A comparative study of native and heterologous enzyme production in *Bacillus subtilis*

Rita A. Leal Cruz

Centre for Bacterial Cell Biology

Institute for Cell and Molecular Biosciences

Thesis submitted for the degree of Doctor of Philosophy

Newcastle University

November 2016

Abstract

Biotechnology plays a central role in setting the course for a resource-efficient and sustainable Bioeconomy. In particular, the industrial production of enzymes using microorganisms offers an environmentally friendly alternative to many traditional chemical processes. *Bacillus subtilis* is a widely recognized microbial cell factory, able to secrete proteins in tens of grams per litre. However, this organism has developed a variety of quality control and stress response mechanisms designed to facilitate efficient growth and survival in a wide range of natural environments. These pose significant challenges for improving the production efficiency of an increasingly large variety of heterologous commercial proteins.

We used two industrially relevant enzymes to study and compare, in parallel, the challenges associated with the production of native and heterologous enzyme by *B. subtilis*, focusing on the later stages of protein secretion, and the potential of a synthetic translocase for optimising production strains. The enzymes showed very distinct production profiles and secretion kinetics under the same growth and gene expression conditions. We investigated the effect of these enzymes on the regulation of the main genes involved in the secretion pathway. To this end, we analysed the impact of over expressing the model enzymes on secretion stress by fusing the GFP reporter gene to the *htrA* promoter and introducing this into a group of strains carrying precise deletions in ten genes encoding extracytoplasmic proteases.

The results showed that increasing the expression of the Sec translocase proteins did not improve productivity, indicating that the number of Sec translocases it not a limiting factor for enzyme production. However, our data show that the complex regulatory architecture associated with the secretion stress and quality control mechanisms, indicate that post-translocational protein folding and proteolysis are not only inter-dependent but also influenced by the enzyme of interest, resulting in different levels of secretion stress induction.

Declaration

I certify that this thesis contains my own work, except where acknowledged, and that no part of this material has previously been submitted for a qualification at this or any other university.

Rita Cruz

Acknowledgements

Firstly, I would like to express my sincere gratitude to Professor Colin Harwood for the strong and timely support offered throughout my PhD. Professor Colin's exemplary supervision, support, impressive insight and kind leadership has made all the difference. It will provide me with an excellent role model of what a great scientist is all about.

I am also grateful to Dr. Tjeerd van Rij for actively contributing to this project and bringing with him a unique perspective while providing solid professional advice. Most importantly, Tjeerd made me feel at home in the Netherlands and at DSM, which contributed to a unique professional and social experience. Dank je wel!

I would also like to thank Professor Leendert Hamoen for the supervision, advice and criticism.

I would like to thank Christopher Sauer, my friend and project partner, with whom I shared a lab bench, a home, laughs, frustrations but mostly a great companionship throughout. Christopher collaborated with me on Chapter 4 and throughout this project, often contributing with great insight and offering practical support. Chris, it's over!

To Kevin Bos from the Hogeschool van Arnhem en Nijmegen, whom under my inexperienced supervision contributed significantly to the Biolector assays of Chapter 4. Kevin's hard work and attention to detail were incredibily helpful.

To Dr. Phillip Aldridge and Dr. Richard Daniel, my annual assessment panel at Newcastle University, a very sincere thank you for challenging me while still giving me the confidence to continue each year.

A very special thank you goes to my colleagues at the CBCB in Newcastle, an amazing place to do science and eat cake! Susanne Pohl, Henrik Strahl, Pamela Gamba, Alex Egan, Adam Lodge, Beth Lawry, Daniela Vollmer, David Webb, Gabriela Henriques, Heath Murray, Helen Glenwright, Ian Selmes, Ismatul Ismail, Jacob Bilboy, Jad Sassine, Katharina Peters, Lauren Drage, Manuel Pazos, Martin Sim, Sean Keith, Sheryl Rowland, Simon Syvertsson, Soren Nielsen, Stuart Davidson, Susan Leightley, Tom Ewen, Tom Fletcher, Victor Hernandez, Waldemar Vollmer, Wendy Smith, and many others.

A sincere thank you to all my colleagues and friends at DSM who made my stay gezellig and with whom I learned so much: Aad Vollebregt, Aloys Teunissen, Brenda Vonk, Diny van Helvert, Eric van den Berg, Esmé Fijn, Graham Whyteside, Hannes Jürgens, Hans Roubos, Janny Bakhüis, Jeffrey van Wijk, Jos van Vugt, Marco van der Weert, Marijke Dirven-Neele, Martine Spaans, Nathalie Buiting-Wiessenhaan, Noël van Peij, Rémon Boer, Rianne van der Hoek, Richard Kerkman, Sacha Dalhuijsen, Stefan Turk, Ulrike Müeller, Ulrike Obst, Wouter Kroes. A special thank you to the great colleague, scientist and friend Marcel Hillebrand.

For their unbeatable, rock solid friendship, I want to thank Ana, Nádia, Sara, David, Ruben, Filipe, Bacalhau, the Joanas, Burgal, Jad, Polly, Oscar, Jonathan, Raquel, Johannes, Yoshi, Sanjay and Annie.

My heart is filled with love and gratitude towards my amazing parents and sister who made me almost believe, like they do, that I am the greatest scientist in the world. Without their support, I would have never been able to leave them and embark on this journey. A special thank you to my nephew Vicente for making our lives happier when he arrived to this world 6 months into my PhD. Obrigada, amo-vos muito!

Finally, I can’t find any words to express my gratitude to my beloved partner Diego. You are truly amazing and we did this together, amo-te.

Contents

List of Figures

List of Tables

List of abbreviations

°C Degree Celsius

aa Amino acids

ABC ATP-binding cassette

ADP Adenosine diphosphate

amp Ampicillin resistance

ATP Adenosine triphosphate

AU Arbitrary units

bleo Bleomycin resistance

bp Base pairs, can be prefixed kilo- (k) or mega- (m)

cat Chloramphenicol acetyltransferase, confers resistance for chloramphenicol

cDNA Complimentary DNA

CpO Codon pair optimised gene

DNA Deoxyribonucleic Acid

dNTPs Deoxyribonucleotide mix of bases A, T, G, C

DO Dissolved oxygen

em Erythromycin resistance

Ffh Fifty-four homologue

g Gram. Can be combined with metric SI prefixes

g G-force

GDP Guanosine diphosphate

GFP Green fluorescent protein

GRAS Generally regarded as safe

GTP Guanosine triphosphate

h hour

hEGF Human Epidermal Growth Factor

IFN Interferon

IL Interleukin

IPTG Isopropyl β-D-1-thiogalactopyranoside

kDa Kilo Dalton; atomic mass unit for describing protein mass

L Litre. Can be combined with metric SI prefixes

LB Nutrient Luria-Bertani medium

m Metre. Can be combined with metric SI prefixes

M Molar. Can be combined with metric SI prefixes

min Minute

mRNA Messenger RNA

neo Neomycin resistance

OD_{600} Optical density at 600 nm

PCR Polymerase chain reaction

PI Proinsulin

qPCR Quantitative PCR

RBS Ribosome binding site

RFU Relative fluorescence units

RNA Ribonucleic acid

rpm Rounds per minute

RT Reverse transcription

RT Room temperature

s Second

scFV Single-chain variable fragment

SD Standard deviation

SDS Sodium dodecyl sulfate

SEM Standar error of the mean

SM Spent medium

SP Signal peptide

spec Spectinomycin resistance

SRP Signal Recognition Particle

TAT Twin-arginine translocation

v/v Volume per volume

w/v Weight per volume

WT Wild-type

Chapter 1 Introduction

1.1. Context and motivation

Biotechnology is one of the most important industries in our modern society. It promotes the development of technology and products based on cellular and biomolecular processes contributing to a resource-efficient and sustainable economy. The European Commission recognizes biotechnology as one of the bio-based industries that contributes to the Bioeconomy. The Bioeconomy is of great importance in Europe; it has an annual turnover of around two trillion euros and employs around 22 million people and approximately 9% of the total EU workforce (European Commission 2012). Although the origins of biotechnology date back thousands of years, it is a field in which constant innovation is required to remain competitive in a rapidly developing market place.

This thesis is the product of the EU-funded European Industrial Doctorate training programme entitled ATRIEM (Advanced Training in Industrial Enzyme Manufacturing) – a collaboration between Newcastle University and DSM, a worldwide biotechnology company. ATRIEM aimed to tackle important questions relating to industrial enzyme production by combining the expertise of both entities. The production of enzymes in the EU corresponds to an annual turnover of 0.8 billion euros and to which DSM is a major contributor (European Commission 2012). The bacterium on which these studies are based, *Bacillus subtilis*, is widely used for the production of enzymes used in the baking, brewing, biofuel and beverage industries.

B. subtilis is well known for its capacity to secrete proteins at concentrations as high as tens of grams per litre. Together with its amenability to genetic manipulation, this organism has been used for decades in the protein production industry. However, the production of heterologous proteins in *B. subtilis* has encountered some difficulties related to the characteristics of the target proteins or the secretion system itself (Pohl & Harwood 2010). It is therefore of potential commercial importance to examine closely all the stages in the protein secretion pathway and to identify bottlenecks that need to be overcome in order to optimise the production of proteins during industrial *B. subtilis* fermentations.

This thesis focuses on examining fundamental research questions related to protein secretion using two major industrial model enzymes. Factors such as protein expression, processing, translocation and stability were evaluated while addressing the issue of production strain optimisation.

1.2. *Bacillus subtilis*: a bacterial model and a protein factory

Bacillus is a very diverse genus of bacteria that can be found in the soil and associated water sources such as rivers, coastal waters and estuaries (Harwood 1992). Members of the genus *Bacillus* are Gram-positive, aerobic, endospore-forming and rod-shaped (Harwood 1992).

Bacillus subtilis has been used as a biological model for many cellular studies. The interest in this bacterium was initially stimulated by the discovery, in 1959, of its ability to transfer chromosomal genes in this species (Harwood 1992). Along with its ease of cultivation and non-pathogenicity, this contributed to a profound understanding of the molecular biology of *B. subtilis* and stimulated studies on many other bacteria with similar biology.

In addition to its position as a model bacterium, *B. subtilis* and close relatives are a source of products with industrial importance such as proteases, α-amylases, and other hydrolytic enzymes used in, for example, the detergents, dairy, baking, brewing and textile industries (Freudl 1992; van Dijl & Hecker 2013). This capability is due mainly to two intrinsic characteristics of this group of bacteria:

(i) their natural habitat, the soil, has necessitated their being able to produce a wide range of hydrolytic enzymes that allows them to breakdown soil-based macromolecules into vital nutrients (Pohl & Harwood 2010).

(ii) they are able to secrete proteins into the culture medium at concentrations as high as grams per litre, due in part to the single membrane system of Gram-positive bacteria, in contrast to the double membrane system of Gram-negative bacteria.

The fact that *B. subtilis* releases secretory proteins directly into the culture medium is of great commercial interest; it reduces downstream processing costs, makes their purification more straightforward and reduces the likelihood of contamination with cytoplasmic proteins.

An efficient and effective industrial protein process relies on strains that produce high extracellular concentrations of target proteins through the optimisation of expression elements,

culture media, and growth regimes. These are normally combined with extensive strain development programmes that screen for strains with better growth, product yield and secretion characteristics. Because of their commercial sensitivity, the results of such programmes are rarely published.

1.3. The *Bacillus subtilis* protein transport pathways

B. subtilis secretes a large variety of proteins, primarily through the major and ubiquitous "Sec" preprotein secretory pathway. In addition, a small number of proteins with specific functions are secreted via the twin-arginine translocation (TAT) pathway, ATP-binding binding cassette (ABC) transporters and a pseudopilin export pathway required for competence development (Dubnau 1997; Quentin *et al.* 1999; Jongbloed *et al.* 2000). Pseudopilin-exported proteins are retained at the membrane/wall interface, which is unlikely to be of value to the heterologous enzyme production industry.

In contrast to the Sec pathway, the TAT pathway does not require proteins to be maintained in an unfolded, secretion-competent format during the early stages of the secretory pathway. Secretory proteins that follow the TAT secretion pathway are distinguished by signal peptides with a twin-arginine motif (RR), essential for recognition by the Tat translocation machinery (Berks *et al.* 2000). Gram-positive bacteria like *B. subtilis* have a minimal Tat translocase consisting of two subunits, TatA and TatC, whereas Gram-negative bacteria contain an additional subunit, TatB (Oates *et al.* 2005). Two minimal TatAC systems with distinct substrate specificities and genomic positions have been reported for *B. subtilis* and named TatAdCd and TatAyCy (Pop *et al.* 2002; Jongbloed *et al.* 2004). Interestingly, it was reported that TatAd can complement the absence of either TatA or TatB in *E. coli*, suggesting that TatA subunits are bifunctional in *B. subtilis* (Jongbloed *et al.* 2006; Barnett *et al.* 2008). Several studies indicate that the *B. subtilis* Tat machinery is capable of translocating heterologous and tightly folded proteins, however, with limited capacity (Kolkman *et al.* 2008).

From an industrial perspective, the Sec-dependent secretory pathway is the most exploited for the secretion of heterologous proteins from *B. subtilis* and its relatives. However, there are major bottlenecks associated with this pathway and the secretion of heterologous proteins is achieved with mix success. Recent reports have implied the presence of alternative secretion pathways, inspired by the presence of cytoplasmic proteins in the culture medium of *B. subtilis* during the stationary phase of growth. While conventionally these were thought to be the result of cell lysis,

more recently this is thought to be a general phenomenon in which protein domain structure is a contributing factor (Yang *et al.* 2011). Wang *et al*. (2015) showed the ability of 'non-classically secreted proteins' to act as signals to export recombinant proteins to the culture medium. *Chen et al.* (2016) identified a heterologous 'non-classically secreted protein', D-psicose 3-epimerase (RDPE), from *Ruminococcus sp*., and showed that this protein could direct various proteins into the culture medium. However, the non-classical secretion pathways are not well understood and do not yet achieve the same yields as the classical Sec-dependent secretion pathway.

1.3.1 Intracellular targeting of secretory proteins

Despite more than 20 years of intense study, it is very surprising how much there is still to be understood about the intracellular stages of the protein secretion pathway. Generally, it is known that secretory proteins require a signal peptide that targets them to the membrane-bound translocase and that is then removed during the later stages of secretion (Freudl, 1992). Secretory proteins are first synthesized as precursors, containing this cleavable amino-terminal signal peptide. When the intention is to produce proteins that are not naturally secreted, a signal sequence needs to be incorporated in-frame with the N-terminus of the target proteins. It is now generally accepted, that there is no single optimal *Bacillus* signal peptide. Instead, the signal peptide needs to be optimised for each protein to be secreted.

Sec signal peptides have characteristics that are common to all bacterial species. They are usually between 20 and 30 amino acids in length, and have a positively charged amino terminal (N) region, followed by a longer hydrophobic (H) central region and a short cleavage (C) region containing the target site for signal peptidase (Figure 1.1) (Tjalsma *et al.* 2000).

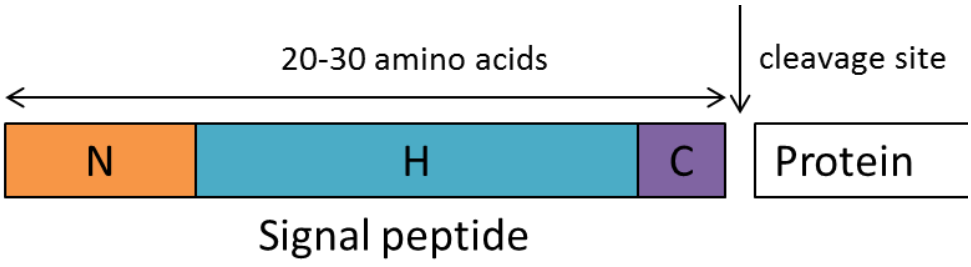


Figure 1.1 Diagrammatic representation of the signal peptide structure in bacteria. The positively charged amino terminal (N) region, the hydrophobic (H) central region and the short cleavage (C) region containing the target site for signal peptidase are represented in different colours.

Sec signal peptides are classified into two different types (Type I and Type II) according to the signal peptidase responsible for their cleaving from the targeted protein. The Type I signal peptidases cleave the majority of the secretory substrates (proteins associated with the cell wall or which are released into the culture medium). There are five Type I signal peptidases in *B. subtilis*: SipS, SipT, SipU, SipV, and SipW (Pohl & Harwood 2010). The Type II signal peptidases recognize and cleave signal peptides of lipoproteins that have to be diacylglycerol-modified at an N-terminal cysteine residue prior to their attachment to the outer surface of the membrane (Pohl & Harwood 2010).

The targeting and piloting of proteins to the Sec translocase in bacteria is poorly understood. It is known that cytosolic chaperones interact with the secretion targets to prevent them from folding or forming aggregates before reaching the membrane-bound secretory translocase (van Wely *et al.* 2001). However, the precise details of these interactions are poorly understood, but are likely to involve the Signal Recognition Particle (SRP), and/or the SecA/SecB complex, either individually or in concert (Figure 1.2).

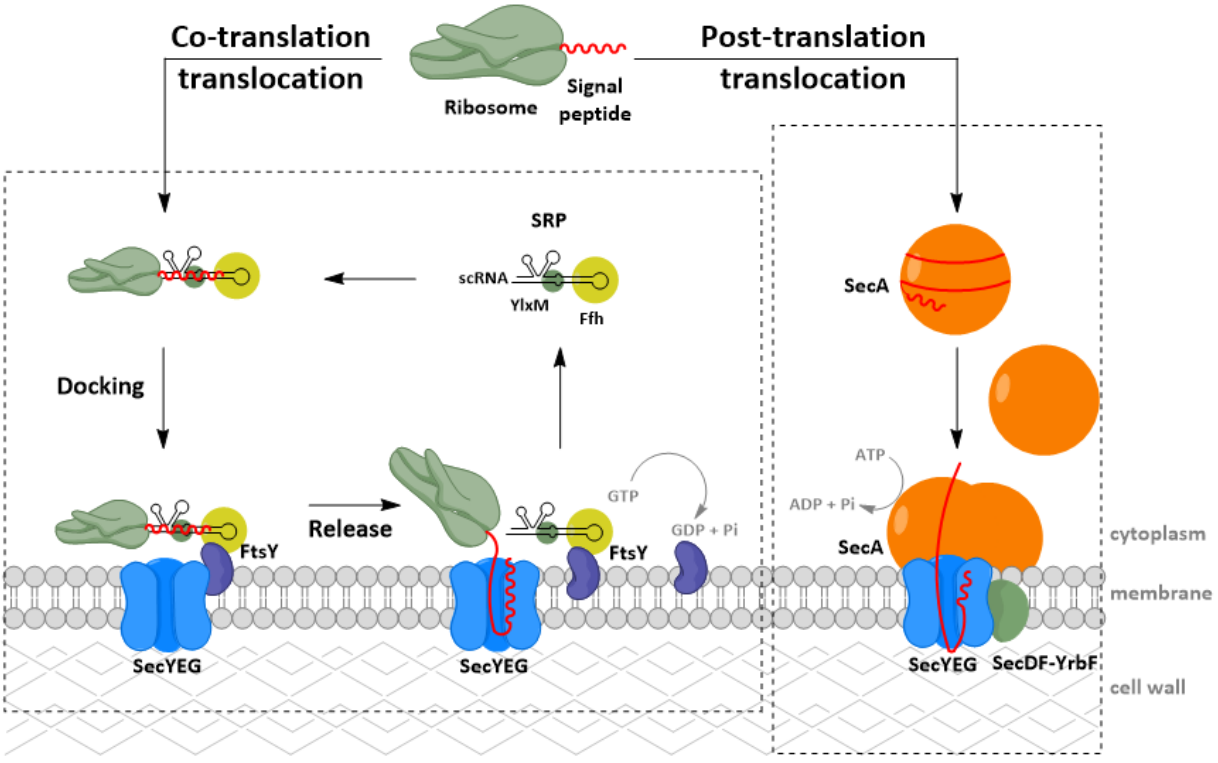


Figure 1.2 Schematic representation of the cytoplasmic chaperoning and targeting pathways for the targeting of secretory preproteins to the SecYEG translocase of *B. subtilis*.

SecA/SecB complex. SecB is a cytoplasmic chaperone that interacts with a subset of secretory proteins as they emerge from the ribosome in Gram-negative organisms. *E. coli* SecB binds to the mature region of SecB-dependent pre-secretory proteins. The resulting complex interacts with SecA+ADP to form a tertiary complex that, in turn, interacts with the membrane-located secretory translocase. Conformational changes occur, resulting in the release of SecB and the replacement of ADP with ATP on SecA (Kumamoto 1989; Harwood & Cranenburgh 2008). SecB is a tetramer, specifically, a dimer of dimers. Each dimer binds one molecule of the SecA-ADP dimer that interacts with the translocase, via the highly conserved C-terminal SecB binding domain (van Wely *et al.* 2001).

The SecB protein is absent from *B. subtilis* and all other Gram-positive bacteria, and many studies have focused on identifying a functional homologue of this protein in this group of organisms. To date these studies have failed to identify a convincing homologue. It might be important to note that SecB is not essential for viability of *E. coli* and it is needed only for the secretion of a specific subset of proteins, many of which are outer membrane proteins (van Wely *et al.* 2001). The absence of outer membrane proteins in Gram-positive bacteria might explain why a homologue of SecB has not been found in this group of organisms.

The only potential protein chaperone that has been identified in *B. subtilis* is the cytosolic protein CsaA. It has been shown to interact with SecA and unfolded precursor proteins, and to restore the growth defect on *E. coli* SecA and SecB mutants (Müller *et al.* 2000). Moreover, deletion of the *csaA* gene is lethal. This suggests that CsaA can either be involved in the translocation of essential proteins or play a vital role in a process other than protein translocation (Müller, Bron, *et al.* 2000).

The Signal Recognition Particle. Since *B. subtilis*, as with other Gram-positive bacteria, apparently lacks a homologue of SecB, the SRP pathway is the only recognised intracellular pathway for recognising, chaperoning and targeting secretory proteins to the translocation sites.

The SRP is a highly conserved and essential RNA-protein complex that is found in all domains of life. In *B. subtilis* the small cytoplasmic RNA (scRNA) is 271 nucleotides in length and is structurally similar to that of archaeal and eukaryotic SRP RNA, with Alu and S domains. *E. coli* and other Gram-negative bacteria have a much shorter scRNA (114 nucleotides) and consequently lack the Alu domain. The scRNA provides a backbone for the attachment of SPR

proteins. In *B. subtilis* two such proteins are identified in the literature, Ffh and HBsu (Nakamura *et al.* 1999). Ffh ('fifty-four homologue') is a GTPase that is homologous to the 54-kDa subunit of the eukaryotic SRP (Honda *et al.* 1993). HBsu is an essential non-specific histone-like DNA-binding protein and one of the most abundant proteins in *B. subtilis* (Ross & Setlow 2000). Recently, unambiguous evidence has shown that HBsu is in fact not a component of the SRP complex and that instead YlxM binds to elements of the Alu and S domains (Ismail, personal communication). The gene encoding YlxM resides directly up-stream of the *ffh* gene in *B. subtilis* and numerous other Gram-positive bacteria and it is thought to be a component of the SRP component with a putative regulatory function (Williams *et al.* 2014).

Very little is known about the functioning of the SRP in bacteria, and what little is known, mainly comes from studies in *E. coli*, that has an SRP with a single domain and single protein, Ffh. It is thought that the SRP complex binds to the signal peptide of integral membrane proteins emerging from the ribosome, targeting them to the cytoplasmic membrane via the Sec translocase by delivering them to a membrane-bound docking protein, FtsY (Du Plessis *et al.* 2011). Both Ffh and FtsY are essential for SRP-dependent protein secretion and cell viability (Ashikaga *et al.* 2003). There is a third SRP-GTPase FlhF with structure similarity with Ffh and FtsY, encoded by a flagellar gene (Carpenter *et al.* 1992). However, Zanen *et al.* (2004) showed that FlhF is not essential for protein secretion or motility, leaving its role to still be identified.

When the SRP/ribosome complex attaches to the translocase, by interaction with the membrane-bound receptor FtsY, the translocase complex is then opened, the SRP is released, GTP is hydrolysed, and the lateral co-translational translocation of the targeted protein into the membrane occurs (Figure 1.2) (Du Plessis *et al.* 2011). Compared to Gram-negative bacteria, the SRP RNA in *B. subtilis* is much longer (271 nucleotides instead of 114).

The absence of detailed studies of the intracellular events involved in *B. subtilis* secretion means that there is a significant gap in our understanding of this stage in the secretion pathway. This has been filled by a combination of extrapolation from studies on *E. coli* and higher organisms and pure speculation. For example, some have suggested that the SRP complex provides the chaperone activity for all integral membrane and secreted proteins in B. subtilis, while others have suggested that SecA alone performs this function for secretory proteins. Only after these early cytoplasmic events are better understood will we have insights needed to devise strategies to improve this stage of the secretion pathway.

1.3.2 The Sec-dependent secretion machinery

After the targeting of the secretory preprotein to the membrane, the positively charged N-domain of the signal peptide interacts with negatively charged phospholipids in the membrane, leading to looping insertion of the H-domain. After the H-domain unloops, the first part of the preprotein is translocated through the membrane via a confined aqueous channel composed of a set of integral membrane proteins (van Wely *et al.* 2001).

The Sec translocases of *E. coli* and *B. subtilis* are similar and studies on both systems have contributed to much of what is known about this complex. In *B. subtilis*, the components of the translocase complex include the SecY, SecE, and SecG proteins, which form the SecYEG complex. This is the core of the translocation channel and these proteins are associated in the membrane with the SecDF protein. SecA is an ATPase located on the *cis* side of the membrane and which is responsible for both targeting and the coupling of the energy required for translocation – it is the motor component of the translocation.

SecYEG – the membrane bound complex. SecY is the largest component of the translocase – it is a 423-amino acid polypeptide – and is essential for translocation and viability. The *secY* gene is located in a ribosomal operon, the *spc* operon. SecY is predicted to span the cytoplasmic membrane 10 times and is found in randomly dispersed foci that dynamically assemble and disassemble in the membrane (Suh *et al.* 1990; Dajkovic *et al.* 2016).

SecE is a small integral membrane protein that is essential for translocation and viability (Jeong *et al.* 1993). SecE associates with SecY preventing it from being degraded by FtsH, an ATP-dependent protease present in the cytoplasmic membrane (Lycklama A Nijeholt *et al.* 2013).

SecG is the third component of this integral membrane complex. It is the only component that is not essential for viability and protein translocation. Some strains become cold-sensitive when the *secG* gene is deleted. In *B. subtilis*, the deletion of this gene only causes cold sensitivity when secretory stress is imposed (van Wely *et al.* 1999). These observations suggest that SecG only contributes to the efficiency of the translocation reaction.

In *E. coli*, pore formation requires the activation of SecA by a secretory preprotein and ATP (see later). Under those conditions, the SecA dimer recruits four SecYEG units to form an oligomer with a large central opening that may function as the protein-conducting pore (Manting *et al.* 2000). The *B. subtilis* SecYEG complex likely functions similarly.

SecA – the motor of translocation. SecA is a homodimer with a subunit mass of about 100 kDa (Takamatsu *et al.* 1992). SecA is responsible for initiating the post-translational translocation through the pore of the translocase, deriving its energy from the binding and hydrolysis of ATP. The SecA protein exhibits a low endogenous ATPase activity, which can be stimulated by the presence of membranes, the SecYEG complex, and precursor proteins (van Wely *et al.* 2001).

In *E. coli*, the cytosolic level of SecA is autoregulated. SecA binds to its own mRNA when it is free, inhibiting its own translation (Dolan & Oliver 1991). In *B. subtilis*, the SecA level gradually increases during the exponential phase of growth, reaching its maximum at the transition stage to the stationary phase, coinciding with the maximum production of exoproteins (Herbort *et al.* 1999).

Studies in *E. coli* suggest that SecA, the motor component of the translocase, mediates the translocation of proteins through the membrane by changing its conformation and insertion in the membrane when ATP is bound. SecA can be divided into several structural domains, the central core DEAD motor consisting of two nucleotide-binding fold (NBF) domains, NBF1 and NBF2, homologous to RecA-like motives found in DNA/RNA helicases (Prabudiansyah & Driessen 2016). Both domains are required for ATPase activity, NBF2 regulates the catalytic activity of NBF1 and the intramolecular regulator of ATP hydrolysis (IRA1) is located in the helical scaffold domain (HSD), immediately downstream of NBF2. SecA also contains the preprotein cross-linking domain (PPXD) with two substrate specific binding sites for the signal peptide and the mature domains of the preprotein, and the so-called two-helix finger (2HF) positioned at the SecA-SecY interface (Pohl & Harwood 2010; Corey *et al.* 2016).

A large number of proteins are translocated via the Sec translocase and it is challenging to envisage one model that recognises and transports them all. The current accepted models can be divided into three categories:

i) Power-stroke models: multiple rounds of ATP hydrolysis transport a short stretch of the peptide at a time (Economou & Wickner 1994). The 2HF is likely to take the role of piston as it is expected to interact with both preprotein and SecY due to its positioning at the SecA-SecY interface (Erlandson *et al.* 2008; Bauer *et al.* 2014).

ii) SecA dimerization models: depending on the experimental conditions, different degrees of SecA dimerization have been found, leading to the suggestion that

translocation might be driven by interaction between multiple SecA dimers in an ATP-dependent fashion, pushing the substrate through the translocation channel (Or *et al.* 2002; Bauer *et al.* 2014).

iii) Diffusional ratchet models: the directional motion of the preprotein through the translocase channel is favoured by chemical asymmetry across the membrane or random diffusion, providing that there is no interaction between the channel and the substrate, and potentially stimulated by the proton motive force (PMF) (Corey *et al.* 2016).

Very recently, Dajkovic *et al.* (2016) used a combination of total internal reflection microscopy (TIRF), scanning fluorescence correlation spectroscopy (FCS), and pair correlation function (pCF) to study the localization of SecA and SecY in *B. subtilis, in vivo.* This study revealed a remarkably dynamic secretion system and proposed a model for the secretion machinery that involves constant reorganization of SecA and SecY clusters throughout the cytoplasmic membrane, fitting a combination of the power-stroke and SecA dimerization models. In this model, SecA delivers the preprotein to the translocase and multiple molecules of SecA on the membrane and in the cytoplasm associate and accumulate during translocation (Dajkovic *et al.* 2016).

SecD and SecF – translocase accessories. Another heterotrimeric complex with a role in the secretion apparatus comprises SecD, SecF and YacJ in *E. coli.* In *B. subtilis*, the YacJ homologue is YrbF. However, there is no known function for YacJ or YrbF in protein translocation (Harwood & Cranenburgh 2008). The SecD and SecF equivalents are present in one single membrane protein in *B. subtilis*, denoted SecDF, predicted to span the membrane 12 times (Bolhuis *et al.* 1998). This heterotrimeric SecDF-YrbF complex is loosely associated with SecYEG, forming a supramolecular translocase complex.

SecDF is thought to only contribute to the efficiency of protein secretion by improving SecA cycling and maintaining the forward momentum of the preprotein (Driessen & Nouwen 2008). Similarly to SecA, the *B. subtilis* SecDF is maximally expressed at early stationary phase in rich medium. On minimal medium, however, expression appears rather constant throughout the growth cycle (van Wely *et al.* 2001).

SPases. The signal peptides of newly secreted preproteins are cleaved by signal peptidases during or shortly after translocation. Type I signal peptides are cleaved by SipS or SipT, while Type II signal peptides associated with lipoproteins are cleaved by LspA (Tjalsma, M. A. Noback, *et al.* 1997; Tjalsma *et al.* 1998). This reaction is essential for the release of the secretory protein from the membrane and peptidases SppA and TepA degrade the resulting signal peptide fragments which would otherwise be inhibitory to protein translocation (Albert Bolhuis *et al.* 1999; Tjalsma *et al.* 2000).

1.3.3 The Bacillus subtilis cell wall and its influence on protein export

It is following translocation of secretory proteins across the membrane that the biggest differences between Gram-positive and Gram-negative bacteria emerge. Gram-negative bacteria have a membrane-enclosed periplasm outside the cytoplasmic membrane that is surrounded by the outer membrane. The periplasm of Gram-negative bacteria is a dynamic and metabolically highly active environment. The outer membrane of Gram-negative bacteria contains lipopolysaccharides in its outer leaflet and phospholipids in the inner leaflet. In contrast, in *Bacillus*, like all Gram-positive bacteria, does not have an outer membrane or a membrane-enclosed periplasm. Although this could be seen as an advantage for the commercial production of proteins, such proteins need to be able to fold rapidly in an environment dominated by the complex physicochemical properties of the peptidoglycan-anionic polymer complex that forms the Gram-positive cell wall.

Gram-positive bacteria compensate for the lack of a periplasm by lipo-modifying homologues of many of the proteins located in the periplasm of Gram-negative bacteria and attaching them to the outer leaflet of the cytoplasmic membrane (Sarvas *et al.* 2004). Moreover, Gram-positive bacteria also immobilize some secretory proteins in the cell wall by ionic or covalent interactions. Proteins active on or at the cell surface include quality control proteases, extracytoplasmic chaperones, autolysins, surface layer proteins, and substrate binding proteins, making the gram-positive cell wall a dynamically flexible structure that is up to 20 times thicker than the peptidoglycan layer of Gram-negative bacteria (Sarvas *et al.* 2004; Pohl & Harwood 2010).

More specifically, the *Bacillus* cell wall consists of a thick, highly cross-linked semi-porous copolymer of peptidoglycan and covalently linked anionic polymers that shield the cytoplasmic membrane from its environment. It also plays key roles in cell division, cell shape maintenance,

metal ion homeostasis, and controls several interactions between the cell and its environment (Vollmer *et al.* 2008).

Peptidoglycan is made of linear glycan strands of alternating *N*-acetylglucosamine (Glc*N*Ac) and *N*-acetylmuramic acid (Mur*N*Ac) residues cross-linked by short peptides (Vollmer *et al.* 2008). In *B. subtilis* these glycan strands are considered to be long (between 50 and 250 disaccharide units) when compared to, for instance, *Staphylococcus aureus* strands (between 3 and 10 disaccharide units) (Vollmer *et al.* 2008).

Many Gram-positive bacteria contain peptidoglycan-attached carbohydrate-based polymers that differ between species or even between strains (Weidenmaier & Peschel 2008). *B. subtilis* has two distinct phosphate rich wall teichoic acids (WTA). The most abundant WTA in strain 168 is poly(glycerol phosphate), with minor amounts of a second polymer, poly(glucosyl *N*-acetylgalactosamine 1-phosphate) (Freymond *et al.* 2006). Teichuronic acid is produced in place of WTA under conditions of phosphate limitation, while the displaced WTA is utilized as a source of phosphate (Allenby *et al.* 2005). These anionic polymers and lipoteichoic acids result in the immobilisation of a high density of negative charge in the cell wall, the very environment into which secretory proteins emerge in an unfolded state from the translocase (Hyyrylainen *et al.* 2000).

Even though at first sight the Gram-positive bacteria seem to offer architecture that is compatible to protein export, a closer look at the cell wall reveals that it offers a very challenging environment for proteins to fold after emerging from the translocase. Unfolded proteins that emerge from the translocase must be able to fold into their native configuration quickly without forming intra or intermolecular interactions that could block the translocation machinery and/or block cell wall synthesis sites, either or which would compromise cell viability.

Stephenson *et al.* (2000) showed that the cell wall has the potential to retard the passage of positively charged secretory proteins to a greater extent than neutral or negatively charged proteins. This confirms the impact of physiochemical properties of the cell wall on protein secretion. It also offers insights for improving the production of certain proteins – tailoring the charge of the cell wall might be a strategy to pursue.

1.3.4 Folding outside the membrane

A central feature of secretion is post-translocational folding. Folding of proteins into their native configuration is crucial not only for their functionality but also for their stability. Misfolded proteins are not only non-functional, but are more susceptible to proteolysis since protease sensitive residues that are normally not exposed on the surface of the protein can be exposed when they are misfolded. *B. subtilis* contains several quality control proteases in the membrane and cell wall, as well as extracellular proteases that rapidly target heterologous proteins or misfolded native proteins with exposed protease recognition sites. Therefore, after crossing the membrane in an essentially unfolded state, it is important that secretory proteins fold rapidly into their active and protease-resistant conformation. For many proteins, this is an assisted process rather than a spontaneous one.

There are intrinsic and extrinsic mechanisms to ensure rapid folding on the *trans* sides of the cytoplasmic membrane. These include chaperones and folding factors, such as propeptides, a peptidyl-prolyl cis/trans isomerase, disulfide isomerases, and metal ions (Pohl & Harwood 2010).

Propeptides. Many secretory proteases in *B. subtilis* are synthesized as precursors prepro-peptides, with and additional peptide located between the signal peptide and mature protein. These propeptides play an important role in folding, firstly by helping to prevent folding in the cytoplasm and then reducing the activation energy for the protein to enter its folding pathway (Shinde *et al.* 1993). Propeptides vary in length and are cleaved proteolytically, sometimes autocatalytically, to generate the active enzymes.

Yabuta *et al.* (2001) proposed a model for the propeptide-mediated protein folding pathway. The unfolded precursor that is translocated through the membrane undergoes a process of folding, autoprocessing and degradation. The folding of the protein requires the presence of the propeptide domain and is a rapid process that occurs through a partially structured non-native folding intermediate. Then this intermediate undergoes structural changes to become a native-like folded precursor. The autoprocessing reaction then occurs to form the mature protein. The protease domain has to somehow free itself of its cognate propeptide domain to promote proteolysis. Once a free protease is formed, it can bind to the propeptide domain in the autoprocessed complex and facilitate *trans* degradation. Propeptides have been explored as tools to enhance the secretion of heterologous proteins with mixed success (Pohl & Harwood 2010).

Metal cations. Metal cations help to counterbalance the negative charge of the teichoic acids in the cell wall, particularly Ca^{2+}, Fe^{3+} and Mg^{2+}. These metal cations are required by some secretory proteins to facilitate their folding and in the case of metalloproteins, are necessary for their stability and activity (Sarvas *et al.* 2004). In fact, it was shown that if the negative charge of the cell wall is increased by inactivation of the *dlt* operon, required for alanylation (neutralization of adjacent phosphoryl residues) of teichoic acids, the yield of many secretory proteins is increased, presumably due to increased affinity of the cell wall for cations (Hyyrylainen *et al.* 2000).

Disulfide isomerases. The oxidized environment into which secretory proteins are translocated favours disulfide bond formation. Furthermore, *B. subtilis* expresses 4 thiol-disulfide oxidoreductase/isomerase enzymes involved in the disulfide bond formation: BdbA, BdbB, BdbC and BdbD, as in *Bacillus d*isulfide *b*ond formation (Sarvas *et al.* 2004). However, secretory proteins from *B. subtilis* and other Gram-positive bacteria generally lack disulfide bonds. Nevertheless, these enzymes are favourable for the production of heterologous proteins that require such bonds.

PrsA. PrsA is an essential membrane-associated hydrophilic 33 kDa lipoprotein, bound to the outer leaflet of the cytoplasmic membrane (Kontinen *et al.* 1991). Based on sequence similarity, PrsA is considered a peptidyl prolyl *cis-trans* isomerase (PPIase) of the parvulin family (Vitikainen *et al.* 2004; Tossavainen *et al.* 2006). PPIases assist proteins with *cis*-prolyl residues to fold. PrsA is found ubiquitously in all Gram-positive species but not in any Gram-negative (Sarvas *et al.* 2004). Interestingly, it is an abundant protein with about 20,000 molecules per cell, in high excess over the number of translocase complexes (Vitikainen *et al.* 2001). Depletion o PrsA results in altered cellular morphology and cell death given that this protein is required for the folding of one or more proteins involved in cell wall synthesis (Kontinen & Sarvas 1993). Moreover, several studies have revealed that there is a linear relationship between the amounts of PrsA and the yield of certain secretary proteins (notably amylases), making PrsA a popular target when engineering host strains dedicated to the production of native as well as heterologous secretory proteins (Vitikainen *et al.* 2005).

1.3.5 Membrane-bound and extracellular proteases

Despite intracellular and extracellular folding factors, secretory proteins still misfold, particularly if the cell is subjected to certain types of stress. Gram-positive bacteria encode membrane- and

cell wall-associated proteases that prevent misfolded or aberrant proteins from blocking the translocase machinery or the cell wall growth sites, thereby providing as a quality control mechanism (Hyyrylainen *et al.* 2001; Westers *et al.* 2006a). Besides these proteases, *B. subtilis* encodes extracellular proteases that provide amino acids and peptides as nutrients by degrading proteins in the media (Pohl & Harwood 2010). Although these are essential processes for cell growth and fitness, they can represent major limitations for the production of heterologous proteins.

The extracytoplasmic quality control mechanisms are regulated by the well-characterised CssRS two-component system, consisting of the CssR response regulator and the CssS sensor kinase, that control the up-regulation of the HtrA and HtrB serine proteases upon hyper-secretion conditions, and prevent potentially fatal obstruction of the secretory translocase and cell wall synthesis (Hyyrylainen *et al.* 2001; Darmon *et al.* 2002; Westers *et al.* 2006a). It was recently suggested that, in addition to their already described roles as quality control proteases, HtrA and HtrB might also function to release lipoproteins from their membrane anchorage, to degrade membrane proteins and also to facilitate the folding of native secretory proteins (Krishnappa *et al.* 2013).

The primary role of the feeding proteases is to degrade proteins and peptides in the environment to provide amino acids and peptides as an important source of nutrients. There are eight feeding proteases in *B. subtilis*, namely, NprB, AprE, Epr, Bpr, NprE, Mpr, Vpr and WprA. AprE and NprE are responsible for 95% of the total extracellular proteolytic activity; however, the remaining five extracellular proteases (NprB, Epr, Bpr, Mpr and Vpr) can still hamper the production of heterologous proteins (Schmidt *et al.* 1993). WprA is a wall-associated protein with two processed products detected in the cell-wall extracts, cell wall binding protein (CWBP) 23 and 52 (Margot & Karamata 1996). CWBP52 is a serine protease, whereas CWBP23 is likely to be a propeptide of CWBP52. *wprA* is regulated by the YvrG-YvrHb two-component signal transduction pathway responsible for the induction of genes encoding the major cell-wall autolysis (Serizawa *et al.* 2005). Furthermore, WprA has been shown to be involved in degradation of PrsA, HtrA and HtrB, as well as certain heterologous proteins (Stephenson *et al.* 2000; Krishnappa *et al.* 2014).

The so-called feeding proteases not only have a negative impact on the production of certain heterologous proteins but also on native *B. subtilis* proteins (Pohl *et al.* 2013). This includes

WapA, a 258 kDa wall-associated protein that is extensively processed in the presence of extracellular proteases (Antelmann *et al.* 2002). Since none of the feeding proteases are essential for cell viability, an obvious strategy to overcome this bottleneck is to engineer mutants where the impact of these proteases is reduced (Pohl *et al.* 2013).

1.4. Optimisation of *Bacillus subtilis* cell factories

B. subtilis is an attractive host for the production of heterologous proteins at an industrial scale mainly due to the absence of an outer membrane and resulting ability to secret proteins directly into the culture medium at high concentrations. Furthermore, *B. subtilis* molecular biology is well understood and it serves as a biological model for many other bacterial systems. The optimisation of *B. subtilis* strains for biotechnology applications often involves a combined approach that targets both the genetic circuits that control the expression of the product of interest, and the secretion capacity of the cell. Nevertheless, attempts to use this bacterium for the production of heterologous proteins are often met with mixed success.

1.4.1 Control of gene expression

Gene expression in prokaryotes is affected by a number of complex factors. Transcriptional regulation is one of the main focuses when engineering cell factories. Transcription of target gene is controlled by the promoter sequence which must interact with the RNA polymerase at the σ subunit to form the holoenzyme and initiate transcription (Browning & Busby 2004). Promoter engineering is one of the strategies used for the control of target gene expression, focusing on the promoter architecture in order to maximise the transcription initiation rate and, consequently, the number of target gene messenger RNA in the cell. Particularly, the -10 and the -35 hexamers located 10 and 35 base pairs upstream from the transcription site play an important role in docking the RNA polymerase, and the consensus sequences of these elements are well-known. However, this is a complicated approach which depends on the complex interplay between the promoter, operators, RNA polymerase, transcription factors and effector molecules. Often, a high throughput random or semi-rational approach facilitates the identification of suitable systems for transcription control. The design of new promoters with biotechnological significance focuses on saturation mutagenesis of spacer regions between these consensus sequences, error-prone PCR and hybrid promoter engineering.

The strength of the promoter does not always correlate with the final product yield. Protein translation efficiency of the target gene mRNA is a crucial factor for protein production. Translation occurs in four phases: initiation, elongation, termination and ribosome recycling. The efficiency of these processes can be influenced by several factors, including the rate of translation initiation, which involves the recognition and binding of the ribosome to the ribosome binding site (RBS), the stability of the mRNA and the host codon usage. *B. subtilis* translation initiation efficiency is very sensitive to the RBS and several studies have addressed the importance of having an efficient *B. subtilis*-like RBS for expression of heterologous genes (Hager & Rabinowitz 1985). RBSs have three features: (1) a Shine-Dalgarno (SD) sequence which interacts with the anti-SD sequence of the 16s subunit of the ribosomal RNA (CCUCC); (2) a spacer region between the SD sequence and the translation initiation codon with optimum length between 7 and 9 nucleotides; (3) the translation initiation codon, AUG, GUG or UUG with AUG being the preferred initiation codon (Shine & Dalgarno 1974; Band & Henner 1984; Vellanoweth & Rabinowitz 1992) . Secondary structure within the RBS domain also plays an important role in determining the efficiency with which a given RBS is recognised by the ribosome. Particularly, a genome wide analysis revealed significantly different codon usage for nearly all aminoacids from positions +2 up to +10 of the gene, in the sense of increasing the number of A terminating codons presumably to avoid mRNA secondary structures (Rocha *et al.* 1999). Furthermore, it is well-known that there is a strong correlation between codon usage and tRNA content in unicellular organisms which correlates to the protein production levels of both native and heterologous genes (Ikemura 1985; Plotkin & Kudla 2011). Expression systems are often engineered to optimise the RBS, include RNA stabilizing elements on the 5' and 3' ends of the mRNA and use codon optimised target gene sequences that are favoured by the host specific pool of tRNAs.

Finally, the nature of expression construct also plays in the control of the target gene expression. Often, replicative plasmids with high copy number are used for the production of high levels of heterologous proteins. However, these systems are known to be unstable and require the presence of antibiotics to assure their maintenance, which is not a favourable feature for the industrial production of proteins for both economic and regulatory reasons. Alternatively, genome integration assures that the heterologous gene is stably maintained at a copy number that reflects that of the chromosome. However, the copy number of the ingrated gene fluctuates due to multiple rounds of replication that originate at the origin of replication before the previous round has been completed. Recently, Sauer *et al.* (2016) showed that this gene dosage is an important

factor that influences heterologous gene expression in *B. subtilis* with a potential 5-fold increase in protein production when the target gene is close to the origin of replication.

Due to their robust mechanism of gene regulation, designing strong and stable expression constructs is often not enough to optimise product production. It is important to take into account the tight control of metabolomic and stress-response pathways and how they influence the control of the target gene expression.

1.4.2 Optimisation of the secretion capacity of Bacillus subtilis protein production strains

The secretion pathway in *B. subtilis* involves multiple stages from targeting to release into the extracellular environment which often constitute secretion bottlenecks for the production of heterologous proteins. Many of the bottlenecks responsible for reductions in yield are well established and these represent promising targets for optimisation strategies. Following the fate of secretory proteins from synthesis to release into the extracellular culture medium, these bottlenecks are as follows: (i) the targeting and transport of the protein to the translocase in a secretion-competent state; (ii) passage through the secretion channel; (iii) post-translational folding; (iv) crossing of the cell wall barrier; (v) stability in the culture medium.

A major issue has been to understand the intracellular chaperoning and targeting mechanisms of secretory proteins, and more studies are need to address the roles of the SRP, CsaA and SecA. However, it is well accepted that the low yield of many heterologous proteins is mainly the result of later stages of the secretion pathway (Harwood & Cranenburgh 2008).

Nevertheless, to address the potential formation of secretion-incompetent (i.e. folded or aggregated) intermediates of secretory preproteins in the cytoplasm, the synthesis of intracellular chaperones such as GroES and GroEL, or DnaK, DnaJ and GrpE has been up-regulated by inactivating the gene encoding HrcA, the negative regulator of the *groE* and *dnaK* operons (Wu *et al.* 1998). The authors showed increased yields of the single-chain antibody (SCA) fragment and reduction of inclusion body formation. Diao *et al.* (2012) constructed an artificial posttranslational protein targeting pathway in *B. subtilis* by co-expressing SecB from *E. coli* and a *B. subtilis* hybrid SecA, where the C-terminal 32 amino acids were replaced by the ones of SecA from *E. coli*. With this strategy, the authors showed a significant improvement of maltose binding protein (MalE11) and alkaline phosphatase (PhoA) secretion.

Several studies have addressed the improvement of the amounts of secreted proteins by attempting to optimise the combination of signal peptide and mature protein. Zanen *et al.* (2005) have suggested that *B. subtilis* is not able to secrete preproteins with signal peptides of low hydrophobicity. It is known that, in *E. coli*, signal peptide hydrophobicity is a determinant for the targeting of secretory proteins to the SRP pathway, rather than to the SecA/B pathway (Lee & Bernstein 2001). It is notable that, compared with those of Gram-negative bacteria, preproteins from Gram-positive bacteria (which lack SecB) have longer and more hydrophobic signal peptides. It would therefore be expected that secretory preproteins with signal peptides with high hydrophobicity would be more efficiently translocated in this bacterium (Zanen *et al.* 2005). Several attempts have been made to identify an optimal "*Bacillus'* signal peptide and libraries of native signal peptides have been constructed to identify the optimal peptide for specific heterologous proteins. Both approaches have revealed large variations in target protein yield when using the same signal peptide (Brockmeier, Wendorff, *et al.* 2006; Degering *et al.* 2010). The optimal signal peptide for one particular recombinant protein is not necessarily the best for the secretion of a different protein. Most likely explanation is that the signal sequence and the immediate downstream amino acid sequences of native proteins have been optimised during evolution to avoid unfavourable interactions. Obviously such optimisation process has not occurred in the case of heterologous proteins and the signal peptide library approach provides the best alternative strategy.

The rate of signal peptide processing by signal peptidases has also been considered has a possible limiting factor for protein production. Vitikainen *et al.* (2001) showed that the proportion of mature AmyQ in cells of the SipT overproducer was higher than in the wild type. Furthermore, multiple type I signal peptidases have different affinities for different secretory precursors, suggesting that controlled overexpression of the optimal combination of signal peptidases can potentially elevate the amounts of a protein of interest (Tjalsma, *et al.* 1997).

Manipulation and optimisation of the translocation capacity of the cell and the translocase complex itself is of great interest for the optimisation of production strains. SecA, the motor protein, is considered to be a key pathway component due to its ability to interact with both the secretory preproteins and membrane translocases. Kakeshita *et al.* (2010) showed that deleting the 61 amino acids at the C-terminal region of SecA led to significant enhanced production of heterologous proteins in *B. subtilis.* Furthermore, aiming to increase the number of translocons,

Mulder *et al.* (2013) constructed an artificial SecYEG operon fused to an inducible promoter and demonstrated that the yield of α-amylase could be improved by increasing production of the SecYEG proteins. Chen *et al.* (2015) took a combinatorial approach to systematically overexpress 23 genes or operons encoding proteins involved in or closely related to the Sec pathway in *B. subtilis*, including the translocase genes. In contrast to the work of Mulder *et al.* (2013), Chen *et al.* (2015) found that overexpressing the translocase genes did not significantly increase of the yield of heterologous proteins.

Following translocation, secretory proteins must rapidly fold in the challenging environment of the cell wall. Increased amounts of the membrane-associated lipoprotein PrsA has led to improved production of several heterologous targets in *B. subtilis*, such as the recombinant protective antigen (rPA) of *B. anthracis*, AmyS of *Geobacillus stearothermophilus*, pneumolysin, a fibrin-specific single-chain antibody fragment, and AmyL and AmyS from *B. licheniformis* (Williams *et al.* 2003; Vitikainen *et al.* 2005; Wu *et al.* 2002; Chen, Fu, *et al.* 2015). As a result, the PrsA chaperone is a well-established target for optimising production strains.

Many heterologous proteins are susceptible to the numerous extracellular quality control and feeding proteases of *B. subtilis*. Controlling the proteolytic activity of a protein production host is a logical and fairly straightforward strategy for maximisation of protein production. Many strains with several different combinations of proteases deletions have been constructed (Kodama *et al.* 2012; Pohl *et al.* 2013). Such strains have been helpful in improving the production of some heterologous proteins but not others, reflecting the fact that the folding and structure characteristics of each heterologous protein determine its individual stability and production yield.

Overall, all these approaches offer important knowledge not only for the development of commercial strains optimised for the production of native and heterologous proteins, but also for the understanding of the molecular biology of the organism. Nevertheless, characterisation of the less understood limiting factors, such as SRP transport, is still a major challenge and further research is needed to elucidate the complicated mechanism and control of the secretion pathway.

1.4.3 Protein products from Bacillus cell factories

Currently, about 60% of the commercially available enzymes are produced by *Bacillus* species (L. Westers *et al.* 2004). Most of them, are native naturally secreted enzymes, such as alkaline

proteases and amylases for the detergent and starch industries. *Bacillus subtilis*, *Bacillus amyloliquefaciens* and *Bacillus licheniformis* are the most popular *Bacillus sp*. for industrial production of enzymes due to their fermentation properties, high product yields (20 to 25 grams per litre) and the complete lack of toxic by-products (van Dijl & Hecker 2013). Table 1.1 summarises the yields obtained for the production of several native or heterologous protein products available in the literature. The production of heterologous proteins from eukaryotes produces the lowest yields which seems to be related to the intracellular accumulation of the precursor protein (Kakeshita *et al.* 2010). *E. coli* is still the most commonly used host for industrial production of pharmaceutical proteins mainly due to being the first choice in the lead finding phase of a drug development project given its genetic acessability (L. Westers *et al.* 2004). However, due to recent advances in functional genomics and Systems and Synthetic Biology analysis of *B. subtilis*, there is the potential for a significant increase in the variety of recombinant enzymes and biopharmaceutical proteins sucessfuly produced by *B. subtilis*.

Table 1.1 Protein products from *B. subtilis* ordered by production yield.

Product	Protein Size (kDa)	Native host	Yield	Reference
Human Interferon (IFN)- α2b	19	*Homo sapiens*	0.5-1 mg/L	(Palva *et al.* 1983)
Poly(30hydroxybutyrate) depolymerase	42	*Paucimonas lemoignei*	1.9 mg/mL	(Braaz *et al.* 2002)
hEGF	134	*Homo sapiens*	7 mg/L	(Lam *et al.* 1998)
scFV	26	*Homo sapiens*	10-15 mg/L	(Wu *et al.* 2002)
Cutinase	20	*Fusarium solani pisi*	20 mg/L	(Brockmeier, Caspers, *et al.* 2006)
Human Interferon (IFN)- γ	19	*Homo sapiens*	20 mg/L	(Rojas Contreras *et al.* 2010)
Streptavidin	19	Streptomyces avidinii	35-50 mg/mL	(Wu & Wong 2002)
Interleukin (IL)-3	17	*Homo sapiens*	100 mg/L	(L. Westers *et al.* 2006)
Lipase A	19	*B. subtilis*	600 mg/L	(Lesuisse *et al.* 1993)
Proinsulin (PI)	12	*Homo sapiens*	1 g/L	(Olmos-Soto & Contreras-Flores 2003)
α-amylase (AmyQ)	58	*B. amyloliquefaciens*	1-3 g/L	(Palva 1982)
Acid-stable α-amylase	59	*B. licheniformis*	3.1 g/L	(Heng *et al.* 2005)

1.5. Aims

This thesis presents the work of a collaboration between industry and academia aimed at investigating important questions related to enzyme production by the industrial workhorse *B. subtilis*. Improving the production efficiency of an increasingly large variety of commercial proteins and enzymes poses significant challenges for industrial-scale fermentations. A variety of strain optimisation strategies have been described, aimed at different stages of the protein synthesis and secretion pathway, however, these are not applicable for every product and further knowledge is need to understand what determines the limitations of production. The aim of this project was to use two industrially-relevant enzymes to study and compare, in parallel, the challenges of native and heterologous enzyme production by *B. subtilis*.

Therefore, this study presents a comprehensive comparative study that looks into the effects of native and heterologous enzyme production with respect to several previously identified bottlenecks at the later stages of the secretion pathway. Specifically, we analysed the enzyme-related effects on the production profiles and kinetics of protein secretion, the effect of native and heterologous signal peptides, the impact of overexpressing the translocase genes, the effect of up-regulation of the quality control mechanisms under secretion stress conditions and the potential of a collection of protease deficient hosts for enzyme production. Furthermore, we investigated the effect of both native and heterologous enzymes on the regulation of the genes expressing relevant proteins involved in the secretion, chaperoning and quality control mechanisms of the secretion pathway. Together, these studies provide a comprehensive analysis of the interplay between a particular product of interest and key elements of the host-mediated secretion pathway with the ultimate aim of optimising productivity.

Chapter 2 Methods

In this chapter, the molecular biology, microbiology and enzymatic methods used throughout this study are described.

2.1. Bacterial strains and plasmids

The bacterial strains and plasmids used in this study are described in the beginning of each chapter. Schematic representations of plasmids are in Appendix B.

2.2. Growth and maintenance of bacterial cultures

Nutrient Luria-Bertani medium (LB) was used for general-purpose growth of *Escherichia coli* and *Bacillus subtilis.* For *B. subtilis* DNA transformation, the Spizizen-plus and Spizizen-starvation media were used. Growth media and buffers are described in Appendix A.

Growth in LB and Spizizen media was at 37 °C with orbital shaking at 250 rpm, unless otherwise stated. Growth supplements (Table 2.1) and antibiotics (Table 2.2) were added to the media when appropriate. Growth was monitored by measuring optical density at 600 nm (OD_{600}) in a spectrophotometer.

For long-term storage of bacterial strains, LB broth cultures were supplemented with 10 % (w/v) glycerol after overnight growth, and stored at -80 °C in cryovials or Nunc™ 96-well polypropylene MicroWell™ Plates.

Table 2.1 Stock and working solutions of growth supplements

Supplement	Solute	Stock concentration	Usual working concentration
IPTG	dH_2O	1 M	1 mM
Xylose	dH_2O	20% (w/v)	0.2%

Table 2.2 Stock and working solutions of antibiotics.

Antibiotic	Solute	Stock concentration ($mg.ml^{-1}$)	Working concentration ($\mu g.ml^{-1}$)
Ampicillin (Amp)	dH_2O	100	100
Spectinomycin (Spec)	dH_2O	100	100
Neomycin (Neo)	dH_2O	50	50
Erythromycin (Em)	100% ethanol (v/v)	10	2
Chloramphenicol (Cat)	50% ethanol (v/v)	35	10

2.3. Bacterial culture in the BioLector® microfermentation system

The BioLector® bench top microfermentation system (m2p-labs) was used to preform high-throughput fermentations coupled to online monitoring of biomass, pH, DO and fluorescent reporters.

Overnight cultures were prepared prior to the BioLector® fermentations in 24-well plates in 2 mL of LB medium, supplemented with the appropriate antibiotics. The cultures were inoculated from a glycerol stock of cell culture and grown overnight at 37ºC with vigorous shaking (550 rpm) in a Multitron Standard Infors HT incubator. Pre-cultures were prepared by diluting the overnight cultures to an OD_{600} of 0.1, followed by incubation in the same conditions as the overnight cultures until mid-exponential phase (OD_{600}: 0.4-0.6) was reached. The pre-cultures were diluted to an OD_{600} of 0.1 in 1.5 ml pre-warmed LB medium in FlowerPlate® 48-well MTPs. The layout of the different cultures in the MTP was randomized. The culture plates were incubated for the indicated times at 37 °C with 95% humidity and strong agitation (800 rpm) in the BioLector® system. Biomass (excitation: 620 nm, gain: 20) and GFP (excitation: 488 nm, emission: 520 nm, gain: 95) signals were monitored by systematic measurements every 15 min.

Samples of 50 µl were manually extracted at given time points when measuring protein production and stored at -20 °C for later determination of the enzyme activity in the culture supernatant fractions (Section 2.10.).

2.4. Chemicals

All chemicals and reagents were obtained from Sigma-Aldrich (Dorset, UK) unless otherwise stated.

2.5. DNA manipulation

In general, all molecular methods involving DNA manipulation were performed according to Sambrook & Russell (2001). More detail is provided in the following sections.

2.5.1 Purification of plasmid DNA

Plasmid DNA was precipitated and purified from *E. coli* or *B. subtilis* using the QIAprep Spin Miniprep Kit (QIAGEN), according to manufacturer's instructions. This kit uses a column-based system that contains a silica membrane to which plasmid DNA can be bound, washed and eluted. For the isolation of plasmid DNA from *B. subtilis*, lysozyme from chicken egg white (Sigma-Aldrich) was included in the resuspension buffer of the kit, at a final concentration of 2 mg/mL and the cell suspensions incubated at 37 °C for 10 min. All plasmid DNA was eluted with sterile deionized H_2O.

2.5.2 Precipitation of genomic DNA from Bacillus subtilis

For precipitation of genomic DNA from *B. subtilis*, 500 µL of an overnight culture were centrifuged (1 min, RT, 16,000 x*g*). The supernatant was discarded and the pellet ressupended in 100 µL of a 0.85% (w/v) solution of NaCl containing 5 µL of 10 mg/mL RNAse (10 mg/mL; ThermoFisher Scientific) and 2 µL of chicken egg white lysozyme (50 mg/mL Sigma-Aldrich). In order to promote cell lysis and RNA degradation, the samples were incubated at 37 ºC for 5 min. Complete cell lysis was ensured by the addition of 300 µL of CLS solution (Appendix A). To precipitate proteins and cell debris, 168 µL of the PPS solution (Appendix A) were added and the samples were vortexed for 10 s. The precipitated proteins were centrifuged (10 min, 4ºC, 16,000 x*g*), and 600 µL of the supernatant transferred to a clean 2 mL Eppendorf tube. DNA was precipitated by the addition of 600 µL of isopropanol and inverting the sample six times. The DNA was separated from solution by centrifugation (5 min, RT, 16,000 x*g*) and washed twice

with 70% (v/v) ethanol. The ethanol was completely removed by drying the samples for 10 min at 55ºC. The purified DNA was solubilized in 100 µL of milliQ water.

2.5.3 Oligonucleotides

Oligonucleotides were designed using the Clone Manager software (SciEd) and purchased from IDT (Leuven, Belgium). All oligonucleotides were stored upon arrival at -20°C as 100 µM solutions in sterilised deionized water.

A detailed list of all primers used in this study is given in Appendix C.

2.5.4 Polymerase chain reaction (PCR)

DNA was amplified *in vitro* via the Polymerase Chain Reaction (PCR) using the High-Fidelity DNA polymerase Phusion® (New England Biolabs) (Mullis & Faloona 1987). The PCR reactions were set up as shown in Table 2.3. The thermocycling conditions for a routine PCR are described in Table 2.4.

Table 2.3 Components of a standard PCR.

Component	**Volume**	**Final Concentration**
5X Phusion HF Buffer	10 µL	1X
10 mM dNTPs	1 µL	200 µM
10 µM Forward Primer	2.5 µL	0.5 µM
10 µM Reverse Primer	2.5 µL	0.5 µM
Template DNA	Variable	< 250 ng
Phusion DNA Polymerase	0.5 µL	1.0 units/50 µL PCR
Nuclease-free water	To 50 µL	

Table 2.4 Thermocycling conditions for a routine PCR.

Step	Temperature (°C)	Time	
Initial Denaturation	98	30 s	
Denaturation	98	10 s	30 cycles
Annealing	45-72	20 s	
Extension	72	15 s per kb	
Final extension	72	300 s	

2.5.5 Purification of PCR products

PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN), according to the manufacturer's instructions. This kit uses a column-based system which contain a silica membrane to which plasmid DNA can be bound, washed and eluted. The DNA was eluted in sterile deionized water.

2.5.6 Electrophoresis of DNA

Electrophoresis of DNA was performed according to standard methods (Sambrook & Russell 2001) using 1% agarose gels. Nancy-520 (Sigma-Aldrich), fluorescent stain for double stranded DNA, was added to the agarose gels while still liquid. Prior to loading the gel, DNA samples were mixed with an appropriate volume of 6X Purple Gel Loading Dye (New England Biolabs). Electrophoresis was preformed in TAE buffer at 90-110 V. After electrophoresis, the gels were viewed by UV transillumination and photographed using a gel documentation system (Bio-Rad).

2.5.7 Purification of DNA from agarose gel

After electrophoresis of DNA (Section 2.5.5) fragments of interest were excised from the gel using a clean scalpel and transferred to a 2 mL Eppendorf tube. The DNA was purified from the agarose gel slices using the QIAquick Gel Extraction Kit (QIAGEN) according to the manufacturer's instructions. The DNA was eluted in sterile deionized water.

2.5.8 Digestion of DNA with restriction endonucleases

All restriction enzymes used in this study were obtained from New England Biolabs. DNA was digested from 30 min up to 3 hours in the conditions suggested by the manufacturer and using the buffers provided with the enzymes. The digestion products were either analysed by DNA electrophoresis (Section 2.5.6) or purified from the remaining components of the restriction digestion reaction using the QIAquick PCR Purification kit (Section 2.5.5).

2.5.9 Ligation of DNA fragments

DNA fragments were ligated following restriction digestion (Section 2.5.8) and DNA purification (Section 2.5.5). The ligation reactions were catalysed by the T4 DNA ligase (New England Biolabs) in the conditions recommended by the manufacturer. The ligation reaction volume was typically 20 μL and contained the vector and insert DNA in an approximate molar ratio of 1:3. The ligation mixtures were incubated overnight at 4°C and then used to transform *E. coli.*

2.5.10 Gibson assembly technology

The Gibson Assembly technology was used to assemble multiple DNA fragments with overlapping regions in a single-tube isothermal reaction (Gibson *et al.* 2009; Gibson 2011). Typically, 20 base pairs of homology between adjacent fragments were designed.

Gibson Assembly® Master Mix was purchased from New England Biolabs and used for all assembly reactions. The reaction master mix was diluted in an equal volume of the DNA fragments to be assembled. When transforming *E. coli*, assemblies were done using 50 ng of vector and, when transforming *B. subtilis*, the amount of vector DNA was 500 ng. Insert DNA was combined with the vector in a 3:1 molar ratio. The total volume of this vector and insert(s) mixture corresponded to the volume of the commercial master mix added to catalyse the assembly reaction.

All reactions were incubated at 50ºC for 15 min when 2 or 3 fragments were being assembled, or 60 min when more fragments were assembled (Biolabs 2012). The total volume of the assembly reaction was used to transform either *E. coli* NEB10 (Section 2.6.2) or *B. subtilis* (Section 2.6.3).

2.5.11 StarGate® cloning system

The StarGate® cloning system (IBA) was used for direct transformation of *B. subtilis* with ligation products. This technology combines the restriction digest and ligation steps in one isothermal single-tube reaction. The cloning system makes use of a type IIS restriction enzyme that cleaves outside of its continuous and asymmetric recognition site, allowing the design of unique overhangs and assembly of multiple DNA fragments in an orchestrated order.

StarGate® vectors were designed with an interchangeable chloramphenicol cassette for counter selection of clones with successful ligation of inserts into the vector. The inserts were designed in order to obtain unique overhangs that allow fragment assembly in the desired order.

StarGate® reactions were prepared on ice as described in Table 2.5. All solutions were provided in the Direct Transfer Cloning kit from IBA (discontinued). Reactions were incubated for 3 hours at 30ºC before being transformed into competent *B. subtilis*.

Table 2.5 Components of a StarGate® reaction.

Component	**Volume (µL)**
Star Solution A1	1
Star Solution A2	1
Star Solution A3	1
Formulation Buffer (10X)	2.5
Destination Vector (4 nM)	4
Insert(s) (4 nM)	4
Deionized H_2O	up to 25

2.5.12 DNA sequencing

Confirmation of DNA sequences was outsourced to Baseclear (Leiden, Netherlands). The Quick Shot Sanger sequencing service was typically used providing a premix of primer (10 pmol/µL) and purified DNA (50-200 ng/µL). The sequencing chromatograms were analysed using the free software Chromas, and the sequences were confirmed using Clone Manager's alignment features.

2.6. DNA transformation

2.6.1 Transformation of Escherichia coli – $CaCl_2$ method

Preparation of E. coli competent cells. *E. coli* TOP10 cells were transformed with DNA using the $CaCl_2$ method. *E. coli* TOP10 competent cells were prepared by inoculating 400 mL LB with 4 mL of a fresh overnight culture and incubating at 37 °C until the culture reached an optical density at 600 nm (OD_{600}) of ~0.3. Then the cells were cooled on ice and harvested by centrifugation (8 min, 4ºC, 3,800 x*g*) in a pre-cooled centrifuge. After the supernatant was discarded, 20 ml of 100 mM $CaCl_2$ was added to each pellet, followed by incubation on ice for 30 min. Afterwards, the pellet was ressuspended slowly with a cut tip. The cells were harvested once again (8 min, 4ºC, 3,800 x*g*) in a pre-cooled centrifuge. The supernatant was discarded and the pellet ressuspended with 2 mL of a 100 mM $CaCl_2$ solution supplemented with15% glycerol. The cells were incubated for 15 min on ice and then 100 µL aliquots were prepared into 1.5 mL Eppendorf tubes. Finally, the competent *E. coli* stocks were snap-frozen in liquid N_2 and stored at -80˚C until needed.

DNA transformation. For DNA transformation of *E. coli* TOP10 competent cells, the ligation mixture or ~100 ng of plasmid DNA was added to the frozen competent stocks. The cells were then incubated on ice for 20 min. After this, the mixture was heat shocked for 90 s at 42˚C and then on ice again for 2 min. After, 400 µL of fresh LB were added to the cells and incubated at 37˚C for between 30 and 60 min. Finally, the cells were harvested by centrifugation (1 min, RT, 12,000 x*g*), resuspended in 100 µL of fresh LB and plated on selective agar medium.

2.6.2 Transformation of Escherichia coli NEB 10-beta

When indicated, commercial NEB 10-beta competent *E. coli* cells (New England Biolabs) were used for transformation, particularly for Gibson Assembly reactions (Section 2.5.10).

A tube of competent cells was thawed on ice for 10 min. Up to 100 ng of DNA was added to the cell mixture followed by gentle mixing and incubation on ice for 30 min. The cells were heat shocked by placing the mixture at 42ºC for 30 s and immediately transfer to ice for 5 min. LB medium was added to the mixture (400 µL) and the cells were placed at 37ºC for 60 min with vigorous shaking (250 rpm). The transformed cells were concentrated by centrifugation (1 min, RT, 8,000 x*g*) and resuspended in 100 µL of fresh LB medium. The entire concentrated cell mixture was plated on selective agar medium and incubated at 37ºC overnight.

2.6.3 Transformation of Bacillus subtilis using natural competence

Transformation of *B. subtilis* was performed using a modified version of the Anagnostopoulos & Spizizen (1961) method which is based on the natural competence of vegetative cells.

An overnight culture was prepared by inoculating 2 mL of LB medium, supplemented with the appropriate selection markers, with a single colony from a selective agar plate or a frozen stock of a cell culture. The bacterial culture was incubated overnight in a shaking incubator (250 rpm) at 37˚C. After overnight incubation, the culture was diluted in 20 mL of Spizizen-plus medium (Appendix A) to an OD_{600} of 0.1 in a 100 mL shake flask. The cells were grown in a shaking incubator (250 rpm) at 37˚C until an OD_{600} of 0.4-0.6 was reached. The exponentially growing culture was diluted 1:1 with Spizizen-starvation medium (Appendix A) and incubated at 37˚C and 250 rpm for further 1.5-2 hours. The competent cells were concentrated by centrifugation (10 min, RT, 3,000 x*g*), 90% of the supernatant removed and the pellet was resuspended in the remaining volume. At this point the cells were either used immediately for transformation or stored for future use. For immediate transformation, the cells were incubated in 2 mL Eppendorf tubes with 0.05 – 1 µg of DNA, at 37˚C and 250 rpm shaking speed for 1 hour. For storage of competent cells, the culture was supplemented with 10% glycerol and 100 µL aliquots were prepared into 2 mL cryotubes. The aliquots were then flash frozen in liquid N_2 and stored at -80 ˚C. Frozen competent cells were used up to 3 months after storage by gentle thawing the stock on ice followed by incubation at 37˚C and 250 rpm, for 1 hour, with 0.05 – 1 µg of DNA. After incubating the fresh or frozen competent cells with DNA for the development of antibiotic resistance, the cells were plated on selective agar plates and incubated overnight at 37˚C.

2.7. Quantitative reverse transcription PCR (RT-qPCR)

Quantitative reverse transcription PCR (RT-qPCR) was used to detect RNA expression of specific genes of interest in *B. subtilis*. The analysis was performed with biological and technical duplicates and using two reference genes (*fbaA* and *sdhA*) for sample standardization, according to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) (Bustin *et al.* 2009). The wild type strain 168 was used as the relative-quantification reference.

2.7.1 Harvesting of samples

In order to be able to relatively quantify and standardise the expression levels of target genes, all samples were collected at the same point of growth by synchronisation of all cell cultures.

Overnight cultures of the target and wild type strains were prepared in LB medium supplemented with the appropriate antibiotics and grown at 37°C with shaking (250 rpm). After overnight growth, pre-cultures were inoculated to an OD_{600} of 0.05 and grown until mid-exponential phase (OD_{600} between 0.4 and 0.8). The pre-cultures were diluted, in triplicate, to pre-warmed LB to an OD_{600} of 0.02. The growth was followed every 15 min until the cells reached mid-exponential phase (OD_{600} between 0.4 and 0.8). Samples were harvested by adding 2 volumes of RNAprotect® (QIAGEN) to 1 volume of culture, vortexing and incubating for 1 min at room temperature. The samples were divided between 2 mL Eppendorf tubes, centrifuged (2 min, RT, 16,000 x*g*), the supernatant discarded, and the pellets stored at -80ºC until proceeding with RNA extraction.

2.7.2 RNA purification

Total RNA from *B. subtilis* cells was extracted using the QIAgen RNeasy Mini Kit (QIAGEN) with an adaptation of the manufacturer's instructions. RNAseZap® (Thermo Fisher Scientific) solution was used to clean pipettes, benches and gloves to avoid contamination of samples with RNase.

The cell pellets were thawed on ice and resuspended in 1 mL of RLT buffer (Appendix A) with added β-mercaptoethanol (10 μL of β-mercaptoethanol per 1 mL of RLT buffer). The suspension was transferred to tubes containing 0.1 mm silica beads (Lysing Matrix B, MP Biomedicals) and the cells were mechanically lysed in the Precellys®24 homogenizer in 3 cycles of 10 s agitation at

6500 rpm. The lysis samples were centrifuged (10 min, 4 ºC, 25,000 x*g*) in a pre-cooled centrifuge, and the resulting cleared lysates transferred to 2 mL Eppendorf tubes. Before adding the samples to individual mini columns from the RNeasy kit, 600 µL of 96-100% ethanol was added to the lysates. The columns, loaded with up to 700 µL of sample, were centrifuged (15 s, 4 ºC, 8,000 x*g*). The remaining lysates were loaded onto the columns and centrifuged (15 s, 4 ºC, 8,000 x*g*), the flow-through discarded and 700 µL of Buffer RW1 (Appendix A) was added to the individual columns. The samples were again centrifuged (15 s, 4 °C, 8,000 x*g*), and washing step with Buffer RW1 (Appendix A) repeated once more. The flow-through and the collection tubes were discarded and the columns were placed in fresh 2 mL collection tubes. Similar to the previous step, 500 µL of the Buffer RPE (Appendix A) were added to each column and centrifuged twice (15 s, 4 °C, 8,000 x*g*). The flow-through was discarded and the individual columns were placed in a fresh collection tube and centrifuged (2 min, 4 °C, 16,000 x*g*). The dry columns were placed in a fresh 1.5 mL Eppendorf tube and the total RNA was eluted from each column by centrifugation (1 min, 4 °C, 8,000 x*g*) with 50 µL RNAse free water. The elution step was repeated with the volume of the first elution. The purified RNA samples were stored at -20°C for further sample process in the following day, or at -80°C for long-term storage.

After total RNA extraction, DNA contamination was removed by treatment with the RNase-Free DNase Set (QIAGEN), with an adaptation of the manufacturer's instructions. The entirety of the eluted RNA was used for this DNase treatment by addition of 10 µL of Buffer RDD (Appendix A), 2.5 µL of DNase I stock solution (Appendix A) and 40.5 µL of RNase free water. The samples were incubated on the benchtop (20-25ºC) for 10 min. The treated RNA was cleaned again using the RNAeasy mini kit columns (QIAGEN), as described above.

2.7.3 Reverse transcription

The cDNA Reverse Transcription (RT) reactions were performed with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems), according to the manufacturer's instructions. For all samples a control without RTase enzyme was included, to assess for genomic DNA amplification during quantitative PCR. The RT reactions were performed in a total volume of 40 µL in a standard thermos cycler. The RT products were diluted 10 times with deionized H_2O and stored at 4ºC for a short-term and at -20ºC for a long-term.

2.7.4 Quantitative PCR

The quantitative PCR (qPCR) reactions were performed using the 2X iQ SYBR® Green Supermix (Biorad) in the iQ5 Real-Time PCR Detection System (Biorad). The components of the qPCR reactions are described in Table 2.6.

Table 2.6 Components of a qPCR reaction.

Component	Volume (µL)
2X iQ SYBR® Green Supermix	12.5
Primer 1	0.5
Primer 2	0.5
Template	5
Deionized H_2O	6.5

As mentioned previously, each analysis was performed with biological and technical duplicates and using two reference genes for sample standardization, according to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) (Bustin *et al.* 2009). The wild type strain 168 was used as the relative-quantification reference. All reactions were performed in parallel with the reverse transcription negative controls (without RTase enzyme) to check for amplification of genomic DNA. A non-template control (NTC) was included for each primer pair to check for DNA contamination in the primer stock.

The reference genes chosen for sample standardization were *fbaA* and *sdhA*, both encoding for proteins involved in the carbon core metabolism (Michna *et al.* 2015). The *fbaA* gene encodes fructose 1,6-biophosphate aldolase, which is a constitutively expressed enzyme involved in glycolysis and gluconeogenesis (Ludwig *et al.* 2001; Commichau *et al.* 2013). The *sdhA* gene encodes for a constitutively expressed succinate dehydrogenase, part of a respiratory supercomplex (Michna *et al.* 2015; Sousa *et al.* 2013). These genes exhibit high expression levels in the conditions of the experiments, according to their dedicated SubtiExpress pages (Michna *et al.* 2015).

Primers for the Quantitative PCR (qPCR) were designed using the Primer3Plus online tool (Untergasser *et al.* 2007). The desired amplicon length was set to 100-150 nucleotides. The

primers' desired melting temperature and length was set at 57-58°C and 18-20 nucleotides, respectively. Two pairs of primers were designed for each target and reference gene and one pair was selected by comparing the primers' calibration curves. All calibration curves were calculated by performing a set of qPCR reactions using a dilution series of genomic DNA from 0 to 0.24 ng/µL, in triplicate. The amplification efficiency of each primer pair was determined from the slope of the log-linear portion of the calibration curve – Cq versus log_{10}DNA (Bustin *et al.* 2009). The selected primer pairs had efficiencies of between 95 and 105%. Additionally, the melting peaks were analysed and only primer pairs with a single peak corresponding to the melting temperature of the product of interest were selected. The selected primers for the target and control genes are listed in Appendix C. The amplification efficiencies and melting temperatures for each qPCR primer pair are listed in Appendix G (Table G.6).

2.7.5 Data processing

Each qPCR reaction yielded a quantification cycle (C_q) for each sample and a melting temperature (T_m) for the amplified product. The C_q values facilitate a comparison of the expression levels of the target genes and the T_m values confirm the amplification of the correct product.

The $\Delta\Delta C_q$ method was used to determine differences in concentrations between samples normalised with a reference gene. The difference in C_q values (ΔC_q) between the target gene and the reference gene was calculated for each pair of averaged technical replicates of the target strain (ΔC_q1) and the relative-quantification reference strain (ΔC_q1) – Equations 2-1 and 2-2, respectively.

$$\Delta C_q = C_q(target\ gene, target\ strain) - C_q(reference\ gene, target\ strain) \quad \text{Equation 2-1}$$

$$\Delta C_q = C_q(target\ gene, 168\ strain) - C_q(reference\ gene, 168\ strain) \quad \text{Equation 2-2}$$

The normalized target gene expression level was calculated for the target and wild type strains using Equation 2-3, which assumes 100% of amplification efficiency.

$$Normalized\ target\ gene\ expression\ level = 2^{-\Delta C_q} \quad \text{Equation 2-3}$$

Finally, the normalized target gene relative expression level was calculated using Equation 2-4.

$$Normalized\ target\ gene\ relative\ expression\ level = \frac{2^{-\Delta C_q 1}}{2^{-\Delta C_q 2}}$$ Equation 2-4

The standard deviation of the final relative expression value was calculated relative to the average of all possible combinations between the biological replicates of the target and wild type strains.

2.8. Protein Analysis by SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

For protein visualization and quantification, secreted proteins were precipitated or collected from the culture supernatant and analysed by SDS-polyacrylamide gel electrophoresis (PAGE).

2.8.1 Protein sample preparation

Secreted proteins were collected of culture supernatants by centrifugation (5 min, 4 ºC, 13,000 x*g*), unless otherwise stated. Protein samples for SDS-PAGE analysis were prepared by either precipitation of the total proteins in the sample or collection of a representative smaller sample.

The total proteins in a given sample were precipitated by overnight incubation on ice with 10% (w/v) trichloroacetic acid. After precipitation the sample was centrifuged (15 min, 4 °C, 20,000 x*g*). The supernatant was discarded and the pellet was washed three times with 100% ethanol, performing a centrifugation (15 min, 4 °C, 20,000 x*g*) each time. The final pellet was dried at 60°C for approximately 20 min. The pellet was dissolved in 200 µL of a urea (8 M) /thiourea (2 M) solution by vigorous vortexing. The samples were centrifuged (30 min, RT, 20,000 x*g*) and the supernatant transferred to a clean tube. At this point the protein concentration was determined using the 2-D Quant Kit (GE Healthcare Life Sciences), according to the manufacturer's instructions.

When stated, smaller samples of the cultures supernatants containing the secreted enzymes were collected. These samples were typically between 15 and 20 µL. Sample buffer containing SDS and a reducing agent was added directly to the sample prior to electrophoresis (Section 2.8.2).

2.8.2 Separation of proteins via SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS- PAGE was performed using the NuPAGE® Bis-Tris system from ThermoFisher Scientific. NuPAGE® Bis-Tris gels with 1.0 mm thickness and a gradient of 4-12% of polyacrylamide concentration. Samples were prepared according to Table 2.7 using the appropriate volume of protein sample each time.

Table 2.7 Sample preparation for SDS polyacrylamide gel electrophoresis using NuPAGE® Bis-Tris gels.

Components	Volume (μL)
Sample	x
NuPAGE® LDS Sample Buffer (4X)	6.25
NuPAGE® Reducing Agent (10X)	2.5
Deionized Water	to 25
Total Volume	25

Samples were boiled for 5 min at 100°C and then loaded onto the gel. The gels were electrophoresed at 200 V for 35 min using NuPAGE® MES SDS running buffer. Typically, the SeeBlue® Plus2 Prestained Standard (ThermoFisher Scientific) was used as protein molecular size marker.

2.8.3 InstantBlue™ protein staining

The InstantBlue™ protein stain (C.B.S. Scientific) was used for staining of SDS gels given its high sensitivity and fast protocol. This staining solution removes SDS, fixes and stains in one quick step after protein electrophoresis, offering a very low background staining.

To stain a SDS gel with the InstantBlue™ solution, the gel was first gently washed with deionized water in a clean plastic container after electrophoresis (2.8.2). The staining solution was added onto the gel, using sufficient to just cover it. The container was then placed on a rocking platform for at least 15 min. After this short staining period, the gel was imaged in the ChemiDoc MP system (BioRad).

2.8.4 SYPRO® Ruby protein staining

For ultrasensitive detection of proteins, the SYPRO® Ruby protein stain was used. This stain has a detection limit as low as 0.25 ng and results in a bright fluorescent signal that can be quantified using the ChemiDoc MP system (BioRad).

After protein electrophoresis (Section 2.8.2) the gel was placed in a clean microwavable container and gently washed with deionized water. In order to fix the proteins on the gel, 100 mL

of the Fix and Wash Solution (Appendix A) were added. The container was placed on a rocking platform for 15 min. After repeating this step with fresh Fix and Wash Solution, 60 mL of SYPRO® Ruby stain solution was added onto the gel. The gel was then microwaved for 30 s at full power, agitated for 30 s to distribute heat evenly, microwaved again for 30 s to 80-85°C and agitated for further five min. The gel was reheated by microwaving a third time for 30 s and then agitated for 23 min in a rocking platform. Finally, to avoid heating the de-stain solution and reduce stain sensitivity, the gel was transferred to a clean container and washed in 100 mL of the Fix and Wash Solution (Appendix A) for 30 min on a rocking platform. Before imaging in the ChemiDoc MP system (Biorad), the gel was washed three times with ultrapure water to prevent corrosion of the imager.

2.8.5 Detection of fluorescent proteins with the laser scanner Typhoon 9410

Fluorescent proteins were detected and quantified after SDS-PAGE (Section 2.8.2) using the laser scanner Typhoon 9410 and the software ImageQuant TL (GE Healthcare Life Sciences).

A green laser (532 nm) was used as the excitation source with intensity between 600 and 900 V. Emission filters were selected accordingly to the emission wavelength of the fluorophore being analysed.

2.8.6 Detection of proteins via Western blotting

Protein detection by Western blotting was executed using the iBlot® 7-Min Blotting System (ThermoFisher Scientific), according to the manufacturer's instructions. The pre-electrophoresed protein gel (Section 2.8.2) was assembled with the iBlot™ Gel Transfer Stacks, containing the nitrocellulose transfer membrane, on the iBlot™ Gel Transfer Device. During the assembling process, trapped air bubbles were carefully removed with the provided Blotting Roller to assure efficient protein transfer to the membrane. The blotting was performed at 25 volts for 7 min.

The blot was then transferred to a 50 mL Falcon® tube placing the front side of the membrane, which was in contact with the gel, to the inside of the tube. The membrane was blocked with 20 mL of PMT (Appendix A) for 2 hours at room temperature in a Tube Roller Mixer (Stuart®). After blocking the membrane, the primary antibody was diluted 1:2000, unless otherwise stated, in 5 mL PMT (Appendix A). The antibody solution was added onto the membrane and incubated for 2 hours at room temperature with constant mixing. The unbound primary antibody was removed with four washes of 20 mL of PMT (Appendix A) each with five min of mixing. The

secondary antibody was diluted 1:10,000 in 5 mL PMT (Appendix A) and incubated for 60 min at room temperature with constant mixing. Finally, the membrane was washed six times with PBS (Appendix A).

The sary antibody bound to the membrane was detected with the SuperSignal Chemiluminescent HRP Substrate (ThermoFisher Scientific) according to the manufacturer's instructions, and visualized in the ChemiDoc MP system (BioRad).

2.9. Pulse chase and immunoprecipitation of secreted enzymes

The secretion kinetics of two industrially relevant model enzymes was studied using radioactively labelled pulse-chase experiments. In these experiments, cells producing a relevant enzyme are grown in modified Spizizen Minimal Medium (SMM, Appendix A) using 1% ribose as the carbon source. Ribose was chosen as the culture carbon source because it did not induce any significant catabolite repression of enzyme expression, and still permitted relatively high growth rates.

At a stage of high enzyme production during growth, all synthesized proteins were labelled with radioactive methionine during a short period of time – the pulse – after which an excess of unlabelled methionine was added – the chase. This strategy limits the radioactive signal to all proteins produced during the short pulse, which gives valuable information about the secretion kinetics and bottlenecks.

Two different protocols were designed to study secretion kinetics of the endo-1,4-β-xylanase, XynA, from *B. subtilis* and the α-amylase, AmyM, from *Geobacillus stearothermophilus.* The protocol designed for XynA involves an immunoprecipitation step and analyses both the cell free supernatant and the entire culture of producing cells. The protocol designed for AmyM focuses solely on the supernatant fraction of the cultures.

2.9.1 Pulse-chase and immunoprecipitation of XynA from Bacillus subtilis

Cells producing XynA were inoculated, either from a single colony or a glycerol stock, into 2 mL of LB supplemented with neomycin (50 μg/mL) and grown at 37°C, 250 rpm, during 8 hours. The culture was diluted 10^{-6}, 10^{-7} and 10^{-8}-fold into separate 15 mL lots of SMM (Appendix A), pre-warmed to 37°C. This series of dilutions was incubated overnight at 37°C, 250 rpm.

The optical density of the overnight cultures was measured, and the experiment proceeded with the culture that was in the exponential phase of growth. The selected culture was incubated in a water bath at 37°C. At OD_{600} 0.8 newly synthesized proteins were labelled for one min by addition of 100 μCi of [^{35}S] labelled L-methionine (185 MBq, 1175 Ci/mmol). Immediately after one min, 600 μL of a solution of 25 mg/mL of L-methionine was added to quench the culture. This step corresponds to the start of the so-called "chase" period and two samples were collected immediately after (0 min) and at time points thereafter (0.5, 1, 3, 5 and 10 min).

One of the samples was added directly into 600 μL of 20% (w/v) trichloroacetic acid (TCA) to precipitate the total protein content of the culture. The s sample was filtered (0.45 μm PVDF filters, Whatman), before being added to 20% (w/v) TCA, to separate the proteins in the extracellular environment of the culture. The samples were incubated on ice for 1 hour followed by centrifugation (15 min, 4ºC, 20,000 x*g*,). The pellets were washed with 1 mL of acetone for removal of residual TCA and dried under vacuum.

The samples were resuspended in 100 μL of Lysis Buffer (Appendix A) and incubated at 37°C for 15 min. To solubilize the precipitated proteins, the samples were boiled for 10 min after adding 10 μL of 10% (w/v) Sodium Dodecyl Sulfate (SDS). The samples were then cooled to room temperature. Cell debris was precipitated by adding 900 μL of 1X STD (Appendix A) and the samples then placed on ice for 30 min. The cellular debris was pelleted by centrifugation (5 min, RT, 20,000 x*g*). The supernatants were transferred to clean 1.5 mL Eppendorf tubes.

XynA was immunoprecipitated at 4°C overnight in the presence of 10 μL of anti-serum and 100 uL of a water suspension of 5 mg/mL of Protein A Sepharose beads. The Sepharose beads were pelleted by centrifugation (2 min, RT, 16,000 x*g*) after overnight incubation and washed four times with 100 uL of 1X STD (Appendix A). After the final wash, the beads were resuspended in 7 μL of Sample Buffer (Appendix A) and the immunoprecipitated xylanases were released from the Sepharose beads by boiling for 5 min. The enzyme sample was analysed by SDS-PAGE (Section 2.8.2) followed by Phosphor Imaging (Section 2.9.3).

2.9.2 Pulse-chase of secreted AmyM from Geobacillus stearothermophilus

Cells producing AmyM were inoculated, either from a single colony or a glycerol stock, into 2 mL of LB supplemented with neomycin (50 μg/mL) and grown at 37°C, 250 rpm, during 8 hours.

The grown culture was diluted 10^{-6}, 10^{-7} and 10^{-8}-fold into separate 15 mL lots of SMM (Appendix A), pre-warmed to 37°C. This series of dilutions was incubated overnight at 37°C, 250 rpm.

The optical density of the overnight cultures was measured, and the experiment proceeded with the culture that was in the late exponential phase of growth. The selected culture was incubated in a water bath at 37°C. At OD_{600} 1.5 newly synthesized proteins were labelled for one min by addition of 100 μCi of [^{35}S] labelled L-methionine (3.7 MBq, 1175 Ci/mmol). Immediately after one min, 600 μL of a solution of 25 mg/mL of L-methionine was added to the culture. In contrast to protocol for the XynA pulse chase, only one sample was taken for analysis immediately after (0 min) and at time points thereafter (0.5, 1, 3, 5 and 10 min).

The sample was filtered (0.45 μm PVDF filters, Whatman), before being added to 20% (w/v) TCA, for extracellular protein precipitation. The samples were incubated on ice for 1 hour followed by centrifugation (15 min, 4ºC, 20,000 x*g*). The pellets were washed with 1 mL of acetone for removal of residual TCA and dried under vacuum.

The samples were ressupended in 74 μL of Lysis Buffer (Appendix A) and incubated at 37°C for 15 min. To solubilise and prepare the proteins for SDS-PAGE, 26 μL of Sample Buffer (Appendix A) was added to each sample followed by boiling for 5 min.

The enzyme samples were analysed by SDS-PAGE (Section 2.8.2) followed by Phosphor Imaging (Section 2.9.3).

2.9.3 Phosphor Imaging of proteins labelled with [^{35}S] methionine

The signals of proteins labelled with [^{35}S] methionine during pulse-chase experiments, and electrophoresed on SDS-PAGE gels, were visualized by phosphor imaging. The gels were thoroughly washed with distilled water to remove any unincorporated of radioactive label. Using a vacuum dryer, the gels were dried into filter paper for one hour. The dried gels were placed in a Phosphor Imaging screen and exposure cassette. The radioactive signal in the phosphor atoms present in the labelled proteins was stored in the imaging screen for at least 16 hours. The scanner Typhoon 9410 (GE Healthcare Life Sciences) was used to stimulate the release of the stored energy with visible light and convert the luminescence resulting signal to a digital image. The

software ImageQuant TL (GE Healthcare Life Sciences) was used to quantify the bands of interest.

2.10. Activity analysis of industrially relevant enzymes

2.10.1 Analysis of 1,4-beta-xylanase activity

Endo-1,4-β-xylanases catalyse the hydrolysis of 1,4-β-D-xylosidic linkages in xylans (Gasteiger *et al.* 2003). Relative enzyme activity of the secreted endo-1,4-beta-xylanase, XynA, from *B. subtilis* was determined using the fluorescence based assay EnzChek® Ultra Xylanase Assay Kit (Thermo Fisher Scientific), according to the manufacturer's instructions. In this assay, the hydrolysis of xylosidic linkages within the included hemicellulose polysaccharide substrate results in the unquenching of the bound fluorescent dyes. Fluorescence is then measured using a microplate reader and is a function of the xylanase activity in the sample (Molecular Probes 2007). All assays were performed in triplicate and three or more biological replicates were used.

The working solution of the Xylanase Reaction Buffer (Appendix A) was prepared using the provided 10X stock solution in the commercial kit. A Substrate Solution of 1 mg/mL was prepared by adding 250 μL of the Xylanase Reaction Buffer to the lyophilised Xylanase Substrate vial (250 μg). Prior to performing the assay, a 50 μg/mL Substrate Solution was prepared in Xylanase Reaction Buffer.

The supernatant of cell cultures producing the secreted xylanase were collected by centrifugation of 20 μL of culture in a 96-well non-skirted clear PCR plate (5 min, RT, 3,000 x*g*,). The supernatant samples were diluted 1:10 in 50 μL of Xylanase Reaction Buffer into a black flat bottom 96-well MTP. Three wells were filled with 50 μL of growth medium to serve as the reaction blanks.

The enzymatic reaction was initiated by simultaneously adding 50 μL of the Xylanase Substrate Solution to each well. The reaction plate was incubated in the dark at 21°C, for 30 min and the reaction stopped by the simultaneous addition of 70 μL of Stopping Buffer (Appendix A) to each well. The fluorescent reaction product was quantified using the Infinite® M200 PRO plate reader (Tecan) with an excitation maxima of 358 nm and an emission maxima of 455 nm, a gain of 65 and 20 fluorescent reads per well. Only fluorescent measurements that fitted the detection limits of the plate reader were considered.

2.10.2 Determination of 1,4-beta-xylanase enzymatic units

One Unit of Xylanase Activity (U) was defined as the amount of enzyme catalysing the conversion of 1 μmol of xylose from xylan per min, at 21°C and in 100 mM sodium acetate buffer with pH 4.6.

In order to estimate xylanase Enzyme Units in *B. subtilis* cultures, a correlation was determined between the xylanase enzymatic assay (Section 2.10.1) and the conversion of xylose in the presence of xylan. Xylose was quantified using the DNS method (Miller 1959). In this method, the enzymatic reaction takes place in the presence of 3,5-dinitrosalicylic acid. This acid reacts with reducing sugars to form 3-amino-5-nitrosalicylic acid which absorbs light at 540 nm (Miller 1959).

A xylanase Enzymatic Units Standard Curve (Appendix D, Figure D.5) was obtained by using a xylanase standard sample and a xylose standard curve to correlate the fluorescent values obtained with the commercial enzymatic assay (RFU) and the amount of converted xylose, in defined conditions.

The strain BRC55 (see Section 3.2. , that carries a high copy number plasmid expressing the gene encoding for XynA of *B. subtilis* was inoculated from a single colony in 30 mL of LB medium supplemented with neomycin (50 μg/mL). After overnight growth at 37 °C, the cells were harvested by centrifugation (10 min, RT, 3,000 x*g*). The cell-free culture supernatant was used as the crude enzyme source for an enzyme standard sample.

A series of 12 dilutions of the xylanase enzyme standard was prepared by consecutive dilutions of the sample in an equal volume of Xylanase Reaction Buffer (Appendix A). The enzymatic activity in each dilution was measured, in triplicate, using the xylanase activity of the commercial assay (Section 2.10.1) and the results were plotted to obtain the EnzChek® Standard Curve (Appendix D, Figure D.1).

The xylanase standard dilutions were also used to quantify the release of xylose from xylan in the presence of the DNS reagent (Appendix A). A substrate solution was prepared by dissolving xylan in Xylanase Reaction Buffer to a final concentration of 1% (w/v). The enzymatic reaction was performed in the same conditions as the commercial assay by adding 50 μL of the xylan substrate solution to 50 μL of the xylanase standard dilutions, in triplicate, in a 96-well flat

bottom MTP. The reaction was incubated at 21 °C for 30 min. The converted xylose was quantified by transferring 50 μL of the reaction to 50 μL of DNS reagent in a 96-well flat bottom MTP. The samples were heated for 10 min at 99 °C in an MTP Eppendorf incubator. Immediately after the heating period, 150 μL of milliQ water was added to each well. The plate was then placed on ice for 10 min before measuring the absorbance at 540 nm in a microplate reader. The results were used to obtain the DNS Standard Curve, which correlates the converted xylose absorbance and the correspondent absorbance values obtained with the DNS method (Appendix D, Figure D.2).

Since the EnzChek® and the DNS Standard Curves were obtained with the same enzyme standard samples, they were combined into a single, correlation denominated, Xylanase Enzymatic Reaction Standard Curve (Appendix D, Figure D.3).

A Xylose Standard Curve (Appendix D, Figure D.4) was calculated by preforming the DNS method in the same conditions as above using 50 μL of xylose solutions in a range of concentrations.

The Xylanase Enzymatic Units Standard Curve (Appendix D, Figure D.5) was derived by combining the Xylose Standard Curve and the Xylanase Enzymatic Reaction Standard Curve through the common variable of absorbance at 540 nm. The resulting value was divided by the duration of the assay in min (30), the duration of the assays in min, to obtain a direct correlation between the fluorescence measured with the commercial enzymatic assay and the Xylanase Enzymatic Units (U):

$$U = 4.83\times10^{-8}\times RFU \qquad \text{Equation 2-5}$$

2.10.3 Analysis of α-amylase activity

α-Amylases catalyse the hydrolysis of 1,4-α-D-gluocosidic linkages in polysaccharides containing three or more D-glucose units, such as starch and glycogen (Gasteiger *et al.* 2003). The relative enzyme activity of the secreted α-amylase from *Geobacillus stearothermophilus* was determined using the colorimetric assay Phadebas® Amylase Test Tablets (Magle Life Sciences). In this assay, a water-insoluble cross-linked starch polymer carrying a blue dye is used as substrate. Substrate hydrolysis in presence of a α-amylase releases the blue dye in proportion to

the activity of α-amylase in the sample. All assays were performed in triplicate and three or more biological replicates were used.

The protocol suggested by the manufacturer was adapted to all the reactions to be performed in a 2 mL Eppendorf and in a 96-well transparent flat bottom MTP. These assays formats served the purpose of routine activity checks and high throughput quantifications of relative enzyme activity, respectively. All assays were performed at pH 5.5 using the Amylase Reaction Buffer (Appendix A).

MTP assay. The substrate suspension was prepared by adding one of the provided test tablets to 5 mL of the Amylase Reaction Buffer (Appendix A). The substrate was suspended homogeneously with the help of a vortex, and 180 μL was added to each well of a 96-well transparent flat bottom MTP. The plate was incubated at 60°C for 5 min in an Eppendorf Thermomixer®, in order to pre-heat the Amylase Reaction Buffer.

The supernatants of cell cultures producing the secreted amylase were collected by centrifugation of 40 μL of culture in a 96-well non-skirted clear PCR plate (5 min, RT, 3,000 x*g*,). The supernatant samples were diluted when necessary, to ensure that the measured absorbance fitted the detection limits of the plate reader and that the enzyme did not saturate the reaction substrate.

To initiate the enzymatic reaction, 20 μL of the supernatant samples containing the secreted α-amylase were simultaneously added to individual wells of the pre-heated reaction plate with vigorous mixing. To three of the wells, 20 μL of growth medium was added to serve as the reaction blanks. The plate was incubated for 20 min at 60°C in an Eppendorf Thermomixer®. To stop the enzymatic reaction, 70 μL of Stopping Buffer was added simultaneously to each well. The plate was then centrifuged (5 min, RT, 3,000 x*g*) to separate the suspended substrate from the dissolved blue product. After centrifugation, 100 μL of the supernatant was transferred to a fresh 96-well transparent flat bottom MTP. The relative amylase activity was quantified by measuring the absorbance at 620 nm in a Multiskan® Ascent 96/384 Plate Reader.

Eppendorf tube assay. To prepare the substrate suspension, one provided test tablet was added to each necessary 5 mL of Amylase Reaction Buffer (Appendix A) and kept homogenously suspended with the help of a vortex. Each individual reaction was performed in a 2 mL

Eppendorf in 300 μL of substrate suspension pre-heated to 60°C for 5 min in an Eppendorf Thermomixer®.

The supernatants of cell cultures containing the secreted α-amylase were collected by centrifugation (1 min, RT, 16,000 x*g*) of 100 uL of culture in a 1.5 mL Eppendorf tube. To initiate the enzymatic reaction, 20 μL of the collected supernatant samples were added to the substrate suspension and incubated at 60°C for 15 min in an Eppendorf Thermomixer®. The reaction was stopped by addition of 300 uL of the Stopping Buffer (Appendix A). The suspended substrate was separated from the reaction product by centrifugation (5 min, RT, 3,000 x*g*). The reaction product was quantified by measuring the absorbance of the supernatant at 620 nm in a spectrophotometer.

2.10.4 Determination of α-amylase enzymatic units

One Unit of Amylase Activity (U) was defined as the amount of enzyme catalysing the conversion of 1 μmol of maltose from starch per min, at 60 °C and in 100 mM phosphate buffer with pH 5.5. In order to estimate Amylase Enzyme Units in *B. subtilis* cultures, a correlation was determined between the amylase enzymatic assay (Section 2.10.3) and the conversion of maltose in the presence of starch. The converted maltose was quantified using the DNS method (Miller 1959) similarly to that described in Section 2.10.2 .

An Amylase Enzymatic Units Standard Curve (Appendix E, Figure E.5) was obtained using an amylase standard sample and a maltose standard curve to correlate the absorbance at 620 nm values obtained with the Phadebas® Amylase Test Tablets (Magle Life Sciences) and the amount of converted maltose, in defined conditions.

Strain BRC39 (see Section 3.2.), carrying a high copy number plasmid expressing the gene encoding for AmyM of *G. stearothermophilus* was inoculated from a single colony in 30 mL of LB medium supplemented with neomycin (50 μg/mL). After overnight growth at 37 °C, the cells were harvested by centrifugation (10 min, RT, 3,000 x*g*). The cell-free culture supernatant was used as the crude enzyme source as an enzyme standard sample.

A series of twelve dilutions of the amylase enzyme standard was prepared by consecutive dilutions of the sample in an equal volume of Amylase Reaction Buffer (Appendix A). The enzymatic activity in each dilution was measured, in triplicate, using the amylase activity using

the Phadebas® commercial assay (Section 2.10.3) and the results were plotted to obtain the Phadebas® Standard Curve (Appendix E, Figure E.1).

The amylase standard dilutions were also used to quantify the conversion of maltose from starch by reacting with the DNS reagent (Appendix A). A substrate solution was prepared by dissolving starch in Amylase Reaction Buffer to a final concentration of 1% (w/v). The enzymatic reaction was preformed in the same conditions as the commercial assay by adding 180 μL of the starch substrate solution to 20 μL of the amylase standard dilutions, in triplicate, in a 96-well flat bottom MTP. The reaction was incubated at 60 °C for 20 min. The converted maltose was quantified by transferring 50 μL of the reaction to 50 μL of DNS reagent in a 96-well flat bottom MTP. The samples were heated for 10 min at 99 °C in an MTP Eppendorf incubator. Immediately after the heating period, 150 μL of milliQ water was added to each well. The plate was then placed on ice for 10 min before measuring the absorbance at 540 nm in a microplate reader. The results were used to obtain the DNS Standard Curve that correlates the converted maltose absorbance and the corresponding absorbance values obtained with the DNS method (Appendix E, Figure E.2).

Since the Phadebas® and the DNS Standard Curves were obtained with the same enzyme standard samples, they were combined into a single correlation denominated the Amylase Enzymatic Reaction Standard Curve (Appendix E, Figure E.3).

A Maltose Standard Curve (Appendix E, Figure E.4) was calculated by preforming the DNS method in the same conditions as above using 50 μL of maltose solutions in a range of concentrations.

The Amylase Enzymatic Units Standard Curve (Appendix E, Figure E.5) was calculated by combining the Maltose Standard Curve and the derived Amylase Enzymatic Reaction Standard Curve. The resulting valus was divided by the duration of the enzymatic assay in min (20), the duration of the enzymatic assay in min, to obtain a direct correlation between the absorbance values measured using the commercial enzymatic assay and the correspondent Amylase Enzymatic Units (U) was obtained:

$$U = 0.2118 \times A_{620}$$ Equation 2-6

2.11. *In silico* data analysis

All data calculations were done using Microsoft Excel. Representation of data in graphs was prepared using GraphPad Prism 6 (GraphPad Software, Inc.). *In silico* DNA manipulation was performed in Clone Manager 9 (SciEd Software). Vector schemes were prepared in Vector NTI® (Thermo Fisher Scientific). DNA chromatograms were interpreted using the software Chromas (Technelysium). Microscopy pictures were analysed and prepared in ImageJ (open source). Protein quantification in SDS-PAGE gels was performed with ImageQuant TL Software (GE Healthcare Life Sciences).

Chapter 3 Characterisation of native and heterologous enzyme production in *Bacillus subtilis* using two industrially relevant model enzymes

3.1. Introduction

The production of enzymes and therapeutic proteins is a global-scale market of major importance in modern society with an annual turnover of over 2 billion Euros (van Dijl & Hecker 2013). *Bacillus* sp. play an important role in this economy, being one of the most widely-used biological platforms for protein production. This is particularly due to its capacity to deliver high yields of product (≥20 grams per litre) secreted directly into the culture medium, reducing downstream processing costs (Harwood & Cranenburgh 2008; van Dijl & Hecker 2013). *Bacillus subtilis* in particular is widely used industrially to produce native and heterologous enzymes such as proteases, α-amylases, xylanases and lipases with great commercial interest (Harwood & Cranenburgh 2008; Pohl & Harwood 2010).

In this study, two industrially relevant model enzymes were chosen to illustrate the well-known capacity of *B. subtilis* to produce industrial enzymes and to investigate the characteristics and limitations of different industrial strains. These are the endo-1,4-β-xylanase XynA from *B. subtilis*, and the maltogenic α-amylase AmyM from *Geobacillus stearothermophilus*. The model enzymes where characterized biochemically, and their production profiles from *B. subtilis* were analysed with respect to gene expression, enzyme activity profiles, secretion kinetics and the impact on their secretion stress responses and cellular secretion mechanisms. In this chapter these comparative studies are described, facilitating a comprehensive characterization of the production strains that are analysed throughout the rest of the study.

3.2. Strains and plasmids

The strains and plasmids used in this chapter are listed in Table 3.1 and Table 3.2., respectively. Appendix C contains a list of the primers used. Appendix B contains the plasmid schemes from Table 3.2. DNA manipulation and transformation methods are described in Sections 2.5. and 2.6. , respectively. *B. subtilis* strain 168 was used as the host for all the strains in this study.

The plasmid pCS73 was obtained from DSM, it is a derivative of plasmid pNAPHB27 (Quax & Broekhuizen 1994) and corresponds to a high copy number plasmid expressing *amyM* from *Geobacillus stearothermophilus* under the control of the *amyQ* promoter of *Bacillus amyloliquefaciens*. It has a neomycin resistance gene for selection in *B. subtilis*. To obtain an identical expression system for XynA, the *amyM* gene in pCS73 was replaced by the *xynA* gene from *B. subtilis* 168 using Gibson Assembly. This work resulted in the plasmid pCS58 and is described by Sauer (2016).

The plasmids pRC67 and pRC68 correspond to the plasmids pCS58 and pCS73 with a signal peptide swop in the precursors of XynA and AmyM. To generate pRC67, the signal peptide of the AmyM precursor, amplified using primers 353 and 355, was assembled with the pCS58 plasmid, amplified with the primers 356 and 357, using Gibson Assembly. Similarly, the signal peptide of the XynA precursor was amplified using the primers 353 and 354 and assembled to pCS73 amplified with the primers 356 and 358. The new XynA and AmyM precursors with swopped signal peptides were designated XynA2 and AmyM2, respectively.

Table 3.1 Summary of the strains used in the study of this chapter.

Strain	Genotype	Source
B. subtilis		
168	*trpC2*	Kunst *et al.* 1997
BWAP	168 pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BWXP	168 pCS58 (P_{amyQ}-*xynA*), *neo*	Sauer (2016)
BWXP2	168 pRC67 (P_{amyQ}-*xynA2*), *neo*	This work
BWAP2	168 pRC68 (P_{amyQ}-*amyM2*), *neo*	This work

neo – neomycin resistance gene

Table 3.2 Summary of the plasmids used in the study of this chapter.

Plasmid	Properties	Source
pCS58	P_{amyQ}-*xynA*, *reppUB*, *neo*, *bleo*	Sauer (2016)
pCS73	P_{amyQ}-*amyM*, *reppUB*, *neo*, *bleo*	DSM
pRC67	P_{amyQ}-*xynA2*, *reppUB*, *neo*, *bleo*	This work
pRC68	P_{amyQ}-*amyM2*, *reppUB*, *neo*, *bleo*	This work

neo – neomycin resistance gene; *bleo* – bleomycin resistance

3.3. XynA and AmyM: two industrially relevant model enzymes

In the enzyme industry, *B. subtilis* and its relatives are used for the production of food grade enzymes, including amylases, glucanases, xylanases and proteases. In this study, *B. subtilis* production strains BWXP and BWAP (Table 3.1), encoding respectively the endo-1,4-β-xylanase XynA from *B. subtilis*, and the maltogenic α-amylase AmyM from *Geobacillus stearothermophilus*, are analysed in detail. The main enzymatic, molecular and functional characteristics of XynA and AmyM are shown in Table 3.3.

Table 3.3 The enzymatic, molecular and functional characteristics of XynA and AmyM (Brenda 2016).

	XynA	**AmyM**
Recommended name	Endo-1,4-β-xylanase	Maltogenic α-amylase
Alternative name	-	Glucan 1,4-α-maltohydrolase
Enzyme Commission number	EC: 3.2.1.8	EC: 3.2.1.133
Gene length (bp)	639	2160
Protein length (aa)	213	720
Protein weight (KDa)	23.0	79.2
Organism	*Bacillus subtilis*	*Geobacillus stearothermophilus*
Catalytic activity	Endohydrolysis of (1->4)-β-D-xylosidic linkages in xylans.	Hydrolysis of (1->4)-α-D-glucosidic linkages in starch so as to remove successive α-maltose residues from the non-reducing ends of the chains.
Application	Baking, pulp and paper industries.	Food, textile, fuel alcohol production, paper and detergent industries.

XynA is an endo-1,4-β-xylanase from *B. subtilis* that catalysis the endohydrolysis of (1->4)-β-D-xylosidic linkages in xylans (Brenda, 2016). This enzyme degrades xylan, the major hemicellulose in cereals and hardwoods, and which is the second most abundant renewable polysaccharide in nature (Polaina & MacCabe 2007). Xylanases have great value in the baking industry as they improve bread volume, crumb structure, reduce stickiness, increase shelf life and

reduce bread staling (Butt *et al.* 2008). These enzymes also have applications in the pulp and paper industries, particularly to increase the bleachability of kraft pulps in an environmentally sensitive manner, increasing the brightness of the pulp and improving fibre properties without the need for harsh chlorine-based chemicals (Buchert *et al.* 1994).

AmyM is a maltogenic α-amylase from *Geobacillus stearothermophilus* that catalysis the hydrolysis of (1->4)-α-D-glucosidic linkages in polysaccharides so as to remove successive α-maltose residues from the non-reducing ends of the chains (Brenda, 2016). α-Amylases degrade starch, the most abundant polymer on earth, which provides them a wide number of applications in industrial processes, such as food, textile, fuel alcohol production, paper and detergent industries (Polaina & MacCabe 2007; de Souza & de Oliveira Magalhães 2010). AmyM is secreted naturally by the thermophile *Geobacillus stearothermophilus* and therefore has potential to be explored in industrial processes that require enzymes to be active at high temperatures. Thermostable amylases have particular value in a number of commercial applications, such as the enzymatic liquefaction and saccharification of starch which are performed at high temperatures (de Souza & de Oliveira Magalhães 2010).

Production strains BWXP and BWAP (Table 3.1) encode the genes for two model enzymes with major commercial relevance since they catalyse the conversion of the most abundant polysaccharides in nature. Furthermore, one of the strains produces a native product (XynA) and the other a heterologous product (AmyM). Together these strains are representative of the contexts in which bacterial production strains are used in industrial processes, where often a host with good secretion capacity is explored for production of not only natively secreted products, but also heterologous products.

The production strains described in this chapter make use of the same plasmid expression system under the control of a strong promoter. A complete study of the physiological properties and production profiles with regards to gene expression, enzyme activity profiles, secretion kinetics and their impact on the secretion and stress response mechanisms of the cell will be presented.

3.3.1 Characterisation of XynA and AmyM enzymatic properties

The thermostability and pH and temperature optima of the enzymatic activities of XynA and AmyM were analysed and compared. Samples of crude enzyme were obtained from the spent media of the production strains BWXP and BWAP grown for 24 hours at 37°C in 10 mL LB with

agitation (250 rpm). The presence of the secreted enzymes was confirmed via SDS-PAGE analysis (Figure 3.1). Assays were performed in triplicate for XynA (Section 2.10.1) and AmyM (Section 2.10.3) at a different range of pHs and temperatures, following crude enzyme incubation for 60 min. These assays revealed the optimal pH and thermostability of the enzymes (Figure 3.2). The fluorescence and absorbance values obtained with the commercial assays were converted to enzymatic units using the derived formulae described in Section 2.10.2 and Section 2.10.4 , respectively. The results of the enzymatic assays are summarised in Appendix F (Table F.1). The analysis of enzyme activity at different pH values show that the pH optimum of AmyM was between 5.5-6.5 and at the higher pH values there was a steep decline in activity (Figure 3.2-B). In contrast, the pH optimum for XynA was around 7.0 and high activity was maintained up to a pH of at least 10.0 (Figure 3.2-A).

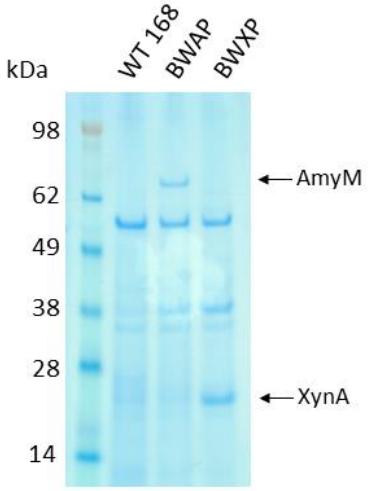


Figure 3.1 SDS-PAGE analysis of BWXP and BWAP spent media showing the secreted enzymes XynA and AmyM, respectively. The cultures were grown for 24 h at 37 °C in 10 mL LB with agitation (250 rpm). The spent media was purified from the culture using centrifugation and the samples for protein analysis were processed according to Section 2.8. .

When assayed at pH 5.5, the temperature optima of AmyM was 90 °C (Figure 3.2-B). This was to be expected as *Geobacillus stearothermophilus* is a thermophilic bacterium and the industrial relevance of AmyM is related to its ability to withstand high processing temperatures (Polaina & MacCabe 2007). In contrast, XynA shows increasing enzymatic activity up until 60ºC in the conditions of the commercial enzymatic assay preformed at pH 4.6 (Section 2.10.1) but above this temperature there is a sharp decrease in activity (Figure 3.2-B). These data were reinforced when experiments were carried out to determine their thermostability (Figure 3.2-C).

Samples of crude enzyme were incubated for one hour at temperatures ranging from 40 to 99 ºC, determining the residual enzyme activity following the heat treatment (Bukhari & Rehman 2015). Enzymatic assays were performed under the conditions described in Section 2.10. Increasing the exposure time to high temperatures decreases the activities of both enzymes significantly. As

expected, AmyM is more thermostable, taking 60 min at 75 ºC to inactivate the enzyme. XynA is inactivated after 60 min at 60.3 ºC.

In general, the study shows that XynA is an alkaline enzyme with pH optima of 7.1 and temperature optima of 60 ºC (Figure 3.2). AmyM is a highly thermostable enzyme with pH optima of 6.5 and temperature optima of 90 ºC (Figure 3.2). These enzymatic properties are, of course, only relevant to the conditions in which the assays were performed, but provide a good guide to their general characteristics.

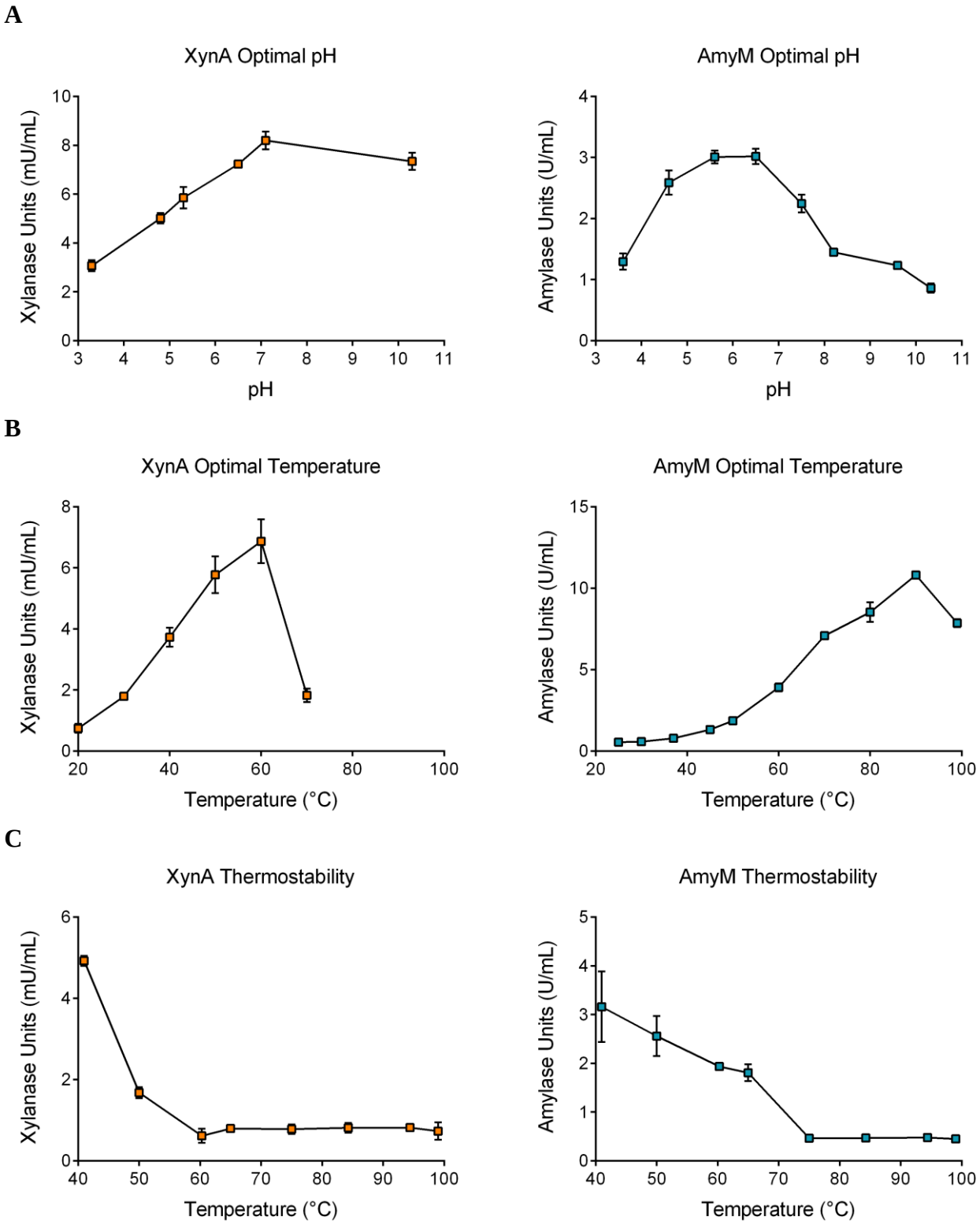


Figure 3.2 Effect of (A) pH, (B) temperature and (C) thermostability on XynA and AmyM enzymatic activities. Enzymatic assays were performed in triplicate according to Section 2.10. Enzymatic units were calculated with the correlations derived in Sections 2.10.2 and 2.10.4 . The error bars show the SD of three technical replicates.

3.3.2 Characterization of XynA and AmyM production strains

In this section, the gene expression, enzyme activity profiles, secretion kinetics and impact on their secretion and cell stress response mechanisms are analysed in production-like strains BWXP and BWAP. This comprehensive analysis follows aims to identify differential characteristics and challenges that are encountered when native and heterologous enzymes are secreted from strains of *B. subtilis*.

3.3.3 Growth and pH profiles in rich medium

The BioLector® bench top microfermentation system (m2p-labs) facilitates the monitoring of biomass, pH and DO, the latter two via optical sensors. These sensors consist of special dyes that respond to environmental conditions (m2p 2015). The BioLector® was used to study the enzyme activity profiles of the production strains BWXP and BWAP in comparison with the wild type 168 strain.

The growth regime, described in detail in Section 2.3. briefly, involves the preparation of both overnight and preliminary cultures to synchronize the growth phases of the cells before inoculating pre-warmed LB medium in 48-well MTP FlowerPlates®. Growth was monitored for 33 hours at 37 °C with 95% humidity and vigorous agitation (800 rpm). Biomass (excitation: 620 nm, gain: 20) and pH were monitored by systematic measurements every 15 min. The wild type strain 168 was used as a negative control and four wells containing only sterile media as blanks to monitor contaminations during the experiment. The results are shown in Figure 3.3.

Under the test conditions, the production strains show similar growth rates and mean generation times, with BWAP strain showing the slowest growth (Table 3.4). The calculated averages of the growth parameters for 8 biological replicates indicate that the production of XynA and AmyM increases the variance of the growth rate between biological replicates. Nonetheless, the growth profiles are similar after incubation for 15 hours (Figure 3.3-A). After this stage, the dramatic metabolic and physical changes that accompany entry into “stationary phase” leads to a decrease of cell density (Figure 3.3-A). It is notable that production of the heterologous enzyme, AmyM, has a higher impact on culture density than the similarly expressed native XynA enzyme.

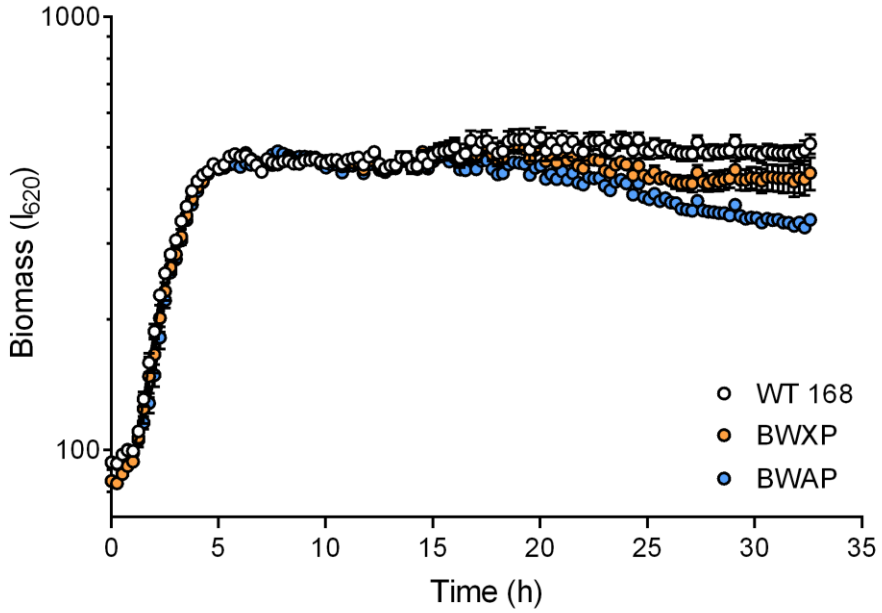
The cultures showed similar pH profiles throughout the growth cycle with a rapid alkalization of the cell culture during exponential phase of growth (Figure 3.3-B). This is likely due to the release of nitrogen waste products (e.g. amines, ammonium) during growth on nitrogen-rich LB.

At a later stage of growth, there was a slow decrease in pH for all cultures, although this was less marked in the case of the amylase producing cells (Figure 3.3-B). LB medium is not buffered which makes it susceptible to changes in pH from released acidic and alkaline compounds generated by the cell growth and cell lysis. Since the intracellular pH of growing *B. subtilis* has been reported between 7.5 and 8.2 (Beilen 2013; Setlow & Setlow 1980), a possible explanation for the reduced pH of the amylase producer is the apparent cell lysis occurring at the same stage in the growth cycle (Figure 3.3-A). Ultimately, the high-level enzyme production has an effect on the pH of the medium, resulting in a more alkaline broth when compared to the wild type 168 cell culture.

Table 3.4 Growth rates and generation times of bacterial cultures of the strains wild type 168, BWXP and BWAP. The strains were grown in 1.5 mL LB in FlowerPlate® 48-well MTPs, and incubated at 37 °C with 95% humidity and strong agitation (800 rpm) using the BioLector® bench-top microfermentation system (m2p-labs). The presented growth rates correspond to the maximum slope of the function ln N $versus$ t, where *N* is the biomass measured at any given time, *t.* The mean generation time (MGT) was calculated using the formula: $MGT = {ln2}/{growth\ rate}$. The average from eight biological replicates was calculated and is shown with the corresponding standard deviation.

	Wild type 168	**BWXP**	**BWAP**
Growth rate (h^{-1})	1.02 ± 0.04	1.05 ± 0.10	0.94 ± 0.07
Mean generation time (min)	40.65 ± 1.46	40.18 ± 4.04	44.51 ± 3.57

A



B

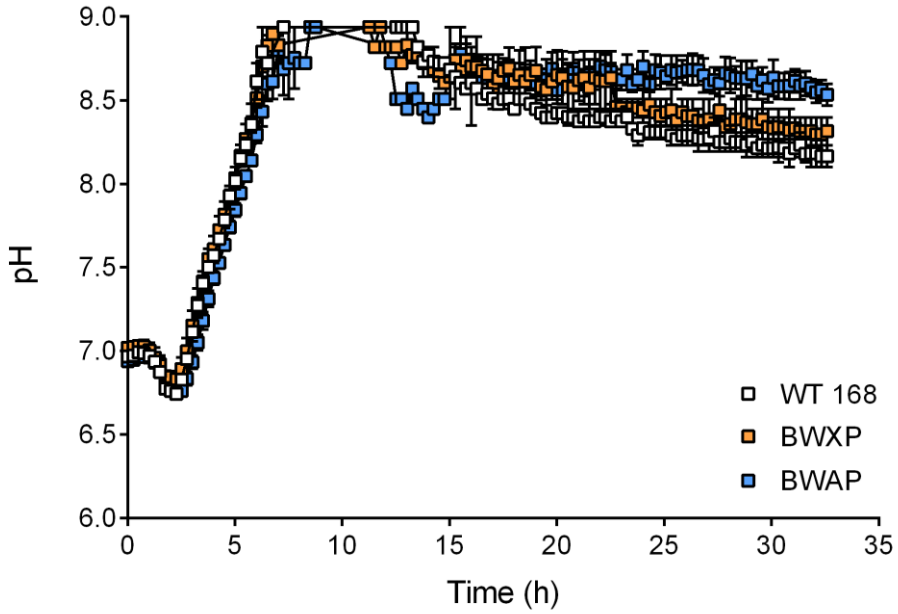


Figure 3.3 Monitoring of (A) biomass and (B) pH during the cultivation of the production strains BWXP and BWAP using the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 1.5 mL LB in FlowerPlate® 48-well MTPs, and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The average of four biological replicates per strain was plotted with the corresponding SEM. Due to the limitations of the optical sensors for detection of pH, pH-values above 9.0 were discarded.

3.3.4 Enzyme production profiles

During the 33-hour study in the BioLector®, the production of enzyme by strains BWXP, BWAP, and 168, was monitored by taking samples (17) at several time points. The samples were immediately frozen and the XynA or AmyM activity assayed up to three days after storage according to the methods described in Section 2.10.2 and Section 2.10.4 , respectively. The fluorescence and absorbance values obtained with the commercial assays were converted to enzymatic units (Section 2.10.2 and Section 2.10.4 , respectively). The results of the enzymatic assays are summarised in Appendix F (Table F.2).

The assays show that both production strains produce high levels of the enzyme of interest when compared to the wild type 168. Interestingly, given that they are expressed from the same promoter system (*amyQ*), the XynA and AmyM production profiles are different. This difference is highlighted in Figure 3.4 where the activity profiles are plotted as relative values based in the maximum production level observed during the first 10 hour of growth. Whereas xylanase production by strain BWXP coincides with the start of exponential growth, amylase production starts during transition from exponential to stationary growth.

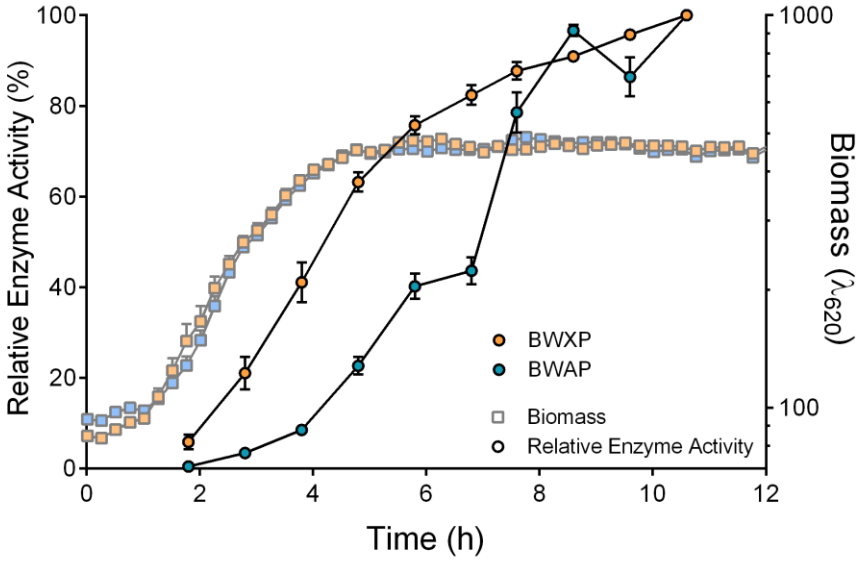
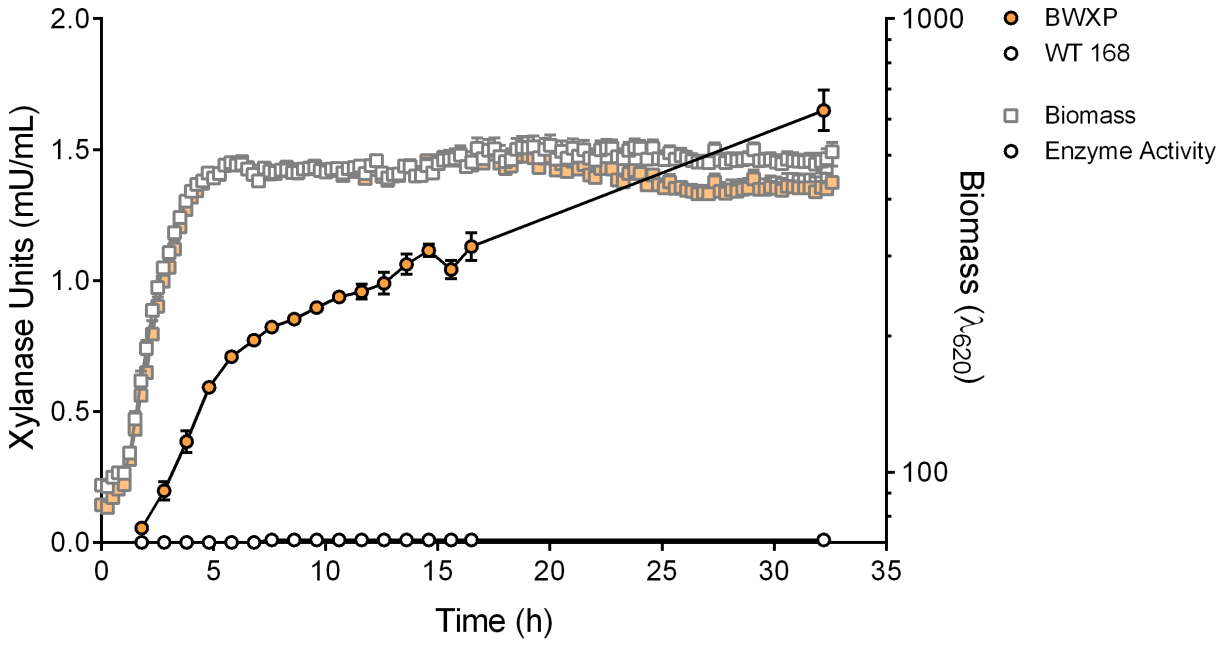


Figure 3.4 Enzyme production profiles of the strains BWXP (XynA) and BWAP (AmyM). The profiles are relative to the maximum production obtained in 10 hours of growth in 1.5 mL LB, at 37 °C with 95% humidity and strong agitation (800 rpm). The cultures were grown in FlowerPlate® 48-well MTPs using the BioLector® bench-top microfermentation system (m2p-labs). The enzymatic activity was measured using commercial assays (Section 2.10.). The average of four biological replicates per strain was plotted with the corresponding SEM.

A



B

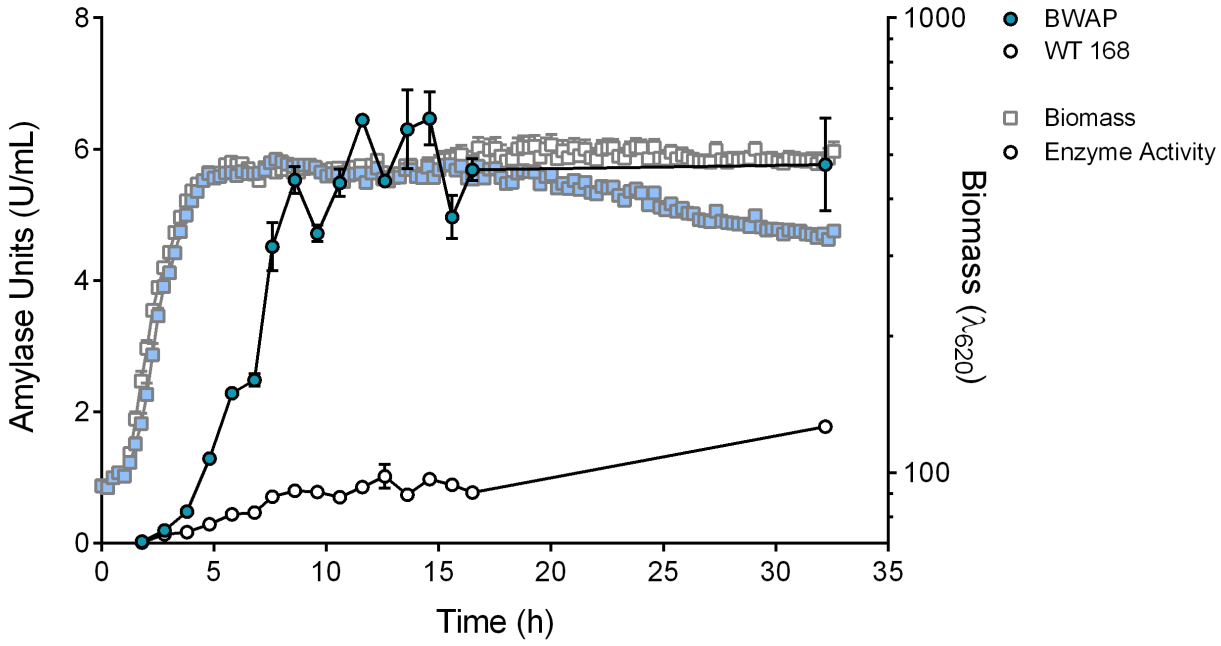


Figure 3.5 (A) XynA and (B) AmyM production during the cultivation of the production strains BWXP and BWAP, respectively, using the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 1.5 mL LB in FlowerPlate® 48-well MTPs and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). During the 33 hours during which growth was monitored, 17 samples were taken for enzymatic activity determination (Section 2.10.). The average of four biological replicates per strain was plotted with the corresponding SEM.

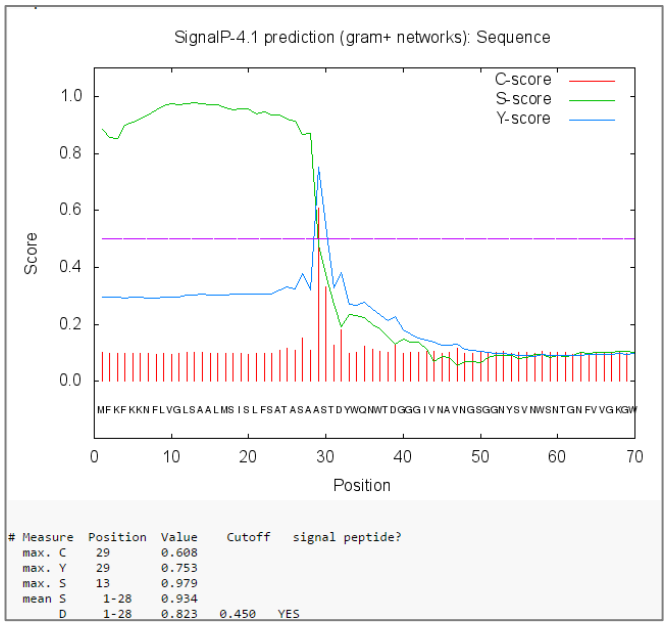
The difference in the production profiles of XynA and AmyM is likely related to different interactions of the enzymes with the secretion machinery of the cell, as both production strains express the enzymes from identical expression systems. This hypothesis was explored by analysing the impact of signal peptide replacement on the production profiles (Section 3.3.5), measuring transcription of the enzyme genes (Section 3.3.6), comparing the kinetics of enzyme secretion (Section 3.3.7) and investigating the up-regulation of the secretion and stress response mechanisms (Section 3.3.8).

3.3.5 Impact of signal peptide replacement on the enzyme production profile

The protein secretion pathway in *B. subtilis* starts with identification of the substrate for targeting to the secretion apparatus via an hydrophobic N-terminal extension, the signal peptide (Harwood & Cranenburgh 2008). Little is known about how signal peptides are used to target secretory proteins, however, Zanen *et al.* (2005) have suggested that *B. subtilis* is not able to secrete precursor proteins with signal peptides of low hydrophobicity. It is known that, in *E. coli*, signal peptide hydrophobicity is a determinant for the targeting of secretory proteins to the SRP pathway, rather than to the SecA/B pathway (Lee & Bernstein 2001). Together with the fact that precursor proteins from Gram-positive bacteria, which lack SecB, have longer and more hydrophobic signal peptides than those of Gram-negative bacteria, it is expected that precursors containing signal peptides with high hydrophobicity would, in general, be more efficiently translocated in *B. subtilis* (Zanen *et al.* 2005).

The SignalP 4.1 server was used to predict the location of signal peptide cleavage sites in the amino acid sequences of XynA and AmyM and subsequently the N-terminal sequence corresponding to the signal peptide of the respective precursors (Figure 3.6) (Petersen *et al.* 2011). The hydrophobicity of the predicted signal peptides was calculated using the online tool Protein GRAVY (Grand Average of Hydropathy) and values of 0.986 and 0.694 were obtained for XynA and AmyM signal peptides, respectively (Stothard 2000). In hydrophobicity scales, the more positive the value, the more hydrophobic are the amino acids of the peptide. Therefore, the XynA precursor contains a more hydrophobic signal peptide then AmyM.

A



B

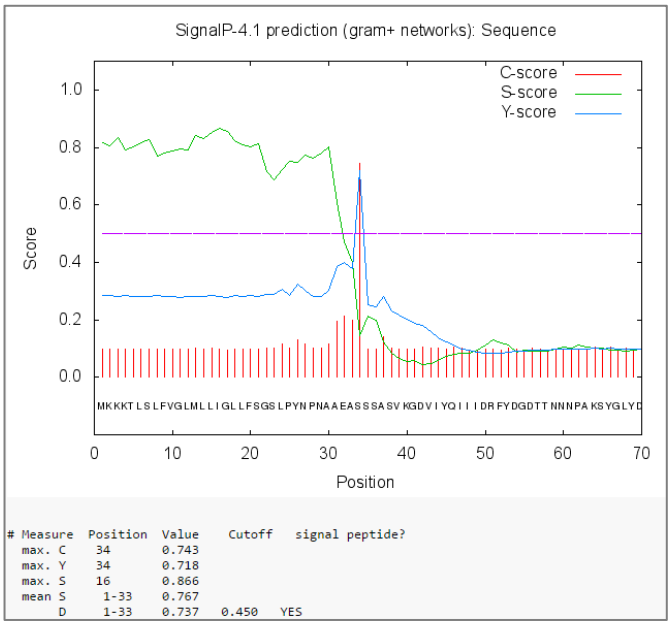


Figure 3.6 Prediction of the location of the signal peptide cleavage sites in the amino acid sequences of (A) XynA and (B) AmyM using the SignalP 4.1 server. The graphical output from SignalP shows the three different scores, C, S and Y, for each position in the sequence, and two additional scores, mean S and D. The C-score (raw cleavage site score): the C-score is trained to be high at the position immediately after the cleavage site (the first residue in the mature protein). The S-score (signal peptide score): the S-score is trained to distinguish positions within signal peptides from positions in the mature part of the proteins and from proteins without signal peptides. The Y-score (combined cleavage site score): combination (geometric average) of the C-score and the slope of the S-score, resulting in a better cleavage site prediction than the raw C-score alone. Mean S: the average S-score of the possible signal peptide (from position 1 to the position immediately before the maximal Y-score). D-score (discrimination score): weighted average of the mean S and the maximum Y score that is used to discriminate signal peptides from non-signal peptides. (Petersen *et al.* 2011)

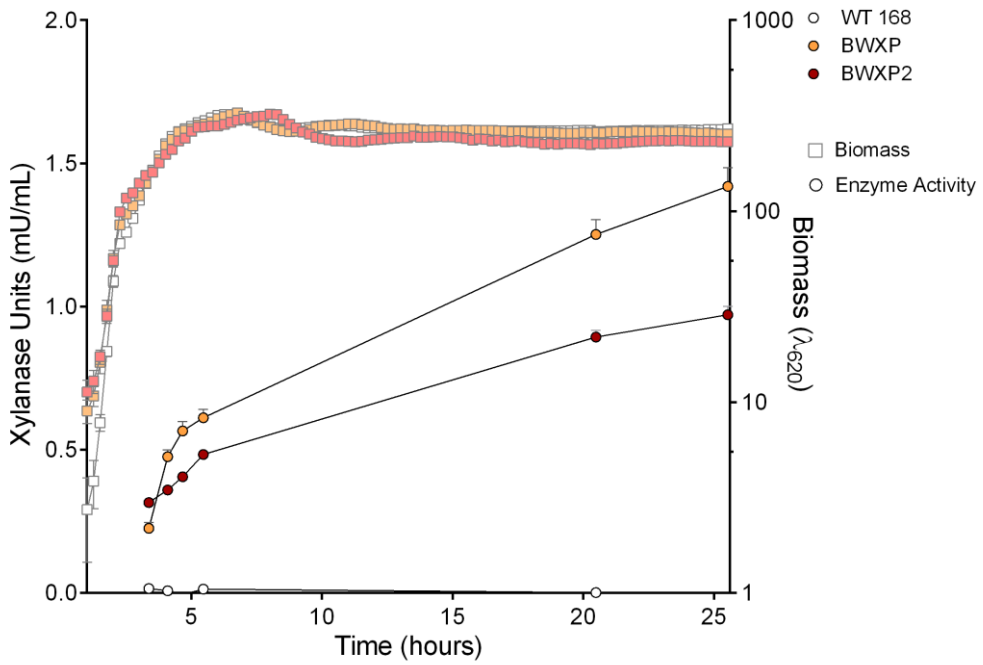
In order to evaluate the influence of the signal peptides on the enzyme production profile, we swapped the signal peptides of the precursors of XynA and AmyM. Plasmids were constructed in which the XynA precursor had an AmyM signal peptide, and the AmyM precursor a XynA signal peptide (Section 3.2.). Hereafter, these precursors are referred to as XynA2 and AmyM2, and their respective production strains are BWXP2 and BWAP2.

The BioLector® microfermentation system (m2p-labs) was used to monitor enzyme production in strains BWXP, BWAP, BWXP2 and BWAP2 during growth in LB medium, using wild type strain 168 to determine endogenous enzyme production. The growth regime was identical to that described in Section 3.3.4 . The cultures were grown for 26 hours at 37 °C with 95% humidity and vigorous agitation (800 rpm). Biomass was monitored every 15 min (excitation: 620 nm, gain: 20) and 6 samples taken manually from each well at the indicated time points to determine enzyme activity. The samples were immediately frozen and the XynA (Section 2.10.1 and AmyM (Section 2.10.3) enzyme activities were determined after up to three days of storage. The fluorescence and absorbance values obtained with the commercial assays were converted to enzymatic units per millilitre of culture (Section 2.10.2 and Section 2.10.4) and the data shown in Figure 3.7 and summarised in Appendix F (Table F.3).

The replacement of the XynA signal peptide with the AmyM signal peptide reduces the production of XynA at all stages of the growth cycle (Figure 3.7), with the ultimate yield being about 30% lower (0.97 ± 0.03 mU/mL *cf.* 1.42 ± 0.07 mU/mL). When the XynA signal peptide was used in place of the native AmyM signal peptide the reduction in amylase production was even more dramatic, being reduced to approximately 20% (0.61 ± 0.02 U/mL *cf.* 2.41 ± 0.15 U/mL). The data is therefore consistent with that of Zanen *et al.* (2005) who showed that the secretion of the AmyQ from *B. licheniformis* was most efficiently directed by its cognate signal peptide rather than heterologous signal peptides of either higher and lower hydrophobicity.

Taken together, these results clearly show that the relationship between signal peptide and passenger protein is likely to be multifactorial and that no single factor, such as the hydrophobicity of the H-region, can necessarily improve productivity. In general, signal peptides and cognate mature proteins have co-evolved to optimise secretion in their native environment, and a better understanding of how these two components interact during targeting and translocation is needed before efficient signal peptide/mature protein combinations can be designed rationally.

A



B

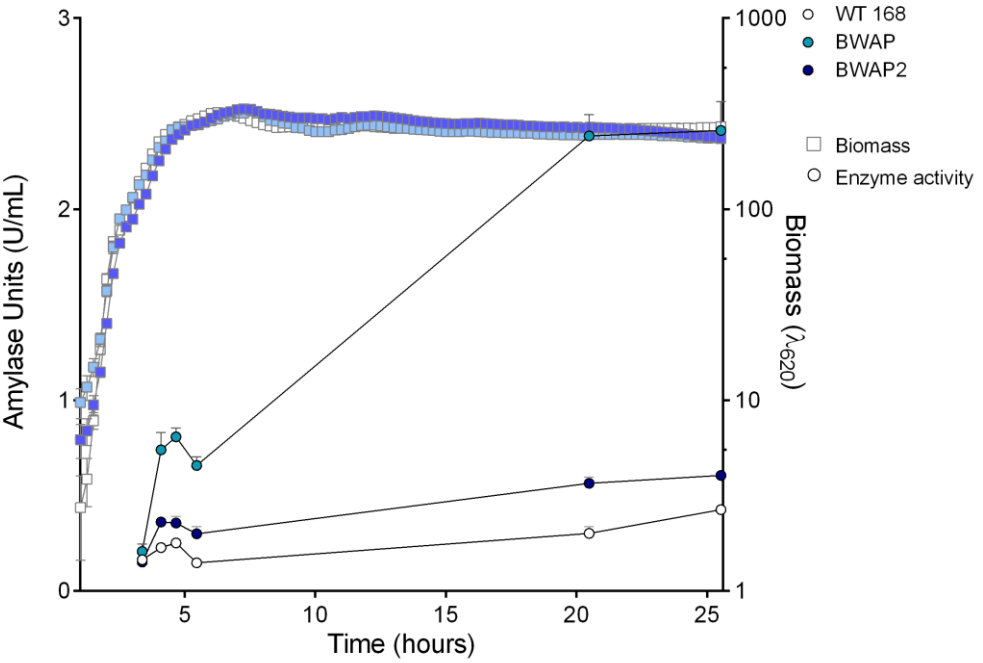


Figure 3.7 Analysis of the effect of swapping XynA and AmyM signal peptides on enzyme production profiles. The production profiles of the strains (A) BWXP2 and (B) BWAP2 corresponding to a signal peptide swap between the XynA and AmyM proteins in the production strains BWXP and BWAP are plotted relative to the maximum production obtained in 25 hours of growth. The cells were cultured in 1.5 mL LB, at 37 °C with 95% humidity and strong agitation (800 rpm) in FlowerPlate® 48-well MTPs using the BioLector® bench-top microfermentation system (m2p-labs). The enzymatic activity was measured using commercial assays (Section 2.10.). The average of four biological replicates per strain was plotted with the corresponding SEM.

3.3.6 Enzyme expression levels

The production strains BWXP and BWAP carry a high copy number plasmid expressing XynA and AmyM, respectively, under the control of the P_{amyQ} promoter and the same ribosome binding site (see Table 3.1 and Appendix B for the plasmid maps). Despite sharing the same expression system, the production profiles of these strains were different (Figure 3.4). Specifically, the main production phase for XynA was during the exponential growth phase while that of AmyM was towards the end of the exponential phase. In order to understand the reason for this discrepancy, the transcription levels of *xynA* and *amyM* were quantified using RT-qPCR (Section 2.7.) and using either the ΔC_t or the $\Delta\Delta C_t$ method to process the data (Figure 3.8) (Livak & Schmittgen 2001).

The expression levels of *xynA* and *amyM* are represented in Figure 3.8-A in terms of ΔC_t values that correspond to the difference between the C_t of the gene of interest and the control gene *fbaA*. In this case, it is not possible to apply the $\Delta\Delta C_t$ method as AmyM is an heterologous enzyme and *amyM* expression in relation to the wild type 168 negative control is an undefined number. Therefore, the ΔC_t values represent the expression levels of *xynA* and *amyM* in the samples normalised solely to the control gene *fbaA.* Appendix G contains a summary of the results in table format, including an alternative normalisation of the samples with the control gene *sdhA*. Together, these results show that *xynA* and *amyM* are transcribed at similar levels, since no significant differences were found, and synchronously at the time points at which the samples were harvested.

The levels of *xynA* transcription present a higher level of variance indicated by the larger SEM in Figure 3.8-A. This might be due to the different pattern of transcription of plasmid and chromosome based genes since *xynA* is present in both the high copy number plasmid and the chromosome of production strain BWXP. Therefore, it is important to confirm that the transcription levels in Figure 3.8-A are partially associated to the expression system introduced in the production strain. Figure 3.8-B confirms this by making use of the $\Delta\Delta C_t$ method to calculate the up-regulation of *xynA* in relation to the wild type strain 168. This figure also shows that target gene expression is down-regulated at 20 hours in comparison to the 4-hour time point. This correlates with the enzyme activity profile in Figure 3.3-A which shows a decrease in the rate of enzyme production at this stage when compared to the earlier time point of growth.

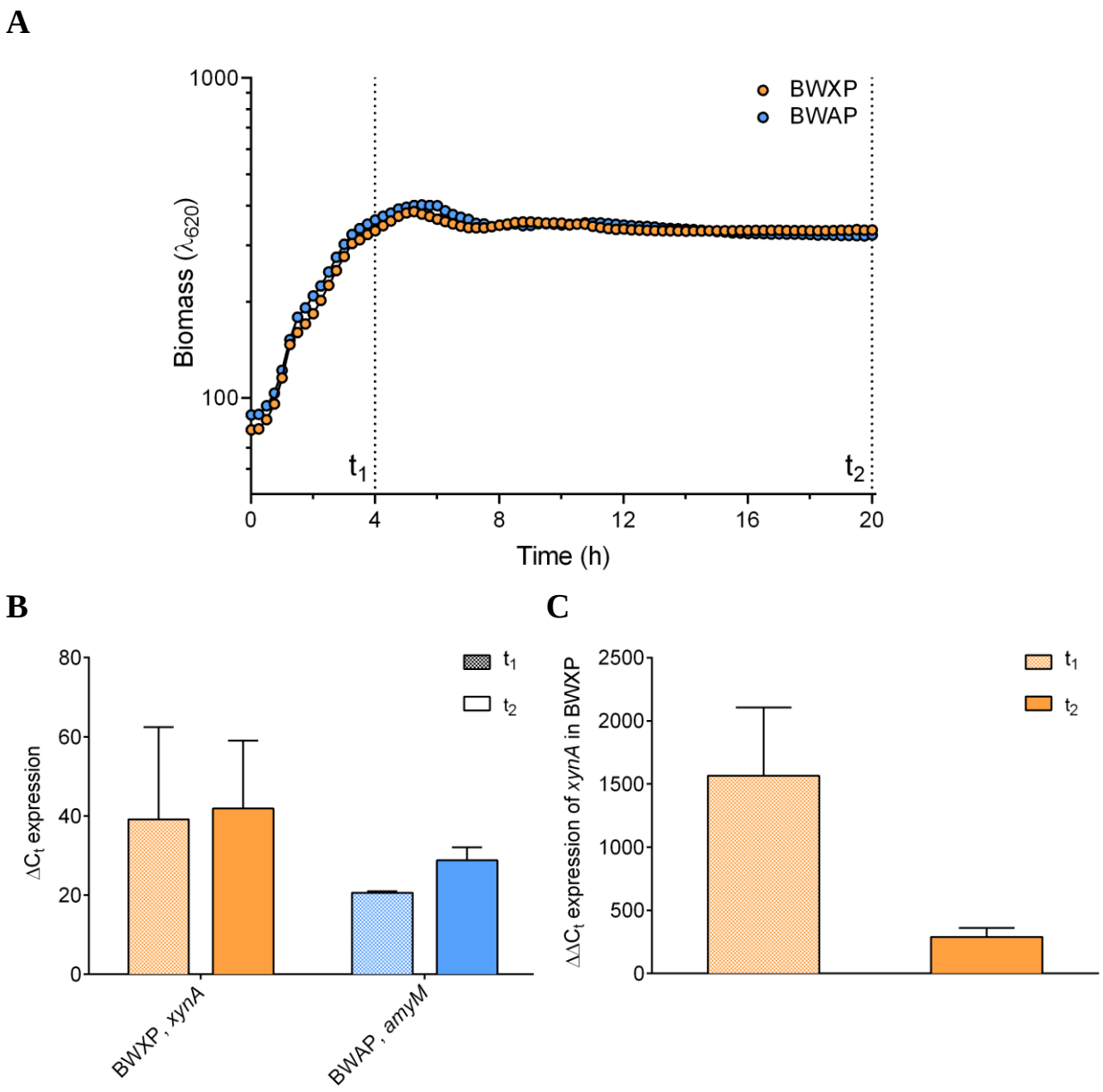


Figure 3.8 Expression of the *xynA* and *amyM* genes in the strains BWXP and BWAP.
A. Cultures of BWXP and BWAP were grown in 1.5 mL LB in FlowerPlates at 37 °C with 95% humidity and vigorous agitation (800 rpm). Samples were taken at two time points; t_1 during transition phase and t_2 in late stationary phase. The cultures were harvested and total cell RNA extracted as described in Section2.7.
B. The expression of the *xynA* and *amyM* genes determined by the qPCR.
C. The expression of the *xynA* gene in strain BWXP shown in relation to the wild type strain 168.
In the case of (B) and (C), the C_t values were averaged from two technical replicates, and the ΔC_t values were calculated using the *fbaA* gene as control. The average of the ΔC_t values of two biological replicates are shown with the corresponding SEM. A summary of the results in a table format can be found in Appendix G.

3.3.7 Enzyme secretion kinetics

Despite sharing the same expression system and similar transcription levels (Figure 3.8-A), the strains BWXP and BWAP show considerably different enzyme production profiles (Figure 3.4). To address this issue, pulse-chase experiments were performed to monitor their secretion kinetics. During a pulse-chase experiment, a radioactive pulse facilitates the monitoring of protein synthesis, translocation and processing for only the proteins synthesised during the pulse (usually 1-2 minutes). If combined with immunoprecipitation, the kinetics of a single protein can be determined. In these experiments, radioactively labelled methionine was added to a culture growing in a methionine-free medium (pulse). After a short period of label incorporation (1 minute), the radioactive signal was quenched by the addition a large excess of unlabeled methionine (chase). The addition of the chase corresponds to time zero. Samples collected at time zero and various time thereafter allow localisation of the protein of interest to be determined. This technology was used to reveal the secretion kinetics and stability of XynA and AmyM from the production strains BWXP and BWAP, respectively.

In the absence of an effective polyclonal antibody for AmyM, two slightly different protocols were used for XynA (Section 2.9.1) and AmyM (Section 2.9.2). In the case of XynA an immunoprecipitation step was included and this facilitated the analyses of both the cell-free and whole cell culture samples. In the case of AmyM, immunoprecipitation was not possible and consequently it was only possible to analysed cell-free culture medium samples. Nonetheless, AmyM was easily identified in the extracellular proteome as this *Geobacillus stearothermophilus* protein is not present in the wild type 168 strain (see Appendix I).

Figure 3.9-A shows the secretion kinetics of XynA in the production strain BWXP. The radioactively labelled protein was quantified (Section 2.9.3) and the activity plotted relative to the maximum protein levels detected in the whole-culture sample. Samples taken immediately following the chase (0 min) show 66.1% of the maximum detected protein in the whole-culture sample and 31.8% in the culture medium. This reveals that significant protein synthesis, processing and release occurs during the minute-long pulse and that it takes less than one minute for XynA to be translated, processed and secreted. The amount of XynA in the whole-culture sample peaked at 30 s, after which it declined until it reached a constant level of approximately 60% of the maximum detected xylanase. The amount of XynA released into the culture medium increased with time until it peaked at 2 min post-chase, corresponding to about 70% of the

maximum of protein detected. These data suggest that only a proportion of the xylanase synthesized during the pulse was released into the growth medium and that the observed degradation of XynA in the first two minutes of the chase occurs during or shortly after translocation across the cytoplasmic membrane and in a cell-associated location. However, despite the apparent susceptibility to cell-associated proteases, this enzyme is stable once it reaches the extracellular environment and is not targeted by the complex cocktail of extracellular proteases naturally secreted by *B. subtilis* (Pohl & Harwood 2010).

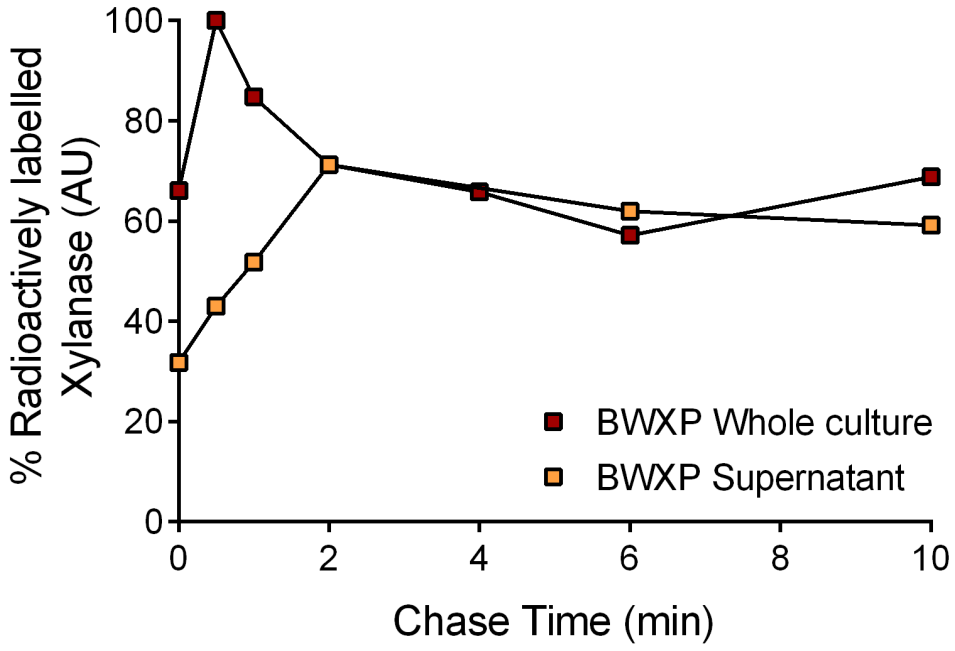
The kinetics of secretion of the heterologous enzyme AmyM are markedly different (Figure 3.9-B). AmyM first makes it appearance in the culture medium at 3 min, indicating that the translation, processing and release of the much larger AmyM proteins takes at least 4 min. Remarkably, at 60 min post-chase, there is still an increase of the radioactive signal corresponding to AmyM in the extracellular environment, indicating the rate and/or efficiency of AmyM processing is limited by an as yet unknown bottleneck. In the absence of whole-cell data, it is not currently possible to establish the nature of this bottleneck. For example, does the precursor protein accumulate in the cytoplasm at a faster rate than it can be translocated across the membrane. This would require the unfolded precursor to be stable in the cytoplasm, and protected from intracellular quality control proteases. However, in the absence of a suitable antibody, our data does not allow us to confirm this possibility. Nonetheless, the current data clearly point to a crucial difference between the kinetics of release of XynA and AmyM into the culture medium.

Besides the native proteolytic activity affecting the yield and kinetics of released enzyme, it has been shown previously that protein charge influences the efficiency of the late stages of secretion (Stephenson *et al.* 2000). The cytoplasmic membrane is surrounded by a thick cell wall, which consists of a heteropolymeric matrix of peptidoglycan and anionic polymers. The anionic polymers teichoic or teichuronic acid confer a high density of negative charge on the wall. As a result, the cell wall has anion-exchange characteristics with which proteins and cations can strongly interact (Harwood & Cranenburgh 2008). Consequently, positively charged and slowly folding proteins can be trapped in this structure. Therefore, the physico-chemical properties of each secretory protein, in particular its net charge, can affect the degree of this interaction and the efficiency of this latter stages of the secretion process, particularly the rate of passage through the wall. Naturally secreted *B. subtilis* proteins tend to have a neutral pI which may serve to limit

interactions with the cell wall (Kunst *et al.* 1997; Coxon 1990). In theory, proteins with an overall positive charge are more likely to interact with the negatively charged cell wall, while proteins with neutral or negative charge are likely to interact weakly with or are repelled by the cell wall. The online tool Compute pI/Mw tool from the ExPASy Bioinformatics Resource Portal was used to compute the theoretical pI of XynA and AmyM, which were estimated to be 9.05 and 5.48, respectively (Gasteiger *et al.* 2005). Our evidence shows that despite having an overall higher net charge, XynA folds and crosses the cell wall faster than AmyM, which has a lower pI and would be predicted to be repelled by the cell wall. However, as indicated above, the structures and sizes of AmyM and XynA are very different and other physicochemical properties like shape and surface charge are likely to contribute for discrepancy in secretion efficiency. Other possible explanation for the differences in secretion kinetics is that the wall acts as a diffusion barrier for the larger AmyM protein, in which case its kinetics will in part reflect the rate of cell wall turnover.

In general, these pulse-chase experiments indicate that XynA is a rapidly translocated and secreted enzyme that suffers from some cell-associated degradation during the later stages of secretion. The slow kinetics of AmyM released into the culture medium suggests an as yet indeterminate bottleneck, but might involve diffusion across the cell wall. A similar studied published by Bolhuis *et al.* (1999) is consistent with our results in the sense that several secretory proteins were associated with individual bottlenecks such as precursor processing, folding limitations and proteolytic degradation. The potential influences of cell-associated and extracellular proteolysis on AmyM and XynA production will be further studied in Chapter 6. For now, these results help to understand the reduced cell viability of the production strain BWAP compared to BWXP (Figure 3.3-A) and, most importantly, the difference in their production profiles (Figure 3.4).

A



B

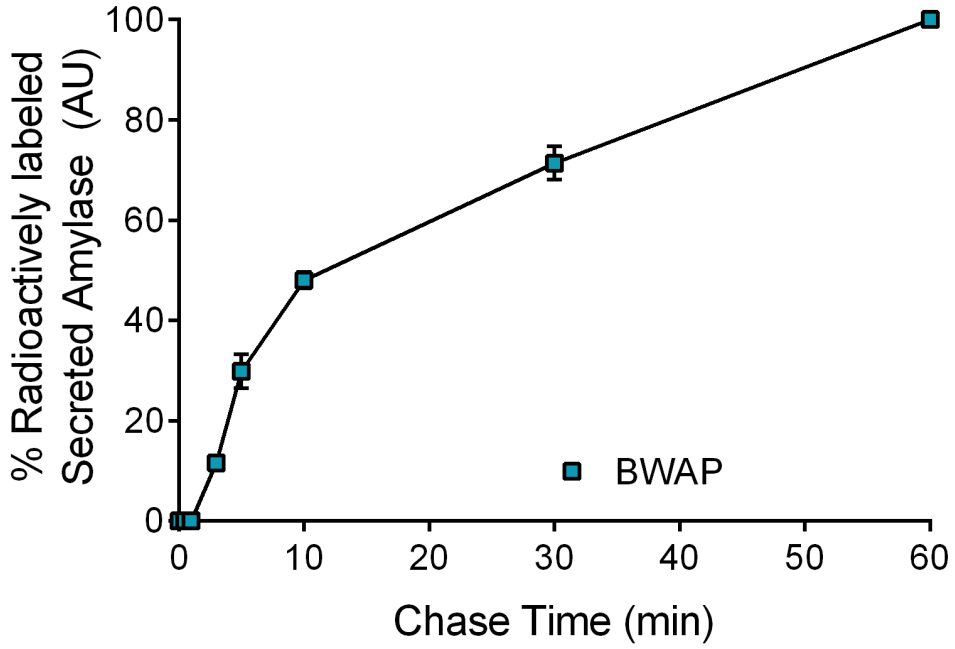


Figure 3.9 Pulse-chase analysis of (A) XynA and (B) AmyM production and secretion in the production strains BWXP and BWAP, respectively. The cultures were grown, harvested and processed according to Section 2.9. The quantified radioactively labelled proteins were plotted relative to the maximum protein levels detected in the whole-culture. In B, the average of two technical replicates of the protein quantification step are plotted with the corresponding SEM.

3.3.8 Up-regulation of the secretion and stress response mechanisms

It has been shown that under similar gene expression conditions (Figure 3.8-A), strains BWXP and BWAP show considerably different production profiles for, respectively, XynA and AmyM (Figure 3.4). These differences are likely to reflect the distinctive secretion kinetics of each enzyme (Figure 3.9). The rate of enzyme release is affected by how each enzyme interacts with the secretion machinery and its final yield is dependent on its stability during this process. The upregulation of the proteins involved in these processes was investigated in both production strains in order to detect any enzyme-specific effects.

B. subtilis translocates secretory proteins through the Sec-dependent translocase consisting of the SecA dimer, the heterotrimeric pore (SecYEG) and the heterodimeric SecDF-YrbF complex (Harwood & Cranenburgh 2008). The number of translocases is related to growth rate since the gene encoding SecY, the main pore-forming component of the translocase, is located in a ribosomal operon that is down-regulated during transition from exponential phase to stationary phase (Yang *et al.* 2013). Additionally, several studies have shown different levels of *secA* expression in the exponential and stationary phase of growth (Herbort *et al.* 1999; Yang *et al.* 2013). However, it has not been addressed whether high levels of expression of a secretory protein affect directly, or indirectly, the regulation of these translocase components, or if this is growth rate/phase dependent.

Transcription analysis of the genes comprising the Sec-dependent translocase was carried in the strains BWXP, BWAP, and 168. Cell cultures were grown in 1.5 mL LB, at 37 °C with 95% humidity and strong agitation (800 rpm), in FlowerPlate® 48-well MTPs using the BioLector® system (m2p-labs). As described previously, samples were harvested at the end of the exponential phase of growth (t_1, Figure 3.8), total RNA extracted and processed (Section 2.7.). The relative expression of the translocase genes in strains BWXP and BWAP relative to strain 168 is shown in Figure 3.10-A, using the $\Delta\Delta C_t$ quantification method (Bustin *et al.* 2009; Livak & Schmittgen 2001). Analysis of significance was carried using an unpaired *t*-test with Welch's corrections for different standard deviations and assuming a Gaussian distribution. Appendix G contains a summary of the results in a table format (Table G.3), including an alternative C_t normalisation using the *sdhA* gene (Table G.4).

The data in Figure 3.10-A (Table G.3, Appendix G) indicate that *secA, secY, secE, secG* and *secDF* do not show significantly different relative gene expression in the production strains,

under the conditions tested. This indicates that the cell does not increase the number of translocase channels in response to higher levels of protein production. However, *yrbF* is significantly downregulated in both BWXP and BWAP, showing a $\Delta\Delta C_t$ expression value of 0.45±0.04 and 0.38±0.07, respectively (Table G.3, Appendix G). YrbF is the orthologue of YajC from *Escherichia coli* and very little is known about its function and expression (Fang & Wei 2011). YrbF is a 9 kDa non-essential membrane protein identified in the SecDF-YrbF complex. Previous studies in *E. coli* have described the involvement of the orthologue SecDF-YajC complex in assisting the interaction between YidC, which facilitates the translocation of membrane proteins into the lipid bilayer, and the SecYEG complex (Sachelaru *et al.* 2013). Such role has not been demonstrated in *B. subtilis* and the exact function of YrbF remains unknown. This data offers the first insight into the regulation of this protein under the conditions of high levels of secretory enzyme production. The down-regulation of YrbF suggests its redundancy or cellular impairment under these conditions.

A similar down-regulation trend was observed for *secDF* relative expression even though insufficient data was obtained to confirm its significance (Figure 3.10-A). SecDF forms a complex with YrbF and it is expected that both proteins are similarly regulated. SecDF has been described to be required for efficient translocation across the membrane and processing of secretory proteins under conditions of hypersecretion (Bolhuis *et al.* 1998). This seems to be in disagreement with our data, particularly since both the literature and our study were performed under similar conditions, namely towards the end of exponential growth in rich media, (Bolhuis *et al.* 1998). However, we suggest that the reduced levels of secreted enzyme in the SecDF mutants described by Bolhuis *et al.* (1998) are a consequence of the role of SecDF in efficient growth, as evidenced by the mutants' growth defect at low temperatures. The level at which SecDF seemed to be required for efficient protein secretion depends on the level of synthesis and on the secretory protein(s) involved, which reveals a complex function and regulation of this protein complex.

After reaching the *trans* side of the membrane, efficient release of secretory proteins into the culture medium relies on rapid folding at the membrane/wall interface before their passage through the cell wall (Stephenson *et al.* 1998; Harwood & Cranenburgh 2008). Protein folding is facilitated by a combination of intrinsic and extrinsic factors, such as chaperones and folding factors. PrsA is a lipoprotein essential for growth and production of some secretory proteins in *B.*

subtilis which functions as an extracellular but cell-associated folding chaperone or foldase, reducing secretory proteins susceptibility to proteolysis (van Wely *et al.* 2001; Pohl & Harwood 2010; Harwood & Cranenburgh 2008). Several studies (e.g. Chen *et al.*, 2015) have addressed the positive effect of PrsA overexpression on enzyme production in *B. subtilis* since the original work by Kontinen & Sarvas (1993). However, it has not yet been addressed whether the overexpression of secretory proteins has an effect on PrsA expression. Figure 3.10-B shows that compared to strain 168, *prsA* is significantly up-regulated in strains BWXP and BWAP with a $\Delta\Delta C_t$ expression value of 2.16±0.37 and 2.80±0.23, respectively (Table G.3 - Appendix G). These data indicate the presence of a regulatory system that up-regulates the expression of this chaperone under conditions of high-level secretion. This hypothesis has been previously suggested by Krishnappa *et al.* (2013), who observed significantly enhanced amounts of PrsA detectable in the cell envelope of protease deficient mutants.

Such a regulatory system has been described by Hyyryläinen *et al.* (2001) and identified CssRS as a potential two-component regulatory system involved. This system is required for the cell to survive the secretion stress resulting from high-level enzyme production. This regulatory system insures that misfolded proteins that accumulate at the membrane/wall interface do not interfere with cell growth by blocking cell wall synthesis or the secretory translocase itself (Harwood & Cranenburgh 2008). To this end the CssRS two component system regulates the expression of the quality control proteases HtrA and HtrB (Hyyrylainen *et al.* 2001; Darmon *et al.* 2002; Westers *et al.* 2006a; Trip *et al.* 2011). Further studies are required to establish whether CssRS or a similar regulatory system is responsible for upregulating the expression of PrsA.

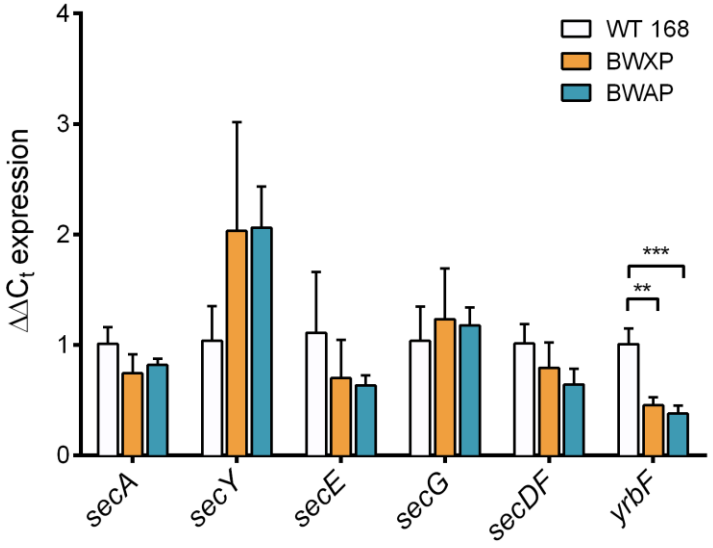
The RT-qPCR analysis confirmed that the CssRS-mediated secretion stress response was activated in strains BWXP and BWAP (Figure 3.10-B). Previous studies have shown that the CssRS-dependent response can be triggered by both native and heterologous secretory proteins (Westers *et al.* 2006a). Accordingly, both BWXP and BWAP show significant up-regulation of these quality control proteases with a higher level of induction observed in response to the production of the heterologous AmyM. Remarkably, *htrA* is more than 20-fold and *htrB* more than 30-fold up-regulated in the production strain BWAP (Table G.3 - Appendix G). These results could be interpreted as indicating that AmyM is sensed as being “foreign” by the cells and likely to be subjected to more extensive proteolytic degradation as part of the quality control processes aimed at maintaining the efficient functioning at the membrane/wall interface. This is

in agreement with the enzyme activity profile of AmyM (Figure 3.4) and slow kinetics of secretion (Figure 3.9-B).

The data in Figure 3.10-B indicate that *htrC* is not up-regulated like its homologues *htrA* and *htrB*. Pohl *et al.* (2013) showed that *htrC* is up-regulated to compensate for the absence of HtrA and HtrB, probably because their absence leads to cell wall stress. The results for *htrC* are as expected since the presence of HtrA and HtrB in the strains BWXP and BWAP has presumably successfully avoided enzyme over production from interfering with cell wall synthesis. Stephenson & Harwood (1998) demonstrated that the wall protease WprA targets the production of the heterologous AmyL from *B. licheniformis* by promoting the proteolysis of misfolded proteins following release form the translocated into the growth medium. The role is likely to be similar to that of HtrA and HtrB. However, our results indicate that WprA is not part of a regulation mechanism that is activated under conditions of high-levels of enzyme production (Figure 3.10-B).

Together, these results reveal that high levels of XynA and AmyM production have similar effects on the regulation of some components of the secretion machinery (YrbF, PrsA, HtrA and HtrB), reflecting a general secretion stress response. The effect is stronger in the case of the heterologous AmyM, presumably reflecting the fact that unlike XynA, this protein has not co-evolved with *B. subtilis*. This indicates that these regulation mechanisms respond similarly but with different intensities to individual secretory proteins, and this well illustrates the challenges faced in designing each production strain, particularly when trying to produce high levels of a heterologous protein.

A



B

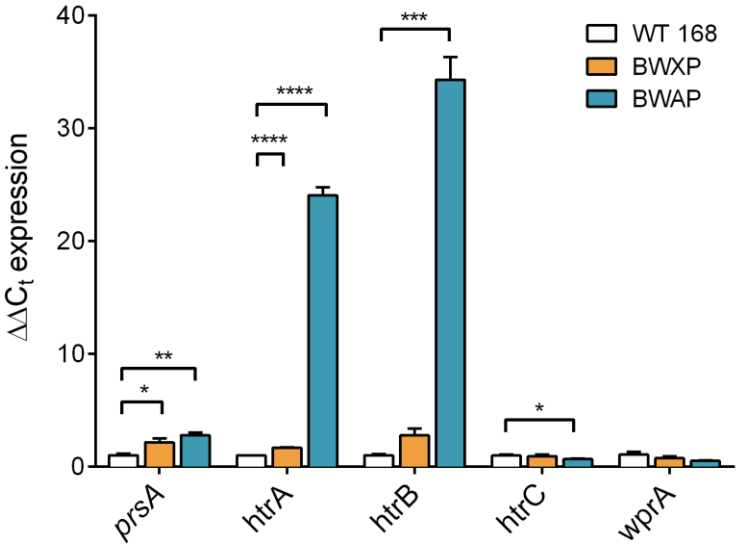


Figure 3.10 Relative expression of components of the secretory apparatus (A) and proteins involved folding (*prsA*) and quality control proteolysis (B) in the strains BWXP and BWAP relative to the wild type 168 strain. The cultures were harvested at the end of exponential phase of growth (Figure 3.8-A) and total RNA extracted (Section 2.7.). The cells were grown in 1.5 mL LB, at 37 °C with 95% humidity and vigorous agitation (800 rpm), in FlowerPlate® 48-well MTPs using the BioLector® bench-top microfermentation system (m2p-labs). The C_t values were averaged from two technical replicates and the ΔC_t values were calculated using the *fbaA* gene as control. The $\Delta\Delta C_t$ expression values are relative to the wild type strain 168 and correspond to the average between two biological replicates. The $\Delta\Delta C_t$ values are plotted with the corresponding SEM and the significance analyses correspond to unpaired *t*-tests with Welch's corrections for different standard deviations, assuming a Gaussian distribution. A summary of the results in a table format can be found in Appendix G.
$P > 0.05$; *: $P \leq 0.05$; **: $P \leq 0.01$; ***: $P \leq 0.001$; ****: $P \leq 0.0001$

3.4. Conclusions

This chapter describes a comprehensive study of two strains, one of which produces the native enzyme XynA and the other the α-amylase AmyM from *G. stearothermophilus.* These model enzymes where characterized biochemically, revealing that XynA is optimally active at pH 7.1 and 60 ºC, while AmyM is a highly thermostable enzyme with optimal pH at 6.5 and optimal temperature at 90 ºC (Figure 3.2).

An analysis of strains BWXP and BWAP showed two very different enzyme production profiles (Figure 3.4). Whereas xylanase production by BWXP coincided with the start of exponential growth, amylase production was initiated during transition from exponential to stationary growth.

These results were unexpected since these strains share an identical expression system and similar expression levels were confirmed by RT-qPCR (Figure 3.8). Therefore, the different production profiles appear to be related to post-transcriptional processes involving either membrane translocation, or interaction with or processing by the protein secretion pathway.

Our results indicate that the native signal peptides do not play a crucial role in these differences, likely due to the fact that each signal peptide has co-evolved with its cognate mature protein (Figure 3.7). However, pulse-chase analysis revealed clear difference between the secretion kinetics of these two enzymes (Figure 3.9); XynA is synthesised and secreted rapidly, taking less than a minute to be translated, folded and secreted, and is subjected to limited proteolysis. In contrast, AmyM, is released slowly into the growth medium, suggesting extensive cell-associated bottlenecks, in addition to a degree of degradation. This was further supported by the very strong upregulation of the quality control proteases HtrA and HtrB in response to AmyM (Figure 3.10-B). A similar but much more limited response was identified during XynA production.

Both enzymes also have similar effects on the regulation of other components of the secretion machinery, namely down-regulation of YrbF, previously associated with secretion efficiency, and up-regulation of PrsA, the major membrane bound extracellular chaperone.

These results provide new insight into the molecular effects of high levels of protein production and illustrate the different challenges involved of designing and optimising a production strain, particularly when the aim is high level synthesis and secretion of a heterologous enzyme.

Chapter 4 The impact of secretion stress on protein production by *Bacillus subtilis*

4.1. Introduction

The exploitation of *Bacillus subtilis* as an industrial workhorse for protein production has inspired several comprehensive studies on the secretion machinery of this organism and, particularly, on the bottlenecks associated with the high-level production of secretory proteins.

Two of the most recognised bottlenecks are the quality control mechanisms, which detect and target slowly folding or misfolded proteins in the secretion pathway and cell wall, and the powerful cocktail of secreted non-discriminatory proteases, the so-called "feeding proteases" (Pohl & Harwood 2010b; Stephenson & Harwood 1998; Westers *et al.* 2006b). Both of these protein production bottlenecks are essential cellular processes that ensure cell fitness during environmentally challenging conditions. The well characterised CssRS two-component system regulates the up-regulation of the HtrA and HtrB proteases upon hyper-secretion conditions, and prevents potentially fatal obstruction of the secretory translocase and cell wall synthesis (Noone *et al.* 2000 ; Darmon *et al.* 2002; Westers *et al.* 2006b; Antelmann *et al.* 2003). On the other hand, the so-called "feeding proteases" facilitate the breakdown of proteins in the extracellular environment, primarily for the uptake of essential amino acids and peptides (Pohl & Harwood 2010).

Recombinant DNA technology, coupled with Synthetic Biology, facilitates the manipulation of these cellular processes in order to overcome production limitations and maximize the secretion of specific products of interest. Moreover, such mechanisms have inspired biosensor-like systems for the hypersecretion of protein products. Trip *et al.* (2011) made use of the CssRS sensing mechanism to engineer a biosensor for the overexpression of heterologous secretory proteins by fusing the secretion stress responsive promoter, P_{htrA} to sf*gfp.* This fluorescent reporter was shown to be specific for extracellular protein accumulation. Similarly, Ploss *et al.* (2016) made use of the *htrB* promoter to demonstrate heterogeneous activation of the secretion stress pathway following overproduction of AmyM from *Geobacillus stearothermophilus*. Such reporter systems

represent a versatile tool for real-time monitoring of important cellular processes during protein production.

Addressing the bottleneck of proteolysis, Pohl *et al.* (2013) engineered a set of strains with systematic deletions of the proteases genes associated with the instability of secreted heterologous proteins (Figure 4.1). These proteases have been classified either as quality control or feeding proteases depending on whether they are associated with the targeting of misfolded or slowly folding proteins, or with the degradation of proteins in the culture medium, respectively.

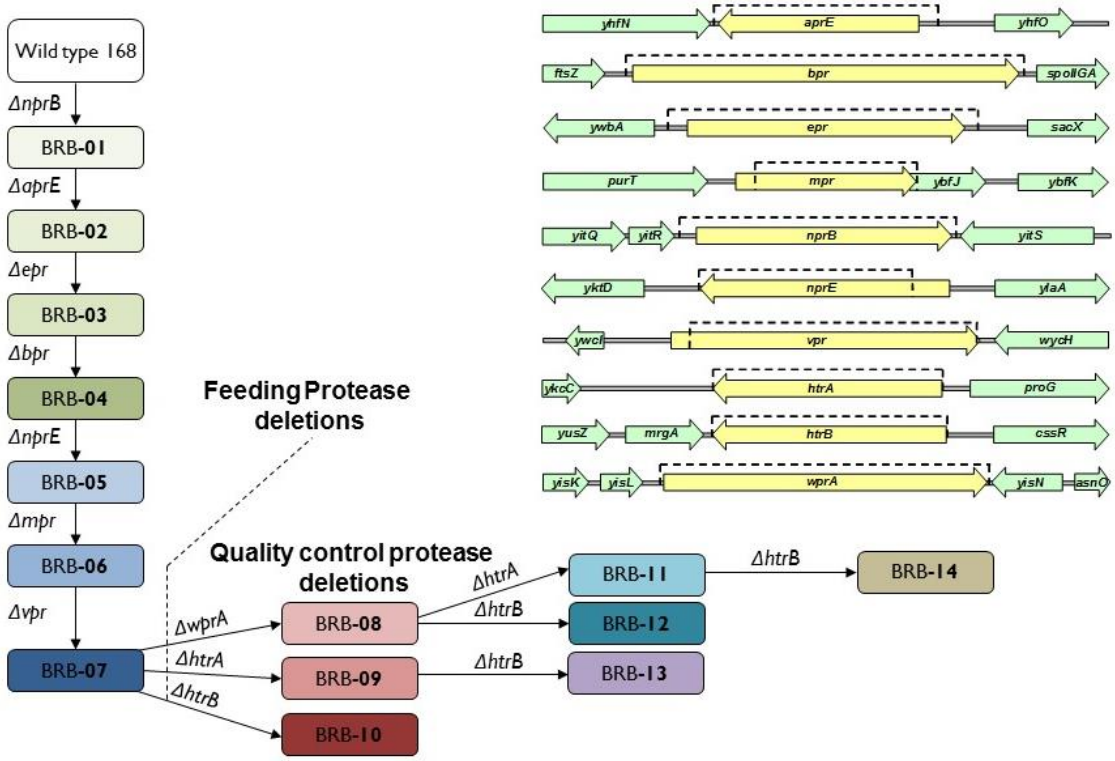


Figure 4.1 Representation of the sequential construction of protease deletion strains. Dashed lines indicate the regions deleted in the wild type loci of the protease genes (Pohl *et al.* 2013).

This collection of strains provides a valuable resource for the optimisation of dedicated hosts for protein production and provides a platform for investigating the relations between proteolysis and other important cellular mechanisms. Krishnappa *et al.* (2013) presented an approach to distinguish the roles of the quality control proteases HtrA and HtrB from other proteases present in the cell wall and extracellular environment by comparing the exoproteomes of the BRB collection of strains. Besides the already described quality control roles in the secretion pathway, the authors suggest other important roles for HtrA and HtrB in the folding of native secretory proteins, release of lipoproteins and degradation of membrane proteins. Additionally, Krishnappa

et al. (2014) demonstrated how HtrA, HtrB and the main membrane bound extracellular chaperone PrsA are substrates of multiple extracytoplasmic proteases, providing additional insights into the complex relations between native secretory proteins and the proteolytic network.

Up until now, all the studies with the BRB collection of extracytoplasmic protease deficient hosts have been done using the native genetic background of each mutant. In this Chapter, we contribute more complex studies on these strains by including a secretion stress reporter system coupled with native and heterologous enzyme production. This allowed us to correlate both the bottlenecks of cellular and extracellular-associated proteases with native and heterologous enzyme production, and the secretion stress response. Furthermore, the analysis of the effects of homologous and heterologous enzyme production on secretion stress in the different hosts provides important information about how each product of interest presents different challenges for maximising production.

4.2. Strains and plasmids

The strains and plasmids used in this chapter are listed in Table 4.1 and Table 4.2, respectively. DNA manipulation and transformation methods are described in Sections 2.5. and 2.6. , respectively. Appendix C contains a list of the primers used. Appendix B contains the plasmid schemes from Table 4.2.

B. subtilis wild type strain 168 was used as the host for all the strains in this study. Strains BRB01-13 were obtained from Cobra Biologics via the Harwood strain collection (Pohl *et al.* 2013). The construction of strains BWAP, BWAP2, BWXP and BWXP2 was described in Section 3.2. The source and construction of plasmids pCS58, pCS73, pRC67 and pRC68 was described in Section 3.2. Plasmid pCS74 was obtained from the plasmid collection at DSM it is a derivative of plasmid pNAPHB27 (Quax & Broekhuizen 1994) and was used as an empty plasmid control for the XynA and AmyM production studies. Plasmid pCS72 was obtained from the DSM plasmid collection and was used as template for the amplification of the chloramphenicol resistance gene.

The strains expressing a copy of proteases *htrA, htrB, htrC* and *wprA* under the control of a xylose inducible promoter in the *amyE* locus (BCS234, BCS291, BCS292 and BCS293, respectively) were constructed by Sauer (2016). Plasmids pCS58 and pCS73 were transformed separately to strains BCS234, BCS291, BCS292 and BCS293 in collaboration with Sauer (2016)

creating strains BCS237, BCS294, BCS295, BCS296, BCS297, BCS298, BCS299 and BCS300. These were used to study the effects of overexpressing the HtrA, HtrB, HtrC and WprA proteases on the production of the model enzymes XynA and AmyM (Section 4.6.).

To construct a library of protein secretion stress reporter strains, the fluorescent reporter P_{htrA}-*sfGFP* was amplified from the chromosome of strain BCS031 obtained from Sauer (2016). The *lacA*::(*spec*, P_{htrA}-*sfGFP*) integration cassette (Figure 4.2) was amplified via PCR using the primers 319 and 320 and the chromosome of BCS031 as template. The purified PCR product was used to transform strains 168 and BRB01-13, creating strains BWR and BRB01R-BRB13R, respectively.

Plasmids pCS58, pCS73, pRC67 and pRC68 were transformed into strain BWR creating strains BWXP, BWAP, BWXP2 and BWAP2, respectively, to study the effect of XynA and AmyM production in the secretion stress response (Section 4.3.).

Plasmids pCS58, pCS73 and pCS74 were transformed into strains BRB01R-BRB13R creating a collection of protease deficient mutants carrying a secretion stress reporter and a high copy number plasmid expressing either XynA, AmyM or no enzyme, respectively. The suffixes XP, AP and EP where added to the strain names BWR and BRB01R-BRB13R when carrying the plasmids pCS58, pCS73 and pCS74, respectively. These strains were used for a comprehensive study of the impact of secretion stress on protein production by different hosts (Section 4.4. and 4.5.).

Gene deletions of *htrA* and *htrB* were created by transforming the wild type 168 strain with a PCR-amplified DNA fragment consisting of the chloramphenicol resistance gene flanked by the upstream and downstream regions of each protease gene. The *htrA* upstream and downstream regions were amplified from the chromosome of wild type 168 with the primers 410 and 411, and 412 and 413, respectively. The *htrB* upstream and downstream regions were amplified with the primers 414 and 415, and the 416 and 417, respectively. The chloramphenicol resistance gene was amplified from plasmid pCS72 with the primers 406 and 407 for construction of the *htrA* deletion cassette, and the primers 408 and 409 for the *htrB* deletion cassette. The chloramphenicol resistance gene was flanked by the upstream and downstream regions of either *htrA* or *htrB* via Gibson Assembly. The resulting assembly reaction was used as template for the amplification of the deletion cassettes using primers 410 and 413 for the *htrA* deletion cassette,

and primers 414 and 417 for the *htrB* deletion cassette. The purified PCR products were transformed to strain BWR creating strains BWRΔHtrA and BWRΔHtrB. Finally, plasmids pCS58 and pCS73 were transformed separately to each of these strains creating strains BWRXPΔHtrA, BWRAPΔHtrA, BWRXPΔHtrB and BWRAPΔHtrB. The resulting strains were used to study the effect of *htrA* and *htrB* deletions in the secretion stress and production of XynA (-XP) and AmyM (-AP) (Section 4.4. and Section 4.5. , respectively).

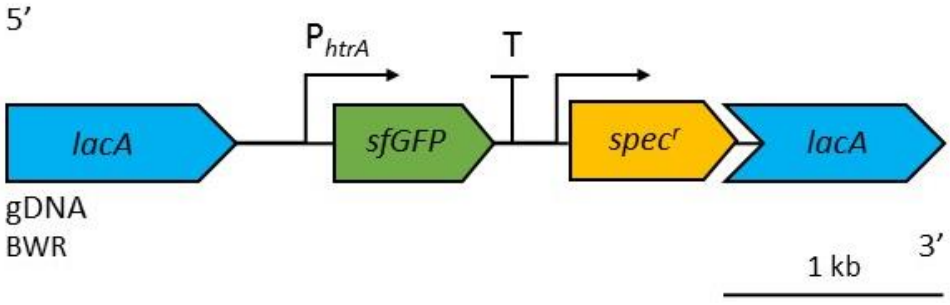


Figure 4.2 Schematic overview of the protein secretion stress reporter cassette. The SigA-dependent *htrA* promoter (P_{htrA}) regulated by CssR is located upstream of sfGFP and followed by the transcription terminator (T) of the maltogenic α-amylase of *Geobacillus stearothermophilus*. The integration of the stress reporter via double homologous recombination into the *lacA* locus of *Bacillus subtilis* was selected by spectinomycin resistance (*spec*^{r}).

Table 4.1 Summary of the strains used in the study of this chapter.

Strain	**Genotype**	**Source**
B. subtilis		
168	*trpC2*	Kunst *et al.* (1997)
BCS031	168 *aprE*::P_{htrA}-*sfGFP*, *spec*	Sauer (2016)
BCS234	168 *amyE*::P_{xyl}-*htrA*, *spec*	Sauer (2016)
BCS291	168 *amyE*::P_{xyl}-*htrB*, *spec*	Sauer (2016)
BCS292	168 *amyE*::P_{xyl}-*htrC*, *spec*	Sauer (2016)
BCS293	168 *amyE*::P_{xyl}-*wprA*, *spec*	Sauer (2016)
BCS237	168 *amyE*::P_{xyl}-*htrA*, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BCS294	168 *amyE*::P_{xyl}-*htrA*, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BCS295	168 *amyE*::P_{xyl}-*htrB*, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BCS296	168 *amyE*::P_{xyl}-*htrC*, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BCS297	168 *amyE*::P_{xyl}-*wprA*, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BCS298	168 *amyE*::P_{xyl}-*htrB*, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work

Strain	Genotype	Source
BCS299	168 *amyE*::P_{xyl}-*htrC*, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BCS300	168 *amyE*::P_{xyl}-*wprA*, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB03RAP	168 Δ*nprB* Δ*aprE* Δ*epr; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB03RXP	168 Δ*nprB* Δ*aprE* Δ*epr; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB03REP	168 Δ*nprB* Δ*aprE* Δ*epr; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BRB04	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr*	(Pohl *et al.* 2013)
BRB04R	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr; lacA*::P_{htrA}-sfGFP, *spec*	This work
BRB04RAP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB04RXP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB04REP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BRB05	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE*	(Pohl *et al.* 2013)
BRB05R	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE; lacA*::P_{htrA}-sfGFP, *spec*	This work
BRB05RAP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB05RXP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB05REP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BRB06	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr*	(Pohl *et al.* 2013)
BRB06R	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr; lacA*::P_{htrA}-sfGFP, *spec*	This work
BRB06RAP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB06RXP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB06REP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BRB07	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr*	(Pohl *et al.* 2013)
BRB07R	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr; lacA*::P_{htrA}-sfGFP, *spec*	This work
BRB07RAP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB07RXP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB07REP	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BRB08	168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr*, Δ*wprA*	(Pohl *et al.* 2013)

Strain	Genotype	Source
BRB08R	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr*, Δ*wprA; lacA*::P_{htrA}-sfGFP, *spec*	This work
BRB08RAP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr*, Δ*wprA; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB08RXP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr*, Δ*wprA; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB08REP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr*, Δ*wprA; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BRB09	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrA*	(Pohl *et al.* 2013)
BRB09R	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrA; lacA*::P_{htrA}-sfGFP, *spec*	This work
BRB09RAP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrA; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB09RXP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrA; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB09REP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrA; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BRB10	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrB*	(Pohl *et al.* 2013)
BRB10R	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec*	This work
BRB10RAP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB10RXP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB10REP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BRB11	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*wprA* Δ*htrA*	(Pohl *et al.* 2013)
BRB11R	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*wprA* Δ*htrA; lacA*::P_{htrA}-sfGFP, *spec*	This work
BRB11RAP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*wprA* Δ*htrA; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB11RXP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*wprA* Δ*htrA; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB11REP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*wprA* Δ*htrA; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BRB12	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*wprA* Δ*htrB*	(Pohl *et al.* 2013)
BRB12R	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*wprA* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec*	This work
BRB12RAP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*wprA* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB12RXP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*wprA* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB12REP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*wprA* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BRB13	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrA* Δ*htrB*	(Pohl *et al.* 2013)

Strain	Genotype	Source
BRB13R	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrA* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec*	This work
BRB13RAP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrA* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRB13RXP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrA* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BRB13REP	168 Δ*nprB* Δa*prE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr* Δ*vpr* Δ*htrA* Δ*htrB; lacA*::P_{htrA}-sfGFP, *spec;* pCS74 (P_{amyQ}), *neo*	This work
BWAP	168 pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BWAP2	168 pRC68 (P_{amyQ}-*amyM2*), *neo*	This work
BWXP	168 pCS58 (P_{amyQ}-*xynA*), *neo*	Sauer (2016)
BWXP2	168 pRC67 (P_{amyQ}-*xynA2*), *neo*	This work
BWR	168 *lacA*::P_{htrA}-sfGFP, *spec*	Sauer (2016)
BWRΔHtrA	168 *lacA*::P_{htrA}-sfGFP, *spec*; *htrA*::*cat*	This work
BWRΔHtrB	168 *lacA*::P_{htrA}-sfGFP, *spec*; *htrB*::*cat*	This work
BWRAP	168 *lacA*::P_{htrA}-sfGFP, *spec*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BWRAPΔHtrA	168 *lacA*::P_{htrA}-sfGFP, *spec*; *htrA*::*cat*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BWRAPΔHtrB	168 *lacA*::P_{htrA}-sfGFP, *spec*; *htrB*::*cat*; pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BWRAP2	168 *lacA*::P_{htrA}-sfGFP, *spec*; pRC68 (P_{amyQ}-*amyM2*), *neo*	This work
BWRXP	168 *lacA*::P_{htrA}-sfGFP, *spec;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BWRXPΔHtrA	168 *lacA*::P_{htrA}-sfGFP, *spec*; *htrA*::*cat;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BWRXPΔHtrB	168 *lacA*::P_{htrA}-sfGFP, *spec*; *htrB*::*cat;* pCS58 (P_{amyQ}-*xynA*), *neo*	This work
BWRXP2	168 *lacA*::P_{htrA}-sfGFP, *spec;* pRC67 (P_{amyQ}-*xynA2*), *neo*	This work

neo – neomycin resistance; *spec*- spectinomycin resistance; *cat*-chloramphenicol resistance *bleo* – bleomycin resistance

Table 4.2 Summary of the plasmids used in the study of this chapter.

Plasmid	Properties	Source
pCS58	P_{amyQ}-*xynA, reppUB, neo, bleo*	Sauer (2016)
pCS72	*npr*::(*cat, spec*); pUC, *bla*	DSM
pCS73	P_{amyQ}-*amyM, reppUB, neo, bleo*	DSM
pCS74	P_{amyQ}, *bla*, $cat_{(a)}$, *reppUB, neo, bleo*	DSM
pRC67	P_{amyQ}-*xynA2, reppUB, neo, bleo*	This work
pRC68	P_{amyQ}-*amyM2, reppUB, neo, bleo*	This work

neo – neomycin resistance gene; *bla* – ampicillin resistance; *cat* – chloramphenicol resistance; *bleo* – bleomycin resistance; (a) - defective

4.3. The impact of enzyme production in the secretion stress response

The BioLector® bench top microfermentation system (m2p-labs) facilitates simultaneous monitoring of biomass and fluorescence while controlling shaking speed, temperature and humidity. The BioLector® system was used to monitor the secretion stress response associated with the production of the endo-1,4-β-xylanase XynA from *B. subtilis*, and the maltogenic α-amylase AmyM from *Geobacillus stearothermophilus*. The secretion stress response was quantified using an ectopically expressed fluorescent reporter consisting of the promoter controlling the transcription of the quality control protease HtrA fused to sfGFP (Cotlet *et al.* 2006; Trip *et al.* 2011).

Biological replicates were grown in triplicate in 48-well MTP FlowerPlates® in a randomised layout. The growth regime is described in Section 3.3.1 . Briefly, overnight cultures were diluted into fresh LB medium to give synchronize precultures which were used to inoculate pre-warmed LB medium in 48-well MTP FlowerPlates®. Growth was monitored for 26 hours at 37 °C with 95% humidity and vigorous agitation (800 rpm). Biomass (excitation: 620 nm, gain: 20) and GFP (excitation: 488 nm, emission: 520 nm, gain: 95) were monitored by systematic measurements every 15 min. Appropriate controls were incorporated by including the wild type strain 168 as a negative control and wells containing only sterile medium, allowing for control of enzyme expression effects and potential contamination during the experiment.

Figure 4.3 shows the secretion stress response induced in the wild type strain 168 when expressing the native XynA and AmyM proteins (Figure 4.3-A) and the precursors of XynA and AmyM with swapped signal peptides (Figure 4.3-B). The raw fluorescent signal generated from the P_{htrA} promoter was plotted together with the biomass values to allow visualisation of the secretion stress throughout the growth cycle.

It was shown previously in Chapter 3 that XynA and AmyM have very different production profiles under the same expression system and conditions, most likely due to different interactions and processing by the protein secretion pathway. The detection of active XynA in the culture medium coincides with the start of exponential growth, whereas amylase production is initiated during the transition from exponential to stationary growth (Figure 3.4). Moreover, pulse-chase experiments monitoring protein synthesis, translocation and processing revealed that XynA is synthesised and secreted rapidly and is subjected to limited proteolysis, whereas AmyM is released very slowly into the growth medium, suggesting extensive cell-associated bottlenecks

and degradation (Figure 3.9). Previous studies have shown that cells producing high levels of an heterologous protein significantly up-regulate the synthesis of the quality control proteases HtrA and HtrB (Antelmann *et al.* 2003; Westers *et al.* 2006b; Trip *et al.* 2011). qPCR analysis of the strains producing the heterologous AmyM and native XynA also revealed a strong up-regulation of the quality control proteases HtrA and HtrB which corroborates the hypothesis of extensive targeting of the secretory proteins by cell-associated proteolytic activity (Figure 3.10-B).

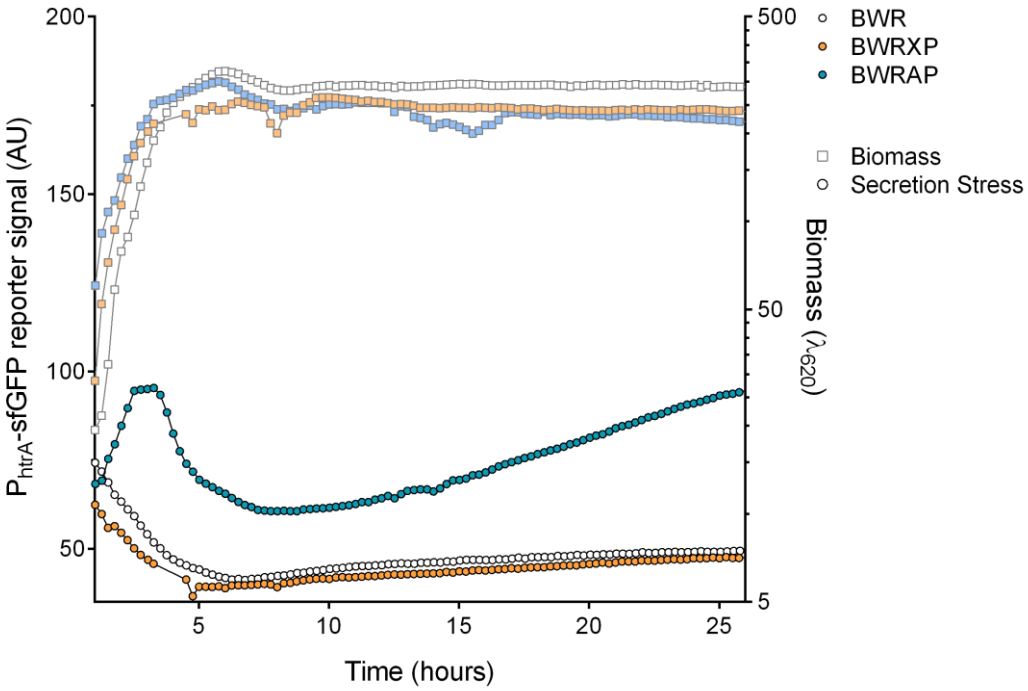
An analysis of the protein secretion stress response, shown in Figure 4.3-A, supports these previous observations. The results show an up-regulation of the secretion stress reporter P_{htrA}-sf*gfp* in the AmyM producing strain in comparison to the wild type strain 168. Interestingly, the fluorescent signal increases significantly in the stationary growth phase of BWRAP. This suggests that AmyM is secreted and causes significant secretion stress even during the late stages of growth. It was previously observed that AmyM is secreted after up to one hour after being synthesised (Figure 3.9-B). This slow kinetics of release might contribute for the increased secretion stress signal at late stages of growth even though there is no apparent increase in active enzyme production at this stage (Figure 3.4-B).

The wild type 168 and BWRXP strains, encoding XynA production, show identical GFP profiles, corresponding to a decrease in the initial fluorescent signal due to the consumption of nutrients present in the LB medium that contribute for background fluorescence (Figure 4.3-A). This indicates that the high levels of production of the native enzyme XynA do not cause a detectable up-regulation of the secretion stress response, despite the higher levels of *htrA* transcript in this strain (Figure 3.10-B). Presumably, the *htrA* up-regulation is not high enough to overcome the sensitivity limits of the reporter.

Strains BWXP2 and BWAP2, corresponding to the wild type strain 168 expressing the proteins XynA and AmyM with swapped signal peptides, were constructed in order to investigate the influence of the signal peptide on the enzyme production profile (Section 3.3.5). It was shown that secretion was most efficiently directed by their native signal peptides, indicating that the signal peptide has co-evolved with its cognate mature protein to optimise secretion (Figure 3.7). It was therefore interesting to determine whether the resulting drastic decreases in protein production associated with the heterologous signal peptides were related to the up-regulation of the quality control mechanisms in response to secretion stress. The data indicates that strain BWRXP2 exhibits virtually no secretion stress up-regulation in comparison to the wild type

strain (Figure 4.3-B), indicating that the reduction in XynA activity was likely associated with elements in the secretion pathway prior to release from the translocase. In the case of the AmyM producing strain, replacing the native signal peptide of AmyM by the one of XynA eliminates the pronounced secretion stress response found when expressing the original precursor form of AmyM (Figure 4.3-A). This indicates that the reduced production levels of the recombinant versions of XynA and AmyM (Figure 3.7) is again likely to be related to the early stages of secretion, rather than the later quality control stage.

A



B

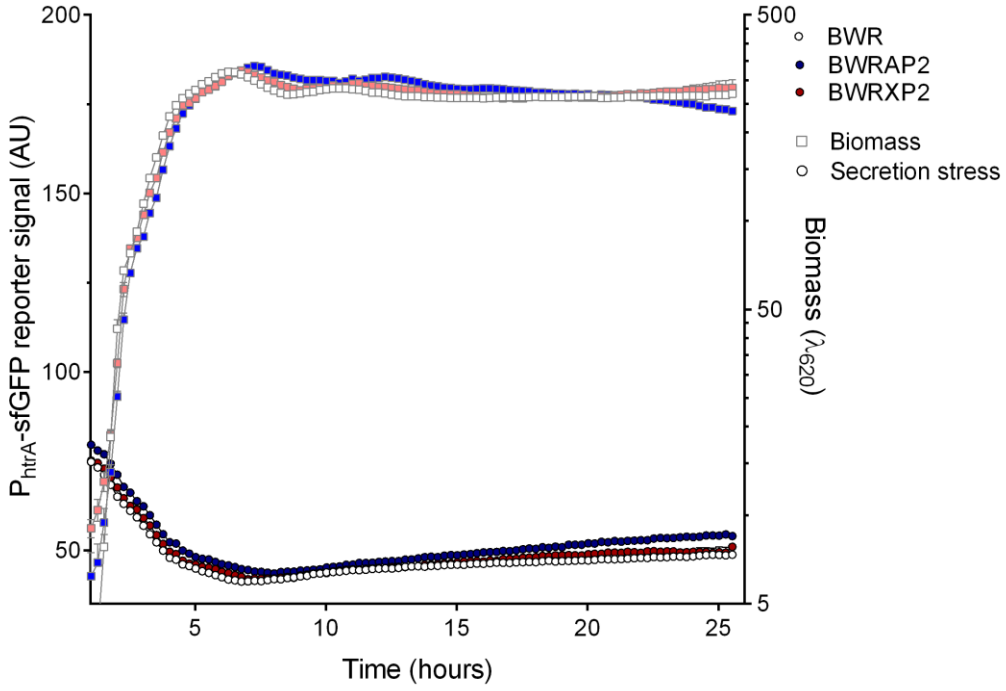


Figure 4.3 The protein secretion stress responses associated with the overexpression of XynA and AmyM using (A) the native and (B) swapped signal peptides, measured by live detection of the reporter P$_{htrA}$-*sfGFP* fluorescence reporter. The data was obtained using the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 1.5 mL LB in FlowerPlate® 48-well MTPs and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The average of either nine (A) or four (B) biological replicates per strain was plotted.

4.4. The impact of protease mutations on the secretion stress associated to enzyme production

The BioLector® microbioreactor system was used to monitor the secretion stress response of a collection of 13 extracytoplasmic protease deficient hosts producing the endo-1,4-β-xylanase, XynA, from *B. subtilis*, and the maltogenic α-amylase, AmyM, from *Geobacillus stearothermophilus* (Table 6.1). This was an extension of the study presented in the previous section, measuring the ectopic expression of the transcription fluorescent reporter P_{htrA}-*sfGFP* in all strains, including the wild type strain 168 as negative controls. Three independent biological replicates were grown, in triplicate (i.e. x9), in 48-well MTP FlowerPlates® in a randomised layout, in multiple identical experiments. Bacterial growth was synchronised as described previously (Section 4.3.) and biomass (excitation: 620 nm, gain: 20) and GFP (excitation: 488 nm, emission: 520 nm, gain: 95) were monitored by systematic measurements every 15 min for 26 hours at 37 °C with 95% humidity and vigorous agitation (800 rpm).

Figure 4.4 shows the secretion stress response induced in the 13 extracytoplasmic protease deficient hosts and wild type strain 168 when expressing the native XynA (Figure 4.4-A) and AmyM (Figure 4.4-B). The plotted fluorescent signal corresponds to the difference between the measured signal and the background signal of wild type 168, divided by biomass. The growth curves and raw fluorescent signals measured for each set of replicates are shown in Appendix J (Figure J.1).

In the case of XynA overexpression, Figure 4.4-A1 shows that the deletion of the seven known extracellular proteases of *B. subtilis* (NprB, AprE, Epr, Bpr, NprE, Mpr and Vpr) did not change the secretion stress response significantly in comparison to the wild type response. The stress signals are very low in comparison to other stress responses observed in Figure 4.4. Since the primary role of these proteases is likely to be nutritional, it was anticipated that their absence would be unlikely to affect the regulation of the quality control mechanisms at the membrane-wall interface. However, BRB07RXP, which lacks the expression of all seven extracellular proteases, did show a low but noticeable up-regulation of the secretion stress response. In contrast, there was a significant increase of the fluorescent secretion stress signal in strain BWRAP during entry into the stationary growth phase (Figure 4.4-B1), but this was attenuated with cumulative deletions of extracellular enzymes. This effect is presumably related to differences in the growth profile of these strains. The deletion of *aprE* in the strain BRB02 causes

a major change in the growth profile during stationary phase, leading to higher cell densities (Figure J.1, Appendix J). AprE is the major extracellular protease of *B. subtilis* and, together with NprE, accounts for ~95% of the extracellular proteolytic activity of the cell (Schmidt *et al.* 1993). Interestingly, previous studies have suggested that mutants deficient in extracellular proteases show increased rates of cell wall turn over and are more susceptible to cellular lysis during exponential growth and following the transition to stationary phase (Jolliffe *et al.* 1980; Coxon *et al.* 1991; Stephenson *et al.* 1999). These showed that multiple protease deficiencies result in increased cellular lysis following transition to stationary phase. These observations imply that at least some of the extracellular proteases, produced mostly during stationary phase, target the autolysins when their important role for growth and separation of dividing cells is no longer needed. Although most of the growth profiles of protease deficient strains shown in Figure 4.4 also show an initial decrease of optical density in the early stages of stationary growth, our experiments monitor the growth profile of protease-deficient strains during a total of 26 hours and up to 24 hours after transition to stationary phase, revealing optical densities higher than the wild type after removing AprE synthesis (Figure J.1, Appendix J). This effect is independent of the overproduction of the model enzyme products (Figure J.1, Appendix J), providing new insights on batch growth of protease deficient hosts for protein production. The apparently higher level of cell viability after the deletion of *aprE* coincides with a significantly lower secretion stress signal in comparison with BWRAP and BRB01RAP which show increasing levels of stress per unit biomass throughout the later stages of stationary phase. The reason for this is currently unknown.

Figure 4.4-A2 and Figure 4.4-B2 show that the cumulative deletions of quality control proteases genes *htrA* and *htrB* in the absence of feeding control proteases results in an up-regulation of the secretion stress response. This is particularly significant in strains producing the heterologous AmyM (Figure 4.4-B2). When either HtrA or HtrB is absent (strains BRB09RAP, BRB10RAP, BRB11RAP and BRB12RAP) there is a significant up-regulation of the secretion stress response in comparison to the BWRAP. This was to be expected because the CssRS system is known to be induced if either HtrA or HtrB is absent, suggesting that both proteins have compensatory roles and reinforcing the importance of both proteases for preventing the accumulation of slow or misfolded proteins at the membrane-wall interface (Pohl *et al.* 2013; Krishnappa *et al.* 2013). This was confirmed with our fluorescent reporter system with or without the overexpression of the model secretory enzymes (Figure J.2, Appendix J).

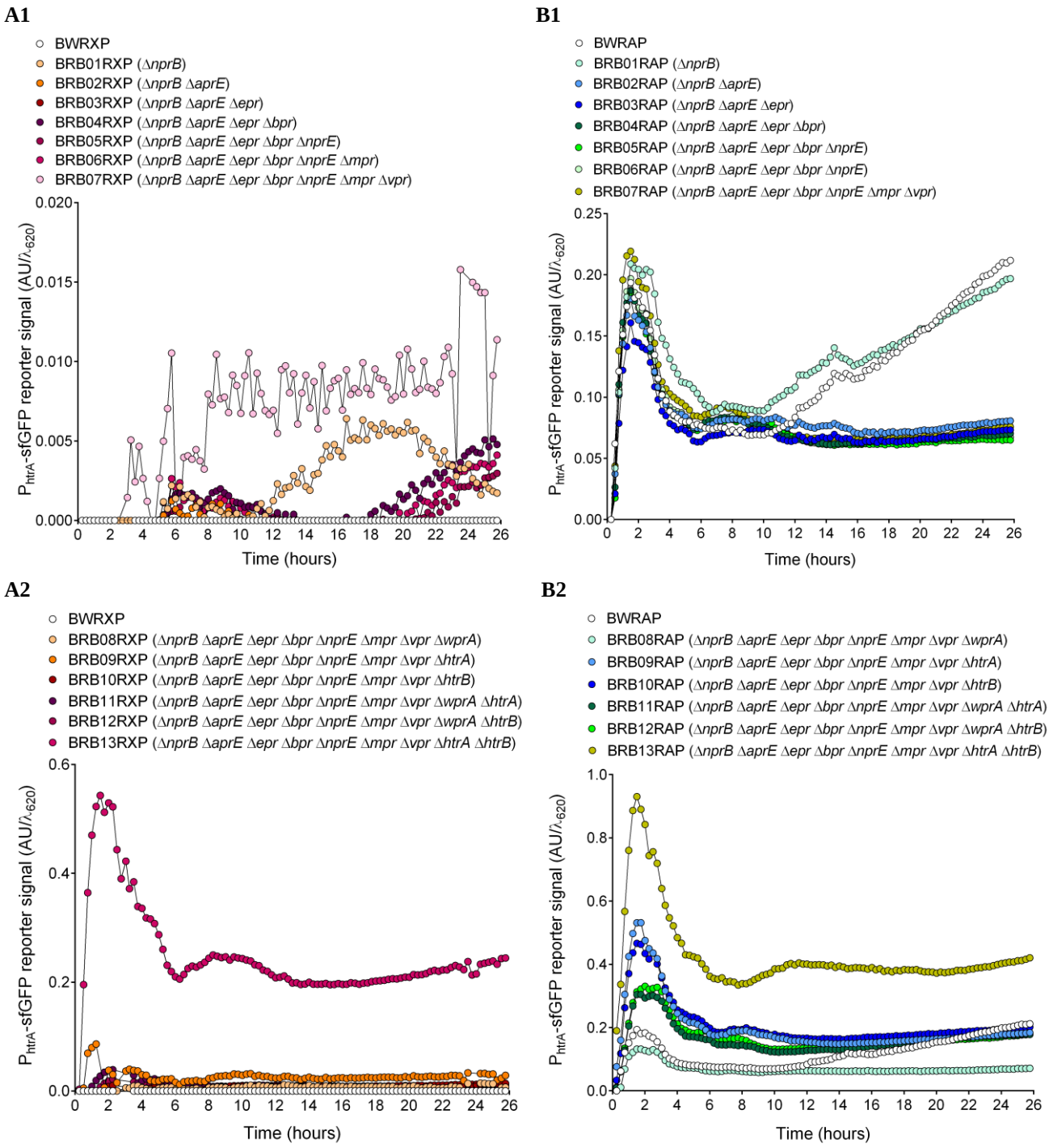


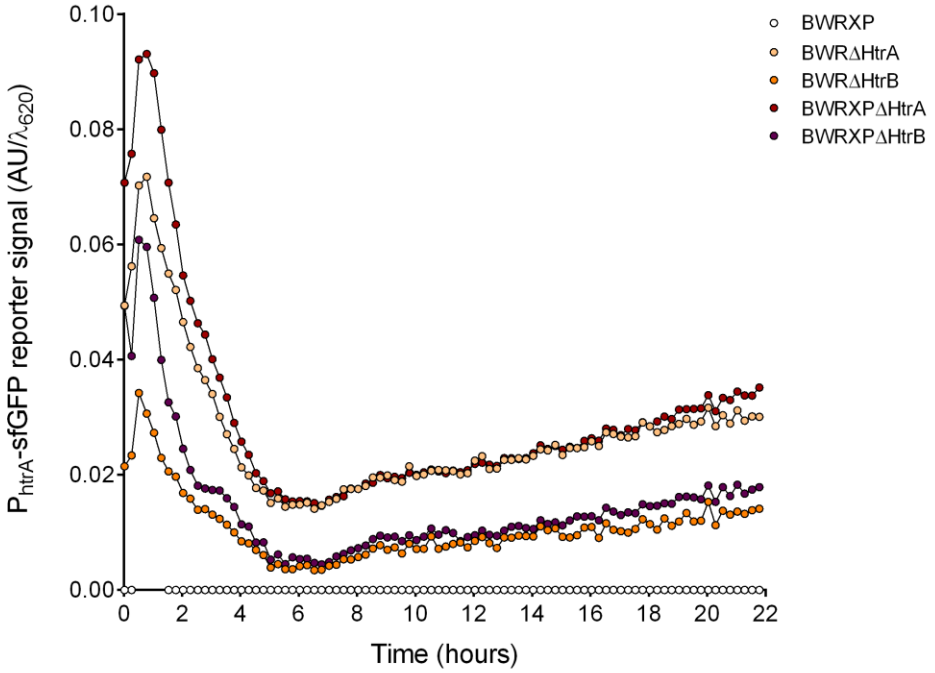
Figure 4.4 Protein secretion stress response upon overexpression of (A) XynA and (B) AmyM in strains lacking multiple extracytoplasmic proteases, measured by live detection of the reporter P_{htrA}-*sfGFP* fluorescence using the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 1.5 mL LB in FlowerPlate® 48-well MTPs and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The plotted fluorescence signal was blanked with the detected signal for the wild type strain 168 and normalised to the biomass. The average of three biological replicates per strain, grown in triplicate, was plotted.

Notably the deletion of WprA in strain BRB08RAP had no significant effect in the secretion stress response. However, when combined with deletions of either HtrA or HtrB (i.e. strains BRB11RAP and BRB12RAP), the absence of WprA reduces the secretion stress level in comparison with strains BRB09RAP and BRB10RAP, which lack either HtrA or HtrB (Figure 4.4-B2). WprA has been associated with the degradation of both HtrA and HtrB when increased amounts of these proteins were found in the cellular and extracellular environment of BRB08 (Krishnappa et al. 2014). Presumably, the increased stability of both HtrA and HtrB when WprA is absent leads to a lower secretion stress signal of strains BRB11RAP and BRB12RAP in comparison to BRB09RAP and BRB10RAP, respectively.

Strain BRB13, which contains deletions of both major control proteases HtrA and HtrB, as well as all seven feeding proteases, shows the highest level of expression of the secretion stress signal in all studied hosts for both XynA and AmyM production. However, it is important to note that this strain also shows higher levels of secretion stress response in its native state, without the overexpression of the model secretory enzymes (Figure J.2, Appendix J). Nevertheless, the fluorescent reporter signal is higher when the strains are producing either XynA or AmyM (Figure 4.4-A2 and Figure 4.4-B2, respectively). This reflects the importance of the compensatory regulation of both HtrA and HtrB, especially during high-level of secretory protein production.

To confirm that the observed up-regulation of the secretion stress response when either HtrA or HtrB are absent was not due to the heavily edited genetic background of the BRB strains, which lack all extracellular proteases, the study was repeated on strains which carried single knockouts of either *htrA* or *htrB* (Figure 4.5). As expected, the single knockouts of the quality control proteases genes resulted in a significant up-regulation of the secretion stress signal. Interestingly, the stress response in these strains is generally higher in comparison to strains BRB09, BRB10, BRB11 and BRB12 which, in addition to knockouts in either *htrA* or *htrB,* were deficient in all of the extracellular feeding proteases and, in the case of the last two strains, WprA as well. This suggests that removing all extracellular proteolytic activity of *B. subtilis* directly or indirectly affects the regulation of the protein secretion quality control mechanisms of the cell, an issue that has not previously been addressed.

A



B

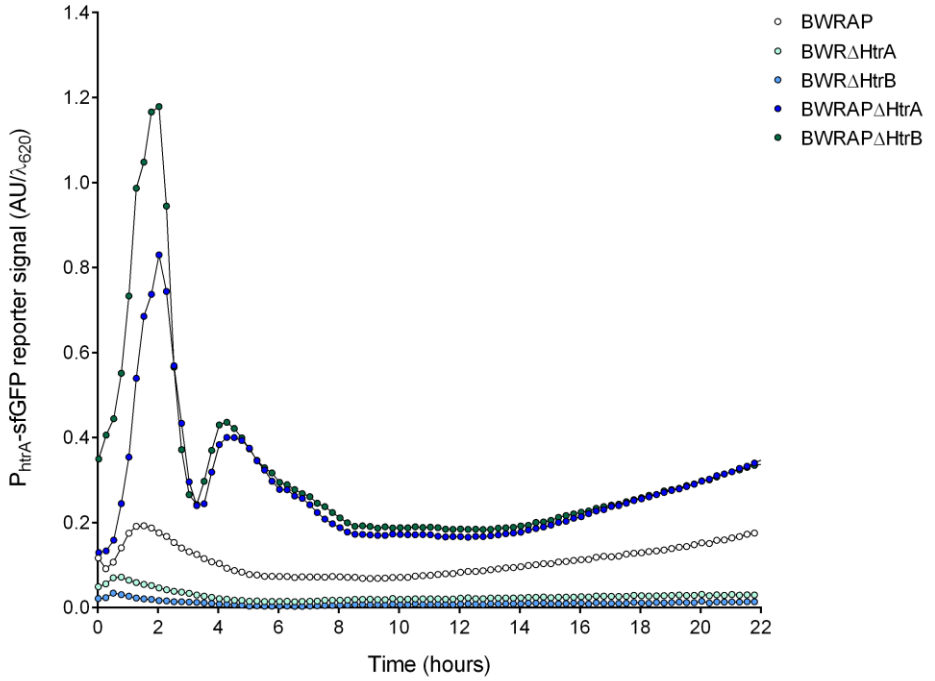


Figure 4.5 Protein secretion stress response upon overexpression of (A) XynA and (B) AmyM in strains lacking either *htrA* or *htrB* expression, measured by live detection of the reporter P$_{htrA}$-*sfGFP* fluorescence using the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 1.5 mL LB in FlowerPlate® 48-well MTPs and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The plotted fluorescence signal was blanked with the detected signal for the wild type strain 168 and normalised to the biomass. The average of six biological replicates per strain was plotted.

The final fluorescent signal divided by the biomass detected after 22 hours of growth of all strains studied in this chapter is plotted inFigure 4.6. By compiling the fluorescent secretion stress response of all studied strains with one or more deletions of genes encoding for extracytoplasmic proteases, it is clear that the secretion stress response of the heterologous protein, AmyM, is significantly higher than that of the native XynA (Figure 4.6). Moreover, Figure 4.6 demonstrates that a variety of stress levels are observed for the same secretory protein in strains with different proteolytic backgrounds. Although the HtrA and HtrB compensatory regulation was expected, the influence of extracellular proteases in the regulation of the quality control mechanisms was not anticipated.

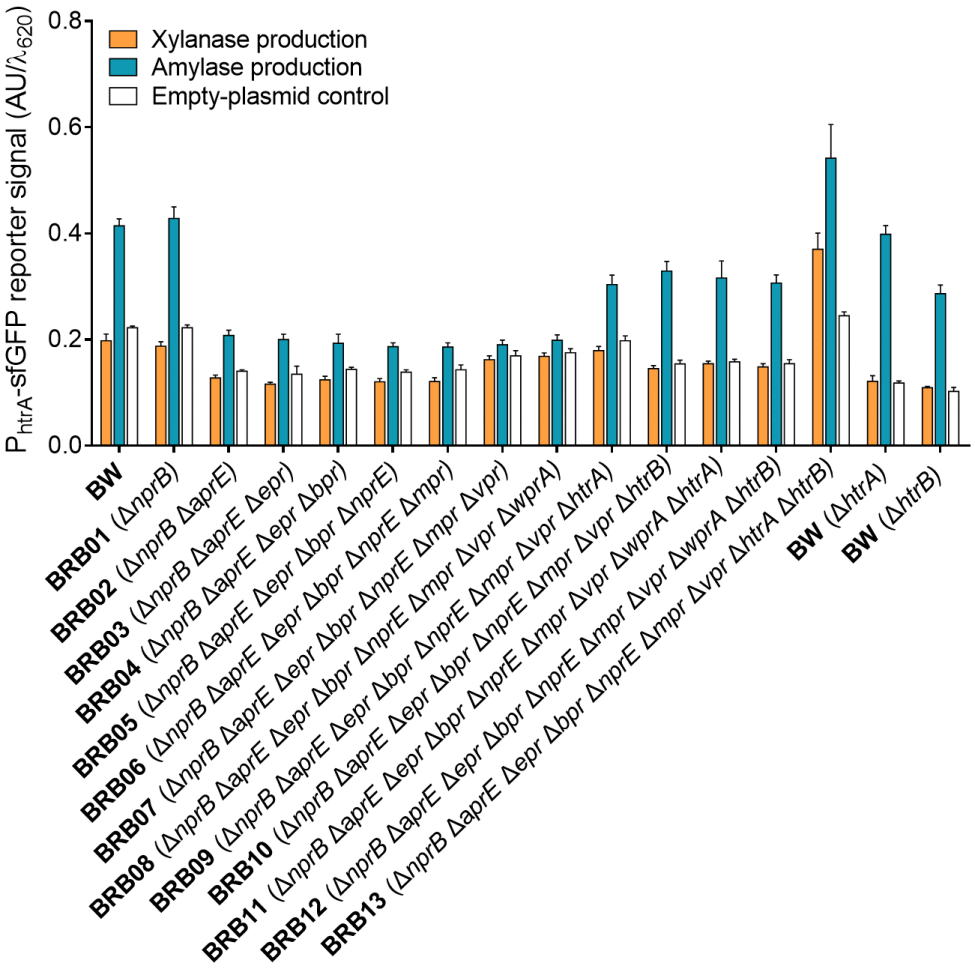


Figure 4.6 Protein secretion stress levels at 22 hours of enzyme production in strains lacking one or multiple extracytoplasmic proteases measured by the fluorescent reporter P_{htrA}-*sfgfp* using the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 1.5 mL LB in FlowerPlate® 48-well MTPs and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The fluorescence signal was blanked with the detected signal for the wild type strain 168 and normalised to the biomass. The average of at least six biological replicates per strain was plotted with the corresponding SEM.

4.5. The impact of protease mutations on enzyme production

The BRB strain collection offers a platform for strain optimisation, tackling the bottlenecks associated with the proteolysis of heterologous targets. Previous studies using the anthrax protective antigen as a model heterologous protein showed that the inactivation of proteases in these strains improved product stability in the culture medium of *B. subtilis* (Pohl *et al.* 2013). Previously, several other studies have used a similar strategy to improve the production of an heterologous target (Wu *et al.* 1991; Murashima *et al.* 2002; Pomerantsev *et al.* 2011; Westers *et al.* 2006b; Feng *et al.* 2015) .

Our results in the previous section suggest that there is a strong up-regulation of the CssRS system upon enzyme production, which triggers the expression of the quality control proteases HtrA and HtrB when multiple protease genes deletions are combined. This prompted us to investigate whether this response had a negative effect on the enzyme production levels.

The study in Section 4.4. , using the BioLector® microbioreactor system, was completed by quantifying enzyme activity after the 26 hours of growth. This was done using the commercial assays as described in Section 2.10. . The activity of each culture was calculated relative to the activity of the wild type 168 host (strains BWRXP and BWRAP for XynA and AmyM production, respectively). The results show that reducing the proteolytic activity of the cell does not significantly improve the production of the model enzymes XynA and AmyM (Figure 4.7). In general, the deletion of genes encoding for extracellular proteases (strains BRB01-BRB07) does not improve the production of XynA. This was to be expected since this is a native enzyme that has presumably evolved to be resistant to the proteases of its native host. Interestingly, there is a significant decrease in XynA production when the *wprA* or *htrA* deletions were added to the strains lacking extracellular proteases (strains BRB08, BRB09 and BRB11). The same effect was not observed when the *htrB* deletion was added to the strain lacking all seven extracellular proteases (BRB10) and, more interestingly, XynA production actually recovered when *htrB* was deleted in strain BRB08 (strain BRB12). These results suggest that XynA might be susceptible to the protease HtrB which, according to the literature, is up-regulated when *htrA* is deleted (strains BRB09 and BRB11) and has increased stability in the absence of *wprA* (strain BRB08 and BRB11). In contrast, when *htrB* is deleted in the wild type strain 168, there is a very significant decrease of XynA production (Figure 4.7-A). Presumably, the apparent susceptibility of XynA to

HtrB in the absence of extracellular proteases is related to the increased stability of the extracellular form of this protease, as shown previously by Krishnappa *et al.* (2014).

When the production of AmyM was analysed in strains with the same genotype, the outcome was different (Figure 4.7-B). The results show there is an increase in AmyM activity in the extracellular medium when various extracellular protease deletions are combined. In particularly, strain BRB04 (168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr*) and BRB06 (168 Δ*nprB* Δ*aprE* Δ*epr* Δ*bpr* Δ*nprE* Δ*mpr*) show a significant increase of AmyM production. Since Krishnappa *et al.* (2013) suggested that BpR is needed to process MpR it would be interesting to verify whether MpR targets AmyM.

Previous results have indicated that AmyM is secreted slowly in comparison to the native XynA. Together with the strong up-regulation of the *htrA* and *htrB* genes (Figure 3.7-B), these observations suggest that AmyM is subjected to strong cell-associated bottleneck(s) and degradation. Therefore, it could be expected that depleting the cell of the quality control proteases would facilitate the secretion of AmyM into the extracellular medium. Figure 4.7-B suggests that is the case when single deletions of either *htrA* or *htrB* are introduced in the wild type background, however, AmyM production is not increased with the deletions of the *htrA*, *htrB* and *wprA* in the absence of extracellular proteolytic activity (strain BRB13). In fact, there is a very significant negative effect on the AmyM production when both *htrA* and *htrB* are deleted. Pohl *et al.* (2013) have shown previously that the absence of *htrA*, *htrB* and *wprA* leads to a marked increase in *htrC* expression, the third HtrA-like protein encoded by *B. subtilis.* The gene *htrC* is a member of the WalR cell-wall stress regulon, which implies that the absence of both *htrA* and *htrB* leads to accumulation of misfolded secretory proteins at the membrane-wall interface which affects cell-wall biosynthesis and induces the WalR operon (Line Fabret & Hoch 1998). The significant decrease in AmyM production in strain BRB13 suggests either that AmyM is susceptible to HtrC proteolytic activity or is not efficiently secreted by cells under cell-wall stress, presumably because the secretory proteins accumulate in the membrane/cell wall interface. A similar but less pronounced effect was observed for XynA production (Figure 4.7-A, strain BRB13).

These results are in accordance with the analysis of stress response up-regulation. It was shown in Section 4.4. that preventing the production of extracellular and quality control proteases results in an up-regulation of the stress response mechanism in almost all strains producing the

model enzymes. Although the engineering of strains with reduced proteolysis is motivated by a potential increase in target stability, our results show that such extensive editing has unforeseen consequences for secretion stress induction and growth kinetics. This can result in reduced rather than increased productivity. This highlights the importance of validating each strain optimisation principle for the target.

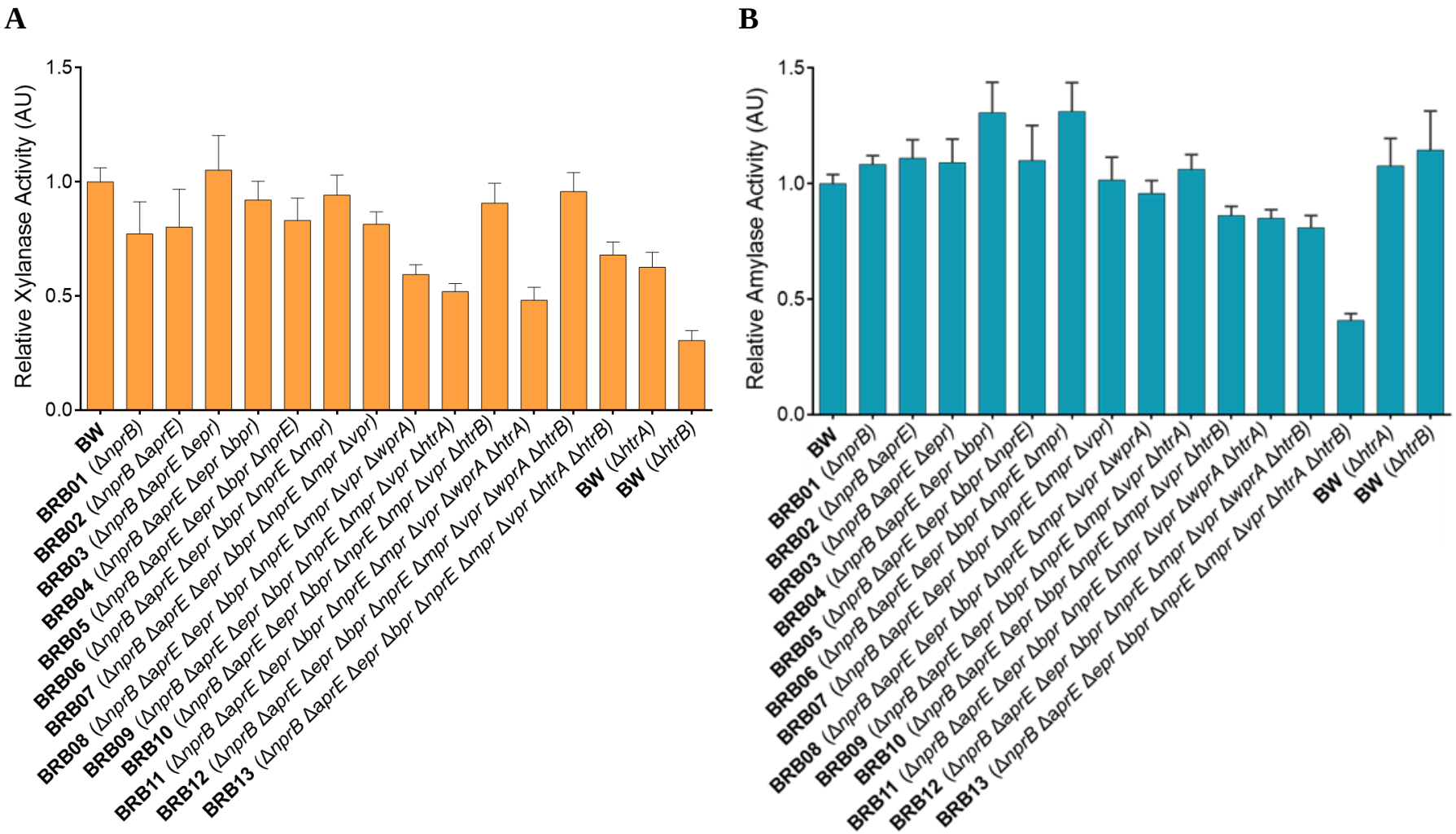


Figure 4.7 Relative enzyme active activity of strains overexpressing (A) XynA and (B) AmyM after 26 hours of growth in the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 1.5 mL LB in FlowerPlate® 48-well MTPs and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The enzymatic activity was measured using commercial assays according to what is described in Section 2.10. The average of nine biological replicates per strain was plotted with the corresponding SEM.

4.6. Detection of extracellular proteins with the P_{htrA}-sfGFP secretion stress reporter

The CssRS sensing mechanism has been used to engineer a biosensor, P_{htrA}-sfGFP, to detect the stress resulting from the overexpression of heterologous secretory proteins in this and other studies (H. Westers *et al.* 2004; Trip *et al.* 2011; Ploss *et al.* 2016). We have shown that this biosensor can detect different levels of secretion stress caused overproduction of the model secretory proteins, or by changes in the genetic background of the production strain (Figure 4.4). Our results also show that the level of GFP does not necessarily reflect the level of the target enzyme activity detected in the extracellular medium. For instance, strain BWRAP shows a significantly higher stress reporter signal in comparison to the wild type strain 168 (Figure 4.3-A) accompanied by a significantly higher level of AmyM production (Figure 3.5-B). In contrast, strain BRB13RAP shows an up-regulation of the secretion stress signal in comparison to strains BWRAP and BRB12RAP (Figure 4.4-B2), but a significantly lower production of AmyM (Figure 4.7-B). This reveals some limitations of the previously described use of this biosensor as a monitor for α-amylase production (H. Westers *et al.* 2004; Trip *et al.* 2011).

This prompted to investigate the sensitivity of the secretion stress reporter P_{htrA}-sfGFP to the presence of high levels of heterologous protein in the extracellular medium. In order to do this, we cultured strain BW (Table 4.1) in LB media containing different proportions of spent medium (SM) obtained following the the cultivation of the production strain BWAP (Table 4.1) for 24 hours. The fluorescent signal of the secretion stress reporter was monitored for 25 hours in the BioLector® following the growth regime described previously in section 4.3. The data in Figure 4.8 shows that the specific fluorescent signal shows a cumulative increase in the secretion stress response with increasing amounts of the SM in the culture media. At this stage we cannot be sure whether the increase in the signal is due to the depletion of nutrients or the presence of high levels of the heterologous protein in the culture medium.

These preliminary results suggest that the previously identified limitations of this reporter system could be overcome by a co-culture system where the secretion stress is monitored by a second strain, particularly when the stress response is likely to be influenced by heavy genetic editing of the production strain, rather than the level of production. Furthermore, these results contribute with a novel insight into the induction mechanism of the CssRS system, which is still unclear. The CssS sensor kinase has been suggested to react to accumulated misfolded proteins at the membrane/wall interface (Harwood & Cranenburgh 2008). Our current results indicate that this

may only reflect part of the story, as pre-synthesised and released proteins appear to be able to elicit a CssRS-mediated stress response. Further studies, using SM from the non-producing strain and the strain overproducing XynA are needed to confirm this interpretation.

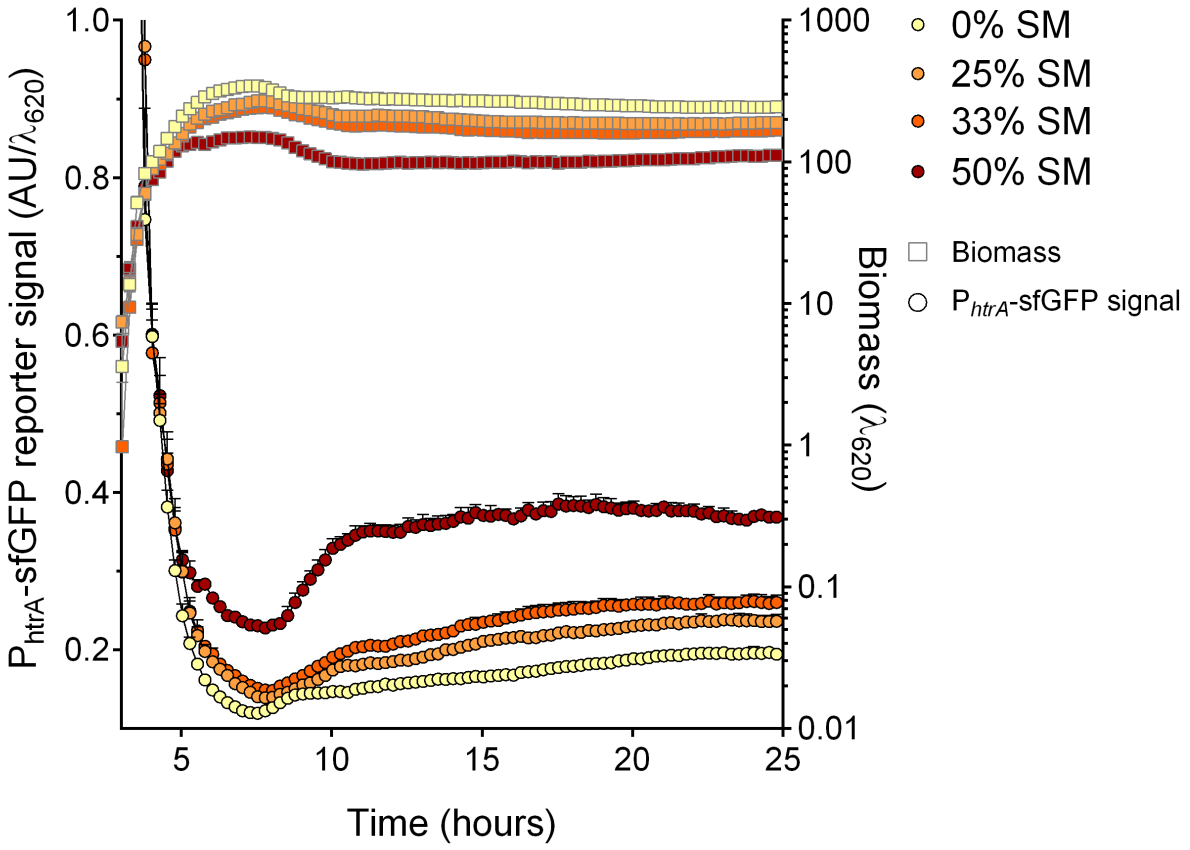


Figure 4.8 Detection of high levels of extracellular heterologous protein by the secretion stress reporter P_{htrA}-sfGFP. Strain BW was cultured in LB media containing 0, 25, 33 or 50% of spent medium (SM) obtained after growing strain BWAP for 24 hours in LB. The signal of the fluorescent reporter was measured using the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 1.5 mL LB supplemented with SM in FlowerPlate® 48-well MTPs and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The plotted fluorescence signal was normalised to the biomass. The average of three biological replicates per strain was plotted with the corresponding SEM.

4.7. The impact on model enzyme production of overexpressing quality control proteases

The up-regulation of quality control proteases is associated with a cellular response targeting misfolded proteins and there was a generally negative effect on enzyme production in protease-deficient mutants. However, we were not able to discount the possibility that increased levels of HtrA and HtrB are actually required for the production of high yields of XynA and AmyM. To test this hypothesis, strains overexpressing the serine proteases HtrA, HtrB and HtrC, as well as WprA, were analysed to determine the effect on XynA and AmyM production.

Strains were constructed which encoded a copy of one of the protease genes under the control of a xylose inducible promoter, inserted in the chromosome at the *amyE* locus (Table 4.1). The plasmids for XynA and AmyM overexpression were transformed in these strains and enzyme production levels were quantified in the culture medium at two time points. The growth regime was as described previously in Section 4.3. . Briefly, preliminary cultures prepared from overnight cultures were inoculated into pre-warmed LB medium in 48-well MTP FlowerPlates®. All strains were grown with and without the addition of xylose (0.2% final concentration). Growth was monitored for 20 hours at 37 °C with 95% humidity and vigorous agitation (800 rpm) in the BioLector® system. Samples were collected at 4 and 20 hours of growth and enzymatic activity was assayed as described in Section 2.10. . In a parallel experiment, strains BCS234, BCS291, BCS292 and BCS293 (Table 4.1) were grown in the same conditions to mid-exponential phase and cells were harvested for reverse transcription and RT-qPCR analysis to quantify the level of overexpression of each xylose-inducible construct (Section 2.7.).

Figure 4.9 shows the growth profiles of the analysed strains and the time points at which samples were collected for enzyme activity assays (Figure 4.9-A), as well as expression levels of the artificial protease expression cassettes (Figure 4.9-B).

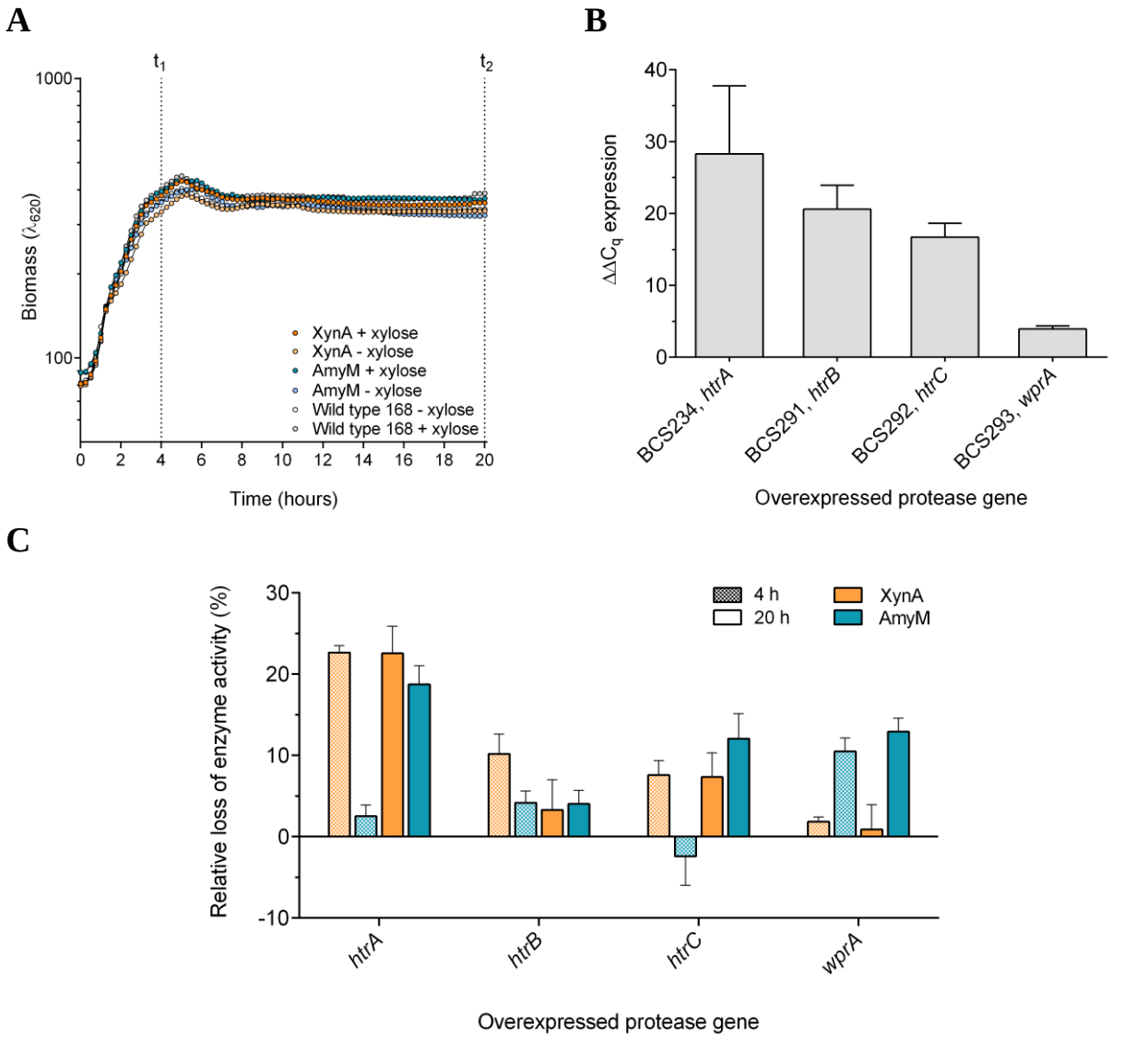


Figure 4.9 The effect of *htrA, htrB, htrC* and *wprA* overexpression on XynA and AmyM production.
A. Cultures of BCS237, BCS294-99 (Table 4.1) were grown in 1.5 mL LB in FlowerPlates at 37 °C with 95% humidity and vigorous agitation (800 rpm). Samples were taken at two time points; t_1 during transition phase and t_2 in late stationary phase. The enzyme activity in the culture medium was assayed as described in Section 2.10. . The growth curves were plotted as the averages of the strains producing either XynA or AmyM in the presence or absence of xylose induction.
B. Expression levels of *htrA*, *htrB*, *htrC* and *wprA* in the strains BCS234, BCS291, BCS292 and BCS293 in the presence of 0.2% xylose, relative to the expression levels in the absence of inducer. The C_t values were averaged from two technical replicates, and the ΔC_t values were calculated using the *fbaA* gene as control. The average of the ΔC_t values of two biological replicates are shown with the corresponding SEM. A summary of the results in a table format can be found in Appendix G (Table G.5).
C. Relative loss of enzyme activity upon overexpression of *htrA*, *htrB*, *htrC* and *wprA*. The relative loss of enzyme activity was calculated as the difference between the enzyme activities measured in the induced and non-induced cultures, relative to the positive control for XynA and AmyM production – strains BWXP and BWAP, respectively.

The expression of the *htrA, htrB, htrC* and *wprA* genes introduced by the addition of 0.2% xylose was calculated with respect to their expression in the absence of inducer (Figure 4.9-B). Although each of the genes was placed under the control of the same inducible promoter and ribosome binding site, and the expression cassettes were integrated into the same chromosome locus (*amyE*), there were significant variations in the relative expression levels of target genes. Nevertheless, it was confirmed that all artificial expression cassettes resulted in a significant overexpression of the protease genes corresponding to 28.3±18.9, 20.6±6.6, 16.7±3.9 and 3.9±0.8 fold more gene transcripts of *htrA*, *htrB*, *htrC* and *wprA*, respectively.

The effect of protease overexpression on the levels of model enzyme activity in the culture medium is shown in Figure 4.9-C as relative enzyme activity. Relative enzyme activity was calculated as the difference between the enzyme activities measured in the induced and non-induced cultures, relative to the positive controls for XynA (strain BWXP) and AmyM (strain BWAP). The results show a general decrease in enzyme production when *htrA*, *htrB*, *htrC* and *wprA* are overexpressed. The loss of enzyme activity is most significant when *htrA* is overexpressed. However, this might be due to higher levels of *htrA* overexpression measured for strain BCS234 in comparison to the other protease overexpressing strains (Figure 4.9-B). Interestingly, *wprA* overexpression has significantly more effect on Amylase production than on XynA production. This is likely to be because XynA has evolved to be resistant to the protease activities of its native host. Nevertheless, the overproduction of *htrA htrB* and *htrC* causes significant loss of XynA, presumably because it remains susceptible to these proteases during the time it takes to fold following its release from the translocase. However, the increased susceptibility to the HtrA-like proteases, compared with that of WprA, could in part be due to their higher level of induction (Figure 4.9-B).

The general decrease in the activities of the model enzymes in response to the induction of the HtrA, HtrB, HtrC and WprA proteases suggests that the up-regulation of the secretion stress response is a consequence of, rather than a requirement for, high-level of enzyme production. Furthermore, these results help to understand the lack of significant optimisation of enzyme production in the protease deficient strains (Figure 4.7).

4.8. The impact of protease mutations on the kinetics of enzyme secretion

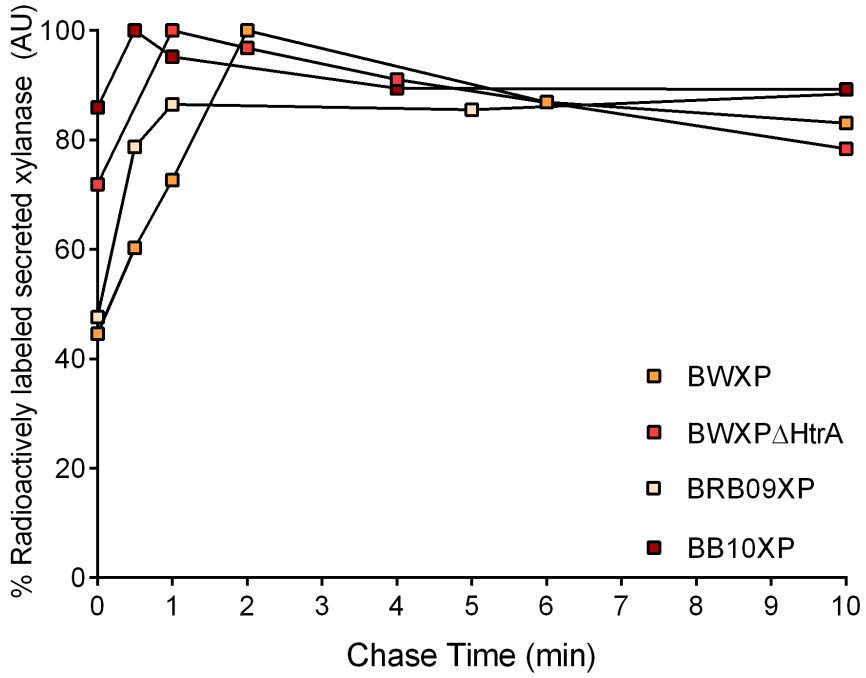
In Section 3.3.7 , pulse-chase experiments indicated that XynA is rapidly processed and secreted by the cell, suffering some cell-associated degradation during the late stages of secretion, while

AmyM is processed and released slowly into the growth medium, suggesting extensive cell-associated bottlenecks and degradation. It was therefore of interest to determine whether the protease deficient strains studied in this chapter affected the secretion kinetics of the model enzymes.

Pulse-chase experiments were carried in selected protease deficient strains expressing high levels of XynA and AmyM (Section 3.3.7) and the data is shown in Figure 4.10. In the case of XynA, strains lacking just HtrA (BWXPΔHtrA), all of the feeding proteases plus HtrA (BRB09XP) and all of the feeding proteases plus HtrB (BRB10XP) were analysed in comparison with the wild type strain (BWXP). The data (Figure 4.10A) indicate that deletion of the quality control proteases HtrA or HtrB increases the rate at which XynA accumulates in the culture medium compared with the wild type. In addition, there was less degradation of XynA in the culture medium in strains BRB09XP and BRB10XP, which lack all seven extracellular proteases. Significantly, however, the production of active XynA was reduced (Figure 4.7A). Taken together, these data indicate that although the deletion of the quality control proteases HtrA and HtrB might contribute to the alleviation of secretion bottlenecks, it does not result in increased production of active enzyme into the culture medium. This highlights the import role of the quality control mechanisms in protecting cell growth by degrading misfolded proteins in membrane-wall interface.

In the case of AmyM (Figure 4.10-B), the absence of the quality control proteases in strains BRB09AP and BRB10AP does not result in a significant change in the secretion kinetics. Similarly to XynA production, these protease deficient strains did not show increased production of the heterologous AmyM. It is not clear why the kinetics of AmyM secretion is so relatively slow, although the kinetics of release may imply possible interactions with the cell wall during its passage to the culture medium.

A



B

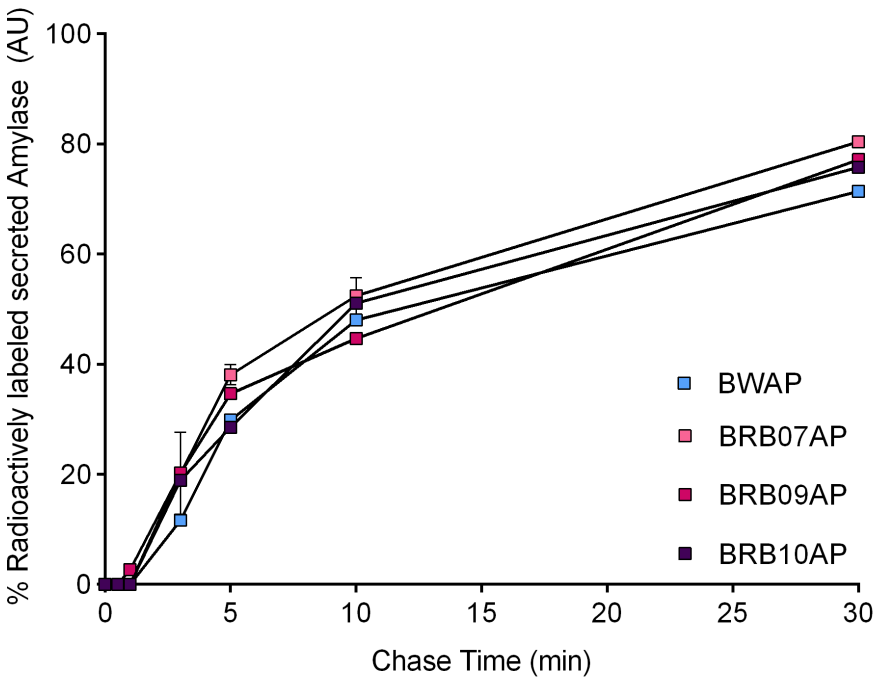


Figure 4.10 Pulse-chase analysis of (A) XynA and (B) AmyM production and secretion in the protease deficient strains BWXPΔHtrA, BRB09XP, BRB10XP, BRB07AP, BRB09AP and BRB10AP. Pulse-chase analysis of strains BWXP and BWAP are shown as controls for corresponding XynA and AmyM production in (A) and (B), respectively. The cultures were grown, harvested and processed according to Section 2.9. The quantified radioactively labelled proteins were plotted relative to the maximum protein levels detected in the culture medium. In B, the average of two technical replicates of the protein quantification step are plotted with the corresponding SEM.

Conclusions

The work in this chapter represents a comprehensive study of the effects of extra-cytoplasmic proteases on the production of two industrially relevant model enzymes, one native to *B. subtilis*, the other a heterologous enzyme. In addition, we also monitored the impact of the production of these enzymes on the secretion stress response. To this end we used a protein secretion stress biosensor based on the *htrA* promoter, as described by Trip *et al.* (2011), to continuously monitor the induction of the secretion stress response upon hyper-production of the model enzymes (Figure 4.1). The BioLector® bench top microfermentation system (m2p-labs) was used to monitor biomass and fluorescence under controlled aeration, temperature and humidity.

The high levels of production of the native enzyme XynA do not cause a detectable induction of the secretion stress response, as monitored via the P_{htrA}-sf*gfp* gene fusion (Figure 4.3-A), despite the higher levels of *htrA* transcript in this strain (Figure 3.10-B). This might in part reflect limitations in the sensitivity of the fluorescent reporter system. On the other hand, AmyM hyper production resulted in a significant increase in the P_{htrA}-*gfp* fluorescent signal throughout the growth of strain BWRAP; expression increases during exponential phase, peaks at transition phase, shows a temporary decline in early stationary phase before steadily increases throughout late stationary phase (Figure 4.3-A). Replacing the native signal peptide of AmyM with that of XynA eliminates this secretion stress response (Figure 4.3-B) which is in agreement with the reduced AmyM production levels of strain BWAP2 (Figure 3.7). This reduction in productivity is likely to be related to early stages of secretion, possibly an intracellular stress response, in contrast to the secretion stress response observed in strain BWRAP (Figure 4.3-A).

Our study of the secretion stress response in protease deficient mutants revealed that the deletion of the extracellular protease gene *aprE* causes a major change in the stationary phase growth profile (Figure J.1, Appendix J). Strain BRB02 (Δ*nprB* Δ*aprE)* and strains with additional deletions of protease genes reach higher optical densities which lead to significantly lower secretion stress signals upon AmyM production in comparison to strain BWRAP and BRB01RAP (Figure 4.4-B1). Interestingly, this reduction of stress levels did not translate into a decrease of AmyM production (Figure 4.7-B). This reveals some limitations of the previously described use of this biosensor as a monitor for α-amylase production (H. Westers *et al.* 2004; Trip *et al.* 2011). Furthermore, these results suggest there is a correlation between the levels of

extracellular proteolytic activity and cell fitness, which impacts on the regulation of the secretion stress response without, in the case of AmyM at least, compromising enzyme production levels.

The deletion of either *htrA* or *htrB* in the presence or absence of all seven extracellular proteases caused an up-regulation of the secretion stress response (Figure 4.4-A2, Figure 4.4-B2 and Figure 4.6). It has been shown previously that inactivating either *htrA* or *htrB,* results in a compensatory increase in the expression of the other (Noone *et al.* 2000; Noone *et al.* 2001), and our data indicates that this is enhanced in the presence of the heterologous enzyme, AmyM (Figure 4.4-B2 and Figure 4.7-B).

Next we investigated the relationship between the secretion stress response and high-levels of enzyme production. We showed that overexpression of the four quality control proteases, HtrA, HtrB, HtrC and WprA, resulted in a reduced level of both XynA and AmyM enzyme production (Figure 4.9). This observation is in good accord with the lower levels of production in strains with deletions of either *htrA* or *htrB,* which is known to result in a compensatory up-regulation of the remaining protease (Figure 4.7) (Noone *et al.* 2001; Pohl *et al.* 2013).

When both *htrA* and *htrB* were deleted in the absence of extracellular proteases, the highest secretion stress levels were detected for both native and heterologous enzyme productions. The resulting reduction of enzyme production in both cases provides further evidence that the fluorescent reporter does not provide a sufficiently quantitative measurement of HtrA-like enzyme production when different hosts are compared. This may reflect differences in the regulation of the *htrA* and *htrB* promoters.

Interestingly, pulse-chase analysis suggested that the deletion of HtrA or HtrB influences the secretion kinetics of XynA, leading to a faster release of this secretory protein. Since these strains did not show an increase in enzyme activity in the culture medium, the importance of the quality control mediated by HtrA and HtrB is highlighted. The HtrA protease has been suggested to have chaperone activity, as in *E. coli*, and its transitory interaction with secretory proteins as they emerge from the translocase could account for these results (Spiess *et al.* 1999; Antelmann *et al.* 2003).

The observations in this chapter have important implications for the use of protease-deficient strains as hosts for protein production and emphasise how production strains must be tailored to the product of interest. Hyper production of the native and heterologous model enzymes

produced very different results with respect to the effect of the various protease gene deletions on the secretion stress response and enzyme production levels. However, in both cases it was evident that heavy editing of the proteolytic capacity of the cell does not, as might have been anticipated, translate into optimal enzyme production. Although proteases constitute a relevant bottleneck for enzyme production, the quality control systems have evolved to maintain cell fitness by maintaining important cellular functions and viability. The other main observation is that, in general, proteins that have co-evolved with a particular strain impose less stress, even when over produced, than do heterologous proteins (Jensen *et al.* 2000).

Chapter 5 Engineering the translocase machine

5.1. Introduction

The widespread use of *Bacillus subtilis* and some of its close relatives for the industrial-scale production of enzymes is based almost exclusively on their potential to secrete proteins at grams per litre concentrations. However, the range of proteins that is capable of this high-level production and secretion is currently limited (Pohl & Harwood 2010). One important challenge is that improvements targeting protein synthesis are not fully reflected in corresponding increase in product yield. It is widely accepted that several factors have the potential to limit production yield. One very obvious limiting factor is the capacity of the secretory translocons, responsible for the translocation of secretory proteins across the cytoplasmic membrane. The saturation of the Sec translocon capacity, caused by high-level gene expression in industrial production strains, has the potential to significant to be a major secretion bottleneck.

B. subtilis secretes a large number of proteins, mainly through the major and ubiquitous Sec preprotein secretory pathway. Even though other secretion pathways have been identified, such as the TAT pathway, the Sec-dependent pathway is the most exploited for the secretion of heterologous proteins from *B. subtilis* and its relatives.

Secretory proteins are directed to membrane-bound Sec translocases by a N-terminal signal peptide. Cytosolic chaperones interact with the secretion targets keeping them in a secretion competent, essentially unfolded, state. After reaching the membrane, the preprotein is translocated via a confined aqueous channel composed of a set of integral membrane proteins (van Wely *et al.* 2001). In *B. subtilis*, the components of the translocase complex include the SecY, SecE, and SecG proteins, which form the SecYEG complex. This is the core of the translocation channel and these proteins are associated in the membrane with the SecDF protein. SecDF is thought to contribute to the efficiency of protein secretion by improving SecA cycling and maintaining the forward momentum of the preprotein (Driessen & Nouwen 2008). SecA is an ATPase located on the *cis* side of the membrane responsible for both targeting and the coupling of the energy required for translocation – it is the motor component of the translocation.

The expression of the Sec translocase machinery proteins, as most proteins in the cell, is related to growth rate. The gene encoding SecY, the main pore-forming component of the translocase, is located in a ribosomal operon that is down-regulated during transition from exponential phase to stationary phase (Suh *et al.* 1996; Yang *et al.* 2013). In the case of SecA synthesis, there are contradictary reports in the literature. Initially, it was suggested that *secA* expression occurs mainly in the exponential growth phase, reaching a maximal value almost precisely at the transition from exponential to the stationary phase (Herbort *et al.* 1999; Leloup *et al.* 1999; Blom *et al.* 2011). More recently, Yang *et al.* (2013) found that, in the presence of glucose, SecA expression increases from exponential to stationary phase while the expression of SecY and SecDF decreases, resulting in at least a tenfold differential reduction of the SecA/SecY and SecA/SecDF ratios in exponential phase. There are conflicting requirements during exponential and stationary phase. During exponential phase the positive nutrient status of the medium means that there is an increased need for housekeeping proteins, such as cell wall synthesis proteins, but a reduced need to synthesize and secrete macromolecular hydrolases to utilize less readily available nutrient sources. During stationary phase the opposite is the case and the secretion of proteins into the extracellular medium occurs mainly in this part of the growth cycle (Yang *et al.* 2011). In the absence of a metabolic flux balance analysis it is not clear why major components of the Sec-secretion machinery are expressed differently during each growth phase.

It has been speculated that the commercial scale secretion capacity of *B. subtilis* is limited by the production of translocase components. It has been reported that the secretion of specific proteins have differential requirements for SecA (Leloup *et al.* 1999). A number of studies have addressed the impact on protein production of increasing the expression of or modifying the translocase components. Kakeshita *et al.* (2010) showed that deleting the 61 amino acids at the C-terminal region of SecA led to significant enhanced production of heterologous proteins in *B. subtilis.* Diao *et al.* (2012) constructed an artificial posttranslational protein targeting pathway in *B. subtilis* by co-expressing SecB from *E. coli* and a *B. subtilis* hybrid SecA, where the C-terminal 32 amino acids were replaced by the equivalent residues from *E. coli* SecA, involved in SecB binding. With this strategy, the authors showed a significant improvement of maltose binding protein (MalE11) and alkaline phosphatase (PhoA) secretion. Furthermore, with the aim of increasing the number of translocons, Mulder *et al.* (2013) constructed an artificial SecYEG operon fused to a strong inducible promoter and demonstrated that this increased the yield of secreted α-amylase. Chen *et al.* (2015) took a combinatorial approach to systematically

overexpress 23 genes or operons encoding proteins involved in or closely related to the Sec pathway in *B. subtilis*, including the translocase genes. In contrast to the work of Mulder *et al.* (2013), Chen *et al.* (2015) found that overexpressing the translocase genes did not significantly increase of the yields of two heterologous proteins, AmyL from *B. licheniformis* and AmyS from *G. stearothermophilus*.

Despite these studies having met mixed success, it is essential to ensure that sufficient translocons are present to sustain large-scale and efficient production of target enzymes. This chapter explores a synthetic biology approach to engineering synthetic translocon operons using codon pair optimised genes. We then analysed the influence of overexpressing these synthetic operons on the production of two industrially relevant model enzymes, the native XynA and the heterologous AmyM.

5.2. Strains and plasmids

The strains and plasmids used in this chapter are listed in Table 5.1 and Table 5.2, respectively. DNA manipulation and transformation methods are described in Section 2.5. and Section 2.6. , respectively. The primers used are shown in Appendix C and the plasmid maps from Table 5.2 are shown in Appendix B.

B. subtilis strain 168 was used as the host for all the strains engineered in this study. The construction of strain BWAP was described in Section 3.2. This strain produces high levels of the heterologous model enzyme AmyM from *G. stereathermophilus.* Strain BWXC was constructed by Sauer (2016). This strain carries an extra copy of the native model enzyme *xynA* in the *aprE* locus of *B. subtilis* chromosome, expressed under the control of the P_{E4} strong promoter from the bacteriophage SPO1 (Stewart *et al.* 1998).

In order to create strains expressing a synthetic translocase machinery, two inducible expression systems were designed for the expression of *secA* and a polycistronic operon expressing *secY, secE* and *secG*. Variants of these expression constructs were designed with combinations of native or synthetic RBS with either the native or codon pair optimised versions of the genes. The design of the combinatorial BioBricks synthesised by SGI-DNA are described in Section 1.3. Strains BRC31P2A, BRC31P2B, BRC31P2C, BRC31P1A and BRC31P1D were created via the StarGate® type II cloning system using, for each construct, two synthetic BioBricks as inserts, the

promoter and coding sequence, and the integrative pRC31 vector (see later). The strains were confirmed by PCR amplification of the *amyE* integration locus and sequencing.

Strains BRC31P2A, BRC31P2B, BRC31P2C, BRC31P1A and BRC31P1D were transformed with either plasmid pCS73 or the genomic DNA of strain BWXC to create strains expressing the synthetic translocase and either a native (XynA) or heterologous (AmyM) model enzyme. The suffixes X and A were added to the strain nomenclature when expressing XynA and AmyM, respectively.

To create the pRC31 integrative vector, plasmid pRC30 was obtained from the DSM plasmid collection and used as template for creating the vector backbone. Plasmid pRC30 is an integrative plasmid targeted to the *amyE* locus of *B. subtilis* 168. Inside the integration cassette there is a chloramphenicol resistance cassette which is exchanged with the inserts introduced via the StarGate® cloning system (IBA). Outside the integration cassette is an erythromycin resistance gene for counter-selection of Campbell-like integration events. In plasmid pRC31, the erythromycin resistance gene was replaced by an expression cassette encoding the red fluorescent mCherry protein under the control of the strong P_{E4} promoter from the bacteriophage SPO1 (Stewart *et al.* 1998). The presence of red coloration allows colonies which have undergone a Campbell-like integration event to be distinguished and discarded, significantly reducing the screening time need to identify the correct clones. The backbone of plasmid pRC31 was amplified using primers 91 and 92. The P_{E4} promoter region was obtained as a gBlock from Integrated DNA Technologies (IDT) and primers 93 and 94 were used to amplify the DNA. Primers 95 and 96 were used to amplify the *mcherry* gene using as template the chromosomal DNA of an in-house DSM strain. These three fragments (plasmid backbone, P_{E4} and *mcherry*) were assembled via Gibson Assembly and transformed to chemically competent *E. coli* NEB10 cells. The correct clones were identified by the red coloration in the colony and liquid culture and confirmed via sequencing and endonuclease digest.

Table 5.1 Summary of the strains used in the study of this chapter.

Strain	Genotype	Source
B. subtilis		
168	*trpC2*	Kunst *et al.* (1997)
BRC31P2A	168 *amyE*::P_{xyl}-*secY-secE-secG*, *spec*	This work
BRC31P2B	168 *amyE*::P_{xyl}-$secY_{CpO}$-$secE_{CpO}$-$secG_{CpO}$, *spec*	This work
BRC31P2C	168 *amyE*::P_{xyl}-*secY-secE-secG*, *spec*	This work
BRC31P1A	168 *amyE*::P_{spac}-*secA*, *spec*	This work
BRC31P1D	168 *amyE*::P_{spac}-$secA_{CpO}$, *spec*	This work
BRCA31P2A	168 *amyE*::P_{xyl}-*secY-secE-secG*, *spec;* pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRCA31P2B	168 *amyE*::P_{xyl}-$secY_{CpO}$-$secE_{CpO}$-$secG_{CpO}$, *spec;* pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRCA31P2C	168 *amyE*::P_{xyl}-*secY-secE-secG*, *spec;* pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRCA31P1A	168 *amyE*::P_{spac}-*secA*, *spec;* pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BRCX31P2A	168 *amyE*::P_{xyl}-*secY-secE-secG*, *spec; nprE*::P_{E4}-*xynA*, *ery*	This work
BRCX31P2B	168 *amyE*::P_{xyl}-$secY_{CpO}$-$secE_{CpO}$-$secG_{CpO}$, *spec; nprE*::P_{E4}-*xynA*, *ery*	This work
BRCX31P2C	168 *amyE*::P_{xyl}-*secY-secE-secG*, *spec; nprE*::P_{E4}-*xynA*, *ery*	This work
BRCX31P1A	168 *amyE*::P_{spac}-*secA*, *spec; nprE*::P_{E4}-*xynA*, *ery*	This work
BRCX31P1D	168 *amyE*::P_{spac}-$secA_{CpO}$, *spec; nprE*::P_{E4}-*xynA*, *ery*	This work
BWAP	168 pCS73 (P_{amyQ}-*amyM*), *neo*	This work
BWXC	168 *nprE*::P_{E4}-*xynA*, *ery*	Sauer (2016)

neo – neomycin resistance; *spec*- spectinomycin resistance; *em* – erythromycin resistance; *CpO* – codon pair optimised

Table 5.2 Summary of the plasmids used in the study of this chapter.

Plasmid	Properties	Source
pCS73	P_{amyQ}-*amyM*, *reppUB, neo, bleo*	DSM
pRC30	*amyE*::(*cat*, *spec*); *em*; *bla*	DSM
pRC31	*amyE*::(*cat*, *spec*); P_{E4}-*mCherry*; *bla*	This work
pAPNC213	*aprE::(*P_{spac}, *spec*); *bla*	(Morimoto *et al.* 2002)
pAX01	*lacA*::(P_{xyl}, *em*); *bla*	(Härtl et al., 2001)

neo – neomycin resistance gene; *bleo* – bleomycin resistance gene; *bla* – ampicillin resistance; *cat* – chloramphenicol resistance; *em* – erythromycin resistance

5.3. Design of synthetic inducible expression constructs of the Sec translocase components

In order to create strains expressing synthetic operons encoding the Sec translocase, two inducible expression systems were designed, one for the expression of *secA* and the other for the expression of the polycistronic operon expressing *secY, secE* and *secG*. Variants of these expression constructs were designed with combinations of native or synthetic RBSs with either the native or codon pair optimised versions of the genes. Table 5.3 summarises the combinatorial nature of synthetic constructs expressed in these strains.

Two inducible promoters were selected to create two BioBricks that can be combined with coding sequences via the StarGate® type II restriction enzyme cloning system (Section 2.5.11). The sequences of the IPTG inducible promoter P_{spac} and the corresponding *lacI* repressor were obtained from the plasmid pAPNC213 (Morimoto *et al.* 2002). The sequences of the xylose inducible P_{xyl} and *xylR* repressor were obtained from the plasmid pAX01 (Härtl *et al.* 2001). The recognition sequence of the type II restriction enzyme *Bsm*BI was added at the 5' and 3' ends of these sequences to generate compatible overhangs with the pRC31 integrative vector (5' end), and with the BioBricks with the coding sequences of the translocase components (3' end). The IPTG and xylose inducible promoters BioBricks were named P1 and P2, respectively (Figure 5.1). The sequences of these BioBricks are available in Appendix H (Table H.3).

Table 5.3 Summary of the strains carrying synthetic operons encoding components of the Sec translocase with combinations of native or synthetic RBS and with either the native or codon pair optimised versions of the genes.

Promoter	RBS	Genes	Strain
P_{spac}	Native	*secA*	BRC31P1A
		$secA_{CpO}$	BRC31P1B*
	Synthetic	*secA*	BRC31P1C*
		$secA_{CpO}$	BRC31P1D
P_{xyl}	Native	*secY-secE-secG*	BRC31P2A
		$secY_{CpO}$-$secE_{CpO}$-$secG_{CpO}$	BRC31P2B
	Synthetic	*secY-secE-secG*	BRC31P2C
		$secY_{CpO}$-$secE_{CpO}$-$secG_{CpO}$	BRC31P2D*

* It was not possible to isolate versions of strains that expressed the model enzymes as well as these synthetic operons and therefore these strains were not used in this study. CpO, codon pair optimised.

The synthetic RBSs were calculated using the online algorithm "RBS calculator" (Salis *et al.* 2010; Borujeni *et al.* 2014). This algorithm predicts the translation initiation rate of a protein coding sequence in bacteria, and designs synthetic ribosome binding site sequences to rationally control the protein expression level, combining a thermodynamic model of ribosome and mRNA interactions with a sequence optimization algorithm (Salis *et al.* 2010). The synthetic RBS for the SecA and SecYEG expression constructs were obtained after inputting the 25 bp region upstream the RBS and the protein coding sequence of each gene variant (Table H.2, Appendix H). The codon pair optimised versions of the *secA, secY, secE* and *secG* genes were obtained using in-house expertise within the bioinformatics department of DSM. A list of the codon pair optimised sequences is in Appenidx H (Table H.1). The sequences of the native and codon optimised versions of the *secA* and *secY, secE* and *secG* genes were combined with either the sequence of the synthetic or native RBS, creating 8 different BioBricks (Figure 5.1). The recognition sequences for the type II restriction enzyme *Bsm*BI were added at the 5' and 3' ends of these BioBricks to allow generating compatible overhangs with the promoter BioBrick (5' end) and with the integrative promoter pRC31 (3' end).

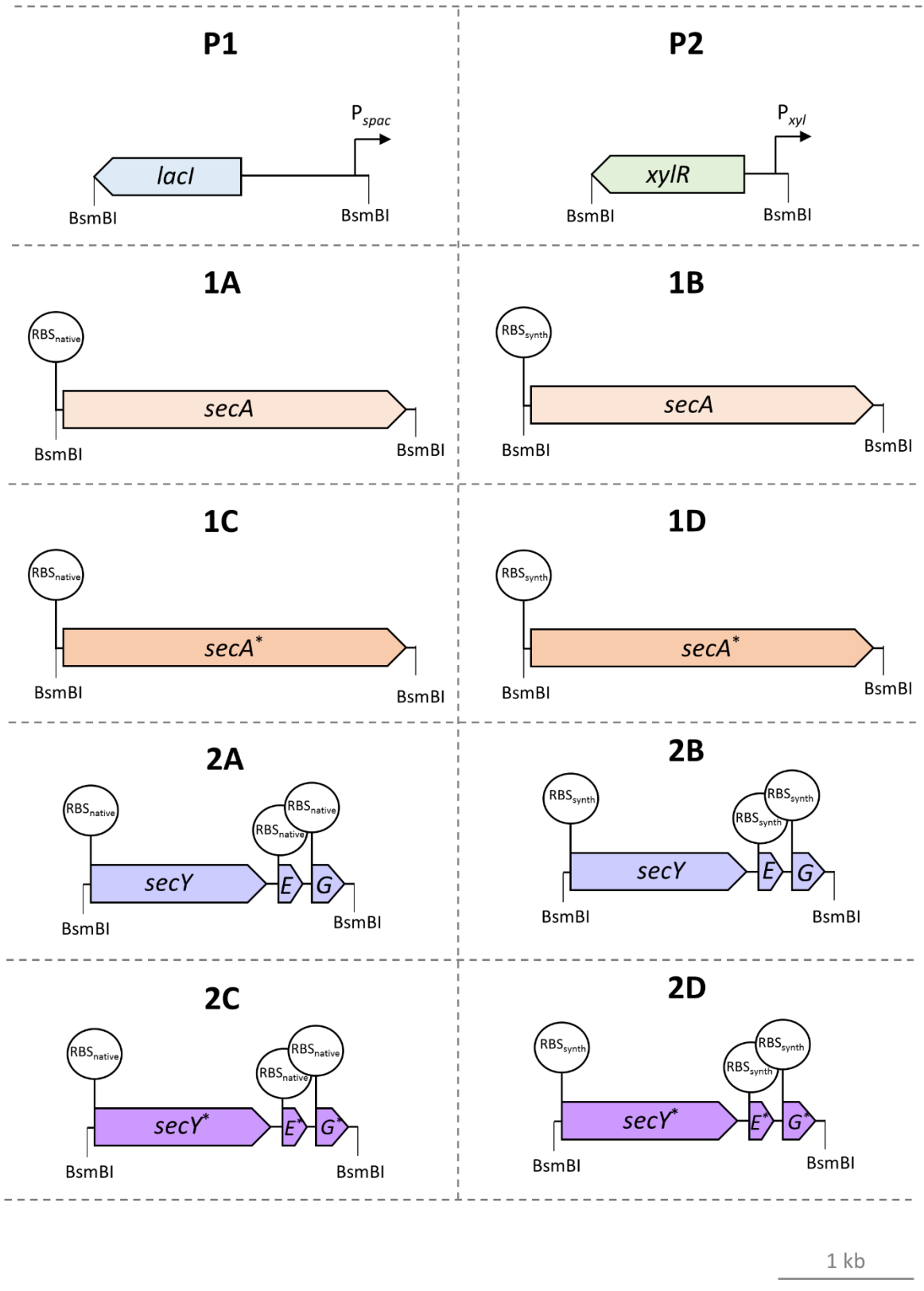


Figure 5.1 Schematic representation of the BioBricks designed to express a synthetic translocase machinery. The DNA sequence of each BioBrick is available in Appendix H (Table H.3).
* Codon pair optimised genes.

The expression of the synthetic constructs integrated into the *amyE* locus of the *B. subtilis* 168 wild type strain was confirmed via qPCR. To quantify the expression of the synthetic constructs, primers specific to the constructs were designed and the calculated expression levels are shown relative to the expression in the absence of inducer. In the case of the polycistronic operons expressing SecYEG, the primers used target the third and final gene of the operon, *secG*. All analysed constructs showed an up-regulation of gene expression in the presence of inducer (Figure 5.2). This up-regulation was more significant for the xylose induced SecYEG operons (Figure 5.2-B). However, the P_{spac} promoter is known to show a basal levels of activity in the absence of IPTG that affects the induction ratio, and it is expected that the absolute expression levels of the SecA synthetic constructs are higher than in Figure 5.2-A in comparison to the wild type strain.

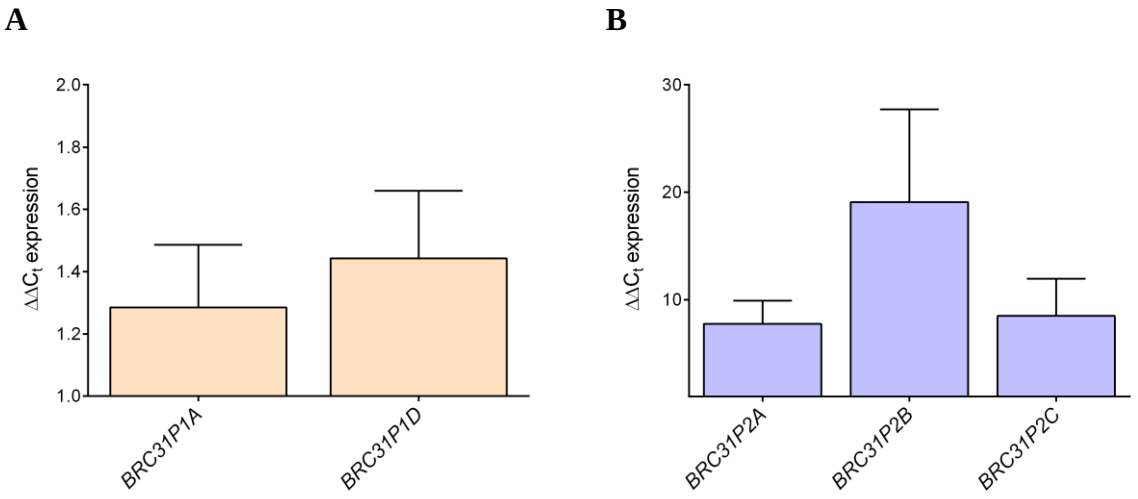


Figure 5.2 Confirmation of gene expression from the synthetic constructs expressing components of the Sec translocase. Cells cultures were grown in 1.5 mL LB in FlowerPlates at 37 °C in a Biolector microfermentor with 95% humidity and vigorous agitation (800 rpm). The cultures were harvested at the end of exponential phase and total RNA extracted (Section 2.7.). The expression levels were quantified in the presence of (A) 1 mM IPTG and (B) 0.3% xylose, relative to the expression levels in the absence of inducer. The C_t values were averaged from two technical replicates, and the ΔC_t values were calculated using the *cheA* gene as control. The average of the ΔC_t values of two biological replicates are shown with the corresponding SEM. A summary of the results in a table format can be found in Appendix G (Table G.6.).

5.4. The effect of expressing synthetic translocase components on the production of industrially relevant model enzymes

Chapter 3 shows that the industrially relevant enzymes, the endo-1,4-β-xylanase XynA from *B. subtilis*, and the maltogenic α-amylase AmyM from *G. stearothermophilus*, have very different production profiles and secretion kinetics, even when expressed under the same expression

system and growth conditions. This is most likely due to their different interactions with, and processing by, the protein secretion pathway. In Chapter 4 we showed that both enzymes impose different levels of stress when overproduced in strains with different levels of proteolytic activity, even though the proteolytic activity of the host did not affect significantly the production of these model enzymes. Here we investigate whether the overexpression of the translocase motor (SecA) and pore forming proteins of the Sec translocase (SecYEG) influence the levels of native and heterologous enzyme production. In fact, overexpression of one or more secretory machinery components has been shown previously to assist the secretion and folding of several proteins (reviewed in Section 1.4.). Our approach aims to enhance these previous studies by a comparative analysis on the production of a native and heterologous protein at different growth phases, using a combinatorial synthetic biology approach.

The BioLector® bench top microfermentation system (m2p-labs) was used to monitor the growth of *B. subtilis* strains producing XynA and AmyM, while expressing components of the Sec-dependent translocase from synthetic constructs integrated in the chromosome of *B. subtilis* (Table 5.1). Biological replicates were grown in triplicate in 48-well MTP FlowerPlates® in a randomised layout. Overnight cultures were diluted into fresh LB medium with and without the addition of xylose or IPTG to give synchronize precultures which were used to inoculate pre-warmed LB medium in 48-well MTP FlowerPlates®, with or without the inducer. Biomass (excitation: 620 nm, gain: 20) was monitored for 50 hours at 37 °C with 95% humidity and vigorous agitation (800 rpm). Samples were collected at the exponential, transition, stationary and late stationary stages of the growth cycle to quantify enzyme activity using commercial assays (Section 2.10.).

The continuous monitoring of biomass revealed that the expression burden of the synthetic constructs has some impact on bacterial growth (Figure 5.3-A1/B1 and Figure 5.4-A1/B1). SecA is a 100 kDa homodimer responsible for both the targeting and coupling of energy required for protein translocation through the membrane channel (Dajkovic *et al.* 2016). SecA ATPase activity is stimulated by the presence of membrane, the SecYEG complex and precursor protein (van Wely *et al.* 2001). The SecYEG complex is the core of the translocation channel associated to the cytoplasmic membrane with SecY predicted to span the membrane 10 times and located in randomized dispersed foci (Suh *et al.* 1990; Dajkovic *et al.* 2016). It was anticipated that the

overexpression of these proteins might impact on cell fitness, although any effects on cell physiology have not been investigated.

The effect on the native and heterologous enzyme production of overexpressing the motor or heterotrimeric pore forming complex components of the translocase at different stages of the growth cycle and from different expression constructs are shown in Figure 5.3 and Figure 5.4.

The strains overexpressing SecA resulted in reduced levels of production of both native and heterologous enzymes (Figure 5.3). For example, the data in Figure 5.3-A2 show that the level of model enzyme production by the version of SecA with its native gene sequence and RBS (BRXC21P1A) was identical to the version of that with codon pair optimised gene sequence and with a synthetic RBS (strain BRCX31P1D). There is an overall decrease in the production of XynA in both strains which is particularly significant during exponential growth. Data in Chapter 3 showed that, at this stage of growth, XynA shows a high rate of production (Figure 3.5-A) and is rapidly synthesised, processed and secreted by the cell (Figure 3.9-A). While it was anticipated that increasing the motor component of the translocase would only increase enzyme production if SecA was a pathway-limiting factor, it is not clear why its over production decreases productivity. Presumably, this reflects physiological differences underlying the slightly different growth profiles observed when SecA is overexpressed (Figure 5.3-A1). Interestingly, the same effect was observed for the heterologous AmyM enzyme, although this was less pronounced, particularly at very late stages of the growth cycle (Figure 5.3-B2), when this enzyme is normally secreted (Figure 3.5-B) and is very slowly released from the cell (Figure 3.9-B). It is possible that the physiological effects of SecA overexpression are less pronounced on AmyM production since the additional SecA copies may help to alleviate the saturation of the translocase caused by the slow release of AmyM.

Figure 5.4 shows the effect of three different versions of the synthetic polycistronic operon SecYEG on XynA (Figure 5.4-A2) and AmyM (Figure 5.4-B2) production. The impact of SecYEG overexpression is dependent on the enzyme, the growth phase and expression construct. The overexpression of the codon pair optimised version of the *secYEG* genes in strain BRCX31P2B is less detrimental for protein production at the later stages of growth than that of the strains carrying the native sequences of these genes (strains BRCX31P2A and BRCX31P2C). The effect of SecYEG overexpression on AmyM production was growth phase dependent (Figure 5.4-B2): during the exponential phase the effects were minor, however, there is a sharp decrease

in production during the transition phase in all of the strains in which the heterotrimeric SecYEG complex is overexpressed. This effect is very different from what was observed at the same growth phase for XynA production. The transition phase represents a readjustment of the several cellular functions associated with arrest of cell division. It is possible that the overexpression of this membrane-associated complex has a negative impact at this stage of growth which particularly impacts on the slowly released AmyM. AmyM production recovers from this effect later on at 24 h in the strains that carry the extra copies of the non-optimised *secYEG* operon, accompanied by a moderate increase of enzyme production. Interestingly, this is contrary to the effect observed for XynA at the same time point, where the induction of the non-optimised *secYEG* operon resulted in a decrease in production. Figure 5.2-B shows that strain BRC31P2B, which encodes the codon pair optimised *secYEG* genes, results in the highest expression values measured via qPCR, at the exponential phase, in comparison to strains BRC31P2A and BRC31P2C. Even though it is not possible to confirm that this difference in expression levels is still true after 24 hours of growth, it is likely that this is the cause for the difference in the effects on enzyme production between the strains expressing native and codon pair optimised versions of the *secYEG* genes.

In a previous study, Chen *et al.* (2015) showed that overexpression of SecA and SecYEG do not increase, and even decrease, the production of two heterologous amylases in *B. subtilis* after growth for 72 hours in rich media, while Mulder *et al.* (2013) showed that an increase in the SecY, SecE and SecG proteins contributes to the release of retained α-amylase inside the cells. However, the increase in protein the amount of AmyL protein in the extracellular fraction detected by Western Blotting was not matched by a corresponding increase in amylase activity. Although not commented on by the authors, this observation points to an important post-translocational bottleneck affecting the correct folding following its emergence from the translocase. Our results are in good agreement with these previous studies and, in addition, reveal that this approach for strain optimization not only effects native and heterologous enzyme production differently, but also has distinct effects at different stages of the growth cycle.

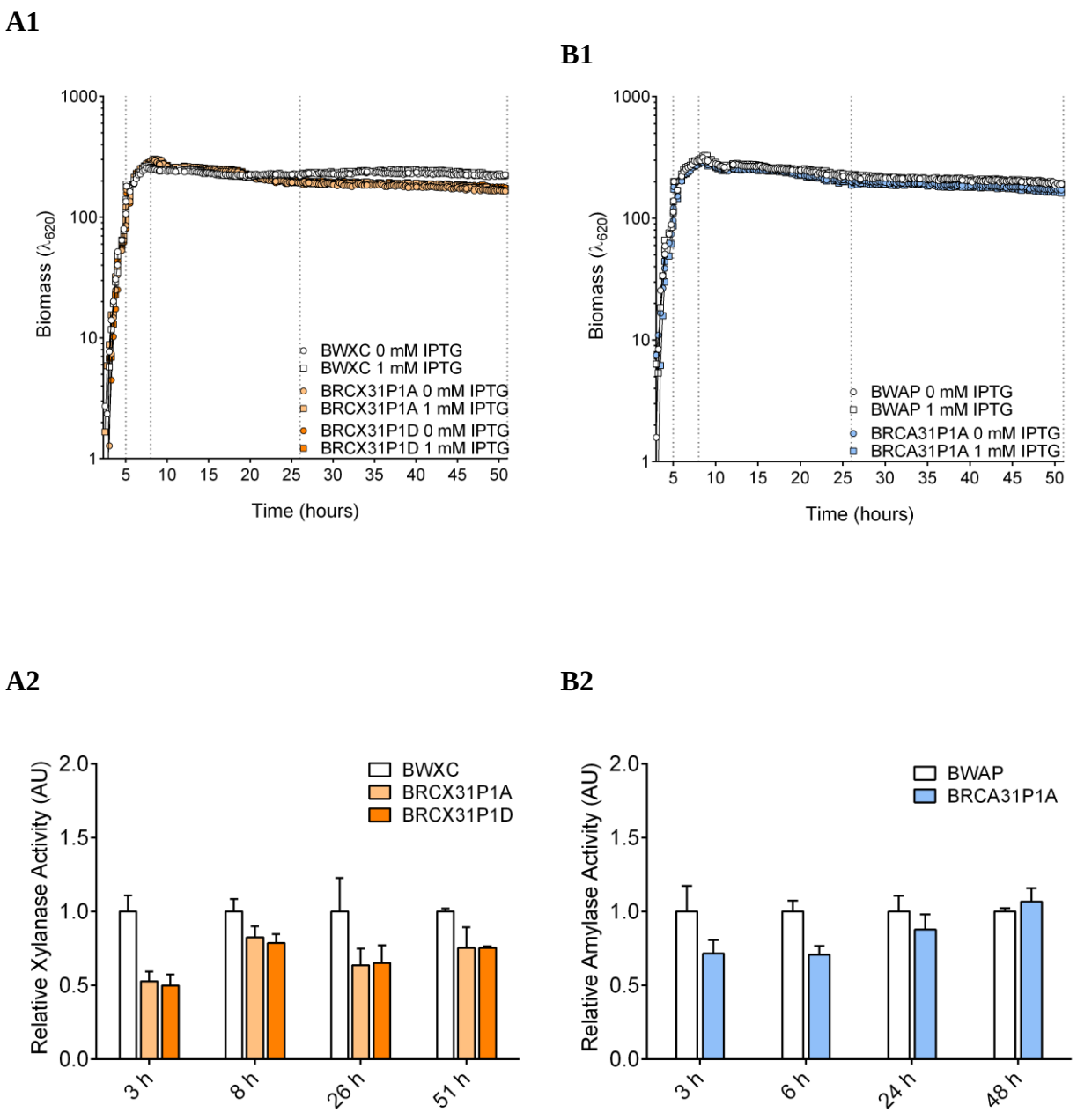


Figure 5.3 The effect of overexpressing synthetic SecA constructs on (1) growth and (2) enzyme production of (A) XynA and (B) AmyM. The cultures were grown in 1.5 mL LB supplemented with or without 1 mM IPTG in FlowerPlate® 48-well MTPs in the BioLector® bench-top microfermentation system (m2p-labs), and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The enzymatic activity was measured at four time points, as indicated by the dotted lined the growth curves (A1 and B1), using commercial assays (Section 2.10.). The enzymatic activities are relative to the strains BWXC and BWAP for XynA and AmyM production, respectively. The average growth and relative enzyme activities values of three biological replicates per strain were plotted, together with the corresponding SEM.

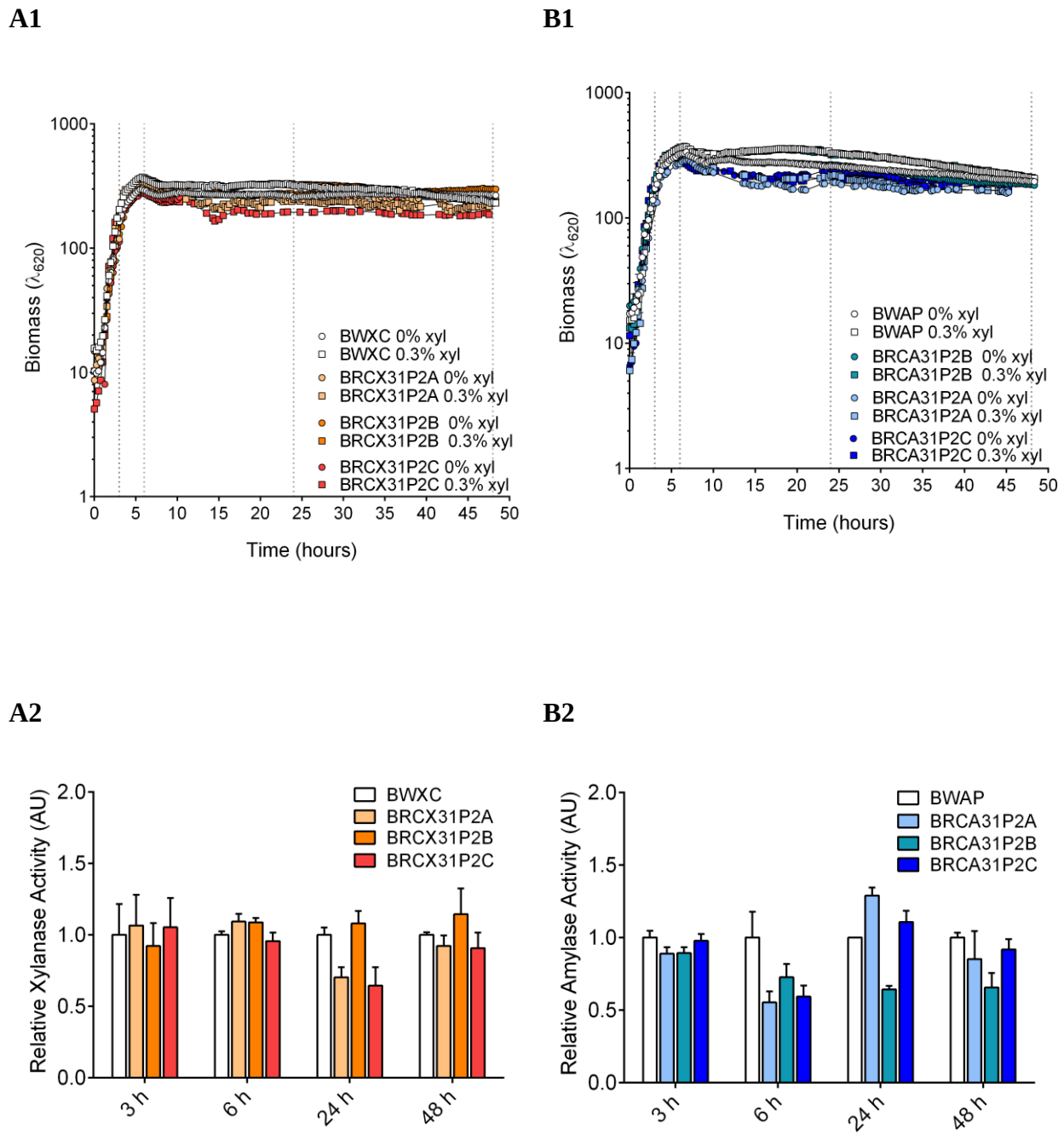


Figure 5.4 The effect of synthetic polycistronic SecYEG operons on (1) growth and (2) enzyme production of (A) XynA and (B) AmyM production strains. The cultures were grown in 1.5 mL LB supplemented with or without 0.3% xylose in FlowerPlate® 48-well MTPs in the BioLector® bench-top microfermentation system (m2p-labs), and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The enzymatic activity was measured at four time points indicated in the growth curves (A1 and B1) using commercial assays (Section 2.10.). The plotted enzymatic activities are relative to the strains BWXC and BWAP for XynA and AmyM production, respectively. The average growth and relative enzyme activities values of three biological replicates per strain were plotted together with the corresponding SEM.

5.5. Conclusions

In this Chapter, we explored a synthetic biology approach to address potential enzyme production bottlenecks associated with the saturation of the Sec translocons with preprotein substrates.

A combinatorial design approach was used to construct and to compare the effects of multiple synthetic expression operons encoding the SecA and SecYEG genes. In summary, two inducible expression systems were designed for the expression of *secA* and a polycistronic operon expressing *secY, secE* and *secG*, using combinations of the native or synthetic RBS with either the native or codon pair-optimised versions of the genes (Table 5.3). The synthetic DNA constructs were assembled using the StarGate® cloning system to combine for each construct two synthetic BioBricks (promoter and coding sequence) in an integrative vector with a colorimetric counter-selection for Campbell-like integration events. Constructs expressing high levels of native XynA and heterologous AmyM were added to these strains to investigate the impact of overexpressing a synthetic translocase on enzyme production.

Our results show that the overexpression of the synthetic constructs has an impact on the cell growth of the industrially relevant enzymes. This was at first evident due to the failure to isolate strains expressing all the combinations of the synthetic constructs for both enzymes, but further confirmed by the different growth profiles of the strains (Figure 5.3-A1/B1 and Figure 5.4-A1/B1).

In agreement with previous studies using synthetic SecA and SecYEG operons, we observed no major improvement on the production of either the native and heterologous enzymes (Leloup *et al.* 1999; Chen, Fu, *et al.* 2015). However, there were more subtle effects throughout growth that were distinct for both enzymes. Similarly, Leloup *et al.* (1999) observed that even though the yield of levansucrase was proportional to the amount of SecA, while α-amylase secretion was insensitive to a large decrease in the SecA level. Chen *et al.* (2015) observed distinct impacts on the production of two heterologous amylases, AmyS (*G. stearothermophilus*) and AmyL (*B. licheniformis*), when systematically expressing several components of the secretion pathway, including SecA and SecYEG. These findings suggest that different precursors exhibit different dependencies on the amounts of SecA and SecYEG. The physico-chemical properties of the signal sequences and mature peptides are also likely to affect the interaction with SecA and the translocase channel.

Furthermore, this study revealed that the effects of overexpressing the motor and pore forming components of the translocase on enzyme production upon depend on both the growth phase and expression construct. This observation suggests that the expression of the translocase components can be optimised to express these components at specific growth phases to maximise production, for example exponential phase for XynA and stationary phase for AmyM. This is an important observation reflecting the complexity of the translocase channel preponderance and the need for fine-tuning each production strain depending on the product of interest and growth conditions.

Overall, this approach reveals a promising strategy for strain optimization. Even though there is no indication for major production improvements, it is likely that a combinatorial approach aimed at relieving several secretion bottlenecks simultaneously can contribute to the development of strains with high levels of secretion capacity. Importantly, a systematic and deep insight into how components of Sec pathway interact with each other may be the key to improving the yield of heterologous proteins.

Chapter 6 General Discussion

Close collaboration between academia and industrial manufacturers has played an important role in the development of Biotechnology and the Bioeconomy. This thesis is the product of such a collaboration, between Newcastle University and DSM, a worldwide biotechnology company. The aim of the collaboration was to tackle a number of important issues relevant to industrial enzyme production. Despite the industrial enzyme industry being a well-established and profitable industry, it faces the constant challenge of producing innovative enzymes with novel substrate specificities and wider activity spectra. This requires optimising the production of an increasingly large variety of heterologous commercial proteins in a high competitive market.

Bacillus subtilis is a well-recognised industrial workhorse due to its capacity to secrete proteins at concentrations as high as tens of grams per litre. The use of this bacterium for decades for industrial enzyme production and for fundamental research has made *B. subtilis* the model organism for Gram-positive bacteria. Therefore, many studies have addressed important cellular processes such as cell division, sporulation, biofilm formation and protein secretion. Despite extensive research, the production of heterologous proteins by *B. subtilis* is still a hit and miss process, with the encountered difficulties relating to incompatibilities associated with the characteristics of both the target protein and the secretion system itself (Pohl & Harwood 2010). Our approach was to carry out a comparative study of two industrially-relevant production strains, one encoding a native industrial enzyme, the other encoding a heterologous protein. Given that background production strains were identical, difference in cell physiology would be expected to point to the impacts of the target enzymes on secretion. By understanding these differences, it was expected that this would lead to new potential strategies to improve production yields. Therefore, factors such as protein expression, processing, translocation and stability were evaluated while addressing the issue of production strain optimisation.

This thesis focuses on two industrial enzymes, the native xylanase, XynA, and the heterologous AmyM from *Geobacillus stearothermophilus.* XynA is a 23 kDa endo-1,4-β-xylanase that catalysis the endohydrolysis of (1->4)-β-D-xylosidic linkages in xylans (Brenda, 2016). This enzyme degrades xylan, the major hemicellulose in cereals and hardwoods, and has great value in

the baking, wood pulp and paper industries. AmyM is a 79.2 kDa maltogenic α-amylase that catalysis the hydrolysis of (1->4)-α-D-glucosidic linkages in polysaccharides so as to remove successive α-maltose residues from the non-reducing ends of the glycan chains (Brenda, 2016). α-Amylases degrade starch which has a wide variety of applications in the food, textile, fuel alcohol production, paper and detergent industries (Polaina & MacCabe 2007; de Souza & de Oliveira Magalhães 2010). AmyM is secreted naturally by the thermophile *Geobacillus stearothermophilus* and therefore it has particular commercial interest in a number of applications, such as the enzymatic liquefaction and saccharification of starch which are performed at high temperatures (de Souza & de Oliveira Magalhães 2010). The two main production strains of this study, BWXP and BWAP, express high levels of each of these enzymes due to a high copy number plasmid-based expression system under the control of a constitutive promoter (Table 3.1).

At first, we carried out a biochemical characterization of the model enzymes, revealing that XynA is optimally active at pH 7.1 and 60 ºC, while AmyM is a highly thermostable enzyme with optimal pH of 6.5 and an optimal temperature of 90 ºC (Figure 3.2). An analysis of strains BWXP and BWAP showed two very different enzyme production profiles (Figure 3.4) despite their identical expression systems. While xylanase production by BWXP coincided with the start of exponential phase, amylase production was initiated during the transition from exponential to stationary phase. Together, these are important observations that clearly illustrate how the characteristics of the target proteins reflect the challenges associated with the optimisation of enzyme production. The distinct production profiles are presumably related to post-transcriptional processes involving either membrane translocation, or interaction with or processing by the upstream or downstream elements of the protein secretion pathway, all of which were analysed further in this study.

Through pulse-chase analysis we identified a clear difference between the secretion kinetics of XynA and AmyM (Figure 3.9). XynA is synthesised and secreted rapidly, taking less than a minute to be translated, folded and secreted. XynA was also subjected to limited proteolysis. In contrast, AmyM is released slowly into the growth medium, suggesting extensive cell-associated bottlenecks in addition to a degree of degradation. The distinct secretion kinetics of XynA and AmyM most likely reflects the different interactions of each target protein with the secretory machinery at the later stages of secretion. Prompted by this observation, we investigated the

effect of the overexpression of XynA and AmyM on the regulation of components of the secretion machinery, via RT-qPCR. Our data showed that both enzymes have similar effects on the regulation of some components of the secretion machinery, namely the down-regulation of YrbF, associated with secretion efficiency, and up-regulation of PrsA, the major membrane bound extracellular chaperon (Figure 3.10). Furthermore, an upregulation of the quality control proteases HtrA and HtrB in response to XynA and AmyM production was observed (Figure 3.10). This response is much more significant in case of the AmyM producing strain which shows an up-regulation of *htrA* and *htrB* expression of up to 24 and 34-fold, respectively, compared to 1.7 and 2.8-fold in the case of XynA production (Table G.3, Appendix G).

The strong up-regulation of the quality control proteins caused by protein secretion stress response inspired us to develop a biosensor systems to monitor stress during secretory protein production. In Chapter 4 we made use of a protein secretion stress fluorescent biosensor based on the *htrA* promoter, as described by Trip *et al.* (2011), to facilitate the continuous monitoring of the secretion stress response during the hyper-production of the model enzymes under a controlled environment, using the BioLector® bench top microfermentation system (m2p-labs) (Figure 4.3). As expected, the biosensor reporter system showed a very distinct secretion stress response upon production of the native and heterologous model enzymes. Despite having detected an up-regulation of *htrA* transcription, the P_{htrA}-sf*gfp* reporter fusion did not show a detectable induction of the secretion stress response upon high levels of production of the native enzyme XynA (Figure 4.3-A). In contrast, AmyM hyper production resulted in a strong fluorescent signal of the P_{htrA}-sf*gfp* gene fusion that peaks at transition phase, shows a temporary decline in early stationary phase before increasing steadily throughout late stationary phase (Figure 4.3-A). At this stage, it is not clear whether the peak at transition phase translates an effective attenuation of the secretion stress, a general down regulation of the cell metabolism due to transition to the stationary phase of growth, or a combination of both. The rate of AmyM production is the highest at this stage (Figure 3.4) and therefore this behaviour is intriguing. A comprehensive comparison of the whole transcriptome of these production strains at different stages of growth could potentially clarify the regulation of the secretion stress reporter and identify new targets for manipulation of the secretion pathway and attenuation of secretion bottlenecks.

In the future, it would be interesting to perform an identical analysis for the production of native and heterologous proteins which are homologous. Our results suggest that AmyM is slowly released by the cell under secretion stress conditions, however it is not clear whether these effects are due to the heterologous nature of the protein, or the properties of the protein itself. Particularly, AmyM is a significantly larger protein (79.2 kDa) when compared to the native XynA model (23 kDa) which is likely to difficult the passage through the cell wall. This highlights the importance of complementary pulse-chase and western blot analysis which were not possible for AmyM due to unavailability of a suitable antibody. A combined analysis of intracellular and extracellular protein location during secretion with enzymatic assays would help to clarify where is the bottleneck for AmyM secretion. However, it is likely that AmyM strongly interacts with the cell wall due to its size and/or surface charge. This could be confirmed, for instance, by performing our enzyme secretion kinetics studies in protoplasts.

Next, we swapped the signal peptides of the precursors of XynA and AmyM and observed that, even though the differences in the production profiles remained unchanged, the production levels of both enzymes were drastically reduced and, consequently, the secretion stress response upon hyperproduction of the AmyM recombinant precursor was eliminated. These observations are in accordance with the view that each signal peptide has co-evolved with its cognate mature protein, and the reduction of productivity is likely to be related to early stages of secretion, possibly an intracellular stress response, in contrast to the secretion stress response observed in strain BWRAP (Figure 4.3-B).

Using a set of strains with systematic deletions of the proteases genes associated with the instability of secreted proteins, we correlated both the bottlenecks of cellular and extracellular-associated proteases with native and heterologous enzyme production, and the secretion stress response. We incorporated the P_{htrA}-sf*gfp* reporter fusion and the high copy number plasmids expressing either XynA or AmyM in each of these strains and observed very different results of hyper production with respect to the secretion stress response and enzyme production levels. However, in both cases it was evident that heavy editing of the proteolytic capacity of the cell does not, as might have been anticipated, translate into optimal enzyme production. This is partially associated with the compensatory regulation of the quality control proteases HtrA and HtrB, and the HtrC protease. An alternative approach control the proteolytic activity of the cell could involve the de-coupling of the CssRS two-component system. Although this is likely to

have a negative effect on the fitness of the cell, it would help to clarify whether this bottleneck is in fact necessary for maintaining important cell functions and viability in the production strains.

Moreover, pulse-chase analysis suggested that the deletion of HtrA or HtrB influences the secretion kinetics of XynA, leading to a faster release of this secretory protein (Figure 4.10-A). However, these strains did not show an increase in enzyme production, highlighting the importance of the quality control mediated by HtrA and HtrB.

Our results showed that biosensor-like systems, such as the reporter P_{htrA}-sf*gfp* fusion, are not necessarily reliable indicators of enzyme production as suggested previously (H. Westers *et al.* 2004; Trip *et al.* 2011). We detected sensitivity limitations in this system and, more importantly, that the highest level of secretion stress did not correspond to the highest level of protein production. In particular, strain BRB13, which has all seven extracellular proteases deleted as well as *wprA* and the quality control proteases *htrA* and *htrB*, showed the highest levels of secretion stress response upon production of XynA and AmyM, but also the lowest level of enzyme activity in the extracellular medium. Overall, the results of Chapter 4 suggest there is a correlation between the levels of extracellular proteolytic activity and cell fitness, which impacts on the regulation of the secretion stress response and may compromise enzyme production levels. However, the is likely to be target protein specific, as shown by Pohl *et al.* (2012).

It was interesting to investigate whether the Sec translocons were saturated during high level of enzyme production and if this was an important bottleneck for one or both of the model enzymes. We took a combinatorial design approach to compare the effects of multiple synthetic expression constructs and optimise gene expression of the motor and pore forming components of the Sec translocase. We observed that the expression of the synthetic constructs has an impact on the cell growth upon overexpression of these industrially relevant enzymes (Figure 5.3-A1/B1 and Figure 5.4-A1/B1). There was generally no major improvement on the production of the native and heterologous enzymes upon overexpression of the translocase components (Figure 5.3 and Figure 5.4), however, there was a range of effects throughout growth that were distinct for both enzymes and, in some cases, specific for each expression construct. Together these observations suggest that there is a potential for optimisation of the expression of the translocase components, however, this must be tailored for each target protein and for the growth phase at which production is optimised. Overall, this illustrates the complexity of the translocase channel and the need to fine-tune each production strain according to the product and growth conditions.

A major conclusion from the studies carried in this thesis is that future work should use a combinatorial approach to optimise the expression of several components of the secretion pathway and translocase and fine-tune their expression to the product of interest. Such an approach would provide new insights into how the components of the Sec pathway interact with each other and their substrates. Furthermore, it would be valuable to routinely combine these strains with biosensors that facilitate continuous monitoring of the secretion stress response. Our studies have focused on the detection of extracellular enzyme activity. Systematic Western Blot analysis of intra and extracellular fractions would allow clarification of whether the overexpression of the translocase components contributed to an increased release of the overexpressed model enzymes. Combined with a quantification of the secretion stress response, these studies would further clarify where, in the secretion pathway, the major limitations for enzyme production occur.

Appendix A - Media and Buffers

10x PBS

NaCl	80 g/L
KCl	2 g/L
Na_2HPO_4	14.4 g/L
KH_2PO_4	2.4 g/L

10x STD

NaCl	9% (w/v)
Triton X-100	10% (w/v)
Sodium deoxycholate	5% (w/v)

1X SDS-PAGE sample buffer

NuPAGE® LDS Sample Buffer (4x)	1x
NuPAGE® Reducing Agent (10X)	1x

1x Spizizen-plus medium

2x Spizizen medium	1x
Glucose	0.5% (w/v)
Tryptophan	20 µg/mL
Casamino acids	0.02% (w/v)
Ferri ammonium citrate	2.2 µg/mL

1x Spizizen-starvation medium

2x Spizizen medium	1x
Glucose 40% (w/v)	0.5% (w/v)
Casamino acids 20% (w/v)	0.02% (w/v)

2x Spizizen medium

K_2HPO_4	28 g/L
KH_2PO_4	12 g/L
Potassium L-glutamate	4 g/L
Na_3-citrate.$2H_2O$	2.3 g/L
$MgSO_4$ .7 H_2O	0.4 g/L
pH	7

3,5-Dinitrosalicylic Acid Reagent (DNS Reagent)

3,5-Dinitrosalicylic Acid	10 g/L
Sodium potassium tartrate tetrahydrate	30 g/L
NaOH	400 mM

Amylase Reaction Buffer

K_2HPO_4	17.42 g/L
KH_2PO_4	13.61 g/L
pH	5.5

Cell Lysis Solution (CLS)

EDTA	25 mM
SDS	2% (w/v)

Fix and Wash Solution

Methanol	10% (v/v)
Acetic Acid	7% (v/v)

Lysis Buffer

Tris.Cl (pH7.2)	10 mM
MgCl2	25 mM
NaCl	200 mM
Lysozyme	5 mg/mL

Nutrient Luria-Bertani medium (LB)

Tryptone	10 g/L
NaCl	10 g/L
Yeast Extract	5 g/L
pH	7.2-7.5

Nutrient Luria-Bertani medium agar (LB agar)

Tryptone	10 g/L
NaCl	10 g/L
Yeast Extract	5 g/L
Agar	15 g/L
pH	7.2-7.5

PMT Buffer

10x PBS	1x
Skim Milk powder	5% (w/v)
Tween	1% (v/v)
pH	>7.6

Protein Precipitation Solution (PPS)

Ammonium acetate	10 M

SMM + Ribose

2x Spizizen medium	1x
Methionine free amino acid solution	20 µg/mL
Ferri ammonium citrate	2.2 µg/mL
Ribose	1% (w/v)

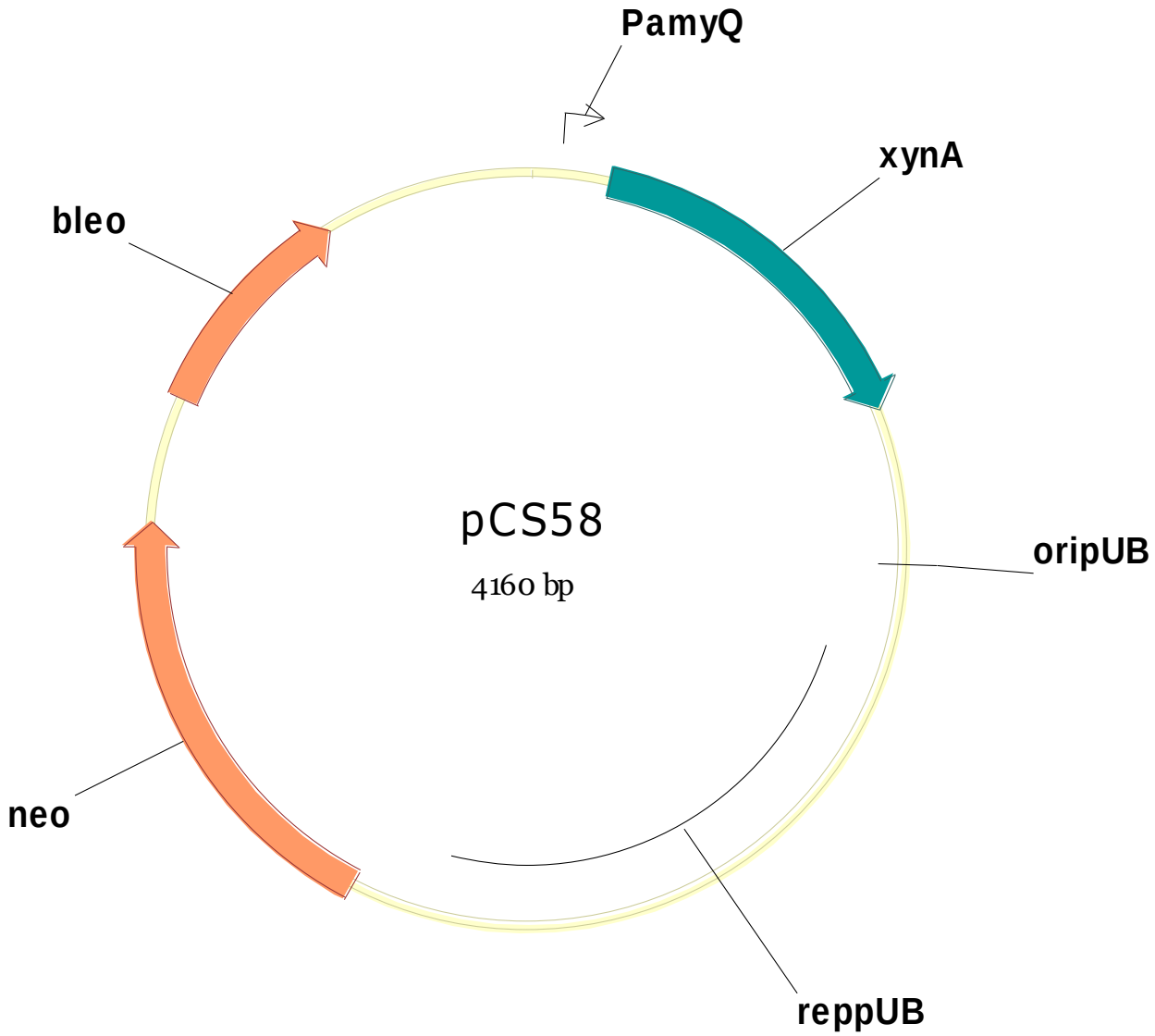
Stopping Buffer

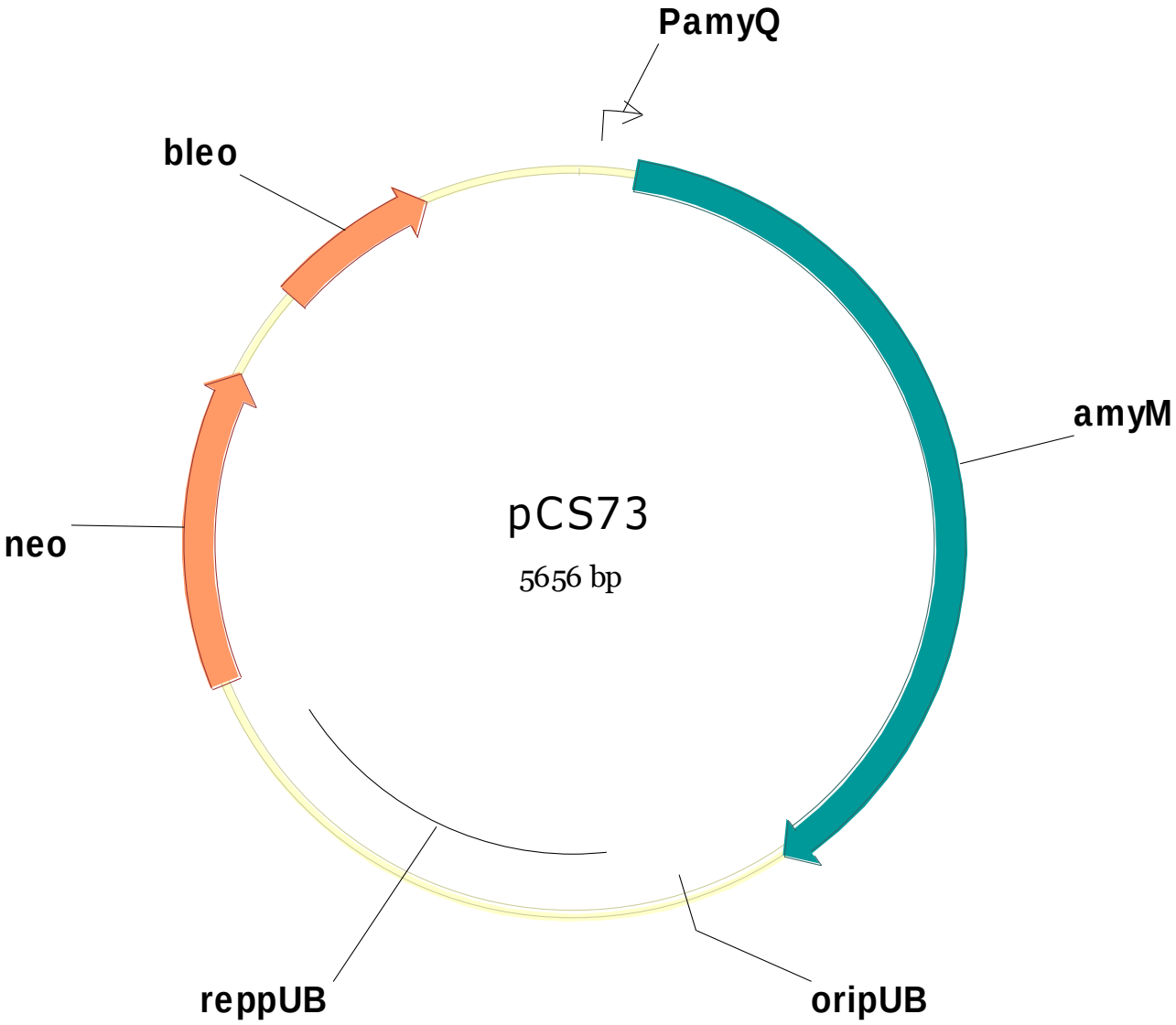
Tris Base	1 M
pH	11

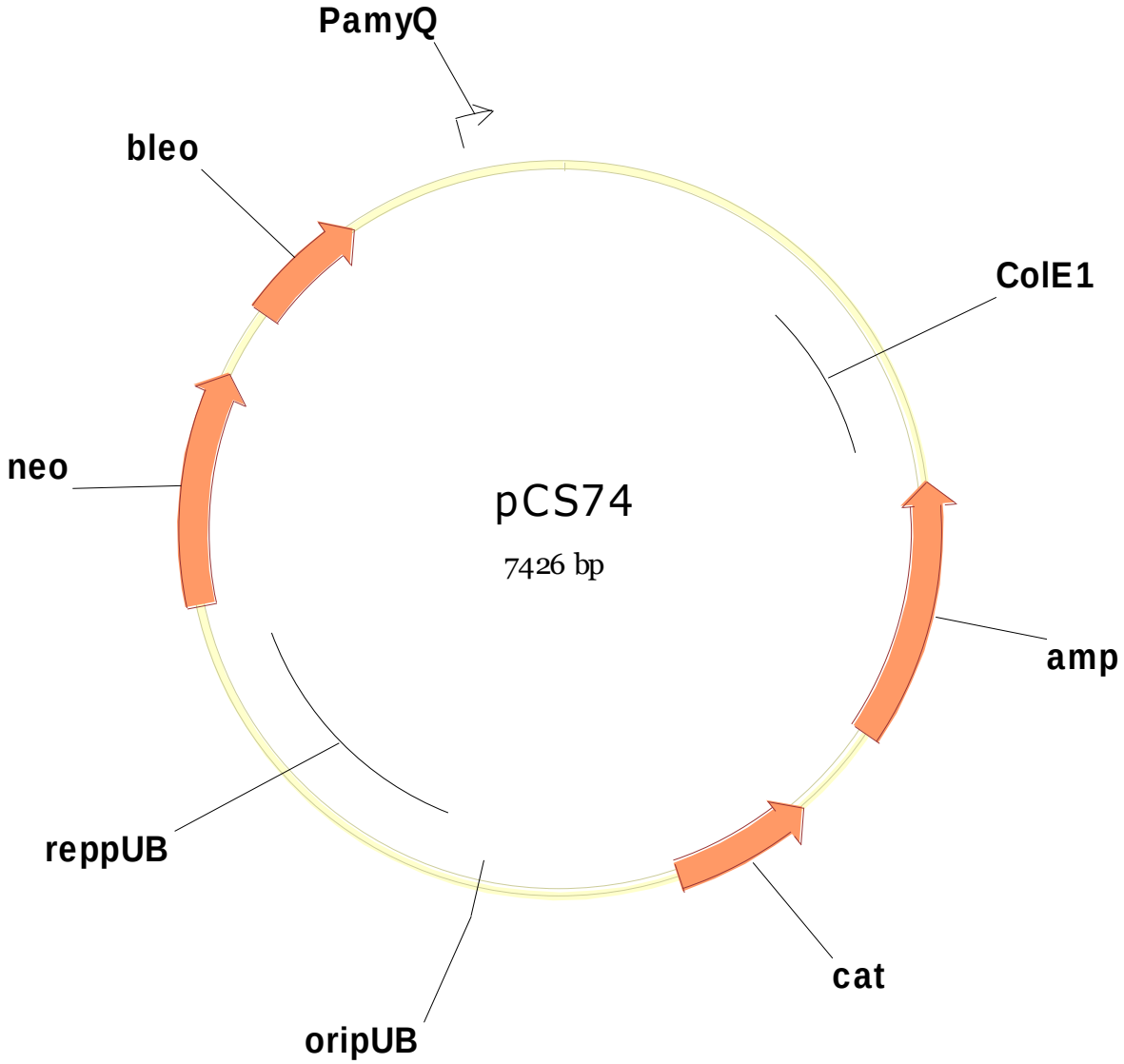
Xylanase Reaction Buffer

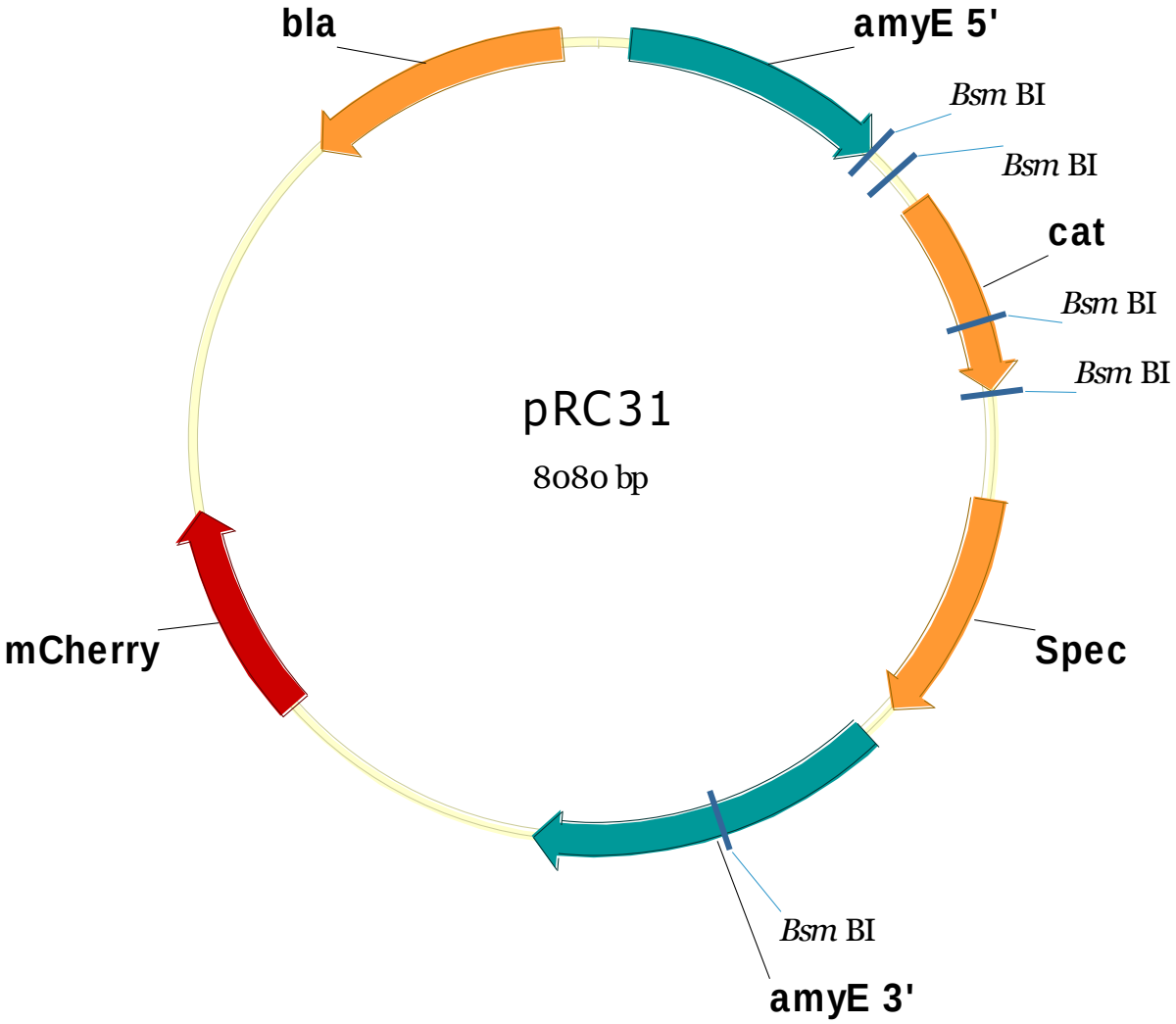
Sodium Acetate	100 mM
pH	4.6

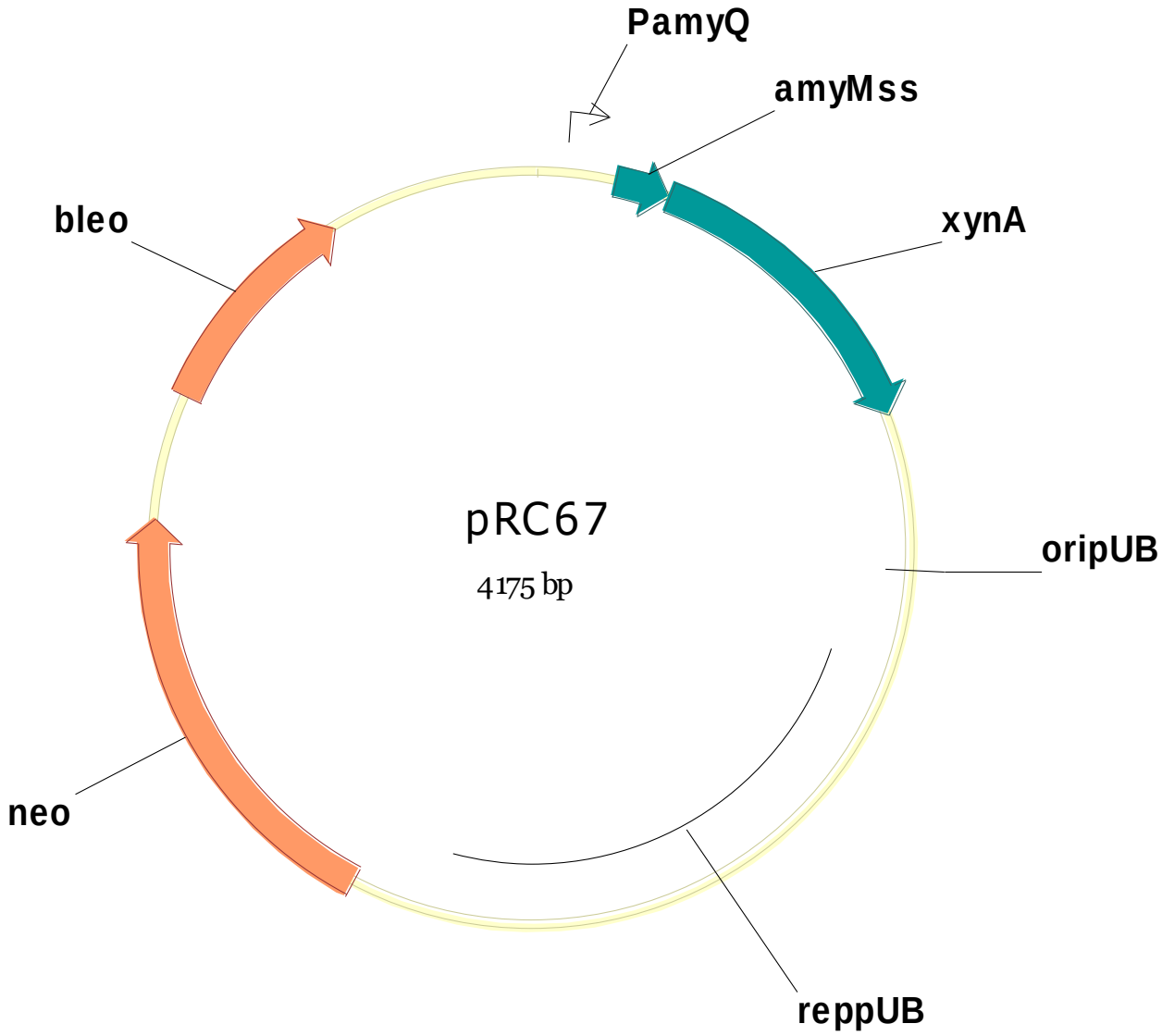
Appendix B – Plasmid maps

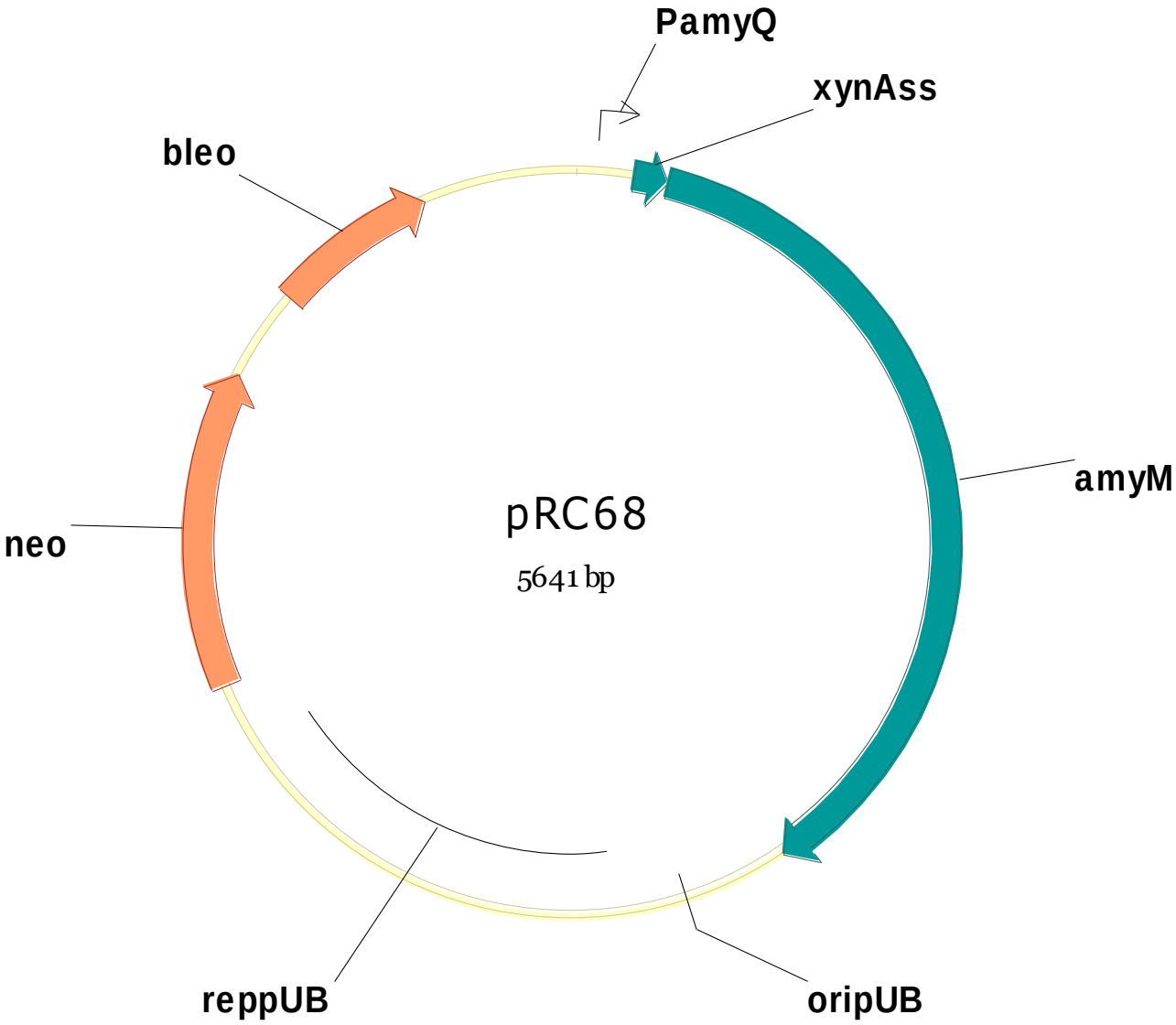



PamyQ
bleo
amyM
pCS73
5656 bp
neo
reppUB
oripUB


PamyQ
bleo
ColE1
neo
pCS74
7426 bp
amp
reppUB
cat
oripUB


bla
amyE 5'
Bsm BI
Bsm BI
cat
Bsm BI
Bsm BI
pRC31
8080 bp
mCherry
Spec
Bsm BI
amyE 3'


PamyQ
amyMss
xynA
oripUB
reppUB
neo
bleo
pRC67
4175 bp


PamyQ
xynAss
bleo
amyM
pRC68
5641 bp
neo
reppUB
oripUB

Appendix C – Oligonucleotides sequences

Figure C.1 Nucleotide sequences of the the oligonucleotides used throughout this study.

Number	Sequence 5'-3'
91	CAT CTA CAC CGC CAT AAA C
92	AGG GTT GCC AGA GTT AAA GG
93	CCT TTA ACT CTG GCA ACC CTG AAT TCA TCC ACG CTG TG
94	CAT ATG TAA ATC GCT CC
95	GGA GCG ATT TAC ATA TGA GCA AAG GAG AAG AAG
96	GTT TAT GGC GGT GTA GAT GAT AAG GCC GCC TCT TAA C
319	AGC GCG TCT CCC CGC CTC AAT AAA TAG CTC ATT CTC
320	AGC GCG TCT CCT TAT CTA TTT GTA GAG CTC ATC CAT G
353	TAA GAA AAT GAG AGG GAG AGG
354	TTG ACG CTT GCG GAA CTG CTT GCA GAG GCG GTT GCC GAA
355	TTG CCA GTA GTC TGT GCT AGC GGC TTC AGC GGC GTT TGG
356	CCT CTC CCT CTC ATT TTC TTA
357	GCT AGC ACA GAC TAC TGG CAA
358	AGC AGT TCC GCA AGC GTC
406	CAG TGA TGA CCA CGT CAC AAC GTC GAA AAT TGG ATA AAG TGG
407	GAA TCA CCG AAG TTT GCC ACA CCG ACT GTA AAA AGT ACA GTC
408	CAG TTT TGG AAA TCA GCG GCC GTC GAA AAT TGG ATA AAG TGG
409	GTG AAG AGT CAC CAA AGC TCA CCG ACT GTA AAA AGT ACA GTC
410	GAA AGG TAA TGA GAA TGA GGT C
411	TTG TGA CGT GGT CAT CAC TG
412	GTG GCA AAC TTC GGT GAT TC
413	CGT TTC GGA GAA TTT TCA CTT C
414	TGG ATT ATC GAC GTG ATG GC
415	GCC GCT GAT TTC CAA AAC TG
416	GAG CTT TGG TGA CTC TTC AC
417	TGC TTT CTG TCT GCT TGG TC
362	CCT AAA TGG CGA ACC TGT AGT C
363	AAA CGC TGT CAA TGG GTC TG
368	ACA TCA GCA ACT GGG ACG AC
369	GCT TCA CCG CAT CAA TCC

Number	Sequence 5'-3'
370	GAT TCA CCT TCG TGC TTA CG
371	TTG TTA TCG TCG CCT TCT TG
374	CAT TCA CCT TCG TGC TTA CG
375	TGT CAT CGC CTT CTT GAG G
378	GGT TAT CGT CAG CAT TGC AC
379	AAC AAG ACT GCC AGC ACT ACC
384	TCG CTT TAA TCA TCG TTG TGC
385	ACG AAG AAC AAT ACA GCA AGC A
386	ATC GCG TCA TTC TTT GGA AC
387	CAG CCA TTT GTT CAG GGT TT
390	AGA TGT TGG GAA AGA AAT GAA AA
391	TCA ACT ATT AAA CGA ATT AAT TGA GAA
394	TAC GGC ATT CAT CCA GTC AA
395	CGC ACT TCC TGT GAA ACT GA
398	ACG ACT TTG CTT TCG TCG AT
399	TGC GGT CCT TTA CTT CTT GC
402	CTG AAC AGC TGC GTT GTC AT
403	TGG ACG AAA CAT TCA GCA AA
570	CGA AAG AGA ACG ATC AGA GC
571	AAA TGC CAA GAG CAA GAC TG
572	GAA GCT GAT TGG ACA TTC TG
573	GAA GCT GAT TGG ACA TTC TG
574	TTA CGA ACG TGA GGA AGA AC
575	TGT ATC CAG CCC TTC ATT TG
576	GAA AAT TCA GCT CGG TTG TG
577	TTT GCC GTT TTC CAG AGC AG
14526	TTG GCG GTT CAG TCA GTG TG
14527	GAA GAA ATC GGA ATG GCA AAT GTC
14528	ACA GTA GGC GGA CAA GAA GAT G
14529	GCC GTG AAC AGA ACC TAA TGC
14536	GCT TGG CGG CAT CAT TGA G
14537	TTG ATC GTA ATC CAC CCA TAA ACC

Appendix D - Determining Xylanase Enzyme Units

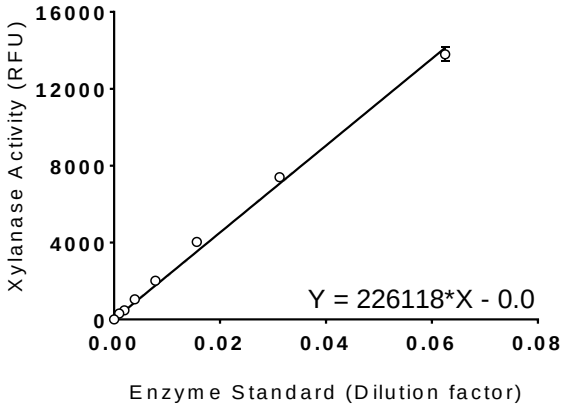


Figure D.1 EnzChek® Standard Curve.

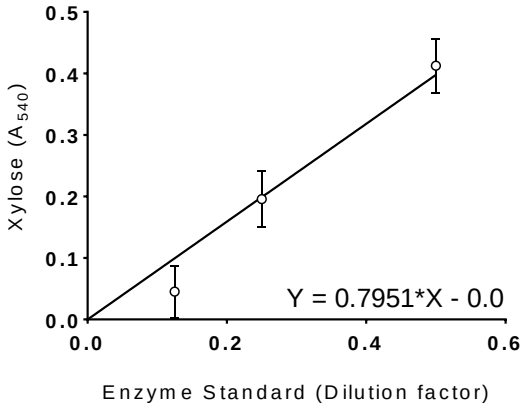


Figure D.2 DNS Standard Curve.

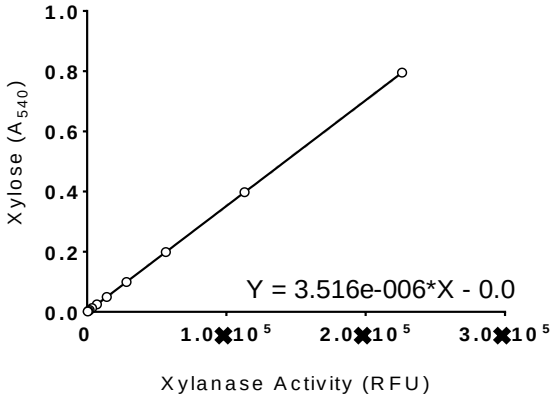


Figure D.3 Xylanase Enzymatic Reaction Standard Curve.

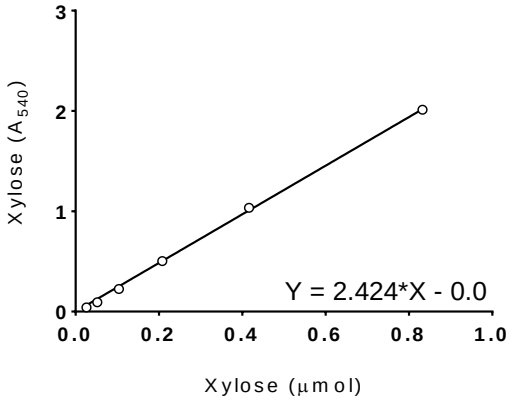


Figure D.4 Xylose Standard Curve.

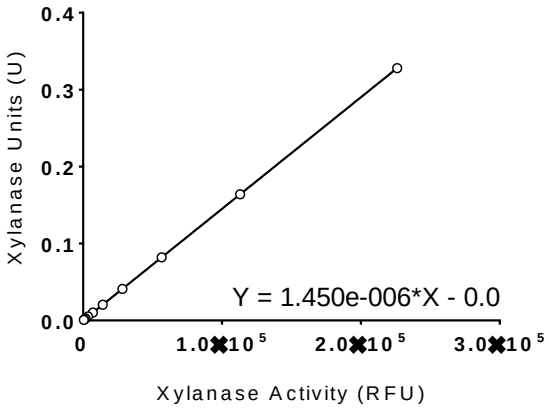


Figure D.5 Xylanase Enzymatic Units Standard Curve.

Appendix E - Determining Xylanase Enzyme Units

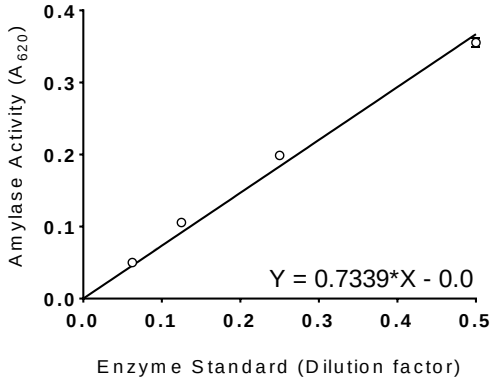


Figure E.1 Phadebas® Standard Curve.

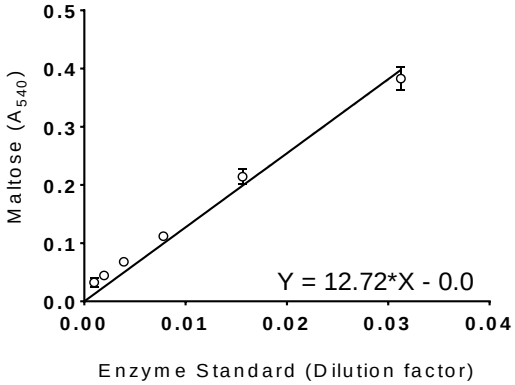


Figure E.2 DNS Standard Curve.

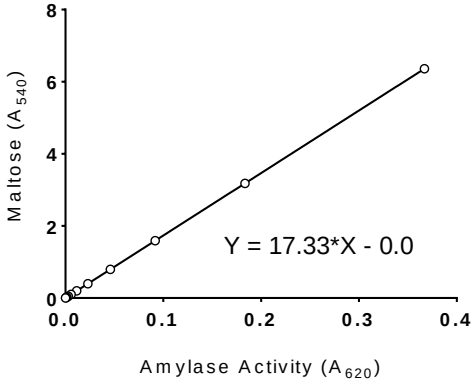


Figure E.3 Amylase Enzymatic Reaction Standard Curve.

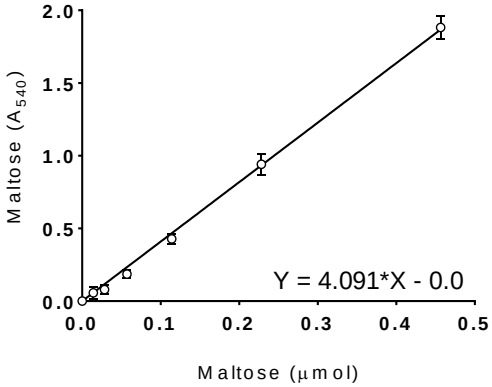


Figure E.4 Maltose Standard Curve.

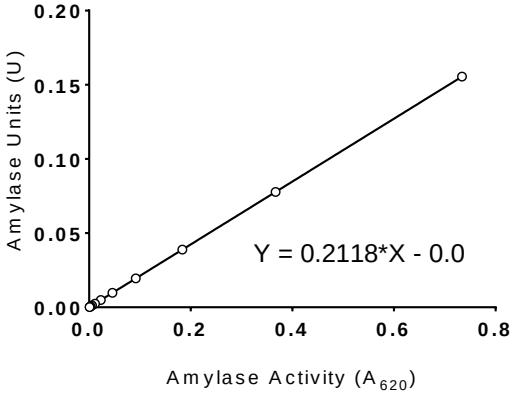


Figure E.5 Amylase Enzymatic Units Standard Curve.

Appendix F – Enzymatic Activity Assays

Table F.1 Effect of pH (A), temperature (B) and thermostability (C) on XynA and AmyM enzymatic activities, presented on Section 3.3.1 . Enzymatic assays were performed in triplicate according to Section 2.10. Enzymatic units were calculated with the correlations derived in

A

pH	Xylanase Activity (mU/mL)		
	mean	SD	N
3.3	3.07	0.22	3
4.8	5.01	0.21	3
5.3	5.85	0.44	3
6.5	7.23	0.10	3
7.1	8.20	0.36	3
10.3	7.34	0.35	3

pH	Amylase Activity (U/mL)		
	mean	SD	N
3.6	1.30	0.13	3
4.6	2.59	0.2	3
5.6	3.01	0.11	3
6.5	3.02	0.12	3
7.5	2.25	0.15	3
8.2	1.45	0.06	3
9.6	1.23	0.05	3
10.33	0.86	0.07	3

B

T (°C)	Xylanase Activity (mU/mL)		
	mean	SD	N
20	0.74	0.14	3
30	1.79	0.10	3
40	3.73	0.31	3
50	5.77	0.60	3
60	6.87	0.72	3
70	1.83	0.22	3

T (°C)	Amylase Activity (U/mL)		
	mean	SD	N
25	0.55	0.03	3
30	0.58	0.01	3
37	0.78	0.03	3
45	1.32	0.06	3
50	1.86	0.14	3
60	3.92	0.24	3
70	7.08	0.17	3
80	8.53	0.60	3
90	10.81	0.14	3
99	7.86	0.24	3

C

T (°C)	Xylanase Activity (mU/mL)		
	mean	SD	N
41	4.92	0.12	3
50	1.68	0.14	3
60.3	0.61	0.17	3
65	0.79	0.06	3
75	0.78	0.12	3
84.3	0.81	0.12	3

T (°C)	Amylase Activity (U/mL)		
	mean	SD	N
41	3.16	0.72	3
50	2.56	0.41	3
60.3	1.93	0.04	3
65	1.81	0.17	3
75	0.47	0.01	3
84.3	0.47	0.01	3
94.4	0.48	0.02	3
99	0.45	0.01	3

Table F.2 Monitoring of XynA (A) and AmyM (B) production during the cultivation of the production strains BWXP and BWAP, respectively, using the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 2 mL LB in FlowerPlate® 48-well MTPs and incubated at 37 °C with 95% humidity and strong agitation (800 rpm). During a total of 33 hours of growth, 17 samples were taken and the enzymatic activity was assayed with commercial assays (Section 3.3.4). The fluorescence and absorbance values obtained with the commercial assays were converted to enzymatic units per millilitre using the derived formulas described in Sections 2.10.2 and 2.10.4 .

A

Time (hours)	**Xylanase Units (mU/mL)**					
	WT 168			BWXP		
	mean	SEM	N	mean	SEM	N
1.8	0.00	0.00	3	0.06	0.02	4
2.8	0.00	0.00	3	0.20	0.03	4
3.8	0.00	0.00	3	0.39	0.04	4
4.8	0.00	0.00	3	0.59	0.02	4
5.8	0.00	0.00	3	0.71	0.02	4
6.8	0.00	0.00	3	0.77	0.02	4
7.6	0.01	0.00	3	0.82	0.02	4
8.6	0.01	0.00	3	0.85	0.01	4
9.6	0.01	0.00	3	0.90	0.01	4
10.6	0.01	0.00	3	0.94	0.01	4
11.6	0.01	0.00	3	0.96	0.03	4
12.6	0.01	0.00	3	0.99	0.04	4
13.6	0.01	0.00	3	1.06	0.04	4
14.6	0.01	0.00	3	1.12	0.02	4
15.6	0.01	0.00	3	1.04	0.03	4
16.5	0.01	0.00	3	1.13	0.05	4
32.2	0.01	0.00	3	1.65	0.08	4

B

Time (hours)	**Amylase Units (U/mL)**					
	WT 168			**BWAP**		
	mean	SEM	N	mean	SEM	N
1.8	0.01	0.00	3	0.03	0.00	4
2.8	0.13	0.01	3	0.20	0.02	4
3.8	0.17	0.01	3	0.48	0.03	4
4.8	0.29	0.01	3	1.29	0.07	4
5.8	0.44	0.02	3	2.29	0.09	4
6.8	0.47	0.03	3	2.48	0.09	4
7.6	0.71	0.02	3	4.52	0.37	4
8.6	0.80	0.03	3	5.53	0.20	4
9.6	0.78	0.02	3	4.72	0.12	4
10.6	0.70	0.01	3	5.49	0.21	4

Time (hours)	Amylase Units (U/mL)					
	WT 168			BWAP		
	mean	SEM	N	mean	SEM	N
11.6	0.85	0.00	3	6.44	0.02	4
12.6	1.02	0.18	3	5.52	0.09	4
13.6	0.74	0.08	3	6.30	0.60	4
14.6	0.97	0.07	3	6.47	0.41	4
15.6	0.89	0.09	3	4.97	0.33	4
16.5	0.77	0.07	3	5.69	0.17	4
32.2	1.78	0.08	2	5.77	0.71	3

Table F.3 Enzymatic units corresponding to the analysis of the effect of swapping XynA and AmyM signal peptides on the enzyme production. The enzymatic units resulting of the production profiles of the strains BWXP2 (A) and BWAP2 (B) are presented in comparison with BWXP, BWAP and wild type 168. The cells were cultured in 2 mL LB, at 37 °C with 95% humidity and strong agitation (800 rpm) in FlowerPlate® 48-well MTPs using the BioLector® bench-top microfermentation system (m2p-labs). The enzymatic activity was measured using commercial assays (2.10.). The fluorescence and absorbance values obtained with the commercial assays were converted to enzymatic units per millilitre using the formulas derived in Sections 2.10.2 and 2.10.4 .

A

Time (h)	**WT 168**			**BWXP**			**BWXP2**		
	mean	SEM	N	mean	SEM	N	mean	SEM	N
3.38	0.02	0.00	5	0.23	0.02	5	0.32	0.02	5
4.10	0.01	0.00	5	0.48	0.02	5	0.36	0.01	5
4.67	-0.01	0.00	5	0.57	0.03	5	0.41	0.01	5
5.46	0.01	0.00	5	0.61	0.03	5	0.48	0.01	5
20.48	0.00	0.00	5	1.25	0.05	5	0.89	0.02	5
25.53	0.00	0.00	5	1.42	0.07	5	0.97	0.03	5

B

Time (h)	**WT 168**			**BWAP**			**BWAP2**		
	mean	SEM	N	mean	SEM	N	mean	SEM	N
3.38	0.16	0.02	5	0.21	0.04	5	0.15	0.04	5
4.10	0.23	0.02	5	0.74	0.09	5	0.36	0.02	5
4.67	0.25	0.01	5	0.81	0.05	5	0.36	0.03	5
5.46	0.15	0.01	5	0.66	0.05	5	0.30	0.04	5
20.48	0.30	0.04	5	2.38	0.11	5	0.56	0.03	5
25.53	0.43	0.02	5	2.41	0.15	5	0.61	0.02	5

Appendix G – RT-qPCR results and amplification efficiencies

Table G.1 Mean and SEM of the ΔC_t expression of the *xynA* and *amyM* genes in the production strains BWXP and BWAP at two different time points of growth t_1 and t_2. The C_t values were averaged from two technical replicates and the ΔC_t values correspond to the difference between the C_t of the gene of interest and the control gene, either *sdhA* or *fbaA*.

		ΔC_t Expression			
		t_1		t_2	
	Ctrl gene	mean	SEM	mean	SEM
BWXP, *xynA*	*sdhA*	77.58	45.10	139.90	65.13
	fbaA	39.19	23.30	41.91	17.18
BWAP, *amyM*	*sdhA*	42.80	17.66	83.11	13.60
	fbaA	20.62	0.36	28.83	3.24

Table G.2 Mean and SEM of the $\Delta\Delta C_t$ expression values of *xynA* in the strain BWXP, relative to the wild type strain 168 and using either *sdhA* or *fbaA* as the control genes.

		$\Delta\Delta C_t$ Expression			
		t_1		t_2	
	Ctrl gene	mean	SEM	mean	SEM
BWXP, *xynA*	*sdhA*	2823.00	961.10	815.10	219.10
	fbaA	1565.00	540.10	289.30	70.48

Table G.3 Mean and SEM of the $\Delta\Delta C_t$ expression values of several genes involved in the secretion machinery of the production strains BWXP and BWAP, relative to the wild type strain 168 and using *fbaA* as the control gene. The significance analyses correspond to unpaired *t*-tests with Welch's corrections for different standard deviations, assuming a Gaussian distribution.

P > 0.05; *: *P* ≤ 0.05; **: *P* ≤ 0.01; ***: *P* ≤ 0.001; ****: *P* ≤ 0.0001

	WT 168		**BWXP**				**BWAP**			
	Mean	SEM	Mean	SEM	P value vs WT 168	P value summary	Mean	SEM	P value vs WT 168	P value summary
secA	1.0090	0.0757	0.7425	0.0858	0.0597	ns	0.8175	0.0554	0.0922	ns
secY	1.0360	0.1569	2.0330	0.4923	0.1339	ns	2.0600	0.3731	0.0643	ns
secE	1.1080	0.2754	0.7000	0.1719	0.2641	ns	0.6325	0.0916	0.1834	ns
secG	1.0350	0.1556	1.2330	0.2289	0.5066	ns	1.1780	0.1628	0.5515	ns
secDF	1.0110	0.0877	0.7925	0.1135	0.1807	ns	0.6425	0.1408	0.0764	ns
yrbF	1.0070	0.0697	0.4525	0.0368	0.0013	**	0.3775	0.0720	0.0008	***
prsA	1.0280	0.1372	2.1630	0.3660	0.0464	*	2.7950	0.2294	0.0013	**
htrA	1.0010	0.0271	1.6880	0.3728	<0.0001	****	24.0600	0.7218	<0.0001	****
htrB	1.0210	0.1185	2.7780	0.6322	0.0667	ns	34.3000	2.0130	0.0005	***
htrC	1.0110	0.0840	0.9400	0.1418	0.6870	ns	0.6850	0.0352	0.0231	*
wprA	1.0900	0.2506	0.7625	0.1524	0.3152	ns	0.5175	0.0695	0.1031	ns

Table G.4 Mean and SEM of the $\Delta\Delta C_t$ expression values of several genes involved in the secretion machinery of the production strains BWXP and BWAP, relative to the wild type strain 168 and using *sdhA* as the control gene. The significance analyses correspond to unpaired *t*-tests with Welch's corrections for different standard deviations, assuming a Gaussian distribution.

P > 0.05; *: *P* ≤ 0.05; **: *P* ≤ 0.01; ***: *P* ≤ 0.001; ****: *P* ≤ 0.0001

	WT 168		**BWXP**				**BWAP**			
	Mean	SEM	Mean	SEM	P value vs WT 168	P value summary	Mean	SEM	P value vs WT 168	P value summary
secA	1.0067	0.0669	1.1650	0.0527	0.1150	ns	1.0350	0.0155	0.7051	ns
secY	1.0177	0.1091	2.6150	0.4464	0.0337	*	2.1250	0.2161	0.0080	**
secE	1.2590	0.4417	1.1425	0.4120	0.8535	ns	0.7025	0.1844	0.3093	ns
secG	1.1024	0.2704	2.1275	0.6252	0.2054	ns	1.5450	0.3354	0.3455	ns
secDF	1[#]	0[#]	1.0300	0.0800	0.7716[#]	ns	0.7900	0.3600	0.6638[#]	ns
yrbF	1.0501	0.1850	0.5300	0.0763	0.0602	ns	0.3550	0.1014	0.0240	*
prsA	1.1091	0.2770	3.3725	0.8994	0.0818	ns	3.4175	0.5717	0.0192	*
htrA	1.0102	0.0825	2.3975	0.3710	0.0303	*	28.4675	1.2864	0.0002	***
htrB	1.0020	0.0362	4.5125	1.1654	0.0570	ns	67.8525	1.7519	<0.0001	****
htrC	1.0000	0.0007	1.5275	0.0075	<0.0001	****	0.8050	0.1068	0.1654	ns
wprA	1.0243	0.1280	0.9700	0.0969	0.7476	ns	0.7000	0.0871	0.0873	ns

[#] Due to a technical error, only one biological replicate of the wild type 168 strain was used in the calculations.

Table G.5 Mean and SEM of the $\Delta\Delta C_t$ expression values of the *htrA*, *htrB*, *htrC* and *wprA* in the strains BCS234, BCS291, BCS292 and BCS293 (Table 4.1), respectively. These strains carry inducible overexpression constructs for each of the proteases genes and the samples were harvested at the end of exponential growth in the absence or presence of 0.2% xylose. The $\Delta\Delta C_t$ values are relative to the respective strain in the absence of xylose using either the *fbaA* or *sdhA* gene as control gene, as indicated. The significance analyses correspond to unpaired *t*-tests with Welch's corrections for different standard deviations, assuming a Gaussian distribution.

P > 0.05; *: *P* ≤ 0.05; **: *P* ≤ 0.01; ***: *P* ≤ 0.001; ****: *P* ≤ 0.0001

		- xylose		**+ xylose**		**Significance**	
	Ctrl gene	Mean	SEM	Mean	SEM	P value	P value summary
htrA	*fbaA*	1.907	0.938	28.310	9.456	0.0675	ns
	sdhA	1.968	0.979	29.250	10.220	0.0751	ns
htrB	*fbaA*	1.011	0.087	20.590	3.324	0.0097	**
	sdhA	1.008	0.075	24.480	3.348	0.0059	**
htrC	*fbaA*	1.080	0.235	16.700	1.926	0.0036	**
	sdhA	1.093	0.255	20.760	2.551	0.0043	**
wprA	*fbaA*	1.067	0.215	3.928	0.416	0.0025	**
	sdhA	1.154	0.332	4.865	0.751	0.0099	**

Table G.6 Mean and SEM of the $\Delta\Delta C_t$ expression values of the *secA*, *secA* (codon pair optimised), *secG* and *secG* (codon pair optimised) in the strains BRC31P1A, BRC31P1D, BRC31P2A (and BRC31P2C) and BRC31P2B (Table 5.1), respectively. Cell cultures were grown in 1.5 mL LB in FlowerPlates at 37 °C in a Biolector microfermentor with 95% humidity and vigorous agitation (800 rpm). The cultures were harvested at the end of exponential phase and total RNA extracted (Sectio 2.7.). The expression levels were quantified in the presence of 1 mM IPTG or 0.3% xylose, relative to the expression levels in the absence of inducer. The C_t values were averaged from two technical replicates, and the ΔC_t values were calculated using the *cheA* gene as control. The average of the ΔC_t values of two biological replicates are shown with the corresponding SEM. The significance analyses correspond to unpaired t-tests with Welch's corrections for different standard deviations, assuming a Gaussian distribution.

P > 0.05; *: *P* ≤ 0.05; **: *P* ≤ 0.01; ***: *P* ≤ 0.001; ****: *P* ≤ 0.0001

			- Inducer		**+ Inducer**		**Significance**	
Strain	Gene	Ctrl gene	Mean	SEM	Mean	SEM	P value	P value summary
BRC31P1	*secA*	*fbaA*	1.000	0.025	1.120	0.027	0.0214	*
		cheA	1.071	0.221	1.285	0.202	0.5018	ns
BRC31P1D	$secA_{CpO}$	*fbaA*	1.003	0.047	1.178	0.028	0.0254	*
		cheA	1.012	0.089	1.443	0.217	0.1404	ns
BRC31P2A	*secG*	*fbaA*	1.057	0.198	2.743	0.276	0.0034	**
		cheA	1.037	0.160	7.763	2.163	0.0525	ns
BRC31P2B	$secG_{CpO}$	*fbaA*	1.005	0.057	4.375	0.216	0.0003	***
		cheA	1.211	0.394	19.08	8.64	0.1305	ns
BRC31P2C	*secG*	*fbaA*	1.044	0.172	2.845	0.367	0.0100	**
		cheA	1.453	0.608	8.497	3.456	0.1330	ns

CpO – Codon pair optimised gene.

Table G.6 A dilution series of know template concentrations was used to establish a standard curve and assess the amplification reaction efficiency for each primer pair used in qPCR. The log of each known DNA concentration was plotted against the C_t value for that concentration. The template DNA was the genomic DNA from *Bacillus subtilis* 168 and each reaction was performed in duplicate. The slop of the curve was used to determine the amplification efficiency for each primer pair using the equation $Efficiency = 10^{(-1/slope)} - 1$. Every primer pair gave a good reaction efficiency between 90 and 110%.

Amplicon	**Primers**	**DNA ng**	**Ct_1**	**Ct_2**	**Average Ct**	**Slope (Ct *versus* *l*ogDNA)**	**Efficiency (%)**
xynA	362 +363	0.00	-	-	-	-3.21	105.02
		0.10	21.02	20.55	20.79		
		0.50	18.33	18.30	18.32		
		1.00	17.34	17.30	17.32		
		3.00	15.93	15.83	15.88		
		6.00	15.10	15.11	15.11		
amyM	368+369	0.00	-	-	-	-3.40	96.77
		0.10	12.93	12.33	12.63		
		0.50	10.29	10.05	10.17		
		1.00	9.36	9.21	9.28		
		3.00	7.55	7.55	7.55		
		6.00	6.63	6.53	6.58		
secA	370+371	0.00	-	-	-	-3.25	103.15
		0.10	20.61	20.41	20.51		
		0.50	18.21	18.05	18.13		
		1.00	17.13	16.97	17.05		
		3.00	15.53	15.41	15.47		
		6.00	14.94	14.75	14.85		
secG	378+379	0.00	-	-	-	-3.29	101.52
		0.10	20.50	20.11	20.31		
		0.50	18.02	17.97	18.00		

Table G.6 A dilution series of know template concentrations was used to establish a standard curve and assess the amplification reaction efficiency for each primer pair used in qPCR. The log of each known DNA concentration was plotted against the C_t value for that concentration. The template DNA was the genomic DNA from *Bacillus subtilis* 168 and each reaction was performed in duplicate. The slop of the curve was used to determine the amplification efficiency for each primer pair using the equation $Efficiency = 10^{(-1/slope)} - 1$. Every primer pair gave a good reaction efficiency between 90 and 110%.

Amplicon	**Primers**	**DNA ng**	**Ct_1**	**Ct_2**	**Average Ct**	**Slope (Ct *versus* *l*ogDNA)**	**Efficiency (%)**
		1.00	17.10	16.86	16.98		
		3.00	15.62	15.29	15.45		
		6.00	14.56	14.31	14.44		
secY	386 +387	0.00	-	-	-	-3.52	92.36
		0.10	18.95	18.52	18.74		
		0.50	16.16	16.12	16.14		
		1.00	15.10	14.81	14.96		
		3.00	13.48	13.29	13.38		
		4.00	13.31	13.07	13.19		
secE	390 + 391	0.00	-	-	-	-3.50	93.08
		0.10	18.87	18.75	18.81		
		0.50	16.22	16.14	16.18		
		1.00	15.22	14.77	14.99		
		3.00	13.81	13.35	13.58		
		4.00	13.22	13.19	13.20		
secDF	394 + 395	0.00	-	-	-	-3.53	92.03
		0.10	18.31	18.01	18.16		
		0.50	15.44	15.33	15.38		
		1.00	14.46	14.32	14.39		
		3.00	12.94	12.76	12.85		
		4.00	12.56	12.39	12.48		

Table G.6 A dilution series of know template concentrations was used to establish a standard curve and assess the amplification reaction efficiency for each primer pair used in qPCR. The log of each known DNA concentration was plotted against the C_t value for that concentration. The template DNA was the genomic DNA from *Bacillus subtilis* 168 and each reaction was performed in duplicate. The slop of the curve was used to determine the amplification efficiency for each primer pair using the equation $Efficiency = 10^{(-1/slope)} - 1$. Every primer pair gave a good reaction efficiency between 90 and 110%.

Amplicon	**Primers**	**DNA ng**	**Ct_1**	**Ct_2**	**Average Ct**	**Slope (Ct *versus* *l*ogDNA)**	**Efficiency (%)**
yrbF	398 + 399	0.00	-	-	-	-3.48	93.73
		0.10	18.39	18.36	18.37		
		0.50	15.78	15.80	15.79		
		1.00	14.97	14.59	14.78		
		3.00	13.41	13.25	13.33		
		4.00	12.98	12.73	12.85		
prsA	402 + 403	0.00	-	-	-	-3.17	106.93
		0.10	17.88	17.56	17.72		
		0.50	15.20	15.02	15.11		
		1.00	14.33	14.12	14.23		
		3.00	12.86	12.81	12.84		
		4.00	12.59	12.43	12.51		
htrA	570 + 571	25.00	12.79	12.82	12.80	-3.37	98.54
		5.00	15.05	15.09	15.07		
		1.00	17.31	17.40	17.36		
		0.25	19.71	19.86	19.79		
		0.04	22.25	22.12	22.19		
		0.01	24.48	24.49	24.48		
htrB	572 +573	25.00	13.29	13.26	13.28	-3.42	97.15
		5.00	15.40	15.42	15.41		
		1.00	17.88	17.83	17.86		

Table G.6 A dilution series of know template concentrations was used to establish a standard curve and assess the amplification reaction efficiency for each primer pair used in qPCR. The log of each known DNA concentration was plotted against the C_t value for that concentration. The template DNA was the genomic DNA from *Bacillus subtilis* 168 and each reaction was performed in duplicate. The slop of the curve was used to determine the amplification efficiency for each primer pair using the equation $Efficiency = 10^{(-1/slope)} - 1$. Every primer pair gave a good reaction efficiency between 90 and 110%.

Amplicon	**Primers**	**DNA ng**	**Ct_1**	**Ct_2**	**Average Ct**	**Slope (Ct *versus* *l*ogDNA)**	**Efficiency (%)**
		0.25	20.19	20.21	20.20		
		0.04	22.74	22.84	22.79		
		0.01	25.11	24.98	25.04		
htrC	574 + 575	25.00	12.37	12.46	12.41	-3.39	97.88
		5.00	14.59	14.63	14.61		
		1.00	16.92	17.05	16.98		
		0.25	19.43	19.45	19.44		
		0.04	21.81	21.76	21.79		
		0.01	24.22	24.10	24.16		
wprA	576 + 577	25.00	13.17	13.05	13.11	-3.42	97.14
		5.00	15.12	15.25	15.18		
		1.00	17.57	17.51	17.54		
		0.25	20.05	19.92	19.98		
		0.04	22.63	22.38	22.51		
		0.01	24.84	24.94	24.89		
fbaA	14528 + 14529	0.00	-	-	-	-3.20	105.49
		0.10	20.52	20.28	20.40		
		0.50	18.14	18.04	18.09		
		1.00	17.26	16.98	17.12		
		3.00	15.81	15.45	15.63		
		6.00	15.14	14.66	14.90		

Table G.6 A dilution series of know template concentrations was used to establish a standard curve and assess the amplification reaction efficiency for each primer pair used in qPCR. The log of each known DNA concentration was plotted against the C_t value for that concentration. The template DNA was the genomic DNA from *Bacillus subtilis* 168 and each reaction was performed in duplicate. The slop of the curve was used to determine the amplification efficiency for each primer pair using the equation $Efficiency = 10^{(-1/slope)} - 1$. Every primer pair gave a good reaction efficiency between 90 and 110%.

Amplicon	**Primers**	**DNA ng**	**Ct_1**	**Ct_2**	**Average Ct**	**Slope (Ct *versus* *l*ogDNA)**	**Efficiency (%)**
sdhA	14536 + 14537	0.00	-	-	-	-3.25	103.05
		0.10	19.90	19.69	19.80		
		0.50	17.23	17.24	17.23		
		1.00	16.51	16.29	16.40		
		3.00	14.92	14.81	14.86		
		6.00	14.54	13.90	14.22		
cheA	14526 + 14527	0	20.61	20.37	20.49	-3.24	103.76
		0.1	18.19	17.95	18.07		
		0.5	17.32	-	17.32		
		1	15.73	15.43	15.58		
		3	15.05	14.66	14.86		
		6	20.61	20.37	20.49		
		0	18.19	17.95	18.07		

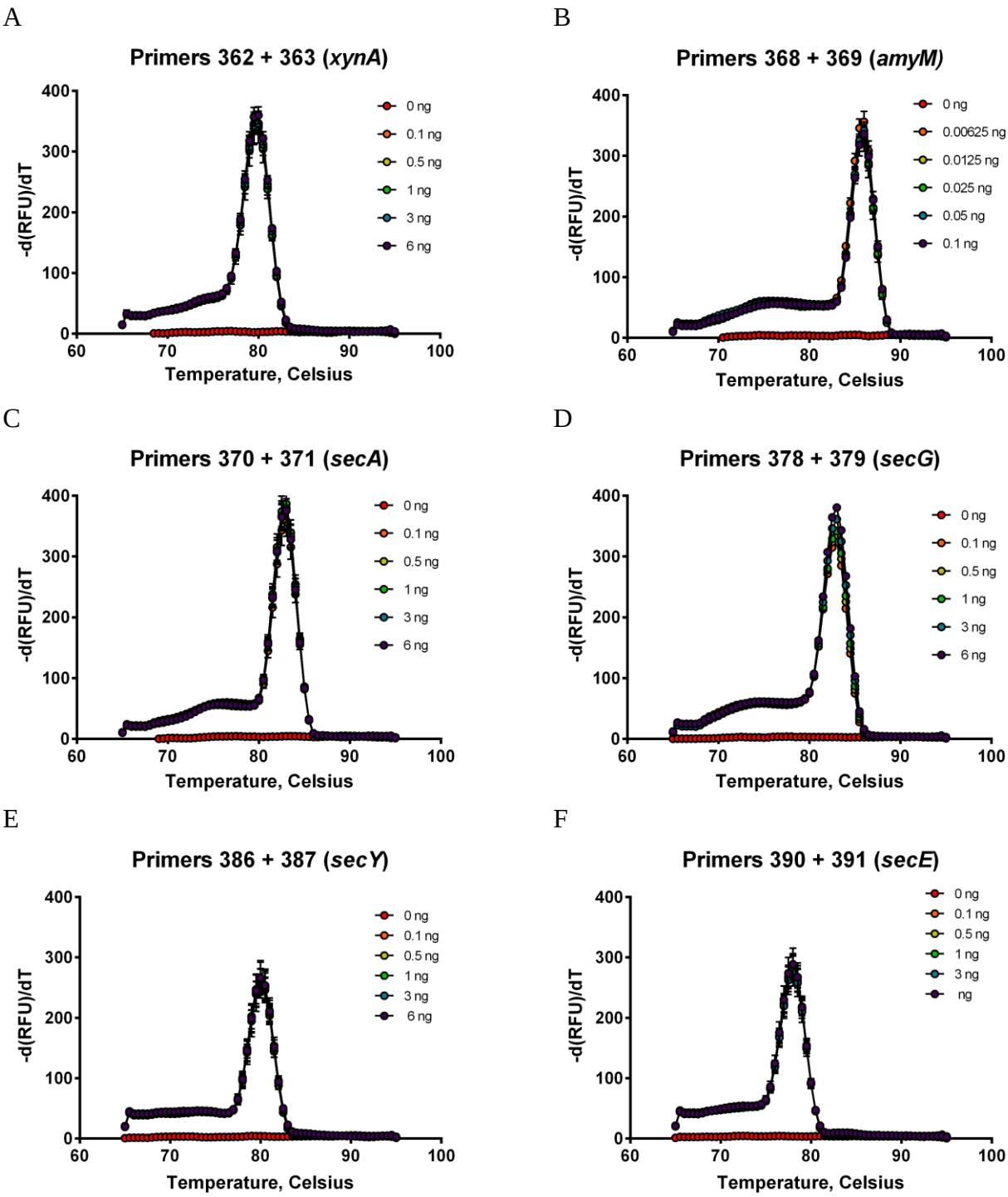


Figure G.1 Melting curves, or dissociation curves, of each primer pair obtained with a dilution series of template DNA. The curves show the change in fluorescence observed when double-stranded DNA (dsDNA) with incorporated dye molecules dissociates into single-stranded DNA (ssDNA) as the temperature increases. The ratio between the change in fluorescence and change in temperature (-dRFU/dT) was plotted against temperature to confirm a single peak corresponding to one single amplification product for each primer pair. The melting curves for the primer pairs corresponding to the amplicons *xynA* (A), *amyM* (B), *secA* (C), *secG* (D), *secY* (E), *secE* (F), *secDF* (G), *yrbF* (H) and *prsA* (I), *htrA* (J), *htrB* (K), *htrC* (L), *wprA* (M), *fbaA* (N), *sdaA* (O) and *cheA* (P) are presented.

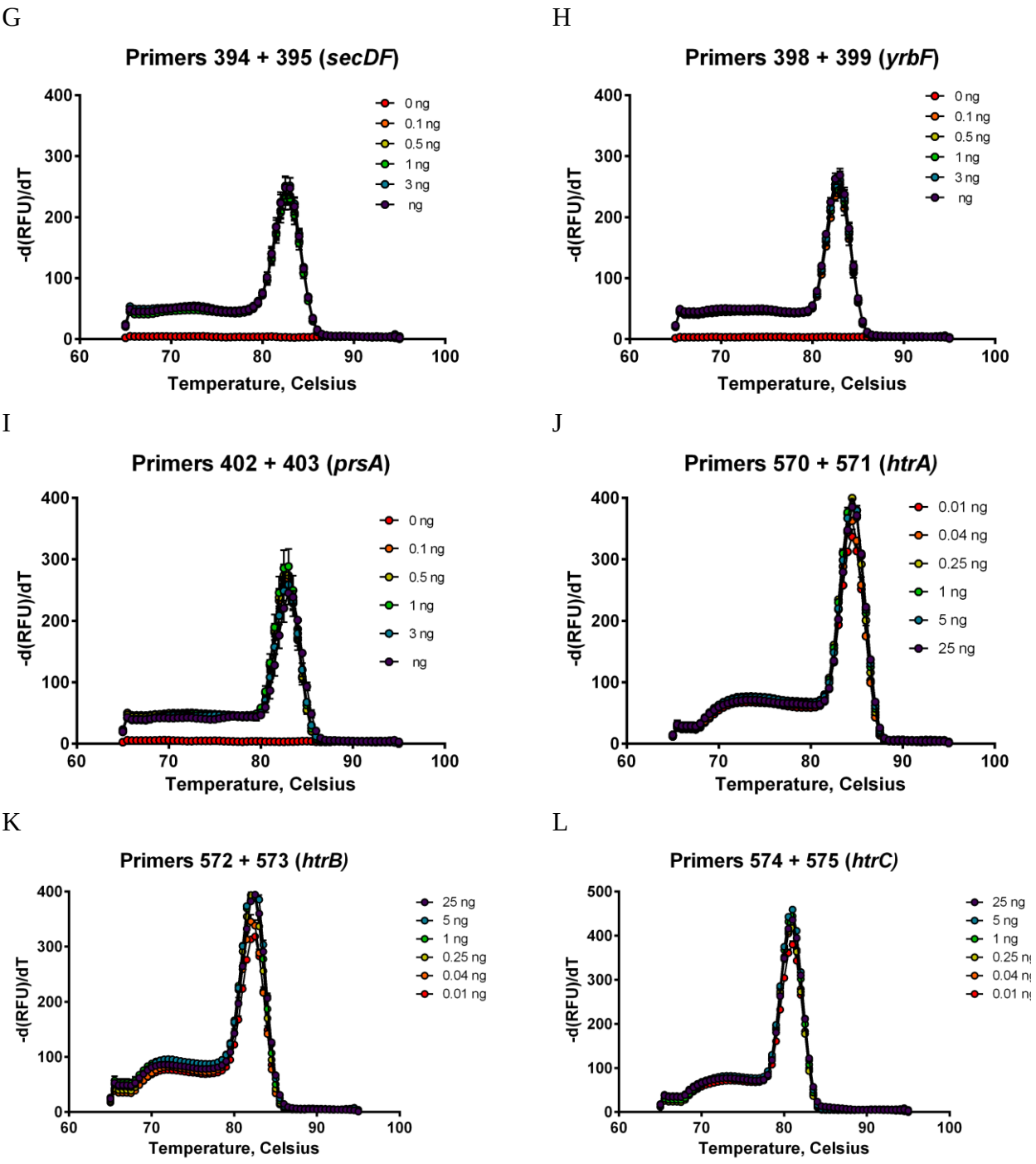


Figure G. 1 Melting curves, or dissociation curves, of each primer pair obtained with a dilution series of template DNA. The curves show the change in fluorescence observed when double-stranded DNA (dsDNA) with incorporated dye molecules dissociates into single-stranded DNA (ssDNA) as the temperature increases. The ratio between the change in fluorescence and change in temperature (-dRFU/dT) was plotted against temperature to confirm a single peak corresponding to one single amplification product for each primer pair. The melting curves for the primer pairs corresponding to the amplicons *xynA* (A), *amyM* (B), *secA* (C), *secG* (D), *secY* (E), *secE* (F), *secDF* (G), *yrbF* (H) and *prsA* (I), *htrA* (J), *htrB* (K), *htrC* (L), *wprA* (M), *fbaA* (N), *sdaA* (O) and *cheA* (P) are presented.

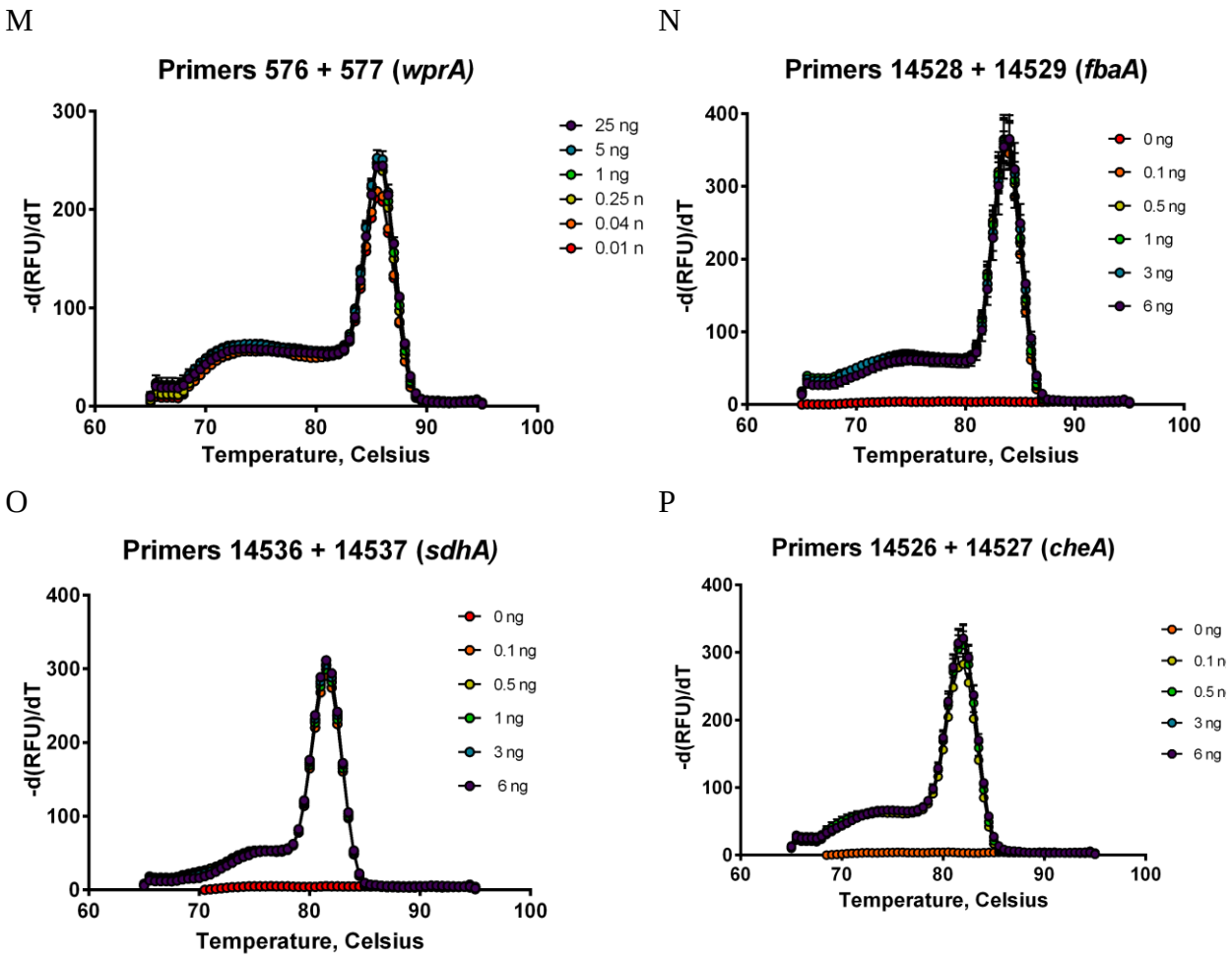


Figure G.1 Melting curves, or dissociation curves, of each primer pair obtained with a dilution series of template DNA. The curves show the change in fluorescence observed when double-stranded DNA (dsDNA) with incorporated dye molecules dissociates into single-stranded DNA (ssDNA) as the temperature increases. The ratio between the change in fluorescence and change in temperature (-dRFU/dT) was plotted against temperature to confirm a single peak corresponding to one single amplification product for each primer pair. The melting curves for the primer pairs corresponding to the amplicons xynA (A), amyM (B), secA (C), secG (D), secY (E), secE (F), secDF (G), yrbF (H) and prsA (I), htrA (J), htrB (K), htrC (L), wprA (M), fbaA (N), sdaA (O) and cheA (P) are presented.

Appendix H – Generation of XynA anti-serum

An animal immunisation was requested from the Comparative Biology Centre of Newcastle University. One mililitre of XynA crude enzyme sample supplied by DSM was dialysed overnight at 4 °C in 2 L of 100 mM Phosphate Buffer. Size exclusion chromatography was used to purify the XynA protein from the crude enzyme sample. After equilibrating the column with 100 mM Phosphate Buffer overnight (0.5 mL/min), the crude enzyme sample was diluted 1:1 with the same buffer and a total of 24 mL was loaded onto the column at a rate of 0.5 mL/min. 1 mL fractions were collected. A peak was detected from fraction 11 until fraction 29. These fractions were analysed via SDS-PAGE, according to Section 2.8. (Figure H.1).

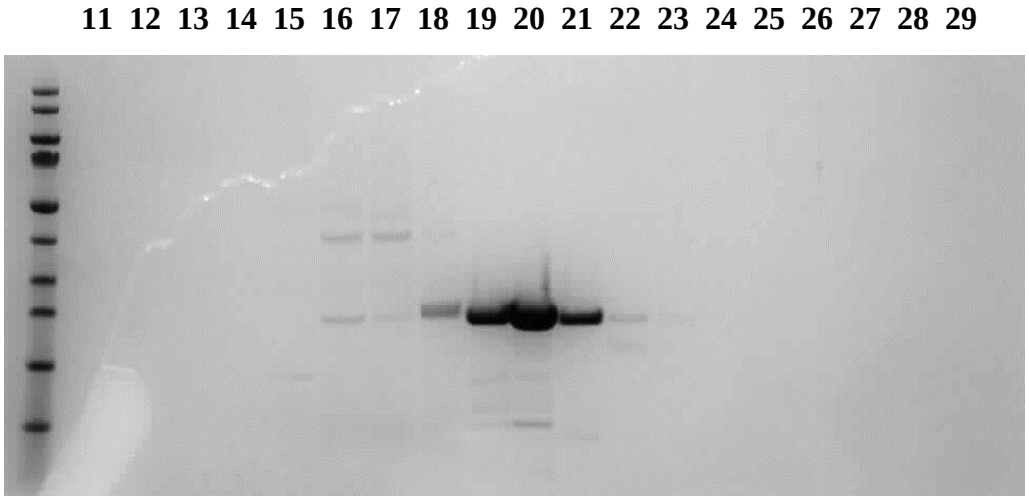


Figure H.1 SDS-PAGE analysis of size exclusion chromatography sample fractions for purification of XynA (23 kDa) from a crude enzyme sample.

The protein concentration of fraction 20 was quantified to be 0.84 g/L with the 2D-quant kit from GE Healthcare, according to the manufacturer's instructions. This sample was used to perform the rabbit immunisations according to the immunisation schedule described on Table H.1.

Table H. 1 Rabbit immunisation schedule used to generate blood serum anti XynA.

Day 0	**Pre-immunisation blood sample** from the marginal ear vein (< 5 mL)
Day(+)1	**Primary immunisation** Immunise with 1 mL of antigen/adjuvant mix made up as follows: • 125 µL of CFA • 200 µg of XynA Make up to 1 mL with PBS
Day 15	**First Boost immunisation** Immunise with 1 mL of antigen/adjuvant mix made up as follows: • 500 µL of IFA • 125 ug of XynA • Make up to 1 mL with PBS
Day 30	**Second Boost immunisation** Immunise with 1 mL of antigen/adjuvant mix made up as follows: • 500 µL of IFA • 125 ug of XynA • Make up to 1 mL with PBS
Day 40	**First Post-immunisation Blood Sample** from the marginal ear vein (~10 mL) Test the antibody response and decide whether to continue with further boosts.
Day 45	**Third Boost immunisation** Immunise with 1 mL of antigen/adjuvant mix made up as follows: • 500 µL of IFA • 125 ug of XynA • Make up to 1 mL with PBS
Day 55	**Second Post-immunisation Blood Sample** from the marginal ear vein (~10 mL)
Day 60	**Fourth (Final) Boost immunisation** Immunise with 1 mL of antigen/adjuvant mix made up as follows: • 500 µL of IFA • 125 ug of XynA • Make up to 1 mL with PBS
Day 70	**Third (Final) Post-immunisation Blood Sample** – the animal will be exsanguinated under terminal general anaesthesia.

CFA - Complete Freund's Adjuvant (CFA); IFA – Incomplete Freund's Adjuvant

The blood serum of the final exsanguination was separated from the rest of the blood components by centrifugation (10,000 xg, 10 min, RT). The serum response to XynA was tested via Western Blot against the supernatants of strains BWXC, BWXP (Table 5.1) and the original crude enzyme sample used for immunisation (Figure H.2). It was concluded that the serum had a good response to XynA and can be used for immunodetection of this protein using a 1:1000 serum dilution.

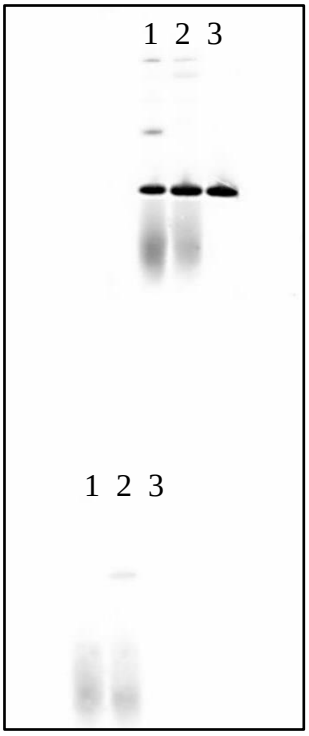


Figure H.2 Western blot analysis of the response of pre and post-immusitation rabitt serum (1:1000) to XynA in the supernatant of strains strains (1) BWXC, (2) BWXP (Table 5.1) and (3) the original crude enzyme sample used for immunisation.

Appendix I – Pulse-chase experiments

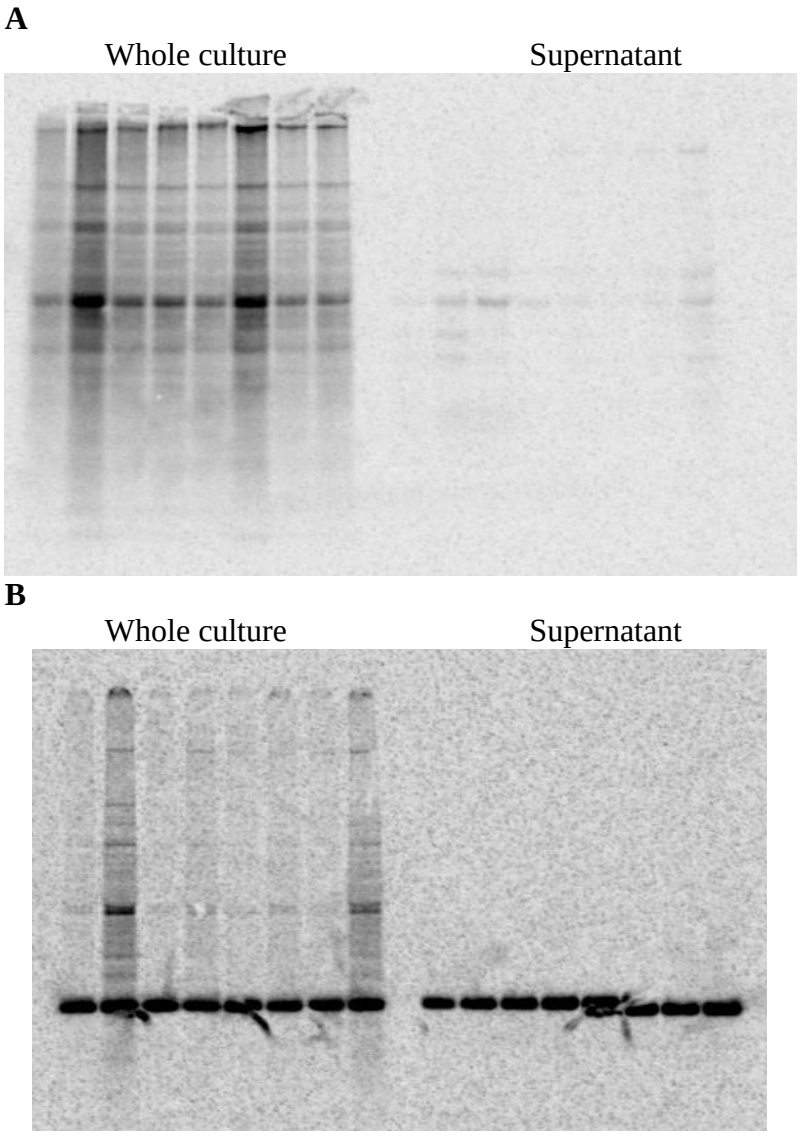


Figure I.1 Phosphor imaging of proteins labelled with [^{35}S] methionine during pulse-chase experiments, and electrophoresed on SDS-PAGE gels after immunoprecipitation of whole culture and supernatant fractions with XynA anti-serum (Section 2.9.1). The whole culture and supernatant fractions of the (A) *B. subtilis* wild type strain 168 and (B) BWXP strain expressing high levels of XynA from a high copy number plasmid are shown.

A **B**

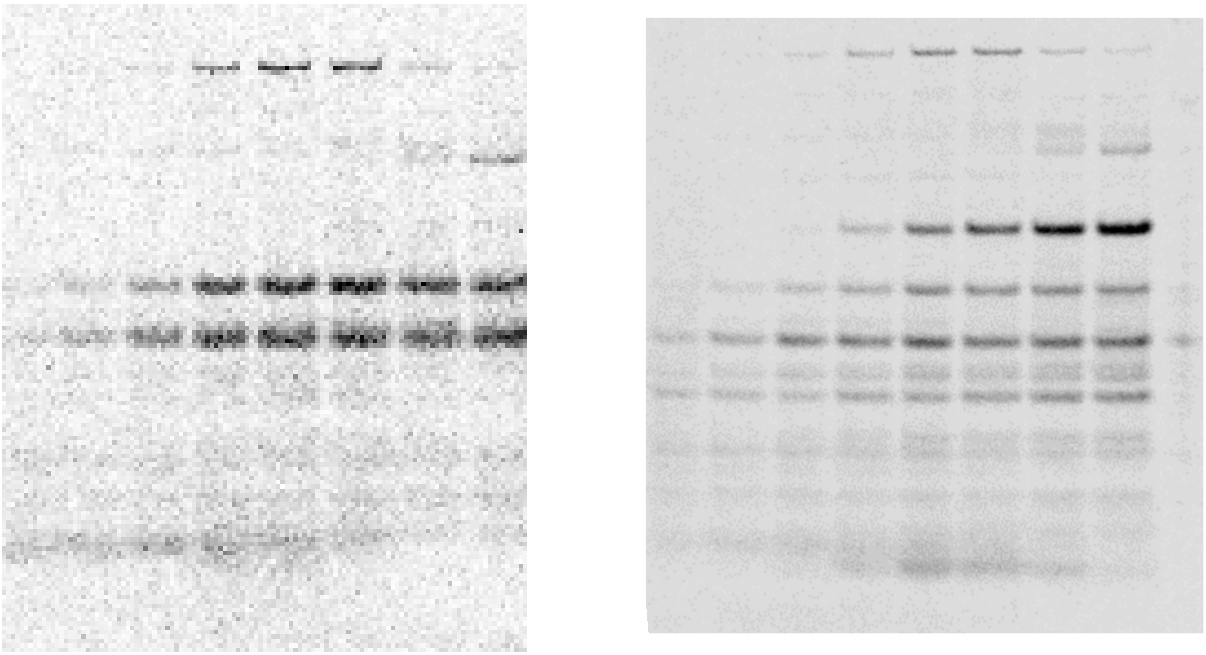
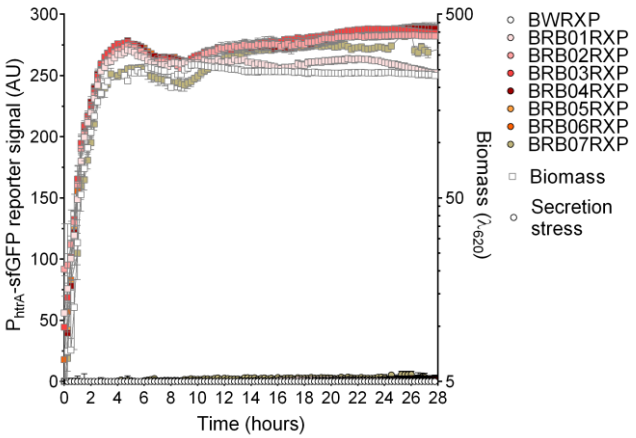


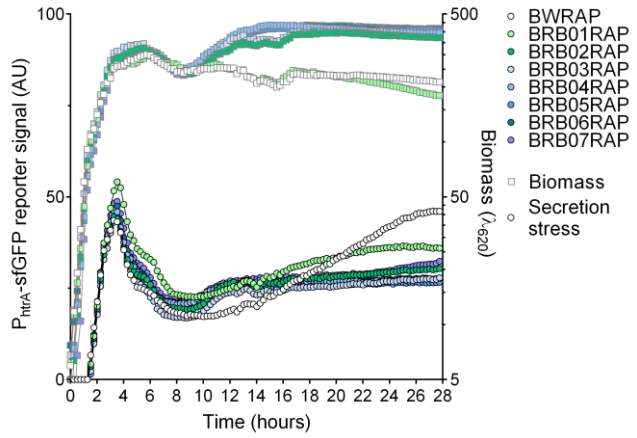
Figure I.1 Phosphor imaging of secreted proteins labelled with [^{35}S] methionine during pulse-chase experiments, and electrophoresed on SDS-PAGE gels (Section 2.9.2). The supernatant of the (A) *B. subtilis* wild type strain 168 and (B) BWAP strain expressing high levels of AmyM from a high copy number plasmid are shown.

Appendix J – Growth and secretion stress monitoring in the Biolector®

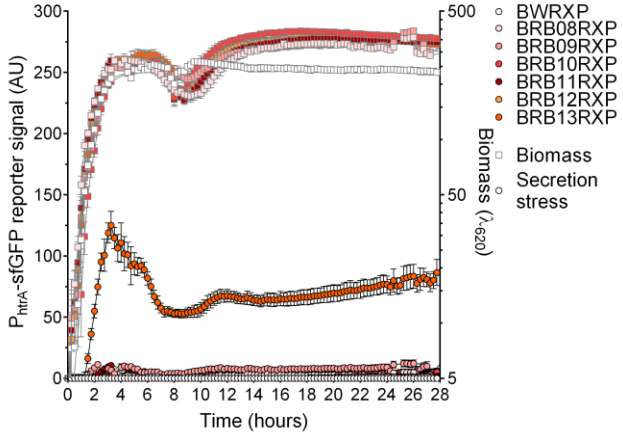
A1



B1



A2



B2

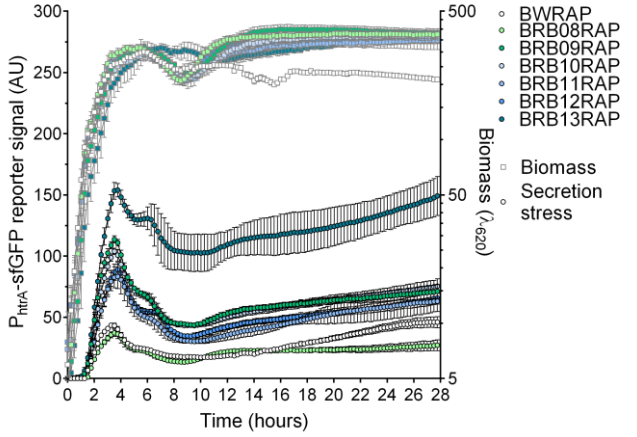
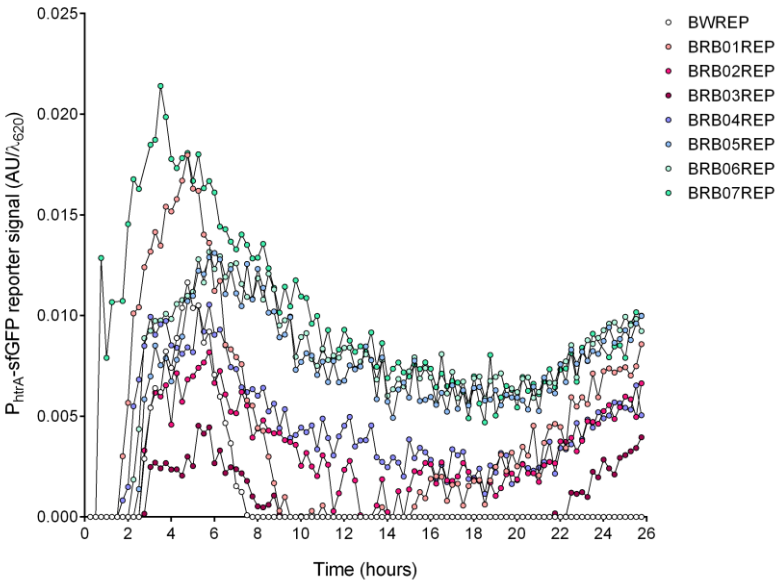


Figure J.1 Protein secretion stress response upon overexpression of (A) XynA and (B) AmyM in strains lacking multiple extracytoplasmic proteases, measured by live detection of the reporter P$_{htrA}$-*sfGFP* fluorescence using the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 1.5 mL LB in FlowerPlate® 48-well MTPs and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The plotted fluorescence signal was blanked with the detected signal for the wild type strain 168. The average of three biological replicates per strain, grown in triplicate, was plotted with the corresponding SEM

A



B

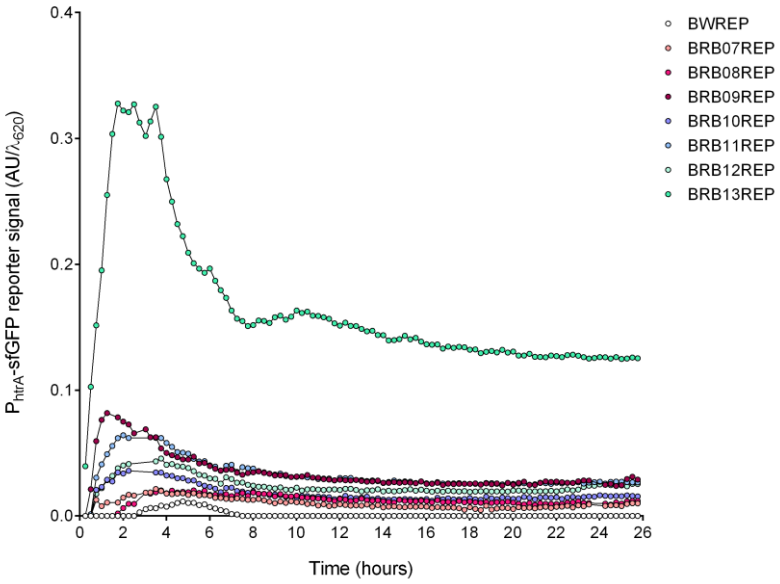


Figure J.2 Protein secretion stress in strains lacking multiple extracytoplasmic proteases and carrying the empty plasmid pCS74, measured by live detection of the reporter P_{htrA}-sfGFP fluorescence using the BioLector® bench-top microfermentation system (m2p-labs). The cultures were grown in 1.5 mL LB in FlowerPlate® 48-well MTPs and incubated at 37 °C with 95% humidity and vigorous agitation (800 rpm). The plotted fluorescence signal was blanked with the detected signal for the wild type strain 168. The average of three biological replicates per strain, grown in triplicate.

Appendix H – Synthetic translocase DNA sequences

Table H.1 Codon pair optimised genes used in Chapter 5.

Gene	Codon pair optimised sequence	Match with wild type (%)
secA	ATGCTTGGCATTTTAAACAAAATGTTCGACCCAACAAAACGTACTTTAAACCGTTATGAGAAAATCGCAAACGATATCGATGCGATCCGCGGAGACTA CGAAAACCTTTCTGATGATGCTTTAAAACACAAAACGATCGAATTCAAAGAACGTCTTGAAAAAGGTGCAACTACTGATGATTTGCTTGTTGAAGCAT TCGCTGTTGTTCGTGAAGCAAGCCGCCGTGTAACTGGTATGTTCCCATTCAAAGTACAGCTGATGGGCGGTGTGGCGCTTCATGACGGAAACATCGCT GAAATGAAAACAGGTGAAGGAAAAACGCTGACAAGCACGCTTCCTGTATACTTAAACGCGCTGACTGGTAAAGGCGTTCACGTTGTAACTGTAAACGA ATACCTTGCTTCTCGTGACGCTGAACAAATGGGTAAAATCTTCGAATTCCTTGGTTTGACTGTAGGTTTGAACCTGAACAGCATGAGTAAAGATGAAA AACGTGAAGCTTACGCTGCTGATATCACTTACAGTACAAACAACGAGCTTGGCTTCGATTATTTGCGTGACAACATGGTGCTTTACAAAGAGCAAATG GTGCAGCGTCCATTGCACTTCGCTGTTATCGATGAAGTTGACAGCATTTTGATTGATGAAGCTCGTACACCGCTGATCATTTCTGGTCAAGCTGCAAA AAGCACAAAGCTTTACGTGCAGGCAAACGCTTTCGTTCGTACGCTGAAAGCTGAAAAAGACTATACGTATGACATCAAAACAAAAGCTGTTCAATTAA CTGAAGAAGGCATGACAAAAGCAGAAAAAGCATTCGGTATCGACAACCTATTCGATGTAAAACACGTTGCGCTGAACCACCACATCAACCAAGCTTTA AAAGCGCACGTTGCAATGCAAAAAGATGTGGATTACGTTGTTGAAGACGGACAAGTCGTTATCGTTGATTCTTTCACTGGACGTTTAATGAAAGGACG CCGTTACTCTGAAGGTCTTCATCAGGCGATTGAAGCAAAAGAAGGTCTTGAAATCCAAAACGAATCAATGACTTTGGCAACCATTACATTCCAAAACT ACTTCCGTATGTACGAAAAGCTTGCTGGTATGACTGGAACTGCAAAAACGGAAGAAGAAGAATTCCGTAATATTTACAACATGCAAGTTGTGACGATT CCGACAAACCGTCCTGTTGTGCGTGATGACCGTCCTGATTTGATTTACCGTACAATGGAAGGAAAATTCAAAGCTGTTGCTGAAGATGTTGCACAGCG TTACATGACTGGACAGCCTGTGCTTGTCGGAACTGTTGCTGTTGAAACTTCTGAGCTGATTTCTAAACTATTGAAAAACAAAGGCATTCCTCATCAAG TGCTGAACGCTAAAAACCATGAGCGTGAAGCTCAAATCATCGAAGAAGCAGGACAAAAAGGCGCTGTAACGATTGCAACAAACATGGCAGGCCGCGGA ACTGATATCAAACTAGGTGAAGGCGTAAAAGAGCTTGGCGGACTTGCTGTTGTAGGTACTGAGCGTCATGAAAGCCGCCGCATTGATAACCAGCTTCG CGGACGCAGCGGACGTCAAGGTGATCCAGGTATCACTCAATTCTACCTTTCCATGGAGGATGAGCTGATGCGTCGTTTCGGCGCTGAACGTACGATGG CGATGCTTGACCGTTTCGGCATGGATGACAGCACGCCAATCCAATCTAAAATGGTCAGCCGCGCTGTAGAATCTTCTCAAAAACGTGTTGAAGGAAAC AACTTCGACAGCCGCAAACAGCTTCTTCAATACGATGATGTGCTTCGCCAGCAGCGTGAAGTGATTTACAAACAGCGTTTCGAAGTGATCGACAGTGA AAACCTTCGTGAAATCGTTGAAAACATGATCAAATCAAGCCTTGAGCGTGCGATCGCTGCTTACACACCGCGTGAAGAACTTCCTGAAGAATGGAAGC TTGACGGACTTGTTGATTTAATCAACACAACTTACCTTGATGAAGGTGCTTTAGAAAAATCTGATATTTTCGGAAAAGAGCCTGATGAGATGCTTGAG CTGATCATGGACCGTATCATCACAAAATACAACGAAAAAGAAGAACAATTCGGCAAAGAACAAATGCGTGAATTTGAAAAAGTGATCGTACTTCGTGC TGTAGACAGCAAATGGATGGATCACATTGATGCGATGGACCAATTGCGTCAAGGCATTCACCTTCGTGCTTACGCTCAAACAAACCCGCTTCGTGAAT ACCAAATGGAAGGCTTCGCAATGTTCGAGCACATGATCGAAAGCATTGAAGATGAAGTGGCAAAATTCGTTATGAAGGCTGAAATCGAAAACAACCTT GAACGTGAAGAAGTGGTTCAAGGCCAAACAACTGCTCATCAGCCTCAAGAAGGCGATGACAACAAGAAAGCGAAGAAAGCTCCTGTACGCAAAGTTGT TGATATCGGACGCAACGCTCCTTGCCATTGCGGAAGCGGCAAAAAGTACAAAAACTGCTGCGGCCGTACTGAA	81

secY	ATGTTCAAAACGATCAGCAACTTCATGCGTGTATCTGATATCCGCAACAAAATCATCTTCACATTGCTCATGCTGATTGTATTCCGTAT CGGTGCTTTCATTCCTGTTCCTTACGTGAACGCTGAAGCGCTGCAAGCTCAATCTCAAATGGGTGTATTCGATCTTTTAAACACATTCG GCGGCGGCGCTCTATACCAATTCAGCATTTTCGCAATGGGCATCACTCCATACATCACTGCTTCTATCATCATTCAGCTATTGCAAATG GATGTTGTGCCGAAATTCACTGAATGGAGCAAACAAGGTGAAGTGGGACGCAGAAAGCTTGCTCAATTCACTCGTTACTTCACGATTGT GCTCGGATTCATTCAAGCTCTTGGTATGTCATACGGCTTCAACAACCTTGCAAACGGTATGCTGATCGAAAAATCAGGTGTATCAACTT ACTTAATCATCGCTCTTGTATTAACTGGCGGAACTGCATTCCTGATGTGGCTCGGTGAACAAATCACTTCTCACGGCGTTGGAAACGGA ATCAGCATCATCATCTTCGCAGGCATCGTTTCTTCTATCCCAAAAACGATCGGACAAATTTACGAAACTCAATTCGTCGGAAGCAACGA CCAATTGTTCATTCACATTGTGAAAGTAGCTTTGCTTGTTATCGCAATCCTTGCTGTTATCGTAGGTGTGATCTTCATTCAACAAGCTG TTCGCAAAATCGCTATTCAATACGCAAAAGGTACTGGCCGTTCTCCTGCTGGCGGCGGACAATCAACTCACCTTCCATTGAAAGTAAAC CCAGCTGGTGTTATCCCAGTTATCTTCGCTGTTGCTTTCCTGATCACTCCAAGAACGATTGCAAGCTTCTTCGGAACAAACGATGTGAC AAAATGGATTCAAAACAACTTCGACAACACTCACCCTGTTGGTATGGCGATTTACGTGGCGCTGATCATCGCTTTCACTTACTTCTACG CATTCGTTCAAGTGAACCCTGAACAAATGGCTGATAACCTAAAGAAACAAGGCGGATATATTCCAGGTGTTCGTCCTGGAAAAATGACT CAAGACCGCATTACAAGCATCCTTTACCGCCTGACGTTCGTTGGAAGCATTTTCTTGGCTGTTATCAGCATCCTTCCGATTTTCTTCAT CCAATTCGCTGGTCTTCCTCAATCTGCTCAAATCGGCGGAACTTCCTTGCTCATTGTTGTCGGCGTTGCGCTTGAAACGATGAAGCAGC TTGAAAGCCAGCTTGTAAAACGCAACTACCGCGGATTCATGAAAAAC	78
secE	ATGCGTATCATGAAATTCTTCAAAGATGTAGGTAAAGAAATGAAGAAAGTATCTTGGCCTAAAGGAAAAGAACTGACTCGTTATACGAT CACTGTTATCAGCACTGTGATTTTCTTCGTTATCTTCTTCGCACTTCTTGACACTGGCATTTCTCAATTAATCCGTTTAATCGTTGAA	81
secG	ATGCACGCTGTGCTGATCACTCTTCTTGTTATCGTATCTATCGCTTTAATCATCGTTGTGCTGCTTCAATCTTCTAAATCAGCTGGTCT TTCTGGTGCAATCAGCGGCGGCGCTGAACAATTATTCGGAAAACAAAAAGCACGCGGACTTGATTTGATCCTTCACCGTATCACTGTTG TGCTTGCTGTATTGTTCTTCGTTTTAACGATCGCTCTTGCTTACATCCTT	72
secDF	ATGAAAAAGGGCCGTTTGATTGCATTCTTCCTGTTTGTGCTTTTAATCGGAACTGGTCTTGGCTACTTCACAAAACCTGCTGCAAACAA CATCACTTTAGGTCTTGATCTTCAAGGCGGATTTGAAGTACTATATGACGTTCAGCCTGTGAAGAAAGGCGACAAAATCACAAAAGATG TGCTTGTATCAACTGTTGAAGCTTTAAACCGTCGTGCAAACGTACTAGGTGTCAGTGAACCAAACATCCAAATCGAAGGAAACAACCGT ATCCGTGTTCAATTGGCAGGCGTAACAAACCAAAACCGTGCTCGTGAAATTTTGGCAACTGAAGCACAGCTAAGCTTCCGTGATGCAAA CGACAAAGAACTATTAAACGGCGCTGATCTTGTTGAAAACGGAGCGAAACAAACGTATGACAGCACGACAAATGAACCGATTGTGACGA TTAAGCTGAAAGACGCTGACAAATTCGGAGAAGTGACAAAGAAAGTGATGAAAATGGCTCCAAACAACCAATTAGTTATTTGGCTTGAT TATGACAAAGGCGATTCTTTCAAGAAAGAAGTTCAAAAAGAGCATCCAAAATTCGTTTCTGCTCCAAACGTATCTCAAGAGCTGAACAC AACTGATGTGAAAATCGAGGGACACTTCACTGCTCAAGAAGCAAAAGATTTGGCAAGCATTTTAAACGCTGGAGCGCTTCCTGTGAAGC TGACTGAAAAATACTCTACAAGCGTAGGTGCTCAATTCGGACAACAAGCGCTTCACGATACTGTATTCGCCGGAATCGTCGGCATCGCA ATCATCTTCTTGTTCATGCTGTTCTACTACCGTCTTCCTGGTTTGATCGCTGTTATCACTCTTTCTGTATACATCTACATCACGCTTCA AATCTTCGACTGGATGAACGCAGTGCTGACGCTTCCTGGTATCGCTGCTTTAATCCTTGGCGTTGGTATGGCTGTTGATGCCAACATCA TCACTTACGAACGTATCAAAGAAGAATTGAAGCTTGGAAAATCTGTTCGTTCTGCTTTCCGCAGCGGAAACCGCCGTTCTTTCGCAACC	77

	ATTTTCGATGCAAACATCACAACAATCATCGCTGCTGTTGTGCTGTTCATCTTCGGAACAAGCAGTGTGAAAGGCTTCGCAACAATGCT GATCCTTTCTATCCTTACTTCTTTCATTACAGCTGTGTTTTTAAGCCGCTTCCTTCTTGCTTTGCTTGTTGAATCCCGCTGGCTTGACC GCAAGAAAGGCTGGTTCGGTGTAAACAAAAAGCACATCATGGATATTCAAGATACTGATGAAAACACTGAGCCTCATACGCCATTCCAA AAATGGGATTTCACAAGCAAACGCAAATACTTCTTCATCTTCTCTTCTGCTGTAACTGTTGCCGGAATCATCATTTTGCTTGTATTCCG TTTAAACCTTGGCATTGATTTCGCTTCTGGTGCTCGTATCGAAGTTCAATCTGACCACAAGCTGACTACTGAACAAGTAGAAAAAGACT TCGAAAGCCTTGGCATGGATCCTGATACTGTTGTTCTTTCTGGTGAAAAATCAAACATCGGTGTAGCTCGTTTCGTTGGTGTTCCTGAT AAAGAAACAATCGCAAAAGTGAAAACATACTTCAAAGACAAATACGGTTCTGATCCAAACGTTTCTACTGTTTCTCCAACTGTTGGTAA AGAGCTTGCTCGTAATGCGCTTTACGCTGTTGCAATCGCTTCTATCGGCATTATCATTTACGTGTCTATCCGTTTCGAATACAAAATGG CAATCGCTGCAATCGCAAGTCTGCTTTACGATGCTTTCTTCATCGTGACGTTCTTCAGCATCACTCGTCTTGAAGTTGATGTGACATTC ATCGCAGCGATTTTAACGATCATCGGATACAGCATCAACGATACAATCGTAACGTTCGACCGTGTTCGTGAGCACATGAAGAAACGTAA ACCAAAAACGTTCGCTGATTTAAACCACATCGTAAACCTAAGCCTTCAACAAACGTTCACTCGTTCTATCAACACTGTGCTGACTGTTG TTATCGTTGTTGTAACGCTATTAATCTTCGGTGCTTCTTCCATTACAAACTTCAGCATTGCGCTTCTTGTTGGTTTGCTGACTGGCGTT TACAGTTCGCTTTACATCGCTGCACAAATTTGGCTGGCTTGGAAAGGCCGTGAGCTGAAGAAAGACAGTGCACAG	
vrbF	ATGACTGGAACACTTGGAACGCTTGTGCCGATCATTTTAATGTTCGCTGTTCTTTACTTCCTGCTGATCCGTCCTCAACAAAAACAACA AAAAGCTGTTCGTCAAATGCAAGAAGAATTAAAGAAAGGCGATTCTGTTGTGACAATCGGCGGTCTTCACGGTACTGTTGACAGCATTG ATGAAAGCAAAGTAGTCATCAAAACAGGTGACAACACTCGTTTAACGTTCGACCGCCGTGCGATCCGTGAAGTATCTGCTGCTGAA	81

Table H.2 Synthetic RBS calculated using the online tool “RBS calculator” for the expression of native or codon pair optimised (CpO) version of genes expressing components of the Sec translocase (Salis *et al.* 2010; Borujeni *et al.* 2014).

Gene	**Synthetic RBS**
secA	AAATAGGCGTGTGATGATAGCGGAAAGGGAAGAGGAAGGAGGTAAAAA
$secA_{CpO}$	AAATAGGCGTGTGATGATAGCGGAAATCAACCGCGCAGAAGAAGGAGGTACAGT
secY	AAGCTGCTGGCGGTACAGCTGCCCTTTAAAAATCCCAGAGGAGGTCTATT
$secY_{CpO}$	AAGCTGCTGGCGGTACAGCTACTCATCAGTAGGGAGGGAAGGAGGTACTCA
seE	TTTTTGCGCTTTTAAATTGTCGGGTCGAAATCGGTCCATTAGAAGGAGGTAACT
$secE_{CpO}$	TTTTTGCGCTTTTAAATTGTGAGGTCAATCGTCTCAATCAAAAAGGAGGTAGGA
secG	GTAATGTAGCCAGTGAGTCTAGAAATTCTAGTAAAAGAAGGAGGTACCAC
$secG_{CpO}$	GTAATGTAGCCAGTGAGTCTGTTCAATTTACTATCATCCGGGGGTTAACTTA

Table H.3 DNA sequences of the biobricks used for the combinatorial design to optimise the expression of components of the Sec translocase in Chapter 5.

Biobrick	Sequence
P1	CGTCTCGCCGCAGCTTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGA GAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGACACGGGCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAG TTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTGACGGCGGGATATAACATGAGCTGTCTTCGGTAT CGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACC AGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGG CTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCT GGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCA AGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGCGTTG CGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAG ATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGT GCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCG GGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCATCAAATAGTCTCCCTCCGTTTGAATATTTGATT GATCGTAACCAGATCAAGCACTCTTTCCACTATCCCTACAGTGTTATGGCTTGAACAATCACGAAACAATAATTGGTACGTACGATCTTTCAGCC GACTCAAACATCAAATCTTACAAATGTAGTCTTTGAAAGTATTACATATGTAAGATTTAAATGCAACCGTTTTTTCGGAAGGAAATGATGACCTC GTTTCCACCGAATTAGCTTGAAATAGTACATAATGGATTTCCTTACGCGAAATACGGGCAGACATGGCCTGCCCGGTTATTATTATTTTTGACAC CGCATGCTGCGGTACCACCGGATCTGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGT AGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTC GTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGG CGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTTGT TTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATCCTGAAGTCGGGGATCTCTGCAGTC GGGGGATCGATCCTCTAGAGTCAATTGTGAGCGCTCACAATTTCTGCTAAAATTCCTGAAAAATTTTGCAAAAAGTTGTTGACTTTATCTACAAG GTGTGGCATAATGTGTGGAATTGTGAGCGCTCACAATTGACCTGCAGGCATGCCTGCAGTTTAGGAGACG
1A	CGTCTCATTTATACATGCCTCTAAAATAGGCGTGTGATGATAGAGGAGCGTTATAAATGCTTGGAATTTTAAATAAAATGTTTGATCCAACAAAA CGTACGCTGAATAGATACGAAAAAATTGCTAACGATATTGATGCGATTCGCGGAGACTATGAAAATCTCTCTGACGACGCATTGAAACATAAAAC AATTGAATTTAAAGAGCGTCTTGAAAAAGGGGCGACAACGGATGATCTTCTTGTTGAAGCTTTCGCTGTTGTTCGAGAAGCTTCACGCCGCGTAA CAGGCATGTTTCCGTTTAAAGTCCAGCTCATGGGGGGCGTGGCGCTTCATGACGGAAATATAGCGGAAATGAAAACAGGGGAAGGGAAAACATTA ACGTCTACCCTGCCTGTTTATTTAAATGCGTTAACCGGTAAAGGCGTACACGTCGTGACTGTCAACGAATACTTGGCAAGCCGTGACGCTGAGCA AATGGGGAAAATTTTCGAGTTTCTCGGTTTGACTGTCGGTTTGAATTTAAACTCAATGTCAAAAGACGAAAAACGGGAAGCTTATGCCGCTGATA TTACTTACTCCACAAACAACGAGCTTGGCTTCGACTATTTGCGTGACAATATGGTTCTTTATAAAGAGCAGATGGTTCAGCGCCCGCTTCATTTT GCGGTAATAGATGAAGTTGACTCTATTTTAATTGATGAAGCAAGAACACCGCTTATCATTTCTGGACAAGCTGCAAAATCCACTAAGCTGTACGT

	ACAGGCAAATGCTTTTGTCCGCACGTTAAAAGCGGAGAAGGATTACACGTACGATATCAAAACAAAAGCTGTACAGCTTACTGAAGAAGGAATGA CGAAGGCGGAAAAAGCATTCGGCATCGATAACCTCTTTGATGTGAAGCATGTCGCGCTCAACCACCATATCAACCAGGCCTTAAAAGCTCACGTT GCGATGCAAAAGGACGTTGACTATGTAGTGGAAGACGGACAGGTTGTTATTGTTGATTCCTTCACGGGACGTCTGATGAAAGGCCGCCGCTACAG TGAGGGGCTTCACCAAGCGATTGAAGCAAAGGAAGGGCTTGAGATTCAAAACGAAAGCATGACCTTGGCGACGATTACGTTCCAAAACTACTTCC GAATGTACGAAAAACTTGCCGGTATGACGGGTACAGCTAAGACAGAGGAAGAAGAATTCCGCAACATCTACAACATGCAGGTTGTCACGATCCCT ACCAACAGGCCTGTTGTCCGTGATGACCGCCCGGATTTAATTTACCGCACGATGGAAGGAAAGTTTAAGGCAGTTGCGGAGGATGTCGCACAGCG TTACATGACGGGACAGCCTGTTCTAGTCGGTACGGTTGCCGTTGAAACATCTGAATTGATTTCTAAGCTGCTTAAAAACAAAGGAATTCCGCATC AAGTGTTAAATGCCAAAAACCATGAACGTGAAGCGCAGATCATTGAAGAGGCCGGCCAAAAAGGCGCAGTTACGATTGCGACTAACATGGCGGGG CGCGGAACGGACATTAAGCTTGGCGAAGGTGTAAAAGAGCTTGGCGGGCTCGCTGTAGTCGGAACAGAACGACATGAATCACGCCGGATTGACAA TCAGCTTCGAGGTCGTTCCGGACGTCAGGGAGACCCGGGGATTACTCAATTTTATCTTTCTATGGAAGATGAATTGATGCGCAGATTCGGAGCTG AGCGGACAATGGCGATGCTTGACCGCTTCGGCATGGACGACTCTACTCCAATCCAAAGCAAAATGGTATCTCGCGCGGTTGAATCATCTCAAAAA CGCGTCGAAGGCAATAACTTCGATTCGCGTAAACAGCTTCTGCAATATGATGATGTTCTCCGCCAGCAGCGTGAGGTCATTTATAAGCAGCGCTT TGAAGTCATTGACTCTGAAAACCTGCGTGAAATCGTTGAAAATATGATCAAGTCTTCTCTCGAACGCGCAATTGCAGCCTATACGCCAAGAGAAG AGCTTCCTGAGGAGTGGAAGCTTGACGGTCTAGTTGATCTTATCAACACAACTTATCTTGATGAAGGTGCACTTGAGAAGAGCGATATCTTCGGC AAAGAACCGGATGAAATGCTTGAGCTCATTATGGATCGCATCATCACAAAATATAATGAGAAGGAAGAGCAATTCGGCAAAGAGCAAATGCGCGA ATTCGAAAAAGTTATCGTTCTTCGTGCCGTTGATTCTAAATGGATGGATCATATTGATGCGATGGATCAGCTCCGCCAAGGGATTCACCTTCGTG CTTACGCGCAGACGAACCCGCTTCGTGAGTATCAAATGGAAGGTTTTGCGATGTTTGAGCATATGATTGAATCAATTGAGGACGAAGTCGCAAAA TTTGTGATGAAAGCTGAGATTGAAAACAATCTGGAGCGTGAAGAGGTTGTACAAGGTCAAACAACAGCTCATCAGCCGCAAGAAGGCGACGATAA CAAAAAAGCAAAGAAAGCACCGGTTCGCAAAGTGGTTGATATCGGACGAAATGCCCCATGCCACTGCGGAAGCGGGAAAAAATATAAAAATTGCT GCGGCCGTACTGAATAGATAACGAGACG
1B	CGTCTCATTTATACATGCCTCTAAAATAGGCGTGTGATGATAGAGGAGCGTTATAAATGCTTGGCATTTTAAACAAAATGTTCGACCCAACAAAA CGTACTTTAAACCGTTATGAGAAAATCGCAAACGATATCGATGCGATCCGCGGAGACTACGAAAACCTTTCTGATGATGCTTTAAAACACAAAAC GATCGAATTCAAAGAACGTCTTGAAAAAGGTGCAACTACTGATGATTTGCTTGTTGAAGCATTCGCTGTTGTTCGTGAAGCAAGCCGCCGTGTAA CTGGTATGTTCCCATTCAAAGTACAGCTGATGGGCGGTGTGGCGCTTCATGACGGAAACATCGCTGAAATGAAAACAGGTGAAGGAAAAACGCTG ACAAGCACGCTTCCTGTATACTTAAACGCGCTGACTGGTAAAGGCGTTCACGTTGTAACTGTAAACGAATACCTTGCTTCTCGTGACGCTGAACA AATGGGTAAAATCTTCGAATTCCTTGGTTTGACTGTAGGTTTGAACCTGAACAGCATGAGTAAAGATGAAAAACGTGAAGCTTACGCTGCTGATA TCACTTACAGTACAAACAACGAGCTTGGCTTCGATTATTTGCGTGACAACATGGTGCTTTACAAAGAGCAAATGGTGCAGCGTCCATTGCACTTC GCTGTTATCGATGAAGTTGACAGCATTTTGATTGATGAAGCTCGTACACCGCTGATCATTTCTGGTCAAGCTGCAAAAAGCACAAAGCTTTACGT GCAGGCAAACGCTTTCGTTCGTACGCTGAAAGCTGAAAAAGACTATACGTATGACATCAAAACAAAAGCTGTTCAATTAACTGAAGAAGGCATGA CAAAAGCAGAAAAAGCATTCGGTATCGACAACCTATTCGATGTAAAACACGTTGCGCTGAACCACCACATCAACCAAGCTTTAAAAGCGCACGTT GCAATGCAAAAAGATGTGGATTACGTTGTTGAAGACGGACAAGTCGTTATCGTTGATTCTTTCACTGGACGTTTAATGAAAGGACGCCGTTACTC TGAAGGTCTTCATCAGGCGATTGAAGCAAAAGAAGGTCTTGAAATCCAAAACGAATCAATGACTTTGGCAACCATTACATTCCAAAACTACTTCC GTATGTACGAAAAGCTTGCTGGTATGACTGGAACTGCAAAAACGGAAGAAGAAGAATTCCGTAATATTTACAACATGCAAGTTGTGACGATTCCG ACAAACCGTCCTGTTGTGCGTGATGACCGTCCTGATTTGATTTACCGTACAATGGAAGGAAAATTCAAAGCTGTTGCTGAAGATGTTGCACAGCG TTACATGACTGGACAGCCTGTGCTTGTCGGAACTGTTGCTGTTGAAACTTCTGAGCTGATTTCTAAACTATTGAAAAACAAAGGCATTCCTCATC AAGTGCTGAACGCTAAAAACCATGAGCGTGAAGCTCAAATCATCGAAGAAGCAGGACAAAAAGGCGCTGTAACGATTGCAACAAACATGGCAGGC CGCGGAACTGATATCAAACTAGGTGAAGGCGTAAAAGAGCTTGGCGGACTTGCTGTTGTAGGTACTGAGCGTCATGAAAGCCGCCGCATTGATAA

	CCAGCTTCGCGGACGCAGCGGACGTCAAGGTGATCCAGGTATCACTCAATTCTACCTTTCCATGGAGGATGAGCTGATGCGTCGTTTCGGCGCTG AACGTACGATGGCGATGCTTGACCGTTTCGGCATGGATGACAGCACGCCAATCCAATCTAAAATGGTCAGCCGCGCTGTAGAATCTTCTCAAAAA CGTGTTGAAGGAAACAACTTCGACAGCCGCAAACAGCTTCTTCAATACGATGATGTGCTTCGCCAGCAGCGTGAAGTGATTTACAAACAGCGTTT CGAAGTGATCGACAGTGAAAACCTTCGTGAAATCGTTGAAAACATGATCAAATCAAGCCTTGAGCGTGCGATCGCTGCTTACACACCGCGTGAAG AACTTCCTGAAGAATGGAAGCTTGACGGACTTGTTGATTTAATCAACACAACTTACCTTGATGAAGGTGCTTTAGAAAAATCTGATATTTTCGGA AAAGAGCCTGATGAGATGCTTGAGCTGATCATGGACCGTATCATCACAAAATACAACGAAAAAGAAGAACAATTCGGCAAAGAACAAATGCGTGA ATTTGAAAAAGTGATCGTACTTCGTGCTGTAGACAGCAAATGGATGGATCACATTGATGCGATGGACCAATTGCGTCAAGGCATTCACCTTCGTG CTTACGCTCAAACAAACCCGCTTCGTGAATACCAAATGGAAGGCTTCGCAATGTTCGAGCACATGATCGAAAGCATTGAAGATGAAGTGGCAAAA TTCGTTATGAAGGCTGAAATCGAAAACAACCTTGAACGTGAAGAAGTGGTTCAAGGCCAAACAACTGCTCATCAGCCTCAAGAAGGCGATGACAA CAAGAAAGCGAAGAAAGCTCCTGTACGCAAAGTTGTTGATATCGGACGCAACGCTCCTTGCCATTGCGGAAGCGGCAAAAAGTACAAAAACTGCT GCGGCCGTACTGAATAAATAACGAGACG
1C	CGTCTCATTTATACATGCCTCTAAAATAGGCGTGTGATGATAGCGGAAAGGGAAGAGGAAGGAGGTAAAAAATGCTTGGAATTTTAAATAAAATG TTTGATCCAACAAAACGTACGCTGAATAGATACGAAAAAATTGCTAACGATATTGATGCGATTCGCGGAGACTATGAAAATCTCTCTGACGACGC ATTGAAACATAAAACAATTGAATTTAAAGAGCGTCTTGAAAAAGGGGCGACAACGGATGATCTTCTTGTTGAAGCTTTCGCTGTTGTTCGAGAAG CTTCACGCCGCGTAACAGGCATGTTTCCGTTTAAAGTCCAGCTCATGGGGGGCGTGGCGCTTCATGACGGAAATATAGCGGAAATGAAAACAGGG GAAGGGAAAACATTAACGTCTACCCTGCCTGTTTATTTAAATGCGTTAACCGGTAAAGGCGTACACGTCGTGACTGTCAACGAATACTTGGCAAG CCGTGACGCTGAGCAAATGGGGAAAATTTTCGAGTTTCTCGGTTTGACTGTCGGTTTGAATTTAAACTCAATGTCAAAAGACGAAAAACGGGAAG CTTATGCCGCTGATATTACTTACTCCACAAACAACGAGCTTGGCTTCGACTATTTGCGTGACAATATGGTTCTTTATAAAGAGCAGATGGTTCAG CGCCCGCTTCATTTTGCGGTAATAGATGAAGTTGACTCTATTTTAATTGATGAAGCAAGAACACCGCTTATCATTTCTGGACAAGCTGCAAAATC CACTAAGCTGTACGTACAGGCAAATGCTTTTGTCCGCACGTTAAAAGCGGAGAAGGATTACACGTACGATATCAAAACAAAAGCTGTACAGCTTA CTGAAGAAGGAATGACGAAGGCGGAAAAAGCATTCGGCATCGATAACCTCTTTGATGTGAAGCATGTCGCGCTCAACCACCATATCAACCAGGCC TTAAAAGCTCACGTTGCGATGCAAAAGGACGTTGACTATGTAGTGGAAGACGGACAGGTTGTTATTGTTGATTCCTTCACGGGACGTCTGATGAA AGGCCGCCGCTACAGTGAGGGGCTTCACCAAGCGATTGAAGCAAAGGAAGGGCTTGAGATTCAAAACGAAAGCATGACCTTGGCGACGATTACGT TCCAAAACTACTTCCGAATGTACGAAAAACTTGCCGGTATGACGGGTACAGCTAAGACAGAGGAAGAAGAATTCCGCAACATCTACAACATGCAG GTTGTCACGATCCCTACCAACAGGCCTGTTGTCCGTGATGACCGCCCGGATTTAATTTACCGCACGATGGAAGGAAAGTTTAAGGCAGTTGCGGA GGATGTCGCACAGCGTTACATGACGGGACAGCCTGTTCTAGTCGGTACGGTTGCCGTTGAAACATCTGAATTGATTTCTAAGCTGCTTAAAAACA AAGGAATTCCGCATCAAGTGTTAAATGCCAAAAACCATGAACGTGAAGCGCAGATCATTGAAGAGGCCGGCCAAAAAGGCGCAGTTACGATTGCG ACTAACATGGCGGGGCGCGGAACGGACATTAAGCTTGGCGAAGGTGTAAAAGAGCTTGGCGGGCTCGCTGTAGTCGGAACAGAACGACATGAATC ACGCCGGATTGACAATCAGCTTCGAGGTCGTTCCGGACGTCAGGGAGACCCGGGGATTACTCAATTTTATCTTTCTATGGAAGATGAATTGATGC GCAGATTCGGAGCTGAGCGGACAATGGCGATGCTTGACCGCTTCGGCATGGACGACTCTACTCCAATCCAAAGCAAAATGGTATCTCGCGCGGTT GAATCATCTCAAAAACGCGTCGAAGGCAATAACTTCGATTCGCGTAAACAGCTTCTGCAATATGATGATGTTCTCCGCCAGCAGCGTGAGGTCAT TTATAAGCAGCGCTTTGAAGTCATTGACTCTGAAAACCTGCGTGAAATCGTTGAAAATATGATCAAGTCTTCTCTCGAACGCGCAATTGCAGCCT ATACGCCAAGAGAAGAGCTTCCTGAGGAGTGGAAGCTTGACGGTCTAGTTGATCTTATCAACACAACTTATCTTGATGAAGGTGCACTTGAGAAG AGCGATATCTTCGGCAAAGAACCGGATGAAATGCTTGAGCTCATTATGGATCGCATCATCACAAAATATAATGAGAAGGAAGAGCAATTCGGCAA AGAGCAAATGCGCGAATTCGAAAAAGTTATCGTTCTTCGTGCCGTTGATTCTAAATGGATGGATCATATTGATGCGATGGATCAGCTCCGCCAAG GGATTCACCTTCGTGCTTACGCGCAGACGAACCCGCTTCGTGAGTATCAAATGGAAGGTTTTGCGATGTTTGAGCATATGATTGAATCAATTGAG GACGAAGTCGCAAAATTTGTGATGAAAGCTGAGATTGAAAACAATCTGGAGCGTGAAGAGGTTGTACAAGGTCAAACAACAGCTCATCAGCCGCA

	AGAAGGCGACGATAACAAAAAAGCAAAGAAAGCACCGGTTCGCAAAGTGGTTGATATCGGACGAAATGCCCCATGCCACTGCGGAAGCGGGAAAA AATATAAAAATTGCTGCGGCCGTACTGAATAGATAACGAGACG
1D	CGTCTCATTTACAAGATAATAACAAGGAAGGAGGAATACAATGCTTGGCATTTTAAACAAAATGTTCGACCCAACAAAACGTACTTTAAACCGTT ATGAGAAAATCGCAAACGATATCGATGCGATCCGCGGAGACTACGAAAACCTTTCTGATGATGCTTTAAAACACAAAACGATCGAATTCAAAGAA CGTCTTGAAAAAGGTGCAACTACTGATGATTTGCTTGTTGAAGCATTCGCTGTTGTTCGTGAAGCAAGCCGCCGTGTAACTGGTATGTTCCCATT CAAAGTACAGCTGATGGGCGGTGTGGCGCTTCATGACGGAAACATCGCTGAAATGAAAACAGGTGAAGGAAAAACGCTGACAAGCACGCTTCCTG TATACTTAAACGCGCTGACTGGTAAAGGCGTTCACGTTGTAACTGTAAACGAATACCTTGCTTCTCGTGACGCTGAACAAATGGGTAAAATCTTC GAATTCCTTGGTTTGACTGTAGGTTTGAACCTGAACAGCATGAGTAAAGATGAAAAACGTGAAGCTTACGCTGCTGATATCACTTACAGTACAAA CAACGAGCTTGGCTTCGATTATTTGCGTGACAACATGGTGCTTTACAAAGAGCAAATGGTGCAGCGTCCATTGCACTTCGCTGTTATCGATGAAG TTGACAGCATTTTGATTGATGAAGCTCGTACACCGCTGATCATTTCTGGTCAAGCTGCAAAAAGCACAAAGCTTTACGTGCAGGCAAACGCTTTC GTTCGTACGCTGAAAGCTGAAAAAGACTATACGTATGACATCAAAACAAAAGCTGTTCAATTAACTGAAGAAGGCATGACAAAAGCAGAAAAAGC ATTCGGTATCGACAACCTATTCGATGTAAAACACGTTGCGCTGAACCACCACATCAACCAAGCTTTAAAAGCGCACGTTGCAATGCAAAAAGATG TGGATTACGTTGTTGAAGACGGACAAGTCGTTATCGTTGATTCTTTCACTGGACGTTTAATGAAAGGACGCCGTTACTCTGAAGGTCTTCATCAG GCGATTGAAGCAAAAGAAGGTCTTGAAATCCAAAACGAATCAATGACTTTGGCAACCATTACATTCCAAAACTACTTCCGTATGTACGAAAAGCT TGCTGGTATGACTGGAACTGCAAAAACGGAAGAAGAAGAATTCCGTAATATTTACAACATGCAAGTTGTGACGATTCCGACAAACCGTCCTGTTG TGCGTGATGACCGTCCTGATTTGATTTACCGTACAATGGAAGGAAAATTCAAAGCTGTTGCTGAAGATGTTGCACAGCGTTACATGACTGGACAG CCTGTGCTTGTCGGAACTGTTGCTGTTGAAACTTCTGAGCTGATTTCTAAACTATTGAAAAACAAAGGCATTCCTCATCAAGTGCTGAACGCTAA AAACCATGAGCGTGAAGCTCAAATCATCGAAGAAGCAGGACAAAAAGGCGCTGTAACGATTGCAACAAACATGGCAGGCCGCGGAACTGATATCA AACTAGGTGAAGGCGTAAAAGAGCTTGGCGGACTTGCTGTTGTAGGTACTGAGCGTCATGAAAGCCGCCGCATTGATAACCAGCTTCGCGGACGC AGCGGACGTCAAGGTGATCCAGGTATCACTCAATTCTACCTTTCCATGGAGGATGAGCTGATGCGTCGTTTCGGCGCTGAACGTACGATGGCGAT GCTTGACCGTTTCGGCATGGATGACAGCACGCCAATCCAATCTAAAATGGTCAGCCGCGCTGTAGAATCTTCTCAAAAACGTGTTGAAGGAAACA ACTTCGACAGCCGCAAACAGCTTCTTCAATACGATGATGTGCTTCGCCAGCAGCGTGAAGTGATTTACAAACAGCGTTTCGAAGTGATCGACAGT GAAAACCTTCGTGAAATCGTTGAAAACATGATCAAATCAAGCCTTGAGCGTGCGATCGCTGCTTACACACCGCGTGAAGAACTTCCTGAAGAATG GAAGCTTGACGGACTTGTTGATTTAATCAACACAACTTACCTTGATGAAGGTGCTTTAGAAAAATCTGATATTTTCGGAAAAGAGCCTGATGAGA TGCTTGAGCTGATCATGGACCGTATCATCACAAAATACAACGAAAAAGAAGAACAATTCGGCAAAGAACAAATGCGTGAATTTGAAAAAGTGATC GTACTTCGTGCTGTAGACAGCAAATGGATGGATCACATTGATGCGATGGACCAATTGCGTCAAGGCATTCACCTTCGTGCTTACGCTCAAACAAA CCCGCTTCGTGAATACCAAATGGAAGGCTTCGCAATGTTCGAGCACATGATCGAAAGCATTGAAGATGAAGTGGCAAAATTCGTTATGAAGGCTG AAATCGAAAACAACCTTGAACGTGAAGAAGTGGTTCAAGGCCAAACAACTGCTCATCAGCCTCAAGAAGGCGATGACAACAAGAAAGCGAAGAAA GCTCCTGTACGCAAAGTTGTTGATATCGGACGCAACGCTCCTTGCCATTGCGGAAGCGGCAAAAAGTACAAAAACTGCTGCGGCCGTACTGAATA AATAACGAGACG
P2	CGTCTCGCCGCACTAACTTATAGGGGTAACACTTAAAAAAGAATCAATAACGATAGAAACCGCTCCTAAAGCAGGTGCATTTTTTCCTAACGAAG AAGGCAATAGTTCACATTTATTGTCTAAATGAGAATGGACTCTAGAAGAAACTTCGTTTTTAATCGTATTTAAAACAATGGGATGAGATTCAATT ATATGATTTCTCAAGATAACAGCTTCTATATCAAATGTATTAAGGATATTGGTTAATCCAATTCCGATATAAAAGCCAAAGTTTTGAAGTGCATT TAACATTTCTACATCATTTTTATTTGCGCGTTCCACAATCTCTTTTCGAGAAATATTCTTTTCTTCTTTAGAGAGCGAAGCCAGTAACGCTTTTT CAGAAGCATATAATTCCCAACAGCCTCGATTTCCACAGCTGCATTTGGGTCCATTAAAATCTATCGTCATATGACCCATTTCCCCAGAAAAACCC TGAACACCTTTATACAATTCGTTGTTAATAACAAGTCCAGTTCCAATTCCGATATTAATACTGATGTAAACGATGTTTTCATAGTTTTTTGTCAT ACCAAATACTTTTTCACCGTATGCTCCTGCATTAGCTTCATTTTCAACAAAAACCGGAACATTAAACTCACTCTCAATTAAAAACTGCAAATCTT

	TGATATTCCAATTTAAGTTAGGCATGAAAATAATTTGCTGATGACGATCTACAAGGCCTGGAACACAAATTCCTATTCCGACTAGACCATAAGGG GACTCAGGCATATGGGTTACAAAACCATGAATAAGTGCAAATAAAATCTCTTTTACTTCACTAGCGGAAGAACTAGACAAGTCAGAAGTCTTCTC GAGAATAATATTTCCTTCTAAGTCGGTTAGAATTCCGTTAAGATAGTCGACTCCTATATCAATACCAATCGAGTAGCCTGCATTCTTATTAAAAA CAAGCATTACAGGTCTTCTGCCGCCTCTAGATTGCCCTGCCCCAATTTCAAAAATAAAATCTTTTTCAAGCAGTGTATTTACTTGAGAGGAGACA GTAGACTTGTTTAATCCTGTAATCTCAGAGAGAGTTGCCCTGGAGACAGGGGAGTTCTTCAAAATTTCATCTAATATTAATTTTTGATTCATTTT TTTTACTAAAGCTTGATCTGCAATTTGAATAATAACCACTCCTTTGTTTATCCACCGAACTAAGTTGGTGTTTTTTGAAGCTTGAATTAGATATT TAAAAGTATCATATCTAATATTATAACTAAATTTTCTAAAAAAAAACATTGAAATAAACATTTATTTTGTATATGATGAGATAAAGTTAGTTTATT GGATAAACAAACTAACTCAATTAAGATAGTTGATGGATAAACTTGTTCACTTAAATCAAAGGGGGAAATGACAAATGGTCCAAACTAGTGATATC TAAAAATCAAAGGGGGAAATGGGATCCGCGATGAGACG
2A	CGTCTCTGCGAAAGAAGCTGTTGAAGCTGCTGGCGGTACAGCTGAGGTGATCTAACTTGTTTAAAACAATCTCCAACTTTATGCGTGTGAGTGAT ATCAGGAATAAAATCATATTCACTTTACTCATGCTTATCGTCTTTCGCATAGGTGCGTTTATTCCTGTGCCTTACGTTAACGCTGAAGCGTTACA GGCACAGTCTCAAATGGGTGTTTTTGATCTCCTTAATACATTTGGCGGCGGTGCGCTTTACCAATTTTCCATTTTCGCAATGGGAATTACTCCTT ATATCACGGCTTCGATCATCATTCAGCTGCTTCAGATGGATGTGGTACCGAAGTTTACCGAGTGGTCTAAGCAAGGTGAAGTTGGCCGCCGTAAA TTAGCTCAGTTCACAAGGTACTTTACGATTGTGCTTGGTTTCATCCAAGCGTTAGGTATGTCATATGGATTCAACAATCTGGCAAACGGTATGCT GATCGAAAAATCCGGTGTATCGACATATCTTATCATTGCTTTAGTGCTCACTGGCGGAACTGCCTTTTTAATGTGGCTTGGGGAACAAATTACTT CTCATGGAGTAGGCAACGGAATATCGATCATTATCTTCGCGGGGATTGTGTCTAGTATTCCAAAAACAATTGGGCAAATATATGAGACTCAATTT GTCGGCAGCAACGATCAGTTGTTTATTCATATTGTGAAAGTCGCACTTCTTGTGATTGCGATTTTAGCAGTTATTGTTGGAGTTATTTTCATTCA GCAAGCCGTACGGAAAATTGCGATTCAATATGCTAAAGGCACAGGTCGTTCACCTGCTGGCGGAGGTCAGTCTACACACCTTCCATTGAAAGTGA ATCCTGCAGGGGTTATTCCGGTAATCTTTGCGGTTGCGTTTTTGATAACGCCGCGGACGATCGCGTCATTCTTTGGAACAAACGATGTGACAAAG TGGATTCAAAACAACTTTGATAATACGCATCCGGTGGGTATGGCGATATATGTTGCGTTGATTATTGCCTTTACGTACTTTTATGCTTTTGTACA GGTAAACCCTGAACAAATGGCTGATAACCTTAAAAAACAGGGTGGCTATATCCCGGGGGTTCGTCCAGGGAAAATGACTCAAGATAGAATTACGA GCATTTTGTATCGACTTACGTTTGTGGGTTCTATATTCTTAGCCGTGATTTCCATTCTTCCTATCTTTTTCATTCAATTCGCTGGATTGCCTCAA AGTGCACAAATTGGCGGAACATCTTTGTTAATTGTTGTCGGGGTAGCCTTGGAGACAATGAAACAACTAGAAAGCCAGTTGGTGAAACGAAACTA CCGTGGATTTATGAAAAACTAGATCTTGAAACAAAATAGTTTTTGCGCTTTTAAATTGTGGAGGTCTTTTACATGCGTATTATGAAATTCTTTAA AGATGTTGGGAAAGAAATGAAAAAGGTAAGCTGGCCTAAAGGAAAAGAGTTAACGCGTTATACCATTACGGTAATTTCAACAGTTATCTTTTTTG TTATCTTTTTTGCCCTCCTTGACACAGGAATTTCTCAATTAATTCGTTTAATAGTTGAATAATTTCATGTAAAATAGAAGTAATGTAGCCAGTGA GTCTGGAGGTGTATGGGATGCACGCAGTTTTGATTACCTTATTGGTTATCGTCAGCATTGCACTTATTATTGTCGTTTTGCTTCAATCCAGTAAA AGTGCCGGATTATCTGGTGCGATTTCAGGCGGAGCGGAGCAGCTCTTCGGGAAACAAAAAGCAAGAGGTCTTGATTTAATTTTGCACCGCATTAC GGTAGTGCTGGCAGTCTTGTTTTTCGTGTTAACGATTGCGCTTGCTTATATCCTATAGATAACGAGACG
2B	CGTCTCTGCGAAAGAAGCTGTTGAAGCTGCTGGCGGTACAGCTGAGGTGATCTAACATGTTCAAAACGATCAGCAACTTCATGCGTGTATCTGAT ATCCGCAACAAAATCATCTTCACATTGCTCATGCTGATTGTATTCCGTATCGGTGCTTTCATTCCTGTTCCTTACGTGAACGCTGAAGCGCTGCA AGCTCAATCTCAAATGGGTGTATTCGATCTTTTAAACACATTCGGCGGCGGCGCTCTATACCAATTCAGCATTTTCGCAATGGGCATCACTCCAT ACATCACTGCTTCTATCATCATTCAGCTATTGCAAATGGATGTTGTGCCGAAATTCACTGAATGGAGCAAACAAGGTGAAGTGGGACGCAGAAAG CTTGCTCAATTCACTCGTTACTTCACGATTGTGCTCGGATTCATTCAAGCTCTTGGTATGTCATACGGCTTCAACAACCTTGCAAACGGTATGCT GATCGAAAAATCAGGTGTATCAACTTACTTAATCATCGCTCTTGTATTAACTGGCGGAACTGCATTCCTGATGTGGCTCGGTGAACAAATCACTT CTCACGGCGTTGGAAACGGAATCAGCATCATCATCTTCGCAGGCATCGTTTCTTCTATCCCAAAAACGATCGGACAAATTTACGAAACTCAATTC GTCGGAAGCAACGACCAATTGTTCATTCACATTGTGAAAGTAGCTTTGCTTGTTATCGCAATCCTTGCTGTTATCGTAGGTGTGATCTTCATTCA

	ACAAGCTGTTCGCAAAATCGCTATTCAATACGCAAAAGGTACTGGCCGTTCTCCTGCTGGCGGCGGACAATCAACTCACCTTCCATTGAAAGTAA ACCCAGCTGGTGTTATCCCAGTTATCTTCGCTGTTGCTTTCCTGATCACTCCAAGAACGATTGCAAGCTTCTTCGGAACAAACGATGTGACAAAA TGGATTCAAAACAACTTCGACAACACTCACCCTGTTGGTATGGCGATTTACGTGGCGCTGATCATCGCTTTCACTTACTTCTACGCATTCGTTCA AGTGAACCCTGAACAAATGGCTGATAACCTAAAGAAACAAGGCGGATATATTCCAGGTGTTCGTCCTGGAAAAATGACTCAAGACCGCATTACAA GCATCCTTTACCGCCTGACGTTCGTTGGAAGCATTTTCTTGGCTGTTATCAGCATCCTTCCGATTTTCTTCATCCAATTCGCTGGTCTTCCTCAA TCTGCTCAAATCGGCGGAACTTCCTTGCTCATTGTTGTCGGCGTTGCGCTTGAAACGATGAAGCAGCTTGAAAGCCAGCTTGTAAAACGCAACTA CCGCGGATTCATGAAAAACTAAATCTTGAAACAAAATAGTTTTTGCGCTTTTAAATTGTGGAGGTCTTTTACATGCGTATCATGAAATTCTTCAA AGATGTAGGTAAAGAAATGAAGAAAGTATCTTGGCCTAAAGGAAAAGAACTGACTCGTTATACGATCACTGTTATCAGCACTGTGATTTTCTTCG TTATCTTCTTCGCACTTCTTGACACTGGCATTTCTCAATTAATCCGTTTAATCGTTGAATAATTTCATGTAAAATAGAAGTAATGTAGCCAGTGA GTCTGGAGGTGTATGGGATGCACGCTGTGCTGATCACTCTTCTTGTTATCGTATCTATCGCTTTAATCATCGTTGTGCTGCTTCAATCTTCTAAA TCAGCTGGTCTTTCTGGTGCAATCAGCGGCGGCGCTGAACAATTATTCGGAAAACAAAAAGCACGCGGACTTGATTTGATCCTTCACCGTATCAC TGTTGTGCTTGCTGTATTGTTCTTCGTTTTAACGATCGCTCTTGCTTACATCCTTTAAATAACGAGACG
2C	CGTCTCTGCGAAAGAAGCTGTTGAAGCTGCTGGCGGTACAGCTGCCCTTTAAAAATCCCAGAGGAGGTCTATTTTGTTTAAAACAATCTCCAACT TTATGCGTGTGAGTGATATCAGGAATAAAATCATATTCACTTTACTCATGCTTATCGTCTTTCGCATAGGTGCGTTTATTCCTGTGCCTTACGTT AACGCTGAAGCGTTACAGGCACAGTCTCAAATGGGTGTTTTTGATCTCCTTAATACATTTGGCGGCGGTGCGCTTTACCAATTTTCCATTTTCGC AATGGGAATTACTCCTTATATCACGGCTTCGATCATCATTCAGCTGCTTCAGATGGATGTGGTACCGAAGTTTACCGAGTGGTCTAAGCAAGGTG AAGTTGGCCGCCGTAAATTAGCTCAGTTCACAAGGTACTTTACGATTGTGCTTGGTTTCATCCAAGCGTTAGGTATGTCATATGGATTCAACAAT CTGGCAAACGGTATGCTGATCGAAAAATCCGGTGTATCGACATATCTTATCATTGCTTTAGTGCTCACTGGCGGAACTGCCTTTTTAATGTGGCT TGGGGAACAAATTACTTCTCATGGAGTAGGCAACGGAATATCGATCATTATCTTCGCGGGGATTGTGTCTAGTATTCCAAAAACAATTGGGCAAA TATATGAGACTCAATTTGTCGGCAGCAACGATCAGTTGTTTATTCATATTGTGAAAGTCGCACTTCTTGTGATTGCGATTTTAGCAGTTATTGTT GGAGTTATTTTCATTCAGCAAGCCGTACGGAAAATTGCGATTCAATATGCTAAAGGCACAGGTCGTTCACCTGCTGGCGGAGGTCAGTCTACACA CCTTCCATTGAAAGTGAATCCTGCAGGGGTTATTCCGGTAATCTTTGCGGTTGCGTTTTTGATAACGCCGCGGACGATCGCGTCATTCTTTGGAA CAAACGATGTGACAAAGTGGATTCAAAACAACTTTGATAATACGCATCCGGTGGGTATGGCGATATATGTTGCGTTGATTATTGCCTTTACGTAC TTTTATGCTTTTGTACAGGTAAACCCTGAACAAATGGCTGATAACCTTAAAAAACAGGGTGGCTATATCCCGGGGGTTCGTCCAGGGAAAATGAC TCAAGATAGAATTACGAGCATTTTGTATCGACTTACGTTTGTGGGTTCTATATTCTTAGCCGTGATTTCCATTCTTCCTATCTTTTTCATTCAAT TCGCTGGATTGCCTCAAAGTGCACAAATTGGCGGAACATCTTTGTTAATTGTTGTCGGGGTAGCCTTGGAGACAATGAAACAACTAGAAAGCCAG TTGGTGAAACGAAACTACCGTGGATTTATGAAAAACTAGATCTTGAAACAAAATAGTTTTTGCGCTTTTAAATTGTCGGGTCGAAATCGGTCCAT TAGAAGGAGGTAACTATGCGTATTATGAAATTCTTTAAAGATGTTGGGAAAGAAATGAAAAAGGTAAGCTGGCCTAAAGGAAAAGAGTTAACGCG TTATACCATTACGGTAATTTCAACAGTTATCTTTTTTGTTATCTTTTTTGCCCTCCTTGACACAGGAATTTCTCAATTAATTCGTTTAATAGTTG AATAATTTCATGTAAAATAGAAGTAATGTAGCCAGTGAGTCTAGAAATTCTAGTAAAAGAAGGAGGTACCACATGCACGCAGTTTTGATTACCTT ATTGGTTATCGTCAGCATTGCACTTATTATTGTCGTTTTGCTTCAATCCAGTAAAAGTGCCGGATTATCTGGTGCGATTTCAGGCGGAGCGGAGC AGCTCTTCGGGAAACAAAAAGCAAGAGGTCTTGATTTAATTTTGCACCGCATTACGGTAGTGCTGGCAGTCTTGTTTTTCGTGTTAACGATTGCG CTTGCTTATATCCTATAGATAACGAGACG
2D	CGTCTCTGCGAAAGAAGCTGTTGAAGCTGCTGGCGGTACAGCTACTCATCAGTAGGGAGGGAAGGAGGTACTCAATGTTCAAAACGATCAGCAAC TTCATGCGTGTATCTGATATCCGCAACAAAATCATCTTCACATTGCTCATGCTGATTGTATTCCGTATCGGTGCTTTCATTCCTGTTCCTTACGT GAACGCTGAAGCGCTGCAAGCTCAATCTCAAATGGGTGTATTCGATCTTTTAAACACATTCGGCGGCGGCGCTCTATACCAATTCAGCATTTTCG CAATGGGCATCACTCCATACATCACTGCTTCTATCATCATTCAGCTATTGCAAATGGATGTTGTGCCGAAATTCACTGAATGGAGCAAACAAGGT

GAAGTGGGACGCAGAAAGCTTGCTCAATTCACTCGTTACTTCACGATTGTGCTCGGATTCATTCAAGCTCTTGGTATGTCATACGGCTTCAACAA
CCTTGCAAACGGTATGCTGATCGAAAAATCAGGTGTATCAACTTACTTAATCATCGCTCTTGTATTAACTGGCGGAACTGCATTCCTGATGTGGC
TCGGTGAACAAATCACTTCTCACGGCGTTGGAAACGGAATCAGCATCATCATCTTCGCAGGCATCGTTTCTTCTATCCCAAAAACGATCGGACAA
ATTTACGAAACTCAATTCGTCGGAAGCAACGACCAATTGTTCATTCACATTGTGAAAGTAGCTTTGCTTGTTATCGCAATCCTTGCTGTTATCGT
AGGTGTGATCTTCATTCAACAAGCTGTTCGCAAAATCGCTATTCAATACGCAAAAGGTACTGGCCGTTCTCCTGCTGGCGGCGGACAATCAACTC
ACCTTCCATTGAAAGTAAACCCAGCTGGTGTTATCCCAGTTATCTTCGCTGTTGCTTTCCTGATCACTCCAAGAACGATTGCAAGCTTCTTCGGA
ACAAACGATGTGACAAAATGGATTCAAAACAACTTCGACAACACTCACCCTGTTGGTATGGCGATTTACGTGGCGCTGATCATCGCTTTCACTTA
CTTCTACGCATTCGTTCAAGTGAACCCTGAACAAATGGCTGATAACCTAAAGAAACAAGGCGGATATATTCCAGGTGTTCGTCCTGGAAAAATGA
CTCAAGACCGCATTACAAGCATCCTTTACCGCCTGACGTTCGTTGGAAGCATTTTCTTGGCTGTTATCAGCATCCTTCCGATTTTCTTCATCCAA
TTCGCTGGTCTTCCTCAATCTGCTCAAATCGGCGGAACTTCCTTGCTCATTGTTGTCGGCGTTGCGCTTGAAACGATGAAGCAGCTTGAAAGCCA
GCTTGTAAAACGCAACTACCGCGGATTCATGAAAAACTAAATCTTGAAACAAAATAGTTTTTGCGCTTTTAAATTGTGAGGTCAATGGTCTCAAT
CAAAAAGGAGGTAGGAATGCGTATCATGAAATTCTTCAAAGATGTAGGTAAAGAAATGAAGAAAGTATCTTGGCCTAAAGGAAAAGAACTGACTC
GTTATACGATCACTGTTATCAGCACTGTGATTTTCTTCGTTATCTTCTTCGCACTTCTTGACACTGGCATTTCTCAATTAATCCGTTTAATCGTT
GAATAATTTCATGTAAAATAGAAGTAATGTAGCCAGTGAGTCTGTTCAATTTACTATCATCCGGGGGTTAACTTAATGCACGCTGTGCTGATCAC
TCTTCTTGTTATCGTATCTATCGCTTTAATCATCGTTGTGCTGCTTCAATCTTCTAAATCAGCTGGTCTTTCTGGTGCAATCAGCGGCGGCGCTG
AACAATTATTCGGAAAACAAAAAGCACGCGGACTTGATTTGATCCTTCACCGTATCACTGTTGTGCTTGCTGTATTGTTCTTCGTTTTAACGATC
GCTCTTGCTTACATCCTTTAAATAACGAGACG

Bibliography

Allenby, N., O 'connor, N., Prágai, Z., Ward, A.C., Wipat, A. & Harwood, C.R., 2005. Genome-Wide Transcriptional Analysis of the Phosphate Starvation Stimulon of *Bacillus subtilis*. *Journal of Bacteriology*, 187(23), pp.8063–8080.

Anagnostopoulos, C. & Spizizen, J., 1961. Requirements for transformation in *Bacillus subtilis*. *Journal of bacteriology*, 81(5), pp.741–6.

Antelmann, H., Darmon, E., Noone, D., Veening, J.-W., Westers, H., Bron, S., et al., 2003. The extracellular proteome of *Bacillus subtilis* under secretion stress conditions. *Molecular Microbiology*, 49(1), pp.143–156.

Antelmann, H., Yamamoto, H., Sekiguchi, J. & Hecker, M., 2002. Stabilization of cell wall proteins in in *Bacillus subtilis*: a proteomic approach. *Proteomics*, 2(5), pp.591–602.

Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S.C., Bron, S., et al., 2003. Essential *Bacillus subtilis* genes. *Proceedings of the National Academy of Sciences of the United States of America*, 100(8), pp.4678–83.

Band, L. & Henner, D.J., 1984. *Bacillus subtilis* requires a "stringent" Shine-Dalgarno region for gene expression. *DNA*, 3(1), pp.17–21.

Barnett, J.P., Eijlander, R.T., Kuipers, O.P. & Robinson, C., 2008. A minimal Tat system from a gram-positive organism: a bifunctional TatA subunit participates in discrete TatAC and TatA complexes. *Journal of Biological Chemistry*, 283(5), pp.2534–42.

Bauer, B.W., Shemesh, T., Chen, Y. & Rapoport, T.A., 2014. A "push and slide" mechanism allows sequence-insensitive translocation of secretory proteins by the SecA ATPase. *Cell*, 157(6), pp.1416–1429.

Beilen, J.W.A. van, 2013. *Molecular physiology of weak organic acid stress in* Bacillus subtilis. University of Amsterdam.

Berks, B.C., Sargent, F. & Palmer, T., 2000. The Tat protein export pathway. *Molecular*

Microbiology, 35(2), pp.260–274.

Biolabs, N.E., 2012. Gibson Assembly Master Mix. *Manual*, pp.1–16.

Blom, E.J., Ridder, A.N.J.A., Lulko, A.T., Roerdink, J.B.T.M. & Kuipers, O.P., 2011. Time-resolved transcriptomics and bioinformatic analyses reveal intrinsic stress responses during batch culture of *Bacillus subtilis*. *PLoS ONE*, 6(11).

Bolhuis, A., Broekhuizen, C.P., Sorokin, A., van Roosmalen, M.L., Venema, G., Bron, S., et al., 1998. SecDF of *Bacillus subtilis*, a Molecular Siamese Twin Required for the Efficient Secretion of Proteins. *Journal of Biological Chemistry*, 273(33), pp.21217–21224.

Bolhuis, A., Matzen, A., Hyyryläinen, H.L., Kontinen, V.P., Meima, R., Chapuis, J., et al., 1999. Signal peptide peptidase- and ClpP-like proteins of *Bacillus subtilis* required for efficient translocation and processing of secretory proteins. *Journal of Biological Chemistry*, 274(35), pp.24585–24592.

Bolhuis, A., Tjalsma, H., Smith, H.E., de Jong, A., Meima, R., Venema, G., et al., 1999. Evaluation of bottlenecks in the late stages of protein secretion in *Bacillus subtilis*. *Applied and Environmental Microbiology*, 65(7), pp.2934–2941.

Borujeni, A.E., Channarasappa, A.S. & Salis, H.M., 2014. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Research*, 42(4), pp.2646–2659.

Braaz, R., Wong, S.-L. & Jendrossek, D., 2002. Production of PHA depolymerase A (PhaZ5) from Paucimonas lemoignei in Bacillus subtilis. *FEMS microbiology letters*, 209(2), pp.237–41.

BRENDA, 2016a. Information on EC 3.2.1.133 - glucan 1,4-alpha-maltohydrolase.

BRENDA, 2016b. Information on EC 3.2.1.8 - endo-1,4-beta-xylanase.

Brockmeier, U., Caspers, M., Freudl, R., Jockwer, A., Noll, T. & Eggert, T., 2006. Systematic Screening of All Signal Peptides from *Bacillus subtilis*: A Powerful Strategy in Optimizing Heterologous Protein Secretion in Gram-positive Bacteria. *Journal of Molecular Biology*, 362(3), pp.393–402.

Brockmeier, U., Wendorff, M. & Eggert, T., 2006. Versatile expression and secretion vectors for *Bacillus subtilis*. *Current microbiology*, 52(2), pp.143–8.

Browning, D.F. & Busby, S.J.W., 2004. The regulation of bacterial transcription initiation. *Nature Reviews Microbiology*, 2(1), pp.57–65.

Buchert, J., Tenkanen, M., Kantelinen, A. & Viikari, L., 1994. Application of xylanases in the pulp and paper industry. *Bioresource Technology*, 50(1), pp.65–72.

Bukhari, D.A. & Rehman, A., 2015. Purification and Characterization of α -Amylase from *Bacillus subtilis* Isolated from Local Environment. *African Journal of Biotechnology*, 47(4), pp.905–911.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., et al., 2009. The MIQE Guidlines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry*, 55(4), pp.611–622.

Carpenter, P.B., Hanlon, D.W. & Ordal, G.W., 1992. flhF, a *Bacillus subtilis* flagellar gene that encodes a putative GTP-binding protein. *Molecular Microbiology*, 6(18), pp.2705–13.

Chen, J., Fu, G., Gai, Y., Zheng, P., Zhang, D. & Wen, J., 2015. Combinatorial Sec pathway analysis for improved heterologous protein secretion in *Bacillus subtilis*: identification of bottlenecks by systematic gene overexpression. *Microbial Cell Factories*, 14(1), p.92.

Chen, J., Gai, Y., Fu, G., Zhou, W., Zhang, D. & Wen, J., 2015. Enhanced extracellular production of α-amylase in *Bacillus subtilis* by optimization of regulatory elements and over-expression of PrsA lipoprotein. *Biotechnology Letters*, 37(4), pp.899–906.

Chen, J., Zhao, L., Fu, G., Zhou, W., Sun, Y., Zheng, P., et al., 2016. A novel strategy for protein production using non-classical secretion pathway in *Bacillus subtilis*. *Microbial Cell Factories*, 15(1), p.69.

Corey, R.A., Allen, W.J., Collinson, I., Song, C., Kumar, A., Saleh, M., et al., 2016. Protein translocation: what's the problem? *Biochemical Society transactions*, 44(3), pp.753–9.

Cotlet, M., Goodwin, P.M., Waldo, G.S. & Werner, J.H., 2006. A comparison of the fluorescence dynamics of single molecules of a green fluorescent protein: One- versus two-photon

excitation. *ChemPhysChem*, 7(1), pp.250–260.

Coxon, R.D., 1990. *Factors affecting protein export from* Bacillus subtilis. University of Newcastle upon Tyne.

Coxon, R.D., Harwood, C.R. & Archibald, A.R., 1991. Protein export during growth of *Bacillus subtilis*: the effect of extracellular protease deficiency. *Letters in Applied Microbiology*, 12(3), pp.91–94.

Dajkovic, A., Hinde, E., MacKichan, C. & Carballido-Lopez, R., 2016. Dynamic organization of SecA and SecY secretion complexes in the *Bacillus subtilis* membrane. *PLoS ONE*, 11(6).

Darmon, E., Noone, D., Masson, A., Bron, S., Kuipers, O.P., Devine, K.M., et al., 2002. A novel class of heat and secretion stress-responsive genes is controlled by the autoregulated CssRS two-component system of *Bacillus subtilis*. *Journal of Bacteriology*, 184(20), pp.5661–5671.

Degering, C., Eggert, T., Puls, M., Bongaerts, J., Evers, S., Maurer, K.-H., et al., 2010. Optimization of protease secretion in *Bacillus subtilis* and *Bacillus licheniformis* by screening of homologous and heterologous signal peptides. *Applied and Environmental Microbiology*, 76(19), pp.6370–6.

Diao, L., Dong, Q., Xu, Z., Yang, S., Zhou, J. & Freudl, R., 2012. Functional Implementation of the Posttranslational SecB-SecA Protein-Targeting Pathway in *Bacillus subtilis*. *Applied and Environmental Microbiology*, 78(3), pp.651–659.

van Dijl, J.M. & Hecker, M., 2013. *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. *Microbial Cell Factories*, 12, p.3.

Dolan, K.M. & Oliver, D.B., 1991. Characterization of *Escherichia coli* SecA protein binding to a site on its mRNA involved in autoregulation. *Journal of Biological Chemistry*, 266(34), pp.23329–33.

Driessen, A.J.M. & Nouwen, N., 2008. Protein translocation across the bacterial cytoplasmic membrane. *Annual Review of Biochemistry*, 77, pp.643–667.

Dubnau, D., 1997. Binding and transport of transforming DNA by *Bacillus subtilis*: The role of

type-IV pilin-like proteins - A review. In *Gene*. pp. 191–198.

Economou, A. & Wickner, W., 1994. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell*, 78(5), pp.835–843.

Erlandson, K.J., Miller, S.B.M., Nam, Y., Osborne, A.R., Zimmer, J. & Rapoport, T.A., 2008. A role for the two-helix finger of the SecA ATPase in protein translocation. *Nature*, 455(7215), pp.984–987.

European Commission, 2012. *Innovating for Sustainable Growth: A Bioeconomy for Europe*, Brussels: European Comission.

Fang, J. & Wei, Y., 2011. Expression, purification and characterization of the *Escherichia coli* integral membrane protein YajC. *Protein and Peptide Letters*, 18(6), pp.601–8.

Feng, J., Gu, Y., Quan, Y., Zhang, W., Cao, M., Gao, W., et al., 2015. Recruiting a new strategy to improve levan production in *Bacillus amyloliquefaciens*. *Scientific reports*, 5, p.13814.

Freudl, R., 1992. Protein secretion in Gram-positive bacteria. *Journal of Biotechnology*, 23(3), pp.231–240.

Freymond, P.P., Lazarevic, V., Soldo, B. & Karamata, D., 2006. Poly(glucosyl-N-acetylgalactosamine 1-phosphate), a wall teichoic acid of *Bacillus subtilis* 168: Its biosynthetic pathway and mode of attachment to peptidoglycan. *Microbiology*, 152(6), pp.1709–1718.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D. & Bairoch, A., 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research*, 31(13), pp.3784–8.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., et al., 2005. Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook*, pp.571–607.

Gibson, D.G., 2011. Enzymatic assembly of overlapping DNA fragments. *Methods in Enzymology*, 498, pp.349–361.

Gibson, D.G., Young, L., Chuang, R., Venter, J.C., Hutchison, C. a & Smith, H.O., 2009.

Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5), pp.343–5.

Hager, P.W. & Rabinowitz, J.C., 1985. Inefficient Translation of T7 Late mRNA by Bacillus subtilis Ribosomes. *The Journal of Biological Chemistry*, 260(28), pp.15163–15167.

Härtl, B., Wehrl, W., Wiegert, T., Homuth, G. & Schumann, W., 2001. Development of a new integration site within the *Bacillus subtilis* chromosome and construction of compatible expression cassettes. *Journal of Bacteriology*, 183(8), pp.2696–2699.

Harwood, C.R., 1992. *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. *Trends in Biotechnology*, 10(7), pp.247–256.

Harwood, C.R. & Cranenburgh, R., 2008. *Bacillus* protein secretion: an unfolding story. *Trends in Microbiology*, 16(2), pp.73–79.

Heng, C., Chen, Z., Du, L. & Lu, F., 2005. Expression and Secretion of an Acid-Stable α-Amylase Gene in *Bacillus Subtilis* by SacB Promoter and Signal Peptide. *Biotechnology Letters*, 27(21), pp.1731–1737.

Herbort, M., Klein, M., Manting, E.H., Driessen, A.J.M., Freudl, R., Arnold, J.M., et al., 1999. Temporal Expression of the Bacillus subtilis secA Gene , Encoding a Central Component of the Preprotein Translocase. *Journal of bacteriology*, 181(2), pp.493–500.

Hyyrylainen, H.L., Bolhuis, A., Darmon, E., Muukkonen, L., Koski, P., Vitikainen, M., et al., 2001. A novel two-component regulatory system in *Bacillus subtilis* for the survival of severe secretion stress. *Molecular Microbiology*, 41(5), pp.1159–1172.

Hyyrylainen, H.-L., Hyyryläinen, H.L., Vitikainen, M., Thwaite, J., Wu, H., Sarvas, M., et al., 2000. D-alanine substitution of teichoic acids as a modulator of protein folding and stability at the cytoplasmic membrane/cell wall interface of *Bacillus subtilis*. *Journal of Biological Chemistry*, 275(35), pp.26696–26703.

Ikemura, T., 1985. Codon usage and tRNA content in unicellular and multicellular organisms. *Molecular biology and evolution*, 2(1), pp.13–34.

Jensen, C.L., Stephenson, K., Jorgensen, S.T. & Harwood, C., 2000. Cell-associated degradation

affects the yield of secreted engineered and heterologous proteins in the *Bacillus subtilis* expression system. *Microbiology*, 146(10), pp.2583–2594.

Jeong, S.M., Yoshikawa, H. & Takahashi, H., 1993. Isolation and characterization of the secE homologue gene of *Bacillus subtilis*. *Molecular Microbiology*, 10(1), pp.133–142.

Jolliffe, L.K., Doyle, R.J. & Streips, U.N., 1980. Extracellular proteases modify cell wall turnover in *Bacillus subtilis*. *Journal of Bacteriology*, 141(3), pp.1199–1208.

Jongbloed, J.D.H., Grieger, U., Antelmann, H., Hecker, M., Nijland, R., Bron, S., et al., 2004. Two minimal Tat translocases in *Bacillus*. *Molecular Microbiology*, 54(5), pp.1319–1325.

Jongbloed, J.D.H., Martin, U., Antelmann, H., Hecker, M., Tjalsma, H., Venema, G., et al., 2000. TatC is a specificity determinant for protein secretion via the twin-arginine translocation pathway. *Journal of Biological Chemistry*, 275(52), pp.41350–41357.

Jongbloed, J.D.H., Van Der Ploeg, R. & Van Dijl, J.M., 2006. Bifunctional TatA subunits in minimal Tat protein translocases. *Trends in Microbiology*, 14(1), pp.2–4.

Kakeshita, H., Kakeshtia, H., Kageyama, Y., Ara, K., Ozaki, K. & Nakamura, K., 2010. Enhanced extracellular production of heterologous proteins in *Bacillus* subtilis by deleting the C-terminal region of the SecA secretory machinery. *Molecular Biotechnology*, 46(3), pp.250–7.

Kodama, T., Manabe, K., Kageyama, Y., Liu, S., Ara, K., Ozaki, K., et al., 2012. Approaches for Improving Protein Production in Multiple Protease-Deficient *Bacillus subtilis* Host Strains. In *Advances in Applied Biotechnology*. InTech.

Kolkman, M., van der Ploeg, R., Bertels, M., van Dijk, M., van der Laan, J., Maarten van Dijl, J., et al., 2008. The Twin-Arginine Signal Peptide of Bacillus subtilis YwbN Can Direct either Tat-or Sec-Dependent Secretion of Different Cargo Proteins: Secretion of Active Subtilisin via the *B. subtilis* Tat Pathway. *Applied and Environmental Microbiology*, 74(24), pp.7507–7513.

Kontinen, V.P., Saris, P. & Sarvas, M., 1991. A gene (*prsA*) of Bacillus subtilis involved in a novel, late stage of protein export. *Molecular Microbiology*, 5(5), pp.1273–1283.

Kontinen, V.P. & Sarvas, M., 1993. The PrsA lipoprotein is essential for protein secretion in *Bacillus subtilis* and sets a limit for high-level secretion. *Molecular Microbiology*, 8(4), pp.727–37.

Krishnappa, L., Dreisbach, A., Otto, A., Goosens, V.J., Cranenburgh, R.M., Harwood, C.R., et al., 2013. Extracytoplasmic proteases determining the cleavage and release of secreted proteins, lipoproteins, and membrane proteins in *Bacillus subtilis*. *Journal of Proteome Research*, 12(9), pp.4101–4110.

Krishnappa, L., Monteferrante, C.G., Neef, J., Dreisbach, A. & Van Dijl, J.M., 2014. Degradation of extracytoplasmic catalysts for protein folding in *Bacillus subtilis*. *Applied and Environmental Microbiology*, 80(4), pp.1463–1468.

Kumamoto, C., 1989. *Escherichia coli* SecB protein associates with exported protein precursors in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 86(14), pp.5320–5324.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, a M., Alloni, G., Azevedo, V., et al., 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature*, 390(6657), pp.249–256.

Lam, K.H., Chow, K.C. & Wong, W.K., 1998. Construction of an efficient *Bacillus subtilis* system for extracellular production of heterologous proteins. *Journal of biotechnology*, 63(3), pp.167–77.

Lee, H.C. & Bernstein, H.D., 2001. The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proceedings of the National Academy of Sciences of the United States of America*, 98(6), pp.3471–3476.

Leloup, L., Driessen, A.J.M., Freudl, R., Chambert, R.G. & Oise Petit-Glatron, M.-F., 1999. Differential Dependence of Levansucrase and α-Amylase Secretion on SecA (Div) during the Exponential Phase of Growth of *Bacillus subtilis*. *Journal of Bacteriology*, 181(6), pp.1820–1826.

Lesuisse, E., Schanck, K. & Colson, C., 1993. Purification and preliminary characterization of

the extracellular lipase of *Bacillus subtilis* 168, an extremely basic pH-tolerant enzyme. *European journal of biochemistry*, 216(1), pp.155–60.

Line Fabret, C. & Hoch, J.A., 1998. A Two-Component Signal Transduction System Essential for Growth of *Bacillus subtilis*: Implications for Anti-Infective Therapy. *Journal of Bacteriology*, 180(23), pp.6375–6383.

Livak, K.J. & Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT Method. *Methods*, 25(4), pp.402–8.

Lycklama A Nijeholt, J.A., De Keyzer, J., Prabudiansyah, I. & Driessen, A.J.M., 2013. Characterization of the supporting role of SecE in protein translocation. *FEBS Letters*, 587(18), pp.3083–3088.

m2p, 2015. Biolector Application Notes.

Manting, E.H., van Der Does, C., Remigy, H., Engel, A. & Driessen, A.J., 2000. SecYEG assembles into a tetramer to form the active protein translocation channel. *The EMBO journal*, 19(5), pp.852–861.

Margot, P. & Karamata, D., 1996. The wprA gene of *Bacillus subtilis* 168, expressed during exponential growth, encodes a cell-wall-associated protease. *Microbiology*, 142, pp.3437–3444.

Miller, G.L., 1959. Use of DinitrosaIicyIic Acid Reagent for Determination of Reducing Sugar. *Analytical Chemistry*, (lll).

Molecular Probes, 2007. EnzChek ® Ultra Xylanase Assay Kit. *Invitrogen detection technologies*.

Morimoto, T., Loh, C., Hirai, T., Asai, K., Kobayashi, K., Moriya, S., et al., 2002. Six GTP-binding proteins of the Era/Obg family are essential for cell growth in *Bacillus subtilis*. *Microbiology*, 148, pp.3539–3552.

Mulder, K.C.L., Bandola, J. & Schumann, W., 2013. Construction of an artificial secYEG operon allowing high level secretion of α-amylase. *Protein Expression and Purification*, 89(1), pp.92–96.

Müller, J.P., Bron, S., Venema, G. & Van Dijl, J.M., 2000. Chaperone-like activities of the CsaA protein of *Bacillus subtilis*. *Microbiology*, 146(1), pp.77–88.

Müller, J.P., Ozegowski, J., Vettermann, S., van Wely, K.H.M. & Driessen, A.J.M., 2000. Interaction of Bacillus subtilis CsaA with SecA and precursor proteins. *Biochemical Journal*, 348(2), p.367.

Mullis, K.B. & Faloona, F.A., 1987. Specific synthesis of DNA in vitro via a polymerase-catalysed chain reaction. *Methods in Enzymology*, 155(155), pp.335–350.

Murashima, K., Chen, C.-L., Kosugi, A., Tamaru, Y., Doi, R.H. & Wong, S.-L., 2002. Heterologous production of *Clostridium cellulovorans* engB, using protease-deficient Bacillus subtilis, and preparation of active recombinant cellulosomes. *Journal of bacteriology*, 184(1), pp.76–81.

Noone, D., Howell, A., Collery, R. & Devine, K.M., 2001. YkdA and YvtA, HtrA-like serine proteases in *Bacillus subtilis*, engage in negative autoregulation and reciprocal cross-regulation of ykdA and yvtA gene expression. *Journal of Bacteriology*, 183(2), pp.654–663.

Noone, D., Howell, A. & Devine, K.M., 2000. Expression of ykdA, encoding a *Bacillus subtilis* homologue of HtrA, is heat shock inducible and negatively autoregulated. *Journal of Bacteriology*, 182(6), pp.1592–1599.

Oates, J., Barrett, C.M.L., Barnett, J.P., Byrne, K.G., Bolhuis, A. & Robinson, C., 2005. The *Escherichia coli* twin-arginine translocation apparatus incorporates a distinct form of TatABC complex, spectrum of modular TatA complexes and minor TatAB complex. *Journal of Molecular Biology*, 346(1), pp.295–305.

Olmos-Soto, J. & Contreras-Flores, R., 2003. Genetic system constructed to overproduce and secrete proinsulin in Bacillus subtilis. *Applied Microbiology and Biotechnology*, 62(4), pp.369–373.

Or, E., Navon, A. & Rapoport, T., 2002. Dissociation of the dimeric SecA ATPase during protein translocation across the bacterial membrane. *EMBO Journal*, 21(17), pp.4470–4479.

Palva, I., 1982. Molecular cloning of alpha-amylase gene from *Bacillus amyloliquefaciens* and its expression in *B. subtilis*. *Gene*, 19(1), pp.81–7.

Palva, I., Lehtovaara, P., Kääriäinen, L., Sibakov, M., Cantell, K., Schein, C.H., et al., 1983. Secretion of interferon by Bacillus subtilis. *Gene*, 22(2–3), pp.229–35.

Petersen, T.N., Brunak, S., von Heijne, G. & Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, 8(10), pp.785–786.

Du Plessis, D.J.F., Nouwen, N. & Driessen, A.J.M., 2011. The Sec translocase. *Biochimica et Biophysica Acta - Biomembranes*, 1808(3), pp.851–865.

Ploss, T.N., Reilman, E., Monteferrante, C.G., Denham, E.L., Piersma, S., Lingner, A., et al., 2016. Homogeneity and heterogeneity in amylase production by *Bacillus subtilis* under different growth conditions. *Microbial Cell Factories*, 15(1), p.57.

Plotkin, J.B. & Kudla, G., 2011. Synonymous but not the same: the causes and consequences of codon bias. *Nature*, 12, pp.32–42.

Pohl, S., Bhavsar, G., Hulme, J., Bloor, A.E., Misirli, G., Leckenby, M.W., et al., 2013. Proteomic analysis of *Bacillus subtilis* strains engineered for improved production of heterologous proteins. *Proteomics*, 13(22), pp.3298–3308.

Pohl, S. & Harwood, C.R., 2010. Heterologous Protein Secretion by *Bacillus* Species: From the Cradle to the Grave. *Advances in Applied Microbiology*, 73, pp.1–25.

Polaina, J. & MacCabe, A.P., 2007. *Industrial Enzymes: Structure, Function and Applications* J. Polaina & A. P. MacCabe, eds., Springer Science & Business Media.

Pomerantsev, A.P., Pomerantseva, O.M., Moayeri, M., Fattah, R., Tallant, C. & Leppla, S.H., 2011. A *Bacillus anthracis* strain deleted for six proteases serves as an effective host for production of recombinant proteins. *Protein Expression and Purification*, 80(1), pp.80–90.

Pop, O., Martin, U., Abel, C. & Müller, J.P., 2002. The twin-arginine signal peptide of PhoD and the TatAd/Cd proteins of *Bacillus subtilis* form an autonomous tat translocation system. *Journal of Biological Chemistry*, 277(5), pp.3268–3273.

Prabudiansyah, I. & Driessen, A.J.M., 2016. The Canonical and Accessory Sec System of Gram-positive Bacteria. *Current topics in microbiology and immunology*.

Quax, W.J. & Broekhuizen, C.P., 1994. Development of a new *Bacillus* carboxyl esterase for use

in the resolution of chiral drugs. *Appiled Microbiology Biotechnology*, 41, pp.425–431.

Quentin, Y., Fichant, G. & Denizot, F., 1999. Inventory, assembly and analysis of *Bacillus subtilis* ABC transport systems. *Journal of Molecular Biology*, 287(3), pp.467–484.

Rocha, E.P., Danchin, A. & Viari, A., 1999. Translation in *Bacillus subtilis*: roles and trends of initiation and termination, insights from a genome analysis. *Nucleic acids research*, 27(17), pp.3567–76.

Rojas Contreras, J.A., Pedraza-Reyes, M., Ordoñez, L.G., Estrada, N.U., Barba de la Rosa, A.P. & De León-Rodríguez, A., 2010. Replicative and integrative plasmids for production of human interferon gamma in *Bacillus subtilis*. *Plasmid*, 64(3), pp.170–176.

Sachelaru, I., Petriman, N.A., Kudva, R., Kuhn, P., Welte, T., Knapp, B., et al., 2013. YidC occupies the lateral gate of the SecYEG translocon and is sequentially displaced by a nascent membrane protein. *Journal of Biological Chemistry*, 288(23), pp.16295–16307.

Sadiq Butt, M., Tahir-Nadeem, M., Ahmad, Z. & Tauseef Sultan, M., 2008. Xylanases and Their Applications in Baking Industry. *Food Technology & Biotechnology*, 46(1), pp.22–31.

Salis, H.M., Mirsky, E. a & Voigt, C. a, 2010. Automated Design of Synthetic Ribosome Binding Sites to Precisely Control Protein Expression. *Nature Biotechnology*, 27(10), pp.946–950.

Sambrook, J. & Russell, D.W., 2001. *Molecular Cloning - Sambrook & Russel*,

Sarvas, M., Harwood, C.R., Bron, S. & Van Dijl, J.M., 2004. Post-translocational folding of secretory proteins in Gram-positive bacteria. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1694(1–3 SPEC.ISS.), pp.311–327.

Sauer, C., Syvertsson, S., Bohorquez, L.C., Cruz, R., Harwood, C.R., van Rij, T., et al., 2016. Effect of Genome Position on Heterologous Gene Expression in *Bacillus subtilis*: An Unbiased Analysis. *ACS Synthetic Biology*, 5(9), pp.942–947.

Schmidt, B.F., Jarnagin, A.S. & Ferrari, E., 1993. Commercial Production of Extracellular Enzymes. In *Bacillus subtilis and Other Gram-Positive Bacteria*. American Society of Microbiology, pp. 917–937.

Serizawa, M., Kodama, K., Yamamoto, H., Kobayashi, K., Ogasawara, N. & Sekiguchi, J., 2005.

Functional analysis of the YvrGHb two-component system of *Bacillus subtilis*: identification of the regulated genes by DNA microarray and northern blot analyses. *Bioscience, Biotechnology, and Biochemistry*, 69(11), pp.2155–2169.

Setlow, B. & Setlow, P., 1980. Measurements of the pH within dormant and germinated bacterial spores. *Proceedings of the National Academy of Sciences of the United States of America*, 77(5), pp.2474–6.

Shinde, U., Li, Y., Chatterjee, S. & Inouye, M., 1993. Folding pathway mediated by an intramolecular chaperone. *Biochemistry*, 90, pp.6924–6928.

Shine, J. & Dalgarno, L., 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America*, 71(4), pp.1342–6.

de Souza, P.M. & de Oliveira Magalhães, P., 2010. Application of microbial α-amylase in industry - A review. *Brazilian Journal of Microbiology*, 41(4), pp.850–61.

Spiess, C., Beil, A., Ehrmann, M., Arslan, E., Schulz, H., Zufferey, R., et al., 1999. A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. *Cell*, 97(3), pp.339–347.

Stephenson, K., Bron, S. & Harwood, C.R., 1999. Cellular lysis in *Bacillus subtilis*; the affect of multiple extracellular protease deficiencies. *Letters in Applied Microbiology*, 29(2), pp.141–145.

Stephenson, K., Carter, N.M., Harwood, C.R., Petit-Glatron, M.F. & Chambert, R., 1998. The influence of protein folding on late stages of the secretion of alpha-amylases from *Bacillus subtilis*. *FEBS Letterstters*, 430(3), pp.385–9.

Stephenson, K. & Harwood, C.R., 1998. Influence of a cell-wall-associated protease on production of alpha-amylase by *Bacillus subtilis*. *Applied and Environmental Microbiology*, 64(8), pp.2875–2881.

Stephenson, K., Jensen, C.L., Jorgensen, S.T., Lakey, J.H., Harwood, C.R. & Jørgensen, S.T., 2000. The influence of secretory-protein charge on late stages of secretion from the Gram-positive bacterium *Bacillus subtilis*. *The Biochemical Journal*, 350, pp.31–39.

Stewart, C.R., Gaslightwala, I., Hinata, K., Krolikowski, K. a, Needleman, D.S., Peng, a S., et al., 1998. Genes and regulatory sites of the "host-takeover module" in the terminal redundancy of *Bacillus subtilis* bacteriophage SPO1. *Virology*, 246(2), pp.329–340.

Stothard, P., 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques*, 28(6).

Suh, J.W., Boylan, S.A., Thomas, S.M., Dolan, K.M., Oliver, D.B. & Price, C.W., 1990. Isolation of a secY homologue from *Bacillus subtilis*: evidence for a common protein export pathway in eubacteria. *Molecular Microbiology*, 4(2), pp.305–314.

Suh, T.W., Boylan, S.A., Oh, S.H. & Price, C.W., 1996. Genetic and transcriptional organization of the *Bacillus subtilis* spc-α region. *Gene*, 169(1), pp.17–23.

Takamatsu, H., Fuma, S., Nakamura, K., Sadaie, Y., Shinkai, A., Matsuyama, S., et al., 1992. *In vivo* and *in vitro* characterization of the secA gene product of Bacillus subtilis. *Journal of Bacteriology*, 174(13), pp.4308–4316.

Tjalsma, H., Bolhuis, A., Jongbloed, J.D., Bron, S. & van Dijl, J.M., 2000. Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiology and Molecular Biology Reviews*, 64(3), pp.515–47.

Tjalsma, H., Bolhuis, A., van Roosmalen, M.L., Wiegert, T., Schumann, W., Broekhuizen, C.P., et al., 1998. Functional analysis of the secretory precursor processing machinery of *Bacillus subtilis*: identification of a eubacterial homolog of archaeal and eukaryotic signal peptidases. *Genes & development*, 12(15), pp.2318–31.

Tjalsma, H., Noback, M.A., Bron, S., Venema, G., Yamane, K. & Van Dijl, J.M., 1997. Bacillus subtilis contains four closely related type I signal peptidases with overlapping substrate specificities. Constitutive and temporally controlled expression of different *sip* genes. *Journal of Biological Chemistry*, 272(41), pp.25983–25992.

Tjalsma, H., Noback, M. a, Venema, G., Yamane, K., Bron, S. & Dijl, J.M. Van, 1997. Cell Biology and metabolism : *Bacillus subtilis* Contains Four Closely Related Type I Signal Peptidases with Overlapping Substrate Specificities. *Journal of Biological Chemistry*, 272(41), pp.25983–25992.

Tossavainen, H., Permi, P., Purhonen, S.L., Sarvas, M., Kilpelainen, I. & Seppala, R., 2006. NMR solution structure and characterization of substrate binding site of the PPIase domain of PrsA protein from *Bacillus subtilis*. *FEBS Letters*, 580(7), pp.1822–1826.

Trip, H., van der Veek, P.J., Renniers, T.C., Meima, R., Sagt, C.M., Mohrmann, L., et al., 2011. A novel screening system for secretion of heterologous proteins in *Bacillus subtilis*. *Microbial Biotechnology*, 4, pp.673–682.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. & Leunissen, J.A.M., 2007. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research*, 35.

Vellanoweth, R.L. & Rabinowitz, J.C., 1992. The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo. *Molecular Microbiology*, 6(9), pp.1105–1114.

Vitikainen, M., Hyyrylainen, H.L., Kivimaki, A., Kontinen, V.P. & Sarvas, M., 2005. Secretion of heterologous proteins in *Bacillus subtilis* can be improved by engineering cell components affecting post-translocational protein folding and degradation. *Journal of Applied Microbiology*, 99(2), pp.363–375.

Vitikainen, M., Lappalainen, I., Seppala, R., Antelmann, H., Boer, H., Taira, S., et al., 2004. Structure-Function Analysis of PrsA Reveals Roles for the Parvulin-like and Flanking N- and C-terminal Domains in Protein Folding and Secretion in *Bacillus subtilis*. *Journal of Biological Chemistry*, 279(18), pp.19302–19314.

Vitikainen, M., Pummi, T., Airaksinen, U., Wahlström, E., Wu, H., Sarvas, M., et al., 2001. Quantitation of the capacity of the secretion apparatus and requirement for PrsA in growth and secretion of α-amylase in *Bacillus subtilis*. *Journal of Bacteriology*, 183(6), pp.1881–1890.

Vollmer, W., Blanot, D. & De Pedro, M.A., 2008. Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*, 32(2), pp.149–167.

Wang, G., Xia, Y., Gu, Z., Zhang, H., Chen, Y.Q., Chen, H., et al., 2015. A new potential secretion pathway for recombinant proteins in *Bacillus subtilis*. *Microbial Cell Factories*, 14(1), p.179.

Weidenmaier, C. & Peschel, A., 2008. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nature Reviews Microbiology*, 6(1740–1534), pp.276–287.

van Wely, K.H., Swaving, J., Freudl, R. & Driessen, a J., 2001. Translocation of proteins across the cell envelope of Gram-positive bacteria. *FEMS Microbiology Reviews*, 25(4), pp.437–454.

Van Wely, K.H.M., Swaving, J., Broekhuizen, C.P., Rose, M., Quax, W.J. & Driessen, A.J.M., 1999. Functional identification of the product of the *Bacillus subtilis* yvaL gene as a SecG homologue. *Journal of Bacteriology*, 181(6), pp.1786–1792.

Westers, H., Darmon, E., Zanen, G., Veening, J.W., Kuipers, O.P., Bron, S., et al., 2004. The *Bacillus* secretion stress response is an indicator for α-amylase production levels. *Letters in Applied Microbiology*, 39(1), pp.65–73.

Westers, H., Westers, L., Darmon, E., Van Dijl, J.M., Quax, W.J. & Zanen, G., 2006. The CssRS two-component regulatory system controls a general secretion stress response in *Bacillus subtilis*. *FEBS Journal*, 273(16), pp.3816–3827.

Westers, L., Dijkstra, D.S., Westers, H., van Dijl, J.M. & Quax, W.J., 2006. Secretion of functional human interleukin-3 from *Bacillus subtilis*. *Journal of Biotechnology*, 123(2), pp.211–24.

Westers, L., Westers, H. & Quax, W.J., 2004. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1694(1), pp.299–310.

Williams, R.C., Rees, M.L., Jacobs, M.F., Prágai, Z., Thwaite, J.E., Baillie, L.W.J., et al., 2003. Production of *Bacillus anthracis* protective antigen is dependent on the extracellular chaperone, PrsA. *Journal of Biological Chemistry*, 278(20), pp.18056–62.

Wu, S.-C. & Wong, S.-L., 2002. Engineering of a *Bacillus subtilis* strain with adjustable levels of intracellular biotin for secretory production of functional streptavidin. *Applied and environmental microbiology*, 68(3), pp.1102–8.

Wu, S.C., Ye, R., Wu, X.C., Ng, S.C. & Wong, S.L., 1998. Enhanced secretory production of a

single-chain antibody fragment from *Bacillus subtilis* by coproduction of molecular chaperones. *Journal of Bacteriology*, 180(11), pp.2830–5.

Wu, S.-C., Yeung, J.C., Duan, Y., Ye, R., Szarka, S.J., Habibi, H.R., et al., 2002. Functional production and characterization of a fibrin-specific single-chain antibody fragment from *Bacillus subtilis*: effects of molecular chaperones and a wall-bound protease on antibody fragment production. *Applied and Environmental Microbiology*, 68(7), pp.3261–9.

Wu, X.C., Lee, W., Tran, L. & Wong, S.L., 1991. Engineering a *Bacillus subtilis* expression-secretion system with a strain deficient in six extracellular proteases. *Journal of Bacteriology*, 173(16), pp.4952–4958.

Yabuta, Y., Takagi, H., Inouye, M. & Shinde, U., 2001. Folding pathway mediated by an intramolecular chaperone: Propeptide release modulates activation precision of pro-subtilisin. *Journal of Biological Chemistry*, 276(48), pp.44427–44434.

Yang, C.-K., Ewis, H.E., Zhang, X., Lu, C.-D., Hu, H.-J., Pan, Y., et al., 2011. Nonclassical protein secretion by *Bacillus subtilis* in the stationary phase is not due to cell lysis. *Journal of bacteriology*, 193(20), pp.5607–15.

Yang, C.-K., Lu, C.-D. & Tai, P.C., 2013. Differential expression of secretion machinery during bacterial growth: SecY and SecF decrease while SecA increases during transition from exponential phase to stationary phase. *Current microbiology*, 67(6), pp.682–7.

Zanen, G., Antelmann, H., Westers, H., Hecker, M., Van Dijl, J.M. & Quax, W.J., 2004. FlhF, the third signal recognition particle-GTPase of *Bacillus subtilis*, is dispensible for protein secretion. *Journal of Bacteriology*, 186(17), pp.5956–5960.

Zanen, G., Houben, E.N.G., Meima, R., Tjalsma, H., Jongbloed, J.D.H., Westers, H., et al., 2005. Signal peptide hydrophobicity is critical for early stages in protein export by *Bacillus subtilis*. *FEBS Journal*, 272, pp.4617–4630.