CHARLES UNIVERSITY

FACULTY OF SCIENCE

Study programme: Biology

Branch of study: Microbiology

Mgr. Volha Ramaniuk

Factors interacting with bacterial RNA polymerase and their effect on the regulation of transcription initiation

Faktory interagující s bakteriální RNA polymerázou a jejich vliv na regulaci iniciace transkripce

DOCTORAL THESIS

Supervisor: Mgr. L. Krásný, Ph.D.

Institute of Microbiology, Czech Academy of Sciences

PRAGUE, 2018

Prohlašují, že tuto práci jsem vypracovala samostatně s uvedením všech použitých informačních zdrojů a literatury. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

Praha, Olga Ramaniuk

This Thesis is dedicated to my parents

ACKNOWLEDGEMENTS

I would like to greatly thank my supervisor Libor Krásný for giving me an opportunity to start my scientific career in an outstanding laboratory and for his excellent mentoring.

Also, I am very thankful to my colleagues and friends for their constant support. Above all, I want to thank all the members of the Laboratory of Microbial Genetics and Gene Expression for fantastic working atmosphere, help and valuable discussions, and for non-working activities we had live through together. It was a great time!

Moreover, I am thankful to Jiří Vohradský and his lab members for fruitful collaboration.

Many other people contributed to the completion of this Thesis: teachers, educators, and scientists who raised my interest in the prokaryotic world, molecular biology and genetics, and showed how exciting science can be.

Finally, my huge thanks go to my parents, brother and his family. I am and always will be grateful for all their support, love and care they were giving me during my Ph.D. study, and especially during my first years in the Czech Republic. All my achievements partially belong to them.

ABSTRACT (ENGLISH)

The bacterial cell needs to regulate its gene expression in response to changing environmental conditions. RNA polymerase (RNAP) is the pivotal enzyme of this process and its activity is controlled by a number of auxiliary factors. Here I focus on RNAP-associating factors involved in the regulation of transcription in G^+ bacteria: σ factors, initiating nucleoside triphosphates (iNTPs), HelD, δ, and small RNA Ms1. The main emphasis is on σ factors from *Bacillus subtilis*.

σ factors allow RNAP to specifically recognize promoter DNA. In my first project, I set up *in vitro* transcription systems with purified alternative σ factors, σ^B, σ^D, σ^H, σ^I from *B. subtilis*. Using these systems, I studied the effect of initiating NTP concentration ([iNTP]) on transcription initiation. I showed that promoters of alternative σ factors are often regulated by [iNTP].

In the next project I comprehensively characterized one of the least explored alternative σ factors from *B. subtilis*, σ^I. I identified ~130 genes affected by σ^I, though only 16 of them were directly affected. Moreover, I discovered that σ^I is involved in iron metabolism. Finally, I showed that σ^I binding requires not only the conserved -35 and -10 hexamers but also extended -35 and -10 elements located in the spacer region.

In collaboration with colleagues-bioinformaticians, I studied the gene expression network created for σ^A-regulated genes in *B. subtilis* during spore germination and outgrowth. They predicted new genes to be controlled by σ^A. Using our *in vitro* system I verified the computationally predicted interactions.

Next, I studied δ and HelD, both proteins binding *B. subtilis* RNAP. I showed that δ enhanced transcription with selected σ factors; I demonstrated that HelD had no effect on RNAP affinity for promoter DNA, consistent with findings that HelD affects elongation/termination.

My final contribution was demonstrating that Ms1, a highly abundant sRNA in mycobacteria, has the same transcription start site both in *M. smegmatis* and *M. tuberculosis,* and contributed to the P_{Ms1} promoter characterization.

Together, the results were published in four papers (in two of them I am the first author), advancing our knowledge of transcription regulation in bacteria.

ABSTRAKT (ČESKÝ)

Bakteriální buňka reguluje svou genovou expresi jako odpověď na změny růstových podmínek. RNA polymeráza (RNAP) je stěžejním enzymem pro tento proces, její aktivita je ovlivňována mnoha transkripčními faktory. Ve své práci se zabývám faktory, které se účastní regulace transkripce u grampozitivních baktérií: faktory σ, iniciačními nukleozid trifosfáty (iNTP), HelD, δ a malou RNA Ms1. Hlavní důraz v této práci je kladen na faktory σ z *Bacillus subtilis*.

Faktory σ rozpoznávají specifickou sekvenci DNA promotoru a umožnují navázaní RNAP na tuto sekvenci. Ve svém prvním projektu jsem vytvořila transkripční systém *in vitro* s purifikovanými alternativními σ faktory z *B. subtilis* – σ^B, σ^D, σ^H, σ^I. Pomocí tohoto systému jsem zkoumala efekt změny koncentrace iNTP ([iNTP]) na iniciaci transkripce. Prokázala jsem, že promotory přepisované pomocí alternativních σ faktorů jsou často regulovány [iNTP].

V dalším projektu jsem charakterizovala jeden z nejméně prozkoumaných alternativních σ faktorů z *B. subtilis*, σ^I. Identifikovala jsem ~130 genů ovlivněných σ^I, přitom 16 z nich bylo přepisováno přímo σ^I. Prokázala jsem, že se σ^I účastní metabolizmu železa v buňce, a že se váže nejenom na klasické sekvence -35 a -10, ale také na "prodloužené" -35 a -10 elementy.

Ve spolupráci s bioinformatickou laboratoři jsem se zúčastnila studie regulace genové exprese σ^A faktorem z *B. subtilis* ve fázi sporulace a výrůstu ze spory. Naši spolupracovnicí navrhli nové geny regulované σ^A ve zmíněné fázi, tyto návrhy jsem potvrdila pomocí transkripce *in vitro*.

Dále jsem studovala interakční partnery RNAP z *B. subtilis*, δ a HelD. Ukázala jsem, že protein δ zesiluje transkripci s vybranými σ faktory; a že protein HelD nemá vliv na afinitu RNAP vůči promotorové DNA. Poslední pozorování odpovídá skutečnosti, že HelD ovlivňuje spíše elongaci a terminaci transkripce.

V posledním projektu jsem prokázala, že Ms1 malá RNA hojně přepisovaná u mykobaktérii, má stejný transkripční počátek v *M. smegmatis* a *M. tuberculosis*, a přispěla jsem k charakterizaci promotoru pro tuto RNA.

Výsledky získané ve výše zmíněných projektech rozšiřují znalosti o regulaci bakteriální transkripce. Tyto výsledky jsou součásti čtyř publikací (na dvou z nich jsem první autorkou).

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

asRNA – antisense RNA

ATP – adenosine triphosphate

CPK models – molecular models designed by chemists R. Corey, L. Pauling, W. Koltun

α-CTD – the alpha-C-terminal domain of RNA polymerase

CTD – C-terminal domain

DEGs – differentially expressed genes

E – RNA polymerase core

Eσ – RNA polymerase holoenzyme

IPTG – Isopropyl β-D-1-thiogalactopyranoside

iNTP – initiating nucleoside triphosphate

[iNTP] – concentration of iNTP

G+/G- bacteria – Gram-positive/Gram-negative bacteria

nt - nucleotides

NTD – N-terminal domain

NCR – non-conserved region of σ

NMR – Nuclear Magnetic Resonance Spectroscopy

RBS – Ribosome Binding Site

RNAP – RNA polymerase

RP_o – open promoter complex

RP_c – closed promoter complex

RNA-seq – RNA sequencing

RT – room temperature

PAGE – Polyacrylamide Gel Electrophoresis

PCR – polymerase chain reaction

sRNA – small RNA

Taq – *Thermus aquaticus*

TEC – transcription elongation complex

TSS – transcription start sites

3D – 3-dimensional

5’ RACE – 5’ Rapid Amplification of cDNA Ends

INTRODUCTION

Initiation of transcription is the first step in gene expression. At this step, live cells can regulate their gene expression in changing environmental conditions. The key enzyme of the transcription process in all domains of life is a multi-subunit enzyme, DNA-dependent RNA polymerase (RNAP). Transcription regulation is a dynamic and sophisticated process that involves many auxiliary transcription factors. In our laboratory, we study transcription regulation in bacteria using Gram-positive (G+) model organisms *B. subtilis* and *M. smegmatis*.

In this Thesis, I focus on the regulation of transcription initiation with RNAP-associated factors. This Thesis comprises five Chapters, each of them represents a project or closely related sub-projects pursuing the same global **aim** – **revealing the particular mechanism of regulation of bacterial transcription initiation under certain conditions**. The main emphasis of this Thesis is on the σ factors from *B. subtilis*.

First, using purified σ factors from *B. subtilis*, I studied the effect of concentration of initiating nucleoside triphosphate [iNTP] on the activity of selected promoters. Further, I comprehensively characterized the σ^I factor from *B. subtilis*.

My next project from systems biology area was developed in collaboration with bioinformaticians. I worked on the validation of predicted *in silico* σ^A-dependent interactions of the newly created σ^A-dependent gene network.

Further, I studied the impact of HelD and δ interacting partners of RNAP on selected aspects of bacterial transcription.

The final project where I participated in dealt with a small RNA Ms1 from *M. smegmatis*.

LITERATURE REVIEW

1. Model organisms

In this Thesis, the majority of experiments were performed with the model organism *Bacillus subtilis*, and several experiments were done with *Mycobacterium smegmatis*.

1.1 *Bacillus subtilis*

In our laboratory, we study bacterial transcription and gene expression using model G+ bacterium *Bacillus subtilis*. Non-pathogenic *Bacillus subtilis* is a rod-shaped, spore-forming microorganism that belongs to the *Firmicutes* phylum. *B. subtilis* is commonly found in the upper layers of the soil, and evidence exist that *B. subtilis* is a normal gut commensal in humans (Hong et al., 2009). *B. subtilis* serves as a model organism for a number of processes in G+ bacteria: replication, transcription, and sporulation (Browning and Busby, 2016; Earl et al., 2008; Helmann, 1995; Nicholson and Park, 2015; Piggot and Hilbert, 2004).

The *B. subtilis* genome has a single circular A-T rich chromosome containing about 4,100 genes, of which only 192 were shown to be essential (Kobayashi et al., 2003). Comparative genomic analyses based on microarray data have revealed that *B. subtilis* exhibit considerable genomic diversity (Earl et al., 2008). *B. subtilis* is a model organism used to study bacterial chromosome replication initiating from a single locus (*oriC*). Its replication proceeds bidirectionally and two replication forks go in clockwise and counterclockwise directions. Chromosome replication is completed when the forks reach the terminus region (*Ter*) placed opposite to *oriC*, as reviewed by Wen gang Xiao (Weng and Xiao, 2014). In the past two decades, *B. subtilis* had been frequently used for studying of bacterial gene expression (Earl et al., 2008; Krásny and Gourse, 2004; Nicolas et al., 2012)

B. subtilis is able to survive extreme environmental conditions of temperature, desiccation and nutrient deprivation, when it undergoes the process of sporulation (Mckenney et al., 2013). *B. subtilis* sporulation ability had been extensively studied. All basic steps of endospore formation have been discovered using *B. subtilis* (reviewed by Hilbert and Piggot (Hilbert and Piggot, 2004; Piggot and Hilbert, 2004).

Next advantage of *B. subtilis* species is its ability to uptake foreign DNA with its potential transformation and insertion into *B. subtilis* DNA. It happens when cells enter a special physiological state called competence, reviewed by Dubnau and Losick (Dubnau and Losick, 2006). Competence in *B. subtilis* is induced just before the end of logarithmic growth, especially under conditions of amino-acid limitation (Anagnostopoulos and Spizizen, 1961).

Finally, excellent fermentation properties of *B. subtilis* and high product yields have been utilized to produce various enzymes, such as amylase and proteases (Dijl and Hecker, 2013). *B. subtilis* also serves as a model for investigating of such important human pathogens as those from the *Bacillus cereus* group.

1.2 *Mycobacterium smegmatis*

M. smegmatis is non-pathogenic, Gram-positive bacterium that belongs to the *Actinobacteria* phylum. Its natural habitat is soil, water or plant surfaces. *M. smegmatis* cells create wrinkled creamy white colonies while it is growing on accessible nutrients. In exceptional cases of weakened immunity, *M. smegm*atis can be an opportunistic pathogen (Piersimoni and Scarparo, 2009). Among *Actinobacteria*, *M. smegmatis* is rapidly growing species (doubling time of approximately 3.5 h). *M. smegmatis mc*^{2} strain, in particular, is very useful for the molecular analysis of other species in the genus *Mycobacterium* (Fujiwara et al., 2012). The genome of *M. smegmatis* contains 6.988.209 base pairs. It has a 67% GC content and 33% AT content and, therefore is classified as a high G-C content G+ bacterium. 90% of the *M. smegmatis* genome represents coding regions that encode 6716 proteins (Gupta et al., 2011). Despite some limitations, *M. smegmatis* is a suitable model for studying pathogenic *Mycobacteria* as *M. tuberculosis*, *M. bovis*, and *M. leprae* (Altaf et al., 2010).

2. Bacterial transcription: its main players and mechanism

The transcription process (the copying of DNA into RNA) is the first step of gene expression. DNA-dependent RNAP is the key enzyme of transcription. Unlike Archaea and Eukarya, members of Bacteria domain have only one form of RNAP. Moreover, the crystal structure of the bacterial core enzyme (Zhang et al., 1999) is similar to that of the archaeal RNAP (Hirata et al., 2008) and the eukaryotic RNAPs (Cramer et al., 2001). Thus, core enzyme (E) is generally conserved among all kingdoms of life, the only differences are on the surface of the complex.

Bacterial DNA-dependent RNAP is a multi-subunit enzyme capable of transcription elongation. The additional "Bacteria-only" subunit essential for transcription initiation is the σ [sigma] factor (Burgess, 1969). σ factor specifically recognizes and binds to the promoter DNA when in complex with RNAP in a holoenzyme form (Eσ). There are many various σ factors in most bacterial species, and thus multiple forms of RNAP holoenzyme (Davis et al., 2016; Gruber and Gross, 2003; Paget, 2015).

2.1 Overview of bacterial transcription

Bacterial transcription consists of 3 main steps: initiation, elongation, and termination. This process requires two main players: RNAP holoenzyme comprising the RNAP core and a σ factor, and promoter DNA. Transcription **initiation** starts with the interaction of Eσ and promoter DNA to form an initial closed complex (**Figures 1, 2**). Several kinetic intermediates are created following with the "open complex" formation [RP_o] (**Figures 1, 2**), as reviewed by Ruff, et al. (2015) and Haugen et al. (2008) (Haugen et al., 2008; Ruff et al., 2015). Eσ separates the two strands of DNA and exposes a portion of the template strand. The section of promoter DNA that is within the open complex is known as a "transcription bubble". Structural studies revealed that the early steps of RP_o formation are often rapidly reversible in comparison to the slower "isomerization" step that includes DNA opening, as reviewed by Haugen et al. (Haugen et al., 2008).

At some promoters, a phenomenon known as "abortive transcription" occurs – repeating synthesis of 9-12 nt long transcripts (Goldman et al., 2009; Hsu, 2009; Hsu et al., 1995).

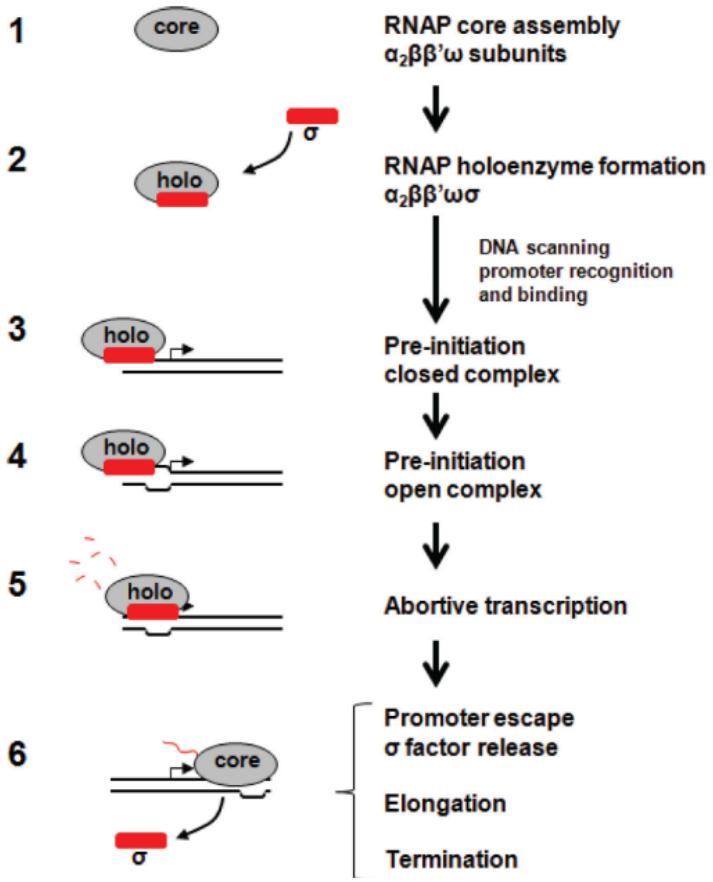


Figure 1. Simplified scheme of bacterial transcription. Transcription start site (+1) is indicated by the angled arrow. The scheme is adapted from Davis et al. (2016).

Transcription initiation is the most crucial phase of transcription from the kinetic point of view. Steps in the transcription initiation can be schematically summarized by the following equation:

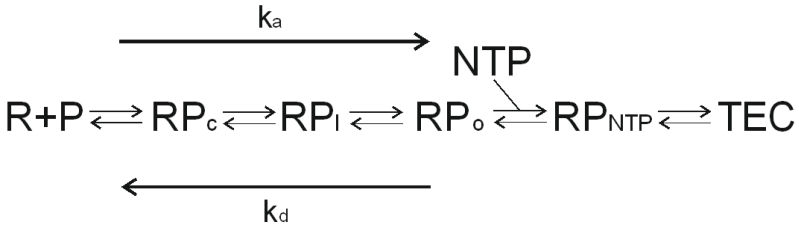


Figure 2. Kinetic steps of transcription initiation. RP_c refers to the earliest promoter complex with RNAP. RP_o refers to the final complex before nucleoside 5'-triphosphate occupancy of RNAP. RP_I is used as an abbreviation for all the intermediates between RP_c and RP_o. R – RNAP; P – promoter; NTP – nucleoside triphosphate; k_a is the composite association-rate constant for RP_o formation; k_d is the composite dissociation constant; TEC – transcription elongation complex. The scheme is adapted from Haugen et al. (2008).

All steps of transcription can be regulated, the rate-limiting step determining the frequency of transcription initiation are sequence and structure of the promoter DNA (McClure, 1985; Ruff

et al., 2015). *In vitro* transcription with purified components is a highly useful tool for studies of gene expression; it can be applied prior to numerous downstream applications in various biochemical and molecular biology studies. For a given promoter sequence, changes in temperature, salt, and solute concentrations (Kontur et al., 2010), as well as additions of protein factors and ligands, can affect kinetics by 10–1000-fold or more. In addition, binding of initiating NTP (iNTP) stabilizes the short-lived open complex at rRNA promoters (Gaal et al., 1997a), shifting the distribution of promoter complexes from closed to open in an NTP concentration-dependent manner (Ruff et al., 2015). It is also known that *B. subtilis* RNAP forms weaker open complexes than *E. coli* RNAP due to a different structure and lack of several DNA-binding domains that helps to form stable open complex (Artsimovitch et al., 2003).

After intermediate stages RNAP proceeds to **elongation** phase via DNA-scrunching mechanism (Hsu, 2002; Kapanidis et al., 2006) during the "**promoter escape**" process. (**Figure 1**). The generally accepted model is that during promoter escape σ dissociates from RNAP and undergoes "recycling" phase. This "σ factor cycle" emerged as σ^{70} is present in chromatographically or electrophoretically isolated RP_o, but is not present in chromatographically or electrophoretically isolated RNAP-DNA elongation complexes (Hansen and McClure, 1980; Straney and Crothers, 1985). This σ cycle model has been questioned in later studies, where it was suggested that σ can translocate to elongation phase in complex with RNAP (Bar-Nahum and Nudler, 2001; Mukhopadhyay et al., 2001). At this point, the transcription elongation complex (TEC) forms and elongation phase of transcription starts. In *E. coli*, GreA and GreB factors enhance the promoter escape phase (Hsu et al., 1995).

During elongation phase RNAP undergoes conformational changes. Newly synthesized nascent RNA leaves RNAP through the RNA-exit channel. RNAP moves along the template strand and synthesize nascent RNA. Thus, nascent RNA is a precise copy of template DNA strand. Based on the complementarity principle, ribonucleotides are attached to the free 3' end of nascent RNA which is created by -OH residue (Belogurov and Artsimovitch, 2015; Washburn and Gottesman, 2015).

There are two modes of transcription **termination** in bacteria: Rho-dependent and intrinsic. In the case of **Rho-dependent** termination, transcription ends when a termination factor ρ (rho factor) binds to a rho utilization site on the nascent RNA strand and moves along the mRNA toward the RNAP. ρ factor destabilizes TEC, and RNAP dissociates from the nascent RNA. This way of transcription termination requires additional ρ-dependent transcription factors NusA, NusB, NusG (Washburn and Gottesman, 2015). **ρ-independent** termination takes place when a stem-loop

structure is created upstream of the terminator region. This structure is composed of guanines and cytosines and followed with several uraciles. This loop causes interruption of the RNA:DNA hybrid and RNAP dissociation from the nascent RNA (Santangelo and Artsimovitch, 2011; Washburn and Gottesman, 2015).

2.2 Bacterial RNA polymerase – a multi-subunit enzyme

The bacterial RNAP core enzyme (~400 kDa) is a multi-subunit enzyme that consists of five subunits: two α, one β, one β', and one ω (α2ββ'ω) (Burgess, 1969; Burgess et al., 1969). The first bacterial RNAP core crystal structure was solved in 1999 (Zhang et al., 1999), the first holoenzyme crystal structure was solved in 2002 (Vassylyev et al., 2002). In 2015, a 3D crystal structure of a promoter open complex (RP_o) containing *Thermus aquaticus* RNAP holoenzyme and promoter DNA that included the full transcription bubble was published (Bae et al., 2015).

In the RNAP core, the α dimer forms the scaffold on which two catalytic subunits β and β′ assemble (**Figure 3a**), and the ω subunit assists β' binding to the α_2 sub-assembly (Minakhin et al., 2001). In both the core and holoenzyme structures of bacterial RNAP, the overall design resembles a crab claw, with the β and β' subunits forming the pincers [**Figure 3**] (Murakami and Darst, 2003; Vassylyev et al., 2002). The RNAP pincers form a 27Å wide internal channel, with the active site of the enzyme (where RNA phosphodiester bond catalysis occurs) positioned at its back wall (**Figure 3**).

RNAP holoenzyme is formed when σ factor binds to the core enzyme. Protein-protein interaction studies showed that σ^{70} (the main vegetative σ factor in *E. coli*) and RNAP interact in a multistep fashion, which suggests a cycle of changes at the interface between these proteins that are associated with progression through transcription initiation (Gruber et al., 2001; Haugen et al., 2008; Murakami et al., 2014; Zuo and Steitz, 2015). After binding of the σ factor and anchoring to the promoter in the holoenzyme-DNA complex, the conformation of the σ factor is extended, which means that the σ domains spread primarily along one face of RNAP. Domain 2 of σ interacts with a portion of the β' pincer, called the β' coiled-coil. This interaction is the most extensive RNAP interaction of the σ domains in terms of the contact area (Gruber et al., 2001; Young et al., 2001). Domain 3 of σ sits above the active site of the enzyme and mainly interacts with the β subunit (**Figure 3b-c**). The linker connecting domains 3 and 4 of σ is buried in the main channel as well as in the RNA exit channel, which passes under the β flap domain (Haugen et al., 2008).

There are two major channels in RNAP, main and secondary, which are bifurcated by a long, evolutionarily conserved feature of the enzyme, the bridge helix (**Figure 3**). These channels provide solvent access to the active site in the absence of DNA. Promoter DNA fills the main channel that is formed by the cleft between the β and β' pincers (Young et al., 2002). The secondary channel provides access for solvent and NTPs to the enzyme's active site and is wide enough to accommodate back-tracked RNA (Korzheva et al., 2000). Moreover, it appeared that the secondary channel is a route by which small molecules and small proteins can access the RNAP active site and cause a wide range of effects on transcription (Landick, 2005; Paul et al., 2004; Perederina et al., 2004). The nascent RNA transcript follows the path of the template strand for several bases and then exits the RNAP underneath a flexible element of β called the flap domain through the RNA exit channel (**Figure 3**).

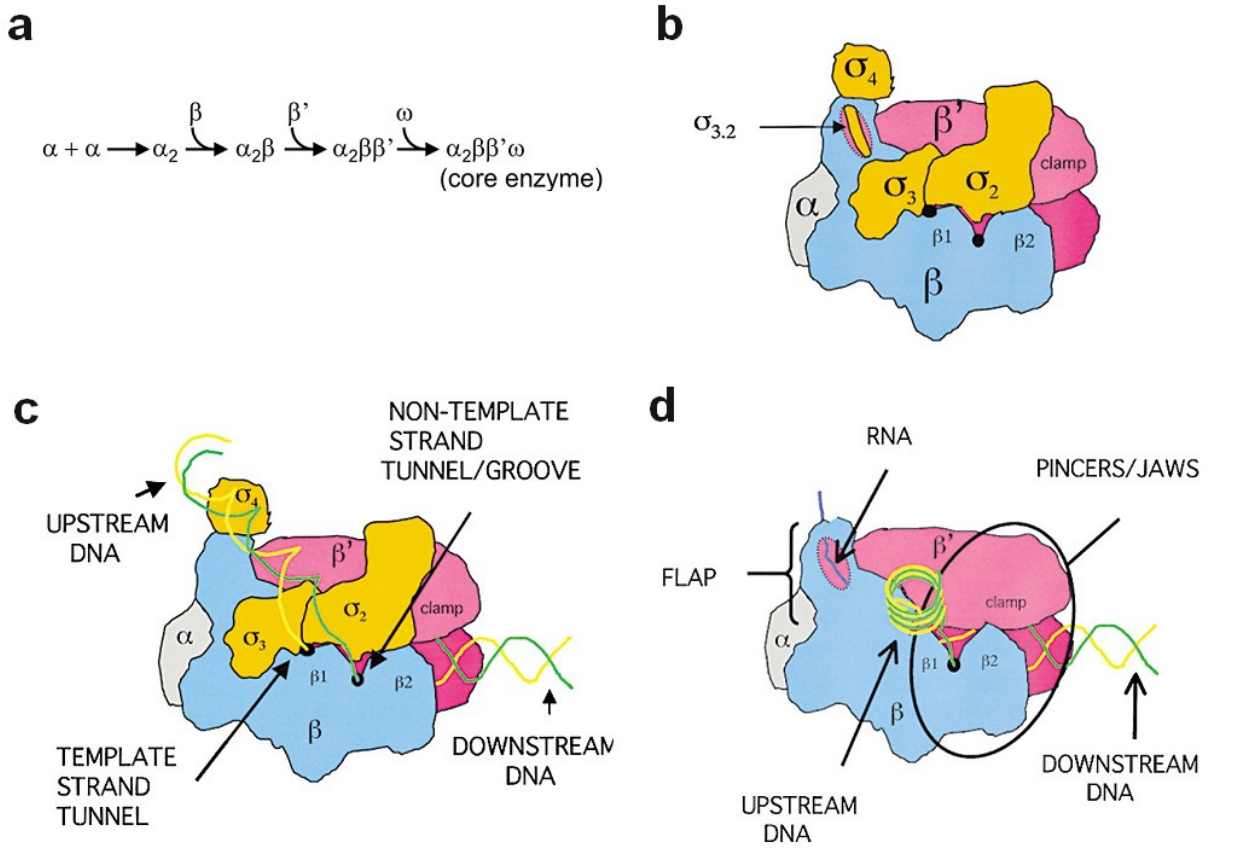


Figure 3. Bacterial RNAP assembly and its schematic structure. (a) Assembly scheme of the RNAP core enzyme (b) RNAP holoenzyme; (c) open complex; (d) elongating core. The β subunit is drawn in blue, the β' subunit in magenta, α subunits in gray (only one visible in this view), and the σ domains are shown in yellow (except 1.1 region). The flexible β flap domain is indicated. The picture was adapted from Sutherland and Murakami (2018) and Young, Gruber and Gross (2002) (Sutherland and Murakami, 2018; Young et al., 2002).

2.2.1 Bacterial σ factor as a specificity factor

σ factor is the master regulator of all gene expression in bacteria; it is an essential transcription factor that reversibly binds RNAP and mediates transcription of all bacterial genes. σ factor was first identified in 1969 as a protein that stimulates transcription by RNAP (Burgess, 1969; Burgess et al., 1969). σ factors play 3 major roles in the RNA synthesis initiation process. They (i) mediate interaction between RNAP holoenzyme and specific promoters, (ii) melt a region of double-stranded promoter DNA and stabilize it as a single-stranded open complex, and (iii) interact with other DNA-binding transcription factors that contribute to regulation of bacterial gene expression, as reviewed by Davis et al. (2017), Feklistov et al. (2014), Paget (2015) (Davis et al., 2017; Feklistov et al., 2014; Paget, 2015).

Primary (also main, principal, housekeeping, vegetative) σ factor recruits RNAP to the majority of the promoters, but nearly all bacteria have one or more alternative σ factors. Alternative σ factors are proteins that serve to reprogram promoter preferences of RNAP and change gene expression in the cell. In some cases, this reprogramming mediates global responses to general stress, while in others, the alternative σ factor participates in driving a developmental pathway such as sporulation. Majority of alternative σ factors are controlled by their availability in the cell, and therefore, target promoters tend not to require transcription activators and contain recognition elements that resemble the consensus (Koo et al., 2009). Under favorable growth conditions, the housekeeping σ factor is more abundant and thus able to outcompete alternative σ factors (Korzheva et al., 2000; Mekler et al., 2002; Murakami and Darst, 2003).

Each σ factor recognizes unique **consensus sequence**; promoters containing this sequence encodes genes that constitute the **regulon** specific for that σ factor (a group of regulated genes under control of the same regulator).

Each bacterial species has a different number of σ factors, *e.g. E. coli* – 7; *Streptomyces coelicolor* – 64, *Mycobacterium smegmatis mc² 155* – 24; *Corynebacterium glutamicum* – 6, as reviewed by Sun et al. (Sun et al., 2017). Number of σ factors vary among bacterial species between 1 and >100 (Feklistov et al., 2014). In most cases, a bacterial species has one primary σ factor. Depending on the conditions, appropriate σ factors reversibly associate with RNAP.

The most prevalent σ factor that is conserved in all bacterial genomes is vegetative σ^{70}-type (σ^{70} in *E. coli*, σ^{A} in *B. subtilis*) σ factor. The σ^{70} family of σ factors is historically named after the *Escherichia coli* housekeeping σ^{70}. The σ^{70}-dependent promoters usually consist of 2 highly

conserved sequences situated -10 and -35 nucleotides upstream of the +1 transcription start site (**Figure 4**).

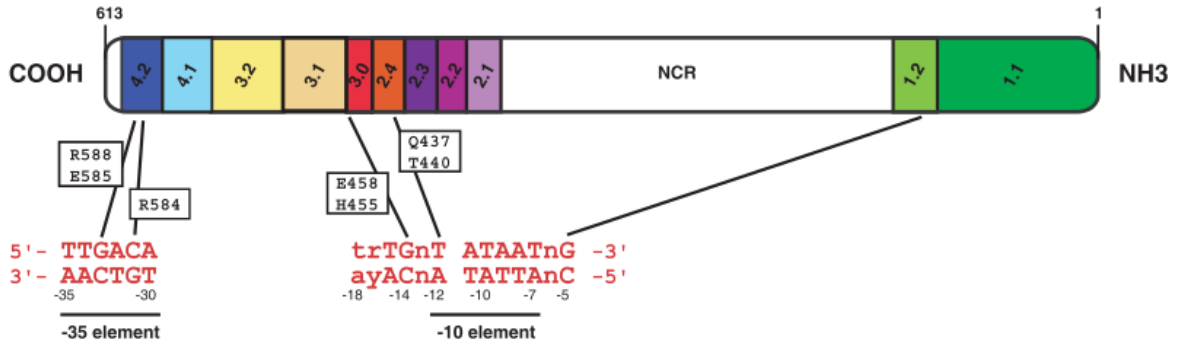


Figure 4. Schematic representation of σ^{70} regions. Subregions (1 - 4) and non-conserved region (NCR) are based on function, structure, and sequence conservation. Canonical promoter sequence elements of *B. subtilis* (upper row) and *E. coli* (bottom row) are shown below σ. Specific σ^{70} residues, which are thought to interact with base determinants, are indicated. The picture is adapted from Hook-Barnard and Hinton, 2007.

There are two families of σ factors: σ^{70} and σ^{54}. The σ^{54} family is structurally and evolutionarily distinct from the σ^{70} family, its representatives are frequently involved in regulating expression of genes involved in nitrogen metabolism. σ^{54} transcription factor forms a transcriptionally silent complex requiring specialized ATP-dependent activators for initiation.

Sequence alignments of the σ^{70} family members reveal four conserved regions that can be further divided into sub-regions 1– 4. The simplest σs have two domains (Group 4 or ECF σs: σ_2, σ_4), some have three domains (Group 3 σs: σ_2, σ_3, σ_4), and the housekeeping σs (Group 1) have four domains ($\sigma_{1.1}$, σ_2, σ_3, σ_4) connected by linker sequences of varying lengths (**Figure 5**) (Campbell et al., 2002; Malhotra et al., 1996; Murakami et al., 2002, 2014; Vassylyev et al., 2002). All of these domains mediate either protein-protein interactions with RNAP core, important interactions with promoter DNA, or catalyze important steps in the transcription initiation process.

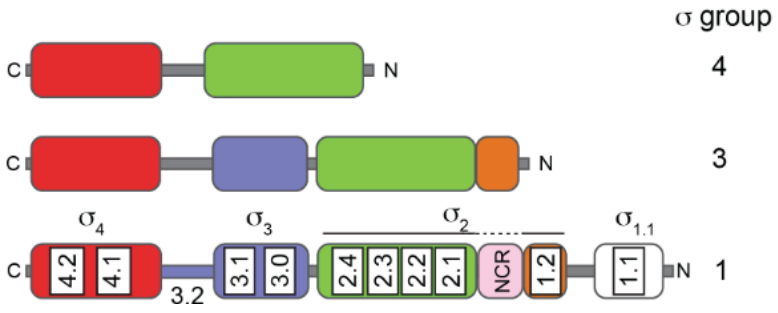


Figure 5. The domain organization of the Groups 1, 3 and 4 σ factors from σ^{70} family. Structural domains are colored: $\sigma_{1.1}$, white; σ_2, green/orange; σ_3, blue; σ_4, red. Within each domain, conserved σ regions are indicated for Group 1 σs. Note that σ_2 is colored green and orange to distinguish σ regions 2.1–2.4 and 1.2. The nonconserved region (NCR; pink) located between 1.2 and 2.1 (pink) is variable in size and structure among Group 1 σ factors. The picture was adapted from Paget 2015.

Primary σ factors of **Group 1** (σ^{70} in *E. coli*, σ^{A} in *B. subtilis*) usually have an extra domain 1.1 that auto-inhibits protein activity in free form, preventing its binding to promoter DNA (**Figure 6**) (Bowers and Dombroski, 1999; Dombroski et al., 1993). This auto-inhibition is relieved when σ binds the core RNAP to form the holoenzyme, thus playing a role in RP_o formation (Murakami and Darst, 2003). Structure of 1.1 from the *B. subtilis* of σ^{A} has been recently solved (Zachrdla et al., 2017): it is highly compact and requires minimal conformational changes for accommodating RNAP in the DNA channel.

Alternative σ factors differ from Group 1 σ factors by the complete absence of $\sigma_{1.1}$, the variable presence of σ_3, promoter specificity, and in some aspects of initiation. **Group 2** σ factors are structurally closely related to Group 1 but are non-essential. Where studied, Group 2 σs are usually involved in adaptation to stress including nutrient limitation and other stresses associated with the stationary phase. Members of **Group 3** σ factors are structurally and functionally diverse, but usually contain σ_2, σ_3, and σ_4 (**Figure 6**). The smallest σ^{70} proteins, the ~20 kDa **Group 4** consist only of domains σ_2 and σ_4 (Campbell et al., 2002), highlighting the functional importance of these domains. Group 4 is also known as the Extra Cytoplasmic Function (ECF) group on account of the frequent role of members in sensing and responding to signals that are generated outside of the cell or in the cell membrane (Lonetto et al., 1994).

Despite their large variation in size, from ~70 kDa for Group 1 to ~20 kDa for Group 4, all members of the σ^{70} family possess the σ_2 and σ_4 domains that include the major RNAP- and promoter-binding determinants (**Figures 4-5**). The extent to which alternative σ factors are used varies enormously between bacteria. In a similar way varies the range of functions of σ factors, from sensing and responding to a wide variety of extracellular and intracellular signals, to the expression of products directly involved in morphological changes in the cell.

The σ recognition strategy is described in details in the next sub-Chapter. Briefly, each domain (except for $\sigma_{1.1}$), has DNA-binding determinants: σ_4, -35 motif; σ_3, extended -10 motif; σ_2, -10 and discriminator motifs (Feklistov et al., 2014) [for details see Literature Review – Chapter 2.3].

Many σ factors are negatively regulated by a membrane or cytoplasmic proteins called **anti-σ factors** (Helmann, 2016a; Missiakas and Raina, 1998; Paget, 2015). Anti-σ factors are proteins that regulate transcription by sequestering one or more of the σ surfaces that bind to core RNAP or to promoter DNA in the absence of a specific environmental cue. With the imposition of a specific physicochemical signal, anti-σ factor function is abolished through targeted proteolysis (Heinrich and Wiegert, 2009), by structural modification (Kang et al., 1999; Paget, 2015), or by

the action of an anti-anti-σ factor (Francez-Charlot et al., 2009; Pané-Farré et al., 2005) [**Figure 6**]. They can be defined also as non-DNA-binding transcription factors that repress transcription by preventing a productive σ factor-core RNAP interaction, rather than by directly blocking RNAP access to promoter DNA, as in the case with most DNA-binding repressors (Browning and Busby, 2004).

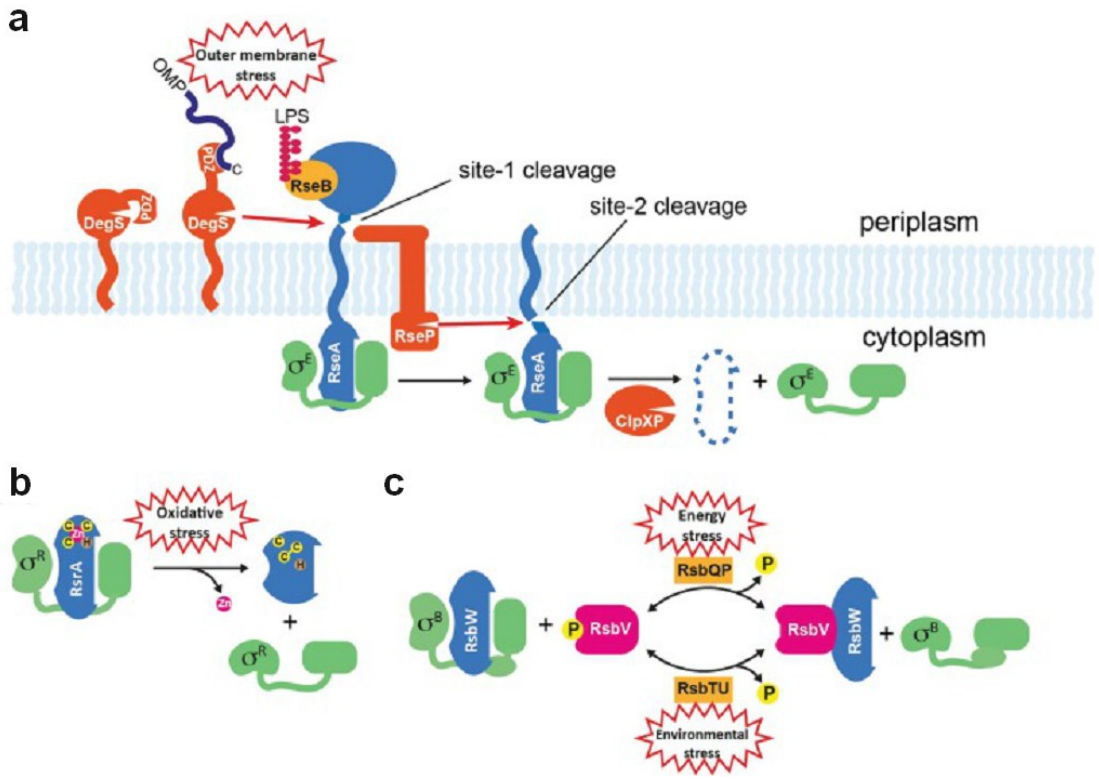


Figure 6. Mechanisms for σ factor release from anti-σ factors – examples. (a) Activation of σ^E in *E. coli*. The membrane-spanning anti-σ RseA binds σ^E through its cytoplasmic ASD; (b) Activation of σ^R in *S. coelicolor*. RsrA binds to and inactivates σ^R via its ZASD domain; (c) Activation of σ^B in *B. subtilis*. RsbW is an anti-σ factor/kinase that binds to σ^B and additionally inactivates its alternative binding partner RsbV by phosphorylating it to RsbV-P. The picture was adapted from (Paget, 2015).

The mechanism for releasing cytoplasmically-located σ factors in response to signals that often stem from the external environment is understood in only a small number of cases. They can be broadly divided into partner-switching, direct sensing and regulated proteolysis mechanisms (**Figure 6**). In the case of partner-switching and regulated proteolysis, an emerging theme is the integration of distinct signals involving separate input pathways that enable σ activation in response to varied environmental and physiological cues, as reviewed by Paget (Paget, 2015).

2.3 Structure of bacterial promoters and its recognition

Through many years of studying, promoter elements have been defined structurally, genetically, and functionally (summarized in **Figure 7**). Previously it was thought that regulons controlled by different σ factors contain distinct sets of genes. Genome-wide studies have now shown that many genes can be served by RNAPs carrying different σ factors, thus regulons of different σ factors are overlapping (Wade et al., 2006), such as for ECF σ factors (Helmann, 2016a).

Despite an estimated 2 billion years since the G+ and G- cell types divergence from a common ancestor (Doolittle et al., 1996; Feng et al., 1997), high conservation exists for the -10 and -35 elements found in bacterial promoters recognized by the primary σ^{70} factor of *E. coli* and *B. subtilis* σ^{A} (**Figures 5, 7**). The -35 and extended -10 motifs are recognized as double-stranded DNA and remain double-stranded throughout the initiation process.

σ factor ($\sigma_{4.2}$) specifically recognizes and binds the **-35** hexanucleotide sequence through multiple helix-turn-helix and major groove specific and nonspecific interactions (Campbell et al., 2002; Gardella et al., 1989; Lane and Darst, 2006; Siegele et al., 1989). The -35 element (**Figure 7**, blue) of primary σ factors has the consensus sequence 5'-TTGACA-3' (Hawley and McClure, 1983). $\sigma_{4.2}$ also forms a protein-protein interaction with the β subunit of core RNAP that is required for holoenzyme formation (Geszvain et al., 2004).

σ_2 recognizes and binds the **-10** element. The -10 element (**Figure 7**, yellow), with an all-AT bp consensus sequence (5'-TATAAT-3') constitutes the upstream half of the region opened by RNAP. Recent structural analyses suggest that there are fewer intermolecular interactions between domain σ_2 and duplex -10 element DNA and these are primarily of a nonspecific nature (Feklistov and Darst, 2011); residues of both template and nontemplate strands interact with $\sigma_{2.4}$ (Feklistov and Darst, 2011; Feng et al., 2016). After double strand opening, the -10 region of the nontemplate strand interacts with conserved residues of σ, with the nearly invariant bases -11A and -7T in pockets of σ^{70} (Feklistov and Darst, 2011; Feng et al., 2016; Zhang et al., 2012).

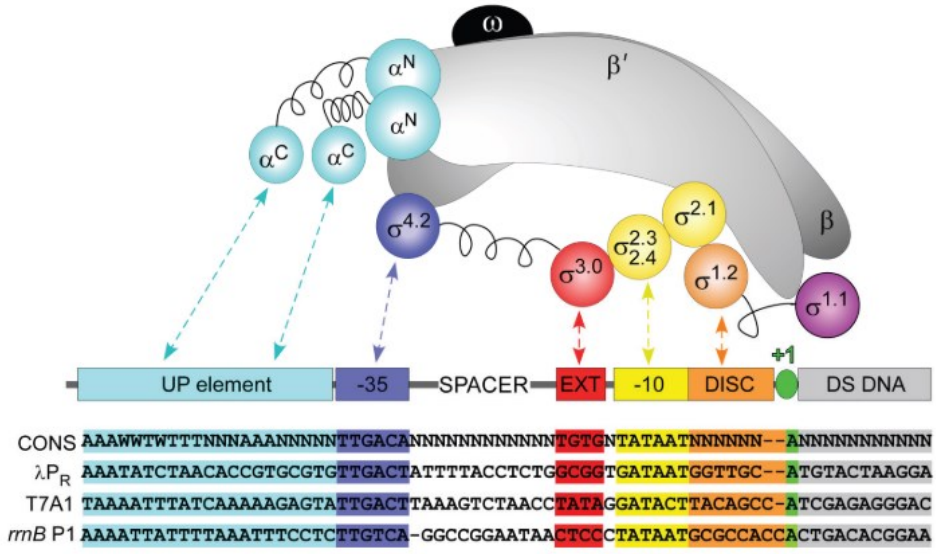


Figure 7. Sequence-specific interactions between σ^{70} RNAP and promoter regions. Non-template strand sequences of a "consensus" and λP$_R$, T7A1 and *rrnB* P1 promoters are shown below; missing bases are indicated by dashes (λP$_R$, T7A1, and *rrnB* P1 promoter sequences are shown as examples). σ^{70} regions interacting with promoter are colored in accordance with colors of promoter binding elements. UP element: cyan; -35 element: blue; extended -10: red; -10 element: yellow; discriminator: orange; transcription start site: green; DNA downstream of the transcription start site: gray. Linker regions in α and σ subunits are shown as springs. The scheme is adapted from Ruff et al. (2015).

There is no consensus sequence for the majority of the **spacer** between the -35 and -10 elements, but there is a consensus length, which is determined largely by the spacing between $\sigma_{4.2}$ and $\sigma_{2.3}$ (Murakami et al., 2002). The most common spacer length for σ^{70} promoters is 17 bp (Hawley and McClure, 1983; Shimada et al., 2014).

In addition to -10 and -35, other promoter elements exist such as the UP element (Aiyar et al., 1998), the extended -10 element (Barne et al., 1997; Keilty and Rosenberg, 1987; Sanderson et al., 2003), and the discriminator element (Haugen et al., 2006; Travers, 1980), which are functionally important in many sequences, as reviewed by Davis et al. (Davis et al., 2016) [**Figure 7**].

The **UP** elements (sequence upstream of the -35 element) are usually recognized by the C-terminal domains of the RNAP α subunits (αCTDs), they bind to the narrow minor groove of UP elements (Feklistov et al., 2014; Ross et al., 1993); **Figure 7**, light blue. The UP element consists of two subsites, proximal and distal, one for each αCTD. Promoters may have one or both, though distal UP element tends to function nearly like a full UP element (Estrem et al., 1999). The upstream sequence-specific interactions between RNAP and promoter DNA are in the UP element region from approximately base -40 to -60 (**Figures 7-8**), as reviewed by Gourse at al. (Gourse et al., 2000).

At some promoters, the "**extended -10**" element (TGn; **Figure 7**, red) increases promoter activity through specific contacts with α-helix of σ_3 (Barne et al., 1997) that recognizes the DNA major groove (Barne et al., 1997; Sanderson et al., 2003). The transcription bubble forms downstream of base -12 (for λP_R and other promoters with six bp discriminators); (**Figure 7**). The TGn motif was found in 20% of the 554 promoters identified in one study (Mitchell et al., 2003), with 44% of these promoters having a G at -14. σ_3 recognizes the extended −10 element TGx (Barne et al., 1997).

The **discriminator** region between the -10 element and the start site (**Figure 7**, orange) is involved in regulation of open complex lifetime. Its upstream end interacts with $\sigma_{1.2}$ (Basu et al., 2014; Feng et al., 2016; Zhang et al., 2012; Zuo and Steitz, 2015). Most discriminators are 6–8 bases in length (Shimada et al., 2014). The -10 and discriminator motifs are recognized upon strand separation as nontemplate-strand, single-stranded DNA (ssDNA, **Figure 8**). Usually, one of the double-stranded motifs (-35 or extended -10) is required for initiation, and these are thought to hold the holoenzyme in the appropriate position and orientation, and subsequently engage the -10 motif upon strand separation (Feklistov, 2013; Feklistov and Darst, 2011). After RP_c formation, the addition of nucleoside triphosphates (NTPs) enables a further transition to the initiating complex RP_o, which synthesizes the RNA transcript.

Although all specific recognition of promoter sequences is mediated by σ, it is generally accepted that σ does not specifically bind promoters on its own (Feklistov et al., 2014). As an exception, in *B. subtilis* the evidence exists that σ^D binds to the promoter without previous binding to RNAP (Chen and Helmann, 1995). It is unclear, however, whether this phenomenon is of any biological significance.

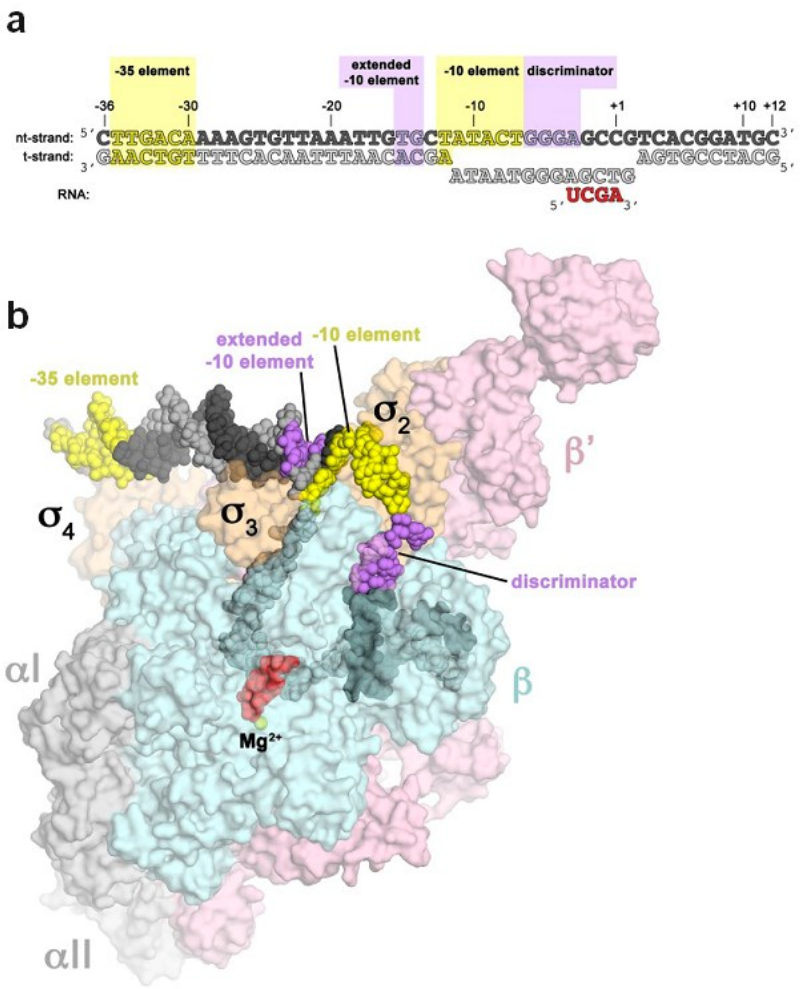


Figure 8. Structure of RP_o. (a) Oligonucleotides used for RP_o crystallization. The numbers above denote the DNA position with respect to the transcription start site (+1). The -35 and -10 elements are shaded yellow, the extended -10 and discriminator elements purple. The nt-strand DNA (top strand) is colored dark grey; t-strand DNA (bottom strand), light grey; RNA transcript, red. (b) Overall structure of RP_o. The nucleic acids are shown as CPK spheres and color-coded as above. The Taq EΔ1.1σ^A is shown as a molecular surface (αI, αII, ω, grey; β, light cyan; β', light pink; Δ1.1σ^A, light orange), transparent to reveal the RNAP active site Mg^{2+} (yellow sphere) and the nucleic acids held inside the RNAP active site channel. The picture was adapted from Bae et al., 2015.

During transcription initiation, RP_o is in equilibrium with intermediates of the transcription reaction. This equilibrium depends on several factors as a promoter sequence, the presence of Mg^{2+} ions and temperature of the reaction (Murakami and Darst, 2003). Based on the bacterial transcription initiation mechanism, bacterial promoters are divided according to their ability to bind RNAP and create RP_c and RP_o, also according to their RP_o stability (half-life of RP_o) (Gaal et al., 1997a; Krásný et al., 2008; Revyakin et al., 2004).

2.4 Interacting partners of RNAP in *B. subtilis*

Bacterial transcription is a highly sophisticated mechanism that requires precise control. In *B. subtilis*, there is a number of transcription factors and accessory proteins interacting with RNAP at different transcription stages (**Table 1**). Some of them are exclusive to G+ bacteria, *e. g*. δ (Weiss and Shaw, 2015), some have homologs in *E. coli*, *e.g*. – PcrA helicase from *B. subtilis* is a homolog of UvrD from *E. coli* (Gwynn et al., 2013).

Table 1. Proteins/transcription factors found associated with *B. subtilis* RNAP

Name	Function	Reference(s)
Small accessory subunits		
RpoE	δ subunit of RNAP, regulates RNAP by the concentration of the initiating nucleoside triphosphate (iNTP)	(Hyde et al., 1986)
RpoY	ε subunit of RNAP, control of RNAP activity	(Spiegelman et al., 1978)
YloH	ω subunit of RNAP, function unknown	(Minakhin et al., 2001)
Other interaction partners		
NusA	essential elongation factor; transcription termination factor of RNAP	(Yakhnin and Babitzke, 2002)
GreA	transcription elongation factor, resolves promoter-proximal pausing of RNAP	(Kusuya et al., 2011)
CshA	DEAD-box RNA helicase	(Hunger et al., 2006)
PcrA	ATP-dependent DNA helicase	(Petit et al., 1998)
Temporary interaction partners		
Spx	transcription regulator, interacts with RpoA	(Antelmann et al., 2000)
MgsR	transcription regulator orthologous to Spx, interacts with RpoA	(Reder et al., 2008)
Btr	transcription activator	(Baichoo et al., 2002)
YlyA	modulates σ^G-dependent transcription	(Nicolas et al., 2012)
Additional interaction partners of the RNAP		
HelD	DNA 3-5 helicase IV, stimulates transcription in an ATP-dependent manner by enhancing transcriptional cycling and elongation	(Delumeau et al., 2011; Wiedermannová et al., 2014)
TopA	DNA topoisomerase I, relaxation of negatively supercoiled DNA behind RNA polymerase	(Aravind et al., 1998)
CssR	two-component response regulator (OmpR family), control of cellular responses to protein secretion stress	(Fabret et al., 1999)
RnhC	Mn^{2+}-dependent RNase HIII, endonucleolytic cleavage of RNA in RNA-DNA hybrid molecules, processing of R-loops	(Wipat et al., 1996)

YpsC	23S rRNA methyltransferase, rRNA modification	(Delumeau et al., 2011)
Mfd	transcription-repair coupling factor, eliminates genetic damage from transcriptionally active genes during sporulation, required for increased mutagenesis of lagging strand genes	(Ayora et al., 1996)
YpiA	unknown	(Delumeau et al., 2011)
YdjO	unknown	(Huang et al., 1999)
ResD	two-component response regulator (OmpR family), regulation of aerobic and anaerobic respiration, activates expression of target genes in response to oxygen limitation	(Sun et al., 1996)

In *B. subtilis*, two subunits or accessory protein of RNAP that are exclusive for Gram-positive microorganisms were found – delta (δ/RpoE) (Lampe et al., 1988) and epsilon (ε/RpoY) (Spiegelman et al., 1978). In addition, later studies discovered a novel interacting partner of *B. subtilis* RNAP – putative helicase HelD (Delumeau et al., 2011). Part of my experimental work is related to δ and HelD proteins, thus next sub-Chapters contain detailed characteristics of these interacting partners of bacterial RNAP.

2.4.1 HelD, an interacting partner of RNAP from *B. subtilis*

In the growing bacterial cell DNA replication and transcription are frequently in conflict with each other (Lang and Merrikh, 2018). In *B. subtilis*, these conflicts are resolved by a molecular machinery involving essential helicase PcrA and accessory helicase DinG (Lang and Merrikh, 2018). One of the RNAP accessory proteins in *B. subtilis* is putative helicase, HelD (YvgS) [**Table 1**]. Understanding the exact mechanism of its functioning and role in the transcription is crucial for understanding transcription and gene expression regulation. In 2011, HelD was identified as a binding partner of *B. subtilis* RNAP (Delumeau et al., 2011). Based on sequence homology, HelD belongs to the superfamily I of DNA and RNA helicases (this work). This means that HelD is most closely related to HelIV helicases from G+ bacteria, as PcrA helicase from *Geobacillus stearothermophilus*. In *E. coli*, there are distantly related UvrD and Rep helicases. All these helicase homologs unwind DNA duplexes in an ATP-dependent manner, inchworming along the nucleic acid (Yang, 2010).

2.4.2 δ subunit of bacterial RNAP

δ protein is a non-essential RNAP subunit from G+ Firmicutes. The mechanism of functioning and role of δ subunit in the cell are comprehensively reviewed by (Weiss and Shaw, 2015). δ was first described in 1975 as ~21.5 kDa protein that copurified with *B. subtilis* RNAP (Pero et al., 1975).

These authors also showed for the first time the effect of δ on transcriptional specificity using *in vitro* assays. δ was shown to be required for maintaining the strand-specific transcription of phage genes. Since 1975 the δ subunit has been extensively studied by several groups.

Different studies showed that *rpoE* (the δ encoding gene) is highly expressed both in exponential and stationary phases under standard laboratory conditions (Watson et al., 1998; Weiss and Shaw, 2015) with a peak in the transition between exponential and stationary phases and it is also highly abundant in the spore extracts (López De Saro et al., 1999). δ is present in relatively equal amounts to other components of core RNAP, therefore suggesting a permanent interaction of this subunit with the transcription machinery (López De Saro et al., 1999).

The structure of δ protein consists of two distinct regions, an ordered and structured N-terminus, and a flexible and unstructured C-terminus whose amino acid composition – stretches of glutamic and aspartic acid residues – makes it a polyanion (Lopez de Saro et al., 1995). The δ protein has a significantly acidic pI of 3.6. Interaction of δ with RNAP is mediated by the N-terminus of the subunit (Lopez de Saro et al., 1995). The 3D structure of δ from *B. subtilis* was determined by NMR. Initially, the main focus was on the N-terminal 100 amino acids: four α-helices each being 5 to 12 amino acids in length, and β-sheet, consisting of three β-strands (Motácková et al., 2010; Papoušková et al., 2013).

Although many of early δ-related studies describe its negative effects on transcription of specific promoters, several groups have demonstrated that δ also has the capacity to increase overall transcriptional activity (Achberger and Whiteley, 1981; Juang and Helmann, 1994b; Spiegelman et al., 1978). The specific mechanism of this phenomenon is not fully understood.

In our group, one of the recent studies of *B. subtilis* δ was focused on defining the exact role of δ in promoter melting and open complex formation, with a specific focus on transcription regulation by the concentration of the iNTP /[iNTP] (Rabatinová et al., 2013). While high levels of iNTPs ensure efficient transcription, lower amounts cause the opposite effect. This work revealed that the δ subunit influences transcription by enhancing the effects of iNTP on RP_o formation. Promoters that possess relatively unstable open complexes require higher amounts of iNTP to initiate transcription (sensitive to [iNTP] promoters, *e.g.* rRNA promoters). δ destabilizes RP_o formation, thus increasing the amount of iNTP required for successful transcription initiation. δ thus potentiates promoter regulation by [iNTP]. This observation supports the concept of iNTP-dependent transcriptional regulation (Rabatinová et al., 2013).

Finally, deletion of δ caused deregulation of a number of genes of the cell and resulted in a decreased ability to adjust to changing environments together with general fitness defects. *rpoE*-

null mutants of *B. subtilis* were readily outcompeted by the wild-type strain when cultured together (Rabatinová et al., 2013).

3. *B. subtilis* σ factors

Most bacteria, except for symbionts and parasites with extremely reduced genomes, encode at least one alternative σ factor. *B. subtilis* has one primary σ factor, σ^A (Price and Doi, 1985), 17 alternative σ factors (Haldenwang, 1995; Helmann, 2016b; MacLellan et al., 2008; Matsumoto et al., 2005; Nicolas et al., 2012; Zuber et al., 2001) and one σ-like factor (Gruber and Gross, 2003; McDonnell et al., 1994). Among all *B. subtilis* σ factors, all but one are of σ^{70}-type, and σ^L is σ^{54}-type. Sequence alignment of *B. subtilis* σ factors is shown in **Figure 9**.

Of *B. subtilis* σ factors, the primary σ factor σ^A and general stress response factor σ^B are the most comprehensively characterized. Other alternative σ factors were studied occasionally, and there is a different amount of information (in some cases very limited) about each of them.

Table 2 demonstrates parameters of each *B. subtilis* σ factor, according to SubtiWiki database (Zhu and Stülke, 2018). Several lines of evidences exist in favor of the view that, in some circumstances, different σ factors compete for a limiting pool of the core enzyme, both in *E. coli* and in *B. subtilis* (Farewell et al., 1998; Hicks and Grossman, 1996; Lord et al., 1999; Rollenhagen et al., 2003).

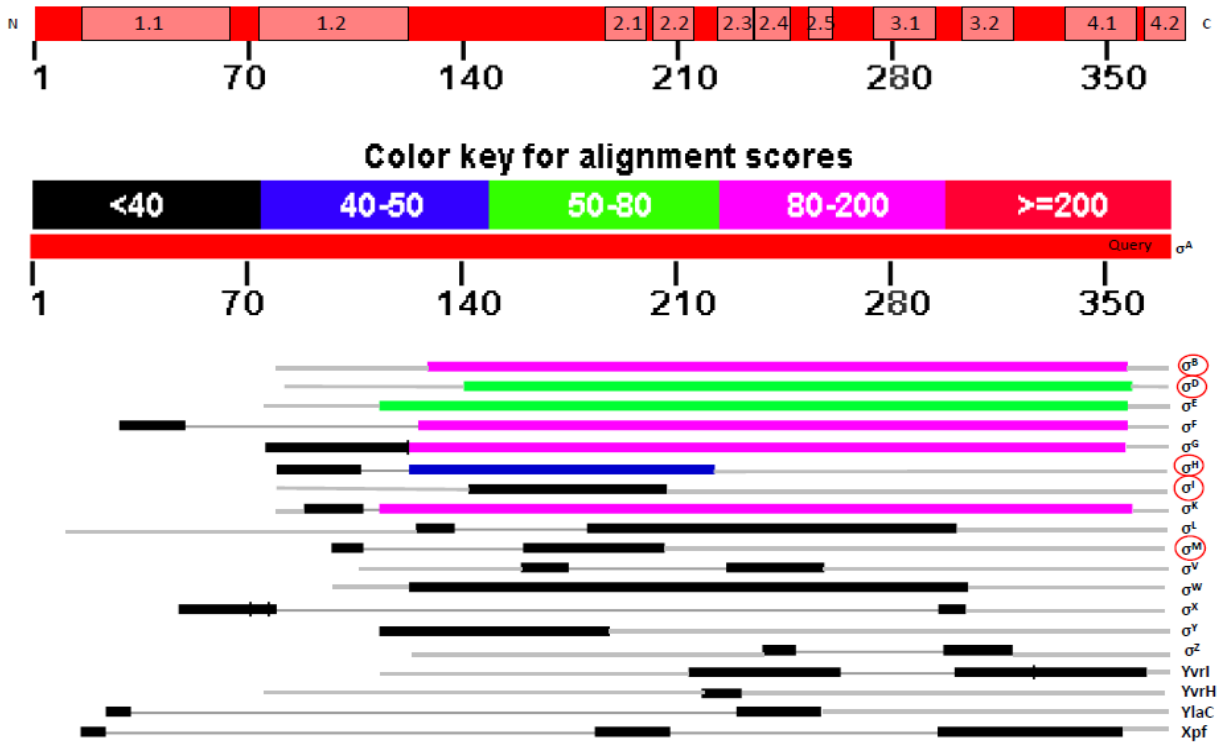


Figure 9. Alignment of *B. subtilis* alternative σ factors to the primary σ^A factor (upper row). The names of the σ factors that were purified for the purpose of this Thesis are put in red circles.

Table 2. Parameters of σ factors from *B. subtilis*

Name	Gene in the context of the operon	Mw, KDa	Number of regulated genes	Function	Reference(s)
σ^A σ^{55}	*yqxD-dnaG-**sigA***	42.80	Over 800	RNAP major σ factor	(Price and Doi, 1985; Shorenstein and Losick, 1973a)
σ^B σ^{37}	*rsbR-rsbS-rsbT-rsbU-rsbV-rsbW-**sigB**-rsbX*	29.99	234	General stress response σ factor	(Haldenwang and Losick, 1979, 1980)
σ^D σ^{28}	*ylxF-fliK-flgD-flgE-swrD-fliL-fliM-fliY-cheY-fliZ-fliP-fliQ-fliR-flhB-flhA-flhF-flhG-cheB-cheA-cheW-cheC-cheD-**sigD**-swrB*	29.32	151	Regulation of flagella, motility, chemotaxis, and autolysis	(Jaehning et al., 1979)

σ^E σ^{29}	*spoIIGA-**sigE**-sigG*	27.55	257	Transcription of sporulation genes (early mother cell)	(Haldenwang et al., 1981)
σ^F	*spoIIAA-spoIIAB-**sigF***	29.22	110	Transcription of sporulation genes (early forespore)	(Partridge et al., 1991)
σ^G	*spoIIGA-sigE-**sigG***	29.92	137	Transcription of sporulation genes (late forespore)	(Sun et al., 1989)
σ^H	***sigH**-rpmGB-secE*	25.30	49	Transcription of early stationary phase genes (sporulation, competence)	(Johnson et al., 1983)
σ^I	***sigI**-rsgI*	29.04	5	Control of a class of heat shock genes	(Zuber et al., 2001)
σ^K	***sigKN / sigKC***	17.00 16.00	75	Late mother cell-specific gene expression	(Zheng and Losick, 1990)
σ^L	***sigL***	49.54	6	σ^{54} -type σ factor. Utilization of arginine, acetoin, and fructose, required for cold adaptation	(Débarbouillé et al., 1991)
σ^M	***sigM**-yhdL-yhdK*	19.26	44	ECF-type, adaptation to inhibitors of peptidoglycan synthesis	(Horsburgh and Moir, 1999; Kunst, 1997)
σ^V	***sigV**-rsiV-oatA-yrhK*	19.57	23	ECF-type, resistance to lytic enzymes	(Horsburgh et al., 2001; Kunst, 1997)
σ^W	***sigW**-rsiW*	21.57	38	ECF-type, adaptation to membrane-active compounds required for the adaptation to membrane-active agents, activated by alkaline shock and by polymyxin B, vancomycin, cephalosporin C, D-cycloserine, and Triton X-100	(Kunst, 1997)
σ^X	***sigX**-rsiX*	23.03	18	ECF-type, resistance to cationic antimicrobial peptides	(Kunst, 1997)
σ^Y	***sigY**-yxlC-yxlD-yxlE-yxlF-yxlG*	21.21	2	ECF-type, maintenance of the SPß prophage, antibiotic production, and resistance	(Kunst, 1997; Mendez et al., 2012)
σ^Z	***sigZ***	20.57	-	ECF-type, transcription	(Kunst, 1997)
YlaC	*ylaA-ylaB-**ylaC**-ylaD*	20.78	1	ECF-type, transcription	(Kunst, 1997)
YvrI-YvrHa	***sigO-rsoA***	22.54-9.00	3	transcription	(MacLellan et al., 2008)
Xpf	***xpf***	19.95	1	transcription of PBSX prophage genes	(McDonnell et al., 1994)

In the experimental part of this Thesis, we focused on σ^B, σ^D, σ^H alternative σ factors, and especially on σ^I – one of the least explored σ factors with almost unknown regulon in *B. subtilis*.

3.1 σ^A

σ^A from *B. subtilis* was purified for the first time in 1973 in the group of R. Losick (Shorenstein and Losick, 1973a, 1973b). σ^A is the primary *B. subtilis* σ factor that is highly expressed in the exponential phase of growth (Shorenstein and Losick, 1973b). The molecular weight of σ^A is 42.80 kDa. In the live cell, the concentration of σ^A in the exponential and stationary phase of growth is approximately the same (Rollenhagen et al., 2003). σ^A has a regulon of more than 800 genes, mostly experimentally verified (Zhu and Stülke, 2018) and regulates the majority of vegetative genes in the *B. subtilis*. The consensus sequence of σ^A was reported to be TTGACA for the -35 region and tgnTATAAT for the -10 region, with the 14 nucleotides long spacer. Same as *E. coli* σ^{70}, σ^A possess auto-inhibitory 1.1 region in the sequence (Camarero et al., 2002; Zachrdla et al., 2017).

3.2 σ^B

σ^B is the general stress-response regulating σ factor, one of the best studied alternative σ factors in *B. subtilis*. It was discovered in 1979 in the group of R. Losick, same as the primary σ factor when they used the *ctc* and *spoVG* genes as templates for *in vitro* transcription assays (Haldenwang and Losick, 1979). In these pioneering *in vitro* studies, σ^B was identified as the first alternative σ factor of bacteria. The molecular weight of σ^B is 29.99 kDa. The σ^B-mediated response is triggered by diverse environmental stress signals and activates expression of a broad range of genes needed for cell survival in these conditions, thereby controlling one of the most comprehensive stresses/starvation regulons in *B. subtilis*, as reviewed by Hecker et al. (Hecker et al., 2007). The activity of σ^B is tightly regulated by a partner-switching network comprising σ^B, its antagonist anti-σ-factor RsbW, and anti-anti-σ-factor RsbV (Narula et al., 2016). The genes encoding σ^B and its regulators lie within a σ^B-controlled operon (Haldenwang, 1995). The consensus sequence of σ^B was reported to be AGGTTT for the -35 region and GGGTAT for the -10 region.

The ability of σ^B to compete successfully with σ^A for the core enzyme increases in the stationary phase. The concentration of σ^B increases fivefold during stress. At its maximum, σ^B concentration becomes at least two-fold higher than that of σ^A, though in exponential phase the

affinity of σ^B for the core enzyme is 60-fold lower than that of σ^A (Delumeau et al., 2002; Rollenhagen et al., 2003). The core of the σ^B regulon has been defined in several independent studies, and it is now clear that σ^B controls over 200 genes (Hecker et al., 2007; Zhu and Stülke, 2018).

3.3 σ^D

σ^D was discovered in 1979 in the M. J. Chamberlin's group (Jaehning et al., 1979). σ^D is present and active in the late exponential phase (Wiggs and Gilman, 1981). The molecular weight of σ^D is 29.32 kDa. The consensus sequence of σ^D was reported to be TAAA for the -35 region and GCCGATAT for the -10 region (Fredrick and Helmann, 1994). σ^D (product of *sigD* gene) is responsible for the expression of many genes encoding cell surface proteins related to flagellar assembly, motility, chemotaxis and autolysis (Mirel and Chamberlin, 1989; Serizawa et al., 2004). The anti-σ^D factor is FlgM (Caramori et al., 1996). Compilation of published data indicated that 151 genes are regulated by σ^D (Arrieta-Ortiz et al., 2015; Serizawa et al., 2004).

3.4 σ^H

σ^H was discovered in 1983 in the group of R. Losick (Johnson et al., 1983). The σ^H regulon comprises 49 genes (Britton et al., 2002). The σ^H (*spo0H*) gene product, directs transcription of several genes that function in the transition from exponential growth to stationary phase, including the initiation of spore formation and entry into the state of genetic competence (Britton et al., 2002). The consensus promoter sequence for σ^H was reported to be AGGTATT for the -35 region and GAATT for the -10 region (Britton et al., 2002). The molecular weight of σ^H is 25.30 kDa. In addition to genes that are under the direct control of σ^H, there are many genes whose transcription is indirectly influenced by σ^H. For example, during sporulation, σ^H stimulates transcription of the master regulator of sporulation, Spo0A (Predich et al., 1992). Spo0A, in turn, activates or represses a large number of genes, many of which are transcribed by RNAP containing σ^A. The regulation of σ^H itself is complex. Transcription of *sigH* is controlled directly by the transcriptional repressor AbrB and indirectly by the phosphorylated form of Spo0A (Britton et al., 2002). σ^A and Spo0A both control the expression of hundreds of genes involved in stationary phase adaptation in addition to those involved in the earliest stages of spore formation; also these regulators are under complex and interconnected regulatory control (Chastanet et al., 2010; Sonenshein, 2000).

3.5 σ^I

σ^I was discovered in 2001 by Zuber and colleagues (Zuber et al., 2001). It is encoded by the non-essential gene *sigI* (*ykoZ*) that is co-transcribed with the *rsgI* gene. *rsgI* encodes the cognate anti-σ^I factor RsgI. Transcription of the *sigI-rsgI* operon is driven from a σ^I-dependent promoter (Asai et al., 2007) and also from a σ^A-dependent promoter (Salzberg et al., 2013) [**Figure 10**]. The consensus promoter sequence for σ^I was reported to be ACCCCC for the -35 region and CGAA for the -10 region; the length of the spacer between -35 and -10 is 19 nucleotides (Muñoz-Gutiérrez et al., 2015; Tseng and Shaw, 2008). In addition, two nucleotides downstream from -35 the extended -35 element (AA) was defined (Tseng and Shaw, 2008). The predicted molecular weight of σ^I is 29.04 kDa, and it has a theoretical isoelectric point of 8.31. RsgI is a transmembrane protein that sequesters σ^I under favorable growth conditions and releases it when appropriate stimuli, for example, heat shock, appear (Asai et al., 2007). Degradation of RsgI by proteases ClpC and ClpP then provides a post-translational layer of regulation of the σ^I activity (Liu et al., 2017).

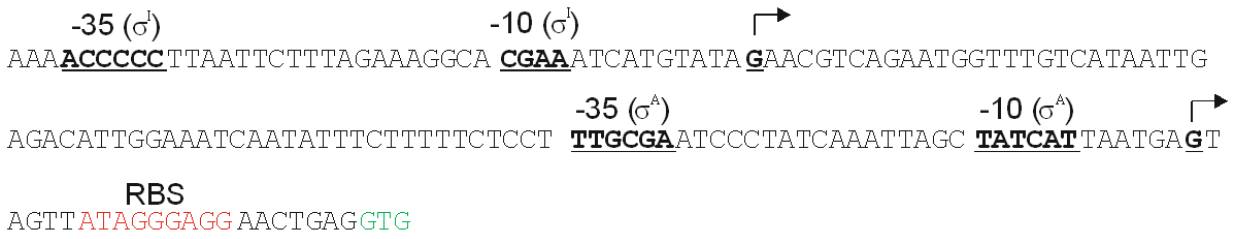


Figure 10. The DNA sequence of the *sigI-rsgI* promoter region. The consensus -35 and -10 regions of the σ^I and σ^A promoters are shown in bold and underlined, the initiation points of transcription are indicated by bent arrows; the ribosome binding site (RBS) and the initiation codon (GTG) are indicated in red and green, respectively. The picture is adapted from Salzberg et al., 2013.

A σ^I deletion strain was reported to be unable to grow on agar plates at 54°-55°C (Zuber et al., 2001). This temperature-sensitive phenotype of a strain lacking σ^I was also implied in other studies (Asai et al., 2007; Huang et al., 2013; Schirner and Errington, 2009; Tseng and Shaw, 2008). Moreover, it was shown that along with the heat shock response, σ^I was involved in the cold shock response (Schirner and Errington, 2009). Currently, there are only seven genes organized in five operons known to be σ^I-dependent in *B. subtilis*: the *sigI-rsgI* operon itself (Asai et al., 2007), *bcrC*, the *mreBH-ykpC* operon (Tseng and Shaw, 2008), *lytE* (Tseng et al., 2011), and *gsiB* (Zuber et al., 2001). Expression of some of these genes was shown to be stimulated by WalR. WalR is the response regulator of the WalRK two-component system that controls cell wall metabolism (Dubrac et al., 2008). BcrC helps protect the cell against bacitracin (Podlesek Z.,

Comino A., Herzog-Velikonja B., Zgur-Bertok D., Komel R., 1995) and paraquat (Cao et al., 2005) and is also needed for the production of the carrier lipid for cell wall synthesis (Cao and Helmann, 2002). MreBH participates in the formation of straight rod-shaped cells, and its depletion or overexpression in *B. subtilis* leads to the appearance of malformed cells (Carballido-López et al., 2006; Defeu Soufo and Graumann, 2003; Jones et al., 2001; Kawai et al., 2009). Moreover, MreBH is required for the lytic activity of a cell wall hydrolase, LytE, that is important for cell elongation and separation (Bisicchia et al., 2007; Hashimoto et al., 2012; Margot et al., 1998). The function of the *ykpC* gene located in the same operon with the *mreBH* gene is still unknown. GsiB is a general stress protein that prevents inactivation of cellular enzymes upon freeze-thaw treatments and is involved in responses to nutrient deprivation (Mueller et al., 1992; Völker et al., 1994).

4. Reconstructions of σ factor regulatory network

In the past decades, rapid advances in genomic technology have generated an enormous wealth of data on which mathematical and statistical tools can be applied to infer qualitative and quantitative relationships between DNA, RNA, proteins and other cellular molecules.

Efforts to integrate high throughput transcriptomic, proteomic and metabolomic data (Becker and Palsson, 2008; Shlomi et al., 2008; Yizhak et al., 2010) have led to the reconstruction and curation of a large number of organism-specific genome-scale metabolic networks (Feist et al., 2009). Much of the information contained in these studies is still under-exploited or disordered. One of such examples is also the amount of available experimental evidence from transcriptomic studies of *Bacillus subtilis* (Nicolas et al., 2012; Zeng et al., 2015). Consequently, our understanding of gene regulatory networks in *B. subtilis* is far from complete. Mining transcriptomic databases offers an opportunity to provide new insights into σ factor-controlled networks. Still, the task of current interest inferring gene regulatory networks in bacteria is the identification of the target genes of σ factors.

There are two main methods to discover such target genes of σ factors in bacteria: chromatin immunoprecipitation (ChIP) experiments and gene expression analysis. ChIP methods (ChIP-chip and ChIP-seq) test for physical interactions between σ factors and gene promoter sequences. However, it has been shown that this static binding information may also include silent binding events that do not directly enhance transcription (Macquarrie et al., 2011; To and Vohradsky, 2010). To increase certainty, ChIP experiments are complemented with RNA-seq experiments in strains with deletions in the σ factors of interest. However, such deletion is not possible for essential σ factors (*e.g.* σ^A in *B. subtilis*), for which a different approach must be employed. One such approach is the **kinetic modelling** of gene expression, as measurements of gene expression over time enable the reconstruction of transcriptional networks. The process of reconstructing biochemical networks using genomic data (also known as network inference or reverse engineering), bringing us one step closer to understanding how genetic background combined with non-genetic, environmental factors influence the characteristics of a living system (Wang and Huang, 2014).

Various methods to infer gene regulatory networks from gene expression data have been suggested, based on ordinary and stochastic differential equations, neural networks, dynamic Bayesian networks, and information theoretic- or correlation-based methods, which have been reviewed by Bansal et al. (2016), Penfold and Wild (2011) and Bar-Joseph (2012) (Bansal et al., 2006; Bar-Joseph et al., 2012; Penfold and Wild, 2011). Similar to ChIPseq, kinetic modelling alone is not sufficient to reliably determine regulation networks, and multiple sources of information have to be combined.

Currently, several databases comprising regulatory interactions in microorganisms have emerged (Zhulin, 2015). The examples are IMG – comprehensive platform for annotation and analysis of microbial genomes and metagenomes (Markowitz et al., 2012); MicrobesOnline – portal for comparative and functional microbial genomics (Alm et al., 2005); GOLD – resource for comprehensive information about genome and metagenome sequencing projects (Mukherjee et al., 2018); and KEGG [Kyoto Encyclopedia of Genes and Genomes] (Kanehisa, 1996). In the laboratories working with *Bacillus*, one of the most frequently used databases is SubtiWiki (Zhu and Stülke, 2018). Though performing a great job in archiving, storing, maintaining, and sharing information on genes, genomes, expression data (also on protein sequences and structures, metabolites and reactions, interactions, and pathways), these databases are non-specific in the sense of particular developmental processes of bacteria. They collect information about regulatory

interactions between the regulator and its targets, regardless of the conditions under which the particular experiment was made.

A combination of kinetic expression modelling obtained for particular experimental conditions with static and databased data may provide new insight into the kinetics of the control of, for example, σ factors and their target genes, and consequently allow modelling the σ factor-controlled network.

Creating of *B. subtilis* gene regulatory network requires the application of **meta-analysis**. A meta-analysis is a statistical analysis that combines the results of multiple scientific studies and/or databases (Paul and Leibovici, 2013; Walker et al., 2008). A key benefit of this approach is the aggregation of information leading to a higher statistical power and more robust point estimate than is possible from the measure derived from any individual study. However, in performing a meta-analysis, an investigator must make choices which can affect the results, including deciding how to search for studies, selecting studies based on a set of objective criteria, dealing with incomplete data, analyzing the data, and accounting for or choosing not to account for publication bias (Walker et al., 2008). Kinetic modelling and meta-analysis were used in our study to create the σ^A factor-controlled network in *B. subtilis*.

5. Regulation of initiation of bacterial transcription

Bacteria need to respond effectively to different stresses to ensure their survival in the changing environment. The most economical way to do it is to adjust gene expression on the level of transcription initiation. In *B. subtilis*, regulation of transcription with transcription factors has been extensively studied. The most prominent mechanisms of regulation of bacterial transcription initiation are:

a) Regulation with factors that interact at specific promoters, thereby affecting transcription directed by these promoters (DNA-binding activators and repressors, *e. g*. Lac operon);

b) Regulation with factors that interact with RNAP and change its preferences for target promoters (*e. g*. σ factors);

c) Regulation with transcription factors and small regulatory molecules that do not interact with DNA or RNAP (anti-σ factors, nucleoside triphosphates [NTPs], ppGpp, DksA);

d) Pervasive (antisense) transcription (asRNAs);

Some mechanisms of bacterial transcription regulation are still understudied, *e.g*. pervasive (or genome-wide) transcription, that was previously dismissed as an artefact at all domains of life (Lybecker et al., 2014; Wade and Grainger, 2014).

5.1 DNA-binding regulators

Transcription factors (trans-acting transcription factors) either enhance or block access to the promoter. For this purpose, transcription factors have structural motifs that bind to promoters that contain cognate "operators", which are specific sequences. The operators for most bacterial DNA-binding structural motifs, such as the helix-turn-helix motif, have 4-5 base pairs. In many transcription factors, operators mediate DNA binding in response to the cues that are sensed by the binding of a small ligand or protein, or covalent modification. The activities of other transcription factors depend on their abundance and availability, which can be regulated by synthesis, turnover or sequestration (Browning and Busby, 2016).

Many DNA-binding transcription factors in G+ and G- microorganisms were found: more than 300 in *E. coli* (Browning and Busby, 2016; Ishihama, 2012); 237 in *B. subtilis* (Moreno-Campuzano et al., 2006). Basically, they share a similar mode of action with the textbook example of DNA-binding regulators – *lac* operon in *E. coli* (Jacob and Monod, 1961). The examples of most prominent of global regulators in *E. coli* – DnaA, Dps, CitB, Fis, Fur, LexA (Ishihama, 2012; Zhu and Stülke, 2018); in *B. subtilis* – AbrB, CodY, ComK, Spo0A, Fur (Moreno-Campuzano et al., 2006; Zhu and Stülke, 2018).

Although some transcription factors regulate only a single promoter, most transcription factors regulate many promoters. In addition, many promoters are regulated by more than one factor. Furthermore, many transcription factors are expressed from promoters that are themselves regulated by other transcription factors. Promoter regulation by transcription factors, therefore, generates a complex regulatory network in which the concerted activities of specific, global and master regulators orchestrate the distribution of RNAP to the various transcription units that are present in the genome (Cho et al., 2011; Ishihama, 2012; Martínez-Antonio and Collado-Vides, 2003).

5.2 RNAP-associated regulators

Many factors (proteins and ligands) affect the activity of RNAP at different promoters through direct interactions. The simplicity and ubiquity of regulation by σ factors (Literature review – Chapter 3) have detracted attention from numerous other regulatory factors that interact with RNAP, thus regulating transcription initiation. Some of these regulators are restricted to particular groups of bacteria, such as RbpA and CarD, which bind to and stabilize open complexes in Mycobacteria (Flentie et al., 2016). The extended coiled-coil protein motif in DksA of *E. coli* inserts into a narrow channel (known as the secondary channel) that leads from the surface of RNAP to the active site (Perederina et al., 2004). Recent studies suggest that DksA together with GreB and other proteins function as "inspectors" that continually probe the enzyme active site by making rapid transient "visits" to the secondary channel in *E. coli* (Zhang et al., 2014).

Other factors also can decrease the number of RNAP molecules that are available for transcription by sequestering the holoenzyme. One example is 6S RNA (Literature review – Chapter 5.4), an approximately 180-nucleotide non-coding small RNA that is synthesized in response to slow growth and forms a 1/1 complex with the RNAP holoenzyme (Cavanagh and Wassarman, 2014). In *E. coli*, 6S RNA is a mimetic for the DNA of promoters that are targets for the housekeeping RNAP holoenzyme (that is, holoenzyme that contains σ^{70}). In Mycobacteria, a structural homolog of 6S was found –sRNA Ms1(Pánek et al., 2011).

The *B. subtilis* protein Spx modulates the function of the RNAP α subunit. Spx functions as an anti-α factor, and it can activate or inhibit transcription at specific promoters during oxidative or disulfide stress conditions, depending on the specific promoter (Nakano et al., 2005; Zuber, 2004).

Small ligands provide an alternative mechanism by which RNAP can respond quickly and efficiently to the environment. The best example is guanosine 3',5' bisphosphate (ppGpp), which is synthesized when amino-acid availability is restricted to the extent that translation is also limited (Chatterji and Kumar Ojha, 2001). ppGpp works by destabilizing open complexes at promoters that control synthesis of the machinery for translation (Barker et al., 2001). It has been proposed that ppGpp controls the expression of the translation machinery in response to sudden starvation (Schneider et al., 2003).

5.3 iNTP and the control of transcription initiation

The activity of RNAP holoenzyme can be regulated by fluctuations in the levels of another small ligand – the four nucleoside triphosphate (NTP) substrates; this phenomenon is an important mechanism of regulation of gene expression in bacteria. The phenomenon of gene expression regulation by [iNTP] was previously studied *in vitro* with *E. coli* promoters (Gaal et al., 1997b; Liu and Turnbough, 1994; Liu et al., 1994; Walker et al., 2004) and *B. subtilis* ribosomal promoters (Sojka et al., 2011).

The concentration of iNTPs (same as ppGpp) changes depending on the growth phase and growth conditions (Jöres and Wagner, 2003). NTPs are substrates of RNAP in the transcription process, their concentration in the cell can affect the intensity of transcription for certain promoters. It is important that the Michaelis Constant for the initiating nucleotide is higher than that for subsequently added NTPs (Mangel, Walter F. and Chamberlin, 1974). It means that, for a given rate of transcription, the required concentration of the iNTP is higher than that of subsequent NTPs. Consequently, the concentration of the initiating NTP is crucial to the activity of the RNAP, as all promoters require higher concentrations of the iNTP than the subsequent NTPs. There are certain promoters in *E. coli* – for example, rRNA, some tRNA promoters (Gaal et al., 1997a; Murray et al., 2003; Paul et al., 2004) – that require even higher concentrations of the iNTP than promoters in general. Regulation by [iNTP] has been documented for different promoters of G- *E. coli* (Liu C 1994; Liu J 1994; Walker 2004) and *Salmonella typhimurium* (Schwarz 1975; Sorensen 1993) or σ^A-dependent ribosomal promoters of *B. subtilis* (Sojka 2011). In general, there are two types of promoters with respect to [iNTP]: sensitive (requiring high iNTP concentration) and insensitive (with a constant high expression even at low iNTP concentration) (Gaal et al., 1997a). K_{NTP} is a constant characterizing iNTP requirements of a promoter – is the iNTP concentration required for half maximum level of transcription under certain conditions with the certain promoter.

As the iNTP for rRNA transcripts in *E. coli* is mainly ATP (at some cases also GTP or CTP), transcription initiation at rRNA promoters is expected to be sensitive to the cellular concentration of ATP (also GTP or CTP), which increases as cells leave stationary phase with a concomitant burst of rRNA synthesis (Murray et al., 2003; Schneider et al., 2002). In *B. subtilis*, the initiating NTP for rRNA is exclusively GTP, which decreases in abundance when it is consumed for the synthesis of ppGpp. Therefore, ppGpp levels in *B. subtilis* indirectly couples metabolism to RNAP activity through corresponding changes in the GTP level (Krásny and Gourse, 2004; Liu et al.,

2014). In extended stationary phase, depletion of NTPs therefore preferentially inhibits these promoters even though little or no ppGpp is present (Murray et al., 2003). The rapid increase in iNTP concentration is responsible for the increase in rRNA transcription that occurs when cells emerge from stationary phase (Murray et al., 2003).

The mechanism by which changes in the iNTP concentration selectively affect rRNA synthesis has been debated. The transient stabilization of the intrinsically short-lived open complex by the pairing of the first NTP (and subsequent NTPs) with the template strand may be sufficient to stimulate the initiation reaction by mass action (Barker et al., 2001; Gaal et al., 1997a; Paul et al., 2004). Alternatively, it was proposed that binding of the first two NTPs results in a conformational change in RNAP and that this conformational change promotes transcript initiation independently of its role in stabilization of promoter-RNAP interactions (Lew and Gralla, 2004).

5.4 6S and Ms1 – small RNAs as transcription regulators

Small RNAs (sRNAs) are well-established regulators of gene expression in bacteria. Most of sRNA function by base-pairing to target RNAs, but there are examples of sRNAs that interact directly with proteins and modulate their activity. Probably the best-described example of such RNA is the ~180 nt long 6S RNA. 6S RNA was discovered almost 50 years ago as a highly abundant, stable RNA, expressed during stationary growth of *E.coli* (Hindley, 1967) and it is one of the first RNAs to be ever sequenced (Brownlee, 1971). Despite of this, the role of 6S RNA had been a mystery till 2000, where it was discovered that *E.coli* 6S RNA specifically forms a complex with the housekeeping holoenzyme of RNAP [$E\sigma^{70}$] (Wassarman and Storz, 2000).

In the exponential phase of growth, gene expression is driven by the complex of RNAP with the primary σ factor ($E\sigma^{70}$). After the entry into stationary phase, many σ^{70}-dependent are downregulated and genes recognized by alternative σ factors are activated. 6S RNA accumulates in the stationary phase of growth and by binding to $E\sigma^{70}$, 6S RNA regulates expression of several hundred genes in *E. coli* (Cavanagh et al., 2008; Neusser et al., 2010).

6S RNAs have a highly conserved secondary structure consisting of a long double-stranded hairpin interrupted with a central single-stranded region (bubble) and this structure is essential for binding to RNAP holoenzyme (Shephard et al., 2010; Trotochaud and Wassarman, 2005). Based on structural similarity, 6S RNA was found in many bacterial species (Barrick et al., 2005; Trotochaud and Wassarman, 2005; Wehner et al., 2014), such as *Bacillus subtilis* and *Legionella pneumophila* that contain two 6S RNA genes per chromosome (Barrick et al., 2005; Faucher et

al., 2010; Trotochaud and Wassarman, 2005; Weissenmayer et al., 2011), cyanobacteria: *Synechococcus* (Watanabe et al., 1997), *Prochlorococcus* (Axmann et al., 2007) *or Synechocystis* (Rediger et al., 2012), hyperthermophilic bacteria *Aquifex aeolicus* (Willkomm et al., 2005), and also bacterial pathogens such as *Bordetella pertussis* (Trotochaud and Wassarman, 2005), *Helicobacter pylori* (Sharma et al., 2010), *Clostridium difficile* (Soutourina et al., 2013), and *Yersinia pestis* (Yan et al., 2013).

So far, 6S RNA has not been found in Mycobacteria (Arnvig et al., 2011; Pánek et al., 2011; Wehner et al., 2014). However, in our lab, we discovered a structural homolog of 6S RNA from *M. smegmatis* (Pánek et al., 2011) and named it Ms1. The same RNA was later found in *M. tuberculosis* (named MTS2823) (Arnvig et al., 2011) and Ms1 homologs were predicted in other mycobacterial and actinobacterial species (*Nocardia* and *Rhodococcus*) (Hnilicová et al., 2014). Ms1 and its homolog MTS2823 are together with the ribosomal RNAs the most abundant RNAs inside the stationary phase mycobacterial cell (Arnvig et al., 2011; Hnilicová et al., 2014).

Ms1 is longer than 6S RNA – Ms1 has the length of ~300 nt, while the usual length of 6S RNAs is between 180-200 nt. The Ms1 predicted structure is a double-stranded hairpin interrupted with the central bubble, which resembles 6S RNA. However, the Ms1 structure has short hairpins at the 5' and 3' ends that are missing in 6S RNA structures.

The proposed 6S RNA definition is that "6S RNA is an RNA that binds the primary form of their cognate RNAP holoenzyme in a manner resembling promoter DNA binding". 6S RNA does not bind to the core RNAP free of σ factors, or to RNAP holoenzymes with alternative σ factors (Cavanagh and Wassarman, 2014). While 6S RNA interacts with RNAP holoenzyme (Eσ), Ms1 binds the RNAP devoid of σ factors (Hnilicová et al., 2014) therefore it is not 6S RNA. Currently, the function of Ms1 in the cell is unknown.

5.5 Regulation by antisense RNAs

Although bacterial genomes are usually densely protein-coding, genome-wide mapping approaches of transcriptional start sites (TSSs) revealed that a significant fraction of the identified promoters drives the transcription of noncoding RNAs, as reviewed by Georg and Hess (Georg and Hess, 2018). A significant fraction of these noncoding RNAs consists of natural antisense transcripts (asRNAs), which overlap other transcriptional units. This phenomenon is called **pervasive transcription** (Wade and Grainger, 2014). In *E. coli* asRNAs consists 37% of all TSSs (Thomason et al., 2015).

In *B. subtilis*, asRNAs are less often expressed from σ^A-dependent promoters than mRNAs and, furthermore, the expression of sense-antisense pairs is often anticorrelated (Nicolas et al., 2012). Although mechanisms of action of asRNAs are often coupled with translation blocking (Georg and Hess, 2018), an interesting hypothesis is that these asRNAs are activated by alternative σ factors at stress conditions to repress or enhance the turnover of mRNAs from σ^A-dependent genes. For example, in *Bacillus,* the stress-dependent σ^B regulon consists of ~200 protein-coding genes and 136 putative regulatory RNAs (Mars et al., 2015). This includes an asRNA that is responsible for the ethanol stress-dependent repression of the essential *rpsD* gene, thereby contributing to the reduction in the number of ribosomes (Mars et al., 2015). Other bacteria respond to severe stress conditions in a similar way (Georg and Hess, 2018).

MATERIALS AND METHODS

1. Bacterial strains and DNA manipulations

Strains and plasmids used for the purpose of this Thesis are listed in **Table 1**. For the study, we used *B. subtilis* 168 trp+ (BaSysBio) as the model organism (Nicolas et al., 2012). Competent *E. coli* cells were prepared as described (Hanahan, 1983); competent *B. subtilis* cells were prepared as described (Dubnau and Davidoff-Abelson, 1971). In several experiments, *M. smegmatis* mc^2 155 was used as the model organism.

PCR was performed using the Expand High Fidelity PCR System (Roche). The list of primers is shown in **Table 2**.

Table 3. List of strains and plasmids

Strain/plasmid	Relevant characteristics [a]	Source or reference
B. subtilis		
LK#1432 (BSB1)	BaSysBio 168 wt trp+	(Nicolas et al., 2012)
LK#1550	Δ*sigI-rsgI::spc*[a], BaSysBio	This work
LK#1435 (MGNA-A781)	Δ*rsgI*::MLS[a]	(Kobayashi et al., 2003)
LK#1456	Δ*rsgI*::MLS, BaSysBio	This work
LK#1275 (MH5636)	Bsu RNAP *rpoC*-10×His	(Qi and Hulett, 1998)
LK#2181 (DK5247)	Δ*mreBH::kan*[a]	(Koo et al., 2017), gift from D. Kearns
LK#2191	Δ*mreBH::kan*, BaSysBio	This work
LK#1032	Bsu RNAP *rpoC*-10His, *rpoE*::kan, *helD*::MLS	(Wiedermannová et al., 2014)
LK#1277	Bsu RNAP *rpoC*-10His, *rpoE*::kan	
LK#782	Bsu RNAP *rpoC*-10His, *helD*::MLS	

E. coli		
LK#278	pET-22b, DH5α	This work
LK#475	pGEX-5X-3	This work
LK#805 (LK22)	pCD2/Bsu_*sigA*, BL21 (DE3)	(Chang and Doi, 1990)
LK#180 (pRLG770)	pRLG770, DH5α	(Ross et al., 1990)
LK#1177 (LK7558)	pRLG770 with P*veg* DH5α (-38/+1, +1G)	(Krásny and Gourse, 2004)
LK#1230	pRLG770 with P*gsiB* DH5α (-248/+11)	This work
LK#1231	pRLG770 with P*trxA* DH5α (-249/+11)	This work
LK#1232	pRLG770 with P*hag* DH5α (-182/+9)	This work
LK#1233	pRLG770 with P*motA* DH5α (-249/+11)	This work
LK#1234	pRLG770 with P*citG* DH5α (-249/+11)	This work
LK#1235	pRLG770 with P*spoVG* DH5α (-94/+11)	This work
LK#1366	pRLG770 with P*kinA* DH5α (-203/+57)	This work
LK#1236	pRLG770 with P*bcrC*, DH5α (-75/+11)	This work
LK#1230	pRLG770 with P*gsiB*, DH5α (-248/+11)	This work
LK#1238	pRLG770 with P*lytE*, DH5α (-248/+12)	This work
LK#1453	pRLG770 with P*mreBH*, DH5α (-239/+24)	This work
LK#1452	pRLG770 with P*sigI*, DH5α (-258/+127)	This work
LK#1766	pRLG770 with P*dhbA*, DH5α	This work
M. smegmatis		
LK#865	Laboratory strain mc^2 155 (wt)	
M. tuberculosis		
H37Rv ATCC 27294	ATCC; Strain derived from E. R. Baldwin's 1905 human-lung isolate H37 by W. Steenken in 1934	

[a] MLS – Macrolide-lincosamide-streptogramin B resistance; *kan* – kanamycin resistance; *spc* – spectinomycin resistance.

Table 4. List of primers

Primer number/name	Sequence 5'→3'
1164/sigI_For	GGAATTCCATATGGTGAAACCAGTGCTTAGCC
1166/sigI_Rev_His	CCGCTCGAGTGAGTGCAGCACCCCTTTAAG
1478/sigI_rsgI_LA_F	CGCGGATCCGCTTGCTTTATATACGCTTGC
1479/sigI_rsgI_LA_R	CCCAAGCTTCTCAGTTCCTCCCTATAACTA
1480/sigI_rsgI_RA_F	CGGTCGACAGACCTGAATTTATTTAGTTGTG
1481/sigI_rsgI_RA_R	AAATATGCGGCCGCCTGATGATGTCATCAGCCCG
1702/Pveg-50+10_6-FAM	CTTCAAGAATTCTATTTGACAAAAATGGGCTCGTGTTGTACAATAAATGTGTCTAAGCTT
1703/RC Pveg-50+10	AAGCTTAGACACATTTATTGTACAACACGAGCCCATTTTTGTCAAATAGAATTCTTGAAG
1758/PsigI-50+10_6-FAM	ACACGCATAAAACCCCCTTAATTCTTTAGAAAGGCACGAAATCATGTATAGAACGTCAGA
1759/PsigI-50+10_RC1758	TCTGACGTTCTATACATGATTTCGTGCCTTTCTAAAGAATTAAGGGGGTTTTATGCGTGT
1067/PgsiB_For	CCGGAATTCCAGATAGTGCCGGTTGCCG
1068/PgsiB_Rev	CCCAAGCTTAATTGGTGTTGGTTGTTGTATTC
1081/PlytE_For	CCGGAATTCAAAGTTTTTCATTTATTTCCTTTATG
1082/PlytE_Rev	CCCAAGCTTATTTTCCTCCCCAAATGTTAAC
1083/PbcrC_For	TTCGAATTCTTCAAGCGCCGTTATTTC
1084/PbcrC_Rev	CCCAAGCTTTACATTTTTATATTTAGTAGAC
1372/PmreBH_For	CCGGAATTCCCCTTCTTCCTTTAAATGTTTC
1373/PmreBH_Rev	CCCAAGCTTCTATCCTAATTTAATATGATTCTAC
1390/PsigI_For_2	CCGGAATTCTGGGGTGTCTTAGCAG
1391/PsigI_Rev_2	CCCAAGCTTACTGGTTTCACCTCAGTTC
1882/PybbA_F	CCGGAATTCCAAATAAAGAAAAAACCATCCTTC
1883/PybbA_R	CCCAAGCTTGCATTCTCCCCCTTTTTTTGC
1884/PfeuA_F	CCGGAATTCCTGACTGTATGATGCTTTTTC
1885/PfeuA_R	CCCAAGCTTCTATAGAGCCTCCTGTTCAATTG
1886/PycdA_F	CCGGAATTCCAAAAAGTAATAGTCTAAAATAC
1887/PycdA_R	CCCAAGCTTTCATTTTTCCTCCTCAAATGG
1888/PmalA_F	CCGGAATTCCGCACAATTGGATGTTTTATATA
1889/PmalA_R	CCCAAGCTTATGACGACCTCCTTGATAACG
1890/PdhbA_F	CCGGAATTCGAATTTTGCGAGTTTTCAGGAG
1891/PdhbA_R	CCCAAGCTTATCATCAATTCCTTTCTTCGCTC
1892/PbesA_F	CCGGAATTCCGAAAGAAAATCACCCTGGC
1893/PbesA_R	CCCAAGCTTGCAGTACGACATTCCTCTCC
1894/PsrfAA_F	CCGGAATTCCATTTAAACTGAACGGTAGAAAG
1895/PsrfAA_R	CCCAAGCTTATTGTCATACCTCCCCTAATC
1896/PxlyA_F	CCGGAATTCCTGCAGCATTAAAGGGGG
1897/PxlyA_R	CCCAAGCTTTTTCATCTCTCCTTATTTCGTC
2525/PykuN_F	CGGTTATGCCATTTTGCCG
2526/PykuN_R	CAATCGATATCCAACTCATATTC

2527/PodhA_F	GTATATTTGTCCCGGATGTTG
2528/PodhA_R	GTTCGGATCCTGGGTATACTG
2529/PyxeB_F	CTATTTCTAACCTTGATGCC
2530/PyxeB_R	GAAGCAGAATCGCTTGAGCTC
2531/PfabI_F	CTTTAACGGCCCAAACGGC
2532/PfabI_R	GTGCACCCGCTTCATGTAAAG
2533/PqueC_F	GCAAAAAAGCTCCTCGGCGG
2534/PqueC_R	GTCACCGTTTCGACTTCTTC
2537/PyxkC_F	CTATGTGTTTTCTCTCGGTC
2538/PyxkC_R	GCATAAAAAAGTTACGTTGATC
2539/PyfmB_F	CCTCTACATCTTTCATGAGC
2540/PyfmB_R	CGTGAATTCCTGGTGAACATC
2541/PxtmA_F	CAGAGATAGAATTCGAGAAGC
2542/PxtmA_R	GTTTTCGCGGAAACACCGATTG
2543/PcysH_F	GAATTGTGCTAAAATTTACTAC
2544/PcysH_R	CGTAATGGCCGTATGCCC
2557/PspoVD_F	CGGCAAAATCAAAATGCCTG
2558/PspoVD_R	GATCAGAAACACGATCACGC
2628/PfhuB_F	GTAATAGCCGTCTGCCATGAC
2629/PfhuB_R	CAATTGCTGTTTCAGTATATCAC
2630/PfhuD_F	CATAGACCTGCGATAAGGAC
2631/PfhuD_R	CGCAGGCTGCCAGCGCTG
2636/PyfiY_F	CTTGATGATGACATGCTCATTTG
2637/PyfiY_R	CTGAACTATTACAGGCAGAAAG
2638/PyhfQ_F	GAAGTATCATTGGAAGAAACGC
2639/PyhfQ_R	GCTTGAAGACGAGCAAGCAG
2507/PsigI_Native_F	CAAAAAACACGCATAAAACCCCCTTAATTCTTTAGAAAGGCACGAAATCATG
2508/PsigI_M-35_F	CAAAAAACACGCATAAAACttCCTTAATTCTTTAGAAAGGCACGAAATCATG
2509/PsigI_M_Ex-35_F	CAAAAAACACGCATAAAACCCCCTTggTTCTTTAGAAAGGCACGAAATCATG
2510/PsigI_2M-35_F	CAAAAAACACGCATAAAACttCCTTggTTCTTTAGAAAGGCACGAAATCATG
2511/PsigI_M-10_F	CAAAAAACACGCATAAAACCCCCTTAATTCTTTAGAAAGGCACGggATCATG
2512/PsigI_M_Ex-10_F	CAAAAAACACGCATAAAACCCCCTTAATTCTTTAGggAGGCACGAAATCATG
2513/PsigI_2M-10_F	CAAAAAACACGCATAAAACCCCCTTAATTCTTTAGggAGGCACGggATCATG
2524/PsigI_M-35/-10_F	CAAAAAACACGCATAAAACttCCTTAATTCTTTAGAAAGGCACGggATCATG
2688/PsigI2M_Ex-35/-10F	CAAAAAACACGCATAAAACCCCCTTggTTCTTTAGggAGGCACGAAATCATG
2705/PsigI_K-_UP-35	CAAAAAACACGCAggAAACCCCCTTAATTCTTTAGAAAGGCACGAAATCATG
2708/PsigI_K-DOWN-10B	CAAAAAACACGCATAAAACCCCCTTAATTCTTTAGAAAGGCACGAAATCATGTggAG
2514/PsigI_PCRpr_R	GGCTAAGCACTGGTTTCACC

Enzymatic restriction with restriction endonucleases (NdeI, XhoI, EcoRI, HindIII, AvrII) was performed for 2 h in 37° C. Reaction was either stopped for 20 min at 65° C (deactivation of heat-deactivated endonucleases), or desired construct was immediately separated on the 1 % agarose gel and purified from the gel using QIAquick Gel Extraction Kit (Qiagen).

Transformation of created constructs was performed into *B. subtilis* 168 trp+ competent cells. Cells were thawed on ice for 15 min, then 100 μl of cells were added to created construct diluted in 10 μl volume. Mix was incubated on ice for 30 min with further heat-shock for 90 sec in 42° C. Cells were incubated 5 min on ice, mixed with 1 ml of fresh LB medium and incubated for 60 min in 37° C. Cells were briefly spinned down, supernatant was discarded and pellet was spread into an agar plate without/with antibiotics.

The **Δ*sigI-rsgI*::spc knockout strain** was prepared via double cross-over. The upstream (primers 1478/sigI_rsgI_LA_F and 1479/sigI_rsgI_LA_R) and downstream (primers 1480/sigI_rsgI_RA_F and 1481/sigI_rsgI_RA_R) regions of the *sigI-rsgI* operon were cloned into the pGEX-5X-3 plasmid vector bearing inserted spectinomycin cassette (LK#475). The resulting plasmid (LK#1549) was transformed into BaSysBio competent cells (**Figure 13**). The final Δ*sigI-rsgI*::spc strain (LK#1550) was selected on spectinomycin plates. In the Δ*rsgI*::MLS strain (MGNA-A781, source - NBRP *B. subtilis*, Japan), the *rsgI* gene is disrupted and not expressed. The Δ*rsgI*::MLS BaSysBio strain (LK#1456) was obtained by transformation of gDNA from the purchased MGNA-A781 strain into BaSysBio competent cells.

For **overproduction of σ^I**, the *sigI* gene was amplified with primers 1164/sigI_For and 1166/sigI_Rev_His and cloned using NdeI and XhoI restriction enzymes into expression vector pET-22b comprising 6×His-tag. The resulting plasmid named pSigI-6×His (LK#1242) was transformed into BL21 (DE3) competent cells. All promoter regions of σ^A- and σ^I-dependent genes were amplified from genomic DNA of *B. subtilis* BaSysBio with primers 1067/PgsiB_For to 1897/PxlyA_R (**Table 2**) and cloned into pRLG770 plasmid using EcoRI and HindIII restriction enzymes. Purified supercoiled plasmids for multiple-round *in vitro* transcription assays were obtained using the Wizard Midipreps Plus DNA Purification System (Promega) with subsequent phenol-chloroform extraction and ethanol precipitation. The plasmids were dissolved in water.

For characterization of P*sigI* promoter sequence and promoter elements enhancing σ^I binding, we used PCR product of P*sigI* promoter region with inserted double substitutions and their combinations in different parts of the promoter region. Mutations were inserted using PCR primers with the changed sequence.

All constructs were verified by sequencing.

2. Media and growth conditions

For all experiments (with the exception of those testing iron requirement), pre-cultures were inoculated from single colonies from LB agar plates and grown overnight in LB at 37°C. Overnight pre-cultures were inoculated to fresh room temperature (RT) media to OD_{600}=0.03 and grown at 37°C and 52°C, respectively. Pre-cultures of Δ*sigI-rsgI*, Δ*rsgI*, and Δ*mreBH* strains were supplemented with spectinomycin 100 μg/ml, lincomycin 12.5 μg/ml + erythromycin 0.5 μg/ml, or kanamycin 10 μg/ml, respectively.

To monitor iron requirement, Δ*sigI-rsgI,* Δ*rsgI* and wt strains were inoculated from a single colony on LB agar plates to a MOPS medium containing/lacking iron and grown overnight at 37°C. Overnight pre-cultures were inoculated to a fresh MOPS medium to OD_{600}=0.03; cultures were grown at 37°C for 24 hours. Afterward, cultures were inoculated into a fresh MOPS medium (RT) at OD_{600}=0.03 and grown at 37°C and 52°C. Growth was monitored at OD_{600}. MOPS medium containing iron: 50 mM MOPS (3-(*N*-morpholino)propanesulfonic acid), 10 mM $(NH_4)_2SO_4$, 1 mM KH_2PO_4, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 5 μM $MnCl_2$, 0.5 μM $FeCl_3$, 50 μg/ml of each amino acid and 0.4% glucose. MOPS medium lacking iron: 50 mM MOPS, 10 mM $(NH_4)_2SO_4$, 1 mM KH_2PO_4, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 5 μM $MnCl_2$, 25 μM 2,2'-bipyridine, 50 μg/ml of each amino acid and 0.4% glucose. 2,2'-bipyridine was added to chelate iron.

M. smegmatis mc2 155 (wt, LK#865) were grown at 37°C in Middlebrook 7H9 medium with 0.2% glycerol and 0.05% Tween 80, and harvested in exponential (OD600 ~0.5; 6 hours of cultivation), early stationary (OD600 ~2.5–3, 24 hours of cultivation) or late stationary (OD600 ~2, 48 hours of cultivation) phase.

For *M. smegmatis* and *M. tuberculosis* cultivation see the Sub-chapters 10 and 11.

2.1 Growth of *B. subtilis* cells for RNA sequencing

Bacillus subtilis Δ*sigI-rsgI* and wt strains (LK#1550 and LK#1432, respectively) were inoculated from a single colony to 10 ml of LB medium. Cultures were grown overnight at 37°C. Next day, the cultures were inoculated into 20 ml of LB medium (RT) to OD_{600}~0.01 and grown at 37°C and 52°C, respectively. Cells were harvested in the exponential phase (OD_{600}~0.5). Three ml of the culture were immediately treated with a double volume of RNAprotect Bacteria reagent (QIAGEN) for 5 min at RT to prevent degradation of RNA. Cells were pelleted and frozen

immediately. In all steps of cultivation (with the exception of the last step of cultivation for total RNA isolation), the medium for Δ*sigI-rsgI* strain was supplemented with spectinomycin 100 μg/ml. The experiment was repeated three times.

2.2 Spot assays

Δ*sigI-rsgI,* Δ*rsgI*, and wt strains were grown in LB broth at 37° C to mid-logarithmic phase (OD_{600}=0.5). Serially diluted aliquots (1 μl; 10x dilutions between spots) were spotted on 1.5% LB agar plates lacking antibiotics and allowed to dry. Plates were incubated at 37°C and 52°C for 40 hours. Bacterial colonies were visualized using an SZX10 stereomicroscope (Olympus) and photographed using an Olympus E-600 digital camera.

3. PCR techniques

For **PCR reaction** Expand High Fidelity PCR system was used (Roche). The reaction was performed in 50 μl, and contained following components: 5x buffer with $MgCl_2$, 200 μM dNTP mix, 1 U/reaction DNA polymerase, 1 μl of each primer, template (1 μg of gDNA, 1-5 ng of plasmid). Following PCR program was used: 1 step 94° C for 2 min, 2 step 94 ° C for 15 sec, 3 step 60° C for 30 sec, 4 step 72° C for 3 min, 5 step 94° C for 15 sec, 6 step 56° C for 30 sec, 7 step 72° C for 3 min, steps 2-7 were repeated 5 times; 8 step 94° C for 15 min, 9 step 52° C for 30 sec, 10 step 72° C for 3 min, steps 8-10 were repeated 24 times. PCR products were separated on 1-2% agarose gel.

For **colony PCR** reaction Biotools DNA polymerase was used. The reaction was performed in 30 μl reaction and contained the following components: 10x buffer with $MgCl_2$, 100 μM dNTP mix, 0.27 U/reaction DNA polymerase, 0.15 μl of each primer. Instead of purified DNA, a small number of bacteria boiled for 5 min in water was added to the reaction on a sterile micropipette tip. Following program was used: 99 for 5 min, 95 for 30 sec, 42 for 45 sec, 72 for 1 min; steps 2-4 were repeated for 24 times. PCR products were separated on 1-2% agarose gel.

Real-time quantitative PCR (RT-qPCR) was performed using LightCycler 480 (Roche) and the LightCycler® 480 SYBR Green I Master mix (Roche). The master mix contains FastStart Taq DNA Polymerase and SYBR Green I dye for detection of DNA double-strand product. Each reaction was prepared from 2.5 µl of the master mix, 0.5 µl of primers (5 µM each) and 2 µl of the cDNA diluted 1:10. The following PCR program was used: step 1 – 95°C for 7 min; step 2 – 95°C for 20 sec, step 3 – 61°C for 20 sec, step 4 – 72°C for 35 sec, steps 2-4 were repeated for 45 times. PCR products were separated on 2% agarose gel.

4. Extraction of nucleic acids

Plasmid DNA extraction for *in vitro* transcription was performed from a pellet of 100 µl cell culture using Wizard Plus Midipreps DNA Purification System Kit (Promega). The sample obtained at the final step of purification (300 µl volume) was further extracted with phenol-chloroform or directly precipitated (see below in this Chapter), depending on the downstream application. When extracted with phenol-chloroform, the volume of the sample was increased to 600 µl, the sample was mixed 1:1 with phenol solution (pH=8.0, Sigma-Aldrich), mixed for 5 min and centrifuged 5 min, 15000x g, RT. The upper layer of the 2-phase solution was transferred to a fresh tube. Next, the sample was mixed with phenol and chloroform 2:1:1, mixed for 5 min and centrifuged 5 min, 15000x g, RT. The upper layer of the 2-phase solution was transferred to a fresh tube. As the last step, the sample was mixed with chloroform 1:1, mixed for 5 min and centrifuged 5 min, 15000x g, RT. Afterwards, the sample was precipitated with the 2.5x volume of 96% ethanol, 0.1x volume of 3M sodium acetate and 1 µl of glycogen for 1h at -20° C, then centrifuged 15 min 15000x g at 4° C. Pellet was dried and dissolved in the desired volume of ultrapure water.

Genomic DNA extraction was performed with the kits: either with Charge Switch gDNA Mini Bacteria Kit (Invitrogen) or with High Pure PCR Template Preparation Kit (Roche).

PCR products as templates for *in vitro* transcription (where indicated in the Results section) were purified with Agencourt AMPure XP PCR Purification system (Beckman Coulter) according to the manufacturer's protocol or were purified from the 1 % agarose gel using QIAquick Gel Extraction Kit (Qiagen). All constructs were verified by sequencing.

Extraction of total RNA with phenol-chloroform was performed from frozen cells pellet (from 15-25 ml of cell culture). The pellet was mixed with 300 µl of acidic phenol solution (pH=4.5, Sigma-Aldrich), 300 µl of chloroform, 240 µl of TE buffer and 60 µl of LETS buffer. The mixture was sonicated for 60 sec, then transferred to 2 ml tube. Ultrapure water was added up to 1.5 ml volume. The mixture was briefly vortexed, centrifuged for 5 min with maximum speed, RT. 700 µl of the supernatant was transferred to a fresh tube and mixed with phenol and chloroform 2:1:1. The procedure was repeated 2x, every time transferred volume was reduced by 100 µl. Extracted RNA was precipitated with the 2.5x volume of 96% ethanol, 0.1x volume of 3M sodium acetate and 1 µl of glycogen for 1h at -20° C, then centrifuged 15 min 15000x g at 4° C. Pellet was dried and dissolved in the desired volume of ultrapure water.

5. Samples preparation for RNA-seq

After treating cell culture with RNAprotect Bacteria reagent cells were frozen. In the following days, a protocol for samples preparation to RNA-seq was applied to frozen cell pellets. **RNA was extracted** using RNeasy Mini Kit 50 (QIAGEN) according to the manufacturer's protocol. Finally, RNA was treated twice with DNase (TURBO DNA-free Kit, Ambion).

RNA library construction was performed in the following way. Two micrograms of total RNA were rRNA-depleted with RiboMinus Transcriptome Isolation Kit, bacteria (Invitrogen). A strand-specific library was prepared for each sample with Illumina compatible NEXTflex Rapid Directional RNA-Seq Kit (Bioo Scientific) according to the manufacturer's protocol.

Transcriptome profiling (**library sequencing**) with RNA-seq was performed at the EMBL Genomics Core Facility (Heidelberg, Germany). Pooled barcoded library (four samples in biological triplicates) was sequenced in a single lane at Illumina HiSeq 2000 (50 bp single-end; ~8-16 million reads per sample regime). The quality of sequencing reads was checked with fastQC 0.11.2 (www.bioinformatics.babraham.ac.uk/projects/fastqc).

Reads were aligned to the *Bacillus subtilis* subsp. *subtilis* strain 168 genome (NCBI Nucleotide acc. no. NC_000964) using BWA 0.7.9a-r786 (Li and Durbin, 2009) and samtools 0.1.19 (Li et al., 2009). Alignment quality was checked using QualiMap 2.1.3 (Okonechnikov et

al., 2015) and IGV 2.3 (Thorvaldsdóttir et al., 2013). *B. subtilis* genome annotation was obtained from NCBI Assembly (acc. no. GCF_000009045.1). Analysis of differential gene expression was performed using R (www.r-project.org) and Bioconductor package DESeq2 (Huber W., Carey V.J., Gentleman R., Anders S., Carlson M., Carvalho B.S., Corrada Bravo H., Davis S., Gatto L., Girke T., Gottardo R., Hahne F., Hansen K.D., Irizarry R.A., Lawrence M., Love M.I., MacDonald J., Obenchain V., Ole A.K., Pages H., Reyes A., 2015; Love et al., 2014) at a 5% False Discovery Rate (FDR). Data from RNA-seq are stored in the ArrayExpress database (https://www.ebi.ac.uk/arrayexpress/) under accession No. ArrayExpress: E-MTAB-6314.

6. Electrophoretic Mobility Shift Assay (EMSA)

60-bp fluorescently labeled P*sigI* and P*veg* promoter fragments (-50/+10) were generated by annealing in a Taq PCR buffer (20mM Tris-HCl pH 8.4, 50mM KCl, 0.5 mM $MgCl_2$). Primers were: for P*sigI* 6-FAM (6-Carboxyfluorescein) labeled 1758/PsigI-50+10_6-FAM and 1759/PsigI-50+10_RC1758; for P*veg* 6-FAM-labeled 1702/Pveg-50+10_6-FAM and 1703/RC Pveg-50+10. Prior to EMSA, RNAP (storage buffer pH 8.0) was reconstituted with σ^A (storage buffer pH 8.0) or σ^I (storage buffer pH 7.3) for 10 min at 30°C (ratio 1:5). Next, 0.25 pmol of 6-FAM-labeled template was added to reconstituted RNAP and incubated for 15 min at 30°C. Samples were loaded onto a native bis-tris gel (4-16%) and electrophoresed on ice for 100 minutes in the buffer of pH 7.0 at 150 V. Dried gels were scanned with Molecular Imager FX (Bio-Rad) and were visualized and analyzed using the Quantity One software (Bio-Rad).

7. Electron Microscopy

/Performed by O. Benada and O. Kofronová at the Laboratory of Molecular Structure Characterization, Institute of Microbiology of the CAS/

Cells were grown to exponential phase (OD_{600}=0.5) at 37°C and 52°C and fixed in buffered 3% glutaraldehyde at 4°C. The extensively washed cells were then sedimented onto poly-L-lysine-coated glass coverslips at 4°C overnight. The coverslips were dehydrated through an alcohol series followed by absolute acetone and critical point-dried in a K850 Critical Point Dryer (Quorum Technologies Ltd, Ringmer, UK). The dried samples were sputter-coated with 3 nm of platinum in a Q150T Turbo-Pumped Sputter Coater (Quorum Technologies Ltd, Ringmer, UK). The final samples were examined in an FEI Nova NanoSEM scanning electron microscope (FEI, Brno, Czech Republic) at 5 kV using CBS and TLD detectors. The cells were imaged according to a protocol described previously (Seydlová et al., 2017).

8. Protein purification

B. subtilis RNAP with a 10×His-tagged β' subunit was purified from the *E. coli* MH5636 strain by affinity chromatography using Ni-NTA agarose as described (Qi and Hulett, 1998). σ^B, σ^M, δ, δN, and HelD proteins were **purified from the supernatant** was performed using **C-terminal 6×His-tag** sequence. Gene of anticipated protein encoding σ factor was cloned using NdeI and XhoI restriction enzymes into expression vector pET-22b under T7 promoter induced by IPTG. Proteins were purified as described from *B. subtilis* expression strain BL21 (DE3)(Qi and Hulett, 1998).

σ^I protein purification was performed from the BL21/pSigI-6×His, BL21 (DE3) strain (LK#1242) as described previously (Qi and Hulett, 1998) with some modifications. Briefly, induction of σ^I was carried out at OD_{600}=0.6 with 0.05 mM IPTG for 3 hrs at RT. Cells were harvested by centrifugation and the pellet was resuspended in P-buffer (30 mM NaCl, 50 mM Na_2HPO_4, 3 mM 2-mercaptoethanol, 1 mM PMSF, pH 7.3) to avoid subsequent precipitation of the protein. The cells were then disrupted by sonication 20x 10 sec on ice. The protein was purified

from the soluble fraction by affinity chromatography using Ni-NTA agarose (Qiagen). After elution with 400 mM imidazole in P-buffer, fractions containing σ^I were pooled and dialyzed against storage buffer.

The σ^A, σ^D and σ^H subunits of RNAP were **purified from the inclusion bodies** (via denaturation with Guanidine-HCl and subsequent renaturation) using weak anion exchanger DE52 cellulose. σ^A was overexpressed from the pCD2 plasmid (Chang and Doi, 1990) (LK22) and purified as described (Juang and Helmann, 1994a) except FPLC step. σ^D and σ^H subunits were cloned into expression vector pET-22b and purified as described (Juang and Helmann, 1994a) except FPLC step.

All purified proteins were dialyzed against storage buffer containing 50 mM Tris-HCl pH 8.0 (with exception of σ^I, pH 7.3), 100 mM NaCl, 50% glycerol, 3 mM 2-mercaptoethanol and stored at -20°C. Proteins were visualized in NuPAGE 4-12% Bis-Tris gels (Invitrogen) with Novex Sharp Pre-Stained Protein Standard as a marker (Invitrogen).

Prior to each protein purification, its **trial overexpression** was performed. Cell culture was grown in 20 ml of LB medium at 37° C under constant shaking (180 rpm), OD_{600} was measured. At OD=0.6-0.8 culture was split into two parts, flasks were transferred to RT, the inductor of protein expression (0.8 mM IPTG) was added to one of them. Control samples (2 ml volume) were taken from each flask. Cells were further cultivated at 37° C for 3 hrs. After 3 hrs final OD_{600} was measured (commonly OD was around 1.5-2.0; cells were at stationary phase). 1 ml sample of cell culture was taken from induced and non-induced cultures and centrifuged 3800x g at 4° C. Obtained pellets were mixed with 200 µl of lysis buffer. Cells were sonicated 15x 10 sec. Lysates were centrifuged 17000x g at 4° C. 180 µl of the supernatant was separated from the pellet, the pellet was dissolved in 180 µl of Lysis buffer. 10 µl samples of pellet and supernatant with/without induction were loaded onto denatured SDS PAGE. 1D protein profiles with/without IPTG induction were compared. Protein concentration was always measured using Bradford protein assay.

9. *In vitro* transcription assays

The *B. subtilis* RNA polymerase core was reconstituted with saturating concentration of σ factor (ratio 1:5) in storage buffer (50 mM Tris-HCl, pH 8.0, 0.1 M NaCl, 50% glycerol) for 10 min at 37° C. Multiple round transcription reactions were carried out in 10-µl reaction volumes with 30 nM RNAP holoenzyme and 100 ng of supercoiled plasmid DNA templates containing specific promoters or 100 ng of linear PCR-product templates containing promoter regions. The transcription buffer contained 40 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 0.1 mg/ml BSA (bovine serum albumin) and 60 mM NaCl. ATP, CTP, and GTP were 200 µM, and UTP was 10 µM plus 0.33 µM radiolabeled [α-^{32}P] UTP. All transcription experiments were done at 37°C. Transcription reaction was started with RNAP and allowed to proceed for 15 min. Transcription was stopped with equal volumes (10 µl) of formamide stop solution (95% formamide, 20 mM EDTA pH 8.0). Samples were loaded onto 7 M urea-7% polyacrylamide gels and electrophoresed. Dried gels were scanned with Molecular Imager FX (Bio-Rad) and visualized and analyzed using the Quantity One software (Bio-Rad). The strong constitutive σ^A-dependent P*veg* promoter cloned into pRLG770 was used as a control. The K_{NTP} value was calculated from the f=a×[1–exp(–b×x)] equation, where f is relative transcription; x – time; a and b – constants.

10. β-galactosidase assays

Annealed oligonucleotides were cloned via the ScaI site preceding the *lacZ* reporter gene in the pSM128 integrative vector. Sequence verified constructs were transformed into *M. smegmatis* mc^2 155 (LK#865). 1 ml of bacterial culture (see growth conditions) was centrifuged (13200x g, 10 min, 4°C) and pellets washed with 500 µl of Z-buffer 2 (Z-buffer with 2.7 µl of 2-mercaptoethanol/ml; Z-buffer: 60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, pH 7.0), centrifuged again and resuspended in 500 µl of Z-buffer 2. Cells were sonicated 3x 20 s (amplitude 50 %) on ice with 1 min pauses between sonications and centrifuged (17900x g, 10 minutes, 4°C). 100 µl of sonicate and 900 µl of Z-buffer were incubated for 5 min at 30ºC and 200 µl of ortho-nitrophenyl-β-D-galactopyranoside (ONPG) solution was added (4 mg/ml ONPG in

Z-buffer). After 20 min, the reaction was stopped with 500 µl of 1M Na_2CO_3, OD_{420} and OD_{550} were measured, and β-galactosidase activity (in arbitrary units) calculated: activity = $1000*(OD_{420}-1.75*OD_{550})/(v*t*c)$, where v is sample volume [ml], t time of reaction, c protein concentration [mg/ml] measured by Bradford protein assay.

11. 5' RACE

M. tuberculosis H37Rv ATCC 27294 strain was cultivated for three weeks on Lowenstein-Jensen agar plates at 35.5°C in aerobic conditions. The cells were harvested and directly lysed in TRI Reagent (Sigma Aldrich) with subsequent RNA isolation according to manufacturer's instructions. Five µg of total RNA was treated with 5U TEX (Terminator 5'-Phosphate-Dependent Exonuclease; Epicentre) for 1 h at 37°C. After the reaction, RNA was extracted with TRIzol and precipitated with ethanol. Purified RNA was treated with 1U of Tobacco Acid Pyrophosphatase (TAP; Epicentre) for 1 h at 37°C, extracted with TRIzol again, precipitated with ethanol, and a 5'-adaptor DNA/RNA oligonucleotide (5'-ATCGTaggcaccugaaa-3', DNA in upper case letters) was ligated to the 5' ends (1 h at 37°C). RNA was then extracted and reversely transcribed into cDNA (SuperScriptIII, Invitrogen) with an MTS2823 (Ms1 homolog) specific reverse primer (5'-CATCTGCTGTTCGCAATTAC-3'). The same reverse primer and the 5'-ATCGTAGGCACCTGAAA-3' forward primer were used for PCR with Taq DNA polymerase (Biotools). The PCR products were sequenced and mapped to the *M. tuberculosis* H37Rv (GenBank #AL123456.2).

12. *In silico* procedures

12.1 Sequence logo creation

The promoter sequence logos were created using WebLogo 3 tool available online (http://weblogo.threeplusone.com/create.cgi).

12.2 Phylogenetic tree creation

Selection of σ^I homologs was based on the BLAST search (https://blast.ncbi.nlm.nih.gov/). The selected sequences were aligned using MUSCLE 3.8.31 (Edgar, 2004) with default parameters (Waterhouse et al., 2009). The phylogenetic tree was inferred using RAxML BlackBox webserver (Stamatakis et al., 2008), settings for protein sequences and maximum likelihood search were chosen, and other settings were left to default (JTT substitution matrix).

12.3 Data acquisition for *in silico* kinetic modelling

We downloaded the *B. subtilis* transcriptomic microarray data from 14 time points (0, 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90 and 100 min) obtained during germination and outgrowth as previously reported (Keijser et al., 2007) from GEO http://d360prx.biomed.cas.cz:2259/geo/query/acc.cgi? acc=GSE6865. Briefly, the generation of *Bacillus subtilis* 168 spores was induced by the depletion of defined MOPS medium during 4 days of shaking at 37°C. Subsequently, spores were activated in the germination medium by thermal treatment at 70°C for 30 min. The release of dipicolinic acid in the medium during spore germination was monitored using the terbium fluorescence assay. During germination and outgrowth, samples for RNA isolation were drawn at regular intervals. RNA was isolated from spores and outgrowing spores and then reverse transcribed to cDNA. Cy-labeled cDNA was made by the direct incorporation of Cy-labelled dUTP. Samples were hybridized to microarray slides, and microarrays were scanned using an Agilent G2505 scanner. Data were averaged over repeated samples. For further processing, the original log2-based data were exponentiated.

Regulons of σ factors were downloaded from databases available on-line. The σ^A regulon genes were downloaded from SubtiWiki (Michna et al., 2016) and DBTBS (Sierro et al., 2008). The database contains a collection of experimentally validated gene regulatory relations of *B. subtilis* genes constructed by surveying literature references. Among the σ^A target genes (850), eight other σ^A factors were found (σ^D, σ^H, σ^M, σ^X, YlaC, σ^E, σ^F and σ^G). σ^E, σ^F, σ^G and YlaC were excluded from the analysis because their expression profiles were too low and could, therefore, be subject to high experimental variance, which could lead to misinterpretations of the modelling results. For the other alternative σ factors, their regulons were downloaded from SubtiWiki — σ^D (73 genes), σ^H (48 genes), σ^M (84 genes) and σ^X (31 genes). Some genes were members of more than one regulon. Altogether, a list of 1087 genes was compiled. The dataset contains time series of 4008 genes.

12.4 Kinetic model of gene expression

/Created by J. Vohradský and colleagues/

A kinetic model of gene expression controlled by a σ factor that was originally developed by J. Vohradský (Vohradský, 2001) and further revised and extended (To and Vohradsky, 2010; Vu and Vohradsky, 2007, 2009) was used. The model was derived from the assumption that the mRNA level of a gene controlled by a σ factor is determined by the concentration of the σ factor binding in complex with RNA polymerase to the promoter region. The probability of the σ factor binding to the gene promoter is determined by the σ factors' binding strength and the number of molecules around the promoter. Transcription is a discontinuous process that depends on the actual binding of the holoenzyme to the promoter. When the number of σ factors molecules is low, the probability of triggering transcription of a given gene is also low. With increasing amounts of σ factor molecules, the probability of a gene transcription event increases until the promoter is saturated and the expression rate becomes constant. The relation between the accumulation of transcribed mRNA and σ factor concentration can thus be described mathematically by a sigmoid with parameters reflecting the strength of binding, reaction delay and mRNA degradation rate. The sigmoidal shape of the function was also confirmed by the results of stochastic simulations (*e.g.*, (Roussel and Zhu, 2006; Vohradsky, 2012)). The model used in this study has the following form:

$$\frac{dy_i}{dt} = \frac{k_{1i}}{1 + \exp[-w_i R(t + \Delta t)_j + b_i]} - k_{2i} y_i \qquad (1)$$

where y_i represents the concentration of the genes mRNA and Rj is the concentration of the j-th σ factor modulated by parameter wi, which corresponds to binding strength to the promoter. The bi and Δt parameters correspond to the reaction delay. The accumulation of the mRNA of the gene i is diminished by the degradation described by the term $k_{2i}y_i$.

Since the expression data were noisy, data were smoothed prior to computation with a piecewise cubic spline with 6. After smoothing, the results were more robust with respect to the low-frequency phenomena expected in gene expression data. A further advantage of smoothing is that it lets us subsample the fitted curve at arbitrary resolution. We subsampled the profiles at 1-minute time steps, which allowed us to integrate (Equation 1) accurately with a computationally cheap Euler method. The parameters of the model for individual σ factor-transcribed gene combinations were optimized using a simulated annealing scheme by minimizing an objective function

$$E = \sqrt{\sum (\mathbf{y} - \tilde{\mathbf{y}})^2} \qquad (2)$$

where y represents the measured mRNA concentration time series proportional value and $\tilde{y}$ represents the time series computed using the model Equation 1. For each profile, optimization was repeated 100 times with random values as estimates of the initial parameters, and those parameters that gave the smallest E were selected from the 100 runs. The expression values of the σ factor-transcribed genes from the 14 time points were provided. This data set was subsequently analyzed. The goal was to identify parameters that would give the best fit of the model to the actual profile of a given regulated gene with the σ^A (or other considered σ factor) profile as the regulator. The regulatory interaction between a σ factor and a gene was accepted; *i.e.*, the control of the transcribed gene by the given σ factor was considered possible, if the profile $\tilde{y}$ computed with the best set of parameters was within the confidence interval of the measured profile (y) in at least 12 measured time points of the profile. This constraint was chosen to minimize the influence of the first and last time points that have the highest experimental and spline fitting errors. When the confidence interval could not be determined from the experimental data, a flat value of 20% of profiles maximum was used as the confidence interval.

RESULTS

This Thesis contains published and unpublished data from five projects I participated in during my Ph.D. study. All these projects are focused on the regulation of bacterial transcription initiation, the main emphasis is on σ factors from *B. subtilis*.

To make the following of the Results section easier, each Chapter is accompanied by its own Discussion; Chapter Summary and Future Prospects finalize the Results section.

1. *In vitro* transcription system with selected alternative σ factors from *B. subtilis* and the impact of [iNTP] on *in vitro* transcription with these factors

Chapter 1 describes two related sub-projects following **the aims: (i) establishment of *B. subtilis in vitro* transcription system with alternative σ factors**, and, using this system, **(ii) characterization of an important mechanism of bacterial transcription regulation – by [iNTP] at alternative σ factors-dependent promoters.**

The objectives of this project were:

- Purification of selected σ factors – σ^B, σ^D, σ^H, σ^I;
- Establishment of *in vitro* transcription systems with purified σ factors and selected templates of these alternative σ factors;
- Analysis of the [iNTP] effect on *in vitro* transcription with purified alternative σ factors.

1.1 Establishment of *in vitro* transcription systems with alternative σ factors

To set up the *in vitro* system, we purified selected alternative σ factors: σ^B, σ^D, σ^H, σ^M, σ^I (**Figure 11**). The *in vitro* system with *B. subtilis* primary σ factor (σ^A) had been established in our laboratory previously. The *B. subtilis* σ^A is overexpressed in and from *B. subtilis* cells, as well as RNAP core. Transcription assays with σ^A served as a control in our experiments.

Alternative σ factors were selected to cover different kinds of stress responses in the cell: σ^B is the main alternative σ factor with multiple roles during stress response (Hecker et al., 2007); σ^D is involved in motility and chemotaxis (Jaehning et al., 1979; Mirel and Chamberlin, 1989); σ^H is an early sporulation σ factor that regulates transition into stationary phase (Britton et al., 2002; Johnson et al., 1983); σ^I is involved in heat stress response (Zuber, Drzewiecki and Hecker, 2001). Genes of corresponding σ proteins were cloned using expression vector pET-22b. Purification processes for each protein were optimized. σ^B and σ^I were purified using a 6×His-tag, while σ^D and σ^H were purified from inclusion bodies (**Figure 11**).

Along with reconstituted RNAP holoenzyme [Eσ], we prepared two DNA promoter templates for each σ factor: P*gsiB* and P*trxA* for σ^B, P*hag,* and P*motA* for σ^D, P*kinA* and P*spoVG* for σ^H, P*sigI* and P*mreBH* for σ^I. Transcription start sites (TSS) in these promoters were +1G or +1A (**Table 1**). Promoter regions were cloned into p770 plasmid; then each modified plasmid bearing a cloned promoter was purified from *E. coli* cells DH5α in the supercoiled form. As controls we used two σ^A-dependent promoters: strong constitutive P*veg* (insensitive to [iNTP]) and ribosomal P*rrnB* P1 (sensitive to [iNTP]). TSS of σ^A-dependent promoters is either +1A or +1G. In the cell, TSS for *veg* is +1A, but in this work we used P*veg* with +1G. This does not affect the results of performed experiments.

All proteins functioned in initial transcription experiments, although transcription signals with σ^M were weak, and we suspended assays with σ^M.

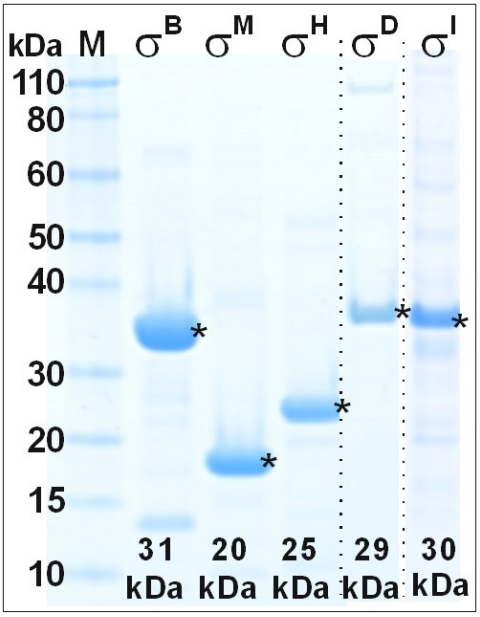


Figure 11. SDS-PAGE of purified recombinant σ factors. Proteins of interest are marked with an asterisk. The molecular weight of each protein (kDa) is indicated at the bottom of appropriate lanes. Protein marker Novex™ Sharp Pre-stain Standard (Invitrogen) is shown on the left side of the gel. Vertical dotted lines indicate non-adjacent lanes containing σ^B and σ^I from other gels electronically positioned to fit marker in the gel with σ^B, $\sigma^{M,}$ and σ^H.

1.2 [iNTP] affects *in vitro* transcription with σ^B, σ^D, σ^H, σ^I alternative σ factors

Having *in vitro* transcription system with alternative σ factors, we characterized the ability of these promoters to be regulated by [iNTP]. We wanted to reveal if the mode of regulation of promoters transcribed in the stationary phase via alternative σ factors is similar to that of σ^A-dependent promoters transcribed in the exponential phase.

In the experiments, we used the *in vitro* system with σ^A and newly established *in vitro* systems with alternative σ factors. We performed series of transcription experiments with σ^A, σ^B, σ^D, σ^H and σ^I and calculated $K_{[iNTP]}$ for each promoter. $K_{[NTP]}$ is a constant characterizing iNTP requirement of a promoter – is a half of the iNTP concentration needed for the maximum level of transcription under certain conditions with the certain promoter. Previously, for the σ^A-dependent sensitive promoter *rrnB* P1 $K_{[iNTP]}$ was defined as 172 μM, for insensitive P*veg* promoter – 18 μM (Sojka et al., 2011). In our experiments $K_{[iNTP]}$ for P*rrnB* P1 is 136 μM, for P*veg* – 36 μM. From this moment and further in this Thesis, we considered promoter with $K_{[iNTP]}$ lower than 100 μM as insensitive, and promoter with $K_{[iNTP]}$ higher than 100 μM as sensitive to [iNTP].

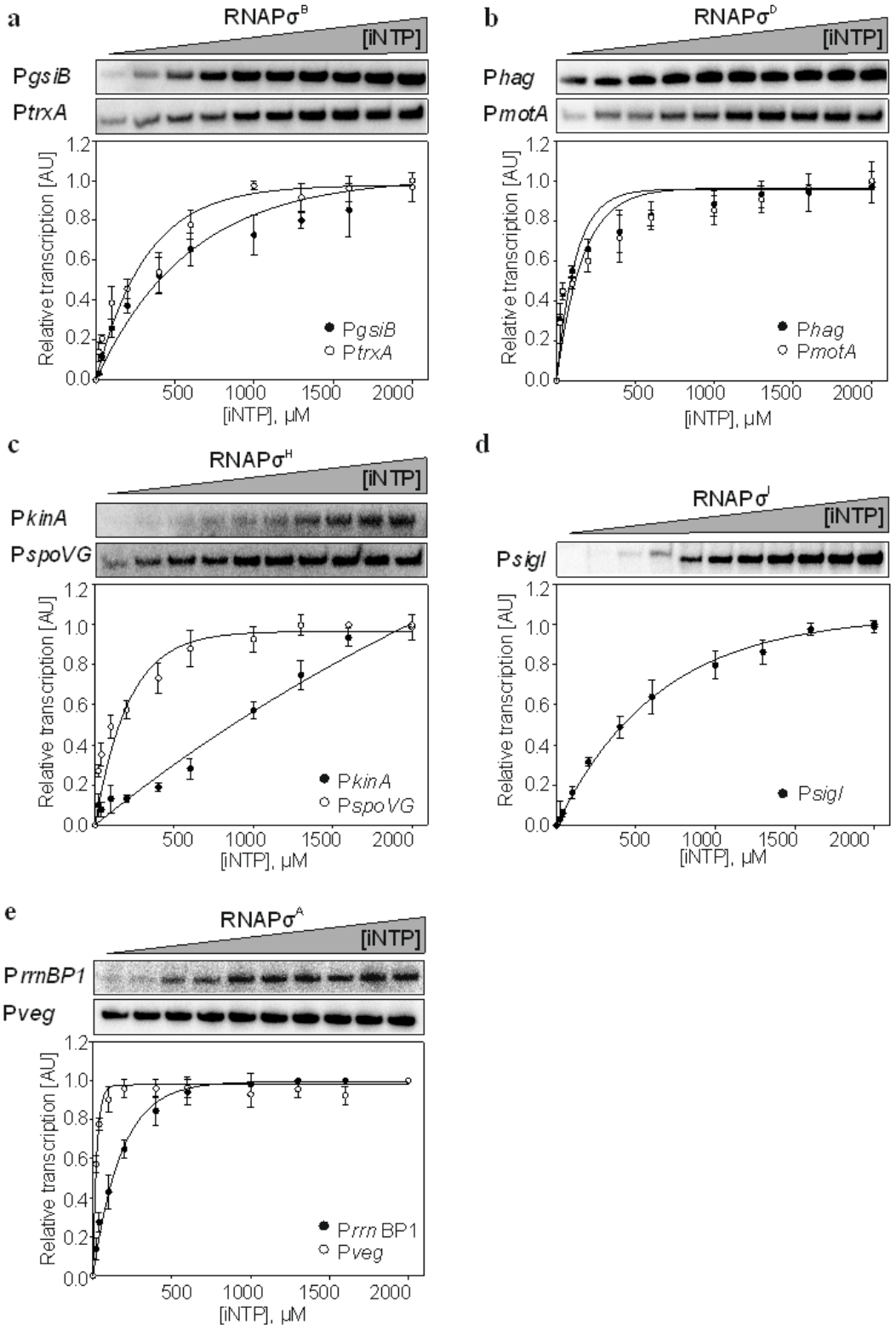


Figure 12. Multiple-round *in vitro* transcription assays with main (σ^A) and alternative (σ^B, σ^D, σ^H, σ^I) σ factors. a) Titration by [iNTP] on σ^B-dependent P*gsiB* (filled circles) and P*trxA* (empty circles) promoters; b) Titration by [iNTP] on σ^D-dependent P*hag* (filled circles) and P*motA* (empty circles) promoters; c) Titration by [iNTP] on σ^H-dependent P*kinA* (filled circles) and P*spoVG* (empty circles) promoters; d) Titration by [iNTP] on σ^I-dependent P*sigI* promoter; e) Control assays: titration by [iNTP] on σ^A-dependent P*rrnB* P1 (filled circles) and Pveg (empty circles) promoters. Values obtained in e) were used as a reference point for sensitive (P*rrnB* P1, filled circles) and insensitive (P*veg*, empty circles) promoters. The error bars represent ±SD of the mean.

After setting up an *in vitro* system, we tested the effect of [iNTP] on promoters selected in Chapter 1.1 (**Figure 2**). The identity of +1 position of these promoters is shown in **Table 1.1**.

Obtained results (**Figure 12**) show that all tested promoters driven by alternative σ factors are sensitive to [iNTP] *in vitro*, similar to σ^A-dependent ribosomal promoters. These results suggest that this type of regulation may be involved in coping with different stresses or nutrition starvation when bacterial cell overexpress alternative σ factors to react quickly to different stresses.

Table 5. $K_{[iNTP]}$ of selected promoters. $K_{[iNTP]}$±SD was calculated for each tested promoter.

σ factor	Tested promoter	iNTP	$K_{[iNTP]}$ [μM] ±SD
σ^B	P*gsiB*	ATP	426±52.39
	P*trxA*	ATP	249±41.29
σ^D	P*hag*	GTP	113±9.68
	P*motA*	ATP	150±31.36
σ^H	P*kinA*	ATP	786±42.31
	P*spoVG*	ATP	167±35.57
σ^I	P*sigI*	GTP	405±32.46
σ^A	P*veg*	GTP	36±13.43
	P*rrnB* P1	GTP	137±27.69

Discussion 1

The main focus of this Thesis is *B. subtilis* σ factors – essential proteins for bacterial transcription initiation. Until now, the vast majority of research in the bacterial transcription area had been performed with primary σ factors that are most highly expressed in the exponential phase of growth. The important role of alternative σ factors is to change the gene expression under stress conditions, thereby helping the cell to survive.

The sequence of a bacterial promoter determines the concentration of the iNTP required for maximal transcription efficiency, thus the [iNTP] in the cell is important for the regulation of gene expression (Gaal et al., 1997b). In our lab, we study this phenomenon using *in vitro* assays, which

belong among widely used research tools. Our *in vitro* systems comprise purified RNAP, alternative σ factors and DNA templates from *B. subtilis*. Previously, a similar system has been documented for several promoters transcribed by primary σ factor – σ^{70} of Gram-negative *E. coli* (Gaal et al., 1997b; Liu and Turnbough, 1994; Liu et al., 1994; Walker et al., 2004). In *B. subtilis*, the *in vitro* system was set up only for σ^{A}-dependent promoters (Krásny and Gourse, 2004; Krásný et al., 2008; Rabatinová et al., 2013; Sojka et al., 2011), but never for *B. subtilis* promoters transcribed with alternative σ factors. It was shown that in general *B. subtilis* RNAP forms less stable complexes with σ^{A}-dependent promoters DNA than *E. coli* RNAP (Ishikawa et al., 2010; Whipple FW, 1992). The main reason is difference in the structures of the two enzymes. Therefore, a high percentage of *B. subtilis* promoters can be modulated by changes in the intracellular concentration of [iNTP] and this type of regulation is important for the *B. subtilis* responses to environmental changes (Sojka et al., 2011; Tojo et al., 2013; Turnbough, 2008).

Here, using newly set up *in vitro* systems with stress-inductive alternative σ factors σ^{B}, σ^{D}, σ^{H}, and σ^{I}, we have analyzed the effect of [iNTP] on transcription with promoters expressed in the stationary phase or under stress conditions. We tested 7 non-ribosomal *B. subtilis* promoters, and all of them appeared to be sensitive to [iNTP] when compared to the control σ^{A}-dependent P*veg*/P*rrnB* P1 promoters (**Figure 12**). An interesting point is that for some promoters (P*gsiB*, P*sigI,* and P*kinA*) the $K_{[iNTP]}$ is extremely high (**Table 5**). Clearly, sensitive promoters display a different level of sensitivity. Originally, we divided tested promoters into two groups – sensitive and insensitive to [iNTP] (Gaal et al., 1997), and consider promoters with $K_{[iNTP]} \leq 100$ μM as insensitive, promoters with $K_{[iNTP]} \geq 100$ μM as sensitive to the [iNTP], without further subdividing. The obtained observation is consistent with the assumption that promoters transcribed with alternative σ factors (despite their iNTP identity) are mainly sensitive to the [iNTP]. This mechanism may contribute to the ability of the bacterial cell to quickly change gene expression, thus control the energy state of the cell during adaptation to environmental changes.

With some modifications, our *in vitro* transcription systems can be used to test the direct effect of different transcription factors on transcription. However, like any artificial system, it has some limitations. It contains only essential components for its functioning (RNAP, σ factor, DNA template, nucleotides, buffer) and lacks factors stimulating or suppressing transcription *in vivo*. Also, the scale of conditions generated at *in vitro* system is limited comparing to those in a living cell. Nevertheless, the obtained results characterize one of the important mechanisms of transcription regulation with [iNTP] at the stationary phase of growth or under stress conditions.

Importantly, the transcription system set up in this project was used (with further adjustments) in the following projects I participated in during my Ph.D. study (Results, Chapters 2-4).

2. σ^I from *Bacillus subtilis*

Chapter 2 represents my main Ph.D. project that was focused on one of the least explored *B. subtilis* alternative σ factors – σ^I. **The goal of this project was a comprehensive characterization of the transcription factor σ^I from *B. subtilis*.**

The aims of this project were:

- Characterization of the effect of σ^I on *B. subtilis* gene expression using RNA-seq approach;
- Definition of prospective σ^I new roles in *B. subtilis*;
- Characterization of σ^I *in vitro* properties, including transcription initiation and DNA binding ability;
- Determination of promoter elements important for promoter recognition by σ^I.

The obtained results were published as my first-author paper titled "σ^I from *Bacillus subtilis*: Impact on Gene Expression and Characterization of σ^I-dependent Transcription that Requires New Types of Promoters with Extended -35 and -10 Elements" in the Journal of Bacteriology (June 2018) [Appendix 1].

2.1 Creation of the Δ*sigI-rsgI* knock out strain

We used *B. subtilis* 168 trp+ (BaSysBio) as the genetic background to create Δ*sigI-rsgI* and Δ*rsgI* strains. In the *B. subtilis* genome the gene encoding the σ^I protein (*sigI*) is located in the same operon with the gene encoding its anti-σ^I factor (*rsgI*). In order to create a Δ*sigI-rsgI* strain, the *sigI-rsgI* operon was replaced with a spectinomycin resistance cassette using integrative plasmid pGex-5x-3 (**Figure 13**). In our experimental assays, we compared properties of the *sigI* knock-out strain with those of the wild-type *B. subtilis* 168 (BaSysBio) strain and a strain with deleted *rsgI*. The last strain was purchased from National BioResource Project for *Bacillus subtilis* (Japan).

Subsequently, genomic DNA from the purchased strain was transformed into *B. subtilis* 168 (BaSysBio) strain (for details see Materials and Methods).

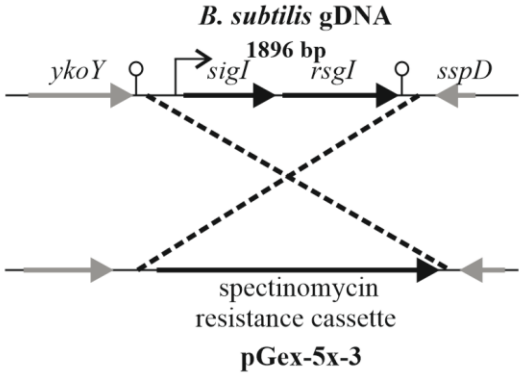


Figure 13. Construction of the Δ*sigI-rsgI* strain. In the *B. subtilis* 168 trp+ chromosome, the *sigI-rsgI* operon (1896 bp long) was replaced with the spectinomycin resistance cassette via a double crossover process. Mutants were selected on the agar plates supplemented with spectinomycin.

2.2 σ^I is important for growth at elevated temperature

As the first step of our study, we tested for the thermo-sensitive phenotype of the newly created strains. Under standard conditions (aerobic cultivation in LB broth at 37°C), the Δ*sigI-rsgI* and Δ*rsgI* strains grew indistinguishably from the wt strain (**Figures 14**, **15a**).

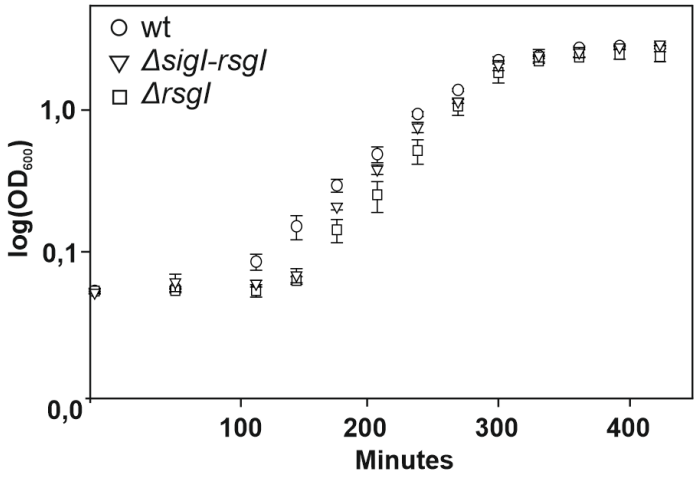


Figure 14. Growth curves (log scale) of the wt (circles), Δ*sigI-rsgI* (triangles) and Δ*rsgI* (squares) strains [#1432, #1550 and #1456 strains, respectively], grown in LB at 37°C.

Subsequently, we determined whether the Δ*sigI-rsgI* strain was sensitive to elevated temperature. We grew the strains in LB at 37°C, spotted serial dilutions of cell suspensions on agar plates, and let them incubate at 37°C and 52°C for 40 hrs (**Figure 15**). All three strains displayed the same pattern of growth when cultivated at 37°C. In contrast, the Δ*sigI-rsgI* strain

displayed impaired growth at 52°C, while wt and Δ*rsgI* strains cultivated at this temperature had the same pattern of growth as at 37°C. Thus, σ^I was important for efficient growth when the cells were chronically exposed to elevated temperature.

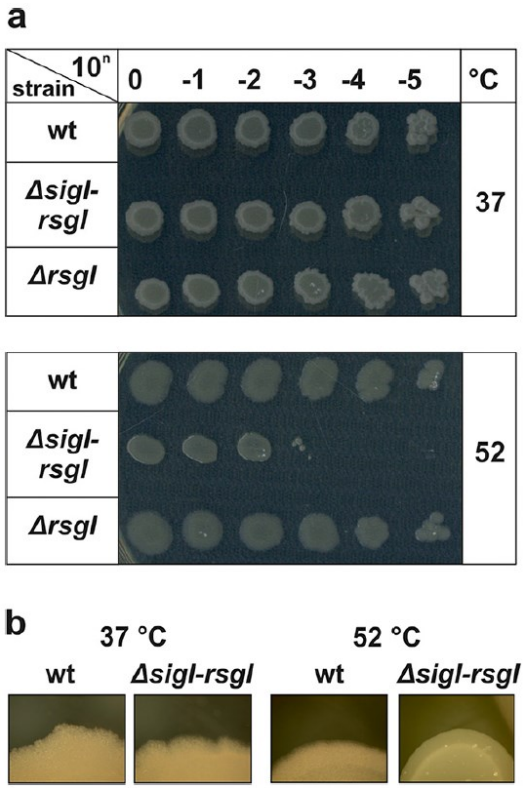


Figure 15. Spot assays of the *B. subtilis* wt, Δ*sigI-rsgI*, and Δ*rsgI* strains on agar plates at 37°C and 52°C. (a) Serial dilutions of mid-logarithmic phase cultures (OD_{600}~0.45) of wt, Δ*sigI-rsgI* and Δ*rsgI* strains were spotted on LB agar plates and incubated at 37°C and 52°C for 40 hours. The experiment was repeated three times with the same result. (b) Colony morphology of wt and Δ*sigI-rsgI* strains.

We hypothesized that the phenotype we observed for Δ*sigI-rsgI* strain at 52°C could be a consequence of a misregulated *mreBH* gene (MreBH is a cell shape-determining protein), previously shown to be under σ^I control. To test this hypothesis we performed the same spot assay as in **Figure 15** with a Δ*mreBH* strain. The results showed that the absence of the *mreBH* gene did not result in impaired growth at any temperature tested, and indicated a more complex cause of the observed phenotype (**Figure 16**).

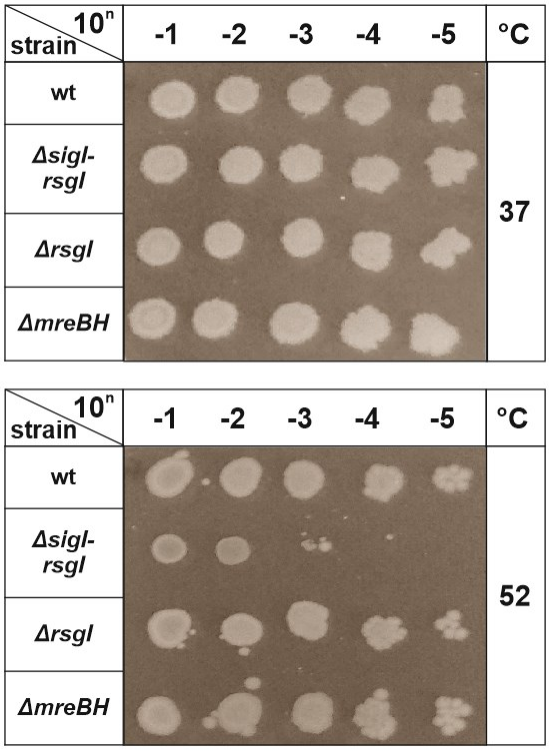


Figure 16. Spot assays of the *B. subtilis* wt, Δ*sigI-rsgI*, Δ*rsgI* and Δ*mreBH* strains on agar plates at 37°C and 52°C. Serial dilutions of mid-logarithmic phase cultures (OD_{600}~0.45) of wt, Δ*sigI-rsgI*, Δ*rsgI,* and Δ*mreBH* strains were spotted on LB agar plates and incubated at 37°C and 52°C for 40 hours. The experiment was repeated three times with the same result.

2.3 Absence of σ^I affects cell morphology during heat stress

During the spot assays described above, we noticed that the Δ*sigI-rsgI* colonies grown at 52°C displayed different morphology in comparison to the wt and Δ*rsgI* strains (**Figure 15b**). Hence, we looked at the cells in close detail using Scanning Electron Microscopy (SEM) of mid-logarithmic phase Δ*sigI-rsgI,* Δ*rsgI*, and wt strains cultivated at 37°C and 52°C (**Figure 17**). The results correlated with the spot assays: cells of all three strains grown at 37°C had the typical rod shape of *B. subtilis*. In contrast, the Δ*sigI-rsgI* cells grown at 52°C displayed a previously unreported phenotype: cells were bent and their shape was irregular even when compared to the same cells grown at 37°C, as well as to Δ*rsgI* and wt cells grown at 37°C and 52°C.

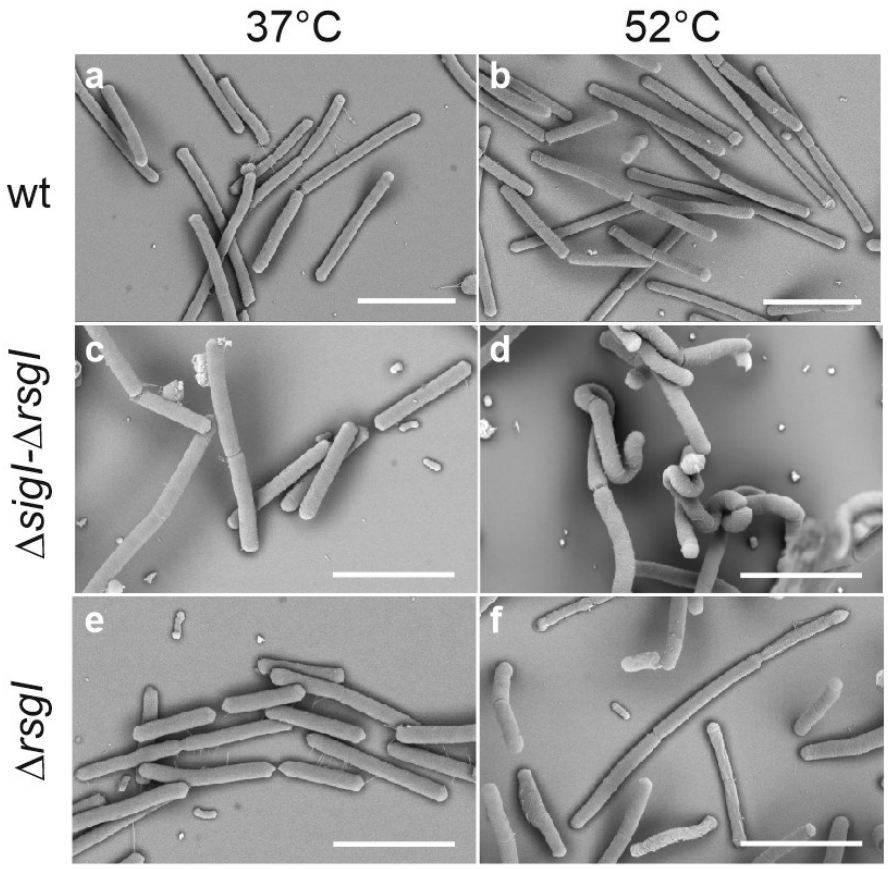


Figure 17. Cell shape of the *B. subtilis* wt, Δ*sigI-rsgI,* and Δ*rsgI* cells imaged by SEM. (a) wt grown in LB at 37°C. (b) wt grown in LB at 52°C. (c) Δ*sigI-rsgI* grown in LB at 37°C. (d) Δ*sigI-rsgI* grown in LB at 52°C. (e) Δ *rsgI* grown in LB at 37°C. (f) *rsgI* grown in LB at 52°C. The experiment was performed twice with identical results. The length of the white horizontal bar represents 5 µm.

2.4 σ^I impact on gene expression

To identify σ^I-affected genes we used Δ*sigI-rsgI* and wt strains for RNA-seq experiments. First, we tried an *in vitro* approach. Our idea was to synthesize *in vitro* specific σ^I-dependent RNAs from chromosomal DNA and detect them via RNA-seq. Unfortunately, the obtained samples contained too low an amount of desired RNAs. Optimization of the protocol did not improve the RNA yield. Therefore, we used the *in vivo* RNA-seq approach.

We cultivated the strains at 37°C and 52°C to mid-logarithmic phase, purified total RNA, depleted it for rRNA, and performed Illumina-based RNA-seq with a subsequent comparison of the Δ*sigI-rsgI* and wt transcriptomes to identify differentially expressed genes (DEGs).

First, we identified DEGs between Δ*sigI-rsgI* and wt cells grown at 37°C (>1.5-fold difference, <5% False Discovery Rate). Expression of 13 genes was decreased in the Δ*sigI-rsgI*

strain relative to wt (**Figure 18a**, blue circle; *i.e.,* genes positively regulated by σ^I). Proteins encoded by these genes are mainly involved in transcription, translation, and coping with stress (**Figure 19**). Moreover, expression of 63 genes was increased (upregulated) in the Δ*sigI-rsgI* strain (**Figure 18b**, blue circle; *i.e.*, genes negatively regulated by σ^I). They are mainly involved in the regulation of cell metabolism and in coping with stress (**Figure 20**). Thus, σ^I positively affects 13 genes and negatively 63 genes at 37° C.

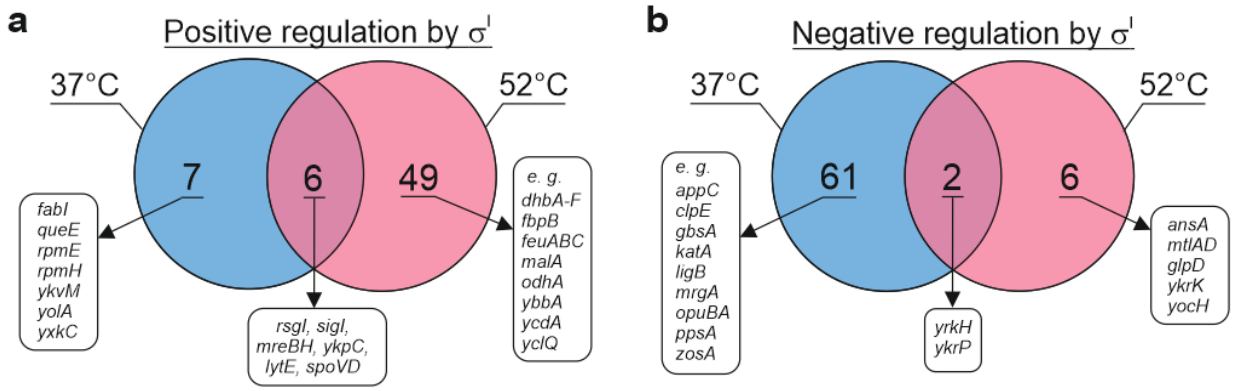


Figure 18. Genes affected by σ^I in *B. subtilis.* Δ*sigI-rsgI* and wt strains were grown at 37°C and 52°C in LB broth to OD_{600}~0.45. RNA was extracted and libraries were prepared for transcriptome sequencing (RNA-seq). (a) Genes positively regulated by σ^I. These genes were downregulated in the Δ*sigI-rsgI* strain compared to the wt strain. (b) Genes negatively regulated by σ^I. These genes were upregulated in Δ*sigI-rsgI* compared to wt.

Second, at 52°C we detected 55 downregulated and eight upregulated genes in Δ*sigI-rsgI* compared to wt (**Figure 18,** red circles). The majority of the affected genes are involved in the regulation of cell metabolism (*e.g.*, iron metabolism). In addition, upregulated genes included those involved in cell wall turnover (**Figures 19 - 20**). Thus, at 52°C, σ^I positively affected 55 genes and negatively eight genes. A considerable portion of these genes encoded membrane-associated proteins (**Figures 19 - 20**).

Six genes were downregulated at both temperatures – *lytE*, *sigI*, *rsgI*, *mreBH*, *ykpC* (previously known to be σ^I-dependent) and *spoVD*. Two genes – *ykrP* and *yrkH* – were upregulated in Δ*sigI-rsgI* compared to wt at both temperatures.

Taken together, based on the RNA-seq data, σ^I affected the expression of 131 genes organized in 90 operons (**Figure 18**). We note here that the identified DEGs could be regulated by σ^I either directly (the *bona fide* σ^I regulon) or indirectly.

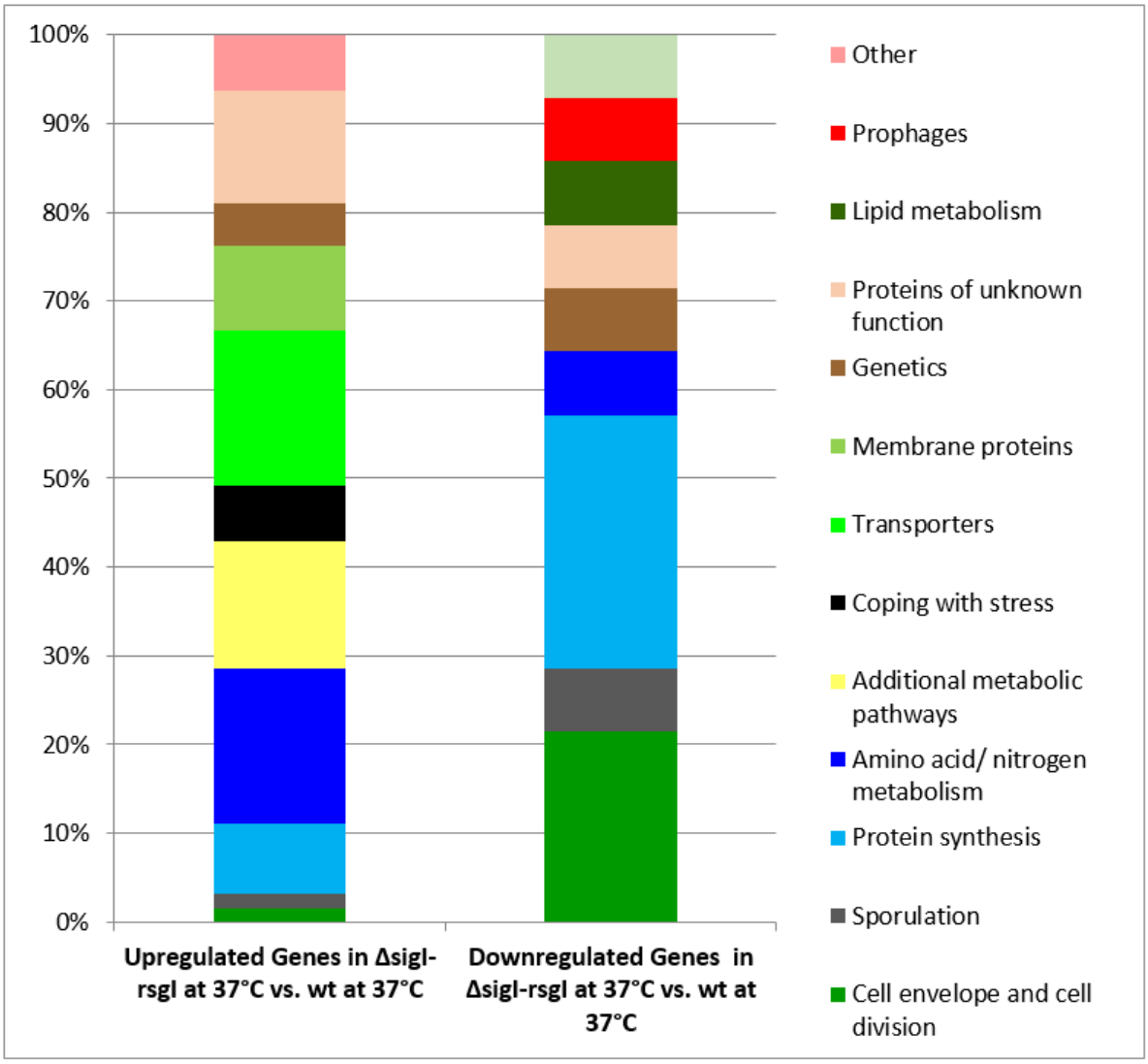


Figure 19. Gene ontology of the DEGs in the Δ*sigI-rsgI* strain at 37°C versus wt strain at 37°C. Data correspond to the Supplementary Tables S2 and S3 of the published article (Appendix 1). In this and following figures indicated gene ontologies were based on SubtiWiki datasets (Michna et al., 2016).

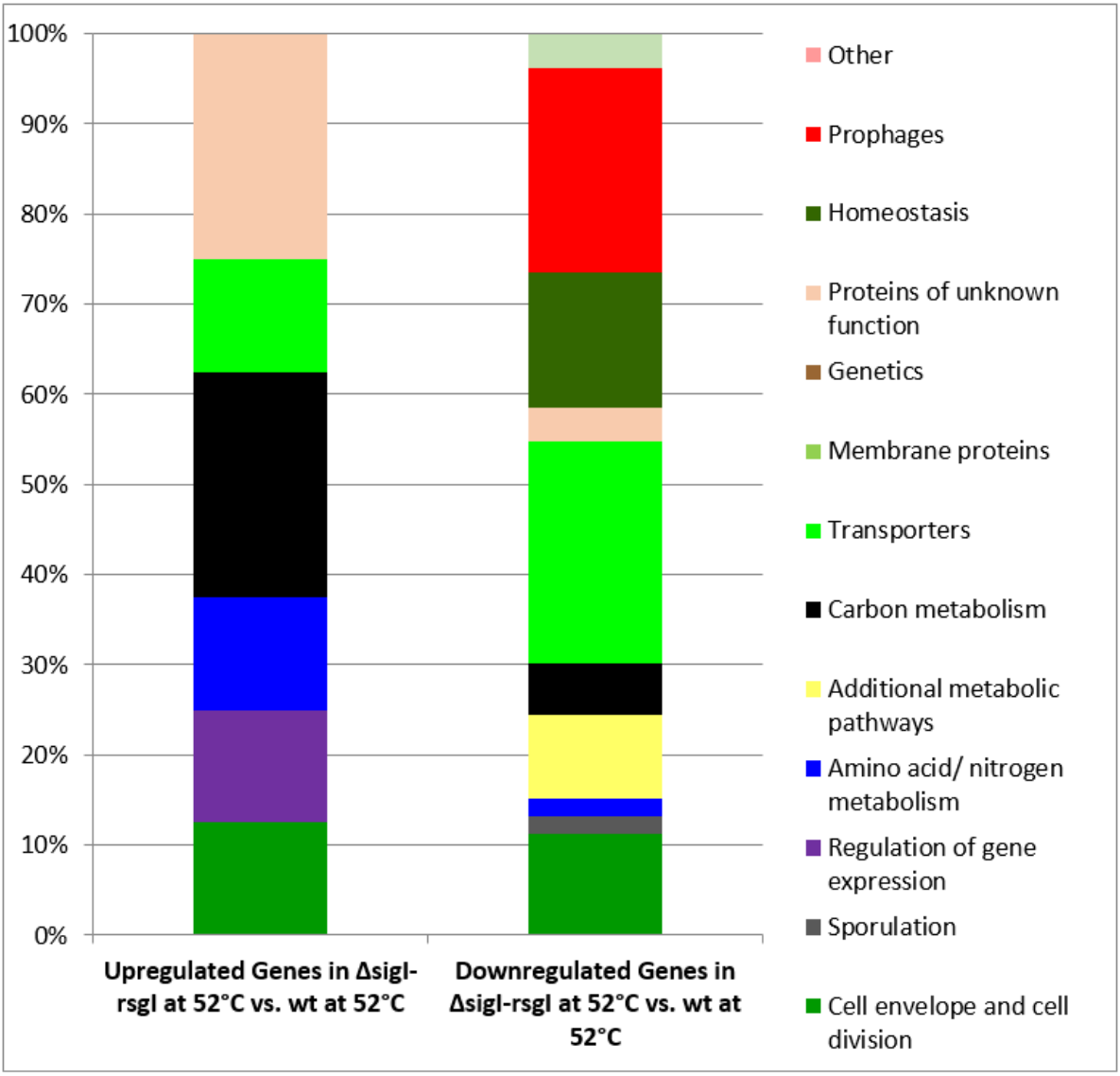


Figure 20. Gene ontology of the DEGs in the Δ*sigI-rsgI* strain at 52°C versus wt strain at 52°C. Data correspond to the Supplementary Tables S4 and S5 of the published article (Appendix 1).

2.5 σ^I is involved in iron metabolism

The RNA-seq data revealed that a significant number of genes influenced by σ^I (23 genes organized in 11 operons) were involved in iron metabolism. Two of these genes were misregulated at 37°C and 21 genes at 52°C.

The absence of σ^I increased expression of *zosA* [*pfeT*] (encodes Fe(II) efflux pump) and *mrgA* (encodes iron storage protein) genes at 37°C. We speculated that a combination of increased iron efflux and increased iron retention by MrgA protein in the Δ*sigI-rsgI* strain could affect iron homeostasis under limiting iron conditions. To test this hypothesis, we cultivated and Δ*rsgI* strains in a defined MOPS medium containing or lacking iron. While the doubling time of all strains was comparable, the results repeatedly showed that at 37°C, the Δ*sigI-rsgI* strain had a markedly prolonged lag phase in the absence of iron (**Figure 21a**).

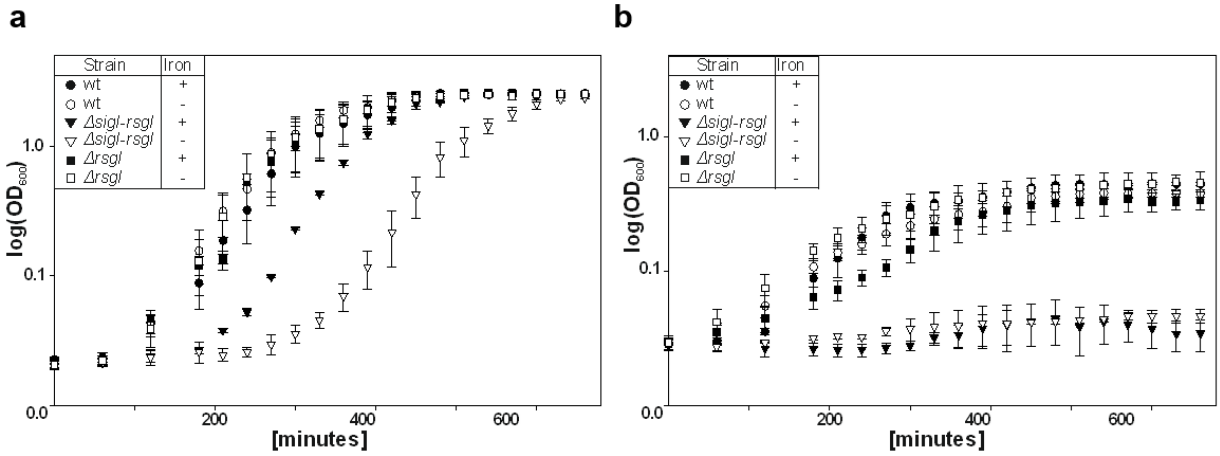


Figure 21. σ^I is involved in iron metabolism in *B. subtilis*. Growth curves of the wt (circles), Δ*sigI-rsgI* (triangles) and Δ*rsgI* (squares) strains grown in defined MOPS medium at (a) 37°C and (b) 52°C in the presence (filled shapes) / absence (empty shapes) of $FeCl_3$. The experiment was repeated three times. The error bars show ±SD.

The RNA-seq data indicated that the presence of σ^I increased expression of 21 genes associated with iron metabolism at 52°C, including genes involved in iron uptake (*e.g.*, enzymes participating in siderophore [iron chelating compounds] synthesis and ABC transporters of siderophores). However, the Δ*sigI-rsgI* strain did not grow at 52°C in the defined MOPS medium regardless of the presence or absence of iron (**Figure 21b**).

2.6 σ^I-affected genes in the context of heat stimulon

To assess the importance of σ^I for the cell, we compared genes affected by σ^I with the genes stimulated by heat in wt (52°C vs. 37°C) and searched where among these genes the σ^I-affected genes ranked (**Figure 22**, Appendix 1). According to our RNA-seq data, the wt heat stimulon contained >370 genes (DEGs > 2-fold difference). Seven genes that were stimulated by σ^I (either directly or indirectly) belonged among the top 10 % of the heat-stimulated genes (*e.g., malA* - carbon metabolism; *mreBH* – cell shape; *dhb* genes – iron metabolism). We concluded that σ^I was involved in the regulation of expression of genes whose stimulation was prominent at elevated temperature.

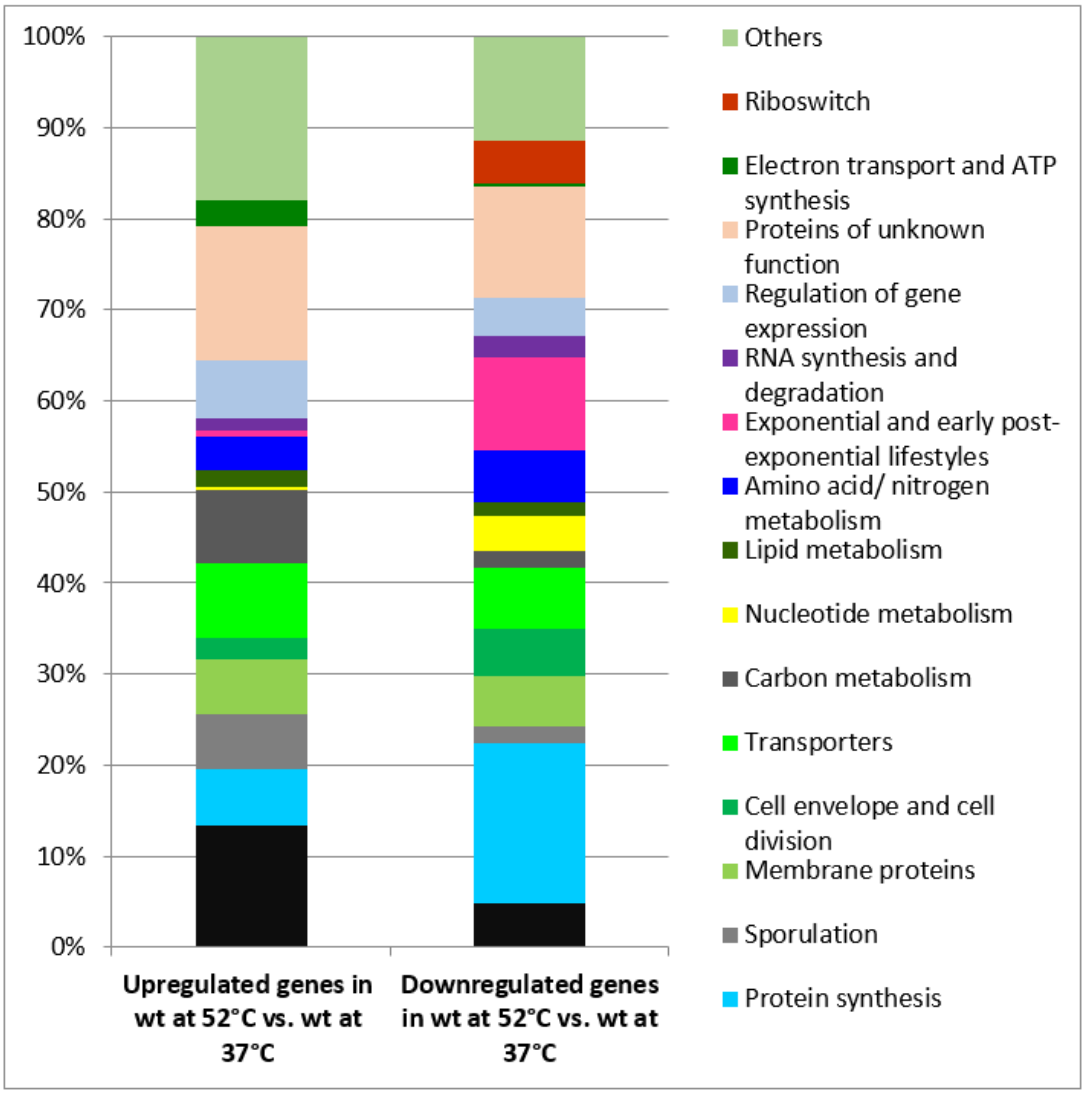


Figure 22. Gene ontology of the DEGs in the wt strain at 52°C versus wt strain at 37°C. Data corresponds to Supplementary Tables S6 and S7 of the published article (Appendix 1).

2.7 DNA binding properties of σ^I

To address DNA binding properties of σ^I, we used purified recombinant σ^I protein by nickel affinity chromatography via the introduced C-terminal His-tag (**Figure 23**; Chapter 1.1). As σ^I lacks domain 1.1 that in primary σ factors prevents their binding to DNA in the absence of RNAP, we tested by electrophoretic mobility shift assay (**Figure 24**) whether free σ^I is able to bind to its cognate promoter P*sigI* alone.

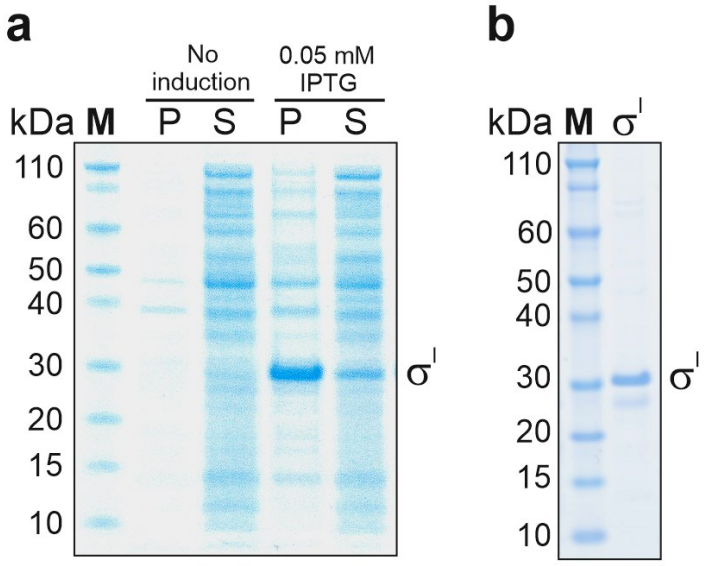


Figure 23. σ^I overexpression. (a) SDS-PAGE of the cell lysates from the #1242 strain (BL21/pSigI-6×His) after 3 hrs (time zero was at OD_{600}=0.6) of growth in the presence/absence of an inducer (0.05 mM IPTG; indicated). P, pellet; S, supernatant. (b) SDS-PAGE of the σ^I protein (1 µg) in a storage buffer. In (a) and (b) M represents protein marker (Novex™ Sharp Pre-stained Protein Standard).

As a control, we used σ^A in combination with the strong and well characterized P*veg* promoter. The results showed, that σ^I forms a specific complex with DNA only in the presence of the RNAP core, thus behaves in the same manner as the primary σ factor, σ^A.

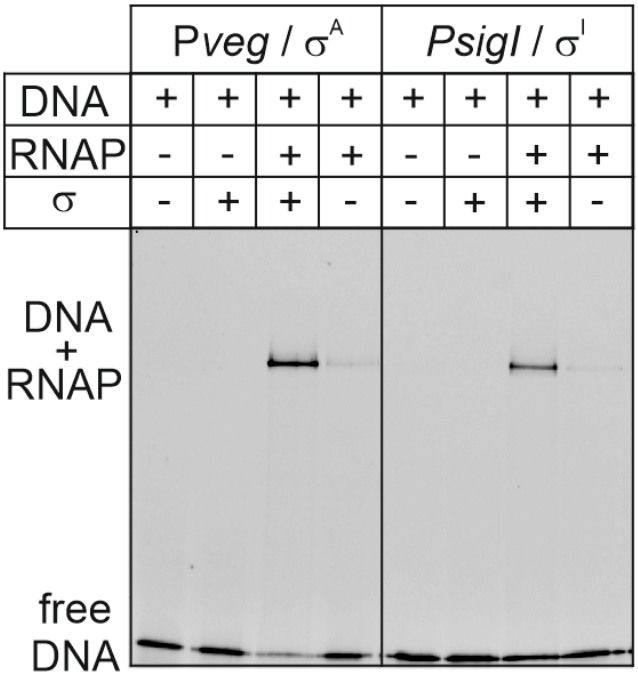


Figure 24. DNA binding properties of σ^I. A representative gel of EMSA performed with purified *B. subtilis* RNAP, $\sigma^{A,}$ and σ^I proteins. Promoter regions of P*veg* (-50/+10) and P*sigI* (-50/+10) labeled with the 6-FAM fluorescent dye served as DNA templates for experiments. The experiment was performed three times with the same result.

2.8 *In vitro* assays: σ^I regulon is small

As the next step, we set up an *in vitro* transcription system to assess the ability of σ^I to directly regulate selected genes. We used 25 promoter regions identified in our RNA-seq (including three promoter regions of genes previously demonstrated to be σ^I-dependent) and the promoter regions of *bcrC* and *gsiB*, which were previously shown to be affected by σ^I but did not appear in our transcriptomic screen (= 27 promoter regions in total, of them 22 putative σ^I-dependent promoter regions). We performed transcription assays with these templates both with RNAPσ^A and RNAPσ^I (**Figure 25**). RNAPσ^A was active on P*veg* (positive control) and also on promoter regions of *lytE*, *sigI*, and *dhb* operon genes. These genes were previously known to contain σ^A-dependent promoters; our results confirmed these findings (**Figure 25**). The experiments were done at 37°C. In addition, we also performed the same experiments at 52°C; the results were similar to those obtained at 37°C (data not shown).

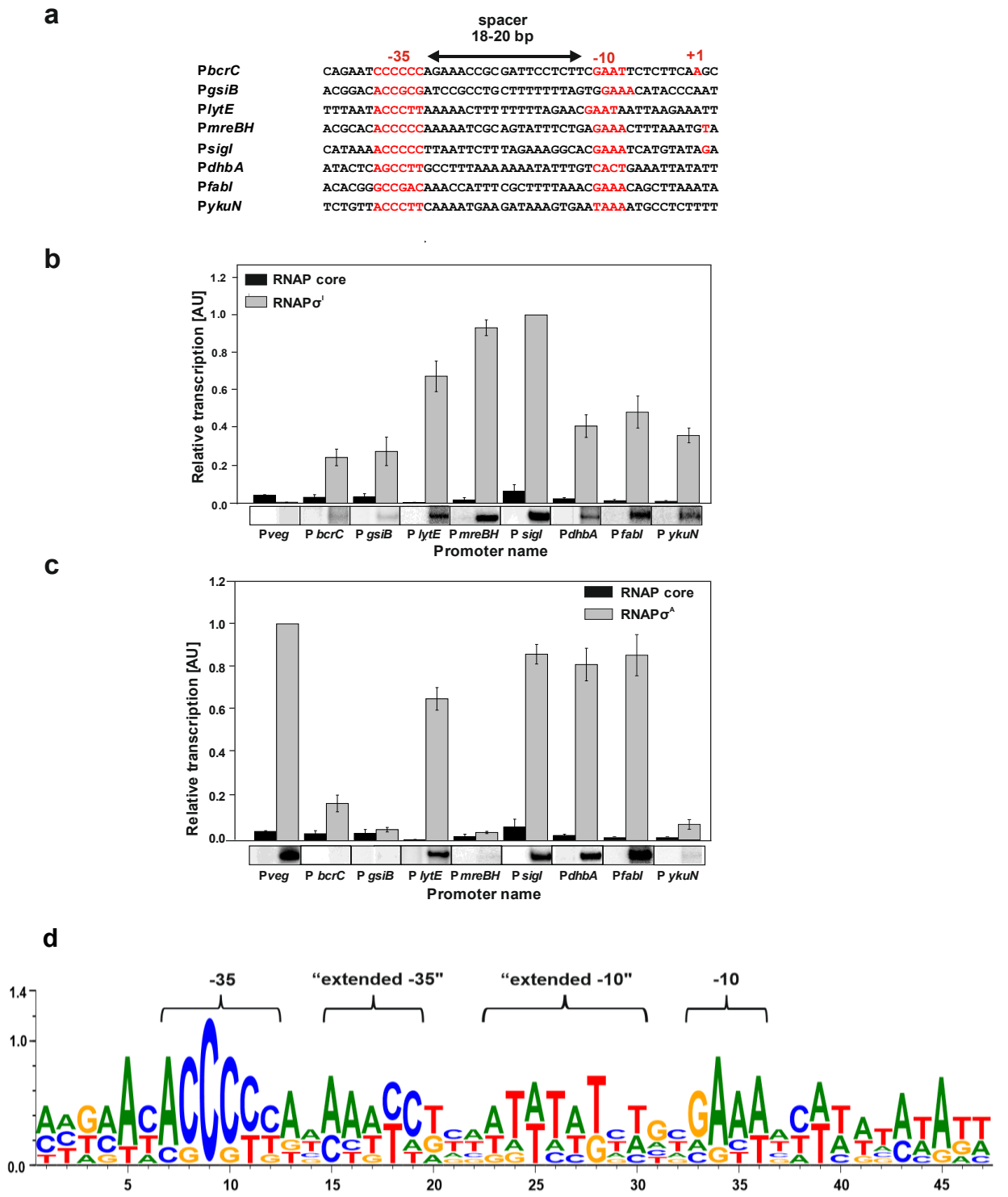


Figure 25. Multiple-round *in vitro* transcription assays with promoter regions of σ^I-regulated genes and RNAPσ^I or RNAPσ^A. (a) Alignment of σ^I-dependent promoters. The -10 and -35 elements and +1 position for P*sigI* (Asai et al., 2007), P*mreBH* and P*bcrC* (Tseng and Shaw, 2008) are in red. (b) Transcription was performed with RNAPσ^I holoenzyme and the RNAP core. Promoter P*sigI* was used as a control and its signal was set as 1. Primary data (radioactively labeled transcripts resolved on PAA gels) are shown below the graph. The error bars show averages from three independent experiments ±SD. (c) Transcription was performed with RNAPσ^A holoenzyme and the RNAP core. The strong constitutive σ^A-dependent promoter P*veg* was used as a control and its transcription was set as 1. (d) σ^I Consensus Logo created from promoter sequences shown in (a). Transcription with the RNAP core was used to assess potential contamination of the RNAP core with σ factors. Conserved promoter elements are indicated above the Logo.

RNAPσ^I was active on eight out of the 27 tested promoter regions: *bcrC*, *gsiB, sigI-rsgI*, *lytE*, *mreBH-ykpC*, *dhb, yku,* and *fabI.* The first five promoter regions corresponded to genes/operons known to be regulated by σ^I. Thus, among the newly identified putative 22 promoter regions, RNAPσ^I was active only on three of them (promoter regions of *dhb* and *yku* operons and the *fabI* gene). Although we did not test the remaining 37 promoter regions positively affected by σ^I according to RNA-seq, this result strongly implied that the σ^I regulon is small. Finally, we aligned all the known and newly identified σ^I promoter regions and created a σ^I promoter sequence logo (**Figure 25d**).

2.9 P*sigI-rsgI* promoter sequence analysis

We compared previously published σ^I promoter logo created from σ^I-dependent promoter regions from *Bacillales* (Muñoz-Gutiérrez et al., 2015; Tseng and Shaw, 2008) with our updated σ^I promoter logo. In addition to the -10, -35 and "extended -35" elements, we also identified an "extended -10" element, dissimilar from the well-studied extended -10 element of σ^A-dependent promoters (Jarmer et al., 2001) [**Figure 25d**]. A bioinformatic search for the σ^I consensus sequence among promoter regions of the σ^I-affected genes yielded no obvious additional hits. As the σ^I logo was information-rich (compared to *e.g.* σ^A), we wished to determine the importance of the individual sequence elements for the promoter activity. For this purpose, we used the P*sigI* promoter and systematically mutated these elements (**Figure 26**).

The results showed that double substitutions in -35 and -10 elements drastically reduced transcription (constructs 2, 5; by ~90% and ~97%, respectively) and their combination (construct 8) virtually abolished transcription. Interestingly, the experiments revealed that the "extended -35" element is highly important, as its mutation canceled transcription almost completely (construct 3). The "extended -10" element then led to an approximately 40% decrease in transcription activity (construct 6), suggesting that it is still required for optimal performance of the promoter (**Figure 26)**. Control "neutral" mutations without expected large effects on transcription confirmed that the "extended" elements are *bona fide* promoter regulatory sequences involved in σ^I-dependent transcription (**Figure 26b**; constructs 10-11).

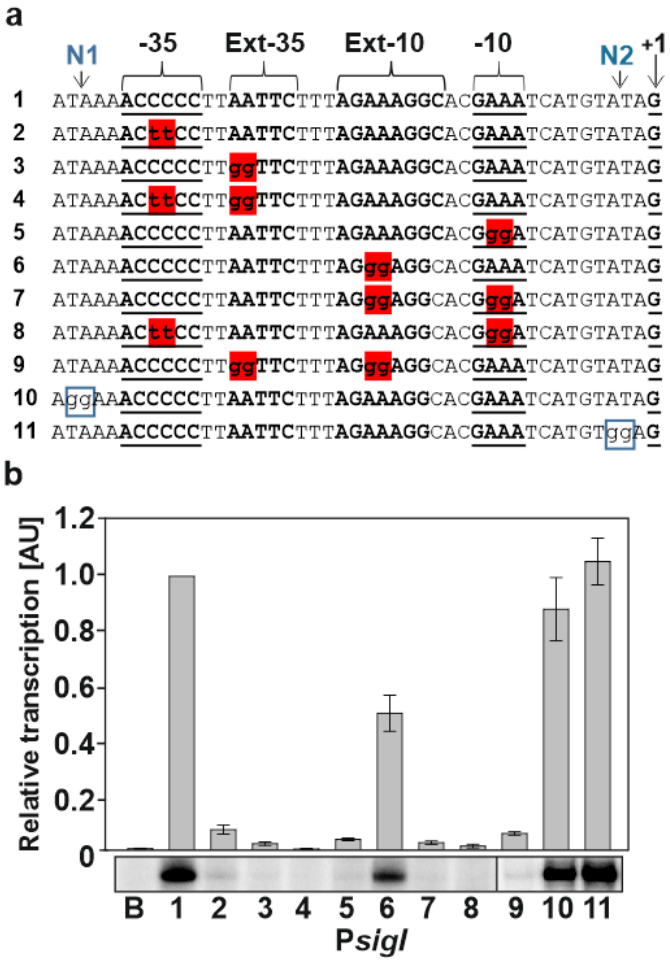
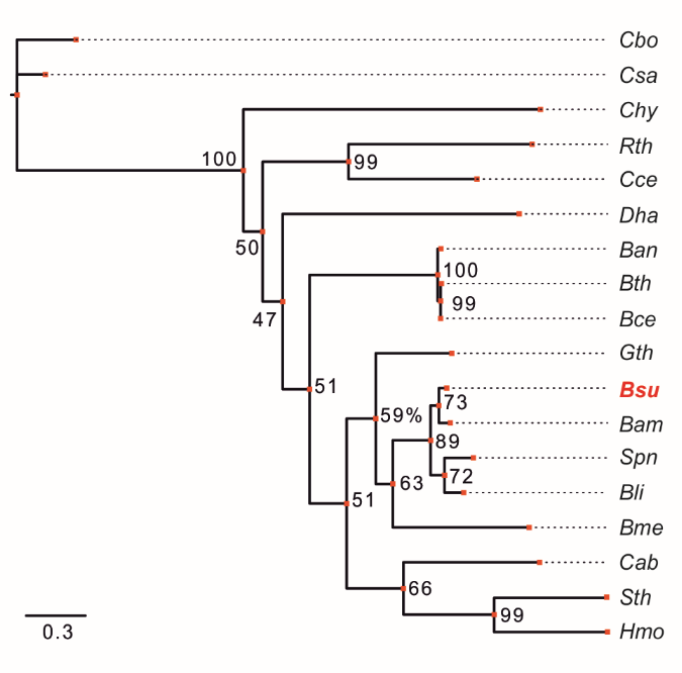


Figure 26. Multiple-round *in vitro* transcription assays with mutated P*sigI* promoter region. (a) Fragments of the P*sigI* promoter region used for evaluating the importance of the identified sequence elements. Fragment 1 contains the native P*sigI* promoter region. Fragments 2-9 contain mutations in -10, -35, "extended -35" (Ext -35), "extended -10" (Ext -10) elements, and their combinations. Mutations are highlighted in red. Fragments 10-11 contain control "neutral" (N) mutations highlighted in blue boxes. -10 and -35 regions and +1 position are in bold and underlined. "Extended" elements are in bold. (b) *In vitro* transcription with specific double-substitutions in P*sigI* -10 and -35 σ^I binding sites, "extended -10" and "extended -35" elements. Transcription was performed with the RNAP core (lane B – blank assay) and RNAPσ^I holoenzyme on PCR-products as a template. The blank assay was used to demonstrate that the RNAP core was devoid of contaminating σ factors. Primary data (radioactively labeled transcripts resolved on PAA gel) are shown below the quantification graph; the vertical black line indicates the border between two PAA gels used for illustration. Error bars show averages from three independent experiments ±SD.

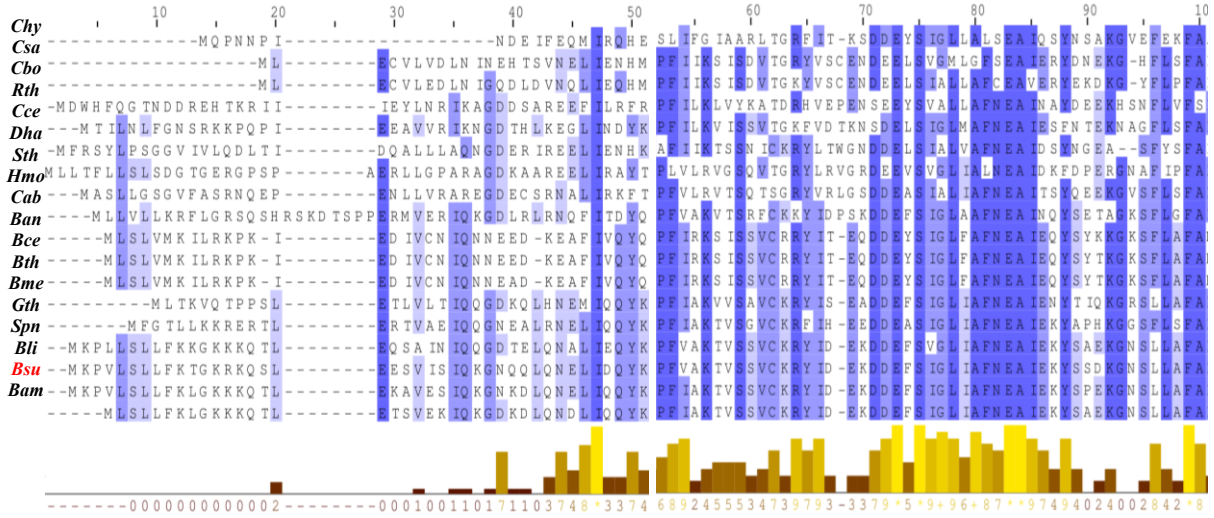
2.10 σ^I in different bacterial species

As a final analysis, we searched for σ^I homologs in other bacterial species, using BLAST (Agarwala et al., 2016) searches and created a phylogenetic tree from the most related amino acid sequences (**Figure 27**). The closest relatives of *B. subtilis* σ^I were proteins from *B. amyloliquefaciens*, a putative σ^I from *Streptococcus pneumoniae*, and σ^I from *B. licheniformis*.

a


Cbo
Csa
Chy
Rth
Cce
Dha
Ban
Bth
Bce
Gth
Bsu
Bam
Spn
Bli
Bme
Cab
Sth
Hmo
100
99
50
100
99
47
51
59%
73
89
72
63
51
66
99
0.3

b


Chy
Csa
Cbo
Rth
Cce
Dha
Sth
Hmo
Cab
Ban
Bce
Bth
Bme
Gth
Spn
Bli
Bsu
Bam

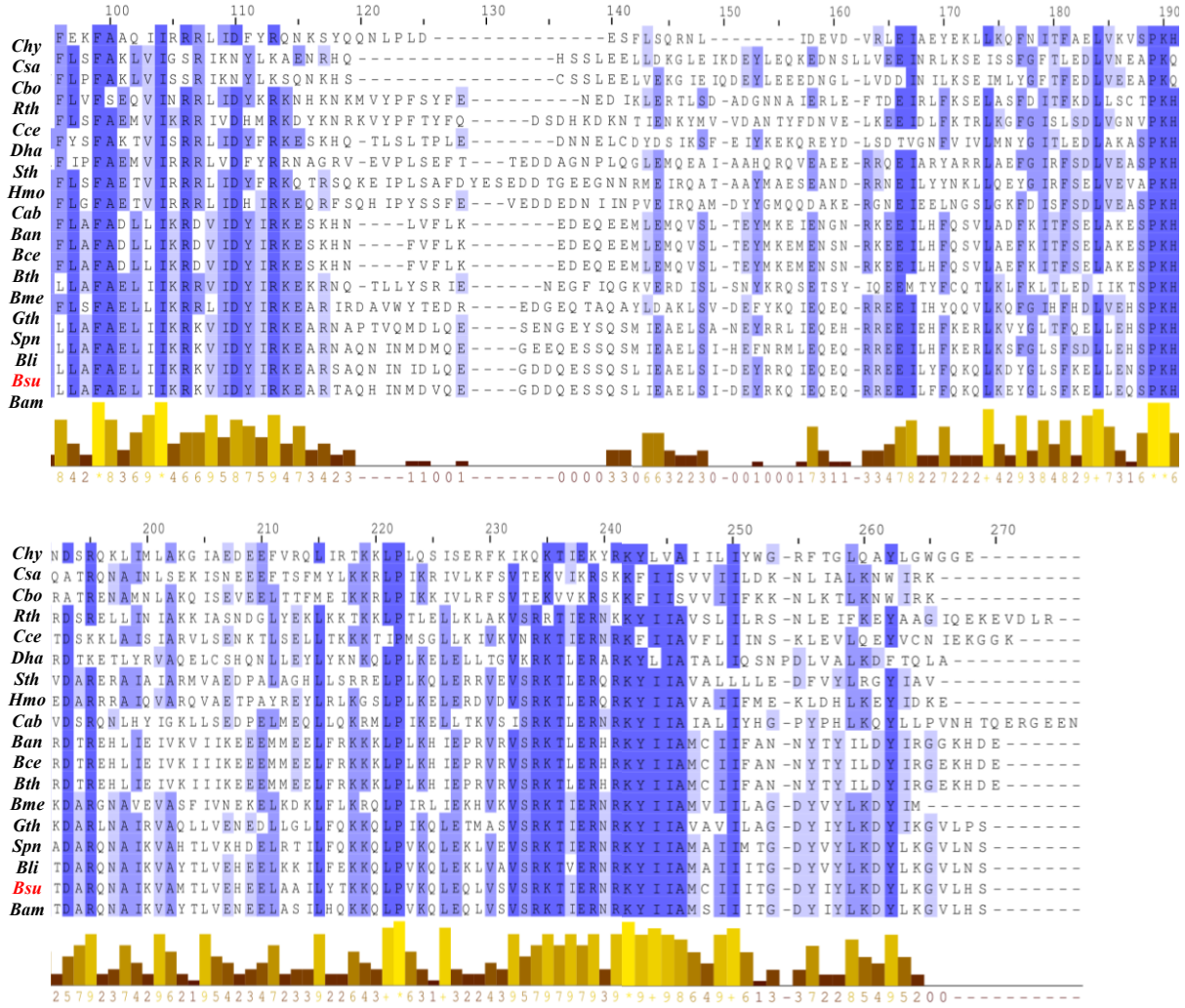


Figure 27. σ^{I} homologs from different bacterial species. (a) Phylogenetic tree of *B. subtilis* σ^{I} homologs. The phylogenetic tree was inferred with RAxML and the best-scoring maximum likelihood tree is shown. Numbers denote bootstrap values in percent as reported by RAxML. The scale bar represents expected number of substitutions per site. *Cbo – Clostridium botulinum, Csa – Clostridium sacharobutylicum, Chy – Carboxydothermus hydrogenoformans, Rth – Rumuniclostridium thermocellum, Cce – Clostridium cellulolyticum, Dha – Desulfitobacterium hafniense, Ban – Bacillus anthracis, Bth – Bacillus thuringiensis, Bce – Bacillus cereus, Gth – Geobacillus thermodenitrificans, Bsu (in red) – Bacillus subtilis, Bam – Bacillus amyloliquefaciens, Spn – Streptococcus pneumoniae, Bli – Bacillus licheniformis, Bme – Bacillus megaterium, Cab – Chlamydia abortus, Sth – Symbiobacterium thermophilum, Hmo – Heliobacterium modesticaldum*. (b) Alignment of σ^{I} homologs. Alignment of selected sequences was performed with MUSCLE(Edgar, 2004). Picture, conservation score, and percentage identity were computed with Jalview(Waterhouse et al., 2009). The coloring denotes percentage identity in each column, only residues which match consensus sequence are colored, ranging from lightest blue to deep blue respective to increasing percentage identity. The consensus sequence is not shown. Conservation annotation score ranges from 0 to 11, where 10 and 11 are denoted by ("+" and "*" respectively), "-" means no score was computed.

Discussion 2

We comprehensively characterized one of the least studied σ factors from *B. subtilis* – σ^I. We determined the effect of its presence/absence on gene expression during exponential growth in a rich medium at 37°C and 52°C, identified its involvement in iron metabolism, defined its DNA binding properties and characterized DNA sequences important for its transcriptional activity.

First, we observed previously unreported malformed cell shape of Δ*sigI-rsgI* at 52°C (**Figure 17**). This malformation might be attributed to the pronounced downregulation of *mreBH* and *lytE*, as it was previously demonstrated that both MreBH and LytE were required for proper cell shape (Carballido-López et al., 2006; Domínguez-Cuevas et al., 2013). Moreover, these two proteins were shown to interact in the cell (Carballido-López et al., 2006; Huang et al., 2013). The downregulation of both *mreBH* and *lytE* in the Δ*sigI-rsgI* strain at 52°C was quite pronounced (~16x↓ for both genes), whereas at 37°C it was rather moderate (~3x↓ for both genes), suggesting why the change in cell morphology was observed only at the higher temperature. In addition, *yfiY* and *yjeA*, which are involved in the cell wall metabolism, were also downregulated, and this might have contributed to the distorted cell shape.

To find σ^I-affected genes we first applied *in vitro* approach, but it failed because of technical difficulties. Next, using the *in vivo* RNA-seq approach, we revealed that σ^I affected 131 genes either in a positive or negative way (**Figure 18**). Nevertheless, the majority of these genes appear to be affected by σ^I indirectly: 69 genes out of 130 were affected by σ^I negatively. This fact itself excluded the possibility of direct regulation by σ^I, as it cannot bind to DNA alone (**Figure 24**). Of the seven genes that had been known or proposed to be regulated by σ^I, we identified five by RNA-seq, which gave us confidence in the credibility of the transcriptomic data. We also demonstrated the direct ability of RNAPσ^I to initiate transcription from upstream regions of these genes. In addition, we detected σ^I-dependent promoter activity for the *bcrC* and *gsiB* genes *in vitro*. However, we did not identify these genes in our RNA-seq. This could be due to the very weak transcriptional activity of the σ^I-dependent promoters of these genes (see **Figure 25**) and, perhaps, to the specific experimental conditions used: *i.e.*, chronic heat stress (this study) vs. temperature shift (Tseng and Shaw, 2008; Zuber et al., 2001).

Out of the 22 tested putative new σ^I-dependent promoter regions, *in vitro* transcription analysis revealed three new σ^I-dependent promoters driving transcription of nine genes (*dhbA-dhbC-dhbE-dhbB-dhbF*, *ykuN-ykuO-ykuP*, and *fabI*) [**Figure 25**]. The first two operons are

involved in iron metabolism. The *dhb* operon is involved in bacillibactin siderophore synthesis (May et al., 2001). The *yku* operon encodes flavodoxins replacing ferredoxin under conditions of iron limitation and catalyzing the O_2-dependent desaturation of the acyl chain of membrane phospholipids (Baichoo et al., 2002). In our transcriptomic data (in SigI-null relative to wt), we observed decreased expression of *ykuO* and *ykuP* genes at 52°C, but their σ^I-dependent promoter was identified at *in vitro* transcription within a DNA fragment preceding the first gene of the operon – *ykuN*. Careful inspection of the RNA-seq data revealed that also the *ykuN* gene was affected (~ 2x ↓) by the absence of σ^I in the same way as *ykuO* and *ykuP*, but this change in expression was not deemed significant by the DESeq2 algorithm due to the low level of *ykuN* expression. The product of the *fabI* gene is enoyl-[acyl-carrier-protein] reductase that participates in fatty acid biosynthesis (Heath et al., 2000). Expression of the *fabI* gene was affected at 37°C, suggesting that this effect is heat-shock independent. A search for σ^I promoter sequences within the upstream regions of genes that did not function at *in vitro* transcriptions revealed, in all cases but one (*feuA*), the -10 GAAA motif and the absence of -35 regions. Hence, it is possible that some of these genes may still be regulated by σ^I and the presence of an unknown regulator is required. Nevertheless, as there are still 37 promoter regions with putative activation by σ^I untested, it is possible that a few more genes may be added to the list in the future. Thus, the σ^I regulon is one of the smallest known *B. subtilis* regulons, currently containing 16 genes. Moreover, as there still are 37 promoter regions with putative activation by σ^I untested, a few more genes may be added to the list in the future. An alternative approach – ChIP-seq – can be applied to determine direct targets of σ^I, as well as to verify already determined interactions.

A significant fraction of the affected genes in the Δ*sigI* mutant were genes involved in iron metabolism. Consequently, a growth defect – prolonged lag phase – was observed in the absence of iron at 37°C (**Figure 21**). This could be due to altered iron homeostasis, as iron uptake during the lag phase is important for preparing the cell for subsequent exponential growth (Rolfe et al., 2012). Moreover, the involvement of *B. subtilis* σ^I in iron metabolism suggests functional similarity of *B. subtilis* σ^I with extracytoplasmic function (ECF) σ factors from various bacterial species that are often involved in this process (Gruber and Gross, 2003).

We used all eight known promoter regions to create the σ^I promoter sequence logo (**Figure 25d**). Our logo revealed that σ^I-dependent promoters are information rich in the spacer region: besides -35 and -10 elements, "extended" -35 and -10 elements were identified. Importantly, the functional analysis of "extended" elements demonstrated their significance for the promoter

activity (**Figure 26**). The importance of the spacer region sequence for efficient promoter utilization was demonstrated in a number of previous studies. In Gram-negative *E. coli* it was shown that an AT-rich spacer could both stimulate and inhibit transcription initiation, depending on the promoter (Gaballa et al., 2017; Hook-Barnard and Hinton, 2009). For extracytoplasmic function (ECF) σ factor-dependent promoters from *B. subtilis* it was demonstrated that the homopolymeric T-tract motif, proximal to the -35 element, functioned in combination with the core promoter sequences to determine the selectivity of ECF σ factors (Gaballa et al., 2017). This homopolymeric T-tract is reminiscent of, and might be analogous to, the σ^I "extended -35" element described here. The sequence and position of the σ^I "extended -10" element then differs from the TRTGN motif that was described for $\sigma^{A/70}$-dependent promoters (Sudalaiyadum Perumal et al., 2018; Voskuil and Chambliss, 1998), where it precedes -10 by 2-5 bp and interacts with domain 3 of $\sigma^{A/70}$ (Mitchell et al., 2003; Ruff et al., 2015). The "extended" -10 of σ^I, however, is positioned 7-8 bp upstream of -10, and may thus not interact with domain 3. It is possible that the A:T pairs within the "extended" promoter elements increase conformational flexibility (Johnson et al., 2013; Okonogi et al., 2002) of the relatively long spacer [18-20 bp; compared to σ^A promoters ~ 17 bp (Helmann, 1995)] and thus contribute to promoter accommodation onto RNAP and proper interactions of -35 and -10 with σ^I. Future studies (including more detailed mutagenesis, changing the length of the spacer region) will be required to address the detailed roles of the "extended" elements in recognition of these promoters by RNAP.

Our data suggest that, despite the small number of directly regulated genes, σ^I is significant for the proper functioning of the cell. None of the σ^I directly regulated genes alone is likely responsible for the growth defect at elevated temperature. Nevertheless, some of the indirectly affected genes could be responsible for this phenotype. The growth defect might be, at least in part, due to downregulation of genes involved in the carbon uptake and central metabolism, such as *malA* that encodes NAD(H)-dependent phospho-α-1,4-glucosidase (Thompson et al., 1998) (the 3rd most upregulated gene in wt 37°C→52°C) and *odhA* that encodes 2-oxoglutarate dehydrogenase that is an essential enzyme of the TCA cycle (Carlsson and Hederstedt, 1989).

In conclusion of our study, we found that homologs of σ^I from *B. subtilis* exist in other *Bacilli* and *Clostridia* as well as in some other species, such as *Heliobacterium* sp. and *Geobacillus* sp. (O'Leary et al., 2016). A phylogenetic tree of σ^I factor homologs is shown in **Figure 27**. Interestingly, these factors in different organisms regulate divergent sets of genes [thermotolerance (Tseng et al., 2011; Zuber et al., 2001), virulence (Kim and Wilson, 2016),

polysaccharide sensing (Muñoz-Gutiérrez et al., 2015)], although the promoter consensus sequences, at least between *Bacilli* and *Clostridia*, are similar (Muñoz-Gutiérrez et al., 2015), illustrating how different species designed the use of a transcription factor to their specific needs.

3. Validation of an *in silico* model generated for *B. subtilis* σ^A regulatory network during spore germination and outgrowth using *in vitro* transcription

Chapter 3 represents the project carried out in collaboration with colleagues-bioinformaticians (Institute of Microbiology of the CAS, Laboratory of Bioinformatics, head – Jiří Vohradský). **The goal of this study was to reveal how many and which genes are directly regulated by primary *B. subtilis* σ factor (σ^A) during spore germination and outgrowth without the need to invoke additional regulatory layers.**

My aims in this project were:

- Comprehensive literature-based analysis of known σ^A interactions and selected alternative σ factor interactions in *B. subtilis*;
- Analysis and experimental verification of newly predicted targets of σ^A- and selected alternative σ factors-dependent promoter regions.

The obtained results were published as my first-author paper "Kinetic modelling and meta-analysis of the *B. subtilis* σ^A regulatory network during spore germination and outgrowth" in BBA – Gene Regulatory mechanisms (2017) [Appendix 2].

3.1 σ^A regulatory network: an overview

We downloaded initial data from 14 time points [microarrays] (0, 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90 and 100 min) obtained during germination and outgrowth of *B. subtilis* cells (Keijser et al., 2007); from GEO database (see Materials and Methods). Based on this data our colleagues created a kinetic model of the σ^A-controlled regulatory network. They predicted new potential targets of σ^A and σ^A-dependent alternative σ factors. For genes that did not match the σ^A kinetic profile, they suggested alternative σ factors capable of modelling their expression profiles. Subsequently, we extracted the known σ^A regulon genes from SubtiWiki and DBTBS databases (Michna et al., 2016; Sierro et al., 2008). Then we analyzed their kinetic profiles based on the gene expression time series. After data processing, these genes were split into two categories: 1. confirmed to be controlled by σ^A (consistent with the literature and the results of modelling) and 2. found to not satisfy the kinetic constraints (**Tables 6** and **7**). The model of gene expression created by Jiří Vohradský and colleagues is presented in Materials and Methods. To discover novel putative σ^A-controlled genes, we used all *B. subtilis* genes that were not identified in databases as members of the σ^A regulon and extracted their kinetic profiles. Then, their expression profiles were modelled with σ^A (*sigA* mRNA) as the regulator and genes with good matching profiles were selected (**Table 8**).

Table 6. List of two groups of genes found to be controlled by σ^A consistent with the literature and the results of modelling. The two groups differ in the sign of the parameter w (positive or negative). Complete operons are listed; genes found by modelling are shown in bold. The genes that are not in bold were either not present in the source dataset, or their profiles were flat or too low to be processed. Hyphens indicate genes in operons.

Positive control w>0
*adhA-**yraA**, aroF-aroB-aroH-trpE-trpD-trpC-trpF-trpB-**trpA-hisC-tyrA**-aroE, **cggR**-gapA-pgk-tpi-pgm-eno*, ***citG**, citZ-**icd-mdh**, ctsR-mcsA-mcsB-**clpC**-radA-yacK*, ***codV-clpQ-clpY-codY**, **clpX**, **comA**-yuxO, comEA-**comEB**-comEB, **comQ-comX**, cymR-yrvO-**trmU**, **cspB**, **ctaA**, **drm-punA**, epsA-epsB-epsC-**yveN**-epsE-epsF-epsG-epsH-epsI-epsJ-epsK-epsL-epsM-epsN-epsO, fadF-acdA-**rpoE**, **rpsD**, **ftsA-ftsZ**, **ftsH**, **fabHA-fabF**, **fbp**, **fur**, **galE**, **gcaD-prs**-ctc, **glnR**-glnA, **gltA-gltB**, **gudB**, **gyrA**, **gyr-recF-yaaB**-gyrB, hemA-hemX-**hemC-hemD-hemB-hemL**, **htpG**, ilvB-ilvH-**ilvC**-leuA-leuB-leuC-leuD, **infC-rpmI-rplT**-ysdA, **lepA-hemN**-hrcA-**grpE**- -dnaK-**dnaJ-yqeT-yqeU-yqeV**, lcfA-**ysiA-ysiB-etfB**-etfA, **lonA-ysxC**, **lytA**-lytB-**lytC**, **lytR**, med-**comZ**, **menF-menD**-ytxM-**menB**-menE-menC, **mtrA-mtrB**, **murE-mraY-murD-spoVE-murG-murB-divIB**-ylxW-**ylxX**-sbp, nifS-**yrxA**, **nfrA**-ywcH, opuAA-**opuAB-opuAC**, pabB-pabA-pabC-**sul-folB**-folK-yazB-yacF-**lysS**, **pheS**-pheT, **ppiB**, **pssA-ybfM-psd**, ptsG-**ptsH-ptsI**, **pucA**-pucB-pucC-pucD-pucE, **pyrG**, **recA**, **resA**-resB-resC-resD-resE, **rnc**-smc-**ftsY**, **rpmH**, **rpoB**, secA-**prfB**, **sigX-rsiX**, **sigM-yhdL**, sinI-**sinR**, **sipS**, **speE-speB**, gapB-**speD**, **spo0A**, **spo0B-obg**, **spo0E**, spoIIIJ-**jag**, spoVFA-spoVFB-asd-**dapG-dapA**, **thrS**, **thyB-dfrA-ypkP**, **veg**, valS-**folC**, **ydjM**, **yfhQ-fabL**-sspE, **yfkJ-yfkI-yfkH**, **ykuJ-ykuK-ykzF-ykuL**-ccpC*, ***ykuN-ykuO-ykuP**, **yocH**, ylpC-plsX-fabD-**fabG**-acpA, ylxM-**ffh**, ylxS-nusA-ylxR-ylxQ-**infB**-ylxP-**rbfA**, **ymaA-nrdE-nrdF-ymaB**, ypuE-ribD-ribE-ribA-ribH-**ribT**, **yusM**-yusL-yusK-yusJ, **yuxH**, **yqxD-dnaG-sigA**, **yvcE***
Negative control w<0
*arsR-**yqcK**-arsB-**arsC**, **citB**, **citZ**-icd-mdh, ctsR-mcsA-**mcsB**-clpC-radA-yacK, **cysH-cysP-sat-cysC-ylnD-ylnE-ylnF**, glpF-**glpK**, iolR-**iolS**, **odhA-odhB**, **pbpD**-yuxK*, ***pckA**, pheS-**pheT**, pucA-**pucB**-pucC-pucD-pucE, **pucR**-pucJ-pucK-pucL-pucM, **purT**, **rsbR**-rsbS-rsbT-rsbU-rsbV-rsbW-sigB-rsbX, **sdhC-sdhA-sdhB**, **xylA**-xylB, ycdH-ycdI-**yceA**, **ykrT-ykrS**, ytrG-ytrA-ytrB-ytrC-ytrD-**ytrE-ytrF**, yusM-yusL-**yusK**-yusJ, **yvgR-yvgQ***

Table 7. List of 121 genes (56 operons) of the σ^A regulon identified in the literature that did not satisfy kinetic constraints. Complete operons are listed, and the genes found by modelling are shown in bold. The genes that are not in bold were either not present in the source dataset, or their profiles were flat or too low to be processed. Hyphens indicate genes in operons.

ackA, ***ahpC-ahpF***, ***ald***, *alsS-***alsD***, **ansA-ansB**, **clpE**, **clpP**, *gcaD-prs-***ctc**, **ctsR-mcsA**-*mcsB-clpC-radA-***yacK**, **cysK**, *degS-***degU**, **dnaA-dnaN**, *gltX-cysE-cysS-***yazC**-*yacO-yacP*, **glyA**, **groES-groEL**, **gtaB**, **guaA**, **guaD**, **hemA-hemX**-*hemC-hemD-hemB-hemL*, **hemZ**, *lepA-hemN-***hrcA**-*grpE-dnaK-dnaJ-yqeT-yqeU-yqeV*, **lmrA-lmrB**, **mecA**, *murE-mraY-murD-spoVE-murG-murB-divIB-***ylxW**-*ylxX-***sbp**, **nadE**, **opuAA**-*opuAB-opuAC*, **pabB-pabA-pabC-***sul-folB-***folK-yazB**-*yacF-lysS*, **pta**, **ptsG**-*ptsH-ptsI*, *pucR-***pucJ**-*pucK-***pucL**-*pucM*, **purA**, **purE-purK-purB-purC-purS-purQ-purL-purF-purM-purN-purH-purD**, **pyrR-pyrP-pyrB-pyrC-pyrAA-pyrAB-pyrK-pyrD-pyrF-pyrE**, *rbsR-rbsK-rbsD-rbsA-rbsC-***rbsB**, *rnc-***smc**-*ftsY*, *rsbR-***rsbS-rsbT-rsbU-rsbV-rsbW-sigB-rsbX**, **secA**-*prfB*, **tagA-tagB**, **tagD-tagE-tagF**, **xpt-pbuX**, **ycdA**, **ycdH-ycdI**-*yceA*, **yceC-yceD**-*yceE-yceF-yceG-yceH*, **yhaG**, **yhcL**, **yjbC-yjbD**, **yjeA**, **ylxM**-*ffh*, **ylxS-nusA-ylxR-ylxQ**-*infB-***ylxP**-*rbfA*, **yoeB**, **ylpC-plsX-fabD**-*fabG-acpA*, *ypuE-***ribD-ribE-ribA-ribH**-*ribT*, **yrrT-mtnN-yrhA-yrhB-yrhC**, **yvbA-yvaZ**, **yxeK-yxeL-yxeM-yxeN-yxeO-yxeP**-*yxeQ*, *wapA-wapI-wapA-***yxxG**

Table 8. Potential new target genes of σ^A. Genes whose expression profiles were best fitted by the model with σ^A as the regulator. Whole operons are listed, and genes that were subjected to modelling are shown in bold. The genes that are not in bold were either not present in the source dataset, or their profiles were flat or too low to be processed. Hyphens indicate genes in operons.

yqzB-yqfL, **fbaA-ywjH**, **yyaF**-*rpsF-ssbA-***rpsR**, **acpA**, **yprA**, **yloC**, **ahrC-recN**, **efp**, **ymfK**, **smpB**, **fmt**, **prpC-prkC-yloQ**, **proS**, **hepS**, **ruvA-ruvB**, **ykpA**, **rpmGA**, **yrrC**, **ypjQ**, **yrhP**, **ykaA-pit**, **yloI**, **dinG**, **yloN**, **mrpA**-*mrpB-mrpC-***mrpD**-*mrpE-mrpF-mrpG*, **yycC-yycB**, **alsT**, **yloH**, **yugI**, **yprB**, **recQ**, **yrrS**, **ywiB**, **yneF**, **rnpA**, **priA**, **yncD**, **mutS-mutL**, **yloU-yloV**, *mobA-***moeB-moeA-mobB**-*moaE-***moaD**, **ylzA**, **def**, **yaaD-yaaE**, **rpsO**, **recG**, **ynbA**, **aroA**, **nudF**, **ylmG-ylmH**, **ylmB**, **rpsT**, **ylbB-ylbC**, **ylbG**, **ylbH**, **ylbF**, **ylbI**, **yqhQ-yqhP**, **yqhL**, **ypbD**, **rluB**, **menA**, **menH**, **ywzC**, **yfkB**, **yclN-yclO-yclP**, **yfmL**, **ygaF**, **pnpA**, **yqfF**, **accD-accA**, **opuD**, **ybbP**, **gmk**, **ytpQ**, **secDF**, **yvoF**, **yocJ**, **yvcI**-*yvcJ-yvcK-yvcL-crh-***yvcN**, **tgt**, **yobF**, **yrzD**, **sdaAB-sdaAA**, **ypdA**, **ytzA**, **ymcB**, **ripX**, **ask**, **hepT**, **ytxC**, **yncF**, **ywiE**, **yrpC**, **yrzL**, **ypmA**, **ybaL**, **yabA-yabB-yazA-yabC**, **mutSB**, **pcrB**-*pcrA-***ligA**-*yerH*, **alaR-alaT**, *rpsB-tsf-pyrH-***frr**, **tmk**, **metS**, **ytcI**, **senS**, **dltA**-*dltB-dltC-dltD-dltE*, **ytpA**, **ytpS**, **ytpT**, **zwf**, **hisJ**, **yqfO**, **yqxC**, **divIC**, **yhaJ**, **yneS**, **rocE-rocF**, **panD**, **yccF**, **yoaA**, **dgk**, **ytxB**, **polC**, **ytoP**, **yrzC**, **mbl**, **yybN**, **ylaI**, **ysmB**, **ytlQ**, **ytqI**, **atpI**, **yjzD**, **kamA**, **ykvY**, **ykkD**, **yheA**, **ypmP**, **proA**, **yojG**, **ykuQ**, **yqzC**, **gid**, **yrrB**, **yckJ**, **yloS**, **ymfL**, **ykbA**, **yhaK**, **yfkA**, **yusV**, **ykvL**, **mlpA**, **yhfQ**, **ytkP**, **queA**, **mutM**, **yrvC**, **serA**, **yshE**, **ytiA**, **ytbJ**, **ywdF**, **ykvK**, **yerB**, **topB**, **uppS**, **ygxA**, **parE**, **rpmB**, **dxs**-*yqxC-ahrC-recN*, **penP**, **spo0J**, **yqfN**, **ykvM**, **yfhC**, **asnB**, **rnmV-ksgA**, **ymzA**, **birA**, **ymfC**, **ymfG**, **ypmT**, **leuS**

3.2 Validation of computationally predicted interactions – σ^A

Our *in vitro* transcription system (see Chapter 1) was used to verify kinetically predicted interactions between σ^A and σ^A-dependent promoters. To test the validity of these predictions, we selected the upstream regions of 10 genes that were predicted to be σ^A-dependent and whose kinetic profiles correlated perfectly with the σ^A profile (**Table 6**). The upstream regions belonged to genes that were either monocistronic or positioned as the first genes of respective operons. For each upstream region, we prepared two PCR fragments. The shorter fragment lacked approximately 30 bp from the 3' end in the direction of expected transcription. This was to distinguish the direction of transcription from the DNA fragments. Finally, we performed *in vitro* assays in a defined system with *B. subtilis* RNAPσ^A holoenzyme with these fragments as templates. As a positive control, we used the strong constitutive σ^A-dependent P*veg* promoter (Fukushima et al., 2003). As a negative control, we performed a blank transcription assay with the templates and the RNAP core only.

Transcription signals were then obtained with RNAP reconstituted with σ^A for 5 out of the 10 tested gene upstream regions, thus suggesting the presence of σ^A-dependent promoters for *acpA*, *fbaA*, *rpmGA*, *ykpA*, and *yyaF*, in the expected orientation in each case (**Figure 28**).

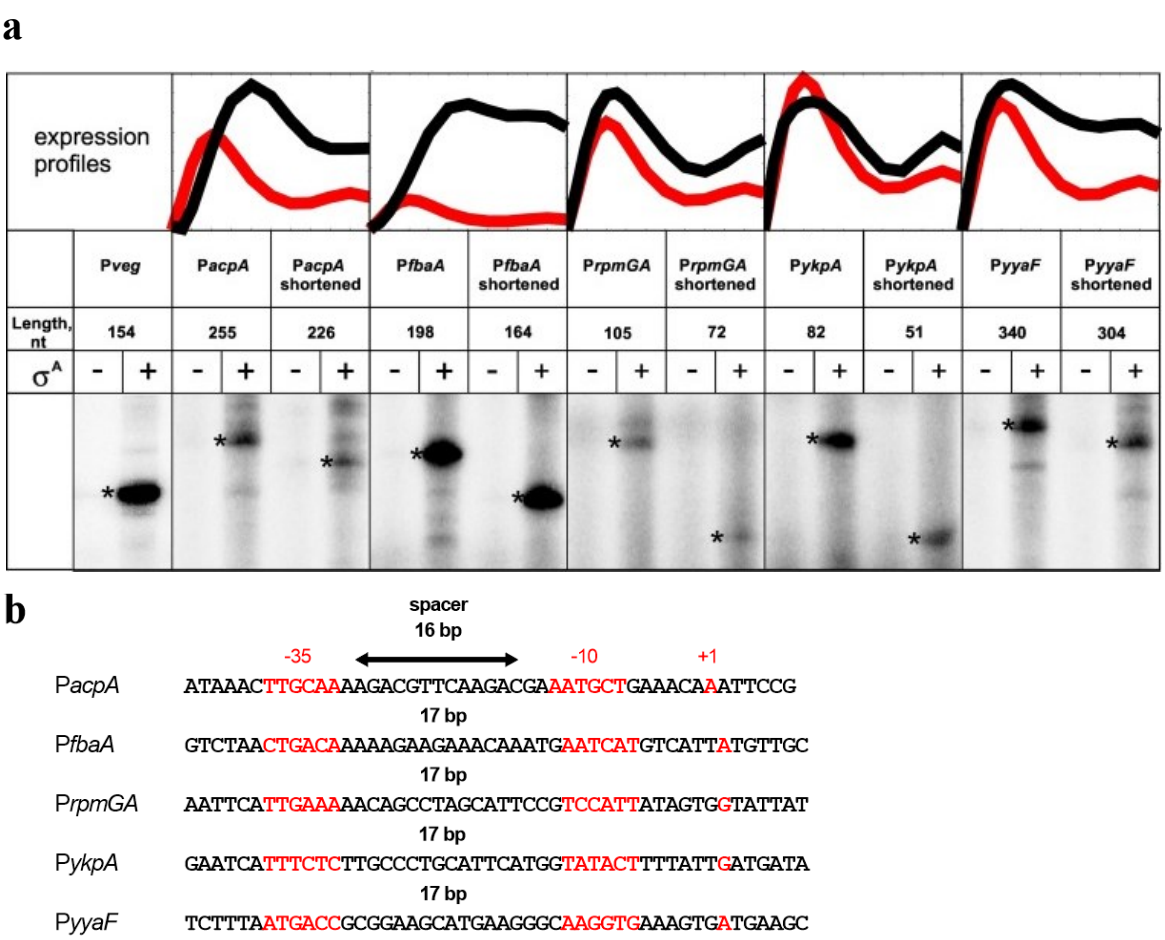


Figure 28. *In vitro* multiple round assays with DNA fragments derived from upstream regions of genes predicted to be σ^A-dependent. (a) Representative primary data. Each reaction was performed with the RNAP core and σ^A-containing holoenzyme to demonstrate that the core had not been contaminated with σ factors prior to its reconstitution with σ^A. Radioactively labeled samples were loaded onto polyacrylamide gels. P*veg* was used as a positive control. Transcript length was calculated with an RNA ladder (data not shown). The differences in length between the long and shortened fragment variants were in the 29–36 bp range. The distance between the long and shortened transcript variants in the gel differed for different promoters because each transcript had a unique length. Asterisks indicate the specific transcripts. The upper part of the figure shows respective kinetic modelling results. Red, σ^A mRNA; black, specific gene. The scaling between boxes varies to accommodate the graphs as the levels of specific transcripts differed over a wide range. (b) Alignment of putative promoter sequences identified by *in vitro* transcription assays. The − 35 and − 10 hexamers and the transcription start sites (+ 1) are indicated in red. Spacer regions between − 35 and − 10 hexamers are indicated.

The lack of transcription from the remaining five upstream regions (data not shown) might be a consequence of a requirement for additional transcription factor(s), different reaction conditions or the absence of σ^A-dependent promoters in these DNA fragments.

3.3 Analysis of σ^A-dependent promoter sequences

We performed sequence analysis of σ^A-dependent promoters of the genes that were either approved or disapproved by kinetic analysis to be regulated by σ^A during germination and outgrowth (**Tables 6** and **7**). Based on 70 known σ^A-dependent promoters for genes from **Table 6** (those consistent with our kinetic analysis) we created a sequence logo (**Figure 29b)**. A sequence logo, based on 43 known promoters for genes from **Table 7** (those inconsistent with our kinetic analysis) is shown in (**Figure 29c**). Both these logos resemble the M14 consensus sequence of σ^A-dependent promoters (**Figure 29a**) as reported by Nicolas and colleagues (Nicolas et al., 2012). A sequence logo created from the five newly identified promoters (**Figure 29d**) is consistent with the M14 logo, but we note that the size of the sample was too small for reliable analysis.

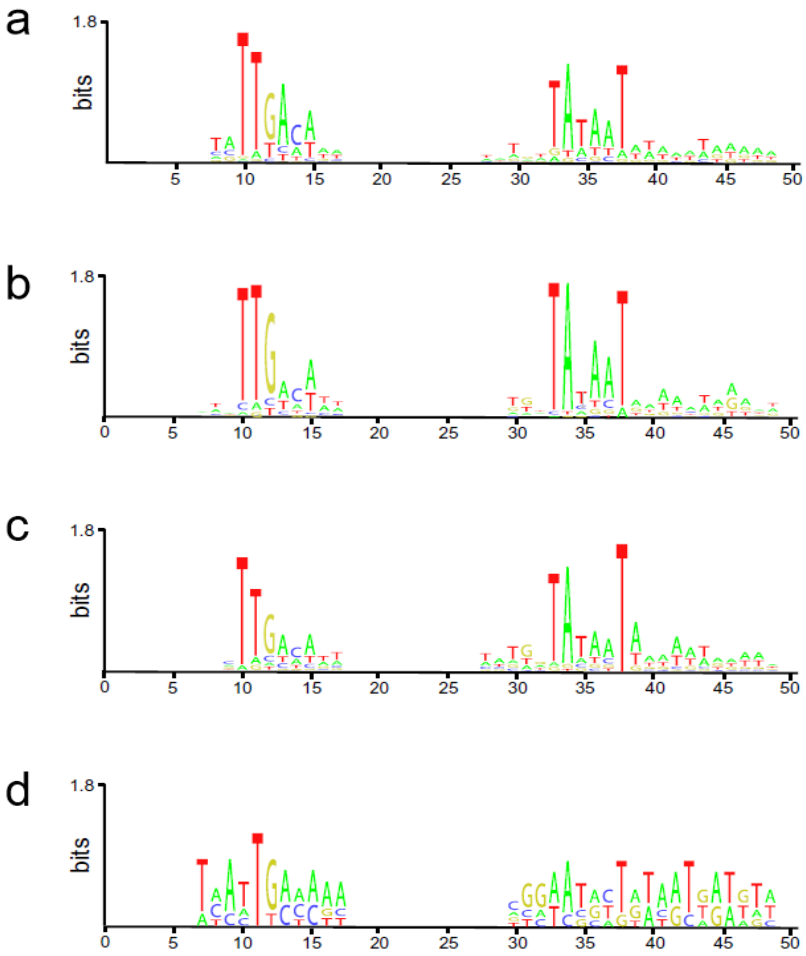


Figure 29. Comparisons of sequence logos. (a) The M14 motif of σ^A-dependent promoters according to Nicolas and colleagues (Nicolas et al., 2012). (b) The sequence motif of known σ^A-dependent promoters verified by the kinetic model. The logo was created from sequences listed in Table 6. The numbering of the horizontal axis was done according to Nicolas and colleagues. (c) The sequence motif of known σ^A-dependent promoters that were not verified by the kinetic model. The logo was created from sequences listed in Table 7. (d) The sequence logo for the experimentally verified σ^A dependent promoters from Figure 28.

3.4 Validation of computationally predicted interactions – σ^B, σ^D, σ^H

The σ^A regulon also contains genes of several σ factors: σ^E, σ^D, σ^H, σ^M, σ^X, σ^F (also σ^H regulon), YlaC, and σ^G (also σ^F regulon). We excluded σ^E, σ^G, σ^F, and YlaC from our analysis because their expression profiles had very low overall values and a high variance that might have caused false results during modelling. The expression profiles (respective mRNAs) of the *sigA*, *sigD*, *sigH*, *sigM*, and *sigX* genes are shown in **Figure 30**, illustrating the prominent role of σ^A in the regulation of their expression.

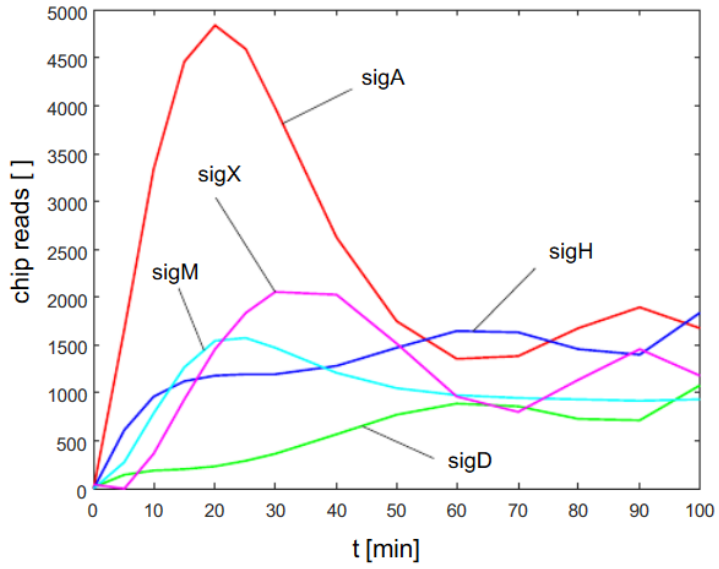
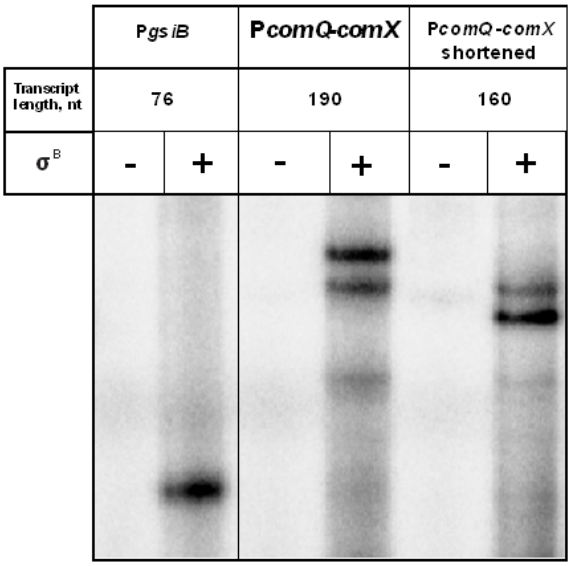


Figure 30. Expression profiles (mRNAs) of σ factors from the σ^A regulon during germination and outgrowth. Red — σ^A, green — σ^D, blue — σ^H, cyan — σ^M, magenta — σ^X. Horizontal axis — time in minutes, vertical axis — microarray units.

Our colleagues analyzed regulons of these σ factors and modelled the kinetic interactions between individual σ factors and their target genes. In a similar way as for putative σ^A-dependent promoters, we performed verification of selected predicted interactions for alternative σ factors σ^B, σ^D, σ^H (Results – Chapter 1). We tested seven putative σ^B-dependent promoters (of the *ald*, *comQ-comX*, *groES*, *lytA*, *xpt-pbuX*, *purE*, *rsbR* genes) at *in vitro* assays. Out of seven the tested promoters we obtained transcription signal only with the P*comQ-comX* template (**Figure 31**). We repeatedly obtained two bands with the P*comQ-comX* template, which most likely indicated two σ^B-dependent promoters in the tested area. In accordance with this result, we found two σ^B-like consensus sequences in the tested DNA fragment (**Figure 31b**).

a



b

GTAAAACAATTGTTATTCCGACTCGAACTTGATATTAAAGAAACGACAGATTCATTACGAAACATTAACAAAAGCATT
GATCAACTCGATAAATACAATTATGCAATGAAAATTTCGTGAAAAAGACTTGGAAACAAGTCTTTTTTTTCGTTCTAC
CGATACAATAAATGGATAAAGTATTATATGATTGTTAAAAAACGAAAAACCTGCTGTCCTTTAAATGTCCCATTTAGT
AAAATGGAATGGGAGGGGGGAAGTCGTTATTGAGCAGATATGTTTAGATTCTGTCCGGATCAAGGAGAAAATGAAGGA
GATTGTGGAGCAAAACATATTTAACGAGGATTTGTCACAACTTC

Figure 31. ***In vitro*** **multiple round assays with DNA fragment derived from the upstream region of the** ***comQ-comX*** **operon.** (a) Representative primary data are shown. Each reaction was performed with the RNAP core and σ^B-containing holoenzyme to demonstrate that the core had not been contaminated with σ factors prior to its reconstitution with σ^B. Radioactively labeled samples were loaded onto polyacrylamide gels. P*gsiB* was used as a positive control. Transcript length was calculated with an RNA ladder (data not shown). The difference in length between the long and shortened fragment variants is 30 bp. (b) The sequence of PCR product of the P*comQ-comX* region used at *in vitro* assay in (a). Two putative σ^B-dependent promoters are indicated: first is underlined, second is in bold. The initiation codon is indicated in green.

We obtained no transcription signals with six promoter regions of putative σ^D-dependent promoters (P*ctaA*, P*cysK*, P*fur*, P*hemZ*, P*ycdA*, P*pucR*), and the result was equally negative with six putative σ^H-dependent promoters (P*mecA*, P*hrcA*, P*fur*, P*hemZ*, P*ycdA*, P*pucR*) [data not shown].

Discussion 3

In this study, we developed a complex approach combining *in silico* (kinetic modeling and meta-analysis) and *in vitro* (transcription system) methods to reveal the σ^A-regulated network of *B. subtilis* during spore germination and outgrowth. The *in silico* part was performed by the J. Vohradský team, while the *in vitro* part and analysis was performed by us. Together, we analyzed the expression of *B. subtilis* genes during spore germination and outgrowth, as measured by microarrays in a unique time course experiment consisting of 14 time points spaced at 5-10-min intervals [see Materials and methods; (Keijser et al., 2007)]. Based on the microarray data, we created a time series of gene expression for σ^A, other σ factors from the σ^A regulon, and their target genes. The time series were then subjected to kinetic analysis based on a computational model of gene expression (see Materials and methods).

Of the 850 genes in the σ^A regulon suggested by SubtiWiki, 311 expression profiles were kinetically analyzed, 190 of which were confirmed as possible target genes of σ^A that also satisfied the kinetics-based criteria for the conditions covered by the microarray time series. Subsequent analysis of the remaining time series data suggested another 214 genes as putative targets of σ^A. Using *in vitro* transcription assays we demonstrated that five of the ten tested newly predicted σ^A-dependent genes are indeed transcribed by RNAP reconstituted with σ^A. The data complemented the information of the σ^A regulation network in *B. subtilis* during germination and outgrowth and extended the list of known σ^A-dependent genes.

Further, we determined a promoter sequence logo for σ^A-dependent genes that were found to be controlled by σ^A during germination and outgrowth (**Figure 29b**), closely matching the M14 logo reported by Nicolas and colleagues (Nicolas et al., 2012). However, the promoter logo of σ^A dependent genes that were found not under control of σ^A was almost identical (**Figure 29c**), indicating additional level of control for the genes whose profiles could not be modelled with σ^A. Taken together, the results show that for the σ^A regulon, ~ 60% of the gene expression profiles analyzed were consistent with the kinetic analysis and directly regulated by σ^A without the requirement for additional regulators or modulators. The remaining 40% displayed more complex kinetics. This kinetics was not dependent on the core promoter sequence but rather could be explained by additional layers of regulation, such as alternative σ or/and other transcription

factors. Thus, the kinetic analysis is clearly able to distinguish between genes where the dominant regulator is the σ factor and where other regulators must be involved.

Altogether, our modelling was able to

- Discriminate for possible regulatory interactions among the σ factors and their putative target genes;
- Optimize parameters of the model that could be used to computationally simulate the accumulation of mRNA of a gene under the control of a specific σ factor;
- Find other possible regulators of specific genes or suggest a new mechanism of control of a gene by computing the profile of an unknown regulator that could explain the observed expression profile;
- Suggest new σ^A-dependent genes (selected predictions were subsequently validated experimentally);
- Create a specific σ^A-controlled gene expression network that is active under the conditions measured by the microarray time series;
- Identify promoter sequence logos associated with the σ^A-dependent promoters that are, or are not, dominantly regulated by σ^A during germination and outgrowth.

However, our approach has its limitations: alternative modes of control such as anti- and anti-anti-σ factors or others (*e.g.*, attenuation, additional transcription factors) were ignored in this study, although such modes of control could help explain some of the observed kinetics. This could be the reason of the absence of the transcription signal in the assays with alternative σ factors σ^B, σ^D and σ^H.

To sum up, we developed a complex approach that involves computational modelling and experimental validation of predicted interactions. This approach can be used for processing/analyses of extensively generated NGS-data from the past decades. Majority of information contained in public repositories remains unexplored. We strongly believe that our analysis combining *in vivo* and *in vitro* approaches will help restore order in this huge amount of data.

4. RNAP binding partners HelD and δ in transcription with primary and alternative σ factors

An interesting, yet poorly explored aspect of bacterial transcription is cooperation between *B. subtilis* σ factors and non-essential binding partners of RNAP from Gram-positive bacteria – HelD and δ. Both – HelD and δ – have been extensively studied in our laboratory. HelD is an ATP-dependent helicase from *B. subtilis* that was discovered in our laboratory and in parallel by Olivier Delumeau (Delumeau et al., 2011). δ is a non-essential RNAP subunit of G+ bacteria (Literature review – Chapter 2.4) that increases transcription driven with σ^A; it is important for cell survival under non-favorable conditions (Achberger and Whiteley, 1981; Rabatinová et al., 2013; Spiegelman et al., 1978; Weiss and Shaw, 2015). Structurally, δ is a polyanion comprising two parts: structured N-terminal (δN) and largely unstructured and highly charged C-terminal (δC) part (Lopez de Saro et al., 1995).

Data presented in this Chapter are either part of an already published article about the HelD protein "Characterization of HelD, an interacting partner of RNA polymerase from *Bacillus subtilis*" in Nucleic Acids Research [Appendix 3] (Wiedermannová et al., 2014) or were generated for a project focused on the δ subunit. The experiments were performed with the main *B. subtilis* σ factor – σ^A – and, where indicated, with two alternative σ factors – σ^B and σ^D. **The aim of this study was to characterize the effect of RNAP binding partners HelD and δ on *in vitro* transcription with main and selected alternative σ factors.**

The objectives of this project were:

- Establishment of a transcription system containing alternative σ factors, HelD and δ proteins;
- Characterization of δ and HelD effects on *in vitro* transcription with alternative σ factors;
- Determining the effects of truncating the δ protein (δN) on *in vitro* transcription;
- Revealing the effect of [iNTP] on transcription with δ and δN.

4.1 HelD does not affect the initial steps of transcription

In our lab we showed that HelD stimulates transcription in an ATP-dependent manner (Wiedermannová et al., 2014); Appendix 3. Consequently, we wanted to know, which part of the transcription process HelD affects. Having established that HelD binds to the RNAP core in the cell, it seemed that it affects elongation and/or termination. Nevertheless, we decided to test for an effect of HelD on an early phase of transcription initiation – affinity of RNAP for promoter DNA (**Figure 32**; Wiedermannová et al., 2014). *In vitro* transcription assays were performed in the presence/absence of HelD with increasing concentration of the σ^A-dependent promoter P*veg*. The results showed that HelD does not affect the affinity of RNAP to promoter DNA, thus does not affect the initial phase of transcription.

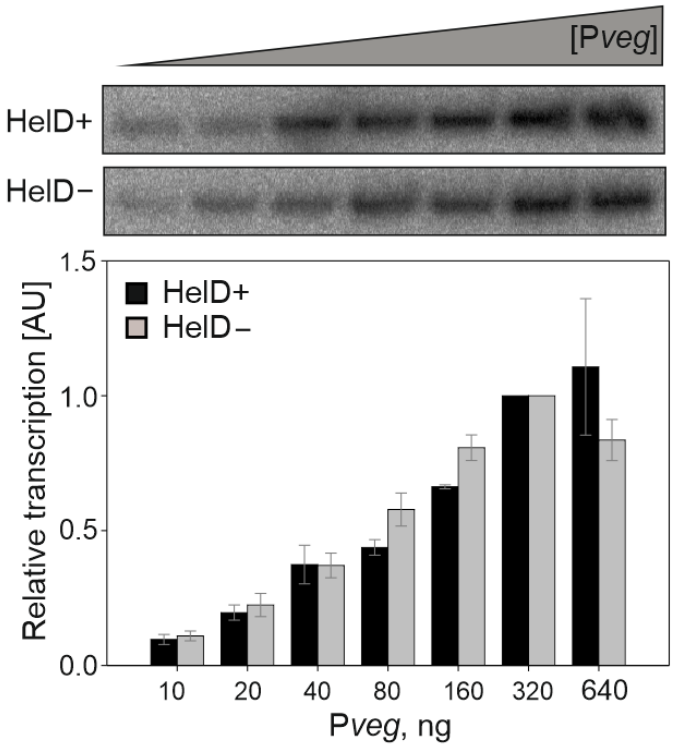


Figure 32. HelD does not affect the affinity of RNAP to promoter DNA. RNAP (LK#782) was reconstituted for 15 min at 30°C with σ^A (molar ratio 1:5). RNAP was (black bars) or was not (grey bars) reconstituted with HelD and used in transcription assays with increasing amounts of template DNA (P*veg*, 10-640 ng). When reconstituted with HelD (15 min at 30°C), molar ratio RNAP:HelD was 1:4. The signal corresponding with 320 ng of P*veg* that was set as 1, the data were normalized to this value. Graphs represent data from three independent experiments ±SD.

4.2 HelD affects *in vitro* transcription with σ^A and alternative σ factors σ^B and σ^D in different ways

After we concluded that HelD does not affect the affinity of the RNAP to promoter DNA, we tested the effect of HelD on transcription with main and alternative σ factors. We performed titration by HelD on σ^A-dependent promoters P*veg*, P*glpD* and P*helD* (**Figure 33a**; Wiedermannova et al., 2013, Figure S5) and σ^B-dependent P*gsiB* and σ^D-dependent P*motA* (**Figure 33b**). The results showed that at a saturated concentration (ratio RNAP:HelD=1:4) HelD stimulated σ^A-dependent transcription from three different promoters in the same manner (**Figure 33a**).

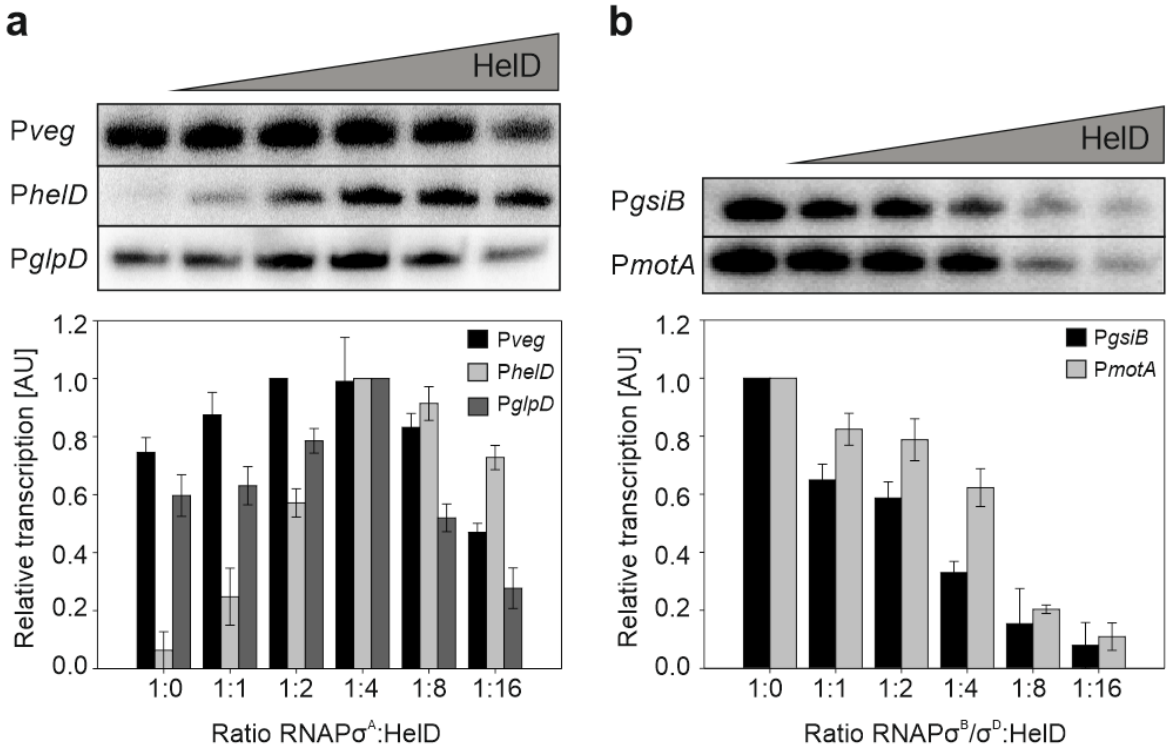


Figure 33. Effect of HelD on *in vitro* transcription with σ^A, σ^B and σ^D. (a) RNAP core (from HelD knock out strain LK#782) was reconstituted with saturating concentrations of σ^A. The holoenzyme was then incubated with increasing amounts (molar ratio from 1:0 to 1:16) of HelD and used to initiate transcription. The template plasmids with either the P*veg* (LK#1177, black bar), P*helD* (LK#1109, light gray bar) or P*glpD* (LK#888, dark grey bar) promoters were used. (b) RNAP core was reconstituted with saturating concentrations of σ^B or σ^D. The holoenzyme was then incubated with increasing amounts (molar ratio from 1:0 to 1:16) of HelD. The template plasmids with either the P*gsiB* (σ^B-dependent promoter, LK#1230, black bar) or P*motA* (σ^D-dependent promoter, LK#1233, light grey bar) were used. Primary data from representative experiment are shown in the upper panel. The data were normalized to the maximum signal that was set as 1. Graphs represent data from three independent experiments ±SD.

To the contrary, the results of assays with alternative σ factors were different: they repeatedly showed that increasing HelD concentration inhibited transcription with σ^B and σ^D, and that excess

of HelD in the assay almost eradicated transcription (**Figure 33b**). These results suggested different roles of HelD in transcription with primary and alternative σ factors.

4.3 δ stimulates σ^B-dependent transcription

To determine the effect of δ on transcription with alternative σ factors, we selected a transcription system containing σ^B. Titration by δ was performed with four σ^B-dependent promoters – P*gsiB*, P*ybyB*, P*trxA*, P*mgsR* (**Figure 34**).

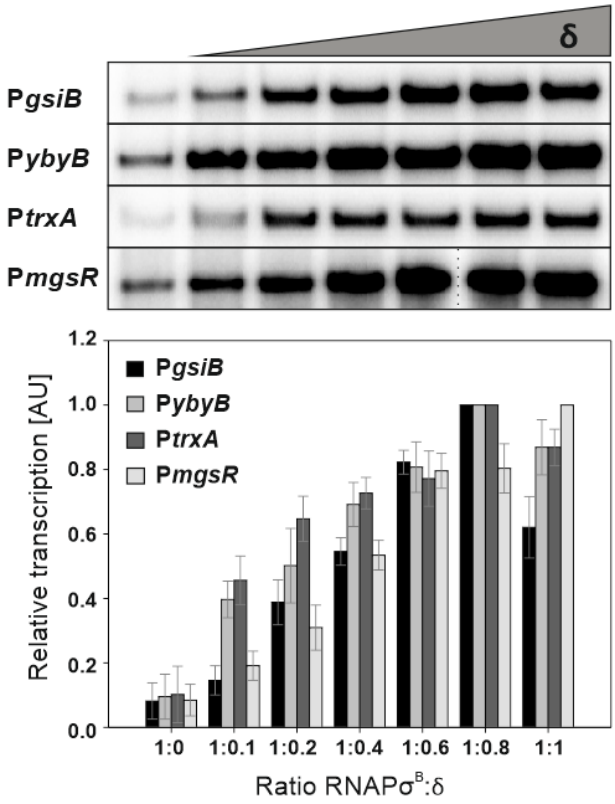


Figure 34. Stimulating effect of δ on σ^B-dependent *in vitro* transcription. RNAP core (from δ knock out strain LK#1277) was reconstituted with saturating concentration of σ^B (ratio RNAP:σ^B = 1:5). The holoenzyme was then incubated with increasing amounts (molar ratio from 1:0 to 1:1) of δ and used to initiate transcription. The template plasmids with the P*gsiB* (LK#1230), P*ybyB* (LK#1585), P*trxA* (LK#1231) and P*mgsR* (LK#1584) promoters were used. Primary data from the representative experiment are shown in the upper panel. The data were normalized to the maximum signal that was set as 1. The vertical dotted line represents non-adjacent lanes on the same gel. Graphs represent data from two independent experiments ±SD.

The obtained results clearly showed that δ stimulates transcription in a concentration-dependent manner. The saturated ratio of RNAP holoenzyme:δ is approximately 1:1. Thus, σ^B-dependent transcription is stimulated by δ in a similar fashion as σ^A-dependent transcription (Achberger and Whiteley, 1981; Dobinson and Spiegelman, 1987).

4.4 The C-terminal domain of δ is crucial for the stimulatory effect of δ on σ^B-dependent promoters

After revealing the stimulatory effect of δ on σ^B-dependent transcription, we decided to identify which domain of the δ protein is essential for this phenomenon – the structured N-terminal (δN) or the unstructured and highly charged C-terminal (δC). For this purpose, we cloned and purified δN – a truncated δ subunit lacking the C-terminal domain and containing only the highly organized N-terminal domain using 6xHis-tag (cloning was performed by A. Rabatinová). In the presence/absence of δ and δN, we performed assays with four σ^B-dependent promoters (P*gsiB*, P*ybyB*, P*trxA*, P*mgsR*). Transcription assay with σ^A and its dependent promoter P*veg* was used as a control.

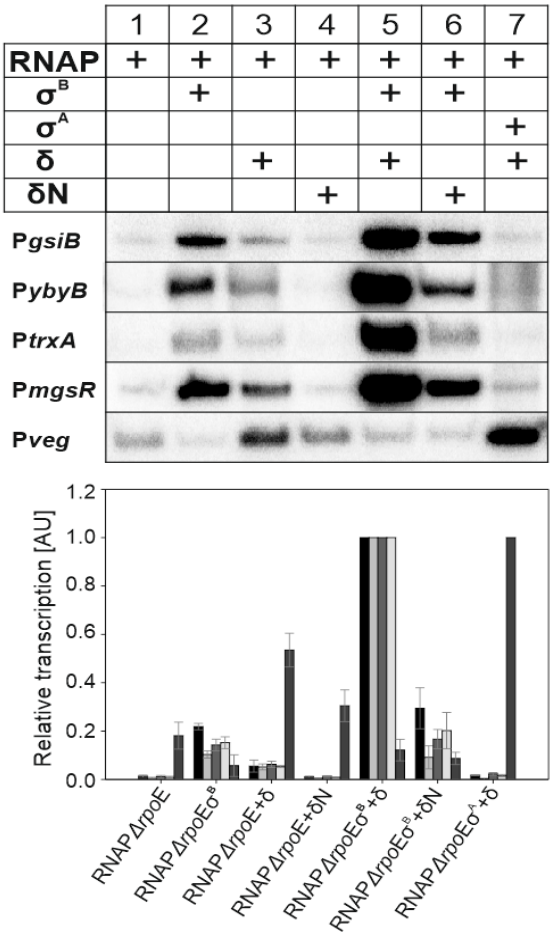


Figure 35. δC is crucial for stimulation of σ^B-dependent transcription. Multiple-rounds *in vitro* transcription with σ^B- (P*gsiB*, P*ybyB*, P*trxA*, P*mgsR*) and σ^A- (P*veg*) dependent promoters. RNAP was reconstituted with σ^A and σ^B at 30°C for 15 min (ratio E:σ was 1:5). Blank (dilution buffer)/δ/δN protein was added to appropriate tubes, reconstitution at 30° C lasted for additional 15 min (ratio Eσ:δ/δN was 1:1). Primary data from the representative experiment are shown in the upper panel. The data were normalized to the maximum signal that was set as 1. Graphs represent data from three independent experiments ±SD.

Transcription signal obtained in the assay containing truncated δN was comparable with those completely lacking δ (**Figure 35,** lanes 2 and 6). To the contrary, we obtained a strong signal in the assay containing full δ protein (**Figure 35**, lane 5). Thus, the unstructured δC domain is crucial for the σ^B-dependent transcription stimulation. The results were consistent with those from Chapter 4.3.

4.5 δ modulates promoter sensitivity to [iNTP] *in vitro*

In our previous study (Rabatinová et al., 2013) we showed that the δ subunit potentiates σ^A-dependent promoter regulation by [iNTP]. After experimental confirmation that the δC domain has a crucial effect on the transcription stimulation, we compared the effects of δ and δN in transcription with [iNTP]-sensitive promoters. In the experiments we used three forms of RNAP*ΔrpoE*: lacking δ, reconstituted with δ or δN proteins. With each type of RNAP we performed titration by [iNTP] at σ^A-dependent P*rrnB* P1 ribosomal promoter (sensitive to [iNTP]) and at P*veg* (insensitive to [iNTP]). Assay with P*veg* served as a control.

The obtained results showed that the presence of the δ increased overall transcription from the P*veg* promoter (**Figure 36a**). This observation was in agreement with previously published results (Rabatinová et al., 2013). More importantly, we observed that δ protein modulates sensitivity of [iNTP]-sensitive promoter P*rrnB* P1 to [iNTP]: in the presence of δ P*rrnB* P1 clearly displayed high sensitivity to [iNTP] (**Figure 36b**). On the other hand, we observed only partial effect on P*rrnB* P1 sensitivity to [iNTP] when RNAP was reconstituted with δN. Thus, δ protein is essential for modulating the activity of [iNTP]-sensitive promoter; δC domain plays important role in this effect. Similar assays with other σ factors remain to be done.

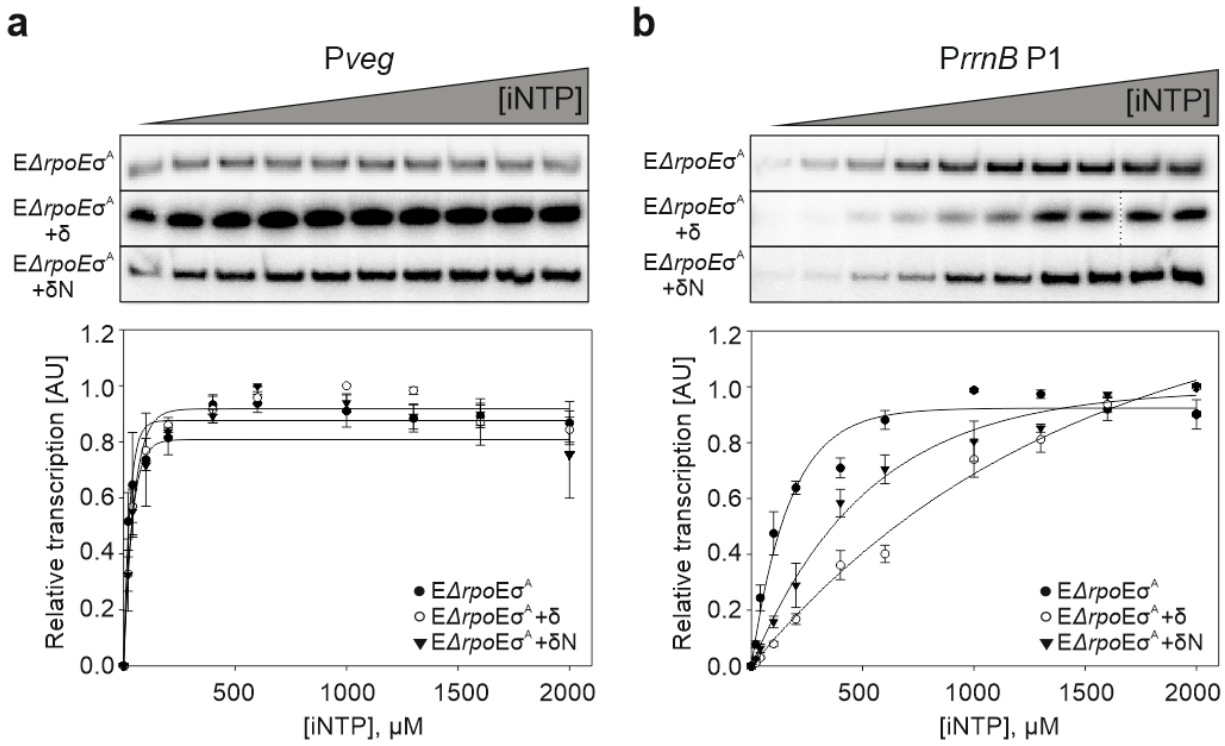


Figure 36. C-terminal domain of δ modulates the sensitivity of the P*rrnB* P1, but not the P*veg* at *in vitro* transcription. RNAP core (from δ knock out strain LK#1277) was reconstituted with saturating concentration of σ^A (ratio RNAP:σA = 1:5). The holoenzyme was then reconstituted with δ and δN proteins, or dilution buffer. Resulting holoenzyme containing either δ or δN, or lacking these proteins, was used to initiate transcription. The template plasmids with the P*veg* (LK#1177) and P*rrnB* P1 (LK#653) promoters were used. Primary data from the representative experiment are shown in the upper panel. The data were normalized to the maximum signal that was set as 1. The vertical dotted line represents non-adjacent lanes on the same gel. E – RNAP core; Eσ – RNAP holoenzyme. Graphs represent data from three independent experiments ±SD.

Discussion 4

In Chapter 4 we describe selected aspects of regulation of *B. subtilis* RNAP by its two binding partners HelD and δ. The main experimental tool was *in vitro* transcription with alternative σ factors.

HelD in transcription with primary and alternative σ factors

First, we showed that HelD does not affect the affinity of RNA polymerase for promoter DNA (**Figure 32**). This result was in agreement with the fact that HelD binds to the RNAP core, not to holoenzyme (Gwynn et al., 2013; Wiedermannová et al., 2014). Nevertheless, our next results were surprisingly inconsistent with previous observations: we revealed that HelD affects *in vitro* transcription with σ^B- and σ^D-dependent promoters differently than with σ^A-dependent promoters (**Figure 33**). Interestingly, an equal concentration of HelD that increased transcription with σ^A, strongly inhibited transcription with σ^B- and σ^D-dependent promoters. By now we do not have an explanation for this phenomenon, as our previous findings showed that HelD increases the transcription rate by stimulating elongation via binding only to RNAP core, but not the σ^A (Wiedermannová et al., 2014). The most possible reason for this contradiction is yet uncovered involvement of HelD in the transcription initiation driven by alternative σ factors. The difference in the molecular weight between primary and alternative σ factors is ~10 kDa (Helmann, 2016b; Narula et al., 2016), alternative σ factors lack auto-inhibitory 1.1 region [**Figure 9**] (Bowers and Dombroski, 1999). Thus, the contradictory effect we observed, might be displayed due to a slightly different conformation of an alternative σ factor attached to the RNAP, comparing to σ^A conformation, and as a consequence – additional properties of HelD in transcription. To test this hypothesis we are planning to perform (i) the binding assays of RNAP$\sigma^{alternative}$/promoters recognizable by alternative σ factors, including gel shifts (ii) promoter escape assays with alternative σ factors, (iii) [iNTP] sensitivity assays. If our hypothesis is correct, this suggests alternative ways of the HelD mode of action at the exponential vs. stationary phase of growth, when bacterial cell copes with various stresses.

Another possibility is that lower HelD concentration (not tested in our experiments) might be needed to enhance alternative σ factor-dependent transcription. Nevertheless, additional experiments are needed to reveal the mechanism of alternative σ factors-dependent transcription deprivation in the presence of HelD.

δ in transcription with primary and alternative σ factors

The stimulatory effect of HelD in σ^A-dependent transcription can be enhanced by the small subunit of RNAP, δ (Wiedermannová et al., 2014). In this study, we made an important observation that δ subunit stimulates σ^B-dependent transcription in a concentration-dependent manner, same as σ^A-dependent transcription. This fact corresponds to the previously made conclusion that δ is significant for a cell survival when coping with changing environmental conditions (Rabatinová et al., 2013). Importantly, we showed that unstructured domain of δ (δC) contributes greatly to the stimulation of σ^A- and σ^B-dependent transcription (**Figures 34 – 36)**. Apparently, this stimulation is a consequence of the δC involvement in the RNAP recycling process during transcription with both σ factors (Juang and Helmann, 1994b).

It was previously shown that δ subunit changes the stability of the RP_o at σ^A-dependent promoters P*rrnB* P1 and P*veg* (Rabatinová et al., 2013). In this study, we showed that both δ domains – unstructured δC and highly structured δN – are important for this phenomenon (**Figures 35** and **36**). The absence of δC domain rapidly decreased observed effect at sensitive promoter P*rrnB* P1 (**Figure 36b)**. This can be due to biophysical properties of unstructured δC, as it was previously shown to be a highly charged polyanion (Lopez de Saro et al., 1995) that facilitates RP_o formation (Prajapati et al., 2015).

We previously showed that HelD and δ act synergistically to stimulate transcription (Wiedermannová et al., 2014). Our further plan is to study the cumulated effect of HelD and δ in transcription with alternative σ factors.

5. Small RNA Ms1 from *M. smegmatis*: transcription start sites of Ms1 and MTS2823, and the activity of the P_{Ms1} promoter

At the end of my Ph.D. study, I participated in the project characterizing the role of the small RNA Ms1 from *M. smegmatis* that binds to the RNAP core (Chapter 5.4 in the Literature review).

My aims in the project were:

- Validation of **the transcription start sites (TSSs) in the MTS2823 sRNA from *M. tuberculosis*** and subsequent comparison with the TSS of its homologous Ms1 sRNA from *M. smegmatis*;
- Determination of the activity of **the P_{Ms1} promoter regions** comparing to the ribosomal promoter.

Data presented in Chapter 5 are part of a published article about the role of Ms1 sRNA from *M. smegmatis* (**Figure 40**): "Ms1 RNA increases the amount of RNA polymerase in *Mycobacterium smegmatis*" in the Molecular Microbiology [Appendix 4] (Šiková et al., 2018).

5.1 Transcription start sites of Ms1 from *M. smegmatis* and MTS2823 from *M. tuberculosis*

In our previous work, we identified the first nucleotide of Ms1 and the putative P_{Ms1} promoter in *M. smegmatis* (Hnilicová et al., 2014). Subsequently, we noticed that the putative promoter sequence of Ms1 is conserved among *M. smegmatis*, *M. bovis*, *M. avium* and *M. tuberculosis* (**Figure 38**). The TSSs of Ms1 homologs in *M. bovis* and *M. avium* (DiChiara et al., 2010; Ignatov et al., 2013) were mapped previously to the same or nearly identical positions; however, the 5' terminus of MTS2823 (the Ms1 homolog in *M. tuberculosis*) was mapped further upstream (**Figure 38**) and no promoter consensus sequence upstream of this nucleotide was reported (Arnvig et al., 2011).

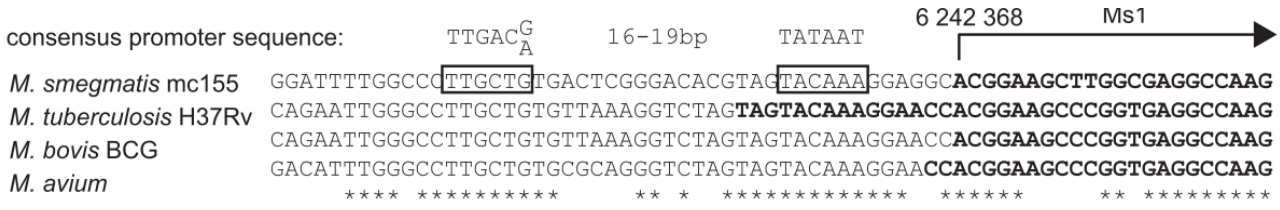


Figure 37. TSSs of Ms1 homologs from *M. tuberculosis*, *M. bovis,* and *M. avium*. The 5' end sequences of previously identified Ms1 homologs in these species are highlighted in bold. The first nucleotide of *M. smegmatis* Ms1 is adenine transcribed from position 6 242 368 in the genome. The putative -10 and -35 promoter sequences are framed in *M. smegmatis* sequence. Asterisks indicate conserved motifs. The picture was created by J. Hnilicová (Hnilicová et al., 2014).

We performed 5' RACE to define the exact TSS position of MTS2823 from pathogenic *M. tuberculosis* and compare it to *M. smegmatis* Ms1 TSS (Hnilicová et al., 2014). The results showed that MTS2823 has TSS identical to the Ms1. Thus we rebut previous data of MTS2823 TSS and showed the exact position of MTS2823 +1 (**Figure 39**).

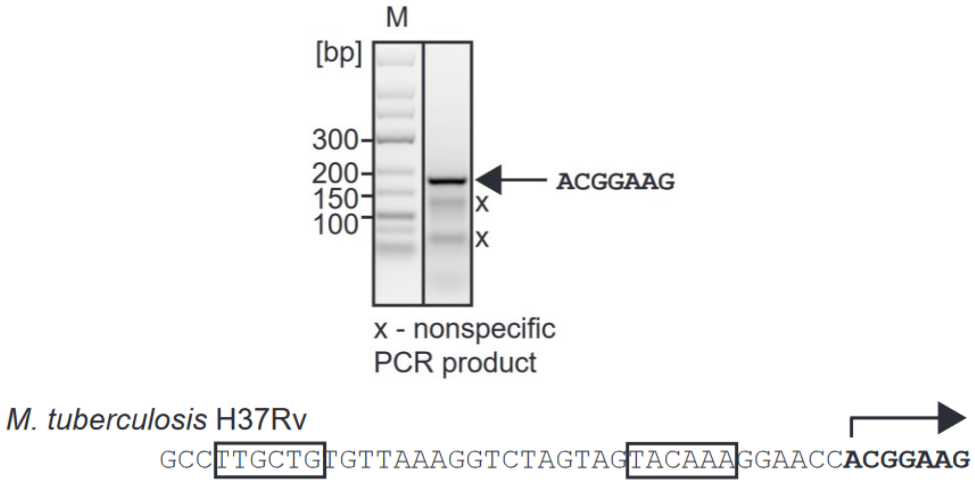


Figure 38. 5' RACE of *M. tuberculosis* MTS2823. PCR products of 5' RACE were resolved on an agarose gel and sequenced. The arrow indicates the specific band and its sequence from the 5' end. Putative promoter elements of the Ms1 homolog MTS2823 in *M. tuberculosis* H37Rv are shown in boxes, the transcription start site (+1) identified by 5' RACE is indicated with the horizontal arrow.

5.2 Activity of the P_{Ms1} promoter

To identify the Ms1 promoter and subsequently define its activity, we cloned several *M. smegmatis* promoter region fragments (-38/+9 - the putative core promoter; -131/+9; -231/+9; -331/+9; and -491/+9, see **Figure 37a**) into the integrative pSM128 vector encoding the *lacZ* reporter gene. Further, we measured the β-galactosidase activity of P_{Ms1} variants in *M. smegmatis* comparing to the ribosomal promoter (**Figure 37b**). The β-galactosidase activity of the -38/+9 region was significantly above the background and was set as 100. This confirmed that the -38/+9 region contained a core promoter sequence. The relative β-galactosidase activities of extended promoter fragments were normalized to the activity of the -38/+9 core promoter in exponential phase.

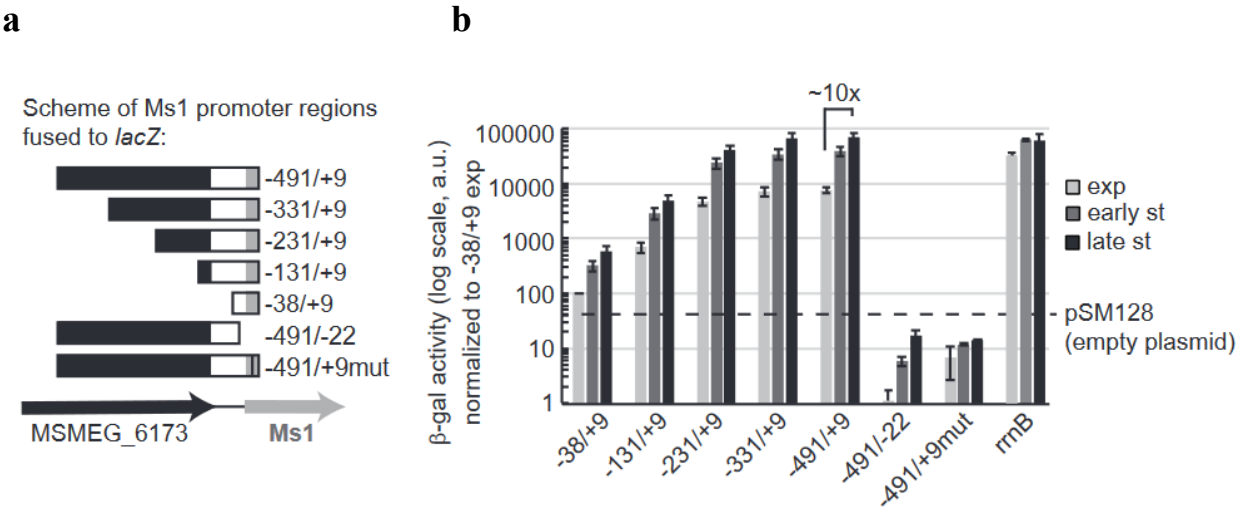


Figure 39. Identification of the Ms1 promoter. (a) Ms1 promoter region fragments differing in length of the upstream region were fused to *lacZ*. Shortened and mutated Ms1 promoter variants (-491/-22; -491/+9mut) were used to show that Ms1 was transcribed from a single promoter, P_{Ms1}. (b) The graph shows results of β-galactosidase assays for constructs shown in (a). β-galactosidase activity was measured in exponential (exp, OD_{600} 0.5, ~6 hrs of growth), early stationary (OD_{600} 2-3, ~24 hrs of growth) and late stationary (OD_{600} 2, ~48 hrs of growth) phases. Data were normalized to the value for the core promoter (-38/+9 construct) in an exponential phase that was set to 100. The averages from at least three independent biological experiments performed in duplicates are shown with ±SEM. The dashed horizontal line represents β-galactosidase activity measured in control cells with an empty pSM128 plasmid. *rrnB* - ribosomal promoter (-72/+20), a.u. - arbitrary units.

The results of the β-galactosidase measurements showed that the activity of the P_{Ms1} core promoter was highest in late stationary phase (after 48 hours of growth). Together with the gradual addition of ~100 bp long sequences of the native promoter upstream from the core (**Figure 37a**), promoter activity increased both in exponential and stationary phases ~10-fold. Addition of 100 bp to the promoter fragment (-231/+9) led to an additional ~8-fold increase promoter activity in both growth phases. Following two sequential extensions by 100 bp (-331/+9 and -491/+9) had only a minor non-specific effect (max. ~1.5-fold transcription increase) [**Figure 37b**]. The

measurements were performed in parallel with the strong ribosomal promoter P_{rrnB}. We concluded that the P_{Ms1} is a a very strong, almost constitutive promoter, and its activity in stationary phase is comparable to the ribosomal promoter P_{rrnB} (- 72/+20 fragment) (**Figure 37b**). Besides, as P_{Ms1} activity was increasing with increasing length of the promoter region tested, we concluded that additional transcription factors probably affect transcription from P_{Ms1}.

Discussion 5

The results described in Chapter 5 are part of the project focused on a novel regulator of gene expression in *Mycobacterium smegmatis* – small RNA Ms1. The research was performed with the non-pathogenic species *M. smegmatis*, and with pathogenic *M. tuberculosis*.

We tested in *M. smegmatis* (i) whether *M. smegmatis* Ms1 and *M. tuberculosis* MTS2823 transcripts have the same transcription start sites (TSSs), and therefore might be similarly regulated by their upstream sequences; (ii) whether the predicted Ms1 core promoter (Hnilicová et al., 2014) was functional, and asked when it was expressed the most and compared in activity to the ribosomal promoter.

First, we compared the TSSs from *M. smegmatis* Ms1 (Hnilicová et al., 2014) and its homologous MTS2823 from *M. tuberculosis*. In *M. tuberculosis*, the 5' end of MTS2823 was mapped 15 nucleotides upstream compared to Ms1 from *M. smegmatis* (**Figure 37**) (Arnvig et al., 2011). The region upstream of the reported MTS2823 TSS contains no obvious promoter consensus sequence. Moreover, no promoter sequence for the Ms1 homolog in *M. tuberculosis* was reported (Arnvig et al., 2011). Nevertheless, a sequence identical with the *M. smegmatis* putative Ms1 promoter is present in the same location also in *M. tuberculosis*. Mapping of the TSS of MTS2823 from *M. tuberculosis* using 5' RACE showed that one dominant MTS2823-specific PCR product (**Figure 38**) corresponded to the same TSS as Ms1 in *M. smegmatis* (**Figure 37**). Thus, both Ms1 and MTS2823, have the same TSS and the promoter sequence we found in *M. smegmatis* is conserved also in *M. tuberculosis* (**Figure 38**).

Second, we measured the β-galactosidase activity of several P_{Ms1} promoter region fragments in *M. smegmatis*. The measurement was performed in different growth phases and in parallel with ribosomal promoter P_{rrnB} (**Figure 39**). The results showed that the activity of the P_{Ms1} core

promoter was highest in late stationary phase. Both, P_{Ms1} and P_{rrnB} core promoters, are rather weak and are activated by the upstream elements more than 300-fold (Arnvig et al., 2005). Nevertheless, the full-length P_{Ms1} appeared to be very strong, and its activity in stationary phase is comparable to the full-length P_{rrnB} (China et al., 2010; Tare et al., 2012). We concluded that transcriptional activators are needed for activation of both promoters and that these proteins are necessary for expression of the most abundant mycobacterial RNAs. Also, we showed that Ms1 in *M. smegmatis* is transcribed from a single promoter and that the -35 and -10 elements of P_{Ms1} correspond to the σ^A binding site. Thus, P_{Ms1} is probably σ^A-dependent, similarly to 6S RNA in *E. coli* (Kim and Lee, 2004).

The overall concept of the Ms1 project is depicted in **Figure 40**. We showed that Ms1 is the most abundant non-rRNA transcript in stationary phase in *M. smegmatis*. The most important result we obtained using transcriptomic study is the mechanism responsible for Ms1 high accumulation: we revealed that Ms1 affects the intracellular level of RNAP through increasing the expression of two RNAP subunits, β and β'. Moreover, we identified the Ms1 promoter, P_{Ms1}, and cis-acting elements important for its activity. Next, we demonstrated that an RNase called PNPase contributes to the differential accumulation of Ms1 during growth. Finally, we showed that MTS2823 from pathogenic *M. tuberculosis* has the same TSS as *M. smegmatis* Ms1 (**Figure 40)**.

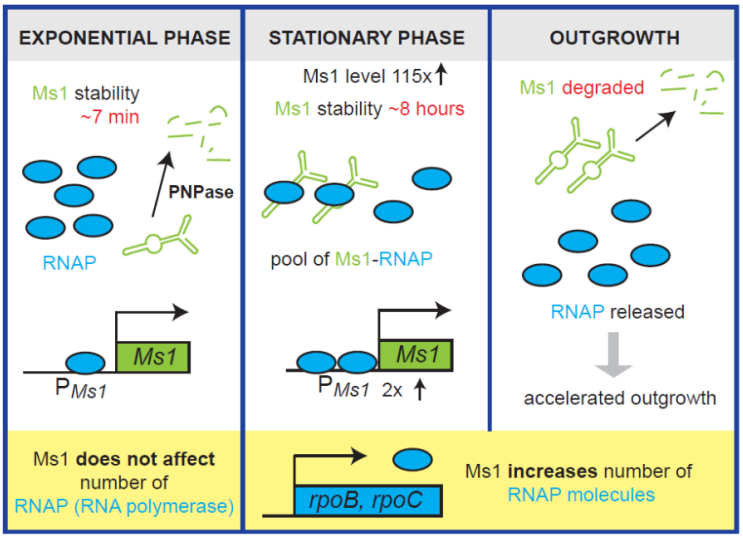


Figure 40. The concept of the paper from Šiková and Janoušková (2018) explaining the role of the Ms1 sRNA in *M. smegmatis* (Šiková et al., 2018).

SUMMARY AND FUTURE PROSPECTS

During my Ph.D. study, I participated in five projects aimed at regulation of initiation of bacterial transcription. The main focus of the research was on *B. subtilis* σ factors, and it also involved selected issues of HelD and δ interacting partners of *B. subtilis* RNAP, and of small RNA from *M. smegmatis*, Ms1.

In the first project, I set up *in vitro* transcription systems with *B. subtilis* alternative σ factors. Using these systems I showed that gene expression from *B. subtilis* promoters driven by alternative σ factors can be regulated by changes in concentrations of iNTPs. With further modifications, the established technique can be used to study various aspects of bacterial transcription *in vitro* (*e.g.* effect of diverse transcription factors on *in vitro* transcription with alternative σ factors).

The second, and the most significant, project was aimed at σ^I, one of the least explored *B. subtilis* alternative σ factors. It resulted in the comprehensive characterization of *B. subtilis* σ^I. I showed that the σ^I regulon is small and consists of 16 genes organized in eight operons, nevertheless, there are more than 100 genes indirectly affected by σ^I. Moreover, I revealed that σ^I is involved in iron metabolism, and that absence of σ^I leads to malformed cell shape at elevated temperature. The last important observation I made for the σ^I project was that σ^I promoter region contains a previously unknown "extended -10" element that is important for σ^I binding to promoter DNA. This project was published as my first-author publication in the Journal of Bacteriology (June 2018; Appendix 1). It will be highly interesting to further characterize the interactions of the extended promoter elements with alternative σ factor. This challenge can be solved in collaboration with bioinformaticians and crystallographers. Moreover, there are still other alternative σ factors from *B. subtilis* (YvrI-YvrHa and ECF-type σ^V, σ^Y, σ^Z, YlaC) with almost unknown regulons and poorly characterized functions, and their potential roles remain to be elucidated. Special attention should belong to the structural studies of alternative σ factors. This task is challenging mainly because of difficulties in their purification. Until now, none of the alternative σ factors from *B. subtilis* was crystallized.

The third project from systems biology area was performed in collaboration with colleagues-bioinformaticians, who are studying σ factors' regulatory network with a novel computational approach combining meta-analysis and kinetic modelling. I experimentally proved that the *in silico* modelling is a useful tool for revealing new σ factor interactions. This collaboration resulted in my first-author paper published in the BBA – Gene Regulatory Mechanisms (June 2017; Appendix

2). The project was extended to other regulons of *B. subtilis* alternative σ factors, I participated in this extension.

Further, I participated in a project aimed at the effect of HelD and δ binding partners of RNAP on transcription with different *B. subtilis* σ factors. I showed that the HelD protein does not affect initiation of transcription in σ^A-dependent assays, but that it can be a regulator of transcription initiation driven by alternative σ factors σ^B and σ^D. Thus, HelD is a putative regulator not only of transcription elongation but also initiation. The possible effect of HelD on transcription initiation needs to be further studied. The outcome of this work is my co-author publication in the Nucleic Acids Research (2014), see Appendix 3.

The experiments with the δ protein showed that δ stimulates σ^B-dependent transcription in the same manner as transcription with the primary σ factor and that unstructured C-terminal part of δ plays a substantial role in this process. In addition, I revealed that δC is essential for modulation of the sensitivity of certain promoters to [iNTP], though both – δC and δN – are involved in this process. Experiments addressing HelD/δ/σ interactions are still under way. Our future plans are to test all transcription systems set up in the laboratory with HelD and δ, and reveal the putative role of HelD in transcription initiation.

Finally, I took a part in the project studying small RNA Ms1 from *Mycobacterium smegmatis*. I showed that P_{Ms1} is a strong, almost constitutive promoter whose activity can be compared to the ribosomal promoter. Moreover, I identified the TSS for *M. tuberculosis* MTS2823 that appeared to be identical to Ms1 TSS (Appendix 4).

The obtained results helped advance our understanding of the regulation of bacterial transcription initiation in the free-living species *Bacillus subtilis* and *Mycobacterium smegmatis*. Importantly, in the future, the knowledge that I gained can be extended to explore pathogenic bacteria (*e.g*. to the *B. cereus* group, *M. tuberculosis*), leading the way to applications in the area of health.

STATEMENT OF THE AUTHORS

I confirm that this Thesis written by Volha Ramaniuk is based on published and unpublished data generated during her Ph.D. study in between 2012-2018.

Volha Ramaniuk (in the publications – Olga Ramaniuk) had published two first-author and two co-author publications in the impacted journals (Appendices 1 – 4).

Chapters 4 ("HelD and δ project") and 5 ("Ms1 project") comprise the results that were generated for projects with the main contribution of other lab members. The contribution of other authors to the Results section represented in this Thesis is the following:

Figure 17 was created in collaboration with **O. Benada**

Figure 18 is based on RNA-sequencing data processed by **M. Převorovský**

Figures 19-20, **22**, **37-40** were created in collaboration with **J. Hnilicová**

Figure 39 was created by **J. Hnilicová** and **P. Páleníková.** P. Páleníková created the constructs in **Figure 39a**. The results summarized in **Figure 39** were generated together with **P. Páleníková**

The phylogenetic tree in **Figure 27** was created by **M. Schwarz**

Figure 30 was created by **J. Vohradský**

Tables 6-8 were created by **J. Vohradský**

Supervisor: Mgr. L. Krásný, Ph.D.

AFTERWORD

Six years of Ph.D. study at the Institute of Microbiology, Czech Academy of Sciences was the finest time in my life. It was not an easy pass, but through it, I had lived full and exciting life that positively affected not only my professional skills but also the development of my personality. Thus and so, I am finishing my Ph.D. fully satisfied. Despite all demotivating moments, quandary and obstacles I went through, I gained reach experience in science, and I believe that I became a person to be called “a young scientist”.

I was proud to put all the collected positive results into this Thesis. This is my brick in the wall of basic science, and I am strongly convinced that I put it in the right place ☺

REFERENCES

Achberger, E.C., and Whiteley, H.R. (1981). The role of the delta peptide of the *Bacillus subtilis* RNA polymerase in promoter selection. J. Biol. Chem. *256*, 7424–7432.

Agarwala, R., Barrett, T., Beck, J., Benson, D.A., Bollin, C., Bolton, E., Bourexis, D., Brister, J.R., Bryant, S.H., Canese, K., et al. (2016). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. *44*, D7–D19.

Aiyar, S.E., Gourse, R.L., and Ross, W. (1998). Upstream A-tracts increase bacterial promoter activity through interactions with the RNA polymerase alpha subunit. PNAS *95*, 14652–14657.

Alm, E.J., Huang, K.H., Price, M.N., Koche, R.P., Keller, K., Dubchak, I.L., and Arkin, A.P. (2005). The MicrobesOnline Web site for comparative genomics. Genome Res. *15*, 1015–1022.

Altaf, M., Miller, C.H., Bellows, D.S., and O'Toole, R. (2010). Evaluation of the *Mycobacterium smegmatis* and BCG models for the discovery of *Mycobacterium tuberculosis* inhibitors. Tuberculosis *90*, 333–337.

Anagnostopoulos, C., and Spizizen, J. (1961). Requirements for Transformation in *Bacillus Subtilis*. J. Bacteriol. *81*, 741–746.

Antelmann, H., Scharf, C., and Hecker, M. (2000). Phosphate starvation-inducible proteins of *Bacillus subtilis*: Proteomics and transcriptional analysis. J. Bacteriol. *182*, 4478–4490.

Aravind, L., Leipe, D.D., and Koonin, E. V. (1998). Toprim - A conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. Nucleic Acids Res. *26*, 4205–4213.

Arnvig, K.B., Gopal, B., Papavinasasundaram, K.G., Cox, R.A., and Colston, M.J. (2005). The mechanism of upstream activation in the *rrnB* operon of *Mycobacterium smegmatis* is different from the *Escherichia coli* paradigm. Microbiology *151*, 467–473.

Arnvig, K.B., Comas, I., Thomson, N.R., Houghton, J., Boshoff, H.I., Croucher, N.J., Rose, G., Perkins, T.T., Parkhill, J., Dougan, G., et al. (2011). Sequence-Based Analysis Uncovers an Abundance of Non-Coding RNA in the Total Transcriptome of *Mycobacterium tuberculosis*. PLoS Pathog. *7*, 494–500.

Arrieta-Ortiz, M.L., Hafemeister, C., Bate, A.R., Chu, T., Greenfield, A., Shuster, B., Barry, S.N., Gallitto, M., Liu, B., Kacmarczyk, T., et al. (2015). An experimentally supported model of the

Bacillus subtilis global transcriptional regulatory network. Mol. Syst. Biol. *11*, 839.

Artsimovitch, I., Svetlov, V., Murakami, K.S., and Landick, R. (2003). Co-overexpression of *Escherichia coli* RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. J. Biol. Chem. *278*, 12344–12355.

Asai, K., Ootsuji, T., Obata, K., Matsumoto, T., Fujita, Y., and Sadaie, Y. (2007). Regulatory role of RsgI in σ^I expression in *Bacillus subtilis*. Microbiology *153*, 92–101.

Axmann, I.M., Holtzendorff, J., Voß, B., Kensche, P., and Hess, W.R. (2007). Two distinct types of 6S RNA in *Prochlorococcus*. Gene *406*, 69–78.

Ayora, S., Rojo, F., Ogasawara, N., Nakai, S., and Alonso, J.C. (1996). The *Bacillus subtilis* chromatin-associated protein Hbsu is involved in DNA repair and recombination. J Mol Biol *256*, 301–318.

Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., and Darst, S. a (2015). Structure of a bacterial RNA polymerase holoenzyme open promoter complex. Elife *4*, 1–23.

Baichoo, N., Wang, T., Ye, R., and Helmann, J.D. (2002). Global analysis of the *Bacillus subtilis* Fur regulon and the iron starvation stimulon. Mol. Microbiol. *45*, 1613–1629.

Bansal, M., Gatta, G. Della, and di Bernardo, D. (2006). Inference of gene regulatory networks and compound mode of action from time course gene expression profiles. Bioinformatics *22*, 815–822.

Bar-Joseph, Z., Gitter, A., and Simon, I. (2012). Studying and modelling dynamic biological processes using time-series gene expression data. Nat. Rev. Genet. *13*, 552–564.

Bar-Nahum, G., and Nudler, E. (2001). Isolation and characterization of σ^{70}-retaining transcription elongation complexes from *Escherichia coli*. Cell *106*, 443–451.

Barker, M.M., Gaal, T., and Gourse, R.L. (2001). Mechanism of Regulation of Transcription Initiation by ppGpp . II . Models for Positive Control Based on Properties of RNAP Mutants and Competition for RNAP. J Mol Biol *305*, 689–702.

Barne, K.A., Bown, J.A., and Minchin, S.D. (1997). Region 2.5 of the *Escherichia coli* RNA polymerase σ^{70} subunit is responsible for the recognition of the “extended -10” motif at promoters. EMBO J. *16*, 4034–4040.

Barrick, J.E., Sudarsan, N., Weinberg, Z., Ruzzo, W.L., and Breaker, R.R. (2005). 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. Rna *11*,

774–784.

Basu, R.S., Warner, B.A., Molodtsov, V., Pupov, D., Esyunina, D., Fernández-Tornero, C., Kulbachinskiy, A., and Murakami, K.S. (2014). Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme. J. Biol. Chem. *289*, 24549–24559.

Becker, S.A., and Palsson, B.O. (2008). Context-Specific Metabolic Networks Are Consistent with Experiments. PLoS Comput. Biol. *4*, 1–10.

Belogurov, G.A., and Artsimovitch, I. (2015). Regulation of Transcript Elongation. Annu Rev Microbiol *91*, 165–171.

Bisicchia, P., Noone, D., Lioliou, E., Howell, A., Quigley, S., Jensen, T., Jarmer, H., and Devine, K.M. (2007). The essential YycFG two-component system controls cell wall metabolism in *Bacillus subtilis*. Mol. Microbiol. *65*, 180–200.

Bowers, C.W., and Dombroski, a J. (1999). A mutation in region 1.1 of Sigma 70 affects promoter DNA binding by *Escherichia coli* RNA polymerase holoenzyme. EMBO J. *18*, 709–716.

Britton, R.A., Eichenberger, P., Eduardo, J., Fawcett, P., Monson, R., Grossman, A.D., Gonzalez-pastor, J.E., and Losick, R. (2002). Genome-Wide Analysis of the Stationary-Phase Sigma Factor Sigma-H Regulon of *Bacillus subtilis* Genome-Wide Analysis of the Stationary-Phase Sigma Factor Sigma-H Regulon of *Bacillus subtilis*.

Browning, D.F., and Busby, S.J.W. (2004). The regulation of bacterial transcription initiation. Nat. Rev. Microbiol. *2*, 57–65.

Browning, D.F., and Busby, S.J.W. (2016). Local and global regulation of transcription initiation in bacteria. Nat. Rev. Microbiol. *14*, 638–650.

Brownlee, G.G. (1971). Sequence of 6S RNA of *E. coli*. Nat New Biol. *229*, 147–149.

Burgess, R.R. (1969). Separation and characterization of the subunits of ribonucleic acid polymerase. J. Biol. Chem. *244*, 6168–6176.

Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K. (1969). Factor stimulating transcription by RNA polymerase. Nature *221*, 43–46.

Camarero, J. a, Shekhtman, A., Campbell, E. a, Chlenov, M., Gruber, T.M., Bryant, D. a, Darst, S. a, Cowburn, D., and Muir, T.W. (2002). Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. Proc. Natl. Acad. Sci. U. S. A. *99*, 8536–8541.

Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz,

M.L., and Darst, S.A. (2002). Structure of the bacterial RNA polymerase promoter specificity σ subunit. Mol. Cell *9*, 527–539.

Cao, M., and Helmann, J.D. (2002). Regulation of the *Bacillus subtilis bcrC* bacitracin resistance gene by two extracytoplasmic function σ factors. J Bacteriol *184*, 6123–6129.

Cao, M., Moore, C.M., and Helmann, J.D. (2005). *Bacillus subtilis* paraquat resistance is directed by σ^M, an extracytoplasmic function σ factor, and is conferred by YqjL and BcrC. J. Bacteriol. *187*, 2948–2956.

Caramori, T., Barillà, D., Nessi, C., Sacchi, L., and Galizzi, A. (1996). Role of FlgM in σ^D-dependent gene expression in *Bacillus subtilis*. J. Bacteriol. *178*, 3113–3118.

Carballido-López, R., Formstone, A., Li, Y., Ehrlich, S.D., Noirot, P., and Errington, J. (2006). Actin Homolog MreBH Governs Cell Morphogenesis by Localization of the Cell Wall Hydrolase LytE. Dev. Cell *11*, 399–409.

Carlsson, P., and Hederstedt, L. (1989). Genetic characterization of *Bacillus subtilis odhA* and *odhB*, encoding 2-oxoglutarate dehydrogenase and dihydrolipoamide transsuccinylase, respectively. J Bacteriol *171*, 3667–3672.

Cavanagh, A.T., and Wassarman, K.M. (2014). 6S RNA, a Global Regulator of Transcription in *Escherichia coli*, *Bacillus subtilis*, and Beyond. Annu. Rev. Microbiol. *68*, 45–60.

Cavanagh, A.T., Klocko, A.D., Liu, X., and Wassarman, K.M. (2008). Promoter specificity for 6S RNA regulation of transcription is determined by core promoter sequences and competition for region 4.2 of Sigma 70. Mol. Microbiol. *67*, 1242–1256.

Chang, B., and Doi, R. (1990). Overproduction, Purification, and Characterization of *Bacillus subtilis* RNA Polymerase σ^A Factor. J Bacteriol *172*, 3257–3263.

Chastanet, A., Vitkup, D., Yuan, G.-C., Norman, T.M., Liu, J.S., and Losick, R.M. (2010). Broadly heterogeneous activation of the master regulator for sporulation in *Bacillus subtilis*. Proc. Natl. Acad. Sci. *107*, 8486–8491.

Chatterji, D., and Kumar Ojha, A. (2001). Revisiting the stringent response, ppGpp and starvation signaling. Curr. Opin. Microbiol. *4*, 160–165.

Chen, Y., and Helmann, J.D. (1995). The *Bacillus subtilis* Flagellar Regulatory Protein σ^D: Overproduction, Domain Analysis and DNA-binding Properties. J Mol Biol 743–753.

China, A., Tare, P., and Nagaraja, V. (2010). Comparison of promoter-specific events during

transcription initiation in mycobacteria. Microbiology *156*, 1942–1952.

Cho, B.-K., Palsson, B., and Zengler, K. (2011). Deciphering the regulatory codes in bacterial genomes. Biotechnol. J. *6*, 1052–1063.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural Basis of Transcription: RNA Polymerase II at 2.8 Angstrom Resolution. Science (80). *292*, 1863–1876.

Davis, M.C., Kesthely, C.A., Franklin, E.A., Affiliations, S.R.M., and Maclellan, S.R. (2016). The Essential Activities of the Bacterial σ Factor. Can. J. Microbiol. *99*, 1–36.

Davis, M.C., Kesthely, C.A., Franklin, E.A., and MacLellan, S.R. (2017). The essential activities of the bacterial sigma factor. Can. J. Microbiol. *63*, 89–99.

Débarbouillé, M., Martin-Verstraete, I., Kunst, F., and Rapoport, G. (1991). The *Bacillus subtilis sigL* gene encodes an equivalent of sigma 54 from gram-negative bacteria. Proc. Natl. Acad. Sci. U. S. A. *88*, 9092–9096.

Defeu Soufo, H.J., and Graumann, P.L. (2003). Actin-like Proteins MreB and Mbl from *Bacillus subtilis* Are Required for Bipolar Positioning of Replication Origins. Curr. Biol. *13*, 1916–1920.

Delumeau, O., Lewis, R.J., and Yudkin, M.D. (2002). Protein-protein interactions that regulate the energy stress activation of Sigma B in *Bacillus subtilis*. J. Bacteriol. *184*, 5583–5589.

Delumeau, O., Lecointe, F., Muntel, J., Guillot, A., Guédon, E., Monnet, V., Hecker, M., Becher, D., Polard, P., and Noirot, P. (2011). The dynamic protein partnership of RNA polymerase in *Bacillus subtilis*. Proteomics *11*, 2992–3001.

DiChiara, J.M., Contreras-Martinez, L.M., Livny, J., Smith, D., McDonough, K.A., and Belfort, M. (2010). Multiple small RNAs identified in *Mycobacterium bovis* BCG are also expressed in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. Nucleic Acids Res. *38*, 4067–4078.

Dijl, J.M. Van, and Hecker, M. (2013). *Bacillus subtilis* : from soil bacterium to super-secreting cell factory. Microb. Cell Fact. 1–6.

Dobinson, K.F., and Spiegelman, G.B. (1987). Effect of the δ Subunit of *Bacillus subtilis* RNA Polymerase on Initiation of RNA Synthesis at Two Bacteriophage ϕ29 Promoters. Biochemistry *26*, 8206–8213.

Dombroski, A.J., Walter, W.A., and Gross, C.A. (1993). Amino-terminal amino acids modulate σ factor DNA-binding activity. Genes Dev. *7*, 2446–2455.

Domínguez-Cuevas, P., Porcelli, I., Daniel, R.A., and Errington, J. (2013). Differentiated roles for

MreB-actin isologues and autolytic enzymes in *Bacillus subtilis* morphogenesis. Mol. Microbiol. *89*, 1084–1098.

Doolittle, R.F., Feng, D., Tsang, S., Cho, G., and Little, E. (1996). Determining Divergence Times of the Major Kingdoms of Living Organisms with a Protein Clock Published by: American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/2890144. Science (80). *271*, 470–477.

Dubnau, D., and Davidoff-Abelson, R. (1971). Fate of transforming DNA following uptake by competent *Bacillus subtilis*: I. Formation and properties of the donor-recipient complex. J Mol Biol *56*, 209–221.

Dubnau, D., and Losick, R. (2006). Bistability in bacteria. Mol. Microbiol. *61*, 564–572.

Dubrac, S., Bisicchia, P., Devine, K.M., and Msadek, T. (2008). A matter of life and death: Cell wall homeostasis and the WalKR (YycGF) essential signal transduction pathway. Mol. Microbiol. *70*, 1307–1322.

Earl, A.M., Losick, R., Kolter, R., Wang, J., Dye, B.T., Rajashankar, K.R., Kurinov, I., and Brenda, A. (2008). Ecology and genomics of *Bactillus subtilis*. Trends Microbiol. *16*, 1–11.

Edgar, R.C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

Estrem, S.T., Ross, W., Gaal, T., Chen, Z.W.S., Niu, W., Ebright, R.H., and Gourse, R.L. (1999). Bacterial promoter architecture: Subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase α subunit. Genes Dev. *13*, 2134–2147.

Fabret, C., Feher, V. a, and Hoch, J. a (1999). Two-Component Signal Transduction in Bacillus subtilis : How One Organism Sees Its World MINIREVIEW Two-Component Signal Transduction in *Bacillus subtilis* : How One Organism Sees Its World. J. Bacteriol. *181*, 1975–1983.

Farewell, A., Kvint, K., and Nyström, T. (1998). Negative regulation by RpoS: A case of sigma factor competition. Mol. Microbiol. *29*, 1039–1051.

Faucher, S.P., Friedlander, G., Livny, J., Margalit, H., and Shuman, H.A. (2010). Legionella pneumophila 6S RNA optimizes intracellular multiplication. Proc. Natl. Acad. Sci. *107*, 7533–7538.

Feist, A.M., Herrgård, M.J., Thiele, I., Reed, J.L., and Palsson, B. (2009). Reconstruction of

biochemical networks in microorganisms. Nat. Rev. Microbiol. *7*, 129–143.

Feklistov, A. (2013). RNA polymerase: in search of promoters. Ann. N. Y. Acad. Sci. *1293*, 25–32.

Feklistov, A., and Darst, S.A. (2011). Structural basis for promoter -10 element recognition by the bacterial RNA polymerase σ subunit. Cell *147*, 1257–1269.

Feklistov, A., Sharon, B.D., Darst, S.A., and Gross, C.A. (2014). Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective. Annu. Rev. Microbiol. *68*, 357–376.

Feng, D.-F., Cho, G., and Doolittle, R.F. (1997). Determining divergence times with a protein clock: Update and reevaluation. Proc. Natl. Acad. Sci. *94*, 13028–13033.

Feng, Y., Zhang, Y., and Ebright, R.H. (2016). Structural basis of transcription activation. Science (80). *91*, 165–171.

Flentie, K., Garner, A.L., and Stallings, C.L. (2016). *Mycobacterium tuberculosis* Transcription Machinery: Ready To Respond to Host Attacks. J B *198*, 1360–1373.

Francez-Charlot, A., Frunzke, J., Reichen, C., Ebneter, J.Z., Gourion, B., and Vorholt, J.A. (2009). Sigma factor mimicry involved in regulation of general stress response. Proc. Natl. Acad. Sci. *106*, 3467–3472.

Fredrick, K.L., and Helmann, J.D. (1994). Dual chemotaxis signaling pathways in *Bacillus subtilis*: A σ(D)-dependent gene encodes a novel protein with both CheW and CheY homologous domains. J. Bacteriol. *176*, 2727–2735.

Fujiwara, N., Naka, T., Ogawa, M., Yamamoto, R., Ogura, H., and Taniguchi, H. (2012). Characteristics of *Mycobacterium smegmatis* J15cs strain lipids. Tuberculosis *92*, 187–192.

Fukushima, T., Ishikawa, S., Yamamoto, H., Ogasawara, N., and Sekiguchi, J. (2003). Transcriptional, functional and cytochemical analyses of the *veg* gene in *Bacillus subtilis*. J. Biochem. *133*, 475–483.

Gaal, T., Bartlett, M.S., Ross, W., Jr, C.L.T., and Gourse, R.L. (1997a). Transcription regulation by initiating NTP concentration - rRNA synthesis in bacteria. Science (80). *278*, 1–6.

Gaal, T., Bartlett, M.S., Ross, W., Jr, C.L.T., and Gourse, R.L. (1997b). Transcription Regulation by Initiating NTP Concentration rRNA Synthesis in Bacteria. Science (80). *278*.

Gaballa, A., Guariglia-Oropeza, V., Dürr, F., Butcher, B.G., Chen, A.Y., Chandrangsu, P., and Helmann, J.D. (2017). Modulation of extracytoplasmic function (ECF) sigma factor promoter

selectivity by spacer region sequence. Nucleic Acids Res. *46*, 134–145.

Gardella, T., Moyle, H., and Susskind, M.M. (1989). A mutant Escherichia coli σ^{70} subunit of RNA polymerase with altered promoter specificity. J. Mol. Biol. *206*, 579–590.

Georg, J., and Hess, W.R. (2018). Widespread Antisense Transcription in Prokaryotes. Microbiol. Spectr. *6*, 1–20.

Geszvain, K., Gruber, T.M., Mooney, R.A., Gross, C.A., and Landick, R. (2004). A hydrophobic patch on the flap-tip helix of *E. coli* RNA polymerase mediates σ^{70} region 4 function. J. Mol. Biol. *343*, 569–587.

Goldman, S.R., Ebright, R.H., and Nickels, B.E. (2009). Direct Detection of Abortive RNA Transcripts *In Vivo*. Science (80). *324*, 1–4.

Gourse, R.L., Ross, W., and Gaal, T. (2000). UPs and downs in bacterial transcription initiation: The role of the alpha subunit of RNA polymerase in promoter recognition. Mol. Microbiol. *37*, 687–695.

Gruber, T.M., and Gross, C.A. (2003). Multiple σ Subunits and the Partitioning of Bacterial Transcription Space. Annu. Rev. Microbiol. *57*, 441–466.

Gruber, T.M., Markov, D., Sharp, M.M., Young, B.A., Lu, C.Z., Zhong, H.J., Artsimovitch, I., Geszvain, K.M., Arthur, T.M., Burgess, R.R., et al. (2001). Binding of the initiation factor σ^{70} to core RNA polymerase is a multistep process. Mol. Cell *8*, 21–31.

Gupta, R., Barkan, D., Redelman-Sidi, G., Shuman, S., and Glickman, M.S. (2011). *Mycobacteria* exploit three genetically distinct DNA double-strand break repair pathways. Mol. Microbiol. *79*, 316–330.

Gwynn, E.J., Smith, A.J., Guy, C.P., Savery, N.J., McGlynn, P., and Dillingham, M.S. (2013). The Conserved C-Terminus of the PcrA/UvrD Helicase Interacts Directly with RNA Polymerase. PLoS One *8*, 1–11.

Haldenwang, W.G. (1995). The sigma factors of *Bacillus subtilis*. Microbiol. Rev. *59*, 1–30.

Haldenwang, W.G., and Losick, R. (1979). A modified RNA polymerase transcribes a cloned gene under sporulation control in *Bacillus subtilis*. Nat Rev Microbiol *282*, 256–260.

Haldenwang, W.G., and Losick, R. (1980). Novel RNA polymerase sigma factor from *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S. A. *77*, 7000–7004.

Haldenwang, W.G., Lang, N., and Losick, R. (1981). A sporulation-induced sigma-like regulatory

protein from *B. subtilis*. Cell *23*, 615–624.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol *166*, 557–580.

Hansen, U.M., and McClure, W.R. (1980). Role of the u Subunit of *Escherichia coli* RNA Polymerase in Initiation. J. BlOLOClCAL Chem. *255*, 95–9570.

Hashimoto, M., Ooiwa, S., and Sekiguchi, J. (2012). Synthetic lethality of the *lytE cwlO* genotype in *Bacillus subtilis* is caused by lack of D, L-endopeptidase activity at the lateral cell wall. J. Bacteriol. *194*, 796–803.

Haugen, S.P., Berkmen, M.B., Ross, W., Gaal, T., Ward, C., and Gourse, R.L. (2006). rRNA Promoter Regulation by Nonoptimal Binding of σ Region 1.2: An Additional Recognition Element for RNA Polymerase. Cell *125*, 1069–1082.

Haugen, S.P., Ross, W., and Gourse, R.L. (2008). Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Microbiol *49*, 1841–1850.

Hawley, D.K., and McClure, W.R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. *11*, 2237–2255.

Heath, R.J., Su, N., Murphy, C.K., and Rock, C.O. (2000). The enoyl-[acyl-carrier-protein] reductases FabI and FabL from *Bacillus subtilis*. J. Biol. Chem. *275*, 40128–40133.

Hecker, M., Pane-Farre, J., and Volker, U. (2007). SigB-Dependent General Stress Response in *Bacillus subtilis* and Related Gram-Positive Bacteria. Annu Rev Microbiol *61*, 215–236.

Heinrich, J., and Wiegert, T. (2009). Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. Res. Microbiol. *160*, 696–703.

Helmann, J.D. (1995). Compilation and analysis of *Bacillus subtilis* δA-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. Nucleic Acids Res. *23*, 2351–2360.

Helmann, J.D. (2016a). *Bacillus subtilis* extracytoplasmic function (ECF) sigma factors and defense of the cell envelope John. Curr. Opin. Microbiol. *97*, 5421–5433.

Helmann, J.D. (2016b). *Bacillus subtilis* extracytoplasmic function (ECF) σ factors and defense of the cell envelope. Curr. Opin. Microbiol. *30*, 122–132.

Hicks, K.A., and Grossman, A.D. (1996). Altering the level and regulation of the major sigma subunit of RNA polymerase affects gene expression and development in *Bacillus subtilis*. Mol

Microbiol. *20*, 201–212.

Hilbert, D.W., and Piggot, P.J. (2004). Compartmentalization of gene expression during *Bacillus subtilis* spore formation. Microbiol. Mol. Biol. Rev. *68*, 234–262.

Hindley, J. (1967). Fractionation of 32P-labelled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting. J Mol Biol *30*, 125–136.

Hirata, A., Klein, B.J., and Murakami, K.S. (2008). The X-ray Crystal Structure of RNA Polymerase from Archaea. Nature *451*, 1–12.

Hnilicová, J., Jirát Matějčková, J., Šiková, M., Pospíšil, J., Halada, P., Pánek, J., and Krásný, L. (2014). Ms1, a novel sRNA interacting with the RNA polymerase core in *Mycobacteria*. Nucleic Acids Res. *42*, 11763–11776.

Hong, H.A., Khaneja, R., Tam, N.M.K., Cazzato, A., Tan, S., Urdaci, M., Brisson, A., Gasbarrini, A., Barnes, I., and Cutting, S.M. (2009). *Bacillus subtilis* isolated from the human gastrointestinal tract. Res. Microbiol. *160*, 134–143.

Hook-Barnard, I.G., and Hinton, D.M. (2009). The promoter spacer influences transcription initiation via 70 region 1.1 of *Escherichia coli* RNA polymerase. Proc. Natl. Acad. Sci. *106*, 737–742.

Horsburgh, M.J., and Moir, A. (1999). Sigma M, an ECF RNA polymerase sigma factor of *Bacillus subtilis* 168, is essential for growth and survival in high concentrations of salt. Mol. Microbiol. *32*, 41–50.

Horsburgh, M.J., Thackray, P.D., and Moir, A. (2001). Transcriptional responses during outgrowth of *Bacillus subtilis* endospores. Microbiology *147*, 2933–2941.

Hsu, L.M. (2002). Promoter clearance and escape in prokaryotes. Biochim. Biophys. Acta - Gene Struct. Expr. *1577*, 191–207.

Hsu, L.M. (2009). Monitoring abortive initiation. Methods *47*, 25–36.

Hsu, L.M., Vo, N. V., and Chamberlin, M.J. (1995). *Escherichia coli* transcript cleavage factors GreA and GreB stimulate promoter escape and gene expression in vivo and in vitro. Proc. Natl. Acad. Sci. *92*, 11588–11592.

Huang, W.Z., Wang, J.J., Chen, H.J., Chen, J.T., and Shaw, G.C. (2013). The heat-inducible essential response regulator WalR positively regulates transcription of *sigI*, *mreBH* and *lytE* in *Bacillus subtilis* under heat stress. Res. Microbiol. *164*, 998–1008.

Huang, X., Gaballa, A., Cao, M., and Helmann, J.D. (1999). Identification of target promoters for the *Bacillus subtilis* extracytoplasmic function σ factor, σ(W). Mol. Microbiol. *31*, 361–371.

Huber W., Carey V.J., Gentleman R., Anders S., Carlson M., Carvalho B.S., Corrada Bravo H., Davis S., Gatto L., Girke T., Gottardo R., Hahne F., Hansen K.D., Irizarry R.A., Lawrence M., Love M.I., MacDonald J., Obenchain V., Ole A.K., Pages H., Reyes A., M.M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods *12*, 115–121.

Hunger, K., Beckering, C.L., Wiegeshoff, F., Graumann, P.L., and Marahiel, M.A. (2006). Cold-induced putative DEAD box RNA helicases CshA and CshB are essential for cold adaptation and interact with cold shock protein B in *Bacillus subtilis*. J. Bacteriol. *188*, 240–248.

Hyde, E.I., Hilton, M.D., and Whiteley, H.R. (1986). Interactions of the *Bacillus subtilis* RNA polymerase with subunits determining the specificity of interaction. Sigma and Delta peptides can bind simultaneously to core. J.Biol.Chem. *261*, 16565–16570.

Ignatov, D., Malakho, S., Majorov, K., Skvortsov, T., Apt, A., and Azhikina, T. (2013). RNA-Seq Analysis of *Mycobacterium avium* Non-Coding Transcriptome. PLoS One *8*, 1–11.

Ishihama, A. (2012). Prokaryotic genome regulation : A revolutionary paradigm. Proc. Japan Acad. Ser. B *88*, 485–508.

Ishikawa, S., Oshima, T., Kurokawa, K., Kusuya, Y., and Ogasawara, N. (2010). RNA polymerase trafficking in *Bacillus subtilis* cells. J. Bacteriol. *192*, 5778–5787.

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. *3*, 318–356.

Jaehning, J.A., Wiggs, J.L., and Chamberlin, M.J. (1979). Altered promoter selection by a novel form of *Bacillus subtilis* RNA polymerase. Proc Natl Acad Sci U S A *76*, 5470–5474.

Jarmer, H., Larsen, T.S., Krogh, A., Saxild, H.H., Brunak, S., and Knudsen, S. (2001). σA recognition sites in the *Bacillus subtilis* genome. Microbiology *147*, 2417–2424.

Johnson, S., Chen, Y.J., and Phillips, R. (2013). Poly(dA:dT)-Rich DNAs Are Highly Flexible in the Context of DNA Looping. PLoS One *8*.

Johnson, W.C., Moran, C.P., and Losick, R. (1983). Two RNA polymerase sigma factors from *Bacillus subtilis* discriminate between overlapping promoters for a developmentally regulated gene. Nature *302*, 800–804.

Jones, L.J.F., Carballido-López, R., and Errington, J. (2001). Control of cell shape in bacteria:

helical, actin-like filaments in *Bacillus subtilis*. Cell *104*, 913–922.

Jöres, L., and Wagner, R. (2003). Essential steps in the ppGpp-dependent regulation of bacterial ribosomal RNA promoters can be explained by substrate competition. J. Biol. Chem. *278*, 16834–16843.

Juang, Y.-L., and Helmann, J.D. (1994a). A promoter melting region in the primary σ factor of *Bacillus subtilis*. J. Mol. Biol. *235*, 1470–1488.

Juang, Y.L., and Helmann, J.D. (1994b). The δ subunit of bacillus subtilis RNA polymerase: An allosteric effector of the initiation and core-recycling phases of transcription. J. Mol. Biol. *239*, 1–14.

Kanehisa, M. (1996). Toward pathway engineering: a new database of genetic and molecular pathways. Sci. Technol. Japan *59*, 34–38.

Kang, J.G., Paget, M.S.B., Seok, Y.J., Hahn, M.Y., Bae, J.B., Hahn, J.S., Kleanthous, C., Buttner, M.J., and Roe, J.H. (1999). RsrA, an anti-sigma factor regulated by redox change. EMBO J. *18*, 4292–4298.

Kapanidis, A.N., Margeat, E., Ho, S.O., Kortkhonjia, E., Weiss, S., and Ebright, R.H. (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. Science (80-.). *314*, 1144–1147.

Kawai, Y., Asai, K., and Errington, J. (2009). Partial functional redundancy of MreB isoforms, MreB, Mbl and MreBH, in cell morphogenesis of *Bacillus subtilis*. Mol. Microbiol. *73*, 719–731.

Keijser, B.J.F., Ter Beek, A., Rauwerda, H., Schuren, F., Montijn, R., Van Der Spek, H., and Brul, S. (2007). Analysis of temporal gene expression during *Bacillus subtilis* spore germination and outgrowth. J. Bacteriol. *189*, 3624–3634.

Keilty, S., and Rosenberg, M. (1987). Constitutive function of a positively regulated promoter reveals new sequences essential for activity. J. Biol. Chem. *262*, 6389–6395.

Kim, J.G.Y., and Wilson, A.C. (2016). Loss of σI affects heat-shock response and virulence gene expression in *Bacillus anthracis*. Microbiology *162*, 564–574.

Kim, K.S., and Lee, Y. (2004). Regulation of 6S RNA biogenesis by switching utilization of both sigma factors and endoribonucleases. Nucleic Acids Res. *32*, 6057–6068.

Kobayashi, K., Ehrlich, S.D., Albertini, A., Amati, G., Andersen, K.K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., et al. (2003). Essential *Bacillus subtilis* genes. Proc.

Natl. Acad. Sci. *100*, 4678–4683.

Kontur, W.S., Capp, M.W., Gries, T.J., Saecker, R.M., and Record, M.T. (2010). Probing DNA binding, DNA opening, and assembly of a downstream clamp/jaw in *Escherichia coli* RNA polymerase-Lambda PR promoter complexes using salt and the physiological anion glutamate. Biochemistry *49*, 4361–4373.

Koo, B.M., Rhodius, V.A., Nonaka, G., DeHaseth, P.L., and Gross, C.A. (2009). Reduced capacity of alternative sigmas to melt promoters ensures stringent promoter recognition. Genes Dev. *23*, 2426–2436.

Koo, B.M., Kritikos, G., Farelli, J.D., Todor, H., Tong, K., Kimsey, H., Wapinski, I., Galardini, M., Cabal, A., Peters, J.M., et al. (2017). Construction and Analysis of Two Genome-Scale Deletion Libraries for *Bacillus subtilis*. Cell Syst. *4*, 291–305.e7.

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S.A. (2000). A structural model of transcription elongation. Science (80). *289*, 619–625.

Krásny, L., and Gourse, R.L. (2004). An alternative strategy for bacterial ribosome synthesis: *Bacillus subtilis* rRNA transcription regulation. EMBO J. *23*, 4473–4483.

Krásný, L., Tišerová, H., Jonák, J., Rejman, D., and Šanderová, H. (2008). The identity of the transcription +1 position is crucial for changes in gene expression in response to amino acid starvation in *Bacillus subtilis*. Mol. Microbiol. *69*, 42–54.

Kunst, F. (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature *390*, 249–256.

Kusuya, Y., Kurokawa, K., Ishikawa, S., Ogasawara, N., and Oshima, T. (2011). Transcription factor GreA contributes to resolving promoter-proximal pausing of RNA polymerase in *Bacillus subtilis* cells. J. Bacteriol. *193*, 3090–3099.

Lampe, M., Binnie, C., Schmidt, R., and Losick, R. (1988). Cloned gene encoding the delta subunit of *Bacillus subtilis* RNA polymerase. Gene *67*, 13–19.

Landick, R. (2005). NTP-entry routes in multi-subunit RNA polymerases. Trends Biochem. Sci. *30*, 651–654.

Lane, W.J., and Darst, S.A. (2006). The structural basis for promoter -35 element recognition by the group IV σ factors. PLoS Biol. *4*, 1491–1500.

Lang, K.S., and Merrikh, H. (2018). The Clash of Macromolecular Titans: Replication-

Transcription Conflicts in Bacteria. Annu. Rev. Microbiol. *72*, 71–88.

Lew, C.M., and Gralla, J.D. (2004). Nucleotide-dependent isomerization of *Escherichia coli* RNA polymerase. Biochemistry *43*, 12660–12666.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Liu, J., and Turnbough, C.L. (1994). Effects of transcriptional start site sequence and position on nucleotide- sensitive selection of alternative start sites at the pyrC promoter in *Escherichia coli*. J. Bacteriol. *176*, 2938–2945.

Liu, C., Heath, L.S., and Turnbough Jr., C.L. (1994). Regulation of pyrBI operon expression in *Escherichia coli* by UTP-sensitive reiterative RNA synthesis during transcription initiation. Genes Dev *8*, 2904–2912.

Liu, K., Bittner, A.N., and Wang, J.D. (2014). Diversity in (p)ppGpp metabolism and effectors Kuanqing. *15*, 34–48.

Liu, T.Y., Chu, S.H., Hu, Y.N., Wang, J.J., and Shaw, G.C. (2017). Genetic evidence that multiple proteases are involved in modulation of heat-induced activation of the sigma factor σ^I in *Bacillus subtilis*. FEMS Microbiol. Lett. *364*, 1–7.

Lonetto, M.A., Brown, K.L., Rudd, K.E., and Buttner, M.J. (1994). Analysis of the *Streptomyces coelicolor sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. Proc. Natl. Acad. Sci. *91*, 7573–7577.

Lopez de Saro, F.J., Woody, A.Y., and Helmann, J.D. (1995). Structural Analysis of the *Bacillus subtilis* Delta Factor: a Protein Polyanion Which Displaces RNA from RNA Polymerase. J Mol Biol *252*, 189–202.

López De Saro, F.J., Yoshikawa, N., and Helmann, J.D. (1999). Expression, abundance, and RNA polymerase binding properties of the δ factor of *Bacillus subtilis*. J. Biol. Chem. *274*, 15953–15958.

Lord, M., Barilla, D., and Yudkin, M.D. (1999). Replacement of vegetative Sigma A by

sporulation-specific Sigma F as a component of the RNA polymerase holoenzyme in sporulating *Bacillus subtilis*. J. Bacteriol. *181*, 2346–2350.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15/550*, 1–21.

Lybecker, M., Bilusic, I., and Raghavan, R. (2014). Pervasive transcription: Detecting functional RNAs in bacteria. Transcription *5*, 1–5.

MacLellan, S.R., Wecke, T., and Helmann, J.D. (2008). A previously unidentified σ factor and two accessory proteins regulate oxalate decarboxylase expression in *Bacillus subtilis*. Mol. Microbiol. *69*, 954–967.

Macquarrie, K.L., Fong, A.P., Morse, R.H., and Tapscott, S.J. (2011). Genome-wide transcription factor binding: beyond direct target regulation. Trends Genet. *27*, 141–148.

Malhotra, A., Severinova, E., and Darst, S.A. (1996). Crystal structure of a Sigma 70 subunit fragment from *E. coli* RNA polymerase. Cell *87*, 127–136.

Mangel, Walter F. and Chamberlin, M.J. (1974). Studies of Ribonucleic Acid Ribonucleic Acid Polymerase T7 Deoxyribonucleic Acid by *Escherichia coli*. J. Biol. Chem. *249*, 2995–3001.

Margot, P., Wahlen, M., Gholamhuseinian, A., Piggot, P., and Karamata, D. (1998). The *lytE* Gene of *Bacillus subtilis* 168 Encodes a Cell Wall Hydrolase. J Bacteriol *180*, 749–752.

Markowitz, V.M., Chen, I.A., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., Ratner, A., Jacob, B., Huang, J., Williams, P., et al. (2012). IMG: the Integrated Microbial Genomes database and comparative analysis system. Nucleic Acids Res. *40*, D115-22.

Mars, R.A.T., Mendonça, K., Denham, E.L., and van Dijl, J.M. (2015). The reduction in small ribosomal subunit abundance in ethanol-stressed cells of *Bacillus subtilis* is mediated by a SigB-dependent antisense RNA. Biochim. Biophys. Acta - Mol. Cell Res. *1853*, 2553–2559.

Martínez-Antonio, A., and Collado-Vides, J. (2003). Identifying global regulators in transcriptional regulatory networks in bacteria. Curr. Opin. Microbiol. *6*, 482–489.

Matsumoto, T., Nakanishi, K., Asai, K., and Sadaie, Y. (2005). Transcriptional analysis of the ylaABCD operon of *Bacillus subtilis* encoding a sigma factor of extracytoplasmic function family. Genes Genet. Syst. *80*, 385–393.

May, J.J., Wendrich, T.M., and Marahiel, M.A. (2001). The *dhb* Operon of *Bacillus subtilis* Encodes the Biosynthetic Template for the Catecholic Siderophore 2,3-Dihydroxybenzoate-

Glycine-Threonine Trimeric Ester Bacillibactin. J. Biol. Chem. *276*, 7209–7217.

McClure, W.R. (1985). Mechanism and control of transcription initiation in prokaryotes. Annu Rev Microbiol *54*, 171–204.

McDonnell, G.E., Wood, H., Devine, K.M., and Mcconnell, D.J. (1994). Genetic control of bacterial suicide : regulation of the induction of PBSX in *Bacillus subtilis*. J Bacteriol *176*, 5820–5830.

Mckenney, P.T., Driks, A., and Eichenberger, P. (2013). The *Bacillus subtilis* endospore: Assembly and functions of the multilayered coat. Nat. Rev. Microbiol. *11*, 33–44.

Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R., and Ebright, R.H. (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. Cell *108*, 599–614.

Mendez, R., Gutierrez, A., Reyes, J., and Márquez-Magaña, L. (2012). The Extracytoplasmic Function Sigma Factor SigY Is Important for Efficient Maintenance of the Spβ Prophage That Encodes Sublancin in *Bacillus subtilis*. DNA Cell Biol. *31*, 946–955.

Michna, R.H., Zhu, B., Mäder, U., and Stülke, J. (2016). *Subti* Wiki 2.0—an integrated database for the model organism *Bacillus subtilis*. Nucleic Acids Res. *44*, D654–D662.

Minakhin, L., Bhagat, S., Brunning, A., Campbell, E.A., Darst, S.A., Ebright, R.H., and Severinov, K. (2001). Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc. Natl. Acad. Sci. U. S. A. *98*, 892–897.

Mirel, D.B., and Chamberlin, M.J. (1989). The *Bacillus subtilis* flagellin gene (hag) is transcribed by the σ28form of RNA polymerase. J. Bacteriol. *171*, 3095–3101.

Missiakas, D., and Raina, S. (1998). The extracytoplasmic function sigma factors: Role and regulation. Mol. Microbiol. *28*, 1059–1066.

Mitchell, J.E., Zheng, D., Busby, S.J.W., and Minchin, S.D. (2003). Identification and analysis of "extended -10" promoters in *Escherichia coli*. Nucleic Acids Res. *31*, 4689–4695.

Moreno-Campuzano, S., Janga, S.C., and Pérez-Rueda, E. (2006). Identification and analysis of DNA-binding transcription factors in *Bacillus subtilis* and other Firmicutes - A genomic approach. BMC Genomics *7*, 1–10.

Motácková, V., Sanderová, H., Zídek, L., Novácek, J., Padrta, P., Svenková, A., Korelusová, J.,

Jonák, J., Krásný, L., and Sklenár, V. (2010). Solution structure of the N-terminal domain of *Bacillus subtilis* delta subunit of RNA polymerase and its classification based on structural homologs. Proteins *78*, 1807–1810.

Mueller, J.P., Bukusoglu, G., and Sonenshein, A.L. (1992). Transcriptional regulation of *Bacillus subtilis* glucose starvation-inducible genes: control of gsiA by the ComP-ComA signal transduction system. J. Bacteriol. *174*, 4361–4373.

Mukherjee, S., Stamatis, D., Bertsch, J., Ovchinnikova, G., Katta, H.Y., Mojica, A., Chen, I.-M.A., Kyrpides, N.C., and Reddy, T. (2018). Genomes OnLine database (GOLD) v.7: updates and new features. Nucleic Acids Res. 1–11.

Mukhopadhyay, J., Kapanidis, A.N., Mekler, V., Kortkhonjia, E., Ebright, Y.W., and Ebright, R.H. (2001). Translocation of σ70 with RNA Polymerase during Transcription. Cell *106*, 453–463.

Muñoz-Gutiérrez, I., Ortiz de Ora, L., Rozman Grinberg, I., Garty, Y., Bayer, E.A., Shoham, Y., Lamed, R., and Borovok, I. (2015). Decoding Biomass-Sensing Regulons of *Clostridium thermocellum* Alternative σ^I Factors in a Heterologous *Bacillus subtilis* Host System. PLoS One *11*.

Murakami, K.S., and Darst, S.A. (2003). Bacterial RNA polymerases: The wholo story. Curr. Opin. Struct. Biol. *13*, 31–39.

Murakami, K.S., Masuda, S., and Darst, S.A. (2002). Structural basis of transcription initiation: RNA polymerase holoenzyme at 4A resolution. Science (80-.). *296*, 1280–1284.

Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2014). Structural Basis of Transcription Initiation : An RNA Polymerase HoLoenzyme-DNA Complex. Science (80-.). *296*, 1285–1290.

Murray, H.D., Schneider, D.A., and Gourse, R.L. (2003). Control of rRNA expression by small molecules is dynamic and nonredundant. Mol. Cell *12*, 125–134.

Nakano, S., Erwin, K.N., Ralle, M., and Zuber, P. (2005). Redox-sensitive transcriptional control by a thiol/disulphide switch in the global regulator, Spx. Mol. Microbiol. *55*, 498–510.

Narula, J., Tiwari, A., and Igoshin, O.A. (2016). Role of Autoregulation and Relative Synthesis of Operon Partners in Alternative Sigma Factor Networks. PLoS Comput. Biol. *12*, 1–25.

Neusser, T., Polen, T., Geissen, R., and Wagner, R. (2010). Depletion of the non-coding regulatory 6S RNA in *E. coli* causes a surprising reduction in the expression of the translation machinery.

BMC Genomics *11*.

Nicholson, W.L., and Park, R. (2015). Anaerobic growth of *Bacillus subtilis* alters the spectrum of spontaneous mutations in the rpoB gene leading to rifampicin resistance. FEMS Microbiol. Lett.

Nicolas, P., Mäder, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., Bidnenko, E., Marchadier, E., Hoebeke, M., Aymerich, S., et al. (2012). Condition-Dependent Transcriptome Reveals High-Level Regulatory Architecture in *Bacillus subtilis*. Science (80). *335*, 1103–1106.

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., et al. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. *44*, D733–D745.

Okonechnikov, K., Conesa, A., and García-Alcalde, F. (2015). Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics *32*, 292–294.

Okonogi, T.M., Alley, S.C., Reese, A.W., Hopkins, P.B., and Robinson, B.H. (2002). Sequence-dependent dynamics of duplex DNA: The applicability of a dinucleotide model. Biophys. J. *83*, 3446–3459.

Paget, M.S. (2015). Bacterial σ factors and anti-σ factors: Structure, function and distribution. Biomolecules *5*, 1245–1265.

Pané-Farré, J., Lewis, R.J., and Stülke, J. (2005). The RsbRST stress module in bacteria: A signalling system that may interact with different output modules. J. Mol. Microbiol. Biotechnol. *9*, 65–76.

Pánek, J., Krásný, L., Bobek, J., Ježková, E., Korelusová, J., and Vohradský, J. (2011). The suboptimal structures find the optimal RNAs: Homology search for bacterial non-coding RNAs using suboptimal RNA structures. Nucleic Acids Res. *39*, 3418–3426.

Papoušková, V., Kadeřávek, P., Otrusinová, O., Rabatinová, A., Šanderová, H., Nováček, J., Krásný, L., Sklenář, V., and Žídek, L. (2013). Structural study of the partially disordered full-length δ subunit of RNA polymerase from *Bacillus subtilis*. ChemBioChem *14*, 1772–1779.

Partridge, S.R., Foulger, D., and Errington, J. (1991). The role of SigF in prespore-specific transcription in *Bacillus subtilis*. Mol. Microbiol. *5*, 757–767.

Paul, M., and Leibovici, L. (2013). Systematic review or meta-analysis? Their place in the evidence hierarchy. Clin. Microbiol. Infect. *20*, 97–100.

Paul, B.J., Ross, W., Gaal, T., and Gourse, R.L. (2004). rRNA Transcription in *Escherichia coli*. Annu. Rev. Genet. *38*, 749–770.

Penfold, C.A., and Wild, D.L. (2011). How to infer gene networks from expression profiles, revisited. Interface Focus *1*, 857–870.

Perederina, A., Svetlov, V., Vassylyeva, M.N., Tahirov, T.H., Yokoyama, S., Artsimovitch, I., and Vassylyev, D.G. (2004). Regulation through the secondary channel - Structural framework for ppGpp-DksA synergism during transcription. Cell *118*, 297–309.

Pero, J., Nelson, J., and Fox, T.D. (1975). Highly asymmetric transcription by RNA polymerase containing phage-SP01-induced polypeptides and a new host protein. Proc. Natl. Acad. Sci. U. S. A. *72*, 1589–1593.

Petit, M.A., Dervyn, E., Rose, M., Entian, K.D., McGovern, S., Ehrlich, S.D., and Bruand, C. (1998). PcrA is an essential DNA helicase of *Bacillus subtilis* fulfilling functions both in repair and rolling-circle replication. Mol. Microbiol. *29*, 261–273.

Piersimoni, C., and Scarparo, C. (2009). Extrapulmonary infections associated with nontuberculous mycobacteria in immunocompetent persons. Emerg. Infect. Dis. *15*, 1351–1358.

Piggot, P.J., and Hilbert, D.W. (2004). Sporulation of *Bacillus subtilis*. Curr. Opin. Microbiol. *7*, 579–586.

Podlesek Z., Comino A., Herzog-Velikonja B., Zgur-Bertok D., Komel R., G.M. (1995). *Bacillus licheniformis* bacitracin-resistance ABC transporter: relationship to mammalian multidrug resistance. Mol. Microbiol. *16*, 969–976.

Prajapati, R.K., Sengupta, S., Rudra, P., and Mukhopadhyay, J. (2015). *Bacillus subtilis* δ functions as a transcriptional regulator by facilitating the open complex formation. J. Biol. Chem. jbc.M115.686170.

Predich, M., Nair, G., and Smith, I. (1992). *Bacillus subtilis* early sporulation genes-kinA, spo0F, and spo0A are transcribed by the RNA polymerase containing Sigma H. J. Bacteriol. *174*, 2771–2778.

Price, C.. W., and Doi, R.H. (1985). Genetic mapping of rpoD implicates the major sigma factor of *Bacillus subtilis* RNA polymerase in sporulation initiation. Mol Gen Genet. *201*, 88–95.

Qi, Y., and Hulett, F.M. (1998). Pho-P and RNA polymerase σ^A holoenzyme are sufficient for transcription of Pho regulon promoters in *Bacillus subtilis*: PhoP-P activator sites within the

coding region stimulate transcription in vitro. Mol. Microbiol. *28*, 1187–1197.

Rabatinová, A., Šanderová, H., Matějčková, J.J., Korelusová, J., Sojka, L., Barvík, I., Veronika Papoušková, Sklenář, V., Žídek, L., and Krásný, L. (2013). The δ subunit of RNA polymerase is required for rapid changes in gene expression and competitive fitness of the cell. J. Bacteriol. *195*, 2603–2611.

Ramaniuk, O., Převorovský, M., Pospíšil, J., Vítovská, D., Kofroňová, O., Benada, O., Schwarz, M., Šanderová, H., Hnilicová, J., and Krásný, L. (2018). PREPRINT σ^I from *Bacillus subtilis*: Impact on Gene Expression and Characterization of σ^I-dependent Transcription that Requires New Types of Promoters with Extended -35 and -10 Elements. J Bacteriol.

Reder, A., Höper, D., Weinberg, C., Gerth, U., Fraunholz, M., and Hecker, M. (2008). The Spx paralogue MgsR (YqgZ) controls a subregulon within the general stress response of *Bacillus subtilis*. Mol. Microbiol. *69*, 1104–1120.

Rediger, A., Geißen, R., Steuten, B., Heilmann, B., Wagner, R., and Axmann, I.M. (2012). 6S RNA - An old issue became blue-green. Microbiol. (United Kingdom) *158*, 2480–2491.

Revyakin, A., Ebright, R.H., and Strick, T.R. (2004). Promoter unwinding and promoter clearance by RNA polymerase: Detection by single-molecule DNA nanomanipulation. PNAS *101*, 4776–4780.

Rolfe, M.D., Rice, C.J., Lucchini, S., Pin, C., Thompson, A., Cameron, A.D.S., Alston, M., Stringer, M.F., Betts, R.P., Baranyi, J., et al. (2012). Lag phase is a distinct growth phase that prepares bacteria for exponential growth and involves transient metal accumulation. J. Bacteriol. *194*, 686–701.

Rollenhagen, C., Antelmann, H., Kirstein, J., Delumeau, O., Hecker, M., and Yudkin, M.D. (2003). Binding of sigma(A) and sigma(B) to core RNA polymerase after environmental stress in Bacillus subtilis. J. Bacteriol. *185*, 35–40.

Ross, W., Thompson, J.F., Newlands, J.T., and Gourse, R.L. (1990). *E.coli* Fis protein activates ribosomal RNA transcription in vitro and in vivo. EMBO J. *9*, 3733–3742.

Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R.L. (1993). A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. Science (80). *262*, 1407–1413.

Roussel, M.R., and Zhu, R. (2006). Stochastic kinetics description of a simple transcription model. Bull. Math. Biol. *68*, 1681–1713.

Ruff, E., Record, M., and Artsimovitch, I. (2015). Initial Events in Bacterial Transcription Initiation. Biomolecules *5*, 1035–1062.

Salzberg, L.I., Powell, L., Hokamp, K., Botella, E., Noone, D., and Devine, K.M. (2013). The WalRK (YycFG) and σI RsgI regulators cooperate to control CwlO and LytE expression in exponentially growing and stressed *Bacillus subtilis* cells. Mol. Microbiol. *87*, 180–195.

Sanderson, A., Mitchell, J.E., Minchin, S.D., and Busby, S.J.W. (2003). Substitutions in the Escherichia coli RNA polymerase σ^{70} factor that affect recognition of extended -10 elements at promoters. FEBS Lett. *544*, 199–205.

Santangelo, T.J., and Artsimovitch, I. (2011). Termination and antiterimation: RNA polymerase runs a stop sign. Nat.Rev.Microbiol. *9*, 319–329.

Schirner, K., and Errington, J. (2009). The cell wall regulator σI specifically suppresses the lethal phenotype of mbl mutants in *Bacillus subtilis*. J. Bacteriol. *191*, 1404–1413.

Schneider, D.A., Gaal, T., and Gourse, R.L. (2002). NTP-sensing by rRNA promoters in *Escherichia coli* is direct. Proc. Natl. Acad. Sci. *99*, 8602–8607.

Schneider, D.A., Ross, W., and Gourse, R.L. (2003). Control of rRNA expression in *Escherichia coli*. Curr. Opin. Microbiol. *6*, 151–156.

Serizawa, M., Yamamoto, H., Yamaguchi, H., Fujita, Y., Kobayashi, K., Ogasawara, N., and Sekiguchi, J. (2004). Systematic analysis of SigD-regulated genes in *Bacillus subtilis* by DNA microarray and Northern blotting analyses. Gene *329*, 125–136.

Seydlová, G., Pohl, R., Zborníková, E., Ehn, M., Šimák, O., Panova, N., Kolář, M., Bogdanová, K., Večeřová, R., Fišer, R., et al. (2017). Lipophosphonoxins II: Design, Synthesis, and Properties of Novel Broad Spectrum Antibacterial Agents. J. Med. Chem. *60*, 6098–6118.

Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiß, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., Reinhardt, R., et al. (2010). The primary transcriptome of the major human pathogen *Helicobacter pylori*. Nature *464*, 250–255.

Shephard, L., Dobson, N., and Unrau, P.J. (2010). Binding and release of the 6S transcriptional control RNA. RNA *16*, 885–892.

Shimada, T., Yamazaki, Y., Tanaka, K., and Ishihama, A. (2014). The Whole Set of Constitutive Promoters Recognized by RNA Polymerase RpoD Holoenzyme of *Escherichia coli*. PLoS One *9*.

Shlomi, T., Cabili, M.N., Herrgård, M.J., Palsson, B., and Ruppin, E. (2008). Network-based

prediction of human tissue-specific metabolism. Nat. Biotechnol. *26*, 1003–1010.

Shorenstein, R.G., and Losick, R. (1973a). Comparative Size and Properties of the Sigma Ribonucleic Acid Polymerase from *Bacillus subtilis* and *Escherichia coli*. 6170–6174.

Shorenstein, R.G., and Losick, R. (1973b). Purification and properties of the sigma subunit of ribonucleic acid polymerase from vegetative *Bacillus subtilis*. J Biol Chem *248*, 6163–6169.

Siegele, D.A., Hu, J.C., Walter, W.A., and Gross, C.A. (1989). Altered promoter recognition by mutant forms of the σ^{70} subunit of *Escherichia coli* RNA polymerase. J. Mol. Biol. *206*, 591–603.

Sierro, N., Makita, Y., De hoon, M., and Nakai, K. (2008). DBTBS: A database of transcriptional regulation in Bacillus subtilis containing upstream intergenic conservation information. Nucleic Acids Res. *36*, 93–96.

Šiková, M., Janoušková, M., Ramaniuk, O., Páleníková, P., Pospíšil, J., Bartl, P., Suder, A., Pajer, P., Kubičková, P., Pavliš, O., et al. (2018). Ms1 RNA increases the amount of RNA polymerase in *Mycobacterium smegmatis*. Mol. Microbiol. 0–2.

Sojka, L., Kouba, T., Barvíkš, I., Šanderová, H., Maderová, Z., Jonák, J., and Krásný, L. (2011). Rapid changes in gene expression: DNA determinants of promoter regulation by the concentration of the transcription initiating NTP in Bacillus subtilis. Nucleic Acids Res. *39*, 4598–4611.

Sonenshein, A.L. (2000). Control of sporulation initiation in *Bacillus subtilis*. Curr. Opin. Microbiol. *3*, 561–566.

Soutourina, O.A., Monot, M., Boudry, P., Saujet, L., Pichon, C., Sismeiro, O., Semenova, E., Severinov, K., Le Bouguenec, C., Coppée, J.Y., et al. (2013). Genome-Wide Identification of Regulatory RNAs in the Human Pathogen *Clostridium difficile*. PLoS Genet. *9*.

Spiegelman, G.B., Hiatt, W.R., and Whiteley, H.R. (1978). Role of the 21.000 Molecular Weight Polypeptide of *Bacillus subtilis* RNA Polymerase in RNA Synthesis. J Biol Chem *253*, 1756–1765.

Stamatakis, A., Hoover, P., and Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. *57*, 758–771.

Straney, D.C., and Crothers, D.M. (1985). Intermediates in transcription initiation from the *E. coli* lac UV5 promoter. Cell *43*, 449–459.

Sudalaiyadum Perumal, A., Vishwakarma, R.K., Hu, Y., Morichaud, Z., and Brodolin, K. (2018). RbpA relaxes promoter selectivity of *M. tuberculosis* RNA polymerase. Nucleic Acids Res. *46*,

10106–10118.

Sun, D., Stragier, P., and Setlow, P. (1989). Identification of a new σ-factor involved in compartmentalized gene expression during sporulation of *Bacillus subtilis*. Genes Dev. *3*, 141–149.

Sun, D., Liu, C., Zhu, J., and Liu, W. (2017). Connecting Metabolic Pathways: Sigma Factors in *Streptomyces spp*. Front. Microbiol. *8*, 1–7.

Sun, G., Sharkova, E., Chesnut, R., Birkey, S., Duggan, M.F., Sorokin, A., Pujic, P., Ehrlich, S.D., and Hulett, F.M. (1996). Regulators of aerobic and anaerobic respiration in *Bacillus subtilis*. J. Bacteriol. *178*, 1374–1385.

Sutherland, C., and Murakami, K.S. (2018). An Introduction to the Structure and Function of the Catalytic Core Enzyme of *Escherichia coli* RNA Polymerase. EcoSal Plus *8*, 1–9.

Tare, P., China, A., and Nagaraja, V. (2012). Distinct and Contrasting Transcription Initiation Patterns at *Mycobacterium tuberculosis* Promoters. PLoS One *7*, 1–12.

Thomason, M.K., Bischler, T., Eisenbart, S.K., Förstner, K.U., Zhang, A., Herbig, A., Nieselt, K., Sharma, C.M., and Storza, G. (2015). Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. J. Bacteriol. *197*, 18–28.

Thompson, J., Pikis, A., Ruvinov, S.B., Henrissat, B., Yamamoto, H., and Sekiguchi, J. (1998). The Gene glvA of *Bacillus subtilis* 168 Encodes a Metal-requiring, NAD(H)-dependent 6-Phospho-α-glucosidase. J Biol Chem *273*, 27347–27356.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.

To, C.C., and Vohradsky, J. (2010). Measurement variation determines the gene network topology reconstructed from experimental data: a case study of the yeast cyclin network. FASEB J. *24*, 3468–3478.

Tojo, S., Hirooka, K., and Fujita, Y. (2013). Expression of kinA and kinB of *Bacillus subtilis*, necessary for sporulation initiation, is under positive stringent transcription control. J. Bacteriol. *195*, 1656–1665.

Travers, A.A. (1980). Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. J. Bacteriol. *141*, 973–976.

Trotochaud, A.E., and Wassarman, K.M. (2005). A highly conserved 6S RNA structure is required for regulation of transcription. Nat. Struct. Mol. Biol. *12*, 313–319.

Tseng, C.-L., and Shaw, G.-C. (2008). Genetic evidence for the actin homolog gene mreBH and the bacitracin resistance gene bcrC as targets of the alternative sigma factor σI of *Bacillus subtilis*. J. Bacteriol. *190*, 1561–1567.

Tseng, C.-L., Chen, J.-T., Lin, J.-H., Huang, W.-Z., and Shaw, G.-C. (2011). Genetic evidence for involvement of the alternative sigma factor σI in controlling expression of the cell wall hydrolase gene *lytE* and contribution of LytE to heat survival of *Bacillus subtilis*. Arch. Microbiol. *193*, 677–685.

Turnbough, C.L. (2008). Regulation of bacterial gene expression by the NTP substrates of transcription initiation. Mol. Microbiol. *69*, 10–14.

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., and Borukhov, S. (2002). Crystal structure of a bacterial /*Thermus thermophilus*/ RNA polymerase holoenzyme at 2.6A˚ resolution. Nature *417*, 0–7.

Vohradsky, J. (2012). Stochastic simulation for the inference of transcriptional control network of yeast cyclins genes. Nucleic Acids Res. *40*, 7096–7103.

Vohradský, J. (2001). Neural network model of gene expression. FASEB J. *15*, 846–854.

Völker, U., Engelmann, S., Maul, B., Riethdorf, S., Völker, A., Schmid, R., Mach, H., and Hecker, M. (1994). Analysis of the induction of general stress proteins of *Bacillus subtilis*. Microbiology *140*, 741–752.

Voskuil, M.I., and Chambliss, G.H. (1998). The -16 region of *Bacillus subtilis* and other gram-positive bacterial promoters. Nucleic Acids Res. *26*, 3584–3590.

Vu, T.T., and Vohradsky, J. (2007). Nonlinear differential equation model for quantification of transcriptional regulation applied to microarray data of *Saccharomyces cerevisiae*. Nucleic Acids Res. *35*, 279–287.

Vu, T.T., and Vohradsky, J. (2009). Inference of active transcriptional networks by integration of gene expression kinetics modeling and multisource data. Genomics *93*, 426–433.

Wade, J.T., and Grainger, D.C. (2014). Pervasive transcription: Illuminating the dark matter of bacterial transcriptomes. Nat. Rev. Microbiol. *12*, 647–653.

Wade, J.T., Roa, D.C., Grainger, D.C., Hurd, D., Busby, S.J.W., Struhl, K., and Nudler, E. (2006).

Extensive functional overlap between σ factors in *Escherichia coli*. Nat. Struct. Mol. Biol. *13*, 806–814.

Walker, E., Hernandez, A. V, and Kattan, M.W. (2008). Meta-analysis: Its strengths and limitations. Cleve. Clin. J. Med. *75*, 431–439.

Walker, K.A., Mallik, P., Pratt, T.S., and Osuna, R. (2004). The *Escherichia coli* fis promoter is regulated by changes in the levels of its transcription initiation nucleotide CTP. J. Biol. Chem. *279*, 50818–50828.

Wang, Y.X.R., and Huang, H. (2014). Review on statistical methods for gene network reconstruction using expression data. J. Theor. Biol. *362*, 53–61.

Washburn, R.S., and Gottesman, M.E. (2015). Regulation of transcription elongation and termination. Biomolecules *5*, 1063–1078.

Wassarman, K.M., and Storz, G. (2000). 6S RNA regulates *E. coli* RNA polymerase activity. Cell *101*, 613–623.

Watanabe, T., Sugiura, M., and Sugita, M. (1997). A novel small stable RNA, 6Sa RNA, from the cyanobacterium *Synechococcus sp*. strain PCC6301. FEBS Lett. *416*, 302–306.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189–1191.

Watson, S.P., Antonio, M., and Foster, S.J. (1998). Isolation and characterization of *Staphylococcus aureus* starvation-induced, stationary-phase mutants defective in survival or recovery. Microbiology *144*, 3159–3169.

Wehner, S., Damm, K., Hartmann, R.K., and Marz, M. (2014). Dissemination of 6S RNA among bacteria. RNA Biol. *11*, 1467–1478.

Weiss, A., and Shaw, L.N. (2015). Small things considered: The small accessory subunits of RNA polymerase in Gram-positive bacteria. FEMS Microbiol. Rev. *39*, 541–554.

Weissenmayer, B.A., Prendergast, J.G.D., Lohan, A.J., and Loftus, B.J. (2011). Sequencing illustrates the transcriptional response of *Legionella pneumophila* during infection and identifies seventy novel small non-coding RNAs. PLoS One *6*.

Weng, X., and Xiao, J. (2014). Spatial organization of transcription in bacterial cells. Trends Genet. *30*, 287–297.

Whipple FW, S.A. (1992). Mechanism of initiation of transcription by *Bacillus subtilis* RNA polymerase at several promoters. J Mol Biol.

Wiedermannová, J., Sudzinová, P., Kovaľ, T., Rabatinová, A., Šanderová, H., Ramaniuk, O., Rittich, Š., Dohnálek, J., Fu, Z., Halada, P., et al. (2014). Characterization of HelD, an interacting partner of RNA polymerase from *Bacillus subtilis*. Nucleic Acids Res. *42*, 5151–5163.

Wiggs, J.L., and Gilman, M.Z. (1981). Heterogeneity of RNA polymerase in *Bacillus subtilis*: evidence for an additional sigma factor in vegetative cells. Proc. Natl. Acad. Sci. USA *78*, 2762–2766.

Willkomm, D.K., Minnerup, J., Hüttenhofer, A., and Hartmann, R.K. (2005). Experimental RNomics in Aquifex aeolicus: Identification of small non-coding RNAs and the putative 6S RNA homolog. Nucleic Acids Res. *33*, 1949–1960.

Wipat, A., Carter, N., Brignell, S.C., Guy, B.J., Piper, K., Sanders, J., Emmerson, P.T., and Harwood, C.R. (1996). The dnaB-pheA (256°-240°) region of the *Bacillus subtilis* chromosome containing genes responsible for stress responses, the utilization of plant cell walls and primary metabolism. Microbiology *142*, 3067–3078.

Yakhnin, A. V., and Babitzke, P. (2002). NusA-stimulated RNA polymerase pausing and termination participates in the *Bacillus subtilis* trp operon attenuation mechanism in vitro. Proc. Natl. Acad. Sci. *99*, 11067–11072.

Yan, Y., Su, S., Meng, X., Ji, X., Qu, Y., Liu, Z., Wang, X., Cui, Y., Deng, Z., Zhou, D., et al. (2013). Determination of sRNA expressions by RNA-seq in *Yersinia pestis* grown in vitro and during infection. PLoS One *8*, 1–13.

Yang, W. (2010). Lessons Learned From UvrD Helicase : Mechanism For Directional Movement. Annu Rev Biophys. *39*, 367–385.

Yizhak, K., Benyamini, T., Liebermeister, W., Ruppin, E., and Shlomi, T. (2010). Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. Bioinformatics *26*, 255–260.

Young, B.A., Anthony, L.C., Gruber, T.M., Arthur, T.M., Heyduk, E., Lu, C.Z., Sharp, M.M., Heyduk, T., Burgess, R.R., and Gross, C.A. (2001). A coiled-coil from the RNA polymerase β' subunit allosterically induces selective nontemplate strand binding by σ^{70}. Cell *105*, 935–944.

Young, B.A., Gruber, T.M., and Gross, C.A. (2002). Views of transcription initiation. Cell *109*, 417–420.

Zachrdla, M., Padrta, P., Rabatinová, A., Šanderová, H., Barvík, I., Krásný, L., and Žídek, L. (2017). Solution structure of domain 1.1 of the σ^A factor from *Bacillus subtilis* is preformed for binding to the RNA polymerase core. J. Biol. Chem. *292*, 11610–11617.

Zeng, X., Zhang, Y., Kwong, J.S.W., Zhang, C., Li, S., Sun, F., Niu, Y., and Du, L. (2015). The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: A systematic review. J. Evid. Based. Med. *8*, 2–10.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Crystal Structure of *Thermus aquaticus* Core RNA Polymerase at 3.3 Å Resolution DNA template, and finally releases itself and the com- pleted transcript from the DNA when a specific termination signal is encountered. The current view is that the. Cell *98*, 811–824.

Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M.X., Arnold, E., and Ebright, R.H. (2012). Structural basis of transcription initiation. Science (80). *338*, 1076–1080.

Zhang, Y., Mooney, R.A., Grass, J.A., Sivaramakrishnan, P., Herman, C., Landick, R., and Wang, J.D. (2014). DksA Guards Elongating RNA Polymerase Against Ribosome-Stalling-Induced Arrest. Mol. Cell *53*, 766–778.

Zheng, L., and Losick, R. (1990). Cascade regulation of spore coat gene expression in *Bacillus subtilis*. J. Mol. Biol. *212*, 645–660.

Zhu, B., and Stülke, J. (2018). SubtiWiki in 2018: From genes and proteins to functional network annotation of the model organism *Bacillus subtilis*. Nucleic Acids Res. *46*, D743–D748.

Zhulin, I.B. (2015). Databases for microbiologists. J. Bacteriol. *197*, 2458–2467.

Zuber, P. (2004). Spx-RNA Polymerase Interaction and Global Transcriptional Control during Oxidative Stress. J. Bacteriol. *186*, 1911–1918.

Zuber, U., Drzewiecki, K., and Hecker, M. (2001). Putative σ Factor σI (YkoZ) of *Bacillus subtilis* Is Induced by Heat Shock. J. Bacteriol. *183*, 1472–1475.

Zuo, Y., and Steitz, T.A. (2015). Crystal Structures of the *E. coli* Transcription Initiation Complexes with a Complete Bubble. Mol. Cell *58*, 534–540.

APPENDICES

Appendices attached to the Thesis contains four published articles:

Appendix 1 – published article:

Ramaniuk O, Převorovský M, Pospíšil J, Vítovská D, Kofroňová O, Benada O, Schwarz M, Šanderová H, Hnilicová J, Krásný L. σ^I from *Bacillus subtilis*: Impact on Gene Expression and Characterization of σ^I-dependent Transcription that Requires New Types of Promoters with Extended -35 and -10 Elements. J Bacteriol. 2018 Jun 18. pii: JB.00251-18. doi: 10.1128/JB.00251-18. PMID: 29914988

Appendix 2 – published article:

Ramaniuk O, Černý M, Krásný L, Vohradský J. Kinetic modelling and meta-analysis of the *B. subtilis* SigA regulatory network during spore germination and outgrowth. Biochim Biophys Acta. 2017 Aug; 1860(8):894-904. doi: 10.1016/j.bbagrm.2017.06.003. Epub 2017 Jun 22. PMID:28648455

Appendix 3 – published article:

Wiedermannová J, Sudzinová P, Kovaľ T, Rabatinová A, Šanderová H, **Ramaniuk O**, Rittich Š, Dohnálek J, Fu Z, Halada P, Lewis P, Krásný L. Characterization of HelD, an interacting partner of RNA polymerase from *Bacillus subtilis*. Nucleic Acids Res. 2014 Apr;42(8):5151-63. doi: 10.1093/nar/gku113. Epub 2014 Feb 11.

Appendix 4 – published article:

Šiková M, Janoušková M, **Ramaniuk O**, Páleníková P, Pospíšil J, Bartl P, Suder A, Pajer P, Kubičková P, Pavliš O, Hradilová M, Vítovská D, Šanderová H, Převorovský M, Hnilicová J, Krásný L. Ms1 RNA increases the amount of RNA polymerase in *Mycobacterium smegmatis*. Mol Microbiol. 2018 Nov 14. doi: 10.1111/mmi.14159. [Epub ahead of print]