Improving Bacillus subtilis as a cell factory for heterologous protein production by adjusting global regulatory networks

Cao, Haojie

Improving *Bacillus subtilis* as a cell factory for heterologous protein production by adjusting global regulatory networks

HAOJIE CAO

university of groningen

The work described in this thesis was carried out in Molecular Genetics Group of the Groningen Biomolecular Sciences and Biotechnology Institute (GBB) of the University of Groningen, the Netherlands. The candidate was financially supported by the China Scholarship Council (CSC).

Printing of this thesis was financially supported by the Graduate School of Science and Engineering and the University of Groningen.

Cover design & layout: Lovebird design
www.lovebird-design.com

Printed: Eikon+

ISBN: 978-94-034-0788-3
978-94-034-0787-6 (ebook)

university of groningen

Improving Bacillus subtilis as a cell factory for heterologous protein production by adjusting global regulatory networks

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. E. Sterken
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Monday 9 July 2018 at 9:00 hours

by

Haojie Cao

born on 16 March 1987
in Henan, China

Supervisor
Prof. dr. O.P. Kuipers

Co-supervisor
Prof. dr. J.W. Veening

Assessment Committee
Prof. J.M. van Dijl
Prof. W.J. Quax
Prof. A.T. Kovacs

CONTENTS

CHAPTER 1

General Introduction

BACILLUS SUBTILIS

Bacillus subtilis, the name of which was coined in 1872 [1], is predominately found as an inhabitant of the soil or is living in association with plants [2]. It has become the most studied species of the Gram-positive bacteria, and is characterized by one cytoplasmic membrane and a thick cell wall [3]. The cells of *B. subtilis* are typically rod-shaped and about 4.0–10.0 μm long and 0.25–1.0 μm wide in diameter [4]. *B. subtilis* has been regarded as an obligate aerobe, until solid evidence proved that it is a facultative anaerobe [5]. Under some special circumstances, *B. subtilis* can activate the cellular process for forming tough and protective endospores, allowing a subpopulation to survive in environmental conditions of extreme temperatures and desiccation [6]. This heavily flagellated bacteria also enables the cells to move quickly in liquids [7]. *B. subtilis* can grow in nutrient-rich media as well as in a chemically defined salt media in which glucose, malate or other simple sugars that provide carbon sources, and ammonium salts or certain amino acids are applied as nitrogen sources [8]. The most commonly used laboratory strain of *B. subtilis,* 168, is a tryptophan auxotroph (trpC2) and therefore requires the extra addition of tryptophan to the growth media, even to media containing acid-hydrolyzed proteins such as those of casein [2]. The genome of *B. subtilis* has been (re)sequenced and annotated, and contains 4,214,630 base pairs (bp) encoding 4,100 proteins with an overall 43.5% GC content [9, 10]. Based on genome-wide gene function studies of *B. subtilis*, 253 genes that account for 6% of the whole genome are considered to be essential that involved in central metabolism, processing information, cell wall synthesis, cell division, shape [11–13], and only one of them remains function unknown [14]. *B. subtilis* is well amenable to genetic manipulation, and genome editing can be performed using a variety of techniques,

such as phage-mediated transduction [15], transformation (natural, electro- and protoplast-), conjugation [2], CRISPR-Cas9 system [16, 17] and nanotubes-mediated molecular exchange [18].

THE MICROBIAL CELL FACTORY

Microbial cell factories have been largely exploited for the high-level production of various industrially relevant products in the fields of food, pharmaceutical, and biotechnology. By far, the most frequently studied and widely applied microorganism hosts are *Escherichia coli* [19], *B. subtilis* [20, 21], lactic acid bacteria [22], yeast (*Saccharomyces cerevisiae*) [23] and fungi (*Aspergilli*) [24]. All of these are excellent work-horses for producing a wide range of high-value biochemical products. In contrast to other well-known cell factories, the best-characterized Gram-positive bacterium *B. subtilis*, is highly favored due to its status of generally recognized as safe (GRAS) and the outstanding natural secretion capacity that facilitates the downstream purification processing [25]. Furthermore, cultivation of *B. subtilis* cells at high densities is relatively easy and inexpensive. Therefore, this bacterium serves as the most popular large-scale prokaryotic expression system in producing pharmaceutical- or food-grade products by biotechnology companies [20]. In industry, about 60% of the commercial enzymes are generated by *B. subtilis* and its close relatives [26, 27]. Moreover, as shown in **Fig. 1**, introducing a heterologous protein synthesis pathway will take up a large proportion of the resources distribution in the host cell [28]. This metabolic drain problem caused by a human-imposed overproduction task in the expression host, also called metabolic burden, will reduce the activities of energy metabolism and of native enzymes for essential cellular processes as well as limit the availability of building

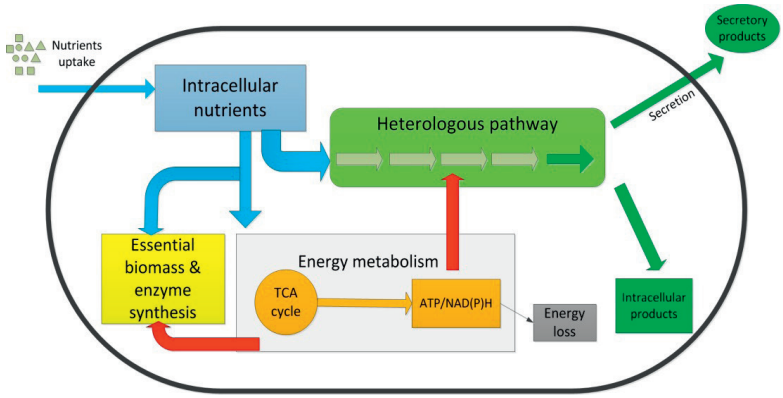


Fig. 1. Schematic representation of metabolic networks in a microbial cell factory.

blocks for the protein overproduction, leading to a lower yield of the target protein [28]. Although high production yields can be achieved for proteins originated from *Bacillus* species by various strain optimization strategies [29], the success rate for overproducing a majority of heterologous proteins is limited [30, 31]. Therefore, modifying the industrial expression host *B. subtilis* to be a better cell factory for heterologous proteins of commercial importance is a research hotspot.

INCREASING PRODUCTIVITY OF *B. SUBTILIS*

Empirical approaches for improving protein production and secretion

In the past, the empirical attempt for optimizing protein production in *B. subtilis* was performed by studying the effects of media compositions on the final product yields [20]. However, the advances in recombinant DNA technology enabled more directed genetic alterations of this production host. Strong promoters were utilized to achieve high transcription levels, and

these promoter sequences that are accompanied by efficient ribosome binding sites (RBS) are commonly used for strain improvement [32, 33], which normally results in higher yields of target proteins [34]. In addition, the widely used isopropylbeta-D-thiogalactopyranoside (IPTG)-inducible promoter [35], the xylose-inducible promoter [36], and the subtilin-inducible promoter [37] offer large advantages over constitutively active promoters in *B. subtilis*. Moreover, the translation efficiency can be enhanced by replacing the rare codons with optimal ones in the open reading frame (ORF) based on the codon usage bias in the host organisms [38, 39].

In *B. subtilis,* the major protein secretion machinery is the Sec-secretion pathway [40, 41]. The overexpression of Sec-components SecDF and SecG and C-terminal modification of SecA are beneficial for obtaining increased levels of heterologous protein secretion in *B. subtilis* [34]. The native pre-protein contains an N-terminal polypeptide sequence, signal peptide, that subsequently directs the to be-secreted target protein across the cytoplasmic membrane [42]. Therefore, suitable signal peptides can effectively enhance the translocation and secretion efficiency of heterologous proteins [25]. Furthermore, the overexpression of thiol-disulfide oxidoreductases BdbB and BdbC is capable of boosting the secretion of disulfide bond-containing proteins [43, 44]. The molecular chaperone PrsA is important for protein folding during the post-translocational phase, and the higher production of this foldase can improve the secretion yields of several model proteins [45, 46]. The deficiency of the *dlt* operon leads to a significant increase of secreted heterologous proteins by altering the cell wall net charge [47]. *B. subtilis* secretes multiple extracellular proteases, and they can rapidly degrade the misfolded, folding or folded protein during or after the secretion process, reducing heterologous protein yields. The use of extracellular protease-deficient

strains that retain less than 5% of total extracellular protease activity compared to the 168 strain can usually tackle this problem [48, 49]. This strategy broadens the application of *B. subtilis* as a super-secreting cell factory for products that are sensitive to these proteases [50, 51]. In **Chapter 2**, several cell surface components were genetically engineered in the protease-inactivated strain *B. subtilis* DB104, and the corresponding influence on the product yield was investigated by analyzing the secretion of codon-optimized α-amylase variants. Our study demonstrated that the composition of the cell membrane phospholipid bilayer and the physicochemical properties of α-amylases play key roles in the final secretion efficiency.

Progress in the metabolic engineering of *B. subtilis*

The classic strategies for improving protein expression systems involve modifying the regulatory elements of homologous or heterologous pathways, such as expression vectors, promoters, RBSs, and terminators, or by varying the availability of the secretion machinery components [52–55]. In order to increase our ability to develop a better production system for a wide range of proteins, a variety of newly established systems and synthetic biology tools has been recently introduced. In the past decades, genetic engineering was mainly based on classic homologous recombination genome editing approaches. However, the advances in novel genome-editing devices, including sRNAs and CRISPR-Cas9 systems, have expanded the genetic engineering scope from specific pathways to the whole genome scale [16, 56, 57]. Engineering strategies of metabolic pathways that integrate systems- and synthetic biology approaches have greatly facilitated unlocking phenotypes with desired cellular properties in *Bacillus* species [58]. The applications of new techniques have significantly enhanced the heterologous production of N-acetylglucosamine in *B. subtilis*, the poly-γ-glutamic

acid in *B. amyloliquefaciens*, and vitamin B12 in *B. megaterium* [59–62]. In light of the fact that physiological properties of *B. subtilis* can seriously affect its robustness of production as a microbial host, researchers also tried to increase the stability and controllability of the *B. subtilis* cell factory by solving the problems of carbon overflow and unfavorable cell lysis [58]. AceA and AceB from *B. licheniformis* were overexpressed in *B. subtilis* to introduce a glyoxylate shunt, which resulted in an improved acetate utilization and strengthened cellular robustness of the host cells [63]. Extending the growth and production period through avoiding cell lysis is another strategy for improving the physiological properties of *B. subtilis* [64]. The absence of a series of lytic genes, namely, *skfA, sdpC, lytC,* and *xpf,* finally caused a 2.6-fold increase of nattokinase production thanks to the drastic decline of the cell lysis rate [64]. In addition, genome reduction is also commonly utilized for strain improvement of established industrial production microorganisms [65]. Using models to design the minimalization programme and predict the impact of genomic alterations on growth and metabolism, *B. subtilis* has been subjected to deletion of large portions of dispensable genome regions, only retaining genes that are responsible for the essential functions to construct a minimal cell factory [12, 13]. This genome engineering did not cause "serious harm" but improved specific heterologous protein yields greatly by decreasing unwanted by-products [65, 66].

Global transcription machinery engineering (gTME)

In a microbial cell factory, overproduction of recombinant proteins is always challenging, and specific pathway optimization for accessing interesting cellular phenotypes usually requires a comprehensive understanding of the cellular metabolic networks under overexpressing conditions [20]. In other words, these strain modification strategies are designed, mainly based

on the known information about metabolic pathways. However, our insufficient insights of multiple layers in metabolic regulation and the complex underlying interaction landscape limits the development of engineered microbial hosts with further improved protein production properties [58]. Metabolic engineering and synthetic biology enable us to manipulate cellular processes by the deletion, depletion or tunable expression of single genes [56], but the exploration of the effects of system-wide pathway modifications is difficult to perform [67]. To overcome these limitations, a concept termed global transcription machinery engineering (gTME) was created that allows multiple and simultaneous perturbations of the whole transcriptome through the engineering of transcription factors that have global effects [52]. gTME, which focuses on the increase of end-products by rerouting metabolic fluxes at a top layer of the regulatory networks, greatly simplifies the strain enhancement strategy [68, 69]. Several proof-of-concept studies of gTME have proved to outperform many other conventional methods in unlocking desired phenotypes with better properties. A variety of global transcription factors, including zinc finger-containing artificial transcription factor [70], RNA polymerase sigma subunit [67, 71], Spt15 [72, 73], H-NS and Hha [74, 75], have been engineered to elicit variants with higher specific metabolic capacity or chemical tolerance. In **Chapter 3**, the gTME-based approach was applied for effectively and quickly unlocking variants with an improved production capacity of the target protein by randomly mutagenizing the global N- and C-regulators CodY and CcpA, respectively. The best cell factory containing crucial mutations that reached an increase of 2-fold in overproduction of β-galactosidase was further demonstrated by the significantly enhanced overexpression of GFP, a xylanase and a peptidase.

GLOBAL REGULATORY PROTEINS INVOLVED IN THE CENTRAL METABOLIC PATHWAYS

To date, many transcriptional regulators that are involved in metabolic processes have been identified and characterized from *B. subtilis* by biochemical and biophysical methods [76]. CcpA and CodY are the two most important global transcriptional regulatory proteins that control key metabolic intersections, orchestrate large regulons for balancing the availability of carbon and nitrogen sources, respectively, and coordinate the intracellular C/N fluxes to maintain cell homeostasis [77].

The global regulators CodY and CcpA

CodY, as a DNA-binding global transcriptional regulator, has homologs that are ubiquitously found in a variety of low G+C Gram-positive bacteria [78]. It was first identified as a repressor of the dipeptide permease (*dppABCDE*) operon in *B. subtilis* and turned out to play a global role in modulating the expression of many other important genes involved in N-metabolism [76]. In *B. subtilis*, CodY exists in the form of a dimer of two 29- kDa subunits, and its cofactors specifically bind to amino acids (aa) 1–155 of the N-terminal region [79]. The conserved helix-turn-helix motif (ASKIADRVGITRSVIVNALR) for the less-conserved CodY-binding sites ('CodY box') has been proposed from analyzing the target gene upstream regions by site-directed mutagenesis [80], which show high affinity to aa 203–222 of the CodY C-terminal domain. The activity of CodY depends on GTP and branched-chain amino acids (BCAAs) (leucine, isoleucine, and valine), and these two ligands have additive effects on CodY-binding efficiency [77]. During growth of cells in rich media, CodY acts as a repressor for hundreds of genes and activates the expression of a few genes in late exponential phase or early stationary phase upon the activation by the intracellular pools of GTP and

BCAAs (**Fig. 2**) [81, 82]. When the nutrients are exhausted, CodY-mediated repression diminishes, and the products of the regulon enable cells to adapt to the suboptimal nutritional conditions [83]. Thus, *B. subtilis* starts an adaptive response to nutrient limitation by inducing a wide variety of cellular processes, such as sporulation, competence development, transport systems, carbon and nitrogen metabolism and biofilm formation [84].

Catabolite control protein A (CcpA), a member of the LacI/GalR family of transcriptional regulators, is highly conserved at the sequence level and widely present in many low-GC Gram-positive bacterial species [44, 46]. CcpA plays a central role in the carbon acquisition and metabolism by modulating the expression of more than 100 relevant genes [85]. As demonstrated in **Fig. 2**, in *B. subtilis*, the sugar phosphotransferase system (PTS) transports the preferred carbohydrate, commonly glucose, into the bacterial cells, and the intracellular PTS-sugar is converted into glucose-6-P (G6P) and fructose-1,6-bisphosphate (FBP) [86]. Subsequently, the accumulation of these two glycolytic intermediates triggers the phosphorylation of HPr (histidine-containing protein) or its homolog protein Crh at Ser-46 in an ATP-dependent reaction catalyzed by HPr kinase/ phosphatase [87]. The P-Ser-HPr/CcpA or P-Ser-Crh/CcpA complex binds to specific upstream DNA regions of regulated genes, causing carbon catabolite control, i.e. carbon catabolite repression (CCR) or carbon catabolite activation (CCA) [88, 89]. The *cis*-acting palindromic sequences located in the promoter or other regions of the regulon are called catabolite-responsive element (*cre*) sites, and two, slightly differently constructed, consensus motifs for *cre* sequences WTGNAANCGNWWNCA and WTGAAARCGYTTWNN have been determined in *B. subtilis* by extensive base substitution analysis [90, 91]. This consensus sequence that slightly differs among different bacteria [92], shows a great variety regarding positioning relative to the transcriptional start site,

which finally results in different regulatory effects of individual genes [76].

The metabolic intersections of CodY and CcpA regulatory pathways in *B. subtilis*

In *B. subtilis*, CcpA and CodY provide a top layer of metabolic regulation that controls the expression levels of the central metabolic genes by sensing the variable availability of key intracellular metabolites (FBP, G6P, GTP, and BCAAs). The products of these regulons direct the distribution of the cellular resources to crucial decision points, through central metabolic intersections [77]. When *B. subtilis* cells are grown in a media that consists of an excess of carbon and nitrogen nutrients, the conversion of pyruvate and acetyl CoA via the *acetate kinase* (AckA)–*phosphate acetyltransferase* (Pta) pathway generates ATP and by-products of overflow metabolism, including lactate, acetate, and acetoin (**Fig. 3**). This pathway is positively regulated by both CcpA and CodY, and the transcription of *ackA* is induced by the binding of these two regulators at two specific neighboring sites among the *ackA* promoter region, and these regulatory effects can be additive [93]. Thus, CcpA and CodY regulate the expression of enzymes that determine the metabolic fate of pyruvate, contributing to the stimulation of the acetate- and lactate-synthesis pathways, while CcpA also activates the synthesis of acetoin (**Fig. 3**). In sum, CcpA and CodY orchestrate the expression of carbon-overflow related genes, affecting the overall *B. subtilis* metabolism. Moreover, there are four transcription units, the *ilv-leu* operon, and the *ilvA*, *ilvD* and *ybgE* genes, devoted to BCAAs biosynthesis in *B. subtilis* [94]. CcpA and CodY showed opposing transcriptional effects on ILV expression by binding to two partially overlapping sites of the *ilvB* promoter region [82, 95]. CodY acts as a direct negative regulator of the *ilv-leu* operon by binding to the overlapping −35-promoter region of

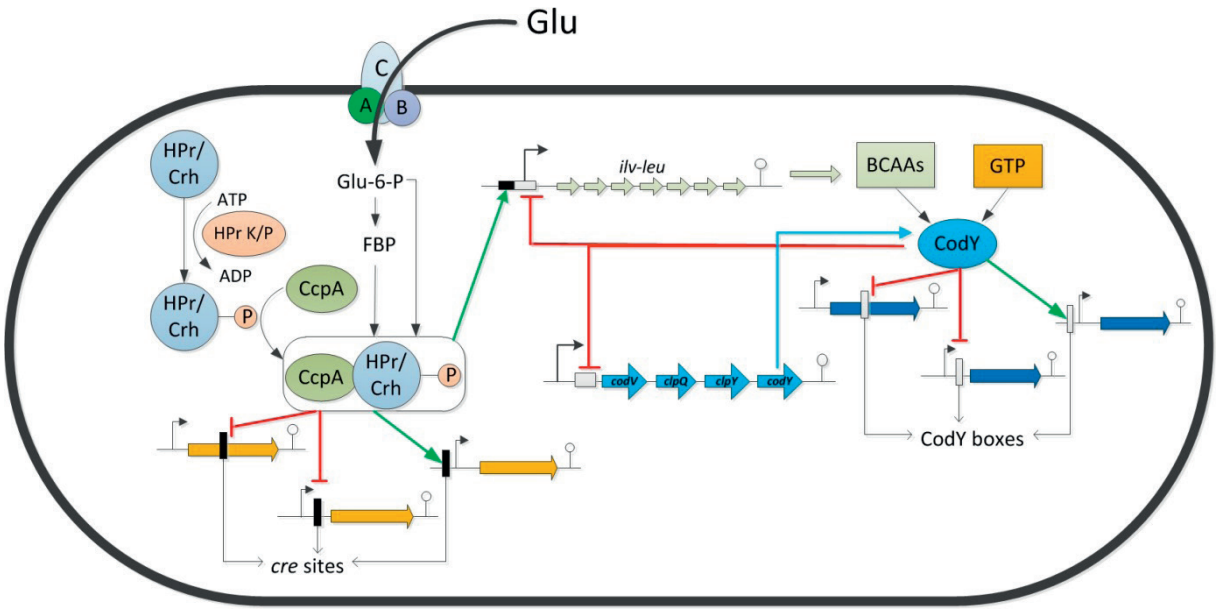


Fig. 2. Schematic diagram of CcpA and CodY regulatory mechanisms in *B. subtilis*. Arrows and perpendiculars represent the positive and negative actions, respectively. Glu denotes glucose; A, B, and C denote a PTS system or permease for glucose.

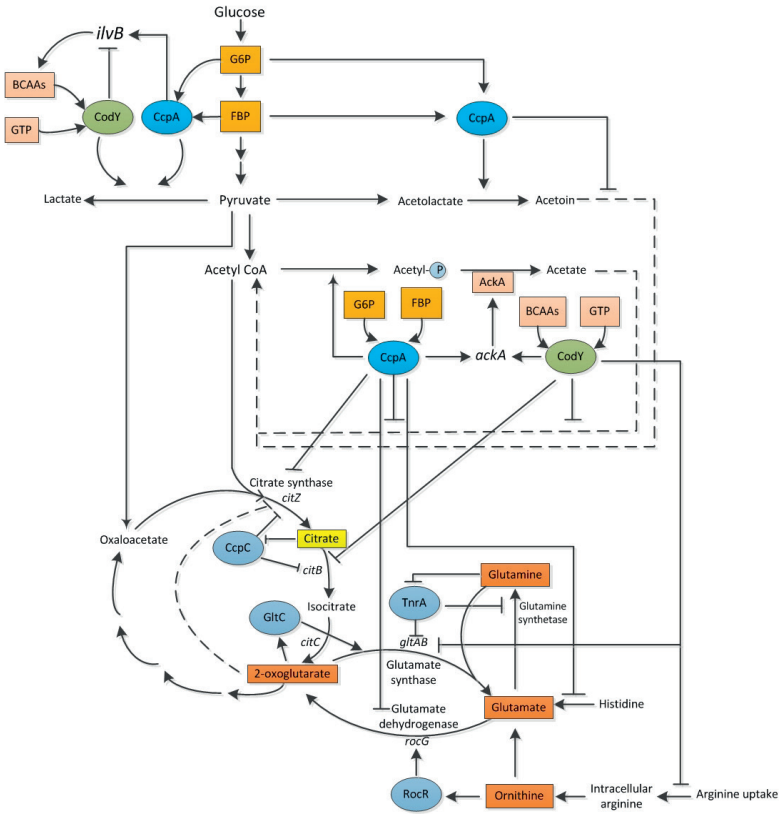


Fig.3. **The metabolic intersections of CodY and CcpA regulatory pathways in *B. subtilis*.** Arrows and perpendiculars represent the positive and negative actions, respectively. This figure was made based on the information from Sonenshein's review [77].

ilvB (**Fig. 2**). Besides the direct positive role, CcpA also acts partly as an indirect positive regulator by interfering with the repression by CodY [95, 96]. As shown in **Fig. 3**, under variable nutrient availability, the two regulators are able to quickly achieve a steady state by mediating the intracellular levels of BCAA biosynthesis. The two proteins coordinately determine the impact on many CodY-regulated genes, that is, CcpA directly regulates some genes that are in the CodY regulon [77].

Moreover, CcpA, CodY and another regulatory protein, CcpC, cooperate on the modulation of a series of genes that encode the enzymes of the citric acid cycle (also known as tricarboxylic acid (TCA) cycle). The generated enzymes citrate synthase (*citZ*), aconitase (*citB*) and isocitrate dehydrogenase (*citC*) function together to determine the extent to which pyruvate and acetyl CoA enter the TCA branch, and regulate the pathway from pyruvate to 2-oxoglutarate (**Fig. 3**). CcpC is a specific regulator of the TCA cycle genes, and the *citZ* and *citB* genes are repressed by CcpC [97], while CcpA and CodY are on a top layer of this primary regulatory mechanism by the respective repression of *citZ and citB* [98, 99]. The TCA cycle intermediate, 2-oxoglutarate, is also an entry point into the central carbon and nitrogen metabolism by providing carbon skeletons for several amino acids (**Fig. 3**). The operon specific regulators GltC and TnrA determine the expression of *gltAB*, the product of which, glutamate synthase, catalyzes the *de novo* synthesis of glutamate from 2-oxoglutarate [100–102]. Notably, the conversion of 2-oxoglutarate from glutamate, which is driven by glutamate dehydrogenase (RocG), depends on another specific activator, RocR [103]. Moreover, the transcription of *gltAB* and *rocG*, are separately repressed by CodY and CcpA [83, 104]. Therefore, the pathway, no matter from or to 2-oxoglutarate, is under control of both specific and global regulation [77]. Hence, CcpA and CodY collaborate with a wide variety of other transcriptional regulators to determine the overall metabolic status of the bacteria by repressing or activating genes, which are involved in the carbon overflow, and citric acid cycle pathways, BCAA biosynthetic pathway, and the interplay between carbon and nitrogen metabolism [77]. In **Chapter 4**, we found that the mutated CodY and CcpA proteins lead to an overall shift of the central metabolic pathways by analyzing the transcriptome and binding affinities in the cell, which is expected to further reveal the intricate metabolic networks.

CELLULAR HETEROGENEITY OF *B. SUBTILIS*

The genetically identical *B. subtilis* cells within a population can display a multitude of distinct phenotypes, even under the same environmental condition [105]. When the nutrients are exhausting, *B. subtilis* in the stationary phase generates a mixed population, in which some cells form spores that are highly resistant to external stresses [106], and a subset of the sporulating cells can secrete an extracellular toxin to cause the lysis of sister cells [107]. Under certain conditions, a subpopulation of the cells can differentiate into a genetically competent state that takes up DNA from the environment [108, 109], or become motile by producing flagellar [110], or generate extracellular matrix material to form a robust biofilm [111]. The diverse developmental pathways that determine distinct cell types, are part of an intricate network that relies primarily on the activity of three major transcriptional regulators: Spo0A, DegU, and ComK [112]. The phenotypes with diverse features that are present in one community are regarded as increasing the chance of the whole population that is better adapted to changing conditions [113]. Interestingly, a previous study demonstrated the coexistence of distinct high- and low-level α-amylase-producing cells in one bacterial population of *B. subtilis*, and the overall secretion yields of α-amylases is highly correlated with the expression homogeneity [114]. The green fluorescent protein (GFP) that is derived from *Aequorea victoria* [115] has been widely utilized for benchmarking gene expression in the study of protein localization or promoter activity in living cells [116, 117]. However, the inherent expression heterogeneity of this most representative reporter protein has rarely been studied. In **Chapter 3**, we show that the mutant $CodY^{R214C}CcpA^{T19S}$ could improve the intracellular synthesis of GFP. To obtain a deeper insight and dynamic pattern of GFP production in various expression hosts, flow cytometry and

fluorescent microscopy that facilitate the analysis of cell behavior both at the population and single-cell levels [118–120], were used in **Chapter 5**. Here, the heterogeneity of GFP expression that is partly dependent on wild-type or mutated global regulators and its role in heterologous protein production at a population scale, were systematically investigated.

SCOPE OF THIS THESIS

Throughout this study, we utilized advanced engineering approaches to improve the product yields of the classic reporter proteins α-amylase and β-galactosidase in *B. subtilis*, thus increasing the application value of this model organism as an industrial production platform. Additionally, the analyses of overproducing cells by a variety of techniques, revealed correlations between cellular regulatory processes and product yields and offered a better understanding of the underlying intersections of diverse metabolic pathways.

Chapter 1 is the general introduction concerning the research progress of studies in *B. subtilis* as a microbial production system for recombinant proteins, and the global transcriptional regulators, which are involved in the key metabolic pathways and the related metabolic intersections, and the cellular heterogeneity of *B. subtilis*.

Chapter 2 explores the possibility of improving the secretion efficiency of heterologous protein-α-amylases in *B. subtilis* by genetically engineering some of its cell surface components and secreted proteins. Lipid analysis was also performed to investigate the effect of cell membrane phospholipid composition alteration on secretion yields.

In **Chapter 3**, we reprogramed the carbon/nitrogen metabolism at a global level by use of random mutagenesis of the

global transcriptional regulators CcpA and CodY in *B. subtilis*. High throughput screening was used for quickly selecting phenotypes with increased production capacity of the heterologous intracellular reporter protein–β-galactosidase.

In **Chapter 4**, we investigated the alterations of the metabolic regulatory networks in the previously identified higher-producing β-galactosidase cells, by a system-wide analysis of the transcriptome and by use of a protein-DNA affinity assay of the mutated regulatory proteins.

Chapter 5 is an extension of **Chapter 3**; in this part, we expressed another reporter protein-i.e. GFP, in the previously obtained modified expression host. The dynamic expression of GFP in *B. subtilis* with mutations in CcpA and/or CodY both at the population, subpopulation, and single-cell levels was monitored. Also, the influence of GFP expression heterogeneity on the overall product yield in different backgrounds of the cell populations was studied.

Chapter 6 summarizes the studies of all the experimental chapters and discusses the future perspectives for improving *B. subtilis* as a microbial cell factory for the production of many useful proteins.

REFERENCES

1. Cohn F. Untersuchungen über Bacterien. *Beiträge zur Biologie der Pflanzen.* 1875;1:pp.127–224.
2. Harwood CR, *et al. Bacillus subtilis*: model Gram-positive synthetic biology chassis. In *Methods in microbiology.* Academic Press. 2013;40:87–117.
3. Harwood CR, Cranenburgh R. *Bacillus* protein secretion: an unfolding story. *Trends Microbiol.* 2008;16:73–79.
4. Yu AC, *et al.* Monitoring bacterial growth using tunable resistive pulse sensing with a pore-based technique. *Appl Microbiol Biotechnol.* 2014;98:855–862.

5. Nakano MM, Zuber P. Anaerobic growth of a "strict aerobe"(*Bacillus subtilis*). *Annu Rev Microbiol.* 1998;52:165–190.
6. Madigan M, Martinko J. Brock biology of microorganisms. *SciELO Espana.* 2005.
7. Guttenplan SB, Shaw S, Kearns DB. The cell biology of peritrichous flagella in *Bacillus subtilis*. *Mol Microbiol.* 2013;87:211–229.
8. Harwood CR, Cutting SM. *Molecular biological methods for Bacillus.* Wiley,1990.
9. Kunst F, *et al.* The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature.* 1997;390.
10. Barbe V, *et al*. From a consortium sequence to a unified sequence: The *Bacillus subtilis* 168 reference genome a decade later. *Microbiology.* 2009;155:1758–1775.
11. Kobayashi K, *et al.* Essential *Bacillus subtilis* genes. *Proc Natl Acad Sci U S A.* 2003;100:4678–4683.
12. Reuss DR, *et al.* The blueprint of a minimal cell: MiniBacillus. *Microbiol Mol Biol Rev.* 2016;80:955–987.
13. Reuss DR, *et al.* Large-scale reduction of the *Bacillus subtilis* genome: consequences for the transcriptional network, resource allocation, and metabolism. *Genome Res.* 2017;27:289–299.
14. Juhas M, *et al. Bacillus subtilis* and *Escherichia coli* essential genes and minimal cell factories after one decade of genome engineering. *Microbiology.* 2014;160:2341–2351.
15. Errington J. Efficient *Bacillus subtilis* cloning system using bacteriophage vector 01Q5J9. *Microbiology.* 1984;130:2615–2628.
16. Westbrook AW, Moo-Young M, Chou CP. Development of a CRISPR-Cas9 tool kit for comprehensive engineering of *Bacillus subtilis*. *Appl Environ Microbiol.* 2016;82:4876–4895.
17. Altenbuchner J. Editing of the *Bacillus subtilis* genome by the CRISPR-Cas9 system. *Appl Environ Microbiol.* 2016;82:5421–5427.
18. Dubey GP, Ben-Yehuda S. Intercellular nanotubes mediate bacterial communication. *Cell.* 2011;144: 590–600.
19. Li J, Neubauer P. *Escherichia coli* as a cell factory for heterologous production of nonribosomal peptides and polyketides. *Nat Biotechnol.* 2014;31:579–585.

20. Westers L, Westers H, Quax WJ. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochim Biophys Acta.* 2004;1694:299–310.
21. van Dijl JM, Hecker M. *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. *Microb Cell Fact.* 2013;12:p.1.
22. Hugenholtz J. The lactic acid bacterium as a cell factory for food ingredient production. *Int Dairy J.* 2008;18:466–475.
23. Kavscek M, *et al.* Yeast as a cell factory: current state and perspectives. *Microb Cell Fact.* 2015;14:94.
24. Meyer V, Wu B, Ram AFJ. *Aspergillus* as a multi-purpose cell factory: current status and perspectives. *Biotechnol Lett.* 2011;33:469–476.
25. Simonen M, Palva I. Protein secretion in *Bacillus* species. *Microbiol Rev.* 1993;57:109–137.
26. Schallmey M, Singh A, Ward OP. Developments in the use of *Bacillus* species for industrial production. *Can J Microbiol.* 2004;50:1–17.
27. Ferreira L, Ferreira RC, Schumann W. *Bacillus subtilis* as a tool for vaccine development: from antigen factories to delivery vectors. *An Acad Bras Cienc.* 2005;77:113–124.
28. Glick BR. Metabolic load and heterologous gene expression. *Biotechnol Adv.* 1995;13:247–261.
29. Harwood CR. *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. *Trends Biotechnol.* 1992;10:247–256.
30. Bolhuis A, *et al.* Evaluation of bottlenecks in the late stages of protein secretion in *Bacillus subtilis*. *Appl Environ Microbiol.* 1999;65:2934–2941.
31. Li W, Zhou X, Lu P. Bottlenecks in the expression and secretion of heterologous proteins in *Bacillus subtilis*. *Res Microbiol.* 2004;155:605–610.
32. Rocak S, Linder P. DEAD-box proteins: the driving forces behind RNA metabolism. *Nat Rev Mol Cell Bio.* 2004;5:232-241.
33. Salis HM, Mirsky EA, Voigt CA. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol.* 2009;27:946–950.
34. Nijland R, Kuipers OP. Optimization of protein secretion by *Bacillus subtilis*. *Recent Pat Biotechnol.* 2008;2:79–87.

35. Joseph P, *et al.* Rapid orientated cloning in a shuttle vector allowing modulated gene expression in *Bacillus subtilis*. *FEMS Microbiol Lett.* 2001;205:91–97.
36. Kim L, Mogk A, Schumann W. A xylose-inducible *Bacillus subtilis* integration vector and its application. *Gene.* 1996;181:71–76.
37. Bongers RS, *et al.* Development and characterization of a subtilin-regulated expression system in *Bacillus subtilis*: strict control of gene expression by addition of subtilin. *Appl Environ Microbiol.* 2005;71:8818–8824.
38. Jana S, Deb JK. Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Appl Microbiol Biotechnol.* 2005;67:289–298.
39. Veening JW, *et al.* Visualization of differential gene expression by improved cyan fluorescent protein and yellow fluorescent protein production in *Bacillus subtilis*. *Appl Environ Microbiol.* 2004;70:6809–6815.
40. Tjalsma H, *et al.* Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol Mol Biol Rev.* 2000;64.
41. Yamane K, Bunai K, Kakeshita H. Protein traffic for secretion and related machinery of *Bacillus subtilis*. *Biosci Biotech Bioch.* 2004;68:2007–2023.
42. von Heijne G. The signal peptide. *J Membrane Biol.* 1990;115:195–201.
43. Kouwen TR, van Dijl JM. Applications of thiol-disulfide oxidoreductases for optimized in vivo production of functionally active proteins in *Bacillus*. *Appl Microbiol Biotechnol.* 2009;85:45.
44. Kouwen TR, *et al.* Modulation of thiol-disulfide oxidoreductases for increased production of disulfide-bond-containing proteins in *Bacillus subtilis*. *Appl Environ Microbiol.* 2008;74:7536–7545.
45. Vitikainen M, *et al.* Secretion of heterologous proteins in *Bacillus subtilis* can be improved by engineering cell components affecting post-translocational protein folding and degradation. *J Appl Microbiol.* 2005;99:363–375.
46. Vitikainen M, *et al.* Quantitation of the capacity of the secretion apparatus and requirement for PrsA in growth and secretion of α-amylase in *Bacillus subtilis*. *J Bacteriol.* 2001;183:1881–1890.
47. Hyyryläinen HL, *et al.* D-Alanine substitution of teichoic acids as a modulator of protein folding and stability at the cytoplasmic membrane/cell wall interface of *Bacillus subtilis*. *J Biol Chem.* 2000;275: 26696–703.

48. Sloma A, *et al.* Cloning and characterization of the gene for an additional extracellular serine protease of *Bacillus subtilis*. *J Bacteriol.* 1991;173:6889–6895.
49. Wu XC, *et al.* Engineering a *Bacillus subtilis* expression-secretion system with a strain deficient in six extracellular proteases. *J Bacteriol.* 1991;173:4952–4958.
50. Wong SL, Ye R, Nathoo S. Engineering and production of streptokinase in a *Bacillus subtilis* expression-secretion system. *Appl Environ Microbiol.* 1994;60:517–523.
51. Wu XC, *et al.* Efficient production of a functional single-chain antidigoxin antibody via an engineered *Bacillus subtilis* expression-secretion system. *Nat Biotechnol.* 1993;11:71–76.
52. Alper H, *et al.* Tuning genetic control through promoter engineering. *Proc Natl Acad Sci U S A.* 2005;102:12678–12683.
53. Curran KA, *et al.* Use of expression-enhancing terminators in *Saccharomyces cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metab Eng.* 2013;19:88–97.
54. Curran KA, *et al.* Short synthetic terminators for improved heterologous gene expression in yeast. *ACS Synth Biol.* 2015;4:824–832.
55. Zelcbuch L, *et al.* Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res.* 2013;41:e98–e98.
56. Tyo KE, Alper HS, Stephanopoulos GN. Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends Biotechnol.* 2007;25:132–137.
57. Peters JM, *et al.* A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell.* 2016;165:1493–1506.
58. Liu Y, *et al.* Metabolic engineering of *Bacillus subtilis* fueled by systems biology: Recent advances and future directions. *Biotechnol Adv.* 2017;35:20–30.
59. Biedendieck R, *et al.* Metabolic engineering of cobalamin (vitamin B12) production in *Bacillus megaterium*. *Microb Biotechnol.* 2010;3:24–37.
60. Feng J, *et al.* Improved poly-γ-glutamic acid production in *Bacillus amyloliquefaciens* by modular pathway engineering. *Metab Eng.* 2015;32:106–115.
61. Liu Y, *et al.* Modular pathway engineering of *Bacillus subtilis* for improved N-acetylglucosamine production. *Metab Eng.* 2014;23:42–52.

62. Biggs BW, *et al.* Multivariate modular metabolic engineering for pathway and strain optimization. *Curr Opin Biotech.* 2014;29:156–162.
63. Kabisch J, *et al.* Metabolic engineering of *Bacillus subtilis* for growth on overflow metabolites. *Microb Cell Fact.* 2013;12:72.
64. Wang Y, *et al.* Deleting multiple lytic genes enhances biomass yield and production of recombinant proteins by *Bacillus subtilis*. *Microb Cell Fact.* 2014;13:129.
65. Moya A, *et al.* Toward minimal bacterial cells: evolution vs. design. *FEMS Microbiol Rev.* 2009;33:225–235.
66. Manabe K, *et al.* Combined effect of improved cell yield and increased specific productivity enhances recombinant enzyme production in genome-reduced *Bacillus subtilis* strain MGB874. *Appl Environ Microbiol.* 2011;77:8370–8381.
67. Alper H, Stephanopoulos G. Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab Eng.* 2007;9:258–267.
68. Lanza AM, Alper HS. Using transcription machinery engineering to elicit complex cellular phenotypes. In *Synthetic Gene Networks*. Humana Press. 2012;229–248.
69. Woolston BM, Edgar S, Stephanopoulos G. Metabolic engineering: past and future. *Annu Rev Chem Biomol Eng.* 2013;4:259–288.
70. Park KS, *et al.* Phenotypic alteration of eukaryotic cells using randomized libraries of artificial transcription factors. *Nat Biotechnol.* 2003;21:1208–1214.
71. Klein-Marcuschamer D, Stephanopoulos G. Assessing the potential of mutational strategies to elicit new phenotypes in industrial strains. *Proc Natl Acad Sci U S A.* 2008;105:2319–2324.
72. Alper H, *et al.* Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science.* 2006;314:1565–1568.
73. Liu JZ, *et al.* A versatile gene-based test for genome-wide association studies. *Am J Hum Genet.* 2010;87:139–145.
74. Hong SH, Wang X, Wood TK. Controlling biofilm formation, prophage excision and cell death by rewiring global regulator H-NS of *Escherichia coli*. *Microb Biotechnol.* 2010;3:344–356.

75. Hong SH, Lee J, Wood TK. Engineering global regulator Hha of *Escherichia coli* to control biofilm dispersal. *Microb Biotechnol.* 2010;3:717–728.
76. Brantl S, Licht A. Characterisation of *Bacillus subtilis* transcriptional regulators involved in metabolic processes. *Curr Protein Pept Sc.* 2010;11:274–291.
77. Sonenshein AL. Control of key metabolic intersections in *Bacillus subtilis*. *Nat Rev Microbiol.* 2007;5:917–927.
78. Sonenshein AL. CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. *Curr Opin Microbiol.* 2005;8:203–207.
79. Levdikov VM, *et al.* Structure of the branched-chain amino acid and GTP-sensing global regulator, CodY, from *Bacillus subtilis*. *J Biol Chem.* 2017;292:2714–2728.
80. Joseph P, Ratnayake-Lecamwasam M, Sonenshein AL. A region of *Bacillus subtilis* CodY protein required for interaction with DNA. *J Bacteriol.* 2005;187:4127–4139.
81. Ratnayake-Lecamwasam M, *et al. Bacillus subtilis* CodY represses early-stationary-phase genes by sensing GTP levels. *Genes Dev.* 2001;15:1093–1103.
82. Shivers RP, Sonenshein AL. Activation of the *Bacillus subtilis* global regulator CodY by direct interaction with branched-chain amino acids. *Mol Microbiol.* 2004;53:599–611.
83. Molle V, *et al.* Additional targets of the *Bacillus subtilis* global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript analysis. *J Bacteriol.* 2003;185:1911–1922.
84. Belitsky BR, Sonenshein AL. Genome-wide identification of *Bacillus subtilis* CodY-binding sites at single-nucleotide resolution. *Proc Natl Acad Sci U S A.* 2013;110:7026–7031.
85. Moreno MS, *et al.* Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. *Mol Microbiol.* 2001;39:1366–1381.
86. Fujita Y. Carbon catabolite control of the metabolic network in *Bacillus subtilis*. *Biosci Biotechnol Biochem.* 2009;73(2):245–59.
87. Mijakovic I, *et al.* Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: A relic of early life? *Proc Natl Acad Sci U S A.* 2002;99:13442–13447.

88. Fujita Y, *et al.* Specific recognition of the *Bacillus subtilis gnt cis*-acting catabolite-responsive element by a protein complex formed between CcpA and seryl-phosphorylated HPr. *Mol Microbiol.* 1995;17:953–960.
89. Galinier A, *et al.* The *Bacillus subtilis crh* gene encodes a HPr-like protein involved in carbon catabolite repression. *Proc Natl Acad Sci U S A.* 1997;94:8439–8444.
90. Miwa, Y, *et al*. Evaluation and characterization of catabolite-responsive elements (*cre*) of *Bacillus subtilis*. *Nucleic Acids Res.* 2000;28:1206–1210.
91. Miwa, Y, and Fujita, Y. Involvement of two distinct catabolite-responsive elements in catabolite repression of the *Bacillus subtilis myo*-inositol (*iol*) operon. *J Bacteriol.* 2001;183,5877–5884.
92. Zomer AL, *et al.* Time-resolved determination of the CcpA regulon of *Lactococcus lactis subsp. cremoris* MG1363. *J Bacteriol.* 2007;189:1366–1381.
93. Shivers RP, Dineen SS, Sonenshein AL. Positive regulation of *Bacillus subtilis ackA* by CodY and CcpA: establishing a potential hierarchy in carbon flow. *Mol Microbiol.* 2006;62:811–822.
94. Fink PS. Biosynthesis of the branched-chain amino acids. In *Bacillus subtilis and other gram-positive bacteria.* American Society of Microbiology; 1993: 307–317.
95. Shivers RP, Sonenshein AL. *Bacillus subtilis ilvB* operon: an intersection of global regulons. *Mol Microbiol.* 2005;56:1549–1559.
96. Tojo S, *et al.* Elaborate transcription regulation of the *Bacillus subtilis ilv-leu* operon involved in the biosynthesis of branched-chain amino acids through global regulators of CcpA, CodY and TnrA. *Mol Microbiol.* 2005;56:1560–1573.
97. Jourlin-Castelli C, *et al.* CcpC, a novel regulator of the LysR family required for glucose repression of the *citB* gene in *Bacillus subtilis*. *J Mol Biol.* 2000;295:865–878.
98. Kim HJ, Roux A, Sonenshein AL. Direct and indirect roles of CcpA in regulation of *Bacillus subtilis* Krebs cycle genes. *Mol Microbiol.* 2002;45:179–190.
99. Kim HJ, *et al.* Complex regulation of the *Bacillus subtilis* aconitase gene. *J Bacteriol.* 2003;185:1672–1680.
100. Belitsky BR, Janssen PJ, Sonenshein AL. Sites required for GltC-dependent regulation of *Bacillus subtilis* glutamate synthase expression. *J Bacteriol.* 1995;177:5686–5695.

101. Picossi S, Belitsky BR, Sonenshein AL. Molecular mechanism of the regulation of *Bacillus subtilis gltAB* expression by GltC. *J Mol Biol.* 2007;365:1298–1313.
102. Wray LV, *et al. Bacillus subtilis* glutamine synthetase controls gene expression through a protein-protein interaction with transcription factor TnrA. *Cell.* 2001;107:427–435.
103. Belitsky BR, Sonenshein AL. Role and regulation of *Bacillus subtilis* glutamate dehydrogenase genes. *J Bacteriol.* 1998;180:6298–6305.
104. Belitsky BR, Kim HJ, Sonenshein AL. CcpA-dependent regulation of *Bacillus subtilis* glutamate dehydrogenase gene expression. *J Bacteriol.* 2004;186:3392–3398.
105. Mars RA, *et al.* Small regulatory RNA-induced growth rate heterogeneity of *Bacillus subtilis*. *PLoS Genet.* 2015;11:e1005046.
106. Higgins D, Dworkin J. Recent progress in *Bacillus subtilis* sporulation. *FEMS Microbiol Rev.* 2012;36:131–148.
107. Gonzalez-Pastor JE, Hobbs EC, Losick R. Cannibalism by sporulating bacteria. *Science.* 2003;301:510–513.
108. Chen I, Dubnau D. DNA uptake during bacterial transformation. *Nat Rev Microbiol.* 2004;2:241–249.
109. Shank EA, Kolter R. Extracellular signaling and multicellularity in *Bacillus subtilis*. *Curr Opin Microbiol.* 2011;14:741–747.
110. Kearns DB, Losick R. Cell population heterogeneity during growth of *Bacillus subtilis*. *Genes Dev.* 2005;19:3083–3094.
111. Vlamakis H, *et al.* Sticking together: building a biofilm the *Bacillus subtilis* way. *Nat Rev Microbiol.* 2013;11:157–168.
112. Lopez D, Vlamakis H, Kolter R. Generation of multiple cell types in *Bacillus subtilis*. *FEMS Microbiol Rev.* 2009;33:152–163.
113. de Jong IG, Veening JW, Kuipers OP. Single cell analysis of gene expression patterns during carbon starvation in *Bacillus subtilis* reveals large phenotypic variation. *Environ Microbiol.* 2012;14:3110–3121.
114. Ploss TN, *et al.* Homogeneity and heterogeneity in amylase production by *Bacillus subtilis* under different growth conditions. *Microb Cell Fact.* 2016;15:57.

115. Chalfie M, *et al.* Green fluorescent protein as a marker for gene expression. *Science.* 1994:802–805.
116. Shaner NC, Steinbach PA, Tsien RY. A guide to choosing fluorescent proteins. *Nat Methods.* 2005;2:905–909.
117. Southward CM, Surette MG. The dynamic microbe: green fluorescent protein brings bacteria to light. *Mol Microbiol.* 2002;45:1191–1196.
118. Veening JW, *et al.* Visualization of differential gene expression by improved cyan fluorescent protein and yellow fluorescent protein production in *Bacillus subtilis*. *Appl Environ Microbiol.* 2004;70:6809–6815.
119. Garcia-Betancur JC, *et al.* Single-cell analysis of *Bacillus subtilis* biofilms using fluorescence microscopy and flow cytometry. *J Vis Exp.* 2012.
120. Trip EN, *et al.* Balanced transcription of cell division genes in *Bacillus subtilis* as revealed by single cell analysis. *Environ Microbiol.* 2013;15:3196–3209.

CHAPTER 2

Cell surface engineering of *Bacillus subtilis* improves production yields of heterologously expressed α-amylases

Haojie Cao[1], Auke J. van Heel[1], Hifza Ahmed[1], Maarten Mols[1], Oscar P. Kuipers[1]

[1]Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands.

Published as:

Haojie Cao, Auke J. van Heel, Hifza Ahmed, Maarten Mols, Oscar P. Kuipers. Cell surface engineering of *Bacillus subtilis* improves production yields of heterologously expressed alpha-amylases. *Microb Cell Fact.* 2017;16:56.

ABSTRACT

Bacillus subtilis is widely used as a cell factory for numerous heterologous proteins of commercial value and medical interest. To explore the possibility of further enhancing the secretion potential of this model bacterium, a library of engineered strains with modified cell surface components was constructed, and the corresponding influences on protein secretion were investigated by analyzing the secretion of α-amylase variants with either low-, neutral- or high- isoelectric points (pI). Relative to the wild-type strain, the presence of overall anionic membrane phospholipids (phosphatidylglycerol and cardiolipin) increased dramatically in the PssA$^-$, ClsA$^-$ and double KO mutants, which resulted in an up to 47% higher secretion of α-amylase. Additionally, we demonstrated that the appropriate net charge of secreted targets (AmyTS-23, AmyBs and AmyBm) was beneficial for secretion efficiency as well. In *B. subtilis*, the characteristics of cell membrane phospholipid bilayer and the pIs of heterologous α-amylases appear to be important for their secretion efficiency. These two factors can be engineered to reduce the electrostatic interaction between each other during the secretion process, which finally leads to a better secretion yield of α-amylases.

Keywords: *Bacillus*, Protein secretion, α-Amylases, Electrostatic interaction, PssA, ClsA, Cardiolipin, Phosphatidylglycerol

BACKGROUND

The Gram-positive bacterium *Bacillus subtilis* is one of the best-characterized microorganisms to date. This non-pathogenic cell factory is commonly used for the large-scale production of industrial enzymes due to its genetic amenability and superb fermentation characteristics [1, 2]. The molecular mechanisms underlying protein targeting and export have been studied extensively. Various classical genetic approaches have been applied to enhance gene expression and protein secretion, which has resulted in the development of efficient strains for high-level protein production and recovery [3–5]. However, the expression of heterologous proteins can still be challenging and unpredictable with respect to yield. Efforts to improve our understanding of this economically important process are therefore useful to society and industry [6–8].

Previously, numerous studies have been done to improve the protein production and secretion. For instance, Kakeshita *et al.* deleted the C-terminus of the SecA secretory machinery to improve the secretion of heterologous proteins [9]. An extracellular α-amylase has been shown to have increased expression in *B. subtilis* by overproduction of PrsA lipoprotein and optimization of regulatory components [10]. In addition, Thwaite *et al.* have found that the modified cell wall microenvironment (the deficiency of D-alanylation) allows 2.5-folds higher production of recombinant *Bacillus anthracis* protective antigen (rPA) [11]. Furthermore, Degering *et al.* managed to raise the yield of extracellular protease significantly both in *B. subtilis* and *B. licheniformis* by a screening of homologous and heterologous signal peptides [12]. Nevertheless, most of these improvement strategies have focused on the modification of the secretion machinery itself. The engineering of the cell envelope, where secretion takes place, is a novel approach. The cell envelope of

B. subtilis, which is composed of the lipid bilayer and cell wall, should be transversed by a protein that is excreted by the bacterium into the extracellular environment. Some important aspects of the lipid bilayer and cell wall in relation to secretion are discussed below.

B. subtilis has a very complex and variable membrane lipid composition; it consists of 20–50% zwitterionic phospholipid phosphatidylethanolamine (PE), 15–45% phosphatidylglycerol (PG), 2–15% lysyl-phosphatidylglycerol (LysPG), 2–25% cardiolipin (CL) and 10–30% mono-, di- and tri-glucosyl diacylglycerol (GL). The lipid composition changes during growth and cross-regulation between lipid synthesis pathways are suggested to occur in order to maintain membrane functionality and integrity, but how this is regulated is currently unknown [13–15]. The presence of PG in the membrane is essential for the survival of *B. subtilis*, and the specific subcellular localization of SecA in spiral-like structures was shown to have a high PG dependence [16]. Additionally, cardiolipin plays an important role in spore formation [15] and the adaptation to high salt concentrations, and the amount of anionic lipids (PG and CL) in the membrane indicated a strong correlation with the osmoresistance of the cells [13]. Furthermore, Tat-dependent translocation in *E. coli* was shown to depend on negatively charged phospholipids [17]. Interestingly, the Tat proteins in *B. subtilis* are localized at the poles, where the membrane is enriched in CL. Hence, this lipid might also be important for activity and/or localization of the Tat machinery in *B. subtilis*, but this has not been investigated yet.

The cell wall of *B. subtilis* is a multilayered structure formed by a copolymer of peptidoglycan and anionic polymers (teichoic and teichuronic acid) and contains lipoteichoic acid and proteins. There are two aspects of the bacterial cell wall that can determine the efficiency of passage by a secretory protein,

i.e. the charge density and the cross-linking index. Generally, proteins that are translocated via the Sec machinery arrive at the trans-side of the membrane in a relatively unfolded state, where they will encounter the cell wall and are efficiently folded into a protease-resistant conformation [2]. In addition, the overall cell wall net charge is modulated by the extent of D-alanylation of teichoic acid by the products of the *dlt* operon [11]. The inactivation of this operon can increase the net negative charge of the cell wall, thus enhancing the folding and stability of a number of secreted proteins [10,18]. Besides the charge density, the amount of crosslinking of the thick peptidoglycan layer of the cell wall that determines the size of the holes in the peptidoglycan network, may have significant effects on the efficiency of secretion [19,20].

In this study, we attempted to weaken the secretion barrier from the cell envelope in order to improve the secretion potential of the *Bacillus subtilis*, realizing that the physicochemical properties of the secreted protein are crucial as well, and the enzyme productivity also depends on the nature of the target protein. The α-amylase from *B. licheniformis* can have a secretion advantage when it is optimized to have a lower isoelectric point (pI) [21]. Knowing this, the cell surface components were genetically engineered, and their effects on protein secretion were systematically investigated by using a variety of α-amylases with different pIs. We finally observed that reduced electrostatic interactions increased secretion efficiency of amylase proteins.

RESULTS

α-Amylase variants

A variety of α-amylases were chosen as secretion targets in this study. They originate from different *Bacillus* species and were

Table 1. Amylases used in this research.

Amylase variants	Organism	Molecular Weight (kD)	pI
Amy#707	*Bacillus sp.* 707	56.4	6.72
AmyTS-23	*Bacillus sp.* TS-23	67.3	6.41
AmyBS	*Bacillus subtilis* 168	70.2	5.88
AmyBm	*Bacillus megaterium*	56.5	5.70
AmyK38	*Bacillus sp.* KSM-K38	56.3	4.77

genetically codon optimized. They have either low-, neutral- or high- pI, and were designed and synthesized by DSM and Genencor (Table 1).

The α-amylase variants have different physicochemical properties. Mature α-amylase proteins (signal sequence cleaved) have molecular masses ranging from 55 kDa to 71 kDa, with pIs ranging between 4.77–6.72. They were expected to differentially interact with the cell envelope, finally resulting in a difference of amylase accumulation in the supernatants. They are all publically available and have a good possibility to be expressed and active in parental strain *B. subtilis* DB104. Their secretion capacity was evaluated with respect to the cell surface modification.

α-Amylases are secreted with different efficiencies in various cell envelope backgrounds

As mentioned previously, the secretory proteins have to get across the two structural hurdles (cell membrane and cell wall) to be released into the media. During this process, the secreted targets will inevitably interact with various components of the cell envelope. Accordingly, the reduction of this kind of interaction for achieving better secretion potential of the cell factory has a high feasibility.

To begin with, we carefully chose six cell surface relevant enzymes as the modulation targets. Among these candidates, the phosphatidylserine synthase (PssA) and cardiolipin synthase

(ClsA) are responsible for the synthesis of two major membrane phospholipids phosphatidylethanolamine (PE) and cardiolipin (CL) [22], respectively. The other four of the candidates all play important roles in cell wall composition and functionality. The teichoic acid linkage unit synthase (TagO) catalyzes the synthesis of wall teichoic acids; the lipid carrier sugar transferase (TuaA) is involved in the teichuronic acids formation; DltA, D-alanyl-D-alanine carrier protein ligase, catalyzes the alanylation of lipoteichoic acid; D-alanyl-D-alanine carboxypeptidase (DacA), mediates the crosslinking of peptidoglycan [20, 23, 24]. A null mutant library ($PssA^-$, $TagO^-$, $ClsA^-$, $DltA^-$, $TuaA^-$, $DacA^-$) was successfully constructed, and the effect of these cell surface alterations on the secretion of target proteins was subsequently analyzed and quantified by enzymatic assays.

In comparison with the wild-type strain JMM8901, most of the deletion mutants could significantly produce more α-amylase TS-23, Bs and Bm into the media, while very low and similar levels of the α-amylases #707 and K38 were secreted from all the hosts, with a yield of around 4 CU/ml (**Fig. 1**). Among various expression hosts, the $ClsA^-$ performed remarkably well in the production of the reporter protein. Compared to WT extracellular concentrations, it generated 47% more of AmyTS-23 and 43% more of AmyBs, while the $PssA^-$ strain showed increased production for two of the α-amylase variants (TS-23-32%, BS-39%). In contrast to WT, all of the other four mutants were capable of releasing the elevated amount of AmyBs, and in the $TagO^-$ and $DacA^-$ background, while AmyTS-23 could also get more than 30% increased extracellular release. To sum up, the enzymatic assay results suggested that inactivation of the aforementioned cell surface components was helpful for excreting the secretory proteins. The $PssA^-$ and $ClsA^-$ mutants were the two best expression hosts with higher yields of amylases and were prepared for further analysis.

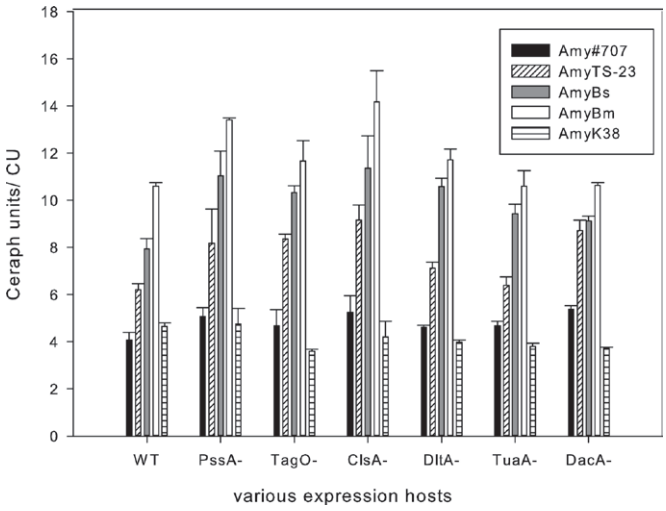


Fig. 1. The amylase activity of different variants from various *B. subtilis* deletion mutants and WT. The same OD equivalents of the culture samples were harvested after overnight growing in LB under 37 °C, 220 rpm and the amylase activities were determined by Ceralpha assays. Different patterns of the columns correspond to various amylases. Each column represents the mean ± SD of three independent experiments, and each assay was performed in duplicate.

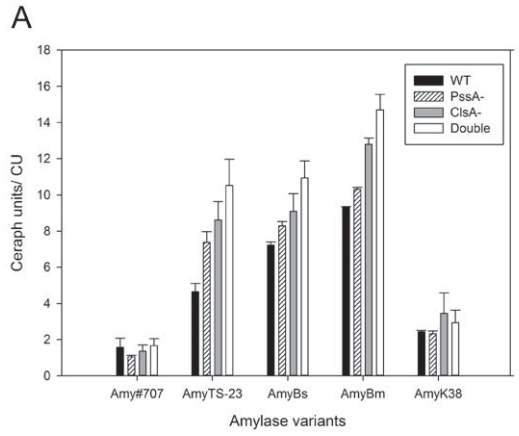
Specific alteration of cell membrane components and more negative charge of α-amylases could facilitate secretion yields

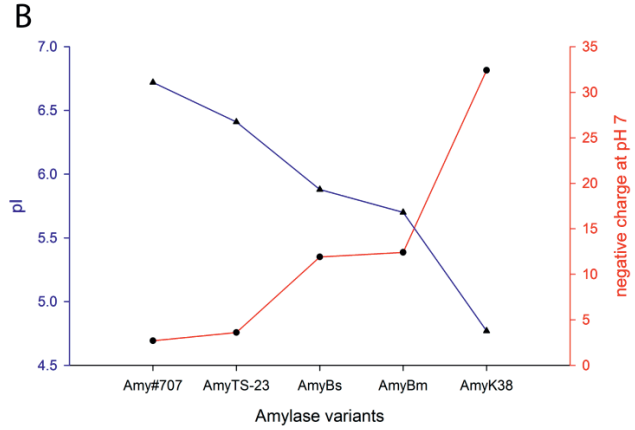
For further studying the involvement of PssA and ClsA on α-amylase secretion, a double knockout was constructed as previously described. Then, the yield of the above five α-amylases in the three mutant hosts was measured in comparison to WT host JMM8901.

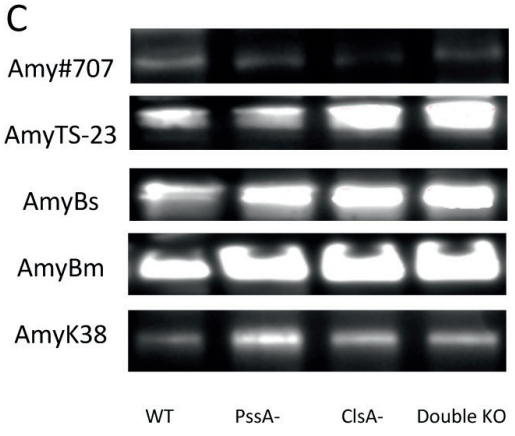
The protein samples that were harvested from the media after induction and expression were subsequently prepared for activity assays and Western blotting. As indicated in **Fig. 2A**, the accumulation of various α-amylases was quite different in the

growth media, and the α-amylase yields showed a high correlation with the net charge of the target proteins. Generally, the higher negative charge, the more corresponding α-amylases were secreted. This might result from the weaker interaction with the cell surface, as the yields of the α-amylases TS-23, Bs and Bm all went up gradually with the increase of associating negative charge. For the similar reason, the lowest producing α-amylase, Amy#707, had the highest pI and lowest negative charge (**Fig. 2B**), and the low production is probably because this α-amylase can strongly interact with cell envelope components, hampering its traverse through. Surprisingly, the secretion of AmyK38, which should be even higher than AmyBm based on pI and charge, turned out to be only a little better than that of Amy#707. This might be owing to the strong binding to some positively charged components or particles causing a severe retardation of AmyK38 in the cell.

Expectedly, when compared with JMM8901, the PssA$^-$ and ClsA$^-$ strains could significantly enhance the secretion of α-amylases, while the double KO even further improved the increase of protein yields for α-amylase TS-23, Bs and Bm. In other words, combining these two beneficial modifications would have an additive effect on the α-amylase production. Furthermore, these improvements can also be visualized by protein immunoblotting. As clearly shown in **Fig. 2C,** from left to right, they are WT, PssA$^-$, ClsA$^-$ and double KO, respectively. The α-amylase TS-23, Bs, Bm, were consistently accumulated in the supernatants at an elevated level, while the production of α-amylase K38 and #707 had little difference among different secretion hosts, which suggested a high consistency with preceding enzymatic activity assays. To conclude, the target proteins α-amylase TS-23, Bs and Bm showed a clear trend for a better secretion in correlation to a higher negative charge, and the double deletion host was the best producer of these α-amylases.


A
Ceraph units/ CU
WT
PssA-
ClsA-
Double
Amy#707
AmyTS-23
AmyBs
AmyBm
AmyK38
Amylase variants


B
pI
negative charge at pH 7
Amy#707
AmyTS-23
AmyBs
AmyBm
AmyK38
Amylase variants


C
Amy#707
AmyTS-23
AmyBs
AmyBm
AmyK38
WT
PssA-
ClsA-
Double KO

Anionic membrane lipids go up greatly in the absence of PssA and ClsA

Considering the direct responsibility of ClsA and PssA for the phospholipid synthesis of CL and PE, the single or double deficiency of them can probably alter the lipid composition and cell envelope net charge greatly, which finally resulted in a better performance of the secreted targets. To reveal how the absence of PssA and ClsA influences the phospholipid composition, lipid analysis was performed on four strains (JMM8901, PssA$^-$, ClsA$^-$ and the double KO) in stationary phase. The extracted lipid samples from the same amount of various cell pellets were loaded in parallel on the TLC plate. Subsequently, the quantification was done by analysis of the TLC plate with software ImageJ [25].

The phospholipids analysis demonstrated that all the mutants showed an obvious increase of PG content during the stationary growth phase when compared to its wild-type control JMM8901. The negatively charged phospholipids PG + CL account for 63.9% of total phospholipids in WT. However, the deletion of *pssA* and *clsA* separately raise this percentage to 77.85% and 82.14% respectively. More interestingly, the lack of these two genes in one strain caused an extreme proportion of PG, i.e. almost 90% of the total lipids (**Fig. 3**). Moreover, the PssA$^-$ and ClsA$^-$ mutants completely abolished the synthesis of PE and CL respectively, while the double KO resulted in a deficiency of both phospholipids. Taken together, in the null mutants, not only the specific corresponding phospholipid but also others

Fig. 2. The α-amylase activity in selected expression hosts grown to stationary phase in LB. (A) From left to right, they are the different α-amylases; for each α-amylase, the column pattern represents different hosts. (B) The pI and negative charge of mature α-amylase proteins. The blue line represents pIs, and the red line represents negative charge at pH 7. (C) Western blotting of α-amylase products in the media using an antibody against *Strep*-tag.

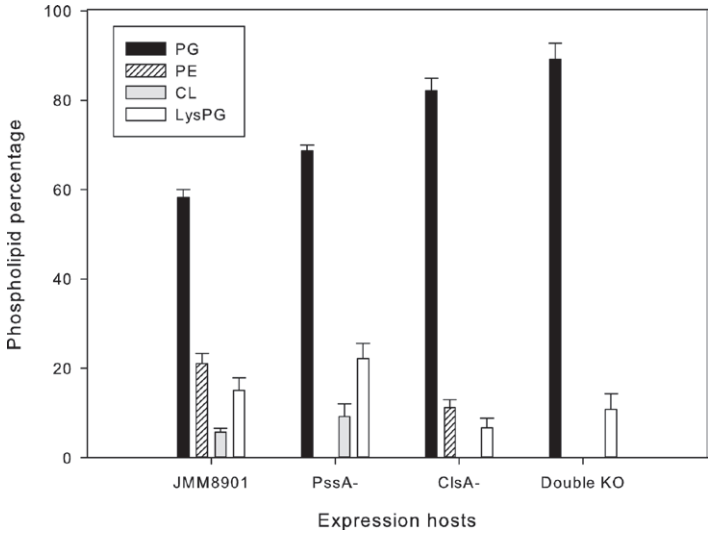


Fig. 3. Phospholipid composition of *B. subtilis* wild-type JMM8901, and PssA-, ClsA-, double KO mutant strains in stationary phase. For each strain, the various column color represents the different single lipid class, phosphatidylglycerol (PG), phosphatidylethanolamine (PE), cardiolipin (CL) and lysyl-phosphatidylglycerol (LysPG). The values represent the mean ± SD of three independent experiments.

were significantly altered, and especially the increase of the anionic phospholipid PG was observed. This suggests that these modifications could improve the availability of these anionic phospholipids, and thereby affect the electrostatic interactions between protein and cell membrane, and hence influence the rate of protein secretion.

DISCUSSION

Over the last few decades, numerous attempts have been made to overproduce heterologous recombinant proteins in *Bacillus subtilis*. The conventional approaches have already more or less

reached their limits, but there is still a great need for improved heterologous protein secretion systems. *B. subtilis* DB104, a derivative of strain *B. subtilis* 168, which only expresses 4% of the extracellular protease activity in comparison to its parental strain, shows an excellent expression and secretion advantage when being used as a cell factory of industrially important extracellular enzymes [28, 29]. Cell envelope engineering has recently been shown to be beneficial to increase yields in industrial processes, such as the production of enzymes, biofuels and chemicals, and this powerful approach is perfectly suitable for novel protein engineering and directed evolution strategies combined high-throughput screening [30, 31]. Inspired by this, cell envelope engineering is employed in this research to alter the properties of cell surface components with the goal of enhancing heterologous protein production in *B. subtilis* DB104. Moreover, our study is also providing new insights into the critical factors of the secretion process, the pIs of the secreted proteins and characteristics of the cell surface that are crucial for protein secretion efficiency.

The correlation between the cell surface and the secretion ability of the bacterium is rarely known and will be explored by modulating the cell surface composition and functionality. Firstly, we genetically modified the cell surface by the deletion of six selected genes, and the functional consequences on the α-amylase yields had been assessed by enzymatic assays. The enzymatic assays suggest that the absence of ClsA/PssA enhances the α-amylase production and these beneficial effects can be additive in the double KO. This secretion advantage was strongly suspected as the proper response to the phospholipids composition altering because previous phospholipid analysis of various isogenic hosts suggested a correlation between anionic lipids and secreted proteins in the supernatants. Consistently, it has been reported that the ClsA$^-$ mutant has a higher negative

charge density on the cell membrane because of the significant rise of PG [13].

Additionally, the importance of target proteins with specific physicochemical properties in the secretion process was also investigated. The secretion efficiency of α-amylases with different pIs was tested in the various mutant strains, and the enzymatic activity assay in combination with Western Blotting indicated that the net charge plays an important role in the protein secretion. This led us to the speculation that in general, a higher negative membrane charge, the more proteins get released into the media. Among the five α-amylase proteins studied, the amyK38 is an exception, with the production level being much lower than expected. That is possibly owing to its net charge and pI being too far away from others, which might either lead to a tight binding to some unknown positively charged intracellular- or cell surface components or local repulsion by the strong negative charges in the membrane. Surprisingly, despite the relatively small difference between Amy#707 and AmyTS-23 in pIs (6.72 and 6.41) and negative charges at pH 7.0 (2.7 and 3.6), the amylase TS-23 can still be accumulated in the supernatant much more than amylase #707. This might be because of other physiochemical properties of these secreted proteins, or other, yet unidentified factors, influence the overall production yield. In fact, more α-amylases than the above five were investigated in this study. We also engineered different α-amylases with increased positive charge, but unfortunately, they showed very low expression and secretion levels and were very difficult to be detected by our enzymatic assay technique, so we could draw no sound conclusions from this part of the study.

In conclusion, we aimed to improve the use of *B. subtilis* as a cell factory by manipulating and characterizing several factors that influence its protein secretion efficiency. Particularly, we explored the role of electrostatic interactions between the

membrane phospholipids and the secreted protein. We managed to design tailor-made *B. subtilis* production strains with enhanced potential for protein secretion, exemplified by various α-amylases and identifying their pI as an important determinant.

CONCLUSIONS

In *Bacillus subtilis,* during the protein secretion process, the characteristics of membrane phospholipid bilayer and the pIs of heterologous α-amylases determine the electrostatic interaction between the cell surface and secreted proteins and hence influence the secretion efficiency of the α-amylases variants. Consequently, the secretion barrier could be lowered effectively by engineering the cell membrane components and the secreted targets, which finally enhance the secretion yield of α-amylases significantly. In other words, the modification of these two factors provides a large advantage for further improving the protein secretion of *B. subtilis* as a cell factory.

MATERIALS AND METHODS

Bacterial strains, plasmids, and media

All the *Bacillus subtilis* and *E. coli* were grown at 37 °C with shaking (220 rpm) in liquid Lysogeny Broth (LB). The antibiotics were added when necessary as follows: 100 mg/ml ampicillin for *E. coli,* 5 mg/ml kanamycin, erythromycin, and chloramphenicol, 100 mg/ml spectinomycin, 6 mg/ml tetracycline for *B. subtilis*. For solid media, 1.5% (wt/vol) agar was added to the LB. *Bacillus subtilis* was naturally transformed using the Spizizen minimal media (SMM) supplemented with 0.5% glucose and 50 μg/ml tryptophan [26]. For induction, 0.1% subtilin-containing supernatant

Table 2. The main strains and plasmids used in this study.

Strain or plasmid	Genotype or properties	Reference or source
B. subtilis		
DB104	*his nprE aprE*	[28]
JMM8900	DB104 *thrC*::*spaRK*, Em^r	This study
JMM8901	DB104 *thrC*::*spaRK*, Em^r, *amyE*::Spc^r	This study
JMM8901- *dacA*	DB104 *thrC*::*spaRK*, Em^r, *amyE*::Spc^r, *dacA*::Km^r	This study
JMM8901- *dltA*	DB104 *thrC*::*spaRK*, Em^r, *amyE*::Spc^r, *dltA*::Km^r	This study
JMM8901- *tagO*	DB104 *thrC*::*spaRK*, Em^r, *amyE*::Spc^r, *tagO*::Km^r	This study
JMM8901- *tuaA*	DB104 *thrC*::*spaRK*, Em^r, *amyE*::Spc^r, *tuaA*::Km^r	This study
JMM8901- *pssA*	DB104 *thrC*::*spaRK*, Em^r, *amyE*::Spc^r, *pssA*::Km^r	This study
JMM8901- *clsA*	DB104 *thrC*::*spaRK*, Em^r, *amyE*::Spc^r, *clsA*::Km^r	This study
JMM8901- *pssA +clsA*	DB104 *thrC*::*spaRK*, Em^r, *amyE*::Spc^r, *pssA*::Tet^r, *clsA*:: Km^r	This study
E.coli		
MC1061	F-, *araD*139, Δ(*ara-leu*)7696, Δ(*lac*)X74, *galU*, *galK*, *hsdR2*, *mcrA*, *mcrB1*, *rspL*	[32]
Plasmids		
pNZ8901	*PspaSpn repC repA*, Cm^r	[27]
pNZ8901-Amy#707	pNZ8901 carrying *amy#707*	This study
pNZ8901-AmyTS-23	pNZ8901 carrying *amyTS-23*	This study
pNZ8901-AmyBs	pNZ8901 carrying *amyBs*	This study
pNZ8901-AmyBm	pNZ8901 carrying *amyBm*	This study
pNZ8901-AmyK38	pNZ8901 carrying *amyK38*	This study
pUC21	Amp^r *lacZ*	[33]
pUC21_DacAKO	pUC21_*dacAUp*_Km^r_*dacADown*	This study
pUC21_DltAKO	pUC21_*dltAUp*_Km^r_*dltADown*	This study
pUC21_TagOKO	pUC21_*tagOUp*_Km^r_*tagODown*	This study
pUC21_TuaAKO	pUC21_*tuaAUp*_Km^r_*tuaADown*	This study
pUC21_PssAKO	pUC21_*pssAUp*_Km^r_*pssADown*	This study
pUC21_ClsAKO	pUC21_*clsAUp*_Km^r_*clsADown*	This study
pUC21_PssAKO(Tet^r)	pUC21_*pssAUp*_Tet^r_*pssADown*	This study

of *B. subtilis* ATCC6633 was added for activation of the subtilin-regulated gene expression (SURE) system [27]. The main strains and plasmids used in this study are listed in Table 2.

DNA manipulations

DNA manipulations (purification, digestion, and ligation) were carried out as previously described [34]. T4 DNA ligase, Fastdigest Restriction enzymes and DNA polymerases (Phusion and DreamTaq) were purchased from Thermo Fisher Scientific (Bremen, Germany). Chromosomal DNA of the *B. subtilis* strain DB104 or plasmids constructed in this research were used as templates for PCR. Oligonucleotides were synthesized from Biolegio (Nijmegen, the Netherlands). Plasmid isolation was performed with the Roche high pure plasmid isolation kit, and all the constructs were sequence-verified by Macrogen Europe (Amsterdam, the Netherlands).

Construction of the null mutant library

To construct the mutant library, a component of the SURE system, *spaRK* was introduced into the *thrC* locus of parental strain *B. subtilis* DB104 by double recombination, and the obtained strain was named JMM8900. In a similar manner, we replaced the original amylase gene (*amyE*) with Spec^r to get the strain JMM8901. Subsequently, a variety of KO plasmids was made as follows: ~1 kb of both upstream and downstream flanking regions were amplified by PCR with Pfux7 DNA polymerase (a kind gift from Bert Poolman, University of Groningen), and the background of plasmid pUC21 and antibiotic resistance genes were also amplified. After that, they were ligated together using the uracil-excision DNA engineering method [35]; the ligation mixture was directly transformed into competent *E.coli* MC1061. Afterward, the positive ones were selected out for deleting corresponding genes in JMM8901, and the target genes (*pssA, tagO, clsA, dltA, tuaA, dacA*) were replaced by an antibiotic resistance gene (Km^r) separately, and thus a library of single null mutants was successfully obtained. In this way, we made the ClsA+PssA double

KO strain as well. All the constructs and strains were verified by PCR and DNA sequencing.

Construction of the expression vector

The vector that we used for expression includes a variety of functional components: a subtilin-inducible promoter, a common RBS sequence and terminator, an identical signal peptide sequence of *amyk38*, various amylase-coding regions without the signal peptide, and a *Strep*-tag at the C-terminus, which can be used for further analysis, like protein purification and Western blotting. All these fragments were combined step by step via overlap PCR and ligation. In this vector, all the factors mentioned above are identical, except the mature α-amylase coding sequences (**Fig. 4**).

α-Amylase activity assays

Strains were grown in LB overnight, after which the cells were collected by centrifugation (10,000 rpm, 1 minute) and the pellets were suspended and diluted in 200 µl LB (1:50) to start the growth in 96 well plates, the inducer subtilin was added after 3 hours' preculture (1:1000). The optical density was read automatically every 15 minutes by Infinite 200 plate reader (Tecan, Switzerland) during the whole incubation process. Subsequently, the supernatants were harvested by high-speed centrifugation (14,000 rpm, 1 minute) and immediately frozen at -20 °C. The next day, the α-amylase activity was assessed based on the Ceralpha HR Kit manual (Megazyme, Wicklow, Ireland) as follows: 110 µl substrate BPNPG7 solution buffer (54.5 mg BPNPG7 substrate was dissolved in dilution buffer to get 40 ml working substrate) was mixed to 10 µl aliquots of collected supernatant, 140 µl stop buffer (200 mM Boric acid/NaOH buffer, pH 10) was added to terminate the reaction after incubating 4 minutes at room temperature. The tube contents were stirred vigorously

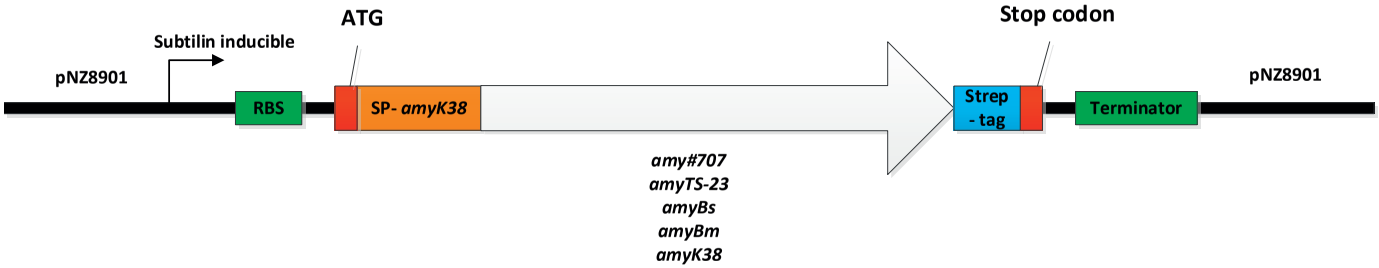


Fig. 4. The α-amylase expression vector. The construct contains a subtilin inducible promoter, RBS sequence, signal sequence (SP-*amyk38*), various α-amylase sequences, a *Strep*-tag before the stop codon, terminator, and this whole combination was ligated to the background plasmid pNZ8901.

before the 405 nm absorbance was read, and the final amylase activity was calculated by the Ceralpha Unit.

SDS-PAGE and immunoblotting

The same OD equivalents of previously harvested samples were prepared for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting. Firstly, the proteins were separated by SDS-PAGE and transferred onto the polyvinyl difluoro (PVDF) membranes (Millipore, Bedford, MA, USA) using a Mini Trans-Blot system (Bio-Rad) [10], which was then followed by 4 °C overnight incubation in 5% BSA. The next day, the membranes were washed three times 10 minutes with PBST buffer before incubated with the *Strep*-tag antibody (1:5000; IBA) at room temperature for 90 minutes, and then the membranes were washed twice with PBST. Finally, the signal density was visualized using the freshly mixed western blotting detection reagents (GE Healthcare Life Sciences) and detected by the Molecular Imager ChemiDoc XRS+ (Bio-Rad).

Total lipid extraction and quantification

The extraction of total lipids was performed as reported previously [13, 36]. The same OD equivalents of overnight cultures were harvested by high-speed centrifugation. Then the cells were subjected to sonication for 30 minutes after washing twice with MilliQ water. Next, the obtained extracts were incubated for 15 minutes at 54 °C with the same volume of methanol (v/v), followed by the addition of 2 volumes chloroform (plus 1 ml 6 M HCl to improve the lysPG extraction) [37]. After 16 hours' incubation at room temperature, the mixture was centrifuged for 10 minutes at full speed. The lower phase containing the lipids was then taken and transferred to a new tube, which was left open and evaporated until dry in the fume hood. The thin-layer chromatography (TLC) with Plate Silica Gel 60 F254 (Merck) was

carried out for the lipid quantification. The previously dried samples were suspended in 500 µl of chloroform (plus 50 µl 6 M HCl) and then developed with the solvent system acetone/hexane (1:6, v/v) to separate the lipid classes. The purified standards CL, PG, PE, and lysPG, which have been reported as the most abundant lipids in *B. subtilis* membranes were also loaded on the plates. Afterward, the phospholipids were visualized with iodine vapor and analyzed by ImageJ (NIH) software.

ABBREVIATIONS

PssA: Phosphatidylserine synthase; ClsA: Cardiolipin synthase; DacA: D-alanyl-D-alanine carboxypeptidase; TuaA: Lipid carrier sugar transferase; DltA: D-alanyl-D-alanine carrier protein ligase; TagO: Teichoic acid linkage unit synthase; SMM: Spizizen minimal media; SURE: Subtilin-regulated gene expression; LB: Lysogeny Broth; Amp: Ampicillin; Spc: Spectinomycin; Em: Erythromycin; Km: Kanamycin; Cm: Chloramphenicol; Tet: Tetracycline; SDS-PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis; PVDF: Polyvinyl difluoro; TLC: Thin layer chromatography; CL: Cardiolipin; PE: Phosphatidylethanolamine; PG: Phosphatidylglycerol; LysPG: Lysyl-phosphatidylglycerol; pI: Isoelectric point.

DECLARATIONS

Authors' contributions

OPK conceived, supervised and coordinated the project. OPK, MM, HC, and AJH conceived and designed the experiments. MM constructed the expression vectors and most of the mutants, HC performed the Western blotting, amylase activity assays, lipid

composition measurement, and data analysis. HA did part of amylase activity assays and data analysis. HC drafted the first manuscript which was later revised and corrected by AJH and OPK. All authors read and approved the final manuscript.

Acknowledgments

We thank Dr. Marc Kolkman (Genencor) and Dr. Tjeerd van Rij (DSM) for providing materials and for helpful discussions. We thank Dr. Ruud Detert Oude Weme and Dr. Lance Keller (Molecular Genetics, University of Groningen) for proofreading the manuscript. We are grateful to Xiaoming Miao (System Chemistry, University of Groningen) for technical assistance with the lipid composition quantification.

Funding

This work was financially supported by a grant from Stichting Technische Wetenschappen (STW), projectnr. 10718.

REFERENCES

1. Harwood CR. *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. *Trends Biotechnol*. 1992;10:247–256.
2. Harwood CR, Cranenburgh R. *Bacillus* protein secretion: an unfolding story. *Trends Microbiol*. 2008;16:73–79.
3. Westers L, Westers H, Quax WJ. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochim Biophys Acta.* 2004;1694:299–310.
4. Brockmeier U, *et al.* Systematic screening of all signal peptides from *Bacillus subtilis*: a powerful strategy in optimizing heterologous protein secretion in Gram-positive bacteria. *J Mol Biol*. 2006;362:393–402.
5. Morimoto T, *et al.* Enhanced recombinant protein productivity by genome reduction in *Bacillus subtilis*. *DNA Res*. 2008;15:73–81.

6. Fu LL, *et al.* Protein secretion pathways in *Bacillus subtilis*: implication for optimization of heterologous protein secretion. *Biotechnol Adv*. 2007;25:1–12.
7. Nijland R, Kuipers OP. Optimization of protein secretion by *Bacillus subtilis*. *Recent Pat Biotechnol*. 2008;2:79–87.
8. Pohl S, Harwood CR. Heterologous protein secretion by *Bacillus* species: From the cradle to the grave. *Adv Appl Microbiol*. 2010;73:1–25.
9. Kakeshita H, *et al.* Enhanced extracellular production of heterologous proteins in *Bacillus subtilis* by deleting the C-terminal region of the SecA secretory machinery. *Mol Biotechnol*. 2010;46:250–257.
10. Vitikainen M, *et al.* Secretion of heterologous proteins in *Bacillus subtilis* can be improved by engineering cell components affecting post-translocational protein folding and degradation. *J Appl Microbiol*. 2005;99:363–375.
11. Thwaite JE, *et al.* Optimization of the cell wall microenvironment allows increased production of recombinant *Bacillus anthracis* protective antigen from *B. subtilis*. *Appl Environ Microbiol*. 2002;68:227–234.
12. Degering C, *et al.* Optimization of protease secretion in *Bacillus subtilis* and *Bacillus licheniformis* by screening of homologous and heterologous signal peptides. *Appl Environ Microbiol*. 2010;76:6370–6376.
13. Lopez CS, *et al.* Role of anionic phospholipids in the adaptation of *Bacillus subtilis* to high salinity. *Microbiology*. 2006;152:605–616.
14. Lopez CS, *et al.* Variations of the envelope composition of *Bacillus subtilis* during growth in hyperosmotic medium. *Curr Microbiol*. 1998;36:55–61.
15. Kawai F, *et al.* Cardiolipin domains in *Bacillus subtilis* marburg membranes. *J Bacteriol*. 2004;186:1475–1483.
16. Campo N, *et al.* Subcellular sites for bacterial protein export. *Mol Microbiol*. 2004;53:1583–1599.
17. Mikhaleva NI, *et al.* Requirement for phospholipids of the translocation of the trimethylamine N-oxide reductase through the Tat pathway in *Escherichia coli*. *FEBS Lett*. 1999;463:331–335.
18. Hyyrylainen HL, *et al.* D-Alanine substitution of teichoic acids as a modulator of protein folding and stability at the cytoplasmic membrane/cell wall interface of *Bacillus subtilis*. *J Biol Chem*. 2000; 275:26696–26703.

19. Ghuysen JM. Molecular structures of penicillin-binding proteins and β-lactamases. *Trends Microbiol*. 1994;2:372–380.
20. Atrih A, *et al.* Analysis of peptidoglycan structure from vegetative cells of *Bacillus subtilis* 168 and role of PBP 5 in peptidoglycan maturation. *J Bacteriol*. 1999;181:3956–3966.
21. Stephenson K, *et al.* The influence of secretory-protein charge on late stages of secretion from the Gram-positive bacterium *Bacillus subtilis*. *Biochem J*. 2000;350:31–39.
22. Salzberg LI, Helmann JD. Phenotypic and transcriptomic characterization of *Bacillus subtilis* mutants with grossly altered membrane composition. *J Bacteriol*. 2008;190:7797–7807.
23. Soldo B, *et al.* Teichuronic acid operon of *Bacillus subtilis* 168. *Mol Microbiol*. 1999;31:795–805.
24. Brown S, *et al.* Wall teichoic acids of gram-positive bacteria. *Annu Rev Microbiol*. 2013;67:313–336.
25. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9:671–675.
26. Anagnostopoulos C, Spizizen J. Requirements for transformation in *Bacillus subtilis*. *J Bacteriol*. 1961;81:741–746.
27. Bongers RS, *et al.* Development and characterization of a subtilin-regulated expression system in *Bacillus subtilis*: strict control of gene expression by addition of subtilin. *Appl Environ Microbiol*. 2005;71:8818–8824.
28. Kawamura F, Doi RH. Construction of a *Bacillus subtilis* double mutant deficient in extracellular alkaline and neutral proteases. *J Bacteriol*. 1984;160:442–444.
29. Vojcic L, *et al.* An efficient transformation method for *Bacillus subtilis* DB104. *Appl Microbiol Biotechnol*. 2012;94:487–93.
30. Saeui CT, *et al.* Cell surface and membrane engineering: emerging technologies and applications. *Funct Biomater*. 2015;6:454–485.
31. Ueda M. Establishment of cell surface engineering and its development. *Biosci Biotechnol Biochem*. 2016;80:1243–1253.
32. Casadaban MJ, Cohen SN. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J Mol Biol*. 1980;138:179–207.

33. Vieira J, Messing J. New pUC-derived cloning vectors with different selectable markers and DNA replication origins. *Gene*. 1991;100:189–194.
34. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.
35. Norholm MH. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnol*. 2010;10:21.
36. Garcia Montes de Oca LY, *et al.* Cardiolipin deficiency causes a dissociation of the *b6c:caa3* megacomplex in *B. subtilis* membranes. *J Bioenerg Biomembr*. 2016;48:451–467.
37. Machado MC, *et al.* Osmotic response in *Lactobacillus casei* ATCC 393: biochemical and biophysical characteristics of membrane. *Arch Biochem Biophys*. 2004;422:61–70.

CHAPTER 3

Boosting heterologous protein production yield by adjusting global nitrogen and carbon metabolic regulatory networks in *Bacillus subtilis*

Haojie Cao[1], Julio Villatoro-Hernandez[1], Ruud Detert Oude Weme[1], Elrike Frenzel[1], Oscar P. Kuipers[1]

[1]Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands.

Based on the manuscript that is under review in *Metab Eng*: Haojie Cao, Julio Villatoro-Hernandez, Ruud Detert Oude Weme, Elrike Frenzel, Oscar P. Kuipers. Boosting heterologous protein production yield by adjusting global nitrogen and carbon metabolic regulatory networks in *Bacillus subtilis*.

ABSTRACT

Bacillus subtilis is extensively applied as a microorganism for the high-level production of heterologous proteins. Traditional strategies for increasing the productivity of this microbial cell factory generally focused on the targeted modification of rate-limiting components or steps. However, the longstanding problems of limited productivity of the expression host, metabolic burden and non-optimal nutrient intake, have not yet been completely solved to achieve significant production-strain improvements. To tackle this problem, we systematically rewired the regulatory networks of the global nitrogen and carbon metabolism by random mutagenesis of the pleiotropic transcriptional regulators CodY and CcpA, to allow for optimal nutrient intake, translating into significantly higher heterologous protein production yields. Using a β-galactosidase expression and screening system and consecutive rounds of mutagenesis, we identified mutant variants of both CcpA and CodY that in conjunction increased production levels up to 290%. Analysis of the obtained phenotypes revealed crucial mutations within the DNA-binding helix-turn-helix (HTH) domain of both CodY (R214C) and CcpA (T19S). The improved cell factory capacity was further demonstrated by the successfully increased overexpression of GFP, xylanase and peptidase in the double mutant strain. Taken together, the amino acid substitutions in the conserved region of the master transcriptional regulators eventually allowed the *B. subtilis* cell factory to synthesize higher amounts of heterologous proteins.

Keywords: *Bacillus subtilis*, microbial cell factory, CcpA, CodY, global transcription machinery engineering (gTME), high-throughput screening, heterologous proteins

INTRODUCTION

Bacillus subtilis is a well-characterized microbial cell factory and is widely used in the production of a variety of proteins for commercial and medical applications [1–4]. Improving the production potential of this classic chassis has been a research focus for several decades. Numerous engineering and biotechnological approaches have been employed in attempts to enhance production yields in industrial strains, for instance by utilizing modified promoters and RBSs, codon-optimization, pathway rerouting or gene disruption [5–7]. However, although remarkable progress has been made in improving the protein overproduction capacity of *B. subtilis*, the space for traditional techniques or strategies to further improve this host organism's productivity is increasingly limited.

In nature, the intracellular distribution of various resources in healthy cells has been 'optimized' by natural evolution over very long periods of time [8]. Introducing an overexpression pathway for heterologous proteins into an engineered organism requires a large proportion of the host cell's resources, including ATP, carbohydrates and amino acids. This imposed metabolic drain has been defined as 'metabolic burden' or 'metabolic load' [9]. In this case, the vast majority of the intracellular metabolic fluxes, including energy resources such as NAD(P)H and ATP and carbon/nitrogen/oxygen building blocks, are forcibly assigned towards the heterologous product biosynthesis [10]. The essential requirements for cellular maintenance, in turn, become imbalanced and insufficient in the engineered microbes [11]. Therefore, the biosynthetic yield of the expressed target product will remain at a relatively low level [10, 12], or even suddenly drop into the 'death valley' (minimal production level) on a 'cliff' under suboptimal growth conditions [8]. Hence, the strategy to reduce the metabolic burden in a microbial host

by enhancing the uptake of required nutrients and balancing heterologous and native metabolic flux demands, could potentially benefit the robust production of large quantities of the target product.

In *B. subtilis*, the molecular mechanisms of nutrient-sensing based central metabolic regulations have become increasingly clear. Two global regulatory proteins, CodY and CcpA, behave as a repressor or inducer of gene expression by specifically binding to consensus sequences located in or near the promoter region of targets [13–15]. These two regulators and their ligands (FBP, GTP, and BCAAs) [16] function together to orchestrate regulons that balance the use of available nutrients sources. Thus, they systemically coordinate the intracellular carbon and nitrogen fluxes and contribute to cell homeostasis by stimulating specific cellular processes [17, 18]. Prior studies showed that global transcription machinery engineering (gTME) elicits a global alteration at the transcriptional level that perturbs the expression of multiple proteins simultaneously, which allows acquisition and selection of phenotypes of interest from a broad library [19, 20]. Some global transcription machinery components, such as sigma factors in bacteria [19, 21], zinc finger-containing artificial transcription factors [22], and Spt15 in yeast [23] were randomly mutagenized for generating phenotypes of biotechnological interest, including improved production capacities and strain tolerance towards elevated end-product levels. We thus hypothesized this strategy could be exploited to rewire the nitrogen and carbon metabolic flux distributions and to optimize nutrient uptake and utilization in *B. subtilis* at the whole-cell level to gain enhanced protein production traits by specific adjustments of the activity of CodY and CcpA. In addition, this study allows for gaining further insights into the interaction between CodY and CcpA by the analysis of the globally rewired nitrogen and carbon metabolic networks.

RESULTS

Construction of randomly mutagenized CodY libraries

The master transcriptional regulator CodY controls hundreds of genes in a large regulon, the products of which are mainly linked to nitrogen metabolism. More specifically, CodY senses the intracellular levels of BCAAs and GTP and represses or activates the transcription of nitrogen metabolic network related genes to trigger varying metabolic effects by binding to consensus function sites called the CodY box [16]. Therefore, any alteration of the CodY amino acid sequence can potentially reprogram the downstream metabolic fluxes and thus influence the production of heterologous proteins. To create a tunable overexpression system, we introduced an expression cassette containing the heterologous β-galactosidase encoding gene (*lacZ*) into an IPTG inducible system. The resulting construct P*hyspank-lacZ* was chromosomally integrated into the *mdr* locus of *B. subtilis* 168 to obtain the reporter strain *B. subtilis* 168_β-gal (**Fig. 1A**).

To construct the mutant libraries of CodY, a specific integration vector for transporting different versions of *codY* into its native locus was made by USER cloning method [24]. The obtained plasmid (pJV54), which consists of ~1k bp flanking regions, an antibiotic resistance marker and a pUC18 backbone, was applied as the template for mutagenesis experiments. First, the pJV54 was utilized as the cloning template for random mutagenesis experiments. The GeneMorph II Random Mutagenesis Kit (Agilent Technologies, United States) was applied to achieve the desired mutation rate (low-, medium-, high-) for three mutant libraries according to the manufacturer's indications (**Fig. 2A**). The mutagenized *codYs* were subsequently cloned into the previous integration vector and transformed into competent *E. coli*. In this way, three pools of randomly mutated *codY* genes with different mutation frequencies (1–4.5, 4.5–9,

9–16 mutations/kb amplified target gene) were constructed in the cloning host *Escherichia coli* (**Table 1**). We scraped off all the fresh transformants and extracted plasmid after overnight incubation. Finally, this fresh plasmid DNA library was directly transformed into competent *B. subtilis* 168_β-gal, and approximately 1,400 bp (*codY*_cmr*) of DNA fragments were introduced into the P*hyspank-lacZ* integrated reporter strain via double homologous recombination, thereby various mutagenized versions of *codY* replaced the wild-type *codY* gene (**Fig. 1B**). Correct integration was tested by colony PCR and confirmed by sequencing. In total, three independent CodY mutant libraries with a respective size of more than 5,000 cells were established. The CodY$^-$ (*codY::cmr*) were also constructed in a similar way.

Selection of CodY mutants with increased capacity of β-galactosidase production

From isolates with a widely varying protein production potential, protein overproducing candidates were selected based on an increase in β-galactosidase activity (**Fig. 2B**). The black-white screening was performed on transparent SMM plates, which allowed direct benchmarking of the color intensity and monitoring of the colony size. Phenotypes with increased β-galactosidase activity were selected out of a total of over 15,000 CodY mutant candidates, and the β-galactosidase production yields were quantified by colorimetric assays. *B. subtilis* mutants with higher reporter enzyme production originated from the low mutation frequency library (1-4.5 mutations/kb), indicating that libraries with more than 4.5 mutations/kb have a higher probability of harboring phenotypic variations that negatively affect the biosynthesis of the reporter protein. Finally, the selected four CodY mutant phenotypes outperformed the WT and CodY$^-$ by producing significantly more enzyme using as media LB supplemented with 1.0% glucose. Under

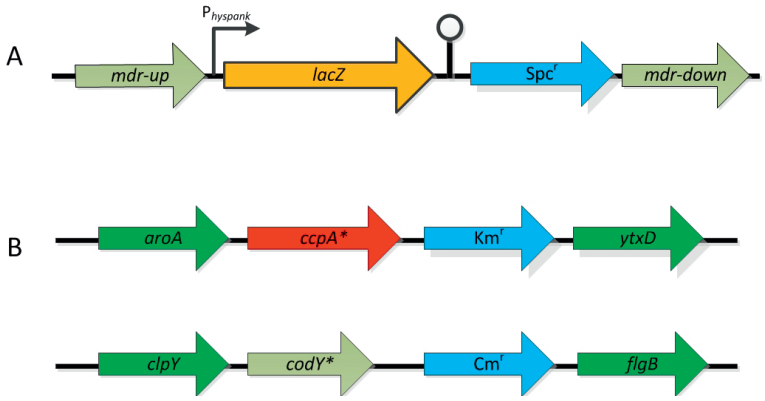
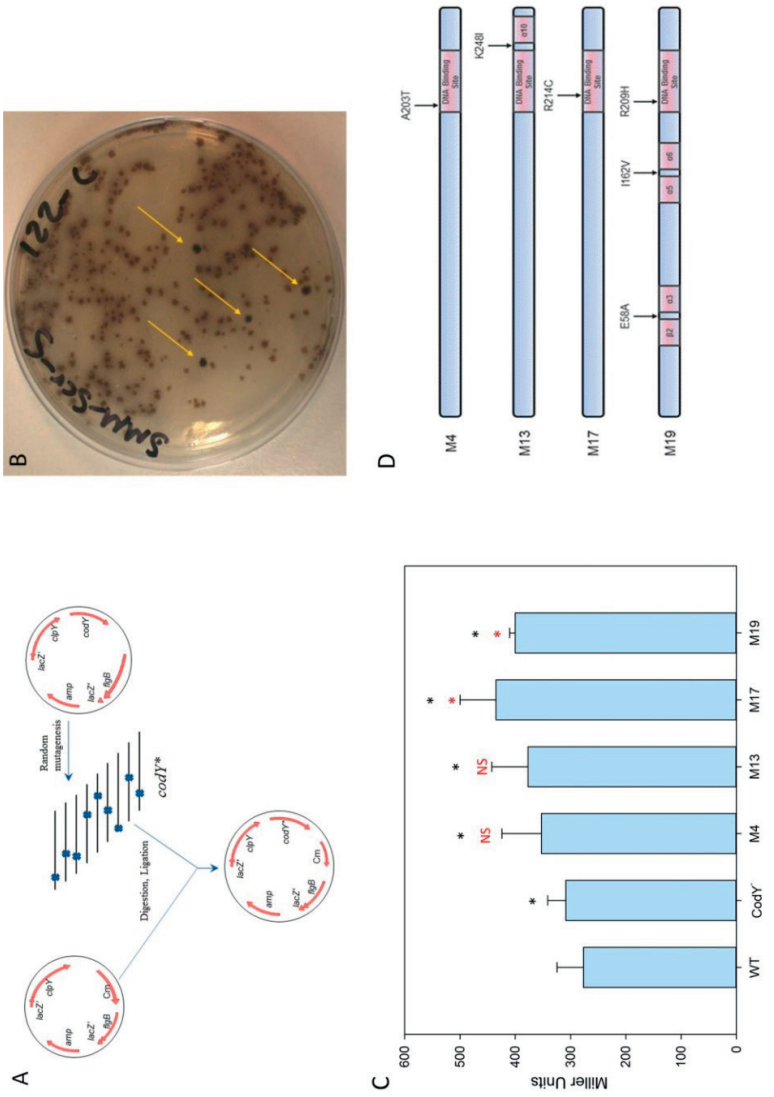


Fig. 1. (A) The composition of the reporter enzyme expression cassettes. A synthetic operon was constructed under the regulation of the IPTG-inducible promoter-P*hyspank* (black arrow) followed by the encoding sequence for the β-galactosidase and a *rho*-independent terminator. This expression construct was flanked by the up- and downstream regions of *mdr* to allow integration of a single copy of the cassette into the *mdr* locus. **(B) Gene replacement of mutant versions of CcpA and CodY.** The mutated versions of *codY* and *ccpA* were cloned in-frame flanked by two homologous regions of their native loci. The genes for the mutated CodY* and CcpA* proteins are under the regulation of their native promoters after integration into the chromosome and replacement of the WT genes.

Table 1. Mutation frequency of each library in *B. subtilis*. The mutation frequency can be determined by the amount of template and the PCR cycle numbers. The final results was verified by DNA sequencing, and the numbers of nucleotides changed per 1,000 base pair of target DNA were calculated.

Library	Mutation rate (mutations/1 kb)
CodY	
-low	1-4.5
-medium	4.5-9
-high	9-16
CcpA	
-low	1-4.5

identical culture and enzymatic assay conditions, the *codY* deficient strain could already produce around 12% higher β-gal than WT, while the most outstanding four isolates M4, M13, M17 and M19 respectively generated 27%, 36%, 52% and 40% more β-gal than the reporter strain equipped with the unmodified,



wildtype *codY* gene (**Fig. 2C**). On the other hand, sequencing analysis illustrated that all the higher β-gal producers possessed at least one amino acid substitution in the DNA-binding domain. In M4, M13, and M17, only the single mutations A203T, K248I, and R214C were detected, respectively, while M19 carried the HTH domain mutation R209H next to the amino acid exchanges E58A and I162V (**Fig. 2D**). Hence, we showed that the gTME approach coupled to an enzymatic protein production capacity screen could be successfully applied in *B. subtilis* and the specific mutations within the DNA binding domain of CodY resulted in a significantly increased product yield of the heterologous protein β-galactosidase.

CodY variants lead to different genetic competence levels

Genetic competence is a well-defined cellular state in *B. subtilis* 168 when the nutrients are depleted. Transcription factor ComK activates the expression of late competence genes and causes the entry into the competent state, allowing the cells to

Fig. 2. (A) The workflow of the mutant libraries construction for CodY. (B) Black-white screening on the selective plate. Mutants of *B. subtilis* were plated on the SMM agar media containing S-Gal in the presence of 0.1 mM IPTG. Phenotypes with a dark color and big size were isolated and further analyzed. **(C) Mutation in CodY leads to phenotypes with higher β-galactosidase production.** Cells were grown in LB media supplemented with 1.0% glucose. β-Galactosidase activities [Miller Units] represent the mean values of three samples taken during exponential growth at OD600~1.0. Strains used were WT (wild-type *codY*), CodY null mutant (*codY::cmr*), and four selected higher producing phenotypes M4, M13, M17, M19. The statistical significance of differences was performed by T-TEST, black symbols represent the comparison with WT, red words/symbols represent the comparison with $CodY^-$ (NS: no significance, *P<0.05). **(D) Genotypic characterization of selected CodY mutants.**

uptake and integrate extracellular DNA [25]. Moreover, CodY can bind to the promoter region of *comK*, resulting in a repression of ComK [26]. Therefore, various alterations in the amino acid sequence of CodY may result in diverse competence ability of the host cells. To identify the natural competence of the selected CodY mutants, a construct P*comG-gfp* was introduced by Campbell-type integration into *B. subtilis*, leaving the original *comG* operon intact. Microplate analysis was performed for the P*comG-gfp* fused strains to monitor the GFP expression in individual cells. As shown in **Fig. 3**, cells with different versions of CodY showed significant differences in the green fluorescence signal strength over time. Specifically, the mutant M19 and wildtype exhibited very similar fluorescence curves to each other, which reflects the very close competence development abilities for these two strains. Besides, M4, M13, and M17 showed lower genetic competence relative to wildtype. Consistent with the fact that CodY negatively regulates the positive competence regulator ComK [27], the deletion of CodY enhanced the corresponding genetic competence obviously (2-fold). Interestingly, the peak of CodY$^-$ occurred two hours later than wildtype, probably owing to the slower growth of the CodY$^-$ strain in the competence media (data not shown). In sum, the CodY variants caused different levels of genetic competence, as was expected.

Screening gTME libraries of CcpA in the two CodY mutants M17 and M19

To explore whether the production potential of the previously engineered expression hosts could be further improved, we additionally reprogrammed the carbon metabolic network in the two CodY mutants by random mutagenesis of the transcriptional regulator CcpA. A *ccpA** library with 1–4.5 mutations/kb mutation frequency was constructed (**Table 1**) and integrated into the chromosome of M17 and M19 to obtain two

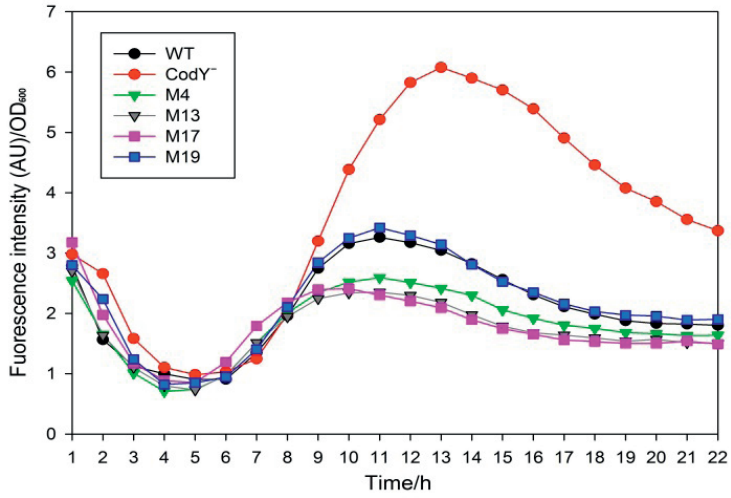
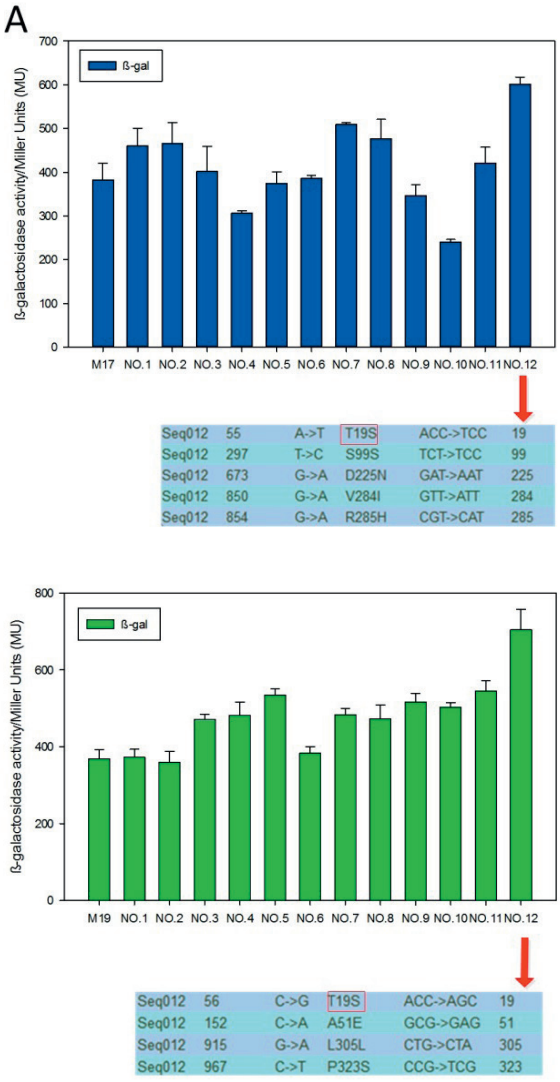


Fig. 3. The expression activity of P*comG-gfp* in various strains. Strains JV156 (wild-type), CodY⁻, M4, M13, M17, M19 with P*comG-gfp* construct were grown in the competence media MC 1X in the 96 well plates at 37 °C, 220 rpm. The OD600 and fluorescence intensity of GFP were read every hour for 22 hours in the plate reader with a GFP filter set (excitation at 485/20 nm, emission 535/25). The values of GFP fluorescence intensity/OD600 of individual time points are presented in the figure. Experiments were performed in triplicates, but for clarity, only one representative line of the mean value is shown.

libraries, M17 (*ccpA**) and M19 (*ccpA**). Subsequently, CcpA mutant strains were selected from each library by the black-white screening based on their higher β-galactosidase product yields. The two best performing phenotypes from the two different libraries showed the same amino acid exchange, T19S, in the DNA-binding HTH motif of CcpA (**Fig. 4A**). To further verify the influence of the single amino acid substitution T19S and to rule out the possibility of acquisition of additional (compensatory) mutations, we performed site-directed mutagenesis of *ccpA* to introduce the mutation T19S in different background strains, thereby obtaining the mutants CcpA^{T19S} and M17CcpA^{T19S}. Since then, the M17 and M17CcpA^{T19S} were renamed as CodY^{R214C} and



CodY^{R214C}CcpA^{T19S} in the following analyses. In comparison to the parental WT control strain 168_β-gal, all transcription factor mutants showed significantly increased β-gal activities (**Fig. 4B**). We observed that the deficiency of CcpA and/or CodY, which was introduced by gene knockouts, could already improve the β-galactosidase yield by 10–20% than that of WT strain. However,

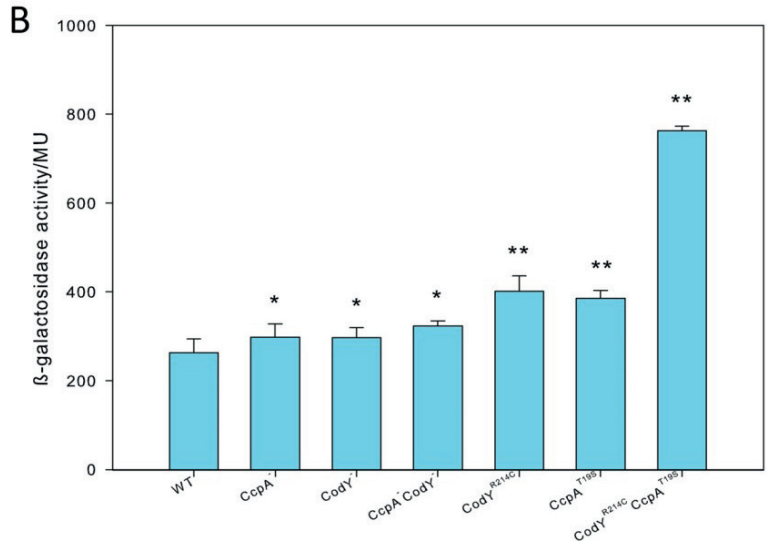
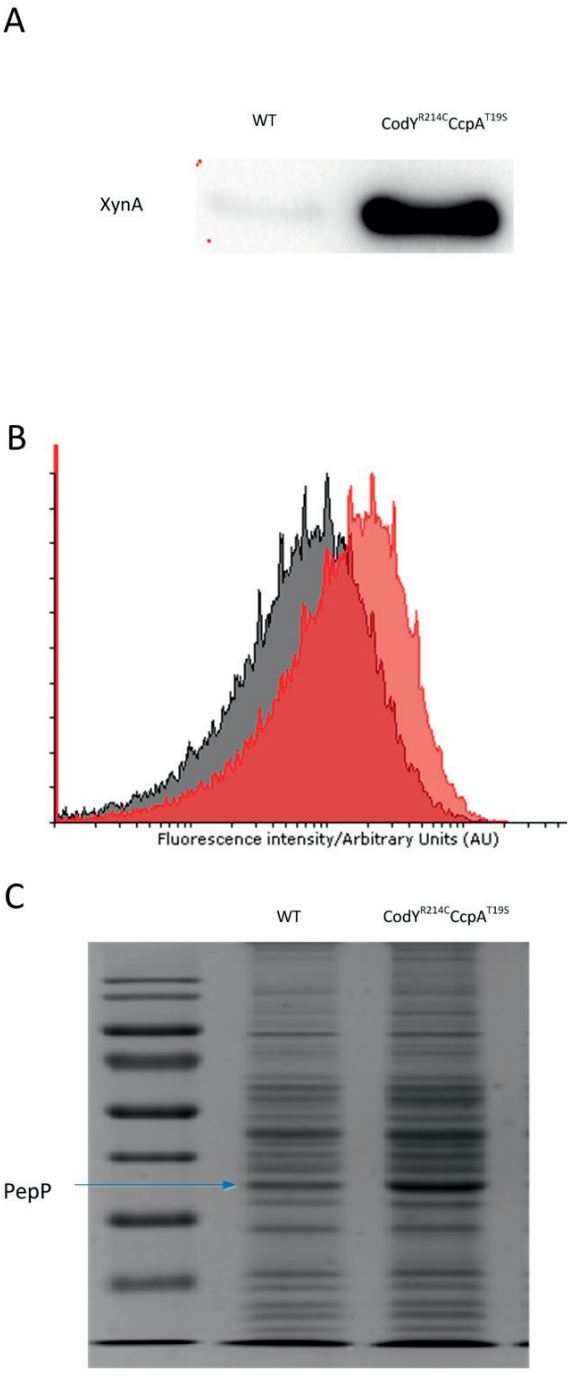


Fig. 4. (A) The screening of CcpA* library in the two CodY mutants M17 and M19. Screening of β-galactosidase higher-producing phenotypes isolated from the M17CcpA* and M19CcpA* libraries. Isolates with improved enzymatic activity were selected from the SMM screening plates, and the enzymatic assay and sequencing analysis of optimized phenotypes were carried out. **(B)** β-galactosidase activities **in *B. subtilis* strains with null mutations or single amino acid substitutions in the DNA-binding HTH domain of CodY and CcpA.** Enzymatic assay of the recombinant strains was carried out in comparison to the control strain carrying a *lacZ* gene as a capacity monitor (WT). The cultures of were sampled at an OD600 of 1.0, and the β-galactosidase activities are shown in Miller Units. Each column represents the mean ± SD of three independent experiments. The statistical significance of differences was performed by the T-TEST, the black symbols represent the comparison with the WT (*P<0.05, **P<0.01).

in the single mutants CodY^{R214C} and CcpA^{T19S}, the respective production of the target enzyme was increased to 152% and 140% relative to the WT. This indicates that single mutations in the DNA-interacting HTH domains of the global transcriptional


A
WT
CodY^{R214C}CcpA^{T19S}
XynA
B
Fluorescence intensity/Arbitrary Units (AU)
C
WT
CodY^{R214C}CcpA^{T19S}
PepP

regulators, rather than complete gene knockouts, were advantageous. Moreover, these production advantages were synergistic when the *codY* and *ccpA* mutations were combined in one strain, because heterologous protein production was enhanced to 290% in the double mutant CodY^{R214C}CcpA^{T19S}.

The production improvement in the genetically modified cell factory is achieved for various proteins

Theoretically, this tailor-made system offers potential to overexpress any heterologous protein by achieving multiple and simultaneous perturbations of the whole transcriptome and metabolome. Next to the model protein β-galactosidase, we also managed to overproduce additional heterologous proteins in the transcription factor mutant. In line with the results presented above, a significant increase of GFP (89%), XynA (xylanase protein from *B. subtilis*) (more than 11-fold) and PepP (aminopeptidase P from *Lactococcus lactis*) (80%) production could be achieved in CodY^{R214C}CcpA^{T19S} relative to the WT host (**Fig. 5**). This demonstrates that productivity improvement in the genetically modified cell factory is true for various proteins, although

Fig. 5. The production of other protein candidates in the strain CodYR214C-CcpAT19S in comparison with wildtype. All the strains were grown in LB media supplemented with 1.0% glucose, and the induction of different protein expression systems was carried out in various ways. **(A)** The expression of XynA was performed on an expression vector pNZ8902, and the inducer subtilin (0.1%) was added when the culture reaches OD600~0.6, the overnight cultures were sampled, and the generated XynA (supernatant) was analyzed by Western Blotting. **(B)** Strains containing P*hyspank-sfgfp* were grown in the production media supplemented with 0.1mM IPTG, and the cultures were collected for flow cytometric analysis, black-WT, red-CodY^{R214C}CcpA^{T19S}, the fluorescence intensity was calculated by Arbitrary Units (AU). **(C)** The PepP was expressed as similarly as GFP**,** and the relevant production level was detected by SDS-PAGE.

the levels of improvement for individual candidates were very different. The cell factory productivity can be further influenced by the intracellular nutrient availability, codon usage and the utilization bias for some specific nitrogen sources. This can also explain why the protein can be differentially expressed when growing in media with different compositions [1, 28].

DISCUSSION

To date, great progress has been achieved for heterologous protein production in *B. subtilis* by modifying the pathway regulatory elements, such as gene copy number, promoters, ribosome binding sites (RBSs), and terminators [7]. These genetic engineering strategies have mainly focused on the regional modification of specific rate-limiting factors or steps, while multiple genetic modifications, which show improved capacity to elicit phenotypes of interest, has rarely been explored. In recent years, the fast accumulation of genetic information in regulatory mechanisms and advances in genome-editing and -omics techniques provided novel engineering tools and strategies for fine-tuning metabolic pathways. An approach termed gTME (global transcription machinery engineering), which focuses on the increase of end-products by rerouting metabolic fluxes at a global level, can remarkably simplify the strain enhancement design even without a thorough understanding of the underlying metabolic regulatory mechanisms [29, 30].

In this study, we developed a combinatorial system involving random mutagenesis and high-throughput selection, which outperform the traditional approaches in expanding the engineering scale from the local pathway to the global metabolic networks [20]. The gTME-based tool demonstrated to effectively and quickly unlock desired mutants with rewired central

metabolic pathways and improved uptake and utilization of specific nitrogen sources. The best phenotype with single amino acid substitutions within the DNA-binding HTH domain of CcpA and CodY could reach an increase of 2-fold in overproduction of β-galactosidase. This was further demonstrated by the successfully increased overexpression of GFP, xylanase, and peptidase in the double mutant strain $CodY^{R214C}CcpA^{T19S}$. Although the level of gTME-based strain improvement differs per protein used, this strategy illustrates the broader application scope in overproducing a variety of proteins. Moreover, the subsequent analyses of the perturbed transcriptome will provide novel insights into the cellular components and their underlying interactions. A further characterization of the mutant strains will be described in the next Chapter.

MATERIALS AND METHODS

Plasmids, bacterial strains, and media reagents

B. subtilis 168 (*trpC2*) is the unique mother strain for all the mutants in this study. *E. coli* MC1061 was used as intermediate cloning host for all the plasmid construction. Both *B. subtilis* and *E. coli* were grown aerobically at 37 °C in Lysogeny Broth (LB) media unless otherwise indicated. When necessary, antibiotics were added in growth media as described previously [36]. The plasmids and strains included in this work are listed in **Table 2**.

DNA manipulation

Procedures for PCR, DNA purification, restriction, ligation and genetic transformation of *E. coli* and *B. subtilis* were carried out as described before [37, 38]. Pfu x 7 DNA polymerase [39] was a kind gift from Bert Poolman, and the USER enzyme was purchased from New England Biolabs. All FastDigest restriction

Table 2. The strains and plasmids used in this study.

Strain or plasmid	Genotype or properties	Reference or source
B. subtilis		
168	*trpC2*	[31]
168_β-gal	*trpC2, mdr::(*P*hyspank-lacZ spcr)*	This study
CodY^{-}	*trpC2, codY::cmr, mdr:(*P*hyspank-lacZ spcr*)	This study
CcpA^{-}	*trpC2, ccpA::ermr, mdr::(*P*hyspank-lacZ spcr)*	This study
CcpA^{-}CodY^{-}	*trpC2, ccpA::ermr, codY::cmr, mdr::(*P*hyspank-lacZ spcr)*	This study
CodY^{R214C}	*trpC2, codY R214C cmr, mdr::(*P*hyspank-lacZ spcr)*	This study
CodY$^{(E58A,I162V,R209H)}$	*trpC2, codY (E58A,I162V,R209H) cmr, mdr::(*P*hyspank-lacZ spcr*)	This study
CodY^{R214C}CcpA^{-}	*trpC2, ccpA::ermr, codY R214C cmr, mdr::(*P*hyspank-lacZ spcr)*	This study
CcpA^{T19S}	*trpC2, ccpAT19S kmr, mdr::(*P*hyspank-lacZ spcr)*	This study
CodY^{R214C}CcpA^{T19S}	*trpC2, ccpAT19S kmr, codYR214C cmr, mdr::(*P*hyspank-lacZ spcr*)	This study
comG-gfp	*trpC2, comG-gfp km*r	[25]
CodY^{-}*_comG-gf*p	*trpC2, comG-gfp kmr, codY::cmr*	This study
M4*_comG-gfp*	*trpC2, comG-gfp kmr, codYA203T cmr*	This study
M13*_comG-gfp*	*trpC2, comG-gfp kmr, codYK248I cmr*	This study
M17*_comG-gfp*	*trpC2, comG-gfp kmr, codYR214C cmr*	This study
M19*_comG-gfp*	*trpC2, comG-gfp kmr, codY(E58A,I162V,R209H) cmr*	This study
168*_sfgfp*	*trpC2, amyE::*P*hyspank- sfgfp spcr*	[32]

Strain or plasmid	Genotype or properties	Reference or source
CodY^{R214C}CcpA^{T19S}_*sfgfp*	*trpC2, codY R214C cmr, ccpAT19S kmr, amyE*::P*hyspank- sfgfp spcr*	This study
168_*pepP*	*trpC2, amyE*::P*hyspank- pepP spcr*	This study
CodY^{R214C}CcpA^{T19S}_ *pepP*	*trpC2, codY R214C cmr, ccpAT19S kmr, amyE*::P*hyspank- pepP spcr*	This study
NZ8900-spc	*trpC2, amyE::spaRK spcr*	This study
NZ8900-spc_ CodY^{R214C}CcpA^{T19S}	*trpC2, codY R214C cmr, ccpAT19S kmr, amyE::spaRK spcr*	This study
E.coli		
MC1061	F^{-}, *araD*139, Δ(*ara-leu*)7696, Δ(*lac*)X74, *galU*, *galK, hsdR2, mcrA, mcrB1, rspL*	[33]
Plasmids		
pJV112	*bla, mdr*::P*hyspank-lacZ spcr*, *lacI*	This study
pUC18	*bla, lacZ, ampr*	[34]
pCH3	pUC18_*aroA_ccpA_kmr_ytxD*	This study
pCH3_CcpAKO	pUC18_*aroA _ermr_ytxD*	This study
pJV54	pUC18_*clpY_codY_cmr_flgB*	This study
pJV55	pUC18_*clpY_cmr_flgB*	This study
pJV152	pUC18_*clpY_codYA203T_cmr_flg*B	This study
pJV151	pUC18_*clpY_codYK248I_cmr_flg*B	This study
pJV153	pUC18_*clpY_codYR214C_cmr_flg*B	This study
pJV173	pUC18_*clpY_codY(E58A,I162V,R209H)_cmr_flg*B	This study
pNZ8902_XynA	pNZ8902 harboring *xynA*	[35]

Table 3. The primer list in this study.

Oligonucleotide	Used for	Sequence (5' -> 3')
J01	Construction of CodY KO plasmid	GCGCGAATTCTGCTGCAGGGCTCGGGTATG
J02		GCGCGGTACCACGATGTATCTCCGCTCGCAAC
J03		GCGCAGATCTTATTTCCTGGGTTGAAAGTC
J04		GCACGCTAGCAAATAATCCTCCTAAACATTCC
J05	Amplify *mdr* coding region	GACGGGTACCTCGGCTGTTGTTTTCACTCTG
J06		CGCTCTGCAGTCTCCTGCTGGGCATCTTGATGTC
J07	Construction of reporter strain with P*hyspank-lacZ*	GGCTGCTAGCCCGCTTGTTTGCCACAGCACAGA
J08		GCGACTCGAGGGCAATGAAGAAAACGGCGAATAC
J09		GCCGAAGCTTAGGAGGAGAGGAAATGGAGGTTACTGACGTAAGATTACGG
J10		TTGGGCTAGCTTGCGGCCGCCCTTATTATTTTTGACACCAGACCAACTGGTAAT
J11		CCAGCGGCCGCAGGGAGAGGAAACATGATTCAAAAAC
J12		CTGGGCTAGCGCAGCGATCCCGATGAACAATCC
J13	Construction of integration vector for CodY	GGTTGGAATTCTGCTTCAAAGCCTGTCGGAATTGG
J14		TTGGGAGCTCAGATGCATTTTATGTCATATTGTA
J15		GGGTTGAGCTCAAATTTAATATGAGGAATGTTTAGG
J16		GTCGGGATCCAAAGACTTTCAACCCAGGAAATA
J17	Amplify *codY* coding region	GCCAGTCACGTTACGTTATTAG
J18		CTTATGCGACTCCTGCATTAG

Oligonucleotide	Used for	Sequence (5' -> 3')
pUC18-F	Construction of integration vector for CcpA	AAGCTTAGAGCUTGGCACTGGCCGTCGTTTTAC
pUC18-R		ATCGTAATCAUGGTCATAGCTGTTTCCTGTGTGAAATTG
aroA-F		ATGATTACGAUGAGCAACACAGAGTTAGAG
aroA-R		ACCGAGCGTTCUGAGGTACCCCTAAAACCACTCCTTTTACTG
gfp-F		AGAACGCTCGGUTGCCGCCGGGCGTTTTTTATG
gfp-R		ACGCACCTTUGCATGCGCTTCACAGTTTCTTCTTC
km-F		AAAGGTGCGUTGAAGTGTTGGTATGTATG
km-R		AAACGTTTCAUTCTAGATTACGCGAAATACGGGCAGAC
ytxD-F		ATGAAACGTTUTGATTATCTTACACCTGTTGGATTTG
ytxD-R		AGCTCTAAGCTUCATGTACAGATCCCTTTTTTG
CcpA-F	Amplify *ccpA* coding region	ATATGGTACCTTGAACAATCCAAAAGGCCGCCGTGC
CcpA-R		ATATGCATGCCATTTCTCCCCATGAAAAAAG
erm-F	Construct *ccpA* KO plasmid	ATATGGTACCTCGATAACCCTAAAGTTATG
erm-R		ATATTCTAGAGATACAAATTCCCCGTAGGC
CcpA-T19S-F	Site-directed mutation of CcpA	CTAATGTAAGCATGGCATCCGTTTCCCGTGTCGTG
CcpA-T19S-R		CACGACACGGGAAACGGATGCCATGCTTACATTAG
CcpA-seq-F	*ccpA* sequencing	GCTCAGCAAATGGCGATTCC
CcpA-seq-R		ACCGCTTACTCCCGATCCTG
CodY-seq-F	*codY* sequencing	GTCGAAGAAAAGCTCGGAACGATAG
CodY-seq-R		AGACTTTCAACCCGAGAAATAAAGC

enzymes, Phusion and Dreamtaq DNA polymerases were acquired from Thermo Fisher Scientific (Landsmeer, Netherlands). The NucleoSpin Plasmid EasyPure and NucleoSpin Gel and PCR Clean-up kits were purchased from BIOKE (Leiden, Netherlands). All the reagents used were bought from Sigma unless otherwise indicated. All the oligonucleotide primers used in this study are listed in **Table 3**, which were synthesized by Biolegio (Nijmegen, Netherlands). Sequencing of all our constructs was performed at MacroGen (Amsterdam, Netherlands).

Black-white screening

All transformants of *B. subtilis* mutant libraries were scraped off the LB agar plates with appropriate antibiotic(s) and collected into one flask with fresh LB liquid media. After 37 °C overnight incubation, the culture was serially diluted and plated on selective agar media -Spizizen's minimal media (SMM) [40] supplemented with 1.0% glucose, 0.1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG), 300 mg/l of S-gal and 500 mg/l of ferric ammonium citrate for black-white screening. After 20 h of incubation at 37 °C, colonies were isolated from the plates based on the color intensity and morphology, followed by sequence analysis and enzymatic assays (**Fig. 6**).

Site-directed mutagenesis

Site-directed mutagenesis of *ccpA* was processed to get the clean mutation site T19S in amino acids sequence, and it was performed using overlap PCR, the previously constructed vector pCH3 was used as the template for allelic exchange [41]. All plasmids were introduced into *B. subtilis* reporter strain 168_β-gal and two CodY mutants (M17 and M19) by transformation after passaging through *E. coli* strain MC1061, sequences were verified as before. The primers were designed via the website: http://bioinformatics.org/primerx/.

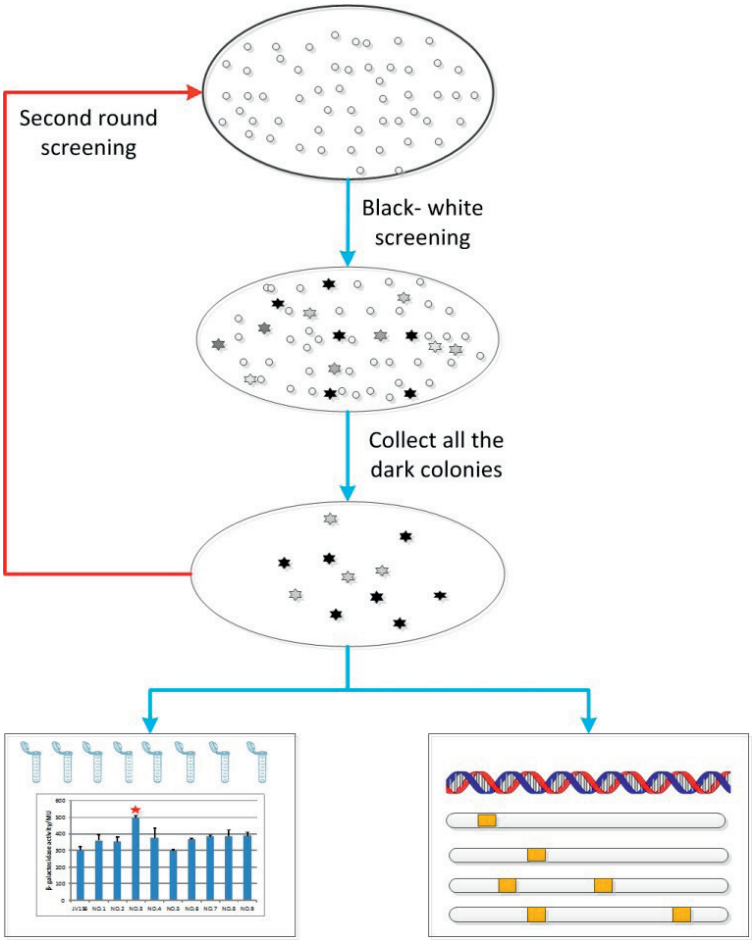


Fig. 6. Overall workflow of the high-throughput screening for β-galactosidase higher-producing phenotypes. The mutant libraries were screened via the black-white selection, and all the dark colonies were collected together for the second round of screening for narrowing down the libraries to a few of good phenotypes with enhanced β-galactosidase production. The genotypes of selected mutants were sequencing identified.

Enzymatic assays

For determination of β-galactosidase activity, strains were grown under identical conditions in LB media supplemented

with 1.0% glucose and 0.1 mM IPTG, shaking at 37 °C and 220 rpm until the mid-exponential growth phase was reached (OD600 of 1.0). The cultures (1 ml) were immediately harvested and frozen in liquid nitrogen. The pellet was processed for β-galactosidase quantitation as previously described [42]. The α-amylase activity was assessed as previously described, and the final amylase activity was calculated by the Ceralpha Unit [36]. Each assay was performed in duplicate, and the mean value from three independent experiments was calculated.

Microplate experiments

We used a GFP fusion under the control of the *comG* promoter (P*comG-gfp*) to monitor the competence development in various strains. First, the cells were grown overnight in the competence media MC1X at 37 °C, 220 rpm, then we diluted the pre-cultures 50 times in fresh MC1X and started 22 hours' incubation at 37 °C and 220 rpm shaking in the plate reader-VarioskanLUX (Thermo Fisher) with a GFP filter set (excitation at 485/20 nm, emission 535/25), and the absorbance was measured at 600 nm. The values of GFP intensity and OD600 were automatically recorded every 15 minutes for 22 hours, data of all samples were collected in triplicates. All the optical density and fluorescence values were corrected for the background of the media by the following formula: (GFPreporter-GFPmedia)/(ODreporter-ODmedia) [43].

Western blotting

We grew the strains overnight in LB media supplemented with 1.0% glucose, the cultures were diluted in fresh media and incubated for overnight. Next morning, 5 ml cells were collected by centrifugation (14,000 rpm, 5 min), the pellets were resuspended in 1 ml solution A [50 mM Tris·Cl (pH 7.5), 5% (vol/vol) glycerol, and 1 mM PMSF] and broken by sonication for 2 min using 70% amplitude with 10-s bursts, with 10-s pauses between

bursts. After centrifugation (14,000 rpm, 5 min) the supernatant was transferred to a clean tube. 15 µl of total protein was heated for 5 min at 95 °C before loading on a 12% SDS-PAGE gel. After the proteins were separated by electrophoresis, the proteins were transferred to a PVDF membrane (80 mA, 90 min). The PVDF membrane was incubated in PBST + 5% (w/v) BSA at 4 °C overnight. Next morning, the membrane was washed three times 10 min with PBST and subsequently incubated at room temperature for 2 hours in PBST + 5% BSA + 1: 10,000 dilution of anti-Histag antibody. The membrane was again washed three times 10 min with PBST and then incubated 1.5 hours in PBST + donkey anti-rabbit IgG-horseradish peroxidase at room temperature. Then the membrane was washed two times 10 min with MQ and two times 10 min with PBST. 1 ml of ECL detection Reagent (GEHealth care) and the Molecular Imager ChemiDoc XRS+ (BioRad) were used for signal visualization.

Flow cytometry

All the strains were streaked on LB agar plates with spec antibiotic, and the single colonies were picked up and grown overnight in LB supplemented with 1.0% glucose and 0.1 mM IPTG at 37 °C, 220 rpm. Next morning, cultures were diluted 20 times in phosphate buffered saline (PBS) and directly measured on the Becton Dickinson FACSCanto (BD BioSciences, USA) with an argon laser (488 nm) [44]. For each sample, the green fluorescent signals of 50,000 cells were collected by a FITC filter. All the captured data was further analyzed using Flowing Software (http://www.flowingsoftware.com/).

ACKNOWLEDGMENTS

We thank Tjeerd van Rij (DSM) and Marc Kolkman (Genencor) for helpful discussions. This research was partially funded by a grant from the former Kluyver Center for Genomics of Industrial Fermentation (Delft/Groningen) to JVH and by an ALW grant to RDOW. HC was supported by a grant from China Scholarship Council (CSC). We thank Dr. Lance Keller (Department of Fundamental Microbiology, University of Lausanne) for proofreading the manuscript.

REFERENCES

1. Westers L, Westers H, Quax WJ. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochim Biophys Acta.* 2004;1694:299–310.
2. Zweers JC*, et al.* Towards the development of *Bacillus subtilis* as a cell factory for membrane proteins and protein complexes. *Microb Cell Fact.* 2008;7:p.1.
3. van Dijl J, Hecker M. *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. *Microb Cell Fact.* 2013;12:p.1.
4. Song Y, Nikoloff JM, Zhang D. Improving protein production on the level of regulation of both expression and secretion pathways in *Bacillus subtilis*. *J Microbiol Biotechnol.* 2015;25:963–977.
5. Kang Z, *et al*. Molecular engineering of secretory machinery components for high-level secretion of proteins in *Bacillus* species. *J Ind Microbial Biot.* 2014;41:1599–1607.
6. Chen J*, et al.* Enhanced extracellular production of alpha-amylase in *Bacillus subtilis* by optimization of regulatory elements and over-expression of PrsA lipoprotein. *Biotechnol Lett.* 2015;37:899–906.
7. Liu Y, *et al.* Metabolic engineering of *Bacillus subtilis* fueled by systems biology: Recent advances and future directions. *Biotechnol Adv.* 2017;35:20–30.

8. Wu G*, et al.* Metabolic Burden: cornerstones in synthetic biology and metabolic engineering applications. *Trends Biotechnol.* 2016;34:652–664.
9. Zou W, Edros R, Al-Rubeai M. The relationship of metabolic burden to productivity levels in CHO cell lines. *Biotechnol Appl Bioc*. 2017.
10. Glick BR. Metabolic load and heterologous gene expression. *Biotechnol Adv.* 1995;13:247–261.
11. Pitera DJ, *et al.* Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metab Eng.* 2007;9:193–207.
12. Colletti PF*, et al.* Evaluating factors that influence microbial synthesis yields by linear regression with numerical and ordinal variables. *Biotechnol Bioeng.* 2011;108:893–901.
13. Levdikov VM*, et al.* Structure of the branched-chain amino acid and GTP-sensing global regulator, CodY, from *Bacillus subtilis*. *J Biol Chem.* 2017;292:2714–2728.
14. Mijakovic I*, et al.* Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: A relic of early life? *Proc Natl Acad Sci U S A.* 2002;99:13442–13447.
15. Schumacher MA*, et al.* Structural basis for allosteric control of the transcription regulator CcpA by the phosphoprotein HPr-Ser46-P. *Cell.* 2004;118:731–741.
16. Sonenshein AL. Control of key metabolic intersections in *Bacillus subtilis*. *Nat Rev Microbiol.* 2007;5: 917–927.
17. Belitsky BR, Sonenshein AL. Genome-wide identification of *Bacillus subtilis* CodY-binding sites at single-nucleotide resolution. *Proc Natl Acad Sci U S A.* 2013;110:7026–7031.
18. Fujita Y. Carbon catabolite control of the metabolic network in *Bacillus subtilis*. *Biosci Biotechnol Biochem.* 2009;73:245–259.
19. Alper H, Stephanopoulos G. Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab Eng.* 2007;9:258–267.
20. Tyo KE, Alper HS, Stephanopoulos GN. Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends Biotechnol.* 2007;25:132–137.

21. Klein-Marcuschamer D, Stephanopoulos G. Assessing the potential of mutational strategies to elicit new phenotypes in industrial strains. *Proc Natl Acad Sci U S A.* 2008;105:2319–2324.
22. Park KS*, et al.* Phenotypic alteration of eukaryotic cells using randomized libraries of artificial transcription factors. *Nat Biotechnol.* 2003;21:1208–1214.
23. Alper H, *et al.* Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science.* 2006;314:1565–1568.
24. Bitinaite J*, et al.* USER friendly DNA engineering and cloning method by uracil excision. *Nucleic Acids Res.* 2007;35:1992–2002.
25. Smits WK*, et al.* Stripping Bacillus: ComK auto-stimulation is responsible for the bistable response in competence development. *Mol Microbiol.* 2005;56:604–614.
26. Hamoen LW, Venema G, Kuipers OP. Controlling competence in *Bacillus subtilis*: shared use of regulators. *Microbiology.* 2003;149:9–17.
27. Serror P, Sonenshein AL. CodY is required for nutritional repression of *Bacillus subtilis* genetic competence. *J Bacteriol.* 1996;178:5910–5915.
28. Chen PT, Chiang CJ, Chao YP. Medium optimization for the production of recombinant nattokinase by *Bacillus subtilis* using response surface methodology. *Biotechnol Progr.* 2007;23:1327–1332.
29. Lanza AM, Alper HS. Using transcription machinery engineering to elicit complex cellular phenotypes. In *Synthetic Gene Networks*. Humana Press. 2012;229–248.
30. Woolston BM, Edgar S, Stephanopoulos G. Metabolic engineering: past and future. *Annu Rev Chem Biomol.* 2013;4:259–288.
31. Kunst F*, et al.* The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature.* 1997;390.
32. Overkamp W*, et al.* Benchmarking various green fluorescent protein variants in *Bacillus subtilis*, *Streptococcus pneumoniae*, and *Lactococcus lactis* for live cell imaging. *Appl Environ Microbiol.* 2013;79:6481–6490.
33. Casadaban MJ, Cohen SN. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J Mol Biol.* 1980;138:179–207.
34. Vieira J, Messing J. New pUC-derived cloning vectors with different selectable markers and DNA replication origins. *Gene.* 1991;100:189–194.

35. Trip H*, et al.* A novel screening system for secretion of heterologous proteins in *Bacillus subtilis*. *Microb Biotechnol.* 2011;4:673–682.
36. Cao H, *et al.* Cell surface engineering of *Bacillus subtilis* improves production yields of heterologously expressed alpha-amylases. *Microb Cell Fact.* 2017;16:56.
37. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.
38. Konkol MA, Blair KM, Kearns DB. Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J Bacteriol.* 2013;195:4085–4093.
39. Norholm MH. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnol.* 2010;10:p.1.
40. Anagnostopoulos C, Spizizen J. Requirements for transformation in *Bacillus subtilis*. *J Bacteriol.* 1961; 81:741–746.
41. Zheng L, Baumann U, Reymond JL. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* 2004;32:e115.
42. Smale ST. β-Galactosidase assay. *Cold Spring Harb Protoc.* 2010;5:pdb.prot5423.
43. van den Esker MH, Kovacs AT, Kuipers OP. YsbA and LytST are essential for pyruvate utilization in *Bacillus subtilis*. *Environ Microbiol.* 2017;19:83–94.
44. Veening JW, Hamoen LW, Kuipers OP. Phosphatases modulate the bistable sporulation gene expression pattern in *Bacillus subtilis*. *Mol Microbiol.* 2005;56:1481–1494.

CHAPTER 4

Stronger repression of carbon metabolic pathways and de-repression of nitrogen metabolic benefit heterologous protein synthesis in *Bacillus subtilis*

Haojie Cao[1], Julio Villatoro-Hernandez[1], Elrike Frenzel[1], Oscar P. Kuipers[1]

[1]Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands.

Based on the manuscript that is under review in *Metab Eng*: Haojie Cao, Julio Villatoro-Hernandez, Ruud Detert Oude Weme, Elrike Frenzel, Oscar P. Kuipers. Boosting heterologous protein production yield by adjusting global nitrogen and carbon metabolic regulatory networks in *Bacillus subtilis*.

ABSTRACT

To increase the yield of heterologous protein in *Bacillus subtilis*, we applied global transcription machinery engineering (gTME) and high-throughput screening, which proved to be able to effectively boost the production of reporter proteins. The transcription factor variants with amino acid substitutions in the DNA-binding HTH domain of CcpA and CodY resulted in a significant increase of β-galactosidase production. Transcriptome and gel mobility shift analyses revealed that these two specific mutations not only altered the overall binding specificity to the respective regulon genes operator sites but also affected the expression of these two regulatory proteins. This accordingly leads to further repression of the carbon core metabolism and dramatic de-repression of nitrogen metabolism. Consequently, these two central metabolic networks, which are intertwined by feedback-regulation of branched-chain amino acids (BCAAs), are rewired and better balanced. Thus, the two reprogrammed metabolic pathways function coordinately to enhance the expression of four specific CodY-regulated operons (*rocABC*, *rocDEF*, *hutPHUIGM* and *appDFABC*), which showed a positive correlation with the production yields of β-galactosidase in *B. subtilis*.

Keywords: CcpA, CodY, metabolic intersection, branched-chain amino acids (BCAAs), transcriptomics, electrophoretic mobility shift assay (EMSA)

INTRODUCTION

Bacillus subtilis, the best-characterized member of the Gram-positive bacteria, has been intensively investigated since the early 1950s [1]. In the last decades, the fast accumulation of genetic information greatly enhanced our ability to further understand the underlying regulatory mechanisms of central metabolic pathways in *B. subtilis*. The global transcriptional regulator CodY either represses or, less frequently induces the transcription of target genes in the late exponential or early stationary phase in the presence of high intracellular levels of GTP and branched-chain amino acids (BCAAs; isoleucine, valine, and leucine) [2, 3]. BCAAs act as a corepressor by sterically triggering conformational changes that lead to altered DNA binding capabilities [4]. This transcriptional regulation enables cells to adapt to various nutrient conditions in different growth environments, inducing a wide variety of cellular processes, e.g. sporulation, competence, nitrogen metabolism and biofilm formation [5, 6]. A second global transcriptional regulator that orchestrates central metabolism, specifically carbon utilization, is the extensively studied catabolite control protein A (CcpA). This transcriptional factor is activated when in complex with phosphorylated histidine-containing protein (HPr) or its paralogous protein Crh [7, 8]. Its activity can be enhanced by fructose-1,6-bisphosphate (FBP) and glucose-6-phosphate (G6P) when the cells are grown with glucose or other preferentially utilized carbon sources [9]. Subsequently, the active CcpA binds to the catabolite repression elements (*cre* sites) of the target regulon, leading to carbon catabolite repression (CCR) or carbon catabolite activation (CCA) [10, 11].

In brief, CodY and CcpA sense diverse intracellular metabolites (GTP, BCAAs, FBP, and G6P) to be active, and then modulate the expression of hundreds of genes (**Fig. 1**), directly or indirectly (5). Thus, these two global regulatory proteins function

together to orchestrate large regulons that balance the uptake and utilization of nutrient sources, and systemically coordinate the intracellular metabolic flux distributions by regulating specific cellular processes [12–14]. In the previous study (Chapter 3), we reprogrammed the central metabolic networks via random mutagenesis of CodY and CcpA, and determined the influence on the expression level of the reporter protein (β-galactosidase) that was apparent in the color intensity of various selected colonies. High-throughput screening was performed to successfully gain interesting phenotypes with a significantly higher production level of β-galactosidase relative to wildtype (WT) strain. The different regulatory effects on the carbon and nitrogen metabolic pathways were investigated by transcriptomics and other analytic approaches, and the analyses of the globally rewired metabolic networks provide new insights into the complex interactions between CodY and CcpA.

RESULTS

Enhanced transcription of specific nitrogen metabolic operons is positively correlated with β-galactosidase production

The target genes that are under the regulation of CodY and CcpA have been identified by genome-wide analyses of the transcriptome and protein-DNA interactions in *B. subtilis* in various studies [3, 6, 15–20]. We opted to investigate the global cellular responses to the amino acid substitutions in the DNA-binding HTH motifs of the transcriptional regulators, i.e. $CodY^{R214C}$ and $CcpA^{T19S}$, respectively (**Chapter 3**), during the mid-exponential growth phase to obtain a better understanding of the regulatory effects. The transcriptome patterns associated with the selected strains implied that the expression of individual genes is differentially affected,

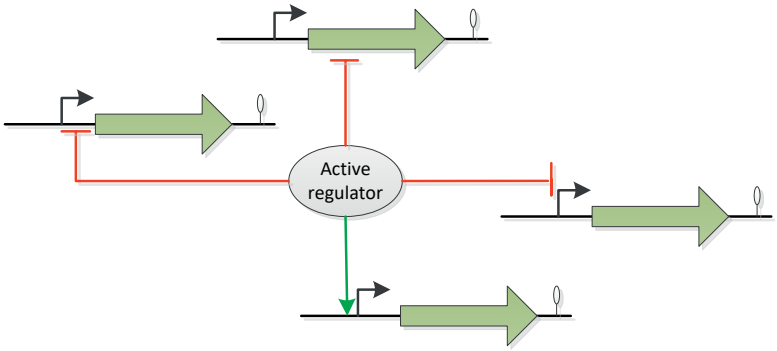
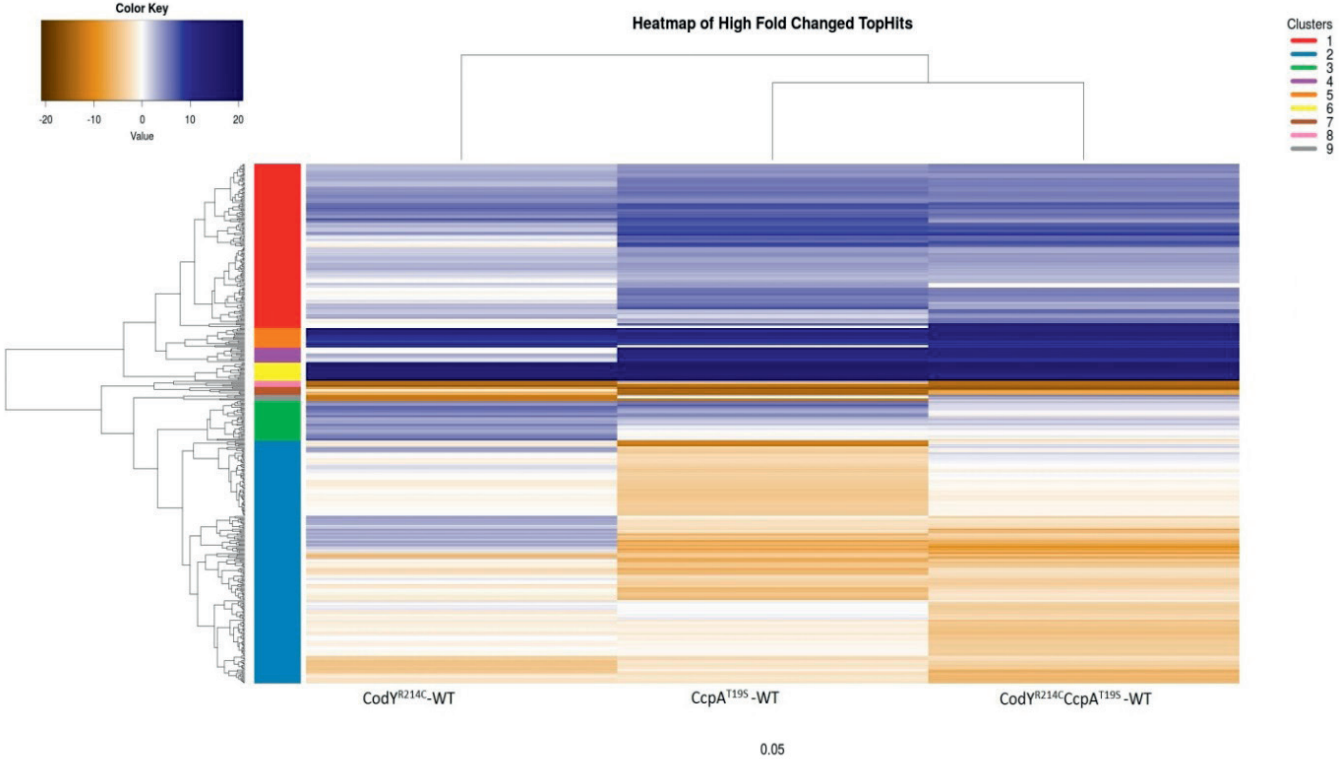


Fig. 1. Schematic diagram of regulatory mechanism. The regulators specifically bind to different locations of various regulons to trigger diverse regulation effects. Green arrow--activation; red perpendicular--repression.

which would display a broad range of sensitivities to the HTH sequence mutations (**Fig. 2**). This would be expected because of the scale of global transcriptional regulation and the complexity of the interplay between diverse metabolic networks [21] and also because CodY binds to different target sites with a varying affinity. Next to genes that are known regulon members and under direct transcriptional control of CodY or CcpA, additional genes involved in nitrogen and carbon core metabolic pathways were transcribed differentially in the mutant strains. These were clustered according to their functional category in the *Subti*Wiki database (**Fig. S1**).

In comparison to the WT strain, the vast majority of CodY regulon members and nitrogen metabolism associated genes were either up-regulated or unchanged in the strains with mutations in DNA-binding regions of CodY and/or CcpA (**Fig. S1A-D**). In contrast, the overall fluctuation in expression levels of genes from the central carbon metabolism was modest; most of the genes were slightly down-regulated, and only a few were expressed 2-fold higher than in the WT (**Fig. S1E**).

Interestingly, a specific set of gene clusters was positively correlated with the β-galactosidase production performance in the


Color Key
-20
-10
0
10
20
Value
Heatmap of High Fold Changed TopHits
Clusters
1
2
3
4
5
6
7
8
9
CodY^{R214C}-WT
CcpA^{T19S}-WT
CodY^{R214C}CcpA^{T19S}-WT
0.05

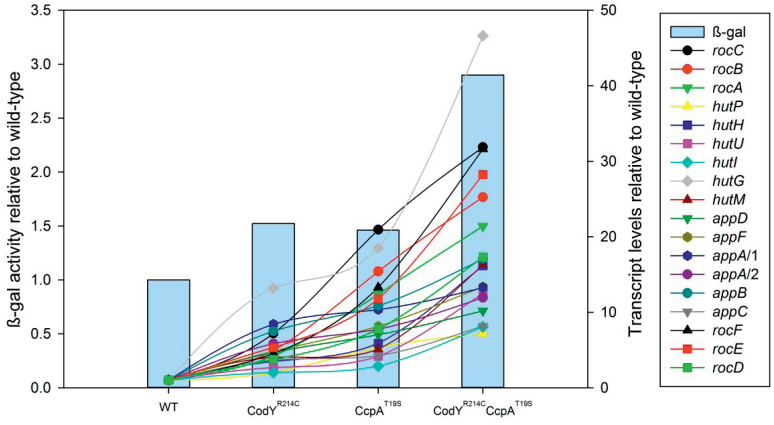


Fig. 3. Transcription levels of the operons *rocABC*, *hutPHUIGM*, *appDFABC* and *rocDEF* are correlated with increased β-galactosidase production. The enzymatic activity of β-galactosidase is shown relative to the enzyme activity in the WT (blue columns). Because of the high differences in transcript levels, the normalized RPKM (Reads Per Kilobase per Million mapped reads) values of indicated genes are shown relative to the WT levels, which were arbitrarily scaled to 1.

HTH domain mutation strains (**Fig. 3**). All of these operons are negatively controlled by CodY, and their corresponding products are involved in the uptake and utilization of specific nitrogen sources. The operons *rocABC* and *rocDEF* encode enzymes that participate in the uptake and utilization of arginine, ornithine, and citrulline [22, 23]. The *hutPHUIGM* operon is involved in histidine metabolism, which is additionally also negatively

Fig. 2. Heatmap of high fold changed genes in various regulator mutants relative to WT. We excluded the genes that did not show statistically significant changes in transcript levels ($p>0.05$) and fold change less than two ($FC<2$) in all contrasts. The differentially expressed genes ($p<0.05$) between mutants and WT libraries were normalized, centered and automatically clustered by web server T-Rex. Brown -- lower expression; blue -- higher expression.

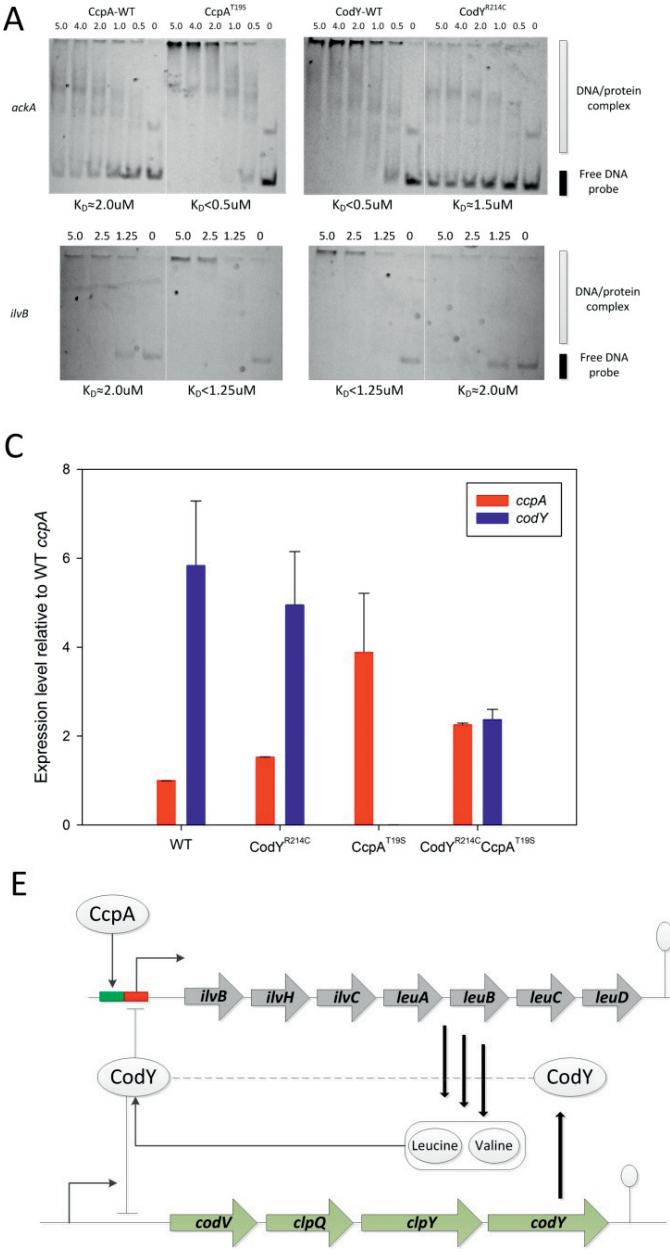
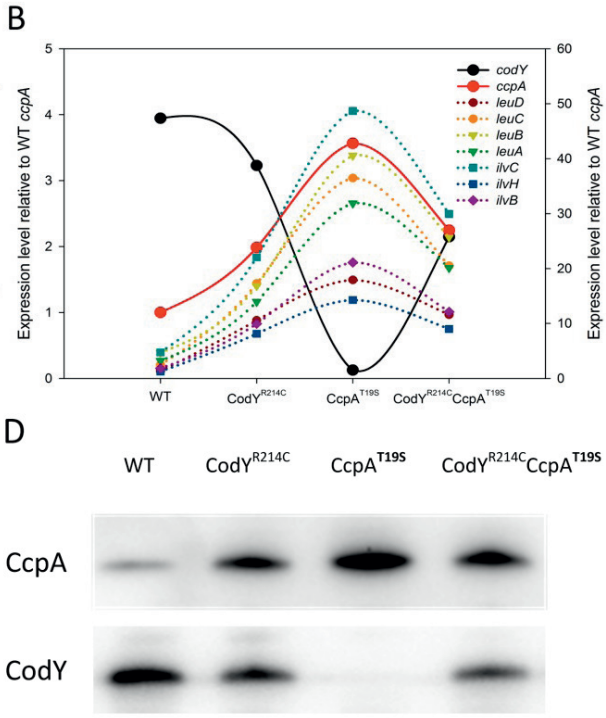


Fig. 4. (A) Gel mobility shift analysis of CodY and CcpA binding to the regulatory regions of *ackA* and *ilvB*. The 3' ends of DNA fragments were labeled with Cy3, and the obtained DNA probes (0.1 μm) were incubated with various



concentrations (μM) of His6-tagged transcriptional factors. KD reflects the protein concentration needed to shift 50% of DNA fragments [26]. **(B) The relative expression levels of CodY, CcpA and the *ilv-leu* operon in the HTH domain mutant strains.** The RPKM values of each strain were normalized by that of *sigA* (internal reference gene). The normalized values of *ccpA*, *codY* and the *ilv-leu* operon genes were related to WT *ccpA*, and the value of WT *ccpA* was finally defined to 1. **(C) qRT-PCR to analyze the expression level of the two proteins.** The transcript level of *sigA* was used as the internal control, and the values obtained were related to that of WT *ccpA*, which was normalized to 1.0. Each column represents the mean ± SD of three independent experiments, and each assay was performed in duplicate. **(D) Immunoblots of lysates from cells grown to an OD600 of 1.0.** Strains were grown in LB media in the presence of 1.0% glucose and 0.1 mM IPTG; the proteins were separately immunoblotted with polyclonal CcpA and CodY antibodies. **(E) Schematic diagram of the interaction between CcpA and CodY mediated by the biosynthesis of BCAAs.** Arrows and perpendiculars represent the positive and negative actions, respectively.

regulated by CcpA [24]. The *appDFABC*operon is directly related to nitrogen source utilization by encoding an ABC transporter for the uptake of peptides [25]. As illustrated in **Fig. 3,** $CodY^{R214C}$ and $CcpA^{T19S}$ separately promoted the changed expression of specific pathways for nutrient uptake and utilization, and these benefits were clearly synergistic in the double mutant host. Thus, we demonstrate that mutations in the conserved DNA binding motifs of CodY and CcpA significantly enhanced the expression of specific operons, and the resulting up-regulated nitrogen source metabolism was identified as a beneficial factor allowing the mutant strains to synthesize more reporter protein during growth in LB supplemented with 1.0% glucose.

Stronger repression of carbon metabolic pathways and de-repression of nitrogen metabolic pathways benefit the synthesis of β-galactosidase

Since the HTH motifs of CodY and CcpA are highly conserved among many low G+C Gram-positive species [27, 28], we next addressed the question whether the T19S and R214C mutations affect the DNA-binding ability of the regulator proteins. Electrophoresis mobility shift assay (EMSA) with purified CcpA and CodY WT and mutant proteins and *ackA* and *ilvB* promoter fragments revealed that all protein variants were capable of binding to the selected DNA probes (**Fig. 4A**), which was in line with previous findings [12, 14]. The $CodY^{R214C}$ and $CcpA^{T19S}$ mutants, however, bound to several regulatory sites with reduced and increased efficiency respectively, in comparison to their WT proteins. Furthermore, the mutation T19S significantly enhanced the binding efficiency of CcpA to the promoter regions of *rbsR* and *treP*, which are under the direct negative control of CcpA (**Fig. S2**).

The transcriptome analysis further revealed that the transcription factors CodY and CcpA themselves were differentially expressed in the reprogrammed HTH domain mutants. Surprisingly,

these two regulatory proteins displayed exactly opposite expression patterns in the strains (**Fig. 4A**). The transcript abundance of CodY was around three times higher than that of CcpA in the WT strain JV156, and this difference in transcript levels was halved in the CodY^{R241C} mutant. Particularly, the CcpA mutation T19S led to an increase in transcription of *ccpA* and substantially decreased the transcription of *codY*, making the CcpA^{T19S} mutant strain nearly behave in the direction of a CodY deficient strain. However, the differential accumulation of the two transcriptional regulators got balanced when these two mutations were combined in one cell (**Fig. 4B**). This is supposedly a consequence of the interplay between these two regulators with altered regulation efficiencies and expression capacities. Importantly, this observation could be confirmed at the transcriptional and translational level by qRT-PCR and Western Blotting (**Fig. 4C** and **Fig. 4D)**.

We thus show that the bacteria tend to alter the overall metabolic network fluxes through the expression variation of these two regulators to meet the demand of resources for the overproduction of heterologous proteins. In other words, the metabolic shifts occurring in *B. subtilis* can be regarded as a fitness adaptation of this microbial cell factory to the global transcriptome perturbations and the requirement of high-yield protein production [30]. Finally, the increased CcpA and decreased CodY protein levels lead to an enhancement of the repression of the carbon metabolism and amplification of the reactions in the nitrogen metabolism networks, and thus the reprogrammed metabolic networks are obviously beneficial for the biosynthesis of β-galactosidase.

CcpA^{T19S} promotes β-galactosidase synthesis by impacting co-factor availability and the auto-regulatory expression loop of CodY

In *B. subtilis*, the two BCAAs leucine and valine, which are biosynthesized through the catalysis of *ilv-leu* encoded enzymes,

effectively increase the affinity of CodY to its DNA binding sites. CodY and CcpA bind to different regulatory regions of the *ilv-leu* gene cluster promoter region and CcpA thus indirectly controls the expression of other CodY-regulated genes by the modulation of the intracellular level of BCAAs [12, 13]. The transcription profile of the HTH mutants reflects that CcpA and CodY regulated metabolic pathways interact at the node of the *ilv-leu* operon (**Fig. 4B**). The two opposite transcriptional regulatory effects cause this operon to be expressed in a similar pattern as the *ccpA* gene but exactly opposite to *codY* (**Fig. 4B**). More specifically, the higher levels of accumulated BCAAs have negative feedback on the expression of CodY itself (**Fig. 4E**). The increased β-galactosidase production in the $CodY^{R214C}$ and/or $CcpA^{T19S}$ strains was strongly correlated with the transcript abundance of several CodY-regulated operons, which is obviously caused by an indirect derepression mediated by $CcpA^{T19S}$ (through indirectly suppressing the transcription and activity of $CodY^{R214C}$ (**Fig. S3**)). Hence, these two global regulatory proteins with altered binding specificities and expression activities cooperate to achieve the synergistic effect of target protein expression in the double mutant strain $CodY^{R214C}CcpA^{T19S}$ by prioritizing the upregulation of specific CodY targets via the modulation of BCAAs biosynthesis.

The gene regulatory network of nitrogen metabolism is more affected in high-capacity production mutants than the network of carbon metabolism

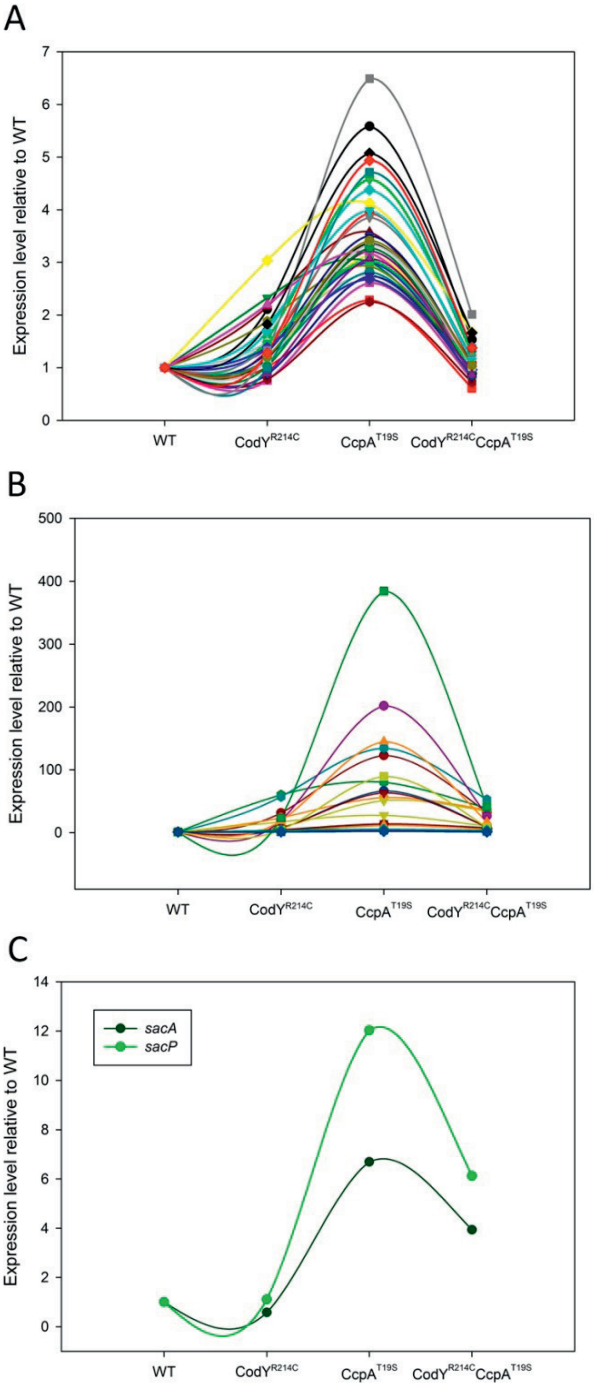
A higher number of CodY-regulated genes than CcpA regulon member genes were altered with respect to their expression levels in various mutants relative to the WT strain (**Fig. S1A** and **Fig. S1B).** Next to the *ilv-leu* operon, 52 CodY-repressed genes including the BCAA biosynthesis related genes *ilvA, ilvD, and ybgE* were differentially expressed [12, 31], while only two

CcpA targets (*sacA* and *sacP*) showed a consistent co-expression with CcpA (**Fig. 5**). We therefore conclude that CodY-regulated genes react more sensitively to the changes of the corresponding regulator levels, while the expression levels of CcpA regulon genes remained rather stable in all HTH domain mutant strains. However, by specifically analyzing genes related to amino acid utilization or biosynthesis/acquisition and carbon core metabolism according to the gene ontology (GO) classification in the *Subti*Wiki database, in **Fig. S1C, Fig. S1D** and **Fig. S1E**, significantly more amino acid metabolism genes were differentially expressed than carbon core metabolic genes. In short, the response of amino acid metabolic pathways to the changed transcriptome is more significant than that of central carbon metabolism in overproduction strains.

DISCUSSION

Decades of research have demonstrated the importance of improving specific target modification to increase the cell factory protein productivity, but engineering the global transcription machinery of central metabolic networks has so far been neglected. In **Chapter 3**, the tool of gTME in combination with high-throughput screening was applied to gain good phenotypes with improved product yields. Which was proved to outperform the traditional approaches for increasing the productivity of microbial cell factories. Most importantly, the system-wide analyses we performed here involving transcriptomics and protein-DNA binding assays provide a better understanding of the complex interactions between central metabolic pathways in *B. subtilis*.

In contrast to the nitrogen metabolic network, the central carbon metabolism was less responsive to the transcriptional


A
Expression level relative to WT
0
1
2
3
4
5
6
7
WT
CodY^{R214C}
CcpA^{T19S}
CodY^{R214C}CcpA^{T19S}
B
Expression level relative to WT
0
100
200
300
400
500
WT
CodY^{R214C}
CcpA^{T19S}
CodY^{R214C}CcpA^{T19S}
C
Expression level relative to WT
0
2
4
6
8
10
12
14
sacA
sacP
WT
CodY^{R214C}
CcpA^{T19S}
CodY^{R214C}CcpA^{T19S}

regulatory perturbations in all the mutants. This reflects that the threshold activities of regulators required for each gene are unequal, and the individual targets for both regulators are subject to the differential, gradual stimulation, and repression [3]. Our data suggest that the higher activity of CcpA^{T19S} exceeded the maximum activation threshold, which was obviously sufficient to keep a stringent inhibition of gene expression of a vast majority of the CcpA regulon. In contrast, the activity of CodY^{R214C} was decreased and obviously lower than the level that full repression of gene transcription demands. Consequently, a set of CodY targets was dramatically upregulated due to the decline or even elimination of the transcriptional and metabolic repression. Although these two transcription factors regulate more than 200 genes directly [32], only one out of ten genes showed an altered expression profile in response to the global transcriptome perturbations. This is likely based on the fact that the vast majority of genes and operons are subject to complex, multiple forms of regulation at different expression levels.

In the natural environment, the availability of nutrients can be highly variable, and the bacteria have evolved sophisticated adaptation systems for making good use of a wide range of sources of essential elements [5]. Therefore, the carbon core

Fig. 5. Grouping of genes with similar transcription patterns in the metabolic reprogramming mutants whose expression correlates with the expression level changes of CcpA and CodY, respectively. The absolute values of transcript abundances from each strain were normalized to that of the constitutively expressed *sigA* gene. The normalized RPKM values of indicated genes are shown relative to the WT levels, which were arbitrarily scaled to 1. **(A)** The *fla-che* operon. **(B)** Other genes which are negatively controlled by CodY. These genes include *ybgE, amhX, ylmA,ilvA,ilvD, yufN, yufO, yufP, yufQ, yuiC, yuiB, yuiA,yurJ, frlD, frlM, frlN, frlO, frlB, ywaA, rocR.* **(C)** The gene *sacA* and *sacP*, which are negatively controlled by CcpA.

metabolism, which can guarantee essential energy and building blocks supply [33], has been well evolved to serve the bacteria in various conditions. The central pathways can be protected against the stochastic fluctuations by the overabundance of relevant enzymes [34]. The generation of buffer space ensures that the transcriptome perturbations will not severely restrict the capabilities of cellular energy metabolism. During the long-lasting natural evolutionary process, surviving under unfavorable or extreme growth environments is the primary task for microorganisms in contrast to the human demand for overproduction of heterologous protein, explaining why a global adjustment of global N- and C- metabolism is effective to support the latter.

In brief, we could significantly improve the productivity of *B. subtilis* by the rewiring of central metabolic regulation, which promotes a good balance of resource distributions between normal cellular processes and needs for heterologous protein production. Undoubtedly, further improvements in our ability to reveal the underlying interactions between transcriptional regulation and dynamic metabolic status will come from future studies [34]. This investigation provides a new approach to improve *B. subtilis* as a cell factory, which is of broad significance for both industrial application and fundamental studies.

MATERIALS AND METHODS

Plasmids, bacterial strains, and growth condition

The *B. subtilis* 168 (*trpC2*) is the unique mother strain for all the derived *B. subtilis* in this study. The *Escherichia coli* MC1061 was used as intermediate cloning host for all the plasmid construction. Both *B. subtilis* and *E. coli* were grown aerobically at 37 °C in Lysogeny Broth (LB) media unless otherwise indicated. When necessary, the antibiotics were added in growth media as

described previously [35]. A detailed list of plasmids and strains included in this work is found in **Table 1**.

Recombinant DNA techniques

Procedures for PCR, DNA purification, restriction, ligation and genetic transformation of *E. coli* and *B. subtilis* were carried out as described before [36, 37]. Pfu x 7 DNA polymerase (38) was a kind gift from Bert Poolman, and the USER enzyme was purchased from New England Biolabs. All FastDigest restriction enzymes, Phusion and Dreamtaq DNA polymerases were acquired from Thermo Fisher Scientific (Landsmeer, Netherlands). The NucleoSpin Plasmid EasyPure and NucleoSpin Gel and PCR Clean-up kits were purchased from BIOKE (Leiden, Netherlands). All the reagents used were bought from Sigma unless otherwise indicated. Oligonucleotides were synthesized by Biolegio (Nijmegen, Netherlands). Sequencing of all our constructs was performed at MacroGen (Amsterdam, Netherlands).

β-Galactosidase assay

For determination of β-galactosidase activity, strains were grown under identical conditions in LB media supplemented with 1.0% glucose and 0.1 mM IPTG, shaking at 37 °C and 220 rpm until the mid-exponential growth phase was reached (OD600 of 1.0). The cultures (1 ml) were immediately harvested and frozen in liquid nitrogen. The pellet was processed for β-galactosidase quantitation as previously described [42]. Each assay was performed in duplicate, and the mean value from three independent experiments was calculated.

Transcriptome analysis and qRT-PCR

Five ml samples were harvested from the culture as in the β-galactosidase assay, and the pellets were immediately shock frozen in liquid nitrogen when they reached OD600 of 1.0. The

Table 1. The main strains and plasmids used in this study.

Strain or plasmid	Genotype or properties	Reference or source
B. subtilis		
168	*trpC2*	[39]
168_β-gal	*trpC2, mdr::(*P*hyspank-lacZ spcr)*	This study
CodY^{-}	*trpC2, codY::cmr, mdr::(*P*hyspank-lacZ spcr)*	This study
CcpA^{-}	*trpC2, ccpA::ermr, mdr::(*P*hyspank-lacZ spcr)*	This study
CcpA^{-}CodY^{-}	*trpC2, ccpA::ermr, codY::cmr, mdr::(*P*hyspank-lacZ spcr)*	This study
CodY^{R214C}	*trpC2, codY R214C cmr, mdr::(*P*hyspank-lacZ spcr*)	This study
CodY$^{(E58A,I162V,R209H)}$	*trpC2, codY (E58A,I162V,R209H) cmr, mdr::(*P*hyspank-lacZ spcr*)	This study
CodY^{R214C}CcpA^{-}	*trpC2, ccpA::ermr, codY R214C cmr, mdr::(*P*hyspank-lacZ spcr)*	This study
CcpA^{T19S}	*trpC2, ccpAT19S kmr, mdr::(*P*hyspank-lacZ spcr)*	This study
CodY^{R214C}CcpA^{T19S}	*trpC2, ccpAT19S kmr, codYR214C cmr, mdr::(*P*hyspank-lacZ spcr)*	This study
E.coli		
MC1061	F-, *araD*139, Δ(*ara-leu*)7696, Δ(*lac*)X74, *galU*, *galK*, *hsdR2*, *mcrA*, *mcrB1*, *rspL*	[40]
BL21(DE3)	F^{-}, *omp*T, *hsd*SB (rB^{-}, mB^{-}), *gal, dcm* (DE3)	Laboratory stock
Plasmids		
pDOW01	*amyE::*P*hyspank spcr, lacI*	[41]
pDOW-CcpA	*amyE::*P*hyspank-ccpAhis8 spcr, lacI*	This study
pDOW-CcpA^{T19S}	*amyE::*P*hyspank-ccpAT19Shis8 spcr, lacI*	This study
pDOW-CodY	*amyE::*P*hyspank-his6codY spcr, lacI*	This study
pDOW-CodY^{R214C}	*amyE::*P*hyspank-his6codYR214C spcr, lacI*	This study

total RNA was extracted [43] and split into two aliquots for RNA sequencing and qRT-PCR. The sequencing of cDNA versions of the RNAs was accomplished by PrimBio (USA), and the data analysis was performed as before [44, 45]. Reverse transcription of the RNA samples was performed by using the SuperScript™ III Reverse Transcriptase kit, and quantitative PCR analysis was performed with the iQ5 Real-Time PCR Detection System (Bio-Rad) as described previously [46]. All the figures were generated by SigmaPlot 12.0 and were prepared for publication.

Electrophoretic mobility shift assay (EMSA)

E. coli strains BL21 (DE3) carrying plasmid pDOW in which different versions of the target genes were respectively cloned, and they were grown until the optical density at 600 nm (OD600) reached 0.7. Inducer IPTG (0.4 mM, final concentration) was added and then continued for six hours of incubation. Cells were harvested by 4 °C full speed centrifugation for 10 min and then lysed with 1 mg/ml lysozyme and sonication. CodY and CcpA proteins with six polyhistidine tag (His6-tag) were purified by Histrap™ excel column by following manufacturer's protocol (GE Healthcare Life Sciences). The purified protein samples were visualized by Coomassie blue stained sodium dodecyl sulfate (SDS)-polyacrylamide gels.

DNA probes were PCR amplified using the Cy3 labeled primers, and the acquired PCR products were purified by using the DNA clean-up kits (BIOKE). The DNA-target protein binding step was carried out in the presence of cofactors (FBP for CcpA, GTP and BCAAs for CodY) with 1X binding buffer, 0.2 µl of 10 mg/ml BSA, 5 nM of labeled DNA fragments and the purified His6-tagged proteins in different concentrations. The total volume was adjusted to 20 µl with MilliQ water and incubated at 30 °C for 20 min to complete the binding reaction. The obtained samples were loaded on a 5% nondenaturing polyacrylamide gel

(750 µl 40% acrylamide, 600 µl 5× TBE, 5 µl TEMED, 50 µl 10% APS, MilliQ up to 6 ml). Electrophoresis was carried out in 0.5% TBE buffer (pH 7.4) at 200 V for 30 min. Afterward, fluorescence signals were recorded using a Fuji LAS-4000 imaging system.

Western blot analysis

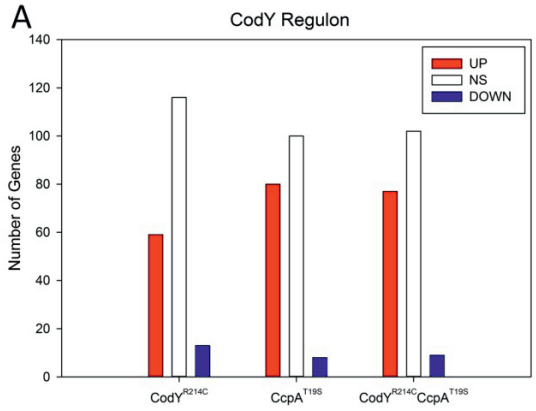
Immunoblot analysis was performed according to previous studies [3, 47]. First, 5 ml of cells were grown to OD600 of 1.0, and the cultures were harvested by centrifugation (14,000 rpm, 5 min). The pellets were resuspended in 1 ml solution A [50 mM Tris·Cl (pH 7.5), 5% (vol/vol) glycerol, and 1 mM PMSF]. Subsequently, the cells were broken by sonication for 2 min using 70% amplitude with 10-s bursts, 10-s pauses, and the total protein samples were collected by centrifugation (14,000 rpm, 5 min). 10 mg of total protein was heated for 5 min at 95 °C before being separated by SDS-PAGE, after which, proteins in the PAGE gel were transferred to a PVDF membrane (Millipore, USA). The membranes were blocked in PBST + 5% (wt/vol) BSA at 4 °C overnight. Subsequently, the membranes were separately subjected to a first incubation (90 min) with a rabbit anti-CodY [48] and rabbit anti-CcpA [49] polyclonal antibody (1:10,000) and a second incubation (90 min) with a donkey anti-rabbit IgG horseradish peroxidase (1:10,000) at room temperature. The signal intensity of bands was visualized using the ECL Prime kit (GE Healthcare Life Sciences) and detected by the Molecular Imager ChemiDoc XRS+ (Bio-Rad).

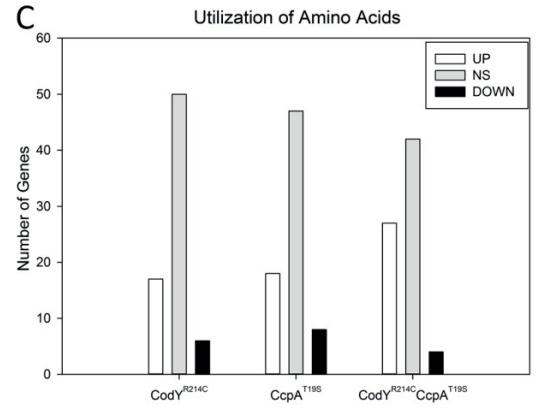
ACKNOWLEDGMENTS

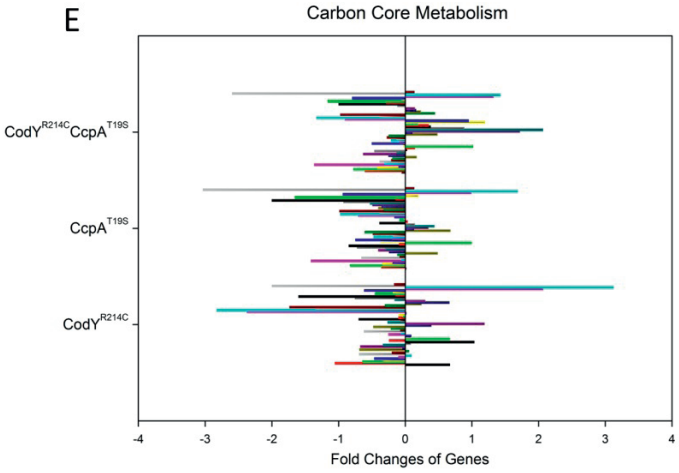
This research was partially funded by a grant from the former Kluyver Center for Genomics of Industrial Fermentation (Delft/Groningen) to JVH. HC was supported by a grant from China

Scholarship Council (CSC). We thank Dr. Lance Keller (Department of Fundamental Microbiology, University of Lausanne) for proofreading the manuscript. We are grateful to Anne de Jong (Department of Molecular Genetics, University of Groningen) for expert technical assistance of transcriptome analysis.

SUPPLEMENTARY MATERIALS







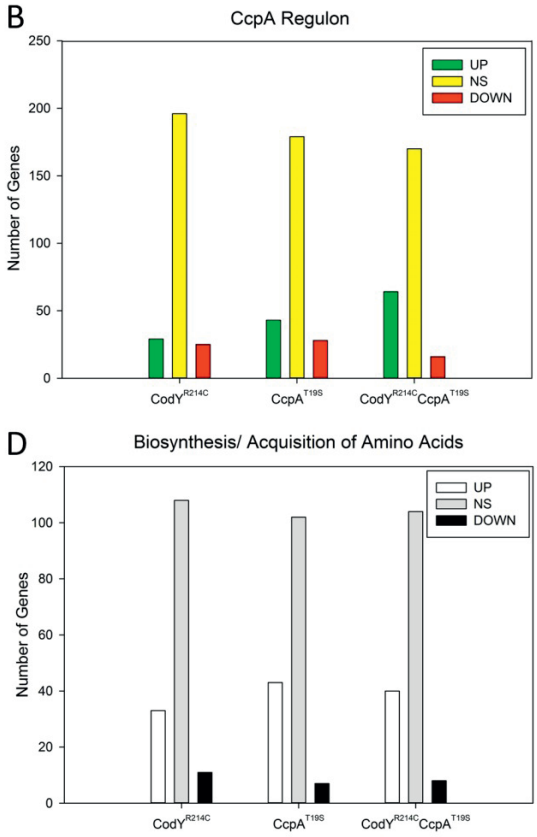


Fig. S1. Expression levels of selected genes in various reprogramming mutants relative to WT. All the CcpA/CodY regulon genes and carbon/nitrogen metabolism-related genes were grouped on the basis of the gene ontology (GO) classification in the *Subti*Wiki database. The up- and down-regulated genes were determined using the cutoff parameters $p < 0.05$ and fold-change > 2.0. Genes with no significant difference in transcript level were also identified ($p > 0.05$). **(A)** CodY regulon. **(B)** CcpA regulon. **(C)(D)** The genes are related to the utilization and biosynthesis/acquisition of amino acids. **(E)** The genes are related to carbon core metabolism (glycolysis, gluconeogenesis, pentose phosphate pathway, TCA cycle and carbon overflow).

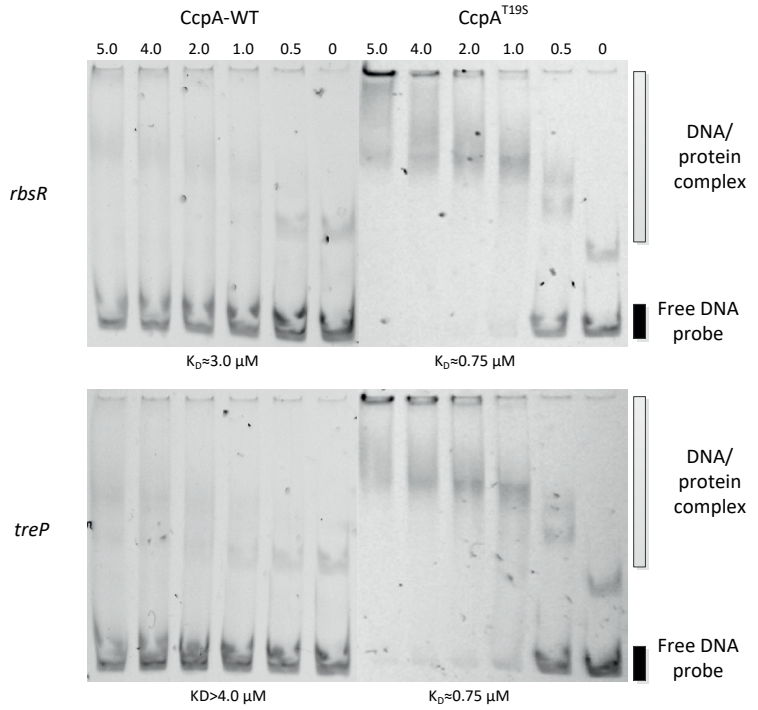


Fig. S2. DNA binding affinity of CcpA and its variant T19S to selected promoter regions. The promoter regions of *rbsR* and *treP*, which are negatively regulated by CcpA, were chosen as DNA probes for the EMSA experiment. The concentrations (μM) of proteins are indicated above each lane.

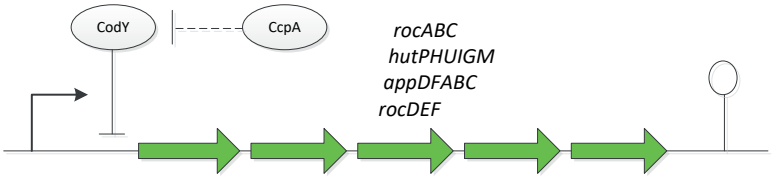


Fig. S3. Schematic representation of CcpA-mediated transcriptional regulation of operons that are repressed by CodY.

Table S1. Oligonucleotide primers used in this study.

Oligonucleotides	Sequence (5' -> 3')
ackA-PRO-F	TTGAAGACCGGACTTGACGAATTG
ackA-PRO-R	GATTGACGCTCCTTTATACTCTG
ilvB-PRO-F	CGAGGGAACAAGAGAAGTGCCTATC
ilvB-PRO-R	ACGGCTTTCCAGCTGTTCAAGAAGG
rbsR-F	AAGACTTTGTCAAAAAAAGAGTGAAAAC
rbsR-R	TAGCCGTTATCATTCAGGTTGC
treP-F	CCAGGGAACTGTCAATAAAGTATATG
treP-R	TCCGCCGTTAAAATGTTATTGATCCC
CcpA-EcoRI-F	ATATGAATTCATGAGCAATATTACGATCTAC
CcpA-His8_XhoI-R	ATATCTCGAGATGGTGATGGTGATGGTGATGGTGTGACTTGG TTGACTTTCTAAG
CodY-EcoRI_His6-F	ATATGAATTCCACCATCACCATCACCATGCTTTATTACAAAAAA CAAGAATTATTAAC
CodY-XhoI-R	ATATCTCGAGTTAATGAGATTTTAGATTTTCTAATTC

REFERENCES

1. Sonenshein AL, Hoch JA, Losick R. *Bacillus subtilis*: from cells to genes and from genes to cells. *Bacillus subtilis and its closest relatives*, (American Society of Microbiology), 2002;pp 3–5.
2. Shivers RP, Sonenshein AL. Activation of the *Bacillus subtilis* global regulator CodY by direct interaction with branched-chain amino acids. *Mol Microbiol.* 2004;53:599–611.
3. Brinsmade SR*, et al.* Hierarchical expression of genes controlled by the *Bacillus subtilis* global regulatory protein CodY. *Proc Natl Acad Sci U S A.* 2014;111:8227–8232.
4. Levdikov VM*, et al.* Structure of the branched-chain amino acid and GTP-sensing global regulator, CodY, from *Bacillus subtilis*. *J Biol Chem.* 2017;292:2714–2728.
5. Sonenshein AL. Control of key metabolic intersections in *Bacillus subtilis*. *Nat Rev Microbiol.* 2007;5:917–927.
6. Belitsky BR, Sonenshein AL. Genome-wide identification of *Bacillus subtilis* CodY-binding sites at single-nucleotide resolution. *Proc Natl Acad Sci U S A.* 2013;110:7026–7031.

7. Mijakovic I*, et al.* Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: A relic of early life? *Proc Natl Acad Sci U S A.* 2002;99:13442–13447.
8. Schumacher MA*, et al.* Structural basis for allosteric control of the transcription regulator CcpA by the phosphoprotein HPr-Ser46-P. *Cell* 2004;118:731–741.
9. Schumacher MA, *et al.* Structural mechanism for the fine-tuning of CcpA function by the small molecule effectors glucose 6-phosphate and fructose 1,6-bisphosphate. *J Mol Biol.* 2007;368:1042–1050.
10. Stulke J, Hillen W. Regulation of carbon catabolism in *Bacillus* species. *Annu Rev Microbiol.* 2000;54:849–880.
11. Fujita Y. Carbon catabolite control of the metabolic network in *Bacillus subtilis*. *Biosci Biotechnol Biochem.* 2009;73:245–259.
12. Shivers RP, Sonenshein AL. *Bacillus subtilis ilvB* operon: an intersection of global regulons. *Mol Microbiol.* 2005;56:1549–1559.
13. Tojo S*, et al.* Elaborate transcription regulation of the *Bacillus subtilis ilv-leu* operon involved in the biosynthesis of branched-chain amino acids through global regulators of CcpA, CodY and TnrA. *Mol Microbiol.* 2005;56:1560–1573.
14. Shivers RP, Dineen SS, Sonenshein AL. Positive regulation of *Bacillus subtilis ackA* by CodY and CcpA: establishing a potential hierarchy in carbon flow. *Mol Microbiol.* 2006;62:811–822.
15. Moreno MS, *et al.* Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. *Mol Microbiol.* 2001;39:1366–1381.
16. Molle V*, et al.* Additional Targets of the *Bacillus subtilis* global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript analysis. *J Bacteriol.* 2003;185:1911–1922.
17. Blencke HM*, et al.* Transcriptional profiling of gene expression in response to glucose in *Bacillus subtilis*: regulation of the central metabolic pathways. *Metab Eng.* 2003;5:133–149.
18. Belitsky BR, Sonenshein AL. Genetic and biochemical analysis of CodY-binding sites in *Bacillus subtilis*. *J Bacteriol.* 2008;190:1224–1236.

19. Marciniak BC*, et al.* High- and low-affinity cre boxes for CcpA binding in *Bacillus subtilis* revealed by genome-wide analysis. *BMC genomics.* 2012;13:401.
20. Geiger T, Wolz C. Intersection of the stringent response and the CodY regulon in low GC Gram-positive bacteria. *Int J Med Microbiol.* 2014;304:150–155.
21. Büscher JM*, et al.* Global network reorganization during dynamic adaptations of *Bacillus subtilis* metabolism. *Science.* 2012;335.
22. Calogero S*, et al.* RocR, a novel regulatory protein controlling arginine utilization in *Bacillus subtilis*, belongs to the NtrC/NifA family of transcriptional activators. *J Bacteriol.* 1994;176:1234–1241.
23. Gardan R, Rapoport G, Debarbouille M. Expression of the *rocDEF* operon involved in arginine catabolism in *Bacillus subtilis*. *J Mol Biol.* 1995;249:843–856.
24. Gopinath SCB*, et al.* Insights into anti-termination regulation of the *hut* operon in *Bacillus subtilis*: importance of the dual RNA-binding surfaces of HutP. *Nucleic Acids Res.* 2008;36:3463–3473.
25. Koide A, Hoch JA. Identification of a second oligopeptide transport system in *Bacillus subtilis* and determination of its role in sporulation. *Mol Microbiol.* 1994;13:417–426.
26. Belitsky BR. Indirect repression by *Bacillus subtilis* CodY via displacement of the activator of the proline utilization operon. *J Mol Biol.* 2011;413:321–336.
27. Pabo CO, Sauer RT. Transcription factors: structural families and principles of DNA recognition. *Annu Rev Biochem.* 1992;61:1053–1095.
28. Joseph P, Ratnayake-Lecamwasam M, Sonenshein AL. A region of *Bacillus subtilis* CodY protein required for interaction with DNA. *J Bacteriol.* 2005;187:4127–4139.
29. Ishii H, Tanaka T, Ogura M. The *Bacillus subtilis* response regulator gene *degU* is positively regulated by CcpA and by catabolite-repressed synthesis of ClpC. *J Bacteriol.* 2013;195:193–201.
30. Goel A, *et al.* Metabolic shifts: a fitness perspective for microbial cell factories. *Biotechnol Lett.* 2012;34:2147–2160.
31. Fink PS. Biosynthesis of the branched-chain amino acids. *Bacillus subtilis and other gram-positive bacteria*, (American Society of Microbiology), 1993;pp 307–317.

32. Michna RH, *et al. Subti*Wiki 2.0—an integrated database for the model organism *Bacillus subtilis*. *Nucleic Acids Res.* 2015;44:654–662.
33. Chubukov V, *et al.* Coordination of microbial metabolism. *Nat Rev Microbiol.* 2014;12:327–340.
34. Kochanowski K, Sauer U, Chubukov V. Somewhat in control–the role of transcription in regulating microbial metabolic fluxes. *Curr Opin Biotech.* 2013;24:987–993.
35. Cao H, *et al.* Cell surface engineering of *Bacillus subtilis* improves production yields of heterologously expressed alpha-amylases. *Microb Cell Fact.* 2017;16:56.
36. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.
37. Konkol MA, Blair KM, Kearns DB. Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J Bacteriol.* 2013;195:4085–4093.
38. Norholm MH. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnol.* 2010;10:p. 1.
39. Kunst F, *et al.* The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature.* 1997;390.
40. Casadaban MJ, Cohen SN. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J Mol Biol.* 1980;138:179–207.
41. Weme RGDO, *et al.* Single cell FRET analysis for the identification of optimal FRET-pairs in *Bacillus subtilis* using a prototype MEM-FLIM system. *PLoS One.* 2015;10:e0123239.
42. Smale ST. β-Galactosidase assay. *Cold Spring Harb Protoc.* 2010;5:pdb. prot5423.
43. Yi Y, *et al.* Comparative transcriptomics of *Bacillus mycoides* strains in response to potato-root exudates reveals different genetic adaptation of endophytic and soil isolates. *Front Microbiol.* 2017;8:1487.
44. de Jong A, *et al.* T-REx: Transcriptome analysis webserver for RNA-seq expression data. *BMC Genomics* 2015;16:663.
45. van der Meulen SB, de Jong A, Kok J. Transcriptome landscape of *Lactococcus lactis* reveals many novel RNAs including a small regulatory RNA involved in carbon uptake and metabolism. *RNA Biol.* 2016;13:353–366.

46. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}$T method. *Methods.* 2001;25(4):402–408.
47. Weme RGDO, Seidel G, Kuipers OP. Probing the regulatory effects of specific mutations in three major binding domains of the pleiotropic regulator CcpA of *Bacillus subtilis*. *Front Microbiol.* 2015;6:1051.
48. Ratnayake-Lecamwasam M, Serror P, Wong KW, Sonenshein AL. *Bacillus subtilis* CodY represses early-stationary-phase genes by sensing GTP levels. *Gene Dev.* 2001;15:1093–1103.
49. Küster E, *et al.* Immunological crossreactivity to the catabolite control protein CcpA from *Bacillus megaterium* is found in many Gram-positive bacteria. *FEMS Microbiol Lett.* 1996;139:109–115.

CHAPTER 5

Influence of global gene regulatory networks on single cell heterogeneity of green fluorescent protein production in *Bacillus subtilis*

Haojie Cao[1], Oscar P. Kuipers[1]

[1]Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands.

This chapter is under review in *Microb Cell Fact*:
Haojie Cao, Oscar P. Kuipers. Influence of global gene regulatory networks on single cell heterogeneity of green fluorescent protein production in *Bacillus subtilis*.

ABSTRACT

In the past decades, the Gram-positive bacterium *Bacillus subtilis* has been extensively studied as a microbial cell factory for the production of industrially and medically relevant products. Green fluorescent protein (GFP) is commonly used as a marker for determining the strength of a given promoter or the subcellular localization of a fusion protein. However, inherent heterogeneity of GFP expression among individual cells that can arise from global regulation differences in the expression host, has not yet been fully assessed. Here, we investigate the dynamic production performance of GFP in *B. subtilis* reporter strains, with single mutation(s) in the two major transcriptional regulators CcpA and/or CodY that were earlier found to improve overall heterologous protein production levels, by flow cytometry and fluorescence microscopy. We discovered that the transcriptome perturbations caused by the mutated global regulators affect the production of superfolder GFP -sfGFP(Sp) during growth and significantly reduce the heterogeneity that is prominent in the wildtype (WT) cells. The mutation R214C in the DNA-binding domain of CodY effectively reduces the extrinsic noise of sfGFP(Sp) synthesis and enhances GFP production at the population level. Single-cell analysis of GFP expression demonstrated that cells harboring the amino acid substitution $CodY^{R214C}$ showed much lower phenotypic heterogeneity of fluorescence signals relative to two other strains, i.e. WT and $CcpA^{T19S}$.

Keywords: *Bacillus subtilis*, superfolder green fluorescent protein (sfGFP), heterogeneous expression, global transcriptional regulation, production level, phenotypic noise

INTRODUCTION

The gradual but very rapid accumulation of genetic information and fast development of experimental approaches have opened up many new frontiers in cellular investigation [1]. The traditional bulk-scale measurements that only investigate the average values for a population of cells give an incomplete picture of what happens in bacterial cultures. The information on individual cells is needed for correctly monitoring biological processes. It has become evident that various subpopulations of bacteria can exist under certain conditions, with cells in distinct physiological or developmental states [2, 3]. Multiple studies have been focused on the development and utilization of single-cell techniques, which aid the research on the cellular behavior of individual cells in bacterial populations [4, 5].

It is widely recognized that bacterial cells with the same genetic information (clonal populations) can display a multitude of distinct phenotypes, even when exposed to the same environment, this phenomenon is known as phenotypic heterogeneity [6]. *Bacillus subtilis*, the best-characterized member of low GC Gram-positive bacterial species, has been studied extensively with respect to phenotypic diversity. When nutrient is limited, *B. subtilis* in the stationary phase generates a mixed population, in which some cells form spores that are highly resistant to external stresses [7]. Additionally, a subset of cells that have entered into the sporulation state can secrete an extracellular 'killing factor' and toxin to block sister cells from sporulating and to stimulate the lysis of them [8]. In certain conditions, a subpopulation of the *B. subtilis* cells can enter into the competent state, enabling them to take up DNA from the environment [9, 10]. Heterogeneity also plays an important role in biofilm formation, which resulted by a subpopulation generating extracellular matrix material that tightly holds the surrounding

cells together to form a robust biofilm [11]. Moreover, during exponential growth, a fraction of cells manages to express *sigD*, which is necessary for flagellar production, resulting in the cells to be motile [2].

Phenotypic heterogeneity, which mostly results from heterogeneous gene expression, increases the survival chance of a subpopulation that is better adapted to changing conditions [12-15]. Three factors are considered as the source of dynamic cellular behavior: i) the circuit architecture or regulatory interaction patterns; ii) quantitative parameters, such as promoter strengths; and iii) stochastic fluctuations or "noise," which depends on the availability of certain cellular components [16]. In general, the noise of gene expression arises from two sources. The "intrinsic" noise is generated by the inherent stochasticity of biochemical processes such as transcription and translation, causing identical copies of a gene to be expressed at different levels. On the other hand, the fluctuations in the states or accumulations of crucial cellular components such as regulatory proteins and polymerases represent "extrinsic" noise, leading indirectly to particular gene expression variation and which has a global effect [4, 17].

A wide variety of proteins have been chosen as reporters for benchmarking gene expression in order to study the mechanisms of phenotypic heterogeneity. In *B. subtilis*, the mostly used reporters include *lacZ*, encoding the β-galactosidase from *E. coli* [18], *luxAB*, encoding the luciferase from *Vibrio harveyi* [19], *mCherry*, encoding an enhanced red fluorescent protein from *Discosoma sp.* [20] and *gfp*, encoding the green fluorescent protein (GFP) from *Aequorea victoria* [21]. GFP and its derivatives have been extensively utilized in the study of protein localization or promoter activity in live cells [22], which has tremendously increased our knowledge of bacterial cell biology [23-25]. These analyses can be carried out using flow cytometry, fluorescent

microscopy or both [26, 27]. Flow cytometry facilitates the rapid analysis of cells in the population, while time-lapse microscopy follows the behavior of individual cells over time and dynamic movements of proteins within a single cell [28–31]. A previous study from our lab benchmarked the expression of a library of GFP variants in three model microorganisms, *i.e. B. subtilis, Streptococcus pneumoniae,* and *Lactococcus lactis* [32]. Surprisingly, the superfolder GFP with codon optimization specifically for *S. pneumoniae* -sfGFP(Sp) displayed the highest fluorescence intensity and relatively low phenotypic noise in *B. subtilis.*

In an earlier study, we explored the heterologous protein production potential of *B. subtilis* by genetically altering its two global transcriptional regulators (Chapter 3), which demonstrated that two mutations, i.e. $CodY^{R214C}$ and $CcpA^{T19S}$ in one cell resulted in the reorganization of metabolic networks, which eventually improved the intracellular synthesis of β-galactosidase (β-gal) and other soluble proteins. In the present study, the robustly folded version of GFP -sfGFP(Sp) was utilized as the reporter protein to quantify the productivity of the obtained mutant $CodY^{R214C}CcpA^{T19S}$ over time, both at the population and single-cell level. Notably, this investigation points to altered production levels of GFP and great variation between single cells, depending on the central regulatory metabolic pathways operating in the WT and mutant cells.

RESULTS AND DISCUSSION

The alteration of global regulatory networks significantly impacts the GFP production in *B. subtilis*

As presented previously, the strain $CodY^{R214C}CcpA^{T19S}$ with rewired metabolic pathways displays a 2-fold increase of β-galactosidase production in comparison to the WT. To investigate

the expression of another classic reporter, GFP, in the genetically modified expression hosts, the sfGFP(Sp) was utilized in this research. Moreover, since the plasmid-based expression systems can cause additional heterogeneity due to copy number variation and polar fixation effects [33, 34], we integrated the expression cassette P*hyspank*-sfGFP(Sp) into the *amyE* locus in *B. subtilis* 168 WT, CodY^{R214C}, CcpA^{T19S}, CodY^{R214C}CcpA^{T19S} to obtain the four reporter strains.

Subsequently, we grew all the strains and induced the GFP expression identically in microtiter plates, and the fluorescence and growth were monitored using a plate reader (Varioskan-LUX, Thermo Fisher) over time. As shown in **Fig. 1A**, during the 22 hour's incubation, the host CodY^{R214C} and CodY^{R214C}CcpA^{T19S} produced higher levels of GFP, while the WT and CcpA^{T19S} generated relatively lower amounts of GFP under identical culture conditions. Since only a rough estimation of the fluorescence intensity at the population level can be determined in the microtiter plate reader, and the corresponding fluorescence signals were getting variable after five hours, the cultures of CodY^{R214C}CcpA^{T19S} and WT at that time point were subjected to fluorescence microscopy for visualizing and comparing the GFP expression at the single-cell level. As illustrated in **Fig. 1B**, there was a clear fluorescence signal variation among the WT cells, which demonstrated that the expression of the sfGFP in *B. subtilis* 168 is heterogeneous. In comparison, the fluorescent signals of individual CodY^{R214C}-CcpA^{T19S} cells were more homogeneous (**Fig. 1B**). Taken together, the overall GFP production was different in individual cells of the *B. subtilis* strains with various versions of CodY and/or CcpA. Compared with the WT control, the hosts containing the mutation CodY^{R214C} could significantly increase green fluorescent protein production, as was the case for β-galactosidase production (Chapter 3). Notably, the superfolder GFP was most heterogeneously expressed in WT cells.

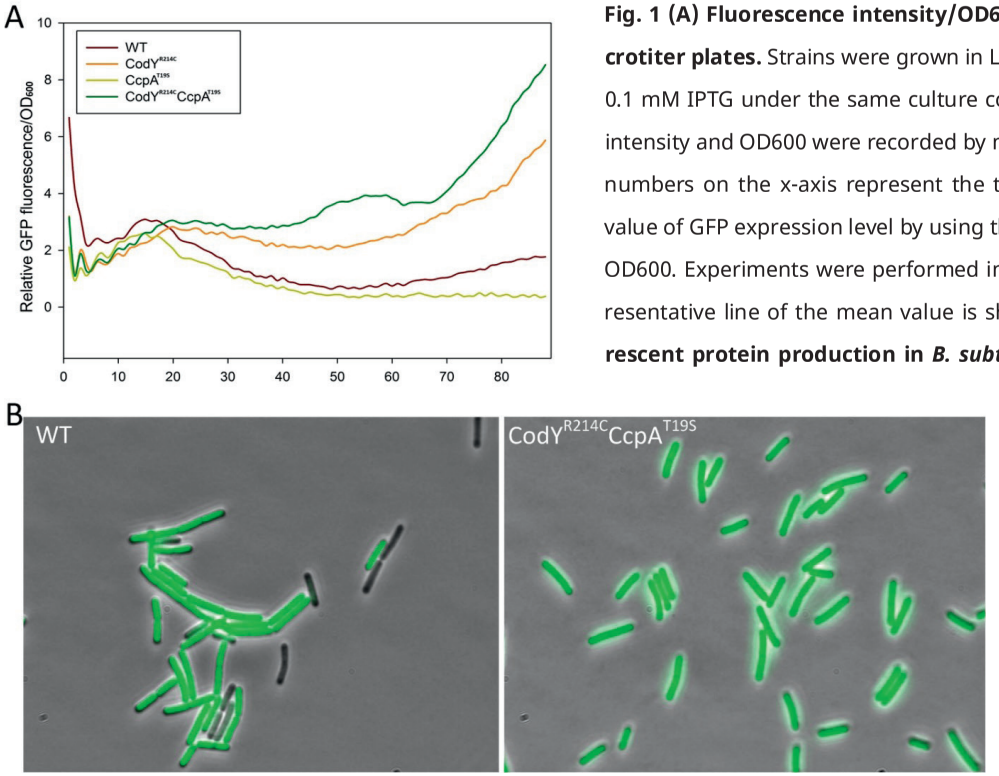
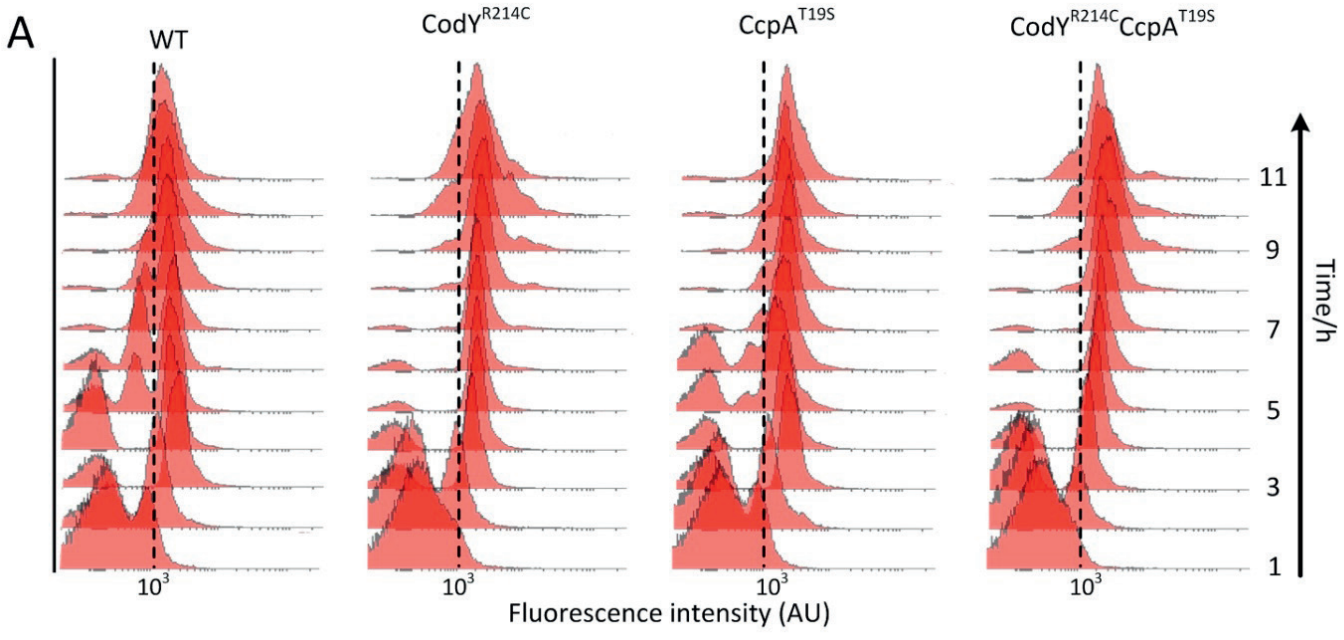


Fig. 1 (A) Fluorescence intensity/OD600 of various *B. subtilis* strains in microtiter plates. Strains were grown in LB supplemented with 1.0% glucose and 0.1 mM IPTG under the same culture condition (37 °C, 220 rpm). Fluorescence intensity and OD600 were recorded by microplate reader every 15 minutes, the numbers on the x-axis represent the time points. We calculated the relative value of GFP expression level by using the formula: GFP fluorescence intensity/ OD600. Experiments were performed in triplicate, but for clarity, only one representative line of the mean value is shown. **(B) Visualization of green fluorescent protein production in *B. subtilis* by fluorescence microscopy.** The overnight pre-culture was diluted to OD600 of 0.1 in fresh production media (LB, 1.0% glucose, 0.1 mM IPTG). Subsequently, the mixture was incubated in flasks at 37 °C, 220 rpm for five hours, and then the culture was immediately taken for fluorescence microscopy.


A
WT
CodY^{R214C}
CcpA^{T19S}
CodY^{R214C}CcpA^{T19S}
11
9
7
5
3
1
Time/h
10^3
10^3
10^3
10^3
Fluorescence intensity (AU)

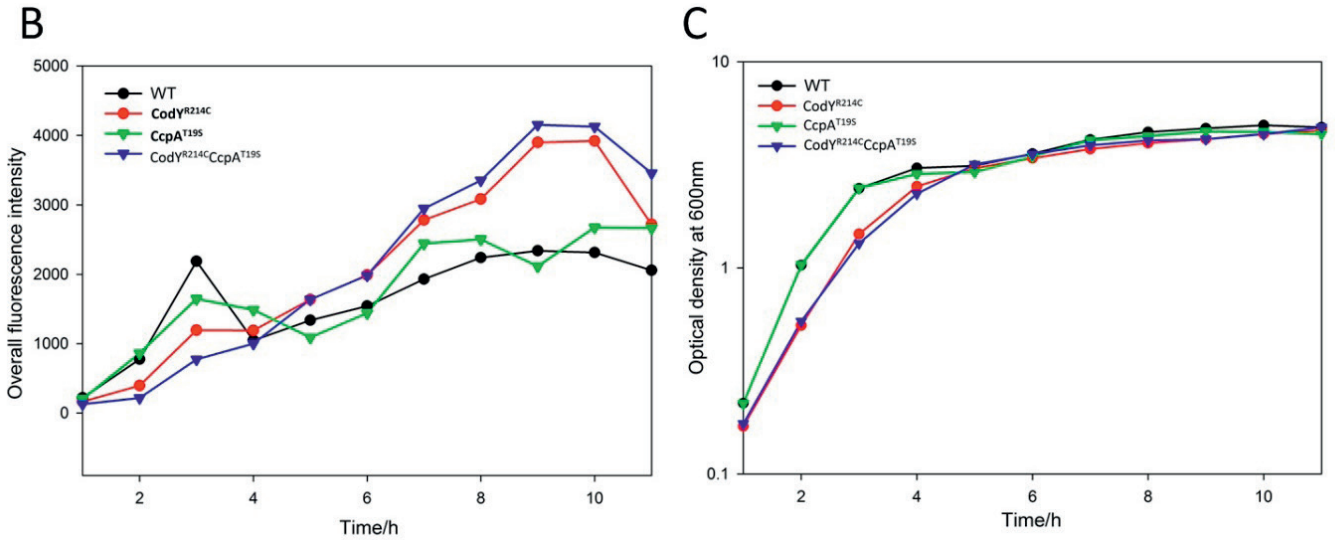


Fig. 2 The expression of sfGFP(Sp) in various *B. subtilis* strains. *B. subtilis* WT, CodY^{R214C}, CcpA^{T19S}, CodY^{R214C}CcpA^{T19S} harboring *amyE*::P*hy-spank-sfgfp(Sp)* were grown in flasks with LB supplemented with 1.0% glucose and 0.1 mM IPTG under the same growth conditions (37 °C, 220 rpm). Samples were harvested for both fluorescence and OD600 measurement per hour. *(A)* Flow cytometric analysis of GFP expression. Dotted lines were placed at 10^3 Arbitrary Units (AU) to serve as a reference of the fluorescence distributions. *(B)* The mean fluorescence intensity of the whole population over time. *(C)* The optical density at 600 nm of various strains was measured by spectrophotometry.

The rewired central nitrogen metabolism plays a crucial role in the GFP production enhancement

To reveal the mechanism behind the upshift of GFP production and to elucidate cellular behavior during expression, fluorescence microscopy and flow cytometric analysis of GFP production in the four strains (168, CodY^{R214C}, CcpA^{T19S}, CodY^{R214C}-CcpA^{T19S}) were performed in parallel. **Fig. 2A** shows the flow cytometry tracings of the four mutants when cultured under the same conditions. The corresponding mean fluorescence intensity and optical density for each time point are presented in **Fig. 2B** and **Fig. 2C,** respectively. In line with the prior observation, the CodY^{R214C} and CodY^{R214C}CcpA^{T19S} showed higher GFP signals than the other strains at the population level. The WT and CcpA^{T19S} exhibited similar curves to each other concerning the growth and the fluorescence intensity, being significantly different from that of CodY^{R214C} and CodY^{R214C}CcpA^{T19S}, which showed similar growth and GFP production to each other. WT and CcpA^{T19S} reached stationary phase one hour earlier than the two strains containing CodY^{R214C} (**Fig. 2C**). The GFP production level in the latter two hosts, especially during the stationary phase, was higher than that of the former two (**Fig. 2B**). Furthermore, there was a detectable decline of mean fluorescence intensity in 50,000 cells of WT and CcpA^{T19S} after the first three hour's gradual rise. In contrast, the accumulation of GFP in CodY^{R214C} and CodY^{R214C}CcpA^{T19S} improved continuously until the late stationary phase. In summary, the amino acid substitution R214C in CodY caused a stronger GFP synthesis ability at a slight expense of growth rate, while the mutation CcpA^{T19S} did not play a positive role in the expression of the reporter protein-sfGFP(Sp) in *B. subtilis*.

Phenotypic noise, related to global regulation, negatively correlates to the overall GFP production level

The distribution of the expression of a single gene can be defined by the mean value of expression level indicated by <*p*> with a standard deviation-*σp* or coefficient of variation (CV) [35]. The phenotypic noise strength (*σp*/<*p*>), is extensively applied for the measure of noise [1, 15, 36]. Based on the data from the flow cytometric analysis, we quantified the spread of GFP fluorescence signals in a population of various strains. Since the different versions of the regulator(s) in the expression hosts are the only variable during the GFP synthesis process, the extrinsic noise that arises from the regulation, should play a crucial role in the final GFP yield. As shown in **Fig. 3A**, the noise strength of the GFP expression in *B. subtilis* is dynamic over time. Overall, the phenotypic noise was high at the beginning of growth and then dropped sharply in the following four hours (**Fig. 3A**). This is probably due to the IPTG induction, which controls the GFP production, does not start simultaneously in different cells [37]. After remaining at a steady state for an extended period, the noise increased again when cultures reached late stationary phase (**Fig. 3B**). In addition, a significant difference with regard to phenotypic noise was observed from the four assessed strains *after 8 hours* of growth. The CcpA^{T19S} strain showed the strongest noise value of GFP expression compared to the other three hosts, and the CodY^{R214C}CcpA^{T19S} strain showed the lowest noise among all the expression hosts. We thus conclude that the strength of noise is opposed to the corresponding mean fluorescence intensity in various strains. This indicates that the different versions of global regulators cause diverse extrinsic noise levels during the overexpression of sfGFP(Sp), which eventually results in different levels of the overall GFP yield.

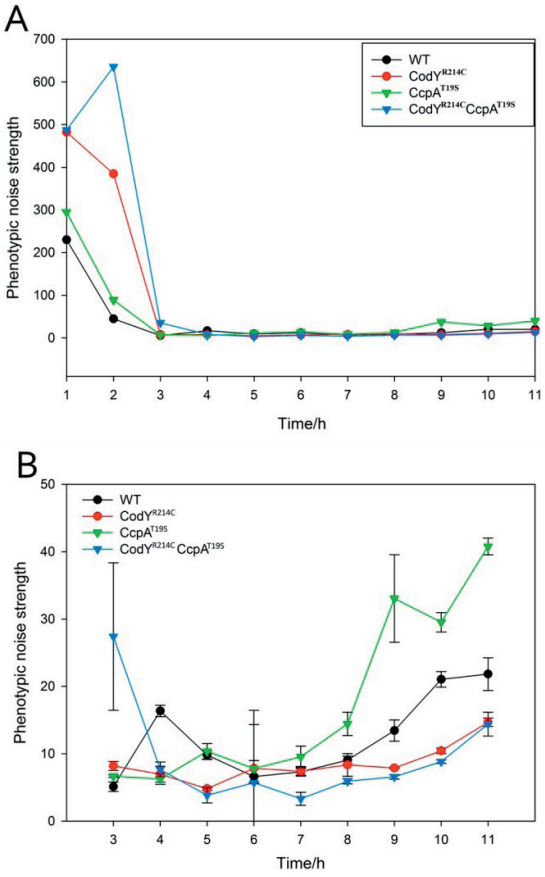


Fig. 3 The phenotypic noise of GFP expression in various hosts. The phenotypic noise was calculated by using the formula: $\sigma p^2/\langle P \rangle$ (variance/mean), σp was also named the coefficient of variation (CV) in the flow cytometric analysis. All the experiments were performed in triplicate, but for clarity, only the average lines of whole 11 hours are shown in A, while the average lines with error bars from 3 to 11 hours are presented in B.

Characterization of GFP production at the single-cell level

Fluorescence microscopy was carried out to visualize the production of sfGFP(Sp) in single cell per hour. Here, we picked three representative images of the cells in exponential,

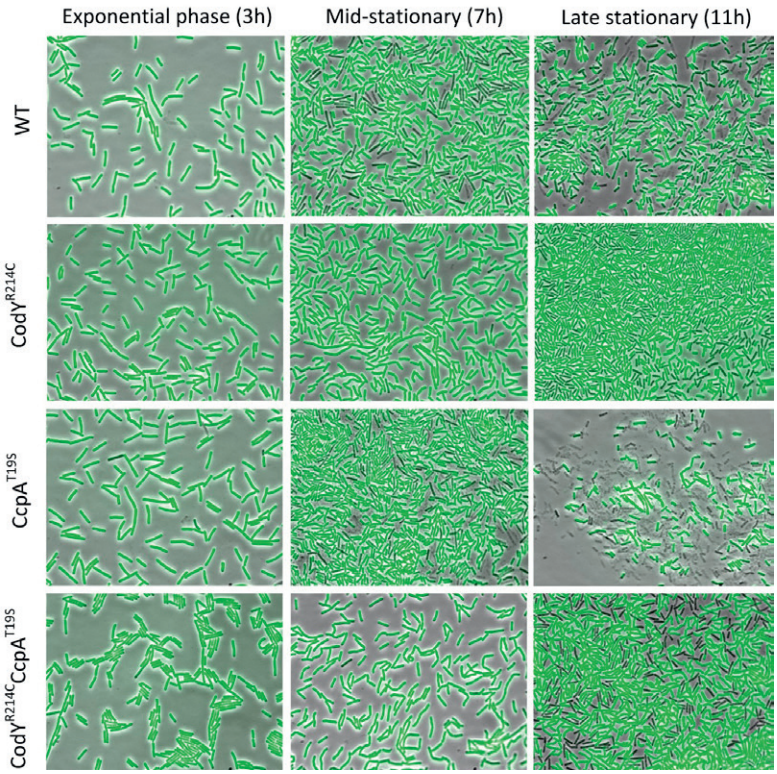


Fig. 4 Phenotypic heterogeneity of various strains during growth. The strains were grown at 37 °C, 220 rpm in LB supplemented with 1.0% glucose and 0.1 mM IPTG for 11 hours. The GFP fluorescence images and phase contrast images of cells at different time points were acquired, and the merged micrographs are presented.

mid-stationary, and late stationary phase for further analysis. As indicated in **Fig. 4**, during the exponential phase, all the cells of the four detected strains show strong signal and similarity in the fluorescence intensity. When the cultures reached the stationary phase, most cellular heterogeneity with respect to fluorescence occurred among the cells of WT and CcpA^{T19S}. This phenotypic diversity is most prominent during mid-stationary growth after 7 hours. Dark cells with low GFP activity co-exist

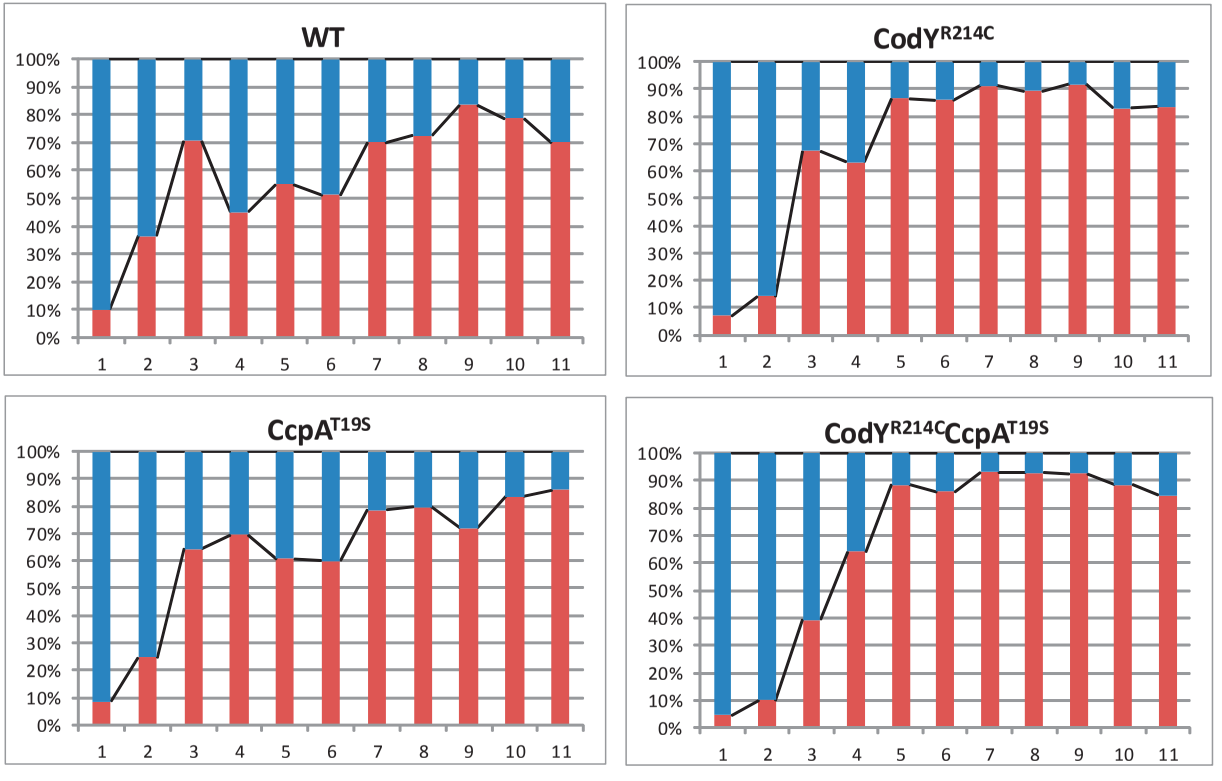


Fig. 5 The dynamic proportion of the two GFP intensity subpopulations. The red bars represent positive subpopulations (>10^3 AU), and the blue bars represent negative subpopulations (<10^3 AU). The numbers on the x-axis represent the time points (hour).

with the cells having strong GFP intensity in the cultures of the above two strains. During the mid-stationary growth phase, cellular heterogeneity of the other two strains, namely, CodY^{R214C} and CodY^{R214C}CcpA^{T19S}, was hardly visible. Finally, the GFP is expressed heterogeneously in the strains with CodY^{R214C} in the late stationary phase, while the cells of the other two strains, especially the CcpA^{T19S}, already lysed severely. This is consistent with the observation in **Fig. 2B**, the GFP intensity in CodY^{R214C} and CodY^{R214C}CcpA^{T19S} reduced at the end of 11 hours' expression. This reflects that the activity of cellular processes decreased owing to the short supply of essential nutrient sources when the strains entered into the late stationary phase. During the same growth phase, the GFP production level in CcpA^{T19S} went up **(Fig. 2B)** because most of the dark cells lysed and only the ones with high GFP intensity survived and could be detected by FACS.

Characterization of GFP production at the subpopulation level

To further study GFP production in subpopulations, we analyzed the flow cytometry results of different strains by Flowing Software. We set the fluorescence intensity 10^3 AU as the cutoff value and defined the subpopulations as negative (<10^3 AU) or positive (>10^3 AU). As displayed in **Fig. 5**, the two strains harboring the WT version of CodY showed similarity in the percentage of the two subpopulations, while the two hosts carrying CodY^{R214C} also shared similar subpopulation proportions. During the stationary growth phase, the overall percentages of positive subpopulations for the CodY^{R214C} and CodY^{R214C}CcpA^{T19S} strains were obviously higher than that of the WT and CcpA^{T19S}. If we combine **Fig. 2** and **Fig. 5**, it is interesting to note that the positive percentages show high consistency with GFP expression performance in expression hosts harboring various versions of

CodY and/or CcpA. The overall fluorescence signal strength depends on the positive subpopulations in various strains.

Metabolic burden might affect the heterologous expression of GFP

Metabolic burden, a known phenomenon for heterologous expression systems, is caused by the fact that the overexpression pathways of foreign proteins can take up a large proportion of the nutrient source fluxes, which then influences the original metabolic distribution in the cell, and cause serious physiological problems and finally results in lower yields of target products [38–40]. In a previous study (Chapter 4), we reprogrammed the metabolic regulatory networks, and found that a more strongly repressed carbon metabolism and de-repressed nitrogen metabolism coordinately contribute to an increase of the reporter protein β-galactosidase production in *B. subtilis*. The production improvements were found to be consistent with upregulation of several nitrogen metabolic operons, and this was regarded to reduce the metabolic burden of β-gal overexpression in the genetically modified strains. The balanced and modified metabolic networks with increased uptake and utilization ability of arginine, ornithine, citrulline, and histidine could also weaken the extrinsic noise of GFP expression in the $CodY^{R214C}CcpA^{T19S}$. Different from the previous observation, strain $CcpA^{T19S}$ does not have an advantage in the expression of sfGFP(Sp), which is slightly lower than the WT control. This is in accordance with the fact that protein production improvement is performed in a protein-specific way [41]. Nevertheless, based on population-scale analysis, the mutation $CcpA^{T19S}$ can still further improve the GFP expression on the basis of the improvement in $CodY^{R214C}$. This shows that the effects of mutation $CodY^{R214C}$ and $CcpA^{T19S}$ on the final production of sfGFP(Sp) are more complex than a simple addition. To sum up, the

CodY^{R214C}CcpA^{T19S} strain displays balanced metabolic flux distributions between essential cellular processes and heterologous over-expression pathway probably has a lower metabolic burden. This not only increased the overall product yield but also decreased the phenotypic heterogeneity of sfGFP(Sp) expression in *B. subtilis*, a property generally useful for overproduction of any soluble intracellular protein.

CONCLUDING REMARKS

In this study, we investigated the production of sfGFP(Sp) in strains with mutation(s) in CodY and/or CcpA and the WT strain as the control. We demonstrated that the mutation CodY^{R214C} improves the overall expression of reporter protein sfGFP(Sp) significantly, with a slight decrease of the growth rate, while the CcpA^{T19S} mutant slightly reduces the GFP synthesis. Nevertheless, when the two amino acid substitutions among the DNA-binding HTH motif of CodY and CcpA were combined, this yielded the best GFP producer - CodY^{R214C}CcpA^{T19S}. Furthermore, the phenotypic noise clearly differs between different mutants of the global regulator(s). This extrinsic noise comes from global regulation and is shown to be negatively correlated with GFP production in our cell factories. In addition, the single-cell and subpopulation analyses demonstrated that the cells of WT and CcpA^{T19S} show stronger heterogeneity during the expression process over time. Although the full understanding of the mechanisms underlying expression heterogeneity is still incomplete, this study provides novel insights into decreasing cellular diversity and directs the way to further increase heterologous protein production in cell factories.

MATERIALS AND METHODS

Plasmids, bacterial strains, and media

The plasmids and bacterial strains used in this study are listed in **Table 1**. All the *Bacillus subtilis* and *E. coli* were grown at 37 °C with shaking (220 rpm) in liquid Lysogeny Broth (LB) unless otherwise indicated. For solid media, 1.5% (wt/vol) agar was added to the LB. Antibiotics were added when necessary as follows: 100 mg/ml ampicillin for *E. coli*, 5 mg/ml kanamycin and chloramphenicol, 100 mg/ml spectinomycin for *B. subtilis*. When required, 0.1 mM IPTG (isopropyl-β-D-thiogalactosidase) was added to the media for activation of the IPTG-inducible expression system.

Recombinant DNA techniques and oligonucleotides

Procedures for DNA purification, restriction, ligation, gel electrophoresis and transformation of *E. coli* were carried out as previously described [44]. *B. subtilis* was naturally transformed as described before [45]. T4 DNA ligase, Fastdigest Restriction enzymes and DNA polymerases (Phusion and DreamTaq) were purchased from Thermo Fisher Scientific (Landsmeer, Netherlands). Chromosomal DNA of the *B. subtilis* 168 and the constructed plasmids in this research were used as templates for PCR. The NucleoSpin® Plasmid EasyPure and NucleoSpin® Gel & PCR Clean-up kits were purchased from BIOKE (Leiden, Netherlands). All the reagents used were bought from Sigma unless otherwise indicated. Oligonucleotides were synthesized by Biolegio (Nijmegen, Netherlands). Sequencing of all our constructs was performed at MacroGen (Amsterdam, Netherlands).

Construction of bacterial strains

B. subtilis strain 168_sfGFP(Sp)_CodY^{R214C} was obtained by homologous double crossover recombination of plasmid pJV153

Table 1. The plasmids and bacterial strains used in this study

Strains and plasmids	Phenotype or property	Source or reference
Stains		
168	*trpC2*	[42]
168_sfGFP(Sp)	*trpC2, amyE::Physpank-sfgfp(Sp) spcr*	[32]
168_sfGFP(Sp)_CodY^{R214C}	*trpC2, codY R214C cmr, amyE::Physpank-sfgfp(Sp) spcr*	This study
168_sfGFP(Sp)_CcpA^{T19S}	*trpC2, ccpAT19S kmr, amyE::Physpank-sfgfp(Sp) spcr*	This study
168_sfGFP(Sp)_CodY^{R214C}CcpA^{T19S}	*trpC2, codY R214C cmr, ccpAT19S kmr, amyE::Physpank- sfgfp(Sp) spcr*	This study
E.coli		
MC1061	F$^-$, *araD*139, Δ(*ara-leu*)7696, Δ(*lac*)X74, *galU*, *galK, hsdR2, mcrA, mcrB1, rspL*	[43]
Plasmids		
pCH3_CcpA^{T19S}	pUC18_*aroA_ccpAT19S_kmr_ytxD*	Chapter 3
pJV153	pUC18 *clpY codYR214C cmr flgB*	Chapter 3

into the flanking region of *codY* in *B. subtilis* 168. Strain 168_sfGFP(Sp)_CcpA^{T19S} was obtained by the integration of plasmid pCH3_CcpA^{T19S} into the specific chromosomal region of *B. subtilis* 168. Transformants were selected on LB agar plates containing appropriate antibiotic(s), after overnight incubation at 37 °C. Correct integration was verified by PCR and sequencing analysis. The strain 168_sfGFP(Sp)_CodY^{R214C}CcpA^{T19S} was constructed in the same way as described above.

Microplates experiments

Single colonies of required strains were picked from LB agar plates with antibiotics and were incubated at 37 °C,220 rpm overnight. The day after, the O/N cultures were diluted in a 96-well microtiter plate to OD600-0.1 with 200 l fresh LB media containing 1.0% glucose and 0.1 mM ITPG. Plates were incubated at 37 °C and 220 rpm shaking in the plate reader-VarioskanLUX (Thermo

Fisher) with a GFP filter set (excitation at 485/20 nm, emission 535/25), and the absorbance was measured at 600 nm. The values of GFP intensity and OD600 were automatically recorded every 15 minutes for 22 hours, data of all samples were collected in triplicates. All the optical density and fluorescence values were corrected for the background of the media by the following formula: (GFPreporter-GFPmedia)/(ODreporter-ODmedia) [46].

Flow cytometry

All the strains were streaked on LB agar plates supplemented with a specific antibiotic, and the single colonies were picked up and grown overnight in LB media at 37 °C,220 rpm. Next morning, the pre-cultures were diluted to OD600-0.1 in fresh LB supplemented with 1.0% glucose and 0.1 mM IPTG and further incubated in a 37 °C shaker. Subsequently, the cultures of each time point were prepared for flow cytometry as described before [3, 12]. Cells were diluted 10–20 times in phosphate buffered saline (PBS) and directly measured on the Becton Dickinson FACSCanto (BD BioSciences, USA) with an Argon laser (488 nm). For each sample, the green fluorescent signals of 50,000 cells were collected by a FITC filter. The fluorescent intensity was calculated in Arbitrary Units (AU). All the captured data was further analyzed using Flowing Software (http://www.flowingsoftware.com/).

Fluorescence Microscopy

In parallel, the above-described cultures of each time point were also prepared for fluorescence microscopy and applied to agarose slides as described before [25]. The expression of the fluorescent protein was analyzed by fluorescence microscopy (Nikon Eclipse Ti, Japan) equipped with a CoolsnapHQ2 CCD camera. Fluorescent signals from cells were visualized using 450–490 nm excitation and 500–550 nm emission for GFP

fluorescence channel and an Intensilight light as phase contrast channel. Software NIS-Elements AR [47] was used for image capturing by 0.2 s exposure, and the final images for publication were generated by ImageJ software [48].

ACKNOWLEDGMENTS

We thank Ard Jan Grimbergen (Molecular Genetics, University of Groningen) for critically reading the manuscript. We are grateful to Yanglei Yi (Molecular Genetics, University of Groningen) for helpful technical assistance in fluorescence microscopy. HC was funded by China Scholarship Council.

REFERENCES

1. Kaern M, Blake WJ, Collins JJ. The engineering of gene regulatory networks. *Annu Rev Biomed Eng.* 2003;5:179–206.
2. Kearns DB, Losick R. Cell population heterogeneity during growth of *Bacillus subtilis*. *Genes Dev.* 2005;19:3083–3094.
3. Veening JW, Hamoen LW, Kuipers OP. Phosphatases modulate the bistable sporulation gene expression pattern in *Bacillus subtilis*. *Mol Microbiol.* 2005;56:1481–1494.
4. Longo D, Hasty J. Dynamics of single-cell gene expression. *Mol Syst Biol.* 2006;2:64.
5. Veening JW, *et al.* Single cell analysis of gene expression patterns of competence development and initiation of sporulation in *Bacillus subtilis* grown on chemically defined media. *J Appl Microbiol.* 2006;101:531–541.
6. Mars RA, *et al.* Small regulatory RNA-induced growth rate heterogeneity of *Bacillus subtilis*. *PLoS Genet.* 2015;11:e1005046.
7. Higgins D, Dworkin J. Recent progress in *Bacillus subtilis* sporulation. *FEMS Microbiol Rev.* 2012;36: 131–148.

8. Gonzalez-Pastor JE, Hobbs EC, Losick R. Cannibalism by sporulating bacteria. *Science.* 2003;301:510–513.
9. Chen I, Dubnau D. DNA uptake during bacterial transformation. *Nat Rev Microbiol.* 2004;2:241–249.
10. Shank EA, Kolter R. Extracellular signaling and multicellularity in *Bacillus subtilis*. *Curr Opin Microbiol.* 2011;14:741–747.
11. Vlamakis H, *et al.* Sticking together: building a biofilm the *Bacillus subtilis* way. *Nat Rev Microbiol.* 2013;11:157–168.
12. Smits WK, *et al.* Stripping *Bacillus*: ComK auto-stimulation is responsible for the bistable response in competence development. *Mol Microbiol.* 2005;56:604–614.
13. Veening JW, *et al.* Transient heterogeneity in extracellular protease production by *Bacillus subtilis*. *Mol Syst Biol.* 2008;4:184.
14. de Jong IG, Veening JW, Kuipers OP. Heterochronic phosphorelay gene expression as a source of heterogeneity in *Bacillus subtilis* spore formation. *J Bacteriol.* 2010;192:2053–2067.
15. de Jong IG, Veening JW, Kuipers OP. Single cell analysis of gene expression patterns during carbon starvation in *Bacillus subtilis* reveals large phenotypic variation. *Environ Microbiol.* 2012;14:3110–3121.
16. Süel GM, *et al.* Tunability and noise dependence in differentiation dynamics. *Science.* 2007;315:1716–1719.
17. Swain PS, Elowitz MB, Siggia ED. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci U S A.* 2002;99:12795–12800.
18. MacGregor GR, *et al.* Use of *Escherichia coli* (*E. coli*) *lacZ* (β-Galactosidase) as a reporter gene. In *Gene transfer and expression protocols.* 1991:217–235.
19. Kirchner G, *et al.* Active bacterial luciferase from a fused gene: expression of a *Vibrio harveyi luxAB* translational fusion in bacteria, yeast and plant cells. *Gene.* 1989;81:349–354.
20. Shaner NC, *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma sp.* red fluorescent protein. *Nat Biotechnol.* 2004;22:1567–1572.
21. Chalfie M, *et al.* Green fluorescent protein as a marker for gene expression. *Science.* 1994:802–805.

22. Shaner NC, Steinbach PA, Tsien RY. A guide to choosing fluorescent proteins. *Nat Methods.* 2005;2: 905–909.
23. Haseloff J, Siemering KR. The uses of green fluorescent protein in plants. *Green Fluorescent Protein: Properties, Applications, and Protocols.* 1998:191–220.
24. Southward CM, Surette MG. The dynamic microbe: green fluorescent protein brings bacteria to light. *Mol Microbiol.* 2002;45:1191–1196.
25. Veening JW, *et al.* Visualization of differential gene expression by improved cyan fluorescent protein and yellow fluorescent protein production in *Bacillus subtilis*. *Appl Environ Microbiol.* 2004;70:6809–6815.
26. Garcia-Betancur JC, *et al.* Single-cell analysis of *Bacillus subtilis* biofilms using fluorescence microscopy and flow cytometry. *J Vis Exp.* 2012.
27. Trip EN, *et al.* Balanced transcription of cell division genes in *Bacillus subtilis* as revealed by single cell analysis. *Environ Microbiol.* 2013;15:3196–3209.
28. Marston AL, Errington J. Dynamic movement of the ParA-like Soj protein of *B. subtilis* and its dual role in nucleoid organization and developmental regulation. *Mol Cell.* 1999;4:673–682.
29. Anderson DE, Gueiros-Filho FJ, Erickson HP. Assembly dynamics of FtsZ rings in *Bacillus subtilis* and *Escherichia coli* and effects of FtsZ-regulating proteins. *J Bacteriol.* 2004;186:5775–5781.
30. Johnson AS, van Horck S, Lewis PJ. Dynamic localization of membrane proteins in *Bacillus subtilis*. *Microbiology.* 2004;150:2815–2824.
31. Doubrovinski K, Howard M. Stochastic model for Soj relocation dynamics in *Bacillus subtilis*. *Proc Natl Acad Sci U S A.* 2005;102:9808–9813.
32. Overkamp W, *et al.* Benchmarking various green fluorescent protein variants in *Bacillus subtilis*, *Streptococcus pneumoniae*, and *Lactococcus lactis* for live cell imaging. *Appl Environ Microbiol.* 2013; 79:6481–6490.
33. Mileyko Y, Joh RI, Weitz JS. Small-scale copy number variation and large-scale changes in gene expression. *Proc Natl Acad Sci U S A.* 2008;105:16659–16664.
34. Munch KM, *et al.* Polar fixation of plasmids during recombinant protein production in *Bacillus megaterium* results in population heterogeneity. *Appl Environ Microbiol.* 2015;81:5976–5986.

35. Ozbudak EM, *et al.* Regulation of noise in the expression of a single gene. *Nat Genet.* 2002;31:69–73.
36. Thattai M, van Oudenaarden A: Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci U S A.* 2001;98:8614–8619.
37. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol.* 2014;5.
38. Glick BR. Metabolic load and heterologous gene expression. *Biotechnol Adv.* 1995;13:247–261.
39. Wu G, *et al.* Metabolic Burden: cornerstones in synthetic biology and metabolic engineering applications. *Trends Biotechnol.* 2016;34:652–664.
40. Zou W, Edros R, Al-Rubeai M. The relationship of metabolic burden to productivity levels in CHO cell lines. *Biotechnol Appl Biochem.* 2017.
41. Nijland R, Kuipers OP. Optimization of protein secretion by *Bacillus subtilis*. *Recent Pat Biotechnol.* 2008;2:79–87.
42. Kunst F, *et al.* The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature.* 1997;390.
43. Casadaban MJ, Cohen SN. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J Mol Biol.* 1980;138:179–207.
44. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.
45. Konkol MA, Blair KM, Kearns DB. Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J Bacteriol.* 2013;195:4085–4093.
46. van den Esker MH, Kovacs AT, Kuipers OP: YsbA and LytST are essential for pyruvate utilization in *Bacillus subtilis*. *Environ Microbiol.* 2017;19:83–94.
47. Eliceiri KW, *et al.* Biological imaging software tools. *Nat Methods.* 2012;9:697–710.
48. Schneider CA, Rasband WS, Eliceiri KW: NIH Image to ImageJ: 25 years of image analysis. *Nat Methods.* 2012;9:671–675.

CHAPTER 6

General Discussion

In today's modern society, aided by the fast development of recombinant DNA technology, large amounts of biotech-based products that are generated by microbial cell factories provide important substances to the traditional food, pharmaceutical and chemistry industries. For instance, enzymes, including proteases, amylases, and lipases, serve as efficient product additives in detergents and as catalysts in biofuel industry [1]. In Europe alone, the proteases, which are commonly used in the detergent industry, account for 900 tons of pure enzymes per year [1]. The global market for industrial proteins is growing rapidly, and therefore, the development of highly efficient production systems is in high demand. *Bacillus subtilis* is widely applied as a microbial cell factory, due to the fact that achieving large-scale production at high cell densities is relatively straightforward and inexpensive [2]. Moreover, its high genetic accessibility and amenability provide excellent possibilities for further modification by molecular genetic techniques [3]. Most importantly, this microbial host that is generally recognized as safe (GRAS), can naturally secrete high amounts of products (up to 20–25 gram per liter) directly into fermentation media [4]. Thus, this production host is highly favored over other famous production organisms such as *Escherichia coli* in view of the relatively simple downstream purification processing. Therefore, *B. subtilis* and its close relatives that can deliver higher yields of industrial enzymes at lower costs, have become of substantial economic importance.

B. subtilis, has been subjected to extensive exploitation for protein production during approximately three decades, which was initiated by the overexpression of proteins derived from *Bacillus* species. The high-level production of more complex proteins or enzymes that originate from Gram-negative bacteria or humans was severely hampered [2]. Numerous attempts, e.g. strong promoters and RBSs, gene disruption (knockout,

mutagenize), regional optimization of specific pathways, have been tried to achieve the overproduction of heterologous proteins in this cell factory [5]. However, these strain modification approaches have already more or less touched the ceiling, and commonly, one specific protein improvement method cannot ensure high-level production of other proteins [6]. To increase the production capacity and also broaden the range of proteins that can be overexpressed in *B. subtilis*, more knowledge about the cellular process and the development of better production systems is still highly needed. Throughout this whole thesis, we further explored the production potential of *B. subtilis* for both secretory and intracellular proteins as well as investigated the underlying reasons, for which both traditional approaches and novel genetic engineering and analytical tools were used.

In **Chapter 2**, we investigated the *B. subtilis* cell surface composition regarding its protein secretion capacity of recombinant α-amylase variants with either low-, neutral- or high- isoelectric points (pI). The protein secretion mechanism is increasingly clear. As shown in **Fig. 1**, pre-protein needs to successively go through the cell membrane lipid bilayer and cell wall, and then be excreted into the media by Sec-type secretion; another system for secretion, but specifically for folded proteins is the Tat system (not further shown and discussed here). During the Sec-secretion process, the secreted target is modified to be a mature protein by interacting with cell surface components and embedded quality-control proteases. We demonstrate that the absence of the cell membrane phospholipid biosynthesis-related enzymes PssA and/or ClsA are beneficial for various α-amylases secretion yields. Although a full understanding of the interaction between the secreted protein and the membrane lipid bilayer is still incomplete, the improved secretion ability shows high correlation with the presence of overall anionic phospholipids (phosphatidylglycerol and cardiolipin) in the engineered

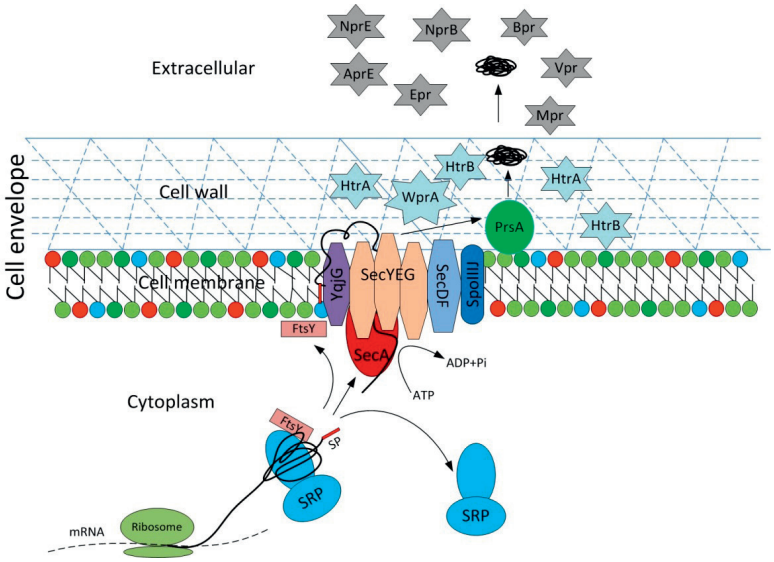


Fig. 1. Schematic representation of the Sec-dependent secretion pathway in *B. subtilis*.

expression hosts. In this study, a combinatorial strain improvement strategy, which consists of codon optimization, tunable expression system, specific modification of the secretion machinery (cell envelope engineering in this case), was employed for improving the secretion efficiency of protein candidates. Here, we specifically studied two variables, that is, cell envelope components and the pIs of secreted proteins. A variety of other relevant factors that also play vital roles in the secretion of end-products, such as signal peptides, translocases, were out of the scope of this study. So, it is foreseeable that, on the basis of the previously engineered strains, huge enhancement of the product yields might be achieved by integrating more efficient modifications of secretion-limiting factors.

In the past twenty years, most of the metabolic optimization methods have focused on the engineering of regional pathways or specific steps, while a novel strategy directed on transcription

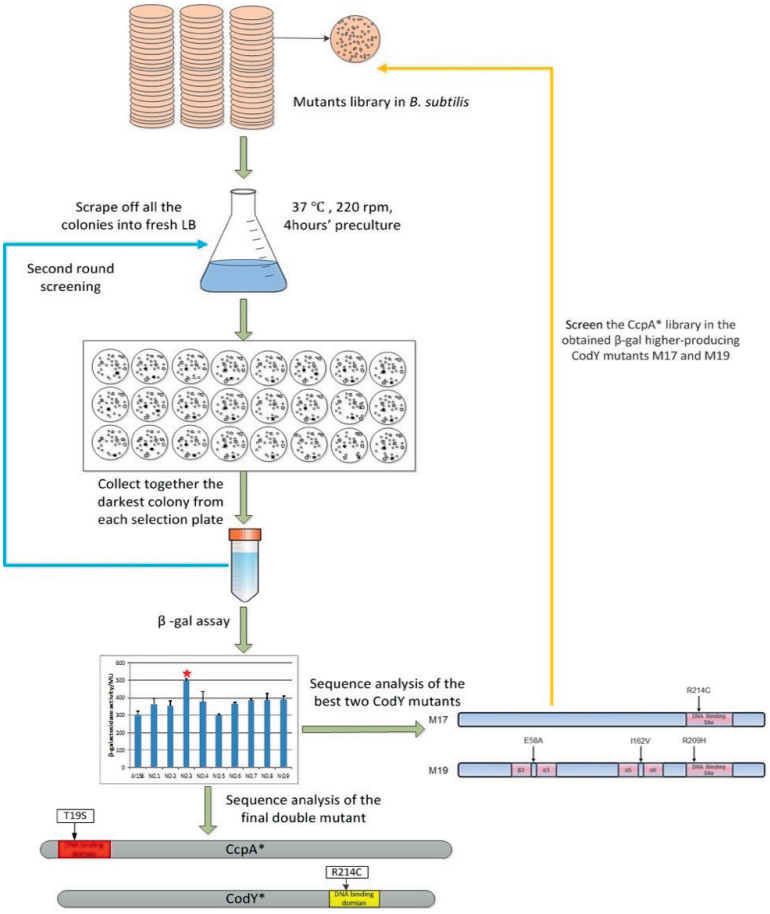


Fig. 2. Workflow of the black-white selection for β-galactosidase higher-producing phenotypes. Desired phenotypes (dark colonies, the substrate S-gal can be degraded by β-galactosidase into black products) were selected. Next, all the collected colonies were subjected to a second screening round, and the finally acquired mutants were analyzed by β-galactosidase assays and DNA sequencing.

regulation that can achieve multiple and simultaneous modifications, has rarely been applied. In **Chapter 3**, the global transcription machinery engineering (gTME) approach was utilized to unlock phenotypes in overexpressing a target protein by

randomly mutagenizing the pleiotropic transcriptional regulators CodY and CcpA. In combination with black-white selection, we efficiently isolated variants with enhanced production of the reporter protein β-galactosidase (**Fig. 2**). The best mutant contained two amino acid substitutions within the DNA-binding HTH domains, $CodY^{R214C}$ and $CcpA^{T19S}$, and increased the production level of β-galactosidase up to 290% relative to the wild-type control. This well-designed toolkit that expands the scale of pathway modification to a global level, can remarkably and straightforwardly lead to strain enhancement, even without the complete understanding of the metabolic networks [7, 8]. Besides the initially used β-galactosidase, some other recombinant proteins, including GFP, xylanase, and peptidase, were also significantly higher-produced in the selected cell factory $CodY^{R214C}CcpA^{T19S}$. Since the overexpression of different proteins probably has varying utilization biases for available intracellular nutrient sources, the effect of the strain improvement differs per protein used. Notwithstanding, our best engineered microbial host obtained from gTME libraries still obviously broadens the application in overproducing a wide range of proteins. Although the direct deletion of CodY and/or CcpA can already improve the production of β-galactosidase by 10–30%, we still utilized the relatively complex gTME-based approach, not only for overexpressing the target protein even more, but also to further study the interesting mutants, uncovering hidden information behind the rewired metabolic networks.

In **Chapter 4,** we analyzed the transcriptome perturbations at a global level in the previously obtained expression host $CodY^{R214C}CcpA^{T19S}$ by using RNA-sequencing. Moreover, DNA-protein binding analysis of the two mutated regulatory proteins was performed by gel electrophoretic mobility shift assays (EMSA). As demonstrated in **Fig. 3**, CcpA can indirectly govern the expression of CodY and its regulon, by the mediation

of the intracellular BCAA pool [9, 10]. Thus, the expression patterns of these two regulators were exactly opposite to each other in various mutants. In addition, the amino acid substitutions within the HTH domains of the two transcription factors altered the overall binding specificity to their direct target genes. In effect, the carbon metabolism was further repressed, while the nitrogen metabolism was obviously de-repressed, which in conjunction resulted in a system-wide metabolic shift allowing for enhanced synthesis capacity of the target protein. Moreover, only a small portion of regulated genes exhibited significant responses to the transcriptome perturbations, reflecting that the vast majority of genes are under control of complex, multiple forms of expression regulation, which probably meet the needs of maintaining the cellular processes at a relatively steady state [11]. The carbon core metabolism, which has been well-evolved to guarantee essential energy and building block supply in the cell [12], is less responsive to the transcriptional regulatory alterations. In comparison, there is still space to further adjust the nitrogen metabolic networks for the overproduction of heterologous recombinant proteins that has been demonstrated in **Chapter 3**. Furthermore, solid relevant evidence can be possibly achieved by further analyzing the metabolic flux and implementing glucose or amino acid consumption assays of the obtained mutants. Taken together, we speculate that the reorganization of cellular metabolic fluxes reduces the metabolic burden by allowing a better balance of resource distributions for both essential native metabolic networks and heterologous protein overproduction pathways.

Under nutrient-limiting growth conditions, *B. subtilis* cells activate a variety of regulatory processes to determine diverse cell fates, such as sporulation, competence development, biofilm formation, and differentiate into a community of multiple subpopulations. This phenomenon is known as phenotypic

heterogeneity [13]. In the study of heterologous protein production in cell factories, major research efforts have focused on the product yield at the population level, while the production activity of individual cells has largely been ignored. A recent study demonstrated that the production of amylases in *B. subtilis* is performed in a non-uniform manner, and a *degU* mutation and optimized growth conditions can significantly improve the overall secretion yields by suppressing the heterogeneity in expression [4]. In **Chapter 3**, we presented that the well-known reporter protein GFP, which is mostly used as a benchmark for visualizing protein localization and promoter activity, can be higher-expressed in the best β-gal producer $CodY^{R214C}CcpA^{T19S}$. In **Chapter 5**, we demonstrate that the fluorescence signals of GFP are dynamic over time and differ in strains with various global regulator mutation backgrounds. Strains harboring $CodY^{R214C}$ showed a high similarity in growth curve and GFP expression performance at the population, subpopulation and single-cell levels. In comparison to the expression hosts carrying wild-type versions of CodY, the $CodY^{R214C}$-containing strains, although showing slightly lower growth rates, had overall higher GFP product yields. We initially regarded this difference as a consequence of the relatively different growth rate, but the following analyses at the single-cell and subpopulation levels showed high consistency of the overall protein yield with the homogeneity of high-expressing populations. In other words, the expression heterogeneity, which can be regarded as phenotypic noise, comes from global regulation, suggesting a negative correlation with overall GFP production in the cell factories we assessed, especially during the late stationary growth phase. We reason that the strains with $CodY^{R214C}$ have reprogrammed metabolic networks that focus more on the target protein synthesis, while in the wild-type with an imbalanced metabolic flux distribution, only a subset of cells can get sufficient nutrient resources for

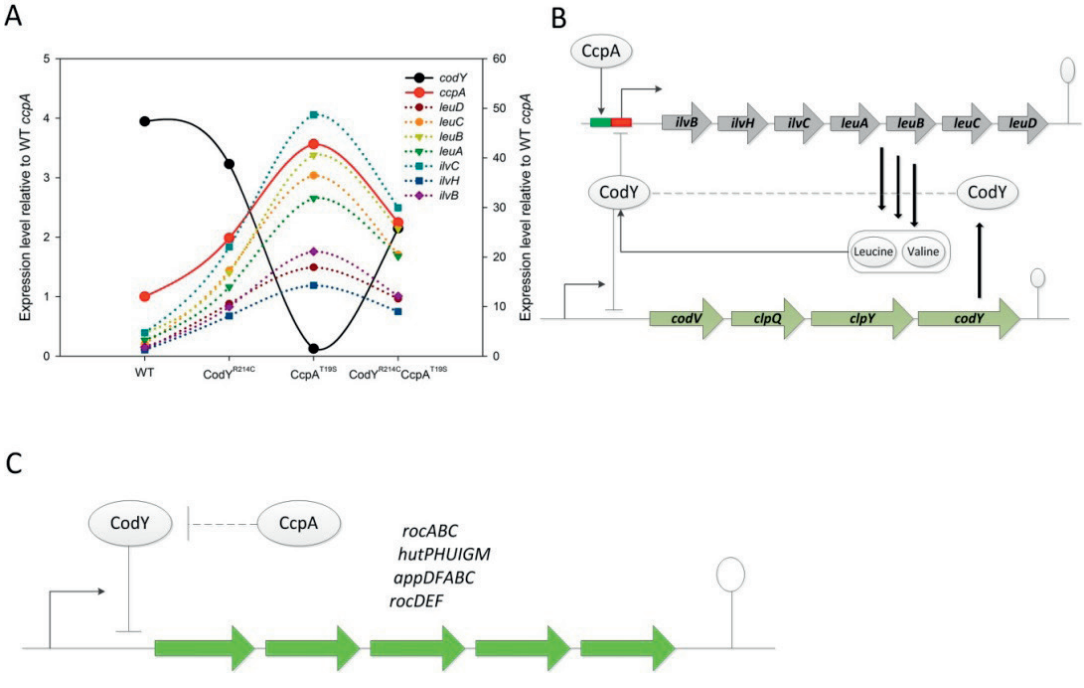


Fig. 3 (A) The relative expression levels of CodY, CcpA and the *ilv-leu* operon in the HTH domain mutant strains. **(B)** Schematic diagram of the interaction between CcpA and CodY mediated by the biosynthesis of BCAAs. **(C)** Schematic representation of CcpA-mediated transcriptional regulation of operons that are repressed by CodY. Arrows and perpendiculars represent the positive and negative actions, respectively.

GFP expression pathway. This can be considered as a population-scale survival strategy for the wild-type strain under suboptimal nutrient conditions by ensuring a small proportion of high-producing cells and another low-producing subpopulation as cost. Although the detailed mechanism underlying the GFP expression heterogeneity is still incomplete, this study offers a new perspective into the overexpression of recombinant proteins and paves the way to further increase the use of *B. subtilis* as a cell factory.

OUTLOOK

In natural environments, wild-type bacterial cells have evolved sophisticated adaptation systems that allow them to take advantage of many kinds of nutrient sources for optimal fitness during changing nutritional conditions [14]. With their intricate regulatory systems, the bacteria know where they are by sensing the availability level of metabolites, and then stimulate the global metabolic regulation and motility mechanisms for shifting to the place they should be as a response [14]. However, the human-imposed task for overexpressing proteins of industrial and commercial interest, especially the heterologous ones, has not been evolved over time spent in nature. Therefore, modifying the global regulatory networks in the existing cell factories like *B. subtilis* to obtain a better expression host appears to be a good solution. We are now able to unlock a comprehensive profile inside the cells by use of 'omics' technologies, and numerous systematic functional studies have considerably enhanced our understanding of the complex metabolic and regulatory pathways in *B. subtilis* [15–18]. Moreover, the newly developed systems and synthetic biology devices, such as CRISPR-Cas9 and gTME, provide an unprecedented level of engineering

possibilities that potentially facilitate strain improvements [19, 20]. Undoubtedly, we are still at the early stages of developing *B. subtilis* as a highly adaptable chassis with both higher yields and a wider range of products. Nevertheless, novel insights into the complicated cellular and global regulatory processes and the great advances in technological tools will lead to the construction of super-producing cell factories.

REFERENCES

1. van Dijl J, Hecker M. *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. *Microb Cell Fact.* 2013;12:p.1.
2. Westers L, Westers H, Quax WJ. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochim Biophys Acta.* 2004;1694:299–310.
3. Harwood CR, *et al. Bacillus subtilis*: model Gram-positive synthetic biology chassis. In *Methods in microbiology.* Academic Press. 2013;40:87–117.
4. Ploss TN, *et al*. Homogeneity and heterogeneity in amylase production by *Bacillus subtilis* under different growth conditions. *Microb Cell Fact.* 2016;15:57.
5. Liu Y, *et al*. Metabolic engineering of *Bacillus subtilis* fueled by systems biology: Recent advances and future directions. *Biotechnol Adv.* 2017;35:20–30.
6. Nijland R, Kuipers OP. Optimization of protein secretion by *Bacillus subtilis*. *Recent Pat Biotechnol.* 2008;2:79–87.
7. Lanza AM, Alper HS. Using transcription machinery engineering to elicit complex cellular phenotypes. In *Synthetic Gene Networks*. Humana Press. 2012;229–248.
8. Woolston BM, Edgar S, Stephanopoulos G. Metabolic engineering: past and future. *Annu Rev Chem Biomol Eng.* 2013;4:259–288.
9. Shivers RP, Sonenshein AL. *Bacillus subtilis ilvB* operon: an intersection of global regulons. *Mol Microbiol.* 2005;56:1549–1559.

10. Tojo S, *et al*. Elaborate transcription regulation of the *Bacillus subtilis ilv-leu* operon involved in the biosynthesis of branched-chain amino acids through global regulators of CcpA, CodY and TnrA. *Mol Microbiol.* 2005;56:1560–1573.
11. Kochanowski K, Sauer U, Chubukov V. Somewhat in control--the role of transcription in regulating microbial metabolic fluxes. *Curr Opin Biotechnol.* 2013;24:987–993.
12. Chubukov V, *et al*. Coordination of microbial metabolism. *Nat Rev Microbiol.* 2014;12:327–340.
13. Lopez D, Vlamakis H, Kolter R. Generation of multiple cell types in *Bacillus subtilis*. *FEMS Microbiol Rev.* 2009;33:152–163.
14. Sonenshein AL. Control of key metabolic intersections in *Bacillus subtilis*. *Nat Rev Microbiol.* 2007;5: 917–927.
15. Pisithkul T, Patel NM, Amador-Noguez D. Post-translational modifications as key regulators of bacterial metabolic fluxes. *Curr Opin Microbiol.* 2015;24:29–37.
16. Fujita Y. Carbon catabolite control of the metabolic network in *Bacillus subtilis*. *Biosci Biotechnol Biochem.* 2009;73:245–259.
17. Büscher JM, *et al*. Global network reorganization during dynamic adaptations of *Bacillus subtilis* metabolism. *Science.* 2012;335.
18. Brinsmade SR, *et al*. Hierarchical expression of genes controlled by the *Bacillus subtilis* global regulatory protein CodY. *Proc Natl Acad Sci U S A.* 2014;111:8227–8232.
19. Alper H, Stephanopoulos G. Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab Eng.* 2007;9:258–267.
20. Jakociunas T, Jensen MK, Keasling JD. CRISPR/Cas9 advances engineering of microbial cell factories. *Metab Eng.* 2016;34:44–59.

Nederlandse samenvatting

In natuurlijke omgevingen hebben wild-type bacteriën verfijnde systemen ontwikkeld om optimaal gebruik te kunnen maken van verschillende nutriënten om zo een optimale fitness te hebben onder veranderende natuurlijke omstandigheden. Door middel van hun regulatie systeem, en verfijnde sensoren zijn ze in staat hun omgeving, en aanwezigheid van nutriënten waar te nemen. Hierop kan gereageerd worden door de metabole regulatie aan te passen, of zich te verplaatsen naar daar waar de nutriënten zijn.

Daarentegen is de door de mens gewilde over expressie van genen voor industriële en commerciële doeleinden, vooral de heterologe genen, niet geëvolueerd onder natuurlijke omstandigheden. Daarom is het aanpassen van de globale regulatie netwerken in bestaande cell-factories zoals *Bacillus subtilis* om zodoende een betere expressie host te krijgen een logische oplossing. In deze scriptie is het potentieel van *B. subtilis* om zowel uitgescheiden als intracellulaire eiwitten te produceren, en de onderliggende mechanismen hiervan onderzocht. Hiervoor zijn zowel nieuwe als en traditionele genetische technieken gebruikt.

In **Hoofdstuk 2** is onderzoek gedaan naar de samenstelling van het cel oppervlak met betrekking tot eiwit secretie van recombinante α-amylase varianten met lage, neutrale en hoge iso-electrische punten (pI). We hebben aangetoond dat de afwezigheid van celmembraan fosfolipide gerelateerde enzymen PssA en/of ClsA een verhoogde opbrengst van verschillende uitgescheiden α-amylases veroorzaken. Ondanks dat we nog niet volledig begrijpen hoe de interactie tussen uitgescheiden eiwitten en de lipid-bilayer werkt, tonen de verhoogde secretie levels aan dat dat er een sterke correlatie is met de aanwezigheid van anionische fosfolipiden (phosphatidylglycerol en cardiolipin) in de expressie stammen. In deze studie is gebruik gemaakt van een stam waarin codon optimalisatie, instelbare expressie systemen en specifieke modificaties van de expressie systemen

(cell-enveloppe engineering in dit geval), om zodoende de secretie van eiwitten te optimaliseren.

In **Hoofdstuk 3** is gebruik gemaakt van global transcription machinery engineering (gTME) om fenotypes te verkrijgen die voor over expressie van eiwitten te krijgen door random mutagenese van de pleiotropische regulatoren CodY en CcpA. Door gebruik te maken van zwart-wit selectie hebben we op een efficiënte manier mutanten kunnen selecteren met een overproductie van β-galactosidase. De beste mutant heeft twee aminozuur substituties in het DNA-binding HTH domein, $CodY^{R214C}$ en $CcpA^{T19S}$. Deze mutaties verhoogde de β-galactosidase productie met 290% in vergelijking met het wild-type. Met deze wel doordachte toolkit waarmee pathway-modificatie wordt uitgebreid naar een globaal niveau, kan opmerkelijk snel en direct leiden tot stam verbetering, zelfs zonder het complete metabole netwerken te begrijpen. Behalve de gebruikte β-galactosidase zijn er ook enkele andere recombinante eiwitten, waaronder GFP, xylanase en peptidase die aanzienlijk meer geproduceerd werden in de gebruikte cell-factory $CodY^{R214C}CcpA^{T19S}$. Doordat de over expressie van verschillende eiwitten voor variatie zorgt in het gebruik van de beschikbare intracellulaire nutriënten, zal de manier waarop een stam verbeterd kan worden afhankelijk zijn van de gebruikte eiwitten.

In **Hoofdstuk 4** is de verstoring van het transcriptoom geanalyseerd van de expressie host $CodY^{R214C}CcpA^{T19S}$ door middel van RNA sequencing. Daarnaast is DNA-proteïne binding analyse van de twee gemuteerde regulatie eiwitten getest door middel van gel electorphoretic mobility shift assays (ESMA). De aminozuur veranderingen in de HTH domeinen van de twee transcrioption factors veranderde de gehele bindings-specificiteit voor hun directe doelwitten. Een ander effect was dat het koolstof metabolisme nog verder onderdrukt werd, in tegenstelling tot het stikstof metabolisme dit werd duidelijk versterkt.

Dit heeft er samen voor gezorgd dat er in het hele systeem een metabolische shift was waardoor bepaalde eiwitten in hogere mate geproduceerd konden worden. Daarnaast was er maar een klein aantal van de gereguleerde eiwitten die een significante respons hadden op de verstoring van het transcriptoom. Dit is in samenspraak met de theorie dat de meeste genen onder controle staan van meerdere vormen van regulatie, dit resulteert waarschijnlijk in het behoud van een constante niveau van de cellulaire processen. Het centrale koolstof metabolisme, dat op een dusdanige manier is geëvolueerd dat het gegarandeerd essentiële bouwstenen en energie kan leveren aan de cel, reageert minder op de verandering van in transcriptoom regulatie. Ter vergelijking, is er nog steeds ruimte voor aanpassingen aan het stikstof metabolisme voor de overproductie van heterologe recombinante eiwitten zoals gedemonstreerd in **Hoofdstuk 3**.

In **Hoofdstuk 5** hebben we laten zien dat fluorescentie signaal van GFP dynamisch zijn door de tijd, en dat dit verschilde in stammen met verschillende mutaties in de globale regulatoren. Stammen met de mutatie $CodY^{R214C}$ hadden veel overeenkomst in groei en GFP expressie in zowel de hele populatie als in single-cell metingen. In vergelijking met de expressie stammen met wild-type varianten van CodY, hebben de stammen met de $CodY^{R214C}$ mutatie hogere GFP opbrengst, ondanks een iets lagere groeisnelheid. We dachten eerst dat dit werd veroorzaakt door de relatief lage groeisnelheid, maar analyses op single-cell niveau en op subpopulaties hebben hoge mate van consistentie laten zien wat betreft eiwit opbrengst met homogeniteit van populaties met hoge expressie. In andere woorden, de heterogeniteit van expressie, die kan worden gezien als fenotypische ruis, wordt veroorzaakt door globale regulatie, dit suggereert dat er een negatieve correlatie is met de algehele GFP productie in de cellen waar wij naar gekeken hebben, vooral tijdens

de laat stationaire fase. Ondanks dat de exacte mechanismes van heterogeniteit in GFP expressie nog niet compleet zijn geeft deze studie een nieuwe kijk op het tot over expressie brengen van recombinante eiwitten, en draagt het bij aan het beter gebruiken van *B. subtilis* als werkpaard voor de industrie.

We zijn nu in staat om de hele cel begrijpelijk te maken door het gebruik van 'omics' technologieën, en vele systematische studies hebben onze kennis van de complexe metabole en regulatoire netwerken van *B. Subtilis* verbeterd. Daarnaast brengen nieuwe synthetische biologie technieken zoals CRISPR-Cas9 en gTME ongekende mogelijkheden met zich mee op het gebied van genetische engineering, waarmee in potentie stammen nog sneller en beter veranderd kunnen worden. Ongetwijfeld staan we nog maar aan het begin van het gebruik van *B. subtilis* als zeer aanpasbaar chassis voor zowel betere, en gevarieerdere producten. Desondanks zullen nieuwe inzichten in de complexe structuur van regulatoire netwerken, en de snelle ontwikkelingen van technieken er voor zorgen dat we super-producerende cellen kunnen maken in de toekomst.

中文总结

在自然环境中，野生型的细菌进化出了精密而又复杂的调控系统，使其能够通过利用各种各样的营养资源，更好地适应不断变化的生长条件。凭借这套错综复杂的调控体系，细菌可以通过感知体内关键代谢物的水平，进而激发相应的代谢调节以及运动机制，趋利避害。然而，细菌对于人为强加的任务—高水平表达具有工业和商业价值的蛋白质，尤其是异源蛋白，其相对应的生产能力并没有在自然进化中得到加强。因此，改良现有的优秀细胞工厂如枯草芽孢杆菌成为更好的表达宿主，以用来高水平表达各种异源重组蛋白。贯穿整个论文，我们使用传统和新型的基因工程方法和分析手段，探索了在枯草芽孢杆菌中进一步提高分泌型和胞内蛋白质的生产能力，并调查了蛋白过表达背后的原因。

在第2章中，我们研究了枯草芽孢杆菌细胞表面（特别是细胞膜）的组成与带有不同等电点（pI）的重组α-淀粉酶变体分泌能力的关系。在本章节中，我们采用了密码子优化，可控表达系统，分泌系统特异性修饰等手段组合而成的菌种改良策略，用于有效提高靶蛋白的分泌效率。细胞膜磷脂双分子层相关酶PssA和/或ClsA的缺失被证明有助于提高α-淀粉酶的分泌量。尽管我们还不能完全了解在蛋白分泌过程中分泌对象和膜脂质双分子层之间相互作用，但本研究表明，表达宿主的分泌能力与其细胞膜中总阴离子磷脂（磷脂酰甘油和心磷脂）含量显著正相关。

在第3章中，我们采用高通量的gTME方法先后对多效转录调控因子CodY和CcpA进行随机诱变，以获取目标蛋白高表达的表型。然后，通过黑白斑筛选的方法，快速有效地分离出靶蛋白—β-半乳糖苷酶的产量大幅增强的突变体。在筛选出的最佳表型$CodY^{R214C}CcpA^{T19S}$中，调控蛋白CodY和CcpA上负责与DNA相结合的HTH结构域内分别有一个氨基酸被替换，相对于野生型对照，其β-半乳糖苷酶的表达水平提高至290%。在此研究中，我们精心设计的异源蛋白表达的增强策略，能将代谢通路优化的范围扩大至全细胞水平，即使没有对代谢网络的完整掌握，也可以直接有效地实现菌种改良。除β-半乳糖苷酶之外，其他重组蛋白（包括绿色荧光蛋白，木聚糖酶和肽酶）也能够在$CodY^{R214C}CcpA^{T19S}$中显著地得到更好地表达。由于不同蛋白质的过表达途径对胞内营养源具有利用差异，因此不同的蛋白质在我们获得改良宿主中的表达增强程度各异。

在第4章中，我们对在第3章中获得的突变菌株CodY^{R214C}，CcpA^{T19S}，CodY^{R214C}CcpA^{T19S}进行了转录组分析（RNA-Seq）。此外，通过凝胶电泳迁移率变动分析（EMSA）分别对两个调控蛋白突变体进行DNA-蛋白结合亲和力的检测。实验表明，两个多效转录调控因子的HTH结构域内的氨基酸突变改变了其对直接作用基因的整体结合特异性，进而重组了整个相对应的代谢网络的活动。具体是，碳代谢通路被进一步抑制，而氮代谢的抑制作用则有明显的减弱，这两个通路变化协调作用导致在全系统水平上的代谢迁移，从而增强了靶蛋白的合成能力。此外，只有一小部分（约1/10）被调控基因表现出对突变CodY和CcpA的显著响应，很可能是由于绝大多数目标基因的表达受制于复杂的，多种形式的调控，以确保细胞内的活动维持在相对稳定的状态。同时，细胞内必需的能量和元素的供应者—中心碳代谢网络，已经在长期的自然进化过程中得到充分优化，其对于转录扰动的响应较弱。相比之下，在第3章中已经证明，氮代谢网络仍然可以被进一步调整以实现异源重组蛋白高水平合成。

在第5章中，我们证明GFP在同一宿主细胞中的合成能力是随时间动态变化的，并且其在具有不同多效调控蛋白突变背景的菌株中荧光强度也有很大不同。在群体水平上，携带CodY^{R214C}的菌株在生长速度和GFP表达方面显示出高度的相似性。与野生型CodY的表达宿主相比，CodY^{R214C}菌株尽管生长速度略慢，但具有更高的GFP总产量。我们最初认为这是生长速度不同造成的结果，但在单细胞和亚群体水平的检测推翻了这一推论，其证明总蛋白质产量与高表达群体的同质性相关。尤其是在稳定生长末期阶段，源自于多效调控蛋白的表达异质性，也被认为是在这个案例中的表型噪声，与我们评估的表达宿主中的总GFP产量呈负相关。尽管我们仍未全面掌握GFP表达异质性的详细机制，本章节仍然为重组蛋白的过表达研究展现了全新的视角，并为进一步增强枯草芽孢杆菌作为细胞工厂的应用奠定基础。

我们如今能够通过组学分析的手段来获得细胞内部重要信息，而许多系统的相关研究也大大增强了我们对枯草芽孢杆菌复杂的代谢和调控网络的了解。同时，新开发的系统和合成生物学技术（如CRISPR-Cas9和gTME）提供了前所未有的菌株改良手段。毫无疑问，我们目前仍处于开发枯草芽孢杆菌成为更高水平表达更多种类的产物的高适应性菌株的早期阶段。尽管如此，我们对复杂的胞内代谢调控网络的更加详尽的掌握以及技术手段的巨大进步将有助于超级细胞工厂的成功构建。

Acknowledgments

In September 2013, I left China and came to Groningen, the Netherlands for my PhD study in Molecular Genetics group. From then on, I started this fun-filled exploration and the journey of self-discovery, during which, happiness and pain, ups and downs are experienced. How time flies, it has been 4 years and 9 months, and the thesis finally comes out as a reward of hard work. Now, it is the time to look back and give my sincere appreciation to all the kind people who make my way here.

Firstly, and most importantly, I would like to express my deepest gratitude to my dear promoter, Prof. Oscar P. Kuipers. Thank you very much for providing me the opportunity to do my research in this great and warm group with unique international atmosphere, and it is really my great honor to be your PhD student. You are an amazing supervisor, without your generous support and helpful guidance, I cannot make the thesis. Your rigorous and optimistic attitude always encourage and inspire me to do better and go further, and this valuable spiritual wealth will benefit me forever.

My grateful thanks also go to Prof. Jan Kok and Prof. Jan-Willem Veening for all the interesting discussions and suggestions on the Monday morning work meetings, which broaden my knowledge and vision. I really appreciate the nice work of the reading committee members, Prof. J.M. van Dijl, Prof. W.J. Quax, Prof. A.T. Kovacs, who spent valuable time on reading through the manuscript, and the thoughtful comments help to improve the quality of this thesis.

In the past a few years, I did good collaborations with some excellent scientists, they are, Dr. Elrike Frenzel, Dr. Julio Villatoro-Hernandez, Dr. Ruud D. O. Weme, Dr. Auke van heel. Elrike, you are very experienced and organized, thank you very much for helping me out with CcpA project and correcting the relevant manuscript, and I wish you all the best with your new job. Julio, thank you for the prior work on the part of CodY, I

cannot make it without the useful information and the β-gal production improved phenotypes you made. Ruud, my daily supervisor in my first year, sincere gratitude for teaching me the basic experimental skills and the helpful discussions and manuscript revision on the CcpA project. I am also very grateful to Auke for contributing to the amylase project. Moreover, many special thanks go to Lance Keller for reading my manuscripts, Ard Jan for helping with the Dutch summary and the Chapter 5, Amanda and Jingqi for being my paranymphs and all the kind help for my defense. My best wishes to all of you for a bright and prosperous future!

I would like to give sincere acknowledgments to Siger, Anne-Hesseling, Harma for all the kind support and help and Anne-de Jong for the expert technical assistance of transcriptome analysis. I also want to gratefully thank our secretaries Manon, Klazien and Jannet for all the patience and time with documents. All the convenience and help that you offered make my work easier in this highly organized research group.

My most special gratitude goes to Yi and Qian for our valuable friendship during these years. We share our ideas and opinions all the time and are always ready to help each other both in the work and life. I really enjoy all the great time we had and wish you good luck and success with your future career!

During my stay in Molgen, many people come while many others go, but once a molgener, forever a molgener. Here, I want to acknowledge all the former and current colleagues in office 0646, Ruud, Andrius, David, Atze, Joaquin, Ard Jan, Putri, Mark, and I really appreciate the very free and cozy working environment we created. And, I am thankful to the great young scientists who are working or used to work in the Bacillus lab 0651, Luiza, Barbara, Amanda, Mirjam, Mariella, Xin, Jeroen, Auke and many students for keeping the lab clean and tidy and making it a pleasant place to work. In addition, I would

like to express my deep appreciation to all the other previous and present colleagues, Robyn, Katrin, Manolo, Ruben, Morten, Martijn, Maike, Angel, Anne-marie, Ana, Afzal, Anna, Irfan, Taketo, Yoshi, Tonia, Lieke, Dongdong, Liang, Sjoerd, Rieza, Jimmy, Jakob, Jhonatan, Chunxu, Jingjing, Xue, Chenxi, Fangfang, Fleur, Xinghong, Zhenhua, Zhibo, Lu, Meishan, Patricia, Eduardo, Clement, Renske, Dimitra, Arnau, Stefano, Robin, Jelle. Thank you all for your concerns and help, and I hope we can meet somewhere in the future.

I also sincerely appreciate the China Scholarship Council (CSC) for providing me 4-year financial support during my study in the Netherlands.

同时，格村“艺术烧烤私人会所”成员，一姐，老李，伊导，刘胖子，阳阳，我们有缘在这个阴雨连绵的荷兰大农村相聚，一起度过许多美好的烧烤和聚餐时光，由衷感谢你们的温暖陪伴，鼓励和支持。希望你们勇敢追梦，大展宏图，也祝你们未来一切顺利，生活幸福美满！

感谢格村老欢足球队的各位球友们，张涛，袁涛，博群，唐振辰，小明，刘彬，刘凌，黄刚，卓华，永壮，李军，岳军，纪全，金涛，国伟，覃伟，孙强，赵培亮，冯琮，何啸，李牧，庆凯，语真，马超，孙静，董京金，杨程皓，田小波，徐光彩，兴洪，二嘎，老姜，虎哥，飞哥，凯哥，峰哥，向南兄，其宏，家文，杨硕等。我们曾并肩作战，欢声笑语，在周末时光里尽情享受放松；也曾觥筹交错，推杯换盏，让单调孤苦的留学生活充满色彩。如今老友们陆续各奔东西，新鲜血液注入，你大爷已不是你大爷，老欢还是老欢，老欢在传承，欢乐在继续。期待将来有缘再会，也祝老铁们在世界各处继续绽放光芒！

还要感谢在荷兰遇见和结识的其他朋友们，柏玉香，孟祥峰，尹慧芳，吕磊，武敏，武静，曹华堂，程亮慧，王淼，杨浦，王元泽，李晅，刘友涛，葛艳萍，邵金凤，孙宇，王鹏，邢丽娟，张胜利，张学文，张铮，张小明，周祺惠，张辑，李延军，盛志高，高永强，高亚军等人；非常珍惜与你们相见的时光，衷心地祝你们工作开心，生活快乐，前程似锦！

吉林大学的赵寿经老师和梁彦龙老师，非常感谢多年来在生活上的关心和学业上的帮助与支持；在此，祝你们在未来的日子里科研顺利，事事如意！同时，十分感谢一直关心我的其他老师和小伙伴们，马研，徐源，黄术强，王乐，胡城，王扬，王聪，四华，李豪，老马，光明，李幸，海宝，狼哥，松辉，周旭，群柱，营凯，邢鹏，贵军，忠哥，世阳，小璞，绍永，开平等人；祝一切安好，也期待与你们国内相聚！

此外，我要诚挚感谢我的亲人们。感谢父母对我的养育之恩和淳淳教诲，如今我也渐渐成长，也越发能感受你们的辛苦和不易；感谢你们的默默付出，使我能顺利地走到今天；在此，祝你们身体健康，永葆青春，开心每一天！非常感谢我的姐姐和姐夫长期以来对我的关心与支持，以及对父母的细心照顾，使我没有后顾之忧，专心于科研；祝你们事业蒸蒸日上，心想事成，也祝我的小外甥能够健康快乐，茁壮成长！同时，还要真诚感谢所有的亲戚朋友们长久以来的帮助，照顾与理解，祝你们所有人幸福安康，事事顺心！

一路走来，风雨兼程，有得有失，冷暖自知。诗和远方，心之所往；赤心山河犹在，未来依然可期。

04/06/2018