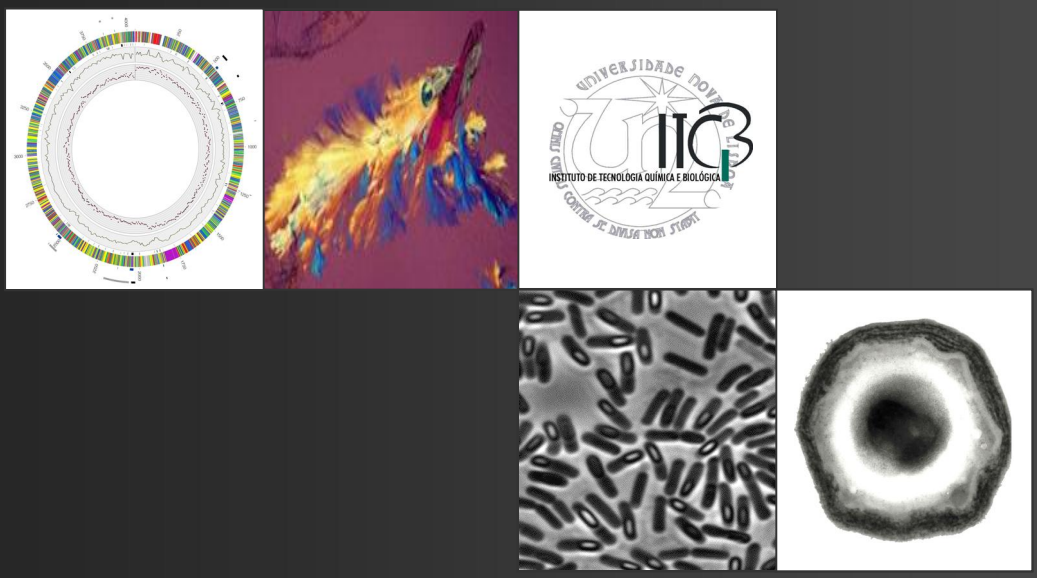
Metabolic and Morphogenetic Engineering of *Bacillus subtilis*: Biotechnology for Industry

Sébastien Potot



Dissertation presented to obtain the Ph.D. degree in "Biology"
at the Institute of Chemical and Biological Technology
of the New University of Lisbon

Dissertação apresentada para obtenção
do grau de Doutor em "Biologia"
pelo Instituto de Tecnologia Química e Biológica
da Universidade Nova de Lisboa

Oeiras, November 2010

Metabolic and Morphogenetic Engineering of *Bacillus subtilis*: Biotechnology for Industry

Cover

From the top-left to the bottom-right, pictures represent: circular map of the chromosome of *B. subtilis* BSP1 (Schyns et *al.*, manuscript in preparation; see references Chapter V); crystals of vitamin B1 (DSM Nutritional Products picture gallery); phase-contrast microcopy of *B. subtilis* sporulating culture (courtesy of A.O. Henriques, ITQB-NL); electron micrograph of a cross-section of *B. subtilis* spore (Henriques and Moran, 2007; see references Chapter I).

ISBN number

Applied for.

Supervisor

Dr Ghislain Schyns
DSM Nutritional Products.
P.O. Box 2676, 203/25. CH 4002 Basel, Switzerland
+41. 618158315, email: ghislain.schyns@dsm.com

Co-Supervisor

Prof. Dr Adriano O. Henriques
Institute of Chemical and Biological Technology
(Instituto de Tecnologia Química e Biológica), Microbial Development.
Universidade Nova de Lisboa. Apartado 127. 2781-901 Oeiras, Portugal.
+351.214449521, email: aoh@itqb.unl.pt

Jury

Dr John B. Perkins
DSM Food Specialties.
P.O. Box 1. 2600 MA Delft, The Netherlands
+31.152792347, email: john.perkins@dsm.com

Prof. Dr Isabel Sá Nogueira
Faculty of Sciences and Technology
(Faculdade de Ciências e Tecnologia)
Universidade Nova de Lisboa. 2829-516 Caparica, Portugal
+351. 212948530, email: isn@fct.unl.pt

Prof. Dr Jan Maarten Van Dijl
University Medical Center Groningen, Department of Medical Microbiology.
Hanzeplein 1, P.O. Box 30001. 9700 RB Groningen, The Netherlands
+31.503615187, email: j.m.van.dijl@med.umcg.nl

Author

Sébastien Potot
1 rue de l'hopital. F 68730 Blotzheim, France
+33.661555016, email : sebastien_potot@yahoo.fr

Acknowledgments

First of all, I would like to thank the Instituto de Tecnologia Química e Biológica (ITQB) in Oeiras, in particular the Microbial Development group, for providing all the conditions leading to this thesis and allowing me to follow the 2008-09 PhD program in Chemical and Biological Sciences and Engineering. The studies described in this thesis were financially supported by DSM Nutritional Products Ldt., Basel (Switzerland). Some aspects of the work in Oeiras were partially financed through Grant PTDC/BIO/73946/2006 from Fundação para a Ciência e a Tecnologia (F.C.T) to A.O.H.

I express my gratitude to my supervisor Ghislain Schyns, for being my mentor during these years of collaboration. Especially, I would like to acknowledge his precious advices, the very helpful discussions, and over all his unbendable support that rose my self-confidence. I also appreciated the number of nice moments (and delectable bottles of wine) we shared so far.

I'm very grateful to Adriano O. Henriques for being my co-supervisor and offering me this unique opportunity to integrate the dynamic ITQB to perform this Ph.D. work. I particularly thank him for sharing his well-recognized expertise in the field of sporulation, and for ensuring the link between academic and industrial research through our close collaboration. I also thank him for the administrative support.

I would like to thank Isabel Sá Nogueira, Jan Maarten Van Dijl, and John B. Perkins for the reading and evaluation of my thesis.

Many thanks also to Roman Imhof, Monique Musch, Zoltán Prágai, Dietmar Laudert, and Hans-Peter Hohmann for supporting the set-up and implementation of my Ph.D work in DSM, and to Jochen Pheiffer and Antoine de Saizieu of the Animal Nutrition and Health cluster for giving me “freedom to operate”.

I acknowledge also the warm welcome of the Microbial Development group in Oeiras, in particular Claudia R. Serra for driving my first steps with the spores. I also appreciated the useful English grammar checks by Jihane Achkar and Delphine Acker-Lagarde, and the editorial tips of Thomas Lapointe.

Last but not least, I would like to thank my wife Céline, for taking care of so many things at home when I was not available, and express my love to her and our little Enzo. Many thanks also to my parents, for their financial efforts to let me access university studies. Finally, my special thanks go to *Bacillus subtilis* for making a pretty good job once again!

Sébastien Potot.

Abstract in English

The bacterium *Bacillus subtilis* plays a major role as model for Gram-positive research, and is widely used as industrial workhorse for the production of high-value enzymes and metabolites. Successful development of commercially-attractive strains and of original products is the key of innovation in microbial industry. Evolving a natural *B. subtilis* strain into an overproducer involves genetic engineering. In spite of a wide range of genetic tools and techniques available for *B. subtilis,* engineering of the metabolic pathways, with the purpose to reach industrially significant amount of products, requires a deep and specific understanding of the metabolism. This thesis describes two aspects of genetic engineering applications in biotechnology research. On the one hand, metabolic engineering aiming to develop a direct "sugar to metabolite" microbial process is described. More precisely, both the exploration of thiamin (vitamin B1) biosynthesis, resulting in the elucidation of a new salvage pathway, and the identification and characterization of relevant mutations in thiamin-regulated genes (*thiN*, *yuaJ*, and *ykoD*) are covered. On the other hand, this thesis gives a glimpse of how structural engineering of *B. subtilis* spores will allow to vehicle recombinant proteins of interest at their surface for putative probiotic applications. The spore display system presented is based on the original use of a coat-associated enzyme (oxalate decarboxylase) as original anchoring motif. Taken together, the described studies constitute a basis to draw perspectives in the future of *B. subtilis* research in biotechnology industry.

Abstract in Portuguese

Bacillus subtilis é um organismo modelo central no estudo de bactérias de Gram-positiva, além da sua importância na produção industrial de enzimas e de metabolitos de valor comercial. O desenvolvimento com sucesso de estirpes de valor comercial acrescido bem como de produtos originais é a chave para a inovação na indústria biotecnológica com base na utilização de microorganismos. O desenvolvimento de uma estirpe natural de *B. subtlis* num organismo super-produtor envolve engenharia genética. Apesar do vasto leque de ferramentas genéticas e outras, disponíveis para a manipulação de *B. subtilis*, a engenharia de uma via metabólica com a finalidade de se atingirem níveis industrialmente atractivos de produtos de interesse, requer em última análise um conhecimento profundo e específico do organismo. Nesta Tese descrevemos duas aplicações da engenharia genética à investigação biotecnológica. Por um lado, são descritos dois exemplos de engenharia metabólica com o objectivo de desenvolver processos de base microbiana de tipo "nutriente a metabolito". Mais especificamente, cobrimos aqui a exploração da biosíntese de tiamina (vitamina B1), um trabalho que resultou na elucidação de uma via de recuperação, e na identificação e caracterização de mutações importantes em genes regulados pela tiamina (*thiN*, *yuaJ*, e *ykoD*). Por outro lado, esta tese ilustra de que modo a engenharia estrutural dos esporos de *B. subtilis* permite a sua utilização enquanto veículos para a apresentação funcional de proteínas recombinantes de interese, com vista a aplicações futuras em probiose. O sistema de apresentação de proteínas à superfície do esporo que

descrevemos tem por base a utilização de uma descarboxilase do oxalato, normalmente um componente da estrutura superficial do esporo (ou manto), como um motivo de ancoramento original. Em conjunto, os estudos aqui descritos constituem uma base para perspectivar investigação e utilizações futuras de *B. subtilis* na indústria biotecnológica.

Contents

CHAPTER I

General Introduction

1. *B. subtilis* origins, ecology and genome features

For more than a century, the rod-shaped endospore forming *Bacillus subtilis* has been a model of the Firmicutes, a major phylum of bacteria. It became a reference for Gram-positive microorganisms, due to its amenability to genetic manipulation (rapid growth, high natural competence for DNA uptake, stable integration of exogenous DNA into the chromosome, and safety for experimenters and environment). The accessibility to the well-annotated genome sequence in addition to the wide availability of physiological and biochemical data make of *B. subtilis* a workhorse for both basic research and industrial purposes. Reflecting its central role in Research, about 25,000 articles in the PubMed reference library (National Center for Biotechnology Information, USA) and more than 1.1 million Google pages refer to *B. subtilis* biology. While sporulation, an archetypical form of cell differentiation, is one of the main topics of *B. subtilis* studies in academic groups, industrial research mainly focuses on the production of commercially attractive levels of small metabolites and enzymes.

Most of the strains used in *B. subtilis* genetics derive from strain 168, a single tryptophan requiring auxotroph bacterium, which is highly competent. Developments of most of the genetic methods and studies of the physiology and sporulation of *B. subtilis* have been based on this strain. Important derivatives from strain 168 are centralized and maintained at the Bacillus Genetic Stock Center (BGSC), created in the seventies in the Ohio State University (USA). Strain 168 was isolated in the late forties at the University of Yale (USA) after mutagenesis of Marburg *B. subtilis* with sublethal doses of X-rays, which resulted in a high frequency of auxotrophy

among the mutants [Burkholder and Giles, 1947]. Although most of those were subsequently lost, five auxotroph strains from this collection, requiring threonine (strain 23), nicotinic acid (strain 122), or tryptophan (strains 160, 166, and 168) are still available [Zeigler *et al*., 2008].

B. subtilis was historically classified as a strict aerobe. However, subsequent research has demonstrated that it can also grow anaerobically, using nitrate or nitrite as alternative terminal electron acceptor [Nakano and Zuber, 1998]. Microbiology textbooks generally indicate soil as the main habitat of *B. subtilis*, like the other aerobic spore-formers. Recent research suggests that this may be an oversimplification. First, studies that examined soil using fluorescent antibodies failed to convincingly prove the existence of vegetative *B. subtilis* in the soil, except if associated with decaying plant matter. Interestingly, *Bacillus* endospores were found in the gut of various insects and animals. A reasonable explanation is that endospores introduced in the gastrointestinal tract by the way of food germinate and proliferate therein [Hong *et al.*, 2009]. Studies performed with murine models demonstrated that *B. subtilis* (strain PY79) is rapidly cleared from mouse gut after only 6 days [Hoa *et al.*, 2001; Le Duc *et al.,* 2004]. Although *B. subtilis* can not rigorously be considered as a gut commensal, the gastrointestinal tract represents an important part of its life cycle. Also supported by the finding of *Bacillus* endospores in diverse other surroundings including rocks, dust, and aquatic environments, the denomination of *B. subtilis* as a pure soil microorganism is not strictly correct anymore.

With the emergence of Genomics in the nineties, many laboratories joined efforts to determine the complete genome

sequence of *B. subtilis* 168, a research model for gram-positive microorganisms. A consortium of 30 groups [Simpson, 2001] sequenced and annotated chromosomal segments covering the whole genome [Harwood and Wipat, 1996]. However, the main drawback of the consortium-generated results was the discrepancies in the sequences. In addition, the methods available at that time required the cloning of *B. subtilis* DNA into *Escherichia coli*, which could often result in the alteration of the original sequence. The regions containing putative errors were therefore re-sequenced, leading to a more accurate genome sequence [Kunst *et al.* 1997]. With the development of fast and accurate sequencing techniques and of advanced annotation platforms, the sequence of *B. subtilis* 168 has been recently updated [Barbe, *et al.* 2009]. First, the Roche 454 technology was used. About 800 peer-reviewed publications reported the use of this high-throughput sequencing technology since its introduction in 2005 [Margulies *et al.*, 2005]. To consolidate the data, the *B. subtilis* genome was then analyzed with the high density sequencing technology developed by Solexa/Illumina in 2006 [Bentley, 2006].

The updated sequence results in more than 78% of identical genes compared to the first one. The 4,216 kb-genome of *B. subtilis* has an average G+C content of 43% and is made of 4,244 protein-coding genes (89.7% of the total size), 30 rRNA and 86 tRNA. 3,662 regions transcriptionally active during mid-exponential growth were reported, accounting for 77.3% of the genes as they are currently annotated [Rasmussen *et al.*, 2009]. A systematic inactivation of *B. subtilis* genes revealed that 271 genes were essential for growth [Kobayashi *et al.*, 2003]. A significant number (171) of new and mostly small genes was annotated in the new sequence. Also, 407 genes of unknown function (whose previous

name started with the letter “y”) were renamed on the basis of experimental evidence from the literature. The coding loci were divided in two categories. The persistent genes encoding ubiquitous functions to survive and perpetuate life are called *paleome*, whereas the genes corresponding to the colonization of particular niches form the *cenome* [Barbe *et al.,* 2009]. A common feature of *Bacillus* spp. is that the G+C content varies considerably throughout the chromosome. In addition, A+T islands are the marker of bacteriophages or transposons. It is therefore hypothesized that the chromosome results from horizontal transfer of DNA from organisms richer in A+T than *B. subtilis* [Tosato and Bruschi, 2004].

The genome sequence of *B. subtilis* has a limited interest *per se*. More interesting are the genomic studies aiming at associating genomic items of the sequence with their functions, which are identified experimentally or predicted *in silico*. Elucidation of the complexity and dynamics of transcriptional activity in the whole-genome notably involves methods based on high-density oligonucleotide tiling array technology. This powerful technology relies on short fragments designed to cover the entire genome or contiguous regions of the genome and fixed onto a solid surface. Labelled target molecules (cDNA) are then hybridized to the unlabeled probes. Relevance of genome-wide investigations with tiling arrays were notably demonstrated with *B. subtilis* [Nicolas *et al.*, 2009].

In the current post-genomic era, studies of DNA topology and secondary structures are also taken in account to understand storage and management of genetic superinformation. Analysis of DNA secondary structure aims to figure out whether the presence of hairpins and bends along the chromosome is correlated to phenomena that primary structure cannot explain, such as preferred

sites of integration in transposon mutagenesis, cellular shape and non-coding DNA. Investigation on DNA supercoiling also adds onto the knowledge of regulation of gene expression. DNA supercoiling is dependant on environmental factors, such as osmolarity, temperature, oxygen, and carbon sources, resulting in activation or repression of gene expression. Gene organization is considered as well, together with DNA secondary structure and supercoiling, because complex interactions exist among different regions of the chromosome [Tosato and Bruschi, 2004]. More specifically, the position of a gene within an operon can have a significant impact on the control of its expression. Exemplifying this, a study on the influence of the position of *sigD*, a gene encoding the sigma D factor of RNA polymerase in *B. subtilis* and driving expression of genes required for cell separation and flagella assembly, was recently published [Cozy and Kearns, 2010]. The part of motile cells reaches 70% when *sigD* is expressed at its original locus within the so-called motility operon, whereas it goes up to 100% when this gene is moved upstream. It is also documented that the position of a gene in the chromosome can have a strong impact on its function [Dworkin and Losick, 2001].

2. Industrial biotechnology with *B. subtilis*

2.1 Advantages/Drawbacks of *B. subtilis* as an industrial workhorse

As a non-pathogenic bacterium (biological hazardous level 1), *B. subtilis* is attractive for industrial purposes for various reasons. First, it enables the development of valuable industrial processes, thanks to its short fermentation cycle times (high growth rates), and relatively inexpensive and easy large-scale fermentation at high cell densities. The single membrane of *B. subtilis* cells enables direct transfer to the growth medium of proteins transported by secretion machineries, which greatly simplify downstream processes. High concentrations of extracellular proteins (up to 25 g/l) can therefore be reached [Perkins *et al.*, 2009], considering that secreted proteins concentrations in the gram per liter range are usually significant. Registration is also simplified, as *B. subtilis* profits of the GRAS (Generally Regarded As Safe) status with the US Food and Drug Administration (FDA). Importantly, the GRAS status is attached to products derived from *B. subtilis* but not the microorganism itself. The QPS (Qualified Presumption of Safety) status, the European equivalent of GRAS, was recently introduced by the EU authorities [EFSA[a], 2007]. Furthermore, proteins produced in *B. subtilis* are free of endotoxin lipopolysaccharide (LPS), a molecule present in the outer membrane of Gram-negative bacteria that must be completely removed before using recombinant proteins for clinical purposes [Zweers *et al.*, 2008].

Despite these relevant advantages, production of heterologous proteins with *B. subtilis* raises some major limitations.

First, expression of at least height cell wall-associated and/or secreted proteases (AprE, Bpf, Epr, Mpr, NprB, NprE, Vpr and WprA; Wu *et al.*, 2002) generates significant degradation of secreted heterologous proteins. Moreover, like many other prokaryotes, *B. subtilis* is not able to glycosylate proteins during their secretion. It is therefore not a producer of choice for pharmaceutical proteins, when total structural authenticity is required (mammalian proteins) [Westers *et al.* 2004].

2.2 Market for industrial compounds produced with *B. subtilis*

Industrial enzymes represent a major area of biotechnology, resulting in the development of a number of new products and improvements in the process and performance of several existing products. Worldwide, about 20 companies share the production of industrial enzymes. The European key players are AB Enzymes GmbH (Germany), BASF AG (Germany), Chr. Hansen AS (Denmark), Danisco AS (Denmark), Direvo Biotech AG (Germany), DSM Food Specialties (The Netherlands), Evonik Industries (Germany) and Novozymes AS (Denmark). It is estimated that *Bacillus* spp. enzymes represent 50% of the total enzyme market. The global market for industrial enzymes is expected to be worth 2 billion euros by 2012. Industrial enzymes produced with *B. subtilis* address various applications. The market is clustered in technical enzymes (*i.e.* enzymes used in textile, detergent, and pulp & paper industries) (50%), food (36%) and feed (14%) enzymes. Current growth of the animal feed enzymes segment is notably supported by the increased use of phytase to improve nutrient value of feedingstuffs and reduce phosphorous pollution in intensive farming

areas [Thakore, 2008]. The main compounds produced from industrial fermentations of *B. subtilis* are summarized in Table 1.

Table 1: Main compounds produced by industrial fermentation of *B. subtilis*. Adapted from Schallmey *et al.*, 2004 and Zeigler and Perkins, 2008.

Products	Industrial applications
Enzymes:	
Alpha-Acetolactate decarboxylase	Beverage
Alpha-Amylase	Food, Paper, Starch, Textile, Brewing
Beta-Glucanase	Beverage
Beta-Glucosidase	Brewing
Cellulase	Detergents
Cyclodextrin glucanotranssferase	Food, Pharma, Cosmetics
Galactomannase	Feed, Beverage
Glutaminase	Food, Flavor
Lipases	Detergent
Neutral (metallo-) Protease	Detergent, Food
Alkaline (serine-) Protease	Detergent, Textile
Penicillin Acylase	Pharma
Pullulanase	Starch, Food, Beverage
Poly-gamma-glutamic acid	Food, Pharma, Cosmetics
Urease	Analysis, Beverage
Xylanases	Baking, Feed, Beverage, Brewing, Food
Fine chemicals:	
D-Ribose	Food, Feed, Cosmetics, Pharma
Poly-gamma-glutamic acid	Food, Feed, Pharma
Purine nucleosides (inosine, guanosine)	Food
Riboflavin	Food, Pharma
Streptavidin	Microarrays
Thaumatin	Food, Pharma
Surfactins	Pharma, Bioremediation
Others:	
Natto (fermented soy)	Food
Animal Probiotics	Feed

Industrial enzymes produced with *B. subtilis* can have interesting properties, like a broad pH optimum and a good thermostability, such as the chimeric cellulase produced with strain KSM-635 [Ito *et al.*, 1989]. In addition, high production titers can be reached, for instance for poly-gamma-glutamic acid produced with strain F-2-01 5 (up to 50 g/l) [Schallmey *et al.* 2004]. Although few vitamins are currently produced by fermentation, extensive strain engineering allows for production of commercially attractive levels of riboflavin (vitamin B2) in *B. subtilis* [Perkins *et al.*, 1999].

3. Genetic engineering in *B. subtilis*

3.1 Targets of metabolic engineering

Metabolic engineering of microorganisms consists in the genetic optimization of biosynthetic pathways toward the overproduction of commercially attractive metabolites. Reaching relevant titers generally involves significant modifications of biosynthetic pathways. This task may be difficult when knowledge is lacking on the exact biochemistry and/or genetics. In addition, severe imbalance in the central metabolism (*e.g.* metabolism of nucleotide bases, aminoacids *etc.*) can have detrimental consequences on the engineered strain. For example, metabolic engineering directed towards high vitamin B5 (pantothenate) production in *B. subtilis* requires manipulation of the genes involved in the synthesis of leucine, valine and isoleucine branched-chain aminoacids [Perkins *et al.*, 2009]. Various strategies of metabolic engineering in *B. subtilis* address the following objectives: (i) overcoming limitations of carbon flux in the metabolic routes, (ii) reducing maintenance metabolism, (iii) improving protein secretion and protein folding.

3.1.1 Carbon limitation in metabolic pathways

On the one hand, rechanneling the carbon flux towards a metabolite of interest generally consists in interfering with specific enzymatic reactions in other metabolic pathways. A gene knockout causes the annihilation of these enzymatic reactions, while the

selection of appropriate mutations only results in bottlenecks that slow down enzymatic turnover.

On the other hand, carbon limitation in the metabolic pathway of interest is overcome by enhancing gene expression. More precisely, either the amount or the stability of messenger RNAs available for ribosomal translation can be increased, or the transcriptional regulation can be reduced. Transcriptional attenuation typically controls genes expression in metabolic pathways. It consists in the binding of regulatory protein or metabolites to a specific segment of the nascent mRNA, resulting in the formation of hairpin secondary structures [Henkin and Yanofsky, 2002]. Premature termination of transcription through this mechanism is for instance involved in the complex regulation of tryptophan biosynthesis in *B. subtilis* [Gollnick *et al.*, 2005]. Specifically, the riboswitch mechanism is a widely distributed way for negative regulation of gene expression in bacteria, independently of any protein cofactor. This sensor-regulatory system relies on genetic elements in the 5'-untranslated regions of mRNAs which are able to specifically bind small metabolites. This results in allosteric rearrangement of the mRNA with subsequent effects on transcription or translation. In *Bacillus* species, riboswitches control the expression of more than 2% of all genes, responding to various metabolites including amino acids, nucleotides and protein cofactors [Montange and Batey, 2006]. Seven riboswitches have been identified in bacteria, involving the following metabolites: coenzyme B12 [Nahvi *et al.*, 2002], flavine mononucleotide [Mironov *et al.*, 2002; Vitreschak *et al.* 2002; Winkler [a] *et al.*, 2002], S-adenosylmethionine [Grundy and Henkin, 1998; Montange and Batey, 2006], adenine and guanine [Mandal *et al.*, 2003], lysine [Blount *et al.*, 2006; Sudarsan *et al.*, 2003], and

thiamin pyrophosphate [Mironov *et al.*, 2002; Rodionov *et al.* 2002; Sudarsan *et al.* 2003; Winkler[b] *et al.*, 2002]. Riboswitches can serve as antimicrobial drug targets [Blount *et al.*, 2006; Sudarsan *et al.*, 2005], but also support the understanding and engineering of high metabolites producing strains. Deregulated mutants of *B. subtilis* can be selected on the basis of their resistance to anti-metabolites which are analogues of the compound to be overproduced. For example, the *B. subtilis* strain currently used for industrial production of vitamin B2 (riboflavin) by DSM Nutritional Products Ltd. (Switzerland) is able to overcome a key transcriptional attenuation in the riboflavin biosynthetic pathway, promoted by a high intracellular concentration of flavine mononucleotide coenzyme. This deregulation is due to a mutation in *ribC* locus, selected after a screening of clones resistant to roseoflavin, a toxic riboflavin analogue [Perkins *et al.*, 1999].

3.1.2 Maintenance metabolism

An important criterion for industrial production of metabolites by high-cell density microbial processes is a low maintenance energy metabolism to reduce non-productive consumption of substrate. Maintenance metabolism is inherent to a given microorganism and represents the energetic cost dedicated to survival, without biomass or product formation. In carbon-rich cultures of *B. subtilis*, the energetic growth efficiency is reduced mainly due to a massive overflow metabolism, the primary dissipating flux of ATP equivalents [Dauner *et al.*, 2001]. Comparatively, carbon-limited fermentation processes, with a low growth rate and high biomass concentration, decrease the maintenance energy demand, which still remains quite high in *B.*

subtilis. Once the specific biosynthetic pathways have been optimized for a given metabolite, reduction of the maintenance metabolism is a valid objective for improvement of commercial process performance. For this purpose, energy respiration can be improved by redirecting electron flow to more efficient proton pumping branches within respiratory chains. Depending on the respiratory chain branch, one to three protons can be translocated by transported electron. This impacts the resulting proton gradient crossing the cytoplasmic membrane in return to generate ATP. A successful example was reported for the production of riboflavin, for which energy metabolism engineering rate was reduced by 40%. The results from this study suggest that increasing respiratory energy generation is a general strategy to improve product yields in industrial bioprocesses with *B. subtilis* [Zamboni *et al.*, 2003].

Another promising area for industrial bioprocesses is the development of bacterial cells with the lowest experimentally determined waste of energy and with industrially relevant phenotypes. Such a purpose is based on the concept of a minimal gene set, referring to the smallest possible group of genes that would be sufficient to sustain a functioning cellular life form under the most favourable conditions [Koonin, 2000]. For industrial biotechnology, the stake is the development of a minimal production host, deprived from the irrelevant or adverse traits and harboring a gene scaffold optimized for high metabolites production. In academic research, studying such a minimal gene set aims to highlight for instance the biological functions which are essential in most, if not all, bacterial species. A major advance in Synthetic Biology was recently achieved by researchers at the J. Craig Venter Institute (USA). Their results constitute the first proof-of-principle that microbial genomes can be designed *in silico*, chemically made

in the laboratory and transplanted into a recipient cell to produce a new self-replicating cell controlled only by the synthetic genome [Gibson *et al.*, 2010].

The non-essential genes in *B. subtilis* may have been acquired during evolution through the perpetually changing ecological niches. Efforts to minimize the *B. subtilis* genome were motivated in the past years by the will to generate a less complex organism, easy to model and understand. Determination of a minimal set of genes combines both computational and experimental approaches. Experimental conditions, such as nutrients in culture media, are major factors of variability. A minimal cellular-function set rather than a minimal gene set was therefore recently explored by computational approaches [Azuma and Ota, 2009]. Results of this study suggest that 41 pathway maps (of the KEGG database; Kyoto Encyclopedia of Genes and Genomes) are essential in *B. subtilis*, including 25 conserved pathway maps and 16 organism-specific pathway maps. The conserved maps include many pathways classified as "genetic information processing", whereas the organism-specific maps mainly include pathways for catabolism, reflecting evolutionary aspects. The genome of *B. subtilis* was experimentally reduced by removing prophages and AT-rich islands, consisting in 332 genes (SP-beta, PBSX prophages, prophage-like regions, and the largest operon *pks*). This resulted in a genome reduction of 7.7% [Westers *et al.*, 2003]. Further work on minimalization of the *B. subtilis* genome led to a collection of 157 interval strains, generated in maintaining 821 genes (essential genes and genes necessary for knockout procedure, competence *etc.*) In total, 77.8% of the genome (3,281 Mbp) was deleted in this collection of strains. These strains were phenotypically characterized and some of the mutants grew

interestingly much faster than the wild-type *B. subtilis* [Henry *et al*., 2010].

3.1.3 Protein secretion and protein folding

In *B. subtilis*, export of proteins requires dedicated transport machineries composed of channels and ATP-dependant or proton gradient-dependant proteins to drive active movements across the membrane (Figure I-1).

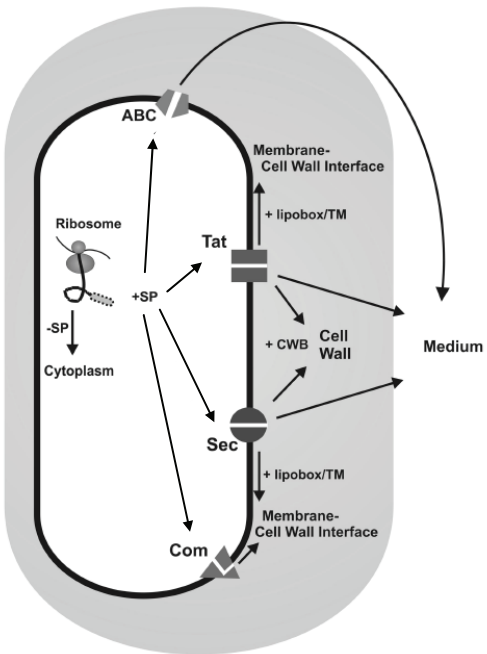


Figure I-1: Protein export pathways in *B. subtilis.* Adapted from Tjalsma *et al.*, 2004. Proteins can be sorted to various destinations depending on the presence or absence of an amino-terminal signal peptide (+/-SP). Depending on the class of their signal peptide, proteins can be secreted into the medium via the Sec or Tat pathway, or by ABC transporters. Proteins that have to be retained at the extracytoplasmic side of the membrane can contain either a transmembrane segment (TM) or a lipid modification (lipobox). Proteins that need to be retained in the cell wall can be exported via either the Sec or Tat pathway. Proteins containing cell wall-binding repeats (CWB) are exported via the Com system in the cell wall interface.

Proteins to be secreted have a signal peptide that directs their transport. The majority of extracytoplasmic proteins are exported via the well characterized Sec translocation pathway, ubiquitously conserved in prokaryotes. Despite the high secretion capacity of the Sec pathway, secretion of heterologous proteins remains difficult. Overexpression of the components of the secretion machinery and of chaperones is therefore a strategy to improve the production of heterologous proteins [Vitikainen *et al.*, 2005; Wu *et al.*, 1998]. An alternative secretion route, the Tat pathway, is also of interest and has been explored by genomic and proteomics approaches [Tjalsma *et al.*, 2004; Van Dijl *et al.*, 2002]. In *B. subtilis*, Tat translocases are made of the two subunits TatA and TatC. Whereas the Sec pathway only transports unfolded proteins, the Tat machinery translocates folded proteins. Another specific feature is a twin-arginine (RR) motif in the signal peptide, essential for recognition by the Tat translocation machinery. To improve the production of recombinant enzymes in industrial bioprocesses (*e.g.* proteases, lipases, alpha-amylases), it is crucial to investigate whether a commercially relevant Sec dependent protein can be re-routed into the Tat pathway. A successful example of engineering of the Tat pathway was recently reported, using the *B. subtilis* serine protease subtilisin as a case study [Kolkman *et al.*, 2008]. To reduce the degradation rate of secreted proteins, it is important to countereffect the extracellular proteases produced by *B. subtilis*. The use of protease-deficient host strains (*e.g.* WB800; Wu *et al.*, 2002) have thus shown to improve the stability of secreted proteins [Murashima *et al.*, 2002; Westers *et al.*, 2004; Wu *et al.*, 1991].

Correct formation of disulfide bounds, happening just after protein translocation, is another bottleneck in the production of heterologous proteins in *B. subtilis*. Disulfide bounds play a key role

in accurate protein folding, and therefore in protein activity and resistance to extracellular proteases. The demand in proteins with disulfide bonds for biopharmaceutical applications (*e.g.* human insulin and growth hormones) is strong, while the ability of *B. subtilis* to produce such molecules is very poor so far. Disulfide bonds consist of intra- or inter-molecular bridges formed through the enzyme-dependant oxidation of thiol groups of two cystein residues. Native *B. subtilis* thiol-disulfide oxidoreductases (BdbA to BdbD) catalyze the required oxidation. This system is however overloaded when recombinant proteins with disulfide bonds are overproduced. Proof-of-principle for the optimized production of disulfide bond proteins within *B. subtilis* was reported in a recent publication in which depletion of the major cytoplasmic disulfide bond reductase (TrxA), together with recombinant expression of an efficient staphylococcal thiol-oxidase (DsbA), results in a 3.5-fold increase in secretion in the growth medium of the disulfide-bond-containing alkaline-phosphatase from *E. coli* [Kouwen *et al.*, 2008]. Exploring both engineering of protein translocation and folding would likely broaden the capacity of *B. subtilis* to secrete a novel class of pharmaceutically relevant proteins with high yields.

3.2 Strategies and tools for directed genetic modifications in *B.subtilis*

3.2.1 Integration or deletion of DNA in the chromosome

Efficient gene insertion, deletion or modification involve techniques of genetic transfer, including bacteriophage-mediated transduction, protoplasts fusions, and transformation with various

concentrations of naked or protoplast-protected DNA [Harwood and Cutting, 1990]. Thanks to the natural competence of this organism, transformation in *B. subtilis* is an easy process, while electroporation and conjugation appear to be comparably inefficient.

B. subtilis 168 does not have any endogenous plasmid. Mechanisms of plasmid replication involving a single-strand circular intermediate DNA are thus not available. Consequently, the significance of plasmid-based expression systems in *B. subtilis* is questionable. Stable chromosomal integration is therefore the preferred method for industrial production purposes. Various methods to insert exogenous DNA fragments into the unique chromosome have been developed for laboratory studies of *B. subtilis* model organism [for a review: Perkins *et al.*, 2009].

Molecular genetic engineering deals with the exchange of a chromosomal gene by a mutated allele or an inactivated copy. Integration of this exogenous DNA into the chromosome of the recipient strain relies on homologous recombination (Figure I-2, panel A). Several strategies have been developed to promote gene replacements in bacterial chromosomes.

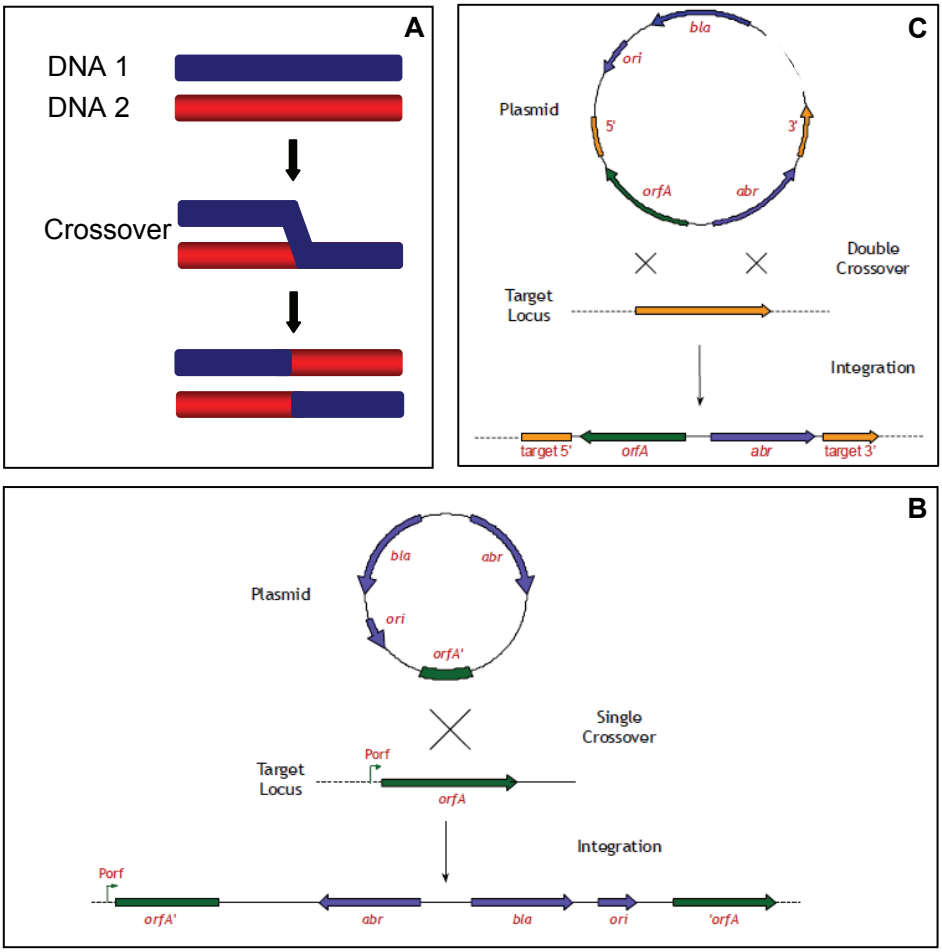


FIGURE I-2: Schematic view of homologous DNA recombination in *B. subtilis*. Panels B and C adapted from D.R. Zeigler. 2002. *Bacillus* Genetic Stock Center (The Ohio State University). **(A)** Recombination by crossover; an exchange of genetic material between homologous DNA molecules. **(B)** Knockout mutation by Campbell-like mechanism (single crossover) using an integration vector in a hypothetical open reading frame *orfA*. **(C)** Ectopic integration by double crossover to insert a hypothetical open reading frame, *orfA*, into a target locus on the chromosome of the recipient strain. *orfA* is flanked by two non-contiguous DNA regions (orange color), where the homologous recombination takes place. *Bla*; encodes β-lactamase for antibiotic selection in *E. coli*, *abr*; antibiotic resistance gene selectable in gram-positives, *or;*: origin of replication that functions in *E. coli* but not in gram-positives. The crossed lines indicate crossover events.

3.2.1.1 Methods using an antibiotic marker

A first set of methods leads to a stable chromosomal modification ensured by an antibiotic marker maintained in the host strain. The most basic strategy deals with the use of recombinant integrative plasmids from *E. coli* harboring an antibiotic cassette next to the DNA fragment to be integrated. These plasmids bear a conditional replication, meaning that replication functions in *E. coli* but not in *B. subtilis.* As the recombinant vector has a single homologous sequence, antibiotic-resistant transformants result from a single crossover by homologous recombination into the target locus, so called “Campbell-type mechanism” (Figure I-2, panel B).

This concept was improved with the development of the pMUTIN vectors collection for systematic gene inactivation. After integration, the gene to be disrupted is placed under the control of an IPTG (isopropyl β-d-1-thiogalactopyranoside)-inducible P*spac* promoter [Vagner *et al.*, 1998]. The main drawback of these two approaches is that antibiotic selection pressure needs to be constantly applied to maintain the vector into the chromosome through generations. This can be applied in laboratories studies but is inappropriate in large-scale fermentations.

Another collection of plasmids enables stable integration in the recipient chromosome by ectopic integration [Guerout-Fleury *et al.*, 1996]. The DNA fragment to be integrated in *B. subtilis* is inserted along a selectable antibiotic marker, between the 5’- and 3’-regions of a *B. subtilis* non-essential gene. These two sequences being homologous to sequences in the recipient chromosome, a double crossover takes place, leading to the integration of the exogenous DNA into the now-disrupted chromosomal target (Figure I-2, panel C). The main loci targeted by the existing plasmids for

ectopic integration include *amyE*, *bpr*, *dif*, *epr*, *gltA*, *lacA*, *pyrD*, *sacA*, *sacB*, *thrC*, and *vpr*, all of which are non essential genes.

A second set of genetic engineering methods has been recently developed. These new strategies are based on long flanking homology (LFH) PCR constructs in *B. subtilis*. LFH-PCR was initially developed in yeast for gene disruption [Walch, 1996]. Instead of being cloned in an *E. coli* vector, a disruption cassette is generated *in vitro* by PCR, in such a way that both sides of the antibiotic marker it contains has LFH regions of several hundred base pairs. The LFH regions are homologous to the targeted DNA sequence in the chromosome. Integration is ensured by a double-crossover event.

The presence of antibiotic resistance genes into genetically engineered strains for the production of commercial products is currently possible with US and EU regulations as long as the final product is strictly deprived of genes encoding antibiotic resistances, harbored by the industrial microorganisms. Nevertheless, registration of the genetically engineered strain is facilitated and perception of consumers is improved in the case of marker-free microorganisms.

3.2.1.2 Marker-free methods

Marker-free engineering consists in a two-steps method: first the insertion of a counter-selectable marker that subsequently loops out of the chromosome target to be knocked out. This method is however quite labor intensive and has a limited efficiency. One example is the Cre/*lox* site-specific recombination system, in which

two flanking regions, homologous to the target to be deleted, and an antibiotic marker are combined by PCR. This cassette is flanked by two *loxP* mutants (*lox*71 and *lox*66). Cre recombinase catalyzes then a reciprocal site-specific recombination between the two *loxP* sites, promoting a loop-out of the DNA located in between. Cre recombinase is entered in the host via a plasmid (pTSC) to be constitutively expressed, or is directly inserted in the chromosome as an IPTG-inducible *cre* expression construct. The whole process can be accomplished in about 2-4 days [Yan *et al.* 2008]. A very similar approach uses two *dif* sites, recognized by the *B. subtilis* native RipX/CodV site-specific recombinases [Bloor and Cranenburgh, 2006]. Another strategy is based on PCR-construction of a cassette with a counter-selection marker and an antibiotic resistance gene [Brans *et al.*, 2004; Fabret *et al.*, 2002; Yu *et al.*, 2009]. This cassette is flanked by two homology regions targeting the chromosomal locus to be deleted. Two short directly repeated (DR) sequences are included in the design of the flanking homology regions. In a first step, the cassette is integrated in the chromosomal target by homologous recombination (antibiotic selection). The counter-selection pressure is then used to promote the loop-out of the cassette by a single-crossover event between the two short DR sequences. The target DNA in between is therefore also excised from the chromosome. Several systems using this strategy have been recently reported with various counter-selection markers: (i) the *E. coli* toxin gene *mazF*, placed under the control of a xylose-inducible promoter to promote its excision [Yu *et al.* 2009], (ii) the product of the *blaI* gene, which represses the *lysA* gene under the control of the *Bacillus licheniformis* P_{blaP} promoter and therefore confers lysine auxotrophy (lysine prototroph mutants, resulting from a loop-out of the cassette,

are selected on minimal medium) [Brans *et al.*, 2004], (iii) the *upp* gene, encoding uracil phosphoribosyl-transferase (eviction of *upp*-cassette from the chromosome is promoted with 5-fluorouracil, a toxic analogue of pyrimidine) [Fabret *et al.*, 2002].

Another method to generate markerless in-frame deletion relies on the pEpUC∆1 *E. coli – B. subtilis* shuttle vector. This vector for conditional integration and excision is based on plasmid pE194 [Weisblum *et al.*, 1979]. The plasmid contains a selectable erythromycin-resistance cassette, a temperature-sensitive origin of replication inhibited over 51°C, and short sequences located upstream and downstream of the target gene to be deleted. Transformation and integration of the plasmid is performed at 51°C, selecting for erythromycin resistance. Excision is then forced at 28°C for 72 hours, in the absence of antibiotic selection. The target gene is excised along with the plasmid, which loops out of the chromosome to be replicated.

When the gene to be modified is involved in a biosynthetic or metabolic pathway, a more straightforward approach can be used. A LFH-PCR product containing an antibiotic marker is first constructed and integrated in the target chromosomal locus. This leads to an auxotrophic phenotype. Same LFH regions are then used to generate another LFH-PCR product, in which the antibiotic marker sequence has been replaced by the modified target locus. Selection for recovery of a prototrophic phenotype allows then to get the desired mutation at the native locus (partial or complete deletion of the ORF or its promoter region, insertion of point mutations).

3.2.2 Enhancement of gene expression

The above-mentioned integrative plasmids can be used to add extra-copies of target genes into the bacterial chromosome. The same recombinant plasmid can be integrated at different loci of the engineered recipient strain. Alternatively, the copy number of the integrated plasmids (*i.e.* of the genes of interest it harbours) can be expanded by increasing concentrations of the selective antibiotics. *B. subtilis* clones containing amplified copies already exist in the bacterial population, and are simply selected with this method. Correct implementation of this strategy requires finding the optimal copy number. For large-scale fermentation, antibiotic selection pressure must be maintained at least during the inoculum fermentation runs in order to stabilize the integrated plasmid(s). This might be associated with important cost and safety issues in industrial processes. Another disadvantage is that the number of copies of these plasmids might vary, and unneeded foreign DNA may be introduced in addition to that of the desired gene.

In a more straightforward strategy, the native promoter region of the target genes is replaced by a strong promoter. Efficient strong promoters result in constitutive expression and harbor binding sites recognized by sigma A RNA polymerase. Sigma A is the major *B. subtilis* sigma factor, so transcription has a high chance to be initiated in vegetative growth conditions, being however maximal during exponential growth. Examples of strong promoters used in *B. subtilis* are: (i) the promoter of *B. subtilis veg* gene (P*veg*), which can be used in a form containing tandem binding sites (for high gene expression) or in a form lacking the upstream site (for medium gene expression) [Le Grice *et al.*, 1982; Moran *et al.*, 1982], (ii) the promoter of *B. subtilis* P43 protein,

active during exponential and lag phases of growth thanks to its recognition by both sigma factors A and B [Wang and Doi, 1984], (iii) other strong constitutive promoters, denominated SP01-15 and its derivative SP01-26 [Lee and Pero, 1981], originating from the bacteriophage SP01.

Following gene overexpression with a strong promoter, messenger RNAs accumulate in the cytoplasm and might be "polluting" the cell compartment. Instead of increasing the amount of mRNA molecules, another strategy consists in extending the half-life of the existing one. The translation rate is therefore improved. A general model in *E. coli* suggests that RNaseE, a 5'-end-dependent endoribonuclease, is responsible for the initiation of the decay of most mRNAs. RNase E binds to the accessible 5'-end and promotes successive cleavage events resulting in mRNA fragments with accessible 3'-ends, which are rapidly degraded by 3'-5'-exoribonucleases [Régnier and Arraiano, 2000]. Although this enzyme is absent in *B. subtilis*, ribonucleases J1 and J2 were recently found to have a dual 5'-3'-exoribonucleolytic/endoribonucleolytic activity, promoting the decay of many different mRNAs and small non-coding RNAs [Arraiano *et al.*, 2010; Evens *et al.*, 2005; Mathy *et al.*, 2007]. The 5'-end is therefore a key determinant of mRNA stability. Stem–loop structures, present at the 5'-end of *aprE, ermA, ermC, glpD, ompA*, and *papA* mRNAs for example, interfere with the accessibility of the 5'-untranslated region of mRNA, conferring stability to the messenger [Sharp and Bechhofer, 2005].

4. B. subtilis endospore and its applications

4.1 Endospore formation

Just like about 200 species among 25 genera of aerobic bacteria, *B. subtilis* vegetative cells have the ability to produce ovoid-shape dormant cells when confronted to nutrient deprivation and high population density. Sporulation in *B. subtilis* constitutes the best studied example of a prokaryotic differentiation process, and has been extensively studied during the past years.

Sporulation leads to a metabolically inactive and extremely resistant structure. Even after a very long period of dormancy, spores can sense when environmental conditions become favourable for growth, possibly thanks to enzymes present on their surface [Dricks, 2003], and convert to a regular vegetative cell cycle through germination. During the three last decades, large efforts focused on understanding the events leading to the induction of sporulation or the sequential events that lead to spore formation [Errington, 2010; Piggot and Hilbert, 2004].

In all endospore forming organisms, sporulation follows a similar morphological sequence and spores have the same concentric architectural plan [Henriques and Moran, 2007]. An asymmetric cell division occurs, resulting in a large mother cell containing a smaller cell (prespore or forespore) that will develop into a mature endospore. The plasma membrane of the mother cell grows around the forespore, resulting in an engulfed core surrounded by two membranes and isolated from the external medium. The core contains chromosomal DNA, ribosomes and other enzymes, but is metabolically inactive. The inner and outer

membranes locate the assembly of peptidoglycans. The inner thin layer is called primordial germ cell wall, and the outer thick layer is defined as cortex. Formation of the spore cortex is accompanied by a dramatic decrease in water content, resulting in the dehydrated state of the core which contains the genetic material. This ensures dormancy and thermostability of the spore. The cortex also provides protection of the spore against extreme desiccation. Closely apposed to the cortex, the multilayer spore coat is then concentrically formed. The undercoat, an elastic layer of amorphous material and undefined composition, ensures the proper adhesion of cortex and coat. Spore coat is mainly composed of about 80 proteins. Although most of these proteins are a part of the coat structure, at least 20 of them have an enzymatic function or show similarity to known enzymes. They can for instance play a role in the assembly process (by posttranslational modifications), modulate germination, or participate in spore protection. For example, the product of the *oxdD* gene, located in the inner coat layer, shows an oxalate decarboxylase activity which is suspected to confer protection to spores against a harmful compound produced by fungi in the soil [Costa *et al.*, 2004]. The successive overlay of these proteins around the cortex takes place in the cytoplasm of the mother cell, covering a period of about 6 hours and resulting in the formation of a lamellar inner coat and an electron-dense outer coat. A collection of about 40 polypeptides from 6 to 70 kDa can be extracted from purified spores by treatment with alkali compounds or reducing agents in presence of detergents. This extraction solubilises proteins of the inner coat and most of the outer coat, but does not damage the cortex. The coat contributes to the resistance of the spore to extreme physical (*e.g.* high pressures, ultraviolet and gamma radiations) and chemical stresses (*e.g.* lysozyme, oxidizing

agents such as ozone, chlorine dioxide, hypochlorite, peroxide nitrite, hydrogen peroxide, and formaldehyde). Although the periphery of *B. subtilis* spores is not surrounded by a real exosporium (like in spores of *Bacillus thuringiensis*, *Bacillus anthracis*, *Bacillus mycoides*, *Bacillus cereus* or some *Clostridia*), a thin glycoprotein layer was recently brought to light and could represent a rudimentary exosporium of unknown function. After completion of this building process, the mother cell undergoes autolysis, resulting in the release of the mature endospore into the surrounding environment (Figure I-3).

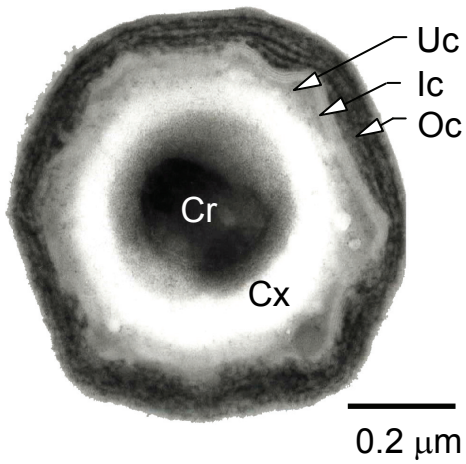


FIGURE I-3: Ultrastructure of *B. subtilis* spores. Adapted from Henriques and Moran, 2007. Sectioning electron micrograph of cross-sections. The spore compartments or structures are recognized in the spore: core (Cr), cortex peptidoglycan layer (Cx), undercoat region (Uc), inner (Ic) and outer (Oc) coat layers.

The major morphological changes taking place in the course of endospore formation (Figure I-4) are the consequence of a complex metabolic cascade.

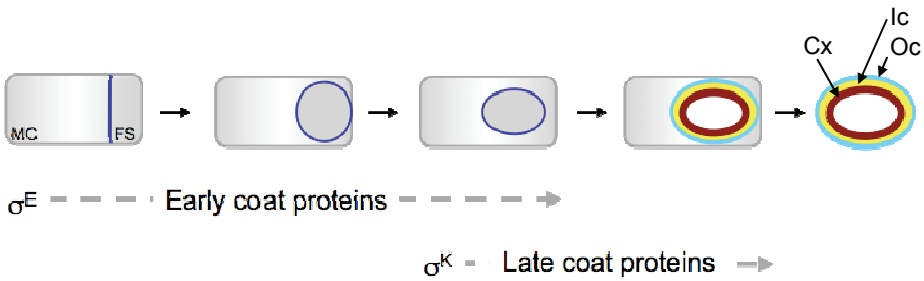


FIGURE I-4: The main stages of coat assembly during *B. subtilis* sporulation. Courtesy of C. Serra and A.O. Henriques (ITQB-NL). Coat assembly covers a period of about 6 hours, starting with the asymmetric division which results in the mother cell (MC) and forespore (FS) compartments. Sporulation then proceeds with engulfment of the forespore by the mother cell. Gene expression in the mother cell is governed in large part by the action of two primary transcription factors, RNA polymerase factors sigma E (σ^E) and sigma K (σ^K). The following spore layers are represented: cortex peptidoglycan layer (Cx), inner (Ic) and outer (Oc) coat layers.

Both external and internal stimuli, including the nutrient status of the cell, cell cycle and cell density signals, help the cell to decide if it should go on with vegetative growth, develop competence, or in the worse case, launch the irreversible process of sporulation. Initiation of sporulation (Figure I-5) involves the activation and expression of hundreds of genes and the expenditure of a large amount of energy. A multicomponent phosphorelay, leading to the phosphorylation of master regulator Spo0A, controls entry in the sporulation process. After sensing a still-unknown starvation factor, five sensor histidine kinases (KinA, KinB, KinC, KinD, KinE) autophosphorylate. They transfer their phosphoryl group to Spo0A through two proteins, Spo0F and Spo0B. The phosphorylated Spo0A activates or represses the transcription of

121 genes in the mother cell. Notably, Spo0A~P represses *abrB* encoding a transcriptional repressor that regulates expression of several sporulation genes and genes for competence, by binding DNA. Expression of the sporulation genes is mainly governed by the action of the two RNA polymerase factors sigma E (SigE) and sigma K (SigK), acting as primary transcription factors, and three accessory regulators SpoIIID, GerR and GerE, acting as DNA binding factors. Expression of the genes encoding structural coat proteins (so called *cot* genes) is governed by a cascade of these expression factors in the mother cell in the chronological sequence SigE>SpoIIID>SigK>GerE.

The complex, and not fully understood network of interactions between spore coat proteins, is crucial for coat assembly and its protective functions. One of the key features of the network is the influence of a small subset of proteins that direct the assembly of most of the coat [Kim *et al.*, 2006].

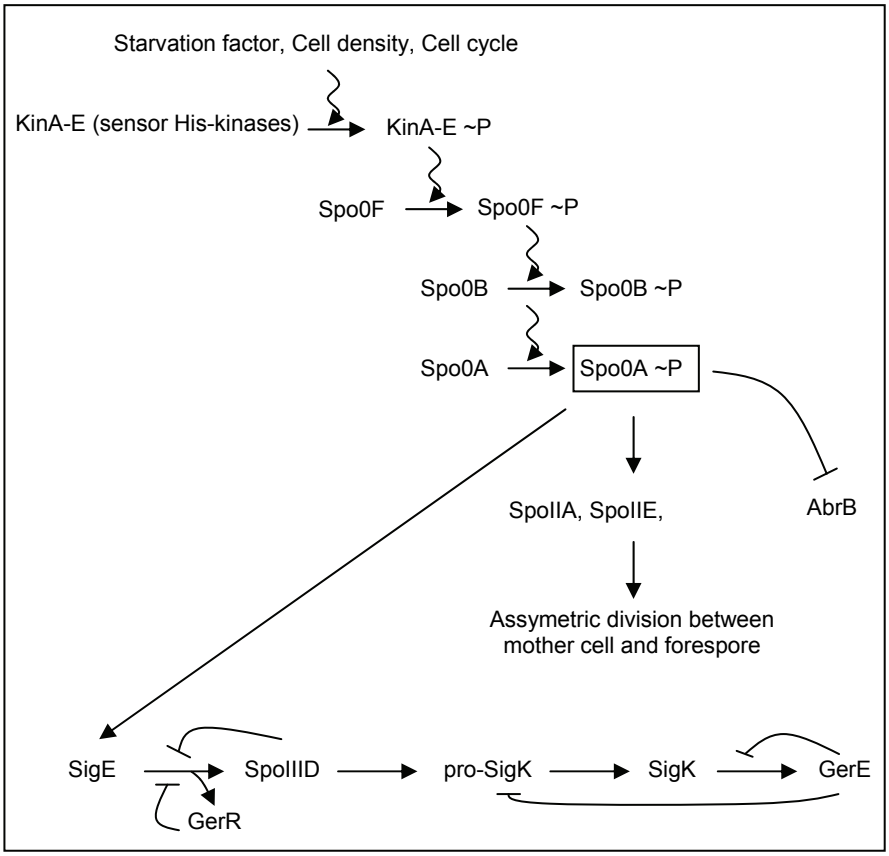


FIGURE I-5: Initiation of the sporulation process in *B. subtilis.* Adapted from Burbulys *et al.*, 1991; Henriques and Moran, 2007; Piggot and Hilbert, 2004. Initiation of sporulation involves master regulator Spo0A, activated by a phosphorelay. Expression of specific sporulation genes is driven by the RNA polymerase factors SigmaE (SigE) and SigmaK (SigK). Sigma K is initially synthesized as a preprotein, whose proteolytic activation is conditioned by a signal from the forespore. SpoIIID, GerR and GerE are the main accessory regulators of the gene expression in the course of endospore formation.

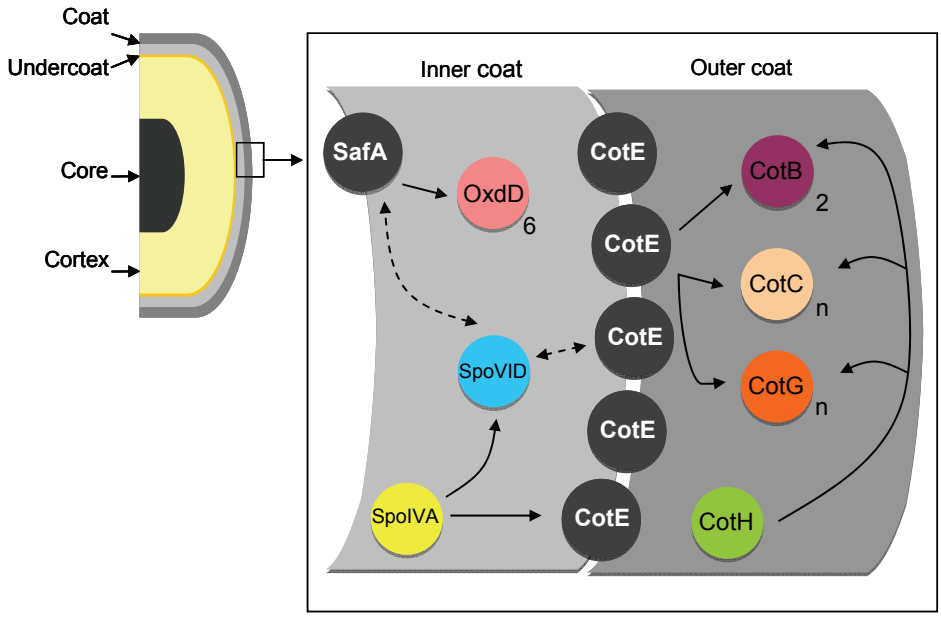


FIGURE I-6. Simplified structure and interactions of *B. subtilis* spore coat. Adapted from Henriques and Moran, 2007. The multilayer coat of *B. subtilis* spores is assembled in a complex order from proteins accumulating in the cytoplasm of the mother cell (see text for more details). SafA, SpoIVA, SpoVID, CotE and CotH are morphogenetic proteins (*i.e.* proteins essential for the assembly but not for the production of other coat proteins). CotB, CotC and CotG are the most abundant structural coat proteins. OxdD is a coat-associated enzyme. SafA brigdes the cortex and the coat, and the CotE-shell locates at the inner/outer-coat interface. Expression of the represented proteins is sequentially driven by RNA polymerase sigmaE (for SpoIVA, SpoVID, SafA, CotE, CotH) and sigmaK (for OxdD, CotB, CotC, CotG) subunits. The subscript n or number indicates multimerization of the protein. Solid arrows represent dependencies for the gene expression and/or protein assembly. Broken arrows indicate protein interactions.

Assembly of the coat layers (Figure I-6) involves interactions between the two essential proteins made under sigma E control: SpoIVA, produced at the early stage of sporulation, and CotE. SpoIVA localizes at the outer forespore membrane and directs the assembly of CotE which forms a shell-like structure surrounding the engulfed forespore. The region delimited by SpoIVA and CotE is

named the precoat. The precoat is then converted into an inner coat layer concomitantly with outer coat assembly outside the CotE ring, both processes depending of sigma K factor. CotE is essential for assembly of the outer coat layer and spore resistance to lysozyme. Several sigma K-controlled proteins are thus targeted to the inner coat layer, presumably including the OxdD oxalate decarboxylase. Three additional proteins play an important role in coat formation: SpoVID, SafA and CotH. The CotE ring is maintained around the forespore by SpoVID that also direct expression of SafA. Both SpoVID and SafA are produced under sigma E control. SpoVID-SafA interaction is essential in lysozyme resistance of the spores. It should be noted that a poorly understood interaction is also postulated between SafA and the OxdD inner coat-associated enzyme. Finally, CotH is involved in the assembly of various outer coat components such as CotB, CotC, and CotG [Barak *et al.*, 2005; Henriques *et al.* 2004; Henriques and Moran, 2007; Wang *et al.*, 2009].

4.2 Spore-display

Display at the surface of *B. subtilis* endospores consists in the fusion of genes encoding respectively an anchoring motif naturally occurring in the spore coat (carrier) and a molecule of interest to be exposed (passenger). Transcription of the fusion gene is driven by the promoter of the anchoring motif, to ensure that both carrier and passenger are expressed in a timely manner, and during endospore formation. Consequently, the fusion accumulates in the cytoplasm of the mother cell before it assembles around the spore cortex together with the other coat proteins (Figure I-7).

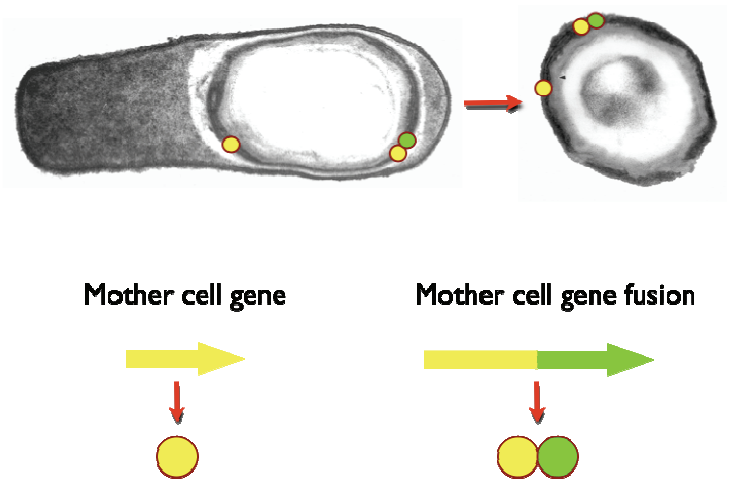


FIGURE I-7: Spore display concept in *B. subtilis*. Courtesy of C. Serra and A.O. Henriques (ITQB-NL). During the sporulation process, the coat protein used as carrier and the fusion protein (carrier-passenger of interest) are accumulated and assembled in the spore coat, in the mother cell. The mature spore is then released with the lysis of the mother cell.

Besides the well-known advantages of *B. subtilis* spores (easy preparation at large scale, indefinite storage time and resistance to harsh environments often met in industrial processes), specific benefits are offered by spore display systems. The display of heterologous proteins on the cell-surface of microorganisms for various biotechnological applications has been reported for viruses [Arbabi-Ghahroudi *et al.*, 2009; Azzazy *et al.*, 2002], yeast [Kondo and Ueda, 2004;Tanaka *et al.*, 2010; Tannino *et al.*, 2009], and bacteria [Daugherty, 2007; Wernérus and Stahl, 2004]. The most recent examples of bacterial display deal with surface proteins of *E. coli* [Kenrick *et al.*, 2010, Yim *et al.*, 2010, Yim *et al.*, 2006] or an exosporium protein of *B. thuringiensis* [Park *et al.*, 2009]. When filamentous phages or bacterial cells are used for surface display, the protein of interest has first to cross the cytoplasmic membrane

before being exposed outside. Some of these proteins, which either fold fast in the cytoplasm or have a high hydrophobic composition, are often trapped in the cytoplasmic membrane and cannot be translocated. In addition, microbial display systems are limited in the size of the proteins (smaller than 80 kDa) that can be displayed [Samuelson *et al.*, 2002; Yim *et al.*, 2006]. These problems can be prevented by using endospores formed by *B. subtilis,* as synthesis of the carrier-passenger fusion occurs in the cytoplasm of the mother cell. Moreover, the full set of ATP-dependant molecular chaperones is present in the cytoplasm to assist in the correct folding of the passenger protein and the multimerization of the passenger protein is possible [Kim and Schumann, 2009; Kwon *et al.*, 2007].

A broad range of potential applications of spore display arose, including the screening of novel binding partners, immobilization of active enzymes for many industrial processes (bioconversion), environmental bioremediation, biosensors, and the delivery of drugs and vaccines. *B. subtilis* spores as vehicles for heterologous antigenic molecules have surely been the most studied application for the past decade. Successful antigen-specific immune responses were observed following oral or nasal administration of recombinant *B. subtilis* spores in murine models [for a review: Barak *et al.*, 2005; Cutting *et al.*, 2009; Li *et al.*, 2009; Ricca and Cutting, 2003].

A few structural proteins (CotG, CotB and CotC) of the outermost layer of the *B. subtilis* spore coat have been selected as anchor for their location and their relative abundance. These proteins are structural tyrosine-rich proteins present as multimers within the spore coat. They were successfully used to display proteins and enzymes. CotB was initially used to display the

carboxyl terminal fragment of the tetanus toxin (TTFC) [Isticato *et al.* 2001], and then the protective antigen (PA) of *B. anthracis* [Duc *et al.*, 2007]. CotC has been used to display the B subunit of the heat-labile toxin (LTB) of *E. coli* [Mauriello *et al.*, 2004], and the passenger antigens above-mentioned; the TTFC fragment [Mauriello *et al.*, 2004] and the anthrax PA [Duc *et al.*, 2007]. Two tegumental proteins of *Clonorchis sinensis* [Zhou *et al.*, 2008] have as well been spore displayed with CotC as anchoring motif. CotG was used as anchoring partner to the bioactive β-galactosidase of *E. coli* [Kwon *et al.*, 2007]. This enzyme is a very large molecule (116 kDa for each monomer) which is only active as a tetramer. This successful example demonstrates that the multimeric nature and the size of the passenger enzyme are not major concerns for spore-based display. It was also reported that streptavidin, a 60 kDa-protein of *Streptomyces avidinii* was displayed at the surface of *B. subtilis* spores fused to CotG. Active streptavidin is formed of four monomers (15 kDa each) and has strong affinity for biotin. Various biotin-coupled biomolecules (*e.g.* proteins, nucleic acids, carbohydrates can therefore be detected by spores displaying streptavidin [Kim *et al.*, 2005]. A variant of the green fluorescent protein (GPuv) was also surface-exposed on *B. subtilis* spores with CotG, offering interesting perspectives for the study of interactions with this coat protein [Kim *et al.*, 2007]. Finally, recent work towards the development of a spore vaccine application describes the use of the three coat proteins CotB, CotC and CotG to display an antigen, the urease subunit (UreA) of *Helicobacter acinonychis*, a major animal pathogen, also recognized as a useful model to study the virulence of *Helicobacter pylori* [Hinc *et al.*, 2010].

5. Probiotics

5.1 Microbial Probiotics

The use of food derived from microbial activity goes back to the foundation of human civilizations with fermented milk, which was probably the first food to contain active microorganisms. The word “probiotic” means “for life”, and indicates microorganisms associated with beneficial effects for humans and animals. The current definition adopted by FAO (Food and Agriculture Organization of the United Nations) and the WHO (World Health Organization) for probiotics is “live microorganisms which, when administered in adequate amounts, confer a health benefit on the host” [FAO/WHO, 2001]. Two main mechanisms for probiotic action have been suggested [Anadon *et al.*, 2006]: (i) nutritional effect: stimulation of indigenous enzymes, and production of vitamins or antimicrobial substances, (ii) health or sanitary effect: increase in colonization resistance, competition for gut surface adhesion, and stimulation of immune response. Bacterial and yeast species used in animal feed can be normal resident of the gastrointestinal tract, or not. A probiotic dose of 10^6 to 10^7 CFU/g of administrated feed is considered as optimal to obtain a balance between probiotic microorganisms and resident bacteria. [Guillot, 2009].

5.2 European regulation

In the EU, the European Food Safety Authority (EFSA) is the keystone for independent scientific risk assessment of food and feed. The European regulation (EC) No 1831/2003 establishes the

rules governing the authorisation of additives in animal nutrition [EFSA[b], 2003]. The assessment of microbial feed additives (enzymes and microorganisms) in feedingstuffs is based on their: (i) identity, characterization, conditions of use and methods of control, (ii) efficacy, (iii) safety under the conditions of use for target species and consumer, including toxicity evaluation, antibiotic resistance profile and transferability, (iv) worker safety assessment, (v) environmental risk assessment.

Commercialisation of genetically modified (GM) probiotics for food/feed applications is not *a priori* excluded in the EU. An authorisation may be granted for food/feed products which consist of, contain or are produced from GM organisms. These products must not promote adverse effects on human/animal health or on the environment, not mislead the consumer (*e.g.* not be nutritionally disadvantageous). The strict approval by EFSA relies on a rigorous safety assessment described in the regulation (EC) No 1829/2003 [EFSA[c], 2003]

5.3 *B. subtilis* as Probiotic

For over a thousand years, *B. subtilis* fermented soybean has been used in human food (traditional Japanese Natto), underlining harmlessness of this bacterium as probiotic. *In vitro* and *in vivo* studies have been performed to support the safety of using *B. subtilis* PY79 (1A747, Bacillus Genetic Stock Center, USA; Youngmann *et al.*, 1984], a prototrophic strain derived from the Marburg 168 type strain, as a potential microbial supplement. An important aspect of safety for the potential microbial supplements is to ensure that no enterotoxin is produced by the bacterium. Neither

genes known to encode enterotoxins in *B. cereus* nor haemolytic activity were detected. In addition, *B. subtilis* PY79 laboratory strain has no antibiotic resistance. No indication of pathogenicity, infection or toxicity can therefore be related to the consumption of *B. subtilis* strains [Hong *et al.*, 2008].

An *in vitro* study held with human intestinal-like cells (Caco-2) showed that *B. subtilis* can also provide health benefits through the production of competence and sporulation factor (CSF), a bioactive pentapeptide involved in quorum-sensing, known to confer to bacteria the ability to communicate and change behaviour of the same or other species in response to conditions and perturbations of the environment (nutrient availability, competition with other potentially pathogenic microorganisms). Once taken up by intestinal epithelial cells, CSF notably induces cytoprotective heat shock proteins (Hsps), which prevent oxidant-induced intestinal epithelial cell injury and loss of barrier function [Fujiya, 2007].

Since the prohibition by the European Union (EU) of antibiotics use in feed as promoter for growth in 2006, a market for innovative growth-promoting or prophylactic products in modern farming has arisen. Table 2 presents a list of the probiotics products based on spores of *B. subtilis* strains. Some of them were authorized by the EFSA and are qualified for QPS status.

Table 2: Commercial probiotics containing spores of *B. subtilis*. Adapted from Hong *et al.* 2005.

Products	Target	Manufacturer	Comments/References
Baozyme-Aqua	Schrimps	Sino-Aqua Corp., Taiwan	*B. subtilis* strains Wu-S and Wu-T
Badisubtilis	Humans	Bidiphar Pharma, Vietnam	
BioGrow®	Poultry, Calves, Swine	Provita Eurotech Ltd., Ireland	Mixture of *B. licheniformis* and *B. subtilis* strains
BioPlus 2B® [(a)]	Piglets, Turkeys [(b)]	Chr. Hansen AS, Denmark	Mixture of *B. licheniformis* (DSM 5749) and *B. subtilis* (DSM 5750) in a 1:1 ratio. Dose of 1.6 x 10^9 CFU/kg feedingstuff [EFSA[d], 2006]
Biosporin®	Humans	Biofarm, Ukraine	Mixture of *B. subtilis* 2335 and *B. licheniformis* 2336 strains (ratio 3:1)
Biostart®	Aquaculture	Microbial Solutions, South Africa	Mixture of two strains of *B. subtilis* and strains of *B. licheniformis*, *B. megaterium*, and *Paenibacillus polymyxa*
B. subtilis 035® [(a)]	Chickens	Chr. Hansen AS, Denmark	Same *B. subtilis* strain than BioPlus 2B® (DSM 5750). Dose range of 1.6 x 10^9 CFU/kg feedingstuff (EFSA, 2006[e])
Biosubtyl DL	Humans	IVAC, Vietnam	Mixture of *B. subtilis* and *Lactobacillus acidophilus*
Calsporin® [(a)]	Chickens	Calpis Co. Ltd., Japan, represented in the EU by Orffa Inter. Holding	Dose of 1 x 10^9 CFU/kg feedingstuff [EFSA[f], 2007]
CloSTAT® [(a)]	Chickens	Kemin Industries, Inc., USA	*B. subtilis* PB6. Dose range of 1 x 10^7 - 5 x 10^7 CFU/kg feedingstuff [EFSA, 2009]
Lactipan Plus [(a)]	Humans	Istituto Biochimico Italiano, Italy	
Medilac	Humans	Hanmi Pharma Ltd., China	Mixture of *B. subtilis* RO179 and *Enterococcus faecium*
Nature's First Food	Humans	Nature's First Law, USA	Mixture of 42 species, including *B. subtilis*
Neolactoflorene [(a)]	Humans	New Pharma, Italy	Mixture of 3 lactic acid bacteria strains and *B. subtilis*
Pastylbio	Humans	Pasteur Institute, Vietnam	
Primal Defense™	Humans	Garden of Life, USA	Mixture of 14 bacterial species, including *B. subtilis*
Promarine®	Schrimps	Sino-Aqua Corp., Taiwan	Mixture of 4 strains of *B. subtilis*

[(a)] Authorized in EU by EFSA
[(b)] Undergoing assessment by EFSA for permanent authorisation in EU

As microoganisms have to be in a viable state to be efficient probiotics, spore-based products are interesting because of their robustness and their long shelf life. *B. subtilis* laboratory strain (PY79) was shown to be able to grow under aerobic and anaerobic conditions. Spores of PY79 are completely resistant to SGF (simulated gastric fluid) and SIF (simulated intestinal fluid). This may ensure that an oral dose of spores can reach and have a potential effect in the gastrointestinal tract [Hong *et al.*, 2008].

Recent *in vivo* studies in murine and poultry models show that ingested *B. subtilis* spores are able to proliferate in the small intestine. Data generated from these studies show evidence that spores of *B. subtilis* administrated as oral dose can survive the harsh acid environment of the stomach and are then able to germinate in the intestines, where the pH of gastric juices is sufficiently reduced. After a safe transit across the stomach, spore germination followed by vegetative growth takes place in the upper part of the small intestine, subsequently ending by sporulation in the lower part of the intestine. Although spore probiotics are metabolically dormant upon administration, they are able to display metabolic activity after germinating in the gastrointestinal tract. A first *in vivo* study performed with a mouse model used a molecular approach based on the detection of mRNA of a chimeric gene fusion in different sections of the intestine by RT-PCR after oral administration of a single dose of *B. subtilis* spores. The chimeric gene was obtained from the fusion of the *ftsH* gene of *B. subtilis* and *lacZ* of *E. coli*. The *ftsH* gene being only expressed during vegetative cell growth, an *ftsH-lacZ* fusion can only be expressed after germination of the administrated spores harbouring the fusion. Negative control experiments were conducted on wild-type strains (without the fusion) and on spores of *B. subtilis* harbouring the

fusion, proving that the RNA detection is specific of the vegetative state. Germination of administrated spores could be observed in the jejunum and ileum, but not in the duodenum of mice [Casula and Cutting, 2002]. Using either *rrnO* or *cotB* genes of *B. subtilis*, which expression is respectively specific of the vegetative state or the sporulation, fused to *tetC* as a reporter gene (encoding a fragment of the tetanus toxin of *Clostridium tetani*), Tam *et al.*, 2006 gave the first indications (immunological and molecular evidence) that *B. subtilis* spores not only germinate in the mouse gut but are also able to grow and resporulate The *B. subtilis rrnO* gene fused to *lacZ* of *E. coli* also demonstrated on a poultry model that spore germination is induced either prior to or upon entry into the small intestine [Cartman *et al.*, 2008].

Persistence experiments with the *B. subtilis* PY79 laboratory strain resulted in the detection of spores released from the mouse gut not more than 12 days after oral administration, suggesting that *B. subtilis* is a transient member of the gut microflora. This can likely be explained by the fact that laboratory *B. subtilis* strains can not form biofilms. Adherence and invasion in various cell culture lines were tested *in vitro* and showed a limited ability to bind for the domesticated *B. subtilis* strains [Hong *et al.*, 2008]. Although *B. subtilis* does not appear to permanently colonize the host microflora, natural gut isolates of *B. subtilis* recovered from human faeces persist in the murine gut for almost twice as long as the laboratory strain PY79, with shedding still detectable up to 27 days after administration. These isolates were shown to form biofilms and were able to adhere to Caco-2 cells [Tam *et al.*, 2006].

6. Scope of the thesis

Besides the interest in its study as a model organism for fundamental research, *B. subtilis* has been a major workhorse for industrial microbial processes. The work presented in this thesis describes various aspects of the development of new commercial applications of *B. subtilis* based on genetic engineering. On the one hand, it aims to show how metabolic engineering, supported by random mutagenesis and environmental chemistry, helps to better understand metabolic pathways and to generate a vitamin overproducer organism. On the other hand, this thesis describes an example of structural engineering applied to the design of innovative applications for *B. subtilis* spores.

Chapter I provides a bibliographic introduction summarizing the origin of *B. subtilis* laboratory strains and associated key features. Due to its intrinsic properties, this microorganism is widely used in the industrial production of enzymes and of metabolites of high commercial value, as well as in the preparation of probiotic mixtures for animals and humans. Use of *B. subtilis* for efficient biotechnological applications often requires genetic modifications in order to optimize the metabolic pathways or some structural elements.

Chapter II illustrates how environmental chemistry and metabolic engineering can impact development of a *B. subtilis* overproducing strain for vitamin B1 (thiamin). More precisely, this chapter reports on the elucidation of an unknown thiamin salvage pathway from thiamin-degraded residues, therefore supplying new building blocks for the biosynthetic pathway independently of the *de novo* synthesis of thiamin. The new salvage pathway highlights in particular the microbial adaptation to the environment and gives

new insights to reconsider the physiological significance of thiaminase II.

Chapter III describes mutagenesis studies performed with *B. subtilis* which result in better understanding the thiamin salvage pathways, and generating a thiamin overproducing strain. Thiamin-deregulated mutations were generated. In our study, they map to *yloS*, *yuaJ*, and *ykoD* loci, and are in agreement with the known riboswitch mechanism of thiamin regulation. When combined in the same *B. subtilis* host, these new mutations result in the significant increase of thiamin production in fermentation, on the way to industrially relevant overproducing strains.

The work reported in **Chapter IV** is an example of genetic engineering of *B. subtilis* spore coat applied to a new concept of spore display. While spore display was mainly oriented in the last decade toward vaccinology applications, using structural proteins of the spore coat as anchoring motif, the use of spore-coat associated enzyme of the inner coat layer was never explored. The results of our feasibility study suggest that oxalate decarboxylase (OxdD) has potential as carrier protein to vehicle bioactive molecules at the spore surface. As a proof-of-concept, spore display of endogenous monomeric (*B. subtilis* phytase) and heterologous oligomeric (*E. coli* β-glucuronidase) passengers was successfully performed.

Finally, **Chapter V** summarizes the topics of this thesis, and puts them in perspective with the future of *B. subtilis* industrial biotechnology.

7. References

Anadon, A., M.R. Martinez-Larranaga, and M.A. Martinez. 2006. Probiotics for animal nutrition in the European Union. Regulation and safety assessment. Regulatory toxicology and pharmacology. 45:91-95.

Arbabi-Ghahroudi, M., J. Tanha, and R. MacKenzie. 2009. Isolation of monoclonal antibody fragments from phage display libraries. Methods Mol. Biol. 502:341-64.

Arraiano, C.M., J.M. Andrade, S. Domingues, I.B. Guinote, M. Malecki, R.G. Matos, R.N. Moreira, V. Pobre, F.P. Reis, M. Saramago, I.J. Silva, and S.C. Viegas. 2010. The critical role of RNA processing and degradation in the control of gene expression. FEMS Microbiol. Rev. 34(5):883-923.

Azuma, Y., and M. Ota. 2009. An evaluation of minimal cellular functions to sustain a bacterial cell. BMC Syst. Biol. (3):111.

Azzazy, H.M., and W.E. Highsmith Jr. 2002. Phage display technology: clinical applications and recent innovations. Clin. Biochem. 35(6):425-45.

Barák, I., E. Ricca, and S.M. Cutting. 2005. From fundamental studies of sporulation to applied spore research. Mol.Microbiol. 55(2): 330-338.

Barbe, S., F.K. Cruveiller, F. Kunst, P. Lenoble, G. Meurice, A. Sekowska, D. Vallenet, T. Wang, I. Moszer, C. Medigue, and A.

Danchin. 2009. From a consortium sequence to a unified sequence: the *Bacillus subtilis* 168 reference genome a decade later. *Microbiology* 155(6):1758-1775.

Bentley, D.R. 2006. Whole-genome re-sequencing. Curr. Opin. Genet. Dev. (6):545-52

Bloor, A.E., and R.M. Cranenburgh. 2006. An efficient method of selectable marker gene excision by Xer recombination for gene replacement in bacterial chromosomes. Appl Environ Microbiol. 72(4):2520-2525.

Blount, K.F., J.X. Wang, J. Lim, N. Sudarsan, and R.R Breaker. 2007. Antibacterial lysine analogs that target lysine riboswitches. Nat. Chem. Biol. 3(1):44-9.

Brans, A., P. Filée, A. Chevigné, A. Claessens, and B. Joris. 2004. Another method to generate *Bacillus subtilis* recombinant strains free of selection markers. Appl Environ Microbiol. 70(12):7241-50.

Burbulys, D., K.A. Trach, and J.A. Hoch. 1991. Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. Cell. 64(3):545-52.

Burkholder, P.R., and N.H. Giles Jr. 1947. Induced biochemical mutations in *Bacillus subtilis*. Am. J. Bot. 34(6):345-8.

Cartman, S.T., R.M. la Regione, and M.J. Woodward. 2008. *Bacillus subtilis* spores germinate in the chicken gastrointestinal tract. Appl. and Environ. Microbiology. 74(16): 5254-5258.

Casula, G., and S.M. Cutting. 2002. *Bacillus* probiotics: spore germination in the gastrointestinal tract. Appl. and Environ. Microbiology. 68(5): 2344-2352.

Costa, T., L. Steil, L.O. Martins, U. Völker, and A.O. Henriques. 2004. Assembly of an oxalate decarboxylase produced under the sigma k control into the *Bacillus subtilis* spore coat. J.Bacteriol. 186(5):1462-1474.

Cozy, L.M., and D.B. Kearns. 2010. Gene position in a long operon governs motility development in *Bacillus subtilis*. Mol. Microbiol. 76(2):273-85.

Cutting, S.M., H.A. Hong, L. Baccigalupi, and E. Ricca. 2009. Oral vaccine delivery by recombinant spore probiotics. Int.Rev.Immunol. 28:487-505.

Daugherty, P.S. 2007. Protein engineering with bacterial display. Curr. Opin. Struct. Biol. 17(4):474-80.

Dauner, M., T. Storni, and U. Sauer. 2001. *Bacillus subtilis* metabolism and energetics in carbon-limited and excess-carbon chemostat culture. J Bacteriol. 183(24):7308-17.

Dricks, A. 2003. The dynamic spore. Proc.Natl.Acad.Sci. 100:3007-3009.

Dworkin, J., and R. Losick. 2001. Differential gene expression governed by chromosomal spatial asymmetry. Cell. 107(3):339-46.

Duc, L.H., H.A. Hong, H.S. Atkins, H.C. Flick-Smith, Z. Durrani, S. Rijpkema, R.W. Titball, and S.M. Cutting. 2007. Immunization against anthrax using Bacillus subtilis spores expressing the anthrax protective antigen. Vaccine. 25:346-355.

Duc L.H., H.A. Hong, T.M. Barbosa, A.O. Henriques, and S.M. Cutting. 2004. Characterization of *Bacillus* probiotics available for human use. Appl Environ Microbiol. 70(4):2161-71.

Errington, J. 2010. From spores to antibiotics via the cell cycle. Microbiology. 156 (1):1-13.

European Food Safety Authority (EFSA). 2009. EFSA Panel on Additives and Products or Substances used in Animal Feed (FEEDAP); Scientific Opinion on the safety and efficacy of *Bacillus subtilis* PB6 as a feed additive for chickens for fattening on request from the European Commission. EFSA Journal. 7(9):1314.

European Food Safety Authority (EFSA[a]). 2007. Introduction of a Qualified Presumption of Safety (QPS) approach for assessment of selected microorganisms referred to EFSA. Opinion of the Scientific Committee. *The EFSA Journal.* 587, 1-16.

European Food Safety Authority (EFSA[d]). 2006. Opinion of the Scientific Panel on Additives and Products or Substances used in Animal Feed on the modification of terms of authorisation of the

micro-organism preparation of *Bacillus licheniformis* and *Bacillus subtilis* (BioPlus 2B) authorised as a feed additive in accordance with Council Directive 70/524/EEC. Compatibility with the coccidiostat maduramicin ammonium. The EFSA Journal. 380: 1-6.

European Food Safety Authority (EFSA[e]). 2006. Opinion of the Scientific Panel on Additives and Products or Substances used in Animal Feed on the safety and efficacy of the microbiological product "035", a preparation of *Bacillus subtilis,* as a feed additive for chickens for fattening in accordance with Regulation (EC) No 1831/2003. The EFSA Journal. 406: 1-11.

European Food Safety Authority (EFSA[b]). 2003. Regulation (EC) No 1831/2003 of the European Parliament and of the Council of 22 September 2003 on additives for use in animal nutrition. Official Journal of the European Union 18.10.2003. L 268:29.

European Food Safety Authority (EFSA[c]). 2003. Regulation (EC) No 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed. Official Journal of the European Union 18.10.2003. L 268:1.

European Food Safety Authority (EFSA[f]). 2007. Scientific Opinion of the Panel on Additives and Products or Substances used in Animal Feed (FEEDAP) on the safety and efficacy of the product Calsporin®, a preparation of *Bacillus subtilis*, as a feed additive for chickens for fattening in accordance with Regulation (EC) No 1831/2003. The EFSA Journal. 543:1-8.

Even, S., O. Pellegrini, L. Zig, V. Labas, J. Vinh, D. Bréchemmier-Baey, and H. Putzer. 2005. Ribonucleases J1 and J2: two novel endoribonucleases in B.subtilis with functional homology to E.coli RNase E. Nucleic Acids Res. 33(7):2141-52.

Fabret, C., S.D. Ehrlich, and P. Noirot. 2002. A new mutation delivery system for genome-scale approaches in *Bacillus subtilis*. Mol Microbiol. 46(1):25-36.

Food and Agriculture Organization of the United Nations (FAO) and the World Health Organization (WHO). 2001. Report of a joint FAO/WHO expert consultation on evaluation of health and nutritional properties of probiotics in food including powder milk with live lactic acid bacteria. Probiotics in food Health and nutritional properties and guidelines for evaluation. Paper 85. ISSN 0254-4725.

Fujiya, M., M.W. Musch, Y. Nakagawa, S. Hu, J. Alverdy, Y. Kohgo, O. Schneewind, B. Jabri, and E.B. Chang. 2007. The Bacillus subtilis Quorum-Sensing Molecule CSF Contributes to Intestinal Homeostasis via OCTN2, a Host Cell Membrane Transporter. Cell Host & Microbe. 1: 299–308.

Gibson, D.G., J.I. Glass, C. Lartigue, V.N. Noskov, R.Y. Chuang, M.A. Algire, G.A. Benders, M.G. Montague, L. Ma, M.M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E.A. Denisova, L. Young, Z.Q. Qi, T.H. Segall-Shapiro, C.H. Calvey, P.P. Parmar, C.A. Hutchison, H.O. Smith, and J.C. Venter. 2010. Creation of a Bacterial Cell

Controlled by a Chemically Synthesized Genome. Science. 329(5987):52-6

Gollnick, P., P. Babitzke, A. Antson, and C. Yanofsky. 2005. Complexity in regulation of tryptophan biosynthesis in *Bacillus subtilis*. Annu. Rev. Genet. 39:47-68.

Grundy, F.J., and T.M. Henkin. 1998. The S box regulon: a new global transcription termination control system for methionine and cystenine biosynthesis genes in gram-positive bacteria. Mol. Microbiol. 30:737–749.

Guérout-Fleury, A.M., N. Frandsen, and P. Stragier. 1996. Plasmids for ectopic integration in *Bacillus subtilis*. Gene.180(1-2):57-61.

Guérout-Fleury, A.M., K. Shazand, N. Frandsen, and P. Stragier. 1995. Antibiotic-resistance cassettes for *Bacillus subtilis*. Gene. 167(1-2):335-6.

Guillot, J.F. 2009. Consequences of probiotics release in the intestine of animals. Université de Tours, IUT, France. 17-21. http://ressources.ciheam.org/om/pdf/c54/01600006.pdf.

Harwood, C.R. 1992. *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. Trends Biotechnol. 10(7):247-56.

Harwood, C.R., and S.M. Cutting. 1990. Molecular biological methods for Bacillus. John Wiley & Sons (eds.), Chichester, England.

Harwood, C.R., and A. Wipat. 1996. Sequencing and functional analysis of the genome of *Bacillus subtilis* strain 168. FEBS Letter. 389(1):84-7.

Henkin T.M., and C. Yanofsky. 2002. Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. Bioessays. 24(8):700-7.

Henriques, A.O., T. Costa, L.O. Martins, and R. Zilãho. 2004. Functional architecture and assembly of the spore coat. In Bacterial spore formers: probiotics and emerging applications. Ricca, E., Henriques A.O. and S.M. Cutting (eds). Norfolk, UK: Horizon Biosci. Press. 65-85.

Henriques, A.O., and C.P. Moran Jr. 2007. Structure, assembly, and function of the spore surface layers. Annu Rev Microbiol. 61:555-588.

Henry, C., K. Tanaka, J. Zinner, M. Cohoon, R. Stevens, and P. Noirot. 2010. Model-driven minimization of the *B. subtilis* Genome. Reported at Systems Biology of Microorganisms Conference. Institut Pasteur, Paris, France. March 2010. From microbial systems to synthetic biology: biotechnology and health.

Hoa, T. T., L. H. Duc, R. Isticato, L. Baccigalupi, E. Ricca, P. H. Van, and S. M. Cutting. 2001. Fate and dissemination of *Bacillus subtilis* spores in a murine model. Appl. Environ. Microbiol. 67:3819–3823.

Hong, H.A., J.M. Huang, R. Khaneja, L.V. Hiep, M.C. Urdaci, and S.M. Cutting. 2008. The safety of *Bacillus subtilis* and *Bacillus indicus* as food probiotics. J. Appl. Micriobiol. 105: 510-520.

Hong, H.A., E. To, S. Fakhry, L. Baccigalupi, E. Ricca, and S.M. Cutting. 2009. Defining the natural habitat of *Bacillus* spore-formers. Res.Microbiol. 160:375-379.

Isticato, R., G. Cangiano, H.T. Tran, A.L. Ciabattini, D. Medaglini, M.R. Oggioni, M. de Felice, G. Pozzi, and E. Ricca. 2001. Surface display of recombinant proteins on *Bacillus subtilis* spores. J. Bacteriol. 183(21):6294-6301.

Ito, S., S.Shikita, K.Ozaki, K. Okamoto, S. Inoue, A. Takei, Y. Ohta, and T. Satoh. 1989. Alkaline cellulase for laundry detergents: production by *Bacillus* sp. KSM-635 and enzymatic properties. Agric. Biol. Chem. 53:1275-81.

Kenrick SA, and P.S. Daugherty. 2010. Bacterial display enables efficient and quantitative peptide affinity maturation. 2010. Protein Eng. Des. Sel. 23(1):9-17.

Kim, H., M. Hahn, P. Grabowski, D.C. McPherson, M.M. Otte, R. Wang, C.C Ferguson, P. Eichenberger and A. Driks. 2006. The *Bacillus subtilis* spore coat protein interaction network. Mol Microbiol. 59(2):487-502.

Kim, J.H., C. Roh, C.W. Lee, D. Kyung, S.K. Choi, H.C. Jung, J.G. Pan, and B.G. Kim. 2007. Bacterial surface display of GFP(uv) on *Bacillus subtilis* spores. J. Microbiol. Biotechnol. 17:677-680.

Kim, J., and W. Schumann. 2009. Display of proteins on *Bacillus subtilis* endospores. Cell.Mol. Life Sci. 66(19):3127-36.

Kobayashi, K., S.D. Ehrlich, A. Albertini, G. Amati, K.K. Andersen, M. Arnaud, K. Asai, S. Ashikaga, S. Aymerich, P. Bessieres, F. Boland, S.C. Brignell, S. Bron, K. Bunai, J. Chapuis, L.C. Christiansen, A. Danchin, M. Débarbouille, E. Dervyn, E. Deuerling, K. Devine, S.K. Devine, O. Dreesen, J. Errington, S. Fillinger, S.J. Foster, Y. Fujita, A. Galizzi, R. Gardan, C. Eschevins, T. Fukushima, K. Haga, C.R. Harwood, M. Hecker, D. Hosoya, M.F. Hullo, H. Kakeshita, D. Karamata, Y. Kasahara, F. Kawamura, K. Koga, P. Koski, R. Kuwana, D. Imamura, M. Ishimaru, S. Ishikawa, I. Ishio, D.L Coq, A. Masson, C. Mauël, R. Meima, R.P. Mellado, A. Moir, S. Moriya, E. Nagakawa, H. Nanamiya, S. Nakai, P. Nygaard, M. Ogura, T. Ohanan, M. O'Reilly, M. O'Rourke, Z. Pragai, H.M. Pooley, G. Rapoport, J.P. Rawlins, L.A. Rivas, C. Rivolta, A. Sadaie, Y.

Sadaie, M. Sarvas, T. Sato, H.H. Saxild, E. Scanlan, W. Schumann, J.F. Seegers, J. Sekiguchi, A. Sekowska, S.J. Séror, M. Simon, P. Stragier, R. Studer, H. Takamatsu, T. Tanaka, M. Takeuchi, H.B. Thomaides, V. Vagner, J.M. van Dijl, K. Watabe, A. Wipat, H. Yamamoto, M. Yamamoto, Y. Yamamoto, K. Yamane, K. Yata, K. Yoshida, H. Yoshikawa, U Zuber, and N. Ogasawara. 2003. Essential *Bacillus subtilis* genes. Proc. Natl. Acad. Sci. USA.100(8):4678-83.

Kolkman, M.A., R. van der Ploeg, M. Bertels, M. van Dijk, J. van der Laan, J.M. van Dijl, and E. Ferrari. 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. Appl. Environ. Microbiol. 74(24):7507-13.

Kondo, A., and M. Ueda. 2004. Yeast cell-surface display-applications of molecular display. Appl. Microbiol. Biotechnol. 64(1):28-40.

Koonin, E.V. 2000. How many genes can make a cell: the minimal-gene-set concept. Annu. Rev. Genomics Hum. Genet. (1):99-116.

Kouwen, T.R., J.Y. Dubois, R. Freudl, W.J. Quax, and J.M. van Dijl. 2008. Modulation of thiol-disulfide oxidoreductases for increased production of disulfide-bond-containing proteins in *Bacillus subtilis*. Appl. Environ. Microbiol. 74(24):7536-45.

Kwon, S.J., H.C. Jung, and J.G. Pan. 2007. Transgalactosylation in a water-solvent biphasic reaction system with beta-galactosidase displayed on the surface of *Bacillus subtilis* spores. Appl Environ Microbiol. **73(7):**2251-2256.

Le Grice, S.F., and A.L. Sonenshein. 1982. Interaction of *Bacillus subtilis* RNA polymerase with a chromosomal promoter. J Mol Biol. 162(3):551-64.

Lee, G., and J. Pero. 1981. Conserved nucleotide sequences in temporally controlled bacteriophage promoters. J. Mol. Biol. 152(2):247-65.

Li, L., X. Hu, Z. Wu, S. Xiong, Z. Zhou, X. Wang, J. Xu, F. Lu, and X. Yu. 2009. Immunogenicity of self-adjuvanticity oral vaccine candidate based on use of *Bacillus subtilis* spore displaying *Schistosoma japonicum* 26 KDa GST protein. Parasitol. Res. 105(6):1643-51.

Mandal, M., B. Boese, J.E. Barrick, W.C. Winkler, and R.R. Breaker. 2003. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. Cell. 113(5):577-86.

Margulies, M., M. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bemben, J. Berka, M.S. Braverman, Y.J. Chen, Z. Chen, S.B. Dewell, L. Du, J.M. Fierro, X.V. Gomes, B.C. Godwin, W. He, S. Helgesen, C.H. Ho, G.P. Irzyk, S.C. Jando, M.L. Alenquer, T.P. Jarvie, K.B., J.B. Kim, J.R. Knight, J.R. Lanza, J.H. Leamon, S.M. Lefkowitz, M. Lei, J. Li, K.L. Lohman, H. Lu, V.B. Makhijani, K.E. McDade, M.P. McKenna, E.W. Myers, E.

Nickerson, J.R. Nobile, R. Plant, B.P. Puc, M.T. Ronan, G.T. Roth, G.J. Sarkis, J.F. Simons, J.W. Simpson, M. Srinivasan, K.R. Tartaro, A. Tomasz, K.A. Vogt, G.A. Volkmer, S.H. Wang, Y. Wang, M.P. Weiner, P. Yu, R.F. Begley, and J.M. Rothberg. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 437(7057):376-80.

Mathy, N., L. Bénard, O. Pellegrini, R. Daou, T. Wen, and C. Condon. 2007. 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. Cell. 129(4):681-92.

Mauriello, E.M., L.H Duc., R. Isticato, G. Cangiano, H.A. Hong, F.M. De, E. Ricca, and S.M. Cutting. 2004. Display of heterologous antigens on the *Bacillus subtilis* spore coat using CotC as a fusion partner. Vaccine. 22: 1177-1187.

Mironov, A.S., I. Gusarov, R. Rafikov, L.E. Lopez, K. Shatalin, R.A. Kreneva, D.A. Perumov, and E. Nudler. 2002. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. Cell. 111(5):747-56.

Montange, R.K., and R.T. Batey. 2006. Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. Nature. 441(7097):1172-5.

Moran, C.P. Jr, N. Lang, S.F. LeGrice, G. Lee, M. Stephens, A.L. Sonenshein, J. Pero, and R. Losick. 1982. Nucleotide sequences

that signal the initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. 186(3):339-46.

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

Nakano, M.M. and P. Zuber. Anaerobic growth of a "strict aerobe" (*Bacillus subtilis*). 1998. Review of Microbiology. 52: 165-190.

Nahvi, A., N. Sudarsan, M.S. Ebert, X. Zou, K.L. Brown, and R.R Breaker. 2002. Genetic control by a metabolite binding mRNA. Chem. Biol. 9:1043.

Nicolas, P., A. Leduc, S. Robin, S. Rasmussen, H. Jarmer, and P. Bessières. 2009. Transcriptional landscape estimation from tiling array data using a model of signal shift and drift. Bioinformatics. 25(18):2341-7.

Park, T.J., S.K. Choi, H.C. Jung, S.Y. Lee, and J.G. Pan. 2009. Spore display using *Bacillus thuringiensis* exosporium protein InhA. J Microbiol Biotechnol. 19(5):495-501.

Perkins, J.B., A. Sloma, T. Hermann, K. Theriault, E. Zachgo, T. Erdenberger, N. Hannett, N.P. Chatterjee, V. Williams II, G.A. Rufo Jr, R. Hatch, and J. Pero. 1999. Genetic engineering of

Bacillus subtilis for the commercial production of riboflavin. J. Ind. Microbiol. Biotechnol. 22:8-18.

Perkins, J.B., M. Wyss, H.P. Hohmann, and U.Sauer. 2009. **Metabolic engineering of** *Bacillus subtilis*. In Smolke, C.D. (eds.). The metabolic pathway engineering handbook: fundamentals. CRC Press, Boca Raton, FL. 23:1-25.

Piggot, P.J., and D.W. Hilbert. 2004. Sporulation of *Bacillus subtilis*. Curr. Opin. Microbiol.
7(6):579-86.

Rasmussen, S., H.B. Nielsen, and H. Jarmer. 2009. The transcriptionally active regions in the genome of *Bacillus subtilis*. Mol. Microbiol. 73(6):1043-57.

Régnier, P., and C.M. Arraiano. 2000. Degradation of mRNA in bacteria: emergence of ubiquitous features. Bioessays. 22(3): 235–244.

Ricca, E., and S.M. Cutting. 2003. Emerging applications of bacterial spores in Nanobiotechnology. J. Nanobiotechnology. 1(1):6.

Rodionov, D.A., A.G. Vitreschak, A.A. Mironov, and M.S. Gelfand. 2002. Comparative genomics of thiamin biosynthesis in prokaryotes, J. Biol. Chem. 277:48949–59.

Sharp, J.S., and D.H. Bechhofer. 2005. Effect of 5'-proximal elements on decay of a model mRNA in *Bacillus subtilis*. Mol. Microbiol. 57(2):484-95.

Samuelson, P., E. Gunneriusson, P.A. Nygren, and S. Ståhl. 2002. Display of proteins on bacteria. J. Biotechnol. 96(2):129-54.

Simpson, A.J. 2001. Genome sequencing networks. Nat. Rev. Genet. 2(12):979-83.

Sonenshein, A.L. 2000. Control of sporulation initiation in *Bacillus subtilis*. Curr.Opin. Microbiol. 3(6):561-6.

Stragier, P., and R. Losick R. 1996. Molecular genetics of sporulation in *Bacillus subtilis*. Annu. Rev. Genet. 30:297-41.

Sudarsan, N., S. Cohen-Chalamish, S. Nakamura, G.M. Emilsson, and R.R. Breaker. 2005. Thiamin pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamin. Chem. Biol. 12:1325-35.

Sudarsan, N., J.K. Wickiser, S. Nakamura, M.S Ebert, and R.R Breaker. 2003. An mRNA structure in bacteria that controls gene expression by binding lysine. Genes Dev. 17(21):2688-97.

Tam, N.K., Q.U. Nguyen, A.H. Huynh, L.H Duc, T.H. Tran, C.R. Serra, A.O. Henriques, and S.M. Cutting. 2006. The intestinal life cycle of *Bacillus subtilis* and close relatives. J. Bacteriol. **188(7)**:2692-2700.

Tanaka, T., S. Masunari, J. Ishii, K. Wakamura, M. Segawa, H. Fukuda, and A. Kondo. 2010. Displaying non-natural, functional molecules on yeast surfaces via biotin-streptavidin interaction. J. Biotechnol. 145(1):79-83.

Tanino, T., T. Aoki, W.Y. Chung, Y. Watanabe, C. Ogino, H. Fukuda, and A. Kondo. 2009 Improvement of a *Candida antarctica* lipase B-displaying yeast whole-cell biocatalyst and its application to the polyester synthesis reaction. Appl. Microbiol Biotechnol. 82(1):59-66.

Thakore, Y. 2008. Enzymes for Industrial Applications. Report BIO030E. BCC Research (Business Communications Company, Inc., Wellesley, MA, USA).

Tjalsma, H., H. Antelmann, J.D. Jongbloed, P.G. Braun, E. Darmon, R. Dorenbos, J.Y. Dubois, H. Westers, G. Zanen, W.J. Quax, O.P. Kuipers, S. Bron, M. Hecker, and J.M. van Dijl. 2004. Proteomics of protein secretion by *Bacillus subtilis*: separating the "secrets" of the secretome. Microbiol. Mol. Biol. Rev. 68(2):207-33.

Tosato, V., and C.V. Bruschi. 2004. Knowledge of the *B. subtilis* genome: impacts on fundamental science and biotechnology. Appl. Microbiol. Biotechnol. 64:1-6.

Vagner, V., E. Dervyn, and S.D. Ehrlich. 1998. A vector for systematic gene inactivation in *Bacillus subtilis*. Microbiology. 144:3097-3104.

Van Dijl, J.M., P.G. Braun, C. Robinson, W.J. Quax, H. Antelmann, M. Hecker, J. Müller, H. Tjalsma, S. Bron, and J.D. Jongbloed. 2002. Functional genomic analysis of the *Bacillus subtilis* Tat pathway for protein secretion. J Biotechnol. 98(2-3):243-54.

Vitikainen, M., H.L. Hyyryläinen, A. Kivimäki, V.P. Kontinen, and M. Sarvas. 2005. Secretion of heterologous proteins in *Bacillus subtilis* can be improved by engineering cell components affecting post-translocational protein folding and degradation. J. Appl.Microbiol. 99(2):363-75.

Vitreschak, A.G., D.A. Rodionov, A.A. Mironov, and M.S. Gelfand. 2002. Regulation of riboflavin biosynthesis and transport genes in bacteria by translational and transcriptional attenuation. Nucleic Acids Res. 30:3141–51.

Walch, A. 1996. PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae*. Yeast **12**.259-265

Wang, P.Z., and D.H. Doi. 1984. Overlapping promoters transcribed by *Bacillus subtilis* sigma 55 and sigma 37 RNA polymerase holoenzymes during growth and stationary phases. J. Biol. Chem. 259(13) 8619-25.

Wang, K.H., A.L. Isidro, L. Domingues, H.A. Eskandarian, P.T. McKenney, K. Drew, P., Grabowski, M.H. Chua, S.N. Barry, M. Guan, R. Bonneau, A.O. Henriques, and P. Eichenberger. 2009.

The coat morphogenetic protein SpoVID is necessary for spore encasement in *Bacillus subtilis*. Mol. Microbiol. 74(3):634-49.

Weisblum, B., M. Y.Graham, T. Gryczan, and D. Dubnau. 1979. Plasmid copy number control: isolation and characterization of high-copy-number mutants of plasmid pE194. J. Bacteriol. 137:635-643.

Wernérus, H., and S. Stahl. 2004. Biotechnological applications for surface-engineered bacteria. Biotechnol. Appl. Biochem. 40(3):209-28.

Westers, H., R. Dorenbos, J.M. van Dijl, J. Kabel, T. Flanagan, K.M. Devine, F. Jude, S.J. Seror, A.C. Beekman, E. Darmon, C. Eschevins, A. de Jong, S. Bron, O.P Kuipers, A.M. Albertini, H. Antelmann, M. Hecker, N. Zamboni, U. Sauer, C. Bruand, D.S. Ehrlich, J.C. Alonso, M. Salas, and W.J. Quax. 2003. Genome engineering reveals large dispensable regions in *Bacillus subtilis*. Mol. Biol. Evol. 12:2076-90.

Westers, L., H. Westers, and W.J. Quax. 2004. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. Biochim. and Biophys. Acta. 1694(1-3): 299-310.

Winkler[a], W., S. Cohen-Chalamish, and R.R. Breaker. 2002. An mRNA structure that controls gene expression by binding FMN. Proc. Natl. Acad. Sci. USA. 99:15908–13.

Winkler[b], W., A. Nahvi, and R.R. Breaker. 2002. Thiamin derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature. 419:952–956.

Wu, X.C., W. Lee, L. Tran, and S.L. Wong. 1991. Engineering a *Bacillus subtilis* expression-secretion system with a strain deficient in six extracellular proteases. J. Bacteriol. 173(16):4952-8.

Wu, S.C., J.C Yeung, Y. Duan, R. Ye R, S.J. Szarka, H.R. Habibi, and S.L. Wong. 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. Appl. Environ. Microbiol. 68(7):3261-9.

Wu, S.C., R.Ye, X.C. Wu, S.C. Ng, and S.L. Wong. 1998. Enhanced secretory production of a single-chain antibody fragment from *Bacillus subtilis* by coproduction of molecular chaperones. J Bacteriol. 1998 Jun;180(11):2830-5.

Yan, X., H.J. Yu, Q. Hong, and S.P. Li. 2008. Cre/*lox* system and PCR-based genome engineering in *Bacillus subtilis*. Appl. Environ. Microbiol. 74(17):5556-5562.

Yim, S.K., H.C. Jung, J.G. Pan, H.S. Kang, T. Ahn, and C.H. Yun. 2006. Functional expression of mammalian NADPH-cytochrome P450 oxidoreductase on the cell surface of *Escherichia coli*. Protein Expr. Purif. 49(2):292-8.

Yim, S.K., D.H. Kim, H.C. Jung, J.G. Pan, H.S. Kang, T. Ahn, and C.H. Yun. 2010. Surface display of heme- and diflavin-containing cytochrome P450 BM3 in *Escherichia coli*: a whole cell biocatalyst for oxidation. J. Microbiol. Biotechnol. (4):712-7.

Youngmann, P., J. Perkins, and R. Losick. 1984. Construction of a cloning site near one end of Tn917 into which foreign DNA may be inserted without affecting transposition in *Bacillus subtilis* or expression of the transposon-borne *erm* gene. Plasmids. 12:1-9.

Yu, H., X. Yan, W. Shen, Y. Shen, J. Zhang, and S. Li. 2010. Efficient and precise construction of markerless manipulations in the *Bacillus subtilis* genome. Microbiol Biotechnol. 20(1):45-53.

Zamboni, N., N. Mouncey, H.P. Hohmann, and U. Sauer. 2003. Reducing maintenance metabolism by metabolic engineering of respiration improves riboflavin production by *Bacillus subtilis*. Metab. Eng. 5(1):49-55.

Zeigler, D.R., and J.B. Perkins. 2008. Bacillus. In Practical Handbook of Microbiology. Goldman, E., and L. Green (eds.). Boca Raton, FL: CRC Press.

Zeigler, D.R., Z. Prágai, S. Rodriguez, B. Chevreux, A. Muffler, T. Albert, R. Bai, M. Wyss, and J.B. Perkins. 2008. The origins of 168, W23, and other Bacillus subtilis legacy strains. J.Bacteriol. 190(21):6983-6995.

Zhou, Z., H. Xia, X. Hu, Y. Huang, Y. Li, L. Li, C. Ma, X. Chen, F. Hu, and Z. Wu. 2008. Oral administration of a *Bacillus subtilis*

spore-based vaccine expressing *Clonorchis sinensis* tegumental protein 22.3 kDa confers protection against *Clonorchis sinensis*. Vaccine. 26:1817-1825.

Zweers, J.C., I. Barák, D. Becher, A.J. Driessen, M. Hecker, V.P. Kontinen, M.J. Saller, L. Vavrová, and J.M. van Dijl. 2008. Towards the development of *Bacillus subtilis* as a cell factory for membrane proteins and protein complexes. Microb. Cell. Fact. 4:7-10.

CHAPTER II

A new thiamin salvage pathway

Amy Haas Jenkins[Ŧ], Ghislain Schyns[¤], Sébastien Potot[¤],
Guangxing Sun[Ŧ] and Tadhg P. Begley[Ŧ]

Nature Chemical Biology. 2007. 3(8):492-7.

[Ŧ] Department of Chemistry and Chemical Biology, 120 Baker Laboratory, Cornell University, Ithaca, New York 14853, USA.

[¤] DSM Nutritional Products, Biotechnology R&D, PO box 3255, Bldg 203/25, CH-4002 Basel, Switzerland.

Contribution:

Sébastien Potot constructed all the *B. subtilis* mutant strains and conducted the *in vivo* confirmation of the newly identified thiamin salvage pathway in *B. subtilis.*

Abstract

The physiological function for thiaminase II, a thiamin-degrading enzyme, has eluded investigators for more than 50 years. Here, we demonstrate that this enzyme is involved in the regeneration of the thiamin pyrimidine rather than in thiamin degradation, and we identify a new pathway involved in the salvage of base-degraded forms of thiamin. This pathway is widely distributed among bacteria, archaea and eukaryotes. In this pathway, thiamin hydrolysis products such as N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine (formylaminopyrimidine; **15**) are transported into the cell using the ThiXYZ transport system, deformylated by the *ylmB*-encoded amidohydrolase and hydrolyzed to 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP; **6**) - an intermediate on the de novo thiamin biosynthetic pathway. To our knowledge this is the first example of a thiamin salvage pathway involving thiamin analogs generated by degradation of one of the heterocyclic rings of the cofactor.

Notes:

Numbers in bold in the text body refer to the compounds (including those represented in the schemes and figures).

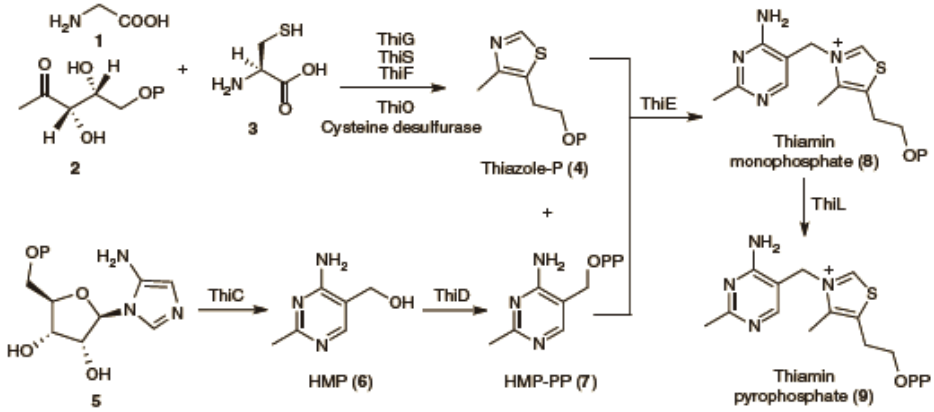
The denomination *thiA*/ThiA in the thiamin biosynthetic pathway, used by the authors in this chapter is a synonym of *thiC*/ThiC, used in both chapters III and V of this thesis (as defined in Subtilist [http://genolist.pasteur.fr/SubtiList] and SubtiWiki [www.subtiwiki.uni-goettingen.de]).

Introduction

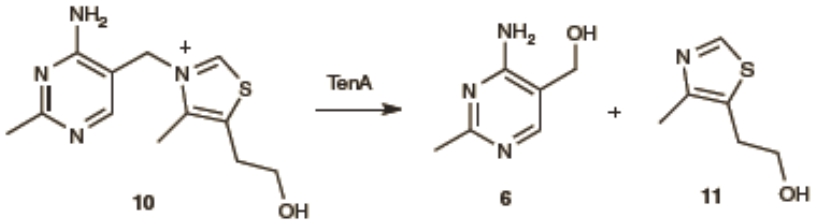
Thiamin is an essential cofactor in all living systems. It is biosynthesized *de novo* in microorganisms and plants via a complex pathway, but it is not synthesized in humans and is therefore an essential nutrient. The biosynthesis of thiamin by microorganisms has been studied extensively and most of the genes involved have been characterized (Scheme 1) [Begley, 2006; Begley and Ealick, 2004].

In *Bacillus subtilis*, HMP-PP (**7**) is produced by rearrangement of aminoimidazole ribonucleotide (AIR; **5**) [Lawhorn *et al*., 2004] followed by phosphorylation [Park *et al*., 2004], and the thiazole phosphate (**4**) is formed by an oxidative condensation of glycine (**1**), 1-deoxy-D-xylulose-5-phosphate (DXP; **2**) and cysteine (**3**) [Dorrestein *et al*., 2004]. Thiamin phosphate (**8**) is then formed by the coupling of the pyrimidine and the thiazole heterocycles [Hanes *et al*., 2007]. A final phosphorylation gives thiamin pyrophosphate (**9**) [Webb and Downs, 1997], the biologically active form of the cofactor. TenA was recently identified as a bacterial thiaminase II, an enzyme that catalyzes the hydrolysis of thiamin (**10**) to give HMP and hydroxyethylthiazole (**11**) (Scheme 2) [Toms *et a*l., 2005].

In many bacteria, the *tenA* gene clusters on the chromosome with or is fused to other thiamin biosynthetic genes (Supplementary Fig. 1) [Haas *et al*., 2005]. This chromosomal clustering is not consistent with a degradative function for this enzyme; therefore, though the structure and previously identified activity of TenA strongly suggested that it catalyzes a pyrimidine substitution reaction [Toms et *al*., 2005], thiamin is unlikely to be the physiological substrate for this enzyme.



Scheme 1: The main pathway for thiamin pyrophosphate biosynthesis in bacteria.



Scheme 2: The previously identified TenA-catalyzed thiamin degradation

Analysis of the genes that cluster on the chromosome with *tenA* in sequenced bacteria provided important clues as to the identity of the TenA substrate (Supplementary Fig. 1). In *B. subtilis*, *tenA* is clustered on the chromosome with a set of genes involved in the biosynthesis of the thiazole moiety of thiamin (*thiO*, *thiS*, *thiG*, *thiF*) and HMP pyrophosphorylation (*thiD*), whereas in *Bacillus halodurans tenA* clusters with a putative amidohydrolase (*ylmB*) and an ABC transporter (*thiX*, *thiY*, *thiZ*) [Rodinov *et al.,* 2002]. In *Bacillus cereus*, *tenA* clusters with *thiX*, *thiY* and *thiZ* and *thiE*, *thiO*, *thiS*, *thiG*, *thiF* and *thiD*. Each of these clusters also contains an

upstream TPP riboswitch, which suggests thiamin regulation [Miranda-Rios, 2007; Winkler, 2002]. The sequence similarity of ThiY to the HMP synthase (Thi5) in *Saccharomyces cerevisiae* suggests that ThiXYZ is involved in the transport of HMP analogs [Rodinov *et al*., 2002]. Given that *B. halodurans* grows in basic soil (pH>10) [Horikoshi and Alkaliphiles, 1999; Takami and Horikoshi, 1999], it seemed likely that these pyrimidine analogs could be generated by base-catalyzed degradation of the thiamin thiazole in soil (Fig. 1a) [Chahine and Dubois, 1983; Maier and Metzler, 1957]. We therefore proposed that YlmB, TenA and ThiXYZ are involved in the salvage of HMP from base-degraded forms of thiamin formed in soil.

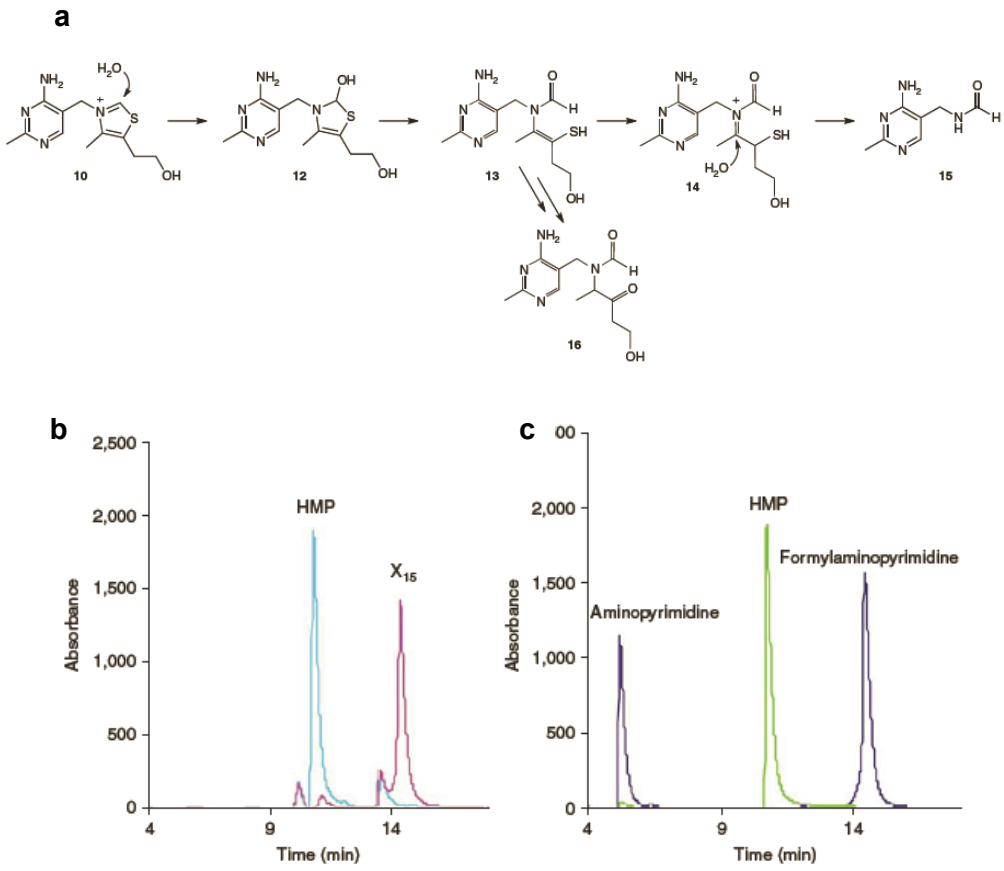


Figure 1: Formation of formylaminopyrimidine and its identification in a mixture of thiamin degraded in the presence of soil. (a) Mechanistic proposal for the formation of formylaminopyrimidine (**15**) by thiamin hydrolysis. **(b)** HPLC analysis of partially purified X_{15} (pink trace) and its conversion to HMP (**6**) catalyzed by YlmB and TenA (light blue trace). **(c)** HPLC chromatogram of the reactions of YlmB and TenA with synthesized formylaminopyrimidine and aminopyrimidine. Dark blue trace: reaction of YlmB (20 μM) with formylaminopyrimidine (2 mM, 30 min) showing its conversion to aminopyrimidine. Green trace: reaction of TenA (1 μM) with aminopyrimidine (5 mM, 30 min) showing its conversion to HMP.

Results

This paper describes the experimental validation of this hypothesis and the identification of the first example of a thiamin salvage pathway involving a heterocycle-degraded form of the cofactor.

Though the high soil pH at which *B. halodurans* grows was an important clue to the function of TenA, the proposed thiamin degradation chemistry must also occur under neutral conditions because *B. subtilis* and *B. cereus* both grow in neutral soil. We therefore focused our attention on thiamin degradation under the milder reaction conditions in neutral soil to simplify the decomposition reactions mixtures. Even under these conditions, HPLC analysis of a reaction mixture generated by treating thiamin with sterilized soil (pH 7, 15 d, 25 °C) revealed the presence of a large number of thiamin degradation products relative to a soil-free control that showed only low levels of decomposition after 6 months (Supplementary Fig. 2). This thiamin decomposition reaction is presumably catalyzed by metal-ion-bound hydroxide on the clay surface, as has been described previously for other clay-catalyzed hydrolysis reactions [Xu *et al*., 2001].

A solution of soil-degraded thiamin was treated with the recombinant purified proteins YlmB and TenA at room temperature for 3 h. HPLC analysis of the resulting reaction mixture demonstrated that in addition to the expected TenA-catalyzed thiamin hydrolysis, a component eluting at 15 min (X_{15}) disappeared from the reaction mixture (Supplementary Fig. 3). In the absence of YlmB, this compound was unaltered under these reaction conditions.

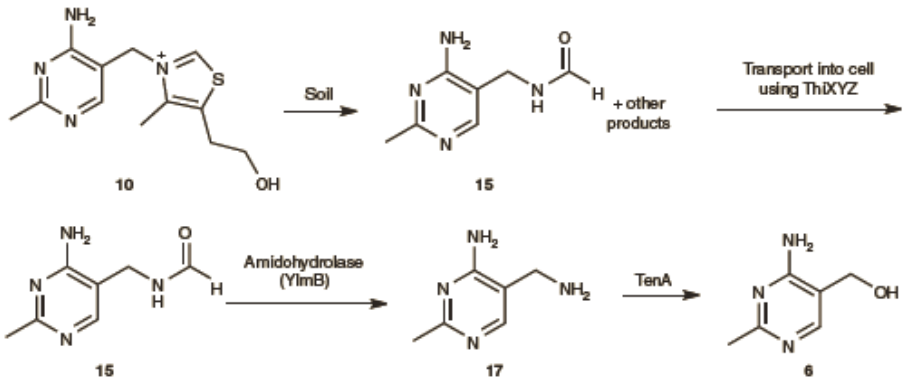
To determine whether any additional YlmB or TenA substrates were buried in the complex mixture of compounds eluting between 18 and 25 min, HPLC fractions containing these compounds were collected, treated with a mixture of TenA and YlmB and reanalyzed by HPLC (Supplementary Fig. 4). None of the components of this mixture were converted to HMP, which demonstrates that these compounds are not relevant to the proposed HMP salvage pathway.

The crude thiamin decomposition mixture was greatly enriched in the YlmB substrate (X_{15}) by chromatography on silica gel to remove the positively charged unreacted thiamin from the neutral degradation products. The resulting mixture was again treated with YlmB and TenA. HPLC analysis of this reaction mixture demonstrated that a combination of YlmB and TenA catalyzes the conversion of X_{15} to HMP (Fig. 1b).

X_{15} was further purified by HPLC. Its structure was determined to be N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine (**15**, Fig. 1a) by mass spectrometry and NMR analysis. A mechanistic proposal for the formation of this product is outlined in Figure 1a [Chahine and Dubois, 1983]. In this mechanism, water addition to the thiazolium heterocycle of **10** gives **12**, which then undergoes ring opening to give **13**. Tautomerization to **14** followed by hydrolysis gives the formylaminopyrimidine (**15**).

Formylaminopyrimidine (**15**) was synthesized by formylation of aminopyrimidine (**17**) and demonstrated to be identical to X_{15} by NMR, mass spectrometry and HPLC analysis (Supplementary Fig. 5). YlmB catalyzed the deformylation of **15** to give aminopyrimidine, and TenA catalyzed the hydrolysis of this aminopyrimidine to give hydroxypyrimidine (**6**) (Fig. 1c). The hydroxypyrimidine formed by TenA can then be phosphorylated by ThiD and incorporated into the

thiamin biosynthetic pathway (Scheme 1). Formylaminopyrimidine is not a substrate for TenA and is stable under the reaction conditions. We have also demonstrated that ThiY (the putative substrate-binding component of the ABC transporter ThiXYZ) binds formylaminopyrimidine and probably delivers it to the ThiXYZ transport system. These identified functions suggest the HMP salvage pathway shown in scheme 3.



Scheme 3: The identified functions for ThiY, YlmB, and TenA, which suggest a new salvage pathway for the thiamin pyrimidine (6).

The kinetic parameters for the YlmB-catalyzed reaction are k_{cat} = 14.0 ± 0.7 min^{-1}, K_m = 5 ± 1.6 μM and k_{cat}/K_m = 2.8 ± 0.97 min^{-1} μM^{-1}, and the kinetic parameters for the TenA-catalyzed reaction are k_{cat} = 22.0 ± 0.48 min^{-1}, K_m = 11.8 ± 1.6 μM and k_{cat}/K_m = 1.9 ± 0.26 min^{-1} μM^{-1}. The YlmB assay was based on formate detection using formate dehydrogenase and monitoring NADH production [Quayle, 1966], whereas the TenA assay was based on ammonia detection using α-ketoglutarate/glutamate dehydrogenase and monitoring NADPH consumption [Day and Keillor, 1999]. The equilibrium binding constant for the formylaminopyrimidine binding

to ThiY was determined to be 200 nM by measuring the change in protein fluorescence upon binding. Thiamin does not bind to ThiY.

To evaluate the *in vivo* relevance of this HMP salvage pathway, we examined the ability of TenA and ThiA (HMP-requiring) mutants in *B. subtilis* to grow on formylaminopyrimidine, aminopyrimidine, and thiamin that was completely base degraded. The results of these experiments are shown in Table 1 and Supplementary Figure 6. The ThiA (also known as ThiC) mutant is HMP-requiring, as expected, and grows on all of the hydroxypyrimidine sources. The TenA mutant does not require a hydroxypyrimidine source as it is able to biosynthesize it using ThiA. The TenA/ThiA double mutant, however, is hydroxypyrimidinerequiring and is unable to salvage the pyrimidine from formylaminopyrimidine, aminopyrimidine, or base-degraded thiamin.

Table 1: Complementation studies on ThiA and TenA mutants in *B. subtilis.*

	***B. subtilis* strain/mutant**			
	ThiA [(1)]	**TenA**	**ThiA [(1)] TenA**	**WT**
Thiamin	+	+	+	+
HMP	+	+	+	+
Aminopyrimidine	+	+	-	+
Formylaminopyrimidine	+	+	-	+
Base-degraded thiamin	+	+	-	+
Nothing	-	+	-	+

[(1)] ThiA is synonym to ThiC (used in Chapter III and V of this thesis).

Although we have not carried out a comprehensive study on the substrate tolerance of TenA or YlmB, it is highly likely that other analogs of thiazole-degraded thiamin are also substrates for this

salvage pathway. In our identification of formylaminopyrimidine as a substrate for YlmB, we deliberately chose mild thiamin degradation conditions (soil, pH = 7) to simplify the degradation reaction mixtures and to mimic biological conditions. Under harsher reaction conditions, a larger set of reaction products is generated and it is likely that several additional substrates for YlmB and TenA exist in such mixtures. We have demonstrated that degradation product **16** [Kurata *et al.*, 1968] (Fig. 1a) is not a substrate for YlmB or TenA, whereas an analog of **15** with a hydroxyethyl substituent (**19**) on the amide nitrogen is a good substrate (data not shown). This demonstrates that some degree of degradation of the C_5 chain of the thiazole is required to generate YlmB substrates.

The previously reported thiaminase II activity of TenA is unlikely to be physiologically relevant because TenA can be overexpressed in *Escherichia coli* at a high level without inducing toxicity or a thiamin requirement. This is most likely due to the substrate selectivity of TenA, which catalyzes the hydrolysis of thiamin but not thiaminphosphate or pyrophosphate, the only forms of biosynthesized thiamin in the cell. In addition, though the TenA catalyzed thiamin hydrolysis does not follow Michaelis-Menten kinetics, it was possible to determine, using a competition assay that TenA catalyzes the hydrolysis of aminopyrimidine 100 times faster than it catalyzes the hydrolysis of thiamin, thereby further adding to the selectivity of TenA for thiamin degradation products.

The three previously identified pathways for thiamin salvage involve phosphorylation of the biosynthetic intermediates HMP, thiamin and hydroxyethylthiazole [Imamura and Nakayama, 1981; Imamura and Nakayama, 1982; Melnick *et al*., 2004; Mizote and Nakayama, 1989]. These pathways were relatively easy to find as an extension of the biosynthetic studies. The identification of a

salvage pathway involving products resulting from the degradation of one of the thiamin heterocyclic rings suggests that additional salvage pathways involving other thiamin decomposition products are likely to exist.

The identification of the physiological function of thiaminase I and II has been a long-standing unsolved problem. Although we do not yet clearly understand the physiological function of thiaminase I, the experiments described in this paper suggest that pyrimidine salvage rather than thiamin degradation is the physiological function of thiaminase II. The wide distribution of TenA in bacteria, archaea and eukaryotes suggests the presence of this salvage pathway in all three kingdoms of life. YlmB is less widely distributed and frequently not present in organisms containing TenA, which suggests that the deformylation of **15** can be catalyzed by a variety of amidohydrolases. The wide distribution of TenA also suggests that destruction of the thiamin thiazole is a commonly occurring thiamin degradation reaction and that this new salvage pathway is widespread and not restricted to soil-growing bacteria.

Materials and Methods

Soil-catalyzed thiamin degradation.

A solution of thiamin (100 ml, 100 mM, pH 7 phosphate buffer) containing 10 g of soil was vigorously stirred at room temperature for 15 d and filtered. The resulting solution was used as a substrate for YlmB and TenA. The soil was obtained from the lawn behind S.T. Olin laboratory at Cornell University in Ithaca, New York, USA and sterilized by autoclaving at 120 °C for 20 min before use. Purification of X_{15} was carried out by slowly pouring 4 ml of this reaction mixture onto a one-inch layer of silica previously equilibrated in 90% chloroform, 10% methanol in a sintered glass funnel. The silica was then washed with 1.5 l of 90% chloroform, 10% methanol, the solvent was removed and the resulting sample of crude X_{15} was dissolved in 3 ml of water. A 250 ml aliquot of this sample was further purified by HPLC as described below. The fractions containing X_{15} in phosphate buffer were collected, the solvent was removed and the sample was dissolved in D_2O. An NMR spectrum was obtained on a Bruker 300 MHz instrument, δ 2.33 (s, 3H), 4.21 (s, 2H), 7.91 (s, 1H), 8.12 (s, 1H). The dried sample was extracted from the phosphate salts with methanol and analyzed by ESI mass spectrometry (Esquire mass spectrometer from Bruker with an ion-trap mass analyzer) to give a single ion with an *m/z* [M^+] ratio of 167.

Synthesis of formylaminopyrimidine (15).

A suspension of aminopyrimidine (**17**) (276 mg, 2.0 mmol) and 1H-benzotriazole-1-carboxyaldehyde (**18**) (Sigma) (294 mg, 2.0 mmol) in acetonitrile (30 ml) was stirred at room temperature for 1 h. Filtration and washing with acetonitrile (15 ml) gave pure product

as a white solid (240 mg, 72%). ^{1}H NMR (300 MHz, D_2O): δ 2.33 (s, 3H, CH_3), 4.21 (s, 2H, CH_2), 7.92 (s, 1H, CH), 8.12 (s, 1H, NCHO). ^{13}C NMR (300 MHz, D2O): δ 23.71, 35.55, 110.68, 153.85, 161.62, 164.74,166.59. ESI mass spectrometry (Esquire mass spectrometer from Bruker with an ion-trap mass analyzer) gave a single ion with (*m/z*): $[M]^+$ calcd. for formylaminopyrimidine, 167.2; found, 167.4.

Synthesis of N-((4-amino-2-methylpyrimidin-5-yl)methyl)-N-(2-hydroxyethyl)formamide (19).

Ethanolamine (**20**) (940 mg, 15.4 μmol) was added to a solution of 5-(chloromethyl)-2-methylpyrimidin-4-amine (**21**) (1.54 mmol) [Dornow and Petsch, 1953] in chloroform (30 ml). The reaction was stirred at room temperature for 12 h. Solvent was removed and chromatography (silica gel, chloroform/methanol 5:1) afforded 2-((4-amino-2-methylpyrimidin-5-yl)methylamino)ethanol (**22**) as a white solid (180 mg, 62%). ^{1}H NMR (300 MHz, CD_3OD): δ 2.38 (s, 3H, CH_3), δ 2.69 (t, J = 5.4 Hz, 2H, CH_2N), δ 3.65 (t, J = 5.4 Hz, 2H, CH_2O), δ 3.69 (s, 2H, CH_2N), δ 7.90 (s, 1H, CH). ^{13}C NMR (500 MHz, DMSO-d6): δ 25.16, 47.28, 50.61, 60.19, 111.87, 153.56, 162.47, 165.14. ESI mass spectrometry (Esquire mass spectrometer from Bruker with an ion-trap mass analyzer) gave a single ion with (*m/z*): $[M]^+$ calcd. for **22**, 183.2; found, 183.4.

A solution of **22** (100 mg, 0.53 mmol) and **18** (100 mg, 0.69 mmol) in acetonitrile (20 ml) was stirred at room temperature for 1 h. Chromatography (silica gel, acetonitrile/methanol 2:1) afforded **19** as a white solid. ^{1}H NMR (300 MHz, CD_3OD): δ 2.38 (s, 3H, CH_3) d 3.33 (t, J = 5.1Hz, 2H, CH_2), δ 3.65 (t, J = 5.1 Hz, 2H, CH_2), δ 4.43 (s, 2H, CH_2), δ 7.99 (s, 1H, CH), δ 8.15 (s, 1H, NCHO). ^{13}C NMR (500 MHz, DMSO-d6): δ 24.90, 40.92, 50.10, 59.82, 110.30,

156.24, 163.61, 166.64, 168.26. ESI mass spectrometry (Esquire mass spectrometer from Bruker with an ion-trap mass analyzer) gave a single ion with (*m/z*): $[M]^+$ calcd. for **19**, 211.2; found, 211.4.

Analysis of the thiamin degradation and the TenA/YlmB reaction mixtures.

All HPLC analysis was performed on a C_{18} column (Supelco, Supelcosil LC-18-T 15 cm_ 4.6 cm, 3 μm) equilibrated in 100 mM phosphate buffer at pH 6.6 using a flow rate of 1 ml min^{-1}. The elution method used a gradient defined as follows: 0–3 min 100% phosphate buffer (100 mM, pH 6.6); 3–4 min 90% phosphate buffer, 10% water; 4–10 min 60% phosphate buffer, 33% water, 7% methanol; 10–15 min 60% phosphate buffer, 33% water, 7% methanol; 15–20 min 10% phosphate buffer, 30% water, 60% methanol; 20–30 min 60% phosphate buffer, 33% water, 7% methanol; 30–35 min 100% phosphate buffer. Reaction products were detected by measuring absorbance at 254 nm.

Bioinformatics tools.

Chromosomal clustering of *tenA* was analyzed using The SEED database (http://theseed.uchicago.edu/FIG/index.cgi).

Overexpression and purification of YlmB.

PCR amplification of the *ylmB* gene (GenBank ID 15615241) from *B. halodurans* genomic DNA was performed in the standard way using Pfx Platinum polymerase from Invitrogen. The *ylmB* gene was then cloned into a pET16b vector using Nde1 and Xho1 restriction sites. After cloning, the *ylmB* plasmid was transformed into the BL21 (DE3) expression strain of *E. coli.* YlmB was obtained by inoculating a 1-l culture of LB medium, containing 100 mg l^{-1}

ampicillin, with a 10-ml starter culture. The cells were grown to an optical density at 600 nm (OD_{600}) of 0.6 at 37 °C, at which point expression was induced with 1 mM IPTG. After induction, the cells were grown for 8 h, pelleted, resuspended in lysis buffer (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 10 mM imidazole), lysed by sonication, and then clarified by centrifugation at 39,000g. The resulting cell-free extract was passed through a Ni-NTA column (Qiagen), which was then washed with 5 column volumes of lysis buffer followed by 300 ml of wash buffer (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 20 mM imidazole). The pure protein was eluted with elution buffer (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 250 mM imidazole) and buffer exchanged by gel filtration into 100 mM phosphate buffer pH 7 containing 100 mM NaCl to prevent protein precipitation.

Overexpression and purification of TenA.

tenA (from *B. subtilis*, GenBank ID 16078230) was cloned into a pDESTF1 plasmid (a Gateway-adapted vector based on the pETsystem from Novagen containing a His_6 tag) and transformed into the BL21 Star (DE3) pRare2 expression strain of *E. coli*. TenA was obtained by inoculating a 1-l culture of LB medium, containing 100 mg l^{-1} ampicillin, with a 10-ml starter culture. The cells were grown to an OD_{600} of 0.6 at 37 °C, at which point expression was induced with 1 mM IPTG. After induction, the cells were grown for 8 h, pelleted, resuspended in lysis buffer (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 10 mM imidazole) and lysed by sonication. The resulting extract was clarified by centrifugation at 39,000g and passed over a Ni-NTA column (Qiagen), which was then washed with 5 column volumes of lysis buffer followed by 300 ml of wash buffer (50 mM NaH_2PO_4 pH 8, 300 mM NaCl, 20 mM imidazole). The pure protein

was eluted with elution buffer (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 250 mM imidazole) and buffer exchanged by dialysis against 100 mM phosphate buffer pH 7. The extinction coefficient of TenA was determined to be 74,280 $M^{-1}cm^{-1}$ using a previously described method [Pace *et al*., 1995].

Overexpression and purification of ThiY.

PCR amplification of the *thiY* gene (GenBank ID 15615245) from *B. halodurans* genomic DNA was performed in the standard way using Pfx Platinum polymerase from Invitrogen. To increase the solubility of ThiY the N-terminal 20 amino acids were truncated to remove a hydrophobic membrane anchor. The truncated *thiY* gene was then cloned into a pET16b plasmid using Nde1 and Xho1 restriction sites. After cloning, the *thiY* plasmid was transformed into the BL21 (DE3) expression strain of *E. coli*. ThiY was obtained by inoculating a 1-l culture of LB medium, containing 100 mg l^{-1} ampicillin, with a 10-ml starter culture. The cells were grown to an OD_{600} of 0.6 at 37 °C, at which point expression was induced with 1 mM IPTG. After induction, the cells were grown for 8 h, pelleted, resuspended in lysis buffer (50 mM NaH2PO_4, pH 8, 300 mM NaCl, 10 mM imidazole) and lysed by sonication. The resulting extract was clarified by centrifugation at 39,000g and passed over a Ni-NTA column (Qiagen), which was then washed with 5 column volumes of lysis buffer followed by 300 ml of wash buffer (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 20 mM imidazole). The pure protein was eluted with elution buffer (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 250 mM imidazole) and buffer exchanged by gel filtration into 100 mM phosphate buffer pH 7.

Assay for YlmB-catalyzed hydrolysis of formylaminopyrimidine (15).

The well-characterized formate dehydrogenase assay [Quayle, 1996] was used to measure the YlmB-catalyzed production of formate from formylaminopyrimidine. Before kinetic analysis, 500 mM EDTA was added to YlmB (1 μM) to remove any metal that may have bound during purification, such as Ni^{+2} from the Ni-NTA resin. After incubation, EDTA was removed by gel filtration. In each assay the concentrations of formate dehydrogenase, NAD^+ and YlmB were held constant and the substrate concentration was varied. The oxidation of formate was not rate limiting in any of the assays. Each reaction contained 1 unit of formate dehydrogenase (a unit is defined as the amount that will oxidize 1.0 μmol of formate to CO2 per min in the presence of NAD at pH 7.6 at 37 °C), 1 mM NAD^+, 1 μM YlmB, 5 μM Co^{+2}. The reactions were started by the addition of formylaminopyrimidine with the final concentrations varying from 5 μM to 400 μM. The YlmB concentration was determined using the Bradford assay, and NADH production was monitored by measuring the increase in absorbance at 340 nm. Kinetic parameters were obtained by fitting the initial rate values (less than 10% conversion) to the Michaelis-Menten equation at varying substrate concentrations using the fitting program GraFit (Erithacus Software Ltd.). The parameters are: k_{cat} = 14.0 ± 0.7 min^{-1}, K_m = 5 ± 1.6 μM and k_{cat}/K_m = 2.8 ± 0.97 min^{-1} μM^{-1}.

Assay for TenA-catalyzed hydrolysis of aminopyrimidine (17).

The well-characterized glutamate dehydrogenase–based assay [Day and Keillor, 1999] was used to measure the production of ammonia from aminopyrimidine catalyzed by TenA. In the kinetic assays the concentrations of NADPH, α-ketoglutarate, glutamate

dehydrogenase and EDTA were kept constant, while the aminopyrimidine concentration was varied from 10 μM to 500 μM. Each reaction contained 5 units of glutamate dehydrogenase (unit is defined as the amount of glutamate dehydrogenase that will reduce 1 μmol of α-ketoglutarate to glutamate per min at pH 8.3 at 30 °C), 5 mM α-ketoglutarate, 0.1 mM EDTA, and 250 μM NADPH. The reduction of α-ketoglutarate to glutamate was not rate limiting in any of the assays. All reactions were performed by adding each reaction component to a solution of aminopyrimidine. The reaction was then started by adding TenA (0.5 μM). The TenA concentration was determined from its extinction coefficient, and NADPH consumption was monitored by measuring the decrease in absorbance at 340 nm. Kinetic parameters were obtained by fitting the initial rate values (less than 10% conversion) to the Michaelis-Menten equation at varying substrate concentrations using GraFit. The parameters are: k_{cat} = 22.0 ± 0.48 min^{-1}, K_m = 11.8 ± 1.6 μM and k_{cat}/K_m = 1.9 ± 0.26 min^{-1} μM^{-1}.

Determination of the selectivity of TenA for aminopyrimidine versus thiamin.

TenA (10 μM) was incubated with equimolar amounts (5 mM) of thiamin (**10**) and aminopyrimidine. The reaction was quenched after 2 min by filtration through a membrane with a molecular weight cutoff of 10 kDa to remove the TenA protein. After the quench, 100 μl of the filtrate was analyzed by HPLC. The molar ratio of the products of the hydrolysis of thiamin and aminopyrimidine were determined using standard curves (peak area as a function of moles of compound). The molar ratio of these products was used as a measurement of relative rate. The extent of

thiamin hydrolysis was determined from the moles of thiazole (**11**) produced. The product of the aminopyrimidine hydrolysis was determined by subtracting the moles of thiazole produced from the total moles of hydroxypyrimidine (**6**) produced (thiamin hydrolysis results in a 1:1 ratio of thiazole to hydroxypyrimidine, Scheme 2), to provide the moles of hydroxypyrimidine that were produced exclusively from aminopyrimidine. Hydrolysis of thiamin resulted in 0.0002 ± 0.000008 μmol of product; hydrolysis of aminopyrimidine resulted in 0.02 ± 0.004 μmol of product. From this data, we calculate a relative rate of aminopyrimidine hydrolysis to thiamin hydrolysis of 99 ± 4.

Bacterial sources of YlmB, TenA and ThiY.

The characterization of the pyrimidine salvage pathway was performed using YlmB and ThiY cloned from *B. halodurans* and TenA cloned from *B. subtilis* because these enzymes were the ones that overexpressed and purified to give the most soluble protein. The YlmB from *B. subtilis* has been overexpressed and catalyzes the same reaction as YlmB from *B. halodurans*, and the TenA from *B. halodurans* has also been overexpressed and catalyzes the same reaction as TenA from *B. subtilis*. Therefore, the pyrimidine salvage pathway has been experimentally characterized in both *B. subtilis* (in which TenA is not clustered with YlmB) and *B. halodurans* (in which TenA is clustered with YlmB).

Determination of the formylaminopyrimidine/ThiY binding constant.

The dissociation constant (K_d) for the binding of formylaminopyrimidine (**15**) to ThiY was determined by fluorescence titration, in which the decrease in protein fluorescence was

measured as a function of substrate concentration. Small aliquots (0.75 μl, 186 μM) of substrate were added to a ThiY solution (3 μM) and the protein fluorescence was measured using a Cary Eclipse fluorescence spectrophotometer (Varian). The protein was excited at a wavelength of 280 nm and emission was detected at 340 nm, with slit widths of 5 mm for both excitation and emission. The K_d value was obtained by fitting the fluorescence data by nonlinear regression to the following quadratic equation:

$$F_{obs} = F_o + \Delta F \cdot \frac{P_o + L_o + K_d - \sqrt{(P_o + L_o + K_d)^2 - 4 \cdot P_o \cdot L_o}}{2 \cdot P_o}$$

where F_o is the *y* intercept, ΔF is the overall change in fluorescence, P_o is the protein concentration and L_o is the ligand concentration added [Anderson *et al*., 1988]. All non-linear regression was performed in the program GraFit 5. The K_d for the binding of formylaminopyrimidine to ThiY is 200 nM.

Construction of *B. subtilis* mutants.

The *thiA* auxotroph TH2 strain was constructed by transforming wild-type strain 1A747 (Bacillus Genetic Stock Center) with a low non-congressional concentration of strain 1A603 (containing the *thiA*::Tn917 mutation) chromosomal DNA, with selection on rich medium containing erythromycin (50 mg ml^{-1}) [Schyns *et al*., 2005]. Strain TH140, which contains a deletion of the *tenA* gene, was constructed as follows: long flanking homology PCR [Wach, 1996] was used to generate a deletion mutation in the coding region of the *tenA* open reading frame, in which a 530-base-pair-long nucleotide region of *tenA* was replaced with the chloramphenicol acetyltransferase (encoded by cat_4) resistance cassette [Schyns *et al*., 2005]. To do this, two PCR fragment 'arms'

were first created: 0.2 ml of a 100 μM solution of primers P1*tenA* and P2*tenA*-*cat* or primers P3*tenA*-cat and P4*tenA* (Supplementary Table 1) were added to 0.1 μg 1A747 chromosomal DNA in a 50-μl reaction volume containing 1 μl of 40 mM dNTPs, 5 μl of 10x buffer and 0.75 μl of PCR enzyme (Taq and Tgo), as described by the manufacturer (Expand high-fidelity PCR system, Roche Applied Science). The PCR reaction was performed for 30 cycles using an annealing temperature of 55.7 °C and an elongation time of 45 s. The resulting fragments, called F1 and F2, respectively, were purified and next used as primers in a second round of PCR. F1 and F2 fragments were diluted 50-fold and 1 μl of each was added to 0.1 μg of chromosomal DNA from strain TH12 in a 50-μl reaction volume to amplify the cat_4 cassette. In the first 10 cycles, an annealing temperature of 63 °C and an elongation time of 3 min were used. In the next 20 cycles, the elongation time was extended by 20 s after each cycle. The resulting products were then used in a third round of PCR as a template. The PCR products were diluted 50-fold and 1 μl was combined with 0.2 μl of a 100 μM solution of primer P1*ten A* and P4*tenA* in a 50-μl reaction volume containing dNTPs, buffer and enzyme as described above. The PCR reaction parameters were identical to those used in the second-round PCR. The finished PCR fragments were next transformed into the wild-type strain 1A747, selecting for 5 mg ml^{-1} chloramphenicol resistance on TBAB (Difco) medium (Cm^r). A single Cm^r colony deleted for *tenA* was isolated and named TH140 (Δ*tenA::cat*). The presence of the *cat* cassette was confirmed by diagnostic PCR using P1*tenA* and P4*ten*A, using standard reaction conditions. Finally, both *thiA* and *tenA* mutations were combined by moving the *tenA*::*cat* deletion into the TH2 background at non-congressional

DNA concentrations. The resulting recombinant strain was named TH141.

Complementation growth conditions.

Complementation studies were performed using minimal medium (Spizizen salts, 0.04% tryptophan, 0.04% sodium glutamate and 1% glucose) supplemented with chloramphenicol (5 μg ml^{-1}), erythromycin (1 μg ml^{-1}) or both depending on the antibiotic resistance of the *B. subtilis* mutant. Thiamin and related supplements (Table 1) were at a final concentration of 5 mM. Base-degraded thiamin was prepared by reacting thiamin (1 M) with 5 M NaOH for 9 d followed by neutralization with HCl and chromatography on a plug of silica equilibrated with 90% chloroform, 10% methanol as described above. This sample was thiamin-free by HPLC analysis.

Accession codes.

GenBank: the following sequences were deposited as part of previous studies: the *ylmB* gene from *B. halodurans* genomic DNA (GenBank ID 15615241), *tenA* from *B. subtilis* (GenBank ID 16078230), and the *thiY* gene from *B. halodurans* genomic DNA (GenBank ID 15615245).

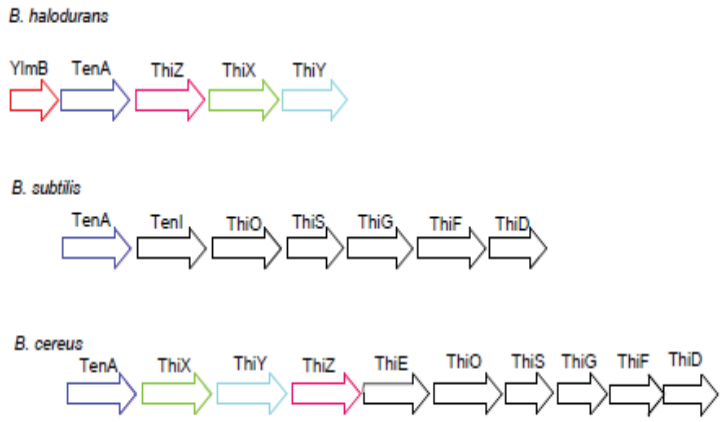
Acknowledgments.

We would like to thank C. Kinsland (Protein Purification Facility, Cornell University) for overexpressing TenA and J. Hanes for his help with the K_d determination. We would also like to thank Roche for providing aminopyrimidine. This research was supported by a grant from the US National Institutes of Health (DK44083).

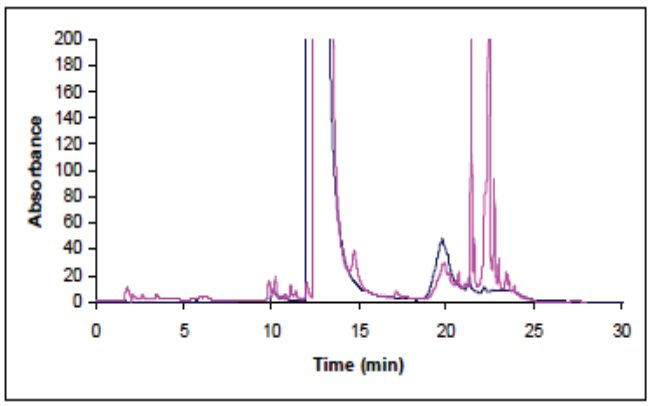
Competing Interests statement.

The authors declare no competing financial interests.

Supplemental Material

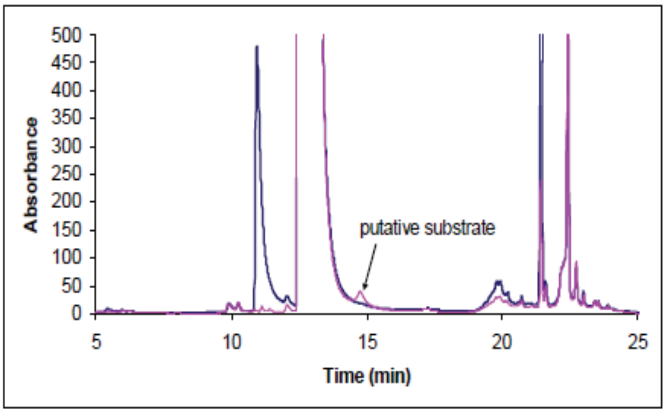


Supplemental Figure 1: Chromosomal clustering of *tenA* in *B. cereus*, B. *halodurans* and *B. subtilis*. The reactions catalyzed by ThiO, ThiS, ThiG, ThiF and ThiD are shown in Scheme 1. The functions of TenA, ThiY, and YlmB are identified in this paper and the function of TenI is still unknown.



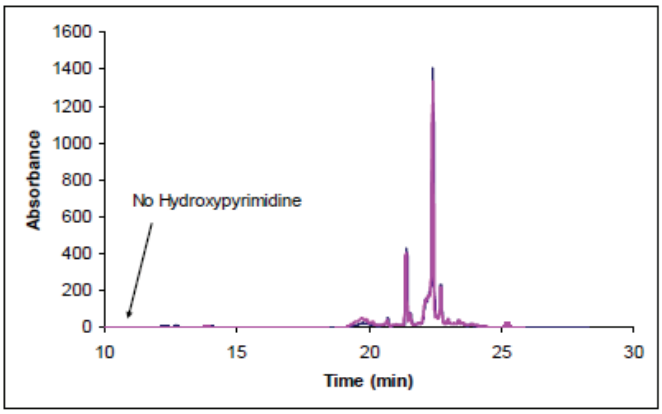
Supplementary Figure 2:

HPLC chromatogram of the thiamin degradation reaction in soil at pH 7. Thiamin in soil undergoes extensive degradation after 15 days (pink trace) while the thiamin that is not exposed to soil does not undergo the same degradation after 3 months (blue trace).



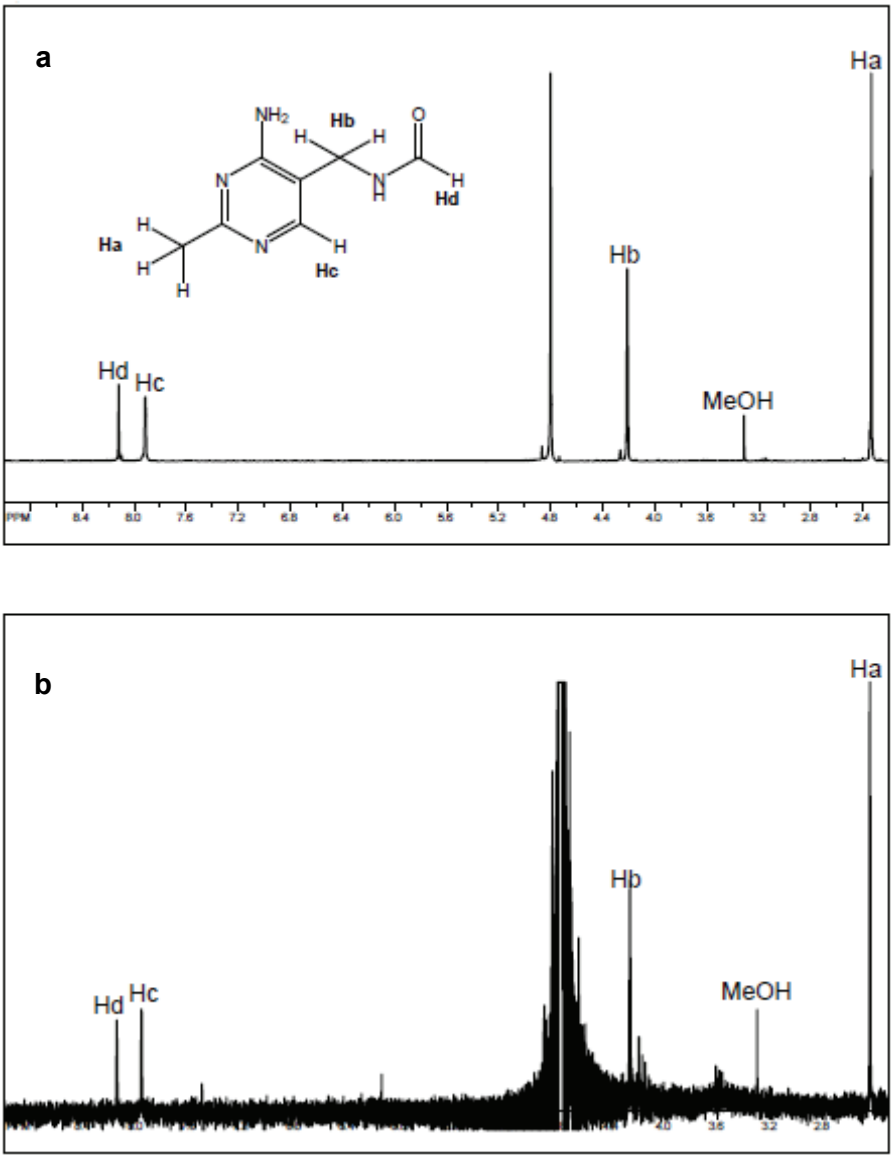
Supplementary Figure 3:

HPLC chromatogram of the reaction of soil-degraded thiamin with TenA and YlmB. Pink Trace: Soil-degraded thiamin. Blue trace: Soil-degraded thiamin treated with YlmB and TenA.



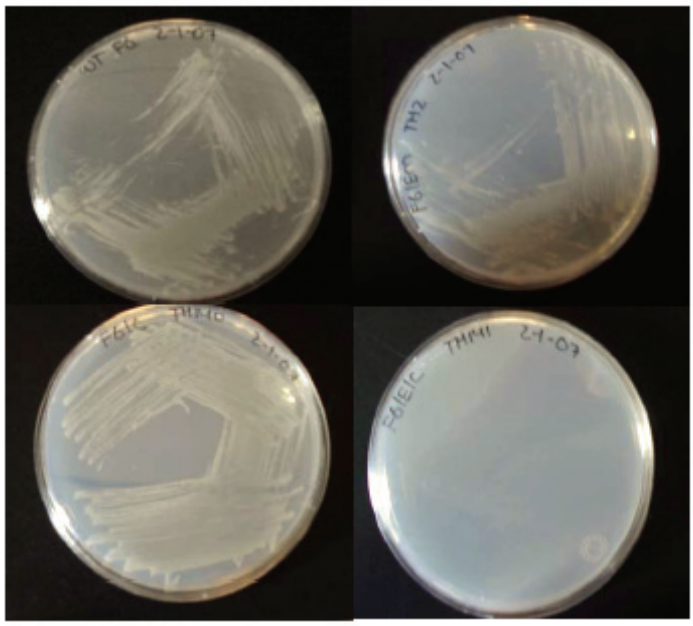
Supplementary Figure 4:

HPLC chromatogram indicating the absence of any substrates for TenA/YlmB in the complex mixture eluting between 18 and 25 minutes. Blue trace is the mixture in the absence of YlmB/TenA, pink trace is the mixture treated with YlmB and TenA



Supplemental Figure 5:

(a) ^{1}H-NMR of synthetic formyl-aminopyrimidine (**15**) and **(b)** the formyl-aminopyrimidine (**15**, X_{15}) obtained from thiamin degraded in soil.



Supplementary Figure 6:

Complementation studies using formylaminopyrimidine (**15**), top left: WT, top right : ThiA-, bottom left: TenA-, bottom left: ThiA-TenA-. Growth is seen on the WT, ThiA- and TenA- strains, while there is no growth on the ThiA-TenA- mutant.

Supplementary Table 1:

Strains and primers used in this work

Name	Genotype or Nucleotide sequence (5'>3')
1A747	Wild type (BGSC)
TH2	1A747 *thiA::Tn917* (this work)
TH12	Δ*thiL::cat4* (Schyns et al., 2005)
TH140	1A747 *tenA::cat4* (this work)
TH141	1A747 *thiA::Tn917 tenA::cat4* (this work)
P1tenA	CGTCATCTTTGTCAGGATCTTCAGAGG
P2tenA-cat	CCCACTTTATCCAATTTTCGGAACGAACGGATGGA CAAAGC
P3tenA-cat	TATGAGATAATGCCGACTGTACTTCCAGCTACTAT GAATATCAATTTTGGG
P4tenA	GATACTCCTCTGCCTTCAAGACC

References

Anderson, K.S., J.A. Sikorski, and K.A. Johnson. 1988. Evaluation of 5-enolpyruvoylshikimate-3-phosphate synthase substrate and inhibitor binding by stopped-flow and equilibrium fluorescence measurements. Biochemistry 27: 1604–1610.

Begley, T.P. 2006. Cofactor biosynthesis: an organic chemist's treasure trove. Nat. Prod. Rep. 23:15–25.

Begley, T.P., and S.E. Ealick. 2004. Mechanistic and structural studies on thiamine biosynthetic enzymes. Oxidative Stress and Disease. 11:15–28.

Chahine, E.H., and J.M. Dubois. 1983. Kinetics and thermodynamics of the structural transformations of thiamine in neutral and basic aqueous media. The UV spectrum of the tetrahedral pseudobase intermediate. J. Am. Chem. Soc. 105:2335–2340.

Day, N., and J.W. Keillor. 1999. A continuous spectrophotometric linked enzyme assay for transglutaminase activity. Anal. Biochem. 274: 141–144.

Dornow, A., and G. Petsch. 1953. Reductions with lithium aluminum hydride V. The preparation of vitamin B1. Chem. Ber. 86:1404–1407.

Dorrestein, P.C., H. Zhai, F.W. McLafferty, and T.P. Begley. 2004. The biosynthesis of the thiazole phosphate moiety of thiamin:

the sulfur transfer mediated by the sulfur carrier protein ThiS. Chem. Biol. 11:1373–1381.

Haas, A.L., N.P. Laun, and T.P. Begley. 2005. Thi20, a remarkable enzyme from *Saccharomyces cerevisiae* with dual thiamin biosynthetic and degradation activities. Bioorg. Chem. 33:338–344.

Hanes, J.W., S.E. Ealick, and T.P. Begley. 2007. Thiamin phosphate synthase: the rate of pyrimidine carbocation formation. J. Am. Chem. Soc. 129:4860–4861.

Horikoshi, K. 1999. Alkaliphiles: some applications of their products for biotechnology. Microbiol. Mol. Biol. Rev. 63:735–750.

Imamura, N., and H. Nakayama. 1981. *thiD* locus of *Escherichia coli*. Experientia. 37:1265–1266.

Imamura, N., and H. Nakayama. 1982. *thiK* and *thiL* loci of *Escherichia coli*. J. Bacteriol. 151:708–717.

Kurata, G.I., T. Sakai, and T. Miyahara. 1968. Antagonists of thiamine XVIII: reaction condition in the formation of desthiothiamine from alkaline thiamine solution with amino acids. Vitamin. 37:398–402.

Lawhorn, B.G., R.A. Mehl, and T.P. Begley. 2004. Biosynthesis of the thiamin pyrimidine: the reconstitution of a remarkable rearrangement reaction. Org. Biomol. Chem. 2:2538–2546.

Maier, G.D., and D.E. Metzler. 1957. Structures of thiamine in basic solution. J. Am. Chem. Soc. 79: 4386–4391.

Melnick, J., E. Lis, J. H. Park, C. Kinsland, H. Mori, T. Baba, J. Perkins, G. Schyns, O. Vassieva, A. Osterman, and T. P. Begley. 2004. Identification of the two missing bacterial genes involved in thiamine salvage: thiamine pyrophosphokinase and thiamine kinase. J Bacteriol.186:3660-3662.

Miranda-Rios, J. 2007. The THI-box riboswitch, or how RNA binds thiamin pyrophosphate. Structure. (15):259–265.

Mizote, T., and H. Nakayama. 1989. The *thiM* locus and its relation to phosphorylation of hydroxyethylthiazole in *Escherichia coli*. J. Bacteriol. 171: 3228–3232.

Pace, C.N., F. Vajdos, L. Fee, G. Grimsley, and T. Gray. 1995. How to measure and predict the molar absorption coefficient of a protein. Protein Sci. 4: 2411–2423.

Park, J.H., K. Burns, K., C. Kinsland, and T.P. Begley. 2004. Characterization of two kinases involved in thiamine pyrophosphate and pyridoxal phosphate biosynthesis in *Bacillus subtilis*: 4-amino-5-hydroxymethyl-2-methylpyrimidine kinase and pyridoxal kinase. J. Bacteriol. 186:1571–1573.

Quayle, J.R. 1966. Formate dehydrogenase. Methods Enzymol. 9:360–364.

Rodionov, D.A., A.G. Vitreschak, A.A. Mironov, and M.S. Gelfand. 2002. Comparative genomics of thiamin biosynthesis in procaryotes. J. Biol. Chem. 277:48949–48959.

Schyns, G., S. Potot, Y. Geng, T.M. Barbosa, A.O. Henriques, and J.B. Perkins. 2005. Isolation and characterization of new thiamine-deregulated mutants of *Bacillus subtilis*. J. Bacteriol. 187(23):8127-36.

Takami, H., and K. Horikoshi. 1999. Reidentification of facultatively alkaliphilic *Bacillus* sp. C-125 to *Bacillus halodurans*. Biosci. Biotechnol. Biochem. (63): 943–945.

Toms, A.V., A.L. Haas, J.H. Park, T.P. Begley, and S.E. Ealick. 2005. Structural characterization of the regulatory proteins TenA and TenI from *Bacillus subtilis* and identification of TenA as a thiaminase II. Biochemistry. 44: 2319–2329.

Wach, A. 1996. PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae.* Yeast. 12:259–265.

Webb, E., and D.Downs. 1997. Characterization of *thiL*, encoding thiamin-monophosphate kinase, in *Salmonella typhimurium*. J. Biol. Chem. 272: 15702–15707.

Winkler, W., A. Nahvi, and R.R. Breaker. 2002. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature. 419: 952–956.

Xu, J.C., J.W. Stucki, J. Wu, J.E. Kostka, and G.K. Sims. 2001. Fate of atrazine and alachlor in redox-treated ferruginous smectite. Environ. Toxicol. Chem. 20:2717–2724.

CHAPTER III

Isolation and characterization of new thiamin-deregulated mutants of *Bacillus subtilis*

Ghislain Schyns[Ŧ], Sébastien Potot[Ŧ], Yi Geng[Ŧ*],

Teresa M. Barbosa[¤], Adriano O. Henriques[¤], and John B. Perkins[Ŧ]

Journal of Bacteriology. 2005. 7(23):8127-36.

[Ŧ] DSM Nutritional Products, Biotechnology R&D, PO box 3255, Bldg 203/25, CH-4002 Basel, Switzerland.

[¤] Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Apartado 127, 2781-901 Oeiras, Portugal.

* Current address: Wyeth Pharmaceuticals, Pearl River, NY, USA.

Contribution:

Sébastien Potot made the mapping of EMS-induced mutations, the strain construction of *B. subtilis* mutants (except EMS-induced thiamin-deregulated mutants), and their *in vivo* characterization for thiamin production.

Abstract

In bacteria, thiamin pyrophosphate (TPP) is an essential cofactor that is synthesized de novo. Thiamin, however, is not an intermediate in the biosynthetic pathway but is salvaged from the environment and phosphorylated to TPP. We have isolated and characterized new mutants of *Bacillus subtilis* that deregulate thiamin biosynthesis and affect export of thiamin products from the cell. Deletion of the *ydiA* gene, which shows significant similarity to the thiamin monophosphate kinase gene of *E. coli* (*thiL*), did not generate the expected thiamin auxotroph, but instead generated a thiamin bradytroph that grew to nearly wild-type levels on minimal medium. From this ΔthiL deletion mutant, two additional EMS-induced mutants were isolated that de-repressed expression of a *thiC-lacZ* transcriptional reporter. One mutant, Tx1, contained a nonsense mutation within the *B. subtilis yloS* (*thiN*) gene that encodes a thiamin pyrophosphokinase, a result which confirmed that *B. subtilis* contains a single-step, yeast-like thiamin-to-TPP pathway in addition to the bacterial TPP de novo pathway. A second mutant was shown to contain two lesions, *tx26-1* and *tx26-2*. Genetic mapping and DNA sequencing indicated that the *tx26-1* mutation is an allele of *yuaJ*, which encodes a thiamin permease. The second mutation, *tx26-2*, was located within the *ykoD* cistron of the *ykoFEDC* operon, which putatively encodes the ATPase component of a unique thiamin-related ABC transporter. Genetic and microarray studies indicated that both mutant *yuaJ* and *ykoD* genes were required for derepression of thiamin-regulated genes. Moreover combination of the four mutations (Δ*thiL*, *thiN*, *yuaJ*, and *ykoD*) into a single strain significantly increased production and excretion of thiamin products into the culture medium. These results

are consistent with the proposed “Riboswitch” mechanism of thiamin gene regulation (Winkler *et al*., 2002, Nature 419:952-956).

Introduction

Vitamin B1 (thiamin pyrophosphate) is a cofactor of a number of important enzymes in carbohydrate and amino acid metabolisms, which can be synthesized by microorganisms, plants and fungi but not by mammals [Schowen, 1998]. Unlike other vitamin biosynthetic pathways in bacteria (*e.g*. riboflavin and biotin), thiamin is not part of the de novo pathway, but is actually part of the salvage pathway. In *B. subtilis*, thiamin pyrophosphate biosynthesis occurs by a complex multistep pathway [Begley *et al*., 1999]. The pyrimidine moiety, 4-amino-2-methyl-5-hydroxymethylpyrimidine pyrophosphate (HMP-PP), is derived from an intermediate in the de novo purine biosynthetic pathway, 5-aminoimidazole ribotide (AIR), in a conversion catalyzed by the *thiC* gene product [Begley *et al*., 1999]. HMP-P is then phosphorylated to HMP-PP by the product of the *thiD* (*yjbV*) gene prior to coupling with the thiazole unit [Park *et al.*, 2004]. The thiazole moiety, 5-(2-hydroxyethyl)-4-methylthiazole phosphate (HET-P), is derived from 1-deoxy-D-xylulose 5-phophate (DXP), glycine and cysteine in a complex oxidative condensation reaction requiring the products of at least five different genes, *thiF*, *thiS*, *,thiO*, *thiG* and a *nifS*-like gene [Park *et al.*, 2003]. Coupling of HMP-PP and HET-P is catalyzed by thiamin phosphate pyrophophorylase encoded by *thiE*, resulting in thiamin monophophate (TMP) [Begley *et al*., 1999]. TMP is then phosphorylated to form thiamin pyrophosphate (TPP) by the action of thiamin monophosphate kinase, encoded by *thiL*. *B. subtilis* contains another potential way to produce TPP, through a thiamin one-step salvage pathway formerly discovered in *Saccharomyces cerevisiae* [Nygaard and Saxild, 2005]. This pathway, initially thought to exist only in lower eukaryotes, is catalyzed by the

product of the *thiN* (*yloS*) gene [Melnick *et al.*, 2004]. Two other salvage kinases, one specific for hydroxyethylthiazole and the other with a broad range of substrates specificity including vitamin B6 compounds and HMP, have been identified in *B. subtilis* and are encoded by the genes *thiM* and *pdxK* (*ywdB*), respectively [Park *et al*., 2004]. However an ortholog to the *E. coli* gene *thiK* (*ycfN*) encoding thiamin kinase, a salvage enzyme that converts thiamin to TMP, has not been identified in *B. subtilis* [Melnick *et al.*, 2004]. The *B. subtilis* thiamin biosynthetic pathway showing all known and proposed genes and intermediates is diagramed in Figure 1 (by recent convention, *B. subtilis thi* gene nomenclature has been changed to the *E. coli* nomenclature; some literature cited here may contain the original *B. subtilis thi* gene designations).

Interestingly, a thiamin regulatory gene has not been identified so far in bacteria [Downs and Petersen, 1994; Lawhorn[a] *et al*., 2004; Webb *et al*., 1996]. Instead, it has been proposed that thiamin pyrophosphate directly regulates expression of the thiamin biosynthetic genes by a novel mechanism called Riboswitch, in which TPP interacts with the nascent mRNA message at a *cis*-acting site within the 5'-leader region (call the *thi*-box) to form a secondary structure that allow formation of a transcription terminator [Mandal *et al*., 2003; Miranda-Rios *et al*., 2001; Mironov *et al*., 2002; Winkler *et al*., 2002]. In addition to TPP, which was shown by transcription profiling to regulate expression of the *thiC* gene and the *tenA-tenI-thiOSGFD* operon [Lee *et al*., 2001], thiazole (HET) was also described to repress the *ywbI-thiME* operon in *B. subtilis* [Zhang and Begley, 1997; Zhang *et al*., 1997].

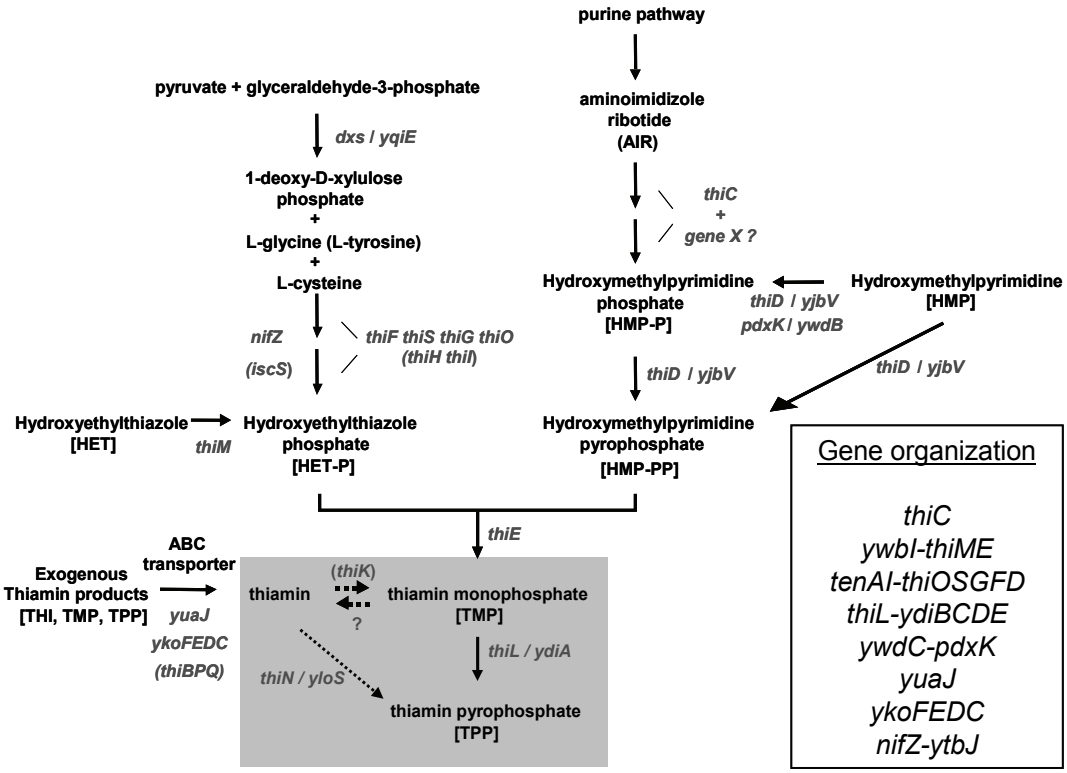


Figure 1: Thiamin biosynthesis and salvage pathway of *Bacillus subtilis*. *B. subtilis thi* genes were previously identified by genetic, biochemical, and/or *in silico* evidence [Begley *et al*., 1999; Melnick *et al*., 2004; Park *et al*., 2003; Park *et al*., 2004; Rodinov *et al*., 2002]; in some cases, their original y gene designation is listed. *E. coli* intermediates or genes not present in *B. subtilis* but with similar functions are given in parentheses [Begley *et al*., 1999]. Abbreviations are in brackets. The thiamin-TMP-TPP salvage pathway is shown in a shaded box. See the text for details. Question marks indicate possible unknown genes involved in the thiamin biosynthesis and salvage pathway [Lawhorn[a] *et al*., 2004; Lawhorn[b] *et al*., 2004; Melnick *et al*., 2004]. The organization of *Bacillus* thiamin biosynthesis genes is shown on the right.

In silico screening of the *B. subtilis* genome with the *thi*-box sequence has resulted in the identification of several new putative thiamin-regulated genes [Rodinov *et al*., 2002]: the *yuaJ* gene and the *ykoFykoE-ykoD-ykoC* operon. *yuaJ* could encode an ATP-independent thiamin transporter containing multiple trans-membrane segments. YkoE and YkoC are likely two trans-

membrane components while YkoD is an ATPase component and YkoF acts as a novel HMP/ thiamin-binding protein [Devedjiev *et al*., 2004]. However, no genetic experiments have confirmed the role of these genes in thiamin transport, and their products show no similarity to the known thiamin ABC transporter genes of *Salmonella typhimurium* (*thiBPQ*) [Webb *et al*., 1998].

In this study, we used mutagenesis and transcriptomics to identify and characterize mutants of *B. subtilis* that deregulate thiamin gene expression. As predicted by the Riboswitch model of gene regulation, none of these mutations were located in genes predicted to encode a regulatory protein. Instead, these mutations mapped to several genes previously identified as encoding a thiamin pyrophosphokinase activity (*thiN*), or encoding thiamin-related transport proteins (*ykoD* and *yuaJ*). Moreover, we also show that these mutations must be combined to deregulate thiamin gene expression, resulting in bacterial strains that are altered in the production and export of thiamin. These results provide genetic support that intracellular TPP levels control thiamin biosynthetic gene expression.

Materials and Methods

Bacterial strains, plasmids, and growth conditions.

Plasmids and strains used in this study are listed in Table 1. Antibiotic resistance genes that confer resistance to neomycin (*neo*), tetracycline (*tet*), and chloramphenicol (*cat*), were obtained from plasmid pBEST501, pDG1514, and pC194, respectively (The Bacillus Genetic Stock Center, The Ohio State University [BGSC]). *B. subtilis* was grown in minimal medium (MM: 1X Spizizen salts, 0.04% sodium glutamate, and 0.5% glucose) or Veal Infusion-Yeast Extract complete medium (VY), or grown on agar plates consisting of Tryptose Blood Agar Base (TBAB, Difco, Maryland, USA) or MM medium. *E. coli* was grown on MM or Luria-Bertani medium without glucose. The compositions of these media are described elsewhere [Harwood and Archibald, 1990]. For testing thiamin production in liquid test tubes cultures, a thiamin-free medium was used [Difco, 1998]. Thiamin supplements (thiamin, TMP, or TPP) required for auxotrophy were added at 1 μM. Resistance to the thiamin anti-metabolite pyrithiamin was tested at 0.1 or 10 μM.

Thiamin biological assays.

Total thiamin compounds were assayed using indicator strains derived from *Salmonella typhimurium* using known methods [Difco, 1998]. Strain DM456 (Ω*thiD906::mudJ*) responds to thiamin, TMP and TPP in minimal medium, whereas strain DM1856 (Ω*thiL934::*Tn*10d*) responds to only TPP [Petersen and Downs, 1997; Webb *et al*., 1998].

Table 1: List of plasmids and strains.

Plasmids	Relevant characteristic or genotype	Reference
pBR322	*E. coli* plasmid cloning vector	Bolivar *et al.,* 1977
pMUTIN2	*B. subtilis* integration plasmid for gene inactivation	Vagner *et al.,* 1998
pDG1728	*B. subtilis* plasmid for ectopic integration of *lacZ* transcriptional fusions at *amyE*	Guérout-Fleury *et al*., 1996
pEpUCΔ1	*E. coli-B. subtilis* shuttle vector to generate in-frame mutations	S. Seror (pers.commu)
pTH1	Internal 348-bp fragment of *thiL* (*ydiA*) cloned into pMUTIN2	This work
pTH12	712-bp *thiC* promoter fragment cloned into pDG1728	This work
pTH30	In-frame Δ*thiL* (*ydiA*) deletion (G_{79}-G_{202}) cloned into pEpUCΔ1	This work
pTH31	Internal 353-bp *yuaJ* fragment cloned into pMUTIN2	This work
Strains		
B. subtilis		
PY79	SPβ^c (cured) prototroph	BGSC (1A747)
TH12	SPβ^c Δ*thiL*::*cat*$_4$	This work
TH21	SPβ^c Ω*amyE*::*thiC-lacZ*	This work
TH22	SPβ^c Δ*thiL*::*cat*$_4$ Ω*amyE*::*thiC-lacZ*	This work
TH48	Δ*thiL*::*cat*$_4$ *yuaJ* (the 1st lesion in Tx26) *ykoD* (the 2^d lesion in Tx26)	This work
TH76	Ω*yuaJ*::pMUTIN2	This work
TH83	Δ*thiL yuaJ ykoD*	This work
TH95	Δ*thiL thiN* (*tx1*) *yuaJ ykoD* Ω*yloA*::Tn*917*	This work
TH105	Δ*thiL*::*cat*$_4$	This work
TH106	Δ*thiL*::*cat*$_4$ Ω*thiL*::pMUTIN2	This work
TH112	Δ*thiL*::*cat*$_4$ Ω*amyE*::*thiC-lacZ ykoD* Ω*yufR*::Tn*917*	This work
TH128	Δ*ispA*::*neo*	This work
TH137	Ω*yuaJ*::pMUTIN2 Δ*ykoD*::*cat*$_4$	This work
TH138	Δ*ykoD*::*cat*$_4$	This work
1A603	(SPβc2*) trpC2* Ω*thiC84*::Tn*917*	Vandeyar and Zahler, 1986
1A631	(SPβc2) *trpC2* Ω*motA*::Tn*917*	
1A633	(SPβc2) *trpC2* Ω*yloA*::Tn*917*	
1A642	(SPβc2) *trpC2* Ω*yufR*::Tn*917*	
Tx1	Δ*thiL*::*cat*$_4$ Ω*amyE*::*thiC-lacZ thiN* (*tx1*)	This work
Tx26	Δ*thiL*::*cat*$_4$ Ω*amyE*::*thiC-lacZ yuaJ ykoD*	This work
RL2066	Δ*spoVM*::*spec*	Chiu *et al.,* 1998
E. coli		
PT-R1	*thiO-35*	*E. coli* Genetic Stock Center
K12	Wild type	*E. coli* Genetic Stock Center
S.enterica serovar Typhimurium		
DM456	Ω*thiD906*::MudJ(Km)	Webb and Downs, 1997
DM1856	Ω*thiL934*::Tn*10d*(Tc)	Peterson and Downs, 1997

The response of DM456 to known amounts of thiamin, TMP, and TPP ranging from 0.0256 to 100 µg/ liter, was similar. In addition, DM456 was found to be more sensitive to TPP than DM1856. To assay *B. subtilis* cultures, supernatants were filter-

sterilized before preparation of dilutions. Intracellular thiamin levels were measured from dilutions of 1 ml, filter-sterilized cellular extracts that were obtained by French press-breaking of the cells and centrifugation at 10,000g for 10 min. Indicator strains were grown overnight at 37°C in thiamin assay medium (TAM). Turbidity readings were made at 600 nm and compared to a range of standard solutions.

Thiamin thiochrome/ HPLC assay.

Individual thiamin compounds, thiamin, TMP, and TPP were measured using a modified thiochrome-HPLC assay procedure described previously [Chiu *et al*., 1999]. Briefly, 10 µl of culture supernatant or intracellular extracts are added to 200 µl of 4M potassium acetate. The sample is then oxidized by the addition of 100 µl fresh 3.8 mM potassium ferricyanide in 7 M NaOH. The mixture is vigorously mixed and then quenched by addition of 100 µl fresh 0.06% H_2O_2 in saturated KH_2PO_4. Samples are transferred to HPLC vials and injected onto a Supelcosil LC-18-T column (15cm X 4.6 mm, 3 µm) (Supelco – Ref. No 58970- U). Elution is made by a 10% - 35% methanol (H_2O 50% - 25%) gradient in the presence of 40% 0.1 M K_2HPO_4 (pH 6.6) and 4 mM tetrabutyl ammonium hydrogen sulfate. Fluorescence is measured at 444 nm after excitation at 365 nm. The chronological order of elution from the column is thiamin, TMP, and TPP. This procedure was utilized to monitor both internal and external thiamin production during fermentation. In our HPLC/DAD assay used to directly measure thiamin and the intermediates HMP and HET in the fermentation broth, chromatography of samples was performed on a Phenomenex LUNA C_{18} column, using an Agilent 1100 HPLC system equipped with a thermo-stated auto-sampler and a diode

array detector (DAD). The column dimensions are 150 x 4.6 mm, particle size 5 micron. The column temperature was kept constant at 20°C. The mobile phase is a mixture of 0.4 g pentane sulfonate in water, pH 2 (A) and methanol (B). Gradient elution is applied, ranging from 2% A (3 min) to 20% A in 20 minutes. The flow rate is 1 ml/min. The detection method is UV absorption at 254 nm. The selectivity of the method was verified by injecting 10 µl standard solutions of the relevant reference compounds, thiamin, HMP, and HET, each at 100 µg/ml. The target compounds were completely separated without special sample preparation.

Molecular and genetic techniques.

Standard genetic and molecular biology techniques used in this study have been previously described by Maniatis *et al*., 1982. DNA transformation, PBS1 generalized transduction, and other standard *B. subtilis* genetic techniques have been described previously [Cutting and Vander Horn, 1990].

Construction of *thiL* (*ydiA*) mutant.

Comparison of the *E. coli* ThiL protein sequence to the protein database of Subtilist database detected significant similarity to only one protein sequence, YdiA (P(N)= 8.1e-15). The gene encoding this protein, *ydiA*, is 975 base pairs in length and is the first gene of a five-gene operon located at 55° on the *B. subtilis* chromosome. *thiL* deletion mutant was generated by inserting a chloramphenicol acetyltransferase cassette lacking (*cat*$_4$) the endogenous *rho*-independent transcription termination site between nucleotide 267 and 272 of *thiL*. Removal of the terminator in *cat*$_4$ results in non-polar insertional mutations. To do this, PCR primer pair #1 - #2 (Table 2) was used to first generate DNA cassette *cat*$_4$

gene. This fragment was then individually ligated to two DNA fragments containing either the 3' or 5' ends of *thiL*, which were also generated by PCR using primers #3 – #4 and #5 - #6 (Table 2). The PCR fragment containing the Δ*thiL*::*cat*$_4$ cassette was then inserted directly into the chromosomal *thiL* gene of strain PY79 by DNA transformation, selecting for colonies resistant to 5 µg/ml chloramphenicol. Strain TH12 containing Δ*thiL*::*cat*$_4$ was recovered.

Table 2: List of PCR primers.

Primer No.	Sequence
1……	5'-GGGGGTACCGAAAATTGGATAAAGTGGG-3'
2……	5'-GGGACGCGTAAGTACAGTCGGCATTATCTCATA-3'
3……	5'-GGGGAATTCTACCAGTTGTTCTGCCAAGGGCAT-3'
4……	5'-GGGGGTACCGCAAGTGATACCAGATAAAACTTAGG-3'
5……	5'-GGGACGCGTCTTCAAAATGGACTGAATCTGAAA-3'
6……	5'-GGGGGATCCTTCAACGAGACAGACACCTTGTCCG-3'
7……	5'-ATGCGGATCCCGTCCGGACCGCC-3'
8……	5'-CGATCCCGGGGCCTCCCATCGCGGC-3'
9……	5'-ATGCCCCGGGATTTGCCTAAGCTTCATCCTAAC-3'
10……	5'-CGATGAATTCAGCCCTTCTGCAAAACCTT-3'
11……	5'-AGCTAAGCTTGGCAGCCGTTATTTTAGACATTG-3'
12……	5'-TGCAGGATCCATAAAAACTGCGCTGACCACTGA-3'

Plasmid pEpUCΔ1 (S. Seror, personal communication) was used to construct in-frame, marker-less deletion mutation in *B. subtilis*. This *E. coli* vector contains a selectable erythromycin-resistance (erm) cassette and the pE194 temperature-sensitive origin of replication that does not function over 51°C. To do this, a DNA fragment containing an in-frame gene deletion was prepared using standard PCR methods and inserted between the BamHI and EcoRI sites of the pEpUCΔ1. Competent cells of the recipient strain were transformed at 51°C with the in-frame containing pEpUCΔ1 plasmid, selecting for erythromycin resistance at 5 µg/ml erythromycin. Em^r colonies that are also Cm^r were recovered and grown overnight at 28°C in the absence of antibiotic selection for 72 hours. Bacteria were then plated onto TBAB agar plates and the

plates incubated overnight at 37°C. Approximately 25% of the colonies were found to be sensitive to both erythromycin and chloramphenicol. PCR analysis of chromosomal DNA from several Em^S Cm^S colonies was used to confirm the presence of the in-frame mutation.

Isolation of EMS-induced thiamin deregulated mutants.

The strategy to isolate thiamin deregulated mutants of *B. subtilis* was to mutagenize bacteria that contained a *thiC-lacZ* fusion and then screen for colonies that were Lac+ in the presence of TPP or thiamin. A 732 bp-long DNA fragment containing 417 bp of the 5' promoter region of *thiC* was prepared by PCR using standard methods and cloned unidirectionally in front of the promoter-less *lacZ* gene of the pDG1728 vector [Guérout-Fleury *et al.*, 1996], resulting in plasmid pTH12. This vector is designed to introduce ectopic transcriptional *lacZ* fusions into the non-essential *amyE* locus of *B. subtilis*. Plasmid pTH12 was linearized by restriction enzyme digestion and transformed into *B. subtilis* PY79, selecting for colonies that were resistant to 100 µg/ml spectinomycin (TH21). The same transcriptional fusion was also introduced into TH12 to generate TH22 (ΔthiL::*cat*$_4$, Ω*amyE::thiC-lacZ*). Based on these results, *thiC-lacZ* reporter strain TH22 (Δ*thiL::cat*$_4$) was used to screen for deregulated mutants under repressing growth conditions. Two methods were used to isolate such mutants. In the first method, MM agar plates were prepared that contained 1 µM thiamin and 25 µg/ ml XGAL. After applying a uniform dilution of logarithmic growth phase TH22 cells, a paper disk containing 3 drops of ethylmethane sulfonate (EMS, *d* = 1.17 g/ml solution) was placed in the center of the plate. Lac+ colonies appeared over a period of 7 days incubation at 37°C. In the second method, banks of EMS-

mutagenized cells were prepared and screened. Accordingly, logarithmic stage TH22 cells were treated with 9.4 mM EMS for 90 minutes and aliquots frozen in 10% glycerol at -90°C. Cells from the frozen stocks were diluted in VY medium, incubated at room temperature for 30 minutes and then plated onto MM medium containing 1 µM thiamin and 25 µg/ ml XGAL. Screening of Lac+ colonies led to mutants Tx11 and Tx26.

Combination of mutations Δ*thiL*, *tx1* and *tx26*.

As a first step, the mutations in Tx26 were transferred into TH12 (Δ*thiL*::*cat*$_4$) by DNA transformation followed by selection for resistance to 0.1 µM pyrithiamin (P-0256, Sigma, Saint-Louis, MO). One pyrithiamin-resistant (Pyr^r) colony that was also Lac+ in the presence of TPP was recovered and designated TH48. Each strain was grown in minimal medium supplemented with micronutrients and 2.5% Difco nutrient broth (NB) for 18 hours at 37°C. The *cat*-interrupted *thiL* gene was next replaced by an in-frame deletion using pEpUCΔ1 (method described above). Using standard PCR methods, an in-frame deletion of *thiL* (removing amino acid residues Gly$_{79}$ through Gly$_{202}$) was first constructed using primers #7 – #8 and #9 – #10 (Table 2) and inserted between the BamHI and EcoRI sites of pEpUCΔ1 creating pTH30. This plasmid was then used to replace the Δ*thiL::cat*$_4$ mutation in TH48 with the in-frame *thiL* mutation (Δ*thiL*), resulting in strain TH83. The *tx1* (*thiN*) mutation was next introduced by transduction into TH83 by PBS1 transduction by standard procedures using linkage to a silent Tn*917* insertion, Ω*yloA*::Tn*917* (strain 1A633 of the Bacillus Genetic Stock Center, also called CU4153 or Ω*zdi-82*::Tn*917*). The resulting strain was called TH95.

Mapping experiments and constructions of mapping strains.

Genetic mapping studies using PBS1 generalized transduction were performed under standard conditions [Cutting and Vander Horn, 1990] using donor strains from the collection of phenotypically-silent Tn*917* mapping insertions [Vandeyar and Zahler, 1986] and strains listed in Table 1. In order to inactivate the *yuaJ* gene, a 324 bp internal fragment of *yuaJ* starting at position 353 was PCR-amplified and inserted between the BamHI and HindIII sites of pMUTIN2 using primers #14 and #15 (Table 2), creating plasmid pTH31. To determine if *tx1* is allelic to *thiN* (*yloS*), a 448 bp deletion mutation, starting at base 124 of the *thiN* gene (Δ*thiN*::*cat*$_4$) was constructed using the *cat*$_4$ gene and cloned into pBR322 [Bolivar et al;, 1977]. Long-flanking homology PCR was used to construct strains PY79 derivatives TH128 and TH138, which contains *cat*$_4$-mediated deletions of the *ispA* and *ykoD* genes, respectively [Wach, 1996].

Microarray profiling.

TH22, Tx1, Tx26, and TH48 were grown in shake-flask cultures containing 50 ml Spizizen minimal medium with added thiamin pyrophosphate (0.34 µg/ml). Overnight cultures were diluted to Klett = 10 units into fresh medium and grown to exponential growth phase (Klett = 100 units). Cells from half of the culture were collected by centrifugation, and the total RNA was immediately extracted as previously described [Lee *et al*., 2001]. The remaining culture was grown to early stationary phase before RNA extraction. Early stationary phase was judged to be 30 min after glucose exhaustion; glucose content in the medium was measured by a glucose analyzer (Beckman, Fullerton, CA) using standard procedures. Preparation of RNA, labeled cDNA targets, microarray

hybridization and staining procedures, and data analysis are described in Lee et *al.*, 2001.

Results

Bacterial thiamin levels.

As a first step in isolating *B. subtilis* thiamin deregulatory mutants, the intra- and extra-cellular levels of thiamin products from wild type and engineered *B. subtilis* strains were determined as described in Material and Methods. As a control, the thiamin deregulated *E. coli* PT-R1 mutant was similarly tested; this mutant is reported to produce only intracellular thiamin (no excretion of thiamin products into the medium) mostly in the form of TPP (>90%). Bioassay results indicated that thiamin products were readily detected from extracts of sonicated cells, but not from the culture medium (data not shown). The intracellular level of thiamin products in logarithmic or stationary phase wild-type *B. subtilis* was calculated to be approximately 100-200 µg/L (Table 3). This thiamin level was lower than for *E. coli* PTR1, which reached approximately 1.6 mg/L in stationary phase cells. Determination of the individual thiamin forms using Thiochrome/HPLC showed only trace amounts of TMP and TPP in the supernatant of PY79 cultures (<0.05 µg/ml). However, intracellular thiamin products were 80% TPP and 20% TMP, in good agreement with the bioassay results. Intra- or extracellular levels of free thiamin were not observed.

Table 3: Thiamin production and expression of *thiA-lacZ* transcriptional fusions in wild-type and thiamin-deregulated mutants of *B. subtilis.*

Strain	Genotype	Accumulation (μg/l)		Regulation (β-galactosidase specific activity [Miller units][c])					
				Log			Stationary		
		Extracellular [a]	Intracellular [b]	None	HMP [d]	TPP [d]	None	HMP	TPP
TH21	*thiC*$^+$ Ω*amyE*::*thiC-lacZ*	< 0.1	190	110	6	0.5	31	3	1.2
TH22	*thiC*$^+$ Ω*amyE*::*thiC-lacZ* Δ*thiL*::*cat*$_4$	0.9	530	140	26	1.3	75	9	1.8
Tx1	*thiC*$^+$ Ω*amyE*::*thiC-lacZ* Δ*thiL*::*cat*$_4$ *thiN*	5	500	NG	135	8	NG	177	5
Tx26	*thiC*$^+$ Ω*amyE*::*thiC-lacZ* Δ*thiL*::*cat*$_4$ *yuaJ ykoD*	15	500	280	35	130	120	20	36

[a] thiamin concentration in MM (30 ml) after 24 h of growth at 37°C.
[b] thiamin concentration of 1-ml supernatant of sonicated cells collected from a 30-ml MM culture after 24 h of growth at 37°C.
[c] calculated according to Miller, 1972. NG, no growth.
[d] final concentration of TPP and HMP were 1μM.

Construction and isolation of thiamin-deregulated mutants.

The strategy to isolate and characterize thiamin deregulated mutants was based on monitoring the activity of a *thiC-lacZ* transcriptional fusion in the presence of TPP or thiamin. As described in Material and Methods, TH21 (Ω*amyE*::*thiC-lacZ*) is a derivative of PY79 in which a transcription *thiC-lacZ* fusion was integrated in single copy into the *amyE* gene. In control studies, TH21 displayed thiamin-regulated expression of the *lacZ* fusion. On minimal medium agar plates containing the indicator X-gal, colonies of TH21 were Lac+ (*i.e.* blue), but were Lac- if the medium was supplemented with thiamin or TPP (1μM). Moreover when grown to early logarithmic phase (OD_{600} = 0.8-0.9) in shake-flask cultures containing 1 μM TPP, β-galactosidase activity was repressed

approximately 200-fold compared to cells grown in minimal medium without TPP (Table 3). In stationary growth phase, the difference caused by TPP repression was less pronounced (30-fold repression). HMP (1 µM) also repressed expression of the fusion, but to a lesser extend. Both thiazole and adenosine (1 µM each) did not show any thiamin repressing activity (data not shown).

The *thiC-lacZ* transcriptional reporter fusion was introduced into a ∆*thiL::cat*$_4$ mutant (TH22) to screen for Lac+ (*i.e.* blue) colonies under thiamin or TPP repressing growth conditions.

In *E. coli*, inactivation of the *thiL* gene produces a thiamin deregulated strain which is a strict thiamin auxotroph. However, the *B. subtilis* ∆*thiL* mutant was not a thiamin or TPP auxotroph, but instead it appeared to be a thiamin bradytroph (data not shown) and was slightly deregulated for thiamin gene expression (Table 3). We choose this strain to generate thiamin-deregulated mutants in order to exclude *thiL* mutants from our screen. Out of 50,000 EMS-mutagenized colonies screened, twenty-six Lac+ mutants were recovered. Amongst these, two (Tx1 and Tx26) were selected for based on their unique phenotypes. Tx1 appeared to be a thiamin bradytroph based on its slow growth profile on minimal medium without TPP. Conversely, Tx26 was a prototroph but conveyed a much strong Lac+ phenotype than Tx1. In addition, Tx26 was found to be resistant to up to 10 µM pyrithiamin (Pyr^r), a toxic thiamin analog.

Thiamin-deregulated mutations are allelic to genes that are involved in thiamin salvage and transport.

PBS1 generalized transduction and DNA transformation were used to determine the genetic map position of *tx1* and *tx26*. In most cases, the Tn*917* mapping strains were used as donors; since

these mapping strains carry the wild-type alleles we reasoned that replacement of the *tx1* or *tx26* mutation with the wild-type allele would result in a concomitant change in Lac phenotype (*i.e.* Lac+ to Lac-). In addition, conversion of Pyr^r to Pyr^s could also be used to map the lesion(s) in the Tx26 mutant.

Initial mapping studies showed that the *tx1* mutation was not linked to either the Δ*thiL*::*cat*$_4$ mutation or the Ω*amyE::thiC-lacZ* fusion. However, the Ω*spoVM*::Tn*917* mutation (strain RL2066), showed the highest transduction (> 99 %) and transformation linkages (72 %) to *tx1*, indicating that this lesion was within or adjacent to *spoVM*. Inspection of this region located a possible candidate, *thiN* (*yloS*), a gene adjacent to *spoVM* which has been previously shown to have thiamin pyrophosphokinase activity (Thiamin-to-TPP activity; [Melnick *et al*., 2004]). To determine if *tx1* is allelic to *thiN* (*yloS*), the DNA sequence of *thiN* in the Tx1 mutant was determined. Results revealed a single base mutation that caused a leucine-to-phenylalanine substitution at amino acid residue 116 (Leu_{116} [CTT] >Phe_{116} [TTT]).

Initial DNA transformation studies of Tx26 indicated that the transformation frequency of the pyrithiamin-resistance marker to sensitive bacteria was non-linear with respect to DNA concentration (data not shown). This result suggested that Pyr^r in Tx26 was caused by more than one mutation. Transduction experiments confirmed that the Tx26 mutant contained two mutations located at different regions of the chromosome. One mutation, designated *tx26-1*, showed 70% linkage to Ω*yufR*::Tn*917*, and the other mutation, designated *tx26-2*, showed 59% linkage to Ω*motA*::Tn*917*. Further transformation mapping studies showed strong linkage of *tx26-1* to *yuaJ*, a thiamin-regulated gene proposed to encode a thiamin permease [Rodinov *et al.*, 2002]. To determine

if this lesion is allelic to *yuaJ*, a knock-out mutation of *yuaJ* was first constructed using pMUTIN. As expected, introduction of Ω*yuaJ*::pMUTIN into wild type strains, (*e.g.* PY79) or *thiC* mutants were without phenotype. However introduction of Ω*yuaJ*::pMUTIN into strain TH112 (*tx26-1* + *tx26-2* Δ*thiL*) resulted in Em^r colonies that were resistant to 0.1 µM pyrithiamin. Finally, DNA sequence analysis of *yuaJ* from Tx26 confirmed that *tx26-1* was an allele of *yuaJ*. YuaJ is a predicted membrane-bound thiamin permease. Sequencing detected a single base mutation that resulted in the change of a glutamine residue at amino acid position to an ocher stop codon (Gln_{35} [CAA] > Stop [TAA]). Introduction of the *tx26-1* mutation is predicted to prematurely terminate translation resulting in an inactive, 35 amino acid truncated protein. In similar mapping studies, the *tx26-2* allele showed strong linkage (65%) to the *ispA* locus (117°). This genetic map position corresponded to a cluster of genes, *ykoFEDC*, previously predicted to encode HMP transport proteins [Rodinov *et al.*, 2002]. DNA sequencing of PCR fragments containing this operon and the promoter region detected a single missense mutation in the *ykoD* gene that resulted in an Asp_{180} [GAC] to Asn_{180} [AAC] substitution.

Mutant *yloS* (*thiN*) has decreased pyrophosphokinase activity.

Growth of Tx1 was found to be slower on minimal medium lacking thiamin or TPP, taking more than three days to form colonies at 37°C. Addition of TPP or thiamin restored full growth within one day, suggesting than Tx1 was a strong thiamin bradytroph. In addition, Tx1 was strongly Lac+ on minimal medium containing thiamin, HMP, or TMP, but showed no Lac activity in presence of TPP. These phenotypes strongly suggested that the lesion in the pyrophosphokinase gene decreased the conversion of

thiamin to TPP whether thiamin is added externally or formed by dephosphorylation of TMP. Reconstruction of the Tx1 mutant confirmed this suspension. To do this, a double mutant (TH106) was constructed that consisted of a stable *thiN* deletion mutation (Δ*thiN*::*cat*$_4$) and a *thiL* disruption mutation (Ω*thiL*::pMUTIN2). Growth on minimal or complex medium of this double mutant was similar to that displayed by the original Tx1 mutant: TH106 could only grow on minimal medium in presence of TPP, addition of thiamin or TMP failed to restore growth. The Δ*thiN*::*cat*$_4$ alone in PY79 (TH105) showed no discernible phenotype in terms of growth on MM.

Simultaneous inactivation of both YkoD and YuaJ is required for thiamin deregulation.

Unlike Tx1, Tx26 containing mutant *ykoD* and *yuaJ* genes was a thiamin prototroph and grew like the parental strain on both minimal and complex media. Reconstruction experiments confirmed that inactivation of both *yuaJ* and *ykoD* were required for resistance to pyrithimin. To do this, a double mutant (TH137) was constructed that consisted of a stable *ykoD* deletion mutation (Δ*ykoD::cat*$_4$) and a *yuaJ* disruption mutation (Ω*yuaJ:*::pMUTIN2). This mutant was found to be resistant to pyrithiamin (Figure 2). However, the *yuaJ* and *ykoD* single mutants were pyrithiamin sensitive. To investigate thiamin deregulation in Tx26 further, β-galactosidase levels of the *thiC-lacZ* reporter were measured when Tx26 was grown under thiamin repressing and derepressing conditions.

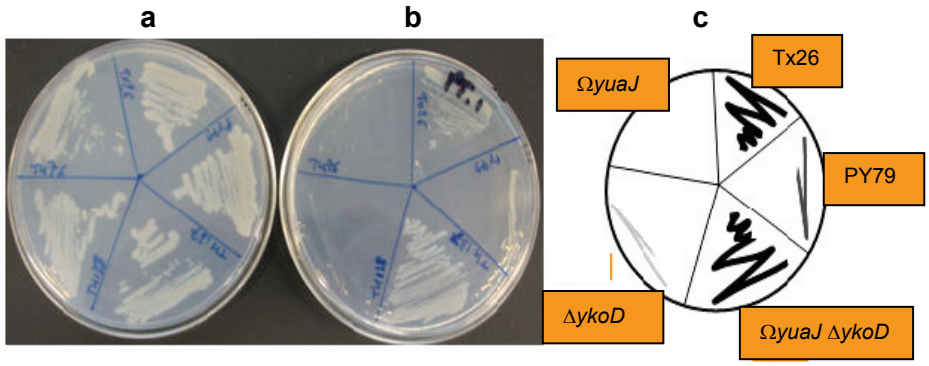


Figure 2: Reconstitution of the Tx26 pyrithiamin resistance phenotype. The strains with the indicated mutations were streaked for single colonies on MM **(a)** or MM containing 0.1μM pyrithiamin and incubated for 3 days at 37°C **(b).** Panel **(c)** is a cartoon rendition of the results in panel b.

As shown in Table 3, β-galactosidase activity was induced more that 200-fold relative to the wild-type control, when cells were grown to logarithmic stage in minimal medium containing TPP. The level of deregulation was reduced to about 30 fold in stationary phase. In addition, β-galactosidase levels were only two-fold higher in cells grown to logarithmic stage in minimal medium without TPP compared to cells grown with TPP (1 μM); this difference in β-galactosidase levels was less pronounced in stationary growth phase. This variability in derepression of the *thiC-lacZ* fusion probably reflects fluctuation in TPP intracellular levels during growth. Interestingly, expression of the fusion was still partially repressed in the presence of HMP, which was likely caused by conversion of HMP to TPP via the salvage pathway.

To determine if expression of other thiamin biosynthetic genes were deregulated in Tx26, transcription profiling was carried out using a custom *Bacillus subtilis* oligonucletotide microarray. A reconstructed, backcrossed thiamin deregulated mutant TH48 was

used in these first transcription profile experiments (see Material and Methods), because the mutation(s) causing thiamin deregulation (*i.e.* pyrithiamin-resistance) could be separated away from the *thiC-lacZ* fusion and any other unrelated EMS-induced mutation from the original Tx26 mutant (see below). Previous profiling results of wild-type *B. subtilis* PY79 observed a 30-90-fold thiamin-specific repression of the two major thiamin biosynthetic operons*, thiC* and t*enAIthiOSGFD*, but not of a third operon *ywbI-thiME* [Lee *et al*., 2001]. As shown in Table 4, high-level, deregulated expression of both *thiC* and *tenAI-thiOSGFD* was observed in TH48. Transcript levels of *thiC* were increased 105-fold in TH48 compared to PY79 when both strains were grown in minimal medium containing repressing levels of TPP. Transcription levels of *tenAI-thiOSGFD* were increased 30 – 100 fold depending on the gene. Transcription of a third biosynthetic operon containing *thiM* and *thiE*, and other known thiamin biosynthetic or salvage genes, *pdxK*, *dxs* (*yqiE*), *thiI* (*ytbJ*), and *thiN* (*yloS*) remained unchanged. A similar result was obtained by comparing transcription profiles from the original Tx26 mutant *vs.* the parental TH21 strain (data not shown). The Δ*thiL::cat*$_4$ mutation in both TH48 and Tx26 had little effect on the transcription profile of thiamin-biosynthetic genes other that *thiC* (2-3 fold increase compared to PY79 when strains were grown to logarithmic stage in the presence of TPP; data not shown).

Table 4: Change in transcript levels of genes in *B. subtilis* in response to TPP.

Gene [a]	*thi*-box element [e]	Enzyme/Function [a]	Ratio [b]	
			WT^-/WT^+ [c]	$Dereg^+/WT^+$ [d]
thiC	+	Biosynthesis of hydroxymethylpyrimidine phosphate	58	105
ywbI	-	Similar to LysR family of transcriptional regulators	NC	NC
thiM	-	Hydroxyethylthiazole kinase	NC	NC
thiE	-	Thiamin phosphate synthase	NC	NC
tenA	+	Thiaminase [f]	68	62
tenI	+	Possible antagonist of TenA	25	27
thiO (*goxB*)	+	Glycine oxidase	67	56
thiS (*yjbS*)	+	Sulfur carrier protein	87	86
thiG (*yjbT*)	+	Thiazole synthase	82	61
thiF (*yjbU*)	+	ThiS adenylyltransferase	87	96
thiD (*yjbV*)	+	Phosphomethylpyrimidine kinase	27	34
pdxK (*thiD*)	-	Pyridoxal/pyridoxal/pyridoxamine kinase	NC	NC
thiL (*ydiA*)	-	TMP kinase	NC	NC
thiN (*yloS*)	-	Thiamin pyrophosphorylase	NC	NC
ytbJ	-	Possible *thiI* orthologs (sulfur transferase)	NC	NC
dxs (*yqiE*)	-	1-deoxy-D-xylulose synthase	NC	NC
ykoC	+	Possible transmembrane component of thiamin-related ABC transporter	10	15
ykoD	+	Possible ATPase component of thiamin-related ABC transporter	11	16
ykoE	+	Possible transmembrane component of thiamin-related ABC transporter	22	32
ykoF	+	Possible ligand-binding protein of thiamin-related ABC transporter	17	23
yuaJ	+	Possible thiamin permease	6.6	8.4
ylmB	+	Similar to acetylornithine deacetylase	15	11

[a] Gene names and functions were described by Perkins and Pero, 2001, and in SubtiList (http://genolist.pasteur.fr/SubtiList).
[b] Transcript ratios were calculated by dividing the average difference values (after normalization) from hybridization experiments of wild-type cells grown to exponential phase in MM without TPP treatment by those with TPP treatment (WT^+/WT^-) or from hybridization experiments of deregulated mutant cells grown to exponential phase in MM with TPP treatment by those of wild-type cells grown under the same condition ($Dereg^+/WT^+$). For some genes (in bold), average difference values were obtained from duplicate probe sets per hybridization experiment: NC, no change in average difference values. Some data were taken from Table 1 in a report by Lee *et al*., 2001.
[c] r value of 0.95 (TPP-treated versus untreated data sets).
[d] r value of 0.85 (TPP-treated TH48 versus TPP-treated PY79 data sets).
[e] Under *thi* box element control according to report by Rodionov et al., 2002.
[f] See Toms *et al*., 2005.

More interestingly, the transcript levels of genes recently shown to contain *thi*-box regulatory sequences [Rodinov *et al.*, 2002] also exhibited thiamin-specific derepression. For example, transcription of the *ykoFEDC* gene cluster was significantly increased in TH48 relative to PY79: *ykoF*, *ykoD*, and *ykoC* transcripts were increased 15- to 23-fold, whereas *ykoE* was increased 32-fold. Other examples include the *yuaJ* gene, and *ylmB* encoding a protein similar to acetylornithine deacetylase. The average difference levels of *yuaJ* and *ylmB* were 8- and 11-fold higher, respectively. A similar pattern of regulated gene expression of *ykoFEDC*, *yuaJ*, and *ylmB*, was also observed in microarray analyses of wild-type cells grown in the presence or absence of thiamin. Finally in HMP-grown TH48, only expression of *thiC*, *thiOSGFD*, and *ykoFEDC*, were repressed relative to cells grown in the absence of HMP (data not shown).

Strains with mutant *yloS*, *yuaJ*, and *ykoD* genes overproduce and excrete thiamin products.

As shown in Table 3, there appeared to be a direct correlation between the degree of thiamin deregulation (*i.e.* expression level of the *thiC-lacZ* fusion in the presence of TPP) and excretion of thiamin products. Strain TH22 with mutant *thiL* alone was only weakly thiamin deregulated and excreted the least amount of thiamin products relative to the TH21 wild-type control strain (9-fold higher). Tx1 with mutant *thiL* and *thiN* genes, which showed partially thiamin deregulation, produced a 50-fold higher level of excreted thiamin products (5 µg/L). Lastly Tx26 containing mutant *thiL*, *yuaJ* and *ykoK* genes, which showed the highest level of thiamin deregulation, produced the highest level of excreted thiamin production (15 µg/L). This level of production was 100- to -150-

times higher than that excreted by the control strain, of which slightly over 50% was in the form of TPP (data not shown). The level of intracellular thiamin production of all mutants was approximately 500 µg/L or about three fold higher than the wild-type strain.

The four thiamin mutations, *tx26-1*, *tx26-2*, *thiL* and *tx1*, were next combined into a single strain (as described in Material and Methods) to determine if higher levels of thiamin overproduction and excretion could be obtained. Although Tx26 is a prototroph, the resulting strain (TH95) grew very poorly on minimal medium unless TPP or Nutrient Broth (reported as containing significant levels of thiamin precursors HMP and Thiazole; ref. 8) was added to the medium. In this respect TH95 was a stronger bradytroph than Tx1 (*i.e.* TH95 colonies took longer to appear on minimal medium than Tx1 colonies). Despite this growth defect, TH95 produced 3-fold more excreted thiamin products in shake-flask cultures compared to Tx26 (Table 5). Moreover, a shift in the excretion profile was observed from predominantly TPP in Tx26 to a combination of thiamin and TMP in TH95. This shift in the thiamin product profile was confirmed by thiochrome/HPLC (data not shown).

Table 5: Thiamin production of TH95.

Strain[a]	OD_{600}[b]	Concentration produced (μg/l)			
		Extracellular		Intracellular	
		Thiamin + TMP + TPP	TPP	Thiamin + TMP + TPP	TPP
Tx26 (Δ*thiL*::*cat*$_4$ *yuaJ ykoD amyE*::*thiC-lacZ*)	9.2	380	390	830 (25)[c]	370
TH95 (Δ*thiL thiN yuaJ ykoD ΩyloA*::Tn*917*)	10.4	1270	120	1150 (40)[c]	270

[a] Bacteria were grown in MM supplemented with 2.5% nutrient broth for 18 hours at 37°C. Culture media and cell-free extracts were measured for total thiamin products (thiamin + TMP + TPP) and TPP by a biological assay using *S. enterica* serovar Typhimurium indicators DM456 (*thiD906*::MudJ) and DMI856 (*thiL934*::Tn*10*), respectively.
[b] OD_{600}, optical density at 600 nm.
[c] Values in parentheses are the estimated extracellular concentrations of thiamin product if all are excreted into the culture medium in μg/liter; see Table 1.

Discussion

It is now well accepted that "*thi* box" genes are regulated by a mechanism referred as "Riboswitch". Using mainly in vitro methods, Winkler et al., 2002 convincingly showed that TPP alone acts to regulate thiamin gene expression by directly interacting with the nascent RNA transcript at the *thi*-box sequence to either block transcription or translation. The work presented here describes mutagenesis studies that identified and characterized new thiamin-deregulated mutations of *Bacillus subtilis* that are located within the coding regions of three genes: *thiN*, *yuaJ*, and *ykoD*. The encoded proteins have been lately suggested by *in silico* analysis to be involved in thiamin salvage and transport. We demonstrate here that indeed an important phenotype of these deregulatory mutants is the ability to excrete significant quantities of thiamin products and therefore that an important role of these proteins seems to be the control of TPP levels. It is important to note that the identification of deregulatory mutations within thiamin transport genes is strikingly similar to the recent results of Johansen *et al*., 2003 and Nygaard and Saxild, 2005, which showed that mutations that increase the activity of the *pbuE* encoded purine– specific pump deregulate expression of the G-box regulon by increasing the efflux of hypoxanthine, a metabolite that regulates purine gene expression by a Riboswitch mechanism.

Melnick *et al*., 2004 first demonstrated that *B. subtilis* contained a thiamin pyrophosphoryation activity normally found in yeast. This activity was encoded by *yloS*, subsequently renamed *thiN*. We present here experimental evidences confirming the function of this enzyme in thiamin biosynthesis. Indirect genetic evidence for this gene came from mutagenesis studies of B. subtilis

thiL gene, which resulted in mutants that were thiamin bradytrophs rather than strict auxotrophs as shown for *thiL* mutants of *S. typhimurium* and *E.coli*. This shows that *B. subtilis* has two pathways to produce TPP. Direct evidence was obtained by showing that the *tx1* mutation mapped to *thiN* and that reconstitution experiments indicated that inactivation of both *thiN* and *thiL* together resulted in a synthetic TPP auxotrophy in *B.subtilis*. In addition, the differential Lac activity displayed by Tx1 in the presence of thiamin (Lac+) and TPP (Lac-) is consistent with the presence of a direct thiamin-to-TPP route (Table 3). By blocking the ThiL reaction (TMP-to-TPP) and reducing the ThiN (YloS) reaction (THI-to-TPP) by the *tx1* mutation, the expected intracellular TPP level would be too low to repress *thi* gene expression. What is not clear, however, is whether these two pathways are linked (Figure 1). Preliminary genetic studies suggest that *B. subtilis* contains both thiamin kinase and TMP phosphatase activities (data not shown). However because the *B. subtilis* genome does not contain genes that display significant similarity to known thiamin kinase (*e.g. E. coli thiK* [*ycfN*]) or TMP phosphatase genes (yeast PHO3), it will be interesting to determine if these putative enzymes are specific to thiamin biosynthesis or represent activities from housekeeping or other biosynthetic pathways.

The mutations in the Tx26 mutant were found to be allelic to two putative thiamin transport genes, *yuaJ* (*tx26-1*) and ykoD (*tx26-2*). The *yuaJ* gene is predicted to encode a thiamin permease (passive transport) that contains six predicted trans-membrane domains [Rodinov *et al*., 2002]. The *ykoD* is predicted to encode an ATPase activity, which together with the *ykoC* and *ykoE* gene products, is predicted to form an ABC transporter [Rodinov *et al*., 2002]. The *tx26* mutations were found to generate three

phenotypes: deregulation of thiamin gene expression, an increase in pyrithiamin resistance, and an increase in extracellular levels of thiamin products. Moreover, the fact that inactivation of both *yuaJ* and *ykoD* was required to cause all three phenotypes suggested that these transporters control influx of TPP and other thiamin products into the cell (mutations that inactivate efflux transporters would not be consistent with our observation of higher extracellular levels of thiamin product with Tx26.). Based on these results we recommend that *yuaJ* be renamed *thiT* and the *ykoFEDC* operon, *thiUVWX*. Although *in silico* analysis does not indicate a function for *ykoF*, we have tentatively assigned it as a *thi* gene because it is co-transcribed with the *thiVWX* genes and the crystal structure of its encoded protein indicates that it binds thiamin molecules [Devedjiev *et al*., 2004]

Both *lacZ* reporter fusion (Table 3) and microarray experiments indicated that expression of all known *thi*-box-containing genes (*e.g. thiC*, *thiOSGFD*, *ykoFEDC*, and *yuaJ*) was highly derepressed in the Tx26 mutant. Since a key phenotype of this mutant is efflux of thiamin products from the cell, gene deregulation is most likely caused by changes in TPP intracellular levels. Interestingly, addition of exogenous HMP appeared to partially regulate thiamin gene expression (Table 3) in Tx26. We do not believe HMP is functioning as a repressor (or co-repressor) independent of TPP because *lacZ* reporter fusion experiments showed that HMP did not repress *thiC* expression in the Tx1 mutant. It seems more likely that HMP enters the cell by a different transport route and is converted to the TPP effector via the salvage pathway (HMP to HMP phosphate to TMP to TPP), which then interacts with thi box sequences of the nascent mRNA to repress thi gene expression via the Riboswitch mechanism.

Finally, an important application of our study was that these thiamin-deregulated mutants could be further manipulated to generate new strains that excreted more thiamin products. Combination of four mutations in *thiL*, *thiN*, *thiT* (*yuaJ*), and *thiW* (*ykoD*) has led to a strain that under 10-liter scale fermentation conditions excreted thiamin products in the mg/liter range (data not shown). Genetic engineering of the thiamin biosynthetic and precursor genes should lead to strains with a further increase in thiamin production.

Acknowledgments

We gratefully acknowledge Werner Bretzel for analytical support, Markus Goese for fermentation support, and Jian-ming Lee for microarray analysis of thiamin deregulated mutants. We thank Hans-Peter Hohmann for helpful discussion and critical reading of the manuscript. We also thank Tadhg Begley and Diana Downs for providing data before publication, and their expertise and support during the course of this work.

References

Begley, T. P., D. M. Downs, S. E. Ealick, F. W. McLafferty, A. P. Van Loon, S. Taylor, N. Campobasso, H. J. Chiu, C. Kinsland, J. J. Reddick, and J. Xi. 1999. Thiamin biosynthesis in prokaryotes. Arch Microbiol. 171:293-300.

Bolivar, F., R. L. Rodriguez, M. C. Betlach, and H. W. Boyer. 1977. Construction and characterization of new cloning vehicles I. Amplicillin-resistant derivatives of the plasmid pMB9. Gene. 2:75-91.

Chiu, H.J., J. J. Reddick, T. P. Begley, and S. E. Ealick. 1999. Crystal structure of thiamin phosphate synthase from Bacillus subtilis at 1.25 Å resolution. Biochemistry. 38:6460-6470.

Cutting, S. M., M. Anderson, E. Lysenko, A. Page, T. Tomoyasu, K. Tatematsu, T. Tatsuta, L. Kroos, and T. Ogura. 1997. SpoVM, a small protein essential to development in *Bacillus subtilis*, interacts with the ATP-dependent protease FtsH. J. Bacteriol. 179:5534-5542.

Cutting, S.M., and P. B. Vander Horn. 1990. Genetic analysis. p74. In C. R. Harwood and S. M. Cutting (ed.), Molecular biological methods for *Bacillus*. John Wiley and Sons, New York.

Devedjiev, Y., Y. Surendranath, U. Derewenda, A. Gabrys, D. R. Cooper, R. Zhang, L. Lezondra, A. Joachimiak, and Z. S. Derewenda. 2004. The structure and ligand binding properties of

the *B. subtilis* YkoF gene product, a member of a novel family of thiamin/HMPbinding proteins. J Mol. Biol. 343:395-406.

Difco. 1998. Manual, 11th ed. Difco Laboratories, Sparks.

Downs, D. M., and L. A. Petersen. 1994. *apbA*, a new genetic locus involved in thiamin biosynthesis in *Salmonella typhimurium*. J. Bacteriol. 176:4858-4864.

Guérout-Fleury, A. M., N. Frandsen, and P. Stragier. 1996. Plasmids for ecotopic integration in *Bacillus subtilis*. Gene. 180:57-61.

Harwood, C. R., and A. R. Archibald. 1990. Growth, maintenance and general techniques, p1-26. In C. R. Harwood and S. M. Cutting (ed.), Molecular Biological Methods for Bacillus. John Wiley & Sons, New York.

Johansen, E. L. P., P. Naygarrd, C. Lassen, Y. Agersø, and H. H. Saxild. 2003. Definition of a second *Bacillus subtilis pur* regulon comprising the *pur* and *xpt-pbuX* operons plus *pbuG, nupG* (yxjA), and pbuE (ydhL). J. Bacteriol. 185:5200-5209.

Kawasaki, T., and Y. Nose. 1969. Thiamine regulatory mutants in *Escherichia coli*. J Biochem. (Tokyo). 65:414-425.

Lawhorn[a], B. G., S. Y. Gerdes, and T. P. Begley. 2004. A genetic screen for the identification of thiamin metabolic genes. J. Biol. Chem. 279:43555-43559.

Lawhorn[b], B. G., R. A. Mehl, and T. P. Begley. 2004. Biosynthesis of the thiamin pyrimidine: the reconstitution of a remarkable rearrangement reaction. Org Biomol Chem 2:2538-2546.

Lee, J.-M., S. Zhang, S. Saha, S. Santa Anna, C. Jiang, and J. Perkins. 2001. RNA expression analysis using an antisense *Bacillus subtilis* genome array. J. Bacteriol. 183:7371-7380.

Mandal, M., B. Boese, J. E. Barrik, W. C. Winkler, and R. R. Breaker. 2003. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. Cell 113:577-586.

Maniatis, T., E. Fritsch, and J. Sambrook. 1982. Molecular cloning. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Melnick, J., E. Lis, J. H. Park, C. Kinsland, H. Mori, T. Baba, J. Perkins, G. Schyns, O. Vassieva, A. Osterman, and T. P. Begley. 2004. Identification of the two missing bacterial genes involved in thiamine salvage: thiamine pyrophosphokinase and thiamine kinase. J Bacteriol.186:3660-3662.

Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Miranda-Rios, J., M. Navarro, and M. Soberson. 2001. A conserved RNA structure (*thi* box) is involved in regulation of thiamin biosynthesis gene expression in bacteria. Proc. Natl. Acad. Sci. USA. 98:9736-9741.

Mironov, A. S., I. Gusarov, R. Rafikov, L. E. Lopez, K. Shatalin, R. A. Kreneva, D. A. Perumov, and E. Nudler. 2002. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. Cell 111:747-756.

Nosaka, K., Y. Kaneko, H. Nishimura, and A. Iwashima. 1993. Isolation and characterization of a thiamin pyrophohokinase gene, THI80, from *Saccharomyces cerevisiae*. J. Biol. Chem. 268:17440-17447.

Nygaard, P., and H. H. Saxild. 2005. The purine efflux pump PbuE in *Bacillus subtilis* modulates expression of the PurR and G-box (XptR) regulatons by adjusting the purine base pool size. J. Bacteriol. 187:791-794.

Park, J. H., K. Burns, C. Kinsland, and T. P. Begley. 2004. Characterization of two kinases involved in thiamine pyrophosphate and pyridoxal phosphate biosynthesis in *Bacillus subtilis*: 4-amino-5-hydroxymethyl-2-methylpyrimidine pyrophosphate kinase activity. J. Bacteriol. 186:1571-1573.

Park, J. H., P. C. Dorrestein, H. Zhai, C. Kinsland, F. W. McLafferty, and T. P. Begley. 2003. Biosynthesis of the thiazole moiety of thiamin pyrophosphate (vitamin B1). Biochemistry. 42:12430-12438.

Perkins, J. B., and J. G. Pero. 2001. Vitamin biosynthesis, p279-293. In A. L. Sonenshein, J.A. Hoch, and R. Losick (ed.), *Bacillus subtilis* and its relatives: from genes to cells. American Society for Microbiology, Washington, D.C.

Petersen, L. A., and D. M. Downs. 1997. Identification and characterization of an operon in *Salmonella typhimurium* involved in thiamine biosynthesis. J. Bacteriol. 179:4399-4402.

Rodionov, D. A., A. G. Viteschak, A. A. Mironov, and M. S. Gelfand. 2002. Comparative genomics of thiamin biosynthesis in prokaryotes. J. Biol. Chem. 277:48949-48959.

Schowen, R. 1998. Thiamin-dependent enzymes. In L. Sinnott (ed.), Comprehensive catalysis. Academic Press, San Diego.

Toms, A. V., A. L. Haas, J. H. Park, T. P. Begley, and S. E. Ealick. 2005. Structural characterization of the regulatory proteins TenA and TenI from Bacillus subtilis and identification of TenA as a thiaminase II. Biochemistry 44:2319-2329.

Vagner, V., E. Dervyn, and S. D. Erhlick. 1998. A vector for systematic gene inactivation in *Bacillus subtilis*. Microbiology 144:3097-3104.

Vandeyar, M. A., and S. A. Zahler. 1986. Chromosomal insertions of Tn*917* in *Bacillus subtilis*. J. Bacteriol. 167:530-534.

Wach, A. 1996. PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae*. Yeast. 12:259-265.

Webb, E., K. Claas, and D. Downs. 1998. t*hiBPQ* encodes an ABC transporter required for transport of thiamine and thiamine

pyrophosphate in *Salmonella typhimurium*. J. Bacteriol. 273:8946-8950.

Webb, E., and D. M. Downs. 1997. Characterization of *thiL*, encoding thiamin-monophosphate kinase, in *Salmonella typhimurium*. J. Biol. Chem. 272:15702-15707.

Webb, E. A., F. Ferbres, and D. M. Downs. 1996. Thiamin pyrophosphate (TPP) negatively regulates transcription of some *thi* genes in *Salmonella typhimurium.* J. Bacteriol. 178:2533-2538.

Winkler, W. C., A. Nahvi, and R. R. Breaker. 2002. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature. 419:952-956.

Zhang, Y., and T. P. Begley. 1997. Cloning, sequencing, and regulation of thiA, a thiamin biosynthesis gene from *Bacillus subtilis*. Gene. 198:73-82.

Zhang, Y., H. J. Chiu, and T. P. Begley. 1997. Characterization of the *Bacillus subtilis thiC* operon involved in thiamine biosynthesis. J.Bacteriol. 179:3030-3035.

CHAPTER IV

Bacillus subtilis spore display of recombinant proteins using a coat-associated enzyme as carrier

Sébastien Potot[Ŧ], Claudia Serra[¤],
Adriano O. Henriques[¤] and Ghislain Schyns[Ŧ]

Applied and Environmental Microbiology. 2010. 76(17):5926-33.

[Ŧ] DSM Nutritional Products, Biotechnology R&D, PO box 3255, Bldg 203/25, CH-4002 Basel, Switzerland.

[¤] Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Apartado 127, 2781-901 Oeiras, Portugal.

Contribution:

Sébastien Potot conducted all the experiments described except the construction of *E. coli* strain SD58, *B. subtilis* strains AH7666, and AH2873, and the validation of the anti-phytase antibodies.

Abstract

The display of proteins such as feed enzymes at the surface of bacterial spore systems has a great potential use for animal feed. Feed enzymes increase the digestibility of nutrients, leading to greater efficiency in the manufacturing of animal products and minimizing the environmental impact of increased animal production. To deliver their full potential in the gut, feed enzymes must survive the harsh conditions of the feed preparation and the gastrointestinal tract. The well documented resistance of spores to harsh environments, together with the ability to use proteins that compose the spore as carriers for the display of passenger proteins, suggests that spores could be used as innovative tools to improve the formulation of bioactive molecules. Although some successful examples have been reported, in which abundant structural proteins of the *Bacillus subtilis* spore outer-coat layer were used as carriers for the display of recombinant proteins, only one convincing example resulted in the display of functional enzymes. In addition, no examples are available about the use of an inner-coat protein for the display of an active passenger enzyme. In our study, we show that the inner-coat oxalate decarboxylase (OxdD) can expose an endogenous phytase, a commonly used feed enzyme for monogastric animals, in an active form at the spore surface. Importantly, despite the higher abundance of CotG outer-coat protein, an OxdD-Phy fusion was more represented at the spore surface. The potential of OxdD as a carrier protein is further documented through the spore display of a bioactive heterologous passenger, the tetrameric β-glucuronidase enzyme from *Escherichia coli.*

Introduction

Under extreme nutrient deprivation, *Bacillus subtilis* has the ability to enter a complex differentiation process that culminates with the formation of an extremely resistant spore. Spores consist of a central core compartment that contains a copy of the chromosome, and is surrounded by a thick layer of a modified form of peptidoglycan known as the cortex. The cortex is covered by a multilayered protein coat, formed by an inner layer apposed to the cortex, and an outer layer. In most characterized strains of *B. subtilis*, the outer-coat layer is the outermost spore structure. The spore coat is formed through synthesis in the mother cell of over 70 polypeptides, which follow an ordered assembly program at the surface of the developing spore [Henriques *et al*., 2004]. Assembly is largely guided by a class of morphogenetic proteins that have important roles in the structural organization of the two main coat layers [Driks, 2002; Henriques and Moran, 2007; Imamura *et al.*, 2010]. SafA, for example, localizes at the interface between the cortex and inner coat, and has an important role in assembly of the inner-coat layer [Costa *et al*., 2006]. CotE, on the other hand, is found at the inner coat-outer coat interface, and is essential for assembly of the outer-coat structure. When conditions become conductive for growth, spores germinate to regenerate vegetative cells, which rapidly resume the normal cyclic pattern of growth and division.

The *B. subtilis* spore coat has recently emerged as a nanostructure offering a novel and interesting surface for the display of biomolecules. Since *B. subtilis* presents a good safety record as an additive in human and animal preparations (GRAS; generally regarded as safe), one potentially valuable use of the spore coat

display system is in the area of probiotics. In animal nutrition, feed enzymes are commonly used to improve the nutrition value of feeds mainly by enhancing their digestibility and/or assimilation [Choct, 2006]. Display of these enzymes at the spore surface could ensure *in situ* efficient enzymatic activity application at moderate cost. Examples of feed enzymes candidate for display are xylanase, hemicellulase, cellulase, protease, glycanase, or phytase. Phytase, in particular, is a commonly used feed enzyme for monogastric animals to improve nutritive value [Rao *et al.,* 2009]. Most of the phosphorus (50-80%) contained in feedstuffs of plant origin exists as the storage form phytate and is indigestible for non-ruminant animals such as poultry and pigs, since they lack the enzyme to free phytate-bound phosphorus. Therefore, sufficient phytase needs to be added to the feed to decrease the supplementation of phosphorous to feedstuffs, thus reducing the environmental pollution in areas with intensive livestock production. However, despite successful spore display examples using the abundant structural coat proteins CotB [Duc *et al.*, 2007; Hinc *et al*., 2010; Isticato *et al.*, 2001], CotC [Hinc *et al*., 2010; Mauriello *et al.*, 2004; Ricca and Cutting, 2003] and CotG [Hinc *et al*., 2010; Kim *et al*., 2005] as an anchoring motif, a very limited number of studies is available regarding spore display of functional enzymes [Kim *et al*., 2004; Kwon *et al*., 2007; Yim *et al*., 2009]. In all reported cases of spore display, the common denominator governing choice of the carrier protein seems to have been its abundance, and its ability to ensure the highest level of surface exposure. CotB, CotC and CotG, possibly the most represented proteins within the coat, are outer-coat proteins, dependent on CotE for assembly (Fig. 1A). All three proteins also undergo extensive multimerization during their assembly at the spore surface [Isticato *et al*., 2004; Zilhão *et al*.,

2004] (Fig. 1A). Thus far, the display of antigens, enzymes or other functional components on spores using inner-coat proteins as carriers, has not been reported.

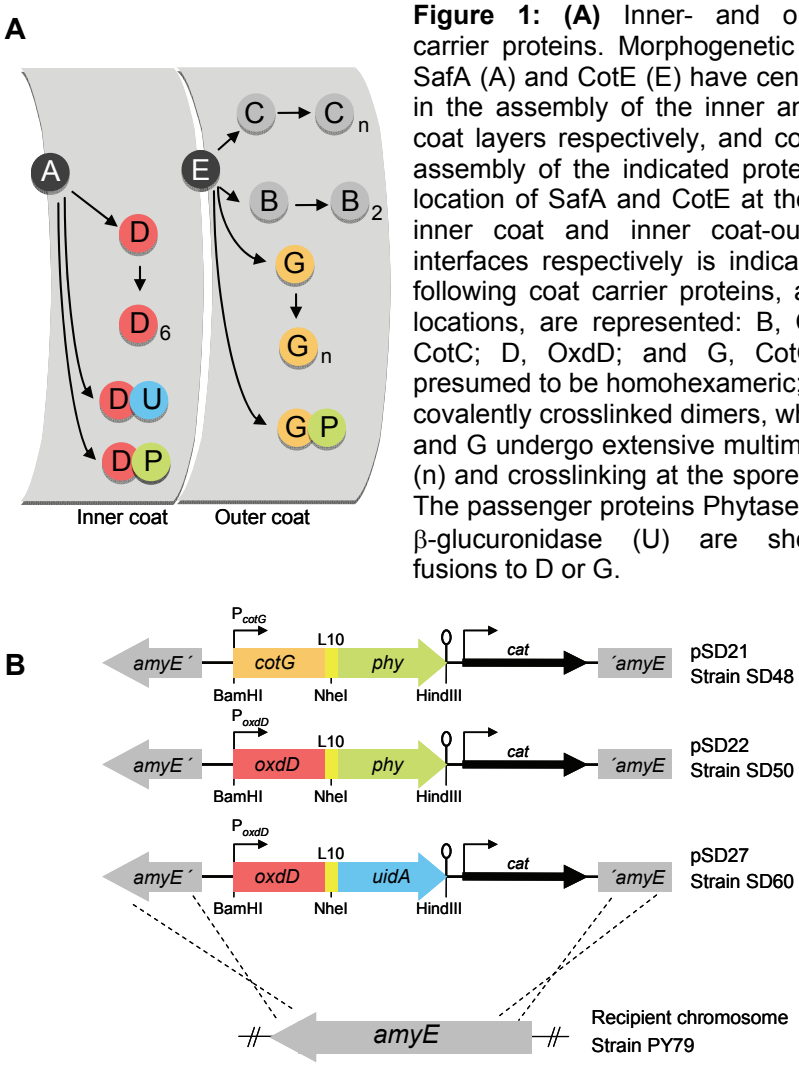


Figure 1: (A) Inner- and outer-coat carrier proteins. Morphogenetic proteins SafA (A) and CotE (E) have central roles in the assembly of the inner and outer-coat layers respectively, and control the assembly of the indicated proteins. The location of SafA and CotE at the cortex-inner coat and inner coat-outer coat interfaces respectively is indicated. The following coat carrier proteins, and their locations, are represented: B, CotB; C, CotC; D, OxdD; and G, CotG. D is presumed to be homohexameric; B forms covalently crosslinked dimers, whereas C and G undergo extensive multimerization (n) and crosslinking at the spore surface. The passenger proteins Phytase (P), and β-glucuronidase (U) are shown as fusions to D or G.

Figure 1: (B) General strategy used to construct the various carrier-passenger fusions. The indicated pDG364 derivatives, carrying the various gene fusions, were linearized and transferred to the non-essential *amyE* locus through a marker replacement (double-crossover) recombination event. Promoters are represented by arrows and transcriptional terminators are represented by hairpin structures. L_{10}, linker made of ten alanine residues; *cat*, gene encoding chloramphenicol acetyltransferase. The various genes, color-coded to match panel A, are not drawn to scale.

The 43.4-kDa product of the *oxdD* gene, OxdD, is a minor component of the spore coat [Costa *et al*., 2004]. OxdD is a component of the inner-coat layers, dependent on morphogenetic protein SafA for assembly [Costa *et al*., 2004; Ozin *et al*., 2001] (Fig. 1A). OxdD is highly similar to OxdC, a homohexameric enzyme (EC 4.1.1.2), which is specifically produced during growth of *B. subtilis* under acidic conditions [Tanner *et al*., 2001]. Both OxdD and OxdC show oxalate decarboxylase activity [Costa *et al*., 2004; Tanner *et al*., 2001].

We show here that *B. subtilis* OxdD can be used as an original anchoring motif to display proteins of biological interest at the surface. Because of its relatively low abundance and internal localization, the use of the coat-associated OxdD enzyme as a carrier protein for spore display could afford a higher degree of protection to a passenger protein while minimizing any impact on the spore coat assembly process. First, as a step in the development of spore display for use in probiotics, OxdD was fused *to B. subtilis* phytase (Phy), a monomeric enzyme (EC 3.1.3.8) encoded by the *phy* gene with optimal activity at pH 7 to 7.5 [Kerovuo *et al*., 1998]. We show that OxdD-Phy fusion protein is exposed at the spore surface and results in recombinant spores with phytase activity. We also show that the spore coat structural protein CotG can expose the endogenous phytase, but despite the higher abundance of this protein, the resulting recombinant spores show reduced levels of phytase compared to those of a strain expressing *oxdD-phy*. The potential use of OxdD as a carrier for the display of functional enzymes at the spore surface is further illustrated by the display of a fusion of OxdD to β-glucuronidase (GusA) from *Escherichia coli*, a homotetrameric enzyme encoded by the *uidA* gene [Jefferson *et al*., 1986; Karow and Piggot, 1995].

Materials and Methods

Bacterial strains, media and general techniques.

The *B. subtilis* strains used in the present study are listed in Table 1. They are derived from strain PY79 (1A747; Bacillus Genetic Stock Center). *E. coli* K-12 was used to PCR-amplify the *uidA* gene, encoding for β-glucuronidase (GusA). The high-fidelity Herculase polymerase (Stratagene) was used to generate PCR products. Tryptone Blood Agar Broth (BD Difco) was used as the standard solid medium for routine propagation of all *B. subtilis* strains. Difco sporulation medium (DSM) was used to induce sporulation by nutrient exhaustion [Nicholson and Setlow, 1990].

Table 1: Bacterial strains

Strain	Genotype / phenotype	Origin or reference
E. coli		
SD58	pET16::*phy* / overproduction of His_6-Phy lacking its signal peptide, under the control of P_{T7lac}	This work
B. subtilis		
PY79	*B. subtilis* wild type	BGSC [(b)]
SD48	PY79 Δ*amyE*::*cotG-phy* / Cm^r [(a)]	This study
SD50	PY79 Δ*amyE*::*oxdD-phy* / Cm^r [(a)]	This study
SD53 (AH2873)	Δ*oxdD*::*oxdD-gfp* / Sp^r [(a)]	Costa *et al.*., 2004
SD60	PY79 Δ*amyE*::*oxdD-uidA* / Cm^r [(a)]	This study
AH7666	PY79 *Δphy* / Sp^r [(a)]	This study

[a] Cm^r, chloramphenicol resistance; Sp^r, spectinomycin resistance

[b] BGSC; *Bacillus* Genetic Stock Center (http://www.bgsc.org).

Construction of *B. subtilis* strains expressing CotG-Phy, OxdD-Phy and OxdD-GusA fusion proteins.

A schematic view of the cloning strategy used for the translational fusions is presented in Fig.1B. The original promoter of the spore-anchoring motif (300 bp of DNA upstream from *oxdD* start codon, or 465 bp of DNA upstream *cotG* start codon) was used to synchronize expression with the passenger protein during sporulation. Passenger proteins were fused in frame with the carboxyl terminus of the anchoring motif (either OxdD or CotG), deleted of its stop codon. *B. subtilis* native phytase Phy (GenBank accession number BG11198), was fused to either OxdD (BG13484) or CotG (BG11017) (numbering according to Genolist, Institut Pasteur, 1999-2001; Subtilist [http://genolist.pasteur.fr/SubtiList] and Colibri [http://genolist.pasteur.fr/Colibri]). The region encoding the signal peptide of Phy (encompassing the 26 first residues; Swiss-Prot P42094) was excluded. A linker made of 10 alanine residues was introduced between the carrier and passenger proteins. Gene fusions *cotG-phy* (2,173 bp long) and *oxdD-phy* (2,594 bp long) were synthesized as synthetic genes and sequenced by DNA2.0 Inc. (USA, CA). They were then cloned between the BamHI and HindIII sites of pDG364 suicide vector [Cutting and Vander Horn, 1990], resulting in plasmids pSD21 (*cotG-phy*) and pSD22 (*oxdD-phy*). To construct a fusion of the *E. coli uidA* gene, coding for β-glucuronidase (GusA; GenBank accession number EG11055) to the 3'-end of the *oxdD* gene, we first introduced a NheI restriction site into the 10-residue linker of pSD22, leading to the replacement of the sixth alanine residue by a serine. Then, the *uidA* gene of *E. coli* was cloned between NheI and HindIII sites of plasmid pSD22, to yield pSD27, carrying a 3,339-bp-open reading frame coding for a OxdD-GusA fusion protein under

the control of the native *oxdD* promoter. After linearization with XhoI restriction endonuclease, pSD21, pSD22 and pSD27 were used to independently transform *B. subtilis* PY79 with selection for chloramphenicol (5 μg/ml). Transformants that resulted from a marker replacement recombination event (double crossover) at the non essential *amyE* locus were identified and designated SD48 (expressing the CotG-Phy fusion), SD50 (OxdD-Phy), and SD60 (OxdD-GusA).

Construction of SD58 (*E. coli* His_6-Phy, N-term).

A 1.1-kb PCR fragment containing the *phy* gene without its first 84 codons (corresponding to the signalling peptide) was amplified by using primers phy-for 5'-GGAATTCCATATGGTGAATGAGGAACATCATTTCAAAG-3' and phy-rev 5'-ATCGCTCGAGGCCGTCAGAACGGTCTTTCAGCTTCCTC-3', which generated NdeI and a XhoI restriction site ends, respectively. Chromosomal DNA of *B. subtilis* 168 was used for amplification. The PCR fragment was digested and inserted into pET16b (Novagen) to create pSD26. Selection of transformants was done on LB plates containing 100 μg of ampicillin /ml. Plasmid pSD26 carries an N-terminal fusion of *B. subtilis* phytase (without its signalling peptide) fused to a His tag. After purification, pSD26 was introduced into *E. coli* BL21 competent cells (Invitrogen) to allow efficient overexpression of the heterologous protein. Selection of transformants was done in LB plates containing 100 μg of ampicillin /ml and 50 μg of chloramphenicol /ml. The resulting strain was named SD58.

Overproduction and purification of *B. subtilis* phytase.

A starter culture of strain SD58, grown overnight at 37°C with shaking, was used to inoculate 100 ml of fresh LB broth supplemented with ampicillin and chloramphenicol, which was incubated at 37°C and 150 rpm until reaching an optical density at 600 nm of ~0.6. At this point, IPTG (isopropyl-β-D-thiogalactopyranoside) was added to a final concentration of 1 mM and the culture was incubated for another 4 hours. The culture was then centrifuged 20 minutes at 8,000 x g and 4ºC, and the cell pellet was resuspended in 4 ml of ice cold lysis buffer (20 mM phosphate, 300 mM NaCl, 1 mM $CaCl_2$, 10 mM imidazole, 1 mM phenylmethylsulfonyl fluoride, 1X anti-protease EDTA free solution). The cell lysate was obtained by passage through a French Press cell at 19,000 lb/in^2, and the supernatant fraction was used to purify the His_6-tagged phytase on a Ni^{2+}-nitrilotriacetic acid (NTA) column (Qiagen). The column was operated according to the manufacturer´s instructions. Bound phytase was eluted at an imidazole concentration of 300 mM. All purification steps were carried out at 4ºC.

Isolation of *B. subtilis* extracellular proteins.

The extracellular proteome was prepared essentially as described by Antelmann et *al.*, 2007. Briefly, cultures (20 ml) of PY79 and its derivative bearing a Δ*phy* mutation (AH7666), were grown in a synthetic medium containing 15 mM $(NH_4)_2SO_4$, 8 mM $MgSO_4.7H_2O$, 27 mM KCl, 7 mM sodium citrate. $2H_2O$, 50 mM Tris-HCl (pH8) supplemented with 0.16 mM KH_2PO_4, 2 mM $CaCl_2.2H_2O$, 1 µM $FeSO_4.7H_2O$, 10 µM $MnSO_4.4H_2O$, 4.5 mM glutamic acid, 780 µM tryptophan, 860 µM lysine, and 0.2% (wt/vol) glucose. Samples were collected 1 h after the end of the exponential phase of growth

by centrifugation for 30 minutes at 21,000 x g and 4ºC. Proteins present on the culture supernatant were then precipitated during 1 h at 4ºC with trichloroacetic acid (TCA) at a final concentration of 5% and recovered by centrifugation during 30 minutes at 21,000 x g and 4ºC. The protein pellet was once washed with 70% ice-cold ethanol and finally resuspended in 200 µl of a buffer containing 10 mM Tris (pH 8), 10 mM $MgCl_2$, 500 mM EDTA, 200 mM NaCl, 10% glycerol, 1mM PMSF, and 1X anti-protease EDTA free solution.

Construction of AH7666 (B. subtilis wild-type with *phy* deleted).

Long flanking homology PCR was used to delete the endogenous copy of *phy* in PY79 background [Walch, 1996]. The primers P1-phy 5'-ATATCTGCCTAAAAAAGTGC-3' and P2-phy-spec 5'-ACATGTATTCACGAACGAAAATCGAAATAGAAAGCAGCTTGTG CAGC-3' were used to amplify a fragment upstream of the *phy* gene, while the primers P3-phy-spec 5'-ATTTTAGAAAACAATAAACCCTTGCACCTTCATTTGGTACCCTC C-3' and P4-phy 5'-CACCTGTTTAGGTGTAAGCAG-3' were used to amplify a fragment downstream of *phy*. Both PCR products were bridged by a fragment containing the spectinomycin cassette [Guérout-Fleury *et al.*, 1995]. The final PCR product was then used to replace the endogenous *phy* gene by the spectinomycin cassette via double crossover.

Spore purification and analysis.

Spores were harvested by centrifugation of DSM cultures 24 h after the onset of sporulation. After washing, spores were purified with a 20 to 50% Gastrografin (Schering) step gradient, as previously described for Renocal-76 gradients [Henriques *et al.*,

1995]. Purified spores were suspended in cold buffers (as specified below) supplemented with protease inhibitors (Complete EDTA-free; Roche). Samples of the purified spore suspension were used to analyze the coat polypeptide composition as described before [Henriques *et al*., 1995]. The total cell count, as well as the titer of heat or lysozyme-resistant spores was determined in samples of DSM cultures 24 h after the onset of sporulation [Henriques *et al*., 1995].

Detection of β-glucuronidase (GusA) activity by fluorescence microscopy.

In situ detection of β-glucuronidase activity was performed using the fluorogenic substrate C_{12}FDGlcU, a lipophilic analog of fluorescein di-β-D-glucuronic acid containing a 12-carbon aliphatic chain supplied with the ImaGene Green C_{12}FDGlcU GUS gene expression kit from Molecular Probes. Cleavage by β-glucuronidase releases the yellow-colored, green-fluorescent compound 5-dodecanoylaminofluorescein (Ex_{max}= 495 nm, Em_{max}= 518 nm). This substrate was used on purified spores of PY79 and SD60, according to the indications of the manufacturer. Before adding the substrate, part of the spore suspension was washed with water and incubated for 1 h at 37°C in a 0.1% trypsin solution (Amimed), as a control. Spores were washed three times with phosphate-buffered saline (PBS) at pH 7.4 (Roche). The spores were then examined by fluorescence microscopy (Eclipse E600; Nikon) using phase-contrast optics and a standard filter for the visualization of green fluorescence emission (Ex = 490 nm, Em = 520 nm), reflecting GusA activity at the spore surface. Images were acquired with a digital camera (CoolSnap EZ; Visitron Systems) and processed with *ImageJ* (W.S Rasband, *ImageJ*, 1997-2008.

[http://rsb.info.nih.gov/ij/]; National Institutes of Health, Bethesda, MD, USA) for quantification of the fluorescence signal. Background was measured for each picture and subtracted from the fluorescence intensity of the samples. The quantification was made on 150 spores per strain.

Assay of *B. subtilis* phytase activity.

Whole-cell extracts were prepared from 24-h DSM cultures of strains SD48, SD50 and PY79. Cells and mature spores were harvested by centrifugation (20 min, 8,000 x g), and suspended in a buffer containing 100 mM Tris-HCl pH 7.4 (Fluka), a cocktail of protease inhibitors (Complete EDTA-free) and 2 mM $CaCl_2$ (Sigma), required to maintain active conformation of *B. subtilis* phytase [Kerovuo *et al*., 2000; Shimizu, 1992]. The mixture was passed twice in French cell press (Sim-Aminco) at 900 lb/in^2 to release both forespores and the cytoplasm content of the mother cells. Spores were purified from whole-cell extracts as described previously [Zilhão *et al*., 2004]. Phytase activity was measured by incubating either 10 µl of whole-cell extracts or 10 µl of pure spores suspensions with 40 µl of 2 mM sodium phytate (Sigma) in the above-mentioned buffer. After 30 min of incubation at 55°C, the reaction was stopped by adding 50 µl of 5% TCA (Merck) [Shimizu, 1992]. After centrifugation (4 min, 8,000 x g), 50 µl of the supernatant was diluted in 500 µl of water. The released orthophosphate (Pi) was photometrically measured at 700 nm by monitoring the production of phosphomolybdate for 15 min at 50°C after the addition of 500 µl of the color reagent. The color reagent was prepared freshly by mixing one volume of 2.5% ammonium molybdate (Fluka), one volume of 10% sodium ascorbate (Fluka), and three volumes of 1 M H_2SO_4 (Fluka) [Kim and Lei, 2005]. One unit of phytase was defined as the amount of enzyme required to

release one micromole of Pi from sodium phytate in one minute. The total amount of protein was assayed with the bicinchoninic acid method (Pierce), in order to calculate specific activity. The dry weight was measured after drying spores suspensions for 24 h at 80°C.

Immunodetection of spore displayed phytase.

An anti-phytase rabbit polyclonal antibody was produced by immunizing rabbits with a mix of 2 synthetic phytase-specific peptides (NH_2-CAEPGGGSKGQVVDRA-COOH and NH_2-CHKQVNPRKLKDRSDG-COOH) (Eurogentec). Before immunolabelling, part of the spores was incubated at 37°C for 1h in a 0.1% trypsin solution (Amimed), as a control. They were washed three times with PBS (pH 7.4) supplemented with 2% bovine serum albumin (Sigma) and then probed with primary and secondary antibodies. Immunostained spores were mounted on agarose-coated microscope slides and observed by fluorescence as described above. Fluorescence signal was quantified on 150 spores per strain.

Results and Discussion

Amongst spore-associated enzymes, OxdD resides in the inner-coat layers of *B. subtilis* spores [Costa *et al.*, 2004] (Fig. 1A). A passenger protein fused to this enzyme could in theory be more protected than when fused to a carrier located in the outer coat. Here, we investigated the behavior of the inner-coat protein OxdD by comparison to the outer-coat protein CotG, as a carrier for the display of monomeric *B. subtilis* phytase. We also assessed the ability of OxdD to display the heterologous multimeric *E. coli* β-glucuronidase.

Spore display of B. subtilis phytase.

Although green fluorescence protein (GFP) fused to OxdD has been successfully exposed at the surface of *B. subtilis* spores [Costa *et al.*, 2004], spore display of a functional enzyme has not been reported. The *B. subtilis* Phy passenger enzyme, truncated for its 26-residues signal peptide, was translationally fused to the C terminus of the anchoring motif OxdD. A linker, made of ten aliphatic amino acids, was introduced between the carrier and passenger proteins to minimize any potential steric effect that could disturb the correct folding of either protein (Fig.1B). Since CotG had been previously described as the carrier for the spore display of recombinant proteins [Hinc *et al*., 2010; Kim *et al*., 2004; Kwon *et al*., 2007], a control fusion CotG-Phy was also designed (Fig. 1B). The translational *oxdD-phy* and *cotG-phy* fusions are carried by strains SD50 and SD48, respectively (Table 1).

Table 2: Heat resistance and Lysozyme resistance to various *B. subtilis* strains

Strain	Relevant genotype	Sporulation (CFU/ml)[a]		
		Viable cells	Heat-resistant cells	Lysozyme-resistant cells
PY79	Wild-type	4.2×10^8	3.2×10^8	2.5×10^8
SD48	Δ*amyE*::*cotG-phy*	4.2×10^8	4.8×10^8	4.0×10^8
SD50	Δ*amyE*::*oxdD-phy*	4.5×10^8	4.3×10^8	3.2×10^8
SD60	Δ*amyE*::*oxdD-uidA*	4.8×10^8	2.6×10^8	2.9×10^8

[a] The total viable, heat-resistant, or lysozyme-resistant cell enumeration was determined 24 hours after the onset of sporulation in liquid medium (DSM), as described in Materials and Methods. CFU: colony forming unit.

Since in spores of *cotG* insertional mutants the normal structural organization of the outer coat is severely disturbed [Henriques *et al*., 1998], we have chosen to allow expression of the wild-type *cotG* allele in strains expressing passenger-carrier fusions (strain SD48). A similar approach was used with OxdD as a carrier (strain SD50), since the function of OxdD in spore assembly is not yet fully elucidated [Costa *et al*., 2004]. Both integrations at the *amyE* locus of the wild-type strain PY79 had no detectable effect on spore viability and resistance to heat and to lysozyme (Table 2). Moreover, examination of the collection of coat polypeptides extracted from spores of strains SD48 or SD50 resulted in a pattern that did not differ from that obtained for the wild-type PY79 spores (data not shown). We concluded that expression of the wild-type

and engineered carrier proteins did not interfere in any detectable way on the assembly and function of the spore coat lattice.

To test whether the fusion proteins were presented at the spore surface, we conducted immunofluorescence microscopy experiments. For these studies, we used a rabbit polyclonal anti-Phy antibody (see Material and Methods). This antibody specifically recognized the purified His_6-Phy protein (Fig. 2A). In addition, it recognized the native Phy protein in extracts prepared from culture supernatants of the wild-type strain PY79, grown under conditions where Phy is known to accumulate [Antelmann *et al.*, 2007], and did not react with any protein in extracts prepared in parallel from a *phy*-deletion mutant (AH7666) (Fig. 2B). Spores were collected 24 h after the onset of sporulation, purified on density gradients (see Material and Methods) and either directly analyzed or treated with trypsin as a control.

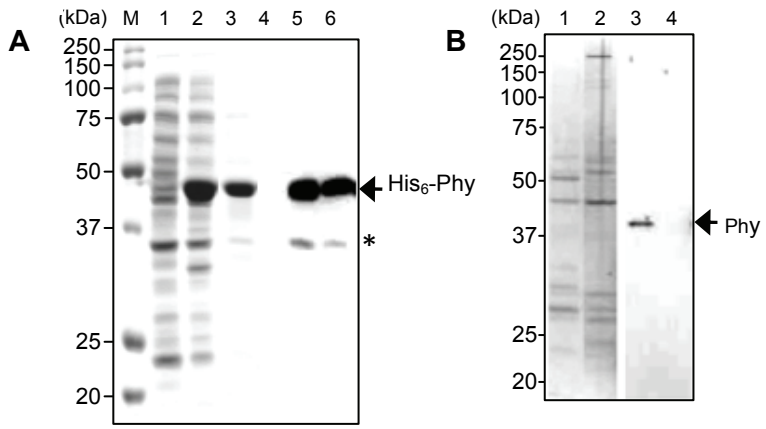


Figure 2. Purification of *B. subtilis* phytase (Phy) and specificity of an anti-Phy antibody. (A) SDS-PAGE analysis (lanes 1 to 3) and immunoblot analysis (lanes 4 to 6) of extracts of *E. coli* strain SD58 noninduced (lanes 1 and 4), induced with 1 mM IPTG (lanes 2 and 5) and of the His_6-Phy protein (lanes 3 and 6), partially purified by Ni^{2+} affinity chromatography. The asterisk refers to a cross-reactive species or to a degradation product of the His_6-phytase **(B)** SDS-PAGE (lanes 1 and 2) and immunoblot analysis (lanes 3 and 4) of the cell culture supernatants of the PY79 wild-type strain of *B. subtilis* (lanes 1 and 3) and of strain AH7666, bearing a Δ*phy* mutation (lanes 2 and 4).

Significant fluorescence was observed on purifed spores harbouring the CotG-Phy or OxdD-Phy fusions (Fig. 3A). As expected, no significant fluorescence was observed at the surface of spores of the wild-type strain PY79. Importantly, treatment with trypsin resulted in loss of the fluorescence signal. Treatment with trypsin also resulted in the complete loss of fluorescence from spores of a strain (AH2873; Table 1) expressing an *oxdD-gfp* fusion (data not shown). As an additional control, the primary antibody was omitted, to test for specific binding of the secondary antibody coupled with the fluorophore. No fluorescence was observed under these conditions. We then quantified the fluorescence signal for spores of strains SD48 and SD50, in comparison to wild-type PY79 spores. The results show that the average signal on SD48 (85 AU) and SD50 (91 AU) were ~3-fold higher compared to the baseline signal observed for PY79 spores (29 AU) (Fig. 3B, 4A). In spite of a similar average intensity for SD48 and SD50, 75% of SD48 spores displayed at least 2-fold more fluorescence than the baseline (> 60 AU), whereas 100% did so for SD50. Together, these results suggest that the OxdD-Phy fusion was more efficiently displayed at the spore surface.

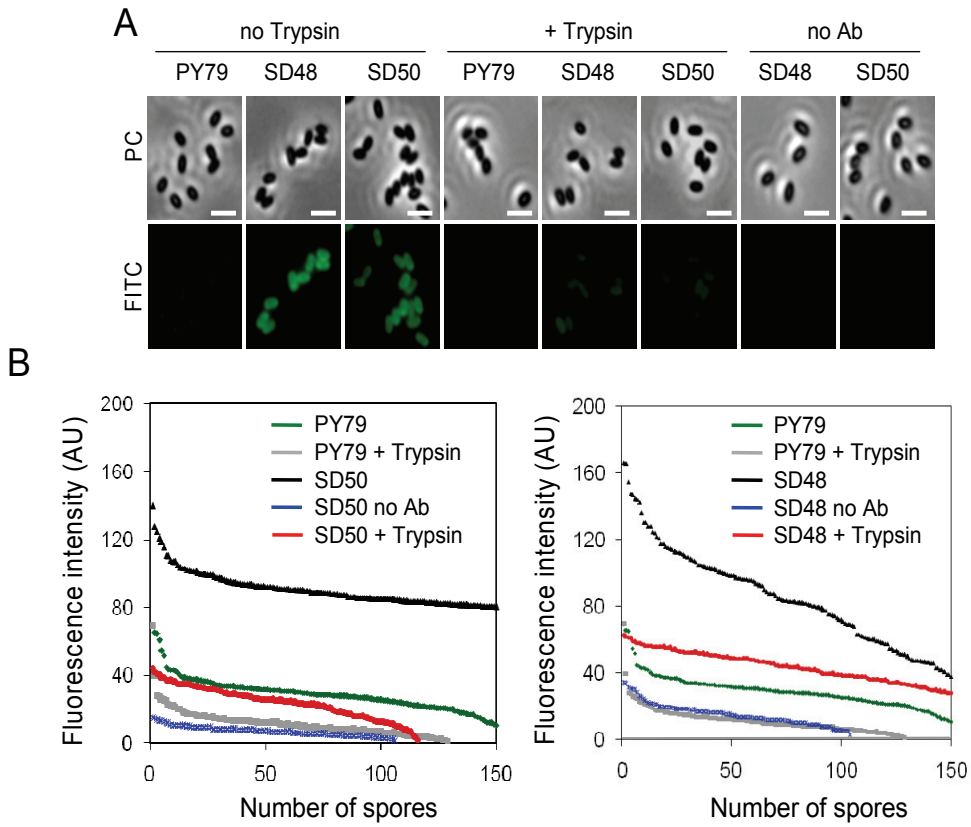


Figure 3. Immunofluorescence detection of CotG-Phy and OxdD-Phy protein fusions at the surface of purified spores. FITC, Fluorescein isothiocyanate **(A)** Phase-contrast microscopy (PC) and fluorescence microscopy (FITC). Scale bars, 2 μm. **(B)** Fluorescence intensity was quantified on 150 spores. AU, arbitrary units. Purified spores were submitted or not to trypsin proteolysis. noAb , control without primary antibody. Spores were purified from cultures of the following strains: PY79, wild-type; SD48, CotG-Phy; and SD50, OxdD-Phy.

Phytase activity at the surface of *B. subtilis* spores.

Having determined that both CotG and OxdD served as carriers for the display of Phy at the surface of *B. subtilis* spores, we then wanted to determine whether the passenger protein was presented in an active form. For this purpose, we conducted measurements of Phy specific activity. We first measured Phy activity in whole-cell extracts prepared from samples directly

collected from DSM cultures of strains SD48, SD50, and PY79, 24 hours after the onset of sporulation (Fig. 4B). Although a specific activity of 3.1 U/mg of protein was obtained for PY79, reflecting the endogenous *B. subtilis* phytase activity, values of 11.4 U/mg and 6.9 U/mg were obtained for SD48 and SD50, respectively (Fig. 4B). For reference, a partially purified His_6-tagged version of the *B. subtilis* phytase missing its signal peptide (Fig. 2A) gave a specific activity of 23 U/mg of protein when produced in *E. coli* (data not shown). The phytase activity detected for SD48 and SD50 extracts over the activity obtained for PY79 cultures clearly demonstrated expression of a functional phytase during sporulation, from the sporulation specific promoters that drive expression of either *cotG* or *oxdD*. Next, we thought to determine whether phytase activity could be detected on density gradient-purified spores - in other words, whether the CotG-Phy or OxdD-Phy fusion detected at the spore surface (see above) represented active enzyme (Fig. 4C). Since not all proteins are extractable from the spore coat, units of phytase activity in purified spores were normalized with respect to spore dry weight. Activities of 5.7×10^3 U/g (dry weight) and of 2.7×10^3 U/g (dry weight) were obtained for SD48 and SD50, respectively. No activity could be detected for PY79 spores or for spores of strains SD48 and SD50 that were treated with trypsin prior to the assay (data not shown). These results indicate that both the CotG or OxdD carriers resulted in the display of active phytase at the spore surface.

Although the OxdD-Phy fusion protein appeared more abundant than CotG-Phy at the spore surface, as assessed by immunofluorescence (see above), spores of SD50 (OxdD-Phy) showed a 2-fold-decreased specific activity compared to those of SD48 (CotG as the carrier). One possibility is that the OxdD-Phy

fusion protein, presumed to be mostly associated with the inner-coat layers, is less accessible to the substrate than is CotG-Phy, thought to be mainly located in the outer coat.

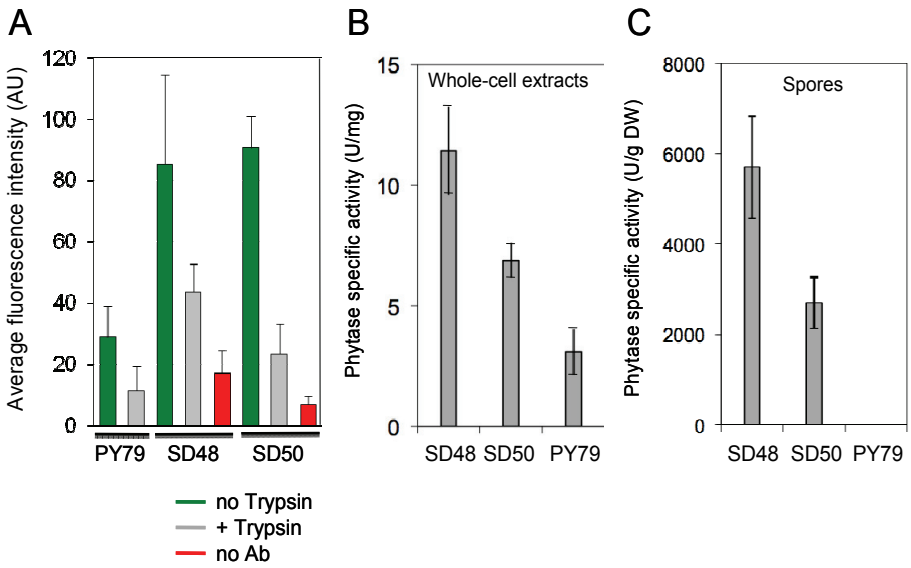


Figure 4. Abundance of the displayed phytase at the spore surface and specific activity of *B. subtilis* phytase. (A) Immunofluorescence detection of CotG-Phy and OxdD-Phy protein fusions at the surface of purified spores (average signal). **(B)** Assay of whole-cell extracts prepared from *B. subtilis* cultures in sporulating medium, 24 h after the onset of sporulation. **(C)** Assay of purified spores from the same cultures. DW, dry weight. Spores were purified from cultures of the following strains: PY79, wild-type; SD48, CotG-Phy; and SD50, OxdD-Phy. One unit (U) of phytase was defined as the amount of enzyme required to release one μmole of inorganic phosphate from sodium phytate in one minute.

Spore display of active *E. coli* β-glucuronidase.

The results presented above show that the inner-coat protein, OxdD, can be used as a carrier for the display of active, monomeric phytase at the spore surface. We then wanted to test whether OxdD could be used for spore display of a larger, oligomeric enzyme. We tested whether *E. coli* β-glucuronidase, encoded by the *uidA* gene, could be presented in active form at the

spore surface as a fusion to OxdD. The crystal structure of human GusA has been determined, and suggests that the enzyme is likely to function as a 273-kDa homotetramer [Cervera, 2005] with the catalytic site formed from a large cleft at the interface of two monomers [Jain *et al*., 1996]. Presumably, the *E. coli* enzyme, which shares 50% amino acid identity with the human version, is also a homotetramer [Flores and Ellington, 2002]. To attempt the functional display of *E. coli* GusA at the surface of *B. subtilis* spores, the *uidA* gene was fused to the 3´-end of the *oxdD* gene with, as for the *oxdD-phy* fusion (see above), an interspacing ten amino acid-linker (Fig. 1B). Ectopic integration of the gene fusion at *amyE* resulted in strain SD60 (Table 1). Expression of the OxdD-GusA fusion at the *amyE* locus of PY79 did not measurably interfere with spore function or assembly (Table 2 and data not shown).

Spores of strain SD60 were collected 24 h after the onset of sporulation in DSM, purified, and assayed for *in situ* detection of β-glucuronidase activity using the fluorogenic substrate C_{12}FDGlcU, (see Material and Methods). In parallel, activity assays were conducted in trypsin-treated SD60 spores, and in spores of the wild-type strain PY79. Cleavage of the colorless C_{12}FDGlcU substrate into 5-dodecanoylaminofluorescein results in the production of a stable green-fluorescent product, allowing β-glucuronidase activity to be monitored by fluorescence microscopy. Fluorescence was emitted by SD60 spores (Fig. 5A). In contrast, no fluorescence was detected for PY79 spores, as well as for SD60 or PY79 spores treated with trypsin prior to the assay (Fig. 5A). Quantification of the fluorescence signal (Fig. 5B, 5C) showed that 40% of SD60 spores emitted at least 2-fold more fluorescence than the baseline (>60 AU). Possible explanations for the occurrence of non-fluorescent spores might be that the fluoregenic product of cleavage diffuses

away from spores or that the high molecular weight (882 g/mol) and lipophilic nature (12-carbon aliphatic chain) of the fluorogenic substrate limits its penetration into the inner coat, where OxdD-GusA presumably resides. It can also be speculated that OxdD-GusA fusion has not assembled into an active form in non fluorescent spores. In any event, our results show that active GusA can be displayed at the spore surface as a fusion to the inner-coat protein OxdD.

Further works will address the optimization of strain engineering to increase activity of the passenger enzyme fused to OxdD. For instance, improvement might result from increased multi-copy driven expression, from expression in protease-deficient *B. subtilis* strains or through the selection of passenger enzymes with higher specific activity (*e.g.* phytase from *Citrobacter braakii*). Nevertheless, based on the current activity measured at the surface of spores displaying *B. subtilis* OxdD-Phy (2.7×10^3 U/g dry weight), about 0.4 g of spores (dry weight) would represent 1,000 U of phytase. This amount of enzyme, added to 1 kg of feed in a daily swine diet, would replace 1 g of inorganic phosphorus supplementation, and therefore could reduce total phosphorus excretion by 30 to 50% [Lei and Porres, 2003].

Based on the results presented here, spore-associated enzymes, such as OxdD, seem to be valuable carriers for spore display of passenger proteins. One question raised by our study is whether exposure of the spores to acid or proteases in the upper parts of the gastro-intestinal tract would result in the inactivation of the displayed carrier-enzyme fusion protein. This may indeed happen to some extent. However, *B. subtilis* spores were recently shown to have the ability to germinate and re-sporulate upon exiting the stomach, leading to persistence of *B. subtilis* strain in the gut

[Cartman *et al.*, 2008; Nguyen *et al.*, 2006] and to constant surface display of freshly synthesized and fully active enzymes. Spore surface display thus provides the basis for novel means of formulation for proteins of interest such as feed enzymes (phytases, lipases, hemicellulases and others).

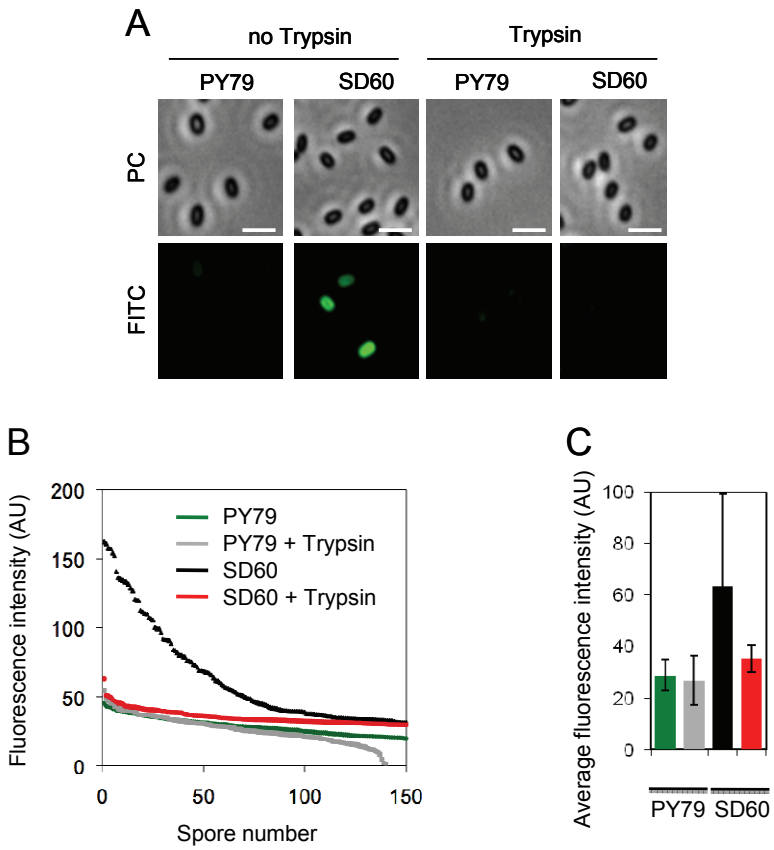


Figure 5. Assay for the activity of GusA at the spore surface, through conversion of the colorless substrate C_{12}FDGlcU (see Materials and Methods) into a yellow fluorescent product. (A) Detection of a functional OxdD-GusA fusion by phase-contrast microscopy (PC) and fluorescence (FITC) microscopy. Bars, 2 µm. **(B)** The fluorescence intensity was quantified on 150 spores. Purified spores submitted or not to trypsin proteolysis. Spores were purified from cultures of the following strains: PY79, wild-type; and SD60, OxdD-GusA. **(C)** Average fluorescence intensity quantified on 150 spores. AU: arbitrary units.

Acknowledgments.

This work was partially financed by grant PTDC/BIO/73946/2006 from Fundacão para a Ciência e Tecnologia (F.C.T.) to A.O.H. C.S. was the recipient of a doctoral fellowship (SFRH/BD/29397/06) from F.C.T. We thank Jochen Pheiffer and Antoine de Saizieu for their support and Zoltan Prágai, John B. Perkins and Charlie Moran Jr for helpful discussion and critical reading of the manuscript.

References

Antelmann, H., S. Töwe, D. Albrecht, and M. Hecker. 2007. The phosphorus source phytate changes the composition of the cell wall proteome in *Bacillus subtilis*. J. Proteome Res. 6(2):897-903.

Cartman, S.T., R.M.L. Ragione, and M.J. Woodward. 2008. *Bacillus subtilis* spores germinate in the chicken gastrointestinal tract. Appl. Environ. Microbiol. 74:5254-5258.

Cervera, M. 2005. Histochemical and fluorometric assays for uidA (GUS) gene detection. Methods Mol Biol. 286:203-214.

Choct, M. 2006. Enzyme for the feed industry: past, present and future. World's Poultry Sci J. 62:6-16.

Costa, T., A. L. Isidro, C. P. Moran Jr., and A. O. Henriques. 2006. The interaction between coat morphogenetic proteins SafA and SpoVID. J. Bacteriol. 188:7731-7741.

Costa, T., L. Steil, L.O. Martins, U. Völker, and A.O. Henriques. 2004. Assembly of an oxalate decaroxylase produced under the sigma k control into the *Bacillus subtilis* spore coat. J. Bacteriol. 186(5):1462-1474.

Cutting, S.M., and P.B. Vander Horn. 1990. Genetic Analysis. P53-54. In Harwood, C.R., and S.M. Cutting eds. Molecular biological methods for Bacillus. Modern Microbiological methods. J. Wiley and Sons ed., Chichester.

Driks, A. 2002. Maximum shields: the assembly and function of the spore coat. Trends Microbiol. 10(6):251-4.

Duc, L.H., H.A. Hong, H.S. Atkins, H.C. Flick-Smith, Z. Durrani, S. Rijpkema, R.W. Titball, and S.M. Cutting. 2007. Immunization against anthrax using *Bacillus subtilis* spores expressing the anthrax protective antigen. Vaccine. 25(2):346-355.

Flores, H., and A.D. Ellington. 2002. Increasing the thermal stability of an oligomeric protein, beta-glucuronidase. J Mol Biol. 315(3):325-337.

Guérout-Fleury, A.M., K. Shazand, N. Frandsen, and P. Stragier. 1995. Antibiotic resistance cassettes for *Bacillus subtilis*. Gene 167: 335-336.

Henriques, A.O., B.W. Beall, K. Roland, and C.P. Moran Jr. 1995. Characterization of *cotJ*, a sigma E-controlled operon affecting the polypeptide composition of the coat of *Bacillus subtilis* spores. J. Bacteriol. 177(12):3394-3406.

Henriques, A.O., T. Costa, L.O. Martins, and R. Zilhão. 2004. Functional architecture and assembly of the spore coat. In Bacterial spore formers: probiotics and emerging applications. Ricca, E., A.O. Henriques, and S.M. Cutting (eds.). London, UK. Horizon Sci. Press. 34-52.

Henriques, A.O., L.R. Melsen, and C.P. Moran Jr. 1998. Involvement of superoxide dismutase in spore coat assembly in *Bacillus subtilis*. J. Bacteriol. 180(9):2285-2291.

Henriques, A.O., and C.P. Moran Jr. 2007. Structure, assembly, and function of the spore surface layers. Annu Rev Microbiol. 61:555-588.

Hinc, K., R. Isticato, M. Dembek, J. Karczewska, A. Iwanicki, G. Peszynska-Sularz, M. De Felice, M. Obuchowski, and E. Ricca. 2010. Expression and display of UreA of Helicobacter acinonychis on the surface of *Bacillus subtilis* spores. Microb. Cell Fact. 982.

Imamura D., R. Kuwana, H. Takamatsu, and K. Watabe. 2010. Localization of proteins to different layers and regions of *Bacillus subtilis* spore coats. J. Bacteriol. 192:518-24.

Isticato, R., G. Cangiano, H.T. Tran, A.L. Ciabattini, D. Medaglini, M.R. Oggioni, M. de Felice, G. Pozzi, and E. Ricca. 2001. Surface display of recombinant proteins on *Bacillus subtilis* spores. J. Bacteriol. 183(21):6294-6301.

Isticato, R., G. Esposito, R. Zilhão, S. Nolasco, G. Cangiano, M. de Felice, A. O. Henriques, and E. Ricca. 2004. Assembly of multiple forms of CotC into the *Bacillus subtilis* spore coat. J. Bacteriol. 186:1129-1135.

Jain, S., W.B. Drendel, Z.W. Chen, F.S. Mathews, W.S. Sly, and J.H.Grubb. 1996. Structure of human beta-glucuronidase reveals candidate lysosomal targeting and active-site motifs. Nat Struct Biol. 3(4):375-381.

Jefferson, R.A., S.M. Burgess, and D. Hirsh. 1986. Beta-Glucuronidase from *Escherichia coli* as a gene-fusion marker. Proc Natl Acad Sci U S A. 83(22):8447-8451.

Karow , M.L., and P.J. Piggot. 1995. Construction of gusA transcriptional fusion vectors for *Bacillus subtilis* and their utilization for studies of spore formation. Gene. 163(1):69-74.

Kerovuo, J., M. Lauraeus, P. Nurminen, N. Kalkkinen, and J. Apajalahti. 1998. Isolation, characterization, molecular gene cloning, and sequencing of a novel phytase from *Bacillus subtilis*. Appl Environ Microbiol. 64(6):2079-2085.

Kerovuo, J., I. Lappalainen, and T. Reinikainen. 2000. The metal dependence of *Bacillus subtilis* phytase. Biochem Biophys Res Commun. 268(2):365-369.

Kim, J.H., B.G. Kim, S.K. Choi, H.C. Jung, and J.G. Pan. 2004. Methods for expression of proteins on spore surface. US patent 7,582,426.

Kim, J.H., C.S. Lee, and B.G. Kim. 2005. Spore-displayed streptavidin: a live diagnostic tool in biotechnology. Biochem Biophys Res Commun. 331(1):210-214.

Kim, T.W., and X.G. Lei. 2005. An improved method for a rapid determination of phytase activity in animal feed. J Anim Sci. 83:1062-1067.

Kwon, S.J., H.C. Jung, and J.G. Pan. 2007. Transgalactosylation in a water-solvent biphasic reaction system with beta-galactosidase displayed on the surface of *Bacillus subtilis* spores. Appl Environ Microbiol. 73(7):2251-2256.

Lei, X.G., and J.M. Porres. 2003. Phytase enzymology, applications, and biotechnology. Biotechnology Letters. 25:1787-1794.

Mauriello, E., L.H. Duc, R. Isticato, G. Cangiano, H.A. Hong, M. de Felice, E. Ricca, and S.M. Cutting. 2004. Display of heterologous antigens on the *Bacillus subtilis* spore coat using CotC as a fusion partner. Vaccine. 22:1177-1187.

Nguyen, K.M.T., Q.U. Nguyen, A.H. Huynh, L.H Duc, T.H. Tran, C.R. Serra, A.O. Henriques, and S.M. Cutting. 2006. The intestinal life cycle of *Bacillus subtilis* and close relatives. J. Bacteriol. 188(7):2692-2700.

Nicholson, W.L., and P. Setlow. 1990. Sporulation, Germination and Outgrowth, p391-450. In Harwood, C.R., and S.M. Cutting eds. Molecular biological methods for Bacillus. Modern Microbiological methods. J. Wiley and Sons ed., Chichester.

Ozin, A.J., C.S. Samford, A.O. Henriques, and C.P. Moran Jr. 2001. SpoVID guides SafA to the spore coat in *Bacillus subtilis*. J. Bacteriol. 183(10):3041-9.

Rao, D.E., K.V. Rao, T.P Reddy, and V.D. Reddy. 2009. Molecular characterization, physicochemical properties, known and potential

application of phytases: an overview. Crit Rev Biotechnol. 29(2):192- 198

Ricca, E., and S.M. Cutting. 2003. Emerging applications of bacterial spores in Nanobiotechnology. J.Nanobiotechnology. 1(1):6.

Shimizu, M. 1992. Purification and characterization of phytase from *Bacillus subtilis* (natto) N-77. Biosci. Biotech. Biochem. 56(8). 1266-1269.

Tanner, A., L. Bowater, S.A. Fairhurst, and S. Bornemann. 2001. Oxalate decarboxylase requires manganese and dioxygen for activity. Overexpression and Characterization of *Bacillus subtilis* YvrK and YoaN. J Biol Chem. 276(47):43627-43634.

Walch, A. 1996. PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in S. cerevisiae. Yeast 12.259-265.

Yim, S.K., H.C. Jung, C.H. Yun, and J.G. Pan. 2009. Functional expression in *Bacillus subtilis* of mammalian NADPH-cytochrome P450 oxidoreductase and its spore-display. Protein Expr Purif. 63(1):5-11.

Zilhão, R, M. Serrano, R. Isticato, E. Ricca, C.P. Moran Jr, and A.O. Henriques. 2004. Interactions among *CotB, CotG,* and *CotH* during assembly of the Bacillus spore coat. J. Bacteriol. 186:1110-19.

CHAPTER V

Summary and Perspectives

During the last decades, industrial microbial processes, known as “white biotechnology”, have been developed and have become economically competitive compared to the chemical processes they have replaced. White biotechnology refers to the production of biochemicals (food/feed ingredients, fine chemicals, pharmaceuticals), biomaterials and biofuels of commercial interest from renewable resources (generally agricultural by-products). White biotechnology relies on enzymes derived from microorganisms (biocatalysis) and on living cells as factories (fermentation technology). In the area of industrial microbiology, microorganisms used (bacteria, yeast, or fungi) may be natural isolates, selected mutants or genetically engineered organisms. Microbial production processes are of an ecological interest compared to traditional chemical processes, as they consume less raw materials, water and energy, enabling economies to become less dependent on fossil fuels. Finally, use of expensive substrates and optically resolved racemic intermediates, which are required for chemical processes, can be avoided.

Bacillus subtilis is a key player among the favorite microorganisms for industrial biotechnology. In the soil, its principal environment*, B. subtilis* is able to secrete degradative enzymes that release nutrients from macromolecules which are too large to be directly taken up. This innate capacity to secrete proteins makes of this organism an efficient microbial factory for high-yield production of commercial enzymes or metabolites such as vitamins. Moreover, the safety of *B. subtilis* and its ability to form resistant endospores constitute major advantages to use this species for probiotic applications. Overproduction of the desired metabolites/enzymes, and the development of innovative products to improve the daily life of consumers by ensuring sustainable profits, involve important

efforts in applied microbiology research of private companies. More specifically, elucidating the molecular basis of microbial physiology is crucial to either modulate the control of metabolic pathways or divert natural physiological features of the microorganisms toward new commercial applications. The results discussed in this chapter illustrate both of these aspects.

1. Toward a *B. subtilis* industrial workhorse for the production of vitamin B1

Microbial processes using *B. subtilis* have been developed and implemented for the industrial production of vitamin B2 (riboflavin). A commercially relevant production yield of riboflavin results from the intensive metabolic engineering of a wild-type strain, otherwise secreting negligible amount of this product (collaboration between Omnigene Bioproducts Inc., USA and Roche Vitamins Ltd., Switzerland). This successful example of biotech-based industrial production opened the way toward the development of *B. subtilis* strains able to secrete high amounts of other water-soluble vitamins. The market for chemically produced thiamin currently consists of several tons per year, for feed (40%), pharma (40%) and food (20%) applications [Perkins *et al.*, 2009].

Vitamin B1 (thiamin) is a nutritional requirement for mammals, being a cofactor for several enzymes in the metabolism of carbohydrates and aminoacids (*e.g.* pyruvate dehydrogenase, pyruvate oxidase and transketolase). Thiamin is synthesized by many microorganisms, plants, and fungi, but is not produced by vertebrates. No complete “glucose to thiamin” microbial process has been described so far. The biologically active form of thiamin is thiamin pyrophosphate (TPP), which represents 98% of the intracellular thiamin content. Biosynthesis of thiamin in *B. subtilis* appears to use the same enzymes and intermediates as in *E. coli* [Perkins and Pero, 2001].

Thiamin biosynthesis results from the assembly of a pyrimidine moiety (HMP-PP: hydroxymethylpyrimidine pyrophosphate), derived from the purine biosynthetic pathway (synthesis of HMP involves the product of *thiC* gene), and a thiazole

moiety (HET-P: hydroxyethylthiazole phosphate pyrimidine), derived from the condensation of aminoacid residues (glycine, cysteine) with an intermediate of the isoprenoid biosynthesis pathway (DXP: 1-deoxy-d-xylulose phosphate). The coupling of HMP-PP and HET-P results in thiamin monophosphate (TMP), which is then phosphorylated into TPP. This reaction is catalyzed by a thiamin pyrophosphokinase encoded by *thiL* gene.

TPP biosynthesis follows a complex multistep pathway with numerous enzymatic reactions and involvement of cofactors, and is associated with large energy expenditure. Therefore, as an alternative to the *de novo* biosynthesis, bacteria have developed ways to salvage residues of thiamin degradation in the environment, released from the decay of various organisms in the soil. These residues uptaken by the bacteria, converted to new building blocks, and are recycled in the thiamin biosynthetic pathway. Figure V-1 summarizes the key genes of the thiamin biosynthesis and the salvage pathways previously known or discovered in our studies.

Salvage pathways described in the past occur either at the level of the pyrimidine moiety with ThiD kinase [Park *et al.*, 2004], of the thiazole moiety with ThiM kinase [Melnick *et al.*, 2004], or of thiamin. During thiamin salvage, thiamin kinase catalyzes the formation of TMP, subsequently converted into TPP by ThiL. Alternatively to this two-step pathway, the product of *yloS* (thiamin pyrophosphokinase) catalyzes the direct conversion of thiamin to TPP [Melnick *et al.*, 2004].

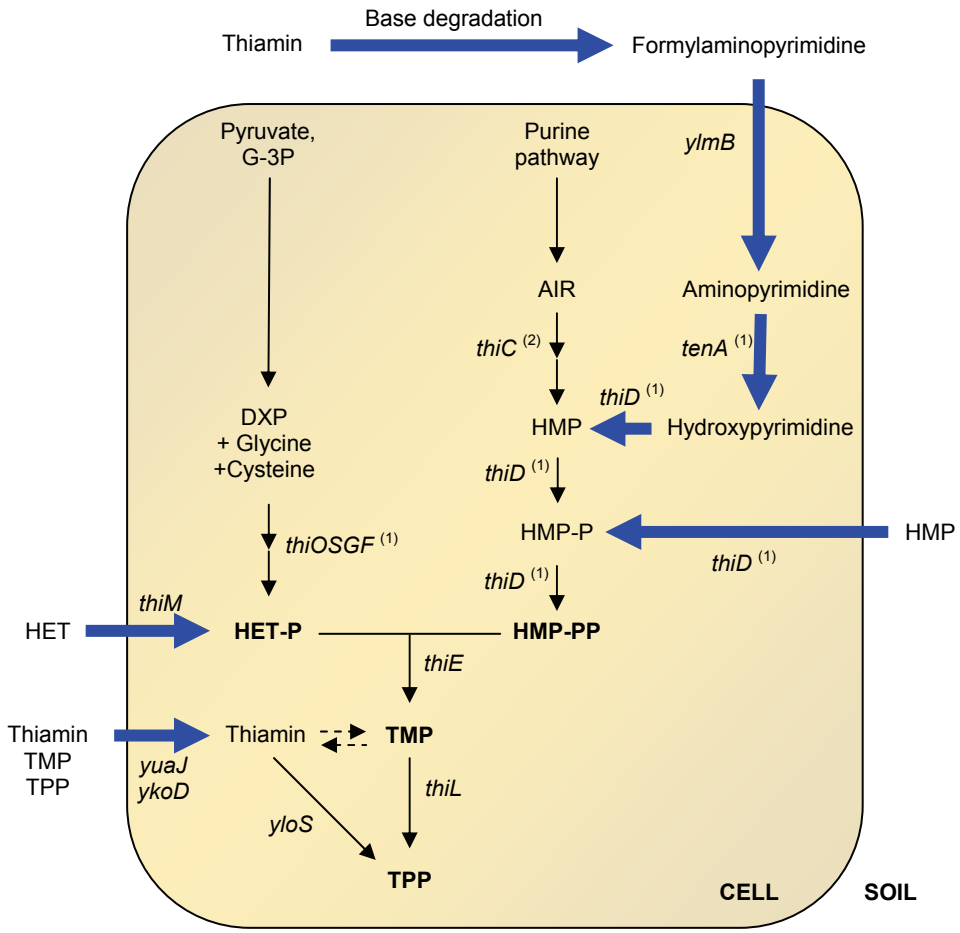


FIGURE V-1: Simplified thiamin biosynthesis and salvage pathways in *B. subtilis.* G-3P (glyceraldehyde-3-phosphate); DXP (1-deoxy-d-xylulose phosphate); AIR (aminoimidazole ribotide); (1) gene organized in *thiB* operon (*tenAI-thiOSGFD*); (2) *thiA* synonym was used in chapter II, referring to *thiC*; black arrows (biosynthetic pathways); blue arrows (salvage pathways)

Chapter II illustrates how the discovery of new metabolic intermediates at the crossroad of thiamin degradation and biosynthesis constitutes the first step in the identification of unknown salvage pathways. The previously reported role of TenA (thiaminase II) in thiamin degradation into HMP and HET [Toms *et al.*, 2005] was difficult to rationalize, since *tenA* clusters with genes

of the thiamin biosynthetic pathway (*thiB* operon) in *B. subtilis*. The gene organization in *Bacillus halodurans* reveals an association of *tenA* with the uncharacterized *ylmB* locus and with *thiXYZ*, which was shown, by comparative genomics of thiamin biosynthesis within prokaryotes, to encode an ABC transporter involved in the transport of HMP analogs [Rodinov *et al.*, 2002]. These elements led to reconsidering the role of TenA, and to propose a new model in which TenA could be involved, together with YlmB and ThiXYZ, in a salvage pathway from thiamin-degraded products of the soil.

In order to validate this hypothesis, environmental chemistry experiments based on HPLC analyses of soil solutions containing degradation products resulted in the identification of a pyrimidine analog (formylaminopyrimidine) which was shown to bind to ThiY. Formylaminopyrimidine appeared to be converted into HMP when the soil solution was supplemented with recombinant purified YlmB and TenA proteins. More specifically, our data suggested that a residue of thiamin degradation in the soil, released by the decay of various organisms in the environment, could be salvaged when taken up by bacteria through its interaction with the ThiXYZ ABC transporter. YlmB catalyzes then its deformylation into aminopyrimidine, which is then hydrolyzed by TenA into hydroxypyrimidine. A detailed mechanism for the TenA-catalyzed amino-pyrimidine hydrolysis was recently proposed [Jenkins *et al.*, 2008]. Further phosphorylation of hydroxypyrimidine by ThiD results in HMP, supplied to thiamin biosynthetic pathway.

TenA is not essential for thiamin biosynthesis, as *tenA* mutants are able to grow in minimal medium. On the contrary, *thiC* mutants display thiamin auxotrophy in minimal medium. Complementation analyses (with either thiamin, HMP, aminopyrimidine, formylaminopyrimidine, or based-degraded

thiamin) on *B. subtilis thiC* and/or *tenA* mutants led to the observation that TenA enables growth of *thiC* mutant in presence of thiamin degradation products. These results confirm the hypothesis of a thiamin salvage pathway from thiamin-degraded residues which involves TenA and results in supplying HMP to the biosynthetic pathway, independently from *de novo* HMP synthesis catalyzed by ThiC activity.

In light of the results presented in **chapter II**, one can speculate that similar salvage pathways do exist for other vitamins of the B group, or even for other metabolites of commercial interest. This surely opens new research directions for the identification of unknown salvage pathways in *B. subtilis*, broadening the possibilities of optimization by metabolic engineering to bring vitamins production levels up to industrial relevance.

Work toward the generation of a *B. subtilis* strain for relevant vitamin B1 industrial production is described in **chapter III**. Mutagenesis studies allowed identifying and characterizing new thiamin-deregulated mutations in *B. subtilis*. Regulation of thiamin biosynthetic genes has been reported to be based on the direct and specific binding of TPP to the 5'-untranslated regions of messenger RNAs (*thi* box). This binding promotes conformational changes that affect the expression of downstream genes [Winkler *et al.*, 2002]. This riboswitch mechanism is highly conserved within thiamin producing microorganisms. Aiming to provide data for the reconstruction of the thiamin metabolic pathway and its regulation, comparative genomic analyses of the *thi* box sequence in all available bacterial genomes resulted in the identification of new putative thiamin-related loci, including the hypothetical transporters YuaJ and YkoEDC. *yuaJ*, predicted to encode a passive thiamin-permease, is the only thiamin-regulated gene in several organisms

known to miss the thiamin biosynthesis pathway [Rodionov *et al*., 2002], suggesting that it is involved in thiamin uptake. *ykoD*, predicted to encode a part of an ATP-dependant ABC transporter, was linked to the thiamin salvage system based on the positional clustering of the *ykoEDC* genes with other thiamin salvage genes [Rodionov *et al.*, 2002]. Until the work described in this manifest, no genetic experiments had confirmed the role of these genes in thiamin transport.

Ethyl methanesulfonate-induced mutants of *B. subtilis* deregulated for thiamin production were selected on the basis of the activity of a *lacZ* transcriptional fusion with the *thiC* biosynthetic gene in the presence of TPP. Screening was performed in a Δ*thiL* mutant in order to exclude *thiL* mutants from the pool of $LacZ^+$ thiamin-deregulated clones, and focus on mutations related to exogenous thiamin uptake (*i.e.* salvage pathway). Interestingly, Δ*thiL* mutants are not strictly auxotroph unlike in *E. coli or Salmonella enterica.* This supports the hypothesis of an alternative way to produce TPP in *B. subtilis*, distinct from the biosynthetic pathway. Based on their unique phenotype, two mutants were selected for further characterization. Clone Tx1 has a bradytroph growth (slow on minimal medium without supplementation in thiamin or its derivatives), and harbored one mutation in the *yloS* locus (L116F). The mutation leads to a reduced pyrophosphokinase activity, since the addition of external TPP or thiamin can restore a normal prototroph phenotype. This result supports the direct thiamin to TPP route. Clone Tx26 grows like the parental strain, is resistant to a toxic analog of thiamin (pyrithiamin) and has two remarkable mutations respectively in *yuaJ* (Q35*; predicted to result in a truncated inactive protein) and *ykoD* genes (D180N). Transcriptomics profiling of mutant Tx26 (*yuaJ*, *ykoD*) showed a

deregulated expression of all genes known to contain a *thi*-box (*i.e. thiC*, *thiOSGFD*, *ykoFEDC*, *yuaJ*). This deregulation was confirmed to be independent from Δ*thiL*. Resistance to pyrithiamin and to higher levels of external thiamin suggested that the two Tx26 mutations control together the uptake of thiamin and of its precursors rather than act as efflux transporters. Combination of the four mutations (*yloS*, *yuaJ*, *ykoD* and Δ*thiL*) in the same *B. subtilis* host resulted in a strain (TH95) able to excrete thiamin in fermentation (ten-liter scale) in the mg per liter range.

Although the expected fermentation yield could not be reached so far to develop a commercially relevant biotechnological process for vitamin B1 production (range of gram per liter being required), results presented in **chapter III** constitute a significant step forward through the isolation and characterization of four mutations in thiamin biosynthetic and salvage genes. Improvement of the basic knowledge of the biosynthetic pathway of thiamin is crucial to move forward to a commercially relevant strain.

A possible way of improvement could consist in increasing the pool of thiamin precursors in the biosynthetic pathway. To test this hypothesis, fermentation experiments of strain TH95 were performed by co-feeding both HMP and HET precursors (0.54 g/l each), leading to a 27-fold increase in thiamin production. A putative bottleneck in the thiamin pathway could also be the coupling reactions between both HMP and HET moieties. In a new TH95-based strain, expression of *thiE* was driven by a strong constitutive promoter derived from *Bacillus* bacteriophage SP01 (SP01-26). Co-feeding HMP and HET in fermentation led to about 40-times more thiamin than TH95. These results suggested that our deregulated strain has a large capacity to synthesize TPP by coupling the

exogenous-added moieties, to dephosphorylate TPP to thiamin, and to excrete thiamin in the growth medium.

Intracellular supply of either HET or HMP precursors was then addressed. Two mutants were generated using SP01-26 promoter, in order to respectively increase the pool of HET (*thiB* operon; *i.e tenAI-thiOSGFD*) or HMP (*thiC*) in the biosynthetic pathway. Fermentation, with either HMP or HET supplementation was performed. Both mutants resulted in a significantly increased thiamin production compared to TH95. However, the *thiC* mutant (HMP supply) produced significantly less thiamin than the *thiB* operon mutant (HET supply). This result suggests a limitation in the formation of HMP in the pathway, possibly due to a low level of ThiC substrate (AIR: aminoimidazole ribotide) or an insufficient additional activity. To confirm this hypothesis, the pool of AIR available for ThiC was increased in TH95 (the purine operon was deregulated through a mutation in *purO*, and the conversion to the carboxy-aminoimidazole ribotide side product was blocked through a mutation in the *purE* locus). Fermentation of the resulting strain with HET supplementation led to more thiamin than TH95 SP01-26 *thiC*. This result supports the hypothesis that HMP formation is a rate-limiting factor. Possibly, another unknown gene (not located in the *thiB* operon) may be involved in HMP biosynthesis and would need to be overexpressed to improve the precursor supply. In any event, our data suggest that increasing the pool of precursors in the biosynthetic pathway successfully improves the production of thiamin in *B. subtilis* strain TH95, even though it does in a limited manner (< 10-fold).

Increase of the HMP pool in the cell could for instance involve genetic engineering of the new salvage pathway described in **chapter II**, or the previously described ThiD kinase pathway

[Park *et al.*, 2004]. The current emergence of metagenomics could also help circumventing the limitation of HMP in the thiamin biosynthetic pathway by finding new biological chemistry to generate HMP.

2. *B. subtilis* for innovative probiotic applications

Probiotics can display growth-promoting, prophylactic or therapeutic attributes. As they also improve weight gain in modern farming, they represent an alternative to antimicrobial drugs commonly employed as growth-promoter to broilers, avoiding drug residues to accumulate in the animal, and therefore in the food meat. When ingested in sufficient amount, probiotics are believed to play an important role in the control of host intestinal microbiota. However, understanding the molecular processes underlying the host-microbe interactions constitute a special challenge, mostly because of the complexity and diversity of the microbial community inhabiting the gastrointestinal tract.

Although their mechanism of action remains poorly understood, spores from several *Bacillus* species are currently used as human and animal probiotic. Criteria to screen strains targeting probiotic application include functional criteria such as the ability to resist environmental conditions found in the digestive tract (low gastric pH and bile salts), to antagonize or competitively exclude pathogens by secretion of antimicrobial compounds or to compete with pathogenic organisms for nutrients and adhesion sites. The use of *B. subtilis* as probiotic relies on its capability to sporulate, because spores can survive to the harsh conditions of gastric juices before reaching the intestine where they can reactivate their germination/sporulation cycle pattern. Proficient use of selected strains isolated from intestinal microbial ecosystems has led to a variety of microbiological products. Therefore, the screening of undomesticated gut-associated bacterial strains that evolved to survive and proliferate in the intestine environment constitutes a source of potential new probiotic candidates. Such a screening was

for instance performed for *Bacillus* isolates in the broiler gastrointestinal tract of poultry raised in organic farms. Interesting isolates exhibited high sporulation ability and antimicrobial activity against food spoilage and pathogenic organisms such as *Clostridium perfringens*, *Staphylococcus aureus* and *Listeria monocytogenes* [Barbosa *et al.*, 2005]. Assuming that these *B. subtilis* strains are deprived of any toxicity and that proof of their safety toward their host can be made, they could be considered to be part of a probiotic product to be used in animal farming. Furthermore, such isolates have not been engineered, which would give them a non-GMO status, in agreement with current regulation.

Social acceptance of GM foods or ingredients is not uniform in developed countries, and is especially difficult in Europe. The introduction of new probiotics based on GM strains would surely raise safety concern. The main issues with GMOs, and above all those of human origin, is the evaluation of the risk for human health of uncontrolled product expression following horizontal transfer of recombinant genes into commensal bacteria. A strain genetically engineered to introduce foreign DNA or even create a punctual mutation is considered as a GMO. The paradox is that a similar mutant occurring naturally (*e.g.* replication errors, plasmid loop-out, transposons, or phage-elements) or generated by classical strain improvement methods (*e.g.* chemical mutagenesis) might not be classified as GMO, even though it is less characterized.

Even if GM probiotics are not yet authorised on the market, research on these potential products for the future should go forward. These new strains would have a great potential when used to maintain health, prevent or treat disease, or be of nutritional value. Because of their intrinsic properties, probiotic spores could be of interest to offer long-shelf stability, avoiding costly formulation

process (*e.g* encapsulation in a polymer matrix, such as alginate or chitosan) to keep beneficial microorganisms alive until they reach the intestine. In this scope, spore display of bioactive molecules with a commercial interest (*e.g.* feed enzymes) could constitute a major advance for their *in situ* delivery in the host intestine. **Chapter IV** provides preliminary data for the development of innovative recombinant probiotics using *B. subtilis* spore display.

The spore display concept was introduced for the first time almost a decade ago [Isticato *et al.*, 2001]. So far, research on spore display was turned toward pharmaceutical applications (vaccinology), as most of the reported studies addressed the presentation of antigenic molecules at the surface of *B. subtilis* endospores in order to promote an immune response after being inhaled or orally administrated. All the successful published examples use anchoring motifs which are abundant structural proteins of the outermost layer of the spore coat (CotB, CotC, CotG). Among the 80 polypeptides of the coat, about 20 show an enzymatic activity or have a sequence highly similar to known enzymes [Henriques and Moran, 2007]. The work presented in **chapter IV** explores the use of one of these coat-associated enzymes, OxdD, as carrier protein.

Although its association with the spore coat still remains unclear, the product of the *yoaN* gene was characterized as an oxalate decarboxylase, targeted in a SafA-dependant manner to the inner coat [Costa *et al.*, 2004]. In order to demonstrate that this original anchoring motif can be used to display bioactive passenger molecules, we first analyzed spores from a culture of a recombinant strain harbouring a fusion of OxdD carrier with *B. subtilis* phytase (Phy) which functions as a monomer. As a positive control, and for comparison, a similar fusion was also made with CotG carrier,

successfully used in several published examples. Results of our immunofluorescence experiments (anti-Phy antibody) indicated that OxdD is a potential anchoring motif for the display of passenger phytase at the spore surface. Importantly, spore-exposed phytase, with one or the other fusion, displays activity (in the range of 10^3 U/g of gradient-purified spores). To further challenge our innovative system, spore display of a fusion between the OxdD carrier and a large oligomeric enzyme was performed. The β-glucuronidase (GusA) of *E. coli* encoded by the *uidA* locus was selected as passenger. Fluorescence microscopy observations on gradient-purified spores harbouring the OxdD-GusA fusion enabled the detection of a green fluorescent compound resulting from the specific hydrolysis of a colourless substrate by GusA. *In situ* activity of the displayed GusA enzyme was therefore clearly confirmed. These successful spore display examples of both a bioactive endogenous monomeric phytase and a heterologous oligomeric β-glucuronidase demonstrate the potential of coat associated oxalate decarboxylase located in the inner coat to be used as anchoring motif [Schyns *et al.*, 2008].

Proof-of-concept described in **chapter IV** open a new way in the design of recombinant spores. Spore display using OxdD requires however to be better characterized and optimized. Interestingly, not all the spores displaying OxdD-GusA were fluorescent. In addition, a 2-fold lower specific phytase activity was observed with OxdD-Phy compared to CotG-Phy. These discrepancies are thought to be due to the position of the OxdD fusion in the inner layer of the coat, being therefore less accessible to either of the substrates. Another explanation could be that OxdD is a less abundant protein than CotG (present in the range of 10^3 copies per spore). This hypothesis is however not consistent with

our immonufluorescence results, where the OxdD-Phy fusion is more abundant than the CotG-Phy. This could be explained by the relatively lower importance of OxdD in the coat structure compared to CotG, a key player of the coat protein assembly. In a *cotG* mutant, the coat structure was shown to be affected, whereas the disruption of the *oxdD* locus had no detectable effect on spore resistance to lysozyme, reflecting a well structured coat [Costa *et al.*, 2004; Henriques and Moran, 2007]. The OxdD-Phy fusion may preferentially be assembled in the coat during spore formation compared to CotG-Phy, because wild-type (non-fused) CotG, that is required in the coat network formation, could compete with the assembly of CotG-Phy. Using a non-structural protein as anchoring motif for spore display, rather than a protein involved in the coat formation, could thus be seen as an advantage. To address this question, more experiments are required, such as for instance the analysis of spores produced from strains with *cotG* deletion or bearing multicopies of *oxdD-phy* fusions in their genome, or in a replicative plasmid (*e.g.* pMK3 shuttle vector, GenBank EU549779).

OxdD is presumed to be a homohexameric enzyme, as suggested by its high similarity to the OxdC molecule crystal whose structure was resolved [Anand *et al.*, 2002]. To improve the efficiency of spore display of a bioactive passenger, it can be hypothesized that the degree of polymerization of OxdD in the fusion is an important determinant of the steric conditions leading to the optimal folding of the displayed enzyme. A possible approach to figure out the degree of multimerization would consist in the expression of the fusion protein in the cytoplasm of *B. subtilis*, preventing its assembly in the spore coat (for instance by driving expression with the promoter of a gene specific of the vegetative state). SDS-PAGE under reducing and non-reducing conditions or

gel filtration analysis could then be carried-out. Information obtained from such experiments could support the selection of a less oligomerized carrier protein of the coat for spore display purpose.

Recent identification of small open-reading frames in intergenic regions of the *B. subtilis* genome could also provide new insights in the development of innovative spore display systems [Schmalish *et al.*, 2010]. Using high-density microarrays covering all intergenic regions larger than 50 nucleotides, the study focused on unknown genes, transcriptionally active during sporulation. The authors reported that two predicted products show a high similarity to the respective products of *ydgB* and *ydzH* genes. Analysis of specific mutants revealed that these genes might be expressed under the control of the late-appearing, mother cell specific sigma K factor. Fluorescence microscopy experiments with GFP fusions gave evidence that YdgB and YdzH localize around the forespore, suggesting that they could be components of the spore coat. These results provide two unexplored spore-associated proteins to be tested in the future as original anchoring motifs for spore display applications. They could be of particular interest due to their low molecular weight (< 10 kDa), which could prevent possible interferences in the folding of the fused passenger protein.

3. Conclusion and future directions on *B. subtilis* research for the biotechnology industry

The studies in this thesis illustrate how exploration of the metabolic pathways involved in the biosynthesis of vitamin B1 contributes to the development of an industrially relevant *B. subtilis* overproducer (**chapters II and III**). The other application of *B. subtilis* genetic engineering (**chapter IV**) gives new input toward the design of an innovative spore-display system, using for the first time a coat-associated enzyme (oxalate decarboxylase) of the inner-coat as anchor. This work also constitutes a successful example of display for enzymes exhibiting activity on spores, which could in particular open perspectives for future feed additives and probiotics. Progresses in the use of the spore as display system for useful enzymes will strongly depend on a better understanding of the molecular mechanisms involved in the assembly pathway of particular carrier proteins. Another topic to explore is the influence of the peptide linker between carrier and passenger proteins (*i.e.* length, flexible *versus* helical structure) on the assembly and activity of the fusion proteins at the spore surface.

Being one of the preferred organisms for the industrial production of secreted enzymes and fine chemicals, and a very popular Gram-positive model for fundamental research, *B. subtilis* has been extensively engineered with a set of diverse and efficient genetic tools which have been developed in the course of the last fifty years. Future developments of *B. subtilis* biotechnological applications involve obviously the combination of classical genetic engineering with recent powerful technologies that allow ultra fast generation and/or identification of genetic diversity.

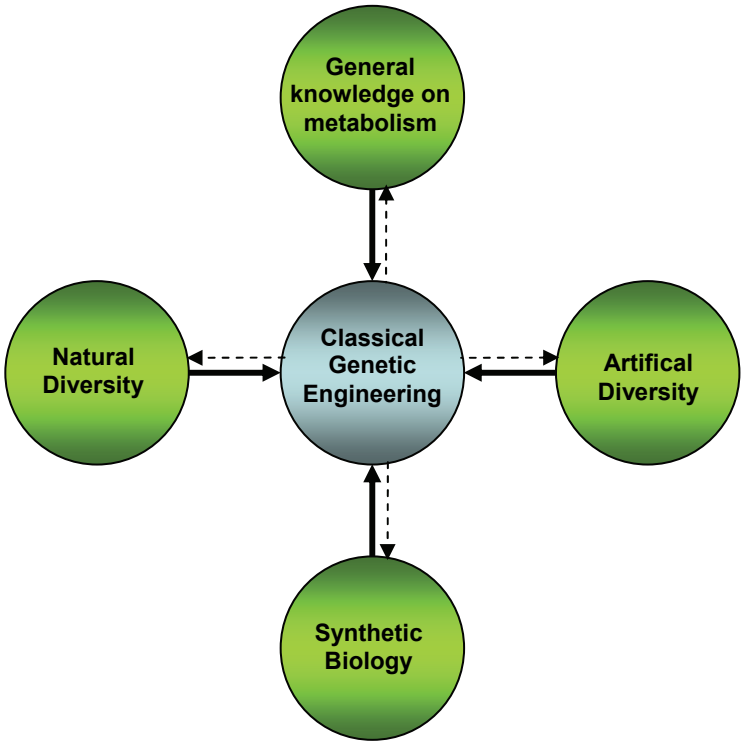


FIGURE V-2: Future perspectives in biotechnology research with *B. subtilis*. "Classical genetic engineering" gathers the strategies and techniques developed in the course of the last fifty years for *B. subtilis* strain improvement (as described in chapter I, 3). This current knowledge is supplied (broken arrows) in emerging research areas (green bullets). Further information from the complex metabolisms, synthetic biology and the selection of well-adapted strains for the biotechnology industry resulting from the exploration of the natural and artificial diversities, may constitute a new input (solid arrows) for further genetic engineering.

Random and directed mutageneses currently stay the most efficient concepts available in number of cases for metabolic engineering. **Chapters II and III** of this thesis demonstrate once again that engineering the metabolic pathways needs a massive and specific knowledge to successfully rechannel metabolic fluxes. The growing recognition of a higher degree of complexity in *B. subtilis* metabolism is a step toward solving this well-defined problematic, which is a prerequisite to overrun the metabolic

limitations. Several examples demonstrate this high level of complexity. For instance, little is understood about the regulatory networks (*e.g.* non-transcriptional mechanisms such as metabolite-protein interaction, protein phosphorylation *etc.*) enabling the metabolism to take control of itself [Heinemann and Sauer, 2010]. Another poorly explored area of metabolic pathways was lately pointed out. For the first time, the critical interactions between the glycolytic enzymes (6-phosphofructokinase, enolase and phosphoglycerate kinase) and essential protein partners were introduced. In addition, some of the glycolytic enzymes (6-phosphofructokinase and enolase) might interact with proteins involved in RNA metabolism, resulting in a complex equivalent to the *E. coli* degradosome [Commichau *et al.*, 2009]. These examples highlight the requirement for a multi-dimensional view toward a successful genetic engineering of metabolic pathways. Although mathematical modelling tools constantly progress, difficulties to integrate the large data-sets produced by “omics” technologies (transcriptomics, proteomics, metabolomics, fluxomics), and to align them with the global regulation of gene transcription also remain.

New research work streams recently emerged to solve these issues. The development of high-throughput genomics tools and new computational methodologies for the quantitative interpretation of the data is the topic of the *BaSysBio* (*Bacillus* Systems Biology) project, launched in 2006 (www.basysbio.eu). The elaboration of *B. subtilis* strains fitting industrial conditions is directly related to industrial research. The current program of genome minimalization in the scope of the *Basynthec* (Bacterial Synthetic Minimal Genomes for Biotechnology) project, aims at reducing the bacterial chromosome to generate a set of strains with industrially relevant phenotypes (*i.e.* good growth rate on various nutrients sources,

genetic stability and low waste of energy). In practice, the expected outcome is a simpler cell in which synthetic functional modules or regulators can be plugged in and their effects can be well monitored.

Another exciting approach for future strain improvement or innovative applications is the naturally occurring diversity of *B. subtilis* strains. Because of the great adaptation skills of *B. subtilis*, and its ability to survive in extreme conditions (*e.g.* extreme pH, temperatures or salt concentrations, high radiation), it is likely that exploring the natural diversity of microbes in unusual or specific environments would lead to the discovery of strains with new abilities fitting with industrial applications. Less than a third of the *B. subtilis* genes (paleome) is conserved across other *Bacilli*. This suggests that the large part of its genome is present in response to environmental cues (cenome), meaning that *B. subtilis* has genetically evolved for the occupation of diverse habitats and niches [Alcaraz *et al.*, 2010].

As an example, genes putatively encoding enzymes related to the degradation of complex polysaccharides, as well as possible collagen-binding adhesins, have been identified in the genome of a gut-associated strain [Schyns *et al.*, manuscript in preparation; C. Serra and A.O. Henriques, personal communication]. This undomesticated *B. subtilis* has been isolated from faecal material of organically reared broilers. The presence of these genes, absent from laboratory strains, together with the enhanced sporulation phenotype of the strain, may represent a considerable advantage in terms of adaptation to the gut environment, and might be of interest for the putative use of this strain in probiotic applications.

Exploration of the mutations occurring in the course of strain engineering (*i.e.* artificial diversity of *B. subtilis*) can also be of

interest for industrial biotechnology. Dramatic decrease in the cost of accurate whole-genome re-sequencing during the five past years (about 1,500 € during the writing of this thesis) drew attention to the large amount of untargeted mutations resulting from random and directed mutagenesis campaigns in the multiple-step construction of industrial production strains. To limit the amount of untargeted mutations, a better characterization of the dose-response of mutagenic agents (*e.g*. nitrosoguanidine, ethylmethane sulphonate *etc*.) in *B. subtilis* would probably be relevant. Nevertheless, part of these mutations greatly contributes to the predisposition of *B. subtilis* to become a high metabolite producer. Hence, identification of the relevant ones is of interest, the beneficial mutations providing new strain improvement leads. To support prediction of relevant untargeted mutations, computational tools such as SNAP (Screening of Non-Acceptable Polymorphisms) appeared, based on integration of several databases [Bromberg *et al.*, 2008]. It can also be speculated that a minor part of untargeted mutations has an adverse effect on the expected phenotype. Correction of these mutations, back to the wild-type situation, can therefore directly result in improvement of products yield. The artificial diversity in *B. subtilis* could also be generated in a closely-controlled way, using recent technologies that can speed up the introduction of specific SNPs (Single Nucleotide Polymorphisms). For instance, the RTDS™ (Rapid Trait Development System) developed by Cibus Global (San Diego, USA), is an oligo-based technique designed to alter a targeted gene by utilizing the cell's own gene repair system, without insertion of foreign DNA [Breyer *et al.*, 2009]. This technology has been successfully used in eukaryotes, and should demonstrate its possible applications in *B. subtilis.* Due to the complex interactions of metabolic pathways, the profile of an ideal

candidate for high metabolites production likely depends on multiple genes broadly distributed throughout the genome. Consequently, direct application of targeted genetic engineering for strain improvement has a relatively low chance of success. Genome mixing of a small number of strains obtained by classical strain improvement methods constitutes a valuable method to promote artificial diversity in *B. subtilis* strains. Resulting in subtle progresses in the expected phenotype, this concept could be beneficial in biotechnology research. Successful examples of genome shuffling by performing few rounds of protoplast fusion with *Streptomyces* [Zhang *et al.*, 2002] and *Lactobacillus* genera [Patnaik *et al.*, 2002] have already been reported. A similar approach with *B. subtilis* would therefore surely be of interest.

The greater knowledge and the well-selected strains gained from the above-described emerging axes (synthetic biology, exploration of both natural and artificial genomic diversities), will support further strain improvement, and reinforce the central role of *B. subtilis* in research for biotechnology industries.

4. References

Alcaraz, L.D., G. Moreno-Hagelsieb, L.E. Eguiarte, V. Souza, L. Herrera-Estrella, and G. Olmedo. 2010. Understanding the evolutionary relationships and major traits of *Bacillus* through comparative genomics. BMC Genomics. 11:332.

Anand, R., P.C. Dorrestein, C. Kinsland, T.P. Begley, and S.E. Ealick. 2002. Structure of oxalate decarboxylase from *Bacillus subtilis* at 1.75 Å resolution. Biochemistry. 41(24):7659-69.

Barbosa, T.M., C.R. Serra, R.M. La Ragione, M.J. Woodward, and A.O. Henriques. 2005. Screening for *Bacillus* isolates in the broiler gastrointestinal tract. Appl. Environ. Microbiol. 71(2):968-78.

Breyer, D., P. Herman, A. Brandenburger, G. Gheysen, E. Remaut, P. Soumillion, J. Van Doorsselaere, R. Custers, K. Pauwels, M. Sneyers, and D. Reheul. 2009. Genetic modification through oligonucleotide-mediated mutagenesis. A GMO regulatory challenge? Environ. Biosafety Res. 8(2):57-64.

Bromberg, Y., G. Yachdav, and B. Rost. 2008. SNAP predicts effect of mutations on protein function. Bioinformatics. 24(20):2397-8.

Commichau, F.M., F.M. Rothe, C. Herzberg, E. Wagner, D. Hellwig, M. Lehnik-Habrink, E. Hammer, U. Völker, and J. Stülke. 2009. Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. Mol. Cell. Proteomics. 8(6):1350-60.

Costa, T., L. Steil, L.O. Martins, U. Völker, and A.O. Henriques. 2004. Assembly of an oxalate decarboxylase produced under the sigma K control into the *Bacillus subtilis* spore coat. J.Bacteriol. 186(5):1462-1474.

Isticato, R., G. Cangiano, H.T. Tran, A.L. Ciabattini, D. Medaglini, M.R. Oggioni, M. de Felice, G. Pozzi, and E. Ricca. 2001. Surface display of recombinant proteins on *Bacillus subtilis* spores. J. Bacteriol. 183(21):6294-6301.

Jenkins, A.L., Y. Zhang, S.E. Ealick, and T.P. Begley. 2008. Mutagenesis studies on TenA: a thiamin salvage enzyme from *Bacillus subtilis*. Bioorg. Chem. 36(1):29-32.

Heinemann, M., and U. Sauer. 2010. Systems biology of microbial metabolism. Curr. Opin. Microbiol. 13(3):337-43.

Henriques, A.O., and C.P. Moran Jr. 2007. Structure, assembly, and function of the spore surface layers. Annu. Rev Microbiol. 61:555-588.

Melnick, J., E. Lis, J.H. Park, C. Kinsland, H. Mori, T. Baba, J. Perkins, G. Schyns, O. Vassieva, A. Osterman, and T.P. Begley. 2004. Identification of the two missing bacterial genes involved in thiamin salvage: thiamin pyrophosphokinase and thiamin kinase. J. Bacteriol. 186(11):3660-2

Park, J.H., K. Burns K, C. Kinsland, and T.P. Begley. 2004. Characterization of two kinases involved in thiamin pyrophosphate

and pyridoxal phosphate biosynthesis in *Bacillus subtilis*: 4-amino-5-hydroxymethyl-2methylpyrimidine kinase and pyridoxal kinase. J Bacteriol. 186(5):1571-3.

Patnaik, R., S. Louie, V. Gavrilovic, K. Perry, W.P. Stemmer, C.M. Ryan, and S.B. del Cardayré. 2002. Genome shuffling of *Lactobacillus* for improved acid tolerance. Nat. Biotechnol. 20(7):707-12.

Perkins, J.B., and J. Pero. 2002. Vitamin biosynthesis. In Sonenshein, A.L., *et al.* (eds.). *Bacillus subtilis* and its closest relatives: from genes to cells. ASM Press, Washington DC. 271-286.

Rodionov, D.A., A.G. Vitreschak, A.A. Mironov, and M.S. Gelfand. 2002. Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms. J. Biol. Chem. 277(50):48949-59.

Schmalish, M., E. Maiques, L. Nikolov, A.H. Camp, B. Chevreux, A. Muffler, S. Rodriguez, J.B. Perkins, and R. Losick. 2010. Small genes under sporulation control in the *Bacillus subtilis* genome. J. Bacteriol. In print during the writing of this thesis.

Schyns, G., J. Pereira-Leal, C. Serra, P. Fickers, B. Joris, S. Potot, T. Barbosa, B. Chevreux, S. Rodriguez, A. de Saizieu, J.B. Perkins, M. Wyss, and A.O. Henriques. The genome of a gut strain of *Bacillus subtilis*. Manuscript in preparation.

Schyns, G., S. Potot, A.O. Henriques, and T.J. Wenzel. 2008. Spore surface displays of bioactive molecules. World Intellectual Property Organization. WO/2008/017483.

Toms, A.V., A.L. Haas, J.H. Park, T.P. Begley, and S.E. Ealick. 2005. Structural characterization of the regulatory proteins TenA and TenI from *Bacillus subtilis* and identification of TenA as a thiaminase II. Biochemistry. 44(7):2319-29.

Winkler, W., A. Nahvi, and R.R. Breaker. 2002. Thiamin derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature. 419: 952–956.

Zhang Y.X., Perry K., Vinci V.A., Powell K., Stemmer W.P., and S.B. del Cardayré. 2002. Genome shuffling leads to rapid phenotypic improvement in bacteria. Nature. 415(6872):644-6.