Charles University

Faculty of Science

Doctoral study program: Molecular and Cellular Biology, Genetics and Virology

Ing. Martina Janoušková

Interaction of nucleic acids with RNA polymerase

Interakce nukleových kyselin s RNA polymerázou

Doctoral Thesis

Supervisor: Mgr. Libor Krásný, Ph.D.

Department of Microbial Genetics and Gene Expression

Institute of Microbiology of the ASCR, v. v. i.

Prague, 2019

Prohlášení:

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

Declaration:

I hereby declare that I have written this Doctoral Thesis independently and that I have cited all sources and literature. This thesis has not been submitted for any other degree or purposes.

April 2019, Prague

…….....

Ing. Martina Janoušková

ACKNOWLEDGEMENT

I would like to thank especially to Libor Krásný for supervision and for his patience during my doctoral study. I am very grateful for the opportunity to be a member of his team in the laboratory of Microbial Genetics and Gene Expression.

Next, I would like to thank Jarmila Hnilicová for leading the Ms1 project and for many brilliant ideas, and to Michal Hocek´s group, especially to Zuzana Vaníková, Veronika Raindlová and Michaela Slavíčková for collaboration on the modified DNA project.

Many thanks go to all members of the "Krásná laborka" for all their help during my studies, very friendly environment, and all the great times we spent together.

Finally, I would like to thank my family and my friends for their great support.

I would like to thank the foundation "Nadání Josefa, Marie a Zdeňky Hlávkových" for a fellowship.

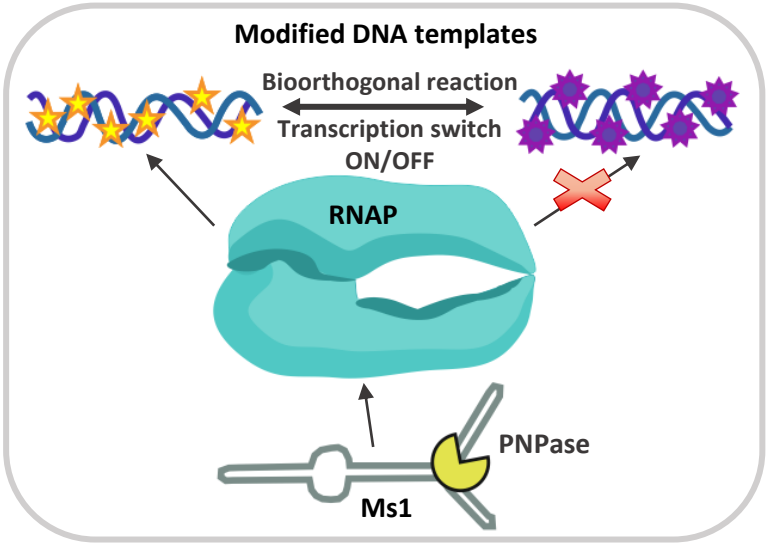
ABSTRACT

Regulation of gene expression by RNA polymerase (RNAP) is an essential ability of living organisms, required for their adaption to a changing environment and ultimately enabling their survival. Interaction of RNAP with ribonucleic acids (DNA or RNA) is crucial for transcription and its regulation. This Doctoral Thesis contains two projects addressing interactions of RNAP with nucleic acids: (i) Transcription of modified DNA templates and (ii) Ms1, a small RNA (sRNA) from *M. smegmatis*.

(i) We investigated the influence of modifications in the major groove of DNA on bacterial transcription *in vitro*. We found out that transcription of modified DNA templates is influenced on the transcription initiation level and that the promoter sequence is important for the effect of the modifications. Furthermore, we successfully performed transcription switch ON and OFF *in vitro* by bioorthogonal reactions. This regulation of transcription by artificial DNA modifications has a future in biotechnologies and/or medical therapy.

(ii) Regulators of transcription are also small non-coding RNAs. These molecules have an important role in gene expression regulation among prokaryotes and eukaryotes. Ms1 is an sRNA found in mycobacteria. It makes a complex with the RNAP core and it is abundant in stationary phase (in amounts comparable to those of ribosomal RNA [rRNA]). We investigated Ms1 synthesis and degradation in *M. smegmatis*. We discovered that the high abundance of Ms1 in stationary phase is caused by its high stability. Next, RNase polynucleotide phosphorylase (PNPase) has an impact on Ms1 degradation. Future studies may utilize this knowledge in the development of specific antimycobacterial drugs.

Graphical abstract:



ABSTRAKT

Regulace genové exprese RNA polymerázou (RNAP) je nezbytnou schopností živých organizmů, která v konečném důsledku umožňuje jejich přežití. Interakce RNAP s nukleovými kyselinami (DNA nebo RNA) je pro transkripci a její regulaci rozhodující. Tato doktorská práce se zabývá dvěma projekty adresovanými k interakci RNAP s nukleovými kyselinami: (i) Transkripcí modifikovaných DNA templátů a (ii) Ms1, malou RNA, v *M. smegmatis*.

(i) Zkoumali jsme vliv modifikací v malém žlábku DNA na bakteriální transkripci *in vitro*. Zjistili jsme, že transkripce modifikovaných DNA templátů je ovlivněna ve fázi iniciace transkripce a efekt modifikací je závislý na sekvenci promotoru. Dále jsme úspěšně provedli zapnutí a vypnutí transkripce *in vitro* pomocí bioortogonálních reakcí. Regulace transkripce pomocí umělých modifikací DNA má budoucnost v biotechnologiích a/nebo v medicíně.

(ii) Dalšími regulátory genové exprese jsou malé nekódující RNA. Tuto důležitou roli mají tyto molekuly jak v prokaryotech, tak v eukaryotech. Ms1 je malá RNA nalezená v mykobakteriích tvořící komplex s jádrem RNAP. Ms1 se v buňkách nachází ve velkém množství ve stacionární fázi (množství Ms1 je porovnatelné s ribozomální RNA). Zkoumali jsme její syntézu a degradaci v *M. smegmatis* a došli jsme k závěru, že vysoké množství Ms1 ve stacionární fázi je způsobeno její zvýšenou stabilitou. Dále jsme zjistili, že RNáza polynukleotid fosforyláza má vliv na degradaci Ms1. Vědomosti získané v tomto projektu mohou být využity k vývoji specifických léčiv proti mykobakteriím.

TABLE OF CONTENTS

1. LIST OF ABREVIATIONS

6mA	6-methyl adenine
A	Adenosine
AID	Activation-induced cytosine deaminase
atc	Anhydrotetracycline
B. subtilis	*Bacillus subtilis*
BER	Base excision repair
bp	Base pair
C	Cytosine
caC	5-carboxy cytosine
C^{hm}	5-hydroxymethyl cytosine
C^{Me}	5-methyl cytosine
C^{NB}	5-nitrobenzyl cytosine
C^{P}	5-propynyl cytosine
CRISPR	Clustered regularly interspaced short palindromic repeats
CuAAC	Copper-catalysed azide-alkyne cycloaddition
DNMT	DNA methyltransferases
E	RNA polymerase core
E. coli	*Escherichia coli*
EC	Elongation complex
EX	Exponential
$E\sigma$	RNA polymerase holoenzyme
fC	5-formyl cytosine
G	Guanosine
M. smegmatis	*Mycobacterium smegmatis*
mRNA	Messenger RNA
NDP	Ribonucleoside diphosphate
NMP	Ribonucleoside monophosphate
NTP	Ribonucleoside triphosphate
P_i	Inorganic phosphate

PNPase	Polynucleotide phosphorylase
RE	Restriction endonuclease
RNAP	DNA-dependent RNA polymerase
RNase	Ribonuclease
RPc	Closed promoter complex
RPo	Open promoter complex
rRNA	Ribosomal RNA
sRNA	Small non-coding RNA
ST	Stationary
T	Thymine
T. thermophilus	*Thermus thermophilus*
TBTA	tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine
TET	Ten eleven translocation
Tet	Tetracycline
TF	Transcription factor
tRNA	Transfer RNA
U	Uracil
U^{DHT}	5-{4-[1-(2,3-dihydroxypropyl)]triazolyl}-2´-deoxy uracil
U^{E}	5-ethynyl uracil
U^{hm}	5-hydroxymethyl uracil
U^{NB}	5-nitrobenzyl uracil

2. INTRODUCTION

This Doctoral Thesis deals with two types of nucleic acids: DNA and RNA. In the first project, the focus is on modified DNA templates and their effects on transcription and the second project focuses on Ms1, a small RNA (sRNA) in *M. smegmatis*. The Thesis consists of five publications.

Natural DNA modifications are epigenetic signals that are vital for the regulation of gene expression. Insights into how RNA polymerase (RNAP) interacts with and reads through DNA with such modifications are critical for our understanding of how these modifications affect gene expression. To study this, we created by PCR DNA templates bearing artificial modifications or modifications present in cells. We used different promoter sequences and *E. coli* or *B. subtilis* RNAPs and investigated transcription *in vitro*. Modifications influenced transcription at the transcription initiation level and their effect was dependent on promoter sequence (**Publications I and II**).

Next, we performed experiments with modified DNA templates to switch transcription ON or OFF by bioorthogonal reactions. One example of a bioorthogonal reaction is the conversion of nitrobenzyl U (U^{NB}) or C (C^{NB}) to hydroxymethylated U (U^{hm}) or C (C^{hm}) by irradiation at the 400 nm wavelength. Transcription switch ON was performed by irradiation of U^{NB}/C^{NB}-modified DNA templates and subsequent transcription switch OFF by phosphorylation of photocaged U^{NB}-modified DNA templates. The second example of a bioorthogonal reaction in this Thesis is a click reaction with ethynyl-uridine (U^{E}) modified DNA templates to switch transcription OFF. This study provided insight into the regulatory potential of these epigenetic marks (**Publications III and IV**).

Ms1 is a highly abundant sRNA (rivalling in amounts those of rRNA). Ms1 forms a complex with the RNAP core and it is a regulator of gene expression, enhancing survival of the cell under various types of stress. Here we showed that Ms1 was highly stable in stationary phase and it was rapidly degraded when the cells were shifted into the nutrient-rich medium. We identified and characterized the Ms1 promoter, the dynamics of Ms1 expression, and revealed the presence of a transcription factor involved in the regulation of its expression. Further, we identified an RNase, polynucleotide phosphorylase (PNPase) to have an effect on the Ms1 level *in vivo*. With recombinant PNPase, we demonstrated that it was able to degrade

Ms1 *in vitro* and identified Ms1 secondary structures that affected its stability. In summary, we provided a comprehensive characterization of how the intracellular level of Ms1 was controlled, paving the way to potential future designs altering its expression in the case of pathogenic species (**Publication V**).

These two projects (DNA modifications and Ms1) deal with basic properties of RNAP and its interactions with DNA or RNA. The DNA modifications project sheds light on artificially modified DNA transcription and the results can be utilized in future *in vivo* approaches. The Ms1 project deals with an sRNA in *M. smegmatis,* which is present also in pathogenic species such as *M. tuberculosis* or *M. bovis*. Understanding the mechanism of action of Ms1 can lead to the development of new drugs against these species.

3. LITERARY REVIEW

3.1 Gene expression

Living cells have to adapt to environmental changes. It means that they have to change expression of their genes to survive. Gene expression entails synthesis of RNA from the DNA template and subsequent synthesis of a protein from the RNA template. The central dogma of molecular biology postulates that DNA is either replicated to a new copy of DNA or transcribed to RNA. RNA can be either reversibly transcribed to DNA or translated to proteins (Snyder et al., 2013). When the central dogma of molecular biology was first proposed, more steps were predicted and the following decades demonstrated that only some of these predictions were confirmed in cells (Figure 1, Crick, 1970).

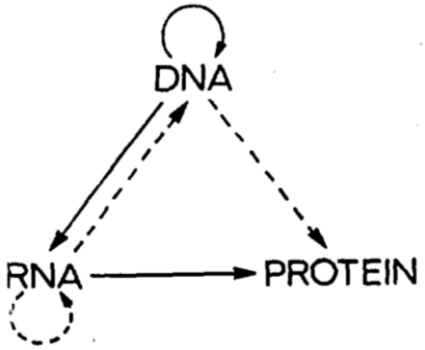


Figure 1. Scheme of the central dogma of molecular biology (Crick, 1970)

3.2 Bacterial transcription

Transcription is a level in gene expression which is regulated by many mechanisms. It can be divided into three steps – initiation, elongation, and termination (Mustaev et al., 2017; Ray-Soni et al., 2016; Whipple and Sonenshein, 1992).

In transcription initiation, the DNA-dependent RNA polymerase (RNAP) core binds to a σ factor and forms RNAP holoenzyme (Eσ). Eσ recognizes promoter sequence and forms the closed promoter complex (RPc) with DNA. The RNAP leading edge then starts to move downstream, while the trailing edge stays bound to the promoter, which causes unwinding of the double-stranded DNA. This process is called scrunching and it helps to create the open promoter complex (RPo). After binding nucleoside triphosphates (NTPs), RNAP makes an initiation complex (IC). In some cases, RNAP can before escaping the promoter perform so-

called abortive transcription, which usually occurs performed in several rounds and short products (approximately 6 nucleotides) are released (Winkelman and Gourse, 2017). After transcription of ~12 nucleotides, the σ factor is released and RNAP escapes from the promoter. Promoter escape causes conformational changes resulting in the transcription elongation complex (EC) formation. After reaching the transcription termination signal, RNAP is released from the DNA template and it can be recycled for a new round of transcription (Figure 2)(Ma et al., 2016). Eukaryotic and prokaryotic transcription systems differ in their composition and mechanisms of regulation.

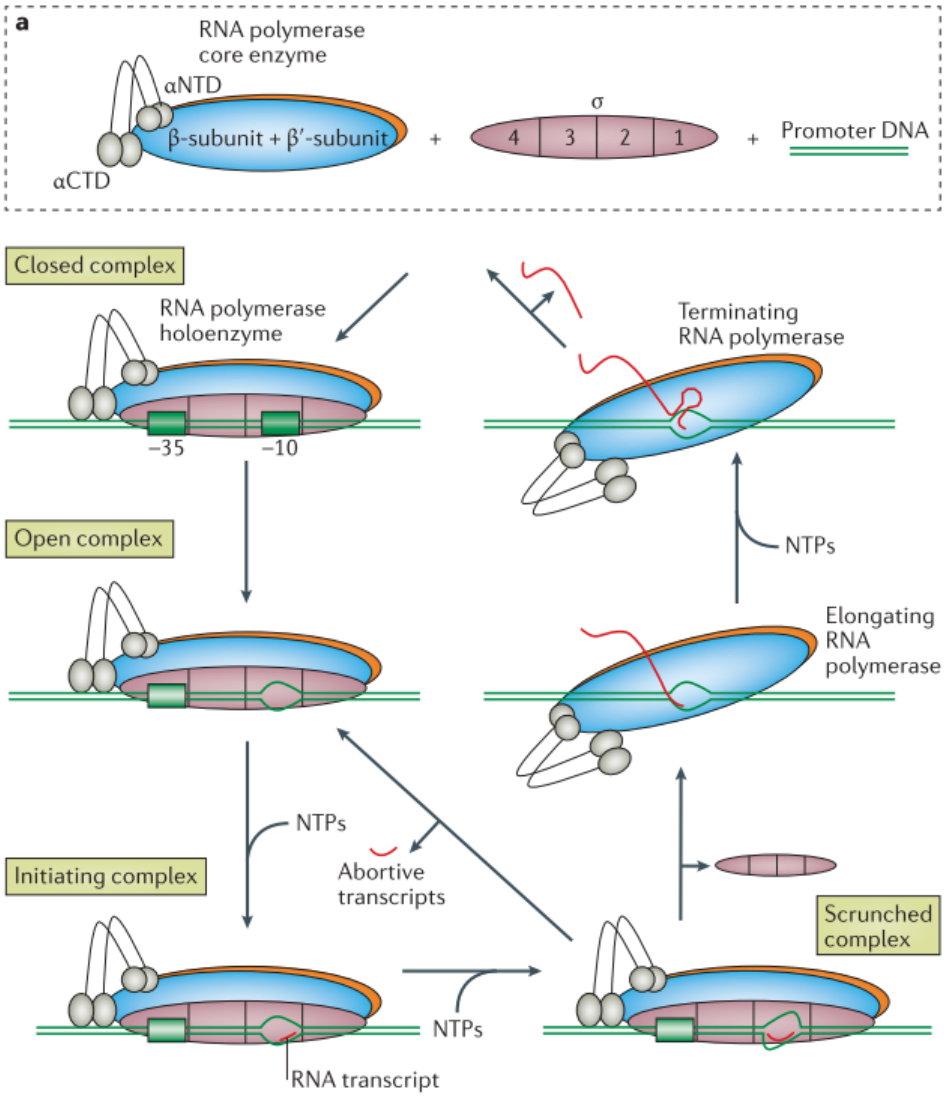


Figure 2. Bacterial transcription cycle. The bacterial transcription cycle. Eσ (RNAP core + sigma factor) recognizes promoter sequence and forms the closed promoter complex (RPc). The RPc transitions to the open promoter complex (RPo) by unwinding the DNA duplex in the region of -10 promoter sequence to the transcription start site (TSS). The addition of NTPs enables the transition to the initiation complex (IC), which is responsible for the RNA transcript synthesis. The template strand of the DNA is pulled into the IC, this process is known as 'scrunching'. The scrunched complex can be either held at the promoter, which results in cycles of abortive transcription that produce small RNA fragments, or the RNAP can escape the promoter. After promoter escape, RNAP enters elongation phase, which leads to the release of the σ factor and elongation of RNA. Transcription proceeds until RNAP reaches a terminator. The RNA transcript is released afterwards and RNAP dissociates from the DNA template (Browning and Busby, 2016).

3.2.1 RNA structure

There are several types of RNAs in bacteria: messenger mRNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and small non-coding RNA (sRNA). mRNA, rRNA, tRNA are involved in protein synthesis, sRNAs are involved in regulation, replication, and protein secretion. RNA is built from ribonucleoside triphosphates (NTP) that bear four bases guanine, cytidine, adenine, and uridine. Uridine is the difference (with respect to the composition of bases) between RNA and DNA because DNA bears mostly thymine instead of uridine. RNA bases can be modified after incorporation into the RNA chain and/or non-canonical nucleotides can be used by RNAP (Barvík et al., 2017). RNA is synthesized from the 5´- to the 3´-end. The primary structure of RNA is its sequence of nucleotides. RNA can be single-stranded; however, the pairing of RNA in the complementary regions and their folding to double-stranded structures is frequently observed too. These double-stranded structures (e. g. hairpins) represent the secondary structure of the RNA. There are also tertiary structures of RNA. Unpaired structures of secondary structure (loop of a hairpin) pairs with another unpaired region of the same RNA molecule and create so-called pseudoknot (Snyder et al., 2013).

3.2.2 Bacterial DNA-dependent RNA polymerase

The key enzyme of transcription is RNAP. It is a multi-subunit enzyme and its catalytically competent bacterial core is composed of α_2, β, β´, ω subunits. For successful transcription initiation, the interaction between the RNAP core and a σ factor is necessary (Browning and Busby, 2016). There are also other subunits, depending on the species, e. g. δ and ε in *B. subtilis* (Keller et al., 2014; Rabatinová et al., 2013), interaction partners, e. g. HelD in *B. subtilis* (Wiedermannová et al., 2014), and small molecules binding to RNAP, e. g. ppGpp, a regulator of stress response in bacteria (Murakami et al., 2015).

In bacteria, one type of RNAP is responsible for transcription of all RNAs, including mRNA, rRNA, tRNA and sRNA. The size of the whole enzyme is approximately 400 kDa. RNAP is responsible for RNA chain synthesis from nucleotide triphosphates (NTPs) by attaching of 5´phosphate of a ribonucleotide to the 3´hydroxyl group of the preceding nucleotide in the nascent chain.

The RNAP structure has been studied by crystallization and subsequent X-rays, by nuclear magnetic resonance, or by cryo-EM (Murakami et al., 2015; Narayanan et al., 2018; Vassylyev et al., 2002).

The first studied and structurally described bacterial RNAP core was from the thermophilic bacterium *Thermus thermophilus* (Tsuji et al., 1976). The RNAP has a form of a crab claw. The claw consists of two pincers: β and β´subunits. The β and β´ subunits assemble with N-terminal domains of the α subunit (αNTD) homodimer and form a cleft containing the active site. At the base of the cleft is located the so-called switch region serving as a hinge of claw pincers. The switch region changes from closed to open claw state (Mukhopadhyay et al., 2008).

Sequences of β and β´ subunits are highly conserved among bacteria, however, there are large inserted sequences in some evolutionary lineages of bacteria. These insertions are highly stable as structures and can be isolated and crystallized, their position is usually on the protein surface. Each α subunit has independently folded N-terminal and C-terminal domains joined by a flexible linker. The C-terminal domain of α subunit (αCTD) of *E. coli* RNAP regulates transcription by interaction with transcription factors and binds to upstream promoter sequence (Figure 3.) (Murakami et al., 2015).

The primary channel of RNAP, where DNA binds, is cleft-shaped and is formed by β and β´subunits. Its structure is conserved. The cleft is mostly positively charged, while the RNAP surface has an overall negative charge. Mg^{2+} catalyzes nucleotide addition to the RNA chain (Steitz et al., 1994) and may contribute to DNA melting. During transcription, NTP substrates enter the active site through the secondary channel. The secondary channel of RNAP is formed by β´ subunit and it is approximately 12 Å in diameter and funnel-shaped. It is so narrow that it allows passage only one NTP or sRNAs called nanoRNAs (oligonucleotides 2- to ~ 5 nt long) (Nickels and Dove, 2011), not double-stranded nucleic acids. The ω subunit is mainly a chaperon of the β´ subunit (Ma et al., 2016). The bacterial and eukaryotic RNAP cores are structurally and functionally homologous.

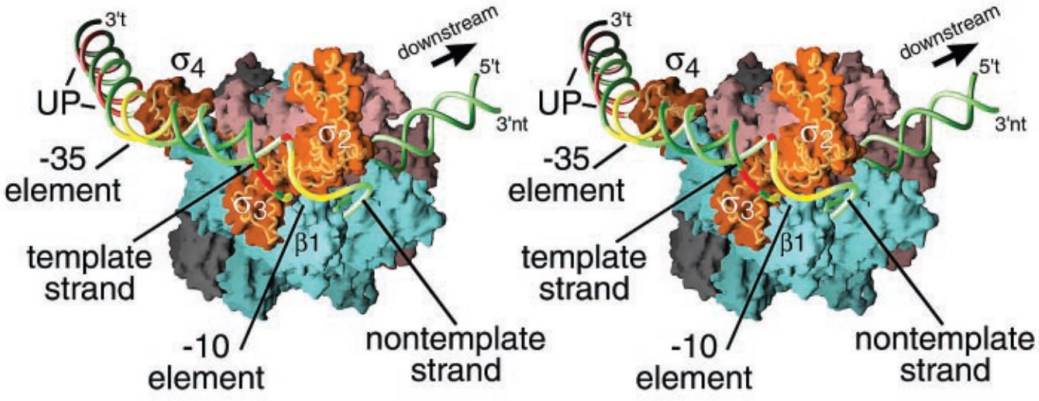


Figure 3. Structure of bacterial RNA polymerase holoenzyme interacting with DNA template. The Eσ is shown as a molecular surface, colour coding is: α_2 and ω grey; β cyan; β´ pink; and σ, orange. The DNA phosphate is shown with the template strand (t) dark green and the non-template strand (nt) light green, except for the promoter sequence (–35 and –10 elements, which are coloured yellow, and the UP elements and extended –10 element, which are in red) (Murakami et al., 2002)

3.2.3 Bacterial promoters

Transcription can start only from certain sites of DNA called promoters. Here, and in the following chapters, I will describe mainly the *E. coli* model system form, which most of the knowledge about transcription has been derived. Bacterial promoters contain several sequence motifs: the -35 element (the numbering applies to the primary/primary-like σ factors), the extended -10 element, the -10 element, and the discriminator region, which are recognized by the σ factor; and the UP element, which is recognized by αCTD (α C-terminal domain) (Figure 4, Browning and Busby, 2016). The -10 motif is AT-rich to enable easier unwinding of double-stranded DNA and formation of the transcription bubble. In some cases, the promoters can be organized in tandem to increase the effectivity of transcription. We can find them, e. g. upstream of genes for rRNA or 6S RNA (Brosius et al., 1981).

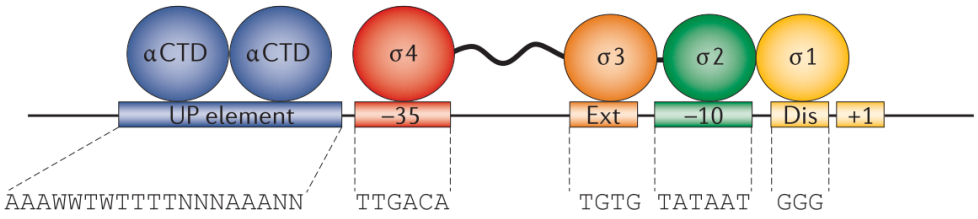


Figure 4. Interaction of RNAP holoenzyme with a promoter. Promoter sequence with highlighted important elements, their typical sequences, and αCTD of RNAP and primary-like σ factor domains interaction with promoter elements are shown. UP element, the –35 element (positions –35 to –30), the extended –10 element, the –10 element and the discriminator element are highlighted (Browning and Busby, 2016)

3.2.4 σ factors

A σ factor makes a complex with the RNAP core. This complex is called RNAP holoenzyme and is capable to recognize a promoter sequence and initiate transcription. Many different σ factors are present in bacterial cells. Different σ factors are then able to recognize different promotor sequences and so regulate transcription of genes in response to changing conditions. The cells contain one main σ factor regulating the majority of genes: σ^A in *B. subtilis* or *M. smegmatis* (Gram-positive bacteria) and σ^{70} in *E. coli* (Gram-negative bacteria) and several alternative σ factors. The number of alternative σ factors differs among bacterial species. Seven σ factors are present in *E. coli* (Maeda et al., 2000). In *M. smegmatis*, 26 σ factors are predicted (Waagmeester et al., 2005). *B. subtilis* contains 18 σ factors (Haldenwang, 1995; Helmann, 2016; MacLellan et al., 2008; Matsumoto et al., 2005; Nicolas et al., 2012; Zuber et al., 2001)and 1 σ-like factor (Gruber and Gross, 2003). Every σ factor has a set of genes under its control called regulon. Genes with connected functions are usually under the control of one σ factor and are transcribed together in one long transcript. These sets of genes are called operons or cistrons.

Sigma factors can be divided into four groups according to the presence or absence of four conserved regions (R1.1, R1.2–2.4, R3.0–3.2, R4.1–4.2) (Figure 5). These regions are structured helical domains ($\sigma_{1.1}$, σ_2, σ_3, σ_4). σ_2, σ_3, σ_4 domains interact with the promoter sequence and the RNAP core. The most conserved domain is σ_2 (R1.2–2.4) and it interacts with the β' subunit of RNAP. Regions R2.3–2.4 make base-specific interactions with the -10 hexamer of the non-template DNA strand during the DNA unwinding process. They stabilize RPo. σR1.2 in Group 1 and 2 consists of two α helices oriented 90° to each other that interact

with the non-template strand “discriminator” element of the promoter (Paget et al., 2015). The discriminator element is located between the -10 element and the transcription start site (TSS). Domain σ_3 consists of a three-helix bundle, which interacts with the major groove of duplex DNA upstream of the -10 element (Mitchell et al., 2003). Interactions with these “extended -10” elements ($T_{-15}G_{-14}$ in *E. coli*) can stabilize initiation complexes so strongly that the -35 element, which is otherwise crucial for promoter recognition, is not required. Promoters of some σ factors (e. g. σ^I from *B. subtilis*) contain also extended -35 region (Ramaniuk et al., 2018). Domain σ_4 (σR4.1–4.2) is composed of four helices with the third and fourth forming a helix-turn-helix motif binding to the -35 element. Domain σ_4 forms a contact point for transcriptional activators that bind DNA upstream of the -35 element. Domain $\sigma_{1.1}$ (σR.1.1) is present only in Group 1 and promotes a compact form of free σ, inhibiting DNA binding in the absence of the RNAP core. $\sigma_{1.1}$ is negatively charged and it mimics DNA upon Eσ formation (Zachrdla et al., 2017). It occludes the RNAP active site, however, it is displaced by the DNA duplex during RPo formation (Paget et al., 2015). Some σ factors from Group 1 contain a non-conserved region between σR1.2 and σR2.1 too. It has a variable length and composition and it has been implicated in core binding and promoter escape in the case of *E. coli* σ^{70} (Leibman and Hochschild, 2007).

Alternative σ factors vary in size from 20 kDa (Group 4) to 70 kDa (Group 1). All members of the σ^{70} family include the σ_2 and σ_4 domains that bear major RNAP- and promoter-binding determinants. Alternative σ factors lack $\sigma_{1.1}$, the presence of σ_3 is variable, and they differ in specificity and in some aspects of transcription initiation. Alternative σ factors vary hugely among bacteria, starting with their range of functions, sensing and responding to different extracellular and intracellular signals and ending with stresses. Alternative σ factors´ active concentrations are controlled by many mechanisms. These mechanisms may act on transcription, translation, or protein turnover levels. Some σ factors

are activated by proteolysis (removal of inhibitory extension) and negative regulation by anti-σ factors binding to σ factors was described too (Paget et al., 2015).

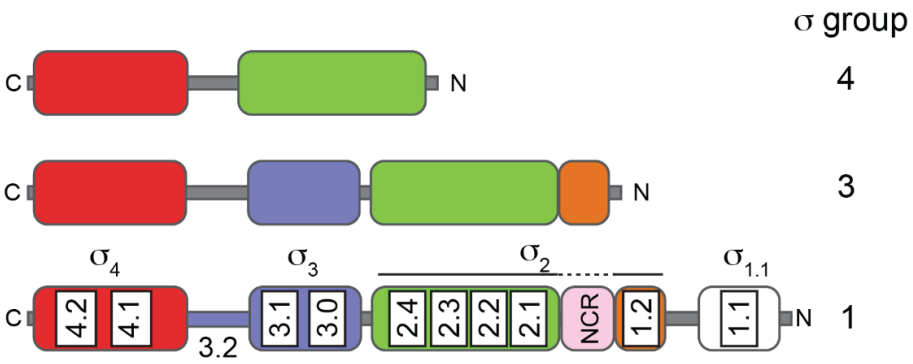


Figure 5. Schemes of groups of σ factors. Organization of σ factor domains from Groups 1, 3 and 4. Structural domains are in these colours: σ1.1, in white; σ2, in green or orange; σ3, in blue; σ4, in red. Each domain and conserved σ regions are indicated for Group 1 σ. The non-conserved region (NCR) is in pink (Paget et al., 2015).

3.2.5 Bacterial transcription initiation

Transcription initiation is an essential step in gene expression regulation and it is most thoroughly described for bacterial cells (Saecker et al., 2011). It is regulated by the presence of transcription factors (TFs). Eσ consisting of the RNAP core and a σ factor recognizes promoter sequence, binds it (forms RPc) and forms a transcription bubble by unwinding ~12-14 base pairs (bps) of the DNA double strand to single strands. This transcription-competent complex is RPo and it is formed when positions +1 (TSS) and +2 of the unwound template strand are able to enter the active site. In the presence of NTPs, RNAP holoenzyme starts from the +1 nucleotide directly or by abortive transcription, generating short transcripts without escaping the promoter. After the promoter escape, Eσ releases the σ factor and the elongation complex is created (Goldman et al., 2009; Murakami et al., 2002). In gram-positive bacterium *B. subtilis* and gram-negative bacterium *E. coli* was found that initiating NTPs are regulators of transcription initiation of a number of promoters (Krásný et al., 2008; Murray et al., 2003; Sojka et al., 2011; Walker et al., 2004).

3.2.6 Bacterial transcription elongation

The next step of transcription is transcription elongation. The nascent RNA remains base-paired with the template DNA strand forming a 9-10 bp RNA: DNA hybrid. The non-template strand is separated from the template strand. The melted region of DNA strands is 10-12 nt long and forms the transcription bubble (Turtola and Belogurov, 2016; Zaychikov et al., 1995).

Transcription can be regulated during elongation and the transcription rate is not uniform. Transcription pausing is utilized by mechanisms such as attenuation or phage λ protein Q anti-terminator (Kingston and Chamberlin, 1981; Yarnell and Roberts, 1992). Some pausing may be short and resolved spontaneously, whereas in other cases it may lead to EC backtracking. Backtracked EC can be restarted by GreA and GreB factors (Toulmé et al., 2000), which cleave from 3'- end of the nascent transcript to the active centre and unblock the active centre to let transcription continue. Elongation rate determined by pausing is influenced by both template sequence and RNA secondary structure (Mustaev et al., 2017).

3.2.7 Bacterial transcription termination

Termination sites are present downstream of operons and single transcribed genes and termination of bacterial transcription requires termination signal. There are two different types of bacterial transcription termination: Factor-independent and Factor-dependent (Peters et al., 2011).

A typical factor-independent terminator site consists of two parts: a GC rich inverted repeat, which gives rise to a hairpin after being transcribed. The hairpin is followed by a short stretch Us. The Us make RNAP to pause and subsequently release the DNA template (Santangelo and Artsimovitch, 2011). Some elongation factors, e. g. NusA can be involved in transcription termination process (Mondal et al., 2017).

The major termination factor in *E. coli* is the ρ factor. The ρ-dependent termination sequence in many RNAs has very little in common. The ρ factor can be found in most bacterial species. The ρ factor usually terminates transcription of untranslated RNAs and it has two main properties: (i) The ρ factor is an RNA-dependent ATPase using ATP as an energy source, however, its activity depends on the presence of RNA. (ii) ρ is RNA-DNA helicase unwinding double helix formed by RNA in one strand and DNA in the second strand (Snyder et al., 2013).

3.2.8 Bacterial transcription inhibitors

Inhibitors of transcription can be divided into several groups: Primary channel inhibitors, mobile elements of the primary channel inhibitors, secondary channel inhibitors, switch region inhibitors and transcription factors inhibitors. Targets of some transcription inhibitors are not known yet. Currently, two RNAP inhibitors are used in clinical practice as antibacterial drugs, rifampin (also known as rifampicin) (Strydom et al., 2019), belonging to primary channel inhibitors, and fidaxomicin (also known as DIFICID® and lipiarmycin), belonging to switch region inhibitors. Rifamycins were the first group antimicrobial drugs targeting RNAP, discovered in *Amycolatopsis mediterranei*. Fidaxomicin is a derivative from fermentation products of *Dactylosporangium aurantiacum*. It is used for the treatment of *Clostridium difficile*-associated diarrhoea (Ma et al., 2016).

3.3 Epigenetics

Epigenetics was defined by Conrad Waddington in early 1940s as “the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being“. It is accepted as “the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence“ now. Currently described modifications comprise covalent modifications of DNA bases, and posttranslational modifications of amino acids (aa) on the amino-terminal tail of histones. Enzymes modifying DNA bases or histone aa are called “writers” (e. g. DNA methyltransferases [DNMT], histone lysine methyltransferases, protein arginine methyltransferases or acetyltransferases [Ernst et al., 2010; Lei et al., 1996]). Proteins recognizing these modifications are called “readers” (e. g. methyl-CpG binding proteins [Moore et al., 2013]), and enzymes removing these modifications are called “erasers” (e. g. Ten-eleven translocation [TET] proteins [Tahiliani et al., 2009]).

Histones are proteins and several variants of them are known: H1, H2A, H2B, H3, and H4. Octamers of histones spooled around by DNA are fundamental blocks of chromatin called nucleosomes. Each histone octamer contains two units of these variants: H2A, H2B, H3, and H4. A variety of enzymes modify histones post-transcriptionally on specific serine, lysine and arginine residues on the amino-terminal tail. The best-characterized histone modifications

are methylations and acetylations of lysine residues in histones H3 and H4. Histone acetylation leads to transcription activation and methylation to both transcription activation and inhibition (Dupont et al., 2009).

3.3.1 DNA modifications

DNA is originally synthesized from four natural types of nucleotides: adenosine, guanosine, thymine, and cytosine. However, these nucleotides can be postreplicationally changed, or modified NTPs can be incorporated into DNA during replication. DNA modifications can be natural (present in cells) or artificial (Chen et al., 2016; Gramlich et al., 2008).

3.3.2 5-methyl cytosine

Methylation of the cytosines 5th carbon is the best-studied DNA modification. In eukaryotes, this epigenetic mark is involved in gene regulation and cell differentiation (Figure 6). 5-methyl cytosine (C^{Me}) represents about 1 % of all DNA bases and this methylation is almost exclusively present in CpG dinucleotides (Ehrlich and Wang, 1981). The majority of methylated CpG are in repetitive elements, probably serving as a defence against transposons or other parasitic elements (Goll and Bestor, 2005). Methylation highly changes in early embryogenesis causing X-chromosome inactivation or expression inhibition of imprinted genes (Reik, 2007). Methylation of CpG in promoter sequences is an important type of gene expression regulation. The methylation in CpG disables interaction of transcription factors with promoter sequence or binds methyl-CpG binding proteins and represses transcription (Bird, 2002). CpG methylation is changing even during ageing or pathological processes like tumorigenesis. DNMTs are responsible for cytosine methylation on carbon 5. There are several mutations of DNMT which cause diseases. C^{Me} is a substrate for deaminases and the product is thymine. They may cause mutations, because 5mC pairs with G, however T pairs with A. CGA can be mutated to TGA and lead to a nonsense mutation. There is even a danger of mutations at other sites than CpG nucleotides; it can be even at CpHpG, where H is A, C or T (Cooper et al., 2010). Simultaneously, $dC^{Me}TP$, can be incorporated into DNA during replication (Zauri et al., 2015).

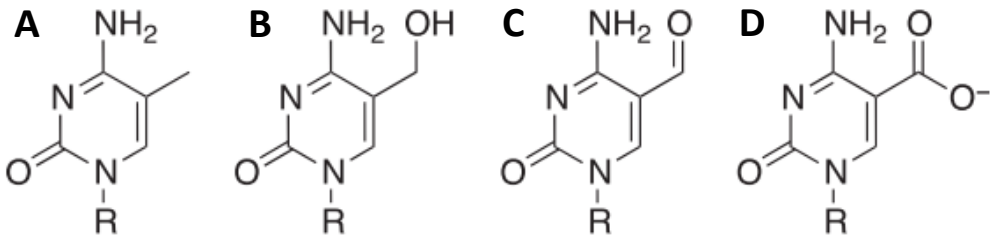


Figure 6. Schemes of (**A**) 5-methyl cytosine, (**B**) 5-hydroxymethyl cytosine, (**C**) 5-formyl cytosine, (**D**) 5-carboxy cytosine; R is a DNA backbone (Booth et al., 2013)

3.3.2.1 DNA methylation in bacteria

In bacteria, C^{Me}, 4-methyl cytosine (4mC), and 6-methyl adenine (6mA) are present in genomes. C^{Me} is present in all taxons, 6mA is present in lower eukaryotes, but not in vertebrates and 4mC is present only in bacteria. The modifications are formed postreplicativelly and they occur at specific sites. Base methylation can affect recognition of specific sites on DNA by DNA-binding proteins, replication control, DNA mismatches repair, transcription, and formation of epigenetic lineages by phase variation. C^{Me}, 4mC, and 6mA formations are catalysed by DNA methyltransferases recognizing specific DNA motifs. DNA methyltransferases transfer a methyl group from S-adenosyl-methionine to DNA. All these described modifications are major-groove oriented. In the major groove, DNA motifs are usually recognised by DNA-binding proteins (Sanchez-Romero et al., 2015).

3.3.2.2 Restriction-modification systems

A restriction-modification system (RMS) is usually described as a primitive immune system protecting the host (bacterium) against phages; however, additional roles as an epigenetic role in gene expression have been described. RMSs consist of DNA (adenine or cytosine) methyltransferase and a restriction endonuclease (RE). In the majority of restriction-modification systems, base methylation serves as protection of host DNA by prevention of DNA cleavage by restriction endonucleases. However, the activity of some restriction endonucleases on modified DNA has been described too. Numbers of RMSs differ even among strains, e. g. 15 – 20 in *Neisseria gonorrhoeae*, and more than 25 in *Helicobacter pylori* (Sanchez-Romero et al., 2015).

3.3.3 5-hydroxymethyl cytosine

In eukaryotes, 5-hydroxymethyl cytosine (C^{hm}) is an intermediate of the demethylation pathway and it is produced by TET (ten eleven translocation) proteins. TET proteins are α-ketoglutarate-dependent dioxygenases, which are responsible for the conversion of C^{Me} to C^{hm} and then to 5-formyl cytosine (fC) and 5-carboxy cytosine (caC) (Figure 6) (Tahiliani et al., 2009). These modifications can be present at as deoxynucleoside triphosphates by reaction of $dC^{Me}TP$ with TET proteins (Zauri et al., 2015). 5fC and 5caC can be removed by thymine DNA glycosylase during base excision repair (BER) and then the non-modified cytosine can be introduced. The biological role of fC and caC is not known. C^{hm} is involved in many processes: transcription, pluripotency, differentiation, development, tumorigenesis, and metastasis. However, its biological relevance is not fully understood yet (Klungland and Robertson, 2017).

Some bacteriophages incorporate $dC^{hm}TP$ into DNA, subsequently, the glucose moiety is loaded onto the base and a highly α- or β- glycosylated phage DNA is created (Sinsheimer, 1954).

3.3.4 5-hydroxymethyl uracil

5-hydroxymethyl uracil (U^{hm}) is present in DNA of some phages (Figure 7), e. g. in SPO1 phage of *B. subtilis* and dinoflagellates (Goodrich-blair and David, 1994; Stewart et al., 2009). It can also occur as a transmutation product of tritiated T in DNA sequence. Finally, exogenous $dU^{hm}TP$ can be accepted by DNA polymerase and incorporated into DNA during semiconservative DNA replication (Herrala and Vilpo, 1989). U^{hm} is present in eukaryotic cells. It can be created by the reaction of T with reactive oxygen species, conversion of C^{Me} by TET enzymes, or by deamination of C^{hm} and cause mispairing with A (Pfaffeneder et al., 2014).

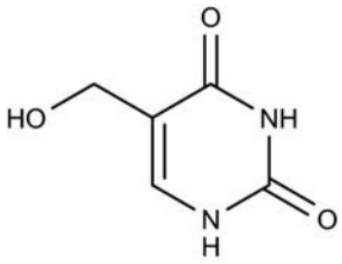


Figure 7. Scheme of 5-hydroxymethyl uracil (Schormann et al., 2014)

3.3.5 Uracil

In eukaryotes, U (Figure 8) modifications in DNA are produced by C reaction catalyzed by enzyme activation-induced cytosine deaminase (AID). AID is expressed in antigen-activated B cells. This is the basis of the phenomenon called antigen-dependent antibody diversification which occurs in two stages: somatic hypermutation and class switch. This leads to U/G mispairing and subsequently, to a generation of antibodies with different effector functions. U/G mispairs are repaired to C/Gs by uracil DNA glycosylase in BER (Bregenhorn et al., 2016). U was found also in the genome of bacteriophage AR9 of *B. subtilis* (Sokolova et al., 2017)

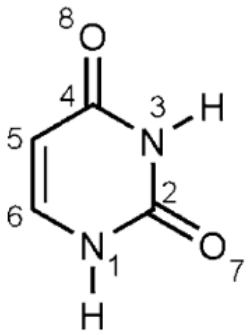


Figure 8. Scheme of uracil with atom´s labelling (Madzharova et al., 2016)

3.4 Artificial DNA modifications and bioorthogonal reactions

The reactivity-based bioorthogonal chemistry approach has been developed to study biomolecules such as nucleic acids, glycans, lipids, small molecule metabolites, and posttranscriptional modifications in living systems (Devaraj, 2018). It is a two-step approach: 1) small reporters such as aldehyde, azide, alkyne or alkene are incorporated into the molecules of interest. 2) bioorthogonal reactions are done *in situ* to perform selective ligation of probes bearing the cognate reactive groups with the pre-tagged biomolecules (step 1) of interest.

This brings new insights into many biological processes, e. g. , glycome imaging, protein lipidation, and lipid trafficking, and activity-based protein profiling. Many bioorthogonal chemistry-based methods have been developed this way (Ramil and Lin, 2013).

3.4.1 Copper-catalyzed azide-alkyne cycloaddition

Copper-catalyzed azide-alkyne cycloaddition (CuAAC), also known as “click chemistry“, is the most popular metal catalyzed reaction that is used in bioorthogonal chemistry. This reaction was first reported in 1963 by Huisgen as 1,3-dipolar cycloaddition reaction between alkyne and azide (Huisgen, 1963, Figure 9). The reaction is accelerated by a Cu(I) salt and by ligands with stabilizing effects such as tris[(1-benzyl-1H-1,2,3-triazole-4-yl)methyl]amine (TBTA) and tris(3-hydroxypropyltriazolylmethyl)amine. CuAAC of 5-ethynyl uracyl (U^E) with organic azides belongs among the most obvious reactions to convert small ethynyl modifications to bulky groups. It is possible to use both nucleotides (Amblard et al., 2009) or DNA in the reaction (Gramlich et al., 2008). U^E-modified DNA is a good substrate for DNA polymerases, it enables to incorporate deoxy-5-ethynyl uracil triphosphate to DNA by PCR. Most REs have no problem to recognize restriction sites in U^E modified DNA and U^E-modified DNA is sufficient template for bacterial transcription (Macíčková-Cahová et al., 2011).

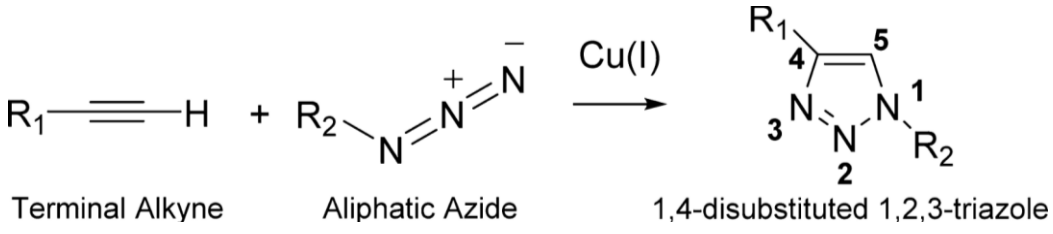


Figure 9. – Scheme of Cu(I)-catalyzed CuAAC mechanism (Castro et al., 2016)

3.5 Regulatory RNAs in bacteria

Bacterial RNAs can have several types of regulatory functions. They can modulate transcription, translation, mRNA stability, DNA maintenance or silencing. They regulate by several mechanisms such as RNA conformation changes, protein binding, base-pairing with other RNAs or interaction with DNA (Waters and Storz, 2009).

3.5.1 Riboswitches

Riboswitches are regulatory elements in the 5´end regions of mRNAs, which can react to environmental signals by conformational changes. These environmental changes can be stalled ribosomes, uncharged tRNAs, elevated temperature, or small molecule ligands (Grundy and Henkin, 2006).

3.5.2 Small non-coding RNAs

Gene expression does not have to yield a protein as its final product, but it can stop at the RNA stage, yielding sRNAs. sRNAs are 50-300 nt long and have various functions. They usually base-pair with mRNA and influence its stability or mRNA translation efficiency (Storz et al., 2011) Bacterial species such as *E. coli* encode hundreds of different sRNAs (Raghavan et al., 2011).

3.5.3 *Cis*-encoded sRNAs

Cis-encoded sRNAs are transcribed from the same region of DNA as their target RNA but in the opposite strand. Therefore, they are fully complementary to these RNAs. Most of the *cis*-encoded RNAs from bacteriophage, transposons or plasmids control a copy number of mobile elements (Brantl et al., 2007). Cis-encoded RNAs have several mechanisms of action. They are involved in premature transcription attenuation by conformational changes in DNA and subsequent terminator loop creation. Another mechanism of action is transcription inhibition by direct blocking of ribosome binding sites. Next, they inhibit maturation of primers necessary for plasmid replication, e. g. ColE1. Some of the *cis*-encoded RNAs prevent the formation of an activator RNA pseudoknot necessary for expression of some genes in plasmids (Brantl et al., 2007; Waters and Storz, 2009). Generally, *cis*-encoded RNAs expressed from bacterial chromosomes are less understood. Some of them are involved in degradation and/or repression of translation of mRNAs encoding proteins which are toxic at a certain concentration (Fozo et al., 2008).

3.5.4 *Trans*-encoded sRNAs

Trans-encoded sRNAs bind mRNAs via imperfect, short base-pairing interactions. In many aspects, they are functionally analogous to eukaryotic miRNAs (Aiba et al., 2007). They usually pair with complementary mRNAs close to the ribosome binding site and they inhibit translation, or they influence mRNA stability (Waters and Storz, 2009). In many cases, Hfq is required for regulation of *trans*-encoded RNAs. The Hfq protein is necessary for some of sRNA's function and/or stability. Hfq function is in many aspects homologous to Sm and Sm-like proteins involved in mRNA decay and splicing in eukaryotes and can remodel the secondary structure of RNA (Aiba et al., 2007).

3.5.5 CRISPR RNAs

Clustered regularly interspaced short palindromic repeats (CRISPR) are adaptable immune systems helping microbes to defend themselves against viruses by targeting their sequences by complementary RNAs. CRISPR RNAs are regulatory RNAs participating in resistance to bacteriophages and prevent plasmid conjugation (Marraffini and Sontheimer, 2008). The CRISPR sequences have been found in approximately 90% of archaea and in 40% of bacteria and they are highly variable. The DNA regions consist of a ~550 bp leader sequence connected with several repeat-spacer units. The repeat-spacer units vary from 24 to 47 nt. The same repeat appears in the CRISPR array 2 – 249 times. CRISPR-associated genes are adjacent to the CRISPR DNA array (Sorek and Hugenholtz, 2008).

3.5.6 sRNAs modulating protein activity

sRNAs can modulate protein activity either by having an intrinsic activity (RNase P), by providing essential functions to ribonucleoprotein particles (4.5S, tmRNA) or by acting as regulators by antagonizing activities of their cognate proteins by mimicking the structures of targeted nucleic acids. These sRNAs are e. g. CsrB, 6S, GlmY, and Ms1 (Hnilicová et al., 2014; Waters and Storz, 2009).

3.5.6.1 6S RNA

The first identified sRNA was 6S RNA in *E. coli*. It was found in the 1960s (Hindley, 1967) and subsequently, its homologs were found in other bacterial species, for example in Gram-positive *B. subtilis,* Gram-negative *Streptomyces coelicolor,* and many other species (Wehner et al., 2014). *B. subtilis* and other gram-positive bacteria have two variants of 6S RNA. It is processed from a longer transcript and it is highly abundant in the stationary phase. 6S RNA has a secondary structure mimicking the open promoter and RNAP holoenzyme recognizes this structure and binds it. 6S RNA makes a complex with $E\sigma^{70}$. Therefore, RNAP holoenzyme is prevented from binding to σ^{70} promoter regions. RNAP is released from 6S RNA by transcription producing a short transcript, which is 12 – 14 nt long and is called product RNA (pRNAs). It was observed especially during outgrowth from stationary phase because of the availability of nutrients and subsequently increased concentration of NTPs in cells that promote escape of RNAP from 6S RNA. The model is that $E\sigma^{70}$ is released from 6S RNA when

the concentration of NTPs increases (Wassarman and Saecker, 2006).

3.5.6.2 Ms1

Ms1 is an sRNA which was predicted *in silico* as a putative 6S RNA in *M. smegmatis* because of a similar predicted secondary structure (Figure 10) containing the central bubble which could mimick open DNA helix (Panek et al., 2011). Subsequently, the expression of Ms1 in *M. smegmatis* was verified. Interestingly, a different mechanism of RNAP binding than that of 6S RNA was found. It binds only the RNAP core, not Eσ^{70} as 6S RNA. Ms1 is 300 nt long and it is highly abundant in the stationary phase. Its abundance is comparable to ribosomal RNAs. Ms1 is conserved among mycobacteria including pathogenic species such as *Mycobacterium tuberculosis* (sRNA name MTS 2823*;* [Arnvig et al., 2011]) and *Mycobacterium leprae*. Two more proteins were found as interacting partners of Ms1 in stationary phase: ρ factor and RNAse polynucleotide phosphorylase (PNPase) (Hnilicová et al., 2014).

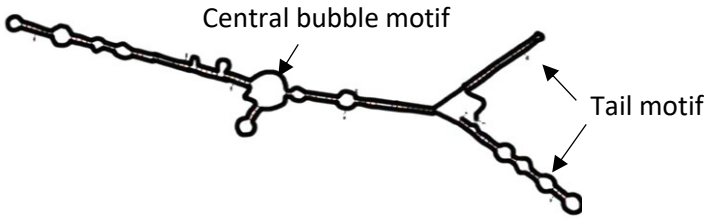


Figure 10. Ms1 secondary structure based on *in silico* prediction. Central bubble and tail motifs are indicated (Hnilicová et al., 2014)

3.6 Stability of RNA

Structured RNAs such as tRNA or rRNA are usually more stable than unstructured RNAs (mRNA). Most stable RNAs are processed from their longer forms to mature forms (Carpousis et al., 2009).

3.6.1 Degradation of RNA

RNA decay in cells is an important part of gene expression regulation. Translation is a highly energy-demanding process. Cells can save a lot of energy by degrading mRNAs whose translation is unnecessary. The process of RNA degradation is done by ribonucleases (RNases) that cleave the phosphodiester linkage of the RNA chain. They are of two basic types:

exonucleases and endonucleases. Exonucleases are responsible for cleavage of nucleoside monophosphates (NMP) or nucleoside diphosphates (NDP) from ends (3´or 5´) of the RNA chain. Endonucleases cleave inside the RNA chain, producing two shorter oligo fragments. There are many types of RNAses and their mechanism of action differ. They may cleave single- or double-stranded RNA or they cleave hydrolytically or phosphorytically etc. Hydrolytic cleavage produces NMPs and phosphorolytic cleavage produces NDPs (Carpousis et al., 2009). There is a suggestion that RNases E, J, and Y are necessary for eubacterial RNA metabolism. All known bacteria and eukaryotes contain at least one of these three ribonucleases and several of them have all three (Laalami and Putzer, 2011).

3.6.2 RNAses in *M. smegmatis*

M. smegmatis express several endonucleases and exonucleases (Table 1)

Table1. Overview of RNases predicted in *M. smegmatis* (Taverniti et al., 2011)

Endonucleases		Exonucleases	
RNase E	+	RNase J	+
RNase G	-	PNPase	+
RNase III	+	RNase T	-
Mini III	-	RNase R	-
RNase M5	-	RNase II	-
RNase P	+	RNase PH	+
RNase Z	+	YhaM	-
RNase I	-	Orn	+
RNase Y	-	NrnA	+?
RNase J	+		

3.6.2.1 Polynucleotide phosphorylase

PNPase or polyribonucleotide nucleotidyltransferase (EC 2.7.7.8) catalyzes 3´→5´phosphorolysis of polyribonucleotides. NDPs are products of phosphorolysis. PNPase is also capable to catalyze reverse template-independent 5´→3´ polymerization of NDPs and inorganic phosphate is released in this reaction.

$$RNA_n + P_i \Leftrightarrow RNA_{n-1} + NDP$$

PNPase is widely conserved among bacteria and in eukaryotic organelles that originated from bacteria such as chloroplasts and mitochondria. It is absent from *Archaea* (Lin-Chao et al., 2007). PNPase differs functionally in different organisms. The key role of PNPase is connected with mRNAs and sRNAs abundance control; Moreover, PNPase is involved in DNA metabolism too. CDP generated by RNA phosphorolysis is the precursor of both CTP and dCTP.

PNPase expression is autogenously regulated at the posttranscriptional level or at the level of its activity by diverse factors. PNPase may be part of macromolecular complexes involved in RNA degradation (e. g. degradosomes). In most analyzed species, PNPase is not an essential protein (Danchin, 1997). In PNPase mutants, phenotypes in cold shock, biofilm formation, oxidative stress, and virulence were found. PNPase catalysis of RNA reactions requires Mg^{2+}; however, it is inhibited by a high concentration of Mn^{2+} (Reiner', 1969). On the other hand, reactions with dNDPs and DNA are stimulated by Mn^{2+} (or other ions such as Fe^{2+}) and inhibited by Mg^{2+} (Cardenas et al., 2011). In bacteria, PNPase is an important part of RNA decay. It is able to process RNA molecules with 10-12 nt long single-stranded 3´-ends; Stable secondary structures block RNA degradation (Coburn and MacKie, 1998).

PNPase is a homotrimeric ring. In *M. smegmatis*, PNPase is encoded by the MSMEG_2656 gene and the protein is 763-aa long. PNPase consists of several domains: two PH, α, KH, and S1. The core of the trimeric ring is created by two RNase PH-like domains. The metal-binding site is located in the PH domain. An α-helical module separating the PH domains is located on the interior surface of the ring. The S1 and C-terminal KH domains are on the opposite side of the ring and are conformationally mobile (Figure 11). Mutations of S1 and KH domains strongly inhibit RNA dependent activities. Deletion of the S1 domain enhances DNA polymerase and DNA phosphorylase activities. Deletion of both S1 and KH domains does not affect DNA synthesis and stimulates DNA synthesis (Unciuleac and Shuman, 2013).

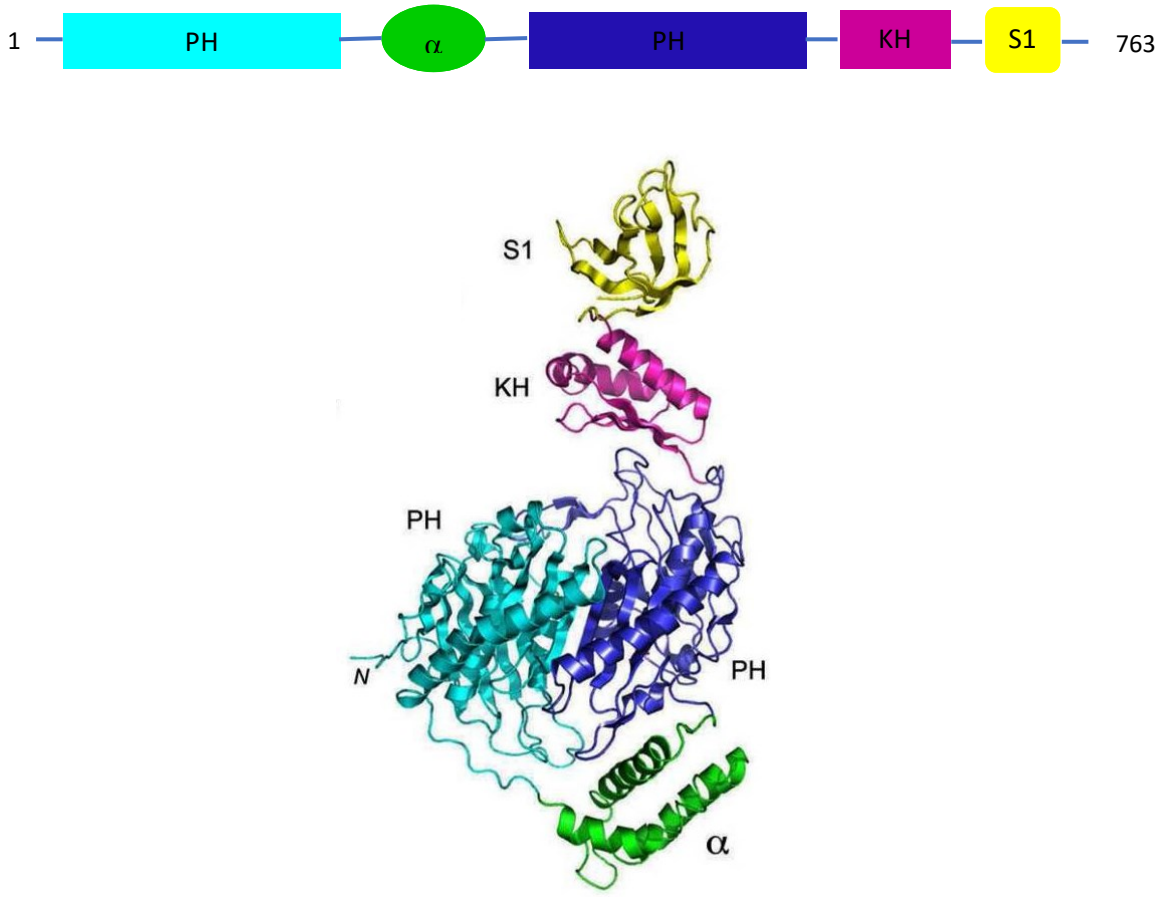


Figure 11. PNPase structure of *M. smegmatis* (**A**) primary structure, PH domains are in blue (cyan and dark blue), α domain is in green, KH domain is in magenta and S1 domain is in yellow. Numbers of aa at the ends are indicated. (**B**) The tertiary structure of monomer with colours are distinguished as in (A) (Unciuleac and Shuman, 2013).

3.6.2.2 RNase E

RNase E is an essential ribonuclease originally identified as involved in 5S rRNA processing (Apirion and Lassar, 1978). In *E. coli*, The RNase E contains 1061 aa. The first half of the enzyme has a ribonuclease activity (Mcdowall and Cohen, 1996). It is encoded by gene MSMEG_4626 in *M. smegmatis*. In *E. coli*, RNase E usually catalyzes endonucleolytic cleavage within an AU rich single-stranded region. Subsequently, the 3´- 5´exonucleolytic degradation of the upstream of the cleavage site is performed. The catalytic region is located within the N-terminus. The C-terminal part serves as a scaffold for association with PNPase, RhlB RNA helicase and the glycolytic enzyme enolase to form the degradosome complex (Carpousis et al., 2009).

Generally, RNase E proteins are found in gram-negative bacteria. They are absent from many gram-positive bacterial species such as many members Firmicutes including *B. subtilis*. These bacteria missing RNase E and contain the dual-active endo-/5´-3´exoribonuclease RNase J and/or the endonuclease RNase Y. RNase E was mostly studied in *E. coli*. It is involved in posttranscriptional metabolism of RNA, such as tRNA maturation, rRNA processing, and sRNA processing and decay and it is key enzyme initiating mRNA decay. Both mRNA decay and processing of primary tRNA transcripts contribute to RNase E´s essentiality (Deutscher, 2009).

The available structure of bacterial RNase E is from *E. coli* (Callaghan et al., 2005; Koslover et al., 2008; Marcaida et al., 2006). Identities of *E. coli* and *M. smegmatis* RNases E are only by 16% based on the alignment of protein sequences (Altschul et al., 1990). The *E. coli* RNase E N-terminal domain is conserved and composed from several subdomains (Figure 12) and C-terminal domain is not conserved and is very poorly structured. Crystallographic structures revealed that RNase E has an S1 RNA-binding domain, RNase H-like subdomains, a 5´ sensor, a DNAse I-like subdomain. The 5´sensor and RNase H-like subdomains form a pocket, where active site is (Callaghan et al., 2005).

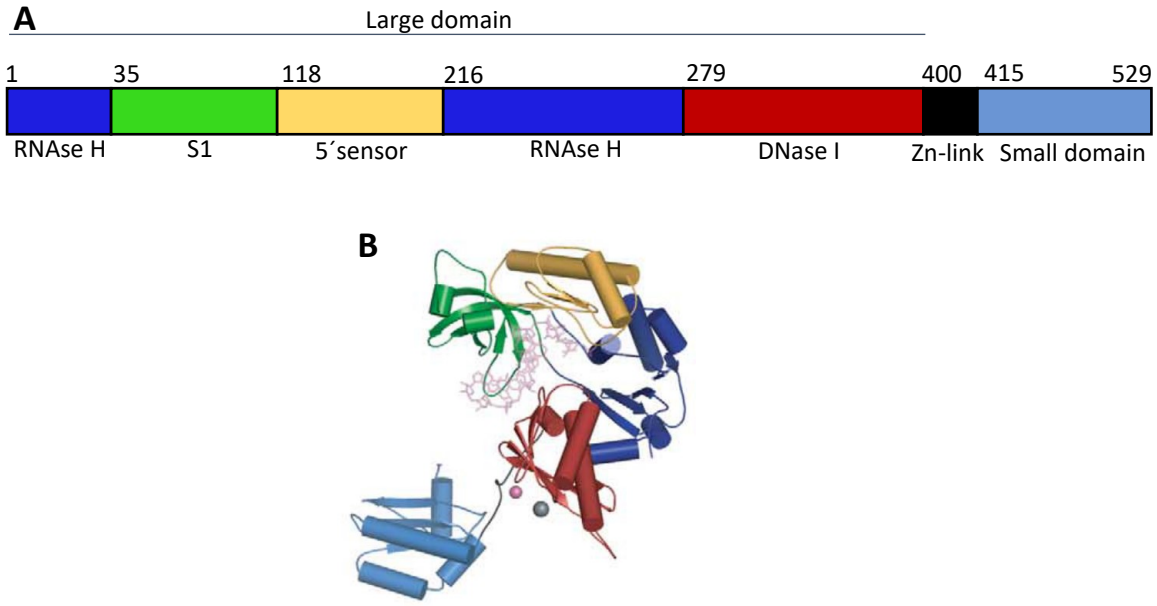


Figure 12. *E.coli* RNase E structure (N-terminal domain) (**A**) Primary structure of N-terminal domain RNase H-like subdomains in dark blue, S1 subdomain in green, 5´sensor in yellow, DNA I-like subdomain in red, Zn-link in black, and small domain in light blue. The S1 RNA recognition fold, a 5´-sensing domain, the catalytic DNase-I-like fold and a small domain that forms one of the dimer interfaces. (**B**) Crystal structure of RNase E. The Zn^{2+} and Mg^{2+} ions are shown as grey and as magenta spheres, respectively (Marcaida et al., 2006).

3.6.2.3 RNase J

RNases J1 and J2 were discovered in *B. subtilis* but orthologous enzymes are present in about half of the bacterial and archaeal genomes (Even et al., 2005). They have 5´-3´exonucleolytic activity. *M. smegmatis* contains a protein with high similarity to J1/J2 from *B. subtilis*. It is encoded by MSMEG_2685 gene. In *M. smegmatis,* it possesses both 5´-3´exonucleolytic and endoribonucleolytic activities (Taverniti et al., 2011).

RNase J has a similar function as XRN RNases in eukaryotes, which have 5´-3´exonucleolytic activity too. They are essential for fidelity control of RNA turnover in cells (Nagarajan et al., 2013).

RNAse J structure is available from bacterium *T. thermophilus* (Figure 13). The identity of *T. thermophilus* and *M. smegmatis* RNases J are 38% based on aligned protein sequences (Altschul et al., 1990). The monomer of *T. thermophilus* RNase J contains globular domains of distinct sizes, which are referred to as the β-lactamase core, β-CASP and C-terminal domains (Li De La Sierra-Gallay et al., 2008).

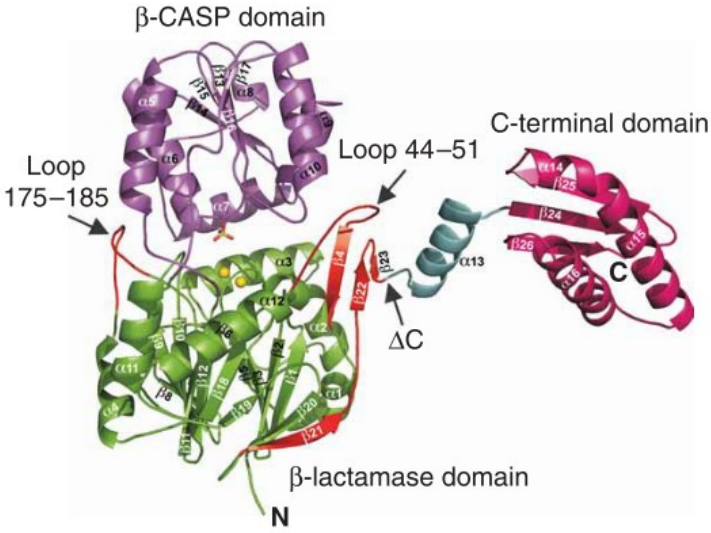


Figure 13. Structure *T. thermophilus* RNase J monomer. The β-lactamase core is in green, β-CASP domain is in purple and C-terminal domain is in magenta. The two zinc ions in the active site are shown as yellow spheres. The sulfate ion is visualised as a stick model. The N- and C-terminal ends are indicated (Li De La Sierra-Gallay et al., 2008).

3.7 Model organism - *Mycobacterium smegmatis*

M. smegmatis was found in 1885 in smegma (genital secret) with *M. wolinskyi* and *M. goodie*. These three species were considered to be one species at the time. However, more than 100 years later, these three species were recognized thanks to developed taxonomic methods such as DNA sequencing. They are called the *M. smegmatis* group now. Characteristic features of this group, which were used to distinguish them from other species were negative 3-day arylsulfatase reaction, growth at 45°C, and in the presence of 5% NaCl, positive nitrate reductase reactions, positive iron uptake, often a very smooth colony type, and utilization of mannitol, inositol, and sorbitol as carbon sources. In late growth, yellow to orange pigmentation appeared on Middlebrook 7H10 agar in 50% cases.

Approximately 50% of clinical isolates and additional ATCC reference strains of *M. smegmatis* were found in different studies and hence were renamed *M. smegmatis* sensu stricto. The mycolic acid composition of these isolates was unique and the isolates were susceptible to tobramycin agar disk diffusion. Sequence analysis of the 16S rRNA and PCR restriction analysis of the 439-bp hsp65 gene sequence was found unique to this species and used for its recognition (Brown et al., 1999). *M. smegmatis* sensu stricto has been incriminated in many cases of lymphadenitis, cellulitis, osteomyelitis, wound infections, and

rarely in respiratory diseases, usually associated with exogenous lipoid pneumonia (Wallace et al., 1988).

M. smegmatis is often used as a model organism for *M. tuberculosis* (Shiloh and Champion, 2010). *M. tuberculosis* causes tuberculosis, which is among the top 10 most deadly diseases in the world. 1.6 million people died in 2017 from tuberculosis according to information from the World Health Organisation.

Mycobacterial transformation with plasmids was performed in the late 1980s. Manipulation with the mycobacterial genome is performed mostly by recombineering or by CRISPR. Regulation of gene expression is mostly done by using the Tet (tetracycline) system inducible by anhydrotetracycline (atc) (Rock et al., 2017).

4. AIMS

The topic of this work is nucleic acid interaction with RNA polymerase. I worked on two projects. The first project was focused on modifications of DNA and their effect on transcription. The second project dealt with an sRNA found in *M. smegmatis* called Ms1.

The main goals of the modified DNA templates project were:

- To perform *in vitro* transcription assays with DNA templates bearing natural or artificial modifications
- To perform experiments dealing with transcription switch ON and OFF

The main goals of the Ms1 project were:

- To prepare an Ms1 knock-out strain and phenotype it
- To purify recombinant PNPase (interaction partner of Ms1) and perform *in vitro* digestion experiments of Ms1
- To prepare knock-down strains of selected RNases and examine their effect on the Ms1 level in *M. smegmatis* cells

5. MATERIALS AND METHODS

The study of Ms1 was performed on the model organism *M. smegmatis* mc^2 155. Plasmids were purified from *E.coli* DH5α, proteins were purified from *E. coli* DE3.

List of methods:

Bacterial cells cultivation

DNA isolation, PCR

Cloning, transformation

CRISPR/dCas9 (knock-down strains preparation)

Bacterial *in vitro* transcription assays (single- and multiple-round)

RNA isolation, reverse transcription, qPCR, Northern Blots

Electromobility shift assays

Protein purification

Western blot analysis

Digestion tests with RNAses

6. LIST OF PUBLICATIONS

Publication I

Influence of major-groove chemical modifications of DNA on transcription by bacterial RNA polymerases

Veronika Raindlová*, Martina Janoušková*, Michaela Slavíčková, Pavla Perlíková, Soňa Boháčová, Nemanja Milisavljevič, Hana Šanderová, Martin Benda, Ivan Barvík, Libor Krásný and Michal Hocek

*These authors contributed equally to the paper as first authors.

Nucleic Acids Research. 2016 Apr 20; 44(7): 3000-3012

PMID: 27001521

IF_{2017}: 11.561

IF_{2016}: 10.162

The contribution of the author: 50%; I performed all multiple-round *in vitro* transcription assays, the majority of single-round *in vitro* transcription assays of DNA templates with different bulkiness of modifications. I designed the DNA templates for electromobility shift assays and I performed all experiments with them to evaluate equilibrium K_Ds of *E. coli* RNAP and modified DNA templates of interest. I performed all the quantitation and statistical analyses.

Publication II

5-(Hydroxymethyl)uracil and -cytosine as potential epigenetic marks enhancing or inhibiting transcription with bacterial RNA polymerase

Martina Janoušková*, Zuzana Vaníková*, Fabrizia Nici, Soňa Boháčová, Dragana Vítovská, Hana Šanderová, Michal Hocek and Libor Krásný

Chem Commun (Camb). 2017 Dec 12; 53(99): 13253-13255

PMID: 29184924

IF_{2017}: 6.29

The contribution of the author: 60%; I performed all quantitation of differently labelled DNA from gels. I performed all *in vitro* transcription assays with fully and partially modifies DNA templates. I performed subsequent RNA quantitation and statistical analyses.

Publication III

Turning Off Transcription with Bacterial RNA Polymerase through CuAAC Click Reactions of DNA Containing 5-Ethynyluracil

Michaela Slavíčková, Martina Janoušková, Anna Šimonová, Hana Cahová, Milada Kambová, Hana Šanderová, Libor Krásný, and Michal Hocek

Chemistry. 2018 Jun 12; 24(33): 8311-8314

PMID: 29655191

IF_{2017}: 5.16

The contribution of the author: 40%; I performed the majority of *in vitro* transcription assays, where conditions of click reactions were tested.

Publication IV

Switching transcription with bacterial RNA polymerase through photocaging, photorelease and phosphorylation reactions in the major groove of DNA

Zuzana Vaníková*, Martina Janoušková*, Milada Kambová, Libor Krásný and Michal Hocek

Chemical Science, 2019 March 4, 10: 3937-3942

IF_{2017}: 9.063

The contribution of the author: 40%; I performed quantitation of DNA with different labels from gels. I performed multiple-round *in vitro* transcription assays with fully modified DNA templates. I performed subsequent RNA quantitation and statistical analyses.

Publication V

Ms1 RNA increases the amount of RNA polymerase in *Mycobacterium smegmatis*

Michaela Šiková*, Martina Janoušková*, Olga Ramaniuk, Petra Páleníková, Jiří Pospíšil, Pavel Bartl, Agnieszka Suder, Petr Pajer, Pavla Kubičková, Ota Pavliš, Miluše Hradilová, Dragana Vítovská, Hana Šanderová, Martin Převorovský, Jarmila Hnilicová, and Libor Krásný

Mol Microbiol. 2019 Feb;111(2):354-372

PMID: 30427073

IF_{2017}: 3.816

The contribution of the author: 40%; I prepared ΔMs1 strain and measured growth curves under different stresses. I prepared knock-down strains of genes for: Ms1, RNase E, RNase J and PNPase using CRISPR/dCas9. I performed *in vivo* tests of Ms1 levels in RNase knock-down strains. I purified PNPase with His-tag and I performed *in vitro* digestion assays with PNPase and different variants of Ms1. I performed tests of Ms1 promoter activity measuring RNA level.

I confirm that this Doctoral Thesis was written by Martina Janoušková and it is based on published data generated during her Ph.D. study in between 2014-2019. I hereby declare that her contribution described above is based on truth.

Supervisor: Mgr. Libor Krásný, Ph.D.

7. SUMMARY OF PUBLICATIONS

Publication I

Influence of major-groove chemical modifications of DNA on transcription by bacterial RNA polymerases

Publication I is a study focusing on which DNA modifications can be tolerated by RNAP and allow transcription. This and the following publications dealing with DNA modifications are results of a collaboration with the Michal Hocek´s group of Bioorganic and Medical Chemistry, where the modified DNA templates were prepared. We used a set of natural and artificial modifications of purines and pyrimidines. 7-modified 7-deazapurines and 5-modified pyrimidines were used in experiments. The modified DNA templates were full-length modified, prepared by PCR. After multiple-round *in vitro* transcription assays with *B. subtilis* and *E. coli* RNAP holoenzyme, the bulkier modifications had a stronger inhibitory effect on transcription. However, there were also some unexpected results, e. g. small modifications U or U^{V} had a strong inhibitory effect on transcription. The question was in which step transcription is inhibited. We performed single-round *in vitro* transcription assays and resolved products on sequencing gels to see shorter products of early transcription terminations caused by DNA modifications; however, no shorter products were observed with modified DNA template. Then, we focused on transcription initiation and performed EMSA to establish equilibrium K_Ds and transcription initiation was found as the critical step for transcription inhibition by DNA modifications.

Publication II

5-(Hydroxymethyl)uracil and -cytosine as potential epigenetic marks enhancing or inhibiting transcription with bacterial RNA polymerase

Publication II is a study of modifications present in cells: C^{Me}, C^{hm}, U, U^{hm}. P*veg* (promoter of *B. subtilis* – our model promoter functional with *E. coli* Eσ^{70}) and *E. coli* RNAP holoenzyme were used in *in vitro* transcription assays. Surprisingly, the C^{hm} and U^{hm} modifications had a strong stimulatory effect on transcription. We wanted to know, whether another promoter with a different sequence had the same effect as P*veg*. We used *rrn*B P1 from *B. subtilis* and the effect was opposite, both modifications C^{hm} and U^{hm}, inhibited transcription. Then, we

focused on the promoter sequence of P*veg* and we performed *in vitro* transcription assays with templates modified either in the promoter or in transcribed sequence. Transcription was stimulated when modifications were present in the promoter sequence, as expected from results published in **Publication I**. Subsequently, *in vitro* transcription assays with modifications in template or non-template strands of the promoter sequence were performed and the template strand was identified to mediate this effect.

Publication III

Turning Off Transcription with Bacterial RNA Polymerase through CuAAC Click Reactions of DNA Containing 5-Ethynyluracil

Publication III is a study of click reactions as a tool for transcription switch OFF. Click reactions required several components to proceed and we first extensively investigated what effects had these components on *in vitro* transcriptions. We performed *in vitro* transcription assays with DNA template in the presence of single components of click reactions and with their combinations. We found that the 1 molar equivalents of reaction components did not inhibit transcription. Then, using these conditions, we succeeded to switch OFF transcription.

Publication IV

Switching transcription with bacterial RNA polymerase through photocaging, photorelease and phosphorylation reactions in the major groove of DNA

Publication IV is a study of transcription of DNA with nitrobenzyl-pyrimidine modifications to first switch transcription ON by conversion of the nitrobenzyl group to hydroxymethyl group by irradiation and then to switch transcription OFF by subsequent phosphorylation of hydroxymethyl groups. Firstly, we performed a kinetic study of DNA irradiation by λ = 400 nm at different times of irradiation. We irradiated nitrobenzyl, non- and hydroxymethyl-modified DNA as DNA damage controls. The kinetic study revealed that U^{NB} and C^{NB} irradiation caused transcription increases, consistent with the release of part of the nitrobenzyl group. Irradiated U^{NB}-modified DNA transcription was on the U^{NB}-modified DNA transcription level after 20 min (350%) and C^{NB}-modified DNA transcription after 10 min (230%). It indicated that the conversion had been completed. Irradiation longer than 30 min led to a decrease of transcription, which was caused probably by DNA damage. Next experiments focused on

transcription switch ON by irradiation and subsequent switch OFF by phosphorylation of U^{hm}. We used 5-hydroxymethyl uracil DNA kinase (5-HMUDK) and ATP in reactions with irradiated samples under optimized conditions. U^{NB}- modified DNA was irradiated for 30 min at 400 nm and the transcription rose to 350%. After phosphorylation of irradiated U^{NB}-modified DNA templates (the part of nitrobenzyl group had been released by the irradiation), transcription decreased to 37% compared to positive control (non-modified DNA), so the transcription was inhibited. However, it was not blocked (original transcription of U^{NB} modified DNA templates was 10-15%) completely. In summary, we have demonstrated a transcription switch ON and subsequent switch OFF *in vitro* proof of concept for the first time.

Publication V

Ms1 RNA increases the amount of RNA polymerase in *Mycobacterium smegmatis*

Publication V is a study of Ms1, an sRNA in *M. smegmatis*. We found out that Ms1 is the most abundant non-rRNA in stationary (ST) phase and its level is 115-fold higher than in exponential (EX) phase. We prepared an ΔMs1 strain to perform phenotypic experiments and RNAseq (EX vs. ST phase). Based on RNAseq data we found out that Ms1 is the most expressed non-rRNA in stationary phase and only 6 genes changed expression in the ΔMs1 strain including Ms1. Genes for β and β´ of RNAP were among them and they were downregulated in the ΔMs1 strain and the RNAP level in the ΔMs1 strain is lower (60-80%) than in the wt strain. Subsequently, we prepared an Ms1 knock-down strain with CRISPR/dCas9 system and ΔMs1 strain with complemented Ms1 and we made phenotypic tests with them. We found out that the Ms1 gives an advantage to the cells during outgrowth. On the other hand, ΔMs1 strain cells reached higher OD_{600} at the end of exponential phase. Next, we focused on Ms1 synthesis and degradation. In Ms1´s synthesis research we verified and characterized the putative Ms1 promoter (P_{Ms1}). Its upstream sequence was found important for its higher activity. To investigate Ms1 degradation, experiments of Ms1 stability were performed and Ms1 was found to be 60-fold more stable in ST phase in comparison to EX phase. 50% of Ms1 was degraded after 18 min after dilution of culture in the ST phase to fresh media. Subsequently, knock-down strains of RNase E, RNase J and PNPase were prepared with CRISPR/dCas9 system and PNPase knock-down caused 30% increase in EX phase. Next, PNPase with 6xHis-tag was purified and digestion experiments with variants of Ms1: Ms1 wt,

Ms1 no bubble, and Ms1 no tail based on predicted secondary structure *in vitro* were performed. The tail motif was found important for digestion of Ms1 by PNPase. Finally, the result of this study is that the Ms1 abundance in ST phase is lack of degradation and Ms1 increase a pool of available RNAP in the ST phase, which gives an advantage to the cells during outgrowth.

8. DISCUSSION

My Doctoral Thesis focuses on transcription in bacteria, namely on transcription of DNA templates modified in the major groove and Ms1, an sRNA found in *M. smegmatis*. Common features of these topics are interactions of nucleic acids with RNAP. My goal in the modified DNA templates project was to characterize the ability of RNAP to accept and transcribe templates with artificial modifications, tests of modifications present in cells and transcription switch ON and OFF *in vitro* by bioorthogonal reactions (Janoušková et al., 2017; Raindlová et al., 2016; Slavickova et al., 2018; Vaníková et al., 2019). The second part of this theses deals with Ms1. Ms1 makes a complex with the RNAP core (Hnilicová et al., 2014) and I participated in a project that revealed that Ms1 had an impact on RNAP level in *M. smegmatis* cell (Šiková et al., 2019).

All mentioned publications were prepared in inter- and intra-lab collaborations. Here, I discuss topics where I performed the majority of experiments.

8.1 Influence of DNA modifications in major-groove

We performed a study on a set of artificially modified templates with a different bulkiness of major groove DNA modifications (5-modified pyrimidines and 7-modified 7-deazapurines, where was N replaced by C in position 7). The modifications displayed an increasing bulkiness (H, methyl, vinyl, ethynyl, and phenyl). It was the first study of a complete set of all four deoxyribonucleotides with these modifications. Influence of these modifications was tested by *in vitro* transcription assays with *B. subtilis* or *E. coli* $E\sigma^{A}$/$E\sigma^{70}$ on full-length modified DNA templates. The trend was mostly as expected – the bulkier modifications had a stronger inhibitory effect on transcription. Nevertheless, some exceptions were found.

The U modification had a strong inhibitory effect. However, it was previously published that dihydrouracil in the transcribed sequence was bypassed by *E. coli* RNAP (Liu and Doetsch, 1998). The explanation could be that our templates contained also modifications in the promoter sequence. Interestingly, AR9 phage of *B. subtilis*, has Us in its genome instead of thymines (Ts) and it encodes RNAP requiring U modifications at certain conserved positions in late phage promoters (Sokolova et al., 2017). Then, the absence of Ts in the -10 hexamer of the TACAAT consensus sequence of the non-template strand might

compromise interaction with σ^A region 2 (Liu et al., 2011) and thereby prevent efficient transcription initiation. U^E partially restored transcription; the ethynyl group could mimic methyl group present in the T base, whereas bulkier modifications of U (vinyl or phenyl) blocked transcription.

Interestingly, naturally present C^{Me} was partially tolerated by *E. coli* RNAP, whereas transcription with *B. subtilis* RNAP was blocked. Bulkier 5-vinyl cytosine modification was tolerated by both *E. coli* and *B. subtilis* RNAP. In previous research (Viswanathan and Doetsch, 1998), 5-propynyl cytosine (C^P, bulky modification) positioned in transcribed sequence blocked transcription, but 5-phenyl cytosine decreased transcription of full-length modified DNA templates and no shorter transcripts caused by blocking of RNAP were found. We performed single-round *in vitro* transcription assays with heparin as a competitor of DNA template to find products of transcription interruption. In (Viswanathan and Doetsch, 1998) they used heparin at 250 μg/ml, whereas we used concentration 12.5 μg/ml (Raindlová et al., 2016). Our concentration of heparin was high enough to saturate $E\sigma^{70}$ and block transcription. Hypothetically, C^P DNA lesion could cause RNAP pausing and destabilization of EC leading to premature transcription termination. EC is usually stable in the presence of heparin, however, in the case of less stable RPo it was shown that it can be destabilized by heparin (Ross and Gourse, 2009).

Generally, G (guanosine) modifications had a more pronounced effect than A (adenosine) modifications on transcription. 7-deaza A and 7-deaza G did not influence transcription with both *E. coli* and *B. subtilis* RNAP. It indicated that N at position 7 of purines is not important for RNAP holoenzyme recognition.

Transcription of modified DNA templates by *E. coli* $E\sigma^{70}$ was stronger than by *B. subtilis* $E\sigma^A$, which was possibly connected with the lower stability of open promoter complex stability of *B. subtilis* $E\sigma^A$ (Whipple and Sonenshein, 1992). Therefore, *E. coli* $E\sigma^{70}$ was used in subsequent experiments. Next, we investigated at what stage transcription was inhibited. Firstly, we examined transcription elongation, where shorter transcripts were expected as products of transcription interruption caused by DNA modifications as described in (Dimitri et al., 2008). We performed single-round *in vitro* transcription assays as described above and we found no shorter transcripts except for short transcripts of the positive control, which were most possibly products of abortive transcription (Henderson et al., 2017). To verify the

fidelity of transcription, transcripts of modified DNA templates were sequenced and no mutations were found. Subsequently, transcription initiation was examined and the equilibrium dissociation constants (K_Ds) (Matos et al., 2010) were determined. The constants are consistent with the inhibitory effect of the modifications we used in *in vitro* transcription assays, so the effect of inhibition/stimulation is at the transcription initiation level.

8.2 5-hydroxymethyl uracil and -cytosine as potential epigenetic marks enhancing or inhibiting transcription

In the next experiments, we examined four DNA modifications present naturally: C^{Me}, 5-hydroxymethyl cytosine (C^{hm}), U, and 5-hydroxymethyl uridine (U^{hm}). *In vitro* transcription assays with full-length modified DNA templates, *E. coli* RNAP holoenzyme, and four *B. subtilis* promoters: P*veg* (vegetative promoter [Fukushima et al., 2003]) or *rrnB* P1 (ribosomal promoter [Krásný and Gourse, 2004]) or their reciprocal chimaeras. U and C^{Me} modifications were used already in our previous research (Raindlová et al., 2016) and they were used as controls in transcription with P*veg*.

Interestingly, the promoter sequence had an important role with respect to DNA modifications effects. In the case of P*veg*, U^{hm} and C^{hm} modified templates transcription was strongly stimulated (cca 350%). Contrary to this was *rrn*B P1, where transcription from its modified templates was inhibited. Transcription of chimeric promoters did not reach the same transcription activity as from P*veg*, therefore, the whole P*veg* sequence was necessary for the strong positive effect on transcription. Then, we were interested in which strand of the promoter was responsible for the strong stimulatory effect. We used templates with modifications in non-template, template, both strands of the promoter sequence or both strands of transcribed sequence. The positive effect was dependent on the non-template strand of promoter sequence consistent with results from **Publication I**, and we made a prediction, which nucleotides were responsible for this effect. However, our predictions were not confirmed (changes of individual Ts for U^{hm}s in non-template strand) and the whole promoter sequence modification of the non-template strand was necessary for the strong stimulatory effect.

Modifications in our study were more frequent in our templates then they are in cells. Modifications in cells are present only at certain sites of the DNA sequence (CpG nucleotides in eukaryotes [Cooper and Krawczak, 1989], restriction sites in bacteria [Cherry, 2018], DNA lesions [Pupov et al., 2019]). Still, this research sheds light on the mechanism of how transcription is regulated with respect to promoter sequence and DNA modifications.

8.3 Transcription switch OFF by CuAAC

We investigated transcription switch OFF on ethynyl uracil (U^E) DNA templates bearing P*veg* with *E. coli* RNAP. For transcription switch OFF we used CuAAC click reactions with U^E-modified DNA templates and azide (3-azidopropane-1,2-diol) to create the corresponding dihydroxypropyltriazoles (U^{DHT}). Click reactions with PCR products (Gierlich et al., 2006) and even click reaction *in vivo* on U^E-modified RNA to detect RNA synthesis in cells had been performed previously (Jao and Salic, 2008). Nevertheless, this was the first study of transcription switch OFF by click reaction. We had already known from U^E modified DNA transcription (Raindlová et al., 2016) that it is decreased by 36% in comparison with positive control. Still, the 36% transcription is strong enough.

The reaction required several reactants to proceed (CuBr, TBTA [(1-benzyl-1H-1,2,3-triazo-4-yl)methyl]amine] and solution). We studied effects of these reagents on *in vitro* transcription. We performed *in vitro* transcription assays with combinations of click reaction components to find out which reagents had negative effects on transcription. In the beginning, we used 100 molar equivalent of T/U^E in the non-modified linear DNA template. We observed transcription inhibition in the presence of CuBr. CuBr at high concentration made probably a complex with DNA (as indicated by anomalous migration in agarose gels) and it subsequently blocked transcription. There was an investigation of Cu(I) and Cu-TBTA in eukaryotic cells and Cu-TBTA complex was found more toxic for cells than Cu alone (Kennedy et al., 2011). This is not consistent with our data, where Cu inhibits transcription, whereas Cu-TBTA combination does not. In bacteria, Cu causes cell envelope damage; there is no evidence of mutations or DNA damage at higher Cu concentration (Espírito Santo et al., 2011).

Then, we used lower equivalents of reactants and we found that 1 molar equivalent of all components to U^E moieties in modified DNA template did not affect the transcription, however, after azide addition, the click reaction was performed and transcription was

blocked. The conversion of U^{E} to U^{DHT} was only about 36%, but it was satisfactory for the transcription blocking.

However, as it was previously shown that the azide structure does not affect click reactions (Gierlich et al., 2006), we wanted to know, whether the click reaction was performed and blocked the transcription even with another azide under these conditions. We used 3-azido-7-hydroxycoumarin in the same conditions and the effect was the same: transcription with all reactants was switched OFF.

Although the click reaction requires chemicals, which are not so living cells-friendly, there is an effort to use forms of reagents, which will be suitable for *in vivo* use (Kennedy et al., 2011).

8.4 Transcription switching ON by photocaging and subsequent switching OFF by phosphorylation

This study is the first demonstration of bioorthogonal reaction in the major groove of DNA with bacterial (*E. coli*) RNAP *in vitro*. The studied process was similar to DNA methylation and its demethylation in the regulation of gene expression, which occurs naturally (Chen et al., 2016). This research was based on our previous studies, where we demonstrated that U^{hm}- and C^{hm}-modified DNA templates were transcribed better (250-350%) than non-modified DNA control (Janoušková et al., 2017).

We used U^{NB} or C^{NB} DNA modifications to convert them to U^{hm} or C^{hm} by irradiation. We used the P*veg* promoter sequence and DNA was irradiated by a 3W photodiode with a maximum λ at 400 nm, as described (Boháčová et al., 2018). 400 nm is a wavelength, which should not be harmful to cells in low doses (Ramakrishnan et al., 2016). We used sodium azide and 1,4-dithiothreitol (Barth et al., 1997) to avoid DNA damage and light absorption by nitrobenzaldehyde, which was a byproduct of nitrobenzyl modified DNA irradiation. U^{NB} and C^{NB} modifications of DNA templates can be converted to U^{hm} and C^{hm} modifications and switch transcription ON. Subsequently, U^{hm} modification was phosphorylated by 5-HMUDK (Lee et al., 2018) and this inhibited transcription, switching it OFF this way. We speculated that 5-HMUDK could protect bacteria against bacteriophages, as some phages contain U^{hm} modification naturally in their genome, e. g. SPO1 (Goodrich-blair and David, 1994). However,

further study would be necessary to confirm this hypothesis. Unfortunately, the 5-HMUDK is U^{hm} specific; transcription of C^{hm} remained switched ON after 5-HMUDK treatment. Another way of switching could be via further dephosphorylation of phosphorylated U^{hm} or glycosylation of C^{hm} (Borst and Sabatini, 2008). All experiments were performed *in vitro* and further applications *in vivo* will require future extensive research because just the introduction of modified dNTPs in cells is challenging (Guixens-Ballardo et al., 2018). Nevertheless, this study was the first step on way to get tools to switch transcription ON or OFF by artificial modifications in cells.

8.5 Synthesis and degradation of Ms1

Ms1 was found as the most abundant non-rRNA transcript in stationary phase in mycobacteria (Arnvig et al., 2011). Its abundance in ST phase was 115-fold higher than in EX phase (Šiková et al., 2019). The question was: What was the reason – higher degradation in EX phase or higher synthesis in ST phase?

To characterize the synthesis of Ms1, β-galactosidase assays were performed to examine the activity of the putative P_{Ms1} promoter with different lengths of upstream sequence (the putative core [-38/+9], -131/+9, -231/+9, -331/+9, -491/+9), a promoter with mutated -10 hexamer and upstream sequence without promoter sequence (-491/-22). The activity of the mutated putative core promoter and the activity of upstream sequence with a partially deleted putative core promoter (-491/-22) was at the level of the empty plasmid without any promoter. Therefore we identified/confirmed the core promoter sequence and we showed that the Ms1 gene had only one core promoter (P_{Ms1}), which seemed to be probably σ^A dependent according to its sequence. On the other hand, 6S RNA in *E. coli* has 2 promoters, one is σ^{70} and the second is σ^S dependent (Kim and Lee, 2004). The P_{Ms1} activity is the highest when the upstream sequence is present, so there must be some enhancer increasing transcription activity of the promoter. In our β-galactosidase assays we used the *rrnB* promoter of *M. smegmatis* as a control, however, it had a higher activity in stationary phase, which was not consistent with the literature (Aviv et al., 1996; Deutscher, 2009; Krásný and Gourse, 2004). Therefore, we decided to perform qPCR to directly measure the level of *lacZ* transcripts. These results for the *rrnB* promoter were then consistent with the literature and the *rrnB* promoter activity was the highest in EX phase and decreased in the ST phase.

With respect to the activity of control *rrnB* promoter in the literature in comparison with our results, the qPCR method was more appropriate than protein level measurement in *M. smegmatis* cells. The LacZ protein probably accumulated in cells during the growth (e. i. the protein was too stable) and the protein level did not correspond to the real activity of the P_{Ms1}. The P_{Ms1} activity was 2-fold higher with -491/+9 construct in late ST phase (48 h of growth) than in EX phase (~6 h of growth); however, the Ms1 level was 115-fold higher in ST phase. Thus, the Ms1 amount in the cells must be regulated mainly at the level of RNA stability/degradation.

To examine the degradation of Ms1, I prepared knock-down strains of RNase E, RNase J, and PNPase based on the CRISPR/dCas9 system (Rock et al., 2017). RNase E and RNase J are suggested as the most important RNases in bacteria (with RNase Y, which is not present in mycobacteria [Taverniti et al., 2011]) (Laalami and Putzer, 2011) and PNPase that was found as an interaction partner of Ms1 (Hnilicová et al., 2014). First, we showed that all three RNases had higher levels of transcripts in EX than in ST phase (8 - 30-fold), so they were promising candidates to be involved in the rapid degradation of Ms1 in EX phase. I measured the Ms1 level in knock-down strains in EX phase and only the PNPase knock-down affected the Ms1 level. The Ms1 level was approximately 30% higher after PNPase depletion compared to control strain. On the other hand, recently published data showed opposite effect of PNPase depletion in *M. tuberculosis* (Płociński et al., 2019). The MTS2823 (homolog of Ms1 in *M. tuberculosis* [Arnvig et al., 2011]) level is decreased with depletion of PNPase. The phase of growth of *M. tuberculosis* is not described in the article and the effect of PNPase depletion, e. g. in ST phase in *M. smegmatis,* has not been investigated yet. Therefore comparison of these two results is discutable. Back to our research, even though the Ms1 level in strain with depleted PNPase increased, it was not as high as in the ST phase. Therefore, other factors or RNases must participate in the rapid Ms1 degradation during EX phase. This could be addressed by knock-downs of other RNases present/predicted in *M. smegmatis* (Taverniti et al., 2011) and measurements of Ms1 levels.

To validate whether PNPase alone was able to fully digest Ms1, I expressed and purified mycobacterial PNPase with 6xHis-tag from *E. coli* and performed *in vitro* digestion tests with three forms of Ms1 coming from its predicted secondary structure – Ms1 wild type, Ms1 no tail, lacking a tail motif, and Ms1 no bubble, lacking central bubble. Ms1 wild type was

digested by PNPase, but not fully (~50%) which corresponded to the *in vivo* data. The Ms1 "no tail" was almost immune to digestion by PNPase, so the tail motif was important for PNPase and Ms1 interaction or PNPase digestion. To prove that the accumulation of Ms1 in ST phase was caused mainly by the lack of degradation, we diluted a ST phase culture to fresh media and monitored the Ms1 level just immediately after dilution. The Ms1 decay was very fast, about 50% of Ms1 accumulated during stationary growth decayed within ~18 min after dilution. The stability of Ms1 was examined too and the Ms1 half-life in ST phase was 8 h whereas in EX phase 8 min. In *E. coli*, 6S RNA (sRNA making a complex with $E\sigma^{70}$) level is regulated in EX phase by RNase BN from the RNase Z family. RNase BN has an 3´exonuclease hydrolytic activity whereas PNPase has an 5´ exonuclease phosphorolytic activity (Chen et al., 2016; Unciuleac and Shuman, 2013). RNase Z is predicted in *M. smegmatis* and in *M. tuberculosis* based on BLAST (Altschul et al., 1990; Taverniti et al., 2011); however, it has not been studied yet.

Taken together, the accumulation of Ms1 in ST phase was not caused by increased synthesis of Ms1 according to promoter activity assays. Then, the fast degradation of Ms1 after dilution of culture in the ST phase to fresh media showed that the low Ms1 level in EX phase was caused by degradation. The assays of Ms1 in knock-down strains and digestions tests *in vitro* showed that PNPase was important for Ms1 degradation, however, there must be another RNase/s or some agonist of the reaction, e. g. helicase similar to one, which is necessary for degradation of sRNAs in *E. coli* (Lin and Lin-Chao, 2005). In *M. smegmatis*, ρ factor was found as an interaction partner of Ms1 and it has helicase activity (Hnilicová et al., 2014; Walmacq et al., 2004). ρ factor could be an agonist of PNPase or another RNase cleaving Ms1.

9. CONCLUSIONS

My Doctoral Thesis brought novel insights into topics of transcription of modified DNA templates, and synthesis and degradation of Ms1.

In the modified DNA templates project, we performed a study of the influence of bulkiness of artificial and non-artificial modifications on transcription. We found their effect on transcription initiation level, and subsequently, we demonstrated the promoter sequence (when modified) had the largest effect on transcription.

Then, we did a study of CuAAC click reaction conditions and performed transcription switch OFF *in vitro*. Next, we investigated transcription switch ON by conversion of bulky modifications to smaller modifications by irradiation and subsequently, we phosphorylated these modifications to switch transcription OFF, both *in vitro*. Transcription switch ON as a proof of principle enables study/application of this approach in cells and it is under investigation in our laboratory. Next investigation under our interest is the influence of modified NTPs on transcription. Modified NTPs are present in cells and they are incorporated into RNA co-transcriptionally, however, their functions are generally not completely clear (Huang et al., 2019).

In Ms1 project, we made a study of synthesis and degradation of Ms1. We found promoter sequence of Ms1 gene and found the important upstream sequence, whose presence increases promoter activity. We found that stability in ST phase is caused by lower degradation, not by higher synthesis. PNPase, the interaction partner of Ms1, digests Ms1 and we found the secondary structure of Ms1 important for its activity. Results from this project shed light on the degradation of Ms1 in *M. smegmatis*. We made the next step in the investigation of Ms1 RNAses in mycobacteria and their research is in progress in our laboratory.

10. REFERENCES

Aiba, H., Storz, G., and Haas, D. (2007). Mechanism of RNA silencing by Hfq-binding small RNAs. Curr. Opin. Microbiol. *10*, 134–139.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic Local Alignment Search Tool. J. Mol. Biol *215*, 403–410.

Amblard, F., Cho, J.H., and Schinazi, R.F. (2009). The Cu(I)-catalyzed Huisgen azide-alkyne 1,3-dipolar cycloaddition reaction in nucleoside, nucleotide and oligonucleotide chemistry. Chem Rev. *109*, 4207–4220.

Apirion, D., and Lassar, A.B. (1978). A Conditional Lethal Mutant of Escherichia coli Which Affects the Processing of Ribosomal RNA*. J. Bol. Chem. *253*, 1738–1742.

Arnvig, K.B., Comas, I., Thomson, N.R., Houghton, J., and Boshoff, H.I. (2011). Sequence-Based Analysis Uncovers an Abundance of Non-Coding RNA in the Total Transcriptome of Mycobacterium tuberculosis. PLoS Pathog. *7*, 1–16.

Aviv, M., Giladi, H., Oppenheim ', A.B., and Glaser, G. (1996). Analysis of the shut-off of ribosomal RNA promoters in Escherichia coli upon entering the stationary phase of growth. FEMS Microbiol. Lett. *140*, 71–76.

Barth, A., Corrie, J.E.T., Gradwell, M.J., Maeda, Y., Ma, W., Meier, T., and Trentham, D.R. (1997). Time-Resolved Infrared Spectroscopy of Intermediates and Products from Photolysis of 1-(2-Nitrophenyl)ethyl Phosphates: Reaction of the 2-Nitrosoacetophenone Byproduct with Thiols. J. Am. Chem. Soc. *119*, 4149–4159.

Barvík, I., Rejman, D., Panova, N., Šanderová, H., and Krásný, L. (2017). Non-canonical transcription initiation: the expanding universe of transcription initiating substrates. FEMS Microbiol. Rev. *41*, 131–138.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6–21.

Boháčová, S., Ludvíková, L., Slavětínská, L.P., Vaníková, Z., Klán, P., and Hocek, M. (2018). Protected 5-(hydroxymethyl)uracil nucleotides bearing visible-light photocleavable groups as building blocks for polymerase synthesis of photocaged DNA †. Org. Biomol. Chem. *16*, 1527–1535.

Booth, M.J., B Ost, T.W., Beraldi, D., Bell, N.M., Branco, M.R., Reik, W., and Balasubramanian, S. (2013). Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. Nat. Protoc. *8*, 1841–1851.

Borst, P., and Sabatini, R. (2008). Base J: Discovery, Biosynthesis, and Possible Functions. Annu. Rev. Microbiol. *62*, 235–251.

Brantl, S., Storz, G., and Haas, D. (2007). Regulatory mechanisms employed by cis-encoded antisense RNAs This review comes from a themed issue on Cell regulation Edited by. Curr. Opin. Microbiol. *10*, 102–109.

Bregenhorn, S., Kallenberger, L., Artola-Borán, M., Borán, B., Peña, J., Peña-Diaz, P., and Jiricny, J. (2016). Non-canonical uracil processing in DNA gives rise to double-strand breaks and deletions: relevance to class switch recombination. Nucleic Acids Res. *44*, 2691–2705.

Brosius, J., Dull, T.J., Sleeter, D.D., and Noller, H.F. (1981). Gene Organisation and Primary Structure of a Ribosomal RNA Operon from Escherichia colli. J. Mo *148*, 107–127.

Brown, B.A., Springer, B., Steingrube, V.A., Wilson, R.W., Pfyffer, G.E., Garcia, M.J., Menendez, M.C., Rodriguez-Salgado, B., Jost, K.C., Chiu, S.H., et al. (1999). Mycobacterium wolinskyi sp. nov. and Mycobacterium goodii sp. nov., two new rapidly growing species related to Mycobacterium smegmatis and associated with human wound infections: a cooperative study from the International Working Group on Mycobacterial Ta. Int. J. Syst. Bacteriol. *49*, 1493–1511.

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

Callaghan, A.J., Marcaida, M.J., Stead, J.A., Mcdowall, K.J., Scott, W.G., and Luisi, B.F. (2005). Structure of Escherichia coli RNase E catalytic domain and implications for RNA turnover. *437*, 1187–1191.

Cardenas, P.P., Carzaniga, T., Zangrossi, S., Briani, F., Garcia-Tirado, E., Dehò, G., and Alonso, J.C. (2011). Polynucleotide phosphorylase exonuclease and polymerase activities on single-stranded DNA ends are modulated by RecN, SsbA and RecA proteins. Nucleic Acids Res. *39*, 9250–9261.

Carpousis, A.J., Luisi, B.F., and McDowall, K.J. (2009). Endonucleolytic Initiation of mRNA Decay in Escherichia coli. Prog. Mol. Biol. Transl. Sci. *85*, 91–135.

Castro, V., Rodríguez, H., and Albericio, F. (2016). CuAAC: An Efficient Click Chemistry Reaction on Solid Phase. ACS Comb. Sci. *18*, 1–14.

Chen, K., Zhao, B.S., and He, C. (2016). Nucleic Acid Modifications in Regulation of Gene Expression. Cell Chem. Biol. *23*, 74–85.

Cherry, J.L. (2018). Methylation-Induced Hypermutation in Natural Populations of Bacteria. J. Bacteriol. *200*, 1–13.

Coburn, G.A., and MacKie, G.A. (1998). Reconstitution of the degradation of the mRNA for ribosomal protein S20 with purified enzymes. J. Mol. Biol. *279*, 1061–1074.

Cooper, D.N., and Krawczak, M. (1989). Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. Hum. Genet. *83*, 181–188.

Cooper, D.N., Mort, M., Stenson, P.D., Ball, E. V, and Chuzhanova, N.A. (2010). Methylation-mediated deamination of 5-methylcytosine appears to give rise to mutations causing human inherited disease in CpNpG trinucleotides, as well as in CpG dinucleotides. Hum. Genomis *4*, 406–410.

Crick, F. (1970). Central Dogma of Molecular Biology. Nature *227*, 561–563.

Danchin, A. (1997). Comparison between the Escherichia coli and Bacillus subtilis genomes suggests that a major function of polynucleotide phosphorylase is to synthesize CDP. DNA Res. *4*, 9–18.

Deutscher, M.P. (2009). Maturation and Degradation of Ribosomal RNA in Bacteria. Prog. Mol. Biol. Transl. Sci. *85*, 369–391.

Devaraj, N.K. (2018). The Future of Bioorthogonal Chemistry. ACS Cent. Sci. *4*, 952–959.

Dimitri, A., Jia, L., Shafirovich, V., Geacintov, N.E., Broyde, S., and Scicchitano, D.A. (2008). Transcription of DNA Containing the 5-Guanidino-4-nitroimidazole Lesion by Human RNA Polymerase II and Bacteriophage T7 RNA Polymerase. DNA Repair *7*, 1276–1288.

Dupont, C., Armant, D.R., and Brenner, C.A. (2009). Epigenetics: Definition, Mechanisms and Clinical Perspective. Semin Reprod Med. *27*, 351–357.

Ehrlich, M., and Wang, R.Y.-H. (1981). 5-Methylcytosine in Eukaryotic DNA. Science. *212*, 1350–1357.

Ernst, J., Kellis, M., and Author, N.B. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. Nat Biotechnol *28*, 817–825.

Espírito Santo, C., Wen Lam, E., Elowsky, C.G., Quaranta, D., Domaille, D.W., Chang, C.J., and Grass, G. (2011). Bacterial Killing by Dry Metallic Copper Surfaces. Appl. Environ. Microbiol. *77*, 794–802.

Even, S., Pellegrini, O., Zig, L., Labas, V., Vinh, J., Bréchemmier-Baey, D., and Putzer, H. (2005). Ribonucleases J1 and J2: Two novel endoribonucleases in B.subtilis with functional homology to E.coli RNase E. Nucleic Acids Res. *33*, 2141–2152.

Fozo, E.M., Hemm, M.R., and Storz, G. (2008). Small Toxic Proteins and the Antisense RNAs That Repress Them. Microbiol. Mol. Biol. Rev. *72*, 579–589.

Fukushima, T., Ishikawai, S., Yamamoto, H., Ogasawara, N., and Sekiguchi, J. (2003). Transcriptional, Functional and Cytochemical Analyses of the veg Gene in Bacillus subtilis. J. Biochem *133*, 475–483.

Gierlich, J., Burley, G.A., Gramlich, P.M.E., Hammond, D.M., and Carell, T. (2006). Click Chemistry as a Reliable Method for the High-Density Postsynthetic Functionalization of Alkyne-Modified DNA. Org. Lett. *8*, 3639–3642.

Goldman, S.R., Ebright, R.H., and Nickels, B.E. (2009). Direct Detection of Abortive RNA Transcripts in Vivo. Science. *324*, 927–928.

Goll, M.G., and Bestor, T.H. (2005). Eukaryotic Cytosine Methyltransferases. Annu. Rev. Biochem *74*, 481–514.

Goodrich-blair, H., and David, A.S. (1994). The DNA polymerase genes of several HMU-bacteriophages have similar group I introns with highly divergent open reading frames. Nucleic Acids Res. *22*, 3715–3721.

Gramlich, P.M.E., Wirges, C.T., Manetto, A., and Carell, T. (2008). Postsynthetic DNA modification through the copper-catalyzed azide-alkyne cycloaddition reaction. Angew. Chemie - Int. Ed. *47*, 8350–8358.

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

Grundy, F.J., and Henkin, T.M. (2006). From ribosome to riboswitch: Control of gene expression in bacteria by RNA structural rearrangements. Crit. Rev. Biochem. Mol. Biol. *41*, 329–338.

Guixens-Ballardo, P., Zawada, Z., Matyašovský, J., Dziuba, D., Pohl, R., and Hocek, M. (2018). Brightly Fluorescent 2'-Deoxyribonucleoside Triphosphates Bearing Methylated Bodipy Fluorophore for in Cellulo Incorporation to DNA, Imaging, and Flow Cytometry. Bioconjugate Chem. *29*, 3906–3912.

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

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

Henderson, K.L., Felth, L.C., Molzahn, C.M., Shkel, I., Wang, S., Chhabra, M., Ruff, E.F., Bieter, L., Kraft, J.E., and Thomas Record Jr, M. (2017). Mechanism of transcription initiation and promoter escape by E. coli RNA polymerase. PNAS *114*, 3032–3040.

Herrala, A.M., and Vilpo, J.A. (1989). Template-Primer Activity of 5-(Hydroxymethyl)uracil-Containing DNA for Prokaryotic and Eukaryotic DNA and RNA Polymerases. Biochemistry *28*, 8274–8277.

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

Hnilicová, J., Jirát Matějčková, J., Šiková, M., 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.

Huang, H., Weng, H., Zhou, K., Wu, T., Simen Zhao, B., Sun, M., Chen, Z., Deng, X., Xiao, G., Auer, F., et al. (2019). Histone H3 trimethylation at lysine 36 guides m 6 A RNA modification co-transcriptionally. Nature *567*, 414–418.

Huisgen, R. (1963). 1,3-Dipolar Cycloadditions Past and Future. Angew. Chern. Interntrf. Ed. *2*, 565–632.

Janoušková, M., Vaníková, Z., Nici, F., Boháčová, S., Vítovská, D., Šanderová, H., Hocek, M., and Krásný, L. (2017). 5-(Hydroxymethyl)uracil and -cytosine as potential epigenetic marks enhancing or inhibiting transcription with bacterial RNA polymerase. Chem. Commun. *53*, 13253–13255.

Jao, C.Y., and Salic, A. (2008). Exploring RNA transcription and turnover in vivo by using click chemistry. PNAS *105*, 15779–15784.

Keller, A.N., Yang, X., Wiedermannová, J., Delumeau, O., Krásný, L., and Lewis, P.J. (2014). ε, a New Subunit of RNA Polymerase Found in Gram-Positive Bacteria. J. Bacteriol. *196*, 3622–3632.

Kennedy, D.C., Mckay, C.S., Legault, M.C.B., Danielson, D.C., Blake, J.A., Pegoraro, A.F., Stolow, A., Mester, Z., and Pezacki, J.P. (2011). Cellular Consequences of Copper Complexes Used To Catalyze Bioorthogonal Click Reactions. J. Am. Chem. Soc *133*, 17993–18001.

Kim, K.-S., and Lee, Y. (2004). Regulation of 6S RNA biogenesis by switching utilization of both sigma factors and endoribonucleases. Nucleic Acids Res. *32*, 1–12.

Kingston, R.E., and Chamberlin, M.J. (1981). Pausing and Attenuation of in Vitro Transcription in the rrnB operon of E. coli. Cell *27*, 523–531.

Klungland, A., and Robertson, A.B. (2017). Oxidized C5-methyl cytosine bases in DNA: 5-Hydroxymethylcytosine; 5-formylcytosine; and 5-carboxycytosine. Free Radic. Biol. Med. *107*, 62–68.

Koslover, D.J., Callaganm, A.J., Marcaida, M.J., Garman, E.F., Martick, M., Scott, W.G., and Luisi, B.F. (2008). The Crystal Structure of the Escherichia coli RNase E Apoprotein and a Mechanism for RNA Degradation. Structure *16*, 1238–1244.

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

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

Laalami, S., and Putzer, H. (2011). mRNA degradation and maturation in prokaryotes: the global players. BioMol Concepts *2*, 491–506.

Lee, Y.-J., Dai, N., Walsh, S.E., Muller, S., Fraser, M.E., Kaufman, K.M., Guan, C., Correa Jr., I.R., and Weigele, P.R. (2018). Identification and biosynthesis of thymidine hypermodifications in the genomic DNA of widespread bacterial viruses. PNAS *115*, 3116–3125.

Lei, H., Oh, S.P., Okano, M., Juttermann, R., Goss, K.A., Jaenish, R., and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development *122*, 3195–3205.

Leibman, M., and Hochschild, A. (2007). A r-core interaction of the RNA polymerase holoenzyme that enhances promoter escape. EMBO J. *26*, 1579–1590.

Li De La Sierra-Gallay, I., Zig, L., Jamalli, A., and Putzer, H. (2008). Structural insights into the dual activity of RNase J. Nat. Struct. Mol. Biol. *15*, 206–212.

Lin-Chao, S., Chiou, N.-T., and Schuster, G. (2007). The PNPase, exosome and RNA helicases as the building components of evolutionarily-conserved RNA degradation machines. J. Biomed. Sci. *14*, 523–532.

Lin, P.-H., and Lin-Chao, S. (2005). RhlB helicase rather than enolase is the-subunit of the Escherichia coli polynucleotide phosphorylase (PNPase)-exoribonucleolytic complex. PNAS *102*, 16590–16595.

Liu, J., and Doetsch, P.W. (1998). Escherichia coli RNA and DNA polymerase bypass of dihydrouracil: mutagenic potential via transcription and replication. Nucleic Acids Res. *26*, 1707–1712.

Liu, X., Bushnell, D.A., and Kornberg, R.D. (2011). Leading Edge Previews Lock and Key to Transcription: s-DNA Interaction How does RNA polymerase recognize a promoter in duplex DNA? How are the DNA strands pried apart to enable RNA synthesis? Cell *147*, 1218–1219.

Ma, C., Yang, X., and Lewis, P.J. (2016). Bacterial Transcription as a Target for Antibacterial Drug Development. Microbiol. Mol. Biol. Rev. *80*, 139–160.

Macíčková-Cahová, H., Pohl, R., and Hocek, M. (2011). Cleavage of Functionalized DNA Containing 5-Modified Pyrimidines by Type II Restriction Endonucleases. ChemBioChem *12*, 431–438.

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

Madzharova, F., Heiner, Z., Gühlke, M., and Kneipp, J. (2016). Surface-Enhanced Hyper-Raman Spectra of Adenine, Guanine, Cytosine, Thymine, and Uracil. J. Phys. Chem. C *120*, 15415–15423.

Maeda, H., Fujita, N., and Ishihama, A. (2000). Competition among seven Escherichia coli σ subunits: relative binding affinities to the core RNA polymerase. Nucleic Acids Res. *28*, 3497–3503.

Marcaida, M.J., Depristo, M.A., Chandran, V., Carpousis, A.J., and Luisi, B.F. (2006). The RNA degradosome: life in the fast lane of adaptive molecular evolution. TRENDS Biochem. Sci. *31*, 359–365.

Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. Science. *332*, 1843–1845.

Matos, R.G., Barbas, A., and Arraiano, C.M. (2010). Comparison of EMSA and SPR for the Characterization of RNA-RNase II Complexes. Protein J *29*, 394–397.

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

Mcdowall, K.J., and Cohen, S.N. (1996). The N-terminal Domain of the rne Gene Product has RNase E Activity and is Non-overlapping with the Arginine-rich RNA-binding Site. J. Mol. Biol *255*, 349–355.

Mitchell, J.E., Zheng, D., Busby, S.J.W., and Minchin, S.D. (2003). Identi®cation and analysis ofèxtended ±10' promoters in Escherichia coli. Nucleic Acids Res. *31*, 4689–4695.

Mondal, S., Yakhnin, A. V, Sebastian, A., Albert, I., Babitzke, P., and Author, N.M. (2017). NusA-dependent transcription termination prevents misregulation of global gene expression. Nat. Microbiol. *1*, 1–7.

Moore, L.D., Le, T., and Fan, G. (2013). DNA Methylation and Its Basic Function. Neuropsychopharmacology *38*, 23–38.

Mukhopadhyay, J., Das, K., Ismail, S., Koppstein, D., Jang, M., Hudson, B., Sarafianos, S., Tuske, S., Patel, J., Jansen, R., et al. (2008). The RNA Polymerase "Switch Region" is a Target for Inhibitors. Cell *135*, 295–307.

Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002). Structural Basis of Transcription Initiation: An RNA Polymerase Holoenzyme-DNA Complex. Science. *296*, 1285–1290.

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

Murray, H.D., Schneider, D.A., and Gourse, R.L. (2003). Control of rRNA Expression by Small Molecules Is Dynamic and Nonredundant. Mol. Cell *12*, 125–134.

Mustaev, A., Roberts, J., and Gottesman, M. (2017). Transcription elongation. Transcription *8*, 150–161.

Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. (2013). XRN 5'→3' exoribonucleases: Structure, mechanisms and functions. Biochim Biolhys Acta *1829*, 590–603.

Narayanan, A., Vago, F.S., Li, K., Zuhaib Qayyum, M., Yernool, D., Jiang, W., and Murakami, K.S. (2018). Cryo-EM structure of Escherichia coli 70 RNA polymerase and promoter DNA complex revealed a role of non-conserved region during the open complex formation. J. Biol. Chem *293*, 7367–7375.

Nickels, B.E., and Dove, S.L. (2011). NanoRNAs: a class of small RNAs that can prime transcription initiation in bacteria. J Mol Biol. *512*, 772–781.

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

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

Panek, J., Krasny, L., Bobek, J., Jezkova, E., Korelusova, J., and Vohradsky, 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.

Peters, J.M., Vangeloff, A.D., and Landick, R. (2011). Bacterial Transcription Terminators: The RNA 3'-End Chronicles. J Mol Biol. *412*, 793–813.

Pfaffeneder, T., Spada, F., Wagner, M., Brandmayr, C., Laube, S.K., Eisen, D., Truss, M., Steinbacher, J., Hackner, B., Kotljarova, O., et al. (2014). Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. Nat. Chem. Biol. *10*, 574–581.

Płociński, P., Macios, M., Houghton, J., Niemiec, E., Płocińska, R., Brzostek, A., Słomka, M., Dziadek, J., Young, D., and Dziembowski, A. (2019). Proteomic and transcriptomic experiments reveal an essential role of RNA degradosome complexes in shaping the transcriptome of *Mycobacterium tuberculosis*. Nucleic Acids Res. 1–14.

Pupov, D., Ignatov, A., Agapov, A., and Kulbachinskiy, A. (2019). Distinct effects of DNA lesions on RNA synthesis by Escherichia coli RNA polymerase. Biochem. Biophys. Res. Commun. *510*, 122–127.

Rabatinová, A., Šanderová, H., Matějčková, J.J., Korelusová, J., Sojka, L.K., Barvík, I., Papoušková, V., Sklenář, V., Žídek, L., and Krásný, L. (2013). The δ Subunit of RNA Polymerase Is Required for Rapid Changes in Gene Expression and Competitive Fitness of the Cell. J. Bacteriol. *195*, 2603–2611.

Raghavan, R., Groisman, E.A., and Ochman, H. (2011). Genome-wide detection of novel regulatory RNAs in E. coli. Genome Res. *21*, 1487–1497.

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.

Ramakrishnan, P., Maclean, M., Macgregor, S.J., Anderson, J.G., and Grant, M.H. (2016). Cytotoxic responses to 405 nm light exposure in mammalian and bacterial cells: Involvement of reactive oxygen species. Toxicol. Vitr. *33*, 54–62.

Ramaniuk, O., Převorovský, M., Pospíšil, J., Vítovská, D., Kofroňová, O., Benada, O., Schwarz, M., Šanderová, H., Hnilicová, J., and Krásný, L. (2018). I from Bacillus subtilis: Impact on Gene Expression and Characterization of I-Dependent Transcription That Requires New Types of Promoters with Extended -35 and -10 Elements. J. Bacteriol. *200*, 1–15.

Ramil, C.P., and Lin, Q. (2013). Bioorthogonal chemistry: strategies and recent development. Chem Commun (Camb). *49*, 11007–11022.

Ray-Soni, A., Bellecourt, M.J., and Landick, R. (2016). Mechanisms of Bacterial Transcription Termination: All Good Things Must End. Annu. Rev. Biochem *85*, 319–347.

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. Nature *447*, 425–432.

Reiner', A.M. (1969). Characterization of Polynucleotide Phosphorylase Mutants of Escherichia coli. J. Bacteriol. *97*, 1437–1443.

Rock, J.M., Hopkins, F.F., Chavez, A., Diallo, M., Chase, M.R., Gerrick, E.R., Pritchard, J.R., Church, G.M., Rubin, E.J., Sassetti, C.M., et al. (2017). Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform HHS Public Access. Nat. Microbiol. *2*, 1–21.

Ross, W., and Gourse, R.L. (2009). Analysis of RNA Polymerase-Promoter Complex Formation. Methods *47*, 13–24.

Saecker, R.M., Record, M.T., Dehaseth, P.L., and DeHaseth, P.L. (2011). Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. J. Mol. Biol. *412*, 754–771.

Sanchez-Romero, M.A., Cota, I., and Casadesús, J. (2015). DNA methylation in bacteria: from the methyl group to the methylome. Curr. Opin. Microbiol. *25*, 9–16.

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

Schormann, N., Ricciardi, R., and Chattopadhyay, D. (2014). Uracil-DNA glycosylases - Structural and functional perspectives on an essential family of DNA repair enzymes. Protein Sci. *23*, 1667–1685.

Shiloh, M.U., and Champion, D. (2010). To catch a killer. What can mycobacterial models teach us about Mycobacterium tuberculosis pathogenesis? Curr. Opin. Microbiol. *13*, 86–92.

Šiková, M., Janoušková, M., Ramaniuk, O., Páleníková, P., Pospíšil, J., Bartl, P., Suder, A., Pajer, P., Kubičková, P., Pavliš, O., et al. (2019). Ms1 RNA increases the amount of RNA polymerase in Mycobacterium smegmatis. Mol. Microbiol. *111*, 354–372.

Sinsheimer, R.L. (1954). Nucleotides from T2r+ bacteriophage. Science. *120*, 551–553.

Slavickova, M., Janouskova, M., Simonova, A., Cahova, H., Kambova, M., Sanderova, H., Krasny, L., and Hocek, M. (2018). Turning Off Transcription with Bacterial RNA Polymerase through CuAAC Click Reactions of DNA Containing 5-Ethynyluracil. Chem. Eur. J. *24*, 8311–8314.

Snyder, L., Peters, J.E., Henkin, T.M., and Champness, W. (2013). Bacterial Gene Expression: Transcription, Translation and Protein Folding. In Molecular Genetics of Bacteria, (Washington, DC: ASM Press), pp. 67–124.

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

Sokolova, M., Borukhov, S., Lavysh, D., Artamonova, T., Khodorkovskii, M., and Severinov, K. (2017). A non-canonical multisubunit RNA polymerase encoded by the AR9 phage recognizes the template strand of its uracil-containing promoters. Nucleic Acids Res. *45*, 5958–5967.

Sorek, R., and Hugenholtz, P. (2008). CRISPR — a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat. Rev. Microbiol. *6*, 181–186.

Steitz, T.A., Smerdon, S.J., Jager, J., and Joyce, C.M. (1994). A Unified Polymerase Mechanism for Nonhomologous DNA and RNA Polymerases. Science. *266*, 2022–2025.

Stewart, C.R., Casjens, S.R., Cresawn, S.G., Houtz, J.M., Smith, A.L., Ford, M.E., Peebles, C.L., Hatfull, G.F., Hendrix, R.W., Huang, W.M., et al. (2009). The Genome of Bacillus subtilis Bacteriophage SPO1. J. Mol. Biol. *388*, 48–70.

Storz, G., Vogel, J., and Wassarman, K.M. (2011). Regulation by Small RNAs in Bacteria: Expanding Frontiers. Mol Cell. *43*, 880–891.

Strydom, N., Gupta, S. V, Fox, W.S., Via, L.E., Bang, H., Lee, M., Eum, S., Shim, T., Barry, C.E., Zimmerman, M., et al. (2019). Tuberculosis drugs' distribution and emergence of resistance in patient's lung lesions: A mechanistic model and tool for regimen and dose optimization. PLoS Med. *16*, 1–26.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science. *324*, 930–935.

Taverniti, V., Forti, F., Ghisotti, D., and Putzer, H. (2011). Mycobacterium smegmatis RNase J is a 5'-3' exo-/endoribonuclease and both RNase J and RNase E are involved in ribosomal RNA maturation. Mol. Microbiol. *82*, 1260–1276.

Toulmé, F., Mosrin-Huaman, C., Sparkowski, J., Das, A., Leng, M., and Rahmouni, A.R. (2000). GreA and GreB proteins revive backtracked RNA polymerase in vivo by promoting transcript trimming. EMBO J. *19*, 6853–6859.

Tsuji, S., Suzuki, K., and Imahori, K. (1976). Crystallisation of DNA-dependent RNA polymerase from Thermus thermophilus HB 8. Nature *261*, 725–726.

Turtola, M., and Belogurov, G.A. (2016). NusG inhibits RNA polymerase backtracking by stabilizing the minimal transcription bubble. Elife *5*, 1–27.

Unciuleac, M.C., and Shuman, S. (2013). Distinctive effects of domain deletions on the manganese-dependent DNA polymerase and DNA phosphorylase activities of Mycobacterium smegmatis polynucleotide phosphorylase. Biochemistry *52*, 2967–2981.

Vaníková, Z., Janoušková, M., Kambová, M., Krásný, L., and Hocek, M. (2019). Switching transcription with bacterial RNA polymerase through photocaging, photorelease and phosphorylation reactions in the major groove of DNA. Chem. Sci. *10*, 3937–3942.

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S., and Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at2.6 Å resolution. Nature *417*, 712–719.

Viswanathan, A., and Doetsch, P.W. (1998). Effects of Nonbulky DNA Base Damages on Escherichia coli RNA Polymerase-mediated Elongation and Promoter Clearance. J. Biol. Chem. *273*, 21276–21281.

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

Walker, K.A., Prabhat, M., Pratt, T.S., and Osuna, R. (2004). The Escherichia coli fis Promoter Is Regulated by Changes in the Levels of Its Transcription Initiation Nucleotide CTP. J. Bol. Chem. *279*, 50818–50828.

Wallace, R.J., Nash, D.R., Tsukamura, M., Blacklock, Z.M., and Silcox, V.A. (1988). Human disease due to Mycobacterium smegmatis. J. Infect. Dis. *158*, 52–59.

Walmacq, C., Rahmouni, A.R., and Boudvillain, M. (2004). Influence of Substrate Composition on the Helicase Activity of Transcription Termination Factor Rho: Reduced Processivity of Rho Hexamers during Unwinding of RNA-DNA Hybrid Regions. J. Mol. Biol *342*, 403–420.

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

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.

Whipple, F.W., and Sonenshein, A.L. (1992). Mechanism of initiation of transcription by Bacillus subtilis RNA polymerase at several promoters. J. Mol. Biol. *223*, 399–414.

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

Winkelman, J.T., and Gourse, R.L. (2017). Open complex DNA scrunching: A key to transcription start site selection and promoter escape. Bioessays *39*, 1–12.

Yarnell, W.S., and Roberts, J.W. (1992). The Phage λ Gene Q Transcription Antiterminator Binds DNA in the Late Gene Promoter As It Modifies RNA Polymerase. Cell *69*, 1181–1189.

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

Zauri, M., Berridge, G., Thezenas, M.-L., Pugh, K.M., Goldin, R., Kessler, B.M., and Kriaucionis, S. (2015). CDA directs metabolism of epigenetic nucleosides revealing a therapeutic window in cancer. Nature *524*, 114–118.

Zaychikov, E., Denissova, L., and Heumann, H. (1995). Translocation of the Escherichia coli transcription complex observed in the registers 11 to 20: "Jumping" of RNA polymerase and asymmetric expansion and contraction of the "transcription bubble." PNAS *92*, 1739–1743.

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