Nonribosomal peptide synthesis in Bacillus subtilis
Duitman, Erwin Hans

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2003

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter one

Nonribosomal peptide synthesis in *Bacillus*

1.1 General

Many microorganisms, both prokaryotes and eukaryotes, produce a variety of biological active compounds. Many of these compounds are of peptide origin and can be synthesized ribosomally as well as nonribosomally. In recent years a growing interest has developed for these peptides because many of them exhibit potentially medically relevant properties. Due to this growing interest, the number of biological active peptides as well as the knowledge about their biosynthesis has increased tremendously during recent years. Numerous microbiological species produce these biologically active peptides and a steadily growing number of these biologically active peptides has been identified and characterized.

The emphasis in this overview will be on nonribosomally synthesized peptides (NSP’s) produced by a number of *Bacillus* species, as most studies on the biosynthesis of these peptides have been performed in *Bacillus* species, especially in *B. brevis*, *B. licheniformis* and *B. subtilis*, due to the relative ease to work with these species. This overview briefly discusses the structural features and secretion mechanism of these NSP’s in these Gram-positive bacteria. This mechanism is often involved in self-resistance of the producer against these predominantly antibiotic peptides. However, the main focus of this overview concerns the structure of the multienzyme complexes responsible for their biosynthesis, the way in which the encoding operons are organized and the regulation of these operons.

1.2 Peptide structure and activities

Although there seems to be a size limitation, probably due to way in which they are synthesized, the structural variation of NSP’s is enormous. Structural determination of numerous NSP’s revealed that, in contrast to ribosomally synthesized peptides, which contain only the 21 proteinogenic amino acids, they are assembled from more than 300 different precursors including pseudo, non-proteinogenic, hydroxy, N-methylated and D-amino acids. In addition, the peptide backbone of NSP’s can be composed of linear, cyclic, or cyclic branched structures, which can be further modified by acylation, glycosylation, or heterocyclic ring formation. NSP’s are often a mixture of several, slightly different, peptides due to lack of specificity of many adenylation domains, which are described below in more detail. The 3-dimensional structure from several NSP’s has been solved using NMR, distance geometry, and molecular dynamics, and the observed structural diversity of NSP’s is reflected in a broad spectrum of biological properties (Bonmatin et al., 1994; Peypoux et al., 1986). Although for most of these peptides the precise function remains unclear they probably play an important role in the struggle for nutrients either by inhibiting competing organisms or by facilitating the uptake of metal ions essential for growth. Often NSP’s are surface active and many exhibit antibiotic, antifungal, antitumor or antiviral activity. In addition, NSP’s can exhibit enzyme-inhibiting or immunosuppressive activity. Due to these interesting properties, in combination with
Chapter 2

**Bacillus brevis**

<table>
<thead>
<tr>
<th>gramicidin S</th>
<th>tyrocidine</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Phe - L-Pro - L-Val - L-Orn - L-Leu</td>
<td>D-Phe - L-Pro - L-Phe - D-Phe - L-Asn</td>
</tr>
<tr>
<td>L-Leu - L-Orn - L-Val - L-Pro - D-Phe</td>
<td>L-Leu - L-Orn - L-Val - L-Tyr - L-Gln</td>
</tr>
</tbody>
</table>

**Bacillus licheniformis**

<table>
<thead>
<tr>
<th>bacitracin A</th>
<th>lichenysin A</th>
</tr>
</thead>
</table>
| L-Ile - L-Cys\_ - L-Leu - D-Glu - L-Ile - L-Lys - D-Orn - L-Ile - D-Phx \_
|                         | CH\_\_ - (CH\_\_)\_ - CH - CH\_\_CO - L-Glu - L-Leu - D-Leu \_
| thiazoline               | O - L-Ile - D-Leu - L-Asn - L-Val |
|                          | L-Asn - D-Asp - L-His \_ |

**Bacillus subtilis**

<table>
<thead>
<tr>
<th>iturin A</th>
<th>mycosubtilin</th>
</tr>
</thead>
</table>
| CH\_\_\_ - CH\_\_\_ - CH - CH\_\_CO - L-Asn - D-Tyr - D-Asn \_
|               | CH\_\_\_ - CH\_\_\_ - CH - CH\_\_CO - L-Asn - D-Tyr - D-Asn \_
| NH - L-Ser - D-Asn - L-Pro - L-Gln \_ | NH - L-Asn - D-Ser - L-Pro - L-Gln \_ |

<table>
<thead>
<tr>
<th>bacillomycin D</th>
<th>surfactin</th>
</tr>
</thead>
</table>
| CH\_\_\_\_ - CH - CH\_\_CO - L-Asn - D-Tyr - D-Asn \_
|               | CH\_\_\_\_ - CH\_\_\_ - CH - CH\_\_CO - L-Glu - L-Leu - D-Leu \_
| NH - L-Thr - D-Ser - L-Glu - L-Pro \_ | CH\_\_\_\_ O - L-Leu - D-Leu - L-Asp - L-Val \_ |

<table>
<thead>
<tr>
<th>fengycin</th>
</tr>
</thead>
</table>
| CH\_\_\_ - CH\_\_\_\_\_\_\_\_\_ - CH - CH\_\_CO - L-Glu - L-Orn - D-Tyr - D-allo-Thr - L-Glu - D-Ala \_
|               | L-Ile - D-Tyr - L-Gln - L-Pro \_ |

Figure 1.1: Schematic structure of a number of well-characterized NSP’s produced by some closely related Bacillus species.

their abundance, many of the NSP’s produced by B. brevis, B. licheniformis and B. subtilis, as well as their synthetases, have been studied and are quite well characterized (Fig. 1.1). This applies especially to bacitracin produced by B. licheniformis, gramicidin produced by B. brevis and surfactin produced by B. subtilis, which have been used extensively as model peptides for investigations concerning their mode of action. **B. brevis:** Both the cyclic peptides gramicidin S and tyrocidine, produced by B. brevis, consist of 10 amino acid residues (Izumiya et al., 1979). Gramicidin S consists of 2 identical pentapeptides, which are linked head to tail, and together form the stable amphiphilic cyclic decapeptide (Izumiya et al., 1979). The first amino acid residue of the 2 pentapeptides is in the D-configuration. The peptide exhibits strong antibiotic activities against a wide variety of Gram-positive and Gram-negative bacteria as well as against several pathogenic fungi (Konjelewski et al., 1996a; Konjelewski et al., 1996b). This cation-specific poreforming peptide also exhibits strong hemolytic activity, which together with the lack of specificity, limits the use of gramicidin S as an antibiotic for medical applications.
The second cyclic decapeptide produced by *B. brevis*, tyrocidine, contains 2 amino acid residues in the D-configuration, the first and fourth. Tyrocidine is actually a mixture of 4 slightly different decapeptides, which differ at the third, fourth and seventh amino acid residues (Ruttenberger and Mach, 1966). It is active against several Gram-positive bacteria and it has been suggested that this peptide plays a role in the regulation of sporulation of *B. brevis* (Lee and lipmann, 1975; Ristow et al., 1975).

**B. licheniformis:** The cyclic branched peptide bacitracin produced by *B. licheniformis* and *B. subtilis* consists of 12 amino acid residues and contains a thiazoline ring (Azevedo et al., 1993; Johnson et al., 1945). It contains 4 amino acid residues in the D-configuration, the fourth, seventh, ninth and eleventh. Bacitracin exhibits a strong antibiotic activity, primarily against Gram-positive bacteria by inhibiting cell wall biosynthesis. The primary mode of action of bacitracin is the formation of a tight ternary complex with C₅₅-isoprenyl pyrophosphate preventing recycling of this peptidoglycan carrier to C₅₅-isoprenyl phosphate (Stone and Strominger, 1971; Storm and Strominger, 1973). In addition, bacitracin also inhibits several proteases such as subtilisin (Pfeffer-Hennig, 1996). Bacitracin requires a divalent metal ion such as Zn(II) for its antibiotic activity, and can bind several other transition metal ions, including Co(II), Ni(II), and Cu(II) (Scogin et al., 1980; Stone and Strominger, 1971).

*B. licheniformis* also produces lichenysins, consisting of 7 amino acid residues of which the third and sixth are in the D-configuration, and a β-hydroxy fatty acid. These lipopeptides are structurally strongly related to surfactin produced by *B. subtilis*, and the amino acid composition of the different lichenysins varies mainly at the first, fifth and seventh amino acid residues. Several variants of the lichenysins were studied in more detail and all exhibit a strong surface tension-reducing activity and an antibiotic activity (Horowitz et al., 1990; Horowitz et al., 1991; Lin et al., 1994; Yakimov et al., 1995).

**B. subtilis:** Almost all of the NSP's produced by *B. subtilis* contain a fatty acid moiety, and of these lipopeptides surfactin has been studied most extensively. Surfactin consists of 7 amino acid residues of which the third and sixth are in the D-configuration, and a β-hydroxy fatty acid (Kakinuma et al., 1969a; Kakinuma et al., 1969b). It strongly reduces the surface tension, is hemolytic, and exhibits strong antiviral and limited antibiotic activities (Bernheimer and Avigad, 1970; Nissen et al., 1997; Vollenbroich et al., 1997). In addition, surfactin inhibits several enzymes such as phospholipase A2 and several metal-ion dependent cAMP-phosphodiesterases (Kim et al., 1998).

The iturin family of lipopeptides, consisting of bacilomycin, iturin and mycosubtilin are, like the lichenysins, closely related. However, in contrast to the members of the lichenysin family, the members of the iturin family have been given different names. All members of this lipopeptide family consist of 7 amino acid residues of which the second, third and sixth are always in the D-configuration, and contain an β-amino fatty acid modification. All members of this family exhibit strong antifungal, hemolytic, and limited antibacterial activities (Marget-Dana and Peypoux, 1994).

The largest lipopeptide produced by *B. subtilis* is fengycin, consisting of 10 amino acid residues and a β-hydroxy fatty acid. It contains 4 amino acid residues in the D-configuration, the second, fourth, sixth and ninth. It exhibits strong antifungal activity against filamentous fungi and inhibits several enzymes such as
phospholipase A2 and aromatase (Vanittanakom et al., 1986).

Finally, B. subtilis produces the dipeptide bacilysin, which exhibits strong antibiotic activity and consists of alanine and anticapsin, an amino acid derived from the aromatic amino acid intermediate prephenate (Walker and Abraham, 1970). Little is known about the synthesis of this dipeptide, and whether it is a real NSP is questionable, as no peptide synthetase seems to be involved in its synthesis.

1.3 Peptide synthetases

NSP's are synthesized by large multienzyme complexes, denoted peptide synthetases, which together with the polyketide synthetases, are the largest proteins present in the cell. Genetic and biochemical analyses of many peptide synthetases have revealed that these multifunctional proteins exhibit a modular structure in which a module is defined as the minimal, independent unit that catalyzes all the necessary reactions to modify and incorporate one specific amino acid into the growing peptide chain. The genetic arrangement of the genes encoding these modules is usually colinear with the amino acid sequence of the peptide product (colinearity rule).

The modules can be further subdivided into domains that perform the various reactions necessary to incorporate, and in some cases modify, the amino acid into the growing peptide chain and which can be recognized by a set of short conserved amino acid motifs (Fig. 1.2). Although all domains contain motifs that strongly resemble the core motifs at the expected positions, slight variations do occur. Because these amino acid motifs are
Figure 1.3: Schematic representation of the first steps in gramicidin S synthesis demonstrating the multiple carrier model of nonribosomal peptide synthesis. In the first step (I) the amino acids are activated by the adenylation domain (A) and subsequently bound as thioesters to the thiolation domain (T). In the second step (II) the epimerization domain (E) converts L-Phe to D-Phe. The third step (III) concerns the coupling of the first 2 amino acids resulting in the formation of the dipeptide D-Phe- L-Pro. Presumably in this stage another phenylalanine is activated by the first adenylation domain and bound as thioester to the first thiolation domain. The following steps (IV - V - - -) are continuous repetitions of similar reactions finally resulting in the synthesis of gramicidin S.

strongly conserved, it has been assumed that they play an important role in the activity or structural integrity of the domains, and many have been studied using site-directed mutagenesis.

Already thirty years ago, the multiple carrier model was proposed for the synthesis of NSP’s in which each module is loaded with a specific amino acid (Fig. 1.3, Lipmann et al., 1971). In this model a 2-step amino acid activation mechanism was proposed in which the first step consists of aminoacyl adenylate performed by the adenylation domain. The second step consists of aminoacyl thioesterification at a specific reactive thiol group of the thiolation domain. After activation, the amino acids are sometimes modified by epimerization- or N-methylation domains, and finally the activated amino acids are coupled via a condensation reaction, performed by the condensation domains.

Adenylation domain: The adenylation domain, consisting of about 550 amino acid residues, forms the sequence-determining core of each module, and has therefore been subject of extensive biochemical studies (Mootz and Marahiel, 1997a). This
domain recognizes and activates a specific amino acid and the order of these domain within the peptide synthetase dictates the primary structure of the synthesized peptide product (Lipmann et al., 1971; Lipmann, 1980; Stachelhaus and Marahiel, 1995a; Stachelhaus et al., 1996). The incorporation of an amino acid residue into the growing peptide chain requires a two-step mechanism for substrate activation (Carter, 1993; Delarue, 1995; Kleinkauf and von Dohren, 1996; Lipmann, 1980; Rapaport et al., 1987; Stachelhaus and Marahiel, 1995). First the adenylation domain activates the cognate amino acid at the expense of Mg\(^{2+}\)-ATP, resulting in the formation of an aminoacyl-adenylate, a reaction similar to the amino acid activation process catalyzed by RNA lipases (Fig. 1.4). Second, the enzyme-attached thiol moiety 4’-phosphopantetheine (4’-PP) attacks the aminoacyl-adenylate to yield the aminoacyl thioester and AMP as leaving group (Fig. 1.5). The latter step of this activation reaction requires the presence of the thiolation domain, which will be discussed below.

The adenylation domain contains 10 conserved core motifs (A1-A10, Fig. 1.2), which functions have been studied in considerable detail for the phenylalanine adenylation domain of the gramicidin S synthetase I (grass; Mootz and Marahiel, 1997a; Stachelhaus and Marahiel, 1995b). The crystal structure of this domain has been determined in the presence of its substrates L-phenylalanine, AMP and Mg\(^{2+}\), and revealed that it consists of two domains, a large N-terminal domain and a smaller C-terminal domain (Connate et al., 1997). Almost all of the conserved core motifs are positioned in the proximity of the active site where the substrates are bound. The core motifs L(ST)exile (A1) and LKAGxAYL(LV)P(IL)D (A2) are probably only conserved for structural reasons as they are far away from the active site. Core motif A1 is part of a large \(\alpha\)-helix, which strongly contributes to the fold of the N-terminal domain. The distance and orientation of the core motif ATSGTTGPKG (A3), which is also the signature sequence of the superfamily of adenylate-forming enzymes, relative to AMP suggests an interaction with the pyrophosphate leaving group. This was corroborated by the observation that introduction of a negative charge by replacing the first glycine (G) by an aspartate (D) resulted in complete inactivation (Saito et al., 1995). Other mutations in this motif resulted in mixed phenotypes. Mutation of the second lysine (K) resulted in drastically decreased activity, whereas mutation of the first lysine had little effect on the activity of the valine-activating domain of the surfactin synthetase (Gocht and Marahiel, 1994; Hamoen et al., 1995; Saito et al., 1995). The core motif FDxS (A4) is strictly conserved in peptide synthetases and the aspartate (D) is presumably involved in interactions with the
\(\alpha\)-amino group of the substrate. This aspartate is not conserved in other adenylate-forming enzymes such as luciferase and acetyl-CoA synthetase, as their substrates do not have an \(\alpha\)-amino group (Marahiel, 1992; Kleinkauf and von Döhren, 1996; Turgay et al., 1992). Part of the core motif NxYGPTE (A5) is involved in contacts with the nucleotide base. The main-chain carbonyl atom of the glycine (G) points towards the \(\alpha\)-amino group of the substrate phenylalanine. This is in line with the observation that substitution of this glycine by an aspartate (D) resulted in loss of activity of the valine activating domain of the gramicidin S synthetase II (GrsB; Saito et al., 1995). The core motif GELxIxGxG(LV)ARGYL (A6), located in the proximity of the active site, could be labeled with 2-azido-ATP and is, therefore, thought to be involved in adenylation, but its precise function remains unclear (Pavela-Vrancic et al., 1994b). The highly conserved core motif Y(KR)TGDL (A7), which is conserved in many ATPases, has also been subjected to site-directed mutagenesis (Gocht and Marahiel, 1994; Serrano et al., 1994). When the aspartate (D) was mutated to an asparagine (N), the enzyme activity was slightly reduced, however when replaced by a serine (S), the activity was severely reduced. In the structure of the phenylalanine adenylation domain, the side chain of this invariant aspartate interacts via hydrogen bonds with the oxygen atoms of the nucleotide ribose moiety. The core motif GRxxxQVKIRGxRIELGEIE (A8) is essential for adenylation and the first arginine (R) possibly interacts with the leaving pyrophosphate group. Mutation of the second glycine (G) resulted in a total loss of activity of the proline-activating domain of GrsB (Tokito et al., 1993). In addition, labeling studies with fluorescein 5'-isothiocyanate and 2-azido-ATP suggested the involvement of at least part of this core motif in the adenylation reaction (Pavela-Vrancic et al., 1994a; b; c). Like the core motif A6, the core motif LPxYM(IV)P (A9), located in the proximity of the active site, could be labeled with 2-azido-ATP and is therefore also thought to be involved in adenylation but, as for core motif A6, its precise function remains unclear (Pavela-Vrancic et al., 1994b). The core motif NGK(LV)DR (A10) is involved in the binding of the substrate phenylalanine as well as to two ribose oxygens of ATP and is thus involved in two key polar interactions with both the adenosine and the amino acid. The strictly invariant lysine (K) of this core motif binds to the \(\alpha\)-carboxylate group. This was confirmed by the observation that conversion of this lysine to a glutamine (Q), in the valine-activating domain of the surfactin synthetase, caused a dramatic reduction in its activity, together with the finding that it could be specifically labeled with fluorescein 5'-isothiocyanate (Hamoen et al., 1995; Pavela-Vrancic et al., 1994a).

The substrate specificity of the adenylation domain of peptide synthetases is moderate compared to that of the aminoacyl tRNA synthetases and, consequently, often a mixture of peptides is produced. In many cases the specificity of amino acid incorporation appears to be dependent on the types of amino acids added to the growth medium (Apka et al., 2001; Galli et al., 1994; Kleinkauf, 1979; Kleinkauf et al., 1991; Lawen and Trabner, 1993; Peypoux et al., 1991; Peypoux et al., 1994; Pieper et al., 1992; Ruttenberg and Mach, 1966; Zocher, 1996), although the specificity differed between different adenylation domains and some adenylation domains exhibit a higher specificity than others (Mootz and Marahiel, 1997a). Based on the relatively high homology between core motif A3 and A6, it was thought that the region spanning these two motifs would accommodate the amino acid binding
Figure 1.5: Schematic representation of acylation of the activated (holo) thiolation domain by an amino adenylated substrate attached to the adenylation domain.

pocket, but sequence alignments showed only a slight tendency towards clustering of adenylation domains activating a specific amino acid (Turgay et al., 1992; Cosmina et al., 1993). Yet, the crystal structure of the phenylalanine adenylation domain of GrsA confirmed that indeed the region between core motif A3 and A6 harboured the binding pocket, and mutational analyses of this region in the surfactin synthetase demonstrated that changing specific residues in this region resulted in an altered substrate specificity (Eppelmann et al., 2002).

**Thiolation domain:** The thiolation domain, also called peptidyl carrier protein (PCP) consists of about 100 amino acid residues (Mootz and Marahiel, 1997a). Like the acyl carrier protein (ACP) in modular fatty acid and polyketide synthases, this domain is an integral part of the multienzyme complex to which the aminoacyl-adenylates are bound as aminoacyl thioesters (Fig. 1.5). Although both domains have similar functions, the degree of overall homology is low, except for the region around the core motif DxFFxxLGG(DH)S(IL) (T, Fig. 1.2). The thiolation domain is usually C-terminally to the adenylation domain, but an exception to this arrangement concerns the modules activating N-methylated amino acids in the cyclosporin A and enniatin synthetase, in which the adenylation- and thiolation domains are separated by the N-methylation domain described below (Haese et al., 1993; Weber et al., 1994).

All thiolation domains contain an invariant serine residue within the core motif DxFFxxLGG(DH)S(IL) (T, Fig. 1.2). This residue is essential for the post-translational binding of a 4'-phosphopantetheine (4'-PP) cofactor to which the activated amino acyladenylate substrates from the adenylation domain are transferred. The amino acyladenylate substrates are bound as carboxy thioesters to the terminal cysteamine thiol group of the cofactor. This has been demonstrated in many studies, including the use of site-directed mutagenesis and affinity labeling (Gocht and Marahiel, 1994; Schlumbohm et al., 1991; Stachelhaus et al., 1996; Stein et al., 1994; Stein et al., 1995; Stein et al., 1996). Biochemical characterization of the thiolation domain revealed that the conversion from the inactive (apo) enzyme to the active (holo) enzyme requires the activity of a specific enzyme, the 4'-PP transferase, which utilizes Coenzyme A (CoA) and the thiolation domain as substrates (Stachelhaus et al., 1996). These studies also showed that a nonintegrated, separately expressed, thiolation domain can be acylated by a separate adenylation domain in the presence of the cognate amino acid, which is strong evidence for the functional independence of the domains and the validity of the multiple carrier model of
nonribosomal peptide synthesis. Further studies on the posttranslational modification of thiolation domains of peptide synthetases, modular fatty acid synthases and polyketide synthases resulted in the discovery of a superfamily of proteins, the 4'-PP transferases, that catalyze the attachment of the 4'PP-cofactor to the thiolation domain resulting in the conversion from the inactive apoproteins to the active holoproteins (Lambalot et al., 1996). These enzymes catalyze the nucleophilic attack of the ß-hydroxy side chain of the conserved serine in the pyrophosphate linkage of CoA, resulting in the transfer of the 4'-PP moiety onto the attacking serine.

Recently, the structure of the third thiolation domain of the B. brevis tyrocidine synthetase subunit, TycB, has been solved using NMR techniques (Weber et al., 2000). The overall topology of this thiolation domain resembles that of the acyl carrier proteins of Escherichia coli and Streptomyces coelicolor, which structures have also been solved, and consists of a distorted four-helix bundle with an extended loop between the first two helices (Crump et al., 1997; Holak et al., 1988). The conserved serine (S) of the thiolation domain of TycB has the same location and is situated within a stretch of seven flexible residues. The structure does not provide any evidence for the presence of a peptide-binding pocket and neither provides information as to how this domain interacts with the adenylation domain and the 4'-PP transferase.

**Condensation domain:** A condensation domain, consisting of about 450 amino acids, is located N-terminally of the adenylation domain, and is required for the coupling of two consecutively bound amino acids (Mootz and Marahiel, 1997a). Peptide synthetases responsible for the biosynthesis of lipopeptides contain an additional condensation domain preceding the first (initiation) module, presumably involved in the coupling of the fatty acid moiety to the first amino acid of the peptide moiety (Duitman et al., 1999). The condensation domain probably also prevents mis-initiation of nonribosomal peptide synthesis at the other (elongation) modules, as deletion of the condensation domain from an elongation module transformed the elongation module to an initiation module (Linne and Marahiel, 2000). In this study it was also postulated that the condensation domain plays a role in the timing of the epimerization reaction, as one of the two presumed amino acid binding sites, the acceptor binding site, is strongly enantioselective. Although in the last couple

![Figure 1.6: Proposed mechanism for the condensation reaction in nonribosomal peptide synthesis.](image)

The amino acid residues attached as thioesters to adjacent thiolation domains via the 4'PP cofactor and the second histidine residue conserved within motif C3 of the condensation domain are shown. A nucleophilic attack of the incoming amino group on the thioester activated carboxyl group of the preceding amino acid is presumed.
of years considerable progress has been made in the characterization of this domain and a mechanism for the condensation reaction has been proposed (Fig. 1.6), still limited biochemical data is available to date. The condensation domain is characterized by a set of 7 conserved core motifs (C1-C7, Fig. 1.2). In recent years the sequence of some of these motifs and of a number of highly conserved amino acid residues have been studied in more detail by means of site-directed mutagenesis and in vitro assays using recombinant peptide synthetase fragments of the gramicidin S and tyrocidine synthetases (Bergendahl et al., 2002). These studies revealed that parts of the core motifs RHExLRTxF (C2) and MHHxISDG(VW)S (C3) are essential for the enzymatic activity of this domain and that mutations of the first lysine (R) of core motif C2, and the second histidine (H) and aspartate (D) of core motif C3 destroyed the activity of the condensation domain. Mutation of the second lysine of core motif C2, and the first histidine of core motif C3 as well as of a highly conserved tryptophane (W) in the proximity of core motif C4 resulted in insoluble proteins and it has been presumed that these amino acid residues are important for the integrity of the structure of the condensation domain. In addition, a number of other highly conserved amino acid residues in the proximity of core motifs C1, C3 and C5 were mutated. However, as these mutations did not affect activity or solubility of the recombinant enzyme fragments, it remains obscure why these amino acid residues are so strongly conserved.

**Thioesterase domain:** The thioesterase domain, consisting of about 200 amino acids, is involved in the release of the synthesized peptide from the multienzyme complex, analogous to the thioesterase domains of type I fatty acid and polyketide synthases, which have been shown to perform a similar reaction (Mootz and Marahiel, 1997a). Although in most cases the thioesterase domain is an integral domain of the synthetase, in the operons encoding the gramicidin S and surfactin synthetases the thioesterases are encoded by a separate gene (Cosmina et al., 1993; Kratzschmar et al., 1989). In contrast to the thioesterase domains of type I fatty acid and polyketide synthases, the activities of the thioesterase domain of peptide synthetases seem to be more complex, as this domain is present in systems producing linear, branched via ester bonds, branched via amide bonds and cyclic peptides. Therefore it was presumed that the thioesterase domains in peptide synthetases are not just simply involved in product release, but perform a cyclization reaction, which results in product release. Research using the heterologously expressed thioesterase domain of the tyrocidine synthetase showed that this domain efficiently catalyzes the cyclization of a decapetide-thioester to form tyrocidine A (Trauger et al., 2000). It was also shown that this thioesterase domain can catalyze the cyclization of other linear peptides as long as the N-terminal amino acid and the penultimate C-terminal amino acid were identical to its natural substrate, and that it is even capable to catalyze pentapeptide-thioester dimerization to form gramicidin S. Additional studies using the heterologously expressed thioesterase domain of the tyrocidine A synthetase, revealed that this domain catalyzes cyclization of peptides of 6-14 residues (Kohli et al., 2001). In the latter studies also the thioesterase domains of gramicidin S and surfactin synthetases were shown to perform cyclization reactions to form gramicidin S and surfactin from their natural substrates.

The thioesterase domain is characterized by the core motif G(HY)SxG TE, (Fig. 1.2), which is present in all
other thioesterases including those in fatty acid and polyketide synthases (Mootz and Marahiel, 1997a). Together with a conserved asparagine (R) and histidine (H) the conserved serine (S) in the core motif forms a catalytic triad. Mutation of the conserved serine resulted in strongly reduced product formation, demonstrating the importance of the conserved serine residue in the catalytic activity of this domain (Kallow et al., 2000). Finally, the crystal structure of several thioesterase domains has been solved and showed that all thioesterases have a comparable tertiary structure (Devedjiev et al., 2000; Tsai et al., 2001).

**Epimerization domain:** The optional epimerization domain, mostly located at the C-terminal end of the corresponding synthetase subunit, consists of about 450 amino acids, and performs the conversion of amino acids from the L- to the D-configuration (Mootz and Marahiel, 1997a). Often these domains are an integral part of the multienzyme complex. However, the synthetases of cyclosporin A and HC-toxin contain modules that incorporate D-amino acids but are devoid of epimerization domains, and it is presumed that in these cases non-integrated racemases perform the reaction (Scott-Craig et al., 1992; Weber et al., 1994). The epimerization reaction mostly takes place when the amino acids are bound as carboxy thioesters on the thiolation domain prior to the condensation reaction, but in some cases it is assumed that this reaction takes place subsequent to the condensation reaction during the peptidyl stage (Schwecke et al., 1992; Shiau et al., 1995; Stindl and Keller, 1993; Stindl and Keller, 1994)

The epimerization domain is characterized by 7 conserved core motifs (E1-E7, Fig. 1.2). The motif HHxISDG(VW)S (E2) is also present in the condensation domain and in the family of acyl transferases and, although little biochemical data is available, epimerization involves a proton abstraction and readdition of the Cα proton of the amino acyl or peptidyl moiety linked to the 4'PP cofactor (Fig. 1.7). Based on the acid/base properties, the catalytic mechanism probably involves the second histidine (H) of this core motif. The assumed involvement of this histidine residue is corroborated by the observed pH dependence of nonribosomal peptide synthesis. However, it remains to be confirmed whether this histidine residue is required for the epimerization reaction.

**Figure 1.7:** Two-base mechanism of substrate epimerization during nonribosomal peptide synthesis. An amino acid attached to the thiolation domain, the location of the second histidine of motif E2, and the conjugate acid of a second enzymic base (indicated with XH) are shown.
**N-Methylation domain:** The optional N-methylation domain consist of about 420 amino acids and methylates specific amino acids (Mootz and Marahiel, 1997a). These domains were first discovered in the fungal cyclosporin A and enniatin synthetases (Haese et al., 1993; Weber et al., 1994). The sequence data of the operons encoding these synthetases revealed a number of novel modules containing an insertion of about 420 amino acids. The occurrence of these modules coincided with the number of N-methylated amino acids in the synthesized product.

The methylation domain contains at least three core motifs (M1-M3, Fig. 1.2), of which the first motif VL(DE)GxGxG (M1) exhibits significant similarity to the common S-adenosylmethionine (SAM) binding site of methyltransferases (Brooks et al., 1991; Haese et al., 1993; Kawakami et al., 1990; Kim et al., 1993; Nwankwo et al., 1994; Weber et al., 1994). Some of the characteristics of the N-methylation domain have been analyzed by overproducing functional fragments of the enniatin synthetase (Haese et al., 1994; Pieper et al., 1995; Zocher, 1996). A 158 kDa C-terminal fragment of the second module of this synthetase, encompassing the adenylation-, N-methylation- and thiolation domain activates L-methylvaline, and can be affinity labeled with [C\(^{14}\)]SAM verifying the presence of an N-methylation domain. Further N- and C-terminal deletions resulted in a 65 kDa fragment of 420 amino acid residues, which harbours the methyltransferase activity. These studies also revealed that N-methylation occurs when the amino acids are bound as carboxy thioesters to the thiolation domain prior to peptide bond formation. Using potent inhibitors such as sinefungin and S-adenosyl-homocysteine, the cosubstrate dependence of the methylation reaction in general and SAM charging of the methylation domain in particular, could be confirmed, as in the presence of these inhibitors a dramatically reduced synthesis of a non-methylated peptide product was observed.

### 1.4 Hybrid synthetases

In recent years increasingly more peptide synthetases have been identified, which contain domains normally present in fatty acid- or polyketide synthases (Du and Shen, 2001; Duitman et al., 1999; Shen et al., 2001; Silakowski et al., 2001; Tsuge et al., 2001a). In these hybrid synthetases all catalytic domains were identified, which are normally present in fatty acid- or polyketide synthases. To date very little biochemical studies have been performed on these hybrid synthetases. However, a considerable quantity of research has been conducted on fatty acid- or polyketide synthases, which share striking architectural and organizational similarities with the peptide synthetases. Like peptide synthetases, fatty acid- or polyketide synthases consist of modules that are made up of a set of three mandatory domains of which the acyltransferase and ketosynthase domains are catalytic and the acyl carrier domain acts as a carrier. Together these domains are responsible for the central chain-building reactions. In addition, fatty acid- or polyketide synthase modules contain optional domains that mediate the modification of the newly extended fatty acid- or polyketide chain such as reduction, dehydration and methylation. Here the details of the domains present in fatty acid- or polyketide synthases will not be discussed, but like the domains of peptide synthetases these also contain conserved amino acid motifs, consisting of amino acid residues essential for their enzymatic
activity or structural integrity, by which they can be easily recognized.

One such hybrid synthetase is the mycosubtilin synthetase of *B. subtilis* ATCC6633, which synthetase subunit, MycA, contains a large N-terminal part consisting of 5 additional domains (Duitman et al., 1999). Presumably, these domains are involved in the conversion of a fatty acid into an ß-amino fatty acid and, based on the reactions normally performed by homologous domains or proteins of this type, a model for this conversion has been proposed. The first domain is an acyl CoA-ligase domain, which is supposed to couple coenzymeA (CoA) to a long chain fatty acid in an ATP dependent reaction. The second and fourth domains are acyl carrier domains, homologous to the thiolation domain in peptide synthetases, to which the activated substrates, acyl-CoA and malonyl-CoA, are presumably coupled via a 4'-PP cofactor. The third domain is a ß-keto acyl synthase domain that probably performs a condensation reaction coupling the bound acyl-CoA and malonyl-CoA molecules. The fifth domain is an amino transferase domain presumably converting the hydroxy group resulting from the condensation reaction into an amino group. As an integral domain, this last domain is rarely found in both peptide synthetases and fatty acid- or polyketide synthases. Also the operon encoding the synthetase of another member of the iturin family, iturin A, has been characterized, and revealed that the iturin A synthetase is almost identical to the mycosubtilin synthetase (Tsuge et al., 2001a). In recent years models for the synthesis of several other modified peptides by hybrid synthetases have been proposed such as for bleomycin (Shen et al., 2001).

In one of the first biochemical studies of a hybrid synthetase the 4 subunits, YbtE, HMWP1, HMWP2 and YbtU, of the yersiniabactin synthetase from *Yersinia pestis* were purified, and yersiniabactin was synthesized *in vitro* using 8 biosynthetic precursors (Miller et al., 2002). In this recently performed study the turnover rate of the synthesis, comprising 22 chemical operations, was determined at 1,4 min⁻¹. In addition, it was shown that during the course of yersiniabactin synthesis, the elongating Acyl-S-enzyme chain transferred first across a nonribosomal peptide synthetase/polyketide synthase interprotein interface between the HMWP2 and HMWP1 subunits, and then across a polyketide synthase/nonribosomal peptide synthetase intraprotein interface within the subunit HMWP1 (Fig. 1.8, I and II, respectively).

**Figure 1.8:** Schematic representation of the 4 proteins, YbtE, HMWP2, HMWP1 and YbtU, of the yersiniabactin synthetase required for *in vitro* synthesis of yersiniabactin. YbtE and YbtU are a salicyl-AMP ligase and a reductase, and HMWP2 and HMWP1 encompass the remaining 15 domains. Abbreviations are as follows: A, adenylation; ArCP, aryl carrier protein; Cy, cyclization; T, thiolation domain; KS, ketosynthase; AT, acyl transferase; MT, methyltransferase; KR, ketoreductase; ACP, acyl carrier protein; TE, thioesterase; NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase. I and II represent the transfer across a nonribosomal peptide synthetase/polyketide synthase interprotein interface (I) and across a polyketide synthase intraprotein/nonribosomal peptide synthetase interface (II).
Chapter 1

1.5 Peptide synthetase operons

The number of operons encoding peptide synthetases increases rapidly as the complete sequence of increasingly more genomes is being determined. Due to the huge size and conserved genetic structure of these operons they can be easily recognized, as they are usually, together with operons encoding polyketide synthetases, the largest operons in the genome, spanning regions up to 50000 basepairs. A number of the peptide synthetase operons of various Bacillus species have been quite well characterized, and their overall structure is similar (Fig. 1.9; Cosmina et al., 1993; Duitman et al., 1999; Mootz and Marahiel, 1997b; Konz et al., 1997; Konz et al., 1999a; Konz et al., 1999b; Kratzschmar et al., 1989; Tsuge et al., 2001a). The structural variation between these Bacillus operons represents the variation of all known peptide synthetase operons. However, the Bacillus operons do not encode peptide synthetases containing modules involved in the incorporation of N-methylated amino acids, as have been identified in the fungal cyclosporin A and enniatin synthetases (Haese et al., 1993; Weber et al., 1994).

Analyses of several operons encoding peptide synthetases revealed that the colinearity rule mentioned earlier not only applies to the modular structure of the peptide synthetases at the translational level but in almost all cases also for the genes encoding peptide synthetases at the transcriptional level. The order of the genes encoding the subunits of the synthetase reflects the order of peptide synthesis: the first gene in the operon, encoding a synthetase subunit, encodes the first subunit containing the initiation module and performs the first step(s) in the synthesis of the peptide. This striking conservation of coordinate transcription and linkage of domains suggests a crucial role in the protein assembly and functioning. However, in studies with the surfactin synthetase in which the concerted expression of the synthetase subunits was destroyed, proper assembly and surfactin production was observed (Guenzi et al., 1998a). These results clearly demonstrated that the coordinate transcription and linkage of domains are not essential for proper assembly and activity of peptide synthetases. In addition, characterization of the syringomycin synthetase operon of Pseudomonas syringae revealed that this operon does not reflect the order of peptide synthesis (Guenzi et al., 1998b). This also applies to the gramicidin synthetase operon, since the first gene of this operon encodes a thioesterase performing the last step in the synthesis of gramicidin (Kratzschmar et al., 1989).

Surprisingly, some genes encoding proteins essential for the production of NSP’s are rarely part of the operons encoding the peptide synthetases. This applies, for instance to the genes encoding the 4'-PP transferases, which have never been found to be part of the peptide synthetase operons. In addition, although most precursors for nonribosomal peptide synthesis are also precursors in other cellular processes, the synthesis of lipopeptides often requires rather specific fatty acid moieties, and to date little is known about the origin of these precursors as the proteins synthesizing these precursors are not part of the peptide synthetase operons. Finally, genes encoding the secretory proteins which facilitate secretion and are often required for the survival of the producer during production, are rarely part of the peptide synthetase operons.
Figure 1.9: Schematic diagram showing the well-characterized operons encoding peptide synthetases as well as the deduced modular organization of these peptide synthetases of some closely related *Bacillus* species. The peptide synthetase operons presented are gramicidin S (*grs*) and tyrocidine (*tyc*) operons of *B. brevis*, the bacitracin (*bac*) and lichenysin A operons of *B. licheniformis* and the fengycin (*pps*), iturin A (*itu*), mycosubtilin (*myc*) and surfactin (*srf*) operons of *B. subtilis*. The activated amino acids are depicted within the adenylation domains, and the abbreviations: Orn and aThr represent the nonproteinogenic amino acids ornitine and allo-threonine, respectively.
1.6 Expression and transcriptional regulation of peptide synthetase operons

Although the last few decades have shown a rapid increase in the knowledge about the NSP's as well as the structure and function of the corresponding peptide synthetases, relatively little research has been performed to date to study expression and transcriptional regulation of the operons involved. The only exceptions are the mycosubtilin synthetase operon, and especially the surfactin synthetase operon of *B. subtilis*, the latter because of its involvement in development of genetic competence (Cosmina *et al.*, 1993; Hamoen *et al.*, 1995). The results of studies performed until now suggest that expression and transcriptional regulation of peptide synthetase operons in various *Bacillus* species are controlled by the interactions of diverse factors which also function in controlling other stationary-phase-induced processes like sporulation and the development of genetic competence.

**B. brevis**: Only *B. brevis* cells in stationary-phase cultures produce the cyclic decapetides gramicidin S and tyrocidine. Studies on the expression and transcriptional regulation of the gramicidin S (*grs*) and tyrocidine (*tyc*) synthetase operons of *B. brevis* in *B. subtilis* also suggested that expression of both operons is induced in the beginning of the stationary growth phase (Marahiel *et al.*, 1993).

Analysis of the *grs* promoter revealed that no σ^A^-like promoter sequence is present in this region. However, 2 regions centered around the -10 (GGAATTCACT) and -45 (AAGATTTT) exhibit some similarity to σ^H^-like promoter sequences recognized by a minor holoenzyme form of RNA polymerase involved in the transcription of several genes, which are expressed at the beginning of the sporulation process. As no σ^H^-like promoter sequence was identified in the vicinity of -35, *grs*-promoter driven β-galactosidase expression was measured in a *B. subtilis* strain containing a mutation in *spo0H*, encoding σ^H^-like (Marahiel *et al.*, 1993). Indeed, β-galactosidase expression was totally absent in this mutant. However, since the *spo0H*-dependent expression of *grs* in a wild-type *B. subtilis* strain is induced 2 hours after the end of the exponential growth phase, an indirect role for *spo0H* or the requirement for an additional activator has been postulated.

The putative promoter of *tyc* shows sequence similarity to σ^A^-like promoter sequences utilized by the major RNA polymerase holoenzyme form. Analyses of *tyc* expression and transcriptional regulation using a *tyc-lacZ* (β-galactosidase) fusion in *B. subtilis* revealed that, in accordance with tyrocidine production in *B. brevis*, expression was induced at the transition from exponential to stationary growth (Marahiel *et al.*, 1987). In addition, *tyc* expression was completely absent in cells carrying mutations in *spo0A*, *spo0B* and *spo0E*, encoding proteins involved in the phosphorelay. These results are in agreement with the observation that mutations in *spo0A* and *spo0B* block antibiotic production in *B. brevis* (Brehm *et al.*, 1973; Schaefer, 1969). Introduction of a mutation in the gene encoding the transition state regulator AbrB in *tyc-lacZ*-containing *B. subtilis* strains bypassed the effects of *spo0A* mutations, and resulted in constitutive *tyc* expression in a *spo0A* mutant as well as in the wild type (Marahiel *et al.*, 1987). Together with the observation that *spo0A* mutations result in overexpression of *abrB*, these results suggest that *tyc* expression is induced under nutrient limitation through a Spo0A-dependent mechanism, which relieves the repression caused by AbrB (Perego *et al.*, 1988; Fürbass *et al.*, 1991). Binding studies,
using purified Spo0A and AbrB supported this assumption, by showing that Spo0A binds to the \textit{abrB} promoter as well as the \textit{tyc} promoter (Fürbass \textit{et al.}, 1991; Robertsen \textit{et al.}, 1989; Strauch \textit{et al.}, 1989). Although little data is available similar regulatory proteins seem to be involved in the transcriptional regulation of \textit{grs} and \textit{tyc} in \textit{B. brevis} as in the transcriptional regulation of \textit{srfA} and \textit{mycS} in \textit{B. subtilis} (Fig. 1.10 and Fig. 1.12).

\textbf{B. licheniformis}: In \textit{B. licheniformis} only some initial studies have been done on the expression and transcriptional regulation of the lichenysin A synthetase operon \textit{lchA}. Analysis of the promoter region revealed the presence of an inverted repeat with similarities to that present in the surfactin synthetase operon promoter to which the response regulator ComA, described below, binds. To study the involvement of this dyad symmetry and ComA in the regulation of

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure10.png}
\caption{Schematic representation of the transcriptional regulation of the operons encoding the gramicidin S and tyrocidine synthetases in \textit{B. brevis}. Part of the work has been performed in \textit{B. subtilis} and, although closely related, the conclusions with respect to the regulation in \textit{B. brevis} remain putative.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure11.png}
\caption{Schematic representation of what is known of the transcriptional regulation of the lichenysin A operon in \textit{B. licheniformis} based on experiments performed in \textit{B. subtilis}.}
\end{figure}
**ichA expression**, a promoter fusion of ichA with a green fluorescent protein (PlchA-gfp) was introduced in *B. subtilis* (Yakimov and Golyshin, 1997). Elimination of the dyad symmetry resulted in loss of fluorescence, whereas improvement of the dyad symmetry resulted in a 2-fold increase in fluorescence. Moreover, ichA transcription in *B. subtilis* was severely reduced when comA and comP, encoding the accompanying histidine kinase of ComA, were deleted. These results suggest that the transcriptional activation of ichA via ComA and ComP is comparable to that of the transcriptional activation of srfA in *B. subtilis* (Fig. 1.11 and Fig. 1.12).

**B. subtilis**: Owing to the easy genetic accessibility of *B. subtilis*, the expression and transcriptional regulation of a number of peptide synthetase operons in this bacterium have been studied in much more detail. Especially, the expression and transcriptional regulation of the surfactin synthetase operon, srfA, in *Bacillus subtilis* strain 168 derivatives has been studied extensively as this operon also encodes a small protein, ComS, essential for development of genetic competence.

Expression of the surfactin synthetase operon, srfA, is dependent on the growth-phase as well as the composition of the growth-medium (van Sinderen et al., 1990). In addition, considerable differences have been observed in the level of expression of srfA in different *B. subtilis* strains (Duitman et al., submitted for publication). The expression increases sharply at the transition from exponential to stationary growth and is highest when cells are cultured in minimal medium. Expression of srfA is mainly dependent on a two-component regulatory system, ComA and ComP (Nakano and Zuber, 1991; Roggiani and Dubnau, 1993). The histidine kinase ComP senses the accumulation of the pheromone ComX in the medium and reacts at a critical ComX concentration by phosphorylating itself (Magnuson et al., 1994; Solomon et al., 1996). Subsequently, the phosphorylated ComP transfers the phosphoryl group to ComA, which stimulates the binding of this response regulator to an inverted repeat in the σ^A-dependent srfA promoter, and induces expression of the srfA operon (Hahn and Dubnau, 1991; Weinrauch et al., 1990). The response regulator aspartyl-phosphate phosphatase, RapC, performs the opposite reaction and dephosphorylates ComA (Grossman, 1995). The activity of RapC depends on the intracellular concentration of the accompanying pentapeptide, PhrC, previously known as the competence-stimulating factor (CSF) owing to its role in development of genetic competence (Lazazzera et al., 1997; Solomon et al., 1996). After processing and secretion into the medium, PhrC is taken up again by the cells by the oligopeptide permease Spo0K and inhibits the activity of RapC. Because the ComX and PhrC concentrations in the medium are dependent on the cell density, expression of srfA is mainly under cell density control. In addition, expression of srfA is repressed by the nutritional repressor CodY, which directly binds to the srfA promoter upon increasing concentrations of casamino acids in the growth medium (Serror and Sonenshein, 1996). Finally, deletion of abrB and a certain mutation in degU, causing hyperphosphorylation of the response regulator DegU, result in decreased expression of srfA (Hahn and Dubnau, 1991; Weir et al., 1991).

Recently, a number of genes were identified that block the biosynthesis of the dipeptide bacilysin in *B. subtilis* 168 (Yazgan et al., 2001). One of the genes was oppA, which is part of the spo0K operon encoding the oligopeptide permease, also involved in srfA transcriptional regulation in this strain. This suggested a comparable
regulation mechanism involving ComA, PhrC and RapC. Mutational analysis confirmed that, indeed, ComA and PhrC are involved in transcriptional regulation of a gene or genes essential for bacilysin biosynthesis. Deletion of the gene encoding another Phr, PhrA, did not influence bacilysin production.

A recent study in which the expression and transcriptional regulation of the mycosubtilin synthetase operon, mycS, in *B. subtilis ATCC6633* was compared with that of the surfactin synthetase operon, *srfA*, in *B. subtilis 168*, revealed that expression and transcriptional regulation of mycS differ substantially from that of srfA (Duitman et al., submitted for publication). Expression of mycS, like srfA, increases sharply at the transition from exponential to stationary growth but reaches higher levels when cells are cultured in richer media. Although mycS expression is not dependent on ComA,
addition of PhrC to the culture medium resulted in increased expression whereas deletion of spo0H resulted in decreased expression. It was also shown that decreased mycS expression, caused by a spo0H deletion, could be bypassed by addition of PhrC to the culture medium. This demonstrated that mycS expression is not σH-dependent and that the observed decreased mycS expression was due to the absence of PhrC, which expression is σH-dependent (Solomon et al., 1996).

Surprisingly, deletion of spo0K did not influence mycS expression, suggesting the presence of another oligopeptide permease in the mycosubtilin producing strain. This was corroborated by the observation that in B. subtilis ATCC6633, deletion of spo0K also did not affect srfA expression. Expression of mycS is also regulated by AbrB, and deletion of the abrB gene resulted in increased expression. Because the mycS expression in cells carrying an abrB deletion continued to be growth-phase dependent, unknown additional regulators are apparently involved in the growth-phase dependent expression of mycS. Although differences between the transcriptional regulation of mycS, srfA and unknown genes encoding proteins involved in bacilysin synthesis in B. subtilis have been found, transcriptional regulation of these genes or operons involves often the same proteins (Fig. 1.12).

1.7 Peptide secretion

Organisms producing nonribosomally synthesized peptides of which many have antibiotic activities need efficient systems to facilitate secretion of the synthesized peptides. Often such secretion systems are also essential for the self-resistance to the antibiotic peptide producer. In many organisms the secretion and self-resistance is mediated by ABC-transporters, which couple ATP hydrolysis to translocation across the membrane, or by membrane proteins, which energize translocation across the cell membrane by proton-dependent transmembrane electrochemical gradients. In contrast to the other aspects of nonribosomal peptide synthesis in the various Bacillus species, relatively little is known about the secretion of these peptide compounds and the resistance mechanism of the producer strains.

B. brevis: Although gramicidin S has been subject of many studies still little is known about the secretion and resistance mechanism. The secretion and resistance mechanism of the other NSP, tyrocidine, presumably involves two tandem ABC-transporters, tycD and tycE, encoded by genes directly downstream of the tyrocidine synthetase genes. Both genes are probably part of the tyrocidine operon and show high similarity with two tandem ABC-transporters identified in a locus with peptide/polyketide synthetases, which are involved in product secretion (Mootz et al., 1997b).

B. licheniformis: Of the NSP’s produced by B. licheniformis only information on the resistance mechanism and secretion of bacitracin is available. Bacitracin exhibits strong antibiotic activities especially against Gram-positive bacteria and because the producer B. licheniformis is not sensitive, the existence of a specific resistance mechanism has been presumed. Several factors by which bacitracin producers as well as other bacteria gain self-resistance are known, such as increased production of C55-isoprenyl pyrophosphate and secretion of exopolysaccharides (Chalker et al., 2000; Pollock et al., 1994). As bacitracin forms a tight ternary complex with C55-isoprenyl pyrophosphate, preventing recycling of this peptidoglycan carrier, the increased resistance is apparently due to increased production of C55-isoprenyl pyrophosphate.
However, the mechanism underlying increased secretion of exopolysaccharides remains obscure. In addition, 3 genes, \textit{bcrA}, \textit{bcrB} and \textit{bcrC}, have been identified in \textit{B. licheniformis}, which encode an ABC-transporter that mediates resistance to bacitracin in \textit{B. subtilis} and \textit{E. coli} by active transport of the peptide (Podlesek \textit{et al.}, 1995). These 3 genes are organized in an operon located 3 Kb downstream of the bacitracin operon in \textit{B. licheniformis}. In between the bacitracin operon and the bacitracin self-resistance operon 2 genes were identified, \textit{bacR} and \textit{bacS}, with high homology to response regulators and sensors of two-component regulatory systems. Further research revealed that BacR and BacS act in negative regulation of the genes encoding the ABC-transporter (Neumuller \textit{et al.}, 2001).

\textbf{B. subtilis}: Although \textit{B. subtilis} is genetically easy accessible and thus extensively studied, virtually nothing is known about the secretion mechanism of the nonribosomally synthesized lipopeptides. The complete genome sequence revealed numerous genes encoding proteins homologous to ABC-transporters or other putative membrane-located proteins but for most of these the function remains unknown (Kunst \textit{et al.}, 1997). Although several genes encoding ABC-transporters were identified in the vicinity of the surfactin synthetase operon, deletion did not affect surfactin production (Rodriguez and Grandi, 1995). However, recently a gene, \textit{yerP}, was identified in \textit{B. subtilis} encoding a protein homologous to the proton motive force-dependent efflux pump family involved in resistance, nodulation and cell division in several bacterial species. The putatively membrane located protein YerP seems to play a role in secretion of surfactin and self-resistance against surfactin of \textit{B. subtilis} (Tsuge \textit{et al.}, 2001b). Deletion of the \textit{yerP} gene resulted in increased sensitivity for surfactin and decreased levels of surfactin production. The expression of \textit{yerP} was at its maximum at the end of the exponential growth phase, and was not induced by surfactin. Whether YerP is involved in self-resistance and secretion of the other lipopeptides produced by \textit{B. subtilis} is not known. Finally, the residual self-resistance to surfactin and remaining surfactin production suggest the presence of other surfactin secretion and self-resistance mechanisms in \textit{B. subtilis}.

1.8 FUTURE PROSPECTS

Recent years have seen a rapid increase in the understanding of nonribosomal peptide synthesis. The conserved modular organization of peptide synthetases provides the means for the creation of new peptide synthetases by genetic engineering strategies, either by recombining modules as has been done for the surfactin synthetase or by changing the specificity of certain modules in existing synthetases as applied to gramicidin synthetase (Conti \textit{et al.}, 1997; Stachelhaus \textit{et al.}, 1995). In addition, the prospect of creating numerous new bioactive peptides, by the engineering of existing peptide synthetases, has dramatically increased by the discovery of numerous hybrid synthetases, consisting of peptide synthetase modules as well as fatty acid- or polyketide synthetase modules (Du and Shen, 2001; Duitman \textit{et al.}, 1999; Shen \textit{et al.}, 2001; Silakowski \textit{et al.}, 2001; Tsuge \textit{et al.}, 2001a). The discovery that many pathogenic microorganisms are dependent for their pathogenicity on peptide synthetases not only makes them a potential source for new antibiotics, but also a potential source for new drug targets.

Still, much remains to be done to fully exploit the potential of peptide synthetases. Although there is an increasing number of
examples in which engineered peptide synthetases are functional, the low yields indicate the presence of numerous "bottle necks" for the efficient production of novel peptides. Also, little is known about the interactions between the various domains within a peptide synthetase, hampering the efficient engineering of peptide synthetases by exchanging specific domains. In addition, little is known about the secretion of many of these nonribosomally synthesized peptides. The same applies to the scarce knowledge of the expression and regulation of the peptide synthetase operons, and the nearly complete absence of information about the stability of the RNA produced from these huge operons.

1.9 Outline of this thesis

The NSP's produced by various strains of *B. subtilis* mainly consist of cyclic lipopeptides of which surfactin is the most well-studied. The main subject of this thesis concerns the lipopeptides mycosubtilin, a member of the iturin family, and surfactin produced by a number of *B. subtilis* strains.

In Chapter 2 the identification, sequence determination and characterization of the mycosubtilin synthetase operon, *mycS*, of *B. subtilis* ATCC6633 is described, and a model is presented for the initiation of mycosubtilin synthesis. This model is based on the structure of this lipopeptide and on the reactions performed by domains or proteins homologous to the various domains present in the mycosubtilin synthetase.

Chapter 3 concerns the development of two methods to facilitate transformation of *B. subtilis*. Because genetic analysis of *mycS* of *B. subtilis* ATCC6633 was frustrated by the exceedingly low levels of competence in this *B. subtilis* strain, a method was developed to increase the competence levels. This method makes use of the autostimulatory effect of ComK on its own transcription to increase competence levels in *B. subtilis* ATCC6633, *B. subtilis* ATCC21332 and probably other *B. subtilis* strains that are poorly transformable due to low levels of competence development. In addition, a method was developed, which enables Campbell-type integration in *B. subtilis* using ligation products, as often plasmids containing *B. subtilis* chromosomal DNA fragments are not stably maintained in the intermediate cloning host *E. coli*. The addition of high concentrations of polyethyleneglycol 8000 to the ligation mixture causes macromolecular crowding and results in the formation of large linear ligation products that appeared to be efficient substrates for genomic integrations in *B. subtilis*.

In Chapter 4 the expression and transcriptional regulation of *mycS* and *srfA* in *B. subtilis* ATCC6633 are described, and a comparison is made with the expression and transcriptional regulation of *srfA* in *B. subtilis* 168. The effects of deletion of several known transcriptional regulators of *srfA* as well as those engaged in other post exponential processes in *B. subtilis* 168 on transcription of *mycS* and *srfA* in *B. subtilis* ATCC6633 was determined using *lacZ* reporter gene fusions. In addition, random mutagenesis was used to identify unknown transcriptional regulators of *mycS*.

Chapter 5 is mainly concerned with the effects of surfactin on the integrity of phosphatidylcholine bilayers. Carboxyfluorescein leakage from loaded vesicles, caused by the addition of increasing surfactin concentrations at different pH values was determined. Also, some additional experiments are described in which the influence of a ψ and ΔpH on the surfactin-induced carboxyfluorescein leakage from vesicles was studied.
1.9 References


Kunst, F., Ogasawara, N., Moszer, I. et al. 1997. The complete genome sequence of...
the Gram-positive bacterium *Bacillus subtilis*. Nature **390**: 249-256.


Neumuller, A. M., Konz, D. and Marahiel, M. A. 2001. The two-component regulatory system BacRS is associated with bacitracin...


Tsuge, K., Akiyama, T. and Shoda, M. 2001a. Cloning, sequencing, and