DEVELOPMENT OF AN *IN VITRO*, COUPLED TRANSCRIPTION-TO-TRANSLATION SYSTEM FOR ANALYSIS OF THE INTERACTIONS BETWEEN THE RIBOSOME AND RNA POLYMERASE

DANIEL AUGUSTO CASTRO-ROA

**A dissertation submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy**

**Institute for Cell and Molecular Biosciences,
Centre for Bacterial Cell Biology
Medical School
Newcastle University**

February 2012

Dedication

I would like to dedicate this thesis to my father, Diego and my mother, Maria Victoria whom have made all my dreams come true throughout their unconditional love and efforts. To my brothers, I wouldn't be who I am without you. I also dedicate this work to my wife Manuela and especially to my son, Danny. Their love and understanding have been the greatest helping hand in times of darkness and sorrow. Danny, my son, I lost at least three years of your life whilst concentrating on becoming a Doctor of Philosophy, I owe you so much...
I also dedicate this work to all my professors from Pontificia Universidad Javeriana, in Colombia, from CINVESTAV in Mexico City, from New York University, in New York and from Newcastle University, Newcastle upon Tyne, for opening my mind through their teachings of the fantastic world of science.

Acknowledgements

I am indebted to my supervisor, Nikolay Zenkin, for his unwavering support in the completion of this thesis. It has been a privilege to work for a scientist of this stature.
I would like to acknowledge Mohammad Roghanian, Yulia Yuzenkova, Aleksandra Bochkareva, Rob Van-Nues, Vitaly Epshtein, Ivan Gusarov, Ilya Shamovsky, Christopher Cardinale, Evgeny Nudler, Kostya and Lena Shatalin, Jaime Garcia-Mena, Francisco Martinez, Bruce Dickinson, Dave Mustaine and James Hetfield for their friendship and support.
The members of my thesis committee, Nicholas Watkins and Zofia M. Chrzanowska-Lightowlers have been immensely helpful. In particular, Zofia was a constant source of support and encouragement. I also would like to thank Kenn Gerdes for invaluable comments and ideas.

DECLARATIONS

a) I declare that this thesis is my own work and that I have correctly acknowledged the work of others. This submission is in accordance with University and School guidance on good academic conduct

b) I certify that no part of the material offered has been previously submitted by me for a degree or other qualification in this or any other University.

c) I confirm that the word length is within the prescribed range as advised by my school and faculty

d) Does the thesis contain collaborative work, whether published or not? **No**

If it does, please indicate what part of the work is your independent contribution on a separate sheet.

Signature of candidate ……..…… Date 02/02/2012..

Abstract

The various properties of actively transcribing RNA polymerase (RNAP) complexes with nucleic acids during different stages involve various types of regulation and different cross-talk with other cellular entities and with RNAP itself. For instance, transcription and translation are coupled in bacteria, meaning that translation takes place co-transcriptionally. The interactions of transcriptional apparatus with the translational machinery have been focused mainly in terms of gene expression, whereas the study of the physical interaction of the ribosome and the RNA polymerase remains obscure due to the lack of a system which allows such observations. In this study we have developed a pure, transcription-coupled-to-translation system in which the translocation of the ribosome can be performed in a step-wise manner towards RNAP allowing the observation of the outcomes of the interactions between the two machineries at colliding and non-colliding distances; the system also allows the positioning of RNAP in any desired elongation complex such as paused, roadblocked, backtracked, etc. We show the study of the interactions of the ribosome on different aspects of transcription elongation and also the effects on translation caused by RNAP.

Table of contents

List of Tables

List of Figures

List of abbreviations

RNAP: RNA polymerase
NTP: ribonucleoside triphosphate
EC: elongation complex
AAEC: artificially assembled elongation complexes
CTT: coupled *in vitro* transcription-translation system
TL-CTT: "translation first" coupled *in vitro* transcription-translation system
TR-CTT: "transcription first" coupled *in vitro* transcription-translation system
TLC: thin layer chromatography
TLE: thin layer electrophoresis
TIR: translation initiation region
ORF: open reading frame
PTR: peptidyl transferase reaction
PCR: polymerase chain reaction
PAGE: polyacrylamide gel electrophoresis
SDS-PAGE: Sodium dodecyl sulphate- polyacrylamide gel electrophoresis
TAP: tandem affinity purification tag
LB: Laura Bertani culture media
OD: optical density
EDTA: Ethylenediaminetetraacetic acid
DTT: Dithiothreitol
TGED Tris-Glycerol-EDTA-DTT containing buffer
σ^{70}: transcription initiation factor sigma 70
IPTG: isopropyl-beta-D-thiogalactopyranoside
RPM: revolutions per minute
mRNA: messenger RNA
tRNA: transfer RNA
rRNA: ribosomal RNA

FPLC: fast performance liquid chromatography
TrLB: translation buffer
ATP: adenosine triphosphate
GTP: guanidine triphosphate
UTP: uracil triphosphate
CTP: cytidine triphosphate
GDP: guanidine diphosphate
Rt: reverse transcriptase
dNTP: deoxynucleoside triphosphate
Pyrac: pyridine and acetic acid containing buffer
cDNA: complementary DNA
RNase H: ribonuclease H
TC: ternary complexes (EF-Tu-GTP-Aminoacyl-tRNA)
TL: trigger loop
PPi: Pyrophosphate

Chapter 1. Introduction

1.1. Enzymology of the RNA polymerase

In all living organisms transcription is accomplished by multi-subunit RNA polymerases (RNAP). RNAP is one of the most ancient enzymes on the planet, and it is extremely conserved in evolution from bacteria to humans. Although there are differences in the mechanisms of initiation and the regulation of transcription, mechanisms of catalysis are remarkably similar in all three domains of life. Eukaryotic and archeal RNAPs involve 12-14 subunits (depending on the polymerase type and organism) with a total molecular weight greater than 500 kDa (Cramer et al., 2001).

The simplified versions found in bacteria are composed of four subunits with a molecular mass of approximately 400 kDa. In *E. coli* the catalytically competent core (subunits composition 2α, β`, β and ω) is evolutionarily conserved in terms of its primary sequence, ternary structure and function. Because of its simplicity bacterial RNAP is an excellent model system for understanding the basic principles of all cellular RNAP. The crystal structure RNAP has revealed that both eukaryotic and prokaryotic RNA polymerases have a shape similar to a crab claw and most importantly, that the active centers in both are essentially alike (Vassylyev et al., 2007a, Vassylyev et al., 2002, Brueckner et al., 2009a, Cramer et al., 2001, Gnatt et al., 2001).

The structure of bacterial RNAP is constituted by both fixed and movable elements.

The fixed elements serve primarily as a scaffold that holds in place the catalytic residues in the active site, forms the cleft that accommodates and positions the RNA:DNA hybrid relative to the active centre, makes up the secondary channel through which the nucleoside triphosphate (NTP) substrates enter and it also creates a structure that serves as an attaching site for sigma factors (Borukhov and Nudler, 2008).

The movable elements are needed for both catalysis and translocation of RNAP on the DNA. For example, the β' clamp holds the downstream DNA tightly conferring stability even in the presence of high salt concentrations. It has been shown that it also contributes to the recognition of pause and termination sequences (Nudler et al., 1996) .

Another motile structures required for the functioning of RNAP are the β lobes 1 and 2. The lobes are needed for two major aspects: (i) accommodation of the non-template DNA in the transcription bubble; and (ii) release of both DNA and RNA upon termination (Burgess and Anthony, 2001, Gnatt et al., 2001, Korzheva et al., 2000).

Some mobile elements such as the β' rudder and β' lid, β' zipper, β' N-terminal Zn finger and β' flap, contribute to the proper threading of the DNA, RNA and DNA:RNA hybrid, the correct docking of sigma factors and also to the recognition of termination, pausing and antitermination signals and factors (Murakami et al., 2002, Vassylyev et al., 2002, Kuznedelov et al., 2002, Naryshkina et al., 2006) (Figure 1).

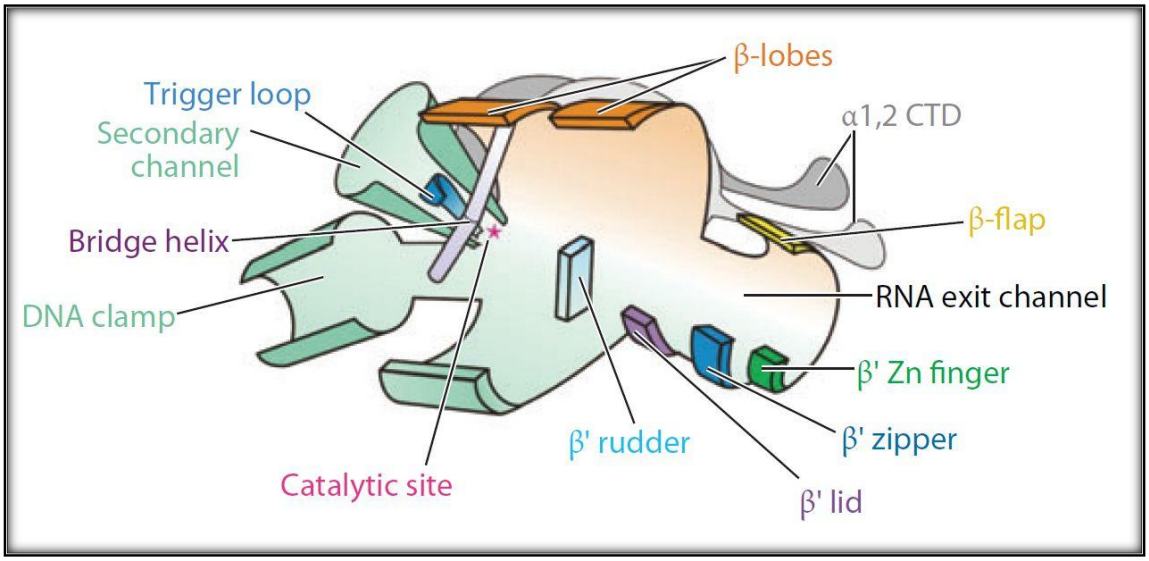

Figure 1. Representation of *Thermus thermophilus* RNA polymerase core enzyme with mobile and fixed elements labelled. (Borukhov, S. Trends Microbiol 2008 Mar;16(3):126-34)

It has been observed that RNAP contains three major nucleic acid binding sites that hold together the elongation complex and also arrange the 3' end of the DNA in the active centre for catalysis to occur. These sites are: the front DNA binding site (DBS) (Gnatt et al., 2001, Kettenberger et al., 2004, Vassylyev et al., 2007a), the RNA:DNA hybrid binding site (HBS) (Korzheva et al., 2000, Wang et al., 2006) and the RNA binding site (Westover et al., 2004, Korzheva et al., 1998) (Figure 2).

It has been shown through biochemical (Nudler et al., 1996, Nudler et al., 1998) and structural (Korzheva and Mustaev, 2001, Korzheva et al., 2000) studies that the DBS interacts with ~9 bp downstream the active centre but at position +2 of the DNA (counting from the NTP binding site (+1)), a sharp kink of ~90 degrees takes place (Rees et al., 1993). The bent DNA facilitates strand separation by structures present in the DNA clamp such as the fork loop. This structure also blocks further advance of the downstream DNA helix into the active centre and keeps the downstream edge of the transcription bubble by preventing re-association of the DNA (Vassylyev et al., 2007a, Wang et al., 2006). Notably, all the contacts in the DNA binding site with the DNA are

not sequence specific. These relatively weak interactions allow RNAP to be highly processive during elongation.

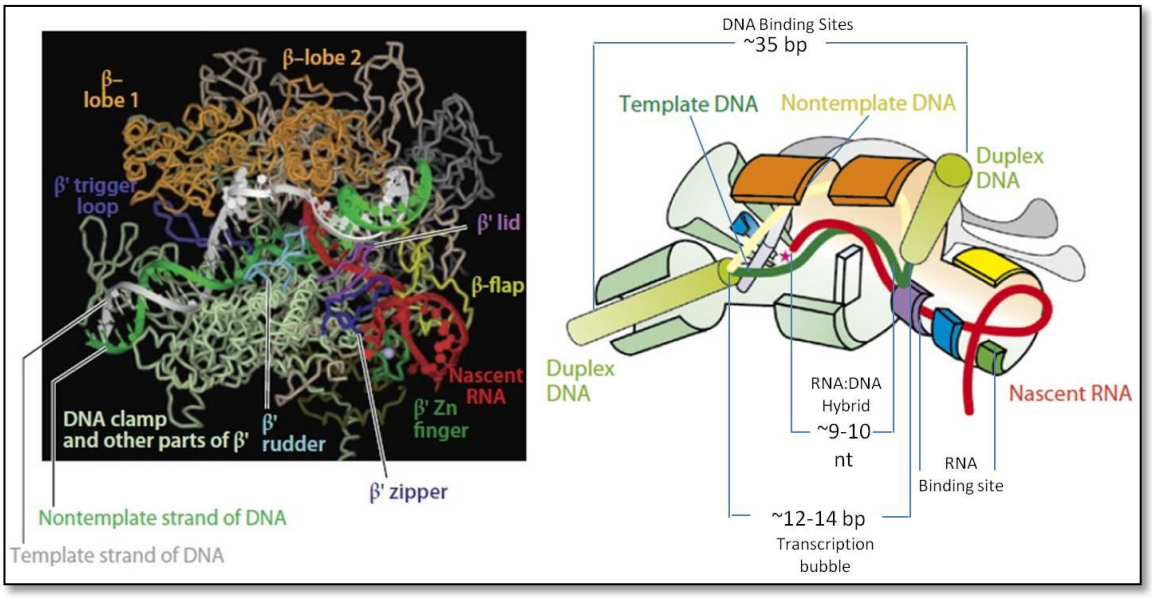

Figure 2. Representation of *T. thermophilus* RNAP with nucleic acid interactions Left: High-resolution crystal structure of *T. thermophilus* RNAP core enzyme. Mobile elements are marked and nucleic acids are shown. Right: Schematic representation of RNAP with nucleic acids interactions. Length of the nucleic acid binding sites is shown. Colours are the same as in left for easier comparison. Red star represents the active centre Modified from Nudler. E. Annu. Rev. Biochem. 2009 78:335-361

The ~9-10 nt (Nudler et al., 1997) long RNA:DNA hybrid is positioned in the main channel formed by β and β'. This cleft is usually referred to as hybrid binding site (HBS) (Gnatt et al., 2001, Kettenberger et al., 2004). The interactions of the RNA:DNA hybrid to the HBS are similar to those found in the DBS. The weak contacts reduce the "friction" facilitating both forward and backward translocation of the enzyme (Nudler et al., 1997). The length of the hybrid and the transcription bubble is kept regularly constant throughout elongation due to the interactions of the nucleic acids with structures such as the β' lid loop, β' rudder and β fork loop-1. The β' lid loop (see Figure 2) acts as a wedge that displaces RNA from the RNA:DNA hybrid binding site into the RNA exit channel disallowing the formation of a longer hybrid (Westover et al., 2004, Naryshkina et al., 2006, Toulokhonov and Landick, 2006). The β' rudder interacts with and stabilizes the RNA-DNA hybrid in the elongation complex. (Vassylyev et al., 2007a, Kuznedelov et al., 2002) . The β fork loop-1 interacts with the RNA phosphates at positions -5, -6 and -7 preventing the unwind of the hybrid (Westover et al., 2004)

RNA is separated from the RNA:DNA hybrid (usually at positions -9 to -10) by the RNA binding site. This site is composed by the β' lid, and β "saddle" structures. These two structures form an aperture where RNA is extended. This narrow aperture directs to a wider one formed by zipper, Zn finger domains of β' and the β-flap domain (Vassylyev et al., 2007a).

These interactions of RNAP with nucleic acids make elongation complex highly stable and resistant to very high ionic strength (1 M KCl) and to competitors such as heparin (Kuznedelov et al., 2002). Besides unusual stability of the elongation complex, transcribing RNAP generates a considerable force. The calculated force of the actively transcribing RNAP is 20 pN per molecule of RNAP and is additive when more RNAP molecules collide with each other (Wang et al., 1998). This allows RNAP to overcome unwanted events and also dislodge proteins bound to DNA template (Epshtein et al., 2003).

This powerful and stably bound to DNA machine is capable to proceed through thousands of base pairs without losing contact with the DNA template and nascent RNA. RNAP is also capable to recognize different signals in the DNA that would make it prone to undergo through different non dissociative states, (paused or backtracked), and can also recognize sequences that lead to the disassociation and therefore termination of transcription (Toulme et al., 2005, Bochkareva et al., 2011). However, core enzyme needs to be assisted by accessory factors in order to recognize promoters that indicate where transcription begins (sigma subunits), and some other auxiliary factors such as GreA/B or Nus that would affect different aspects of elongation (Borukhov et al., 1993).

The active centre of RNAP is positioned on the back wall of the β' subunit where three aspartate residues embedded within the absolutely conserved NADFDGD motif coordinate the magnesium ion (Mg-A). The position of the catalytic site relative to the 3' end of the RNA:DNA hybrid is known as register. For the NTP binding cycle two states have been observed: pre-translocated and post-translocated (reviewed in (Brueckner et al., 2009b). In the pre-translocated state, a new nucleotide has been added to the growing RNA chain and still occupies the nucleotide insertion site (referred to as i+1 site). In order to add a new NTP to the RNA, RNAP must first move the RNA 3' nucleotide from the i+1 site to the product site (referred to as i site) generating enough

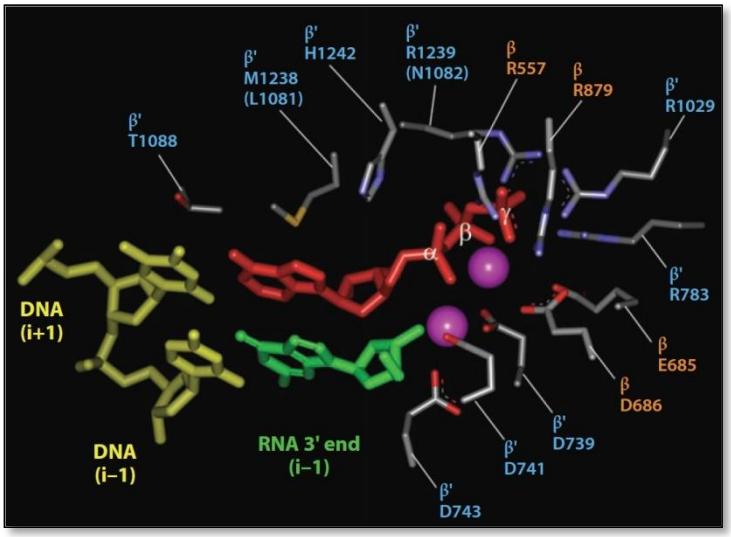

Figure 3 Structure of RNAP active centre of *Thermus thermophilus*. Evolutionarily conserved amino acid residues, Mg^{2+} ions (purple), substrate NTP (red), template DNA (gold), nascent RNA (green) are shown. Residues responsible for the coordination of the Mg^{2+} ions and the orientation of the NTP α-, β- and γ phosphates are shown. (Nudler, E. 2009)

space for the incoming nucleotide through a process known as translocation. This movement of RNAP by 1 nt changes the register into the post-translocated state, where the incoming NTP can enter the i+1 site and align with the RNA 3' OH followed by catalysis of the phosphodiester bond. The catalytic activity of RNAP relies on two Mg^{2+} ions. Besides the magnesium ion afore mentioned, another magnesium ion (Mg-B) is stabilized by the incoming NTP. Mg-A is responsible to the activation of the 3'-terminal OH group and for proper arrangement of the α-phosphate in the incoming NTP. Mg-B stabilizes the α, β and γ phosphates of the NTP keeping it aligned for phosphodiester bond formation (Castro et al., 2009, Kireeva et al., 2009, Vassylyev et al., 2007b, Sosunov et al., 2005) (Figure 3). Catalysis results in the covalent attachment of one NMP to the RNA 3' end generating as a by-product a pyrophosphate (PPi) molecule.

Release of pyrophosphate from the active centre completes the cycle leaving the active centre ready for another round of NTP addition.

Catalysis in the active centre is further modulated by two motile elements. These elements are known as the β'F(bridge)-helix (BH) and β'G(trigger)-loop (TL). In crystal structures of RNAP elongation complexes, the TL was observed in two distinct states, folded and unfolded (Zhang et al., 1999, Cramer et al., 2001, Gnatt et al., 2001, Wang et al., 2006). In the unfolded conformation, NTP is allowed to enter into the active centre. The NTP base-pairs with the template DNA in the i+1 site, inducing folding of the TL. The folded TL not only positions the phosphates of the NTP close to the active centre to stimulate phosphodiester bond synthesis, but also through interactions with the NTP moieties (the base, sugar and phosphates) ensures the correct chemical nature of the NTP and its pairing with the template DNA maintaining transcriptional fidelity. Once the phosphodiester bond has been formed, the TL returns to the unfolded conformation allowing pyrophosphate release (Vassylyev et al., 2007b, Kaplan et al., 2008, Yuzenkova et al., 2010).

It is important to mention that the conformational changes in the active centre with those of the TL have important implications for RNAP forward translocation (Bar-Nahum et al., 2005). The proposed mechanism in which RNAP translocates is based in a nucleotide-biased Brownian-ratchet mechanism (Guajardo and Sousa, 1997, Bai et al., 2004) . In this model (reviewed in (Cramer and Arnold, 2009)), in the absence of NTPs, RNAP in the elongation complex (EC) oscillates between the pre and post translocated states; when the NTP binds to the i+1 site, a pre-insertion state is formed which acts as a ratchet that stabilizes the pre-translocated state and bias the EC towards forward translocation. Therefore, translocation is considered to be driven by thermal motion where binding of the cognate substrate would favor the post-translocated state (Sydow and Cramer, 2009) (Figure 4). The oscillation between the pre and post-translocated state also implies that RNAP is prone to non-catalytically slide backwards by one or more nucleotides along the DNA when elongation is blocked via NTP starvation or when particular DNA sequences are present in the RNA:DNA hybrid (Guajardo and Sousa, 1997, Komissarova and Kashlev, 1997, Abbondanzieri et al., 2005, Bai et al., 2004, Tadigotla et al., 2006, Bochkareva et al., 2011), a phenomenon known as backtracking. During backtracking, RNAP shifts backwards in a manner of a zipper: the 3' end of RNA disengages from the template DNA strand and the active centre, while the rear end of RNA-DNA hybrid RNA anneals back to the template. This keeps the length of the RNA-DNA hybrid the same as in active elongation complex, which

means that this arrested complex is as stable as the active one. In the backtracked complexes, the 3' end of the RNA exits through the RNAP secondary channel losing its register with the active centre (Kireeva and Kashlev, 2009a, Shaevitz et al., 2003, Hogan et al., 2002, Nudler et al., 1997). These arrested complexes require auxiliary factors such as cleavage factors GreA\GreB which upon cleavage of the RNA in the secondary channel will return the 3' end of the RNA back in register with the active centre (Borukhov et al., 1993) for their resolution and other transcript cleavage mechanisms such as intrinsic hydrolysis (see below).

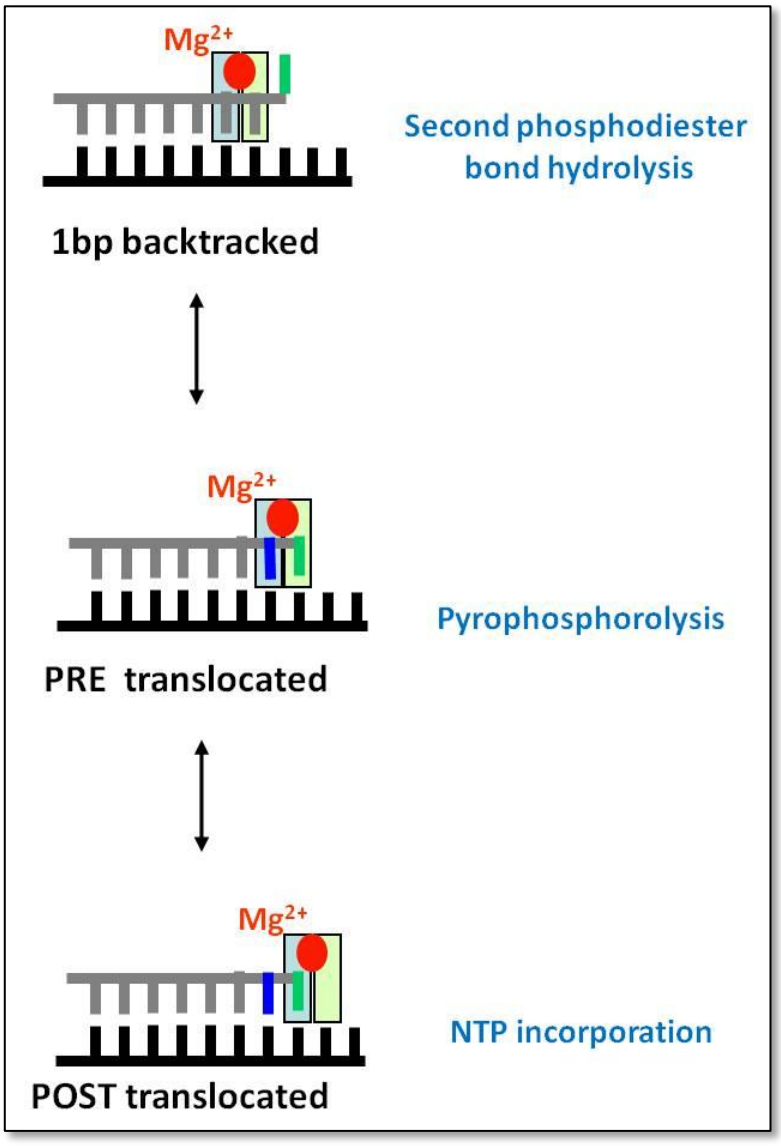

Figure 4. Oscillations of the transcription elongation complex. Scheme of RNAP oscillation in translocation equilibrium and the architecture of the nucleic acid scaffold of the EC in backstepped (1 nt backtracked), pre-translocated and post-translocated states. Template DNA (black), RNA (gray) are shown. Catalytic Mg^{2+} ion (red circle) and the activities that take place in the active centre during different states (blue text) are shown. Modified from (Bochkareva, A. *et al*, 2011)

Since translocation of RNAP is driven by thermal motion, it is possible that external forces applied on RNAP could lead to a shift on the equilibrium between pre, post-translocated and backtracked states (Shaevitz et al., 2003); this could explain the positive mutual aid observed amongst RNAP molecules capable of rescuing backtracked and roadblocked EC (Epshtein and Nudler, 2003, Epshtein et al., 2003) and also between the ribosome and backtracked RNAP (Proshkin et al., 2010); these interactions will be discussed later.

The actively transcribing RNAP is also capable to perform different reactions in its active centre. For instance, in the post-translocated state RNAP can incorporate a new NTP to the nascent RNA chain, whereas in the pre-translocated state (before the movement by one nucleotide) it can perform both exopyrophosphorylase and exonuclease activities and whilst backtracked it can exert endonuclease and endopyrophosphorylase reactions (Figure 5). It's been recently shown that RNAP uses its hydrolytic capacity as a mechanism to ensure fidelity (Zenkin et al., 2006b, Yuzenkova and Zenkin, 2010a) .

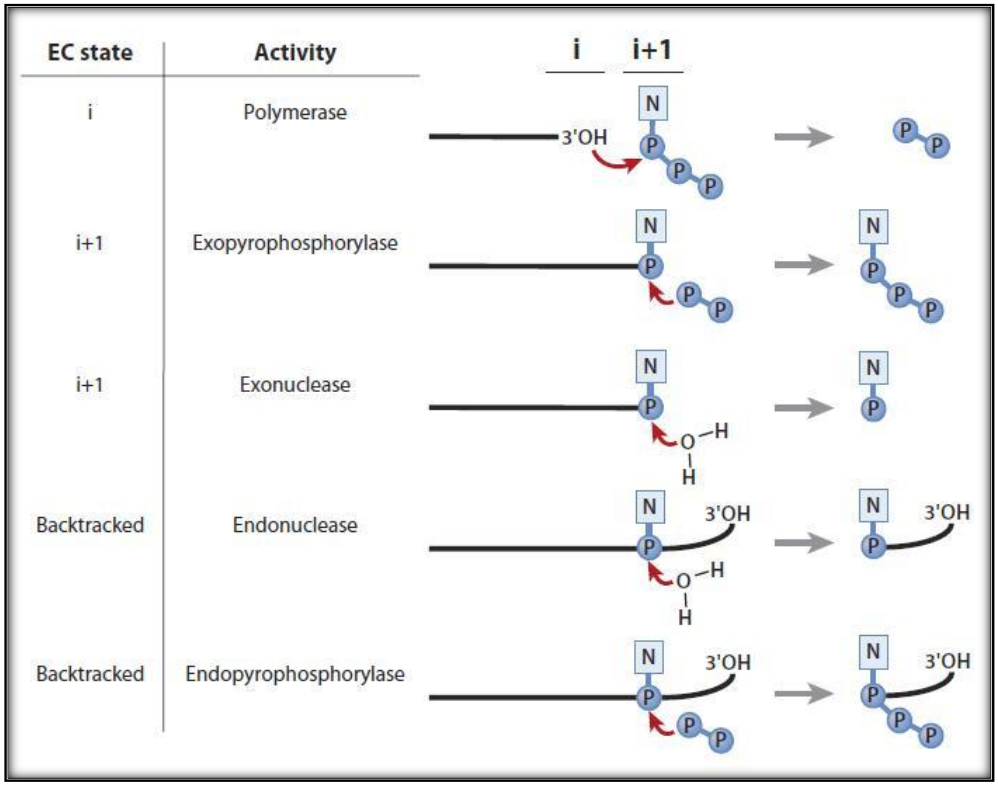

Figure 5. Schematic representation of the different activities in the RNAP active centre. Note that the activities depend on the elongation complex (EC) state. Sosunov, V. 2003

1.2. *The Transcription cycle*

Being at the very beginning of the path of gene expression transcription is a complex and heavily regulated process in the cell. Such regulation occurs during the three stages of the transcription cycle: **initiation**, **elongation** and **termination** (Figure 6)**.**

The catalytic properties of RNA polymerase at these stages are the same but are regulated differently. Although great progress has been achieved recently in understanding its catalytic and structural properties by biochemical and especially by crystallographic studies, much still remain obscure.

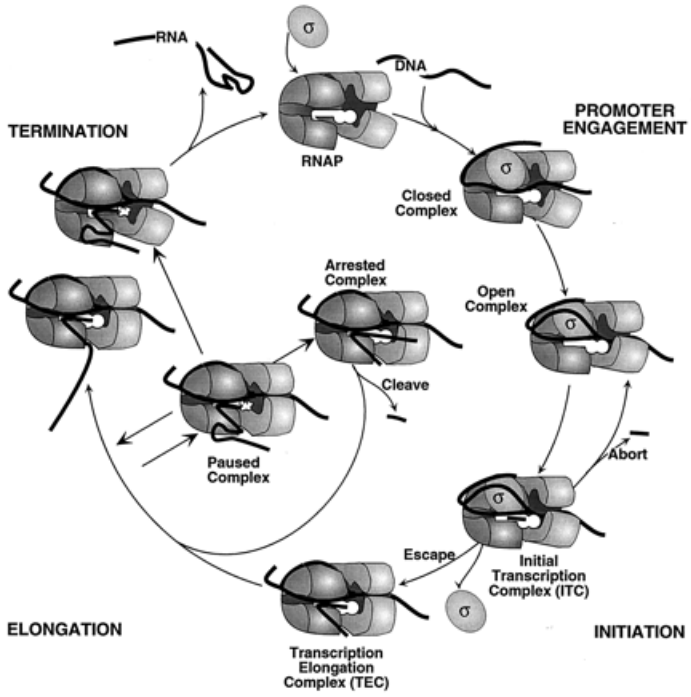

Figure 6. The transcription cycle. Representation of the three stages of transcription. 1) Initiation. RNAP equipped with sigma factor recognizes a promoter sequence on the DNA and binds to it forming promoter closed complex. Through a series of interactions with the DNA and both sigma and RNAP, the DNA is melted and the transcription bubble is formed. This complex is known as open promoter complex. RNAP begins RNA synthesis in an abortive manner until an RNA of 8 – 9 nt long is formed and stabilised in the RNA:DNA binding site, forming the initial transcription complex. After formation of such complex, RNAP escapes the promoter, releasing sigma factor rendering a highly processive state known as transcription elongation complex. 2) Elongation. During elongation, RNAP is capable of synthesising RNA in an almost uninterrupted manner. Transcription elongation is regulated by backtracking and pausing (see text). 3) Termination. RNAP finalises transcription upon interactions with an inverted repeat palindromic sequence followed by a poly A sequence on the DNA. Such sequences are known as terminators and cause destabilisation of the RNA:DNA hybrid which leads to RNA release and disengagement of RNAP from the DNA. The free RNAP is then able to enter the transcription cycle again. **Mooney, R. 2005**

1.2.1 *Transcription initiation*

The first stage of transcription is initiation and it is the first step in gene regulation. This phase begins with the search of the promoter by the core enzyme equipped with the transcription initiation σ factor through a scanning mechanism (Park et al., 1982, von Hippel, 2007). In *E. coli*, there are seven sigma factors, each one of them believed to recognize a large subset of promoters and are needed for the expression of specific genes under environmental stimulus (Ghosh et al., 2010). Besides recognizing and

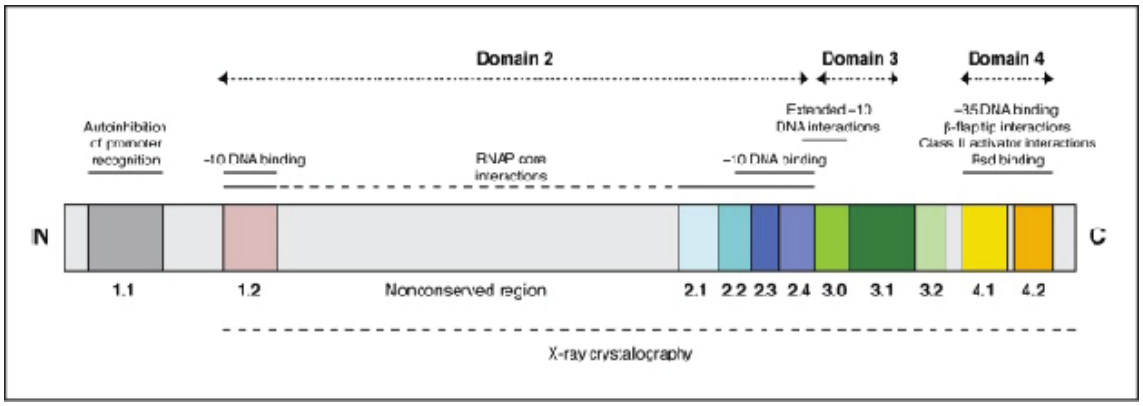

Figure 7. Protein map of σ^{70} factor**.** Conserved regions 1.1-4.2 (coloured) is differentiated from non conserved regions (grey). Function of each region is indicated. The regions resolved by crystallography are also shown. Ghosh T, Bose D, Zhang X. FEMS Microbiol Rev. 2010 Sep;34(5):611-27. Epub 2010 Jun 7. Review

binding to promoters, σ factors have other roles such as serving as targets for regulation of transcription initiation, promote DNA melting and inhibit nonspecific initiation (Buck et al., 2000, Zenkin and Severinov, 2004). In *E. coli*, σ^{70} is the predominant σ factor and binds to the promoters of housekeeping genes (consensus sequence -35 TTGACA and -10 TATAAT) (Gruber and Gross, 2003). σ^{70} is composed of four domains (σ1, σ2, σ3 and σ4) connected by flexible linkers (Malhotra et al., 1996, Campbell et al., 2002). Through crystallographic (Murakami et al., 2002, Vassylyev et al., 2002) and biochemical studies (Severinov et al., 1994, Waldburger and Susskind, 1994, Barne et al., 1997, Wilson and Dombroski, 1997, Zenkin et al., 2007) it has been shown that these domains are involved in promoter recognition, core binding and DNA melting (reviewed in (Murakami and Darst, 2003) (Figure 7).

The means by which RNAP is thought to find the promoter involves the tracking of a groove of the double helix throughout electrostatic interactions reinforced by the entropy that result from the displacement of the counterion cloud that surrounds the DNA (Sakata-Sogawa and Shimamoto, 2004). These initial interactions of the enzyme

with DNA are weak and therefore unstable. σ subunit recognizes a promoter, which usually is comprised of two hexameric sequences around -10 and -35 positions relative to transcription start site, to which it has high affinity (Travers, 2004). The interactions of RNAP with promoter DNA are expanded by the wrapping of the DNA (demonstrated by DNA footprinting experiments) on the surface of the enzyme making the complex more stable (Rogozina et al., 2009). Wrapping also facilitates further rearrangements in both, RNAP and DNA. During this state, known as promoter closed complex (Li and McClure, 1998), the DNA helix remains double stranded and the complex, though being relatively stable, remains sensitive to high ionic strength and competitors such as heparin (Coulombe and Burton, 1999). The stability of this complex depends on the sequence of the promoter. Generally, the farther the sequence of a promoter from the consensus, the less stable promoter complex will be formed on it (Fenton and Gralla, 2001). Promoter sequence also determines the capacity of RNAP to compete with repressors and nucleoid proteins (Grainger et al., 2006).

After formation of the closed complex the double helix of DNA is destabilized by action of specific residues in the σ factor on a precise region of the promoter (Aiyar et al., 1994, deHaseth and Helmann, 1995). Then, the enzyme melts the double helix of the DNA and forms a stretch ~17 nucleotides (nt) of unwound DNA known as the transcription bubble (for some particular σ factors, the energy required is obtained by ATP hydrolysis (Merrick, 1993)). The melting generates further rearrangements of the DNA inside RNAP, placing the downstream DNA into the enzyme's DNA-binding clamp and positioning the template DNA in register with the active centre making the complex catalytically competent (Craig et al., 1998). This new configuration of RNAP is referred to as open promoter complex (Li and McClure, 1998, Mekler et al., 2002). At this point RNAP occupies a total of ~35 bp of the DNA and has undergone several structural rearrangements that provide higher stability compared to afore mentioned closed promoter complexes. Open promoter complex is capable to withstand higher ionic strength (200 mM KCl) and becomes resistant to competitors (Reppas et al., 2006, von Hippel et al., 1984).

Initiation of transcription starts with synthesis of short RNA transcripts (2-9 nucleotides long). Given that RNAP remains anchored to the promoter by σ factor, synthesis involves pulling of a stretch of downstream DNA of the same size inside the main channel of RNAP. This phenomenon is referred to as "scrunching" (Revyakin et al.,

2006). Scrunching results in increase of the size of the transcription bubble (given that its upstream edge is kept at the same position by σ subunit). Initiation ends when RNAP escapes the promoter and releases the σ factor leaving RNAP ready to elongate the RNA chain (Figure 8).

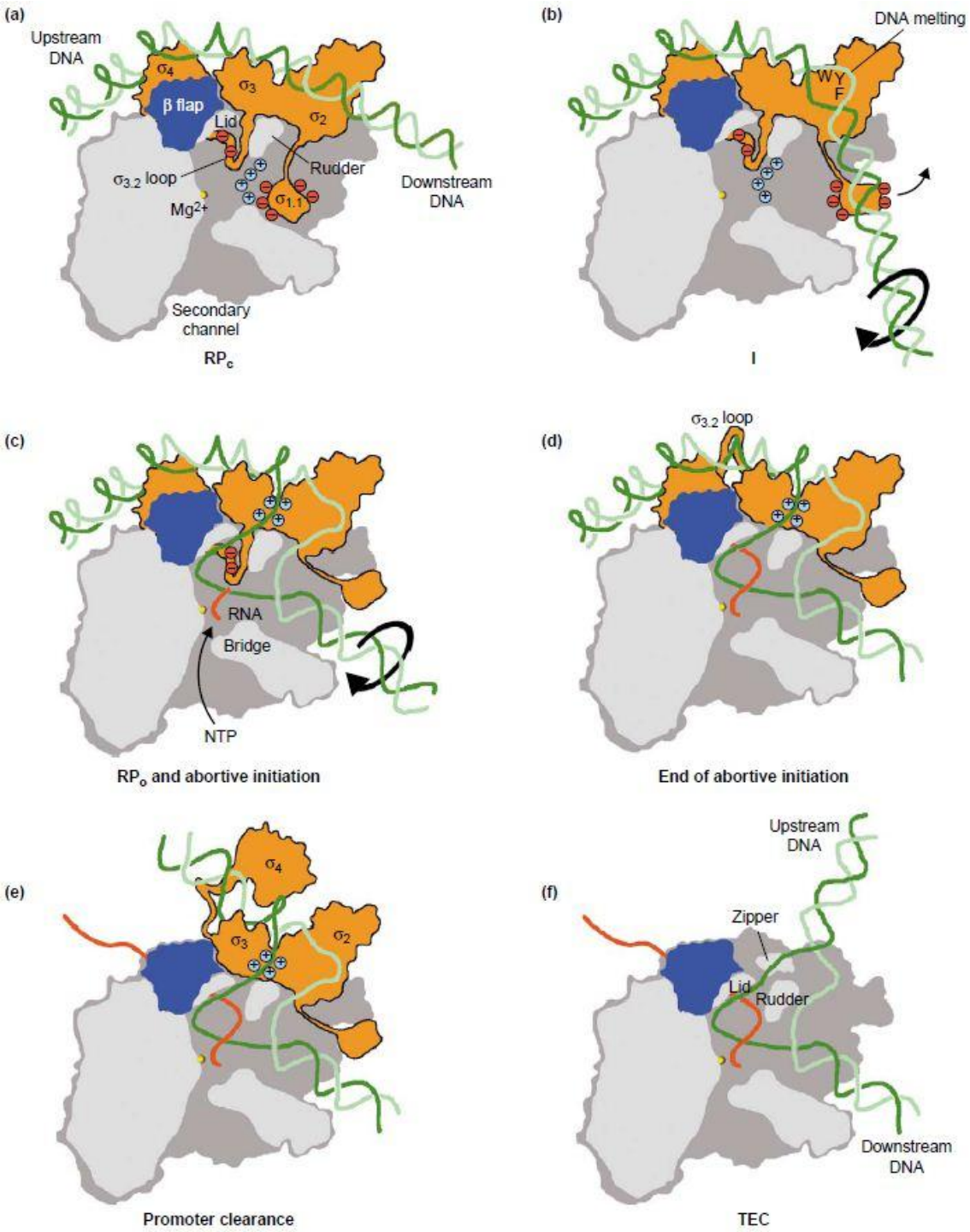

Figure 8. Cross-sectional view of the structural shifts throughout transcription initiation. RNAP grey (dark and light), β flap (blue), σ^{70} (orange), template DNA (dark green), non-template DNA (light green), Mg^{2+} ion (yellow circle)

and nascent RNA (red) are shown. A) Docking of σ^{70} with the core enzyme through interactions of $\sigma_{1.1}$ with β' rudder and $\sigma_{3.2}$ loop with β' lid (shown with its charges) and β flap in σ_4 pocket takes place. Upon binding to the promoter -35 and -10 regions through σ_4 and σ_2 respectively, the closed promoter complex is formed. B) Towards the formation of the open promoter complex, an intermediate (I) has been observed. The DNA is kinked towards the RNAP main channel by interactions with $\sigma_{1.1}$. Simultaneously, melting of the DNA takes place by action of aromatic residues in $\sigma_{2.3}$ (WYF) on the -10 region of the non-template strand of the DNA forming the upstream region of the transcription bubble. C) Fully unwound transcription bubble is observed and it is kept by interactions with σ_2 and the template DNA and with $\sigma_{3.2}$ loop. The template DNA is positioned in the active centre allowing RNAP to enter into abortive synthesis of RNA through a "scrunching" mechanism. This state is called the open promoter complex D) After elongation of a ~12 nt long RNA the 5' end of the RNA displaces the $\sigma_{3.2}$ region and finds its way out through the RNA exit channel. E) The displacement of $\sigma_{3.2}$ region weakens the binding of σ_4 region with β flap leading to a total destabilization of the σ factor. Consequently, σ is released and RNAP enters into elongation. F) RNAP in elongation state. Note that the transcription bubble is not longer being held by σ^{70} but by the zipper, rudder and lid elements. Katsuhiko, S. Curr Opin Struct Biol. 2003 Feb;13(1):31-9.)

1.2.2 *Transcription Elongation and Termination*

The experiments performed for the analysis of the functioning of our transcription-translation system were focused on the effect of the collision of the ribosome with RNAP positioned on transcriptional states that take place during elongation such as hairpin independent pausing and backtracking; such transcriptional events will be discussed below.

The transition from initiation to elongation could be defined as the point at which abortive initiation ceases and σ is released, leaving a stable ternary complex that consists of core, RNA and DNA. In simple terms, if using a comparison of RNAP to a crab claw, the claw in initiation is more open, while in elongation it closes on the RNA-DNA hybrid almost fully surrounding it, making it more stable and resistant to high ionic strength and competitors such as heparin (Kuznedelov et al., 2002). As mentioned above, elongation involves the recognition of the correct NTP, its incorporation to the 3' OH of the nascent RNA chain through formation of the phosphodiester bond followed by the release of pyrophosphate and finally, the translocation of RNAP by one nucleotide. Though RNAP is highly processive, it is subject to stringent regulation. The transcription elongation complex is capable of uninterrupted synthesis of RNA chains thousands of nucleotides long at an approximate rate of 70 nt·s-1 (in *E. coli* (Gotta et al., 1991)), yet it is still capable to recognize regulatory signals embedded on the DNA and RNA and also to be regulated by transcription factors. Those regulatory signals and factors will be discussed below.

1.2.2.1 *Pausing, backtracking and termination*

Like in initiation, the elongation phase of the transcription cycle is also under rigid control, which is achieved through **pausing** and **termination** (Uptain et al., 1997). The main distinction between the two mechanisms is that pausing does not cause disassociation of the elongation complex whereas termination does (McDowell et al., 1994).

RNAP allows processive and uninterrupted transcription elongation, mostly because it is indifferent to the sequences of the nucleic acids which are being transcribed; this observation could indicate the reason why RNAP core enzyme requires auxiliary factors in order to recognise specific sequences during both initiation and elongation (Zenkin and Severinov, 2008). However, during elongation, RNAP often pauses at different locations along the template both *in vitro* and *in vivo* (Jin et al., 2010, Landick, 2006, Toulme et al., 2005) regulating the rate of transcription through specific sequences encoded in the template DNA. Pause is a reversible elongation-incompetent state which is strongly sequence dependent (Landick, 2006). However, pause signals are encoded indirectly and no consensus pause sequence has been identified so far. Pause durations follow an exponential distribution with the majority of pauses lasting less than 10 s (Herbert et al., 2006). In *E. coli*, the average rate of pausing is 0.55 s^{-1} (i.e., approximately once in every 100 bases) (Neuman et al., 2003) and their average duration is 3 s (Greive and von Hippel, 2005). It is widely accepted that all types of pauses (below) result from an elemental pause (Landick, 2009, Zhang et al., 2010, Landick, 2006). The elemental pause is caused by a small range active-site rearrangement that does not involve translocational change (Artsimovitch and Landick, 2000, Markovtsov et al., 1996). The elemental pause can be stabilized in longer-lived pauses by one or more additional events. Different modes of pause stabilization generate different types of pauses, but all of them eventually will result in RNAP escape into the normal elongation pathway. All transcription pauses are considered to be off pathway (non-mandatory) events (Herbert et al., 2008, Landick, 2009), meaning that upon pause signal recognition, the EC branches off the nucleotide addition cycle leading to a pause of transcription. This branching can occur through backtracking of the EC (when the 3' end of the RNA disengages from the RNAP active centre) or through conformational changes in RNAP active centre which slows down catalysis of the phosphodiester bond (Landick, 2009). Pauses cause kinetic partitioning of RNAPs between active and paused

states, meaning that not all transcribing RNAPs are uniformly affected (Landick, 2009), suggesting that some RNAPs would stay in the pathway being not responsive to the pausing signals. However, Bochkareva and co-workers described an "in pathway" pause which is caused by delays in the steps within the NTP addition cycle, which affects all RNAPs uniformly; in their study they observed that RNAP can sense the identity of the RNA-DNA hybrid of the EC, restricting the oscillation between the translocation states (pre and post-translocated), and if the translocation equilibrium favors the pre-translocated state it may lead to pause of transcription; this situation leads to a slow down of the rate of translocation which limits the rates of NTP binding and phosphodiester bond formation (Bochkareva et al., 2011).

Pausing mechanisms have been arranged into two major groups: RNA hairpin-dependent and hairpin-independent. The hairpin independent pauses can be further divided into three groups: pre-translocated stabilised, backtracking and non-backtracking and RNAP pauses (Landick, 1997, Bochkareva et al., 2011)

The hairpin-dependent pauses depend on the formation of an RNA hairpin ~14 nt upstream of the active site (Artsimovitch and Landick, 2000). It has been observed that upon folding, the hairpin reversibly freezes the catalytic center in the pre-translocated state through interactions with the β−flap domain (Toulokhonov et al., 2001) . Through cross-linking and X-ray crystallography experiments, the trigger loop has been observed to be in its closed (folded) conformation, contacting the 3' end of the RNA blocking the access of incoming NTPs substrates, rendering the RNAP inactive (Toulokhonov et al., 2007). It has also been observed that not only the hairpin is important for the pause to occur. Sequences downstream of the hairpin also dictate the efficiency and the site of the pause (Toulokhonov and Landick, 2003). The hairpin dependent pauses are often regulated by transcription factors such as Nus factors which may enhance or reduce the dwelling time of the pause (Yakhnin and Babitzke, 2010) (table 1) .

Observed non-backtracking hairpin independent mechanisms include misalignment of the incoming nucleotide (Kireeva and Kashlev, 2009a), sequences resembling promoter elements which are recognized by σ-factor which fails to disassociate from RNAP core enzyme (Ring et al., 1996, Brodolin et al., 2004) and the RNA:DNA hybrid "sensing" mechanism described above (pre-translocated stabilised paused RNAP) (Bochkareva et

al., 2011). It was shown recently that non-backtrack pausing can occur in prokaryotes (Kireeva and Kashlev, 2009a) and in eukaryotes as observed *in vitro* for the pre-translocated stabilized paused RNAP (Bochkareva et al., 2011).

Backtracking pauses are the most common mechanism of pausing for both pro and eukaryotes (Landick, 2006, Landick, 2009). It was suggested that backtrack pauses are thermodynamically governed and occur when the free energy of backtracked elongation complex is lower than that of the corresponding non-backtracked EC (Erie et al., 1992). Theoretical models based on the thermodynamics of transcription elongation complexes were able to predict pause sites on previously characterized DNA templates with high accuracy (Bai et al., 2004, Tadigotla et al., 2006). During backtracking, RNAP shifts backwards disengaging the 3' end of RNA from the template DNA strand and the active centre, keeping the length of the RNA-DNA hybrid the same as in active elongation complex, meaning that the backtracked complex is as stable as the active one. In the backtracked complexes, the 3' end of the RNA exits through the RNAP secondary channel losing its register with the active centre (Nudler et al., 1997). These arrested complexes require auxiliary factors such as cleavage factors GreA\GreB which upon cleavage of the RNA in the secondary channel will return the 3' end of the RNA back in register with the active centre (Borukhov et al., 1993) for their resolution; backtracked complexes are assessed *in vitro* analyzing the size of the GreB mediated cleavage product which reflects on the extent of backtracking by RNAP. Transcription elongation complexes resistant to GreB cleavage are thought to be stabilized in the post-translocated state where there is no phosphodiester bond in the i+1 site (Zhilina et al., 2011)

Regulatory Molecules for transcription elongation	**Size**	**Target**	**Mechanism**
NusA	55 kDa	RNAP (α-CTD-NusA CTD); nascent RNA hairpins	On its own, enhances hairpin dependent pausing and termination. It is a component of antitermination complexes. It modulates ρ-dependent termination
NusB	15.7 kDa	BoxA nascent RNA and NusE	NusB bound to NusE forms a heterodimer which binds boxA sequence in antitermination complexes
NusE	11.7 kDa	BoxA nascent RNA and NusE	NusB bound to NusE forms a heterodimer which binds boxA sequence in antitermination complexes. NusE is also S10 protein in small ribosome subunit
NusG	20.5 kDa	RNAP and ρ.	On its own reduces backtrack pausing. It is also a component of antitermination complexes. Increases ρ termination.
GreA	17 kDa	RNAP secondary	It stimulates transcript cleavage and anti-arrest.

		channel and nascent RNA	It is involved in promoter escape.
GreB	17 kDa	RNAP secondary channel and nascent RNA	It stimulates transcript cleavage and anti-arrest. It is involved in promoter escape.
HepA	110 kDa	RNAP	Binds tightly to RNAP. It assists transcription of compacted DNA.
Mfd	130 kDa	RNAP and duplex DNA.	Terminates arrested elongation complexes. Involved in transcription-coupled DNA repair.
(p)ppGpp	676 Da	RNAP E site	Slows down elongation. Increases pausing.
ρ	6X 47 kDa	Nascent RNA and RNAP	Transcription termination. ATP-dependant helicase. It translocate on RNA to disassociate transcription elongation complex.

Table 1. Transcription elongation regulatory factors. (Taken from Bact. 612 Lecture #3 & #6 notes. Transcriptional attenuation R. Landick. Oct. 4 & 6. 2004, and references therein).

Regulatory factors binding to a single-stranded element of non-template strand as in the case of sigma-associated (lambda *plac* promoter) and RfaH-associated (*ops*) pausing also cause backtracking (Zhang et al., 2007, Ha et al., 2010, Herbert et al., 2010), although it has been recently observed that the *ops* pause could also be due to stabilisation of the pre-translocated state through sequence recognition of the RNA:DNA hybrid by RNAP (Bochkareva et al., 2011).

Transcription regulation through pausing is an important mechanism in prokaryotes and eukaryotes (Landick, 2006). It has been observed that promoter-proximal pausing plays a crucial role in the regulation of phage development (Hatoum and Roberts, 2008) and recently, promoter-proximal pausing has also been suggested to be a common way of bacterial gene expression regulation (Perdue and Roberts, 2011). Pausing also mediates the regulation of expression of bacterial operons such as *his* (Chan and Landick, 1989), *pyr* (Donahue and Turnbough, 1994), *trp* (Landick et al., 1987), amongst other operons (Yanofsky, 1981) in response to nutritional conditions through hairpin-dependent pausing mechanism. Pausing plays an important role in coupling transcription and translation by slowing down RNAP to allow a translating ribosome to catch up to RNAP and then release the enzyme from the paused state which is the case for the *trp* operon leader region (Landick et al., 1985); this type of regulation is known as transcription attenuation. Transcription attenuation can be defined as any mechanism that uses transcription pausing or termination to modulate expression of downstream genes (Yanofsky, 1981). Attenuation allows bacteria to control the transcriptional read through in response of a metabolite or an environmental stress.

Some examples of attenuation based on translation of a leader peptide, will be discussed in the RNAP interaction with other cellular machineries chapter.

In order to maintain transcription-translation coupling, pauses are present frequently enough to avoid the exposure of more than 100 nucleotides of unstructured RNA, which is the minimal loading space for the termination factor Rho (ρ) (Ciampi, 2006) .
In bacteria, halting RNAP at key locations through pauses also allows the interaction and recruitment of regulators such as lambda Q (Roberts et al., 1998).

Backtracking, like hairpin dependent pauses, have several biological functions, for example, the *ops (Operon Polarity Suppressor)* pause, in *E. coli* is a 12 nt regulatory sequence present in the 5'-proximal transcribed sequence of operons under RfaH control. RfaH is a bacterial transcriptional factor that regulates the expression of secreted toxins such as hemolysin (Bailey et al., 1997). It's been shown that RfaH binds to RNAP but only has an effect on it when interacts with the *ops* sequence in the non-template strand DNA. It was proposed, that the *ops* sequence makes RNAP to backtrack (although more recent data suggests that the *ops* pause could be due to stabilisation of the pre-translocated state through sequence recognition of the RNA:DNA hybrid by RNAP (Bochkareva et al., 2011)) allowing the RfaH factor to interact with RNAP increasing the rate of transcription making RNAP less prone to pause or terminate (Artsimovitch and Landick, 2000) (Figure 9).

.

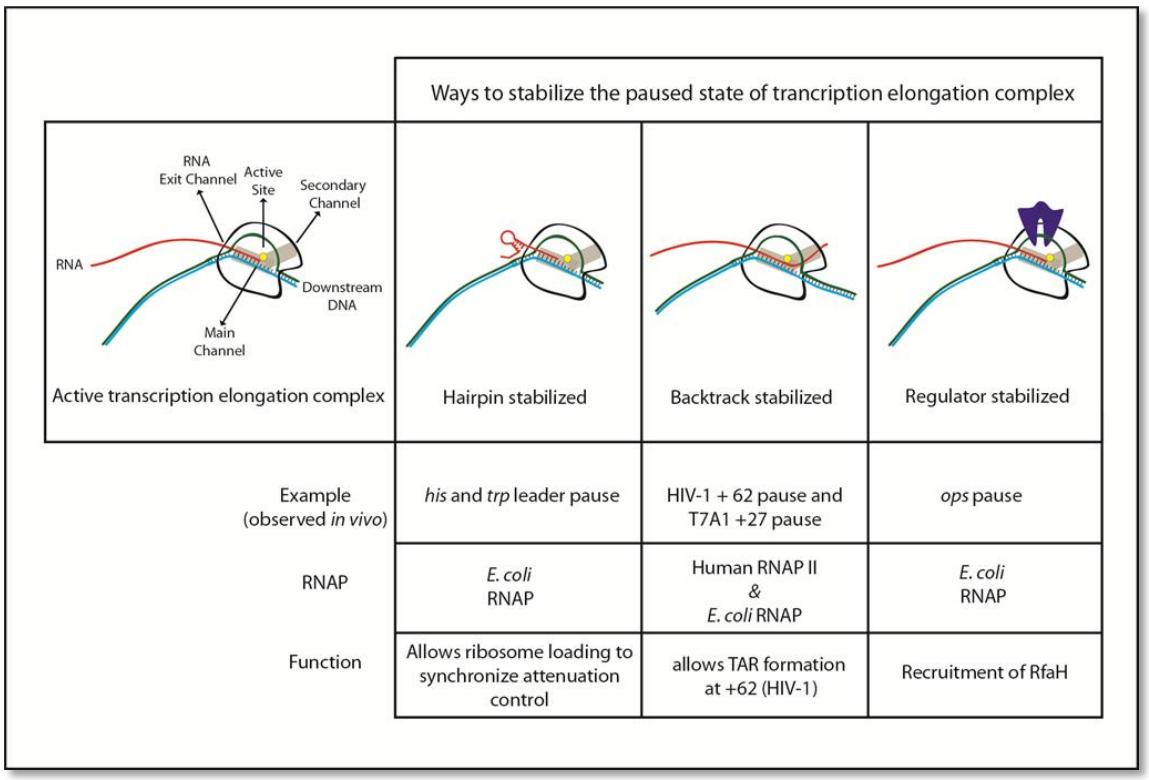

Figure 9 Stabilization of transcription elongation complex in the paused state. Schematic representation of three ways in which RNAP is paused. RNA core enzyme (black claw like structure), nascent RNA (red), template DNA (blue line), non-template-DNA (green line), catalytic Mg^{2+} (yellow circle) are shown.

Pausing is also important for the termination event. Hairpin dependent terminators (intrinsic terminators) are composed of a GC-rich palindromic sequence followed by an oligo T sequence (d'Aubenton Carafa et al., 1990). When RNAP transcribes this sequence, a hairpin-loop structure is formed in the RNA followed by 7-9 uridine residues. This track of poly uridines makes a weak RNA:DNA hybrid, which induces pausing of RNAP, giving time to the nascent hairpin to fold in the RNA exit channel (Gusarov and Nudler, 1999). Currently, there are two prevalent proposed models on how termination occurs: "forward translocation" and the "allosteric (hairpin invasion)" models. In the forward translocation model, the hairpin formed in the RNA forces the EC to translocate forward by 4-5 bp without synthesis of RNA and without the requirement of significant conformational changes in the RNAP; the forward translocation causes the template DNA to rewind at the back and unwind at the front of the transcription bubble whilst the 3' RNA is pulled through the RNA exit channel leading to transcription termination (Santangelo and Roberts, 2004, Yarnell and Roberts, 1999, Larson et al., 2008). The allosteric (hairpin invasion) model, proposes that the formed hairpin in the RNA leads to extensive conformational changes in the RNAP without forward translocation of the enzyme (Gusarov and Nudler, 1999). The conformational changes occur in the DNA, hybrid and RNA binding sites along with

changes in the active centre which render the enzyme inactive; these changes are the result of the hairpin accommodating inside the RNA exit and main channels leading to a quick release of both RNA and DNA (Epshtein et al., 2007) (Figure 10). It's been shown through genomic analysis that intrinsic terminators specify the end of about 50% of annotated protein encoding transcription units and 70% of non-coding RNA transcription units (Lesnik et al., 2001).

It is known that pausing is also important for hairpin independent termination. The hairpin independent termination mechanism rely on ρ protein which is a trimer of dimers with ATPase and RNA helicase activity (Walstrom et al., 1997). There are two proposed pathways for ρ termination: ρ on RNA and ρ on RNAP pathways. In the ρ on RNA pathway, ρ loads onto a stretch of RNA (the rho utilization site RUT) which has little or no secondary structure and a high C content (Richardson and Richardson, 1996). The length of RUT is at least 40 nt long but the minimal required for termination is 80 nt (Nudler and Gottesman, 2002). Once ρ loads onto the RUT, it translocates along the RNA towards the RNAP. In order to catch up with it, ρ has to be faster than RNAP but it has been demonstrated that it is not the case. What allows ρ to catch up RNAP is the pausing signals which halt RNAP (Nehrke et al., 1993); the exact mechanism in which ρ terminates RNAP remains obscure. Some observations postulate that when Rho reaches RNA polymerase, it pulls out the transcript from the active center (shearing or translocation) (Richardson and Richardson, 1996), or that it induces allosteric changes in RNAP which lead to destabilization of the enzyme followed by release of the RNA (allosteric termination) (Epshtein et al., 2010). On the proposed ρ on RNAP pathway, ρ is loaded directly onto RNAP from the beginning of the transcription cycle (as an extra subunit) in the absence of nucleic acid scaffold. After transcription initiation, the nascent RNA chain is wrapped around it. When ρ termination signals are found it either interacts directly with the active centre generating conformational changes that finally lead to termination or it "pulls" the RNA out of the core enzyme (Epshtein et al., 2010) (Figure 10). However preloaded ρ on RNAP is incompatible with the present model of transcription polarity (reviewed in (Banerjee et al., 2006)) in which nascent transcripts that are not simultaneously translated are subject to rho factor-dependent transcription termination, because no mechanism of ρ sensing the lack of translating ribosome was proposed; in the other hand, polarity in the absence of rho factor has been seen mediated

by the overexpression of the protein YdgT in mutant strains where rho has been deleted (Saxena and Gowrishankar, 2011) suggesting alternative mechanism for polarity.

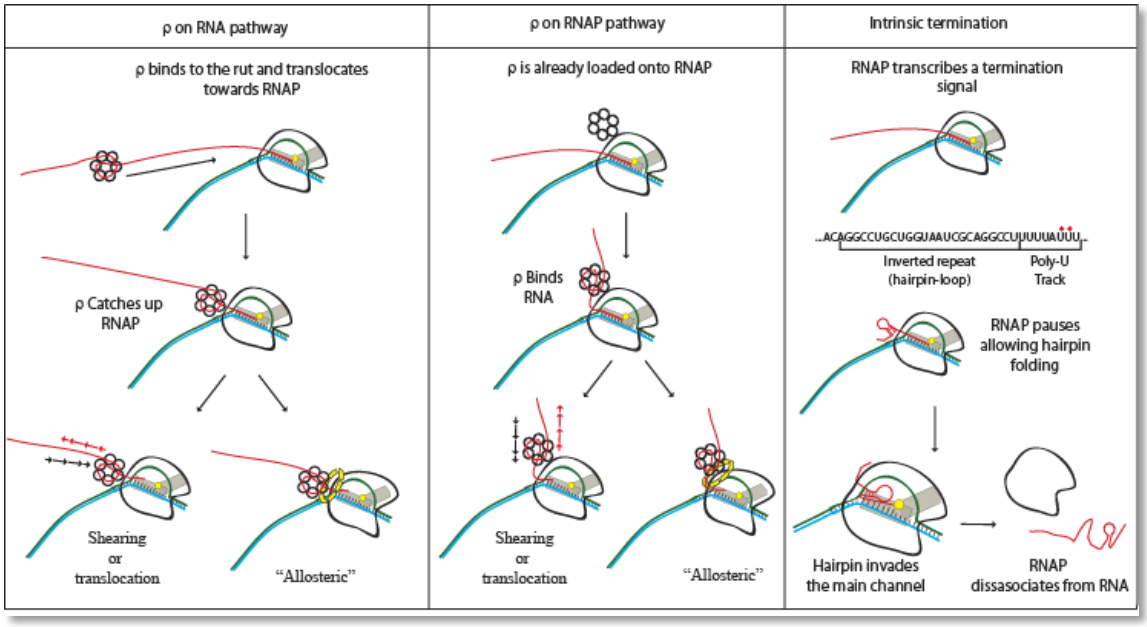

Figure 10. Pathways for termination. Three proposed mechanisms for termination are described. In all panels: RNAP core enzyme (black crab-claw like structure), RNA (red line), template DNA (blue line), non-template DNA (green line), active centre (yellow circle), main and secondary channels (gray paths). First two panels depict the ρ dependent termination proposed pathways and the two possible mechanism of termination (shearing or allosteric mechanisms) the yellow discontinuous oval represents the point where r interacts with RNAP to generate isomerisation of the core RNAP which may lead to release of the RNA. Note that the actual interaction region has not been yet elucidated. Third panel represents the intrinsic termination model. The termination sequence shown is that of TR2 terminator.

1.3. Interactions of RNA polymerase with other cellular machineries

1.3.1 *RNAP-RNAP interactions*

Bacteria utilize a wide range of regulatory mechanisms to control gene expression. Lately, researchers have focused on the interactions of RNAP with other machineries in order to further understand how the systems are interrelated. The various properties of RNAP complexes with nucleic acids during different stages involve various types of regulation and different cross-talk with other cellular entities and with RNAP itself. In prokaryotes and eukaryotes, the genome is arranged into condensed structures. This organisation is due to, in its majority, to histone-like proteins and histones respectively. Besides the continuous presence of such nucleoproteins, a vast number of DNA binding proteins are also present throughout the cell cycle. Notoriously, the rate of transcription is not altered when RNAP is challenged with these obstacles on the DNA *in vivo.* Interestingly, the opposite effect was observed *in vitro* when RNAP encountered roadblocks such as the lac repressor, LexA protein, and DnaA positioned downstream a promoter (Pavco and Steege, 1990). In 2003, Epshtein *et al*, observed that *in vitro* efficient transcription through roadblocks depended on multiple rounds of transcription

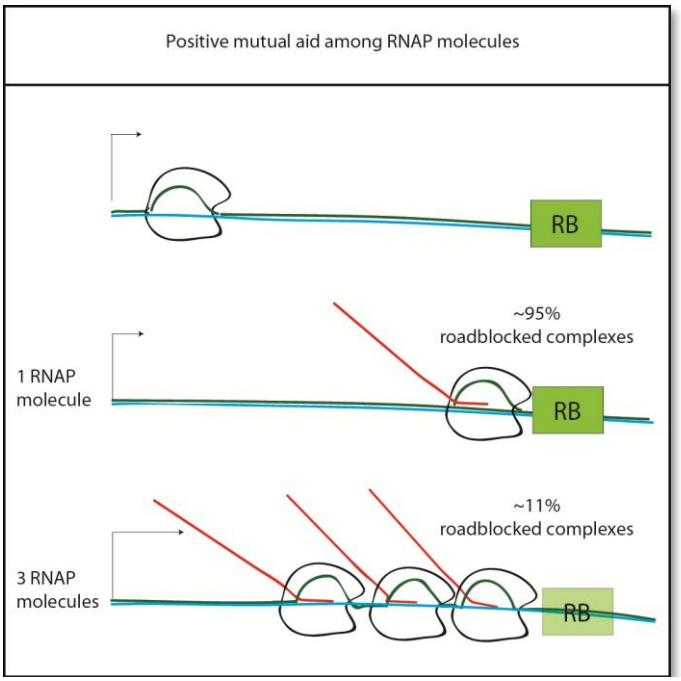

Figure 11. Cooperative effect among RNAP molecules. Schematic representation of the cooperative mechanism of RNAP molecules. Template DNA (blue line), non-template DNA (green line), RNA (red line), promoter (arrow) and roadblock (green box with RB written inside) are shown. When 1 RNAP molecule is allowed to collide with a DNA bound roadblock, few molecules are capable to pass through it. When three RNAP are allowed to initiate and to collide, a cumulative effect takes place which allows greater read-through of the roadblock.

initiation (Epshtein et al., 2003) (Figure 11). In their study, the EcoRI mutant EcoRQ111 which is a mutant version of the EcoRI restriction enzyme which is capable to bind to the DNA but it is unable to restrict it, was used as a roadblock. When only one RNAP molecule was allowed to collide with the roadblock, no transcriptional read-through was seen and the outcome of such collision was RNAP backtracking by 4 nt. When a second RNAP was allowed to transcribe and collide with the already roadblocked complex subtle read through of the roadblock was observed. The outcome of the second RNAP molecule colliding with the first (backtracked and roadblocked) RNAP molecule was the repositioning of the backtracked complex into the post-translocated state (which can be judged by GreB cleavage pattern), suggesting that the second RNAP "pushed" the first one rescuing it from the backtracked state. Upon transcription initiation and translocation of a third RNAP molecule towards the previously roadblocked RNAPs full dislodge of the EcoRQ111 roadblock was observed suggesting a positive mutual aid among RNAP molecules. Similar observations were obtained using the lac repressor instead of the EcoRQ111 mutant demonstrating that the effect was not depending on the binding properties of the roadblock and suggesting that a cooperative model amongst RNAP molecules are responsible for the continuous rate of transcription elongation *in vivo*.

.

1.3.2 *RNAP-replication interactions*

The fact that RNAP is actively transcribing on the DNA at the same time as the chromosome is being replicated implies that collisions between the two machineries are often inevitable and can interfere with replication fork progression. The effect on collisions between RNAP and the replisome depends greatly in their directionality. Co-directional interactions occur on the leading strand whereas head-on collisions take place on the lagging strand. Interestingly, analyses of genome organisation have shown that most of the essential genes, highly transcribed genes, and longer genes (Price et al., 2005, Omont and Kepes, 2004, Huvet et al., 2007, Rocha and Danchin, 2003) are oriented in the same direction as replication on the leading strand (McLean et al., 1998). Although the evolution pressure that resulted in this organisation is still unclear, it is thought to be, at least in part, determined by differences of the interactions of RNAP with the replication fork during co-directional versus head-on collisions (Brewer, 1988). This could be detrimental for RNAP completion of transcripts affecting the production

of correct, full length proteins or most likely, because it could stall the replisome inhibiting the replication fork progression (Deng et al., 2005, Srivatsan et al., 2010). Such collisions have been documented *in vivo* and *in vitro* in *E. coli* (French, 1992). Effects of such collisions have been observed to have a measurable impact on both replication initiation and completion of the nascent DNA and on transcription (Pomerantz and O'Donnell, 2010). Regulatory functions of transcription on replication initiation have also been reviewed in (Castro-Roa and Zenkin, 2011). Other studies about how replication fork progression is affected by transcription-initiation complexes and transcription termination signals showed that transcription-initiation complexes inhibit replication fork progression during head-on collisions (Mirkin and Mirkin, 2005, Mirkin et al., 2006). Transcription termination signals also appeared to attenuate replication but in the opposite co-directional orientation. Most recently, Pomerantz and O'Donnell studied the biological role of co-directional interactions of the RNAP and DNAP. In their paper, they demonstrated that after collisions, the replisome uses the RNA transcript from the displaced RNAP, as a primer to continue the leader strand synthesis (Pomerantz and O'Donnell, 2008). RNAP acting as a primase was also described by N. Zenkin and colleagues in 2006 (Zenkin et al., 2006a). The different outcomes of the interactions between RNAP and the replisome are summarized in Figure 12.

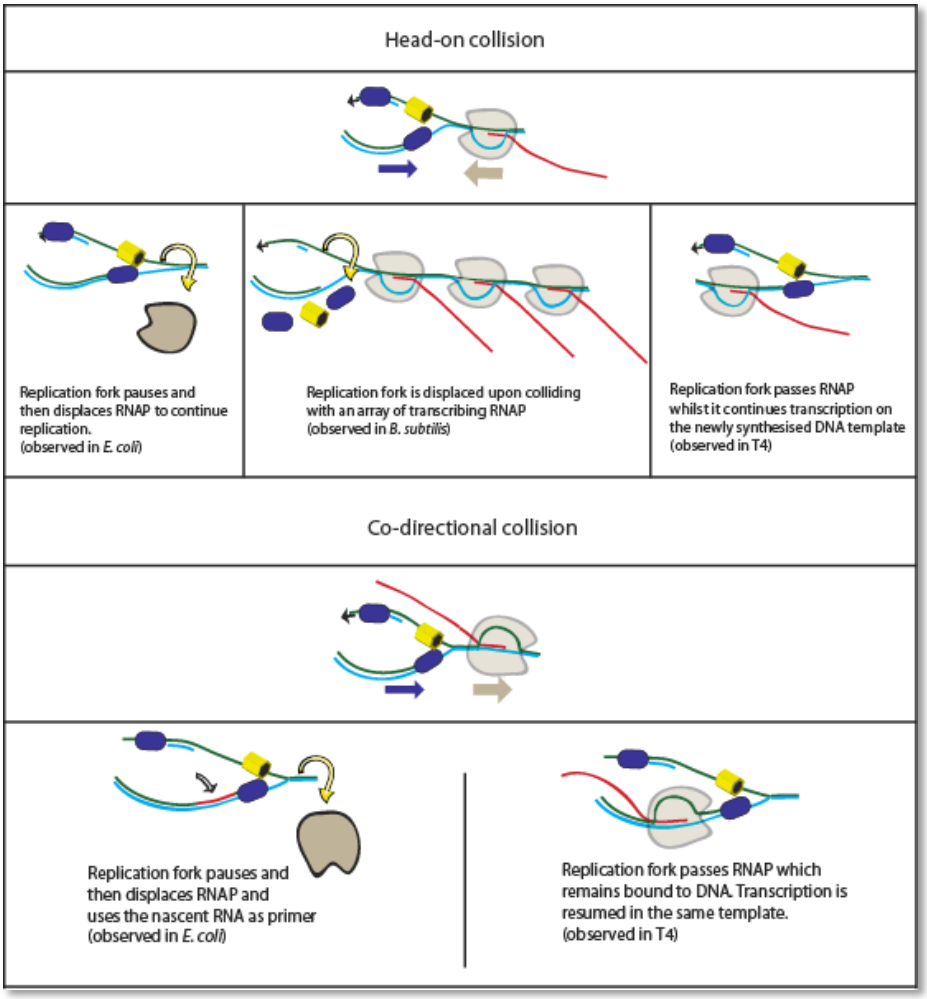

Figure 12 Outcomes of the interactions between RNAP and the replisome depending in their orientation. Daniel Castro-Roa and Nikolay Zenkin (2011). Relations Between Replication and Transcription. Fundamental Aspects of DNA Replication, Jelena Kušić-Tišma (Ed.), ISBN: 978-953-307-259-3 Book chapter.

1.4. Translation

Translation is the process in which the messenger RNA is “read” and “interpreted” into proteins. It is thought that almost half of the dry weight of prokaryotic cells and more than 70% of its energy is used to synthesize proteins. The main component in translation is the ribosome: a massive multisubunit machine with a mass greater than two million Daltons composed of more than fifty different proteins that make up the 40% of it and the remaining 60% is rRNA. The ribosome is considered to be the most ancient enzyme; its function and structure are highly conserved in all life forms (Harris et al., 2003). The ribosome is in charge of both reading the RNA and the synthesis of peptides. In addition to the ribosome, mRNA, tRNA and aminoacyl-tRNA synthetases complement the translational apparatus. In its simplest form, the mechanism of

translation must account for three fundamental events: codon-anticodon recognition, peptide bond formation and movement of tRNA and mRNA relative to the ribosome.

1.4.1. *The prokaryotic ribosome*

The ribosome consists of two subunits in all species. In bacteria, the subunits have been designated as 30S and 50S, and together make up the 70S ribosome (Figure 13). The small subunit is composed of twenty-one proteins assembled with 16S rRNA (Held et al., 1974). The large (50S) subunit is made up by the 23S rRNA, 5S rRNA and over 30 proteins. The association of the 30S and 50S is formed through a network of intermolecular non-covalent bridges.

The intersubunit space is occupied by tRNAs, whose anticodons base-pair with messenger RNA (mRNA) in the 30S subunit, whereas their 3`-CCA ends, which carry the growing polypeptide chain and the incoming amino acid, reach into the 50S subunit, the location of the peptidyl transferase centre, where peptide bond formation is catalysed (Yusupov et al., 2001). Interestingly, all fundamental functions of ribosome are catalyzed by RNA (Nissen et al., 2000).

In order to decipher the genetic code, the ribosome uses tRNAs; in the cell, tRNA exists in three forms: aminoacyl-tRNA, peptidyl-tRNA and deacylated tRNA. These forms differ one from the other by their acylation state, if an amino acid, peptide or no group is attached to the 3' end of tRNA. As it is to expect, the acylation state determines the specificity in which the ribosome would use it. The ribosome provides tRNA access to mRNA codons through unique tRNA binding sites: The A site recognizes the aminoacyl-tRNA and positions the aminoacyl moiety for the peptidyl transferase reaction. The P site that binds peptidyl-tRNAs, and finally the E site that has

affinity for deacylated tRNAs (Green and Noller, 1997)(Clemons et al., 1999, Ban et al., 1998). In the P site the initiator tRNA binds (which in bacteria, is formyl methionyl-tRNA, or F-Met-tRNA^{fMet}) during the initiation of protein synthesis, keeping the proper translational reading frame and preventing loss of nascent polypeptide chain and also binds the peptidyl-tRNA during translation initiation. In contrast to the A site, the P site

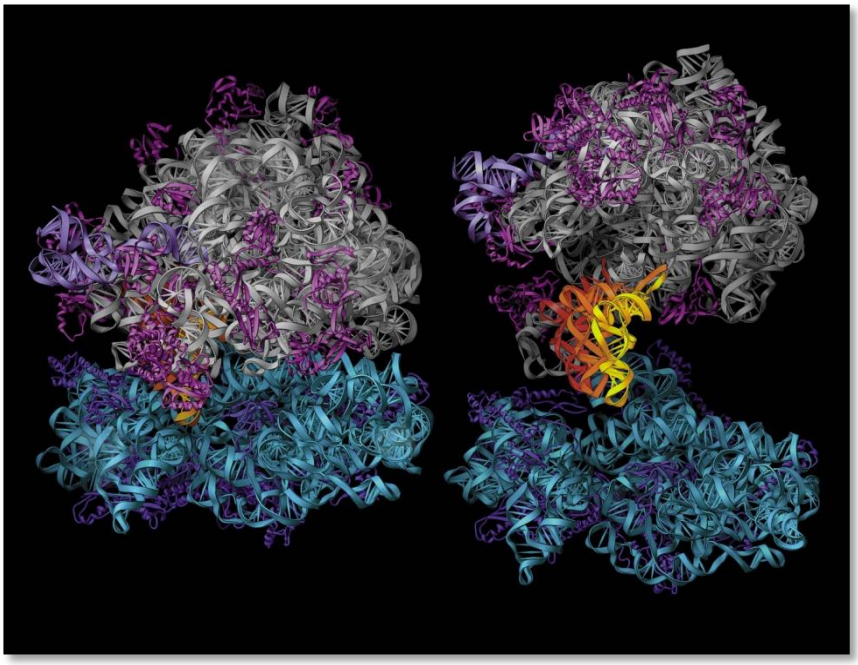

Figure 13. Crystal structure of *T. thermophilus* 70S ribosome. The 70s Ribosome (left) and the detached subunits (right) with tRNAs in the A-site (yellow) P-Site (orange) and E-Site (red).63

has several interactions between the small subunit and tRNA whereas in the A site contacts the tRNA and mRNA with only four nucleotides of 16S rRNA (Clemons et al., 1999). Those minimal interactions between the 30S A site and the third-position codon and anti-codon nucleotides explain the prevalence of third-position wobble in the genetic code; the acceptors ends of A site an P site are located in a the 50S subunit where the peptidyl-transferase centre is. Based on biochemical and genetic studies, 23S rRNA has long been suspected to be the functional component of the peptidyl transferase centre (Mougel et al., 1987). This was demonstrated when the high-resolution structure of the 50S subunit showed that there is no protein within 18 Å of the site of catalysis in the *Haloarcula marismortui* ribosome (Garrett, 1987). After peptide bond formation, the deacylated tRNA translocates form the P site to the E site. The E site provides a favorable free-energy change for movement of the deacylated tRNA out of the 50S P site (Sergiev et al., 2005). This movement occurs concomitantly with the translocation of the aminoacyl-tRNA occupying the A site to the P site for

catalysis to take place. The mechanism on which the ribosome initiates and performs translation elongation will be described below.

1.4.2. *The translation cycle*

Just like transcription, translation is a cyclic event and also can be divided into three main stages: initiation, elongation and termination (Figure 14).

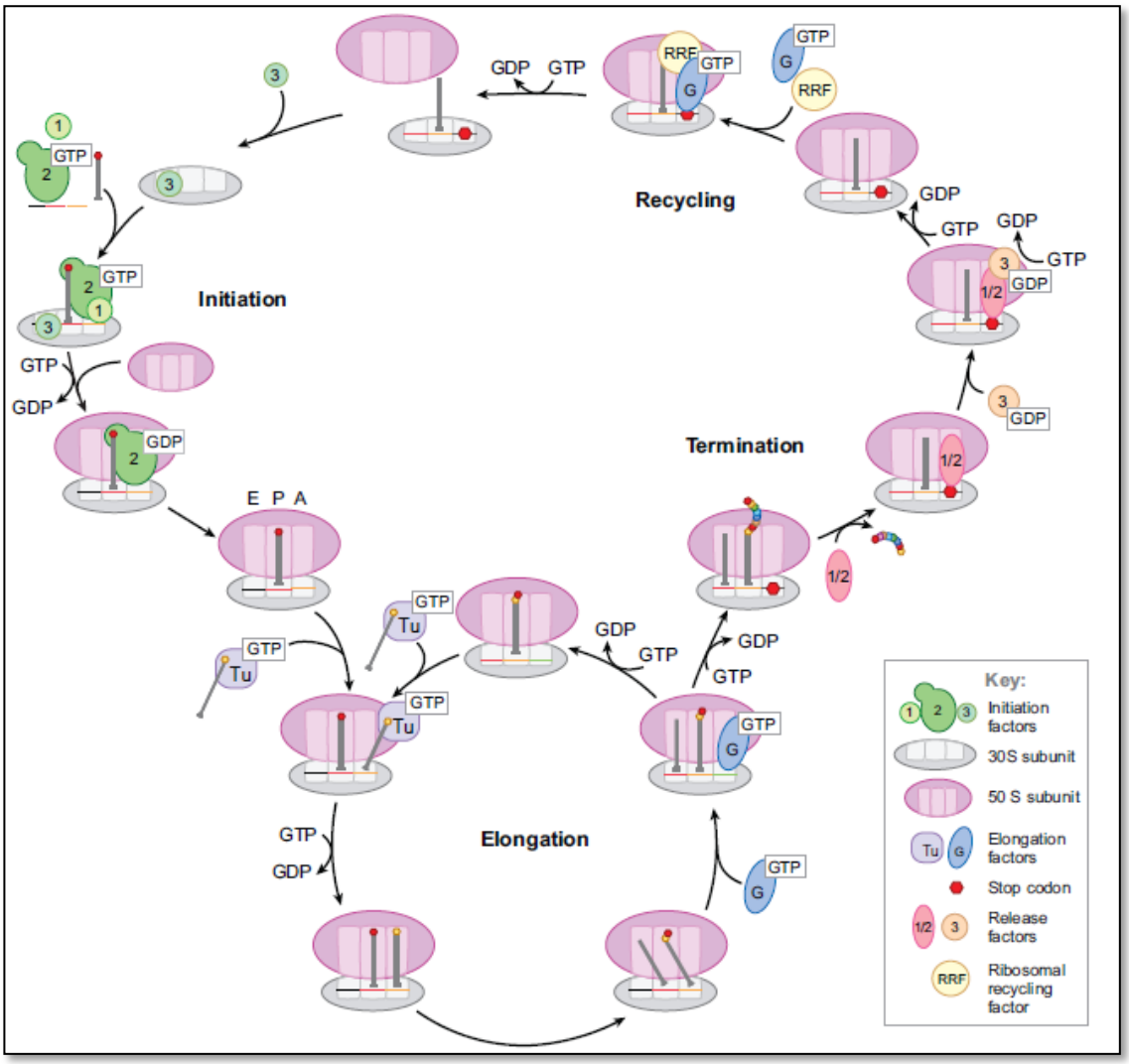

Figure 14. The transcription cycle. Representation of the three stages of transcription: initiation, elongation and termination.

1.4.2.1. *Translational initiation*

The translation initiation reaction plays a major role in gene expression. It starts with the formation of a complex between the small ribosomal unit (30S) with the Shine-Dalgarno sequence in the mRNA (complementary to the 3' end of 16S) and the initiator tRNA (F-Met-tRNA^{fmet}). At least three initiation factors IF1, IF2 and IF3 are required for efficient translation initiation (Laursen et al., 2005). IF3 binds to the 30S subunit of the ribosome and promotes disassociation into 30S and 50S, coupling ribosome recycling with translation initiation. IF1 binds specifically to the base of the A-site in the 30S subunit and directs the initiator tRNA to the ribosomal P site by occupying the A-site. It also interacts with IF3 supporting the disassociation of the ribosome. Once the ribosome is dissembled into two subunits, IF2, mRNA and F-Met-tRNA^{fmet} bind to the 30S unit. Shine-Dalgarno sequence interacts with 16S and the initiation codon is adjusted to the P-site of the ribosome (Pon et al., 1985). It has been shown that both IF2 and IF3 directs the positioning and the stability of the initiator tRNA in the P-site. Once the binding of the 30S subunit to the mRNA is stable, initiation factors 1 and 3 are expelled from the complex, while IF2 stimulates the binding of the 50S subunit. When the initiator tRNA is properly bound to the 50S P site, IF2 is released in a GTP dependent manner. At this stage, initiation has been fully achieved and the ribosome is ready to enter into elongation (Figure 15) (Yusupova et al., 2001). *In vitro* translation initiation can be artificially conducted using the peptidyl-tRNA analogue N-acetyl-met-tRNA^{fmet} which is capable of binding to the P site of the ribosome in the absence of initiation factors; therefore it is known as non-enzymatic initiation.

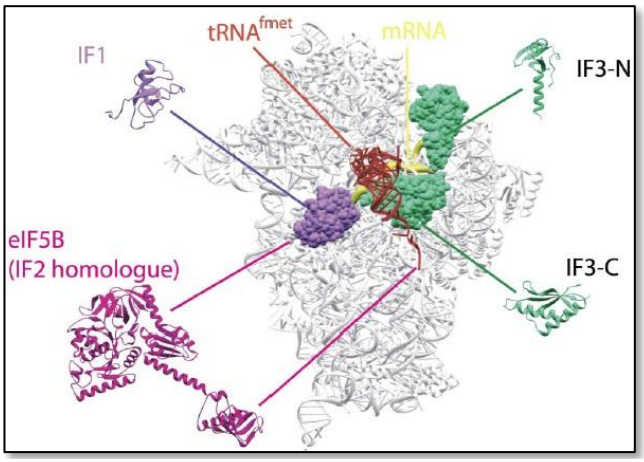

Figure 15. The 30S subunit and its interactions with initiation factors. The structures of IF1 (magenta), IF2 (purple) and IF3 (green) are shown interacting with the 30S subunit (gray). IF1 and IF3 orientation and interactions are shown on the 30S (magenta and green space-fill) and for easier observation outside in ribbons, IF2 interactions are indicated; fmet-tRNA^{fmet} (red) occupying the subunit's P site is shown. Note that the crystal structure of a IF2 homologue (eIF5B) was used *Ramakishnan, V Cell 109,557-572,2002*

1.4.2.2. *Translation elongation and termination*

The cycle of elongation is the central event in protein synthesis. It begins with a peptidyl-tRNA in the P-site and an empty A-site, where the next codon to be translated is exposed. If the exposed codon corresponds to a sequence which codifies an amino acid and not a stop codon, the A site has a high affinity for an aminoacyl-tRNA in complex with the elongation factor Tu (EF-Tu-GTP). During initial selection, a ternary complex of EF-Tu-GTP and aminoacyl-tRNA binds reversibly to the ribosome. The action of EF-Tu depend whether GTP or GDP is bound to it, attributing to GTP hydrolysis a regulatory function, being EF-Tu-GTP the active form. If the aminoacyl-tRNA does not match with the codon in the A site, the aminoacyl-tRNA will dissociate and return to the initial state. In correct match, the codon-anticodon interaction triggers the hydrolysis of the bound GTP. Because of the low affinity for aminoacyl-tRNA and the ribosome, EF-Tu-GDP is released, allowing the positioning of the aminoacyl-tRNA and the formation of the peptide bond between the amino group of it and the C-terminal ester group of the peptidyl-tRNA located at the P-Site. This state, where both A-site and P-site are occupied, is known as pre-translocated state. In order to expose an empty A site to the following codon, the ribosome has to translocate. Translocation involves the precise and coordinated movement of two tRNAs, the tRNA in the P site to the E site, and the one occupying the A site to the P site by one codon. In bacteria, translocation is catalyzed by Elongation Factor G (EF-G) in a GTP dependant manner and requires both the P and A sites to be occupied (Spirin, 1985, Kaziro, 1978).

Translocation of the ribosome can be subdivided into two different stages. The first stage which occurs spontaneously and independent from EF-G and GTP, is the movement of the acceptor ends of tRNA relative to the 50S subunit following peptide bond formation. The second one, which is GTP and EF-G dependent, is the translocation of the anticodon arms of the tRNAs along with the mRNA relative to the 30S subunit. After EF-G binding, translocation takes place and the GTP is hydrolyzed into GDP. EF-G is released leaving the ribosome ready for the incoming aminoacyl-tRNA (Holschuh et al., 1980) (Figure 16).

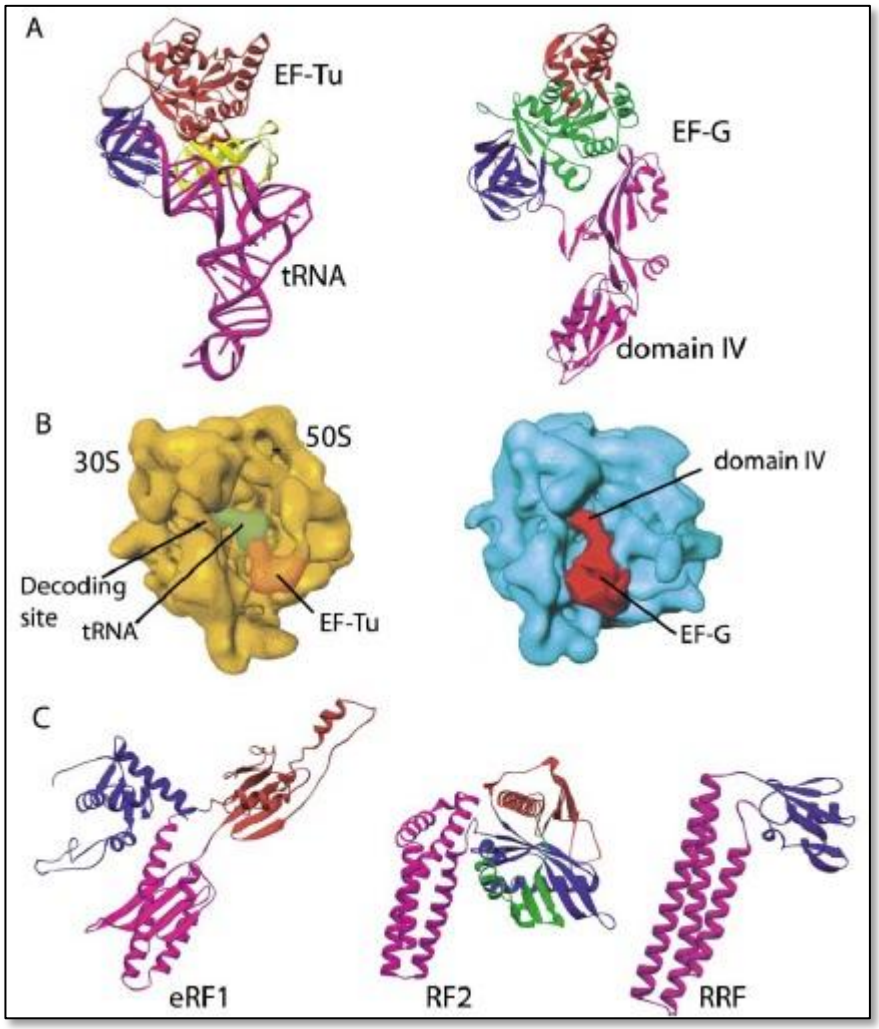

Figure 16. Crystal structures of translation elongation and termination factors. In all structures coloured in magenta represent the domain which mimics the tRNA when interacting with the ribosome A) Ternary complex: EF-Tu (red) in complex with tRNA (magenta) and GdPNP (yellow) and EF-G. B) CryoEM (right). C) Structures of release factors and ribosome recycling factor. *Ramakishnan, V Cell 109,557-572,2002*

When the exposed codon in the A site is a stop codon, one of the release factors (RF1, RF2) with GTP, will bind leading to peptide release. The termination is completed when the ribosome recycling factor along with EF-G and IF3 stimulate the dissociation of the ribosome, the deacylated-tRNA and mRNA to recycle the translational machinery (Weixlbaumer et al., 2008).

1.4.3. Ribosomal proteins utilized by both the transcriptional and translational machineries.

Research on the identification of shared components of the transcriptional and translational machineries has pin-pointed three ribosomal proteins (r-proteins) involved in transcription anti-termination and on regulation of transcription initiation of rRNA operons. Transcription anti-termination is a process in which RNAP is allowed to read-through specific RNA secondary structures which normally would terminate transcription. Such event is mediated by a protein or a set of proteins which interact with RNAP and the termination hairpin. Throughout mutational analysis on *E. coli* to find host factors involved in the bacteriophage lambda N-protein-directed transcription

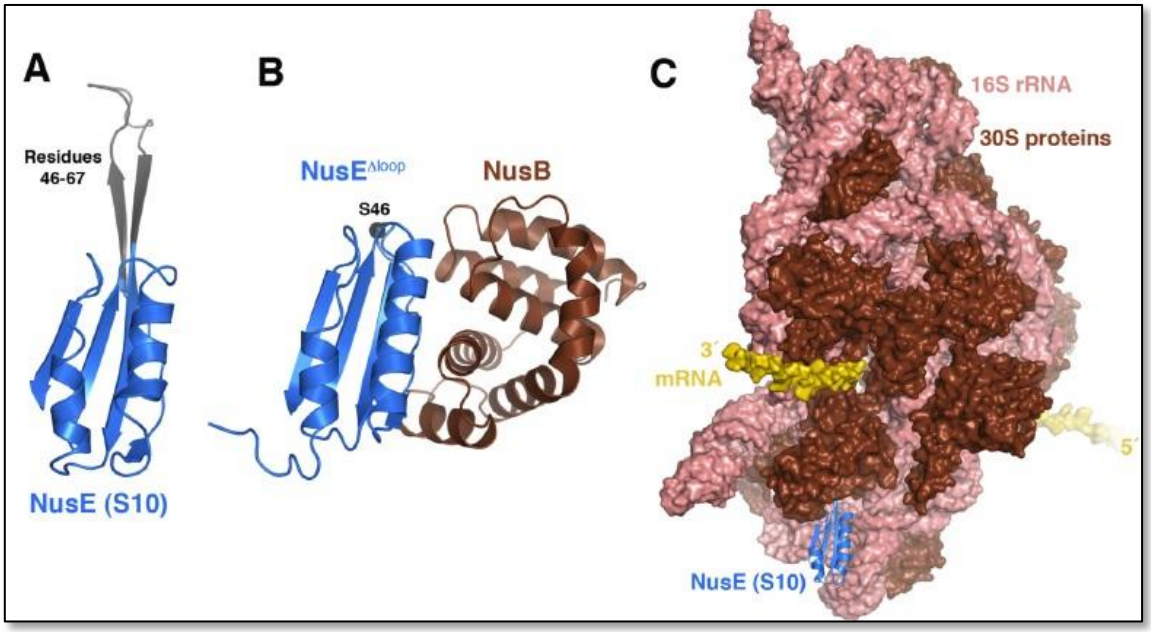

Figure 17. Crystal structures of NusE (S10), NusB and ribosomal small subunit 30S. A) Structure of S10. Residues 46-67 (grey) interact with 30S ribosomal subunit. B) Crystal structure of S10 and NusB. Residues which interact with the ribosome have been deleted. The black dot (S46 residue) replaces the ribosome-binding loop. C) 30S ribosomal subunit. S10 r-protein (blue ribbon cartoon) is orientated with the NusB binding domain exposed. 16S rRNA (pink) other r-proteins (brown) and mRNA (yellow) are shown**.** Modified *from Björn M. Burmann, et al. Science 328, 501 (2010);*

antitermination it was found that NusE (S10 r-protein) defective mutant was deficient in N-antitermination (Friedman et al., 1976). It has been shown that S10 is involved in the proper assembly and maturation of the small ribosomal subunit (Green and Noller, 1997) therefore it is essential for cell survival. *In vitro* studies have demonstrated that the involvement of S10 in antitermination does not require its presence in complex with the entire ribosome or the 30S subunit alone (Das, 1993, Mason and Greenblatt, 1991). It has been described that S10 needs in order to participate in lambda transcription

antitermination is a complex interaction network with the N protein, NusA, NusG and NusB with RNA elements BoxA and BoxB (Figure 17).

These elements form a stable complex which interacts with RNAP causing antitermination. Although there is no direct evidence on the involvement of S10 on the antitermination of rRNA it has been suggested that S10-NusB interactions may aid the stabilization of antitermination complexes acting on both rRNA and mRNA terminators, proposing S10 as a general antitermination factor (Mason and Greenblatt, 1991).

Most recently, through nuclear magnetic resonance spectroscopy (Burmann et al., 2010), it has been observed that S10 is capable to interact with the carboxyl-terminal domain of NusG and that NusG interacts through its amino-terminal domain with RNAP. Therefore it has been suggested that S10 links translation to transcription through NusG which could act as a bridge between the two machineries (Figure 18)

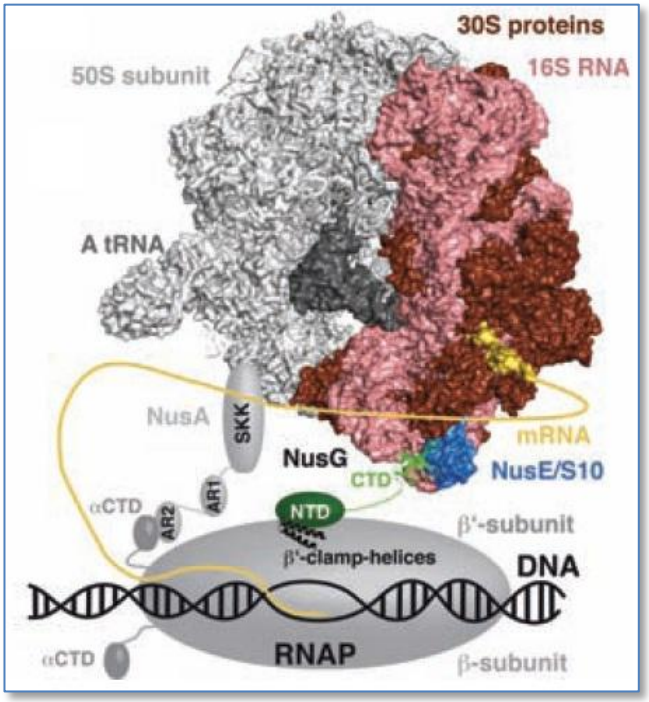

Figure 18. S10-NusG-RNAP bridging model. Representation of the 70S ribosome being linked to RNAP through interactions between S10 and NusG. Ribosome 50S subunit (light gray space-fill) with tRNA (dark gray space-fill) occupying the A site is shown. 16S rRNA (pink) interacting with mRNA (yellow) and r-proteins (brown) is shown. S10 (blue space-fill) is shown interacting with NusG carboxyl-terminal domain (light green) whilst the amino-terminal domain (dark green oval) interacts with RNAP β'-clamp-helices. RNAP (gray oval) is depicted and its subunits shown along with DNA double helix and transcription bubble with nascent mRNA (gold line). NusA (grey labelled oval) interacting with alpha carboxyl-terminal domain (αCTD) domain is proposed to act as a stabilizer of the bridged system. Modified *from Björn M. Burmann, et al. Science 328, 501 (2010);*

Up to this date, there is no biochemical evidence that proves the bridging between the ribosome and RNAP through S10-NusG interactions.

It has been observed that r-proteins are also involved in the transcriptional regulation of rRNA. The majority of r-protein-encoding operons are regulated to keep the levels of r-proteins directly proportional to the levels of rRNA produced in the cell. Therefore it is to expect that the regulator of an r-protein operon is an r-protein encoded in the same operon (Gourse et al., 1996). The S10 operon is composed of 11 genes. Among those 11 genes, 1 gene encodes the r-protein L4 which is involved in the transcriptional and translational regulation of the entire operon (Olins and Nomura, 1981, Zengel et al., 1980). It has been observed that L4 in collaboration with NusA induces transcription termination. The mechanism in which the regulation occurs involves a 172 nt long leader region which can fold to form 5 hairpin-like structures. The first 3 hairpins can be deleted without altering the L4 mediated termination. The fourth hairpin has been

identified as the main precursor for termination to take place on the fifth hairpin. Biochemical studies have shown that RNAP pauses after transcribing the fifth hairpin and this pause is stabilized through NusA. Then, L4 interacts with the fourth loop and further stabilizes the pause leading to termination of transcription (Sha et al., 1995). The actual mechanism in which L4 interacts with NusA and with RNAP and whether or not it acts as a termination factor for other operons remains obscure.

Another possible r-protein which is shared by the translational system is S1. S1 is one of the largest r-proteins in the ribosomal small subunit (70 kDa) and it plays a role in translation initiation. Although it has not been seen interacting with rRNA, its function involves interaction and melting of highly structured mRNA facilitating ribosome binding. S1 has been described as a functioning part of the phage Q β-replicase and as a part of the lambda recombination protein β. Interestingly, S1 has also been observed to have a function in transcription antitermination (Mogridge and Greenblatt, 1998). Throughout competition assays it was shown that S1 displays a 200 fold higher affinity for the ribosomal RNA antitermination BoxA sequence than NusB-S10. To this date evidence for direct participation in rRNA or lambda N-mediated antitermination has not been described.

1.4.4 *RNAP-translational machinery interactions*

The interaction of transcriptional apparatus with the translational machinery has been reviewed over two decades focusing mainly in terms of gene expression regulation rather than in the physical interaction of the ribosome and the RNAP. In 1985 Robert Landick described how translation aids paused RNAP to restore its elongating state in the *trp* operon (tryptophan biosynthesis pathway) leader region. His findings show that the hairpin-dependant pause at position +92 (relative to the transcriptional start site) is diminished as the ribosome synthesizes the leader peptide; it was observed that binding of the ribosome to the mRNA generated changes in the secondary structure of the mRNA thus melting the hairpin allowing RNAP to escape from the paused state. These observations postulated the idea that in the *trp* operon transcription and translation are tightly coupled to attenuation control (Landick et al., 1985).

Another example of ribosome mediated gene expression is the case of the regulation of the tryptophanase (*tnaA)* operon of *E. coli*. The operon contains two major structural genes, a promoter-proximal gene, *tnaA* which encodes tryptophanase and a distal gene *tnaB* that encodes a low affinity tryptophan permease. *tnaC* is a regulatory transcribed region that encodes a 24 amino acid (with a high content of tryptophan residues) peptide leader localized upstream of *tnaA*. From *tnaC* stop codon downstream to the start codon of *tnaA* there are 220 nucleotides that contain several transcriptional pause signals that can be used by the termination factor ρ to load onto the RNA (Stewart et al., 1986). The induction of the entire operon depends on both the TnaC synthesis and high amounts of free tryptophan. The combined action of the uncleaved nascent TnaC-peptidyl-tRNA (retained in the translating ribosome) and ribosome-bound tryptophan inhibits peptidyl-transferase cleavage of the nascent TnaC-peptidyl-tRNA at the *tnaC* stop codon (Gong et al., 2001). Thus, the uncleaved TnaC-peptidyl-tRNA remains bound to the ribosome stalling it at the stop codon. The stalled ribosome blocks access of rho factor permitting RNAP to resume transcription; in the absence of tryptophan transcription is subject to ρ dependent termination (Gong and Yanofsky, 2003).

Besides preventing the loading of the termination factor ρ, transcription – translation coupling is also considered important to prevent the formation of R-loop structures. An R-loop is a structure in which RNA forms a heteroduplex with one strand of the DNA upstream of the transcription bubble (Massé and Drolet, 1999). The formation of this

structure is detrimental or even lethal for bacteria because they act as roadblocks for the following RNA polymerases (Hraiky et al., 2000). It is important to underline that the effects described above do not involve direct, physical interaction of the two machineries.

Lately, research has also been done on the mechanistic effects of the ribosome-RNAP interaction *in vivo*. It has been proposed that the rate of transcription depends on the rate of translation. This was observed using 70S ribosome mutants which are slower than the wild type counterpart. Under such conditions, transcription was observed to be slowed down dramatically. Another observation in the same study was that the translation apparatus could aid backtracked transcriptional complexes (Dutta et al., 2011, Proshkin et al., 2010). In a previous study, it was shown that when RNAP collides with the *lac* repressor (and with other roadblocks such as EcoRI mutant variants which are able to bind to DNA but unable to restrict it), RNAP backtracks; backtracking of the roadblocked RNAPs was overcome by actively transcribing RNAP molecules colliding with the backtracked RNAP, suggesting a cooperative mechanism amongst RNAP which permits rescuing of stalled transcriptional complex (Epshtein and Nudler, 2003). Therefore, Proshkin and co-workers decided to investigate if the ribosome could generate the same effect on backtracked RNAP. The effect on the recovery of backtracked RNAP was performed by initiating transcription from constitutive promoter, allowing transcription elongation until RNAP would be roadblocked by the DNA-bound *lac* repressor; the position of RNAP on the DNA with and without translation was mapped by *in vivo* footprinting. Through this technique, they observed that the ribosome prevented RNAP from further backtracking and presumably “pushed” it forward. The mRNA synthesis of this experiment was followed by reverse transcription; through this technique they observed that the ribosome was not only capable of rescuing RNAP from its backtracked state but it also exerted enough force onto the RNAP allowing it to overcome the *lac* repressor. (Proshkin et al., 2010). Recently, it has been proposed that backtracked RNAP can cause DNA double-strand breaks (DBS), a factor which impedes replication fork progression leading to genome instability (Dutta et al., 2011). It was observed that the ribosome impeding backtracking and also rescuing RNAPs stalled in the backtracked state links transcriptional backtracking to translation in order to maintain genome stability.

Unfortunately, no mechanism can be concluded from these observations due to the fact that it was done *in vivo*. It is quite possible that the described effects are not only due to the direct physical contact between the two machineries but it could also be attributed to factors (both known and unknown) generating those effects.

It is becoming clearer that big, complex molecular machineries act in concert. The understanding of the interactions among them may not only complement what we know today about their cross-talk but may also elucidate new regulatory pathways which will serve to understand the cell as a dynamic entity.

The lack of knowledge on the outcomes on both transcription and translation upon actual physical and mechanistic interactions was the determining point that made us undertake the task of designing a "stripped-down", pure, transcription coupled to translation system, with the capacity of performing step-wise translocation of both apparatuses for the analysis of effects in both machineries with minimized background noise from unwanted regulators. However, the system should permit the addition of such unwanted regulators to determine if they could play an essential role on the cross-talk amongst the machineries.

Chapter 2. Aims

After reviewing the published data and conclusions that exist today on the interactions of RNAP with other cellular machineries particularly the cross-talk of the transcriptional machinery with the translational apparatus, it is clear that many aspects remain obscure. The observed mechanistic interactions have been performed *in vivo* using indirect approaches (such as *in vivo* DNA footprinting of RNA polymerase and reverse transcription)(Dutta et al., 2011, Proshkin et al., 2010); even-though the results obtained in such experiments are valuable, it cannot be disregarded that the observed behaviour of RNAP on backtracking and roadblocking after colliding with the translating ribosome could not be due to transcriptional factors present in their *in vivo* system. Furthermore, the present studies look only at the outcome on the backtracking phenomenon, leaving different pausing mechanisms such as pre-translocated stabilised pausing without characterisation. Another subject which needs to be elucidated is the actual interface amongst transcription and the translational machinery. Efforts to evaluate this topic have been done using nuclear magnetic resonance and molecular modelling but no biochemistry has been performed due the lack of a system to do so (Burmann et al., 2010). Understanding the interface needed for the correct cross-talk between RNAP and the ribosome could be of highly importance for the development of antibiotics which target such interactions. The actual data also lacks information about the effect on the ribosome after colliding with RNAP in different elongation complexes; it could be possible that strong paused RNAP could be a regulatory signal for the ribosome translating behind it, for instance, the possibility of translational frameshifting caused by collisions with RNAP have not been addressed; it is becoming more obvious that an intricate interplay amongst the transcriptional and translational machineries is an extremely important regulatory mechanism in bacteria. For its complete characterisation and understanding, the design of an *in vitro*, pure system, where the translocation of the ribosome towards RNAP and the positioning of RNAP in different transcriptional states can be controlled allowing the analysis of the possible effects of the interactions of the ribosome on transcription became the focus of my Ph.D. studies.

The specific aims of the project are:

1) Design an *in vitro* transcription – translation system complying with the following characteristics:

a. It needs to be assembled from purified components in order to reduce possible interference in the interpretation of interactions between the machineries caused by known and unknown factors.
b. It needs to permit step-wise control over the translocation of the ribosome. Having the capacity of "walking" the ribosome in a stepwise manner towards RNAP will permit the collection of data at different distances among them before collisions take place.
c. Have simultaneous step-wise control of the translocation of RNAP. By capable of "walking" RNAP on the DNA will allow us to place it in any desired transcription elongation complex such as paused, backtracked or roadblocked.
d. The system needs to permit the analysis of the outcomes of the interactions on both transcription (by following the mRNA synthesised) and on translation (by following the nascent peptide synthesised).

2) By using the developed system to assess the effects of the interactions of the ribosome and RNAP whilst in different, defined transcription elongation complexes (paused, backtracked, etc.)

3) Assess the system for its effectiveness on the analysis of the outcome on translation after colliding with RNAP.

Chapter 3. Materials and Methods

3.1 Transcriptional machinery

3.1.1 *Purification of RNAP*

The purification of His-tagged RNAP (strain RL721, made and provided by R. Landick University of Wisconsin, Madison, Wisconsin, U.S.A) and TAP-tagged RNAP (strain RpoBtap, made and provided by A. Emili, Department of Medical Research, University of Toronto, Toronto, Ontario, Canada) was performed exactly as described in (Nudler et al., 2003). Briefly, 5 L of LB are inoculated with 1/100 volumes from a 100 mL overnight culture. The cells are grown in orbital shaker at 37°C until an OD_{600} of 1.3 is reached (wet mass should be about 1.6 g/L). The cells are centrifuged at 5000 rpm in JLA-8100 rotor (Beckman) for 10 min. The resulting cell pellet is resuspended in 100 mL of grinding buffer (50 mM Tris-HCl pH 7.9, 5% glycerol, 10 mM EDTA, 10 mM DTT, 300 mM NaCl) and homogenized using glass homogenizer. After resuspension, the cells are transferred into two stainless steel tubes (50 mL). Lyophilized lysozyme (sigma) is added (20 μg/mL) and incubated on ice for 30 min. The samples are then sonicated (power output=6, duty cycle 60%) in ice-water bath for 15 minutes each. The cell debris is separated from the crude extract by two centrifugations at 15000 rpm in a JA-25.50 rotor (Beckman) for 30 min each. The supernatant is transferred into a 250 mL Erlenmeyer flask and polyethyleneimine (polymine-P sigma) is added to a final concentration of 0.40% (stock solution is 10% pH 7.9). The sample is stirred in cold room for 10 min followed by centrifugation at 5000 rpm in JA-10 rotor (Beckman) for 5 minutes. The pellet is dissolved in 150 mL of ice-cold TGED (10 mM tris-HCl pH 7.9, 5% glycerol, 0.1 mM EDTA, 0.1 mM DTT) containing 0.3 M NaCl. The solution is homogenized thoroughly using glass homogenizer followed by stirring in cold room for 15 minutes. The washing procedure is repeated using TGED + 0.5 M and RNAP is eluted by dissolving the pellet in TGED + 1.25 M NaCl with stirring in the cold room for 1 hour. The sample is centrifuged as above and the supernatant is retained. Ammonium sulphate is added to a saturation of 65% (60 g per 150 mL of supernatant) slowly (in 15g increments with 5 min in between) and it is left stirring overnight in the cold room. The sample is distributed in 30 mL tubes and is spun down at 15000 rpm in JA-25.5o rotor (Beckman) for 30 min. The resulting pellet is dissolved in 50 mL TGED without salt and a second centrifugation is performed. The crude protein is then applied to a heparin column equilibrated with buffer A (TGED + 0.050 M NaCl) at a flow rate of 0.5 mL/min. The protein is eluted with a step gradient using buffer B (TGED +1 M

NaCl) as follows: 25% B wash, 60% B (RNAP elutes at this step), 100% B wash and 0% B wash. The peaks are analysed through SDS-PAGE stained with coomassie blue (see below). The fractions containing RNAP are pooled and concentrated using Centricon 20 device (100.000 MW cut-off) (Millipore) to 0.5 mL. The fractions are then applied to a superose 6 size exclusion column (GE-healthcare) equilibrated in buffer A (same as for heparin column). The elution of the column is performed in the presence of 20% B, at a flow rate of 0.2 mL/min. The peak containing RNAP is pooled, concentrated and stored in storage buffer (50% glycerol, 20 mM Tris-HCl, 0.3 M NaCl, 0.3 mM EDTA, 2 mM β-mercaptoethanol). SDS-PAGE and chromatograms of purification of RNAP can be seen in Appendix 4, Figure A1.

3.1.2 *Purification of σ^{70}*

For the purification of σ^{70}, the *rpoD* gene was cloned into pET21 vector (Novagen) and transformed into BL21 DE3 cells. The purification was carried out as in (Borukhov and Goldfarb, 1993) with minor modifications. Briefly, cells are grown in LB at 37°C in an orbital shaker to an OD_{600}=0.6 is reached. Overexpression of σ^{70} is the induced by addition of IPTG to a final concentration of 1 mM. After 4h incubation, the cells are collected by centrifugation at 5000 rpm in a JLA-8100 for 10 min. The pellet (9 g wet mass total) is resuspended in 50 mL lysis buffer (40 mM Tris-HCl pH 7.9, 0.3 M KCl, 10 mM EDTA, 0.1 mM DTT) containing 0.2 mg/mL lysozyme and 2 protease inhibitor cocktail tablets (Roche). The suspension is then homogenized using glass homogenizer. After resuspension, the cells are transferred into two stainless steel tubes (50 mL) and disrupted by sonication as described for RNAP. The lysate is then centrifuged in a JA-20.25 at 15000 rpm for 30 min. The supernatant is discarded and the pellet is resuspended in 50 mL of lysis buffer containing 0.5% Triton X100 and 1 mM DTT. Another round of sonication is performed followed by centrifugation. The pellet is dissolved in 10 mL denaturing buffer (50 mM Tris-HCl pH 7.9, 6 M guanidine-HCl, 10 mM DTT, 10% glycerol, 10 mM $MgCl_2$, 1 mM EDTA, 10 μM $ZnCl_2$) and incubated for 30 min at room temperature. The solution is centrifuged at 15000 rpm for 30 min in a JA-25.50 rotor. The supernatant is diluted 2 fold with denaturing buffer and dialyzed overnight against two L of reconstitution buffer (50 mM Tris-HCL pH 7.9, 0.2 KCl, 20% glycerol, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 1 mM EDTA, 1 mM DTT). The dialyzate is cleared by centrifugation (15000 rpm for 30 min in a JA-25.50 rotor) and diluted 4 times with buffer E (40 mM Tris-HCl pH 8.3, 1 mM EDTA, 5% glycerol, 0.1 mM

DTT). The protein is loaded onto a 1.6 x 15 cm DEAE toyo pearl TSK 650 M column (Millipore) pre-equilibrated with buffer E. After injection, the column is washed with 40 mL of buffer E. To elute, a linear gradient of buffer F (40 mM tris-HCl pH 8.3, 1 mM EDTA, 5% glycerol, 0.1 mM DTT, 0.6 M NaCl) (0 - 10% length: 10 mL) – (60 – 100% length: 20 mL) is done. σ^{70} normally elutes between 30 – 50%. The fractions containing the protein are pooled, concentrated and dialysed against storage buffer (40 mM Tris-HCl pH 7.9, 0.2 M KCl, 50% glycerol, 1 mM EDTA, 1 mM DTT).

3.1.3 *Purification of GreB*

Plasmid (based on pCA24N, -*gfp*, *cat*) encoding 6XHis-GreB was obtained from the ASKA clone (-) collection (*E. coli* Strain National BioResource Project, Japan). The plasmid was transformed into T7 express Iq competent *E. coli* cells (High Efficiency, NEB). 4 L of LB media are inoculated with 1/100 volumes of a 100 mL LB overnight culture. Cells are allowed to grow to an OD_{600}=0.4 in orbital shaker at 37°C. *greB* is induced by addition of IPTG (1 mM final concentration) for 4 hours. The cells are centrifuged at 5000 rpm in a JLA-8100 rotor for 15 minutes. The pellet is resuspended in 40 mL of buffer A (7M guanidine-HCl, 40 mM Tris-HCl pH 7.9, 0.8 M NaCl, 0.1 mM DTT) plus two protease inhibitor cocktail tablets EDTA-free (Roche).The cells are homogenized thoroughly on ice, sonicated and the lysate cleared as described for purification of RNAP. The supernatant is then applied to 5 mL of Ni-NTA agarose beads (GE-healthcare) equilibrated in buffer A in a 50 mL falcon tube. Histidine-tagged GreB protein is allowed to bind to the Ni-NTA resin by incubation at room temperature for 30 min in rotary shaker. After binding, the resin is washed 4 times with buffer A, followed by 10 washings with buffer B (40 mM Tris-HCl pH 7.0, 0.8 M NaCl, 0.1 mM DTT). GreB is eluted by addition of 10 mL of buffer C (40 mM Tris-HCl pH 7.9, 0.8M NaCl, 0.1 mM DTT, 0.6 M imidazole) with gentle shaking for 10 min at 4°C. After recovery, the protein is concentrated in Centricon 10 (Millipore) and dialyzed against 2 L of buffer B plus 50% glycerol.

3.1.4 *Purification of EcoRI mutant EcoRQ111*

The plasmid carrying the *EcoRI* variant *EcoRQ111* was obtained as a generous gift from Dr P. Modrich (Duke University). An overnight culture of the strain M5248 (grown at 30˚) is diluted 1/100 in 100 mL LB media. The cells are grown at 30˚C until an A_{600} =0.4 is reached. 1 mL cells culture is pelleted at 5000 rpm using table top centrifuge and the supernatant is discarded. The pellet is resuspended in 200 µL of TSB buffer (10% PEG 3350 (PEG 8000), 5% DMSO 20 mM $MgCl_2$ diluted in LB Filter sterile) with the addition of 3 µL of plasmid pQ111 and incubated on ice for 30 min. The cells are heat shocked at 30˚C for 30 seconds followed by the addition of 800 µL LB media and incubation at 30˚C for 1 hour. The cells are centrifuged at 4000 rpm in table top centrifuge. 900 µL of the supernatant are removed and the remaining 100 µL are used to resuspend the cells. The whole culture is plated in LB+Amp plates and incubated overnight at 30°C. A 100 mL overnight culture is prepared using 1 single colony as starter. From this culture, 2 L of LB+AMP media are inoculated (1/500 dilution). Cells are grown until an OD600=0.8 is reached. The culture is transferred to a shaker set at 42°C and allowed to continue growing for 4 hours. The cell pellet is obtained by centrifugation in JA10 at 5000 rpm for 10 min and resuspended in Qlysis buffer: QLysisBuffer (0.02 M KPO_4 pH 7.4, 1 M EDTA, 15 mM mercaptoethanol, 2 protease inhibitor cocktail tablets (Roche)). The cells are homogenized, disrupted and the crude extract cleared as described for RNAP purification. 20 g of ammonium sulphate are added to the supernatant (50 mL) slowly over a period of 30 min in the cold room with constant stirring. After 60 min, precipitated proteins are recovered by centrifugation at 30000 rpm in JA20 rotor for 30 min. The pellet is resuspended in 10 mL of Resusp buffer (0.02 M KPO_4 pH 7.4, 1 mM EDTA, 5 mM mercaptoethanol) and dialysed against 1.5 L of Qbuffer A (0.02 M KPO_4 pH 7.4, 0.5 mM EDTA, 5 mM mercaptoethanol, 10% glycerol, 0.2 M KCl) overnight. The dialyzate is recovered and cleared by centrifugation 15000 rpm JA 25-50 rotor for 30 min at 4° followed by the addition of DTT to a final concentration of 2 mM. The protein solution is diluted (in small batches to avoid precipitation of EcoRQ111) 7.2 times with QbufferA1 (0.02 M KPO_4 pH 7.4, 0.5 mM EDTA, 5 mM mercaptoethanol, 10% glycerol, 0.1 M KCl) and applied to phosphocellulose column (bio-rad) equilibrated in QbufferA1. The column is washed until absorbance (A_{280}) returns to baseline. The protein is then eluted from the column with a linear gradient of KCL (0.15 to 1M in QbufferA2 (0.02 M KPO_4, 5 mM mercaptoethanol, 0.1 mM EDTA, 10% glycerol). Normally EcoRI methylase elutes at

0.47 and EcoRIQ111 elutes at 0.61 M KCl. The fractions containing EcoRQ111 (test by SDS-PAGE or transcription assay) are pooled, concentrated and dialyzed against storage buffer (20 mM KPO_4 pH 7.4, 0.2 M KCl, 0.2 mM DTT, 1 mM EDTA, 50% glycerol)

3.1.5 *Purification of RelE*

Plasmid encoding 6his-RelE and RelB was a kind gift from Professor Kenn Gerdes (Newcastle University). The plasmid (hisRelE/RelB) is transformed into T7 express cells and an overnight culture is prepared. Two L of LB media are inoculated (1/100 dilution) and the cells are grown at 37°C until an OD_{600}=0.5 is reached. Expression of the proteins is induced for 4 hours by addition of IPTG to a final concentration of 1 mM. C centrifugation at 5000 rpm in a JLA-8100 for 10 min is done and the pellet is resuspended in ice-cold buffer A (50 mM NaH_2PO_4, 0.3 M NaCl, 10 mM imidazole, 5 mM mercaptoethanol, 2 tablets of protease inhibitor cocktail (Roche)). The cells are disrupted by 3 consecutive passes through French press at 18000 psi. The resulting crude extract is centrifuged at 15000 rpm at 4°C in a JA-25.50 rotor for 30 min. The supernatant is applied onto 10 mL Ni-NTA beads and incubated with rotation for 1 hour at 4°C. The beads are washed 5 times with buffer B (50 mM NaH_2PO_4, 0.3 M NaCl, 35 mM imidazole, 1 mM mercaptoethanol). RelB is separated from the Ni-His-RelE:RelB complex by incubation at room temperature with 2 volumes of buffer b (100 mM NaH_2PO_4, 10 mM Tris-HCL pH 7.9, 9.8 M urea, 1 mM mercaptoethanol) for 2 hours. RelE is eluted from the beads by washing with 50 mL of buffer D (100 mM NaH_2PO_4, 10 mM Tris-HCL pH 7.9, 9.8 M urea, 0.5 M imidazole, 1 mM mercaptoethanol). The supernatant is diluted to a protein concentration <0.2 mg/mL. The proteins are refolded by a series of dialysis. Dialyze the supernatant against 2 L of buffer d1 (PBS 1X, 1 mM DTT, 0.1% Triton X100) for 8 hours followed by an overnight dialysis against 2 L of bufferd2 (PBS 1X, 1 mM DTT). Perform one last dialysis against 2 L of buffer d (PBS 1X, 1 mM DTT, 20% glycerol). Take one aliquot for analysis, snap-freeze and store at -80. The PBS 10X recipe is: dissolve in 800 mL H_2O: 80g NaCl, 2g KCl, 14.4 g Na_2HPO_4 and 2.4 g KH_2PO_4. Adjust pH to 7.4

3.1.6 *Preparation of DNA and RNA templates*

DNA templates for promoter borne transcription and for *in vitro* mRNA synthesis were obtained by PCR using Phusion DNA polymerase (Finnzymes) and were subsequently purified from 2% agarose gels using Gel Extraction Kit (Qiagen). Oligonucleotides for PCR and for formation of assembled elongation complexes were purchased from IDT. RNA templates were prepared using T7 RNA polymerase (Purified by N. Zenkin). For the *in vitro* synthesis of RNA, 30 pmol of template DNA are mixed with 12 pmol of T7 RNAP in T7 transcription buffer (200 mM Tris-HCl (pH 7.9 at 25°C), 30 mM MgCl2, 50 mM DTT, 50 mM NaCl, 10 mM spermidine, 2 mM all four NTP's) at 37°C for 6 hours. 10 µl of RQI DNase I (Roche) are added and the reaction is incubated for 30 minutes at 37 °C. To stop the reaction, 50 µl of 0.5 M EDTA pH 8.0 and 50 µl of 3 M sodium acetate pH 5.2 are added. The RNA is partially purified by the addition of an equal volume of chloroform followed by vortex for 30 s and subsequent centrifugation in table top centrifuge (three chloroform extractions in total). The RNA is precipitated by the addition of 100% ethanol and incubated overnight at -80°C. After centrifugation at 4 °C for 30 minutes, the pellet is desalted by one washing with 70% ethanol. The RNA is resuspended in 60 µL of 2XRNAbuffer loading dye (90% formamide + 0.02% bromophenol blue) and loaded on a 6% polyacrylamide denaturing gel. The electrophoresis is run at 50W constant. The gel is placed between two sheets of Saran wrap and placed on top of a fluorescent TLE plate (Merck). UV light is used to detect the RNA which is excised from the gel using a clean razor blade. The gel strip containing the RNA is put in a 1 mL tube and 600 µL of 0.3 M sodium acetate pH 5.2 are added. The mixture is vortexed for 12-14 hours at 4 °C using eppendorf platform shaker. The eluted RNA is extracted by the addition of 300 µl of chloroform and vortex for 10 minutes. The sample is spun down at max speed for 2 minutes. The liquid is transferred to a new tube being careful of not transferring any gel pieces. 600 µl of chloroform are added and a total of 3 chloroform extractions are performed. The RNA is then ethanol precipitated, desalted and its concentration is measured using a nanodrop spectrophotometer.

3.1.7 *Transcription from templates that contain promoters*

Although the protocol for the *in vitro* transcription of templates containing a promoter may vary from one template to another (due to the utilization of different sequences) the general protocol can be outlined here. In this study, all DNA templates carried the *E. coli* T7A1 promoter until the +11 position. From this point on, the translation initiation region of the gene 32 of the T4 phage was used until the second codon (F).

For experiments carried out on templates containing promoters the transcription buffer (10 mM Tris HCl pH 7.4, 10 mM $MgCl_2$) differed only in the salt concentration: 50 mM KCl (TB50) for transcription initiation and elongation; 1 M NaCl (TB1000) for high salt washings. Transcription is carried out on solid phase by immobilising the DNA which contains a biotin tag located at the 5' end on streptavidin beads (Sigma). RNA polymerase (1 pmol) is incubated in 20 µL TB50 with σ^{70} (4 pmol) in the presence of biotinolated DNA template (2 pmol) for 5 minutes at 37°C. Then, the initiation mixture 15 pmol CpApUpC, 25 µM of GTP, ATP and CTP is added to a total volume of 40 µL and incubated for 7 min at 37°C (resulting in formation of elongation complex containing 11 nucleotides long RNA; EC11). Streptavidin beads (10 µL, equilibrated in TB50) are added and the tube is shaken gently for 5 min. Two 1 mL washings with TB1000 containing 100 µg/mL heparin (Sigma) are performed, followed by five 1 mL washes with TB50. In order to translocate RNAP to the next positions incomplete sets of NTPs are added to the complexes to a final concentration of 25 µM, incubated for 5 minutes at room temperature (RT) and washed five times with TB50. For instance, to walk EC11 to position +56 mer (labelling step), UTP, GTP and α-[^{32}P]-ATP (4 pmol) are added for 5 minutes at RT. After that, ATP is added, and incubated for another 5 minutes and then washed. To continue "walking" RNAP rounds of transcription with incomplete sets of NTPs cognate to the template DNA are added until RNAP is positioned in the desired elongation complex. After washing of last elongation complex, the volume of the reaction is reduced to 18 µL and transferred to the binding reaction (below). Whenever RNAP is walked, the walking steps will be indicated in the text according to the used DNA template.

3.1.8 *Transcription from artificially assembled elongation complexes*

Although the protocol for the assembly of artificial elongation complex is the same regardless of the sequence, the RNAP walking reaction varies according to the sequence

in the template DNA which is being used. The general procedure of assembly is outlined here and the walking steps will be explained in detail in the text.

Transcription in artificial elongation complexes (AAEC) was performed as described in (21-24) with some modifications. For the "transcription first" system, RNA (5 pmol), RNAP (7.5 pmol), and 5'-biotinolated template DNA (20 pmol) is incubated at 37°C in Low magnesium Low pH buffer (LmLpH) (10 mM Tris pH 7.4, 40 mM KCl and 5 mM $MgCl_2$) for 15 min, followed by the addition of 100 pmol of non-template DNA to a total volume of 40 µL. After 15 min incubation at 37°C, streptavidin beads (10 µL equilibrated in LmLpH) are added and gently shaken for 5 min at RT followed by 2 washings with TB1000 and 5 washings with LmLpH. 4 pmol α-[32^P]-GTP are added and the mixture is incubated for 5 min at RT. The reaction is then washed 5 times with LmLpH buffer and RNAP is translocated by addition of the next cognate NTP's at a final concentration of 25 µM. The reaction is incubated at RT for 7 minutes and washed 5 times with 1 mL of LmLpH buffer. As above, the reaction volume is reduced to 18 µL and used for the ribosome binding reaction.

For analysis of peptidyl transferase activity in the coupled system, 75 pmol of TAP-tagged RNAP is first immobilised on 10 µL of Ig-G beads (GE healthcare) equilibrated in TB100, by gentle shaking at 30°C for 30 min followed by 5 washings with LmLpH buffer. Transcription elongation complexes are assembled as above, except for 50 pmol of RNA, 200 pmol of template DNA and 1000 pmol of non-template DNA were used, followed by addition of translation machinery components (see below).

For the formation of "translation first" artificial elongation complexes, translocated ribosomes (see below) were applied onto a sucrose cushion and subjected to ultracentrifugation as described in (25). The resulting pellet is washed 3 times with TrLB (see components below) and resuspended in 10 µL of the same buffer and then mixed with 60 pmol of template DNA and 30 pmol of RNAP. The reaction is incubated for 15 min at 37°C followed by the addition of 300 pmol of non-template DNA and subsequent incubation for 15 min at 37°C. α-[32P]-GTP is used to label the 3' end of the mRNA and RNAP is walked by addition of the cognate NTPs at a final concentration of 25 µM.

All transcription reactions are terminated by addition of an equal volume of stop buffer (EDTA 20 mM, 7 M urea, 100 µg/mL heparin, 0.02% bromophenol blue, 0.03% xylencianol saturated in formamide) and the products are resolved by sequencing 6% PAGE and analysed using ImageQuant software (GE-Healthcare).

Note that for transcriptional complexes immobilised through biotinylated DNA onto streptavidin beads 44% of such complexes were lost during washings. For transcriptional complexes immobilised through a histidine tag or TAP-tag on RNAP, loses of 13% were observed. The list of sequences used for the experiments conducted in this thesis which are not shown within the text and figures can be find in Appendix 3 table A3

3.2 Translation machinery

3.2.1 *Purification of 70S ribosome*

Ribosomes were purified as described in (Moazed D, 1989) except that two high salt washings were performed instead of one. 15 g of *E. coli MRE* 600 are resuspended in buffer A (20 mM Tris 7.6, 10 mM $MgCl_2$, 100 mM NH_4Cl, 6 mM 2 mercaptoethanol) and then lysed by two passes through a French press (C-019 constant systems UK) at 30000 psi. DNase I is added to a final concentration of 20 µg/mL and the lysate incubated on ice for 30 min, before the volume is adjusted to 45 mL with buffer A. After 2 clearing spins in JA-25.50 rotor at 15000 rpm, the supernatant is loaded onto two 35 mL sucrose cushions in 75 mL polycarbonate tubes. Ultracentrifugation is carried out for 22 hours at 35000 rpm at 4°C in a Ti-45 Beckman rotor. The pellet is washed with buffer A and resuspended in 5 mL of the same buffer. After a clearing spin at 15000 rpm in a JA-25.50 Beckman rotor, the volume is adjusted to 100 mL with buffer A containing 0.5 M NH_4Cl. After ultracentrifugation for 7 hours at 22000 rpm the pellet is washed again with buffer A and then resuspended in 100 mL of buffer A containing 0.5 M NH_4Cl, before a final ultracentrifugation step (7 hours, 22000 rpm). Purified ribosomes are resuspended in 1.6 mL of buffer A containing 50 mM Tris pH 7.6, frozen in liquid nitrogen and stored at -80°C. Ribosome concentration is calculated according to: A_{260} = 1 equals to 23 pmol of 70s ribosomes/mL.

3.2.2 *Purification of EF-G, EF-Tu, IF-1, IF-2, IF-3, F-Met-tRNA^{fmet} aminoacyl tRNA synthetase (MetRS) and methionyl-tRNA^{fMet} formyltransferase (FTM)*

Plasmids (based on pCA24N, -*gfp*, *cat*) encoding 6XHis tagged IF-1, IF-2, IF-3, EF-Tu, EF-G MetRS and FTM, were obtained from the ASKA clone (-) collection (*E. coli* Strain National BioResource Project, Japan). Plasmids were transformed into T7 express Iq competent *E. coli* cells (High Efficiency, NEB). A 100 mL overnight culture is used to inoculate 4 L of LB media supplemented with 25 µg/mL chloramphenicol. Cells are grown in an orbital shaker at 37°C until an OD_{600}=0.4 is reached. IPTG (0.250 mM final) is added and induction is carried out at 30°C for *IF-3, EF-G* and *EF-Tu* and at 37°C for other proteins for a total of 4 hours. After induction, cells are pelleted and washed twice with translation buffer (TrLB) (10 mM Tris-HCl pH 7,4, 60 mM NH_4Cl,

10 mM $Mg(OAc)_2$ and 6 mM β-mercaptoethanol). Cells containing over-expressed EF-Tu are washed with TrLB buffer containing 1 mM GTP to avoid precipitation of the enzyme. Pellets were resuspended in TrLB buffer + 10% glycerol and EDTA-free protease inhibitor cocktail (Roche), and incubated on ice with lysozyme (0.1 mg/mL) for 30 minutes. Cells are disrupted by sonication (Power output = 6 Duty cycle 60%) in stainless steel tubes in an ice-water bath for 15 min, followed by two clearing centrifugation steps at 15000 rpm in a JA-25.50 Beckman rotor. An ultracentrifugation step is done in polycarbonate tubes for 22 hours in a Ti-45 Beckman rotor at 33000 rpm. The supernatants are then applied onto a His-Trap column (GE healthcare) connected to an AKTA Explorer FPLC (GE healthcare). Bound proteins are eluted with a linear gradient of imidazole (from 10 mM to 200mM) in elution buffer (20 mM Tris pH 7.4 600 mM NaCl). Peak fractions were pooled and analysed by SDS-PAGE (10%). Fractions containing the proteins of interest were dialyzed overnight against 2 L of TrLB buffer supplemented with 50% glycerol. All proteins are purified to homogeneity of at least 90% and proved to be RNase and DNase free. Activity of the purified proteins is tested in the ribosome walking experiment which will be described later. SDS-PAGE showing purified components can be seen in Appendix 4, Figure A2.

3.2.3 *Purification of S100 extracts as a source of aminoacyl synthetases*

S100 extracts are obtained as described in (Bourdeau et al., 2001). *E. coli MRE 600* cells are grown in LB media until an OD_{260}=0.6 was reached. The cells are disrupted in grinding buffer (20 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 5% glycerol, 50 mM NaCl) by two 5 minute rounds of sonication (Power output = 6 Duty cycle 60%) on ice. The crude extract is cleared by centrifugation at 15000 rpm in a JA-25.50 Beckman rotor for 30 minutes. Then, ultracentrifugation of the supernatant is done at 30000 rpm for 22h in a Ti-45 rotor. The resulting S100 crude extract is purified on a 16 mL DEAE-cellulose column (Whatman) equilibrated with buffer S100 (50 mM Tris-HCl pH 7.4, 70 mM NH_4Cl, 30 mM KCl and 7 mM $MgCl_2$). A linear gradient from 0 to 300 mM NaCl (in the same buffer) is applied and the eluted peak fraction is dialyzed against storage buffer (50 mM Tris-HCl pH 7.4,10 mM $MgCl_2$, 50% glycerol, 50 mM KCl).

The list of strains and plasmids used for protein purification can be found in Appendix 1, table A1 and Appendix 2 table A2 respectively.

3.2.4 *Preparation of the N^5-N^{10}-methenyltetrahydrofolic acid as precursor of N^{10}-tetrahydrofolate*

In order to obtain the peptidyl-analogue F- Met- $tRNA^{fmet}$ (initiator aminoacyl-tRNA in bacteria), the compound N10-formyl-tetrahydrofolate donates the formyl group to the Met- $tRNA^{fmet}$ in a FTM dependent manner. Due to the chemical instability of the donor compound, an stable precursor is prepared (N_5,N_{10}-methenyltetrahydrofolic acid). This precursor is then transformed into N10-formyl-tetrahydrofolate just before the N-formyltransferase reaction takes place. To prepare the stable form, 25 mg of folinic acid (Ca^+ Salt sigma) are dissolved in 2 mL 50 mM β-mercaptoethanol. Then, 220 μL of 1 M HCl are added and incubated for 3 hours at room temperature. When the incubation is finished, the reaction is diluted with 1 mL of 100 mM HCl and stored in 200 μL aliquots. Right before the preparation of F-Met-$tRNA^{fmet}$ by formylation and aminoacylation reactions (which are conducted simultaneously), N_5,N_{10}-methenyltetrahydrofolic is neutralized by the addition of 10 μL 1 M Tris-HCl pH 7.9 and 20 μL of 1 M KOH. After 15 minutes of incubation at room temperature, the neutralized form (N10-formyl-tetrahydrofolate) is added to the charging/formylation reaction.

3.2.5 *Formylation, aminoacylation and N-acetylation of $tRNA^{fMet}$*

The reaction for the formylation and aminoacylation of $tRNA^{fmet}$ can be performed simultaneously. 1 unit of $tRNA^{fmet}$ (sigma) is mixed with 2 mM L-methionine, 10 mM ATP, 50 mM HEPES, 10 mM KCl, 1 mM DTT, 20 mM $MgCl_2$, 50 pmol MetRS and 50 pmol FTM and incubated at 37°C for 25 minutes. The reaction is quenched by addition of 50 μL of 3 M sodium acetate pH 5.3 and 28 μl of 10% SDS. The reaction is extracted by vortexing in presence of phenol for 10 min. After a 5 min centrifugation at full speed in table top centrifuge, the supernatant is separated from the phenol phase. 500 μl of 0.3 M sodium acetate pH 5.3 are added to the phenol, vortexed and centrifuged. The two supernatants are treated twice with chlorophorm, ethanol precipitated, desalted by washing with 70% ethanol and dissolved in 60 μL of 2 mM sodium acetate pH 5.3. The

resulting F-Met-tRNA^{fmet} is gel filtrated 4 times using Bio-Rad Bio-Spin 6 columns equilibrated in 2 mM sodium acetate to remove traces of ATP. All tRNAs were purchased from Sigma and amino acids from USB. The preparation of the peptidyl analogue N-acetyl-met-tRNA^{fmet} is carried out as described above, but the formylation step is omitted. The acetylation is done right after the ethanol precipitation step. The pellet is dissolved in 200 mM sodium acetate followed by the addition of 2.5 μL of acetic anhydride (sigma). After incubation for 1 hour, another 2.5 μL of acetic anhydride are added and incubated for 1 hour. The volume of the solution is increased to 500 μL of 200 mM sodium acetate, precipitated, desalted, resuspended in 2 mM sodium acetate pH 5.3 and gel filtrated.

3.2.5 *General aminoacylation procedure (for tRNAs other than tRNA^{fMet})*

The aminoacylation of tRNA is carried out the same way as for tRNA^{fmet}. The difference is that the procedure does not require FMT or formyl group donor and also, SDS is not added at the end of the reaction. The reaction is catalyzed by aminoacyl synthetases present in the purified S100 extract. Briefly, 1 unit of tRNA (sigma) is mixed with 2 mM (amino acid), 10 mM ATP, 50 mM HEPES, 10 mM KCl, 1 mM DTT, 20 mM $MgCl_2$ and 50 pmol S100 extract. The reaction is incubated at 37°C for 25 minutes and quenched by addition of 50 μL 3 M sodium acetate pH 5.3. Purification of the aminoacyl-tRNA is carried out exactly as described for tRNA^{fmet}.

3.2.6 *Acid electrophoresis and thin layer chromatography for the evaluation of the extent of formylation, N-acetylation and aminoacylation*

The extent of aminoacylation can be determined by acid gel electrophoresis due to the migratory differences in the aminoacylated and deaminoacylated states of the tRNA. The samples are diluted in loading dye (100 mM sodium acetate pH 5.3, 7 M urea, 0.05% bromophenol blue) and loaded onto a large, 6.5% acrylamide/bisacrylamide (19:1) sequencing gel containing 100 mM sodium acetate (pH 5.3) and 8 M urea. The electrophoresis is run for 22 hours at 200 Volts in the cold room. The gel is stained with methylene blue 0.025% with gentle shaking and distained with water. The aminoacylation state (and extent) can also be assessed by thin layer chromatography. In order to deacylate the aminoacylated tRNA, ammonium hydroxide is added to a final

concentration of 3% and incubated at 55°C for 20 min. The samples (aminoacylated and deaminoacylated) are spotted in a silica gel thin layer chromatography plate (Merck) at 1.5 cm from the bottom. The mobile phase of the chromatography is composed by a mixture of butanol:water:acetic acid at a ratio of 4:1:1. The chromatography is allowed to run for 2.5 hours and then dried under the fume hood. If the samples are radioactive, the chromatography plate is wrapped up in Saran wrap, placed in a phosphorscreen cassette (GE-healthcare) and scanned with a Typhoon scanner. If the samples are not radioactive, after the silica gel plate has dried out, ninhydrin (sigma) is sprayed to the plate using chromatography atomizer (Merck). The amino acids which have reacted with ninhydrin are revealed by heating up the plate to a 100°C (using heating plate) for 7 min.

3.2.7 *EF-Tu-GTP-Aminoacyl-tRNA ternary complex formation*

In order to exchange the GDP bound to purified EF-Tu with GTP, EF-Tu·GDP (400 pmol) is incubated with GTP (600 pmol) in the presence of phosphoenol pyruvate (800 pmol) and phosphoenol pyruvate kinase (200 µg/mL) in 30 µL ternary complex buffer (50 mM Tris-HCl pH 7.4, 40 mM NH_4Cl, 10 mM $MgCl_2$ and 1 mM DTT). The reaction is incubated for 10 min at 37°C. Equimolar amounts of aminoacyl-tRNA are then added and further incubated at 37°C for 5 minutes. 6 µL of the resulting ternary complexes are added in both, peptidyl-transferase and translocation assays (below).

3.2.8 *In vitro translation initiation and elongation*

For translocation of the ribosome in both coupled and uncoupled systems a method similar to (Takyar et al., 2005b) is used. For ribosome binding and initiation: 2 µM 70S ribosomes are mixed with 2 pmol of either mRNA or transcription elongation complexes in a final volume of 49 µL of TrLB. The mixture is incubated for 10 minutes at 37°C followed by the addition of either N-acetyl-met-$tRNA^{fmet}$ (non enzymatic initiation) or F-met-$tRNA^{fmet}$, IF-1, IF-2, IF-3 to final concentration of 5 µM and 200 µM GTP (enzymatic initiation) to final volume of 60 µL and incubation at 37°C for 10 min.

For elongation, 10 µL aliquots of the initiated translation initiation complexes, are withdrawn and transferred to tubes containing 1.2X TrLB buffer, 5 µM EF-G, 200 µM

GTP and 1 µM ternary complexes (translocation reaction). This mixture is incubated for 7 minutes at 37°C. For RelE printing of the translocated complexes, 12 pmol of RelE are added and the reactions incubated for 10 min at the same temperature.

3.2.9 *Toeprinting assay*

To detect translocation by the inhibition of reverse transcription (toeprinting), a radiolabeled primer is annealed to the 3' end of the mRNA prior to translation initiation by heating at 65°C for 10 min in TrLB without magnesium followed by quick cooling in the presence of Mg^{2+}. After translation initiation, 10 µL aliquots are added to the translocation mix which had the same components described above plus four dNTPs (to final concentration of 300 µM). After translocation, 1 unit of SuperScript III reverse transcriptase (Invitrogen) (diluted in the buffer provided by supplier) is added and the reaction is incubated for another 10 min at 37°C. The products from both RelE cleavage and toeprint reactions are terminated by addition of an equal volume of stop buffer (EDTA 20 mM, 7 M urea, 100 µg/mL heparin, 0.02% bromophenol blue, 0.03% xylencianol saturated in formamide) and the products are resolved by sequencing 6% (19:1) PAGE and analysed using ImageQuant software (GE-Healthcare)

3.2.10 *Peptidyl transferase assay*

The peptidyl transferase assay is conducted as described above (translation initiation and translocation) with the following differences. [^{35}S]-F-met-tRNA^{fmet} is used to visualise the peptides. The concentration of ribosomes and mRNA (for the analysis of peptide synthesis in the absence of transcription) or immobilised transcriptional complexes (to analyse peptidyl transferase activity in the presence of the transcriptional machinery) is ~1 µM and all components are scaled up accordingly. After the translocation reaction, KOH is added to a final concentration of 100 mM to deacylate the tRNA, and incubated at 37°C for 30 min. The products are resolved by thin layer electrophoresis (Brunelle et al., 2006).

3.2.11 *Thin layer electrophoresis for the resolution of mono, di, tri, tetra and penta peptides.*

Cellulose chromatography plates (plastic backed) (Merck) are cut in rectangles measuring 17 cm L x 7 cm W. A line is drawn right in the middle of the plate. The samples (from peptidyl transferase assay) are spotted right on this line (referred to as the origin). The plate is then placed in an electrophoresis chamber (with a footprint 15 x 9.5 x 22 cm) with separated buffer chambers. In each chamber Pyrac buffer (200 mL Acetic acid (sigma) 5 mL pyridine (sigma), in one L of water) is added (30 mL each). The buffer is allowed to run over the plate for 25 min. The Pyrac buffer is added to the plate slowly without touching the sample. Once the buffer has concentrated the samples, the plate is covered with Stoddard solvent (Sigma) and the electrophoresis is run at 1200 V, 13 mA, 13 W for 30 min. The buffers are discarded and the plate is dried for 3 hours. Then it is carefully wrapped in Saran wrap, exposed to a phosphorscreen and scanned using Typhoon scanner.

Results Chapter 4. Setting up a controlled *In vitro* translation system

4.1 Introduction

The first step towards the set up of a transcription coupled to translation system consisted in the purification of all minimal components needed by the translational machinery to function and the preparation of substrates for it. The second step involved the establishment of an active *in vitro* translation system in which translation initiation and stepwise translocation of the ribosome was achieved. This chapter will focus on the different approaches we took in order to build an *in vitro* translation system suitable for the purpose of the analysis of the interaction between transcription and translation.

4.2 Attempting to generate a transcription-translation system using purified mRNA from an *E. coli* RNAP transcription reaction.

4.2.1 *Protein purification and aminoacylation of tRNA(s)*

In order to recreate the translational system *in vitro*, 70S ribosomes (from *E. coli* MRE600 cells), initiation factors 1, 2, 3, elongation factors EF-Tu and EF-G were purified. In order to achieve a higher level of purity, some methods were optimized. For example, the standard ribosome purification protocol involves one high salt wash to remove loosely associated non ribosomal proteins; we used two washings instead giving as result pure, vacant, 70S ribosomes. Since we were attempting to develop a system where the ribosome could be translocated in a step-wise manner, the preparation of individual, highly pure substrates for translation (aminoacyl-tRNA(s)) was a requirement. Therefore, the aminoacyl tRNA synthetase MetRS (for preparation of Met-$tRNA^{fmet}$) and mixed ones (present in DEAE S100 extracts) were purified along with other factors needed for the modification of certain aminoacyl-tRNAs (such as methionyl-$tRNA^{fmet}$ formyltransferase (FTM)). Deaminoacylated tRNA's purified from *E. coli*: $tRNA^{fmet}$, $tRNA^{phe}$, $tRNA^{val}$, $tRNA^{tyr}$ and $tRNA^{lys}$ were obtained from Sigma. To reduce the amount of components present in the coupled system the aminoacylation procedure was carried out as a separate experiment so the aminoacyl-tRNA could be further purified. The resulting aminoacyl-tRNA was phenol-chlorophorm extracted, ethanol precipitated, desalted and finally was gel filtrated four times to completely remove traces of ATP to circumvent the possibility that RNAP would incorporate those

NTPs into the nascent RNA affecting the interpretation of the assessment of the effects on transcription-translation coupling itself.

As mention in the introductory chapter, there are two ways in which translation can be initiated *in vitro*. The **enzymatic pathway** in which initiation factors 1, 2 and 3 and GTP are responsible for 70S ribosome assembly (on the translation initiation region of the mRNA) and for positioning of the initiator tRNA f-met-tRNA^{fmet} in the ribosome's P site. The alternative way does not require initiation factors and f-met-tRNA^{fmet}. Instead, the peptidyl analogue N-acetyl-met- tRNA^{fmet} which is capable to bind to the ribosome's P site unaided by enzymatic factors (henceforth, this type of initiation will be referred to as **non-enzymatic initiation**) is used to initiate translation.

The aminoacylation, formylation (for enzymatic initiation) or N-acetylation of tRNA^{fmet} (for non-enzymatic initiation) and the aminoacylation of all other tRNA's was evaluated by acid gel electrophoresis (Walker and Fredrick, 2008). As can be seen in (Figure 19), the migration of deaminoacylated tRNA differs from its aminoacylated counterpart due to the presence of the amino acid which retards the migration of the tRNA. Acid gel electrophoresis was also used to assess the formylation and N-acetylation of Met-tRNA^{fmet}. In our preparations, the extent of aminoacylation, formylation and N-acetylation was always greater than 90%.

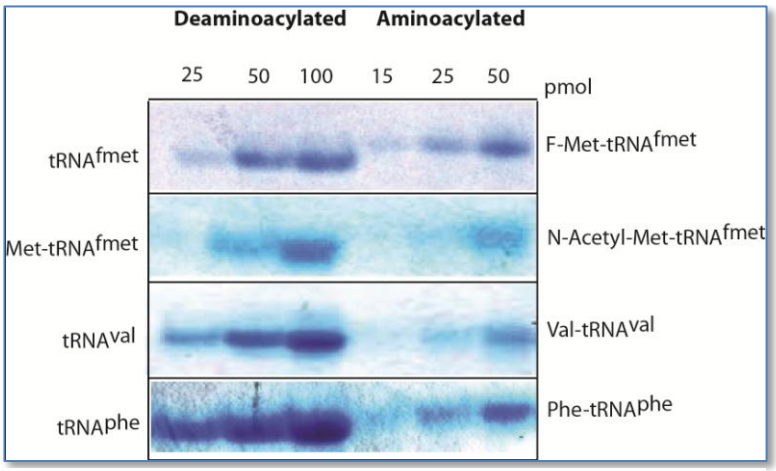

Figure 19. Acid gel electrophoresis. The extent of aminoacylation was tested by the difference in migration of the uncharged versus the charged state of the tRNA. Different amounts of tRNA's were loaded onto the gel to estimate the concentration and the extent of the aminoacylation. For all preparations the aminoacylation was >90% as tested by TLC (not shown). The tRNA's shown in these gels were stained with 0.025% methylene blue and distained with water. This gel was used as a qualitative observation of the aminoacylation reaction.

The elongation factor EF-Tu purifies in a GDP bound form (EF-Tu-GDP) which is the inactive form of the factor. In order to switch it to its active form, the GDP is exchanged for GTP using pyruvate kinase and phosphoenol pyruvate in the presence of GTP; then

aminoacyl-tRNA is added to the reaction to form a ternary complex (TC) composed of EF-Tu-GTP-aminoacyl-tRNA leaving the enzyme ready to participate in the elongation reaction (see below). For a complete list of factors and substrates for the translational machinery see table 2.

Translation		**Transcription**
70S ribosomes	Val-RS	RNAP
IF-1	DEAE-S100	σ70
IF-2	EF-G	GreB
IF-3	EF-Tu	EcoRQ111
MetRS	F-met-tRNA^{fmet}	NusA
FMT	N-Acetyl-met-tRNA^{fmet}	NusG
Phe-RS	Phe, Val, Tyr, Lys-tRNA	
Pyruvate Kinase	Phosphoenol pyruvate	

Table 2 List of factors and substrates needed for the assembly and functional characterisation of a transcription-translation system.

4.2.2 *Programming the ribosome with mRNA purified straight from transcription reaction.*

In order to assemble the translational system, we decided to program the ribosome with mRNA purified from an *in vitro E. coli* RNAP transcription reaction. In this way, once the translational system was operational, to assemble transcription to translation, the mRNA purification step would be omitted. In order to be able to follow both translation initiation and elongation a technique known as toeprinting (Moazed and Noller, 1989) was chosen.

Toeprinting is based on the inhibition of the progression of the reverse transcriptase (rt) by an obstacle in the RNA (in this case the ribosome). This method allows the detection of ribosomal complexes in initiation and elongation states with nucleotide resolution. As can be seen in Figure 19, a radiolabeled oligonucleotide is annealed to the 3' end of the mRNA and rt is allowed to synthesise cDNA. If the mRNA is naked, the rt will generate full length cDNA. If translation is initiated, the rt will collide with the ribosome generating a shorter cDNA compared to the full length control. If the

ribosome is translocated by one or more codons (forward ribosome translocation occurs towards the 3' end of the mRNA), the cDNA synthesised by the rt will be even shorter

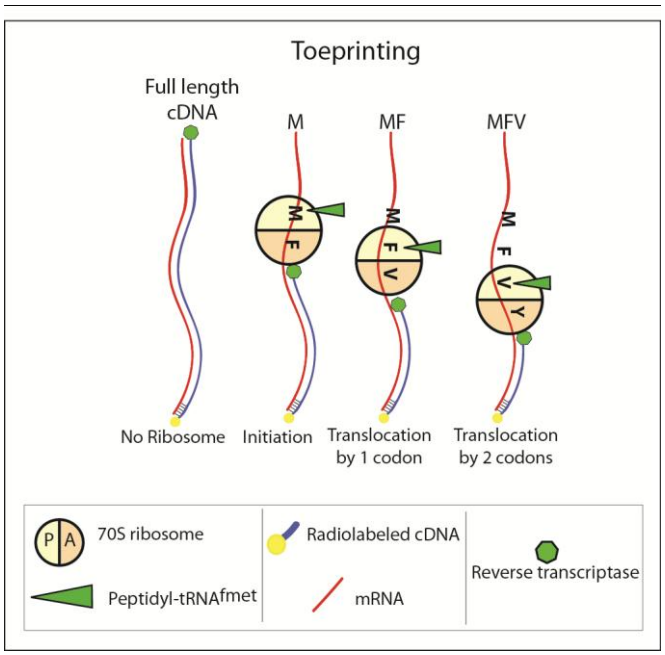

Figure 20. Toeprinting technique. Schematic representation of toeprints. The reverse transcriptase (rt) (green hexagon) synthesises full length cDNA (blue line) of RNA (red line) in the absence of ribosome. If the ribosome (black circle) is initiated the rt progression is inhibited, synthesising a shorter cDNA. As ribosome is translocated to different positions, the distance with the 3' end of the RNA is shortened. Therefore the resulting cDNA species are shorter.

compared to the cDNA generated in initiated translation complexes only (Figure 20).

For the transcription reaction, a PCR generated, 5' end biotinylated DNA template (referred to as: T7A1-g32) was used. This template contained the T7A1 promoter until the +11 position (relative to the transcriptional +1 site) fused to the bacteriophage T4 gene 32 sequence (from -42 to +86 being Met, Phe, Val, Tyr and Lys the first codons respectively). This particular translation initiation region (TIR) has been used for the analysis of different aspects of translation *in vitro* due to the efficient translation initiation achieved on it (Takyar et al., 2005a, Fredrick and Noller, 2002) .

The transcription reaction was initiated by addition of RNAP and σ^{70}. In order to be able to position RNAP in a desired elongation complex (EC) a method known as RNAP "walking" was used (Nudler et al., 2003). The method requires the transcriptional system to be immobilised through RNAP (6xHis, biotin or TAP tagged) or as mentioned above, through the DNA template through a biotin tag on beads carrying the specific ligand for a given tag (therefore the name "solid phase" for immobilised systems). Working on solid phase makes washings possible, giving control over the

presence of substrates (NTPs), the salt concentration or the composition of the buffer which is in the transcription reaction. After immobilisation, in order to "walk" RNAP an incomplete set of NTPs cognate to the sequence in the template DNA is added. This allows RNAP to transcribe a stretch of DNA and then stall in front of the nucleotide which was omitted. The reaction is then washed to remove the unincorporated NTPs leaving the system ready to enter into a new round of elongation upon the addition of another set of NTPs. After transcription initiation, RNAP was walked to EC11 followed by washing with high salt (1 M NaCl) and low salt buffer (40 mM KCl). These washings ensured that all abortive products, promoter complexes, non specific complexes with DNA and unused RNAP were removed assuring that the synthesised full length mRNA from washed EC11 is of higher purity compared to mRNA obtained from an unwashed reaction. After EC11 formation, transcription was allowed to elongate by addition of all four NTPs synthesising full length mRNA (Figure 21).

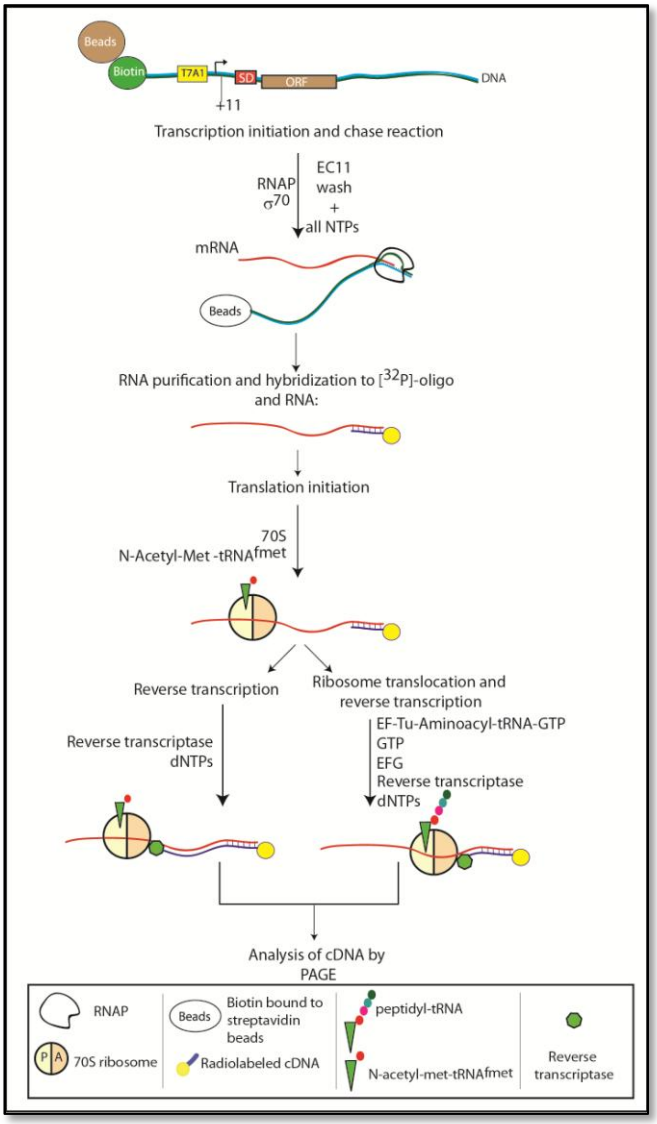

Figure 21. Workflow for the translation system programmed with purified mRNA from *E. coli* RNAP transcription reaction approach (see text)

The mRNA synthesised was purified from RNAP (see below) and other components present in the reaction. The purification of mRNA, hybrid formation and reverse

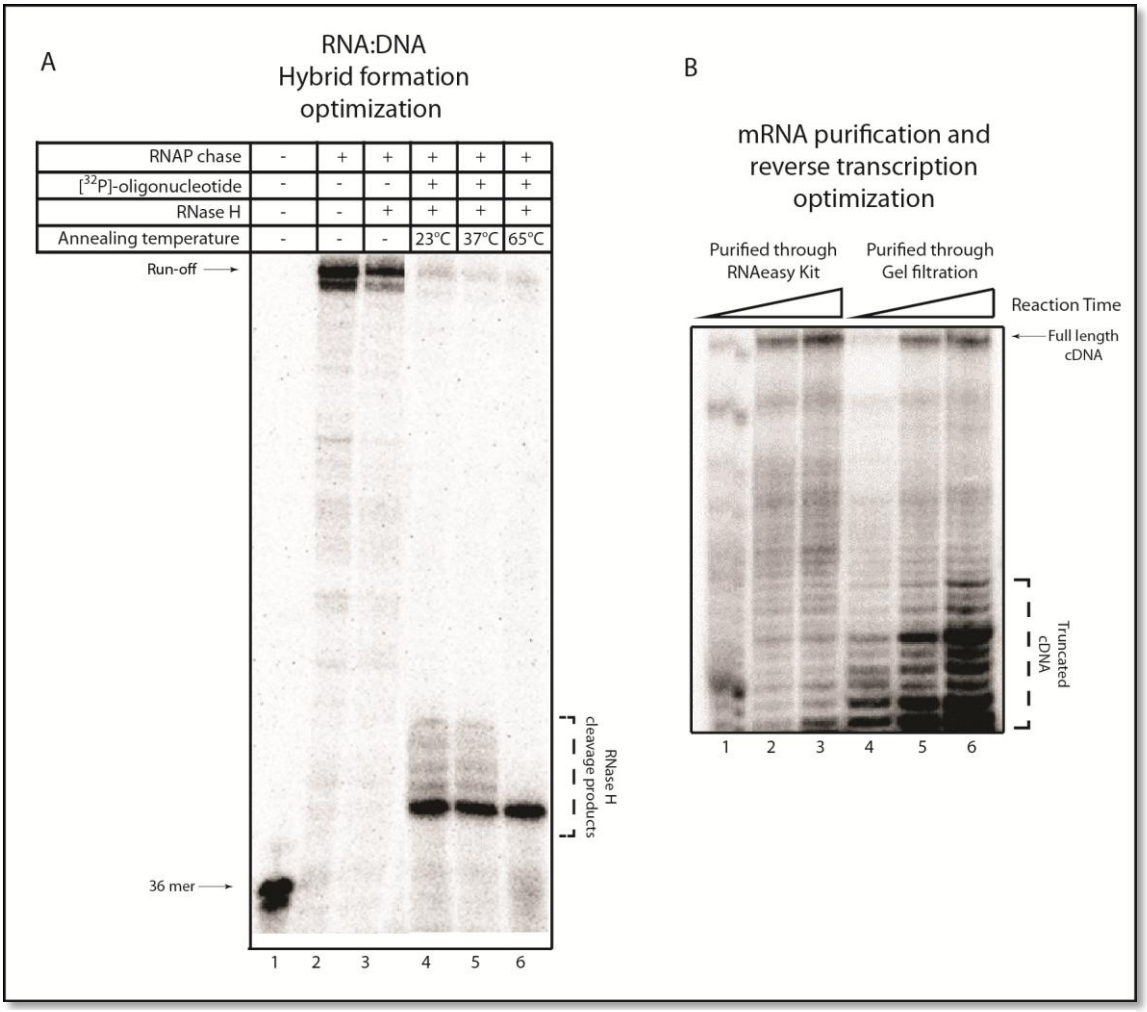

Figure 22. mRNA Purification, RNA:DNA hybrid formation and reverse transcription optimization. A) Denaturing PAGE of mRNA product of *in vitro* transcription reaction. RNAP was walked until elongation complex (EC) 36 with radio labelling of the mRNA in the same complex (lane 1). RNAP was allowed to synthesise full length mRNA (run-off reaction) by addition of all 4 NTPs (lane 2). In the absence of oligonucleotide, no cleavage products were obtained in the presence of RNase H (lane 3). Three different annealing temperatures (but equal incubation times (10 min)) were tested for oligonucleotide binding. The extent and specificity of the hybrid was assessed through RNase H cleavage (lanes 4-6). **B)** Denaturing PAGE of cDNAs synthesized from purified mRNA by two different methods. The reverse transcription reaction was quenched in different time points. mRNAs purified using kit (lanes 1-3) generated a cleaner cDNA compared to the size exclusion purified preparation where truncated cDNAs were obtained (lanes 4-6). Note that for the experiment shown in panel (A), RNAP was walked until EC11 as described in the text. However RNAP was also walked to EC36 and the RNA labeled at its 3' end. This was done primarily to use this sample (lane A1) as a marker to pin-point the RNase H cleavage products.

transcription were performed under different conditions (Figure 22 B and A respectively). For hybrid formation, the optimum conditions were: incubation of mRNA with radiolabeled oligo for 65°C for 10 min followed by incubation on ice for 5 minutes. This procedure increased the binding specificity of the oligo to the mRNA judged by the disappearance of unspecific RNase H cleavage products; RNase H is an

endonuclease which cleaves the RNA in a RNA:DNA hybrid between mRNA and reverse transcription oligo, therefore if more than one cleavage product is observed it could be indication of unspecific oligonucleotide binding (figure 22A compare lanes 4-6). mRNA was purified by two different methods. Before any purification method was assayed, the reaction was separated from the streptavidin beads by DNAse I treatment (which cleaved the biotinylated DNA template), followed by centrifugation and recovery of the solid-phase-free supernatant. The first method tried was the commercially available RNA purification kit (RNAeasy kit Qiagen) was used following the manufacturer's instructions. In parallel, the transcription reaction was gel filtrated (using Bio-Rad biospin 6 columns) equilibrated in transcription buffer (for composition see materials and methods). Because the exclusion limit of this column permits the separation of nucleic acids up to 5 bp, the filtrated mRNA (131 nt long) would be separated from NTPs but not from RNA species greater than 5 nt. The purity and integrity of the mRNA was assessed by reverse transcription, which was also optimized; as seen in figure 21B lanes 1-3 the RNAeasy Kit (Qiagen) gave a cleaner mRNA compared with the gel filtration preparation (figure 22B lanes 4-6). Even though we were capable of synthesising full length cDNA from these mRNA preparations, background noise was always very high regardless of the purification method used and the yield of cDNA was always very low. Next, we decided to initiate translation and detect it by toeprinting.

Due to the lack of a method where we could test the activity of the purified ribosome and proteins (translation initiation) we decided to initiate translation through the non-enzymatic manner, eliminating the possibility of not being able to detect translation initiation complexes due to failure of the initiation factors.

As can be seen in (Figure 23 lanes 1-3), in the presence of N-acetyl-met-tRNA^{fmet} but in the absence of 70S ribosomes, full length cDNA was obtained. In the presence of initiated ribosomes (Figure 23 lanes 4-6), several cDNAs of different lengths were obtained and a dramatic decrease of full length cDNA was observed. Some of those truncated cDNAs were also observed in the absence of 70S suggesting that the secondary structure of the mRNA could have been responsible for these species and the interaction of the ribosome with the mRNA could have led to the stabilisation of the inhibiting structures.

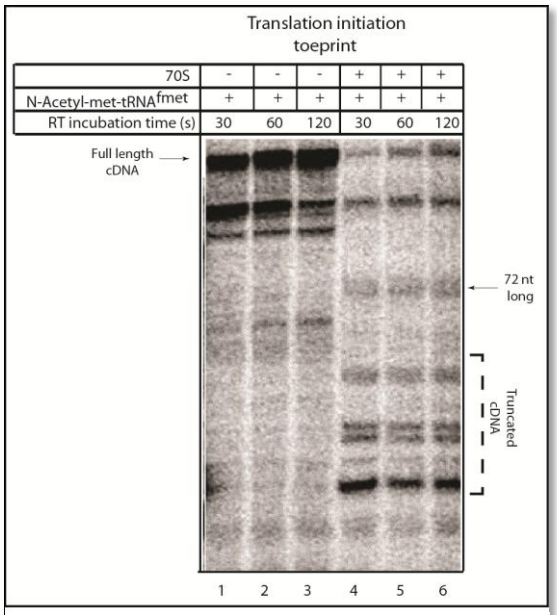

Figure 23 Toeprinting translation initiation complexes formed on purified mRNA from *E. coli* RNAP transcription reaction. Denaturing 12% PAGE resolving cDNAs obtained during toeprinting analysis. Lanes 1-3 time course of reverse transcription in the absence of ribosomes. Lanes 4-6 reverse transcription in the presence of non-enzymatic initiated ribosomes. Note how the full length cDNA is decreased and different cDNA species are formed. Expected toeprint for initiating complexes is 72 nt long.

A weak 72 nt long cDNA was also observed coinciding with the expected toeprint for initiating ribosomes which is normally situated ~15 nt downstream of the initiation codon AUG. However, we could not disregard the possibility that this particular cDNA could be due not to the initiating ribosome but also to the secondary structure of the mRNA. The only way in which we could relate a particular cDNA with ribosome toeprint was by comparing toeprints of ribosome initiated complexes with toeprint of translocated ribosomes which would result in shorter cDNA species.

In order to translocate the ribosomes, the GDP in the purified EF-Tu • GDP form was exchanged for GTP by incubation of the factor with phosphoenolpyruvate and pyruvate kinase in the presence of GTP. EF-Tu • GTP (the factor's active state) was then incubated with corresponding aminoacyl tRNA cognate to the codon to which the ribosome will be translocated to. This forms a ternary complex (EF-Tu • GTP • aminoacyl-tRNA) which upon incubation with an initiated ribosome will fill the ribosome's A site. Ternary complexes (TC) can be prepared for each and every codon, i.e. if the reading frame is F V K, there will be a TC(F), TC(V) and TC(K), allowing the

translocation of the ribosome by one codon at a time through the omission of the following TC (similar to transcription "walking" reaction); incomplete TC can also be prepared such as TC(FV) stalling the ribosome right before K in one reaction. After the ribosome A site has been filled with the appropriate TC, EF-G and GTP catalyze the translation elongation reaction.

For the translation translocation experiment, ribosomes purified from *E. coli* DH5α strain (a kind gift from Dr. Nicholas Watkins) were used as control for our 70S ribosomes preparation. As can be seen in Figure 24 (lanes 1,2), in the absence of ribosomes and in the presence of initiator tRNA, full length cDNA was obtained along with several short cDNA species of weak intensity.

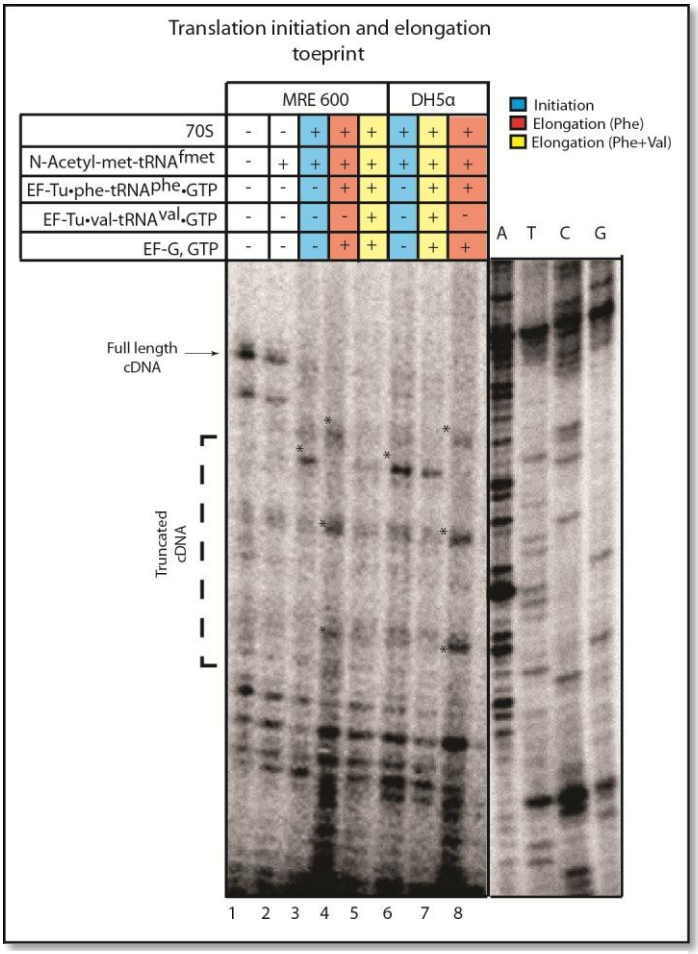

Figure 24. Toeprinting of translation initiation and elongation complexes formed on purified mRNA from *E. coli* RNAP transcription reaction. Different cDNAs were obtained depending on the translational state of the ribosome. Blue lanes: initiation complexes. Orange lanes: translocation by one codon (Phe). Yellow lanes: translocation by two codons (Phe+Val). Right panel: sequencing reaction. Note that ribosomes purified from DH5a (obtained as a gift from Dr. N. Watkins Newcastle University) were used only as a control of the preparation of 70S ribosomes prepared for this study. The aim of using ribosomes isolated from DH5α was to see if the same toeprints would be obtained in this preparation (which was known to be active through experiments carried out in Dr. Watkins laboratory), compared to the toeprints obtained using MRE600 isolated ribosomes.

For the toeprinting of translation initiated complexes a distinctive cDNA was obtained and the reverse transcription product corresponding to the full length cDNA was diminished (for both MRE600 and control ribosomes (lanes 3, 6)). When ribosomal translocation by one codon (Phe) (lanes 4, 8) was allowed to take place, the initiation toeprint was no longer detected and the appearance of other three cDNA's was observed in both ribosome preparations. In the case of ribosome translocation by two codons (Phe and Val), similar toeprints to those obtained for initiated ribosomes were observed (lanes 5, 7). Analysis of the sequence of the toeprints for initiation and translocation suggested that the obtained cDNAs were due to secondary structure of the mRNA formed upon unspecific ribosome binding to the mRNA which inhibited reverse transcriptase progression; this behaviour was the same for the control ribosomes suggesting that it was not due to the ribosome preparation. We took a closer look to the secondary structure of the mRNA used in our set-up (T7A1-g32) and compared it with the structure of the wild-type gene 32 mRNA (Figure 25).

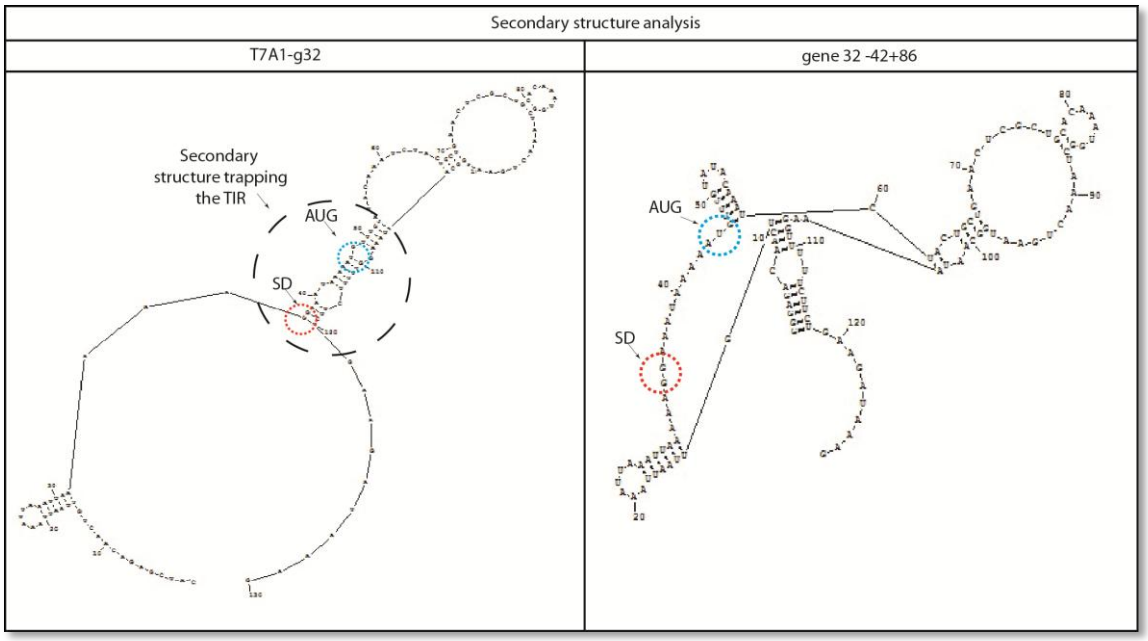

Figure 25. Secondary structure analysis of the transcribed region of T7A1-g32 mRNA and wild-type gene 32 -42+86 mRNA. Dashed large circle shows the region which compromised the translation initiation region (TIR) in the T7A1-g32 template. Red circle shows the location of the Shine-Dalgarno sequence. Blue circle indicates the position of the AUG start codon. T7A1-g32 ΔG=-17.6. Wild-type gene 32 -42+86 ΔG=-11.3. Secondary structure was predicted using RNA structure 5.3 bioinformatic package.

The analysis suggested that the TIR of the T7A1-g32 template was compromised in a hairpin loop structure, whereas in the wild-type gene 32 mRNA the TIR was free of inhibiting secondary structures.

4.3 Attempting to generate a translation system using mRNA synthesised by T7 RNAP

Three observations led us to completely redesign the approach for the set-up of the translational system: i) the transcription from the *E. coli* promoter T7A1 wouldn't yield high concentrations of mRNA. This factor combined with the purification of the mRNA could lead to a major loss of mRNA leaving us very little working amounts of mRNA. ii) The quality of the reverse transcription products (even in the absence of any translational factors) was very poor, judged by the large amount of unspecific short cDNAs obtained (truncated cDNAs). This could also relate directly to the quality of the mRNA preparation. iii) The observation of the unspecific ribosome binding probably due to the compromised TIR in the secondary structure of the mRNA. In order to overcome these obstacles we decided to program the ribosomes with purified mRNA synthesised by T7 RNAP.

4.3.1 *Programming the ribosome with purified mRNA from T7 RNAP transcription reaction.*

In order to synthesise the mRNA needed to program the translational system, the gene 32 region -42+86 (same region used in the previous experimental set-up) was put directly under the control of the T7 phage promoter which is recognised by T7 RNAP. Transcription reactions using T7 RNAP instead of *E. coli* RNAP produce higher concentrations of mRNA due to its efficient initiation and high processivity (considered to be 5 times faster than E. coli RNAP) (Chamberlin et al., 1970, Chamberlin and Ring, 1973, Golomb and Chamberlin, 1974). The DNA template (T7-g32) was used for i*n vitro* transcription reaction using T7 RNAP. The mRNA resulting from this reaction, was chlorophorm extracted, precipitated in ethanol followed by purification from denaturing PAGE (see methods) resulting in highly pure mRNA. The ribosomes were programmed to initiate both enzymaticaly and non enzymatic ways. Initiation and translocation (by one and two codons Phe and Val, respectively) events were monitored by toeprinting exactly as performed in the previous set-up.

In the absence of ribosomes a clean, full length cDNA product was obtained (Figure 26B lane 1) with only minor cDNA species which could relate to secondary structure of the RNA. For ribosomes initiated both enzymaticaly and non-enzymaticaly, a marked toeprint was obtained (Figure 26B lanes 2 and 5). This initiation toeprint was located 15

nt downstream the initiation codon AUG judged by sequencing markers, suggesting that the ribosome binding and initiation was specific. Translocation of the ribosome by one codon (Phe) was performed (Figure 26B lanes 3 and 6).

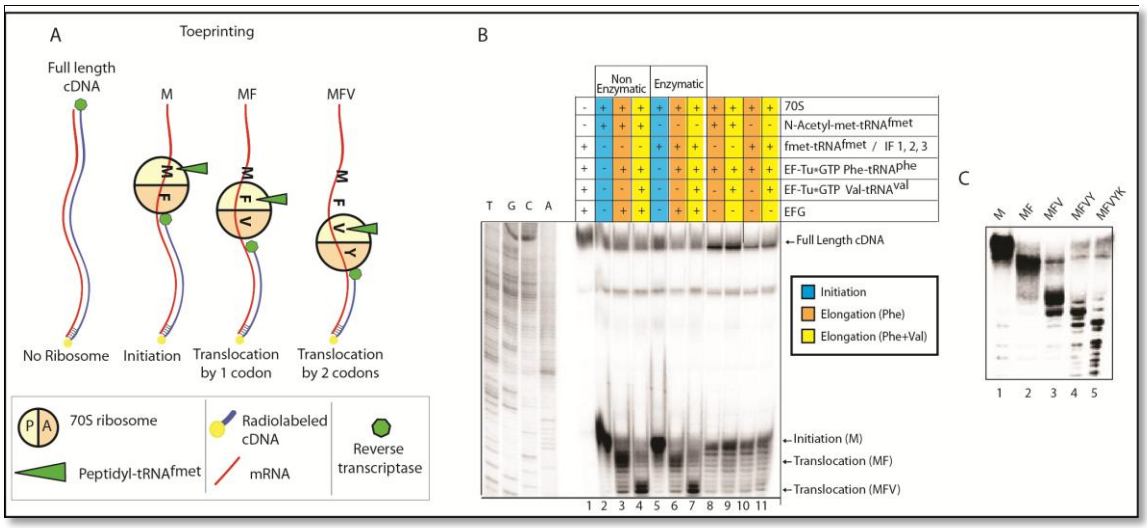

Figure 26. Toeprinting of translation initiation and elongation complexes formed on purified mRNA uncoupled from transcription by *E. coli* RNAP. A) Schematic representation of toeprints. B) Denaturing PAGE resolving cDNA's obtained in the toeprinting of Non-enzymatic and Enzymatic initiated translational complexes (blue lanes), one codon (Phe) translocated complexes (orange lanes) and by two codons (Phe+Val) (yellow lanes). C) Denaturing PAGE of the cDNA's obtained in the toeprinting of initiated (lane 1) and stepwise translocated ribosomes (lanes 2-5) for up to 5 codons. Note that the MFVYK toeprint (lane 5) has several truncated cDNAs because of the short distance between the front edge of the ribosome and the 3' end of the mRNA making the progression of the reverse transcriptase difficult. These gels were used for the qualitative assessment of translation initiation and step-wise translocations. These gels were not used for quantification purposes.

The obtained toeprint was exactly 3 nt downstream from the initiation toeprint suggestive of ribosome translocation by one codon. The one codon translocation toeprint was not obtained in the absence of EF-G (Figure 26B lanes 8 and 10), indicating that the translocation of the ribosome was EF-G dependent and it was not due to ribosome slippage, reassuring the specificity of the system. The ribosome was translocated by two codons and a toeprint 6 nucleotides downstream from the initiation toeprint was obtained (Figure 26B lanes 4 and 7); this translocation event was proven to be EF-G dependant (Figure 26B lanes 9 and 11). Further stepwise translocation of the ribosome by three and four codons was done obtaining toeprints corresponding to the correct translocated ribosomes (Figure 26C).

4.4. Discussion

During the development of the *in vitro* translation system from purified components we found key elements that not only allowed us to get the system to work but also provided

parameters to take into account for its coupling with transcription. For instance, the purity, integrity and secondary structure of the mRNA is fundamental for the system to work. Even though the ribosome has helicase activity (Takyar et al., 2005a), this property is absent during translation initiation when the TIR is being recognised. Techniques such as toeprinting are very reliable but require highly pure mRNA. If the mRNA is degraded or if it is highly structured, the reverse transcriptase will generate truncated cDNA's obscuring the actual toeprint.

The data obtained in the toeprint experiments (Figure 26 B, C) proved that the translational system was active and that we could initiate translation enzymaticaly and non-enzymaticaly, providing flexibility to the system i.e. when translation is coupled to transcription and GTP is not desired in the reaction, non-enzymaticaly initiation can be performed. Through toeprinting we also observed that we could perform step-wise translocation of the ribosome a tool which, once translation coupled to transcription, will enable us to take "snapshots" of the interactions between translation and transcription at different distances until collisions are allowed to take place.

Results chapter 5. Setting up an *in vitro* transcription system coupled to translation from DNA templates containing promoter.

5.1 Introduction

Having our translation system functional and complying with the specific requirements for our study, we were ready to couple it to transcription. In an attempt to keep the system close to the *in vivo* situation, we decided to couple translation on nascent mRNA during transcription by *E. coli* RNAP on a DNA template containing an *E. coli* promoter. In this chapter the strategy taken for the coupling of the two machineries will be described. The methods used to detect the translating ribosome in the coupled system and some interactions observed amongst the two apparatuses will be discussed.

5.2 Transcription from DNA templates that contain a promoter

For the analysis of the effects of translation on transcription in a coupled transcription translation system (CTT) we required to have the ability of placing RNAP in defined transcriptional states such as paused, roadblocked, backtracked or pre-translocated. To

be able to do so, a biotinylated DNA template carrying the strong T7A1 promoter followed by the gene 32 TIR and further downstream one of such signals was engineered (template T7A1-g32ORF9 (Figure 27). This is the parental template used for the studies regarding the assembly and fine tuning of the system and was modified depending on the desired transcriptional signal needed for studies concerning the outcomes of the interactions between the two machineries. To reach and position RNAP near or on such signals, first, a stable elongation complex (EC) was obtained by transcription from the T7A1 promoter in the presence of an RNA primer (CAUC), GTP, ATP and CTP, which allowed elongation only until position +11 (relative to transcription start site, +1). This walking ensures proper transcription initiation and also the formation of high salt and competitors (such as heparin) resistant EC, permitting the washing of the initial EC to ensure purity of the mRNA synthesised during the next walking steps. Transcription elongation complexes containing 11 mer RNA (EC11) were then immobilized on streptavidin beads and were washed to remove unused

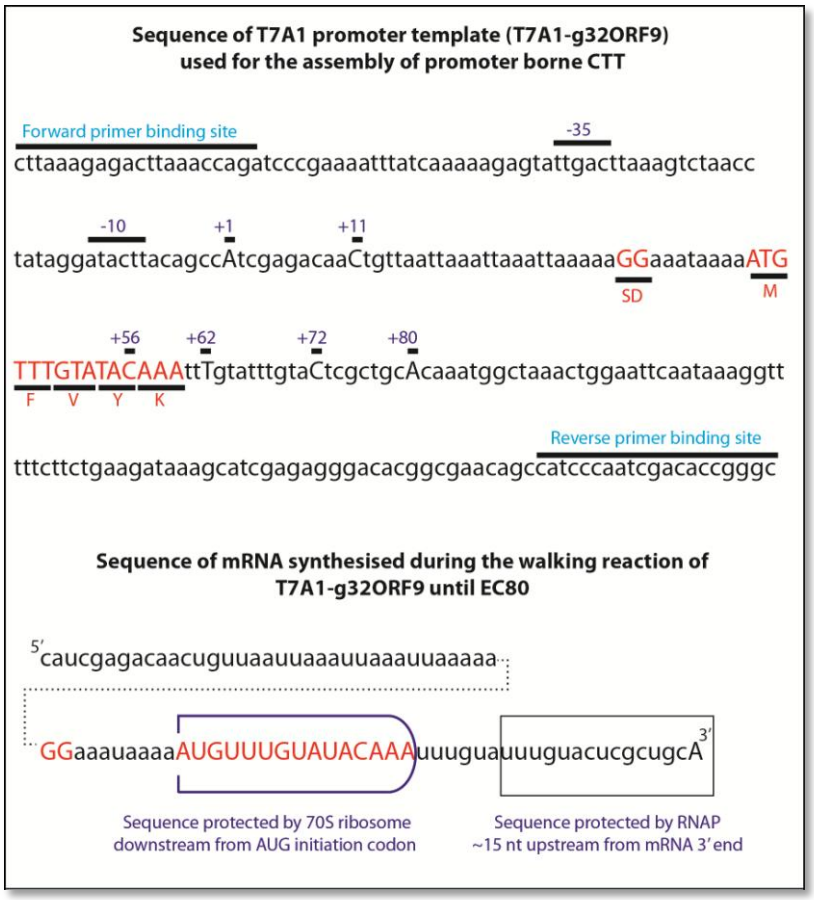

Figure 27. DNA and mRNA sequences of T7A1 promoter template used for the assembly of promoter borne CTT. Top panel: PCR generated DNA template sequence showing oligonucleotide binding sites (for PCR reaction), -35 and -10 promoter elements, transcriptional +1 and every transcriptional walk needed to get to EC80. TIR and five first codons are shown. Bottom panel: Transcript of T7A1-g32ORF9 template until position EC80. TIR and five first codons are shown in red. Ribosome front edge is based on observations made through toeprinting experiments. Rear end of RNAP (black square) has been elucidated by footprinting experiments.

RNAP, σ^{70}, promoter and non-specific complexes and unincorporated NTPs from the reaction. Next, RNAP was walked to a particular position. From position +11 RNAP can be walked to any desired point on the template (Figure 27). RNA in transcription EC can be labelled in the body during 11 mer synthesis or during further walking, or alternatively at the 3' end, by incorporation of radiolabeled NTPs. As a test, we walked RNAP in a stepwise manner by using subsets of NTPs, which were washed away after each step, to position +80 (EC80). Taking into account the observations made in our toeprint experiments, the front edge of the ribosome could be protecting 15 nt from the first nucleotide in the initiation codon AUG towards the 3' end of the mRNA, suggesting that the ribosome could need ~15 nt (from the initiation codon AUG) in order to bind to the mRNA and initiate translation. The boundary of the rear end of RNAP has been identified to be localised 15 nt upstream the 3' end of the mRNA (Korzheva et al., 2000). EC80 was chosen because: i) it could provide enough room for both machineries to be bound to the mRNA at the same time (according to the proposed parameters) (see Figure 27 bottom panel); there are 6 nt in between both the ribosome and RNAP (for a total of 21 nt) providing extra space in case either system could require more space and, ii) according to the analysis of the secondary structure of this complex, the TIR is not compromised.

In order to assemble the coupled transcription-translation system (CTT), once EC80 was obtained, the transcriptional system was washed with translation buffer preparing the system to be coupled to translation. 70S ribosomes were added and initiated non-enzymaticaly by the addition of N-acetyl-met-$tRNA^{fmet}$. To analyse if translation initiation had taken place, we could not use toeprinting as we did for the uncoupled translational system. As described previously, toeprint experiments are based on reverse transcription which requires the 3' end of the mRNA to be free in order to bind a radiolabeled oligonucleotide which will serve as a primer for the reverse transcriptase to start cDNA synthesis. In the previous toeprint experiments the 3' end of the mRNA was free allowing us to follow translation initiation and elongation by toeprinting, but once translation has been coupled to transcription, this technique is no longer possible due to the fact that RNAP occupies the 3' end of the mRNA. Therefore, we needed to search for a method which would allow us to follow the ribosome independently of reverse transcription.

5.3 Assessing translation initiation by RNase H probing

To identify if coupling of translation on transcription had taken place, ribosome initiation was probed using RNase H cleavage. RNase H (Ribonuclease H) is an endoribonuclease that specifically hydrolyzes the phosphodiester bonds of RNA which is hybridized to DNA but does not digests free single stranded or double stranded RNA or single or double-stranded DNA.

To perform RNase H probing in our system, after translation had been initiated, a DNA oligonucleotide complementary to a specific stretch of the TIR of the mRNA would be allowed to anneal followed by the addition of RNase H. If translation initiation had taken place, the ribosome would not allow the binding of the oligonucleotide preventing RNase H cleavage; if the translation initiation had not taken place, RNA:DNA hybrid would be formed and RNase H cleavage could occur (Figure 28A). To test this approach we walked RNAP to EC80 (labelling the mRNA in the EC11 with the incorporation of a radiolabeled NTP) (Figure 28B) and assembled the CTT by the addition of 70S ribosomes which were initiated non-enzymaticaly. Due to the high AU content of the TIR of the mRNA used for these experiments, two oligonucleotides were used to minimise the possibility of obtaining false positive results due to the failure of the oligo to anneal to the mRNA. Oligo1 which hybridises the mRNA three nucleotides upstream the SD until the initiation codon AUG, and Oligo2 which hybridises from the SD to the second codon (F). As can be seen in (Figure 28C lane 1), in the presence of the translational machinery and the oligonucleotide but in the absence of RNase H no degradation of the mRNA was observed. In the presence of initiated ribosomes and oligonucleotides 1 and 2, very subtle RNase H mediated cleavage was observed (Figure 28C lanes 2,4) compared to the cleavage observed in the absence of initiated ribosomes (Figure 28C lanes 3,5). No RNase H cleavage was observed in the presence of the translational apparatus and RNase H (Figure 28C lane 7).

This experiment was the first indication of assembly of translational complexes on active transcription elongation complex, however several questions arose. The fact that we observed slight cleavage in the presence of ribosomes could be indication of either poor ribosome binding (oligo competing with the 70S ribosome for the TIR) or poor occupancy of the ribosomes in our system (some translational complexes occupying transcriptional complexes). These two possibilities would make our system sub-optimal

for the analysis of the interactions between the two machineries due to the existence of a mixed population of EC occupied and unoccupied by ribosomes would obscure the analysis of interactions.

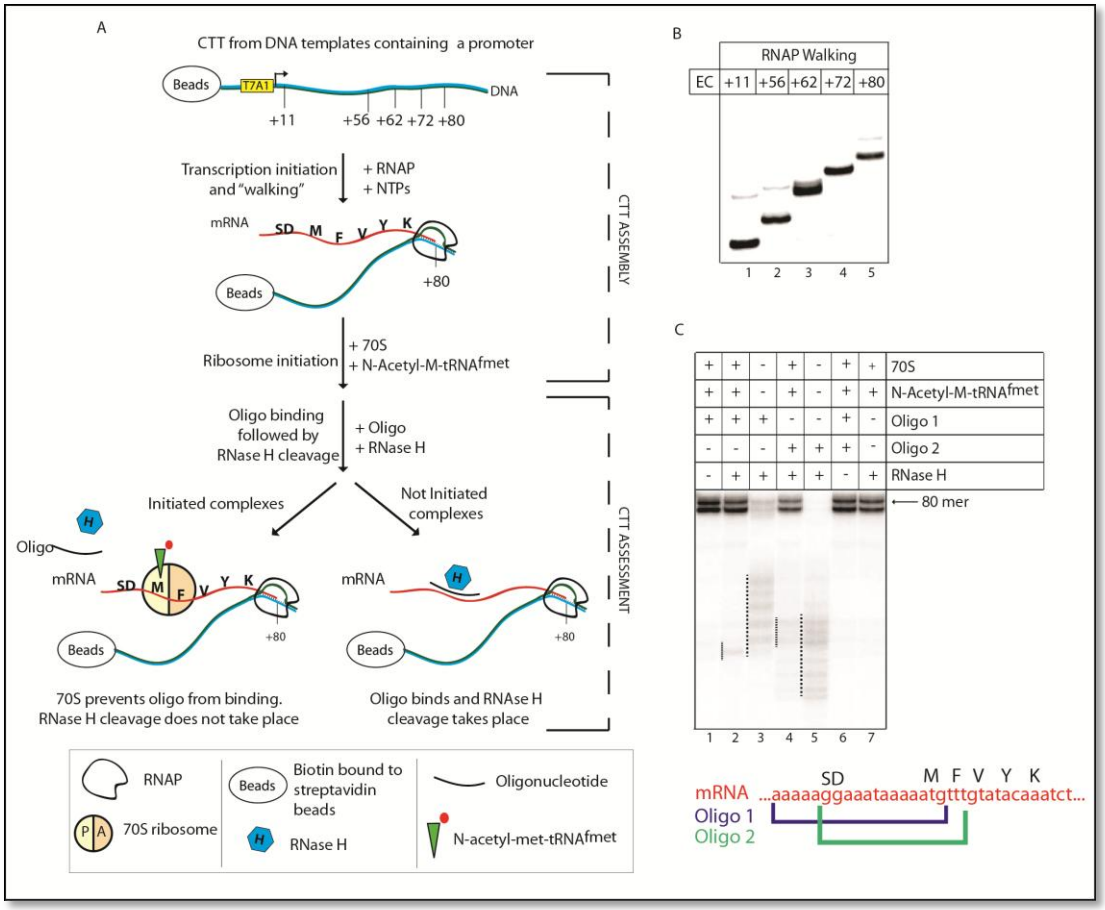

Figure 28. CTT from DNA templates containing a promoter: Assembly and translation initiation assessment through RNase H probing. A) Schematic representation of CTT from DNA templates containing promoter assembly and RNase H probing. B) Denaturing PAGE of the mRNA obtained during the transcriptional walk to synthesise EC80. C) Denaturing PAGE of the radiolabeled mRNA in the CTT. RNase H probing of translation initiation complexes using Oligo 1(lanes 2,3) and oligo 2 (lanes 4,5) indicate the presence of the ribosome bound to the mRNA. Sequence shows the TIR and 4 codons. The binding sites for oligo1 and 2 are indicated (see text) Note that in the main complex EC80 two bands can be observed, this is due to read-through of the complex by residual ATP present in the preparation of N-Acetyl-met-tRNA^{fmet}. Note that for the EC80 two bands are observed due to phosphorolysis of the complex caused by presence of phosphates in the RNase H buffer.

This approach only allows us to identify ribosomes bound to mRNA but it does not determine whether or not the ribosome had been initiated. We also realised that ribosome translocation could not be followed using RNase H probing because the ribosome has an intrinsic helicase activity which would easily displace the oligo rendering the probing ineffective; if addition of the oligonucleotide would be done after ribosome translocation only ribosomal complexes which have translocated by long stretches could be probed due to the limiting size of the oligonucleotide but no step-wise translocation of the ribosome could be monitored.

The system required a direct and specific tool to assess the ribosome during both initiation and translocation.

5.4 Assessing ribosomal translation initiation and translocation by RelE printing

In bacteria, specific genes have been found to encode peptides (generally ~100 amino acids long) which can regulate cell growth and death under various stress conditions. Such peptides are commonly known as toxins and can target a main component of several cellular pathways such as ATP synthesis, DNA replication, RNA stability, cell-wall and protein synthesis (Yamaguchi et al., 2011). These toxins are co-transcribed and co-translated with its regulatory cognate antitoxin from an operon referred to as toxin-antitoxin (TA) operon. During normal conditions, the toxin and its anti-toxin form a highly stable complex which prevents the toxin from acting on its target. This interaction is destabilised during environmental stress or starvation conditions through the action of proteases such as Lon which degrade the antitoxin leaving the toxin free to exert its function (Winther and Gerdes, 2011).

Out of the 33 TA which have been identified in *E. coli*, 12 have been characterised. 8 of those 12 TA cleave mRNA affecting its stability (Christensen-Dalsgaard et al., 2008). This group of 8 interferases has been divided into two types depending on the manner on which they cleave mRNA: ribosome-independent and ribosome dependent mRNA interferases being the latter of particular interest for our experimental approach. One of the ribosome dependant interferases, the toxin RelE, binds to the 30S subunit through the vacant A site of the ribosome and cleaves RNA between the first and second nucleotide of the codon (Overgaard et al., 2008). *In vivo,* RelE cleavage has been observed for ribosomes reaching a termination codon interfering with translation termination and peptide release (Yamamoto et al., 2002) and also for ribosomes stalled through amino acid deprivation (Christensen et al., 2001). *In vivo,* RelE cleavage has been successfully used for the investigation of paused and stalled ribosomes (Pedersen et al., 2003) and the analysis of ribosomal frameshifting in mitochondria (Temperley et al., 2010). We, therefore, decided to use RelE in our experimental set-up as a tool for the localisation of the initiating and translating ribosome. RelE cleavage was first assessed by reverse transcription. As explained in the scheme in Figure 29A, in the

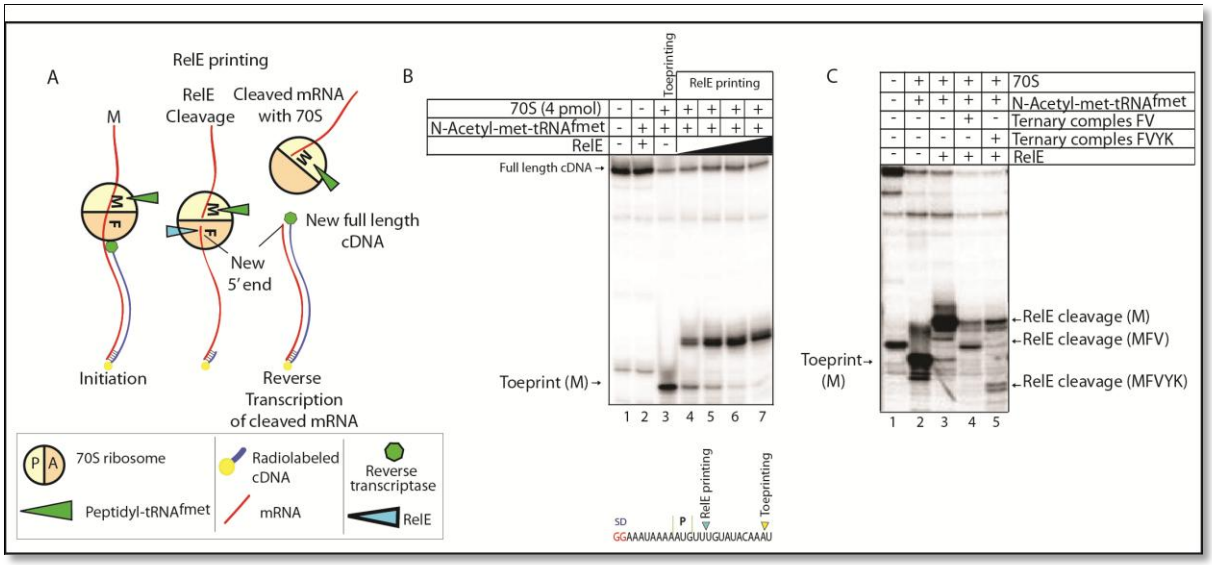

Figure 29. Following translation initiation and elongation by RelE printing. A) Schematic representation of reverse transcription in the absence and in the presence of RelE. B) Denaturing PAGE resolving the cDNAs obtained in the presence of initiated ribosomes in the absence of RelE (lane 3) and in incrementing concentrations of RelE cleaving translation initiation complexes (lanes 4-7) the concentrations of RelE were 1, 4, 8, 14 pmol respectively. C) RelE printing of initiated (lane 3) and translocated ribosomes by 3 codons MFV (lane 4) and 5 codons MFVYK (lane 5). Note that the concentration of mRNA used in lanes 4 and 5 is lesser than the concentration used in lanes 1-3, explaining the difference in intensity.

presence of an initiated ribosome but in the absence of RelE, a cDNA of the toeprint of initiation complex is obtained. When RelE is added and has cleaved the mRNA, a new 5' end is formed which is not longer protected by the boundaries of the ribosome. When reverse transcription is made on the cleaved mRNA, the reverse transcriptase is capable to proceed further downstream due to the absence of the physical blockage caused by the ribosome generating a longer cDNA compared to the toeprint for initiating ribosomes.

We first analyzed RelE cleavage in the absence of the transcription machinery by primer extension using exactly the same experimental setup used for toeprinting. RelE cleavage was tested in the translation initiation complex, which had the phenylalanine codon UUU in the A-site. As can be seen in Figure 29B, in the absence of ribosomes, no cleavage was detected, demonstrating the specificity and dependency of the cleavage event on the ribosome (Figure 29B lane 2). In the absence of RelE a toeprint for the translation initiation complex was observed (Figure 29B lane 3). With increasing concentrations of RelE we observed the diminishing of the toeprint and the appearance of cDNA corresponding to the location of the A-site, where RelE cleavage had occurred (compare Figure 29B lane 4 with lane 3). As higher concentrations of RelE were used, complete disappearance of the toeprint band was obtained (Figure 29B lane 7). We also tested RelE cleavage on translocated ribosomes by three and five codons (Figure 29C lanes 4,5). Although we were capable to locate the position of the translocated

ribosomes the cleavage was very weak due to the different affinity which RelE has to different codons. Taking these results into account, we chose RelE as one of the tools to follow translation in the CTT.

5.5 Assessing ribosomal translation initiation and translocation by RelE printing in CTT from DNA template containing promoter

To test if the translational system was being coupled to the transcriptional apparatus, we initiated transcription and walked RNAP to EC80 on T7A1-g32ORF9 template (the same template used for the experiments described for RNase H assay), radiolabeling the mRNA in its 5' proximal end. EC80 were transferred into translation buffer by washing the immobilized complexes, and 70S ribosomes were added. After ribosome binding, translation was initiated by the addition of N-acetyl-met-tRNA^{fmet}, followed by the addition of RelE (Figure 30A). In the presence of initiated ribosomes, we observed a 49 nucleotide long RelE cleavage product, which corresponded to the expected position of the vacant A-site of the initiating ribosome (Figure 30B, lane 5). The intrinsic cleavage (phosphorolysis) of the RNA (Orlova et al., 1995) in the EC80 by RNAP occurred due to a high concentration of phosphate in the RelE buffer, which was required to maintain the solubility of RelE. The extent of RelE cleavage demonstrated that the occupancy of the mRNA by the ribosome was greater than 50% (Figure 30B, lane 5). Ribosome translocation was also tested by adding a mixture of ternary complexes (allowing synthesis of MFVY tetrapeptide) in the presence of GTP and EF-G. We observed appearance of a 58 nucleotide long RelE cleavage product, corresponding to the A-site of the ribosome translocated by 3 codons (Figure 30C, lanes 3 and 7). Note however, that because RelE has a different affinity to various codons (32), the ratio between RelE cleavage products in initiated and translocated complexes does not reflect the actual proportion of ribosomes that escaped into elongation upon ternary complex addition. Note that some read-through of EC80 occurred in the presence of GTP used during ribosome translocation (band above +80 in Figure 30C, lanes 3 and 4). This meant that DNA template sequences that allow efficient read-through from stalled ECs in the presence of GTP may not be suitable for assembly of CTT.

To test if RNAP remained active in the CTT after ribosome initiation and translocation, EC80 was allowed to be elongated in the presence of all four NTP's (before RelE addition). As seen in Figure 30C, lanes 6 and 7, all transcription elongation complexes were active and resumed elongation. RelE cleavage products for both initiated and

translocated ribosomes observed after chasing of transcription complexes were the same as in CTT assembled on EC80 (Figure 30C, lanes 6 and 7). Additionally, similar results were obtained when using f-met-tRNAfmet and IF-1, IF-2, IF-3 or N-acetyl-met-tRNA^{fmet} for the initiation of translation (will be shown in an upcoming chapter), suggesting that both ways of initiation could be used according to our experimental needs. Having the CTT from DNA templates containing a promoter assembled, we decided to conduct a series of experiments to explore its possible use in the exploration

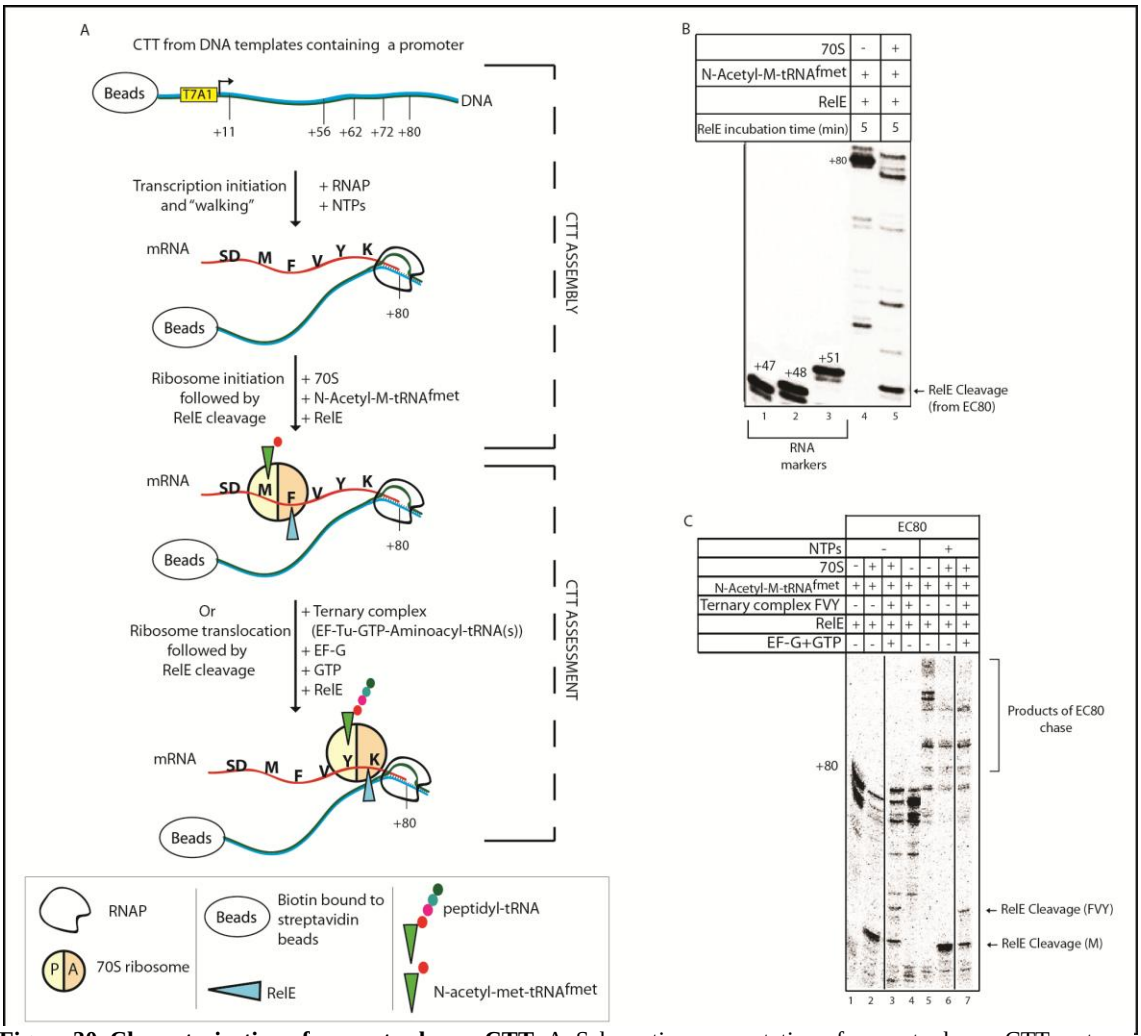

Figure 30. Characterization of promoter borne CTT. A. Schematic representation of promoter borne CTT system assembly with non-enzymatic translation initiation. Biotinylated DNA containing T7A1 promoter utilized by E. coli is immobilized on streptavidin beads. *E. coli* RNAP initiated from the promoter is "walked" to a desired position. Translation initiation and elongation complexes formed on the mRNA can be analyzed by RelE cleavage. B. RelE mapping of translation initiation complexes in CTT system assembled with promoter borne transcription EC80 (lanes 4-5). EC80 was radiolabeled at its 5' end proximal part. RNAs from transcription ECs stalled at positions +47, +48 and +51 (lanes 1-3) were used as size markers to determine the size of the RelE cleavage product. Note that degradation of EC80 is caused not only by RelE cleavage but also by RNAP dependent phosphorolysis by high phosphate of RelE storage buffer. C. PAGE of radiolabeled mRNA of CTT (assembled with stalled EC80), in which translation initiation (lanes 2 and 6) and elongation (in the presence of EF-G; lanes 3 and 7) are probed by RelE cleavage. In lanes 5-7 EC80 was chased in the presence of all NTPs after translation initiation and elongation but before RelE cleavage. EC80 was obtained as in panel B, and the transcript was radiolabeled at its 5' end proximal part. RelE cleavage products were identified as in panel C. A weaker band of RelE cleavage product after ribosome translocation is explained by different activity of RelE on various codons. Note some transcription read through from EC80 (lanes 3, 4) in the presence of GTP required for translocation. Black vertical lines separate lanes originating from one gel which were brought together.

of interactions between the two machineries. Two sets of experiments which serve as a general example of the use of promoter borne CTT will be described: i) the effect on RNAP backtracking by the translating ribosome and ii) the effect on roadblocked transcription elongation complexes by the translational machinery.

5.6 Assessing the effect on RNAP backtracking by the translating ribosome in CTT from DNA template containing promoter.

As explained in the introductory chapter, RNAP has the ability to move backwards along the DNA template, disengaging the 3' end of RNA from the template DNA strand and the active centre, while the rear end of RNA-DNA hybrid RNA anneals back to the template, a phenomenon called backtracking. In the backtracked complexes, the 3' end of the RNA exits through the RNAP secondary channel losing the register with the active centre (Shaevitz et al., 2003). These arrested complexes require auxiliary factors such as cleavage factors GreA or GreB which upon cleavage of the RNA (through the secondary channel of RNAP) return the 3' end of the RNA back in register with the active centre (Borukhov et al., 2001, Borukhov et al., 1993); GreB can also be used *in vitro* to analyse the extend of backtracking of RNAP by looking at GreB-mediated cleavage products (Roghanian et al., 2011a, Epshtein and Nudler, 2003) .

To test if actively transcribing ribosomes could stop RNAP from backtracking in our CTT, we constructed a template which would allow ribosome binding, initiation and translocation towards a backtracking-prone transcription complex (Figure 31A). The backtracking signal used in our template (referred to as T7A1-backtrack template) is a signal naturally found in the position +27 (relative to the transcriptional +1 site) of the T7A1 gene. It has been demonstrated that RNAP backtracks by 16 nt when it passes or is stalled on this particular site by omission of the next cognate NTPs (Kireeva and Kashlev, 2009a). In T7A1-bactrack template, the backtracking signal was placed 28 nt downstream EC80. Upstream EC80 the sequence was exactly the same as in the T7A1-g32ORF9 template used for RNase H and RelE in CTT experiments. Unlike normal elongation complexes, backtracked complexes are susceptible to GreB mediated cleavage thus the extent of backtracking can be determined following the GreB-mediated cleavage products. As can be seen in Figure 31B (lane 1), no apparent degradation of the RNA in the backtracked complex was observed even after 20 min of incubation at 37°C in the absence of GreB factor; the experiment was performed at low

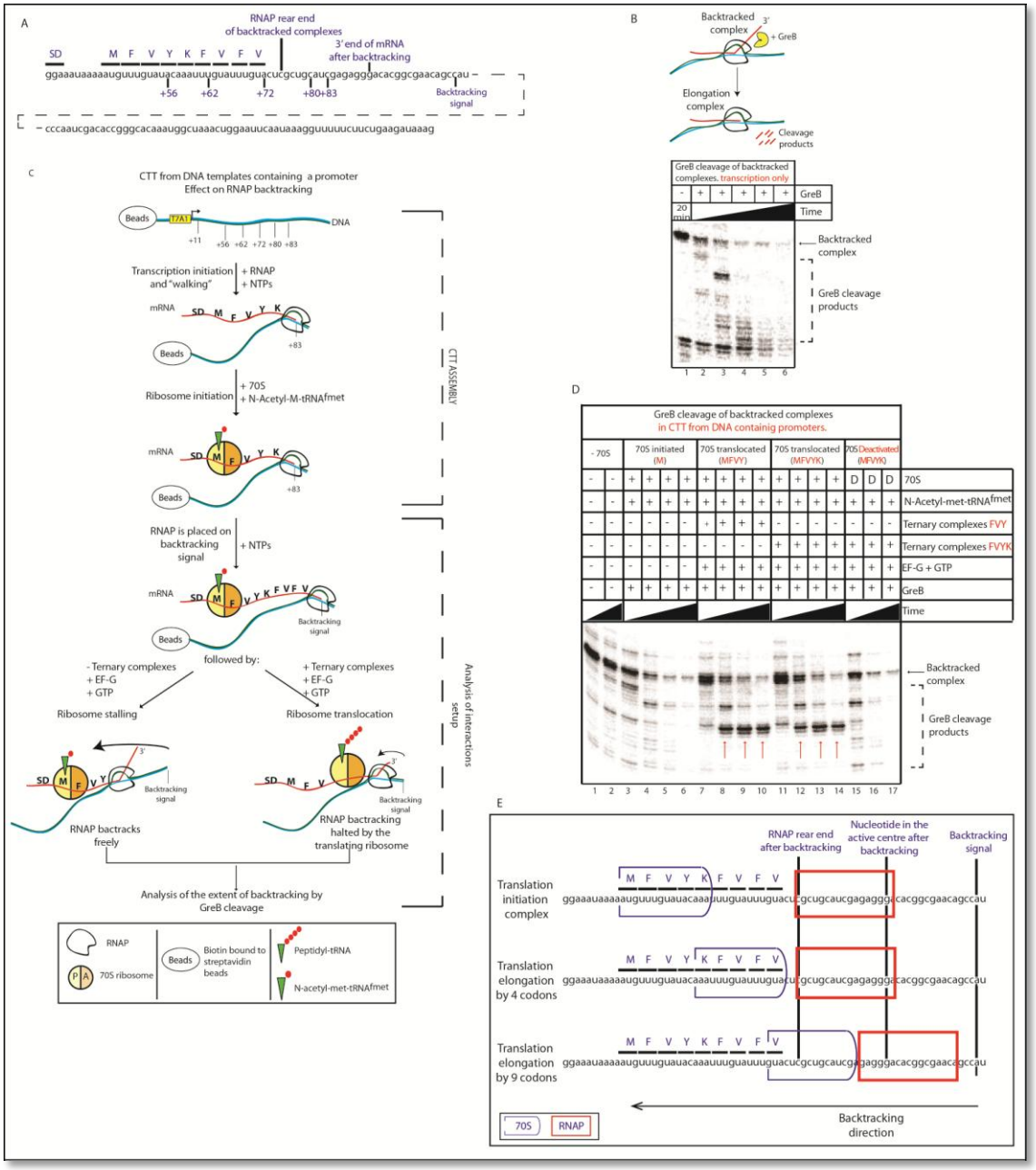

Figure 31. Assessment of effect on RNAP backtracking by the translating ribosome in CTT from DNA template containing promoter. A) Sequence of T7A1-backtracking template used for the analysis of the effects on RNAP backtracking by the translation machinery. Translational features (translation initiation and used codons) are indicated by name and a black horizontal line. Transcriptional steps taken are shown along with the position where RNAP was walked to for backtracking to take place (backtracking signal). 3'end of mRNA after backtracking is annotated along with the rear end boundary of RNAP after backtracking. B) PAGE of GreB cleavage products of backtracked complexes in the absence of the transcriptional machinery. RNAP was walked to the backtracking signal and allowed to backtrack. Backtracked complexes are not degraded in the absence of GreB factor (lane 1). Upon addition of GreB, degradation of the backtracked complexes can be seen occurring through time. C) Schematic representation of experimental set-up to test the effect on RNAP backtracking by the translational machinery (see text). D) PAGE of backtracked complexes in CTT. In the absence of GreB (lanes 1-2) no degradation of the backtracked complex was observed. Upon addition of GreB in initiated complexes (lanes 3-6), similar cleavage pattern was observed compared to the one obtained in the absence of the translational system (compare with figure 30B) and when heat deactivated 70S ribosomes were used (lanes 15-17). In GreB treated backtracked complexes with translocated ribosomes by four codons (MFVY) and nine codons (MFVYK) a GreB resistant transcription elongation complex was obtained (lanes 7-14 indicated by red arrows). No change in GreB cleavage was observed when translational components necessary for translocation of the ribosome by 9 codons was observed when heat deactivated 70S (D) ribosomes were used (lanes 15-17). Deactivation was performed by incubating 70S ribosomes for 15 min at 90 °C in translation buffer. The amount of GreB used in this experiment was 5 pmol/lane. E) Schematic representation of the distances between the ribosome (purple box) and backtracked RNAP (red rectangle) in different translational complexes. Top panel , distance in translation initiation (~14 nt) allows RNAP to backtrack further; middle panel, ribosomes translocated by 4 codons, are ~2 nt away from backtracked RNAP possibly interfering with further backtracking. Bottom panel, situation where ribosomes are translocated by 9 codons; ribosomes would "push" backtracked RNAP by 11 nt. As seen in D), no difference is observed between 4 and 9 codon translocation

pH (7.4) and low Mg2+ concentration (5 mM) conditions to avoid intrinsic cleavage. In the presence of GreB, cleavage of the mRNA was observed (Figure 31B lanes 2-6). Backtracking in this particular signal can continue until RNAP finds a thermodynamically favourable transcription elongation complex which impedes it from backtracking.

We tested the effect on RNAP backtracking by the translating ribosome in CTT from promoters using T7A1-backtracking template. To do so, the mRNA was labelled in EC11 by addition of a radiolabeled nucleotide. RNAP was then walked until EC83 (Figure 31C). At this position 70S ribosomes were added and initiated non-enzymaticaly. After translation initiation, RNAP was allowed to walk to the backtracking signal (this step was performed on ice to avoid premature backtracking). Then, translating ribosomes were either left in the initiation configuration or were allowed to translocate by 4 or 9 codons with concomitant RNAP backtracking. The backtracked complexes were then treated with GreB. As can be seen in Figure 31D (lanes 1, 2), no cleavage was observed in the absence of GreB. When CTT was assembled and initiated and 70S ribosomes were not translocated (left in translation initiation complex Figure 31E top panel), GreB cleavage pattern was similar to the one observed in the absence of the translational machinery (Figure 31B). When the ribosomes were allowed to translocate by 4 and 9 codons, formation of a stable, GreB-resistant transcription elongation complex was obtained (lanes 7-14 indicated by red arrows). This complex was located 16 nt upstream of the backtracking signal and coincides with the position where the active centre should be after backtracking by the expected 16 nt (see figure 30E middle panel). The formation of such complex might be due to the translocated ribosomes acting as a barrier which impeded further backtracking of RNAP and stabilised it in the post-translocated state; the effect of barriers in the mRNA which affect RNAP backtracking has previously been observed when DNA oligonucleotides were allowed to anneal to the mRNA right behind RNAP in the same backtracking sequence used in this experiment (Komissarova and Kashlev, 1997). Interestingly, the GreB resistant elongation complex was not obtained when conditions for translocation by 9 codons were given using heat deactivated 70S ribosomes (Figure 31D, lanes 15-17). However, the fact that exactly the same effect was obtained when the active ribosome was translocated by 4 and 9 codons could not be explained. If the ribosome were able to physically exert strength over the backtracked

RNAP and “push” it, little or no backtracking in the 9 codon translocated scenario was to be expected (Figure 31E bottom panel).

On the other hand, it is possible that the formed transcription elongation complex was stable enough not to be affected by the translating ribosome. Nevertheless, the data in this experiment might suggest that the translating ribosome was the cause for the formation of the GreB resistant elongation complex. To further assess the effectiveness of our promoter borne CTT we decided to analyse the effect on roadblocked RNAP by the translating ribosome.

5.7 Assessing the effect on roadblocked transcription elongation complexes by the translating ribosome in promoter borne CTT

It has been observed *in vivo* that the rate of transcription is not altered when RNAP is challenged with obstacles on the DNA such as histone-like proteins, repressors, etc. Interestingly, the opposite effect was observed *in vitro* when RNAP encountered roadblocks such as the lac repressor and LexA protein (Pavco and Steege, 1990). As mentioned in RNAP-RNAP interactions chapter, Epshtein and co-workers, observed that *in vitro* efficient transcription through roadblocks (EcoRI mutant EcoRQ111 and the lac repressor) depended on multiple rounds of transcription initiation. They observed that the first RNAP encountering the roadblock backtracks (by 5 nt judged by GreB cleavage) after colliding with it; then a second RNAP molecule rescues the backtracked RNAP but is not capable to exert enough force to alter the binding of the roadblock, needing a third RNAP molecule to dislodge the roadblock (Epshtein et al., 2003).

These observations led us to test if the translational system could also have any effects after colliding with roadblocked RNAP observing both roadblock dislodgment and RNAP backtracking behaviour caused by the collision between RNAP and the roadblock in our promoter borne CTT system.

For this experiment, a DNA template (referred to as T7A1-EcoRI) containing an EcoRI binding site 16 nt downstream of EC80 was constructed (Figure 32A). Upstream EC80 the template was identical as T7A1-Backtracking template used in the previous chapter (5.6). The template was designed taking into account the following parameters: i) the distance between RNAP active site and the first nucleotide in the EcoRI binding site

(GAATTC) is 15 nucleotides after collision (Nudler et al., 2003) . ii) Knowing the exact 3' end of the roadblocked RNAP complex allowed us to determine the boundary of the rear end of RNAP, by counting 15 nt upstream the 3' end of the RNA or active centre; giving a total distance of 24 nt downstream from the initiation codon AUG, a distance which is sufficient for translation to initiate as was observed in the characterisation of the CTT from promoters (chapter 5.2) where 21 nt were amongst the initiation codon and the rear end of RNAP when positioned in the EC80. The EcoRQ111 roadblock was purified as described in materials and methods (table 2). Before assembling the CTT, the formation of roadblocked complexes was analysed in the absence of the translational apparatus. Firstly, transcription from T7A1-Roadblock template was initiated and RNAP was walked to EC11 (Figure 32A). In order to ensure that only one molecule of RNAP enters into elongation, the system was immobilized onto streptavidin beads and then washed with transcription buffer containing heparin (100 μg/mL). Heparin binds to free RNAP and competes with weakly DNA bound RNAP but does not affect transcriptional complexes which have already formed stable transcription elongation complexes (Nudler et al., 1996) thus leaving the EC11 intact. After heparin treatment, high salt and low salt washings were performed. High salt and heparin washings were performed to ensure that all transcriptional complexes to be analysed originated from EC11 only; this washings prevent reinitiation by RNAP molecules which are weakly bound to the promoter. The mRNA was radiolabeled in the body during walking from EC11 to EC56 by addition of α-[^{32}P]-ATP (Figure 32B lane 1). At EC56, the reaction was divided into two aliquots; to one aliquot, all NTPs were added synthesising full length mRNA (Figure 32B lane 2). To the second aliquot, EcoRQ111 was added (2 fold over DNA concentration) and allowed to bind to the EcoRI site at 37°C for 5 min followed by addition of all NTPs. We observed a very strong truncation of full length mRNA synthesis indicating the formation of the transcriptional roadblocked complex (compare Figure 32B lanes 2 and 3[red arrow]) along with very little transcriptional read-through. One of the characteristics of the EcoRQ111 enzyme is that it's binding onto the DNA can easily be destabilised by the presence of high salt concentration (1 M NaCl). This property allowed us to "wash" the roadblock away. According to the DNA sequence, after roadblock removal three ATPs can be incorporated in the growing mRNA chain. In the CTT, GTP is present because it is the source of energy for EF-G driven ribosome translocation. Therefore, as a test, after removal of the EcoRQ111 enzyme by high salt washing, GTP and ATP were added into the transcription reaction. As can be seen in Figure 31B lane 4, RNAP walked by 3 nt without effects caused by

GTP. Due to the possibility of some RNAP pausing before the roadblock, we decided to observe roadblock read-through by RNAP walking instead of chase reaction, so the read-through could only come from the EC

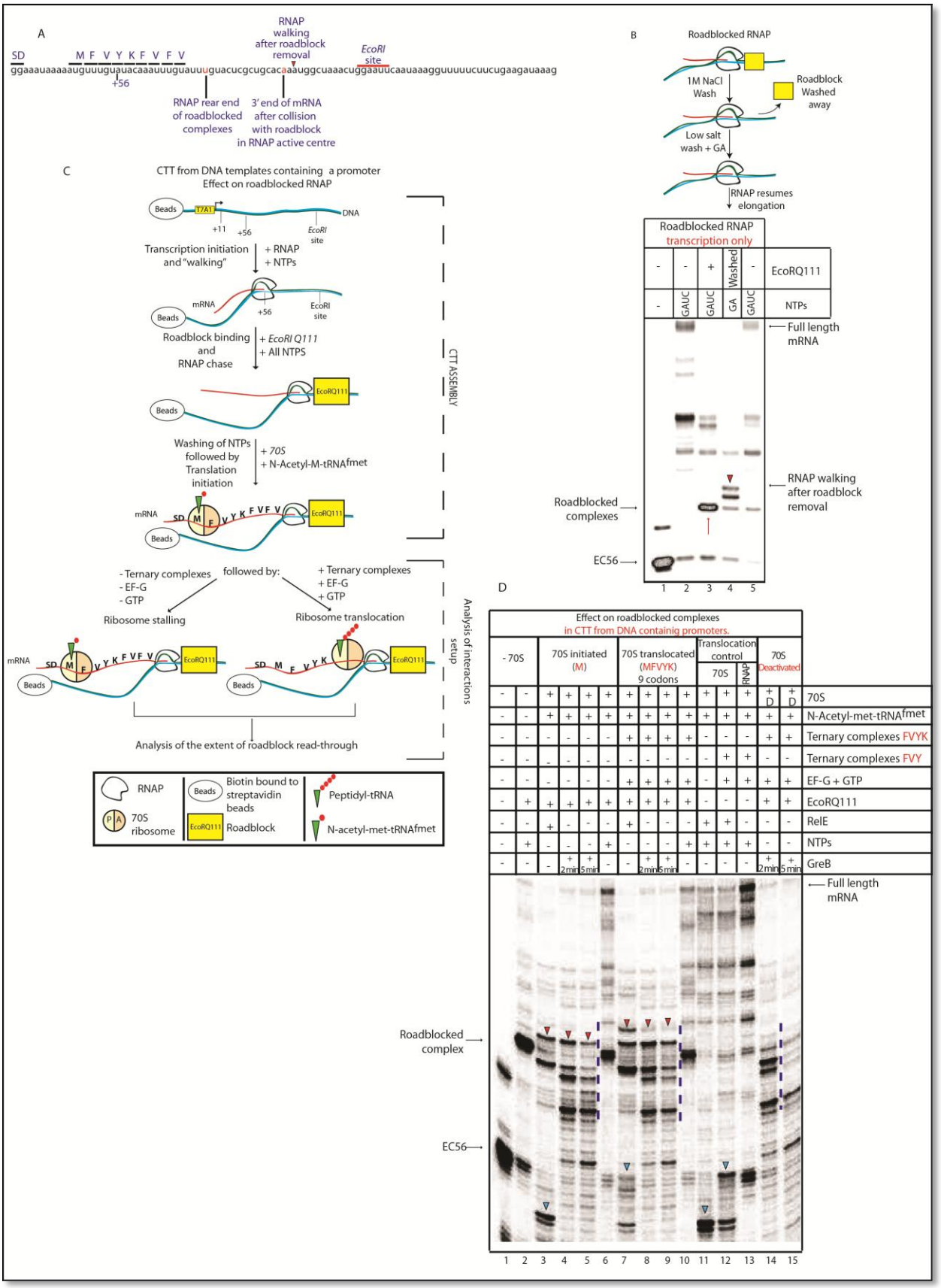

Figure 32 Assessment of the effect on roadblocked transcription elongation complexes by the translating ribosome in promoter borne CTT. A) Sequence of mRNA synthesised from T7A1-Backtracking template. TIR and first nine codons are shown. EcoRI site is shown and transcriptional features are depicted in red, showing the position of the 3' end of the RNA in the RNAP active centre in the roadblocked complex as well as the rear end boundary of the enzyme. Position to where RNAP walks after read-through of the roadblock in the presence of ATP is indicated by a triangle. B) PAGE of radiolabeled mRNA showing the characterisation of the formation of the roadblocked complex in the absence of the translational machinery. Lane 1. Radiolabeled EC56 from which RNAP was either chased (lane 2) or roadblocked and then chased (lane 3). Roadblocked complexes are indicated by a red arrow (note very little read-through above the roadblocked complexes). Lane 4. After roadblock removal (see cartoon depicting formation of roadblocked EC followed by washing off the EcoRQ111 enzyme by high salt washing), RNAP was walked in the presence of GTP and ATP.

Figure 32 legend continuation. Although GTP is not needed for the walking reaction, it is brought into the system by the translational apparatus as energy source for ribosome translocation. RNAP walks by 3 nucleotides incorporating ATP, shown by red triangle. Lane 5. After removal of the roadblock, RNAP is still active and capable of synthesising full length mRNA (note that the radioactivity in lane 5 is lesser than the one in other lanes due to the loss of beads during washing and partitioning of the material) C) Schematic representation of the experimental set-up used to assemble the CTT and the analysis of the interactions between the roadblocked complexes and the translating ribosome (see text). D) PAGE of radiolabeled mRNA obtained in promoter borne CTT. Lanes 1 and 2, formation of EC56 and roadblocked complexes respectively. Lane 3. RelE cleavage of initiated ribosomes. RelE cleavage was obtained indicating that specific translation initiation had taken place (indicated by a green triangle). Lanes 4-5, GreB cleavage of roadblocked complexes in translation initiation. GreB cleavage pattern is indicated by dashed lines. Red arrows indicate read-through of the roadblocked complexes. Lane 6. Transcription elongation control for translation initiation CTT. Lane 7. RelE cleavage of ribosomes translocated by 9 codon (indicated by green triangle) see text. Lanes 8-9 GreB cleavage of roadblocked complexes in 9 codon translocated 70S complexes. GreB cleavage pattern is indicated by dashed lines. Red arrows indicate read-through of the roadblocked complexes. Lane 10 Transcription elongation control for roadblocked complexes in 9 codon translocated 70S CTT. Lanes 11-12 Translation initiation and elongation controls through RelE cleavage of translation initiation complexes and 4 codon translocated complexes respectively in chased elongation complexes with no roadblock added. Lane 13. Transcription elongation control through chase reaction from EC56 in the absence of roadblock. Lanes 14-15. GreB cleavage of roadblocked complexes in the presence of the translational machinery necessary for 9 codon translocation but with heat-deactivated ribosomes. GreB cleavage pattern is indicated with dashed lines.

which given the appropriate NTPs could walk, minimising the possibility of observing chase products from paused RNAPs located before the roadblock.

Having the roadblock set-up in the absence of the translational machinery, we decided to couple it to translation.

In order to assemble the CTT (Figure 32C), RNAP was first waked to EC11 and then walked to EC56 (Figure 32D lane 1) and roadblocked exactly in the same conditions as above. Note that after EC11, RNAP has escaped the promoter therefore it has ejected sigma factor which is washed away during cleansing of the EC11 and EC56 complexes. Once RNAP was roadblocked GTP and ATP were added and no read-through was observed (Figure 32D lane 2). To the roadblocked complexes, 70S ribosomes were added and translation was initiated non-enzymaticaly followed by RelE printing to detect the initiated complexes. As seen in (Figure 32D lane 3 [blue triangle]) RelE cleavage was obtained and the occupancy of the ribosome in the CTT seemed to be ~40%; phosphorolysis of the roadblock band was obtained due to high concentration of phosphate present in RelE preparation needed to keep it soluble. Interestingly, read-through of the roadblock was observed in the translation initiated complexes (Figure 32D lane 3 [red triangle]). Read-through at the 70S initiated complexes was not expected due to the distance between the ribosome and RNAP was 24 nt. For the experiment where the effect on backtracked complexes by the ribosome was observed (chapter 5.9), no outcome was detected for the translation initiation complexes with a distance between RNAP and the ribosome (counting from the initiation codon) of 30 nt. Our proposed ribosome front boundary was 15 nt (from first nucleotide in the initiation

codon AUG to the front edge of the ribosome) due to the data obtained in toeprinting experiments, which pin-points the front edge of the 30S small subunit of the ribosome. This length gives an approximate idea of the space occupied by the 30S subunit of the initiated ribosome bound to the mRNA but does not reflect in the actual distance needed for the initiation process to take place and also for the correct spacing between the ribosome and RNAP. Translation initiation in the roadblocked CTT had a control experiment where ribosomes were initiated and RelE treated in the absence of roadblock in full length mRNA where RNAP could not possibly affect translation initiation because RNAP had reached the end of the DNA template. Ribosome occupancy in such conditions was ~70% proving to be higher than in its roadblocked counterpart (Figure 32D lane 11 [green triangle]). The observed read-through was not detected in the control experiment where heat deactivated ribosomes mixed with all the components of the translation machinery needed for ribosome initiation and translocation by 9 codons were challenged with the roadblocked EC (compare figure 32D lanes 4, 8 with lane 14). These observations might suggest that read-through of the EcoRQ111 roadblocked EC could be due to the ribosome aiding the blocked RNAP. These observations could also suggest that the optimal distance between RNAP and the ribosome for optimal initiation of translation could be greater than 24 nt. It is unlikely to attribute this observation to disassociation of the roadblock from the DNA. EcoRQ111 enzyme is known to bind tightly to DNA and can only be dislodged by high salt washings (1 M NaCl). However, as can be seen in (Figure 32 lanes 6 and 10) after translation initiation and elongation, RNAP was allowed to elongate by addition of all 4 NTP's. Those transcriptional complexes which read-through the roadblock (in a ribosome assisted manner) formed full length mRNA whereas those complexes which were below the roadblock were roadblocked indicating that the roadblock was still bound in some of the complexes suggesting that it did not dissociate from the DNA.

The roadblocked CTT was also treated with GreB as a way to determine the effect of the translational machinery on the RNAP which backtracked as a consequence of colliding with the roadblock. For the 70S initiated complexes, GreB treatment generated a cleavage pattern where some EC seemed to be GreB resistant (Figure 32D lanes 4-5 [dashed lines]) when compared to the pattern obtained for the deactivated ribosomes control (compare figure 32D lanes 5 with lanes 15). Interestingly, some of the read-through complexes were GreB resistant even after 5 min incubation, suggesting that these complexes were stabilised in the post-translocated state (Figure 32D lanes 4-5 [red

arrows]). To evaluate if the resistance of the read-through complexes was not due to the inactivation of RNAP, the complexes were chased by addition of all NTPs. As can be observed in (Figure 32D lane 6), the read-through complexes were capable to continue elongation until full length mRNA was obtained; note that complexes below the road block are roadblocked again after chase.

In the case of the 9 codon translocated 70S complexes, ribosome localisation mapped by RelE cleavage was detected to be in the position of ribosomes translocated only by 3 codons (Figure 32D lane 7 [green triangle]) as compared to the control experiment where the ribosome was allowed to translocate by 3 codons in the absence of roadblock on full length mRNA (Figure 32D lane 12 [green triangle]). This could be due to translational pausing after ribosome translocation by 3 codons or that the 9^{th} codon was never reached by the ribosome, a situation where the A site of the 70S would remain occupied by an aminoacyl-tRNA forbidding RelE to enter and excise the mRNA. Notably, the read-through of the roadblock was observed to a greater extent compared to the read-through obtained for initiation complexes (Figure 32D lane 7 [red triangle]) similar to when EcoRQ111 is removed (by high salt washings) and walked by addition of ATP (for comparison observe Figure 32B lane 4). GreB cleavage pattern also showed GreB resistant ECs (Figure 32D lanes 8-9 [dashed lines]), specially the read-through complex located 1 nt downstream the roadblock became even more resistant than in the initiated 70S counterpart, compare (Figure 32D lane 9 with lane 5 and 15). The read-through complex obtained in the 9 codon translocation set-up (3 nt downstream the roadblock) was GreB sensitive in comparison with the read-through located 1 nt downstream the roadblock. It is important to underline that this experiment was designed to have only 1 molecule of RNAP being roadblocked having only 1 70S ribosome molecule actively translating behind it; this was achieved by removing free RNAP through washing EC11 before synthesis of EC56 and coupling it to translation (avoiding transcription reinitiation); also the distance of the translocated ribosome does not provide enough space for another ribosome to bind and reinitiate translation. It could be expected that a greater outcome would result when more molecules are interplaying amongst them suggesting a more complex cross-talk regulation.

5.8 Discussion

Throughout the process of assembling an *in vitro* transcription system coupled to translation from DNA templates containing promoter, we identified specific factors

required for the validation of the interactions amongst translation and transcription. Based on observations made during the setting up of the translational system such as the measured distances by toeprinting experiments and also on previous experiences regarding mRNA secondary structures along with published data (Komissarova and Kashlev, 1997, Nudler et al., 2003, Wang et al., 2009, Korzheva et al., 2000) which provided the boundaries of RNAP, we designed a DNA template where transcription and translation could coexist. Since our initial goal was to keep the CTT as closed to the *in vivo* counterpart as possible, we used naturally occurring promoters and transcriptional signals (such as the backtracking sequence).

A key question which had to be answered was if the ribosome was being initiated and translocated on the nascent mRNA in the CTT. Although we were certain of the activity of our translational system tested on purified mRNA in the absence of transcriptional factors and additional NTPs (besides GTP), we could not take for granted that the ribosome would behave in the same way once coupled to transcription. Our first approach, the use of RNase H probing, gave us the first indication that indeed the ribosome was blocking the TIR from oligonucleotide binding, suggesting ribosome binding, but we could not withdraw any further information about the transcriptional state i.e. initiated or translocated. Therefore, we looked for a direct way to probe the ribosome. The interferase RelE shone light into our CTT. It allowed us to directly assess translational states of the ribosome and also provided us with a measurable tool to determine ribosome occupancy on the mRNA. The downside of RelE is that it doesn't cleave all codons with the same efficiency. The codon preference of RelE is particularly troublesome when the identification of ribosomal translocated complexes is needed i.e. when in the A site a low cleavage efficiency codon is present. Although the cleavage is still detectable (indicating that the ribosome has undergone translocation), we cannot quantify the exact number of ribosomes which have entered elongation by looking at the translocated complexes, we can only get a qualitative observation. Nonetheless, throughout RelE cleavage we observed activity of both translation and transcription simultaneously, co-existing in our CTT.

During the set-up of our promoter borne CTT, we observed the effect on different transcriptional states by the ribosome. Although the main focus of this thesis is not the elucidation of the outcome of translation on transcription, but the design of the CTTs constructed for such task, the best way to analyse and propose the usefulness of the system is through the description of punctual examples where the CTT is used for the study of the interaction amongst the two machineries. In this chapter, two of such test-

experiments, were described. The effects observed for both transcriptional backtracking and roadblocked ECs, was one of the first demonstrations of the possible capacity of the system to try to unveil the intricate cross-talk amongst the ribosome and RNAP.

In terms of the observed outcomes on backtracked EC in the promoter borne CTT, a preliminary idea of a possible positive aid was obtained. We observed that upon interaction of the ribosome with backtracked RNAP (due to the collision with the roadblock), some EC became resistant to GreB cleavage compared to the scenario where heat-deactivated ribosomes where used. This result could suggest that the ribosome could act as a physical barrier which stops RNAP backtracking, stabilising the EC in the post-translocated state (which is resistant to GreB). The observed effect in backtracking caused by a destabilising sequence in the DNA (backtracking signal) and in backtracking-induced (as a co-effect of RNAP colliding with the roadblock) ECs, was the impediment of continuous backtracking but a "pushing" of RNAP was not observed. In despite of this observation being preliminary but reproducible, it contradicts, at least in part, a mechanism proposed in the laboratory of E. Nudler, which suggests that, *in vivo*, the ribosome is capable of physically "push" backtracked complexes forward reducing the impact caused by stalled RNAP molecules on DNA replication (Dutta et al., 2011) and also the "pushing" effect was observed on RNAP which backtracked as a side-effect of colliding with the lac repressor (Proshkin et al., 2010). It is possible that the "pushing" effect observed *in vivo* could be driven not only by the action of the ribosome over the backtracked complexes but in concert with an auxiliary factor(s) not detected in their experimental approach. Our CTT is based on the minimal factors needed for both transcription and translation to function, so any detected outcome can be regarded only to the direct interaction of the ribosome and RNAP. However, the CTT is open for the addition of transcriptional and translational factors such as GreB and the EcoRQ111 roadblock, generating a plethora of scenarios to analyse the effects on transcription and translation.

The experiment regarding roadblocked EC assisted us in the understanding of the permissive distances between RNAP and the ribosome, suggesting that the distance between RNAP and the ribosome for optimal translation initiation could be greater than 24 nt (from initiation codon to the rear end of RNAP). It also demonstrated that promoter borne CTT is not an optimal experimental set-up due to the background present in the experiment made difficult the assignment of the corresponding band(s) with the obtained effect, interfering with the analysis and comprehension of the outcomes of the collision. Although this preliminary result showed in a reproducible

manner, a different extent of read-through of the roadblock assisted by both initiated and translocated ribosome, and an effect on roadblock-induced backtracking, more experiments are required to elucidate the mechanism of ribosome mediated roadblocked RNAP rescue.

The promoter borne CTT, besides confirming that both transcription and translation were capable to co-exist in this CTT system, also proved to be inefficient for the following reasons: (i) For the experiments described here, RNAP had to be walked by 6 steps. Walking step required at least 5 washings each, so in order to be able to walk RNAP to EC83 (for the study of the effect on backtracked complexes) a total of 30 washings were done. If the experimental needs required us to position RNAP in a remote position, the number of walking steps would increase accordingly. Besides being effort and time consuming, transcriptional walking may generate a lot of background noise, such as undesirable pausing of RNAP, obscuring the analysis of interactions; (ii) besides the fact that the yield of mRNA generated in transcription from promoters is low, in every washing of the RNAP walking reaction the loss of material is unavoidable. Low mRNA concentration used to program the ribosome results in a low yield of peptide synthesis. This is a limiting factor if a proposed experiment requires the monitoring of the synthesised peptides. For these reasons afore mentioned, we decided to look for alternative experimental approaches that would allow us to overcome these obstacles.

Promoter borne CTT allowed us to develop the foundations needed for the design of a more efficient CTT, such as basic measurement of distances between the two machineries and also the fine-tuning of a method for the direct localisation of the ribosome within the coupled system. This first approach led us to a deeper understanding of the requirements for the construction of a system aimed to the investigation of transcription-translation coupling, making us aware of how important the diminishing of the background noise is in order to be able to identify the outcome of collisions. Finally, promoter borne CTT also made us realise that our general approach was aimed to monitor only the effects on transcription disregarding the possibility that the translation machinery could also be affected by the interaction with RNAP, for that reason, the scope of our study broadens involving not only the outcome on transcription but also on translation.

Results chapter 6. Setting up a transcription-translation system with artificially assembled transcription elongation complexes.

6.1 Introduction

The promoter borne CTT was a useful approach to identify the specific requirements needed to develop a system to analyse the interactions between transcription and translation. Therefore, in this chapter the strategies taken for the assembly of two new CTTs will be described. The first system to be discussed on this chapter was designed for the analysis of the outcome of the collision amongst the ribosome and RNAP on translation by looking at the synthesised peptides; the second CTT was designed to overcome all the flaws identified in promoter borne CTT and it is aimed to the analysis of the outcome on transcription upon the collision of RNAP and the ribosome.

6.2 Artificially assembled elongation complex formation

We explored a different approach to couple translation to transcription by using artificially assembled transcription elongation complexes (AAEC). AAEC is usually formed with synthetic RNA oligonucleotides but we used mRNA synthesised *in vitro* using T7 RNAP and then purified by PAGE in exactly the same way mRNA was obtained for the characterisation of the translational system (chapter 4.3.1 and materials and methods). To assemble the AAEC an oligonucleotide containing a sequence complementary to the 3' end of the mRNA surrounded by not complementary (to the mRNA) short and long overhangs (at its 3' and 5' ends respectively) is mixed along with the mRNA forming an RNA:DNA hybrid (Figure 33A). This oligonucleotide serves as "template DNA", because the sequence immediately downstream the hybrid (the long overhang) serves as template which dictates the nucleotides which RNAP incorporates in the growing mRNA chain. Upon hybridisation, RNAP is added into the reaction and the RNA-DNA hybrid is recognized and bound by RNAP positioning the 3' end of the mRNA in the active centre. To complete the assembly of the AAEC, a second DNA oligonucleotide (non-template DNA), fully complementary to the template DNA, is allowed to anneal with the overhanging stretches of the template DNA and it is then accommodated into the complex resulting in the formation of fully active elongation complex. AAEC are indistinguishable from transcription elongation

complexes formed on double-stranded DNA of the same sequence by transcription from a promoter (Sidorenkov et al., 1998, Yuzenkova and Zenkin, 2010b, Daube and von Hippel, 1992).

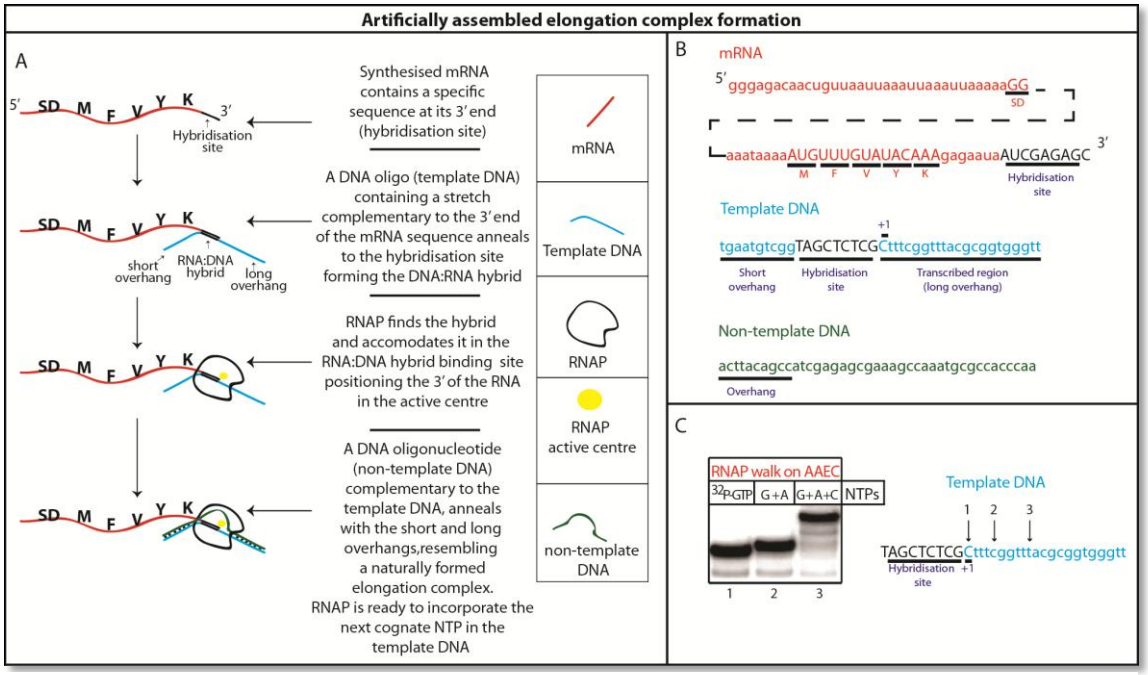

Figure 33. Artificially assembled elongation complex formation. A) Schematic representation of the assembly of artificial transcription elongation complex (see text). mRNA (red) with TIR, a coding sequence and a specific sequence (black line) are shown. Template DNA (blue line) with a complementary sequence to the 3' end of mRNA (black line) is shown along with non-template DNA (green line). RNAP (black circle) with its active centre (yellow circle) is also shown. B) Sequence of synthesised mRNA (mRNA-AAEC1) used for AAEC formation is shown (red) depicting translational features and the hybridisation site. The sequence of the hybridisation site corresponds to the first 8 nt of the *E. coli* T7A1 gene transcribed region plus an additional CTP for a total length of 9 nt. The template DNA (blue) is annotated with the overhangs and hybridisation site. Note that +1 represents the first NTP to be incorporated upon AAEC formation. Non-template DNA (green) complementary to the template DNA is also shown. C) PAGE of AAEC mRNA radiolabeled by RNAP. Upon AAEC formation, His-tagged-RNAP was immobilised to Ni^{+2} – agarose beads and washed with transcription buffer. Lane 1. RNAP incorporates α-[^{32}P]-GTP labelling the mRNA. RNAP was then walked by 3 (lane 2) and 8 nt (lane 3). The template DNA sequence shows the positions to which RNAP was walked to (arrows). Note that RNAP used for experiments where AAEC were used did not contain σ factor. However the presence of σ factor has no effect on the assembly, formation or behaviour of the AAEC.

This technique has been widely used for the investigation of RNAP properties during elongation (Roghanian et al., 2011, Yuzenkova and Zenkin, 2010b, Kireeva and Kashlev, 2009b, Kassavetis et al., 2010, Grunberg et al., 2010).

In order to form the AAEC, an mRNA (referred to as mRNA-AAEC1) containing the same TIR and open reading frame used for the characterisation of the translational system was synthesised (Figure 33B). The hybridisation sequence on the mRNA 3' end corresponds to the first 8 nt of the transcribed region of the T7A1 gene plus an additional CTP, making the hybrid of a total length of 9 nt, the length of the DNA:RNA

hybrid in a EC. mRNA-AAEC1 was incubated with template DNA in the presence of 6xHis-tagged RNAP, followed by the addition of non-template DNA. 6xHis-tagged RNAP was immobilised on Ni^{+2}-agarose beads and washed with high salt and low salt to remove the excess of mRNA, template, non-template DNA and possible unspecific AAEC. Being GTP the first nucleotide to be incorporated according to the template-DNA sequence (Figure 32B), α-[^{32}P]-GTP was added to the AAEC. As can be seen in (Figure 32C lane 1), RNAP in the AAEC was capable to incorporate the radiolabeled NTP into the mRNA, indicating that the AAEC was active. To further test the AAEC, RNAP was walked by 3 and 9 nt (Figure 33 lanes 2-3). Through this experiment we observed that the formed AAEC was capable to withstand high salt washings and also that RNAP could be walked to any desired position demonstrating its resemblance to EC formed from promoters. The particular usefulness of working with AAEC is that, using the same mRNA, many EC states (backtracked, pre-translocated, paused, etc) can be studied by simply exchanging the template and non-template DNA to sequences suiting the experimental needs. AAEC also eliminates the need of walking RNAP through long distances diminishing the background noise and facilitating the execution of the experiment. Since the mRNA is added into the system, AAECs can be easily scaled-up, overcoming the limiting mRNA concentration observed in promoter borne CTT, providing enough material to be capable of observing the synthesised peptides.

6.3 Assessment of ribosome occupancy in AAEC by RelE cleavage.

Firstly, we aimed to develop a system for monitoring the effect on translation upon its interaction with transcription by looking at the synthesised peptides. The first step consisted in the analysis of the occupancy of the ribosome in AAEC; to form the CTT, mRNA-AAEC1 was used and the AAEC was assembled exactly as it was described for the characterisation setup (chapter 6.2) but using TAP tagged RNAP (instead of 6xHis=tagged RNAP) immobilised on immunoglobulin-G-sepharose (Ig-G) beads because the purified translation factors and RelE contained a hexa-histidine tag which would be bound to the Ni^{+2}-agarose beads disabling the system. Upon AAEC formation, the mRNA was labelled on its proximal 3' end by addition of α-[^{32}P]-GTP followed by addition of ATP and CTP, walking RNAP by 9 nucleotides forming a distance of 27 nt amongst the first nt in the initiation codon AUG and the rear-end of RNAP (Figure 34A). The system was washed ensuring that only mRNA forming AAEC is present in the reaction so, once translation is initiated, all the translating ribosomes are coupled to

transcription. After analysing the previous CTT, we opted for having the translational system closer to its *in vivo* environment; therefore, in order to initiate translation, we decided to use formylated met-tRNA^{fmet} (F-met-tRNA^{fmet}), the naturally occurring form

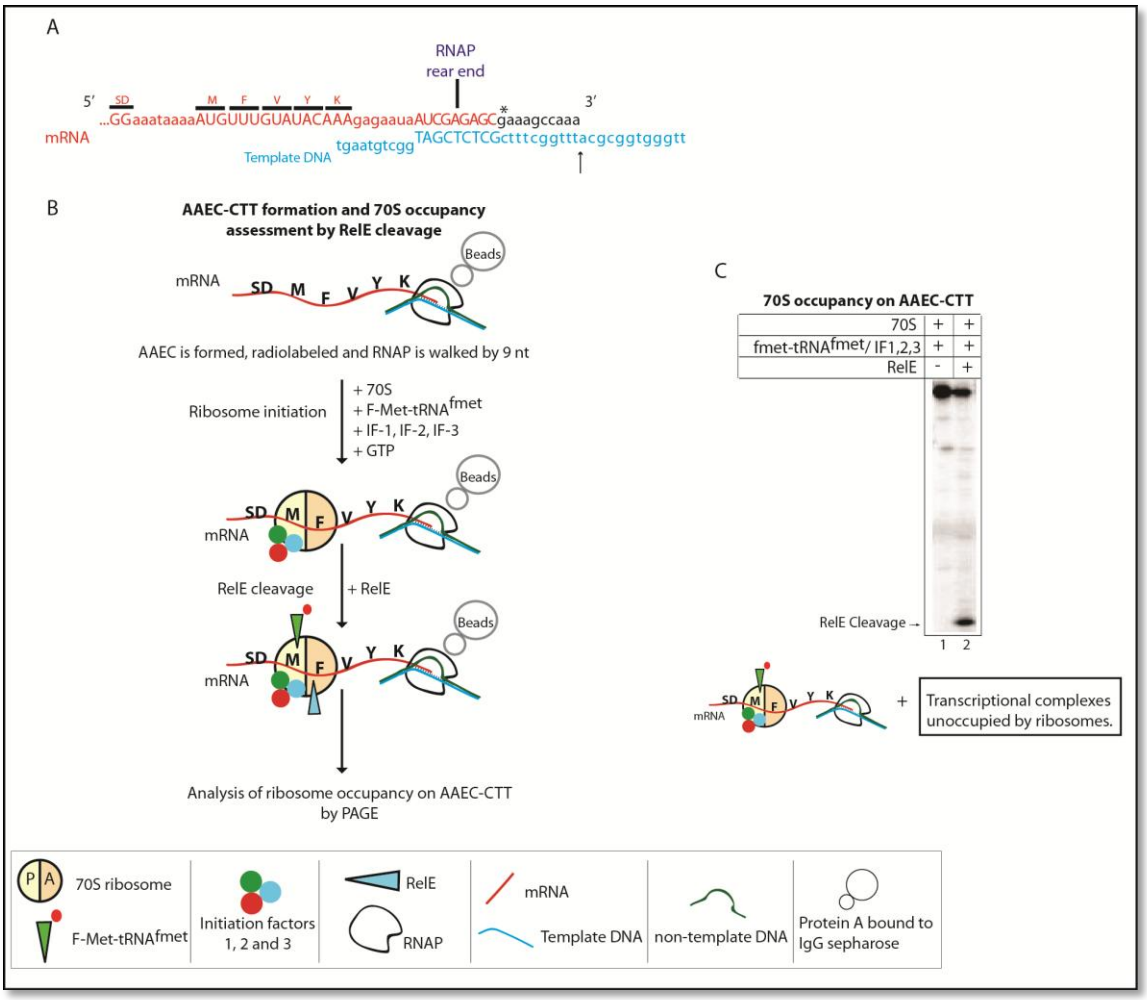

Figure 34. Assessment of ribosome occupancy on AAEC-CTT. A) Sequence of AAEC formed with mRNA-AAEC1 (red), showing TIR and open reading frame. Nucleotides incorporated by RNAP walk by 9 nt are shown in black. Asterisk represents the radiolabeled NTP (GTP). RNAP rear end boundary is depicted. Template DNA (blue) is shown. The arrow indicates the position to where RNAP was walked. Note that even though non-template was used in the experiment it is not shown for clarity purposes. B) Schematic representation of AAEC-CTT assembly and assessment by RelE. C) PAGE of mRNA obtained from AAEC-CTT. Lane 1 mRNA labelled by RNAP during transcriptional walking. Lane 2) Assessment of occupancy by RelE cleavage demonstrated a low occupancy but also indicated that all translational complexes contained a AAEC.

of the initiator tRNA^{fmet} and drive the initiation pathway through the initiation factors IF-1, IF-2 and IF-3 (enzymatic initiation), instead of initiating through the peptidyl analogue N-Acetyl-met-tRNA^{fmet} (non-enzymatic initiation) procedure. After enzymatic initiation (Figure 34B), the 70S occupancy on the AAEC-CTT was assessed through RelE cleavage. As seen in (Figure 34C lane 2), RelE cleavage was ~30%. Even though we tried to increase the occupancy by adding greater amounts of ribosomes, initiator tRNA and initiation factors, no changes were observed. This phenomenon could have

been probably caused by wrapping of some mRNA molecules around RNAP (Rivetti et al., 2003) limiting ribosome binding and initiation. Having a CTT with such low occupancy meant that experiments designed to look at the effects on transcriptional complexes were not possible because of the mixed population present in the assay would compromise the analysis of the resulting interactions.

To be able of analysing the effects on translation we had to be capable of following peptide synthesis. Therefore, we explored different techniques which would allow us to monitor the peptidyl transferase activity of the ribosome.

6.4 Exploring peptidyl transferase activity by thin layer chromatography techniques.

The first step towards the observation of the synthesised peptides was the aminoacylation of $tRNA^{fmet}$ with [^{35}S]-Methionine with its concomitant formylation to obtain [^{35}S]-F-met-$tRNA^{fmet}$, generating a radiolabeled substrate which would allow us to visualise the peptidyl transferase reaction (PTR). Since our translational system was designed to take small steps towards RNAP, the peptides which would be observed would be in the range, in amino acid composition, from dipeptides to hexapeptides, excluding the use of SDS-PAGE because of the low resolution that it provides. Thin layer chromatography (TLC) has been routinely used as a tool for the identification and characterisation of the primary structure of polypeptides. The technique is based on the affinity of a given compound (in this case amino acids) to the stationary and mobile phases; the affinity determines the extent of migration on the stationary phase, allowing the separation of compounds. For the resolution of peptides, the stationary phase is usually composed of cellulose or silica-gel, layered on an inert backing such as plastic or aluminium. The mobile phase is composed of organic solvents mixed in different ratios depending on the sample which is to be analysed.

As a test, we performed PTR in the absence of the transcriptional machinery; the experiment was executed using mRNA-AAEC1 (Figure 35A). 70S ribosomes were added and initiated enzymaticaly in the presence of [^{35}S]-F-met-$tRNA^{fmet}$ for the radiolabeling of the resulting peptides. Step-wise translocation of the ribosome was performed by addition of EF-G, GTP and ternary complexes (EF-Tu·GTP·aminoacyl-tRNA) which permitted the synthesis of mono, di, tri, tetra and pentapeptides (Figure 35B). Aliquots of the resulting peptides from each translocation step were deacylated

through alkaline treatment (100 mM KOH) which hydrolyses the ester bond which links the peptide (or amino acid in the case of monopeptide) to the 3' terminal adenosine of

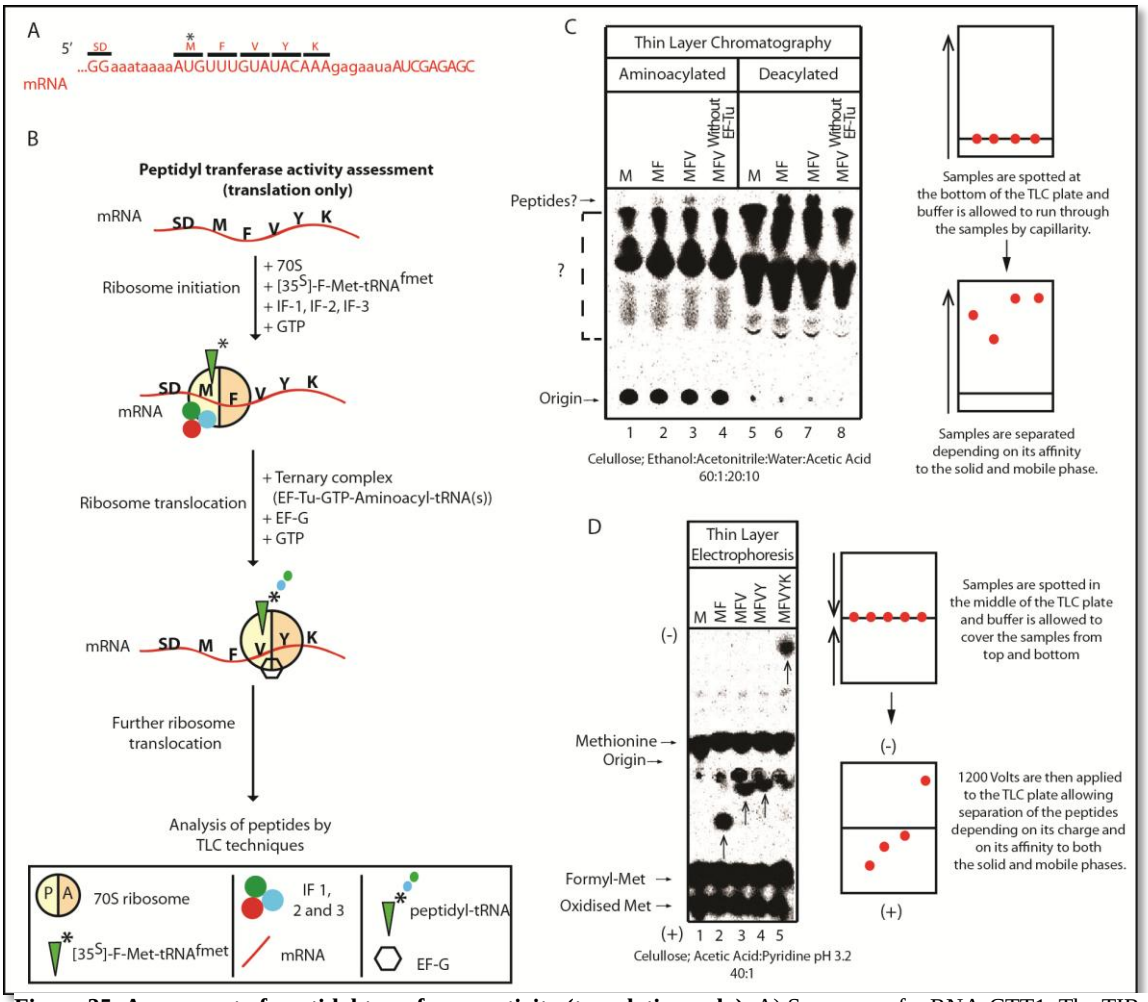

Figure 35. Assessment of peptidyl transferase activity (translation only). A) Sequence of mRNA-CTT1. The TIR and open reading frame are shown. The asterisk represents the radiolabeled amino acid residue in the synthesised peptides. B) Schematic representation of the PTR performed for obtaining the radiolabeled peptides. Note that only two ribosome steps are shown. C) TLC of the obtained peptides (see text). Lanes 1-4 aminoacylated samples were spotted. Lanes 5-8 deacylated samples were spotted. D) Thin Layer Electrophoresis (TLE) of the PTR synthesised peptides. Arrows on gel indicate the spots corresponding to the peptide. (-) and (+) denote the anode and cathode.

the tRNA (Walker and Fredrick, 2008). Aminoacylated and deaminoacylated samples (corresponding to mono (M), di (MF) and tripeptide (MFV)) were spotted on an aluminium backed cellulose TLC plate as solid phase and a mixture of ethanol:acetonitrile:water:acetic acid in a 60:1:20:10 ratio as mobile phase (N. Zenkin personal communications) were used. As for the aminoacylated samples, some spots migrated along with the mobile phase, suggesting spontaneous deacylation of the obtained peptidyl-tRNA and [^{35}S]-Met-tRNA^{fmet} (which co-purifies with [^{35}S]-F-Met-tRNA^{fmet}); the deacylation product of [^{35}S]-Met-tRNA^{fmet} ([^{35}S]-Met) undergoes oxidation forming sulfoxide which also migrates on the TLC plate (Walker and Fredrick, 2008) (Figure 35C lanes 1-4). However, radioactive spots were also detected

at the origin suggesting that the synthesised peptides could still be bound to the tRNA, which does not allow peptide migration on the TLC plate. Looking at the deacylated samples, we observed disappearance of the origin spots and appearance of a new spot in the samples corresponding to dipeptide (MF) and tripeptide (MFV) synthesis (Figure 35B lanes 6-7). Interestingly, this spot was not detected in samples containing only initiating 70S complexes (M) nor it was observed in the control situation where the tripeptide MFV was synthesised in the absence of the elongation factor EF-Tu. These observations could suggest that PTR had taken place and also indicated that TLC under these conditions could not resolve the synthesised peptides.

A TLC-based technique known as Thin Layer Electrophoresis (TLE) takes advantage of the affinity of the samples to the solid and mobile phases and also separates them according to the charge of the peptides. This technique has been used for the elucidation of the mechanism of translation fidelity (Zaher and Green, 2009); on a plastic backed, cellulose TLE plate, the samples are spotted in the middle and a buffer (mobile phase) is allowed to run over the samples from the top and the bottom of the plate. The plate is then placed into an electrophoresis chamber and voltage is applied. PTR was done in the absence of the transcriptional machinery using mRNA-AAEC1. 70S ribosomes were added and initiated enzymaticaly in the presence of [^{35}S]-F-met-tRNA^{fmet} and step-wise translocation of the ribosome was done by addition of EF-G, GTP and ternary complexes (EF-Tu·GTP·aminoacyl-tRNA), synthesising peptides corresponding to M, MF, MFV, MFVY and MFVYK. After alkaline treatment, the samples were spotted in a cellulose TLC plate and TLE was performed. As shown in (Figure 35D) separation of the synthesised peptides was observed. With the help of Dr. Hani Zaher from the laboratory of Dr. Rachel Green from the John Hopkins Institute, the corresponding spots (Figure 35D) were identified and confirmed to be the product of our step-wise ribosomal PTR.

6.5. Setting-up a "Transcription first" CTT (TR-CTT) for analysis of coupling effects on translation.

Being capable of visualising the synthesised peptides in the absence of the transcriptional apparatus through TLE, we took the next step which consisted on the development of a system where the effects on the PTR upon collisions with RNAP could be monitored. To do so, AAEC was formed using mRNA-AAEC1 and the same

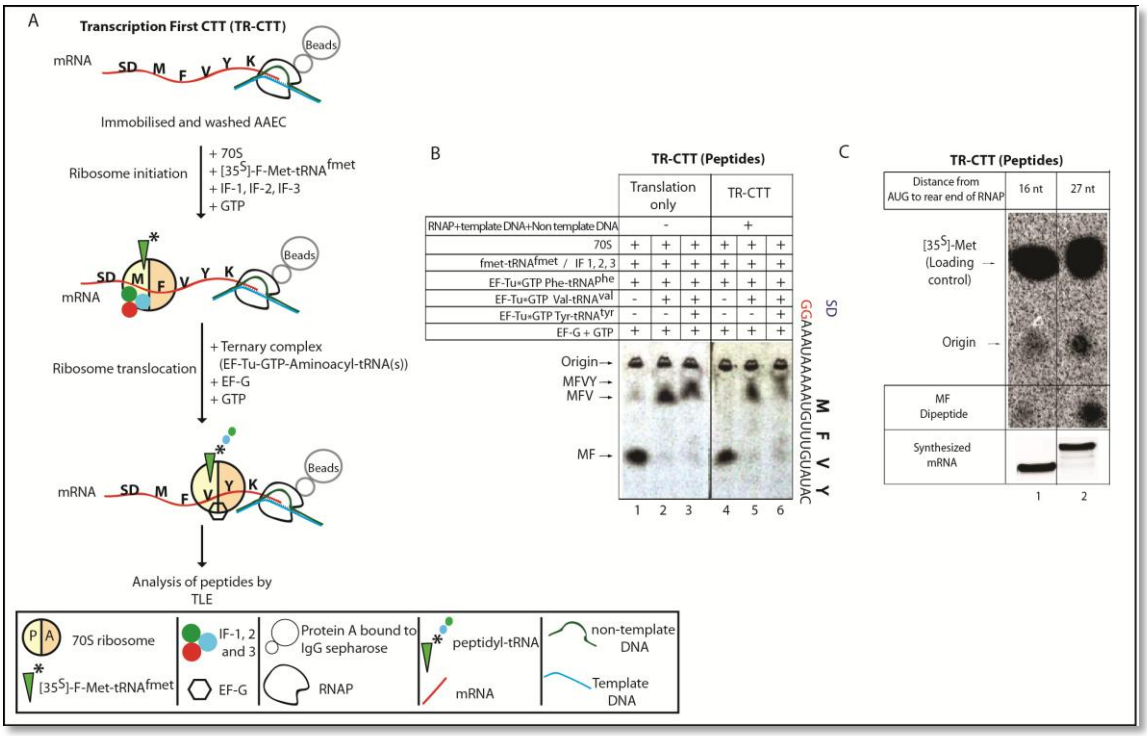

Figure 36. Characterization of "transcription first" TR-CTT. A) Schematic representation of assembly of TR-CTT. Solid phase immobilized AAEC is washed to remove unincorporated mRNA, after that translation is initiated and allowed to elongate on the mRNA of the AAEC. B) Peptidyl transferase activity in TR-CTT. Ribosome elongation by F, FV or FVY codons is followed by thin layer electrophoresis (TLE) of synthesized peptides labelled with [35^S]-F-Met-$tRNA^{fmet}$. Comparison of the peptidyl transferase activity of the ribosome on naked (without AAEC) mRNA (lanes 1-3) and in the TR-CTT (lanes 4-6). C) Rough estimation of mRNA length behind transcribing RNAP sufficient for translation initiation as an example of the usefulness of TR-CTT. MF dipeptide formation was used as a measure of translation initiation efficiency (given that it does not require translocation) on mRNAs containing 16 and 27 nucleotide spacers between the AUG start codon and the rear end of RNAP.

template and non-template DNAs used for the characterisation of AAEC assembly were used. In order to remove template DNA, non-template DNA and mRNA which was not used in the formation of the AAEC, we used TAP-tagged RNAP immobilised on IgG-sepharose beads; even though biotinylated template or non-template DNA can also be used for the immobilisation of AAEC, TAP-tagged RNAP was used because a greater amount of AAECs remained bound to the solid phase after washings compared to biotin-streptavidin immobilised complexes. After AAEC formation and immobilisation,

RNAP was walked by 9 nt by addition of GTP, ATP and CTP (with no radiolabeling of the mRNA) generating a distance of 27 nt between the initiation codon AUG and the rear end of RNAP. The system was then washed with transcription buffer to remove unused template, non-template DNA, NTPs and unused mRNA (Figure 36A) excluding the possibility of monitoring peptides synthesised by translational complexes formed on AAEC-free mRNA. After the washing procedure, 70S ribosomes were added and translation was initiated enzymaticaly in presence of [^{35}S]-F-met-tRNA^{fmet}. Due to the order of events (formation of transcriptional complex followed by coupling with translation), we named the system "transcription first"-CTT (TR-CTT)

As a test, PTR was executed in a step-wise manner allowing the formation of MF, MFV and MFVY peptides. The samples were deacylated and analysed on TLE. As control PTR was also performed in parallel on naked mRNA (without AAEC). Peptides corresponding to dipeptide (MF), tripeptide (MFV) and tetrapeptide (MFVY) were observed in the absence of the transcription machinery (Figure 36B lanes 1-3) and in the TR-CTT (Figure 36 lanes 4-6), suggesting that the efficiency of peptide synthesis in the TR-CTT was similar to that observed in the uncoupled translation system in the absence of RNAP.

We tested if TR-CTT could be used to determine the distance between RNAP and the ribosome that would allow efficient translation initiation. In (Figure 36B), RNAP was far enough from the AUG codon (27 nucleotides between AUG and the rear end of RNAP) to allow translation initiation as efficient as on the naked mRNA. Therefore, we compared the effect on translation initiation on TR-CTT having 16 and 27 nt between AUG codon and the rear end of RNAP. As can be seen from (Figure 36C), the 16 nt distance was already too short for efficient translation initiation (compare lanes 1 and 2) judged by the synthesis of the dipeptide MF (given that ribosome translocation is not needed for dipeptide synthesis). The poor translation initiation observed in this experiment cannot be attributed to a difference on mRNA concentration used to program the ribosomes (Figure 36C mRNA panel) suggesting a direct physical interference of RNAP on ribosome binding and initiation. The TR-CTT will be used to determine precise distances in both initiation and elongation.

6.6 Setting-up a "Translation first" CTT (TL-CTT) for analysis of coupling effects on transcription.

In order to be able to use AAEC for the analysis of the outcome on transcription by the translating ribosome, we needed to achieve complete ribosomal occupancy on the AAEC because the presence of a mixed population of occupied and unoccupied transcriptional complexes would obscure the outcome of the collisions on RNAP. Observation of the low occupancy of the ribosome on the previously mentioned promoter-borne and TR-CTT suggested that it was due to unspecific mRNA interactions with RNAP which blocked the access of the ribosome to bind and initiate. For that reason, we decided to explore an experimental setup where translation was initiated in the absence of the transcriptional machinery (hence "translation first" [TL-CTT]) followed by purification of translation elongation complexes from unused mRNA (see below) and AAEC formation and simultaneous translocation of both ribosome and RNAP to analyse the interactions on transcription. The mRNA template used for this experiment (referred to as mRNA-AAEC2), had 31 nt between AUG and its 3' end (note that the distance here is between the initiation codon and the 3' end of the mRNA; it does not involve the rear end of RNAP because at the moment of the formation of the translational system, RNAP is absent; the distance after TL-CTT formation between AUG and the rear end of RNAP is 28 nt, see below) (Figure 37A). The experiment was executed incubating mRNA-AAEC2 with 70S ribosomes initiating translation enzymaticaly in the presence of [35^S]-F-met-$tRNA^{fmet}$. Initiated complexes were allowed to elongate by one codon (MF dipeptide containing translation elongation complexes). In the translation machinery, there are not tagged proteins to immobilise the system on a solid phase disallowing us from the possibility of washing the system to remove unused mRNA (as done on the TR-CTT). Therefore, the MF dipeptide containing translation elongation complexes were purified by ultracentrifugation through a sucrose cushion (Zaher and Green, 2010) (Figure 37B). During centrifugation, translation elongation complexes are separated from initiation and elongation factors, unused ternary complexes and unused mRNA. The resulting pellet containing only mRNA occupied by translocated ribosomes was washed to remove traces of supernatant and resuspended in translation buffer. This mRNA, fully occupied with elongating ribosomes, was resuspended in translation buffer, incubated with template DNA and RNAP followed by the addition of non-template DNA; After TL-CTT assembly, RNAP was walked by 12 nt by addition of α-[^{32}P]-GTP (labelling the

mRNA at its 3' end thus we observe only translation elongation complexes coupled to transcription elongation complexes), ATP and CTP, leaving a distance of 28 nt between the initiation codon to the rear end of RNAP (Figure 37A). To further confirm the occupancy of mRNA with ribosomes in TL-CTT, we performed RelE probing. As can be seen from (Figure 37C lane 2), RelE cleaved most of mRNA in TL-CTT, indicating that majority of AAECs contained elongating ribosomes making it optimal for the analysis of the effects of coupling on transcription. RelE cleavage after translocation of ribosome by two codons (after addition of EF-Tu·GTP·Val-tRNA^{val}, EF-G and GTP) further revealed that most of these ribosomes were active (Figure 37C, lane 4).

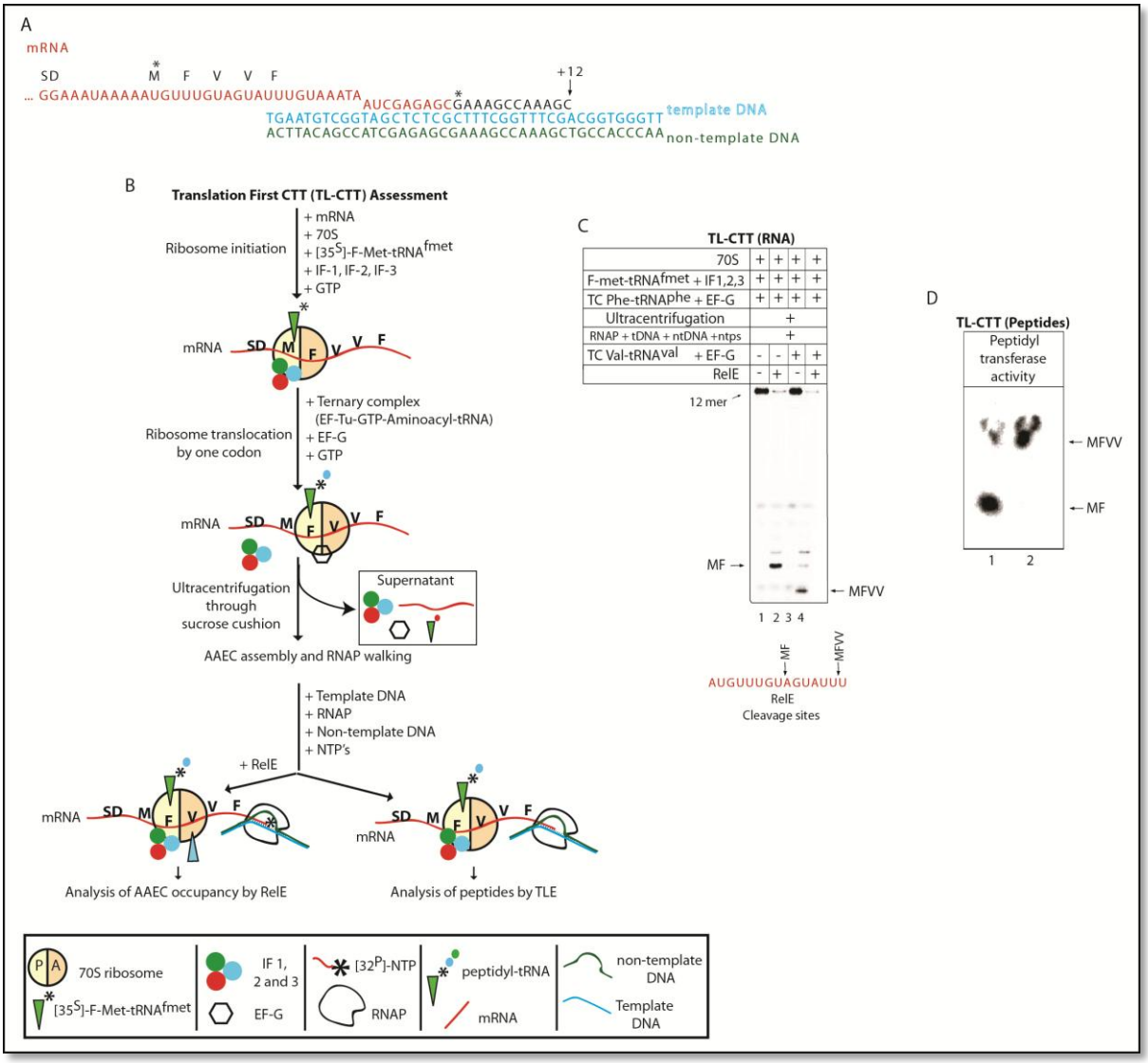

Figure 37. "Translation First" (TL-CTT) system assembly and assessment. A) Sequence of mRNA-AAEC2 based AAEC. mRNA (red) with TIR and open reading frame is shown. Nucleotides incorporated by RNAP are shown in black. The position to where RNAP is walked to is shown (+12 arrow). Asterisks represent radiolabeled nucleotide (GTP) for mRNA labelling (for RelE tested occupancy experiments) and radiolabeled amino acid ([35S]-methionine) for monitoring of PTR. Note that when radiolabeling of mRNA is performed for RelE probing experiments only. Template and non template used for this experiment are shown. B) Schematic representation of assembly of TL-CTT. Translation is initiated with [35S]-F-Met-tRNA^{fmet} (green triangle) and initiation factors 1, 2 and 3 (green, blue and red circles). After translation initiation, ribosomes are allowed to elongate by one codon with synthesis of a dipeptide (MF).

Figure 37 legend continuation The translation elongation complexes are separated from unused factors and unused mRNA by ultracentrifugation through a sucrose cushion. The mRNA carrying elongating ribosome is used in assembly of AAEC. mRNA in TL-CTT is labelled by incorporation of radiolabeled NMP during RNAP walking, ensuring that mRNA is labelled only in coupled complexes. The ribosome can then be probed with RelE or the peptides can be deacylated for the observation of the synthesised peptides (in separate experiments). C) Occupancy of mRNAs with ribosomes in TL-CTT revealed by RelE cleavage. Shown is PAGE of mRNA. Note that only coupled complexes are visible since mRNA is labelled during RNAP transcription after TL-CTT assembly. RelE cleavage was performed in the translation elongation complex containing dipeptide MF (formed prior purification) and after this complex was allowed to elongate by two codons (to tetrapeptide MFVV). The sequence below shows where RelE cleavage takes place. D) PTR assay to analyze the activity of the ribosome after purification and TL-CTT assembly. MF dipeptide (formed prior purification with [35S]-F-Met-tRNA^{fmet}; lane 1) was allowed to be extended to tetrapeptide (MFVV; lane 2). Products were resolved by TLE.

The activity of the ribosomes in TL-CTT was further evaluated by their ability to extend MF dipeptide to tetrapeptide MFVV, i.e. to translocate further by two codons, after addition of EF-Tu·GTP·Val-tRNA^{val} ternary complex, EF-G and GTP. The peptides, radiolabeled at [35^{S}]-F-Met-tRNA^{fmet}, were analyzed by TLE. As seen from (Figure 37 D), the MF dipeptide was readily elongated into the MFVV tetrapeptide (Figure 37D lane 2) demonstrating that the translational system was capable to endure not only the ultracentrifugation through the sucrose cushion but also the incubation times and components needed for the assembly of the AAEC. Note however that, although most of elongating ribosomes in TL-CTT set-up are active, it is possible that some of them may not be coupled to RNAP, making TL-CTT unsuitable for following outcomes of coupling by peptides analysis.

To test the system, we used an AAEC containing a sequence that makes RNAP prone to pause in a hairpin independent manner, (pre-translocated stabilised and backtracked pauses) (Bochkareva et al., 2011) (Figure 38A). The aim of this set-up was to analyze the effect of colliding ribosomes on transcriptional pausing. For the formation of the AAEC, mRNA-AAEC3 was used. mRNA-AAEC3 contained 40 nt between the initiation codon and the 3' end of mRNA and in the absence of RNAP and permitted the synthesis of a decapeptide. The template and non-template DNAs contained a sequence which makes RNAP so pause-prone that, in low NTPs concentrations (5 μM) RNAP hardly reaches the end of template (RO) even after several minutes.

After formation of TL-CTT (Figure 38B), RNAP was walked to the position of first pause (P1) by addition of an incomplete set of NTP's forming a distance between the ribosome and the rear end of RNAP of 29 nt. Then, the ribosome was either left in one codon translocated configuration after synthesis of MF dipeptide (stalled translation elongation complex, left panel in Figure 38C), which is unable to interact with RNAP

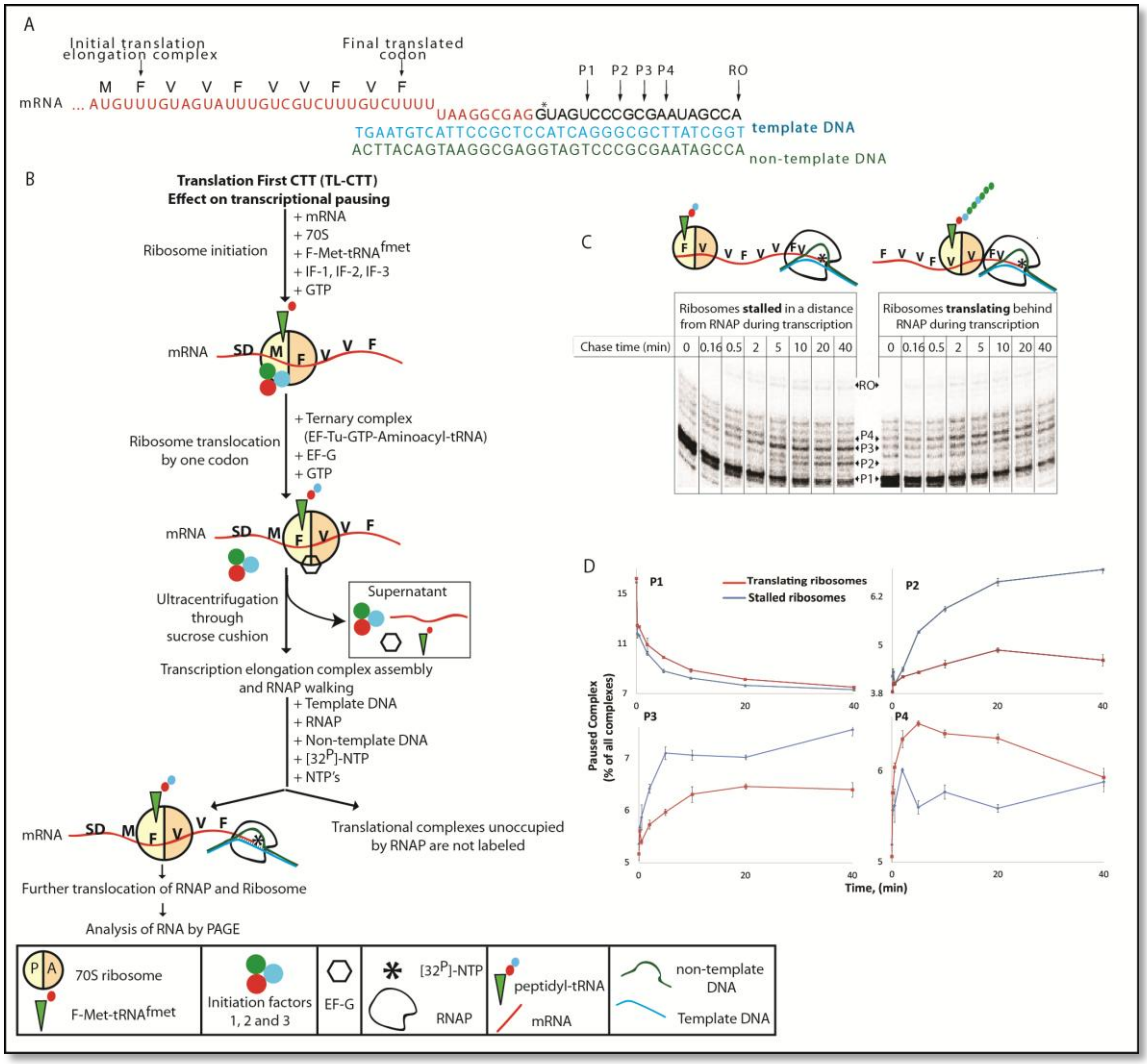

Figure 38. Complex effects of coupling on pausing of transcription as an example of using TL-CTT. A) mRNA-AAEC3 based AAEC. mRNA (red) with TIR and open reading frame is shown. MF initial elongation complex, final translated codon and the positions of the pauses are indicated by arrows, asterisk represents the radiolabeled nucleotide. Incorporated NTPs are shown in black. B) Schematic representation of TL-CTT formation aimed for the study on the effects on transcriptional pausing. C) Shown is PAGE of mRNA. After TL-CTT formation, RNAP was walked to P1 by addition of incomplete set of NTP's. Then, translation elongation complex was either left with dipeptide MF ('stalled' complexes) or was allowed to elongate behind RNAP by addition of F and V ternary complexes in presence of EF-G and GTP for 3 min ('translating' complexes). After this step, RNAP was allowed to transcribe by addition of four NTPs. While the stalled ribosomes remain at their initial position, the translating ribosomes follow transcribing RNAP. Gel shows effects on transcriptional pausing by coupled ribosome. D) Plots show quantification of some pauses as a fraction (in percent) of all complexes in the lane versus time. Error bars are standard deviation from two independent experiments.

due to the distance between the two machineries, or allowed to translocate until colliding with RNAP by addition of Phe-tRNA^{phe} and Val-tRNA^{val} ternary complexes, EF-G and GTP for 3 minutes (translating complex, right panel in Figure 38C). After that, all four NTPs were added allowing RNAP elongation until the end of template. The ribosome stalled in MF configuration remains in this position and does not follow transcribing RNAP (left panel in Figure 38C), whilst the translocated ribosome follows transcribing RNAP until decapeptide is synthesized (right panel in Figure 38C). Note however, that decapeptide synthesis allows the front edge of the ribosome to travel

almost as far as the rear edge of RNAP, which reached the end of the DNA template (the distance from the 10th codon to the rear end of RNAP after reaching the end of the DNA template is 17 nt). This is explained taking into account the observed distances in previous experiments where a distance of 16 nt in between the ribosome and RNAP proved to be limiting for translation initiation and their footprints (15 nt upstream the 3' end of mRNA for RNAP rear-end and 15 nt downstream the first nucleotide in the codon occupying the P site for the ribosome's front edge) suggesting that at this distance the two machineries could be physically interacting. As can be seen in Figure 38D, there was a complex response of transcription pausing to the presence of ribosome translating the mRNA co-transcriptionally (compare "stalled" and "translating" panels of Figure 38D). Read-through of pauses P2 and P3 was observed when the ribosome was allowed to translate behind RNAP. In contrast, no effect caused by coupling was observed for pause P1. Interestingly, P4 was slightly increased in the presence of co-transcriptionally translating ribosome. Given the absence of the secondary structure of the mRNA, the observed effects can be attributed to direct physical interactions between the two machineries. Decrease of pausing at P2 and P3 could be explained by recently proposed pushing of RNAP by the ribosome. However, these effects, as well as the opposite effect on P4 and absence of effect on P1, require further investigation.

6.6.1 Use of TL-CTT system to elucidate the interface between the ribosome and RNAP.

After observing the response caused by the translating ribosome on transcriptional pausing, we decided to investigate if such effects could be affected by modifying the way in which the ribosome approaches RNAP; since ribosome translocates by 3 nt and RNAP translocates by 1 nt, we wondered if altering the "phase" between the colliding ribosome with RNAP through the introduction of 1, 2 and 3 nt in the mRNA between two machineries would have any consequences on the effect on hairpin-independent paused RNAP.

To investigate this, we used mRNA-AAEC3 (Figure 39A top panel) which, in the absence of RNAP, allows translocation of the ribosome by 10 codons; the template DNA (henceforth refer to as t-DNA) used for the assembly of AAEC had signals for pausing transcription elongation through two different mechanisms: pre-translocated stabilised and 1 bp backtracking, permitting the evaluation of effects caused by the translating ribosome on two different pauses simultaneously; the pausing signals were

positioned in tandem (pre-translocated stabilised first followed by 1 bp backtracking) with no spacing in between and have been thoroughly characterised (Bochkareva et al., 2011) (Figure 39A). In order to test if any effects could be observed on these paused EC by action of the colliding ribosome, first TL-CTT ("phase-TL-CTT") was formed on mRNA-AAEC3 by initiating translation enzymaticaly followed by ribosome translocation by one codon. This complex was purified by ultracentrifugation through a sucrose cushion and the resulting pellet (containing one codon translocated translational complexes) was washed and resuspended in translation buffer; RNAP and t-DNA were added followed by incubation with non-template DNA. Once the AAEC was formed, RNAP was walked by three nucleotides by the addition of [^{32}P]-GTP, UTP and ATP, labelling the mRNA at its 3' end and generating a distance between the ribosome and rear end of RNAP of 28 nt. This walking step positioned RNAP right before the signal for pre-translocated stabilised pause (see Figure 39A top panel and Figure 39B); after RNAP walking by the afore mentioned 3 nt, ribosomes were either stalled in the one codon translocated complex or allowed to translocate towards RNAP by addition of Phe-tRNA^{phe} - Val-tRNA^{val} ternary complexes, EF-G and GTP for 3 minutes; after ribosome translocation, all 4 NTPs were added allowing RNAP to resume transcription elongation, pausing at the pre-translocated stabilising and backtracking sites. As for ribosomes stalled in the one codon translocation state, appearance of both pre-translocated and 1bp backtracked pauses was seen through time (Figure 39 C1, lanes 2-8); in the case of ribosomes translating behind RNAP, diminishing of the 1 bp backtracked pause was observed but no apparent effect was seen on the pre-translocated stabilised pause (Figure 39 C1, lanes 9-16). Knowing that in "phase-TL-CTT" outcomes on transcriptional pausing could be monitored, we proceeded to modify the mRNA sequence so the phase of the approaching ribosome would be different (see below); when RNAP is localised on the 1 bp backtracked signal, the mRNA sequence in "phase-TL-CTT" allows collisions to occur with no nucleotides in between the ribosome and RNAP (see Figure 29A bottom panel). Therefore, in order to modify the phase in which collisions occur insertions of 1, 2 and 3 nt were done right before the site where the mRNA hybridises with the t-DNA during AAEC formation, in a way that that wouldn't interfere with the open reading frame. As a test, we inserted 1 and 2 nt in the mRNA so when "phase-TL-CTT+1 and +2" are formed, ribosomes collide with RNAP having 1 and 2 nt more in between (see Figure 29A bottom panels "phase-TL-CTT+1" and "phase-TL-CTT+2"); phase-TL-CTT+1 and TL-CTT+2 were assembled and the experiment was carried out in exactly the same way as for phase-TL-CTT. As seen in

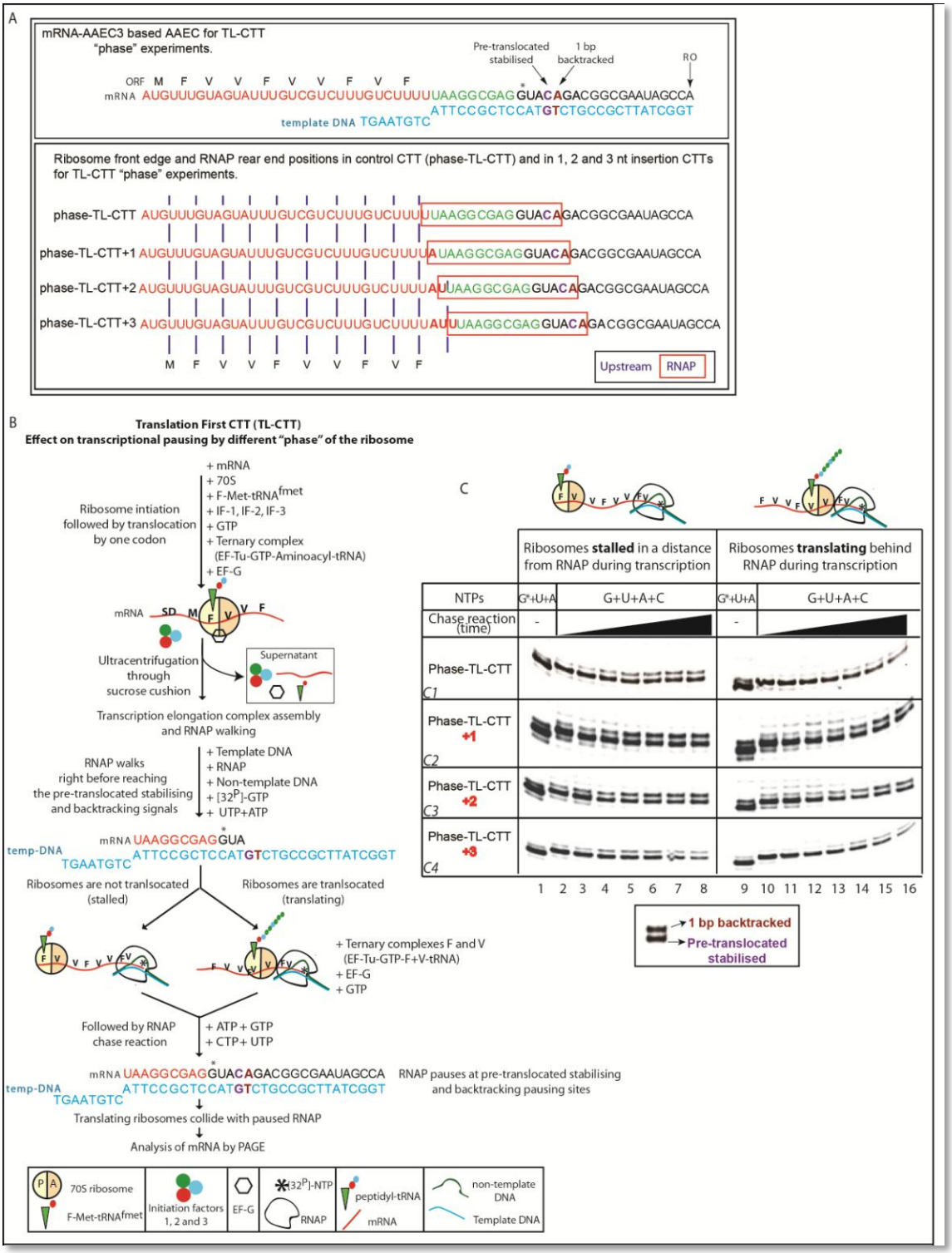

Figure 39. **Effect on transcriptional pausing on TL-CTTs with different phases between the ribosome and RNAP.** A) Top panel: mRNA-AAEC3 based AAEC for TL-CTT phase experiments. mRNA (red) with TIR and open reading frame is shown. Template DNA is shown (blue) with positions of pre-translocated and 1 bp backtracked pause signals (in purple and red correspondingly). Asterisk represents the radiolabeled nucleotide and incorporated NTPs are shown in black. Bottom panel, sequences of phase-TL-CTTs used. Ribosome front edge and RNAP rear end position are depicted (violet lines and red box respectively). Note that the distance from the codon in the P site to the front edge of the ribosome is 15 nt. Inserted nucleotides are in bold red. B) Schematic representation of phase-TL-CTT formation aimed for the study on the effects on transcriptional pausing by different phase of the colliding ribosome (see text). C) Shown is PAGE of mRNA. After phase-TL-CTTs formation, RNAP was walked right before the pre-translocated stabilised pause by addition of incomplete set of NTP's (lanes 1 and 9). Then, translation elongation complex was either left with dipeptide MF ('stalled' complexes) or was allowed to elongate behind RNAP by addition of F and V ternary complexes in presence of EF-G and GTP for 3 min ('translating' complexes). After this step, RNAP was allowed to transcribe by addition of four NTPs and aliquots were withdrawn and stop through time (in minutes: 0, 0.16, 0.5, 2, 5, 10, 20 and 40). While the stalled ribosomes remained at their initial position, the translating ribosomes followed the transcribing RNAP. Gel shows effects on transcriptional pausing (pre-translocated stabilised and 1 bp backtracked) by the ribosome colliding with different phases.

pauses for the stalled translational complexes occurred in a similar way as in “phase-TL-CTT”.

Interestingly, when the ribosomes were allowed to collide with RNAP, the read-through of the 1bp backtracked pause in both “phase-TL-CTT +1” and “phase-TL-CTT+2” was lessen compared to the effect observed for “phase-TL-CTT” (compare figure 39C panels C2, C3 and C1 lanes 10-16). However, the effect on the pre-translocated stabilised pause was not influenced by the change of the phase by 1 or 2 nt. To test if the observed changes in the read-through of the 1 bp backtracked pause could be dependent on the phase on which the ribosome collides with RNAP, we synthesised an mRNA containing an insertion of 3 nt at the same place where previous insertions were made (Figure 39A bottom panel “phase-TL-CTT+3”). By inserting 3 nt, a full codon is restored, returning the phase at the original condition (“phase-TL-CTT”). “Phase-TL-CTT+3” was assembled and the outcome on transcriptional pausing was assessed under exact conditions as for all other phase-TL-CTTs afore mentioned. As seen in (Figure 39C panel C4 lanes 2-8), in conditions where the ribosome was stalled in the one codon translocated configuration, formation of both pre-translocated and 1 bp backtracked pauses was observed in the same way as for previous “phase-TL-CTTs”; notably, when the ribosomes were allowed to translate behind RNAP and collide, the effect on the diminishing of the 1 bp backtracked paused complex differed from the observed in +1 and +2 phase-TL-CTTs becoming similar to the outcome observed for “phase-TL-CTT” (for comparison see Figure 39C panels C4 and C1). Effect on the pre-translocated stabilised complex was not seen as in the rest of the “phase-TL-CTT” tested. As an indirect way to analyse the effect of the collision with paused RNAP on the ribosome, we performed RelE probing on phase-TL-CTT having RNAP paused on both pre-translocated and 1 bp backtracked sites (Figure 40A). RelE cleavage for the stalled (1 codon translocated) ribosomal complexes was obtained (Figure 40 B lane 2), whereas the RelE cleavage corresponding to ribosomes which were allowed to translocate and collide with RNAP was not obtained (compare Figure 40B lanes 2 and 4). This result may suggest that upon colliding with paused RNAP, the ribosome is halted disallowing the transfer of the aminoacyl-tRNA from the A site to the P site (which depends on the translocation of the ribosome), which in turn does not allow a vacant A site to form for RelE cleavage to bind and cleave the mRNA (Figure 40A); another possible explanation for this result is the different affinity of RelE to cleave certain codons but normally, even in low affinity codons some cleavage is observed.

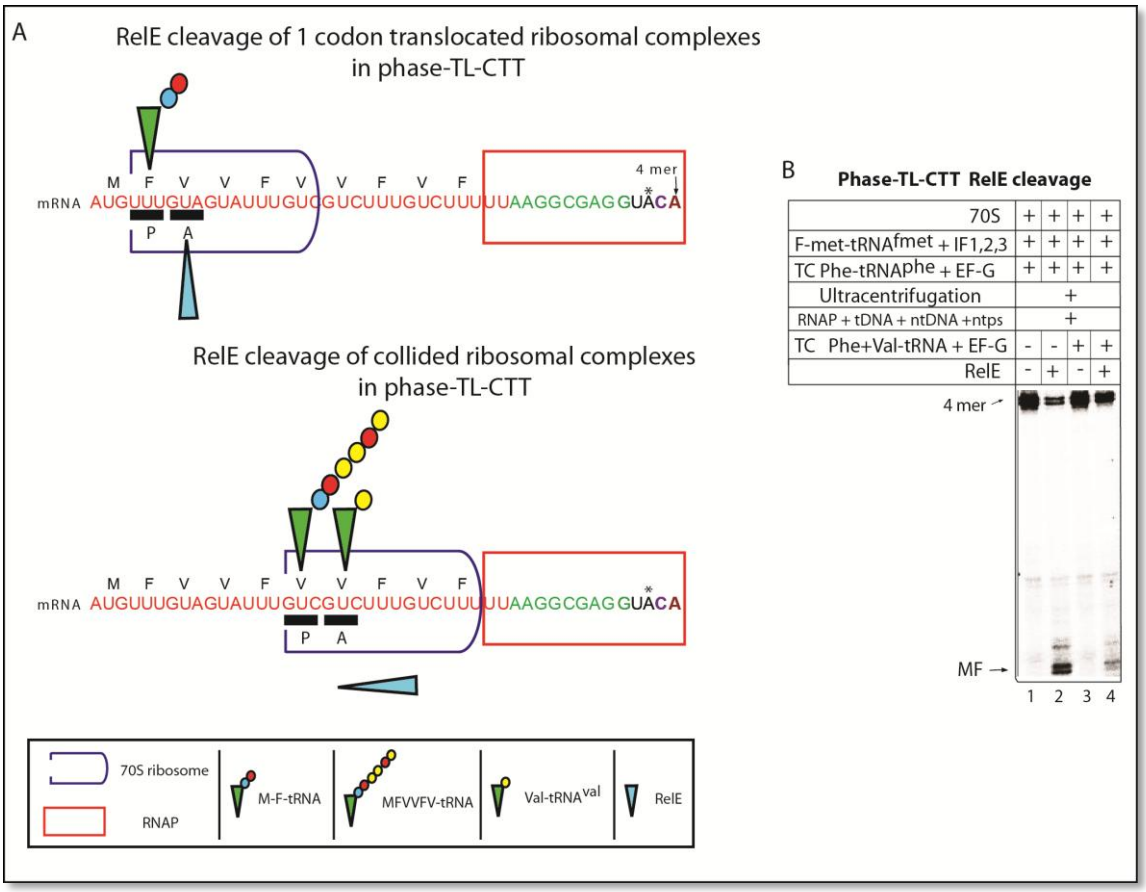

Figure 40. RelE cleavage on "1 codon translocated" and "collided" ribosomal complexes in phase-TL-CTT. A) Schematic representation of the RelE cleavage on phase-TL-CTT. Top panel: RelE cleavage of 1 codon translocated ribosomal complexes; mRNA (light red), RNA:t-DNA hybridisation site (green), pre-translocated stabilised signal (purple), 1 bp backtracking site (dark red), open reading frame are shown. Black rectangles indicate the ribosome sites P and A. Violet figure represents the ribosome; it covers 15 nt from the first nucleotide of the codon occupying the P site. In the P site, the peptidyl-tRNA composed by the dipeptide MF is shown. Since Val-tRNA^{val} is not added to the reaction, the ribosome is stalled. The A site remains unoccupied allowing RelE binding and cleavage of the mRNA. Bottom panel: Conditions were given to the ribosome to collide with RNAP paused in the pre-translocated stabilised and 1bp backtracking pausing sites. Ribosome translocates by four codons and collides with RNAP. The P site contains a peptidyl-tRNA composed by the hexapeptide MFVVFV-tRNA; in the A site of the ribosome, the following cognate aminoacyl-tRNA to be incorporated into the nascent peptide (Val-tRNA^{val}) is bound. Paused RNAP impedes further translocation of the ribosome which doesn't allow the A site to become vacant interfering with RelE binding. B) Shown is PAGE of mRNA from phase-TL-CTT RelE cleavage experiment. Translation initiation followed by 1 codon translocation complexes were purified by ultracentrifugation. After formation of phase-TL-CTT RNAP was walked 4 nucleotides allowing RNAP to pause in the pre-translocated stabilised and 1 bp backtracking sites. The 1 codon translocated complexes were treated with RelE (compare lane 2 and 1) and cleavage for this complex was obtained. Ribosomes were allowed to transcribe and collide with RNAP; In the presence of RelE, no cleavage for the translocated complexes was observed (compare lane 4 and 3) suggesting that the A site was occupied by an aminoacyl-tRNA blocking the access to RelE.

Taking into account that the four "phase-TL-CTTs" tested in this experiment collide with RNAP, restoring of the effect on 1 bp backtracked paused RNAP on "Phase-TL-CTT+3" suggested that the interaction between the ribosome and RNAP does not depend simply on an external force applied on RNAP which "pushes" it forward but might indicate that a more intricate mechanism of cross-talk is involved; the experiments performed in this section were done having only the minimum components for both transcription and translation to function, so the observed outcomes on

transcriptional pausing can only be attributed to the direct physical interaction amongst the ribosome and RNAP; this suggests that an interface could exist and that perhaps it needs to be in a certain alignment with RNAP for the two machineries to interplay. The lack of effect on pre-translocated stabilised pause regardless of the phase of the collision could not be explained in this way; however the existence of strong transcriptional pauses which are not affected by the colliding ribosome might serve as a regulatory signal on translation, which may consist in pausing of the ribosome (as suggested by RelE experiments on collided ribosomal complexes) to allow for instance, *in situ* modifications and folding of the nascent peptide. Experiments regarding the effects on ribosome-unaffected-pauses on translation are currently being executed.

6.7 Discussion

Throughout the formation and characterisation of AAECs we have set the base for the development of experimental set-ups aiming for the elucidation of the apparently complex interactions between RNAP and the ribosome. For the achievement of intelligent design of our AAEC we took into account all the observations made on the early experimental set-ups, from the analysis of our *in vitro* translation system to the assembly of the promoter-borne CTT. Incorporating AAEC as the primary vehicle for the assembly of CTT eliminated the need of long RNAP walking reactions reducing the background noise. AAEC also provided flexibility to the CTT by permitting the analysis of different transcription elongation states by simply exchanging template and non-template DNAs on the same mRNA. It also allowed us to overcome the limitation of low mRNA yields we encountered on the promoter borne CTT, permitting the visualisation of the synthesised peptides in the CTT. The "transcription first" configuration, permitted us to develop a system where the effects on translation could be monitored. Firstly, we tested peptidyl transferase reaction in the translational system alone. The visualisation of the peptides was achieved by testing TLC-based chromatography techniques until we found that TLE had the resolving capacity needed for our purposes, permitting us to continue the development of the CTT. As a result, TR-CTT was used to roughly find permissive and non-permissive distances between RNAP and the initiating ribosome. These types of measurement have never been performed before, simply because an appropriate system was inexistent.

Besides determining distances, the TR-CTT can be used to analyse the effects on translation caused by specific transcription elongation complexes, such as paused

RNAP which are not affected by the translating ribosome that suggest that transcription may block translation (as indirectly observed by RelE cleavage of collided ribosomes with paused RNAP) possibly revealing a more intricate regulatory mechanism. We complemented the analysis of transcription-translation coupling by the development of a system designed for the study of the effects of coupling on transcription. For the "translation first" (TL-CTT) assembly, the biggest challenge was obtaining an AAEC fully occupied by translating ribosomes. Ultracentrifugation of the one codon elongated translational complexes through a sucrose cushion permitted almost complete occupancy of the ribosomes on the AAEC and also removed initiation factors and unused tRNAs and mRNA allowing us to work with pure translation elongation complexes. RelE probing and PTR assays demonstrated that the translational system could survive not only the purification process but also remained active after the incubation times and in the presence of RNAP and nucleic acids needed for the formation of the CTT.

TL-CTT usage was exemplified by the observation of the effects caused on hairpin-independent transcriptional pausing by ribosomes non-colliding and colliding with paused RNAP. The various outcomes observed on paused RNAP by the translating ribosome suggest an intricate mechanism needed for the cross-talk amongst the machineries. The observation of some hairpin independent pauses being diminished, unaffected and even strengthen by the translating ribosome could indicate that the outcome of the interaction between the ribosome and RNAP does not only depend on the oscillation of RNAP between the pre and post-translocated states which makes RNAP prone to be repositioned into different states by external forces, as observed for RNAP-RNAP interactions (Epshtein and Nudler, 2003), but possibly in a specific, interface-mediated, allosteric way which could cause isomerisation of RNAP leading to different effects on transcription elongation. The observation of the read-through of the 1b backtracked complex caused by the ribosome could also be interpreted as if the ribosome could "hold" RNAP back disallowing it from continuing transcription elongation, therefore instead of diminishing of the 1 bp backtracked complex what could be observed is the strengthen of the pre-translocated stabilised paused RNAP. The observed phase-dependent effect could indicate that the ribosome-RNAP interface requires an exact "position" of the ribosome in order to be capable of interact. It is possible that an allosteric mechanism combined with external force brought by the ribosome could be responsible for the diminishing RNAP pausing.

Pauses which were not affected by the translating ribosome (such as the pre-translocated stabilised pause) and even transcriptional pauses which are strengthen (such as P4), could serve as regulatory signals for translation; it is possible that these pauses contribute in protein folding or could also induce ribosomal frameshifting. Experiments carried out using TL-CTT could be aimed for the understanding of the effects of translation on any transcriptional estate and also can contribute to the finding of the interface on the ribosome and RNAP needed for such interactions to occur. The variety of outcomes on transcriptional pausing observed with the use of TL-CTT is just an example of the plethora of scenarios which can be studied using the system described. These experiments are just preliminary approaches towards the understanding of the mechanism of the interplay between the ribosome and RNAP which could be a major gene expression regulatory pathway in prokaryotes.

Chapter 7. Concluding discussion.

Over the past 20 years, as mentioned in the introductory chapter, data has been produced on the effect of coupling on the topics of gene regulation and attenuation; most recently, efforts to identify the mechanistic consequences of such interactions on transcription have been performed. Although the data generated in such studies is highly valuable, the experiments were performed *in vivo* and the conclusions derived from them do not disregard the possibility of the observations being affected by both known and unknown transcriptional and/or translational factors present in the living cell. Therefore, the idea for the development of an *in vitro* transcription-coupled-to-translation systems for the analysis of the interactions between the ribosome and RNAP parts from the need of answering the question of what happens with the transcript and protein synthesis (in terms on the consequences on specific RNAP elongation complexes) when the machineries collide.

In order to couple translation to transcription we first established the translational system taking into consideration two characteristics: purity and the capacity of step-wise translocation of the ribosome. Therefore, we purified all minimal factors required for translation to function; each aminoacyl-tRNA was prepared and purified independently and separately from the translation reaction, minimising the presence of factors and NTPs which could affect the CTT. Throughout techniques such as toeprinting we were capable of monitoring the ribosome translating one codon at a time suggesting that our system was active. Having the translational system operational we

coupled it to transcription. The first approach consisted in the assembly of the CTT from transcription originating from a promoter. During the development of the promoter-borne CTT, we incorporated RelE in the system as a tool to probe the initiating and elongating ribosome; throughout RelE cleavage we observed co-existence of the transcriptional and translational machinery in the promoter-borne CTT. We also observed that we were capable of translocating the ribosome in a step-wise manner towards RNAP, satisfying one of the specific aims of this study. In this CTT we were also capable of positioning RNAP in any transcriptional state through RNAP walking reactions. As a test, we used promoter borne CTT to monitor the effects on transcriptional backtracking by the transcribing ribosome. The observed effects on transcriptional backtracking analysed through GreB cleavage suggested that the ribosome acted as a physical barrier which impeded backtracking of RNAP stabilising it in the post-translocated state. This was the first observation in which the system would allow, in a direct way, the analysis of the outcome of the interaction between RNAP and the ribosome. Through the use of promoter borne CTT we also evaluated the effect on roadblocked RNAP by the translating ribosome; we observed read-through of the EcorQ111 roadblock in both initiating and translating ribosomes. For initiating ribosomes and translocated ribosomes, read-through of the roadblock was observed, suggesting the participation of the ribosome in the dislodgment of DNA bound proteins which might interfere with transcription elongation. Backtracking of RNAP caused by its collision with the roadblock was also assessed by GreB cleavage; the results suggested that the ribosome blocked RNAP from backtracking (as observed in the previous experiment with RNAP backtracked from a defined signal on the DNA). These results are in partial agreement with observations made *in vivo* where RNAP blocked by the lac repressor was aided by the ribosome to overcome it (seen by reverse transcription) (Proshkin et al., 2010) with the exception that *in vitro* (by looking directly at the transcript) we do not detect the "pushing" effect that they detected through foot-printing of RNAP to observe its localisation on the DNA. The effects on backtracking and roadblocking obtained using the CTT *in vitro* were comparable with data obtained *in vivo* by Proshkin and co workers using a complete different strategy, suggested that the CTT system could be used for the analysis of the outcomes on transcription by the ribosome, fulfilling one of the aims proposed for this study. Although the promoter borne CTT could be use for further studies, it had intrinsic flaws such as high background noise and low mRNA concentration for visualisation and analysis of

synthesised peptides; therefore we explored different ways to couple the systems to improve the analysis of the interactions amongst them.

As mentioned in the introductory chapter, the effects on translation caused by the collision of the ribosome with RNAP remains unanswered due to the lack of a system which would allow the study of such event. Therefore, we developed a system where the effects on translation could be monitored; TR-CTT was based on AAEC instead of promoter-borne transcription, allowing us to use higher concentrations of mRNA to program the ribosome, permitting the synthesis of peptides which could be visualised by TLE. TR-CTT usage was exemplified by monitoring the effect caused by RNAP (localised at two different distances) on translation initiation judged by dipeptide formation. We observed that a distance between the ribosome (from the first nt of the initiation codon AUG) to the rear end of RNAP of 16 nt had a detrimental effect on translation initiation compared to a distance of 27 nt where normal initiation took place; TR-CTT can be used for the precise measurement of permissive and non-permissive distances during translation initiation and elongation which is currently under way; this system could also be used for the elucidation of the interface between the ribosome and RNAP and also to study the effect both known and unknown transcriptional and translational factors on peptide synthesis in the CTT.

We developed an AAEC based CTT system aimed for the analysis of different aspects on transcription elongation which could be influenced by the translating ribosome such as hairpin independent pausing. Through the use of TL-CTT we monitored the outcome on transcriptional pausing. Interestingly, we observed different effects caused by the actively transcribing ribosome, including diminishing and strengthen of paused ECs as well as pauses which were unaffected; this result suggests a more complex means of interaction than just a physical “push” of RNAP by the ribosome. The elucidation of the mechanism on which the different pauses respond to the interacting ribosome opens a new field of research for the understanding of the regulatory pathway involving the interplay of transcription and translation. To further analyse the complexity of the effects caused on transcriptional pausing we used TL-CTT to explore different phases on which the ribosome collide with paused RNAP. We observed, using two defined pausing signals (pre-translocated stabilised and 1 bp backtracked) that the read-through of the 1bp backtrack pause depended on the manner in which the ribosome collided with RNAP. Such result supported the idea that the mechanism of interaction not only depends on an interface but also may require an appropriate alignment in order to generate an effect on transcription. Notably, the pre-translocated stabilised pause was

not affected by the ribosome in any of the tested phases. We hypothesise that such strong pauses may function as regulatory signals for translation.

Taking into account the observed behaviour on transcriptional pausing upon collisions amongst the two machineries, it could be hypothesised that if there are no "bridging" factors which regulate the cross-talk between the ribosome and RNAP, it would be possible that the ribosome (either the rRNA or ribosomal proteins) may interacts with β-flap domain which is known to regulate hairpin-dependent pausing; however, the interaction can also occur with β' zinc finger, lid and zipper which can also generate allosteric modifications to RNAP. Cryo-EM Experiments to elucidate the contacting surface between the two machineries are currently being performed.

We tested the status of the translational machinery in the TL-CTT through RelE which was used to probe ribosomes stalled far apart from paused RNAP. We observed RelE cleavage typical for such complex and when the probing was performed in collided ribosomes with paused RNAP no RelE cleavage was obtained, suggesting that the ribosome contained the A site occupied a situation which occurs if translocation is impeded, in this case by paused RNAP, suggesting that pausing could regulate the rate of translation as well. Currently, experiments to directly observe the effect on the synthesised peptides of collided ribosomes with paused RNAP are being carried out.

To conclude, the translation coupled to transcription systems described in this thesis will, for the first time, allow direct assessment of the effects on both transcription and translation upon physical contacts of both machineries. The system permits investigation of the effects of the ribosome on various transcription complexes (paused, backtracked, etc.), as well as the possible control of the rate of translation by RNAP. While eliminating possible interference from other cellular components, the CTT systems described here will also allow exploration of the effects of individual transcription/translation or other factors on the cross-talk between the two machineries. CTTs will be used for measuring distances between transcribing RNAP and the ribosome translating the nascent RNA and can also be used to see if translational events such as frameshifting could occur after collisions with stalled RNAP occur. Finally, CTTs will also be useful for understanding contact interfaces between the two machineries.

Chapter 8. Conclusions.

Lately, a major importance has been given to the interplay amongst cellular machineries such as transcription, translation and replication; the analysis of interactions of RNAP with the replication machinery has revealed the existence of a complex regulatory mechanism which affects cellular activities ranging from gene expression to genome organisation and stability. However, the complete understanding of the cross-talk amongst the major cellular machineries is not complete due to the lack of characterisation of the interactions between RNAP and the translational apparatus. Recently, *in vivo* approaches have been designed to study the effects on transcription elongation caused by the ribosome and also on the relationship of these effects on replication fork progression. Nevertheless, the mechanistic interactions of the colliding ribosome with RNAP and its outcome on peptide synthesis and on transcriptional elongation complexes stabilised in different states remains obscure. Therefore, as an attempt to elucidate these outcomes we decided to design an *in vitro* reconstituted system which would permit the monitoring of the outcomes of the interactions between the ribosome and RNAP.

In this work we have purified all components needed for the assembly of both transcription and translation machineries to function *in vitro*, along with some transcriptional factors needed for the analysis of the outcomes on transcription.
Using the purified components, we have developed three CTT systems in which step-wise translocation of the ribosome towards RNAP polymerase has been achieved. The systems also permit the step-wise translocation of RNAP for its positioning on different transcription elongation states (paused, backtracked, roadblocked, etc). The systems which were assembled and the outcomes monitored with them are:

1. The promoter-borne CTT: it allowed preliminary observations on the effects on backtracking; the observed outcome suggested that the ribosome could act as a barrier which stopped the RNAP backtracking. The system also allowed monitoring of the effect on roadblocked RNAP; the obtained outcome suggested a mechanism for the dislodgment of the roadblock mediated by the ribosome.
2. "Transcription first-CTT" (TR-CTT) system: In this system all translational complexes are coupled to transcription; therefore it can be

used to analyse the outcomes of the collision between the ribosome and RNAP on translation by the direct observation of the nascent peptide synthesised in the CTT. The system was used to roughly analyse the effect on translation initiation depending on the distances between RNAP and the ribosome. Precise measurements of distances between the ribosome and RNAP are currently under way. The system can also be used to analyse if paused RNAP can function as a regulatory signal for the translational machinery.

3. "Translation first-CTT" (TL-CTT) system: in this system all transcriptional complexes are coupled to translation; therefore, it can be used to analyse the effects on the collision between the ribosome and RNAP on transcription by following the nascent, radiolabeled mRNA. The system was used to observe the effects on transcriptional pausing caused by the translating ribosome. The outcomes obtained suggested a complex influence on paused RNAP where some pauses appeared to be diminished, unaffected and strengthen, proposing an intricate mechanism of cross-talk amongst the two machineries. Moreover, TL-CTT was also used to evaluate the effect of the phase in which the ribosome approaches RNAP influences the outcome transcription; the results suggested that the phase in which the ribosome interacts with RNAP is a determinant key for interactions to take place. TL-CTT can be used to analyse the effects on diverse transcription elongation complexes; it can also be used for the analysis of the function on transcription-translation coupling of known and unknown factors.

List of meetings and publications

Publications from this work

1) Castro-Roa, D, Zenkin, N. 2012. In Vitro Experimental system for the analysis of transcription-translation coupling. *Nucleic Acids Res. 40(6):e45.*

2) Daniel Castro-Roa and Nikolay Zenkin. 2011. Relations Between Replication and Transcription, Fundamental Aspects of DNA Replication, Jelena Kušić-Tišma (Ed.), ISBN: 978-953-307-259-3, InTech, Available from: http://www.intechopen.com/articles/show/title/relations-between-replication-and-transcription

Previous publications

1) Joseph T. Wade, Daniel Castro-Roa, David C. Grainger, Douglas Hurd, Stephen J W Busby, Kevin Struhl & Evgeny Nudler. 2006. Gene regulation: The logic of sharing. Research Highlights. *Nature Reviews Microbiology* 4, 724

2) Joseph T. Wade, Daniel Castro Roa, David C. Grainger, Douglas Hurd, Stephen J W Busby, Kevin Struhl & Evgeny Nudler. 2006. Extensive functional overlap between sigma factors in *Escherichia coli*.. *Nature Structural & Molecular Biology*. 13:806-14

3) Ekaterina V. Mirkin, Daniel Castro Roa, Evgeny Nudler & Sergei M. Mirkin. 2006. Transcription regulatory elements are punctuation marks for DNA replication. *PNAS* 103, 7276-81

List of meetings

1) FASEB SRC Summer conferences Mechanism & Regulation of Prokaryotic Transcription. June 19 - 24, 2011 in Vermont Academy, Saxtons River, Vermont. "Effects on transcriptional pause by the translational machinery" Oral presentation

2) RNA polymerase workshop, University of Leeds, UK. April 2011. "Outcomes of coupling of transcription and translation on transcriptional pausing" Oral presentation

3) FASEB SRC Summer conferences Mechanism & Regulation of Prokaryotic Transcription. June 19 - 24, 2009 in Vermont Academy, Saxtons River, Vermont. "Setting up an *in vitro* transcription-translation system for the analysis of the outcomes on transcription elongation" Poster presentation

4) RNA polymerase workshop, University of Nottingham, Nottingham, UK. March 2010. "Assessing interactions of the translational machinery wit E. coli RNA Polymerase" Oral presentation

5) RNA polymerase workshop, Bristol University, Bristol, UK March 2009. "Setting up the system to study RNA polymerase and its interaction with translational machinery" Oral presentation

Appendix 1.

Table A1. List of strains

Strain	antibiotic resistance	Characteristic	Inducing agent	Genotype	Reference
MG1655	N/A	Wild type	N/A	*F⁻ λ⁻ ilvG- rfb-50 rph-1*	Guyer, M.S. *et al.* 1981
DH5 α	N/A	Used for plasmid maintenance and segregation	N/A	*F⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(r_K^- m_K^+), λ–*	Meselson M. and Yuan R. 1968
BL21 (DE3)	N/A	Used for expression of genes cloned under the control of T7 promoter.	N/A	*F⁻ ompT gal dcm lon $hsdS_B$(r_B^- m_B^-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	Studier, FW. 1986
RL721	Kanamycine	RNAP-beta-(his)X6	N/A	*poC3531(His6) zja::kan thr-1 araC14 leuB6(Am) Δ(gpt-proA)62 lacY1 sbcC201 tsx-33 qsr'-0 glnV44(AS) galK2(Oc) LAM- Rac-0 sbcB15 hisG4(Oc) rfbC1 rpoS396(Am) recB21 recC22 rpsL31(strR) kdgK51 xylA5 mtl-1 argE3(Oc) thi-1*	Wang, D. *et al.* 1995
RpoBtap	Kanamycine	RNAP-beta-TAP tagged	N/A	*TAP tagged fused to rpoB (MG1655 background)*	Butland, G. *et al* 2005.
ME5305	N/A	Used for plasmid maintenance and segregation (ASKA strain collection)	N/A	*recA1 endA1 gyrA96 thi-1 hsdR17(rk- mk+) supE44 relA1*	Kitagawa, M. *et al* 2005
m5248	N/A	Used for EcoRQ111 plasmid segregation and expression	Heat shock	*Lysogen of Xbio-275 cI857 AHI*	Modrich, P. 1986
MRE600	N/A	Used for purification of 70S ribosomes	N/A	*F-, rna*	Wehr, C.T. 1973

Appendix 2.

Table A2. List of plasmids

Plasmid	Resistance marker	Characteristics	Inducing agent	Reference
pCAN	Chloramphenicol	Overexpression of N terminal 6XHis proteins. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
Pet21-*rpoD*	Ampicillin	Overexpression of C terminal 6XHis sigma 70 factor. T7 RNA polymerase promoter	IPTG	Castro-Roa, D. 2012
pCAN-*greB*	Chloramphenicol	Overexpression of N terminal 6XHis GreB transcription factor. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
pCAN-NusA	Chloramphenicol	Overexpression of N terminal NusA transcription factor. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
pCAN-NusG	Chloramphenicol	Overexpression of N terminal NusG transcription factor. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
pQ111	Ampicillin	Overexpression of EcoRI mutant (EcoRQ111 Gln111Glu). Expresses both EcoRQ111 and Methylase.	Heat shock	Modrich, P. 1986.
pHisRelE/RelB	Ampicillin	Overexpression of Toxin (His-RelE) and RelB antitoxin.	IPTG	Christensen S. *et al* 2001.
pCAN-EF-G	Chloramphenicol	Overexpression of N terminal EF-G translation factor. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
pCAN-EF-Tu	Chloramphenicol	Overexpression of N terminal EF-Tu translation factor. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
pCAN-IF1	Chloramphenicol	Overexpression of N terminal IF-1 translation factor. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
pCAN-IF2	Chloramphenicol	Overexpression of N terminal IF-2 translation factor. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
pCAN-IF3	Chloramphenicol	Overexpression of N terminal IF-3 translation factor. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
pCAN-MetRS	Chloramphenicol	Overexpression of N terminal metRS aminoacyl synthetase. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
pCAN-PheRS	Chloramphenicol	Overexpression of N terminal pheRS aminoacyl synthetase. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.
pCAN-FMT	Chloramphenicol	Overexpression of N terminal Formyl methionyl transferase. T7 RNA polymerase promoter. *cat.* Obtained from ASKA collection	IPTG	Kitagawa, M et al. 2005.

Appendix 3.

Table A3. List of DNA, RNA and AAEC templates not shown in the figures/text.

Name	Sequence	Characteristics
T7 g32 -42 +86	tcccgcgaaattaatacgactcactatagggagacaactgttaattaaattaa attaaaaaggaaataaaaatgtttgtatacaaatctactgctgaactcgctgc acaaatggctaaactgaatggcaataaaggtttttcttctgaagataaag	Used for in vitro synthesis of g32 -42 + 86 RNA for toeprinting and RelE printing experiments. Red: T7 promoter.
T7 g32 -42 +87	gggagacaacuguuaauuaaauuaaauuaaaaaggaaauaaaaauguuuguauacaaaucuacugcugaacucgcugcacaaauggcuaaacugaauggcaauaaagguuuuucuucugaagauaaag	Red: Shine-Dalgarno sequence, green first 5 codons M,F,V,Y,K. Blue: oligo binding site for reverse transcription
T7A1 g32ORF9	cttaaagagacttaaaccagatcccgaaaatttatcaaaaagagtattgacttaaagtctaacctataggatacttacagccatcgagacaactgttaattaaattaaattaaaaaggaaataaaaatgtttgtatacaaatttgtatttgtactcgctgcacaaatggctaaactggaattcaataaaggtttttcttctgaagataaagcatcgagagggacacggcgaacagccatcccaatcgacaccgggc	Red: T7A1 promoter -35, -10 and +1 elements. Green: SD and ORF used. Blue: EcoRI site. Template was used for the characterisation of promoter-borne CTT and roadblock experiments
T7A1 backtracking	cttaaagagacttaaaccagatcccgaaaatttatcaaaaagagtattgacttaaagtctaacctataggatacttacagccatcgagacaactgttaattaaattaaattaaaaaggaaataaaaatgtttgtatacaaatttgtatttgtactcgctgcatcgagagggacacggcgaacagccatcccaatcgacaccgggcacaaatggctaaactggaattcaataaaggtttttcttctgaagataaag	Red: T7A1 promoter -35, -10 and +1 elements. Green: SD and ORF used. Blue: backtracking signal. Template was used for the backtracking experiments on promoter-borne CTT
mRNA-AAEC1	gggagacaacuguuaauuaaauuaaauuaaaaaggaaauaaaaauguuuguauacaaagagaauaaucgagagc	Red: Shine-Dalgarno sequence, green first 5 codons M,F,V,Y,K. Blue template DNA hybridisation site. mRNA used for AAEC formation and CTT characterisation
template DNA1	tgaatgtcggtagctctcgctttcggtttacgcggtgggtt	Formation of AAEC
non-template DNA1	acttacagccatcgagagcgaaagccaaatgcgccacccaa	Formation of AAEC

Appendix 4.

Purification of RNA polymerase and translation factors.

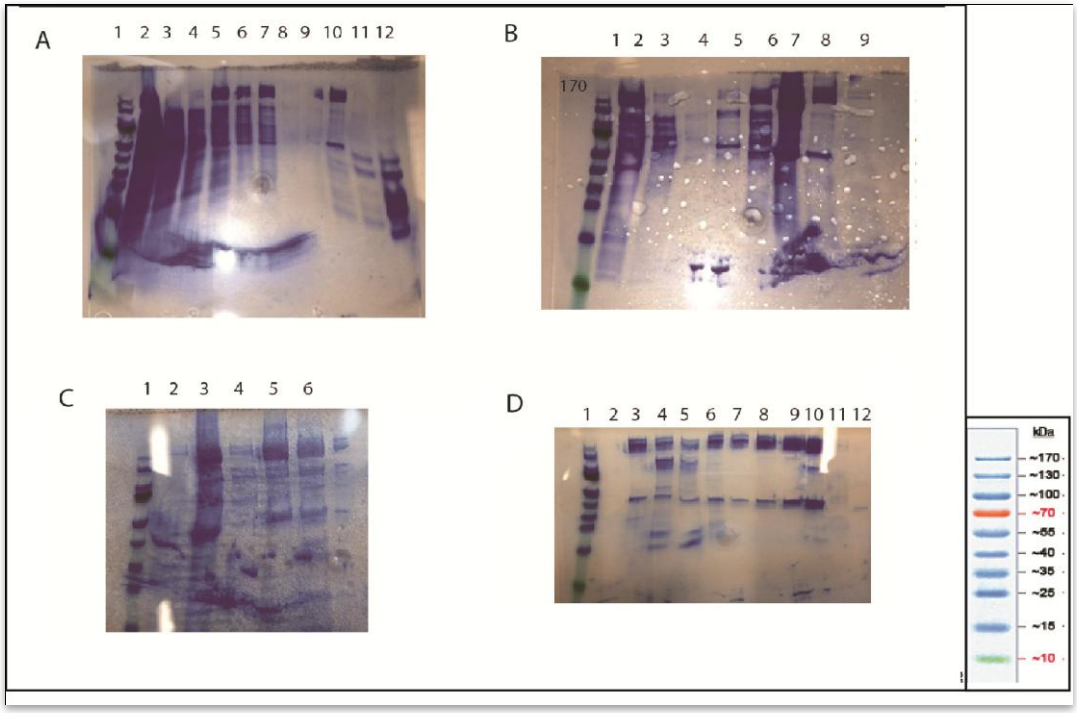

Figure A1. *E. coli* RNA polymerase purification. Coomassie R stained, SDS-PAGE containing samples obtained during purification of RNAP. A) After lysis (lane 2) and polymine steps, enrichment of RNAP can be observed (lane 5). This sample was subjected to affinity chromatography (Heparin) where partially purified RNAP was obtained (lane 10). B) Affinity chromatography was repeated for the other half of the samples obtained from polymine P. Partially purified RNAP is obtained (lanes 7, 8) C) Partially purified RNAP obtained from affinity chromatography was applied onto a size exclusion column (superose 6). Lanes 5 and 6 show the pure RNAP holo-enzyme. D) To separate core enzyme from σ^{70}, the samples were applied onto an ion exchange column (Mono Q). Lanes 6-9 correspond to core enzyme RNAP. Bands in the lanes (6-9) correspond to (from top to bottom): β, β' and α (alpha subunit).The extra band observed between β and α in lanes 4 and 5 correspond to σ^{70}. Core enzyme was pooled, concentrated and dialysed against storage buffer. Typical concentration is 2 mg/mL of RNAP. Note that the molecular weight ladder shown as reference is the same used in the SDS-PAGE.

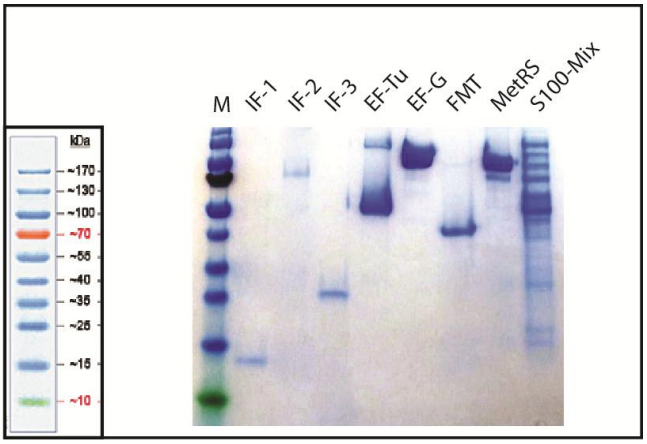

Figure A2. Purification of translation factors. Coomassie R stained, SDS-PAGE showing some of the purified factors for the *in vitro* assembly of the translational machinery. Concentrations are: IF-1, IF-2 and IF-3 (1 mg/mL). EF-Tu and EF-G: 3 mg/mL. Note that the upper band on EF-Tu is result of EF-Tu self-assembly which takes place during protein concentration. Note that the molecular weight ladder shown as reference is the same used in the SDS-PAGE.

List of references

ABBONDANZIERI, E. A., GREENLEAF, W. J., SHAEVITZ, J. W., LANDICK, R. & BLOCK, S. M. 2005. Direct observation of base-pair stepping by RNA polymerase. *Nature,* 438**,** 460-5.

AIYAR, S. E., HELMANN, J. D. & DEHASETH, P. L. 1994. A mismatch bubble in double-stranded DNA suffices to direct precise transcription initiation by *Escherichia coli* RNA polymerase. *J Biol Chem,* 269**,** 13179-84.

ARTSIMOVITCH, I. & LANDICK, R. 2000. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc Natl Acad Sci U S A,* 97**,** 7090-5.

BAI, L., SHUNDROVSKY, A. & WANG, M. D. 2004. Sequence-dependent kinetic model for transcription elongation by RNA polymerase. *J Mol Biol,* 344**,** 335-49.

BAILEY, M. J. A., HUGHES, C. & KORONAKIS, V. 1997. RfaH and the *ops* element, components of a novel system controlling bacterial transcription elongation. *Mol. Microbiol,* 26**,** 845-851.

BAN, N., FREEBORN, B., NISSEN, P., PENCZEK, P., GRASSUCCI, R. A., SWEET, R., FRANK, J., MOORE, P. B. & STEITZ, T. A. 1998. A 9 A resolution X-ray crystallographic map of the large ribosomal subunit. *Cell,* 93**,** 1105-15.

BANERJEE, S., CHALISSERY, J., BANDEY, I. & SEN, R. 2006. Rho-dependent transcription termination: more questions than answers. *J Microbiol,* 44**,** 11-22.

BAR-NAHUM, G., EPSHTEIN, V., RUCKENSTEIN, A. E., RAFIKOV, R., MUSTAEV, A. & NUDLER, E. 2005. A ratchet mechanism of transcription elongation and its control. *Cell,* 120**,** 183-93.

BARNE, K. A., BOWN, J. A., BUSBY, S. J. & MINCHIN, S. D. 1997. Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *Embo J,* 16**,** 4034-40.

BOCHKAREVA, A., YUZENKOVA, Y., TADIGOTLA, V. R. & ZENKIN, N. 2011. Factor-independent transcription pausing caused by recognition of the RNA-DNA hybrid sequence. *EMBO J,* 29;31(3):630-9. doi: 10.1038/emboj.2011.432.

BORUKHOV, S. & GOLDFARB, A. 1993. Recombinant *Escherichia coli* RNA polymerase: purification of individually overexpressed subunits and *in vitro* assembly. *Protein Expr Purif,* 4**,** 503-11.

BORUKHOV, S., LAPTENKO, O. & LEE, J. 2001. *Escherichia coli* transcript cleavage factors GreA and GreB: functions and mechanisms of action. *Methods Enzymol,* 342**,** 64-76.

BORUKHOV, S. & NUDLER, E. 2008. RNA polymerase: the vehicle of transcription. *Trends Microbiol,* 16**,** 126-34.

BORUKHOV, S., SAGITOV, V. & GOLDFARB, A. 1993. Transcript cleavage factors from *E. coli*. *Cell,* 72**,** 459-66.

BOURDEAU, V., PAQUIN, B., CEDERGREN, R. & DELIHAS, N. 2001. Cleavage of mitochondria-like transfer RNAs expressed in *Escherichia coli*. *FEBS Lett,* 497**,** 90-4.

BREWER, B. J. 1988. When polymerases collide: replication and the transcriptional organization of the *E. coli* chromosome. *Cell,* 53**,** 679-86.

BRODOLIN, K., ZENKIN, N., MUSTAEV, A., MAMAEVA, D. & HEUMANN, H. 2004. The sigma 70 subunit of RNA polymerase induces lacUV5 promoter-proximal pausing of transcription. *Nat Struct Mol Biol,* 11**,** 551-7.

BRUECKNER, F., ARMACHE, K. J., CHEUNG, A., DAMSMA, G. E., KETTENBERGER, H., LEHMANN, E., SYDOW, J. & CRAMER, P. 2009a. Structure-function studies of the RNA polymerase II elongation complex. *Acta Crystallogr D Biol Crystallogr,* 65**,** 112-20.

BRUECKNER, F., ORTIZ, J. & CRAMER, P. 2009b. A movie of the RNA polymerase nucleotide addition cycle. *Curr Opin Struct Biol,* 19**,** 294-9.

BRUNELLE, J. L., YOUNGMAN, E. M., SHARMA, D. & GREEN, R. 2006. The interaction between C75 of tRNA and the A loop of the ribosome stimulates peptidyl transferase activity. *RNA,* 12**,** 33-9.

BUCK, M., GALLEGOS, M. T., STUDHOLME, D. J., GUO, Y. & GRALLA, J. D. 2000. The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *J Bacteriol,* 182**,** 4129-36.

BURGESS, R. R. & ANTHONY, L. 2001. How sigma docks to RNA polymerase and what sigma does. *Curr Opin Microbiol,* 4**,** 126-31.

BURMANN, B. M., SCHWEIMER, K., LUO, X., WAHL, M. C., STITT, B. L., GOTTESMAN, M. E. & ROSCH, P. 2010. A NusE:NusG complex links transcription and translation. *Science,* 328**,** 501-4.

BUTLAND, G. PEREGRÍN-ALVAREZ JM, LI J, YANG W, YANG X, CANADIEN V, STAROSTINE A, RICHARDS D, BEATTIE B, KROGAN N, DAVEY M, PARKINSON J, GREENBLATT J & EMILI A. 2005 Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* 3; 531-7.

CAMPBELL, E. A., MUZZIN, O., CHLENOV, M., SUN, J. L., OLSON, C. A., WEINMAN, O., TRESTER-ZEDLITZ, M. L. & DARST, S. A. 2002. Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell,* 9**,** 527-39.

CASTRO-ROA, D. & ZENKIN, N. 2011. Relations Between Replication and Transcription. *Fundamental Aspects of DNA Replication* [Online], 1. Available: http://www.intechopen.com/articles/show/title/relations-between-replication-and-transcription.

CASTRO-ROA D & ZENKIN, N. 2012 *In vitro* experimental system for analysis of transcription-translation coupling. *Nucleic Acids Res,* 40(6):e45.

CASTRO, C., SMIDANSKY, E. D., ARNOLD, J. J., MAKSIMCHUK, K. R., MOUSTAFA, I., UCHIDA, A., GOTTE, M., KONIGSBERG, W. & CAMERON, C. E. 2009. Nucleic acid polymerases use a general acid for nucleotidyl transfer. *Nat Struct Mol Biol,* 16**,** 212-8.

CHAMBERLIN, M., MCGRATH, J. & WASKELL, L. 1970. New RNA polymerase from *Escherichia coli* infected with bacteriophage T7. *Nature,* 228**,** 227-31.

CHAMBERLIN, M. & RING, J. 1973. Characterization of T7-specific ribonucleic acid polymerase. 1. General properties of the enzymatic reaction and the template specificity of the enzyme. *J Biol Chem,* 248**,** 2235-44.

CHAN, C. L. & LANDICK, R. 1989. The *Salmonella typhimurium his* operon leader region contains an RNA hairpin-dependent transcription pause site. Mechanistic implications of the effect on pausing of altered RNA hairpins. *J Biol Chem,* 264**,** 20796-804.

CHRISTENSEN-DALSGAARD, M., OVERGAARD, M., WINTHER, K. S. & GERDES, K. 2008. RNA decay by messenger RNA interferases. *Methods Enzymol,* 447**,** 521-35.

CHRISTENSEN, S. K., MIKKELSEN, M., PEDERSEN, K. & GERDES, K. 2001. RelE, a global inhibitor of translation, is activated during nutritional stress. *Proc Natl Acad Sci U S A,* 98**,** 14328-33.

CIAMPI, M. S. 2006. Rho-dependent terminators and transcription termination. *Microbiology,* 152**,** 2515-28.

CLEMONS, W. M., JR., MAY, J. L., WIMBERLY, B. T., MCCUTCHEON, J. P., CAPEL, M. S. & RAMAKRISHNAN, V. 1999. Structure of a bacterial 30S ribosomal subunit at 5.5 A resolution. *Nature,* 400**,** 833-40.

COULOMBE, B. & BURTON, Z. F. 1999. DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms. *Microbiol Mol Biol Rev***,** 457-78.

CRAIG, M. L., TSODIKOV, O. V., MCQUADE, K. L., SCHLAX, P. E., JR., CAPP, M. W., SAECKER, R. M. & RECORD, M. T., JR. 1998. DNA footprints of the two kinetically significant intermediates in formation of an RNA polymerase-promoter open complex: evidence that interactions with start site and downstream DNA induce sequential conformational changes in polymerase and DNA. *J Mol Biol,* 283**,** 741-56.

CRAMER, P. & ARNOLD, E. 2009. Proteins: how RNA polymerases work. *Curr Opin Struct Biol,* 19**,** 680-2.

CRAMER, P., BUSHNELL, D. A. & KORNBERG, R. D. 2001. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science,* 292**,** 1863-76.

D'AUBENTON CARAFA, Y., BRODY, E. & THERMES, C. 1990. Prediction of rho-independent *Escherichia coli* transcription terminators. A statistical analysis of their RNA stem-loop structures. *J Mol Biol,* 216**,** 835-58.

DAS, A. 1993. Control of transcription termination by RNA-binding proteins. *Annu Rev Biochem,* 62**,** 893-930.

DAUBE, S. S. & VON HIPPEL, P. H. 1992. Functional transcription elongation complexes from synthetic RNA-DNA bubble duplexes. *Science,* 258**,** 1320-4.

DEHASETH, P. L. & HELMANN, J. D. 1995. Open complex formation by *Escherichia coli* RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA. *Mol Microbiol,* 16**,** 817-24.

DENG, S., STEIN, R. A. & HIGGINS, N. P. 2005. Organization of supercoil domains and their reorganization by transcription. *Molecular microbiology,* 57**,** 1511-21.

DONAHUE, J. P. & TURNBOUGH, C. L., JR. 1994. Nucleotide-specific transcriptional pausing in the pyrBI leader region of *Escherichia coli* K-12. *J Biol Chem,* 269**,** 18185-91.

DUTTA, D., SHATALIN, K., EPSHTEIN, V., GOTTESMAN, M. E. & NUDLER, E. 2011. Linking RNA polymerase backtracking to genome instability in *E. coli*. *Cell,* 146**,** 533-43.

EPSHTEIN, V., CARDINALE, C. J., RUCKENSTEIN, A. E., BORUKHOV, S. & NUDLER, E. 2007. An allosteric path to transcription termination. *Mol Cell,* 28**,** 991-1001.

EPSHTEIN, V., DUTTA, D., WADE, J. & NUDLER, E. 2010. An allosteric mechanism of Rho-dependent transcription termination. *Nature,* 463**,** 245-9.

EPSHTEIN, V. & NUDLER, E. 2003. Cooperation between RNA polymerase molecules in transcription elongation. *Science,* 300**,** 801-5.

EPSHTEIN, V., TOULME, F., RAHMOUNI, A. R., BORUKHOV, S. & NUDLER, E. 2003. Transcription through the roadblocks: the role of RNA polymerase cooperation. *EMBO J,* 22**,** 4719-27.

ERIE, D. A., YAGER, T. D. & VON HIPPEL, P. H. 1992. The single-nucleotide addition cycle in transcription: a biophysical and biochemical perspective. *Annu Rev Biophys Biomol Struct,* 21**,** 379-415.

FENTON, M. S. & GRALLA, J. D. 2001. Function of the bacterial TATAAT -10 element as single-stranded DNA during RNA polymerase isomerization. *Proc Natl Acad Sci U S A,* 98**,** 9020-5.

FREDRICK, K. & NOLLER, H. F. 2002. Accurate translocation of mRNA by the ribosome requires a peptidyl group or its analog on the tRNA moving into the 30S P site. *Mol Cell,* 9**,** 1125-31.

FRENCH, S. 1992. Consequences of replication fork movement through transcription units *in vivo*. *Science,* 258**,** 1362-5.

FRIEDMAN, D. I., BAUMANN, M. & BARON, L. S. 1976. Cooperative effects of bacterial mutations affecting lambda N gene expression. I. Isolation and characterization of a nusB mutant. *Virology,* 73**,** 119-27.

GARRETT, R. A. 1987. *The ribosome,* Washington D.C., American Society for Microbiology

GHOSH, T., BOSE, D. & ZHANG, X. 2010. Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiol Rev,* 34**,** 611-27.

GNATT, A. L., CRAMER, P., FU, J., BUSHNELL, D. A. & KORNBERG, R. D. 2001. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science,* 292**,** 1876-82.

GOLOMB, M. & CHAMBERLIN, M. 1974. Characterization of T7-specific ribonucleic acid polymerase. IV. Resolution of the major *in vitro* transcripts by gel electrophoresis. *J Biol Chem,* 249**,** 2858-63.

GONG, F., ITO, K., NAKAMURA, Y. & YANOFSKY, C. 2001. The mechanism of tryptophan induction of tryptophanase operon expression: tryptophan inhibits release factor-mediated cleavage of TnaC-peptidyl-tRNA(Pro). *Proc Natl Acad Sci U S A,* 98**,** 8997-9001.

GONG, F. & YANOFSKY, C. 2003. A transcriptional pause synchronizes translation with transcription in the tryptophanase operon leader region. *J Bacteriol,* 185**,** 6472-6.

GOTTA, S. L., MILLER, O. L., JR. & FRENCH, S. L. 1991. rRNA transcription rate in *Escherichia coli*. *J Bacteriol,* 173**,** 6647-9.

GOURSE, R. L., GAAL, T., BARTLETT, M. S., APPLEMAN, J. A. & ROSS, W. 1996. rRNA transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Annu Rev Microbiol,* 50**,** 645-77.

GRAINGER, D. C., HURD, D., GOLDBERG, M. D. & BUSBY, S. J. 2006. Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res,* 34**,** 4642-52.

GREEN, R. & NOLLER, H. F. 1997. Ribosomes and translation. *Annu Rev Biochem,* 66**,** 679-716.

GREIVE, S. J. & VON HIPPEL, P. H. 2005. Thinking quantitatively about transcriptional regulation. *Nat Rev Mol Cell Biol,* 6**,** 221-32.

GRUBER, T. M. & GROSS, C. A. 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol,* 57**,** 441-66.

GRUNBERG, S., REICH, C., ZELLER, M. E., BARTLETT, M. S. & THOMM, M. 2010. Rearrangement of the RNA polymerase subunit H and the lower jaw in archaeal elongation complexes. *Nucleic Acids Res,* 38**,** 1950-63.

GUAJARDO, R. & SOUSA, R. 1997. A model for the mechanism of polymerase translocation. *J Mol Biol,* 265**,** 8-19.

GUSAROV, I. & NUDLER, E. 1999. The mechanism of intrinsic transcription termination. *Mol Cell,* 3**,** 495-504.

GUYER, M.S., R.E. REED, T. STEITZ & K.B. LOW 1981. *Spr.Harb.Symp.Quant.Biol.* 45:135-140

HA, K. S., TOULOKHONOV, I., VASSYLYEV, D. G. & LANDICK, R. 2010. The NusA N-Terminal Domain Is Necessary and Sufficient for Enhancement of Transcriptional Pausing via Interaction with the RNA Exit Channel of RNA Polymerase. *J Mol Biol*. 5,708-25

HARRIS, J. K., KELLEY, S. T., SPIEGELMAN, G. B. & PACE, N. R. 2003. The genetic core of the universal ancestor. *Genome Res,* 13**,** 407-12.

HATOUM, A. & ROBERTS, J. 2008. Prevalence of RNA polymerase stalling at *Escherichia coli* promoters after open complex formation. *Mol Microbiol,* 68**,** 17-28.

HELD, W. A., BALLOU, B., MIZUSHIMA, S. & NOMURA, M. 1974. Assembly mapping of 30 S ribosomal proteins from *Escherichia coli*. Further studies. *J Biol Chem,* 249**,** 3103-11.

HERBERT, K. M., GREENLEAF, W. J. & BLOCK, S. M. 2008. Single-molecule studies of RNA polymerase: motoring along. *Annu Rev Biochem,* 77**,** 149-76.

HERBERT, K. M., LA PORTA, A., WONG, B. J., MOONEY, R. A., NEUMAN, K. C., LANDICK, R. & BLOCK, S. M. 2006. Sequence-resolved detection of pausing by single RNA polymerase molecules. *Cell,* 125**,** 1083-94.

HERBERT, K. M., ZHOU, J., MOONEY, R. A., PORTA, A. L., LANDICK, R. & BLOCK, S. M. 2010. *E. coli* NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. *J Mol Biol,* 399**,** 17-30.

HOGAN, B. P., HARTSCH, T. & ERIE, D. A. 2002. Transcript cleavage by *Thermus thermophilus* RNA polymerase. Effects of GreA and anti-GreA factors. *J Biol Chem,* 277**,** 967-75.

HOLSCHUH, K., BONIN, J. & GASSEN, H. G. 1980. Mechanism of translocation: effect of cognate transfer ribonucleic acids on the binding of AUGUn to 70S ribosomes. *Biochemistry,* 19**,** 5857-64.

HRAIKY, C., RAYMOND, M. A. & DROLET, M. 2000. RNase H overproduction corrects a defect at the level of transcription elongation during rRNA synthesis in the absence of DNA topoisomerase I in *Escherichia coli*. *J Biol Chem,* 275**,** 11257-63.

HUVET, M., NICOLAY, S., TOUCHON, M., AUDIT, B., D'AUBENTON-CARAFA, Y., ARNEODO, A. & THERMES, C. 2007. Human gene organization driven by the coordination of replication and transcription. *Genome Res,* 17**,** 1278-85.

JIN, J., BAI, L., JOHNSON, D. S., FULBRIGHT, R. M., KIREEVA, M. L., KASHLEV, M. & WANG, M. D. 2010. Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. *Nat Struct Mol Biol,* 17**,** 745-52.

KAPLAN, C. D., LARSSON, K. M. & KORNBERG, R. D. 2008. The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. *Mol Cell,* 30**,** 547-56.

KASSAVETIS, G. A., PRAKASH, P. & SHIM, E. 2010. The C53/C37 subcomplex of RNA polymerase III lies near the active site and participates in promoter opening. *J Biol Chem,* 285**,** 2695-706.

KAZIRO, Y. 1978. The role of guanosine 5'-triphosphate in polypeptide chain elongation. *Biochim Biophys Acta,* 505**,** 95-127.

KETTENBERGER, H., ARMACHE, K. J. & CRAMER, P. 2004. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol Cell,* 16**,** 955-65.

KIREEVA, M., NEDIALKOV, Y. A., GONG, X. Q., ZHANG, C., XIONG, Y., MOON, W., BURTON, Z. F. & KASHLEV, M. 2009. Millisecond phase kinetic analysis of elongation catalyzed by human, yeast, and *Escherichia coli* RNA polymerase. *Methods,* 48**,** 333-45.

KIREEVA, M. L. & KASHLEV, M. 2009a. Mechanism of sequence-specific pausing of bacterial RNA polymerase. *Proc Natl Acad Sci U S A,* 106**,** 8900-5.

KIREEVA, M. L. & KASHLEV, M. 2009b. Mechanism of sequence-specific pausing of bacterial RNA polymerase. *Proc Natl Acad Sci U S A,* 106**,** 8900-5.

KITAGAWA, M, TAKESHI, A, MOHAMMAD, A, TOMOKO, I, EIJI INAMOTO & HIROTADA, M. 2005. Complete set of ORF clones of *Escherichia coli* ASKA library (A Complete Set of E. coli K-12 ORF Archive): Unique Resources for Biological Research. *DNA Res* 12 (5): 291-299.

KOMISSAROVA, N. & KASHLEV, M. 1997. RNA polymerase switches between inactivated and activated states By translocating back and forth along the DNA and the RNA. *J Biol Chem,* 272**,** 15329-38.

KORZHEVA, N. & MUSTAEV, A. 2001. Transcription elongation complex: structure and function. *Curr Opin Microbiol,* 4**,** 119-25.

KORZHEVA, N., MUSTAEV, A., KOZLOV, M., MALHOTRA, A., NIKIFOROV, V., GOLDFARB, A. & DARST, S. A. 2000. A structural model of transcription elongation. *Science,* 289**,** 619-25.

KORZHEVA, N., MUSTAEV, A., NUDLER, E., NIKIFOROV, V. & GOLDFARB, A. 1998. Mechanistic model of the elongation complex of *Escherichia coli* RNA polymerase. *Cold Spring Harb Symp Quant Biol,* 63**,** 337-45.

KUZNEDELOV, K., KORZHEVA, N., MUSTAEV, A. & SEVERINOV, K. 2002. Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. *Embo J,* 21**,** 1369-78.

LANDICK, R. 1997. RNA polymerase slides home: pause and termination site recognition. *Cell,* 88**,** 741-4.

LANDICK, R. 2006. The regulatory roles and mechanism of transcriptional pausing. *Biochem Soc Trans,* 34**,** 1062-6.

LANDICK, R. 2009. Transcriptional pausing without backtracking. *Proc Natl Acad Sci U S A,* 106**,** 8797-8.

LANDICK, R., CAREY, J. & YANOFSKY, C. 1985. Translation activates the paused transcription complex and restores transcription of the *trp* operon leader region. *Proc Natl Acad Sci U S A,* 82**,** 4663-7.

LANDICK, R., CAREY, J. & YANOFSKY, C. 1987. Detection of transcription-pausing *in vivo* in the *trp* operon leader region. *Proc Natl Acad Sci U S A,* 84**,** 1507-11.

LARSON, M. H., GREENLEAF, W. J., LANDICK, R. & BLOCK, S. M. 2008. Applied force reveals mechanistic and energetic details of transcription termination. *Cell,* 132**,** 971-82.

LAURSEN, B. S., SORENSEN, H. P., MORTENSEN, K. K. & SPERLING-PETERSEN, H. U. 2005. Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev,* 69**,** 101-23.

LESNIK, E. A., SAMPATH, R., LEVENE, H. B., HENDERSON, T. J., MCNEIL, J. A. & ECKER, D. J. 2001. Prediction of rho-independent transcriptional terminators in *Escherichia coli*. *Nucleic Acids Res,* 29**,** 3583-94.

LI, X. Y. & MCCLURE, W. R. 1998. Stimulation of open complex formation by nicks and apurinic sites suggests a role for nucleation of DNA melting in *Escherichia coli* promoter function. *J Biol Chem,* 273**,** 23558-66.

MALHOTRA, A., SEVERINOVA, E. & DARST, S. A. 1996. Crystal structure of a sigma 70 subunit fragment from *E. coli* RNA polymerase. *Cell,* 87**,** 127-36.

MARKOVTSOV, V., MUSTAEV, A. & GOLDFARB, A. 1996. Protein-RNA interactions in the active center of transcription elongation complex. *Proc Natl Acad Sci U S A,* 93**,** 3221-6.

MASON, S. W. & GREENBLATT, J. 1991. Assembly of transcription elongation complexes containing the N protein of phage lambda and the *Escherichia coli* elongation factors NusA, NusB, NusG, and S10. *Genes Dev,* 5**,** 1504-12.

MASSÉ, E. & DROLET, M. 1999. *Escherichia coli* DNA topoisomerase I inhibits R-loop formation by relaxing transcription-induced negative supercoiling. *J Biol Chem,* 274**,** 16659-64.

MCDOWELL, J. C., ROBERTS, J. W., JIN, D. J. & GROSS, C. 1994. Determination of intrinsic transcription termination efficiency by RNA polymerase elongation rate. *Science,* 266**,** 822-5.

MCLEAN, M. J., WOLFE, K. H. & DEVINE, K. M. 1998. Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes. *J Mol Evol,* 47**,** 691-6.

MEKLER, V., KORTKHONJIA, E., MUKHOPADHYAY, J., KNIGHT, J., REVYAKIN, A., KAPANIDIS, A. N., NIU, W., EBRIGHT, Y. W., LEVY, R. & EBRIGHT, R. H. 2002. Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell,* 108**,** 599-614.

MESELSON M. & YUAN R. 1968. DNA restriction enzyme from *E. coli*. *Nature* 217:1110

MERRICK, M. J. 1993. In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N). *Mol Microbiol,* 10**,** 903-9.

MIRKIN, E. V., CASTRO ROA, D., NUDLER, E. & MIRKIN, S. M. 2006. Transcription regulatory elements are punctuation marks for DNA replication. *Proc Natl Acad Sci U S A,* 103**,** 7276-81.

MIRKIN, E. V. & MIRKIN, S. M. 2005. Mechanisms of transcription-replication collisions in bacteria. *Mol Cell Biol,* 25**,** 888-95.

MOAZED D, N. H. 1989. Interaction of tRNA with 23S rRNA in the ribosomal A, P and E sites. *Cell,* 57**,** 585-597.

MOAZED, D. & NOLLER, H. F. 1989. Intermediate states in the movement of transfer RNA in the ribosome. *Nature,* 342**,** 142-8.

MODRICH, P. *Escherichia coli* mutS-encoded protein binds to mismatched DNA base pairs. 1986. *Proc Natl Acad Sci U S A* 83,178-184

MOGRIDGE, J. & GREENBLATT, J. 1998. Specific binding of *Escherichia coli* ribosomal protein S1 to boxA transcriptional antiterminator RNA. *J Bacteriol,* 180**,** 2248-52.

MOONEY, R. Sigma and RNA polymerase: An on-again, off-again relationship? 2005 *Mol. Cell*. 3, 20, 335 (Review)

MOUGEL, M., EYERMANN, F., WESTHOF, E., ROMBY, P., EXPERT-BEZANCON, A., EBEL, J. P., EHRESMANN, B. & EHRESMANN, C. 1987. Binding of *Escherichia coli* ribosomal protein S8 to 16 S rRNA. A model for the interaction and the tertiary structure of the RNA binding site. *J Mol Biol,* 198**,** 91-107.

MURAKAMI, K. S. & DARST, S. A. 2003. Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol,* 13**,** 31-9.

MURAKAMI, K. S., MASUDA, S. & DARST, S. A. 2002. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. *Science,* 296**,** 1280-4.

NARYSHKINA, T., KUZNEDELOV, K. & SEVERINOV, K. 2006. The role of the largest RNA polymerase subunit lid element in preventing the formation of extended RNA-DNA hybrid. *J Mol Biol,* 361**,** 634-43.

NEHRKE, K. W., ZALATAN, F. & PLATT, T. 1993. NusG alters rho-dependent termination of transcription *in vitro* independent of kinetic coupling. *Gene Expr,* 3**,** 119-33.

NEUMAN, K. C., ABBONDANZIERI, E. A., LANDICK, R., GELLES, J. & BLOCK, S. M. 2003. Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell,* 115**,** 437-47.

NISSEN, P., HANSEN, J., BAN, N., MOORE, P. B. & STEITZ, T. A. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science,* 289**,** 920-30.

NUDLER, E., AVETISSOVA, E., MARKOVTSOV, V. & GOLDFARB, A. 1996. Transcription processivity: protein-DNA interactions holding together the elongation complex. *Science,* 273**,** 211-7.

NUDLER, E. & GOTTESMAN, M. E. 2002. Transcription termination and anti-termination in *E. coli*. *Genes Cells,* 7**,** 755-68.

NUDLER, E., GUSAROV, I., AVETISSOVA, E., KOZLOV, M. & GOLDFARB, A. 1998. Spatial organization of transcription elongation complex in *Escherichia coli*. *Science,* 281**,** 424-8.

NUDLER, E., GUSAROV, I. & BAR-NAHUM, G. 2003. Methods of walking with the RNA polymerase. *Methods Enzymol,* 371**,** 160-9.

NUDLER, E., MUSTAEV, A., LUKHTANOV, E. & GOLDFARB, A. 1997. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell,* 89**,** 33-41.

OLINS, P. O. & NOMURA, M. 1981. Regulation of the S10 ribosomal protein operon in *E. coli*: nucleotide sequence at the start of the operon. *Cell,* 26**,** 205-11.

OMONT, N. & KEPES, F. 2004. Transcription/replication collisions cause bacterial transcription units to be longer on the leading strand of replication. *Bioinformatics,* 20**,** 2719-25.

ORLOVA, M., NEWLANDS, J., DAS, A., GOLDFARB, A. & BORUKHOV, S. 1995. Intrinsic transcript cleavage activity of RNA polymerase. *Proc Natl Acad Sci U S A,* 92**,** 4596-600.

OVERGAARD, M., BORCH, J., JORGENSEN, M. G. & GERDES, K. 2008. Messenger RNA interferase RelE controls *relBE* transcription by conditional cooperativity. *Mol Microbiol,* 69**,** 841-57.

PARK, C. S., WU, F. Y. & WU, C. W. 1982. Molecular mechanism of promoter selection in gene transcription. II. Kinetic evidence for promoter search by a one-dimensional diffusion of RNA polymerase molecule along the DNA template. *J Biol Chem,* 257**,** 6950-6.

PAVCO, P. A. & STEEGE, D. A. 1990. Elongation by *Escherichia coli* RNA polymerase is blocked *in vitro* by a site-specific DNA binding protein. *J Biol Chem,* 265**,** 9960-9.

PEDERSEN, K., ZAVIALOV, A. V., PAVLOV, M. Y., ELF, J., GERDES, K. & EHRENBERG, M. 2003. The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell,* 112**,** 131-40.

PERDUE, S. A. & ROBERTS, J. W. 2011. Sigma(70)-dependent transcription pausing in *Escherichia coli*. *J Mol Biol,* 412**,** 782-92.

POMERANTZ, R. T. & O'DONNELL, M. 2008. The replisome uses mRNA as a primer after colliding with RNA polymerase. *Nature,* 456**,** 762-6.

POMERANTZ, R. T. & O'DONNELL, M. 2010. Direct restart of a replication fork stalled by a head-on RNA polymerase. *Science,* 327**,** 590-2.

PON, C. L., PACI, M., PAWLIK, R. T. & GUALERZI, C. O. 1985. Structure-function relationship in *Escherichia coli* initiation factors. Biochemical and biophysical characterization of the interaction between IF-2 and guanosine nucleotides. *J Biol Chem,* 260**,** 8918-24.

PRICE, M. N., ALM, E. J. & ARKIN, A. P. 2005. Interruptions in gene expression drive highly expressed operons to the leading strand of DNA replication. *Nucleic Acids Res,* 33**,** 3224-34.

PROSHKIN, S., RAHMOUNI, A. R., MIRONOV, A. & NUDLER, E. 2010. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science,* 328**,** 504-8.

REES, W. A., KELLER, R. W., VESENKA, J. P., YANG, G. & BUSTAMANTE, C. 1993. Evidence of DNA bending in transcription complexes imaged by scanning force microscopy. *Science,* 260**,** 1646-9.

REPPAS, N. B., WADE, J. T., CHURCH, G. M. & STRUHL, K. 2006. The transition between transcriptional initiation and elongation in *E. coli* is highly variable and often rate limiting. *Mol Cell,* 24**,** 747-57.

REVYAKIN, A., LIU, C., EBRIGHT, R. H. & STRICK, T. R. 2006. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science,* 314**,** 1139-43.

RICHARDSON, L. V. & RICHARDSON, J. P. 1996. Rho-dependent termination of transcription is governed primarily by the upstream Rho utilization (rut) sequences of a terminator. *J Biol Chem,* 271**,** 21597-603.

RING, B. Z., YARNELL, W. S. & ROBERTS, J. W. 1996. Function of *E. coli* RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell,* 86**,** 485-93.

RIVETTI, C., CODELUPPI, S., DIECI, G. & BUSTAMANTE, C. 2003. Visualizing RNA extrusion and DNA wrapping in transcription elongation complexes of bacterial and eukaryotic RNA polymerases. *J. Mol. Biol,* 326**,** 1413-26.

ROBERTS, J. W., YARNELL, W., BARTLETT, E., GUO, J., MARR, M., KO, D. C., SUN, H. & ROBERTS, C. W. 1998. Antitermination by bacteriophage lambda Q protein. *Cold Spring Harb Symp Quant Biol,* 63**,** 319-25.

ROCHA, E. P. & DANCHIN, A. 2003. Essentiality, not expressiveness, drives gene-strand bias in bacteria. *Nat Genet,* 34**,** 377-8.

ROGHANIAN, M., YUZENKOVA, Y. & ZENKIN, N. 2011a. Controlled interplay between trigger loop and Gre factor in the RNA polymerase active centre. *Nucleic Acids Res,* 39, 4352-9.

ROGOZINA, A., ZAYCHIKOV, E., BUCKLE, M., HEUMANN, H. & SCLAVI, B. 2009. DNA melting by RNA polymerase at the T7A1 promoter precedes the rate-limiting step at 37 degrees C and results in the accumulation of an off-pathway intermediate. *Nucleic Acids Res,* 37**,** 5390-404.

SAKATA-SOGAWA, K. & SHIMAMOTO, N. 2004. RNA polymerase can track a DNA groove during promoter search. *Proc Natl Acad Sci U S A,* 101**,** 14731-5.

SANTANGELO, T. J. & ROBERTS, J. W. 2004. Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Mol Cell,* 14**,** 117-26.

SAXENA, S. & GOWRISHANKAR, J. 2011. Modulation of Rho-dependent transcription termination in *Escherichia coli* by the H-NS family of proteins. *J Bacteriol,* 193**,** 3832-41.

SERGIEV, P. V., LESNYAK, D. V., KIPARISOV, S. V., BURAKOVSKY, D. E., LEONOV, A. A., BOGDANOV, A. A., BRIMACOMBE, R. & DONTSOVA,

O. A. 2005. Function of the ribosomal E-site: a mutagenesis study. *Nucleic Acids Res,* 33**,** 6048-56.

SEVERINOV, K., FENYO, D., SEVERINOVA, E., MUSTAEV, A., CHAIT, B. T., GOLDFARB, A. & DARST, S. A. 1994. The sigma subunit conserved region 3 is part of "5'-face" of active center of *Escherichia coli* RNA polymerase. *J Biol Chem,* 269**,** 20826-8.

SHA, Y., LINDAHL, L. & ZENGEL, J. M. 1995. RNA determinants required for L4-mediated attenuation control of the S10 r-protein operon of *Escherichia coli*. *J Mol Biol,* 245**,** 486-98.

SHAEVITZ, J. W., ABBONDANZIERI, E. A., LANDICK, R. & BLOCK, S. M. 2003. Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. *Nature,* 426**,** 684-7.

SIDORENKOV, I., KOMISSAROVA, N. & KASHLEV, M. 1998. Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Mol Cell,* 2**,** 55-64.

SOSUNOV, V., ZOROV, S., SOSUNOVA, E., NIKOLAEV, A., ZAKEYEVA, I., BASS, I., GOLDFARB, A., NIKIFOROV, V., SEVERINOV, K. & MUSTAEV, A. 2005. The involvement of the aspartate triad of the active center in all catalytic activities of multisubunit RNA polymerase. *Nucleic Acids Res,* 33**,** 4202-11.

SPIRIN, A. S. 1985. Ribosomal translocation: facts and models. *Prog Nucleic Acid Res Mol Biol,* 32**,** 75-114.

SRIVATSAN, A., TEHRANCHI, A., MACALPINE, D. M. & WANG, J. D. 2010. Co-orientation of replication and transcription preserves genome integrity. *PLoS Genet,* 6**,** e1000810.

STEWART, V., LANDICK, R. & YANOFSKY, C. 1986. Rho-dependent transcription termination in the tryptophanase operon leader region of *Escherichia coli* K-12. *J Bacteriol,* 166**,** 217-23.

STUDIER, FW. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*, 5, 113-30.

SYDOW, J. F. & CRAMER, P. 2009. RNA polymerase fidelity and transcriptional proofreading. *Curr Opin Struct Biol,* 19**,** 732-9.

TADIGOTLA, V. R., D, O. M., SENGUPTA, A. M., EPSHTEIN, V., EBRIGHT, R. H., NUDLER, E. & RUCKENSTEIN, A. E. 2006. Thermodynamic and kinetic modeling of transcriptional pausing. *Proc Natl Acad Sci U S A,* 103**,** 4439-44.

TAKYAR, S., HICKERSON, R. P. & NOLLER, H. F. 2005a. mRNA helicase activity of the ribosome. *Cell,* 120**,** 49-58.

TEMPERLEY, R., RICHTER, R., DENNERLEIN, S., LIGHTOWLERS, R. N. & CHRZANOWSKA-LIGHTOWLERS, Z. M. 2010. Hungry codons promote frameshifting in human mitochondrial ribosomes. *Science,* 327**,** 301.

TOULME, F., MOSRIN-HUAMAN, C., ARTSIMOVITCH, I. & RAHMOUNI, A. R. 2005. Transcriptional pausing *in vivo*: a nascent RNA hairpin restricts lateral movements of RNA polymerase in both forward and reverse directions. *J Mol Biol,* 351**,** 39-51.

TOULOKHONOV, I., ARTSIMOVITCH, I. & LANDICK, R. 2001. Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science,* 292**,** 730-3.

TOULOKHONOV, I. & LANDICK, R. 2003. The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. *Mol Cell,* 12**,** 1125-36.

TOULOKHONOV, I. & LANDICK, R. 2006. The role of the lid element in transcription by *E. coli* RNA polymerase. *J Mol Biol,* 361**,** 644-58.

TOULOKHONOV, I., ZHANG, J., PALANGAT, M. & LANDICK, R. 2007. A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol Cell,* 27**,** 406-19.

TRAVERS, A. A. 2004. The structural basis of DNA flexibility. *Philos Transact A Math Phys Eng Sci,* 362**,** 1423-38.

UPTAIN, S. M., KANE, C. M. & CHAMBERLIN, M. J. 1997. Basic mechanisms of transcript elongation and its regulation. *Annu Rev Biochem,* 66**,** 117-72.

VASSYLYEV, D. G., SEKINE, S., LAPTENKO, O., LEE, J., VASSYLYEVA, M. N., BORUKHOV, S. & YOKOYAMA, S. 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature,* 417**,** 712-9.

VASSYLYEV, D. G., VASSYLYEVA, M. N., PEREDERINA, A., TAHIROV, T. H. & ARTSIMOVITCH, I. 2007a. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature,* 448**,** 157-62.

VASSYLYEV, D. G., VASSYLYEVA, M. N., ZHANG, J., PALANGAT, M., ARTSIMOVITCH, I. & LANDICK, R. 2007b. Structural basis for substrate loading in bacterial RNA polymerase. *Nature,* 448**,** 163-8.

VON HIPPEL, P. H. 2007. From "simple" DNA-protein interactions to the macromolecular machines of gene expression. *Annu Rev Biophys Biomol Struct,* 36**,** 79-105.

VON HIPPEL, P. H., BEAR, D. G., MORGAN, W. D. & MCSWIGGEN, J. A. 1984. Protein-nucleic acid interactions in transcription: A molecular analysis. *Ann. Rev. Biochem.,* 53**,** 389-446.

WALDBURGER, C. & SUSSKIND, M. M. 1994. Probing the informational content of *Escherichia coli* sigma 70 region 2.3 by combinatorial cassette mutagenesis. *J Mol Biol,* 235**,** 1489-500.

WALKER, S. E. & FREDRICK, K. 2008. Preparation and evaluation of acylated tRNAs. *Methods,* 44**,** 81-6.

WALSTROM, K. M., DOZONO, J. M., ROBIC, S. & VON HIPPEL, P. H. 1997. Kinetics of the RNA-DNA helicase activity of *Escherichia coli* transcription termination factor rho. 1. Characterization and analysis of the reaction. *Biochemistry,* 36**,** 7980-92.

WANG, D., BUSHNELL, D. A., HUANG, X., WESTOVER, K. D., LEVITT, M. & KORNBERG, R. D. 2009. Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. *Science,* 324**,** 1203-6.

WANG, D., BUSHNELL, D. A., WESTOVER, K. D., KAPLAN, C. D. & KORNBERG, R. D. 2006. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell,* 127**,** 941-54.

WANG, M. D., SCHNITZER, M. J., YIN, H., LANDICK, R., GELLES, J. & BLOCK, S. M. 1998. Force and velocity measured for single molecules of RNA polymerase. *Science,* 282**,** 902-6.

WANG D., MEIER T., CHAN C., FENG G., LEE D. & LANDICK R. 1995. Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex. *Cell* 81:341–350.

WEIXLBAUMER, A., JIN, H., NEUBAUER, C., VOORHEES, R. M., PETRY, S., KELLEY, A. C. & RAMAKRISHNAN, V. 2008. Insights into translational termination from the structure of RF2 bound to the ribosome. *Science,* 322**,** 953-6.

WEHR, C.T. 1973 Isolation and properties of a ribonuclease-deficient mutant of *Salmonella typhimurium*. *J Bacteriol*, 114:96-102.

WESTOVER, K. D., BUSHNELL, D. A. & KORNBERG, R. D. 2004. Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science,* 303**,** 1014-6.

WILSON, C. & DOMBROSKI, A. J. 1997. Region 1 of sigma70 is required for efficient isomerization and initiation of transcription by *Escherichia coli* RNA polymerase. *J Mol Biol,* 267**,** 60-74.

WINTHER, K. S. & GERDES, K. 2011. Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proc Natl Acad Sci U S A,* 108**,** 7403-7.

YAKHNIN, A. V. & BABITZKE, P. 2010. Mechanism of NusG-stimulated pausing, hairpin-dependent pause site selection and intrinsic termination at overlapping pause and termination sites in the Bacillus subtilis *trp* leader. *Mol Microbiol,* 76**,** 690-705.

YAMAGUCHI, Y., PARK, J. H. & INOUYE, M. 2011. Toxin-antitoxin systems in bacteria and archaea. *Annu Rev Genet,* 45**,** 61-79.

YAMAMOTO, T. A., GERDES, K. & TUNNACLIFFE, A. 2002. Bacterial toxin RelE induces apoptosis in human cells. *FEBS Lett,* 519**,** 191-4.

YANOFSKY, C. 1981. Attenuation in the control of expression of bacterial operons. *Nature,* 289**,** 751-8.

YARNELL, W. S. & ROBERTS, J. W. 1999. Mechanism of intrinsic transcription termination and antitermination. *Science,* 284**,** 611-5.

YUSUPOV, M. M., YUSUPOVA, G. Z., BAUCOM, A., LIEBERMAN, K., EARNEST, T. N., CATE, J. H. & NOLLER, H. F. 2001. Crystal structure of the ribosome at 5.5 A resolution. *Science,* 292**,** 883-96.

YUSUPOVA, G. Z., YUSUPOV, M. M., CATE, J. H. & NOLLER, H. F. 2001. The path of messenger RNA through the ribosome. *Cell,* 106**,** 233-41.

YUZENKOVA, Y., BOCHKAREVA, A., TADIGOTLA, V. R., ROGHANIAN, M., ZOROV, S., SEVERINOV, K. & ZENKIN, N. 2010. Stepwise mechanism for transcription fidelity. *BMC Biol,* 8**,** 54.

YUZENKOVA, Y. & ZENKIN, N. 2010a. Central role of the RNA polymerase trigger loop in intrinsic RNA hydrolysis. *Proc Natl Acad Sci U S A,* 107**,** 10878-83.

ZAHER, H. S. & GREEN, R. 2009. Quality control by the ribosome following peptide bond formation. *Nature,* 457**,** 161-6.

ZAHER, H. S. & GREEN, R. 2010. Hyperaccurate and error-prone ribosomes exploit distinct mechanisms during tRNA selection. *Mol Cell,* 39**,** 110-20.

ZENGEL, J. M., MUECKL, D. & LINDAHL, L. 1980. Protein L4 of the *E. coli* ribosome regulates an eleven gene r protein operon. *Cell,* 21**,** 523-35.

ZENKIN, N., KULBACHINSKIY, A., YUZENKOVA, Y., MUSTAEV, A., BASS, I., SEVERINOV, K. & BRODOLIN, K. 2007. Region 1.2 of the RNA polymerase sigma subunit controls recognition of the -10 promoter element. *Embo J,* 26, 955-64

ZENKIN, N., NARYSHKINA, T., KUZNEDELOV, K. & SEVERINOV, K. 2006a. The mechanism of DNA replication primer synthesis by RNA polymerase. *Nature,* 439**,** 617-20.

ZENKIN, N. & SEVERINOV, K. 2004. The role of RNA polymerase sigma subunit in promoter-independent initiation of transcription. *Proc Natl Acad Sci U S A,* 101**,** 4396-400.

ZENKIN, N. & SEVERINOV, K. 2008. RNA polymerase -the third class of primases. *Cell Mol Life Sci,* 65**,** 2280-8.

ZENKIN, N., YUZENKOVA, Y. & SEVERINOV, K. 2006b. Transcript-assisted transcriptional proofreading. *Science,* 313**,** 518-20.

ZHANG, G., CAMPBELL, E. A., MINAKHIN, L., RICHTER, C., SEVERINOV, K. & DARST, S. A. 1999. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution. *Cell,* 98**,** 811-24.

ZHANG, J., PALANGAT, M. & LANDICK, R. 2010. Role of the RNA polymerase trigger loop in catalysis and pausing. *Nat Struct Mol Biol,* 17**,** 99-104.

ZHANG, Z., KLATT, A., GILMOUR, D. S. & HENDERSON, A. J. 2007. Negative elongation factor NELF represses human immunodeficiency virus transcription by pausing the RNA polymerase II complex. *J Biol Chem,* 282**,** 16981-8.

ZHILINA, E., ESYUNINA, D., BRODOLIN, K. & KULBACHINSKIY, A. 2011. Structural transitions in the transcription elongation complexes of bacterial RNA polymerase during sigma-dependent pausing. *Nucleic Acids Res* 40, 3078-91.