Charles University in Prague
Faculty of Science

Study programme:
Biology
Branch of study:
Genetics, Molecular Biology and Virology

Bc. Petra Páleníková

Regulation of expression of Ms1, a sRNA from *Mycobacterium smegmatis*

Regulace exprese Ms1, sRNA z *Mycobacterium smegmatis*

Diploma thesis

Supervisor: Mgr. RNDr. Libor Krásný, Ph.D.

Prague, 2016

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 29. 04. 2016 Podpis ……

Poďakovanie:

Ďakujem Mgr. RNDr. Liborovi Krásnému, Ph.D. za vedenie práce a cenné rady pri jej vypracovávaní. Tiež ďakujem všetkým kolegom z Laboratória molekulárnej genetiky baktérií za príjemnú atmosféru a pomoc pri práci.

Abstrakt:

V průběhu růstu musí bakterie čelit různým vlivům prostředí, jako jsou změna teploty, nedostatek živin a jiné. Při adaptaci na nepříznivé podmínky dochází u bakterií k přizpůsobení genové exprese. Jedním typem regulace genové exprese je regulace pomocí malých nekódujících RNA (sRNA). Dobře studovanou bakteriální sRNA je 6S RNA, která se váže na holoenzym RNA polymerázy. Avšak tato RNA byla nalezena jen u některých druhů baterií. Jedním z druhů, které pravděpodobně nemají 6S RNA jsou mykobakterie. Místo toho byla u *Mycobacterium smegmatis* nalezena jiná malá RNA – Ms1. Ms1 má strukturu podobnou 6S RNA a skutečně byla původně nalezena jako strukturní homolog 6S RNA. Ukázalo se však, že Ms1 se neváže na holoenzym RNA polymerázy, ale váže se na jádro RNA polymerázy. Jedná se proto o malou RNA významně odlišnou od 6S RNA.

Tato práce popisuje regulaci exprese Ms1. Z regulační oblasti před genem pro Ms1 byli připraveny promotorové fragmenty různé délky. Tyto fragmenty byly vloženy před reportérový gen *lacZ*. Aktivita těchto promotorů byla testována v exponenciální, časné a pozdní stacionární fázi růstu a v průběhu stresu. Na základě provedených experimentů byla určena sekvence promotorového jádra a regulační oblasti, do kterých by se mohl vázat transkripční faktor. Aktivita promotoru pro Ms1 se zvýšila při prodloužení promotorové oblasti a při přechodu do stacionární fáze růstu. Po vystavení bakterií hyperosmotickému stresu došlo k mírnému snížení promotorové aktivity. Etanolový stres způsobil nárůst aktivity promotoru. Nakonec byl navržen model regulace exprese Ms1

Klíčová slova:

sRNA, RNA polymeráza, regulace transkripce, RNA-proteinové interakce, transkripční faktor, mykobakterie

Abstract:

Bacteria are exposed to various environmental conditions during their growth. They have to cope with rapid changes in temperature, lack of nutrition, etc. To survive, bacteria alter their gene expression. One type of regulation of gene expression is regulation by small RNAs (sRNAs). In bacteria, a well-studied sRNA is 6S RNA that binds to the RNA polymerase holoenzyme. However, 6S RNA has not been identified in several bacterial species. Mycobacteria are a genus that probably does not have 6S RNA. Instead, *Mycobacterium smegmatis* possess another sRNA – Ms1. Ms1 structurally resembles 6S RNA and indeed it was first identified as a 6S RNA structural homologue. However, Ms1 binds to RNAP devoid of any sigma factor, and, therefore, is significantly distinct from 6S RNA.

This work describes regulation of expression of Ms1. DNA fragments of different length from the region upstream of the Ms1 gene were prepared. These fragments were fused to the *lacZ* reporter gene and their activity was tested in different growth phases and under stress. This allowed identification and characterization of the core promoter sequence and regulatory sequences that might interact with transcription factor(s). Promoter activity increased with increased length of the promoter fragment and after transition into stationary phase. Transcription slightly decreased after hyperosmotic stress and increased after ethanol stress. Finally, a model of Ms1 expression regulation is proposed.

Keywords:

sRNA, RNA polymerase, transcription regulation, RNA-protein interactions, transcription factor, mycobacterium

Content

Abbreviations

bp	Base pair
CBB	Coomassie Briliant Blue
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
ddH_2O	Double-distilled water
dGTP	Deoxyguanosine triphosphate
DMSO	Dimethyl sulfoxide
dsDNA	Double stranded DNA
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
EDTA	Ethylenediaminetetraacetic acid
Eσ	RNA polymerase holoenzyme
GFP	Green fluorescent protein
ChIP-Seq	Chromatin immunoprecipitation followed by sequencing
iNTP	Initiating NTP
IP	immunoprecipitation
MALDI-TOF	Matrix Assisted Laser Desorption Ionization Time-of-Flight
mRNA	Messenger RNA
Ms1nb	Ms1 without internal bubble
nt	Nucleotide
NTP	Nucleoside triphosphate
OD	Optical density
ONPG	ortho-Nitrophenyl-β-galactoside
PCR	Polymerase chain reaction
PMSF	Phenylmethane sulfonyl fluoride
ppGpp	Guanosine tetraphosphate
pRNA	Product RNA
RNAP	RNA polymerase
RNA-Seq	RNA sequencing
RP_c	Promoter closed complex
RP_{itc}	Promoter initial transcribing complex
RP_o	Promoter open complex
rRNA	Ribosomal RNA
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIPHT	sRNA Identification Protocol using High-throughput Technologies
sRNA	Small RNA
SSB	Single-stranded DNA-binding protein
TopoI	DNA topoisomerase I
TopoI RS	DNA topoisomerase I recognition site
TSS	Transcription start site
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside

1. Introduction

Mycobacteria are a bacterial genus that comprises lethal human pathogens *Mycobacterium tuberculosis* and *Mycobacterium leprae*. A considerable part of human population is infected with tuberculosis and 1.5 million of people died from tuberculosis in 2014 (World Health Organization, 2015). To be able to fight these diseases and identify potential targets for treatment, it is crucial that we understand molecular mechanisms of bacterial cells.

Bacterial cells have to cope with different environments during their growth. To adapt to these changes bacteria constantly alter gene expression. The process of regulation of gene expression is highly complex. To study these processes a non-pathogenic species is often used as it is easier to handle and still provides valuable information of how these processes work. In mycobacteria, such a non-pathogenic model is *Mycobacterium smegmatis*.

Various types of regulation of gene expression exist. In recent years, the importance of small regulatory RNAs (sRNAs) in gene expression is becoming increasingly clear. *M. smegmatis* possesses a highly abundant sRNA – Ms1. This sRNA binds to the RNA polymerase core and alters transcription of many genes. A Ms1 homologue, MTS2823, is also present in *M. tuberculosis*.

Ms1 functions as a transcription regulator, however, regulation of Ms1 transcription itself is unknown. In this diploma thesis I characterized several aspects of regulation of expression of Ms1.

2. Aims of the thesis

The main aim of my thesis was to characterize expression of Ms1 at the level of transcription:

- Characterize Ms1 expression in different growth phases
- Characterize Ms1 expression under stress
- Identify the promoter for Ms1
- Identify transcription factor(s) interacting with the Ms1 promoter region

3. Literature review

3.1 Transcription

Transcription is a process in which a new strand of ribonucleic acid (RNA), complementary to the DNA template strand, is synthesized. RNA is synthesized by an enzyme called RNA polymerase (RNAP). Transcription can be divided into 3 steps – initiation, elongation and termination. Each one of these steps can be regulated.

3.1.1 RNA polymerase

Bacteria contain only one type of RNAP. RNAP is a multi-subunit enzyme consisting of five subunits – two α (~35 kDa), β (~124 kDa), β' (~171 kDa) and ω (~10.5 kDa) subunits (Zhang et al., 1999). These five subunits constitute the core enzyme. To initiate transcription in a sequence dependent manner, an additional subunit needs to be bound to RNAP. This subunit is called sigma. Usually the bacterial cell contains several types of sigma factors. The main sigma factor is responsible for transcription of housekeeping genes and alternative sigma factors are involved in stress response and/or adaptation of the bacterium. The RNAP core enzyme with a bound sigma factor is called RNA polymerase holoenzyme (Eσ). **Table 1** shows sizes of RNAP core subunits for different bacteria.

Table 1. Number of amino acids in RNAP subunits for different bacteria.

	Number of amino acids in RNAP subunit					
Organism	**α**	**β**	**β'**	**ω**	**δ**	**ε**
E. coli	329	1342	1407	91	0	0
B. subtilis	314	1193	1199	67	173	69
M. smegmatis	350	1169	1317	107	0	0
M. tuberculosis	347	1172	1316	110	0	0

Data were obtained from databases. *E. coli* data from http://ecogene.org/ (Zhou and Rudd, 2013), *B. subtilis* data from http://bsubcyc.org/ (Caspi et al., 2014), *M. smegmatis* data from http://mycobrowser.epfl.ch/smegmalist.html (Kapopoulou et al., 2011), *M. tuberculosis* data from http://tuberculist.epfl.ch/ (Lew et al., 2011).

The alpha subunit of RNA polymerase is encoded by the *rpoA* gene. It contains structured N-terminal and C-terminal domains. These two domains are interconnected by an unstructured linker. The C-terminal domain can interact with promoter upstream elements (Blatter et al., 1994). Alpha subunit from *M. tuberculosis* is 347 amino acids long and displays 48% protein identity with α subunit of *B. subtilis* (Healy et al., 1998).

The beta subunit is encoded by the *rpoB* gene. Mutations in *rpoB* can cause resistance to rifampicin – a drug mainly used to treat tuberculosis. *M. tuberculosis* strains with mutations in *rpoB*, and therefore resistant to rifampicin, usually have also mutations in *rpoA* and *rpoC* genes. These mutations appear to be fitness-compensatory for the bacterium (Brandis and Hughes, 2013).

The beta' subunit is the largest subunit of RNAP. It is encoded by the *rpoC* gene. Together with the β subunit they create the catalytic center of RNAP (Murakami, 2015).

The omega subunit is encoded by the *rpoZ* gene. In *M. smegmatis*, this subunit is not essential for the bacterium as deletion mutants are viable. However, they possess many morphological changes. They are deficient in extracellular matrix production resulting in defective biofilm formation and also have reduced sliding motility (Mathew et al., 2006).

3.1.2 Promoter

During transcription initiation RNA polymerase binds DNA at a promoter located upstream of the transcription start site. Within the promoter, we can distinguish several conserved regions. Most significant are -10 and -35 conserved regions. Some promoters also contain the extended -10 region. In mycobacteria the extended -10 region consists of the TGN motif present directly upstream of the -10 region (Bashyam and Tyagi, 1998). Sequence of these regions is recognized by a σ subunit of RNAP. Different σ subunits recognize different promoter sequences enabling changes in transcription according to environmental conditions. Another conserved region might be present further upstream of the -35 region. This region is called the UP element. UP elements are AT rich and enhance promoter strength as they are recognized by α subunits of RNAP (Ross et al., 1993). The region between the -10 region and the +1 transcription start site is called discriminator. It makes direct interactions with sigma factor region 1.2 (Haugen et al., 2006). A schematic view of RNAP interactions with promoter are shown at **Figure 1**.

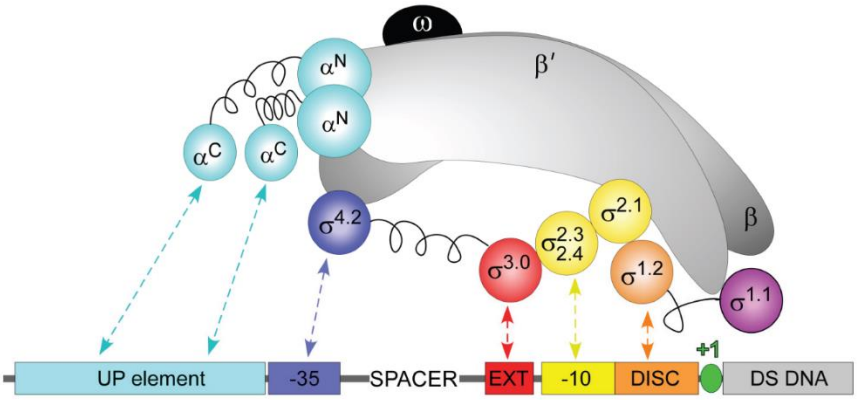

Figure 1. Schematic view of RNAP interactions with promoter DNA (Ruff et al., 2015). DISC – discriminator, DS DNA – downstream DNA.

Various methods can be used to identify promoters. Before the availability of genome sequences of bacteria and without known promoter consensus sequences, it was possible to screen for promoters using a promoter selection vector. After isolation of genomic DNA from the bacterium, the DNA was digested by an enzyme and fragments of appropriate length were cloned into the vector. This vector had to contain an origin of replication and a promoterless reporter gene which could be used to evaluate activity of cloned fragments. Using this technique, several mycobacterial promoters have been identified (Gupta et al., 1993).

Promoter DNA fragments acquired as described above can be further sequenced and alignment of these sequences can be used to predict -10 and -35 consensus sequences of the promoter. From randomly selected *M. smegmatis* and *M. tuberculosis* promoters the -10 consensus sequence was determined as TATAaT and TAYgAT, respectively (lower case letter represent conservation less than 50 %, Y stands for a pyrimidine base) (Bashyam et al., 1996).

3.1.3 Steps of transcription

Transcription imitation is the first step in transcription. It is a complex process that can be further divided into several steps. The first step of transcription initiation is formation of the closed complex (RP_C) between RNAP holoenzyme and promoter DNA. DNA enters RNAP as a double-strand. During isomerization, the promoter region from -11 to +2 is unwound, positioning the +1 nucleotide into the active site (Gries et al., 2010). After formation of the transcription bubble, RNAP and promoter form the open complex (RP_o). In RP_o the downstream end of DNA interacts with the β' subunit and the upstream end with both β and β' subunits of RNAP. Template strand positions +4 to +1 and non-template strand position +4 bind deep within the active-centre cleft (Naryshkin et al., 2000). In the presence of NTPs, RNAP forms initial transcribing complex (RP_{itc}) and starts synthesizing RNA. At this step, abortive transcripts can be formed. After synthesis of 9-11 nucleotides RNAP escapes promoter and continues to elongate the transcript. Both abortive initiation and promoter escape involve scrunching when RNAP pulls downstream DNA into itself. During productive initiation the extent of scrunching is 9 ± 2 bp (Revyakin et al., 2006).

During **elongation** RNAP moves along the DNA molecule and synthesizes RNA strand in the 5' → 3' direction. Elongation rates differ for different transcripts. For *rrn* operons in *E. coli*, the average rate is 80-90 nt/s, ranging from 65 nt/s within 16S and 23S rRNA genes, 250 nt/s in the 5' leader to 400 nt/s near the end of the operon (Dennis et al., 2009). It was assumed that the release of sigma factor from RNAP is necessary for transition into elongation. However,

σ^{70} from *E. coli* can remain bound to RNAP after initiation and throughout elongation (Harden et al., 2016).

When RNAP reaches the end of a gene or operon it finishes transcription of the nascent RNA strand and dissociates from DNA. This process is called **transcription termination**. There are two types of transcription termination. One is mediated by sequence of DNA and is called intrinsic transcription termination. Second is mediated by transcription termination factor Rho.

Intrinsic termination sites consist of a GC rich region forming a hairpin followed by a T stretch. Formation of the hairpin is accompanied by melting of the downstream A:U DNA-RNA hybrid. The hairpin destabilizes and irreversibly inactivates the ternary elongating complex at the termination site (Gusarov and Nudler, 1999).

Rho is a hexameric protein formed by six peanut-shaped identical subunits. Rho has two nucleotide binding sites and an ATP binding pocket (Skordalakes and Berger, 2003). Rho binds to the *rut* site on nascent RNA transcript, this site is necessary for transcription termination mediated by Rho protein and is generally GC rich (Graham and Richardson, 1998).

3.2 Transcription regulation in mycobacteria

Transcription regulation is a complex process which cannot be fully described within this literature review. In this chapter I will list main types of transcription regulation and give examples of transcription regulators in mycobacteria.

3.2.1 Sigma factors

Sigma factors regulate transcription by changing the specificity of RNA polymerase binding to promoters. Different sigma factors recognize different promoter sequences and alter transcription according to environmental conditions. Sigma factors can be divided into two families – σ^{70} and σ^{54}. In the primary structure of σ^{70} we can distinguish four conserved regions. According to the presence or absence of individual conserved regions, the σ^{70} family is divided into four groups. Alternative sigma factors belong to groups 2, 3 and 4 and can be regulated by binding of anti-sigma factors (Reviewed in Paget, 2015).

The number of sigma factors differ among bacterial species. *M. tuberculosis* has 13 sigma factors (Cole et al., 1998), *M. smegmatis* possess 26 sigma factors and 19 sigma factors were found in the genome of *Mycobacterium avium paratuberculosis* (Waagmeester et al., 2005).

In mycobacteria the primary sigma factor is σ^{A}. This sigma factor is essential for the bacterium. The amount of σ^{A} in *M. smegmatis* cells of the SM18 strain is constant during exponential phase of growth and increases in stationary phase reaching threefold higher value after 30 hours of cultivation. In mc^2 155 strain expression of σ^{A} is constant during exponential phase and decreases threefold during stationary phase (Gomez et al., 1998).

3.2.2 Transcription factors

Transcription factors can influence transcription both positively as activators and negatively as repressors. Transcription factors vary in their mechanism of action. They can interact with DNA, RNA and also RNAP. Bacteria with larger genomes encode more transcription factors compared to bacteria with smaller genomes. The number of encoded transcription factors is also affected by bacterial lifestyle. *M. tuberculosis* has fewer transcription factors compared to free-living bacteria with similar genome size (Cases et al., 2003).

Recently a new type of transcription regulator was found in *M. tuberculosis*. Protein Rv1222 from *M. tuberculosis* inhibits transcription by simultaneous binding to RNAP core and DNA. It anchors RNAP onto DNA and slows down mRNA synthesis. This inhibition occurs during promoter escape and transcription elongation steps of transcription. The interaction of Rv1222

with DNA is not sequence specific. When Rv1222 was expressed in *M. smegmatis* and *E. coli* it reduced growth rate of these bacteria (Rudra et al., 2015).

Apart from traditional studies of transcription regulation involving knock-out or overexpression of a single transcription factor and assessing its role in the cell, it has become possible to study complex regulatory networks. A study characterizing global DNA-protein interactions network was carried out on *M. tuberculosis*. Chromatin immunoprecipitation followed by sequencing (ChIP-Seq) analysis after induction of 154 transcription factors revealed ~16 000 binding events. Consensus motives for more than 50 transcription factors have been identified. Approximately 1/3 of binding events occurred within the promoter region and are probably involved in regulation of proximal gene expression (Minch et al., 2015).

3.2.3 Non-coding RNAs

RNAs can carry regulatory functions in the cell. There are various types of RNA regulators (reviewed in Waters and Storz, 2009). Riboswitches, elements at the 5' end of mRNA, can adopt different secondary structures based on the presence or absence of their ligand. These changes can trigger termination or anti-termination, induce or repress translation. Most sRNAs function by base pairing with their target mRNAs. sRNAs encoded *in-cis* have a long complementary region and very limited number of targets; sRNAs encoded *in-trans* have a broader range of target mRNAs with less complementarity. Small RNAs can also modulate protein activity. An example of such a sRNA is 6S RNA from *E. coli*, which will be further described in the next chapter.

To date, many sRNAs have been identified in mycobacterial species. Using SIPHT (sRNA Identification Protocol using High-throughput Technologies) algorithm, several sRNA candidates were predicted in *M. smegmatis* and *M. bovis* BCG. Northern blot analysis confirmed 37 sRNAs in *M. bovis* BCG (DiChiara et al., 2010). Further study revealed 17 and 23 novel sRNA in *M. smegmatis* and *M. bovis* BCG, respectively (Tsai et al., 2013). Sequencing of total RNA from exponentially growing *M. tuberculosis* cells showed high abundance of non-coding RNAs in this bacterium. 28% of reads was mapped to intergenic regions and antisense transcripts (Arnvig et al., 2011).

3.3 6S RNA

6S RNA, encoded by the *ssrS* gene in *E. coli*, was identified about half a century ago but its function remained unclear for a long time. First, it was discovered that 6S RNA is present in an amount of 1000 RNA molecules per DNA molecule and it is in complex with a protein (Lee et al., 1978). The amount of 6S RNA in the cell increases in the stationary phase of growth (Wassarman and Storz, 2000).

6S RNA interacts with RNA polymerase holoenzyme comprising σ^{70} subunit (Wassarman and Storz, 2000). By this interaction 6S RNA blocks transcription by $E\sigma^{70}$ and mediates transcriptional changes between exponential and stationary phase of growth.

When wild-type *E. coli* cells were cocultured with 6S RNA null mutants, the mutant cells displayed decreased viability after 2 days. When mutant cells were grown independently their growth curve was indistinguishable from wild-type cells during the first few days of cultivation. However, the viable-cell counts for null mutants decreased 5-fold compared to wild-type cells from 3 to 5 weeks (Trotochaud and Wassarman, 2004).

3.3.1 6S RNA can be found in many bacterial species

A recent study identified 1750 6S RNA candidate genes distributed among 1610 bacteria. In phylum *Actinobacteria* 6S RNAs have been predicted only among *Coriobacteridae*. No 6S RNA was predicted for phyla *Deinococcus*, *Elusimicrobia*, *Fibrobacteres*, *Tenericutes*, *Thermus* and *Thermotogae*. On the other hand, 95 species in the phylum *Firmicutes* have two copies of 6S RNA gene. Multiple 6S RNA genes have been identified also in other phyla (Wehner et al., 2014).

A sequence similarity search for homologs of the *E. coli ssrS* gene encoding 6S RNA identified several potential 6S RNA encoding genes in γ-proteobacteria. To identify 6S RNA in more distant species, biochemical methods including isolation of total RNA and immunoprecipitation were used. In *B. subtilis* two 6S RNAs (6S-1 and 6S-2 RNAs) interacting with $E\sigma^{A}$ were found (Trotochaud and Wassarman, 2005).

3.3.2 Structure of 6S RNA and its interactions with $E\sigma^{70}$

6S RNA has a structure resembling an open promoter containing double-stranded regions with a central single-stranded bubble. It has been shown that this structure is important for the function of 6S RNA both *in vitro* and *in vivo* (Trotochaud and Wassarman, 2005). It was concluded that interaction of 6S RNA with RNA polymerase is specific for $E\sigma^{70}$ (Trotochaud and Wassarman, 2005). However, further investigation showed that there might be weak specific interactions between 6S RNA and $E\sigma^{38}$ as well as with core RNA polymerase (Gildehaus et al., 2007).

Several conserved residues have been identified in the sequence of 6S RNA located in the -35 region. Mutations of these residues led to decreased binding of $E\sigma^{70}$. Phylogenetically conserved residues in the -10 region influence the 6S RNA release rate. The secondary structure of 6S RNA with colour indication of conservancy of individual nucleotides is shown in **Figure 2** (Shephard et al., 2010).

6S RNA is in direct contact with σ^{70}, β and β' subunits of RNAP (Gildehaus et al., 2007; Wassarman and Storz, 2000). Sigma factor region 4.2 is required for stable interaction between 6S RNA and $E\sigma^{70}$ (Cavanagh et al., 2008). This interaction is mediated mainly by charge-charge interactions between positively charged amino acids of σ^{70} and negatively charged backbone of RNA. Basic amino acids involved in interaction are located at C-terminal helix of region 4.2 of σ^{70} (Klocko and Wassarman, 2009).

6S RNA sites close to σ^{70} were identified by the chemical nuclease FeBABE tethered to different σ^{70} single-cysteine mutants in RNAP holoenzyme. Region 4.2 is in close proximity to the terminal 6S RNA hairpin helix (cleavage at positions 77/78, 90 and 101-103). U44 of 6S RNA is located close to sigma domain 3.2 which is adjacent to RNAP active site. Internal helix structure of 6S RNA is proximal to σ^{70} domains 2.1, 2.3 and 3.1 (cleavage at positions 124-126 and 65-67, 126-133 and 63-65, 71/72 and 119-121 respectively). **Figure 3** shows a model of 6S RNA bound to RNAP holoenzyme constructed based on cleavage data and known structure of *E. coli* RNA polymerase (Steuten et al., 2013).

A study of 6S RNA structure in solution suggested possible dimerization of 6S RNA molecules. However, it is not clear whether dimerization is biologically relevant or whether it was the result of used experimental procedures (Fadouloglou et al., 2015).

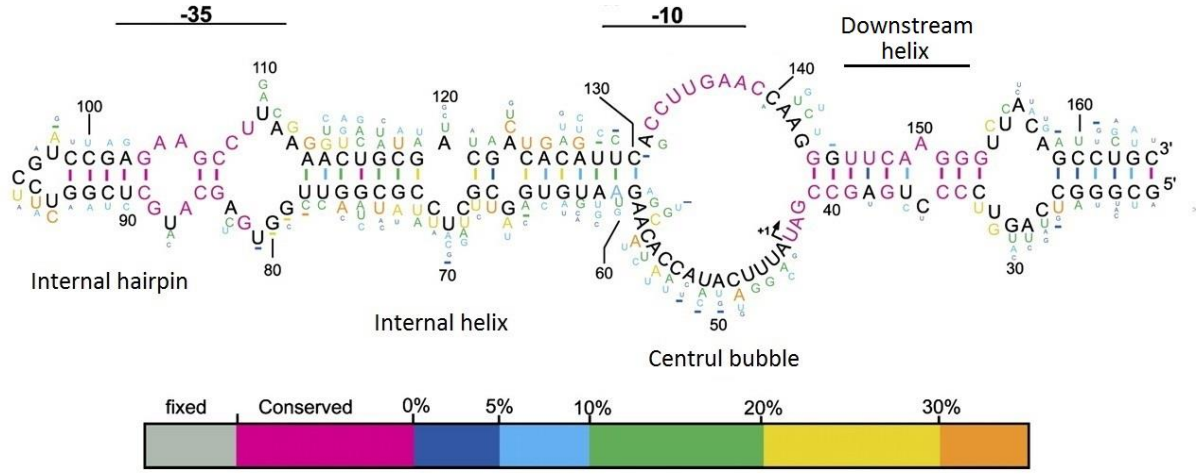

Figure 2. Secondary structure of 6S RNA. Different colours represent conservation of individual nucleotides and base-pairs. This model was constructed based on 6S RNA sequence found in Enterobacteriaceae, Alteromonadales and Vibrionales orders of γ-proteobacteria (Shephard et al., 2010).

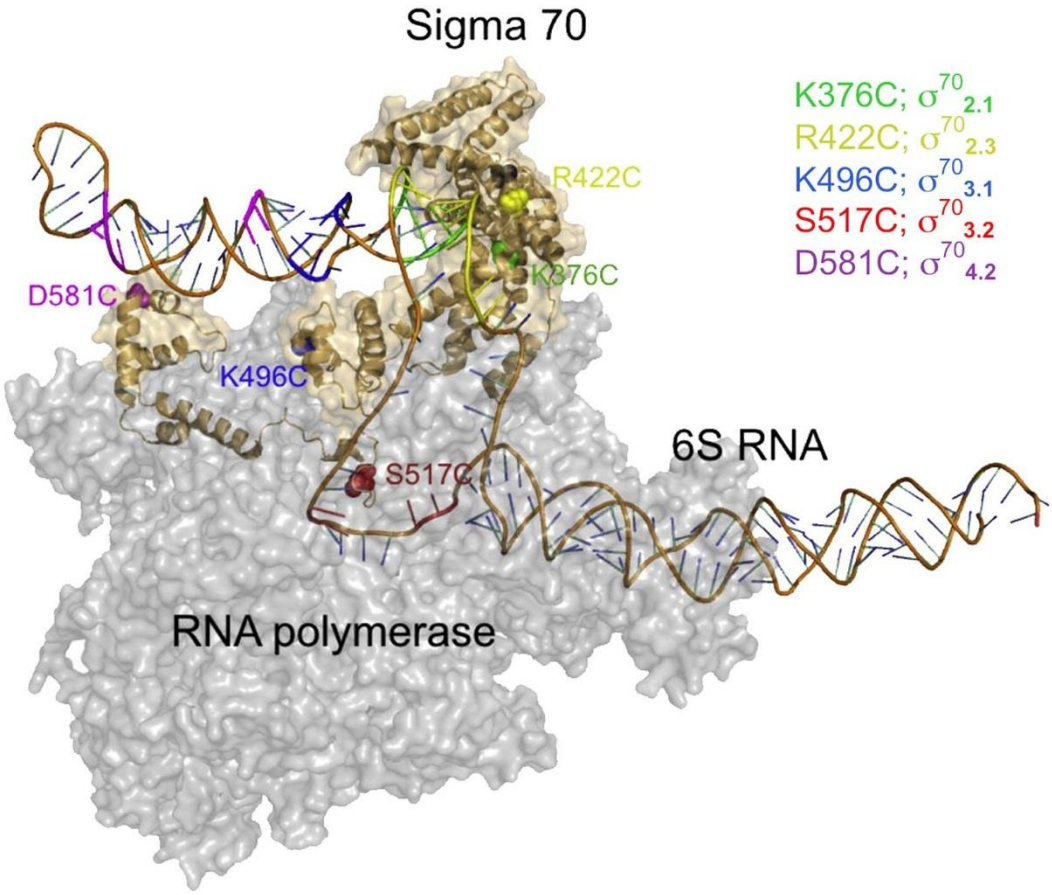

Figure 3. 3D model of 6S RNA bound to RNAP holoenzyme in complex with σ^{70}. Coloured spheres represent single-cysteine mutations in σ^{70} subunit. Cleavage sites within 6S RNA are coloured accordingly (Steuten et al., 2013).

3.3.3 Transcription of 6S RNA in *E. coli*

In *E. coli* 6S RNA is transcribed from two promoters P1 and P2. P1 is the proximal promoter to the mature 6S RNA. It is σ^{70}-dependent and produces a transcript with the 5' end at position -9G. Transcription from P2 promoter starts at -224A and this promoter is both σ^{70} and σ^{S} dependent. Mature 6S RNA is produced by cleavage by endonucleases. The longer transcript (P2 transcript) is processed exclusively by RNase E. The shorter transcript is cleaved by both RNase E and RNase G. Cleaved transcripts have a 5' ends in the range between -1 to +2 (Kim and Lee, 2004). A study of processing of precursor 6S RNA into the mature 6S RNA by RNase E cleavage showed that high substrate concentration inhibits RNase E activity on the precursor 6S RNA (Fadouloglou et al., 2015).

Using P1-l*acZ* and P2-*lacZ* fusion constructs it was determined that P1 transcription is 2 to 3 times higher in exponential phase compared to P2 transcription. Transcription from both promoters increases until late-exponential phase when P1 transcription rapidly decreases. Activity of P2 promoter also declines but to lesser extent. Transcription from P1 and P2 promoters was not affected in Δ*ssrS* cells. Changes in expression are shown in **Figure 4**. Overexpression of 6S RNA resulted in increase of P1 activity and decrease of P2 activity. This effect was probably mediated by the Fis transcription factor, as in Δ*fis* strains this changes in expression disappeared (Lee et al., 2013).

Transcription of 6S RNA is terminated by Rho factor. Transcription termination occurs at more than one site, but predominantly at position +279. Transcription termination sites are located within the open reading frame of the *ygfA* gene. The 3' end of the transcript is further processed by removing ~90 nucleotides most probably by exonuclease activity (Chae et al., 2011).

Expression of 6S RNA is regulated by H-NS, LRP and StpA transcription factors. Direct inhibition of 6S RNA transcription by these proteins was confirmed by *in vivo* experiments in bacterial strains mutated in individual transcription factors and by *in vitro* transcription. The transcription factor Fis has a dual effect on *in vitro* transcription. The presence of Fis in the reaction activated transcription from the P1 promoter and decreased transcription from the P2 promoter (Neusser et al., 2008).

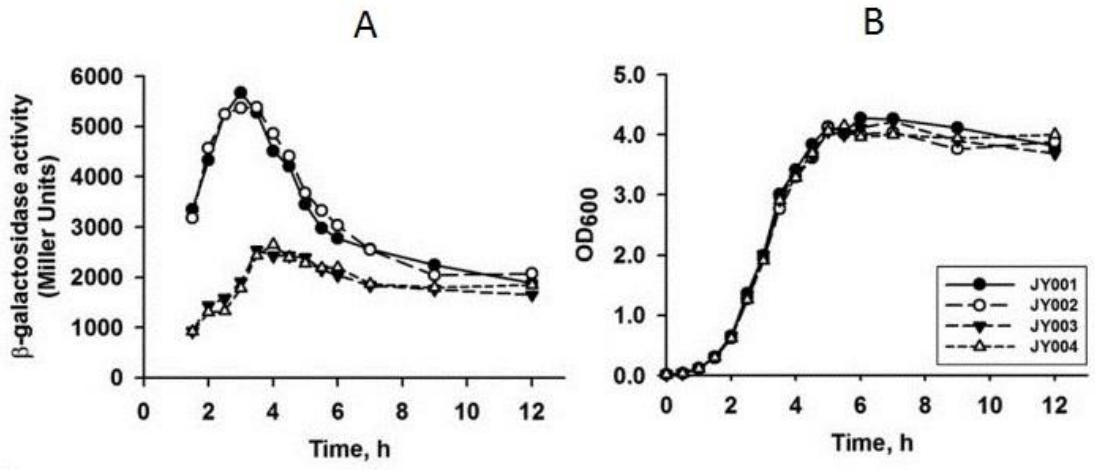

Figure 4. β-galactosidase activity (A) and growth curves (B) of P1-*lacZ* and P2-*lacZ* fusion constructs in *E.coli*. Strains: JY001 – *ssrS*$^+$(P1)-*lacZ*, JY002 – Δ*ssrS*(P1)-*lacZ*, JY003 – *ssrS*$^+$(P2)-*lacZ*, JY004 – Δ*ssrS*(P2)-*lacZ*. Cells were cultivated in rich LB medium (Lee et al., 2013).

3.3.4 6S RNA as a template for pRNA synthesis

6S RNA can serve as a template for RNA polymerase to synthesise short RNA products (pRNAs). Product RNAs are synthesized in outgrowth phase and this process enables release of RNAP from 6S RNA. Their length is either ~20 nucleotides or 170 to 220 nucleotides and products initiate at the U44 template position. Sigma subunit is released from the 6S RNA:Eσ^{70} complex during pRNA synthesis (Gildehaus et al., 2007; Wassarman and Saecker, 2006).

Synthesis of pRNAs occurs immediately after nutritional upshift. Short products of length 13-14 nucleotides are detectable in less than 15 seconds after addition of nutrients. Products up to 24 nucleotides appear within the next 4 minutes. Newly synthesized RNA can form complex with 6S RNA (Wurm et al., 2010). 6S RNA in complex with pRNA is not able to rebind RNAP (Cavanagh et al., 2012; Steuten and Wagner, 2012).

In *E. coli*, four structural states were identified during pRNA mediated release of 6S RNA from Eσ^{70}. The S1 state is the unbound 6S RNA. 6S RNA bound to Eσ^{70} represent the S2 state. In this state no pRNAs can be detected. In the S3 state, 6S RNA is bound to the RNAP core enzyme and contains 9-nt-long pRNA. The S4 state is the final state with 6S RNA:pRNA complex released from RNAP. The length of pRNA in the S4 state is 13 nt. Release of sigma factor is triggered by secondary structure rearrangement by formation of a hairpin between the downstream helix and the -10 region of 6S RNA. The release process is illustrated in **Figure 5** (Panchapakesan and Unrau, 2012; Steuten and Wagner, 2012).

In *B. subtilis* one of two 6S RNAs, the 6S-2 RNA, is a poor template for pRNA synthesis. Only a minimal amount of pRNA can be detected after incubation of 6S-2 RNAP:σ^A complex with nucleotides. Cells expressing 6S-2 RNA and not 6S-1 RNA exhibit delayed outgrowth from stationary phase. The presence of a mutant 6S (M68) RNA that cannot direct pRNA synthesis in *E. coli* cells lacking the endogenous 6S RNA has the same result. Therefore it is assumed that synthesis of pRNA is needed for efficient outgrowth in both *B. subtilis* and *E. coli* (Cavanagh et al., 2012).

Transcription by Eσ^A from 6S RNA templates in *B. subtilis* is sensitive to [iNTP]. For efficient synthesis of pRNA it is important that transcription initiates at a guanine. This preference of Eσ^A can explain why 6S-1 RNA serves as a good template as it has C at the +1 template strand position. In contrast, 6S-2 RNA has U at the +1 position and only a minimal amount of pRNAs is produced from 6S-2 RNA. RNA polymerase from *E. coli* is not sensitive to [iNTP] and can initiate pRNA synthesis from all four nucleotides with comparable efficiency. The nucleotide at the -1 position influences efficiency of pRNA synthesis in both *E. coli* and *B. subtilis* (Cabrera-Ostertag et al., 2013).

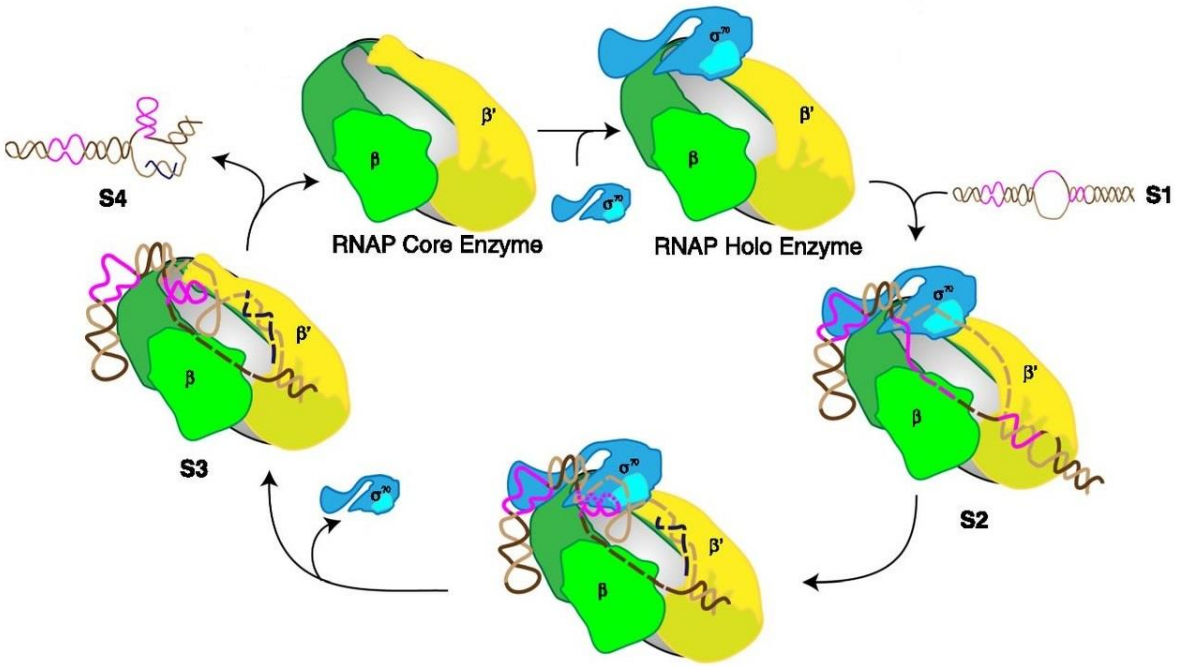

Figure 5. Model of 6S RNA release in *E. coli*. S1 – unbound 6S RNA; S2 – 6S RNA bound to RNAP holoenzyme; after synthesis of 9 nucleotides of pRNA, release hairpin can be formed between upstream helix and region -10, RNA is "scrunched" into the holoenzyme leading to state S3 – 6S RNA bound to RNAP core enzyme; S4 – 6S RNA in complex with pRNA is released from RNA polymerase. Conserved regions of 6S RNA are shown in purple; pRNA is in black; template strand is light brown and non-template strand is dark brown (Adapted from Panchapakesan and Unrau, 2012).

3.3.5 Targets of 6S RNA regulation

Promoters affected by 6S RNA have moderate to week conservation of the -35 region. The presence of the extended -10 promoter region causes promoters with moderately conserved -35 regions to be sensitive to 6S RNA downregulation. Promoters with week -35 elements are further downregulated if they have the extended -10 region (Cavanagh et al., 2008).

One of the genes directly regulated by 6S RNA is *pspF*. PspF is a transcription activator involved in cell response to high pH. In the absence of 6S RNA transcription of the *pspF* is increased. This increase in transcription gives mutant cells higher ability to survive in elevated pH compared to wild-type cells (Trotochaud and Wassarman, 2006).

6S RNA regulates transcription from both promoters of the *relA* gene, which encodes ppGpp synthase I. Although changes in expression of *relA* in 6S RNA null mutants are modest, they are sufficient to see 1.5- to 1.6-fold increase of ppGpp in early stationary phase. The level of transcription from promoters sensitive to ppGpp was also altered. In late stationary phase 6S RNA downregulates expression from some amino acids promoters by a mechanism independent of ppGpp (Cavanagh et al., 2010).

Ribosome biogenesis is inhibited in 6S RNA deficient cells as 17 ribosomal protein mRNAs and genes involved in translation were found to be downregulated in early stationary phase. Among downregulated genes were also *rpoB* and *rpoC* genes encoding RNAP subunits (Neusser et al., 2010).

3.4 Ms1 – novel type of sRNA

There are several bacterial species including mycobacteria in which 6S RNA has not been identified. Using a bioinformatic approach, a new small RNA designated Ms1 was found in *M. smegmatis* as a structural homologue of previously identified 6S RNAs from other bacteria. The predicted secondary structure of Ms1, a long double-stranded hairpin with internal bubble, is shown in **Figure 6** (Pánek et al., 2011).

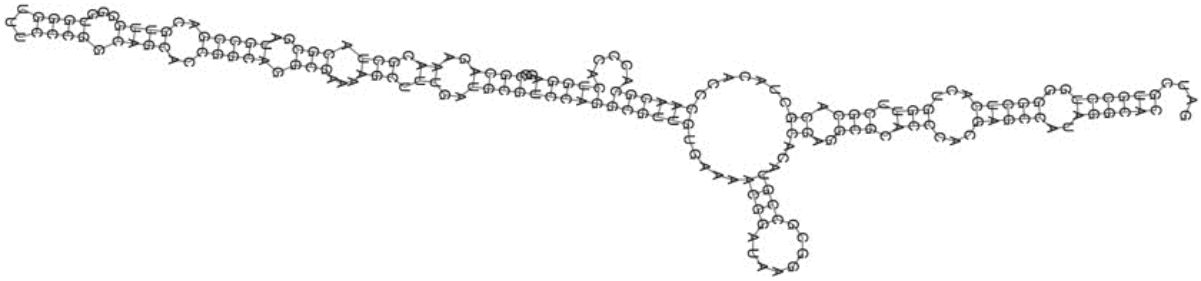

Figure 6. Predicted secondary structure of Ms1 (Pánek et al., 2011).

Ms1 is ~300nt long sRNA. Its transcription start site was mapped to adenine at position 6 242 368 in the *Mycobacterium smegmatis* genome (NCBI Reference Sequence: NC_008596.1). Upstream of this nucleotide, a putative promoter has been identified. In contrast to 6S RNA, Ms1 does not interact with RNA polymerase in complex with primary sigma factor (**Figure 7**). Ms1 interacts with the RNAP core and overexpression of the primary sigma factor in stationary phase of growth decreases the amount of Ms1 bound to RNAP. The amount of Ms1 in the cell increases ~130 times in the stationary phase of growth compared to the exponential phase. Apart from RNAP, polyribonucleotide nucleotidyltransferase and transcription termination factor Rho are other potential interaction partners of Ms1 (Hnilicová et al., 2014; Pánek et al., 2011).

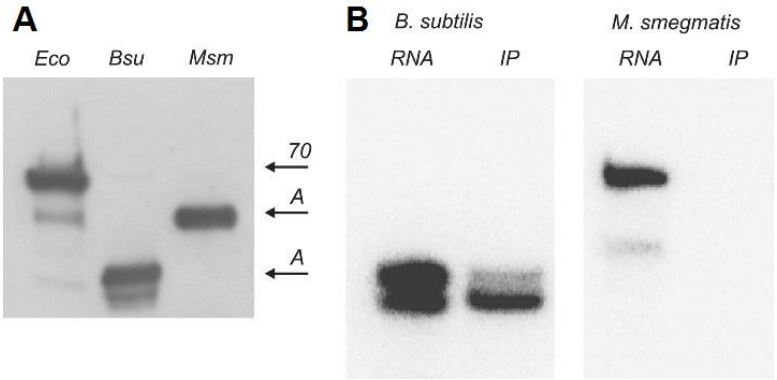

Figure 7. A) Western blot analysis of *E. coli*, *B. subtilis* and *M. smegmatis* cell lysate probed with antibody against σ^{70} from *E. coli*. This experiment was done to test whether σ^{70} antibody reacts also with sigma subunits from *B. subtilis* and *M. smegmatis*. B) Northern blot of total RNA (RNA) and the product of immunoprecipitation (IP). The *B. subtilis* blot was probed with an oligonucleotide against 6S RNA. This RNA is present in RNAP holoenzyme. The *M. smegmatis* blot was probed with an oligonucleotide against Ms1. Ms1 is present only in total RNA, not in immunoprecipitated fraction (Pánek et al., 2011).

To test whether the central bubble is important for binding to RNAP, a mutant Ms1 with deleted internal bubble was prepared. Using biotinylated RNA attached to streptavidin beads, considerably less RNAP core was pulled down with mutant Ms1 compared to wild-type (**Figure 8**). This finding indicates that the internal bubble is important for interaction between Ms1 and RNAP (Hnilicová et al., 2014).

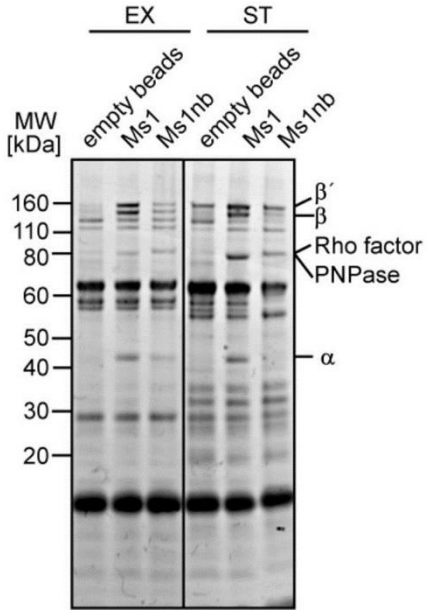

Figure 8. Protein pull-down with biotinylated Ms1 and Ms1 without internal bubble (Ms1nb). Streptavidin beads with bound Ms1/Ms1nb were incubated with *M. smegmatis* lysates from exponential or stationary phase. Pulled down proteins were separated on SDS-PAGE, stained with Coomassie blue and analysed by mass spectrometry. Considerable less RNAP interacted with Ms1nb compared to Ms1 (Hnilicová et al., 2014).

Homologues of Ms1 are present in other mycobacterial species. In *M. tuberculosis* it is MTS2823. MTS2823 is ~300 nucleotides long and has the highest expression among non-coding RNAs in *M. tuberculosis*. The level of MTS2823 further increases ten-fold in stationary phase and is comparable to levels of 16 S rRNA or 23S rRNA. Upon overexpression of this sRNA, 301 genes were downregulated (Arnvig et al., 2011). Likewise, RNA sequencing (RNA-Seq) data from ΔMs1 *M. smegmatis* strain show a huge effect on transcription (Šiková M., Krásný L., unpublished data).

Ms1 is a novel type of regulatory sRNA interacting with the RNA polymerase core. However, how the transcription of Ms1 is regulated itself is not known. In my thesis I characterized some aspects of regulation of Ms1 transcription.

4. Materials and methods

4.1 Used chemicals and enzymes

The following table contains the most frequently used chemicals and enzymes. All chemicals, enzymes, media and buffers are listed in appropriate descriptions of methods.

Table 2. List of used chemicals

10 x ligase buffer	Takara	KCl (Potassium chloride)	Penta
10x Taq PCR buffer	Sigma	MgCl2 (Magnesium chloride)	Penta
5x GC buffer	NEB	$MgSO_4*7H_2O$ (Magnesium sulfate heptahydrate)	Penta
5x GoTaq buffer	Promega	Na_2CO_3 (Sodium carbonate)	Lachema
Aqua pro injection H_2O	Braun	$Na_2HPO_4*12H_2O$ (Disodium phosphate dodecahydrate)	Lach-ner
BSA (Bovine serum albumine)	Sigma	NaCl (Sodium chloride)	Lach-ner
Buffer H	Takara	$NaH_2PO_4*2H_2O$ (Monosodium phosphate dihydrate)	Applichem
Coomassie brilliant blue R - 250	Thermo scientific	ONPG (ortho-Nitrophenyl-β-galactoside)	Sigma
DMSO (Dimethyl sulfoxide)	NEB	PBS (Phosphate-buffered saline)	Sigma
dNTP - ATP, CTP, GTP, UTP	Roche	Phusion® High-Fidelity DNA Polymerase	NEB
DTT (Dithiothreitol)	Thermo scientific	PMSF (Phenylmethylsulfonyl fluoride)	P - LAB
EDTA (Ethylenediaminetetraacetic acid)	Lachema	Protease Inhibitor Cocktail Set III, EDTA Free	Calbiochem
Ethanol 96%	Penta	ScaI restriction endonuclease	Takara
Formaldehyde	Applichem	Spectinomycin	Serva
GelRed	Bio-Rad	Streptomycin	Serva
Glycerol	Sigma	T4 DNA ligase	Takara
Glycine	Serva	Tris-HCl	Serva
GoTaq Polymerase	Promega	Tween80	Sigma
H_2O_2 (Hydrogen peroxide)	Lach-ner	β-mercaptoethanol	Serva
H_3PO_4 (Phosphoric acid)	Applichem		

4.2 Bacterial strains

Table 3. List of used *E. coli* strains

Escherichia coli			
Number	**Strain**	**Genotype**	**Source**
LK1249	1249/pSM128	DH5α pSM128	(Dussurget et al., 1999)
LK1439	1439/pSM12+PMs1 core	DH5α pSM128 PMs1(-38 +9)	This work
LK1420	1420/pSM128+PMs1 (-131+9)	DH5α pSM128 PMs1(-131 +9)	This work
LK1409	1409/pSM128+PMs1 (-231+9)	DH5α pSM128 PMs1(-231+9)	This work
LK1411	1411/pSM128+PMs1 (-331+9)	DH5α pSM128 PMs1(-331 +9)	This work
LK1421	1421/pSM128+PMs1 (-491+9)	DH5α pSM128 PMs1(-491 +9)	This work
LK1547	1547/pSM128+PMs1 mut -10	DH5α pSM128 PMs1(-491 +9) mutations: T-12G, A-11C,C-10A,A-9G	This work
LK1551	1551/pSM128+PMs1 (-491 -22)	DH5α pSM128 PMs1(-491 -22)	This work
LK1771	1771/pSM128+PMs1 (-131+9) mut TopoI RS	DH5α pSM128 PMs1(-131 +9) mutations: T-126C, T-125G, T-104C,T-105G	This work

Table 4. List of used *M. smegmatis* strains

Mycobacterium smegmatis			
Number	**Strain**	**Genotype**	**Source**
LK865	LK865	wt mc^2 155	Šárka Nezbedová Bobková
LK1442	1442/pSM128	mc^2 155 pSM128	This work
LK1443	1443/pSM128+PMs1 core	mc^2 155 pSM128 PMs1(-38 +9)	This work
LK1444	1444/pSM128+PMs1 (-131+9)	mc^2 155 pSM128 PMs1(-131 +9)	This work
LK1445	1445/pSM128+PMs1 (-231+9)	mc^2 155 pSM128 PMs1(-231 +9)	This work
LK1446	1446/pSM128+PMs1 (-331+9)	mc^2 155 pSM128 PMs1(-331 +9)	This work
LK1447	1447/pSM128+PMs1 (-491+9)	mc^2 155 pSM128 PMs1(-491 +9)	This work
LK1578	1578/pSM128+PMs1 mut -10	mc^2 155 pSM128 PMs1(-491 +9) mutations: T-12G, A-11C, C-10A,A-9G	This work
LK1590	1590/pSM128+PMs1 (-491 -22)	mc^2 155 pSM128 PMs1(-491 -22)	This work
LK1781	1781/pSM128+PMs1 (-131+9) mut TopoI RS	mc^2 155 pSM128 PMs1(-491 +9) mutations: T-126C, T-125G, T-104C,T-105G	This work

4.3 Primers

Table 5. List of used primers (5’ to 3’)

Oligonucleotide name	Sequence
Oligonucleotide sequences of Ms1 core promoter	
1320/Pms1_CU	CCCTTGCTGTGACTCGGGACACGTAGTACAAAGGAGGCACGGAAGCT
1321/Pms1_CB	AGCTTCCGTGCCTCCTTTGTACTACGTGTCCCGAGTCACAGCAAGGG
1353/Pms1_CU.A	CCCTTGCTGTGACTCGGGACACGTAGTACAAAGGAGGCACGGAAGCTA
1354/Pms1_CB.A	AGCTTCCGTGCCTCCTTTGTACTACGTGTCCCGAGTCACAGCAAGGGA
Primes for PCR amplification of promoter of Ms1	
1322/Pms1_R	AGCTTCCGTGCCTCCTTTGTAC
1323/Pms1_140F	GCCGGTTCGCGCTCTGACG
1324/Pms1_240F	CGTGACCGGTTCCCCGCGC
1325/Pms1_340F	GGTGGGGCACCCCACCGCG
1326/Pms1_500F	TGGTGGTGGAGGACGGCCG
Primers for sequencing insert in ScaI site on pSM128	
1335/pSM128-F	AATTGGGGACCCTAGAGGTC
1336/pSM128-R	ATTCGTAGAGCCTCGTTGCG
1360/pSM128-F2	CCAGATCTTTGTGGATGACC
Primers for PCR amplification of modified Ms1 promoter	
1431/Pms1_mut	AGCTTCCGTGCCTCCTTCTGCCTACGTGTCCCG
1432/Pms1_no(-10)	CCGAGTCACAGCAAGGGCC
Primers for pull-down assay	
1359/N-HelD_C-His_R	CGGGCTCGAGCGTTTCAATCGCTTCATG
1433/Pms1_240F_biotin	[Btn]CGTGACCGGTTCCCCGCGC
1434/HelD_F_biotin	[Btn]ATGAATCAGCAGGATAAGGAATGG
1280/MS_RpoC_IVT_R	GTCGTTCAGGTCGGAGGTC
1658/MS_RpoC_bio_F	[Btn]CGACGAGATCTGGAACACCT

4.4 Vectors

Plasmid pSM128 was constructed by Dussurget et al. (1999). pSM128 contains oriE and therefore is replicative in *E. coli*. After transformation to *M. smegmatis* cells this plasmid integrates at the *attB* site in the chromosome. It contains the streptomycin/spectinomycin resistance gene, KpnI and ScaI restriction sites and promoterless *lacZ* gene. Scheme of pSM128 is shown in **Figure 9**.

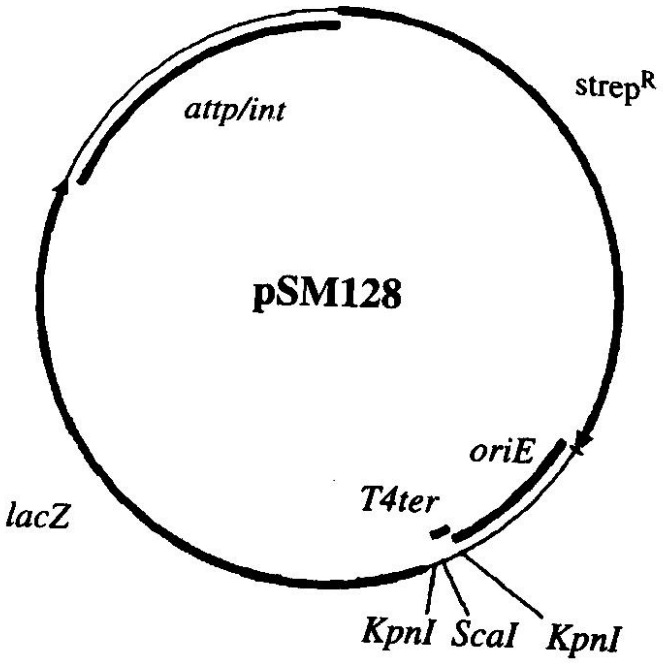

Figure 9. Scheme of pSM128 (Carroll and James, 2009).

4.5 Bacterial growth

4.5.1 *E. coli*

Equipment:

- Horizontal shaker HS250 BS1 (IKA LABORATORTECHNIK)
- Universal 320R centrifuge (Hettich)

Cultivation media:

Luria – Bertoni (LB) mediu: Dissolve 10 g of tryptone, 5 g of yeast extract and 10 g of NaCl in 1l of ddH_2O

Cultivation protocol:

- Bacteria from the glycerol stock strains (stored at -80°C) were inoculated to 7 ml (Miniprep) or 100 ml (Midiprep) of LB medium with spectinomycin (100 μg/ml) and incubated at 37°C overnight with shaking
- Cultures were centrifuged at 9000 x g, 10 minutes, 4°C

4.5.2 *M. smegmatis*

Equipment:

- Biological thermostat BT 120
- Horizontal shaker HS250 BS1 (IKA LABORATORTECHNIK)
- Microfuge 20R (Beckman Coulter)

Cultivation media:

7H9 liquid medium with glycerol: dissolve 4.7 g Difco^{tm} Middlebrook 7H9 Broth and 2 ml of 100% glycerol in 900 ml of ddH_2O

7H10 agar medium: dissolve 19 g Difco^{tm} Midlebrook 7H10 Agar and 5 ml of 100% glycerol in 900 ml of ddH_2O

Both liquid 7H9 medium and 7H10 agar were sterilized by autoclaving.

Difco^{tm} Middlebrook 7H9 Broth contains per 900 ml:

0.5 g Ammonium sulphate, 0.5 g L-Glutamic acid, 0.1 g Sodium citrate, 1.0 mg Pyridoxine, 0.5 mg Biotin, 2.5 g Disodium phosphate, 1.0 g Monopotassium phosphate, 0.04 g Ferric ammonium citrate, 0.05 g Magnesium sulphate, 0.5 mg Calcium chloride, 1.0 mg Zinc sulphate, 1.0 mg Copper sulphate

Difco^{tm} Midlebrook 7H10 Agar contains per 900 ml:

0.5 g Ammonium Sulphate, 1.5 g Monopotassium phosphate, 1.5 g Disodium phosphate, 0.4 g Sodium citrate, 25.0 mg Magnesium sulphate, 0.5 mg Calcium chloride, 1.0 mg Zinc sulphate, 1.0 mg Copper sulphate, 0.5 g L-Glutamic acid (sodium salt), 0.04 g Ferric ammonium citrate, 1.0 mg Pyridoxine hydrochloride, 0.5 mg Biotin, 250.0 µg Malachite green, 15.0 g Agar

Cultivation protocol:

- From 7H10 agar medium agar plates with appropriate antibiotic were prepared. Bacteria from glycerol stock stored in -80°C were streaked out on these plates. The agar plates were incubated at 37°C for 3 days.
- 7 ml of 7H9 with 17.5 µl Tween80 and 3.5 µl streptomycin (10 µg/ml) were added to falcon tube and inoculated with bacteria from plates; the culture was incubated at 37°C with shaking overnight
- OD_{600} of the starter culture was measured and fresh 30 ml of 7H9 with 75 µl Tween80 without antibiotic was inoculated with starter culture to OD_{600} 0.1
- Culture was cultivated at 37°C with shaking, OD_{600} was measured regularly
- At specific OD_{600} or time points samples were collected (time points are specified in description of particular experiment)
    - 1 ml of bacterial culture was collected in duplicates
    - Samples were centrifuged at 13200 x g, 10 minutes, 4°C
    - Pellets were stored at -20°C until further use

Testing of stress response protocol:

Protocol was adapted from the study of Gebhard et al. (2008) and modified. Starvation conditions were adapted from Smeulders et al. (1999).

- 110 ml of cell culture in 300 ml Erlenmeyer flasks was cultivated according to the cultivation protocol to OD_{600} 0.5, further manipulation differed based on used stress
- Heat-stress and cold-stress:
    - 25 ml of culture was pipetted into 100 ml Erlenmeyer flask and placed on shaker in thermostat set to 50°C (heat-stress) or left at room temperature (cold-stress)
- Alcohol, oxidative, hyperosmotic stresses:
    - 25 ml of culture was pipetted into 100 ml Erlenmeyer flask and:
        - 2.5 ml of 5 M NaCl (0.5 M final concentration) was added to the medium – hyperosmotic stress

        - 19 µl of 9.8 M H_2O_2 (7.5 mM final concentration) was added to the medium – oxidative stress
        - 1.315 ml of 96% EtOH (5% final concentration) was added to the medium – alcohol stress
- Hypo osmotic, alkaline, acid stresses and starvation:
    - 25 ml of culture was filtered using MF^{TM} – Membrane Filters 0.45 µm HA (Millipore) cells were resuspended in 25 ml of preheated
        - 25 ml of distilled water – hypo osmotic stress
        - 25 ml of 7H9 medium with pH adjusted to 4
        - 25 ml of 7H9 medium with pH adjusted to 9
        - 25 ml of PBS
- Apart from heat- and cold- stress, all cultures were cultivated at 37°C
- Samples were collected in duplicates at 2 hours and 4 hours after stress, respectively

4.6 Preparation of glycerol stock cultures

Bacteria can be stored for long periods of time (years) in glycerol at -80°C.

Protocol:

- Liquid bacterial culture was grown from single colony overnight
- 850 µl of culture was added to 150 µl of glycerol and mixed thoroughly by vortexing
- The mixture was stored at -80°C

4.7 Polymerase chain reaction

Polymerase chain reaction (PCR) was used to amplify the promoter sequence of Ms1. As the templates for the reaction, genomic DNA from *Mycobacterium smegmatis* mc^2 155 or plasmid DNA with particular insert were used. Used primers are listed in **Table 5** in chapter 4.3. The same reverse primer was used for all reactions. Composition of reaction solution and a typical PCR programme is shown below. Annealing temperature was adjusted according to the calculated t_m of primers. Reaction components and steps of PCR programme are listed in **Table 6** and **Table 7**, respectively.

Equipment:

- MJ Research PTC-200 Thermal Cycler

Table 6. PCR reaction components:

amount	chemical	final concentration
10 µl	5x GC buffer (NEB)	1x
1 µl	dNTPs 10nM	200 µM
1.5 µl	DMSO (NEB)	3%
1 µl	template	5 ng
0.5 µl	Forward primer	1 µM
0.5 µl	Reverse primer	1 µM
0.5 µl	Phusion® High-Fidelity DNA Polymerase (NEB)	1.0 units /50 µl reaction
35 µl	*Aqua pro injection* H_2O (Braun)	

Table 7. Steps of PCR programme:

step	temperature	time	number of repeats
initial denaturation	98°C	30 sec	1x
denaturation	98°C	10 sec	30x
annealing	50°C	30 sec	
elongation	72°C	2 min	
final elongation	72°C	6 min	1x
	4°C	hold	

4.8 Annealing

The shortest Ms1 promoter fragment containing only the core promoter sequence was obtained as a pair of complementary oligonucleotides. The double strand was prepared by an annealing reaction. 5 µl of both oligonucleotides were mixed with 5 µl of 10x Taq PCR buffer (Sigma) including $MgCl_2$ and 35 µl of ddH_2O and placed into the thermal cycler (MJ Research PTC-200). The program for annealing started at 95°C and every minute decreased by 1°C until it reached 45°C.

4.9 Horizontal gel electrophoresis

This method was used to separate DNA of different lengths (after PCR) or state of superhelicity (after restriction analyses). DNA has a negative charge and therefore moves towards anode when placed in the electrical field.

Equipment:

- Owl™ EasyCast™ B1A Mini Gel Electrophoresis Systems
- PowerPack 3000 (Bio-Rad)
- InGenius (Syngene)

Protocol:

- 0.5 g of agarose was dissolved in 50 ml 1x TAE buffer (2M Tris-acetate, 50 mM EDTA, pH 8.0), agarose was melted by heating in microwave
- Agarose was cooled down to ~50°C and 5 µl of GelRed (Bio-Rad) was added
- Agarose was mixed by swirling and poured into gel casting tray and comb was inserted into agarose
- After the agarose turned solid (~30 min) the electrophoresis chamber was filled with 1x TAE so the gel and electrodes were covered and the comb was carefully removed
- Samples with added sample buffer were loaded into wells
- Power supply was turned on to 80 V (~5V/cm), separation ran for approximately 1 hour
- Gel with separated DNA was analysed using InGenius (Syngene) gel documentation system

4.10 Isolation from agarose gel

Although PCR primers are designed to be specific for DNA region of interest there might be minor non-specific interaction of primers with DNA. This results in more products from PCR reaction. After horizontal gel electrophoresis, the band of correct length was isolated from the gel.

For the extraction QIAquick Gel Extraction Kit was used (Qiagene)

Equipment:

- TS-100C Thermoshaker with cooling (P-lab)
- Universal 320R centrifuge (Hettich)

Protocol:

- Band of interest was cut out from the agarose gel using a scalpel and placed in colourless clean microcentrifugation tube
- The gel slice was weighted
- 3 volumes of Buffer QG were added to the gel
- The tube was incubated at 50°C for 10 minutes with occasional vortexing
- After the gel slice dissolved, 1 volume of isopropanol was added and the sample was mixed
- A QIAquick spin column was placed into 2 ml collection tube

- The sample was pipetted to QIAquick column and centrifuged at 17900 x g, 1 min, room temperature
- Flow-through was discarded, 0.5 ml of Buffer QG was added to QIAquick column and centrifuged at 17900 x g, 1 min, room temperature
- Flow-through was discarded, 0.75 ml of Buffer PE was added to QIAquick column to was the sample and centrifuged at 17900 x g, 1 min, room temperature
- Flow-through was discarded and the QIAquick column was centrifuged at 17900 x g, 1 min, room temperature
- QIAquick column was placed into clean 1.5 ml microcentrifuge tube and 30 µl of ddH_2O was pipetted to the centre of the QIAquick membrane
- The column with water was incubated for 1 minute and then centrifuged at 17900 x g, 1 min, room temperature
- Eluted DNA was stored in -20°C until further use

4.11 TA cloning

For cloning DNA inserts into pSM128 the ScaI restriction site was used. The ScaI restriction enzyme produces blunt ends. For easier and more effective creation of desired constructs, TA cloning was used. Taq polymerase is capable of adding a single non-templated nucleotide to the 3' end of double stranded DNA. One thymine was added to 3' ends of linearized plasmid and one adenine to 3' ends of DNA inserts. This should prevent closing of the cleaved plasmid without incorporating a DNA insert.

4.11.1 Production of linearized pSM128 vector

To be able to clone a DNA insert into a plasmid this plasmid needs to be linearized. Plasmid DNA was isolated from *E. coli* strain 1241/pSM128 by Midiprep isolation. Afterwards, the plasmid was cleaved with ScaI at 37°C for 1 hour. The reaction product was separated on the agarose gel and the linear form of plasmid was isolated. Reaction components are listed in **Table 8**.

Equipment:

- MJ Research PTC-200 Thermal Cycler

Table 8. Reaction components:

amount	chemical
6 µl	sample
2 µl	Buffer H (Takara)
0.5 µl	Sca I (Takara)
11.5 µl	ddH_2O

4.11.2 Production of T- A- 3' overhangs

A single adenine was added to the 3' ends of inserts and a single thymine was added to the 3' ends of linearized plasmid. Reaction composition is specified below. Samples were placed at thermal cycler (MJ Research PTC-200) for 30 minutes at 72°C. Reaction components are listed in **Table 9.**

Table 9. Reaction components:

amount	chemical
4 µl	5x GoTaq buffer (Promega)
0.5 µl	GoTaq Polymerase (Promega)
1 µl	dATP/dTTP 10 nM
14.5 µl	DNA solution

4.11.3 Ligation

DNA inserts with 3' A overhangs were ligated into linear pSM128 plasmid with 3' T overhangs. As a control, one ligation reaction contained only plasmid DNA and no DNA insert. Ligation reaction was carried at 16°C overnight at thermal cycler (MJ Research PTC-200). Reaction components are listed in **Table 10.**

Table 10. Ligation reaction components:

amount	chemical
1 µl	plasmid
2 µl	DNA insert
1 µl	10 x ligase buffer (Takara)
0.9 µl	T4 DNA ligase (Takara)
5.1 µl	ddH_2O

4.12 Transformation of *E. coli* cells

Plasmid DNA was integrated into *E. coli* cells by transformation mediated by heat-shock. It is important that all steps are carried out on ice under sterile conditions.

Equipment

- Microfuge 20R (Beckman Coulter)
- Biological thermostat BT 120
- Horizontal shaker HS250 BS1 (IKA LABORATORTECHNIK)

Protocol:

- Competent cells (stored at -80°C) were left to thaw on ice for ~15 minutes
- Whole ligation mixture was pipetted into cooled microcentrifugation tubes and 100 µl of competent cells were added, one tube only with competent cells served as a control
- Tubes were gently mixed and incubated on ice for 30 minutes
- Heat-shock was performed by placing tubes at 42°C for 90 seconds
- Tubes were incubated on ice for additional 5 minutes
- 1 ml of LB medium without antibiotics was added to all tubes
- Tubes were incubated at 37°C with shaking
- After incubation tubes were shortly centrifuged, supernatants were discarded
- Cells were resuspended in the rest of the medium and plated on agar plates with LB medium with spectinomycin (100 µg/ml)
- Plates were incubated at 37°C overnight

4.13 Isolation of plasmid DNA

pSM128 is a replicative plasmid in *E.coli* therefore it is possible to isolate it from the bacterium using a method specific for isolation of plasmid DNA. For plasmid isolation two methods were used. To isolate small amounts of plasmid to be able to check whether the correct insert was incorporated into the plasmid, QIAprep Spin Miniprep Kit (Qiagen) was used. To isolate enough plasmid for *M. smegmatis* transformation, Wizard® *Plus* Midipreps DNA Purification System was used.

4.13.1 Miniprep isolation

Equipment:

- Universal 320R centrifuge (Hettich)
- Horizontal shaker HS250 BS1 (IKA LABORATORTECHNIK)

Protocol:

- A single colony from agar plate was inoculated into 7 ml of LB medium with 7 µl of spectinomycin (100 µg/ml) in falcon tubes
- Tubes were incubated at 37°C overnight with shaking
- Tubes were centrifuged at 9000 x g, 10 minutes, 4°C
- Supernatant was discarded and bacterial pellets resuspended in 250 µl of buffer P1 and transferred to a microcentrifuge tube
- For lysis, 250 µl of buffer P2 was added and tubes were mixed by inverting 6 times. Lysis cannot exceed 5 minutes
- For neutralization, 350 µl of buffer N3 was added and tubes were mixed immediately by inverting 6 times
- Tubes were centrifuged at 17900 x g, 10 minutes, room temperature
- Supernatant was pipetted to the QIAprep spin column
- Columns were centrifuged for 1 minute at 17900 x g, room temperature, flow-through was discarded
- To wash, 0.5 ml of buffer PB was added to columns
- Columns were centrifuged for 1 minute at 17900 x g, room temperature, flow-through was discarded
- To wash, 0.75 ml of buffer PE was added to columns
- Columns were centrifuged for 1 minute at 17900 x g, room temperature, flow-through was discarded
- Columns were centrifuged for 1 minute at 17900 x g, room temperature
- QIAprep spin column was placed into clean 1.5 ml microcentrifuge tube and 30 µl of ddH_2O was pipetted to the centre of the QIAprep spin column
- The column with water was incubated for 1 minute and then centrifuged at 17900 x g, 1 min, room temperature
- Eluted DNA was stored in -20°C until further use

4.13.2 Midiprep isolation

Equipment:

- AVANTI J-26XPI (Beckman) centrifuge, JA – 25.50 rotor
- Universal 320R centrifuge (Hettich)
- Vacuum pump

Protocol:

- Starter culture was prepared by inoculating stock culture into 100 ml of LB medium with spectinomycin (100 μg/ml) and grown overnight with shaking (210 rpm)
- Cells were pelleted by centrifugation (320 R) at 9000 x g, 10 minutes, 4°C
- Cells were resuspended in 3ml Cell Resuspension Solution
- To lyse, 3 ml of Cell Lysis Solution was added and tubes were mixed by inverting 4 times. Lysis cannot continue for more than 5 minutes
- 3 ml of Neutralization Solution was added to stop the lysis and samples were mixed by inverting tube 4 times
- Tubes were centrifuged (AVANTI J-26XPI) at 27000 x g, 15 minutes, 4°C
- Supernatant was carefully decanted into a clean falcon tube
- Midicolumn was attached to the vacuum aspirator
- Supernatant was mixed with 7.5 ml of resin (important to thoroughly mix and resuspend resin before use)
- Resin/DNA mixture was transferred to Midicolumn and vacuum was applied
- When the sample passed through the column, vacuum was disconnected
- Sample was washed with 2 x 15 ml of Column Wash Solution under vacuum
- Resin was dried by vacuum for 30 seconds after the solution passed through the column
- Midicolumn was cut from Reservoir and Midicolumn was placed into a microcentrifuge tube and centrifuged (320R) at 10000 x g, 2 minutes, room temperature
- To elute, 300 μl of ddH_2O preheated to 70°C was added to the Midicolumn, column was incubated for 1 minute and centrifuged (320R) at 10000 x g for 20 seconds
- To remove the residual resin, the microcentrifuge tube was centrifuged (320R) at 10000 x g for 5 minutes and the cleared DNA solution was pipetted into a new 1.5 ml microcentrifuge tube

4.14 Restriction analysis

To check whether DNA insert was present in the plasmid, restriction analysis was carried out. From each isolated plasmid two samples were prepared. One of the samples was subject of cleavage reaction. Cleavage reaction had same conditions and reaction composition as production of linearized vector (section 4.10.1). Cleaved and non-cleaved samples were run on agarose gel. When there was an insert at the ScaI site, no cleavage was visible. A representative result from restriction analysis is shown in **Figure 10**.

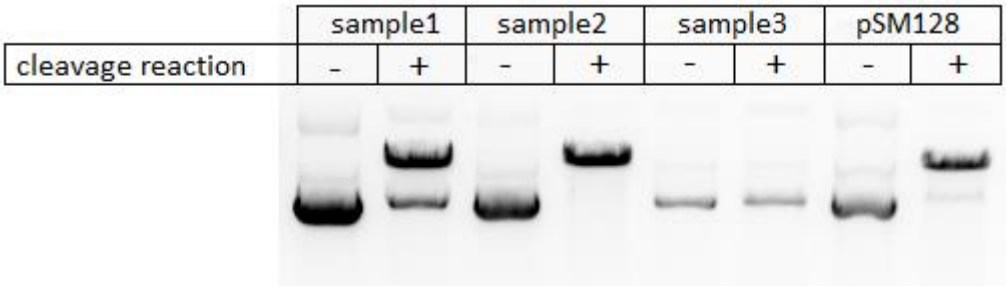

Figure 10. Photograph of agarose gel. Samples with no DNA insert were susceptible to restriction enzyme (samples 1 and 2). Samples with DNA fragment inserted to ScaI site were resistant to ScaI cleavage (sample 3). pSM128 served as a marker.

4.15 DNA sequencing

When TA cloning is used, DNA fragment can be inserted in both orientations. To check whether DNA insert is in correct orientation and if there are no mistakes in the sequence, plasmid DNA was sequenced using 1336/pSM128-R and 1360/pSM128-F2 primers. Sequencing reactions were done at Centre for DNA Sequencing at Institute of Microbiology of the Czech Academy of Science.

4.16 Transformation of *M. smegmatis* cells

After the plasmid construct sequence had been verified a glycerine stock strain was prepared from this *E. coli* colony. Plasmid from stock strain was subsequently isolated by Midiprep and transformed into *M. smegmatis*. *M. smegmatis* cells were transformed by electroporation. All steps were carried out on ice under sterile conditions.

Equipment:

- Universal 320R centrifuge (Hettich)
- Gene Pulser XCell (Bio-Rad)
- Horizontal shaker HS250 BS1 (IKA LABORATORTECHNIK)

Protocol:

- Competent cells (stored at -80°C) strain mc^2 155 were left to thaw on ice for ~15 minutes
- Cells were centrifuged at 15500 x g, 3 minutes, 4°C
- Supernatant was removed by pipetting and cells were resuspended in 200 µl of ice-cold sterile 10% glycerol
- 5 µl (0.5 – 5 ng) of plasmid DNA was added
- Cells were transferred into 2 mm electroporation Gene Pulser® Cuvetes (Bio-Rad) and let stand on ice for 10 minutes
- Electroporation – the desired program was selected on Gene Pulser XCell (2.5 kV, 25 µF, 1000 Ω) cuvettes were placed in a ShockPod cuvette chamber and by pressing pulse button electricity was applied
- Cuvettes were left on ice for 10 minutes
- Cells were pipetted into falcon tubes with 5 ml 7H9 medium with 12.5 µl Tween80 and incubated at 37°C with shaking for 3 hours
- Tubes were centrifuged at 9000xg, 10 minutes, room temperature
- Supernatant was discarded
- Cells were plated on 7H10 plates with streptomycin (20 µg/ml) and incubated at 37°C for 3-5 days
- A single colony was transferred onto a new 7H10 plate and a glycerine stock strain was prepared

4.17 Work with proteins

4.17.1 Measuring protein concentration

Protein concentration was measured using the Bradford protein assay (Bradford, 1976).

Equipment:

- UV – 1601PC UV – VISIBLE – UV-vis spectrophotometer (Shimadzu)

Protocol:

- Standard solutions of bovine serum albumin (BSA) of known concentration (0.4, 1, 2, 4, 8 and 16 µg/ml) were prepared in a total volume of 40 µl
- Samples were prepared from bacterial lysate after sonication – from exponential phase 40 µl, from stationary phase 20 µl of sample + 20 µl of water
- As a blank, 40 µl of water was used

- To all standards, samples and blank, 960 µl of Bradford solution (100 mg Coomassie Brilliant Blue G – 250 diluted in 50 ml ethanol; 8.5% H_3PO_4; H_2O fill up to 1 l) was added, tubes were mixed by vortexing and incubated at room temperature for 5 minutes
- Absorbance of samples was measured at 595 nm on a spectrophotometer
- Calibration curve was constructed using equation:

$$OD = slope * c[\mu g/\mu l]$$

- Concentration of proteins in samples was calculated using equation:

$$c[\mu g/\mu l] = \frac{OD_{595}/V[\mu l]}{slope}$$

V – sample volume, c – protein concentration

4.17.2 β-galactosidase assay

For every promoter construct, β-galactosidase activity was determined. Amount of β-galactosidase corresponds to the promoter strength. This enzyme can cleave lactose and its analogues. Experimentally used analogues usually produce a coloured product which can be detected by colorimetry and promoter activity can be calculated.

Equipment:

- Universal 320R centrifuge (Hettich)
- UP200S ultrasonic device (Hielscher)
- UV – 1601PC UV – VISIBLE – UV-vis spectrophotometer (Shimadzu)

Solutions:

- Z buffer - 0.06 M $Na_2HPO_4*12H_2O$, 0.04 M $NaH_2PO_4*2H_2O$, 0.01 M KCl, 0.001 M $MgSO_4*7H_2O$, pH 7.0, store at 4°C up to one year
- Z buffer 2 – 2.7 µl of β-mercaptoethanol/1 ml of Z buffer
- 1 M Na_2CO_3
- ONPG solution – 4 mg/ml ONPG in Z-buffer, make fresh for every assay

Protocol:

- Bacterial pellets were washed with 500 µl of Z buffer 2
- Cells were centrifuged at 17900 x g, 10 minutes, 4°C
- Supernatant was discarded and cell were resuspend in 500 µl of Z buffer 2
- Cells were sonicated on ice – sonicate for 20 seconds, 1 minute pause and repeat 3 times for each sample

- Cells were centrifuged 17900 x g, 10 minutes, 4°C
- 100 µl of sonicate and 900 µl of Z buffer were added into 2 ml microcentrifuge tubes
- Tubes were incubated for 5 minutes at 30°C
- 200 µl of ONPG solution was added to the tubes – write down time of addition
- After the sample turned yellow (or after 20 minutes) the reaction was stopped with 500 µl of Na_2CO_3 - write down time of reaction
- OD_{420} and OD_{550} were measured
- Ass the blank Z buffer 2 with ONPG and Na_2CO_3 was used

The activity of β-galactosidase was calculated using the following equation:

$$activity = \frac{1000 * (OD_{420} - 1.75 * OD_{550})}{V[ml] * T\ [min] * c\ [mg/ml]}$$

V – sample volume, T – time of reaction, c – protein concentration

The activity was measured in arbitrary units.

4.17.3 Protein pull-down assay

To identify protein interaction partners within the Ms1 promoter region, a pull-down assay using a biotinylated DNA fragment 240 nucleotides long was used. Method adapted from (Hnilicová et al., 2014).

Method principle:

Biotin was attached to the 5' end of DNA fragment of interest. DNA was incubated with streptavidin beads, so that biotin could form stable interaction with streptavidin. Cell lysate was added to the beads. Proteins interacting with DNA stayed bound to the DNA-beads complex, other proteins were washed away. Proteins interacting with DNA were released by boiling at 95°C and analysed on SDS-PAGE.

Equipment:

- Multi Bio RS-24, Programmable rotator (Biosan)
- DYNAL® magnetic rack (Invitrogen)

Solutions:

- Lysis buffer A – 20 mM Tris-HCl pH 7.9, 150 mM KCl, 1 mM $MgCl_2$
- Lysis buffer B – 20 mM Tris-HCl pH 7.9, 150 mM KCl, 1 mM $MgCl_2$, 1 mM DTT, 0.5 mM PMSF, 5 µl/ml Protease Inhibitor Cocktail Set III, EDTA Free (Calbiochem)

Every wash step in the protocol was done as follows:

- Appropriate volume of Lysis buffer was added to the beads in a microcentrifuge tube
- Tubes were inserted into magnetic rack and let stand for 1-3 minutes until beads attached to the site of the tubes
- The liquid was pipetted out and beads were resuspended in solution as specified in the protocol

Protocol:

- Preparation of cell lysate :
    - 25 ml of *M. smegmatis* cells were collected in stationary phase (OD_{600} 2.5) as described in Bacterial growth section (4.5.2)
    - Cells were resuspended in 800 µl of Lysis buffer B
    - Cells were sonicated on ice – 15 x 10 seconds, after each sonication round the cells were left on ice for 1 minute
    - After sonication, the cells were centrifuged at 17900 x g, 4°C, 10 minutes
    - Supernatant was pipetted into clean tubes
- DNA binding to streptavidin beads:
    - 200 µl of Streptavidin−Iron Oxide Particles from *Streptomyces avidinii* (Sigma-Aldrich) were pipetted into microcentrifuge tubes
    - Tubes were inserted into a magnetic rack, after the beads had attached to the site of the tube, the liquid was pipetted out and the beads were washed with 1 ml of Lysis buffer B
    - Beads were resuspended in 200 µl of Lysis buffer B with 2.5 µg of biotinylated DNA
    - DNA was allowed to bind to beads for 30 minutes at room temperature, tubes were placed in rotator, 10 orbital rpm
    - Beads with bound DNA were washed with 1 ml of Lysis buffer A
- Pull-down assay
    - Beads with bound DNA were resuspended in 250 µl of cell lysate
    - Beads were incubated with cell lysate for 1 hour at 4°C (in cold room) with shaking on rotator
    - Beads were washed 4 times with Lysis buffer B, after second washing mirocentrifuge tube was changed

- Beads were resuspended in 10 µl of Lysis buffer B and either stored in -20°C until further use or ran on SDS PAGE and stained with Coomassie brilliant blue (CBB)

Method optimization:

Several rounds of pull-down assay were carried out with several modifications to optimize the protocol. These modifications are discussed below:

1. The binding temperature
   - Originally, the incubation of beads with bound DNA with lysate was carried out at 4°C; incubation temperature was increased to room temperature. It resulted in no significant change of binding and further experiments were carried out again at 4°C to reduce protein degradation during incubation
2. Volume of cell culture, streptavidin beads and biotinylated DNA
   - The volume of used cell culture was increased to 300 ml, the amount of streptavidin beads to 300 µl and the amount of biotinylated DNA to 4 µg, at the final step beads were resuspended in 50 µl of Lysis buffer B instead of 10 µl
   - The volume of the used cell culture was increased to 500 ml, the amount of streptavidin beads to 400 µl and the amount of biotinylated DNA to 5 µg, at the final step beads were resuspended in 30 µl of Lysis buffer B
3. Gel staining
   - Protein gels were stained with silver instead of CBB to visualise smaller amounts of proteins
4. Preclearing
   - Preclearing step was added to the protocol:
   - 300 µl of streptavidin beads was washed with 1 ml of Lysis buffer A
   - Beads were resuspended in 1.2 ml of cell lysate and incubated for 1 hour at 4°C with shaking on rotator
   - Cleared lysate was used for pull-down assay
   - After pull-down assay the beads were resuspended in 50 µl of Lysis buffer B and ran on SDS-PAGE to see efficiency of preclearing step
5. Pull-down assay also under DNA-protein crosslinking conditions
   - Pull-down assay under crosslinking conditions protocol:

- Beads with bound DNA were resuspended in 1.375 ml of cell lysate and 1% formaldehyde
- Beads were incubated with cell lysate for 1 hour at 4°C (in cold room) with shaking on rotator
- Crosslinking was stopped by addition of glycine (final concentration 125 mM), incubation with glycine continued for additional 5 minutes
- Beads were washed 4 times with Lysis buffer B, after second washing mirocentrifuge tube was changed
- Beads were resuspended in 30 µl of Elution buffer (1% SDS, 50 mM Tris-HCl pH 8, 10 mM EDTA, 0.5 mM PMSF, 5 µl/ml Protease Inhibitor Cocktail Set III, EDTA Free (Calbiochem))
- Decrosslink – NaCl of 200 mM final concentration was added to the sample and was incubated at 65°C for 6 hours to overnight
- Samples were either stored in -20°C until further use or ran on SDS-PAGE and stained with silver or CBB

4.17.4 SDS-PAGE

Electrophoretic analysis of proteins was done by SDS-PAGE. Precast NuPAGE® Novex 4-12% Bis-Tris Gels (Thermo Fisher) were used.

Equipment:

- XCell SureLock™ Mini-Cell Electrophoresis System (Thermo Fisher)
- PowerPack 3000 (Bio-Rad)

Solutions:

- NuPAGE® LDS Sample Buffer (4x)
- NuPAGE® MES SDS Running Buffer (20x)

Protocol:

- NuPAGE® Novex 4-12% Bis-Tris Gel was placed into XCell SureLock™ Mini-Cell
- Buffer chambers were filled with 1x running buffer
- Sample was mixed with appropriate volume of sample buffer and boiled at 95°C for 5 minutes
- Samples and marker were loaded into wells
- Electrophoresis was ran at 200 V for 35 minutes

4.17.5 SimplyBlue™ SafeStain staining

Equipment:

- MR-1 Mini Rocker Shaker (Biosan)

Protocol:

- The gel was placed into a clean container with 100 ml of ultrapure water and microwaved for 1 minute on High
- The gel was placed on a shaker for 90 seconds
- Water was removed and steps 1 and 2 were repeated twice
- 20 ml of SimplyBlue™ SafeStain were added to the gel and microwaved for 45 seconds
- The gel was placed on a shaker for 5 minutes
- The dye was discarded and gel was washed with 100 ml of ultrapure water for 10 minutes
- The washing step was repeated once or twice

4.17.6 Silver staining

Some gels were stained with silver using ProteoSilver™ Silver Stain Kit (Sigma-Aldrich)

Solutions:

- Fixing solution – 40 ml of ultrapure water, 10 ml of acetic acid and 50 ml of ethanol
- 30% ethanol – 70 ml of ultrapure water and 30 ml of ethanol
- Sensitizer – 99 ml of ultrapure water and 1 ml of ProteoSilver Sensitizer (from the kit), stable for max. 2 hours
- Silver solution – 99 ml of ultrapure water and 1 ml of ProteoSilver Silver Solution (from the kit), stable for max. 2 hours
- Developer – 95 ml of ultrapure water, 5 ml of ProteoSilver Developer 1 (from the kit) and 0.1 ml of ProteoSilver Developer 2 (from the kit), stable for max. 20 minutes

Protocol:

- The gel was placed into a clean container with 100 ml of Fixing solution and left on shaker overnight
- Fixing solution was discarded
- 30% ethanol was added to the gel and gel was placed on a shaker for 10 minutes
- Fixing solution was discarded

- The gel was washed with 100 ml of ultrapure water and was placed on a shaker for 10 minutes
- Water was discarded
- Sensitizer solution was added to the gel and gel was placed on a shaker for 10 minutes
- The gel was washed twice with 100 ml of ultrapure water and was placed on a shaker for 10 minutes, water was discarded
- Silver solution was added to the gel and gel was placed on a shaker for 10 minutes
- Silver solution was discarded
- The gel was washed with 100 ml of ultrapure water and was placed on a shaker for 1.5 minutes
- Developer solution was added to the gel and gel was placed on a shaker for 3-5 minutes – until band were clearly visible
- Reaction was stopped by the addition of 5 ml of the ProteoSilver Stop solution (from the kit) and incubated for 5 minutes
- The gel was washed with 200 ml of ultrapure water and was placed on a shaker for 15 minutes
- Water was discarded and the gel was stored in fresh ultrapure water

4.17.7 Mass spectrometry

Bands present in sample with Ms1 promoter fragment and absent in control samples were sent to analysis. Proteins were identified by mass spectrometry done in the Laboratory of Molecular Structure Characterization at Institute of Microbiology of the Czech Academy of Science.

Protein identification was done by peptide mass mapping approach using MALDI-TOF mass spectrometry. Silver-stained protein spots were excised from the gel, destained, washed and then digested with trypsin overnight in a cleavage buffer containing 25 mM 4-ethylmorpholine acetate and 5% acetonitrile. MALDI mass spectra were measured in the mass range of 700-5000 Da on an Ultraflex III instrument (Bruker Daltonics, Bremen, Germany) and calibrated internally using the monoisotopic $[M+H]^+$ ions of autoproteolytic fragments of trypsin (842.5 and 2211.1 Da).

5. Results

5.1 Preparation of *M. smegmatis* strains with Ms1 promoter inserts

A putative Ms1 promoter was predicted upstream of the Ms1 +1 transcription start site (Hnilicová et al., 2014). Primers specific for this region of DNA were designed (see **Table 5** in Materials and Methods) and a series of Ms1 promoter fragments gradually increasing in length was prepared. These fragments were cloned into the pSM128 ScaI site preceding the *lacZ* reporter gene, using competent *E. coli* cells as a tool. After the sequence of the construct was verified, plasmids were transformed into *M. smegmatis* cells (LK 865 strain). **Figure 11** shows a schematic view of used promoter fragments (A) and a gel with PCR amplified promoter fragments and linearized pSM128 plasmid (B).

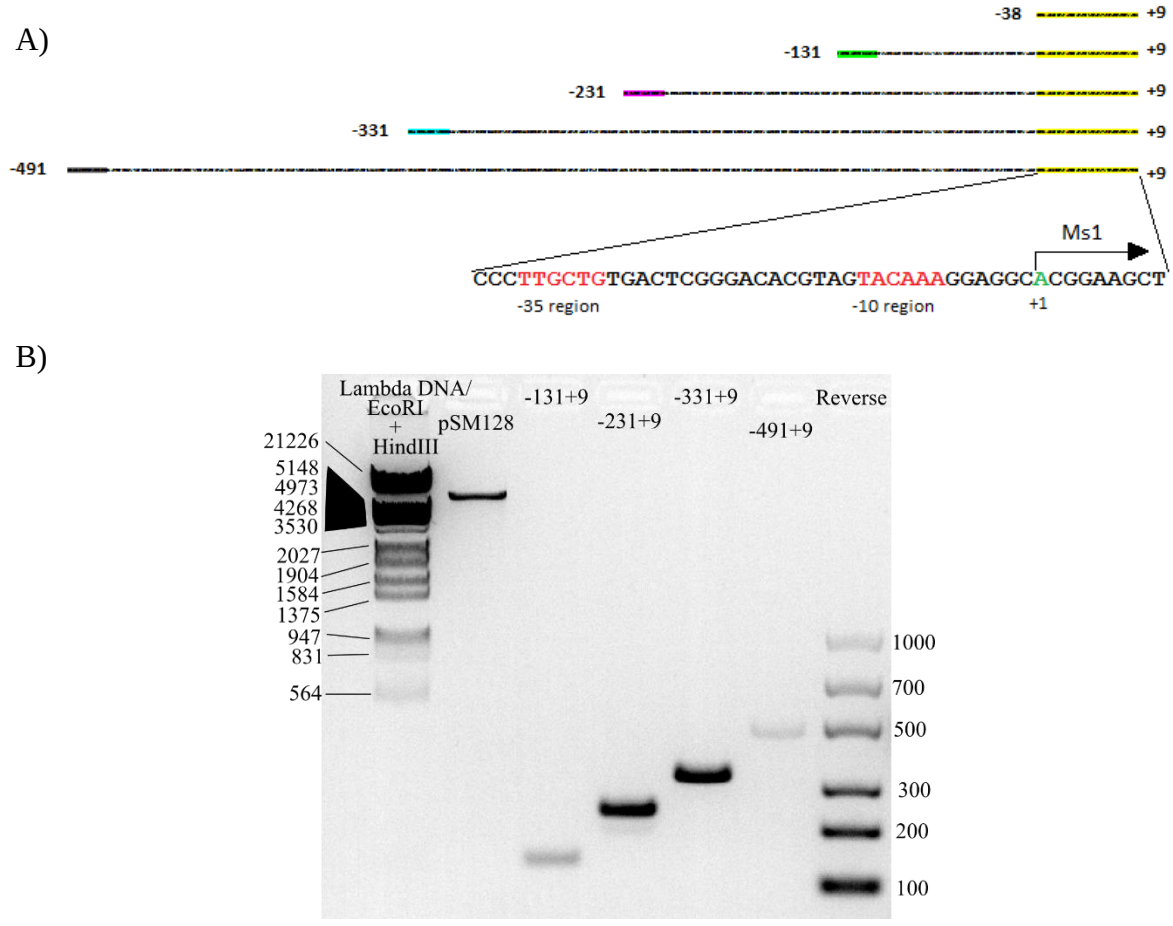

Figure 11. A) Ms1 promoter fragments for testing effect of the promoter length on the level of transcription. Each fragment contain core promoter region from -38 to +9 nucleotide. B) Agarose gel with linearized pSM128 plasmid on the left side and promoter fragments -131 +9, -231 +9, -331 +9 and -491 +9 on the right side. Used markers: Lambda DNA/EcoRI+HindIII (ThermoFisher), MassRuler™ Express LR Reverse DNA Ladder (ThermoFisher). The molecular weights of the markers fragments are indicated next to them.

5.2 Ms1 transcription in different growth phases and effect of promoter length on transcription

Strains LK1442-LK1447 were cultured until late stationary phase as described in Materials and Methods. Samples were collected in duplicates in exponential phase (OD_{600} 0.5, ~6 hours of cultivation), early stationary phase (OD_{600} 2-3, 24 hours of cultivation) and late stationary phase (OD_{600} ~2, 48 hours of cultivation). β-galactosidase activity was assessed at each time point. For each strain, three biological replicates were used. Results are shown in **Figure 12**. The promoterless construct LK1442 is not shown as it was used only as a control and no significant activity was observed. Promoter activity had an increasing tendency during growth. Changes in expression between exponential and stationary phases were similar for all used constructs. A more significant increase in activity was visible with the increased length of the promoter fragment. Between strains LK1443 (promoter region -38 +9) and LK1445 (promoter region -231 +9), promoter activity increased ~100 times. Further extension of the promoter region had only minor effects on promoter activity. This suggested there might either be a stronger promoter present in the region between -38 and -231 or the transcription from the predicted promoter was affected by a transcription activator binding within this region. Both these hypotheses were subsequently tested.

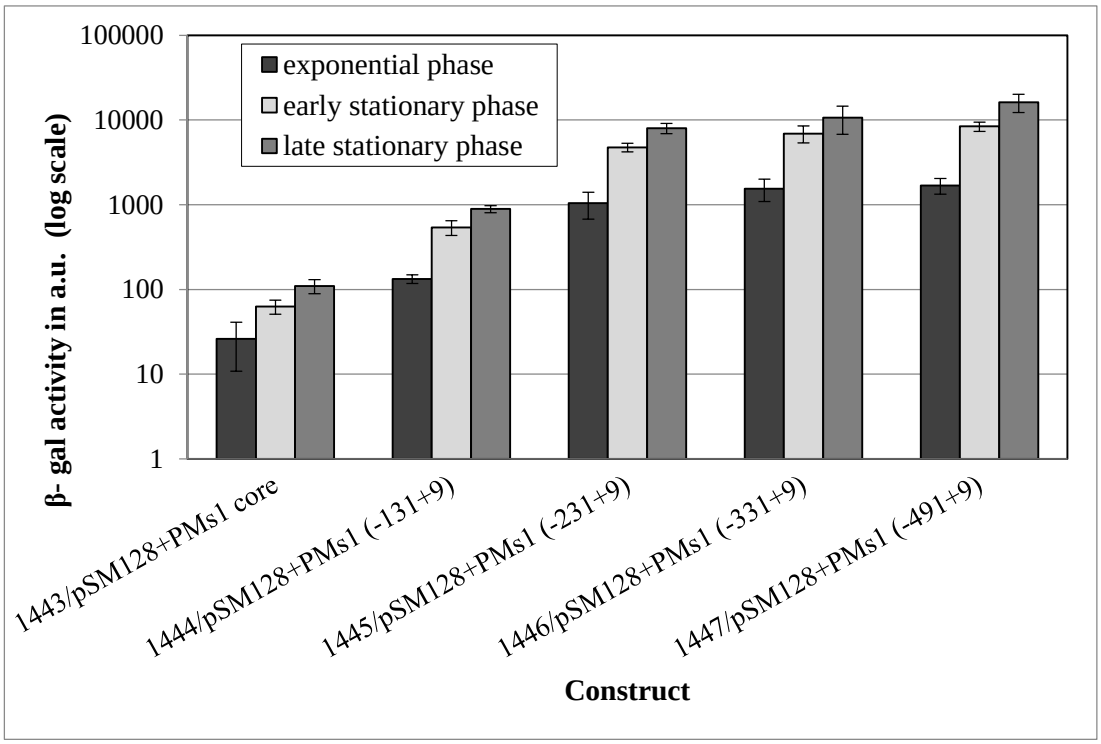

Figure 12. β-galactosidase activity of *M. smegmatis* strains LK1443-LK1447 in exponential, early and late stationary phase measured in arbitrary units. Graph shows mean values calculated from three biological replicates with samples collected in duplicates. Error bars represent standard deviation.

5.3 Identification of the Ms1 promoter

To test whether Ms1 is transcribed from more than one promoter, modified promoter fragments bearing mutations in the predicted promoter were prepared (**Figure 13**). These fragments were cloned into pSM128 and transformed into *M. smegmatis*. The two resulting *M. smegmatis* strains contained a mutated -10 consensus sequence (LK1578) and a fragment without the -10 region (LK1590).

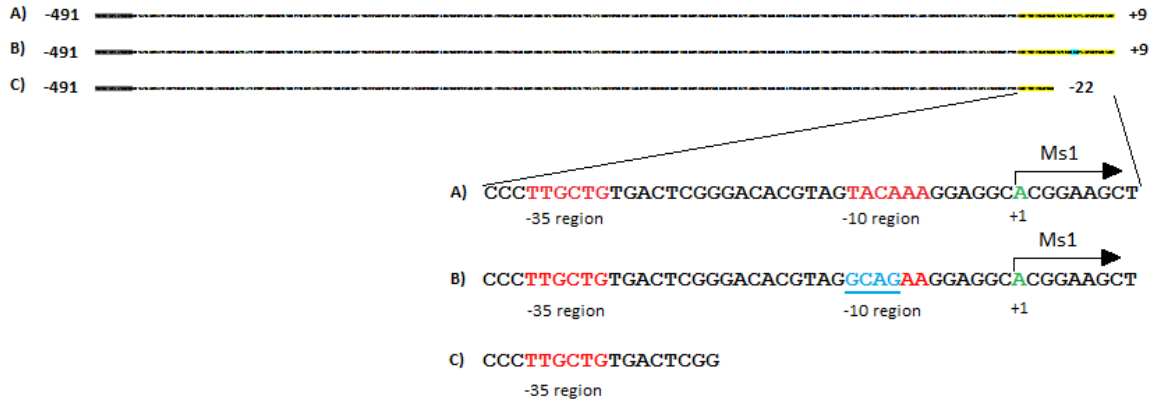

Figure 13. A) Ms1 promoter fragment of the length 500 nt B) Ms1 promoter fragment of the length 500 nt with mutated -10 conserved region C) Ms1 promoter fragment of the length 470 nt with deleted -10 conserved region

β-galactosidase activities of strains LK1578 and LK1590 were measured in exponential, early stationary and late stationary phase of growth. Results confirmed that there is no other promoter further upstream of transcription star site as no β-galactosidase activity was detectable in these strains (**Figure 14**).

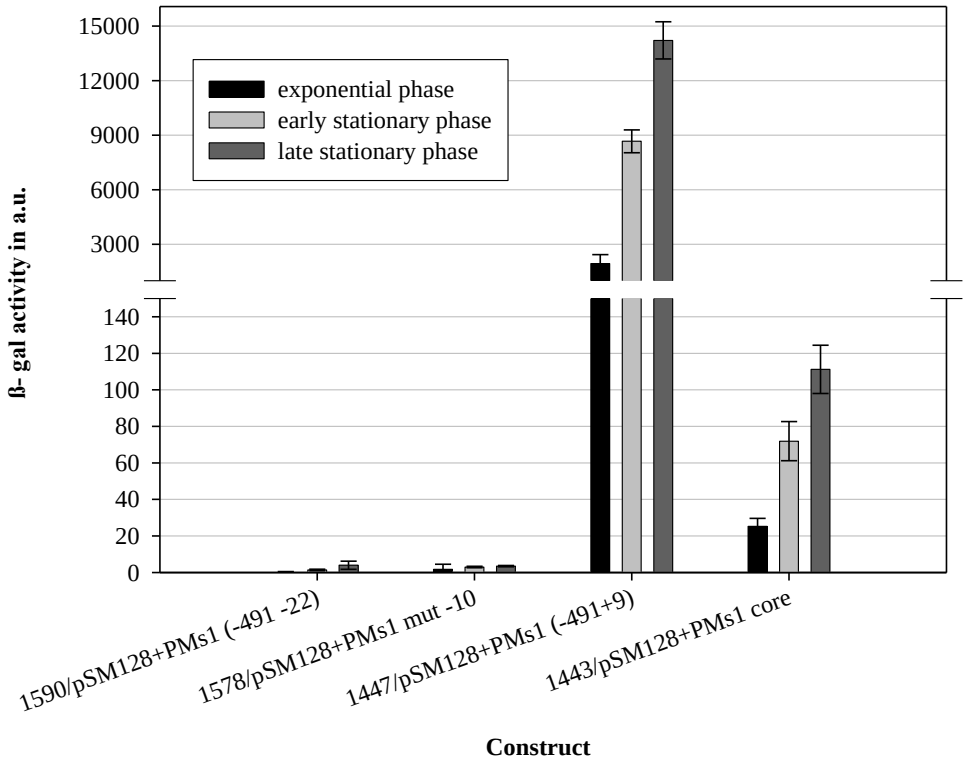

Figure 14. β-galactosidase activity of *M. smegmatis* strains LK1443, LK1447, LK1578 and LK1590 in exponential, early and late stationary phase measured in arbitrary units. Strains with modified Ms1 core promoter (LK1578, LK1590) showed no β-galactosidase activity. Graph shows mean values calculated from two biological replicates with samples collected in duplicates. Error bars represent standard deviation.

5.4 Search for transcription factor(s)

The level of expression of the reporter gene transcribed from the Ms1 promoter increased with increased length of the promoter region. As Ms1 is transcribed from a single promoter adjacent to the transcription start site, it is probable that this change is mediated by binding of a transcription activator. To identify proteins that bind to Ms1 promoter region, DNA fragment from -231 to +9 was prepared by PCR amplification using the forward primer with biotin at the 5' end. As a negative control, non-related DNA sequences from the N-terminal fragment of HelD from *B. subtilis* (202 nucleotides long) and a part of the *rpoC* gene from *M. smegmatis* (370 nucleotides long) were used. The sequences are shown in **Table 11**. DNA was bound to streptavidin coated magnetic beads and incubated with cell lysate. To identify proteins that had weaker interaction with DNA, some incubations were done under crosslinking conditions. **Figure 15** shows polyacrylamide gels with proteins. Bands marked with a rectangle were sent for mass spectrometry analysis. The results of mass spectrometry analysis are summarized in **Table 12**.

Table 11. Sequences of DNA fragments used for pull-down of DNA binding proteins. Sequences of primers are shown in red.

Ms1 promoter	[Btn]CGTGACCGGTTCCCCGCGCCGTCGGGTGCGGCGGTCGCCACCACCGC GGCCGTCGGCATCAGCGCACTGGCCGCCGGCGCTCTGACGTACTCGCTGCT GCGCCGGTTCGCGCTCTGACGCAATCGCTTCGGCGCCGTTTCGTAAAGGTA GCTGGACACCGCTGCGTAACGATCGGACGACGCGCGGATTTTGGCCCTTG CTGTGACTCGGGACACGTAGTACAAAGGAGGCACGGAAGCT
RpoC gene fragment	[Btn]CGACGAGATCTGGAACACCTTCACCAAGCTCGCTCCCAAGCAGCTCA TCGTCGACGAGGTGCTCTACCGCGAGCTGCAGGACCGCTACGGCGAGTAC TTCACCGGCGCCATGGGCGCGGAGTCGATCAAGAAGCTCATCGAGAACTT CGACATCGACGCCGAGGCCGAGTCGCTGCGCGAGGTCATCCGCAGCGGCA AGGGCCAGAAGAAGCTGCGTGCGCTCAAGCGCCTGAAGGTCGTCGCGGCC TTCCAGCAGTCGGGCAACTCGCCCATGGGCATGGTGCTCGACGCCGTTCCG GTGATCCCGCCGGAGCTGCGCCCGATGGTCCAGCTCGACGGTGGCCGCTTC GCGACCTCCGACCTGAACGAC
HelD gene fragment	[Btn]ATGAATCAGCAGGATAAGGAATGGAAGGAAGAGCAGTCGAGAATAG ATGAGGTGCTGAAGGAGCTCGAAAAAAAAAGAGCGGTTTCTGGAAACGTCT GCGGGCGGGTTGAAGCATGATATCATCGGATTAAGAAAAAGCTTTTGGGA GGATGTTAAGGTTAATTTTGATGATGCCCATGAAGCGATTGAAACGCTCGA GCCCG

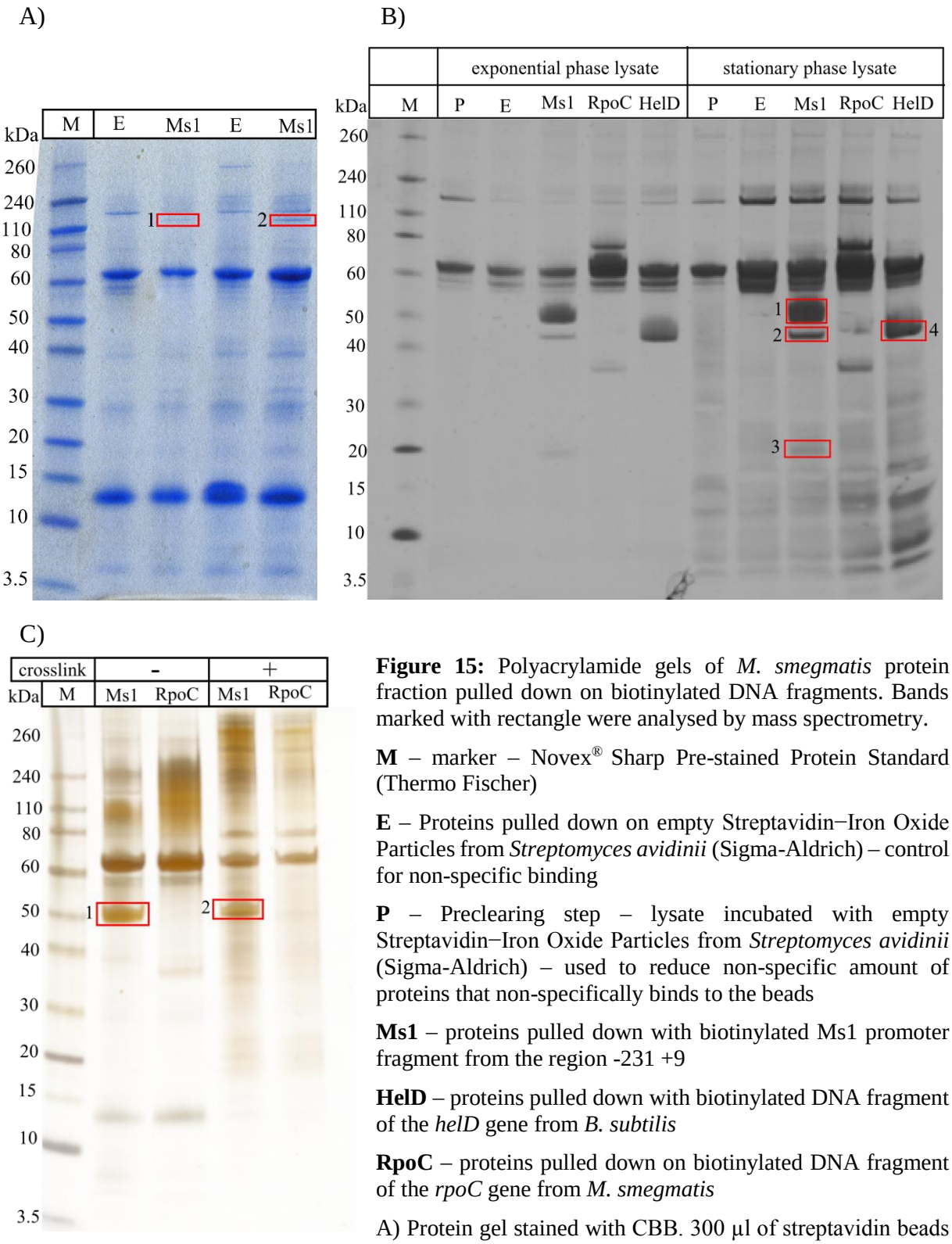

Figure 15: Polyacrylamide gels of *M. smegmatis* protein fraction pulled down on biotinylated DNA fragments. Bands marked with rectangle were analysed by mass spectrometry.

M – marker – Novex® Sharp Pre-stained Protein Standard (Thermo Fischer)

E – Proteins pulled down on empty Streptavidin−Iron Oxide Particles from *Streptomyces avidinii* (Sigma-Aldrich) – control for non-specific binding

P – Preclearing step – lysate incubated with empty Streptavidin−Iron Oxide Particles from *Streptomyces avidinii* (Sigma-Aldrich) – used to reduce non-specific amount of proteins that non-specifically binds to the beads

Ms1 – proteins pulled down with biotinylated Ms1 promoter fragment from the region -231 +9

HelD – proteins pulled down with biotinylated DNA fragment of the *helD* gene from *B. subtilis*

RpoC – proteins pulled down on biotinylated DNA fragment of the *rpoC* gene from *M. smegmatis*

A) Protein gel stained with CBB. 300 µl of streptavidin beads and 300 ml of cell lysate from stationary phase were used. B) Silver stained protein gel. 300 µl of streptavidin beads and 300 ml of cell lysate from exponential and stationary phase were used. C) Silver stained protein gel. 400 µl of streptavidin beads and 500 ml of cell lysate from stationary phase were used. For Ms1 and RpoC bands on the right side, incubation of streptavidin coated beads with bound DNA with cell lysate was carried out under crosslinking conditions.

Table 12. Results of mass spectrometry analysis of bands specified in **Figure 15**. Letters and numbers of samples correspond to the labels in **Figure 15**.

	Band number	Protein name	Database number	Number of peptides	Coverage [%]	MSMS analysis	MW protein
A	1	DNA topoisomerase I	500051163	15	24	GAQLAVTSVEQKPYTR	102
	2	DNA topoisomerase I	500051163	16	25	GAQLAVTSVEQKPYTR	102
		Streptavidin	SAV_STRAV	1	10	NAHSATTWSGQYVGGAEAR	19
B	**Band number**	**Protein name**	**Database number**	**Number of peptides**	**Coverage [%]**	**MSMS analysis**	**MW protein**
	1	Keratin, type II cytoskeletal 1	K2C1_HUMAN	5	13	No	66
		Keratin, type I cytoskeletal 10	K2C1_HUMAN	4	8	No	59
	2	keratins, autotrypsin					
	3	30S ribosomal protein S9	RS9_MYCS2	3	27	LVPGTGQFNLDGR EPVIIDRPIQTVGR	17
		Single-stranded DNA-binding protein	SSB_MYCS2	5	40	FTPSGAAVANFTVASTPR	17
	4	Streptavidin	SAV_STRAV			NAHSATTWSGQYVGGAEAR	19
C	**Band number**	**Protein name**	**Database number**	**Number of peptides**	**Coverage [%]**	**MSMS analysis**	**MW protein**
	1	Serum albumin	ALBU_BOVIN	4	7	No	66
	2	Elongation factor Tu	EFTU_MYCS2	12	39	LLDQGQAGDNVGLLLR	44
		Keratin, type I cytoskeletal 16	K1C16_HUMAN				51

5.5 Effect of topoisomerase I on Ms1 transcription

Among proteins identified by mass spectrometry, topoisomerase I encoded by the *topA* gene was the best candidate for being a transcription factor. Therefore, topoisomerase I was selected for further analysis. DNA topoisomerase I from *M. smegmatis* binds DNA in a sequence specific manner. It recognizes the CG/TCTTC/G hexanucleotide sequence (Sikder and Nagaraja, 2000). This sequence is absent in the Ms1 core promoter sequence but present in all longer promoter fragments. To test the effect of topoisomerase I on Ms1 expression, promoter sequence from -131 to +9 was mutated using 1853/Pms1-mutTopoI-F forward and 1322/Pms1_R reverse primers (**Figure 16**). This modified promoter was cloned into pSM128 and transformed into *M. smegmatis* resulting in a strain LK1781. β-galactosidase activity was measured in exponential, early and late stationary phases of growth. Promoter without the topoisomerase I binding sites showed a small decrease in activity in exponential phase and a minor increase in activity in early and late stationary phase (**Figure 17**).

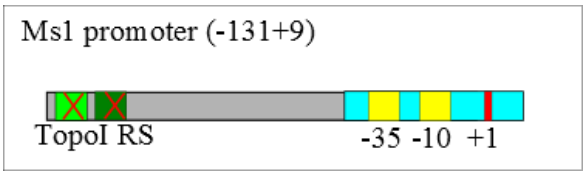

Figure 16. Scheme of Ms1 promoter region from -131 to +9 nucleotide. This promoter region contains core promoter (blue with consensus -10 and -35 regions in yellow and +1 transcription start site in red) and two topoisomerase I recognition sites (RS). Dark green box indicate topoisomerase I consensus sequence at recognition site (RS: CGCTT) as defined in Sikder and Nagaraja (2000); light green boxes represent sequence CGGTT. Red crosses indicate where mutations were made.

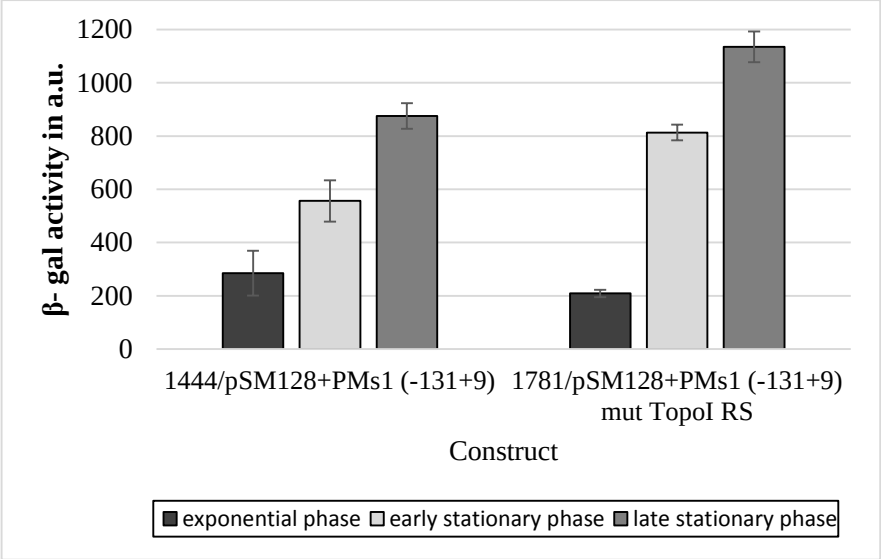

Figure 17. β-galactosidase activity of *M. smegmatis* strains LK1444 and LK1781 in exponential, early and late stationary phase measured in arbitrary units. Strain LK1781 with modified recognition site for topoisomerase I showed a small decrease in exponential phase and a slight increase of β-galactosidase activity in early and late stationary phase. Graph shows mean values calculated from two biological replicates with samples collected in duplicates. Error bars represent standard deviation.

5.6 Effects of stress on Ms1 promoter activity

Using the LK1447 strain with the Ms1 promoter region from -491 to +9, effects of various stress conditions on Ms1 promoter activity were tested. Nine stress conditions were selected. For heat and cold stress, cell cultures were transferred into 50°C and room temperature, respectively. For hyperosmotic, oxidative and alcohol stresses, the stressor was added to the medium. For alkaline, acid, hypo-osmotic stresses and starvation cells were filtered and resuspended in appropriate medium. To test whether filtration itself could affect the level of β-galactosidase activity, a control experiment was performed: cells were filtered and resuspended in fresh, preheated 7H9 medium; β-galactosidase activity was determined and it was the same as for the culture cultivated continuously without filtration. Hence, the filtration did not affect the promoter activity. For each stressor, three biological replicates were performed and samples were collected in duplicates. Promoter activity was measured 2 and 4 hours after stress, respectively. Results are shown in **Figure 18**. No significant change in activity was visible for most stress conditions. After hyperosmotic stress, there was a slight decrease in activity compared to control samples. Alcohol stress resulted in an increase in activity, similar to the increase in activity during transition to stationary phase.

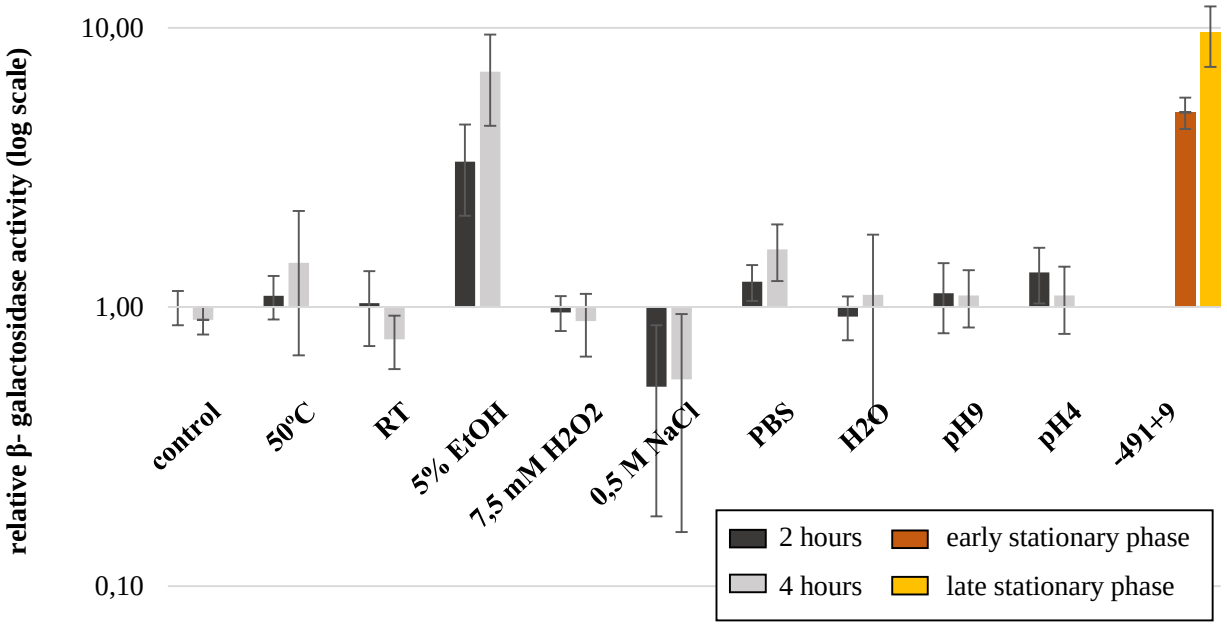

Figure 18. Relative β-galactosidase activity of *M. smegmatis* strain LK1447 under stress. Graph shows average values of expression of three biological replicates. Values were normalized to activity of control sample at time 0 (before stress); The y-axis is shown in logarithmic scale. Error bars represent standard deviation. Values of LK1447 strain β-galactosidase activity in early and late stationary phase normalized to expression in exponential phase were added to the graph for easier comparison to what extent the expression of β-galactosidase was affected by stress.

6. Discussion

6.1 Preparation of *M. smegmatis* strains

6.1.1 Preparation of Ms1 promoter fragments

Transcription is mediated by binding of RNAP holoenzyme to promoter sequence preceding the gene. Transcription from a promoter can be affected by additional elements. In bacteria these elements are usually present in close proximity to the core promoter. Therefore, a series of Ms1 promoter fragments with increasing length was prepared. Each fragment contained the core promoter region and differed in length of the upstream region. The shortest fragment was from the -38 nucleotide to the +9 nucleotide and the longest from -491 to +9. The first nine nucleotides of Ms1 were included in promoter fragments as this region may have an effect on transcription initiation (Hsu et al., 2006).

6.1.2 Cloning Ms1 promoter fragments into pSM128

DNA fragments of interest are usually cloned into plasmids by ligation into a multiple cloning site. The pSM128 plasmid does not contain a polylinker. Therefore, we decided to insert an artificial polylinker, similarly as previously described (Krásný et al., 2000). However, restriction analysis of this plasmid with commonly used restriction enzymes showed there was no possibility to create such multiple cloning site. All enzymes used for restriction analysis cleaved pSM128 at least once elsewhere. pSM128 contains a single restriction site for ScaI, which creates blunt ends. Ms1 promoter fragments were cloned into this site using TA cloning. This method takes advantage of the ability of *Taq* polymerase from *Thermus aquaticus* to add a non-templated nucleotide to the 3' end of dsDNA. From four bases, *Taq* polymerase adds adenine most efficiently, but it is capable of addition of any of the bases (Hu, 1993). Single A was added to the inserts and single T to the plasmid. By this method, however, the insert can be ligated into the plasmid in both orientations. Because of this, a larger number of transformed bacteria must have been tested before the correct clone for each construct was found.

6.1.3 Transformation of bacteria

The ligation mixtures containing the DNA fragments and pSM128 plasmids were transformed into *E. coli* using heat shock. Subsequently, the correct, sequence-verified plasmids were transformed into *M. smegmatis* cells by electroporation. To enhance the selection, X-gal was added to the agar, resulting in blue colonies (Gurvitz et al., 1993).

6.2 Effect of promoter length on Ms1 expression

Fusion of a promoterless reporter gene with the promoter of a gene of interest is a widely used technique to assess strength of the promoter. Different reporter genes can be used and various reporter vectors were created for use in mycobacteria. Gupta and colleges created the pSD7 replicative vector containing chloramphenicol acetyltransferase reporter gene. They used this vector to detect promoter elements from *M. smegmatis* and *M. tuberculosis* (Gupta et al., 1993). Another reporter gene is the gene for green fluorescent protein (GFP). An example of mycobacterial reporter vectors using GFP are pFlame replicative plasmids containing unstable GFP variants (Blokpoel et al., 2003). For this study, the integrative pSM128 plasmid containing the *lacZ* reporter gene was used (Dussurget et al., 1999). The *lacZ* gene encodes β-galactosidase, a tetrameric enzyme which cleaves lactose into glucose and galactose. β-galactosidase can also cleave X-gal producing blue product or o-nitrophenyl-β-D-galactopyranoside (ONPG) resulting in yellow colour of reaction product (Juers et al., 2012). The advantage of the integrative vector is the only copy of the construct per cell, obviating the need to take into account multiple copies in the case of replicative plasmids (Guérout-Fleury et al., 1996).

Expression form the Ms1 promoter differed between growth phases. For strain LK1443 with the shortest promoter fragment, expression between exponential and early stationary phase increased ~3.2 times and between exponential and late stationary phase ~5.7 times. For strain LK1447 with the longest promoter fragment, expression between exponential and early stationary phase increased ~5.4 times and between exponential and late stationary phase ~10.1 times. In the RNA-Seq experiments done in our laboratory, the amount of Ms1 in exponential phase was ~370 times lower compared to stationary phase (unpublished data). In northern blot analysis this difference was ~130 times (Hnilicová et al., 2014). However, the amount of the Ms1 homologue in *M. tuberculosis*, MTS2823, increased in stationary phase only 10 times (Arnvig et al., 2011). RNA-Seq and northern blot measure the amount of RNA present in the cell which is affected by both promoter activity and RNA stability. Promoter-*lacZ* fusion measures the level of transcription from a particular promoter. Our results suggest that difference in the amount of Ms1 in exponential and stationary phase is a combined result of increased promoter activity and higher Ms1 stability or a lowered degradation rate.

β-galactosidase activity experiments revealed significant differences between tested promoter constructs. Promoter activity between strain LK1443 and strain LK1445 differed almost 100

times, independently of growth phase. It can be assumed that some kind of a transcription factor binds to the promoter region between -38 and -231 nucleotide. Other possibility was that there is stronger promoter present upstream of the predicted promoter. This hypothesis was tested and disproved (see Discussion section 6.3).

6.3 Identification of Ms1 promoter

In a previous study the +1 transcription start site (TSS) of Ms1 was identified (Hnilicová et al., 2014). Upstream of this TSS a sequence resembling consensus sequence of primary sigma factor from *M. smegmatis* was found. In this work, it was tested whether this was the only Ms1 promoter or if there was another promoter present further upstream in the chromosome.

The Ms1 promoter sequence from -491 nucleotide to +9 nucleotide was tested. When the sequence corresponding to the predicted -10 consensus sequence was modified or deleted this promoter fragment lost its ability to initiate transcription, in agreement with known deleterious effects of consensus sequence mutations on promoter binding by RNAP (Raindlová et al., 2016). Therefore, it may be assumed that Ms1 is transcribed from a single promoter present directly upstream of TSS. In contrast, 6S RNA in *E. coli* is transcribed from two promoters (Kim and Lee, 2004).

6.4 Search for transcription factor(s)

The activity of Ms1 promoter fragments increased with increased length of the upstream sequence. A probable explanation is that this sequence binds a transcription activator. The method selected to identify this transcription factor was protein pull-down with biotinylated DNA. The same method was used to identify Ms1 interaction partners using biotinylated RNA (Hnilicová et al., 2014). The Ms1 promoter region from -231 to +9 nucleotide was used. This region was selected because the increase in promoter activity between the core promoter and the 240 nt long promoter was more significant than the change between 240 nt and 500 nt long promoter fragments. The usage of a longer DNA fragment could also lead to more non-specific binding events and make it more challenging to identify specific transcription activators.

Several rounds of protein pull-down experiments were carried out. The amount of used cell lysate as well as biotinylated DNA and streptavidin beads were optimized during the experiments. Proteins were separated on gradient polyacrylamide gels and visualized by Coomassie blue. To increase sensitivity and visualize lower protein amounts, some gels were silver stained.

Protein bands specific for the Ms1 promoter region were sent for mass spectrometry analysis. Most samples contained streptavidin. Streptavidin contamination must have appeared in the bands as a result of using streptavidin iron-oxide particles for protein isolation. Other identified protein was human keratin which is one of common contaminants in mass spectrometry analyse (reviewed in Keller et al., 2008).

Four mycobacterial proteins that bound to the Ms1 promoter region were identified by mass spectrometry:

- Ribosomal protein S9 – a component of the small ribosomal subunit; S9 was predicted to be essential for *M. tuberculosis* (Griffin et al., 2011).
- Elongation factor Tu (EF-Tu) – a highly abundant protein in the cell; EF-Tu from *M. smegmatis* and *E. coli* are very similar both in DNA sequence and in function (Bruell et al., 2008).
- Topoisomerase I (TopoI) – an enzyme primarily responsible for negative DNA supercoil relaxation; TopoI is an essential protein for *M. smegmatis*. Conditional knock-down of TopoI resulted in many changes in *M. smegmatis* cell. Among them was also an altered distribution of RNAP on the transcription units (Ahmed et al., 2015).
- Single-stranded DNA-binding protein (SSB) – a protein involved in processes of DNA replication, repair and recombination; SSB interacts with many other proteins in the cell and it also stimulates TopoI activity (Sikder et al., 2001).

Ribosomal protein S9, EF-Tu and SSB are all primarily involved in cellular processes other than transcription. It is not probable that one of them is the sought for transcription factor, although SSB may cooperate with TopoI. The role of TopoI in Ms1transcription was tested and it is discussed in the next section.

6.5 Role of topoisomerase I in Ms1 transcription

Topoisomerase I encoded by the *topA* gene was one of the proteins identified to bind to Ms1 promoter. TopoI from *M. smegmatis* is a single subunit enzyme of 110 kDa. TopoI relaxes negatively supercoiled DNA, catenates and knots ssDNA. It retains relaxation activity even at 65°C (Bhaduri et al., 1998). TopoI from *M. smegmatis* can interact with both dsDNA and ssDNA and it recognizes the CG/TCTTC/G hexanucleotide sequence (Sikder and Nagaraja, 2000). Strain LK1444 prepared in this work contains CGCTTC and CGGTTC sequences in the Ms1 promoter fragment. The CGGTTC sequence is infrequent among TopoI binding sites, but we decided to mutate it as well to prevent even weak binding of TopoI. Strain LK1781 contains

the Ms1 promoter with mutant TopoI recognition sites. If TopoI was the transcription factor responsible for the observed increase in Ms1 promoter activity dependent on the upstream DNA, this modification would decrease its β-galactosidase activity to the level comparable to the core promoter. However, β-galactosidase activity of strain LK1781 decreased only slightly in exponential phase and in early and late stationary phase we observed a minor increase compared to strain LK1444. Hence, it is possible to conclude that topoisomerase I does not function as a transcription activator of the Ms1 promoter.

The increase in the promoter activity of the fragment with the mutated TopoI recognition site can be explained by a direct inhibitory effect of topoisomerase cleavage or change in DNA topology. Another explanation might be that binding sites of transcription activator and topoisomerase I partially overlap. Therefore, preventing binding of topoisomerase I to DNA could improve binding of transcription activator to the promoter. Whether either of these hypothesis is true needs to be tested.

6.6 Changes in promoter activity under stress

Effects of various stress conditions on Ms1 promoter activity were tested. Tested stress conditions included heat and cold shock, acid, alkali, oxidative and ethanol stress, starvation and hyper and hypo osmotic stress.

The stress conditions testing protocol was adapted from the study of Gebhard et al. (2008) and modified. Gebhard and colleagues were testing the role of alternative sigma factor SigF in stress response of *M. smegmatis*. Modifications were made for alkali and acidic stresses. Instead of a citrate/phosphate and sodium borate buffers, the 7H9 medium with pH adjusted to pH4 and pH9 were used. This was done to restrict the difference between control and stressed cells only to different pH. Some stress conditions required cells to be collected and resuspended in a new medium. The collection of cells was done by filtering instead of centrifugation and this did not affect activity of the Ms1 promoter, analogously to previous experiments with promoters monitored after amino acids starvation (Krásný et al., 2008).

The Ms1 promoter activity remained unchanged for most stress conditions. A slight decrease in activity was visible for hyperosmotic stress. When the cells were exposed to 5% ethanol Ms1 promoter activity increased to the level comparable with promoter activity in stationary phase. Ethanol stress is shown to induce changes in expression of various genes in bacteria (Bermudez et al., 1993; Kendall et al., 2004; Thackray and Moir, 2003).

7. Conclusions

Ms1 expression was characterized at the level of transcription

- Ms1 is transcribed from a single promoter directly upstream of the +1 transcription start site.
- The level of Ms1 transcription increases ~10 times between exponential and late stationary phase.
- Ms1 promoter activity increases ~100 times when the promoter sequence is lengthened.
- The increase in promoter activity with increased length is probably the result of binding of a transcription activator.
- The Ms1 promoter activity is not affected by heat shock, cold shock, alkali and acidic environment, hypo osmotic stress, oxidative stress and starvation.
- Ms1 promoter activity increases after ethanol stress almost to the level of transcription in stationary phase; after hyperosmotic stress, the activity slightly decreases.
- Topoisomerase I recognition sites are present in the Ms1 promoter region, their modification led to a subtle increase in Ms1 promoter activity in stationary phase.

The most significant results and an overall scheme of regulation of Ms1 expression are shown in **Figure 19**.

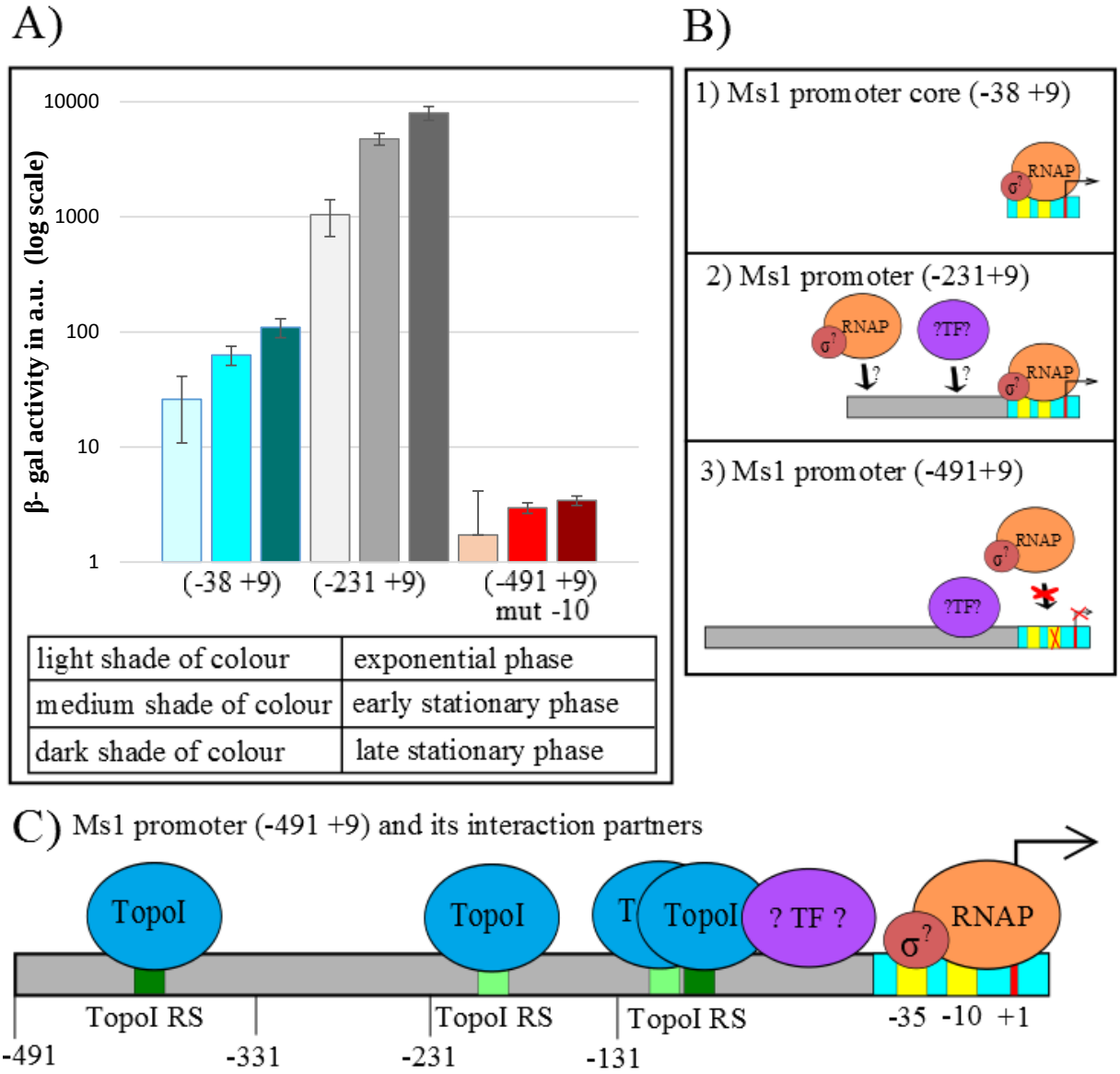

Figure 19. Summary of regulation of Ms1 expression.
A) β-galactosidase activity of *M. smegmatis* strains containing selected promoter-*lacZ* fusions. The graph shows β-galactosidase activity of three mycobacterial strains with Ms1 promoters corresponding to the schemes in B. **B) Schemes of selected Ms1 promoter regions used in this work**. Scheme 1 – Ms1 core promoter region from -38 to +9 was sufficient to initiate transcription. Scheme 2 – Addition of ~200 bp upstream of the promoter resulted in a ~100-fold increase in promoter activity. It was not known whether this activation was mediated by another stronger promoter in the upstream region or by binding of a transcription activator. Scheme 3 – RNA polymerase was not able to initiate transcription from the -491 to +9 promoter region fragment when the -10 region in the core promoter was mutated. Therefore, the increase in promoter activity with increased length of the promoter fragment is mediated by binding of an unknown transcription factor. **C) Model of the Ms1 promoter region from -491 to +9 and its interaction with protein factors**. Ms1 is transcribed from a single promoter. It is not known which sigma factor initiates transcription from the Ms1 promoter. The upstream region contains multiple topoisomerase I recognition sites (TopoI RS), however, their mutation had insignificant effects on Ms1 transcription. Promoter activity is also affected by another transcription factor (TF). The core promoter region is shown in blue with consensus -10 and -35 regions in yellow and the +1 transcription start site in red. Dark green boxes indicate topoisomerase I consensus recognition sequence (CG/TCTT) as defined in Sikder and Nagaraja (2000); light green boxes represent the CGGTT sequence. The schematic representations are not drawn to scale.

8. References

Ahmed, W., Menon, S., Karthik, P.V., and Nagaraja, V. (2015). Reduction in DNA topoisomerase I level affects growth, phenotype and nucleoid architecture of Mycobacterium smegmatis. Microbiol.-Sgm *161*, 341–353.

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*, e1002342.

Bashyam, M.D., and Tyagi, A.K. (1998). Identification and Analysis of "Extended −10" Promoters from Mycobacteria. J. Bacteriol. *180*, 2568–2573.

Bashyam, M.D., Kaushal, D., Dasgupta, S.K., and Tyagi, A.K. (1996). A study of mycobacterial transcriptional apparatus: identification of novel features in promoter elements. J. Bacteriol. *178*, 4847–4853.

Bermudez, L.E., Young, L.S., Martinelli, J., and Petrofsky, M. (1993). Exposure to Ethanol Up-Regulates the Expression of Mycobacterium avium Complex Proteins Associated with Bacterial Virulence. J. Infect. Dis. *168*, 961–968.

Bhaduri, T., Bagui, T.K., Sikder, D., and Nagaraja, V. (1998). DNA Topoisomerase I from Mycobacterium smegmatis An Enzyme With Distinct Features. J. Biol. Chem. *273*, 13925–13932.

Blatter, E.E., Ross, W., Tang, H., Gourse, R.L., and Ebright, R.H. (1994). Domain organization of RNA polymerase α subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. Cell *78*, 889–896.

Blokpoel, M.C.J., O'Toole, R., Smeulders, M.J., and Williams, H.D. (2003). Development and application of unstable GFP variants to kinetic studies of mycobacterial gene expression. J. Microbiol. Methods *54*, 203–211.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248–254.

Brandis, G., and Hughes, D. (2013). Genetic characterization of compensatory evolution in strains carrying rpoB Ser531Leu, the rifampicin resistance mutation most frequently found in clinical isolates. J. Antimicrob. Chemother. *68*, 2493–2497.

Bruell, C.M., Eichholz, C., Kubarenko, A., Post, V., Katunin, V.I., Hobbie, S.N., Rodnina, M.V., and Böttger, E.C. (2008). Conservation of Bacterial Protein Synthesis Machinery: Initiation and Elongation in Mycobacterium smegmatis. Biochemistry (Mosc.) *47*, 8828–8839.

Cabrera-Ostertag, I.J., Cavanagh, A.T., and Wassarman, K.M. (2013). Initiating nucleotide identity determines efficiency of RNA synthesis from 6S RNA templates in Bacillus subtilis but not Escherichia coli. Nucleic Acids Res. *41*, 7501–7511.

Carroll, P., and James, J. (2009). Assaying Promoter Activity Using LacZ and GFP as Reporters. In Mycobacteria Protocols, T. Parish, and A.C. Brown, eds. (Humana Press), pp. 265–277.

Cases, I., de Lorenzo, V., and Ouzounis, C.A. (2003). Transcription regulation and environmental adaptation in bacteria. Trends Microbiol. *11*, 248–253.

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C.A., Holland, T.A., Keseler, I.M., Kothari, A., Kubo, A., et al. (2014). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Res. *42*, D459–D471.

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 σ70. Mol. Microbiol. *67*, 1242–1256.

Cavanagh, A.T., Chandrangsu, P., and Wassarman, K.M. (2010). 6S RNA regulation of relA alters ppGpp levels in early stationary phase. Microbiology *156*, 3791–3800.

Cavanagh, A.T., Sperger, J.M., and Wassarman, K.M. (2012). Regulation of 6S RNA by pRNA synthesis is required for efficient recovery from stationary phase in E. coli and B. subtilis. Nucleic Acids Res. *40*, 2234–2246.

Chae, H., Han, K., Kim, K., Park, H., Lee, J., and Lee, Y. (2011). Rho-dependent Termination of ssrS (6S RNA) Transcription in Escherichia coli IMPLICATION FOR 3′ PROCESSING OF 6S RNA AND EXPRESSION OF DOWNSTREAM ygfA (PUTATIVE 5-FORMYL-TETRAHYDROFOLATE CYCLO-LIGASE). J. Biol. Chem. *286*, 114–122.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. v., Eiglmeier, K., Gas, S., Barry III, C. e., et al. (1998). Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. (Cover story). Nature *393*, 537.

Dennis, P.P., Ehrenberg, M., Fange, D., and Bremer, H. (2009). Varying Rate of RNA Chain Elongation during rrn Transcription in Escherichia coli. J. Bacteriol. *191*, 3740–3746.

Dussurget, O., Timm, J., Gomez, M., Gold, B., Yu, S., Sabol, S.Z., Holmes, R.K., Jacobs, W.R., and Smith, I. (1999). Transcriptional Control of the Iron-ResponsivefxbA Gene by the Mycobacterial Regulator IdeR. J. Bacteriol. *181*, 3402–3408.

Fadouloglou, V.E., Lin, H.-T.V., Tria, G., Hernández, H., Robinson, C.V., Svergun, D.I., and Luisi, B.F. (2015). Maturation of 6S regulatory RNA to a highly elongated structure. FEBS J. *282*, 4548–4564.

Gebhard, S., Hümpel, A., McLellan, A.D., and Cook, G.M. (2008). The alternative sigma factor SigF of Mycobacterium smegmatis is required for survival of heat shock, acidic pH and oxidative stress. Microbiology *154*, 2786–2795.

Gildehaus, N., Neußer, T., Wurm, R., and Wagner, R. (2007). Studies on the function of the riboregulator 6S RNA from E. coli: RNA polymerase binding, inhibition of in vitro transcription and synthesis of RNA-directed de novo transcripts. Nucleic Acids Res. *35*, 1885–1896.

Graham, J.E., and Richardson, J.P. (1998). rut Sites in the Nascent Transcript Mediate Rho-dependent Transcription Termination in Vivo. J. Biol. Chem. *273*, 20764–20769.

Gries, T.J., Kontur, W.S., Capp, M.W., Saecker, R.M., and Record, M.T. (2010). One-step DNA melting in the RNA polymerase cleft opens the initiation bubble to form an unstable open complex. Proc. Natl. Acad. Sci. *107*, 10418–10423.

Griffin, J.E., Gawronski, J.D., DeJesus, M.A., Ioerger, T.R., Akerley, B.J., and Sassetti, C.M. (2011). High-Resolution Phenotypic Profiling Defines Genes Essential for Mycobacterial Growth and Cholesterol Catabolism. PLoS Pathog. *7*.

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

Gupta, S.K.D., Bashyam, M.D., and Tyagi, A.K. (1993). Cloning and assessment of mycobacterial promoters by using a plasmid shuttle vector. J. Bacteriol. *175*, 5186–5192.

Gurvitz, A., Coe, J.G.S., and Dawes, I.W. (1993). Use of reporter genes for the isolation and characterisation of different classes of sporulation mutants in the yeast Saccharomyces cerevisiae. Curr. Genet. *24*, 451–454.

Gusarov, I., and Nudler, E. (1999). The Mechanism of Intrinsic Transcription Termination. Mol. Cell *3*, 495–504.

Harden, T.T., Wells, C.D., Friedman, L.J., Landick, R., Hochschild, A., Kondev, J., and Gelles, J. (2016). Bacterial RNA polymerase can retain σ70 throughout transcription. Proc. Natl. Acad. Sci. *113*, 602–607.

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.

Healy, J.M., Lesoon, A., Bodorova, J., Lam, K., and Wobbe, C.R. (1998). Isolation and characterization of the rpoA gene of Mycobacterium tuberculosis. Protein Pept. Lett. *5*, 185–191.

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.

Hsu, L.M., Cobb, I.M., Ozmore, J.R., Khoo, M., Nahm, G., Xia, L., Bao, Y., and Ahn, C. (2006). Initial transcribed sequence mutations specifically affect promoter escape properties. Biochemistry (Mosc.) *45*, 8841–8854.

Hu, G. (1993). DNA Polymerase-Catalyzed Addition of Nontemplated Extra Nucleotides to the 3′ of a DNA Fragment. DNA Cell Biol. *12*, 763–770.

Juers, D.H., Matthews, B.W., and Huber, R.E. (2012). LacZ β-galactosidase: Structure and function of an enzyme of historical and molecular biological importance. Protein Sci. Publ. Protein Soc. *21*, 1792–1807.

Kapopoulou, A., Lew, J.M., and Cole, S.T. (2011). The MycoBrowser portal: A comprehensive and manually annotated resource for mycobacterial genomes. Tuberculosis *91*, 8–13.

Keller, B.O., Sui, J., Young, A.B., and Whittal, R.M. (2008). Interferences and contaminants encountered in modern mass spectrometry. Anal. Chim. Acta *627*, 71–81.

Kendall, S.L., Movahedzadeh, F., Rison, S.C.G., Wernisch, L., Parish, T., Duncan, K., Betts, J.C., and Stoker, N.G. (2004). The Mycobacterium tuberculosis dosRS two-component system is induced by multiple stresses. Tuberculosis *84*, 247–255.

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

Klocko, A.D., and Wassarman, K.M. (2009). 6S RNA binding to Eσ70 requires a positively charged surface of σ70 region 4.2. Mol. Microbiol. *73*, 152–164.

Krásný, L., Vacík, T., Fučík, V., and Jonák, J. (2000). Cloning and Characterization of the str Operon and Elongation Factor Tu Expression in Bacillus stearothermophilus. J. Bacteriol. *182*, 6114–6122.

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.

Lee, J.Y., Park, H., Bak, G., Kim, K., and Lee, Y. (2013). Regulation of Transcription from Two ssrS Promoters in 6S RNA Biogenesis. Mol. Cells *36*, 227–234.

Lee, S.Y., Bailey, S.C., and Apirion, D. (1978). Small Stable RNAs from Escherichia coli: Evidence for the Existence of New Molecules and for a New Ribonucleoprotein Particle Containing 6S RNA. J. Bacteriol. *133*, 1015–1023.

Lew, J.M., Kapopoulou, A., Jones, L.M., and Cole, S.T. (2011). TubercuList – 10 years after. Tuberculosis *91*, 1–7.

Mathew, R., Mukherjee, R., Balachandar, R., and Chatterji, D. (2006). Deletion of the rpoZ gene, encoding the ω subunit of RNA polymerase, results in pleiotropic surface-related phenotypes in Mycobacterium smegmatis. Microbiology *152*, 1741–1750.

Minch, K.J., Rustad, T.R., Peterson, E.J.R., Winkler, J., Reiss, D.J., Ma, S., Hickey, M., Brabant, W., Morrison, B., Turkarslan, S., et al. (2015). The DNA-binding network of Mycobacterium tuberculosis. Nat. Commun. *6*, 5829.

Murakami, K.S. (2015). Structural Biology of Bacterial RNA Polymerase. Biomolecules *5*, 848–864.

Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V., and Ebright, R.H. (2000). Structural Organization of the RNA Polymerase-Promoter Open Complex. Cell *101*, 601–611.

Neusser, T., Gildehaus, N., Wurm, R., and Wagner, R. (2008). Studies on the expression of 6S RNA from E. coli: involvement of regulators important for stress and growth adaptation. Biol. Chem. *389*, 285–297.

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*, 165.

Paget, M.S. (2015). Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. Biomolecules *5*, 1245–1265.

Panchapakesan, S.S.S., and Unrau, P.J. (2012). E. coli 6S RNA release from RNA polymerase requires σ70 ejection by scrunching and is orchestrated by a conserved RNA hairpin. RNA *18*, 2251–2259.

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.

Raindlová, V., Janoušková, M., Slavíčková, M., Perlíková, P., Boháčová, S., Milisavljevič, N., Šanderová, H., Benda, M., Barvík, I., Krásný, L., et al. (2016). Influence of major-groove chemical modifications of DNA on transcription by bacterial RNA polymerases. Nucleic Acids Res. *44*, 3000–3012.

Revyakin, A., Liu, C., Ebright, R.H., and Strick, T.R. (2006). Abortive Initiation and Productive Initiation by RNA Polymerase Involve DNA Scrunching. Science *314*, 1139–1143.

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 alpha subunit of RNA polymerase. Science *262*, 1407–1413.

Rudra, P., Prajapati, R.K., Banerjee, R., Sengupta, S., and Mukhopadhyay, J. (2015). Novel mechanism of gene regulation: the protein Rv1222 of Mycobacterium tuberculosis inhibits transcription by anchoring the RNA polymerase onto DNA. Nucleic Acids Res. *43*, 5855–5867.

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

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

Sikder, D., and Nagaraja, V. (2000). Determination of the recognition sequence of Mycobacterium smegmatis topoisomerase I on mycobacterial genomic sequences. Nucleic Acids Res. *28*, 1830–1837.

Sikder, D., Unniraman, S., Bhaduri, T., and Nagaraja, V. (2001). Functional cooperation between topoisomerase I and single strand DNA-binding protein1. J. Mol. Biol. *306*, 669–679.

Skordalakes, E., and Berger, J.M. (2003). Structure of the Rho Transcription Terminator. Cell *114*, 135–146.

Smeulders, M.J., Keer, J., Speight, R.A., and Williams, H.D. (1999). Adaptation of Mycobacterium smegmatis to Stationary Phase. J. Bacteriol. *181*, 270–283.

Steuten, B., and Wagner, R. (2012). A conformational switch is responsible for the reversal of the 6S RNA-dependent RNA polymerase inhibition in Escherichia coli. Biol. Chem. *393*, 1513–1522.

Steuten, B., Setny, P., Zacharias, M., and Wagner, R. (2013). Mapping the Spatial Neighborhood of the Regulatory 6S RNA Bound to Escherichia coli RNA Polymerase Holoenzyme. J. Mol. Biol. *425*, 3649–3661.

Thackray, P.D., and Moir, A. (2003). SigM, an Extracytoplasmic Function Sigma Factor of Bacillus subtilis, Is Activated in Response to Cell Wall Antibiotics, Ethanol, Heat, Acid, and Superoxide Stress. J. Bacteriol. *185*, 3491–3498.

Trotochaud, A.E., and Wassarman, K.M. (2004). 6S RNA Function Enhances Long-Term Cell Survival. J. Bacteriol. *186*, 4978–4985.

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.

Trotochaud, A.E., and Wassarman, K.M. (2006). 6S RNA Regulation of pspF Transcription Leads to Altered Cell Survival at High pH. J. Bacteriol. *188*, 3936–3943.

Tsai, C.-H., Baranowski, C., Livny, J., McDonough, K.A., Wade, J.T., and Contreras, L.M. (2013). Identification of Novel sRNAs in Mycobacterial Species. PLoS ONE *8*, e79411.

Waagmeester, A., Thompson, J., and Reyrat, J.-M. (2005). Identifying sigma factors in Mycobacterium smegmatis by comparative genomic analysis. Trends Microbiol. *13*, 505–509.

Wassarman, K.M., and Saecker, R.M. (2006). Synthesis-Mediated Release of a Small RNA Inhibitor of RNA Polymerase. Science *314*, 1601–1603.

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

Waters, L.S., and Storz, G. (2009). Regulatory RNAs in Bacteria. Cell *136*, 615–628.

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

World Health Organization (2015). Global tuberculosis report 2015 (World Health Organization).

Wurm, R., Neusser, T., and Wagner, R. (2010). 6S RNA-dependent inhibition of RNA polymerase is released by RNA-dependent synthesis of small de novo products. Biol. Chem. *391*, 187–196.

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. Cell *98*, 811–824.

Zhou, J., and Rudd, K.E. (2013). EcoGene 3.0. Nucleic Acids Res. *41*, D613–D624.