A genetic analysis of determinants for efficient protein secretion in Bacillus subtillis

Bolhuis, Albert

CHAPTER 9

Summary and General Discussion

In all living organisms proteins are transported across membranes to reach their final destination and fulfil their function. The transport mechanisms for these proteins, which are synthesised as precursor proteins with an amino-terminal signal peptide, involves complex machineries. In the Gram-positive bacterium *Bacillus subtilis*, the translocation machinery in the cytoplasmic membrane consists of the SecA protein, the force generator, and a translocation channel, formed by the membrane proteins SecY, SecE, SecG, SecDF, and (possibly) YrbF. Several other factors are required for efficient translocation across the membrane. Firstly, targeting of preproteins to the translocase is aided by soluble cytosolic factors, called chaperones. Secondly, during, or shortly after the translocation, the signal peptide is removed by signal peptidases (SPases). Thirdly, a number of factors assist in the folding of the secretory protein after its release from the translocase. At the different stages of protein secretion several proteases are present that play a role in the quality control of the process by degradation of, for instance, misfolded proteins.

B. subtilis is an attractive host for the commercial production of secretory proteins. Firstly, this bacterium is able to secrete proteins directly into the growth medium, which greatly facilitates their downstream processing. Secondly, this organism has a huge capacity for protein secretion. Thirdly, *B. subtilis* is a genetically highly amenable host organism, for which a large variety of genetic tools are available, and efficient fermentation technology has been developed. Fourthly, the complete sequence of the *B. subtilis* genome is known. Finally, *B. subtilis* is non-pathogenic, and free of endotoxins. Notwithstanding these advantages, the secretion of various heterologous proteins by *B. subtilis*, in particular proteins of eukaryotic origin, is frequently inefficient, which limits the application potential of this organism. Various bottlenecks for protein secretion in *B. subtilis* have been identified in recent years, such as poor targeting to the translocase, low rate of SPase processing, slow or incorrect folding, poor release from the cell wall, and sensitivity to proteolytic degradation of secretory proteins.

The goal of the research described in this thesis was to analyse determinants that are required for efficient protein secretion in *B. subtilis*. Factors that determine the efficiency of protein secretion represent potential bottlenecks. These factors are basically present at any stage of the protein secretion process. Therefore, several aspects of the secretion pathway, ranging from gene expression in the cell to protein degradation outside the cell, were studied.

In chapter 2, the analysis of bottlenecks in the secretion of four heterologous proteins is described. The proteins examined were *Bacillus amyloliquefaciens* α-amylase (AmyL), *Escherichia coli* TEM-β-lactamase (Bla), human pancreatic α-amylase (HPA), and a lysozyme-specific single-chain antibody (SCA-Lys). All identified bottlenecks related to late stages in the secretion pathway, following translocation across the cytoplasmic membrane. These included processing by signal peptidase, passage through the cell wall,

and degradation in the wall and growth medium. Strikingly, all translocated HPA was misfolded, its stability depending on the formation of disulphide bonds. This suggests that the thiol-disulphide oxido-reductases of *B. subtilis* can not form the disulphide bonds in HPA correctly. As the secretion bottlenecks differed for each heterologous protein tested, it was concluded that the efficient secretion of particular heterologous proteins requires the engineering of specifically optimised host strains.

In chapter 3, the identification and characterisation of the SecDF protein of *B. subtilis* is described. Surprisingly, the *secDF* gene represents a natural gene fusion, which is in contrast to most known *secD* and *secF* genes from other organisms, where these genes are separate. SecDF is a subunit of the preprotein translocase complex of *B. subtilis* that was shown to be required for maintaining a high capacity for protein secretion. Unlike the SecD subunit of the pre-protein translocase of *E. coli*, SecDF of *B. subtilis* was not required for the release of a mature secretory protein from the membrane, indicating that SecDF is involved in early translocation steps. Strains lacking intact SecDF showed a cold-sensitive phenotype, which was exacerbated by high-level production of secretory proteins, indicating that protein translocation in *B. subtilis* is intrinsically cold-sensitive. Comparison with SecD and SecF proteins from other organisms revealed the presence of ten conserved regions in SecDF, some of which appear to be important for SecDF function. Interestingly, the SecDF protein of *B. subtilis* has twelve putative transmembrane domains. Thus, SecDF does not only show sequence similarity, but also structural similarity to secondary solute transporters. These data suggest that SecDF of *B. subtilis* represents a novel type of the SecD and SecF proteins, which seems to be present in at least two other organisms.

In chapter 4, analyses on the biological function of the type I SPase SipS and the regulation of its synthesis are described. Unlike the type I SPase of *E. coli*, SipS was neither essential for protein secretion, nor for viability of the cell. However, in the absence of SipS, the rate of processing of several pre-proteins was reduced, and four of the seven major secreted proteins of *B. subtilis* were hardly detectable in the growth medium. Surprisingly, the processing of *B. amyloliquefaciens* α-amylase, and the secretion of at least two endogenous *B. subtilis* proteins was improved in the absence of SipS. These findings indicate that the substrate preference of SipS differs from that of the other SPases from *B. subtilis*, and that SipS is an important factor in determining the efficiency of protein secretion in *B. subtilis*. The *sipS* gene is transcribed in a growth phase- and medium-dependent manner. In minimal medium, the growth phase-dependent transcription of *sipS* is controlled by the DegS-DegU two-component regulatory system, indicating that the expression of *sipS* is regulated by the same factors that control the expression of most genes for secreted degradative enzymes. These observations suggest that *B. subtilis* can modulate its capacity and specificity for protein secretion through the controlled expression of *sipS*.

After cleavage by SPase, signal peptides are degraded in the membrane and cytoplasm. As described in chapter 5, *B. subtilis* contains two proteins, denoted SppA and TepA, with similarity to the signal peptide peptidase A of *E. coli*. SppA of *B. subtilis* was only required for efficient processing of pre-proteins under conditions of hyper-secretion. This suggests that protein translocation was inhibited by the accumulation of signal peptides. TepA depletion had a much stronger effect on protein secretion, which, in contrast to SppA

depletion, was not limited to conditions of hyper-secretion. In contrast to SppA, which is a typical membrane protein, TepA appeared to have a cytosolic localisation. This is consistent with the observation that TepA is involved in early stages of the secretion process. Interestingly, TepA also shows sequence similarity to the cytosolic ClpP protease. Thus, TepA is possibly a novel determinant for protein secretion in *B. subtilis* that is either involved in regulation of the secretion process, or the cytosolic degradation of signal peptide fragments that are inhibitory to protein translocation. Alternatively, TepA might act as a secretion-specific chaperone.

In chapter 6, analyses are described concerning the thermal inactivation of five temperature-sensitive SipS mutants. These mutants were used to assay proteolysis immediately after protein translocation. The results indicated that two of these mutants, L74A and Y81A, were structurally stable but strongly impaired in catalytic activity at 48°C, demonstrating the involvement of the conserved leucine 74 and tyrosine 81 residues in the catalytic reaction of type I SPases. In contrast, the SipS mutant proteins R84A, R84H, and D146A were inactivated by proteolytic degradation, indicating that the conserved arginine 84 and aspartic acid 146 residues are required for obtaining a protease-resistant conformation. The cell wall-bound protease WprA was shown to be involved in the degradation of SipS D146A, which is in accord with the fact that SipS has a large extracytoplasmic domain. As WprA was not involved in the degradation of the SipS mutant proteins R84A and R84H, we conclude that multiple proteases are responsible for the thermal inactivation of temperature-sensitive SipS mutants.

The formation of disulphide bonds, which is of major importance for the stability and/or activity of many proteins, is catalysed by thiol-disulphide oxido-reductases. In chapter 7 experiments are described concerning the identification of three *B. subtilis* genes, denoted *bdbA*, *bdbB* and *bdbC*, for putative thiol-disulphide oxido-reductases. The BdbB and BdbC proteins both show sequence similarity to the membrane protein DsbB of *Escherichia coli*. *E. coli* alkaline phosphatase, containing two disulphide bonds, was unstable when secreted by *B. subtilis* cells lacking BdbB or BdbC. Cells lacking BdbC also showed decreased stability of cell- associated forms of *E. coli* TEM-β-lactamase, which contains one disulphide bond. BdbB and BdbC appeared to be integral membrane proteins, suggesting that they promote the folding of disulphide bond-containing proteins at the membrane-cell wall interface. Interestingly, the rate of pre-β-lactamase processing was stimulated in cells lacking BdbC, indicating that the unfolded form of this precursor is a preferred substrate for signal peptidase. Surprisingly, cells lacking BdbC were unable to develop competence for DNA uptake, indicating the involvement of disulphide bond-containing proteins in this process. The BdbA protein was shown not to be required for the stability of alkaline phosphatase or β-lactamase, and may be dedicated to the formation of disulphide bonds in a specific (subset) of protein(s). Unlike *E. coli* and yeast, none of the putative thiol-disulphide oxido-reductases of *B. subtilis* were required for growth in the presence of reducing agents.

In chapter 8 the identification and characterisation of a protein, denoted MlpA, for **M**PP-**l**ike **p**rotein, are described. MlpA and its homologues found in mycobacteria and *Rickettsia prowazekii* are closely related to the mitochondrial processing peptidases (MPP), which remove targeting signals from imported proteins in the mitochondrial matrix. MlpA of *B.*

subtilis was shown to be involved in transcriptional regulation of the *aprE* gene for subtilisin, suggesting that MPP proteins have evolved from eubacterial regulators of gene expression.

In summary, the results presented in this thesis show that several factors in different stages of the secretion pathway determine the efficiency of protein secretion by *B. subtilis*. Several gaps in our knowledge on the secretion pathway of *B. subtilis* were filled in. A number of factors described in this thesis were not identified or characterised before, such as SecDF, SppA, TepA, and BdbB/C. The function of some of these factors is still not fully clear. For example, it remains to be investigated whether the membrane protein SecDF indeed functions as a transporter of, for instance, signal peptides or misfolded proteins (chapter 3). Furthermore, the function of the cytoplasmically located TepA protein is not clear (chapter 5). The research described in this thesis was performed mainly by genetic approaches. To be able to understand the function of all the factors in the secretion machinery in more detail, the genetics should be complemented with biochemical approaches.

With our increasing knowledge of the secretion apparatus, *B. subtilis* host strains can be optimised for the secretion of a wide range of (heterologous) proteins at commercially interesting levels. As shown in chapter 2, depending on the heterologous protein, different bottlenecks can be encountered. Some of these can be overcome rather easily. For example, limited availability of SPase activity can be solved by overproduction of SPases (chapters 2 and 4). Such improvements do not necessarily lead to increased levels of secretion, as still other factors can be present that limit the levels of protein accumulating in the growth medium. Other bottlenecks might be more difficult to overcome, such as the cell wall barrier that is a major block for secretion of human α-amylase (chapter 2). In this case, the major problem seems to be the formation of correct disulphide bonds. Increasing the levels of secreted human α-amylase might require a modulation of the redox-balance at the *trans* side of the membrane. This could be achieved by, for instance, modulation of the expression levels of the thiol-disulphide oxido-reductases of *B. subtilis*, site-directed mutagenesis of these proteins in order to change their redox-state, or the introduction of the human protein-disulphide isomerase (PDI).

In conclusion, the optimisation of the secretion of individual heterologous proteins requires engineered *B. subtilis* strains in which specific secretion bottlenecks have been removed. It is anticipated that the engineering of dedicated *B. subtilis* host strains for the secretion of a wide range of heterologous proteins at commercially significant levels will soon be possible, because of the increasing knowledge of the secretion apparatus, the availability of the complete genome sequence, and the genetic amenability of this organism.