Control of Gene Expression Through Coupling of Transcription and Translation

Flint Stevenson-Jones

Centre for Bacterial Cell Biology

Institute for Cell and Molecular Biosciences

Medical School

Newcastle University

January 2017

Abstract

Transcription and translation form the basis of gene expression in all cells. In prokaryotes they are linked both spatially and temporally as the ribosomes begin translation of the RNA before the RNAP has finished transcribing the entire region, a process known as coupling. Interplay between the two machineries is highly complex and plays an important role in gene expression. To date, most of the studies into transcription-translation coupling have been carried out *in vivo*, and have focused on the indirect interactions such as attenuation. Due to the many accessory factors for both transcription and translation present within the cell, there is currently no known technique to study direct interactions between the RNAP and the ribosome. Recently, an *in vitro* transcription-translation system was developed in our lab that is formed from only the pure components required for transcription and translation. This allows the stepwise control of the RNAP and the ribosome. The aim of this study was to determine how close the RNAP and the ribosome can become on the same nascent RNA. The coupled *in vitro* system was redesigned and optimised to measure the distance between the actively transcribing RNAP and the ribosome translating the same transcript. We show that the ribosome can approach the RNAP as close as 26 nts between the A-site of the ribosome and the active site of the RNAP. This distance is far shorter than was previously thought and reveals a very close contact between the two machineries.

Acknowledgements

I would like to thank my Supervisor, Prof. Nikolay Zenkin, for his continued supervision, support and expertise throughout this PhD and my second supervisor, Dr. Yulia Yuzenkova, for her support and guidance. I would also like to thank Dr. Daniel Castro-Roa for teaching me most of the techniques used throughout this thesis, Dr. Soren Nielsen for his expertise with laboratory equipment and Dr. Pamela Gamba for her invaluable input in writing this thesis.

A big thank you goes to Pippa for listening, offering advice and coffee when needed. I would also like to thank the three musketeers for providing entertainment and scintillating discussion. Nunzia and Lucia, my favourite Italians, thanks for keeping me sane and fed throughout the challenges of this PhD.

Finally I would like to thank my family for their encouragement, love and support.

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.

Signature of candidate ..... Date 04/01/2017

Contents

List of Figures

List of Tables

List of Abbreviations

aa	amino acid
AAEC	artificially assembled elongation complex
Amp	ampicillin
ASL	acceptor stem loop
ATP	adenosine triphosphate
BH	bridge helix
CIAP	calf intestinal alkaline phosphatase
CTP	cytosine triphosphate
DBS	DNA binding site
DNA	deoxyribonucleic acid
DNAP	DNA polymerase
dNTP	deoxynucleoside triphosphate
DS	downstream
DTT	1-4, dithiothreitol
E. coli	Escherichia coli
EF-G/Ts/Tu	elongation factor G/Ts/Tu
ePEC	elemental paused elongation complex
fmet	formyl-methionine
FMT	formyl-methyltransferase
GDP	guanosine diphosphate
GMP	guanosine monophosphate
GTP	guanosine triphosphate
GTPase	GTP-dependent protease
HBS	hybrid binding site
hr/s	hour/s
IF-1/2/3	initiation factor 1/2/3
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria broth
M	molar
Mg^{2+}	magnesium ion
metRS	methionine- tRNA synthetase
mL	millilitre
min/s	minute/s
mM	millimolar

mRNA	messenger RNA
NAC	nucleotide addition cycle
NER	nucleotide excision repair
NET-seq	native elongating transcript sequencing
ηg	nanogram
NMP	nucleoside monophosphate
nt/s	nucleotide/s
ntDNA	non-template DNA
NTP	nucleoside triphosphate
Nus factors	N-utilisation factors
Oligo	oligonucleotide
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PE	pause element
Pep	2-polyenolpyruvate
PK	pyruvate kinase
pmol	picomole
Pol III	DNA polymerase III
PPi	inorganic pyrophosphate
(p)ppGpp	guanosine (penta-)tetra-phosphate
PTC	peptidyl transferase centre
RBS	RNA binding site
rut site	rho utilisation site
RNA	ribonucleic acid
RNAP	RNA polymerase
RF-1/2	release factor 1/2
RRF	ribosomal recycling factor
rRNA	ribosomal RNA
RT	room temperature
SD	shine dalgarno
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
tDNA	template DNA
TL	trigger loop
tRNA	transfer RNA
aa-tRNA^{aa}	amino-acyl-tRNA

TB	transcription buffer
TC/s	ternary complex/es
TCR	transcription coupled DNA repair
TEC/s	transcription elongation complex/es
TIR	translation initiation region
TLC	thin layer chromatography
TL-CTT	translation first coupled transcription-translation
TLE	thin layer electrophoresis
TR-CTT	transcription first coupled transcription-translation
TSS	transcription start site
UP	upstream
UTP	uracil triphosphate
UTR	untranslated region
UV	ultraviolet

1. Introduction

Transcription and translation form the basis of gene expression and are highly conserved throughout all domains of life. Together they produce a protein product using the cell's deoxyribonucleic acid (DNA). The DNA acts as a template for the synthesis of complementary ribonucleic acid (RNA) by the DNA dependent enzyme RNA polymerase (RNAP), using nucleoside triphosphate molecules (NTPs) as the substrate. The ribosomes use the RNA transcript as a template to synthesise a polypeptide protein chain made from individual amino acids. In prokaryotes, transcription and translation take place in the same space within the cell. The ribosomes bind and begin to translate the nascent RNA as soon as the ribosome-binding site emerges from the RNAP (Figure 1.1). This is in contrast to eukaryotic cells, where the entire length of RNA is transcribed in the nucleus before being transported to the cytoplasm for translation to take place.

The fundamentals of transcription and translation are conserved across all domains of life. This work focussed on transcription and translation in prokaryotes, specifically that of the model organism *Escherichia coli* (*E. coli*).

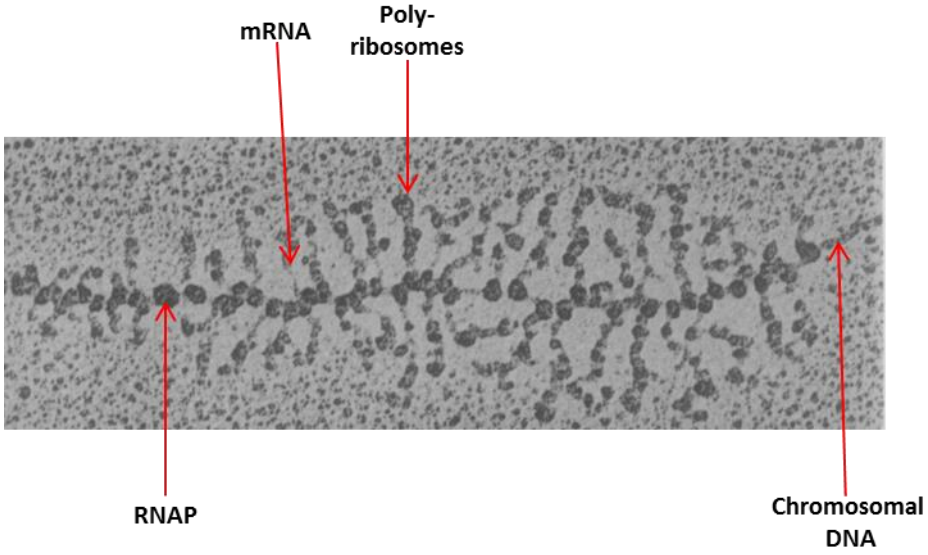


Figure 1.1 Electron micrograph showing simultaneous transcription and translation. Simultaneous transcription and translation form a branched structure, with the template DNA running horizontally and the RNA strands branching from the main stem. Modified from Paul *et al.* 2004.

1.1 Transcription

Transcription is the very first step in gene expression and, although the basic transcription process is conserved throughout all domains of life, the exact mechanism and the specific RNAP enzymes required for transcription vary between

prokaryotes, eukaryotes, archaea and bacteriophages/viruses. In all cases, however, RNA synthesis is achieved through nucleophilic attack of the phosphodiester bond of an NTP in the presence of a catalytic magnesium ion (Mg^{2+}) and incorporation of the resulting nucleoside monophosphate (NMP) at the end of the growing RNA chain. NTPs contain an invariant triphosphate group and a ribose sugar with one of four bases attached to the ribose sugar ring (Figure 1.2 A).

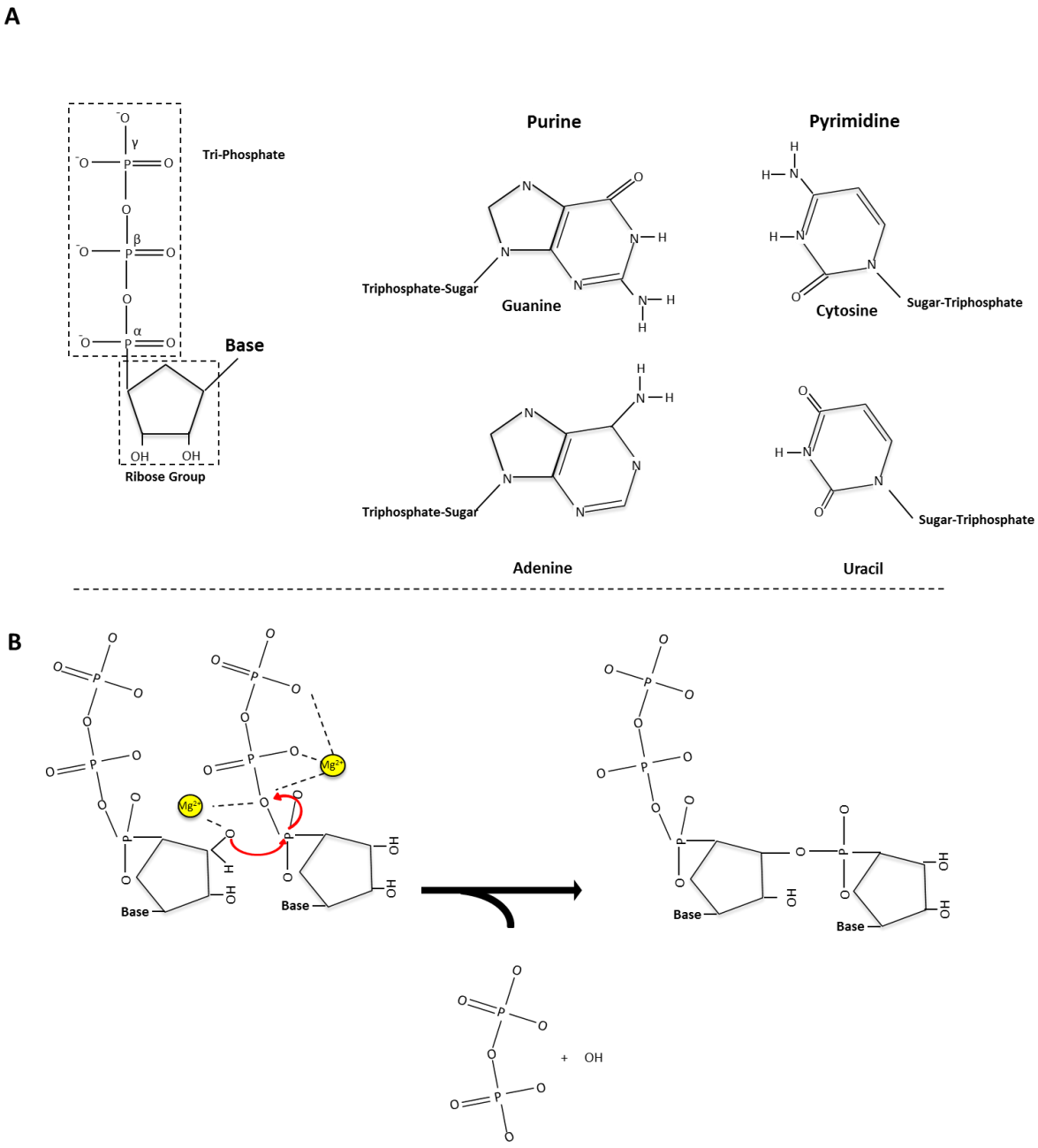


Figure 1.2 NTP structure and catalysis during RNA synthesis. A) Left: Structure of an NTP showing the triphosphate group and ribose group. Right: Structure of the four bases. B) Catalysis of nucleotide addition.The red arrows indicate direction of attack and the magnesium ions are represented in yellow.

1.1.1 RNAP structure and function

Prokaryotic RNAP enzymes are multi-subunit proteins. The *E. coli* RNAP core enzyme is made up of a total of five subunits, two alpha (α) subunits, one each of a beta (β), and beta-prime (β') subunit and an omega (ω) subunit ($\alpha_2\beta\beta'\omega$) (Burgess 1969). The core enzyme alone is sufficient for transcription elongation, but for transcription initiation, an extra subunit, the sigma (σ) factor, is required (Burgess *et al.* 1969). The addition of the σ factor to the core enzyme forms the RNAP holoenzyme. All of the subunits together contain the RNAP catalytic centre and three channels: the main channel, the RNA exit channel and the secondary channel (Figure 1.3). The overall structure of the RNAP is comprised of movable and immovable elements and the conformation of the movable elements varies throughout the transcription cycle.

During the early stages of initiation, the RNAP resembles an open, crab claw like structure, but, during the later stages of initiation and elongation, the structure closes around the DNA and RNA to form a more tightly bound complex (Chakraborty *et al.* 2012). Throughout transcription elongation, the RNAP contains the transcription bubble consisting of a locally melted 12-14 base region of DNA with 8-9 nucleotides of the 3' end of the nascent RNA base paired to the template DNA strand.

Each subunit within the RNAP holoenzyme contains specific domains and structures that have unique roles in the different aspects of transcription. The region around the catalytic site contains specific structural elements to stabilise the catalytic Mg^{2+} ion, select the correct NTP for incorporation and properly orient the NTP for hydrolysis. Other structures direct the nascent RNA out of the exit channel, separate the DNA strands and ensure the DNA template is correctly positioned within the RNAP. The individual subunits themselves fulfil diverse roles within the transcription cycle. The two α subunits of the RNAP bind to the β and β' subunits during RNAP assembly (Severinov *et al.* 1995) and also to the upstream promoter region of the DNA during initiation (Murakami *et al.* 1997). The β and β' subunits form the three channels in the RNAP, with each channel fulfilling a specific role. The main channel (also known as the active site cleft) accommodates the DNA and the 3' end of the RNA and also contains the active site. The nascent 5' RNA is extruded out through the RNA exit channel and the NTPs enter into the RNAP active centre through the secondary channel (Mukhopadhyay *et al.* 2004) (Figure 1.3).

The main channel contains a number of specific elements such as the bridge helix (BH) that functions to separate the main and secondary channels (Zhang *et al.* 1999) and the trigger loop (TL) that is involved in RNA synthesis and cleavage (Zhang *et al.* 2010; Yuzenkova & Zenkin 2010; Wang *et al.* 2006). The β subunit contains the clamp domain, a movable domain that opens and closes to form the open and closed crab claw like structures of the RNAP during transcription initiation and elongation respectively. The β' subunit contains the β' rudder for separation of the template and non-template DNA strands within the active centre of the RNAP by forcing a 90° kink in the DNA (Rees *et al.* 1993).

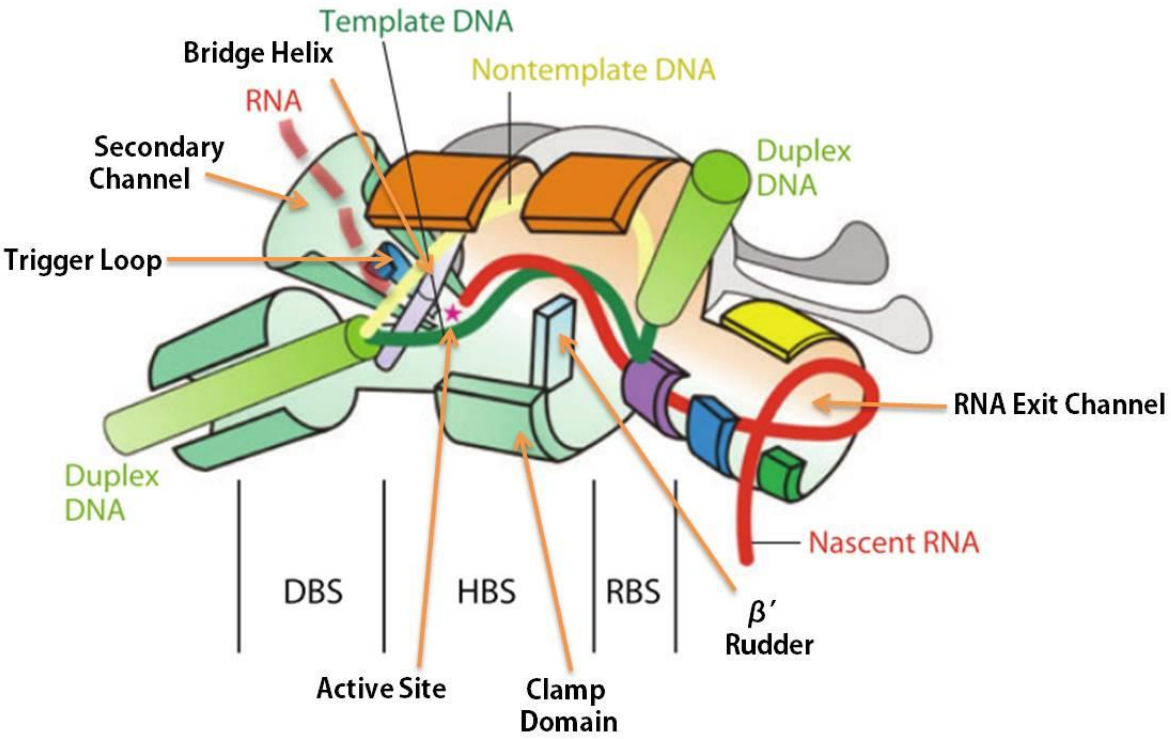


Figure 1.3 Schematic overview of RNAP. The DNA duplex and RNA:DNA hybrid are contained in the mian channel whilst the nascent RNA is extruded through the RNA exit channel. The NTPs enter the active site (represented by the red star) via the secondary channel. Modified from Nudler 2009.

The role of the ω subunit is less well defined and unlike the other subunits, it is not essential (Ishihama 1981; Saitoh & Ishihama 1976). The core enzyme without this subunit has been observed to be fully functional *in vitro* and *in vivo* and the $\alpha_2\beta\beta'$ enzyme is able to self-assemble into the active RNAP core enzyme, even in the absence of the ω subunit, after the subunits were purified individually and reconstituted *in vitro* (Ishihama 1981; Saitoh & Ishihama 1976). The role of the ω subunit appears to be predominantly for the recruitment and stabilisation of the β' subunit during RNAP assembly and also as a target for regulators of RNAP and transcription (Ghosh *et al.* 2001; Igarashi *et al.* 1989).

The σ factor is only required during transcription initiation as part of the RNAP holoenzyme and is essential for promoter recognition and binding during the early stages of initiation (Burgess *et al.* 1969).

The RNAP catalytic centre is located in the β' subunit and consists of 3 universally conserved aspartate residues (D) in a NADFDGD amino acid motif. These residues are essential for catalysis (Sosunov *et al.* 2005) as they co-ordinate the catalytic Mg^{2+} ion (Mg^{2+}A) during RNA synthesis (Zaychikov *et al.* 1996). and degradation. A second Mg^{2+} ion (Mg^{2+}B) is also required for binding of the NTPs in the active centre (Wu & Goldthwait 1969) and is brought into the active site by the incoming NTP. The incoming NTP also plays a part in retaining Mg^{2+} B (Sosunov *et al.* 2003).The Mg^{2+} ions are required for coordinating the phosphate groups of the RNA and NTPs and directing the nucleophilic attack of the hydroxyl group during RNA synthesis (Zaychikov *et al.* 1996). To add the incoming NTP to the 3' end of the RNA, the 3' hydroxyl (3'OH) group of the NMP at the 3' end of the acts as a nucleophile and attacks the triphosphate group of the incoming NTP between the α and β phosphate (Figure 1.2 B). The NTP loses two phosphate groups (PPi, pyrophosphate) and the resulting NMP is incorporated onto the 3' end of the RNA by forming a phosphodiester bond with the attacking 3'OH group. The Mg^{2+} ion is released along with the PPi molecule after formation of the phosphodiester bond (Yang *et al.* 2006).

In addition to the role of the Mg^{2+} ions during RNA synthesis (Sosunov *et al.* 2005), other functions of the RNAP catalytic centre require the presence of the catalytic Mg^{2+} ion, including cleavage of the RNA by the intrinsic hydrolysis capabilities of the RNAP (Sosunov *et al.* 2003).

1.1.2 Transcription Initiation

Transcription is a cyclic process in which the RNAP undergoes multiple rounds of transcription initiation, elongation and termination, altogether known as the transcription cycle. To initiate transcription, the RNAP holoenzyme recognises and binds to the promoter region of the DNA template, a region of the DNA comprised of two conserved motifs called the -35 and -10 elements. These motifs are located 35 and 10 nucleotides respectively upstream of the transcription start site (TSS), which is itself designated +1.The -10 element is made up of the consensus TATAAT nucleotide sequence (Paget & Helmann 2003; Gruber & Gross 2003). Recognition of the promoter region by RNAP requires the presence of the sigma factor in the RNAP holoenzyme (Burgess *et al.* 1969). In *E. coli*, different σ factors recognise the promoter regions of different classes of genes that are expressed under a variety of conditions (Gruber & Gross 2003). For instance, promoters for the housekeeping genes (genes expressed during normal cell growth) in *E. coli* are recognised by σ^{70}.

The RNAP holoenzyme locates the promoter region via a 3D diffusion mechanism (Wang *et al.* 2013; Friedman *et al.* 2013). Upon binding of the promoter region, a 12 base section of the DNA double strand starting at the -10 element and extending a few bases beyond the TSS is melted to expose the +1 TSS (Siebenlist 1979). During the early phase of initiation, the RNAP adopts an open promoter complex conformation (RP_o) resembling a crab claw like structure with the clamp domain in the open conformation (Finn *et al.* 2000). To begin transcription, the RNAP undergoes a few rounds of abortive initiation in which short strands of RNA up to 12 nucleotides long are synthesised and then released without the RNAP leaving the promoter region (Grachev & Zaychikov 1980; Munson & Reznikoff 1981). Escape of the RNAP from the promoter requires synthesis of a longer stretch of RNA. To do this, RNAP employs a mechanism called ‘scrunching’, during which a small section of the downstream DNA template is pulled into the RNAP holoenzyme to enable a longer stretch of RNA to be synthesised without the RNAP moving away from the promoter region (Kapanidis *et al.* 2006; Revyakin *et al.* 2006). The build-up of force generated by progressively pulling more DNA into the active centre and the synthesis of a longer length of RNA causes the eventual release of the RNAP from the promoter (Kapanidis *et al.* 2006). This, along with dissociation of the σ factor and closing of the clamp domain, forms the closed elongation complex (RP_c) and leaves the RNAP free to begin transcription elongation. Although the RNAP is able to bind

DNA in the absence of Mg^{2+}, synthesis of RNA requires the presence of the Mg^{2+} ion (Suh *et al.* 1993).

1.1.3. Transcription Elongation

When the RNAP enters elongation, it forms a transcription elongation complex (TEC) with the clamp domain tightened in the closed conformation (Finn *et al.* 2000). The TEC contains the transcription bubble, a 12 base region of locally melted DNA with a 8-9 base RNA:DNA hybrid (Nudler *et al.* 1997; Sidorenkov *et al.* 1998). The downstream double stranded DNA is accommodated in the DNA binding site (DBS) of the RNAP (Nudler 2009), the RNA:DNA hybrid is situated in the hybrid binding site (HBS) and the nascent RNA is contained in the RNA binding site (RBS) with the 5' end of the nascent RNA exiting the RNAP via the RNA exit channel. The HBS is formed by the β and β' subunits and contains the RNAP catalytic centre with the conserved aspartate residues (Nudler 2009) (Figure 1.3).

During the first step of the nucleotide addition cycle, the RNA 3' NMP is located in the i site of the RNAP catalytic centre, with the RNAP in the post-translocated state (Figure 1.4, step 1) (Komissarova & Kashlev 1997a; Bar-Nahum *et al.* 2005). An open trigger loop allows an incoming NTP to enter into the insertion site (i+1) via the secondary channel. If the incoming NTP is correct, the trigger loop will close and catalysis will occur between the 3' NMP and the new NTP aided by Mg^{2+} A and B (Vassylyev *et al.* 2007). At this point in the nucleotide addition cycle, the RNAP is stabilised in the pre-translocated state (Figure 1.4, step 2). Once the correct NTP has been successfully incorporated, the RNAP shifts forwards by one base along the DNA template, simultaneously unwinding the downstream DNA duplex by one bp and allowing the next bp of the upstream DNA duplex to re-anneal (Zaychikov *et al.* 1995). The NTP paired with the DNA template at the upstream edge of the RNA:DNA hybrid breaks contacts with the DNA template and is directed into the RNA exit channel. The RNAP returns to the post-translocated state with the new 3' end of the RNA once again situated in the i site (Figure 1.4, step 3.).

The RNAP core enzyme is highly processive and proceeds in a 5'-3' direction along the DNA template adding NMPs at a rate of 50 nts/sec (Vogel & Jensen 1994). The interaction of the RNAP with the DNA template is not sequence specific and the contacts formed between the RNAP and the DNA are strong enough that the RNAP remains bound to the template, but not too strong that the RNAP cannot translocate along the DNA at the high speed required for efficient transcription and the high

degree of processivity. RNAP translocates along the DNA template during transcription using the Brownian ratchet motion model of translocation (Bar-Nahum *et al.* 2005) during which RNAP oscillates locally by one nucleotide between the pre and post-translocated positions (Komissarova & Kashlev 1997a) on the DNA template using the free energy generated by the random movement of molecules (Figure 1.4, two way arrow between steps 2 and 3). The cognate NTP entering into the active site acts as a ratchet and anchors the RNAP in the forward, post-translocated position, allowing catalysis to occur (Bar-Nahum *et al.* 2005).

As well as the local oscillation between the pre and post-translocated states the RNAP is also prone to moving backwards along the DNA template by one or more nucleotides after synthesis of the nascent RNA. In this process, known as backtracking, the 3' end of the RNA is extruded out through the secondary channel (Komissarova & Kashlev 1997b). The upstream DNA is unwound and the downstream DNA allowed to rewind as the transcription bubble and associated RNA:DNA hybrid shifts backwards (Figure 1.4 step 4). In order to restart transcription, the 3' end of the RNA needs to be restored in the RNAP active site, either by the intrinsic hydrolysis property of the RNAP alone (Orlova *et al.* 1995) or with the assistance of transcription factors, for instance GreA and GreB in *E. coli* (Borukhov *et al.* 1993). The RNA in the secondary channel is cleaved between the NMP in the i site and the NMP located in the i+1 site, creating a new RNA 3' end in the i site of the active centre. The 5' RNA fragment is released and the i+1 site is now empty and ready for the next incoming NTP for continuation of transcription elongation (Borukhov *et al.* 1993). GreA is involved in the hydrolysis of RNA transcripts after backtracking by up to 3 nts, whereas GreB is involved in the resolution of TEC complexes backtracked by up to 9 nts (Borukhov *et al.* 1993).

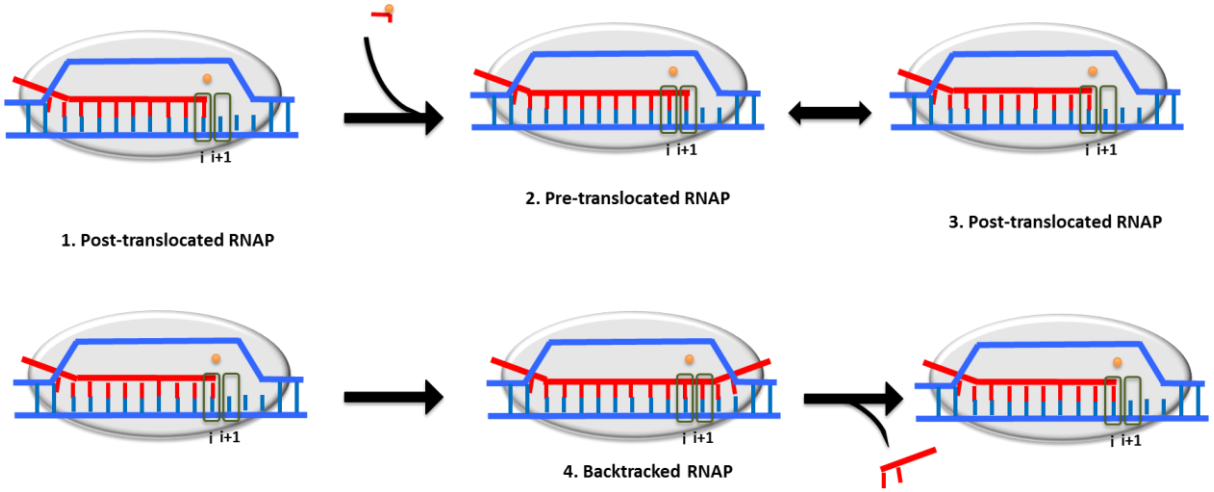


Figure 1.4 Schematic representation of RNA synthesis. Top: 1) Post-translocated RNAP with the 3' RNA end in the i site and the i+1 site empty. 2) The RNAP becomes pre-translocated by binding of the correct NTP in the i+1 site. 3) RNAP shifts along the DNA template by one base and once again adopts the post-translocated form. Bottom: 4) From the post-translocated position, the RNAP can move backwards along the DNA template and become backtracked. Cleaving the RNA in the active centre (either through intrinsic hydrolysis by the RNAP or with assistance from transcription factors) restores the 3' terminal NMP in the i site.

1.1.3.1 Proofreading and fidelity

The RNAP misincorporates NTPs at a rate of approximately 1 nt out of every 10^5 nts (Blank *et al.* 1986). Misincorporation of the incorrect NTP during transcription can have a detrimental impact on the cell as the RNA sequence is used by the ribosomes to produce the correct, functional protein. RNAP employs a variety of techniques to prevent misincorporation of the incorrect NTP, to sense the incorrect NTP either before or after misincorporation and to remove the incorrect NTP if misincorporated. The RNAP prevents misincorporation by displacing the non-cognate NTP from the active centre through folding of the trigger loop (Yuzenkova *et al.* 2010). If misincorporation does still occur, it can cause the RNAP to backtrack by one or more nucleotides, before being resolved by transcription factor assisted RNA hydrolysis to restore the 3' end of the RNA in the RNAP active site (Erie *et al.* 1993; Orlova *et al.* 1995). The misincorporated NTP itself is also capable of stimulating hydrolysis of the phosphodiester bond to remove the 3' dinucleotide portion of the RNA containing the incorrect NMP (Zenkin *et al.* 2006).

1.1.4 Pausing during transcription elongation

Despite the highly processive nature of the RNAP, pausing often occurs, either during the transcription initiation or elongation stages. In *E. coli* the RNAP pauses at a rate of roughly once in every 100 bases transcribed (Neuman *et al.*

2003). Although pausing of the RNAP decreases the overall rate of transcription, it plays an important role in transcription regulation.

The pause state of the RNAP is defined as a reversible state in which the RNAP is unable to catalyse the addition of the next nucleotide at the 3' end of the nascent RNA. There are two main classes of RNAP pause: off-pathway pauses and in-pathway pauses. Most RNAP pauses are considered to be 'off-pathway' events during which the RNAP temporarily deviates from the nucleotide addition cycle of transcription elongation and enters into a catalytically incompetent state. Pausing during elongation begins with the formation of an elemental paused elongation complex (ePEC) (Weixlbaumer *et al.* 2013). The RNAP is stabilised with the clamp domain in the open conformation, in contrast to the closed conformation of the actively transcribing TEC. Additional conformational changes in the RNAP active centre prevent the incoming cognate NTP from base pairing with the nucleotide in the template DNA and cause a loss of contacts between the RNAP and template DNA. Opening of the clamp domain widens the RNA exit channel and potentially allows formation of an RNA hairpin within the exit channel. Formation of a hairpin in the exit channel would sterically block closing of the clamp, leading to stabilisation of the pause state of the RNAP and prolonging the duration of the pause (Weixlbaumer *et al.* 2013). Backtracking of the RNAP can also increase the length of the pause.

Hairpin dependent pauses are an example of an off-pathway pause and are characterised by the formation of a hairpin roughly 10/11 nucleotides upstream of the RNAP active centre (Artsimovitch & Landick 1998; Vassylyev *et al.* 2002). The DNA template sequence downstream of the pause site determines the efficiency and duration of the pause (Chan & Landick 1993). The location of the hairpin sequence in the nascent RNA causes it to form in the RNA exit channel of the RNAP (Kolb *et al.* 2014) and temporarily prevents the RNAP from adding NMPs and translocating along the DNA template (Chan *et al.* 1997). Disruption of the hairpin structure restores the single stranded RNA in the RNAP exit channel and active centre to allow the RNAP to restart transcription elongation.

The vast majority of the interactions between the RNAP and the DNA:RNA hybrid during transcription elongation occur completely independently of the DNA or RNA sequence. Generally, recognition of the DNA template or RNA:DNA hybrid sequences by the RNAP requires the association of accessory factors. One exception to this is a specific type of pause in which the identity of the RNA:DNA

hybrid sequence strongly affects the pausing of the RNAP in the pre-translocated state. This pause is considered to be an in-pathway event as the TEC does not branch from the nucleotide addition cycle but instead the rate of translocation of the RNAP is substantially reduced (Bochkareva *et al.* 2012).

In the last few years, the development of a high throughput sequencing technique called native elongating transcript sequencing (NET-seq) has enabled the analysis of RNAP pausing on a much larger scale (Churchman & Weissman 2011) and was originally developed in *Saccharomyces cerevisiae*. Net-Seq was used to identify a consensus pause sequence in the *E. coli* genome consisting of a G at position -10 relative to the pause nucleotide, a pyrimidine (Y) residue at -1 (the nucleotide at the 3' end of the RNA) and a G at position +1 ($G_{-10}Y_{-1}G_{+1}$) (Vvedenskaya *et al.* 2014; Larson *et al.* 2014; Imashimizu *et al.* 2015). The identity of the NMP at the RNA 3' end (C/U) and the identity of the incoming NTP (G) both strongly affected the pausing of the RNAP within this particular region (Larson *et al.* 2014). The identification of this pause sequence led to the theory that the strength of the base pairing between the RNA and template DNA at these particular bases are strongest and therefore RNAP favours the pre-translocated state at this site, causing the RNAP to enter into a paused state (Vvedenskaya *et al.* 2014). This pause sequence was found to be enriched in the translation start sites of well transcribed genes in both the early and late operons (Larson *et al.* 2014).

After transcription initiation and promoter escape, the sigma factor is released from the RNAP holoenzyme. However, occasionally, σ^{70} is retained after promoter escape (Harden *et al.* 2016) and causes promoter proximal pausing at sequences resembling the -10 element (generally ANNNT nucleotide sequences) (Ring *et al.* 1996; Nickels *et al.* 2005; Goldman *et al.* 2015; Strobel & Roberts 2015). Upon pausing, the RNAP can begin the scrunching of the template DNA seen during initiation in order to continue RNA synthesis. This can subsequently lead the RNAP to enter into an ePEC and the RNAP then becomes prone to backtracking along the DNA template. If backtracking occurs, Gre factors are required to resolve the backtracked complex for RNA synthesis to restart. With sigma dependent RNAP pausing, the RNAP can go through multiple rounds of pausing, backtracking and forwards translocation caused by the same pause site before eventually escaping the promoter-like pause sequence, similar to promoter escape during transcription initiation (Strobel & Roberts 2015).

1.1.5 Transcription Termination

E. coli employs two mechanisms to terminate transcription, Rho dependent and hairpin dependent transcription termination. Rho dependent transcription termination requires the accessory factor Rho, whereas hairpin dependent transcription termination occurs without the need for any transcription termination factors. Rho is a hexameric protein composed of a trimer of dimers that together form a ring shaped unit. Rho recognises the rho utilisation (rut) site on the RNA, characterised by long regions (≥80 bases) of single stranded GC rich RNA devoid of secondary structures (Boudvillain *et al.* 2010) and for which no consensus sequences have been identified. Rho terminates transcription by binding to the rut site and forming a ring around the RNA that leads to the eventual disassociation of the TEC and release of the RNA by the RNAP (Boudvillain *et al.* 2010). The exact mechanism as to how Rho terminates transcription is unknown but the current accepted theory is that Rho disrupts the TEC and causes dissociation of the RNAP from the DNA and RNA (Brennan *et al.* 1987; Epshtein *et al.* 2010). Two separate theories posit that Rho is either associated with the RNAP throughout transcription (Epshtein *et al.* 2010) and binds to the RNA when it encounters a rut site or alternatively that Rho binds to the rut site at a distance from the RNAP and translocates along the RNA towards the RNAP by using its ATP dependent translocase activity (Richardson 2002). If Rho is already associated with the RNAP, it is ideally located to disrupt the TEC and cause dissociation of the RNAP from the RNA upon encountering a rut site. Pausing of the TEC is a useful tool during Rho dependent transcription termination as it allows time for Rho to translocate along the RNA and catch up with the RNAP during the Rho on RNA pathway of transcription termination (Jin *et al.* 1992).

Rho is an essential protein in *E. coli* and Rho dependent transcription termination is important for maintaining cell viability by suppressing expression of intragenic regions of the genome containing horizontally acquired genes that are not normally expressed, either due to their toxicity or because they are not translated (Cardinale *et al.* 2008). Rho is also required for the termination of antisense transcripts that are also not translated (Peters *et al.* 2012).

Recently, small RNAs (sRNAs) have been identified as a factor to prevent Rho dependent transcription termination in the 5’ untranslated region (5’ UTR) of many *E. coli* genes (particularly those with long 5’ UTRs, > 80 nts) during conditions of cell

stress (Sedlyarova *et al.* 2016). sRNAs complementary to the 5' UTR sequence pair with the RNA as it exits the RNAP, preventing Rho from either binding to the RNA or from translocating along the RNA towards the RNAP (Sedlyarova *et al.* 2016).

Hairpin dependent transcription termination occurs via the formation of an RNA hairpin in the RNAP exit channel due to a palindromic sequence in the RNA. The palindrome sequence is followed by a stretch of poly-Us that result in a weak DNA:RNA hybrid in the RNAP active centre due to weak base pairing between the nts in the DNA template and RNA (Peters *et al.* 2011). Formation of the hairpin in the RNA exit channel causes a change in conformation of the RNAP and, combined with a weak RNA:DNA hybrid, leads to disruption of the TEC and release of the RNAP, DNA and RNA (Gusarov & Nudler 1999).

1.1.6 Interactions during transcription elongation

During transcription, strong promoters of highly transcribed genes will undergo multiple rounds of transcription initiation with the next RNAP being initiated as soon as the lead RNAP has cleared the promoter region, before it has finished transcribing the entire gene (Hamming *et al.* 1981). These multiple rounds of transcription initiation and elongation produce many TECs transcribing the same region of the genome. Trailing TECs have been demonstrated *in vivo* to re-start the leading TEC after stalling due to backtracking and without the need for accessory factors (Epshtein & Nudler 2003), a process known as cooperation. The greater the number of trailing TECs, the stronger this co-operative effect is on the stalled TEC. This cooperation has also been observed when TECs were stalled in the backtracked position upon encountering a roadblock bound to the downstream region of the DNA template (Epshtein *et al.* 2003). Multiple rounds of initiation lead to multiple TECs that, together restarted the backtracked TEC and displaced the roadblock by pushing the initial TEC through it (Epshtein *et al.* 2003). Mathematical modelling has also shown that stalling of the leading RNAP can lead to traffic jams of the trailing elongation complexes (Yuzenkova *et al.* 2014).

During transcription, the RNAP will come into contact with many obstacles in the form of other proteins or protein complexes interacting with the DNA. Replication of the DNA by the DNA-dependent DNA polymerase (DNAP) also uses the DNA as a template to produce a new copy of the genome during cell division. The DNAP replicates the DNA bi-directionally and at a much faster rate than the RNAP transcribes the RNA. As a result, the replisome and the TEC will inevitably collide,

both in a co-directional and head on orientation (Figure 1.5). The exact method of resolution depends on the direction in which the collision occurred. When a co-directional collision occurs, the replication fork and TECs collapse and both DNA Polymerase III (Pol III) and RNAP dissociate from the DNA. *In vitro*, Pol III has been shown to use the RNA synthesised by the RNAP that remains bound to the DNA as a primer to restart replication (Pomerantz & O'Donnell 2008) (Figure 1.5). Head on collisions, on the other hand, can lead to disintegration of the TEC and removal of only the RNAP from the DNA template, whilst the replisome continues on replicating the DNA unimpeded (Pomerantz & O'Donnell 2010) (Figure 1.5). Rho dependent transcription termination plays a role in maintaining genome integrity by removing TECs that are transcribing in front of the replisome in an effort to reduce the occurrence of co-directional collisions (Washburn & Gottesman 2011).

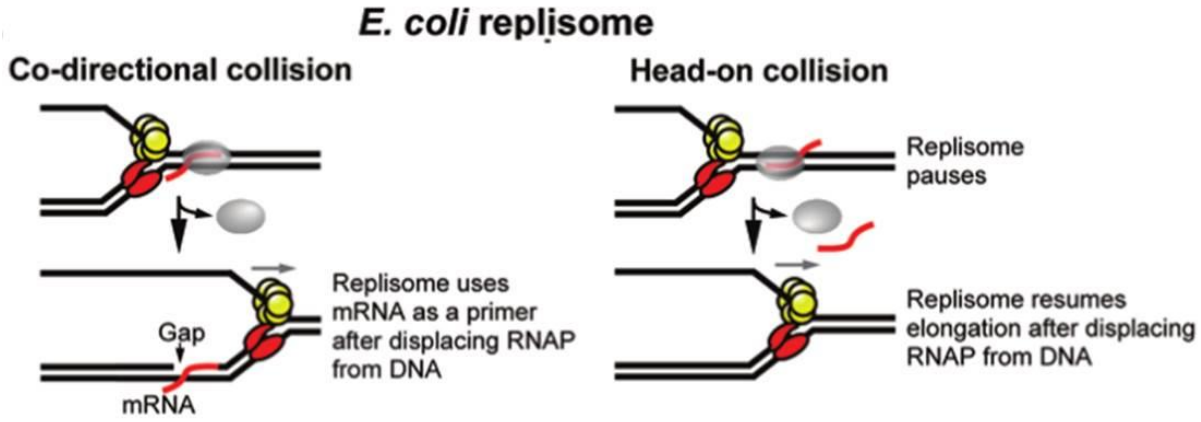


Figure 1.5 Consequences of replisome and TEC collisions. During a co-directional collision, the replisome uses the RNA left on the template after RNAP is displaced to re-start replication. In head-on collisions, the replisome displaces the RNAP to continue replication. Modified from Pomerantz & O'Donnell 2010.

Another group of DNA binding proteins that the RNAP encounters during transcription elongation are the proteins involved in the DNA repair pathways, most notably those involved in nucleotide excision repair (NER). The NER pathway detects lesions in the DNA strand that can consist of either a single damaged base or an entire region of damaged DNA (Kisker *et al.* 2013). In NER, the lesion is removed and the opposite strand is used as a template to repair the damaged area. In the absence of RNAP, the NER protein, UvrA, in the dimeric form detects DNA damage with the help of a UvrB dimer (Figure 1.6). When a lesion is detected, UvrA is released and UvrB binds tightly to the DNA, recruiting UvrC. UvrC excises the damaged base, before UvrD is recruited and DNA polymerase I fills in the gap and a

DNA ligase heals the single strand break (Kisker *et al.* 2013) (Figure 1.6). If the RNAP encounters a lesion during transcription elongation, stalling of the RNAP on the DNA template at the site of DNA damage will stimulate the NER pathway via one of two different pathways, a process known as transcription-coupled repair (TCR, Figure 1.6). In the Mfd pathway, the ATP dependent DNA translocase protein Mfd, removes the RNAP from the damaged region by pushing the RNAP through the site that caused the RNAP to stall (Park *et al.* 2002). This exposes the lesion for UvrA (recruited by Mfd) to begin the NER process (Fan *et al.* 2016) (Figure 1.6). Recent evidence suggests that Mfd, once recruited by stalling of RNAP, can also stimulate NER at sites up to 100 bp upstream from where RNAP originally stalled by utilising its translocase activity to scan ahead along the DNA (Haines *et al.* 2014).

An alternate pathway for TCR relies on the NER helicase protein, UvrD and occurs independently of Mfd where stalling of the RNAP at a site of DNA damage begins the process of TCR. First, UvrD binds to RNAP stalled at sites of DNA damage and induces backtracking of the RNAP away from the damage site (Epshtein *et al.* 2014) (Figure 1.6). UvrD then recruits the proteins required for NER. The UvrD TCR pathway has an advantage over the Mfd TCR pathway in that it is faster and the RNAP is not displaced during repair of the DNA and therefore can resume transcription immediately after NER is completed and the proteins removed.

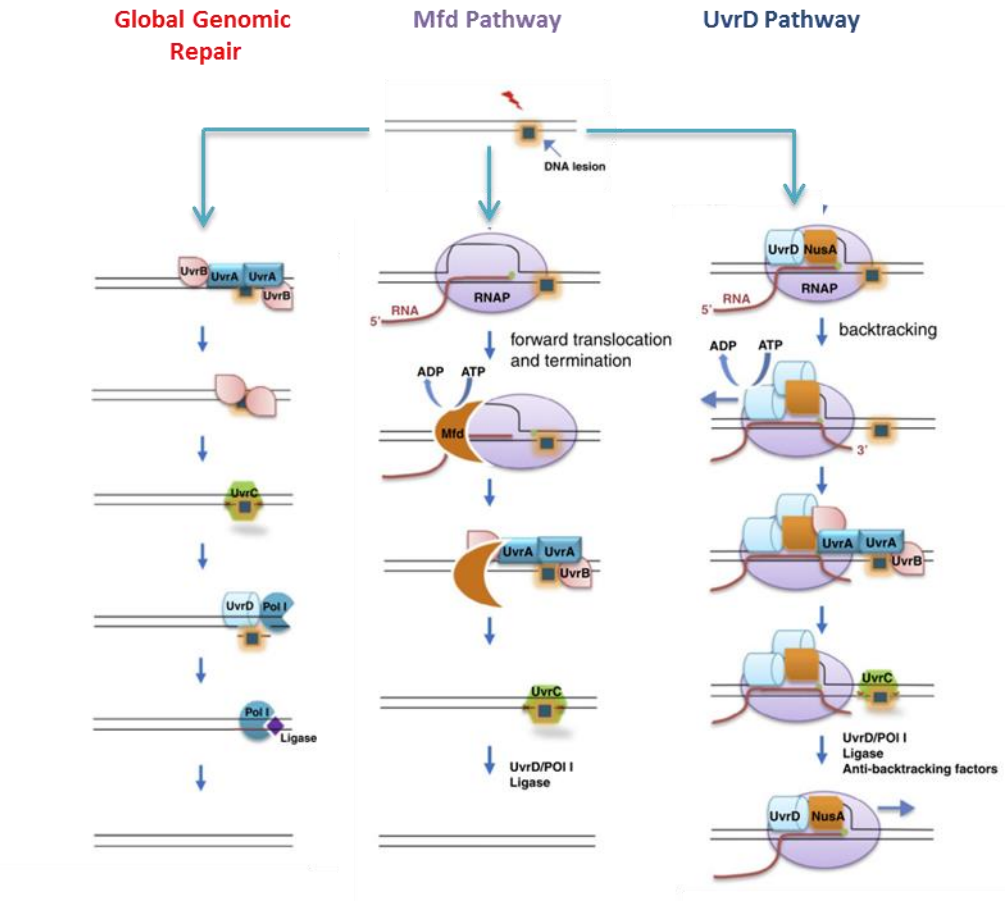


Figure 1.6 The three main Nucleotide Excision Repair pathways. Figure taken from Kamarthapu & Nudler 2015. Global genomic repair occurs in the absence of RNAP, whilst the Mfd pathway and UvrD pathway both use RNAP to detect the DNA lesion. The Mfd pathway causes the RNAP to be pushed through the paused site to allow access to the NER enzymes. In the UvrD pathway, the RNAP is pulled backwards to reveal the DNA lesion. In the Mfd pathway, transcription is terminated whereas in the UvrD pathway transcription is restarted once the DNA lesion is repaired.

1.1.7 Accessory factors

Transcription is one of the fundamental processes within the cell and is subjected to a high level of regulation during all stages of the transcription cycle. As well as the previously mentioned Mfd, UvrD, GreA and GreB proteins, there is a group of four N-utilisation factors, NusA, NusB, NusE and NusG. As a co-ordinated group of proteins, they play a role in transcription anti-termination of the lambda phage genes, but as individual proteins, they also have a wide range of other functions (Figure 1.7). NusA alone facilitates hairpin dependent transcription anti-termination (Yang & Lewis 2010). There is also evidence suggesting that NusA may play a role in TCR by facilitating UvrD induced RNAP backtracking and recruitment of the NER proteins themselves (Epshtein *et al.* 2014). NusB can bind single stranded RNA (Burmann, Luo, *et al.* 2010) and interacts with NusE (Luo *et al.* 2008; Mason *et al.* 1992), a multi-functional protein that is also known as ribosomal protein S10

(Wimberly *et al.* 2000). The interaction between NusE and NusB is thought to facilitate the loading of the ribosomes on the RNA during translation initiation (Luo *et al.* 2008). NusG is the most universally conserved of the Nus factors. Alone, it decreases the rate of RNAP pausing during transcription elongation and also plays a role in Rho-dependent transcription termination (Burmann, Schweimer, *et al.* 2010). An interaction between NusE as part of the translation elongation complex and NusG as part of the TEC has been proposed to play a role in the coupling of translation to transcription (see section 1.3 for a more detailed description) (Burmann, Schweimer, *et al.* 2010).

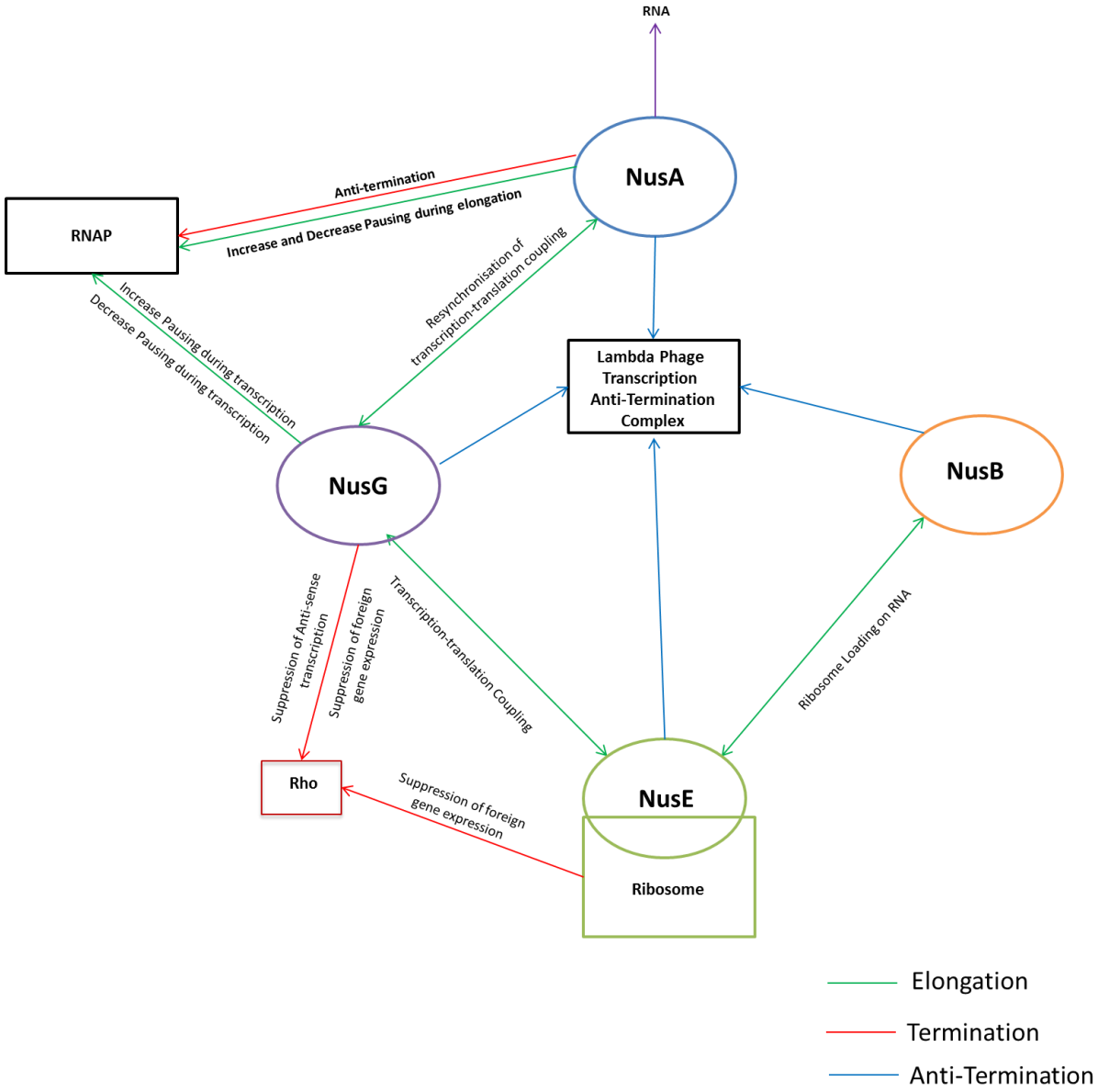


Figure 1.7 Interaction network of the Nus Factors during transcription elongation, termination and anti-termination. The main interactions are shown with the arrow indicating the direction of the interaction. The colour of the arrow corresponds to the role of the interaction in transcription elongation, termination or anti-termination.

1.2 Translation

The second step of gene expression after synthesis of the messenger RNA (mRNA) by the RNAP is translation of the mRNA by the ribosomes. The ribosomes use the nucleotide sequence within the mRNA to sequentially add amino acids to produce a growing polypeptide chain, the protein primary structure. Ribosomes are among the most universally conserved macromolecules across all domains of life and are formed from two subunits, the large and the small subunit, both of which are composed of ribosomal RNA (rRNA) and proteins. In *E. coli*, the 30S small ribosomal subunit and the 50S large ribosomal subunit make up the full 70S ribosome, with the

catalytic peptidyl transferase (PTC) centre located within the 50S subunit (Ramakrishnan *et al.* 2002) (Figure 1.8 A). The 30S subunit is composed of the 16S rRNA and 21 proteins, whilst the 50S subunit is composed of the 5S and 23S rRNA and 31 proteins (Selmer *et al.* 2006).The 16S rRNA of the 30S subunit is highly conserved across all domains of life and is often used to determine the relationship of different species to one another.

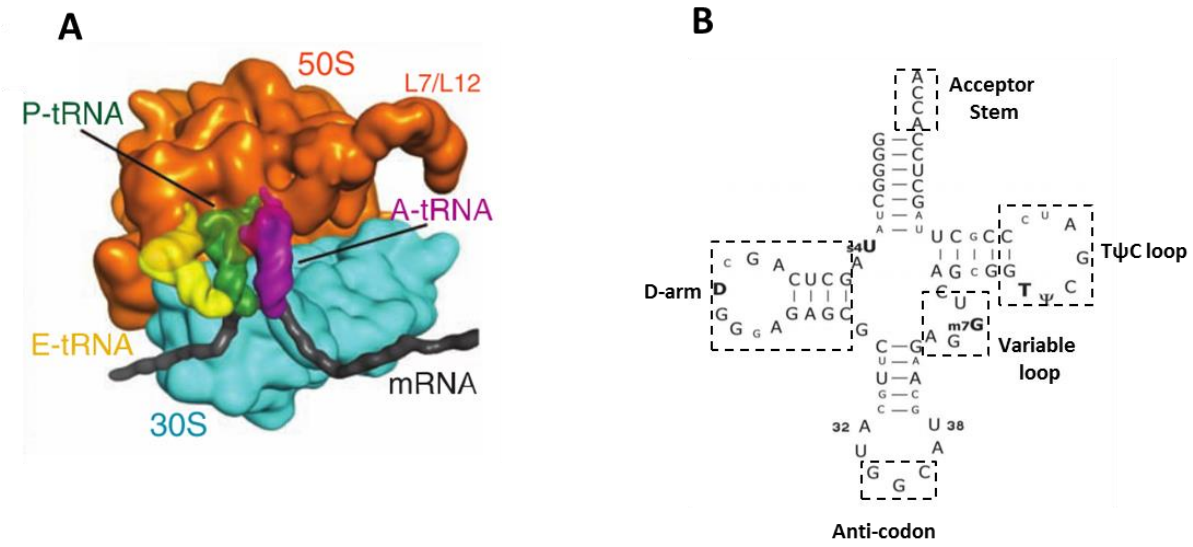


Figure 1.8 Structure of the 70S ribosome and tRNA. A) Structure of the 70S ribosome in the translation elongation form, adapted from Schmeing *et al.* 2009**.** B) Structure of an example tRNA (Alanine-tRNA) adapted from Ledoux *et al.* 2009**.**

During translation, the amino acids are brought into the ribosome's active centre attached to a transfer RNA (tRNA) molecule. tRNAs are a specific type of RNA molecule that form a secondary structure resembling a clover leaf shape made up of 3 stem loops (Figure 1.8 B). Each tRNA is unique to the specific amino acid attached to the CCA-3' RNA sequence (the acceptor stem loop (ASL)) and has an anti-codon composed of a nucleotide triplet sequence located at the bottom of a stem loop structure. The anti-codon is complementary to the tri-nucleotide codon sequence on the RNA and forms base pairs with the mRNA located at the interface between the 30S and 50S subunits. Binding of the tRNA anti-codon to the mRNA codon positions the amino acid attached to the ASL in the PTC for catalysis (Yusupov *et al.* 2001). Because the possible permutations for all of the anti-codon sequences far exceeds the number of amino acids available, there is often more than one anti-codon triplet for each amino acid. In most cases, it is the 3rd nucleotide of the triplet sequence that differs between anti-codons of the same amino acid. In fact, the ribosome can tolerate a lack of base pairing between the 3rd nt of the codon and anti-codon, a phenomenon known as third position wobble (Ogle *et al.* 2001).

When the 70S ribosome is bound to the mRNA, the mRNA is accommodated at the interface between the 50S and 30S subunits (Figure 1.8 A). The ribosome catalytic centre contains three sites, the amino-acyl (A) site, the peptidyl-transferase (P) site and the exit (E) site. The three tRNA binding sites span both of the ribosomal subunits to position the tRNAs such that the anti-codon loop is located in the 30S subunit and the acceptor stem is located within the peptidyl transferase centre of the 50S subunit (Figure 1.8 A) (Yusupov *et al.* 2001). The tRNAs exist in three forms within the ribosome: as amino-acyl-tRNA (aa-tRNA^{aa}) with one amino acid attached to the acceptor stem, as peptidyl-tRNA with the peptide bound to the acceptor stem and without either an amino acid or peptide bound to the acceptor stem (uncharged-tRNA) (Ramakrishnan *et al.* 2002). The incoming aa-tRNA^{aa} binds to the A-site, whilst the P-site contains the peptidyl-tRNA and the uncharged tRNA is located in the E-site (Ramakrishnan *et al.* 2002; Steitz 2008).

Translation, similar to transcription, is a cyclic process with three main stages: initiation, elongation and termination.

1.2.1 Translation Initiation

Translation is initiated by assembly of the ribosomes on the mRNA synthesised by the RNAP during transcription and requires the involvement of three translation initiation factors 1, 2 and 3. The 16S rRNA of the 30S subunit of the ribosome recognises the Shine Dalgarno (SD) site located in the translation initiation region (TIR) on the RNA through base pairing between the SD site and a sequence of complementary nucleotides at the 3' end of the 16S rRNA (Jacob *et al.* 1987; Shine & Dalgarno 1974). The TIR consists of the SD site and the AUG start codon separated by an RNA spacer sequence. Once the SD site has been located binding of the 30S subunit to the RNA requires initiation factor three (IF-3) (Figure 1.9 A step 2) (Paci *et al.* 1985). Once the 30S subunit is correctly positioned on the RNA, initiation factor one (IF-1) directs the initiator-tRNA (fMet-tRNA^{fMet}) directly into the P-site of the ribosome by blocking access to the A site (Figure 1.9 A step 3). The initiator-tRNA contains an anti-codon complementary to the AUG start codon and has a modified methionine with a formyl group (Figure 1.9 B). This modified methionine is only used during translation initiation and is not always retained in the final protein product (Giglione *et al.* 2004). The third IF, IF-2 is a GTPase enzyme that binds to the 30S subunit in its active, GTP bound form at some point before the initiator tRNA is recruited. Binding of the initiator tRNA into the A-site stabilises IF-2:GTP on the

30S subunit (Figure 1.9 A step 3) (Antoun *et al.* 2003). Once the 30S subunit is bound to the mRNA and the initiator-tRNA is correctly located in the P-site, IF-3 is released, as release of IF-3 is required before the 50S subunit can bind (Antoun *et al.* 2006). The presence of IF-2 in its GTP bound form increases the rate of 50S subunit joining to the ribosome to form the 70S subunit (Marshall *et al.* 2009) (Figure 1.9 A step 4). Once the 70S ribosome has been formed, hydrolysis of the GTP bound to IF-2 causes a conformational change in IF-2 and leads to dissociation of IF-2:GDP from the 70S ribosome (Hauryliuk *et al.* 2009) (Figure 1.9 A step 5). Release of IF-2:GDP completes the translation initiation stage of the translation cycle by producing a fully formed translation elongation complex (Figure 1.9 A step 5) (Marshall *et al.* 2009).

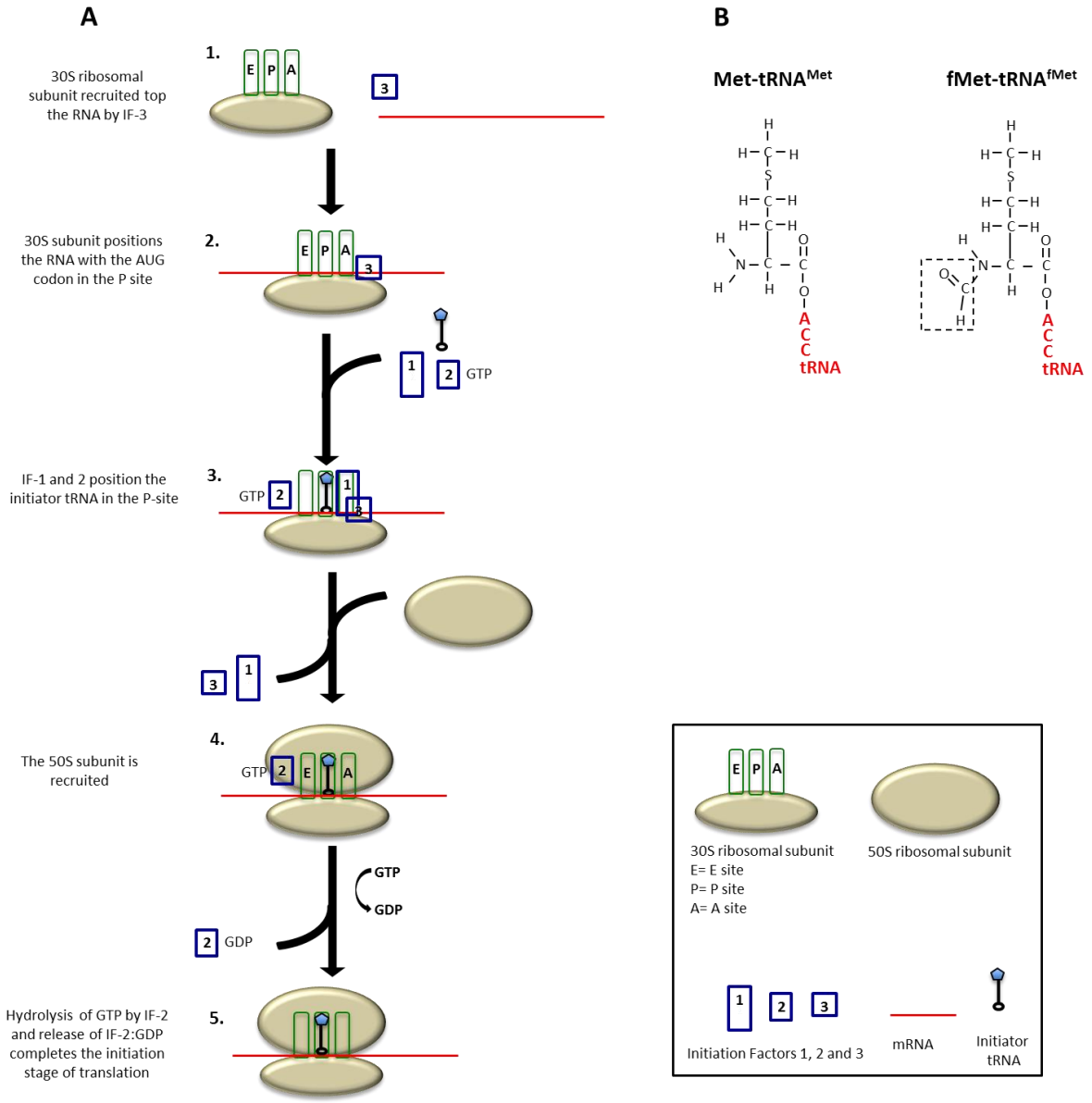


Figure 1.9 Schematic representation of translation initiation. A) Schematic representation of translation initiation showing the ribosomes, mRNA , IFs and GTP hydrolysis. B) Structure of Met-tRNA^{Met} and formylMet-tRNA^{fMet}. ACC-tRNA represents the acceptor stem of tRNA to which the amino acids are bound.

1.2.2 Translation Elongation

Translation elongation is a multi-step process requiring the involvement of elongation factors in recruiting the aa-tRNA^{aa}, proofreading, positioning the aa-tRNA^{aa} within the catalytic site and translocating the ribosome along the RNA after catalysis. The catalytic reaction itself however, does not require any input from additional factors but involves only the RNA within the catalytic centre, specifically a adenosine residue within the 23S rRNA (Muth *et al.* 2000).

To recruit aa-tRNA^{aa} to the ribosome, elongation factor-Tu (EF-Tu), a GTP dependent protease protein (GTPase), binds to aa-tRNA^{aa} to form a ternary complex (TC) when bound by GTP (EF-Tu:GTP:aa-tRNA^{aa}). EF-Tu, like all GTPases, is only active when bound by a GTP molecule. EF-Tu brings the TC to the ribosomal A site. If the aa-tRNA^{aa} is correct, the TC will be positioned in the A-site with the anti-codon loop bound to the cognate codon in the 30S subunit and the 5' CCA ASL in the peptidyl transferase centre of the 50S subunit (Figure 1.10 A, step 2). Upon correct positioning of the cognate aa-tRNA^{aa}, the GTP associated with EF-Tu will be hydrolysed to GDP and EF-Tu released (Figure 1.10 C). If the aa-tRNA^{aa} is not the correct one, the GTP bound to EF-Tu will not be hydrolysed and the non-cognate aa-tRNA^{aa} will exit the A-site to make way for another TC to bind (Figure 1.10 C). After GTP hydrolysis and release, the EF-Tu enzyme is regenerated by exchanging the GDP with GTP by the elongation factor EF-Ts (Green & Noller 1997).

Upon binding of the correct aa-tRNA^{aa} in the ribosomal A-site, the catalytic RNA within the 23S rRNA of the 50S subunit PTC catalyses the formation of a peptide bond between the amino acid on the aa-tRNA^{aa} in the A site and the peptide bound to the peptidyl-tRNA in the P-site, resulting in transfer of the peptide to the A-site tRNA (Figure 1.10 A, step 3) (Green & Noller 1997).

Once the catalytic reaction has occurred, translocation of the ribosomes proceeds via a two-step mechanism. The first step occurs independently of GTP and elongation factors, whereas the second step requires hydrolysis of GTP by elongation factor G (EF-G) (Moazed & Noller 1989). After peptide bond formation and transfer of the peptide to the A-site tRNA, the acceptor stem of the A-site tRNA is shifted relative to the 50S subunit from the ribosomes A-site to the P-site (Figure 1.10 A, step 4). Simultaneously, the de-acylated acceptor stem of the tRNA in the P-site is transferred to the E-site. This occurs by rotation of the 30S subunit relative to the 50S subunit (Valle *et al.* 2003) and forms a hybrid state of the ribosome with the ends of

the tRNAs in different sites within the 30S subunit compared to the 50S subunit (Figure 1.10 A, step 4) (Moazed & Noller 1989). Once the hybrid state of the ribosome has been formed, EF-G:GTP binds to the ribosome (Figure 1.10 A, step 5) (Wilden *et al.* 2006) and upon catalysis of the GTP, shifts the 30S subunit of the ribosome relative to the mRNA and the base-paired anti-codon loop of the tRNA by one codon to realign the 30S subunit with the 50S subunit (Figure 1.10 A, step 6) (Rodnina *et al.* 1997; Agrawal *et al.* 1999). The 30S subunit moves via a clockwise rotation mechanism (Frank & Agrawal 2000; Ramrath *et al.* 2013; Holtkamp *et al.* 2014; Ermolenko & Noller 2011) and shifts the anticodon loop of the now peptidyl-tRNA simultaneously from the A-site to the P-site (Green & Noller 1997).

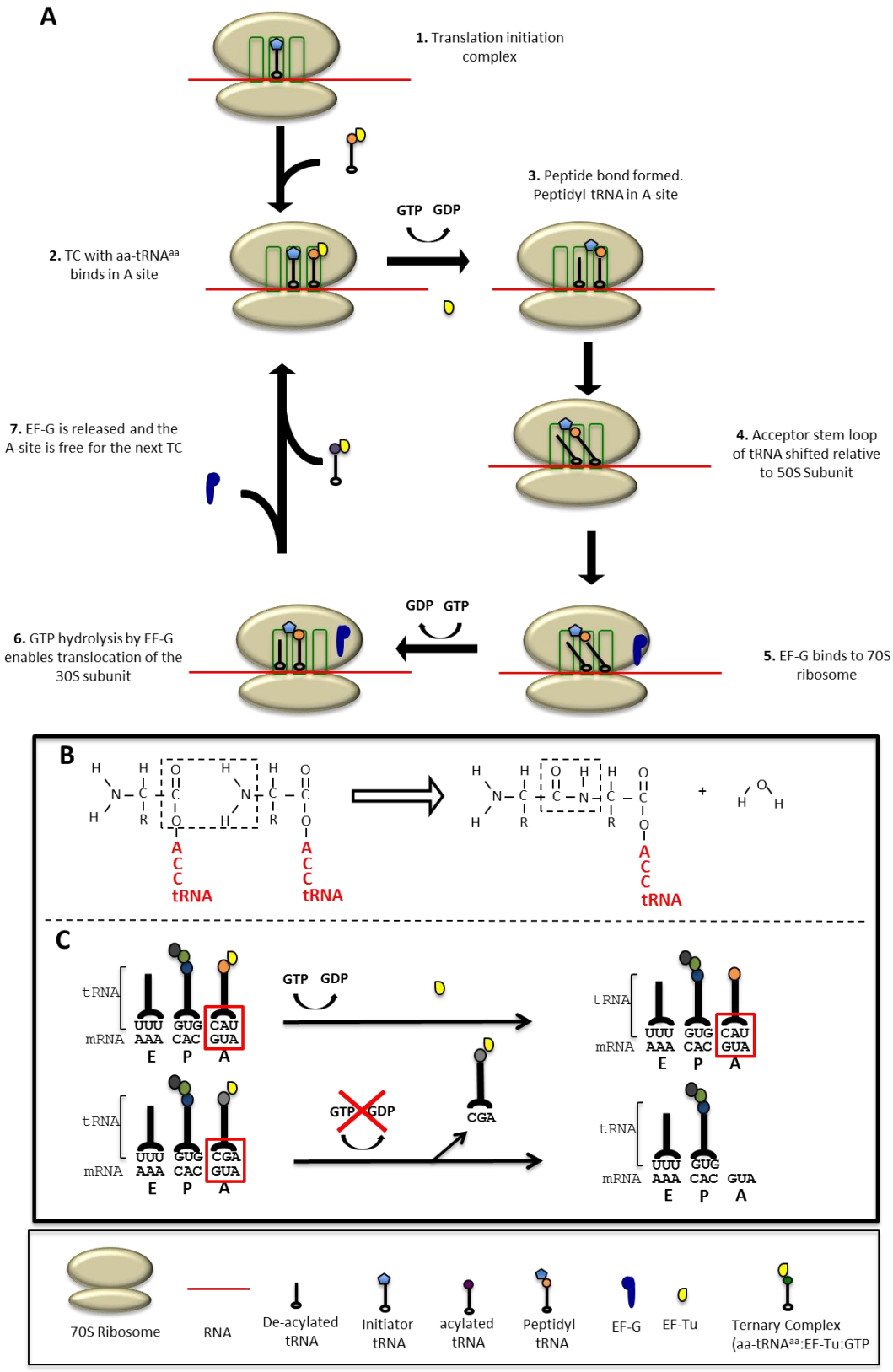


Figure 1.10 Schematic representation of translation elongation. A) Schematic overview of translation elongation including translocation. B) Peptide bond formation. C) Schematic representation of proofreading by EF-Tu:GTP hydrolysis.

1.2.3 Accessory factors

All naturally occurring amino acids contain the same core structure and are distinguishable by the differences in composition of the R-group side chain. The chemical make-up of the R-group differs considerably between amino acids and can affect the efficiency of hydrolysis by the ribosome and potentially decrease the overall rate of translation elongation. This is particularly noticeable during the synthesis of peptides containing a series of amino acids for which the ribosome is especially slow in forming the peptide bond between peptide, for instance within poly-proline regions of an amino acid sequence (Woolstenhulme *et al.* 2013). Due to the ring structure of the proline side chain, the proline amino acid fits awkwardly into the PTC, making it harder for the catalytic RNA to access the carboxyl and amine groups (Melnikov *et al.* 2016). As a result of this, the ribosome is prone to stalling within these particular regions. Recently, the elongation factor, EF-P, was identified as being able to increase the efficiency of peptide bond formation between consecutive proline residues (Ude *et al.* 2013; Doerfel *et al.* 2013). EF-P binds to the ribosome between the E and P sites and interacts with both the CCA acceptor stem and the anticodon stem loop of the tRNA (Blaha *et al.* 2009) and is ideally located to interact with the amino acid in the PTC. EF-P specifically recognises the D-arm loop of Pro-tRNA^{Pro} (Katoh *et al.* 2016) and is thought to lead to an increase in efficiency of catalysis by orientating the amino acid within the PTC for optimum hydrolysis (Katoh *et al.* 2016). Although EF-P appears to be important for efficient synthesis of poly-proline regions within peptides, it was originally identified due to its ability to increase the rate of peptide bond formation by up to 10 fold between fMet-tRNA^{fMet} and the synthetic amino-acyl-tRNA analogue puromycin (Glick & Ganoza 1975) in the absence of all other elongation factors. The D-arm of fMet-tRNA^{fMet} is structurally very similar to that of Pro-tRNA^{Pro} and this could explain the affinity of EF-P for peptide bond formation of the attached amino acids (Katoh *et al.* 2016).

1.2.4 Translation Termination

Termination of translation occurs upon recognition of one of three stop codons (UAA, UAG or UGA) and is facilitated by several ribosomal release and recycling factors. IF-3 is also involved in translation termination as well as initiation. When the ribosome positions a stop codon in the A-site, release factors 1 or 2 (RF-1/2) binds in the A-site, and along with IF-3 and EF-G, stimulates hydrolysis of the ester bond to release the peptide from the peptidyl–tRNA. The ribosomal subunits dissociate from

the RNA and also from each other to recycle the individual subunits for another round of translation on a new strand of RNA (Weixlbaumer *et al.* 2008). IF-3 ensures that the 30S subunit is separated from the 50S subunit. The ribosome recycling factor 3 (RRF-3), recycles RF-1 and 2 by freeing them from the ribosome upon hydrolysis of GTP (Mora *et al.* 2003).

1.3 Transcription-Translation Coupling

Within the prokaryotic cell, the RNA strand is both transcribed and translated simultaneously. As soon as the SD site has exited the RNAP, the ribosome can bind to the nascent RNA, long before the RNAP has finished fully transcribing the entire region. Coupling between transcription and translation has been shown to occur both through direct contacts between the RNAP and the ribosomes and also indirectly via intermediary factors (McGary & Nudler 2013). Transcription-translation coupling is important for the regulation of gene expression both during normal cell growth and throughout the many changes in environmental conditions that the cell can be subjected to.

Whilst bound to the same RNA strand, the RNAP and the ribosomes each cover an expanse of the mRNA (described in detail in section 1.3.2). The RNAP and ribosomes are usually in fairly close proximity on mRNA that is being simultaneously transcribed and translated (Miller *et al.* 1970). This, along with the interaction between NusE and NusG proteins (Figure 1.11), prevents secondary structures from forming in the nascent RNA. If the RNAP and ribosomes become separated, for instance due to translation being impeded by a lack of charged tRNA during amino acid starvation, an extended region of RNA between the RNAP and the ribosomes can becomes accessible to other enzymes and/or secondary structures can form within the exposed region of the RNA. For instance, polarity and transcription attenuation both lead to premature termination of transcription when transcription and translation become uncoupled. Polarity describes the occurrence of rho-dependent transcription termination due to stalling of translation but not transcription (Richardson *et al.* 1975). As only the rate of translation elongation is affected, whilst the RNAP continues to synthesise the RNA, the region of uncovered RNA between the RNAP and the ribosome becomes extended. If this region of RNA extends to >80nts, Rho can bind to the rut site and terminate transcription prematurely before the RNAP has finished transcribing the operon (Adhya & Gottesman 1978).

The translation elongation factor EF-P has been shown to decrease polarity during transcription and translation at operons encoding proteins that contain a large number of proline residues (Elgamal *et al.* 2016). The presence of poly-proline sequences within the coding region slows the rate of translation elongation as catalysis of the ester bond between proline residues is not as fast compared to that between the majority of the other amino acids (Doerfel *et al.* 2013; Ude *et al.* 2013). By increasing the rate of catalysis of the proline amino acid, EF-P decreases the occurrence of Rho dependent transcription termination due to polarity by decreasing the pausing of the ribosome during translation of proteins containing poly-proline regions (Elgamal *et al.* 2016).

A classic example of transcription attenuation occurs in the *trp* operon leader region in the *E. coli* genome and is thought to be used to 'resynchronise' transcription and translation in the event of translation stalling (Landick *et al.* 1985). In the absence of translation, a hairpin structure forms in the nascent RNA and causes the RNAP to enter into a pause state. If translation resumes, the translating ribosome disrupts the hairpin and allows the RNAP to restart transcription (Landick *et al.* 1985).

The above examples of coupling between transcription and translation are all instances where coupling occurs via factors influencing either transcription and/or translation. However, there are also instances where the coupling occurs through direct interactions between the RNAP and the ribosome. Cooperation between trailing TECs after multiple rounds of transcription initiation has been shown to rescue a stalled or backtracked RNAP (Epshtein *et al.* 2003; Epshtein & Nudler 2003). Subsequently, a trailing ribosome was shown to aid a stalled TEC in overcoming a roadblock and also to restart backtracked RNAP *in vivo* (Schweimer *et al.* 2010). The translating ribosome has also been shown to rescue paused RNAP *in vitro* (Castro-Roa & Zenkin 2012). Conversely, reducing the rate of translation elongation produces a corresponding effect on the rate of transcription elongation of the same mRNA (Schweimer *et al.* 2010). This led to the proposal that the ribosomes are able to slow down the rate of transcription elongation by physically pulling on the RNAP. In support of this, a direct link between transcription and translation via NusE and NusG has been demonstrated using NMR imaging with isolated fragments of NusG and NusE. The NMR data showed the N-terminal of NusG interacted directly with NusE (Burmann *et al.* 2010) (Figure 1.11 A). As the NusG C-terminal is bound to the transcribing RNAP and NusE is associated with the ribosome as the ribosomal

protein S10, this provides a theoretical physical bridge between the ribosome and RNAP (Figure 1.11 B). This interaction however is based purely on NMR data using only the purified interaction domains of NusE and NusG and is only a weak interaction. There is some debate as to whether this data is an accurate reflection of what occurs *in vivo* with the whole proteins. There is also a lack of biochemical data to corroborate this interaction.

As well as the Nus factors NusE and NusG providing a direct link between the RNAP and the ribosome, there is evidence to suggest that NusA could also facilitate transcription-translation coupling by 'resynchronising' the RNAP and ribosome in the event of the release of the NusG-CTD from the RNAP. If the NusG-NTD disengages from the RNAP during transcription/translation but NusG remains in a complex with NusE and the ribosome through the CTD, binding of the NusG-NTD by the AR2 domain of NusA whilst NusA is bound to the α subunit of the RNAP could re-establish the connection between RNAP and NusG and reform the NusE:NusG complex to directly couple translation and transcription (Strauß *et al.* 2016).

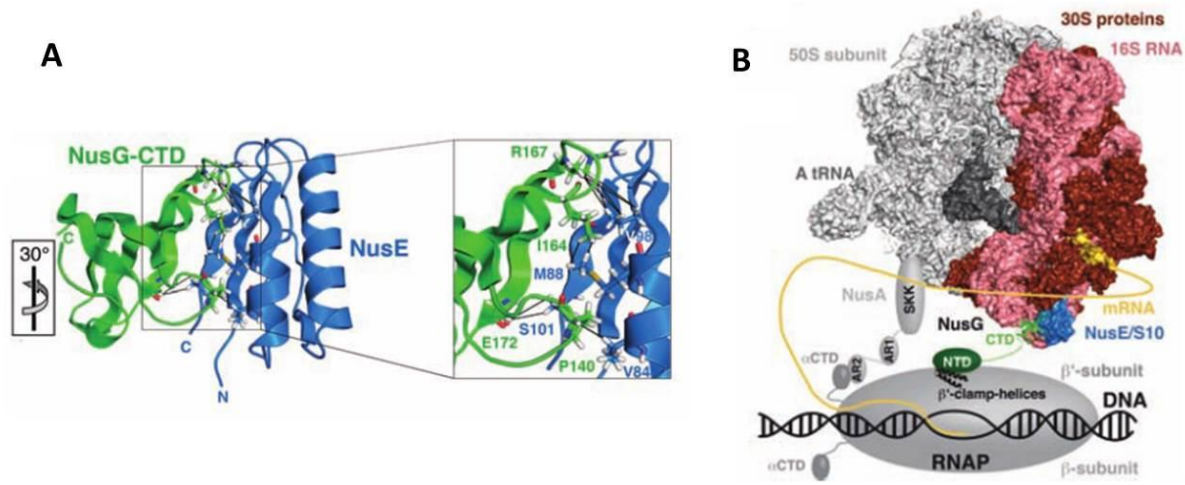


Figure 1.11 Interaction of NusG-CTD and NusE. A) The direct interactions of the NusG-CTD fragment (green) and the NusE fragments (blue). B) Model of the interaction shown as part of the TEC and translation elongation complex. The NusG-NTD (dark green) is bound to the RNAP and the NusG-CTD (light green) is bound to NusE/S10 (blue) whilst in the ribosome bound form (adapted from Burmann *et al.* 2010).

1.3.1 RNAP and ribosomal boundaries

Although it is a well-known fact that the RNAP and ribosomes are often bound to the same nascent RNA strand, it is currently unknown exactly how close they can become. Hydroxyl radical footprinting of the ribosomes revealed that the ribosome

covers a 19 nt region of the mRNA from the first nucleotide of the codon in the P-site to the front edge of the ribosome (or 16 nt from the first nucleotide of the A-site codon) during translation elongation (Hüttenhofer & Noller 1994)(Figure 1.12 A). Based on RNAse digestion of mRNA, the RNAP, on the other hand, covers a 15 nt stretch of RNA from the 3' end of the RNA in the active centre to the rear end of the RNAP (Komissarova & Kashlev 1998) (Figure 1.12 B). Theoretically this means that if the ribosome and RNAP are positioned directly next to each other on the same RNA strand, the first nucleotide of the codon in the ribosomes A-site will be located 31 nucleotides away from the 3' end of the RNA in the RNAP active centre, or 16 nucleotides from the proposed rear end of the RNAP (Figure 1.12 C).

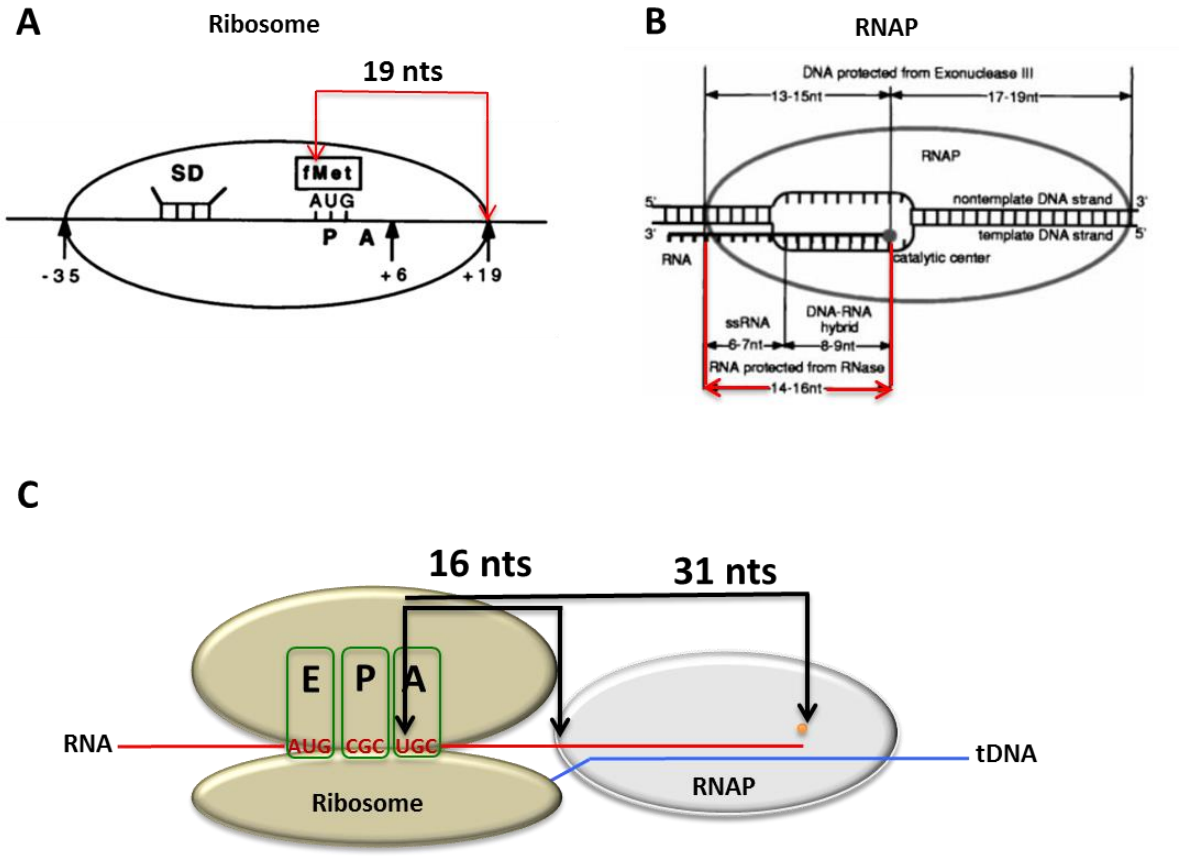


Figure 1.12 RNAP and ribosome boundaries. A) Hydroxyl radical footprinting of the ribosome initiated on the RNA shows a protected region of 19 nucleotides from the first nucelotide of the P-site codon to the ribosome front edge (Hüttenhofer & Noller 1994). B) During transcription elongation the RNAP protects a region of RNA 15 nucleotides from the 3' RNA in the active centre to the predicted rear end of the RNAP (Komissarova & Kashlev 1998). C) Schematic representation of the distance between the the first nucelotide of the codon in the ribosomal A-site and the rear end (16 nts) or active site (31 nts) of the RNAP.

1.3.2 In vitro transcription coupled to translation system

To date, most of the research into transcription and translation coupling has been carried out *in vivo.* Although this is potentially a better reflection of how transcription and translation normally interact within the cell, there are limitations to this approach. For example, in the cell there are numerous accessory factors that interact with the transcription and/or translation machinery that could influence the results of any study looking closely at direct interactions between the RNAP and ribosomes. Our laboratory has previously developed an *in vitro* transcription coupled to translation assay made from only the minimal components necessary for transcription and translation (Castro-Roa & Zenkin 2012; Castro-Roa & Zenkin 2015). This assay has one major advantage over *in vivo* assays in that there is no interference from either known or unknown accessory factors as all the individual components are purified before use.

The transcription coupled to translation *in vitro* technique has two main pathways. *Transcription first coupled transcription to translation* (TR-CTT) is designed to study more closely the effect of transcription upon translation whilst *translation first coupled transcription to translation* (TL-CTT) is designed to examine the effects of translation on transcription. In TR-CTT, an artificially assembled elongation complex (AAEC) is formed to bypass the need for transcription initiation by the RNAP holoenzyme at a double stranded DNA promoter region to produce a TEC. Instead, the AAEC is formed using only an RNA oligonucleotide (oligo), template and non-template DNA and core RNAP enzyme. Importantly, the AAEC formed is as efficient during transcription as the TEC produced after transcription initiated from a promoter (Daube *et al.* 1992; Yuzenkova & Zenkin 2010; Sidorenkov *et al.* 1998). To form the AAEC, a pre-synthesised RNA oligo is annealed to the DNA template by base pairing between two complementary 9 base sequences, one located at the 3’ end of the RNA and the other towards the 5’ end of the DNA template (Figure 1.13 step 1). This forms an RNA:DNA hybrid which is recognised and bound by the RNAP to form a AAEC. Addition of non-template DNA completes the assembly of a stable AAEC (Figure 1.13 step 2). The RNAP is walked along the DNA template by the addition of NTPs to extend the RNA and position the RNAP on the DNA template (Figure 1.13 step 3). The RNAP is modified to contain a biotin tag on the β’ subunit and enables the AAEC to be immobilised on streptavidin beads via the RNAP. Once immobilised, the AAECs are washed with buffer containing a high

salt concentration (1M) to remove all free RNA, NTPs and DNA (Figure 1.13 step 4). Only the RNA that is part of a stable AAEC will be retained by the immobilised RNAP (Figure 1.13 step 5).

After washing of the AAECs, translation is initiated on the RNA bound by the RNAP using ribosomes, the initiator formyl methionine tRNA (fmet-tRNA^{fmet}) and IF-1, 2 and 3 (Figure 1.13 step 6). After translation initiation, ternary complexes (TC) containing the charged tRNA bound by EF-Tu·GTP are added, along with EF-G, to allow the ribosomes to synthesise the peptide encoded within the RNA (Figure 1.13 step 7). Addition of specific TCs influences which peptide the ribosome is able to synthesise based on the peptide-coding region of the RNA.

Once translation elongation has been accomplished, the reaction is stopped, the peptide released from the tRNA (Figure 1.13 step 8) and analysed using thin layer electrophoresis (TLE), which allows separation of peptides based on their size, charge and hydrophobicity. During TLE, the peptide samples are spotted across the middle of a cellulose chromatography plate (Figure 1.13 step 9), which is then placed in buffer and the samples allowed to interact with the mobile and stationary phase, before a charge is applied (Figure 1.13 step 10). The peptides synthesised can be identified based on their migration patterns during TLE (Figure 1.13 step 11). To visualise the peptide, ^{35}S-radiolabelled formyl-methionine is used during translation initiation. Although the RNAP is positioned on the DNA template prior to ribosome binding and translation initiation/elongation, the RNAP can be further manipulated during translation by the addition of NTPs. As transcription is assembled first, the TR-CTT technique can be used to study the effect of transcription on translation and particularly the direct effects of the RNAP on the ribosomes.

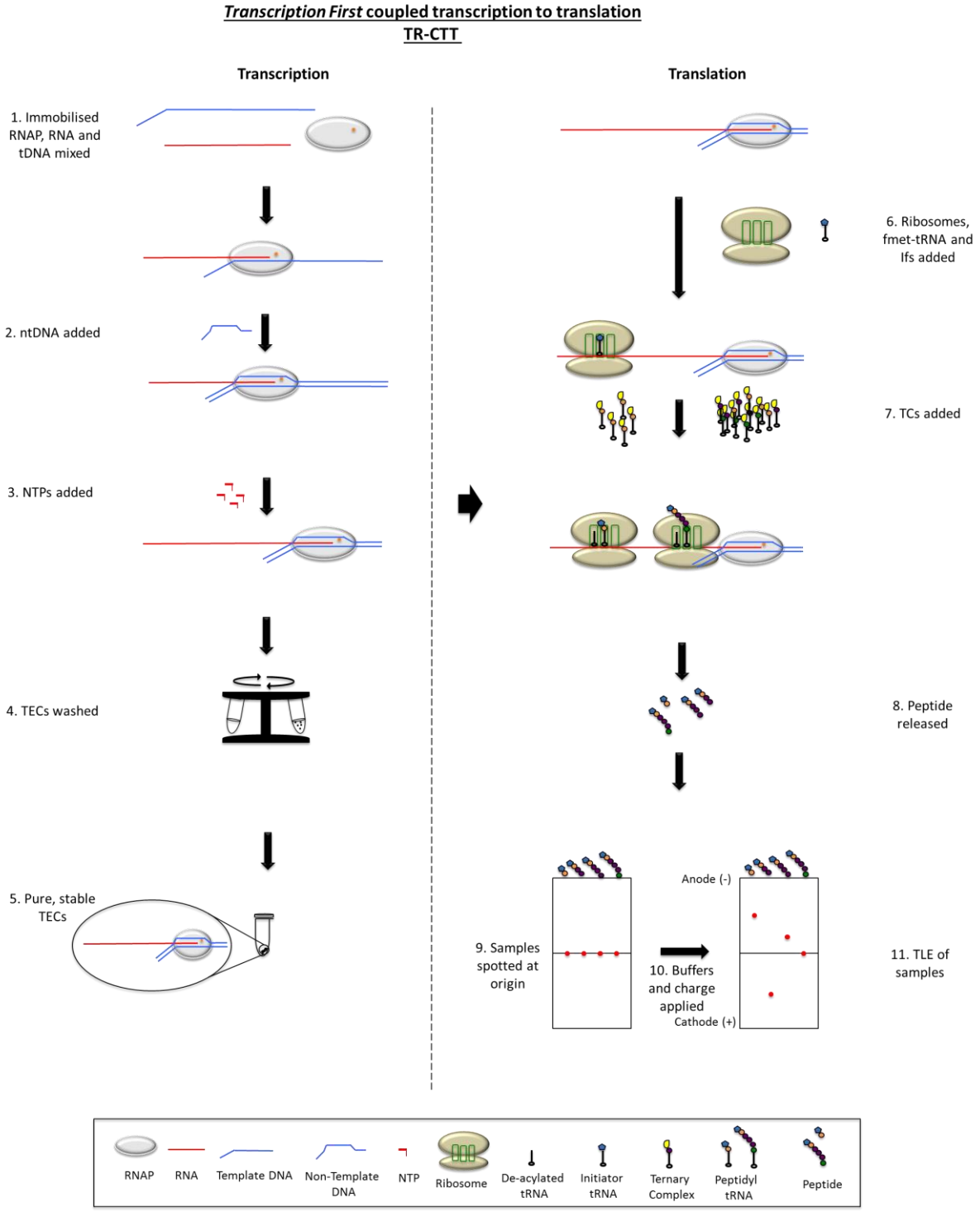


Figure 1.13 Schematic representation of theTranscription first coupled transcription-translation technique (TR-CTT). Left: Assembly of the AAEC and washing of the EC. Right: Translation initiation on the RNA bound by RNAP and elongation, peptide release and analysis by TLE.

TL-CTT is the second *in vitro* coupled technique that our lab has developed and begins with the initiation of translation (Figure 1.14 step 1) and synthesis of a dipeptide made up of the initiator formyl-methionine and the following amino acid (Figure 1.14 step 2). The ribosomes are translocated after dipeptide synthesis to position the third codon in the A-site as the translation elongation complex is more

stable than the initiation complex. The translation reaction is then layered onto a sucrose cushion and centrifuged at high speed (Figure 1.14 step 3). The ribosomes, along with the RNA and peptidyl-tRNA bound by the ribosomes, are able to pass through the sucrose cushion and form a pellet, but the free RNA, GTP (required for translation initiation) and aa-tRNA^{aa} remain on top of (or trapped within) the sucrose cushion. This ensures that all the RNA in the ribosome pellet and therefore subsequently used in transcription, is bound by a ribosome (Figure 1.14 step 4). The ribosome pellet is resuspended and the AAEC is assembled on the RNA by addition of tDNA, RNAP (Figure 1.14 step 5) and ntDNA (Figure 1.14 step 6). Addition of NTPs allows the RNAP to transcribe along the tDNA (Figure 1.14 step 7). Addition of TCs and EF-G after transcription elongation enables the ribosomes to translocate further along the RNA and synthesise a longer peptide (Figure 1.14 step 8). The RNA is radiolabelled at the 5’ end either prior to TL-CTT or at the 3’ end during the transcription stage through incorporation of radiolabelled NTPs by the RNAP. After the reaction is stopped, the RNA is analysed by separation on a denaturing polyacrylamide gel. As translation is initiated first and the end product being visualised in this instance is the RNA, TL-CTT is used primarily to study the effects of translation and the ribosome on transcription and the RNAP.

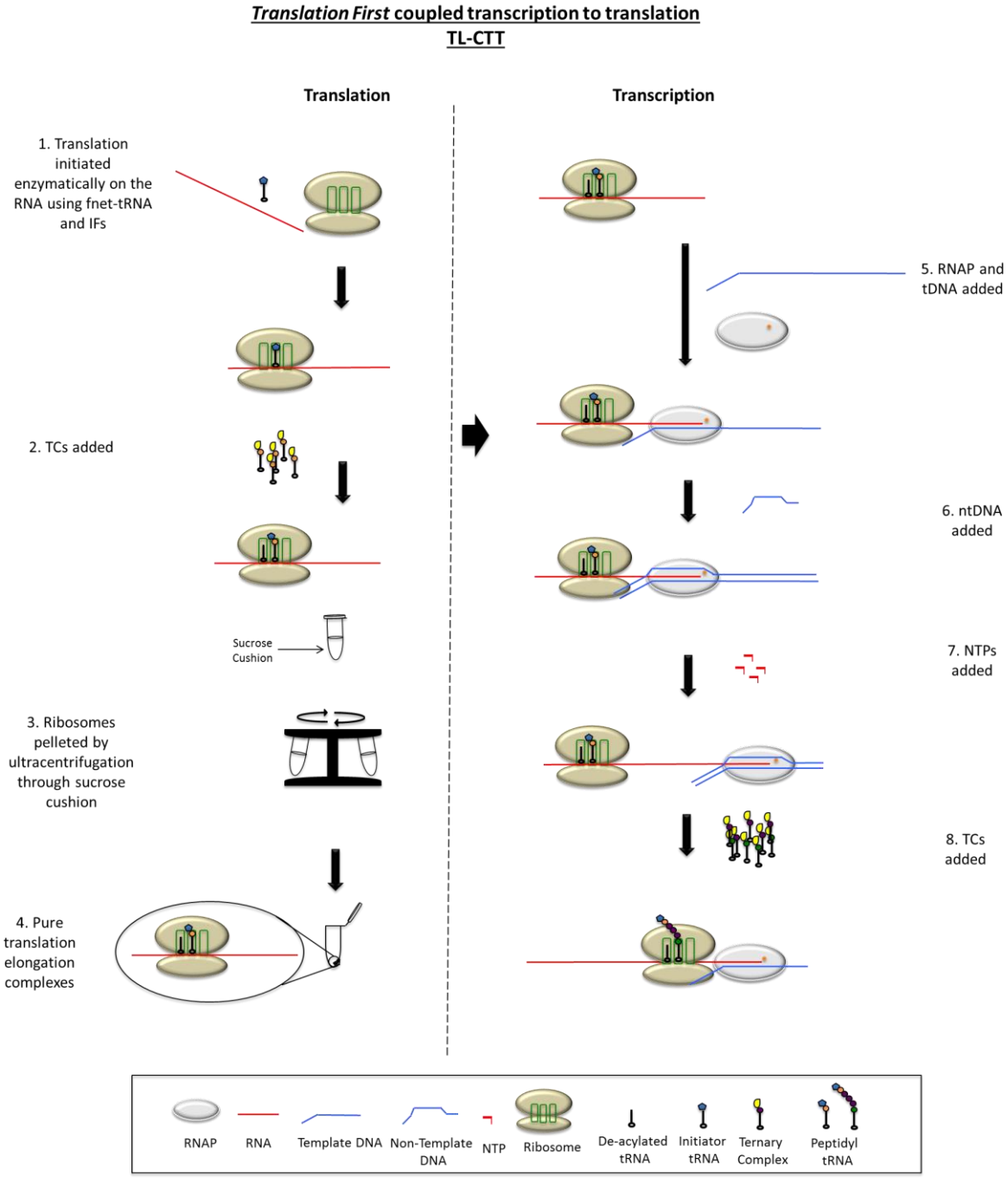


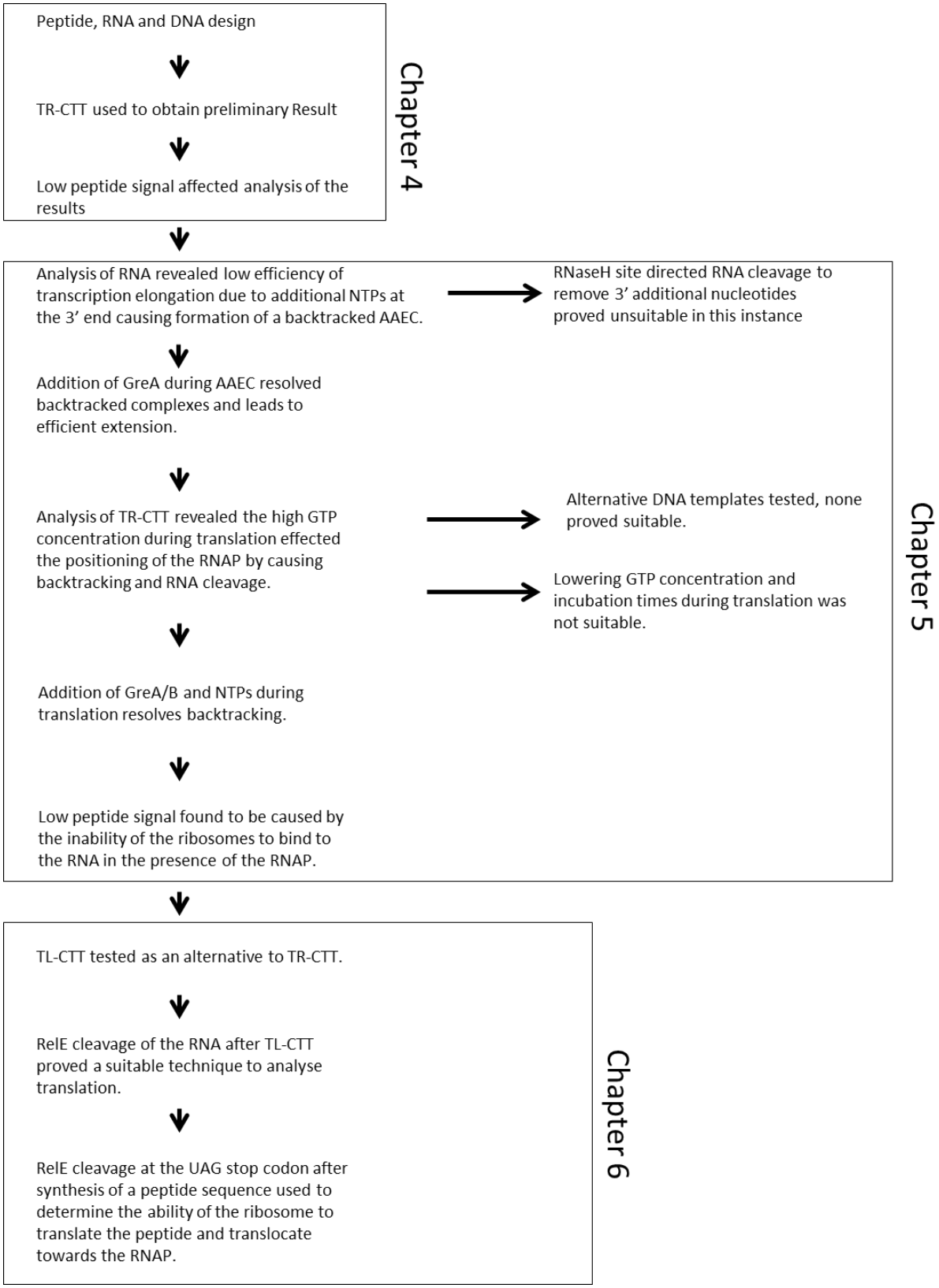
Figure 1.14 Schematic representation of the Translation first coupled transcription to translation technique (TL-CTT). Left: Translation initiation and elongation to form the dipeptide followed by filtering though the sucrose cushion. Right: Assembly of the AAEC on the ribosome bound RNA followed by transcription and translation elongation.

2. Aims

Transcription-translation coupling in bacteria, specifically *E. coli*, is an on-going area of investigation. To-date, most of the research into the interactions between the two has been carried out *in vivo* due to the lack of a suitable *in vitro* system. Studying transcription and translation coupling *in vivo* better reflects the events as they occur in their natural environment, however, the presence of the many transcription/translation factors, both known and unknown, that influence both machineries does not allow the direct, physical interactions to be studied in depth. The design of the transcription-coupled-to-translation system produced within the lab by Castro-Roa and Zenkin (Castro-Roa & Zenkin 2012) now enables a more in depth analysis of coupling in the presence of only the most minimal, pure components required for transcription and translation.

Previously, RNAse digestion and hydroxyl radical footprinting has suggested how many nucleotides of the RNA is covered by the RNAP and the ribosome respectively, but how close to each other the two machineries can become whilst actively transcribing/translating is currently unknown. The aim of this project was to use the coupled *in vitro* technique to determine how close the ribosome was able to translocate towards the RNAP whilst still remaining active. The coupled technique was modified for use by designing suitable peptides, RNA and DNA. A more comprehensive overview of the project is outlined on the following page.

Overview of the Project



3. Materials and Methods

3.1 Protein Purification

3.1.1 RNAP Purification

Grinding Buffer:	50 mM Tris-HCl pH 7.9, 300 mM NaCl, 10 mM EDTA, 10 mM DTT
10 X TGED stock:	100 mM Tris-HCl pH 7.9, 1 mM EDTA, 1 mM DTT and 50% glycerol
Buffer A:	1X TGED, 0.05 M NaCl
Buffer B:	1X TGED 1 M NaCl
Storage Buffer:	20 mM Tris-HCl pH 7.9, 300 mM NaCl, 0.05 mM EDTA, 2mM β-mercaptoethanol, 50% glycerol

Six litres (L) of LB medium was inoculated 1/200 with an overnight culture of NEB T7 express *E. coli* cells transformed with plasmid pIA468 encoding all of the RNAP subunits, including a C-terminal biotin tagged β' subunit (Svetlov & Artsimovitch 2015). The cells were grown in the presence of chloramphenicol (25 µg/ml final concentration) and 50 µg/mL final biotin (Sigma). IPTG was added to a final concentration of 0.4 mM when OD_{600}=0.4 was reached. Once the cells reached an OD_{600} of 1.3, they were harvested by centrifugation at 5,000 rpm using the Beckmann F500 rotor for 7 minutes (min) at 4˚C. The pellet was resuspended in 40 mL of grinding buffer containing two Roche protease inhibitor cocktail tablets, homogenised and 20 µg/mL of lysozyme was added, followed by incubating for 30 min on ice. The cells were lysed by sonication (amplitude of 30%, 1.5 seconds on, 1 second off for a total of 25 min) and centrifuged for 20 min at 15,000 rpm and 4˚C. After centrifugation, polyethyleneimine (Polymin-P, sigma) was added to a final concentration of 0.4% and stirred at 4˚C for 10 min. Following centrifugation at 15,000 rpm for 5 min, the pellet was resuspended by homogenisation in 150 mL of ice-cold 1X TGED with 0.3 M NaCl. The sample was centrifuged at 15,000 rpm with the Beckman JA25.50 rotor for 5 min and the pellet resuspended in 150 mL ice-cold 1X TGED with 0.5 M NaCl, followed by a third centrifugation and a final re-suspension in 150 mL ice-cold 1X TGED with 1.25 M NaCl by stirring at 4˚C for 1 h. After one more centrifugation as previously, the supernatant was collected and ammonium sulphate ($(NH4)_2SO4$) added to a final saturation of 65% to precipitate

the protein by stirring overnight at 4˚C. The precipitated protein was pellet by centrifugation at 15,000 rpm for 30 min and dissolved in 50 mL 1X TGED without salt, then centrifuged again at 15,000 rpm for 30 min. The supernatant was collected, filtered through a 0.45 µM PVDF filter (Merck) and loaded onto a heparin column (Hitrap Heparin HP 5ml, GE healthcare) equilibrated in buffer A. The column was washed with buffer A and RNAP eluted in a stepwise gradient of buffer B (20%, 60% and 100%) with RNAP eluting at 60% buffer B. After the heparin column, the RNAP was further purified using a Superose 6 column (10/30 GL, GE healthcare), using the same buffers as for heparin. The Biotin tagged RNAP was separated in 20% buffer B at a low flow rate of 0.2 mL/min. The fractions containing the core RNAP were combined and further purified using Mono-Q ion exchange column (5/50 GL, GE Healthcare). The sample was loaded at a flow rate of 0.5 mL/min and eluted in a gradient of 30-60% buffer B over 1 hour. The peak fractions were analysed by SDS-PAGE (4-20% Expedeon pre-cast run blue SDS gel, Appendix A11, Figure 9.7 B), concentrated (amicon centrifuge filter, Merck) and dialysed against 2 L of storage buffer at 4˚C overnight before being quantified using BioRad protein assay and stored at -20˚C.

3.1.2 EF-G, IF-1, 2 and 3, FMT, MetRS and EF-Ts

Lysis Buffer:	10 mM Tris-HCl pH7.4, 60 mM NaCl, 10 mM $MgCl_2$, 5% glycerol
Nickel Column Buffer A:	10 mM Tris-HCl pH7.4, 60 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, 10 mM imidazole
Nickel Column Buffer B:	10 mM Tris-HCl pH7.4, 60 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, 200 mM imidazole
Storage Buffer:	10 mM Tris-HCl pH7.4, 60 mM NH_4Cl, 10 mM MgOAc, 6 mM β-mercaptoethanol, 50% glycerol

The strains containing the proteins with an N-terminal His-Tag were taken from the ASKA collection (Kitagawa *et al.* 2005), the plasmids miniprepped (Qiagen) following manufacturers protocol and transformed into NEB T7 express *E. coli* cells. 4 L of LB was inoculated with a 1/500 dilution of overnight culture and the cells grown to OD_{600}=0.4 at 37˚C in the presence of 25 µg/mL chloramphenicol. Expression was induced by adding IPTG to a final concentration of 0.4 mM and the cells were grown for a further 4 h, before harvesting by centrifugation at 5,000 rpm, 4˚C for 7 min.

Samples were taken just prior to, and at 1 h intervals after, induction and analysed by SDS-PAGE to check for overexpression of the protein. The pellet was resuspended in lysis buffer containing two Roche protease inhibitor cocktail tablets, incubated on ice for 30 min with 20 µg/mL lysozyme and sonicated (amplitude 30%, 1.5 seconds on, 1 second off for a total of 25 min) to lyse followed by a clearing spin at 15,000rpm for 20 min at 4˚C. The supernatant was filtered using a 0.45 µM PVDF filter and applied to a Ni-NTA column (HisTrap 5mL, GE healthcare) prewashed in lysis buffer. After loading, the column was washed with buffer A to remove unbound protein. The protein was eluted in a stepwise gradient of buffer B (10%, 30%, 50%, 75% and 100%). The peak fractions were collected and analysed by SDS-PAGE (Appendix A11, Figure 9.7 A). The fractions containing the pure protein were pooled and dialysed overnight at 4˚C into storage buffer and stored at -20˚C.

3.1.3 GreA Purification

Buffer A:	7 M Guanidine-HCl, 40mM Tris-HCl pH 7.5, 0.8 M NaCl, 1 mM EDTA, 1 mM DTT
Buffer B:	40 mM Tris-HCl pH 7.5, 0.8 M NaCl, 1 mM EDTA, 1 mM DTT
Buffer C:	40 mM Tris-HCl pH 7.5, 0.8 M NaCl, 0.6 M imidazole, 1 mM EDTA, 1 mM DTT
Storage Buffer:	40 mM Tris-HCl pH 7.5, 0.8 M NaCl, 1 mM EDTA, 1 mM DTT, 50% Glycerol

C-terminal His-tagged GreA was purified as Koulich *et al.* 1997. GreA was amplified from genomic DNA and inserted into the pET21 plasmid using the primers and restriction sites in Koulich *et al.* 1997. An overnight culture of T7 express cells transformed with pET21-GreA was used to inoculate 2 L of LB media containing 50 µg/mL ampicillin. The cells were grown at 37˚C to OD_{600}=0.4 and induced using 0.4 mM IPTG for 4 hrs, with samples taken before induction and every hour after. The cells were harvested by centrifugation at 5,000 rpm, 4˚C for 7 min and resuspended in 40 mL denaturing buffer A. The cells were centrifuged at 25,000 g for 20 min and the supernatant added to Ni-NTA beads (Quiagen) (1 mL of slurry pre-equilibrated in buffer A) and gently rotated for 20 min at room temperature (RT). The beads were washed three times with 10 mL buffer A using a gravity flow column, then five times with 10ml refolding buffer B. GreA was eluted three times with 2 mL of buffer C. The

fractions containing pure GreA (as shown by SDS-PAGE, Appendix A11, Figure 9.7 A) were dialysed against storage buffer overnight at 4˚C, quantified and stored at -20˚C.

3.1.4 Ribosome Purification

Lysis Buffer:	20 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 150 mM KCl, 30 mM NH_4Cl
Buffer A:	20 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 150 mM KCl, 30 mM NH_4Cl, 5 mM imidazole
Buffer B:	20 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 150 mM KCl, 30 mM NH_4Cl, 150 mM imidazole
Sucrose Cushion:	1.1 M sucrose, 20 mM Tris-HCl pH 7.6, 500 mM NH_4Cl, 10 mM MgOAc, 6 mM β-mercaptoethanol, 0.5 mM EDTA
Buffer C:	20 mM Tris-HCl pH 7.6, 10 mM MgOAc, 100 mM NH_4Cl, 6 mM β-mercaptoethanol, 0.5 mM EDTA
Buffer D:	20 mM Tris-HCl pH 7.6, 10 mM MgOAc, 500 mM NH_4Cl, 6 mM β-mercaptoethanol, 0.5 mM EDTA
Re-suspension Buffer:	50 mM Tris-HCl pH 7.6, 10 mM MgOAc, 100 mM NH_4Cl, 6 mM β-mercaptoethanol, 0.5 mM EDTA

The strain JE28 contains ribosomes with a 6-histidine-tag on the C-terminal of the ribosomal L12 protein (Ederth *et al.* 2009). 6 litres of LB media was inoculated 1/500 with an overnight culture and the cells grown at 37˚C in the presence of 50 µg/mL kanamycin to OD_{600}= 1. The cells were cooled on ice for 1 hr to allow the ribosomes to run off the RNA and then harvested by centrifugation at 4,000 rpm for 30 min at 4˚C. The pellets were resuspended in lysis buffer containing two Roche protease inhibitor cocktail tablets and lysed by passing twice through a French Press cell disruptor at 18 Kpsi. After a clearing spin at 15,000 rpm using a JA-25.50 Beckman rotor for 20 min at 4˚C, the supernatant was filtered through a PVDF 0.45 µM filter, (Merck) and loaded onto a 5 mL Ni-NTA column (GE Healthcare) previously equilibrated in buffer A. After loading, the column was washed with buffer A and the ribosomes eluted in 100% buffer B. The peak fraction was collected and passed through a 35 mL sucrose cushion by ultracentrifugation at 35,000 rpm for 22 h at 4˚C with the Ti-45 Beckman rotor. The supernatant was discarded and the pellets washed with buffer C and resuspended in 10 mL of the same buffer. After a clearing

spin (15,000 rpm for 20 min at 4˚C, Beckman JA25.50 rotor), the supernatant was added to buffer D in a total volume of 100 mL. The ribosomes were re-pelleted by ultracentrifuging at 22,000 rpm for 7 h at 4˚C in a Ti-45 Beckman rotor and the pellets washed as previously with buffer C, then re-suspended in 800 µL re-suspension buffer. The re-suspended ribosomes were aliquoted into 20 µL aliquots, frozen in liquid nitrogen and stored at -20˚C. The ribosomes were quantified by measuring the absorbance at A_{260}. An A_{260} of 1 is equal 20 pmol/mL.

3.1.5 RelE Purification

Lysis Buffer:	50 mM NaH_2PO_4, 0.3 M NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol
Buffer A:	50 mM NaH_2PO_4, 0.3 M NaCl, 35 mM imidazole, 1 mM β-mercaptoethanol
Buffer B:	100 mM NaH_2PO_4, 10 mM Tris-HCl pH 7.9, 9.8 M Urea, 1 mM β-mercaptoethanol
Dialysis Buffer 1:	1X PBS, 1 mM DTT, 0.1% Triton X-100
Dialysis Buffer 2:	1X PBS, 1 mM DTT
Dialysis Buffer 3:	1X PBS, 1 mM DTT, 20% Glycerol

Untagged RelE was purified using the plasmid pKP3077 encoding 6xhis-RelB and the plasmid pKP3067 encoding untagged RelE (Pedersen *et al.* 2002). An overnight culture of T7 express cells containing both plasmids was diluted 1/500 in 2 litres of 2x YT media containing 50 µg/mL chloramphenicol, 100 µg/mL ampicillin and 1mM IPTG. The culture was incubated at 30˚C and at OD_{600}=0.4 a final concentration of 0.2% arabinose added. The cells were grown for a further 4 hrs, before harvesting by centrifugation at 5,000 rpm for 7 min at 4˚C. The pellet was resuspended in 40 mL lysis buffer containing 2 protease inhibitor cocktail tablets (Roche) and 20 µg/mL lysozyme added followed by a further incubation on ice for 30 min. The cells were lysed by sonication (amplitude of 30%, 1.5 seconds on, 1 second off for a total of 25min) then centrifuged at 15,000 rpm for 30 min at 4˚C using the JA25.50 Beckman rotor to pellet any cell debris. The supernatant was added to 5 ml Ni-NTA beads equilibrated in lysis buffer, followed by gentle rotation at 4˚C for 2 hrs. The beads were washed four times on a gravity flow column with two column volumes of buffer A before RelE was removed from the RelE:RelB complex by washing the bound complexes 4 times with 2 ml of denaturing buffer B. The fractions were analysed by

SDS-PAGE (Appendix A11, Figure 9.7 A). and dialysed overnight against 2 L dialysis buffer 1 (the first half hour at room temperature, then 4˚C for the remainder), then for 8 hours against 2 L dialysis buffer 2 at 4˚C and finally overnight against dialysis buffer 3 at 4˚C. RelE is quantified and stored as aliquots flash frozen in liquid nitrogen and placed at -20˚C.

3.2 Transcription

3.2.1 RNA Synthesis and Purification

RNA Synthesis Buffer:	200 mM Tris-HCl pH 7.9, 30 mM $MgCl_2$, 50 mM NaCl, 50 mM DTT, 10 mM Spermidine
TBE buffer:	89 mM Tris, 89 mM Boric Acid and 2 mM EDTA pH 8.0
2X RNA loading dye	90% formamide and 0.02% bromophenol blue

The DNA templates used for RNA synthesis were generated by PCR from a DNA template obtained from Castro-Roa *et al.,* 2012, appendix A1, or a DNA template generated during this work. The Aptataq DNA polymerase (Roche) was used for all PCR reactions. The forward primer (T7Short forward, IDT (appendix A1)) was constant for all the DNA templates and contained the T7 promoter sequence, while the reverse primer (appendix A1 Table 9.1) contained the peptide and RNA:DNA hybrid sequences and was specific for each individual RNA being synthesised. The PCR reaction was assembled according to manufacturer's instructions and the DNA from 8 PCR reactions was pooled, ethanol precipitated and purified from a 2% agarose TBE gel using the Qiagen gel extraction kit and following manufacturer's protocol. The PCR products were sequenced prior to use in RNA synthesis.

The RNA was synthesised by mixing the purified DNA template with 12 pmol of T7 RNAP polymerase in RNA synthesis buffer containing 2mM of each NTP in a final reaction volume of 500 µL. The reaction was incubated for 5.5 hours at 37˚C before 10µl DNAse I (Roche) is added and incubated for a further 30 min. 50 µL 0.5 M EDTA was added to stop the reaction, along with 50 µL of 3 M Sodium Acetate (NaOAc). The RNA was extracted by three chloroform extractions using an equal volume of chloroform and then ethanol precipitated at -20˚C overnight using 2 volumes of 100% ethanol. The sample was centrifuged at 4˚C for 30 min at 15000 rpm, desalted by washing with 70% ethanol, then re-suspended in 60 µL 2X RNA

loading dye. The RNA was run on a 10% denaturing polyacrylamide-sequencing gel with 1X TBE buffer at 40 W for 2 h. The gel was placed between two layers of cling-film and the RNA band visualised by placing the gel on top of a fluorescent TLE plate (Merck) and exposing to UV light. The RNA band was excised and placed into a 0.5 ml clean Eppendorf tube pierced at the bottom with a 0.2 G needle and placed into a 1.5ml Eppendorf tube. This was then centrifuged at 14,000 rpm for 2 min, 200 µL milli-Q water added to the gel pieces and the RNA extracted by incubating at 70˚C for 10 min. The gel slurry was transferred to a Costar Spin-X column and centrifuged at 14,000 rpm for 2 min to remove the gel pieces. 17 µL 3M NaOAc and two volumes of ethanol was added to the flow-through and the RNA was precipitated at -20˚C overnight. The RNA was pelleted by centrifugation at 4˚C for 30 min at 15,000 rpm, desalted by washing with 70% ethanol, then re-suspended in 50 µL 100 mM Tris-HCl, pH 8 and quantified using a nano-drop.

3.2.2 Labelling of RNA using radio-isotopes

3.2.2.1 Labelling RNA at the 5' end

Radiolabelling of RNA at the 5' end with ^{32}P radiolabeled phosphate can be achieved by swapping the 5' γ-ATP with a ^{32}P radiolabelled phosphate. To remove the 5' phosphate, 100 pmol of RNA was incubated with 5 µL10X calf intestinal alkaline phosphatase (CIAP) buffer (NEB) and 1 µL CIAP enzyme in a total volume of 50 µL. The reaction was incubated at 37˚C for 30 min. The reaction volume was increased to 150 µL by the addition of 100 µL of milli-Q water and the RNA extracted by vortexing for 10 min at RT with 150 µL phenol, centrifuging for 5 min at 14,000rpm at RT and performing two chloroform extractions using 150 µL chloroform. The RNA was pelleted and dried using the refrigerated centrivap lab concentrator (Labconco). To add the γ-[^{32}]-phosphate, the dephosphorylated RNA pellet was resuspended in 43.5 µL milli-Q water and 5 µL 10X T4 polynucleotide kinase (PNK) buffer A (thermo scientific), 1 pmol γ-[^{32}P]-ATP (Hartmann Analytic) and 1 µL PNK enzyme added to a final volume of 50 µL. The reaction was incubated at 37˚C for 30 min before the reaction was stopped by addition of 2.5 µL 0.5 M EDTA and heating to 65˚C for 10 min. The enzyme and free ATP were removed by gel filtration using two BioRad P6 columns equilibrated in milli-Q water.

3.2.2.2. Labelling RNA at the 3' end

The RNA can be radiolabelled at the 3' end during transcription by incorporation of radiolabeled α-[^{32}P]-NTPs (Hartmann Analytic) by the RNAP. The radiolabelled NTP was added during transcription elongation, along with any other NTPs required.

3.2.3 RNAseH Site-directed cleavage of RNA

The modified oligonucleotide (IDT) used for RNAseH site directed cleavage of RNA was designed for use with the MFVVVR RNA. The oligo contained four dNTPs at the 5' end, followed by 14 2'o-methly NMPs. The oligo was complementary to the RNA sequence and the 5' dNMPs direct RNAseH to cleave the two (or three if the RNA contained the extra NMP) NMPs at the 3' end of the MFVVVR RNA sequence. 100 pmol RNA and 300 pmol oligo were heated to 95˚C for 3min and then cooled to room temperature. 13 U of RNaseH (thermo scientific) and 5 µL 10X RNaseH buffer were added, the final reaction volume made up to 50 µL with MilliQ water and the reaction incubated at 37˚C for 30 min. To stop the reaction, urea crystals were added to saturation followed by an equal volume of transcription buffer and the sample was heated to 100˚C for 3 min. The sample was loaded onto a pre-run 10% denaturing polyacrylamide gel and the RNA purified as in 3.2.1.

3.2.4. Transcription using artificially assembled transcription elongation complexes (AAEC)

Transcription Buffer:	10 mM Tris-HCl pH7.4, 5 mM $MgCl_2$, 40 mM KCl
High Salt Wash Buffer:	10 mM Tris-HCl pH7.4, 5 mM $MgCl_2$, 1 M NaCl
Transcription Stop Buffer:	20 mM EDTA, 7 M urea, 100 µg/mL heparin, 0.02% bromophenol blue, 0.03% xylencyanol saturated in formamide

The following was the standard protocol for transcription using an artificially assembled elongation complex. Any modifications are explained in the results section. RNA (10 pmol), template DNA (15 pmol, IDT) and biotin-tagged-core-RNAP (20 pmol) were mixed with transcription buffer in a 1:1.5:2 molar ratio and incubated at room temperature for 15 min. If required, 10 µL streptavidin beads (pre-equilibrated in transcription buffer (GE Healthcare)) were added and the reaction incubated at 26˚C with shaking at 1,400 rpm for 10 min to immobilise the biotin-

tagged-core-RNAP. Where relevant, 100 pmol GreA was added before non-template DNA. Non-template DNA (150 pmol) was added in 10 fold molar excess over template DNA and the mixture incubated for a further 15 min at 37˚C. NTPs (GE Healthcare), including any modified NTPs were added at a final concentration of 100 µM and transcription elongation allowed to proceed for 10 min at 37˚C. If the elongation complexes had been immobilised, the reaction was washed once with high salt buffer, then five times in transcription buffer (low salt wash). The reaction was stopped by the addition of an equal volume of transcription stop buffer. The products were separated by PAGE using a 10% sequencing gel, visualised by phosphorimaging and analysed using ImageQuantTL™ software.

3.2.5 Transcriptional Pause Characterisation

3.2.5.1 Pyrophosphorolysis

Pyrophosphorolysis was used to determine if the RNAP was paused in the pre-translocated position. The AAEC was assembled as 3.2.4 and the complexes washed and resuspended to a final volume of 100 µL after transcription elongation. The reaction was divided into 10 µL aliquots and the reactions stopped by the addition of an equal volume of stop buffer 0, 5, 10, 20, 40, 80, 160, 320 or 640 seconds after the addition of pyrophosphate (PPi) to a final concentration of 500 µM (1 µL of 5 mM stock).

3.2.5.1 GreA/GreB Cleavage

GreA and GreB cleavage were used to characterise the halted RNAP to determine if the RNAP was stabilised in the backtracked state. The AAEC was assembled as 3.2.4 and the complexes washed and resuspended to a final volume of 100 µL after transcription elongation. The reaction was divided into 10 µL aliquots and 100 pmol of GreA or 5 pmol of GreB (obtained from laboratory stock) added. The cleavage reaction was stopped by the addition of an equal volume of transcription stop buffer after 0, 5, 10, 20, 40, 80, 160, 320 or 640 seconds.

3.3 Translation

3.3.1 tRNA Purification

3.3.1.1 Purifying total tRNA from culture

Re-suspension Buffer: 0.3 M Sodium Acetate pH 4.5, 10 mM EDTA

Uncharged total tRNA was purified from the *E. coli* MRE600 strain (Varshney *et al.* 1991). 4 mL of overnight culture was pelleted by centrifugation at 3,000g for 5 min at 4˚C and resuspended in 300 µL ice-cold re-suspension buffer. A phenol extraction was performed using 300 µL of cold acid phenol. The cells were vortexed for 30 secs followed by a 60 sec interval a total of 3 times. The sample was centrifuged at 15,000g for 10 min at 4˚C, the aqueous phase transferred to a new Eppendorf tube and 300 µL cold acid phenol added. The sample was vortexed for 60 secs and the centrifuge step repeated, including transfer of the aqueous phase and addition of more cold acid phenol. 1.4 mL of alcohol was added and the nucleic acid precipitated by incubating on ice for 1 hr 15 min. The sample was centrifuged at 15,000 g for 15 min at 4˚C and the supernatant discarded. The pellet was resuspended in 300 µL ice cold 10 mM NaOAc pH 4.5 and 30 µL 8M LiCl added before centrifuging again. The supernatant was removed, frozen in liquid nitrogen and stored at -20˚C. The amount and purity of the sample was determined by measuring the OD at A_{260} and A_{280}.

3.3.1.2. Purifying specific tRNA from total tRNA

2X Hybridisation Buffer: 20 mM Tris-HCl pH 7.6, 1.8 M NaCl, 0.2 mM EDTA
Wash/ Elution Buffer: 10 mM Tris-HCl pH 7.6

To purify specific tRNAs from the total tRNA mix, the technique developed by Yokogawa *et al.* 2010 was used. 20 µL of streptavidin beads in 1X hybridisation buffer were mixed with 500 pmols of biotin tagged DNA probe (IDT) designed for the specific tRNA (Yokogawa *et al.* 2010) (appendix A3 Table 9.3). The DNA was immobilised on the beads by agitating at 1,400 rpm at 26˚C for 5 min. 115 µL of hybridisation buffer was added along with 5 units (U) of total tRNA in a final reaction volume of 300 µL and the tRNA annealed to the DNA probes by heating to 65˚C for 10 min with agitation at 1,400 rpm. A unit of tRNA is defined as the amount of nucleic acid contained in 1 mL and producing an OD=1 when the absorbance is measured at A_{260}.

After annealing, the sample was pelleted by briefly spinning and the supernatant removed. The pellet was washed four times with 400 µL wash buffer at room temperature before the tRNA was eluted twice by adding 200 µL elution buffer, incubating at 65˚C for 5 min, pelleting the beads and removing the supernatant

containing the tRNA. The elutions were flash frozen in liquid nitrogen and stored at -20˚C.

3.3.2 tRNA Aminoacylation

3.3.2.1 fMet-tRNA

The formyl donor group for the initiator fMet- $tRNA^{fMet}$ was prepared directly before amino-acylation of the tRNA, as the formyl donor group was transferred from the very unstable compound N_{10}-formyl-tetrahydrofolate. The stable precursor, N_5-N_{10}-methenyltetrahydrofolate, was prepared in advance by dissolving 25 mg of folinic acid in 2 mL of 50 mM β-mercaptoethanol, adding 220 µL of 1 M HCl and incubating at RT for 3 hours. The reaction was diluted with 1 mL of 100 mM HCl and the N_5-N_{10}-methenyltetrahydrofolate was aliquoted into 200 µL aliquots and stored at -20˚C until needed.

To prepare the N_{10}-formyl-tetrahydrofolate directly before aminoacylation, 10 µL of 1 M Tris-HCl pH 7.9 and 20 µL 1 M KOH was added to a 200 µL aliquot of N_5-N_{10}-methenyltetrahydrofolate and left at RT for 15 min, or until the sample becomes colourless.

3.3.2.2 General aminoacylation procedure

Aminoacylation Buffer: 50 mM HEPES, 10 mM KCl, 1 mM DTT, 20 mM $MgCl_2$

With the exception of the initiator tRNA, the aminoacylation procedure was the same for all the tRNAs. 3 U of uncharged tRNA (sigma Aldrich or tRNA probes) was incubated at 37˚C for 25 min in aminoacylation buffer (20 min for initiator tRNA) with 6 mM amino acid (30 µL ^{35}S-methionine, Hartman analytic for initiator tRNA), 10 mM ATP and 50 pmols of S100 extract (laboratory stock) (plus 50 pmols formyl-methyltransferase and 120 µL neutralised N_{10}-formyl-tetrahydrofolate for amino-acylation of initiator tRNA). The S100 extract is the cytosolic supernatant fraction that contains the aminoacyl synthetase enzymes.

The reaction was stopped by the addition of 5 µL 3 M NaOAc (plus 28 µL 10% SDS for initiator tRNA) and 500 µL phenol followed by vortexing for 10 min at RT. The sample was centrifuged at 14,000 rpm for 5 min in a table top centrifuge at RT and the supernatant collected. 500 µL of 300 mM NaOAc was added to the bottom fraction and the sample vortexed for 30 sec, centrifuged at 14,000 rpm for 2 min and the supernatant collected. Two chloroform extractions using 500 µL of chloroform

were performed on each of the two supernatants, followed by ethanol precipitation with 1.3 mL of ethanol. The sample was centrifuged at 4˚C for 30 min at 15,000 rpm in a table top centrifuge, desalted by washing with 70% ethanol, then re-suspended in 50 µL 3 mM NaOAc (plus 1 mM DTT for initiator tRNA). The charged tRNA was filtered through four BioRad Bio Spin columns (equilibrated in 3 mM NaOAc (plus 1 mM DTT for initiator tRNA) by washing four times with buffer) to remove any remaining ATP and stored at -20˚C. The charged tRNA was tested in translation (or thin layer chromatography (TLC) for the initiator tRNA).

3.3.2.3 Thin layer chromatography analysis of tRNA aminoacylation

TLC Buffer: 2.5 mL water, 2.5mL glacial acetic acid and 10mL butanol

The efficiency of the fMet- $tRNA^{fMet}$ was assessed by thin layer chromatography by comparing the amino-acylated versus the de-acylated forms. 2 µL of the fMet- $tRNA^{fMet}$ was added to 5 µL 3 mM NaOAc and 1 mM DTT (from previous reaction). The amino-acylated sample was left on ice, whilst 2 µL of 15% NaOH was added to de-acylate the second sample by incubating at 55˚C for 20 min. 2.5 µL of each sample was dotted onto an aluminium backed silica chromatography plate (Sigma) 1.5cm from the bottom, dried and placed in TLC buffer for 1hr or until the aqueous phase reached a few centimetres from the top of the chromatography plate. The plate was dried and visualised by exposing to a storage phosphor screen (GE healthcare) overnight.

3.3.3 Translation

Translation Buffer: 10 mM Tris-HCl pH 7.4, 60 mM NH_4Cl, 10 mM $Mg(OAc)_2$, 6 mM β-mercaptoethanol

Ternary Complex Buffer: 50 mM Tris-HCl pH 7.4, 40 mM NH_4Cl, 10 mM $MgCl_2$, 1 mM DT

The following was the standard protocol for translation. Any modifications were explained in the results section. 50 pmol of ribosomes were added to 100 pmol RNA in translation buffer with 2.4 mM GTP in a final volume of 42 µL. The ribosome binding reaction was incubated at 37˚C for 10 min, before 252 pmol of fmet-$tRNA^{fmet}$ was added along with 210 pmol each of IF-1, 2 and 3 and the GTP concentration

was increased to 3.4 mM with the final volume now 54 µL. The initiation reaction was incubated at 37˚C for a further 15 min whilst the ternary complexes were formed.

The ternary complexes were prepared by first exchanging GDP for GTP on EF-Tu by incubating 400pmol ET-Tu:GDP for 10 min at 37˚C with 200 µg/mL phosphoenol pyruvate kinase (PK) and 60 pmols EF-Ts in ternary complex buffer , 1.4 mM GTP and 800 pmol phosphoenol pyruvate (PEP) in a total volume of 32 µL. 6 µL of aa-tRNA^{aa} was added and the reaction incubated for a further 5 min for the ternary complex (EF-Tu:GTP:aa-tRNA^{aa}) to form. For translation elongation, 210 pmol EF-G, 6 µL of each ternary complex and 7.5 µL of the initiation reaction was mixed in with a final GTP concentration of 4.3 mM in 1X translation buffer in a final volume of 42 µL. The translation elongation mix was incubated at 37˚C for 10 min. The reaction was stopped by the addition of 4.5 µL of 1M KOH and a further incubation at 37˚C for 25 min to de-acylate the peptidyl-tRNA and release the peptide.

3.3.3.1 Thin Layer Electrophoresis

Pyrac Buffer: 200 mL Acetic acid and 5 mL pyridine in one L of water

The peptide produced by translation elongation was assessed by thin layer electrophoresis. 2 µL of each sample was spotted across the middle of a nitro-cellulose backed thin layer chromatography plate (17 cm length, 7 cm width, TLC cellulose F, Merck). The plate was placed in a gel electrophoresis chamber with each end in 30 mL pyrac buffer. The plate was left for 25 min to allow the buffer time to cover the plate, then the chamber was filled with Stoddard solvent and a charge applied (1300 V). The length of time of electrophoresis depended on the peptide being analysed. The plate was dried then visualised by exposing to a phosphor screen overnight.

3.3.4 GTP Hydrolysis Assay

Stop Buffer: 5% formic acid in 1X translation buffer

GTP hydrolysis during translation was analysed by making a stock of GTP to be used in translation consisting of 1 µL α-[^{32}P]-GTP in 20 µL of 100 mM unlabelled GTP. The standard protocol for translation was followed except no PEP or PK was used to prevent GTP regeneration and the 100 mM GTP stock containing α-^{32}P-GTP is used for all reactions. 5 µL samples were taken at various time points and an equal

volume of stop buffer was added. The samples were analysed by thin layer chromatography with 2.5 µL of 10 mM samples of GTP, GDP and GMP run alongside as controls.

3.3.4.1 Analysis of GTP Hydrolysis Assay using Thin Layer Chromatography

Running Buffer: 1.3 M KH_2PO_3 pH 4.

1 µL of each sample was spotted across a TLC PEI cellulose F plate (Merck) 1.5 cm from the bottom and air-dried. The TLC plate was placed in enough running buffer to cover the bottom of the container and left at room temperature for 1 h or until the buffer front reached 1 cm from the top of the plate. The plate was air-dried, the control samples visualised by UV light and marked by spotting α-[^{32}P]-GTP onto the plate. The plate was then air-dried again, wrapped in cling film and exposed onto a phosphor screen (GE healthcare).

3.4 Transcription First coupled Transcription-Translation (TR-CTT)

Transcription was assembled as for transcription alone, except the amount of RNA (200 pmol), DNA (300 pmol tDNA, 3000 pmol nt DNA), RNAP (400 pmol) and GreA (400 pmol, where applicable) were increased. After transcription elongation, immobilisation and washing of the AAECs, the reaction was resuspended to a final volume of 15 µL. Translation was assembled as in translation alone, with the 15 µL transcription reaction used in place of the RNA. GreA was added during translation elongation where specified. After translation elongation and de-acylating the peptidyl-tRNA, the reaction was dried in the Labconco refrigerated centrivap concentrator for 40 min at room temperature. The pellet was resuspended in 10 µL translation buffer and analysed using TLE as described previously.

3.5 Translation First Coupled Transcription-Translation (TL-CTT)

Sucrose Cushion: 1.2 M Sucrose, 20 mM HEPES pH 7.4, 500 mM NH_4Cl, 10 mM $MgCl_2$, 0.5 mM EDTA

Translation was initiated as in translation alone and the dipeptide synthesised. The entire elongation reaction was layered onto a 1.3 mL sucrose cushion and centrifuged at 78,000 rpm at 4˚C for 2 hrs using the JLA 100.3 Beckman rotor. Only the ribosomes and associated RNA bound in the translation elongation complex would pass through the sucrose cushion and form a pellet. The pellet was washed

twice with ice-cold translation buffer and resuspended in 10 µL of the same buffer. Transcription was assembled on the resuspended ribosome bound RNA using one pellet per transcription reaction and the same ratios of RNA:DNA:RNAP as for transcription alone. GreA was added where required. The ribosomes were further translocated and a longer peptide synthesised by addition of specific TCs and EF-G after the transcription reaction if required. The amounts of TC and EF-G used were the same as for translation alone and the reaction was stopped by the addition of an equal volume of transcription stop buffer. Where applicable, GreB was added during translation elongation. The RNA products were analysed as for transcription alone, The RNA was visualised by either incorporating radiolabelled NTPs during transcription elongation, or by labelling the RNA at the 5' end prior to TL-CTT.

3.5.1 Calculating RelE cleavage efficiency

To determine the % of RelE cleavage of both the valine stop codon after MF dipeptide synthesis and the UAG stop codon after full length peptide synthesis, ImageQuant™ software was used. The RNA products were quantified and the RelE cleavage products taken as a percentage of the sum of the full length RNA product plus the RelE cleavage product. To compare the efficiency of RelE cleavage between the UAG stop codon and the Val codon for each RNA, the percentage cleavage of the UAG stop codon was taken as percentage of the cleavage of the Val codon.

4. Setting up transcription first coupled transcription-coupled-to-translation: RNA, peptide and DNA template design

4.1 Introduction

The transcription-coupled-to-translation system designed and developed in our lab (Castro-Roa & Zenkin 2012; Castro-Roa & Zenkin 2015) is an *in vitro* system based on *E. coli* and made up of only the minimal components necessary for transcription and translation. All the components added are purified before use therefore this technique can be used to facilitate the study of the interactions between RNAP and the ribosomes without interference from all other factors that influence gene expression in the cell. As described in section 1.3.1, the coupled system has two main techniques, translation first (TL-CTT) and transcription first (TR-CTT). During transcription first coupled transcription to translation (TR-CTT), the artificially assembled transcription elongation complex (AAEC) is assembled first and the RNAP positioned on the DNA template by adding specific NTPs into the system. The RNAP is immobilised on streptavidin beads and the AAECs washed to remove all free RNA, DNA, NTPs and RNAP. This ensures all the RNA remaining in the system is bound by RNAP. Translation is initiated on the RNA bound by the RNAP during transcription and used by the ribosomes as a template to synthesise the peptide. The peptide is radiolabelled using ^{35}S-methionine during translation initiation and analysed by separation using thin layer electrophoresis (TLE) after the translation elongation reaction.

For translation first coupled transcription to translation (TL-CTT), the translation elongation complex is initiated on the RNA and the reaction filtered through a sucrose cushion by ultracentrifugation. Only the ribosomes (and the associated RNA and peptide) can migrate through the cushion, ensuring that the only RNA in the reaction during the transcription stage is bound by a ribosome. The AAEC is assembled on this RNA and the RNAP walked by the addition of NTPs. The ribosomes can also be further translocated by the addition of charged tRNA and elongation factors. TL-CTT was designed primarily to study the effect of the ribosomes/translation on RNAP/transcription and the RNA is analysed by separation with denaturing PAGE. The RNA is visualised by either radiolabeling prior to use in TL-CTT or by incorporation of radiolabelled NTPs during the transcription stage.

The length of RNA predicted to be covered by both the RNAP and the ribosome if they were to be located directly next to each other on the same nascent

RNA has been proposed to be 31 nts from the 1st nt of the codon in the ribosomal A-site to the RNAP active centre (Figure 1.12 C). This prediction is based on RNAse digestion of the TEC (Komissarova & Kashlev 1998) and hydroxyl radical footprinting of the translating ribosome (Hüttenhofer & Noller 1994). It is currently unknown, however, how close the RNAP and the ribosome are actually able to become during active transcription and translation. Preliminary, unpublished data from our lab obtained by Dr. Daniel Castro Roa suggested that the ribosome was able to translocate closer to the RNAP than deduced by the hydroxyl radical footprinting and RNase digestion. This data was obtained using TR-CTT, however the peptide used for analysis degraded rapidly and the signal of the synthesised peptide was low, therefore it was difficult to obtain a conclusive result.

4.1.2 Aims and Objectives

The aim of this chapter was to optimise the *in vitro* TR-CTT technique to determine how close the transcribing RNAP and translating ribosome can become.

4.2 RNA design and Peptide identification

To investigate how close the translating ribosome can become to the transcribing RNAP, the ability of the ribosome to synthesise a peptide product in the presence of the RNAP will be determined. The TR-CTT method was chosen as it is designed for studying the effect of the RNAP on the ribosome due to the order in which transcription and translation are assembled. The basic principle is to position the RNAP at a specific location on the DNA template, immobilise and wash the AAECs to remove free RNA (and NTPs) and then assemble translation on the RNA bound by the RNAP. Once translation is initiated, adding charged tRNA along with elongation factors will allow the ribosomes to synthesise a six amino acid peptide product encoded in the RNA. The peptide produced by the ribosome is then analysed by TLE. The migration of the peptides during TLE is dependent on the size, hydrophobicity and overall charge of the peptide. By positioning the RNAP at varying distances from the peptide-coding region and then analysing the ribosomes ability to synthesise the full peptide product, the distance required between the RNAP and the ribosome for productive translation can be determined. A peptide six amino acids in length was chosen for analysis after TR-CTT as this allows enough space between the ribosomes and RNAP for translation initiation but ensures the peptide was small enough for separation by TLE. If the ribosome was able to incorporate the 6th amino

acid into the peptide in the presence of the RNAP, the distance between the two was deemed to be ‘permissive’, whereas if the ribosome was unable to incorporate the final amino acid, this distance was deemed to be ‘non-permissive (Figure 4.1). For each TR-CTT reaction, synthesis of the dipeptide product was used alongside as a control of translation initiation. Depending on the exact amino acid sequence of the peptide, intermediary products were or were not visible.

Although the TR-CT technique had already been established in our lab (Castro-Roa & Zenkin 2012; Castro-Roa & Zenkin 2015), it still needed to be optimised for use in this particular project. In particular, the peptide sequence, RNA sequence and DNA templates needed to be well characterised and the RNA and peptide products easily analysed. The peptide needed to satisfy a number of requirements: first, the peptide needed to be 6 amino acids long, second, it had to be made up of a number of different amino acids that enable the ribosome to be progressively walked along the RNA template towards the RNAP, and third, the peptide had to consist of amino acids whose uncharged tRNAs were commercially available. Most importantly, however, there had to be a pronounced difference between the five amino acid peptide (penta-peptide) and the six amino acid peptide (hexa-peptide) when analysed by TLE. This last requirement was essential as the outcome of the effect of transcription on translation was determined by the ability of the ribosome to incorporate the 6^{th} amino acid into the peptide chain.

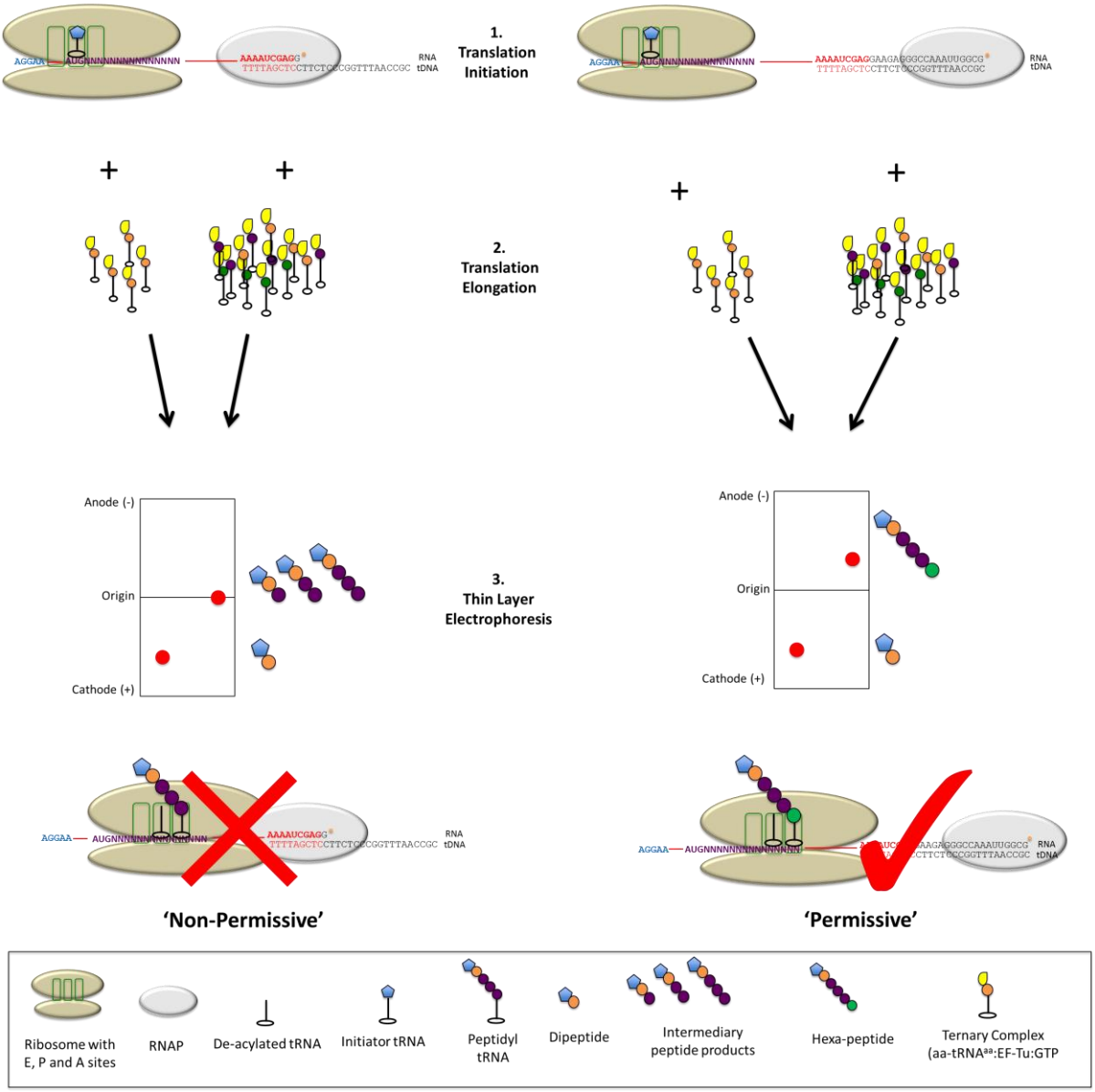


Figure 4.1 Schematic representation of permissive and non-permissive distances. 1) Translation is initiated on the RNAP bound RNA. Although omitted for simplicity, non-template DNA is also present in the transcription elongation complex. 2) Ternary complexes are added to translocate the ribosome and synthesise the dipeptide or hexa peptide. 3) The peptide products are analysed by TLE. Samples are spotted at the origin and then buffer and a charge are applied. The dipeptide control indicates that the ribosome has initiated on the RNA and the presence of absence of the hexapeptide indicates if the ribosome is able to incorporate the $6^{t,h}$ amino acid (permissive) or not (non-permissive). Depending on the composition of the intermediate peptide products, there may or may not be a visible signal if only intermediates are synthesised.

The RNA used in our transcription and translation protocol was synthesised by the single subunit T7 RNAP from a double stranded DNA template generated by PCR using a forward primer containing the T7 promoter region and a reverse primer containing the hybrid sequence and the peptide-coding region (Figure 4.2 A). The DNA template and primer sequences for the PCR reaction are shown in appendix A1. After synthesis, the RNA was purified by gel extraction (Figure 4.2 A). The RNA sequence was designed based on the RNA previously used in our lab (Castro-Roa &

Zenkin 2012; Castro-Roa & Zenkin 2015). The 70 nt long RNA consists of a translation initiation region (TIR), a peptide coding region and a hybrid site (Figure 4.2). The TIR contains the shine dalgarno (SD) site for recognition by the ribosome, in this case the AGGAA sequence positioned 11 nts upstream of the AUG translation start codon (Figure 4.2 B), followed by the peptide coding region. The hybrid site contains a 3' 9 nt sequence complementary to a 9 nt region on the DNA template that formed the RNA:DNA hybrid essential for assembly of the transcription elongation complex (Figure 4.2 B).

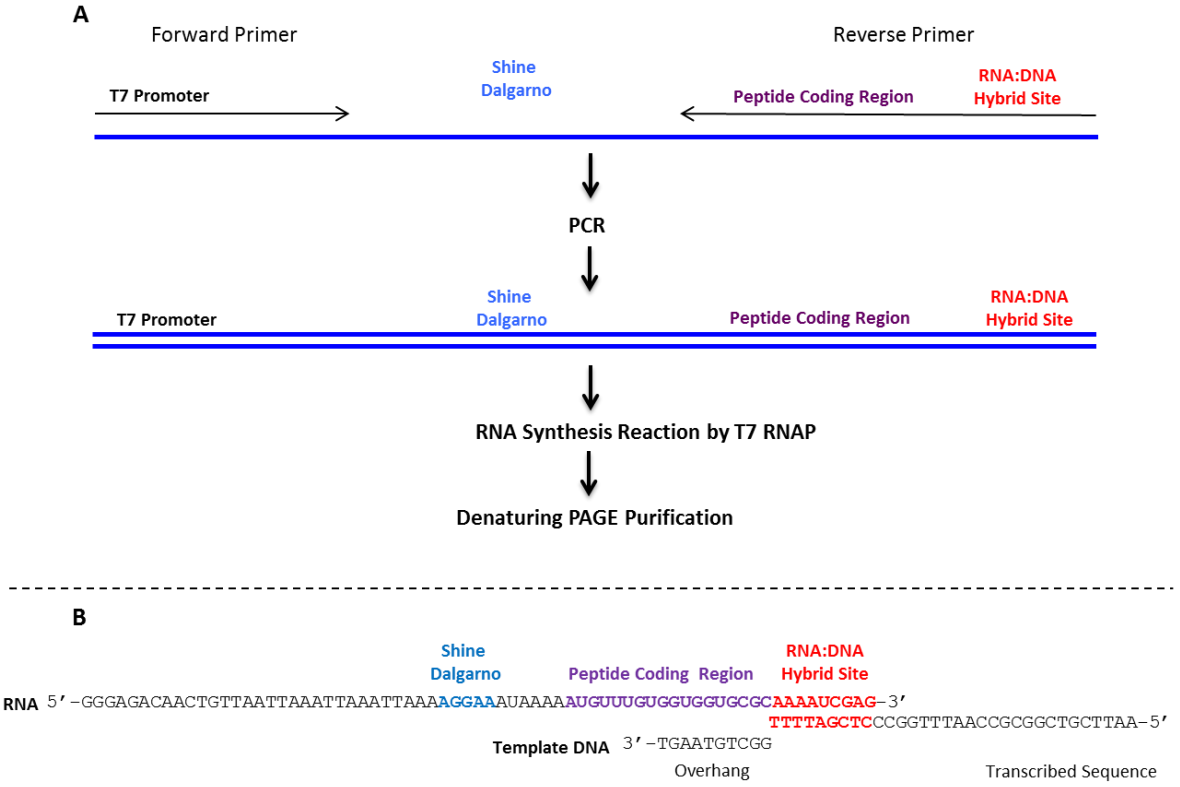


Figure 4.2 RNA Synthesis Schematic representation of RNA synthesis. A) A PCR reaction was used to generate double stranded DNA using a forward primer containing the T7 promoter and a reverse primer containing the peptide coding region and RNA:DNA hybrid. The double stranded DNA was used by the T7 RNAP to synthesise the RNA. The synthesised RNA was then purified by gel extraction. B) The RNA sequence is shown with the SD site (AGGAA), the RNA:DNA hybrid site and the peptide coding region indicated. An example DNA template sequence is shown with the 3' overhang, RNA:DNA complementary hybrid and the transcribed sequence used to position the RNAP also indicated. The non-template DNA is fully complementary to the template sequence.

In our assay, translation is initiated by mixing RNA and ribosomes to allow the ribosomes to bind the SD sequence (Figure 4.3, step 1). Initiator-formyl-methionine (fMet-tRNA^{fMet}) and initiation factors (IF-1, 2 and 3) are added to complete initiation (Figure 4.3, step 2). All charged tRNAs are aminoacylated in a separate aminoacylation reaction prior to use in translation and stored at -20°C. Whilst the translation initiation reaction is incubating, the ternary complexes (TC) containing the aa-tRNA^{aa} are prepared. EF-Tu is purified and stored in the GDP bound, inactive

form and is activated during translation by incubating with GTP and the guanine nucleotide exchange factor, EF-Ts, to replace the GDP with GTP (Figure 4.3, step 3). aa-tRNA^{aa} is added to the active EF-Tu:GTP reaction to form the TC (EF:Tu:GTP:aa-tRNA^{aa}) (Figure 4.3, step 4). Each aa-tRNA^{aa} is added to a separate TC reaction. Once translation has been initiated and the TCs formed, the TCs are mixed with the translation initiation reaction along with EF-G and GTP to allow the ribosome to synthesise the peptide encoded in the RNA (Figure 4.3, step 5 and 6). Addition of specific TCs into the elongation reaction determines the peptide that the ribosome is able to synthesise based on the amino acid sequence encoded in the RNA. For instance adding only phe-TCs into a translation elongation reaction with RNA encoding the peptide MFV will only allow the ribosomes to synthesise the MF dipeptide. After translation elongation, the peptide is released from the peptidyl-tRNA by addition of KOH to increase the pH enough to cause spontaneous hydrolysis of the ester bond between the peptidyl-tRNA and the peptide (Castro-Roa & Zenkin 2012; Castro-Roa & Zenkin 2015). The reaction is spotted onto a chromatography plate and analysed by separation using TLE (Castro-Roa & Zenkin 2012; Castro-Roa & Zenkin 2015). The peptides produced during translation are identified from the migration patterns of the intermediary and full-length peptides.

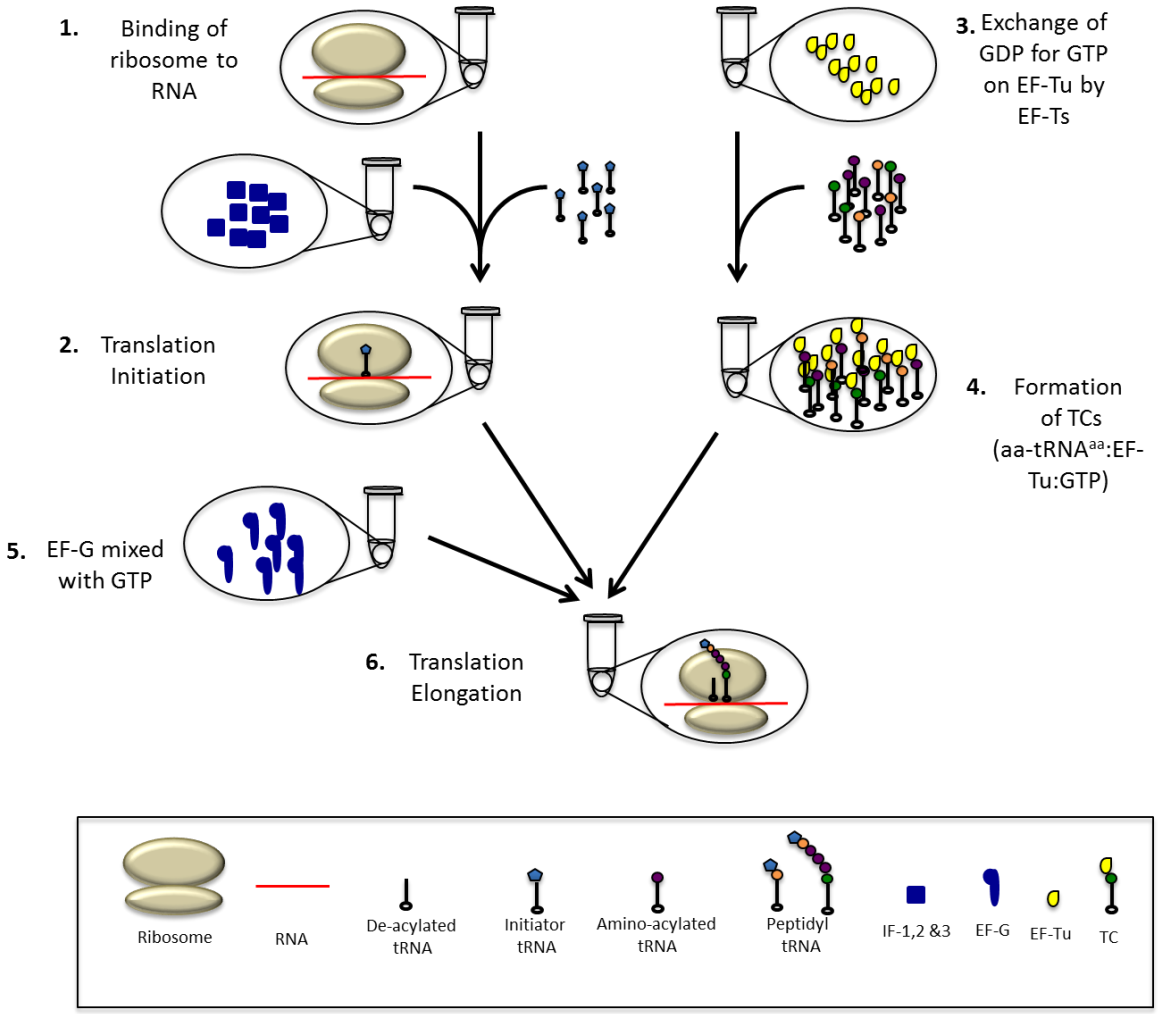


Figure 4.3 Schematic representation of translation on free RNA. The assembly of translation on free RNA is shown. Translation is initiated whilst the TCs are formed, then both reactions are combined along with EF-G and GTP for translation elongation. The specific TCs added into the elongation reaction determine the peptide that the ribosomes synthesised based on the amino acid sequence encoded in the RNA.

The first step in the optimisation of TR-CTT was to identify a suitable peptide. As formyl-methionine (fMet) was required for initiation of translation, methionine was naturally the first amino acid in the peptide chain. Methionine radiolabelled with ^{35}S was used to visualise the peptide on the TLE plate. Phenylalanine (Phe, F) incorporated after methionine produced a peptide with a distinct migration pattern when analysed on a TLE (Figure 4.4 A lanes 1, 4, 7 10 and 13). When valine (Val, V) was added next, there was another clear shift in the migration pattern on the TLE (Figure 4.4 A lane 2). The purified uncharged tRNAs for initiator formyl-methionine, phenylalanine and valine were commercially available and as such, all three were suitable for use in the analysis of translation. Multiple Phe and Val amino acids can be used to extend the amino acid sequence. As previously mentioned, the most important characteristic of the peptide needed for our experiments was the difference

in migration during TLE between the 5 and 6 amino acid long peptides. Adding a charged amino acid as the final amino acid to the peptide chain generally leads to a pronounced shift, as the charge applied across the chromatography plate will have a greater effect on the migration of the peptide. First we tested glutamate/glutamic acid (Glu, E), an amino acid with an overall negative charge and readily available pure uncharged tRNA. When one or more glutamate residues were incorporated at the end of a tri-peptide made up of the MFV residues, the migration pattern of the peptide on the TLE was altered (Figure 4.4 A). Although the peptides with one, two and three glutamates (MFVE, MFVEE and MFVEEE respectively) had a slightly different migration pattern compared to the MFV peptide, the difference was not distinct enough (Figure 4.4 A, lanes 3, 6 and 9 and 2, 5 and 8). The addition of a glutamate residue to the end of a MFVVV peptide produced no change in migration pattern between MFVVV and MFVVVE (Figure 4.4 A lane 11 and 12 respectively). While the ME, MEF and MEFV peptides displayed a difference between their migration patterns, the change on addition of each successive amino acid was again not pronounced enough (Figure 4.4 A lanes 13-15). Therefore glutamate was not a suitable amino acid for use in our experiments, as it did not result in an obvious difference in migration between peptides when incorporated either within the peptide or as the final amino acid.

We then tested a second amino acid, arginine (Arg, R). Arginine is another charged amino acid whose uncharged tRNA is readily available, but in contrast to glutamate, arginine has an overall positive charge. This could cause an arginine containing peptide to migrate towards the anode, the opposite way to MF and MFV and potentially lead to a very distinct shift when added as the final amino acid in the hexapeptide. We therefore produced arginine-containing peptides in translation and analysed by TLE. As expected, incorporation of one arginine residue onto the end of a MFV peptide showed a pronounced shift on the TLE (Figure 4.4 B lanes 2 and 3).

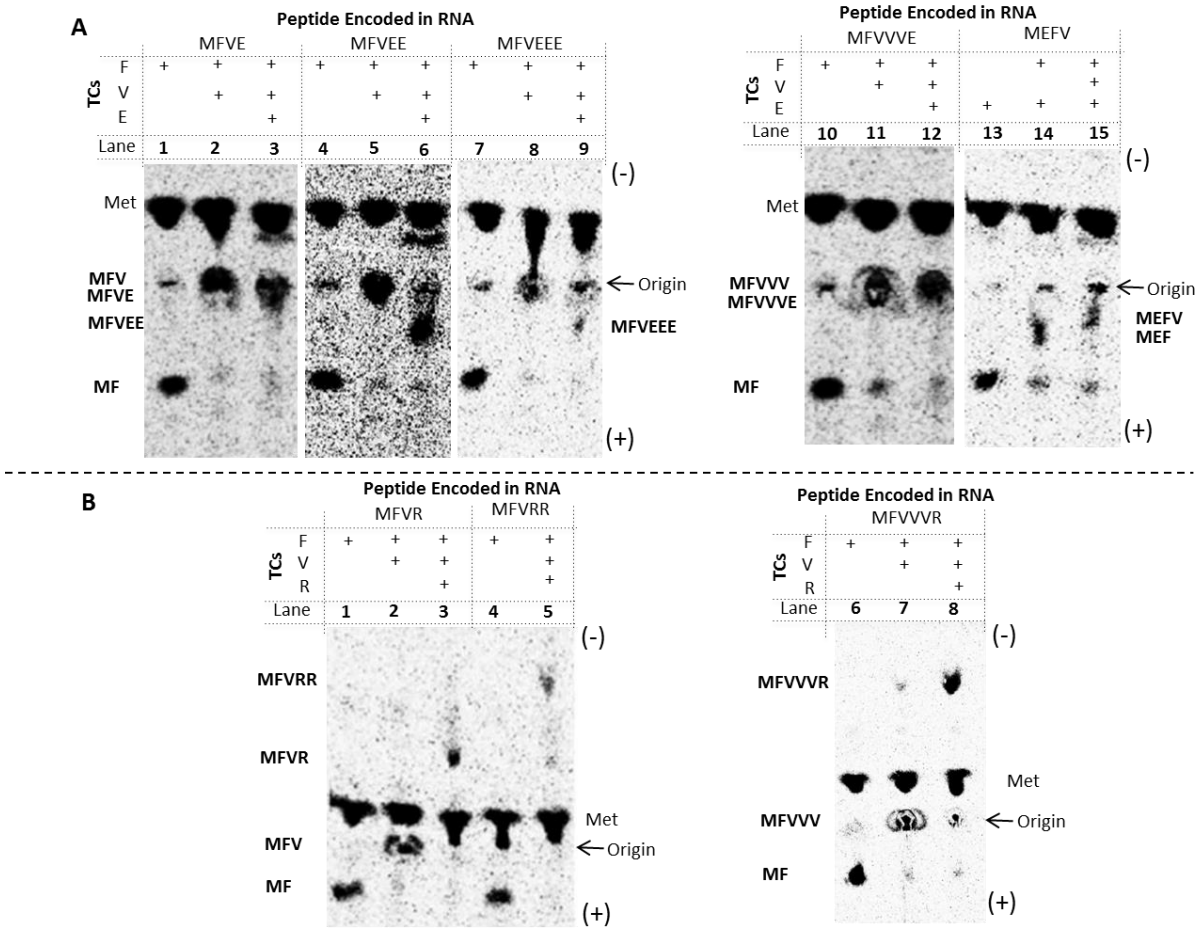


Figure 4.4 Peptide, RNA and DNA template design. A) TLE analysis of glutamate containing peptides. Lanes 1, 4, 7, 10 and 13 show the MF di-peptide control for each RNA template. Lanes 2, 5 and 8 show the MFV tri-peptide control. Lane 3 shows the MFVE peptide, lane 6 the MFVEE and lane 9 the MFVEEE. Comparison between lanes 3, 6 and 9 shows that there is little difference in migration of the peptide upon addition of each successive glutamate residue. Comparison between lane 11 (MFVVV) and 12 (MFVVVE) reveals that addition of the glutamate residue has little effect on the migration pattern between the penta and hexa-peptide. Lanes 13, 14 and 15 show the migration pattern of the MEFV peptide on addition of each amino acid. B) TLE analysis of arginine containing peptides. Lanes 1, 4 and 6 show the MF di-peptide of each respective RNA. Lane 2 shows the MFV tri-peptide. Lanes 3 and 5 show the difference in migration between 1 and 2 arginine residues after MFV. Lanes 6, 7 and 8 show the peptide migration patterns using the MFVVVR RNA template for peptide synthesis after the addition of each TC. Note the clear difference in migration patter between the penta-peptide which does not leave the origin (lane 7), and the hexa-peptide that migrates towards the anode (lane 8). TCs=ternary complexes Met=free methionine

Adding an additional arginine residue produced an even bigger change in migration between the MFV and MFVRR peptides, but there was also an obvious difference in the migration of the MFVR peptide compared to MFVRR (Figure 4.4 B lanes 3 and 5). A third arginine containing peptide, MFVVVR, was also tried out. The MFVVV peptide does not leave the origin during TLE (Figure 4.4 B lane 7), but addition of arginine to produce MFVVVR caused the peptide to migrate towards the anode (Figure 4.4 B lane 8), making the difference in migration with an arginine residue in the peptide distinct to the shorter MF and MFVVV peptides. These results showed that the

MFVRR and MFVVVR peptides may be suitable peptides for analysing the effect of transcription on translation as they fulfilled the necessary requirements.

4.3 Initial use of TR-CTT

The TR-CTT system is designed such that transcription is assembled first by creating an artificially assembled elongation complex (AAEC) using an RNA oligo, template DNA (tDNA) and non-template DNA (ntDNA) and RNAP. The tDNA and ntDNA are important for accurate positioning of the RNAP during TR-CTT. The tDNA contains a 9 nt sequence complementary to the RNA with a short (10 base) overhang sequence upstream of the hybrid site and the DNA sequence that is used by the RNAP as a template to synthesise the RNA situated downstream of the hybrid site (Figure 4.2 B). The ntDNA sequence is complementary to the template DNA sequence.

To assemble the AAEC, both in TR-CTT and transcription only, the RNA:DNA hybrid is formed by annealing the RNA to the tDNA through base pairing of the complementary hybrid site on each oligonucleotide. The RNAP binds to the RNA:DNA hybrid and positions the RNA 3' end in the RNAP active centre ready for transcription elongation. The RNAP is modified with a biotin tag attached to the β' subunit for immobilisation of the AAEC on streptavidin beads. Biotin tagged RNAP is used as the ribosomes are his-tagged and therefore his-tagged RNAP is not suitable. At this stage, the transcription elongation complex is fully active, but not stable enough to withstand high salt washing. ntDNA is required to form the fully stable transcription elongation complex that is resistant to high salt washing (1M) and heparin (100 µg/ml). Addition of NTPs after assembly of the elongation complex allows the RNAP to begin transcription elongation and move forwards along the DNA template. By adding only certain NTPs the RNAP can be 'walked' along the DNA template to specific locations as the RNAP can only synthesise RNA using the specific NTPs present in the reaction. The length of the tDNA that the RNAP is walked along determines by how much the initial RNA is extended and therefore the final length of the RNA and the distance from the RNA 3' end to the peptide coding region. Addition of all four NTPs allows the RNAP to transcribe to the end of the DNA template, known as a chase reaction. After assembly of the AAEC, immobilisation on the streptavidin beads and transcription elongation, the complexes are washed to remove all of the unbound RNA. This ensures that all the RNA molecules left in the

reaction are bound by RNAP in a stable elongation complex and there is no free RNA for the ribosomes to bind and translate.

After washing of the AAECs, translation can be initiated on the RNA bound by the immobilised RNAP through addition of ribosomes, initiation factors and fMet-tRNA^{fMet} to the reaction. As with walking of the RNAP, the ribosomes can also be 'walked' along the RNA template by addition of specific TCs and EF-G to produce the peptide. Once synthesised, the peptide is released by addition of KOH, the sample concentrated by drying, re-suspended and analysed by TLE.

The RNAP was positioned at specific locations on the DNA template during transcription by adding only certain NTPs to the transcription elongation reaction. As the tDNA sequence is comprised of only four different bases, there was a limit to the variety of locations the RNAP can be positioned at on one template and thus different lengths by which the RNA can be extended. It must also be taken into account that GTP was required during translation and so will be added after walking of the RNAP and washing of the AAEC. For these reasons, more than one DNA template sequence was required during TR-CTT to position the RNAP at a wide range of distances from the peptide-coding region (and therefore the ribosome) after extension of the RNA.

Initially we tested the peptide MFVRR in TR-CTT. However, use of this peptide soon proved to be problematic, as, in contrast to what was observed previously in translation alone, it was not possible to distinguish between the MFVR and MFVRR peptides after TLE (Figure 4.5 A lanes 2 and 4 compared to lanes 6 and 7). In TR-CTT, as well as all the components required for translation, there is the addition of RNAP, DNA and streptavidin beads. The translation elongation reaction was also concentrated after KOH treatment to increase the amount of peptide loaded onto the chromatography plate. These extra components, along with concentrating the sample, appeared to affect the migration of the MFVRR peptide and produced a smear on the TLE plate after TLE (Figure 4.5 A lanes 2, 4 and 6). For this reason, the MFVRR peptide was unsuitable for use in TR-CTT and instead the MFVVVR peptide was chosen as it migrated well during TLE even after TR-CTT (Figure 4.5 A lanes 3, 4 and 6).

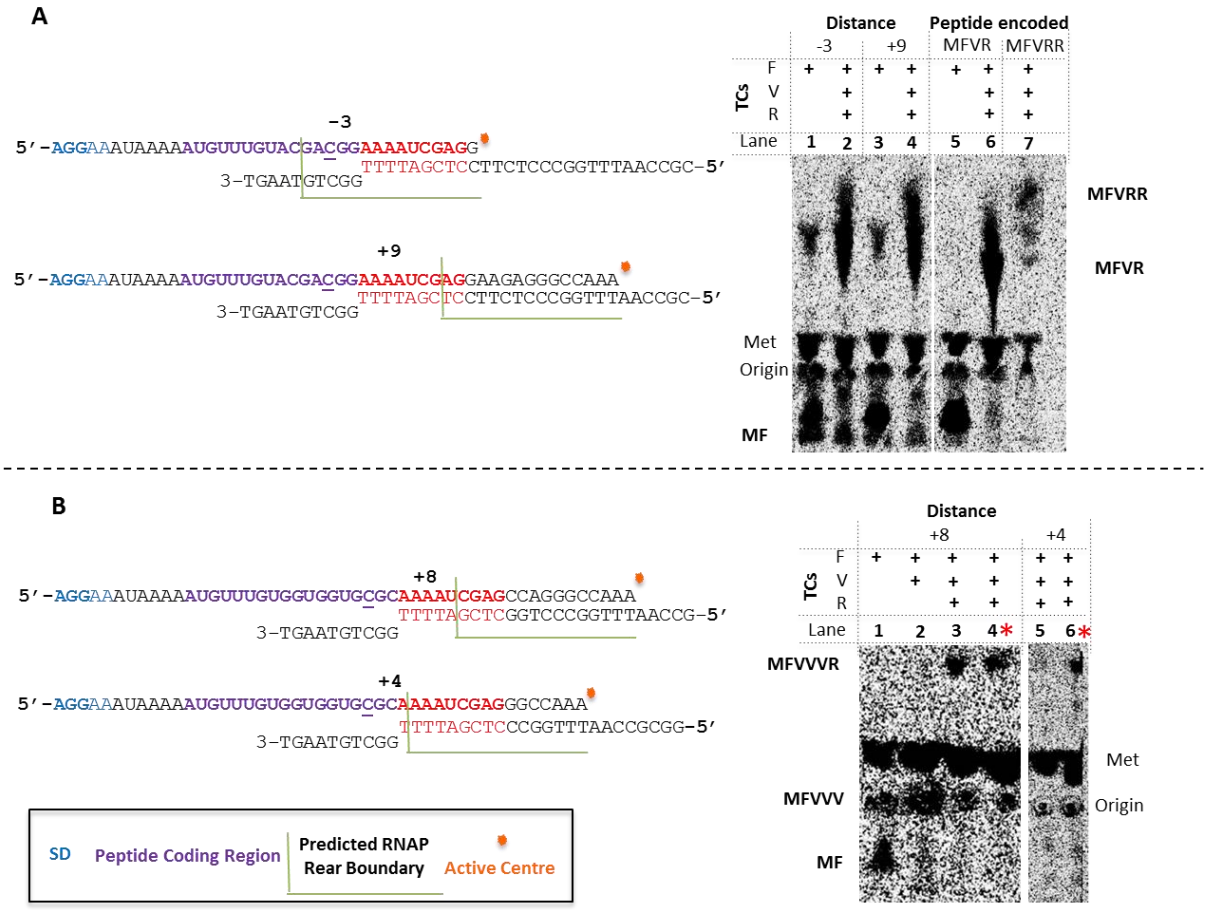


Figure 4.5 TR-CTT using RNA encoding peptides MFVRR and MFVVVR. The location of the RNAP proposed rear end on the DNA template is indicated by the green lines and the RNAP active centre is represented by the orange dot (Mg^{2+} ion). The distance from the RNAP rear end to the 1st nt of the arginine is indicated on each RNA:DNA schematic. A) TR-CTT using MFVRR peptide coding RNA. Lanes 1-4 show the peptide produced in TR-CTT and Lanes 5-7 show the uncoupled peptide controls. The signal in lane 7 below the labelled peptide is likely due to some degradation of the peptide sample. B) TR-CTT using MFVVVR peptide RNA. Lanes 1-4 indicate the peptide produced with the template that positions the RNAP rear end 8nts from the 1' nt of the arginine codon. Lanes 5 and 6 show the peptide after addition of F, V and R TCs when the RNAP is positioned 4nt away. The red asterisk (lanes 4 & 6) indicate the reactions in which the RNAP was chased to the end of the template by the addition of all NTPs.

TR-CTT was performed using the RNA containing the MFVVVR coding region designed in section 4.1 and the DNA templates shown in Figure 4.5. We determined that when there was a distance of 4 nts from the 1st nt of the codon in the ribosomal A-site to the predicted RNAP rear end, the ribosome could not incorporate the final arginine residue (Figure 4.5 B lane 5). When the RNAP was chased to the end of the template after the addition of all four NTPs, the arginine became incorporated into the peptide (Figure 4.5 B lane 6). This suggested that the inability of the ribosome to incorporate arginine was due to the positioning of the RNAP on the template blocking the ribosome from synthesising the hexa-peptide.

We then tested a second RNA template that allowed the RNAP to be translocated to position the predicted rear end of the RNAP 8 nts from the 1st nt of

the arginine codon. At this distance, the ribosome was able to incorporate the arginine residue into the peptide and synthesise the full-length peptide (Figure 4.5 B lane 3), suggesting that the 8 nt distance was permissive for incorporation of arginine, contrary to what was observed with a distance of 4 nt.

These results however, proved problematic to confirm. Often the presence or absence of the full-length peptide was difficult to clearly distinguish from the background due to low peptide signal after TR-CTT (data not shown).

During the transcription stage of TR-CTT, RNAP is halted at a specific site on the DNA template by adding only the NTPs required to translocate the RNAP to that exact location and the AAECs are washed with high salt to remove all loosely formed complexes. This should result in a homogenous population of RNAP molecules positioned at the correct location after synthesis of the designated length of RNA. To analyse transcription on the DNA templates, transcription was carried out on all of the DNA templates and the RNA radiolabeled through incorporation of a radiolabelled α-^{32}P-NTP by the RNAP during transcription (the exact NTP added was dependent on the identity of the first NMP to be incorporated). The AAECs were immobilised and washed with high salt buffer and transcription buffer (low salt). After washing, NTPs were added to translocate the RNAP to the desired location and the transcription reaction stopped with the addition of an equal volume of transcription stop buffer (containing formamide). The samples were denatured by boiling for 2 min and separated on a 10% denaturing polyacrylamide gel and visualised by phosphor-imaging. The results from this analysis revealed that on many of the DNA templates, the RNAP was heterogeneously located after the addition of NTPs. On some of the DNA templates, not all of the AAECs transcribed fully to the correct site in the presence of NTPs but instead paused prematurely, resulting in the production of a shorter length of RNA after transcription elongation (Figure 4.6, green asterisk). On other DNA templates, the RNAP used the NTPs present within the reaction to transcribe beyond the site on the DNA template at which the RNAP was expected to halt, by misincorporating the NTPs added into the reaction (Figure 4.6, red asterisk). Read-through by the RNAP will position the RNAP further away from the peptide coding sequence, increasing the distance between the ribosome and the RNA 3’ end, therefore allowing incorporation of arginine and production of a false positive result. Conversely, premature halting of the RNAP could lead to a false negative result, as the shorter RNAs would potentially block incorporation of the arginine.

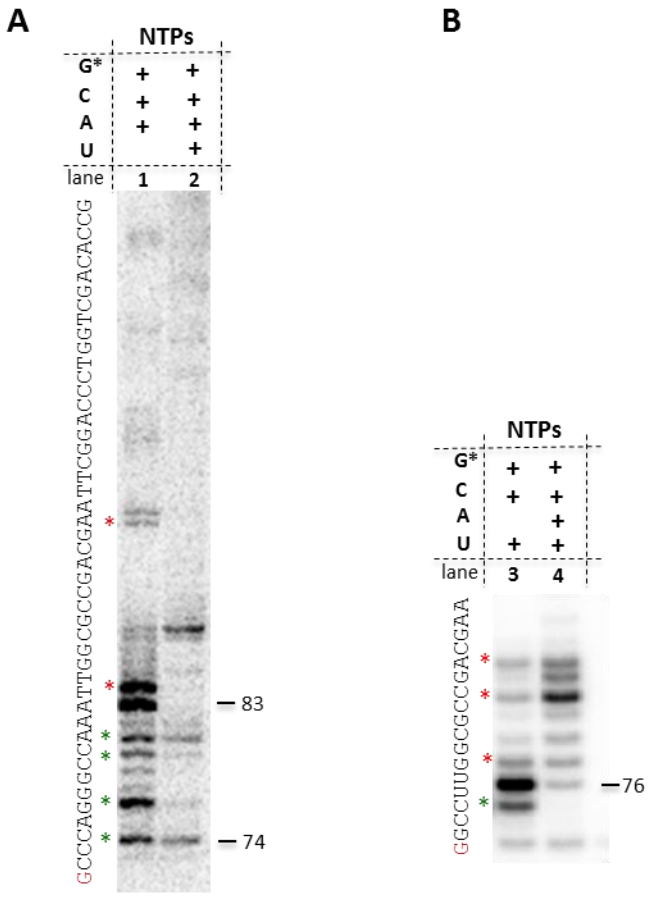


Figure 4.6 Pausing and read-through of the RNAP during transcription. Examples of pausing of the RNAP before the desired halt site are indicated by the green asterisk and read-through beyond the halt site is indicated by the red asterisk. The RNA sequences are shown to the left of the gel images with the radiolabelled α-P^{32}-GTP shown in red. A) Addition of the NTPs GTP, CTP and ATP in the transcription elongation reaction in lane 1 results in extension of the RNA to the desired length (83 nts) but there is also pausing of the RNAP after 74 nts and read-through by one base to produce a 84 nt RNA product. B) Transcription of a second template shows one RNA species after elongation with GTP. CTP and UTP, but RNA products of a variety of lengths are synthesised in the presence of all NTPs.

Another consequence to using a collection of different DNA templates to position the RNAP at a range of distances from the peptide-coding region during TR-CTT resulted in the RNAP transcribing a variety of different sequences. The sequence of the RNA:DNA hybrid in the active centre when the RNAP is halted also differs between the templates. A change to the sequence in the active centre of the RNAP can cause the RNAP to change its behaviour, for instance induce backtracking or pause in the pre-translocated state (Bochkareva *et al.* 2012). Using a variety of DNA sequences as templates for transcription was therefore not ideal as it could potentially change the way RNAP behaves during transcription elongation and even translation.

In order to eliminate the potential complications that arose from having a variety of DNA templates, the approach was changed slightly. Instead of using

different templates to increase or decrease the distance between the RNAP and the ribosome, only one DNA template was used to keep the transcription sequence and RNAP halt site constant, in order to minimise any difference in RNAP behaviour. The length of the RNA used to form the AAEC was instead changed by successively adding nucleotides between the peptide coding sequence and the RNA:DNA hybrid site of the original MFVVVR RNA (Figure 4.7 A and B). The template DNA that allowed the predicted RNAP rear end to be positioned 4 nts away from the 1st nt of the ribosome A-site after walking by addition of GTP, CTP and ATP to produce an RNA of 77 nt was chosen. This template was chosen on the basis that the RNAP transcribed to the halt site well with minimal pausing and was also the template used to position the RNAP rear end 4 nts away from the 1' nt of the ribosomal A-site in Figure 4.5 B.

Characterisation of the pause state of RNAP at this position by pyrophosphorolysis and GreB cleavage showed that the RNAP was stabilised in the post-translocated state and was not backtracked (Figure 4.7 D, top). In this new approach of adding up to 5 nucleotides to the RNA between the peptide coding sequence and the RNA:DNA hybrid site allowed the predicted RNAP rear end to be positioned up to 9 nucleotides away from the 1' nt of the ribosome A-site. This spanned the distances identified as being permissive and non-permissive (8 nts and 4 nts respectively) based on the preliminary results obtained using the different DNA templates and shown in Figure 4.5 B. The RNAs were named MFVVVR +1, +2, +3, +4 and +5 according to the number of nucleotides added before the RNA:DNA hybrid site (RNAs of lengths 71 to 76 nts before and 77 to 82 nts after transcription elongation).

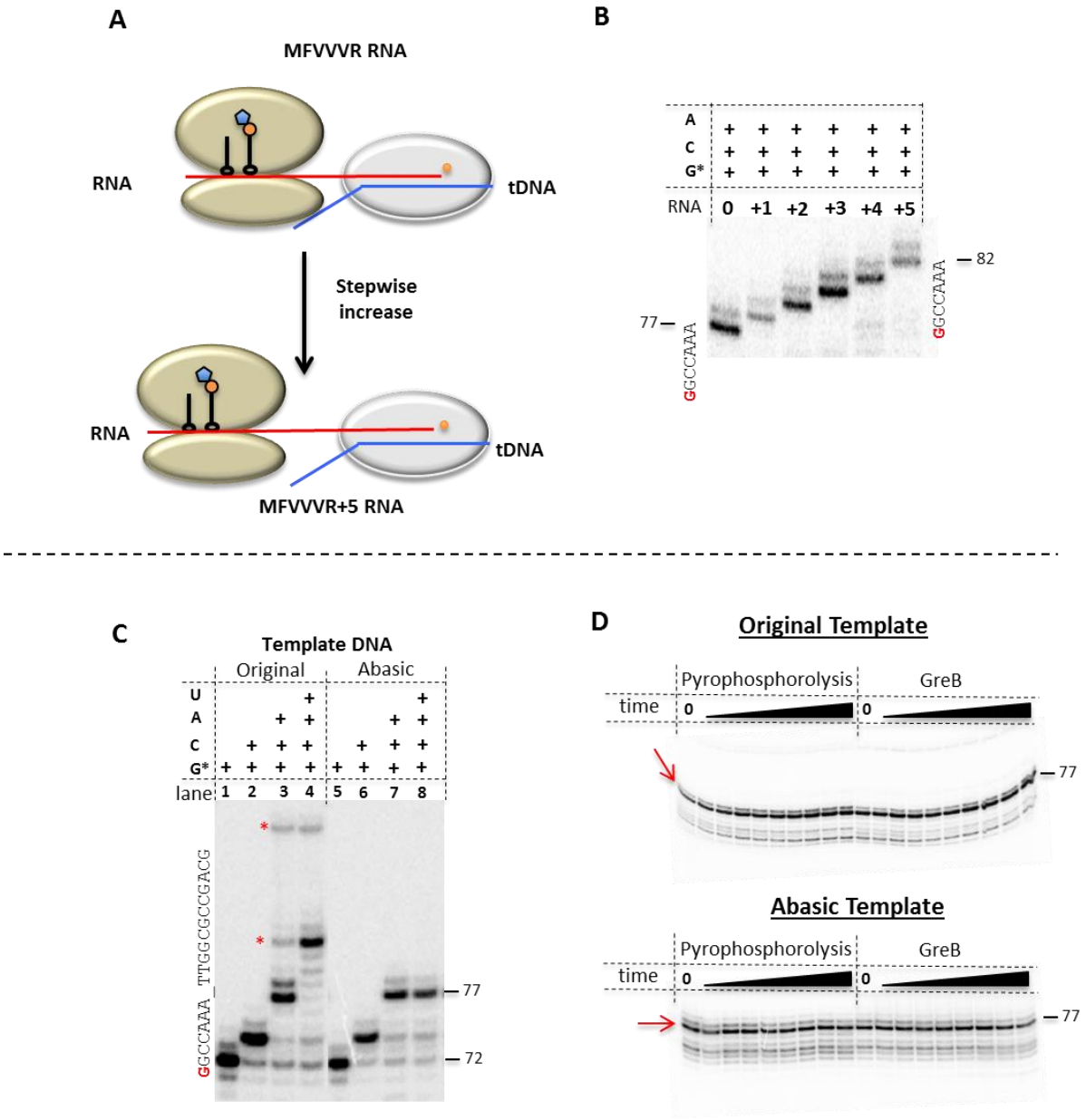


Figure 4.7 Analysis of RNA and tDNA. All RNA was radioactively labelled by incorporation of α-P^{32}-GTP during transcription A) Extended RNAs schematic. Although not shown, non-template DNA was included in the reaction. B) PAGE showing the increasing size of the RNA after addition of 100 µM GCA NTPs to walk the RNAP to the halt site. The template DNA sequence is shown on the left and the size of the RNA indicated on the right. C) Comparison of transcription using MFVVVR RNA and the DNA templates without (lanes 1-4) and with (lanes 5-8) the abasic site. Template sequence and RNA length indicated as in B. The red dots indicate RNA products produced by read-through of the RNAP (experiment performed by Daniel Castro-Roa) D) Pause characterisation of RNAP after transcription from the original DNA template (top) and the abasic template (bottom) after synthesis of the 77nt RNA product. Full-length product indicated by red on the left arrow and the size marker on the right. + indicates presence of the specific NTP in the reaction. A=ATP, G=GTP, C=CTP and U=UTP.

The DNA template used to halt the RNAP after synthesis of the 77 nt RNA product by adding only G, C and A nucleotides to the reaction still showed a small amount of read-through of the halt site (Figure 4.7 C lane 3, red dots). To counteract this, an abasic site was added in place of dATP on the DNA template after the RNAP

halt site. An abasic site consists of a 3' hydroxyl and 5' deoxyribosephosphate but no pyrimidine or purine base and should prevent the RNAP from incorporating a nucleotide opposite the abasic site and continuing transcription. When tested in transcription, the abasic site in the template eliminated the read-through by RNAP seen on the standard template, even in the presence of all four NTPs (Figure 4.7C lanes 7 and 8 compared to 3 and 4). Characterisation of the state of the halted RNAP on this new template revealed it to be in the post-translocated state and not backtracked (Figure 4.7 D bottom).

Once the MFVVVR RNA sequence was lengthened with the successive addition of NTPs before the RNA:DNA hybrid site and tested in transcription and translation, and a suitable DNA template identified and characterised, the TR-CTT system was ready for use again. Unfortunately, even under the new conditions, the appearance of false positives and false negatives persisted with the full-length peptide appearing inconsistently at all distances tested and the level of peptide signal remained low (data not shown).

As well as overextension of the RNA by the RNAP due to read-through, another possibility for the appearance of false positives could be due to release of RNA by the RNAP during translation. Free RNA would be translated along with RNA bound by RNAP, but as there was no RNAP at the 3’ end to impede translation, the ribosome was able to translate the full-length peptide. This could account for the possible hexa-peptide signal seen when the RNAP is obscuring the arginine codon (data not shown). TR-CTT was carried out as usual using MFVVVR RNA and the shortest template as standard with radiolabeling of the RNA at the 3’ end by incorporation of α-P^{32}-GTP during transcription. After the translation elongation stage, the streptavidin beads were pelleted and beads and supernatant separated (Figure 4.8 A). The volumes were equilibrated to make the samples comparable and the samples separated by denaturing PAGE using a 10% sequencing gel (Figure 4.8B). The absence of RNA in the supernatant showed that the RNA was not released by the RNAP during translation (Figure 4.8 B lanes 2 and 4).

The gel images produced from TR-CTT were often difficult to interpret due to the high level of background signal and the low level of signal from any of the peptides. To remove background noise, two techniques were tried in translation alone with free RNA. The purified ribosomes used during translation contained a

histidine tag and so could be immobilised on Ni-NTA beads during translation alone (in the absence of biotin tagged RNAP). After translation elongation but before

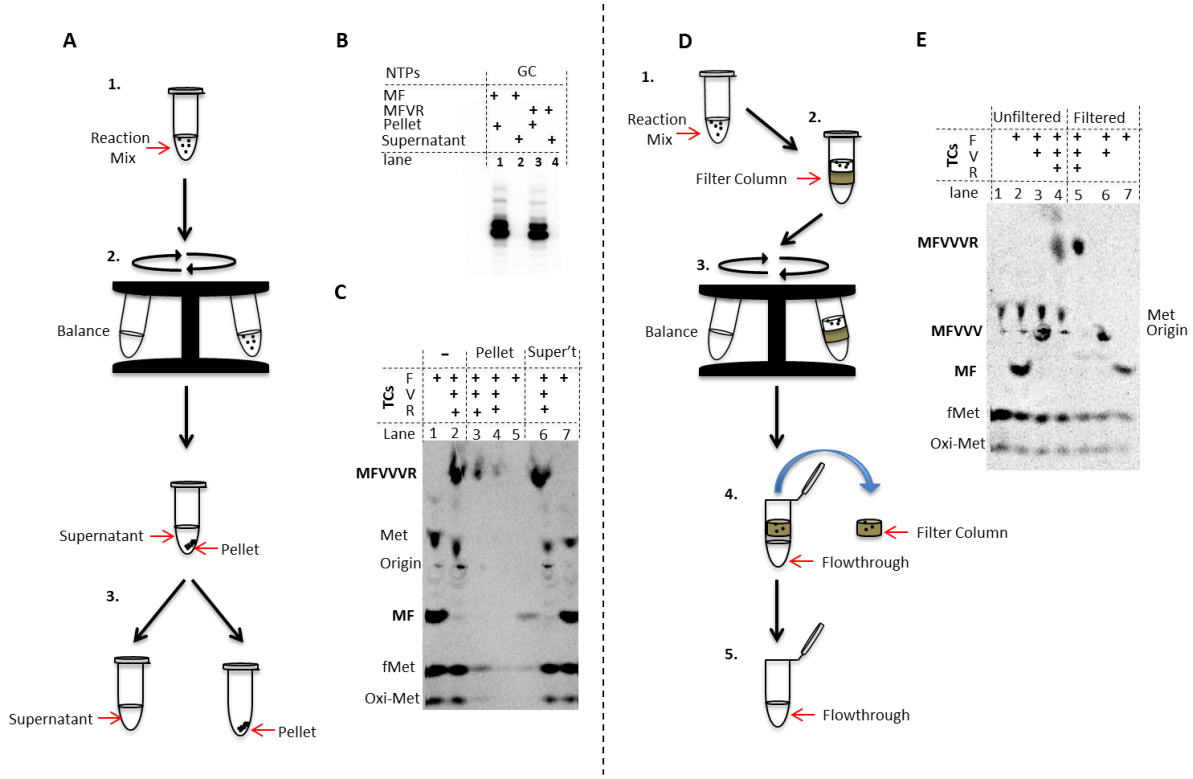


Figure 4.8 Analysis of pellet vs supernatant for RNA and Peptide and filtering of the peptide. A) Schematic representation of the procedure used to separate the pellet and supernatant. The reaction is assembled (step 1), centrifuged briefly in a table top centrifuge (step 2) and the supernatant removed (step 3). An equal volume of the relevant buffer is added to the pellet to equilibrate the volumes. The reactions are stopped with the addition of an equal volume of stop buffer. B) PAGE of the RNA after TR-CTT. The RNA is labelled by incorporation of radiolabelled α-P^{32}-GTP during transcription. C) TLE of the MF and MFVVVR peptide after translation alone. Lanes 1 and 2 contain the control reaction and lanes 3-5 contain the pellet, with lanes 4 and 5 showing the pellet sample after washing with translation buffer. Lanes 6 and 7 contain the supernatant samples. D) Schematic of filtering of the peptide after translation alone using free RNA. The reaction is assembled, translation elongated, and then 10 mM EDTA added (step 1). The reaction is pipetted onto a BioRad gel filtration column equilibrated in translation buffer and filtered by centrifuging at 3,000g for 4 min (step 2). The column is removed and discarded (step 3) and KOH added to the flow through to release the peptide. E) MF, MFVVV and MFVVR peptides before (lanes 2-4) and after (lanes 5-7) filtering.

release of the peptide, Ni-NTA beads were added to the reaction and the pellet and supernatant separated (Figure 4.8 C lanes 3-7). The pellet was also washed with translation buffer (Figure 4.8 C lanes 4-5). The supernatant contained more peptide than the pellet and there was very little peptide remaining after washing. This procedure was therefore not suitable for removing the background noise. The second method tested was to filter the translation products through BioRad chromatography spin columns (Figure 4.8 D). Filtering the products from translation alone and using the flow-through for analysis by TLE produced a much clearer image (Figure 4.8 E compare lanes 1-4 and 5-7). The free methionine normally present near the origin was filtered out, along with some of the formyl-methionine and oxidised methionine.

Applying the technique to the products of TR-CTT before concentrating the samples did clean up the image slightly, but significant background noise was still present and the peptide signal remained low.

4.4 Conclusion and Discussion

In this chapter, the TR-CTT system was modified with the aim of determining the minimum distance required between the RNAP and the ribosome. We showed that the peptide MFVVVR fulfilled the necessary requirements, most notably due to the difference in migration during TLE of the partial peptides and the six amino acid peptide. Initially, a variety of DNA templates were used to position the RNAP at different distances from the 1st nt of the ribosomal A-site, but this technique resulted in low peptide signal and inconsistencies due to read-through or early pausing of the RNAP on the different templates. The method was modified to keep the DNA template sequence constant and instead increasing the length of the RNA used to form the AAEC. The new approach reduced the inconsistencies in transcription elongation observed between different DNA templates. The template DNA sequence was chosen based on the positioning of the RNAP and the stable state of the halted RNAP and modified to introduce an abasic site to prevent any read through during transcription. Preliminary results obtained using the different DNA templates suggested that the distance of 4 nts was too close for the ribosome to incorporate arginine, but when this was increased to 8 nts, the ribosome was able to synthesise the full-length peptide. This result however, proved difficult to confirm. Visualisation of the peptide during TR-CTT proved to be more problematic than expected due to the amount of background signal on the TLE plate. Filtering of the reaction after translation elongation reduced the background slightly, but it was still difficult to clearly detect the different peptides, hindering the interpretation of our results. Although the RNAs and DNA templates fulfilled their expected roles in transcription and translation only, it became apparent that further optimisation was necessary in order to obtain reliable results from our coupled system.

5. Troubleshooting the TR-CTT method

5.1 Introduction

In the previous chapter, the TR-CTT method was adapted for the purpose of studying the effect of the RNAP on the ribosome in an *in vitro* system and was used to obtain preliminary results. These results, however, proved to be difficult to confirm and the images were not clear with respect to distinguishing the peptide signal from the background, even after filtering of the products. Therefore, a further, more in depth analysis of both transcription and translation was required before the TR-CTT method was suitable.

5.1.2 Aims and Objectives

The aim of this chapter was to re-analyse in more depth the TR-CTT set up to determine the cause of the low peptide signal. In particular, the synthesised RNA, the status of the RNAP during transcription and translation and the efficiency of the ribosomes needed to be examined more thoroughly.

5.2 Analysis of the RNA

During the initial screening and testing of the RNA and DNA templates in transcription and TR-CTT, the nascent RNA was radiolabelled at the 3' end by RNAP incorporating an α-[^{32}P]-NTP during transcription. This is a generally accepted method of visualising the RNA in the system, but, however, radiolabelled NTPs will only be incorporated into the RNA that is being actively transcribed. Therefore, any RNA species in the reaction that are not bound by the RNAP and/or are not being extended during transcription will not be labeled using this approach. In order to visualise all of the RNA in the reaction, we decided to use an alternative method, known as kination that is performed prior to transcription. The tri-phosphates of the 5' terminal nucleotide of the RNA oligo are first removed by dephosphorylating the RNA with calf intestinal alkaline phosphatase (CIAP). Then, T4 polynucleotide kinase (PNK) is added to transfer the radiolabelled γ-phosphate group from γ-[^{32}P]-ATP onto the 5' nucleotide of the RNA oligo.

Using this method, we radiolabeled the original MFVVVR RNA at the 5' end. The AAEC was then assembled in the absence of ntDNA with 5' end labeled RNA, RNAP (in 1.5x excess over the RNA) and tDNA (2x excess over the RNA) and the complexes were immobilised on streptavidin beads (Figure 5.1 A, step 1). After AAEC assembly, two aliquots were taken. To one aliquot, an equal volume of stop

buffer was added, whilst to the other aliquot, NTPs were added at a final concentration of 100 µM and transcription was allowed to proceed before the addition of an equal volume of stop buffer (Figure 5.1 A, step 2). To the remainder of the reaction, ntDNA was added in 10x excess over template DNA (Figure 5.1 A, step 3), incubated then aliquots was taken as above (Figure 5.1 A, step 4). The main reaction was centrifuged briefly to separate the pellets and supernatant. Aliquots of the supernatant were taken as above (Figure 5.1 A, step 5 & 6). The pellet was washed with high salt and transcription buffer or just transcription buffer, the supernatant removed and discarded and the pellet resuspended in transcription buffer. To one aliquot, NTPs were added as previously and to the other an equal volume of stop buffer was added. The samples were denatured by boiling for 2 min, before analysis by 10% denaturing PAGE.

Analysis of 5' end labeled MFVVVR RNA in transcription revealed that the RNA being used to form the AAEC contained a mixed population of RNAs, even after purification (Figure 5.1 B). This could not be observed previously as the RNA was radiolabelled by incorporation of α-[^{32}P]-NTP during transcription so only RNA that was extended during transcription elongation was visible. Two species of RNA were most prevalent, the RNA species of the correct length (70 nt) and an RNA species one nucleotide longer (71 nt) (Figure 5.1 B, lane 1). The RNA sizes were determined by comparison to the RNA radiolabeled by α-[^{32}P]-NTP incorporation during transcription. After the addition of NTPs, both the 70 and 71 nt RNA species were elongated to produce one species 77 nt in length (Figure 5.1 B lane 2). This indicates that the difference in size of the RNAs produced during RNA synthesis was due to the addition of NTPs at the 3' end. In fact, T7 RNAP is known to add 1 or more random NTPs to the 3' end of the RNA during synthesis in a seemingly sequence independent manner (Triana-Alonso *et al.* 1995; Cazenave & Uhlenbeck 1994; Arnaud-Barbe *et al.* 1998; Nacheva & Berzal-Herranz 2003; Lapham & Crothers 1996).

5' end labeling of the RNA also revealed the presence of an RNA species 68 nts in length (Figure 5.1B lane 1). This was later proven to be produced by intrinsic hydrolysis of the 70 or 71 nt long RNA by the RNAP during AAEC assembly since these products were only observed after incubation with RNAP (data not shown). Surprisingly, this RNA species appeared to be elongated more successfully than the 70 nt RNA and the 71nt RNA (Figure 5.1B lane 2). Although the 70 nt RNA is of the

correct length, it is not extended as efficiently as the 68 nt RNA produced by intrinsic hydrolysis. This could be due to the T7 RNAP incorporating the incorrect NMP at the 3’ end and causing backtracking of the RNAP.

All three RNA species (68, 70 and 71 nts) were present in the supernatant after immobilisation of the AAEC (Figure 5.1 B lane 5). Addition of the NTPs to the supernatant resulted in the extension of the 68 nt RNA to 77 nt, indicating that either some of the RNAP had been released into the supernatant or some of the streptavidin beads had been released from the pellet (Figure 5.1 B lane 6). High salt washing of the AAEC complexes removed the 71 nt RNA and addition of NTPs after high salt washing resulted in the majority of the RNA being elongated to produce the 77 nt product. However, the proportion of full-length RNA remaining after high salt washing (lane 10) compared to the total amount of RNA originally entered into the reaction was greatly reduced. More RNA was retained after only washing with transcription buffer (lane 7) but a lot of this RNA was not extended in the presence of NTPs (lane 8). These results suggested that the RNA that was not extended during transcription did not form as stable a complex as the RNA that was extended.

Preliminary experiments in chapter 4 suggested that no RNA that was part of the AAEC was released during translation (Figure 4.8 B). However, the method by which the RNA was labeled could not show if there was any RNA in the reaction that was bound by the RNAP and carried over into translation but not extended, because this RNA would not have been radioactively labeled. To visualise the total RNA during TR-CTT, TR-CTT was re-analysed using 5’ end labeled RNA and samples were taken at each step to analyse the total RNA throughout the reaction. The AAECs were assembled with 5’ end labeled RNA, RNAP and template DNA only (Figure 5.1 C lane 1) and then NTPs were added (Figure 5.1 C, lane 2). Again, the 68 nt RNA species was extended more efficiently than the two longer RNA species. The reaction was then split into two and ntDNA added to one reaction (Figure 5.1 C lane 5). A sample of supernatant was taken before washing of the AAECs (with tDNA only, lane 3 and with ntDNA, lane 6). The AAECs without ntDNA were washed with transcription buffer only (Figure 5.1 C lane 4) as AAECs assembled with only tDNA are not stable enough to withstand high salt washing, whilst the AAECs formed with both tDNA and ntDNA were washed with high salt then transcription buffer (Figure 5.1 C lane 7).

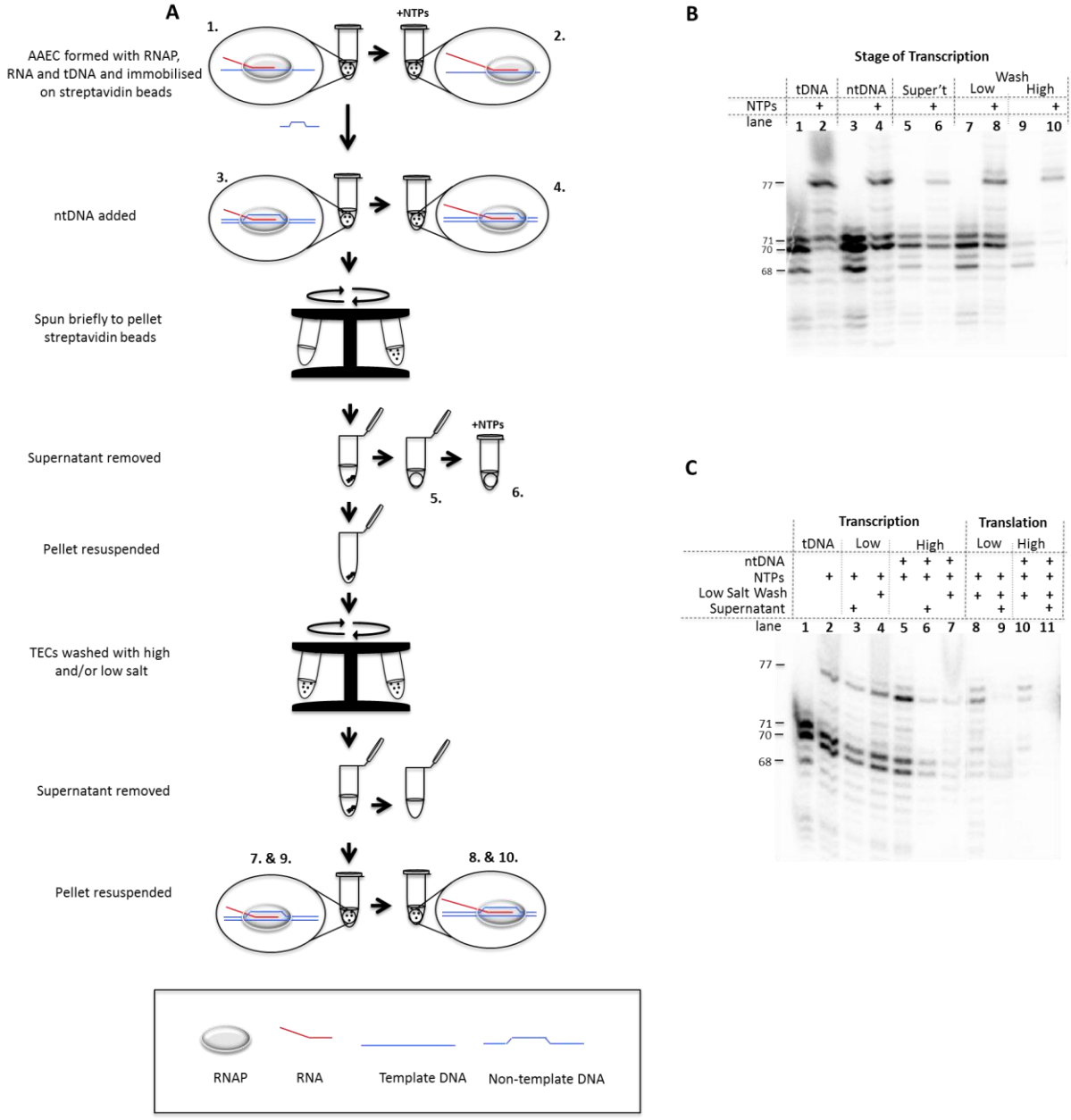


Figure 5.1 Analysis of the RNA during transcription and translation. A) Diagram showing the transcription reaction and the stages at which two aliquots were taken. In one aliquot the reaction was stopped whilst to the other NTPs were added then the reaction stopped. Numbers correspond to the lanes in B. **B)** 5' end labelled RNA after use in transcription. The stage after which the sample was taken is shown across the top. Odd numbered lanes show the RNA before addition of NTPs and even numbered lanes show the RNA after addition of 100 µM final ATP, CTP and GTP **C)** Gel image showing the RNA in the samples taken during TR-CTT. The samples taken after the addition of tDNA and low and high salt washes are indicated across the top. All RNA is radiolabelled at the 5' end and the sizes are indicated to the left of each gel image.

After washing of the AAECs, translation was initiated on the RNA bound by the RNAP. Translation was carried out as described in chapter 4 and the MF dipeptide synthesised. After translation elongation, the pellet and supernatant were separated and samples taken (Figure 5.1 C lanes 8 & 9 for TR-CTT with tDNA only and lanes 10 & 11 for TR-CTT with tDNA and ntDNA). A small proportion of the un-extended RNAs were observed in the supernatant after TR-CTT with tDNA and

washing of the AAEC with transcription buffer (Figure 5.1 C, lane 9) but the majority of the RNA remained in the pellet (lane 8). Only a very small amount of RNA was detected in the supernatant after TR-CTT with tDNA, ntDNA and high salt washing of the AAEC (Figure 5.1B, lane 11) and was visible on the gel only when the contrast was greatly increased (data not shown). Since only the AAECs formed with both tDNA and ntDNA and washed with high salt were used in translation, release of the RNA by the AAECs during translation after only washing with transcription buffer was not a cause for concern. It remained to be determined whether the RNA released from the AAECs during translation of the AAECs washed with high salt would be of a high enough concentration for the ribosomes to synthesise enough peptide to be seen after TLE.

5.2.1 RNaseH Site-directed RNA Cleavage

As previously described, the RNA used for transcription and translation contained an additional NMP at the 3' end as a result of the T7 RNAP adding NMPs during RNA synthesis. As a result of this, the majority of the purified RNA was not extended by the RNAP during transcription. This phenomenon is well documented (Triana-Alonso *et al.* 1995; Cazenave & Uhlenbeck 1994; Arnaud-Barbe *et al.* 1998; Nacheva & Berzal-Herranz 2003; Lapham & Crothers 1996) and methods to remove the additional NTPs have been developed. For instance, RNaseH site directed cleavage of the RNA is used to modify the RNA after synthesis and purification (Lapham & Crothers 1996; Yu 2012; Lapham *et al.* 1997; Inoue *et al.* 1987). The RNaseH enzyme only digests RNA hybrised to DNA (Clarke *et al.* 1973). Short oligos comprised of a sequence of 14 modified 2' o-methyl RNA bases at the 3' end followed by four dNTPs at the 5' end can be used to direct RNaseH to cut a complementary RNA strand at a specific site directly 5' to the dNTPs (Figure 5.2 A). A modified oligo was designed to direct RNaseH to cut the MFVVVR RNA two nucleotides from the 3' end and produce a homogenous population of RNA 68 nts in length (Figure 5.2 B, lanes 2-6). The RNA was re-purified from a 10% denaturing polyacrylamide gel as for the standard purification of synthesised RNA. The pure 68 nt RNA was labeled at the 5' end and tested in transcription. NTPs were added to the reaction either before or after ntDNA and samples were taken at all stages during the transcription reaction (Figure 5.2 C, left pathway is ntDNA followed by NTPs and right pathway is NTPs followed by ntDNA). After transcription elongation, the supernatant was removed and the complexes washed with either high salt and transcription buffer

or with transcription buffer only. Although more of the complexes elongated in the presence of ntDNA, ntDNA also increased the read-through by RNAP beyond the abasic site (Figure 5.2 D, compare lanes 2 (without ntDNA) and 3 (with ntDNA)). Regardless of the order of AAEC assembly and transcription elongation, not all of the RNA was extended and some unextended RNA remained after the wash with transcription buffer only (Figure 5.2 D lane 6 and 11) but was removed during the high salt wash (Figure 5.2 D lanes 5 & 10).

For the 68 nt RNA to be have been able to withstand even washing with transcription buffer it must have been part of a complex. Even after gel purification, there may have been some of the modified oligo still bound to the RNA that could have resembled an RNA:DNA hybrid. If this was the case, the RNAP may have bound to the RNA:oligo complex and formed an AAEC that was stable enough to withstand washing with transcription buffer but not high salt. As there was no DNA template for the RNAP to use to elongate the RNA, any RNA species that were part of such complexes would not have been extended. The modified oligo was very difficult to remove from the RNA after RNaseH treatment and required heating to 100°C for 10 min in the presence of a saturating amount of UREA immediately before loading onto a denaturing 10% polyacrylamide gel during purification to separate the oligo from the RNA (data not shown).

The addition of non-template DNA is not essential for the formation of an active AAEC, but these AAECs are not stable enough to withstand a high salt wash. When the AAECs were formed with the RNaseH digested RNA but without ntDNA, there was no read-through of the abasic site, but the complexes with the 68 nt RNA were not removed during the transcription buffer wash, therefore AAECs without ntDNA were not suitable for use in TR-CTT, as there would be a mixed population of RNA bound by the RNAP during translation. The AAECs assembled with both tDNA and ntDNA were able to withstand the high salt wash that removes the 68 nt RNA but there was read-through of the abasic site in the presence of the ntDNA. For this reason AAECs assembled with both tDNA and ntDNA were also not suitable for use in TR-CTT, again due to the mixed population of RNA species that would be used in translation. In light of the above, site directed RNaseH digestion of the RNA was not a suitable method for removing the extra NTPs added during RNA synthesis.

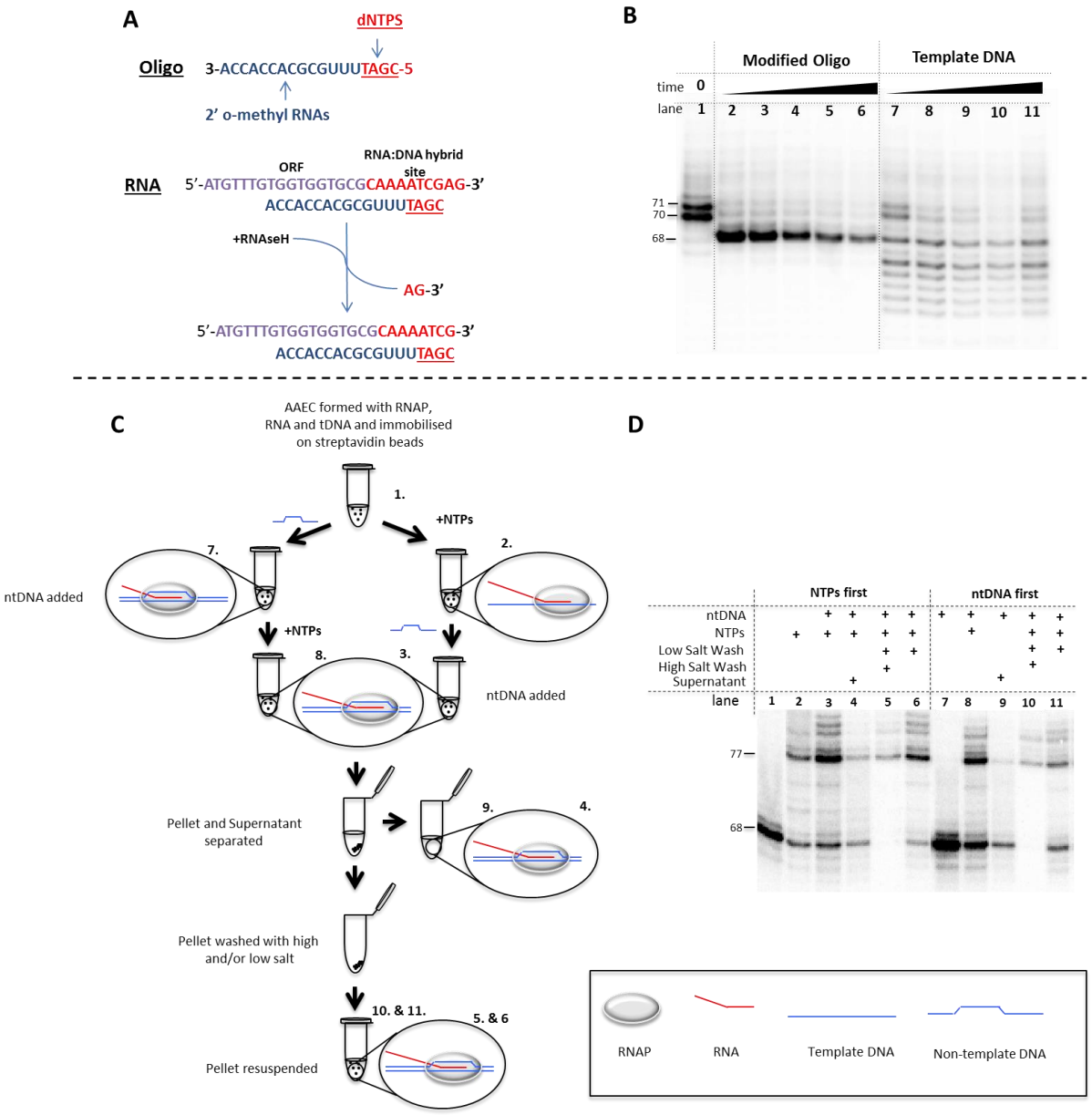


Figure 5.2 RNaseH cleavage of RNA and analysis in transcription. A) Schematic of RNaseH directed cleavage of the RNA. B) PAGE of RNAseH digestion of the MFVVVR original RNA. Lane 1 shows the RNA before RNaseH digestion. Lanes 7-11 contains the control samples of RNAseH digestion of the RNA in the presence of tDNA. Samples were taken 10 min, 30 min, 1 h, 2 h or 3 h after addition of RNAse H. C) Schematic to show the stages during transcription with the RNaseH digested RNA at which the samples were taken. Number corresponds to lane numbers in D. D) Lanes 1-6 show transcription with NTPs added before ntDNA. Lanes 7-11 show transcription with ntDNA added before NTPs. All RNA was radiolabelled at the 5' end and the sizes are indicated to the left of each gel image.

5.2.2 GreA

Analysis of the 5' labeled RNA in transcription revealed that approximately half of the RNA synthesised by T7 RNAP and used in transcription, translation and TR-CTT was a nucleotide longer than expected (Figure 5.1 B). As this extra NTP appeared to be added randomly, the chances are it was not complementary to the DNA template. During AAEC assembly the RNAP appeared to remove up to 3 nts

from the 3’ end of the RNA and the resulting shorter RNA species were extended efficiently by the RNAP during transcription (Figure 5.1 C and D). Based on these observations, it was therefore possible that the AAEC resembled a backtracked elongation complex directly after assembly. As the 68 nt product was fully elongated, it would seem that cleavage of the 3’ 2 or 3 nts of the RNA in the active centre by the intrinsic hydrolysis property of the RNAP created a new 3’ RNA end in the active centre to which the RNAP added nucleotides and thus began transcription. The elongation factor, GreA, resolves RNAP backtracked by up to 3 nts along the DNA template (Borukhov *et al.* 1993) and therefore the addition of GreA during AAEC assembly could result in better extension of the RNA during transcription elongation. To determine if the addition of GreA during AAEC does have an effect, the AAEC was formed as standard using 5’ end labeled RNA but in the absence of ntDNA (Figure 5.3 A, left branch of schematic, step 1). After AAEC assembly, either GreA was added, incubated for 10 min at 37°C (Figure 5.3 A, left branch, step 2) then 100 µM final concentration of NTPs added (step 3), or NTPs were added directly after AAEC (Figure 5.3 A, left branch, step 4). The same reaction setup was also used with ntDNA added to the AAEC before GreA (step 5) and samples taken after GreA and NTPs (steps 6 & 8) or just NTPs (step 7). After assembly of the AAEC with tDNA, RNA and RNAP, GreA induced cleavage of the RNA to produce a population of predominantly 68 and 69 nt long RNA species (Figure 5.3 A lane 2). This RNA is then extended by the RNAP to 77 nt upon addition of NTPs (lane 4). The effect of GreA was not as strong with the presence of ntDNA in the scaffold (Figure 5.3 A, compare lanes 2 with 6 and 4 with 8). A second transcription reaction revealed that the order in which NTPs and ntDNA were added after GreA cleavage did not affect transcription elongation (Figure 5.3 B). GreA was added after AAEC assembly (Figure 5.3 B step/lane 3), then either only NTPs added (step/lane 4) or ntDNA then NTPs (steps/lanes 5 & 6). A control reaction with non-template DNA but without GreA was carried out alongside (step/lanes 1 & 2). The RNAP extended the RNA as efficiently either with or without ntDNA when the ntDNA was added after incubation with GreA. GreA was also equally as effective at increasing the extension of the RNA when the AAEC was assembled using the longer RNAs that contained the additional nucleotides before the hybrid sequence (data not shown).

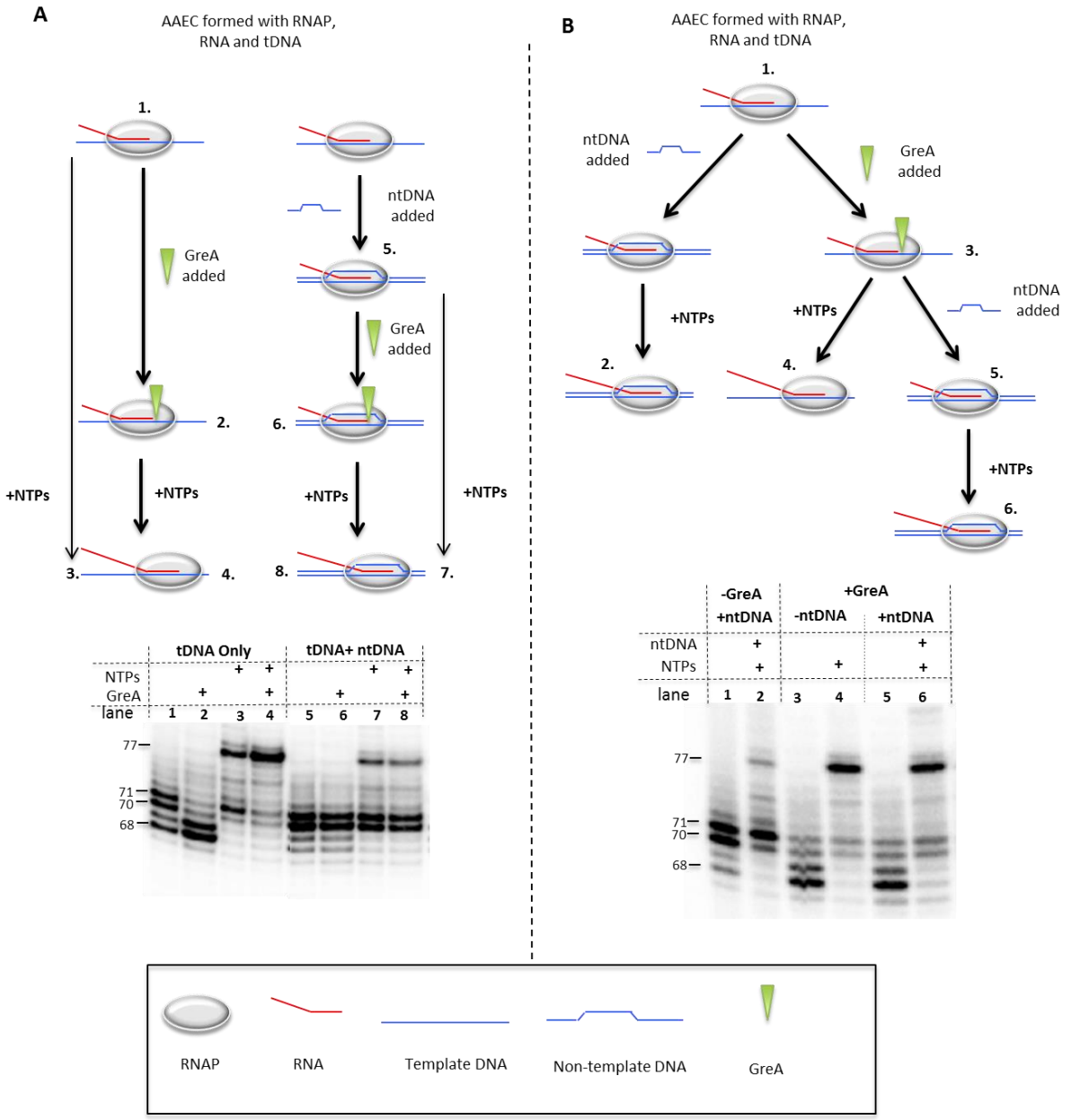


Figure 5.3 Use of GreA during transcription. All RNA is radiolabelled at the 5' end and the sizes are indicated to the left of each gel image. A) The effect of GreA on the AAEC and transcription elongation in the presence (lanes 2 and 4) and absence (lanes 6 and 8) of ntDNA. Numbers in the schematic correspond to lane numbers in the gel image. Lanes 1 and 3 and 5 and 7 show transcription controls without GreA and with and without non-template DNA respectively. B) Numbers in the schematic correspond to the lanes in the gel image. PAGE image of transcription with GreA (lane 3-6) and without (lanes 1and 2). Lanes 3 and 4 show transcription in the absence and 5-6 presence of ntDNA.

Out of the two methods tested to overcome the effect of the additional NTPs added to the RNA 3' end by T7 RNAP during synthesis, GreA proved to be more reliable than using site directed RNA cleavage. Also, GreA was effective on all the RNAs (data not shown) and did not require additional treatment of the RNA after purification, therefore the original AAEC assembly method was modified to include the addition of GreA after initial complex assembly (RNAP, RNA and tDNA), but before stabilisation of the AAEC with ntDNA (Figure 5.3 B, steps 1, 3, 5 & 6).

5.3 The high GTP concentration used during translation impacts upon the AAEC

The use of GreA during AAEC assembly resulted in a vast increase in the amount of the RNA that was elongated during transcription and retained after the high salt washing for use in translation. TR-CTT was therefore repeated with unlabeled RNA and using the modified method containing GreA to determine if this would increase the peptide signal. All six RNAs (MFVVVR original and +1 to +5, Figure 4.7 A and B) were used and the amount of RNA increased to 400 pmols. The RNAP and DNA concentrations were also increased to maintain the same ratios to ensure that a sufficient amount of RNA was retained after high salt washing. However, even after these modifications, the lack of peptide signal persisted (data not shown).

As the addition of GreA during AAEC assembly increased the amount of RNA retained after washing and subsequently used in translation, it was possible that an increased amount of RNA was being released by the RNAP during translation compared to TR-CTT in the absence of GreA (Figure 5.1 C). To determine if this was the case the experiment to separate pellet and supernatant after translation was repeated using the updated TR-CTT method with 5' end labeled MFVVVR RNA. Transcription was assembled as standard and samples taken after transcription elongation (Figure 5.4 A, step 1). Translation was initiated on the RNA extended during transcription elongation (step 2) and further samples taken after translation elongation to produce either the MF (step 3, left ribosome) or MFVVVR (step 3, right ribosome) peptides. A 'mock translation' reaction (as translation but without ribosomes, the translation proteins and aa-tRNA^{aa}) was carried out alongside, as well as a reaction with only translation buffer. After each reaction, the sample was spun briefly (step 4) to separate the pellet and the supernatant (steps 5 & 6). The pelleted beads were equilibrated to the same volume as the supernatant by adding translation buffer and an equal volume of stop buffer was added to both the resuspended pellet and supernatant (Figure 5.4 B).

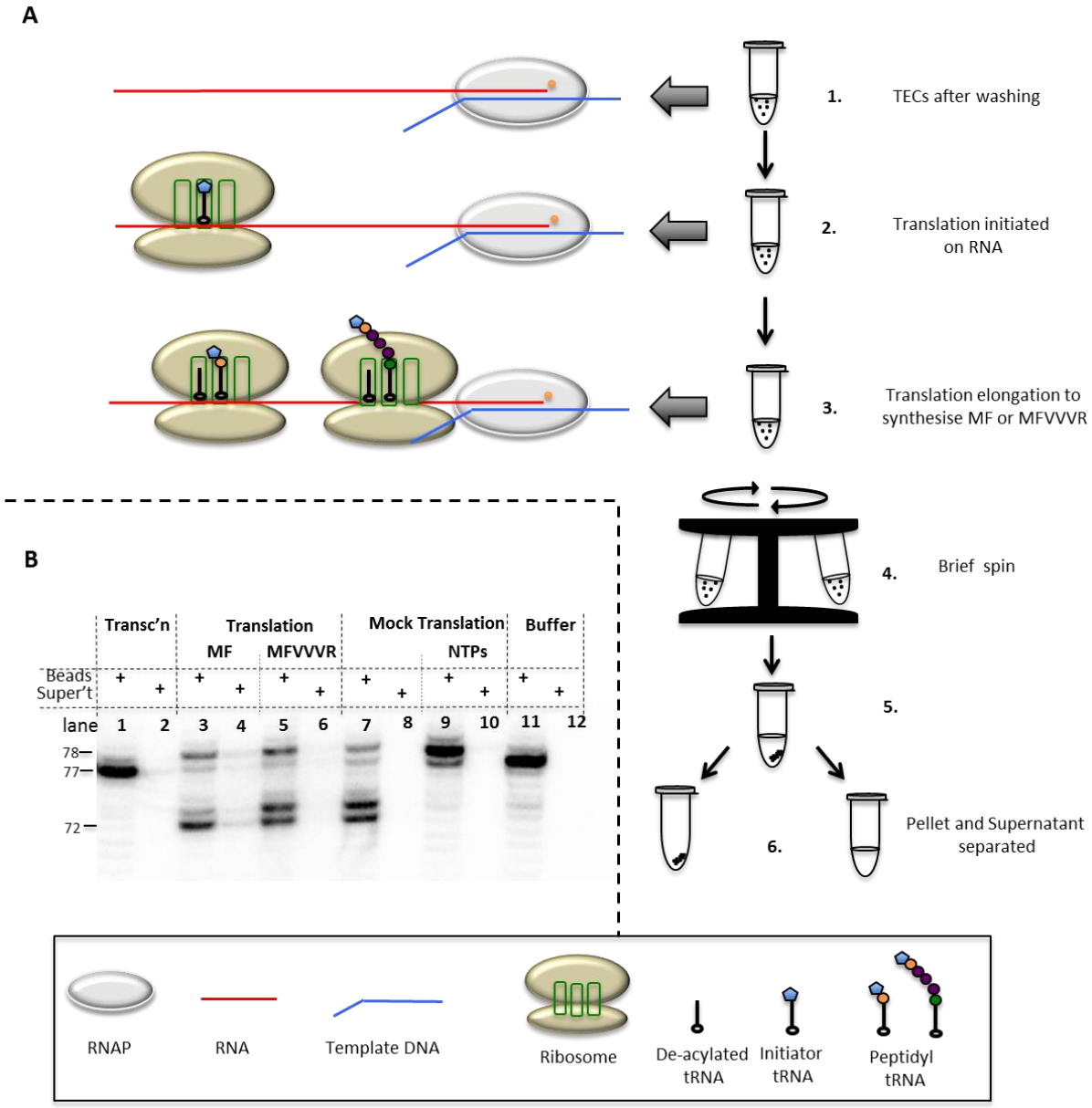


Figure 5.4 The effect of high GTP concentration on RNAP during TR-CTT. A) Outline of the TR-CTT reaction. Although not shown in the diagram, non-template DNA was present in the reaction. 1. AAECs after washing with high salt and transcription buffer. 2. Initiation of translation. 3. Translation elongation to synthesise either MF or MFVVVR peptide. 4. Centrifugation to pellet beads. 5/6. Separation of pellet and supernatant. **B)** PAGE of samples taken during TR-CTT using 5' end labeled RNA. Odd numbered lanes contain the immobilised samples and even numbered lanes contain the supernatant samples. Lanes 1 and 2 contain the samples taken immediately after washing the transcription reaction with high salt and transcription buffer. Lanes 3-6 contain the samples taken after translation of either the MF di-peptide (3 and 4) or the hexa-peptide MFVVVR (lanes 5 and 6). In mock translation (lanes 7-10), 4.3 mM GTP and translation buffer were added to samples and the reactions incubated as for translation. In lanes 9 and 10, 100 µM final of all NTPs was added during 'translation elongation'. Lanes 11 and 12 contain the reaction in which only translation buffer was added but with the same incubation times. The RNA size is indicated by the size markers on the left.

This analysis revealed a very surprising result. After translation and mock translation, the amount of the full length (77 nt) RNA was diminished and two or three new species of RNA of lengths 78 nts (1 nucleotide longer), 73 nts or 72 nts (4 or 5 nucleotides shorter respectively) emerged (Figure 5.4 B lanes 3, 5 & 7). In the

sample with buffer only, the RNA species did not change from those observed in the sample taken before translation (Figure 5.4 B lane 11 compare to lane 1). To one sample from the mock translation reaction, the NTPs CTP and ATP were added at a final concentration of 100 μM (the reaction already contained 4.3 mM GTP from translation). In this sample, the most prominent RNA species seen was the species 78 nt in length (Figure 5.4 B lane 9). Based on these observations, it would appear that, in the presence of the high concentration of GTP used during translation, the RNAP may have been misincorporating the GTP into the RNA opposite the abasic site in the tDNA (suggested by the appearance of the 78 nt long RNA product in the samples containing GTP). GTP was the only NTP in all of the samples apart from the sample with CTP and ATP added after mock translation, therefore it must have been GTP that was being misincorporated.

In addition to the misincorporation event, the RNA was also being cleaved to produce the two shorter, 72 and 73 nt, products in both the translation and mock translation reactions. Since these shorter RNAs occurred irrespective of whether the translation machinery was present or not, the RNA cleavage could not have been due to any component of the translation machinery or any contaminants in the purified proteins/ribosomes. These shorter RNAs did not appear in the sample in which the AAECs were incubated for the same length of time as for translation but with only translation buffer, so was not solely due to the long incubation time either. The only difference between the mock translation reaction, where the shorter species were observed, and the samples containing translation buffer only, where the shorter RNAs were not observed, was the presence or absence of GTP, respectively. This would suggest that the cleavage of the RNA is somehow a consequence of the high amount of GTP in the reaction. It is in fact known that a high concentration of the non-cognate NTP can cause intrinsic hydrolysis of the RNA by the RNAP (Sosunov *et al.* 2003). When NTPs were added to the reaction after mock translation, the full length RNA product (77 nt) was restored, along with the 78 nt product, suggesting that the RNAP remained in an active AAEC. The lack of RNA in the supernatant after cleavage also suggested the RNA remained bound by the RNAP.

The above experiment only revealed the state of the RNA before and after translation. However, the GTP concentration increased during the translation reaction as new components were added to the main reaction for initiation and elongation (Figure 5.5 A). To examine more closely the effect of GTP throughout the course of

translation, a time course experiment of mock translation (translation buffer and GTP only) was carried out. The TR-CTT reaction was repeated using 5' labeled RNA and samples taken at various time points throughout the reaction (Figure 5.5 B). In the experiment shown in Figure 5.5 B, lane 1 shows the RNA species in the reaction after transcription and washing of the complexes. Lane 2 shows the sample taken 5 minutes after the ribosomes and GTP had been added to the AAECs. In this sample, the 78 nt RNA product had appeared but there was no appearance of the shorter RNA products. As the incubation time and the concentration of GTP increased, the shorter RNA products became more abundant and the longer RNA products began to disappear. In the sample taken after translation elongation (lane 6), the longer RNA products had all but disappeared, but they reappeared in the presence of 100 μM NTPs (lane 7). Taken together, these observations suggested the long incubation in the presence of a high concentration of GTP caused RNAP to misincorporate GTP into the RNA using the abasic site of the DNA as the template (Figure 5.5 B step 2), leading to backtracking of the RNAP along the DNA template (step 3). The RNAP then cleaved the RNA in the active centre by intrinsic hydrolysis to produce a new 3' end (step 4). Addition of NTPs allowed the RNAP to transcribe up to the abasic site once again (step 1). In all likelihood, in the reaction containing NTPs, the RNAP was cycling rapidly through all steps, but, as the main RNA species observed when the reaction was stopped were the full length 77 nt product and the 78 nt misincorporation product, the forward transcription reaction appeared to be faster than the backtracking and RNA cleavage reactions.

The above results could explain the lack of signal from TR-CTT, even after translation initiation on the longest RNA. Currently, it would seem that the RNAP was actually located closer to the peptide-coding region than initially thought therefore the RNA product itself was also a lot shorter. This may have resulted in the RNAP being too close to the Shine Dalgarno region for translation initiation and elongation, even on the longer RNAs.

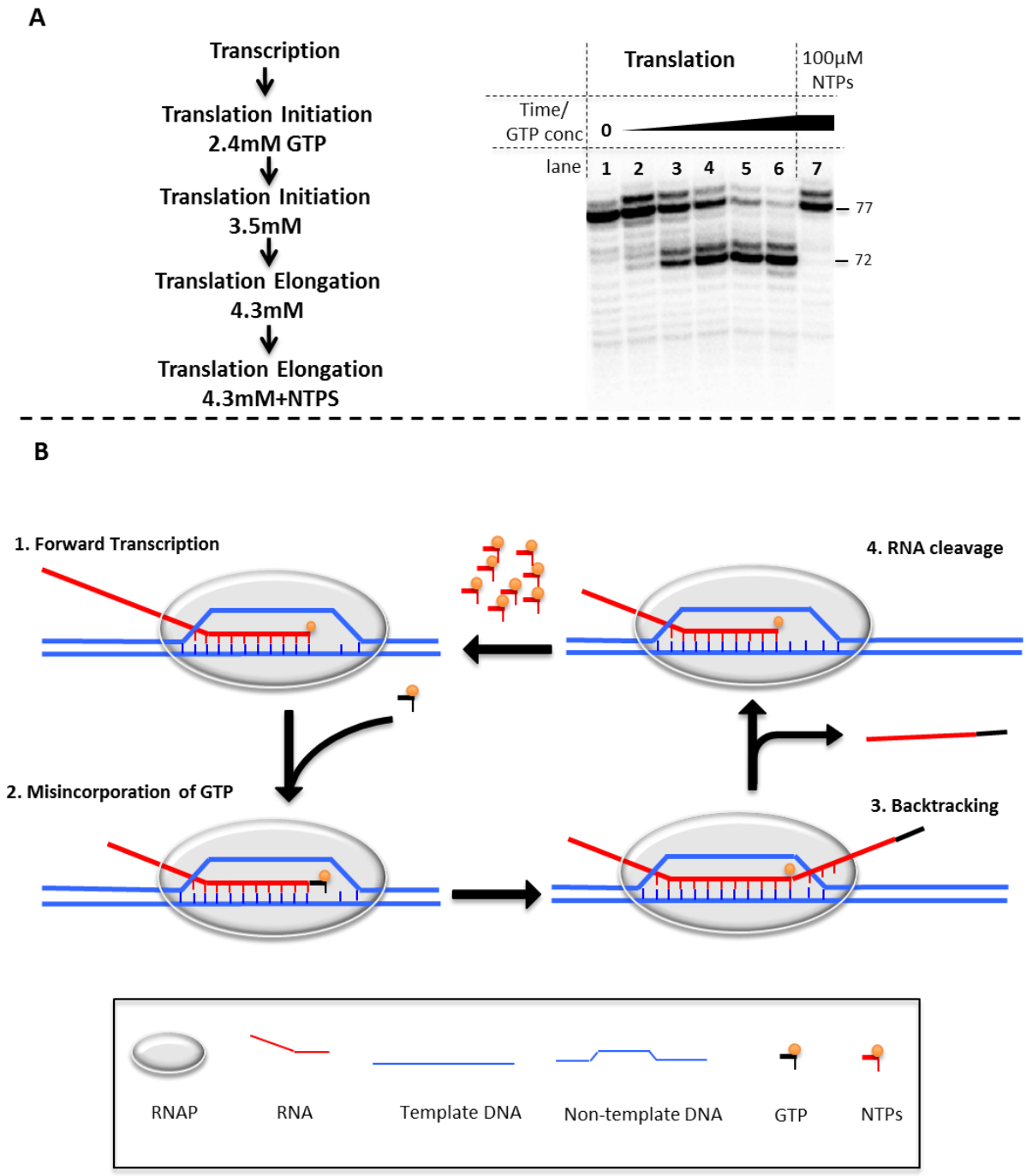


Figure 5.5 The effect of a high GTP concentration on the AAECs. A) PAGE of the 5' end labeled RNA samples taken over time during the translation stage of TR-CTT. Samples were taken at 1 min, 5 min, 10 min (translation initiation 2.4 mM GTP), 20 min (translation initiation 3.5 mM) and 40 min (translation elongation 4.3 mM) time points during the translation stage. Lane 1 contains the RNA sample taken just after transcription. Lane 7 contains the sample to which all NTPs were added after translation elongation. The RNA size is indicated on the right of the gel image. **B)** Schematic representation of the proposed cycle of the AAEC in high GTP conditions in the presence of 100 µM final CTP and ATP concentrations. 1. RNAP transcribes to position the abasic site in the i+1 site. 2. Misincorporation of GTP (black) opposite the abasic site by the RNAP. 3. Backtracking of the RNAP along the DNA template. 4. Cleavage of the RNA in the active centre to produce a new 3' end of the RNA in the i site and an empty i+1 site for NTP incorporation.

In order to fully understand what effect the high GTP concentration had on the RNAP and the AAEC, the paused state of the RNAP after transcription to the abasic

site was characterised. The AAEC was assembled using 5’ labeled RNA, immobilised on streptavidin beads and the NTPs ATP, CTP and GTP added to a final concentration of 100 μM. After transcription elongation, the AAECs were washed to remove all free RNA and NTPs, as standard for the transcription stage of TR-CTT. The reaction was made up to a final volume of 350 μL, before being divided into two 170 μL reactions. One reaction was aliquoted into 10 μL aliquots and 1 μL of either transcription buffer, 5 mM pyrophosphate (PPi, 500 μM final), 5 pmol GreB, 100 pmol GreA or 40 mM GTP (4 mM final) were added and the reactions stopped after 0, 5, 10 and 30 min incubations at 37°C. This characterised the state of the RNAP immediately after transcription elongation. GreA/B stimulate the cleavage of RNA in short or long backtracked AAECs respectively and PPi stimulates the removal of the 3’ NMP in the i+1 site when the RNAP is in the pretranslocated state. 4 mM GTP was also added to one aliquot as a positive control and transcription buffer added to another as a negative control.

The results are shown in Figure 5.6. In the absence of GTP, a proportion of the RNAP molecules were backtracked after elongation, as shown by the cleavage of RNA in the presence of GreB (Figure 5.6, left, lanes 12-14). In the samples containing 4mM GTP, the appearance of the 1 nt longer RNA followed by the 4-5 nucleotide shorter RNA pattern was observed, as seen previously (Figure 5.6, left, lanes 15-18). This is in contrast to previous analysis of the AAEC that indicated the RNAP was stabilised in the post-translocated state after transcription elongation to the abasic site (section 4.3, Figure 4.7 D, bottom). However, during these earlier experiments, the longest time point taken was 10 min and the RNA was radiolabelled during transcription elongation. During the analysis shown in Figure 5.6, samples were taken after 5 min, 10 min and 30 min and GreB cleavage did not occur at a noticeable amount until after 10 min, which explains the discrepancy.

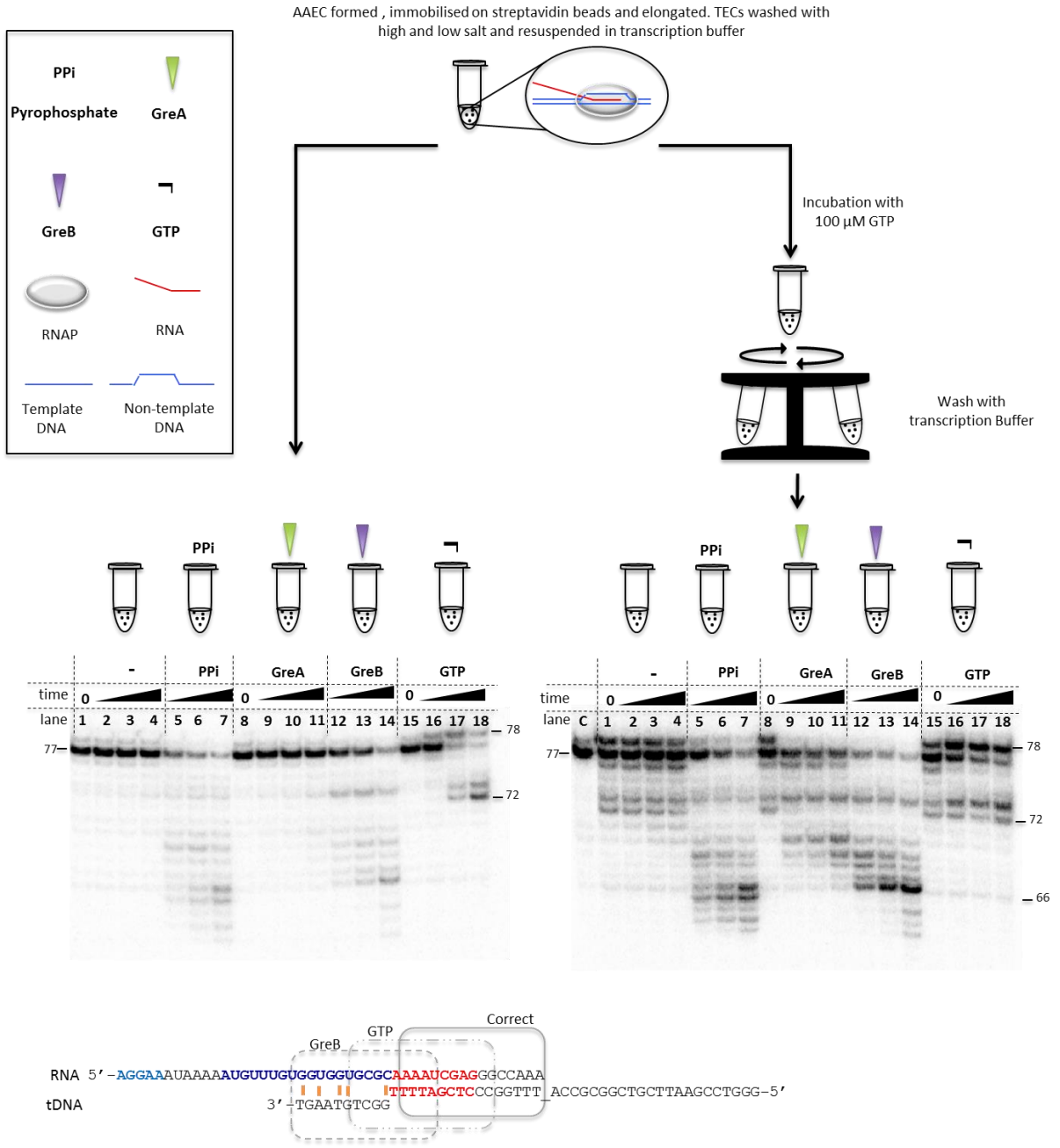


Figure 5.6 Characterisation of the RNAP in the presence and absence of GTP. Pause characterisation of the RNAP on the original abasic DNA template. The gel on the left shows the characterisation of the RNAP after transcription and washing of the AAECs and the gel on the right shows the characterisation after incubation at 37°C for 40 min in the presence of 100 µM final GTP concentration. Samples were taken at 0, 5, 10 and 30 min timepoints. Lanes 1, 8 and 15 contain the control sample taken before addition of buffer, PPi, GreA/B or 4mM GTP. The lane c in the incubated gel contains the sample taken before incubation with GTP (the same sample as in lanes 1, 8 and 15 on the left hand gel). The _ in the DNA template indicates the abasic site. The RNA and DNA template sequences are shown with the base pairing to create the extended hybrid indicated by orange horizontal lines. The grey boxes indicate the portion of the RNAP from the active centre to the proposed rear end (the 15 nts of RNA covered by the RNAP). All RNA is radiolabelled at the 5' end and the RNA sizes indicated.

To the second 170 µL aliquot, GTP was added to a final concentration of 100 µM and the reaction incubated at 37°C for 40 min. The reaction was washed with transcription buffer, aliquoted and characterised as with the other half of the

transcription elongation reaction and the results shown in Figure 5.6, right. A low concentration of GTP combined with a long incubation time caused the RNAP to misincorporate the GTP opposite the abasic site, but did not result in RNA cleavage. This allowed characterisation of the RNAP after misincorporation of the GTP but before RNA cleavage. After incubation of the transcription elongation reaction for 40 min with 100 μM GTP, the 78 nt product appeared, as well as a small amount of the shorter RNAs (Figure 5.6, right, lane 1). Characterisation of this complex showed that the AAEC on the 78 nt RNA was backtracked, as demonstrated by the cleavage of the RNA in the presence of both GreA and GreB (Figure 5.6 A, right, lanes 9-14 compared to lane 8). In contrast to the cleavage in the presence of GreB when the complexes were not pre-incubated with GTP, the cleavage after incubation occurred at a much faster rate. Without pre-incubation with 100 μM GTP, most of the full length RNA was only cleaved after 30 min, whereas after the AAECs were incubated with 100 μM GTP, the full-length products have mostly disappeared by 5 min. In contrast to with GreB, only the 78 nt complexes were cleaved in the presence of GreA but at a similarly fast rate to GreB. The 78 nt RNA was also rapidly cleaved in the presence of PPi by endopyrophosphorolysis. Based on the size of the RNA cleavage products after addition of GreB and PPi, it would appear that the RNAP was backtracked along the DNA template, up to 12 nts.

The misincorporation of GTP into the RNA opposite the abasic site by itself did not have a significant impact on the system, other than that the RNA was one nucleotide longer than planned, which has to be taken into account when calculating the distances. The real problem occurred when this misincorporation event caused the RNAP to backtrack heterogeneously to different locations on the DNA template. As the precise location of the RNAP on the DNA template could not be determined, the distance between the RNAP and ribosome was unknown. The subsequent cleavage of the RNA in the high GTP conditions also presented an obstacle but this was resolved by the addition of NTPs.

Based on the pattern of GreB cleavage, the RNAP appeared to backtrack along the DNA template to the 3’ end of the template DNA, upstream of the original DNA:RNA hybrid region (Figure 5.6). Analysis of the upstream region of the abasic DNA template and the RNA revealed positions at which base pairing could occur between the RNA and DNA and result in an extended complementary region upstream of the hybrid, along which the RNAP could backtrack (Figure 5.6). It may

be that, although the misincorporation event caused the backtracking, the formation of this extended complementary upstream region allowed the RNAP to backtrack all the way to the 3' end of the DNA template.

5.4 Design, characterisation and testing of new DNA templates

To eliminate the complementary region between the DNA and RNA upstream of the RNA:DNA hybrid, new DNA templates were designed and the state of the RNAP before and after extended incubation with GTP was assessed, as done with the original abasic template. First, the template sequence upstream of the RNA:DNA hybrid sequence was adjusted and the RNA:DNA hybrid site reduced by one base pair at the upstream edge to reduce potential formation of the extended hybrid and backtracking of the RNAP (template 1, Table 5.1). Analysis of the state of the AAEC on this new template revealed that changing this sequence alone did not eliminate the backtracking after misincorporation opposite the abasic site in the presence of 4 mM GTP, suggesting that the sequence of the RNA and DNA downstream of the hybrid site also contributed to the backtracking of the RNAP (data not shown).

The template sequence was again modified and the state of the RNAP characterised as for the original abasic template. The new upstream sequence from template 1 was kept and instead the downstream sequence was changed to produce template 2 (Table 5.1), as it was possible that this 7 nt sequence contributes to the occurrence of backtracking in the first place. Indeed, the new downstream template DNA sequence eliminated the backtracking of the RNAP in the absence of GTP (Figure 5.7, left). However, during the 100 μM GTP incubation and in the presence of 4 mM GTP, the RNAP misincorporated the GTP opposite the abasic site and beyond to extend the RNA by up to an additional 5 nts (Figure 5.7, left, lanes 16-18 (4 mM GTP) and right lane 1 (100 μM incubation)). GreA/B analysis of these AAECs showed that the RNAP backtracked to the abasic site after over-extension of the RNA by up to 4 nts beyond the abasic site (Figure 5.7, right, lanes 9-14). Although the AAECs extended heterogeneously beyond the abasic site, all AAECs were located homogeneously in the correct position on the DNA template after backtracking. However, this DNA template was also not suitable for use in the TR-CTT system because, *in vivo*, the ribosomes have been proposed to be able to rescue backtracked AAEC complexes (Schweimer *et al.* 2010) and thus, in our TR-CTT assays, they could affect the positioning of the backtracked RNAP by pushing it forwards, beyond the abasic site.

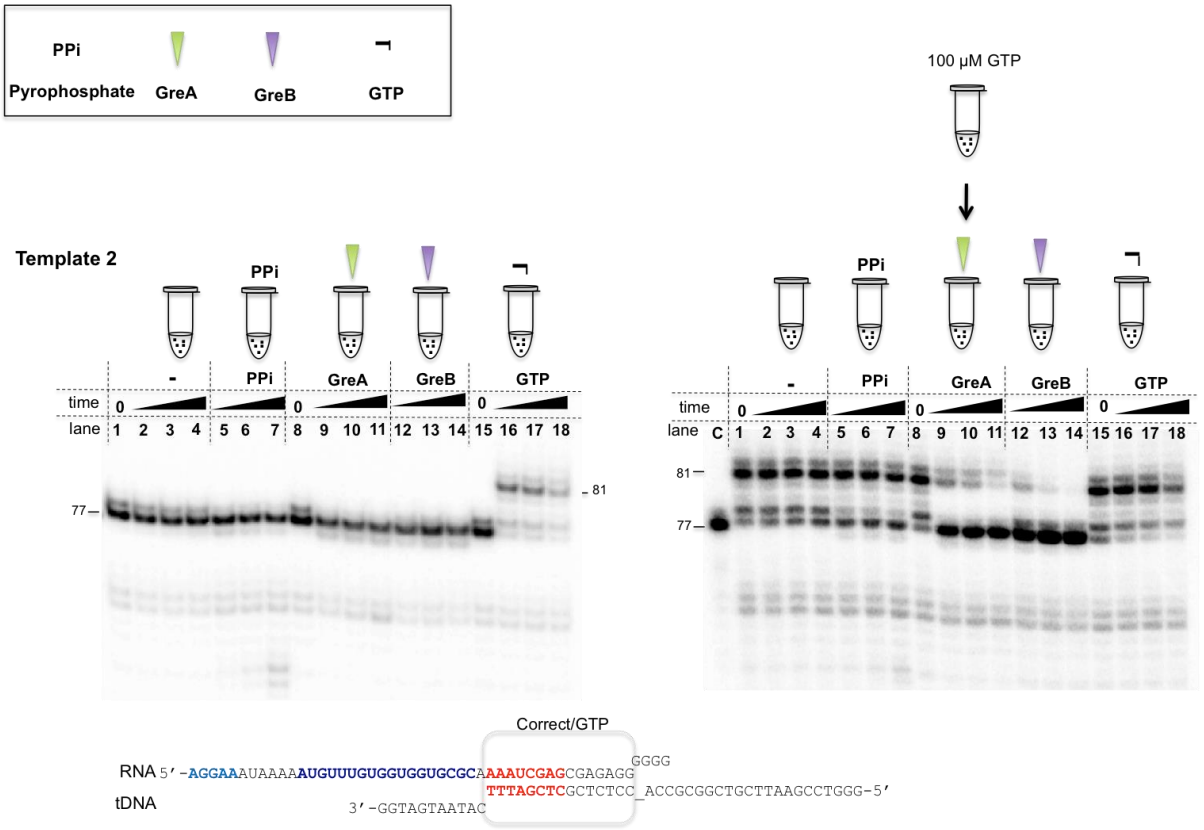


Figure 5.7 Pause characterisation of RNAP on DNA template 2. The gel on the left shows the characterisation of the RNAP after transcription and washing of the AAECs and the gel on the right shows the characterisation after incubation at 37°C for 40 min in the presence of 100 µM final GTP concentration. Samples were taken at 0, 5, 10 and 30 min timepoints. Lanes 1, 8 and 15 contain the control samples taken before addition of buffer, PPi, GreA/B or 4mM GTP. The lane c in the incubated gel contains the sample taken before incubation with GTP (the same sample as in lanes 1, 8 and 15 on the left hand gel). The _ in the DNA template indicates the abasic site. The RNA and DNA template sequences are shown and the grey box indicates the portion of the RNAP from the active centre to the proposed rear end (the 15 nts of RNA covered by the RNAP). All RNA is radiolabelled at the 5' end and the RNA sizes indicated.

Template	Sequence	Characteristics	AAEC state after elongation
Original	(RNA) GGUGGUGCGCAAAAUCGAGGGCCAAA TTTTAGCTCCCGGTTT_ACCGCGG (DNA) TGAATGTCGG	Original Abasic Template	Backtracked
Template 1	(RNA) GGUGGUGCGCAAAAUCGAGGGCCAAA (DNA) TTTAGCTCCCGGTTT_ACCGCGG GGTAGTAATAC	Original DS New UP with hybrid reduced to 8 nt	Backtracked
Template 2	(RNA) GGUGGUGCGCAAAAUCGAGCGAGAGG TTTAGCTCGCTCTCC_ACCGCGG (DNA) GGTAGTAATAC	New DS and Template 1 UP	Backtracked after read-through RNAP located at abasic site
Template 3	(RNA) GGUGGUGCGCAAAAUCGAGCGAGAGG TTTAGCTCGCTCTCC_AGGGCGG (DNA) GGTAGTAATAC	Template 1 DS, Template 2 UP AGG after abasic site	Stabilised in post-translocated position No backtracking or read-through
Template 4	(RNA) GGUGGUGCGCAAAAUCGAGCGUGUGA TTTAGCTCGCACACT_ACCGCGG (DNA) GGTAGTAATAC	Template 2 DS and after abasic site, New UP for base pairing with cordycepin (modified ATP)	No Read-through Post-translocated RNAP at abasic site
Template 5	(RNA) GGUGGUGCGCAAAAUCGAGCGAGAGG TTTTAGCTCGCTCTCC_ACCGCGG (DNA) GGTAGTAATA	Based on template 3 Extended UP edge of hybrid site-9nt in total	No increase in % of RNA extended after AAEC assembly, RNAP Pretranslocated
Template 6	(RNA) GGUGGUGCGCAAAAUCGAGCGAGAGG GTTTTAGCTCGCTCTCC_ACCGCGG (DNA) GGTAGTAAT	Based on template 3 Extended UP edge of hybrid site-10nt in total	Increase in % of RNA extended after AAEC assembly, RNAP Pretranslocated
Template 7	(RNA) GGUGGUGCGCAAAAUCGAGCGAGAGG CGTTTTAGCTCGCTCTCC_ACCGCGG (DNA) GGTAGTAA	Based on template 3 Extended UP edge of hybrid site-11nt in total	Most RNA species extended RNAP Pretranslocated
Template 8	(RNA) GUGGUGCGCAAAAUCGAGCGAGAGG TTTAGCTCGCTCTCC_AGGGCGG (DNA) TGAATGTCGG	Template 3 UP Original template DS	Does not restore extension of RNA Stabilised in pre-translocated position

Table 5.1 Summary of DNA templates. DS=downstream sequence UP=upstream sequence (both DS and UP relative to RNA:DNA hybrid sequence. RNA:DNA hybrid site shown in red. The abasic site is represented by _. The peptide-coding region of the RNA is shown in purple.

As the new upstream and downstream sequences used in template 2 eliminated the backtracking of the RNAP both in the presence and absence of GTP but resulted in overextension of the RNA beyond the abasic site, the next step was to modify the bases before and after the abasic site to reduce read-through (and subsequent backtracking). New templates were produced through modification of template 2. A number of different combinations of nucleotides before and after the abasic site were trialed to reduce the misincorporation of GTP. These new templates were characterised as for all of the previous templates and, based on this analysis, two templates were identified as potentially suitable for use in TR-CTT (Table 5.1, templates 3 and 4). Template 3 has the same sequence as template 2, but with the region after the abasic site modified to contain two dGTP residues to try to reduce the misincorporation after the abasic site that was seen in template 2. As GTP is the nucleotide that was misincorporated, the template nucleotide that it is least likely to pair with is dGTP. When characterised, the RNAP remained stabilised in the post-translocated state in the absence of GTP (Figure 5.8, left). However, in the presence of a high concentration of GTP, the RNAP misincorporated GTP opposite the abasic site (Figure 5.8, left lanes 15-18). Characterisation of the AAEC after the extended incubation with 100 µM GTP revealed that the RNAP was stabilised in the pre-translocated state after misincorporation of GTP to produce the 78 nt RNA product (Figure 5.8, right). In the presence PPi, GreA and GreB, the 78 nt product was cleaved to the 77 nt product. Template 3 was characterised after transcription with all the MFVVVR extended RNAs (MFVVVR+1 to +5) and the RNAP was shown to remain stable in the post-translocated state in the absence of GTP (after 77 nt RNA synthesis), and in the pre-translocated state (after 78 nt RNA synthesis) in the presence of GTP (appendix A5 Figure 9.1).

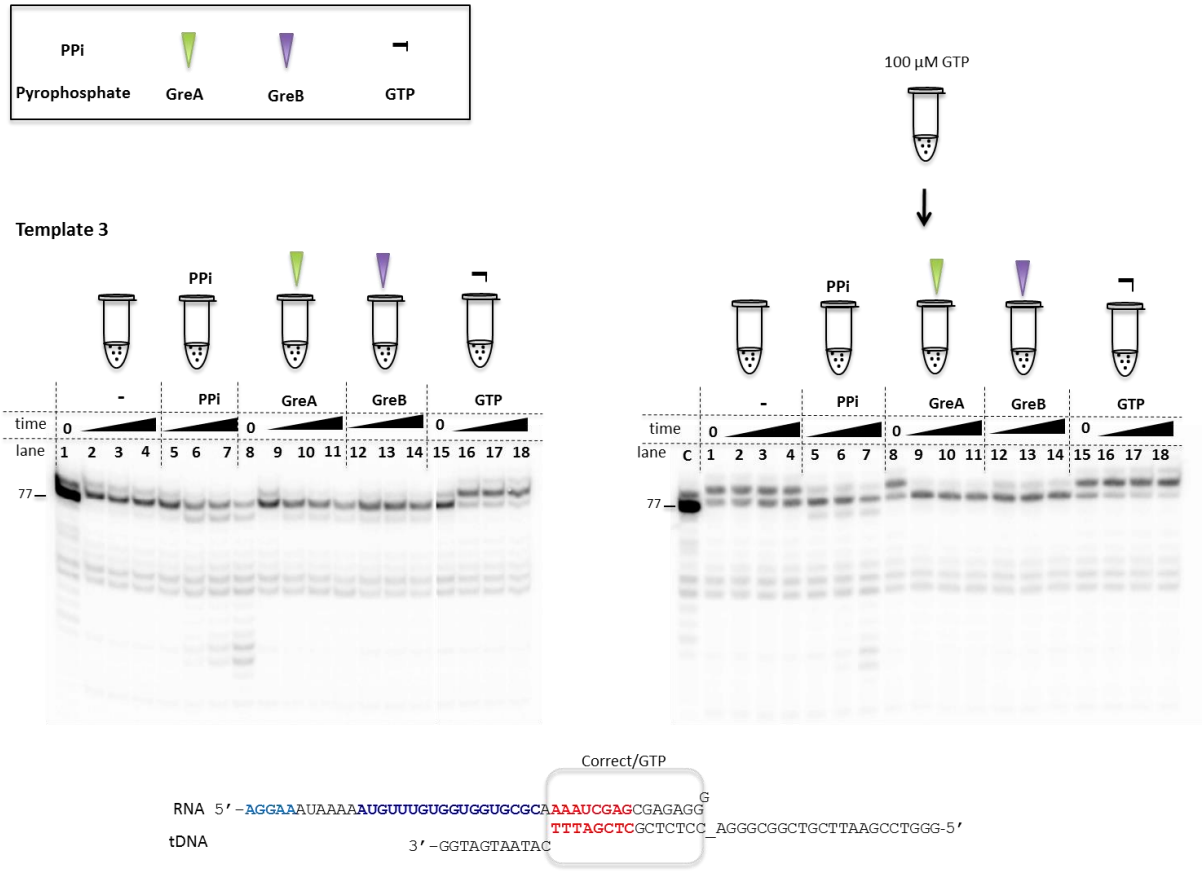


Figure 5.8 Pause characterisation of the RNAP on DNA template 3. Pause characterisation of the RNAP on DNA template 3.The gel on the left shows the characterisation of the RNAP after transcription and washing of the AAECs and the gel on the right shows the characterisation after incubation at 37°C for 40 min in the presence of 100 µM final GTP concentration. Samples were taken at 0, 5, 10 and 30 min timepoints. Lanes 1, 8 and 15 contain the control samples taken before addition of buffer, PPi, GreA/B or 4mM GTP. The lane c in the incubated gel contains the sample taken before incubation with GTP (the same sample as in lanes 1, 8 and 15 on the left hand gel). The _ in the DNA template indicates the abasic site. The RNA and DNA template sequences are shown and the grey box indicates the portion of the RNAP from the active centre to the proposed rear end (the 15 nts of RNA covered by the RNAP). All RNA is radiolabelled at the 5' end and the RNA sizes indicated.

Template 4 contains the same sequence upstream of the RNA:DNA hybrid and after the abasic site as template 2, but the sequence between the RNA:DNA hybrid site and the abasic site was modified to allow incorporation of cordycepin as the last nucleotide of the RNA during elongation. Cordycepin is a modified ATP analogue that lacks the 3' OH group required for catalysis and incorporation of the next nucleotide and so leads to stalling of transcription. Figure 5.9 shows characterisation of template 4 both before and after incubation with GTP. After transcription elongation to synthesise the 77 nt RNA product and washing with high salt and transcription buffer, the RNAP appeared to be stably positioned with the 3' RNA in the i site (i.e. not backtracked or pretranslocated) (Figure 5.9, left). On the original abasic template, incubation with 4 mM GTP of up to 30 min resulted in misincorporation of the GTP, backtracking of the RNAP and cleavage of the RNA

after transcription elongation. On the contrary, on template 4, the presence of GTP did not affect the state of the AAEC after transcription elongation (Figure 5.9, lanes 15-17). The 40 min incubation with 100 μM GTP resulted in misincorporation and backtracking but no RNA cleavage on the original abasic template (Figure 5.6, left) while there was no effect on the state of the AAEC using template 4 (Figure 5.9, right).

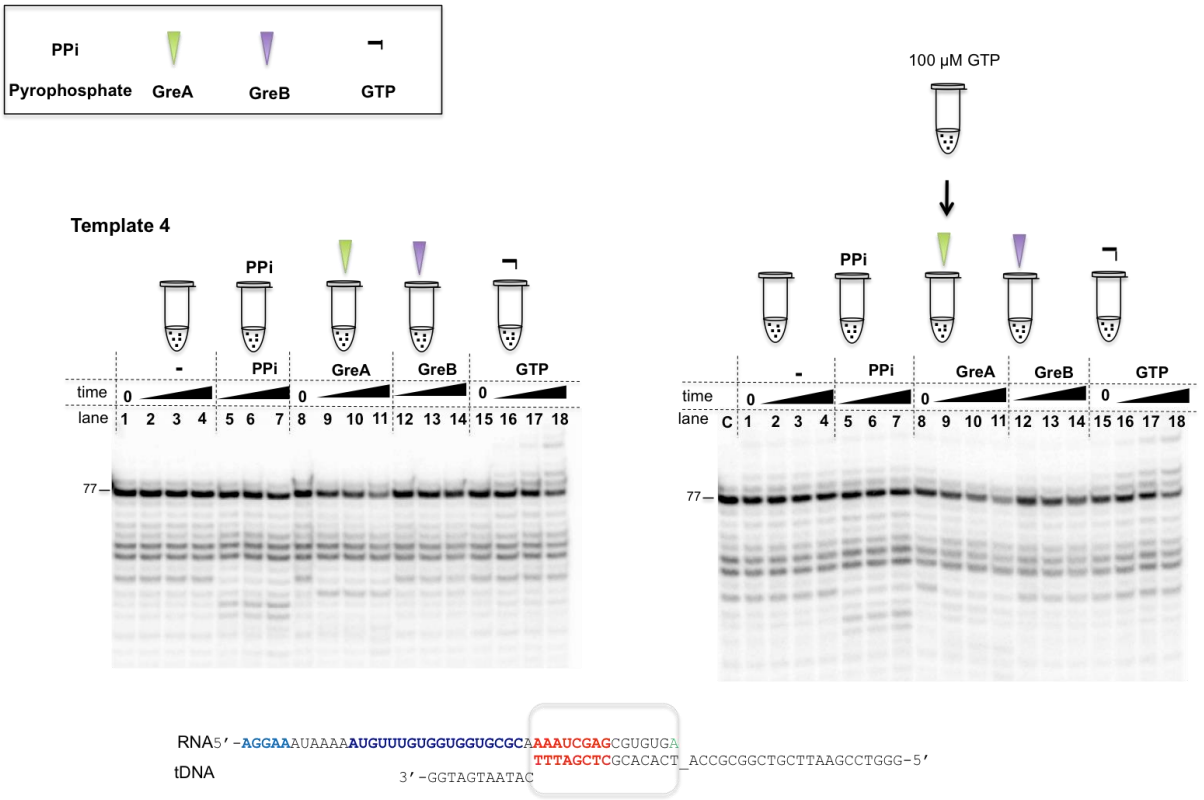


Figure 5.9 Pause characterisation of the RNAP on DNA Template 4. The gel on the left shows the characterisation of the RNAP after transcription and washing of the AAECs and the gel on the right shows the characterisation after incubation at 37°C for 40 min in the presence of 100 μM final GTP concentration. Samples were taken at 0, 5, 10 and 30 min timepoints. Lanes 1, 8 and 15 contain the control sample taken before addition of buffer, PPi, GreA/B or 4mM GTP. The lane c in the incubated gel contains the sample taken before incubation with GTP (the same sample as in lanes 1, 8 and 15 on the left hand gel). The _ in the DNA template indicates the abasic site. The RNA and DNA template sequences are shown with the base pairing to create the extended hybrid indicated by orange horizontal lines. The grey box indicates the portion of the RNAP from the active centre to the proposed rear end (the 15 nts of RNA covered by the RNAP). All RNA is radiolabelled at the 5' end and the RNA sizes indicated.

Based on the above results, both templates 3 and 4 had the potential for use in TR-CTT. Template 3 was suitable for use with all RNAs and did not require the use of modified NTPs during transcription elongation to prevent read-through by the RNAP. Although the RNAP misincorporated GTP opposite the abasic site, the AAEC did not backtrack, in contrast to the original abasic template. Characterisation of the RNAP after incubation with GTP showed that the RNAP was stabilised in the pre-

translocated state with the abasic site located in the i+1 site. Although, for this reason this template is not ideal, it is still compatible for use in TR-CTT because the location of the RNAP is known and is homologous within the reaction.

As described in the previous sections, the TR-CTT method was modified to include the addition of GreA during AAEC assembly to resolve the backtracked AAECs initially formed due to the extra NMP at the RNA 3' end and the extended complementary region between the RNA and DNA upstream of the RNA:DNA hybrid. This was necessary because, without the addition of GreA during AAEC assembly, only a proportion of the RNA in the system was extended. Templates 3 and 4 eliminated the backtracking of the RNAP and the misincorporation past the abasic site, but analysis of transcription using both of these templates revealed that they also reduced the effect of GreA during AAEC formation (appendix A10 Figure 9.6). The RNA:DNA hybrid formed with both templates 3 and 4 is one base shorter than with the original abasic DNA template (8 nt instead of 9). Although this length of hybrid is still sufficient for formation of the AAEC and transcription elongation (Sidorenkov *et al.* 1998), it is possible that the shortening of the initial RNA:DNA hybrid may reduce the backtracking of the RNAP during AAEC assembly.

Three more templates were designed that extended the RNA:DNA hybrid at the upstream edge by up to 3 nucleotides, creating a hybrid 9, 10 and 11 nucleotides in length based on the MFVVVR RNA upstream sequence and template 3 (template 5, 6 and 7 respectively, Table 5.1).These templates were designed with the aim of restoring the backtracking during AAEC assembly required for efficient transcription elongation. Templates 5, 6 and 7 were tested in transcription, and an increase in hybrid length was observed to result in an increase in sensitivity of the AAEC to the effect of GreA (appendix A6 Figure 9.2 A). This led to an increase in the proportion of the RNA being actively transcribed after AAEC assembly. The state of the RNAP after transcription elongation and in the presence of GTP on all three templates was characterised as for the previous templates. On incubation with 4 mM GTP, the RNAP misincorporated GTP opposite the abasic site but did not transcribe beyond (appendix A7 Figure 9.3). After misincorporation, the RNAP was still backtracked by one base and/or in the pre-translocated state but was stable, as was observed previously on template 3 before modification to increase the hybrid length (appendix A7 Figure 9.3). Although the RNA was extended more efficiently by RNAP after AAEC with the extended hybrid, there were also shorter RNA cleavage products

retained, even after washing with high salt and transcription buffer. Templates 5, 6 and 7 were then tested in transcription with each of the extended MFVVVR RNAs (MFVVVR+1 to +5, appendix A8 Figure 9.4). As the RNA length was increased by adding nucleotides to the RNA sequence before the hybrid site, the sequence directly upstream of the hybrid varied between the RNAs. The three new templates were designed based on the original RNA and as such, transcription elongation was not as efficient with all of these RNAs compared to with the original RNA. Because of this, and due to the presence of the shorter RNAs after high salt washing of the AAECs, templates 5, 6 and 7 were deemed unsuitable for use in TR-CTT for the purpose of this project.

The complementarity between the RNA and DNA appeared to be required for backtracking of the RNAP during initial AAEC assembly and this backtracking was essential for hydrolysis of the 3' end of the RNA to removed additional or incorrect NTPs that prevented transcription elongation (Figure 5.3). A new template, template 8 was designed based on the sequence of template 3 but with the entire upstream sequence modified to restore the complementary region upstream of the RNA:DNA hybrid by switching back to the original upstream sequence (Table 5.1). Transcription using this template however revealed that switching the upstream sequences did not restore the effect of GreA during AAEC assembly (appendix A9 Figure 9.5). This would suggest that it was the unique combination of the upstream and downstream DNA sequence of the original template that caused the backtracking of the RNAP necessary for hydrolysis and extension of the RNA during AAEC formation.

RNaseH cleavage of the RNA after synthesis by T7 RNAP to remove the 2/3 3' terminal NMPs was previously tested as a potential solution to increase transcription elongation by RNAP after AAEC formation but was discarded in favour of adding GreA during AAEC assembly (section 5.2.1, Figure 5.2). After the discovery that GreA was no longer effective during AAEC and that, because of this, the RNA was not extended properly after AAEC on templates 3 and 4, the use of RNaseH digested RNA was tested once again. The AAECs were assembled with RNaseH digested MFVVVR RNA and template 3 or 4, before NTPs were added for transcription elongation. When the amount of extension by the RNAP during transcription elongation was compared between the original abasic template and templates 3 or 4, it revealed that there was barely any extension of the RNA with these two templates (3 and 4) (appendix A10 Figure 9.6). Taking into account the shortened hybrid

sequence of the DNA template at the upstream edge and the cleavage of the two nucleotides of the RNA at the downstream edge, the RNA:DNA hybrid was now only 6 bases in length, too short for proper formation of a stable AAEC (Sidorenkov *et al.* 1998). DNA templates 5, 6 and 7 increase the length of the RNA:DNA hybrid at the upstream edge (Table 5.1). Although the RNaseH digested RNA did not form a stable AAEC when assembled on template 3, extending the RNA:DNA hybrid to 7, 8 or 9 nts may restore the stability and result in transcription elongation. The RNaseH digested RNA was also tested in transcription with DNA templates 5, 6 and 7 in the presence of GreA (appendix A6, Figure 9.2). The results revealed that the amount of transcription elongation increased along with the increase in hybrid length, but not to the same level as with the original abasic template. The results also showed that the presence of GreA was required for transcription elongation and shorter RNA products from GreA cleavage remained even after high salt washing, therefore the RNaseH digested RNA had no advantage over the undigested MFVVVR RNA in this instance.

Design and characterisation of DNA templates with modified upstream and downstream sequences did not lead to the identification of a new DNA template on which the RNAP was able to extend all of the RNA species after AAEC and was stably positioned in the post-translocated state after transcription elongation. Therefore, none of the redesigned DNA templates were deemed suitable for use in TR-CTT and the original abasic template still seemed to be the most appropriate template to be used for subsequent TR-CTT experiments. To counteract the effect of backtracking and RNA cleavage due to misincorporation of the GTP by the RNAP during translation, CTP and ATP were added at a final concentration of 25 μM during translation, along with 100 pmols of GreA and GreB. It must be kept in mind that during the cycle of backtracking, RNA cleavage and forward transcription that the RNA produced and used in translation will be one nucleotide longer than anticipated due to the misincorporation of the GTP opposite the abasic site.

5.5 Modifying translation to reduce the GTP concentration

An alternative method to alleviate the effect of GTP on the RNAP during translation would be to reduce the GTP concentration and incubation times. First, GTP hydrolysis assays were performed to analyse the GTP usage by the GTPase enzymes throughout translation. A 50 μL volume of 100 mM GTP stock was supplemented with 2 μL radiolabeled α-^{32}P-GTP to visualise hydrolysis of GTP to GDP and GMP. In this assay, translation was carried out as usual but with the α-^{32}P-

GTP supplemented GTP and without phosphoenolpyruvate (pep) or pyruvate kinase (PK) to prevent the GTP hydrolysis products from being regenerated. 5 µL samples were taken at each stage of translation (Figure 5.10 A) and an equal volume of translation stop buffer was added to stop the reaction (5% formic acid in translation buffer). The products were analysed using thin layer chromatography (TLC) with PEI cellulose TLC plates pre-run with H_2O for the stationary phase and 1.3M KH_2PO_3 pH 4 as the mobile phase. 1 µL of each sample was spotted 1.5 cm from the bottom of the TLC plate and dried. 1 µL of cold 10 mM GTP, GDP and GMP were spotted alongside as controls. The plate was placed in a chromatography chamber and left to run for 1 hour or until the buffer reached 1 cm from the top of the plate. The plates were dried and the control GTP, GDP and GMP samples visualised by exposure to UV light and marked by spotting of a radiolabeled sample. After drying, the plate was exposed to a storage phosphor screen and the image analysed using ImageQuant™ software. The results showed that even after translation elongation in the presence of EF-G, where there was the most GTP hydrolysis, there was still a large amount of un-hydrolysed GTP remaining in the sample (Figure 5.10 A, lane 14). This indicated that the GTP concentration could potentially be lowered.

In the original translation reaction the incubation times and GTP concentration are in excess compared to those required by the enzymes involved. The translation protocol was therefore modified to reduce the incubation times along with the GTP concentration. In the translation protocol used so far, the ribosomes are incubated with GTP and RNA for 10 min (Figure 5.10 A, step 1), then translation is initiated by adding fMet-tRNA^{fMet} and IF-1, 2 and 3 and incubating for a further 15 min (step 2) whilst the ternary complexes are formed (steps 3 & 4). The ternary complexes are formed by incubating EF-Tu, PK and EF-Ts with GTP and pep for 10 min (Figure 5.10 A step 3) before addition of the aa-tRNA^{aa} and incubation for a further 5 min (step 4). For the translation elongation reaction, the TCs and initiation reaction are combined and EF-G is added. To reduce the GTP concentration and incubation times, an alternative method of translation was trialed. All components for assembly of the ternary complexes were mixed simultaneously and the reaction incubated for 10 min (Figure 5.10 B step 1). Once the ternary complexes had been formed, all the components required for the initiation reaction (ribosomes, RNA, IFs and fMet-tRNA^{fMet}) were added to the TCs (Figure 5.10 B steps 2 & 3). After a further 5 min incubation, EF-G was added and translation elongation allowed to proceed for 5 min

(steps 4 & 5). This new method was first used to assess GTP usage in the hydrolysis assay with the final concentration of GTP lowered to either 1 mM or 0.5 mM in all stages of the translation reaction. Samples were taken throughout the translation reaction as above. Even at a final concentration of 0.5 mM GTP, there was still a large amount of unhydrolysed GTP remaining even after translation elongation (Figure 5.10 B, lane 10).

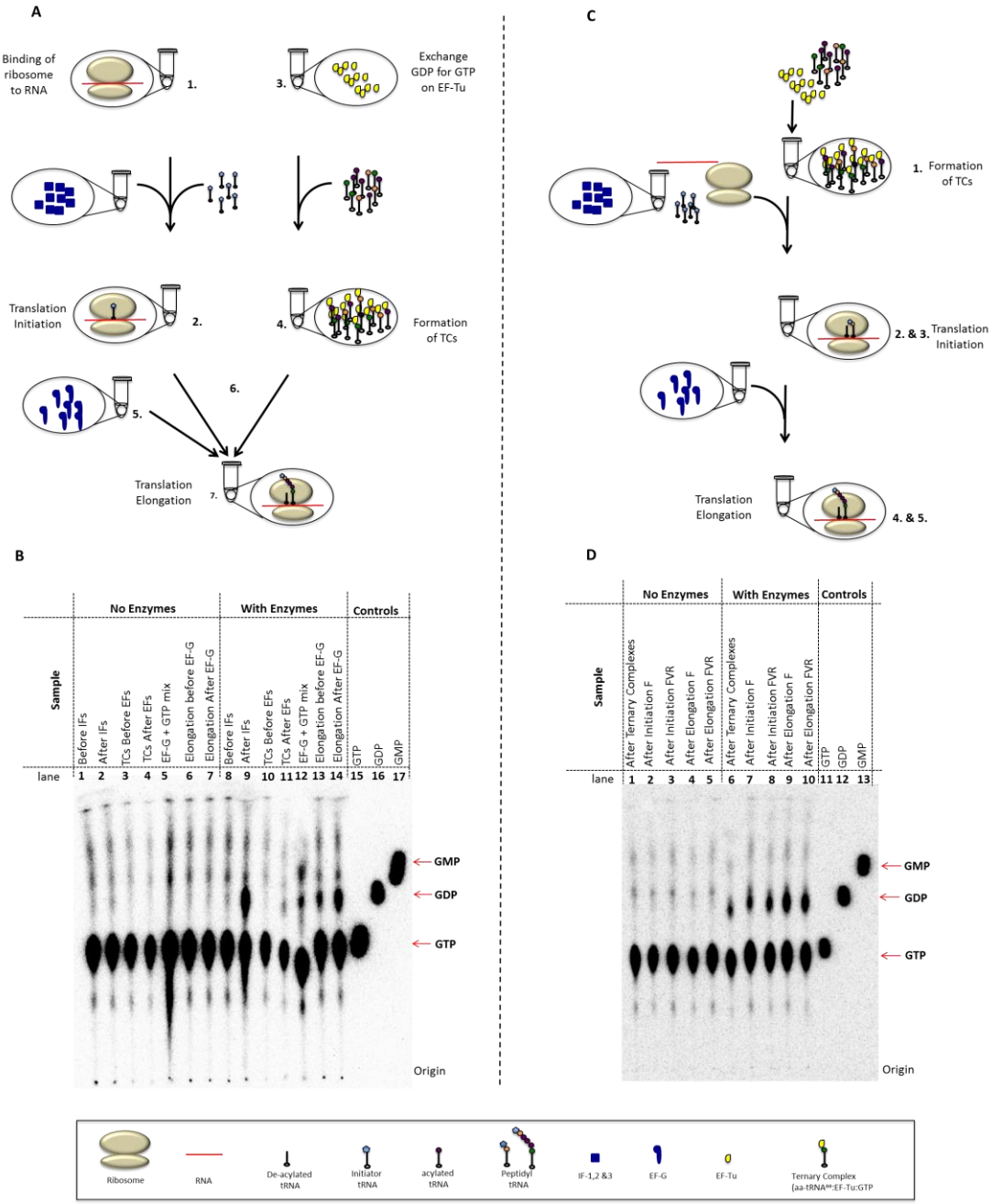


Figure 5.10. GTP hydrolysis assay of translation. A) Schematic representation of the original translation technique. B) GTP hydrolysis assay for the original translation method. C) Schematic representation of the alternative translation technique D) GTP hydrolysis assay using the modified translation method and 0.5 mM final GTP concentration at all stages. α-[^{32}P]-GTP supplemented GTP was used during translation and 10 mM GTP, GDP and GMP run alongside as controls, visualised by

UV light and marked by spotting with α-[^{32}P]-GTP. The lanes designated 'No enzymes' indicates the control translation reaction carried out in the absence of IFs and EF-G.

Based on the GTP hydrolysis assays described above, lowering the GTP concentration to 0.5 mM should still be sufficient for translation. The translation products produced using the new method of translation and with 1 mM GTP final concentration in all stages were analysed by TLE (Figure 5.11 A). The MF dipeptide (lane 6) was synthesised but not the full length MFVVVR peptide (lane 5). Increasing the GTP concentration to the original level (4.3 mM) during elongation increased the amount of MFVVVR produced slightly, but not to the levels produced using the original protocol (compare lane 4 with lanes 2 & 3). Increasing the GTP concentration in the ternary complex reactions to the original GTP concentration of 1.4 mM did not lead to production of the MFVVVR peptide either (Figure 5.11 B, lane 6). This would suggest it was not the GTP concentration during elongation or ternary complex formation that affected the synthesis of MFVVVR. As the MF dipeptide was synthesised, the ternary complex reaction was not the cause either, however the block must have occurred somewhere between formation of the ternary complexes and translation elongation. In the updated method, the TCs were incubated with the initiation reaction for 10 min, the first 5 min of which occurred before the addition of EF-G. The MF dipeptide can be formed in the absence of EF-G as no translocation of the ribosome is required (fMet-tRNA^{fMet} enters directly into the P site during initiation), but any peptides longer than that require EF-G to translocate the ribosome (Wilden *et al.* 2006; Rodnina *et al.* 1997; Agrawal *et al.* 1999). The TCs may be degrading in the presence of the ribosomes before EF-G was added the reaction, and this could explain why MF was produced, but not MFVVVR. Indeed, pre-initiating the transcription reaction before simultaneously adding TCs and EF-G restored production of the full-length peptide (Figure 5.11 C, lanes 3 & 4). Initiating and elongating translation in the same reaction (ternary complexes pre-formed then RNA, ribosomes, initiator-tRNA, initiation factors and EF-G added simultaneously) resulted in not even the MF dipeptide being synthesised (Figure 5.11 C, lane 5 & 6). Initiating translation, adding the TCs and incubating for 1 min before adding EF-G led to synthesis of the MF dipeptide, although at a lower amount compared to the control (compare lane 8 with 1), but no synthesis of the MFVVVR peptide (lane 7). Taken together, this would suggest that translation needed to be initiated and the TCs formed before adding to the elongation reaction. EF-G must be added at the same time as the TCs as, based on the above, it would appear that the aa-tRNA^{aa} rapidly

degrades if left unused for too long in the translation elongation reaction. Under these conditions, it was possible for MFVVVR to be synthesised when the GTP concentration in all reactions was lowered to as little as 0.5 mM (Figure 5.11 D, lane 5). In all reactions, there was synthesis of a product that was not MF or MFVVV. It was initially thought this was maybe an intermediate MFV or MFVV product, but later experiments excluded this possibility (data not shown).

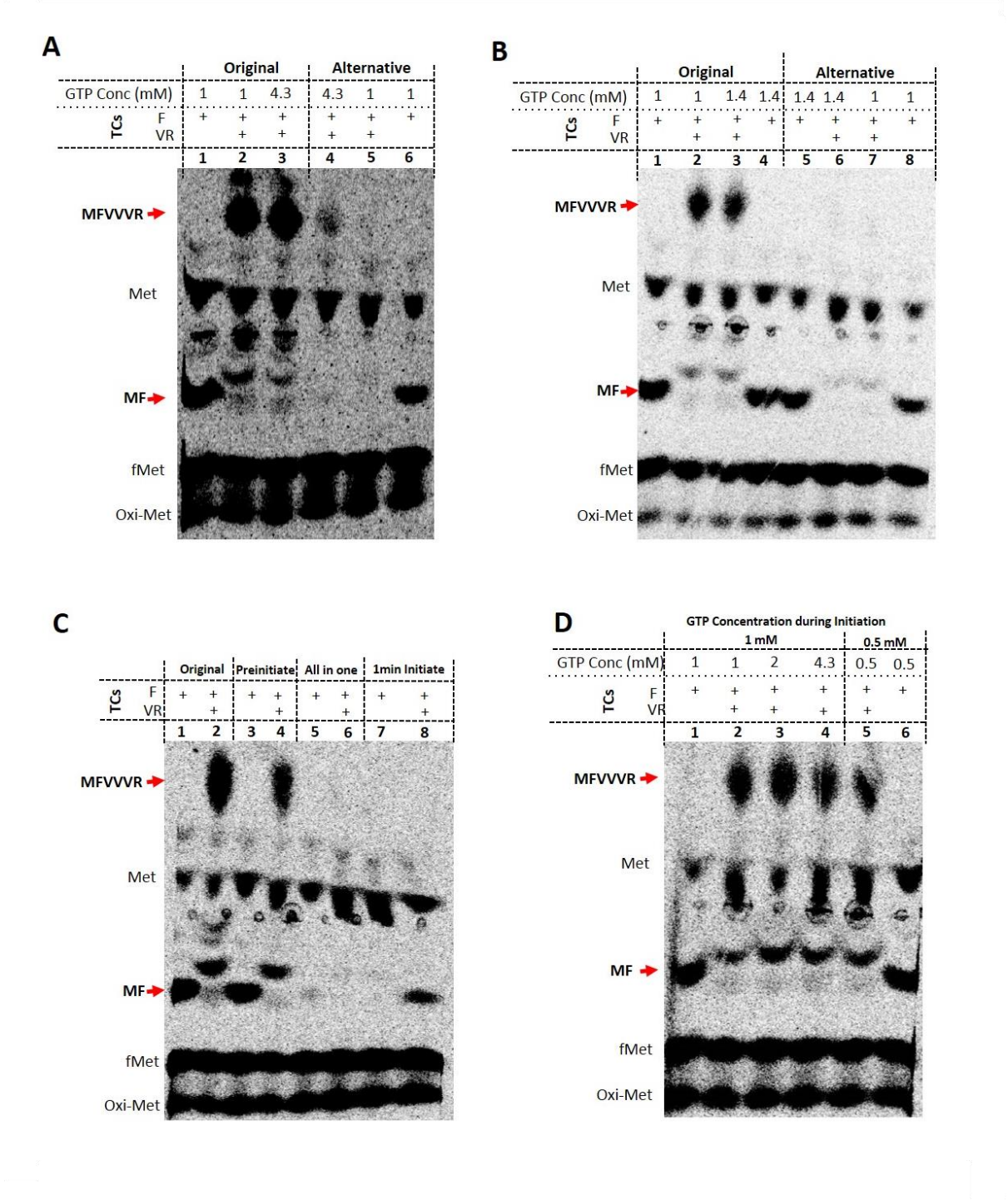


Figure 5.11 TLE analysis of GTP concentration and translation method. A) TLE comparing the original and alternative translation methods with 1 mM final GTP in all stages (lanes 1, 2, 5 and 6) and increasing the GTP concentration to 4.3 mM GTP during translation elongation (lanes 3 & 4). B) TLE

comparing the original and alternative methods with 1 mM GTP in all stages (lanes 1, 2, 7 & 8) and increasing the GTP concentration to 1.4 mM during TC formation (3-6). C) TLE comparing the original translation method (lanes 1 & 2) to the modified method with pre-initiation (lanes 3 & 4), all in one (lanes 5 & 6) and with 1 min initiation (lanes 7 & 8). D) TLE comparing the use of 1 mM (lanes 1-4) with 0.5 mM (lanes 5 & 6) final GTP concentration using the modified method with pre-initiation of translation before TC and EF-G addition. Also comparison of increasing the GTP concentration during elongation from 1 mM (lane 2) to 2 (lane 3) and 4.3 mM (lane 5) is shown.

Lowering the GTP concentration and modifying the translation technique resulted in a decrease in the efficiency of translation. Since the TR-CTT technique already resulted in a low level of peptide signal, the original translation method still seemed the most appropriate for use in TR-CTT. Moreover, using the original abasic DNA template and adding GreA, GreB and NTPs during translation appeared to resolve the effects of the high GTP concentration on the RNAP in TR-CTT (Figure 5.5). TR-CTT with the latest modifications (addition of GreA/B and NTPs during translation elongation) was repeated but still the peptide signal remained low and difficult to distinguish from the background (data not shown). It was therefore clear that further investigation was required to address the inconsistencies of our technique

5.6 The Ribosomes are unable to bind the RNA after transcription

From the initial use of the MFVVVR encoding RNA in TR-CTT and throughout the troubleshooting, the signal of the peptide after TLE of the TR-CTT samples remained low, even after the technique had been successively modified. One explanation for this may be that the amount of RNA that is retained after washing of the AAEC and subsequently used during translation was still too low. After the transcription elongation stage of TR-CTT, the AAECs were washed with both high salt and transcription buffer. High salt washing destroys any unstable complexes and removes a large proportion of the RNA. When previous transcription reactions using radiolabeled RNA were quantified, it was estimated that only 10-30% of the complexes remained after the wash steps (data not shown). Determining the minimum amount of RNA required in translation for the ribosomes to synthesise enough peptide to be visible by TLE would indicate if the low peptide signal was caused by too low an amount of RNA remaining after high salt washing of the AAECs for synthesis of a clearly identifiable peptide during translation. Translation alone was therefore carried out using a decreasing amount of RNA, the samples dried and resuspended as for TR-CTT and analysed by TLE. Even using as little as 2.5 pmols of RNA in the translation initiation reaction was enough for the ribosomes to synthesise a visible MF dipeptide product (Figure 5.12 A, lane 4). The reactions did

not contain any of the extra components from transcription that are present in the reaction after TR-CTT but any background caused by these is unlikely to alter how much RNA is required in translation initiation for the peptide to be visible.

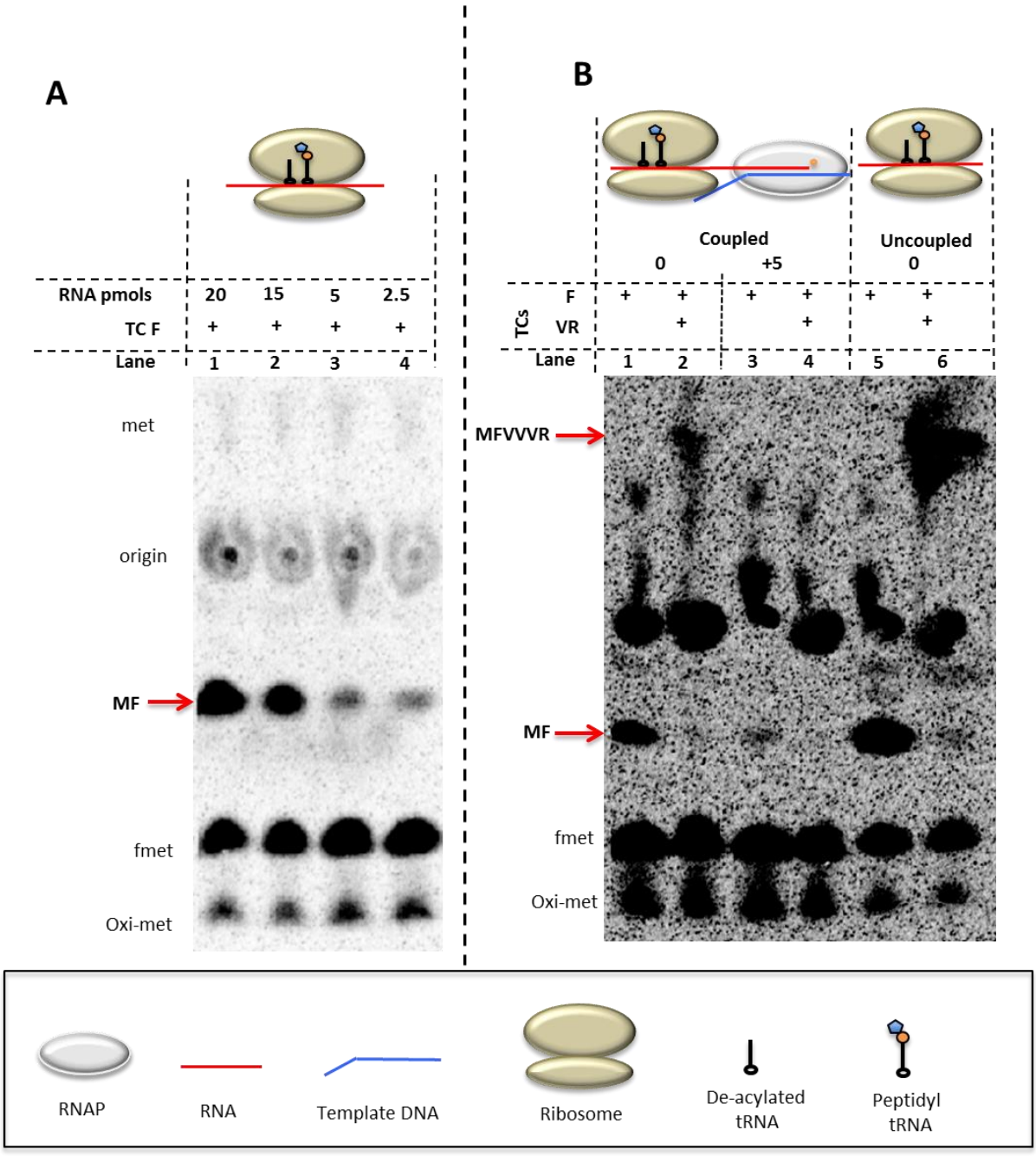


Figure 5.12 The amount of RNA required for visualisation by TLE. A) MF dipeptide synthesis using a decreasing amount of RNA during translation initiation. B) TLE analysis of MF and MFVVVR peptide synthesis during TR-CTT using the MFVVVR original RNA (lanes 1 & 2) and MFVVVR+5 extended RNA (lanes 3 & 4). Uncoupled samples using 20% of the RNA concentration of MFVVVR RNA were run alongside for comparison (lanes 5 & 6). Although not shown, non-template DNA was included in the reaction.

Based on the above observation that 20 pmols of RNA during translation produces a very clear peptide after TLE (Figure 5.12 A lane 1), TR-CTT was repeated using 200 pmols of both MFVVVR and MFVVVR+5 RNA. Translation alone

was performed alongside using the amount of RNA estimated to be retained after washing with high salt and transcription buffer (20%, 40 pmols). The concentrated samples were loaded and run on the same TLE plate to compare the amount of peptide produced in each reaction. The amount of full-length peptide produced in the translation alone reaction was vastly increased compared to that produced in TR-CTT (Figure 5.12 B, compare lanes 2 and 4 with 6). Even if the amount of RNA remaining after washing of the AAECs had been overestimated, there should still have been a visible peptide since, theoretically, only 1.25% (2.5 pmols) of the RNA used in transcription is required for a visible peptide signal after translation. More of both the MFVVVR and MF peptides were observed after TR-CTT with MFVVVR RNA than with MFVVVR+5 RNA, while preliminary results showed that no MFVVVR was produced in TR-CTT of the MFVVVR RNA but could be produced from TR-CTT using the MFVVVR+5 RNA.

The fact that the peptide signal was much lower than anticipated with the estimated amount of RNA remaining after transcription suggested that the amount of RNA entering into translation was not the cause of the lack of signal. After all the analysis and troubleshooting of the TR-CTT system outlined within this chapter and the previous chapter, the only remaining potential cause for the lack of peptide could be that the ribosomes were not initiating on the RNA as efficiently in TR-CTT compared to in translation alone.

5.6.1 RelE analysis of TR-CTT

To establish whether the ribosomes were able to initiate on the RNA, the occupancy of the ribosomes on the RNA after translation initiation needed to be determined. So far, the ability of the ribosomes to translate the RNA was assessed by analysing the peptide, but an alternative approach is achieved through analysis of the RNA. The bacterial toxin, RelE, binds ribosomes without any tRNA in their A-site and cleaves the RNA between the second and third bases of the A-site codon (Pedersen *et al.* 2003; Neubauer *et al.* 2009; Hayes & Sauer 2003). RelE can only cleave the RNA when bound to the ribosome as part of a translation complex (Pedersen *et al.* 2003). This means that the proportion of the RNA bound by the ribosomes can be assessed by looking at the percentage of the RNA cleaved by RelE under conditions in which the ribosomal A-site does not contain aa-tRNA^{aa}. Using RelE cleavage, the occupancy of the ribosomes on the RNA was compared between TR-CTT and translation only using 5’ radiolabeled RNA to analyse the

efficiency of translation initiation. The AAEC was assembled using MFVVVR radiolabelled RNA on both the original abasic template and also on the template with the same DNA sequence but without the abasic site. Translation was initiated on the RNA bound by the AAECs (Figure 5.13 A, 1) and samples taken before (initiation, Figure 5.13 A, 2) and after (elongation, Figure 5.13 A, 3) the ribosomes synthesised the MF dipeptide and translocated in the presence of EF-G to position the next codon in the A-site. EF-G is required in the elongation reaction as, although the MF dipeptide will be synthesised without EF-G, the ribosomes require EF-G to translocate and position the peptidyl tRNA in the P-site and clear the A-site for the next aa-tRNA^{aa} (Green & Noller 1997). EF-G is not required for RelE analysis of the initiation reaction as fMet-tRNA^{fMet} enters directly into the P-site and the A-site remains vacant. The absence of Val-tRNA^{Val} in the reaction leaves the A-site available for RelE to bind. NTPs (either CTP and ATP, or chase at 100 μM final) and 100 pmols of GreA was added into the TR-CTT reaction during translation. 12 pmol of RelE were added after translation elongation, the reaction incubated at 37°C for 30 min and an equal volume of transcription stop buffer was added to the samples to stop the reaction. 10% of the amount of radiolabelled RNA used in TR-CTT was used in the translation only control reaction and samples were taken as for TR-CTT. The results are shown in Figure 5.13 B. Since the RNA is labeled at the 5' end, as the ribosome translocates along the RNA towards the RNAP, the RelE cleavage product will become longer. When the amount of RelE cleavage after translation was compared between the coupled and uncoupled reactions, the cleavage product was only observed in the reaction of translation alone (Figure 5.13 B lanes 28 & 30). Even when the RNAP was chased with the addition of all NTPs on the template lacking the abasic site, there was no detectable cleavage by RelE (Figure 5.13 B lanes 24 & 26).

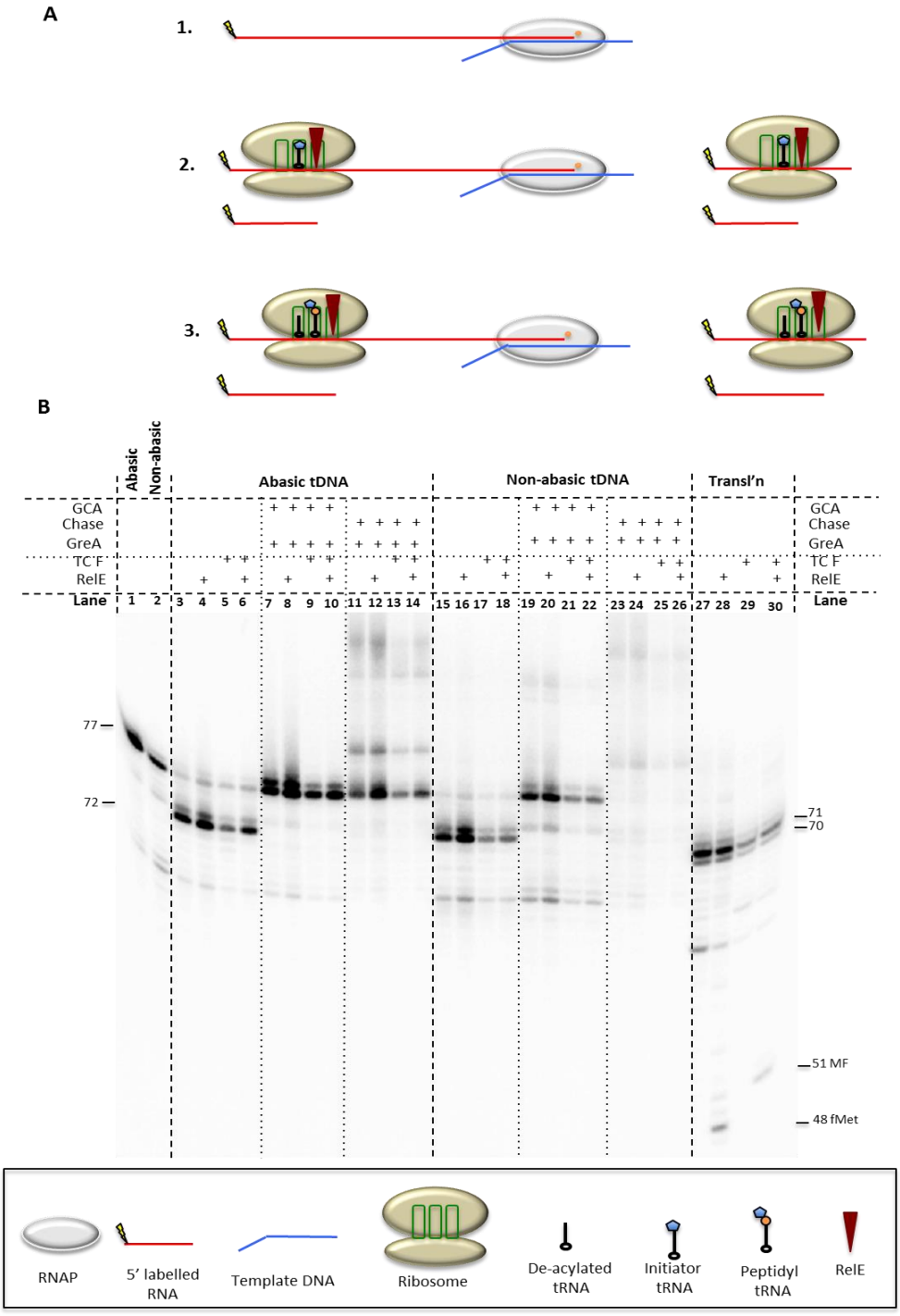


Figure 5.13 Ribosome occupancy after TR-CTT analysed by RelE cleavage. **A**) Schematic of the samples. 1. Before translation initiation. 2. Translation initiation and RelE cleavage product. 3. RelE cleavage after MF dipeptide synthesis. Although not shown in the schematic, non-template DNA was included in all reactions. **B**) PAGE of RelE cleavage after TR-CTT using MFVVVR RNA and either the abasic DNA template (lanes 1, 3-14) or the non-abasic DNA template (lanes 2, 15-26) with RelE cleavage of MFVVVR in translation alone for comparison (lanes 27-30). The non-abasic DNA template contains the same sequence as the abasic template but with a dAMP nucleotide in place of the abasic site. Lanes 1 and 2 show the RNA after washing of the AAECs formed on the abasic template (lane 1) and non-abasic template (lane 2). Translation in TR-CTT was carried either without the addition of NTPs and GreA (lanes 2-6 and 15-18), the presence of GreA and the NTPs G, C and A (lanes 7-10 and 19-22) or in the presence of GreA and all NTPs (Chase, lanes 11-14 and 23-26). RelE cleavage was carried out after either translation initiation or elongation to the MF dipeptide. Transl'n = translation alone. All RNA is 5' end labelled.

With the MFVVVR original RNA, after walking of the RNAP to the abasic site and misincorporation of GTP to produce the 78 nt RNA, there was a distance of 19 nts from the 1st nt of the AUG start codon to the RNAP proposed rear end. It may be that the start codon (and SD) is not far enough away from the RNAP for the ribosomes to bind the RNA and initiate transcription. To determine if this was indeed the case, two new RNAs were designed and synthesised that positioned the RNAP rear end either 43 or 51 nts from the 1st nt of the AUG start codon. These RNAs were named MFVVVR+24 and MFVVVR+32 respectively, based on the number of nucleotides added to the original MFVVVR RNA between the peptide coding region and the RNA:DNA hybrid sequence. After determining that these RNAs worked well in transcription and translation (data not shown), the occupancy of the ribosomes after TR-CTT and translation alone was assessed as with the MFVVVR RNA, with a few modifications. The AAECs were assembled using the abasic DNA template and no Gre factors or NTPs were added during the translation stage. RelE cleavage was also assessed after the synthesis of MF, MFVVV and MFVVVR peptides and the results are shown in Figure 5.14. In TR-CTT, after synthesis of the MF dipeptide, there was no visible RelE cleavage on either RNA (Figure 5.14). This is in direct contrast to translation only where the majority of the RNA was cleaved (Figure 5.14 lanes 10 & 23). After synthesis of the MFVVV peptide during translation only, there was very little cleavage of the RNA (Figure 5.14 lanes 12 & 25) and none after synthesis of the MFVVVR peptide (Figure 5.14 lanes 14 & 27). The disappearance of the cleavage product after synthesis of the MF dipeptide however suggested that the ribosome have synthesised the MFVVV and MFVVVR peptides in these reactions. The difference in RNA cleavage efficiency by RelE is due to the particular codon located in the ribosomal A-site after peptide synthesis and translocation. Although RelE cleaves the RNA in the A-site at all codons, the efficiency with which it cleaves is very codon specific (Pedersen *et al.* 2003). RelE has a particularly low cleavage efficiency for the codons located in the A-site after synthesis of these two peptides (the arginine codon for MFVVV and the isoleucine codon for MFVVVR). The clear difference in RelE cleavage after MF synthesis between TR-CTT and translation only would suggest that, in the presence of the RNAP, the ribosomes were no longer able to initiate translation, irrespective of the distance of the SD and initiation codon from the RNAP rear end.

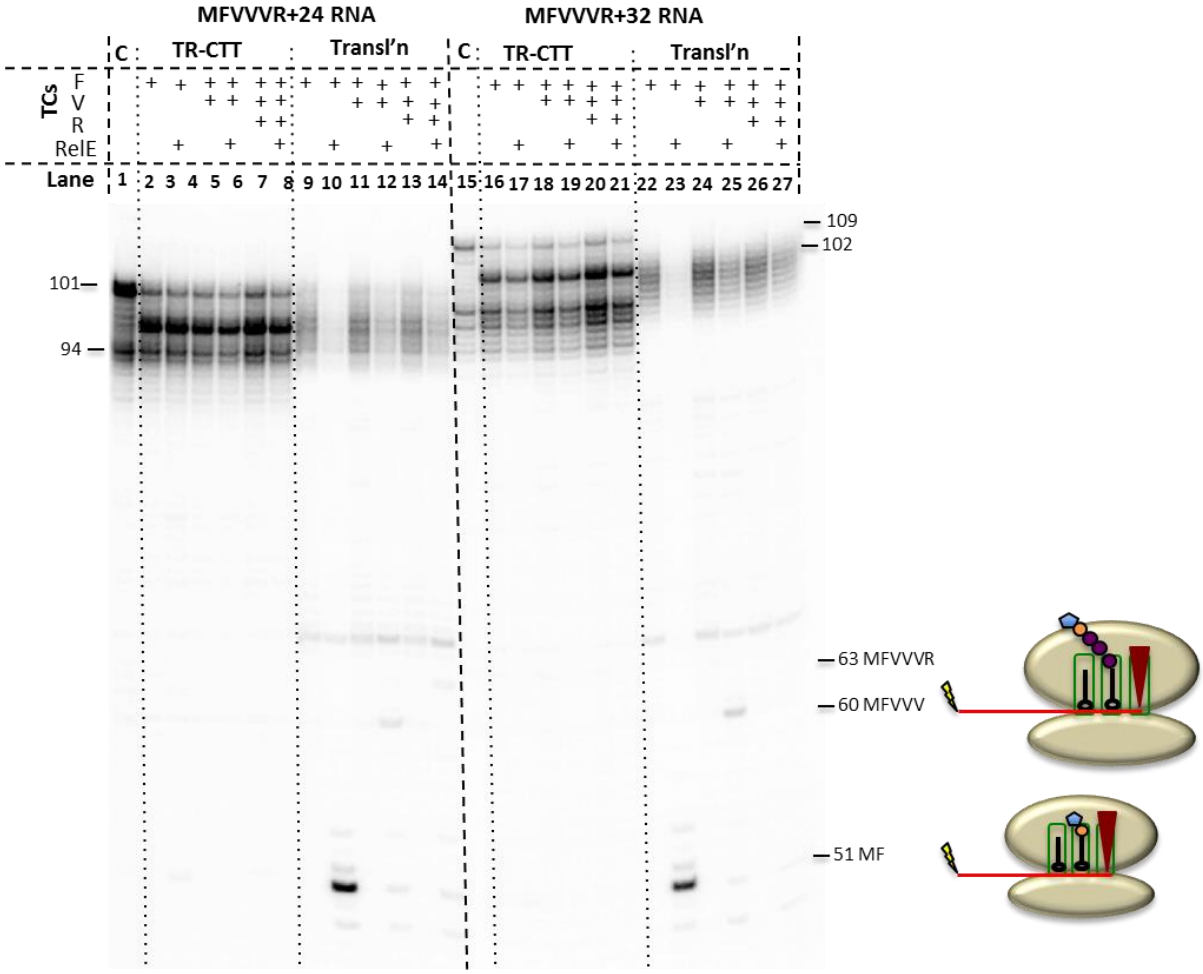


Figure 5.14 RelE analysis of ribosomal occupancy after TR-CTT on long RNA. RelE cleavage after TR-CTT or translation alone using the longer RNAs MFVVVR+24 (lanes 1-14) and MFVVVR+32 (lanes 15-27). Samples were taken after washing of the AAECs (lanes 1&15) and translation elongation to synthesise the MF, MFVVV and MFVVVR peptides. No Gre factors or NTPs were added during the translation stage hence the appearance of the shorter RNA species during translation (lane 1 compared to 2-8 and lane 15 compared to 16-21). The multiple bands seen during translation alone are due to the mixed RNA population produced during RNA synthesis.

5.7 Conclusion and Discussion

Analysis of the purified RNA during TR-CTT after radiolabeling at the 5' end to visualise the entire RNA population revealed that further modification of the TR-CTT method was required. T7 RNAP is prone to adding NTP(s) to the 3' end of the RNA during synthesis (Triana-Alonso *et al.* 1995; Cazenave & Uhlenbeck 1994; Arnaud-Barbe *et al.* 1998; Nacheva & Berzal-Herranz 2003; Lapham & Crothers 1996). In TR-CTT, these extra nucleotides caused the RNAP to backtrack during AAEC assembly and prevented the majority of the RNA from being extended during transcription elongation. Addition of GreA during AAEC assembly resolved these backtracked AAECs and led to the majority of the RNA being elongated by the RNAP in the presence of NTPs.

To position the RNAP at a specific location on the DNA template and to prevent the RNAP transcribing beyond this point, an abasic site was introduced into the DNA template. This site reduced the read-through during transcription but during translation was used by the RNAP as a template to misincorporate the GTP introduced for use by the GTPase translation factors. This misincorporation, combined with the specific tDNA sequence that allowed formation of an extended complementary region upstream of the RNA:DNA hybrid, caused the RNAP to backtrack to the 3' end of the DNA template. The high concentration of GTP also led to cleavage of the RNA by intrinsic hydrolysis after backtracking. Two template sequences were identified on which the RNAP did not misincorporate GTP and backtrack, but these templates reduced the backtracking of the RNAP during AAEC. GreA was no longer as effective during AAEC assembly using these templates and the amount of RNA actively transcribed was diminished compared to with the original DNA template. As it was initially thought that the lack of peptide after TR-CTT was due to the reduced level of the correct length of RNA, these templates were deemed to be unsuitable for use in TR-CTT. Instead, because the addition of NTPs and GreA during translation caused the AAEC to go through a cycle of forward transcription and backtracking, with the RNAP predominantly located in the forward position, the original DNA was judged to be the most suitable for use in TR-CTT.

Despite the optimisation of the TR-CTT technique, the amount of peptide synthesised during the translation stage remained consistently low compared to that produced during translation alone using a comparative amount of RNA. RelE cleavage of 5' end labeled RNA after TR-CTT to determine the occupancy of the ribosomes after translation initiation revealed that the ribosomes were unable to initiate translation in the presence of RNAP, irrespective of the distance between the RNAP and the SD. The reason for this is unclear, but could be due to the RNA wrapping around the RNAP or interacting with the RNAP in an unknown and unexpected manner, either during or after transcription, that results in blocking access of the ribosomes to the SD and/or the initiation codon. Regardless of the reason, the inability of the ribosomes to bind to the RNA in TR-CTT presents a significant obstacle. In light of everything discovered during this chapter and the previous chapter, particularly with respect to these latest results, we concluded that the approach needed to be changed, as TR-CTT did not seem reliable enough for

the purpose of determining the distance between the RNAP and the ribosome that prevents the ribosome incorporating the next amino acid into the peptide chain.

6. TL-CTT

6.1 Introduction

TR-CTT was originally thought to be the most suitable technique for this project, however, it was discovered during extensive analysis (outlined in chapters 4 and 5) that this was not the case and a different approach was required. Translation first coupled transcription-translation (TL-CTT) is an alternative transcription-coupled-to-translation technique and is based on the same core reactions as TR-CTT (Castro-Roa & Zenkin 2012; Castro-Roa & Zenkin 2015).

As previously mentioned (see section 1.3.2, Figure 1.14), in the TL-CTT method, translation is initiated on the RNA and the whole initiation reaction used to synthesise the stable MF dipeptide. The reaction is then layered onto a 1.2 M sucrose cushion and centrifuged at high speed. Only the ribosomes and the associated RNA and peptides are able to migrate through the sucrose cushion and form a pellet. All free RNA, GTP and aa-tRNA^{aa} remain on top of, or trapped within, the sucrose cushion. This ensures that only the RNA that is bound by a ribosome will be pelleted. The pellet is then resuspended and the AAEC assembled on the ribosome bound RNA. The RNA is either labeled at the 3' end during transcription using α-[^{32}P]-NTPs or 5' end labeled RNA is used in the initial translation reaction. After the AAEC is assembled, the RNAP and/or the ribosomes can be translocated by addition of NTPs and/or aa-tRNA^{aa} respectively. As the system is similar to TR-CTT, with the only main difference being that translation is initiated first before the AAEC is assembled, it is possible that this technique could be modified to replace TR-CTT for use in this project. As the TL-CTT technique was originally designed for studying the effect of translation and the ribosomes on transcription and the RNAP, the technique was designed so that all AAECs are coupled to translation, however, not all translation complexes are coupled to an AAEC. As the effect of the RNAP on the ribosome was central to this project, the TL-CTT technique needed to be modified to ensure all translation complexes were coupled to AAECs.

6.1.2 Aims

The aim of this chapter was to modify TL-CTT for the purpose of determining how close to a halted RNAP the ribosomes could translocate whilst actively translating a small peptide product.

6.2 Testing and Modifying TL-CTT

To determine if TL-CTT had the potential for use in this project, TL-CTT was carried out using 5' end labeled MFVVVR+5 RNA (the longest of the original RNAs) and the original abasic template, as this template was previously deemed the most suitable for TR-CTT based on the analysis outlined in chapter 5, section 5.4. Translation was initiated using fMet-tRNA^{fMet}, the MF dipeptide synthesised (Figure 6.1 A, step 1) and the reaction filtered through a sucrose cushion by ultracentrifugation (step 2). After the ribosomal pellet was resuspended (Figure 6.1 A, step 3), the AAEC was assembled using the original abasic template (step 4) in the presence of GreA (step 4), as for the transcription stage in TR-CTT. Finally, ntDNA and NTPs were added to form a stable AAEC and elongate the RNA. Samples were taken at each stage of the transcription reaction and either stop buffer added or RelE (the ribosome associated RNA cleavage enzyme) then stop buffer added to ensure that the ribosomes remained bound to the RNA throughout transcription. Transcription alone using free RNA was assembled alongside as a control.

In TL-CTT compared to transcription alone, the RNAP was not able to elongate the RNA, as shown by the lack of the 82 nt product in TL-CTT Figure 6.1 B lane 7) compared to transcription alone (lane 22). These results also showed a lack of RNA cleavage to produce the 2/3 nt shorter RNA species in the presence of GreA during AAEC assembly after translation (lane 5) compared to with transcription alone (lane 10), suggesting that the RNAP was not able to backtrack after AAEC assembly. As this initial backtracking of RNAP and resolution in the presence of GreA is essential for the RNAP to be able to elongate all of the different RNA species (see section 5.2.2, Figure 5.3), this would explain why not all of the RNA was elongated to the 82 nt product during transcription elongation in TL-CTT. After initiation of translation, synthesis of the MF dipeptide and translocation of the ribosome by one codon on the MFVVVR+5 RNA, the ribosome was potentially covering a 16 nt region of RNA from the 1st nt of the first valine codon in the A-site to the upstream edge of the hybrid site, based on hydroxyl radical footprinting (Hüttenhofer & Noller 1994). The position of the ribosome may therefore have prevented the RNAP from binding to the RNA:DNA hybrid, or even have inhibited formation of the RNA:DNA hybrid in the first place. If either of these was indeed the case, use of the longer RNA, MFVVVR+32, in TL-CTT should allow the RNAP to elongate the RNA, as the distance between the 1st nt of the AUG codon and the 3' end of the RNA would be 53

nts, exceeding the 31 nts of RNA proposed to be covered by the RNAP and the ribosomes together (Hüttenhofer & Noller 1994; Komissarova & Kashlev 1998, see also Figure 1.11). TL-CTT was repeated using 5’ end labeled MFVVVR+32 RNA as above (Figure 6.1 A) and the results shown in Figure 6.1 C. From these results it was clear that the RNAP was able to elongate the RNA after AAEC with both ribosome bound and free RNA (Figure 6.1 C, compare lane 9 and 19). As the ribosomes were positioned further away from the hybrid site, the RNAP was able to bind the RNA:DNA hybrid, backtrack, cleave the RNA in the presence of GreA (lane 7 and 18) and elongate the RNA to produce the full length 109 nt product in the presence of NTPs. RelE cleavage of the RNA during AAEC assembly and transcription elongation also showed that the ribosomes remained bound to the RNA. These results revealed that TL-CTT was potentially a viable method, as the ribosomes and the RNAP were both able to bind to the same RNA oligo, providing the RNA:DNA hybrid was a suitable distance from the translation initiation codon.

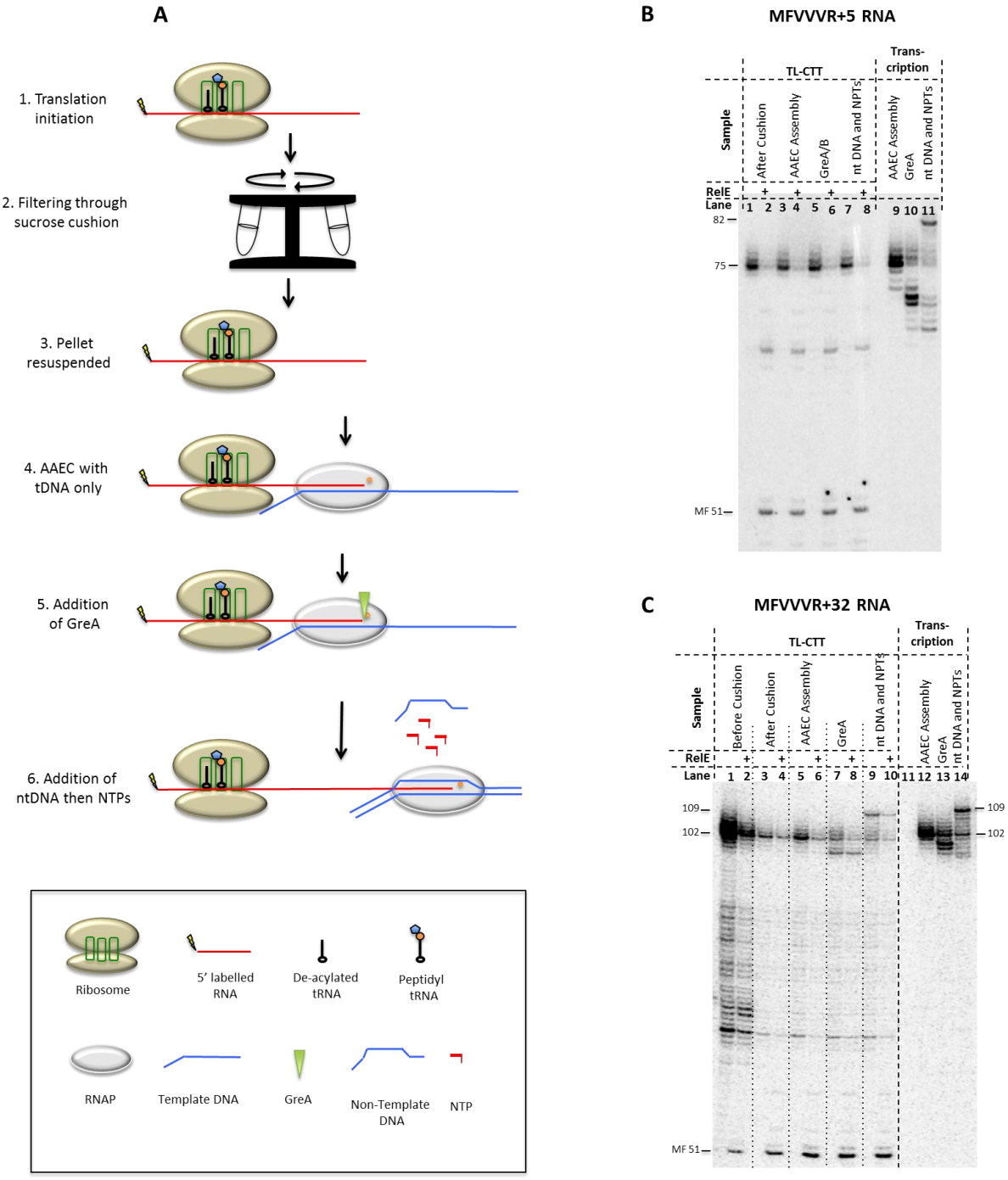


Figure 6.1 RelE analysis of TL-CTT using 5' end labelled RNA. A) Schematic diagram of the TL-CTT reaction. B) 10% PAGE of the products of TL-CTT with MFVVVR+5 RNA (lanes 1-8) compared to transcription alone (9-11). Samples were taken as indicated and the reaction stopped (odd numbered lanes) or RelE added, the reaction incubated further then stopped (even numbered lanes). C) 10% PAGE of the products of TL-CTT with MFVVVR+32 RNA (lanes 1-10) and transcription alone (lanes 12-14). Samples were taken as indicated and the reaction stopped (odd numbered lanes) or RelE added, the reaction incubated further then stopped (even numbered lanes).

6.2.1 RelE Cleavage of 3' Radiolabeled RNA

In TR-CTT, the ability of the ribosomes to incorporate the arginine into the peptide was determined by visualising and identifying the peptide synthesised during translation. As only RNA bound by the RNAP was present during translation due to

washing of the AAECs, all the RNA used as a template for peptide synthesis was associated with an RNAP enzyme. In TL-CTT, however, this is not the case. The AAEC is assembled after translation of the MF dipeptide and, even though the RNAP is added in molar excess over the RNA, it cannot be guaranteed that the RNAP will bind all of the RNA. Visualisation of the full-length peptide after AAEC would reveal all the peptide synthesised by ribosomes translating RNA, either with or without an associated AAEC. This would affect the end result, as it is not possible to distinguish between peptides produced from the coupled versus uncoupled RNA.

RelE footprinting is an alternative method to assess the location and occupancy of the ribosome on the RNA. Although RelE cleavage is not normally used in either TR-CTT or TL-CTT, it was used previously to analyse translation initiation during TR-CTT (see section 5.6.1). In order for RelE cleavage of the RNA to be analysed, however, the RNA needs to be radiolabeled. During the initial analysis of TL-CTT, the RNA was visualised by 5' end labeling prior to TL-CTT but use of 5' labeled RNA in TL-CTT has limitations with regards to determining the coupling of the RNAP and the ribosomes. The proportion of RNA species extended by the RNAP during transcription indicates the proportion of the RNA containing an actively transcribing AAEC, while RelE cleavage of the RNA shows the proportion of RNA bound by an active ribosome. It doesn't, however, reveal the proportion of RNA actively transcribed and also bound by an active translation elongation complex. An alternative method to label the RNA is by incorporation of radiolabelled NTPs during transcription elongation, as was used during the initial analysis of the RNA for TR-CTT (see chapter 4, section 4.7). This method only radiolabels the elongated RNA during the transcription stage of TL-CTT after the translation elongation complexes are filtered through the sucrose cushion, therefore all the visible RNA will be bound by both an AAEC and a translation elongation complex. RelE footprinting after transcription (and further translation elongation if applicable) would then allow the location of the ribosome to be determined.

To test this method, TL-CTT was assembled using unlabeled MFVVVR+24 RNA (Figure 6.2 A step 1), the dipeptide reaction passed through a sucrose cushion (step 2) and the ribosome pellet resuspended (step 3). The AAEC was formed using the original abasic tDNA in the presence of GreA (step 4). α-[^{32}P]-CTP was used to radiolabel the RNA during transcription and was added along with cold GTP and ATP after stabilisation of the AAEC with ntDNA (step 5). After transcription elongation, the

translation reaction was either kept as the MF dipeptide, Val-tRNA^{Val} and EF-G added to produce the MFVVV peptide or Val-tRNA^{Val}, Arg-tRNA^{Arg} and EF-G added to synthesise the MFVVVR peptide (step 6). RelE was then added to each reaction, before the reaction was stopped by the addition of transcription stop buffer (step 7). As the RNA was labeled at the 3' end and the ribosome was translocating towards the RNAP, the size of the visible RNA fragment after RelE cleavage became progressively shorter as a longer peptide was synthesised (Figure 6.2 B). RelE cleavage showed that the ribosomes were able to translocate towards the RNAP and synthesise the peptide in the presence of Val or Val and Arg TCs after transcription elongation (Figure 6.2 B, compare lane 2 with RelE cleavage after MF, to lanes 4 and 6, showing RelE cleavage after MFVVV and MFVVVR). The difference in RelE cleavage efficiency after synthesis of the different peptides was due to the codon bias of RelE (Pedersen *et al.* 2003), however, the disappearance of the 51 nt RelE cleavage product after MF dipeptide synthesis in the samples with arg and/or val TCs indicated that the ribosome was able to synthesise the longer peptides. Together, these results demonstrated that transcription and translation coupling can be achieved by initiating translation, synthesising the dipeptide, radiolabeling the actively transcribing AAECs during transcription then further translating the RNA and finally cleaving the RNA with RelE. The TL-CTT method modified with RelE cleavage to analyse translation, therefore had the potential to be a suitable technique for use in determining the distances between the RNAP and the ribosome. As visualisation of the peptide was no longer required, unlabeled methionine was used from now on in place of ^{35}S-methionine during translation initiation in TL-CTT.

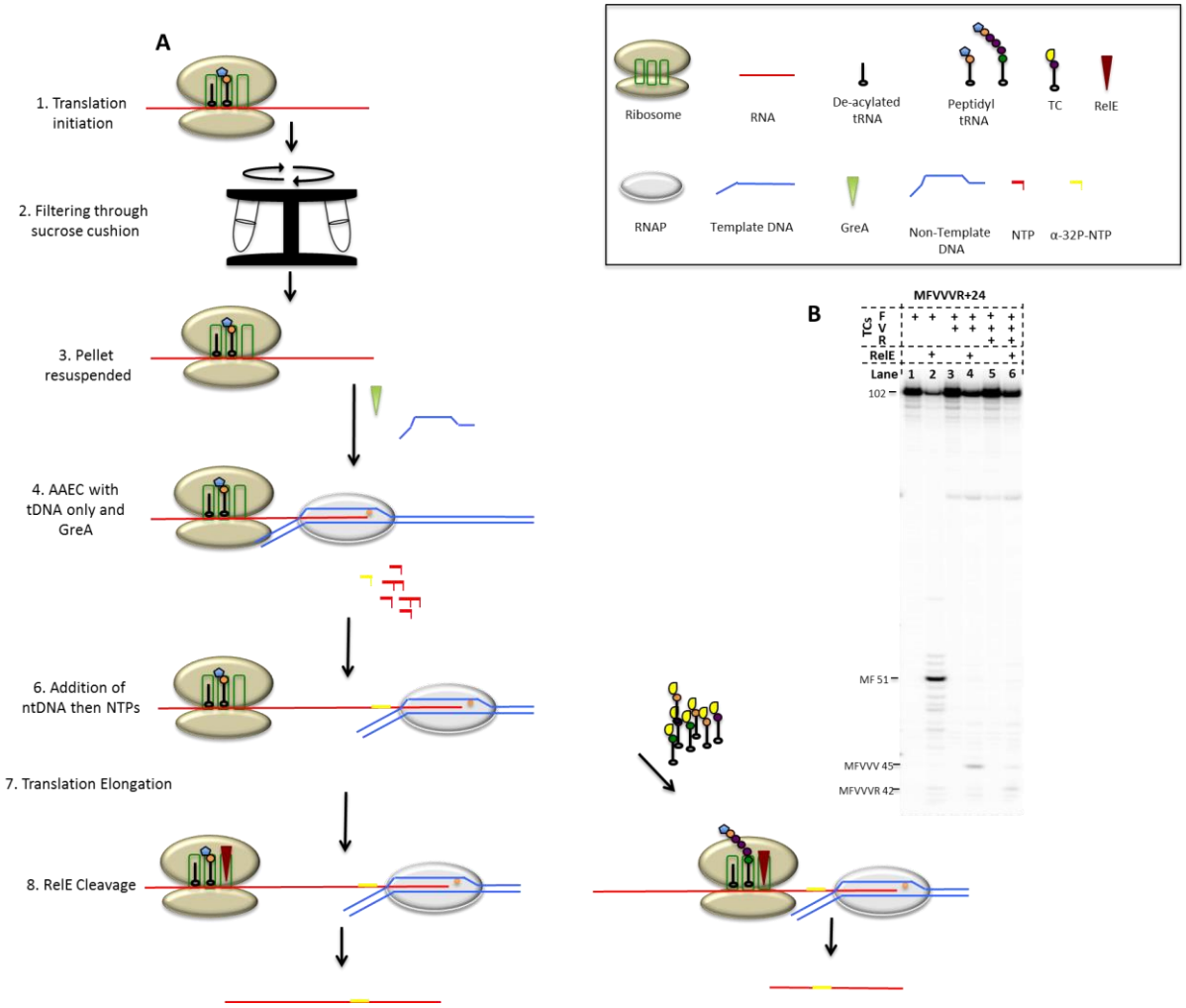


Figure 6.2 TL-CTT using MFVVVR+24 RNA. A) Diagram of TL-CTT and RelE cleavage after translation elongation. B). PAGE of samples taken after TL-CTT and the reaction stopped (odd numbered lanes) or RelE added (even numbered lanes). Lane 2 shows RelE cleavage of the RNA at the valine codon after MF dipeptide synthesis, lane 4 at the arginine codon after MFVVV peptide synthesis and lane 6 at the isoleucine codon after MFVVVR. The RNA was radiolabeled at the 3' end by incorporation of α-[32P]-CTP.

The codon bias of RelE did not allow comparison after MF and MFVVVR peptide synthesis as there was a marked difference in cleavage efficiency at the specific codons located in the ribosomal A-site. In fact, RelE displays a marked bias towards the three translation stop codons, particularly that of the second most abundant stop codon, UAG (Pedersen *et al.* 2003). The efficiency of cleavage at this codon is up to 1000s of times that of valine, arginine and isoleucine (Pedersen *et al.* 2003). To eliminate the variation in RelE cleavage due to codon bias, we decided to integrate the UAG stop codon within the RNA, then analyse the ability of the ribosome to place the UAG stop codon in the A-site after synthesising the peptide by comparing cleavage efficiency between different positions of the stop codon with respect to RNAP. To test whether this approach was viable, the UAG stop codon was

integrated into the MFVVVR+24 RNA after the Met and Phe codons (in place of the Val codon, named MF.+24 RNA) or after the Met, Phe and three Val codons (in place of the Arg codon, named MFVVV.+24 RNA) to compare RelE cleavage before (MF) and after (MFVVV) translation elongation and translocation of the ribosome towards the RNAP. The RNAs containing the stop codon were tested in TL-CTT with radiolabeling of the RNA at the 3' end during transcription elongation (see schematic in Figure 6.3 A). After transcription elongation, either no TCs or EF-G, or Arg-tRNA^{Arg} and/or Val-tRNA^{Val} TCs, along with EF-G, were added in separate reactions and the cleavage of the RNA by RelE analysed as previously. The results are shown in Figure 6.3 B. RelE cleavage of the MF.+24 RNA revealed that the UAG codon was located in the ribosomal A-site, regardless of the TCs present in the reaction (Figure 6.3 B lanes 2, 4 & 6). After TL-CTT with the MFVVV.+24 RNA, cleavage was seen at the first valine codon when no TCs were present (after MF dipeptide synthesis) and at the UAG stop codon in the presence of either Val TCs only, or both Val and Arg TCs. This showed that the ribosome was able to position the UAG stop codon in the ribosomal A-site in the presence of the correct TCs. These results confirmed that analysing the location of the ribosome in the presence of RNAP after TL-CTT using RelE cleavage of the UAG stop codon was a suitable technique for the aim of this project. The results also revealed that the ribosome was able to translocate along the MFVVV.+24 RNA towards the RNAP and position the UAG stop codon (originally the Arg codon) in the A-site. This suggested that a distance of 29 nts between the RNAP rear end and the 1st nt of the codon in the A-site was sufficient for the ribosome to position the UAG stop codon in the A-site for RelE cleavage. It must be taken into account when calculating the distances between the RNAP and the ribosome that after transcription elongation, the RNA was one nt longer than was originally expected and the RNAP located one nt further from the peptide coding region. This was due to the misincorporation of GTP opposite the abasic site by the RNAP during the translation stage of TL-CTT (or TR-CTT, outlined in section 5.3).

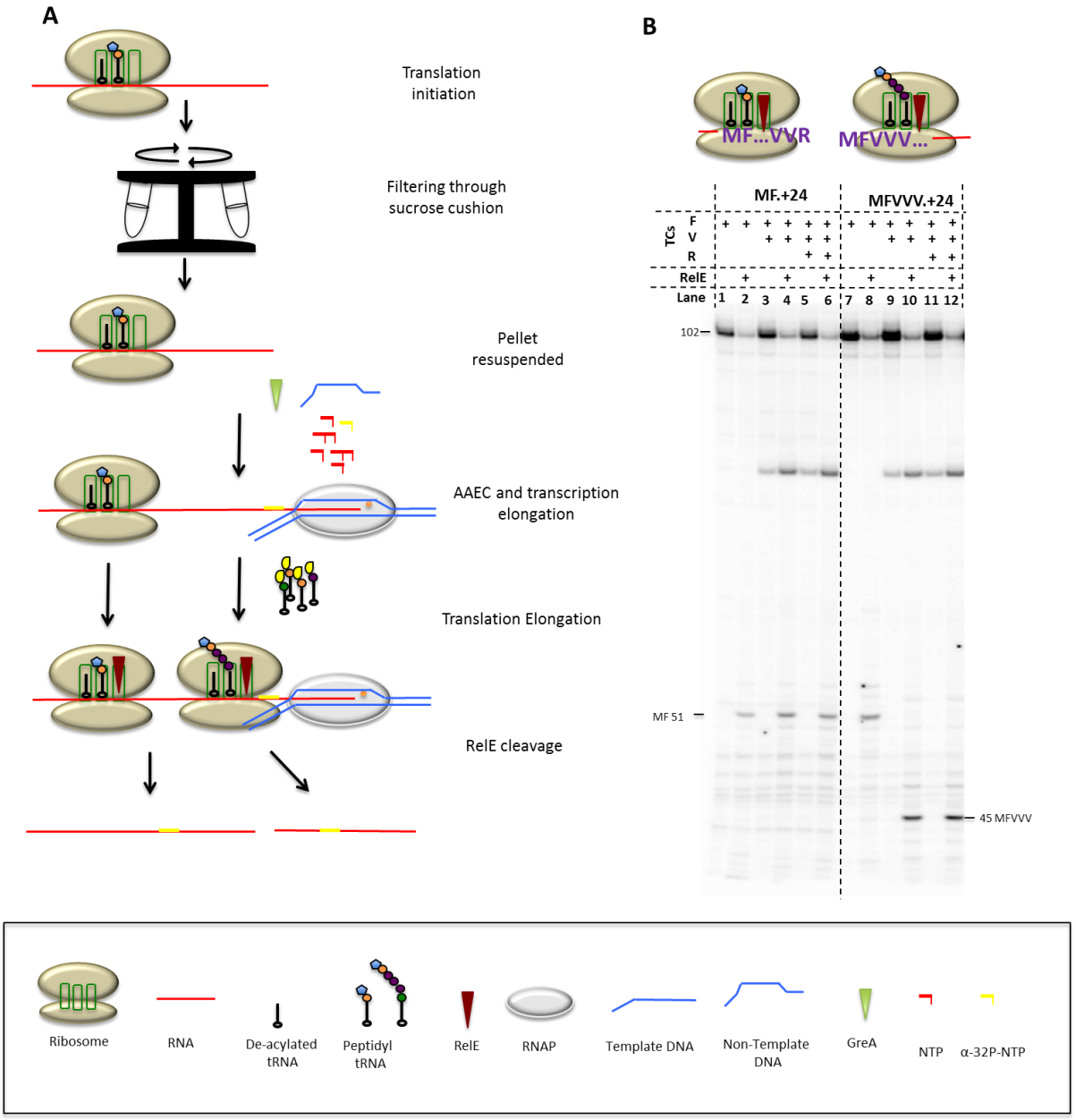


Figure 6.3 Translation of the MFVVVR+24 stop codon containing RNA. A) Schematic representation of TL-CTT and RelE cleavage. B) PAGE of translation and RelE cleavage at the UAG stop codon located after either MF (lanes 1-6) or MFVVV (lanes 7-12). Even in the presence of all TCs, there was no read-through of the stop codon on either RNA. The RNA was radiolabeled at the 3' by incorporation of α-[32P]-CTP.

6.2.2 Decreasing the Distance by Increasing the Peptide

Once it had been determined that the UAG stop codon in the RNA allowed comparison of RelE cleavage efficiency at different locations on the RNA, the next step was to decrease the distance between the RNAP rear end and the stop codon. Reducing the length of the RNA was not an option, as too short an RNA does not allow transcription elongation after AAEC (section 6.1, Figure 6.1 B). Instead, we decided to maintain the same length of RNA, but increase the length of the peptide synthesised by the ribosome by extending the peptide coding region. Coincidentally,

the MFVVVR+24 RNA contained the strongly cleaved UAG stop codon after a nine amino acid peptide sequence (MFVVVRIDG). Although pure tRNAs for the amino acids Met, Phe, Val and Arg of this sequence are commercially available, pure, uncharged tRNAs for Isoleucine (I, Ile, Aspartic Acid (Asp, D) and Glycine (Gly, G) are not. To date, only amino acids with commercially available uncharged tRNAs have been used in either TR-CTT or TL-CTT. In order to extend the peptide sequence using amino acids without commercially available uncharged tRNAs, a method to produce these tRNAs needed to identified.

6.2.3 tRNA Purification

To obtain the pure, uncharged tRNAs not commercially available, we decided to purify the specific tRNA required from a mix of total tRNA isolated from the *E. coli* strain, MRE600 (see materials and methods, section 3.3.1) The uncharged tRNAs were purified from the total tRNA using DNA probes complementary to 25-30 nts of the D-loop and anti-codon loop of the target tRNA and containing a biotin tag at the 3' end (Yokogawa *et al.* 2010). The DNA probes were immobilised on streptavidin beads (Figure 6.4 A, step 1) and the total tRNA mix added and heated to 65°C for 10 min to anneal the tRNA to the DNA probe (step 2). The sample was centrifuged briefly (step 3) and the supernatant removed to remove all free, non-specific tRNA (step 4). The remaining pellet containing the DNA probe hybridized with the specific tRNA and bound to the streptavidin beads (step 5) was washed multiple times to remove any non-specific tRNA. The hybridized tRNA was eluted by heating the sample to 65°C for 5 min to release the tRNA into the supernatant. The samples were centrifuged (step 6) to separate the supernatant from the pellet and the supernatant containing the tRNA removed (step 7). The amount and purity of the tRNA was determined by measuring the absorbance at A_{260} and A_{280} and analysing the tRNA samples by separation on a 5% short denaturing polyacrylamide gel run at 14W for 1 hr. The nucleic acid was visualised by staining with SyBr gold (Thermo Fisher) (Figure 6.4 B).

The purified Ile-, Asp- and Gly-tRNAs were aminoacylated as standard and tested in translation alone using 5' end labeled MFVVVR+24 RNA and RelE cleavage. The results revealed that the purified and aminoacylated tRNA did not allow the ribosomes to translate the full peptide, as judged by RelE cleavage of the

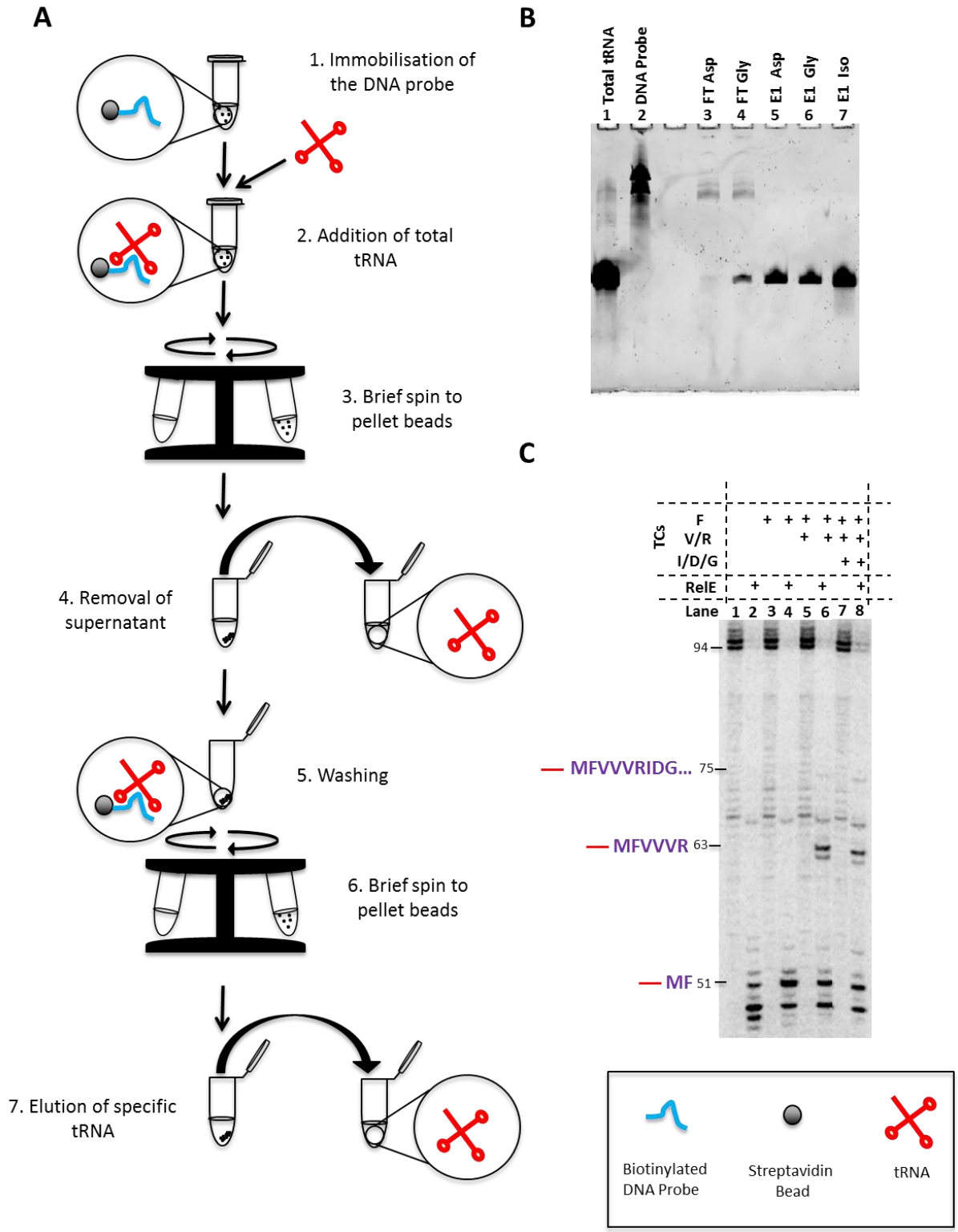


Figure 6.4 Purification of specific tRNA. A) Schematic of tRNA purification from total tRNA using biotin-tagged DNA probes. B) PAGE analysis of tRNA purification visualised by SyBr gold. FT= flowthrough E1=elution 1. C) PAGE of translation alone to test the purified tRNA before (odd numbered lanes) and after RelE cleavage (even numbered lanes). RelE cleavage was carried out after translation initiation (M, lane 2), after translation elongation of the MF dipeptide (lane 4), the MFVVVR peptide (lane 6) and and the MFVVVRIDG peptide (lane 8). In the presence of all TCs, the main product seen after RelE cleavage was at the Iso codon after MFVVVR peptide synthesis, not the UAG codon after MFVVVRIDG synthesis.

UAG codon after the glycine codon (Figure 6.4 C, lane 8). The ribosomes appeared to stall with an empty A-site after MFVVVR synthesis allowing RelE access, suggesting that the aa-tRNA^{aa} concentration was too low. This would suggest that either the tRNA was not as pure as expected or was not aminoacylated efficiently. Either way, an alternative method of producing the required aa-tRNA^{aa} was needed.

6.2.4 Total tRNA aminoacylation

In order to produce aa-tRNA^{aa} for the additional amino acids (I, D and G), we decided to try an alternative method not previously used in TR-CTT or TL-CTT. Instead of using pure, specific tRNAs for each amino acid in separate aminoacylation reactions, uncharged total tRNA from the *E. coli* strain MRE600 was used and the specific amino acids (Ile, Asp and Gly) added into one reaction. The aminoacylation reaction was modified to allow for the use of uncharged total tRNA and multiple aminoacylation reactions in one by using 9 units of uncharged total tRNA and increasing the amount of S100 extract to 100 µL but keeping the final concentration of each amino acid the same.

The aa-tRNA^{aa} mix was tested in translation using 5' end labelled MFVVVR+24 RNA with RelE cleavage of the RNA. Translation was carried out as standard and samples taken after translation elongation with EF-G and either Phe-tRNA^{Phe} only, Phe-tRNA^{Phe}, Val-tRNA^{Val} and Arg-tRNA^{Arg} TCs or with all TCs including the mixed I, D, G TCs to synthesise the 9 amino acid peptide (Figure 6.5 A). RelE cleavage after translation elongation showed that the ribosome was able to fully translocate along the RNA and position the UAG stop codon in the A-site after the MFVVVRIDG peptide (Figure 6.5 A, lane 6). This showed that using total uncharged tRNA but only amino-acylating specific tRNAs was a viable method for producing charged tRNAs for amino acids whose purified uncharged tRNAs were not commercially available.

Translation of the longer peptide (MFVVVRIDG) positions the 1st nt of the ribosomal A-site 17 nts from the proposed rear end of the RNAP when coupled to transcription. To decrease the distance between the ribosomal A-site and the RNAP rear end even further, the stop codon was moved progressively closer to the RNA 3' end one codon at a time. This produced RNAs encoding 10 (MFVVVRIDGT, MFVVVR+24_10aa), 11 (MFVVVRIDGTN, MFVVVR+24_11aa), 12 (MFVVVRIDGTNG, MFVVVR+24_12aa) and 13 (MFVVVRIDGTNGL, MFVVVR+24_13aa) amino acid peptides followed by the UAG stop codon. These

RNAs positioned the 1st nt of the UAG stop codon 14, 11, 8 and 5 nts away from the RNAP rear end respectively after transcription and translation elongation during TL-CTT (see summary in Table 6.1).

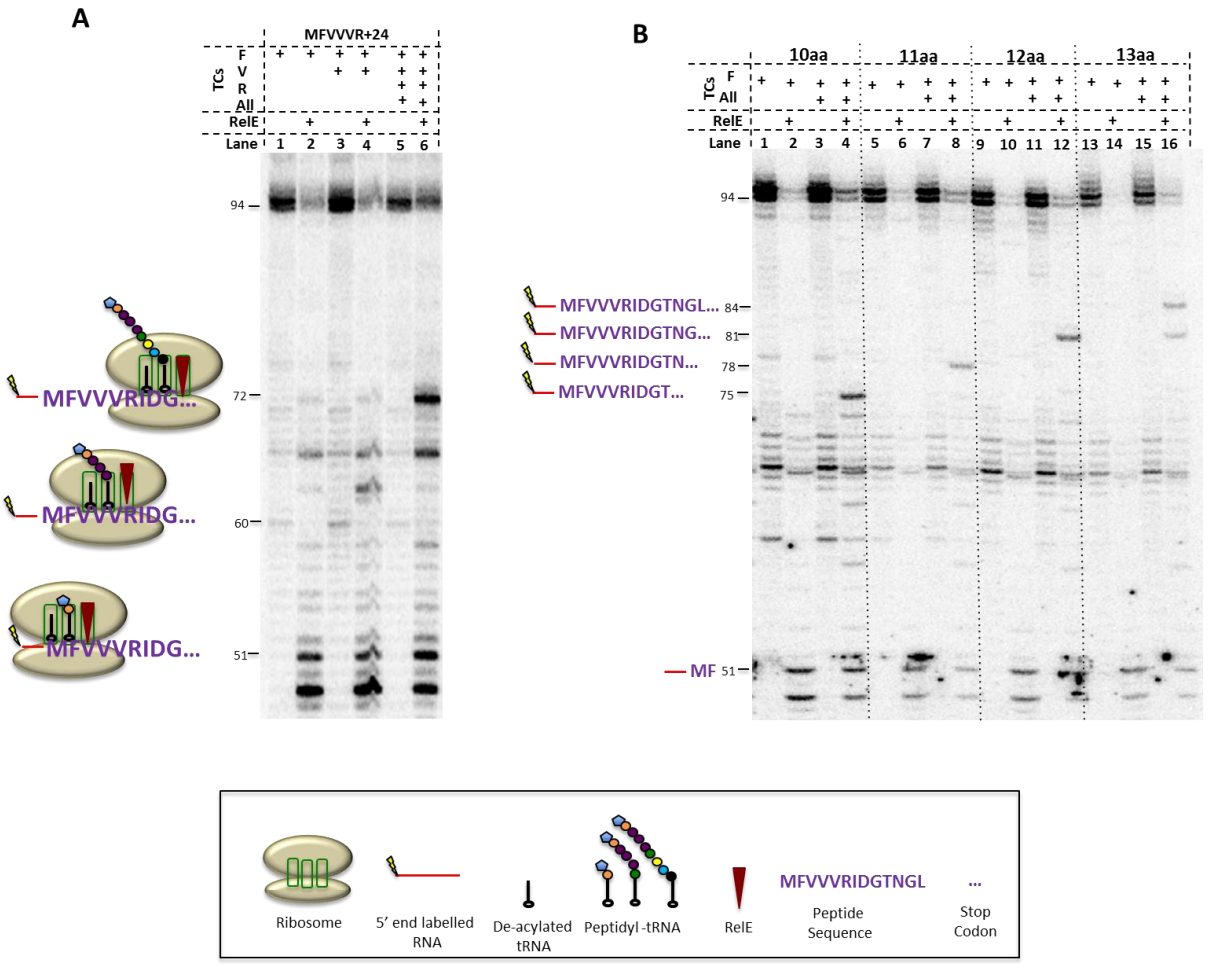


Figure 6.5 Translation analysis by RelE cleavage of longer peptides. A) Translation of the MFVVVRIDG peptide analysed by RelE cleavage. The aa-tRNA^{aa} for the IDG amino acids were produced from aminoacylation of total tRNA. The RelE cleavage of the UAG stop codon after MFVVVRIDG peptide synthesis (lane 6) indicates the aminoacylation reaction was successful. B) Translation of the longer peptides (10aa to 13aa) analysed by RelE cleavage of the UAG stop codon. As the peptide length is increased and the stop codon moved towards the 3' end, the cleavage product becomes longer. The RNA was labelled at the 5' prior to translation.

New aminoacyl-tRNAs for all the extra amino acids required for the 13 amino acid peptide (Ile, Asp, Gly, Thr, Asn and Leu) were made by amino-acylating all 6 aminoacyl-tRNAs in one reaction using total uncharged tRNA as above and this aa-tRNA^{aa} mix was used during translation. For each RNA, in translation alone using 5' end labelled RNA, the ribosomes stopped translocating with the UAG stop codon positioned fully in the A-site in the presence of all of the TCs (Figure 6.5 B). The ribosomes were able to use the TC mix to synthesise the peptides encoded within the RNA and did not read through the stop codon even when the aminoacyl-tRNAs for the following codons were present in the reaction.

RNA Name	Peptide Length (aa)	Distance to RNAP Rear End (nts)	Distance to RNAP Active Centre (nts)
MF.+24	2	38	53
MFVVV.+24	5	29	44
MFVVVR+24	9	17	32
MFVVVR+24_10aa	10	14	29
MFVVVR+24_11aa	11	11	26
MFVVVR+24_12aa	12	8	23
MFVVVR+24_13aa	13	5	20
MFVVVR+25_12aa	12	9	24
MFVVVR+26_12aa	12	10	25
MFVVVR+27_12aa	12	11	26

Table 6.1 Summary table of the RNA and distances.

6.3 TL-CTT to determine the distance

Now that aa-tRNA^{aa}s for all the amino acids in the extended peptide sequence were available, TL-CTT was carried out as standard with translation assembled using all 5 RNAs (MFVVVR+24 and MFVVVR+24_10aa to 13aa, Figure 6.6 A). After ultracentrifugation and resuspension of the ribosome pellet, the AAEC was assembled using the original abasic template DNA and GreA was added before stabilisation of the AAEC with ntDNA. The RNA was labelled at the 3' end during transcription elongation by addition of α-[^{32}P]-CTP, along with unlabeled ATP and GTP. After transcription elongation, GreB was added, along with either RelE or Val-tRNA^{Val}, Arg-tRNA^{Arg} and the aa-tRNA^{aa} mix were added, followed by RelE.

The results of TL-CTT analysed by RelE cleavage are shown in Figure 6.6B. RelE cleavage after the addition of all TCs to all the different RNAs revealed that the ribosomes were able to synthesise the 9, 10 and 11 amino acid peptides and position the stop codon in the ribosomal A-site (17, 14 and 11 nts respectively to the RNAP rear end from the 1st nt of UAG, (Figure 6.6 A lanes 5, 10 &15) but were unable position the UAG stop codon in the ribosomal A-site after the longer 12 and 13 amino acid peptides (8 and 5 nts from RNAP, lanes 20 & 25).

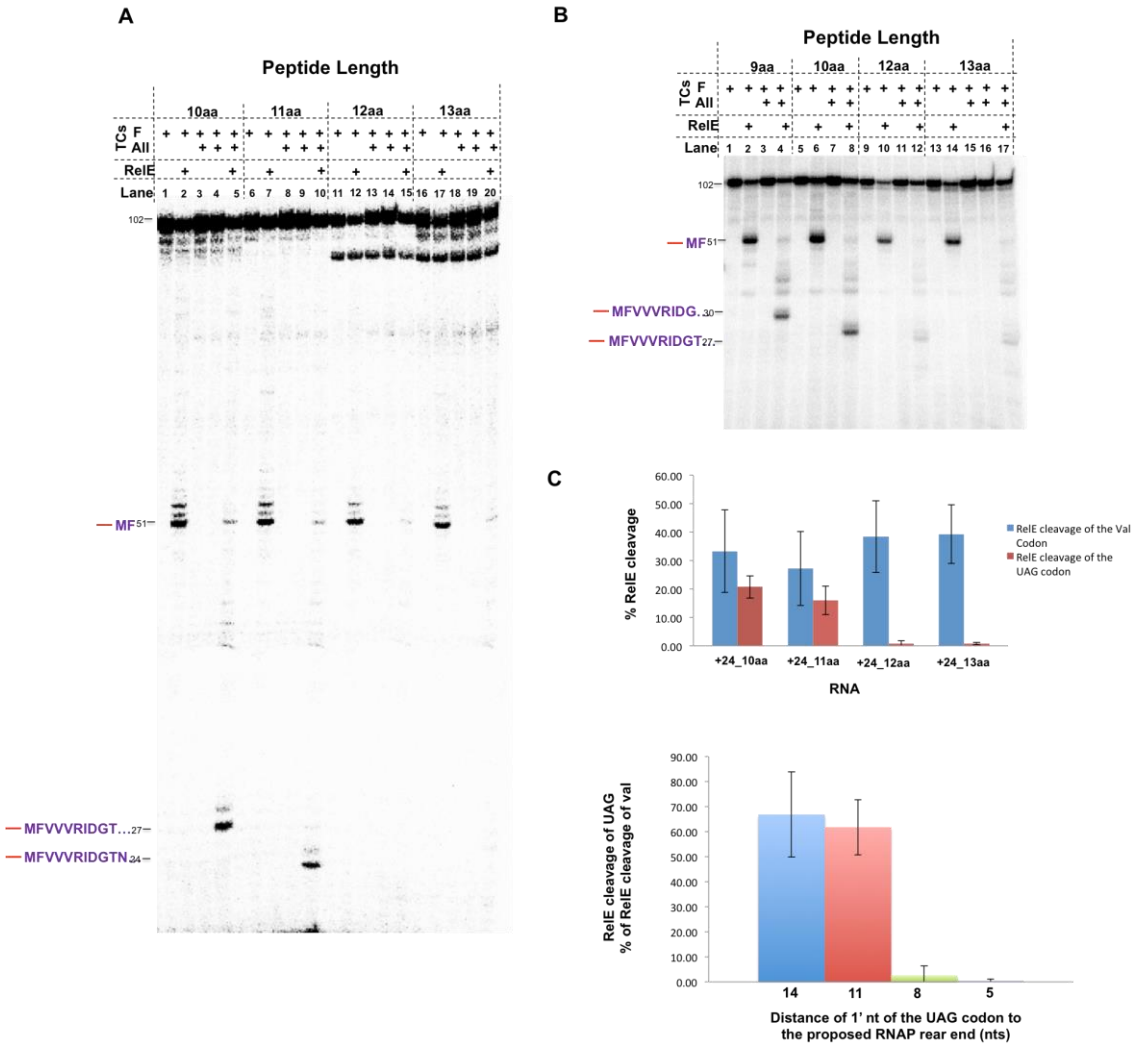


Figure 6.6 TL-CTT and RelE cleavage to determine the distances (1). **A**) PAGE of RelE cleavage at the UAG stop codon after TL-CTT using MFVVVR+24_9aa (lanes 1-5), 10aa (6-10), 11aa (lanes11-15), 12aa (lanes 16-10) and 13aa (lanes 21-25) RNAs. RelE cleavage of the UAG stop codon is seen after TL-CTT of the 9, 10 and 11 aa peptides (lanes 5, 10 and 15 respectively) but not after the 12 or 13 aa peptides. As RelE cleavage is seen after MF dipeptide synthesis of all RNAs, this would suggest the RNAP is preventing the ribosome positioning the UAG stop codon in the ribosomal A-site after peptide synthesis with the MFVVVR+24_12 and 13 RNAs. The gel image was adjusted using the sigmoidal setting for clarity. **B**) Example of RelE cleavage of the UAG stop codon after separation by 20% denaturing PAGE. **C**) Quantification of RelE cleavage efficiency. Top: the % of RelE cleavage of both the Val codon (after MF dipeptide synthesis) and the UAG stop codon (after full length protein synthesis) are shown in blue and red respectively. Bottom: the cleavage efficiency of the UAG stop codon was taken as a percentage of cleavage of the Val codon after MF synthesis. The values are mean ± standard deviation from 3 independent experiments.

The RelE cleavage efficiency of the Val codon and the UAG stop codon was quantified (Figure 6.6 C, top). The efficiency of RelE cleavage of the Val codon after MF dipeptide synthesis varied between the different RNAs, therefore the amount of cleavage after translation elongation was calculated as a percentage of the cleavage after the translation initiation control (after MF synthesis, Figure 6.6 C, bottom). This allowed comparison of the cleavage efficiency of the UAG stop codon relative to the Val codon between the different RNAs, despite the variation in cleavage after MF

synthesis. As RelE was able to cleave the valine codon after MF dipeptide synthesis on all the RNAs, lack of RelE cleavage after synthesis of the longer peptides would suggest that when the 1st nt of the UAG stop codon is 8 nts or less from the RNAP proposed rear end the ribosome is unable to position the stop codon in the A-site that allows RelE cleavage. A distance of 11 nts was however, sufficient. As the ribosome was able to translocate fully in the absence of the RNAP on all RNAs, the lack of RelE cleavage after the longer peptides in the coupled system suggested that the presence of the RNAP on the RNA did not allow the ribosomes to translocate fully. This result proved to be reproducible.

To narrow down the distances between the RNAP and the ribosome further, MFVVVR+24_12aa RNA was used as a basis to add up to three nts between the stop codon and the RNA:DNA hybrid to increase the distance between the peptide and RNAP in 1 nt increments. The peptide length and sequence was kept the same, but the distance between the stop codon and the RNAP rear end was increased sequentially from 8 to 11 nts. The RNAs were named MFVVVR+25_12aa, +26_12aa and +27_12aa based on the number of additional nucleotides added between the peptide coding region and the RNA:DNA hybrid sequence compared to the original MFVVVR RNA. The longest RNA, MFVVVR+27_12aa, had the same distance as MFVVVR+24_11aa from the 1st nt of the UAG codon to the RNAP rear end when coupled (11nts). These RNAs were used in TL-CTT as previously and the results shown in Figure 6.7 A. The RelE cleavage efficiency was quantified as previously and the results shown in Figure 6.7 B. The difference in cleavage efficiency with the addition of successive nts between the peptide coding region showed a gradual increase. There did not appear any defined, exact distance at which the ribosome was suddenly unable to incorporate the amino acid and translocate, as judged by RelE cleavage of the UAG stop codon. The results from all of the MFVVVR+24 extended peptides were collated and are shown in Figure 6.7 C. These results suggest that as the ribosome translocated towards the RNAP, it became less able to position the UAG codon in the A-site for optimum cleavage by RelE.

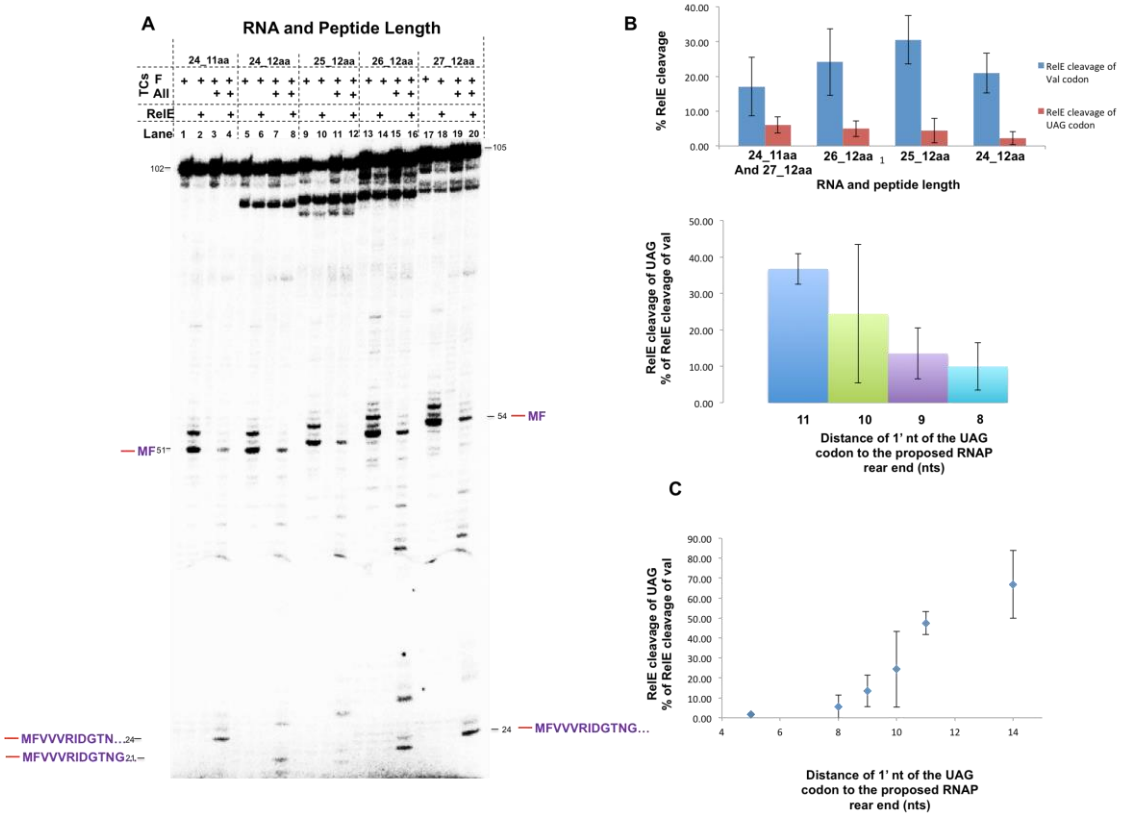


Figure 6.7 TL-CTT and RelE cleavage to determine the distances (2). A) PAGE after TL-CTT and RelE cleavage of the UAG stop codon on the RNAs MFVVVR+24_11 (lanes 1-4) and 12aa (lanes 5-8) and MFVVVR+25_12 (lanes 9-12), 26_12 (lanes 13-16) and 27_12 (lanes 17-20). **B)** Quantification of the RelE cleavage efficiency of the UAG stop at the decreasing distances from the RNAP rear end. Top: the % of RelE cleavage of both the Val codon (after MF dipeptide synthesis) and the UAG stop codon (after full length protein synthesis) are shown in blue and red respectively. Bottom: the cleavage efficiency of the UAG stop codon is shown as a percentage of cleavage of the Val codon after MF synthesis. The values are the mean with the variance from 2 separate experiments. **C)** Summary of the RelE cleavage efficiency at the UAG codon on all RNAs. The values are mean ± standard deviation from 3 or more independent experiments except distances 9 and 20 where the values are the mean with the variance from 2 separate experiments.

6.3.1 Optimising Transcription

The abasic original template was used to position the RNAP and synthesise the full length RNA during TL-CTT. As seen previously with the shorter RNAs (see section 5.3), the RNAP backtracked and cleaved the RNA during translation when the MFVVVR+24 RNAs were used (Figure 5.14). In TL-CTT, during the final translation elongation and RelE cleavage stages, the reaction contained both GreA and B and NTPs, including the α-[^{32}P]-NTP used to label the RNA. Under these conditions, the RNAP was undergoing successive rounds of forward translocation, misincorporation of GTP, backtracking and cleaving of the RNA, either due to the Gre factors or the GTP concentration, restoring the 3' end of the RNA in the active centre for RNAP to transcribe to the abasic site to begin the cycle again. Analysis of the RNA species after TL-CTT revealed that all the control samples not treated with RelE contain the full length RNA suggesting the vast majority of the AAECs were located

in the forward position after misincorporation of the GTP to produce the 101 nt long RNA.

RelE cleavage analysis of translation after TL-CTT revealed that the ribosome was able to fully position the UAG stop codon in the ribosomal A-site after synthesis of the 11 amino acid peptide and when positioned 11 nts from the proposed rear end of the RNAP, but, when the UAG stop codon was positioned one codon closer, the RelE cleavage efficiency was significantly reduced. However, the exact distance was not determined using the original abasic template (Figure 6.7). Cycling of the RNAP between the post-translocated state and the backtracked position could potentially impact upon on the ability of the ribosome to translate the RNA efficiently and position the UAG stop codon in the A-site correctly for RelE cleavage.

The original abasic template was chosen over all the other templates tested in section 5.4 due to the greatest amount of backtracking of the RNAP that occurred after AAEC assembly and is required for efficient extension of the RNA during transcription. In TR-CTT, full extension of all of the RNA in the reaction was very important, as the distance between all of the RNAP molecules and the ribosomes were analysed based on the total peptide synthesised in the reaction. Any AAECs located in the incorrect position would have affected the overall result. A high concentration of full length RNA was also required for synthesis of enough peptide for it to be visible after TLE. In TL-CTT, where only the active AAECs were visualised by incorporation of a radiolabeled NTP during transcription, the presence of unextended AAECs did not interfere with the result, as they were not labeled and therefore not visible.

During Chapter 5, modification of the original abasic template resulted in alternative templates also containing the abasic site on which the RNAP was located stably at the abasic site after transcription elongation, both in the presence and absence of the high concentration of GTP used in translation (see section 5.3, Table 5.1). These templates were deemed to be unsuitable for TR-CTT due to the proportion of RNA that was not extended during transcription. However, as this is not an issue in TL-CTT using incorporation of α-[^{32}P]-NMPs to radiolabel and visualise the RNA, it is possible that these templates will be suitable for use in TL-CTT.

Three templates (templates 3, 4 and 8, Table 5.1) were chosen for analysis in TL-CTT using MFVVVR+24_10aa RNA. TL-CTT was performed as standard using the 3 different DNA templates and non-templates to assemble the AAEC, as well as

the abasic t/ntDNA for comparison. RelE cleavage was carried out after TL-CTT and the results shown in Figure 6.8 B. Out of the three templates tested, template 4 produced the clearest result with the least amount of background bands (Figure 6.8, lanes 13-16) and also contained the most stable AAEC after transcription elongation, based on characterisation of the AAECs in section 5.4. Template 4 contained a dTMP residue before the abasic site that allowed incorporation of the modified ATP analogue, cordycepin, at the 3' end of the RNA. This modified NTP lacks the 3' OH group required for addition of the next nucleotide and therefore acted as a terminator NTP and prevented RNAP adding an extra NMP to the RNA opposite the abasic site.

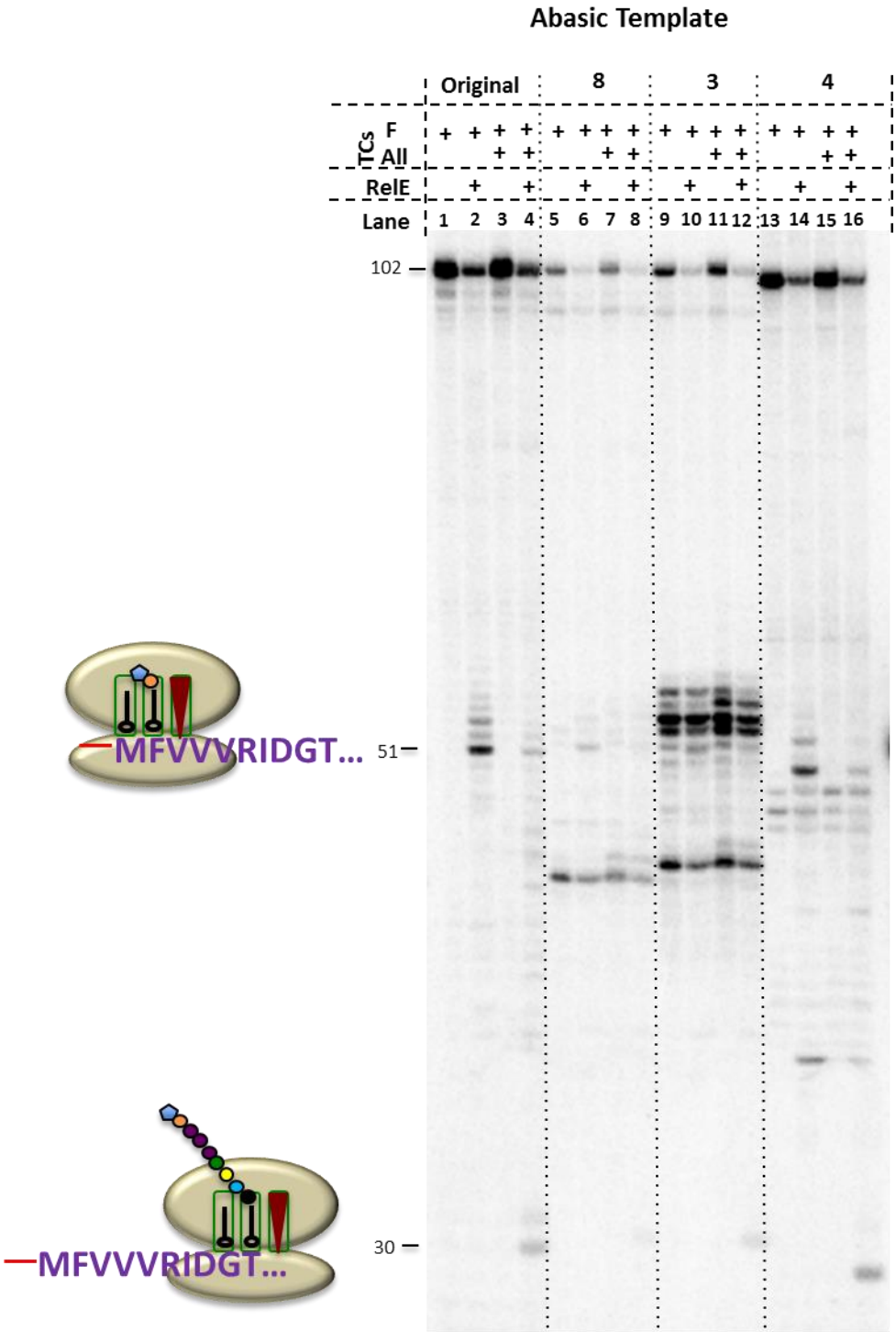


Figure 6.8 Comparison of DNA templates in TL-CTT. TL-CTT using MFVVVR+24_10aa RNA and the abasic original DNA template (lanes 1-4), template 8 (lanes 5-8), template 3 (lanes 9-12) and template 4 (lanes 13-16). RelE cleavage at the valine codon (after MF synthesis, lane 14) and at the UAG stop codon (after MFVVVRIDGT synthesis, lane 16) on template 4 produces RNA that is one base shorter than on the other templates. This is due to the RNA being one base shorter after synthesis as the RNAP does not incorporate GTP opposite the abasic site as it does on the other templates.

After determining that template 4 produced the least amount of background after TL-CTT with the MFVVVR+24_10aa RNA, TL-CTT was repeated using all of the MFVVVR+24 based RNAs (MFVVVR+24, MFVVVR+24_10aa-13aa), the AAEC

assembled with template 4 t/ntDNA and labeling of the RNA by incorporation of α[-^{32}P]-CTP during transcription. After transcription and translation of either the dipeptide or the full-length peptide, RelE was added then the reaction stopped and the samples analysed. The results are shown in Figure 6.9 A. Although TL-CTT using DNA template 4 resulted in the least amount of background, the band produced from the cleavage of the RNA at the valine codon after MF dipeptide synthesis was still obscured by background bands from unspecific transcription.

In the standard protocol for TL-CTT the RNA is labeled by incorporation of α-[^{32}P]-CTP during transcription elongation. The NTPs required for transcription elongation on template 4 were GTP, CTP, UTP and ATP (or cordycepin). As the background bands seen after TL-CTT with template 4 were due to unspecific transcription, it is possible that using a different radiolabelled NTP during transcription will reduce the visibility of these products. α-[^{32}P]-GTP is not suitable for use in transcription elongation as it will be diluted by the cold GTP used in translation and cordycepin is not available as a radiolabeled NTP, but α-[^{32}P]-UTP is a potential alternative.

To determine if the use of α-[^{32}P]-UTP during TL-CTT reduced the appearance of the background bands, TL-CTT was repeated using MFVVVR+24_10aa RNA and template 4 t/ntDNA, but with α-[^{32}P]-UTP added along with cold CTP, GTP and cordycepin during transcription elongation. After RelE cleavage, the reactions were stopped, visualised and the results shown in Figure 6.9 B. Use of α-[^{32}P]-UTP in TL-CTT with template 4 greatly reduced the background bands of a similar size to the RelE cleavage product after MF dipeptide synthesis (Figure 6.9 B lane 3) by reducing the signal from unspecific transcription due to assembly of the AAEC on impurities in the RNA. Use of α-[^{32}P]-UTP, however, increased the level of background signal from RNA products of approximately the same size as those produced from RelE cleavage of the UAG stop codon. The reaction in lane 4 was stopped immediately after translation elongation, whereas the reactions in lanes 5 and 6 where incubated for a further 10 min with (lane 6) or without (lane 5) RelE and show that the intensity of background signal also increased with the length of incubation. As these background products were seen in the absence of RelE, they were not due to cleavage of the RNA by RelE or the presence of the ribosomes, as this would have affected the dipeptide samples too. This suggested it was due to the presence of TCs and EF-G in the reaction, as the background bands were seen only after the addition

of TCs and EF-G to the reaction. It is unlikely that EF-G had any impact, but it was possible that the presence of a large amount of uncharged tRNA in the reaction was the cause. The template DNA and tRNA could somehow have formed an RNA:DNA hybrid on which the RNAP then assembled an AAEC and was able to transcribe and incorporate α-[^{32}P]-UTP into an RNA product. The presence of the large amount of GTP in the system could also have been a contributing factor. Decreasing the amount of RNAP in transcription to reduce the amount of free RNAP after AAEC assembly did not reduce the amount of unspecific transcription, whereas decreasing the amount of GTP added during translation elongation lead to an increase in background RNA (data not shown).

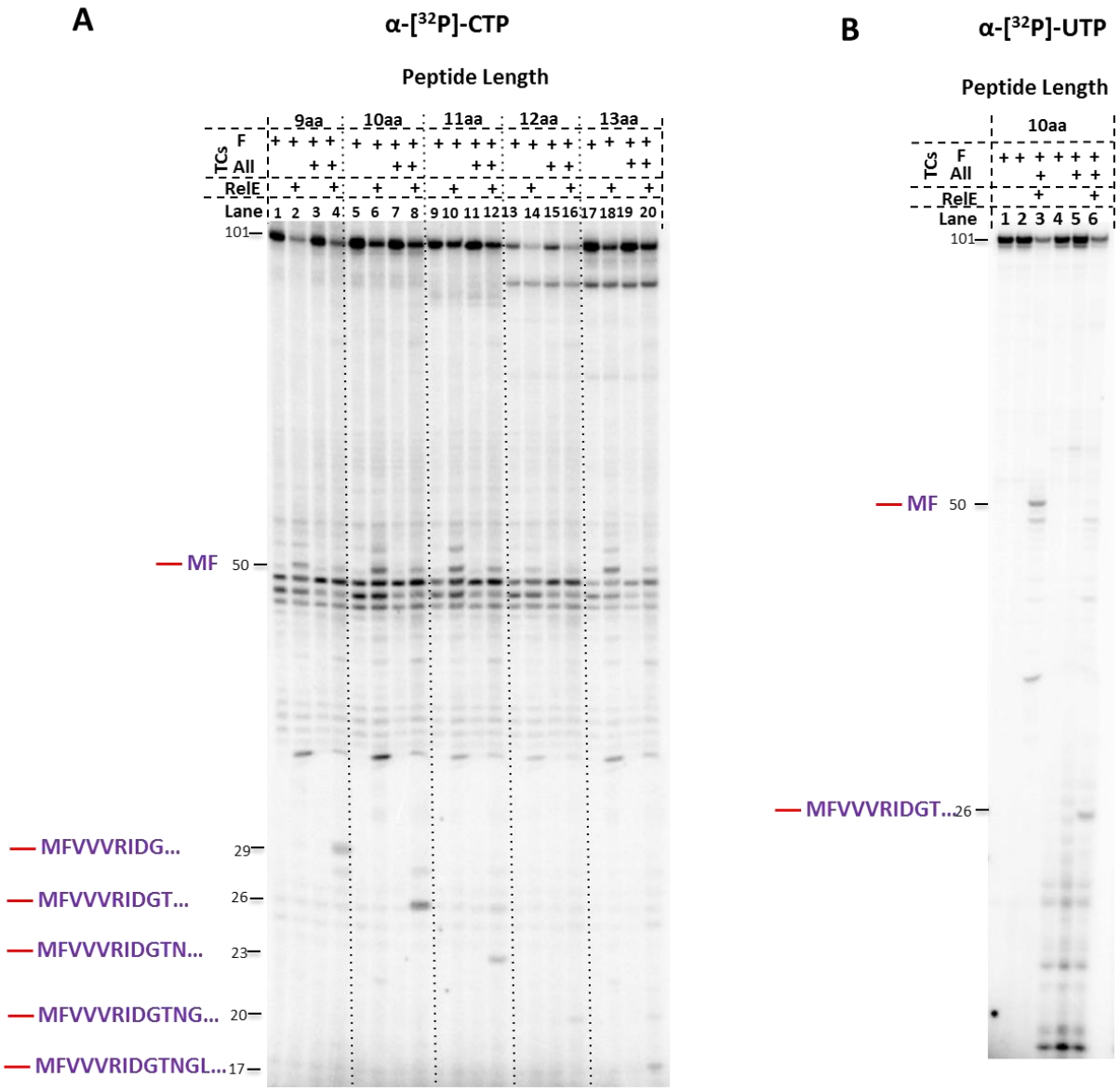


Figure 6.9 Comparison of α-[^{32}P]-NTP during TL-CTT on template 4. A) RelE cleavage analysis TL-CTT with all extended peptide sequence +24 RNAs (MFVVVR+24 and MFVVVR+14_10 to 13)with radiolabelling by incorporation of α-[^{32}P]-CTP. Background signal obscures the band produced by RelE cleavage of the Val codon after MF dipeptide synthesis (50 nt in size). B) TL-CTT using

MFVVVR+24_10aa RNA and radiolabelling with α-[^{32}P]-UTP. The background signal around the band produced by RelE cleavage after MF synthesis (lane 3) disappears but background bands appear around the band produced by RelE cleavage of the UAG stop codon (lane 6). This background signal only appears in the samples taken after translation elongation (compare, elongation lanes 4 to 6 to initiation, lanes 1-3)

If the tRNA and tDNA were forming hybrids for the RNAP to bind, form an AAEC and synthesise RNA, these AAECs would most likely have been unstable. Heparin is a compound that binds to the active centre of RNAP and disrupts unstable TECs, as well as sequestering free RNAP (Walter *et al.* 1967). If the background signal was due to unspecific transcription, the addition of heparin after transcription but before translation elongation may reduce the background signal. However, as heparin is not normally a component of TL-CTT, the effect of an increasing concentration of heparin on translation elongation and RelE cleavage was unknown. Heparin was first tested in translation alone using MFVVVR+24_10aa RNA to determine if it disrupts translation elongation complexes and/or the ability of RelE to bind the ribosome and cleave the RNA. Translation was assembled as for TL-CTT but using 5' end labeled RNA and the MF dipeptide reaction passed through a sucrose cushion to ensure that the only RNA in the system was bound by a ribosome. After the ribosome pellet was resuspended, the translation elongation reactions were assembled with an increasing concentration of heparin added before the TCs and EF-G. The heparin concentration tested ranged from 10 µg/mL to 89 µg/mL. After translation elongation, RelE was added and the results, shown in Figure 6.10 A, indicated that even the highest concentration, heparin did not interfere with translation elongation and/or RelE cleavage (Figure 6.10 A, compare lane 21 to 6).

Next, the effect of heparin was analysed in TL-CTT to ensure that heparin did not affect translation and/or transcription, both of the full-length transcription product and the shorter background products after TL-CTT. TL-CTT was performed as standard using MFVVVR+24_10aa RNA, template 4 t/ntDNA, both GreA and B and labeling of the RNA with α-[^{32}P]-UTP. Heparin was added to the reaction after AAEC assembly and transcription elongation but just before the TCs and EF-G were added for translation elongation. The same concentration range was used as for translation alone (10-89 µg/mL). Samples were taken directly after translation elongation and after incubation with and without RelE and the results shown in Figure 6.10 B, lanes 1 to 14. Increasing the concentration of heparin had no effect on the full-length transcription product or translation elongation, even at the highest concentration (89 µg/mL). The presence of heparin during translation elongation and RelE cleavage did,

however, have the desired effect of reducing the background RNA products by reducing the background signal to a barely visible level in the presence of the highest concentration of heparin (Figure 6.10 B lanes 13 & 14).

Alongside the TL-CTT experiment carried out above, a TL-CTT reaction in which Gre factors were omitted during AAEC and translation elongation was also performed and the highest concentration of heparin added during translation elongation. The results from this experiment are shown in Figure 6.10 B lanes 15 to 20 and revealed that the amount of background RNA around the 20 to 30nt size was diminished in the absence of Gre factors compared to the when GreA/B were added during AAEC assembly, even in the absence of Heparin (Figure 6.10 B compare lanes 2 & 3 to 17 & 18). Addition of heparin at the highest concentration (89 µg/mL) reduced the background noise even more to a point where the RelE cleavage product was very clearly distinguishable from any remaining background noise (Figure 6.10 B lanes 19 & 20).

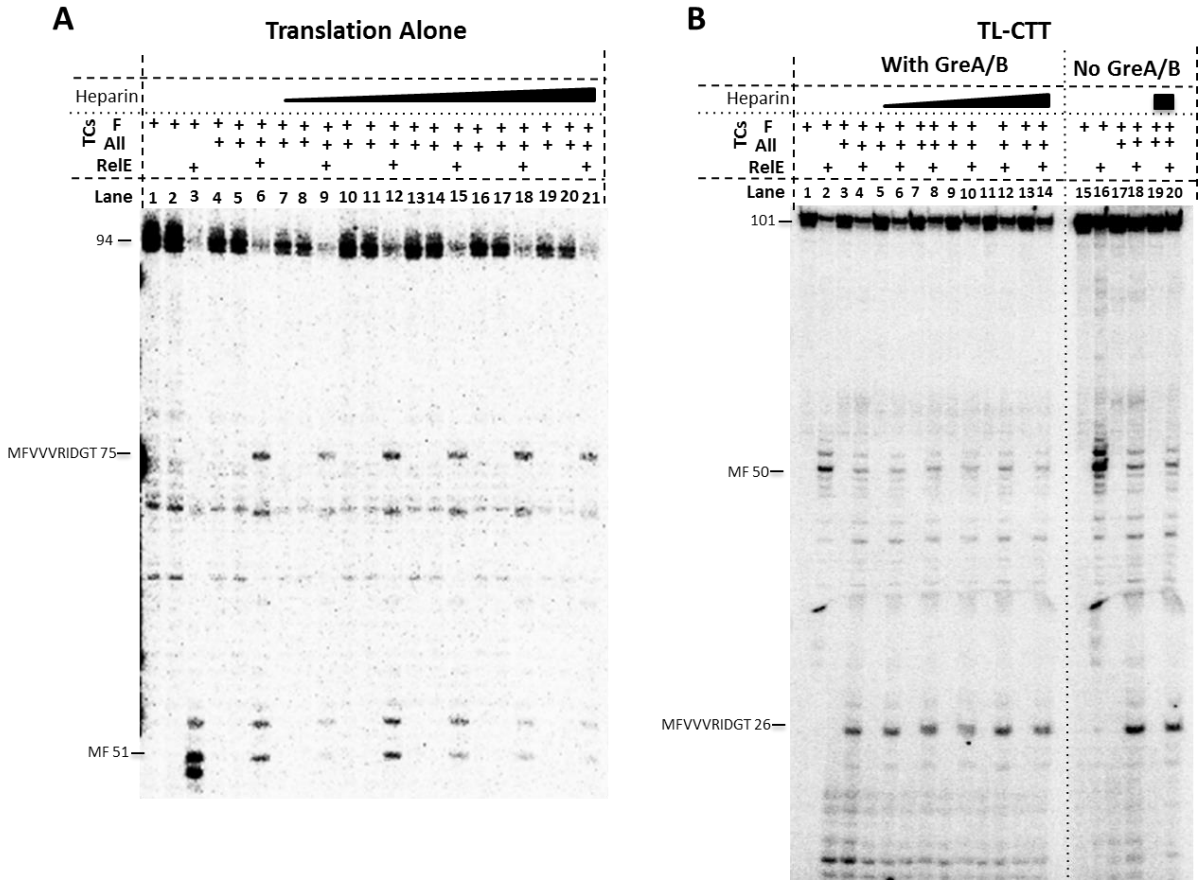


Figure 6.10 Test of Heparin concentration during translation and TL-CTT. A) Increasing heparin concentration in translation alone on MFVVVR+24_10aa RNA. A sample was taken before heparin addition as a RelE cleavage control after MF dipeptide synthesis. Heparin was added at concentrations of 0, 10, 20, 40, 75 and 89 µg/mL just prior to translation elongation and RelE cleavage. Heparin at the highest concentration (89 µg/mL, lane 21) appeared to have no negative effect on translation elongation compared to without heparin (lane 5) B) TL-CTT with MFVVVR+24_10aa RNA and an increasing concentration of heparin (as A). Lanes 1-14 contain GreA/B, NTPs during

transcription and translation and heparin during translation where indicated. Lane 15-20 contains the reaction without GreA/B during transcription and with heparin (lanes 19-20, 89 µg/mL) and without heparin (lanes 15-18).

The TL-CTT technique was once again modified taking into account these latest observations. The Gre factors were removed and the AAEC assembled with template 4 tDNA/ntDNA, labeling of the RNA with α-^{32}P-UTP and the addition of heparin at a final concentration of 89 µg/mL after transcription but before translation elongation. Using this updated method, TL-CTT was repeated with the MFVVVR+24 RNA and MFVVVR+24_10-12aa RNAs to compare the RelE cleavage efficiency of the UAG stop codon at different distances from the RNAP proposed rear end buffer. The results from this TL-CTT are shown in Figure 6.11 A.

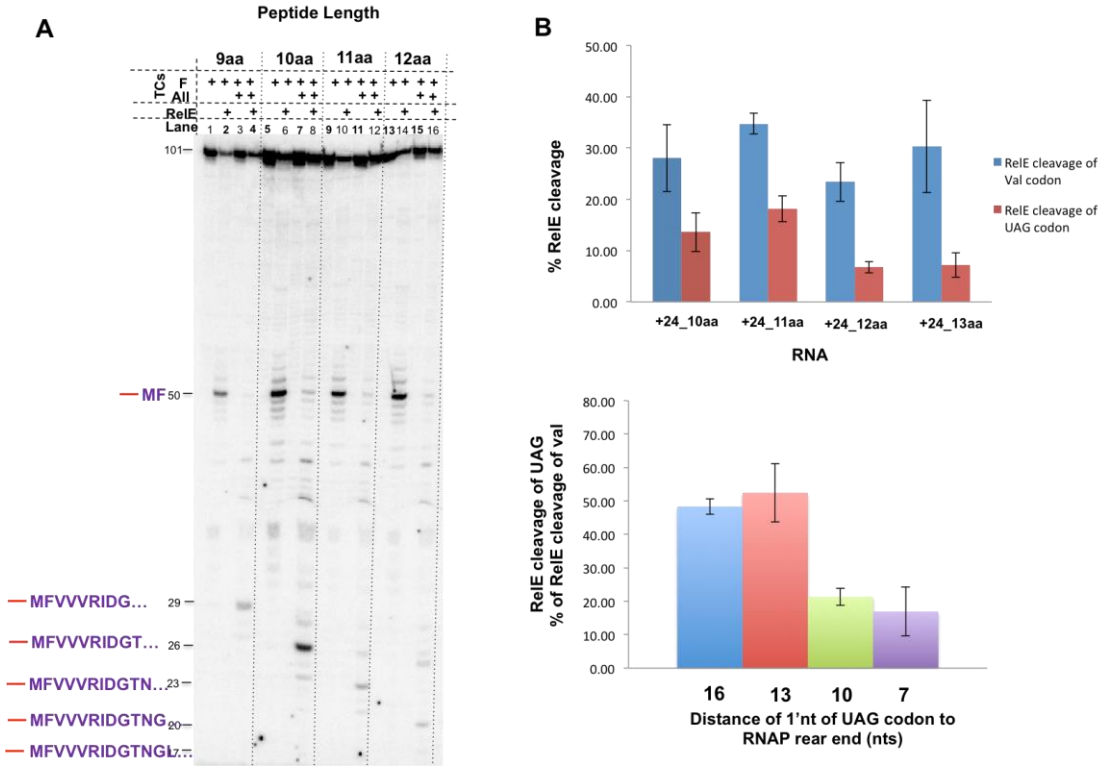


Figure 6.11 TL-CTT using template 4 and α-[^{32}P]-UTP. A) PAGE of RelE cleavage at the UAG stop codon after TL-CTT using MFVVVR+24_9aa (lanes 1-4), 10aa (5-8), 11aa (lanes 9-12), 12aa (lanes 13-16) RNAs. RelE cleavage of the UAG stop codon is seen after TL-CTT of all peptides (lanes 4, 8, 12, 16 and 20) but more is seen after the 9 and 10 aa peptides compared to the 11 and 12 aa peptides. As RelE cleavage is seen after MF dipeptide synthesis of all RNAs, this would suggest the RNAP is preventing the ribosome positioning the UAG stop codon in the ribosomal A-site after peptide synthesis with the MFVVVR+24_11, and 12 RNAs. **B)** Quantification of RelE cleavage efficiency. Top: the % of RelE cleavage of both the Val codon (after MF dipeptide synthesis) and the UAG stop codon (after full length protein synthesis) are shown in blue and red respectively. Bottom: the cleavage efficiency of the UAG stop codon shown as a percentage of cleavage of the Val codon after MF synthesis at varying distances from the RNAP rear end. The values are mean ± standard deviation from 3 or more independent experiments except MFVVVR+24 where the values are the mean with the variance from 2 separate experiments.

With template 4, RelE cleavage of the UAG stop codon after translation elongation is seen with all the RNAs, including the MFVVVR+24_12aa RNA. After

TL-CTT, RelE cleavage was quantified as for the results obtained using the original abasic template (Figure 6.11 B). There is a clear reduction in RelE cleavage efficiency of the UAG stop codon after synthesis of the 11 amino acid peptide compared to the 10 amino acid peptide. This would suggest that when the distance between the RNAP and the ribosome is 13 nts, the ribosome is better able to synthesise the peptide and position the stop codon in the A-site compared to a distance of 10 nts.

6.4 Conclusion and Discussion

In this chapter, TL-CTT was modified for use in determining how close the RNAP and the ribosome could become in an *in vitro* coupled system. The ability of the ribosome to translocate and position a UAG stop codon in the A-site was analysed by RelE cleavage and the efficiency of cleavage compared between stop codons located at different distances from the proposed RNAP rear end.

During the analysis and modification of TL-CTT, a number of limitations to the original TL-CTT method were discovered and modifications not previously used were made to the technique, as occurred with TR-CTT. Firstly, positioning the ribosome too close to the RNA:DNA hybrid site affected the ability of the RNAP to elongate the RNA in the presence of NTPs. Whether the ribosomes prevented the RNA:DNA hybrid forming, the RNAP binding or the backtracking of the RNAP required for GreA stimulated RNA cleavage shown to be essential for the RNAP to elongate the RNA was not identified. Neither was the minimum distance required between the ribosomes and the hybrid to allow correct formation of AAEC determined, as these were outside the remit of this project.

To translocate the ribosome closer to the RNAP, a longer peptide was synthesised. Previously, only amino acids with commercially available uncharged tRNAs had been used in either TR-CTT or TL-CTT, but, in order to extend the peptide sequence during this project, additional amino acids that did not have commercially available uncharged tRNAs were used. Initially, we tried to purify these specific tRNAs, but this approach proved unsuccessful. Instead, a total tRNA mix was aminoacylated with specific amino acids and these aa-tRNA^{aa}s used in translation elongation, an approach designed and tested exclusively during this project.

Using this technique results in aminoacylation of only a subset of the total tRNA and there would have been a large amount of uncharged free tRNA also present, a condition not normally encountered by the ribosome, therefore it is

possible that the uncharged tRNA will compete with the charged-tRNA for binding in the ribosome and may slow down the incorporation of the correct aa-tRNA^{aa}. Another consequence of the high amount of tRNA in the system was the production of background signal that interfered with the detection of the RelE cleavage products. However, this was overcome by the addition of heparin after transcription elongation as it was determined during this project that heparin did not have an effect on translation elongation during TL-CTT. An alternative strategy to produce pure tRNA is to synthesise specific tRNAs from a DNA template using T7 RNAP. These tRNAs would not have the modified nucleotide structures present within tRNAs isolated from culture, but they may still be suitable for use in TL-CTT.

Initially, the original abasic DNA template on which the RNAP cycled between a backtracked and fully elongated state was used to determine that there was a very distinct difference in cleavage efficiency at a distance of 11 nts from the 1st nt of the UAG stop codon to the proposed RNAP rear end, compared to when this distance was reduced to 8 nt. The efficiency of RelE cleavage of the UAG codon was shown to reduce gradually as the distance was decreased from 11 nts to 8 nts. This was not entirely unexpected due to the cycle of forward transcription and backtracking of the RNAP caused by the presence the misincorporation of GTP opposite the abasic site during translation. Although the majority of the RNAP molecules were located in the forward translocated position, some of the RNAP molecules will have been located homogenously on the DNA template. This could explain why a precise distance could not be identified, as not all of the RNAP was located homogenously in the post-translocated position. The RNAP oscillates locally between the pre and post-translocated state during transcription (Komissarova & Kashlev 1997a), which may also prevent the distance being narrowed down to a specific nucleotide.

Another DNA template, template 4, was also used in TL-CTT. On this template, characterisation showed that the RNAP was located stably in the post-translocated state after transcription elongation both in the absence and presence of a high GTP concentration. RelE cleavage of the UAG codon after TL-CTT revealed that cleavage at a distance of 13 nts (from the 1st nt of the UAG stop codon to the proposed RNAP rear end) was much more efficient compared to a distance of 10 nts. Although the cleavage efficiency of the UAG codon by RelE after TL-CTT varied slightly between the two templates, the overall result was shown to be consistent. As the RNAP is stably located in the post-translocated state after transcription elongation on template

4, it may be possible to define the distance more precisely using RNA that increases the distance one nt at a time from 10 to 13 nts.

Hydroxyl radical footprinting of the ribosome and RNAse footprinting of the RNAP suggest that together they cover a 31 nt region of the RNA from the RNAP active site, to the 1st nt of the A-site codon, or 16 nts from the proposed RNAP rear end (Hüttenhofer & Noller 1994; Komissarova & Kashlev 1998). The data obtained during this project using RelE cleavage suggests that the ribosome can translocate close enough to position the 1st nt of the A-site codon as close as 11 nts to the RNAP rear end (or 26 nts to the active centre). This is closer than predicted and would suggest that the RNAP and the ribosome are able to interact in such a way that they are in closer proximity than expected.

7. Final Conclusion and Discussion

Initially, the transcription first coupled transcription-translation (TR-CTT) technique was chosen to determine how close the RNAP and ribosome were able to become to the RNAP in an *in vitro* system. The technique had already been established in our lab and was used at the beginning of this project to obtain preliminary data that indicated the ribosome could translocate close enough to position the A-site 8 nts from the proposed rear end of the RNAP (Komissarova & Kashlev 1998; Hüttenhofer & Noller 1994). However, due to the low level of peptide signal, this result could not be confirmed. Therefore, the system was analysed in order to identify the cause of the low peptide signal and find a solution.

First, analysis of transcription using 5' end labelled RNA revealed that the RNA contained an additional NMP, which was added to the 3' end by the T7 RNAP during synthesis. This extra NMP caused the RNAP to backtrack during AAEC and prevented the RNAP from extending the majority of the RNA during transcription. This result was completely unexpected but was overcome through modifying the original TR-CTT protocol to include the addition of the transcript cleavage factor, GreA, before the ntDNA during AAEC assembly. This resolved the initial backtracking and lead to the extension of the majority of the RNA during transcription elongation. However, the peptide signal was still low after this modification.

Further analysis revealed that the high concentration of GTP added during translation caused the RNAP to misincorporate GTP opposite the abasic site in the DNA template, leading to backtracking of the RNAP and cleavage of the RNA to produce an RNA species 5 nts shorter than anticipated. Although it is known that misincorporation of the incorrect NTP can cause backtracking the presence of a high concentration of the non-cognate NTP can cause intrinsic cleavage by the RNAP (Laptenko *et al.* 2003) it was another unexpected occurrence in this system that presented an additional obstacle to overcome. A variety of alternative DNA templates were designed and tested with the aim of eliminating the effect of the high GTP concentration but none were deemed to be suitable, either because the RNAP was not stably positioned in the post-translocated state in the presence of GTP or because the initial backtracking of the AAEC that is required for GreA stimulated RNA cleavage and extension of the RNA during transcription no longer occurred. The effect of decreasing the GTP concentration and incubation times during translation was also examined as a possible solution. An initial GTP hydrolysis assay suggested

that the concentration of GTP could be decreased, however, this subsequently proved not to be the case as reducing the GTP concentration and incubation times led to a decrease in the efficiency of translation elongation. Instead, the effect of the high concentration of GTP was overcome by using the original abasic template on which the RNAP misincorporated the GTP, backtracked and cleaved the RNA. However, this was solved by the addition of GreA, GreB and NTPs during translation to enable the RNAP to transcribe back to the abasic site after misincorporation and backtracking, Even after these latest modifications, however, the peptide signal after TR-CTT still remained low.

To test whether too little RNA present after washing of the AAECs was causing the lack of signal, TR-CTT was once again analysed. This was investigated by comparing the peptide signal produced from TR-CTT to the peptide signal produced from translation alone using a similar amount of RNA estimated to remain after washing of the AAECs. This revealed that the signal from TR-CTT was greatly reduced compared to translation alone, suggesting that the ribosomes had a reduced ability to initiate translation in the presence of RNAP. This was subsequently confirmed when the ability of the ribosomes to translate the RNA bound by the RNAP was assessed by the toxin RelE. RelE cleaves the RNA between the 2nd and 3rd nt of the codon in an empty ribosomal A-site and will only cleave RNA bound by the ribosome. Using this method, it was determined that the ribosome was not able to initiate translation on RNA bound by the RNAP. This occurred regardless of the distance between the RNAP and the SD site and initiation codon and was another surprising and unforeseen result. Nevertheless, this did explain the lack of peptide signal produced throughout the project, even after all the modifications to the TR-CTT technique. In light of this, it was concluded that TR-CTT was not a suitable approach for the purpose of this specific project.

Translation first coupled transcription-translation (TL-CTT) is the companion technique to TR-CTT and was initially designed and used to study the effect of translation on transcription (Castro-Roa & Zenkin 2012). In TL-CTT, translation is assembled first and the reaction filtered through a sucrose cushion. The AAEC was then assembled on the ribosome bound RNA. Therefore, this technique had the potential to be suitable for the purpose of this project but required modifications because, due to the order in which translation and transcription are assembled, it was not possible to analyse the interactions between the ribosome and the RNAP by

TLE of the peptide produced, as it could not be guaranteed that all of the translation elongation complexes were coupled to an AAEC. Instead, the TL-CTT protocol was modified to analyse translation by RelE cleavage of radiolabeled RNA. To visualise the RNA, α-[^{32}P]-NTPs were incorporated during transcription elongation to ensure that only the coupled complexes were observed during analysis of the RNA because only the RNA actively extended by the RNAP was labeled. The ability of the ribosomes to translocate towards the RNAP was determined by the efficiency of RelE cleavage after peptide synthesis. However, RelE cleavage of the RNA is very codon specific and was particularly inefficient at the isoleucine codon located in the A-site after synthesis of the MFVVVR peptide. To eliminate variation in the efficiency of RelE cleavage due to the difference in codons, a UAG stop codon was introduced after the peptide sequence to allow accurate comparison of RelE cleavage at different distances between the ribosome and the RNAP.

During TR-CTT, the distance between the ribosome and the RNAP was adjusted by changing the length of the RNA. For TL-CTT, however, this approach was not viable since reducing the distance between the AUG start codon and the RNA:DNA hybrid site reduced the ability of the RNAP to elongate the RNA when the AAEC was assembled after translation initiation. Instead, the length of the peptide synthesised by the ribosome was increased to move the stop codon progressively towards the RNA:DNA hybrid one codon at a time. The efficiency of RelE cleavage of the stop codon after peptide synthesis was then analysed and compared between the different distances. To extend the peptide sequence, amino acids with no commercially available, pure uncharged tRNAs were used. Purification of these specific tRNA using a biotin tagged DNA probe was trialed, but this approach proved unsuccessful. Instead, aminoacylation of specific tRNAs within a mix of total tRNA was tested as an alternative way of producing the required aa-tRNA^{aa} and was in fact shown to be a successful approach that had not previously been used in either the TR-CTT/TL-CTT systems. This technique also means the range of amino acids possible for use during TR-CTT and or TL-CTT is no longer limited to those with commercially available uncharged tRNAs. Charged tRNAs can now be produced for use in the synthesis of any peptide sequence, however, codon bias should however be taken into account when designing the peptide sequence, as the total uncharged tRNA will reflect the tRNA availability found within the cell they are isolated from.

The updated TL-CTT method incorporating the above modifications was used with two different DNA templates on which the RNAP was either cycling through forwards transcription, backtracking and RNA cleavage (the original abasic DNA template) or was positioned stably in the post-translocated state after transcription elongation (template 4). Using the original abasic template, a distance of 11 nts appeared to be sufficient for the ribosome to synthesise a peptide and position the UAG stop codon in the ribosomal A-site for cleavage of the RNA by RelE. When this distance was reduced to 8 nts however, the RelE cleavage efficiency was reduced, suggesting that the ribosome was no longer able to position the UAG stop codon in the ribosomal A-site in the correct conformation for RelE cleavage. The distance at which there was no longer any RelE cleavage of the UAG stop codon was not determined to a single nucleotide distance but instead a gradual decrease in cleavage efficiency occurred as the stop codon was moved closer to the RNAP rear end one nucleotide at a time from 11 nts to 8 nts. This was not entirely unexpected, as the RNAP is not completely static, particularly on this specific DNA template.

TL-CTT with template 4 indicated that a distance of 13 nts from the proposed RNAP rear end allowed RelE cleavage but reducing the distance to 10 nts and 7 nts progressively reduced, but did not entirely eliminate, RelE cleavage of the UAG stop codon. This gradual decrease in RelE cleavage efficiency as the stop codon is moved closer to the RNAP rear end reflects the results obtained using the original abasic DNA template. However, even at the shorter distances after TL-CTT using DNA template 4, there was a higher amount of RelE cleavage of the UAG stop codon compared to similar distances after TL-CTT with the original abasic DNA template. This may be due to the close proximity of the ribosomes causing a sub-population of the elongation complexes to be dislodged or to hyper-translocate forwards along the DNA template without extending the RNA. The transcription elongation complexes on the original abasic DNA template are undergoing a cycle of forward transcription, misincorporation of GTP, backtracking and RNA cleavage and this may stabilise the elongation complexes compared to template 4 where the RNAP is stalled at the abasic site.

The slightly higher amount of RelE cleavage occurring after TL-CTT with template 4 could also be a side effect of the heparin added during translation. Heparin disrupts less stable elongation complexes and as the system is not washed after transcription elongation, some of the elongation complexes may not contain

non-template DNA and therefore won't be as stable and resistant to heparin compared to the complexes containing both template and non-template DNA. As heparin is added after transcription elongation the less stable complexes will still extend and label the RNA, but these complexes may be disrupted during translation elongation and lead to free, radiolabeled RNA during translation elongation.

These distances, along with those obtained using the original abasic DNA template, contradict the preliminary results obtained using TR-CTT which suggest suggested that a distance of 8 nts from the ribosomal A-site to the RNAP was sufficient for the ribosome to incorporate the amino acid into the peptide. In light of the many unexpected features revealed through the extensive analysis of TR-CTT, it was concluded that these results were most likely an artefact of the technique rather than an accurate representation of the majority of the coupled complexes.

The results from this project indicated that the ribosomes could translocate closer to the RNAP that previously thought (Hüttenhofer & Noller 1994; Komissarova & Kashlev 1998). In order to achieve this, the front edge of the ribosome and the rear end of the RNAP must be overlapping, most likely with the ribosome front edge encroaching on the RNAP rear end based on the individual structures (Figure 7.1). Structural modelling of the interaction interface would be an exciting addition to this project.

The ability of the ribosome to synthesise the peptide and translocate towards the RNAP was determined by the efficiency of RelE cleavage of the UAG stop codon, with a lack of RelE cleavage thought to be due to the position of the RNAP physically blocking translocation of the ribosome to the next codon (the stop codon) after peptide bond formation. However, there are other possible explanations for the lack of RelE cleavage at the shorter distances. The A-site of the ribosome is located towards the front. Depending on the exact interaction interface between the RNAP and the ribosome, access by RelE may have been partially or fully blocked even if the ribosome was able to position the stop codon in the A-site. This could also block access of the next aa-tRNA^{aa} and lead to stalling of translation.

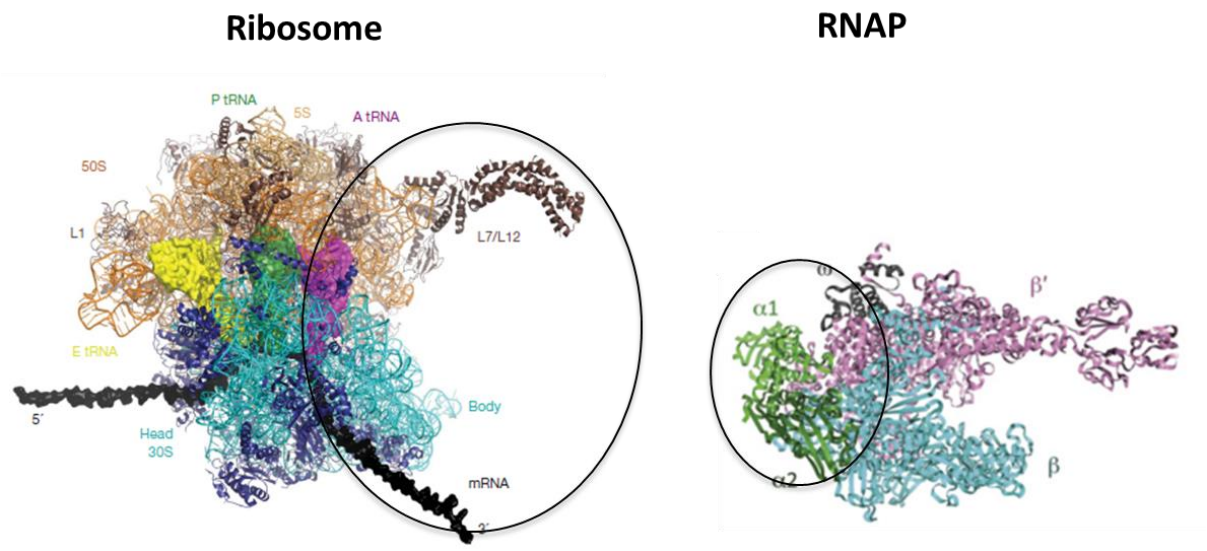


Figure 7.1 Ribosome and RNAP structures. The circles indicate the areas of potential interaction if they were bound to the same nascent RNA. Modified from Schmeing & Ramakrishnan 2009 and Weixlbaumer *et al.* 2013.

The EF-G binding site is also located at the front edge of the ribosome and a tight interaction between the ribosome and the RNAP could potentially block EF-G from binding. As binding of EF-G and hydrolysis of its associated GTP is essential for full translocation of the ribosome, any obstruction to EF-G binding would prevent translocation to the next codon.

If obstruction of the A-site does occur when the RNAP and ribosome are tightly associated, by blocking access of the aa-tRNA^{aa}, EF-G or other translation factors to their respective binding sites on the ribosome, this could have implications during translation *in vivo.*

The region of the RNAP potentially covered by the ribosome contains several important structural elements, including the RNA exit channel, the β flap domain, β' zipper domain, the β' lid domain and the β' zinc finger domain. All of these structures play various roles during transcription elongation and pausing. As these elements are subject to regulation by accessory factors or form interaction surfaces, the close proximity of the ribosome could impact on transcription regulation. This could be examined further using mutant RNAP lacking one or more of these structures in the TL-CTT system and adding the corresponding accessory factors.

Previous studies have already shown that translocating ribosome can affect stalled RNAP molecules both *in vivo* (Schweimer *et al.* 2010) and *in vitro* (Castro-Roa & Zenkin 2012). The strength of the effect on the paused RNAP may be dependent on how close to the RNAP the ribosome actually is. Due to the modifications designed, tested and implemented during this project, the amino acid sequence of

the peptide used in TL-CTT is no longer limited due use of total tRNA during aminoacylation. The ribosome can be positioned and the location determined using the stop codon and RelE cleavage analysis. This means it would be possible to examine the strength of the effect of the ribosome on the stalled RNAP as the interaction between the two machineries becomes progressively tighter.

The beauty of the coupled system is that only the minimal components required for transcription and translation are used. This means that additional factors can be added, if required, to study further the interactions between the RNAP and the ribosome in the presence of a wide variety of combinations of transcription and/or translation factors that are not essential for the basic processes.

8. List of Meetings

RNA Polymerase Workshop, Milton Keynes, UK, March 2013

RNA Polymerase Workshop, Newcastle University, UK, April 2014. **Poster Presentation**

RNA Polymerase Workshop, Birmingham University, UK, March 2015, **Poster Presentation**

FASEB SRC Summer Conferences. June 21 - 26, 2015 in Vermont Academy, Saxtons River, Vermont, USA. **Poster presentation**

RNA Polymerase Workshop, University of York, UK, March 2016

9. Appendix

A1 RNA Synthesis

The DNA template was generated by PCR using either of the parental DNA sequences below, the T7A1 forward primer and the reverse primer specific for each RNA.

Parental template sequence:

TCCCGCGAAATTAATACGACTCACTATAGGGAGACAACTGTTAATTAAATTAAATTAAAAAGGAAATAAAAATGG
AGTTTGTAGGGAAAATCGAG

or

TCCCGCGAAATTAATACGACTCACTATAGGGAGACAACTGTTAATTAAATTAAATTAAAAAGGAAATAAAAATGT
TTGTGGTGGTGCGCAAAATCGAG

T7 Forward Primer: TCCCGCGAAATTAATACGACTC

RNA	Reverse Primer Sequence
MFVE	CTCGATTTTCCCTACAAACTCCATT
MFVEE	CTCGATTTTCTCTTCTACAAACATT
MFVEEE	CTCGATTTTCCTCTTCTACAAACATT
MFVVVE	CTCGATTTTCTCCACCACCACAAACATT
MEFV	CTCGATTTTCCCTACAAACTCCATT
MFVR	CTCGATTTTCCCTCGTACAAACATT
MFVRR	CTCGATTTTCCGTCGTACAAACATT
MFVVVR	CTCGATTTTGCGCACCACCACAAACATT
MFVVVR+1	CTCGATTTTCGCGCACCACCACAAACATT
MFVVVR+2	CTCGATTTTCCGCGCACCACCACAAACATT
MFVVVR+3	CTCGATTTTCCCGCGCACCACCACAAACATT
MFVVVR+4	CTCGATTTTGCCCGCGCACCACCACAAACATT
MFVVVR+5	CTCGATTTTCGCCCGCGCACCACCACAAACATT
MFSt+5	CTCGATTTTCGCCCGCGCACCACCTAAAACATT
MFVVVSt+5	CTCGATTTTCGCCCCTACACCACCACAAACATT
MFSt+11	CTCGATTTTTATCCGTCAATCTACACCACCACAAACATT
MFVVVSt+11	CTCGATTTTTATCCGTCAATCTACACCACCACAAACATT
MFSt+17	CTCGATTTTTTAGTCTATCCGTCAATGCGCACCACCTAAAACATT
MFVVVSt+17	CTCGATTTTTTAGTCTATCCGTCAATCTACACCACCACAAACATT
MFSt+24	CTCGATTTTCAATCCATTAGTCTATCCGTCAATGCGCACCACCTAAAACAT
MFVVVSt+24	CTCGATTTTCAATCCATTAGTCTATCCGTCAATCTACACCACCACAAACAT
MFVVVR+24	CTCGATTTTCAATCCATTAGTCTATCCGTCAATGCGCACCACCACAAACAT
MFVVVR+24_10aa	CTCGATTTTCAATCCATTCTAGGTTCCGTCAATGCGCACCACCACAAACAT
MFVVVR+24_11aa	CTCGATTTTCAATCCCTAATTGGTTCCGTCAATGCGCACCACCACAAACAT
MFVVVR+24_12aa	CTCGATTTTCAACTATCCATTGGTTCCGTCAATGCGCACCACCACAAACAT
MFVVVR+25_12aa	CTCGATTTTCAATCTATCCATTGGTTCCGTCAATGCGCACCACCACAAACAT
MFVVVR+26_12aa	CTCGATTTTCAATCCTATCCATTGGTTCCGTCAATGCGCACCACCACAAACAT
MFVVVR+27_12aa	CTCGATTTTCAATCCCTATCCATTGGTTCCGTCAATGCGCACCACCACAAACAT
MFVVVR+24_13aa	CTCGATTTTCTACAATCCATTGGTTCCGTCAATGCGCACCACCACAAACAT
MFVVVR+32	CTCGATTTTAATTAGTTATGTGAATGAATATTATTATGTATGCGCACCACCACA AACAT

Table 9.1 RNA Synthesis Primers and Template. The reverse primers used for generating the DNA template for synthesis of the RNA.

A2. DNA Template and Non-template Sequences

Template	Sequence
Original	T: TGAATGTCGGTTTTAGCTCCCGGTTTAACCGCGGCTGCTT NT:ACTTACAGCCAAAATCGAGGGCCAAATTGGCGCCGACGAA
Abasic Original	T: TGAATGTCGGTTTTAGCTCCCGGTTT_ACCGCGGCTGCTT
Template 1	T: GGTAGTAATACTTTAGCTCCCGGTTT_ACCGCGGCTGCTTAAGCCTGGG NT:CCATCATTATGAAATCGAGGGCCAAATTGGCGCCGACGAATTCGGACCC
Template 2	T: GGTAGTAATACTTTAGCTCGCTCTCC_AACCGCGGCTGCTTAAGCCTGGG NTCCATCATTATGAAATCGAGCGAGAGGTTTGGCGCCGACGAATTCGGACCC
Template 3	T: GGTAGTAATACTTTAGCTCGCTCTCC_AGGGCGGCTGCTTAAGCCTGGG NT:CCATCATTATGAAATCGAGCGAGAGGTTCCCGCCGACGAATTCGGACCC
Template 4	T: GGTAGTAATACTTTAGCTCGCACACT_ACCGCGGCTGCTTAAGCCTGGG NT:CCATCATTATGAAATCGAGCGTGTGATTGGCGCCGACGAATTCGGACCC
Template 5	T: GGTAGTAATATTTTAGCTCGCTCTCC_AGGGCGGCTGCTTAAGCCTGGG NT:CCATCATTATAAAATCGAGCGAGAGGTTCCCGCCGACGAATTCGGACCC
Template 6	T: GGTAGTAATGTTTTAGCTCGCTCTCC_AGGGCGGCTGCTTAAGCCTGGG NT:CCATCATTACAAAATCGAGCGAGAGGTTCCCGCCGACGAATTCGGACCC
Template 7	T: GGTAGTAACGTTTTAGCTCGCTCTCC_AGGGCGGCTGCTTAAGCCTGGG NT:CCATCATTGCAAAATCGAGCGAGAGGTTCCCGCCGACGAATTCGGACCC
Template 8	T: TGAATGTCGGTTTAGCTCGCTCTCC_AGGGCGGCTGCTTAAGCCTGGG NT:ACTTACAGCCAAATCGAGCGAGAGGTTCCCGCCGACGAATTCGGACCC

Table 9.2 DNA Template and non-template Sequences. The full template (T) and non-template (NT) sequences for all of the DNA templates used. The Abasic original DNA template was used in transcription with the original non-template DNA. Indicate 5' and 3'

A3. tRNA Purification Probe Sequences

tRNA	Probe Sequence
Isoleucine	ACCCTTATCAGGGGTGCGCTCTAACCACCT-bio
Aspartate	TGCGTGACAGGCAGGTATTCTAACCGACTG-bio
Glycine	ATCAGCTTGGAAGGCTGAGGTAATAGCCAT-bio

Table 9.3 tRNA Purification Probes. The probes contained a biotin tag on the 3' end of the DNA.

A4. List of strains and Plasmids

E. coli Strain	Antibiotic	Characteristic	Genotype	Source

	Resistance			
MRE600	N/A	Used for purification of tRNA	*F-, rna*	(Cammack & Wade 1965)
JE28	Kanamycin	6X his-tagged ribosome	MG1655 (rplL-his6):kan:rpoB+	(Ederth *et al.* 2009)
T7 Express	N/A	Used for the overexpression of recombinant proteins		NEB
DH5α	N/A	Used for plasmid propagation		NEB

Table 9.4 List of strains used in this work.

Plasmid	Resistance	Characteristic	Inducer	Reference
pIA468	ampicillin	Overexpression of RNAP. T7P-α-β-β':BCCP; residues 71-156 of the biotin carboxyl carrier protein (BCCP) fused to the C-terminus of *rpoC*	IPTG	Laboratory plasmid
pET21-GreA	Ampicillin	Overexpression of 6X C-terminal His-GreA	IPTG	Koulich *et al.* 1997
pKP3077	Chloramphenicol	Overexpression of 6Xhis-tagged RelB	IPTG	(Pedersen *et al.* 2002)
pKP3067	Ampicillin	Overexpression of RelE	Arabinose	(Pedersen *et al.* 2002)
		6X his-tagged T7 RNAP		
pCAN-EF-G	Chloramphenicol	Overexpression of N terminal EF-G translation factor. T7 RNA polymerase promoter. cat. Obtained from ASKA collection	IPTG	(Kitagawa *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 *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 *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 *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 *et al.* 2005)
pCAN-MetRS	Chloramphenicol	Overexpression of N terminal metRS aminoacyl synthetase. T7 RNA polymerase promoter. cat. Obtained from ASKA collection	IPTG	(Kitagawa *et al.* 2005)
pCAN-FMT	Chloramphenicol	Overexpression of N terminal Formyl methionyl transferase. T7 RNA polymerase promoter. cat. Obtained from ASKA collection	IPTG	(Kitagawa *et al.* 2005)

Table 9.5 List of plasmids used in this work.

A5. Characterisation of MFVVVR+1 to +5 RNAs with template 3

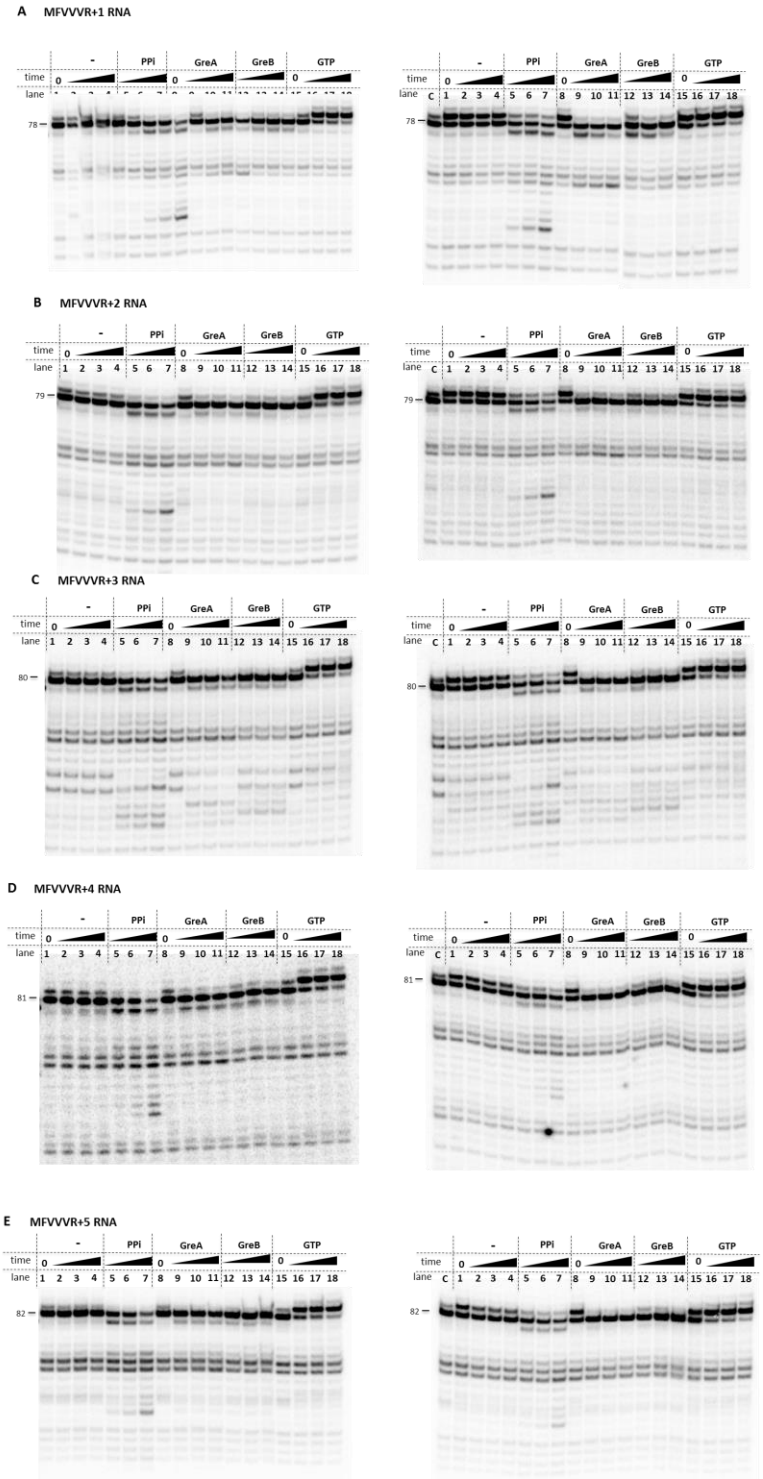


Figure 9.1 Characterisation of MFVVVR+1 to +5 RNAs with template 3. Pause characterisation of the RNAP on DNA template 3. The gel on the left shows the characterisation of the RNAP after transcription and washing of the AAECs and the gel on the right shows the characterisation after incubation at 37°C for 40 min in the presence of 100 µM final GTP concentration. Samples were taken at 0, 5, 10 and 30 min timepoints. Lanes 1, 8 and 15 contain the control samples taken before addition of buffer, PPi, GreA/B or GTP. The lane c in the incubated gel contains the sample taken before incubation with GTP (the same sample as in lanes 1, 8 and 15 on the left hand gel). All RNA is radiolabelled at the 5' end and the RNA sizes indicated.

A6. MFVVVR and RNAseH Digested RNA in transcription with extended hybrid templates

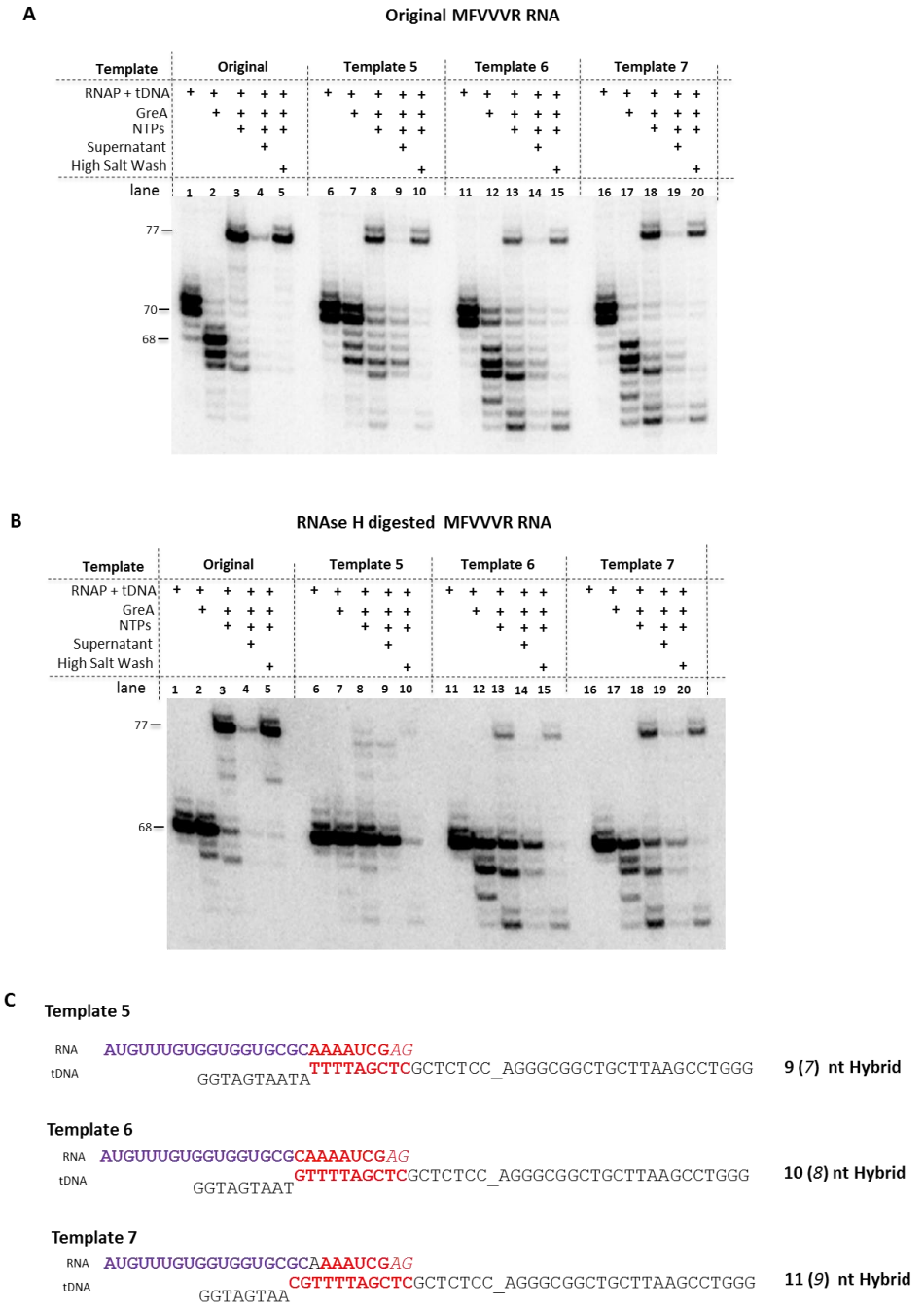


Figure 9.2 MFVVVR and RNAseH Digested RNA in transcription with extended hybrid templates. **A)** Transcription was carried out as standard with 3' end labelled MFVVVR RNA, templates 5-7 and the original abasic template for comparison. Samples were taken at each step of the transcription reaction and an equal volume of stop buffer added to stop the reaction. The samples were separated with 10% PAGE and visualised. **B)** Transcription was carried out as standard with 3' end labelled MFVVVR RNAseH digested RNA, templates 5-7 and the original abasic template for comparison. Samples were taken at each step of the transcription reaction and an equal volume of stop buffer added to stop the reaction. The samples were separated with 10% PAGE and visualised. **C)** The RNA and DNA sequences showing the extended hybrid (red) and peptide coding sequence (purple). The nts removed from the RNA 3'end during RNaseH digestion as indicated by italics.

A7. Characterisation of the extended hybrid templates

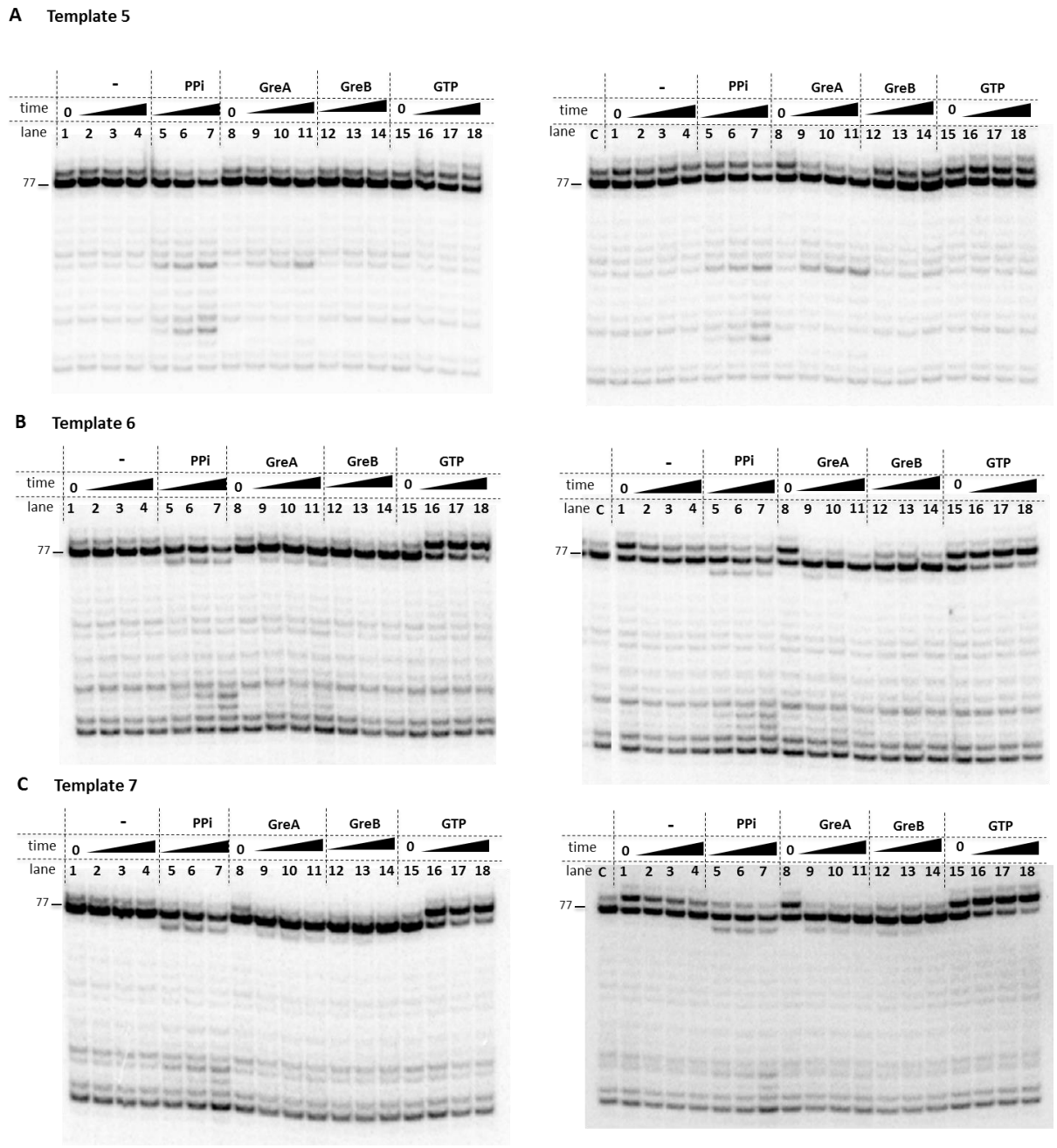


Figure 9.3 Characterisation of the extended hybrid templates. Pause characterisation of the RNAP on DNA templates 5 **(A)**, 6 **(B)** and 7 **(C)**. The gel on the left shows the characterisation of the RNAP after transcription and washing of the AAECs and the gel on the right shows the characterisation after incubation at 37°C for 40 min in the presence of 100 μM final GTP concentration. Samples were taken at 0, 5, 10 and 30 min timepoints. Lanes 1, 8 and 15 contain the control samples taken before addition of buffer, PPi, GreA/B or GTP. The lane c in the incubated gel contains the sample taken before incubation with GTP (the same sample as in lanes 1, 8 and 15 on the left hand gel). All RNA is radiolabelled at the 5' end and the RNA sizes indicated.

A8. Transcription with extended MFVVVR RNA (+1 to +5) and extended hybrid templates (5-7)

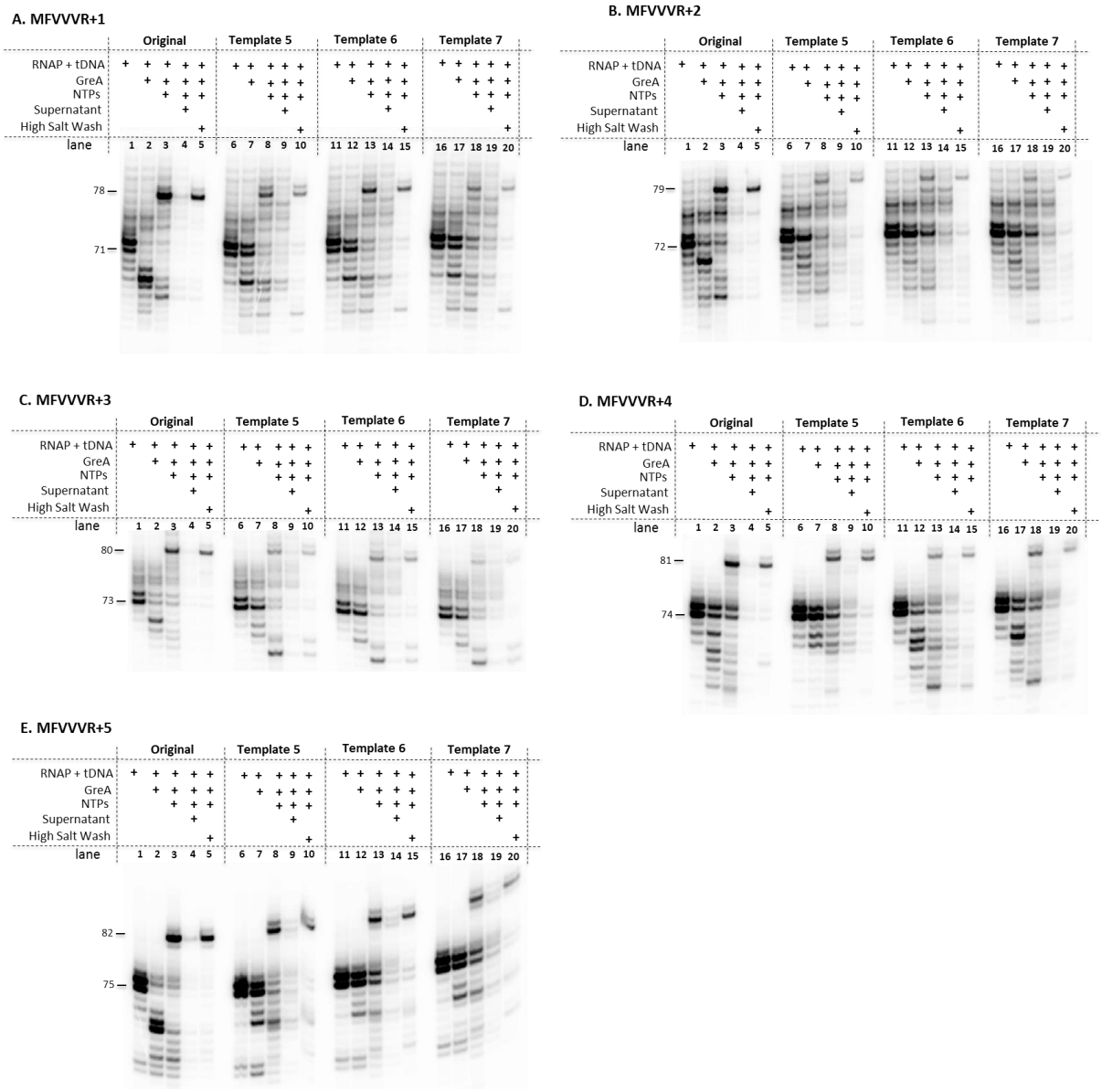


Figure 9.4 Transcription with extended MFVVVR RNA (+1 to +5) and extended hybrid templates (5-7). Transcription was carried out as standard using the 3' end labelled RNA indicated and DNA templates/non-templates. Samples were taken at each stage of transcription, an equal volume of transcription stop buffer added and the samples separated by denaturing 10% PAGE.

A9. Transcription of Template 8

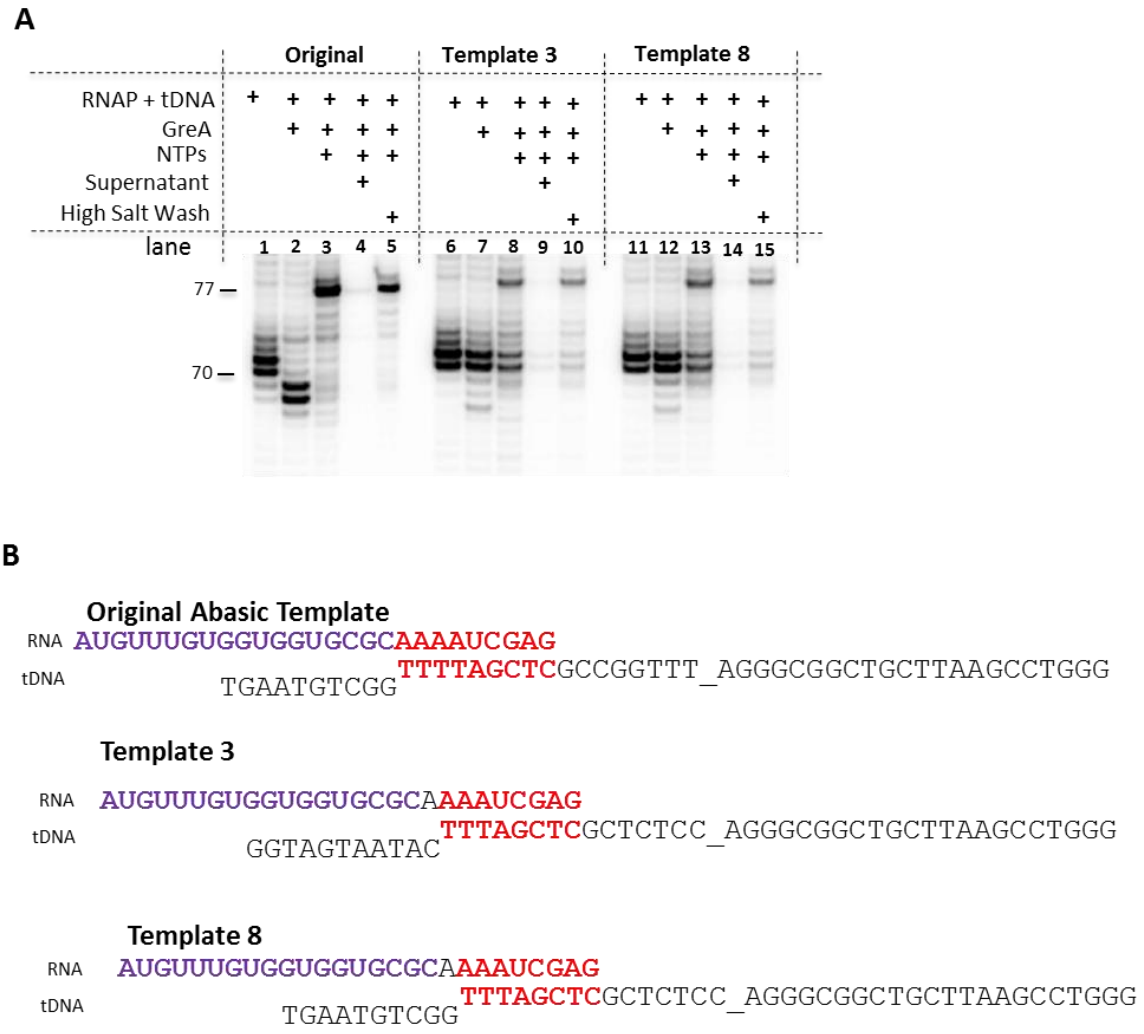
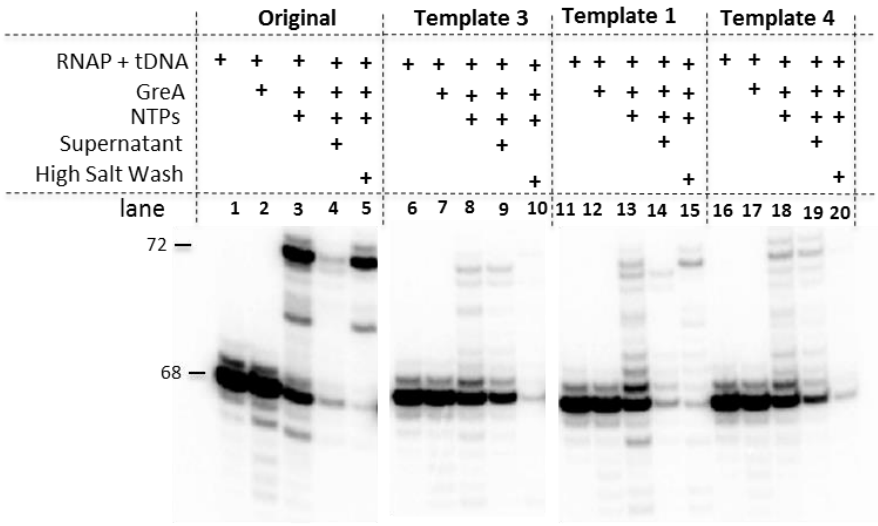


Figure 9.5 Transcription of template 8. Transcription was carried out as standard using either the abasic original DNA template, template 3 or template 8 and 3' end labelled MFVVVR RNA. Samples were taken at each stage of transcription and the reaction stopped by the addition of an equal volume of stop buffer. The samples were separated by 10% PAGE and visualised. B) DNA Template sequences with the peptide coding region (purple) and RNA:DNA hybrid region (red).

A10. Transcription with templates 1, 3 and 4 and MFVVVR RNAseH digested RNA.

A



B

Template 3

RNA AUGUUUGUGGUGGUGCGCAAAAUCG

tDNA GGTAGTAATAC TTTAGCTCGCTCTCC_AGGGCGGCTGCTTAAGCCTGGG

Template 4

RNA AUGUUUGUGGUGGUGCGCAAAAUCG

tDNA GGTAGTAATAC TTTAGCTCGCACACT_ACCGCGGCTGCTTAAGCCTGGG

Template 1

RNA AUGUUUGUGGUGGUGCGCAAAAUCG

tDNA GGTAGTAATAC TTTAGCTCCCGGTTT_ACCGCGGCTGCTTAAGCCTGGG

Figure 9.6 Transcription with templates 1, 3 and 4 and MFVVVR RNAseH digested RNA. A) Transcription was carried out as standard with 3' end labelled MFVVVR RNAseH digested RNA, templates 1, 3, 4 and the original abasic template for comparison. Samples were taken at each step of the transcription reaction and an equal volume of stop buffer added to stop the reaction. The samples were separated with 10% PAGE and visualised. B) The RNA and DNA sequences showing the shorter hybrid (red) and peptide coding sequence (purple).

A11. SDS-PAGE of the proteins purified during this project

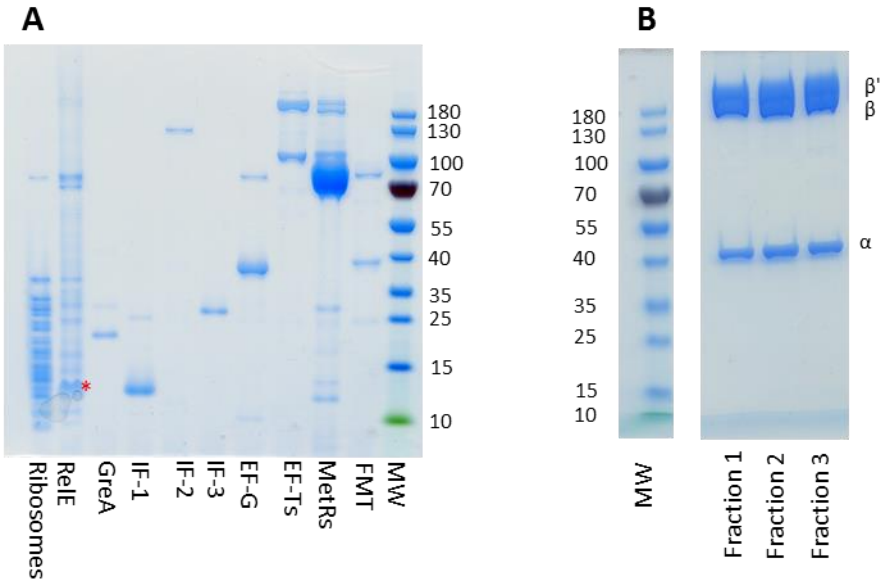


Figure 9.7 SDS-PAGE of the proteins purified during this work. A) 4-12% gradient SDS-PAGE of ribosomes, RelE, GreA, IFs-1, 2 and 3 EF-G, EF-Ts, MetRs, and FMT. 10-20 pmol of each protein was loaded except MetRs where approximately 50 pmol was loaded. MW indicates the molecular weight marker and the sizes (in kDa) are indicated to the left of the MW. RelE is indicated by the red asterisk. B) 4-20% gradient SDS-PAGE of biotin tagged core RNAP after ion exchange chromatography (Mono-Q). Fractions 2 and 3 were combined and dialysed after analysis by SDS-PAGE. MW indicates the molecular weight marker and the sizes (in kDa) are indicated to the right of the MW. The three fractions and MW were run on the same gel.

10. References

Adhya, S. & Gottesman, M., 1978. Control of Transcription Termination. *Annual review of biochemistry*, 47, pp.967–996.

Agrawal, R.K. *et al.*, 1999. EF-G-dependent GTP hydrolysis induces translocation accompanied by large conformational changes in the 70S ribosome. *Nature Structural Biology*, 6(7), pp.643–647.

Antoun, A. *et al.*, 2006. How initiation factors tune the rate of initiation of protein synthesis in bacteria. *The EMBO Journal*, 25(11), pp.2539–2550.

Antoun, A. *et al.*, 2003. The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis. *EMBO Journal*, 22(20), pp.5593–5601.

Arnaud-Barbe, N. *et al.*, 1998. Transcription of RNA templates by T7 RNA polymerase. *Nucleic acids research*, 26(15), pp.3550–3554.

Artsimovitch, I. & Landick, R., 1998. Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. *Genes and Development*, 12(19), pp.3110–3122.

Bar-Nahum, G. *et al.*, 2005. A ratchet mechanism of transcription elongation and its control. *Cell*, 120(2), pp.183–193.

Blaha, G., Stanley, R.E. & Steitz, T.A., 2009. Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. *Science*, 325(5943), pp.966–970.

Blank, A. *et al.*, 1986. An RNA polymerase mutant with reduced accuracy of chain elongation. *Biochemistry*, 25(20), pp.5920–8.

Bochkareva, A. *et al.*, 2012. Factor-independent transcription pausing caused by recognition of the RNA-DNA hybrid sequence. *The EMBO journal*, 31(3), pp.630–9.

Borukhov, S., Sagitov, V. & Goldfarb, A., 1993. Transcript cleavage factors from *E. coli*. *Cell*, 72(3), pp.459–466.

Boudvillain, M., Nollmann, M. & Margeat, E., 2010. Keeping up to speed with the transcription termination factor Rho motor. *Transcription*, 1(2), pp.70–75.

Brennan, C, Dombroski, J. & Platt, T., 1987. Transcription termination factor rho is an RNA-DNA helicase. *Cell*, 48(6), pp.945–52.

Burgess, R.R. *et al.*, 1969. Factor stimulating transcription by RNA polymerase. *Nature*, 221, pp.43–46.

Burgess, R.R., 1969. Separation and characterization of the subunits of ribonucleic acid polymerase. *The Journal of biological chemistry*, 244(22), pp.6168–6176.

Burmann, B.M., Schweimer, K., *et al.*, 2010. A NusE:NusG complex links transcription and translation. *Science*, 328(5977), pp.501–504.

Burmann, B.M., Luo, X., *et al.*, 2010. Fine tuning of the *E. coli* NusB:NusE complex affinity to BoxA RNA is required for processive antitermination. *Nucleic Acids Research*, 38(1), pp.314–326.

Cammack, K.A. & Wade, H.E., 1965. The sedimentation behaviour of ribonuclease-active and -inactive ribosomes from bacteria. *The Biochemical journal*, 96(3), pp.671–680.

Cardinale, C.J. *et al.*, 2008. Termination Factor Rho and Its Cofactors NusA and NusG Silence Foreign DNA in E.coli. *Science*, 320(5878), pp.935–8.

Castro-Roa, D. & Zenkin, N., 2012. In vitro experimental system for analysis of transcription-translation coupling. *Nucleic acids research*, 40(6), p.e45.

Castro-Roa, D. & Zenkin, N., 2015. Methodology for the analysis of transcription and translation in transcription-coupled-to-translation systems in vitro. *Methods*, 86, pp.51–59.

Cazenave, C. & Uhlenbeck, O.C., 1994. RNA template-directed RNA synthesis by T7 RNA polymerase. *Proceedings of the National Academy of Sciences*, 91(15), pp.6972–6976.

Chakraborty, A. *et al.*, 2012. Opening and Closing of the Bacterial RNA Polymerase Clamp. *Science*, 337(6094), pp.591–595.

Chan, C.L. & Landick, R., 1993. Dissection of the his Leader Pause Site by Base Substitution Reveals a Multipartite Signal that Includes a Pause RNA Hairpin. , pp.25–42.

Chan, C.L., Wang, D. & Landick, R., 1997. Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3’ end spacing distinguishes pause and termination pathways. *Journal of molecular biology*, 268(1), pp.54–68.

Churchman, L.S. & Weissman, J.S., 2011. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature*, 469(7330), pp.368–73.

Clarke, S.J., Low, B. & Konigsberg, W., 1973. Isolation and Characterization of an Endonuclease from Escherichia coli Specific for Riobonucleic Acid in Ribonucleic Acid-Deoxyribonucleic Acid Hybrid Structures. *Journal of Biological Chemistry*, 113, pp.1096–1103.

Daube, S.S. *et al.*, 1992. Functional transcription elongation complexes from

synthetic RNA-DNA bubble duplexes. *Science*, 258(5086), pp.1320–1324.

Doerfel, L.K. *et al.*, 2013. EF-P Is Essential for Rapid Synthesis of Proteins Containing Consecutive Proline Residues. *Science*, 339(6115), pp.85–88.

Ederth, J. *et al.*, 2009. A single-step method for purification of active His-tagged ribosomes from a genetically engineered *Escherichia coli*. *Nucleic Acids Research*, 37(2).

Elgamal, S., Artsimovitch, I. & Ibba, M., 2016. Maintenance of Transcription-Translation Coupling by Elongation Factor P. *mBio*, 7(5), pp.1–11.

Epshtein, V. *et al.*, 2010. An allosteric mechanism of Rho-dependent transcription termination. *Nature*, 463(7278), pp.245–9.

Epshtein, V. *et al.*, 2003. Transcription through the roadblocks: the role of RNA polymerase cooperation. *The EMBO journal*, 22(18), pp.4719–27.

Epshtein, V. *et al.*, 2014. UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature*, 505(7483), pp.372–7.

Epshtein, V. & Nudler, E., 2003. Cooperation between RNA polymerase molecules in transcription elongation. *Science*, 300(5620), pp.801–805.

Erie, D. *et al.*, 1993. Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. *Science*, 262(5135), pp.867–73.

Ermolenko, D.N. & Noller, H.F., 2011. mRNA translocation occurs during the second step of ribosomal intersubunit rotation. *Nature structural & molecular biology*, 18(4), pp.457–62.

Fan, J. *et al.*, 2016. Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. *Nature*, 536(7615), pp.234–237.

Finn, R.D. *et al.*, 2000. *Escherichia coli* RNA polymerase core and holoenzyme structures. *EMBO Journal*, 19(24), pp.6833–6844.

Frank, J. & Agrawal, R.K., 2000. A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature*, 406(6793), pp.318–322.

Friedman, L.J., Mumm, J.P. & Gelles, J., 2013. RNA polymerase approaches its promoter without long-range sliding along DNA. *Proceedings of the National Academy of Sciences*, 110(24), pp.9740–5.

Ghosh, P., Ishihama, A. & Chatterji, D., 2001. *Escherichia coli* RNA polymerase subunit ω and its N-terminal domain bind full-length β′ to facilitate incorporation into the $\alpha_2\beta$ subassembly. *European Journal of Biochemistry*, 268(17), pp.4621–4627.

Giglione, C., Boularot, a & Meinnel, T., 2004. Protein N-terminal methionine excision. *Cellular and molecular life sciences : CMLS*, 61(12), pp.1455–1474.

Glick, B.R. & Ganoza, M.C., 1975. Identification of a soluble protein that stimulates peptide bond synthesis. *Proceedings of the National Academy of Sciences*, 72(11), pp.4257–4260.

Goldman, S.R. *et al.*, 2015. The primary σ factor in *Escherichia coli* can access the transcription elongation complex from solution *in vivo*. *eLife*, 4, pp.1–17.

Grachev, M.A. & Zaychikov, E.F., 1980. Initiation by Escherichia coli RNA-polymerase: transformation of abortive to productive complex. *FEBS Letters*, 115(1), pp.23–26.

Green, R. & Noller, H.F., 1997. Ribosomes and translation. *Annual review of biochemistry*, 66, pp.679–716.

Gruber, T.M. & Gross, C.A., 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual review of microbiology*, 57(1), pp.441–66.

Gusarov, I. & Nudler, E., 1999. The mechanism of intrinsic transcription termination. *Molecular Cell*, 3(4), pp.495–504.

Haines, N.M. *et al.*, 2014. Stalled transcription complexes promote DNA repair at a distance. *Proceedings of the National Academy of Sciences*, 111(11), pp.4037–42.

Hamming, J. *et al.*, 1981. Electron microscopic analysis of transcription of a ribosomal RNA operon of *E. coli*. *Nucleic acids research*, 9(6), pp.1339–1350.

Harden, T.T. *et al.*, 2016. Bacterial RNA polymerase can retain σ 70 throughout transcription. *Proceedings of the National Academy of Sciences*, 113(3), pp.602–607.

Hauryliuk, V. *et al.*, 2009. Thermodynamics of GTP and GDP Binding to Bacterial Initiation Factor 2 Suggests Two Types of Structural Transitions. *Journal of Molecular Biology*, 394(4), pp.621–626.

Hayes, C.S. & Sauer, R.T., 2003. Cleavage of the A site mRNA codon during ribosome pausing provides a mechanism for translational quality control. *Molecular Cell*, 12(4), pp.903–911.

Holtkamp, W. *et al.*, 2014. GTP hydrolysis by EF-G synchronizes tRNA movement on small and large ribosomal subunits. *EMBO Journal*, 33(9), pp.1073–1085.

Hüttenhofer, A. & Noller, H.F., 1994. Footprinting mRNA-ribosome complexes with chemical probes. *The EMBO journal*, 13(16), pp.3892–901.

Igarashi, K., Fujita, N. & Ishihama, A., 1989. Promoter selectivity of *Escherichia coli* RNA polymerase: Omega factor is responsible for the ppGpp sensitivity. *Nucleic Acids Research*, 17(21), pp.8755–8765.

Imashimizu, M. *et al.*, 2015. Visualizing translocation dynamics and nascent transcript errors in paused RNA polymerases in vivo. *Genome Biology*, 16(1), pp.1–17.

Inoue, H. *et al.*, 1987. Sequence-dependent hydrolysis of RNA using modified oligonucleotide splints and RNase H. *FEBS Letters*, 215(2), pp.327–330.

Ishihama, A., 1981. Subunit of assembly of *Escherichia coli* RNA polymerase. *Advances in biophysics*, 14, p.1—35.

Jacob, W.F., Santer, M. & Dahlberg, E., 1987. A single base change in the Shine-Dalgarno region of 16S rRNA of *Escherichia coli* affects translation of many proteins. *Proceedings of the National Academy of Sciences*, 84(14), pp.4757–4761.

Jin, D.J. *et al.*, 1992. Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. *Proceedings of the National Academy of Sciences*, 89(4), pp.1453–1457.

Kamarthapu, V. & Nudler, E., 2015. Rethinking transcription coupled DNA repair. *Current Opinion in Microbiology*, 24, pp.15–20.

Kapanidis, A.N. *et al.*, 2006. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*, 314(5802), pp.1144–1147.

Katoh, T. *et al.*, 2016. Essential structural elements in tRNAPro for EF-P-mediated alleviation of translation stalling. *Nature Communications*, 7, p.11657.

Kisker, C., Kuper, J. & Van Houten, B., 2013. Prokaryotic nucleotide excision repair. *Cold Spring Harb Perspect Biol*, 5(3), pp.1–18.

Kitagawa, M. *et al.*, 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 research*, 12(5), pp.291–9.

Kolb, K.E., Hein, P.P. & Landick, R., 2014. Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH. *Journal of Biological Chemistry*, 289(2), pp.1151–1163.

Komissarova, N. & Kashlev, M., 1998. Functional topography of nascent RNA in elongation intermediates of RNA polymerase. *Proceedings of the National*

Academy of Sciences, 95(25), pp.14699–704.

Komissarova, N. & Kashlev, M., 1997a. RNA polymerase switches between inactivated and activated states By translocating back and forth along the DNA and the RNA. *The Journal of Biological Chemistry*, 272(24), pp.15329–15338.

Komissarova, N. & Kashlev, M., 1997b. Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proceedings of the National Academy of Sciences*, 94(5), pp.1755–60.

Koulich, D. *et al.*, 1997. Domain organization of Escherichia coli transcript cleavage factors GreA and GreB. *Journal of Biological Chemistry*, 272(11), pp.7201–7210.

Landick, R., Carey, J. & Yanofsky, C., 1985. Translation activates the paused transcription complex and restores transcription of the trp operon leader region. *Proceedings of the National Academy of Sciences*, 82(14), pp.4663–4667.

Lapham, J. *et al.*, 1997. The position of site-directed cleavage of RNA using RNaseH and 2'-O-methyl oligonucleotides is dependent on the enzyme source. *RNA*, 3, pp.950–951.

Lapham, J. & Crothers, D.M., 1996. RNase H cleavage for processing of in vitro transcribed RNA for NMR studies and RNA ligation. *RNA*, 2(3), pp.289–296.

Laptenko, O. *et al.*, 2003. Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase. *EMBO Journal*, 22(23), pp.6322–6334.

Larson, M.H. *et al.*, 2014. A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science*, 344(6187), pp.1042–7.

Ledoux, S., Olejniczak, M. & Uhlenbeck, O.C., 2009. A sequence element that tunes Escherichia coli tRNA(Ala)(GGC) to ensure accurate decoding. *Nature structural & molecular biology*, 16(4), pp.359–64.

Luo, X. *et al.*, 2008. Structural and Functional Analysis of the *E. coli* NusB-S10 Transcription Antitermination Complex. *Mol. Cell*, 32(6), pp.791–802.

Marshall, R.A., Aitken, C.E. & Puglisi, J.D., 2009. GTP Hydrolysis by IF2 Guides Progression of the Ribosome into Elongation. *Molecular Cell*, 35(1), pp.37–47.

Mason, S.W., Li, J. & Greenblatt, J., 1992. Direct interaction between two *Escherichia coli* transcription antitermination factors, NusB and ribosomal protein S10. *Journal of molecular biology*, 223(1), pp.55–66.

McGary, K. & Nudler, E., 2013. RNA polymerase and the ribosome: the close relationship. *Current opinion in microbiology*, 16(2), pp.112–117.

Melnikov, S. *et al.*, 2016. Molecular insights into protein synthesis with proline

residues. , pp.1–9.

Miller, O.L., Hamkalo, B. a & Thomas, C. a, 1970. Visualization of bacterial genes in action. *Science*, 169(943), pp.392–395.

Moazed, D. & Noller, H.F., 1989. Intermediate states in the movement of transfer RNA in the ribosome. *Nature*, 342(6246), pp.142–8.

Mora, L. *et al.*, 2003. Stop codon recognition and interactions with peptide release factor RF3 of truncated and chimeric RF1 and RF2 from Escherichia coli. *Molecular Microbiology*, 50(5), pp.1467–1476.

Mukhopadhyay, J. *et al.*, 2004. Antibacterial peptide Microcin J25 inhibits transcription by binding within and obstructing the RNA polymerase secondary channel. *Molecular Cell*, 14(6), pp.739–751.

Munson, L.M. & Reznikoff, W.S., 1981. Abortive initiation and long ribonucleic acid synthesis. *Biochemistry*, 20(8), pp.2081–2085.

Murakami, K. *et al.*, 1997. The two alpha subunits of *Escherichia coli* RNA polymerase are asymmetrically arranged and contact different halves of the DNA upstream element. *Proceedings of the National Academy of Sciences*, 94(5), pp.1709–1714.

Muth, G.W., Ortoleva-Donnelly, L. & Strobel, S.A., 2000. A Single Adenosine with Neutral pKa in the Ribosomal Peptidyl Transferase Center. *Science*, 289(5481), pp.947–950.

Nacheva, G.A. & Berzal-Herranz, A., 2003. Preventing nondesired RNA-primed RNA extension catalyzed by T7 RNA polymerase. *European Journal of Biochemistry*, 270(7), pp.1458–1465.

Neubauer, C. *et al.*, 2009. The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE. *Cell*, 139(6), pp.1084–95.

Neuman, K.C. *et al.*, 2003. Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell*, 115(4), pp.437–47.

Nickels, B.E. *et al.*, 2005. The interaction between sigma70 and the beta-flap of *Escherichia coli* RNA polymerase inhibits extension of nascent RNA during early elongation. *Proceedings of the National Academy of Sciences*, 102(12), pp.4488–93.

Nudler, E., 2009. RNA polymerase active center: the molecular engine of transcription. *Annual Review of Biochemistry*, 78(1), pp.335–361.

Nudler, E. *et al.*, 1997. The RNA-DNA hybrid maintains the register of transcription

by preventing backtracking of RNA polymerase. *Cell.*, 89(1), pp.33–41.

Ogle, J.M. *et al.*, 2001. Recognition of Cognate Transfer RNA by the 30S Ribosomal Subunit. *Science*, 292(5518), pp.897–902.

Orlova, M. *et al.*, 1995. Intrinsic transcript cleavage activity of RNA polymerase. *Proceedings of the National Academy of Sciences*, 92(10), pp.4596–4600.

Paci, M., Pon, C. & Gualerzi, C., 1985. The interaction between initiation factor 3 and 30 S ribosomal subunits studied by high-resolution 1H NMR spectroscopy. *The Journal of biological chemistry*, 260(2), pp.887–92.

Paget, M.S.B. & Helmann, J.D., 2003. The sigma70 family of sigma factors. *Genome biology*, 4(1), p.203.

Park, J.S., Marr, M.T. & Roberts, J.W., 2002. *E. coli* transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell*, 109(6), pp.757–767.

Paul, B.J. *et al.*, 2004. rRNA transcription in Escherichia coli. *Annual review of genetics*, 38(20), pp.749–770.

Pedersen, K. *et al.*, 2003. The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell*, 112(1), pp.131–140.

Pedersen, K., Christensen, S.K. & Gerdes, K., 2002. Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Molecular Microbiology*, 45(2), pp.501–510.

Peters, J.M. *et al.*, 2012. Rho and NusG suppress pervasive antisense transcription in Escherichia coli. *Genes & Development*, 26(23), pp.2621–2633.

Peters, J.M., Vangeloff, A.D. & Landick, R., 2011. Bacterial transcription terminators: the RNA 3’-end chronicles. *Journal of molecular biology*, 412(5), pp.793–813.

Pomerantz, R.T. & O’Donnell, M., 2008. The replisome uses mRNA as a primer after colliding with RNA polymerase. *Nature*, 456(7223), pp.762–766.

Pomerantz, R.T. & O’Donnell, M., 2010. What happens when replication and transcription complexes collide? *Cell Cycle*, 9(13), pp.2537–2543.

Ramakrishnan, V. *et al.*, 2002. Ribosome structure and the mechanism of translation. *Cell*, 108(4), pp.557–72.

Ramrath, D.J.F. *et al.*, 2013. Visualization of two transfer RNAs trapped in transit during elongation factor G-mediated translocation. *Proceedings of the National Academy of Sciences*, 110(52), pp.20964–20969.

Rees, W.A. *et al.*, 1993. Evidence of DNA Bending in Transcription Complexes

Imaged by Scanning Force Microscopy. *Science*, 260(5114), pp.1646–1649.

Revyakin, A. *et al.*, 2006. Abortive Initiation and Productive Initiation by RNA Polymerase Involve DNA Scrunching. *Science*, 314(5802), pp.1139–1143.

Richardson, J.P., 2002. Rho-dependent termination and ATPases in transcript termination. *Biochimica et Biophysica Acta - Gene Structure and Expression*, 1577(2), pp.251–260.

Richardson, J.P., Grimley, C. & Lowery, C., 1975. Transcription termination factor rho activity is altered in *Escherichia coli* with *suA* gene mutations. *Proceedings of the National Academy of Sciences*, 72(5), pp.1725–8.

Ring, B.Z., Yarnell, W.S. & Roberts, J.W., 1996. Function of *E. coli* RNA Polymerase σ Factor σ^{70} in Promoter-Proximal Pausing. *Cell*, 86(3), pp.485–493.

Rodnina, M. V *et al.*, 1997. Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature*, 385(6611), pp.37–41.

Saitoh, T. & Ishihama, A., 1976. Subunits of RNA polymerase in function and structure. *Journal of Molecular Biology*, 104(3), pp.621–635.

Schmeing, T.M. *et al.*, 2009. The Crystal Structure of the Ribosome Bound to EF-Tu and Aminoacyl-tRNA. *Science*, 326(5953), pp.688–695.

Schmeing, T.M. & Ramakrishnan, V., 2009. What recent ribosome structures have revealed about the mechanism of translation. *Nature*, 461(7268), pp.1234–1242.

Schweimer, K. *et al.*, 2010. Cooperation Between Translating Ribosomes and RNA Polymerase in Transcription Elongation. *Science*, 328(504), pp.504–508.

Sedlyarova, N. *et al.*, 2016. sRNA-Mediated Control of Transcription Termination in E. coli. *Cell*, 167(1), p.111–121.e13.

Selmer, M. *et al.*, 2006. Structure of the 70S Ribosome Complexed with mRNA and tRNA. *Science*, 313(5795), pp.1935–1942.

Severinov, K. *et al.*, 1995. Assembly of functional *Escherichia coli* RNA polymerase containing beta subunit fragments. *Proceedings of the National Academy of Sciences*, 92(10), pp.4591–5.

Shine, J. & Dalgarno, L., 1974. The 3’-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America*, 71(4), pp.1342–6.

Sidorenkov, I., Komissarova, N. & Kashlev, M., 1998. Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Molecular cell*, 2(1), pp.55–64.

Siebenlist, U., 1979. RNA polymerase undiwnds an 11-base pair segment of a phage T7 promoter. *Nature*, 279, pp.651–652.

Sosunov, V. *et al.*, 2005. The involvement of the aspartate triad of the active center in all catalytic activities of multisubunit RNA polymerase. *Nucleic Acids Research*, 33(13), pp.4202–4211.

Sosunov, V. *et al.*, 2003. Unifed two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO Journal*, 22(9), pp.2234–2244.

Steitz, T. A, 2008. A structural understanding of the dynamic ribosome machine. *Nature Reviews Molecular Cell Biology*, 9(3), pp.242–253.

Strauß, M. *et al.,* 2016. Transcription is regulated by NusA:NusG interaction. *Nucleic Acids Research*, 44(12), pp.5971–5982.

Strobel, E.J. & Roberts, J.W., 2015. Two transcription pause elements underlie a σ 70 -dependent pause cycle. *Proceedings of the National Academy of Sciences*, 112(32), pp.4374–4380.

Suh, W.C., Ross, W. & Record, M.T., 1993. Two open complexes and a requirement for Mg2+ to open the lambda PR transcription start site. *Science*, 259(5093), pp.358–61.

Svetlov, V. & Artsimovitch, I., 2015. Purification of Bacterial RNA Polymerase: Tools and Protocols. *Bacterial Transcriptional Control : Methods and Protocols*, 1276, p.13–29c.

Triana-Alonso, F.J. *et al.*, 1995. Self-coded 3'-extension of run-off transcripts produces aberrant products during in vitro transcription with T7 RNA polymerase. *Journal of Biological Chemistry*, 270(11), pp.6298–6307.

Ude, S. *et al.*, 2013. Translation Elongation Factor EF-P Alleviates Ribosome Stalling at Polyproline Stretches. *Science Translational Medicine*, 339(6115), pp.82–85.

Valle, M. *et al.*, 2003. Locking and unlocking of ribosomal motions. *Cell*, 114(1), pp.123–134.

Varshney, U., Lee, C.P. & RajBhandary, U.L., 1991. Direct analysis of aminoacylation levels of tRNAs in Vivo. *Journal of Biological Chemistry*, 266(36), pp.24712–24718.

Vassylyev, D.G. *et al.*, 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature*, 417(6890), pp.712–719.

Vassylyev, D.G. *et al.*, 2007. Structural basis for substrate loading in bacterial RNA polymerase. *Nature*, 448(7150), pp.163–168.

Vogel, U. & Jensen, K., 1994. The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *Journal of Bacteriology*, 176(10), pp.2807–2813.

Vvedenskaya, I.O. *et al.*, 2014. Transcription. Interactions between RNA polymerase and the 'core recognition element' counteract pausing. *Science*, 344(1285), pp.1285–1289.

Walter, G. *et al.*, 1967. Initiation of DNA-dependent RNA synthesis and the effect of heparin on RNA polymerase. *European journal of biochemistry / FEBS*, 3(2), pp.194–201.

Wang, D. *et al.*, 2006. Structural Basis of Transcription: Role of the Trigger Loop in Substrate Specificity and Catalysis. *Cell*, 127(5), pp.941–954.

Wang, F. *et al.*, 2013. The promoter-search mechanism of *Escherichia coli* RNA polymerase is dominated by three-dimensional diffusion. *Nature structural & molecular biology*, 20(2), pp.174–181.

Washburn, R.S. & Gottesman, M.E., 2011. Transcription termination maintains chromosome integrity. *Proceedings of the National Academy of Sciences*, 108(2), pp.792–797.

Weixlbaumer, A. *et al.*, 2008. Insights into translational termination from the structure of RF2 bound to the ribosome. *Science*, 322(5903), pp.953–956.

Weixlbaumer, A. *et al.*, 2013. Structural basis of transcriptional pausing in bacteria. *Cell*, 152(3), pp.431–441.

Wilden, B. *et al.*, 2006. Role and timing of GTP binding and hydrolysis during EF-G-dependent tRNA translocation on the ribosome. *Proceedings of the National Academy of Sciences*, 103(37), pp.13670–13675.

Wimberly, B.T. *et al.*, 2000. Structure of the 30S ribosomal subunit. *Nature*, 407, pp.327–339.

Woolstenhulme, C.J. *et al.*, 2013. Nascent peptides that block protein synthesis in bacteria. *Proceedings of the National Academy of Sciences*, 110(10), pp.E878–E887.

Wu, C.W. & Goldthwait, D.A., 1969. Studies of nucleotide binding to the ribonucleic acid polymerase by equilibrium dialysis. *Biochemistry*, 8(11), pp.4458–4464.

Yang, W., Lee, J.Y. & Nowotny, M., 2006. Making and Breaking Nucleic Acids: Two-Mg2+-Ion Catalysis and Substrate Specificity. *Molecular Cell*, 22(1), pp.5–13.

Yang, X. & Lewis, P.J., 2010. The interaction between RNA polymerase and the elongation factor NusA. *RNA biology*, 7(3), pp.272–275.

Yokogawa, T. *et al.*, 2010. Optimization of the hybridization-based method for purification of thermostable tRNAs in the presence of tetraalkylammonium salts. *Nucleic Acids Research*, 38(6), p.e89.

Yu, C.H. and Y.-T., 2012. Synthesis and Labelling of RNA In Vitro. *Current Protocols in Molecular Biology*, 29(6), pp.997–1003.

Yusupov, M.M. *et al.*, 2001. Crystal structure of the ribosome at 5.5 A resolution. *Science*, 292(5518), pp.883–896.

Yuzenkova, Y. *et al.*, 2014. Control of transcription elongation by GreA determines rate of gene expression in *Streptococcus pneumoniae*. *Nucleic acids research*, 42(17), pp.10987–99.

Yuzenkova, Y. *et al.*, 2010. Stepwise mechanism for transcription fidelity. *BMC biology*, 8, p.54.

Yuzenkova, Y. & Zenkin, N., 2010. Central role of the RNA polymerase trigger loop in intrinsic RNA hydrolysis. *Proceedings of the National Academy of Sciences*, 107(24), pp.10878–10883.

Zaychikov, E. *et al.*, 1996. Mapping of catalytic residues in the RNA polymerase active center. *Science*, 273(5271), pp.107–109.

Zaychikov, E., Denissova, L. & Heumann, H., 1995. Translocation of the *Escherichia coli* transcription complex observed in the registers 11 to 20: 'jumping' of RNA polymerase and asymmetric expansion and contraction of the 'transcription bubble'. *Proceedings of the National Academy of Sciences*, 92(5), pp.1739–1743.

Zenkin, N., Yuzenkova, Y. & Severinov, K., 2006. Transcript-assisted transcriptional proofreading. *Science*, 313(5786), pp.518–520.

Zhang, G. *et al.*, 1999. Crystal Structure of *Thermus aquaticus* Core RNA Polymerase at 3.3 A Resolution. *Cell*, 98, pp.811–824.

Zhang, J., Palangat, M. & Landick, R., 2010. Role of the RNA polymerase trigger loop in catalysis and pausing. *Nature structural & molecular biology*, 17(1), pp.99–104.