Interconnections between transcription and pre-mRNA splicing

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Oksana Gonchar

Department of Molecular and Cell Biology

University of Leicester

2016

ABSTRACT

Interconnections between transcription and pre-mRNA splicing − **Oksana Gonchar**

Eukaryotic gene expression involves many processes some of which are transcription and pre-mRNA splicing. It has been shown that the majority of splicing events are functionally or physically coupled to RNA polymerase II transcription machinery. This suggests that transcription and splicing processes might influence one another. For example, a number of studies have implicated snRNPs, in particular U1 snRNP, in transcription initiation and elongation. In this work, interconnections between transcription and splicing were tested using an *in vitro* RNA polymerase II transcription/splicing assays (Pol II-TSRs). The results obtained in this study showed that inhibition of U1 but not U2 or U6 snRNPs led to a major reduction in transcript levels using different DNA templates. However, interference with initiation was excluded because it was found that this effect was the result of reduced RNA stability. Moreover, similar results were observed both with transcription by T7 RNA polymerase and with purified transcripts added to the extract. These results allow to conclude that the U1 snRNP has a novel function in protecting RNA from degradation. Further investigations showed that the RNA is protected by the U1 snRNP against 5' exonucleases and 3' exonucleases and possibly endonucleases. It was found that the presence of 5' splice site (5'SS) is necessary for RNA stability. These data suggest that the U1 snRNP through a direct interaction with the 5'SS of the pre-mRNA protects it from degradation and only U1 snRNP but not the active spliceosome is required for RNA stability. It was observed that under splicing conditions, the RNA level of a transcript lacking the intronic consensus 5'SS but having it at the 3' end was significantly higher compared to that for a transcript lacking any consensus 5'SSs, suggesting that the U1 snRNP protects transcripts from degradation along their length. It is proposed that rapid binding of the U1 snRNPs to the nascent transcripts induces the assembly of RNA binding proteins that protect the RNA. Interestingly, any transcripts tested regardless of the presence of the consensus 5'SSs were stable in the reaction with ongoing transcription, suggesting that Pol II transcription has an additional effect on RNA stability.

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my supervisors, Professor Ian C Eperon and Dr Cyril Dominguez, for their guidance and support throughout the duration of my PhD. My sincere thanks go to my thesis committee member, Dr Olga Makarova, for her insightful comments and encouragements. I thank all members of Lab 3/19, past and present, for their help advice and friendship. Also, I would like to thank Dr Dmitry Cherny for supportive scientific discussions.

I owe much thankfulness to my parents for their endless love and encouragements. Finally, I express my unfailing gratitude to my husband, Dr Andrey Morozov, who has supported and constantly encouraged me throughout my PhD, without whom it would have not been possible.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 3

LIST OF TABLES

LIST OF ABBREVIATIONS

A – Adenosine

AMPS – Ammonium persulfate

AMO – Antisense morpholino oligonucleotides

AS – Alternative splicing

ATP – Adenosine-5'-triphosphate

bp – Base pair

BPS – Branch point sequence

CDK – Cyclin-dependent kinases

CrP – Creatine phosphate

CTD – Carboxy-terminal domain

CTP – Cytidine 5γ-triphosphate

DMEM – Dulbecco's modified Eagle's medium

DMSO – Dimethyl sulfoxide

DNA – Deoxyribonucleic acid

dNTPs – Deoxyribonucleotide triphosphates

E. coli – Escherichia coli

EDTA – Ethylenediaminetetraacetic acid

eGFP – Enhanced green fluorescent protein

Fmol – Femtomole

ESE – Exonic splicing enhancer

ESS – Exonic splicing silencer

FBS – Fetal bovine serum

GTFs – general transcription factors

GFP – Green fluorescent protein

GTP – Guanosine-5'-triphosphate

h – Hour

HEK 293T – Human embryonic kidney 293T

Hepes – N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid

hnRNP – Heterogeneous nuclear ribonucleoprotein

ISE – Intronic splicing enhancer

ISS – Intronic splicing silencer

kDa – Kilodalton

KGlu – Potassium glutamate

l – Litre

LMP – Low melting point

µ – Micro

m – Milli

min – Minute

mL – Millilitre

MMLV-RT – Moloney Murine Leukemia Virus reverse transcriptase

M – Molar

mRNA – Messenger RNA

NE – Nuclear extract

nM – Nanomolar

NMD – Nonsense-mediated Decay

NP-40 – Nonidet P-40

nt – Nucleotide

Oligo – Oligonucleotide

O/N – Overnight

PAGE – Polyacrylamide gel electrophoresis

PAS – Polyadenylation signal

PBS – Phosphate buffered saline

PCPA – Prematurely cleaved and polyadenylated

PCR – Polymerase chain reaction

PIC – Pre-initiation complex

PK – Proteinase K

pmol – Picomole

PNACL – Protein Nucleic Acid Chemistry Laboratory

Pol I, II and III– RNA polymerase I, II and III

pre-mRNA – Precursor messenger RNA

Pro – Proline

PPT – Polypyrimidine tract

PTB – Polypyrimidine tract binding protein

P-TEFb – Positive transcription elongation factor

RNA – Ribonucleic acid

rRNA – Ribosomal RNA

RNase – Ribonuclease

rNTPs – Ribonucleotide triphosphates

RRM – RNA recognition motif

RT – Reverse transcription

RON – Receptor d'origine nantais

Sam68 – Src-associated in mitosis 68 kDa protein

SDS – Sodium dodecyl sulfate

sec – second(s)

Ser – Serine

SL – Stem loop

SMN – Survival of motor neuron

snRNA – Small nuclear ribonucleic acid

snRNP – Small nuclear ribonucleoprotein

SR – Splicing reaction

SRSF – Serine/Arginine-rich splicing factor

SS – Splice site

T7 Pol – T7 RNA polymerase

TBP – TATA-binding protein

TBE – Tris-borate-EDTA buffer

TEMED – N,N,N',N'-tetramethylethylenediamine

Thr – Threonine

tRNA – Transfer RNA

TSR- Transcription/splicing reaction

TSS – Transcription start site

Tyr – Tyrosine

U2AF – U2 snRNP auxiliary factor

U snRNP – Uridine-rich small nuclear ribonucleoprotein particle

UTP – Uridine-5'-triphosphate

V – Volts

v/v – Volume per volume

W – Watts

1 INTRODUCTION

1.1 Expression of eukaryotic genes

1.2 Transcription

1.3 Capping of pre-mRNA

1.4 Pre-mRNA splicing

1.5 U1 snRNP

1.6 Mechanisms of RNA degradation

1.7 Aims

1.1 Expression of eukaryotic genes

Genes comprise deoxyribonucleic acid (DNA), the sequence of which is a source of genetic information. The production of proteins from genes is not straightforward and requires many intermediate steps (Figure 1.1). The flow of information from DNA is recorded into precursor messenger ribonucleic acid (pre-mRNA) during transcription which is carried out in eukaryotes by RNA polymerase II (Pol II). This is followed by the maturation of pre-mRNA by processing factors, leading to messenger RNA (mRNA) production. The pre-mRNA undergoes 5' end capping, splicing and 3'-end formation. During capping a guanine nucleotide is added to the 5' end of the pre-mRNA to protect it from degradation. Transcription produces pre-mRNA containing exon and intron sequences. The exon is a sequence that remains in the mature mRNA, whereas an intron is a sequence located between exons that gets removed. Pre-mRNA splicing removes introns and joins the exons together. Therefore, only exons are retained in mature mRNA. After Pol II has reached the end of a gene, the nascent transcript undergoes cleavage at the 3' end which is followed by synthesis of a poly(A) tail during a process termed polyadenylation.

Initially, it was assumed that transcription and pre-mRNA processing were independent processes (Figure 1.1) because *in vitro* RNA processing could occur in isolation and in the absence of transcription. However, it is now well established that RNA processing events are linked, initiated and coordinated co-transcriptionally (reviewed by Bentley, 2014). The coupling of transcription and pre-mRNA processing can be mechanical, when the processing complexes are attached to the transcription apparatus. Functional coupling also takes place, when RNA processing factors interact with the transcription machinery in the vicinity of the nascent transcript when RNA exits from Pol II (Pandya-Jones and Black, 2009). These tight interactions serve as an important timing mechanism necessary to protect transcripts from degradation (reviewed by de Almeida and Carmo-Fonseca, 2008).

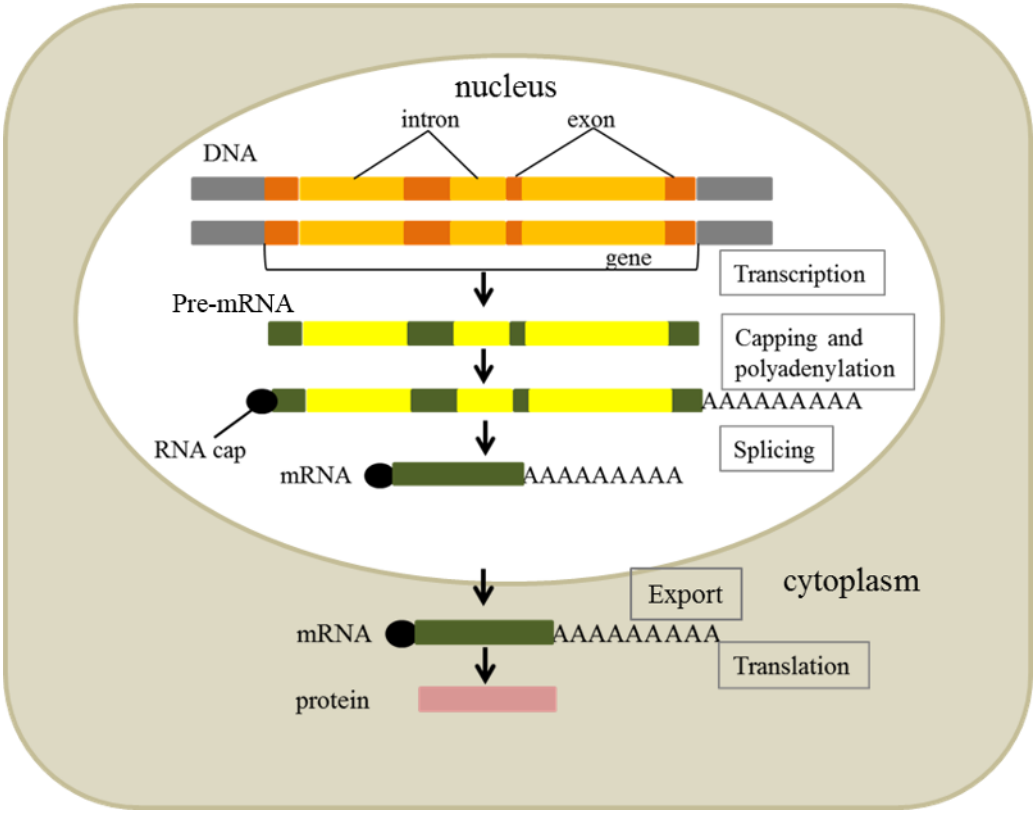


Figure 1.1. Schematic representation of the flow of information from DNA to proteins in eukaryotes. During transcription, coding (exons) and noncoding (introns) regions of DNA are transcribed into precursor messenger RNA (pre-mRNA). The pre-mRNA undergoes modifications from 5' and 3' ends during processes termed capping and polyadenylation. Introns of the pre-mRNA are removed while the remaining exons joined together during splicing. This leads to the formation of messenger RNA (mRNA), which is exported out of the nucleus to the cytoplasm. Finally, the mRNA is used for protein synthesis during translation.

Cross-talk between transcription and RNA processing events and their effects on one another have been widely investigated. It was found that transcription influences capping, splicing and polyadenylation, while splicing, in turn, affects transcription efficiency (reviewed by Bentley, 2014). This evidence led to a whole field of research, ultimately revealing that Pol II is a key player in coupling transcription and RNA maturation processes (Cho et al., 1997; Hirose and Manley, 1998; Hirose et al., 1999; reviewed by Proudfoot et al., 2011). RNA processing factors are recruited to the transcription machinery. This is mostly regulated by posttranslational modifications of

the carboxy-terminal domain (CTD) of the largest subunit (Rpb1) of Pol II (reviewed by Egloff et al., 2012). Capping and polyadenylation events always occur co-transcriptionally, when pre-mRNA is still being synthesised by Pol II (Ho and Shuman, 1999; Moteki and Price, 2002). Although coupling of transcription with splicing has being widely investigated (reviewed by Bentley, 2014), the full extent and the precise mechanisms of co-transcriptional splicing regulation remain unclear. It has been shown that splicing factors interact with transcription components, but their role in transcription has not been well characterized.

1.2 Transcription

Transcription is a highly controlled process which ensures correct cell growth. In eukaryotes, transcription is carried out by three polymerases: Pol I, Pol II and Pol III (Roeder and Rutter, 1969; Roeder et al., 1970). It has been shown that each polymerase has specific gene targets. For example, Pol I synthesises ribosomal RNAs (rRNAs) while Pol III produces noncoding RNAs such as transfer RNAs (tRNAs), 5S rRNA and U6 small nuclear RNA (snRNA; Reddy et al., 1987;Vannini and Cramer, 2012). Pol II is responsible for transcription of protein-coding genes and noncoding RNAs including small nuclear RNAs (U1, U2, U4, U5 and U7 snRNA) and microRNA (miRNA) precursors (reviewed by Arimbassery et al., 2014).

1.2.1 Transcription by RNA polymerase II

Transcription driven by Pol II is a highly controlled process and requires numerous factors. With a help of general transcription factors (GTFs), Pol II is recruited to a promoter region, leading to the formation of a pre-initiation complex (PIC; reviewed by Luse, 2014). Once the first nucleotides of the pre-mRNA have been synthesised, Pol II escapes from the promoter and enters the elongation phase, which comprises the lengthening of a transcript. The transcription is completed when Pol II and its factors are released from the transcription complex and recycled for further transcription events (Zawel et al., 1995; reviewed by Selth et al., 2010). The transcription of highly

expressed genes can be carried out by several polymerase molecules simultaneously (reviewed by Selth et al., 2010). Pol II is a large enzyme (500 kDa) that consists of 12 subunits (reviewed by Coulombe et al., 1999). The CTD domain of the largest Pol II subunit (Rpb1) is implicated in linking transcription and pre-mRNA processing, which is achieved via recruitment of factors required for capping, splicing, cleavage and polyadenylation (Cho et al., 1997; Hirose and Manley, 1998; Hirose et al., 1999; Moteki and Price, 2002; reviewed by Proudfoot et al., 2011).

1.2.1.1 Roles of the carboxy-terminal domain of Pol II

The CTD of Pol II is unique to eukaryotic organisms and conserved from fungi to humans (reviewed by Hsin and Manley, 2012). This domain is absent in Pol I and Pol III (Allison et al., 1985; reviewed by Buratowski et al., 2009). In vertebrates, the CTD contains 52 tandem heptapeptide repeats with the consensus sequence Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 (Corden et al., 1985; Figure 1.2). The number of repeats differs between organisms and it is correlated with their genomic complexity (reviewed by Egloff et al., 2012). Studies in mouse, *Drosophila* and yeast have shown that the total deletion of the CTD is lethal (McCracken et al., 1997). This indicates that the CTD plays vital roles in gene expression.

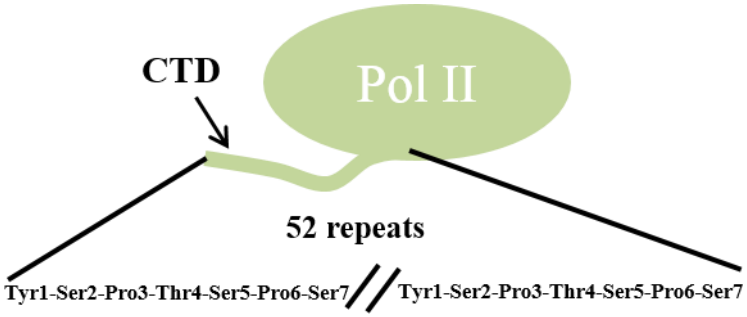


Figure 1.2. The CTD of Pol II consists of 52 heptapeptide repeats with the consensus sequence.

The amino acids of the CTD repeats undergo different post-translational modifications independently of one another (reviewed by Egloff et. al., 2012). The CTD modifications play crucial roles in transcription initiation (Kim et al. 1994), promoter escape (Kugel and Goodrich, 1998), elongation (Marshall and Price 1995; Marshall et al., 1996) and

termination (Hirose and Manley, 1998; Gudipati et al. 2008) as they coordinate the recruitment of many transcription-related factors. The CTD is also involved in co-transcriptional processing, since upon its modifications numerous factors involved in RNA biogenesis are loaded to the CTD (Cho et al., 1997; Hirose and Manley, 1998; Hirose et al., 1999; reviewed by Proudfoot et al., 2011). Different residues of the CTD are responsible for different functions in pre-mRNA processing (Fong and Bentley, 2001).

1.2.1.2 The carboxy-terminal domain modifications

The CTD undergoes phosphorylation at Tyr1, Ser2, Thr4, Ser5 and Ser7 (reviewed by Egloff et al., 2012 and Heidemann et al., 2012), peptidyl-prolyl bond isomerisation at Pro3 and Pro6 (Yogesha et al., 2014) and glycosylation at Thr4, Ser5 and Thr7 residues (Ranuncolo et al., 2012). Phosphorylation and glycosylation are mutually exclusive processes which influence the function of Pol II (reviewed by Bentley, 2014). The CTD modifications are sequential and synchronized during all transcription phases.

Phosphorylation is the most studied post-translational modification of the CTD (reviewed by Egloff et al., 2012). Pol II binds to promoter when it is in a hypophosphorylated state (Lu et al., 1991). Over the course of transcription, kinases and phosphatases add and remove phosphate groups from the CTD repeats. The major CTD phosphorylation occurs at Ser2 and Ser5 residues (Corden el al., 1985; reviewed by Price, 2000) which are essential for the expression of most Pol II transcribed genes. The pattern of CTD phosphorylation and dephosphorylation is correlated with transcription initiation, elongation and termination stages. Ser5 phosphorylation peaks near promoter and promoter proximal regions, indicating that this modification of the CTD predominates at transcription initiation (Komarnitsky et al., 2000; Figure 1.3). As transcription progresses, phosphorylation at Ser2 increases further down towards the end of genes (Komarnitsky et al., 2000; Figure 1.3). Ser7 and Thr4/Tyr1 residues are also phosphorylated during initiation and elongation stages of transcription, respectively (reviewed by Heidemann and Eick, 2012).

The CTD phosphorylation events play crucial functions in RNA biogenesis. For example, Ser5 mediates the recruitment of capping enzymes (Cho et al.,

1997; McCracken et al., 1997; Ho et al., 1999) while Ser2 is required for pre-mRNA splicing (Bird et al., 2004; Gu et al., 2013). Efficient 3' end RNA processing depends on phosphorylation of the CTD at Ser2 and Thr4 residues (Meinhart and Cramer, 2004; Davidson et al., 2014; reviewed by Hsin and Manley, 2012). Phosphorylation at Thr4 and Ser7 residues can display gene-specific features (reviewed by Egloff et al., 2012). For example, phosphorylated Ser7 plays a crucial role in transcription of snRNAs (reviewed by Egloff et al., 2007).

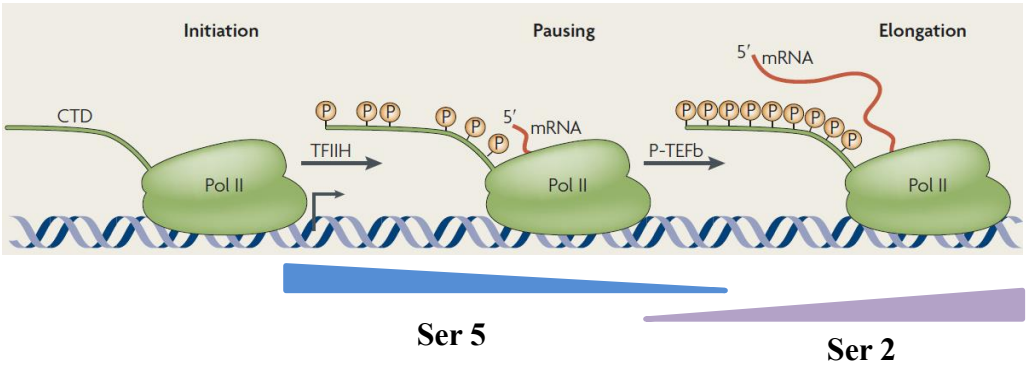


Figure 1.3. Phosphorylation pattern of Ser5 and Sr2 residues of the CTD (Cole and Cowling, 2008). Transcription factors TFIIH and P-TEFb phosphorylate the CTD at Ser5 and Ser2 residues during transcription initiation and elongation stages, respectively. The level of Ser5 phosphorylation is increased during transcription initiation while that for Ser2 is increased towards the end of a gene.

The order and combination of CTD phosphorylation events are regulated by different enzymes called Cyclin-dependent kinases (CDKs). CDK7 has a predominant role in phosphorylating Ser5 residue (Feaver et al., 1991; Lu et al., 1992) but also it phosphorylates Ser7 residue (Akhtar et al. 2009). Until recently it was believed that CDK9 was the main kinase which phosphorylates Ser2 (reviewed by Price, 2000). However, it has been shown that CDK11, CDK12 and CDK13 are additional Ser2 kinases (Pak et al., 2015; Bartkowiak et al., 2010; Greifenberg et al., 2016). Apart from Ser2, CDK9 was determined to be able to phosphorylate Ser5, Ser7 and Thr4 (Hsin et al., 2011; reviewed by Hsin and Manley, 2012). A dual functionality of CDK8 in phosphorylating both Ser2 and Ser5 residues has also been demonstrated *in vitro* (Liao et al., 1995; Sun et al., 1998). It is apparent that some CDKs exhibit functions in phosphorylating more than one residue, indicating that the CTD code is highly complex.

1.2.2 Transcription process

1.2.2.1 Transcription initiation

Transcription of protein-coding genes is regulated at multiple steps and it is one of the most highly regulated processes in gene expression. The first step of transcription is initiation (Figure 1.4) which begins with the binding of an activator to a promoter mediating an ordered recruitment of GTFs. The promoter elements are binding platforms for GTFs and define transcription start site (TSS; reviewed by Luse, 2014). The engagement of Pol II to a promoter locus is facilitated by core GTFs leading to the assembly of the pre-initiation complex (PIC; Sainsbury et al., 2015). The factors of PIC facilitate promoter opening, initiate RNA production and stimulate the promoter escape by Pol II.

The sequence composition of promoters can be different, although many but not all have an element called the TATA box (Mathis and Chambon, 1981). The composition of transcription complexes on TATA-containing and TATA-less promoters differs in the factors that recognise the promoter (Shandilya and Roberts, 2012). In mammalian TATA-containing genes, the TATA box is located about 30 bp upstream of TSS (Ponjavic et al., 2006). The interaction of TATA-binding protein TBP, a component of the multisubunit TFIID factor complex (Nikolov et al., 1995), with the TATA box is a crucial step for PIC assembly and it is regulated by activators and repressors (Kuras and Struhl, 1999; Cang et al., 1999). TBP binds to the TATA element and initiates a DNA bend (Figure 1.4; Starr et al., 1995). TFIID binding facilitates the ordered assembly of other GTFs, leading to PIC formation. TFIIA and TFIIB are recruited to the complex leading to the stabilization of the interaction between TBP and the TATA box (Shandilya et al., 2012). The loading of Pol II and TFIIF onto an active promoter is facilitated by TFIIB (Bushnell et al., 2004) which also has a function in initiating RNA synthesis. TFIIF is involved in the stabilisation of TFIIB within the PIC, transcription start site (TSS) selection and prevention of pausing by Pol II (Shandilya and Roberts, 2012). The assembly of the PIC is completed when TFIIE and TFIIH are recruited. These transcription factors are required for promoter DNA opening ("transcription bubble"). CDK7, the kinase subunit of the general transcription factor TFIIH (Lu et al. 1992; Akoulitchev et al., 1995), mediates the addition of phosphate groups to the Ser5

and Ser 7 residues of the CTD (Feaver et al, 1994; Akhtar et al., 2009). This leads to transcription initiation and promoter clearance (Kugel and Goodrich, 1998). After Pol II has escaped from the promoter it enters the transcription elongation phase.

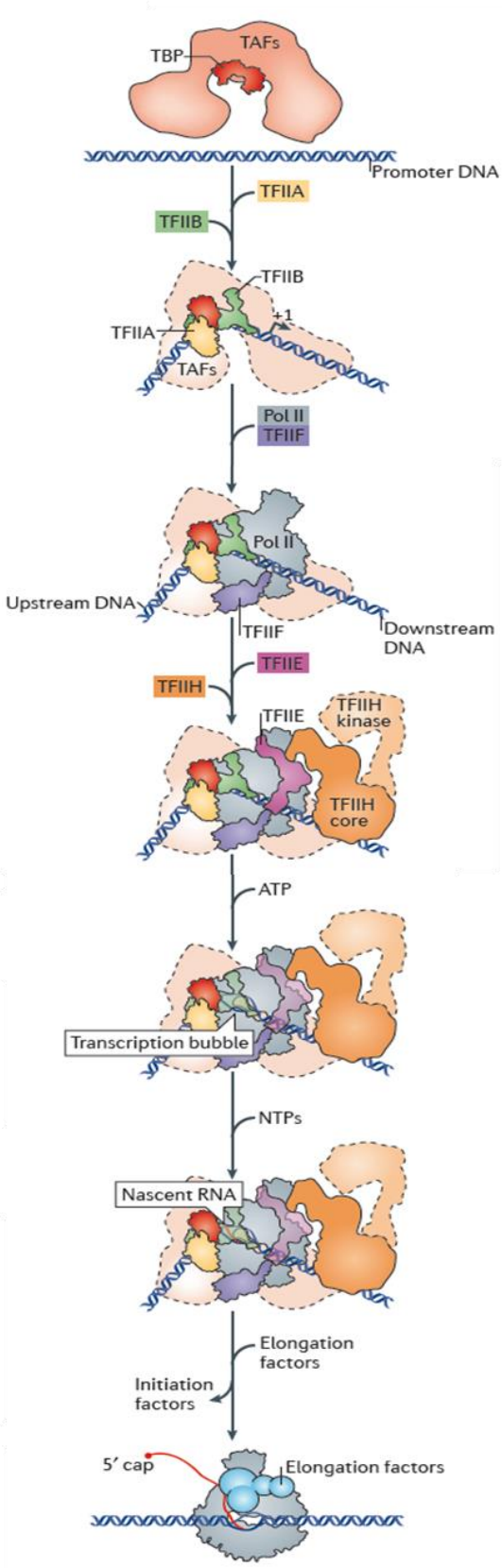


Figure 1.4. Schematic representation of Pol II transcription initiation (Sainsbury et al., 2015). A stepwise assembly of pre-initiation complex (PIC) by general transcription factors (GTFs) and RNA polymerase II on a promoter is depicted. After the completion of the PIC formation, in an ATP-dependent manner DNA is opened and RNA synthesis commences. This is followed by Pol II interacts with elongation factors which enable the polymerase to enter transcription elongation phase.

1.2.2.2 Transcription elongation

Upon the successful promoter release, Pol II transcription becomes processive. In general, the polymerase moves in a 5' to 3' direction, but is capable of moving backward for several nucleotides. This process is important for removing mispaired nucleotides (Thomas et al., 1998; Izban and Luse, 1992). Between early elongation and productive elongation, the negative elongation factor (NELF) in conjunction with the 5, 6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) sensitivity inducing factor (DSIF) cause Pol II pausing in the promoter-proximal region (Coppola et al., 1983; Yamaguchi et al., 1999; Wu et al., 2003), 25–60 bp downstream of the transcription start site. This process is not a repressive step but acts as a quality checkpoint for capping (reviewed by Adelman and Lis, 2012).

The release of Pol II from its pausing site is facilitated by the positive transcription elongation factor (P-TEFb), which is the main regulator of early elongation steps (Marshall and Price 1995; Marshall et al., 1996). P-TEFb contains CDK9, which has a predominant role in phosphorylating the CTD at Ser2 residues, leading to transcription elongation (reviewed by Price, 2000). In addition, the phosphorylation of NELF by P-TEFb promotes NELF to leave the complex, allowing Pol II to proceed the elongation (Fujinaga et al. 2004; Cheng and Price 2007). A balance between negative elongation factors and P-TEFb determines the level of pausing of Pol II (reviewed by Jonkers and Lis, 2015).

After Pol II is released from the pausing, Ser2 phosphorylation levels increase towards the end of genes leading to productive elongation. At this stage the transcription machinery mediates the recruitment of factors required for co-transcriptional splicing (Hirose and Manley, 1998), cleavage and polyadenylation (McCracken et al., 1997;

Zhao et al. 1999; reviewed by Proudfoot, 2011). This indicates that Pol II is accompanied not only by transcription elongation factors but those required for RNA processing events. During transcription elongation, Pol II faces significant obstacles. For example, histone modifications affect the strength of DNA binding around nucleosomes forcing Pol II to pause frequently (reviewed by Selth et al., 2010). The Pol II pauses are considered to be regulatory steps for transcription and RNA processing including pre-mRNA splicing.

1.2.2.3 Transcriptional termination and 3' end processing of RNA

The last step of the transcription cycle is termination which leads to the production of stable transcripts that are consequently exported to the cytoplasm for protein production (translation). The removal of the pre-mRNA from the transcription machinery is a highly regulated process, where cleavage and polyadenylation are key events (Colgan et al., 1997; reviewed by Proudfoot, 2011). The CTD plays an essential role in recruiting cleavage and polyadenylation factors to the elongating Pol II machinery (McCracken et al., 1997; Zhao et al., 1999). This indicates that termination and 3'end processing events of the pre-mRNA occur during transcription (reviewed by Proudfoot, 2011).

There are some important sequences located at the end of the gene that are required for the recruitment of the cleavage and polyadenylation machinery. In particular, a polyadenylation signal (PAS) that comprises a hexanucleotide AAUAAA sequence (Lim and Canellakis, 1970; Edmonds et al., 1971; Adesnik et al., 1972; Mendecki et al., 1972; Proudfoot, 1989). It has been shown that this site is not well conserved (Tian et al., 2005). In addition, two terminator elements are located downstream of PAS such as GU-rich (Gil and Proudfoot, 1984, 1987; McLauchlan et al., 1985) and AT-rich sequences (White et al., 2013). After transcription has passed these elements, these sites are exposed to the cleavage and polyadenylation machinery. As a result, cleavage of the nascent RNA occurs between PAS and GU-rich sequences. This implies that the signals are recognized co-transcriptionally.

The main factors that mediate the key events of transcription termination are cleavage and polyadenylation specificity factor (CPSF; Dantonel et al., 1997), cleavage

stimulatory factor (CstF; Takagaki et al., 1996), cleavage factor I (CFI; reviewed by Proudfoot, 2011) and Poly (A) polymerase (PAP; Winters and Edmonds, 1973 a, b). Once PAS appears in the pre-mRNA, CPSF binds to it leading to Pol II pausing (Yonaha and Proudfoot, 2000; Nag et al., 2007). This event is necessary for efficient transcription termination. The GU-rich sequence is recognized by CstF, while the role of CFI is to stimulate the recognition of PAS (Brown and Gilmartin, 2003; Venkataraman et al., 2005). The cleavage of RNA is facilitated by CPSF73 (a component of CPSF complex) leading to RNA break (Mandel et al., 2006). The RNA tethered to Pol II undergoes degradation by the 5'-3' exonuclease Xrn2 (Gromak et al., 2006; West et al., 2008). This results in Pol II termination and its dissociation from the chromatin template. PAP is then recruited to the RNA lying to the 5' from the cleavage site to add 250–300 adenine nucleotides (in humans) to the 3' end of the RNA (Kaufmann et al., 2004; Mandel et al., 2008). After, the poly(A) tail is bound by the poly(A)-binding protein (PABP) which serves as a scaffold to protect the transcript from exonuclease degradation (Blobel, 1973; Minvielle-Sebastia et al., 1997). The polyadenylated RNA ensures efficient mRNA export and translation (Lewis et al., 1995).

It has been discovered that many genes contain more than one PAS, indicating the existence of alternative polyadenylation (APA; Tian et al., 2005). *Trans*-acting factors are involved in APA regulation, leading to the production of mRNAs with alternative 3' ends. It has been found that cryptic PASs located in close proximity to the 5'SS are silenced (Guo et al., 2010). It was shown that the usage of these sites is subject to regulation by splicing factor U1 snRNP (Kaida et al., 2010). Based on this evidence it becomes clear that all RNA processing events are cross correlated and to some extent affect one other.

1.3 Capping of pre-mRNA

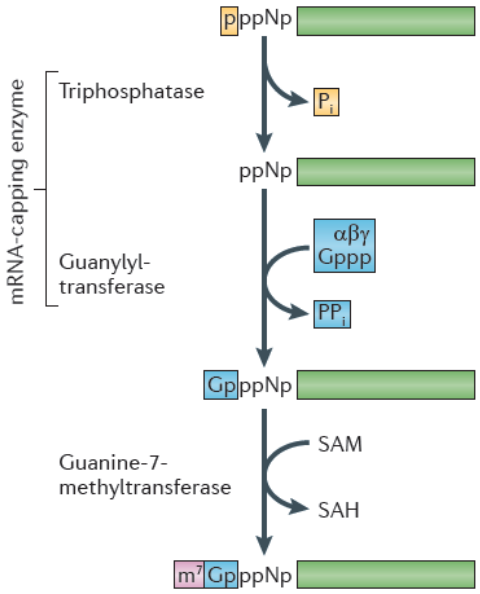
In eukaryotes, the modification of the 5' end of Pol II transcripts is the first pre-mRNA processing event which occurs as soon as transcripts are emerging from Pol II (Furuichi et al., 1975; Shimotohno et al., 1997; Moteki and Price, 2002). The promoter-proximal pause of Pol II coincides with capping and serves as a checkpoint to ensure proper capping of the nascent pre-mRNA (reviewed by Adelman and Lis, 2012). RNA capping entails three enzymatic modifications (Figure 1.5 A) which occur when the length of RNA reaches about 20-40 nt (Coppola et al., 1983; Rasmussen and Lis, 1993; Yamaguchi et al., 1999). Once pre-mRNA emerges from the exit channel of Pol II, RNA 5' triphosphatase removes the terminal phosphate group creating a diphosphate RNA end (reviewed by Decroly et al., 2012). Next, RNA guanylyltransferase facilitates the addition of a guanosine monophosphate (GMP) to the diphosphate RNA forming an unusual 5'-5' triphosphate linkage (GpppRNA). GpppRNA cap is methylated at position 7 by methyltransferase. As a result, the 5' end of the nascent pre-mRNA is modified (Figure 1.5 B). The cap binding complex (CBC) is recruited to the 7-methylguanosine triphosphate cap (m7 GpppRNA) to stabilize and protect RNA from decapping and degradation (reviewed by Gonatopoulos-Pournatzis and Cowling, 2014). Capping mediates mRNA stability and protects it from degradation. The cap structure of pre-mRNA positively affects splicing, polyadenylation, mRNA export and translation (Konarska et al., 1984; Edery and Sonenberg, 1985; Ohno et al., 1987; reviewed by Bentley, 2014).

The recruitment of CBC and the capping enzymes to the CTD is mediated by phosphor-Ser5 residues (Cho et at., 1997; McCracken et al., 1997; Ho et al., 1999). The recruitment of the capping enzymes to the transcription complex disables NELF-mediated stalling of transcription, resulting in progressive transcription elongation (Mandal et al., 2004). It has been shown that capped pre-mRNA enhances elongation rates through an association between CBC and P-TEFb (Lenasi et al., 2011).

The removal of the 5' cap structure occurs in the cytoplasm which is facilitated by the decapping enzymes Dcp1 and Dcp2 (Dunckley and Parker, 1999; Wang et al., 2002). This results in the exposure of the 5' mono-phosphorylated RNA to the cytoplasmic 5'

exonuclease Xrn1 (Stevens, 1978; reviewed by Houseley and Tollervey, 2009). Decapping is an important process in gene expression and it is implicated in mRNA decay, turnover and quality control.

A



B

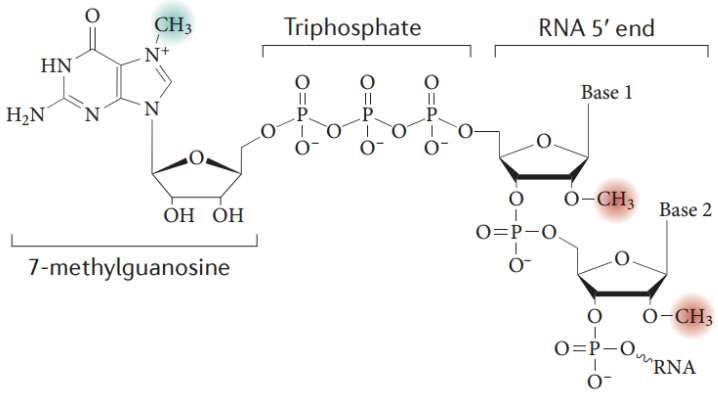


Figure 1.5. The mechanism of pre-mRNA capping and the cap structure (Decroly et al., 2012). (A) Illustration showing the cap structure synthesis which is facilitated by three enzymes. (B) The mRNA cap structure contains a 7-methylguanosine which is connected to the mRNA through a 5'–5' triphosphate linker. The methyl group at the guanosine N7 is depicted in green. Methylation reactions also occur on the first and second nucleotides of mRNA (cap-1 and cap-2). These methyl groups are depicted in red.

1.4 Pre-mRNA splicing

Splicing was discovered in 1977 (Berget et al., 1977; Chow et al., 1977) and it was determined that intron removal is catalysed by a macromolecular enzymatic complex termed the spliceosome. This process occurs in coordinated and regulated sequences of RNA-RNA, RNA-protein and protein-protein interactions. During splicing, the spliceosome is assembled in a stepwise manner by the ordered recruitment of five small uridine-rich nuclear ribonucleoprotein particles (snRNPs) to the pre-mRNA (reviewed by Will and Lührmann, 2011). Each snRNP is a complex itself consisting of a small nuclear RNA (snRNA), specific proteins and seven Sm proteins. The latters are shared among the snRNPs (Hermann et al., 1995; Raker et al., 1996; Mayes et al., 1999). The role of snRNPs is to identify the exons and introns in the pre-mRNA to promote a splicing reaction. Some of the snRNPs base pair with short consensus sequences termed splice sites (SSs). These elements are located at the boundaries between exons and introns of pre-mRNA. In addition to the core splicing factors, the spliceosome contains up to 150 auxiliary factors, including RNA helicases which act as motors to drive the rearrangements of RNA-RNA and RNA-protein interactions (reviewed by Will and Lührmann, 2011).

1.4.1 The mechanism of splicing

In humans the 5’SS contains about eight nucleotides with a GU dinucleotide being located at the beginning of an intron (Breathnach et al., 1978; Staley and Guthrie, 1998). The 3’SS includes the 3’SS itself, the branch point sequence (BPS) and a polypyrimidine tract (PPT) (Staley and Guthrie, 1998). The 3’SS comprises a AG dinucleotide being located at the end of an intron. BPS harbouring adenosine lies approximately 15–50 nt upstream of the 3’SS. PPT having a sequence of pyrimidines is located between the 3’SS and the BPS (reviewed by Will and Lührmann, 2011). In addition to these key sequence elements, there are numerous *cis*-acting regulatory motifs which are embedded over the whole length of the pre-mRNA (reviewed by Kornblihtt et al., 2013). The *cis*-acting sequences are required for the recruitment of *trans*-acting factors which are able to enhance or repress the spliceosome assembly.

Upon base pairing to the splice sites, snRNAs mediate spliceosome assembly which catalyses the intron removal via two sequential transesterification reactions (Figure 1.6; Moore and Sharp, 1993). In the first step, the 2'-OH group of the branch point adenosine carries out a nucleophilic attack on the 5'-3'phosphodiester bond at the 5'SS, connecting the 5' end of the intron to the branch site (Figure 1.6). This reaction leads to the formation of the 3' exon lariat intermediate and release of the 5' exon. In the second step, the 3'-OH group of the upstream exon attacks the 5'-3'phosphodiester bond at the 3'SS, displacing the intron (Figure 1.6). This leads to the splicing of two exons together and the release of introns having a lariat structure (Grabowski et al., 1984; Ruskin et al., 1984).

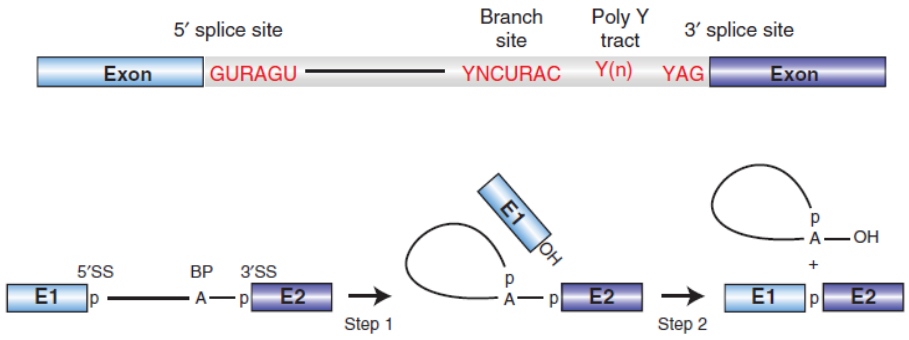


Figure 1.6. Schematic representation of two sequential transesterification reactions of pre-mRNA splicing (Will and Lührmann, 2011). Boxes and lines illustrate exons and introns, respectively. The branch site adenosine, 5' and 3' splice sites are represented. In the first reaction, the 2'-hydroxyl group of the adenosine branch point performs a nucleophilic attack on the phosphate group at the 5'SS resulting in the production of lariat intermediate. In the second reaction, the 3'-hydroxyl group of the 5'SS attacks the phosphate at the 3'SS. This results in the fusion of two exons and release of the intron in a form of lariat.

Efficient *in vitro* splicing assay using nuclear extracts (NEs) allowed characterisation of composition of the spliceosomal complexes (Reed, 2000). It was found that within the spliceosome, snRNPs rearrange, some interact with each other and some undergo structural intramolecular reorganisations (Hilliker et al. 2007; Perriman and Ares, 2007). This leads to the formation of different complexes named in order of appearance as complex A, B, B^{act} and C (Figure 1.7; Jamison et al., 1992; reviewed by Will and Lührmann, 2011). Apart from these complexes there is also a heterogeneous complex

(complex H) formed on pre-made pre-mRNA, which comprises RNA binding proteins of NE bound to the added pre-mRNA. However, it is not known if this complex is formed in cells because it has been suggested that complex H is not a functional precursor to the spliceosome (Das et al., 2006).

In vitro splicing assays have shown that the early splicing complex (complex E) is assembled in an ATP-independent manner when the U1 snRNP base pairs to the 5'SS through the 5' terminus of the U1 snRNA (Figure 1.7; Michaud and Reed, 1991; Jamison et al., 1992; Eperon et al., 2000). This interaction is stabilized by binding of serine-arginine-rich (SR) proteins (Eperon et al., 1993; Kohtz et al., 1994), proteins of the U1 snRNP (Du and Rosbash, 2002) and the cap-binding complex (Lewis et al., 1996). In addition, complex E includes SF1, known as branch point binding protein (BBP; Kramer and Utans, 1991; Berglund et al., 1998) and U2 auxiliary factor (U2AF), which consists of two subunits U2AF35 and U2AF65 (reviewed by Will and Lührmann, 2011). The 3'SS and PPT are recognised by U2AF35 and U2AF65, respectively (Merendino et al., 1999; Wu et al., 1999; Zorio and Blumenthal, 1999; Shao et al., 2014). Both of these factors are able to interact with each other (Zamore and Green, 1989). In an ATP-dependent manner complex A is formed, in which the U2 snRNP replaces SF1 from the BPS and tightly binds to it (reviewed by Will and Lührmann, 2011). The precatalytic complex B is formed when the tri-snRNP is recruited. The tri-snRNP consists of U4 and U6 snRNPs, which are base-paired and the U5 snRNP. At this stage the spliceosome contains all the necessary components for catalysis but is not active yet, because some internal reorganisations are required for splicing reactions to begin. Helicases disrupt the connection between the U4 and U6 snRNPs, resulting in the release of the U1 and U4 snRNPs from the spliceosome (reviewed by Will and Lührmann 2011). This leads to the formation of the catalytically active complex B^{act} in which the U6 snRNP binds to the free 5'SS (Wassarman and Steitz, 1992; Kandels-lewis and Seraphin, 1993) and base pairs with the U2 snRNP (Madhani and Guthrie, 1992). Together these snRNPs form the active core of the spliceosome. Complex C is formed when the U5 snRNP binds to the 5' and 3' exons and brings them into close proximity allowing the second transesterification reaction to occur (Yan et al., 2015). Consequently, the exons are joined together while the lariat intron is removed. Upon the spliceosome cycle completion, splicing components are displaced and recycled for a new splicing process.

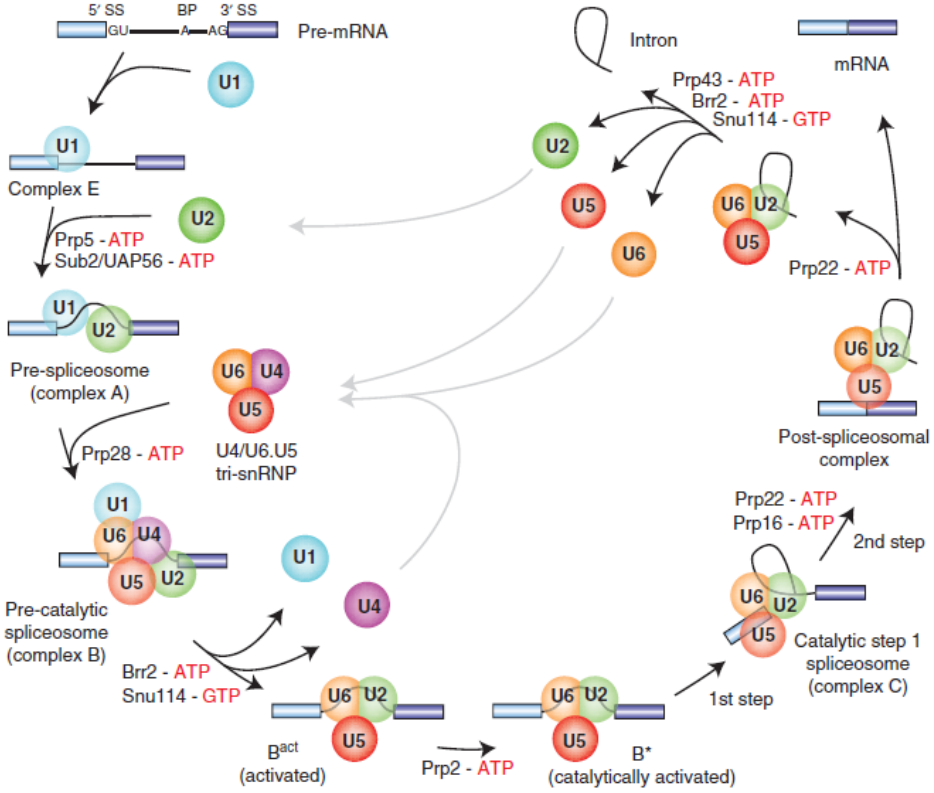


Figure 1.7. Schematic representation of the splicing cycle (Will and Lührmann, 2011). The order of five U snRNPs recruitment to pre-mRNA during the spliceosome assembly is represented. The key additional factors and main spliceosome complexes are depicted.

1.4.2 Alternative splicing

Initially, it was thought that the number of genes would correspond to the number of proteins. In the mid-1990s it was estimated that the number of genes present in the human genome was between 50,000 and 100, 000 (reviewed by Pertea and Salzberg, 2010). However, later it was discovered that humans have only about 25, 000 genes and up to 95% of these genes undergo alternative splicing (AS; Wang et al., 2008; Pan et al., 2008). With the help of AS one gene is capable of producing many different mRNAs, leading to the synthesis of different protein isoforms (Sharp, 2005; Berget et al., 1977; Chow et al., 1977). On average, there are 3.9 AS events per human gene (Su et al., 2006). This means that apart from constitutively spliced exons, which are always included into mRNA, the pre-mRNA has variable exons and introns. AS comprises a

reaction, in which the recognition of 5'SS and 3'SS is in competition with other splice sites. This indicates that splice sites have different "strength" which depends on complementarity with the spliceosomal snRNPs. Usually, variable exons are surrounded by at least one "weak" SS. In different tissues variable exons are spliced in different ways, outlining tissue-specific expression pattern of a gene.

AS is a highly regulated process in which numerous splicing factors play roles as enhancers or inhibitors (reviewed by Kornblihtt et al., 2013). For example, the serine/arginine-rich proteins (SR) represent a group of activators, while heterogeneous nuclear ribonucleoproteins (hnRNPs) have negative function in splicing. These two groups of splicing regulators bind to sequences within the pre-mRNA and work antagonistically on the spliceosome assembly (Figure 1.8). These sequences can be enhancers or silencers and they are located within exons and introns.

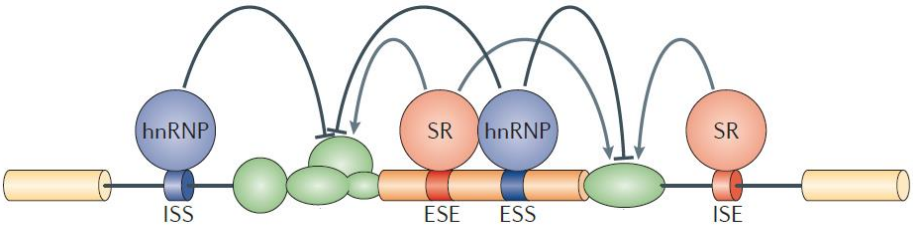


Figure 1.8. Sequences and factors important for alternative splicing (Kornblihtt et al., 2013). Splicing is regulated by *cis*-regulatory sequences located within exons and introns of the pre-mRNA. These motifs are exonic splicing enhancers (ESEs), exonic splicing silencers (ESSs), intronic splicing enhancers (ISEs) and intronic splicing silencers (ISSs). There are two main families of alternative splicing regulatory proteins such as SR and hnRNPs. These factors either mediate or prevent recruitment of components of the spliceosome (green) which interact with the 5'SS and 3'SS.

1.4.3 Co-transcriptional splicing

Initially, based on *in vitro* studies, it was believed that splicing was a transcription-independent process. However, a growing body of evidence obtained for the last couple of decades has shown that the majority of splicing events are catalysed in a transcription-dependent manner (reviewed by Bentley, 2014). On average a splicing event is accomplished within 0.5-2.5 min (Huranová et al., 2010), suggesting that pre-mRNA splicing can occur before Pol II has reached the end of a gene. This allows multiple spliceosomes to be assembled on a pre-mRNA simultaneously.

1.4.3.1 Evidence of co-transcriptional splicing

The term "co-transcriptional splicing" is used not only when the transcription and splicing machineries are physically coupled, but also when splicing is completed post-transcriptionally meaning that splicing can, at least, be initiated during transcription. The first and most convincing data on co-transcriptional splicing was obtained by Beyer and Osheim in 1988. The electron micrographs of *Drosophila melanogaster* chorion genes revealed intron loops on the nascent pre-mRNAs that are still attached to the transcribing Pol II (Figure 1.9). Later it was shown that splicing factors are loaded to transcription sites during active transcription (Spector et al., 1991; Jimenez-Garcia and Spector, 1993; Misteli and Spector, 1997; Neugebauer and Roth, 1997). The accumulation of splicing factors on intron-containing genes has also been observed (Huang and Spector, 1996; Listerman et al. 2006). RNA *in situ* hybridization also provided evidence that spliced RNA is retained at the chromatin until splicing is completed (Zhang et al., 1994; Lacadie and Rosbash, 2005).

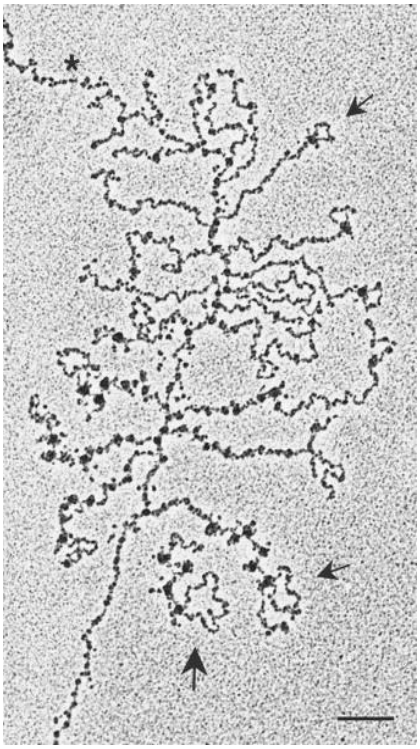


Figure 1.9. Electron micrograph of a *Drosophila melanogaster* embryonic gene showing co-transcriptional splicing (Beyer and Osheim, 1988). The nascent pre-mRNAs emerge from the DNA template have splicing loops (arrows). The length of the gene is 6 kb. Bar, 200 nm.

It has been revealed that the majority of introns are excised during transcription especially those that flank constitutive exons (Neugebauer, 2002; reviewed by Merkhofer et al., 2014). Introns flanking alternative exons are removed before Pol II reaches the end of a gene. The 3' end terminal introns are excised after the transcript release (Baurén G, and Wieslander L, 1994; Wetterberg et al., 1996). It has been shown that Pol II is forced to make a pause at the terminal exon giving time for spliceosome assembly on the upstream introns (Carrillo Oesterreich et al., 2010; Alexander et al., 2010). Regarding AS, it has been found that the decision on the variable exon inclusion or exclusion is made during transcription (Singh and Padgett, 2009). The promoter structure, regardless of the promoter strength, can influence AS pattern, where the sensitivity to splicing regulators like SR proteins strongly depends on the promoter architecture (Cramer et al. 1997; Cramer et al. 1999; Kadener et al., 2002; Nogues et al., 2002). Promoter sequence also influences the efficiency of constitutive splicing (Rosonina et al., 2003).

1.4.3.2 “The recruitment model” of co-transcriptional splicing

The recruitment coupling model suggests that splicing regulators and spliceosome components associate with the transcription machinery leading to their increased local concentration in the locus of transcription (Ujvari and Luse, 2004). As soon as the nascent transcript emerges from the transcribing Pol II, it becomes easier for accumulated splicing factors to be recruited from the transcription apparatus to their binding sites within the pre-mRNA (Greenleaf, 1993; Pandya-Jones and Black, 2009; reviewed by Hsin and Manley, 2012). This essential property of the CTD to recruit splicing factors enables the transcription and splicing machineries to be functionally coupled. Some studies have shown that splicing factors do not only associate with the polymerase but can directly interact with GTFs (Fong and Zhou, 2001; reviewed by Bentley, 2014). The rapid co-transcriptional recruitment of splicing factors to the pre-mRNA stabilizes RNA preventing it from degradation and ensures proper maturation (Hicks et al., 2006).

Splicing factors such as U2, U5, U6 snRNPs and U2AF65 have been shown to be recruited to the transcription machinery if genes contain introns (Listerman et al., 2006; Tardiff and Rosbash, 2006). In comparison, U1 snRNP is preloaded onto genes regardless of the presence of introns (Spiluttini et at., 2010; Brody et al., 2011). There is a co-transcriptional accumulation of U4/U6·U5 snRNPs on active transcription units which is mediated by CBC (Pabis et al., 2013). Splicing factor U2AF65 is capable to interact directly with the CTD (David et al., 2011; Spiluttini et al., 2010). It has been shown that phosphorylated Ser2 is responsible for the recruitment of splicing factor U2AF65 (Gu et al., 2013; Figure 1.10).

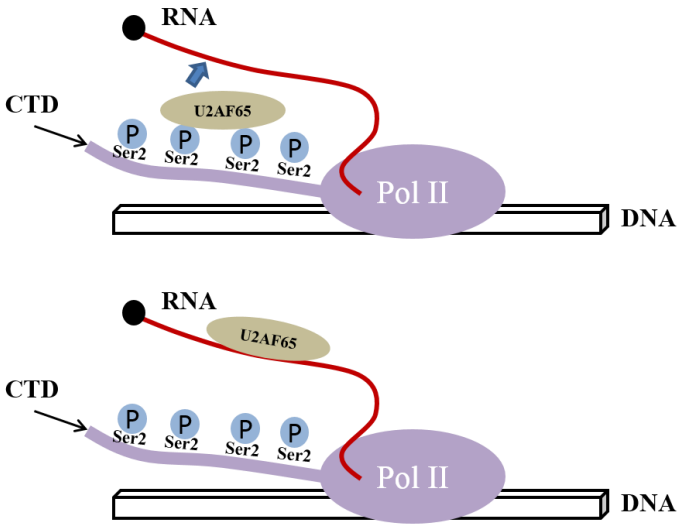


Figure 1.10. The recruitment of U2AF65 to phosphorylated Ser2 residues of the CTD. Ser2 phosphorylation mediates recruitment of U2AF65 and U2 snRNP to the CTD. As soon as pre-mRNA appears from the exit pore of Pol II, the splicing factors are recruited to the splice sites. The cap structure of the pre-mRNA is illustrated in black.

Splicing regulators such as SR proteins and hnRNPs associate with the pre-mRNA in a transcription-dependent fashion (Listerman et al., 2006), suggesting that the competition between SR proteins and inhibitory hnRNP proteins can occur co-transcriptionally (Das et al., 2007). SR proteins were found to be efficiently recruited to both intron-containing and intronless genes (Huang and Steitz 2001; Pozzoli et al., 2004; Spiluttini et al., 2010). SR proteins are able to interact with the transcription machinery before transcription starts and during transcription complexes assembly (Kornblihtt et al., 2004; Lin et al., 2008; Ji et al., 2013; Das et al., 2007). These proteins preferentially bind to the phosphorylated CTD (Vincent et al., 1996; Mortillaro et l., 1996; Kim et al., 1997). It has been shown that the direct interaction between SR proteins and the CTD can affect the efficiency of splicing *in vitro* (Das et al., 2007). Strong support for the recruitment model was obtained from experiments showing that splicing factor SRp20 is recruited from the CTD to the fibronectin pre-mRNA and inhibits the inclusion of a cassette exon in the mRNA (de la Mata et al., 2006; Figure 1.11). Some proteins interact with the CTD transiently (Sapra et al. 2009; reviewed by Bentley, 2014). Although, it has been shown that numerous splicing factors are recruited to the transcribing

polymerase complex, but it is not easy to understand if these interactions are direct or mediated by the nascent pre-mRNA (Listerman et al., 2006).

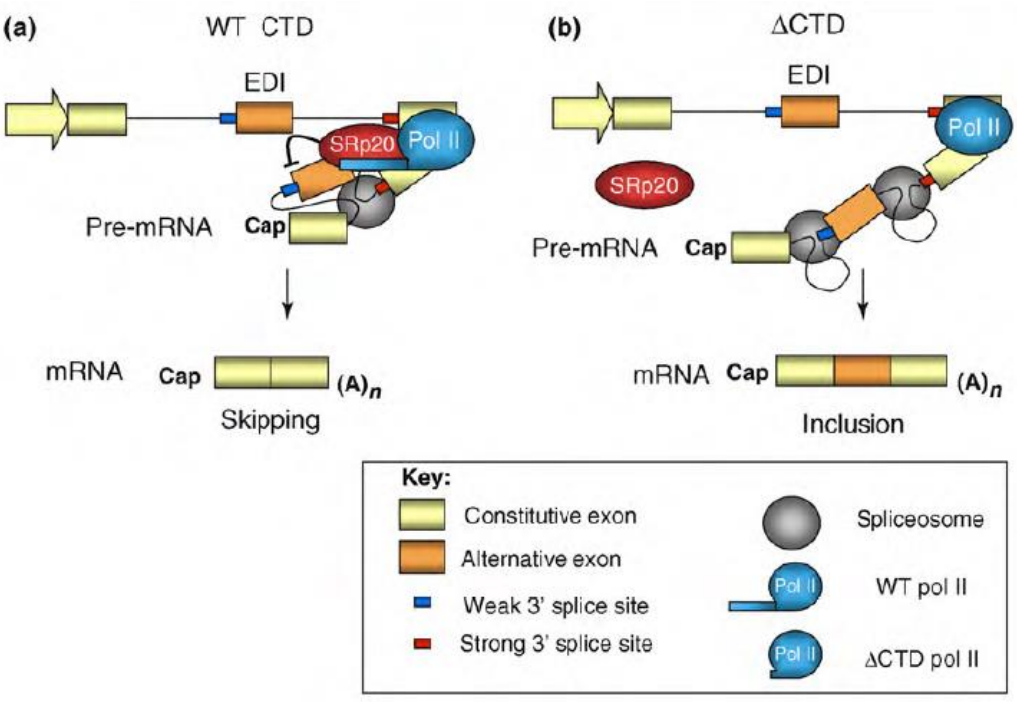


Figure 1.11. The recruitment coupling model of co-transcriptional splicing (Munoz et al., 2010). (A) Upon recruitment to the CTD, SRp20 facilitates skipping of extra domain I (EDI) alternative exon of the human fibronectin gene. (B) Mutated CTD fails to recruit SRp20 to the EDI exon resulting in the production of mRNA containing the variable exon.

1.4.3.3 “The kinetic model” of co-transcriptional splicing

A kinetic coupling model was originally proposed by Eperon et al. (1988). It was suggested that the rate of transcription can influence RNA secondary structure, which consequently can affect splicing. Transcription elongation rates are different across the chromatin because of the presence of the nucleosomes (reviewed by Kornblihtt et al., 2004). The analysis of genome-wide datasets has shown that nucleosomes are mainly positioned in exons, while there is a random distribution of nucleosomes in introns (Andersson et al., 2009; Schwartz et al., 2009; Kolasinska-Zwierz et al., 2009). This suggests that during transcription Pol II makes transient pauses over exons, allowing the splicing machinery to identify SSs on pre-mRNA (Berget, 1995; Nahkuri et al., 2009). It has been shown that Pol II pauses on terminal exons and this facilitates the spliceosome assembly on the upstream introns (Carillo Oesterreich et al., 2010). Therefore, transcription elongation rates affect splicing outcome (Roberts et al., 1998; reviewed by Kornblihtt et al., 2004).

The kinetic model supports the notion that the rates of Pol II-mediated elongation determine the time for weak and strong splice sites to be exposed to splicing factors. It implies that the speed of transcription modulates the recognition of weak splice sites. This provides the opportunity for competition between splicing factors which assemble on the upstream and downstream splice sites (de la Mata et al., 2003). This, in turn, determines if they promote exon inclusion or skipping (reviewed by Munoz et. al., 2010). In a situation when the elongation is fast, the skipping of variable exon occurs if the 3’SS of the downstream intron is stronger than the 3’SS of the upstream intron (Figure 1.12; reviewed by Munoz et al., 2010). Since transcription occurred rapidly and the spliceosome has not assembled yet, the strong 3’SS of the downstream intron will be favoured. This leads to the exon skipping as the 5’SS is joined to the strong downstream 3’SS rather than the weak upstream 3’SS. In contrast, when the elongation rate is reduced, for example, because of the polymerase pausing between these two splice sites, the upstream intron is spliced first, followed by the splicing of the downstream intron. This results into the inclusion of an alternative exon into the mRNA. In both models, the CTD plays a key function as the recruitment of splicing factors and the modulation of the transcription elongation speed are controlled by the modifications of the CTD (reviewed by Merkhofer et al., 2014).

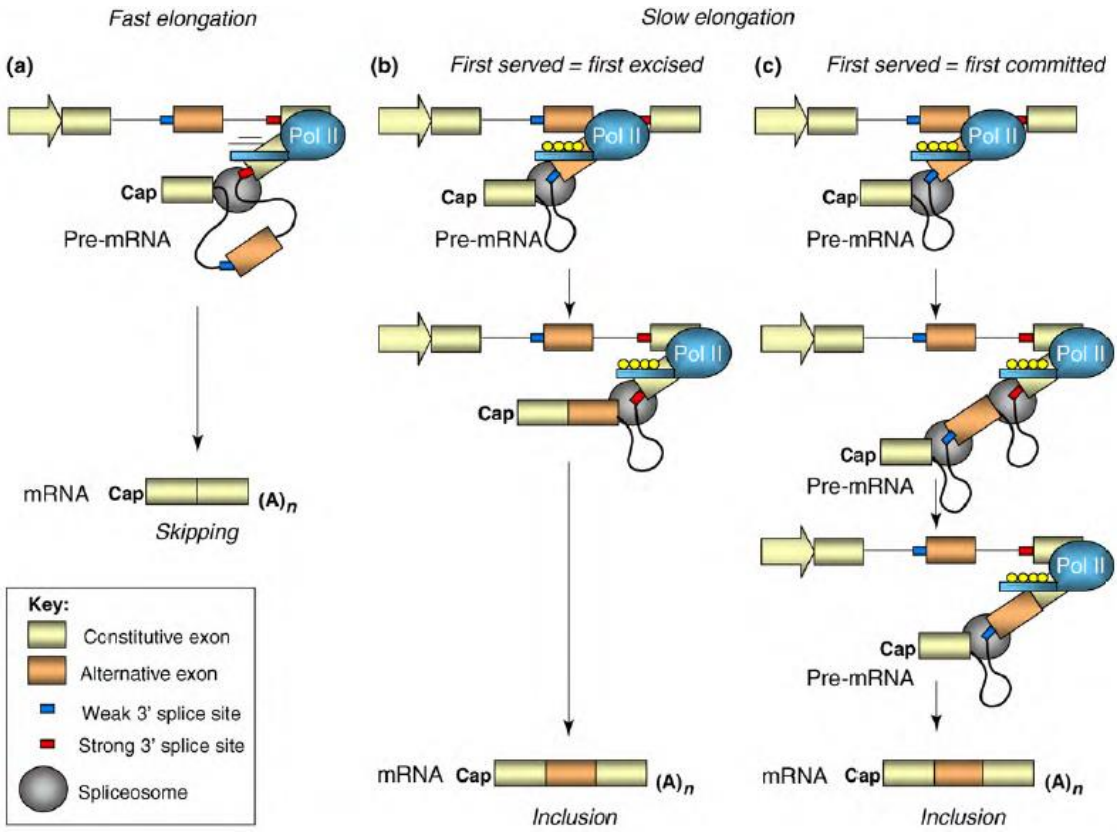


Figure 1.12. The kinetic model of co-transcriptional splicing in alternative splicing regulation (Munoz et al., 2010).

1.4.3.4 Splicing effects transcription

The "reverse coupling" model suggests that transcription efficiency is correlated with splicing activity. Early studies demonstrated that the presence of an intron within a gene had a positive effect on transcription efficiency (Brinster et al., 1988; Choi et al., 1991). Later it was shown that Pol II activity is enhanced if genes containing introns in comparison to intronless genes (Furger et al., 2002). Also a promoter proximal 5'SS has a transcription stimulatory effect (Furger et al., 2002; Damgaard et al., 2008). A template harbouring mutated PPT, the 5' and 3' splice sites showed impaired ability to stimulate transcription (Fong and Zhou, 2001). These data confirm an intron-dependent stimulatory effect on transcription (Fong and Zhou, 2001; Damgaard et al., 2008). It has been revealed that GTFs, which participate in PIC assembly, are more abundant over

promoter region and the distribution of Pol II is higher across the whole length of a gene which contains an active 5'SS (Damgaard et al., 2008). The authors proposed that the 5'SS close to the promoter is required for transcription stimulation by specifically affecting the recruitment of transcription initiation factors.

A splicing regulator SRSF2 (originally known as SC35) was found to play a crucial role in transcription elongation (Lin et al., 2008). SRSF1 and SFSF2 proteins are able to associate with both promoter sequence and RNA near the TSS (Ji et al., 2013). It has been shown that SRSF2 promotes the escape of Pol II from pausing near promoter element by triggering the release of P-TEFb from the 7SK complex. Previously it has been known that inactive P-TEFb is sequestered in the 7SK complex (Peterlin and Price, 2006), and upon its release, P-TEFb becomes active and stimulates transcription elongation. In an RNA-dependent manner, SRSF2 stimulates the recruitment of P-TEFb and other key transcription elongation factors to promoters. The spliceosome component U2 snRNP has been demonstrated to be responsible for a transcription elongation checkpoint as Pol II accumulates over introns if the U2 snRNA is mutated (Chathoth et al., 2014). This study has proposed that the proper co-transcriptional formation of spliceosome triggers efficient transcriptional elongation.

1.4.3.5 *In vitro* co-transcriptional splicing

After the discovery of splicing, attempts were made to study the mechanism of splicing of pre-mRNA transcribed by Pol II in the NE (Weingartner and Keller, 1981; Kole and Weissman, 1982; Padgett et al., 1983), but the results were inconsistent and low splicing efficiencies were observed (Hernandez and Keller, 1983). The basic mechanisms of splicing reactions were studied using *in vitro* splicing systems in which pre-made pre-mRNA, previously transcribed by bacteriophage RNA polymerases (T7 or SP6), was added to HeLa NE for splicing to occur (Krainer et al., 1984). Using this assay splicing was studied independently of other processes including transcription. Significantly later an *in vitro* transcription/splicing reactions (TSRs) were set up to investigate the mechanisms of coupling between Pol II transcription and pre-mRNA splicing (Ghosh and Garcia-Blanco, 2000; Das et al., 2006; Hicks et al., 2006; Lazarev and Manley, 2007). Since the NE contain all necessary transcription and RNA processing factors, the

system allows both Pol II transcription and splicing to occur in the same extract. Previous studies tested three systems: Pol II-TSR supplemented with a DNA template carrying an appropriate promoter for Pol II; conventional splicing reaction (SR) in which pre-made transcripts are added; and T7-TSR in which transcription of a DNA template having a T7 promoter is driven by T7 RNA polymerase (T7 Pol) in the NE. T7 Pol was chosen as a control model for Pol II because it is a small (99 kDa) and fast (40 nucleotides per sec) enzyme which lacks any domains responsible for the recruitment of RNA processing factors (Skinner et at., 2004).

It was found that splicing of pre-mRNA transcribed by Pol II was faster and more efficient than that in either T7-TSR or conventional SR (Ghosh and Garcia-Blanco, 2000; Das et al., 2006). The yield of spliced mRNA, if pre-made pre-mRNA had been directly added to the NE or produced by T7 Pol in the NE, was low in comparison to Pol II counterparts. A fraction of splicing products was found to be associated with a DNA template in Pol II-TSR, while in the reaction where the NE was supplemented with T7 Pol splicing products were not determined to be associated with a DNA template (Ghosh and Garcia-Blanco, 2000; Yu et al., 2010). The authors concluded that splicing occurred while the nascent transcripts were still tethered to Pol II in *in vitro* TSR. The effect of Pol II transcription on spliceosome assembly was investigated (Das et al., 2006). There was a large spliceosome complex detected in both Pol II-TSR and SR, but only complex H was observed in T7-TSR. There was no complex H assembly found in Pol II-TSR. It was concluded that by coupling transcription to splicing, the pre-mRNA was not accessible for highly abundant splicing inhibitors such as hnRNP proteins and as a result complex H was not formed in Pol II-TSR. However, rather fast and efficient spliceosome assembly was found on transcripts produced by Pol II. A late study demonstrated a contradictory result where neither kinetics nor splicing efficiency were improved by Pol II in comparison to that for standard SR (Lazarev and Manley, 2007).

1.5 The U1 snRNP

1.5.1 The U1 snRNP structure and function in splicing

As all snRNPs, the U1 complex contains snRNA and a number of proteins (Figure 1.13; Pomeranz Krummel et al., 2009). The U1 snRNA has a sequence of 164 nt forming four stem loops (SL). The U1-specific proteins are U1-A, U1-70K, U1-C (Mount et al., 1983; Bringmann and Lührmann, 1986) and Sm proteins such as SmB/SmB′, SmD1, SmD2, SmD3, SmE, SmF and SmG which are shared between other snRNPs (reviewed by Guiro and O'Reilly, 2015). The U1-specific proteins have direct and indirect interactions with the U1 snRNA. Sm proteins form a ring around a portion of U1 snRNA between SL3 and SL4 (Pomeranz Krummel et al., 2009; Weber et al., 2010). With the help of RNA recognition motif (RRM), U1-70K and U1-A proteins bind to SLI and SLII, respectively (Pomeranz Krummel et al., 2009). U1-C protein does not have a direct interaction with the U1 snRNA, but it becomes a part of the complex through the protein-protein interactions with U1-70K and SmD3 (reviewed by Will and Lührmann, 2011). The maturation of the U1 snRNA occurs in the cytoplasm by the assembly of Sm proteins which is coordinated by the survival of motor neuron (SMN) complex (Fischer et al., 1997), while the U1-specific proteins are recruited to the U1 snRNA in the nucleus (reviewed by Guiro and O'Reilly, 2015).

The U1-associated proteins play crucial roles in 5'SS recognition (Krämer et al., 1984; Zhuang and Weiner, 1986). It has been shown that through RRM of U1-70K protein, U1 snRNP interacts with RRMs of splicing enhancers such as SR proteins (reviewed by Graveley, 2000; Roca et al., 2013). The disruption of U1-C interaction with the U1 snRNP complex prevents the U1 snRNA to bind to the 5'SS, suggesting that U1C contributes to 5' SS sequence specificity (Du and Rosbash, 2002). These results suggest that the function of U1-C is to stabilize the base pairing between the U1 snRNA and the 5'SS. The U1 snRNP and SR proteins have been found to be loaded onto the CTD (Das et al., 2007), suggesting their cooperation function in co-transcriptional spliceosome assembly.

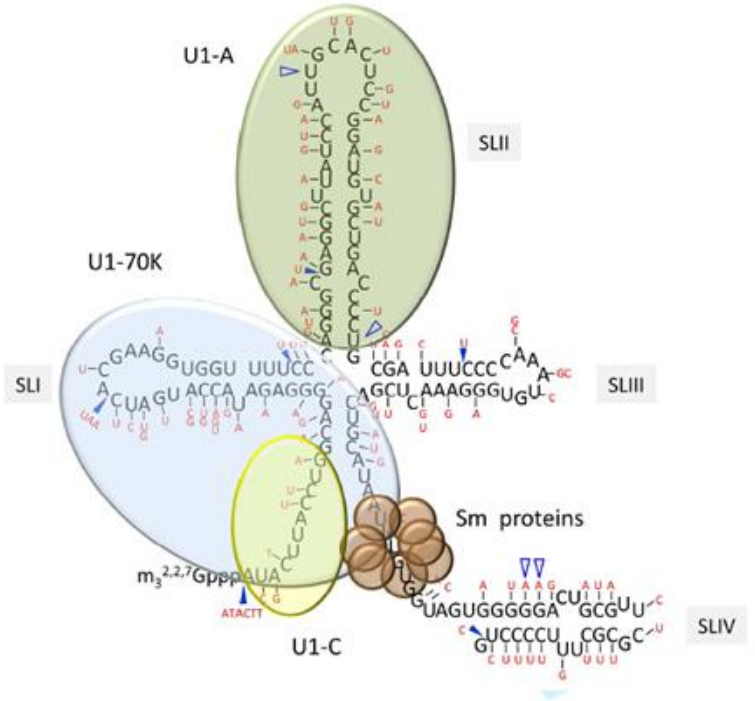


Figure 1.13. Secondary structure of the U1 snRNP (Guiro and O'Reilly, 2015). The U1 snRNA is depicted in black. Base changes and additional bases which can be found in U1 snRNA variants are shown in red and by blue triangles, respectively. A specific trimethyl cap structure of the U1 snRNA is shown as m32,2,7Gppp. The U1-specific proteins such as U1-70K, U1-A and U1-C are depicted in blue, green and yellow, respectively. Sm proteins are illustrated in brown.

Apart from the U1 snRNA gene located on the short arm of chromosome 1, 1p36 (Naylor et al., 1984), there is a large number of U1 snRNA pseudogenes which are spread throughout the genome (reviewed by O'Reilly et al., 2013). The U1 snRNA pseudogenes encode incomplete copies of the U1 snRNA leading to the production of U1 snRNA variants (Denison et al., 1981). A recent study has found that some U1 snRNA variants contribute to the regulation of 3' end processing (reviewed by O'Reilly et al., 2013).

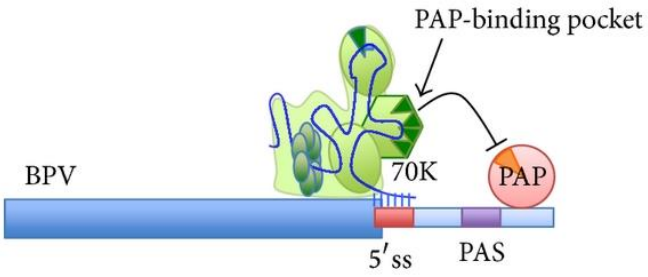
1.5.2 Splicing-independent roles of the U1 snRNP

The U1 snRNP complex exerts its main function in 5'SS selection where the 5' end of the U1 snRNA binds to the 5'SS of pre-mRNA (Krämer et al., 1984; Zhuang and Weiner, 1986). This interaction mediates a stepwise assembly of the other snRNPs that form the spliceosome complex (reviewed by Will and Lührmann, 2011). However, evidence that the abundance of the U1 snRNA far exceeds that of the other spliceosomal snRNAs (reviewed by Guiro and O'Reilly, 2015) led to the proposition that the U1 snRNP can have additional roles in RNA processing than just a single function in pre-mRNA splicing. Indeed, it has been shown that the U1 snRNP has a crucial non-splicing role, which is the regulation of transcript length (Gunderson et al., 1998; Kaida et al., 2010). For premature termination to occur, cryptic PASs located every 500–800 nucleotides have to be used (Kaida et al., 2010). A genomic tiling microarrays analysis and sequencing study have shown that the functional knockdown of the U1 snRNA by using antisense morpholino oligonucleotides (AMO) led to the accumulation of unspliced and short polyadenylated transcripts (Kaida et al., 2010). It implies that transcription did not produce full length pre-mRNAs because they had been prematurely cleaved and polyadenylated (PCPA). This effect was not observed when the U2 snRNP activity had been inhibited, suggesting that only U1 snRNP plays a vital role in the suppression of cryptic PAS usage and premature cleavage. This has been proven by showing the production of long mRNAs upon overexpression of the U1 snRNA (Berg et al., 2012). When the 5'SS upstream of the alternative PAS has been blocked the PAS was activated (Vorlova´ et al., 2011). In contrast, the inhibition of the downstream 3'SS had no effect on PAS activity. A recent study has shown a dramatic increase of Ser2 phosphorylation level of the CTD close to PAS upon the inhibition of U1 snRNA (Davidson et al., 2014). These data supports the notion that the U1 snRNP has a crucial role in regulating the 3' end processing in a transcription-dependent manner.

It has been reported that U1-70K protein through its RRM domain physically interacts with PAP abolishing its ability to add a poly(A) chain to the 3' end of pre-mRNA (Figure 1.14 A; Gunderson et al., 1998; Vagner et al., 2000). Upon knockdown of U1-70K the usage of alternative PASs occurred. Also, U1-70K protein has been found to associate with components of the cleavage factor1 (CF1; Awasthi and Alwine, 2003).

In addition to U1-70K protein, the PCPA can be controlled by U1-A protein (Boelens et al., 1993; van Gelder et al., 1993). The interaction of U1-A with pseudo U1 snRNA binding elements, which are located between two GU-rich motifs, prevents pre-mRNA cleavage at pseudo PASs (Figure 1.14 B; Spraggon and Cartegni, 2013). Based on the fact that there are numerous 5'SS variants spread across the human genome (Roca et al., 2012) and that high complementarity of the U1 snRNA to bind pre-mRNA is not required, it has been suggested that the U1 snRNP targets pre-mRNA sequences and enables suppression of alternative PAS usage by competing with factors of the polyadenylation machinery (Kaida et al., 2010; Almada et al., 2013).

A



B

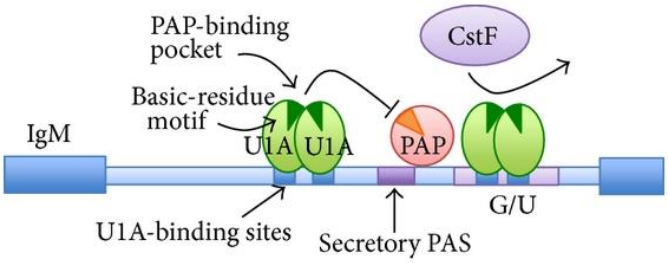


Figure 1.14. Schematic representation of the mechanisms how the U1 snRNP prevents premature cleavage and polyadenylation (Spraggon and Cartegni, 2013). (A) Once the U1 snRNP is bound to the 5'SS, U1-70K is able to inhibit the function of Poly(A) polymerase (PAP). (B) U1-A specific protein of the U1 snRNP complex prevents PAP and CstF to bind to a poly(A) signal (PAS) and G/U sequence, respectively.

Since transcription driven by Pol II occurs both in the sense and antisense directions from active promoters (Core et al., 2008; Seila et al., 2008; Grzechnik et al., 2014), it has been proposed that the U1 snRNP might regulate the synthesis of the upstream antisense RNAs. Indeed, a number of U1 snRNA binding sites is smaller in the antisense direction in comparison to their abundancy in the sense direction (Almada et

al., 2013). This indicates that the U1 snRNP is less recruited to RNA produced upstream from the promoter (Figure 1.15), leading to transcription termination and polyadenylation. It has been determined that such short transcripts are not used for further process because they are targeted for degradation by exosomes (Preker et al., 2008). There is an abundant amount of U1-snRNA-like sequences across the downstream part of a pre-mRNA from its 5' end. This allows the U1 snRNP be recruited to these sequences. Therefore, a suppression of cryptic PASs usage is regulated by the U1 snRNP, allowing transcription downstream from the promoter to be carried out efficiently. Together, the studies demonstrated that the U1 snRNP appears to play splicing-independent roles in facilitating proper gene expression.

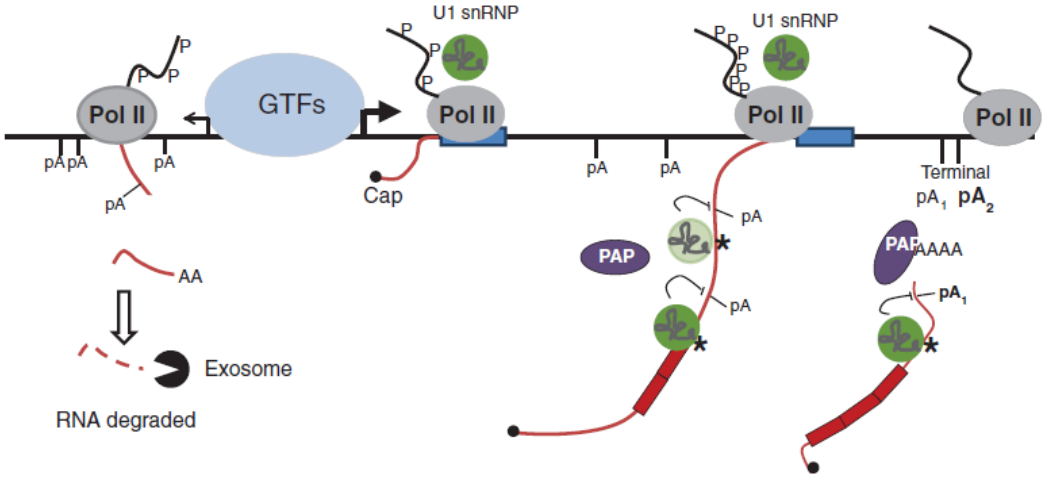


Figure 1.15. The function of the U1 snRNP to inhibit polyA site selection (Guiro and O'Reilly, 2015). U1 snRNP binding sites (*) are located not only at exon/intron junctions but within introns and untranslated regions. There are alternative poly(A) (pA) sites in the pre-mRNA. Through the interaction with the pre-mRNA, the U1 snRNP prevents the recognition of cryptic poly(A) sites leading to a normal transcript synthesis. There are less U1 snRNP binding sits in the antisense direction of a gene in respect to a promoter position. Transcripts are prematurely cleaved if transcription occurs in the antisense direction because there is less bound U1 snRNP.

1.5.3 The U1 snRNP and transcription factors

Some studies have shown links between the U1 snRNP and Pol II transcription machinery (Figure 1.13), and suggested possible roles of the U1 snRNP in transcription regulation. One of the first observations revealed that the U1 snRNP via U1-70K protein interacts with the human transcription elongation factor TAT-SF1 (Fong and Zhou, 2001). TAT-SF1 can interact with P-TEFb complex which is a component of transcription elongation apparatus functioning in phosphorylation of the CTD at Ser2 (Zhou and Sharp, 1996; reviewed by Price, 2000). The study showed that transcript level was reduced upon inhibition or depletion of the U1 snRNA from TAT-SF1-U1 snRNP complex. Therefore, the data was interpreted as the U1 snRNP could stimulate Pol II transcription elongation. A recent *in vivo* study showed that the U1 snRNP inhibition led to a reduced level of transcripts (Koga et al., 2015), representing a reproducibility of the effect between *in vivo* and *in vitro*.

The next study showed that the U1 snRNA associates with the general transcription factor TFIIH (Kwek et al., 2002; Figure 1.13), which phosphorylates the CTD of Pol II during transcription initiation (Feaver et al., 1994; Akoulitchev et al., 1995). It was shown that the removal of the U1snRNA from the TFIIH-U1 snRNP complex led to the reduction of TFIIH activity and the efficiency of transcription reinitiation was dependent on both the U1 snRNA and the promoter proximal 5'SS (Kwek et al., 2002). Based on these results the authors proposed a potential role of the U1 snRNA in transcription initiation. Later it was highlighted that the CDK7-associated cyclin H (cyclin H), a component of TFIIH, interacts with the U1 snRNA (O'Gorman et al., 2005). The SLI and SLII of the U1 snRNA were responsible for this association. The study showed that the U1 snRNA played a role in supporting the kinase activity of TFIIH complex.

In addition, it has been found that the U1 snRNA and U1 snRNP-specific components U1-70K, U1-C, and Sm can associate with TBP-associated factor 15 (TAF15; Jobert et al., 2009; Leichter et al, 2011), which is a component of TFIID complex (Bertolotti et al, 1996). However, a functional property of the U1 snRNP in transcription regulation was not shown in these studies. The association between the U1 snRNP and Pol II has also been found (Tian, 2001; Robert et al., 2002; Yu and Reed, 2015). The protein

fused-in-sarcoma (FUS) is responsible for bringing the U1 snRNP to Pol II (Yu and Reed, 2015; Figure 1.16). Another study has shown that the distribution of Pol II across genes was not changed upon U1 snRNP inhibition (Davidson et al., 2014). Also, there was no change in the CTD phosphorylation level of either Ser5 or Ser7 which are responsible for transcription initiation. Together these data suggest that the U1 snRNP might not be required for transcription regulation but transcript stability.

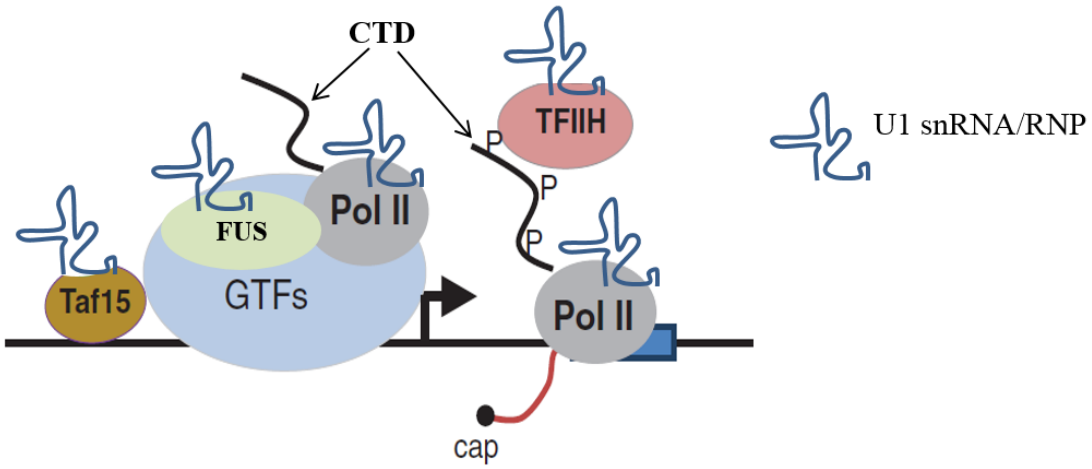


Figure 1.16. The U1 snRNA/RNP interacts with different transcription factors (Guiro and O'Reilly, 2015). GTFs stands for general transcription factors.

1.6 MECHANISMS OF RNA DEGRADATION

Gene transcripts are generated as large precursors which are processed extensively to the functional forms. Finally, all RNAs undergo degradation. Messenger RNAs are destroyed in the cytoplasm. However, errors can occur during RNA maturation steps, folding or mRNA export. Surveillance machineries search for such defective transcripts and degrade them in both the cytoplasm and nucleus. This indicates that cells have acquired quality control mechanisms to ensure that only high-quality RNAs are used for protein synthesis. However, RNA degradation occurs not only to discard RNA after translation or those which failed processing but it is a key mechanism in transcription termination (Gromak et al., 2006; West et al., 2008). Also, degradation of some introns leads to production of small nucleolar RNAs and microRNAs (Schmid and Jensen, 2008).

General RNA turnover through degradation is ubiquitous (Schmid and Jensen, 2008). There are three classes of RNA-degrading enzymes: endonucleases, 5' exonucleases and 3' exonucleases. Exonucleases function on transcripts with accessible single-stranded termini. Endonucleases catalyse the internal cleavage of RNA, generating substrates subject to digestion by 5' and 3' exonucleases. The activity of 5' exonucleases is stimulated by the presence of a 5'-monophosphate (Schmid and Jensen, 2008). Each class of nucleases contains numerous factors. Some of them function during transcription. If one enzyme loses its activity, this does not prevent RNA degradation, indicating that the same RNA can be a substrate for other enzymes (Vanacova et al., 2007). The main degradation pathways are described in this section.

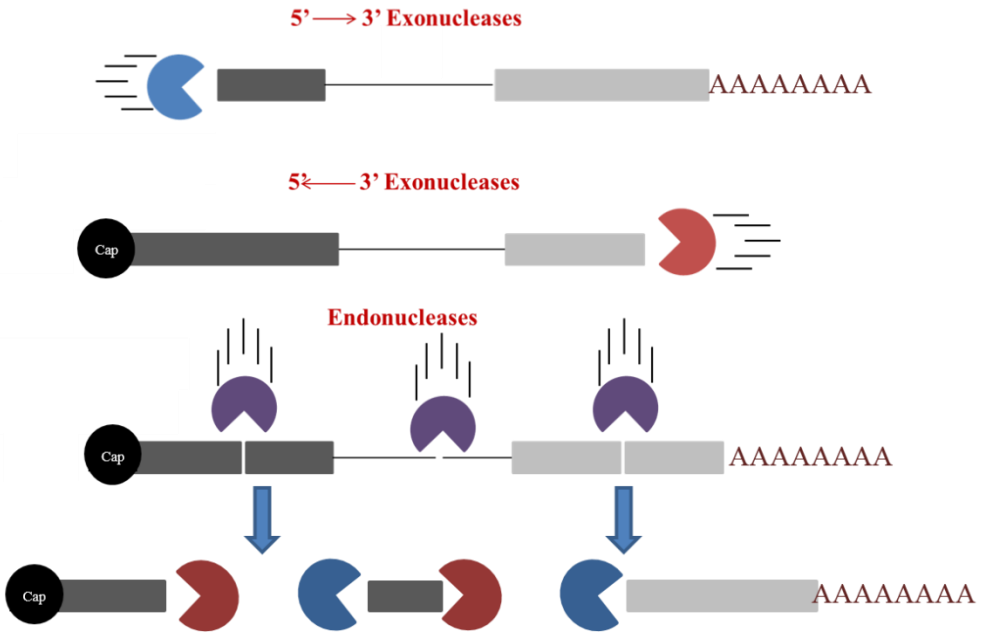


Figure 1.17. RNA degradation pathways. 5' and 3' exonucleases target transcripts through their accessible ends while endonucleases perform internal cleavage.

One of the most studied nuclear 5'-3' exonucleases is Xrn2 (exoribonuclease 2; Zhang et al., 1999). There are different mechanisms that can create entry sites for Xrn2. For example, during 3' end processing and transcription termination, when cleavage occurs, RNA that is attached to the transcribing polymerase is degraded by Xrn2 (Gromak et al., 2006; West et al., 2004). The removal of the cap structure usually occurs in the cytoplasm after translation is completed (Dunckley and Parker, 1999; Wang et al., 2002) but in a situation of capping failure, transcription elongation does not proceed (Guiguen et al., 2007) and decapping is carried out in the nucleus (Liu et al., 2004; Brannan et al., 2012). At this stage RNA undergoes degradation in a 5' to 3' direction by Xrn2 (Brannan et al., 2012). Xrn2 has also been implicated in degradation of aberrantly processed transcripts in the nucleus (reviewed by Houseley and Tollervey, 2009). If pre-mRNAs harbour mutations in splice sites and splicing is invalid it has been shown that such transcripts are subject to degradation by Xrn2 (Davidson et al., 2012). Transcripts cleaved by endonucleases are potential Xrn2 substrates (reviewed by Houseley and Tollervey, 2009).

In eukaryotes 3’ end degradation is facilitated by the exosome complex which also exhibits endonuclease activity (Mitchell et al., 1997). The exosome plays a role in RNA biogenesis (Andrulis et al, 2002) and is responsible for degradation of the majority of transcripts generated by RNA polymerases I, II and III (Allmang et al., 1999a; reviewed by Houseley and Tollervey, 2009). The exosome is functional in the cytoplasm and nucleus. Aberrant pre-mRNAs that result from failed splicing, 3’ end processing or export are substrates for the exosome (Bousquet-Antonelli et al, 2000; Libri et al, 2002; Wyers et al., 2005; Das et al., 2006). In the nucleus the exosome localizes at the sites of transcription (Andrulis et al, 2002). This complex has a core that is composed of nine protein subunits (Liu et al., 2006; Hernandez et al., 2006). The core components themselves are inactive and the exosome becomes functional when exonucleases Rrp44 and Rrp6 join the complex (Dziembowski et al., 2007; Lorentzen et al., 2008; Allmang et al., 1999b). Rrp44 degrades structured and unstructured transcripts (Lee et al., 2012) while Rrp6 has a preference for unstructured substrates (Schmid and Jensen, 2008).

There is an additional surveillance mechanism called nonsense-mediated decay (NMD). NMD targets newly transcribed mRNAs that have acquired premature translation termination codons (PTCs) as a result of a point mutations or inclusion of an alternative exon containing PTC (He and Jacobson, 2015; Lewis et al., 2003). NMD prevents translation of defective mRNAs that could be used for synthesis of abnormal proteins. This translation-dependent process requires NMD-specific *trans*-acting factors such as UPF1, UPF2 and UPF3 (Culbertson et al., 1980; Conti and Izaurralde, 2005). Initially, it was thought that NMD exclusively occurs in the cytoplasm but it has been shown that NMD can occur in the nucleus (Muhlemann et al., 2001).

1.7 AIMS

The aims of this project were to establish *in vitro* transcriptional/splicing assays in which transcription would be carried out by RNA polymerase II (Pol II) or T7 RNA polymerase in nuclear extract. Test whether Pol II transcriptional/splicing system (Pol II-TSR) is able to recapitulate coupled reaction between the two processes. Investigate whether the *in vitro* Pol II-TSR is a suitable system to address questions such as: is Pol II transcription machinery able to affect splicing yield of constitutively and alternatively spliced subtracts, splicing kinetics and spliceosome assembly. Identify factors that mediate functional coupling between transcription and splicing.

2 MATERIALS AND METHODS

2.1 General molecular biology techniques

2.2 *In vitro* RNA techniques

2.3 Mammalian cell biology techniques

2.1 General molecular biology techniques

2.1.1 Exonuclease III cloning

Cloning procedure involved amplification of insert DNA by polymerase chain reaction (PCR), gel purification of PCR product, double restriction enzyme digestion of EFGP-C1 vector, Exonuclease III treatment of vector and PCR product. An insert and the vector were ligated prior transformation. To determine the presence of insert DNA in the plasmid, a colony screening PCR was done. This followed by plasmid isolation from bacterial cultures. Each procedure is outlined in details below.

2.1.1.1 Oligonucleotide design

All genes of interest were incorporate into AgeI and HindIII sites of a EGFP-C1 vector (Clontech). Overlapping oligonucleotides were designed to be about 50 bp in length. The insert complimentary part was designed to have approximately AT:CG ratio of 50%, ensuring an annealing temperature to be around 55-60°C. Oligonucleotides were synthesized and purified by Invitrogen. They were dissolved in TE.1 buffer (10 mM Tris pH 7.5, 0.1 mM EDTA pH 8) to stock concentrations of 100 µM and stored at -20 ºC. Oligonucleotides used to amplify Bg, RON and Bcl-X are listed in Table 2.1. To clone CD44v5 and Bg mutants Bg2 F VAP and Bg3 R primers were used.

Name	Sequence
Bg2 F VAP	TAGTGAACCGTCAGATCCGCTAGCGCTACCGGTGGGCTGCTG GTTGTCTACCCA
Bg3 R	GTCGACTGCAGAATTCGAAGCTAACTTACCTGCCAAAATG ATGAGACAGC
RON10 F	TAGTGAACCGTCAGATCCGCTAGCGCTACCGGTTGTGAGAGG CAGCTTCCAGAG
RON12 R	TCCTCGCCCTTGCTCACCATGGTGGCGACCGGTCTAGCTGCT TCCTCCGCCA
Bcl-X F	TAGTGAACCGTCAGATCCGCTAGCGCTACCGGTGGGAGGTGA TCCCCATGGCAG

Bcl-X R	GTCGACTGCAGAATTCGAAGCTTACTTACCTGGCCACAGTCAT

Table 2.1 List of primers for PCR amplification of DNA templates for Exonuclease III cloning.

2.1.1.2 DNA amplification by PCR

All amplifications were carried out in a Storm PCR thermocycler. To produce insert DNA for cloning, Phusion high-fidelity DNA polymerase (NEB, UK) was used. A 50 µl volume reaction contained 50 ng of highly pure plasmid, 1 µM each forward and reverse primers, 1x buffer, 0.4 mM dNTPs (Promega) and 1 unit (U) DNA polymerase. Cycle conditions were set up according to manufacturer's recommendations. Annealing temperature varied depending on oligonucleotides sequence composition. PCR was done during 35 cycles. Gel electrophoresis was done to confirm the presence of PCR products.

2.1.1.3 Agarose gel electrophoresis

The total PCR reaction was combined with 6x loading buffer (NEB) in appropriate proportions. Samples were resolved by electrophoresis on 1% agarose gel (Helena) containing ethidium bromide (0.5 mg/ml) in 1x TAE buffer. To determine the size of amplified DNA fragments, 2 µl of 1 kb DNA ladder (NEB) were used. Electrophoresis was performed at 80 V for 45 min. PCR products were analysed under UV light.

2.1.1.4 Agarose gel extraction of PCR products

The PCR products were excised from gel with a clean scalpel. The DNA fragments were extracted from gel slices by using QlAquick gel extraction kit (Qiagen) following the manufacturer's instructions. Purified DNA templates were eluted in 30 µl of TE.1 Buffer and their concentrations were determined by measuring A_{260} using nanodrop.

2.1.1.5 Double restriction enzyme digestion

3 µg of EGFP-C1 plasmid was digested with 3 U of each restriction endonucleases AgeI and XindIII (NEB) in 30 µl reaction mixture in the presence of 1x Buffer I (NEB). The

reaction was incubated at 37 ºC for 2 h. The enzyme inactivation was achieved by incubating the reaction at 65 ºC for 20 min. The size of the digestion products was confirmed by agarose gel electrophoresis as described in Section 1.1.3.

2.1.1.6 Exonuclease III digestion and products ligation

Exonuclease III reaction was carried out according to previously optimised conditions in the laboratory. A standard reaction mixture was 10 µl in volume and contained 300 ng vector and purified insert in a molar ratio of 1:1-1:3 and 1x Buffer I (NEB). 1 U of freshly diluted Exonuclease III enzyme (NEB) in Buffer I was added last. To allow step-wise release of nucleotides from the 3' end of DNA double strands by the enzyme, the reaction was incubated at 14 ºC for 1 min. The reaction was terminated with "Stop" buffer containing 50 µl of 10 mM Tris pH7.5 and 4 mM EDTA. Samples were span with 1/10 volume 3M NaAc and 3 volume 100% ethanol at room temperature (RT), 13000 rpm for 15 min. A pellet was washed with 200 µl of 70% ethanol and span again under the same conditions. After drying DNA was resuspended in 10 µl TE.1 buffer. To anneal the products, the reaction was incubated at 55 ºC for 3 min and then at 37 ºC for 3 min. The ligation reaction was transformed into competent cells.

2.1.2 TOPO vector cloning

Zero Blunt TOPO PCR Cloning Kit (Invitrogen) was used to clone PCR amplified U1 snRNA gene. 3 µl reaction containing 50 ng PCR product and 0.5 µl of each TOPO vector and salt solution (supplied) was incubated at RT for 5 min. This was followed by transformation.

2.1.3 Transformation of chemically competent cells

5 µl of cloning reaction was added to 50 µl of chemically competent TOP10 *E. coli* cells (Invitrogen). After incubation on ice for 5 min, the sample was heat shocked at 42 ºC for 40 sec and immediately placed on ice. After 2 min, 200 µl of Luria Bertani medium (LB; 10 g Bacto-tryptone, 5 g Yeast extract, 10 g NaCl in 1 L H_2O) was added to the cells. The tubes were incubated at 220 rpm, 37 ºC for 1 h. 200 µl of each sample were spread onto pre-warmed selective LB plates (antibiotic at 30 µg/ml) following the

incubation at 37 ºC for 12-14 h. Colonies were tested for the presence of a DNA insert by PCR.

2.1.4 Colony screening PCR

Colony screening PCR was done to identify colonies which contained the insert. Red Taq DNA polymerase (Sigme) was used. The same primers designed to amplify mini-genes for *in vitro* transcription were used for colony screening PCR (Table 2.2). A 10 µl reaction mixture contained 1 mM of each forward and reverse primers, 1x buffer, 0.4 mM dNTPs and 0.5 U of Red Taq DNA polymerase. Individual colonies were streaked on a selective agar plate. The remaining colony was added directly to the PCR reaction. Cycle conditions were set up according to manufacturer's recommendations, primer annealing temperature and a length of the insert. The number of cycles was 35. On completion PCR reactions were analysed by agarose gel electrophoresis (Section 1.1.3). The plate was incubated for 12 h at 37 ºC. Successful colonies were used for plasmid isolation.

2.1.5 Small scale plasmid isolation

A successful colony was inoculated into 5 ml of LB medium containing the appropriate antibiotic (30 µg/ml). The cells were cultured at 37 ºC, 220 rpm for 12 h. The cells were harvested by centrifugation at 4 ºC, 4600 rpm for 10 min. 200 µl of Solution 1 (25 mM Tris-HCl pH 8.0, 10 mM EDTA, 50 mM glucose and 6 µ/ml RNase A (NEB) was added to resuspend a cell pellet following incubation at RT for 10 min. 400 µl of fresh Solution 2 (1% SDS and 0.2M NaOH) was added. The samples were mixed and incubation on ice for 5 min. 300 µl of Solution 3 (3M KAc pH 4.8) was added, mixed and incubated on ice for 10 min. After centrifugation at 4 ºC, 13 000 rpm for 10 min, 400 µl of the supernatant was taken to precipitate plasmid with 100% ethanol. Plasmid pellet was washed with 70% ethanol. DNA was resuspended in 100 µl of water prior phenol chlrophorm extraction and ethanol precipitation (Section 2.1.11). Plasmid was resuspended in 50 µl of TE.1 buffer. Plasmid concentration was determined by measuring A_{260} using nanodrop machine with further sending for sequence verification.

2.1.6 Sequence analysis

250 ng of each plasmid DNA was used per one sequencing reaction. Primers complementary to the vector (Frw: GCCACCTCTGACTTGAGCGT; Rev: AGGTCAGGGT GGTCACGAG) were supplied at the concentration of 1 pmol/µl. Sequencing was done by Protein Nucleic Acid Chemistry Laboratory (PNACL; University of Leicester).

2.1.7 Growing calcium competent cells

5 ml bacterial cell culture was incubated at 37 ºC, 220 rpm overnight (O/N). The culture was diluted 1:100 in fresh LB medium (100 ml) and grown until OD_{600} reached 0.5. Cells were collected by centrifugation for 10 min at 3000 rpm and 4 ºC. This was followed by resuspending a cell pellet in 25 ml of 0.1 M $CaCl_2$ and incubated on ice for 30 min. Cells were collected by centrifugation under the same conditions and resuspended in 4 ml of 0.1 M $CaCl_2$ and 15% glycerol. 50 µl aliquots were frozen using dry ice and industrial methylated spirit and stored at -80 ºC.

2.1.8 Routine PCR

Red Taq DNA Polymerase (Sigma) was used to optimize PCR conditions, amplify a gene of interest after reverse transcription and generate DNA templates for *in vitro* transcription and transcription/splicing assays. A 50 µl reaction mixture contained 1 mM of each forward and reverse primers, 1x buffer, 0.4 mM dNTPs, 50 ng of DNA and 1 U of Red Taq DNA polymerase. The reaction was prepared on ice. Cycle conditions were set up according to manufacturer's recommendations, primer annealing temperature and a length of a gene. The number of cycles was 35. On completion PCR reactions were analysed by agarose gel electrophoresis (Section 1.1.3).

Oligonucleotide name	Oligonucleotide sequence
P16	AAATTAATACGACTCACTATAGGGCTGCTGGTTGTC TACCCA
P17	AACTTACCTGCCAAAATGATGAGACAGCAC
CMV Frw	ATAGTAATCAATTACGGGGTCATTAGTT
Bg 3 Rev	TGCCAAAATGATGAGACAGCACAA
RON 10 Frw	AAATTAATACGACTCACTATAGGGTGTGAGAGGCA GCTTCCAGA
RON 12 Rev	CTAGCTGCTTCCTCCGCCAC
T7-GFP	AAATTAATACGACTCACTATAGGGCGCCACCATGGT GAGCAA
GFP Rev	CAGCTCGATGCGGTTCACCA
GFP-U1 BS Rev	AACTTACCTCAGCTCGATGCGGTTCA
T7-Bcl-X Frw	AAATTAATACGACTCACTATAGGGCGTGATCCCCAT GGCA
Bcl-X Rev	TACTTACGTCATGCCCTGGCCACA
T7-U1 Frw	AAATTAATACGACTCACTATAGGATACCTTACCTGG CAGGGGAGATACCA
U1 Rev	CAGGGGAAAGCGCGAACGCA

Table 2.2. List of primers used for PCR amplification of DNA templates for *in vitro* transcription and co-transcriptional splicing.

2.1.9 PCR with radiolabelled nucleotides

PCR with radiolabelled dTTP nucleotide was done to label DNA templates for DNA-protein complexes analysis. 0.5 µl of [α-^{32}P] dTTP (10 mCi/ml, 3000 Ci/mmol; PerkinElmer) was added to a 50 µl Red Taq PCR reaction. The components of the reaction mixture are described in Section 2.1.8. On completion PCR reactions were analysed by agarose gel electrophoresis (Section 1.1.3).

2.1.10 Phenol-chloroform extraction and ethanol precipitation of PCR products

The procedure was performed at RT. PCR products were diluted in autoclaved ultra-pure water reaching 100 µl. 1 volume of Phenol: Chloroform: Isoamyl alcohol (25:24:1) was added to the samples, mixed by vortexing and centrifuged for 3 min at 13000 rpm. The upper aqueous phase was collected in a fresh tube and this step was repeated. In a fume hood 100 µl of chloroform was added to the sample, mixed by vortexing and span under the same conditions. This step was repeated. The volume of the samples was assessed. 1/10 volume of 3M NaAc pH 5.2 and 2 volumes of 100% ethanol were added

and mixed by shaking. After centrifugation at 13000 rpm for 15 min, the supernatant was removed. A DNA pellet was washed with 100 µl of 70% ethanol and span applying the same conditions. After drying, DNA was resuspened in 20 µl of TE.1 buffer. The samples were stored at -20 ºC.

2.2 *In vitro* RNA techniques

2.2.1 *In vitro* transcription with radiolabelled nucleotides ("Hot transcription")

Transcription reactions were prepared at RT. DNA fragments containing the T7 promoter at the 5' end were generated by PCR. The homemade T7 RNA polymerase was used to generate transcripts for *in vitro* splicing assays. The reaction mixture (10 µl) contained transcription buffer (40 mM Tris pH 7.5, 6 mM $MgCl_2$, 10 mM NaCl), 0.5 mM ATP, 0.5 mM CTP, 0.5 mM UTP, 0.05 mM GTP (Promega), 50 ng DNA template, 10 U RNaseOut (Invitrogen), 5% T7 Pol (1:20 dilution), 1 mM Ribo m7G cap analogue (Promega) and 0.5 µM [α-^{32}P] GTP (10 mCi/ml, 3000 Ci/mmol; Pelkin Elmer). After incubation at 37 °C for 2 h, samples were mixed with 10 µl formamide dyes (90 % v/v formamide, 50 mM EDTA, bromophenol blue and xylene cyanol) and fractionated on 6% denaturing polyacrylamide gel electrophoresis (Section 2.2.3). RNA was visualized using x-ray film (Fujifilm) and excised with a clean scalpel. RNA was eluted O/N at 4 ºC using 300 µl RNA Elution buffer (0.5 M NaCl pH 4, 0.2 % SDS, 1 mM EDTA 0.5 M pH 8).

2.2.2 *In vitro t*ranscription without radiolabelled nucleotides ("Cold transcription")

Transcription reactions were prepared at RT. DNA fragments containing the T7 promoter at the 5' end were generated by PCR. The home-made T7 RNA polymerase was used. The reaction mixture (100 µl) contained 40 mM Tris (pH 7.5), 6 mM MgCl2, 10 mM NaCl, 0.5 mM ATP, 0.5 mM rCTP, 0.5 mM rUTP, 0.05 mM rGTP, 1 µg DNA template, 10U RNaseOut (Invitrogen), 5% T7 Pol (1:20 dilution) and 1 mM Ribo m7G cap analogue (Promega). After incubation at 37 °C for 4 h, DNA template was digested by RNase A (10 mg/ml). Samples were purified using S-300 column (GE Healthcare)

according to manufacturer's recommendations. RNA was phenol-chloroform extracted and ethanol precipitated (Section 2.1.11). The pellet was dissolved in 30-50 µl TE.1 buffer adding 10U RNaseOut. RNA was stored at -80 °C.

2.2.3 Denaturing polyacrylamide gel

20 ml and 40 ml gels contained 7 M Urea, 6 % acrylamide 19:1 bisacrylamide solution (National Diagnostics) and 1X TBE (89 mM Tris base, 89 mM boric acid, 2 mM EDTA). Polymerisation was achieved using 0.1% ammonium persulfate (AMPS) and 0.002% TEMED (Sigma). Formamide dyes (10µl) were added to samples prior loading onto the gel. The electrophoresis was carried out for various times depending on the length of RNA. Splicing gels were dried and exposed to Phosphor screens (Packard) for analysis.

2.2.4 RNA purification from polyacrylamide gel

RNA eluted O/N was incubated at 30 ºC for 5 min. The liquid phase was transferred into a fresh tube. This was followed by the addition of 900 µl 100% ethanol and centrifugation at 13000 rpm, RT for 15 min. The pellet was washed with 1 ml 100% ethanol and spun again. After drying, RNA was dissolved in 30-50 µl TE.1 buffer adding 10U RNaseOut. Transcripts were stored at -80 °C.

2.2.5 *In vitro* standard *s*plicing assay

Samples contained 1.5 mM ATP, 20 mM creatine phosphate, 3.2 mM $MgCl_2$, 20 mM Hepes (pH 7.5), 20 mM CrP, 50 mM KGlu, 0.05% NP-40, 30% nuclear extract (Cilbiotech) and 1 µl of RNA (10-20 cps). For time-course and triplicate assays a reaction volume was 10 µl and 5 µl, respectively. The reaction mixtures were set up on ice and incubated at 30 °C for 2 h. Aliquots of 2 µl were taken into a microtitre plate on dry ice at 0, 30, 60 and 120 min. Samples were proteinase K treated, ethanol precipitated, dissolved in 10 µl of formamide dyes and heated at 80 °C for 30 sec. RNA was fractionated on 6% denaturing polyacrylamide gel electrophoresis (Section 2.2.3). Quantifica-

tion of pre-mRNA and splicing products was done using OptiQuant software. The intensities of each band was normalized according to the number of guanines in each RNA specie.

2.2.6 *In vitro* Pol II transcription/splicing assay

Samples contained 1.5 mM ATP, 20 mM creatine phosphate, 3.2 mM $MgCl_2$, 20 mM Hepes (pH 7.5), 20 mM CrP, 50 mM KGlu, 0.05% NP-40, 30% nuclear extract (Cilbiotech), 10 U RNaseOUT (Invitrogen), 0.2 μM [α-^{32}P] GTP and 1 μl of CMV-DNA. The concentration of DNA template was 10 ng/μl. For time-course and triplicate assays a reaction volume was 10 μl and 5 μl, respectively. The reaction mixtures were set up on ice and incubated at 30 °C for various times. Aliquots of 2 μl were taken into a microtitre plate on dry ice. The rest of the procedure was the same as described in Section 2.2.5.

2.2.7 *In vitro* T7 transcription/splicing assay

Samples contained 1.5 mM ATP, 0.5 mM rCTP, 0.5 mM rUTP, 0.05 mM rGTP, 20 mM creatine phosphate, 3.2 mM $MgCl_2$, 20 mM Hepes (pH 7.5), 50 mM KGlu, 20 mM CrP, 0.05% NP-40, 30% nuclear extract (Cilbiotech), 10U RNaseOUT (Invitrogen), 0.2 μM [α-^{32}P] rGTP (10 mCi/ml, 3000 Ci/mmol; Pelkin Elmer), 5% T7 Pol (1:20 dilution) and 1 μl of T7-DNA. The concentration of DNA template was 10 nM. For time-course and triplicate assays a reaction volume was 10 μl and 5 μl, respectively. The reaction mixtures were set up on ice and incubated at 30 °C for various times. Aliquots of 2 μl were taken into a microtitre plate on dry ice. The rest of the procedure was the same as described in Section 2.2.5.

2.2.8 Inhibition of snRNAs using antisense 2'-*O*-methyl oligonucleotides

To block the U1, U2 and U6 snRNAs 2'-*O*-methyl modified oligonucleotides were used at concentrations 5 μM, 4 μM and 2 μM, respectively. These oligonucleotides were added to either standard splicing or Pol II/T7 transcription/splicing assays. Samples were incubated for 15 min at 30 °C following the addition of RNA or DNA. The procedures for SR, Pol II- and T7-TSR were the same as described in the above sections. The

concentrations of anti-U1/U2 and U6 2'-*O*-methyl oligonucleotides (in bold; Table 2.4) were 5 μM, 4 μM and 2 μM, respectively. Some batches of the NEs required to increase the concentrations of these oligos.

2.2.9 Oligonucleotide-directed RNase H digestion

DNA antisense oligonucleotides against the U1 (1-14), U2 (L15) and U6 (MB) snRNAs (Table 2.4) were used in HeLa NEs. 10 μl of reaction mixtures (SR or Pol II-TSR) including the NE were supplemented with 0.2 U RNase H (NEB). Samples were incubated for 15 min at 30 °C following the addition of RNA or DNA. The procedures for SR and Pol II-TSR were the same as described in the above sections.

Name	Sequence
anti-U1	**UGCCAGGUAAGUAU**
anti-U2	**CAGAUACUACAGUUG**
anti-U6	**CUGUGUAUCGUUCCAAUUUU**
1-14	TGCCAGGTAAGTAT
L15	CAGATACTACAGTTG
MB	CTTCTCTGTATCGTTCCAATT

Table 2.3. Oligonucleotides complementary to the U1, U2 and U6 snRNAs. Sequences in bold indicate 2'-*O*-methyl modification.

2.2.10 Pulse chase of transcription/splicing assays

Pol II-TSR and T7-TSR assays were set up as described in Section 2.2.6 and Section 2.2.7. After incubation at 30 °C, unlabelled nucleotides (GTP) were added to reach the final concentration of 5 mM. The reactions were mixed well following further incubation.

2.2.11 Antarctic Phosphatase and XRN-1 treatment

10 ng of radiolabelled RNA were added to a mixture containing 3.2 mM $MgCl_2$, 20 mM Hepes (pH 7.5), 50 mM KGlu, 0.05% NP-40, 0.1 mM $ZnCl_2$, 1U of XRN-1 and 1.5U of Antarctic Phosphatase enzymes. The reaction volume was 10 μl. Samples were incubated for 1 h at 30 °C. The efficiency of RNA digestion was analysed by 6% PAGE.

This was followed by phenol-chloroform extraction and ethanol precipitation of RNA (Section 2.1.11). Dry RNA pellet was resuspended in TE.1 buffer and stored at -80 °C.

2.2.12 Proteinase K treatment and ethanol precipitation

Time points were treated with 50 µl Proteinase K buffer (100 mM Tris-HCL pH 7.5, 12.5 mM 0.5 M EDTA, 150 mM NaCl, 1 % SDS) supplemented with proteinase K (0.4 mg/ml; Roche) for 15 minutes at 37 °C. 100 µl of 100% ethanol was added to the samples following centrifugation at 62000 rpm, RT for 15 min (Rotonta 460 R). The liquid was discarded using water pump. RNA was washed with 150 µl 100% ethanol and span again. 10 µl of formamide dyes were added to each sample prior loading on denaturing polyacrylamide gel.

2.2.13 Analysis of spliceosomal complexes

To analyze heparin-resistant complexes in SR and Pol II-TSR, the reaction mixtures were prepared as described in Section 2.2.5 and Section 2.2.6, respectively. The samples were incubated at 30 °C. Aliquots of 2 µl were taken at different time points into a microtitre plate on dry ice. Heparin at the final concentration of 0.8 mg/ml was added to each aliquot and incubated at RT for 30 min. This followed by the addition of 2 µl of TG loading buffer (1X TG buffer, 40% glycerol, bromophenol blue and xylene cyanol dyes) and loading the samples on 2% LMP agarose gel (Invitrogen) containing 1x TG buffer (50mM tris base and 50mM glycine). The gel electrophoresis was done at 4 °C, 80V for 5 h for 2 exon construct or 8 h for 3 exon construct.

2.2.14 Analysis of DNA-protein complexes in transcription/splicing assays

The reaction mixtures were prepared as described in Section 2.2.6. The radiolabelled DNA templates under either the CMV promoter or T7 promoter were used. The procedure was the same as for spliceosomal complexes and it is described in Section 2.2.14.

2.2.15 Synthesis of radioactive ladder

To prepare a DNA ladder for radioactive labelling, the plasmid pBR322 (NEB) was digested with the restriction enzyme HpaII (NEB). A sample (10 µl) contained 1 µg of plasmid DNA and 1 U of the restriction enzyme. Incubation was carried out at 37 °C for 30 min. 1 µl of the digested DNA templates was added to a reaction mixture containing 1 µl of T4 polynucleotide kinase (PNK; NEB), 1x PNK buffer and 1 µl of [γ-^{32}P] ATP. The volume of the mixture was 10 µl. The sample was incubated at 37 °C for 30 min. Free nucleotides were removed by centrifugation through a G50 column (GE Healthcare) according to manufacturer's recommendations. 50 µl of formamide dyes were added. The radiolabelled marker was heated at 80 °C for 8 min prior loading on a denaturing polyacrylamide gel.

Sizes of DNA ladder fragments: 622, 527, 404, 307, 242, 238, 217, 201, 190, 180, 160, 147, 123, 110, 90, 76 and 67.

2.2.16 RNA extraction

10 cm dish of HEK 293T cells was washed twice wish 5 ml ice cold 1x PBS buffer (136.89 mM NaCl, 2.68 mM KCl, 10.14 mM Na2HPO4, 1.76 mM KH2PO4). 2 ml of 1x PBS buffer were used to collect cells in a fresh eppendorf tube. RNA was extracted by using RNeasy Mini Kit (Qiagen) according to manufacturer's recommendations. RNA pellet was dried and resuspended in 30 µl of TE.1 buffer and stored at -80 °C.

2.2.17 Reverse transcription

2 µg of cell extracted RNA was mixed with 0.5 µg random primers (Promega) and made up to 17 µl with water. This was heated to 70 °C for 5 min and incubated on ice for 10 min. A 17 µl reaction mixture contained 1 µl MMLV reverse transcriptase (RNase H-, Promega), 1x MMLV reverse transcriptase buffer and 0.5 mM dNTPs. The reaction volume was 25 µl. The samples were incubated at 37 °C for 1 h. 50 ng of cDNA was used for PCR using Phusion High-Fidelity DNA Polymerase (Thermo Scientific). PCR procedure was done according to manufacturer's recommendations.

2.3 Mammalian cell biology techniques

2.3.1 Establishing HEK 293T cells from a frozen stock

An aliquot (1 ml) of HEK 293T cells was warmed up in hands. Cells were placed in a 10 cm plate containing 9 ml of DMEM (Gibco) supplemented with 1% PenStrep (Gibco) and 10% FBS (Gibco). The plate was incubated at 37 ºC, 0.5% CO_2. The confluence of cells was assessed under Olympus Tokyo 10x microscope. The cell number was measured using a heamocytometer. When cells reached 70-80% confluence they were split.

2.3.2 Freezing mammalian cells

Cells were harvested using trypsin and centrifuged for 8 min at 1000 RPM. A mixture of 10 ml of 70% DMEM, 20% FBS and 10% DMSO was used to resuspend cells. The concentration of cells was measured by a haemocytometer. After diluting cells to achieve the concentration of $2x10^6$/ml, they were frozen at -20 °C for 2 h, then -80 °C O/N and transferred to liquid nitrogen.

2.3.3 Calcium chloride transfection

HEK 293T cells were seeded at $5x10^5$ cells per well of a six-well plate supplemented with 2 ml complete DMEM. When cells reached 30-40% confluence, the media was replaced and cells were incubated for 4 h. 10 µl of plasmid DNA (0.5-1 µg) was added to 60 µl of ice cold $CaCl_2$ buffer (1 mM Tris-HCL pH 7.5, 0.1 mM EDTA, 300 mM CaCl2) and 60 µl 1x HBS buffer (135 mM NaCl, 5 mM KCl, 0.7 mM mM Na2HPO4, 5.5 mM glucose, 21 mM Hepes pH 7.05). The incubation was carried out on ice for 10 min. The mixture was gently added to cells. After 24 hours, pre-warmed DMEM:DMSO solution (75% DMEM:25% DMSO) was added to cells for 3 min. Cells were washed twice with 100% DMEM and left for 24-48 h of incubation. After transfection, fluorescence was detected using a Nikon Eclipse TS100 microscope at 40x and 10x magnifications. A eGFP filter 460 nm-500 nm was used to detect eGFP fluorophores.

3 ESTABLISHING *IN VITRO* TRANSCRIPTION/SPLICING ASSAYS

3.1 Introduction

3.2 Establishing an in vitro Pol II transcription/splicing assay

3.3 Comparison of splicing between Pol II-TSR and SR assays

3.4 Establishing an in vitro T7 transcription/splicing assay

3.5 Effects of Pol II transcription on splicing

3.6 Characterisation of transcription complexes in Pol II-TSR

3.7 Conclusions

3.1 Introduction

Understanding of the basic mechanisms of splicing reaction was achieved by the development of an in vitro splicing assay in which pre-made transcripts are spliced in HeLa nuclear extract (NE; Krainer et al., 1984). In this assay transcripts are generated by RNA T7 polymerase (T7 Pol) following their purification. After, naked pre-mRNA is added to the NE for splicing. This procedure is applicable if a study addresses questions about the mechanisms of splicing regulation by factors which have direct or indirect interactions with RNA sequences and their function does not depend on transcription. In living cells the molecular events of gene expression are dynamic and factors involved in one process can interact with those required for another process (reviewed by Bentley, 2014). To study the mechanisms of RNA processing at least reactions which follow one another do not have to be separated, as transcription and RNA maturation take place within or involve the Pol II transcription complexes. Differences in splicing outcomes obtained from in vitro and in vivo studies supported the model of co-transcriptional splicing regulation (reviewed by Kornblihtt et al., 2013). For example, in living cells introns of 55,000 nt in length are efficiently removed, while in vitro they have to be significantly shortened (for example Bcl-X; Cloutier et al., 2008; Weldon et al, 2016). Intron excision takes several minutes *in vivo*, while up to one hour is needed for splicing reaction *in vitro* (Mueller and Hertel, 2014). Also, it has to be taken into account that after *in vitro* transcription RNA is kept in buffers suitable for storage, but these conditions can greatly affect the structure of RNA, which to some extent can be different to that in the nucleus. Proteins bound to RNA can change its structure which, in turn, makes an impact on splicing outcome (Glisovic et al., 2008). Therefore, due to limitations of *in vivo* technologies, splicing regulation has to be studied in a combined *in vitro* reaction. In a recapitulated *in vitro* system even if RNA has been released from Pol II and it enters the splicing pathway in the same manner as it happens in the conventional *in vitro* splicing assay but structural properties of the newly synthesised RNA could be different. This can happen because RNA might be trapped by proteins as soon as it emerges from the polymerase exit pore. This can result in the natural RNA structure organisation. Therefore, to address questions about potential effects of Pol II transcription on splicing and test plausible roles of splicing factors in Pol II transcription it would be helpful to validate a model system that would support

transcription and splicing in the NE. Factors that mediate coupling between Pol II transcription and splicing can be elucidated by this approach.

To investigate whether two processes affect one another, they have to be studied in a comparison with an uncoupled reaction. The conventional splicing assay is an uncoupled reaction as pre-made transcripts are added to the NE. However, an assay in which transcription is carried out by a polymerase which could not be implicated in coupling transcription to RNA processing is a better system. T7 RNA polymerase (T7 Pol) has no domain responsible for the recruitment of RNA processing factors (Skinner et al., 2004). It is widely used to generate transcripts *in vitro* which are efficiently spliced in a separate reaction in the presence of the NE. Therefore, an assay in which transcripts are synthesized by T7 Pol in the presence of the NE is necessary to test. Although T7-TSR and the conventional splicing reaction (SR) are both uncoupled reactions and transcripts are produced by T7 Pol but there is a significant difference between these two systems. In T7-TSR released RNA in the NE may exhibit different secondary structure to that of pre-synthesised RNA. Therefore, to study effects of Pol II transcription machinery on splicing, a reaction in which transcripts are synthesized by T7 Pol in the presence of NE is ideal. The establishment and optimization of Pol II and T7 transcription/splicing assays will be described in this chapter. The effects of Pol II transcription on constitutive and alternative splicing and spliceosome assembly were investigated in this study and will be described below.

3.2 Establishing an *in vitro* Pol II transcription/splicing assay

In previous studies, significant progress was achieved in developing an *in vitro* transcriptional splicing assay that allowed transcription and splicing to occur simultaneously (Ghosh et al., 2000; Das et al., 2006; Hicks et al., 2006). Previously, DNA constructs with various Pol II promoters were tested (Lazerev and Manley, 2007). The most efficient transcription and splicing were obtained with substrates containing a cytomegalovirus (CMV) promoter. Based on this, the CMV promoter having the core elements and enhancers was used in this study. To establish an *in vitro* transcription/splicing reaction (TSR), a β-globin (Bg) mini-gene was used, comprising a

sequence of the second exon, a shortened second intron and the beginning part of the third exon (Skordis et al., 2003; Hodson et al., 2012). Previously in the laboratory, the 5'SS of Bg was replaced with the consensus 5'SS (CAGGGTGAGT---CAGGTAAGT) in order to achieve full complementary with the 5' end of the U1 snRNA (Hodson et al., 2012). Prior to testing this substrate in Pol II-TSR, the Bg mini-gene was cloned into EGFP-C1 plasmid containing the CMV promoter (Materials and Methods). As a result, the gene was fused downstream from the CMV promoter. Bg mutants with inactivated splice sites and other constructs having different gene origins were cloned into the EGFP-C1 plasmid to attach the CMV promoter to the 5' end of mini-genes.

CMV-Bg DNA template was amplified by PCR using a CMV Frw primer and either primer P17 or Bg 3 Rev (Table 2.2; Materials and Methods). PCR fragments amplified by CMV Frw and P17 primers contained the 5'SS at the 3' end because primer P17 contained the consensus 5'SS (Figure 3.1 A). This was used with the aim to increase a stability of pre-mRNA, demonstrated in the laboratory earlier (Hodson, 2011). This construct will be further called Bg C. PCR fragments amplified by CMV Frw and Bg 3 Rev primers did not contain the 5'SS at the 3' end. This construct will be further called Bg K. The majority of experiments were carried out using the Bg C construct. P16 and P17 primers (Materials and Methods) were used to amplify a Bg C mini-gene containing a T7 promoter (from P16). This DNA fragment was used for transcribing pre-mRNA by T7 Pol for further use in a conventional *in vitro* splicing reaction (SR). The efficiency of Pol II-TSR system will be compared with standard SRs in which the optimal potassium, magnesium and HeLa NE concentrations were previously established in the laboratory.

HeLa NE has been widely used to study the mechanisms of splicing regulation using *in vitro* assays, in which pre-made pre-mRNA is spliced with high efficiency. Since the NE also contains transcription factors including Pol II, it is also widely used for studying transcription. Therefore, the protocol for *in vitro* TSR was adapted from a standard SR by adding the components required for transcription such as a DNA template containing a CMV promoter, radiolabelled nucleotides and RNase inhibitor (Folco and Reed, 2014). The exact conditions and salt concentrations used in this study are described in detail in Materials and Methods. 100 ng of the CMV-Bg C PCR product of 977 bp (Figure 3.1) and 2 fmol of pre-made Bg C pre-mRNA (385 nt) were

incubated in two different commercial HeLa NEs (A: batch number 5317 and B: batch number 5343). RNA transcribed in TSR was visualised by adding trace amount of [α-^{32}P] GTP (0.66 μM). It was found that efficient transcription varied among NEs, even though they supported splicing with equal efficiency (Figure 3.1 B). Splicing products such as mRNA and intermediates were assigned in conventional SR in the laboratory previously. There were three bands detected in TSR.

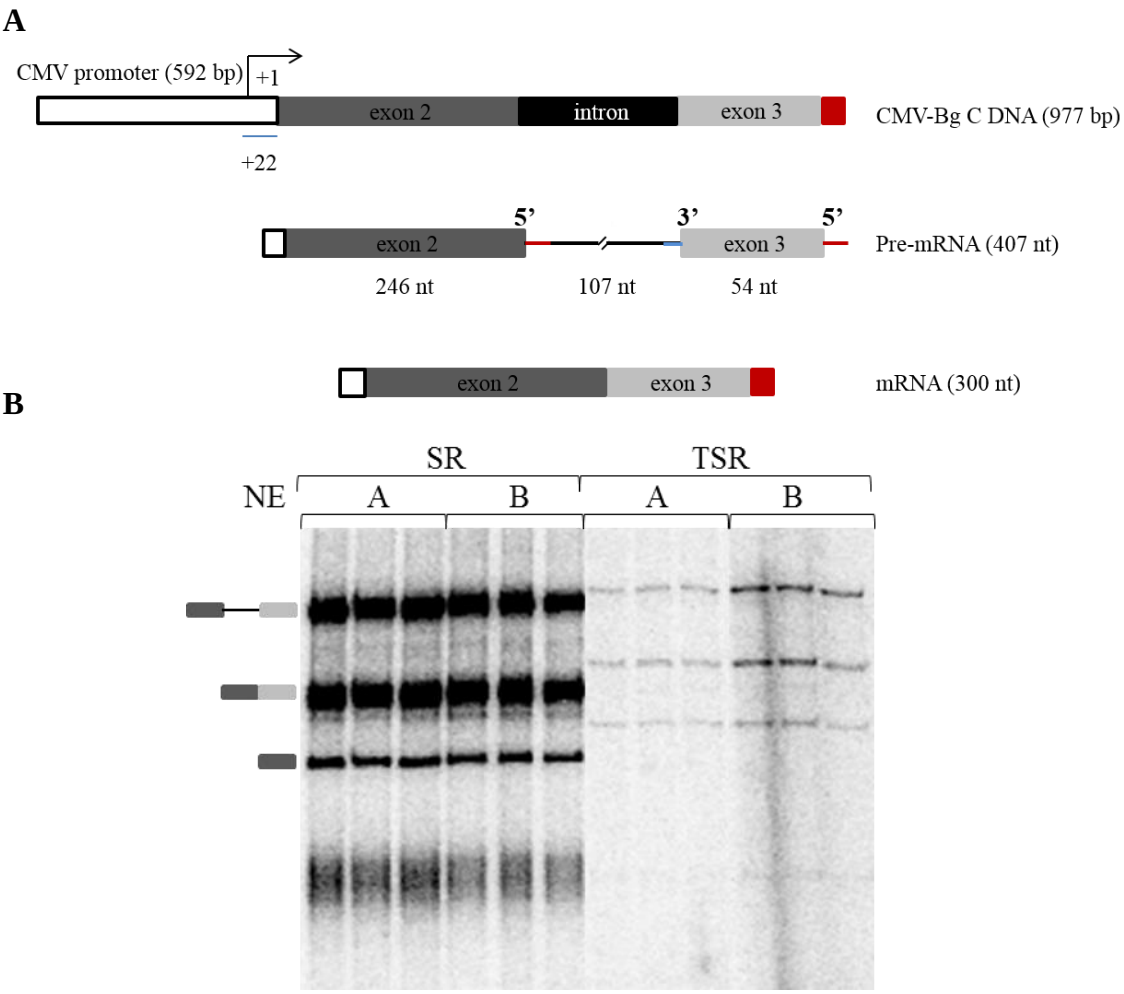


Figure 3.1. Schematic of the Bg C DNA template used for Pol II polymerase-driven RNA synthesis. CMV promoter (open box) and the position of the transcription start site are shown. The grey boxes represent exons 2 and 3 of rabbit Bg gene. The intron is shown in the black box. Transcription from the CMV promoter starts at 22 nt upstream of the 5' end of the exon 2. Therefore, a smaller open box represents extra 22 nt transcribed from the CMV promoter and this piece of RNA becomes a part of the exon 2. The positions of the 5' and 3' splice sites are indicated. The consensus 5'SS at the 3' end is shown as a red box. The sizes of exons, intron and CMV promoter are indicated. (B) Testing different commercial nuclear extracts (A and B) in splicing and transcription/splicing reactions (SR and TSR). The concentration of CMV-Bg C DNA template in each sample was 10 ng/μl (reaction volume was 10 μl). The experiment was

done in triplicate. Incubation was carried out for 90 min. RNA was fractionated on 6% denaturing polyacrylamide gel. Radiolabelled nucleotides (GTP) get incorporated into the pre-mRNA during transcription and allow detection of the splicing products using phosphorimager screens. Icons illustrated to the left hand side of the gel indicate the positions of pre-mRNA and species produced as a result of splicing (mRNA and 5' exon) in SR. Pre-mRNA and splicing products in TSR were not assigned at this point.

In the time-course experiment transcripts having different length were visualized (Figure 3.2 A). These products were not present when the general transcription inhibitor Actinomycin D (Act D), which is known to bind to DNA leading to transcription inhibition (Sobell, 1985), was used at the start of the reaction. Since not all reaction products appeared simultaneously and some were accumulated over the course of incubation, it was possible to differentiate pre-mRNA from mRNA. Also, assignments can be validated by staling splicing at certain steps of the spliceosome formation. Therefore, splicing was inhibited by a 2'-*O*-methy oligonucleotide that anneals to a part of the U6 snRNA (27–46 nt) that forms a duplex with the U2 snRNA during splicing (Materials and Methods; Table 2.3; Dönmez et al., 2007). The NE was incubated with this oligonucleotide for 15 min at 30^{o}C prior to the addition of the CMV-Bg C DNA. This was followed by further incubation for 120 min. As a result, the precursor was observed as one band accumulating with time while other species were absent in this sample (Figure 3.2 B). This suggests that the observed bands in the uninhibited reaction were the mRNA and the 5' end exon intermediate. It is apparent that transcription and splicing occur sequentially.

The length of precursor and splicing products in TSR was determined based on their electrophoretic mobility relatively to the radiolabelled markers (Figure 3.2 C). The origin of splicing products in SR was assigned in the laboratory previously. However, all transcripts in TSR were longer by 22 nt because transcription starts at the position +22 from the 5' end of the gene (Figure 3.1 A). Therefore, Pol II transcription yields Bg C pre-mRNA of 407 nt in length. In this reaction splicing of Bg C precursor produces mRNA of 300 nt and the intermediate 5' exon is 246 nt. In some experiments, the intermediate 3' exon lariat (161 nt) and intron lariat (107 nt) were not detected, but they could be observed in others (for example, Figure 3.10 C). From two systems the

products have similar electrophoretic mobilities, suggesting that the expected sites were selected.

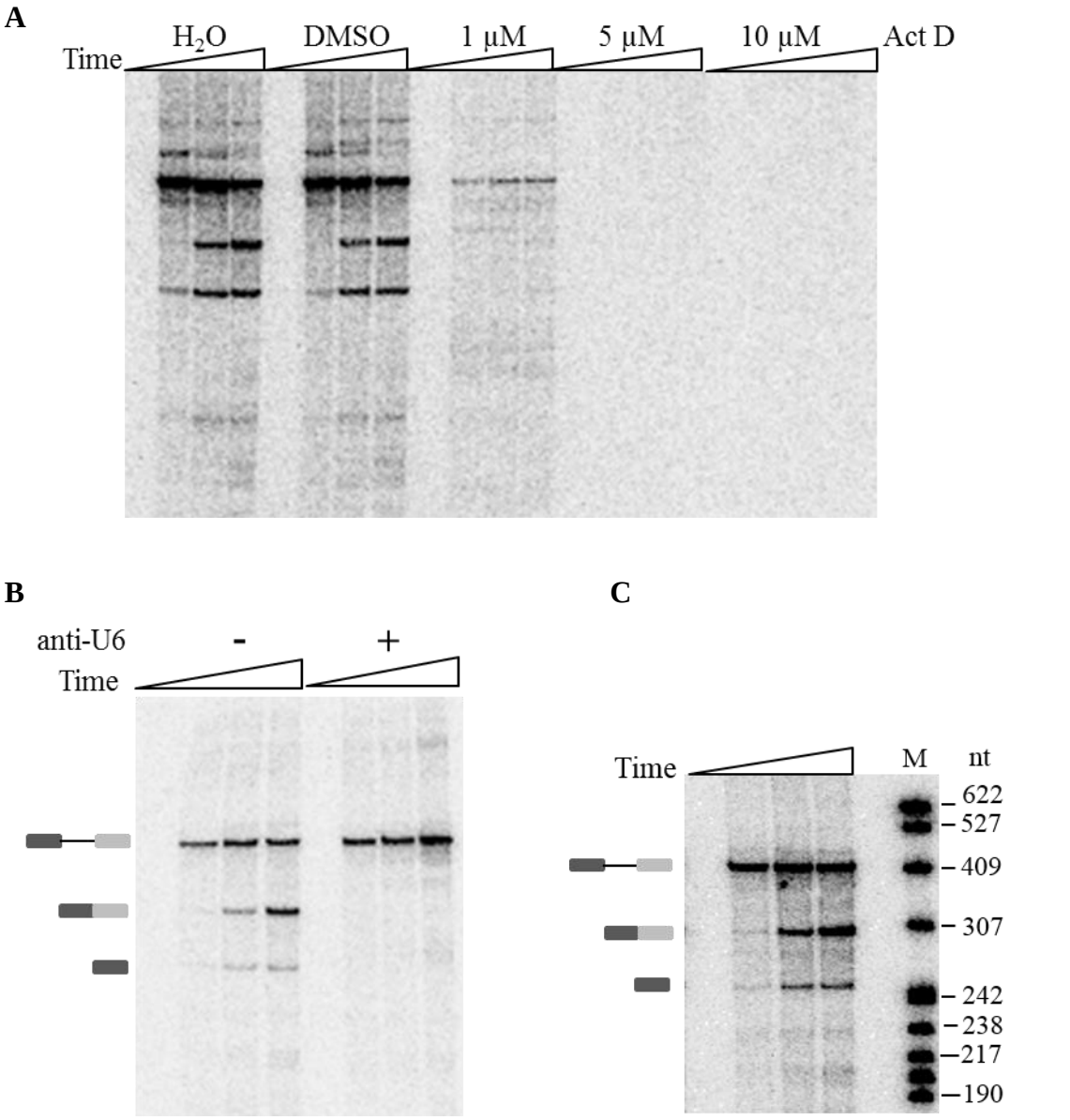


Figure 3.2. An *in vitro* Pol II transcription/splicing system. (A) Effect of Actinomycin D on Pol II transcription. Actinomycin D (Act D) was titrated with increasing concentrations. DMSO was used as a control as Act D was dissolved in 10% DMSO. The final concentration of DMSO in each reaction was 1%. DMSO and Act D were added at the beginning of the reaction. (B) Anti-U6 snRNA 2'-*O*-methyl oligonucleotide-mediated inhibition of splicing in Pol II-TSR. The oligonucleotide at the concentration of 2 µM was added to inhibit splicing with the aim to assign splicing products. (C) PAFE of time-course Pol II-TSR with the marker. The MspI digest fragments of pBR322 DNA were used as size markers. Icons illustrated to the left hand side of the gel indicate the positions of species produced as a result of transcription and splicing. M refers to the marker with indicated sizes. (A, B and C) Time points of 0, 30, 60 and 120 min were taken.

To confirm that Pol II was responsible for the observed transcription, 5, 6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) compound was used as it has been shown that Pol II is the only polymerase sensitive to this molecule (Yankulov et al., 1995). DRB diluted in 10% DMSO was added at the start of the reaction at different concentrations (Figure 3.3). Quantification of pre-mRNA and splicing products was calibrated according to the number of GTP. The level of transcription was calculated as the sum of the pre-mRNA, mRNA and 5' exon intermediate molecules. The result showed that the level of transcription was reduced, indicating that pre-mRNA synthesis is carried out by Pol II. There was a threefold decrease in transcription level in the sample with 50 μM DRB. Interestingly, transcription at 60 min incubation was of the same level in the samples with the two concentrations of DRB, while at 120 min it was different, fitting the progressive inhibition of Pol II elongation.

3.2.1 CMV-Bg C DNA titration

Although conditions for Pol II-TSR have been optimized previously by other groups (Ghosh et al., 2000; Das et al., 2006; Hicks et al., 2006), there could be differences in the efficiency of the NE used in this study. The most obvious starting point for optimization was to titrate DNA with the aim of obtaining optimal Pol II transcription and splicing efficiency. Initially, 100 ng of CMV-DNA was used. CMV-Bg C DNA template was used to optimize TSR conditions in further experiments. The amount of DNA was optimized in a range between 50 and 200 ng per 10 μl reaction. The result showed a linear increase in the level of RNA during incubation for all samples (Figure 3.4). There were equal amounts of transcripts made in the reactions with 100 ng and 200 ng of DNA, suggesting a limiting capacity of the NE. The graph shows that the highest level of RNA was generated in the reaction with 150 ng of DNA. Based on the data obtained from the samples with 100 ng and 200 ng it is expected that the level of transcripts should be similar in the sample with 150 ng. The experiment was not done in triplicate and the result from the sample with 150 ng was attributed to loading error. Since Pol II-TSR with 100 ng of CMV- Bg C DNA represented high levels of transcription, this amount of DNA was used in further experiments.

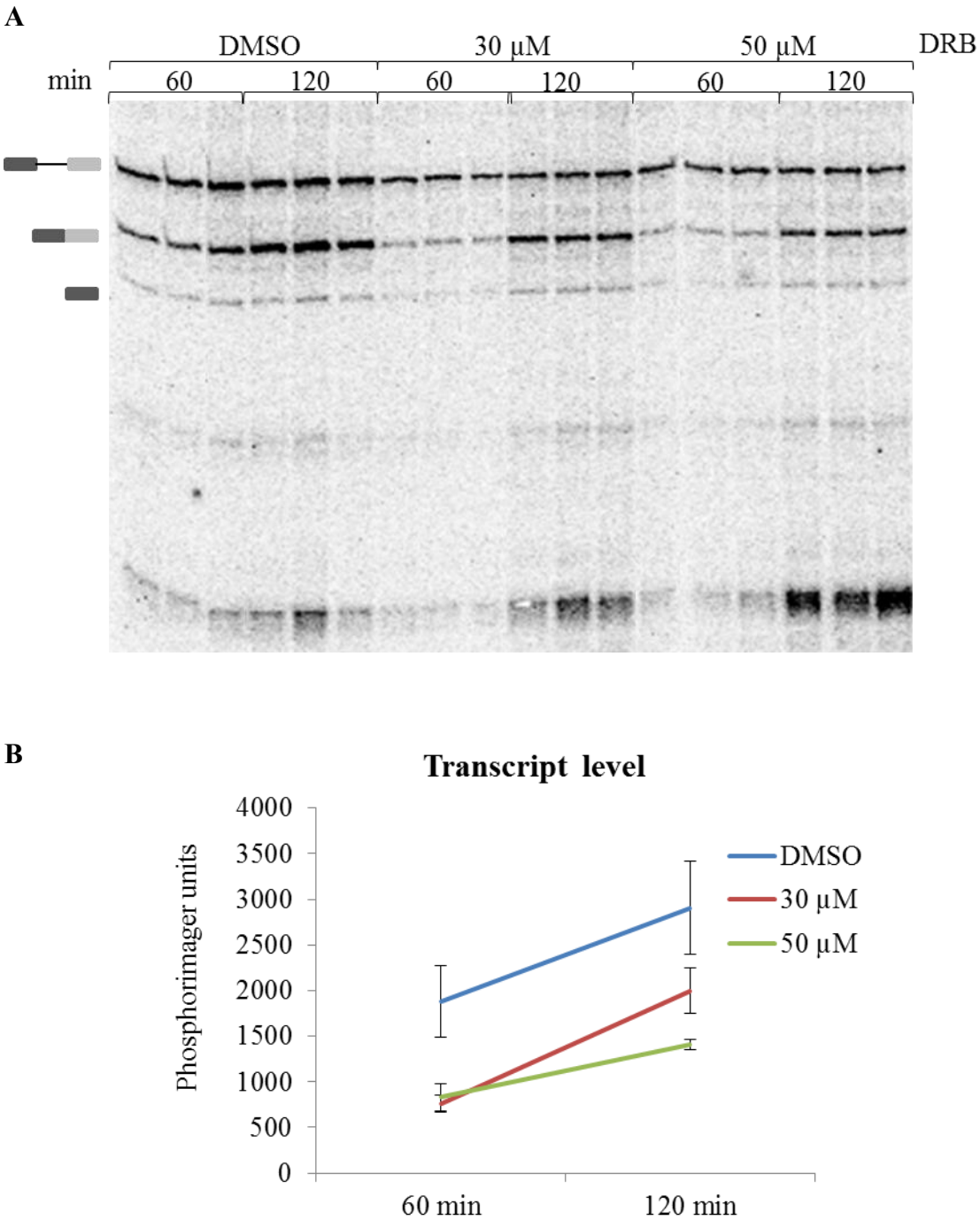
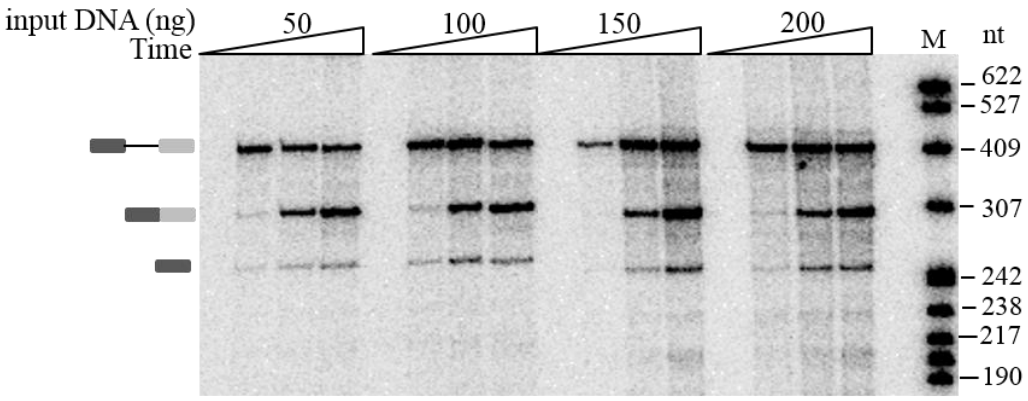


Figure 3.3. Effect of DRB on transcription. (A) DRB titration in transcription/splicing reaction. The experiment was done in triplicate and incubated for the indicated time. DRB was dissolved in DMSO and added to the samples to produce the final concentration indicated. Final concentration of DMSO was 1%. Icons illustrated to the left hand side of the panel represent pre-mRNA, mRNA and 5' exon. (B) Graph shows quantification of transcription levels (sum of the pre-mRNA, mRNA and 5' exon) of the data (A) versus time.

A



B

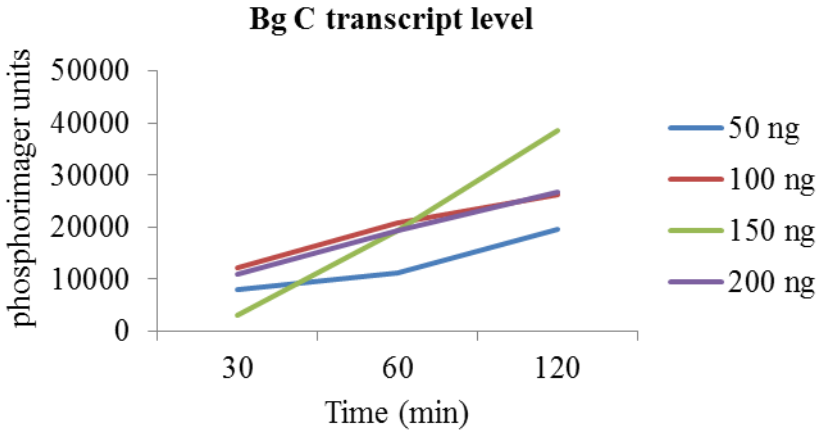
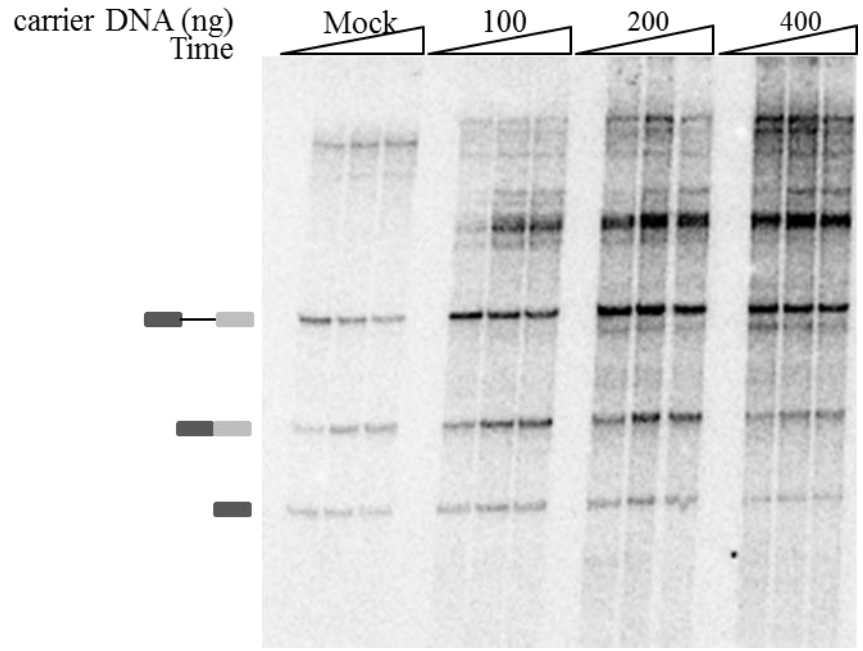


Figure 3.4. Optimisation of Pol II transcription/splicing system. (A) CMV-Bg C DNA template titrated with increasing concentrations. The amount of input DNA in 10 µl reaction is represented. M refers to the marker with indicated sizes. The MspI digest fragments of pBR322 DNA were used as size markers. Time points of 0, 30, 60 and 120 min were taken. (B) Transcription level (sum of the pre-mRNA, mRNA and 5' exon) from (A) was plotted versus time. The amount of RNA in each band was calculated as total counts divided by the number of GTP located in RNA sequences.

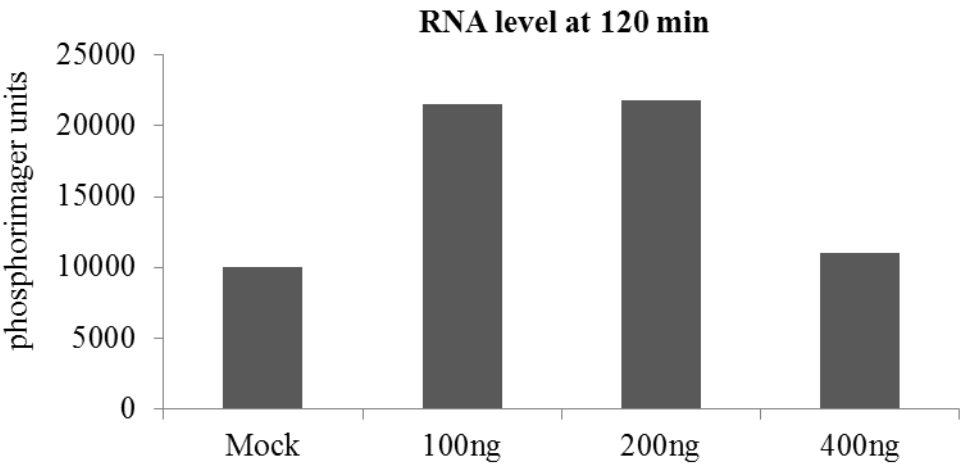
3.2.2 Carrier DNA titration

It was found previously in the laboratory that not all RNA molecules are committed to splicing as some either undergo degradation or get adsorbed onto the plastic walls of a test tube. To investigate the proportion of DNA being transcribed, the amount of CMV-DNA in the presence of carrier DNA which lacks the CMV promoter was reoptimized. To obtain a carrier DNA, TOPO plasmid containing RON mini-gene was used for PCR amplification using RON 10 Frw and RON 12 Rev primers (Materials and Methods; Smith et al., 2013). This DNA template was titrated while the amount of input CMV-Bg C DNA was 10 ng in all reactions (Figure 3.5). The experiment showed increased transcription in the samples with 100 ng and 200 ng of carrier DNA. Transcription was inhibited by carrier DNA at the amount of 400 ng. It is suggested that transcription factors could associate with a sequence of carrier DNA. There was a band above the pre-mRNA in the reactions with carrier DNA but its nature was not determined. It is clear from the appearance of weak bands in the mock sample (10 ng CMV-Bg C DNA) in comparison to other samples that a significant proportion of CMV-Bg C DNA did not participate in transcription as transcripts are barely detectable in the mock sample. Based on the levels of transcripts, twice as much DNA became available for transcription if the carrier DNA was present (samples: 100 ng or 200 ng; Figure 3.5 B), suggesting that DNA undergoes degradation in the NE. Although the transcription level was equal in the reactions containing 100 ng and 200 ng RON DNA but the mRNA/pre-mRNA ratio was higher in the reaction with 100 ng (Figure 3.5 C). Splicing was inhibited by the abundant RON DNA (400 ng). To explain this, it is possible that splicing regulators could associate with transcription factors (Fong and Zhou, 2001), forming complexes and, if these factors could interact with the abundant DNA templates, this could deprive pre-mRNA of splicing factors. In conclusion, since, in all experiments 100 ng CMV-DNA was used, it is assumed that only 50% of DNA template was used for transcription.

A


carrier DNA (ng)
Time
Mock
100
200
400

B


RNA level at 120 min
phosphorimager units
25000
20000
15000
10000
5000
0
Mock
100ng
200ng
400ng

C

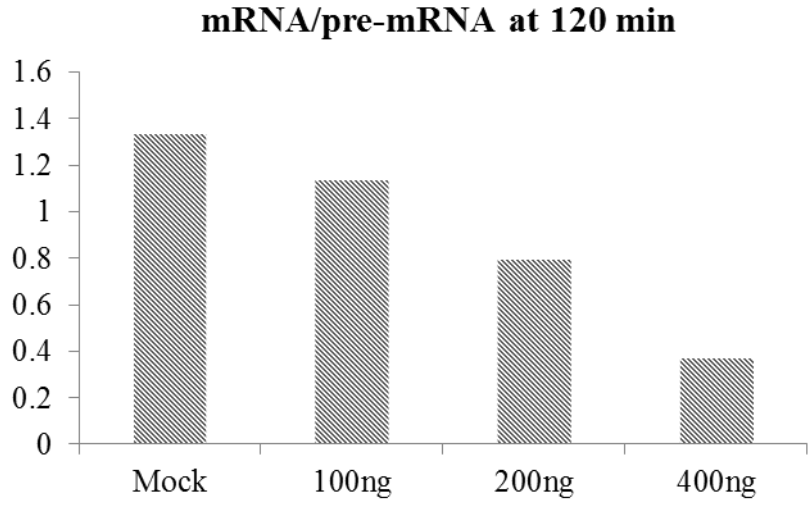
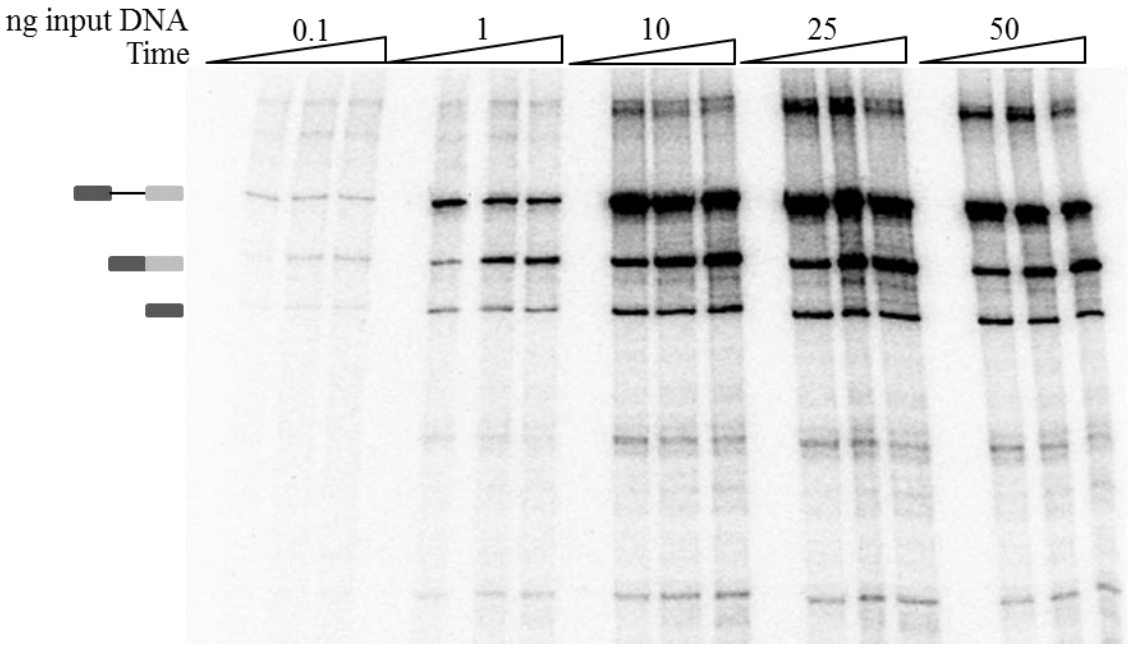

mRNA/pre-mRNA at 120 min
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Mock
100ng
200ng
400ng

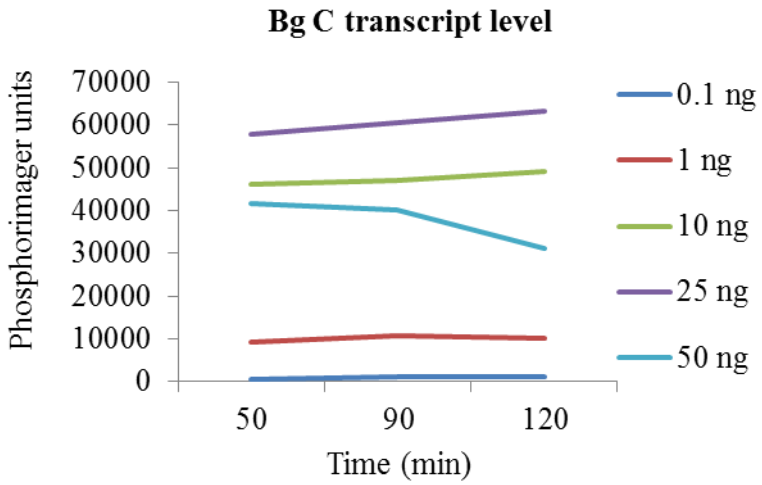
Figure 3.5. Transcription by Pol II is enhanced in the presence of carrier DNA (I). (A) Time course of Pol II transcription with carrier DNA. The amount of input carrier DNA (RON PCR fragment) in 10 µl reaction is represented. The amount of CMV-Bg C DNA was constant in each reaction (10 ng). Time points of 0, 30, 60 and 120 min were taken. (B) Quantification of transcription level at 120 min time point. (C) The ratio between mRNA and pre-mRNA at 120 min time point is shown in the diagram.

In the next experiment, equal amounts of carrier DNA (150 ng) were used and the concentration of CMV-Bg C DNA template was titrated in 10 µl of Pol II-TSR (Figure 3.6 A). Time points at 50, 90 and 120 min were taken. The result showed that transcription was enhanced proportionally to the increased amount of CMV-Bg C DNA (Figure 3.6 B). The most efficient transcription was under conditions in which the amount of CMV-Bg C DNA was 25 ng. Interestingly, there was no significant increase of transcription with time in all samples. There are two reasons behind this. First, RNA might not be steady stable. Second, at later time of incubation ATP become limited. This experiment allowed conditions to be identified in which the highest transcription level can be achieved with relatively similar mRNA/pre-mRNA ratio (Figure 3.6 C). Pol II-TSR conditions comprising 150 ng of carrier DNA and 25 ng of CMV-DNA were used in experiments in which the spliceosome complexes assembly was investigated (Section 3.5). These conditions allowed the detection of complexes with the highest resolution.

A


ng input DNA
Time
0.1
1
10
25
50

B


Bg C transcript level
Phosphorimager units
70000
60000
50000
40000
30000
20000
10000
0
50
90
120
Time (min)
0.1 ng
1 ng
10 ng
25 ng
50 ng

C

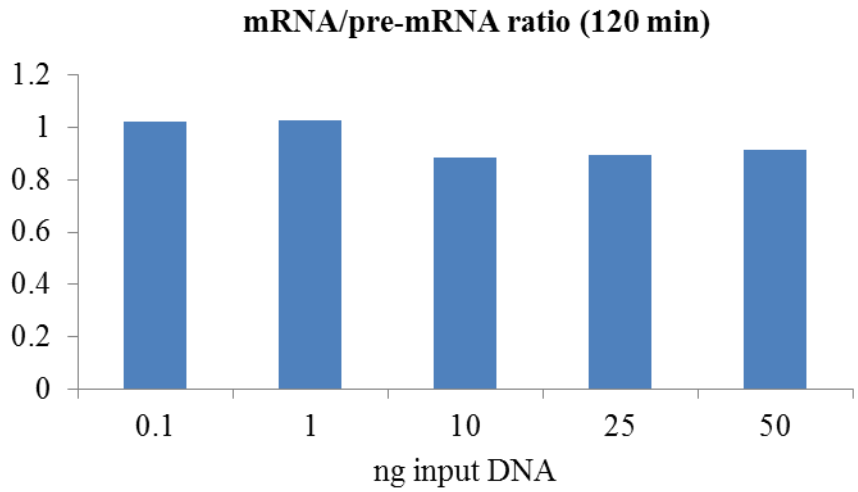

mRNA/pre-mRNA ratio (120 min)
1.2
1
0.8
0.6
0.4
0.2
0
0.1
1
10
25
50
ng input DNA

Figure 3.6. Transcription by Pol II is enhanced in the presence of carrier DNA (II). (A) Time course of Pol II transcription with carrier DNA. The amount of input CMV-Bg C DNA in 10 µl reaction is represented. The amount of carrier DNA (RON PCR fragment) was constant in each reaction (150 ng). Time points of 0, 50, 90 and 120 min were taken. (B) The graph shows quantification of transcription levels from (A). (C) The ratio between mRNA and pre-mRNA at 120 min time point is shown in the diagram.

3.2.3 Effects of exogenous nucleotides on Pol II-TSR

In the original protocol (Materials and Methods), exogenous nucleotides were used in TSR. By measuring the amount of radioactivity of nucleotides incorporated into pre-mRNA during transcription, we would be able to infer the concentration of pre-mRNA in TRS. However, the presence of endogenous NTPs would complicate the analysis. To test this, reactions were done in the absence of exogenous CTP and UTP, such that transcription could only take place if there were endogenous rNTPs that had survived dialysis. Each sample contained equal amounts of [α-^{32}P] GTP. As shown in Figure 3.7, the intensity of bands was increased in the reaction supplemented with ATP and lacking exogenous CTP and UTP nucleotides. This experiment also suggests that nucleotides present in the NE are not limiting factors for efficient transcription. In all further experiments unlabelled CTP, GTP and UTP nucleotides were not used, but ATP and [α-^{32}P] GTP. The concentration of radiolabelled RNA being synthesised in Pol II-TSR was not measured because the amount of endogenous nucleotides in the NE was not known.

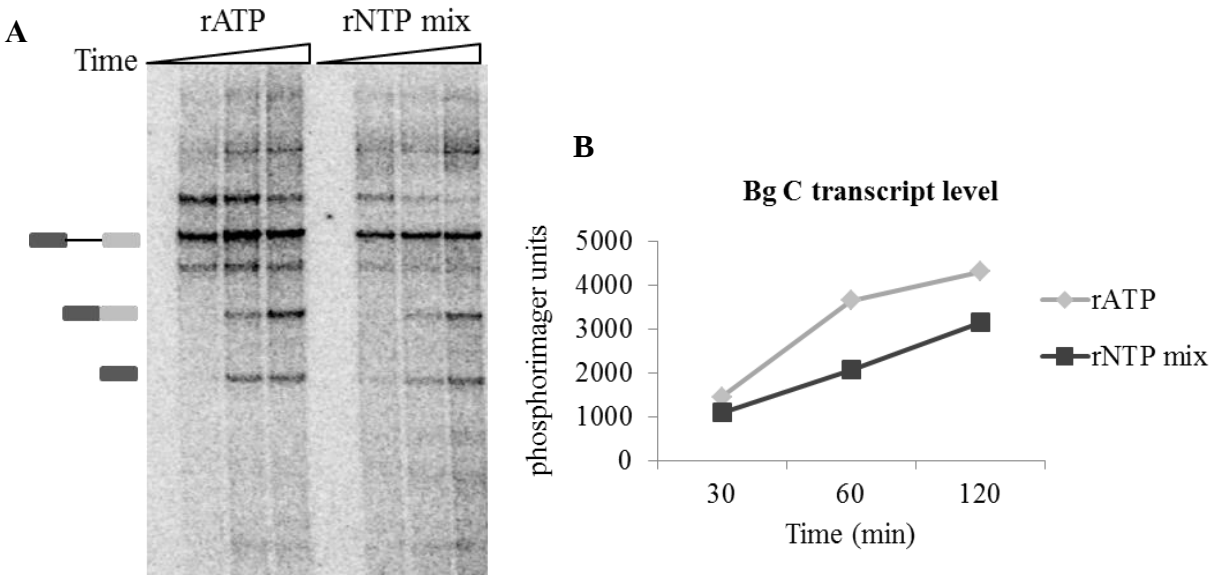


Figure 3.7. Effect of exogenous nucleotides on transcription level. Testing ATP and rNTP mix in Pol II-TSR. 0.66 μM [α-^{32}P] GTP was used in each sample. The final concentration of exogenous ATP in each sample was 0.5 mM. In the sample in which exogenous NTPs were added, the concentrations of three UTP/CTP and GTP were 0.5 mM and 0.05 mM, respectively. (B) Transcription level from (A) was quantified and plotted versus time.

3.2.4 Testing long constructs in Pol II-TSR

To investigate whether Pol II-TSR can also be efficient on long DNA templates than 407 bp (Bg C), we tested under the same conditions as before a DNA template three times longer than Bg C. Plasmid CPXJ41, containing a Nrxn2 mini-gene downstream of the CMV promoter, was used (provided by Professor D Elliott, Newcastle). The Nrxn2 construct contains three exons and two introns (Figure 3.8 A). Exon 2 of Nrxn (90 bp) with flanking intronic sequences was cloned into the second intron of rabbit β-globin gene (Ehrmann et al., 2013). The middle exon is subject to alternative slicing, so, the Nrxn2 precursor (1166 nt) produces three-exon (inclusion, 276 nt) and two-exon (skipping, 186 nt) mRNA isoforms. For Pol II-TSR, Nrxn 2 mini-gene was PCR amplified using CMV Frw and Bg 3 Rev primers (Materials and Methods). The results showed efficient production of pre-mRNA but no mRNA isoforms were detected during the course of incubation (Figure 3.8 B). To test if the bands migrating above the pre-mRNA are lariat intermediates of splicing, splicing was blocked by sequestering the U6 snRNA with 2'-*O*-methyl oligonucleotide. The top bands were lost, suggesting that they correspond to lariat intermediates. Since mRNA isoforms were not detected it is concluded that splicing was not completed under these conditions, but the splice sites were recognized. The spliceosome was able to assemble on the pre-mRNA progressing to B^{act} complex and the first transesterification reaction occurred. This suggests that some splicing factors required for the progression of the spliceosome complex C could be limiting or more time was needed for splicing to be completed. Since the Nrnx2 pre-mRNA generated by Pol II was inefficient, it was not used in further experiments.

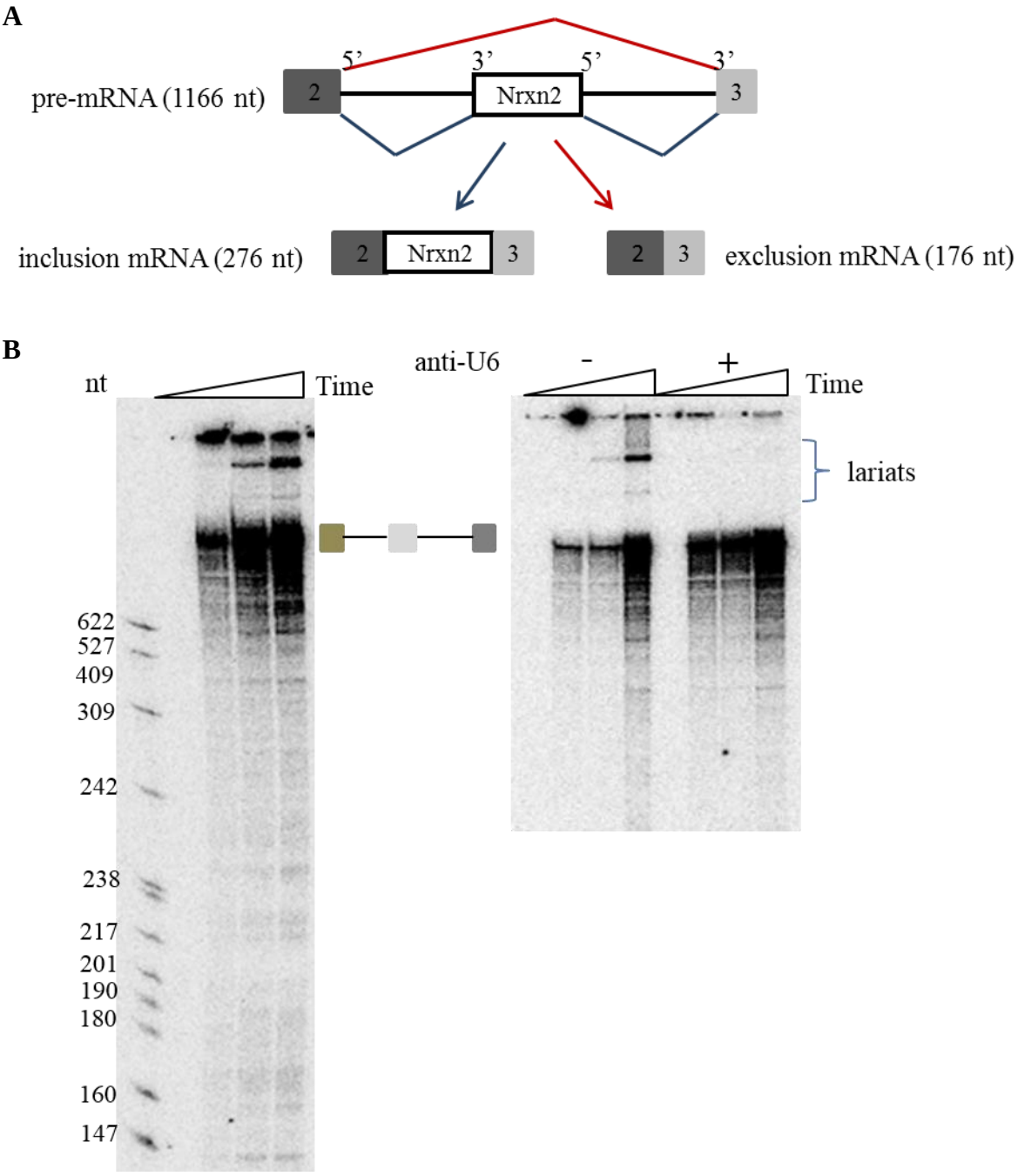


Figure 3.8. 1.2 kb transcript is efficiently synthesised by Pol II but splicing is not completed. (A) Diagram of chimeric Nrxn2 mini-gene and an alternative splicing pattern. The variable exon 2 (open box) of Nrxn gene was cloned into the full length intron of rabbit Bg gene. The grey boxes represent exons 2 and 3 of rabbit Bg. The Bg exons were shortened. The length of exon 2 and 3 was 140 and 46 bp, respectively. The size of Nrxn 2 exon was 90 nt. The positions of the 5' and 3' splice sites are indicated. The sizes of pre-mRNA and mRNA isoforms are indicated. (B) Pol II transcription/splicing of Nrxn2 DNA fragment under CMV promoter. Time points of 0, 30, 60 and 120 minutes were taken. An icon illustrated indicates the positions of pre-mRNA. Ani-U6 snRNA 2'-*O*-methyl oligonucleotide was used to inhibit splicing in order to determine if the bands above pre-mRNA are lariat intermediates.

3.2.5 Testing homemade nuclear extracts in Pol II-TSR

One of the possible applications of transcription/splicing system could be to test effects of different proteins in transcription and splicing. Commercial HeLa NE is often used for studying the mechanisms of splicing regulation *in vitro.* If a protein of interest has to be introduced with the aim of investigating its function, then a recombinant protein can be added directly to the NE. However, this can be challenging as full-length proteins are not always soluble and it is not easy to purify sufficient amounts of proteins. As alternative, the desired proteins can be overexpressed in human cells, prior to NE preparation. Unfortunately, in my experiments any "homemade" NEs prepared in the laboratory and tested in TSR did not support transcription (Figure 3.9), even though they were efficient in conventional *in vitro* splicing assays (data not shown). This suggests that transcription factors were lost during NE preparation. There are two possible reasons behind this. First, transcription factors were less positively charged and they were lost from the nucleus during the stages when cells are swollen and disrupted. Second, they are highly charged and stay bound to the chromosomal DNA and, therefore, not eluted in the high salt buffer C.

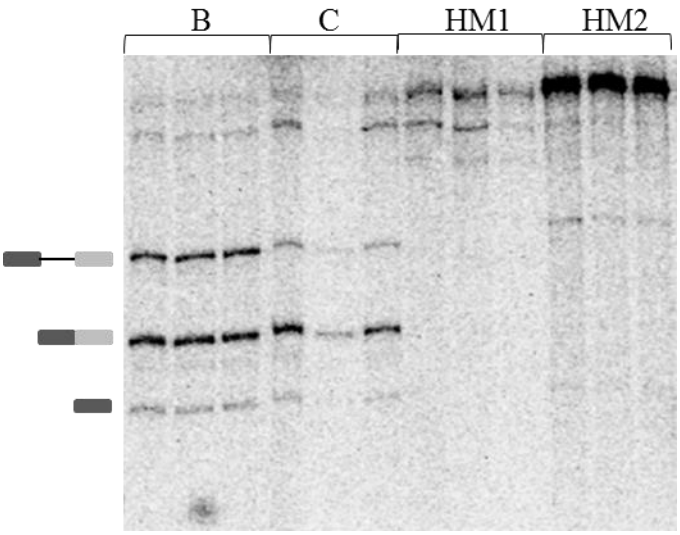


Figure 3.9. Homemade HeLa nuclear extracts do not support efficient transcription. B and C stand for commercial NE (batches numbers 5343 and 5314).
Two different homemade NE were tested (HM1 and HM2). The experiment was done in triplicate. Linear CMV-Bg CDNA was used. The samples were incubated for 90 min.

3.3 Comparison of splicing between Pol II-TSR and SR assays

3.3.1 Pol II transcription does not improve splicing efficiency

The amount of pre-mRNA which undergoes splicing was compared between two optimised SR and Pol II-TSR systems to determine whether transcription driven by Pol II is able to affect splicing efficiency. The concentration of pre-made Bg C pre-mRNA was 2.3 fmol in each reaction. Quantification of pre-mRNA and splicing products was converted for number of GTPs in each specie. Splicing efficiency was calculated as the percentage of (mRNA+5'exon)/(pre-mRNA+mRNA+5'exon). The experiment showed that 70% pre-mRNA was spliced in Pol II-TSR while up to 90% of pre-made pre-mRNA was spliced in standard SR (Figure 3.10 A and B). However, these measurements were not precise as pre-mRNA was still transcribing in the transcription/splicing system. To establish a "zero" time point for Pol II-TSR, transcription was allowed for 30 min and pulse chase was done by adding an excess of unlabelled GTP nucleotides (5 mM). This was followed by the incubation for another 80 min. The experiment revealed that at 80 min time point, 65% pre-mRNA underwent splicing (Figure 3.10 C). Based on many experiments done in this study, splicing efficiency of Bg C in Pol II transcriptional/splicing system appeared to be maximum of 70% 2h after the start of the reaction. The data suggests that Pol II transcription does not increase the rate of intron removal in comparison to standard SR assay.

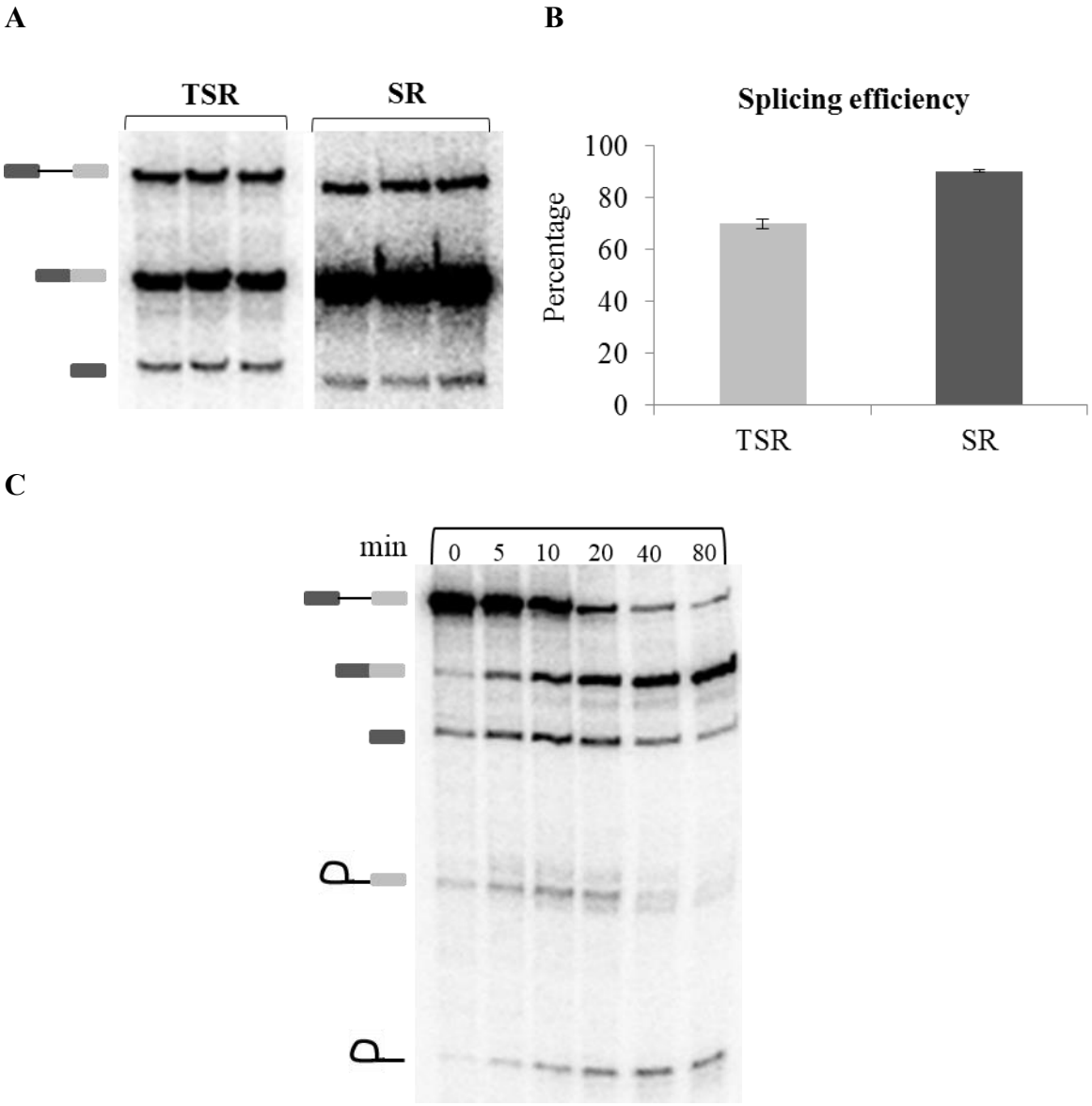


Figure 3.10. Splicing efficiency is not enhanced in the Pol II transcription/splicing system. (A) Denaturing PAGE of conventional splicing reaction (SR) and Pol II transcription/splicing reaction (TSR) in triplicates. Reaction conditions were identical excepting that some components required for transcription were added to TSR. Nuclear extract comprised 30% of the total volume. Pol II transcription and the splicing of presenthysised pre-mRNA were carried out for 90 min. (B) Splicing efficiency was calculated as the percentage by the following formula: (mRNA+5'exon)/(pre-mRNA+mRNA+5'exon) and shown in the diagram. Error values are standard deviation values. (C) Pol II transcription/splicing pulse chase assay. Transcription from CMV-Bg C DNA was allowed to proceed for 30 min and unlabelled GTP was added to a final concentration of 5mM. The samples were incubated further to the times shown. Apart from pre-mRNA, mRNA and 5' exon, intermediate 3' exon lariat and intron lariat are indicated.

3.3.2 Pol II transcription does not enhance splicing kinetics

The "recruitment model" of co-transcriptional splicing suggests that splicing factors are recruited to the nascent pre-mRNA by the transcription machinery leading to splicing enhancement (reviewed by Bentley, 2014). We hypothesised that if splicing were coupled to transcription in this *in vitro* system, then splicing kinetics would be enhanced resulting to faster intron excision as observed *in vivo* where it occurs on a timescale of 0.5-2.5 min (Bauren and Wieslander, 1994), while in a conventional *in vitro* splicing assay it takes up to 60 min (Mueller and Hertel, 2014). We, therefore, investigated whether Pol II transcription can improve the kinetic characteristics of splicing. The splicing of pre-synthesized RNA added to HeLa NE and RNA produced by Pol II in the same extract were compared in time-course (Figure 3.11) which was run with evenly spaced time points (every 3 min). RNA was first clearly detectable at the 3 min time point. The 5' exon intermediate and mRNA appeared in the Pol II-TSR at 18 min and 24 min after the start of the reaction, respectively. In the corresponding SR, the 5' exon was first detected at 9 min time point and mRNA was observed at 12 min. The rate of splicing was twice slower in transcription/splicing system in comparison to that in the conventional splicing assay. The apparent difference in the splicing kinetics allowed concluding that in a detectable way Pol II transcription does not enhance the *in vitro* rate of intron removal.

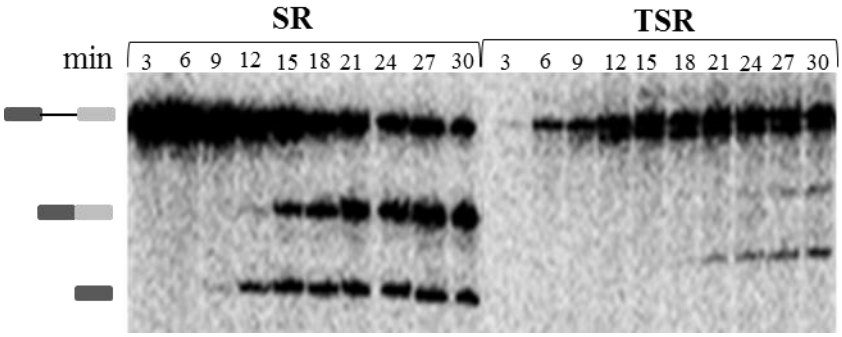


Figure 3.11. Splicing kinetics is not enhanced by Pol II transcription. (A) Denaturing PAGE of conventional splicing reaction (SR) and Pol II transcription/splicing reaction (TSR) set side-by-side. Time points are represented. Icons comprising pre-mRNA, mRNA and 5' exon intermediate are indicated.

3.4 Establishing an *in vitro* T7 transcription/splicing assay

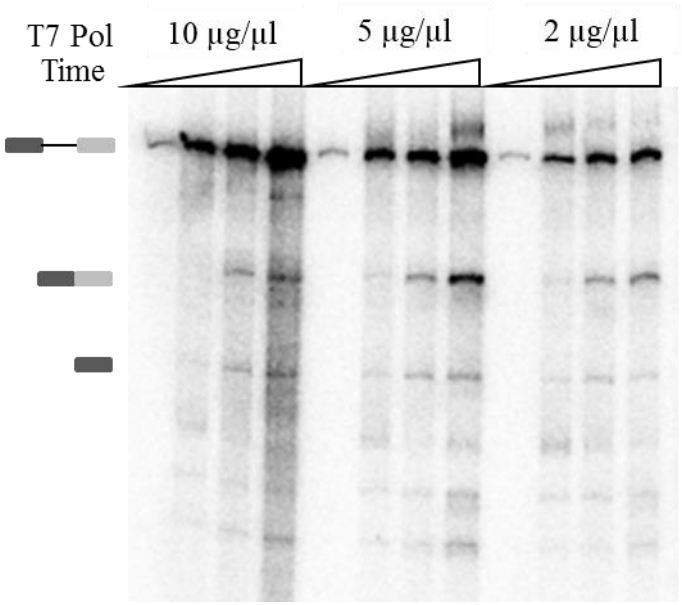
Conventional SR is supplemented with pre-made splicing substrate, whereas in TSR a precursor has to be transcribed by Pol II. Therefore, a more valid comparison for the Pol II system has to be a reaction in which transcription occurs by a polymerase that is not implicated in coupled transcription and splicing. For this, HeLa NE can be supplemented with phage T7 polymerase (T7 Pol), and T7 transcription/splicing reaction (T7-TSR) would be able to carry out two processes. Some studies compared three systems (Ghosh and Garcia-Blanco, 2000; Das et al., 2006) and used their comparison as a basis for claims. It was found that transcripts generated by Pol II are spliced more efficiently than the T7 Pol counterparts. Also, the splicing rate was enhanced by Pol II in comparison to that in T7-TSR. Thus, whether pre-made T7 transcripts were directly added to the NE or produced in the NE, the yield of spliced mRNA was low in comparison to Pol II counterparts. Based on these data it was concluded that *in vitro* Pol II-TSR system recapitulates functional *in vivo* coupling (Ghosh and Garcia-Blanco, 2000; Das et al., 2006). However, it was not clear why pre-mRNA being transcribed by T7 Pol in a separate reaction and then added to the NE splices efficiently while pre-mRNA transcribed by T7 Pol in the presence of NE is barely spliced. To answer this question, a T7-TSR assay in which transcription is carried out by T7 Pol in the NE was optimized in this study. For this, a DNA construct containing the Bg C gene under the T7 promoter was used.

3.4.1 Splicing does not occur in T7-TSR

To establish T7 transcription/splicing assay, first a titration of T7 Pol concentration was done. 10 ng/µl of T7-Bg C DNA was used. The result showed that T7 transcription occurred in the presence of NE as some transcription products were detected (Figure 3.12 A). Since more mRNA was detected in the reaction with 5 µg/µl of T7 Pol, this concentration of T7 Pol was used in further experiments. To differentiate the precursor from splicing products, splicing was prevented by using anti-U6 2'-*O*-methyl oligonucleotide (Figure 3.12 B).The result was consistent with the previous findings (Ghosh and Garcia-Blanco, 2000; Das et al., 2006l; Hicks et al., 2006), showing

efficiently transcribed pre-mRNA by T7 Pol in the presence of NE but barely any spliced mRNA was detected (3%).

A



B

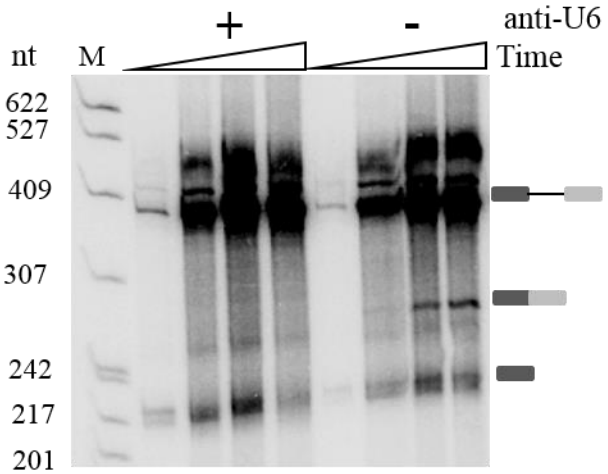


Figure 3.12. RNA transcribed by T7 polymerase does not splice in HeLa nuclear extract. (A) Titration of T7 polymerase *in vitro* T7 transcription/splicing assay. The concentration of the polymerase is indicated. (B) Time course of T7 transcription/splicing system. Ani-U6 snRNA 2'-*O*-methyl oligonucleotide was used to inhibit splicing and assign splicing products. (A and B) Time points of 0, 30, 60 and 120 min were taken.

3.4.1.1 Spliceosome assembly in T7-TSR

To characterize the spliceosome complexes in T7-TSR, they can be compared with complexes formed on pre-made pre-mRNA in SR using electrophoresis in native gels (Hodson et al., 2012). As described in Chapter 1, the spliceosome assembles on a pre-mRNA in a series of sequential reorganizations, and complexes can be characterized by using inhibitors to stop spliceosome formation at different points.

Based on data on the concentration of oligonucleotides needed to efficiently inhibit snRNPs (Hodson et al., 2012), the progression of splicing complexes was prevented at certain points by blocking snRNPs by antisense oligonucleotides in SR and T7-TSR. The inhibition of the 5' end of the U1 snRNA using 2'-*O*-methyl oligonucleotide (Materials and Methods; Table 2.3) led to a complete inhibition of complexes in T7-TSR (Figure 3.13 A), suggesting that these are the spliceosome complexes. It is concluded that complex assembly was abolished as it depends on the U1 snRNP (Reed, 1990). Splicing was not completely inhibited in SR as some complexes were still visualized. To stall the spliceosome at complex E, the NE has to be depleted of ATP before addition to a splicing reaction because complex E is defined as being ATP-independent (Michaud and Reed, 1991). However, this method is not applicable in transcription/splicing assay, because transcription requires ATP. The function of the U2 snRNP was inhibited by incubating extracts with anti-U2 2'-*O*-methyl oligonucleotide which prevents the U2 snRNA base-pairing with BP (Materials and Methods; Black et al., 1985). This led to staling the spliceosome in both reactions. By using antisense U6 snRNA 2'-*O*-methyl oligonucleotide (Donmez et al., 2007), the spliceosome was sequestered and this complex is defined as complex A, as the transition from complex A to B depends on tri-snRNP (Chapter 1). This allows to assign complexes in the mock reaction in T7-TSR. There were three bands observed in the mock reaction (SR). Previously these were assigned as complexes A, B and C (Hodson et. al., 2012). Based on the appearance and electrophoretic mobility of the complexes formed on the pre-mRNA transcribed by T7 Pol, it is concluded that the spliceosome progresses through complexes A and B. Complex C was not identified in T7-TSR but was formed after 10 min of incubation in SR. It is suggested that the reaction conditions of T7-TSR are suitable for splice site recognition by the core splicing factors but the spliceosome does not progress through step 1 to complex C. Our results disagree with previous study (Das

et al., 2006) which showed that complex H is formed in T7-TSR as this complex was not detected in T7-TSR (Figure 3.13).

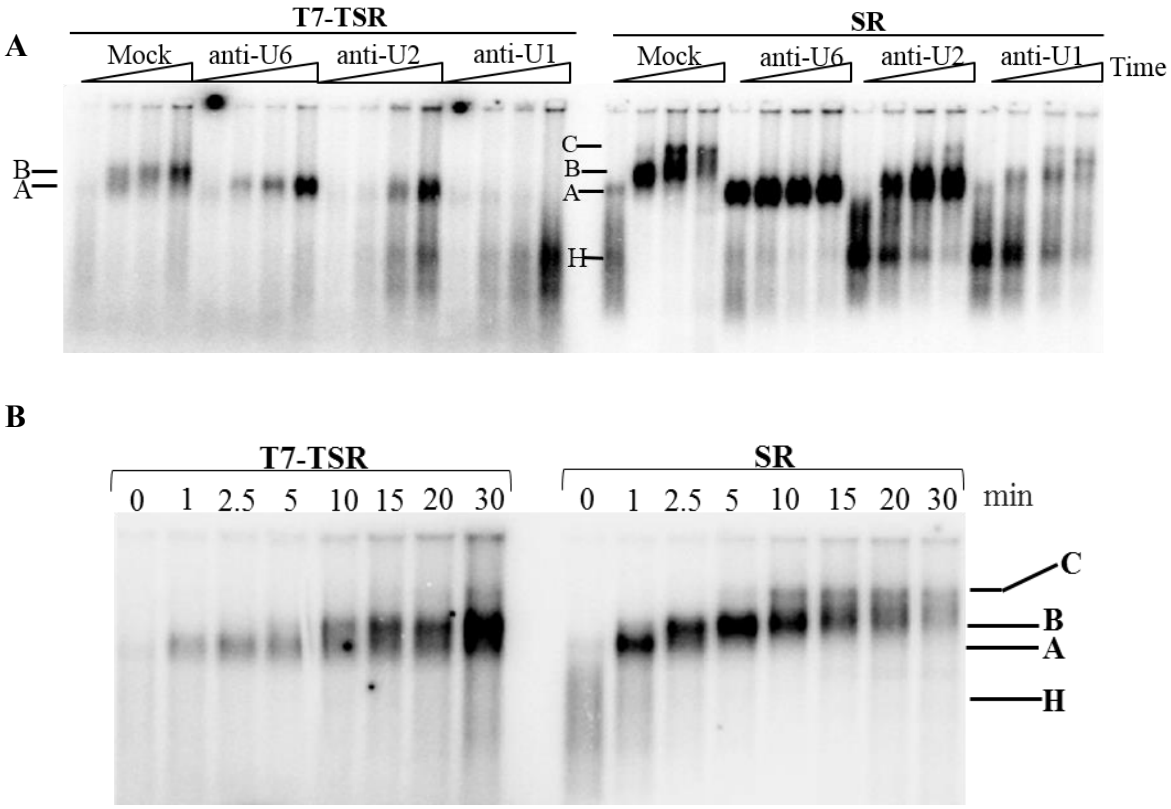


Figure 3.13. The spliceosome complexes are able to assemble on Bg C pre-mRNA generated by T7 polymerase but the spliceosome is stalled at complex B. (A) Assignment of splicing complexes in T7-TSR and SR by sequestering the U1, U2 and U6 snRNAs with 2'-*O*-methyl oligonucleotides. Samples were incubated with the oligonucleotides, following by the addition of T7-Bg C DNA template and naked Bg C pre-mRNA. T7-TSR and standard SR were set up side by side. Time points of 0, 5, 10 and 15 min were taken. (B) Same as (A) but different time points. The reactions were incubated for the times indicated. (A and B) Heparin was added to each sample. Complexes were fractionated on 2% low-melting-point agarose gel. The nonspecific complex H and the spliceosome complexes A, B and C are represented in letters on the sides of the panels.

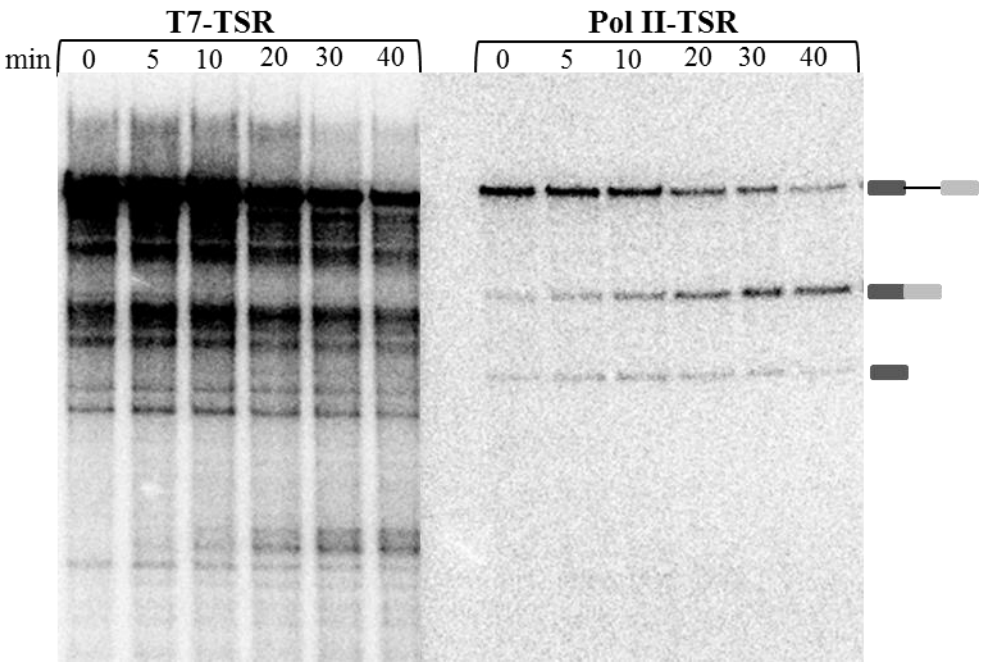
A more detailed time-course provided evidence of complex C's existence in T7-TSR (Figure 3.13 B). It was found that the kinetics of the spliceosome formation in T7-TSR was slowed down compared to SR. Complex A was formed by 1 min in both samples. This suggests that the spliceosome is assembled on transcribing pre-mRNA as fast as on added pre-made RNA. The transition from complex A to B occurred between 5 and 10 min in T7-TSR, while complex B was organised after 2.5 min of the start of the reaction. Complex C was assembled after 10 min in SR while 30 min were required for its assembly in T7-TSR. The efficiency of complex B to complex C transition could not be estimated because the complexes were not well separated and overlapped. However, it is clear that the transition of complex B to complex C was more efficient in SR than in T7-TSR. It is suggested that the spliceosome was stalled in complex B in T7-TSR. This observation is in line with the hypothesis that splicing factors are titrated out among abundant transcripts produced by T7 Pol leading to inability of the spliceosome to fully progress through the spliceosome cycle.

The electrophoretic mobility of complexes A, B and C suggested that transcription by T7 Pol did not make an effect on the protein composition of the spliceosome. Remarkably, there was no complex H formed in T7-TSR. The observation that complex H is missing in T7-TSR suggests that, during splicing complex assembly *in vitro*, binding of non-spliceosomal proteins occurs (Reed, 1990), indicating that newly transcribed RNA is committed to splicing while pre-made pre-mRNA is bound by numerous proteins as soon as it has been added to the NE. This might happen because pre-made RNA is added to the NE in excess. Our results demonstrate that spliceosome assembly occurs in T7-TSR. This observation differs from previous study in which only complex H was detected on transcripts generated by T7 Pol in the NE (Das et al., 2006).

3.4.1.2 Degradation profiles of T7 transcripts

Pulse chase experiment in which Pol II- and T7-TSR were assayed side-by-side showed that the amount of transcripts generated by T7 Pol was 3-5-fold higher in comparison to Pol II counterparts. The result revealed that RNA produced by T7 Pol was not committed to splicing but displayed significant degradation (Figure 3.14). In contrast, a degradation profile for Pol II transcripts showed that Pol II transcripts were more stable and committed to splicing. To test a hypothesis that T7 Pol has abilities to inhibit splicing, T7 Pol was added to Pol II-TSR (data not shown). The experiment showed that splicing was unaffected.

A



B

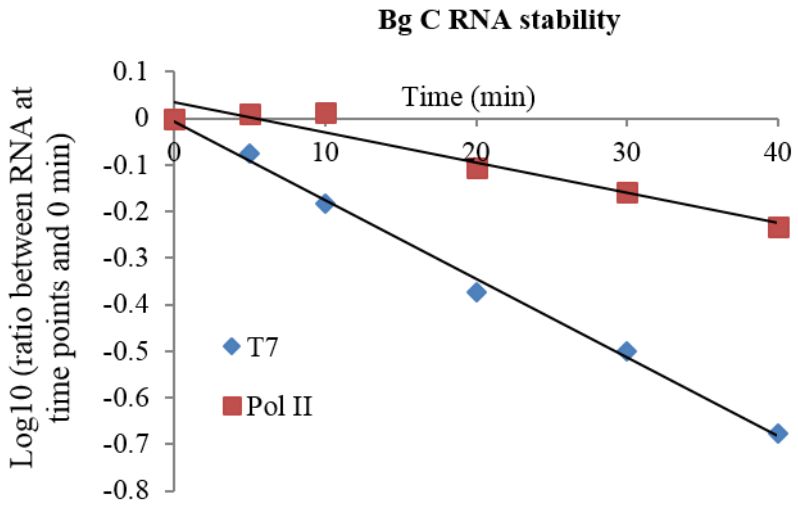


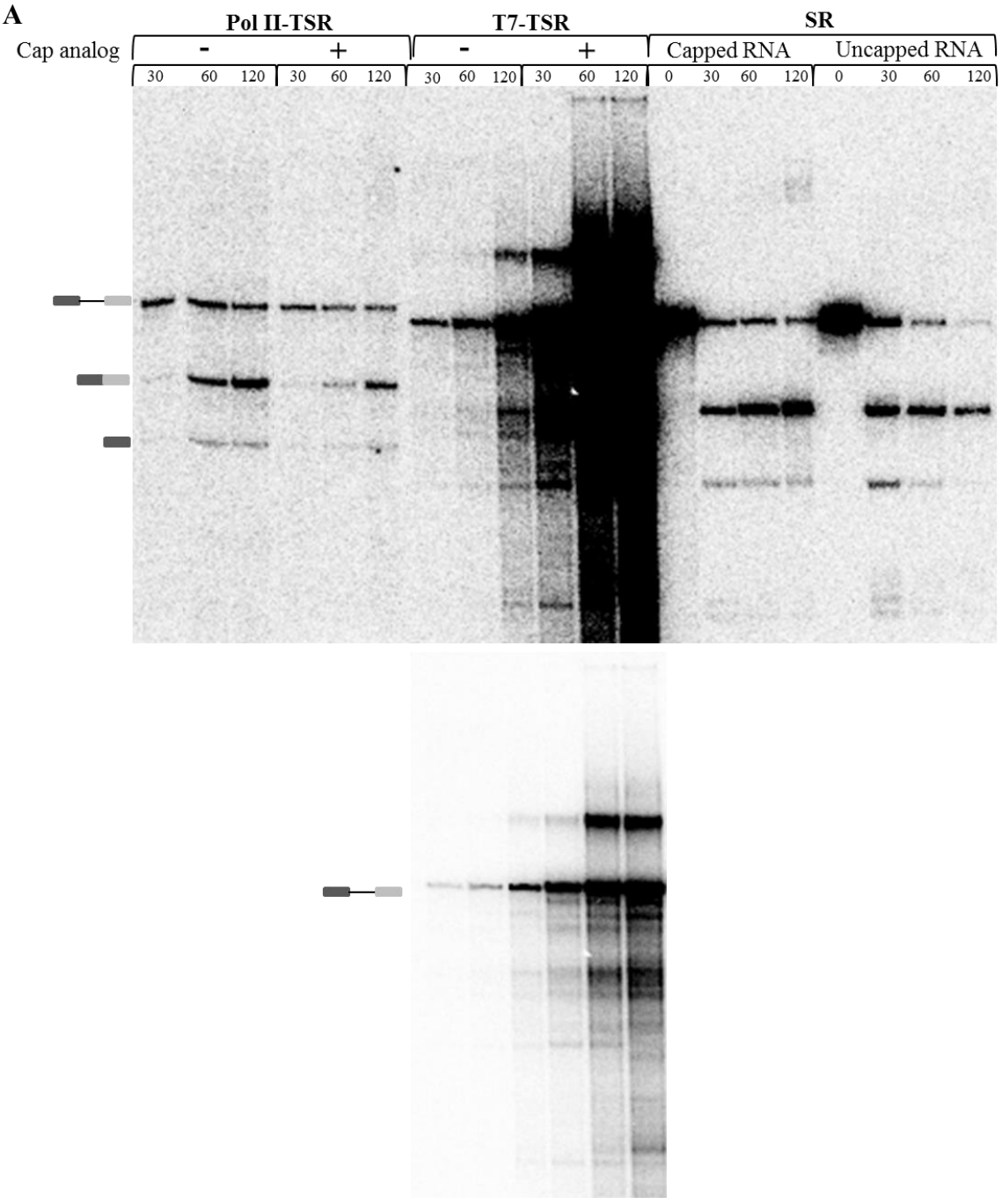
Figure 3.14. RNA transcribed by T7 polymerase undergoes degradation in HeLa nuclear extract. (A) Pol II and T7 transcription/splicing pulse chase assays. The zero time point refers to the addition of excess unlabelled GTP 30 min after the start of incubation. Icons illustrated to the right hand side of the gel indicate the positions of pre-mRNA, mRNA and 5' exon in Pol II-TSR. T7 splicing products were not indicated as it was difficult to differentiate them from products of degradation. (B Quantification of the degradation profiles versus time is shown in the graph.

3.4.2 Optimization of T7-TSR

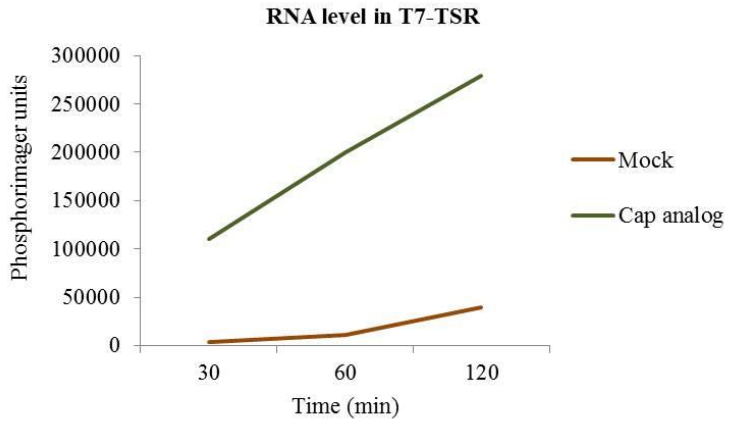
3.4.2.1 Testing cap analogue in T7-TSR

The cap structure of RNA does not only provide resistance to 5'-3'exonucleases (Furuichi et al., 1977; Shimotohno et al., 1977) but is a critical determinant of RNA metabolism including pre-mRNA splicing (Konarska et al., 1984; Edery and Sonenberg, 1985; Ohno et al., 1987; Patzelt et al., 1987; Jzaurralde et al., 1994; Lewis et al., 1996). Based on these data and the fact that transcripts generated by Pol II spliced efficiently (Figure 3.14), we suggested that capping, which occurs during Pol II transcription (Rasmussen and Lis, 1993), enhanced splicing whereas pre-mRNAs produced by T7 Pol did not undergo capping and, therefore, spliced poorly or degraded. Cap analogue (m^7G(5')ppp(5')G) is usually used in an *in vitro* transcription assay for incorporation at the 5' end of RNA. Further experiment to enhance mRNA yield in T7-TSR was done by using cap analogue. Cap analogue at a concentration used in a standard transcription assay (1mM) was added to T7- and Pol II-TSRs. The amount of input DNA in Pol II and T7 transcription/splicing assays was the same (10 ng/µl), but the concentration was 16 nM and 38 nM, respectively. As can be seen in Figure 3.15 A and B, transcription by T7 Pol was remarkably improved (7-fold), suggesting that RNA was capped in the reaction. A reason for the observed elevated levels of transcription can be explained that the cap structure protects RNA from 5' end degradation (Furuichi et al., 1977; Shimotohno et al., 1977). Some products of degradation but not mRNA were detected. In conclusion, the rate of splicing in this sample was not improved by the cap structure of the pre-mRNA.

There are two plausible reasons behind the observed poor splicing of transcripts generated by T7 Pol in the reaction with cap analogue. First, cap analogue could compete for cap binding complex (CBC), in line with work showing that CBC is involved in splicing (Izaurralde et al., 1994). Second, since the amount of RNA generated in T7-TSR was three-fold higher than that for Pol II-TSR (Figure 3.15 B and C), it is suggested that in T7-TSR splicing factors were titrated out among abundant transcripts leading to inability of the spliceosome to assemble on all pre-mRNAs. The latter is in line with the evidence on inefficient spliceosome assembly (Figure 3.13 B).



B



C

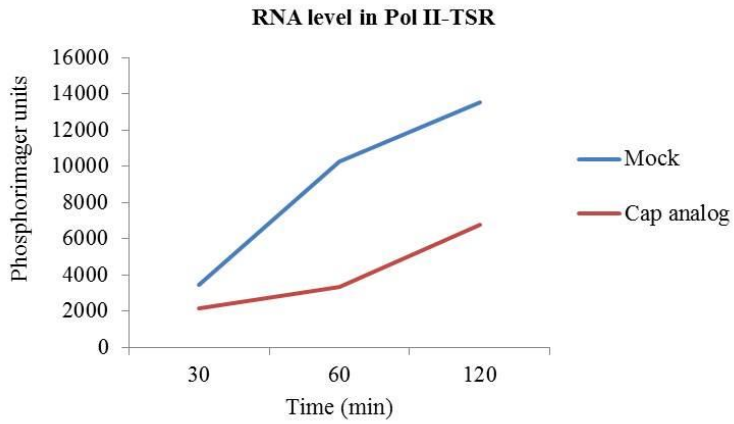


Figure 3.15. Yield of Bg C splicing is not increased if pre-mRNA is capped in T7 transcription/splicing system. Pol II and T7 transcription/splicing reactions either containing or lacking the cap analog (m7GpppG) were incubated during the indicated time. Pre-made capped and uncapped Bg C pre-mRNAs were spliced in standard splicing assay (SR). The mount of input DNA was 10 ng/µl (38 nM and 16 nM in T7-and Pol II-TSRs, respectively). All reactions were set up side by side. For better visualisation a part of the gel showing T7-TSR was adjusted. The graphs show RNA levels in Pol II-TSR (B) and T7-TSR (C).

RNA stability was compared in SR using capped and uncapped pre-synthesised Bg C pre-mRNAs. The experiment showed that 23% of pre-made capped transcripts and 7% of the uncapped ones were left after 120 min of incubation, suggesting that capped RNA is more stable than uncapped (Figure 3.15 A). Interestingly, cap analogue had a negative effect at the level of transcription in Pol II-TSR (Figure 3.15 C). It is known that the protecting role of the cap structure in RNA biogenesis is mediated by CBC which inter-

acts with the cap structure of pre-mRNA (reviewed by Gonatopoulos-Pournatzis and Cowling, 2014). It is possible that cap analogue competed for CBC, inhibiting the binding of CBC to capped RNA *in vitro*. These data fit the model that the pre-mRNA generated by Pol II *in vitro* undergoes capping and CBC interacts with the cap structure. If it was not the case the reaction (Pol II-TSR with cap analogue) would not be inhibited.

3.4.2.2 Magnesium chloride titration and changing T7 Pol storage buffer

Experiments were carried out to optimize T7-TSR assay with the aim of improving the mRNA yield. Usually the polymerase is kept in storage buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM DTT and 50% Glycerol). To exclude any inhibitory effect of the buffer on splicing, T7 Pol was added undiluted keeping the same absolute amount of it as it was in the previous experiments (5 ng/µl). Previously in the laboratory, the concentration of $MgCl_2$ was titrated showing that splicing required $MgCl_2$ but was not enhanced by increasing the concentration above 3.2 mM. In the experiments described above 3.2 mM $MgCl_2$ was used. To investigate whether $MgCl_2$ could improve splicing in T7-TSR, $MgCl_2$ concentrations of 1.6 mM and 3.2 mM were compared. Undiluted T7 Pol was used in all samples. The experiment showed that splicing efficiency was improved by 20% in T7-TSR with both $MgCl_2$ concentrations (Figure 3.16) in comparison to the previous experiment (Figure 3.12). This suggests that the storage buffer had a negative effect on splicing. An effect of T7 storage buffer on splicing was not tested in a separate experiment but it is possible that NaCl (50 mM final concentration) could negatively influence splicing. No difference in the efficiency of splicing was observed between two samples with different concentrations of $MgCl_2$ but the level of transcription was significantly increased with 3.2 mM $MgCl_2$. The concentration of $MgCl_2$ appeared to have no effect on splicing efficiency in Pol II-TSR (Figure 3.16). 3.2 mM $MgCl_2$ in all three assays was used in further experiments.

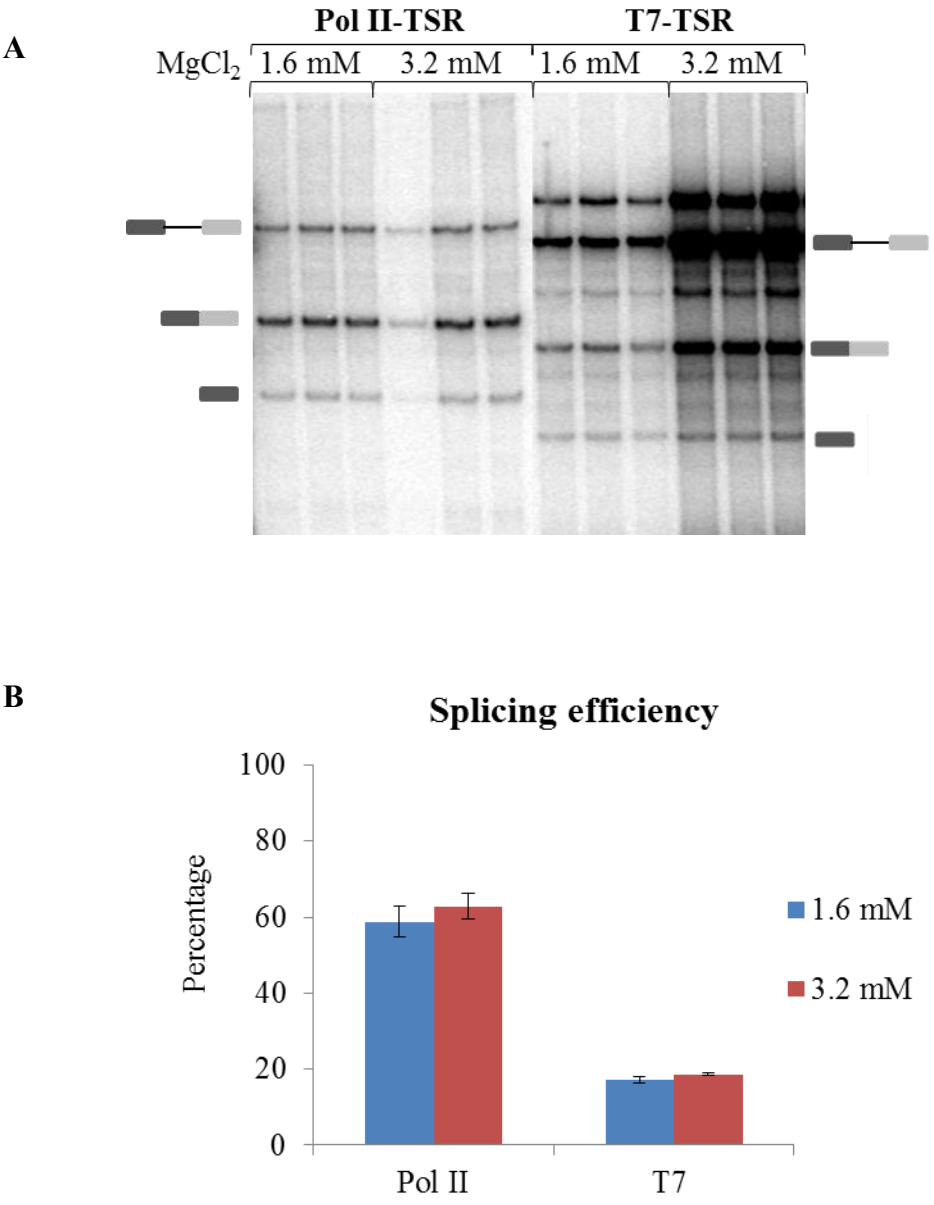
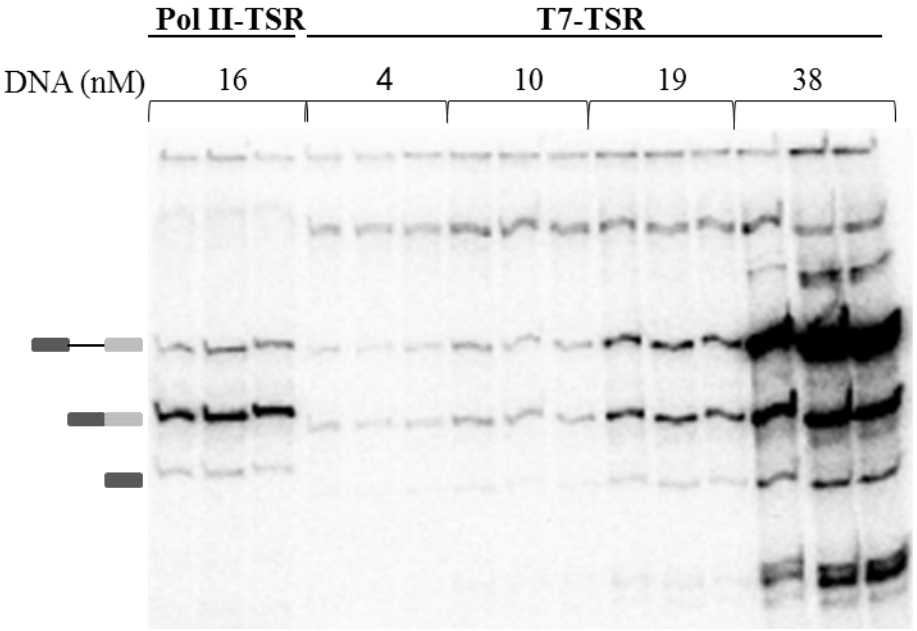


Figure 3.16. Magnesium does not enhance splicing but increases transcription in T7 transcription/splicing assay. (A) Denaturing PAGE shows titration of $MgCl_2$ in Pol II-TSR and T7-TSR in triplicates. The incubation was carried out for 90 min. (B) Quantification of splicing efficiency in two systems is shown in the graph. Error values are standard deviation values.

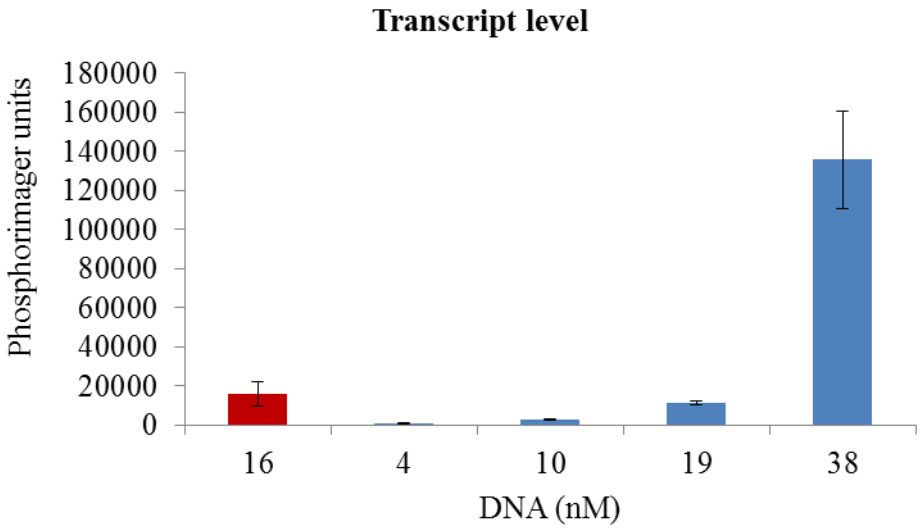
3.4.2.3 T7-Bg C DNA titration

The next step towards the improvement of splicing in T7-TSR was to optimize the concentration of T7-Bg C DNA. In the experiments characterised above, the amount of input DNA was 10 ng/µl (38 nM and 16 nM in T7-and Pol II-TSR, respectively). The concentration of the T7-Bg C DNA template was titrated from 38 nM to 4 nM. A sharp increase in the level of transcription was observed in the reaction with the highest DNA concentration (38 nM; Figure 3.17). This effect was consistent with results from some other experiments (data not shown), suggesting that it is reproducible. It is clear that when transcription reaches equal levels in both assays (19 nM and 16 nM of DNA in T7-and Pol II-TSR, respectively), the splicing efficiency is almost similar. These data in conjunction with experiments on the spliceosome complexes allow us to conclude that the reason behind poor splicing efficiency in T7-TSR was that splicing factors were titrated out among the abundant pre-mRNA molecules generated by T7 Pol. In conclusion, based on splicing efficiency Pol II does not affect splicing.

A


Pol II-TSR
T7-TSR
DNA (nM)
16
4
10
19
38

B


Transcript level
Phosphorimager units
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
16
4
10
19
38
DNA (nM)

C

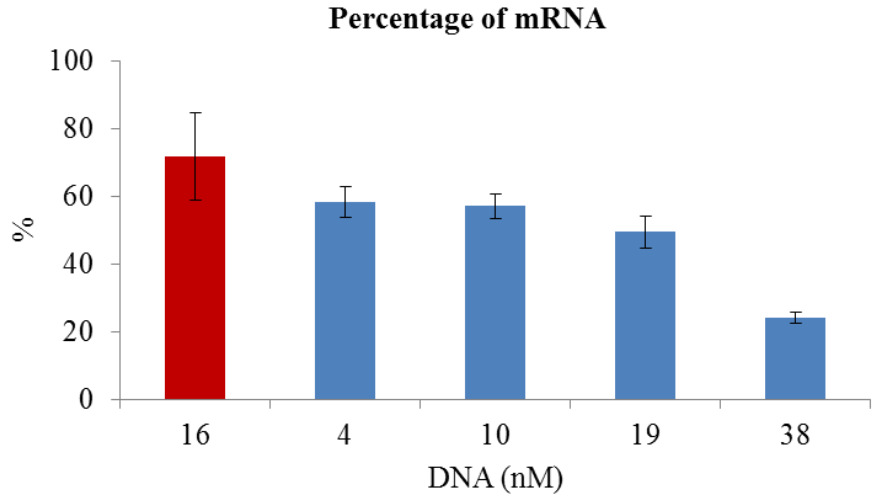

Percentage of mRNA
100
80
60
40
20
0
%
16
4
10
19
38
DNA (nM)

Figure 3.17. Yield of splicing is remarkably enhanced if less DNA is used in T7 transcription/splicing system. (A) Titration of T7-Bg C DNA template in transcription/splicing assay (sample volume was 5 µl). The concentration of CMV-Bg C and T7-Bg C DNA templates is represented. The concentration of input DNA in Pol II (16 nM) and T7 (38 nM) transcription/splicing assays was the same. The reactions were carried out for 90 min. (B) Quantification of RNA synthesis in Pol II (red) and T7 (blue) transcription/splicing systems is shown in the graph. Transcription level was quantified as sum of the pre-mRNA, mRNA and 5' exon. (C) The yield of spliced mRNA was calculated as the percentage by the following formula: (mRNA)/(pre-mRNA+mRNA+5'exon) and is shown in the graph. Error values are standard deviation values.

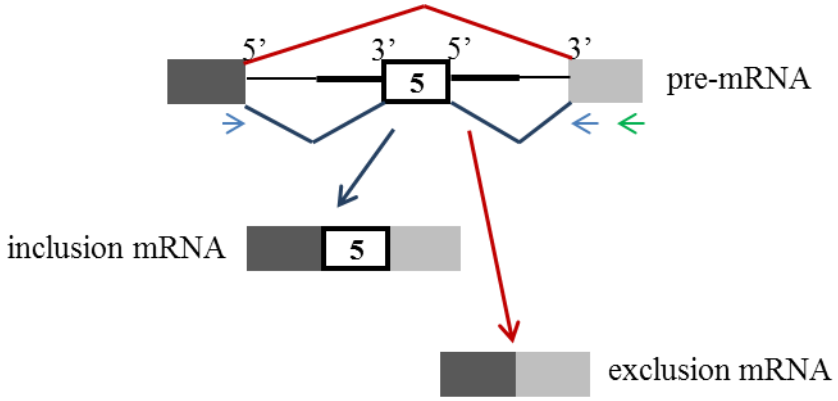
3.5 Effects of Pol II transcription on splicing

3.5.1 Splicing pattern of an alternatively spliced transcript is changed in Pol II-TSR

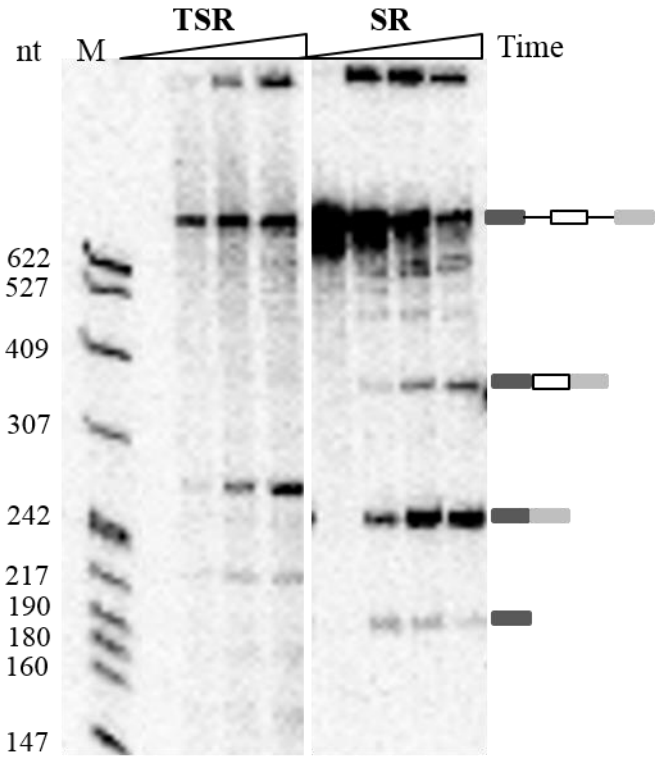
There is a large amount of data showing that the kinetics of transcription by Pol II and factors recruited to the CTD can influence alternative splice site choice (reviewed by Munoz et al., 2010; Kornblihtt et al., 2013). To examine whether the presence of Pol II transcription affects alternative splicing patterns *in vitro*, a mini-gene containing two constitutive exons and a variable exon in the middle was tested. It has been shown previously that the inclusion of exon 5 into mRNA is mediated by a splicing factor, Sam68, which is a downstream target in the MAPK pathway (Matter et al., 2002). In cells, this exon is usually skipped, but in case when the MAPK pathway is overactivated, Sam68 binds to ESE and ISE sequences in the pre-mRNA. This promotes the exon 5 inclusion. It has been shown that Sam68 associates with Brm, which is a component of the chromatin remodeling complex (SWI/SNF; Batsché et al., 2006). This led to a model in which Brm recruits Sam68 during transcription of CD44 pre-mRNA (Batsché et al., 2006). Therefore, CD44v5 substrate was of a particular interest to test in this study as a model for transcription-dependent alternative splicing.

Previously in the laboratory exon 5 (117 bp) with flanking intron sequences (238 bp upstream and 270 bp downstream) had been cloned into the intron of rabbit β-globin. To be able to use the CD44v5 construct in a transcription/splicing assay it was fused to a CMV promoter as described in Materials and Methods. The AS pattern of CD44v5 pre-mRNA was determined using Pol II-TSR and standard SR (Figure 3.18 A and B). The experiment showed that in SR CD44v5 precursor (930 nt) produced two mRNA isoforms: the inclusion mRNA (421 nt) which contains exon 5 between two Bg exons, and the skipping mRNA (304 nt) which lacks exon 5. A significant difference in the extent of variable exon inclusion was observed between the two systems (Figure 3.18 B). In particular, in SR 20 % of mRNA containing the inclusion isoform, whereas in TSR only the exclusion mRNA isoform was generated. Splicing efficiency was compared between SR and TSR samples, showing 68% and 52%, respectively (Figure 3.18 C). To test whether the result of Pol II-TSR is consistent with splicing outcome in cells, HEK 293T cells were co-transfected with plasmids containing CD44v5 mini-gene and either GFP or Sam68. Sam68 was overexpressed in cells to show that the CD44v5 mini-gene was functional and the pre-mRNA was able to splice producing two mRNA isoforms. The experiments were not performed in triplicate. It was found that exon inclusion reached only 7% in cells with GFP, whereas 61% was reached in cells with overexpressed Sam68 protein (Figure 3.18 D and E). By comparing the results obtained from Pol II-TSR and *in vivo* assays (sample of GFP and CD44v5 plasmids), it is apparent that the "weak" 3'SS flanking the variable exon is not favored and the distal 5'SS and 3'SS are recognized by the spliceosome resulting to skipping of the exon 5. These data show that in contrast to SR, transcription/splicing system recapitulates the *in vivo* outcome. Our data is in line with previous study (Hicks et al., 2006) which tested an effect of Pol II transcription on AS showing that an alternative exon was skipped in the reaction with Pol II transcription in comparison to the conventional splicing assay. It is suggested that the process of transcription could affect alternative splice site choice in the *in vitro* assay. Alternatively, the secondary structure of the newly transcribed pre-mRNA may be different to that of pre-made pre-mRNA leading to a different splicing pattern.

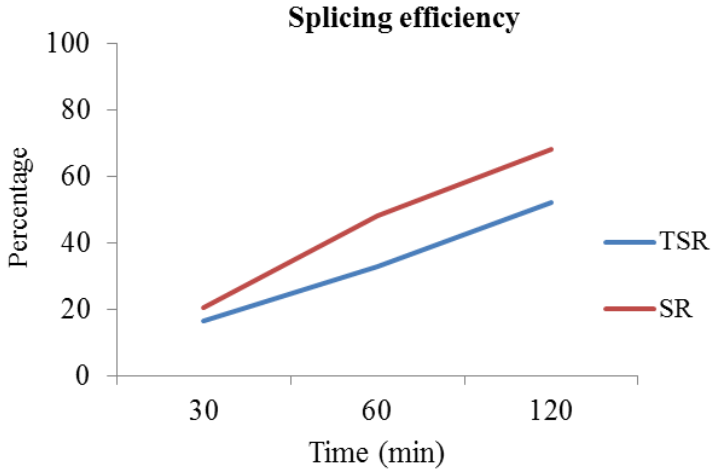
A


5'
3'
5'
3'
5
pre-mRNA
inclusion mRNA
5
exclusion mRNA

B


TSR
SR
nt
M
Time
622
527
409
307
242
217
190
180
160
147

C


Splicing efficiency
100
80
60
40
20
0
Percentage
30
60
120
Time (min)
TSR
SR

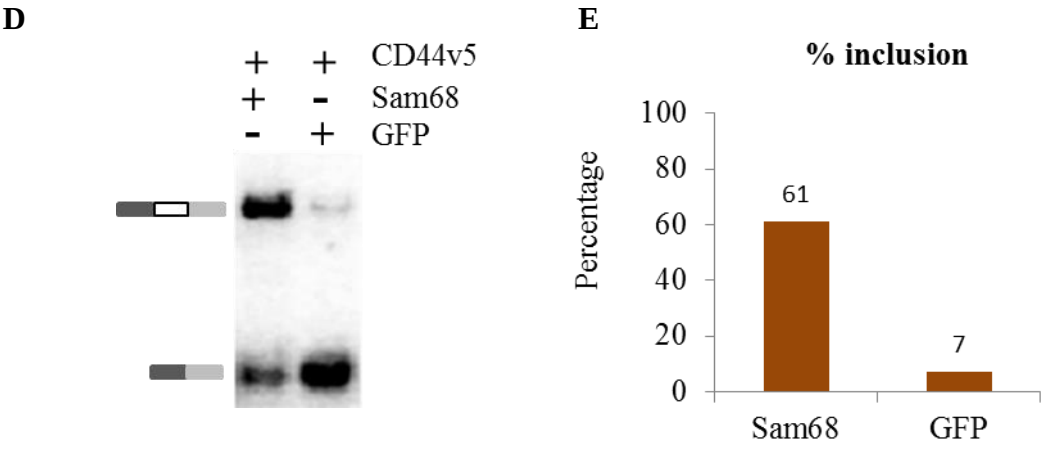


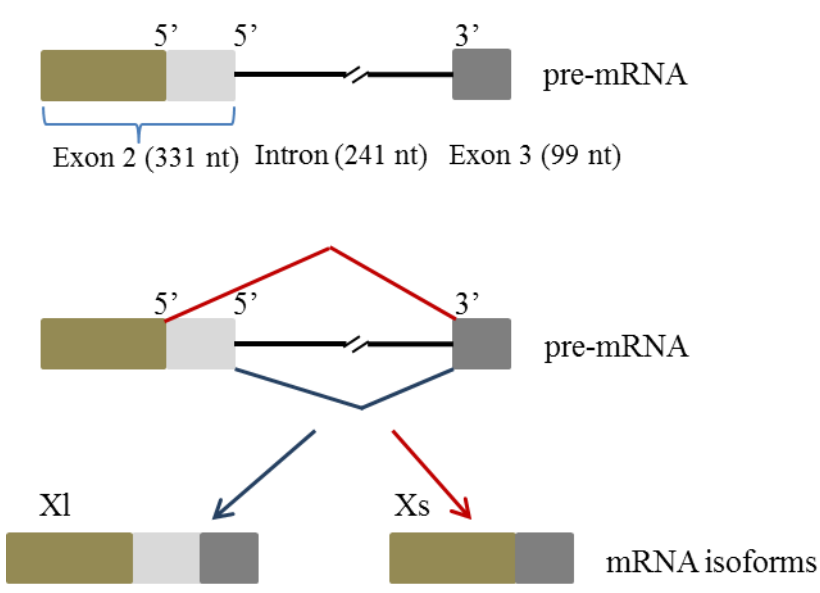
Figure 3.18. Pol II transcription affects alternative splice site choice. (A) Diagram shows alternative splicing pattern of CD44v5/Bg chimeric pre-mRNA. Bg exons (grey boxes), introns (black lines), CD44 exon v5 (open box), splicing alternatives (red and blue lines) and PCR primer positions (arrows) are illustrated. Thick parts of the line indicate flanking intronic regions of the exon 5. Thin parts of the line indicate intronic regions of Bg. (B) Time-courses of *in vitro* splicing and Pol II transcription/splicing assays using pre-made CD44v5 pre-mRNA and CMV-CD44v5 DNA template, respectively. Time points of 0, 30, 60 and 120 min were taken. The pre-mRNA, inclusion, exclusion and 5' exon products are indicated to the left side of the phosphorimage. (C) Splicing efficiency at each time point was measured as the percentage by the following formula: (inclusion +exclusion)/(pre-mRNA+ inclusion +exclusion +5'exon) and is shown in the graph. (D) RT–PCR analysis of HEK 293T cells co-transfected with CD44v5–mini-gene construct and either an expression construct for human Sam68 fusion to GFP (pLEICS25) or the empty expression vector (pEGFP-C1) which contained GFP. Empty pLEICS25 was not used as a control as it is lethal for cells. The experiment was not done in triplicate. cDNA was synthesized by use of a downstream Bg 3 Rev oligonucleotide (green arrow). PCR was done using radiolabelled primers (blue arrows) and samples were electrophoresed through native 6% PAGE. (E) Diagram shows the percentage of inclusion mRNA isoform from (D).

CD44v5 has a pattern of AS in which an alternative cassette exon can be either included or skipped. However, there are other patterns of AS such as usage of an alternative 5'SS, where there is a competition between two 5'SS located within one exon. If the upstream 5'SS is favoured by the spliceosome recognition and it is joined to the 3'SS, then splicing leads to the inclusion of the upstream portion of the exon, while the downstream part of it is removed together with an intron. As a result, a short mRNA isoform is produced. In contrast, the usage of the downstream 5'SS causes the full-

length exon to be included into mRNA. To investigate the ability of the transcription machinery to influence the 5'SS recognition *in vitro*, a construct having alternatively spliced 5'SSs was used. For this, Bcl-X construct containing two 5' SSs within the upstream exon was used (Weldon et al., 2016; Figure 3.19 A). The distance between the two 5'SS is 189 nt. The substrate contains a shortened intron (originally 55,000 bp long) of 241 bp in length having 180 nt from the 5' end of intron and 61nt from 3' end of intron. The length of the pre-mRNA, Xl and Xs was 671 nt, 430nt and 250nt, respectively. The Bcl-X construct was cloned into pEGFP-C1 plasmid (Materials and methods). CMV-Bcl-X and T7-Bcl-X DNA templates were used in Pol II-TSR and SR, respectively.

Pulse chase of transcription/splicing assay was done to determine the relative use of 5' SSs (Figure 3.19 B). Bcl-X splicing products and intermediates were assigned by *in vitro* SR (Weldon et al, 2016). The experiment showed no difference in the Xl/Xs ratio between Pol II-TSR and SR assays at the last time point (Figure 3.19 C). Importantly, the splicing pattern of the Bcl-X substrate *in vivo* had been tested in the laboratory (Weldon et al., 2016) and the result showed a similar ratio between two mRNA isoforms. Since the three systems are consistent in the Bcl-X splicing outcome, it is suggested that in the case of Bcl-X, transcription by Pol II does not influence the recognition of 5'SSs.

A



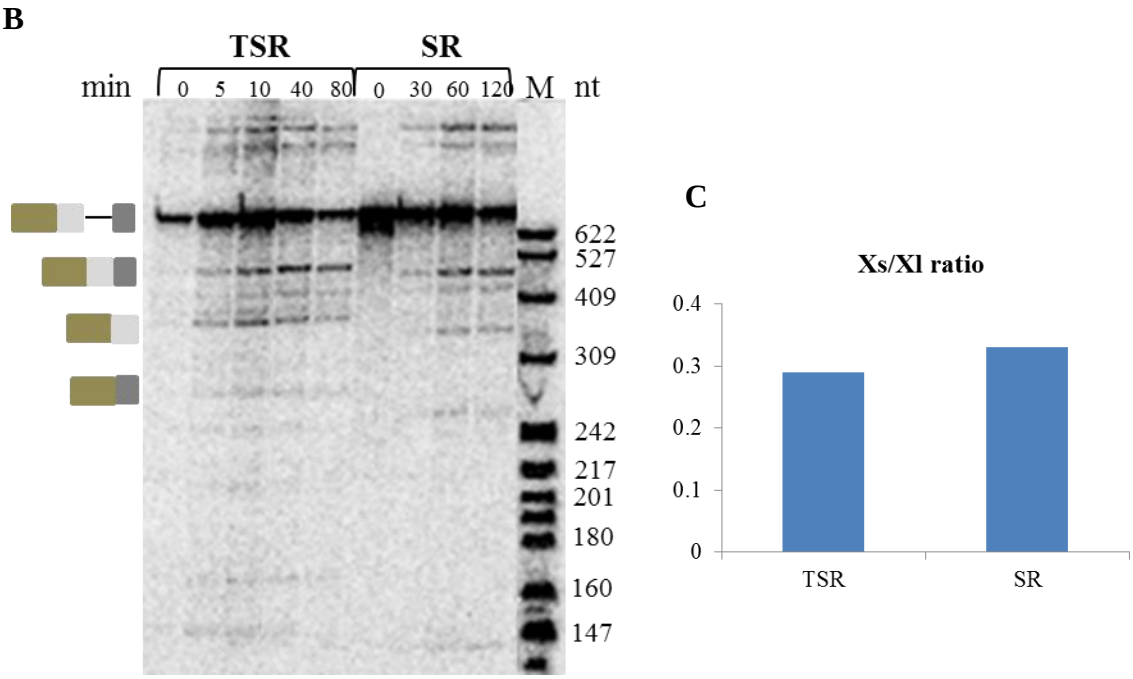


Figure 3.19. Transcription does not affect splicing of Bcl-X substrate. (A) Diagram shows alternative splicing pattern of Bcl-X pre-mRNA. Exons (boxes), intron (black line), splicing alternatives (red and blue lines) are illustrated. The alternative 5' splice sites are indicated. (B) Time-courses *in vitro* Pol II transcription/splicing and splicing assays using CMV-Bcl-X DNA and pre-made Bcl-X pre-mRNA, respectively. Pulse chase of Pol II-TSR was done after 40 min of incubation and this time was referred as zero min time point. The pre-mRNA, long and short products are annotated. (C) Diagram shows the ratio between short and long mRNAs at the last time points (120 min).

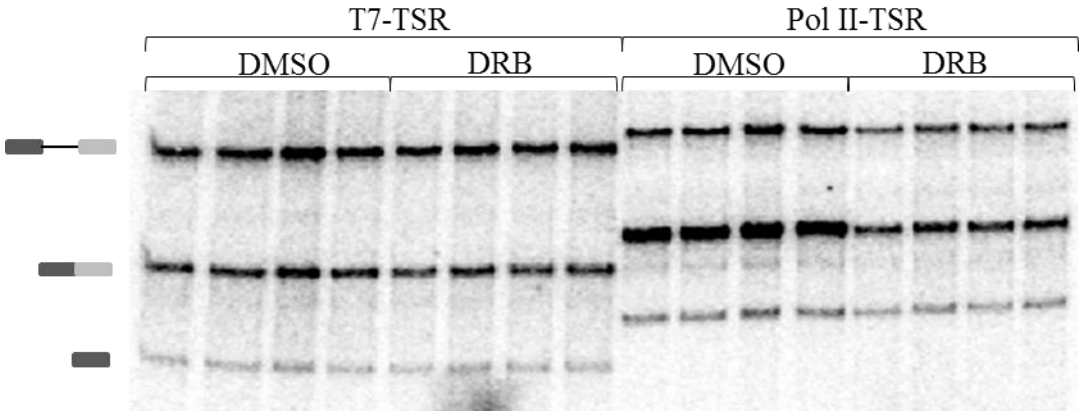
3.5.2 Effects of CTD phosphorylation on splicing

The transcription repressor DRB is a molecule that is able to inhibit the elongation stage of transcription (Zandomeni et al., 1986) through targeting several kinases that phosphorylate the CTD. DRB does not exhibit a particular selectivity as it has been found to inhibit the activity of CDK7 (Yankulov et al., 1995), CDK8 (Rickert et al., 1999) and CDK9 (Marshall et al., 1996; Saunders et al., 2006). An effect of Pol II transcription on splicing by inhibiting transcription elongation by DRB *in vivo* was tested (Bird et al., 2004). It was found that efficient splicing required CTD phosphorylation at Ser2 when coupled to transcription but not when splicing occurred in a transcription-independent manner. However, another *in vivo* study showed that DRB inhibited transcription but splicing was not affected (Koga et al., 2015).

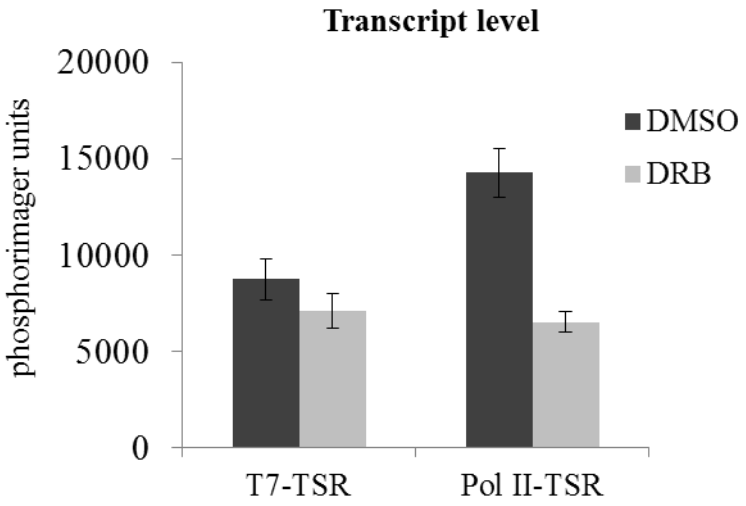
To test whether the phosphorylation state of the CTD affects splicing outcome in our *in vitro* Pol II-TSR assay, DRB was used. The control systems in which splicing is not linked to transcription (T7-TSR and SR) were also treated with this molecule. We hypothesised that if splicing efficiency would be affected by DRB treatment in Pol II-TSR but remained unchanged in T7-TSR and SR then the result would allow to suggest that Pol II transcription affects splicing co-transcriptionally. By treating cells with DRB as it was done in previous study co-transcriptional splicing was negatively affected (Bird et al., 2004). However, it is possible to deactivate proteins other than CDK7, CDK8 and CDK9 that are also required for splicing activity. Therefore, we also aimed to determine whether kinases necessary for phosphorylation of splicing factors are affected by DRB. We hypothesised if they were not sensitive to DRB then splicing would not be changed in assays such as T7-TSR and SR.

The result revealed that Pol II transcription level was decreased upon DRB treatment (2.5-fold), while T7 transcription efficiency was not changed (Figure 3.20). There was 14% less mRNA production in the reaction in which transcription was carried out by Pol II (Figure 3.20). No difference in splicing efficiency in either T7-TSR or SR was detected. It is suggested that in Pol II-TSR the activity of splicing regulators was not affected by DRB treatment.

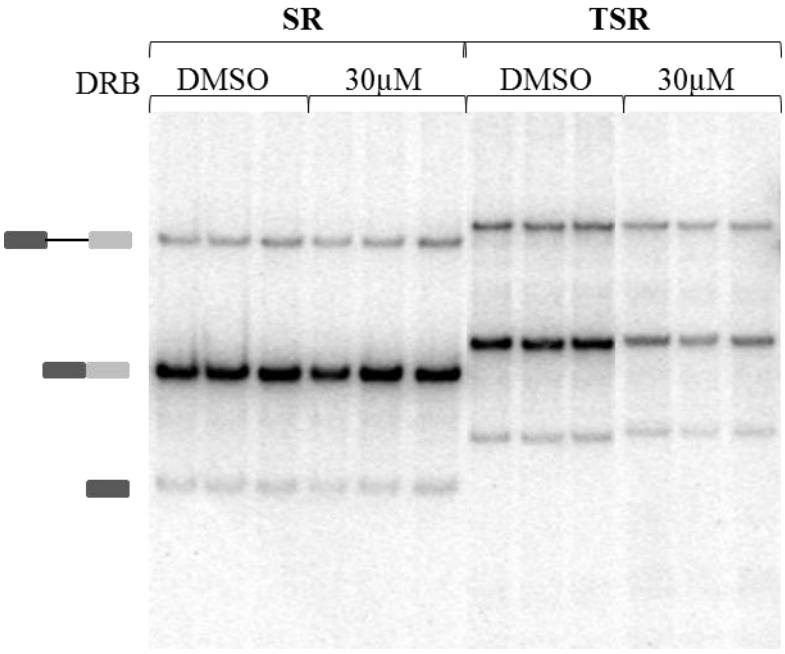
A


T7-TSR
Pol II-TSR
DMSO
DRB
DMSO
DRB

B


Transcript level
20000
15000
10000
5000
0
phosphorimager units
DMSO
DRB
T7-TSR
Pol II-TSR

C


SR
TSR
DRB
DMSO
30µM
DMSO
30µM

D

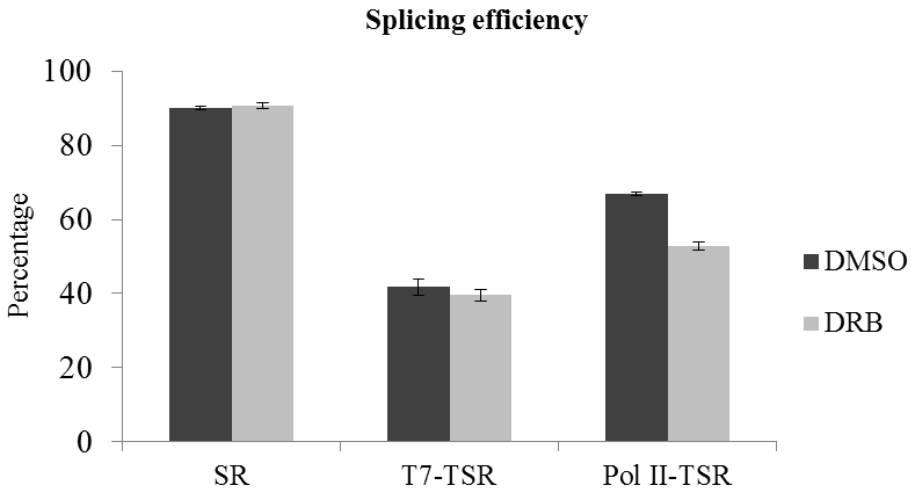


Figure 3.20. Inhibition of Pol II transcription elongation by DRB results in reduced splicing efficiency in Pol II-TSR. Effect of DRB on splicing efficiency in splicing, Pol II transcription/splicing (A) and T7 transcription/splicing (B) assays. DRB was dissolved in DMSO and added to reactions to produce the final concentration shown. (A) T7 and Pol II transcription/splicing reactions were incubated in the presence of DRB at the concentration of 30 μM. (A and B) The reactions were done in triplicate and incubated for 90 min. (C) Diagram shows quantification of RNA synthesis in Pol II and T7 transcription/splicing systems. (D) Diagram shows splicing efficiency which was measured using data from (A) and (C). Error values are standard deviation values.

Since the amount of the pre-mRNA synthesised by Pol II was decreased after DRB treatment, it was necessary to test whether less transcripts are poorly spliced. Two- and three-fold less CMV-Bg C DNA compared to the previous experiment (10 ng/μl) was used. Carrier DNA was added to keep the same overall amount of DNA in each sample (50 ng/μl). The experiment showed no difference in the level of transcription between mock-treated samples regardless of DNA concentration. The efficiency of splicing was reduced by 13% in all treated samples (Figure 3.21). This is in line with our previous results (Figure 3.5 and Figure 3.6) which showed that not all DNA templates were available as transcription substrates in Pol II-TSR. Together the results suggest that about three fourth of CMV-Bg C DNA is either degraded or get adsorbed onto the plastic walls of a test tube.

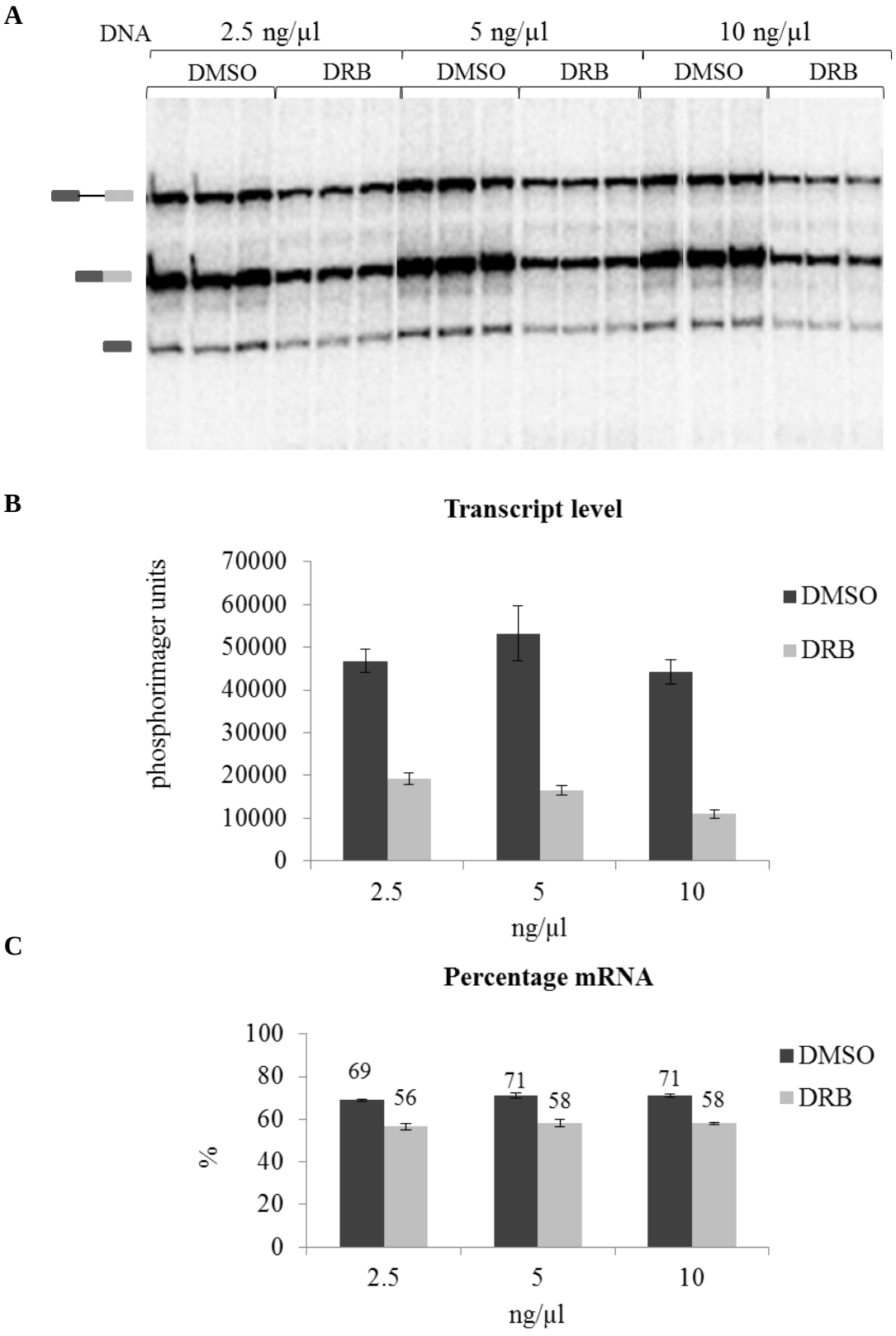


Figure 3.21. Testing DRB in Pol II-TSR using different concentrations of CMV-Bg C DNA. The experiment was done in triplicate and incubated for 90 min. DRB was dissolved in DMSO (1% final concentration). DRB and DMSO were added prior the beginning of incubation. Diagrams show quantification of transcription level (B) and mRNA splicing (C).

To investigate if splicing efficiency depends on the amount of transcripts, the concentration of premade pre-mRNA was varied in SR (Figure 3.22). It has to be noted that the amount of RNA generated in transcription/splicing assay was not determined. The result showed that the NE supports better splicing if fewer pre-mRNA is present. In conclusion, the observed decrease in splicing efficiency in Pol II-TSR upon DRB treatment (Figure 3.20) was not due to the decrease of transcript amounts.

Although splicing was affected only by 14 % in Pol II-TSR (Figure 3.20) but the fact that splicing remained unchanged in either SR or T7-TSR systems, it is suggested that the change in splicing outcome in Pol II-TSR was not due to the inhibition of proteins necessary for splicing. This result is in line with previous findings which showed that CTD phosphorylation on Ser2 residue ensures efficient pre-mRNA splicing when it is linked to Pol II transcription but not when splicing occurs post-transcriptionally (Bird et al., 2004). It is suggested that splicing efficiency decrease is a result of inhibited CTD Ser2 phosphorylation and the two reactions could be functionally coupled in Pol II-TSR.

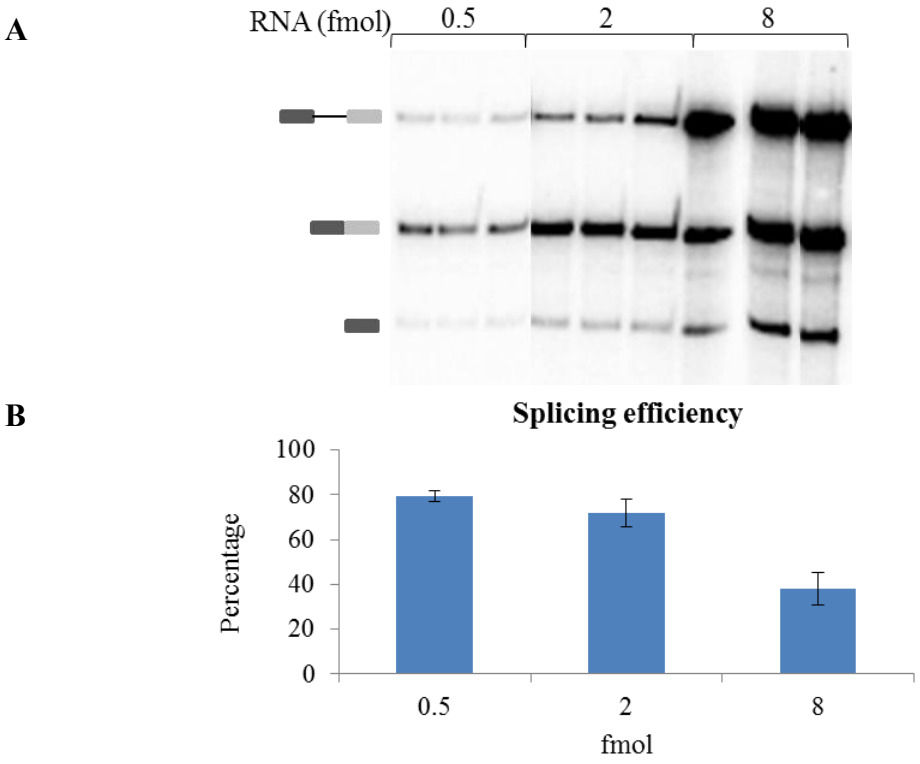


Figure 3.22. Splicing is more efficient if less pre-mRNA is present. (A) Pre-made Bg C pre-mRNA titration in the *in vitro* splicing assay. The experiment was done in triplicate and incubated for 90 min. The concentration of pre-mRNA is indicated. (B) Diagram shows quantification of splicing efficiency from (A).

3.5.3 Characterization of spliceosome complexes

In order to investigate whether Pol II transcription affects spliceosome assembly, the complexes formed on Bg C RNA in this reaction need to be defined. SR is supplemented with pre-made transcripts at high concentration and this leads to the synchronised formation of the complexes and easy technical detection. In contrast, transcripts have to be synthesised in the transcription/splicing assay and their amount should be high enough for detection. In TSR the precursor was observed as early as 3 min (Figure 3.11) making spliceosome investigations possible at the earliest stages of incubation. The spliceosome complexes formed in optimized Pol II transcription/splicing and conventional splicing systems were characterised side-by-side. Since the highest transcription efficiency was obtained when 150 ng of carrier DNA and 25 ng of CMV-DNA were used (Figure 3.6), these conditions were applied in the spliceosome complexes assay.

In SR complexes A, B and C were assembled at 1, 5 and 10 min time points, respectively (Figure 3.13), suggesting that the pattern of the spliceosome complexes obtained in this study was in line with previous results (Hodson et al., 2012). Complexes formed in Pol II-TSR will be compared to those in SR. The inhibition of the 5' end of U1 snRNA with anti-U1 oligonucleotide (Materials and Methods) at 5 μM led to a loss of some complexes (Figure 3.23 A). This experiment allows to determine complexes which have splicing origin as they are dependent on the presence of the U1 snRNA (Reed, 1990). Apart from the spliceosome complexes, other bands were visualised which were unaffected by the inhibition of the U1 snRNA. This suggests that these are splicing-unrelated complexes. There was a possibility that they were either transcription or capping complexes. We hypothesised if the complexes were related to transcription then they would still be present regardless of any radiolabelled nucleotides used, or if they were capping complexes then they would be lost as capping factors can trap GTP (Chapter 1). To test this hypothesis, the reaction was supplemented with [α-^{32}P] UTP. As a result, the complexes were lost (Figure 3.23 B). This observation suggests that bands observed in TSR sample treated with anti-U1 oligonucleotide (Figure 3.23 A) were capping complexes.

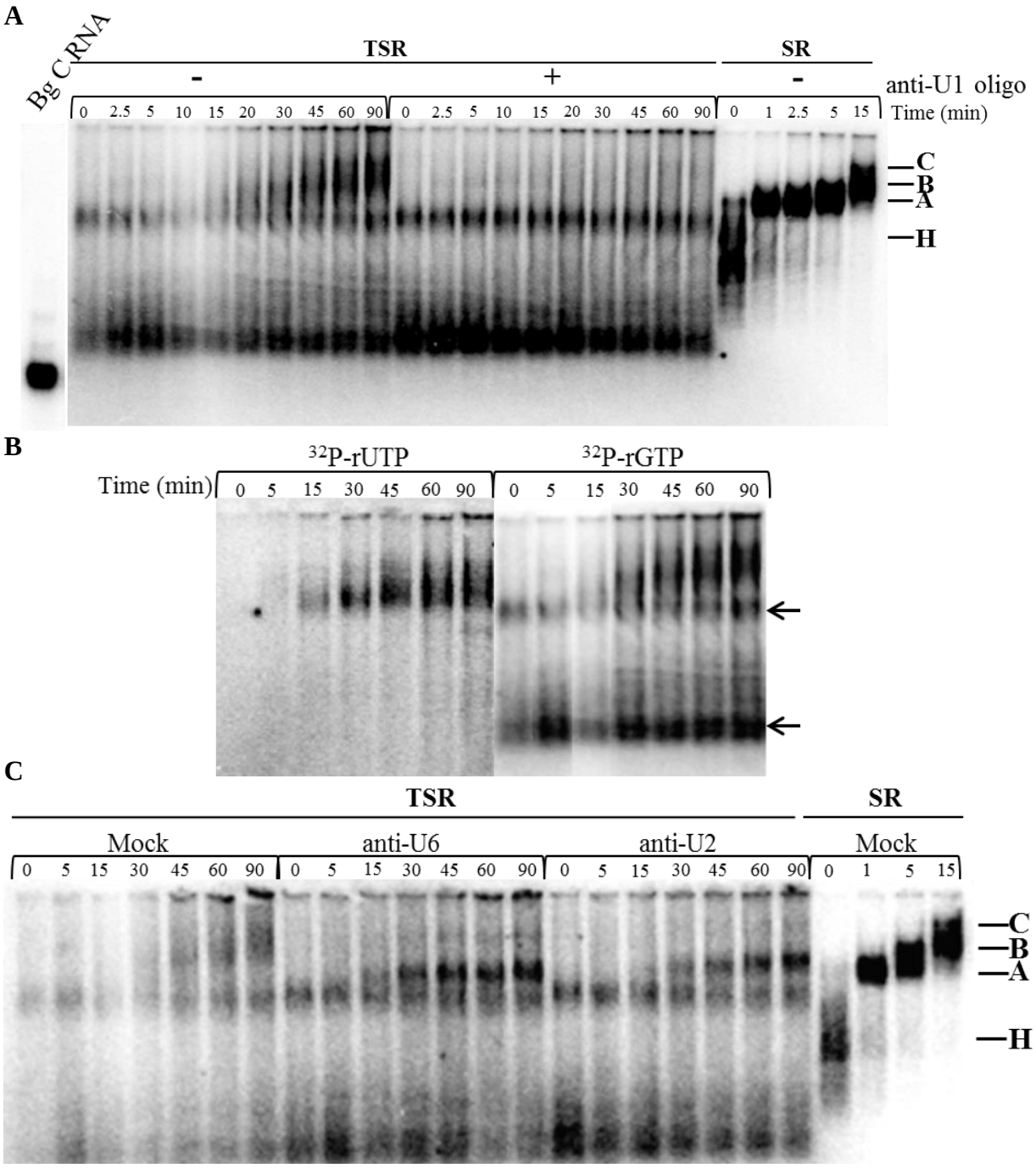


Figure 3.23. Analysis of ATP-dependent spliceosomal complexes assembly in Pol II transcription/splicing and splicing systems. (A) 2% LMP agarose gel shows time course of spliceosome complex assembly in Pol II transcription/splicing (TSR) and splicing (SR) assays. Splicing complexes assembled on radioactively labelled pre-made Bg C pre-mRNA in SR. Anti-U1 snRNA 2'-*O*-methyl oligonucleotide was used to identify the spliceosome complexes. Naked Bg C pre-RNA was loaded to assign the precursor produced in TSR. (B) Capping-related complexes (shown by arrows) were assigned in the sample with ^{32}P-GTP as complexes were missing in the sample with ^{32}P-UTP. (C) The spliceosome was inhibited by using 2'-*O*-methyl oligonucleotides complementary to the U2 and U6 snRNAs.

This result allowed the spliceosomal complexes to be assigned (Figure 3.23 A). Since transcription was not inhibited in the mock reaction, it was expected that the similar bands would be detected during the whole incubation time. However, this was not the case as different bands were detected. It is possible that transcripts were not continuously generated. Our result showed that from 50 to 120 min almost steady-state levels of RNA were observed (Figure 3.6). It is possible that nucleotides or ATP might me limiting factors for transcription. The experiment showed that complexes sensitive to anti-U1 oligonucleotide were detected at 15 min and their qualitative and quantitative characteristics changed over time. Also the intensity of bands was increasing, indicating that the complexes were accumulating over the course of incubation time. Complexes at 45, 60 and 90 min in Pol II-TSR comigrated with those formed at 15 min in SR. It is known that the speed of Pol II transcription *in vivo* is approximately 20-40 nt per sec (Jonkers et al., 2014). Even if transcription is slower *in vitro* but, still, there is more time needed for the splicesome complexes progression in Pol II-TSR in comparison to SR. This data suggests that transcription by Pol II affects the kinetics of spliceosome assembly and it is possible that splicing occurs when the pre-mRNA is released from Pol II in TSR.

To investigate if the spliceosome complexes in Pol II-TSR contain the core components, snRNPs were blocked by antisense 2'-*O*-methyl oligonucleotides. The spliceosome was inhibited by incubating the NE with anti-U2 and anti-U6 oligonucleotides (Donmez et al., 2007). The result showed stalling of the spliceosome (Figure 3.23 C). The complex in the sample with anti-U6 oligonucleotide had the same size as complex A in SR. In conclusion, the U2 and U6 snRNAs participate in spliceosome assembly in the reaction with Pol II transcription and complex A is assembled in Pol II-TSR.

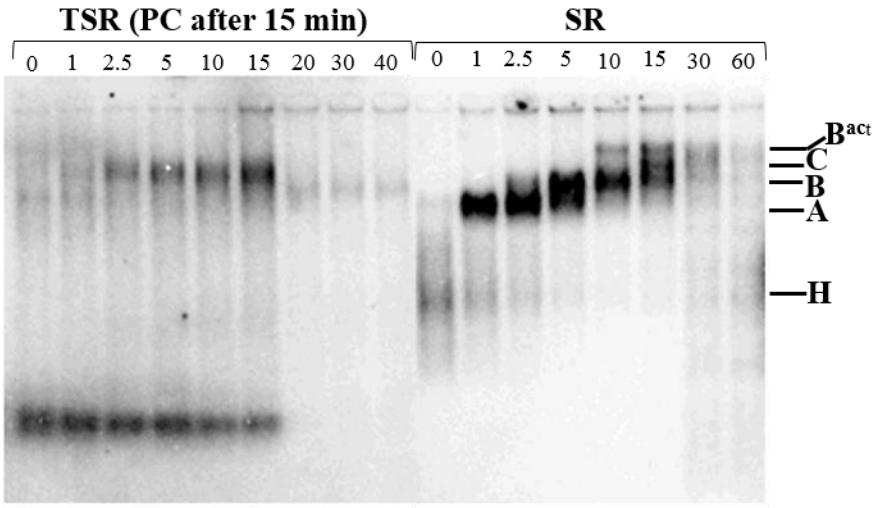
The experiments (Figure 3.23 A and C) showed that early splicing complexes are detected at 15 min. In the next experiment pulse chase was done after 15 min of the reaction start to follow the spliceosome progression. After pulse chase two bands are lost. Based on the previous experiment they were assigned as capping complexes (Figure 3.23 B). As can be seen from Figure 3.24 A, there was just one complex detected after pulse chase having the same size as complex B in SR. Any other complexes including complex H formed in SR were not detected in Pol II-TSR. This was consistent with the result obtained from the experiment in which the CMV-CD44v5

DNA template was used (Figure 3.24 B). These observations were in line with previous study in which complex H was absent in Pol II-TSR (Das et al., 2006). An attempt to detect other spliceosome complex at later time by using pulse chase assay was not successful as still just one complex was observed at 30, 45 and 60 min (data not shown).

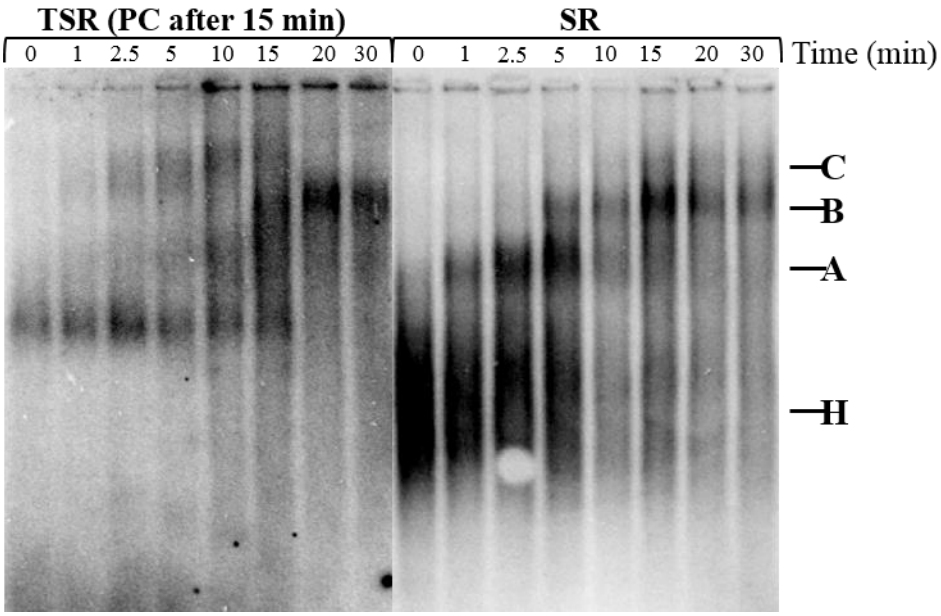
In the next experiment we wanted to detect complexes at early time of incubation. The spliceosome was stalled at complex A using anti-U6 oligonucleotide. Pulse chase after 5 min was done. Time points at 6, 10 and 15 min were taken. The result revealed that the spliceosome was stalled at complex A (Figure 3.24 C). It is concluded that in the mock reaction the single band is a splicing complex because it was not present in the reaction with inhibited U6 snRNA. Since just one bigger complex was detected in the mock-treated sample and the fact that complex A is able to assemble in Pol II-TSR (Figure 3.24 C) it is suggested that complex A was formed before 6 min in the mock sample. Although the single complex comigrates with complex B in SR (Figure 3.24 A and B), it is not possible to assign it as complex B because any bigger complexes were detected. The nature of this complex can be taken for further investigation. For example, Bg construct which lacks the 3'SS can be used in parallel. It is expected that on the mutant pre-mRNA the spliceosome will be stalled at complex C. Therefore, the comparative mobility of bands can indicate if this is complex B or C.

In conclusion, transcription by Pol II affects the spliceosomal complex assembly rate. Also, a remarkable feature was observed, in particular, there was no complex H detected in Pol II-TSR. This is in line with our previous result showing that complex H was not formed on the pre-mRNA generated by T7 Pol (Figure 3.13). In SR pre-made pre-mRNA is added to the NE in excess, it is bound by numerous proteins forming complex H before the spliceosome recognises the splice sites (Bennett at al., 1992). This suggests that newly transcribed pre-mRNAs generated by either Pol II or T7 Pol are committed to splicing as soon as they have been transcribed. Out results agree with previous study (Das et al., 2006) which showed that complex H is not formed in Pol II-TSR but disagree regarding complex H formation in T7-TSR.

A


TSR (PC after 15 min)
SR
0 1 2.5 5 10 15 20 30 40
0 1 2.5 5 10 15 30 60
Bact
C
B
A
H

B


TSR (PC after 15 min)
SR
0 1 2.5 5 10 15 20 30
0 1 2.5 5 10 15 20 30
Time (min)
C
B
A
H

C

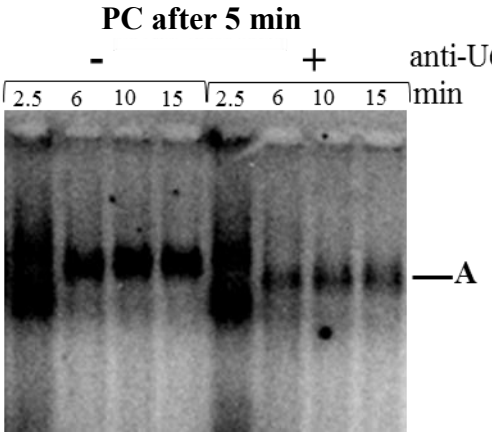

PC after 5 min
-
+
anti-U6
2.5 6 10 15 2.5 6 10 15
min
A

Figure 3.24. A single spliceosome complex is detected in Pol II transcription/splicing assay. (A) Analysis of the spliceosome complexes by electrophoresis in a 2% LMP agarose gel. CMV-Bg C DNA template and pre-made Bg C pre-mRNA were used in TSR and SR, respectively. (B) CMV-CD44v5 DNA template and pre-made CD44v5 pre-mRNA were used in TSR and SR, respectively. (A and B) To visualise complexes in Pol II transcription/splicing system pulse chase (PC) was done after 15 min of the start of the reaction. The complexes formed in splicing reaction are assigned in letters. (C) To detect complexes in Pol II transcription/splicing system with CMV-Bg C DNA, PC was done after 5 min of the start of the reaction. The spliceosome complex A was stalled by using 2'-*O*-methyl oligonucleotide to block the U6 snRNA. (A, B and C) Incubation times are represented at the top of each phosphorimager. After pulse chase two bands are lost.

3.6 Characterisation of transcription complexes in Pol II-TSR

In the present study, I wanted to investigate whether splicing factors affect transcription complex formation on the CMV promoter-containing DNA. Non-denaturing gel electrophoresis is used to great advantage in defining spliceosome complex assembly and RNA-protein interactions. Thus, we hypothesized that the same method would be applicable to study transcription complexes. For this purpose, native agarose gel electrophoresis was done using samples with radiolabeled DNA. The transcription complexes needed to be characterized first.

A series of experiments were performed using Bg C DNA template containing the CMV promoter that contained radiolabeled nucleotides incorporated during PCR (Materials and Methods). To assign transcription elongation complexes, Actinomycin D was used to inhibit transcription (Figure 3.25). The DNA template was incubated under optimized conditions for the Pol II transcription/splicing assay. Native agarose gel electrophoresis was done using the same conditions as for the spliceosome complexes (Materials and Methods). Time points were also taken for PAGE, to prove transcription inhibition (Figure 3.25 B). The distribution of species containing DNA in native gel had the appearance of a ladder (Figure 3.25 A). As expected, the electrophoretic mobility of bands was different in the reaction with Actinomycin D. At 0 min time point the bands in both reactions looked similar, suggesting that at the beginning of incubation DNA molecules can be free from proteins and transcription has not started yet. Shorter ladder consistent with reduced elongation. As transcription progresses, different proteins are recruited to the transcription machinery, therefore, this leads to DNA shifts during electrophoresis. However, the appearance of discrete bands has yet to be accounted for.

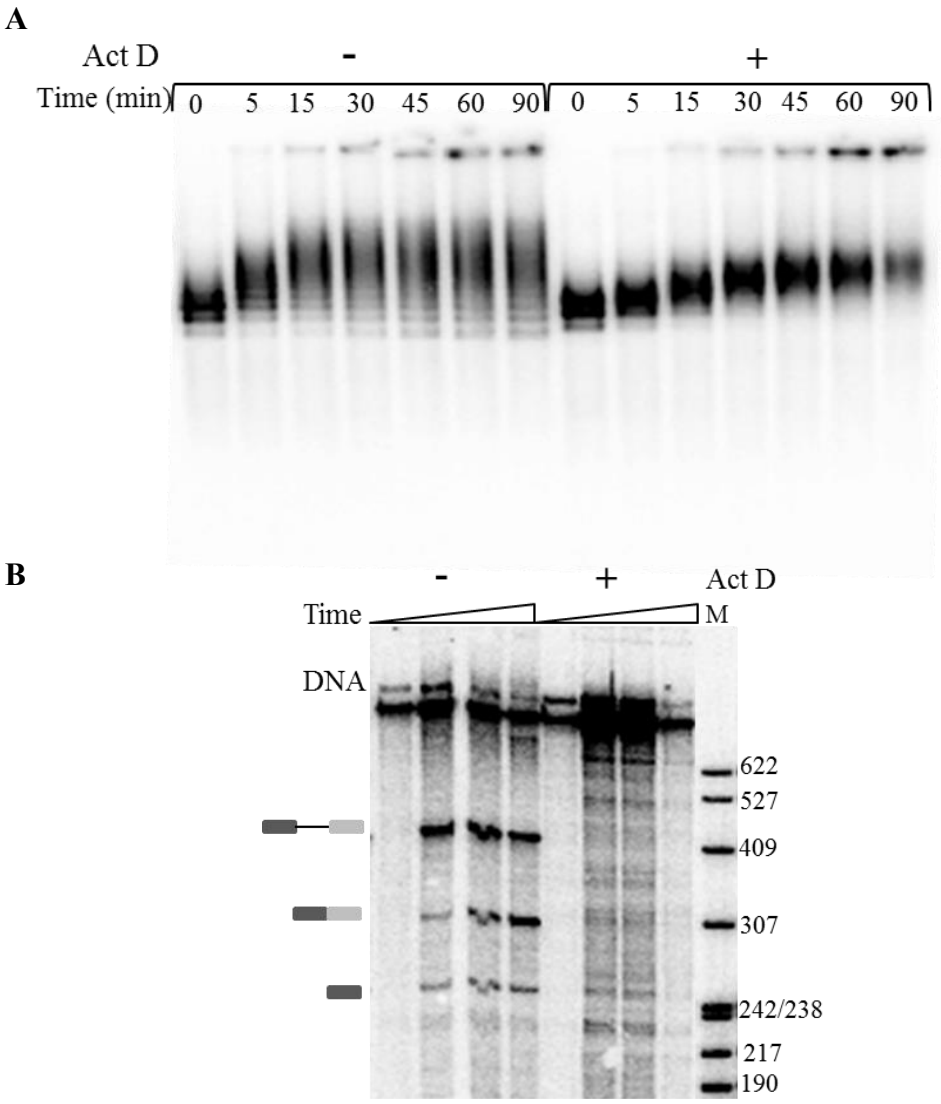


Figure 3.25. Characterisation of DNA-protein complexes formed in Pol II transcription/splicing system on native agarose gel. (A) DNA-protein complexes were fractionated on 2% low-melting-point agarose gel. [α-^{32}P]dTTP was incorporated into the Bg C DNA template during PCR to obtain a radiolabeled product. Actinomycin D (Act D) at a concentration of 5 μM was used to inhibit transcription. DMSO was added to the control sample as Act D had been dissolved in DMSO. (B) Denaturing PAGE of the same samples from (A). DNA, transcription and splicing products are indicated on the left side of the panel. Time points of 15, 30, 60 and 120 min were taken.

To explore the possibility that splicing can affect transcription complexes formation, a DNA template lacking an intron was used. The CMV-GFP DNA template was generated by PCR using EGFP-C1, which contains GFP gene downstream of the CMV promoter. Primers were designed in a way to amplify the DNA template to be similar in length as the Bg C DNA template was. To perform native gel electrophoresis radiolabeled CMV-DNAs were added to Pol II-TSR assay (Figure 3.26 A; samples

[32P]Bg C DNA and [32P]GFP DNA). These reactions did not contain radiolabeled nucleotides. In the absence of the NE, naked DNAs and RNA were detected as single bands. This allowed to determine that in the NE DNA and RNA are bound by proteins. The result showed that transcription of the CMV-GFP DNA was carried out but splicing did not occur (Figure 3.26 B). There was no difference in the pattern of DNA migration in the reactions which had either CMV-Bg C or CMV-GFP DNA templates. In a parallel reaction unlabeled CMV-Bg C DNA and [α-^{32}P] GTP nucleotide were added to detect splicing complexes after pulse chase at 15 min of incubation ([32P]rGTP sample). After pulse chase two major bands were lost. Based on the previous result (Figure 3.23), these are assigned as capping complexes. A single splicing complex was observed at 20 and 30 min time points ([32P]rGTP sample). Although splicing and transcription complexes comigrated (samples: [32P]Bg C DNA and [32P]rGTP), based on these observations it was not possible to determine whether the spliceosome was assembled on the transcribing or released pre-mRNA.

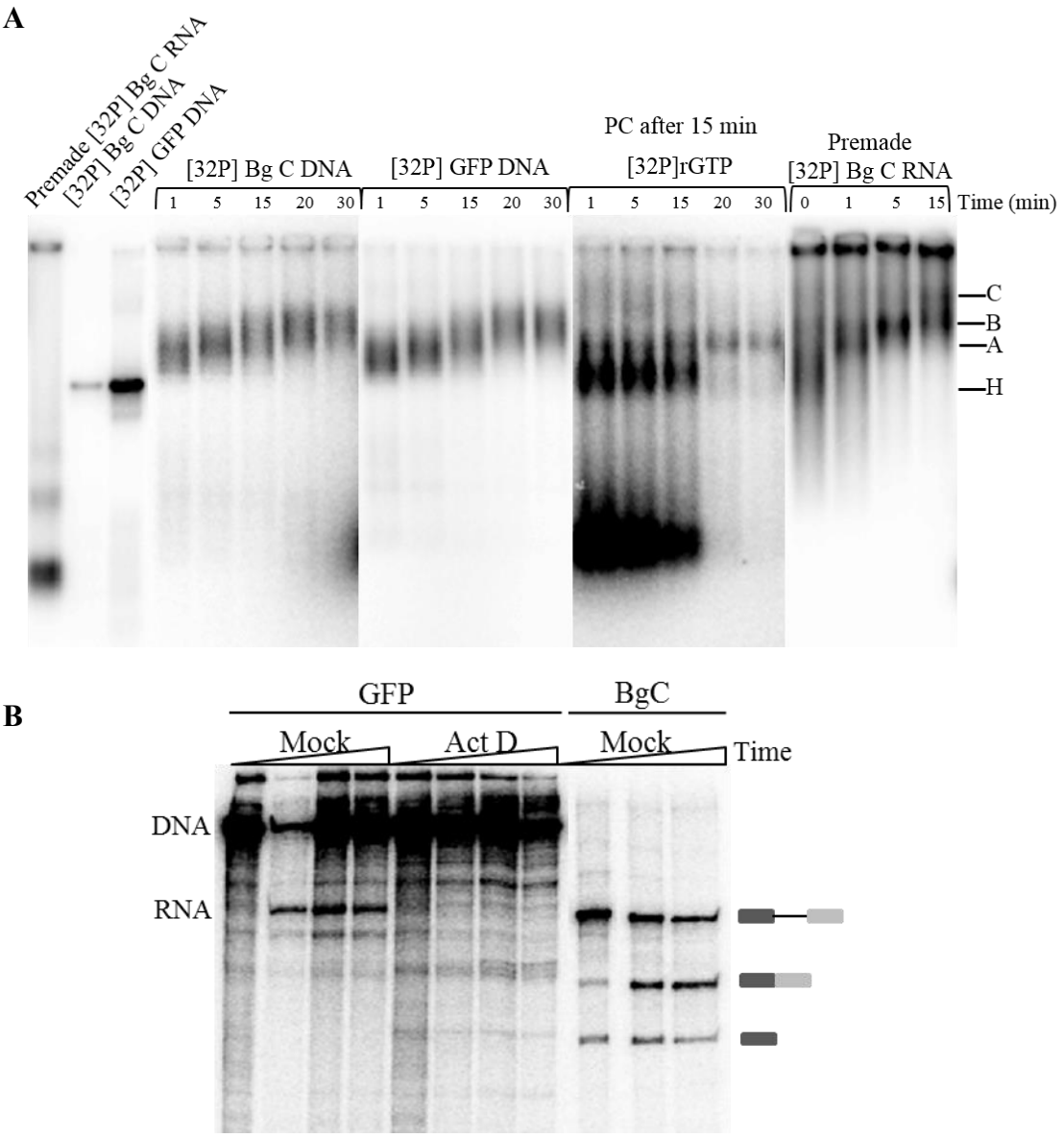


Figure 3.26. Characterisation of DNA-protein complexes of genes with and without an intron. (A) Effect of splicing on the DNA-protein complexes in Pol II-TSR. Radiolabeled CMV-Bg C and CMV-GFP DNA templates were added to the samples which are indicated as [32P] BgC DNA and [32P] GFP DNA, respectively. Unlabeled CMV-Bg C DNA template and [α-^{32}P]GTP nucleotide were added to the sample ([32P]rGTP) to visualize splicing complexes. PC stands for pulse chase. SR was set up side-by-side and it is indicated as pre-made Bg C RNA. Naked DNA and RNA fragments were loaded on the gel. Splicing complexes are not fully resolved in this experiment. After pulse chase two bands (capping complexes) are lost. (B) Inhibition of Pol II transcription by Actinomycin D (Act D) using radiolabelled CMV-GFP DNA. DMSO (1%) was added to the mock samples. A control sample of Pol II-TSR contained unlabeled CMV-Bg C DNA template.

In T7-TSR protein-DNA complexes had the appearance of a ladder too, but the distance between bands was greater than that in Pol II-TSR (Figure 3.27). This can be explained that the length of T7-Bg C DNA is shorter than CMV-Bg C DNA because of the absence of the CMV promoter. Therefore, CMV-Bg C DNA migrated slower and there was smaller separation of the bands in Pol II-TSR.

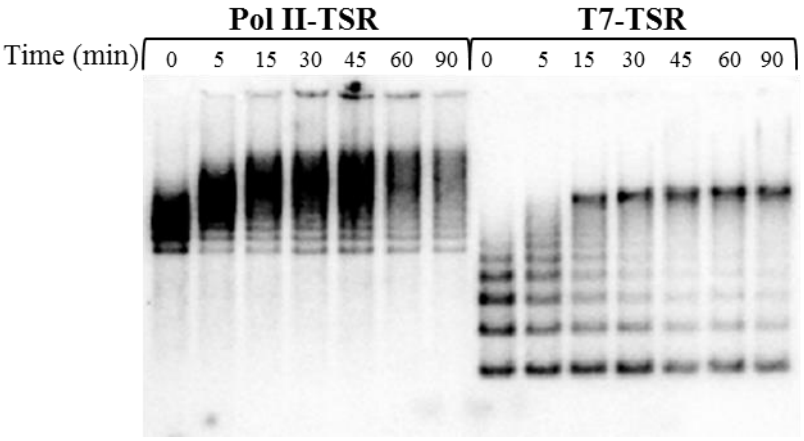
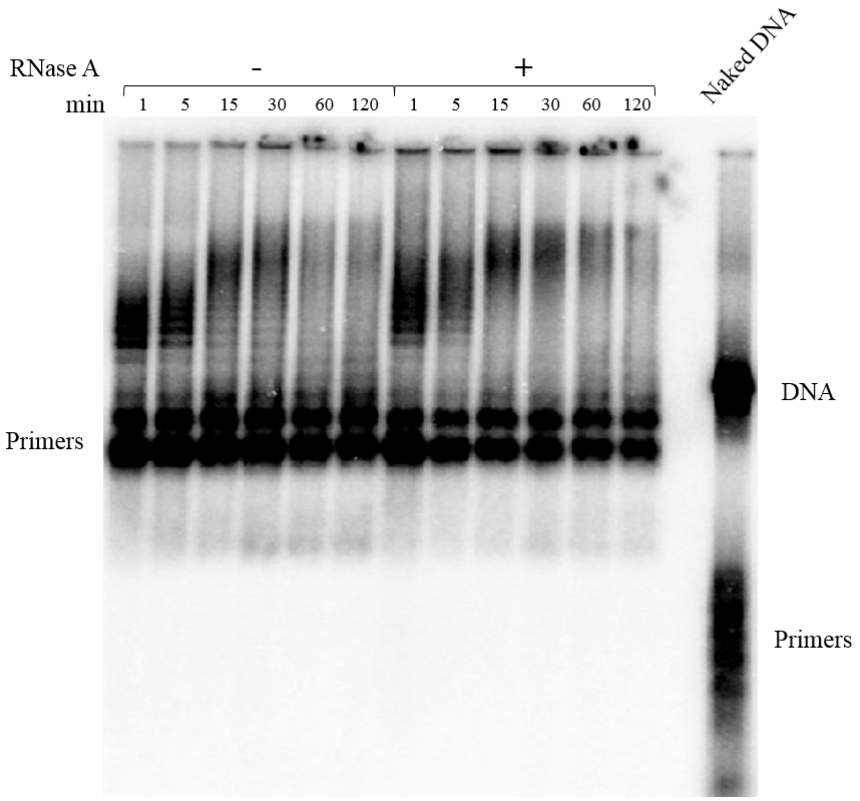


Figure 3.27. Characterisation of DNA-protein complexes formed in Pol II and T7 transcription/splicing assays. Protein-DNA complexes were fractionated on 2% low-melting-point agarose gel. PCR radiolabeled CMV- and T7-Bg C DNA templates were used.

The effect of RNA production on the DNA-protein complexes was tested by using RNase A to degrade transcripts (Figure 3.28). After incubation of Pol II-TSR, RNase A at a concentration of 1 µg/µl was added to each time point and left for 10 min at 37 ºC. Time points were also taken for PAGE, to prove the absence of RNA upon RNase A treatment (Figure 3.28 B). Although RNA was degraded by RNase A, complexes were not sensitive to RNase A treatment. Since RNA degradation did not lead to DNA-protein shift, it is suggested that the spliceosome was assembled on the released pre-mRNA. This result in conjunction with our previous result (Figure 3.23 A) suggest a model that the pre-mRNA undergoes splicing after transcription is completed and the pre-mRNA is released from Pol II.

A



B

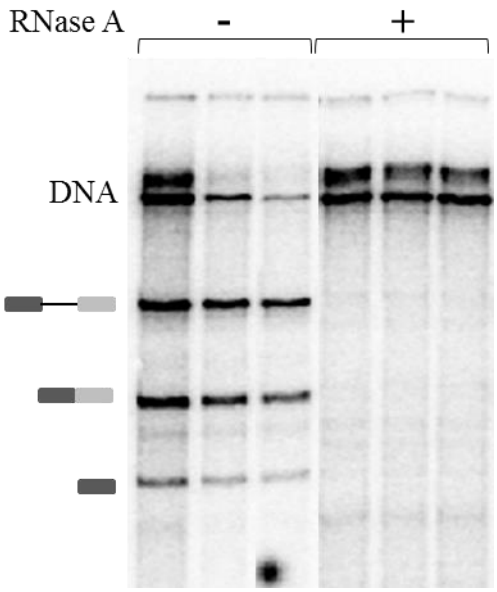


Figure 3.28. Effect of RNase A on DNA-protein complexes in Pol II-TSR. (A) DNA-protein complexes were fractionated on 2% low melting point agarose gel. In this experiment PCR radiolabeling was not done but PCR product was radiolabeled by 5' end labelling using [γ-^{32}P] ATP. Therefore primers were also labelled. Naked DNA template was loaded on the gel. (B) 6% PAGE shows an effect of RNase A in Pol II-TSR.

It is known that during transcription elongation by Pol II different transcription proteins are recruited to the polymerase. We wanted to test if transcription elongation affects the appearance of DNA-protein complexes in our study. An effect of DRB on DNA-protein complexes was tested (Figure 3.29). Since DRB inhibits transcription elongation and we know how complexes appear when transcription does not occur (Figure 3.25 A), we expected to observe different electrophoretic mobilities of complexes upon DRB treatment. DRB at a concentration of 50 μM was added to Pol II-TSR prior incubation. The result showed that the appearance of complexes was not changed upon DRB treatment (Figure 3.29). It may be that transcription elongation was not inhibited completely as complete transcription inhibition was not achieved and some transcripts were still produced (Figure 3.3). Thus, transcription complexes were not affected.

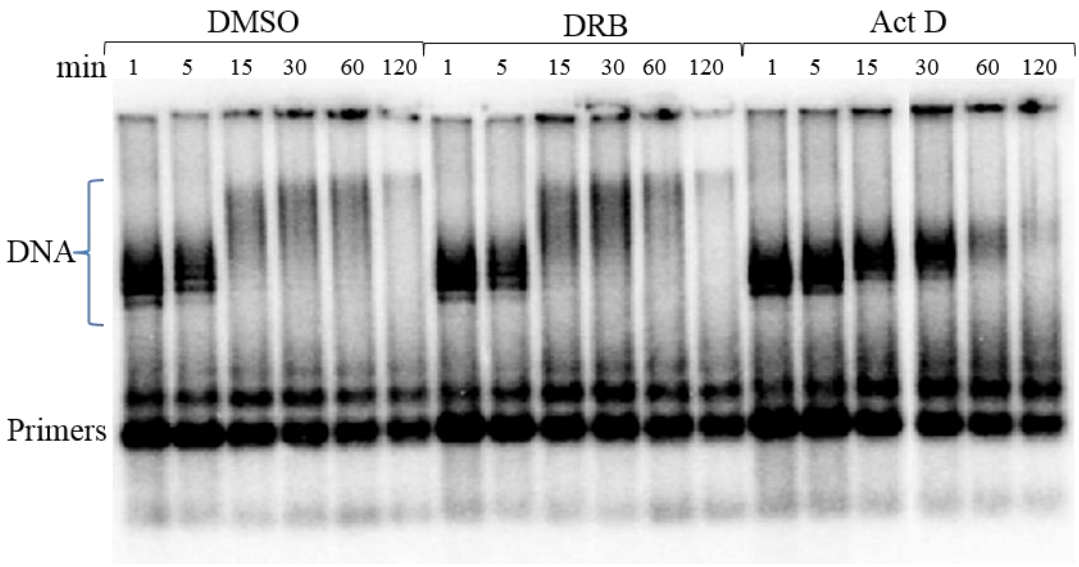


Figure 3.29. Effect of DRB on DNA-protein complexes in Pol II-TSR. DRB and Actinomycin D were used at 50 μM and 5 μM, respectively. DNA-protein complexes were fractionated on 2% low-melting-point agarose gel. PCR radiolabeled CMV-Bg C DNA template was used.

Based on the appearance of the protein-DNA complexes we wanted to test whether histones can interact with DNA forming complexes. A common experimental approach to map nucleosome positions involves the use of micrococcal nuclease (MNase). MNase digests DNA in regions which are not bound by proteins; these are linkers in DNA connecting two nucleosomes. As a result 147 bp DNA fragments are produced.

We hypothesised that if nucleosomes are formed under transcription/splicing conditions then short DNA fragments would be detected in samples treated with MNase. Pol II-TSR having radiolabeled CMV-Bg C DNA was incubated for 15 min at 30 ^{o}C. This was followed by the addition of $CaCl_2$ (1 mM) and MNase and incubation for 10 min. To inactivate MNase, EGTA at a concentration at 2 mM was added to the samples. After PK treatment, the samples were loaded on polyacrylamide gel for electrophoresis. The result showed that DNA was digested but bands of the expected size were not detected, suggesting that nucleosomes are not formed (Figure 3.30). In conclusion, as the nature of protein-DNA complexes was not determined by approaches described above, the potential effects of splicing factors in transcription regulation could not be explored by this method. However, an interesting question can be addressed in future studies such as whether splicing regulators interact with the transcription machinery under Pol II-TSR conditions. For this, the CMV-Bg C DNA template can be biotinylated, incubated under transcription/splicing conditions, pulled down and the sample can be sent for mass spectrometry analysis.

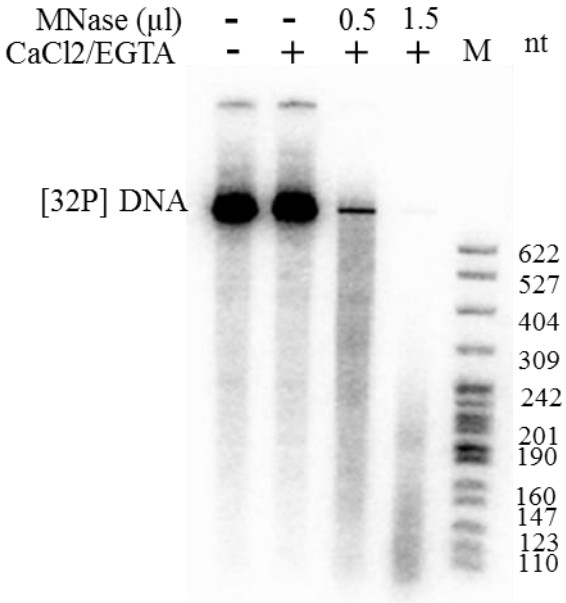


Figure 3.30. Effect of micrococcal nuclease (MNase) on CMV-Bg C DNA incubated under optimizes Pol II transcription/splicing assay. The final concentration of $CaCl_2$ and EGTA was 1 mM and 2 mM, respectively. The samples were electrophoresed through 6% PAGE. The position of radiolabelled CMV-Bg C DNA ([32P]DNA) is shown at the left of the image.

3.7 Conclusions

To establish an efficient Pol II transcription/splicing assay the Bg C DNA fused to the CMV promoter was used for optimization of conditions which can support transcription and splicing in HeLa NE. We proved that transcription was carried out by Pol II. RNA species synthesized by Pol II were assigned based on comigration with a known marker. The same splice sites were recognized as in SR assay. Two optimized conditions regarding the usage of carrier DNA were found. It was determined that the presence of carrier DNA increases transcript level and improves better spliceosome complexes detection. The usage of ATP and one type of radiolabelled nucleotides improves detection of RNA.

Different mini-genes were tested. It was found that our *in vitro* Pol II-TSR system is capable of supporting efficient transcription of all tested constructs. However, the length of a construct has an impact on splicing outcome. For example, 50% of produced pre-mRNA of 930 nt (CD44v5) are spliced, but no splicing products are detected if a precursor is 1.2 kb (Nrxn2), although the splice sites are recognized by the spliceosome. In conclusion, splicing of a long construct can be initiated but is not completed. Pol II-TSR assay was used to test the properties of Pol II transcription to affect splicing efficiency and kinetics in comparison to the conventional SR assay. It was noticed that splicing efficiency was not improved using either Bg C or CD44v5. The rate of intron removal was not enhances by Pol II.

Initially, in this study splicing efficiency in T7-TSR was low (3%) consistent with previous observations (Ghosh and Garcia-Blanco, 2000; Das et al., 2006; Hicks et al., 2006). Pulse chase experiment showed that RNA produced by T7 Pol was not committed to splicing but degraded. A reason behind poor splicing was found. Due to the abundant transcripts generated by T7 Pol, splicing factors were titrated out leading to inability of the spliceosome to fully progress through the spliceosome cycle. Based on electrophoretic mobility of complexes formed in T7-TSR and SR, it is concluded that T7 Pol did not make any qualitative effects on complex A and B. After optimization of T7-Bg C DNA concentration in T7-TSR assay, the conversion of the pre-mRNA to the spliced products occurred with similar efficiency in Pol II-TSR assay.

Therefore, it is concluded that transcription by Pol II does not increase splicing efficiency. There was no complex H detected in either Pol II-TSR or T7-TSR.

The established assays allowed to explore potential properties of Pol II transcription machinery to influence alternative splicing. The CD44v5 construct consisting of three exons and two introns was used. It was found that the splicing profile of CD44 v5 was different between Pol II-TSR and SR, showing that the distal 3'SS is much more preferred over the proximal 3'SS in the reaction with the ongoing transcription. This led to exon 5 skipping in TSR. However, this system recapitulated splicing pattern obtained from *in vivo* study. It is concluded that this change in AS choice is transcription-dependent. Bcl-X substrate, harbouring two alternative 5'SSs, was tested showing similar splicing pattern between Pol II-TSR and SR. In particular, the Bcl-X precursor was able to produce two possible mRNA isoforms. Since the ratio between short and long mRNA isoforms was similar not only between the two *in vitro* systems but *in vivo*, it is suggested that Bcl-X splicing occurs in a transcription-independent fashion.

We asked whether Pol II transcription is coupled to splicing in the *in vitro* assay and whether functional coupling can be studied *in vitro*. DRB was used to inhibit phosphorylation of the CTD in Pol II-TSR. This was compared with two uncoupled systems such as SR and T7-TSR. A small decrease (14%) in splicing efficiency was detected in Pol II-TSR assay while splicing was not affected in either SR or T7-TSR. This data fits previous model (Bird et al., 2004) which suggested that splicing efficiency depends on CTD phosphorylation when splicing occurs co-transcriptionally but not post-transcriptionally. Although DRB treatment made a small effect on splicing efficiency in Pol II-TSR, it is suggest that transcription is linked to splicing.

The property of Pol II to influence spliceosome assembly was tested. It was determined that the rate of spliceosome assembly was significantly slower in comparison to that for SR. Based on the observations that transcripts were visualized as early at 3 min and splicing complexes were detected at 15 min, it is suggested that splicing occurs post-transcriptionally. By doing pulse chase at any time after 5 min after the start of incubation (Pol II-TSR), a single complex was detected and it was shown that it was not complex A. It was detected that the single complex comigrated with complex B in SR.

At this stage this complex is not assigned. In the future the nature of this complex can be taken for investigation.

4 U1 SNRNA PROTECTS RNA FROM DEGRADATION

4.1 Introduction

4.2 Transcript level depends on U1 but not U2 or U6 snRNAs

4.3 Investigation of RNA degradation pathways

4.4 Effects of 5'SS and 3'SS on RNA level

4.5 Complementation assays using U1 snRNP/snRNA

4.6 Conclusions

4.1 Introduction

After the discovery of splicing (Berget et al., 1977; Chow et al., 1977), research was focused on the mechanisms of spliceosome formation and intron removal. When evidence of co-transcriptional splicing was obtained (Beyer et al., 1989), researchers started investigating the potential associations between the core spliceosomal components and Pol II transcription factors. It was found that all five snRNPs interact with human transcription elongation factors (Fong and Zhou, 2001). When the spliceosomal snRNAs were depleted from NE and this NE was used in *in vitro* Pol II transcription, the latter was negatively affected (Fong and Zhou, 2001). Another result provided *in vivo* evidence on decreased transcript levels upon inhibition of the U1, U2, U4 and U6 snRNP by antisense morpholino oligonucleotides (Koga et al., 2015). The association between the U1 snRNP and the transcription apparatus in the absence of transcription was shown (Das et al., 2007). The U1 snRNA was found to interact with Pol II (Kim et al., 1997; Tian, 2001) and general transcription factors such as TFIIH (Kwek et al., 2002) and TAF15 (Jobert et al., 2009). The disruption of the essential base-pairing between the U1 snRNA and the 5'SS of pre-mRNA led to Pol II transcription inhibition *in vitro* (Tian, 2001; Yu and Reed, 2015). When the U1 snRNA was blocked *in vivo*, its vital role in the suppression of cryptic poly(A) site usage and premature cleavage was discovered (Kaida et al., 2010). Therefore, these studies suggested that the U1 snRNP plays essential roles in transcription initiation, elongation and termination. However, ChiP data demonstrated unchanged Pol II occupancy on genes upon inhibition of the 5' terminus of the U1 snRNA (Davidson et al., 2014). Another *in vivo* studies showed that the levels of transcripts are unchanged if mutated U1 snRNA interacts with intronic sequences by complementarity (Lu et al., 1990; Dal Mas et al., 2015). Thus, it is not clear if the U1 snRNP participates in transcription regulation or it is required for RNA stability. To investigate the properties of the core spliceosome snRNPs to regulate Pol II transcription or play roles in other RNA processing events, an assay with simultaneous transcription and splicing was considered ideal (Folco and Reed, 2014). The effects of U1, U2 and U6 snRNPs on Pol II transcription were studied using the *in vitro* Pol II transcription/splicing system characterized in Chapter 3.

4.2 Transcript level depends on U1 but not U2 or U6 snRNAs

4.2.1 Effects of U1, U2 and U6 snRNAs in Pol II-TSR

To investigate the effect of individual snRNPs on Pol II transcription in an *in vitro* transcription/splicing assay, the activity of the splicing machinery was targeted by 2'-*O*-methyl oligonucleotides which interact with regions of U1, U2 and U6 snRNAs that are important at different stages of splicing. The effects of these oligonucleotides on the spliceosome assembly were tested and the results are shown in Chapter 3. The NEs were treated with oligonucleotides prior to the addition of CMV-Bg C DNA. As predicted, inhibition of functional U1, U2 and U6 snRNAs resulted in loss of splicing activity as splicing products were not detected (Figure 4.1). Interestingly, the U1 snRNP functional reduction had an additional and marked effect on Pol II transcription showing a five-fold decrease in the amount of transcripts. A slight inhibition of transcription was detected in the sample with blocked U6 snRNA and non with anti-U2. These results are in line with previous *in vitro* studies (Tian, 2001; Kwek et al., 2002) which showed that transcription by Pol II was abolished by RNA oligonucleotides harbouring sequence of the 5'SS. Also these data fits previous results obtained from *in vivo* studies (Kaida et al., 2010; Koga et al., 2014) showing that the level of transcripts drops down if the U1 snRNP has been blocked.

2'-*O*-methyl oligonucleotides are used to block certain regions of snRNAs indicating that a target RNA region is not destroyed but blocked. To confirm that permanent inhibition of snRNPs has the same effect on splicing and Pol II transcription, I wanted to degrade the same regions of the U1, U2 and U6 snRNAs which were blocked in the previous experiment. For this, DNA-directed RNA cleavage by RNase H was done. This method is used to degrade the RNA strand of an RNA:DNA duplex (Donis-Keller, 1979). Interference between DNA oligonucleotides and target sequences was not tested in this study, but oligonucleotides were of the same sequence as DNA oligonucleotide used for RNase H-mediated degradation of the U1, U2 and U6 snRNPs in previous studies (Kramer et al., 1984; Krainer and Maniatis, 1985; Black et al., 1985) proving an interaction between the probes and targets. The procedure to deplete snRNAs was the same as for 2'-*O*-methyl oligonucleotides but the reaction mixture was supplemented

with RNase H (0.2 U; Materials and Methods). The 5' terminus (1-14 nt) of the U1 snRNA was degraded using DNA oligonucleotide (called 1-14; Table 2.3) at 1 μM and 5 μM. As shown in Figure 4.2, splicing was significantly reduced at 5 μM. The decrease in transcription indicates the consistency in the transcription outcome upon the functional inhibition of the U1 snRNA by two types of oligonucleotides (Figure 4.1 and 4.2).

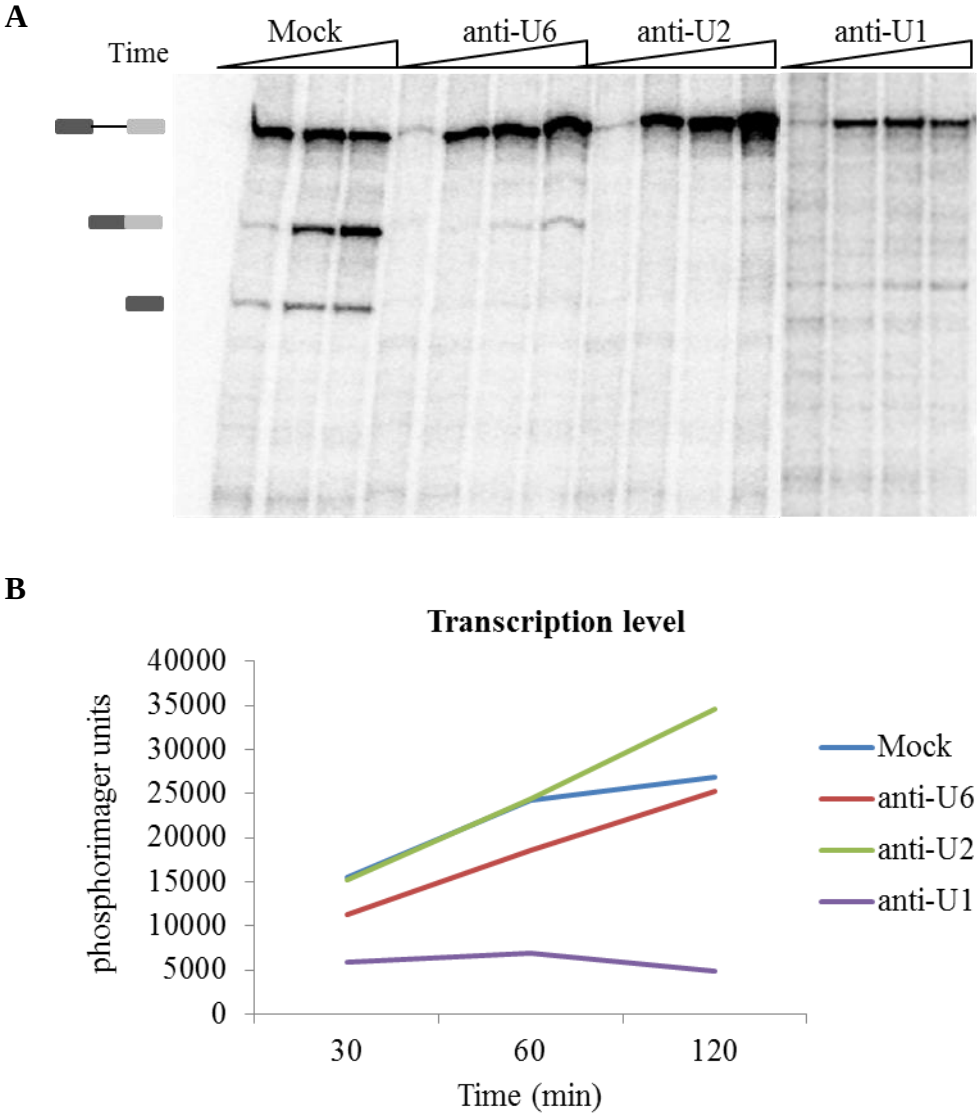


Figure 4.1. Pol II transcription of Bg C is decreased upon inhibition of U1 but not U2 or U6 snRNAs. (A) Inhibition of *in vitro* Pol II transcription/splicing reaction by anti-U1, U2 and U6 snRNAs 2'-*O*-methyl oligonucleotides. The mock-treated reaction contained water. The reaction mixtures with oligonucleotides were pre-incubated at 30°C for 15 min before the addition of CMV-Bg C DNA template. Time points of 0, 30, 60 and 120 min were taken. (B) Transcription level (sum of the pre-mRNA, mRNA and 5' exon) from (A) was plotted versus time.

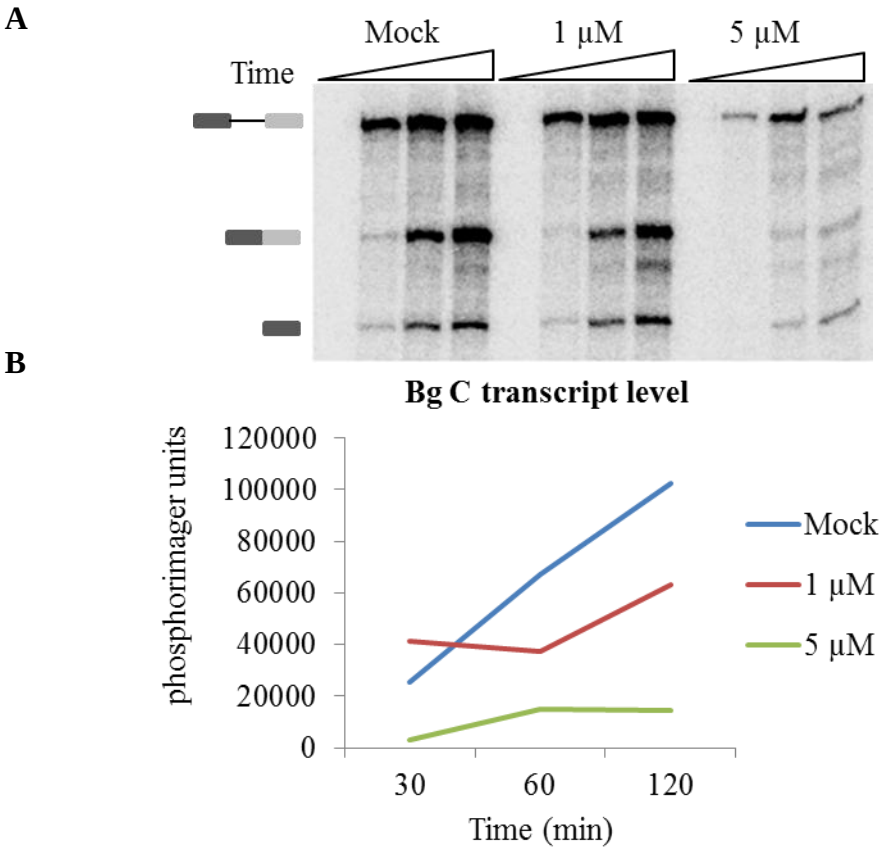


Figure 4.2. Transcription is decreased upon RNase H-mediated degradation of the 5' terminus of the U1 snRNA. (A) Effect of targeted depletion of the 5' terminus (1-14 nt) of the U1 snRNA. The 5' end sequence of the U1 snRNA was degraded by RNase H cleavage directed by a DNA oligonucleotide. The reaction mixture including DNA oligonucleotide and RNase H (0.2 U) was pre-incubated at 30°C for 15 min prior to the addition of CMV-Bg C DNA. As oligonucleotides were diluted in water, the mock-treated reaction contained water too. The 5' end oligonucleotide was added in concentrations denoted within the figure. Time points of 0, 30, 60 and 120 min were taken. (B) Transcription level from (A) was plotted versus time.

RNase H-mediated degradation of the U2 snRNA region (28-42 nt) that interacts with the branch site was done using antisense DNA oligonucleotide at 5 μM (L15; Table 2.3) (Black et al., 1985). The U6 snRNA region for base-pairing with the U2 snRNA was cleaved by RNase H using DNA oligonucleotide (MB; Table 2.3) at 5 and 10 μM. The treatments blocked splicing, indicating functionality of all oligonucleotides (Figure 4.3). Surprisingly, transcription was slightly increased upon digestion of the U2 snRNA (Figure 4.3 A). There were either different transcription or degradation products produces in the sample with MB (Figure 4.3 C). The relative number of molecules in each band was not calculated by normalising the intensity according to the number of GMP in the corresponding RNA molecules because the nature of the bands was not determined. Therefore, transcription levels were not measured. However, it is clear that the level of transcription was increased upon degradation of the U6 snRNA. This is different from the results obtained from experiments using two types of antisense oligonucleotides against the U2 and U6 snRNAs. Since the experiment was not done in triplicate, it is difficult to assess the effects of the U2 and U6 snRNPs on transcription. However, the effect on transcription by inhibiting the U1 snRNP was reproducible and in line with previous observations (Tian, 2001; Kwek et al., 2002).

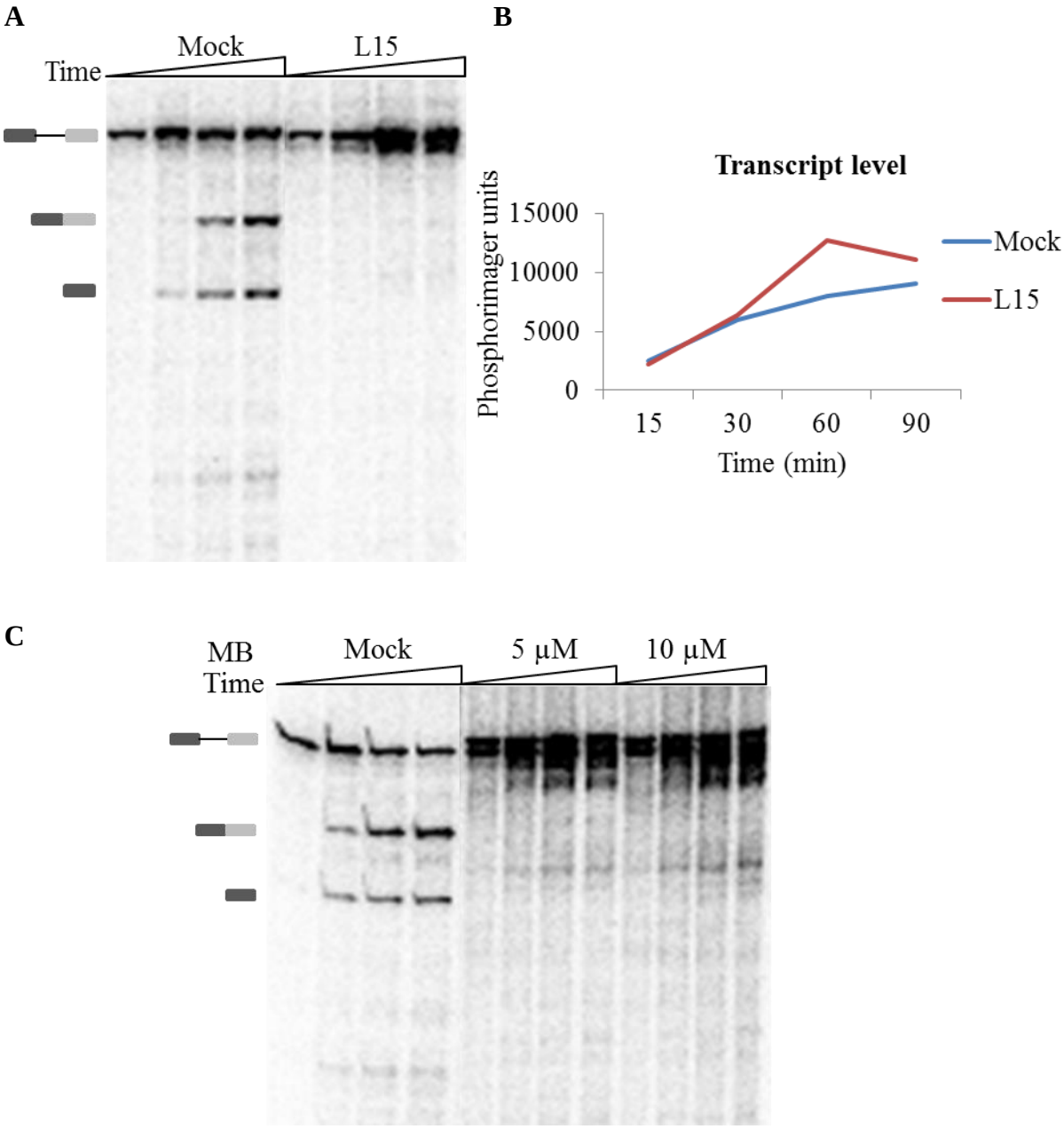


Figure 4.3. Transcription is increased upon RNase H-mediated degradation of the U2 and U6 snRNAs. Effects of targeted depletion of U2 snRNA (A) and U6 snRNA (C) using DNA antisense oligonucleotides L15 and MB, respectively. (B) Transcription level from (A) was plotted versus time. The concentration of oligonucleotides was 5 µM. The reaction mixtures including DNA oligonucleotides and RNase H (0.2 U) were pre-incubated at 30°C for 15 min prior to the addition of the CMV-Bg C DNA. Time points at 0, 30, 60 and 120 min were taken. (C) Due to the presence of many transcripts, transcription level was not measured.

4.2.2 The anti-U1 snRNA oligonucleotide effect is not substrate specific

2'-*O*-methyl oligonucleotides to block U1, U2 and U6 snRNAs were added to Pol II-TSR supplemented with the CMV-CD44v5 DNA. The experiment showed that transcription was sensitive to all oligonucleotides (Figure 4.4). Regarding the U1 snRNA effect on transcription, this result is consistent with the data obtained from the experiment with CMV-Bg C DNA. However, transcription outcome upon inhibition of the U2 and U6 snRNAs using Bg and CD44v5 constructs was different (Figure 4.1). To validate a contribution of the U2 and U6 snRNAs in Pol II transcription, the experiment should be done in triplicate.

Since the CD44v5 construct has sequence similarities with Bg C, as CD44v5 was cloned into the intron of Bg C (Appendix), substrate with different origin was also test. A three exon-two intron construct derived from RON (Smith et al., 2013) was cloned into pEGFP-C1 and fused to the CMV promoter (Materials and Methods; Figure 4.5 A). It was previously shown that RON does not undergo splicing *in vitro* (Smith et al., 2013), so, we wanted also to test whether Pol II transcription could stimulate splicing of RON pre-mRNA. A time-course of Pol II-TSR showed that RON pre-mRNA did not undergo splicing in the mock-treated reaction (Figure 4.5 B), suggesting that splicing was not stimulated by Pol II transcription. There was a two-fold decrease in the level of transcript upon U1 snRNA inhibition suggesting that the efficiency of transcription was dependent on the U1 snRNA. The results obtained from the experiments with CD44v5 and RON substrates were consistent with the experiments in which Bg C was used. In conclusion, transcription of DNA constructs with different sequence compositions is dependent on the functional activity of the U1 snRNA.

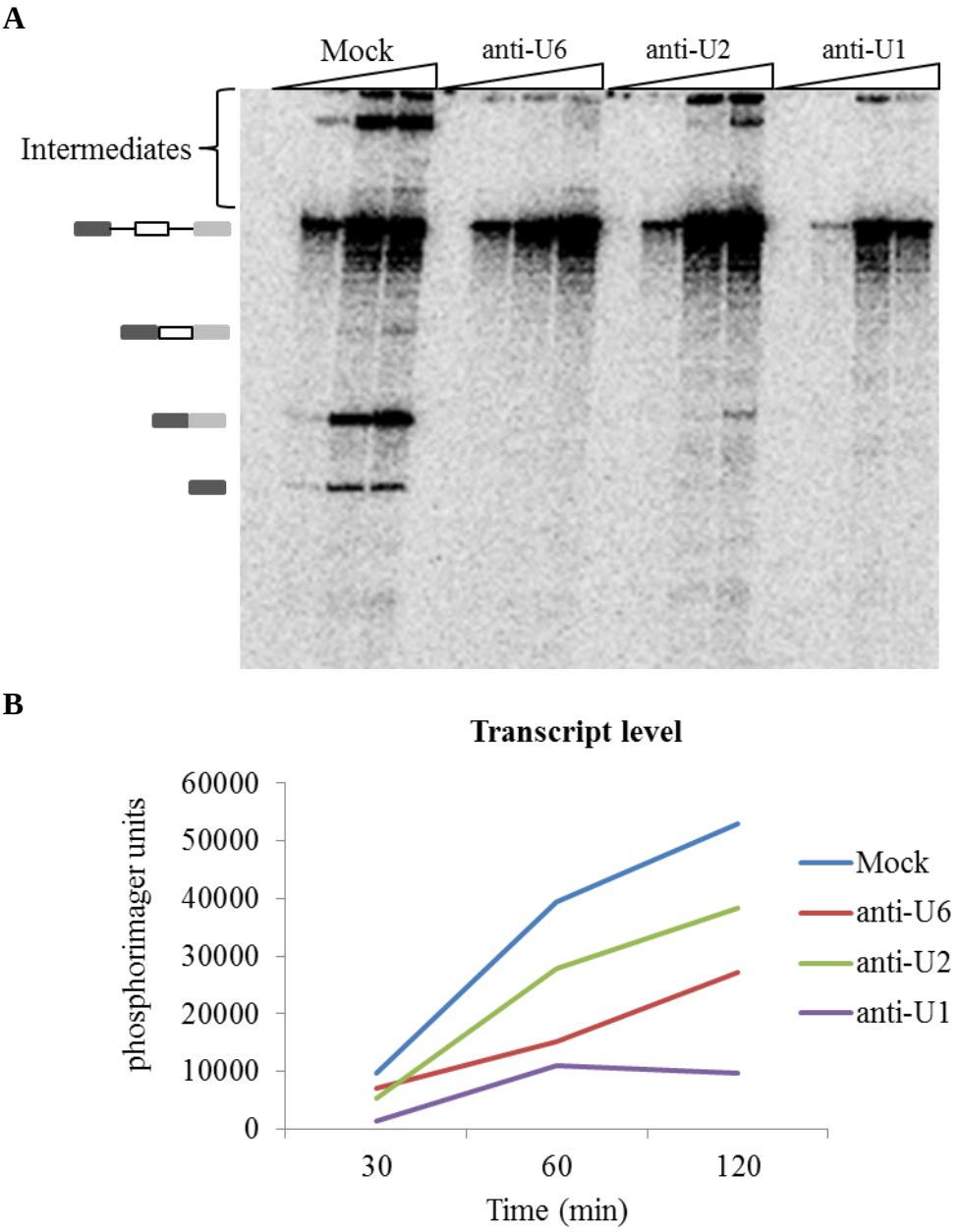


Figure 4.4. The level of Pol II transcription of CD44v5 substrate is decreased upon inhibition of the U1, U2 and U6 snRNAs. (A) Inhibition of the U1, U2 and U6 snRNAs by 2'-*O*-methyl oligonucleotides using CMV-CD44v5 DNA. The pre-mRNA, RNA splicing products (inclusion and exclusion mRNAs) and intermediates are diagrammed at the right of the phosphorimage. Time points at 0, 30, 60 and 120 min were taken. (B) Transcription level from (A) was plotted versus time.

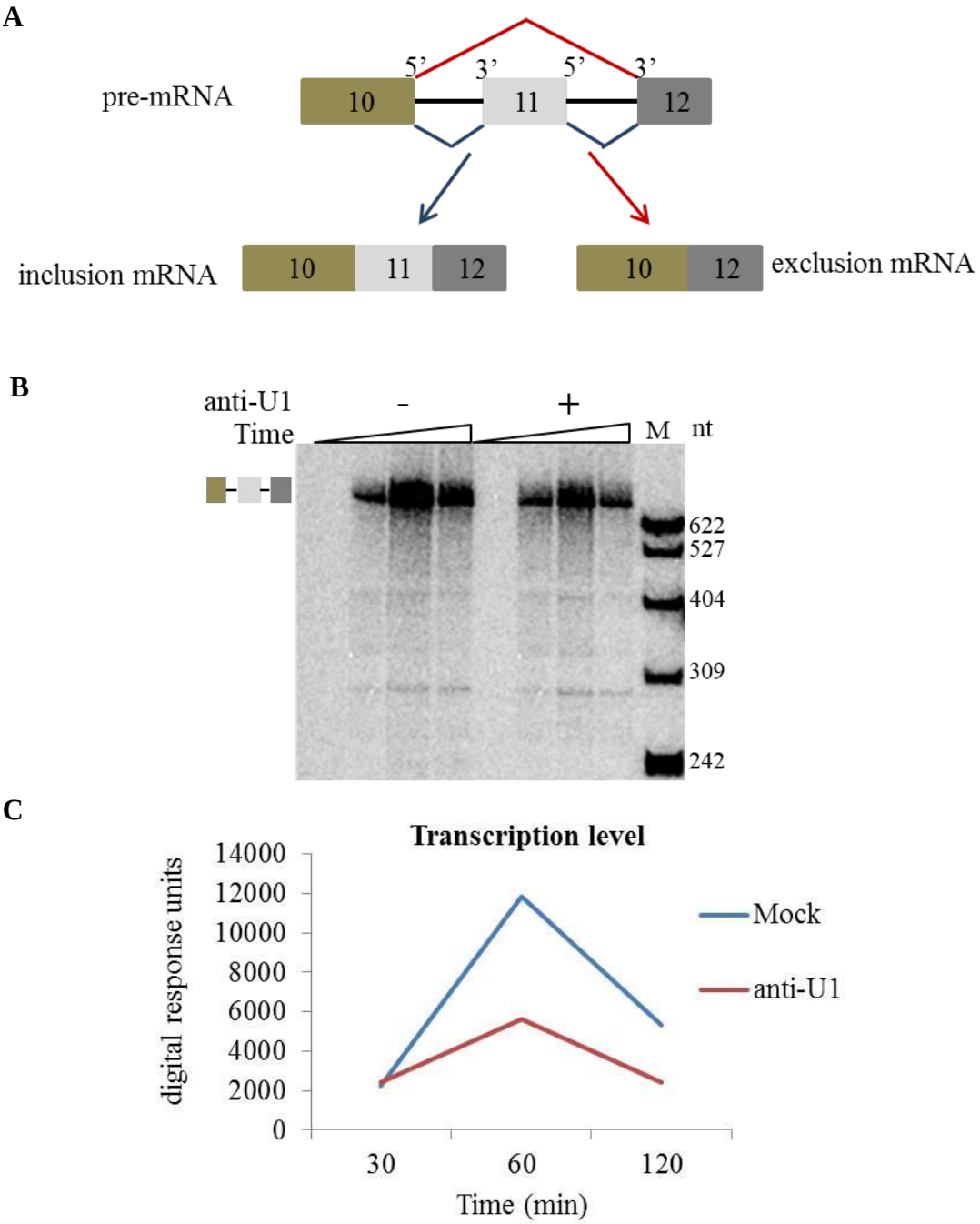


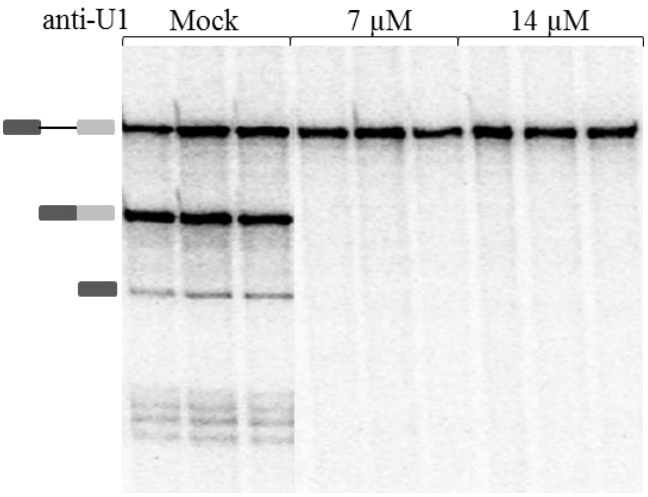
Figure 4.5. The level of Pol II transcription of RON substrate is decreased upon inhibition of the U1 snRNA. (A) Diagram shows alternative splicing pattern of RON pre-mRNA. Exons 10-12 (boxes), introns (black lines), splice sites and splicing alternatives (red and blue lines) are represented. (B) Inhibition of U1 snRNA by 2'-*O*-methyl oligonucleotide using CMV-RON DNA. The pre-mRNA is diagrammed at the right of the phosphorimage. Time points at 0, 30, 60 and 120 min were taken. (C) Transcription level from (A) was plotted versus time.

4.2.3 Inhibition of U1, U2 and U6 snRNAs in T7-TSR

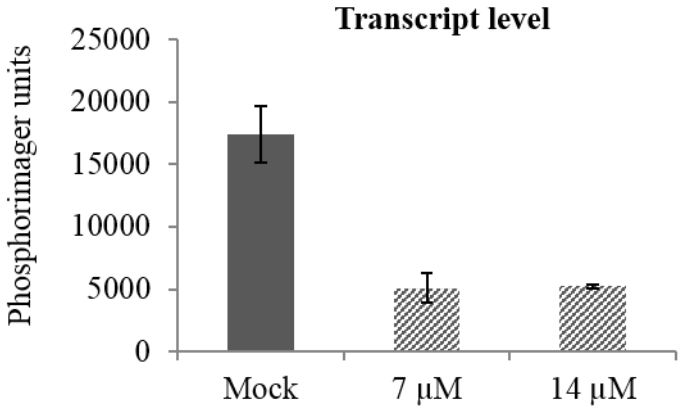
The studies that showed that the U1 snRNA associates with Pol II (Kim et al., 1997; Tian, 2001) or general transcription factors such as TFIIH (Kwek et al., 2002) and TAF15 (Jobert et al., 2009) suggested that the U1 snRNA plays roles in transcription initiation and elongation. Previous work had shown that this effect was not observed for transcription by RNA Pol III, and it was inferred that it was restricted to RNA Pol II (Tian, 2001). To test this, the U1 snRNA was targeted in T7-TSR. Anti-U1 2'-*O*-methyl oligonucleotides at 7 and 14 μM was used. The procedure to inactivate the U1 snRNA in T7-TSR was the same as before. The experiment was done in triplicate and incubation at 30^{o}C was carried out for 90 min. The yield of RNA was decreased by 3.5-fold (Figure 4.6 A and B). The oligonucleotide at the highest concentration had no additional effect on transcript level, suggesting that 7 μM oligonucleotide was sufficient to influence both splicing and transcription. It is concluded that the U1 snRNA effect on the level of transcription is polymerase-independent. One possibility is that the binding of the U1 snRNPs confers stability of transcripts.

Inhibition of the U2 and U6 snRNPs by 2'-*O*-methyl oligonucleotides had a converse effect: the level of RNA was increased by 3-fold. This result is similar for Pol II-TSR in which the U2 and U6 snRNAs were targeted by RNase H-mediated degradation (Figure 4.3). Although, this increase was not seen with the 2'-O-methy oligos (Figure 4.1). It is suggested that the difference in the level of RNA between the untreated and treated samples is that spliced products may not be stable and could undergo degradation after the completion of splicing.

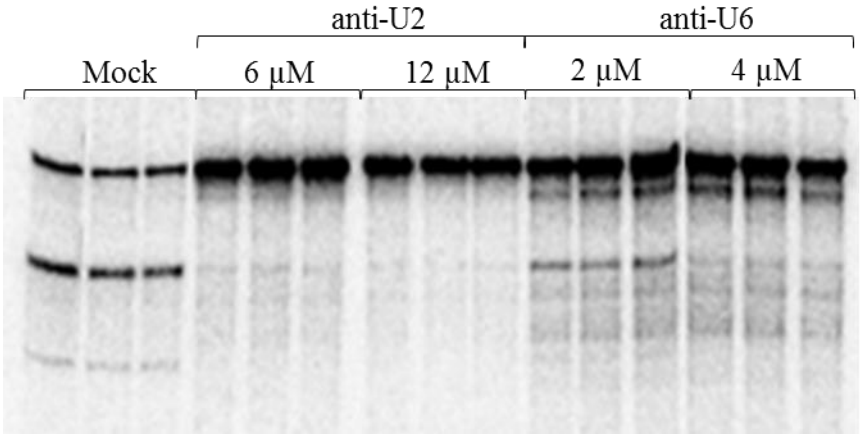
A


anti-U1
Mock
7 µM
14 µM

B


Transcript level
Phosphorimager units
25000
20000
15000
10000
5000
0
Mock
7 µM
14 µM

C


anti-U2
anti-U6
Mock
6 µM
12 µM
2 µM
4 µM

D

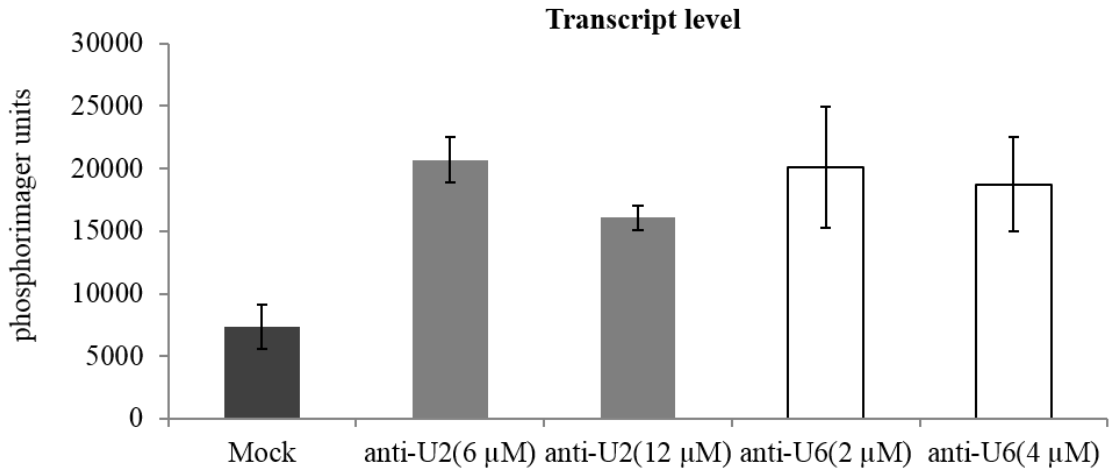


Figure 4.6. Transcript level is decreased upon inhibition of U1 but not U2 or U6 snRNAs in T7 transcription/splicing system. Effect of anti-U1 (A), U2 and U6 snRNAs (B) using 2'-*O*'methyl oligonucleotides. The reaction mixtures with oligonucleotides were pre-incubated at 30°C for 15 min prior to the addition of T7-Bg C DNA. After, the samples were incubated for further 90 min. (B) Transcription level from (A) was plotted in the diagram. (D) Transcription level from (C) was plotted in the diagram. Error values are standard deviation values.

4.2.4 Inhibition of U1, U2 and U6 snRNAs in SR

The above experiments suggest that a T7 Pol-generated transcript in the NE is not stable if the U1 snRNP is not functional. To further evaluate U1 snRNP-dependent effect on RNA stability, we tested the RNA stability of pre-synthesised transcripts under splicing conditions. Since transcription is not functionally coupled to splicing in T7-TSR or SR, we hypothesised that the protective RNA stability effect might be similar in these two assays.

Prior to the addition of the pre-mRNA to SR mixtures, the NEs were treated with anti-U1, anti-U2 and anti-U6 oligonucleotides at 7 µM, 6 µM and 4 µM, respectively (Figure 4.7). To measure the level of RNA after incubation, 0 and 90 min time points were taken. The experiment was done in triplicate, excepting the sample with anti-U6 oligonucleotide.

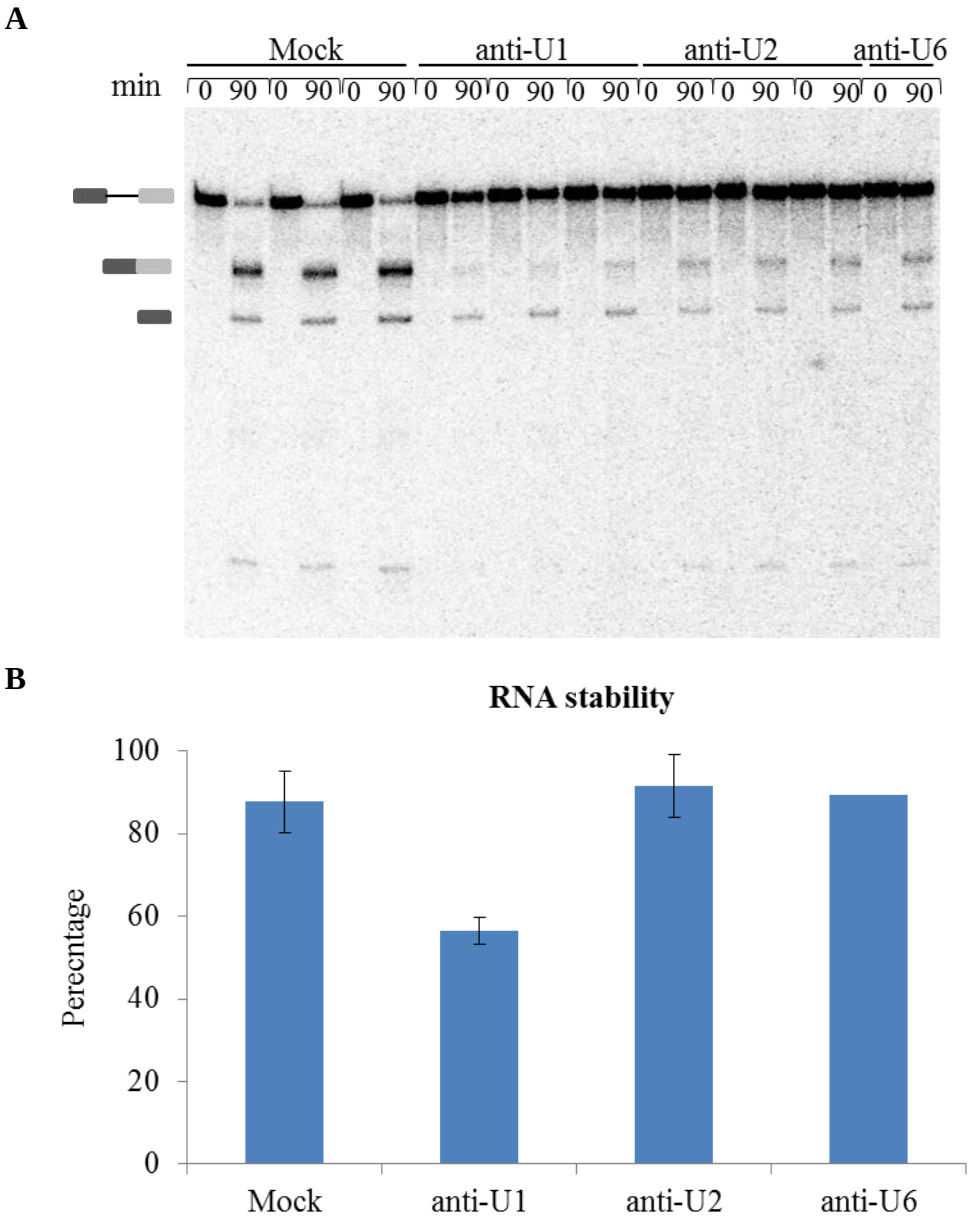


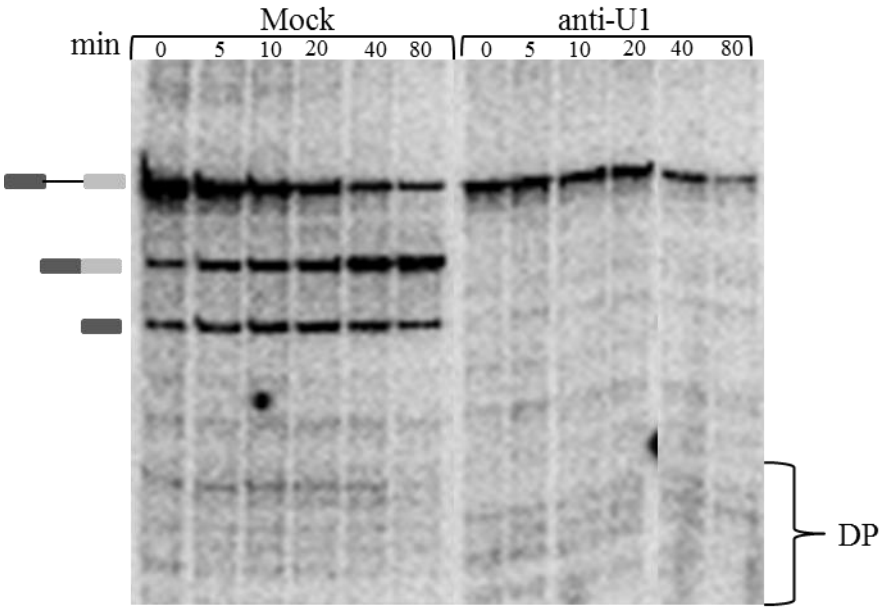
Figure 4.7. Pre-made pre-mRNA is degraded in nuclear extract under conditions with not functional U1 snRNA. In splicing conditions HeLa nuclear extracts were treated with anti-U1, U2 and U6 snRNAs 2'-*O*-methyl oligonucleotides at 30°C for 15 min prior to the addition of 32P-labelled Bg C pre-mRNA. This followed by the incubation for further 90 min. All reactions were done in triplicates, excepting that with anti-U6 snRNA oligonucleotide. The pre-mRNA, mRNA and 5' exon are indicated. (B) The diagram shows transcript levels which were measured as the percentage of RNA at 90 min to the input RNA at 0 min. Error values are standard deviation values.

RNA levels were measured as the percentage of the sum of the pre-mRNA and splicing products at 90 min relatively to the input pre-mRNA at 0 min time point. Complete splicing inhibition was not achieved as splicing products were detected. As expected, the level of RNA was decreased under the conditions with inhibited U1 snRNA, but was unchanged in the samples with either blocked U2 or U6 snRNAs. This suggests that the RNA underwent degradation upon U1 snRNA inhibition. However, no degradation products were detected. In conclusion, among the tested spliceosomal snRNPs, U1 is the only snRNP that clearly affected RNA stability. In conclusion, the U1 snRNP is required for RNA surveillance and this phenomenon is transcription-independent.

4.2.5 RNA stability in Pol II-TSR

The above experiments suggest that RNA stability depends on the activity of the U1 snRNA. To further prove this, we made use of pulse chase assays to measure the degradation profiles for Bg C transcripts produced by Pol II under conditions with inhibited U1 snRNA. The procedure for the U1 snRNA inactivation by anti-U1 oligonucleotide was the same as before. After incubation for 30 min, an excess of unlabelled GTP was added to quench the production of radiolabelled transcripts and this was referred as 0 min time point (Figure 4.8 A). The sum of the pre-mRNA and splicing products at each time point was related to the amount of pre-mRNA at the 0 min time point. The result is shown in Figure 4.8 B. The quantification of the degradation profiles showed a logarithmic decrease in RNA levels in both treated and untreated samples. However, a significant increase in the degradation rate of the pre-mRNA under conditions with blocked U1 snRNA was observed, which can be seen from the comparison of the corresponding slopes. RNA degradation products were detected in the treated sample. Also, the RNA is not fully stable in the mock-treated sample, suggesting that U1 snRNP-independent degradation also takes place in Pol II-TSR. In conclusion, the U1 snRNP has a property to stabilize RNA.

A



B

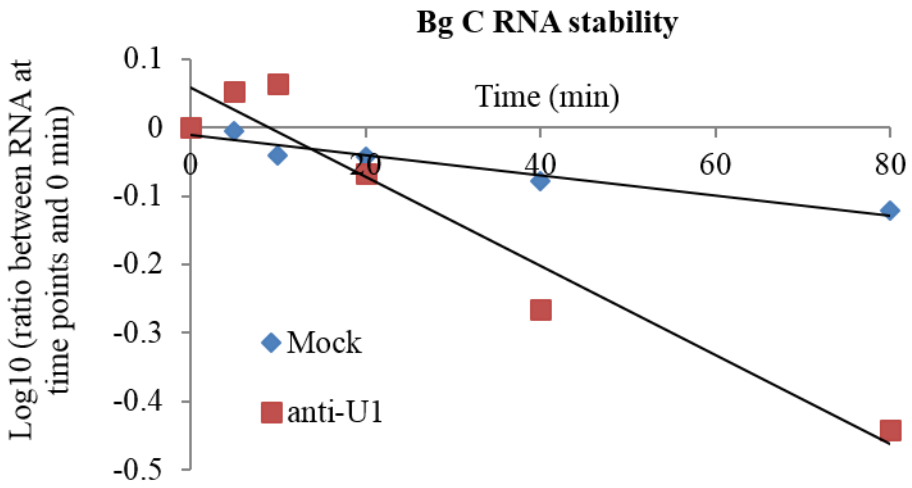


Figure 4.8. Bg C RNA transcribed by Pol II undergoes degradation under conditions with not functional U1 snRNA. (A) Inhibition of the U1 snRNA by 2'-*O*-methyl oligonucleotide using CMV-Bg C DNA in Pol II-TSR. The zero time point refers to the addition of excess unlabelled GTP nucleotides 30 min after the start of incubation. DP stands for degradation products. (B) Quantification of the degradation profiles versus time is shown in the graph.

4.3 Investigation of RNA degradation pathways

The described above experiments revealed that the U1 snRNP mediates a greater stability of transcripts in both uncoupled reactions (Figure 4.6 and Figure 4.7) and a reaction in which transcription is linked to splicing (Figure 4.8). These observations indicate that the U1 snRNP plays a crucial role in protecting transcripts from degradation. To investigate the mechanisms by which the U1 snRNP stabilises RNA and protects it from degradation, it was important to understand how RNA is degraded *in vitro* in the absence of U1 snRNP.

4.3.1 RNA protection in Pol II-TSR

In vitro RNA can be degraded by three major classes of RNA-degrading enzymes: endonucleases, 5' exonucleases and 3' exonucleases. By protecting the ends of transcripts, it would be possible to determine which classes of enzymes are involved in RNA degradation and, therefore, what kind of RNA degradation pathways become activated when RNA is not protected by the U1 snRNP. 2'-*O*-methyl modified oligonucleotides were used to protect RNA from exonuclease cleavage. Oligonucleotides complementary to the ends of Bg C pre-mRNA were designed and used in Pol II transcription/splicing and splicing assays to render the RNA ends to be inaccessible to 5' and 3' exonucleases (Figure 4.9 A). Further, these oligonucleotides will be called 5' E and 3'E. We hypothesised that if the 5' and 3' ends of transcripts become degraded when the U1 snRNP does not interact with RNA then the level of transcripts would be increased in the reactions with the oligonucleotides. First, the U1 snRNA was blocked with anti-U1 oligonucleotide during a pre-incubation step. This was followed by the addition of protective oligonucleotides and the CMV-Bg C DNA template. The concentration of oligonucleotides was titrated from 0.01 μM to 1 μM. Contrary to our expectations, the 5' E and 3'E oligonucleotides did not show an enhancement of transcript levels in Pol II-TSR (Figure 4.9 B). A slight increase of transcript was observed in the reaction with both 3'E (0.1 and 1 μM) and anti-U1 oligos in comparison to the reaction with only anti-U1 oligo (Figure 4.9 C).

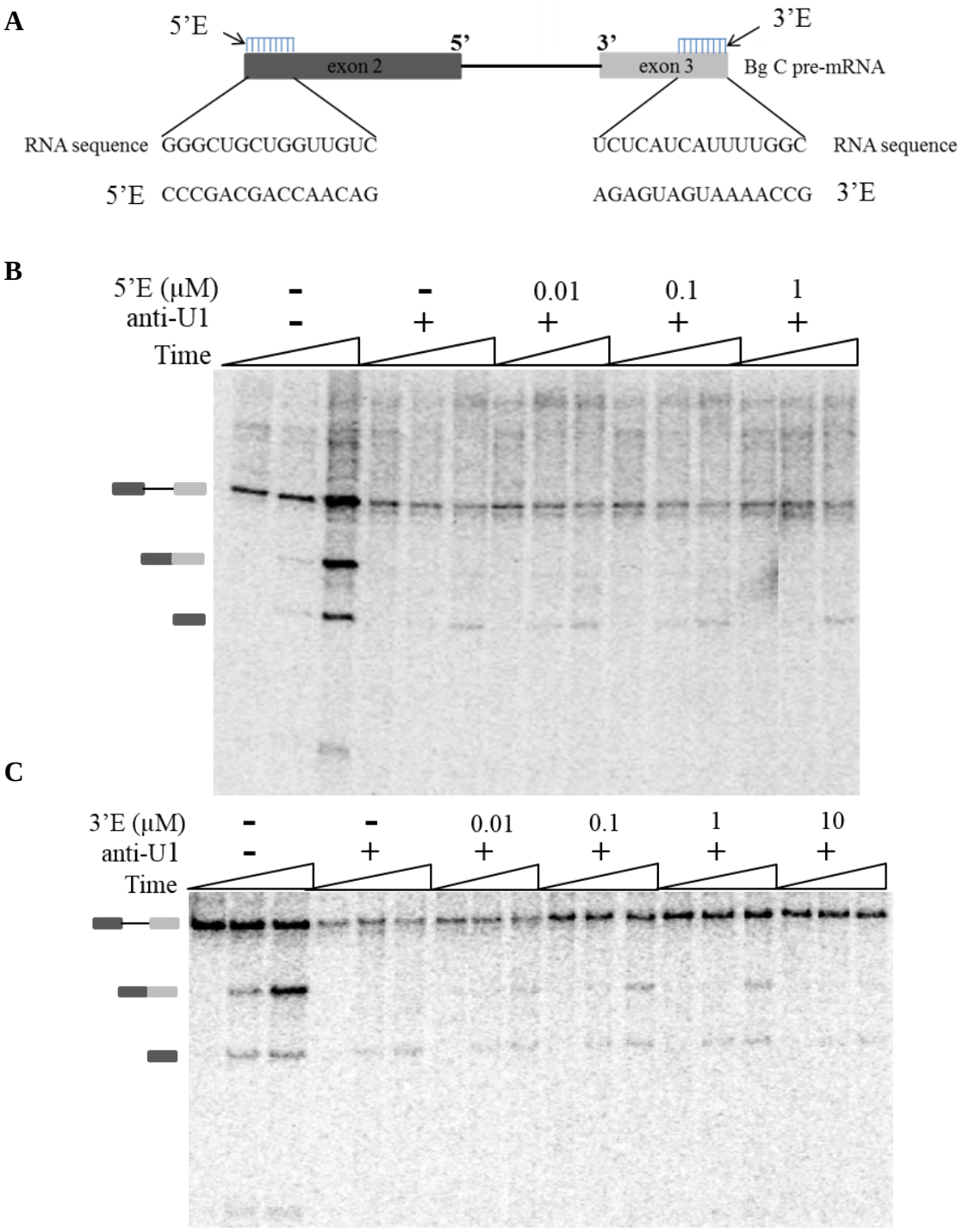
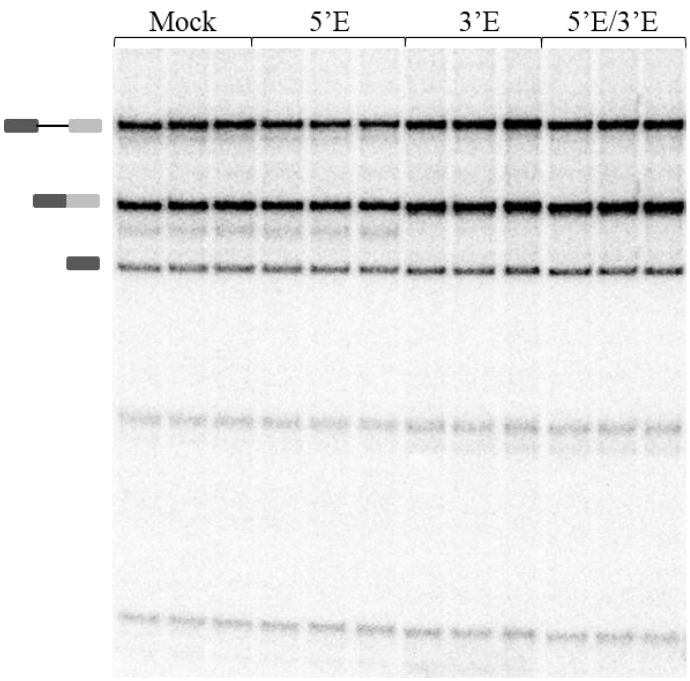


Figure 4.9. Protection of pre-mRNA ends by complementary oligonucleotides in Pol II-TSR. (A) Diagram shows the position of 2'-*O*-methyl oligonucleotides annealing (called 5'E and 3'E). The oligonucleotide sequences are indicated. *In vitro* time-courses of Pol II transcription/splicing with 5'E (B) and 3'E (C) oligonucleotides. CMV-Bg C DNA template was used. In Pol II transcription/splicing conditions HeLa nuclear extracts were treated with and without anti-U1 oligonucleotide (5 μM) at 30 °C for 15 min. After, 5'E, 3'E oligonucleotides and DNA template were added. The concentrations of the protective oligonucleotides are represented. Time points at 30, 60 and 90 min were taken.

However, the level of RNA was significantly less compared to the mock sample. From this experiment, it is not clear if 3' end degradation is specific to the absence of the functional U1 snRNP or it is a general effect. To test this idea, Pol II-TSR was supplemented with 5'E and 3'E oligonucleotides. The U1snRNA was not inhibited in this experiment. The result showed a small increase (1.2-fold) of transcript level in the reaction with 3'E oligo (Figure 4.10). This result indicates that the newly synthesised RNA is not fully protected in the NE and even in the presence of the functional U1 snRNP some pre-mRNA undergoes degradation from the 3' end. Since none of the protective oligonucleotides were effective in the reaction with inhibited U1 snRNA (Figure 4.9), it is suggested that RNA could undergo degradation during transcription and, therefore, the RNA was not released form the polymerase. Alternatively, transcripts could be degraded by endonucleases and there was no substrates for the oligonucleotides to anneal to.

A



B

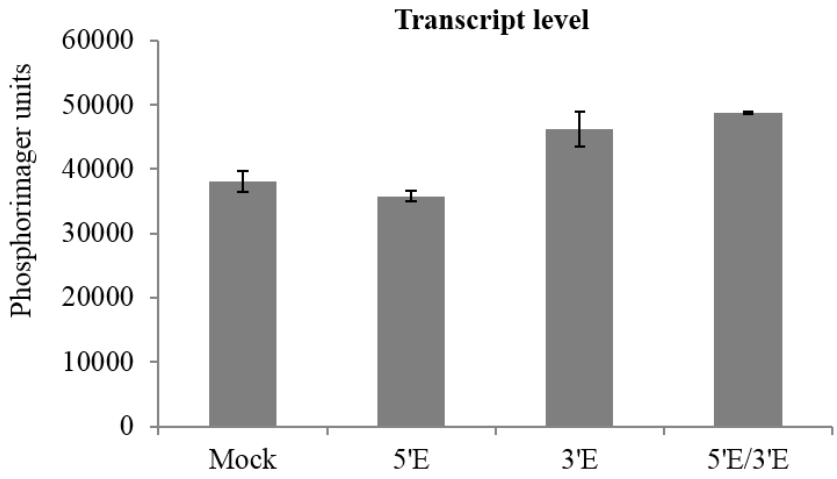


Figure 4.10. Pre-mRNA undergoes degradation through the 3’ end in Pol II-TSR. (A) *In vitro* Pol II transcription/splicing assay with 5’E and 3’E oligonucleotides. CMV-Bg C DNA template was used. The 5’E, 3’E oligonucleotides and the DNA template were added all at once. The experiment was done in triplicate. The samples were incubated for 90 min. (B) The diagram shows transcript levels. Error values are standard deviation values.

4.3.2 RNA protection in SR

To investigate the functionality of 5'E oligonucleotide, it was used in a reaction with pre-made Bg C pre-mRNA under splicing conditions. Based on the result from Figure 3.15, which suggested that the nascent RNA is capped in Pol II-TSR, capped pre-made pre-mRNA was used in SR. The experiment was done in triplicate. The 5' terminus of the U1 snRNA was blocked by anti-U1 oligonucleotide during pre-incubation, following by the addition of 5'E oligo together with the Bg C pre-mRNA. Time points at 0 and 90 min were taken. Just 34% of RNA was present in the "anti-U1" sample. The analysis revealed a slight effect of the 5'E oligo on transcript level (Figure 4.11). Since 48 % of the pre-mRNA was still subject to degradation upon inhibition of the U1 snRNA, it is concluded that capped pre-mRNA was not protected by the 5'E oligo even at the highest concentration. Since the cap structure of transcripts plays a role in protecting RNA from 5' exonuclease digestion, it is unlikely that pre-made pre-mRNA was decapped in the reaction with inactive U1 snRNP. If the pre-mRNA was decapped upon inhibition of the U1 snRNP then the 5'E oligo would protect it from 5' exonucleases.

The majority of pre-mRNA in SR was not degraded by 5' exonucleases, suggesting that the pre-mRNA might undergo degradation by 3' exonucleases as the 3' end might be accessible for 3' exonuclease hydrolysis. This was tested by using 3'E oligo in the samples with blocked U1 snRNP. The concentration of this oligo was titrated from 0.01 µM to 1 µM using capped Bg C pre-mRNA. The experiment was not done in triplicate. The result showed that in a dose-dependent manner the 3'E oligo protected the pre-mRNA from degradation (Figure 4.12). A 1.5-fold increase of transcript level at 1 µM was achieved in comparison to the sample with just anti-U1 oligonucleotide. The experiment demonstrated that the U1 snRNP plays a role in protecting transcripts against 3' exonuclease digestion. However, the 3'E oligo did not protect capped RNA from degradation completely, it is possible that some internal degradation by endonucleases could take place.

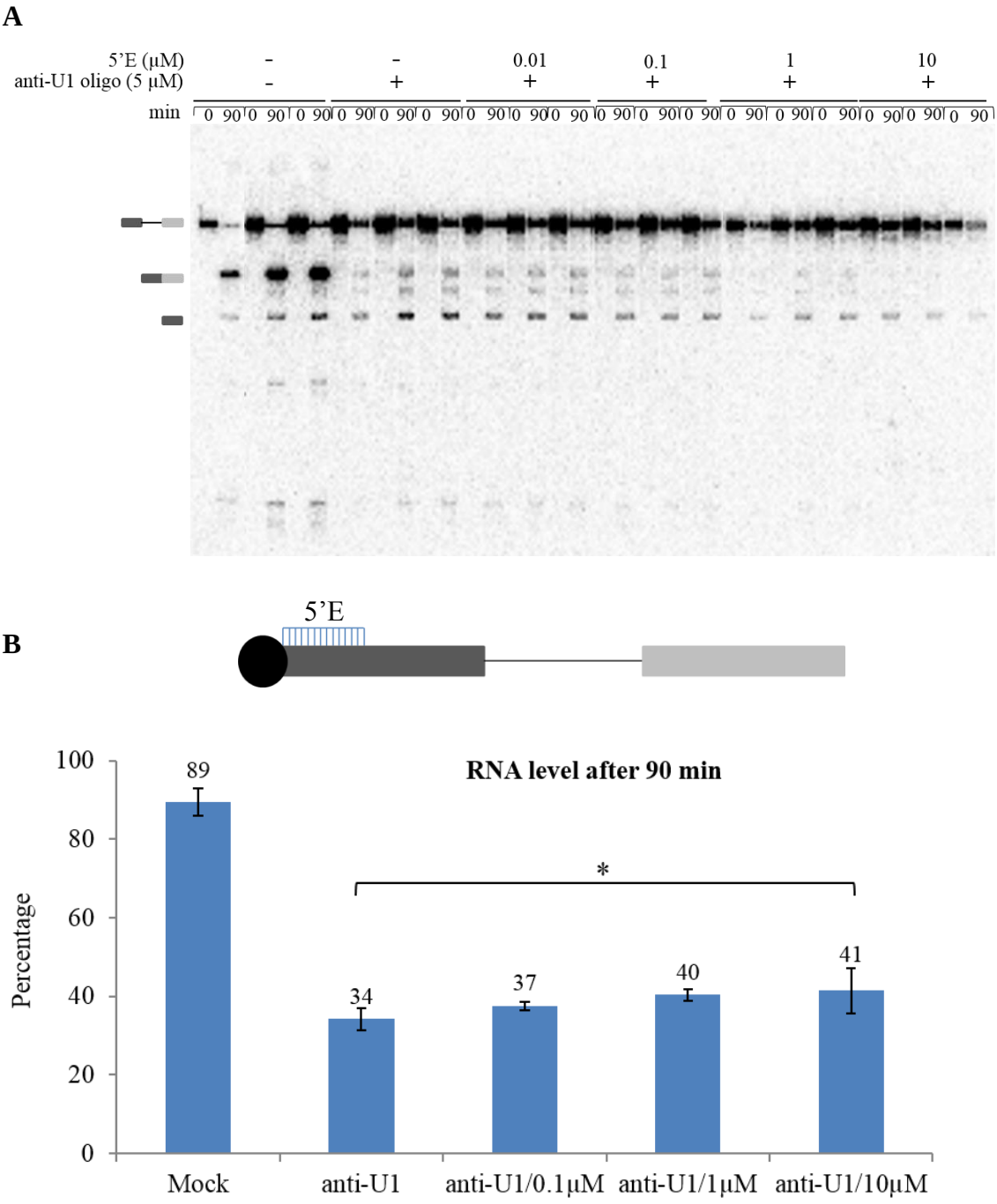


Figure 4.11. The 5'E oligonucleotide does not protect pre-made capped pre-mRNA from degradation in SR. (A) *In vitro* splicing reactions in triplicate with anti-U1 (5 µM) and 5'E oligonucleotides. Capped Bg C pre-mRNA was used. The concentrations of 5'E oligonucleotide are represented. The time points are indicated. (B) The diagram shows transcript levels which were measured as the percentage of RNA at 90 min to the input RNA at 0 min. (*) $P < 0.05$. Error values are standard deviation values.

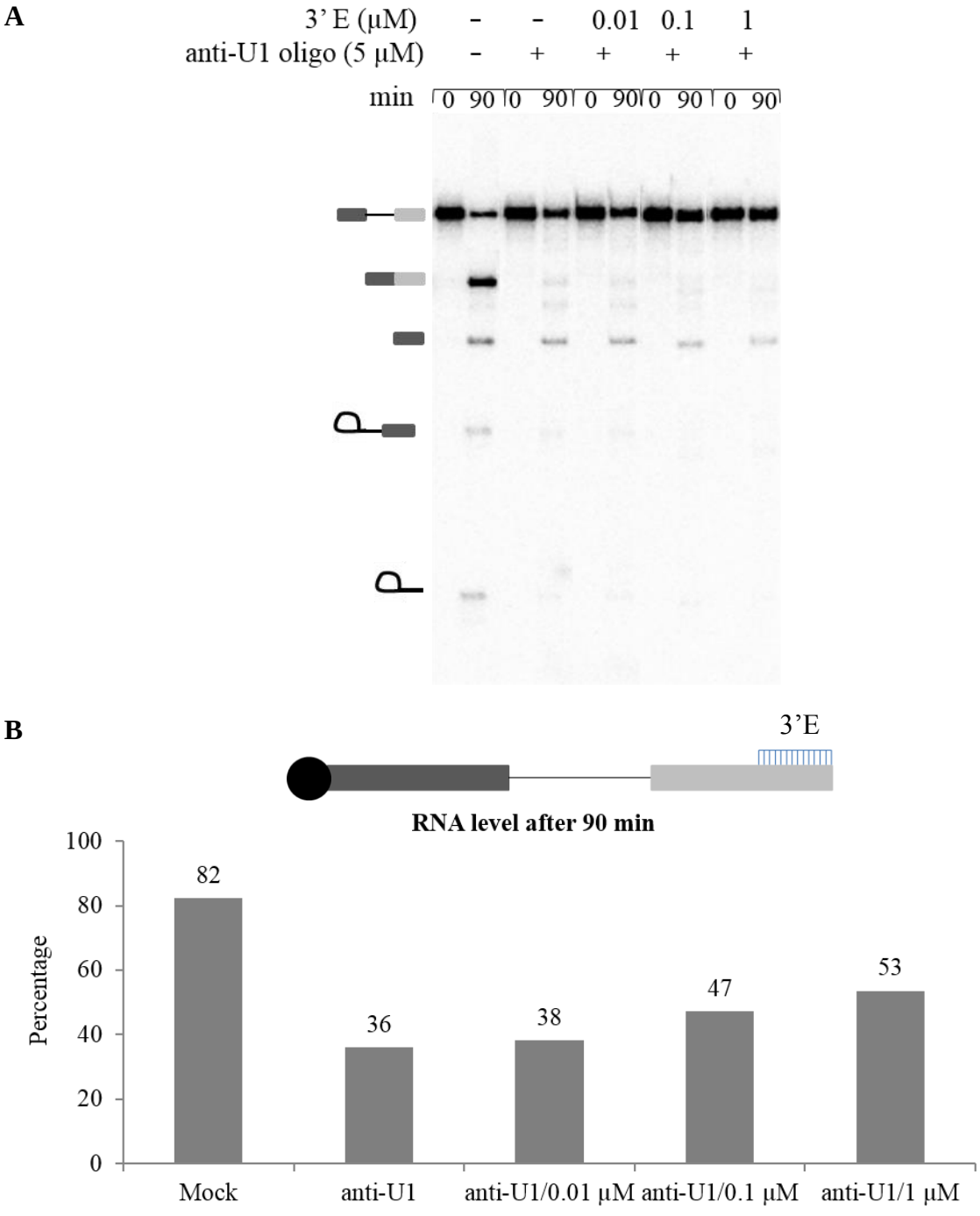


Figure 4.12. The 3'E oligonucleotide protects capped pre-mRNA from degradation in SR. (A) *In vitro* splicing reactions with anti-U1 (5 μM) and 3'E oligonucleotides. Capped Bg C pre-mRNA was used. The experiment was not done in triplicate. The concentrations of 3'E oligonucleotide are indicated. (B) The diagram shows transcript levels which were measured as the percentage of RNA at 90 min to the input RNA at 0 min.

Since the U1 snRNP is able to somehow protect a transcript from degradation if the transcript has a free 3' end, we hypothesized that the U1 snRNP could protect transcripts from 5' exonuclease activity if the 5' end were accessible. To test this idea, uncapped pre-synthesised Bg C pre-mRNA was used. Based on the previous experiment (Figure 4.11), the 5' E oligo at a concentration of 1 μM was used. The result showed that in the mock sample 64% of the pre-mRNA was degraded if the pre-mRNA did not contain the cap structure (Figure 4.13). However, the 5' end protection of RNA using the 5'E oligo only did not prevent degradation (sample "5'E"). Just 10% of the pre-mRNA was left in the sample with blocked U1 snRNA. Interestingly, uncapped pre-mRNA was protected by the 5'E oligo in the sample with blocked U1 snRNA. It is clear that RNA in the "anti-U1/5'E" sample was restored to the same level in the mock-treated sample suggesting that the U1 snRNP protects uncapped RNA from 5' end degradation.

The cooperative function of the 5'E and 3'E oligonucleotides was tested at a concentration of 1 μM. Uncapped Bg C RNA was used. This experiment with two protective oligonucleotides proved that the U1 snRNP has the property of protecting uncapped RNA from degradation, as the level of RNA was even higher in the sample with anti-U1, 5'E and 3'E oligos in comparison to the mock sample (Figure 4.14). Based on the results demonstrating that uncapped and not polyadenylated transcripts are subjected to degradation through the free ends, it is concluded that the U1 snRNP stabilises RNA via protecting it from degradation by 5' and 3' exonucleases.

The difference between the levels of transcripts in "anti-U1/5'E/3'E" and "5'E/3'E" reactions is that in "5'E/3'E" reaction the pre-mRNA was spliced while in "anti-U1/5'E/3'E" reaction splicing was prevented. This suggests that unspliced and spliced transcripts might undergo different degradation pathways. Therefore, both 5'E and 3'E oligos were ineffective to protect spliced RNA in "5'E/3'E" sample.

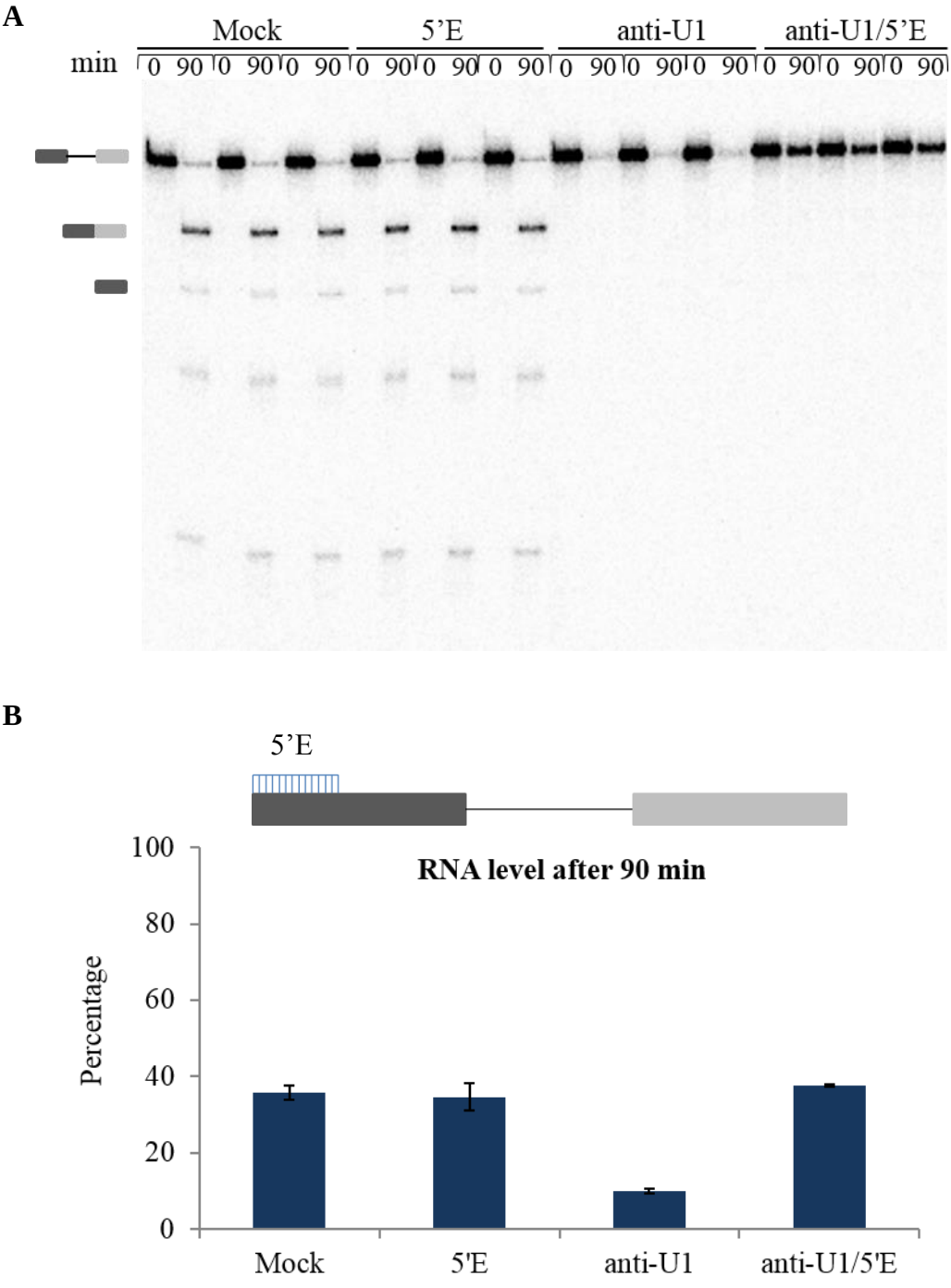


Figure 4.13. The 5'E oligonucleotide protects uncapped RNA from degradation in SR. (A) *In vitro* splicing reactions in triplicate with anti-U1 (5 µM) and 5'E (1 µM) oligonucleotides. Uncapped Bg C pre-mRNA was used. (B) The diagram shows transcript levels which were measured as the percentage of RNA at 90 min to the input RNA at 0 min. Error values are standard deviation values.

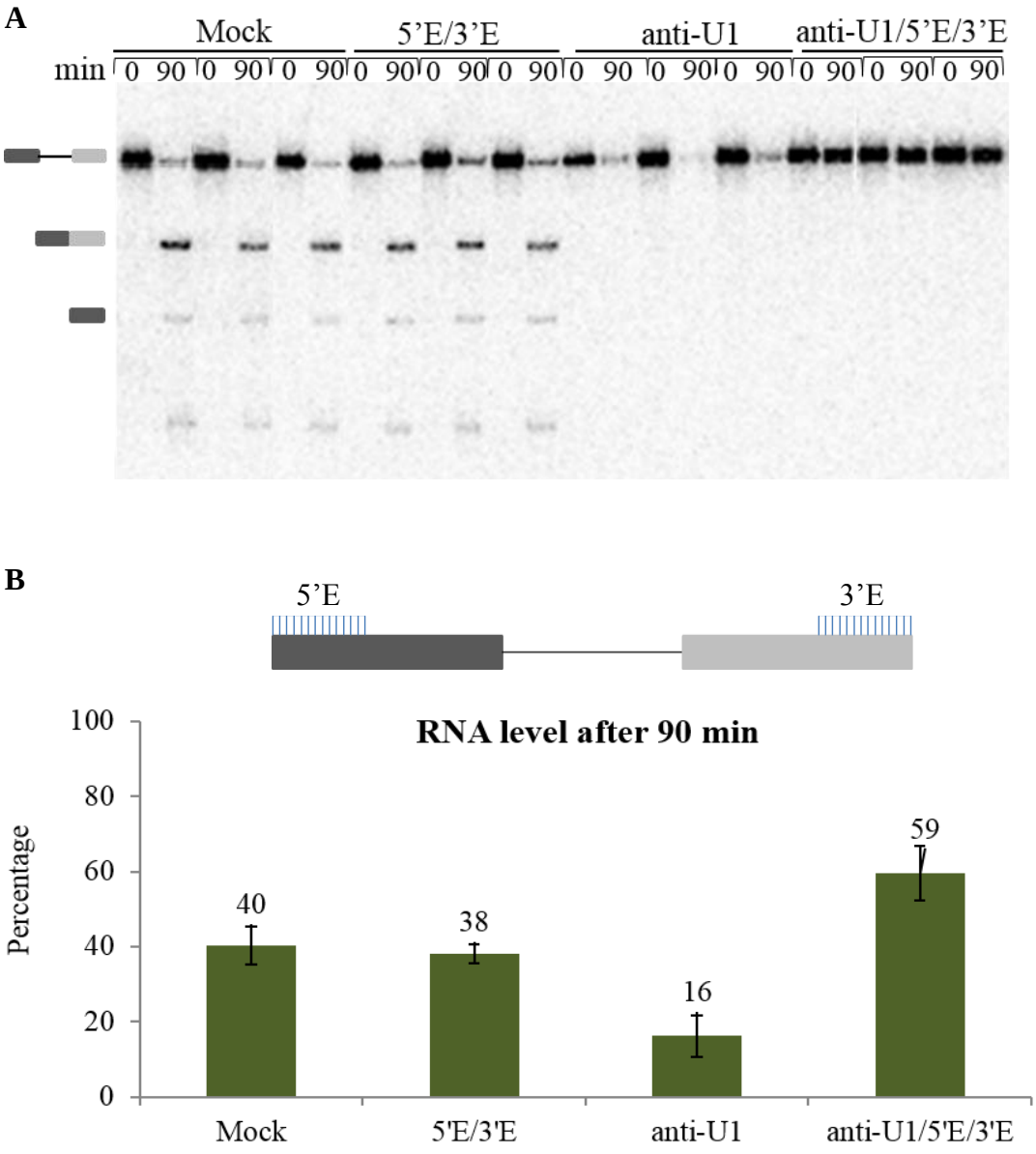


Figure 4.14. The 5'E and 3'E oligonucleotides protect uncapped RNA from degradation. (A) *In vitro* splicing reactions in triplicate with anti-U1 (5 µM), 5'E and 3'E oligonucleotides. Uncapped Bg C pre-mRNA was used. The protective oligonucleotides were used at the concentration of 1 µM. (B) The diagram shows transcript levels which were measured as the percentage of RNA at 90 min to the input RNA at 0 min. Error values are standard deviation values.

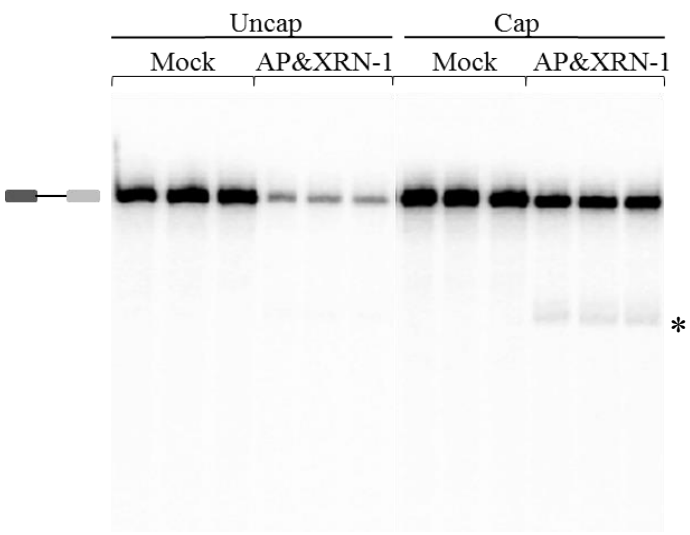
Interestingly, RNA level was not changed in the sample with the 5'E and 3'E oligonucleotides compared to the mock reaction (Figure 4.14). Another independent experiment was done to reproduce this phenomenon showing consistency in the results (data not shown). These data fit our previous observations which showed that just 5'E oligo was not able to protect uncapped RNA from degradation (sample "5'E"; Figure 4.11). If the cap structure protected RNA just from 5' end degradation then the 5'E oligo (sample "5'E") would rescue RNA leading to increased levels of RNA. This implies that the protection of free ends of the transcript does not prevent it from degradation. Therefore, it is suggested that the cap structure protects RNA not only from 5' exonucleases but other nucleases. It is possible that CBC promotes the recruitment of RNA binding proteins to RNA mediating the stability of RNA (reviewed by Gonatopoulos-Pournatzis and Cowling, 2014).

4.3.3 Stability of capped RNA

The experiments (Figure 4.7; Figure 4.11; Figure 4.12) showed that capped pre-made RNA is not stable after inhibiting the U1 snRNA. One possible explanation is that some of this RNA had not been capped during *in vitro* transcription. It is not known whether all transcripts contained cap after *in vitro* transcription. It is possible that a proportion of RNA can contain GTP at the 5' end. This can happen because GTP can be incorporated as the first nucleotide of the RNA strand too. Thus, after *in vitro* transcription some RNA species are capped while some might be uncapped, indicating the presence of heterogeneous RNA species in the sample. We hypothesised, if a proportion of transcripts do not contain the cap structure but GTP then this RNA can be a substrate for 5' end degradation. The next experiment was done to test this. RNA was treated with Antarctic Phosphatase (AP) and 5' exonuclease XRN-1 (Materials and Methods). It was expected that AP would remove two phosphate groups from the terminal nucleotide of the RNA and when the monophosphate is left the RNA would be subject to XRN-1 digestion. As a control for the efficiency of AP and XRN-1 treatment, uncapped RNA was used. Thus, capped and uncapped Bg C transcripts were used. After the treatment by AP and XRN-1 (Materials and Methods), 82% of the uncapped RNA was found to have been degraded (Figure 4.15). Since 18% of the uncapped RNA was

not degraded it is suggested that the AP and XRN-1 treatment was not completely efficient.

A



B

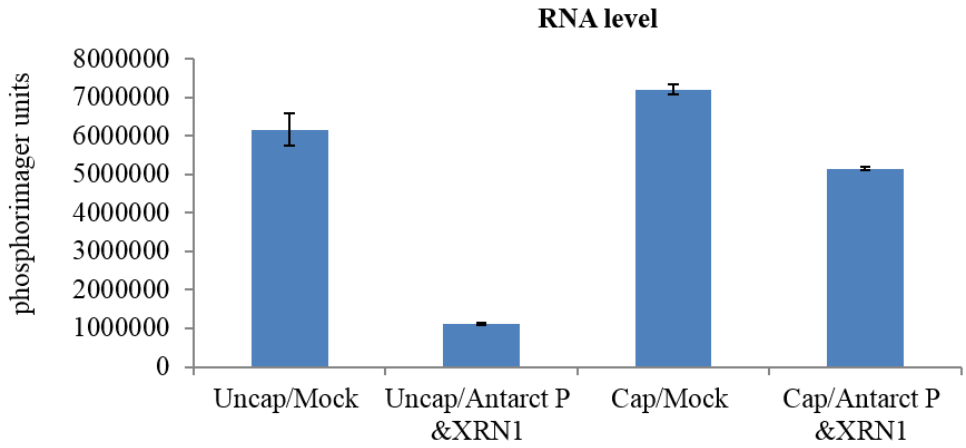


Figure 4.15. 1/3 of pre-made pre-mRNA does not contain the cap structure. (A) PAGE shows pre-made capped (cap) and uncapped (uncap) Bg C pre-mRNAs after treatment by Antarctic Phosphatase (AP) and XRN-1. The experiment was done in triplicate. A band (*) is a product of degradation. (B) The level of RNA is shown in the diagram. Error values are standard deviation values.

As we expected, the efficiency of cap analogue incorporation during *in vitro* transcription was not 100%, as 29% capped RNA was degraded. This indicates that in the sample with capped RNA, most of the remaining 71% of the RNA contained the cap structure. This RNA was considered to be fully capped and taken for further purification (Materials and Methods).

An *in vitro* splicing assay was done using purified fully capped pre-made pre-mRNA. Anti-U1 oligonucleotide was added. The result showed that the capped RNA suffered from degradation in the sample with blocked U1 snRNA (Figure 4.16). This implies that both types of RNAs (capped and uncapped) suffer from degradation if the U1 snRNA is not functional. This result together with the previous data (Figure 4.11) suggest that capped RNA does not become a substrate for 5' exonucleases when the U1 snRNA has been blocked. Based on our results (Figure 4.11) it is suggested that capped RNA is targeted by 3' exonucleases and might be degraded by endonucleases.

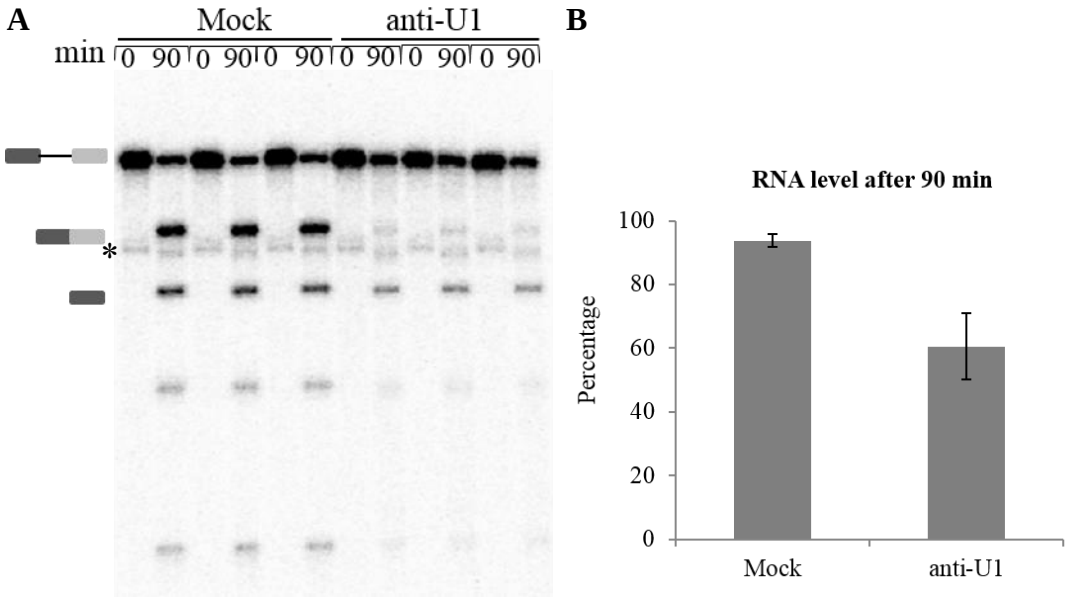


Figure 4.16. Fully capped pre-mRNA undergoes degradation if the U1 snRNP is not functional. (A) *In vitro* splicing reactions in triplicate with and without anti-U1 oligonucleotide. Capped Bg C pre-mRNA after Antarctic Phosphatase and XRN-1 treatment was used. After purification, RNA was added to SR. A band (*) at 0 min time point is a product which remains after the treatment. (B) The diagram shows transcript levels which were measured as the percentage of RNA at 90 min to the input RNA at 0 min. Error values are standard deviation values.

4.3.4 Evidence of RNA degradation

Although the experiments showed that the U1 snRNA protects transcripts from degradation but there was no clear evidence of degradation, except in Figure 4.8. We thought that small pieces of RNA were lost during gel electrophoresis. To test this idea, electrophoresis of Pol II-TSR with blocked U1 snRNA was carried out till radiolabelled nucleotides reached the bottom of the gel. Surprisingly, no degradation products were found (data not shown). It is possible that nucleases degrade RNA rapidly and completely.

To detect degradation products, 2'-*O*-methyl oligonucleotides complementary to the centres of the intron and 5' exon of Bg C pre-mRNA were designed (E and I; Figure 4.17 A). We hypothesised if the RNA were degraded by 5' and 3' exonucleases then a band of approximately 200 nt would be visualised (Figure 4.17 A). After incubating the NE with anti-U1 oligonucleotide under splicing conditions, E and I oligonucleotides were added. This followed by the addition of capped Bg C pre-mRNA. To increase the visibility of degradation products, 10 times as much pre-mRNA was added in comparison to previous SR assays (20 fmol and 2 fmol, respectively). Time-course experiment was done. The result showed that E and I oligos had no effect on splicing (Figure 4.17 B). A complete inhibition of splicing was not achieved in the sample with anti-U1 oligonucleotide only as the mRNA and 5' exon intermediate were visualised. A band of bigger size than the mRNA (278 nt) was detected in the sample with anti-U1, E and I oligos (shown in arrow). Based on the results from Figure 4.11 and Figure 4.12 which showed that capped RNA is digested by 3' but not 5' exonucleases, it is suggested that the RNA is degraded by 3' exonucleases until the position of I oligo on the RNA. Thus, 294 nt RNA product was detected (Figure 4.17 C). Also, this result proves that the capped RNA is not degraded by 5' exonucleases if the U1 snRNA is blocked, as if it was the case a band of 199 nt would be observed (Figure 4.17 A). Also, by increasing the concentration of transcripts some other degradation products (DP) were detected in the samples with inhibited U1 snRNA. This observation suggests that degradation products were not visualised in the previous experiments because low amounts of transcripts were present in SR, Pol II-TSR and T7-TSR assays.

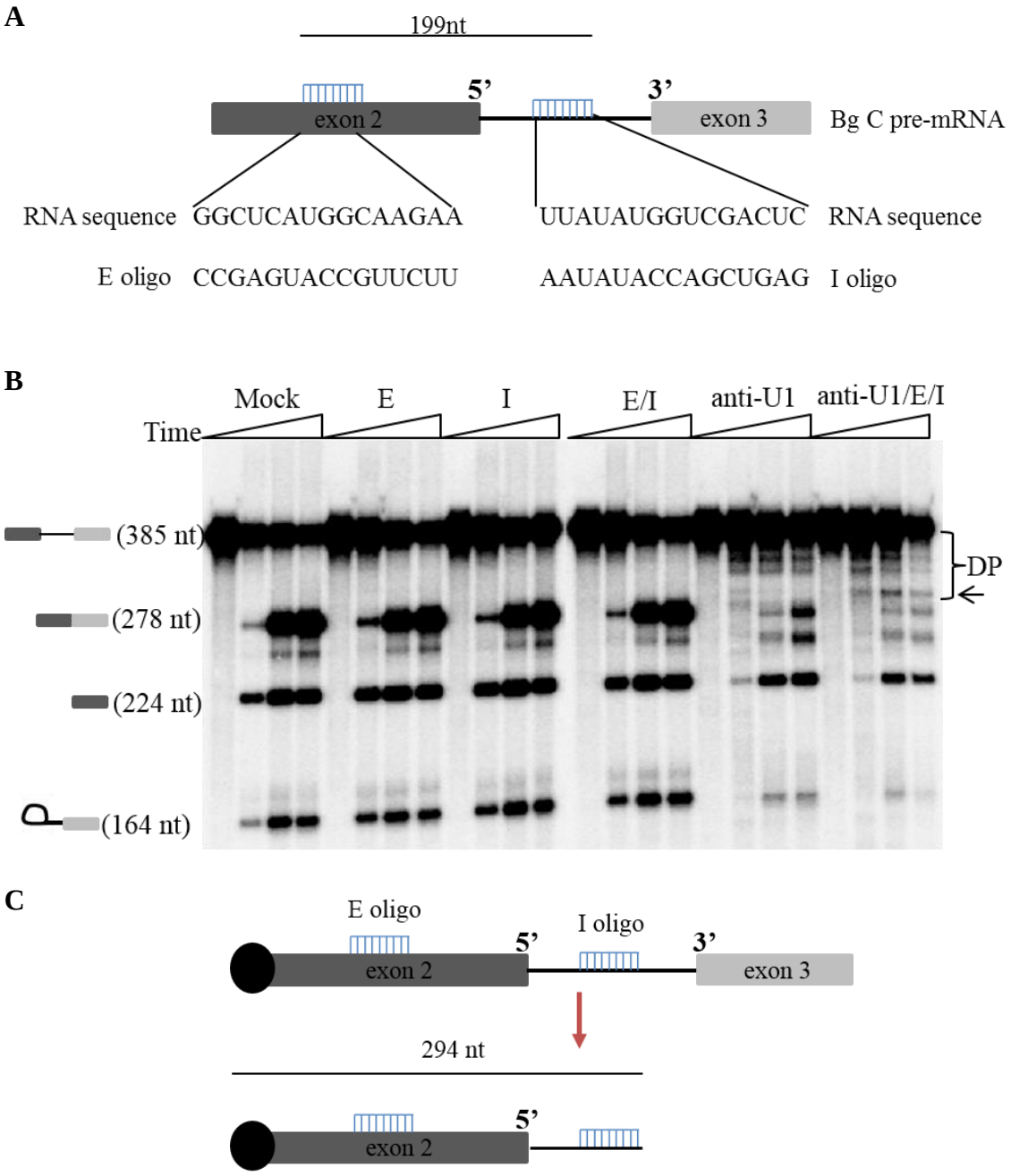


Figure 4.17. Evidence of RNA degradation. (A) Diagram shows the position of 2'-*O*-methyl oligonucleotides (E and I) annealing to exonic and intronic regions of Bg C pre-mRNA. The sequences of oligonucleotides are indicated. The cap structure is shown in black circle. 199 nt band product is expected if the RNA is degraded through the 5' and 3' ends. (B) *In vitro* time-course splicing assays with oligonucleotides complementary to the exon and intron of Bg C pre-mRNA. Under splicing conditions HeLa NEs were treated with and without anti-U1 2'-*O*-methyl oligonucleotide at 30 °C for 15 min prior to the addition of E and I oligonucleotides. In 10 min capped labelled Bg C pre-mRNA was added. The concentration of Bg C pre-mRNA was 20 fmol in each reaction. This followed by the incubation for further 90 min. Time points at 0, 30, 60 and 120 min were taken. The samples containing oligonucleotides are assigned. Complete inhibition

of splicing by anti-U1 snRNA oligonucleotide was not achieved as mRNA and 5' exon bands were detected. A band of degradation product in the sample with anti-U1 snRNA, E and I oligonucleotides is indicated by arrow. DP stands for degradation products, which can be seen in the anti-U1 and anti-U1/E/I samples. The sizes of the pre-mRNA, mRNA, 5' exon intermediate and lariat intermediate are represented. (C) Diagram shows full length Bg C pre-mRNA and a product with its size after 3' end degradation. The product of degradation is shown by arrow in (B).

4.4 Effects of 5'SS and 3'SS on RNA level

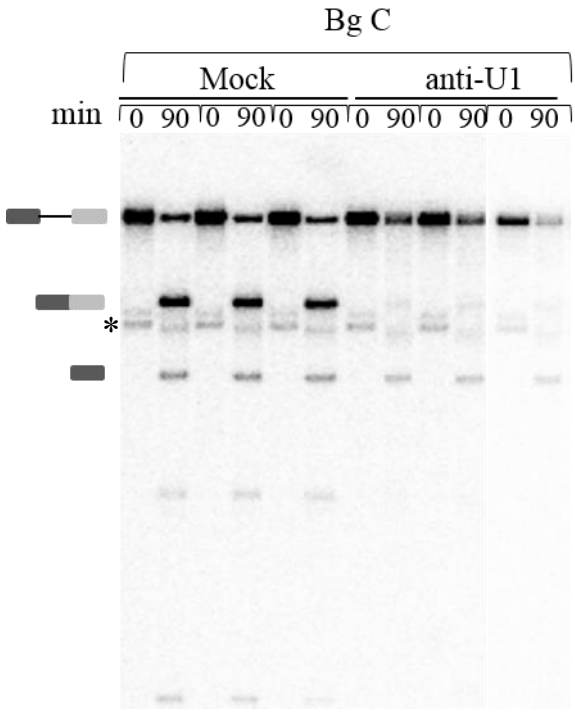
4.4.1 Effect of 5'SS on RNA level in SR

The described experiments revealed that inhibition or digestion of the 5'end of the U1 snRNA led to decreased stability of RNA regardless of the RNA polymerase used or whether RNA was transcribed in the NE or added to the NE. This suggests that the 5' terminal portion of the U1 snRNA enables the U1 snRNP to somehow prevent pre-mRNA from entering the degradation pathway. To test if RNA stability depends on the direct interaction between the U1 snRNA and pre-mRNA, a mutant construct that had no 5'SS was used (Hodson et al., 2012). This substrate is called Bg M. To exclude a possibility that other splice sites are required for RNA stability, a mutant version of Bg lacking any SSs could be used. However, to eliminate a possibility of not removing all necessary sites required for the binding of the spliceosome components, an intronless construct GFP was used. It was expected that neither Bg M RNA nor GFP RNA would be stable in the presence of the functional U1 snRNP. It is important to note that Bg C pre-mRNA contained two consensus 5'SSs: the first one was located at the beginning of the intron and the second one flanked the 3' exon. This construct was used in the majority of experiments as it was shown previously in the laboratory that Bg C pre-mRNA having an extra 5'SS at the 3' end of pre-mRNA was more stable in comparison to the pre-mRNA containing just the intronic 5'SS (Hodson et al., 2011).

The stabilities of pre-made Bg C, Bg M and GFP transcripts were tested in a conventional splicing assay with and without anti-U1 oligonucleotide (Figure 4.18). Each RNA was 100% capped prior to addition to the SR, as the RNAs had been treated with AP and XRN-1 (Materials and Methods). Each reaction was done in triplicate. To

measure the amount of RNA present after incubation, 0 and 90 min time points were taken (Figure 4.18 B). As expected, after 90 min of incubation 85% of the input pre-mRNA carrying 5'SSs remained stable (Bg C), while just 30% of transcript containing mutant 5'SS was present (Bg M). Compared to Bg C, Bg M displayed significantly higher rates of degradation. It is clear that the observed reduction in transcript level is due to the absence of the consensus 5'SS. GFP RNA was not stable suggesting that the presence of the 5' SS is important for RNA stability. Degradation rates of Bg M and GFP were not increased in the reactions with blocked U1 snRNA, suggesting that the consensus 5'SS is important for RNA stability. 60% of GFP RNA remained after 90 min, twice as much as for Bg C pre-mRNA. It is possible that the difference arises from the secondary structure of these two RNAs.

A



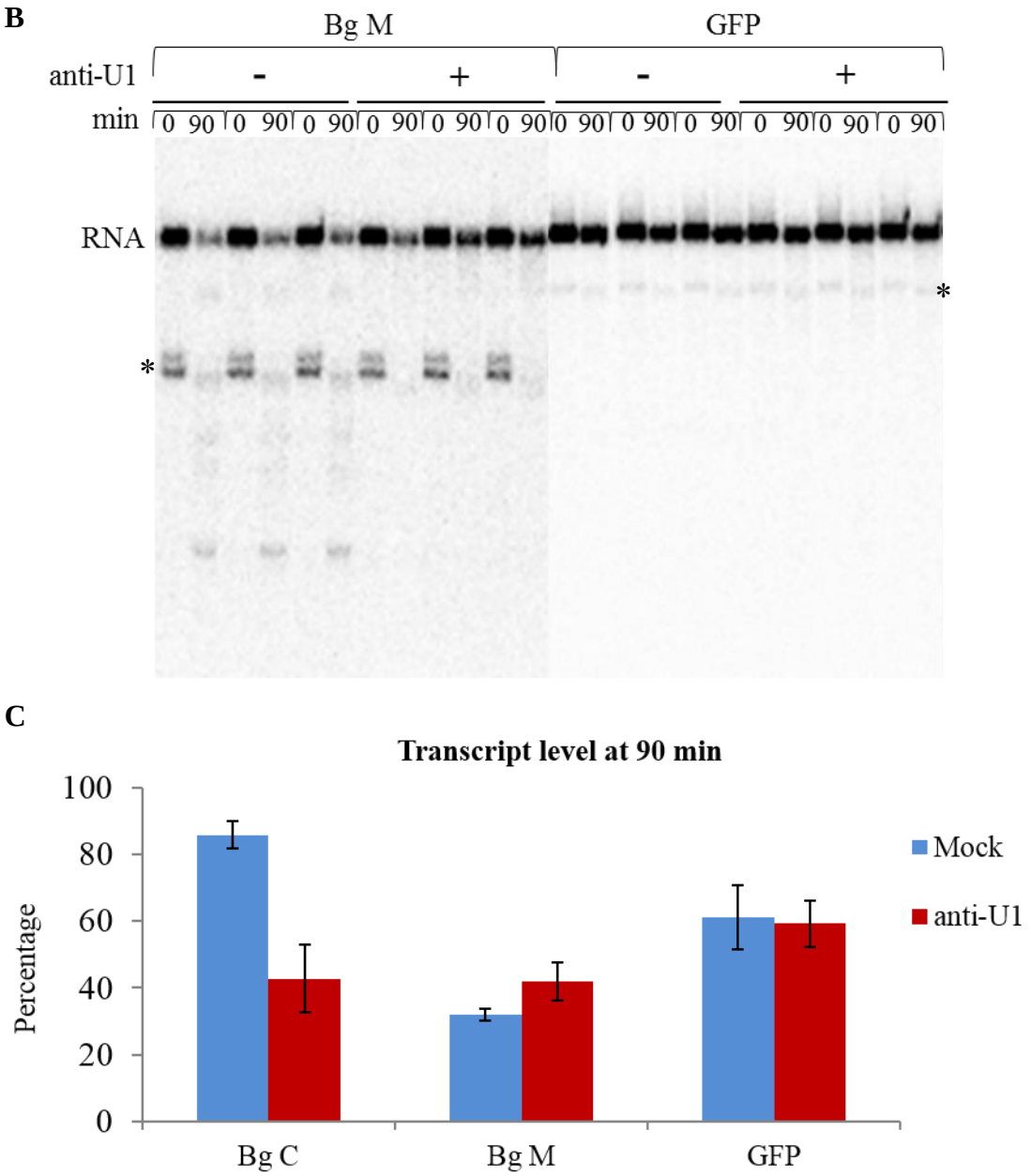


Figure 4.18. RNA stability depends on the functional 5'SS. (A and B) *In vitro* splicing reactions with and without anti-U1 oligonucleotide. The experiment was done in triplicate. A band (*) at 0 min time point is a product which remains after Antarctic Phosphatase and XRN-1 treatment. (C) The diagram shows transcript levels which were measured as the percentage of RNA at 90 min to the input RNA at 0 min. Error values are standard deviation values.

To test if RNA stability is 5’SS-position dependent, the consensus 5’SS was fused to the 3’ end of the Bg M pre-mRNA. This construct will be further called Bg K. It was expected that the overall Bg K transcript level would be increased compared to that of Bg M, as the strong U1 snRNA binding site was available. Surprisingly, the construct having the 5’SS at the 3’ end of the pre-mRNA (Bg K) was able to splice very efficiently (Figure 4.19). It is known that the genome contains a large number of cryptic 5’SS (Roca et al., 2012). Depending on the complementarity between the U1 snRNA and U1-like sequence the interaction can be “weak” or “strong” (Roca et al., 2013). This is known as the splice site strength. Bg C substrate was inspected for the presence of cryptic U1 snRNP binding sites. Eleven potential cryptic 5’SSs were found to be located in the 5’ exon of the Bg C pre-mRNA (Figure 4.20; umd.be, 2016). Since Bg K pre-mRNA produced splicing products it is suggested that cryptic 5’SSs were recognized by the U1 snRNP leading to spliceosome assembly. It is possible that the U1 snRNP bound to the 5’SS at the 3’ end of the Bg K pre-mRNA plays a role in stabilization of the interaction between a cryptic 5’SS and the U1 snRNP. Although the position of cryptic sites was known but it was not possible to determine which sites were recognised by the spliceosome to produce these splicing products. These splicing products could be assigned in the future. Relative number of molecules in each band was not calculated by normalising the intensity according to the number of GMP in the corresponding RNA molecules because the nature of splicing products was not determined. Therefore, the intensity of each band at 120 min time point was measured and relative RNA levels were calculated. The result showed that Bg K pre-mRNA was four-fold more stable than its mutant counterpart, but was 1.5-fold less stable than Bg C which harbours two “strong” 5’SSs (Figure 4.19). In conclusion, RNA stability depends on the presence of a “strong” 5’SS regardless of its position.

In the reaction with Bg M pre-mRNA, a band of 180 nt was detected. It is the same splicing product as that in the sample with Bg K pre-mRNA. It is suggested that the same cryptic 5’SS located in the 5’ exon was recognised by the U1 snRNP. This data suggests that at least one cryptic 5’SS can be bound by the U1 snRNP but this interaction was not sufficient to stabilize the Bg M pre-mRNA. In conclusion, only strong base-pairing between the U1 snRNA and pre-mRNA stabilizes RNA under SR conditions and a direct binding of the U1 snRNP to the 5’SS of pre-mRNA is required.

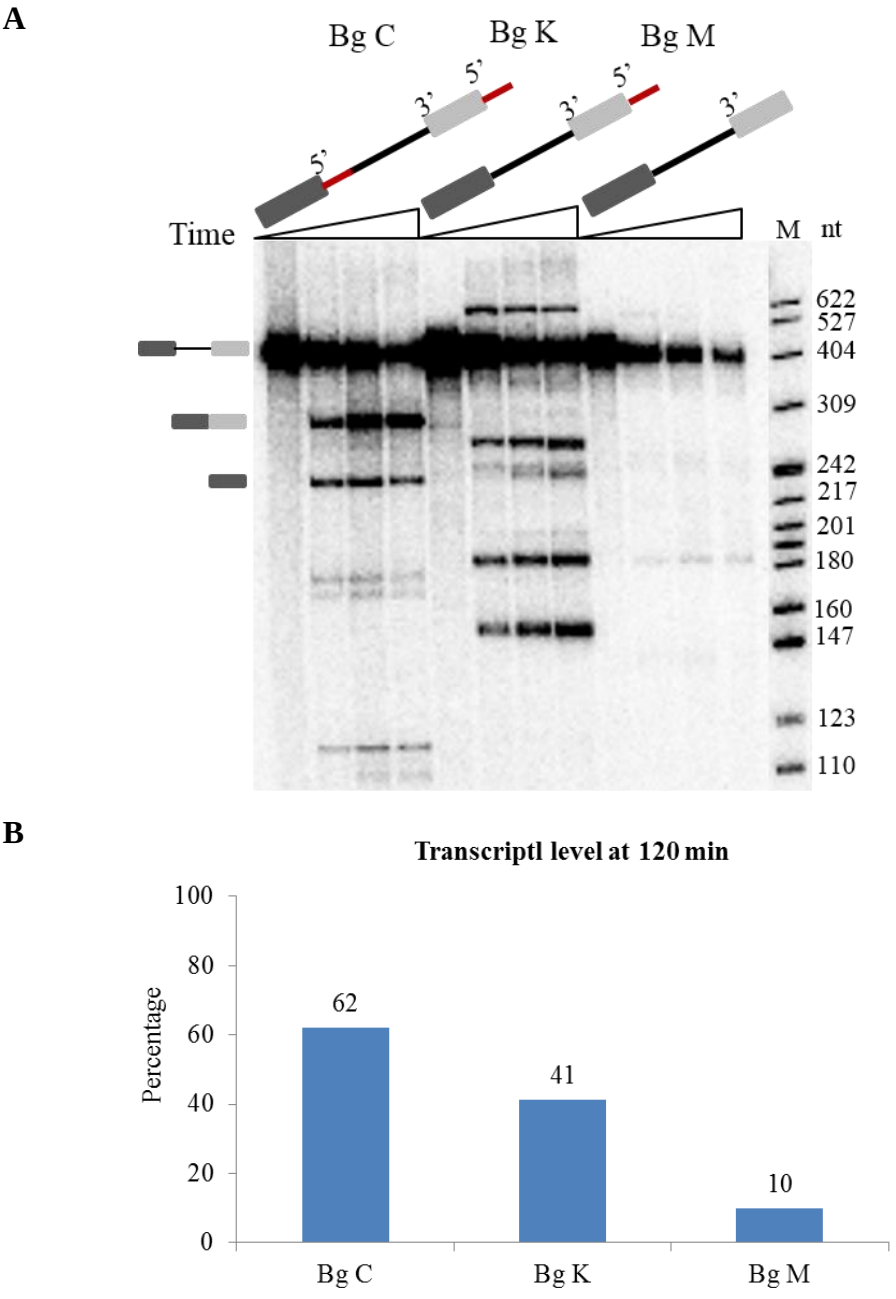


Figure 4.19. The presence of the 5'SS regardless its position is crucial for RNA stability in the conventional splicing assay. (A) *In vitro* time-course splicing of Bg C, Bg K and Bg M pre-mRNAs. The constructs are represented on the top of the phosphorimage. Time points at 0, 30, 60 and 120 min were taken. The 5'SS is indicated as the red line. (B) Transcript level based on the intensity of each band at 120 min time point is shown in the diagram. Relative number of molecules in each band was not calculated by normalising the intensity according to the number of rGTP in the corresponding RNA molecules because the nature of the bands in the sample with Bg K RNA was not determined.

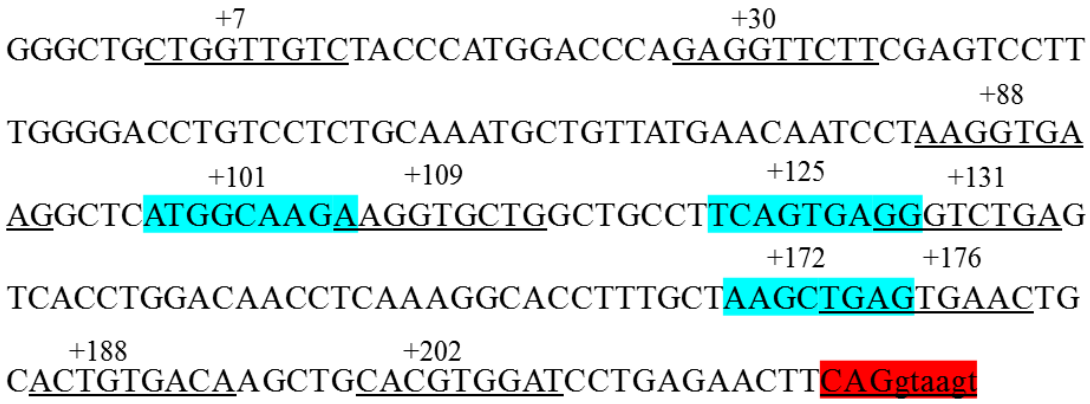
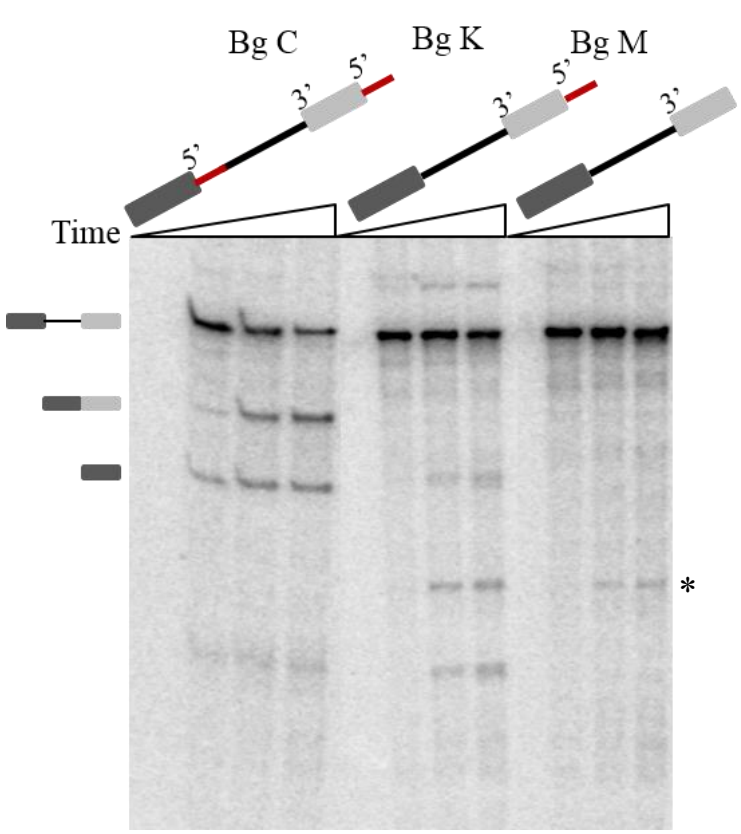


Figure 4.20. Sequence of the 5' exon of Bg C mini-gene with cryptic 5'SSs. The sequence in red indicates the consensus 5'SS located at the exon/intron junction. Eleven cryptic 5'SSs are represented as underlined and coloured to differentiate them as some sites overlap. The position of each site relatively to the start of the exon is indicated.

4.4.2 Effect of the 5'SS on RNA level in Pol II-TSR

The above experiments provided clear evidence of the dependence of stability on the U1 snRNP and a functional 5'SS. We next investigated if the consensus 5'SS is required for RNA stability in the reaction in which Pol II transcription is linked to splicing. Time-course Pol II transcription/splicing assays were done using the equal amounts of Bg C, Bg M and Bg K templates under the CMV promoter. The results showed that cryptic 5'SSs of Bg K transcript were recognised by the spliceosome as some splicing products were detected (Figure 4.21 A). Strikingly, there was a small difference between the levels of Bg C, Bg M and Bg K transcripts (Figure 4.21 B).

A



B

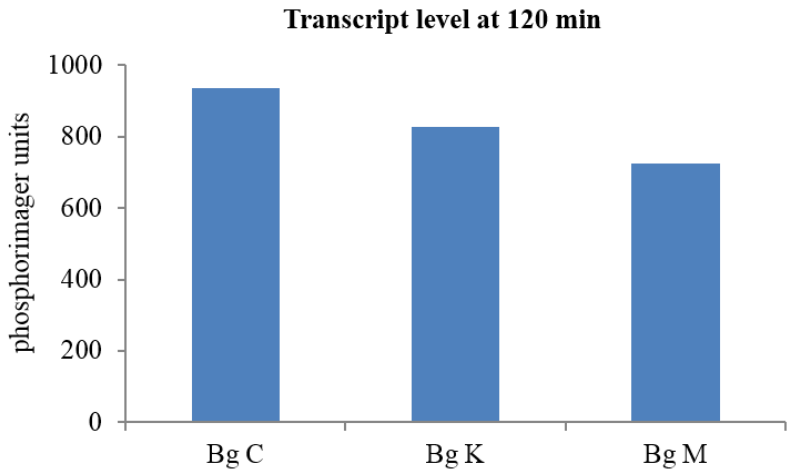


Figure 4.21. Effects of the consensus 5'SS on RNA level in Pol II-TSR assay. (A) *In vitro* time-course Pol II transcription/splicing using equal amounts of Bg C, Bg K and Bg M DNA templates under the CMV promoter. Time points of 0, 30, 60 and 120 min were taken. (*) One splicing product is produced in the sample with Bg M suggesting that a cryptic 5'SS is recognized by the spliceosome. (B) Transcript level is shown in the diagram.

Although equal amounts of DNA templates were used before (Figure 4.21), this experiment was repeated to exclude a possibility that different concentrations of DNA were taken. Bg C, Bg M and Bg K construct and Bg C without the extra 5'SS at the 3'end (called Bg A) were used. The experiment was done in triplicate. The effect of the U1 snRNP on these substrates was tested using anti-U1 oligonucleotide. Surprisingly, all transcripts were sensitive to the absence of the functional U1 snRNA as RNA levels of all transcripts were decreased (Figure 4.22), contrary to the data from SR assay (Figure 4.18). The experiment showed that in the mock-treated samples of Pol II-TSR the levels of Bg C and Bg M pre-mRNAs were similar, confirming the previous result (Figure 4.21). Strikingly, the ratio between RNA levels in the treated and untreated samples differed among the transcripts (Figure 4.23 C).

Comparing the results between SR and Pol II-TSR systems (Figure 4.18 and Figure 4.22, respectively), it is apparent that transcript levels between Bg C and Bg M significantly differ. In particular, in SR assay the level of Bg M pre-mRNA is 2.7-fold less than that of Bg C, but there is no difference in Pol II-TSR assay. In Pol II-TSR Bg M pre-mRNA lacking a strong 5'SS is stable to the same extent as Bg C pre-mRNA, which has two "strong" 5'SSs. Thus, the presence, position and number of the strong 5'SSs do not affect RNA stability in the reaction with ongoing Pol II transcription. The hypothesis is that the U1 snRNP could interact with cryptic 5'SS in a transcription-dependent manner resulting to stability of newly produced transcripts. In contrast to SR, this does not happen as splicing is impaired form transcription.

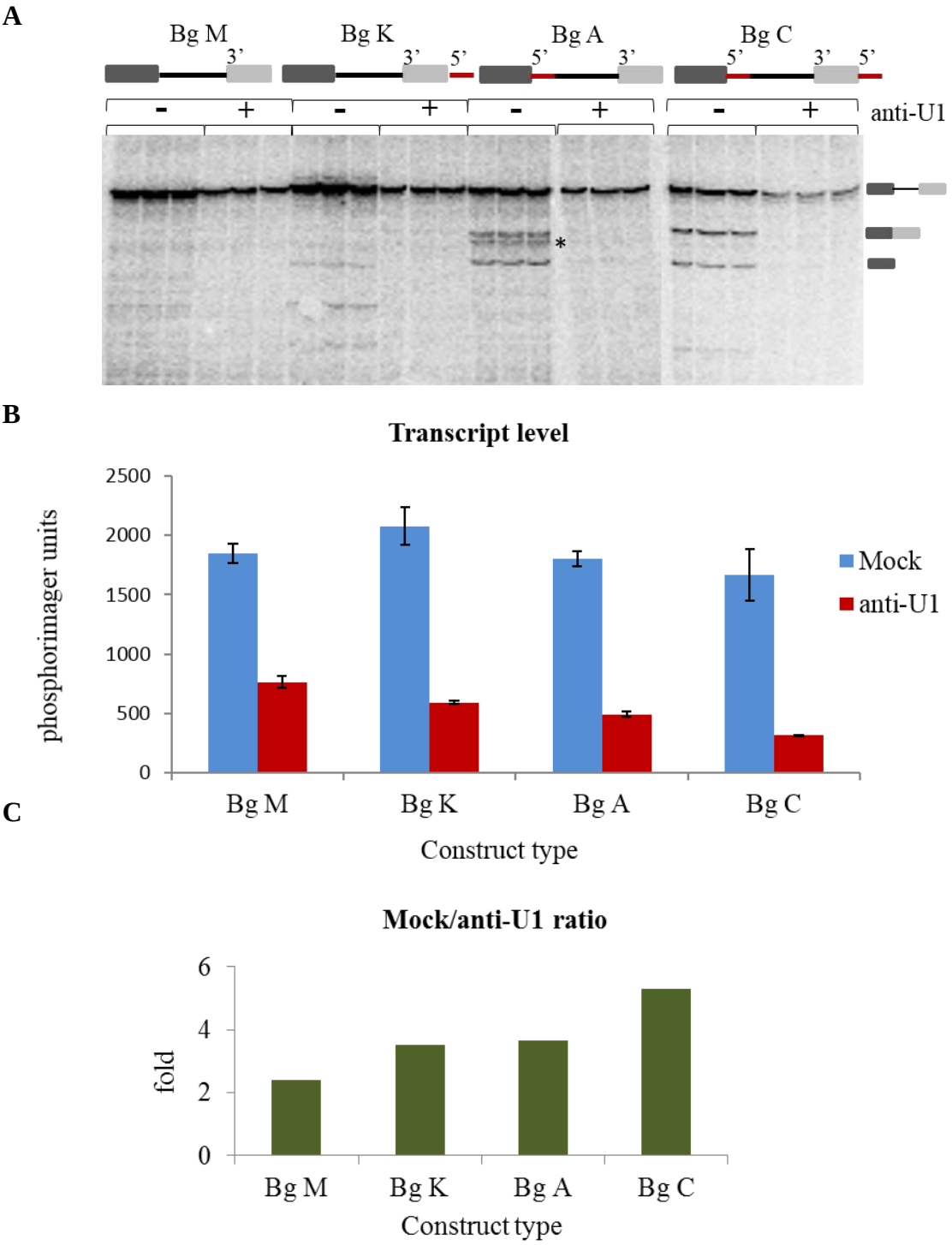


Figure 4.22. Regardless of the presence or position of the functional 5'SS, pre-mRNA is stable under conditions with ongoing Pol II transcription. (A) *In vitro* Pol II transcription/splicing reactions with and without anti-U1 oligonucleotide using different Bg constructs. The experiment was done in triplicate. The constructs are represented on the top of the phosphorimage. The incubation was carried out for 90 min. The 5'SS is indicated as the red line. One band (*) migrates between the mRNA and the 5' exon in the sample with Bg A substrate. This product is not detected in the reaction with the anti-U1 oligonucleotide. The origin of this product can be explained

based on previous study (Hodson, 2011) which suggested that a product of mRNA degradation occurs if there is no 5'SS at the 3' end of the pre-mRNA. (B) Transcript level is shown in the diagram. Error values are standard deviation values. (C) Quantitation shows a fold change between the mock and anti-U1-treated samples.

4.4.3 The stability of an intronless transcript depends on the presence of 5'SS

Together the above experiments showed that at least one cryptic 5'SS within the 5'exon can be bound by the U1 snRNP, but its usage depends on the presence of the 5'SSs at the 3' end of Bg C pre-mRNA. Also, the stability of Bg M transcript appeared to be different between the reaction in which splicing is separated from transcription and the reaction in which Pol II transcription occurs (Figure 4.18 and Figure 4.22). Based on previous findings that the core spliceosome components can associate with the transcription machinery (Fong and Zhou, 2001), we hypothesised that the U1 snRNP can be recruited to the nascent pre-mRNA which exits from Pol II and the U1 snRNP binds stably to not only the consensus but cryptic 5'SSs. Therefore, transcription mediates a strong association between the U1 snRNP and cryptic 5'SSs enabling RNA to avoid entering the degradation pathway. To test this hypothesis, the intronless GFP construct was inspected for the presence of any cryptic U1 snRNP binding sites. It was found that GFP contained 15 U1 snRNP-motifs (Figure 4.23; umd.be, 2016). Pol II-TSR with and without anti-U1 snRNA oligonucleotide was done in triplicate (Figure 4.24). The incubation was carried out for 90 min. The result showed that the level of GFP transcript in the reaction with blocked U1 snRNA was reduced by 1.4-fold in comparison to the mock-treated sample.

Figure 4.23. Sequence of GFP gene with cryptic 5'SSs. Fifteen cryptic 5'SSs are represented as underlined and coloured to differentiate them as some sites overlap. The position of each site relatively to the start of the gene is indicated.

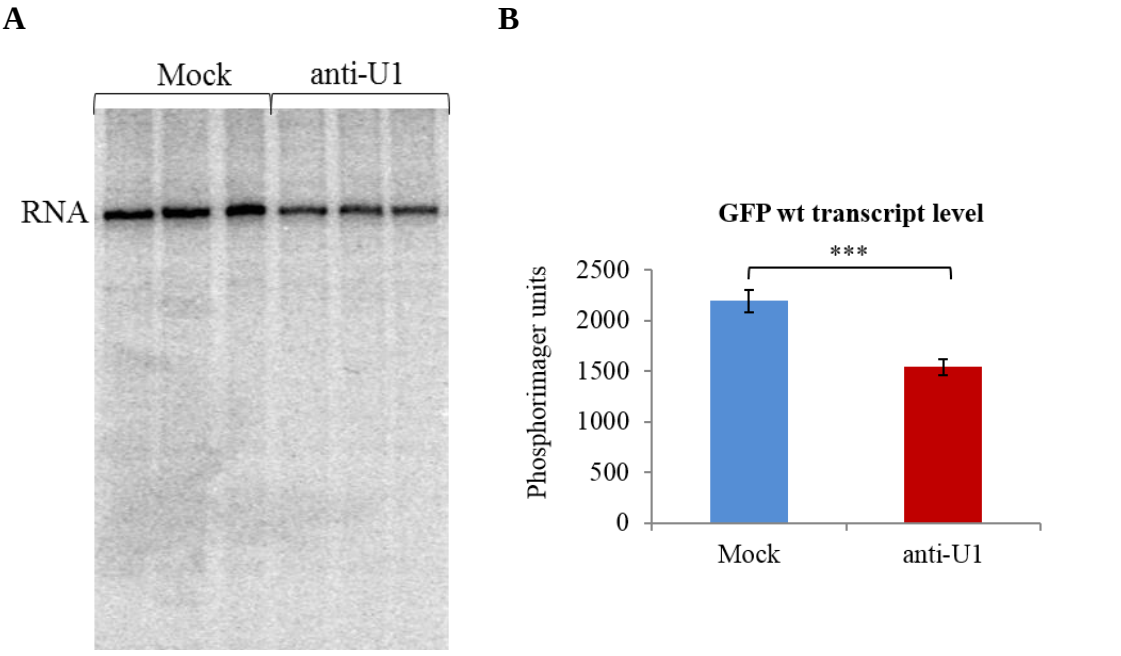
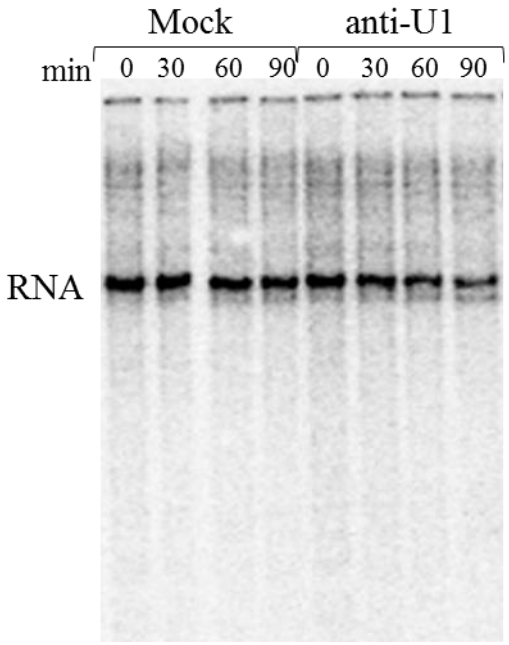


Figure 4.24. The stability of intronless RNA depends on the U1 snRNA. (A) *In vitro* Pol II transcription/splicing assay with and without anti-U1 oligonucleotide using the CMV-GFP DNA template. The experiment was done in triplicate. The samples were incubated for 90 min. (B) Transcript level is shown in the diagram. (*) $P < 0.01$. Error values are standard deviation values.

To test whether the U1 snRNP mediates RNA stability of the intronless gene, a pulse chase was done after 30 min of incubation. Time points at 30, 60 and 90 min were taken. RNA at each time point was measured relative to the amount of RNA at 0 min time point. The rates of RNA degradation is represented in Figure 4.25. The quantification of the degradation profiles showed an exponential decrease with time in both treated and untreated samples. However, a significant increase in the degradation rate of the newly synthesised pre-mRNA under conditions with blocked U1 snRNA was observed, which can be seen from the comparison of the corresponding slopes.

A



B

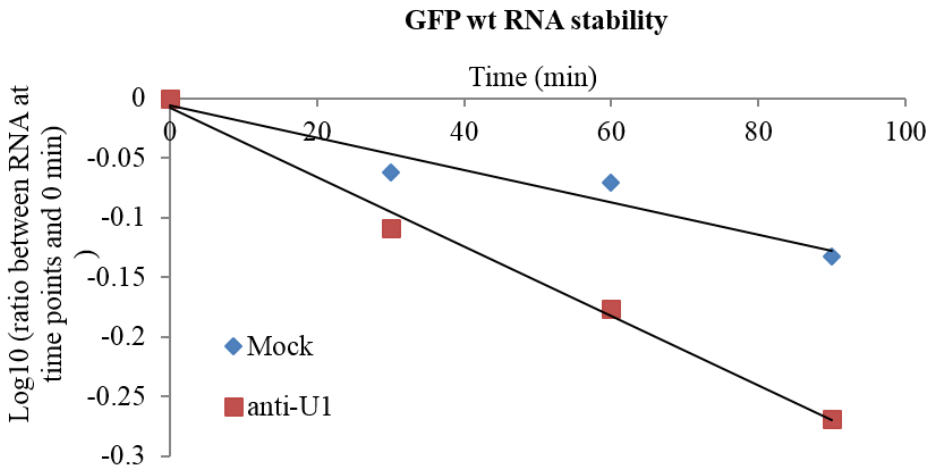
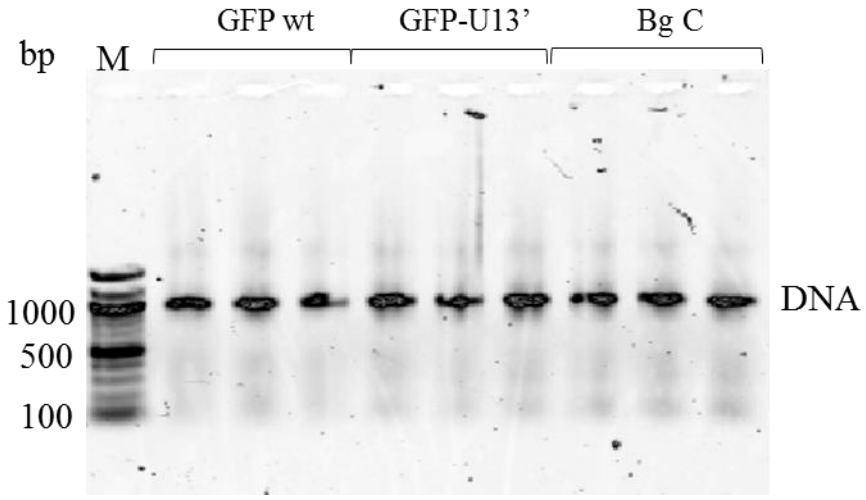


Figure 4.25. Intronless RNA transcribed by Pol II undergoes degradation if the U1 snRNA is blocked. (A) Inhibition of the U1 snRNA by 2'-*O*-methyl oligonucleotide using the CMV-GFP DNA in Pol II-TSR. The zero time point refers to the addition of excess unlabelled rGTP (pulse chase) 30 min after the start of incubation. (B) Quantification of the degradation profiles versus time is shown in the graph.

These observations allow to conclude that the intronless transcript also suffers from degradation if the U1 snRNA is not functional. We found that GFP gene contains 15 cryptic 5'SSs (Figure 4.23). It is suggested that cryptic 5'SSs might be bound by the U1 snRNP in a transcription-dependent manner. These interactions might increase the stability of the GFP RNA. These data fit the model about the U1 snRNA recruitment both to normally spliced and splicing-deficient transcription units (Spiluttiny, 2010), possibly through the association of the U1 snRNP with the transcription machinery (Fong and Zhou, 2001; Kim et al., 1997; Tian, 2001; Kwek et al., 2002; Jobert et al., 2009). These data allow to conclude that in Pol II-TSR RNA stability depends on the U1 snRNP but not on presence of a strong active 5'SS, unlike in SR.

To test if the consensus 5'SS affects the level of intronless RNA transcript in Pol II-TSR, it was fused to the 3' end of GFP. This construct was called GFP-U1(3'). Usually, before each experiment, the concentration of DNA templates (PCR products) was carefully estimated by spectrophotometer. To exclude a possibility that the concentration of DNA templates differed between samples, according to spectrophotometer measurements the amount of each DNA template was estimated by agarose gel electrophoresis. The experiment proved that equal amounts of DNA were added to each sample (Figure 4.26). The result from Pol II-TSR assay showed that the level of GFP transcript harbouring the 5'SS at the 3' end was increased compared to GFP wt with an observed fold change of 1.6 (Figure 4.27). This experiment was repeated showing the same result (data not shown). These data fit previous studies which showed that the 5'SS itself positively affects the level of transcripts *(*Barret et al., 1995; Damgaard et al., 2008*).* The difference in the level of RNA between samples with the GFP wt and Bg C DNA templates (Figure 4.27) allows to hypothesise that spliced products of Bg C pre-mRNA are not stable and undergo degradation after the completion of splicing. This fits our previous observations (Figure 4.14) suggesting that spliced and unspliced transcripts could undergo different pathways.

A



B

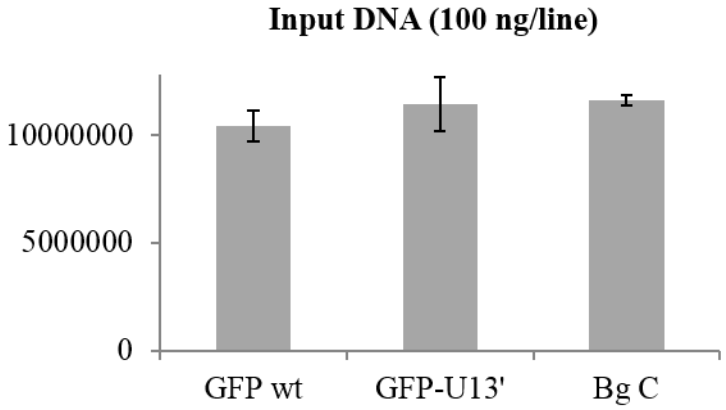
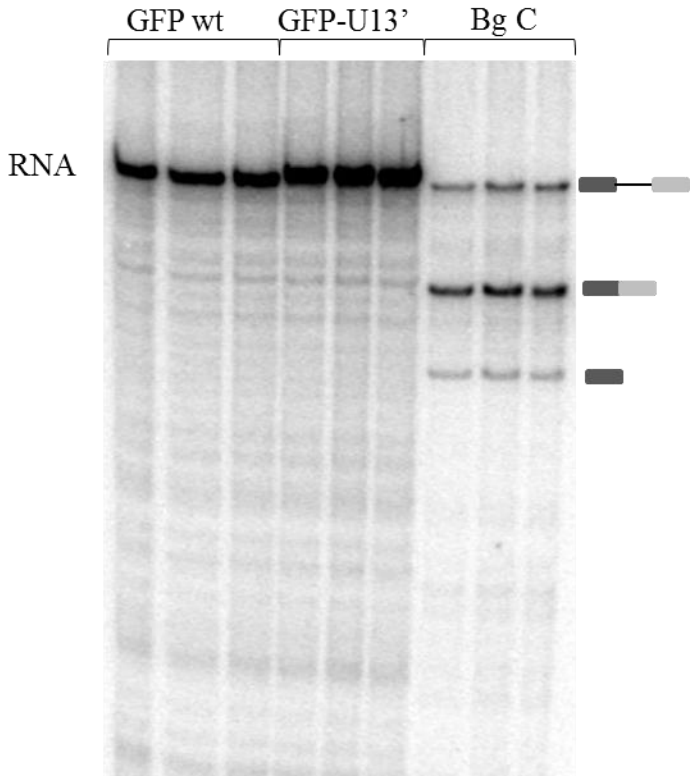


Figure 4.26. Testing if equal amounts of RNA are added to Pol II-TSR. (A) Agarose gel electrophoresis was done to test the amount of CMV-DNA used in Pol II-TSR (Figure 4.27). 100 ng of each DNA template was used. The experiment was done in triplicate. DNA substrates are indicated. (B) Quantification of the input DNA is shown in the diagram. Error values are standard deviation values.

A



B

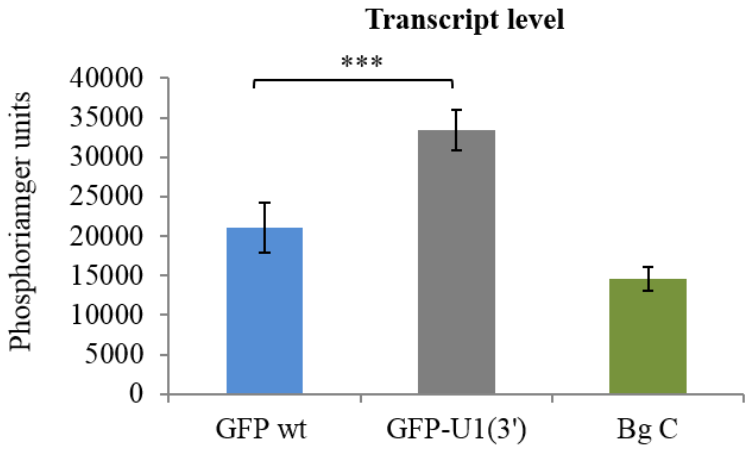


Figure 4.27. The level of intronless RNA is increased if RNA has the consensus 5'SS at the 3' end. (A) Effect of the consensus 5'SS at transcript level of an intronles substrate (GFP) in Pol II transcription/splicing assay. The reactions were done in triplicate. The samples were incubated for 90 min. (B) Quantification shows transcript levels. (*) $P < 0.01$. Error values are standard deviation values.

4.4.4 Effect of 3'SS on RNA level in Pol II-TSR

A connection between a functional 3' SS and transcription has been investigated *in vivo* (Fong and Zhou, 2001; Barret er al., 1995). It was shown that templates harbouring mutated 3'SS showed decreased Pol II transcription in comparison to a wt substrate (Fong and Zhou, 2001). The effect of the 3'SS on the level of transcript generated by Pol II polymerase was also tested in this study. Previously in the laboratory the AG dinucleotide of the 3'SS was mutated to GG (construct called Bg gg). To test Bg gg in Pol II-TSR system, this gene was cloned into pEGFP-C1 (Materials and Methods). The CMV-Bg C and CMV-Bg gg DNA templates were used in a time-course assay. As it was expected, there was an accumulation of the 5' exon and it was the only product observed in the sample (Figure 4.28 A). It indicates that the spliceosome was assembled on Bg gg pre-mRNA progressing to complex C. The level of transcripts at the 120 min time point was estimated showing a slight increase of RNA level in the sample with Bg gg construct (Figure 4.28 A). This result is not in line with previous observations (Fong and Zhou, 2001) but fits data obtained by Barret er al. (1995).

Since the above result showed that the 5'SS flanking the 3' exon facilitated splicing of pre-mRNA with the inactivated intronic 5'SS (Figure 4.19), this experiment was used as an opportunity to test the ability of the consensus 5'SS flanking the 3' exon to affect splicing of Bg gg. For that the CMV-Bg gg DNA template with the consensus 5'SS at the 3' end was used. This construct is called Bg gg U1BS. The experiment revealed that the 5'SS at the 3' end of the Bg gg pre-mRNA promoted splicing as the mRNA and 5' exon intermediate were detected. Based on the size of mRNAs in the samples with Bg C and Bg gg constructs, it is apparent that any alternative 3'SS were not targeted by the spliceosome but the mutated 3'SS. However, splicing of Bg gg U1BS was less efficient compared to Bg C as the percentage of mRNAs was 8 and 40, respectively (Figure 4.28 C). Nonetheless, it is concluded that the presence of the 5'SS at the 3' end of the pre-mRNA stimulated the complete pathway of the spliceosome resulting to mRNA production. The mechanism behind this phenomenon is not clear but it is suggested that through the binding to the 5'SS at the 3' end, the U1 snRNP could stabilise the interaction between U2AF35 and the mutated 3'SS.

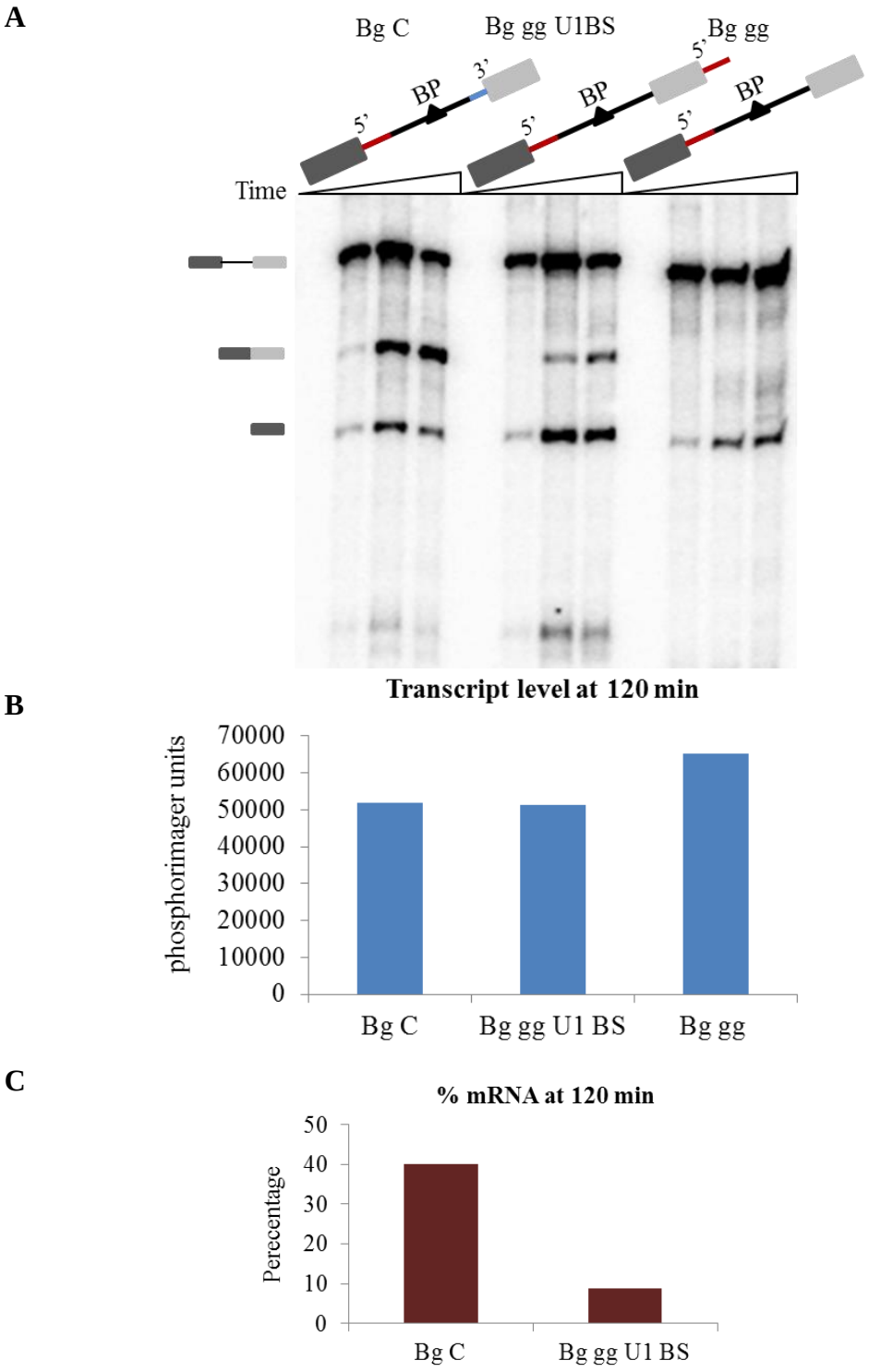


Figure 4.28. The 3' splice site does not affect RNA stability. (A) *In vitro* time-courses of Pol II transcription/splicing assay using Bg constructs with mutated 3'SS. The constructs are represented on the top of the phosphorimage. The 5'SS and 3'SS are indicated as the red and blue lines, respectively. BP stands for branch point. Time points at 0, 30, 60 and 120 min were taken. (B) The diagram shows transcript level at the last time point (120 min). (C) The percentage of spliced mRNA is shown in the diagram. Splicing was measured as the percentage by the following formula: [mRNA]/[pre-mRNA+mRNA+5'exon].

4.5 Complementation assays using U1 snRNP/snRNA

The next aim was to restore the ability of the U1 snRNP to stabilise RNA. This could be achieved by using NE with overexpressed U1 snRNA, the 5' terminus of which is mutated enabling the U1 snRNA to be complementary to other parts of Bg C. Unfortunately, efficient *in vitro* Pol II transcription/splicing was only supported by commercial nuclear extract and homemade extracts were not active (Figure 3.9). Thus, I was not able to use NE with overexpressed U1 snRNP. As an alternative, a complementation assay could be done using transcribed U1 snRNA or purified U1 snRNP. A previous study showed that splicing of U1-depleated NE was restored when either purified U1 snRNA or synthetically transcribed U1 snRNA were added to *in vitro* conventional SR (Will and Luhrmann 1996). It was suggested that regardless of the origin of the U1 snRNA (native or transcribed), the U1 snRNP can be reassembled in NEs if all complex components are present there. In another study splicing was affected by exogenous stem loop IV of the U1 snRNA (Sharma et al., 2014). Together these data indicate that the exogenous U1 snRNA/snRNP can be functional *in vitro*. We expected that transcripts could be stabilized in our *in vitro* assays when exogenous U1 snRNP and U1 snRNA are added.

The 5' terminus of the endogenous U1 snRNA was inhibited by antisense oligonucleotide and an excess of transcribed U1 snRNA or purified U1 snRNP were added, then the restoration of both transcript level and splicing could be achieved. The priority was given to 2'-*O*-methyl oligonucleotide to inhibit the U1 snRNA. The native U1 snRNP used in this study had been provided by Lührmann's laboratory. Although the previous study showed that the addition of 200 ng of the U1 snRNP restored splicing to normal levels (Will and Luhrmann 1996), the concentration of the U1 snRNP in Pol II-TSR was titrated. After the inhibition of the endogenous U1 snRNA, purified U1 snRNP was added and the samples were pre-incubated for 30 min. This was followed by the addition of the CMV-Bg C DNA template and further incubation for 90 min. The experiment revealed an inhibitory effect of purified U1 snRNP on transcription (Figure 4.29).

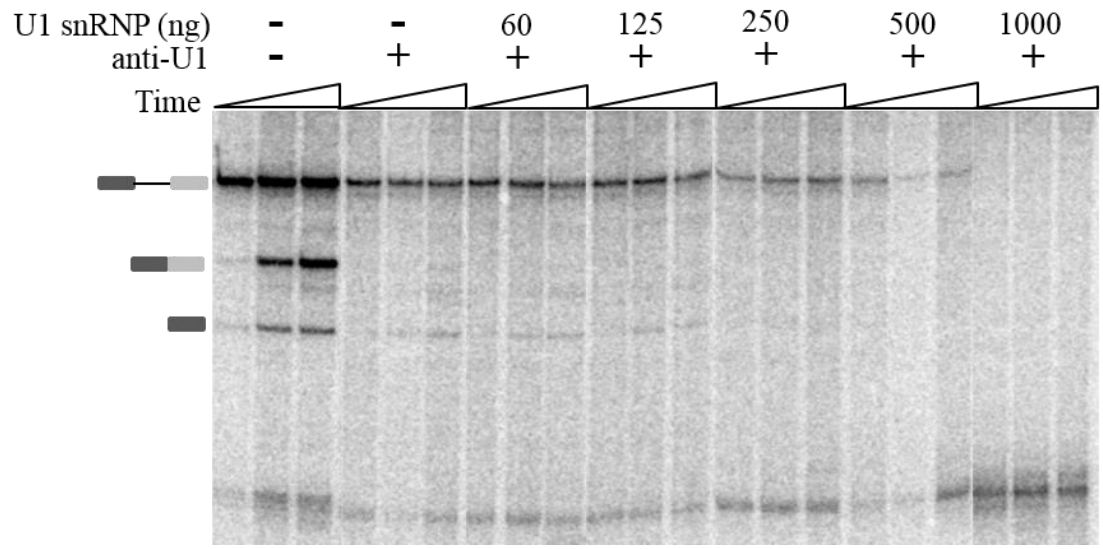


Figure 4.29. The addition of purified U1 snRNP did not restore either transcription or splicing in extracts with blocked endogenous U1 snRNA. Time-course Pol II-TSR assay with purified U1 snRNP. The samples were preincubated with anti-U1 oligonucleotide at the concentration at 5 μM for 15 min for 30°C. The U1 snRNP was added following the incubation for 15 min. After, Bg C DNA template was used. Time points at 30, 60 and 120 min were taken.

To test the exogenous U1 snRNA in the complementation experiments, the U1 snRNA gene was cloned. For that, RNA extracted from HEK293T cells was reverse transcribed by using random primers to convert to cDNA. After PCR amplification, the U1 DNA template was cloned into Zero Blunt TOPO vector. Prior to the addition of the U1 snRNA to Pol II-TSR, transcription was done. These procedures are described in more details in Materials and methods. In transcription and splicing conditions after the endogenous U1 snRNA was inhibited, the exogenous U1 snRNA was added at different concentrations and the samples were pre-incubated for 30 min. This was followed by the addition of CMV-Bg C DNA fragment and further incubation for 90 min. The experiment revealed an inhibitory effect of transcribed U1 snRNA on transcription (Figure 4.30). The attempts to restore splicing with transcribed U1 snRNA in the conventional *in vitro* splicing system with blocked endogenous U1 snRNA were also unsuccessful (Figure 4.31).

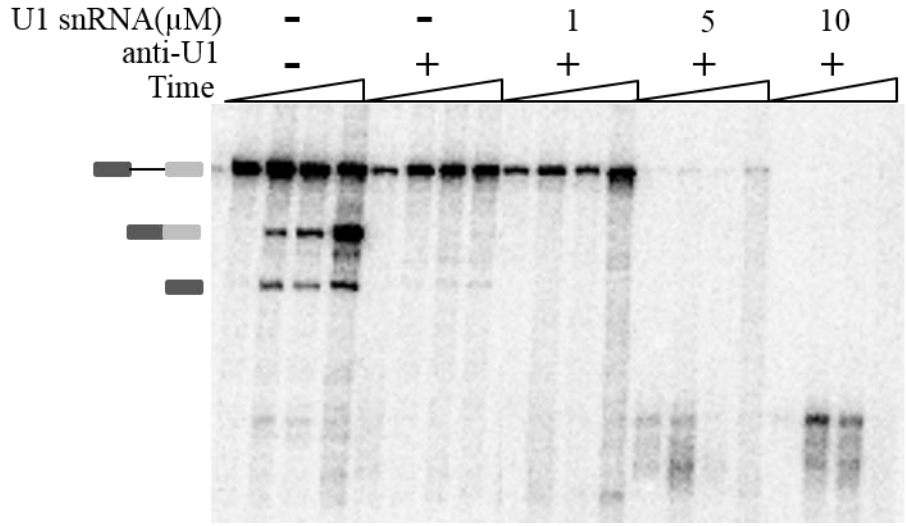


Figure 4.30. Effect of capped transcribed U1 snRNA in Pol II-TSR. The samples were preincubated with the anti-U1 oligonucleotide at the concentration at 5 μM for 15 min at 30 °C. The exogenous U1 snRNA was added at different concentrations following the incubation for 30 min. After, the CMV-Bg C DNA template was used and time points at 15, 30, 60 and 90 min were taken.

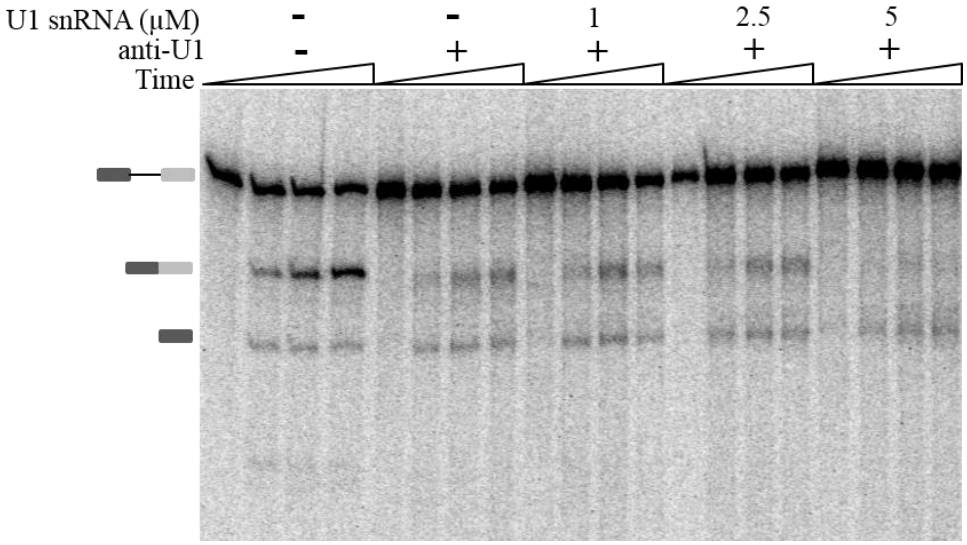
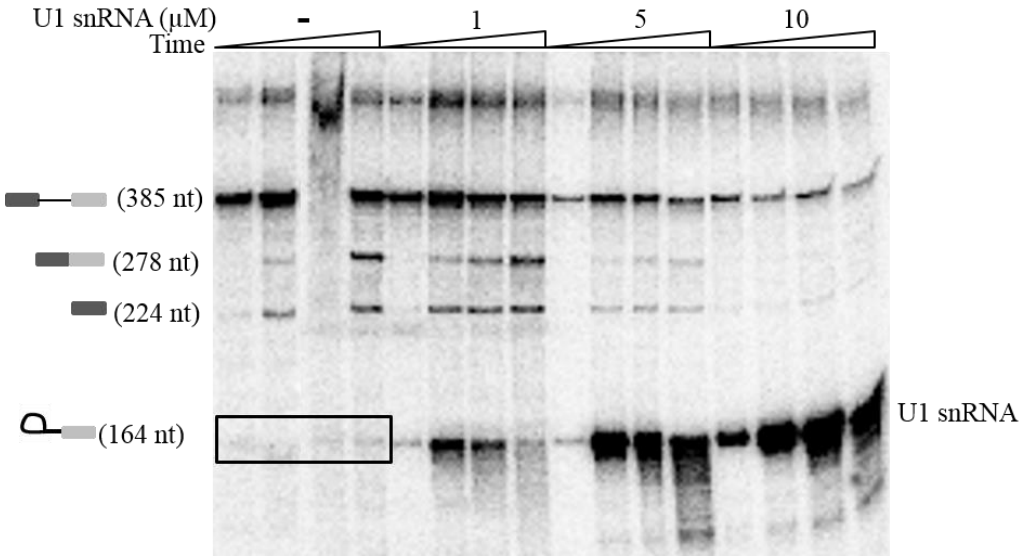


Figure 4.31. Exogenous U1 snRNA does not stimulate splicing in a splicing assay. Effect of transcribed U1 snRNA on splicing in the samples with block endogenous U1 snRNA. The splicing reactions were preincubated with the anti-U1 oligonucleotide at the concentration of 5 μM for 15 min at 30°C. The exogenous U1 snRNA was added at different concentrations following the incubation for 30 min. After, Bg C pre-mRNA was used. Time points of 0, 30, 60 and 90 min were taken.

To test if transcribed U1 snRNA had any side effects on Pol II transcription, a control experiment was done by adding the U1 snRNA only. The concentration of transcribed U1 snRNA was 1 µM, 5 µM and 10 µM. In a dose-dependent manner, an inhibitory effect on transcription was observed with a maximum concentration at 10 µM (Figure 4.32 A). Unexpectedly, a band of small size was detected in the samples with added U1 snRNA and it became bigger as the amount of the U1 snRNA was increased. To investigate whether exogenous RNA has a property to inhibit transcription, RNA extracted from yeast (Thermo Fisher Scientific) was taken for control purpose. The result showed a lack of transcription in a concentration-dependent manner (Figure 4.32 B). This is consistent with the experiment in which the exogenous U1 snRNA was used (Figure 4.32 A). It is concluded that exogenous RNA of specific sequence or heterogeneous RNA species added to the reaction with ongoing transcription negatively affect transcription.

A



B

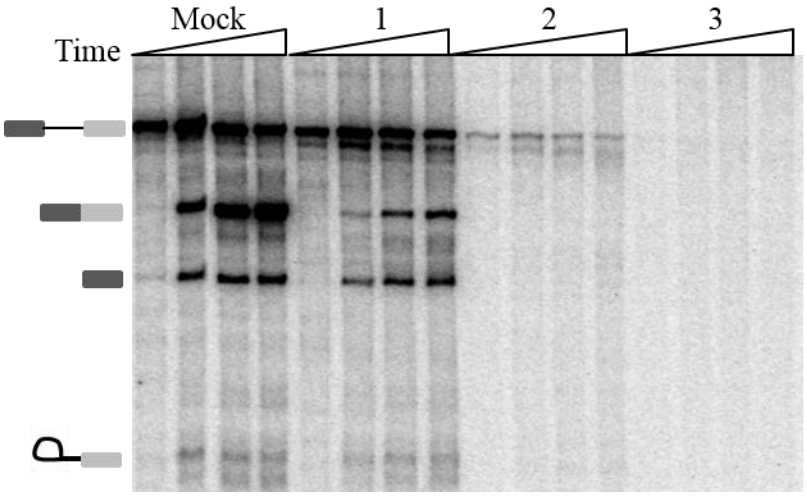


Figure 4.32. Reconstitution of transcription is not achieved when exogenous U1 snRNA is added. (A) *In vitro* Pol II transcription/splicing reactions were performed with transcribed U1 snRNA. The U1 snRNA was transcribed by T7 polymerase *in vitro* and gel purified. The cap analog (m^7GpppG) was not used in *in vitro* T7 transcription assay. The concentration of U1 snRNA was titrated as shown. A band of low size appeared in the samples with transcribed U1 snRNA. The lariat intemediate in the mock sample is shown in the open box. The sizes of each product is represented. (B) Testing yeast RNA in Pol II-TSR assays. (A and B) Time points at 15, 30, 60 and 120 min were taken.

The band in the samples with the exogenous U1 snRNA (164 nt) comigrated with the lariat intemediate (164 nt) in the mock sample. It is suggested that the exogenous U1 snRNA was radiolabelled. It is likely that unlabelled U1 snRNA became radiolabelled in Pol II-TSR, as the samples contained [α-^{32}P]GTP nucleotide. Since [α-^{32}P]GTP could not be internally incorporated into the U1 snRNA, we hypothesised that transcription-independent capping could occur because the U1 snRNA did not contain the cap structure and the sample had an abandon [α-^{32}P]GTP nucleotide. This coincided with an opportunity to test if capping can occur post-transcriptionally and in a transcription-independent manner. To test this, the U1 snRNA was transcribed in the presence of 1 mM chemically synthesised cap analogue (m7GpppG). The experiment showed a remarkable reduction of transcription but the radiolabelling of the U1 snRNA still occurred (Figure 4.33 A). However, the intensity of the U1 snRNA band was lower in comparison to the previous experiment (Figure 4.32 A). Based on the result which showed that about 30% RNA does not contain the cap structure after *in vitro* T7 transcription (Figure 3.15), and the fact that less U1 snRNA was labelled (Figure 4.33 A), it is possible that uncapped species of the U1 snRNA sample could undergo capping. There were several ways to test this. First, if 100% capped U1 snRNA was used in Pol II-TSR the U1 snRNA should not be labelled. Second, capping of U1 snRNA would not be detected if [α-^{32}P]GTP was substituted by any [α-^{32}P]rNTPs. To test the latter, [α-^{32}P]rUTP was used in Pol II-TSR. Interestingly, the exogenous U1 snRNA was still labelled in the samples with ^{32}P-rUTP (Figure 4.33 B). It is suggested that RNA 3' end tailing could occur (Gu et al., 1997) but the basis of these observations remains unclear.

The poor transcription observed in the samples with either the U1 snRNP or U1 snRNA can be explained in two ways. First, low transcription was detected not due to the inhibition of transcription process but rather due to the competition for ^{32}P-rGTP which was incorporated in the exogenous U1 snRNA. Second, factors required for RNA processing and transcription may have been recruited to the U1 snRNA or yeast RNA. In conclusion, an *in vitro* complementation assay cannot be exploited because exogenous RNA has a negative effect on transcript visibility. Also, to investigate effects of exogenous U1 snRNP in Pol II-TSR, NE depleted from the U1 snRNP should be used.

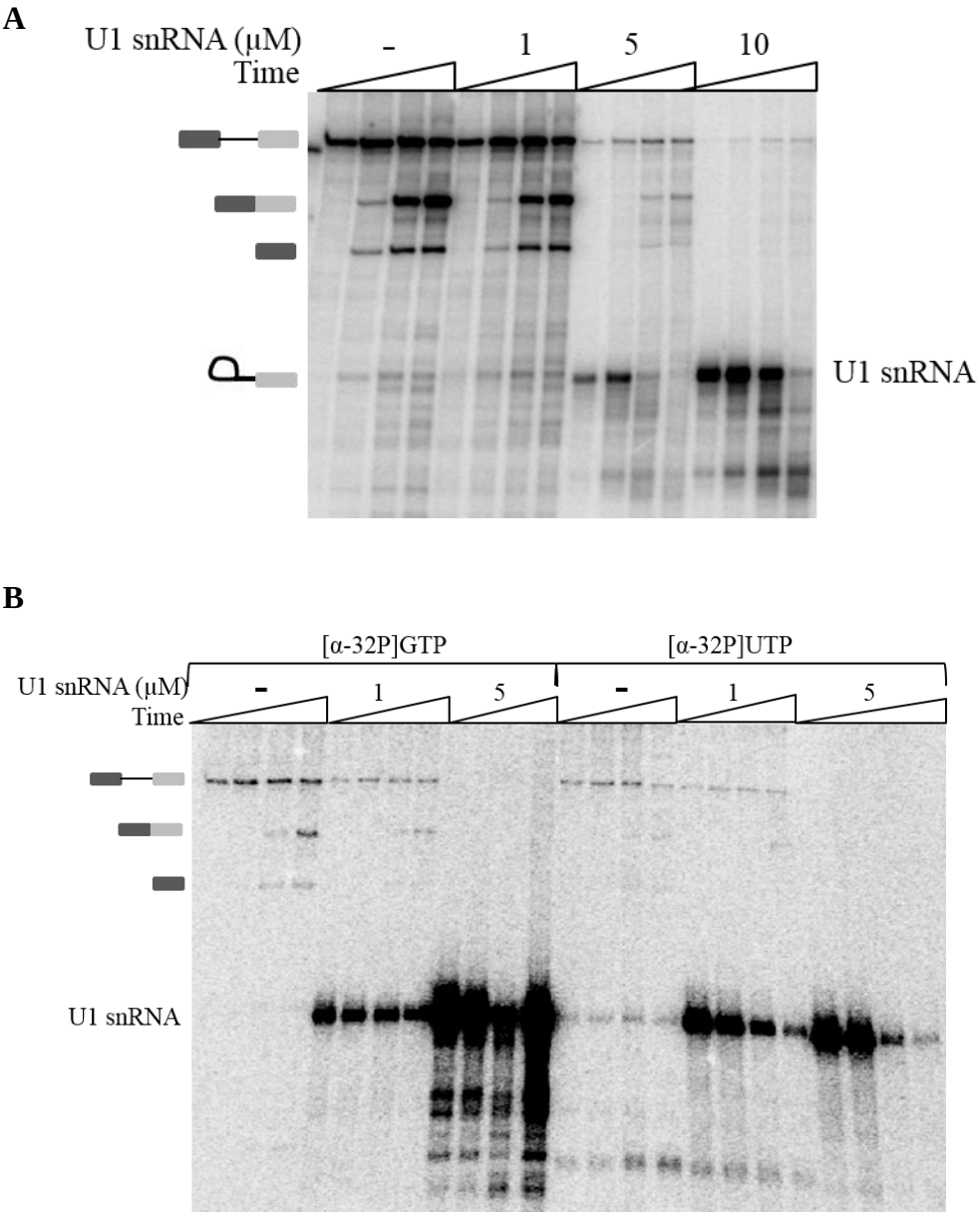


Figure 4.33. The exogenous U1 snRNA becomes radiolabelled regardless which nucleotide is used in Pol II transcription/splicing assay. The radiolabelled nucleotides such as GTP and UTP were used. The samples were preincubated with uncapped U1 snRNA at 1 and 5 μM for 15 min at 30°C. This followed by the addition of Bg C DNA template and radiolabelled nucleotides. Time points at 15, 30, 60 and 120 min were taken.

4.6 Conclusions

The activities of the U1, U2 and U6 snRNPs in Pol II transcription assay was tested by inactivating these factors using a method widely used in conventional splicing assays. The results showed that the level of transcripts was significantly decreased if the 5' end of the U1 snRNA was blocked or degraded. This phenomenon is substrate-independent as transcripts haing different sequence composition were sensitive to the functional inhibition of the U1 snRNA. Even the level of the intronless RNA, GFP, was dependent on the U1 snRNA. Transcripts synthesised in T7-TSR or added to the conventional splicing reaction were also not stable if the U1 snRNA had been blocked. Together the results provide strong evidence that the U1 snRNA exhibits a crucial role in RNA stability in a polymerase-independent manner. Indeed, this was proven by the pulse chase experiments which showed that in Pol II-TSR both Bg C and GFP transcripts underwent degradation in the reactions with blocked U1 snRNA. There was inconsistency in comparable experiments involving inhibition or digestion of the U2 and U6 snRNAs in Pol II-TSR system. However, the result from T7-TSR showed increased level of RNA upon inhibition of the U2 and U6 snRNAs. Under splicing conditions, pre-synthesised transcripts were stable if the U2 and U6 snRNAs were blocked. Therefore, it is suggested that the U1 snRNA is the only snRNP that plays a role in RNA stability in both uncoupled reactions (SR and T7-TSR) and the reaction in which transcription is linked to splicing (Pol II-TSR). Based on these data and the experiments with GFP which lacks any splice sites, it is suggested that the assembly of the complete spliceosome is not required to maintain pre-mRNA stability but only the U1 snRNP is essential for that.

The pathways of RNA degradation were investigated in this study. Two oligonucleotides complementary to the ends of Bg C pre-mRNA were used in SR. The results showed that the U1 snRNA protects uncapped pre-mRNA from degradation by 5' and 3' exonucleases. As the complete protection of RNA by the oligonucleotides was not achieved using capped RNA, it seems likely that the internal cleavage by endonucleases also took place. However, none of the protective oligonucleotides were efficient in Pol II-TSR, suggesting that a different RNA degradation pathway is active in Pol II-TSR compared to SR.

Investigation of the requirements for the 5'SS in transcripts showed that its presence is necessary for maintaining pre-made RNA stable under splicing conditions. However, this was not the case in Pol II-TSR. We found that Bg C pre-mRNA and its construct having mutated 5'SS were both stable in the reaction with ongoing Pol II transcription. It is suggested that some cryptic splice sites could be bound by the U1 snRNP during transcription mediating the satiability of transcripts which lack the consensus 5'SS. Possibly the U1 snRNA stabilizes newly produced transcripts by Pol II during transcription. These data allows to suggest that functional coupling between transcription and splicing occurs in *in vitro* Pol II-TSR. However, a different pattern was observed using the intronless GFP construct. It was found that the GFP-U1(3') transcript having the consensus 5'SS at the 3'end was more stable than GFP wt RNA. In conclusion, the U1 snRNA has a property to protect RNA from degradation but there are two distinct mechanisms. The mechanisms depend on the type of the reaction regarding the presence or absence of ongoing Pol II transcription.

5 DISCUSSION

5.1 Effects of Pol II transcription on splicing

5.2 Pol II-TSR recapitulates splicing pattern observed *in vivo*

5.3 DRB affects splicing in Pol II-TSR but not either in T7-TSR or SR

5.4 Role of U1 snRNP in RNA stability

5.1 Effects of Pol II transcription on splicing

Early studies established conditions in which transcription by Pol II and splicing occurred in one reaction (Ghosh and Garcia-Blanco, 2000; Das et al., 2006; Hicks et al., 2006; Folco and Reed, 2014). Previously splicing efficiency was compared between Pol II-TSR and conventional SR assays. However, the overall efficiency of splicing was low in the conventional SR assay (Ghosh and Garcia-Blanco 2000; Das et al., 2006). Despite this, splicing yield was still compared between both assays, concluding that Pol II transcription enhances splicing efficiency, kinetics and spliceosome assembly. Previous attempts to establish transcription/splicing system dependent on T7 RNA polymerase (T7-TSR) were unsuccessful (Ghosh and Garcia-Blanco 2000; Das et al., 2006; Hicks et al., 2006). It was shown that T7 transcripts were not stable and degraded (Hicks et al., 2006). Therefore, it was concluded that splicing was faster, the yield was increased and spliceosome assembly was accelerated in Pol II-TSR in comparison to T7-TSR (Ghosh and Garcia-Blanco 2000; Das et al., 2006). Based on these observations the authors concluded that transcription was functionally coupled to splicing in Pol II-TSR.

In this study an *in vitr*o Pol II transcription/splicing assay was developed using HeLa NE. Initially, poor splicing in T7-TSR was observed in this study. The results showed that the level of transcripts generated by T7 Pol was 3-5-fold higher in comparison to that for Pol II counterparts (Figure 3.14). Possibly due to highly productive transcription by T7 Pol, splicing factors could be saturated due to the abundancy of pre-mRNA. Therefore, the spliceosome staling at complex B observed in the experiments (Figure 3.13) could occurred leading to transcript degradation (Figure 3.14). To achieve similar levels of RNA in Pol II-TSR and T7-TSR assays, the amount of T7-Bg C DNA template was varied. When similar amounts of transcripts were generated in Pol II-TSR and T7-TSR, the pre-mRNAs spliced with almost equal efficiency (Figure 3.17). It is possible that splicing factors were titrated out by the abundant T7 pre-mRNAs and the spliceosome did not progress through step 1 to complex C. Consequently transcripts underwent degradation. This is the first report of almost equal splicing yields in both transcription/splicing assays. These result differ from previous observations (Das et al., 2006; Hicks et al., 2006) in which similar amounts of pre-mRNA were produced by two polymerases but splicing was poor in the reaction with T7 Pol. In this study pre-mRNA

transcribed in T7-TSR degraded only when RNA concentration was significantly higher than that in Pol II-TSR. The results from this work showed clearly that the splicing efficiency was similar regardless which polymerase generated pre-mRNAs. The kinetic parameters of splicing in the presence of Pol II transcription were not tested in comparison to T7-TSR in this study. This can be analysed in future.

The Pol II-TSR system was employed to address questions regarding the ability of Pol II to influence spliceosome assembly. The results show that transcripts are visualized at 3 min (Figure 3.11) while splicing complexes appear at 15 min (Figure 3.23 A). Also, it is clear that the splicing complexes are accumulating with time suggesting a synchronized pattern of pre-mRNAs to undergo splicing. Therefore, it is suggested that splicing occurs post-transcriptionally. This hypothesis is supported by the data obtained from the experiment in which DNA-protein complexes were treated with RNase A (Figure 3.28 A). It was found that the migration of DNA-protein complexes was not changed upon RNase A treatment, suggesting that the pre-mRNA was degraded being released from Pol II. If it was not the case and the spliceosome was assembled on the pre-mRNA tethered to Pol II then RNA degradation would make an effect on the mobility of DNA-protein complexes as the spliceosome is a multi megadalton complex (reviewed by Will and Lührmann, 2011).

Early studies showed that complex H is not a functional precursor of the spliceosome (Michaud and Reed 1991, 1993). The work by Das et al. (2006) showed that complex H was formed on pre-made pre-mRNA and the nascent T7 transcripts, but it was not detected in reactions in which transcription was driven by Pol II. Based on these results the authors (Das et al., 2006) concluded that Pol II is able to direct immediate spliceosome assembly rather than stalling at complex H. Therefore, it was concluded that Pol II transcription is functionally coupled to splicing *in vitro*. Since the CTD has a property to interact with splicing factors (reviewed by Bentley, 2014), it was suggested that they might be recruited to the pre-mRNA as it emerges from Pol II exit pore. The data obtained in this study did not show the presence of complex H in either T7 or Pol II transcription/splicing assays (Figure 3.13 and Figure 3.23). In contrast, complexes A and B were formed on the nascent T7 transcripts (Figure 3.13). The results obtained in this study suggest that as soon as RNA appears from either Pol II or T7 Pol it might be bound by splicing factors which then mediate spliceosome assembly. Since the effect is

polymerase-independent, the CTD cannot be implicated in functional coupling. Complex H may be an artefact of the conventional *in vitro* splicing assay. Its presence can be explained as the abundant hnRNP proteins immediately interact with pre-made pre-mRNA (Bennett at al., 1992). It is possible that this happens because transcripts are added to the NE in excess and all at once.

In summary, the results showed that splicing yield (Figure 3.10), kinetics (Figure 3.11) and spliceosome assembly (Figure 3.23) were not improved by Pol II. Therefore, based on these parameters, Pol II does not have a property to stimulate splicing *in vitro*. These data is in line with some previous studies (Lazarev and Manley, 2007; Hicks et al., 2006) but contrary to others (Ghosh and Garcia-Blanco 2000; Das et al., 2006).

5.2 Pol II-TSR recapitulates splicing pattern observed *in vivo*

To determine whether transcription has any effects on pre-mRNA splicing, alternatively spliced substrates were tested. One of them was a CD44v5 construct consisting of three exons and two introns. Two mRNA isoforms (inclusion and skipping) were generated in SR (Figure 3.18 B) while in Pol II-TSR just one mRNA isoform (skipping) was detected. This implies that in the reaction with ongoing transcription the distal 3'SS was preferred to the proximal 3'SS. Interestingly, this system recapitulated the splicing profile obtained from an *in vivo* study (Figure 3.18 D and E). Therefore, it is concluded that the change in the alternative splice site choice is transcription-dependent. If it was a parallel reaction, in which splicing is initiated after the transcript release, then the splicing pattern would be similar to the reaction in which pre-mRNA is added to the extract (SR). Based on the recruitment model it is suggested that splicing enhancers were not recruited from the CTD to the nascent pre-mRNA. Alternatively, the local concentration of splicing inhibitors in the transcription site might be increased (de la Mata et al., 2006). Therefore, the alternative exon was not incorporated into the mRNA. From the kinetic model of co-transcriptional splicing, slow transcription would favour inclusion. If the rate of transcription were fast and SSs of the variable exon 5 were not recognized by splicing enhancers, then the pre-mRNA would be spliced producing only the skipping isoform. Since the two models of co-transcriptional coupling are not mutually exclusive (reviewed by Bentley, 2014), the competition of splicing enhancers and repressors could occur during fast transcription, leading to exon skipping. To test

this hypothesis recombinant Sam68 protein could be added to Pol II-TSR and SR assays.

Alternatively, the absence of the inclusion isoform can be explained as the secondary structure of the newly transcribed pre-mRNA may be different to that of pre-made pre-mRNA leading to different splicing patterns between the two assays. Since, in this study T7-TSR was optimized, splicing pattern of CD44v5 in T7-TSR could be compared with Pol II-TSR in future.

The result observed from the experiment with a Bcl-X substrate harbouring two alternative 5' SSs showed similar splicing patterns between Pol II-TSR and SR (Figure 3.18). In particular, Bcl-X precursor was able to produce two possible mRNA isoforms (Xl and Xs) in both assays. Since the ratio between the short and long mRNA isoforms was similar not only between the two *in vitro* systems (Figure 3.18) but *in vivo* (Weldon et al, 2016), it is likely that Bcl-X splicing is post-transcriptional.

5.3 DRB affects splicing in Pol II-TSR but not either in T7-TSR or SR

It is known that the CTD of Pol II RPB1 domain serves as a binding platform for numerous proteins involved in RNA processing, including splicing (Chapter 1). Early studies showed that phosphorylated recombinant CTD enhanced spliceosome assembly even in the absence of transcription (Hirose et al., 1999; Zeng and Berget, 2000). Later, contradictory data showed that inhibition of phosphorylation of the CTD by DRB (an inhibitor of elongation that inhibits phosphorylation of P-TEFb) *in vivo* had no effect on splicing uncoupled from transcription (Bird et al., 2004) but decreased co-transcriptional splicing. Therefore, it is not clear whether the phosphorylated CTD affects splicing when it is linked to transcription or is also able to influence post-transcriptional splicing. To test this, transcription elongation inhibitor, DRB, was included in our optimized *in vitro* systems.

Decreased splicing efficiency upon DRB treatment was detected in Pol II-TSR assay (Figure 3.19) but not T7-TSR or SR. There are two ways in which the phosphorylated CTD could influence co-transcriptional splicing. First, via the rate of transcriptional elongation and, second, by affecting the recruitment of splicing factors. Based on

previous study which showed that the association of U2AF65 with the CTD is mediated through phosphorylated Ser2 (Gu et al., 2013), it is likely that U2AF65 is less recruited to the transcribing pre-mRNA when the CTD is hypophosphorylated.

These results disagree with data on the abilities of the recombinant CTD to influence post-transcriptional splicing (Hirose et al., 1999; Zeng and Berget, 2000) and they are in line with the observations obtained by Bird et al. (2004). To test functional coupling, biotinylated CMV-Bg C DNA could be used in Pol II-TSR treated with DRB. After pulling-down the DNA template, the samples could be used for mass spectrometry analysis. Since DRB is able to inhibit activities of CDK7 and CDK9 which, in turn, phosphorylate Ser5 and Ser2 residues, respectively (Feaver et al., 1991; Lu et al., 1992; reviewed by Price, 2000), at this stage it is not possible to conclude which residues are important for splicing regulation. For this western blotting using antibodies against specific residues of the CTD could be done in the future.

5.4 Role of U1 snRNP in RNA stability

A number of transcription factors have been shown to associate with the U1 snRNP, including initiation factors TFIIH and TAF15 (Kwek et al., 2002; Jobert et al., 2009), and the elongation factor TAT-SF1 (Fong and Zhou, 2001). However, these reports described only associations and no evidence was shown to support a functional activity of the U1 snRNP in transcription initiation and elongation. The U1 snRNP was also shown to associate directly and indirectly with Pol II (Kim et al., 1997; Tian, 2001; Yu and Reed, 2015). Disruption of the essential base-pairing between the U1 snRNA and the 5'SS of pre-mRNA led to inhibition of Pol II transcription *in vitro* and loss of Pol II binding to the CMV promoter (Tian, 2001). Based on these results the author concluded that the U1 snRNP participates in the pre-initiation complex (PIC) assembly (Tian, 2001). However, contradictory ChiP data demonstrated that Pol II occupancy was unchanged on genes upon inactivation of the U1 snRNA using the same method (Davidson et al., 2014). Furthermore, *in vivo* studies showed that the levels of transcripts were unchanged if mutated U1 snRNA interacts with intronic sequence by complementarity (Lu et al., 1990; Dal Mas et al., 2015). However, these data can be interpreted as via a direct interaction with pre-mRNA, the U1 snRNP can stimulate transcription initiation by enhancing PIC assembly as it was suggested by Damgaard et

al. (2008). Therefore, it is not clear whether the U1 snRNP regulates transcription or it protects RNA from degradation.

In this study the effect of U1 snRNP on Pol II transcription was investigated using the *in vitro* Pol II transcription/splicing assay. Possible functions of the U2 and U6 snRNPs in transcription were also explored. In previous studies the effects of these core spliceosomal components on Pol II transcription were studied using antisense oligonucleotides which interact with regions of snRNAs that are important at different stages of splicing (Tian, 2001; Koga et al., 2015). The same method was exploited in this study. It was found that the level of transcripts is significantly decreased if the 5' end of the U1 snRNA is blocked or degraded (Figure 4.1; Figure 4.2). This phenomenon is substrate-independent as all constructs tested, regardless of the presence of an intron, were sensitive to the functional inhibition of the U1 snRNA. At this stage our results are consistent with previous data (Lu et al., 1990; Tian, 2001; Yu and Reed, 2015; Koga et al., 2015; Dal Mas et al., 2015). However, we found that transcripts generated in T7-TSR or added to splicing reactions undergo degradation if the U1 snRNA but not U2 or U6 snRNAs have been blocked (Figure 4.6 and Figure 4.7). Since the levels of transcripts were decreased in polymerase- and transcription-independent manners, the results reported previously (Tian, 2001; Lu et al., 1990) could be attributed to effects on RNA stability rather than the level of transcription. The pulse chase experiments of Pol II-TSR (Figure 4.8 and Figure 4.25) support the notion that the U1 snRNA is important for RNA surveillance. However, our results do not disprove a possible function of the U1 snRNP in Pol II transcription regulation. It is possible that the U1 snRNP could participate in transcription regulation but it is difficult to estimate how important it is. Since we detect severe degradation upon inhibition of the U1 snRNP, this might mask an additional role of U1 in transcription regulation. Based on our results it is clear that the importance of the U1 snRNP in RNA stability is greater than any role in transcription regulation.

It is known that the U1 snRNP has a property to protect transcripts from premature cleavage and polyadenylation (Koga et al., 2010; reviewed by Guiro and D O'Reilly, 2015). However, in our study the U1 snRNP cannot be implicated in suppression of premature cleavage and polyadenylation as any poly(A) sites were found within the substrates tested.

Under splicing conditions the stability of both capped and uncapped transcripts was dependent on the presence of functional U1 snRNP (Figure 4.11 and Figure 4.13). The pathways of RNA degradation were investigated. The experiments with protective oligonucleotides complementary to the terminal parts of Bg C pre-mRNA showed that uncapped transcripts are subject to degradation through the free ends by 5' and 3'exonucleases if the U1 snRNP has been blocked (Figure 4.13 and Figure 4.14). When the 3' end of capped pre-mRNA was protected, only one third of RNA was rescued (Figure 4.12). Since two third of RNA was still degraded it is suggested that capped RNA could be cleaved by endonucleases also.

In the Pol II-TSR the protective oligonucleotides did not protect RNA from degradation in the samples with blocked U1 snRNA (Figure 4.9). Thus, it is unlikely that the RNA was degraded through the 5' and 3' ends. Another explanation is that RNA could be degraded during transcription. Previously it was shown that splicing-defective pre-mRNAs undergo degradation during transcription (Davidson et al., 2012). It was found that one of the 5' exonucleases, Xrn2, which physically associates with Pol II, was responsible for degrading such transcripts during transcription. Based on these data the model (Figure 5.1) suggests that the pre-mRNA could be degraded co-transcriptionally by Xrn2 and, therefore, the protective oligonucleotides would be unable to restore RNA levels. However, it is not known how RNA becomes accessible for 5' exonucleases. There is a possibility that endonucleases might cleave RNA internally, creating free 5' ends for Xrn2 digestion. Alternatively, capping might not occur, resulting in open 5' termini of the RNA. Since Xrn2 seems a more likely candidate, its responsibility for RNA degradation could be tested *in vivo* by using RNA interference to target Xrn2 with further analysis of RNA levels with and without inhibition of the U1 snRNA by complementary morpholino oligonucleotides. It is expected that RNA would be accumulated upon knocking down this nuclease.

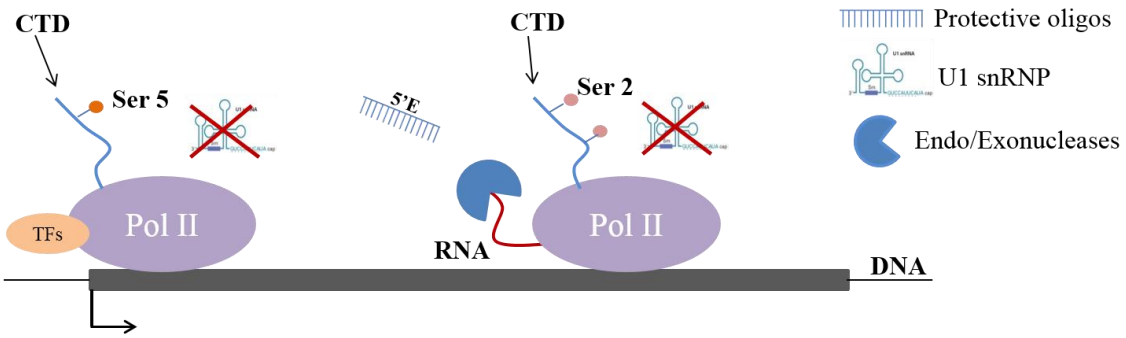


Figure 5.1 Schematic representation of a co-transcriptional RNA degradation pathway under conditions when the U1 snRNA is not functional. The protective oligonucleotides complementary to Bg C RNA ends (5'E and 3'E) are not able to anneal to RNA because it is not produced. Ser5 and Ser2 represent phosphorylation modifications of the CTD during transcription initiation and elongation, respectively. Inhibition of U1 snRNA is depicted.

The effects of an intron on gene expression have been investigated previously showing that transcription efficiency is increased if genes contain introns (Brinster et al., 1988; Choi et al., 1991; Fong and Zhou, 2001; Furger et al., 2002). However, contradictory studies have shown that the 5'SS but not the 3'SSs, branch point or polypyrimidine tract is essential for gene expression (Barret et at., 1995; Damgaard et al., 2008). Other studies provided evidence that the interactions between mutated U1 snRNA and complementary sequences led to improved transcription levels of spliced and unspliced RNAs (Lu et al., 1990; Kammler et al., 2001; Dal Mas et al., 2015). It was suggested that the interactions between the U1 snRNP and pre-mRNAs through the 5'SSs stabilize transcripts in a splicing independent-manner. However, since these studies have not provided any evidence for that, the data could also be interpreted as a stimulatory effect of the 5'SS on transcription (Damgaard et al., 2008). This can be supported by findings that 5'SS mutated substrates are not subject to degradation (Damgaard et al., 2008) and Pol II activity is increased over intron-containing genes (Furger et al., 2002). Since these results were obtained from *in vivo* studies and it is known that the majority of genes are expressed co-transcriptionally, it is not entirely clear whether the U1 snRNP-5'SS binding is essential for protecting the RNA from nucleases or it is an important contributor to transcription activation. To determine whether the U1 snRNP participates in transcription regulation or it is required for RNA stability, an *in vitro* system in which splicing is separated from transcription was necessary to use. We found that RNA

stability depends on the presence of a consensus 5'SS in transcripts in SR (Figure 4.19; Figure 5.2 A) showing that a direct interaction between the U1 snRNP and pre-mRNA is essential to obtain steady-state levels of RNA in a transcription-independent manner. We found that pre-made RNA is stable if the position of the 5'SS is in the middle or at the 3' end. This fits previous observations which showed that the efficient gene expression is 5'SS-dependent (Barret et al., 1995; Damgaard et al., 2008). Based on these observations it is possible to suggest a model of how the U1 snRNP stabilizes RNA (Figure 5.2 B). This model suggests that regardless of the position of the 5'SS, once the U1 snRNP binding has occurred, it mediates the recruitment of RNA binding proteins to RNA which propagate across RNA. As a result, RNA is protected from degradation because it is inaccessible to nucleases as RNA binding proteins serve as a scaffold making RNA stable.

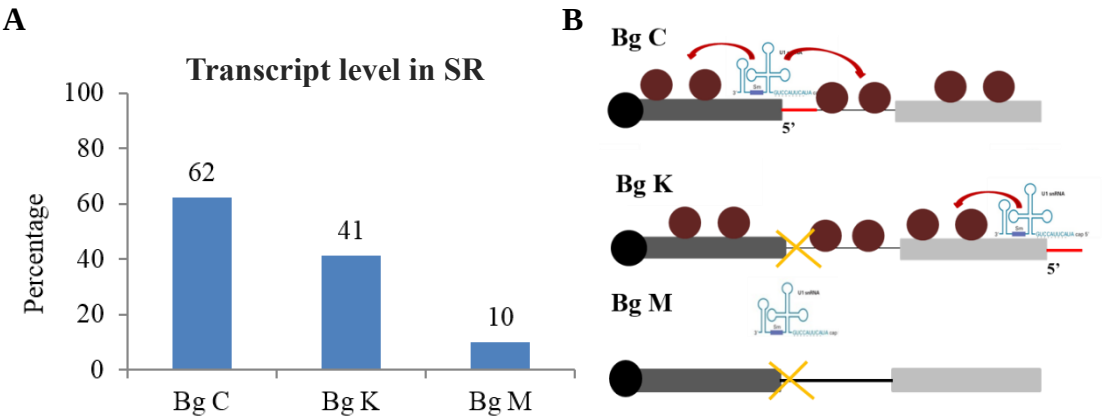
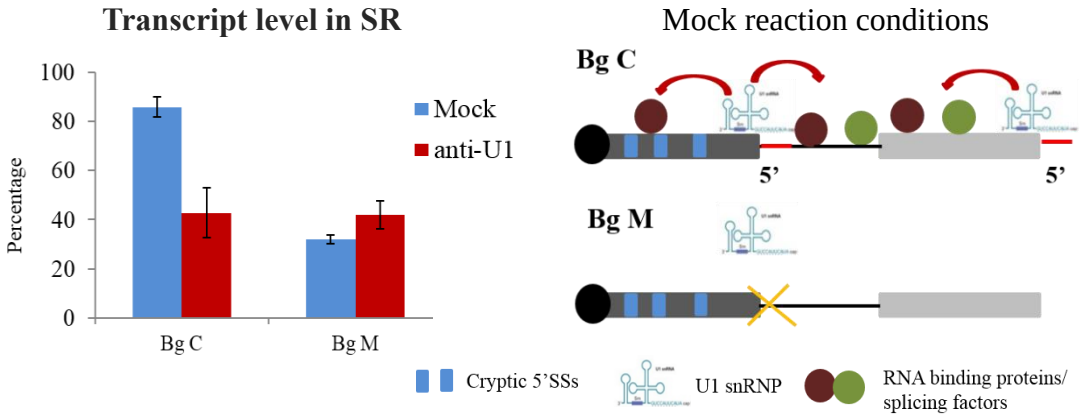


Figure 5.2. The presence of the 5'SS regardless of its position is crucial for RNA stability in the conventional splicing assay. (A) Quantification of transcript levels in SR. (B) Illustration shows that the U1 snRNA mediates the recruitment of RNA binding proteins to RNA upon binding to the consensus 5'SS in reactions in which splicing is impaired form transcription. Bg M pre-mRNA contains mutated 5'SS (yellow lines) leading to inability of the U1 snRNP to associate with the RNA. Consequently, the Bg M RNA becomes unprotected and this leads to degradation. The 5'SS is shown as a red line. RNA binding proteins are represented as coloured circles. The cap structure of RNA is illustrated as a black circle.

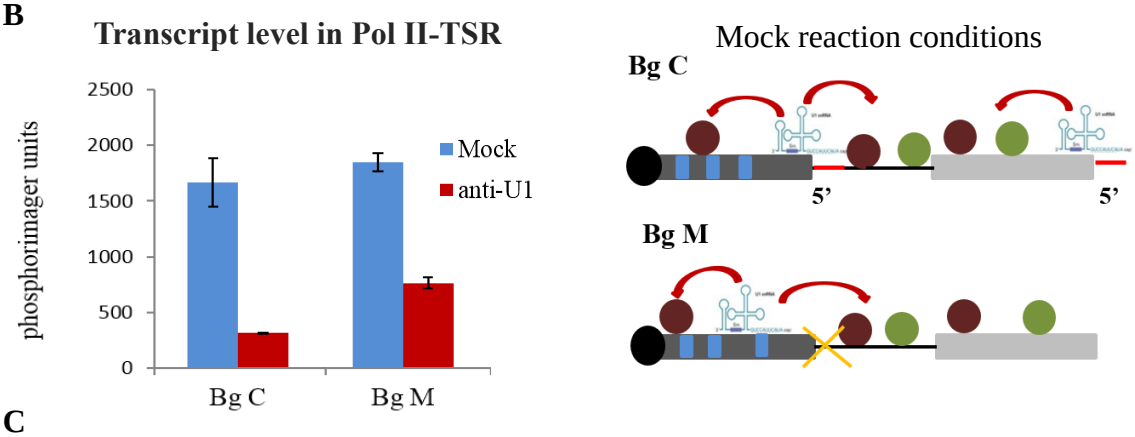
There was a striking difference between RNA levels of Bg C and its mutant version, Bg M, which lacks the consensus 5'SSs in Pol II-TSR and SR (Figure 5.3 A and B). In particular, under splicing conditions Bg M pre-mRNA is not stable in comparison to Bg C (Figure 5.3 A), but both of these transcripts exhibit similar levels in Pol II-TSR (Figure 5.3 B). Based on the results which show that at least one cryptic 5'SS located in the 5' exon can be recognized by the spliceosome (Figure 4.21 A) as a splicing product is produced in Pol II-TSR with Bg M, it is possible that the stability of RNA is mediated by the interaction between the U1 snRNP and this cryptic site. Although the same cryptic site is recognized in SR too (Figure 4.19), this interaction is probably not strong enough to stabilize the Bg M pre-mRNA. In contrast to the reaction in which splicing is separated from transcription (Figure 5.3 A), the scaffolding model suggests that a strong interaction between the U1 snRNP and a cryptic 5'SS of Bg M occurs during transcription, leading to similar RNA levels between Bg C and Bg M (Figure 5.3 B). In contrast, only strong base-pairing between the U1 snRNA and pre-mRNA stabilizes RNA under SR (Figure 5.3 A).

In addition, as with Bg M, GFP wt RNA is not sensitive to inhibition of the U1 snRNP under splicing conditions (Figure 5.3 C) while it is less stable if the U1 snRNP has been blocked in Pol II-TSR (Figure 5.3 D). This fits previous data showing that the U1 snRNP is preloaded onto transcripts regardless of presence of introns (Spiluttini et at., 2010; Brody et al., 2011). Since the GFP RNA contains abundant cryptic 5'SS spread across the whole length (Figure 4.23), the model suggests that through the base pairing to cryptic 5'SSs the U1 snRNA stabilizes RNA via the recruitment of RNA binding proteins (Figure 5.3 D). If it was not the case then GFP wt RNA would not be sensitive to the functional inhibition of the U1 snRNA. To prove this an immunoprecipitation assay with and without anti-U1 oligonucleotide can be done to pull down the DNA with further analysis for the presence of the U1 snRNP.

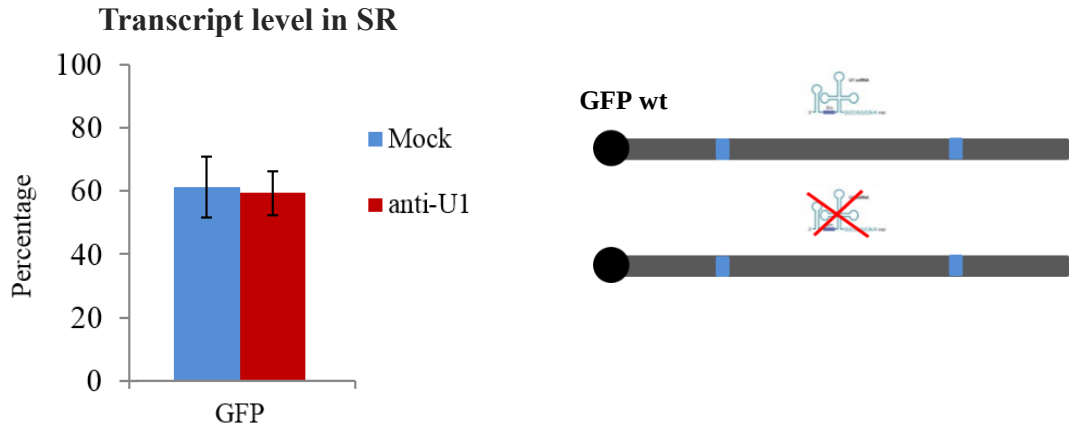
A


Transcript level in SR
Percentage
100
80
60
40
20
0
Bg C
Bg M
Mock
anti-U1
Mock reaction conditions
Bg C
5'
5'
Bg M
Cryptic 5'SSs
U1 snRNP
RNA binding proteins/
splicing factors

B


Transcript level in Pol II-TSR
phosphorimager units
2500
2000
1500
1000
500
0
Bg C
Bg M
Mock
anti-U1
Mock reaction conditions
Bg C
5'
5'
Bg M

C


Transcript level in SR
Percentage
100
80
60
40
20
0
GFP
Mock
anti-U1
GFP wt

D

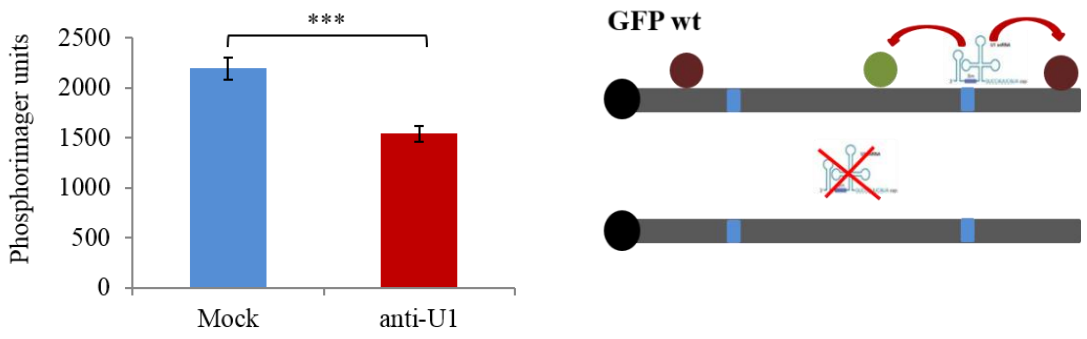

GFP transcript level in Pol II-TSR
Phosphorimager units
2500
2000
1500
1000
500
0
***
Mock
anti-U1
GFP wt

Figure 5.3. Suggested models how the U1 snRNP protects intron-containing and intronless RNAs from degradation in SR and Pol II-TSR. (A) Under splicing conditions the stability of RNA is 5'SS-dependent, possibly mediated via the recruitment of proteins to RNA by the U1 snRNP. Bg M pre-mRNA contains mutated 5'SS (yellow lines). The position of the 5'SS is shown as a red line. The 5'SS is shown as a red line. RNA binding proteins are represented as coloured circles. The cap structure of RNA is illustrated as a black circle. (B) In Pol II-TSR Bg M RNA lacking the consensus 5'SS in the intron is stable to the same extent as Bg C which has 2 sites. (C) GFP RNA is not stable under splicing conditions regardless of the presence of functional U1 snRNP. (B and D) The stability of RNA in Pol II-TSR is U1-dependent suggesting that the U1 snRNP binds to cryptic 5'SSs recruiting RNA binding proteins to RNAs.

In Pol II-TSR the level of GFP RNA which contains the consensus 5'SS at the 3' end (GFP-U1(3')) is increased compared to GFP wt (Figure 5.4). These data fit our previous results (Figure 5.3) and in is in line with *in vivo* data (Barret et al., 1995; Damgaard et al., 2008) showing that the 5'SS itself is sufficient for stable RNA accumulation. The stability of pre-made GFP-U1(3') RNA could be tested under splicing conditions in future studies. Based on our observations and model (Figure 5.2), it is expected that this transcript will be more stable in comparison to the wt.

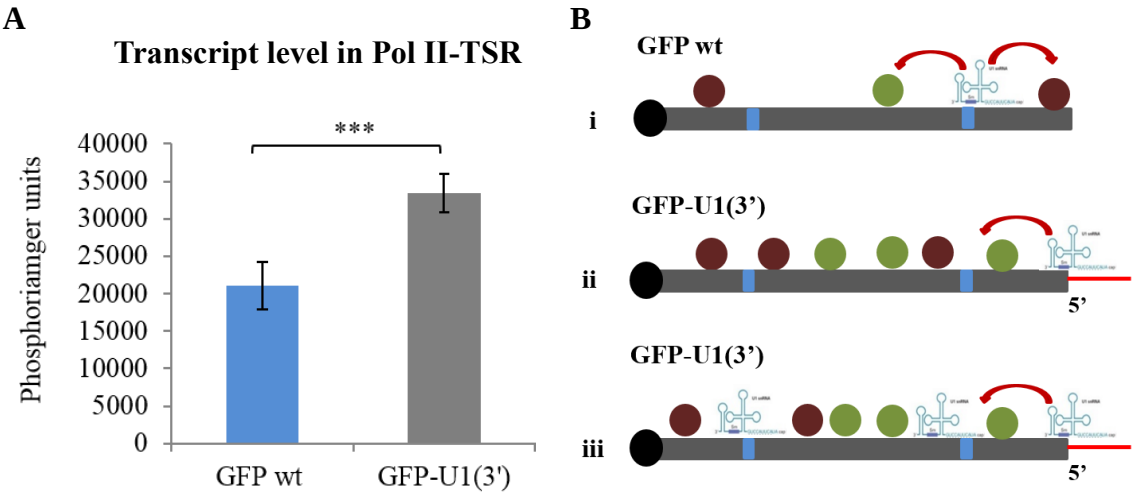


Figure 5.4. The consensus 5'SS stabilizes intronless RNA. (A) Quantification of transcript levels. Cryptic 5'SSs are represented in blue boxes. The consensus 5'SS is shown as a red line. RNA binding proteins are represented as coloured circles. The cap structure of RNA is illustrated as a black circle. (B) Illustration showing that the GFP wt RNA is stable when the U1 snRNP binds to cryptic 5'SSs (i). The stability of GFP RNA is increased if the consensus 5'SS is available. Upon binding to the "strong" 5'SS the U1 snRNP promotes the recruitment of RNA binding proteins (ii) and possibly other U1 snRNP (iii).

Due to the differences between RNA levels in SR and Pol II-TSR, and based on the facts that the CTD is a binding platform for numerous splicing factors (Greenleaf, 1993; reviewed by Bentley, 2014), it is possible that the U1 snRNP is recruited from the CTD to cryptic sites during transcription (Figure 5.3 B and D). This implies that the U1 snRNA stabilizes newly produced transcripts by Pol II in a transcription-dependent manner and not only the functional 5'SS but cryptic sites are required for that. Based on evidence that the U1 snRNP interacts with SR proteins (Eperon et al., 1993; Kohtz et al., 1994; Roca et al., 2013) and the facts that these splicing factors directly associate with Pol II transcription machinery (Vincent et al., 1996; Mortillaro et l., 1996; Kim et al., 1997; Chabot et al., 1995; Lin et al., 2008; Ji et al., 2013; Das et al., 2007), it is possible to speculate on mechanisms by which the U1 snRNP stabilizes RNA. It is possible that the U1 snRNP and its partners such as SR proteins (Eperon et al., 2000; Roca et al., 2013; Das et al., 2007; Pozzoli et al., 2004; Spiluttini 2010) are recruited to the pre-mRNA as soon as it emerges from the exit pore of Pol II enabling transcripts to be protected from nucleases and undergo RNA maturation processing (Figure 5.5). Based on the experiments with the GFP wt construct which lacks any splice sites, it is likely that assembly of the complete spliceosome is not required, but only U1 snRNP binding.

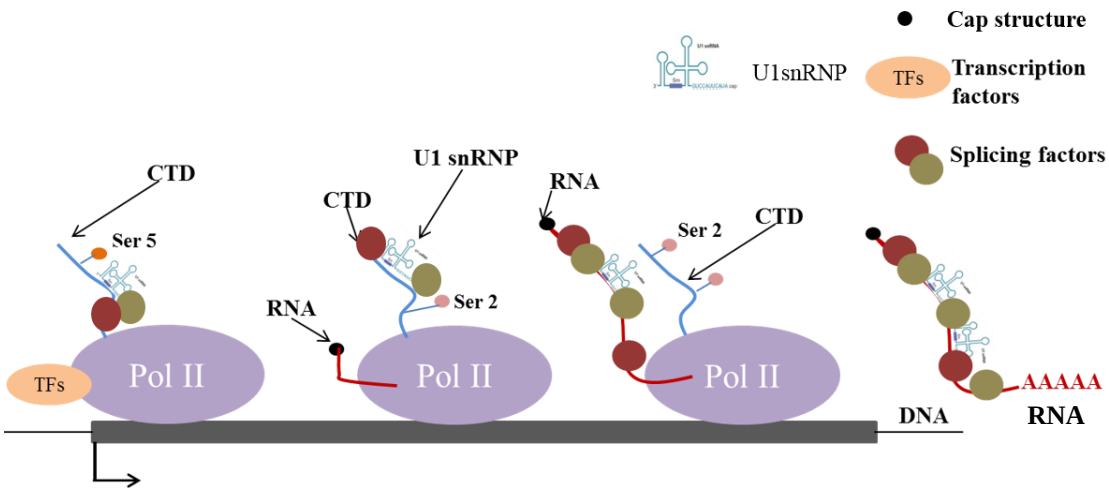


Figure 5.5. Schematic representation of RNA binding proteins and U1 snRNP recruitment to transcripts that emerge from the exit pore of Pol II in a 5'SS-independent manner.

APPENDIX

CD44v5 sequence

The β-globin sequence is shown in italic. The exon 5 of CD44 is shown in grey.

GGGCTGCTGGTTGTCTACCCATGGACCCAGAGGTTCTTCGAGTCCTTTGGGGAC
CTGTCCTCTGCAAATGCTGTTATGAACAATCCTAAGGTGAAGGCTCATGGCAAGAA
GGTGCTGGCTGCCTTCAGTGAGGGTCTGAGTCACCTGGACAACCTCAAAGGCAC
CTTTGCTAAGCTGAGTGAACTGCACTGTGACAAGCTGCACGTGGATCCTGAGAAC
TTCAGGGTGAGTTTGGGGACCCTGATTGTTCTTTCTTTTTCGCTATTGTAAAATTCA
*TGTTATATGGTCGAC*AGCCAACAGCCCTACAAATGTTAGTCATGTATGTGTATC
ACGTATGATATCTACTCCTCGTGTAGGAACAGATGTATTTACAAACTGTCCTGT
GTCTGTCTCTTAGGTGATGAATACAGCTCACTGACTTGTTGTGTGTCCATCCC
TTTTCTGTCCACTTATAGATGTACTCCAACACAGTTGCTTTAAGTGACGAGTT
CATCGCACTGCAGCCATTGTCGAGGGCAAGGAATCAGCTGAGCAAACTTAGG
TAAGAAGGATGATGCACTGAAGTTGTGGTACCCAACTTATCAGAGATGTTGT
GTAGCTGTTAGCTACACCCTTTTGTGTCTGCACAGTAAATTTATCTGCAGAGC
CTGAAAGAGGACATTTTTAAAAATAAGGAAGAAAGATGCAAGTTGTCAGATT
TTCCAACAGACAAGAATGCCTGCAAAGTAGATGTGAGACACTGGTTTTGTGT
TAATGGTTCCTAAGGGT*GTCGACTCTGCTAACCATGTTCATGCCTTCTTCTTTTTC*
CTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAGGTA
AGTT

Bg C sequence

Bg C exon 2 and 3 are underlined.

GGGCTGCTGGTTGTCTACCCATGGACCCAGAGGTTCTTCGAGTCCTTTGGGG
ACCTGTCCTCTGCTGCTGTTATGAACAATCCTAAGGTGAAGGCTCATGGCAA
GAAGGTGCTGGCTGCCTTCAGTGAGGGTCTGAGTCACCTGGACAACCTCAA
AGGCACCTTTGCTAAGCTGAGTGAACTGCACTGTGACAAGCTGCACGTGGA
TCCTGAGAACTTCAGGTAAGTTTGGGGACCCTTGATTGTTCTTTCTTTTTCGC
TATTGTAAAAATTCATGTTATATGGTCGACTCTGCTAACCATGTTCATGCCTT
CTTCTTTTTCCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATC
ATTTTGGCAGGTAAGTT

BIBLIOGRAPHY

ADELMAN, K. & LIS, J. T. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet,* 13**,** 720-31.

ADESNIK, M., SALDITT, M., THOMAS, W. & DARNELL, J. E. 1972. Evidence that all messenger RNA molecules (except histone messenger RNA) contain Poly (A) sequences and that the Poly(A) has a nuclear function. *J Mol Biol,* 71**,** 21-30.

AKHTAR, M. S., HEIDEMANN, M., TIETJEN, J. R., ZHANG, D. W., CHAPMAN, R. D., EICK, D. & ANSARI, A. Z. 2009. TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Mol Cell,* 34**,** 387-93.

AKOULITCHEV, S., MAKELA, T. P., WEINBERG, R. A. & REINBERG, D. 1995. Requirement for TFIIH kinase activity in transcription by RNA polymerase II. *Nature,* 377**,** 557-60.

ALEXANDER, R. D., INNOCENTE, S. A., BARRASS, J. D. & BEGGS, J. D. 2010. Splicing-dependent RNA polymerase pausing in yeast. *Mol Cell,* 40**,** 582-93.

ALLISON, L. A., MOYLE, M., SHALES, M. & INGLES, C. J. 1985. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell,* 42**,** 599-610.

ALLMANG, C., KUFEL, J., CHANFREAU, G., MITCHELL, P., PETFALSKI, E. & TOLLERVEY, D. 1999. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J,* 18**,** 5399-410.

ALMADA, A. E., WU, X., KRIZ, A. J., BURGE, C. B. & SHARP, P. A. 2013. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature,* 499**,** 360-3.

ANDERSSON, R., ENROTH, S., RADA-IGLESIAS, A., WADELIUS, C. & KOMOROWSKI, J. 2009. Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res,* 19**,** 1732-41.

ANDRULIS, E. D., WERNER, J., NAZARIAN, A., ERDJUMENT-BROMAGE, H., TEMPST, P. & LIS, J. T. 2002. The RNA processing exosome is linked to elongating RNA polymerase II in Drosophila. *Nature,* 420**,** 837-41.

ARIMBASSERI, A. G., RIJAL, K. & MARAIA, R. J. 2014. Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. *Transcription,* 5**,** e27639.

AWASTHI, S. & ALWINE, J. C. 2003. Association of polyadenylation cleavage factor I with U1 snRNP. *RNA,* 9**,** 1400-9.

BARTKOWIAK, B., LIU, P., PHATNANI, H. P., FUDA, N. J., COOPER, J. J., PRICE, D. H., ADELMAN, K., LIS, J. T. & GREENLEAF, A. L. 2010. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev,* 24**,** 2303-16.

BATSCHE, E., YANIV, M. & MUCHARDT, C. 2006. The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol,* 13**,** 22-9.

BAUREN, G. & WIESLANDER, L. 1994. Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. *Cell,* 76**,** 183-92.

BENNETT, M., MICHAUD, S., KINGSTON, J. & REED, R. 1992a. Protein components specifically associated with prespliceosome and spliceosome complexes. *Genes Dev,* 6**,** 1986-2000.

BENNETT, M., PINOL-ROMA, S., STAKNIS, D., DREYFUSS, G. & REED, R. 1992b. Differential binding of heterogeneous nuclear ribonucleoproteins to mRNA precursors prior to spliceosome assembly in vitro. *Mol Cell Biol,* 12**,** 3165-75.

BERG, M. G., SINGH, L. N., YOUNIS, I., LIU, Q., PINTO, A. M., KAIDA, D., ZHANG, Z., CHO, S., SHERRILL-MIX, S., WAN, L. & DREYFUSS, G. 2012. U1 snRNP determines mRNA length and regulates isoform expression. *Cell,* 150**,** 53-64.

BERGET, S. M. 1995. Exon recognition in vertebrate splicing. *J Biol Chem,* 270**,** 2411-4.

BERGET, S. M., MOORE, C. & SHARP, P. A. 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A,* 74**,** 3171-5.

BERGLUND, J. A., ABOVICH, N. & ROSBASH, M. 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev,* 12**,** 858-67.

BERTOLOTTI, A., LUTZ, Y., HEARD, D. J., CHAMBON, P. & TORA, L. 1996. hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J,* 15**,** 5022-31.

BEYER, A. L. & OSHEIM, Y. N. 1988. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev,* 2**,** 754-65.

BIRD, G., ZORIO, D. A. & BENTLEY, D. L. 2004. RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'-end formation. *Mol Cell Biol,* 24**,** 8963-9.

BLACK, D. L., CHABOT, B. & STEITZ, J. A. 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. *Cell,* 42**,** 737-50.

BLOBEL, G. 1973. A protein of molecular weight 78,000 bound to the polyadenylate region of eukaryotic messenger RNAs. *Proc Natl Acad Sci U S A,* 70**,** 924-8.

BOELENS, W. C., JANSEN, E. J., VAN VENROOIJ, W. J., STRIPECKE, R., MATTAJ, I. W. & GUNDERSON, S. I. 1993. The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. *Cell,* 72**,** 881-92.

BOUSQUET-ANTONELLI, C., PRESUTTI, C. & TOLLERVEY, D. 2000. Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell,* 102**,** 765-75.

BRANNAN, K., KIM, H., ERICKSON, B., GLOVER-CUTTER, K., KIM, S., FONG, N., KIEMELE, L., HANSEN, K., DAVIS, R., LYKKE-ANDERSEN, J. & BENTLEY, D. L. 2012. mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. *Mol Cell,* 46**,** 311-24.

BREATHNACH, R., BENOIST, C., O'HARE, K., GANNON, F. & CHAMBON, P. 1978. Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proc Natl Acad Sci U S A,* 75**,** 4853-7.

BRINGMANN, P. & LUHRMANN, R. 1986. Purification of the individual snRNPs U1, U2, U5 and U4/U6 from HeLa cells and characterization of their protein constituents. *EMBO J,* 5**,** 3509-16.

BRINSTER, R. L., ALLEN, J. M., BEHRINGER, R. R., GELINAS, R. E. & PALMITER, R. D. 1988. Introns increase transcriptional efficiency in transgenic mice. *Proc Natl Acad Sci U S A,* 85**,** 836-40.

BRODY, Y., NEUFELD, N., BIEBERSTEIN, N., CAUSSE, S. Z., BOHNLEIN, E. M., NEUGEBAUER, K. M., DARZACQ, X. & SHAV-TAL, Y. 2011. The in vivo kinetics of RNA polymerase II elongation during co-transcriptional splicing. *PLoS Biol,* 9**,** e1000573.

BROWN, K. M. & GILMARTIN, G. M. 2003. A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. *Mol Cell,* 12**,** 1467-76.

BURATOWSKI, S. 2009. Progression through the RNA polymerase II CTD cycle. *Mol Cell,* 36**,** 541-6.

BUSHNELL, D. A., WESTOVER, K. D., DAVIS, R. E. & KORNBERG, R. D. 2004. Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. *Science,* 303**,** 983-8.

CANG, Y., AUBLE, D. T. & PRELICH, G. 1999. A new regulatory domain on the TATA-binding protein. *EMBO J,* 18**,** 6662-71.

CARRILLO OESTERREICH, F., PREIBISCH, S. & NEUGEBAUER, K. M. 2010. Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Mol Cell,* 40**,** 571-81.

CHAPMAN, K. B. & BOEKE, J. D. 1991. Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell,* 65**,** 483-92.

CHATHOTH, K. T., BARRASS, J. D., WEBB, S. & BEGGS, J. D. 2014. A splicing-dependent transcriptional checkpoint associated with prespliceosome formation. *Mol Cell,* 53**,** 779-90.

CHENG, B. & PRICE, D. H. 2007. Properties of RNA polymerase II elongation complexes before and after the P-TEFb-mediated transition into productive elongation. *J Biol Chem,* 282**,** 21901-12.

CHO, E. J., TAKAGI, T., MOORE, C. R. & BURATOWSKI, S. 1997. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev,* 11**,** 3319-26.

CHOI, T., HUANG, M., GORMAN, C. & JAENISCH, R. 1991. A generic intron increases gene expression in transgenic mice. *Mol Cell Biol,* 11**,** 3070-4.

CHOW, L. T., GELINAS, R. E., BROKER, T. R. & ROBERTS, R. J. 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell,* 12**,** 1-8.

CLOUTIER, P., TOUTANT, J., SHKRETA, L., GOEKJIAN, S., REVIL, T. & CHABOT, B. 2008. Antagonistic effects of the SRp30c protein and cryptic 5'splice sites on the alternative splicing of the apoptotic regulator Bcl-x. *J Biol Chem,* 283**,** 21315-24.

COLE, M. D. & COWLING, V. H. 2008. Transcription-independent functions of MYC: regulation of translation and DNA replication. *Nat Rev Mol Cell Biol,* 9**,** 810-5.

COLGAN, D. F. & MANLEY, J. L. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes Dev,* 11**,** 2755-66.

CONTI, E. & IZAURRALDE, E. 2005. Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr Opin Cell Biol,* 17**,** 316-25.

COPPOLA, J. A., FIELD, A. S. & LUSE, D. S. 1983. Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. *Proc Natl Acad Sci U S A,* 80**,** 1251-5.

CORDEN, J. L., CADENA, D. L., AHEARN, J. M., JR. & DAHMUS, M. E. 1985. A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc Natl Acad Sci U S A,* 82**,** 7934-8.

CORE, L. J., WATERFALL, J. J. & LIS, J. T. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science,* 322**,** 1845-8.

COULOMBE, B. & BURTON, Z. F. 1999. DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms. *Microbiol Mol Biol Rev,* 63**,** 457-78.

CRAMER, P., CACERES, J. F., CAZALLA, D., KADENER, S., MURO, A. F., BARALLE, F. E. & KORNBLIHTT, A. R. 1999. Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. *Mol Cell,* 4**,** 251-8.

CRAMER, P., PESCE, C. G., BARALLE, F. E. & KORNBLIHTT, A. R. 1997. Functional association between promoter structure and transcript alternative splicing. *Proc Natl Acad Sci U S A,* 94**,** 11456-60.

CULBERTSON, M. R., UNDERBRINK, K. M. & FINK, G. R. 1980. Frameshift suppression Saccharomyces cerevisiae. II. Genetic properties of group II suppressors. *Genetics,* 95**,** 833-53.

DAMGAARD, C. K., KAHNS, S., LYKKE-ANDERSEN, S., NIELSEN, A. L., JENSEN, T. H. & KJEMS, J. 2008. A 5' splice site enhances the recruitment of basal transcription initiation factors in vivo. *Mol Cell,* 29**,** 271-8.

DANTONEL, J. C., MURTHY, K. G., MANLEY, J. L. & TORA, L. 1997. Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. *Nature,* 389**,** 399-402.

DAS, B., DAS, S. & SHERMAN, F. 2006a. Mutant LYS2 mRNAs retained and degraded in the nucleus of Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A,* 103**,** 10871-6.

DAS, R., DUFU, K., ROMNEY, B., FELDT, M., ELENKO, M. & REED, R. 2006b. Functional coupling of RNAP II transcription to spliceosome assembly. *Genes Dev,* 20**,** 1100-9.

DAS, R., YU, J., ZHANG, Z., GYGI, M. P., KRAINER, A. R., GYGI, S. P. & REED, R. 2007. SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol Cell,* 26**,** 867-81.

DAVID, C. J., BOYNE, A. R., MILLHOUSE, S. R. & MANLEY, J. L. 2011. The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev,* 25**,** 972-83.

DAVIDSON, L., MUNIZ, L., WEST, S. 2014. 3’ enf formation of pre-mRNA and phosphorylation of Ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells. *Genes Dev,* 28**,** 342-56.

DE ALMEIDA, S. F. & CARMO-FONSECA, M. 2008. The CTD role in cotranscriptional RNA processing and surveillance. *FEBS Lett,* 582**,** 1971-6.

DE LA MATA, M. & KORNBLIHTT, A. R. 2006. RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. *Nat Struct Mol Biol,* 13**,** 973-80.

DECROLY, E., FERRON, F., LESCAR, J. & CANARD, B. 2012. Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol,* 10**,** 51-65.

DONIS-KELLER, H. 1979. Site specific enzymatic cleavage of RNA. *Nucleic Acids Res,* 7**,** 179-92.

DENISON, R. A., VAN ARSDELL, S. W., BERNSTEIN, L. B. & WEINER, A. M. 1981. Abundant pseudogenes for small nuclear RNAs are dispersed in the human genome. *Proc Natl Acad Sci U S A,* 78**,** 810-4.

DONMEZ, G., HARTMUTH, K., KASTNER, B., WILL, C. L. & LUHRMANN, R. 2007. The 5' end of U2 snRNA is in close proximity to U1 and functional sites of the pre-mRNA in early spliceosomal complexes. *Mol Cell,* 25**,** 399-411.
DU, H. & ROSBASH, M. 2002. The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. *Nature,* 419**,** 86-90.
DUNCKLEY, T. & PARKER, R. 1999. The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif. *EMBO J,* 18**,** 5411-22.
DZIEMBOWSKI, A., LORENTZEN, E., CONTI, E. & SERAPHIN, B. 2007. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat Struct Mol Biol,* 14**,** 15-22.
EDERY, I. & SONENBERG, N. 1985. Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc Natl Acad Sci U S A,* 82**,** 7590-4.
EDMONDS, M., VAUGHAN, M. H., JR. & NAKAZATO, H. 1971. Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly-labeled polyribosomal RNA of HeLa cells: possible evidence for a precursor relationship. *Proc Natl Acad Sci U S A,* 68**,** 1336-40.
EGLOFF, S., DIENSTBIER, M. & MURPHY, S. 2012. Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends Genet,* 28**,** 333-41.
EGLOFF, S., O'REILLY, D., CHAPMAN, R. D., TAYLOR, A., TANZHAUS, K., PITTS, L., EICK, D. & MURPHY, S. 2007. Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science,* 318**,** 1777-9.
EHRMANN, I., DALGLIESH, C., LIU, Y., DANILENKO, M., CROSIER, M., OVERMAN, L., ARTHUR, H. M., LINDSAY, S., CLOWRY, G. J., VENABLES, J. P., FORT, P. & ELLIOTT, D. J. 2013. The tissue-specific RNA binding protein T-STAR controls regional splicing patterns of neurexin pre-mRNAs in the brain. *PLoS Genet,* 9**,** e1003474.
EPERON, I. C., IRELAND, D. C., SMITH, R. A., MAYEDA, A. & KRAINER, A. R. 1993. Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. *EMBO J,* 12**,** 3607-17.
EPERON, I. C., MAKAROVA, O. V., MAYEDA, A., MUNROE, S. H., CACERES, J. F., HAYWARD, D. G. & KRAINER, A. R. 2000. Selection of alternative 5' splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1. *Mol Cell Biol,* 20**,** 8303-18.
EPERON, L. P., GRAHAM, I. R., GRIFFITHS, A. D. & EPERON, I. C. 1988. Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell,* 54**,** 393-401.
FEAVER, W. J., GILEADI, O., LI, Y. & KORNBERG, R. D. 1991. CTD kinase associated with yeast RNA polymerase II initiation factor b. *Cell,* 67**,** 1223-30.
FEAVER, W. J., SVEJSTRUP, J. Q., HENRY, N. L. & KORNBERG, R. D. 1994. Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. *Cell,* 6, 1103-9.
FISCHER, U., LIU, Q. & DREYFUSS, G. 1997. The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell,* 90**,** 1023-9.
FOLCO, E. G. & REED, R. 2014. In vitro systems for coupling RNAP II transcription to splicing and polyadenylation. *Methods Mol Biol,* 1126**,** 169-77.
FONG, N. & BENTLEY, D. L. 2001. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev,* 15**,** 1783-95.

FONG, Y. W. & ZHOU, Q. 2001. Stimulatory effect of splicing factors on transcriptional elongation. *Nature,* 414**,** 929-33.
FUJINAGA, K., IRWIN, D., HUANG, Y., TAUBE, R., KUROSU, T. & PETERLIN, B. M. 2004. Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol Cell Biol,* 24**,** 787-95.
FURGER, A., O'SULLIVAN, J. M., BINNIE, A., LEE, B. A. & PROUDFOOT, N. J. 2002. Promoter proximal splice sites enhance transcription. *Genes Dev,* 16**,** 2792-9.
FURUICHI, Y., LAFIANDRA, A. & SHATKIN, A. J. 1977. 5'-Terminal structure and mRNA stability. *Nature,* 266**,** 235-9.
GHOSH, S. & GARCIA-BLANCO, M. A. 2000. Coupled in vitro synthesis and splicing of RNA polymerase II transcripts. *RNA,* 6**,** 1325-34.
GIL, A. & PROUDFOOT, N. J. 1984. A sequence downstream of AAUAAA is required for rabbit beta-globin mRNA 3'-end formation. *Nature,* 312**,** 473-4.
GIL, A. & PROUDFOOT, N. J. 1987. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit beta-globin mRNA 3' end formation. *Cell,* 49**,** 399-406.
GLISOVIC, T., BACHORIK, J. L., YONG, J. & DREYFUSS, G. 2008. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett,* 582**,** 1977-86.
GONATOPOULOS-POURNATZIS, T. & COWLING, V. H. 2014. Cap-binding complex (CBC). *Biochem J,* 457**,** 231-42.
GRABOWSKI, P. J., PADGETT, R. A. & SHARP, P. A. 1984. Messenger RNA splicing in vitro: an excised intervening sequence and a potential intermediate. *Cell,* 37**,** 415-27.
GRAVELEY, B. R. 2000. Sorting out complexity of SR protein functions. *RNA,* 9, 1197-211.
GREENLEAF, A. L. 1993. Positive patches and negative noodles: linking RNA processing to transcription? *Trends Biochem Sci,* 18**,** 117-9.
GREIFENBERG, A. K., HONIG, D., PILAROVA, K., DUSTER, R., BARTHOLOMEEUSEN, K., BOSKEN, C. A., ANAND, K., BLAZEK, D. & GEYER, M. 2016. Structural and Functional Analysis of the Cdk13/Cyclin K Complex. *Cell Rep,* 14**,** 320-31.
GROMAK, N., WEST, S. & PROUDFOOT, N. J. 2006. Pause sites promote transcriptional termination of mammalian RNA polymerase II. *Mol Cell Biol,* 26**,** 3986-96.
GRZECHNIK, P., TAN-WONG, S. M. & PROUDFOOT, N. J. 2014. Terminate and make a loop: regulation of transcriptional directionality. *Trends Biochem Sci,* 39**,** 319-27.
GU, B., EICK, D. & BENSAUDE, O. 2013. CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. *Nucleic Acids Res,* 41**,** 1591-603.
GUIGUEN, A., SOUTOURINA, J., DEWEZ, M., TAFFOREAU, L., DIEU, M., RAES, M., VANDENHAUTE, J., WERNER, M. & HERMAND, D. 2007. Recruitment of P-TEFb (Cdk9-Pch1) to chromatin by the cap-methyl transferase Pcm1 in fission yeast. *EMBO J,* 26**,** 1552-9.
GUIRO, J. & O'REILLY, D. 2015. Insights into the U1 small nuclear ribonucleoprotein complex superfamily. *Wiley Interdiscip Rev RNA,* 6**,** 79-92.

GUNDERSON, S. I., POLYCARPOU-SCHWARZ, M. & MATTAJ, I. W. 1998. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell,* 1**,** 255-64.

GUO, J., GARRETT, M., MICKLEM, G. & BROGNA, S. 2011. Poly(A) signals located near the 5' end of genes are silenced by a general mechanism that prevents premature 3'-end processing. *Mol Cell Biol,* 31**,** 639-51.

HE, F. & JACOBSON, A. 2015. Nonsense-Mediated mRNA Decay: Degradation of Defective Transcripts Is Only Part of the Story. *Annu Rev Genet,* 49**,** 339-66.

HEIDEMANN, M. & EICK, D. 2012. Tyrosine-1 and threonine-4 phosphorylation marks complete the RNA polymerase II CTD phospho-code. *RNA Biol,* 9**,** 1144-6.

HERMANN, H., FABRIZIO, P., RAKER, V. A., FOULAKI, K., HORNIG, H., BRAHMS, H. & LUHRMANN, R. 1995. snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J,* 14**,** 2076-88.

HERNANDEZ, H., DZIEMBOWSKI, A., TAVERNER, T., SERAPHIN, B. & ROBINSON, C. V. 2006. Subunit architecture of multimeric complexes isolated directly from cells. *EMBO Rep,* 7**,** 605-10.

HICKS, M. J., YANG, C. R., KOTLAJICH, M. V. & HERTEL, K. J. 2006. Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns. *PLoS Biol,* 4**,** e147.

HILLIKER, A. K., MEFFORD, M. A. & STALEY, J. P. 2007. U2 toggles iteratively between the stem IIa and stem IIc conformations to promote pre-mRNA splicing. *Genes Dev,* 21**,** 821-34.

HIROSE, Y. & MANLEY, J. L. 1998. RNA polymerase II is an essential mRNA polyadenylation factor. *Nature,* 395**,** 93-6.

HIROSE, Y., TACKE, R. & MANLEY, J. L. 1999. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev,* 13, 1234-9.

HO, C. K. & SHUMAN, S. 1999. Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol Cell,* 3, 405-11.

HODSON, 2011. Use of single molecule methods to reveal the mechanisms of splice site selection.

HODSON, M. J., HUDSON, A. J., CHERNY, D. & EPERON, I. C. 2012. The transition in spliceosome assembly from complex E to complex A purges surplus U1 snRNPs from alternative splice sites. *Nucleic Acids Res,* 40**,** 6850-62.

HOUSELEY, J. & TOLLERVEY, D. 2009. The many pathways of RNA degradation. *Cell,* 136**,** 763-76.

HSIN, J. P. & MANLEY, J. L. 2012. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev,* 26**,** 2119-37.

HUANG, S. & SPECTOR, D. L. 1996. Intron-dependent recruitment of pre-mRNA splicing factors to sites of transcription. *J Cell Biol,* 133**,** 719-32.

HURANOVA, M., IVANI, I., BENDA, A., POSER, I., BRODY, Y., HOF, M., SHAV-TAL, Y., NEUGEBAUER, K. M. & STANEK, D. 2010. The differential interaction of snRNPs with pre-mRNA reveals splicing kinetics in living cells. *J Cell Biol,* 191**,** 75-86.

IZAURRALDE, E., LEWIS, J., MCGUIGAN, C., JANKOWSKA, M., DARZYNKIEWICZ, E. & MATTAJ, I. W. 1994. A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell,* 78**,** 657-68.

IZBAN, M. G. & LUSE, D. S. 1992. The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5' direction in the presence of elongation factor SII. *Genes Dev,* 6**,** 1342-56.
JAMISON, S. F., CROW, A. & GARCIA-BLANCO, M. A. 1992. The spliceosome assembly pathway in mammalian extracts. *Mol Cell Biol,* 12**,** 4279-87.
JI, X., ZHOU, Y., PANDIT, S., HUANG, J., LI, H., LIN, C. Y., XIAO, R., BURGE, C. B. & FU, X. D. 2013. SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell,* 153**,** 855-68.
JIMENEZ-GARCIA, L. F. & SPECTOR, D. L. 1993. In vivo evidence that transcription and splicing are coordinated by a recruiting mechanism. *Cell,* 73**,** 47-59.
JOBERT, L., PINZON, N., VAN HERREWEGHE, E., JADY, B. E., GUIALIS, A., KISS, T. & TORA, L. 2009. Human U1 snRNA forms a new chromatin-associated snRNP with TAF15. *EMBO Rep,* 10**,** 494-500.
JONKERS, I., KWAK, H. & LIS, J. T. 2014. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife,* 3**,** e02407.
JONKERS, I. & LIS, J. T. 2015. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol,* 16**,** 167-77.
KADENER, S., FEDEDA, J. P., ROSBASH, M. & KORNBLIHTT, A. R. 2002. Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation. *Proc Natl Acad Sci U S A,* 99**,** 8185-90.
KAIDA, D., BERG, M. G., YOUNIS, I., KASIM, M., SINGH, L. N., WAN, L. & DREYFUSS, G. 2010. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature,* 468**,** 664-8.
KANDELS-LEWIS, S. & SERAPHIN, B. 1993. Involvement of U6 snRNA in 5' splice site selection. *Science,* 262**,** 2035-9.
KAUFMANN, I., MARTIN, G., FRIEDLEIN, A., LANGEN, H. & KELLER, W. 2004. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J,* 23**,** 616-26.
KIM, E., DU, L., BREGMAN, D. B. & WARREN, S. L. 1997. Splicing factors associate with hyperphosphorylated RNA polymerase II in the absence of pre-mRNA. *J Cell Biol,* 136**,** 19-28.
KIM, Y. J., BJORKLUND, S., LI, Y., SAYRE, M. H. & KORNBERG, R. D. 1994. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell,* 77**,** 599-608.
KOGA, M., HAYASHI, M. & KAIDA, D. 2015. Splicing inhibition decreases phosphorylation level of Ser2 in Pol II CTD. *Nucleic Acids Res,* 43**,** 8258-67.
KOHTZ, J. D., JAMISON, S. F., WILL, C. L., ZUO, P., LUHRMANN, R., GARCIA-BLANCO, M. A. & MANLEY, J. L. 1994. Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature,* 368**,** 119-24.
KOLASINSKA-ZWIERZ, P., DOWN, T., LATORRE, I., LIU, T., LIU, X. S. & AHRINGER, J. 2009. Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet,* 41**,** 376-81.
KOLE, R. & WEISSMAN, S. M. 1982. Accurate in vitro splicing of human beta-globin RNA. *Nucleic Acids Res,* 10**,** 5429-45.
KOMARNITSKY, P., CHO, E. J. & BURATOWSKI, S. 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev,* 14**,** 2452-60.

KONARSKA, M. M., PADGETT, R. A. & SHARP, P. A. 1984. Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell,* 38**,** 731-6.
KORNBLIHTT, A. R., DE LA MATA, M., FEDEDA, J. P., MUNOZ, M. J. & NOGUES, G. 2004. Multiple links between transcription and splicing. *RNA,* 10**,** 1489-98.
KORNBLIHTT, A. R., SCHOR, I. E., ALLO, M., DUJARDIN, G., PETRILLO, E. & MUNOZ, M. J. 2013. Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat Rev Mol Cell Biol,* 14**,** 153-65.
KRAINER, A. R., MANIATIS, T., RUSKIN, B. & GREEN, M. R. 1984. Normal and mutant human beta-globin pre-mRNAs are faithfully and efficiently spliced in vitro. *Cell,* 36**,** 993-1005.
KRAMER, A., KELLER, W., APPEL, B. & LUHRMANN, R. 1984. The 5' terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors. *Cell,* 38**,** 299-307.
KRAMER, A. & UTANS, U. 1991. Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs. *EMBO J,* 10**,** 1503-9.
KUGEL, J. F. & GOODRICH, J. A. 1998. Promoter escape limits the rate of RNA polymerase II transcription and is enhanced by TFIIE, TFIIH, and ATP on negatively supercoiled DNA. *Proc Natl Acad Sci U S A,* 95**,** 9232-7.
KURAS, L. & STRUHL, K. 1999. Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. *Nature,* 399**,** 609-13.
KWEK, K. Y., MURPHY, S., FURGER, A., THOMAS, B., O'GORMAN, W., KIMURA, H., PROUDFOOT, N. J. & AKOULITCHEV, A. 2002. U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nat Struct Biol,* 9**,** 800-5.
LACADIE, S. A. & ROSBASH, M. 2005. Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. *Mol Cell,* 19**,** 65-75.
LAZAREV, D. & MANLEY, J. L. 2007. Concurrent splicing and transcription are not sufficient to enhance splicing efficiency. *RNA,* 13**,** 1546-57
LEWIS, B. P., GREEN, R. E. & BRENNER, S. E. 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A,* 100**,** 189-92..
LEE, G., BRATKOWSKI, M. A., DING, F., KE, A. & HA, T. 2012. Elastic coupling between RNA degradation and unwinding by an exoribonuclease. *Science,* 336**,** 1726-9.
LEICHTER, M., MARKO, M., GANOU, V., PATRINOU-GEORGOULA, M., TORA, L. & GUIALIS, A. 2011. A fraction of the transcription factor TAF15 participates in interactions with a subset of the spliceosomal U1 snRNP complex. *Biochim Biophys Acta,* 1814**,** 1812-24.
LENASI, T., PETERLIN, B. M. & BARBORIC, M. 2011. Cap-binding protein complex links pre-mRNA capping to transcription elongation and alternative splicing through positive transcription elongation factor b (P-TEFb). *J Biol Chem,* 286**,** 22758-68.
LEWIS, J. D., GUNDERSON, S. I. & MATTAJ, I. W. 1995. The influence of 5' and 3' end structures on pre-mRNA metabolism. *J Cell Sci Suppl,* 19**,** 13-9.
LEWIS, J. D., IZAURRALDE, E., JARMOLOWSKI, A., MCGUIGAN, C. & MATTAJ, I. W. 1996. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev,* 10**,** 1683-98.

LIAO, S. M., ZHANG, J., JEFFERY, D. A., KOLESKE, A. J., THOMPSON, C. M., CHAO, D. M., VILJOEN, M., VAN VUUREN, H. J. & YOUNG, R. A. 1995. A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature,* 374**,** 193-6.
LIBRI, D., DOWER, K., BOULAY, J., THOMSEN, R., ROSBASH, M. & JENSEN, T. H. 2002. Interactions between mRNA export commitment, 3'-end quality control, and nuclear degradation. *Mol Cell Biol,* 22**,** 8254-66.
LIM, L. & CANELLAKIS, E. S. 1970. Adenine-rich polymer associated with rabbit reticulocyte messenger RNA. *Nature,* 227**,** 710-2.
LIN, S., COUTINHO-MANSFIELD, G., WANG, D., PANDIT, S. & FU, X. D. 2008. The splicing factor SC35 has an active role in transcriptional elongation. *Nat Struct Mol Biol,* 15**,** 819-26.
LISTERMAN, I., SAPRA, A. K. & NEUGEBAUER, K. M. 2006. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat Struct Mol Biol,* 13**,** 815-22.
LIU, S. W., JIAO, X., LIU, H., GU, M., LIMA, C. D. & KILEDJIAN, M. 2004. Functional analysis of mRNA scavenger decapping enzymes. *RNA,* 10**,** 1412-22.
LORENTZEN, E., BASQUIN, J., TOMECKI, R., DZIEMBOWSKI, A. & CONTI, E. 2008. Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol Cell,* 29**,** 717-28.
LU, H., FLORES, O., WEINMANN, R. & REINBERG, D. 1991. The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proc Natl Acad Sci U S A,* 88**,** 10004-8.
LU, H., ZAWEL, L., FISHER, L., EGLY, J. M. & REINBERG, D. 1992. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature,* 358**,** 641-5.
LU, X. B., HEIMER, J., REKOSH, D. & HAMMARSKJOLD, M. L. 1990. U1 small nuclear RNA plays a direct role in the formation of a rev-regulated human immunodeficiency virus env mRNA that remains unspliced. *Proc Natl Acad Sci U S A,* 87**,** 7598-602.
LUSE, D. S. 2014. The RNA polymerase II preinitiation complex. Through what pathway is the complex assembled? *Transcription,* 5**,** e27050.
MADHANI, H. D. & GUTHRIE, C. 1992. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell,* 71**,** 803-17.
MANDAL, S. S., CHU, C., WADA, T., HANDA, H., SHATKIN, A. J. & REINBERG, D. 2004. Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proc Natl Acad Sci U S A,* 101**,** 7572-7.
MANDEL, C. R., BAI, Y. & TONG, L. 2008. Protein factors in pre-mRNA 3'-end processing. *Cell Mol Life Sci,* 65**,** 1099-122.
MANDEL, C. R., KANEKO, S., ZHANG, H., GEBAUER, D., VETHANTHAM, V., MANLEY, J. L. & TONG, L. 2006. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature,* 444**,** 953-6.
MARSHALL, N. F., PENG, J., XIE, Z. & PRICE, D. H. 1996. Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J Biol Chem,* 271**,** 27176-83.
MARSHALL, N. F. & PRICE, D. H. 1995. Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J Biol Chem,* 270**,** 12335-8.

MATHIS, D. J. & CHAMBON, P. 1981. The SV40 early region TATA box is required for accurate in vitro initiation of transcription. *Nature,* 290**,** 310-5.
MATTER, N., HERRLICH, P. & KONIG, H. 2002. Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature,* 420**,** 691-5.
MAYES, A. E., VERDONE, L., LEGRAIN, P. & BEGGS, J. D. 1999. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J,* 18**,** 4321-31.
MCCRACKEN, S., FONG, N., YANKULOV, K., BALLANTYNE, S., PAN, G., GREENBLATT, J., PATTERSON, S. D., WICKENS, M. & BENTLEY, D. L. 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature,* 385**,** 357-61.
MCLAUCHLAN, J., GAFFNEY, D., WHITTON, J. L. & CLEMENTS, J. B. 1985. The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. *Nucleic Acids Res,* 13**,** 1347-68.
MEINHART, A. & CRAMER, P. 2004. Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature,* 430**,** 223-6.
MENDECKI, J., LEE, S. Y. & BRAWERMAN, G. 1972. Characteristics of the polyadenylic acid segment associated with messenger ribonucleic acid in mouse sarcoma 180 ascites cells. *Biochemistry,* 11**,** 792-8.
MERENDINO, L., GUTH, S., BILBAO, D., MARTINEZ, C. & VALCARCEL, J. 1999. Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature,* 402**,** 838-41.
MERKHOFER, E. C., HU, P. & JOHNSON, T. L. 2014. Introduction to cotranscriptional RNA splicing. *Methods Mol Biol,* 1126**,** 83-96.
MICHAUD, S. & REED, R. 1991. An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes Dev,* 5**,** 2534-46.
MINVIELLE-SEBASTIA, L., PREKER, P. J., WIEDERKEHR, T., STRAHM, Y. & KELLER, W. 1997. The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. *Proc Natl Acad Sci U S A,* 94**,** 7897-902.
MISTELI, T. & SPECTOR, D. L. 1999. RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol Cell,* 3**,** 697-705.
MITCHELL, P., PETFALSKI, E., SHEVCHENKO, A., MANN, M. & TOLLERVEY, D. 1997. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. *Cell,* 91**,** 457-66.
MOORE, M. J. & SHARP, P. A. 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature,* 365**,** 364-8.
MORTILLARO, M. J., BLENCOWE, B. J., WEI, X., NAKAYASU, H., DU, L., WARREN, S. L., SHARP, P. A. & BEREZNEY, R. 1996. A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proc Natl Acad Sci U S A,* 93**,** 8253-7.
MOTEKI, S. & PRICE, D. 2002. Functional coupling of capping and transcription of mRNA. *Mol Cell,* 10**,** 599-609.
MOUNT, S. M., PETTERSSON, I., HINTERBERGER, M., KARMAS, A. & STEITZ, J. A. 1983. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. *Cell,* 33**,** 509-18.
MUELLER, W. F. & HERTEL, K. J. 2014. Kinetic analysis of in vitro pre-mRNA splicing in HeLa nuclear extract. *Methods Mol Biol,* 1126**,** 161-8.

MUHLEMANN, O., MOCK-CASAGRANDE, C. S., WANG, J., LI, S., CUSTODIO, N., CARMO-FONSECA, M., WILKINSON, M. F. & MOORE, M. J. 2001. Precursor RNAs harboring nonsense codons accumulate near the site of transcription. *Mol Cell,* 8**,** 33-43.

MUNOZ, M. J., DE LA MATA, M. & KORNBLIHTT, A. R. 2010. The carboxy terminal domain of RNA polymerase II and alternative splicing. *Trends Biochem Sci,* 35**,** 497-504.

NAG, A., NARSINH, K. & MARTINSON, H. G. 2007. The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. *Nat Struct Mol Biol,* 14**,** 662-9.

NAHKURI, S., TAFT, R. J. & MATTICK, J. S. 2009. Nucleosomes are preferentially positioned at exons in somatic and sperm cells. *Cell Cycle,* 8**,** 3420-4.

NAYLOR, S. L., ZABEL, B. U., MANSER, T., GESTELAND, R. & SAKAGUCHI, A. Y. 1984. Localization of human U1 small nuclear RNA genes to band p36.3 of chromosome 1 by in situ hybridization. *Somat Cell Mol Genet,* 10**,** 307-13.

NEUGEBAUER, K. M. 2002. On the importance of being co-transcriptional. *J Cell Sci,* 115**,** 3865-71.

NEUGEBAUER, K. M. & ROTH, M. B. 1997. Distribution of pre-mRNA splicing factors at sites of RNA polymerase II transcription. *Genes Dev,* 11**,** 1148-59.

NIKOLOV, D. B., CHEN, H., HALAY, E. D., USHEVA, A. A., HISATAKE, K., LEE, D. K., ROEDER, R. G. & BURLEY, S. K. 1995. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature,* 377**,** 119-28.

NOGUES, G., KADENER, S., CRAMER, P., DE LA MATA, M., FEDEDA, J. P., BLAUSTEIN, M., SREBROW, A. & KORNBLIHTT, A. R. 2003. Control of alternative pre-mRNA splicing by RNA Pol II elongation: faster is not always better. *IUBMB Life,* 55**,** 235-41.

O'GORMAN, W., THOMAS, B., KWEK, K. Y., FURGER, A. & AKOULITCHEV, A. 2005. Analysis of U1 small nuclear RNA interaction with cyclin H. *J Biol Chem,* 280**,** 36920-5.

O'REILLY, D., DIENSTBIER, M., COWLEY, S. A., VAZQUEZ, P., DROZDZ, M., TAYLOR, S., JAMES, W. S. & MURPHY, S. 2013. Differentially expressed, variant U1 snRNAs regulate gene expression in human cells. *Genome Res,* 23**,** 281-91.

OHNO, M., SAKAMOTO, H. & SHIMURA, Y. 1987. Preferential excision of the 5' proximal intron from mRNA precursors with two introns as mediated by the cap structure. *Proc Natl Acad Sci U S A,* 84**,** 5187-91.

OTSUKA, Y., KEDERSHA, N. L. & SCHOENBERG, D. R. 2009. Identification of a cytoplasmic complex that adds a cap onto 5'-monophosphate RNA. *Mol Cell Biol,* 29**,** 2155-67.

PABIS, M., NEUFELD, N., STEINER, M. C., BOJIC, T., SHAV-TAL, Y. & NEUGEBAUER, K. M. 2013. The nuclear cap-binding complex interacts with the U4/U6.U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. *RNA,* 19**,** 1054-63.

PADGETT, R. A., HARDY, S. F. & SHARP, P. A. 1983. Splicing of adenovirus RNA in a cell-free transcription system. *Proc Natl Acad Sci U S A,* 80**,** 5230-4.

PAK, V., EIFLER, T. T., JAGER, S., KROGAN, N. J., FUJINAGA, K. & PETERLIN, B. M. 2015. CDK11 in TREX/THOC Regulates HIV mRNA 3' End Processing. *Cell Host Microbe,* 18**,** 560-70.

PAN, Q., SHAI, O., LEE, L. J., FREY, B. J. & BLENCOWE, B. J. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet,* 40**,** 1413-5.
PANDYA-JONES, A. & BLACK, D. L. 2009. Co-transcriptional splicing of constitutive and alternative exons. *RNA,* 15**,** 1896-908.
PATZELT, E., THALMANN, E., HARTMUTH, K., BLAAS, D. & KUECHLER, E. 1987. Assembly of pre-mRNA splicing complex is cap dependent. *Nucleic Acids Res,* 15**,** 1387-99.
PERRIMAN, R. J. & ARES, M., JR. 2007. Rearrangement of competing U2 RNA helices within the spliceosome promotes multiple steps in splicing. *Genes Dev,* 21**,** 811-20.
PETERLIN, B. M. & PRICE, D. H. 2006. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell,* 23**,** 297-305.
POMERANZ KRUMMEL, D. A., OUBRIDGE, C., LEUNG, A. K., LI, J. & NAGAI, K. 2009. Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. *Nature,* 458**,** 475-80.
PONJAVIC, J., LENHARD, B., KAI, C., KAWAI, J., CARNINCI, P., HAYASHIZAKI, Y. & SANDELIN, A. 2006. Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters. *Genome Biol,* 7**,** R78.
POZZOLI, U., RIVA, L., MENOZZI, G., CAGLIANI, R., COMI, G. P., BRESOLIN, N., GIORDA, R. & SIRONI, M. 2004. Over-representation of exonic splicing enhancers in human intronless genes suggests multiple functions in mRNA processing. *Biochem Biophys Res Commun,* 322**,** 470-6.
PREKER, P., NIELSEN, J., KAMMLER, S., LYKKE-ANDERSEN, S., CHRISTENSEN, M. S., MAPENDANO, C. K., SCHIERUP, M. H. & JENSEN, T. H. 2008. RNA exosome depletion reveals transcription upstream of active human promoters. *Science,* 322**,** 1851-4.
PRICE, D. H. 2000. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol Cell Biol,* 20**,** 2629-34.
PROUDFOOT, N. J. 1989. How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem Sci,* 14**,** 105-10.
PROUDFOOT, N. J. 2011. Ending the message: poly(A) signals then and now. *Genes Dev,* 25**,** 1770-82.
RAKER, V. A., PLESSEL, G. & LUHRMANN, R. 1996. The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *EMBO J,* 15**,** 2256-69.
RANUNCOLO, S. M., GHOSH, S., HANOVER, J. A., HART, G. W. & LEWIS, B. A. 2012. Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription in vitro and in vivo. *J Biol Chem,* 287**,** 23549-61.
RASMUSSEN, E. B. & LIS, J. T. 1993. In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. *Proc Natl Acad Sci U S A,* 90**,** 7923-7.
REDDY, R., HENNING, D., DAS, G., HARLESS, M. & WRIGHT, D. 1987. The capped U6 small nuclear RNA is transcribed by RNA polymerase III. *J Biol Chem,* 262**,** 75-81.
REED, R. 1990. Protein composition of mammalian spliceosomes assembled in vitro. *Proc Natl Acad Sci U S A,* 87**,** 8031-5.

REED, R. 2000. Mechanisms of fidelity in pre-mRNA splicing. *Curr Opin Cell Biol,* 12**,** 340-5.
PERTEA, M. & SALZBERG, S. L. 2010. Between a chicken and a grape: estimating the number of human genes. *Genome Biol,* 11**,** 206.
RICKERT, P., CORDEN, J. L. & LEES, E. 1999. Cyclin C/CDK8 and cyclin H/CDK7/p36 are biochemically distinct CTD kinases. *Oncogene,* 18**,** 1093-102.
ROBERT, F., BLANCHETTE, M., MAES, O., CHABOT, B. & COULOMBE, B. 2002. A human RNA polymerase II-containing complex associated with factors necessary for spliceosome assembly. *J Biol Chem,* 277**,** 9302-6.
ROBERTS, G. C., GOODING, C., MAK, H. Y., PROUDFOOT, N. J. & SMITH, C. W. 1998. Co-transcriptional commitment to alternative splice site selection. *Nucleic Acids Res,* 26**,** 5568-72.
ROCA, X., AKERMAN, M., GAUS, H., BERDEJA, A., BENNETT, C. F. & KRAINER, A. R. 2012. Widespread recognition of 5' splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides. *Genes Dev,* 26**,** 1098-109.
ROCA, X., KRAINER, A. R. & EPERON, I. C. 2013. Pick one, but be quick: 5' splice sites and the problems of too many choices. *Genes Dev,* 27**,** 129-44.
ROEDER, R. G. & RUTTER, W. J. 1969. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature,* 224**,** 234-7.
ROEDER, R. G. & RUTTER, W. J. 1970. Specific nucleolar and nucleoplasmic RNA polymerases. *Proc Natl Acad Sci U S A,* 65**,** 675-82.
ROSONINA, E., BAKOWSKI, M. A., MCCRACKEN, S. & BLENCOWE, B. J. 2003. Transcriptional activators control splicing and 3'-end cleavage levels. *J Biol Chem,* 278**,** 43034-40.
RUSKIN, B., KRAINER, A. R., MANIATIS, T. & GREEN, M. R. 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell,* 38**,** 317-31.
SAINSBURY, S., BERNECKY, C. & CRAMER, P. 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol,* 16**,** 129-43.
SAPRA, A. K., ANKO, M. L., GRISHINA, I., LORENZ, M., PABIS, M., POSER, I., ROLLINS, J., WEILAND, E. M. & NEUGEBAUER, K. M. 2009. SR protein family members display diverse activities in the formation of nascent and mature mRNPs in vivo. *Mol Cell,* 34**,** 179-90.
SAUNDERS, A., CORE, L. J. & LIS, J. T. 2006. Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol,* 7**,** 557-67.
SCHMID, M. & JENSEN, T. H. 2008. The exosome: a multipurpose RNA-decay machine. *Trends Biochem Sci,* 33**,** 501-10.
SCHWARTZ, S., MESHORER, E. & AST, G. 2009. Chromatin organization marks exon-intron structure. *Nat Struct Mol Biol,* 16**,** 990-5.
SEILA, A. C., CALABRESE, J. M., LEVINE, S. S., YEO, G. W., RAHL, P. B., FLYNN, R. A., YOUNG, R. A. & SHARP, P. A. 2008. Divergent transcription from active promoters. *Science,* 322**,** 1849-51.
SELTH, L. A., SIGURDSSON, S. & SVEJSTRUP, J. Q. 2010. Transcript Elongation by RNA Polymerase II. *Annu Rev Biochem,* 79**,** 271-93.
SHANDILYA, J. & ROBERTS, S. G. 2012. The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. *Biochim Biophys Acta,* 1819**,** 391-400.

SHAO, C., YANG, B., WU, T., HUANG, J., TANG, P., ZHOU, Y., ZHOU, J., QIU, J., JIANG, L., LI, H., CHEN, G., SUN, H., ZHANG, Y., DENISE, A., ZHANG, D. E. & FU, X. D. 2014. Mechanisms for U2AF to define 3' splice sites and regulate alternative splicing in the human genome. *Nat Struct Mol Biol,* 21**,** 997-1005.

SHARP, P. A. 2005. The discovery of split genes and RNA splicing. *Trends Biochem Sci,* 30**,** 279-81.

SHIMOTOHNO, K., KODAMA, Y., HASHIMOTO, J. & MIURA, K. I. 1977. Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis. *Proc Natl Acad Sci U S A,* 74**,** 2734-8.

SINGH, J. & PADGETT, R. A. 2009. Rates of in situ transcription and splicing in large human genes. *Nat Struct Mol Biol,* 16**,** 1128-33.

SKINNER, G. M., BAUMANN, C. G., QUINN, D. M., MOLLOY, J. E. & HOGGETT, J. G. 2004. Promoter binding, initiation, and elongation by bacteriophage T7 RNA polymerase. A single-molecule view of the transcription cycle. *J Biol Chem,* 279**,** 3239-44.

SKORDIS, L. A., DUNCKLEY, M. G., YUE, B., EPERON, I. C. & MUNTONI, F. 2003. Bifunctional antisense oligonucleotides provide a trans-acting splicing enhancer that stimulates SMN2 gene expression in patient fibroblasts. *Proc Natl Acad Sci U S A,* 100**,** 4114-9.

SMITH, L. D., LUCAS, C. M. & EPERON, I. C. 2013. Intron retention in the alternatively spliced region of RON results from weak 3' splice site recognition. *PLoS One,* 8**,** e77208.

SOBELL, H. M. 1985. Actinomycin and DNA transcription. *Proc Natl Acad Sci U S A,* 82**,** 5328-31.

SPECTOR, D. L., FU, X. D. & MANIATIS, T. 1991. Associations between distinct pre-mRNA splicing components and the cell nucleus. *EMBO J,* 10**,** 3467-81.

SPILUTTINI, B., GU, B., BELAGAL, P., SMIRNOVA, A. S., NGUYEN, V. T., HEBERT, C., SCHMIDT, U., BERTRAND, E., DARZACQ, X. & BENSAUDE, O. 2010. Splicing-independent recruitment of U1 snRNP to a transcription unit in living cells. *J Cell Sci,* 123**,** 2085-93.

SPRAGGON, L. & CARTEGNI, L. 2013. U1 snRNP-Dependent Suppression of Polyadenylation: Physiological Role and Therapeutic Opportunities in Cancer. *Int J Cell Biol,* 2013**,** 846510.

STALEY, J. P. & GUTHRIE, C. 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell,* 92**,** 315-26.

STARR, D. B., HOOPES, B. C. & HAWLEY, D. K. 1995. DNA bending is an important component of site-specific recognition by the TATA binding protein. *J Mol Biol,* 250**,** 434-46.

STEVENS, A. 1978. An exoribonuclease from Saccharomyces cerevisiae: effect of modifications of 5' end groups on the hydrolysis of substrates to 5' mononucleotides. *Biochem Biophys Res Commun,* 81**,** 656-61.

Su, Z., WANG, J., YU, J., HUANG, X. & GU, X. 2006. Evolution of alternative splicing after gene duplication. *Genome Res,* 16**,** 182-9.

SUN, X., ZHANG, Y., CHO, H., RICKERT, P., LEES, E., LANE, W. & REINBERG, D. 1998. NAT, a human complex containing Srb polypeptides that functions as a negative regulator of activated transcription. *Mol Cell,* 2**,** 213-22.

TAKAGAKI, Y., SEIPELT, R. L., PETERSON, M. L. & MANLEY, J. L. 1996. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell,* 87**,** 941-52.
TARDIFF, D. F. & ROSBASH, M. 2006. Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly. *RNA,* 12**,** 968-79.
THOMAS, M. J., PLATAS, A. A. & HAWLEY, D. K. 1998. Transcriptional fidelity and proofreading by RNA polymerase II. *Cell,* 93**,** 627-37.
TIAN, B., HU, J., ZHANG, H. & LUTZ, C. S. 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res,* 33**,** 201-12.
TIAN, H. 2001. RNA ligands generated against complex nuclear targets indicate a role for U1 snRNP in co-ordinating transcription and RNA splicing. *FEBS Lett,* 509**,** 282-6.
UJVARI, A. & LUSE, D. S. 2004. Newly Initiated RNA encounters a factor involved in splicing immediately upon emerging from within RNA polymerase II. *J Biol Chem,* 279**,** 49773-9.
UBD.BE, 2016 [online]. Available at: http://www.umd.be/HSF3/ [accessed 15.03. 2016].
VAGNER, S., RUEGSEGGER, U., GUNDERSON, S. I., KELLER, W. & MATTAJ, I. W. 2000. Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. *RNA,* 6**,** 178-88.
VAN GELDER, C. W., GUNDERSON, S. I., JANSEN, E. J., BOELENS, W. C., POLYCARPOU-SCHWARZ, M., MATTAJ, I. W. & VAN VENROOIJ, W. J. 1993. A complex secondary structure in U1A pre-mRNA that binds two molecules of U1A protein is required for regulation of polyadenylation. *EMBO J,* 12**,** 5191-200.
VANACOVA, S. & STEFL, R. 2007. The exosome and RNA quality control in the nucleus. *EMBO Rep,* 8**,** 651-7.
VANNINI, A. & CRAMER, P. 2012. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol Cell,* 45**,** 439-46.
VENKATARAMAN, K., BROWN, K. M. & GILMARTIN, G. M. 2005. Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev,* 19**,** 1315-27.
VINCENT, M., LAURIAULT, P., DUBOIS, M. F., LAVOIE, S., BENSAUDE, O. & CHABOT, B. 1996. The nuclear matrix protein p255 is a highly phosphorylated form of RNA polymerase II largest subunit which associates with spliceosomes. *Nucleic Acids Res,* 24**,** 4649-52.
VORLOVA, S., ROCCO, G., LEFAVE, C. V., JODELKA, F. M., HESS, K., HASTINGS, M. L., HENKE, E. & CARTEGNI, L. 2011. Induction of antagonistic soluble decoy receptor tyrosine kinases by intronic polyA activation. *Mol Cell,* 43**,** 927-39.
WANG, E. T., SANDBERG, R., LUO, S., KHREBTUKOVA, I., ZHANG, L., MAYR, C., KINGSMORE, S. F., SCHROTH, G. P. & BURGE, C. B. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature,* 456**,** 470-6.
WANG, Z., JIAO, X., CARR-SCHMID, A. & KILEDJIAN, M. 2002. The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proc Natl Acad Sci U S A,* 99**,** 12663-8.
WASSARMAN, D. A. & STEITZ, J. A. 1993. A base-pairing interaction between U2 and U6 small nuclear RNAs occurs in > 150S complexes in HeLa cell extracts:

implications for the spliceosome assembly pathway. *Proc Natl Acad Sci U S A,* 90**,** 7139-43.
WEBER, G., TROWITZSCH, S., KASTNER, B., LUHRMANN, R. & WAHL, M. C. 2010. Functional organization of the Sm core in the crystal structure of human U1 snRNP. *EMBO J,* 29**,** 4172-84.
WEINGARTNER, B. & KELLER, W. 1981. Transcription and processing of adenoviral RNA by extracts from HeLa cells. *Proc Natl Acad Sci U S A,* 78**,** 4092-6.
WELDON, C., BEHN-ANSMANT, I., BURLEY, G. A., HURLEY, L. H., BRANLANT, C., EPERON, C. I. & DOMINGUEZ, C. 2016. Identification of G-quadruplexes using 7-deaza-guanine substituted RNA. *Nat.Chem.Biol,* in press (21.08.2016)
WEST, S., GROMAK, N. & PROUDFOOT, N. J. 2004. Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature,* 432**,** 522-5.
WEST, S., PROUDFOOT, N. J. & DYE, M. J. 2008. Molecular dissection of mammalian RNA polymerase II transcriptional termination. *Mol Cell,* 29**,** 600-10.
WETTERBERG, I., BAUREN, G. & WIESLANDER, L. 1996. The intranuclear site of excision of each intron in Balbiani ring 3 pre-mRNA is influenced by the time remaining to transcription termination and different excision efficiencies for the various introns. *RNA,* 2**,** 641-51.
WHITE, E., KAMIENIARZ-GDULA, K., DYE, M. J. & PROUDFOOT, N. J. 2013. AT-rich sequence elements promote nascent transcript cleavage leading to RNA polymerase II termination. *Nucleic Acids Res,* 41**,** 1797-806.
WILL, C. L. & LUHRMANN, R. 2011. Spliceosome structure and function. *Cold Spring Harb Perspect Biol,* 3.
WINTERS, M. A. & EDMONDS, M. 1973a. A poly(A) polymerase from calf thymus. Characterization of the reaction product and the primer requirement. *J Biol Chem,* 248**,** 4763-8.
WINTERS, M. A. & EDMONDS, M. 1973b. A poly(A) polymerase from calf thymus. Purification and properities of the enzyme. *J Biol Chem,* 248**,** 4756-62.
WU, C. H., YAMAGUCHI, Y., BENJAMIN, L. R., HORVAT-GORDON, M., WASHINSKY, J., ENERLY, E., LARSSON, J., LAMBERTSSON, A., HANDA, H. & GILMOUR, D. 2003. NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. *Genes Dev,* 17**,** 1402-14.
WU, S., ROMFO, C. M., NILSEN, T. W. & GREEN, M. R. 1999. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature,* 402**,** 832-5.
WYERS, F., ROUGEMAILLE, M., BADIS, G., ROUSSELLE, J. C., DUFOUR, M. E., BOULAY, J., REGNAULT, B., DEVAUX, F., NAMANE, A., SERAPHIN, B., LIBRI, D. & JACQUIER, A. 2005. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell,* 121**,** 725-37.
YAMAGUCHI, Y., TAKAGI, T., WADA, T., YANO, K., FURUYA, A., SUGIMOTO, S., HASEGAWA, J. & HANDA, H. 1999. NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell,* 97**,** 41-51.
YAN, C., HANG, J., WAN, R., HUANG, M., WONG, C. C. & SHI, Y. 2015. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science,* 349**,** 1182-91.

YANKULOV, K., YAMASHITA, K., ROY, R., EGLY, J. M. & BENTLEY, D. L. 1995. The transcriptional elongation inhibitor 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole inhibits transcription factor IIH-associated protein kinase. *J Biol Chem,* 270**,** 23922-5.

YOGESHA, S. D., MAYFIELD, J. E. & ZHANG, Y. 2014. Cross-talk of phosphorylation and prolyl isomerization of the C-terminal domain of RNA Polymerase II. *Molecules,* 19**,** 1481-511.

YU, Y., DAS, R., FOLCO, E. G. & REED, R. 2010. A model in vitro system for co-transcriptional splicing. *Nucleic Acids Res,* 38**,** 7570-8.

YU, Y. & REED, R. 2015. FUS functions in coupling transcription to splicing by mediating an interaction between RNAP II and U1 snRNP. *Proc Natl Acad Sci U S A,* 112**,** 8608-13.

ZAMORE, P. D. & GREEN, M. R. 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc Natl Acad Sci U S A,* 86**,** 9243-7.

ZANDOMENI, R., ZANDOMENI, M. C., SHUGAR, D. & WEINMANN, R. 1986. Casein kinase type II is involved in the inhibition by 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole of specific RNA polymerase II transcription. *J Biol Chem,* 261**,** 3414-9.

ZAWEL, L., KUMAR, K. P. & REINBERG, D. 1995. Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev,* 9**,** 1479-90.

ZENG, C. & BERGET, S. M. 2000. Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing. *Mol Cell Biol,* 20**,** 8290-301.

ZHANG, G., TANEJA, K. L., SINGER, R. H. & GREEN, M. R. 1994. Localization of pre-mRNA splicing in mammalian nuclei. *Nature,* 372**,** 809-12.

ZHANG, M., YU, L., XIN, Y., HU, P., FU, Q., YU, C. & ZHAO, S. 1999. Cloning and mapping of the XRN2 gene to human chromosome 20p11.1-p11.2. *Genomics,* 59**,** 252-4.

ZHAO, J., HYMAN, L. & MOORE, C. 1999. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev,* 63**,** 405-45.

ZHOU, Q. & SHARP, P. A. 1996. Tat-SF1: cofactor for stimulation of transcriptional elongation by HIV-1 Tat. *Science,* 274**,** 605-10.

ZHUANG, Y. & WEINER, A. M. 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell,* 46**,** 827-35.

ZORIO, D. A. & BLUMENTHAL, T. 1999. Both subunits of U2AF recognize the 3' splice site in Caenorhabditis elegans. *Nature,* 402**,** 835-8.