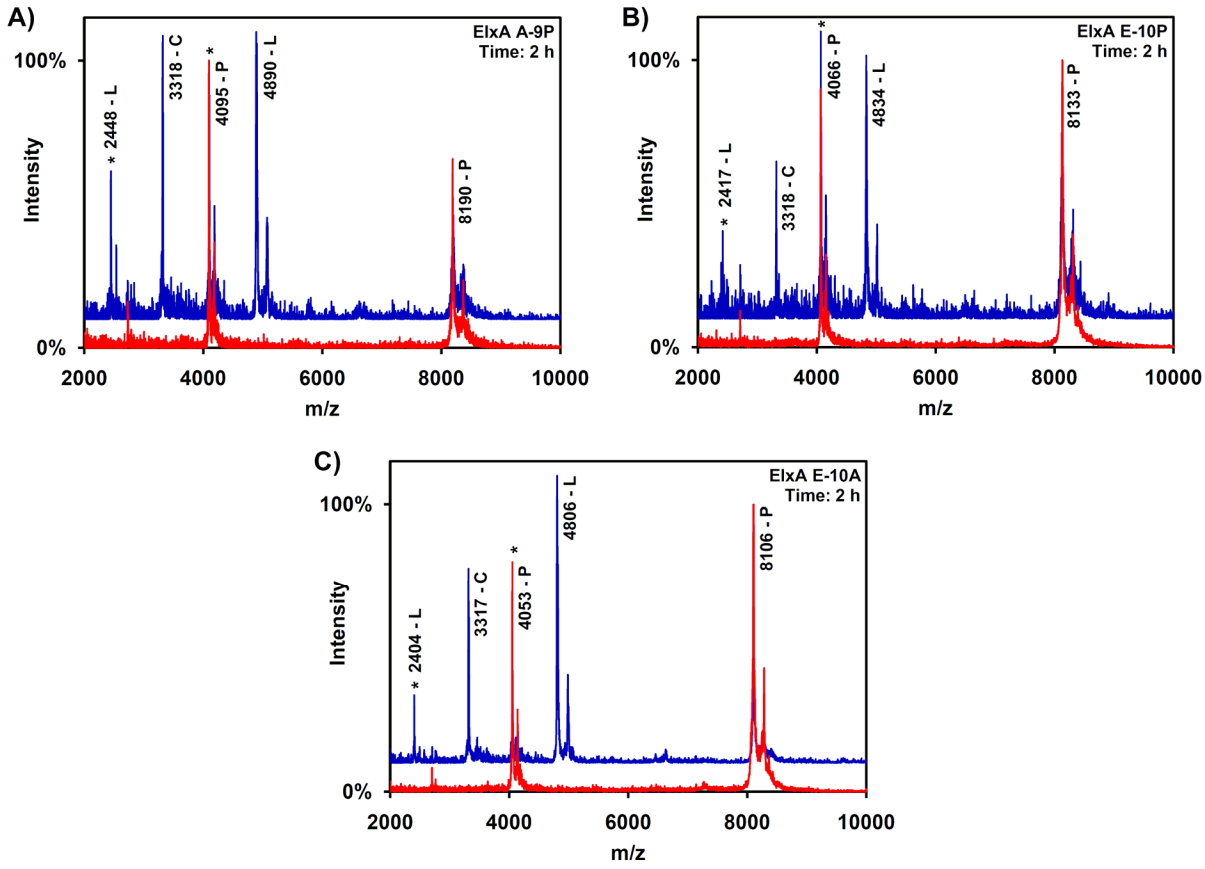


Figure 4.9. MALDI-TOF MS analysis of proteolytic processing of His_6-ElxA A−9P (A), His_6-ElxA E−10P (B), and His_6-ElxA E−10A (C) after incubation with (blue) or without (red) ElxP for 2 h. P, L, and C denotate precursor, leader, and core peptide, respectively. The asterisk (*) indicates a peak corresponding to a double-charged ion.

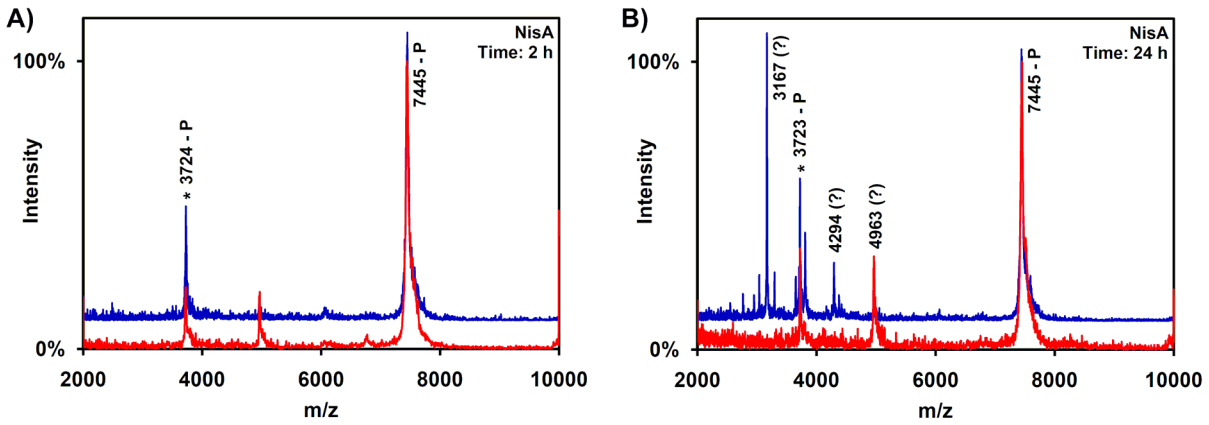


Figure 4.10. MALDI-TOF MS analysis of proteolytic processing of His_6-NisA after incubation with (blue) or without (red) ElxP after 2 h (A) or 24 h (B). P denotates precursor peptide. The single asterisk (*) indicates a peak corresponding to a double-charged ion. New peaks at *m/z* = 3167 and 4294 appeared after long incubation periods and may correspond to degradation products of ElxP.

4.4. DISCUSSION

Leader peptides seem to assist in several processes during the biosynthesis of lantibiotics and other ribosomally synthesized metabolites (*4*). In some cases, the leaders keep the precursor molecules inactive inside the host until secretion and proteolysis, as in nisin A biosynthesis (*25, 37*). However, in epilancin 15X, as well as in several other lantibiotics processed by intracellular LanP proteases (Figure 4.2), the leader sequence is removed inside the cytoplasm and the mature bioactive peptide is produced before secretion. Since the leader peptides usually contain several hydrophilic residues, they may also increase the precursor peptide solubility. Indeed, a hydrophobicity plot of ElxA evidences that the leader peptide is considerably more hydrophilic than the core region (Figure 4.11). Finally, the leaders may also act as a scaffold for the modification enzymes, allowing the recognition and binding of the entire precursor peptide. For instance, ABC transporters can use the leaders as a recognition signal. In fact, the transporter in nisin A biosynthesis, NisT, can secret various

engineered nonlantibiotic cargo attached to the leader (*27, 38*). However, this role of the leader peptide is not relevant for the ABC transporter ElxT, since the leader peptide is removed before secretion (*22*). Furthermore, the leaders can guide the interaction of the precursor peptide with lanthionine synthetases (LanM) (*39-42*), as well as with dehydratases (LanB) and cyclases (LanC), as demonstrated by the *in vivo* processing of nonlantibiotic therapeutic peptides attached to the nisin A leader peptide (*43, 44*).

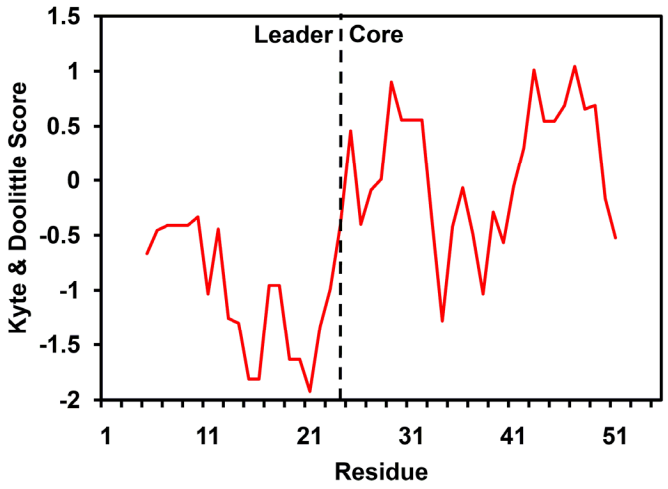


Figure 4.11. Hydrophobicity plot of ElxA using the Kyte & Doolittle scale (*45*). The values were calculated using a nine amino acid window for every position. Negative values correspond to hydrophilic regions.

For class I lantipeptides, four partially conserved sequences are present in the leader peptides: An N-terminal LF(D/N)L(D/N)(I//L/V) or FDLD motif, an Asn at position −8, a Ser at position −6, and a Pro at position −2 (Figure 4.4B). *In vivo* studies with NisB demonstrated that the FDLD motif is required for substrate recognition and efficient dehydration (*25, 26, 46*). Indeed, when this motif was replaced by AAAA, the interaction of NisA with NisB was abolished and no dehydration of the precursor peptide was observed (*26, 46*). Furthermore, a short NisA leader peptide, containing only the FDLD

motif and its N-terminal residues, allows the partial dehydration of the core region by NisB (*26*). A different study suggested that the conserved Ser−6 in NisA, in addition to the FDLD motif, is also important for modification of the core peptide (*25*). Similarly, NisC, the cyclase in nisin A biosynthesis, recognizes the leader peptide as suggested by *in vitro* experiments and contains a specific domain that likely interacts with the leader (*37*). A different *in vivo* study on the biosynthesis of Pep5 suggested that the FDLD motif seems to be important for optimal biosynthesis rates, but not strictly required for ring formation (*47*).

Regarding lantibiotic peptidases, some experimental evidence indicates that the C-terminal region of the class I leader peptides is important for substrate recognition (Table 4.4). However, this hypothesis has only been investigated through a limited set of mutations and a comprehensive analysis of the substrate scope is not available. Furthermore, the role of the FDLD motif in proteolysis is currently unknown. The sequence conservation of LanP proteins facilitated a phylogenetic analysis and the determination of their relatedness (Figure 4.2). The resulting phylogenetic tree highlights the existence of three distinct clades of LanP proteases that are not grouped based on cellular localization (intracellular, cell-wall anchored, or extracellular) or lantibiotic class (class I or class II). Interestingly, the residue N-terminal to the cleavage site in the peptidase substrates is conserved within each phylogroup and the clades were named based on that amino acid (clade Q, clade D/E, or clade R). Furthermore, sequence alignments of leader peptides processed by proteases from the same clade evidenced the conservation of some additional residues at the C-terminus of the leaders that may act as recognition signals for the proteases (Figure 4.4B). In this work, we

obtained some insights into the importance of leader peptide residues for processing of the epilancin 15X precursor peptide by the serine-type protease ElxP. This peptidase is a member of the Q-clade (Figure 4.2) and the conserved sequence DLXPQ at the C-terminus of the leader peptide (Figure 4.4A) could be essential for substrate recognition.

Substrate specificity of ElxP

Because the production of fully modified epilancin 15X precursor peptide has proven to be challenging as described in chapter 2 (*22*), we performed the substrate specificity studies on ElxP using linear peptides that were overexpressed and purified from *E. coli*. ElxP was able to cleave the unmodified precursor peptide His_6-ElxA and several mutants generating the expected proteolytic products (Table 4.4). Mutation of Gln−1 to Ala or Arg did not prevent cleavage but did reduce considerably the efficiency, as only small amounts of cleavage products were observed under the assay conditions, in agreement with previous *in vivo* results for NisP and EpiP that evidenced that the C-terminal residue of the leader peptide is important for proteolysis (Table 4.1) (*19, 24, 25*). In contrast, the replacement of the non-ionic Gln residue with the isosteric, negatively charged Glu residue allowed ElxP to retain good peptidase activity, suggesting that the selectivity for this position is based on steric effects more than electrostatic interactions. Furthermore, after mutation of Pro−2 to Lys or Asp, no significant proteolytic activity was observed and its substitution for Ala considerably affected the hydrolysis of the peptide, indicating that this partially conserved residue is important for proteolysis, in contradiction with previous *in vivo* results for NisP (Table 4.1). However, NisP is located in a different clade of the phylogenetic tree of LanPs

(Figure 4.2) and Pro−2 is not as well conserved in the precursor peptides processed by this group of proteases (Figure 4.4A).

Single mutations across the FDLD motif did not dramatically alter proteolytic processing by ElxP, suggesting that this region of the leader peptide is only important for recognition by the enzymes involved in dehydration and ring formation. Furthermore, only a small reduction of activity was observed after disruption of the predicted helical region by replacement of Ala−9 or Glu−10 with Pro, indicating that the secondary structure of the peptide is not essential. Interestingly, although ElxP cleaved to some extent the leader from His_6-ElxA Q−1R, His_6-NisA completely prohibits proteolysis (Figure 4.10). Thus, other residues in addition to the PQ cleavage site, such as Leu−4 or Asp−5 in the consensus sequence DLXPQ of precursor peptides processed by the Q-clade LanP proteases (and that are replaced by Ala−4 and Gly−5 in His_6-NisA), may be essential for substrate recognition.

The data discussed in this study was obtained at relatively high substrate concentrations (50 μM) and the reported activities might only consider the effect on k_{cat}. Thus, some of the peptides might still be poor substrates under physiological conditions of low concentration where the effect of the substrate binding affinity for the enzyme on the catalytic efficiency (k_{cat}/K_M) is more relevant. Future studies in our laboratory will focus on the determination of the kinetic parameters k_{cat} and K_M for some of the substrates. Interestingly, the formation of the (methyl)lanthionine rings and the dehydrated amino acids Dha at positions P1' and P3' and Dhb at position P7' and P8' are not required for proteolysis by ElxP, suggesting that the fully modified core region is not essential for substrate recognition. However, the N-terminus of dehydroepilancin

15X has a linear sequence of eleven residues (Figure 4.1) that might be important for substrate recognition and further investigation of the effect of this core region on proteolysis is recommended.

LanP proteases in the biosynthesis of class II lantipeptides and genome mining for novel lantipeptides

The LanP serine-type lantipeptide proteases can cleave the leader region from precursor peptides in the biosynthesis of not only class I lantibiotics, as commonly recognized, but also class II lantipeptides. The LanP-mediated cleavage of class II leader peptides occurs either in a single intracellular proteolytic event, as expected in the biosynthesis of lactocin S (*48*) (Figure 4.2), or in a more complex double cleavage process, as exemplified by the biosynthesis of cytolysin L and S (*49*) and probably lichenicidin (*50, 51*). In this case, the N-terminal sequence of the leader peptide containing the GG motif (Figure 4.12) is likely removed by the cysteine protease domain of a LanT (CylB and LicT for cytolysins and lichenicidin, respectively), concomitant with secretion, and followed by removal of the C-terminal amino acids of the leader region by a specific LanP peptidase (CylA and LicP for cytolysins and lichenicidin, respectively) (*49*). Interestingly, PneP, a protease likely involved in pneumococcin biosynthesis (*52*), has 49% amino acid identity with NisP (Figure 4.2) and might cleave off the C-terminus of the leader region of the precursors PneA1 and PneA2 (Figure 4.12), which have homology to the corresponding sequence in NisA (SDGAEPR in PneA1 compared to S–GASPR in NisA). In another example, a gene encoding for a putative LanP protease with 49% sequence similarity to LicP is present in the genome of *Bacillus halodurans* C-

125 (BH1491) (*50*). This protein may be responsible for the removal of the C-terminal amino acids of the leader peptide from fully modified HalA2 generating mature and biological active Halβ, a process that has remained elusive in the biosynthesis of the lantibiotic haloduracin (*53*). Thus, the double cleavage event in some class II lantipeptides further supports the hypothesis that LanP proteases may only recognize the C-terminal region of the leader peptide.

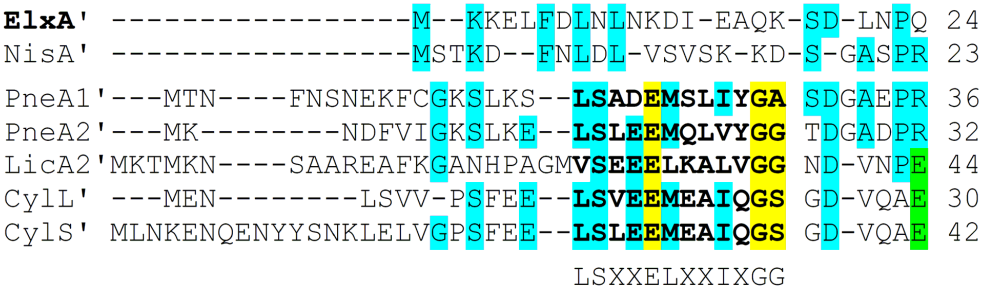


Figure 4.12. Sequence alignment of the leader regions of selected class I and class II lantibiotic precursor peptides. Only leaders encoded by gene clusters containing a *lanP* gene were included. The GG motif recognized by AMS proteins is shown in bold letters and the residues known to be important for leader peptide recognition by the protease domain of LanT proteins are highlighted in yellow (*54*). ElxA': leader peptide of epilancin 15x, NisA': leader peptide of nisin A, PneA1' and PneA2': leader peptide of pneumococcin, LicA2': leader peptide of lichenicidin LicA2, CylL' and CylS': leader peptides of cytolysin L and S.

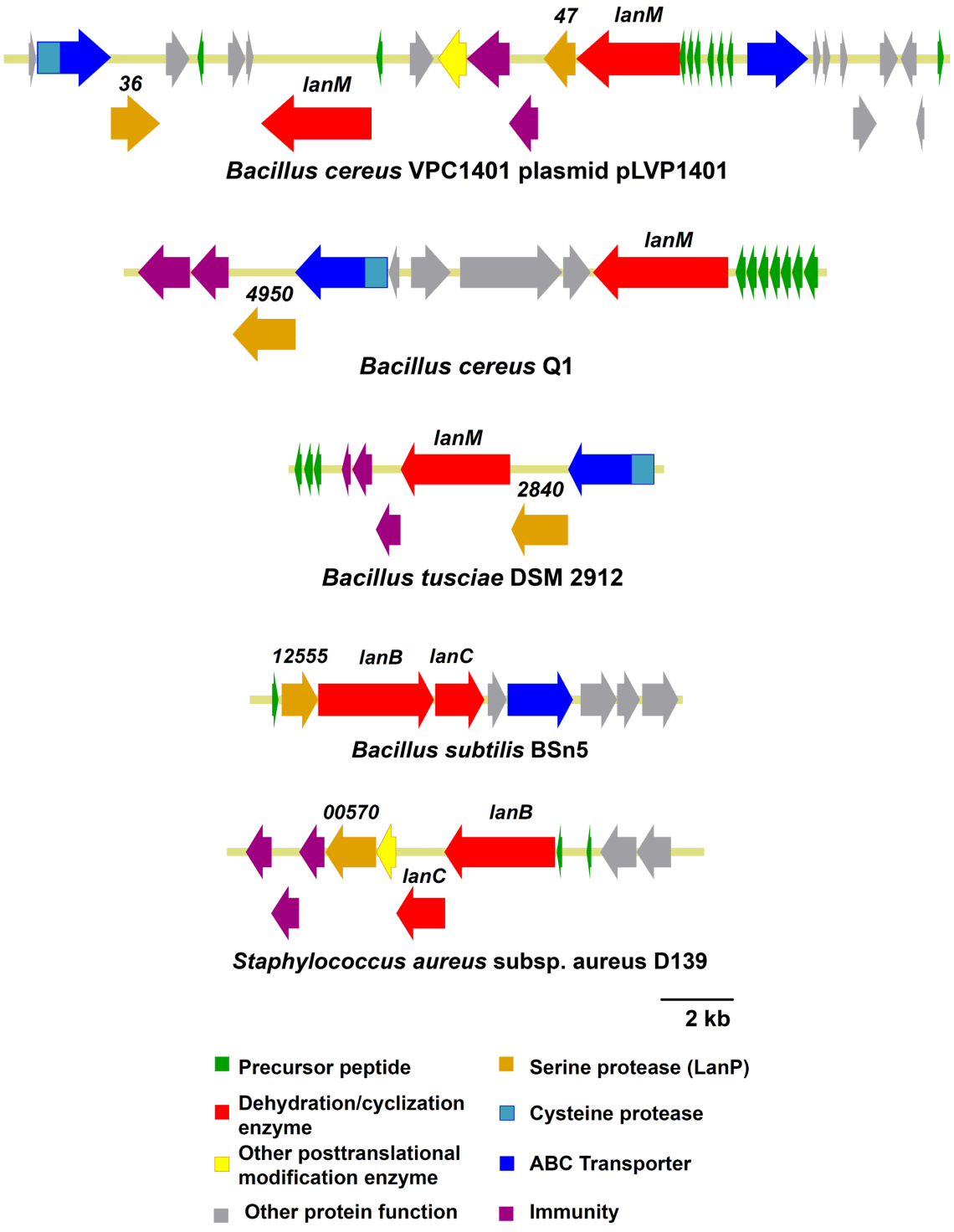


Figure 4.13. Novel biosynthetic gene clusters encoding for lantipeptides discovered by genome mining using the known sequences of LanP proteases. The numbers over the open reading frames encoding for putative proteases correspond to the locus tags in genome databases.

In recent years, genome mining for lanthionine synthetase homologs has guided the discovery of several lantipeptides (*55*), such as haloduracin (*53, 56*), lichenicidin (*50, 51*), Bsa (*57*), venezuelin (*2*), and prochlorosins (*58*). Additional diversity could be discovered by identifying gene clusters encoding for LanP homologs in genome databases. Indeed, ortholog neighborhood analysis at the integrated microbial genome website, with known sequences of LanP proteases, allowed the identification of novel gene clusters encoding for putative class I and class II lantipeptides (Figure 4.13). In particular, three class II biosynthetic clusters from *B. tusciae* DSM 2912, *B. cereus* Q1, and *B. cereus* VPC1401 encoding for LanT proteins with N-terminal cysteine protease domains, in addition to LanP proteins, were identified, suggesting that the double cleavage mechanism of the leader peptide is more widespread than currently recognized. However, the reason(s) behind the double proteolytic processing still remains unknown.

Interestingly, no genes encoding peptidases have been reported in the lantipeptide gene clusters from Actinobacteria, including the clusters of several class II lantibiotics such as cinnamycin (*59*), actagardine (*60*), deoxyactagardine (*61*), and michiganin A (*62*), and the class III and IV lantipeptides (*2, 3, 63*). Indeed, mining of a library of 127 genome sequences of Actinobacteria, composed of published *Streptomyces* spp. sequences and the private Actinobacteria genome database from the Institute for Genomic Biology (University of Illinois at Urbana-Champaign) did not result in the identification of any lantipeptide gene cluster encoding for a subtilisin-like LanP protease (Dr. James Doroghazi, personal communication). Thus, the cleavage of the leader in this phylum is likely performed by a different mechanism, not related to

LanT or LanP-like proteins. In the case of cinnamycin, proteolysis may be performed by the general secretory *sec*-dependent export machinery (*59*). However, the conserved AXA motif preceding the cleavage site of the type I signal peptidases (*64*) is absent from most of the other leader sequences from Actinobacteria.

Ortholog neighborhood analysis of class III and IV lanthionine synthetases (RamC-like and LanL) unrevealed some lantipeptide gene clusters encoding for different classes of peptidases that may cleave off the leader regions from the encoded LanA peptides (Figure 4.14). The proteins encoded by the *Kribbella flavida* and *Streptococcus pneumoniae* gene clusters, which will be designated LanP2, are putative intracellular serine peptidases from the prolyl oligopeptidase family. Similar proteins are involved in leader peptide removal from amatoxins and phallotoxins, bicyclic octa- and heptapeptides, respectively, containing a Trp-Cys crossbridge (tryptathionine) that are produced by poisonous mushrooms of the genera *Amanita*, *Galerina*, *Lepiota*, and *Conocybe* (*65, 66*). The structural fold of members of this peptidase family resembles that of carboxypeptidase C, instead of subtilisin as predicted for LanP proteins, suggesting that both groups of peptidases have evolved by independent convergent evolution (*67*). Furthermore, the order of the catalytic triad residues (Ser, His, and Asp) in the amino acid sequence in both groups of serine peptidases is predicted to be different (DHS for LanPs versus SDH for LanP2s) (*67*). Although members of the prolyl oligopeptidase family have different substrate specificity, amino acid sequence homology of the identified proteases to previously characterized proteins allows the prediction of tentative cleavage sites in the precursor peptides encoded by the corresponding gene clusters (Figure 4.14, underlined red letters in LanAs).

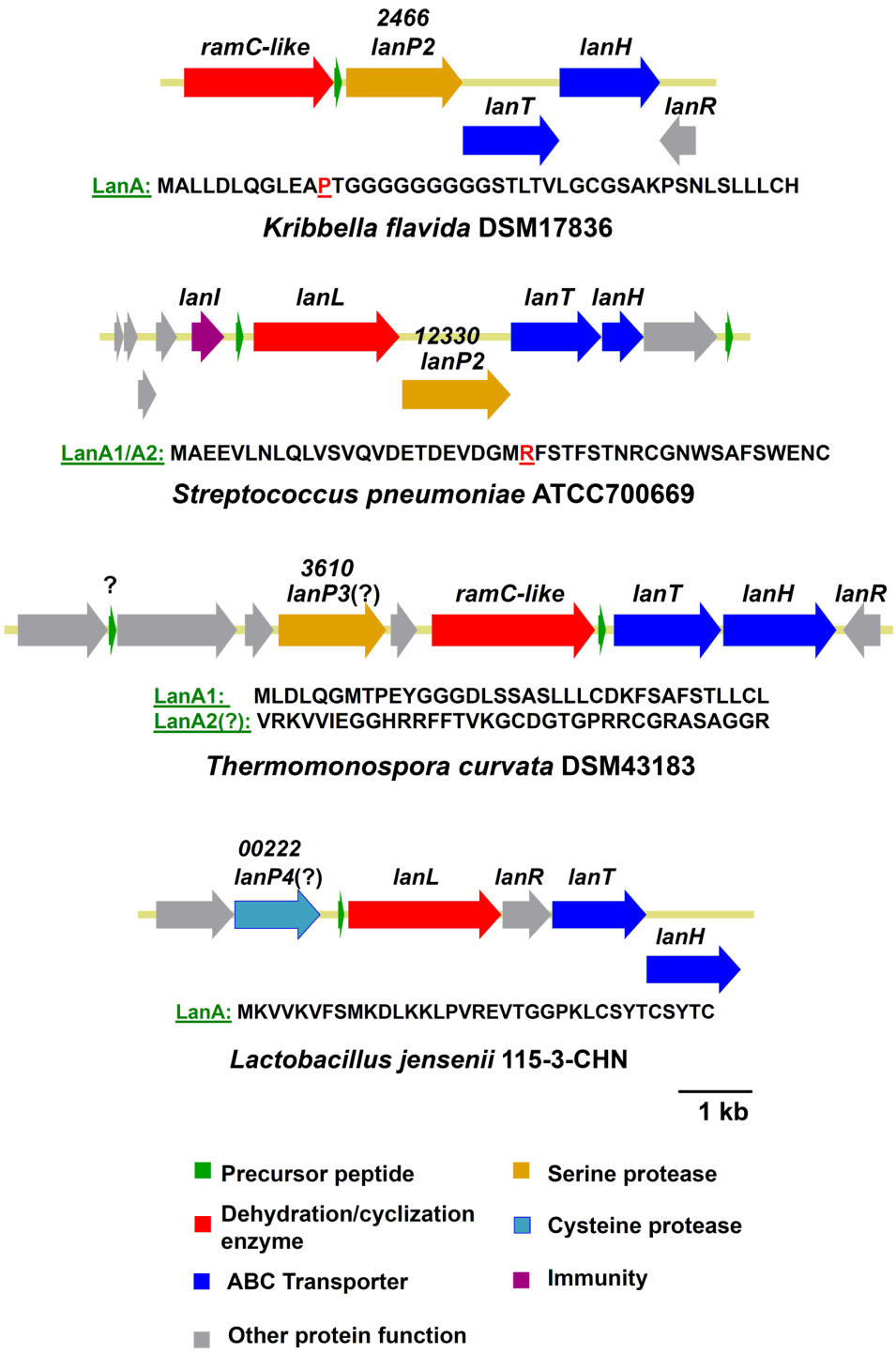


Figure 4.14. Class III and IV lantipeptides gene clusters potentially encoding for novel classes of leader peptide peptidases. For LanP2 proteins (prolyl endopeptidases), the most likely residues at the C-terminus of the cleavage site in the precursor peptides are shown in underlined red letters. The numbers over the open reading frames encoding for the putative peptidases correspond to the locus tags in genome databases.

Open reading frames (ORFs) encoding for two additional classes of putative peptidases in the neighborhood of *lanL*/*ramC*-like genes were identified in the genomes of *Thermomonospora curvata* and *Lactobacillus jensenii*. However, their involvement in lantipeptide biosynthesis is not as clear, since the ORFs are located at the predicted boundaries of the clusters. The putative peptidase LanP3 is a transmembrane serine protease with a Ser-Lys catalytic diad that may process peptides concomitant with secretion, while the protease LanP4 is a predicted intracellular cysteine peptidase. Whether the LanP2, LanP3, and LanP4 peptidases are involved in leader peptide cleavage during lantipeptide biosynthesis or not is currently uncertain. However, heterologous expression of the proteins and *in vitro* reconstitution of their enzymatic activity, as described above, may provide the answer to this question.

Homology modeling of ElxP

To further understand the substrate specificity of ElxP, homology modeling of the protease was performed using as template the recently released X-ray crystal structure of BsaP (pdb ID: 3T41) (*68*), the peptidase involved in the biosynthesis of a group of lantibiotics designated Bsa (*57*). The model was constructed using Phyre2 (Figure 4.15A) (*69*). To validate the results, a second model was generated by using the "multiple template modeling" tool from the same website and six protein templates automatically selected based on heuristics (pdb ID: 2OXA, 1P8J, 1R6V, 3I74, 3QFH, and 3T41). This second model was predicted to have more than 90% confidence for 98% of the amino acid residues. The predicted structures are remarkably similar (Figure 4.15B), especially near the active site and the substrate-binding region (S4-S1', based

on a notation previously defined for proteases (*70*)). Therefore, only the model based on the BsaP structure will be considered for further discussion (Figure 4.15A).

ElxP is predicted to have an α/β fold with the active site located between a three-stranded central parallel β-sheet and two α-helices. The model also supports the role of the predicted catalytic triad and oxyanion hole residues previously identified by sequence alignment (Asp27, His62, Ser240, and Asn154) (*22*) (Figure 4.15C). Based on the crystal structures of subtilisin BPN' and thermitase in complex with the inhibitor eglin (*71, 72*) and previous models for NisP (*73*) and CylA (*49*), the substrate-binding region in ElxP is predicted to be a surface channel or crevice that can accommodate at least seven amino acid residues (P5-P2') (Figure 4.15D). In subtilisin and thermitase, the specificity appears to be mainly determined by interactions of the substrate P4 and P1 residue side chains with the large S4 and S1 subsites (*73, 74*). In ElxP, the S1 site is predicted to be a large, elongated pocket with residues Ser162 and Ser153 on one side and Asp155 at the rim, likely providing H-bonding interactions with the ElxA residue Gln−1(P1) (Table 4.5 and Figure 4.16). Importantly, the corresponding residues in a mutant subtilisin BPN' (Glu156 and Asp166), based on sequence alignment, were shown to be determinant for the substrate specificity of the enzyme for P1 residues (*75*). Furthermore, when Gln−1 in ElxA was substituted for Ala or Arg, the efficiency of the reaction was considerably reduced (Table 4.4 and Figure 4.5), probably due to loss of the essential interactions with the predicted ElxP binding residues. In contrast, when Gln−1 was substituted for its isosteric P1 homolog Glu, the mutant was processed qualitatively with very similar efficiency as the wild-type, suggesting that Glu−1 may still interact with Ser162 and Ser153 by H-bonding interactions. As mentioned above, this

result may be caused by the high peptide concentrations used in the assay and the mutant Q−1E might still be a poor substrate under physiological conditions. The corresponding residues in NisP and CylA were previously predicted to be Asp436 and Asp408 interacting with Arg−1 in NisA (*73*) and His275 interacting with Glu−1 in $CylL_S$' and $CylL_L$' (Figure 4.12) (*49*), respectively (Table 4.5).

Sequence alignment allowed the identification of the residues located at the predicted S1 subsite in other LanP proteases (Table 4.5). In general, for the proteases located in the R-clade of the phylogenetic tree (Figure 4.2), acidic residues (Asp or Glu) in the S1 subsite are expected to interact with Arg−1 in the precursor peptides. Additionally, for the D/E-clade proteases, basic residues (Lys or His) in the S1 subsite usually interact with the Asp−1 or Glu−1 residues of the substrates. Furthermore, the residue Ser162 (and Asp155) in ElxP, or their equivalents for other LanP, likely provide the dominant interactions with the P1 residue of the substrate and may determine the substrate specificity as suggested for subtilisin BPN' (*75*). Importantly, determination of the location of novel proteases in the phylogenetic tree (Figure 4.2) and the S1 subsite residues can provide some insights into the cleavage site of precursor peptides identified by *in silico* methods and the final structure of lantipeptides. For instance, the cleavage site of the precursor peptides encoded by the gene clusters identified by ortholog neighborhood analysis in the present study (Figure 4.13) can be easily predicted using this information (Figures 4.2 and 4.16). Furthermore, for the *B. cereus* VPC1401 gene cluster that encodes two putative LanP proteases, the predicted cellular localization of the proteases and the motifs conserved in the leader peptides allowed determination of the most likely substrates for every protease (Figure 4.17).

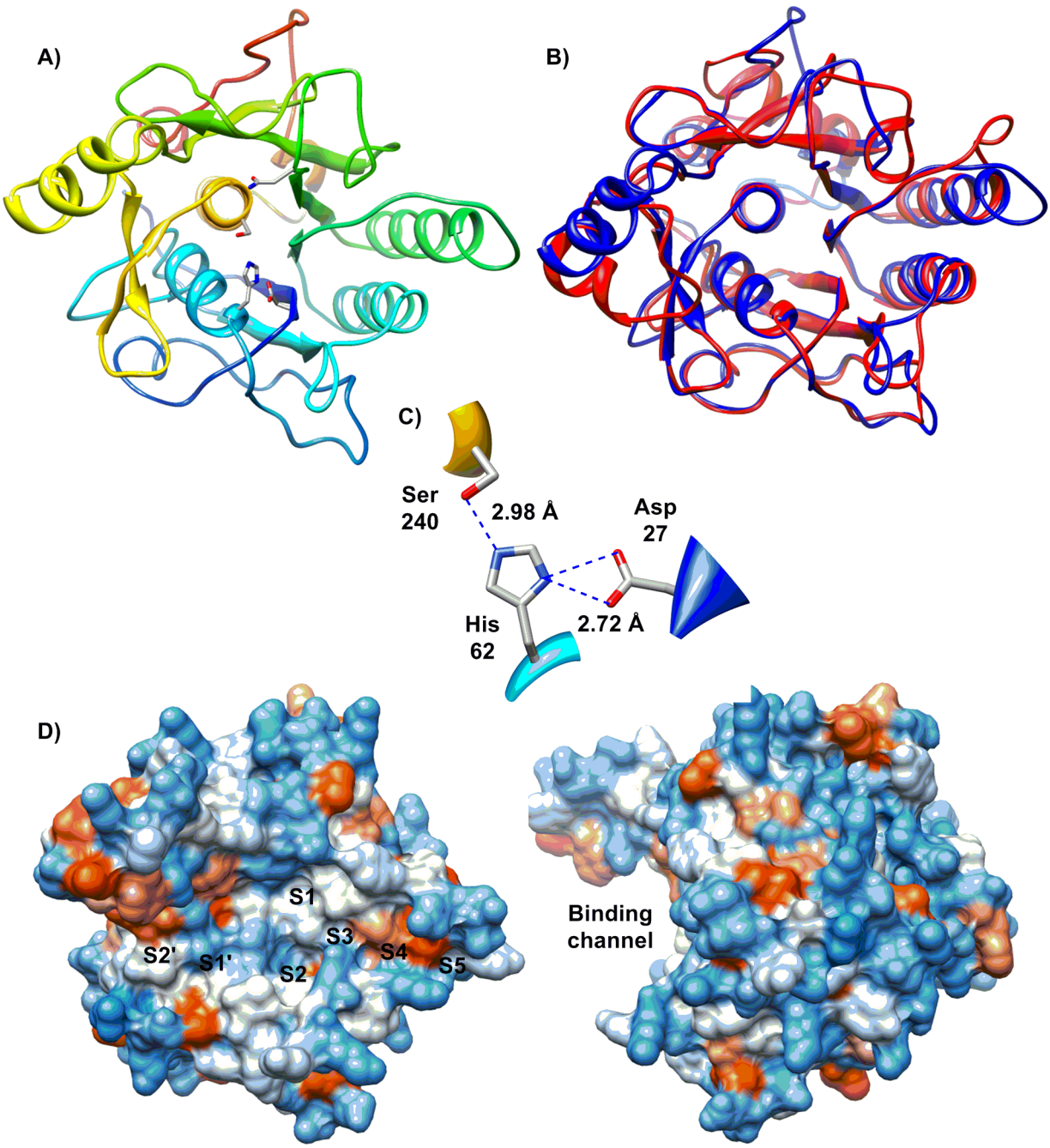


Figure 4.15. Model of the epilancin 15X serine-type protease (ElxP). The model was constructed based on the structure of BsaP (A and B, blue) and compared with the model obtained by multiple template modeling (B, red) using Phyre2 (*69*). The predicted residues of the catalytic triad (C) and the potential substrate binding cleft accommodating residues P5-P2' were identified (D, left: front view; right: side view after 90° rotation). The hydrophobicity surface structure is shown in colors from blue (hydrophilic) to red (hydrophobic) based on the Kyte and Doolittle scale (*45*).

Table 4.5. Predicted S1 site residues for different lantibiotic serine proteases. The enzyme/substrate interaction is probably dominated by S162/D155 in ElxP (shown in red) or the equivalent residues in other proteases.

Clade	Protease	P1 residue	S1 residues					
	ElxP	**Q**	**S162**	**D155**	S122	G123	Y124	S153
Q	EciP	**Q**	**L**	**N**	S	G	A	M
	PepP	**Q**	**M**	**N**	S	G	E	M
D/E	CylA	**E**	**H**	**E**	G	S	Y	G
	LicP	**E**	**H**	**K**	G	S	Y	G
R	BsaP	**R**	**D**	**D**	G	S	Y	G
	EpiP	**R**	**D**	**D**	G	N	Y	G
	GdmP	**R**	**D**	**D**	G	N	Y	G
	MutP.I	**R**	**D**	**D**	G	Q	Y	G
	MutP.III	**R**	**D**	**D**	G	Q	Y	G
	NiqP	**R**	**D**	**D**	G	Q	Y	G
	NisP	**R**	**D**	**D**	G	Q	Y	G
	NsuP	**R**	**D**	**E**	G	E	Y	G
	PneP	**R**	**D**	**D**	G	Q	Y	G

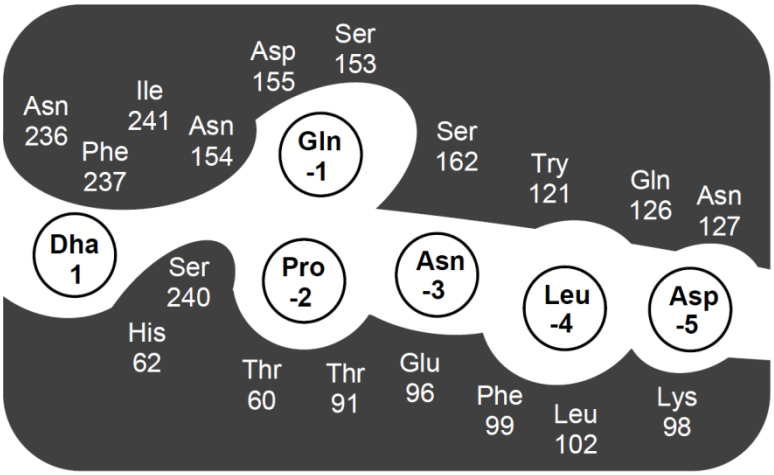


Figure 4.16. Schematic 2-D representation of the binding cleft of ElxP according to the model in Figure 4.15.

The S2 site of ElxP is predicted to be a shallow cleft bounded by Ser28, Thr60, Thr91, and the active site residue His62 and can probably accommodate only small residues such as Pro−2(P2) in ElxA (Figure 4.16). Indeed, when this amino acid was

mutated for the larger residues Lys or Asp, the cleavage reaction was greatly prohibited probably due to steric hindrance (Table 4.4 and Figure 4.6). As in previous models for NisP and CylA (*49, 73*), S3 is not a distinct site and the residue Asn−3(P3) in ElxA might point away from ElxP towards the solvent, although Glu96 may provide some interaction (Figure 4.16). The S4 site is predicted to be a small pocket bounded by large hydrophobic side chains (Tyr97, Phe99, Leu102, Tyr121, and Phe134) and is expected to accommodate hydrophobic residues, such as Leu−4(P4) in ElxA (Figure 4.16). Finally, a potential S5 site with residues Lys98, Gln126, Asn127, and Asn128 pointing towards the solvent may provide crucial interactions with Asp−5(P5) and may explain the lack of cleavage activity of ElxP with NisA that contains a Gly residue at this position (Figure 4.16). Regarding the binding of the segment C-terminal to the cleavage site, the S1' pocket is predicted to be a less-defined small hydrophobic site bounded by His62, Asn236, Phe237, and Ile241 and likely provides weak interactions. Interestingly, a hydrophobic patch was identified in a region near the predicted S2' site (Figure 4.13D). This patch might interact with the hydrophobic core peptide (Figure 4.11) further providing additional binding energy. Based on the discussed experimental results and the simplified structural model, lantipeptide leader peptidases with a desired specificity could be designed by protein engineering as reported before for subtilisin BPN' (*75*). Furthermore, the high fidelity of ElxP suggests that this peptidase and other LanP proteins may have interesting potential for application in biotechnology to remove expression tags similar to the use of Tobacco Etch Virus (TEV) protease, enterokinase, thrombin, and Factor Xa.

In summary, the PQ sequence, but not the FDLD motif, in the leader peptide of ElxA is critical for substrate recognition by ElxP, as most of the mutants generated at these positions were processed at a considerably lesser degree than the wild-type peptide. Furthermore, homology modeling of ElxP, phylogenetic analysis of LanP proteases, and the lack of NisA cleavage by the peptidase suggest that additional residues at the C-terminus of the leader peptide, such as the negatively charged Asp−5 or the small hydrophilic Leu−4, are involved in substrate recognition. These findings also indicate that in class I lantipeptide biosynthesis, different amino acid residues of the leader peptide have dissimilar effects on the recognition by the modification enzymes, in agreement with previous results on the specificity of LanB dehydratases (*25, 26, 46*). Further investigation, including determination of kinetic parameters and evaluation of additional mutants, will contribute to the elucidation of the rules for substrate processing by LanP proteases.

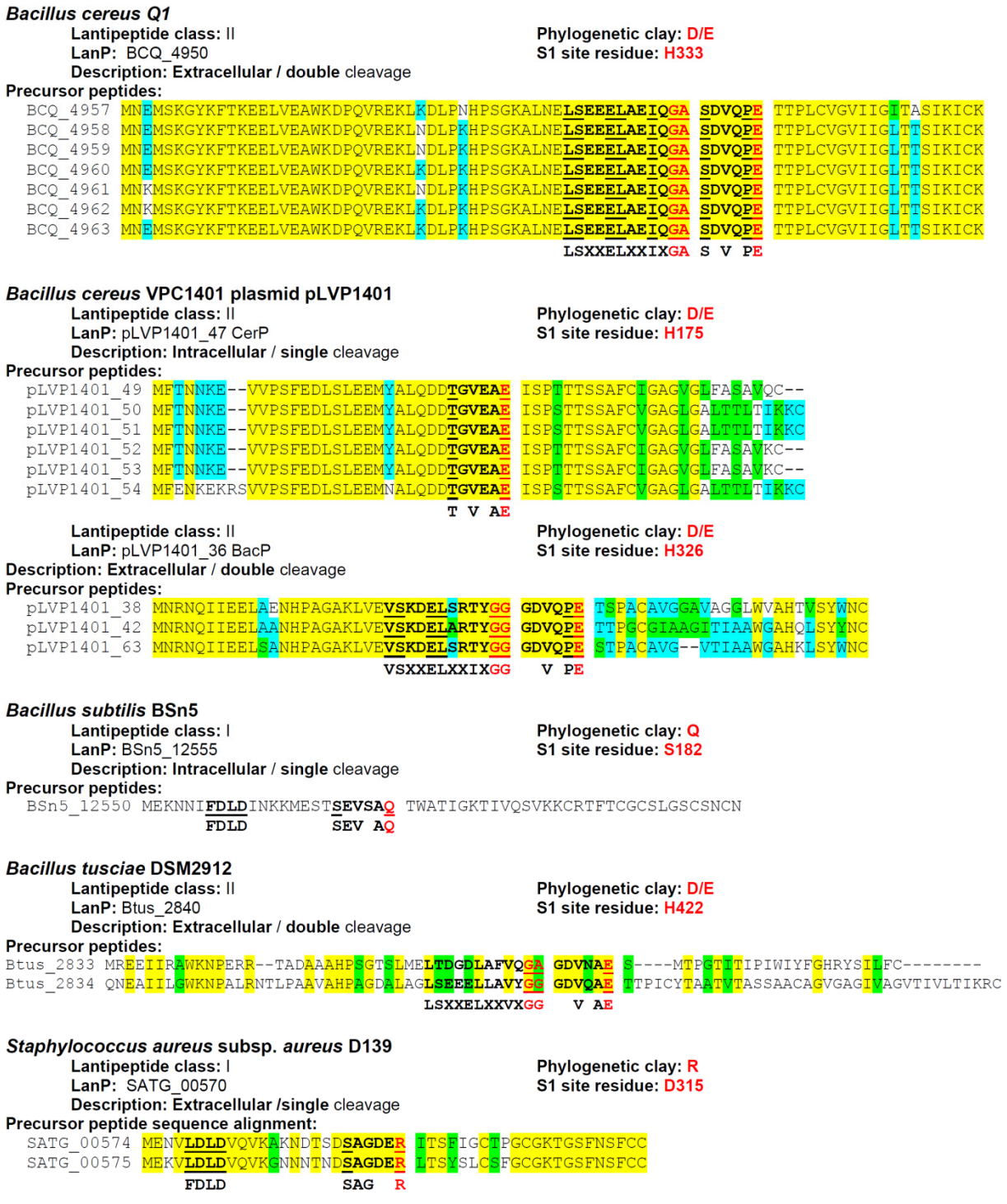


Figure 4.17. Determination of proteolytic cleavage sites of novel precursor peptides discovered by genome mining. The location of the proteases in the phylogenetic tree (Figure 4.2) and the S1 site residues corresponding to S162 in ElxP were used to determine the P1 residues (in red bold letters) in the predicted precursor peptides. The conserved sequences from class I and II lantipeptides are in bold underlined letters.

4.5. REFERENCES

1. Willey, J. M., and van der Donk, W. A. (2007) Lantibiotics: Peptides of diverse structure and function, *Annu. Rev. Microbiol. 61*, 477-501.

2. Goto, Y., Li, B., Claesen, J., Shi, Y., Bibb, M. J., and van der Donk, W. A. (2010) Discovery of unique lanthionine synthetases reveals new mechanistic and evolutionary insights, *Plos Biol. 8*, e1000339.

3. Meindl, K., Schmiederer, T., Schneider, K., Reicke, A., Butz, D., Keller, S., Guhring, H., Vertesy, L., Wink, J., Hoffmann, H., Bronstrup, M., Sheldrick, G. M., and Sussmuth, R. D. (2010) Labyrinthopeptins: A new class of carbacyclic lantibiotics, *Angew. Chem. Int. Ed. 49*, 1151-1154.

4. Oman, T. J., and van der Donk, W. A. (2010) Follow the leader: the use of leader peptides to guide natural product biosynthesis, *Nat. Chem. Biol. 6*, 9-18.

5. Dirix, G., Monsieurs, P., Dombrecht, B., Daniels, R., Marchal, K., Vanderleyden, J., and Michiels, J. (2004) Peptide signal molecules and bacteriocins in Gram-negative bacteria: a genome-wide *in silico* screening for peptides containing a double-glycine leader sequence and their cognate transporters, *Peptides 25*, 1425-1440.

6. Furgerson Ihnken, L. A., Chatterjee, C., and van der Donk, W. A. (2008) *In vitro* reconstitution and substrate specificity of a lantibiotic protease, *Biochemistry 47*, 7352-7363.

7. Nishie, M., Shioya, K., Nagao, J., Jikuya, H., and Sonomoto, K. (2009) ATP-dependent leader peptide cleavage by NukT, a bifunctional ABC transporter, during lantibiotic biosynthesis, *J. Biosci. Bioeng. 108*, 460-464.

8. Ishii, S., Yano, T., and Hayashi, H. (2006) Expression and characterization of the peptidase domain of *Streptococcus pneumoniae* ComA, a bifunctional ATP-binding cassette transporter involved in quorum sensing pathway, *J. Biol. Chem. 281*, 4726-4731.

9. Kotake, Y., Ishii, S., Yano, T., Katsuoka, Y., and Hayashi, H. (2008) Substrate recognition mechanism of the peptidase domain of the quorum-sensing-signal-producing ABC transporter ComA from Streptococcus, *Biochemistry 47*, 2531-2538.

10. Ishii, S., Yano, T., Ebihara, A., Okamoto, A., Manzoku, M., and Hayashi, H. (2010) Crystal structure of the peptidase domain of *Streptococcus* ComA, a bifunctional ATP-binding cassette transporter involved in the quorum-sensing pathway, *J. Biol. Chem. 285*, 10777-10785.

11. Wu, K. H., and Tai, P. C. (2004) Cys32 and His105 are the critical residues for the calcium-dependent cysteine proteolytic activity of CvaB, an ATP-binding cassette transporter, *J. Biol. Chem. 279*, 901-909.

12. Håvarstein, L. S., Diep, D. B., and Nes, I. F. (1995) A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export, *Mol. Microbiol. 16*, 229-240.

13. Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Mol. Biol. Evol. 4*, 406-425.

14. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences, *Comput. Appl. Biosci. 8*, 275-282.

15. Felsenstein, J. (1985) Confidence-limits on phylogenies - an approach using the bootstrap, *Evolution 39*, 783-791.

16. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol. Biol. Evol. 28*, 2731-2739.

17. Corvey, C., Stein, T., Dusterhus, S., Karas, M., and Entian, K. D. (2003) Activation of subtilin precursors by *Bacillus subtilis* extracellular serine proteases subtilisin (AprE), WprA, and Vpr, *Biochem. Biophys. Res. Commun. 304*, 48-54.

18. Schnell, N., Engelke, G., Augustin, J., Rosenstein, R., Ungermann, V., Götz, F., and Entian, K.-D. (1992) Analysis of genes involved in the biosynthesis of the lantibiotic epidermin, *Eur. J. Biochem. 204*, 57-68.

19. Geißler, S., Götz, F., and Kupke, T. (1996) Serine protease EpiP from *Staphylococcus epidermidis* catalyzes the processing of the epidermin precursor peptide, *J. Bacteriol. 178*, 284-288.

20. van der Meer, J. R., Polman, J., Beerthuyzen, M. M., Siezen, R. J., Kuipers, O. P., and de Vos, W. M. (1993) Characterization of the *Lactococcus lactis* nisin A operon genes *nisP*, encoding a subtilisin-like serine protease involved in precursor processing, and *nisR*, encoding a regulatory protein involved in nisin biosynthesis, *J. Bacteriol. 175*, 2578-2588.

21. Meyer, C., Bierbaum, G., Heidrich, C., Reis, M., Süling, J., Iglesias-Wind, M. I., Kempter, C., Molitor, E., and Sahl, H.-G. (1995) Nucleotide sequence of the lantibiotic Pep5 biosynthetic gene cluster and functional analysis of PepP and PepC, *Eur. J. Biochem. 232*, 478-489.

22. Velásquez, J. E., Zhang, X., and van der Donk, W. A. (2011) Biosynthesis of the antimicrobial peptide epilancin 15X and its N-terminal lactate, *Chem. Biol. 18*, 857-867.

23. Rink, R., Wierenga, J., Kuipers, A., Kluskens, L. D., Driessen, A. J., Kuipers, O. P., and Moll, G. N. (2007) Dissection and modulation of the four distinct activities of nisin by mutagenesis of rings A and B and by C-terminal truncation, *Appl. Environ. Microbiol. 73*, 5809-5816.

24. Kuipers, O. P., Rollema, H. S., de Vos, W. M., and Siezen, R. J. (1993) Biosynthesis and secretion of a precursor of nisin Z by *Lactococcus lactis*, directed by the leader peptide of the homologous lantibiotic subtilin from *Bacillus subtilis*, *FEBS Lett. 330*, 23-27.

25. van der Meer, J. R., Rollema, H. S., Siezen, R. J., Beerthuyzen, M. M., Kuipers, O. P., and de Vos, W. M. (1994) Influence of amino acid substitutions in the nisin leader peptide on biosynthesis and secretion of nisin by *Lactococcus lactis*, *J. Biol. Chem. 269*, 3555-3562.

26. Plat, A., Kluskens, L. D., Kuipers, A., Rink, R., and Moll, G. N. (2011) Requirements of the engineered leader peptide of nisin for inducing modification, export, and cleavage, *Appl. Environ. Microbiol. 77*, 604-611.

27. Kuipers, A., De Boef, E., Rink, R., Fekken, S., Kluskens, L. D., Driessen, A. J., Leenhouts, K., Kuipers, O. P., and Moll, G. N. (2004) NisT, the transporter of the lantibiotic nisin, can transport fully modified, dehydrated and unmodified prenisin and fusions of the leader peptide with non-lantibiotic peptides, *J. Biol. Chem. 279*, 22176-22182.

28. Beck-Sickinger, A. G., and Jung, G. (1991) Synthesis and conformational analysis of lantibiotic leader-, pro-, and pre-peptides, In *Nisin and Novel Lantibiotics* (Jung, G., and Sahl, H.-G., Eds.), pp 218-230, ESCOM, Leiden.

29. Weil, H. P., Beck-Sickinger, A. G., Metzger, J., Stevanovic, S., Jung, G., Josten, M., and Sahl, H. G. (1990) Biosynthesis of the lantibiotic Pep5. Isolation and characterization of a prepeptide containing dehydroamino acids, *Eur. J. Biochem. 194*, 217-223.

30. Sen, T. Z., Jernigan, R. L., Garnier, J., and Kloczkowski, A. (2005) GOR V server for protein secondary structure prediction, *Bioinformatics 21*, 2787-2788.

31. Rost, B., Yachdav, G., and Liu, J. (2004) The PredictProtein server, *Nucleic. Acids Res. 32*, W321-326.

32. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0, *Bioinformatics 23*, 2947-2948.

33. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator, *Genome Res. 14*, 1188-1190.

34. Grant, S. G., Jessee, J., Bloom, F. R., and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants, *Proc. Natl. Acad. Sci. U. S. A. 87*, 4645-4649.

35. Li, B., Cooper, L. E., and van der Donk, W. A. (2009) In vitro studies of lantibiotic biosynthesis, *Methods Enzymol. 458*, 533-558.

36. Geoghegan, K. F., Dixon, H. B., Rosner, P. J., Hoth, L. R., Lanzetti, A. J., Borzilleri, K. A., Marr, E. S., Pezzullo, L. H., Martin, L. B., LeMotte, P. K., McColl, A. S., Kamath, A. V., and Stroh, J. G. (1999) Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in *Escherichia coli*: the cause of extra mass of 258 or 178 Da in fusion proteins, *Anal. Biochem. 267*, 169-184.

37. Li, B., Yu, J.-P. J., Brunzelle, J. S., Moll, G. N., van der Donk, W. A., and Nair, S. K. (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis, *Science 311*, 1464-1467.

38. Izaguirre, G., and Hansen, J. N. (1997) Use of alkaline phosphatase as a reporter polypeptide to study the role of the subtilin leader segment and the SpaT transporter in the posttranslational modifications and secretion of subtilin in *Bacillus subtilis* 168, *Appl. Environ. Microbiol. 63*, 3965-3971.

39. Levengood, M. R., Patton, G. C., and van der Donk, W. A. (2007) The leader peptide is not required for post-translational modification by lacticin 481 synthetase, *J. Am. Chem. Soc. 129*, 10314-10315.

40. Patton, G. C., Paul, M., Cooper, L. E., Chatterjee, C., and van der Donk, W. A. (2008) The importance of the leader sequence for directing lanthionine formation in lacticin 481, *Biochemistry 47*, 7342-7351.

41. Nagao, J., Morinaga, Y., Islam, M. R., Asaduzzaman, S. M., Aso, Y., Nakayama, J., and Sonomoto, K. (2009) Mapping and identification of the region and secondary structure required for the maturation of the nukacin ISK-1 prepeptide, *Peptides 30*, 1412-1420.

42. Chen, P., Qi, F. X., Novak, J., Krull, R. E., and Caufield, P. W. (2001) Effect of amino acid substitutions in conserved residues in the leader peptide on biosynthesis of the lantibiotic mutacin II, *FEMS Microbiol. Lett. 195*, 139-144.

43. Rink, R., Wierenga, J., Kuipers, A., Kluskens, L. D., Driessen, A. J. M., Kuipers, O. P., and Moll, G. N. (2007) Production of dehydroamino acid-containing peptides by *Lactococcus lactis*, *Appl. Environ. Microbiol. 73*, 1792-1796.

44. Kluskens, L. D., Kuipers, A., Rink, R., de Boef, E., Fekken, S., Driessen, A. J., Kuipers, O. P., and Moll, G. N. (2005) Post-translational modification of therapeutic peptides by NisB, the dehydratase of the lantibiotic nisin, *Biochemistry 44*, 12827-12834.

45. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol. 157*, 105-132.

46. Mavaro, A., Abts, A., Bakkes, P. J., Moll, G. N., Driessen, A. J., Smits, S. H., and Schmitt, L. (2011) Substrate recognition and specificity of the NisB protein, the lantibiotic dehydratase involved in nisin biosynthesis, *J. Biol. Chem. 286*, 30552-30560.

47. Neis, S., Bierbaum, G., Josten, M., Pag, U., Kempter, C., Jung, G., and Sahl, H. G. (1997) Effect of leader peptide mutations on biosynthesis of the lantibiotic Pep5, *FEMS Microbiol. Lett. 149*, 249-255.

48. Skaugen, M., Abildgaard, C. I., and Nes, I. F. (1997) Organization and expression of a gene cluster involved in the biosynthesis of the lantibiotic lactocin S, *Mol. Gen. Genet. 253*, 674-686.

49. Booth, M. C., Bogie, C. P., Sahl, H.-G., Siezen, R. J., Hatter, K. L., and Gilmore, M. S. (1996) Structural analysis and proteolytic activation of *Enterococcus faecalis* cytolysin, a novel lantibiotic, *Mol. Microbiol. 21*, 1175-1184.

50. Dischinger, J., Josten, M., Szekat, C., Sahl, H. G., and Bierbaum, G. (2009) Production of the novel two-peptide lantibiotic lichenicidin by *Bacillus licheniformis* DSM 13, *Plos One 4*, e6788.

51. Begley, M., Cotter, P. D., Hill, C., and Ross, R. P. (2009) Identification of a novel two-peptide lantibiotic, lichenicidin, following rational genome mining for LanM proteins, *Appl. Environ. Microbiol. 75*, 5451-5460.

52. Majchrzykiewicz, J. A., Lubelski, J., Moll, G. N., Kuipers, A., Bijlsma, J. J., Kuipers, O. P., and Rink, R. (2010) Production of a class II two-component lantibiotic of *Streptococcus pneumoniae* using the class I nisin synthetic machinery and leader sequence, *Antimicrob. Agents. Chemother. 54*, 1498-1505.

53. McClerren, A. L., Cooper, L. E., Quan, C., Thomas, P. M., Kelleher, N. L., and van der Donk, W. A. (2006) Discovery and in vitro biosynthesis of haloduracin, a two-component lantibiotic, *Proc. Natl. Acad. Sci. U. S. A. 103*, 17243-17248.

54. Furgerson Ihnken, L. A., Chatterjee, C., and van der Donk, W. A. (2008) In vitro reconstitution and substrate specificity of a lantibiotic protease, *Biochemistry 47*, 7352-7363.

55. Velásquez, J. E., and van der Donk, W. A. (2011) Genome mining for ribosomally synthesized natural products, *Curr. Opin. Chem. Biol. 15*, 11-21.

56. Lawton, E. M., Cotter, P. D., Hill, C., and Ross, R. P. (2007) Identification of a novel two-peptide lantibiotic, haloduracin, produced by the alkaliphile *Bacillus halodurans* C-125, *FEMS Microbiol. Lett. 267*, 64-71.

57. Daly, K. M., Upton, M., Sandiford, S. K., Draper, L. A., Wescombe, P. A., Jack, R. W., O'Connor, P. M., Rossney, A., Götz, F., Hill, C., Cotter, P. D., Ross, R. P., and Tagg, J. R. (2010) Production of the Bsa lantibiotic by community-acquired *Staphylococcus aureus* strains, *J. Bacteriol. 192*, 1131-1142.

58. Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P. J., Joewono, I., Rusch, D., Chisholm, S. W., and van der Donk, W. A. (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria, *Proc. Natl. Acad. Sci. U. S. A. 107*, 10430-10435.

59. Widdick, D. A., Dodd, H. M., Barraille, P., White, J., Stein, T. H., Chater, K. F., Gasson, M. J., and Bibb, M. J. (2003) Cloning and engineering of the cinnamycin biosynthetic gene cluster from *Streptomyces cinnamoneus cinnamoneus* DSM 40005, *Proc. Natl. Acad. Sci. U. S. A. 100*, 4316-4321.

60. Boakes, S., Cortés, J., Appleyard, A. N., Rudd, B. A., and Dawson, M. J. (2009) Organization of the genes encoding the biosynthesis of actagardine and engineering of a variant generation system, *Mol. Microbiol. 72*, 1126-1136.

61. Boakes, S., Appleyard, A. N., Cortes, J., and Dawson, M. J. (2010) Organization of the biosynthetic genes encoding deoxyactagardine B (DAB), a new lantibiotic produced by *Actinoplanes liguriae* NCIMB41362, *J. Antibiot. 63*, 351-358.

62. Holtsmark, I., Mantzilas, D., Eijsink, V. G., and Brurberg, M. B. (2006) Purification, characterization, and gene sequence of michiganin A, an actagardine-like lantibiotic produced by the tomato pathogen *Clavibacter michiganensis* subsp. *michiganensis*, *Appl. Environ. Microbiol. 72*, 5814-5821.

63. Kodani, S., Hudson, M. E., Durrant, M. C., Buttner, M. J., Nodwell, J. R., and Willey, J. M. (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene *ramS* in *Streptomyces coelicolor*, *Proc. Natl. Acad. Sci. U. S. A. 101*, 11448-11453.

64. Paetzel, M., Karla, A., Strynadka, N. C., and Dalbey, R. E. (2002) Signal peptidases, *Chem. Rev. 102*, 4549-4580.

65. Hallen, H. E., Luo, H., Scott-Craig, J. S., and Walton, J. D. (2007) Gene family encoding the major toxins of lethal *Amanita* mushrooms, *Proc. Natl. Acad. Sci. U. S. A. 104*, 19097-19101.

66. Walton, J. D., Hallen-Adams, H. E., and Luo, H. (2010) Ribosomal biosynthesis of the cyclic peptide toxins of *Amanita* mushrooms, *Biopolymers 94*, 659-664.

67. Szeltner, Z., and Polgár, L. (2008) Structure, function and biological relevance of prolyl oligopeptidase, *Curr. Protein Pept. Sci. 9*, 96-107.

68. Minasov, G., Kuhn, M., Ruan, J., Halavaty, A., Shuvalova, L., Dubrovska, I., Winsor, J., Bagnoli, F., Falugi, F., Bottomley, M., Grandi, G., and Anderson, W. F. (2011) 1.95 Angstrom resolution crystal structure of epidermin leader peptide processing serine protease (EpiP) S393A mutant from *Staphylococcus aureus*. PDB ID: 3T41, Center for Structural Genomics of Infectious Diseases (CSGID).

69. Kelley, L. A., and Sternberg, M. J. (2009) Protein structure prediction on the Web: a case study using the Phyre server, *Nat. Protoc. 4*, 363-371.

70. Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. Papain, *Biochem. Biophys. Res. Commun. 27*, 157-162.

71. Gros, P., Betzel, C., Dauter, Z., Wilson, K. S., and Hol, W. G. (1989) Molecular dynamics refinement of a thermitase-eglin-c complex at 1.98 A resolution and comparison of two crystal forms that differ in calcium content, *J. Mol. Biol. 210*, 347-367.

72. Heinz, D. W., Priestle, J. P., Rahuel, J., Wilson, K. S., and Grutter, M. G. (1991) Refined crystal structures of subtilisin novo in complex with wild-type and two mutant eglins. Comparison with other serine proteinase inhibitor complexes, *J. Mol. Biol. 217*, 353-371.

73. Siezen, R. J., Rollema, H. S., Kuipers, O. P., and de Vos, W. M. (1995) Homology modeling of the *Lactococcus lactis* leader peptidase NisP and its interaction with the precursor of the lantibiotic nisin, *Prot. Eng. 8*, 117-125.

74. Gron, H., Meldal, M., and Breddam, K. (1992) Extensive comparison of the substrate preferences of two subtilisins as determined with peptide substrates which are based on the principle of intramolecular quenching, *Biochemistry 31*, 6011-6018.

75. Wells, J. A., Powers, D. B., Bott, R. R., Graycar, T. P., and Estell, D. A. (1987) Designing substrate specificity by protein engineering of electrostatic interactions, *Proc. Natl. Acad. Sci. U. S. A. 84*, 1219-1223.

APPENDIX. INVESTIGATIONS ON NOVEL PHOSPHONATE NATURAL PRODUCTS

A.1. INTRODUCTION

Over the last several decades, bacterial resistance to known classes of antibacterial agents has been rising at an alarming rate, mainly due to the widespread overuse of antibiotics (*1*). Currently, more than 70% of bacteria that give rise to hospital-acquired infections in the United States are resistant to at least one of the main antimicrobial agents (*2*). Furthermore, the rate of discovery of antibiotics and natural products has decreased over the past years, in part because current screening methods result in a high rate of rediscovery of previously identified compounds (*3*). Thus, there is an urgent need for new antibacterial drug classes not affected by resistance mechanisms already present in pathogenic bacteria and for new approaches to discover them in an efficient way. This chapter focuses on the discovery of novel phosphonate metabolites and on the study of their biosynthetic pathways.

Several natural phosphonates with biological activity have been reported. For instance, fosfomycin (Figure A.1), an antibiotic produced by some species of *Pseudomonas* and *Streptomyces* that is effective against a broad spectrum of Gram-negative and Gram-positive bacteria, has become the first choice for treatment of certain types of infections (*4*). Fosfomycin inactivates UDP-*N*-acetylglucosamine-3-*O*-enolpyruvyltransferase (MurA), an essential enzyme that catalyzes the first step in the biosynthesis of the bacterial cell wall. In addition to fosfomycin, several other characterized phosphonates have shown biological activity (*5*). The bioactivity of this

class of compounds can be explained by the structural similarity of phosphonic acids with natural phosphate esters and carboxylic acids, as well as by the chemical stability of the C-P bond, which allows them to act as potent enzyme inhibitors.

The most abundant and first natural phosphonate to be isolated was 2-aminoethylphosphonic acid (AEP). Although this compound has been found in many species from bacteria to lower plants and animals, the molecule has no apparent biological function (*6*). However, AEP is also found as a head group in membrane phosphonolipids or as a side group in polysaccharides and glycoproteins with extracellular location. The inert character of phosphonate-containing compounds may generate structural components that are resistant to hydrolysis, oxidation, or other metabolic processes (*6, 7*).

After the discovery of fosfomycin, several other antibiotics have been isolated from other species of *Streptomyces, Pseudomonas,* and *Bacillus*, among others. The compound FR-900098 was isolated from cultures of *S*. *rubellomurinus (8)*, followed by the discovery of its homologs FR-31564 (fosmidomycin) and FR-32863 isolated from *S. lavendulae,* and FR-33289 from *S. rubellomurinus* subs. *indigoferis* (Figure A.1) (*9*). These compounds were found to have potent activity against *Trypanosoma cruzi* and various bacterial pathogens. Other isolated natural phosphonates include phosphonothrixin produced by *Saccharothrix sp*. (*10*) and phosphinothricin tripeptide (or bialaphos) isolated from *S. hygroscopicus* and *S. viridochromogenes*, both of which are being broadly used in agriculture as herbicides (*11*). Furthermore, plumbemycin from *S. plumbeus* (*12*), fosfazinomycin from *S. lavedofoliae* (*13*), SF-2312 from *Micromonospora* sp. (*14*), and dehydrophos from *S. luridus* (*15*) have also shown

antimicrobial activity. Finally, the phosphonates denoted as K-4 (*16*) and K-26 (*17*), isolated from *Actinomadura spiculosospora* and *Astrosporangium hypotensionis,* respectively, are potent ACE (angiotensin-converting enzyme) inhibitors that show promise in the treatment of high blood pressure (Figure A.1).

Despite the evident utility of phosphonates, the information about their biosynthesis is limited. Currently, the biosynthetic gene clusters of only a few C-P bond-containing natural products has been reported, including 2-aminoethylphosphonate (*18*), fosfomycin (*4*), phosphinothricin tripeptide (*19, 20*), FR-900098 (*21*), and dehydrophos (*22*). Considering the biological properties of phosphonates and the ubiquitous role of their analog phosphate-esters and carboxylic acids in biology, the discovery of many other novel phosphonates with inhibitory activity is highly anticipated. In addition, knowledge of their biosynthetic pathways could lead to production systems with higher yields and to engineered pathways that generate novel antibiotics.

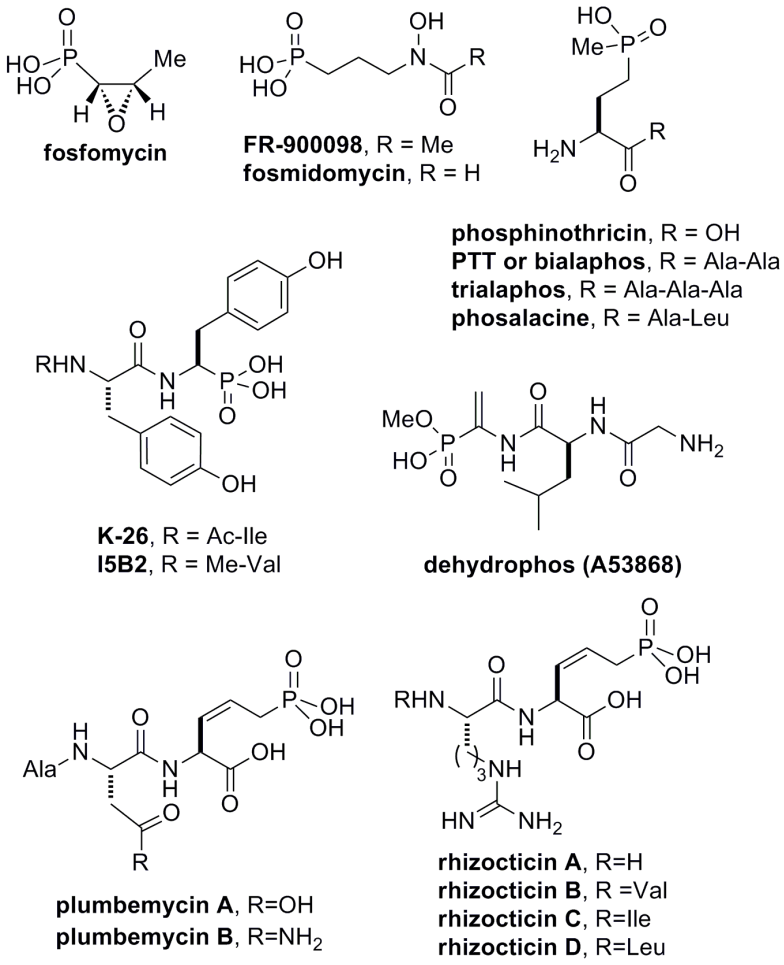


Figure A.1. Chemical structures of several representative phosphonate antibiotics.

All the known natural phosphonates, with the possible exception of K-26 and its analogs (*23*), are believed to have a common biosynthetic step, which is the rearrangement of phosphoenolpyruvate (PEP) to phosphonopyruvate (PnPy) catalyzed by the enzyme PEP mutase (Ppm) (Figure A.2). As a consequence, screening genomic DNA for the *ppm* gene by PCR using degenerate primers can be an efficient strategy to identify strains that produce phosphonates and thus, to discover novel bioactive compounds. Moreover, the generally observed clustering of secondary metabolite genes allows for identification of entire sets of biosynthetic gene clusters. Previously,

Prof. William Metcalf and coworkers identified a set of microorganisms that potentially produce phosphonate antibiotics using this methodology. The following section describes efforts in the construction of genomic DNA libraries and in the identification of gene clusters involved in the biosynthesis of novel phosphonates. Based on this information, two strains containing gene clusters that potentially encode enzymes for the biosynthesis of small molecules were selected for further study. By screening for production of phosphonates under different growth conditions and analyzing the concentrated culture extracts by ^{31}P NMR spectroscopy, two potentially novel phosphonates were identified.

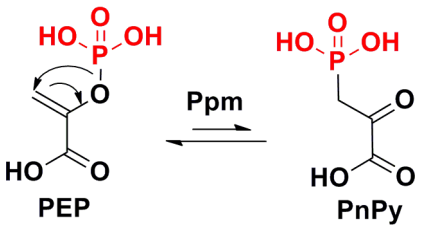


Figure A.2. Rearrangement of phosphoenolpyruvate (PEP) to phosphonopyruvate (PnPy) catalyzed by PEP mutase (Ppm).

A.2. EXPERIMENTAL METHODS

Materials, organisms, media, and general methods

Chemical reagents and media components used in this study were purchased from Sigma-Aldrich, Thermo Fisher Scientific, or Difco, unless otherwise specified, and were used without further purification. The strains are listed in Table A.1. *Escherichia coli* and *Bacillus subtilis* strains were routinely grown in LB solid agar or broth at 37 °C or 30 °C, respectively. *Streptomyces* strains were propagated in malt extract-yeast extract-glucose medium (MYG) or ISP medium 2 (ISP2) at 30 °C, unless otherwise specified. For agar diffusion bioactivity assays, aliquots of 25 mL of agar medium

inoculated with an overnight culture (1/100 dilution) were poured into a sterile plate. Aliquots of antibacterial compounds were placed into wells made on the solidified agar and the plates were incubated at 37 or 30 °C overnight. The bioactivity of the sample was confirmed if a clear growth inhibition zone surrounding the well was formed.

The construction of the genomic library of strain *S. mauvecolor* WM6391 and sequencing of a *ppm*-positive fosmid were performed using the protocol described for *Staphylococcus epidermidis* 15X154 in chapter 2. The degenerate forward CHIpepmutF1 (5'-CGC CGG CGT CTG CNT NGA RGA YAA-3') and reverse CHIpepmutR2 (5'-GGC GCG CAT CAT GTG RTT NGC VYA-3') primers that recognize conserved regions in the *ppm* gene were used for library screening. The bioinformatic analysis of the gene clusters were performed using several tools, including SMART (*24*), NCBI-CD Search (*25*), PROSITE Web Server (*26*), Phyre 3D-PSSM (*27*), NRPSPredictor (*28*), TMHMM server (*29*), and SignalP 3.0 (*30*). The putative phosphonate producer WM6391 was further studied by Mr. Joel Cioni, a current graduate student in the Metcalf laboratory.

Screening of conditions for production of phosphonates

Single colonies of the *Streptomyces* strains WM6372 and WM4235 were inoculated in MYG media for 2 to 4 days at 30 °C. For testing of phosphonate production in different solid media, aliquots of MYG cultures were spread on agar plates and incubated at 30 °C for 5 to 15 days. ISP2, ISP medium 4 (ISP4), Difco nutrient agar (DNA), mannitol soya flour medium (MS) (*31*), supplemented minimal medium (SMMS) (*32*), and minimal medium (MM) with glucose or glycerol (*33*) were used. For the

extraction of phosphonates, the agar was liquefied by freezing at −20 °C followed by thawing at room temperature, and the liquid fractions were decanted. See also *Notebook IV, page 05*.

Table A.1. Microorganisms and plasmids used in this study

Strain	Relevant characteristics	Source or reference
Escherichia coli		
WM4489	*E. coli* DH10B derivative: *mcrA* Δ(*mrr hsdRMS mcrBC*) *ϕ80*(Δ*lacM15*) Δ*lacX74 endA1 recA1 deoR* Δ(*ara-leu*)*7697 araD139 galU galK nupG rpsL* *λattB*::pAE12(*PrhaB*::*trfA33* Δ*oriR6K-cat*::*frt5*)	(*21*)
WM2785	Indicator strain *E. coli* BW26678 derivative: *lacIq, rrnB3,* Δ(*lacZ4787*), *hsdR514, attP22*(*EcoB*), Δ(*araBAD)567,* Δ(*rhaBAD*)*568, rph−1,* Δ(*phnC-P),* Δ(*phoA*) Expresses the phosphonate uptake system encoded by *phnCDE* under the control of a Ptac promoter	W. Metcalf, Univ. Illinois
MMG127	Derivative of WM4489 containing a *ppm* positive fosmid isolated from the *S. mauvecolor* WM6391genomic library	B. Li and W. Metcalf
Bacillus subtilis ATCC6633	Indicator strain	ATCC[a]
Streptomyces		
mauvecolor WM6391	*ppm* positive strain, Illinois soil isolate	W. Metcalf, Univ. Illiniois
verne WM4235	*ppm* positive strain, Alaskan soil isolate	W. Metcalf, Univ. Illinois
sp. WM6372	*ppm* positive strain, Illinois soil isolate	W. Metcalf, Univ. Illinois

[a]ATCC, American Type Culture Collection, Manassas, VA.

For liquid cultures, aliquots of MYG cultures were diluted (1/100) in the trial media and incubated at 30 °C for 2 to 7 days with shaking. MYG, yeast extract-malt extract medium (YEME) (*31*), Difco nutrient broth (DNB), and minimal liquid medium (NMMP) (*34*) were tested. For the induction of phosphonates production with D/L-2-

amino-3-phosphonopropionic acid (D/L-phosphonoalanine, D/L-PnAla), ISP4 media without agar and substituting K_2HPO_4 for D/L-PnAla (0.1 g/L, Sigma) was used for the initial trials and the bacterial cultures were incubated for 14 days at 30 °C with shaking. In the following trials, a modular media composed of freshwater base (NaCl 17.1 mM, $MgCl_2 \cdot 6H_2O$ 1.97 mM, $CaCl_2 \cdot 2H_2O$ 0.15 mM, KCl 6.71 mM), glucose (55 mM) or soluble starch (29 mM), NH_4Cl (10 mM), Na_2SO4 (1 mM), DL-PnAla (4 mM), AEP (10 mM), trace elements ($FeSO_4 \cdot 7H_2O$, 2.1 ppm; $MnCl_2 \cdot 4H_2O$, 0.1 ppm; $CoCl_2 \cdot 6H_2O$, 0.19 ppm; H_3BO_3, 0.03 ppm; $NiCl_2 \cdot 6H_2O$, 0.024 ppm; $CuCl_2 \cdot 2H_2O$, 0.002 ppm; $ZnSO_4 \cdot 7H_2O$, 0.144 ppm; $Na_2MoO_4 \cdot 2H_2O$, 0.036 ppm; $Na_2WO_4 \cdot 2H_2O$, 0.025 ppm; $Na_2SeO_3 \cdot 5H_2O$, 0.006 ppm), and MOPS (50 mM, pH 7.2) as buffering reagent was used and the cultures were incubated for 7 days at 30 °C with shaking. See also *Notebook IV, page 50*.

The presence of phosphonates was determined using ^{1}H-decoupled ^{31}P-NMR spectroscopy. Liquid fractions or culture supernatants were passed through a 0.2 µm filter, concentrated (4× to 30×) by lyophilization, and resuspended in H_2O supplemented with 20% D_2O as a lock solvent. The spectra were externally referenced to an 85% phosphoric acid standard (0 ppm) and acquired at room temperature on a Varian Unity Inova-600 spectrometer equipped with a 5-mm Varian 600DB AutoX probe with ProTune accessory and tuned at 242.789 MHz at the Varian Oxford Center for Excellence in NMR laboratory at the University of Illinois, Urbana-Champaign.

To reduce the complexity of the sample, phosphates and cyclophosphate nucleotides, which may generate false positive signals in the phosphonate range in ^{31}P NMR spectroscopy, were hydrolyzed by incubation with calf intestinal alkaline phosphatase (New England Biolabs) and phosphodiesterase I from *Crotalus atrox*

(Sigma) in reaction buffer (HEPES, 25 mM, pH 7.5) at 37 °C for 4 h. Finally, since the C-P bond in phosphonates, but not the C-O-P bond in phosphate esters, is remarkable stable under acidic conditions, samples were treated with hydrochloric acid (1 N) at 100 °C under reflux for 4 h, followed by analysis by ^{31}P NMR spectroscopy as described above to confirm the presence of phosphonates. See also *Notebook IV, page 62*.

A.3. RESULTS AND DISCUSSION

Genomic DNA library construction of a *ppm*-positive *Streptomyces* strain, sequencing, and annotation of phosphonate gene clusters

In collaboration with Dr. Bo Li, a former graduate student in the van der Donk laboratory, a fosmid-based genomic DNA library for *S. mauvecolor* WM6391 in *E. coli* was constructed under the direction of Prof. William Metcalf and Mr. Junkai Zhang. Briefly, genomic DNA was isolated in fragments longer than 200 kb and partially digested. Then, cloning of 30-50 kb fragments into a cosmid followed by lambda phage transduction resulted in a library of fosmid-containing clones in *E. coli* WM4489. Finally, a fraction of the fosmid library was screened by PCR for the *ppm* gene using degenerate primers that recognize conserved nucleotide sequences on the gene and three positive clones were isolated. To obtain the sequence of one of the fosmids, a library of transposon insertions into the isolated fosmid was prepared and a set of 192 clones were sequenced using primers binding at each end of the transposon. After assembling of the DNA sequences, two contigs were obtained.

The hypothetical function and location of the putative proteins were determined using several bioinformatic tools (Figure A.3). The biosynthetic gene cluster in WM6391

contains three genes with high homology to genes already identified in the FR-900098 cluster and that encode a Ppm (FrbD), a homocitrate synthase for the second step of the biosynthesis (FrbC) and an aconitase isomerase for the third step (FrbA) (Figure A.4) (*21*). Thus, this cluster may encode the biosynthesis of a small metabolite and it was further characterized by Mr. Joel Cioni, a graduate student in the Metcalf laboratory.

In a collaborative effort, genomic libraries of six additional *Streptomyces sp*. that potentially contain the genetic information for phosphonate biosynthesis were obtained by other members of the Mining Microbial Genomes Theme at the Institute for Genomic Biology. After sequencing, five of these clusters were assembled and the open reading frames encoding proteins involved in the biosynthesis of phosphonates were identified (Figure A.6 to A.10). The remaining gene cluster was annotated by Dr. Tyler Johannes, a former graduate student in the Zhao laboratory (Figure A.5).

The gene clusters from the strains *S. goshikiensis* WM6368, *S. virginiae* WM6349, and *Streptomyces sp.* WM6373 are similar and may produce identical phosphonates (Figures A.5 to A.7). These clusters are likely to produce AEP, a common component of the bacterial cell wall (*18*), that may be tethered to a polyketide chain as suggested by the presence of polyketide synthase (PKS) I and II genes in the vicinity of the phosphonate cluster (Figure A.5). The cluster from *Streptomyces* sp. WM6352, in addition to phosphonate-related genes encoding for AEP, contains genes for the biosynthesis of TDP-L-rhamnose, a common component of exo-polysaccharides and glycoproteins (Figure A.8). Two interesting gene clusters from the strains *Streptomyces* sp. WM6372 and *S. verne* WM4235, presumably involved in the

synthesis of non-ribosomal peptides, were selected for further analysis (Figures A.9 and A.10, Tables A.2 and A.3). For both clusters, the gaps between the sequences were filled by primer walking.

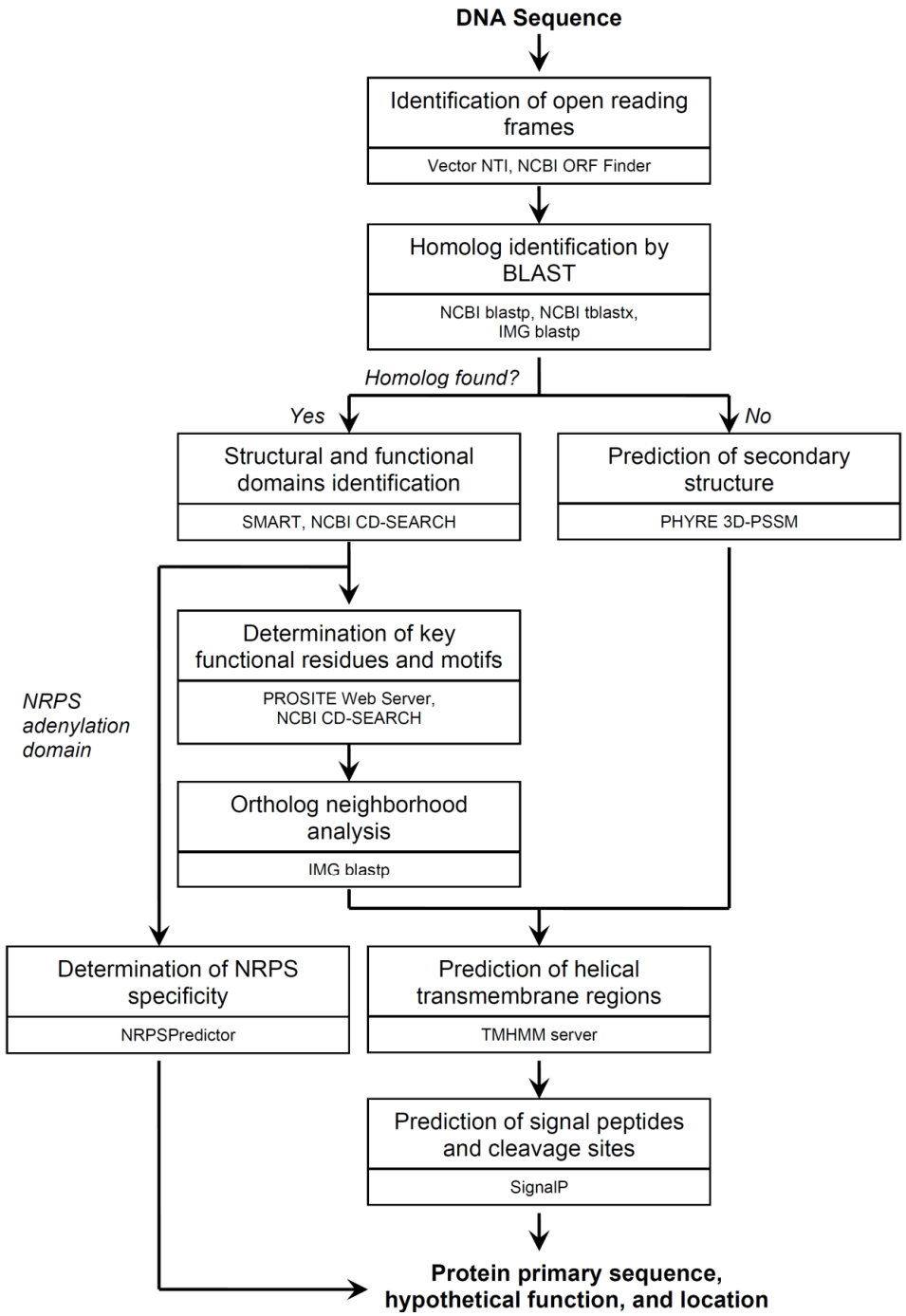


Figure A.3. Workflow for the bioinformatic analysis of the open reading frames identified after assembling the DNA sequences (*24-30*).

.

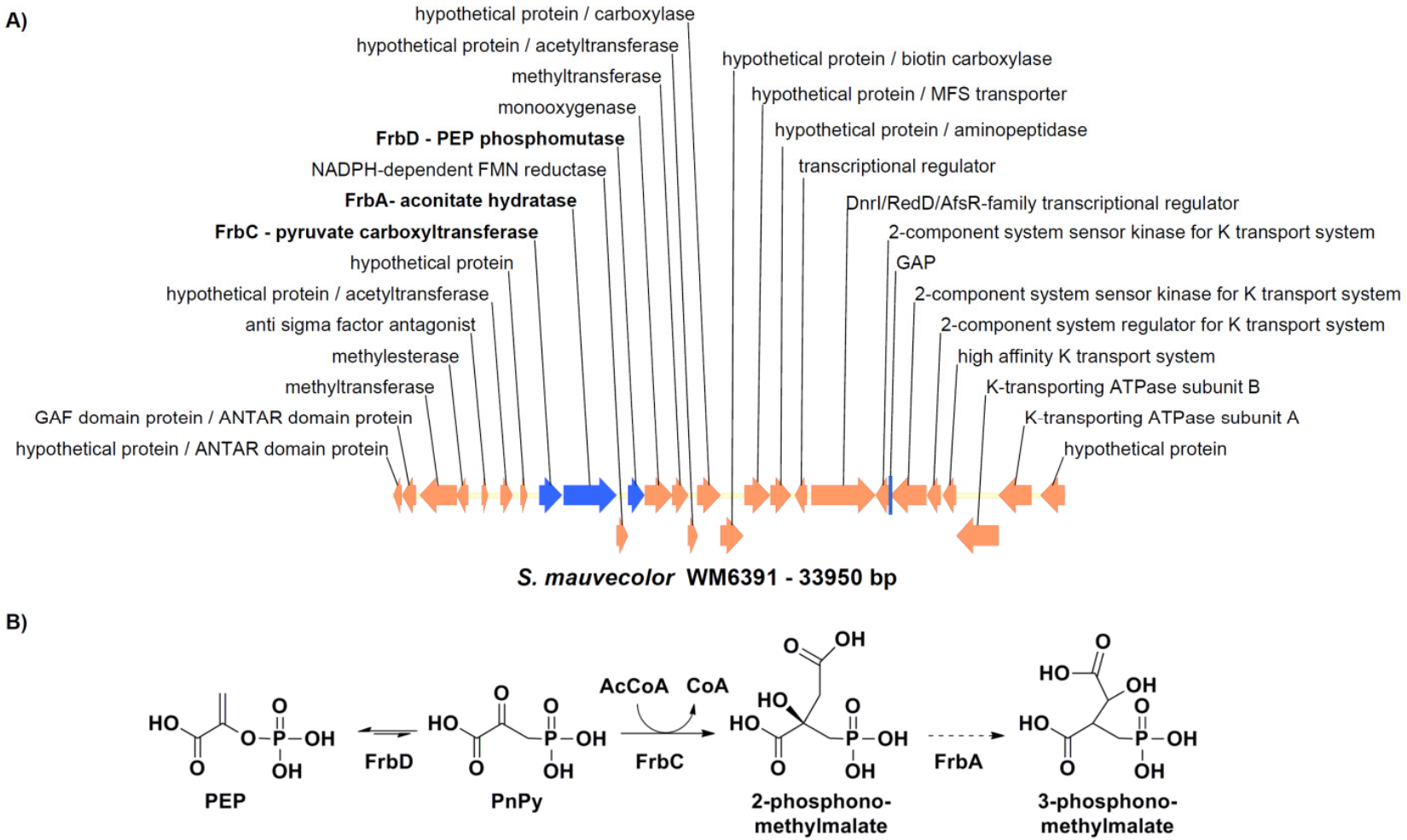


Figure A.4. Open reading frame map of the sequenced fosmid containing the phosphonate gene cluster from *S. mauvecolor* WM6391 (or *E. coli* MMG127 fosmid) (A) and proposed reactions for the first three biosynthetic steps (B).

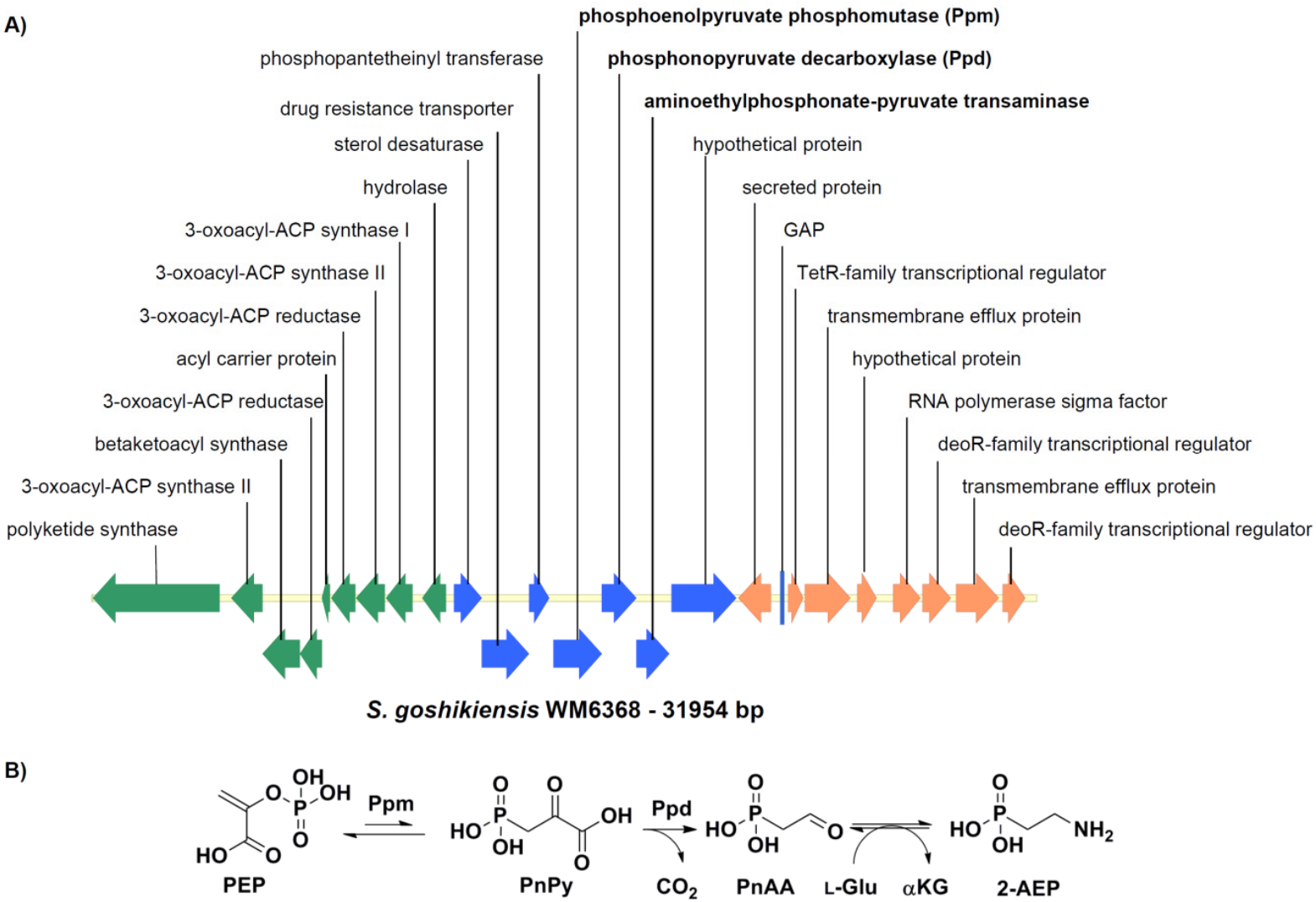


Figure A.5. Open reading frame map of the sequenced fosmid containing the phosphonate gene cluster from *S. goshikiensis* WM6368 (A) and proposed reactions for the first three biosynthetic steps (B). Annotation by Dr. Tyler Johannes. Green: PKS genes, blue: putative phosphonate gene cluster, orange: other genes.

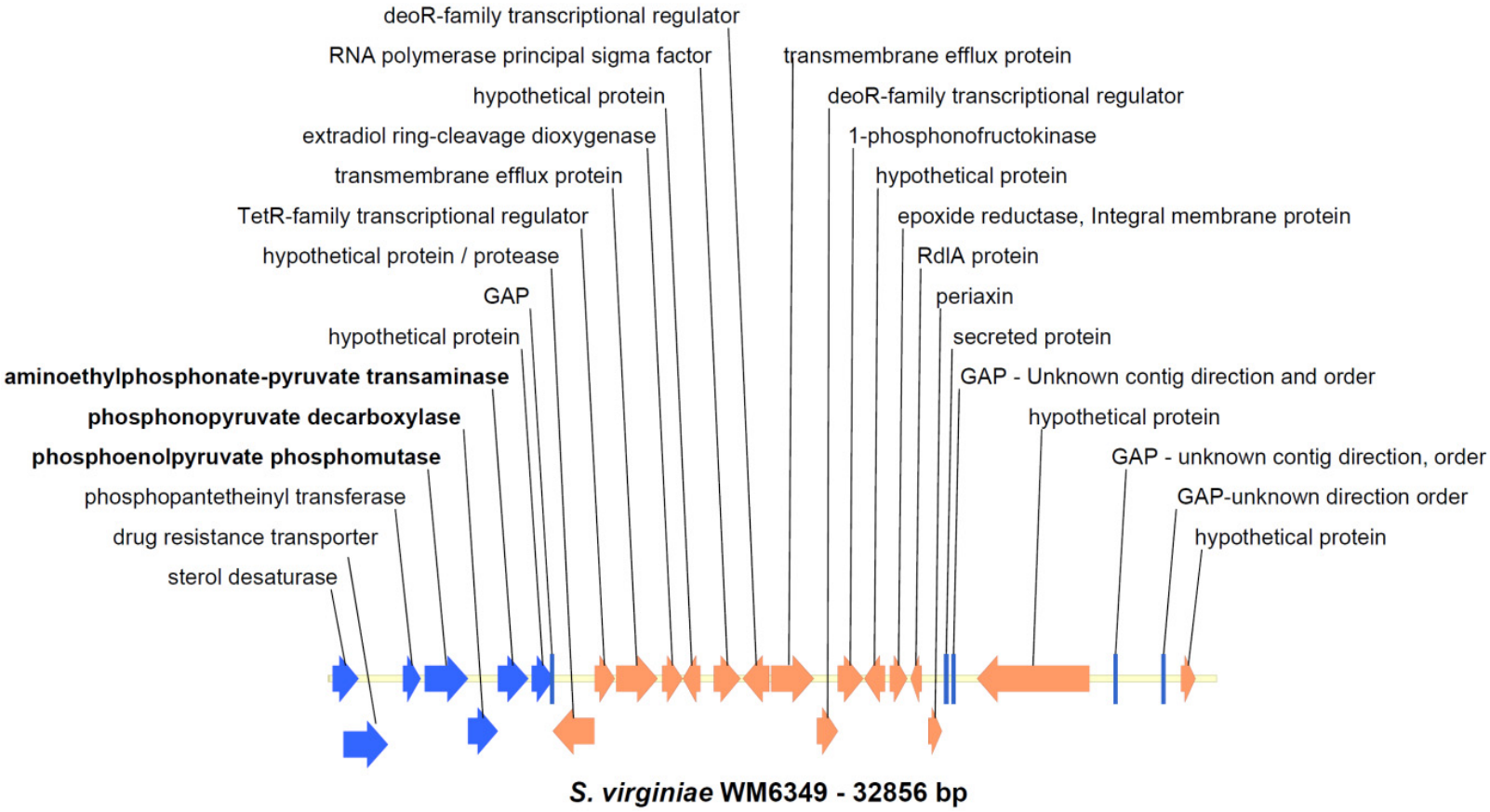


Figure A.6. Open reading frame map of the sequenced fosmid containing the phosphonate gene cluster from *S. virginiae* WM6349. Open reading frames were assigned using the BLAST program at the NCBI website. Blue: putative phosphonate gene cluster, orange: other genes.

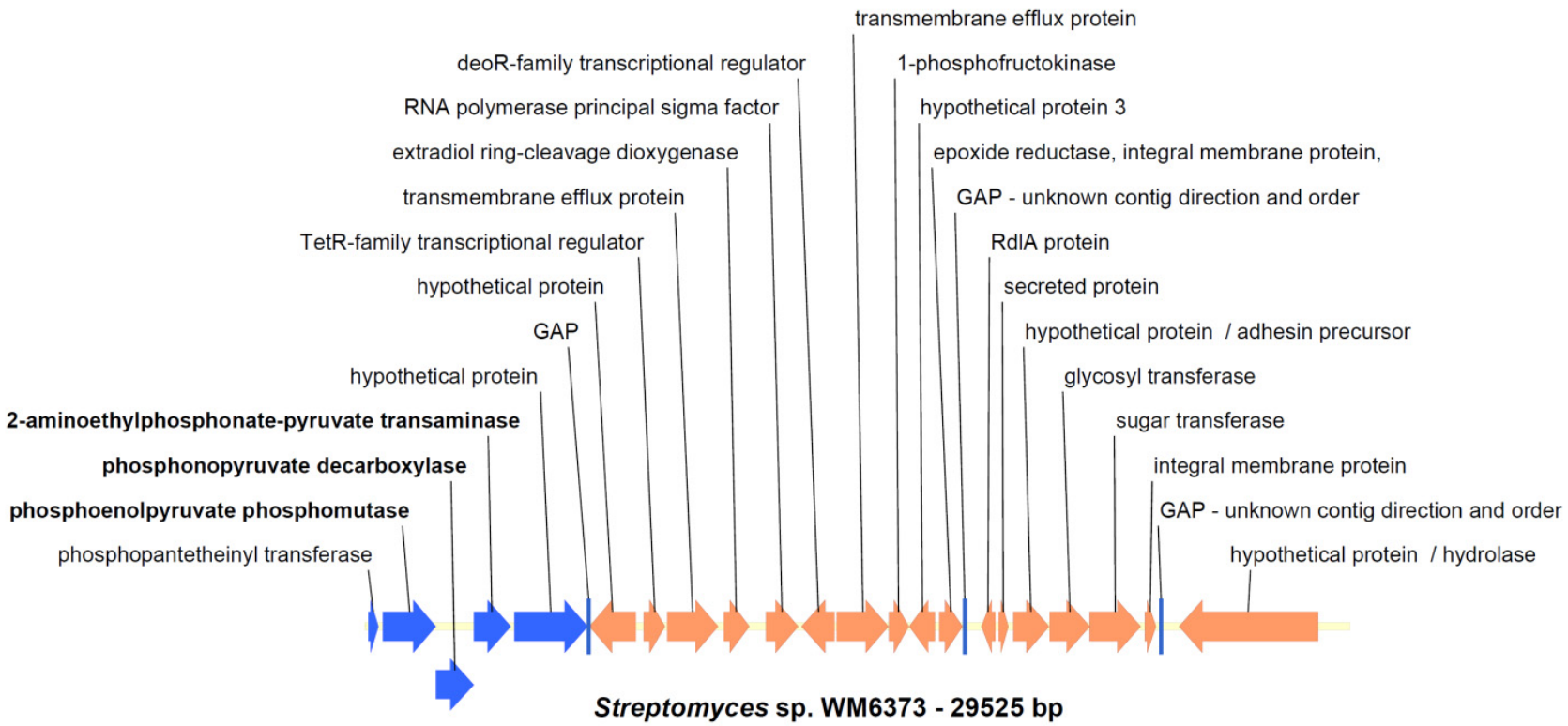


Figure A.7. Open reading frame map of the sequenced fosmid containing the phosphonate gene cluster from *Streptomyces* sp. WM6373. Open reading frames were assigned using the BLAST program at the NCBI website. Blue: putative phosphonate gene cluster, orange: other genes.

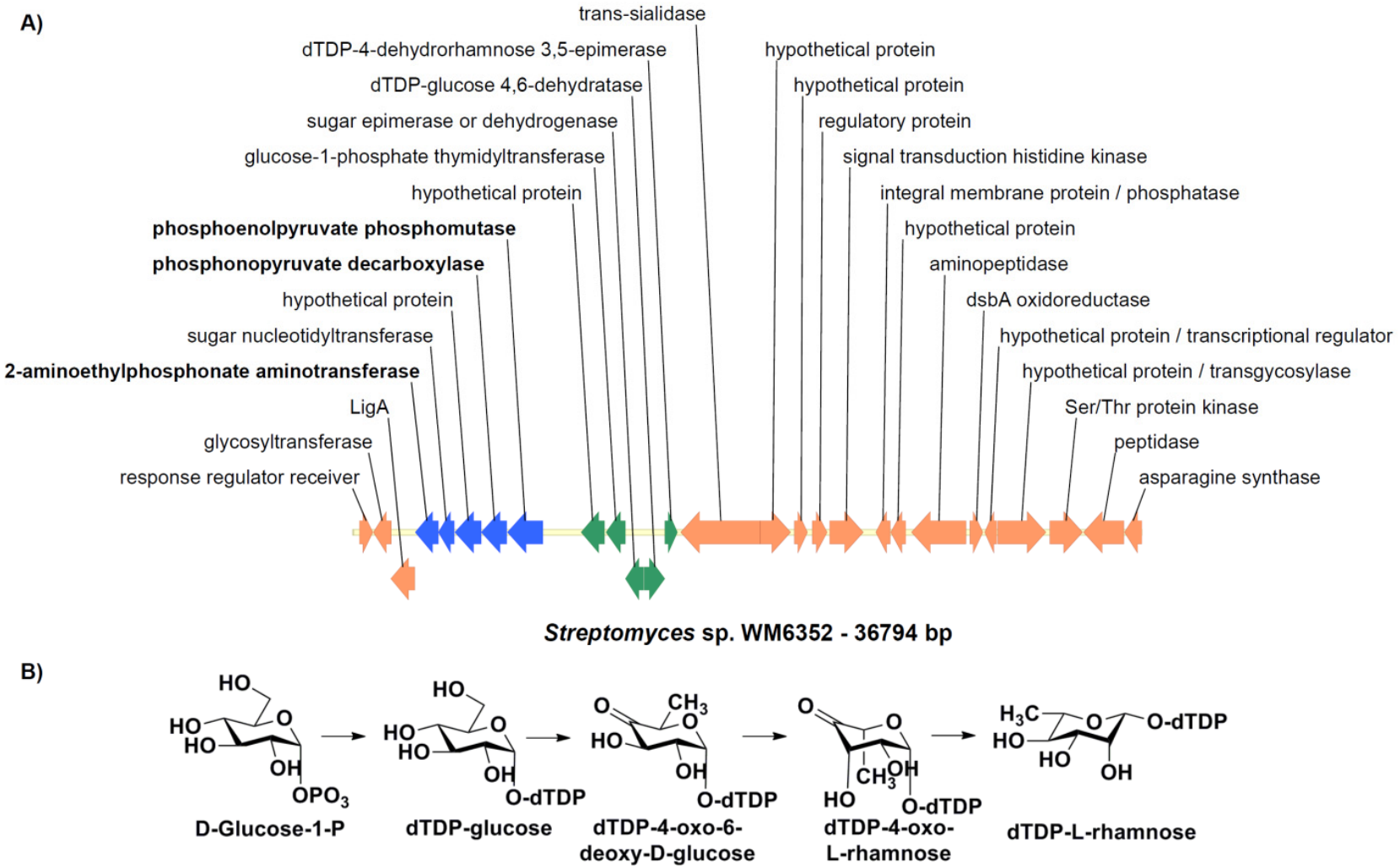


Figure A.8. Open reading frame map of the sequenced fosmid containing the phosphonate gene cluster from *Streptomyces* sp. WM6352. Open reading frames were assigned using the BLAST program at the NCBI website. Blue: putative phosphonate gene cluster, green: TDP-L-rhamnose biosynthesis, orange: other genes.

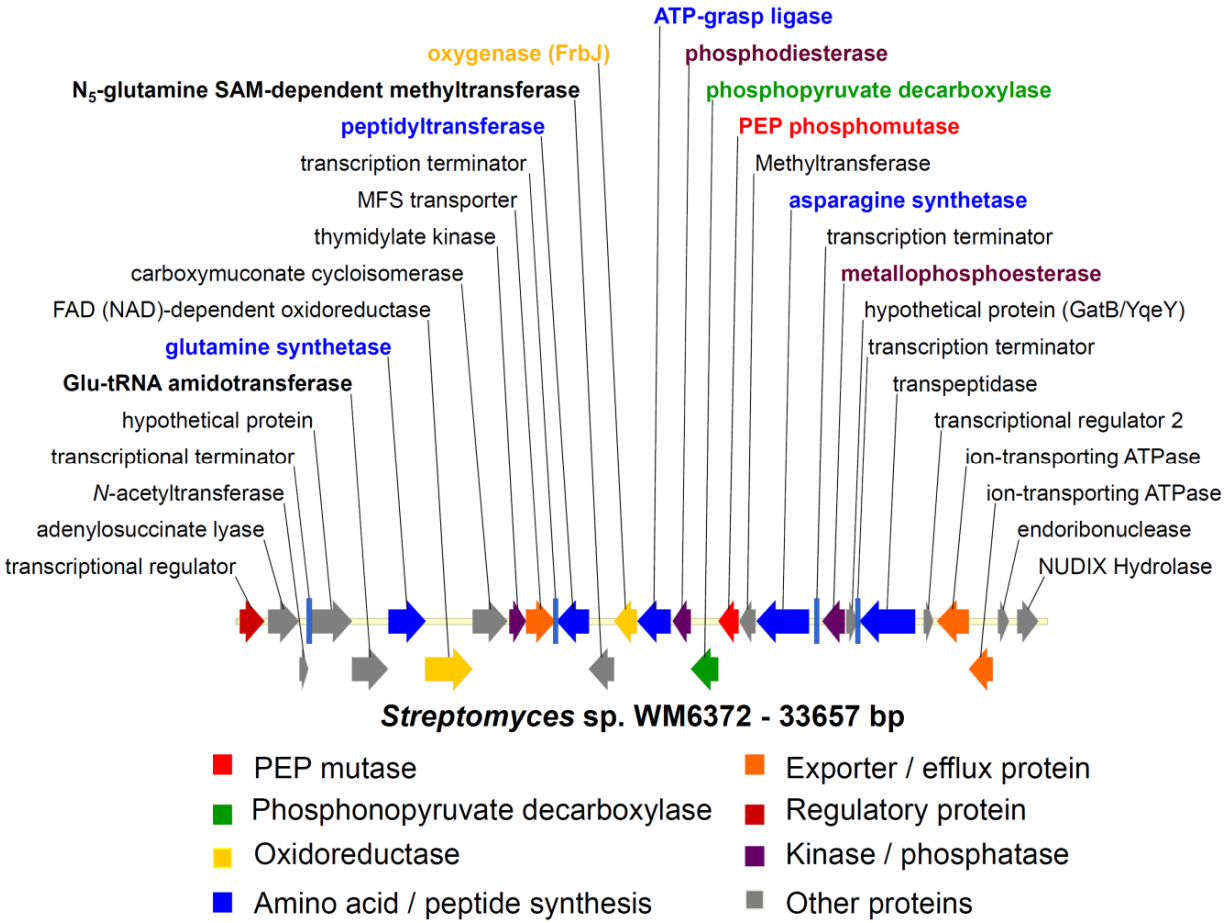


Figure A.9. Open reading frame map of the sequenced fosmid containing the phosphonate gene cluster from *Streptomyces* sp. WM6372. Open reading frames were assigned using the BLAST program at the NCBI website and other bioinformatic tools.

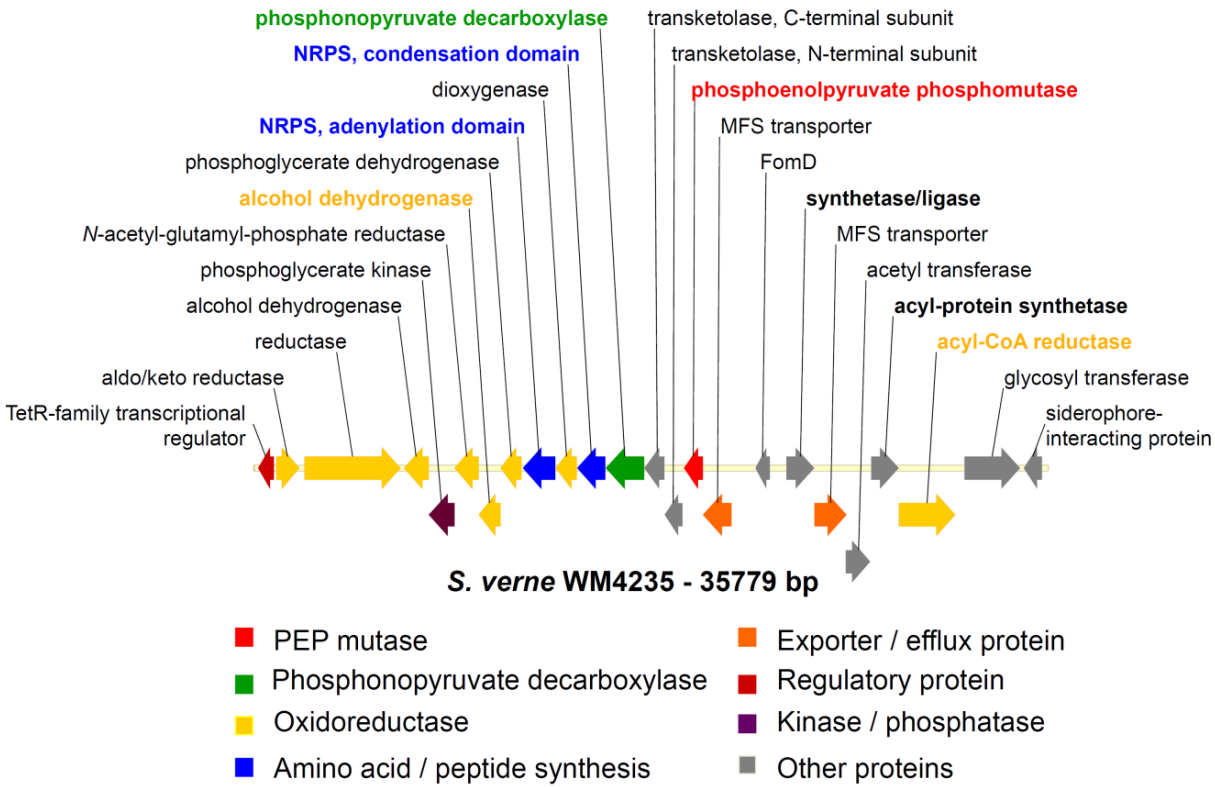


Figure A.10. Open reading frame map of the sequenced fosmid containing the phosphonate gene cluster from *S. verne* WM4235. Open reading frames were assigned using the BLAST program at the NCBI website and other bioinformatic tools.

Table A.2. Open reading frames analysis of the *Streptomyces* sp. WM6372 gene cluster using BLASTp at the NCBI website.

Predicted protein	Protein homology	Identity
Transcriptional regulator	dbj|BAC71516.1| putative transcriptional regulator [*S. avermitilis* MA-4680]	47%
Adenylosuccinate lyase	emb|CAJ89132.1| putative adenylosuccinate lyase [*S. ambofaciens* ATCC 23877]	63%
N-Acetyltransferase	emb|CAJ89131.1| putative acetyltransferase [*S. ambofaciens* ATCC 23877]	64%
Hypothetical protein	emb|CAJ89135.1| hypothetical protein [*S. ambofaciens* ATCC 23877]	33%
Glu-tRNA amido transferase	gb|ABP54656.1| Amidase [*Salinispora tropica* CNB-440]	51%
Glutamine synthetase	emb|CAK51247.1| putative glutamine synthetase [*S. ambofaciens*]	54%
FAD(NAD)-dependent oxidoreductase	dbj|BAC73537.1| hypothetical protein [*S. avermitilis* MA-4680] gb|ABJ69048.1| Predicted FAD(NAD)-dependent oxidoreductase [*Lactobacillus casei* ATCC 334]	58% 35%
Carboxymuconate cycloisomerase	emb|CAK51228.1| putative 3-carboxymuconate cycloisomerase [*S. ambofaciens*]	61%
Thymidylate kinase	emb|CAA83856.1| unknown [*Pseudomonas syringae*] gb|ABE54453.1| Thymidylate kinase-like protein [*Shewanella denitrificans* OS217]	37% 34%
MSF transporter	gb|ABM09775.1| putative major facilitator superfamily (MFS) transporter [*Arthrobacter aurescens* TC1]	33%
Peptidyltransferase	emb|CAA60475.1| orfD - rapamycin gene cluster [*S. hygroscopicus*]	30%
N_5-glutamine SAM-dependent methyltransferase	gb|AAK64442.1|AF377339_3 protoporphyrinogen oxidase HemK [*Myxococcus xanthus*] emb|CAJ93170.1| predicted methyltransferase [*Ralstonia eutropha* H16]	39% 39%
Oxygenase (FrbJ)	gb|EDP64847.1| hypothetical protein BAL199_05284 [alpha *proteobacterium* BAL199] gb|ABB90399.1| FrbJ [*S. rubellomurinus*]	34% 28%
ATP-grasp ligase	gb|AAU23668.1| Carbamoyl-phosphate synthase L chain [*Bacillus licheniformis* ATCC 14580] ref|YP_091723.1| acetyl-CoA carboxylase biotin carboxylase subunit [*B. licheniformis* ATCC 14580]	29% 29%
Phosphodiesterase	emb|CAJ89340.1| putative phosphodiesterase [*S. ambofaciens* ATCC 23877]	54%
Phosphonopyruvate decarboxylase	emb|CAB45023.1| putative phosphonopyruvate decarboxylase [*Amycolatopsis orientalis*]	48%
Phosphoenolpyruvate phosphomutase	emb|CAN97474.1| Phosphoenolpyruvate mutase [*Sorangium cellulosum* 'So ce 56']	47%

Table A.2. *Continuation.*

Predicted protein	Protein homology	Identity
Methyltransferase	gb|EAV40562.1| hypothetical protein SIAM614_16737 [*Stappia aggregata* IAM 12614]	46%
	gb|EDQ31853.1| methlytransferase, UbiE/COQ5 family protein [*Hoeflea phototrophica* DFL-43]	43%
Aspargine synthetase	emb|CAB16028.1| asparagine synthetase (glutamine-hydrolyzing) [*B. subtilis* subsp. *subtilis* str. 168]	37%
Metallophospho-esterase	emb|CAB41730.1| putative secreted protein [*S. coelicolor* A3(2)]	82%
	gb|ABL94102.1| metallophosphoesterase [*Mycobacterium* sp. KMS]	52%
Hypothetical protein (GatB/YqeY)	emb|CAB41729.1| conserved hypothetical protein [*S. coelicolor* A3(2)]	83%
	emb|CAL99661.1| GatB/YqeY [*Saccharopolyspora erythraea* NRRL 2338]	56%
Transpeptidase	emb|CAD55325.1| putative transpeptidase [*S. coelicolor* A3(2)]	71%
	dbj|BAC72295.1| putative penicillin-binding protein [*S. avermitilis* MA-4680]	70%
Transcriptional regulator	emb|CAB43030.1| hypothetical protein [*S. coelicolor* A3(2)]	88%
	dbj|BAC72296.1| putative WhiB-family transcriptional regulator [*S. avermitilis* MA-4680]	87%
Ion-transporting ATPase	emb|CAB45557.1| putative ion-transporting ATPase [*S. coelicolor* A3(2)]	82%
Ion-transporting ATPase	emb|CAB45559.1| putative ion-transporting ATPase [*S. coelicolor* A3(2)]	76%
Endoribonuclease	dbj|BAC72300.1| hypothetical protein [*S. avermitilis* MA-4680]	84%
	gb|ABY24518.1| translation initiation inhibitor [*Renibacterium salmoninarum* ATCC 33209]	74%
NUDIX hydrolase	dbj|BAC72301.1| hypothetical protein [*S. avermitilis* MA-4680]	78%
	gb|ABL79881.1| NUDIX hydrolase [*Nocardioides* sp. JS614]	54%

Note: Results are from a BLASTp search of the GenBank protein database on January 2008.

Table A.3. Open reading frames analysis of the *S. verne* WM4235 gene cluster using BLASTp at the NCBI website.

Predicted protein	Protein homology	Identity
TetR-family transcript. regulator	emb|CAJ88304.1| putative TetR-family transcriptional regulator [*S. ambofaciens* ATCC 23877]	74%
Aldo/keto reductase	dbj|BAB50827.1| oxidoreductase; MocA [*Mesorhizobium loti* MAFF303099]	81%
	gb|ABR61565.1| aldo/keto reductase [*Sinorhizobium medicae* WSM419]	79%
Reductase	dbj|BAC70041.1| putative assimilatory nitrate reductase [*S. avermitilis* MA-4680]	69%
Alcohol dehydrogenase	gb|ABN51632.1| iron-containing alcohol dehydrogenase [*Clostridium thermocellum* ATCC 27405]	32%
Phosphoglycerate kinase	gb|EDO62266.1| hypothetical protein CLOLEP_01127 [*C. leptum* DSM 753]	29%
	dbj|BAD39227.1| phosphoglycerate kinase [*Symbiobacterium thermophilum* IAM 14863]	33%
N-acetyl-glutamyl-phosphate reductase	gb|EAZ54588.1| N-acetyl-gamma-glutamyl-phosphate reductase [*P. aeruginosa* C3719]	42%
Alcohol dehydrogenase	gb|EAX46939.1| Alcohol dehydrogenase GroES domain protein [*Thermosinus carboxydivorans* Nor1]	35%
Phosphoglycerate dehydrogenase	gb|AAL81518.1| phosphoglycerate dehydrogenase [*Pyrococcus furiosus* DSM 3638]	48%
NRPS, adenylation domain	ref|ZP_00111185.1| COG1020: Non-ribosomal peptide synthetase modules and related proteins [*Nostoc punctiforme* PCC 73102]	45%
Dioxygenase	gb|ABD89411.1| Phytanoyl-CoA dioxygenase [*Rhodopseudomonas palustris* BisB18]	35%
NRPS, condensation domain	gb|ABJ11576.1| pyoverdine synthetase D [*P. aeruginosa* UCBPP-PA14]	24%
Phosphonopyruvate decarboxylase	gb|ABH03014.1| IlvB [*Spirochaeta aurantia*]	35%
	gb|AAQ61550.1| acetolactate synthase [*Chromobacterium violaceum* ATCC 12472]	36%
Transketolase, C-terminal subunit	gb|ABR61955.1| Transketolase domain protein [*S. medicae* WSM419]	49%
Transketolase, N-terminal subunit	gb|ABW09872.1| Transketolase domain protein [*Frankia* sp. EAN1pec]	55%
Phosphoenolpyruvate phosphomutase	emb|CAJ65000.1| putative Phosphoenolpyruvate phosphomutase [*Frankia* alni ACN14a]	62%
MFS transporter	gb|ABM09775.1| putative major facilitator superfamily transporter [*Arthrobacter aurescens* TC1	34%
FomD	dbj|BAA32492.1| FomD [*S. wedmorensis*]	48%
Synthetase/ligase	emb|CAK13564.1| putative AMP-dependent synthetase/ligase [*P. entomophila* L48]	41%
MFS transporter	gb|AAD36408.1|AE001788_3 permease, putative [*Thermotoga maritima* MSB8]	22%
Acetyl transferase	emb|CAJ88185.1| putative modular polyketide synthase [*S. ambofaciens* ATCC 23877]	32%

Table A.3. *Continuation.*

Acyl-protein synthetase	emb|CAE16553.1| unnamed protein product [*Photorhabdus luminescens* subsp. *laumondii* TTO1]	51%
Acyl-CoA reductase	gb|ABA74299.1| conserved hypothetical protein [*P. fluorescens* PfO-1]	44%
Glycosyl transferase	dbj|BAC72261.1| putative glycosyltransferase [*S. avermitilis* MA-4680]	58%
Siderophore-interacting protein	dbj|BAC74424.1| putative siderophore-interacting protein [*S. avermitilis* MA-4680]	61%

Note: Results are from a BLASTp search of the GenBank protein database on January 2008.

In addition to Ppm and phosphonopyruvate decarboxylase (Ppd), an enzyme that catalyzes the decarboxylation of PnPy to produce phosphonoacetaldehyde (PnAA), the gene cluster from *Streptomyces* sp. WM6372 contains an ORF encoding a member of the ATP-grasp ligase superfamily. This protein might catalyze an ATP-dependent ligation of a phosphonate-containing amino acid to another amino acid to produce a dipeptide, with the formation of an acylphosphate intermediate as suggested for rhizocticins (Figure A.11) (*35, 36*). The alignment of the predicted protein sequence with that of D-alanine D-alanine ligase (Ddl) demonstrates that the most relevant residues involved in ATP and metal binding are conserved (*37*). Two other genes in the cluster encode ATP-dependent enzymes that catalyze the formation of carbon-nitrogen bonds: a glutamine synthetase homolog and an asparagine synthetase homolog. Gln synthetase is involved in the formation of L-Gln from L-Glu and ammonia, while Asn synthetase catalyzes the formation of L-Asn from L-Asp and L-Gln (Figure A.12). Different roles in the pathway can be proposed for these proteins. For instance, they might supply L-Asn for the formation of a peptide with a phosphonate containing amino acid; alternatively, they might have a role in self-protection if they encode resistant forms of target enzymes as suggested for FR-900098 (*21*), rhizocticins (*36*), and

cycloserine (*38*). If this is the case, the phosphonic acid could be an analog of L-Gln as in the case of phosphinothricin (*20*). Interestingly, a gene designated as peptidyltransferase, which has no significant primary sequence homology to proteins of known function, was found to have high predicted secondary structure similarity to FemX, a non-ribosomal peptidyl transferase that catalyzes the transfer of L-Ala from aminoacylated tRNA to the ε-amino group of L-Lys of UDP-*N*-acetyl-muramyl-pentapeptide (*39*). This peptidyl transferase could be involved in the formation of an additional amide bond as demonstrated before for other enzymes with structural homology to FemX proteins in the biosynthesis of pacidamycin (*40*) and as suggested for dehydrophos (*22*) (Figure A.12). Upstream of this gene, a homolog to an Asp-tRNA/Glu-tRNA amidotransferase was identified, further supporting the hypothesis of the incorporation of L-Gln, L-Asn or an analog into the final metabolite (Figure A.12). Additionally, a gene encoding an N_5-glutamine SAM-dependent methyltransferase was also found, suggesting the incorporation of a methyl group into the molecule (Figure A.12). Thus, the putative phosphonate biosynthetic cluster from *Streptomyces* sp. WM6372 encodes for significantly diverse proteins that potentially perform several steps involving C-N bond formation. Based on the analysis above, some biosynthetic steps can be speculated (Figure A.12). Although experimental evidence is lacking, the bioinformatic analysis performed on the gene cluster and the proposed pathway may help in the purification and structure elucidation of the final metabolite. Interestingly, two genes in the cluster are predicted to encode secretory and non-secretory metallophosphoesterases. These proteins could be involved in self-protection/reactivation mediated by phosphorylation as previously reported for

fosfomycin (Figure A.13) (*41*). The formation of a phosphonate phosphate is consistent with the observation of a doublet of doublets in ^{31}P NMR spectra of culture supernatants of this strain as described below.

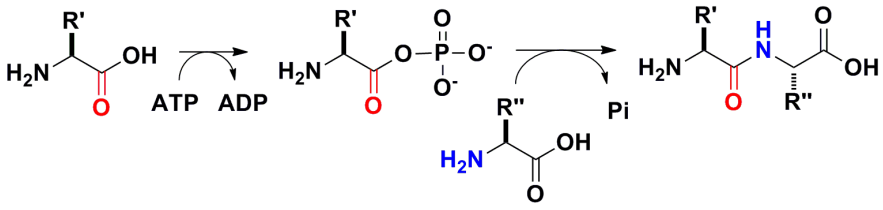


Figure A.11. Formation of a dipeptide catalyzed by an ATP-grasp ligase domain with the generation of an acylphosphate intermediate.

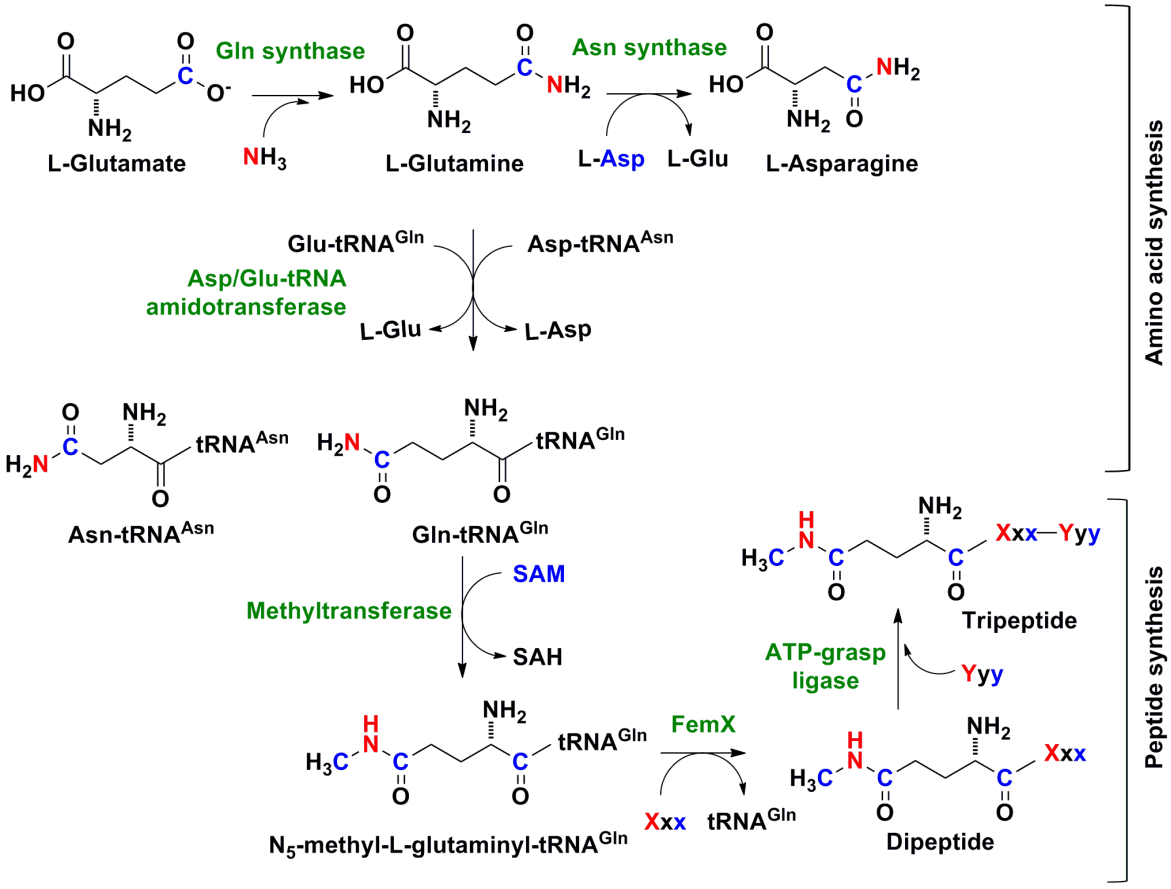


Figure A.12. A potential biosynthetic pathway for the production of a tripeptide phosphonate by *Streptomyces* sp. WM6372. Several proteins catalyzing the formation of C-N bonds are encoded by the putative gene cluster.

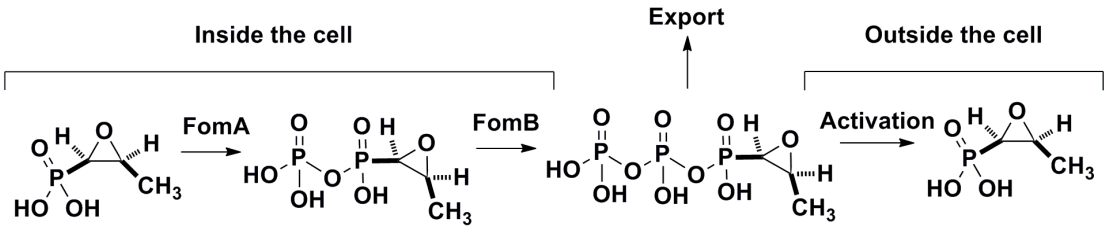


Figure A.13. Phosphorylation of the phosphonate group in fosfomycin as a self-protection mechanism by FomA and FomB in *S. wedmorensis*.

The gene cluster from *S. verne* WM4235 also contains genes encoding for Ppm and Ppd, as expected. Two genes encoding the adenylation and the condensation domains of a non-ribosomal peptide synthetase (NRPS) were also identified, suggesting the formation of a peptide phosphonate. Alignment of nine of the ten residues predicted to define the specificity of the adenylation domain (*42*) suggests the incorporation of L-Thr or Dhb (2,3-dehydroaminobutyric acid, Dht, or Dhbu) into the peptide (Figure A.14). The downstream gene designated as synthetase/ligase encodes a protein with an AMP-binding site domain that could be the adenylation domain for activation of a phosphonate intermediate before incorporation into the peptide. Interestingly, an alcohol dehydrogenase is predicted to be a zinc-dependent oxidoreductase that contains a threonine dehydrogenase domain, suggesting that Thr or an analog might be incorporated into the peptide by the NRPS and subsequently modified by the dehydrogenase.

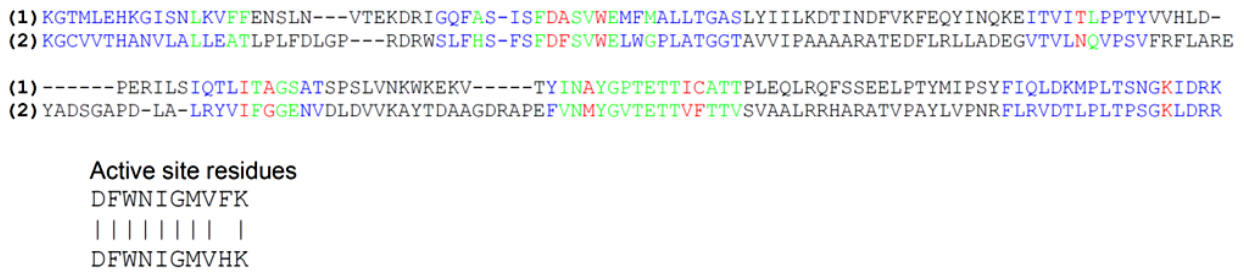


Figure A.14. Amino acid sequence alignment of PheA (1) and the adenylation domain of WM4235 (2). Shown in red are the 10 amino acids defined to determine the specificity of the adenylation domain (*42*). Nine of the ten residues in the active site established to recognize for Thr or Dhb in other NRPS proteins are conserved.

Screening of conditions for production of novel phosphonates

The genomic approach described above allows for the identification of strains that carry the genetic information for the biosynthesis of potentially novel phosphonates. However, to discover new natural products, the biosynthetic pathways have to be activated and the metabolites identified. In *Streptomyces*, the production of antibiotics is often temporally correlated with other changes in colony morphology, such as spore formation and development of aerial mycelia (*43*). A variety of conditions, including nutrient limitation, can induce *Streptomyces* to produce secondary metabolites and to sporulate in order to compete with other organisms and guarantee preservation under hostile conditions. However, the networks of regulatory signals that control these changes are quite complex, highly variable among strains and pathways and only partially understood (*43*). For this reason, the isolation of novel phosphonates from native host cultures can be quite challenging and requires testing many different conditions.

To induce the production of phosphonates by *Streptomyces* sp. WM6372 and *S. verne* WM4235, a broad set of conditions (including solid and liquid, rich and minimal

media) were tested. In the case of solid media, the agar plates were incubated from 7 to 15 days at which time the formation of spores and/or the production of pigments were normally observed. The agar plates were frozen and then thawed before the liquid phase was removed and concentrated tenfold. For liquid media, the cultures were grown from 2 to 5 days and aliquots of cell-free supernatant were concentrated 4- to 30-fold for further analysis. Initially, the samples were analyzed for bioactivity in conventional agar diffusion bioactivity assays using a variety of indicator strains, including Gram-positive and Gram-negative bacteria and an *E. coli* strain that has the three component phosphonate transporter *phnCDE* integrated into the chromosome under an IPTG inducible promoter (*21*). For both *Streptomyces* species, the production of at least two different antibiotics was observed (Figure A.15).

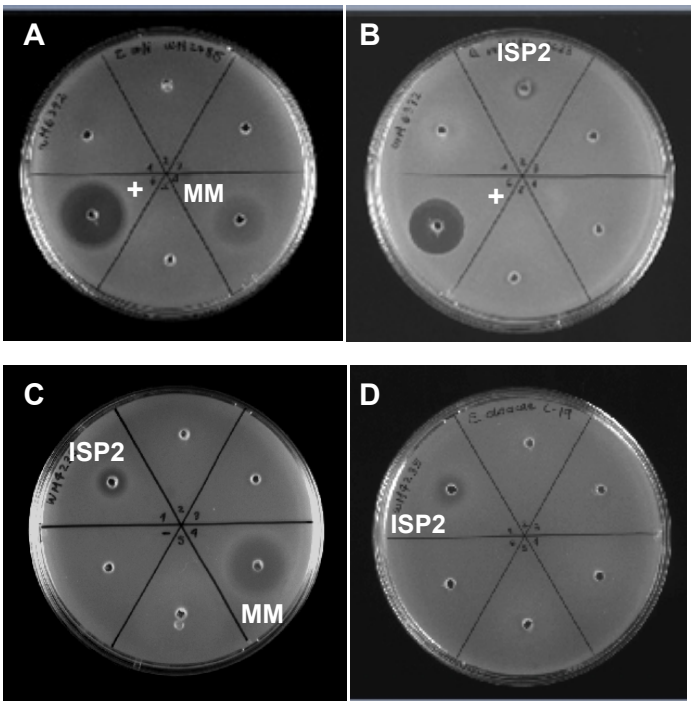


Figure A.15. Agar diffusion bioactivity assay of samples of cell-free extracts after tenfold concentration. Top: Extracts from *Streptomyces* sp. WM6372 tested on LB agar plates with *E. coli* WM2785 (A) and *B. subtilis* ATCC6633 (B) as indicator strains. Extracts from minimal medium (MM) and from ISP2 were bioactive, ampicillin (+) was used as a positive control. Bottom: Extracts from *S. verne* WM4235 tested on LB agar plates with *E. coli* WM2785 (C) and *Enterobacter cloacae* C19 (D) as indicator strains. Extracts from minimal medium (MM) and from ISP2 were bioactive.

Phosphorus-31 nuclear magnetic resonance spectroscopy (^{31}P NMR) is an effective technique to determine the presence of phosphonates in cell-free extracts. In ^{31}P NMR spectra, the signals due to phosphate esters and orthophosphate are normally observed at chemical shifts below 5 ppm, while those of the C-P compounds are observed between 6 and 45 ppm. The cell-free extracts of cultures of the strains WM6372 and WM4235 were analyzed by ^{31}P NMR spectroscopy and various peaks, likely corresponding to phosphonates, were detected. In the case of the strain WM6372 (Figure A.16), the ^{1}H-decoupled ^{31}P NMR spectrum of an ISP2 extract presents a signal

at 10.6 ppm and a smaller peak at 11.2 ppm. Additional phosphorus-containing compounds, most of them not present in the media, were observed, including a compound that generates a doublet of doublets at −11.5 ppm (dd, J = 387, 19.2 Hz). This compound may be a phosphonate phosphate or a diphosphate produced by one of the phosphodiesterases present in the cluster (see above). The ^{1}H-coupled ^{31}P NMR spectrum of the same sample was also recorded, and a doublet at 10.6 ppm (d, J = 17 Hz) was observed, suggesting that only one proton is bound to the α-carbon (Figure A.16). After enzymatic treatment of the sample with a phosphodiesterase (which can break down the phosphodiester bond in 2',3'-cyclophosphate nucleoside derivatives that produce a ^{31}P NMR peak in the range of phosphonates (*44*)) and with an alkaline phosphatase to hydrolyze phosphate-containing molecules, the characteristic peak at 10.6 ppm was still observed in the ^{31}P NMR spectrum. Furthermore, the peak survived concentrated hydrochloric acid treatment confirming the expected chemical stability of the C-P bond. Thus, this phosphorus-containing compound is most likely a phosphonate. Similarly, for the strain WM4235 grown in ISP4 for 15 days the ^{1}H-decoupled and coupled ^{31}P NMR spectra were recorded and a multiplet around 7.8 ppm was observed (Figure A.17). This compound was not detected when the strain was grown for shorter incubation periods or in different media and its production may be correlated with spore formation.

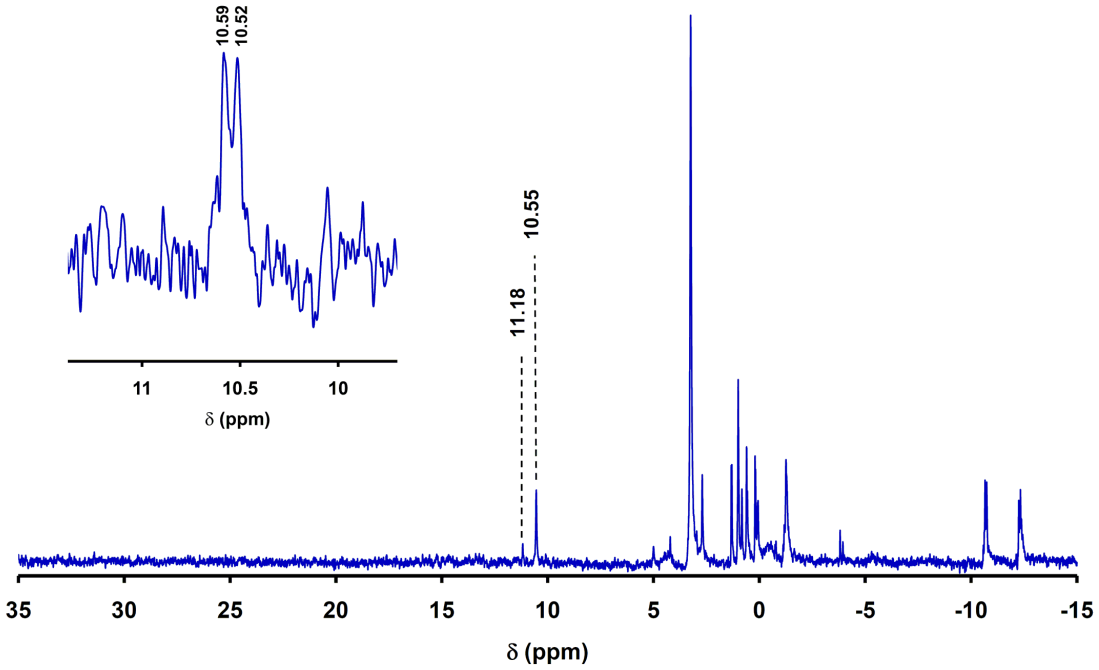


Figure A.16. ^{31}P NMR spectra of concentrated crude extract from strain WM6372 grown on ISP2 agar plates. The ^{1}H-decoupled and ^{1}H-coupled (insert) spectra are shown. The signals at 10.5 ppm at 11.2 ppm are presumably generated by phosphonates.

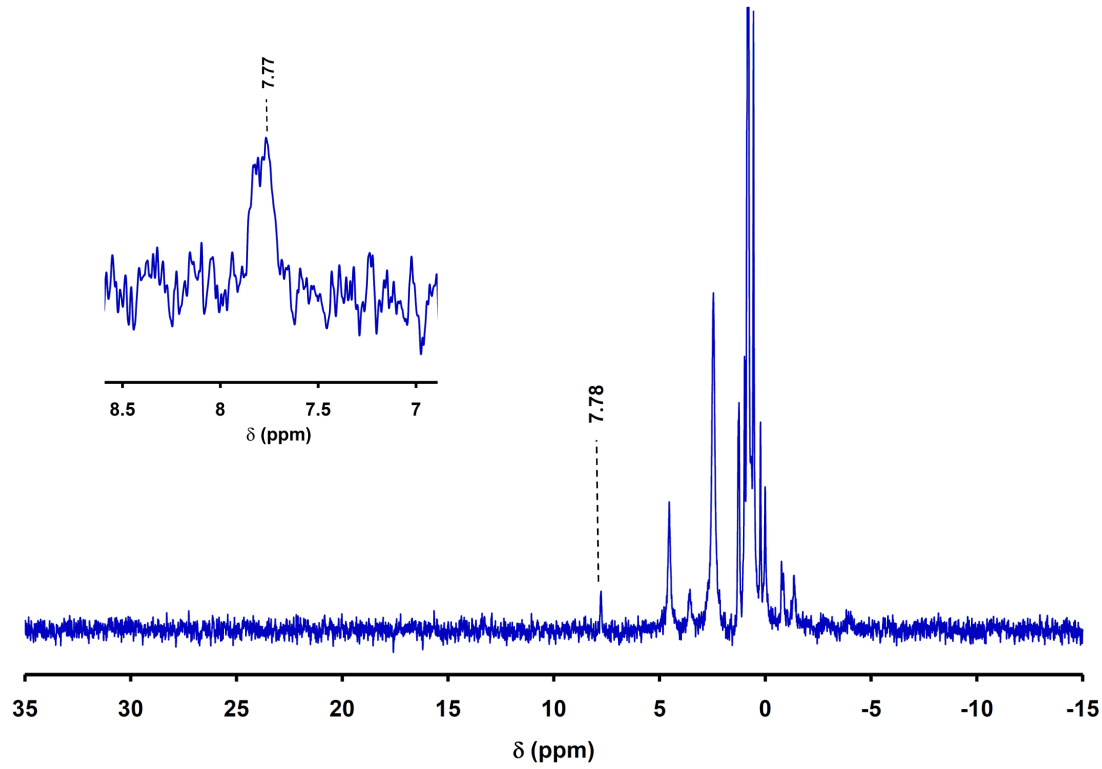


Figure A.17. ^{31}P NMR spectra of concentrated crude extract from strain WM4235 grown on ISP4 agar plates. The ^{1}H-decoupled and ^{1}H-coupled (insert) spectra are shown. The signal at 7.8 ppm is presumably generated by a phosphonate.

Production of phosphonates in the presence of phosphonoalanine

As mentioned above, the identification of suitable conditions for phosphonate production can be quite challenging. To facilitate the discovery of novel metabolites from *ppm*-positive strains, the development of a medium that chemically induces the production of phosphonates is highly desirable. Since the target strains have the genetic potential to convert reversibly PnPy into PEP, a media containing PnPy as the only source of phosphorus or carbon may force the expression of the Ppm-encoding gene operon or cluster, allowing the host to access a source of PEP and other essential phosphates (Figure A.18). The activation of the pathway may simultaneously lead to the production of the phosphonate metabolite or its intermediates. Since PnPy is not commercially available, D/L-PnAla can be alternatively used, provided non-specific aminotransferases in the host catalyze the conversion between them (Figure A.18). Indeed, the Ppm activity, which favors the reverse reaction, was successfully used to isolate several *ppm*-positive microorganisms after growing them in an enrichment media containing D/L-PnAla as a sole phosphorus source (*45, 46*).

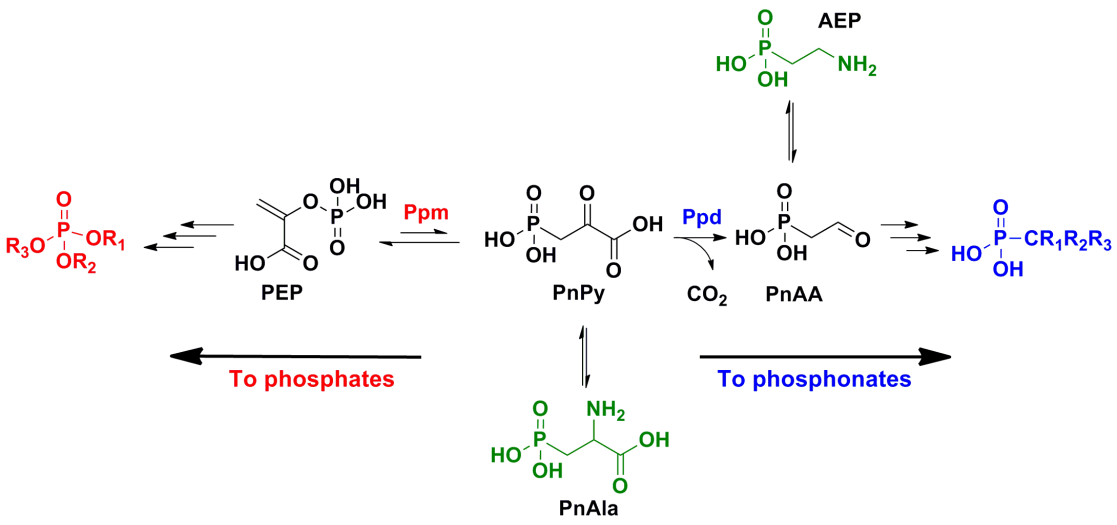


Figure A.18. Conversion of PnAla and AEP into phosphates and phosphonates. The expression of PEP mutase (Ppm) and PnPy decarboxylase (Ppd) encoding genes may allow the conversion of PnAla and AEP (green) into phosphates (red) and the expected phosphonates (blue).

When the strain WM6372 was grown in modified ISP4 medium, without agar and with D/L-PnAla as the sole source of phosphorus, two signals at 10.5 ppm and 10.0 ppm were observed by ^{1}P NMR spectroscopy (Figure A.19A). A spiking experiment indicated that these signals overlap with the peaks generated by the phosphonates produced in ISP2 medium (Figure A.19B). One additional peak at 18.5 ppm, which may correspond to a biosynthetic intermediate or a degradation product, was also detected. Furthermore, no signals in the phosphonate range were observed by ^{1}P NMR spectroscopy when K_2HPO_4 was the source of phosphorus (Figure A.19C), confirming that the production of the phosphonates depends on the presence of D/L-PnAla. As expected, when both, K_2HPO_4 or PnAla, were absent in the media, no significant bacterial growth was detected.

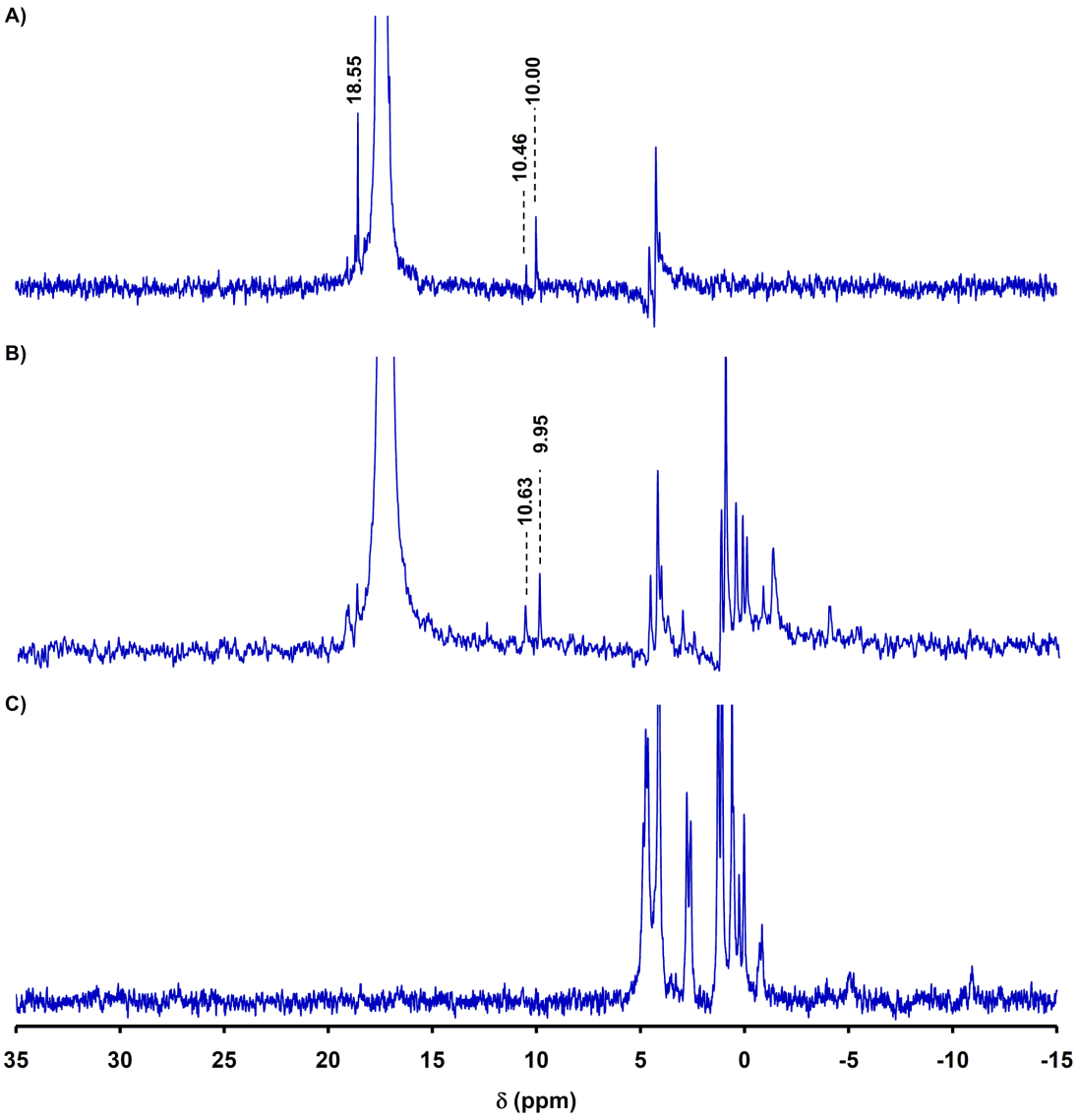


Figure A.19. ^{1}H-decoupled ^{31}P NMR spectra of concentrated crude extract from strain WM6372 grown on ISP4 medium without agar and with PnAla as the sole source of phosphorus (A). After spiking with a sample produced in ISP2 medium, only the signals at 10.5 ppm and 10.0 ppm were observed (B). No signals for phosphonates were detected when PnAla was substituted for K_2HPO_4 in the medium (C). The signal at 17.4 ppm corresponds to D/L-PnAla, while the signal at 18.5 ppm may correspond to a biosynthetic intermediate or a degradation product.

Although the previous results are promising, PnAla could be preferentially utilized for the production of PEP and other phosphates that are essential for bacterial growth and the yield of phosphonates may be considerably low. However, an open reading frame encoding for Ppd, which catalyzes the conversion of PnPy into PnAA, is present

in several of the phosphonate gene clusters, including the strains WM6372 and WM4235. Thus, growing the bacteria in a medium containing PnAA (or its equivalent AEP), in addition to PnAla, as the sources of phosphorus, may further facilitate the production of the desired phosphonates (Figure A.18). Indeed, when a minimal media containing glucose or starch as the main carbon sources, NH_4Cl as nitrogen source, and PnAla with AEP as sources of phosphorus was utilized to grow the strain WM6372, significant production of a phosphonate with a chemical shift of 10.5 ppm was observed (Figure A.20). Furthermore, due to the higher concentration of phosphonate and the relatively simplicity of the sample compared with the sample generated from ISP2 medium, it was possible to record a ^{1}H-^{31}P gHMBC NMR spectrum using a tenfold concentrated sample without any additional treatment. The spectrum revealed only one proton coupled to the phosphorus generating a doublet with a chemical shift of 2.7 ppm (d, *J* = 12.1 Hz), although a more complex splitting pattern cannot be discarded at this point (Figure A.21). From the previously discovered phosphonate natural products (*5*), fosfomycin shows NMR chemical shifts around 11.9 ppm for the phosphorus (*4*) and 2.7 ppm (dd, *J* = 17-21, 5 Hz) for the α-proton (*47*), close to the recorded values for the main phosphonate produced by WM6372. However, spiking experiments performed by Dr. Jiangtao Gao, a current postdoctoral researcher in the van der Donk laboratory, confirmed that both compounds are not identical, although similar chemical moieties are possible. Alternatively, the phosphonate SF-2312 that contains an oxopyrrolidine ring was reported to show NMR chemical shifts at 13 to 18 ppm for the phosphorus and 2.9 ppm to 3.1 ppm (ddd, *J* = 21, 10, 6 Hz) for the α-proton (*48, 49*). Interestingly, the gene cluster for WM6372 encodes for a putative 3-carboxy-*cis,cis*-muconate cycloisomerase,

an enzyme that catalyzes the lactonization of *cis,cis*-muconate. Thus, the phosphonate moiety may be bound to a lactone ring consistent with the NMR spectroscopy data.

The designed medium containing PnAla and AEP as sources of phosphorus can be utilized to obtain relatively large quantities of phosphonates from strain WM6372 (and maybe others) using small-scale cultures, facilitating the production of isotopically labeled compounds and the elucidation of the chemical structures. Importantly, phosphonates, including PnAla and AEP, can be utilized as source of phosphorus by a variety of microorganisms via several pathways, including the C-P lyase, phosphonatase, phosphonopyruvate hydrolase, and the recently characterized phosphonoacetaldehyde dehydrogenase-dependent mechanisms (*7, 50*). Thus, the production of phosphonates under the stimulus of PnAla and AEP may rely on the absence of alternative degradation pathways in the producer organisms. Further optimization of the media composition, including studies with PnPy or enantiomerically pure L-PnAla and different strains, are required to develop a general phosphonate-inducing media that could be utilized to discover novel metabolites.

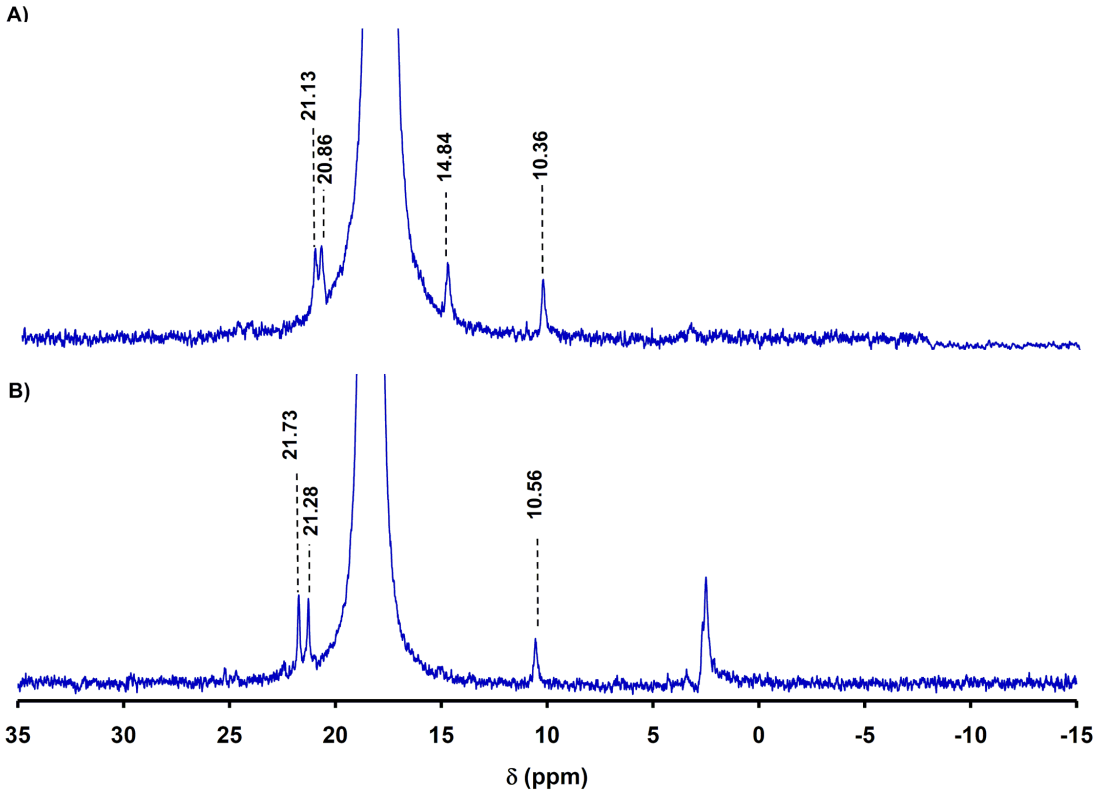


Figure A.20. ^{1}H-decoupled ^{31}P NMR spectra of concentrated crude extract from strain WM6372 grown on a minimal media containing glucose (A) or starch (B) as the main carbon sources, NH_4Cl as nitrogen source, and PnAla with AEP as sources of phosphorus. No phosphonate signals were observed when K_2HPO_4 was used instead of PnAla and AEP. The signals at 18.4 ppm and 18.1 ppm correspond to D/L-PnAla and AEP, while the signals at 21.1 ppm, 20.9 ppm, and 14.8 ppm may correspond to biosynthetic intermediates or degradation products.

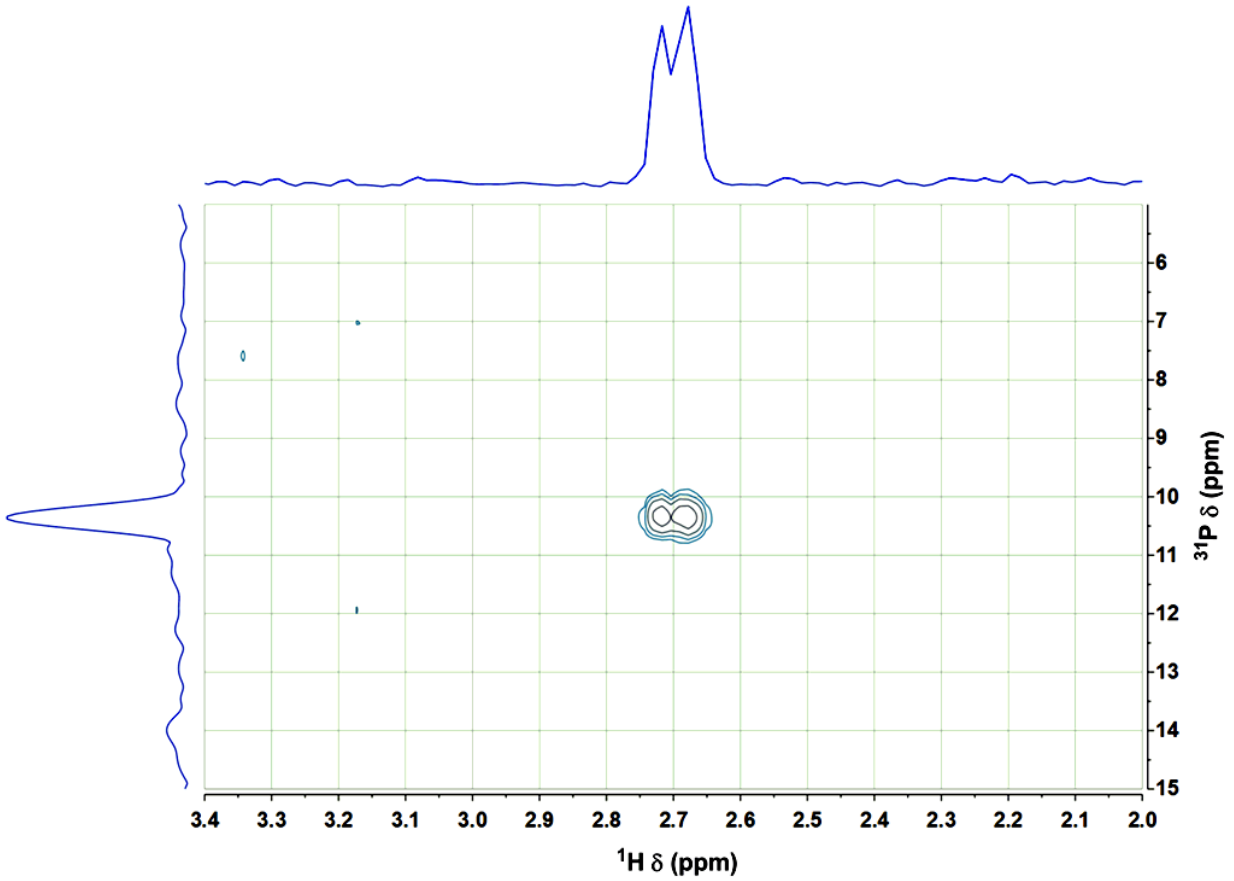


Figure A.21. ^{1}H-^{31}P gHMBC spectrum of a concentrated crude extract from strain WM6372 grown on a minimal media containing glucose, NH_4Cl, PnAla, and AEP. The phosphorus is coupled to a single proton in the α-carbon and with a chemical shift of 2.7 ppm.

In summary, natural phosphonates are a relatively unexplored group of metabolites that include several biologically active compounds. The screening for *ppm*-positive strains by PCR and the construction of genomic DNA libraries are promising tools for the identification of organisms that produce phosphonate antibiotics. Utilizing this methodology, several phosphonate gene clusters were sequenced and characterized. Furthermore, the production conditions for two novel potential phosphonates were determined. Although, the ISP2 medium extract of WM6372 was shown to inhibit the growth of *B. subtilis* ATCC6633, whether the inhibition zone is produced by a phosphonate or not is not certain. Once purified, the chemical structures

and the biological activity can be determined. Currently, it is not possible to assure that the sequenced gene cluster encodes the biosynthesis of the detected phosphonate. Trials for heterologous production of the phosphonate by *S. lividans* may provide the answer to this question and may also allow the identification of the minimal gene cluster as previously performed for fosfomycin (*4*), phosphinothricin tripeptide (*19*), FR-900098 (*21*), and dehydrophos (*22*). Deletion of genes, feeding experiments with identified intermediates, *in vivo* complementation experiments, and *in vitro* reconstitution of enzymes in combination with ^{31}P NMR spectroscopy and mass spectrometry may allow the elucidation of the biosynthetic route.

A.4. REFERENCES

1. Brundtland, G., World Health Organization Report on Infectious Diseases 2000: Overcoming Antimicrobiol Resistance. http://www.who.int/infectious-disease-report/2000/index.html. Accessed on Sept 30, 2008.

2. Coates, A., Hu, Y. M., Bax, R., and Page, C. (2002) The future challenges facing the development of new antimicrobial drugs, *Nat. Rev. Drug Discov. 1*, 895-910.

3. Baltz, R. H. (2006) Marcel Faber Roundtable: is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration?, *J. Ind. Microbiol. Biotechnol. 33*, 507-513.

4. Woodyer, R. D., Shao, Z. Y., Thomas, P. M., Kelleher, N. L., Blodgett, J. A. V., Metcalf, W. W., Van der Donk, W. A., and Zhao, H. M. (2006) Heterologous production of fosfomycin and identification of the minimal biosynthetic gene cluster, *Chem. Biol. 13*, 1171-1182.

5. Metcalf, W. W., and van der Donk, W. A. (2009) Biosynthesis of phosphonic and phosphinic acid natural products, *Annu. Rev. Biochem. 78*, 65-94.

6. Hilderbrand, R. L. (1983) *The role of phosphonates in living systems*, CRC Press, Boca Raton, Fla.

7. White, A. K., and Metcalf, W. W. (2007) Microbial metabolism of reduced phosphorus compounds, *Annu. Rev. Microbiol. 61*, 379-400.

8. Okuhara, M., Kuroda, Y., Goto, T., Okamoto, M., Terano, H., Kohsaka, M., Aoki, H., and Imanaka, H. (1980) Studies on new phosphonic acid antibiotics. I. FR-900098, isolation and characterization, *J. Antibiot. 33*, 13-17.

9. Okuhara, M., Kuroda, Y., Goto, T., Okamoto, M., Terano, H., Kohsaka, M., Aoki, H., and Imanaka, H. (1980) Studies on new phosphonic acid antibiotics .3. Isolation and characterization of FR-31564, FR-32863 and FR-33289, *J. Antibiot. 33*, 24-28.

10. Takahashi, E., Kimura, T., Nakamura, K., Arahira, M., and Iida, M. (1995) Phosphonothrixin, a novel herbicidal antibiotic produced by *Saccharothrix* Sp. St-888 .1. Taxonomy, fermentation, isolation, and biological properties, *J. Antibiot. 48*, 1124-1129.

11. Bayer, E., Zahner, H., Konig, W. A., Jessipow, S., Gugel, K. H., Hagele, K., and Hagenmai.H. (1972) Metabolites of microorganisms 98. Phosphinothricin and phosphinothricyl-alanylalanine, *Helv. Chim. Acta 55*, 224-&.

12. Park, B. K., Hirota, A., and Sakai, H. (1977) Studies on new antimetabolite produced by microorganism .3. Structure of plumbemycin-A and plumbemycin-B, antagonists of L-Threonine from *Streptomyces plumbeus*, *Agr. Biol. Chem. 41*, 573-579.

13. Ogita, T., Gunji, S., Fukazawa, Y., Terahara, A., Kinoshita, T., Nagaki, H., and Beppu, T. (1983) The structures of fosfazinomycin-A and fosfazinomycin-B, *Tetrahedron Lett. 24*, 2283-2286.

14. Watabe, H., JunkoYoshida, Tanaka, E., Ito, M., Miyadoh, S., and Shomura, T. (1986) Studies on a new phosphonic acid antibiotic, SF-2312. I. Screening method, taxonomy and biological activity., *Meiji Seika Kenkyu Nenpo 25*, 12.

15. Hunt, A. H., and Elzey, T. K. (1988) Revised structure of A53868a, *J. Antibiot. 41*, 802-802.

16. Koguchi, T., Yamada, K., Yamato, M., Okachi, R., Nakayama, K., and Kase, H. (1986) K-4, a novel Inhibitor of angiotensin-I converting enzyme produced by *Actinomadura spiculosospora*, *J. Antibiot. 39*, 364-371.

17. Yamato, M., Koguchi, T., Okachi, R., Yamada, K., Nakayama, K., Kase, H., Karasawa, A., and Shuto, K. (1986) K-26, a novel inhibitor of angiotensin-I converting enzyme produced by an *Actinomycete* K-26, *J. Antibiot. 39*, 44-52.

18. Barry, R. J., Bowman, E., McQueney, M., and Dunaway-Mariano, D. (1988) Elucidation of the 2-aminoethylphosphonate biosynthetic pathway in *Tetrahymena pyriformis*, *Biochem. Biophys. Res. Commun. 153*, 177-182.

19. Blodgett, J. A., Zhang, J. K., and Metcalf, W. W. (2005) Molecular cloning, sequence analysis, and heterologous expression of the phosphinothricin tripeptide biosynthetic gene cluster from *Streptomyces viridochromogenes* DSM 40736, *Antimicrob. Agents Chemother. 49*, 230-240.

20. Blodgett, J. A. V., Thomas, P. M., Li, G., Velásquez, J. E., van der Donk, W. A., Kelleher, N. L., and Metcalf, W. W. (2007) Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide, *Nat. Chem. Biol. 3*, 480-485.

21. Eliot, A. C., Griffin, B. M., Thomas, P. M., Johannes, T. W., Kelleher, N. L., Zhao, H., and Metcalf, W. W. (2008) Cloning, expression, and biochemical characterization of *Streptomyces rubellomurinus* genes required for biosynthesis of antimalarial compound FR900098, *Chem. Biol. 15*, 765-770.

22. Circello, B. T., Eliot, A. C., Lee, J. H., van der Donk, W. A., and Metcalf, W. W. (2010) Molecular cloning and heterologous expression of the dehydrophos biosynthetic gene cluster, *Chem. Biol. 17*, 402-411.

23. Ntai, I., Phelan, V. V., and Bachmann, B. O. (2006) Phosphonopeptide K-26 biosynthetic intermediates in *Astrosporangium hypotensionis*, *Chem. Commun.*, 4518-4520.

24. Letunic, I., Copley, R. R., Pils, B., Pinkert, S., Schultz, J., and Bork, P. (2006) SMART 5: domains in the context of genomes and networks, *Nucleic Acids Res. 34*, D257-260.

25. Marchler-Bauer, A., Anderson, J. B., Derbyshire, M. K., DeWeese-Scott, C., Gonzales, N. R., Gwadz, M., Hao, L., He, S., Hurwitz, D. I., Jackson, J. D., Ke, Z., Krylov, D., Lanczycki, C. J., Liebert, C. A., Liu, C., Lu, F., Lu, S., Marchler, G. H., Mullokandov, M., Song, J. S., Thanki, N., Yamashita, R. A., Yin, J. J., Zhang, D., and Bryant, S. H. (2007) CDD: a conserved domain database for interactive domain family analysis, *Nucleic Acids Res. 35*, D237-240.

26. Hulo, N., Sigrist, C. J., Le Saux, V., Langendijk-Genevaux, P. S., Bordoli, L., Gattiker, A., De Castro, E., Bucher, P., and Bairoch, A. (2004) Recent improvements to the PROSITE database, *Nucleic Acids Res. 32*, D134-137.

27. Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J., and Kelley, L. A. (2008) Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre, *Proteins 70*, 611-625.

28. Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W., and Huson, D. H. (2005) Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs), *Nucleic Acids Res. 33*, 5799-5808.

29. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, *J. Mol. Biol. 305*, 567-580.

30. Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools, *Nat. Protoc. 2*, 953-971.

31. Hobbs, G., Frazer, C. M., Gardner, D. C. J., Cullum, J. A., and Oliver, S. G. (1989) Dispersed growth of *Streptomyces* in liquid culture, *Appl. Microbiol. Biotechnol. 31*, 272-277.

32. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000) *Practical Streptomyces genetics*, John Innes Foundation, Norwich (England).

33. Hopwood, D. A. (1967) Genetic analysis and genome structure in *Streptomyces coelicolor*, *Bacteriol. Rev. 31*, 373-403.

34. Hodgson, D. A. (1982) Glucose repression of carbon source uptake and metabolism in *Streptomyces coelicolor* A3(2) and its perturbation in mutants resistant to 2-deoxyglucose, *J. Gen. Microbiol. 128*, 2417-2430.

35. Kino, K., Kotanaka, Y., Arai, T., and Yagasaki, M. (2009) A novel L-amino acid ligase from *Bacillus subtilis* NBRC3134, a microorganism producing peptide-antibiotic rhizocticin, *Biosci. Biotechnol. Biochem. 73*, 901-907.

36. Borisova, S. A., Circello, B. T., Zhang, J. K., van der Donk, W. A., and Metcalf, W. W. (2010) Biosynthesis of rhizocticins, antifungal phosphonate oligopeptides produced by *Bacillus subtilis* ATCC6633, *Chem. Biol. 17*, 28-37.

37. Galperin, M. Y., and Koonin, E. V. (1997) A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity, *Protein Sci. 6*, 2639-2643.

38. Noda, M., Kawahara, Y., Ichikawa, A., Matoba, Y., Matsuo, H., Lee, D. G., Kumagai, T., and Sugiyama, M. (2004) Self-protection mechanism in D-cycloserine-producing *Streptomyces lavendulae*. Gene cloning, characterization, and kinetics of its alanine racemase and D-alanyl-D-alanine ligase, which are target enzymes of D-cycloserine, *J. Biol. Chem. 279*, 46143-46152.

39. Maillard, A. P., Biarrotte-Sorin, S., Villet, R., Mesnage, S., Bouhss, A., Sougakoff, W., Mayer, C., and Arthur, M. (2005) Structure-based site-directed mutagenesis of the UDP-MurNAc-pentapeptide-binding cavity of the FemX alanyl transferase from *Weissella viridescens*, *J. Bacteriol. 187*, 3833-3838.

40. Zhang, W., Ntai, I., Kelleher, N. L., and Walsh, C. T. (2011) tRNA-dependent peptide bond formation by the transferase PacB in biosynthesis of the pacidamycin group of pentapeptidyl nucleoside antibiotics, *Proc. Natl. Acad. Sci. U. S. A. 108*, 12249-12253.

41. Kobayashi, S., Kuzuyama, T., and Seto, H. (2000) Characterization of the *fomA* and *fomB* gene products from *Streptomyces wedmorensis*, which confer fosfomycin resistance on *Escherichia coli*, *Antimicrob. Agents Chemother. 44*, 647-650.

42. Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases, *Chem. Biol. 6*, 493-505.

43. Walsh, C. (2003) *Antibiotics : actions, origins, resistance*, ASM Press, Washington, D.C.

44. Lebedev, A. V., and Rezvukhin, A. I. (1984) Tendencies of 31P chemical shifts changes in NMR spectra of nucleotide derivatives, *Nucleic Acids Res. 12*, 5547-5566.

45. Ternan, N. G., McGrath, J. W., and Quinn, J. P. (1998) Phosphoenolpyruvate phosphomutase activity in an L-phosphonoalanine-mineralizing strain of *Burkholderia cepacia*, *Appl. Environ. Microbiol. 64*, 2291-2294.

46. Nakashita, H., Shimazu, A., and Seto, H. (1991) New screening method for C-P compound producing organisms by the use of phosphoenolpyruvate phosphomutase, *Agric. Biol. Chem. 55*, 2825-2829.

47. Voncarstennlichterfelde, C., Fernandezibanez, M., Galvezruano, E., and Bellanato, J. (1983) Structural study of fosfomycin [(-)-*cis*-1,2-epoxypropylphosphonic acid] salts and related-compounds, *J. Chem. Soc. Perkin Trans. 2*, 943-947.

48. Hanaya, T., and Itoh, C. (2011) An efficient synthesis of antibiotic SF-2312 (3-dihydroxyphosphoryl-1,5-dihydroxy-2-pyrrolidone), *Heterocycles 82*, 1675-1683.

49. Ohba, K., Sato, Y., Sasaki, T., and Sezaki, M. (1986) Studies on a new phosphonic acid antibiotic, SF-2312. II. Isolation, physico-chemical properties and structure, *Sci. Rep. Meiji Seika Kaisha 25*, 18-22.

50. Borisova, S. A., Christman, H. D., Metcalf, M. E., Zulkepli, N. A., Zhang, J. K., van der Donk, W. A., and Metcalf, W. W. (2011) Genetic and biochemical characterization of a pathway for the degradation of 2-aminoethylphosphonate in Sinorhizobium meliloti 1021, *J. Biol. Chem. 286*, 22283-22290.