Pre- and post-translocational stages in protein secretion by Bacillus subtilis
Zanen, Geeske Elisabeth

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:
2005

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.
Introduction

Parts of this chapter were published in:

1. The cell factory *Bacillus subtilis*

*Bacillus subtilis* and related Bacillus species are well known in industry as host organisms. These bacterial species are very attractive for commercial production of proteins, because the cultivation in large-scale production systems is relatively easy and usually inexpensive. For the economical production of food products and pharmaceuticals, the system of production, including the appropriate hosts, is crucial for the downstream processing of these (recombinant) proteins. *Bacillus* species have a high capacity to secrete proteins directly into the growth medium (Simonen and Palva, 1993) resulting in a natural separation of the product from cell components. This simplifies the downstream processing of the protein of interest and may provide better folding conditions compared to the reducing environment of the cytoplasm (Moks et al., 1987).

The use of *B. subtilis* for the production of food products and pharmaceuticals is highly favored over the use of *E. coli*, because *B. subtilis* is considered as a GRAS organism (generally regarded as safe). In contrast to *E. coli*, *B. subtilis* does not produce endotoxins, which are pyrogenic in humans and other mammals and thus have to be removed and complicate the product purification (Bredmose et al., 2001; Petsch and Anspach, 2000). In industry, about 60% of the commercially available enzymes are produced by *Bacillus* species, for example proteases for detergents and amylases used in the starch industry (Palva, 1982; Quax, 2003; Schweder and Jurgen, 2001; Simonen and Palva, 1993). However, most of these enzymes are endogenous to *Bacillus* and naturally secreted into the growth medium and at this moment none of the approved recombinant protein therapeutics is being produced by this organism.

The use of *Bacillus* species for production of biopharmaceuticals of eukaryotic origin is often severely hampered as the result of several bottlenecks that arise during the secretion process, such as poor targeting to the translocase, degradation of the secretory protein, and slow or incorrect post-translocational folding. Genomic and proteomic analyses facilitated by the availability of the complete genome sequence of *B. subtilis* (Kunst et al., 1997) have greatly increased our knowledge of protein transport pathways and their (putative) substrates (Antelmann et al., 2001; Tjalsma et al., 2000a; van Dijl et al., 2001; 2002). Exploiting this knowledge might in the near future result in a more widespread use of *B. subtilis* as a cell factory for the production of pharmaceuticals.

2. Introduction on protein secretion

In all three kingdoms of life (prokaryotes, eukaryotes and archaea), proteins are exported from the cytoplasm to destinations outside the cell. In bacteria, most proteins destined to leave the cytoplasm are exported *via* the highly conserved Sec pathway. More specialized bacterial export pathways are used for the export of
specific subsets of extracellular proteins. The basic principles of protein transport across membranes, as described below, apply to most eukaryotic and prokaryotic organisms (Economou, 1998; Pohlschröder et al., 1997; Riezman, 1997; Schatz and Dobberstein, 1996; Tjalsma et al., 2000).

Most exported proteins are synthesized as precursors with an N-terminal signal peptide (von Heijne, 1990b; 1998). These signal peptides contain information about the destination of these proteins and the transport machinery that should be used to reach this destination (von Heijne, 1998). Precursors are first recognized by soluble cytoplasmic targeting factors that support their transport to the translocation machinery in the cell membrane. Next, the polypeptide chain is transported through a translocase, a proteinacious channel in the membrane. Finally, the signal peptide is removed, resulting in the release of the mature protein from the membrane. Depending on the nature of the export signal, and the presence or absence of retention signals, the protein is either retained in the membrane (Tjalsma et al., 1999a), or released from the membrane due to processing by specialised signal peptidases (SPases; Dalbey et al., 1997; Tjalsma et al., 2001). When the protein is translocated in an unfolded state, it folds into its final conformation after leaving the translocase.

3. Protein sorting in *Bacillus subtilis*

Although the soil bacterium *B. subtilis* has a relatively simple cell structure, proteins can at least be delivered to, or retained at, five (sub)cellular locations: the cytoplasm, the cytoplasmic membrane, the membrane-cell wall interface, the cell wall and the growth medium (Tjalsma et al., 2000). As depicted in Figure 1, the final destination of a protein is defined by the presence or absence of signal peptides and/or retention signals.

Nearly all proteins of *B. subtilis* lacking transport signals are retained in the cytoplasm and fold, with or without the aid of chaperones, into their native conformation. Other proteins contain membrane-spanning domains that are required for their insertion into the cytoplasmic membrane. Most proteins that are completely transported across the cytoplasmic membrane are synthesized with N-terminal signal peptides. As *B. subtilis* lacks an outer membrane, many of these proteins are secreted directly into the growth medium. Other exported proteins involved in processes such as cell wall turnover, substrate binding, or the folding and modification of translocated secretory proteins, have to be retained at the membrane-cell wall interface to fulfil their function. In the following sections, signal peptides, translocation pathways, and accessory components that act before or after translocation, that are known to be involved in protein sorting in *B. subtilis* will be discussed in the light of recent findings.
a) Signal Peptides

When emerging from the ribosome, signal peptides are recognised by chaperones and/or targeting factors, which keep the protein unfolded and translocation-competent. This is essential since premature folding precludes protein export via the Sec pathway, which might be harmful as it could result in the jamming of this pathway (Laminet and Pluckthun, 1989; Liu et al., 1989; Park et al., 1988).

Three distinct domains, the N, H, and C domains, are generally present in signal peptides (von Heijne, 1984; 1989; 1990a; 1990b). The N-domain contains at
least one arginine or lysine residue which, at least for *E. coli*, has been suggested to interact with the translocation machinery and/or the negatively charged phospholipids in the lipid bilayer of the membrane (Akita *et al.*, 1990, de Vrije *et al.*, 1990). As a result, the N-domain of the signal peptide serves as the topological determinant for the precursor in the membrane (von Heijne, 1986; 1992; 1994). The H-region, following the N-region, is formed by a stretch of hydrophobic residues that can adopt an α-helical conformation in the membrane (Briggs *et al.*, 1986). Helix-breaking glycine or proline residues are often present in the middle of this hydrophobic core, which allows the formation of a hairpin-like structure that can insert into the membrane. It has been proposed that subsequent unlooping of this hairpin results in the insertion of the complete signal peptide into the membrane (de Vrije *et al.*, 1990). Analyses of the role of the H-domain showed that the hydrophobic core is the dominant structure in determining signal peptide function (Gennity and Inouye, 1991; Hikita and Mizushima, 1992a; 1992b; Chapter 4). Helix-breaking residues at the end of the H-domain are involved in creating a β-stranded conformation, which facilitates cleavage by SPases (Paetzel *et al.*, 1998). The C-domain, following the H-domain, contains the cleavage site for these SPases, which remove signal peptides from the mature part of the exported protein during or shortly after translocation. Finally, the signal peptide is degraded by signal peptide peptidases (SPPases) and removed from the membrane. Despite the similar structure of signal peptides, apparently small variations can result in transport to different destinations and/or export via different pathways. On the other hand, the same signal peptides can behave differently in different hosts, as shown by Collier (1994) and also described in Chapter 3. In this respect, it is interesting to note that signal peptides of Gram-positive bacteria are usually longer and more hydrophobic than those of Gram-negative bacteria.

b) Signal Peptide Classification

Computer-based predictions showed that 300 proteins with the potential to be exported can be distinguished in *B. subtilis* (Tjalsma *et al.*, 2000). On the basis of SPase cleavage sites and the export pathways via which these precursors are (predicted to be) exported, signal peptides can be divided into five distinct classes: I) secretory (Sec-type) signal peptides, II) lipoprotein signal peptides, III) twin-arginine (RR-type) signal peptides, IV) pseudopilin-like signal peptides, and V) signal peptides of bacteriocins and pheromones. These different classes of signal peptides are depicted in Fig. 2 and are described in more detail below.

I. Secretory (Sec-type) Signal Peptides. The most abundant class is composed of “typical” secretory signal peptides. These signal peptides direct proteins into the Sec pathway. Subsequently, they are cleaved by so-called type I signal peptidases (SPase I) and, as a result, the corresponding mature proteins are
released into the growth medium. The 135 predicted signal peptides of this type have an average length of 28 residues and contain 2-3 positively charged lysine or arginine residues in their N-domains. The hydrophobic core (H-domain) has an average length of 19 residues and about 60% of these predicted Sec-type signal peptides contain a helix-breaking residue in the middle of this domain. The C-domain carries the type I SPase cleavage site, with the consensus sequence A-S-A at positions -3 to -1 relative to the cleavage site. Notably, differences between Sec-type signal peptides might underlie differences in recognition by targeting factors. In this respect, the most prominent difference is that signal peptides directing proteins into the Signal Recognition Particle (SRP)-dependent pathway have a significantly more hydrophobic H-domain than those mediating SRP-independent targeting in *E. coli* (Lee and Bernstein, 2001; Valent *et al.*, 1997) and in *B. subtilis* (Chapter 3).

II. Lipoprotein Signal Peptides. The second class of signal peptides is present at the N-terminus of pre-lipoproteins that are exported via the Sec pathway, lipid-modified, and cleaved by the so-called type II Signal Peptidase (SPase II). Lipoprotein signal peptide predictions resulted in the identification of 114 potential substrates for the lipoprotein-specific (type II) SPase (Tjalsma *et al.*., 1999a; 2001). Signal peptides from lipoproteins have an average length of 19 residues. Both the N-domain (average of 4 residues) and the H-domain (average of 12 residues) are shorter than the corresponding domains in the Sec-type signal peptides that are cleaved by type I SPases. Furthermore, helix-breaking residues are not conserved in the H-region of lipoprotein signal peptides. The C-domain contains a so-called “lipobox” with the consensus sequence L-[A/S]-[A/G]-C. The invariable cysteine residue of this lipobox is the target for lipid modification and the first residue of the mature lipoprotein after cleavage by SPase II (Tjalsma *et al.*, 2000). In fact, the lipid modification of this cysteine is indispensable for signal peptide cleavage by SPase II and anchors the lipoprotein to the membrane.

III. Twin-Arginine (RR-type) Signal Peptides. Signal peptide predictions resulted in the identification of ~180 potential substrates for type I SPases. A twin-arginine motif, containing at least the double arginine in combination with one hydrophobic residue (#) of the consensus sequence R-R-X-##-## was found in 15 of these signal peptides (Jongbloed *et al.*, 2002). Another 13 proteins containing a transmembrane domain or a signal peptide with a Type II SPase cleavage site also contain this RR-X-##-## motif. The presence of such twin-arginine motifs was initially interpreted as an indication that the corresponding pre-proteins could be directed into the Tat pathway for protein export. The predicted twin-arginine signal peptides with a consensus R-R-X-##-## motif containing a predicted type I SPase cleavage site, have an average length of 36 amino acid residues, which is
Figure 2. N-terminal signal peptide classification

On the basis of the presence or absence of signal peptidase cleavage sites and the export pathways via which the precursors are exported, predicted signal peptides (Tjalsma et al., 2000) were divided into five distinct classes: twin-arginine (RR) signal peptides, secretory (Sec-type) signal peptides, lipoprotein signal peptides, pseudopilin-like signal peptides, and bacteriocin and pheromone signal peptides. Most signal peptides have an over-all similar structure: a positively charged N-domain (N), containing Lysine and/or Arginine residues (indicated with ‘+’), a hydrophobic H-domain (H, indicated by a gray box), and a C-domain (C) that carries the cleavage site for their specific signal peptidases (X). The average length of signal peptides and their sub-domains is indicated (numbers between brackets). Furthermore, helix breaking residues, mostly glycine or proline (G/P), in the H-domain of Sec-type signal peptides are indicated. These residues are thought to facilitate loop-wise membrane insertion and cleavage by type I signal peptidase, respectively (Tjalsma et al., 2000). Finally, where appropriate, the most frequently occurring first amino acid of the mature protein (+1) is indicated. Note that the average values for twin-arginine signal peptides are based on the signal peptides of PhoD and YwbN, the only B. subtilis proteins that are known to be secreted in a strictly Tat-dependent manner.
significantly larger than the length of typical Sec-type signal peptides. This is mainly due to the fact that the N-domains of these signal peptides are twice as long as the N-domains of Sec-type signal peptides and have an average length of 14 amino acid residues. Furthermore, these N-domains contain, on average, more positively charged residues than those of Sec-type signal peptides (Tjalsma et al., 2000).

Interestingly, at this moment, transport of only two of the 28 precursors containing consensus RR-signal peptides has been demonstrated to be strictly Tat-dependent: PhoD and YwbN (Jongbloed et al., 2002; 2004). Notably, the length of the signal peptides of PhoD (51 residues) and YwbN (44 residues) is much longer than the average of Sec-type signal peptides (Fig. 2). Remarkably, transport of the predicted “RR-protein” LipA of B. subtilis, a protein containing a signal peptide with a length of “only” 31 residues, is SecA-dependent (Jongbloed et al., 2002). This suggests that many predicted twin-arginine signal peptides might be recognised by components of other export machinery, for example by SecA. Taken together, this could indicate that the length of the signal peptide, especially the N-domain, is also critical for Tat-dependent transport. Another possibility is that the nature of the mature protein plays a role in Tat-dependent translocation, since a fusion of the PhoD signal peptide to the mature part of the Sec-dependent beta-lactamase protein is transported in a Tat-independent manner (Jan Jongbloed, personal communication).

IV. Pseudopilin-like Signal Peptides. The fourth class is formed by signal peptides of pseudopilins, which, in B. subtilis, are cleaved by the SPase ComC (Lory, 1994). Only four proteins with pseudopilin like signal peptides have been identified in B. subtilis; ComGC, GD, GE and GG (Tjalsma et al., 2000). These pseudopilin signal peptides have an average length of 33 residues. Strikingly, the C-domain of these signal peptides, with the consensus sequence K-G-F at positions -2 to +1 relative to the SPase cleavage site, is located between the N- and H-domains. This is in line with the observation that the pseudopilin signal peptidase (ComC) acts at the cytoplasmic side of the membrane (Lory, 1994). In addition to processing, ComC is also responsible for the amino-methylation of the phenylalanine residue at position +1 relative to the cleavage site. Although pseudopilin signal peptides show structural similarities to the previously described signal peptides, pseudopilin precursors are believed to bypass the Tat and Sec pathways and are probably transported via the specific Com pathway (Chung and Dubnau, 1995; Chung et al., 1998; Tjalsma et al., 2000).

V. Signal Peptides of Pheromones and Bacteriocins. The fifth class of signal peptides is found on ribosomally-synthesized pheromones and bacteriocins that are exported and cleaved by ATP Binding Cassette (ABC) transporters (Nakano and
Zuber, 1990). This specific class of signal peptides is often referred to as “leader peptides”. These leader peptides only consist of N-, and C-domains, and completely lack a hydrophobic H-domain. It has been described that parts of the mature protein are also required for export by a dedicated ABC transporter. Moreover, leader peptides are known to have important functions in the prevention of premature antimicrobial activities, and are required for the post-translational modification of pheromones and bacteriocins (van der Meer et al., 1994; van Kraaij et al., 1999).

c) Retention Signals
The membrane-cell wall interface of B. subtilis defines a cellular area that is analogous to the Gram-negative periplasm and contains many proteins that fulfil important functions (Merchant et al., 1995; Pooley et al., 1996). Proteins retained at the membrane-cell wall interface include substrate-binding proteins, chaperones and folding factors involved in the protein secretion process, RNases, DNases, enzymes involved in the synthesis of peptidoglycan, and cell wall hydrolases, which are involved in cell wall turnover during cell growth, cell division, sporulation and germination (Babe and Schmidt, 1998; Blackman et al., 1998; Foster, 1993; Murray et al., 1997; Popham et al., 1996; Tjalsma et al., 2000). To prevent the loss of these proteins, due to putative release from the membrane upon signal peptide cleavage, various retention mechanisms and/or signals are employed by the cell, such as (additional) transmembrane domains, lipid modification, pseudopilin assembly, cell wall-binding repeats, and/or covalent attachment to the cell wall. For an overview see Tjalsma and co-workers (2000).

4. Pre-translocational accessory components
In general, proteins that are destined to be translocated in a more or less unfolded conformation are recognised and bound by cytosolic factors. Such factors, called chaperones, keep the precursors of secretory proteins in a translocation-competent state by preventing their folding and aggregation. Some of these chaperones assist in the targeting to the translocase. A large subset of exported proteins of E. coli utilise the chaperone SecB for their targeting to the membrane, after which they are transported in a post-translational manner (Kumamoto and Francetic, 1993; Watanabe and Blobel, 1989). In B. subtilis a putative SecB analog has been identified, called CsaA. CsaA was shown to be a suppressor of an E. coli SecA temperature-sensitive mutant. Moreover, CsaA was able to partly complement SecB deficiency in an E. coli mutant strain (Müller et al., 1992). CsaA might play a role in protein targeting of precursors to the B. subtilis Sec-translocase, since CsaA has an affinity for SecA and several precursors like YvaY (Kawaguchi et al., 2001; Müller et al., 2000a; 2000b). Although CsaA is essential for life, this putative chaperone appears to be of less importance for protein secretion in B. subtilis than
its analog, SecB, in *E. coli*.

In *B. subtilis*, the signal recognition particle (SRP) is the only secretion-specific chaperone identified so far. SRP is a ribonucleoprotein complex consisting of the Ffh protein assembled onto a 7S RNA scaffold. The histone-like protein HbsU is most likely also part of this complex. However, chaperones with a more general function in protein folding/unfolding might also assist in protein targeting and/or translocation. In *E. coli*, it was shown that GroEL, GroES, DnaK, DnaJ, and GrpE are involved in the secretion of SecB-independent precursors (Bochkareva et al., 1988; Kusukawa et al., 1989; Wild et al., 1992; 1996). In addition, GroEL was shown to interact with SecA and GroEL could even modulate a cycle of association of SecA with the membrane by triggering the release of SecA from the membrane (Bochkareva et al., 1989). Thus far, a role in protein secretion has not been demonstrated for these general chaperones in *B. subtilis*, but GroEL-ES and DnaK-DnaJ-Grp (Wolfgang Schumann and Thomas Wiegert, personal communication) may facilitate secretion by preventing proteins from adopting a translocation-incompatible conformation in the cytoplasm (Fekkes et al., 1999). For *E. coli* it has been shown that the chaperone Trigger Factor (TF) interacts with both secretory and cytosolic proteins (Hesterkamp et al., 1996; Valent et al., 1995; 1997). In contrast to the situation in *B. subtilis*, TF is an essential component for nascent polypeptide folding in *E. coli* (Reyes and Yoshikawa, 2002). Nevertheless, cross-linking studies described in Chapter 4 show that TF of *B. subtilis* does interact with the signal peptide of α-amylase and seems to compete with SRP. This suggests that in *B. subtilis* TF is also involved in pre-protein targeting to the membrane.

5. Protein translocation in *Bacillus subtilis*

Ribosomally synthesised proteins can be transported to various extracytoplasmic destinations, depending on the presence of a signal peptide and/or a specific retention signal. Proteins lacking such signals remain in the cytoplasm. Within the *B. subtilis* cell, four independent export pathways have been recognised; the Sec pathway, the Twin-Arginine Translocation pathway (Tat), the ATP-binding cassette (ABC) transporters and the Com system (see also Fig. 1). Most precursors are translocated across the cytoplasmic membrane in an unfolded state via the Sec pathway (Driessen et al., 2001).

a) Sec-dependent protein translocation

When mRNA is translated into a protein, the signal peptide exits first from the ribosomes. In the cytoplasm different chaperones can bind to this signal peptide depending on the features of the signal peptide (see section 4). On the basis of proteomic studies, it has been proposed that the majority of secretory proteins of *B. subtilis* are targeted to the Sec translocase by the signal recognition particle (SRP)
Therefore, it is of high importance to understand SRP pathway functioning in order to allow exploitation of this pathway for industrial use. In the following sections the current knowledge of the Sec-translocase and the SRP-dependent pathway are discussed.

I. Signal Recognition Particle Pathway. In prokaryotes and eukaryotes, the transport of secretory and membrane proteins to the plasma membrane or the endoplasmic reticulum, respectively, is mediated by the signal recognition particle (SRP; Keenan et al., 2001; Luiirink and Sinning, 2004). At present, the model for SRP-dependent translocation of B. subtilis is based on studies that addressed SRP-dependent translocation in E. coli. When a hydrophobic N-terminal signal peptide emerges from the ribosome, the E. coli SRP binds to this signal peptide and forms a so-called SRP/RNC (ribosome nascent chain) complex. This complex interacts with the membrane-bound SRP Receptor (FtsY) allowing the delivery of the RNC to the translocation channel in the membrane. This leads to a dissociation of the SRP/FtsY complex and the components can take part in another round of translocation. In B. subtilis, the SRP receptor-like protein FtsY is most probably involved in preprotein targeting by the SRP complex (Ogura et al., 1995). For E. coli FtsY has been shown to interact with the SecYEG translocon (Angelini et al., 2005).

SRP is a protein–RNA complex present in all organisms but with different composition. The B. subtilis SRP complex consists of the Ffh (Fifty-four homolog) protein (Honda et al., 1993), a small cytoplasmic RNA (scRNA), and a histone-like protein (HBsu) (Nakamura et al., 1992; 1994; 1999). The most conserved part of the SRP RNA, the domain IV (helix 8 in eukaryotes), binds to Ffh (Bernstein et al., 1989; Römisch et al., 1989). HBsu is a DNA binding protein (Klein and Marahiel, 2002) that is essential for normal growth (Nakamura et al., 1999). HBsu can bind to the Alu domain of scRNA, as shown by in vitro experiments (Nakamura et al., 1999).

Both B. subtilis Ffh and FtsY are essential for viability of the cell and belong to the widely conserved family of SRP-GTPases (Eichler and Moll, 2001; Chapter 2; Chapter 3). The interaction with ribosomes, the presence of the translocation channel, as well as the binding and release of the signal peptide, induce a reaction and conformational change of these SRP-GTPases (Luirink and Sinning, 2004). Ffh consists of three domains; the N domain (α-helical domain), the G domain (nucleotide-binding domain), and the M domain (methionine-rich domain, responsible for RNA interaction and signal peptide binding) (see Fig. 3). FtsY consists of the N and G domains that are homologous to the N and G domains of Ffh. E. coli FtsY contains also a third domain at the N terminus, an acidic A domain, which is homologous to the A domain of SRα in the mammalian SRP receptor. Notably, however, FtsY of B. subtilis lacks this so-called A domain.
Although *B. subtilis* Ffh and FtsY are indispensable for protein secretion, growth and viability (Hirose *et al*., 2000; Kobayashi *et al*., 2003), cells can be depleted of these proteins to levels below detection without causing an absolute secretion block (Chapter 2). However, under these conditions the secretion of a specific subset of secretory proteins is apparently reduced. Remarkably, the production of secretory proteins impacts on the cellular Ffh and FtsY levels in *B. subtilis*. Although the mechanistic details remain to be elucidated, it seems most likely that increased cellular levels of Ffh and FtsY upon high-level secretory protein production is due to a post-transcriptional regulatory event. This could, for

---

**Figure 3. Conserved domains in proteins of the SRP-GTPase family.**

The SRP-GTPase family members of yeast (SRP54, SRα), *E. coli* (P48, FtsY_Ec), and *B. subtilis* (Ffh, FlhF, FtsY_Bs) are represented schematically. Different domains that can be distinguished are the acidic “A” domain; the basic “B” domain, the conserved N-domain, the M-domain involved in RNA and pre-protein binding; and the GTP-binding “G” domain. The five conserved boxes G1 – 5 in the G-domain, as defined by Eichler and Moll (2001), are shown.

Although *B. subtilis* Ffh and FtsY are indispensable for protein secretion, growth and viability (Hirose *et al*., 2000; Kobayashi *et al*., 2003), cells can be depleted of these proteins to levels below detection without causing an absolute secretion block (Chapter 2). However, under these conditions the secretion of a specific subset of secretory proteins is apparently reduced. Remarkably, the production of secretory proteins impacts on the cellular Ffh and FtsY levels in *B. subtilis*. Although the mechanistic details remain to be elucidated, it seems most likely that increased cellular levels of Ffh and FtsY upon high-level secretory protein production is due to a post-transcriptional regulatory event. This could, for
example, be due to an as yet hypothetical protection of Ffh and FtsY from degradation through the formation of SRP complexes with overproduced secretory precursors. However, alternative mechanism can presently not be excluded (Chapter 2).

Interestingly, *B. subtilis* and several other bacterial species (but not *E. coli*) contain a third gene encoding a protein belonging to the SRP-GTPase family. In *B. subtilis*, this paralogue of Ffh and FtsY was named FlhF (flagellum-associated protein). FlhF is not essential for viability of the cell (Carpenter *et al.*, 1992). Specifically, the *B. subtilis* FlhF protein has 46% identical residues and conservative replacements in a stretch of 175 residues with *B. subtilis* Ffh and 37% identical residues and conservative replacements in a stretch of 318 residues with *B. subtilis* FtsY. As shown by sequence alignments and domain searches, FlhF contains the conserved N and G domains of the SRP-like GTPases (Fig. 3). However, it lacks the so-called M domain typical for the C termini of Ffh-like proteins. At the N-terminus FlhF contains a basic B domain instead of the acidic A domain of FtsY-like proteins of bacteria and yeasts. Notably, the mammalian SRP receptor SRα contains a more basic N-terminal domain, like FlhF of *B. subtilis*. On the basis of this similarity between FlhF and Ffh/FtsY, Carpenter *et al.* (1992) proposed that FlhF might be involved in protein secretion. However, by comparing the extracellular proteomes of *flhF* mutant strains with that of the parental strain, it is shown that FlhF is dispensable for protein secretion, even under conditions of overexpression of a secretory protein of *Bacillus amyloliquefaciens* AmyQ (Chapter 3).

In *E. coli*, the effectiveness of SRP recognition is decreased by a reduction of either the net positive charge or the hydrophobicity of certain signal peptides. Importantly, a high degree of H-domain hydrophobicity can compensate for the loss of basic residues in the N-domain, thereby restoring SRP binding (Peterson *et al.*, 2003). For *B. subtilis* cross-links between the Nascent Chain and Ffh could only be demonstrated for a very hydrophobic signal peptide (Chapter 4).

**II. Sec Translocase.** The pre-protein translocation machinery of the *B. subtilis* Sec pathway consists of at least four proteins: SecA, which is the translocation motor, and the integral membrane proteins SecE, SecG and SecY (Fig. 4A). In the current model for pre-protein translocation in *B. subtilis*, which has many similarities with that of *E. coli*, several successive steps in the translocation of secretory proteins occur (Economou, 2002; Fekkes and Driessen, 1999; Mori and Ito, 2001; Tjalsma *et al.*, 2000; van Wely *et al.*, 2001). First, SecA binds to the SecYEG translocase in the cytoplasmic membrane. Next, pre-proteins are transferred from a targeting factor (e.g. SRP; CsaA) to SecA dimers that are bound to the SecYEG complex. The binding of ATP by SecA leads to insertion of the C-terminus of SecA into the pore that is formed by the SecYEG complex in the...
membrane, resulting in the translocation of a short stretch of the pre-protein. Next, ATP is hydrolyzed by SecA, leading to the release of the pre-protein and de-insertion of SecA. Further translocation is driven by both repeated cycling of SecA through ATP binding and hydrolysis, and the proton-motive force.

In addition to the heterotrimeric SecYEG sub-complex, the E. coli Sec machinery contains a second heterotrimeric sub-complex that is composed of the SecD, SecF, and YajC proteins. A second sub-complex is likely to form a part of the B. subtilis Sec machinery as well, although this has not been demonstrated experimentally. In B. subtilis, this complex would be composed of the SecDF “siamese twin” (a natural fusion protein of SecD and SecF; Bolhuis et al., 1998), and YrbF (the homologue of E. coli YajC). The precise role of SecDF-YajC in protein export is presently not clear, but a variety of possible functions have been proposed. These include: 1) the removal of cleaved signal peptides or transmembrane segments from the SecYEG translocation channel; 2) the release of translocated proteins from the translocation channel; 3) regulation of SecA cycling; and 4) the prevention of pre-protein backsliding (Nouwen and Driessen, 2002). Unlike SecD and SecF of E. coli, SecDF of B. subtilis 168 was shown to have little impact on cell viability and protein export, at least under standard laboratory conditions (Bolhuis et al., 1998). A secretion defect in a secDF mutant strain was only observed under conditions of high-level expression of secretory proteins, such as AmyQ of B. amyloliquefaciens.

A final component that can associate with the Sec translocase of E. coli is
the YidC protein, which is involved in the membrane insertion of newly synthesized membrane proteins (Luirink et al., 2001; Samuelson et al., 2000; Scotti et al., 2000). Interestingly, YidC seems to be linked to the SecYEG sub-complex of the translocase through the SecDF-YajC sub-complex (Nouwen and Driessen, 2002). *B. subtilis* contains two homologues of YidC, known as SpoIIIJ and YqjG. Remarkably, the biogenesis of a variety of integral membrane proteins in *B. subtilis* is only mildly affected in cells depleted of both SpoIIIJ and YqjG (Tjalsma et al., 2003). In contrast, the simultaneous depletion of SpoIIIJ and YqjG has a severe impact on, as yet undefined, post-translocational stages in the secretion of proteins, such as AmyQ, LipA, and *E. coli* PhoA (Tjalsma et al., 2003).

**b) Sec-independent translocation**

Although most proteins are translocated via the Sec-pathway, at least three other very specific pathways are active in *B. subtilis*. Proteins that already fold in the cytoplasm, for example because they have to incorporate co-factors in this cellular compartment, are translocated *via* the Tat-pathway. In addition, the Com system deals with a special class of proteins that are involved in the uptake of DNA during the development of natural competence. Next to these, some ABC transporters are involved in the signal peptide-dependent export of peptide antibiotics and pheromones.

**I. Tat-pathway.** Some proteins synthesised with a twin-arginine (RR) signal peptide can be transported *via* the Tat translocase in a folded state. The Tat pathway is presently best characterised for *E. coli*, although it was first described for thylakoids (called the ΔpH pathway in chloroplasts; Chaddock et al., 1995; Mould and Robinson, 1991; Robinson et al., 1994). Four proteins, TatA, B, C, and E, are involved in Tat-dependent protein secretion in *E. coli*. TatA and TatE proteins are paralogous and functionally interchangeable (Sargent et al., 1999). A tatAE double mutation results in a total block in the export of proteins bearing twin-arginine signal peptides. Furthermore, the TatB and TatC proteins play important roles in Tat-dependent protein secretion. Accordingly, disruption of the tatB gene or a tatC mutation lead to a block in the secretion of known Tat substrates (Bogsche et al., 1998; Sargent et al., 1999).

In contrast to most other bacteria, *B. subtilis* contains three TatA-like components: TatAd, TatAy, and TatAc. In addition, two TatC proteins, TatCd and TatCy were identified. Two proteins, PhoD and YwbN, containing a twin-arginine signal peptides were shown to be translocated in a Tat-dependent manner (Jongbloed et al., 2000; 2004). Interestingly, two minimal and independently operating Tat-machines seem to be present in *B. subtilis*. TatAd and TatCd are involved in the translocation of PhoD, under phosphate starvation conditions, whereas TatAy and TatCy are involved in the secretion of YwbN (Jongbloed et al., 2004).
II. Com system. Type IV pilin-like proteins or “pseudopilins” carry signal peptides that show similarity to Sec-type signal peptides (Lory, 1994; Pugsley, 1993, Fig. 2) and their export is dependent on the dedicated Com system. The pseudopilins of *B. subtilis* (ComGC, ComGD, ComGE, and ComGG) are necessary for DNA binding and uptake during the development of genetic competence and are localised at the membrane-cell wall interface and in the cell wall (Chung and Dubnau, 1998).

III. ABC-transporters. *B. subtilis* produces peptide antibiotics and pheromones, which are synthesised either ribosomally or non-ribosomally (Duitman *et al.*, 1999; Galli *et al.*, 1994; Nakano and Zuber, 1990). Some of these peptides contain leader peptides that are removed during their export via ABC transporters (Havarstein *et al.*, 1995; Paik *et al.*, 1998; Sahl *et al.*, 1995). In *B. subtilis* the *sunS-sunT* operon encodes for the lantibiotic sublancin 168 and the ABC transporter SunT (Dorenbos *et al.*, 2002; Paik *et al.*, 1998). The *sunT* gene encodes a protein possessing features of a dual-function ATP-binding cassette transporter with a proteolytic domain and an ATP-binding cassette. These domains are common among lantibiotic transporters (McAuliffe *et al.*, 2001).

6. Post-translocational accessory components

After translocation across the cytoplasmic membrane, the preprotein will be recognised by post-translocational accessory components, which are depicted in Figure 4B. These components play a role in post-translational processes that are important for the production of fully functional mature proteins.

a) Signal Peptidases

Signal peptidases are membrane proteins, which remove the signal peptides from secretory precursors when the C-domain of the signal peptide, containing the signal peptidase cleavage site, emerges at the extracytoplasmic side of the membrane. This enzymatic reaction is necessary for the release of the mature secretory protein from the membrane (Dalbey and von Heijne, 1992; Dalbey *et al.*, 1997; Tjalsma *et al.*, 2000). *B. subtilis* contains different types of signal peptidases as described in the following sections.

I. Type I Signal Peptidases. One of the most remarkable features of the *B. subtilis* protein secretion machinery is the presence of multiple, paralogous, type I SPases. In many other bacteria, archaea and the Endoplasmic Reticulum (ER) of yeast, just one type I SPase seems to be sufficient for the processing of secretory pre-proteins (Tjalsma *et al.*, 1998; 2000; van Roosmalen *et al.*, 2004). In *B. subtilis* five *sip* genes for type I SPases are located on the chromosome (denoted *sipS*, *sipT*, *sipU*, *sipV* and *sipW*; Tjalsma *et al.*, 1997; 1998). Interestingly, SipW is homologous
to SPases found in sporulating Gram-positive bacteria, archaea and the ER membrane of eukaryotes, which together form the sub-family of ER-type SPases. In contrast, all other \textit{B. subtilis} SPases are of the prokaryotic-type (P-type). Such P-type SPases are typically present in eubacteria, mitochondria and chloroplasts (Tjalsma \textit{et al.}, 1998). Although all chromosomally-encoded SPases in \textit{B. subtilis} can process secretory precursors, only SipS and SipT are of major importance for pre-protein processing and cell viability. In contrast, SipU, SipV and SipW have a minor role in protein secretion and have substrate specificities that differ at least partly from those of SipS and SipT (Tjalsma \textit{et al.}, 1997; 1998).

\section*{II. Type II Signal Peptidases.} Although lipoproteins are transported \textit{via} the general Sec pathway, \textit{B. subtilis} contains only one gene for a type II SPase (\textit{lspA}; Prágai \textit{et al.}, 1997, Tjalsma \textit{et al.}, 1999a), which is specifically required for the processing of lipid-modified pre-proteins. Processing occurs N-terminally of the lipid-modified cysteine residue at the +1 position of the mature protein. Interestingly, \textit{B. subtilis} cells lacking SPase II are viable under standard laboratory conditions. This indicates that processing of lipoproteins by SPase II is not strictly required for lipoprotein function, as at least the lipoprotein PrsA is essential for viability (Kontinen and Sarvas, 1993). As shown for \textit{Lactococcus lactis}, uncleaved lipoproteins might also be active (Venema \textit{et al.}, 2003). In \textit{B. subtilis} cells lacking SPase II, lipoprotein precursors are subject to alternative N-terminal processing by, as yet, unidentified proteases (Tjalsma \textit{et al.}, 1999a; 1999b). However, the cumulative activity of unprocessed and alternatively processed (mature-like) lipoproteins is in many cases strongly reduced compared to their corresponding mature form (Bengtsson \textit{et al.}, 1999; Tjalsma \textit{et al.}, 2003).

\subsection*{b) Folding Catalysts}
After the translocation in an unfolded state through the Sec channel, Sec-dependent secretory proteins have to fold into their native conformation. Unfolded proteins are highly susceptible to proteolysis by one of several proteases associated with the cell wall of \textit{B. subtilis}. Therefore, fast and accurate folding of the translocated protein is crucial (Sarvas \textit{et al.}, 2004; Tjalsma \textit{et al.}, 2000; van Dijl \textit{et al.}, 2001; van Wely \textit{et al.}, 2001). Even though proteins can fold spontaneously \textit{in vitro}, their folding \textit{in vivo} is frequently assisted by folding catalysts.

An important folding catalyst involved in protein secretion is the lipoprotein PrsA, which shows homology to peptidyl-prolyl \textit{cis/trans} isomerases (Rudd \textit{et al.}, 1995) and is essential for protein secretion and cell viability of \textit{B. subtilis} (Kontinen and Sarvas, 1993; Kontinen \textit{et al.}, 1991). Strains containing mutant forms of PrsA show impaired secretion of degradative enzymes (Jacobs \textit{et al.}, 1993; Kontinen and Sarvas, 1993). It has been suggested that PrsA is required
for the prevention of unproductive interactions of unfolded secretory proteins with the cell wall shortly after translocation (Wahlström et al., 2003). A similar role in post-translocational protein folding was recently postulated for SpoIIIJ and YqiG, as depletion of both of these proteins affected the stability of at least three secretory proteins during the post-translocational stage in protein secretion (Tjalsma et al., 2003). However, the action of SpoIIIJ and YqiG in protein folding is likely to be indirect in view of the well-documented role of the homologues of these proteins (e.g. YidC of *E. coli*) in membrane protein assembly. The importance of extracytoplasmic folding catalysts is underscored by the fact that the membrane-cell wall interface and extracellular environment of *B. subtilis* are highly proteolytic (Tjalsma et al., 2000; see the section on Quality Control Factors). This results in the rapid degradation of exported proteins of homologous or heterologous origin that fold too slowly, or incorrectly, after translocation (Bolhuis et al., 1999c; 1999d).

Immediately after translocation of a subset of cysteine containing secretory proteins, disulphide bridge formation will occur. In *B. subtilis*, only a limited number of proteins contain disulphide bridges in their native conformation, like ComGC and ComGG. Disulphide bridges in these proteins are important for their activity and stability. To accomplish disulphide bridge formation immediately after translocation, thiol-disulphide oxidoreductases are located in close proximity to the membrane. Bolhuis and co-workers described three thiol-disulphide oxidoreductases in *B. subtilis*, denoted BdbA, BdbB, and BdbC with (limited) similarity with known thiol-disulphide oxidoreductases (Bolhuis et al., 1999). A fourth thiol-disulphide oxidoreductase, BdbD, is highly similar to DsbA of *Staphylococcus aureus* and was discovered and described by Meima et al. (Meima et al., 2001).

Because of limited availability of native substrates for *B. subtilis* thiol-disulphide oxidoreductases, PhoA of *E. coli* was used as a model protein. This protein contains two disulphide bridges in its native conformation and can be produced by *B. subtilis*. For the correct folding and activity of *E. coli* PhoA, the thiol-disulphide oxidoreductases BdbC and BdbD are important (Bolhuis et al., 1999; Meima et al., 2002). In addition, the absence of BdbC and BdbD was shown to have a strong impact on competence development, most likely because they act as a redox pair catalyzing disulphide bond formation in the ComGC protein (Meima et al., 2002). BdbA does not play a major role in disulphide bridge formation, since disruption of the *bdbA* gene did not significantly affect the secretion of active *E. coli* PhoA (Bolhuis et al., 1999).

c) Extracellular Proteases

*B. subtilis* produces at least 27 proteases that end up in the membrane, cell wall, or growth medium. These proteases can cleave (partially) unfolded polypeptide chains, which underscores the importance of efficient folding of secretory proteins.
into their native, protease-resistant conformation (Tjalsma et al., 2000). Secretory proteins that are not correctly folded by membrane or cell wall-attached folding catalysts, described in the previous section, are potential substrates for one of these many proteases. This can be concluded from the observation that extracellular and cell wall-associated proteases are responsible for the degradation of various heterologous proteins secreted by *B. subtilis* (Bolhuis et al., 1999c, Wu et al., 1991). However, it should be kept in mind that even correctly folded heterologous proteins can be subject to degradation.

To investigate the impact of secreted proteases, the extracellular proteome of the *B. subtilis* mutant strains WB700 (Ye et al., 1999) and WB800 (Wu et al., 2002), lacking the genes for seven or eight extracellular proteases, respectively, was investigated by two-dimensional (2D) gel electrophoresis (Antelmann et al., 2002; 2003; Chapter 5 Appendix). As expected, the proteases AprE, Bpr, NprE, Mpr, and Vpr were lacking from the extracellular proteome of these protease mutant strains in comparison to the parental strain. In the WB800 strain, the WprA protease is also absent from the extracellular proteome due to the integration of a *hyg* cassette in the corresponding gene (Wu et al., 2002). Importantly, the levels of most extracellular proteins were not affected by the mutations or deletions of these proteases, which is consistent with the idea that homologous secretory proteins must be largely resistant to the activity of extracellular proteases of *B. subtilis*.

### 7. Quality control factors and secretion stress

The production of high levels of secretory proteins by *Bacillus* species requires quality control mechanisms that are active during the secretion process. Proteins with a known role in the quality control of secretory proteins are the two HtrA-like proteases/chaperones, HtrA and HtrB (Noone et al., 2001). These proteins are believed to have proofreading capabilities for the folding-state of secretory proteins, as demonstrated for HtrA of *E. coli* (Spiess et al., 1999). When a secretory protein is not properly folded, HtrA and/or HtrB can either assist in its folding or degrade the malfolded secretory protein.

Another protease, WprA, seems also to be involved in the quality control of secretory proteins. It has been proposed that, in addition to a protease domain (CWBP52), this protein also contains a chaperone domain (CWBP23) (Babe and Schmidt, 1998; Margot and Karamata, 1996). Notably, HtrA and WprA have a dual localization, being present both in the cell wall proteome and the extracellular proteome (Antelmann et al., 2002; 2003). In contrast to WprA, which is synthesised with an N-terminal signal peptide, the HtrA protein has an N-terminal membrane anchor. As shown by N-terminal sequencing, the HtrA variant that was found in the growth medium lacks this membrane anchor domain (Antelmann et al., 2000). So far, the localization of HtrB, which has a predicted N-terminal membrane anchor like HtrA, has not been documented.
Interestingly, high-level protein secretion induces the transcription of the *htrA* and *htrB* genes, which is controlled by the CssR-CssS two-component regulatory system (Fig. 5). All stimuli that trigger a CssRS-dependent cellular response are presently defined as secretion stress (Antelmann *et al.*, 2003; Chapter 5; Chapter 5 Appendix). This two-component regulatory system consists of the CssS protein, a sensor histidine kinase, and the CssR protein, a cognate response regulator. Most likely, CssS senses secretion and heat stress at the membrane-cell wall interface by stimuli that are directly or indirectly caused by malfolded proteins (Hyyrylainen *et al.*, 2001). As a result, CssS is believed to autophosphorylate itself, after which the phosphate group is transferred to CssR. Next, the transcription of specific genes like *htrA* and *htrB* are regulated by CssR. Notably, the disruption of either *htrA* or *htrB* causes aCssRS-dependent secretion stress response that strongly induces the activity of the promoters of both genes, and this response is mediated by the CssRS two-component system (Darmon *et al.*, 2002; Hyyryläinen *et al.*, 2001; Noone *et al.*, 2001). Furthermore, the transcription of both *htrA* and *htrB* can be strongly reduced by a disruption of the *cssS* gene. To monitor the
impact of HtrA and HtrB on protein secretion, the extracellular proteomes of htrA and/or htrB mutant strains have been analyzed by 2D gel electrophoresis (Antelmann et al., 2003). The spot corresponding with the HtrA in the medium was increased upon deletion of the htrB gene.

Under conditions of severe secretion stress due to overproduction of, for example, AmyQ, this two-component regulatory system is essential for cell viability (Darmon et al., 2002; Hyyrylainen et al., 2001). The level of htrB-lacZ expression correlates with the levels of AmyQ production (Chapter 5). Interestingly, high-level production of a non-translocated variant of AmyQ did not provoke a secretion stress response, whereas the high-level production of human interleukin3 (IL3) and the homologous lipaseA (LipA) did trigger a CssRS-dependent stress response (Chapter 5 Appendix). Accordingly, this CssRS-dependent response can be regarded as a general secretion stress response in B. subtilis.

8. Scope of this thesis

The major aims of the work described in this thesis were (1) to understand the role of the Signal Recognition Particle (SRP) pathway, the major secretion pathway in B. subtilis, and (2) to describe the secretion stress response as the result of high-level protein production, in order to further optimize B. subtilis as a cell factory. The industrial application of B. subtilis might be extended further when bottlenecks in the secretion and production of heterologous proteins can be overcome. For this purpose, the roles of the three SRP-GTPases of B. subtilis, Ffh, FtsY, and FlhF in protein secretion, as well as the role of signal peptide recognition, were studied. Eventually, exploiting this knowledge might contribute to the use and optimization of (components of) the SRP-dependent pathway for the production of commercially and/or pharmaceutically interesting proteins. Moreover, making use of the secretion stress response of B. subtilis for monitoring protein production levels might also help to improve protein production processes.

Ffh, the major player of SRP, is essential for the cell. To obtain more insight in the SRP pathway, the regulation of Ffh was studied (Chapter 2). FtsY, the proposed receptor for SRP-RNC complexes, is also essential for the cell and seems to be partly involved in the secretion of proteins (Chapter 2). To investigate whether the third SRP-GTPase, FlhF, which is not essential for the cell, is involved in protein secretion, its role in this process was studied (Chapter 3). To exploit the SRP pathway of B. subtilis for efficient heterologous protein secretion, it was necessary to get more insights in the way SRP recognizes signal peptides. Since the hydrophobic domain is proposed to be the dominant structure in signal peptide function, the importance of the level of hydrophobicity of the H-domain of signal peptides for recognition by the B. subtilis SRP complex was studied in Chapter 4. Furthermore, the provoked secretion stress response in B. subtilis at high-level protein production was studied in Chapter 5. Its potential use as a monitor for the
protein production levels was investigated. However, since the same secretion stress response can be a bottleneck in the production of (heterologous) proteins in *B. subtilis*, the secretion stress response was also further characterised (Chapter 5 Appendix).

Taken together, in this thesis the study on early stages in protein secretion of *B. subtilis*, like the roles of the SRP-GTPases and the hydrophobicity of signal peptides, are described. In addition, the late stages in protein secretion in *B. subtilis*, in particular the general secretion stress response, are presented. In the near future, the results of these studies might be used to exploit *B. subtilis* as a cell factory for the production of heterologous proteins.