THE EXOSOME AND HUMAN RIBOSOME BIOGENESIS

Katherine Sloan

B.Sc. (Hons) Biochemistry,

University of Manchester

Submitted for Doctor of Philosophy Degree

January 2012

Faculty of Medical Sciences

University of Newcastle upon Tyne

Abstract

Exoribonucleases have many important functions in the cell including RNA processing, turnover and quality control. One of the key 3'-5' exonucleases is the exosome, a multiprotein complex that has been extensively characterised in yeast. Many substrates that undergo maturation and/or degradation involving the yeast exosome have been identified and these include tRNAs, mRNAs, snRNAs, snoRNAs and rRNAs. By comparison, the human exosome is poorly understood and it is not clear whether functions of the yeast exosome are conserved in higher eukaryotes.

We show that the human exosome has degradation functions including the turnover, but not the processing, of snoRNAs and the recycling of excised pre-rRNA fragments. We and others have shown that the human exosome also participates in pre-rRNA processing to form the mature 3' end of 5.8S rRNA. Here we identify a novel role for the exosome in the processing of the pre-rRNA internal transcribed spacer 1 (ITS1). The small (18S) and large (5.8S and 28S) subunit rRNAs are co-transcribed as a single precursor. Processing of ITS1 is a key step in ribosome biogenesis as it separates 18S from the large subunit rRNAs and in higher eukaryotes it involves an additional processing step compared to yeast.

We define alternative ITS1 processing pathways in human cells. In the major pathway, following an endonucleolytic cleavage to separate the small and large subunit rRNAs, the exosome, which is not involved in ITS1 processing in yeast, processes to within 25 nucleotides of the 3' end of 18S. Our data highlight significant differences between the nucleases involved in ITS1 processing in yeast and humans. However, it appears that the roles of several yeast biogenesis factors are conserved in higher eukaryotes. Further, we have investigated mechanisms by which exonucleolytic processing of ITS1 may be regulated and suggest how this could be coordinated with the final maturation steps of the pre-40S complex.

Declaration and acknowledgements

I certify that this thesis contains my own work, except where acknowledged, and that no part of this material has been previously submitted for a degree or any other qualification at this or any other University.

I would like to thank my supervisor, Dr Nick Watkins, for his support, advice and encouragement throughout this project. My thanks also go to all the members of the Watkins lab, present and past, Dr Rob van Nues, Dr Kenneth McKeegan, Dr Amy Turner, Dr Andrew Knox, David Colvin, Dr Charles Debieux and Dr Hannah Richardson, for their valuable contributions. I am grateful to Dr Jeremy Brown for helpful discussions about this work. I would thank Professor David Tollervey (University of Edinburgh), Professor Ger Pruijn (Radbound University, Nijmegen) and Dr Henning Urlab (Max-Planck Institute, Gottingen) who collaborated with us on aspects of this research.

I particularly wish to acknowledge the Faculty of Medical Sciences and the Institute for Cell and Molecular Biosciences for their support, which enabled me to undertake this PhD. The BBSRC and Wellcome Trust funded this work.

Table of Contents

Chapter Two: Materials and methods

Chapter Three: The roles of human exonucleases in RNA processing and degradation in the nucleus

Chapter Four: Major and minor pre-rRNA processing pathways for internal transcribed spacer 1 in humans

Index of Figures

Index of Tables

Abbreviations and Acronyms

Amino acids

A	Alanine
C	Cysteine
D	Aspartic acid
H	Histidine
K	Lysine
L	Leucine
M	Methionine
P	Proline
S	Serine
T	Threonine
Y	Tryosine

Units

h	Hour
min	Minutes
M	Molar
m	Milli
µ	micro
k	kilo
Da	Dalton
S	Svedburg unit
b	Base
v/v	Volume per volume
w/v	Weight per volume
V	Volts

Proteins

GST-tag	Glutathione S transferase
His-tag	6x histidine residues
NES	Nuclear export sequence
C-terminal	Carboxyl terminal
N-terminal	Amino terminal
PIN	PilT N terminus
NLS	Nuclear localisation signal

Methods

IF	Immunofluorescence
IP	Immunoprecipitation
PCR	Polymerase chain reaction
UV	Ultraviolet
PAGE	Polyacrylamide gel electrophoresis

Nucleotides

nt	Nucleotide(s)
RNA	Ribose nucleic acid
DNA	Dioxyribose nucleic acid
A	Adenine
T	Thymine
C	Cytosine
G	Guanine
ATP	Adenosine triphosphate
ADP	Adenosine diphosphate
GTP	Guanosine triphosphate
GDP	Guanosine diphosphate
cDNA	Complementary DNA
siRNA	Small interfering RNA
miRNA	Micro RNA
mRNA	Messenger RNA
tRNA	Transfer RNA
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
PCR	Polymerase chain reaction
m_3G	2,2,7-trimethylguanosine
m_7G	7-monomethylguanosine
rRNA	Ribosomal RNA
Pre-rRNA	Precursor ribosomal RNA
rDNA	Ribosomal DNA
RNAi	RNA interference

Ribosome

SSU	Small subunit
LSU	Large subunit
ETS	External transcribed spacer
ITS	Internal transcribed spacer

Chemicals

LMB	Leptomycin B
5FU	5-Fluorouracil
DAPI	4’,6-diamidino-2-phenylindole

Chapter One

Introduction

1.1 Ribosomes

Ribosomes are highly conserved RNA-protein complexes that are responsible for translating the information encoded in mRNAs into proteins. Eukaryotic ribosomes have a sedimentation coefficient of 80S and are composed of two subunits, a large 60S subunit and a small 40S subunit. The large subunit (LSU) contains the sites of peptidyl transferase activity while the decoding function of the ribosome is carried out in the small subunit (SSU) (Lafontaine & Tollervey, 2001). The LSU of eukaryotic ribosomes contains three ribosomal RNA (rRNA) species, 28S (25S in yeast), 5.8S and 5S, along with approximately 46 ribosomal proteins (r-proteins) while the SSU is composed of a single rRNA, 18S, and 32 r-proteins. Crystal structures of eukaryotic ribosomes and subunits have recently been published showing the spatial organisation of rRNAs and r-proteins (Ben-Shem et al, 2011; Ben-Shem et al, 2010; Klinge et al, 2011; Rabl et al, 2010). Ribosome biogenesis is a complex process and involves the coordinated actions of more than 200 trans-acting protein factors including nucleases, chaperones, helicases, methyltransferases and transport factors (Henras et al, 2008). Ribosome biogenesis requires a great deal of the cells' energy with approximately 70 % of all transcription being directed for ribosome production (Warner, 1999).

The number and production rate of ribosomes determines the protein synthesis capacity of a cell, thereby coupling ribosome biogenesis and cellular growth rate. As a result of this, ribosome biogenesis is down-regulated during cellular differentiation and the levels of many ribosome biogenesis factors are also decreased following differentiation (Bowman & Emerson, 1977; Knox et al, 2011). In contrast, ribosome biogenesis is up-regulated in most cancer cells and several proto-oncogenes and tumour suppressor genes are able to directly regulate ribosome biogenesis (Oskarsson & Trumpp, 2005; Ruggero & Pandolfi, 2003). p53 has also been shown to be activated by ribosome dysfunction and cellular stress (Holzel et al, 2010). Inhibiting ribosome production increases the pool of free ribosomal proteins, such as RPL11, which then interact with HDM2, a key regulator of p53 (Sasaki et al, 2011). More specifically, the ribosome biogenesis factors, BOP1 and ENP, are upregulated in hepatocellular carcinoma and BOP1 also in colorectal carcinoma (Chung et al, 2011; Killian et al, 2006; Wang et al, 2009).

Defects in ribosome biogenesis also occur in several genetic diseases (Narla & Ebert, 2010). Germline mutations in various proteins found in the mature ribosome have been described in diseases characterised by erythroid deficiency including Diamond Blackfan anaemia (DBA) and 5-q syndrome, which are particularly associated with mutations in RPS19, RPS10, RPS24, RPS26 and RPS14 (Ebert et al, 2008). It has been shown that haploinsufficiency of ribosomal protein genes leads to specific p53 activation in erythroid progenitors and cell cycle arrest (Dutt et al, 2011). Mutations in other ribosome biogenesis factor genes have also been implicated in rare congenital syndromes including dyskeratosis congenita and Schwachman-Diamond syndrome (Burwick et al, 2011; Gupta & Kumar, 2010). The craniofacial disorder, Treacher Colins syndrome, is caused by mutations in the TCOF1 gene or mutations in the genes encoding subunits of RNA polymerase I or III (Dauwerse et al, 2011; Dixon et al, 2007; Valdez et al, 2004). The TCOF1 gene product, Treacle, is required both for transcription of the ribosomal DNA (rDNA) and modification of the rRNAs which are important for ribosome function. Other proteins which play important roles in the maturation of the RNA components of ribosomes have also been linked to genetic diseases. The RNA-protein enzyme, RNase MRP, which is important for rRNA processing in yeast, has been linked to CHH syndrome (cartilage-hair hypoplasia) (Mattijssen et al, 2010b; Ridanpaa et al, 2001) and a similar syndrome; ANE (alopecia, neurological defects and endocrinopathy), is caused by mutation of RBM28, the human homologue of a yeast ribosome biogenesis factor (Nousbeck et al, 2008; Spiegel et al, 2010).

1.2 rRNA organisation, transcription and processing

In eukaryotes, the 18S, 5.8S and 28S (25S in yeast) ribosomal RNAs are co-transcribed as a single precursor by RNA polymerase I while 5S is independently transcribed by RNA polymerase III (Nazar, 2004). In yeast, ~150 rDNA genes containing the co-transcribed rRNAs are found in tandem repeats on chromosome 12 whereas in human cells approximately 400 tandem rDNA repeats have been found on the short arms of acrocentric chromosomes (Kobayashi et al, 1998; Prieto & McStay, 2005). In yeast, 5S sequences are found immediately downstream of rDNA repeats while in higher eukaryotes, 5S genes are often found clustered in separate regions of the genome (Douet & Tourmente, 2007). Pre-rRNA transcription occurs within the nucleolus, a sub-compartment of the nucleus formed around nucleolar organiser regions (NORs) containing the rDNA repeats. Following rDNA transcription, the early steps of pre-rRNA processing and ribosome assembly also occur in the nucleoli (Hernandez-Verdun et al, 2010).

All of the rRNAs are synthesised as precursors and undergo processing to form the mature rRNAs. The precursor transcript containing 18S, 5.8S and 28 (25S in yeast) is 13kb or 47S long (35S in yeast). In addition to the mature rRNAs this transcript contains 5' and 3' external transcribed spacer regions (5'ETS and 3'ETS) as well as two internal transcribed spacers (ITS1 and ITS2). These spacer regions are removed by an ordered series of endonucleolytic cleavages and exonucleolytic processing steps to release the mature rRNAs. In yeast, 5S rRNA has been shown to undergo processing at both the 5' and 3' ends by exonucleases (van Hoof et al, 2000).

1.2.1 Pre-rRNA processing in *Saccharomyces cerevisiae*

The pathway of pre-rRNA processing has been studied most extensively in the yeast, *Saccharomyces. cerevisiae (S. cerevisiae)* (Granneman & Baserga, 2004; Henras et al, 2008) (Figure 1.1). Processing of the initial 35S precursor begins as the transcript is being produced in the nucleolus and the first steps that occur are endonucleolytic cleavages at the B_0 site in the 3'ETS, which is carried out by Rnt1 (Kufel et al, 1999), and at the A_0 site in the 5'ETS by an unknown endonuclease. This is followed by a cleavage at the 5' end of 18S (A_1) to remove the remaining 5'ETS and generate the 32S intermediate. Cleavages at sites A_2 and A_3 in ITS1 separate the large subunit (LSU) and small subunit (SSU) pre-rRNAs producing $27SA_2$/$27SA_3$ and 20S pre-rRNAs, respectively. 20S is transported to the cytoplasm where a final endonucleolytic cleavage at site D, carried out by Nob1, generates the mature 3' end of 18S (Lamanna & Karbstein, 2009; Pertschy et al, 2009).The A_2 cleavage site is an AC-rich sequence to the 5' side of a conserved stem-loop structure approximately 217 nucleotides away from the 3' end of 18S and cleavage at this site has recently been proposed to be carried out by Rcl1 (Horn et al, 2011). In addition to A_2 cleavage, the majority (approximately 85%) of the pre-rRNA is cleaved at a second site in ITS1, A_3, by RNase MRP (Lygerou et al, 1996; Schmitt & Clayton, 1993). The A_3 cleavage site is also an AC-rich sequence to the 3' side of a stem-loop structure and is located approximately 60 nucleotides upstream of the A_2 cleavage site. A_3 cleavage is followed by exonucleolytic processing by the 5'-3' exonucleases, Rat1 and Rrp17, to the B_{1S} site which represents the mature 5' end of $5.8S_S$ rRNA (Henry et al, 1994; Oeffinger et al, 2009). The remaining pre-rRNA, which is cleaved at A_2 but not A_3, is processed to produce an alternative form of 5.8S, $5.8S_L$, probably through an additional endonucleolytic cleavage at the B_{1L} site. The 3' ends of both $5.8S_S$ and $5.8S_L$ are produced by a common pathway. An endonucleolytic cleavage in ITS2 at site C_2, which occurs in the nucleus, produces 7S and $27SB_S$/$27SB_L$ (Michot et al, 1999) This is followed by further exonucleolytic processing

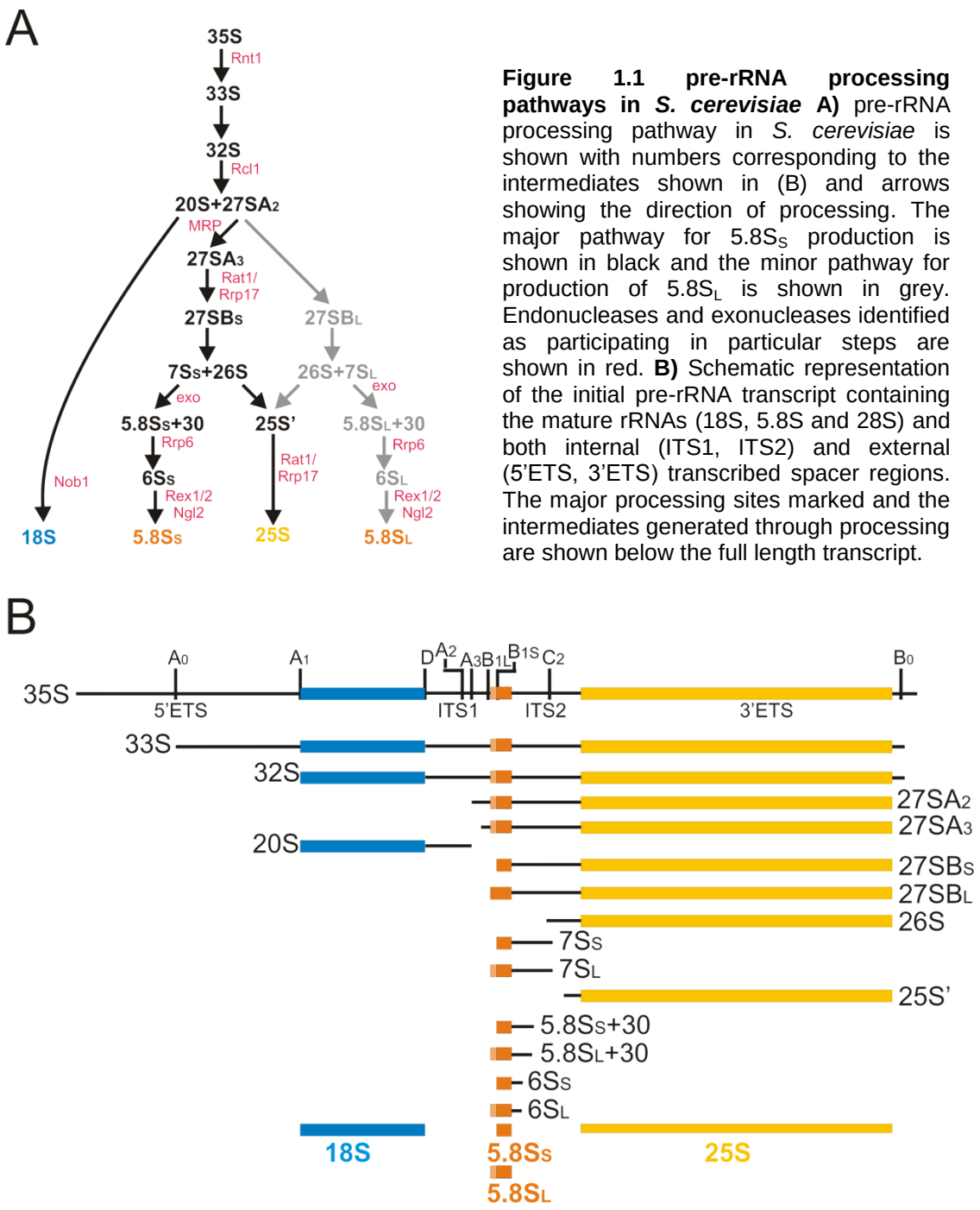


Figure 1.1 pre-rRNA processing pathways in *S. cerevisiae* A) pre-rRNA processing pathway in *S. cerevisiae* is shown with numbers corresponding to the intermediates shown in (B) and arrows showing the direction of processing. The major pathway for $5.8S_S$ production is shown in black and the minor pathway for production of $5.8S_L$ is shown in grey. Endonucleases and exonucleases identified as participating in particular steps are shown in red. **B)** Schematic representation of the initial pre-rRNA transcript containing the mature rRNAs (18S, 5.8S and 28S) and both internal (ITS1, ITS2) and external (5'ETS, 3'ETS) transcribed spacer regions. The major processing sites marked and the intermediates generated through processing are shown below the full length transcript.

in the 5'-3' direction by Rat1 and Rrp17 to generate the mature 5' end of 25S (Geerlings et al, 2000; Oeffinger et al, 2009). The 5.8S containing precursor generated by cleavage at site C_2 is processed by a 3'-5' multiprotein exonuclease complex called the exosome to produce a form of 5.8S with a 30 nucleotide 3' extension (5.8S+30). Exonucleolytic processing in both directions following C_2 cleavage has recently been shown to require Las1 (Schillewaert et al, 2011). 5.8S+30 is then converted to 6S by Rrp6, the RNase D-like subunit of the exosome, and the mature 3' ends of $5.8S_S$ and $5.8S_L$ are formed in the cytoplasm by the exonucleases Rex1, Rex2 and Ngl2 (Faber et al, 2002; Thomson & Tollervey, 2010). The mature 3' end of 25S is also produced by

exonucleolytic processing following Rnt1 cleavage (Oeffinger et al, 2009). In addition to their roles in the formation of the mature 5' and 3' ends of several rRNAs, exonucleases also function to degrade the spacer fragments released by serial endonucleolytic cleavages.

1.2.2 pre-rRNA processing in *Xenopus laevis*

Pre-rRNA processing has been studied in other eukaryotes, more complex than yeast and the processing pathway in the vertebrate, *Xenopus laevis (X. laevis)*, is well characterised (Savino & Gerbi, 1990). As in yeast, the 18S, 5.8S and 28S rRNAs, flanked by internal and external transcribed spacer regions are co-transcribed as a single precursor transcript (40S). Compared to yeast, an additional cleavage in the 5'ETS at a site called A' has been observed and cleavage at this site generates a 40S intermediate (Borovjagin & Gerbi, 2005). Removal of the internal transcribed spacers is proposed to occur by endonucleolytic cleavages at the mature 3' and 5' ends of 18S, 5.8S and 28S but the order of these cleavages is variable and two alternative pre-rRNA processing pathways have been shown to co-exist in a single oocyte (Figure 1.2) (Borovjagin & Gerbi, 2001).

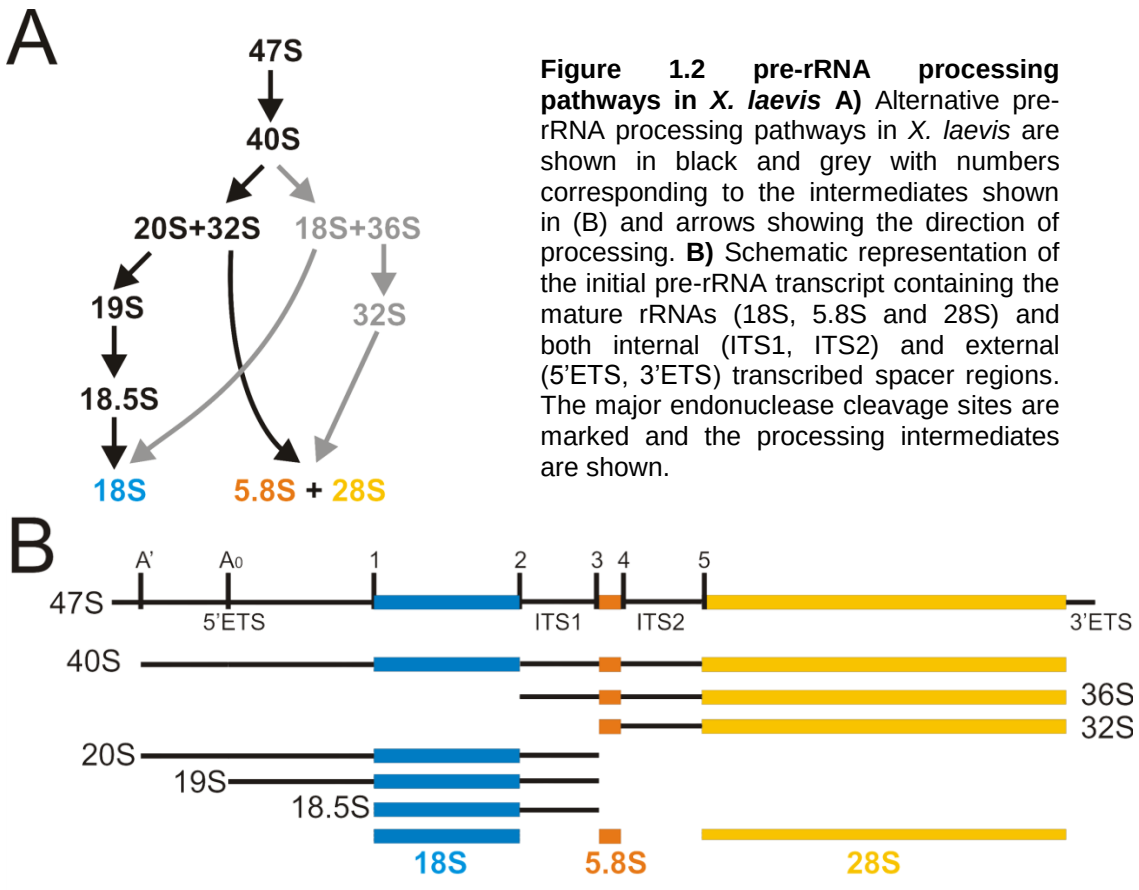


Figure 1.2 pre-rRNA processing pathways in *X. laevis* A) Alternative pre-rRNA processing pathways in *X. laevis* are shown in black and grey with numbers corresponding to the intermediates shown in (B) and arrows showing the direction of processing. **B)** Schematic representation of the initial pre-rRNA transcript containing the mature rRNAs (18S, 5.8S and 28S) and both internal (ITS1, ITS2) and external (5'ETS, 3'ETS) transcribed spacer regions. The major endonuclease cleavage sites are marked and the processing intermediates are shown.

Some *Xenopus* oocytes use only a single pathway in which cleavage at site 3 at the 5' end of 5.8S precedes cleavage at sites 1 and 2, which produce the mature ends of 18S, and cleavages 4 and 5 which release 5.8S and 28S from the 32S precursor. Alternatively, cleavages at sites 1 and 2 can occur prior to site 3 cleavage and this directly generates 18S and a different LSU precursor, 36S. Cleavages at sites 3, 4 and 5 then produce the mature 5.8S and 28S rRNAs.

1.2.3 Pre-rRNA processing in mammals

Less is known about pre-rRNA processing in mammals than in yeast. Early experimental data indicated a processing pathway broadly similar to that used in yeast and *Xenopus* (Hadjiolova et al, 1984) (Figure 1.3) but most of the cleavage sites in human pre-rRNA have not been accurately mapped nor have the enzymes required for specific steps been identified. In higher eukaryotes, two cleavage steps occur in the 5'ETS of the primary 47S transcript; A' and A_0 (Hadjiolova et al, 1993). A' cleavage is the only processing event thought to happen co-transcriptionally and this generates 45S. Following A' processing, cleavage at site 2 in ITS1 separates the SSU and LSU rRNAs producing the 30S and 32S pre-rRNAs, respectively. The 5'ETS is then removed from 30S to generate 21S either directly through simultaneous cleavages at sites A_0 and A_1, or, if A_0 precedes A_1, via a 26S intermediate. The order of these processing steps is variable and cleavages to remove the 5'ETS can occur prior to ITS1 cleavage, generating 41S which is then cleaved at site 2 to produce 21S and 32S. It is thought that two cleavage events occur in ITS2 (C_1, C_2) generating two precursors of 5.8S, 12S and 8S (Michot et al, 1999). As in yeast, the mature 3' end of 5.8S is generated by exonucleolytic processing involving the exosome. Similarly, exonucleolytic processing is thought to be required for formation of the 5' and 3' ends of 28S in higher eukaryotes (Wang & Pestov, 2011).

Production of 18S in mammals involves an additional processing step compared to yeast. An 18S precursor, 18SE, which has an extension beyond the 3' end of 18S of approximately 25 nt to site 2a, has been identified (Rouquette et al, 2005) (Figure 1.3B). However, it is unclear how this precursor is generated and different mechanisms have been postulated. In human cells, depletion of a late-acting SSU component, ENP1 (bystin), causes accumulation of an intermediate called 21SC, just smaller than 21S (Carron et al, 2011). A similar intermediate has also been observed upon depletion of the ribosomal protein, RPS19 (Idol et al, 2007). 21SC is also seen at very low levels in normal cells suggesting that it is a natural processing intermediate that is significantly accumulated in the absence of ENP1 or RPS19. The 3'

end of 21SC is very heterogeneous implying that it and possibly 18SE are generated by exonucleolytic processing (Carron et al, 2011). However, it has also been suggested that 18SE is produced directly by an endonucleolytic cleavage. Depletion of XRN2 from mouse cells caused a significant accumulation of 36S, an intermediate generated by an endonucleolytic cleavage at site 2a before site 2. Also observed when XRN2 was depleted was a released fragment of ITS1 generated by sequential cleavages at sites 2 and 2a (Wang & Pestov, 2011). It is unclear, therefore, how ITS1 is processed and 18SE is formed. Cleavage at site 2 appears to be the initial cleavage in ITS1 that separates the SSU and LSU rRNAs and early data using RNase protection mapping of pre-rRNA from rat liver cells located this cleavage to a poly(T) tract approximately 160 nucleotides from the 5' end of 5.8S (Hadjiolova et al, 1984) but this sequence is not conserved in human pre-rRNA. In human cells, depletion of Dicer, the endonuclease linked to miRNA and siRNA production, influenced cleavages in this region but the effects observed were minor so it seems unlikely that Dicer is the endonuclease responsible for site 2 cleavage (Liang & Crooke, 2011).

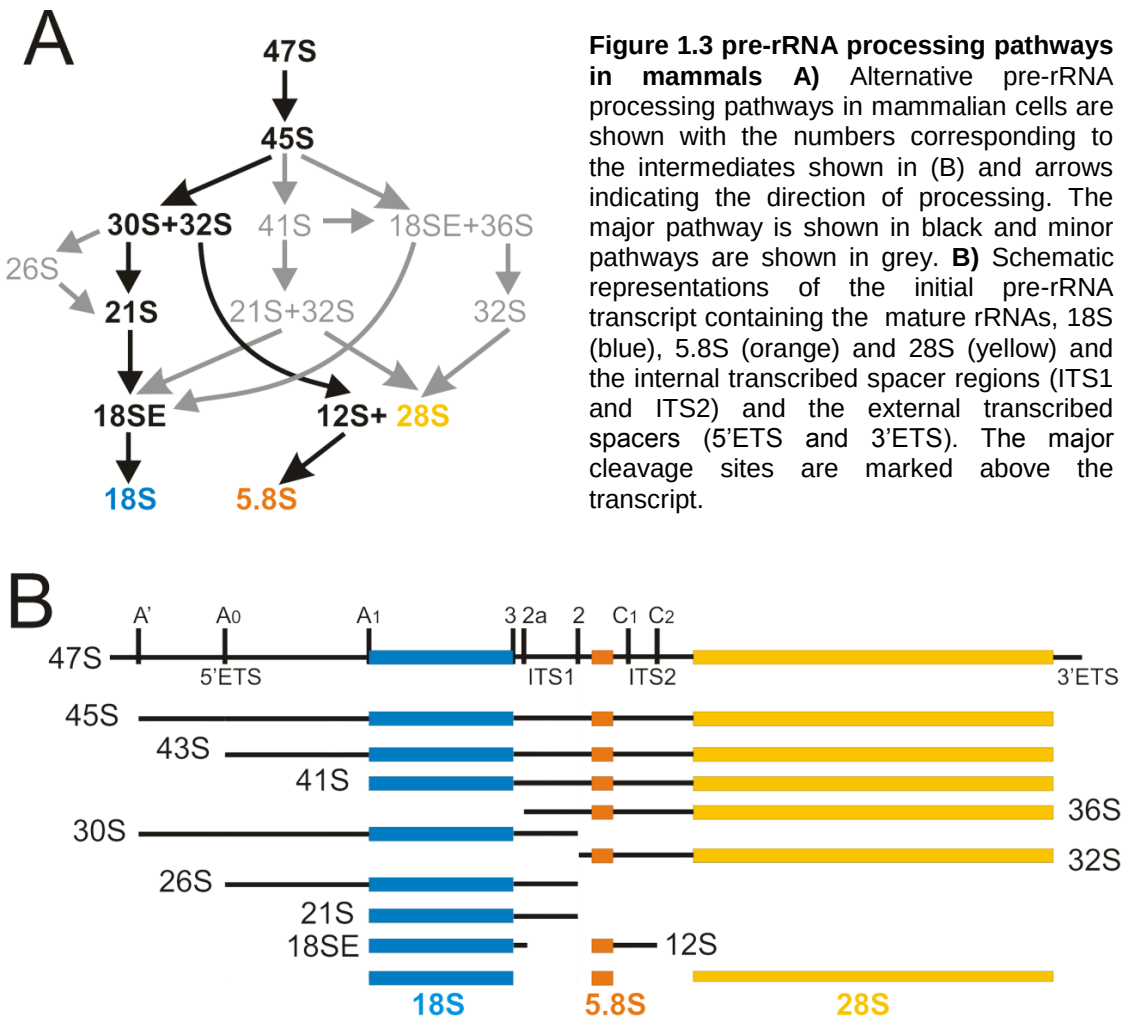


Figure 1.3 pre-rRNA processing pathways in mammals A) Alternative pre-rRNA processing pathways in mammalian cells are shown with the numbers corresponding to the intermediates shown in (B) and arrows indicating the direction of processing. The major pathway is shown in black and minor pathways are shown in grey. **B)** Schematic representations of the initial pre-rRNA transcript containing the mature rRNAs, 18S (blue), 5.8S (orange) and 28S (yellow) and the internal transcribed spacer regions (ITS1 and ITS2) and the external transcribed spacers (5'ETS and 3'ETS). The major cleavage sites are marked above the transcript.

1.2.4 ITS1 processing

In all eukaryotes, cleavages in ITS1 are particularly important in ribosome biogenesis as they separate the small subunit rRNA, 18S, from the large subunit rRNAs, 5.8S and 28S (25S in yeast). After cleavage in ITS1, biogenesis of the small and large subunits follows separate pathways until the mature 40S and 60S subunits become associated in the cytoplasm to form the ribosome. Preliminary data suggest that processing of ITS1 in higher eukaryotes may be significantly different from yeast and how separation of the LSU and SSU rRNAs is achieved may represent a key change in pre-rRNA processing through evolution.

1.3 rRNA modifications and snoRNPs

In addition to processing from precursors, rRNAs undergo extensive covalent modification both co-transcriptionally and post-transcriptionally (Kos & Tollervey, 2010; Terns & Terns, 2002). In excess of 200 nucleotides in 18S and 25S (28S in humans) have been shown to undergo modification. Such modifications involve either 2'-O-methylation of specific nucleotides or conversion of uridine residues to pseudouridine (Figure 1.4A). Methylation can stabilize base-pairing interactions within the ribosome and is important for mediating RNA folding while pseudouridinylation has been shown to help maintain RNA secondary structures and aid interactions between the rRNA and particular ribosomal proteins (Decatur & Fournier, 2002; Decatur & Fournier, 2003; Helm, 2006). Individual modifications increase the cell's sensitivity to certain antibiotics and stress; particular groups of modifications have been shown to be important for cell growth (Baudin-Baillieu et al, 2009; Baxter-Roshek et al, 2007; Esguerra et al, 2008; Piekna-Przybylska et al, 2008). Modification sites are generally highly conserved. They are found in functionally important domains of the mature rRNAs such as the peptidyl transferase domain and the interfaces between the large and small subunits, underlining the key regulatory role they can play (Baudin-Baillieu et al, 2009; Decatur & Fournier, 2002; Liang et al, 2007). In the absence of these modifications, non-functional ribosomes are produced.

rRNA modifications are guided by a number of small nucleolar ribonucleoprotein complexes (snoRNPs) (Watkins & Bohnsack, 2011). snoRNPs are basically comprised of four core proteins and a small RNA component which guides interactions with the pre-rRNA by base-pairing with the target sites to direct the modification (Henras et al, 2004; Kiss-Laszlo et al, 1998; Terns & Terns, 2002). The core snoRNP proteins provide a scaffold for snoRNA folding and the catalytic activities are responsible for carrying out the modification. Most snoRNPs can be divided into

one of two classes depending on the type of modification carried out. Box C/D snoRNPs direct 2'-O-methylation and box H/ACA snoRNPs catalyse pseudouridinylation (Figure 1.4B, C) (Reichow et al, 2007). Instead of catalysing such nucleotide modifications, several snoRNPs are also required for particular pre-rRNA processing steps (Section 1.3.5) (Henras et al, 2008). It has recently been proposed that snoRNPs may also have roles in other cellular processes such as regulating gene expression and mediating stress responses (Bratkovic & Rogelj, 2010).

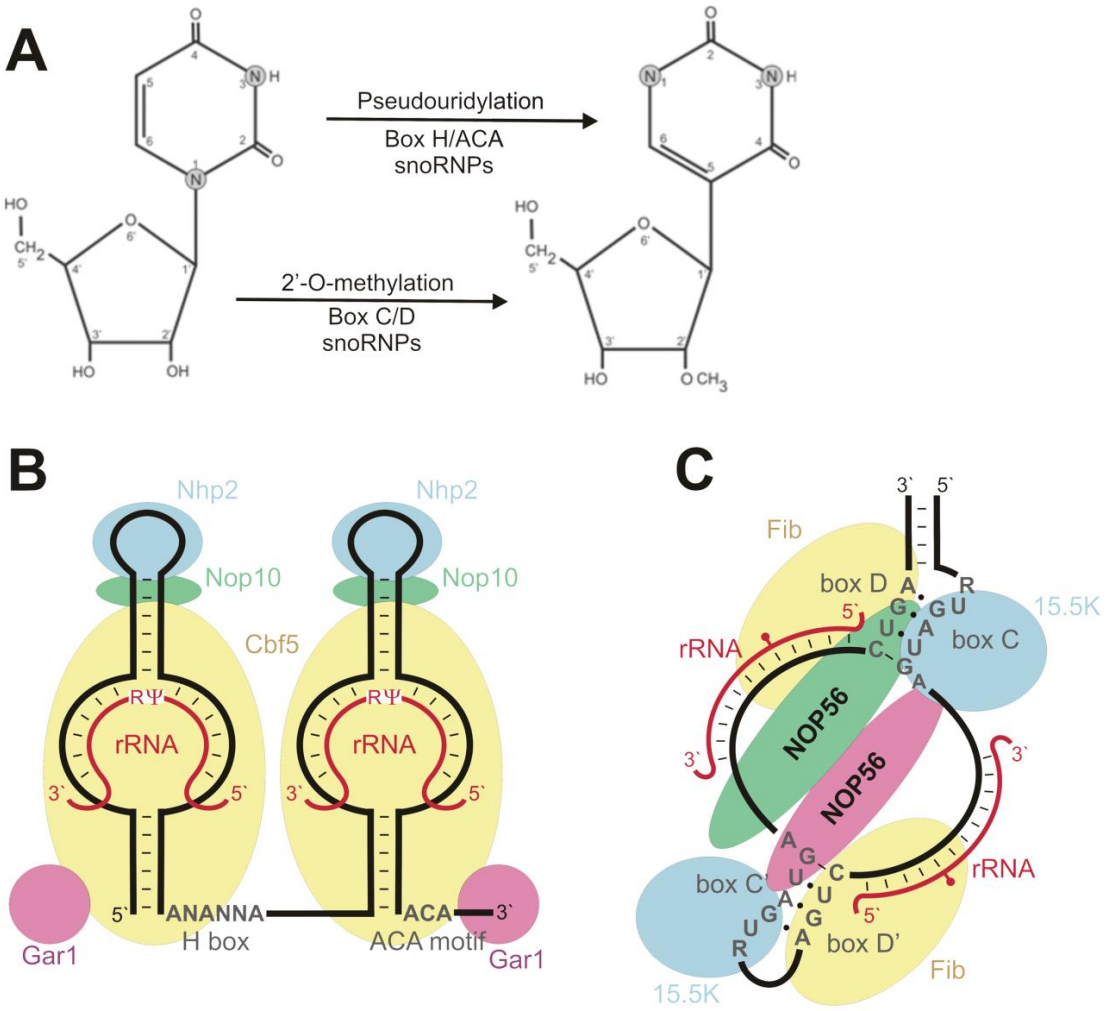


Figure 1.4 rRNA modifications and snoRNP structures A) Conversion of uridine to pseudouridine and 2'-O-methylation are shown. Only the sugar moiety of the ribose is shown without the associated phosphates at the 3' and 5' carbons. Figure based on previous work (Henras et al., 2008). **B)** Schematic representation of the eukaryotic box H/ACA snoRNP which forms a double stem-loop structure. The snoRNA and rRNA are represented by black and red lines respectively, with the nucleotide converted to pseudouridine indicated (Rψ). The H and ACA motifs are shown and the relative binding sites of the core snoRNP proteins are given. **C)** Schematic representation of the eukaryotic box C/D snoRNP which forms stem-loop structure. The snoRNA and rRNA are represented by black and red lines respectively, with the methylated nucleotide indicated. The C, D, C' and D' motifs are shown and the relative binding sites of the core snoRNP proteins are given. Adapted from Watkins and Bohnsack, 2011.

1.3.1 Box H/ACA snoRNPs

Box H/ACA snoRNPs are composed of the snoRNA and four common core proteins, NHP2, GAR1, NOP10 and Dyskerin (Cbf5 in yeast) (Figure 1.4B) (Henras et al, 1998; Lafontaine et al, 1998a). Dyskerin is the active component responsible for catalysing the isomerisation of the targeted uridine residue (Lafontaine et al, 1998a). Box H/ACA snoRNAs are identified by the presence of two conserved sequence elements which form a distinctive double hairpin structure. Box H contains the motif, ANANNA (with N representing any nucleotide) and is found in a single stranded region between the two stem-loops. The ACA motif is found downstream of the second hairpin, 3 nucleotides from the 3' end of the mature snoRNA (Figure 1.4B). Box H/ACA snoRNAs target modifications by base-pairing of sequences within either or both of the stem-loops to the pre-rRNA. This exposes the target uridine, enabling Dyskerin to carry out isomerisation (Reichow et al, 2007). The four core proteins can be associated with either or both of the stem-loops enabling box H/ACA snoRNPs to catalyse one or two modifications approximately 14-16 nucleotides upstream of either the H and/or the ACA boxes (Ganot et al, 1997; Ni et al, 1997).

1.3.2 Box C/D snoRNP structure

Box C/D snoRNAs also associate with four core proteins to form snoRNP complexes; 15.5K (Snu13 in yeast), NOP56, NOP58 and fibrillarin (Nop1 in yeast) (Figure 1.4C) (Baserga et al, 1991; Lafontaine & Tollervey, 1999; Lafontaine & Tollervey, 2000; Lyman et al, 1999; Tyc & Steitz, 1989; Watkins et al, 2000). Fibrillarin is the methyltransferase responsible for transferring a methyl group from the snoRNP cofactor, S-adenosyl-L-methionine (AdoMet), to the 2' hydroxyl of the targeted RNA nucleotide (Galardi et al, 2002). Box C/D snoRNAs are identified by conserved sequence motifs; the C box (RUGAUGA) which is found at the 5' end, the D box (CUGA) at the 3' end and the related but less well conserved C' and D' boxes. An overall stem-loop structure is normally formed as the 5' and 3' ends of the snoRNA base-pair to bring the C and D boxes into close proximity (Figure 1.4C). Upstream of the D box is a sequence capable of base-pairing with the target RNA to direct methylation of the nucleotide bound 5 bases upstream of the D box (Kiss-Laszlo et al, 1996; Kiss-Laszlo et al, 1998). Additional conserved sequences have also been identified in both yeast and human snoRNAs that are complementary to regions adjacent to rRNA modification sites and this extra-base pairing has been shown to support methylation (van Nues et al, 2011).

The C, D, C' and D' motifs are binding sites for the core snoRNP proteins which assemble asymmetrically on the snoRNA backbone. 15.5K is a part of the L7Ae protein family of which the box H/ACA core protein, NHP2 is also a member. 15.5K is the first core protein associated with the snoRNA and its binding to the box C/D and possibly C'/D' regions is thought to bring about a conformational change that generates binding sites for the other core snoRNP proteins (Watkins et al, 2002). Specific interaction between 15.5K and the box C/D of a typical snoRNA has been demonstrated *in vitro* (Szewczak et al, 2002; Watkins et al, 2000). NOP56 and NOP58 contain NOP domains which have been proposed to be important for RNP association (Gautier et al, 1997). Using site-specific cross-linking in *Xenopus*, NOP56 has been shown to associate with the C box and NOP58 with the C' box leading to a model in which these two proteins form a bridge between the box C/D and box C'/D' motifs (Cahill et al, 2002). snoRNP complexes are well conserved and several crystal structures of archaeal snoRNP complexes have been recently been published giving insight into the likely structure of eukaryotic snoRNPs. Different models have been proposed but the structure that fits most experimental data depicts a monomeric complex containing a single sRNA, two molecules of fibrillarin and 15.5K and a single copy of NOP56 and NOP58 (Bleichert et al, 2009; Lin et al, 2011; Xue et al, 2010; Ye et al, 2009). In addition to these core proteins, the eukaryotic box C/D snoRNA, U3 is also associated with another protein, hU3-55K. This has been shown to interact with a U3-specific box B/C motif and is required for recruiting the U3 snoRNP for its role in pre-rRNA processing (Granneman et al, 2004; Knox et al, 2011).

1.3.3 snoRNA maturation

snoRNA genes are found throughout the genome and are transcribed in a variety of different ways (Dieci et al, 2009). snoRNAs can be transcribed independently from a designated promoter either individually or as clusters of multiple snoRNAs. Alternatively, snoRNA genes are found within the introns of protein coding genes that are transcribed by RNA Polymerase II and the pre-snoRNAs are released during splicing. In yeast, the majority of snoRNA genes are independently transcribed, which is consistent with the lack of intronic DNA in lower eukaryotes and most are monocistronic with only a few polycistronic clusters having been identified. In mammals, however, approximately 90% of snoRNAs are intronic, although some, particularly those required for pre-rRNA processing rather than modification, are independently transcribed.

snoRNAs encoded within the introns of protein coding genes are processed from pre-mRNAs in the nucleus by either a splicing dependent or splicing independent pathway. In the splicing-dependent pathway, the snoRNA is released as an intron-lariat excised by the splicesome bringing the two exon-intron junctions together. Endonucleolytic cleavages in the intronic regions can also release the snoRNAs independently of splicing. The released intron fragment then undergoes further processing at both the 5' and 3' ends to produce the mature snoRNA. In yeast, this processing is mediated by exonucleases including Rat1 (XRN2) and the exosome (Allmang et al, 1999a) but it is not yet clear if this is also the case in human cells.

Independently transcribed snoRNAs contain a co-transcriptionally added 7-methylguanosine (m^7G) cap at the 5' end. During the maturation of many snoRNAs, including U3 and U8, this cap is hypermethylated by the methyltransferase, Tgs1, to form a tri-methylguanosine (m_3G) cap (Verheggen et al, 2002). In yeast, some pre-snoRNAs are, however, decapped and undergo 5' processing by exonucleases including Rat1 and Xrn1 (Petfalski et al, 1998). Both mono- and polycistronic yeast snoRNAs are cleaved by the endonuclease, Rnt1, leaving short 3' precursor extensions (Chanfreau et al, 1998b; Kufel et al, 1999). These often contain poly(U) tracts which are stabilised by Lhp1 (yeast homologue of the human La protein) and the LSm complex (Kufel et al, 2003). When these proteins dissociate, the 3' end is processed by the exosome to form the mature 3' end of the snoRNA (Allmang et al, 1999a). The mature snoRNA accumulates in the nucleolus (Boulon et al, 2004). It is also not clear if this pathway of snoRNA maturation is conserved from yeast to humans.

In human cells, it has been demonstrated that the methyltransferase, Tgs1, the La protein, the LSm complex and the exosome are stably associated with the U3 and U8 pre-snoRNPs in HeLa cells (Watkins et al, 2004; Watkins et al, 2007). La was found to associate with early pre-snoRNP complexes whereas the LSm complex and the exosome subunit, RRP46, were found to interact with later pre-snoRNAs suggesting that La binds first, stabilising the 3' end and is then replaced by the LSm complex. The exosome is then recruited presumably to carry out processing of the 3' extended sequence (Watkins et al, 2004).

1.3.4 snoRNP biogenesis

Concurrent with snoRNA processing, snoRNP biogenesis involves recruitment of the core proteins and nucleolar localisation. This involves dynamic association and dissociation of a number of protein factors which are not found in the mature snoRNP.

In addition to the RNA processing factors discussed above, this biogenesis complex includes proteins required for snoRNP assembly (TIP48, TIP49, NUFIP, NOP17, TAF9 and BCD1) and chaperones (HSP90 and HSC70) (Boulon et al, 2008; McKeegan et al, 2007; McKeegan et al, 2009; Watkins et al, 2004; Watkins et al, 2007). These assembly factors have been shown to form a scaffold for core protein assembly, with NUFIP regulating TIP48 and TIP49 bridging the interaction between the NOP proteins and 15.5K (McKeegan et al, 2007; McKeegan et al, 2009). Formation of the pre-snoRNP complex begins at the site of transcription in the nucleus and many factors become associated in the Cajal bodies before the snoRNP is localised to the nucleolus. This requires the involvement of many transport factors and PHAX, CRM1, RAN, NOPP140 and Snurportin1 have been shown to be required for snoRNP biogenesis (Boulon et al, 2004; Watkins et al, 2004).

1.3.5 snoRNPs in pre-rRNA processing

Instead of a role in mediating base modifications, some snoRNPs are required for specific pre-rRNA processing events. Unlike snoRNPs required for post-transcriptional modifications, most snoRNPs required for processing are essential.

U3 snoRNP is the most abundant snoRNP and is important for 18S production as it is required for cleavage steps in the 5'ETS (A_0 and A_1) and ITS1 (A_2) in yeast (Borovjagin & Gerbi, 2001; Granneman et al, 2004; Kass et al, 1990; Savino & Gerbi, 1990; Sharma & Tollervey, 1999). In human cells, the U3 snoRNP is also required for the additional 5'ETS cleavage at A' (Prieto & McStay, 2005). U3 snoRNA is classified as a box C/D snoRNA but instead of the usual secondary C'/D' box, it contains a unique B/C box which provides binding sites for a U3-specific protein, hU3-55K (Granneman et al, 2002; Lubben et al, 1993; Pluk et al, 1998; Tyc & Steitz, 1989; Venema et al, 2000). Due to a poorly conserved C' box (equivalent to the C box of other snoRNAs), the levels of U3 are regulated by association with hU3-55K (Knox et al, 2011). The 5' end of U3 snoRNA also contains an A'/A motif which mediates base-pairing with sequences in 18S rRNA. Additional interactions between U3 and the pre-rRNA are made between the 3' and 5' U3 hinges and sequences near the A' and A_0 pre-rRNA cleavages sites, respectively. It is suggested that these interactions may be required to keep the pre-rRNA in a conformation exposing cleavage sites to the action of endonucleases (Borovjagin & Gerbi, 2001). Further, base-pairing of U3 brings the pre-rRNA into a conformation in which the 5' and 3' ends of 18S are in close proximity which may be essential in co-ordinating their processing. U3-18S base pairing also provides a level of regulation by preventing premature formation of key structural folds

such as the central pseudo-knot found in the mature rRNA (Borovjagin & Gerbi, 2001; Gerbi et al, 2003; Hughes, 1996; Sharma et al, 1999).

The box C/D snoRNP, U14, is highly conserved across eukaryotes. In yeast, U14 has been shown to base-pair with 18S and is required for cleavage steps either side of 18S (A_1 and A_2) (Lempicki et al, 1990; Li et al, 1990). Similarly, the box H/ACA snoRNP, U17 (snR30 in yeast, E1 in *X. laevis*) has been shown to be required for processing in the 5'ETS and ITS1 (A_0, A_1 and A_2) (Enright et al, 1996; Lemay et al, 2011). Another box H/ACA snoRNA which is required for these cleavages is snR10. Although this snoRNA is not essential, deletion of this gene is synthetically lethal with mutations in Rrp5 or Rok1 implying its function is important (Morrissey & Tollervey, 1993).

Other snoRNPs which have roles in pre-rRNA processing specifically in higher eukaryotes are U8 and U22. In *X. laevis* U8 has been shown to be essential for 5.8S and 28S production. U8 is proposed to base pair with pre-rRNA both in 28S and at the 5.8S-ITS2 junction (Michot et al, 1999) and depletion of U8 leads to defects in removal of the 3'ETS (Peculis, 1997; Peculis & Steitz, 1993; Peculis & Steitz, 1994). The helicase, DDX51 is responsible for the dissociation of U8 from the pre-rRNA enabling formation of interactions between 5.8S and 28S that are found in the mature ribosome (Srivastava et al, 2010). Also in *X. laevis*, U22 has been shown to be required for processing events at both the 5' and 3' ends of 18S (Tycowski et al, 1994; Tycowski et al, 1996). Two further snoRNPs specific to higher eukaryotes, E2 and E3 in *X. laevis*, are implicated in the processing of ITS1

Most of the snoRNPs involved in pre-rRNA processing are devoid of any modification activity although U14 is thought to have maintained both functions (Atzorn et al, 2004; Torchet & Hermann-Le Denmat, 2002). In addition, RNase MRP is an unusual snoRNP complex that is also required for pre-rRNA processing rather than modification and is discussed in greater detail in section 1.9.2.

1.4 Assembly of ribosomal subunits

In the nucleolus, the pre-rRNA transcript becomes associated with a subset of ribosomal proteins and many trans-acting factors not found in the mature ribosome to form a 90S pre-ribosomal particle (Bernstein et al, 2004; Dragon et al, 2002; Grandi et al, 2002). Many of these trans-acting factors, including proteins and snoRNPs are recruited as part of a large complex called the SSU processome, which associates with the pre-rRNA concurrent with its transcription by RNA polymerase I. Compaction of the

5’ end of the transcript and association of these early processing factors with the nascent pre-rRNA transcript can be visualised by electron microscopy of chromatin spreads as terminal balls (Miller & Beatty, 1969; Mougey et al, 1993; Osheim et al, 2004). Approximately 30% of pre-rRNAs are fully transcribed with cleavage in the 3’ETS bringing about their release; alternatively, cleavages in the ITS1 can occur co-transcriptionally releasing a pre-40S complex (Kos & Tollervey, 2010; Osheim et al, 2004). Consistent with the majority of pre-rRNAs being cleaved co-transcriptionally, most of the processing factors associated with initial 90S particles are involved in the biogenesis of the small subunit. Many of the factors required for formation of the large subunit via the pre-60S complex are only assembled onto the pre-rRNA after cleavage at site A_2 in ITS1 (Grandi et al, 2002; Schafer et al, 2003; Tschochner & Hurt, 2003). Following cleavage in ITS1, the pre-40S and pre-60S complexes follow separate biogenesis pathways, with numerous factors associating and dissociating as the complexes are translocated through the nucleolus, the nucleoplasm and into the cytoplasm. The final maturation steps of both subunits occur in the cytoplasm before the subunits interact to form the mature ribosome.

1.4.1 Formation of the 90S particle and the SSU processome

The SSU processome is a dynamic complex of more than 40 proteins that is assembled onto the pre-rRNA as it is transcribed forming the 90S complex. The yeast SSU processome is required for cleavages on either side of 18S in both the 5’ETS and ITS1 yielding the pre-40S complex that matures into the small ribosomal subunit. The SSU processome from *S. cerevisiae* has been purified and key complexes identified but more recently, large scale proteomic studies have revealed the protein composition (Bernstein et al, 2004; Dragon et al, 2002; Grandi et al, 2002; Schafer et al, 2003). The proteins comprising the SSU processome in human cells are less well characterised although many of the factors involved appear to be conserved (Gerus et al, 2010; Granneman et al, 2003; Granneman et al, 2002; Rouquette et al, 2005; Turner et al, 2009; Turner et al, 2012).

In both yeast and higher eukaryotes, the SSU processome is composed of several major sub-complexes; t-UTP (transcriptional associated U three protein) complex, b-UTP complex, c-UTP complex, U3 snoRNP and the MPP10 (M phase phosphoprotein 10) complex. These are each independently assembled prior to hierarchal association with the pre-rRNA transcript (Figure 1.5). The first of these sub-complexes to become associated with the nascent transcript is the t-UTP complex containing five to seven proteins: UTP10, UTP4, UTP5, UTP15 and UTP17 with UTP8

and UTP9 additionally found in the yeast complex (Gallagher et al, 2004; Krogan et al, 2004; Prieto & McStay, 2007). The t-UTP complex is associated with the rDNA and this is required for efficient transcription of the pre-rRNA (Prieto & McStay, 2007). In HeLa cells, depletion of t-UTP proteins, UTP4 or UTP10, leads to formation of a novel processome intermediate, 50S, which is also observed when rRNA transcription is blocked by treating cells with actinomycin D supporting the role of the t-UTP complex in enabling transcription (Turner et al, 2009). Binding of the t-UTP complex to the pre-rRNA in yeast is a pre-requisite for the recruitment of other complexes which follow two independent assembly branches (Perez-Fernandez et al, 2011; Perez-Fernandez et al, 2007).

In one branch, the b-UTP complex, composed of PWP2, UTP6, UTP12, UTP13, UTP18 and UTP21 proteins, becomes associated with the pre-rRNA (Dosil & Bustelo, 2004; Perez-Fernandez et al, 2007). It is proposed that the b-UTP complex may be primarily a structural component of the SSU processome since many of its proteins contain common protein-protein interaction domains (WD or TPR) (Dosil & Bustelo, 2004). Following association of the b-UTP complex, the U3 snoRNP is recruited to the pre-ribosomal complex. Base pairing between the U3 snoRNA and the pre-rRNA is not required for this incorporation but instead hU3-55K is proposed to play an important role in recruitment (Granneman & Baserga, 2004). As mentioned in section 1.3.5, base pairing between the 5' end of 18S and the U3 snoRNA prevents premature formation of the central 18S pseudoknot and the MPP10 complex is proposed to enhance this base pairing (Borovjagin & Gerbi, 2001; Gerczei & Correll, 2004; Gerczei et al, 2009; Granneman et al, 2003; Hughes, 1996; Lafontaine & Tollervey, 2001). In both yeast and humans, MPP10 has been shown to associate with IMP3 and IMP4 (Granneman & Baserga, 2003; Lee & Baserga, 1999). This heterotrimeric complex is able to bind directly to the U3 snoRNA and *in vitro* this has been shown to alter the structure of the U3 snoRNA such that specific base pairing interactions with the pre-rRNA are enhanced (Gerczei & Correll, 2004; Gerczei et al, 2009). Recruitment of the MPP10 complex, but not of the U3 snoRNP, to the pre-ribosome is dependent on the presence of the GTPase, Bms1 (Perez-Fernandez et al, 2011). Bms1 associates with the 90S particle downstream of the U3 snoRNP and the b-UTP complex and the incorporation of Bms1 has been shown to stabilise the interaction of the b-UTP complex with the pre-rRNA. Bms1 is proposed to form a stable sub-complex with the endonuclease, Rcl1, and together they are able to interact with the U3 snoRNP (Karbstein & Doudna, 2006; Karbstein et al, 2005; Wegierski et al, 2001). Rcl1 has recently been proposed as the endonuclease responsible for A_2 cleavage in ITS1 and is discussed in more detail in Section 1.9.3 (Horn et al, 2011). A mechanism by which these proteins are recruited to

the early pre-ribosome is proposed: GTP-bound Bms1 interacts with Rcl1 and they are recruited to the pre-rRNA through interactions with U3, which in turn alters the conformation of Bms1 causing GTP hydrolysis, enabling Bms1 to dissociate, leaving Rcl1 deposited on the pre-rRNA (Karbstein & Doudna, 2006; Karbstein et al, 2005; Wegierski et al, 2001). The presence of Bms1 has also recently been shown to be required for the association of a number of other proteins including Utp20, Kre33 and Enp2 (Perez-Fernandez et al, 2011).

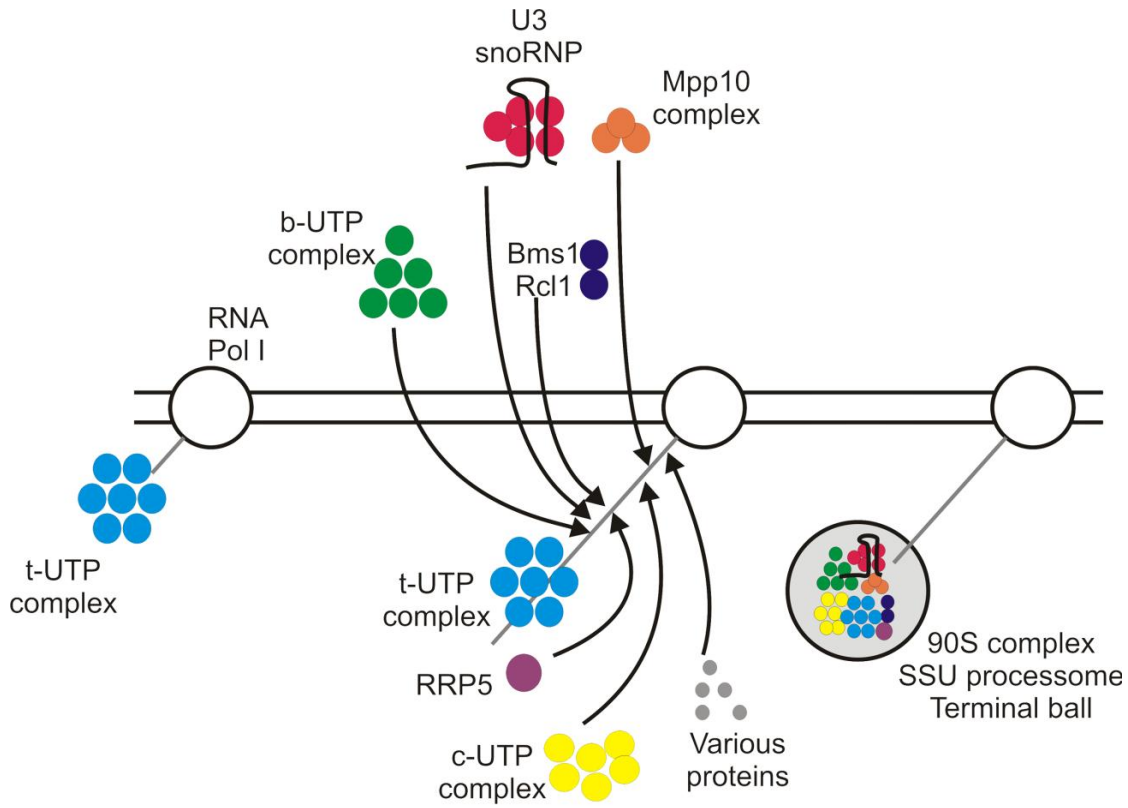


Figure 1.5 pre-rRNA transcription and SSU processome assembly Schematic model of transcription of the pre-rRNA transcript by RNA polymerase I (white circle and grey line) and early steps in the assembly of the 90S pre-ribosomal particle in *S. cerevisiae*. The t-UTP complex is associated first and then assembly occurs in two branches; the first involves association of the b-UTP complex followed by U3 snoRNP, Bms1-Rcl1 and finally the Mpp10 complex while in the second branch Rrp5 is recruited followed by the c-UTP complex. In addition to these sub-complexes a number of other proteins, mostly required for formation of the SSU, become associated and these include Pno1, Mrd1, Dbp8, Utp2, Utp11, Utp19, Utp23, Dim1, Krr1, Emg1, Has1, Utp3, Utp14, Utp20, Utp24, Rio1, Rio2, Nob1, Esf2, Sof1, Utp7, Utp16, Utp22, Dhr1 and Nsr1 (nucleolin). These proteins form the 90S complex which undergoes compaction to form terminal balls visible by electron microscopy. Based on data from references within sections 1.4 and 1.4.1.

The second branch of factors assembled following incorporation of the t-UTP complex into the pre-ribosomal particle is initiated by the association of Rrp5 which facilitates the incorporation of the c-UTP complex. The c-UTP complex consists of Rrp7, Utp22 and the casein kinase II subunits, Cka1, Cka2, Ckb1 and Ckb2 (Krogan et al, 2004). The role of these CKII subunits in ribosome biogenesis is not yet understood but in a yeast, *T. cutaneum*, CKII has been shown to phosphorylate ribosomal proteins

(Wojda et al, 2002). Rrp5 is also required for recruitment of the Rok1 helicase to the 90S complex (Vos et al, 2004).

1.4.2 RRP5

Recruitment of Rrp5 to the pre-ribosome following t-UTP incorporation is important for the subsequent association of several proteins required for formation of the SSU processome. Rrp5 has also been shown to associate with the Noc1/Noc2 protein complex (Merl et al, 2010). This complex is incorporated into the 90S particle but plays an important role in the nuclear transit of pre-LSU complexes (Milkereit et al, 2001; Milkereit et al, 2003). Consistent with this dual association with pre-SSU and pre-LSU complexes, Rrp5 has been shown to be required for formation of both 18S and $5.8S_S$ rRNAs and in yeast is important for cleavages at A_0, A_1, A_2 and A_3 (Venema & Tollervey, 1996). Rrp5 is one of only a small set of proteins which are required for both ITS1 cleavages at A_2 and A_3. Rrp5 contains 12 S1 RNA binding domains at the N-terminus and binds to the pre-rRNA between the A_2 and A_3 cleavage sites (Young & Karbstein, 2011). These S1 domains are required for the functions of Rrp5 in A_2 and A_3 cleavage but not for the cleavages in the 5'ETS which are dependent on the seven tetratricopeptide (TPR) repeats found at the C-terminus of the protein (Eppens et al, 2002; Vos et al, 2004). Mutation of particular S1 domains of Rrp5 blocks A_3 cleavage and causes ITS1 cleavage at a novel site between A_2 and A_3 with a concomitant increase in $5.8S_L$ formation relative to that of $5.8S_S$. Interestingly, these processing defects can be rescued by deletion of the exonuclease, Rex4, in these cells, although this leads to the synthesis of defective ribosomes and is, therefore, lethal (Eppens et al, 2002). Rrp5 is also suggested to be important in exonucleolytic processing to generate the 5' end of 5.8S following A_3 cleavage. A homologue of Rrp5 was identified in human cells through its ability to interact with the transcription factor, NF-kB (p50) (Sweet et al, 2003). It is suggested that RRP5 is required for multiple pre-rRNA processing steps in the 5'ETS and ITS1 in human cells (Sweet et al, 2008). In human cells, RRP5 is also found associated with a 50S SSU processome complex that accumulates when early assembly steps are inhibited (Turner et al, 2009)

1.4.3 A_3-cluster proteins

The 90S particle is divided into a pre-40S complex and a pre-60S complex by cleavages at A_2 and A_3 in ITS1. In a similar way to the recruitment of sub-complexes of the SSU processome to the 90S particle, some trans-acting factors required for

formation of the large subunit rRNAs are recruited to the pre-ribosome prior to cleavage in ITS1. A group of proteins called the A_3-cluster proteins have been defined as they are all required for A_3 cleavage or the processing downstream of this cleavage to convert $27SA_3$ into $27SB_{1S}$. The A_3-cluster proteins include Erb1, Nop7, Ytm1, Rlp7, Nop15, Cic1, Rrp1 and Nop12 (Fatica et al, 2003b; Granneman et al, 2011; Miles et al, 2005; Oeffinger et al, 2002; Oeffinger & Tollervey, 2003; Pestov et al, 2001a; Sahasranaman et al, 2011; Wu et al, 2001). The recruitment of most of these factors to the pre-rRNA is interdependent. Erb1, Nop7 and Ytm1 form a stable trimeric complex that is proposed to form a scaffold onto which other A_3-cluster proteins, exonucleases and ribosomal proteins assemble (Miles et al, 2005; Sahasranaman et al, 2011; Tang et al, 2008). The A_3-cluster proteins have been shown to cross-link to a variety of different sites spanning 5.8S, ITS2 and 25S (Granneman et al, 2011).The major binding site for the Erb1/Nop7/Ytm1 complex is at the 5' end of 25S near the ITS2-25S junction. Nop15 and Cic1 cross-link to sequences at the 5' end of ITS2 *in vivo* although Cic1 was also observed to cross-link to sequences in 25S. During maturation of the 60S complex, ITS2 undergoes major structural rearrangements with the 5' end of 25S and the 3' end of 5.8S forming contacts in the mature ribosome, which are not present in early pre-60S complexes. The presence of both Cic1 and Nop15 in these early complexes is proposed to be important for preventing premature formation of these base pairing interactions (Granneman et al, 2011). Similarly, Nop4, another protein associated with the A_3-cluster (see section 1.4.4), was found to cross-link to many sites in domains I and II of 25S and also to the 5' end of 5.8S. These rRNA regions are close proximity in the mature ribosome suggesting a role for Nop4 in long-range base pairing interactions (Granneman et al, 2011). Four ribosomal proteins (rpL17, rpL26, rpL35 and rpL37) are recruited to the pre-ribosome by the A_3-cluster proteins and bind to this region of the mature ribosome. It is, therefore, suggested that the long-range interactions formed by Nop4 in the pre-ribosome are maintained by these r-proteins in the mature ribosome (Granneman et al, 2011; Sahasranaman et al, 2011).

The A_3-cluster proteins also play important roles in pre-rRNA processing and quality control. The Erb1/Ytm1/Nop7 complex is required for processing from $27SA_2$ to $27SA_3$ and Ytm1 is also important for mediating nucleolar export of pre-66S complexes in a reaction driven by the ATPase, Rea1 (Bassler et al, 2010; Miles et al, 2005; Pestov et al, 2001a). Incorporation of the core A_3-cluster proteins is also required for recruitment of the exonuclease, Rrp17, but interestingly not Rat1, although both of these proteins carry out processing to form the mature end of $5.8S_S$ rRNA (Granneman et al, 2011). The secondary structures of 5.8S, ITS2 and 25S formed by the A_3-cluster proteins are also proposed to regulate the timing of Rat1 processing of the 5' end of

5.8S. In the absence of A_3-factors, short Rat1-dependant RNAs are accumulated suggesting that in addition to processing, the Rat1 exonuclease degrades aberrant pre-rRNAs formed if ribosome assembly is aborted at this stage (Granneman et al, 2011). It is suggested that the exonucleolytic processing to the mature 5' end of 5.8S may normally be arrested by the presence of the A_3-cluster proteins bound near the 5' end of 5.8S but if these proteins are not recruited, then exonucleolytic processing turns into degradation of the aberrant complex.

Homologues of some of these A_3-cluster proteins have been identified in higher eukaryotes. Erb1, Nop7 and Ytm1 homologues BOP1, PES1 and WDR12, respectively, have been characterised in both mouse and human cells and also form a stable trimeric complex, PeBoW (Holzel et al, 2005; Lapik et al, 2004). BOP1 (Erb1) was first identified in mouse as a nucleolar protein that is essential for the production of both 28S and 5.8S rRNAs (Strezoska et al, 2000). *In vitro* reconstitution of the PeBoW complex and *in vivo* protein studies in mouse show that the integrity and stability of the complex is largely dependent on BOP1 levels (Rohrmoser et al, 2007). BOP1 is critical for recruitment of PES1 to the pre-ribosome (Lapik et al, 2004) but interestingly, PES1 is required for transport of BOP1 into the nucleolus following translation (Rohrmoser et al, 2007). In mouse, BOP1 is proposed to be required for four different cleavage steps in both internal transcribed spacers and in the 3'ETS (Strezoska et al, 2002). First, 3' extended forms of 36S, 32S and 41S were observed to accumulate after BOP1 depletion confirming a role for BOP1 in efficient removal of the 3'ETS. Second, depletion of BOP1 revealed an accumulation of 32S and a decrease in 12S suggesting BOP1 is necessary for ITS2 processing that separates the 5.8S precursors from pre-28S. Third, BOP1 has an additional role in ITS2 removal as it is required for processing of 12S to produce 5.8S (Strezoska et al, 2002). Finally, an increase in the levels of 36S, an intermediate of an alternative ITS1 processing pathway, suggests that BOP1 is also required for late ITS1 processing steps. Mutation or depletion of members of the PeBoW complex lead to cell cycle arrest (Strezoska et al, 2002) and expression of these proteins is up-regulated by the proto-oncogene, c-Myc, suggesting that this complex provides a link between ribosome biogenesis and cell proliferation (Holzel et al, 2005).

By homology searches, we have identified the human homologues of other A_3-cluster proteins and of particular interest may be Mki67 (Nop15). This protein interacts with the Ki67 antigen which is localised in the dense fibrillar compartment (DFC) of the nucleolus and has been suggested to play an important role in coordinating ribosome biogenesis with mitosis and cell cycle progression (MacCallum & Hall, 2000).

1.4.4 RBM28 (Nop4)

Nop4 (also known as Nop77) is absolutely required for cleavage at A_3 by RNase MRP but is also thought to be important for cleavage at A_2 and processing from A_3 to B_{1L} (Berges et al, 1994; Granneman et al, 2011; Sun & Woolford, 1994; Sun & Woolford, 1997). A second function of Nop4 in mediating post-transcriptional methylation of the pre-rRNA has been described by one group (Berges et al, 1994) but was not observed in another study (Sun & Woolford, 1994). The human homologue of Nop4 is RBM28 and although this protein has been linked to ribosome biogenesis, little is known about its role in human pre-rRNA processing (Nousbeck et al, 2008). ANE syndrome, which is characterised by alopecia, neurological defects and endocrine deficiency, is caused by a point mutation in the gene coding for RBM28 which alters the protein structure and stability leading to decreased expression levels (Nousbeck et al, 2008; Spiegel et al, 2010). The homologue of RBM28 in *C. elegans* is involved in regulating the expression of the miRNA, lin4, which has important functions in coordinating the timing of development in this organism (Bracht et al, 2010).

1.4.5 Formation of the pre-60S complex

Following cleavage in ITS1, the 90S pre-ribosomal complex is separated into the pre-40S and pre-60S complexes. In addition to exonucleases involved in pre-rRNA processing to produce mature 5.8S and 25S (28S) rRNAs and recruitment of the A_3 cluster proteins, a number of other trans-acting protein factors required for biogenesis of the large subunit are also recruited. In yeast, these include a number of helicases (Mak5, Dbp2, Dbp3, Dbp6, Dbp7, Dbp9, Dbp10, Mtr4, Dsr1, Spb4, Has1 and Prp43), ATPases (Rix7 and Rea1), GTPases (Nug1, Nog1 and Nog2) and methyltransferases (Spb1 and Nop2). The precise roles of many of these proteins have not yet been defined. Association and dissociation of members of a group of proteins, Noc1, Noc2 and Noc3, are proposed to mediate movement of the pre-60S from the nucleolus and across the nucleoplasm prior to export into the cytoplasm (Milkereit et al, 2001). The ATPase, Rea1, is thought to function at multiple steps during pre-60S maturation causing the dissociation of various proteins including the Erb1-Ytm1-Nop7 complex and Rsa4 causing structural rearrangements which enable export both from the nucleolus and the nucleoplasm (Bassler et al, 2010; Galani et al, 2004). The 5S rRNA is independently transcribed by RNA polymerase III in the nucleus and through its interaction with the ribosomal protein, RPL5, is re-located into the nucleolus where it is incorporated into the early pre-60S complex (Zhang et al, 2007). In the mature ribosome, 5S is situated at the interface between the LSU and SSU. The incorporation

of 5S into the ribosome is required for processing of the LSU rRNAs and potentially regulates the stoichometry of the rRNAs in the ribosome (Dechampesme et al, 1999). Export of the pre-60S complex into the cytoplasm will be discussed in more detail in section 1.6 but several non-ribosomal proteins are co-exported into the cytoplasm where the final maturation occurs. The pre-60S complex and the co-exported factors, Alb1, Arx1, Nog1 and Tif6, are joined by cytoplasmic biogenesis factors including Lsg1, Sqt1, Drg1, Efl1, Jjj1 and Rei1. Together these factors release the co-exported biogenesis factors, provide quality control checkpoints, cause conformational rearrangements and promote incorporation of the final ribosomal proteins including RPL24.

1.5 Ribosomal proteins

The organisation of the ribosomal proteins in the mature ribosome has recently been defined through crystal structures of eukaryotic ribosomes (Ben-Shem et al, 2011; Ben-Shem et al, 2010). The ribosomal proteins of the eukaryotic small subunit have been classified into two groups based on their roles in pre-rRNA processing. First, initiation-RPS (iRPS), RPS3a, RPS4, RPS5, RPS6, RPS7, RPS8, RPS9, RPS11, RPS13, RPS14, RPS16, RPS15a, RPS23, RPS24, RPS28 and RPS27, are associated with early 90S complexes. Second, progression-RPS (p-RPS), RPS19, RPS18, RPSA, RPS21 RPS2, RPS3, RPS17, RPS20, RPS27a, RPS29 RPS10, RPS12, RPS15, RPS26, and RPS25, associate later and are required for downstream processing events (O'Donohue et al, 2010). Depletion of i-RPS causes accumulation of 45S and 30S with no downstream pre-rRNAs being produced indicating a complete block in the pathway following the primary cleavages in the 5'ETS and ITS1. It is thought that binding of the i-RPS to the pre-ribosome is important for RNA folding to generate a stable conformation in which processing can be initiated (O'Donohue et al, 2010). Depletion of p-RPS leads to various different phenotypes involving either partial or complete removal of the 5'ETS and inhibition of ITS1 processing (O'Donohue et al, 2010). The class can be subdivided into those proteins which affect 21S-18SE processing preventing formation of 18SE (RPS17, RPS18, RPS19, RPS21 and RPSA) and those which alter the levels of either 21S or 18SE without affecting the conversion process. Of particular interest is the observation of a novel processing intermediate previously called 20S (later referred to as 21SC) between 21S and 18SE upon depletion of RPS19 and weakly following depletion of RPS17, indicating that these proteins are important for the later steps of ITS1 processing (Idol et al, 2007; Robledo et al, 2008). Depletion of many of the p-RPS uncouples the 5'ETS cleavages at A_0 and A_1 suggesting that their association into pre-ribosomes may normally be important for

coordinating these steps. Further, p-RPS, RPS15 and RPS17 are required for nuclear export of pre-40S complexes (Leger-Silvestre et al, 2004) while the presence of many other p-RPS is essential for the final cytoplasmic cleavage at site 3 (O'Donohue et al, 2010). In yeast, however, it has been shown that some RPS such as RPS5 and RPS17 play roles in both early and late processing steps. The role of ribosomal proteins of the large subunit (RPLs) is less well studied but of particular note is the observation that depletion of RPL26 (and to a lesser extent RPL35a) leads to defects in both the 5.8S precursor (12S) and also 18S precursors (21S and 30S) resulting in less mature 40S subunits. This indicates that these LSU proteins are required for processing steps following separation of the SSU and LSU precursors (Robledo et al, 2008).

1.6 Nuclear export and cytoplasmic maturation

Following nucleolar assembly and nuclear maturation of the pre-40S and pre-60S complexes, these pre-ribosomal subunits are separately exported into the cytoplasm where they undergo final maturation steps before they rejoin to form the mature ribosome (Panse & Johnson, 2010; Rouquette et al, 2005; Zemp & Kutay, 2007). Export of both subunits is a dynamic process involving the exportin, Xpo1/CRM1, the RanGTPase system and various nuclear pore components. Each pre-ribosomal subunit contains one or more adaptors which contain leucine-rich nuclear export sequences (NES) through which they interact with CRM1 (Fukuda et al, 1997; Zemp & Kutay, 2007). Until recently, the criteria for defining and identifying NES sequences was based on four key hydrophobic residues (Φ) following the consensus Φ^1-$(x)_{2-3}$-Φ^2-$(x)_{2-3}$-Φ^3-x-Φ^4, with "x" denoting small polar amino acids. This has been extended to include a fifth hydrophobic residue and recent structures of CRM1 bound to classic NES cargos highlight other structural features which are recognised by CRM1 and define a pocket in CRM1 into which the NESs must fit (Guttler & Gorlich, 2011; Guttler et al, 2010; Mattaj & Muller, 2010; Monecke et al, 2009). CRM1 binds to RanGTP and its cargo protein(s) in the nucleus and translocates through a nuclear pore complex to the cytoplasm. Here, hydrolysis of RanGTP to RanGDP causes the complex to dissociate releasing the pre-ribosomal complexes into the cytoplasm (Figure 1.6) (Guttler & Gorlich, 2011; Zemp & Kutay, 2007). It is thought that additional factors chaperone the subunits through the nuclear pore complex. For example, the HEAT repeat-containing protein, Rrp12, contributes to the nuclear export of both the pre-40S and pre-60S complexes (Oeffinger et al, 2004) and could coordinate their interactions with the transport receptors.

Export of the pre-60S complex into the cytoplasm in both yeast and humans is mediated by the NES-containing adaptor protein, Nmd3 (Ho et al, 2000). This protein

participates in nonsense mediated decay (NMD) of mRNAs in the cytoplasm but also shuttles between the cytoplasm and nucleus to export pre-60S complexes. In yeast, export of pre-60S complexes involves additional factors that interact directly with the nuclear pore complex and these include the mRNA export factor, Mtr2/Mex67 and the shuttling trans-acting factors Arx1 and Ecm1 (Bradatsch et al, 2007; Hung et al, 2008; Yao et al, 2007; Yao et al, 2010). Once in the cytoplasm factors such as Lsg1 and Sqt1 participate in the recycling of Nmd3 into the nucleus. In higher eukaryotes, export of pre-60S complexes has also been shown to be achieved by a CRM1-independent pathway involving Exportin 5 (Exp5), an export factor previously associated with translocation of miRNA precursors (Bohnsack et al, 2004; Lund et al, 2004; Wild et al, 2010). Depletion of this protein causes strong defects in 60S biogenesis and Exp5 has been shown to interact specifically with pre-60S complexes in a RanGTP dependant manner clarifying its role as an export adaptor (Wild et al, 2010).

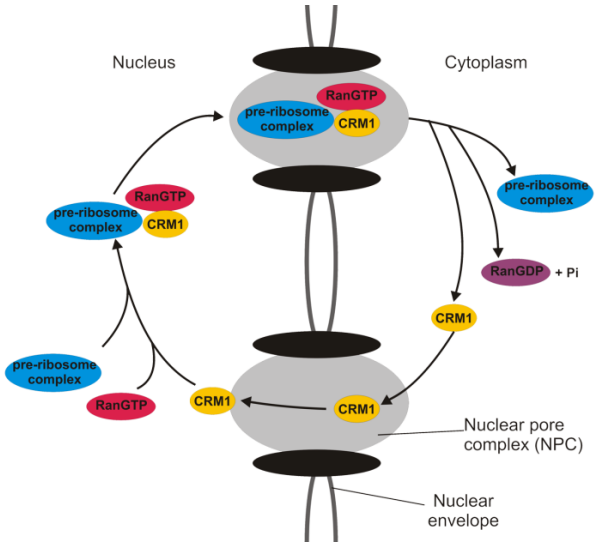


Figure 1.6 Export of pre-ribosomal complexes into the cytoplasm Schematic overview of the nuclear export cycle. Pre-ribosomal complexes contain nuclear export adaptor proteins which interact with CRM1 in a RAN dependant manner. These complexes are recruited to the nuclear pores and following exit into the cytoplasm, the RAN associated GTP is hydrolysed causing a conformational change which releases the pre-ribosomal complexes into the cytoplasm. CRM1 is then recycled back into the nucleus. Adapted from Guttler and Gorlich, 2011.

The export and cytoplasmic stages of pre-40S complex maturation involve a small number of proteins because many of the trans-acting factors required for earlier steps dissociate in the nucleus. The composition of late cytoplasmic pre-40S complexes in yeast has been determined by affinity purification and mass spectrometry. These identified key proteins which are required for late processing steps in yeast: Dim1, Pno1 (Dim2), Enp1, Ltv1, Hrr25, Nob1, Rio2, Rio1 and Tsr1. While Nmd3 is the only known export adaptor of yeast pre-60S complexes, Ltv1 and Pno1 both function as adaptors for export of pre-40S complexes (Seiser et al, 2006; Vanrobays et al, 2008). Ltv1 is a non-essential ribosome biogenesis factor suggesting redundancy between export proteins (Seiser et al, 2006). In human cells, RIO2 has been shown to interact with CRM1 *in vitro* in a RanGTP dependant manner through a conserved NES (Zemp et al, 2009). Disruption of this NES inhibits nuclear export of

pre-40S complexes identifying RIO2 as a nuclear export adaptor for the small subunit (Zemp et al, 2009). It is not clear whether PNO1 is also an export adaptor of pre-40S complexes in human cells as, although this protein interacts with CRM1 *in vitro* and a putative NES has been identified, mutation of this sequence did not affect CRM1 binding (Zemp et al, 2009). In human cells, TSR1 has been shown to be involved in nuclear export of pre-40S complexes but does not contain a leucine-rich NES and is, therefore, not proposed to be a CRM1 adaptor protein. TSR1 depletion does, however, cause nuclear retention of pre-40S complexes although TSR1 is not required for pre-rRNA processing, suggesting the role of TSR1 must be in assembly of an export-competent complex (Carron et al, 2011).

To investigate nuclear export, an inhibitor of CRM1, leptomycin B (LMB), is often used. LMB covalently binds to a cysteine residue (528) in a conserved pocket of CRM1 preventing interaction with the export sequences of cargo proteins. It is important to note that CRM1 has functions other than mediating nuclear export which include coordinating the nucleolar localisation of snoRNPs including U3 (Boulon et al, 2004). Treatment of HeLa cells with LMB prevents the shuttling of the pre-40S adaptor, RIO2 and other co-exported proteins such as NOB1 between the nucleus and cytoplasm causing them to accumulate in the nucleus (Zemp et al, 2009) and Watkins lab unpublished data). It has also been reported that treating HeLa cells with LMB affects early pre-rRNA processing steps causing a significant accumulation of 26S, a pre-rRNA processing intermediate extending from the A_0 cleavage site in the 5'ETS to site 2 in ITS1. It is, therefore, suggested that inhibiting CRM1 uncouples the normally simultaneous cleavages at A_0 and A_1 in the 5'ETS (Rouquette et al, 2005).

1.7 Late steps in pre-40S maturation

Cytoplasmic maturation of the pre-40S complex in yeast involves two major events; structural rearrangements of the complex and pre-rRNA processing to produce mature 18S rRNA (Fromont-Racine et al, 2003; Henras et al, 2008). In yeast, pre-rRNA processing in the cytoplasm only involves cleavage of 20S at site D by the endonuclease Nob1 (Fatica et al, 2003a; Fatica et al, 2004). Other proteins found in late cytoplasmic complexes including Pno1, Rio2 and Prp43 are required for this cleavage step. In higher eukaryotes, pre-rRNA processing to form the mature 3' end of 18S involves an additional processing step (Rouquette et al, 2005) implying that the functions of some of the protein factors involved in late maturation steps may not be conserved from yeast. Binding sites in 18S pre-rRNA for most of the proteins found in yeast late pre-40S complexes (except Rio1 and Pno1) have been defined by *in vivo*

cross-linking (CRAC) enabling a greater understanding of their functions in ribosome assembly (Granneman et al, 2010). Based on the crystal structure of the mature eukaryotic small subunit, a model of the relative positions of these yeast late acting biogenesis factors was developed using cryo-electron microscopy (Rabl et al, 2010; Strunk et al, 2011).

1.7.1 Structural reorganisation of pre-40S complexes

Mature 40S particles contain a structural feature called the “beak”, which consists of a protrusion of 18S rRNA helix 33 and the ribosomal protein Rps3. This structure is lacking in pre-40S complexes indicating that a large-scale reorganisation of this region is an important maturation step. However, premature formation of this beak structure impairs nuclear export of the pre-40S complex indicating that regulating timing of this event is important (Schafer et al, 2006; Seiser et al, 2006). It has been shown that Rps3 is weakly associated with the pre-40S complex in the nucleus but phosphorylation of this protein and the SSU biogenesis factors, Enp1 and Ltv1 by the casin kinase I isoform, Hrr25, causes their dissociation from the complex (Schafer et al, 2006). Rps3 is subsequently dephosphorylated and reforms a more rigid association with the pre-rRNA to form the beak in the cytoplasm (Schafer et al, 2006). Consistent with this model, Enp1 was found to cross-link directly to helix 33 of 18S and Ltv1 was also bound to the beak region (Granneman et al, 2010).

1.7.2 ENP1

In yeast, Enp1 is found associated with both early and late pre-40S complexes suggesting it may play multiple roles in small subunit assembly. The early role of ENP1 involves the phosphorylation of Enp1 by Hrr25 and its subsequent dissociation from pre-40S complexes described above but following dephosphorylation in the cytoplasm, Enp1 is re-associated with late pre-40S complexes to perform its later functions (Schafer et al, 2006). Enp1 is required for efficient production of 20S in yeast and is associated with the U3 and U14 snoRNPs (Chen et al, 2003). A human homologue of Enp1 has been identified and is also known as bystin or BYSL (Fukuda et al, 2008; Miyoshi et al, 2007). Depletion of ENP1 from human cells causes accumulation of 21SC, an intermediate between 21S and 18SE which is thought to be generated by exonucleolytic processing due to its heterogeneous 3’ end (Carron et al, 2011). This implicates ENP1 in a step of human ribosome biogenesis which is not conserved from yeast. In higher eukaryotes, ENP1 has been shown to interact with throphinin, and

through this plays an important role in embryo implantation (Suzuki et al, 1998). It has also been shown that ENP1 is essential for early embryonic stem cell survival (Aoki et al, 2006). More recently, it was reported that ENP1 is over-expressed in hepatocellular carcinomas (Wang et al, 2009). These authors also suggested that ENP1 is a target of c-Myc possibly indicating a role for ENP1 in coupling ribosome biogenesis and cell growth rate or tumour development.

1.7.3 NOB1 and PNO1

In yeast, Nob1 is the endonuclease responsible for cleavage at site D to form the mature 3' end of 18S. The nuclease activity of this protein and its regulation will be discussed in more detail in section 1.9.1. A cofactor protein that binds to Nob1 is Pno1 (Partner of Nob1, also known as Dim2). Pno1 has functions both in early pre-rRNA processing steps in the nucleolus and also in late steps in the cytoplasm. In yeast, Pno1 is associated with the nascent transcript and is essential for early cleavages at A_1 and A_2. Further, an aberrant intermediate extending from A_0 to A_3 (22S) is seen to accumulate when the G207 residue in the KH domain of Pno1 is modified (Vanrobays et al, 2004). This conserved KH domain of Pno1 has been shown to bind to the 5' end of ITS1 (Vanrobays et al, 2008). In yeast, Pno1 has been shown to shuttle between the nucleus and cytoplasm as part of the pre-40S complex. Pno1 contains a leucine-rich NES which is required for efficient export of pre-ribosomal complexes into the cytoplasm indicating that Pno1 is an export adaptor for pre-40S complexes (Vanrobays et al, 2008) but it is not clear if this is also the case for the human protein. The localisation of Pno1 has also been shown to be regulated by the mTOR pathway potentially coupling external nutritional state to the rate of ribosome biogenesis (Vanrobays et al, 2008). Pno1 has been shown to interact with Nob1 both in yeast two hybrid studies and also *in vitro* and more specifically, the KH domain of Pno1 is also proposed to mediate this interaction (Tone & Toh, 2002; Woolls et al, 2011). This interaction increases the affinity of Nob1 for the pre-rRNA and is important for enabling Nob1 to cleave at the 3' end of 18S rRNA (Woolls et al, 2011).The interactions of Nob1 with the pre-rRNA and its activity will be discussed in more detail in section 1.9.1.

1.7.4 DIM1

In yeast, an important step in the maturation of the small subunit is methylation of two conserved adenosine residues (A1179 and A1780) at the 3' end of 18S pre-rRNA. This modification is carried out by the dimethyltransferase, Dim1, which is an

essential component of the SSU processome (Lafontaine et al, 1995). It has been demonstrated that in yeast, dimethylation occurs following export of the pre-40S complex into the cytoplasm although Dim1 is associated with early pre-ribosome complexes (Brand et al, 1977; Lafontaine et al, 1995; Lafontaine et al, 1998b; Schafer et al, 2003). Consistent with these observations, yeast Dim1 is localised predominantly in the nucleolus but can also be detected in the cytoplasm. Dim1 is essential for cleavages at A_1, A_2 and D in yeast but these steps are not dependant on dimethyl modification of 18S (Lafontaine et al, 1995; Lafontaine et al, 1998b). This suggests that the presence of Dim1 and therefore, the ability to methylate is a pre-requisite for early pre-rRNA processing thereby preventing formation of aberrant, unmethylated 40S subunits. Ribosomes which do not contain these modifications have been shown to be inactive in *in vitro* translation assays underlining the importance of such control mechanisms (Lafontaine et al, 1998b).

1.7.5 PRP43

RNA helicases are proposed to have several different functions in ribosome biogenesis including, unwinding RNA secondary structure enabling it to be degraded, RNA remodelling to facilitate pre-rRNA processing steps and the recruitment and/or dissociation of the snoRNP complexes involved in both pre-rRNA processing and modification. There are 19 helicases that have been identified as playing important roles in ribosome biogenesis and most have been shown to be required specifically for either large or small subunit synthesis. Yeast Prp43, however, has been demonstrated to be involved in both SSU and LSU production (Bohnsack et al, 2009; Combs et al, 2006; Granneman et al, 2006; Lebaron et al, 2005; Pertschy et al, 2009). Consistent with this, using *in vivo* cross-linking, Prp43 has been shown to associate with sites both at the 3' end of 18S rRNA and also at multiple sites throughout 25S rRNA (Bohnsack et al, 2009). Prp43 is thought to have distinct functions in the biogenesis of each subunit. Prp43 is implicated in the association and dissociation of a number of different snoRNPs which modify 25S rRNA (Bohnsack et al, 2009). In contrast, in SSU formation, Prp43 is involved in the final processing step which converts 20S into 18S. This step was also shown to require a cofactor of Prp43, Pfa1 and another late processing factor, Ltv1 (Lebaron et al, 2009; Pertschy et al, 2009). It is suggested that the helicase activity of Prp43 may be required to re-structure the pre-rRNA around site D to expose the cleavage site to Nob1. In addition to its role in ribosome biogenesis, Prp43 is also a pre-mRNA splicing factor involved in the release of the intron lariat from the spliceosome (Arenas & Abelson, 1997)

1.7.6 RIO2

RIO2 is a member of a family of atypical serine kinases that are conserved from archaea to mammals (LaRonde-LeBlanc & Wlodawer, 2005a; LaRonde-LeBlanc & Wlodawer, 2005b). Compared to other typical protein kinases (ePKs), the kinase domains of RIO proteins contain conserved catalytic amino acids but many of the sequences and structures normally involved in substrate binding and recognition are missing (LaRonde-LeBlanc & Wlodawer, 2005a). In archaea and lower eukaryotes two Rio kinases are found, Rio1 and Rio2 whereas in higher eukaryotes, three RIO kinases have been identified, RIO1, RIO2 and RIO3. Compared to RIO1, both RIO2 and RIO3 have an additional N terminal domain, the function of which is not known. In yeast, Rio1 and Rio2 have been shown to be capable of serine phosphorylation in vitro (Angermayr & Bandlow, 2002; Geerlings et al, 2003; Vanrobays et al, 2003; Vanrobays et al, 2001). It has also been demonstrated that Rio1 from *Archeoglobus fulgidus* is able to autophosphorylate serine 108 in a flexible loop region (LaRonde-LeBlanc et al, 2005). No substrates of the kinase activity of Rio2 have been identified in yeast or higher eukaryotes. The binding site of yeast Rio2 in 18S has led to the suggestion that ribosomal proteins Rps16 and Rps18 may be potential targets for Rio2 phosphorylation but there is no direct evidence for this (Granneman et al, 2010).

Both RIO1 and RIO2 are required for 18S production in both yeast and human cells (Vanrobays et al, 2003; Zemp et al, 2009). Depletion of either protein inhibits the final maturation step of 18S and in yeast causes accumulation of 20S (Geerlings et al, 2003). However, only RIO2 has been found to be stably associated with pre-40S complexes (Schafer et al, 2003). It is currently not clear what the role of RIO1 in ribosome biogenesis is although recent data show that the kinase activity of RIO1 is required for 18S production in higher eukaryotes (Widmann et al, 2011). In human cells RIO2 has been shown to be an export adaptor for transport of pre-40S complexes into the cytoplasm but this function does not appear to rely on the kinase activity of the protein (Zemp et al, 2009). However, the kinase activity of human RIO2 is necessary for the conversion of 18SE into 18S but it is not clear how this kinase activity enables site 3 cleavage. The kinase activity of both RIO2 and RIO1 is necessary for release of LTV1, PNO1 and NOB1 but not ENP1 from pre-40S complexes (Widmann et al, 2011; Zemp et al, 2009).

1.8 Quality control of ribosome biogenesis

Ribosome assembly is tightly regulated to prevent the formation of defective ribosomes. In yeast, deletion or mutation of ribosomal proteins can cause pre-rRNA processing to stall (Moritz et al, 1990; Moritz et al, 1991; Rotenberg et al, 1988). However, mutations in several ribosomal proteins in human cells have been linked to genetic diseases such as Diamond Blackfan anaemia (Ellis & Gleizes, 2011). The assembly of active biogenesis factors onto the pre-rRNA much earlier in the processing pathway than they are required could provide a quality control system. As described earlier, Dim1 is recruited to early pre-ribosomal complexes and its presence rather than its methylation activity is required for cleavages at A_0 and A_1 (Lafontaine et al, 1998b). Similarly, mutation of factors involved in pre-60S maturation such as Dbp4, Rok1 and Rrp3 has been shown to affect early pre-rRNA cleavages (Kressler et al, 1999). However, when ribosome biogenesis is disrupted, pre-rRNAs are not accumulated very significantly implying they are unstable and that they are rapidly targeted for degradation (Houseley & Tollervey, 2009). These abortive pre-rRNAs are degraded by the exosome. Defective pre-rRNAs are identified by the TRAMP complex which adds short polyadenylated tails which recruits the exosome to these substrates (this will be discussed in more detail in section 1.10.3.4.3). Quality control mechanisms also function once mature ribosomes have been formed in the cytoplasm. Ribosomes containing mutations in functionally important sites such as the peptidyltransferase centre or the decoding site are much less stable than functional ribosomes and are turned over in a non-functional rRNA decay (NRD) pathway (LaRiviere et al, 2006).

1.9 Endonucleases in pre rRNA processing

Although many cleavage steps in pre-rRNA processing in *S. cerevisiae* have been identified and some of the cleavage sites have been mapped, the majority of the endonucleases responsible for these cleavages remain elusive. Other proteins that have been identified as playing important roles in pre-rRNA processing have been putatively suggested to be endonucleases based on protein homology to known endonucleases. Although cleavages comparable to those identified in yeast are thought to occur in human cells, no endonucleases have thus far been shown to be responsible for mediating these cleavages. However, all the endonucleases identified in yeast as important for ribosome biogenesis have homologues in higher eukaryotes making it possible that these functions are conserved.

1.9.1 NOB1

NOB1 is a highly conserved protein that is essential for production of the small subunit. NOB1 is a component of late pre-40S complexes in both yeast (Schafer et al, 2003) and human cells (Rouquette et al, 2005; Zemp et al, 2009) and in yeast is associated with 20S pre-rRNA. Nob1 has been identified as the endonuclease responsible for cleavage at site D at the 3' end of 18S in yeast (Fatica et al, 2003a; Fatica et al, 2004; Lamanna & Karbstein, 2009; Pertschy et al, 2009). It remains unclear if this function is conserved in human NOB1.

NOB1 contains a PIN (PilT N-terminus) domain which has been shown to confer endonuclease activity on a number of proteins. PIN domains are motifs of approximately 100 amino acids containing up to four highly conserved acidic amino acid residues. The spatial arrangement of the four acidic residues is consistent with chelation of divalent metal cations, which are often required for ribonuclease activity. Other PIN domain proteins, such as Rrp44 and Smg6, have been shown to require Mn^{2+} and Mg^{2+}, respectively, for their activities (Huntzinger et al, 2008; Schneider et al, 2009). Structural studies of PIN domain proteins from archaea show that PIN domain proteins have a similar fold to 5'-3' exonucleases, T4 phage RNase H and flap endonucleases (Arcus et al, 2004; Veith et al, 2011). PIN domain proteins, including recombinant yeast Nob1, have been shown to assemble as tetrameric complexes creating a central pore through which only single stranded nucleotides are able to pass and be cleaved (Lamanna & Karbstein, 2009).

The PIN domain of yeast Nob1 has been shown to be required for cleavage at site D both *in vivo* (Fatica et al, 2004) and *in vitro* (Pertschy et al, 2009). Nob1 has been shown to be associated with pre-40S particles in the nucleus and is exported to the cytoplasm as part of the complex where it performs site D endonucleolytic cleavage (Fatica et al, 2003a). This suggests that the activity of Nob1 is suppressed until export to the cytoplasm. Different models have been proposed for how this regulation is achieved and the first suggests that ITS1 sequences base pair to highly conserved sequences near the 3' end of 18S and block the cleavage site (Lamanna & Karbstein, 2011). The second model, based on *in vivo* cross-linking data showing that the primary binding site of Nob1 is in the 3' major domain of 18S RNA and distinct from the cleavage site, proposes that structural changes as processing occurs bring Nob1 into the proximity of its cleavage site (Granneman et al, 2010). The region around site D has been shown to be highly flexible (Granneman et al, 2010) and major structural rearrangement of the pre-40S complex once it is exported to the cytoplasm has been demonstrated giving credence to this model (Schafer et al, 2006). An NMR structure of archaeal Nob1 has recently been published and this identifies key amino acid residues

within the PIN domain which are essential for Nob1 interaction with site D while a zinc ribbon domain mediates the interaction of Nob1 with its primary binding site on helix 40 of 18S rRNA (Veith et al, 2011).

1.9.2 RNase MRP

RNase MRP is an RNA-protein complex that is conserved from prokaryotes through to higher eukaryotes and in yeast both the RNA and protein components have been shown to be essential (Chang & Clayton, 1987; Schmitt & Clayton, 1992). The RNA component of human RNase MRP is 267 nucleotides long and highly structured forming a scaffold for interaction with at least seven proteins: POP1, POP5, RPP20, RPP25, RPP30, RPP38 and RPP40 (Figure 1.7) (Pluk et al, 1999; Welting et al, 2006; Welting et al, 2004). Biogenesis of RNase MRP, like that of other snoRNPs, is an ordered process and at least two of these proteins, RPP20 and RPP25, preassemble into a complex before they are able to interact with the P3 domain of the RNA component of RNase MRP (Hands-Taylor et al, 2010). It is further thought that RPP20 and RPP25 may be only transiently associated with the RNase MRP complex and as they bind near the catalytic centre (P4 and nucleotides 68-71) their dissociation may be required for endonucleolytic cleavage of pre-rRNA (Hands-Taylor et al, 2010; Welting et al, 2007). As well as these seven core proteins, three additional proteins associate with a subset of RNase MRP complexes: POP4, RPP21 and RPP14 (Welting et al, 2006). These proteins have been shown to bind RNase P preferentially and are thought to dissociate from RNase MRP when it interacts with pre-rRNA. RNase MRP is found to co-sediment in two peaks of 12S and 60-80S in both HeLa and Hep2 cells, similar to the profile of U3 snoRNP complexes. The 12S peak corresponds to the free-complex while the 60-80S peak indicates that RNase MRP is associated with pre-ribosomal complexes. In yeast, two RNase MRP-specific proteins have been identified, Snm1 and Rmp1 (Salinas et al, 2005; Schmitt & Clayton, 1994).

RNase MRP is localised to both mitochondria and the nucleolus and contains localisation signals for both these compartments in the central region and the Th/To domain, respectively (Jacobson et al, 1995; Li et al, 1994). RNase MRP was first identified as an endonuclease in mitochondria where it cleaves mitochondrial RNA required for mitochondrial DNA replication (Chang & Clayton, 1987). Localisation of RNase MRP in the nucleolus suggested a role in ribosome biogenesis and in *S. cerevisiae* deletion of the RNA component of MRP leads to a change in the ratio in which the long and short forms of 5.8S rRNA accumulate (Schmitt & Clayton, 1993). In normal cells, 85% of 5.8S is a short form ($5.8S_S$) and only 15% contains an extended 5'

end ($5.8S_L$) but after RNase MRP depletion, the levels of the short form significantly decrease while the long form accumulates. RNase MRP was, therefore, identified as the endonuclease responsible for carrying out cleavage at site A_3 in yeast ITS1 which generates the precursors of $5.8S_S$ (Lygerou et al, 1996). Mutation of various residues within the key subunit of the RNase MRP complex, Pop1, blocks A_3 cleavage without affecting the overall levels of the RNase MRP complex (Xiao et al, 2006). More recently, it has been suggested that RNase MRP has an additional role in mediating entry of early pre-rRNA transcripts into the canonical processing pathway. Pre-rRNA transcription was found to be unaffected by mutations in the RNA component of RNase MRP, but accumulation of long precursors (35S/33S/32S) was severely reduced. This suggests that early cleavage steps (A0, A1, A2 and A3) are inhibited by these mutations and this prevents pre-rRNAs from entering the normal processing pathway (Lindahl et al, 2009).

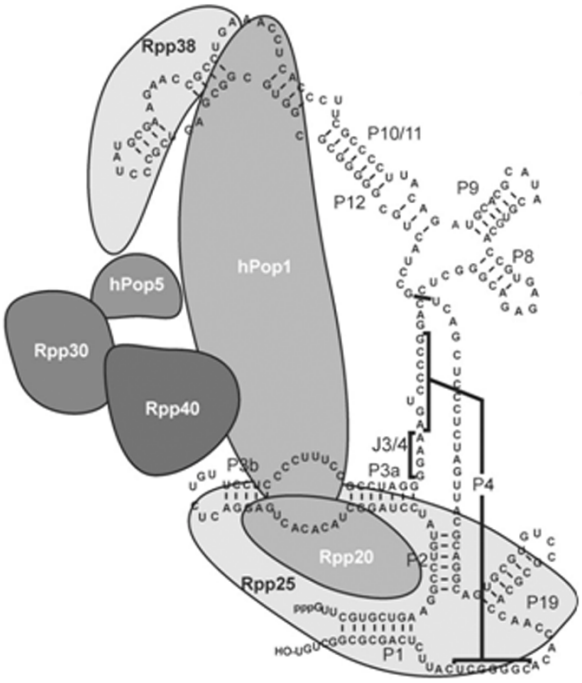


Figure 1.7 Structure of RNase MRP Schematic model of the human RNase MRP complex with the RNA shown in black and particular structural elements number according to Welting et al., 2004. Core proteins are shown in shades of grey. Adapted from Welting et al., 2004.

RNase MRP is also present in human cells although it is not clear whether the role in ribosomal RNA processing is conserved. siRNA mediated depletion of the RNase MRP/RNase P subunit, RPP38, causes defects in $tRNA_i^{Met}$ processing, which are likely to result from RNase P inhibition. However, the ratio of the long and short forms of 5.8S was unaffected, suggesting that RNase MRP may not be responsible for a cleavage analogous to A_3 in human cells (Cohen et al, 2003). RNase protection assay data, however, indicate that cleavage of sequences at the 3' end of ITS1 is

aberrant in cells over-expressing RPP38 potentially indicating a subtle or alternative role for RNase MRP in ITS1 processing in human cells (Cohen et al, 2003).

Although RNase MRP RNA and proteins are essential in yeast, A_3 is not an essential step in pre-rRNA processing suggesting that the mitochondrial activity of RNase MRP is vital or that RNase MRP has other critical roles within the cell. Additional roles for human RNase MRP in cell cycle regulation and disease have also been identified. Cleavage of CLB2 (cyclin B2) mRNA by RNase MRP provides an entry point for the exonuclease, XRN1, which degrades this mRNA causing down-regulation of cyclin B2 and exit from mitosis (Gill et al, 2004). Further, depletion of RNase MRP subunits, POP1 and RPP40, causes accumulation of the anti-viral protein, viperin (Mattijssen et al, 2011). This suggests a role for RNase MRP in regulating the levels of a protein required for the immune response. Finally, the RNA component of RNase MRP forms a complex with telomerase-associated reverse transcriptase (TERT) leading to the formation of a double stranded MRP-RNA. This double stranded RNA is then processed by the RISC pathway (Dicer) to generate siRNAs which down-regulate expression of RNase MRP RNA. This provides a mechanism by which MRP levels can be regulated or controlled within the cell (Maida et al, 2009).

1.9.3 RCL1

Rcl1 is a nucleolar protein that has general sequence homology to 3'-terminal phosphate cyclases. The recombinant yeast protein has been shown not to possess this activity and the lack of critical amino acid residues in the active site of the eukaryotic protein lead to the suggestion that Rcl1 has an alternative catalytic activity as an endonuclease (Billy et al, 2000). More recently, it has been demonstrated that recombinant yeast Rcl1 can act as an endonuclease *in vitro* (Horn et al, 2011). A crystal structure of Rcl1 identified a catalytic pocket of residues likely to coordinate the divalent cations that would be expected to be important for this activity. However, mutation of these was not found to affect Rcl1 activity *in vivo* (Tanaka et al). RCL1 in both yeast and humans associates with the U3 snoRNP-containing SSU processome (Billy et al, 2000; Turner et al, 2012). Rcl1 forms a sub-complex with the GTPase, Bms1, and this interaction is required for its association with U3 (Karbstein & Doudna, 2006; Karbstein et al, 2005; Wegierski et al, 2001). A model of how Rcl1 is recruited to the pre-rRNA has been proposed in which GTP bound Bms1 interacts with Rcl1 and this complex associates with the pre-ribosome. This then induces a conformational change which hydrolyses the GTP and releases Bms1, leaving Rcl1 *in situ* on the pre-rRNA (Karbstein & Doudna, 2006; Karbstein et al, 2005). In yeast, Rcl1 is required for

pre-rRNA processing at the A_0, A_1 and A_2 sites (Billy et al, 2000). A_0 cleavage is less dependent on Rcl1 than A_1 and A_2 are and it is currently proposed that Rcl1 is the endonuclease responsible for carrying out cleavage at the A_2 site. Recombinant Rcl1 has been shown to cleave pre-rRNA transcripts containing A_2 sequences in a concentration dependant manner and RNA mutations that have been shown to affect *in vivo* processing at A_2, impair this (Horn et al, 2011).

1.9.4 UTP24 and UTP23

Like Rcl1, Utp23 and Utp24 are components of the small subunit (SSU) processome in yeast and are proposed to be endonucleases as they contain PIN domains. Both proteins are essential in yeast and are required for the A_0, A_1 and A_2 cleavages needed to produce 18S RNA (Bleichert et al, 2006; Rempola et al, 2006). The PIN domain of Utp24, but not Utp23, is essential for the A_1 and A_2 cleavages suggesting that Utp24 may be the nuclease responsible for carrying out either or both of these steps in yeast. The activity of neither Utp24 or Utp23 is required for A_0 cleavage indicating that these proteins may be part of a processing complex required for the first 5'ETS cleavage (Bleichert et al, 2006). The PIN domain of yeast Utp23 contains only two of the four conserved acidic amino acid residues. Thus, Utp23 may only be a cofactor of Utp24 and enhance the activity of Utp24 by forming hetero-tetrameric complexes as other PIN domain proteins do (Bleichert et al, 2006). Interestingly, three of the four conserved amino acid residues of the PIN domain are found in the human UTP23 protein.

1.9.5 RNT1

Rnt1 is a double strand-specific endonuclease with homology to bacterial RNase III (Elela et al, 1996). Many different substrates of yeast Rnt1 have been identified and Rnt1 plays important roles in releasing pre-snoRNAs from introns or snoRNA clusters and in the formation of the 3' ends of splicosomal snRNAs (Chanfreau, 2003; Chanfreau et al, 1998a; Chanfreau et al, 1998b). Early pre-rRNA processing steps are inhibited in yeast strains expressing mutant Rnt1 and recombinant Rnt1 was found to cleave stem-loop RNAs containing both 5' and 3'ETS sequences *in vitro* indicating a role for Rnt1 in rRNA processing (Elela et al, 1996). Deletion of Rnt1 in *S. cerevisiae* was subsequently shown to cause slowed kinetics of early processing steps in the 5'ETS and ITS1 (A_0, A_1 and A_2) but was not essential for any of them (Kufel et al, 1999). However, cleavage in the 3'ETS is completely blocked

in the absence of Rnt1 and the 3' extended pre-25S intermediates normally generated by this cleavage are not observed. Rnt1 was shown to cleave twice in the 3'ETS on either side of a predicted stem-loop structure which is consistent with the consensus target site of Rnt1 (Kufel et al, 1999). This cleavage site is now called B_1 and endonucleolytic cleavage by Rnt1 provides a substrate end for exonucleolytic processing by Rex1 to generate the mature 3' end of 25S. In plant cells, the homologue of Rnt1, AtRTL2, also cleaves in the 3'ETS (Comella et al, 2008).

1.10 Exonucleases

1.10.1 5'-3' exonucleases

Several different conserved 5'-3' exonucleases have been identified and two of the major proteins are XRN2 (Rat1 in yeast) which is found in the nucleus and XRN1, a closely homologous protein which is localised to the cytoplasm. Another key and conserved 5'-3' exonuclease is NOL12 (Rrp17 in yeast). These and other 5'-3' exonucleases have important functions in a range of RNA degradation and processing pathways including pre-rRNA, snRNA and snoRNA maturation, turnover of RNA fragments and RNA surveillance mechanisms.

The exonuclease activity of Rat1 is stimulated by a cofactor, Rai1 (Xue et al, 2000). A crystal structure of these two proteins in complex from *Schizosaccharomyces pombe* has been published and the key residues which mediate this interaction were identified (Xiang et al, 2009). Although a homologue of Rai1, DOM3Z, is found in human cells, the amino acids which mediate the Rat1-Rai1 interaction are not conserved in XRN2 raising questions about whether DOM3Z interacts with and stimulates the activity of XRN2 in higher eukaryotes (Xiang et al, 2009). Interestingly, the crystal structure of Rai1 showed a separate, highly conserved pocket capable of coordinating a divalent cation. This region was then shown to endow Rai1 with pyrophosphohydrolase activity (Xiang et al, 2009). This activity converts the 5' triphosphate group of RNA substrates to monophosphates which can be degraded by Rat1 (Poole & Stevens, 1997; Xiang et al, 2009). Further, Rai1 possesses endonuclease activity which is used in the decapping of mRNAs. This activity is preferentially targeted towards unmethylated caps implying that Rai1 and Rat1 play an important role in degradation of mRNAs with aberrant or absent 5' caps (Jiao et al, 2010).

In yeast, both Xrn1 and Rat1 can function in pre-rRNA processing that produces the mature 5' end of $5.8S_S$ rRNA although Rat1 is normally responsible for

this (Henry et al, 1994). Another 5'-3' exonuclease, Rrp17, was identified in proteomic analysis of late pre-60S complexes and also participtes in $5.8S_S$ processing (Oeffinger et al, 2009). The catalytic domain of Rrp17 does not share any sequence homology to other known exonucleases but this protein is conserved and homologues have been identified in many species including a homologue, NOL12 (also called NOP25), in human cells (Fujiwara et al, 2006; Suzuki et al, 2007). NOL12 is able to functionally compensate for Rrp17 depletion suggesting that the roles of Rrp17 are conserved (Oeffinger et al, 2009). There is functional redundancy between the exonucleases involved in the 5' processing of $5.8S_S$ and deletion of any individual nuclease does not lead to detectable defects (Oeffinger et al, 2009). Due to its cytoplasmic localisation, Xrn1 does not, however, normally participate in $5.8S_S$ formation but can compensate for the absence of Rat1 in this process (El Hage et al, 2008; Henry et al, 1994) probably because processing is delayed until the complex reaches the cytoplasm. The activities of Rat1 and Rrp17 are probably coordinated during $5.8S_S$ processing as different regions of the ITS1 sequence are preferentially degraded by each of the exonucleases (Oeffinger et al, 2009). Rrp17 but not Rat1 is recruited to the pre-rRNA by the presence of proteins involved in A_3 cleavage (section 1.4.3) and which are required for the downstream exonuclease processing (Granneman et al, 2011).

Rat1 is also responsible for the degradation of fragments of pre-rRNA spacer regions which are released by consecutive endonucleolytic cleavages such as the A_0-A_1 and A_2-A_3 fragments (Petfalski et al, 1998). The interaction between Rat1 and pre-rRNA transcripts has recently been investigated using *in vivo* cross-linking. Consistent with its degradation and processing functions, Rat1 was found to cross-link robustly to the 3' end of the 5'ETS, upstream of the A_1 site, and to the region immediately 5' of the A_3 site but neither Rrp17 nor Rai1 could be cross-linked to pre-rRNA (Granneman et al, 2011).

The function of the Rat1 homologue, XRN2, in degradation of fragments of the 5'ETS is conserved in mouse cells (Wang & Pestov, 2011). Compared to yeast, the 5'ETS in higher eukaryotes contains an additional cleavage site, A', and as well as the A_0-A_1 fragment, mouse XRN2 is responsible for degradation of the released 5' end fragment generated by A' cleavage. Interestingly, in both mouse and plant cells, XRN2 is also required for A' cleavage (Wang & Pestov, 2011; Zakrzewska-Placzek et al, 2010). In mouse cells, this defect in A' cleavage is suggested to arise because other proteins required for this cleavage are not recycled from the 5'-A' fragment when XRN2 is not present to degrade it. In contrast, in plant cells, XRN2 is proposed to degrade the 5' end of the pre-rRNA, exposing the A' site enabling the endonucleolytic cleavage to occur (Zakrzewska-Placzek et al, 2010). In mouse cells, depletion of XRN2 also

causes the accumulation of a fragment of ITS1 extending between the 2a and 2 sites. Depletion of XRN2 alters the balance of ITS1 cleavages leading to an increased proportion of pre-rRNA cleaved at 2a (Wang & Pestov, 2011). Like Rat1 in yeast, XRN2 in mouse cells is required for exonucleolytic processing to form the mature 5' end of 28S rRNA (Wang & Pestov, 2011).

Both Rat1 and XRN2, in yeast and mammals respectively, have important functions in RNA surveillance and quality control of pre-rRNA transcripts. The presence of A_3-cluster proteins arrests Rat1 processing at the 5' end of 5.8S rRNA and in their absence, pre-rRNAs are completely degraded (Granneman et al, 2011). In mouse cells, XRN2 contributes to the degradation of truncated RNA polymerase I transcripts. These abortive transcripts are also targeted for 3'-5' degradation by the exosome (Wang & Pestov, 2011).

Another function of Rat1 in yeast is the 5' processing of small non-coding RNAs such as snRNAs and snoRNAs. Rat1 (XRN2) also participates in transcription termination of mRNAs produced by RNA polymerase II by a mechanism known as "torpedo" (Tollervey, 2004). Poly(A) sequences or gene-specific co-transcriptional cleavage sites trigger the release of transcribed mRNAs from the RNA polymerase. Rat1 (XRN2) is recruited to the carboxy-terminal domain (CTD) of the polymerase (Kim et al, 2004; West et al, 2004). Following release of the mRNA, the RNA still being transcribed by the polymerase is degraded by Rat1 (XRN2) until the exonuclease catches up with the polymerase and terminates transcription by "torpedoing" the polymerase off the template.

1.10.2 REX proteins

The REX proteins represent a poorly characterised family of 3'-5' exonucleases that were identified based on their sequence homology to other 3'-5' exonucleases including RNase D and RNase T (Moser et al, 1997). In yeast, four Rex proteins have been described: Rex1, Rex2, Rex3 and Rex4. These proteins appear to be well conserved with homologues of the yeast proteins REX1, REX2 and REX4 identified in *Xenopus* and humans. *Xenopus* REX4 (XPMC2) has a role in cell cycle progression (Su & Maller, 1995) but no RNA processing defects were detected in yeast upon deletion of the Rex4 gene (van Hoof et al, 2000). PMC2 (Rex4) in human cells has recently been shown to have 3'-5' exonuclease activity *in vitro* and this catalytic activity is required for the repair of estrogen-induced DNA damage (Krishnamurthy et al, 2011). Rex1, 2 and 3 have both specific and overlapping functions in RNA processing in yeast. Rex1 alone is required for the formation of the proper 3' ends of 5S rRNA and

tRNA–Arg3 while Rex2 has a specific function in the fine trimming of the 3' end of U4 snRNA (van Hoof et al, 2000). Defects in the formation of the 3' end of RNase MRP RNA are detected upon deletion of the Rex3 gene while redundancy is observed between Rex1, Rex2 and Rex3 in the formation of the 3' end of the related RNase P RNA (van Hoof et al, 2000). Rex1 and Rex2 have also been shown to be functionally redundant in the maturation of 5.8S rRNA with co-deletion of these two genes causing accumulation of 3' extended precursors (van Hoof et al, 2000). Interestingly, this implicates the Rex proteins in RNA processing events also attributed to the exosome, suggesting there may be functional redundancy between multiple exonucleases performing the same 3' processing events. This redundancy is supported by the observation that yeast strains lacking both Rrp6 and Rex1 (van Hoof et al, 2000) or Rrp47 and Rex1 (Costello et al, 2011) are not viable. However, deletion of specific exonucleases gives rise to subtle variations in the accumulation of aberrant precursors suggesting that different exonucleases may be required for particular steps within the processing or degradation pathways of a single precursor.

1.10.3 The exosome

The exosome is a multi-protein complex that functions as a 3'-5' exonuclease and is highly conserved from archaea to humans (Hartung & Hopfner, 2009; Mitchell et al, 1997). The exosome was first discovered in *S. cerevisiae* as an enzyme required for formation of the mature 3' end of 5.8S rRNA (Mitchell et al, 1996). Many of the proteins of the human exosome complex were first identified as autoantigens that give rise to the autoimmune disorders of polymyositis, scleroderma (Scl) and the PM/Scl overlap syndrome (Brouwer et al, 2001; Gelpi et al, 1990). Since then, the exosome has been shown to be responsible for the processing and/or degradation of a range of other RNA substrates both in the nucleus and cytoplasm (Figure 1.8) (Houseley et al, 2006; Kiss & Andrulis, 2011; Lykke-Andersen et al, 2009; Schmid & Jensen, 2008b; Tomecki et al, 2010a; Vanacova & Stefl, 2007). The exosome is involved in the 3' processing of pre-sno/snRNAs, pre-tRNA as well as pre-rRNA (5.8S) in the nucleus (Allmang et al, 1999a). In the nucleus the exosome also participates in the recycling of the by-products of RNA processing such as the fragments of pre-rRNA released by serial endonucleolytic cleavages. The exosome plays an important role in RNA surveillance and defective sn/snoRNAs, pre-rRNAs, tRNAs or mRNAs which are not correctly processed, spliced or assembled into RNP complexes are targeted for degradation by the exosome. Other RNA polymerase II transcripts such as cryptic unstable transcripts

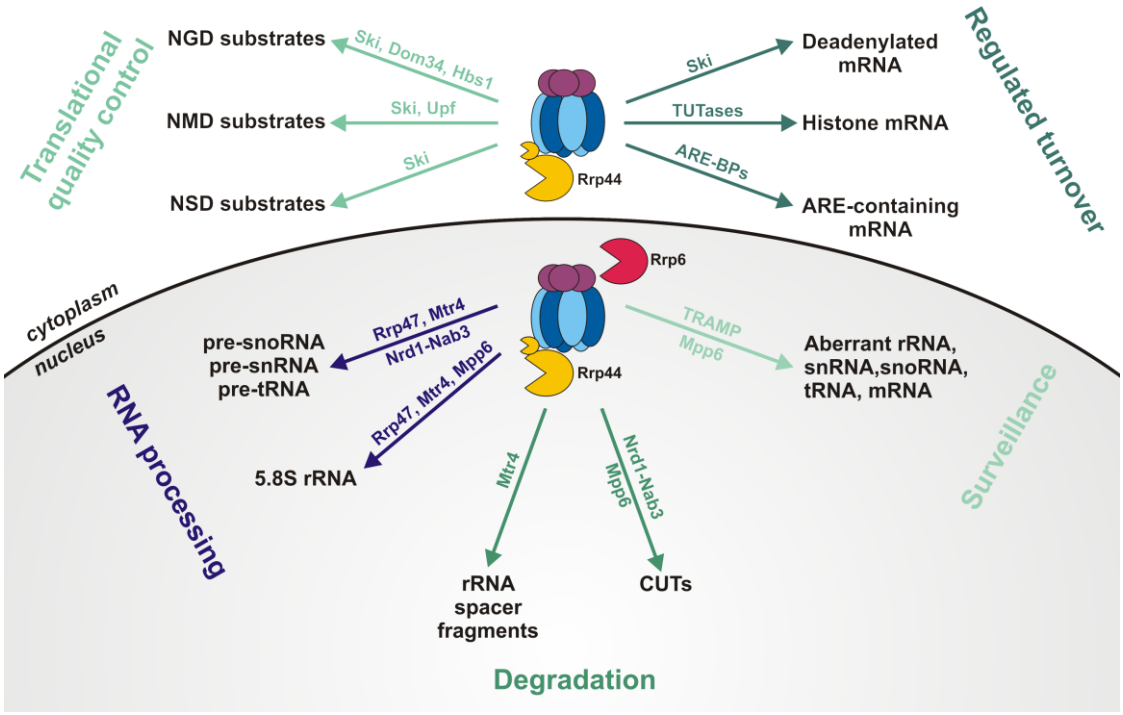


Figure 1.8 The yeast exosome participates in various pathways of RNA metabolism The core exosome (blue/purple ovals) is associated with Rrp44 (yellow) and Rrp6 (red) in the nucleus. Various substrates of the nuclear and cytoplasmic exosomes are shown in black. Cofactors identified as participating in each pathway are shown along the arrow.

(CUTs), which regulate gene expression in yeast, and promoter upstream transcripts (PROMPTs) which are found in human cells are turned over by the exosome in the nucleus (Chekanova et al, 2007; Milligan et al, 2008; Neil et al, 2009; Preker et al, 2008). In the cytoplasm the exosome also has a quality control function in preventing the translation of defective mature mRNAs and as such is involved in the nonsense mediated decay (NMD) pathway which eliminates transcripts containing a premature stop codon, the nonstop decay (NSD) pathway which removes mRNAs lacking a stop codon and the no-go decay (NGD) pathway which degrades mRNAs that could not be translated by the ribosome (Isken & Maquat, 2007). The cytoplasmic exosome is involved in more specific degradation pathways including the degradation of transcripts containing AU-rich elements (AREs) to which the exosome is directy recruited, the turnover of viral RNAs containing zinc-finger protein response elements (ZREs) and that of mRNA byproducts produced by the RNA interference pathway (Chen et al, 2001; Gherzi et al, 2004; Guo et al, 2007; Mukherjee et al, 2002; Orban & Izaurralde, 2005). The cytoplasmic exosome is important for maintaining the levels of histone mRNAs and TUTases are proposed to add oligouridine tails to these mRNAs which recruits both LSm proteins and the exosome bringing about their subsequent degradation (Mullen & Marzluff, 2008). It is still not clear how the exosome is recruited

specifically to each of its substrates or how the activity of the exosome is regulated to perform the different functions of processing and degradation.

1.10.3.1 Core exosome structure

The overall structural organisation of eukaryotic exosomes is conserved from prokaryotic polynucleotide phosphorylases (PNPases) which are part of the bacterial machinery for RNA degradation (Figure 1.9) (Buttner et al, 2005; Liu et al, 2006; Lorentzen et al, 2005). PNPases form a six membered homotrimeric ring composed of proteins which are homologous to RNase PH. Each protein contains two PH domains and only one molecule in each dimer has ribonuclease activity (Figure 1.9) (Carpousis, 2002; Symmons et al, 2000).

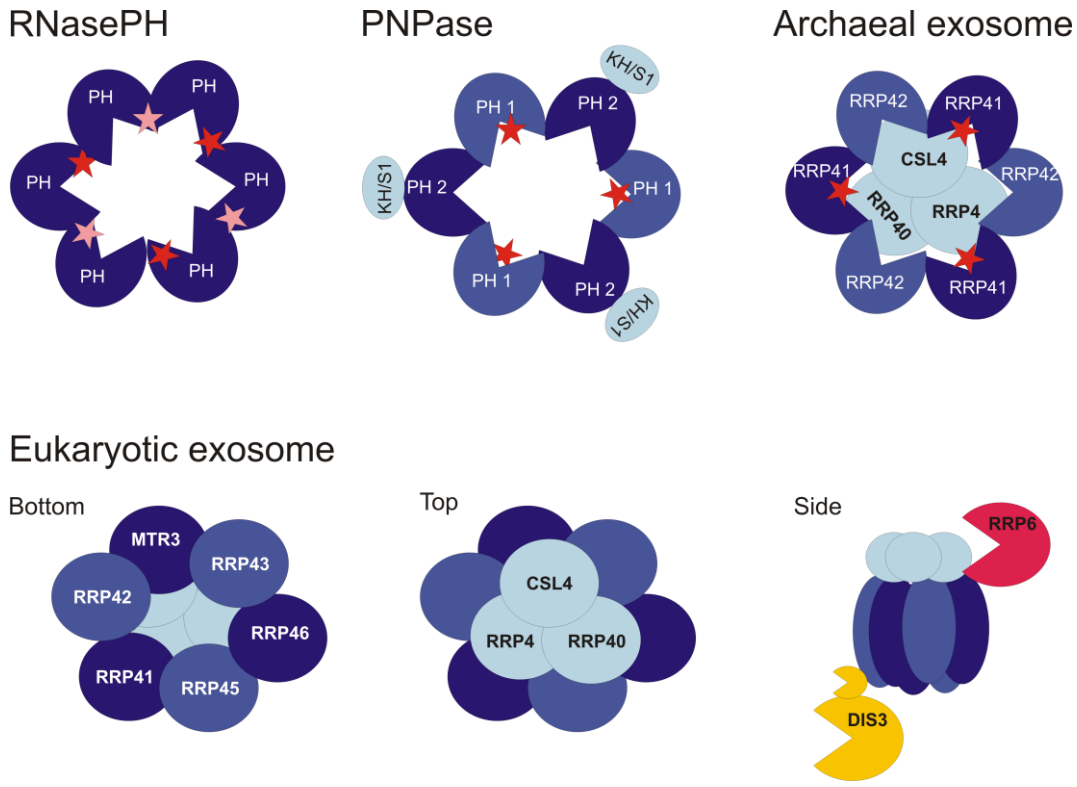


Figure 1.9 RNA degradation complexes are structurally conserved from bacteria to eukaryotes Schematic representations of RNase PH and PNPase from bacteria, the archaeal exosome and the eukaryotic exosome. Dark blue and mid-blue ovals represent PH domain proteins homologous to archaeal RRP41 and RRP42, respectively. Pale blue circles correspond to the trimeric cap formed by the RNA-binding domain proteins containing KH and S1 domains. Red and pink stars represent the active sites of the core proteins in bacteria and archaea (pink stars indicate that the active site is on the opposite face of the hexameric ring). Eukaryotic core exosomes are inactive and associate with RNase D-like RRP6, shown in red, which probably interacts in the proximity of the trimeric cap and the exo-/endonuclease, DIS3 (yellow) which is anchored to the lower face of the core exosome via its PIN domain (small yellow shape).

Similarly, the archaeal exosome is a hexameric ring composed of three dimers of RNase PH homologues, Rrp41 and Rrp42. Only the Rrp41 subunits have retained catalytic activity and as such, three active sites are present in archaeal exosomes which are directed towards a central core channel (Figure 1.9) (Lorentzen et al, 2005). This channel is only compatible with single stranded substrates and is thought to direct RNA substrates toward the active sites of the complex. In archaeal exosomes, additional proteins, Rrp4 and Csl4, bind to the Rrp42 subunits of the core exosome. Rrp4 and Csl4 proteins contain KH and S1 domains which function in RNA binding and therefore, probably substrate recruitment (Hartung & Hopfner, 2009). These domains are also found in the bacterial PNPase proteins.

Eukaryotic exosomes also consist of nine core proteins which share homology with their archaeal and bacterial counterparts and are organised structurally into a hexameric ring and cap (Figure 1.9). Crystal structures of both the yeast and human exosomes have been published (Liu et al, 2006). Similar to archaeal RRP41 are the eukaryotic subunits, RRP41, RRP46 and MTR3 and related to RRP42 from the archaeal complex are the eukaryotic proteins, RRP42, RRP43 (OIP2 in humans) and RRP45 (PM/Scl75 in humans). These are arranged in RRP41-RRP42-like dimers consisting of RRP41-RRP45, RRP42-MTR3 and RRP43-RRP46, which assemble to form the doughnut ring shape. Eukaryotic exosomes contain three KH/S1 domain proteins, RRP4, RRP40 and CSL4 which in higher eukaryotes are required for the stability of the hexameric ring as well as for RNA binding and substrate recognition. Crystal structures of the exosome show that these proteins are clustered at the entrance to the channel on one side of the ring enabling substrates to be targeted through the central pore (Liu et al, 2006).

The catalytic amino acid residues of the archaeal PH domain proteins are not conserved in five out of the six eukaryotic PH domain proteins and both yeast and human core exosomes have lost nuclease activity (Dziembowski et al, 2007; Liu et al, 2006; Schneider et al, 2007) which instead is derived by association with exonuclease subunits, RRP6 and DIS3/DIS3L (Rrp44 in yeast).

1.10.3.2 DIS3/DIS3L (Rrp44)

In yeast, Rrp44 is considered a tenth core exosome subunit as it is constitutively associated with exosome complexes in both the nucleus and cytoplasm and is essential for cell viability (Allmang et al, 1999b; Dziembowski et al, 2007; Liu et al, 2006; Mitchell et al, 1997). Rrp44 is homologous to *E. coli* RNase R, a member of the RNase II family of hydrolytic 3'-5' exonucleases (Vincent & Deutscher, 2006). Like

RNase R, Rrp44 contains a cold shock RNA binding domain (CDS), an RNase II catalytic domain (RNB) and a C-terminal S1 RNA binding domain. The RNB domain contains four highly conserved acidic residues responsible for chelating a divalent metal ion (Mg^{2+}) required for nuclease activity. Rrp44 is a highly processive enzyme and is able to degrade both structured and unstructured RNA substrates although its activity has been shown to be significantly decreased by its association with the core exosome (Lorentzen et al, 2008). Rrp44 requires 3' single stranded overhangs of approximately 14 nucleotides to initiate degradation of a duplex substrate (Lorentzen et al, 2008). Rrp44 additionally contains a PIN domain at its N terminus (Lebreton et al, 2008; Schaeffer et al, 2009; Schneider et al, 2007). It was recently shown by three different groups that this PIN domain endows Rrp44 with *in vitro* endonuclease activity in addition to its well characterised 3'-5' exonuclease activity (Lebreton et al, 2008; Schaeffer et al, 2009; Schneider et al, 2009). It is not yet clear what the specific functions of the endonuclease activity are but Rrp44 mutants deficient in either exo- or both exo- and endonuclease activity show different pre-rRNA processing defects particularly in the degradation of cleaved 5'ETS fragments accumulating after A_0 cleavage (Lebreton et al, 2008; Schaeffer et al, 2009; Schneider et al, 2009). As well as this catalytic role, the PIN domain of Rrp44 is essential for tethering Rrp44 to the core exosome (Schneider et al, 2009). Rrp44 associates with the lower face of the core exosome, opposite to the trimeric cap (Figure 1.9). A model of RNA processing is proposed in which single stranded substrates pass though the central pore of the hexameric ring and are directed towards the active site of the exonuclease domain (Bonneau et al, 2009). Structured RNA substrates are excluded from the core exosome and are thought to be targeted directly to the PIN domain where they are cleaved into fragments which can be degraded exonucleolytically.

While only one form of Rrp44 is present in yeast, three different human homologues have been identified: DIS3, DIS3L1 and DIS3L2. Of these, only DIS3 and DIS3L are associated with the exosome (Staals et al, 2010; Tomecki et al, 2010b). Both these proteins were found to be less stably associated with the core exosome than yeast Rrp44 is (Tomecki et al, 2010b). The PIN domain of DIS3L2 is poorly conserved and as this PIN domain is required for interaction with the exosome this is likely to explain why DIS3L2 is not found in purified exosome complexes (Tomecki et al, 2010b). Distinct subcellular localisation of these proteins was observed with DIS3 being found predominantly in the nucleus but excluded from nucleoli, whereas DIS3L was found exclusively in the cytoplasm (Tomecki et al, 2010b). While both DIS3 and DIS3L have 3'-5' exonuclease activity *in vitro*, only DIS3 was found to have endonuclease activity like its yeast counterpart (Tomecki et al, 2010b). Exosome

associated RNA processing activities of both DIS3 and DIS3L have been demonstrated *in vivo* and RNAi mediated depletion of DIS3 but not DIS3L leads to the stabilisation of PROMPTS in the nucleus and accumulation of extended 3' precursors of 5.8S rRNA (Tomecki et al, 2010b). In the cytoplasm, depletion of DIS3L causes significant stabilisation of cMyc and cFos mRNAs (Tomecki et al, 2010b) and the accumulation of polyadenylated 28S rRNA fragments (Staals et al, 2010) which are both normally degraded by the exosome.

1.10.3.3 RRP6

The other exoribonuclease associated with the exosome is RRP6 (also called PM/Scl100 in humans). RRP6 shows homology to *E. coli* RNase D and, as such, is a member of the DEDD superfamily of 3'-5' ribonucleases and deoxyribonucleases (Midtgaard et al, 2006; Steitz & Steitz, 1993). In addition to the exonuclease (RNB) domain which contains the four conserved catalytic amino acid resides that constitute the DEDD box motif, RRP6 also contains a single HRDC domain. The precise function of this domain is unknown but it may be important for regulating substrate binding by guiding substrates to the catalytic site (Liu et al, 2006; Midtgaard et al, 2006). Other important regions of Rrp6 include the PMC2NT domain which mediates interactions with the cofactor, Rrp47 and a region to the C-terminal side of the HRDC domain which is important for interactions with the core exosome (Callahan & Butler, 2008). Rrp6 also contains a putative nuclear localisation signal (NLS) (Phillips & Butler, 2003). A recent crystal structure of human RRP6 shows that these domains are conserved in the human protein (Januszyk et al, 2011). It is not clear how Rrp6 interacts with the core exosome and different models are proposed. A cryo-EM structure from *Leishmania tarentole* reports Rrp6 binding to the trimeric cap of the core exosome (Cristodero et al, 2008). This places Rrp6 at the mouth of the central pore and close to the RNA binding proteins of the core exosome. Alternatively, yeast two hybrid data suggest that human RRP6 interacts on the opposite face, in direct contact with the hexameric ring (Lehner & Sanderson, 2004). Rrp6 is also able to act independently of the core exosome. Rrp6 lacking the C-terminal region which tethers it to the core exosome, is still able to participate in 5.8S rRNA processing and snoRNA maturation in yeast although other functions, such as degradation of long polyadenylated substrates, are inhibited (Callahan & Butler, 2008). Both yeast and human RRP6 have processive exonucleolytic activity and in yeast, Rrp6 is only capable of degrading unstructured RNA substrates. In contrast, human RRP6 is able to degrade structured substrates, probably due to a more open conformation of the catalytic site in the human protein (Januszyk et al, 2011). While both yeast and human RRP6 are localised in the nucleus,

a key difference is that the human protein is concentrated in the nucleoli and can also be detected in the cytoplasm (Tomecki et al, 2010b).

1.10.3.4 Cofactors

1.10.3.4.1 C1D (Rrp47)

Rrp47 (or Lrp1 in yeast and C1D in humans) is a nuclear protein that in yeast co-purifies with exosome complexes containing Rrp6 (Mitchell et al, 2003; Schilders et al, 2007). Deletion of Rrp47 causes RNA processing defects similar to those caused by the lack of Rrp6 and indeed Rrp47 has been shown to interact with the PMC2NT domain of Rrp6 (Mitchell et al, 2003; Stead et al, 2007). Rrp47 (C1D) is known to participate in 3' processing of 5.8S in both yeast and human cells and is required for maturation of both snoRNAs and snRNAs in yeast (Costello et al, 2011; Schilders et al, 2007). Yeast Rrp47 is an RNA binding protein that *in vitro* can exist as a hexameric complex and binds preferentially to structured RNA substrates (Stead et al, 2007). Rrp47 is able to bind concurrently to both Rrp6 and to nucleotides suggesting that it may function as an adapter recruiting the Rrp6-exosome to its substrates. This model is further supported by evidence of helical or double stranded regions at the 3' ends of known exosome substrates, such as pre-5.8S and CUTs. Rrp47 contains an unusual RNA binding domain called a Sas10/C1D domain which is also found in pre-rRNA binding factors, UTP3 and Lcp5 (Costello et al, 2011). As well as its role in RNA binding, the C-terminal domain of Rrp47 mediates important protein-protein interactions during snoRNP biogenesis. Recombinant Rrp47 interacts with core snoRNP proteins Nop56 and Nop58. A model is proposed in which mis-assembly of snoRNP particles inhibits recruitment of Rrp47, which in turn promotes oligoadenylation of the snoRNA targeting it for degradation by the exosome. In this way, Rrp47 is thought to function in snoRNP quality control (Costello et al, 2011).

1.10.3.4.2 MPP6

Another protein which co-purifies with nuclear exosome complexes is the M-phase phosphoprotein 6, MPP6 (Milligan et al, 2008). MPP6 is an RNA binding protein that interacts most strongly with pyrimidine-rich sequences and localises predominantly to the nucleoli of human cells (Schilders et al, 2005). MPP6 is required for 3' processing of 5.8S rRNA in both yeast and human cells (Milligan et al, 2008; Schilders et al, 2005). In yeast, Mpp6 is also involved in exosome degradation of unstable non-coding RNAs produced by RNA polymerase II (Milligan et al, 2008).

1.10.3.4.3 TRAMP complex

A cofactor of the nuclear exosome that has been well characterised in yeast is the Trf4/5-Air1/2-Mtr4 polyadenylation (TRAMP) complex (LaCava et al, 2005; Vanacova et al, 2005). TRAMP recruits the exosome to its RNA substrates by adding short poly (A) tails and stimulating exonucleolytic degradation by the exosome, thereby playing a key role in RNA surveillance and quality control. The TRAMP complex in *S.cerevisiae* is composed of a non-canonical poly(A) polymerase (either Trf4 or Trf5), an RNA binding protein (either Air1 or Air2) and the RNA helicase, Mtr4 (LaCava et al, 2005; Vanacova et al, 2005). Two different TRAMP complexes have been identified; the first containing Trf4-Air1/2-Mtr4 and the second comprised of Trf5-Air1-Mtr4 (LaCava et al, 2005; Vanacova et al, 2005).

Trf4/5 are distributive polymerases that are only able to add short poly(A) tails to RNA substrates in the context of the TRAMP complex but are inactive by themselves. The polyadenylation activity of Trf4/5 is regulated by the associated helicase, Mtr4, which modulates the relative kinetics of adenylation and TRAMP dissociation to restrict the number of adenosines added to approximately four to five (Jia et al, 2011; Wlotzka et al, 2011). Trf4 and Trf5 are highly similar (48% identical) and are functionally redundant although Trf5 activity is restricted to the nucleolus while Trf4 only functions in the nucleus. Interestingly, evidence has been presented that suggests that Trf4 is able to target aberrant RNAs for exosome degradation independently of its polymerase activity (Rougemaille et al, 2007; San Paolo et al, 2009; Wyers et al, 2005). Similarly, the two Air proteins found in TRAMP complexes are highly related (45% identical) and functionally redundant, with only deletion of both genes leading to a growth defect. The Air proteins contain five zinc knuckle motifs proposed to function in RNA binding making it likely that the Air proteins are critical for substrate specificity of the TRAMP complex. A recent crystal structure of a partial TRAMP complex and mutational studies showed that ZnK4 and ZnK5 are, however, responsible for mediating protein-protein interactions between the Air protein and the central domain of Trf4 (Fasken et al, 2011; Hamill et al, 2010).

Mtr4 (also known as Dob1) is an ATP-dependant RNA helicase which belongs to the DExH/D superfamily of RNA helicases which also includes the cytoplasmic exosome cofactor, Ski2 and Brr2, an essential splicing factor. *In vitro* Mtr4 is able to unwind RNA duplexes with 3’ to 5’ polarity (Wang et al, 2008). The interaction of Mtr4 with the TRAMP complex proteins, Trf4-Air2, is mediated by the central DExH/D core (Weir et al, 2010).

In addition to recruitment of the exosome to substrates targeted for degradation by the addition of polyadenylated tails, the TRAMP complex acts as an exosome cofactor by directly stimulating its exonuclease activity. In yeast, TRAMP has been shown to stimulate specifically the activity of Rrp6 *in vitro* but it does not increase the activity of the core exosome containing Rrp44 (Callahan & Butler, 2010). However, an alternative report shows that degradation of $tRNA_i^{Met}$ by Rrp44 without the core exosome is enhanced in the presence of the TRAMP complex *in vitro* (Schneider et al, 2007). This stimulation does not require ATP implying that the polyadenylation activities of Trf4/5 do not participate in this (Callahan & Butler, 2010).

It is thought that the TRAMP complex is conserved in higher eukaryotes. The human homologue of the helicase, Mtr4, has been characterised as a nuclear protein that is strongly concentrated in nucleoli. Putative homologues of the Trf4/5 and Air1/2 proteins have so far been identified. ZCCHC7 is a strictly nucleolar protein that contains zinc knuckles and an IWRxY motif found in yeast Air proteins and, as such, is proposed as a human homologue (Fasken et al, 2011; Lubas et al, 2011). Based on sequence homology, a human counterpart of yTrf4 is proposed, hTRF4-2/PAPD5. ZCCHC7 and hTRF4-2 interact directly and have been shown to be precipitated by both MTR4 and RRP6 showing that they are associated with the exosome (Fasken et al, 2011; Lubas et al, 2011). RNAi depletion of either ZCCHC7 or hTRF4-2 leads to a decrease in the accumulation of polyadenylated 5'ETS pre-rRNA fragments normally turned over by the nuclear exosome confirming that these proteins are functional cofactors of the human exosome (Lubas et al, 2011). Analogous to the different TRAMP complexes found in the nucleus and nucleolus of yeast, another protein complex has been identified in the nucleus of human cells called the nuclear exosome targeting (NEXT) complex. In addition to MTR4, this complex also contains ZCCHC8 and RBM7 and is involved in the degradation of PROMPTs but does not appear to participate in 3' processing of 5.8S rRNA (Lubas et al, 2011).

1.10.3.4.4 MTR4 and the exosome

Mtr4 has distinct functions not only as part of the TRAMP complex but also as a cofactor of the nuclear exosome. The TRAMP complex stimulates the exonuclease activity of Rrp6 *in vitro* and recruits the exosome to its substrates but the helicase activity of Mtr4 is also proposed to unwind RNA secondary structures and displace RNA-associated proteins ahead of exonucleolytic processing by the exosome. Crystal structure models of *S. cerevisiae* Mtr4 show a helicase domain which is highly similar to that of other DExH/D helicases, such as Prp43. Interestingly, the most conserved

feature of Mtr4 (and the related exosome cofactor, Ski2) forms an interface below the helicase core, at the putative RNA exit site which appears to mediate the interaction with the exosome (Jackson et al, 2010; Weir et al, 2010). A model is proposed in which the ATPase activity of Mtr4 drives RNA substrates through the helicase domain, analogous to the mechanism used by the proteosome, and as these RNAs exit they are directly channelled towards Rrp6 and the core exosome for processing (Jackson et al, 2010). The crystal structure also revealed a novel arch domain that is not found in canonical helicases. This arch structure is essential for the function of Mtr4 in the 3' end processing of 5.8S and disruption of this arch causes accumulation of 3' extended RNA fragments similar to those seen following deletion of Rrp6. In addition to its functions in RNA surveillance as part of the TRAMP complex, MTR4 has been shown to be an independent exosome cofactor in the processing of the 3' end of 5.8S rRNA in yeast, plants and human cells (de la Cruz et al, 1998; Lange et al, 2011; Schilders et al, 2007).

1.10.3.4.5 Nrd1/Nab3 complex

In yeast, the Nrd1/Nab3 complex coordinates the termination of RNA polymerase II transcription of snRNAs and snoRNAs with initiation of precursor processing by the exosome and TRAMP complexes. This complex consists of the RNA binding proteins, Nrd1 and Nab3 and the helicase, Sen1. First, this complex recognises specific RNA sequences and triggers transcription termination. The exosome is then recruited to the nascent 3' ends of the RNAs enabling processing to form mature snRNA and snoRNAs. The exosome is also recruited to the newly formed 3' end of other RNA polymerase II transcripts such as CUTs by the Nrd1-Nab3 complex, but in these cases the exosome completely degrades the substrate rather than performing limited precursor processing (Arigo et al, 2006; Thiebaut et al, 2006). The proteins of the Nrd1-Nab3 complex have homologues in higher eukaryotes although it is not clear whether the function of this complex is conserved. Alternative transcription termination mechanisms for snRNA genes have been identified (Baillat et al, 2005) and the majority of human snoRNAs are not produced by transcription as they are in yeast but are instead released from spliced introns (Kiss et al, 2006) possibly negating a role for this complex in higher eukaryotes.

1.10.3.4.6 Other exosome cofactors

The Ski2 helicase is an essential cofactor of the cytoplasmic exosome. This helicase, which is related to Mtr4 is part of the SKI complex which also contains Ski3, Ski8 and Ski7 which is homologous to the GTP-binding translation elongation factor eF1A. Ski7 interacts with both the SKI complex and the exosome and as part of the NGD pathway is thought to bind to stalled ribosomes at their A site, thereby recruiting the exosome to initiate degradation of the aberrant mRNAs which have caused the ribosome to arrest. The NMD and NSD pathways which degrade mRNAs that contain either premature or no stop codons, respectively, involve 5'-3' exonucleolytic processing by Xrn1, but also the 3'-5' exosome is recruited to these transcripts by the SKI complex and significantly contributes to their turnover. Homologues of Ski proteins are found only in yeasts but a SKI-like complex has also been identified in human cells. It is suggested that the human SKI complex may have additional functions as it is found in the nucleus as well as the cytoplasm and more particularly at the site of actively transcribed genes (Zhu et al, 2005).

Another multiprotein complex which has been identified as a cofactor of the yeast nuclear exosome is the Ccr4-Not complex. This complex was initially characterised for its roles in mRNA degradation in the cytoplasm and for regulation of transcription by RNA polymerase II in the nucleus. Mutations in certain subunits of this complex cause the accumulation of polyadenylated snRNA and snoRNAs similar to that observed in mutants of the nuclear exosome. Synthetic growth defects occur when mutations in Ccr4-Not subunits are combined with deletion of exosome factors suggesting that these complexes are functionally linked (Azzouz et al, 2009). The Ccr4-Not complex physically interacts with several subunits of the exosome and with the TRAMP helicase, Mtr4. It is suggested the Ccr4-Not complex protein Caf4 is necessary for the association of Mtr4 with Rrp6 (Azzouz et al, 2009). It is not clear whether this complex simply provides a scaffold through which the exosome and TRAMP complexes interact or whether the deadenylation and transcriptional regulation functions of the Ccr4-Not complex are also linked to exosome function. It is interesting to note that a nuclease involved in 3' processing of 5.8S rRNA, Ngl2, is closely related to the deadenylase, Ccr4, and both contain a magnesium-dependant endonuclease motif (Faber et al, 2002).

1.11 Aims and Objectives

The exosome was first discovered in the yeast, *S. cerevisiae*, and has been extensively characterised in this organism. The catalytic activities of Rrp44 and Rrp6 and the specific roles of a number of cofactors which regulate the activity of the exosome have been investigated. Many different RNA substrates of the yeast exosome have been identified and models of RNA processing by the exosome outlined. However, it has become clear that there are key differences between the yeast and human exosomes. Three ribonucleases, DIS3, DIS3L and RRP6, are associated with distinct exosome complexes in particular compartments of human cells whereas the yeast DIS3 counterpart, Rrp44, which is constitutively associated with all exosome complexes throughout the cell. This suggests that different subunits and cofactors may be involved in processing particular exosome substrates in human cells and possibly implies that RRP6 may play a more prominent role in human cells. Some of the functions of the exosome identified in yeast have been shown to be conserved in human cells including the role of the exosome and its cofactors in processing the 3' end of 5.8S rRNA. In yeast, other functions of the exosome in pre-rRNA processing, including the degradation of excised spacer fragments have been identified. In higher eukaryotes, removal of the 5'ETS involves an additional processing step, A' cleavage, and in mouse cells, exonucleases have been shown to recycle the cleaved fragments and also participate in A' cleavage. It is not known whether this is also the case in human cells.

In yeast, one of the functions of the exosome is the processing of precursor snoRNAs into their mature form and this involves two cofactors of the nuclear exosome, Rrp47 and the TRAMP complex. As described in section 1.3.3, the human exosome has been shown to be stably associated with two snoRNP complexes, pre-U3 and pre-U8. Since this is the only example of the human exosome directly interacting with a likely substrate, pre-snoRNA processing and snoRNP biogenesis provide a good model by which exosome function in human cells can be studied.

It is likely that the human exosome has additional functions in higher eukaryotes due to the increased complexity of RNA processing. It is suggested that exonucleases may be responsible for a key step in pre-rRNA processing of ITS1. Compared to yeast, processing to form the 3' end of 18S in human cells includes an additional step not found in yeast but currently two different models are proposed for how this is achieved. The first directly involves exonucleolytic processing following cleavage to separate the large and small subunit rRNAs (Carron et al, 2011) and if this is the case then the exosome is a likely candidate to be involved. The alternative pathway involves serial

endonucleolytic cleavages in ITS1 and, in mouse cells, it is suggested that the exonuclease, XRN2, may regulate this pathway (Wang & Pestov, 2011).

The objectives of this study were to characterise the human nuclear exosome and to investigate whether some of the roles of the yeast exosome are conserved in human cells. A major aim was to identify novel roles for the human exosome in RNA processing with particular focus on a potential function in ITS1 pre-rRNA processing. This study therefore aimed to;

- Characterise the structure and activity of the human nuclear exosome and cofactors
- Evaluate the roles of the exosome in pre-snoRNA processing and investigate how this is coordinated with snoRNP biogenesis
- Determine if exonucleases are required for turnover of byproducts of pre-rRNA processing in human cells
- Clarify the pre-rRNA processing pathways used to remove ITS1 in human cells and investigate the potential roles of exonucleases in mediating this step.

Chapter Two

Materials and Methods

2.1 PCR and cloning

2.1.1. PCR

Polymerase chain reactions (PCRs) were used to amplify target sequences from cDNA or plasmid templates using different polymerases depending on the length of the region to be amplified. For genes ≤ 1kb, reactions were carried out using HotStar Taq polymerase (Qiagen) and for templates > 1kb Phusion Polymerase (Finnzymes) was used. Reactions were set up as detailed in Table 2.1A and thermal cycling carried out according to programmes as shown in Table 2.1B.

Table 2.1 PCR reaction mixtures and cycling conditions

A

HotStar Taq		**Phusion**	
Component	**Final Concentration**	**Component**	**Final Concentration**
10x Qiagen PCR Buffer	1 x	5 x HF Buffer	1 x
dNTP (10mM of each)	200 µM	dNTP (10 mM of each)	200 µM
Forward primer	1 µM	Forward primer	0.5 µM
Reverse Primer	1 µM	Reverse Primer	0.5 µM
Hot Start Taq	2.5 U	Phusion Polymerase	0.02 U
Template DNA	≈100 ng	Template DNA	10 ng
Water	to 100 µl	Water	To 50 µl

B

	Hot Star Taq			**Phusion**		
	Temp (^{o}C)	**Duration (min)**	**Cycles**	**Temp (^{o}C)**	**Duration (s)**	**Cycles**
Initial activation	95	15	x 1	98	30	x 1
Denaturation	94	1	x 35	98	10	x 35
Annealing	T_m-5^{o}C	1		T_m+3^{o}C	30	
Extension	72	1 min/kb		72	30 sec/kb	
Final Extension	72	10	x 1	72	300	x 1

2.1.2 Agarose gel electrophoresis

6 x DNA Loading Buffer (0.4 % orange G, 0.03 % bromophenol blue, 0.03 % xylene cyanol FF, 15 % Ficoll400, 10 mM Tris-HCl pH 7.5, 50 mM EDTA pH 8.0) was added to DNA samples to a final 1 x concentration. Samples were analysed on a 1 % (> 500 bp) or 2 % (≤ 500 bp) agarose-TBE gel containing 1 x SYBERSafe (Invitrogen) which was run in 1 x TBE (90 mM Tris-HCl, 90 mM boric acid, 2 mM EDTA pH 8.0) at

120 V for 30-60 min. DNA in the gel was visualised using a Syngene ultraviolet transilluminator.

2.1.3 DNA extraction from agarose gels and purification

DNA bands were excised from agarose gels and purified using a Spin Column Gel Extraction Kit (NBS Biologicals) according to the manufacturer's instructions.

2.1.4 Restriction digest

Digests of 0.5-2 µg DNA were carried out using 10 U of restriction enzyme(s) (Promega), 1 x Enzyme Buffer (as recommended by manufacturer) in a final volume of 20 µl at 37 ^{o}C for 2-4 h.

2.1.5 Ligation reactions

PCR products and vectors were digested with restriction enzymes as required and reactions set up to contain insert and vector in a 5:1 molar ratio respectively, 0.3 U T4 DNA Ligase (Promega) and 1 x T4 DNA Ligase Buffer (Promega) in a total volume of 10 µl. Reactions were incubated at 10 ^{o}C for 16-24 h.

When cloning using the pGEM-T Easy cloning system (Promega), ligation reactions were set up using 1 x Rapid Ligation Buffer, 50 ng pGEM-T Easy Vector, 5-10 ng PCR product (with A overhangs) and 0.3 U T4 DNA Ligase in a total volume of 10 µl. Reactions were incubated at room temperature for at least 1 h.

2.1.6 Transformation of chemically competent *E. coli*

Chemically competent TOP10 *Escherichia coli (E. coli)* were used during cloning ligation reactions and amplification of existing constructs. 50 µl of cells were thawed on ice and ligation reactions (section 2.1.5) or 5 ng plasmid DNA was added before incubation on ice for 30 min. Cells were heat shocked at 42 ^{o}C for 60 seconds and placed back on ice before addition of 500 µl Luria Broth (LB) and shaking at 37 ^{o}C for 1 h. Successful transformants were grown on selective plates containing relevant antibiotics overnight at 37 ^{o}C. Single colonies were used to inoculate 2.5 ml LB cultures which were grown overnight at 37 ^{o}C with agitation. DNA was extracted from the cells

in these cultures using the Spin Column Plasmid Miniprep Kit (NBS Biologicals) following the manufacturer's instructions.

2.1.7 DNA Sequencing

Sequencing of PCR products or plasmids was carried out by GATC Biotech. Sequencing primers annealing to vector sequences upstream and downstream of the cloned gene were generally used and in the case of RRP6 primers annealing to internal gene sequences were used to enable complete sequencing coverage of the open reading frame **(Table 2.2)**.

Table 2.2 RRP6 sequencing primers

Primer	Direction	Sequence
RRP6 seq1	Forward	GTACACCGCTTGTTAATGGCAGCC
RRP6 seq2	Forward	CTTGAACTCTATAGGAAGCAGAAG
RRP6 seq3	Forward	CCTCACTGGGACACCGTGCTCCCG

2.1.8 Site directed mutagenesis

Mutagenesis of single or multiple bases within various constructs was carried out using the QuikChange Site Directed Mutagenesis Kit (Stratagene). Clean DNA preparations of the plasmid with the target mutagenesis site and primers encoding the required modifications were used in PCR reactions. Individual PCR reactions were set up containing 50 ng, 20 ng and 5 ng of template DNA with 125 ng of forward and reverse primers, 0.4 mM dNTPs, 1 x *Pfu* Reaction Buffer and 1.25 U *Pfu*Ultra in a total volume of 50 µl. Reactions were heated to 95 ^{o}C for 30 s and then thermocycled through 95 ^{o}C for 30 s, 55 ^{o}C for 60 s and 68 ^{o}C for 1 min per kilobase of plasmid for 18 cycles. Reactions were cooled on ice and incubated at 37 ^{o}C for 2 h with 10 U DpnI to degrade the methylated DNA template. PCR reactions were then pooled and ethanol precipitated by addition of NaCl to a concentration of 250 mM and three reaction volumes of ethanol. Reaction samples were stored at -20 ^{o}C for 2-16 h before centrifugation at 13,000 rpm in a benchtop centrifuge to pellet the DNA. The pellet was washed in 70 % ethanol to remove contaminating salt and dried. DNA was resuspended in 5 µl water and transformed into 40 µl XL1-Blue Super Competent Cells (Stratagene) which are able to repair the nicked circular DNA generated by the earlier PCR. DNA was extracted from colonies produced and sequenced to confirm the presence of the desired mutation.

2.1.9 Exonuclease constructs

Plasmids for the expression of the nine core human exosome proteins, RRP4, RRP40, CSL4, RRP42, MTR3, RRP43, RRP46, RRP41 and RRP45, as N-terminal-His-tagged recombinant proteins in bacteria were kindly provided by Dr Christopher Lima (Liu et al, 2006). These constructs are based on the pRSFDuet vector (Novagen). Clones of human C1D (yeast Rrp47) and MPP6 in the pGEX4T3 vector for expression of human GST-C1D and GST-MPP6 proteins were kindly provided by Dr Ger Prujin (Schilders et al, 2005; Schilders et al, 2007). RRP6 was amplified from a plasmid template and cloned into the pGEX6P3 vector using the primers and restriction sites in Table 2.3 to enable expression of human RRP6 with an N-terminal GST-tag. A clone for the expression of yeast RRP6 with an N-terminal GST- tag using the vectors pGEX2T, was kindly provided by Dr Phil Mitchell (Stead et al, 2007).

Recombinant human RRP6 was used in exonuclease assays and it was therefore necessary to generate mutant proteins in which the exonuclease activity was inhibited, to demonstrate that any RNA degradation observed in such assays is being carried out by the exonuclease. In yeast, a mutation targeting a conserved aspartic acid residue of the DEDD motif of RNase D nucleases, D303A has been shown to abolish the exonuclease activity of RRP6 (Burkard & Butler, 2000). RRP6 is highly conserved and the corresponding amino acid in the human sequence is aspartic acid 313. The coding sequence in the RRP6-pGEX6P3 clone was modified by site directed mutagenesis to convert this amino acid to alanine using the primers shown in Table 2.3, to enable expression of inactive hRRP6.

To generate constructs for inducible, stable RNAi resistant expression of FLAG-tagged RRP6 from HEK293 cells, the gene coding for RRP6 was cloned into the pcDNA5-FLAG vector. This vector encodes a His-tag, a PreScission Protease site and a FLAG-tag upstream of the genes of interest in frame with a tetracycline inducible promoter. PCR reactions were carried out using the primers shown in Table 2.3 and the RRP6-pGEX6P3 plasmid as a template. The restriction sites detailed in Table 2.3 were used for cloning of the genes into the pcDNA5 vector. From these basic constructs, mutations were introduced that enable RNAi resistant expression of either wild-type or inactive RRP6. Depletion of RRP6 by RNAi is achieved by treatment with three siRNAs which all target the open reading frame (ORF) of RRP6. Site directed mutagenesis was carried out to introduce five sequence modifications within the target site of each of these three siRNAs while not affecting the amino acid sequence of the expressed protein. This mutagenesis was carried out serially in three rounds using the RRP6-pcDNA5 plasmid as a template and the primers shown in Table 2.3. At each stage presence of the mutations and fidelity of the remaining sequence was confirmed

by sequencing. Once this construct had been generated a further mutagenesis step was carried out to introduce the RRP6 D313A mutation that abolishes the exonuclease activity of RRP6 into the RNAi resistant plasmid.

Table 2.3 Primers used for generating RRP6 expression constructs

Construct	Primer direction	Restriction site	Primer Sequence
GST-RRP6-pGEX6P3	Forward	BamHI	GCGCGGATCCATGGCGCCACCCAGTACCCGGGAG
	Reverse	SalI	GCGCGTCGACTCATCTCTGTGGCCAGTTGTACCTGAA
RRP6D313A (exo)	Forward	Mutagenesis	GAATTGTCAGGAATTTGCAGTTAACTTGGAGCACCACTCTTACAGG
	Reverse		CCTGTAAGAGTGGTGCTCCAAGTTAACTGCAAATTCTGACAATTC
RRP6-pcDNA5	Forward	BglII/BamHI NotI	CGCGAGATCTATGGCGCCACCCAGTACCCGGG
	Reverse		CGCGGCGGCCGCTCTCTGTGGCCAGTTCTACCTGAAGCCC
RRP6 RNAi_1	Forward	Mutagenesis	GCTGTCAAGAAGAAAGCGGCGGAACAGGCTGCCCGGGAACAGGCAAAGGAGG
	Reverse		CGGGCAGCCTGTTCCGCCGCTTTCTTCTTGACAGCTTCTTTAGAGTCTTTTTG
RRP6 RNAi_2	Forward	Mutagenesis	GTGCATGAGTCGAGTGATGCAAATCATGGGTGTCGCAGCAACATTAAGG
	Reverse		CCATGATATTGCATCACTCGACTCATGCACTGAAGCAACCTGTCTCCCTG
RRP6 RNAi_3	Forward	Mutagenesis	GCAGGGAAGACGAGAGCTATGGCTATGTACTGCCAAACCACATGATGCTG
	Reverse		GCAGTACATAGCCATAGCTCTCGTCTTCCCTGCGAGCTGTTTTATCCCTCCAGG

2.1.9 SSU biogenesis factor constructs

Constructs containing the gene sequences for human DIM1, NOB1 and PNO1 in the pcDNA5 vector were kindly provided by Dr Andrew Knox (Watkins lab) and these were used as templates for PCR amplification of each of these genes using the primer shown in Table 2.4. cDNA clones containing the full length human RIO2 and ENP1 sequences in the pCMV-SPORT6 vector were purchased from imaGENES (IRATp970E122D and IRATp970B1277D, respectively). These were used as PCR templates with the primers detailed in Table 2.4. The PCR products generated were cut using the restriction enzymes shown and cloned into the pGEX6P1 vector to generate constructs from which these genes can be expressed as N-terminal, GST-tagged recombinant proteins in *E. coli*.

RIO2 has been shown to be a protein kinase in both yeast and humans. Two combined amino acid changes in RIO2 have been identified in the kinase active site that decrease this kinase activity (Zemp et al, 2009). Site directed mutagenesis of the RIO2-pGEX6P1 plasmid using the primers given in Table 2.4 in two sequential rounds

was used to generate a construct for expression of GST-tagged RIO2 K123AD246A (RIO2kd).

Further, to produce a construct for the inducible expression of FLAG-tagged RIO2 protein in HEK293 cells, the RIO2 gene was sub-cloned from RIO2-pGEX6P1 into the pcDNA5 vector by restriction digest using BamHI and XhoI and re-ligation with the new vector.

Table 2.4 Primers used in cloning SSU biogenesis factor genes for protein expression

Gene	Primer direction	Restriction site	Primer Sequence
DIM1	Forward	BamHI	CACCGGATCCATGCCGAAGGTCAAGTCGGGGGC
	Reverse	XhoI	CGCGCTCGAGCTAGGAAAAATGAATACCTTCTGC
NOB1	Forward	BamHI	CAC CGG ATC CATGGCTCCAGTGGAGCACGTTGTGG
	Reverse	XhoI	CGCGCTCGAGTCACCTTTTCTTTCACAAACTTCTTTC TGGAAGC
PNO1	Forward	BamHI	CACCGGATCCATGGAATCCGAAATGGAAACG
	Reverse	XhoI	CGCGCTCGAGTCAGAATCGATCTGCTGATCTGCTA GC
RIO2	Forward	BamHI	CGCGGGATCCATGGGAAAGTGAATGTGGCC
	Reverse	XhoI	CGCGCTCGAGTCATTCTCCCCAAAAGCTGGCTGC
ENP1	Forward	BglII	GCGCAGATCTATGCCCAAATTCAAGGCGGCCCCG
	Reverse	SalI	GCGCGTCGACCTCCACGGTGATGGGAACATCTTCC
RIO2 K123A	Forward	Mutagenesis	GAAGGACAACAATTTGCATTAAAGCTTCACAGACTA GGAAGAACC
	Reverse		GGTTCTTCCTAGTCTGTGAAGCTTTAATGCAAATTGT TGTCCTTC
RIO2 D246A	Forward	Mutagenesis	GTGACCATATCACCATGATTGATTTTCCACAGATGG TTTCAAC
	Reverse		GTTGAAACCATCTGTGGAAAATCAATCATGGTGATA TGGTCAC

2.1.10 snoRNP protein and biogenesis factors constructs

Constructs enabling expression of various GST- and His-tagged snoRNP proteins were kindly provided by Dr Kenneth McKeegan (Watkins lab) (McKeegan et al, 2007). GST-tagged fibrillarin, NOP58, 15.5K, TIP48, TIP49, NUFIP, BCD1 (residues 1-360) and NOP17 were expressed using clones of these genes in the pGEX6P1 vector. Further, plasmids for the expression of His-tagged proteins; fibrillarinΔRGG (residues 82-321) expressed from pET200a vector and 15.5K, TIP48 and TIP49 expressed from pETDuet were provided. Expression of N-terminally thioredoxin-tagged, C-terminally His-tagged NOP56 (residues 1-458) and NOP58 (1-435) was carried out using constructs in the pBAD/Thio-TOPO vector. It is not possible to express full length NOP56 and NOP58 but proteins lacking these C-terminal regions have been shown to function normally with respect to snoRNP formation in *S. cerevisiae* (Gautier et al, 1997; Lafontaine & Tollervey, 2000).

2.2 Protein over-expression and purification

2.2.1 GST-tagged proteins

Plasmids for the expression of GST-tagged proteins were transformed into BL21 Codon Plus *E. coli* cells which were then grown on selective LB agar plates containing the appropriate antibiotics (100 µg/ml ampicillin, 12.5 µg/ml chloramphenicol or 25 µg/ml kanamycin). Overnight cultures were then set up by inoculating a single colony into LB Broth (10 ml/L of expression culture) and grown, shaking at 37 ^{o}C for 16 h. These cultures were then used to inoculate 1 L (or multiple 1 L cultures) of 2 x YT broth containing the appropriate antibiotic which were then grown at 37 ^{o}C, shaking until an $O.D._{600} \approx$ 0.3-0.4 was reached. Cultures were then cooled to 18 ^{o}C, induced by adding isopropyl-beta-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and grown for a further 16 h. To check for protein over-expression 1 ml samples were taken pre- and post-induction and analysed using SDS-PAGE. Cells were harvested by centrifugation at 4000 rpm at 4 ^{o}C in a Beckman J6-HC centrifuge for 30 min. The supernatants were discarded and the pellets resuspended in either 20 ml Exosome Buffer (50 mM Tris-HCl pH 8.0, 10 % glycerol, 300 mM NaCl, 5 mM $MgCl_2$, 0.1 % Tween20) or 20 ml Purification Buffer (20 mM Tris-HCl pH 8.0, 10 % glycerol, 300 mM KCl, 5 mM $MgCl_2$, 0.1 % Tween20). Exonuclease proteins and cofactors were purified in Exosome Buffer and SSU biogenesis factors and snoRNP associated proteins were purified in Purification Buffer. A Complete EDTA Protease Inhibitor Cocktail Tablet (Roche) and Tris(hydroxypropyl)phosphine (THP) to a final concentration of 0.5 mM was added. Samples were sonicated on ice for 3 min, 0.5 cycle time, 90 % power to lyse the cells. Cell debris was then pelleted by centrifugation at 18,500 rpm at 4 ^{o}C for 45 min in a JA20 rotor and the clarified lysate used for protein purification. 1 ml Glutathione Sepharose Fast Flow beads suspended in 50 % ethanol were washed by precipitating by centrifugation at 3000 rpm and resuspension in 1 ml Exosome Buffer/Purification Buffer three times. The sepharose beads was then added to the clarified *E.coli* lysate and rotated slowly on a wheel at 4 ^{o}C for 2 h. The beads with bound protein were then sedimented by centrifugation at 4000 rpm for 3 min in a Fisher accuSpin R1 centrifuge and resuspended in 50 ml Exosome Buffer/Purification Buffer. This wash step was repeated and then the beads resuspended in 2ml of Exosome/Purification Buffer + 50 mM Glutathione pH8.0. The sample was rotated slowly at 4 ^{o}C for a further 1 h before pelletting the beads as previously and collecting the supernatant as the purified protein sample. Protein purification was confirmed by analysis of the sample by SDS-PAGE.

2.2.2 Purification of His-tagged proteins

Proteins were expressed from the plasmids described above in BL21 Codon Plus cells as for GST-tagged proteins and cells harvested, sonicated and a clarified lysate derived as described above. Tagged proteins were purified by immobilised metal (Ni^{2+}) affinity chromatography using 1 ml HiTrap Ni Sepharose columns (GE Healthcare) and the ÄKTA purification system (GE Healthcare). The 1ml HiTrap Ni Sepharose column was equilibrated with approximately 10ml Exosome/Purification Buffer using a peristaltic pump. The cleared lysate was loaded onto this column using the peristaltic pump. The column was then attached to the ÄKTA and washed with 10 ml 50 mM immidazole in Exosome/Purification Buffer, 5 ml 75 mM immidazole in Exosome/Purification Buffer, and then eluted using a gradient of 50-500 mM immidazole over 10 ml followed by 5 ml of 500 mM immidazole in Exosome/Purification Buffer. 1 ml fractions of both the wash step and elution stage were collected, snap frozen in liquid nitrogen and stored at -80 ^{o}C. The absorbance of fractions at 280 nm was followed on the ÄKTA with peaks indicating elution of proteins which was confirmed by SDS-PAGE.

2.2.3 Thioredoxin-/ His- tagged proteins

Recombinant NOP56 and NOP58 were expressed from the pBAD/Thio constructs in BL21 Codon Plus cells grown at 37 ^{o}C and cooled to 18 ^{o}C at an $O.D._{600} \approx$ 0.3-0.4 and expression induced by addition of 0.02 % arabinose for 16 h. Cells were harvested, sonicated and the proteins purified using HiTrap Ni Sepharose columns as described for His-tagged proteins.

2.2.4 Desalting

To remove the immidazole or glutathione used to elute proteins during their purification which would interfere with *in vitro* protein function, protein fractions were desalted using HiTrap Desalting columns (GE Healthcare). A desalting column was equilibrated with 25 ml of the appropriate protein purification or assay buffer and samples were applied to the column using the ÄKTA via an injecting loop. Samples were eluted in 500 µl fractions from the column over 10 ml with proteins and salt being eluted in different fractions. Protein elution was followed using an A_{280} trace and fractions corresponding to peaks were checked for the presence of protein using SDS-PAGE. Protein samples were aliquoted, snap frozen and stored at -80 ^{o}C.

2.2.5 PreScission protease cleavage

Proteins expressed from pGEX vectors are produced as GST fusions but contain a PreScission Protease site between the protein of interest and the GST-tag. PreScission Protease was expressed and purified via its own GST-tag as described above. Desalted protein was produced at a concentration of 1.8 mg/ml and 0.5 µl of this was used to cleave GST-tags off 100 µg protein at 4 ^{o}C for 1 h. PreScission Protease and cleaved GST-tags were then removed from the protein sample using 50 µl of Glutathione Sepharose Fast Flow beads.

2.3 Protein analysis methods

2.3.1 SDS-PAGE and Western Blot transfer

Protein samples were prepared by adding the necessary amount of Protein Loading Buffer (final concentrations; 74 mM Tris-HCl pH6.8, 1.25 mM EDTA, 20 % glycerol, 2.5 % SDS, 0.125 % bromophenol blue, 50 mM DTT). Protein samples were boiled at 95 ^{o}C for 2 min and vortexed prior to gel loading. Denaturing SDS polyacrylamide gels (SDS-PAGE) consisting of a 10 % or 13 % acrylamide resolving gel and a 4 % acrylamide stacking gel were used for large (≥ 70kDa) and small (<70kDa) proteins respectively. Electrophoresis was carried out on these gels at 200 V in 1 x Protein Running Buffer (25 mM Tris-HCl pH8.3, 250 mM glycine, 0.1 % SDS (w/v)). Proteins were visualised by incubation with Coomassie Blue Stain (0.1 % Coomassie Blue (w/v), 40 % methanol (v/v) and 10 % acetic acid (v/v)) for 1-16 h with gentle agitation. Gels were then de-stained by serial washes in 40 % methanol (v/v) with 10 % acetic acid (v/v) for as long as necessary.

2.3.2 Determination of protein concentration

The concentration of a protein sample was determined using a Nanodrop (Thermo Scientific). 1 µl of the protein sample was applied to the machine and the absorbance at 280 nm is used to determine the concentration.

2.3.3 Western blotting

Western blot analysis was used to detect specific proteins or protein tags in samples that had been separated by SDS-PAGE. Proteins were transferred from the polyacrylamide gel onto a nitrocellulose membrane (Protran, GE Healthcare) in

Table 2.5 Antibodies used in Western Blotting

Antibody	Raised in	Source/Reference
RRP46	Rabbit	G Pruijn
CSL4	Rabbit	Eurogentec (Watkins) Raised against full length recombinant protein – this study
RRP6	Rabbit	G Pruijn
C1D	Rabbit	Eurogentec (Watkins) Raised against full length recombinant protein – this study
MPP6	Rabbit	G Pruijn
DIS3	Mouse	Abnova
MTR4	Rabbit	Eurogentec (N Watkins/J Brown) Raised against peptide fragments: GDQKGRKGGTKGPSNV and TDEPIFGKKPRIEES
XRN2	Rabbit	Bethyl A301-103A
PNO1	Rabbit	Eurogentec (Watkins) Raised against full length recombinant protein
NOB1	Rabbit	Eurogentec (Watkins) Raised against full length recombinant protein
RIO2	Goat	Santa Cruz E-14
ENP1	Mouse	Santa Cruz C-19
PRP43-C	Rabbit	Bethyl A300-390A
PRP43-N	Rabbit	R. Luhrmann
NOP56	Rabbit	Bethyl A302-720A
NOP58	Rabbit	Eurogentec (Watkins)
Fibrillarin	Rabbit	Sigma Aldrich H140
UTP24	Rabbit	Eurogentec (Watkins) Raised against full length recombinant protein
RCL1	Rabbit	Eurogentec (Watkins) Raised against full length recombinant protein
BMS1	Rabbit	A.Turner (Turner et al., 2009)
RPS19	Rabbit	P. Mason (Idol et al., 2007)
RRP5	Rabbit	A.Turner (Turner et al., 2009)
mBOP1	Rabbit	D. Pestov (Strezoska et al. 2000)
RBM28	Rabbit	Santa Cruz Z-22 (sc-102075)
POP1	Rabbit	G. Pruijn/S. Altman (Mattijssen et al., 2010)
RRP40	Rabbit	G. Pruijn/S. Altman (Mattijssen et al., 2010)
Viperin	Mouse	G. Pruijn/P.Cresswell (Mattijssen et al., 2010)
Poly-HIS	Mouse	Santa Cruz H-3
GST	Goat	Santa Cruz B14
Thioredoxin	Mouse	Sigma-Aldrich T0803
FLAG	Mouse	Sigma Aldrich F1804
FLAG	Rabbit	Sigma Aldrich F7425
Rabbit IgG-HRP	Donkey	Santa Cruz sc-25397
Mouse IgG-HRP	Donkey	Santa Cruz sc-358914
Goat IgG-HRP	Donkey	Santa Cruz sc-2020

Transfer Buffer (25 mM Tris-HCl pH 8.3, 150 mM glycine, 10 % methanol) using a BioRad Western blotting transfer tank for 2 h at 60 V. Membranes were stained with Ponseau S Solution (Sigma) to verify the presence of proteins on the membrane. Membranes were then incubated in Blocking Buffer (PBS, 0.05 % Triton X-100 (v/v), 2 % Marvel skimmed milk power (w/v)) for 1 h at room temperature or overnight 4 ^{o}C to minimise non-specific binding of antibodies to the membrane. Primary antibodies were diluted to their optimal concentration in Blocking Buffer. A list of antibodies used in this study is given in Table 2.5. Membranes were incubated with the primary antibody for 1 h at room temperature before washing 3 x 5 min and 1 x 15 min with PBS, 0.05 % Triton-X100 (v/v). Secondary antibodies which were conjugated to horseradish

peroxidise (HRP) were diluted in Blocking Buffer to the appropriate concentration and applied to membranes for 1 h at room temperature. Washes were then repeated as previously. Proteins were then developed by enhanced chemiluminescence (ECL) detection (Pierce) and exposure of hyperfilm (GE Healthcare) according to the manufacturer's instructions.

2.3.4 Creation of custom antibodies

1mg of purified His-CSL4 and GST-C1D (Figure 2.1A) were sent to Eurogentec for production of polyclonal antibodies in rabbits. For each protein, two rabbits were immunised and serial bleeds taken (pre-Immume, PPI; small bleed, PP; large bleed, GP; final bleed, SAB). The crude serum antibodies were tested in preliminary Western blotting experiments on HeLa cell extract and the results of the preferred antibodies for each protein are shown in Figure 2.1B.

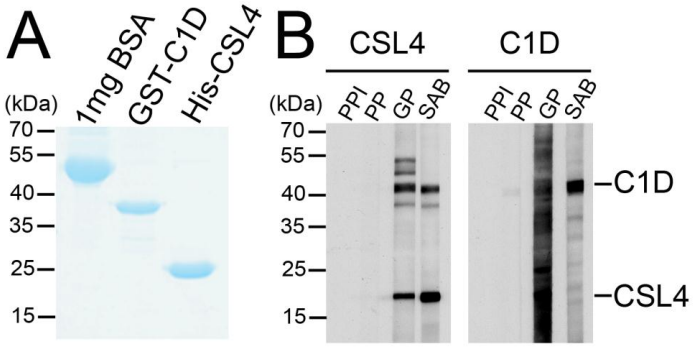


Figure 2.1 CSL4 and C1D antibodies A) Recombinant CSL4 and C1D were purified using their His- and GST-tags respectively. Proteins were separated by SDS-PAGE and visualised by Coomassie staining. 1mg of bovine serum albumin (BSA) indicates the relative concentrations of each protein. **B)** Mouse nuclear extract was separated by SDS-PAGE and transferred for Western blotting. Crude serum received from Eurogentec, pre-immune (PPI), small bleed (PP), large bleed (GP) and final bleed (SAB) were diluted 1:1000 (CSL4) or 1:250 (C1D) and used in Western blotting. The positions of marker proteins are indicated to the left of the panel and bands identified as CSL4 and C1D detected are indicated to the right of the panel.

2.3.4 Proteomics

Protein reactions to be analysed by proteomics were mixed with NuPAGE LDS Sample Buffer (26 mM Tris-HCl pH 8.5, 35 mM Tris Base, 0.5 % LDS (v/v), 2.5 % glycerol (v/v), 0.1275 mM EDTA, 0.05 mM SERVA Blue H250, 0.04 mM Phenol Red) and separated on 4-12 % NuPAGE Novex® Pre-cast Bis Tris Mini Gels in MOPS Running Buffer (2.5 mM MOPS, 2.5 mM Tris Base, 0.005 % SDS (w/v), 0.05 mM EDTA pH 7.7). Gels were then stained with fresh Coomassie Blue Stain overnight. After destaining in 50 % methanol, 10 % acetic acid, gels were stored in 5 % acetic acid. Proteomic analysis was carried out by Dr Henning Urlaub, (Max-Planck Institute for Biophysical Chemistry, Gottingen, Germany.

2.4 RNA analysis methods

2.4.1 RNA extraction

2.4.1.1 Phenol:chloroform extraction

The RNA was extracted from HeLa cell pellets (1 x 10^6 cells) by resuspending in 500 µl RNA extraction buffer as described above and heating to 95 °C for 2 min. An equal volume of phenol:chloroform:isoamylalcohol (25:24:1) was added and samples shaken vigorously for 5 min before centrifugation at 13,000 rpm in a benchtop microfuge. The upper aqueous layer containing RNA was removed to a clean microfuge tube and a second phenol:chloroform extraction performed in the same way. To remove traces of phenol in the sample 500 µl chloroform:isoamylalcohol (24:1) was added and mixing and centrifugation repeated as previously. To the final supernatant NaOAc pH 5.3 was added to a final concentration of 150 mM along with 1 ml ethanol. Samples were cooled either at -80 °C for 30 min or overnight at -20 °C prior to centrifugation at 13,000 rpm in a microfuge to pellet the RNA. The supernatant was removed, the pellet air dried at 37 °C and thoroughly resuspended in an appropriate volume of Nanopure water.

A similar method was employed for RNA extraction during *in vitro* assays such as immunoprecipitations and glycerol gradients. In these cases a single phenol:chloroform:isoamylalcohol (25:24:1) step was used and the chloroform:isoamylalcohol step was omitted. 1 µg of tRNA was also included into each sample during ethanol precipitation to act as a carrier to aid precipitation of the RNA of interest.

2.4.1.2 TRI reagent

RNA was extracted from cell pellets of HEK293 and TC7 cells using TRI Reagent (Sigma) which is a guanidine thiocyanate and phenol based solution. Cell pellets were resuspended in 500 µl TRI reagent and incubated at room temperature for 5 min to allow cell lysis. After addition of 100 µl chloroform, samples were vortexed and allowed to stand at room temperature for 15 min before centrifugation at 12,000 x g for 15 min. The top aqueous layer containing RNA was then removed and to this, 250 µl isopropanol was added. Samples were mixed and allowed to stand at room temperature for 10 min before centrifugation at 12000 x g for 10 min. The supernatant was then removed to leave a white pellet. The pellet was washed in 75 % ethanol, air dried and resuspended in an appropriate volume of water shaking at 55 ^{o}C for 5 min. Corresponding protein samples were also obtained from the lower layers; DNA was

removed from the interface and pelleted by addition of 150 µl ethanol and centrifugation at 8000 x g for 5 min, proteins were then precipitated by addition of 750 µl isopropanol, incubation at room temperature for 10 min and centrifugation at 12000 x g for 10 min. The protein pellet was then washed three times in 0.3 M guanidine hydrochloride in 95 % ethanol, dried and resuspended as required.

2.4.1.3 Determination of RNA concentration

The concentration of RNA samples was determined using a Nanodrop (Thermo Scientific). The absorbance at 260 nm of 1 µl of the RNA sample was used to determine the concentration. The ratio of absorbance at 260 nm relative to 280 nm gave an indication of the purity of the sample.

2.4.2 Northern blotting

2.4.2.1 Acrylamide gel electrophoresis for small RNAs

RNA samples were mixed with 2 x RNA Loading Dye (80 % formamide, 10 mM EDTA pH 8.0, 1 mg/ml xylene cyanol FF, 1 mg/ml bromophenol blue) and denatured at 95 °C for 2 min. RNA was separated by electrophoresis on an 8-12 % acrylamide/ 7 M urea gel in 1 x TBE buffer, at 350 V for 1-2 h depending on the required separation of RNA species. RNAs were then transferred onto Hybond N membrane (Amersham Biosciences) using a Trans-Blot Cell (BIO-RAD) in 0.5 x TBE buffer at 60 V for 2 h. RNA was cross-linked to the membrane using a Stratalinker UV Crosslinker (Stratagene).

2.4.2.2 Agarose-glyoxal gel electrophoresis for pre-rRNAs

Agarose-glyoxal gels were used for the analysis of long pre-rRNA molecules. 3-6 µg of total RNA was mixed in a 1:5 ratio of RNA:Glyoxal Loading Buffer (61.2 % DMSO (v/v), 20.4 % glyoxal (v/v), 12.2 % 1 x BPTE buffer (28.7 mM Bis-Tris, 9.9 mM PIPES, 1 mM EDTA) (v/v), 4.8 % glycerol (v/v), and 0.02 mg/ml ethidium bromide) and heated to 55 °C for 1 h. The electrophoresis gel system was used and thoroughly cleaned with RNAse Zap (Ambion) to remove RNases prior to use. Samples were separated by electrophoresis on a 1.2 % agarose / 1 x BPTE (30 mM Bis-Tris free acid pH 7.0, 10 mM PIPES free acid, 1 mM EDTA) gel run in 1 x BPTE buffer at 165 V for 4 h. RNA in the gel could then be detected using UV light. The gel was washed in 75 mM NaOH for 20 min, twice in Tris-Salt buffer (0.5 mM Tris-HCl pH 6.4, 1.5 M NaCl)

for 15 min each and finally in 6 x SSC (1 M NaCl, 0.1 M $Na_3C_6H_5O_7$) for 20 min. RNA was then transferred to a Hybond N membrane (Amersham Biosciences) overnight using capillary action in 6 x SSC buffer and RNA was cross-linked to the membrane as described in section 2.4.2.1.

2.4.3 Radio-labelled probes and Northern Blot hybridisation

2.4.3.1 Random Prime labelling

Probes for the detection of sno/snRNAs, pre-rRNA 5'ETS fragments, and SRP RNA were produced by random prime labelling using a PCR product amplified from a plasmid template containing the full length RNA sequence, details are given in Table 2.6.

25-50 ng DNA template was mixed with nanopure water to a volume of 9 µl and incubated at 95 ^{o}C for 5 min to denature the DNA. 3 µl of random hexamer mix, (250 mM Tris pH 7.5, 50 mM $MgCl_2$, 5 mM DTT, 500 µM dATP, 500 µM dGTP, 500 µM TTP) was then added along with 2 µl of ^{32}P labelled dCTP and 1 µl of Klenow polymerase and the labelling reaction incubated at 37 ^{o}C for 30 min. The volume of the reaction was made up to 50 µl with water and it was passed through a G-50 spin column (GE Healthcare) by centrifugation at 2000 rpm to remove any unincorporated nucleotides. Probes were denatured by heating to 95 °C for 5 min and then added directly to 10 ml Hybridisation Buffer (25 mM $NaPO_4$ pH 6.5, 6 x SSC, 5 x Denhardts, 0.5 % SDS (w/v), 50 % deionised formamide, 100 µg/ml denatured salmon sperm DNA).

Table 2.6 Random-prime labelled probes

Probe Target	Vector	Amplification Primers	Sequence
U3	U3 pBS+SP6	U3 Forward	TAATACGACTCACTATAGGG
		U3 Reverse	ATTAACCCTCACTAAAGGGA
U1	U1 pUC18	U1 Forward	GGGGAAAGCGCGAACGCAG
		U1 Reverse	TACTTACCTGGCAGGGGAG
ETS1	5.8S pUC19	ETS1_fwd	GCTGACACGCTGTCCTCTGGCGA
		ETS1_rev	CGGACAACCCCGCGGAGACGAGA
ETS2	5.8S pUC19	ETS2_altup	CCGCCTTCGCTTCGCGGGTGCG
		ETS2_dw	GGAAGCGGAGGAGGGTCCTCTGCG
ETS3	5.8S pUC19	ETS3_fwd	TCGTGTCTGTGGCGGTGGGAT
		ETS3_rev	TTCGGAAGAGCGGGCCGGGAGAA
U14	U14 pGEMT easy	T7	TAATACGACTCACTATAGGG
		SP6	CATTTAGGTGACACTATAG
7SL (S domain	7SL pUC19	7SL_fwd	CTATGCCGATCGGGTGTCCGC
		7SL_rev	CAGCACGGGAGTTTTGACCTGC
MRP	MRP pGEMT easy	T7	TAATACGACTCACTATAGGG
		SP6	CATTTAGGTGACACTATAG

2.4.3.2 Transcription

To detect U8, U13 and U15 snoRNAs, probes were made by T7 RNA polymerase transcription using a radio-labelled nucleotide and a PCR template containing the full length snoRNA sequence (Table 2.7). Transcription reactions were carried out by combining 50 ng of PCR product with 1 x Transcription Buffer (40 mM Tris-HCl pH 7.9, 6 mM $MgCl_2$, 10 mM NaCl, 10 mM DTT, 2 mM Spermidine), 1 mM ATP, 1 mM GTP, 1 mM CTP, 0.1 mM UTP, 0.5 µl RNasin (Ambion), 1 µl T7 RNA Polymerase (NEB) and 2 µl α-^{32}P UTP (1.32 µM, 1.48 MBq, PerkinElmer) in a total volume of 10 µl. Reactions were incubated at 37 ^{o}C for 2 h before addition of 1 µl Turbo DNase (Ambion) for a further 30 min at 37 ^{o}C to degrade the template DNA. The reaction volume was increased to 50 µl and centrifuged serially through two G50 Spin Columns (GE Healthcare) to remove any unincorporated nucleotides. Probes were denatured and resuspended in Hybridisation Buffer as above.

Table 2.7 Probes made by transcription

Probe Target	Vector	Amplification Primers	Sequence
U8	U8 pBS+SP6	T7 Forward	TAATACGACTCACTATAGGG
		T3 Reverse	ATTAACCCTCACTAAAGGGA
U13	U13 pGem9Zf	T7 Forward	CATTTAGGTGACACTATAG
		SP6 Reverse	CATTTAGGTGACACTATAG
U15	U15a pGEMTeasy	T7 Forward	TAATACGACTCACTATAGGG
		SP6 Reverse	CATTTAGGTGACACTATAG

2.4.3.3 Kinase labelled oligonucleotide probes

Oligonucleotide probes were designed using the reverse complementary sequence of specific regions of pre-rRNAs and are shown in Table 2.8. Oligonucleotides, 19-40 bases long, were radiolabelled at the 5' end by the addition of a ^{32}P labelled phosphate group using T4 polynucleotide kinase (PNK). 10 pmol oligonucleotide was added to 1 x T4 PNK Buffer (Promega), 10 U T4 PNK (Promega) and 5 µl ^{32}P-ATP in a 10 µl reaction which was incubated at 37 ^{o}C for 1 h. The reaction volume was increased to 50 µl with water and the probe passed through a G-50 spin column (GE Healthcare). Oligonucleotide probes were denatured at 95 ^{o}C for 2 min and added to 10 ml of SES1 (0.5 M sodium phosphate pH 7.2, 7 % SDS (w/v), 1 mM EDTA) for hybridisation.

Table 2.8 Oligonucleotide probe sequences

Probe	Pre-rRNA	Sequence
18S	18S	GGGCGGTGGCTCGCCTCGCG
5520 (5'ITS1)	ITS1	CCTCGCCCTCCGGGCTCCGTTAATGATC
5633	ITS1	GAACGAACGGGCACGC
5687	ITS1	TCTCCCTCCCGAGTTCTCGGCTCT
5958	ITS1	AATCCGGCCGGCCCCGAAGA
6112	ITS1	CGCGCTAGGTACCTGGACGGC
6121 (ITS1)	ITS1	AGGGGTCTTTAAACCTCCGCGCCGGAACGCGCTAGGTAC
6248	ITS1	GAGTCCGCGGTGGAG
6318	ITS1	GCGACGGCCGCCGGGTAAAG
6396	ITS1	GACACCACCCCACAGG
6448	ITS1	GGTCGGAAGGTTTCACACCAC
6508	ITS1	GGTTGCCTCAGGCCG
6603	ITS1	AGGTCGATTTGGCGAG
6773	ITS2	ATTGATCGGCAAGCGAC
3'ITS1	ITS1	CGAGGTCGATTTGGCGAGGGC
5.8S	5.8S	CAATGTGTCCTGCAATTCAC
5'ITS2	ITS2	CCGGGGCGATTGATCGGCAAGCGAC
ITS	ITS2	GCGCGACGGCGGACGACACCGCGGCGTC
3'ITS2	ITS2	GCTCGGCGGACGGACGGACGGAATC
28S-2	28S	TGGTCCGTGTTTCAAGACGGGT
28S-3	28S	CAAGACCTCTAATCATTCGCTT

2.4.3.4 Hybridisation

Membranes to be treated with random prime or transcription labelled probes were pre-hybridised for 1 h at 42 °C in Hybridisation Buffer (25 mM $NaPO_4$ pH 6.5, 6 x SSC (0.9 M NaCl, 90 mM $Na_3C_6H_5O_7$ pH 7.0), 5 x Denhardts, 0.5 % SDS (w/v), 50 % deionised formamide, 100 µg/ml denatured salmon sperm DNA). Hybridisation was performed overnight at 42 °C. The membrane was washed twice for 5 min in 2 x SSC, 0.5 % SDS (w/v) at room temperature, following by two washes in 2 x SSC, 0.1 % SDS (w/v) and then a final wash in 2 x SSC, 0.1 % SDS (w/v) at 50 °C for 30 min.

Membranes to be probed with oligonucleotide probes were pre-hybridised for 30 min at 37 °C in SES1 (0.5 M sodium phosphate pH 7.2, 7 % SDS (w/v), 1 mM EDTA). Hybridisation was carried out overnight at 37 °C following which membranes were washed twice in 1 x SSC, 0.1 % SDS (w/v).

All membranes were dried and exposed to a phosphorimager screen. Signals were then detected using a Typhoon Scanner and analysed using ImageQuant software (GE Healthcare).

2.4.5 RNaseH cleavage of pre-rRNA

Total RNA was extracted from RNAi treated, metabolically labelled HeLa cells (see section 2.6.4). 4 µg RNA was mixed with 0.35 nmol of DNA primer

(GCTTATGACCCGCACTTACTCG) complementary to 18S rRNA. The primer was annealed to the RNA by heating to 95 ^{o}C for 3 min and cooling slowly to room temperature for 10 min. RNA with annealed primer was divided equally into two reactions, one to be treated with RNaseH and a control in which the enzyme is omitted. RNaseH Buffer (20 mM Tris-HCl pH 7.5, 20 mM KCl, 20 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM DTT, 0.8 U/ml RNasin) was added to a give a final volume of 20 µl and RNaseH added to one sample. RNaseH was prepared in the laboratory by Dr Rob van Nues and the amount of enzyme added to obtain optimal cleavage was empirically determined. Reactions were incubated at 37 ^{o}C for 30 min and then reactions terminated by addition of 80 µl of 0.3 M NaOAc, 2 mM EDTA and 100 µl phenol:chloroform. Samples were vortexed vigorously for 5 min and centrifuged at 13,000 rpm in a benchtop centrifuge. The upper aqueous layer was removed and 2.5 volumes of 100 % ethanol added. Samples were incubated at -80 ^{o}C for 1 h before the RNA was pelleted by centrifugation at 13,000 rpm in a microfuge. The supernatant was removed and the pellet dried. RNA pellets were resuspended directly in 10 µl Glyoxal Loading Buffer and analysed by agarose-glyoxal gel electrophoresis and Northern blotting as described above.

2.4.6 Mapping RNA 3' ends

RNA was extracted from either control HeLa cells or from HeLa cells in which exonuclease components had been depleted (see section 2.6.2). 2 µg RNA was DNase treated with 1 µl/2U Turbo DNase (Ambion) in 1 x Turbo Buffer (Ambion) at 37 ^{o}C for 1 h to remove any DNA contamination in the samples and then the enzyme was inactivated by denaturing at 80 ^{o}C for 5 min. An adaptor ddC oligonucleotide (12 µM) (5' AGATCTAGAGGATGGATATGGTGTTCAGGC 3') was ligated to all the RNA molecules in the samples using 1 µl T4 RNA Ligase (New England Biolabs) in 1 x T4 RNA Ligase Buffer (New England Biolabs) for 16 h at 4 ^{o}C, after which the enzyme was deactivated by incubation of samples at 65 ^{o}C for 10 min. A reverse transcription was then performed using a reverse primer that anneals the sequence of the adaptor oligo (pJETrev 5' GCCTGAACACCATATCCATCC 3') and SuperScript III Reverse Transcriptase (Invitrogen). 2 µl pJETrev primer was added to the reaction and heated to 95 ^{o}C for 10 min to anneal the primer to the template. Reactions were divided into two and the following steps carried out in the presence or absence of SuperScript reverse transcriptase to verify that products detected in later steps are derived specifically from the RNA template. 1 x 1st Strand Buffer (Invitrogen), 0.05 M DTT, 0.5 mM dNTP, 20 U Superasin (Ambion) and 100 U SuperScript III RT was added to the sample which was incubated at 50 ^{o}C for 30 min then 55 ^{o}C for a further 30 min to

generate cDNAs. PCR reactions were then carried out using GoTaq polymerase (Promega) and the pJETrev primer. Forward PCR primers depended on the specific RNA that was to be mapped. The preU3 primer overlaps 2 nucleotides into the precursor sequence, so will amplify only the U3 precursors and not mature U3.

PCR products were in the size range 61-250 bp and were analysed on an 8 % non-denaturing acrylamide gel (acrylamide:bisacrylamide 19:1 in 1 x TBE). The gel was stained by washing with SYBER Safe DNA stain (1:50,000) for 10 min before visualisation of DNA by UV light. Sections of the gel containing PCR fragments of interest were excised from the gel and immersed in 1 x TBE + 0.1 %SDS with agitation for 16 h to elute the DNA from the gel slices. Ethanol precipitation as previously described was then used to purify and concentrate the DNA. Pellets were thoroughly washed in cold 70 % ethanol to remove traces of EDTA from the TBE that would inhibit later ligation steps. One-third of the generated material was ligated into the pGEMTeasy vector (section 2.1.5) and colonies containing vectors with inserted fragments selected by blue:white screening using interruption of the β-galactosidase gene on selective plates containing XGal as a marker.

For sequencing of preU3 in control HeLa cells, DNA was extracted from 100 white colonies using the Spin Column Plasmid Miniprep Kit (NBS Biologicals) following the manufacturer's instructions. A sample of 20 clones was manually sequenced by chain termination sequencing using USB Sequense Version 2.0 DNA Polymerase. Plasmid DNA was first denatured and purified by treatment with 0.4 M NaOH for 20 min at 37 ^{o}C, neutralisation by addition of 0.5 M NaOAc pH 5.2 and ethanol precipitation. DNA was resuspended in 7 µl water and to this Sequenase Reaction Buffer (USB) was added to a 1 x concentration along with 1 pmol of 5' ^{32}P labelled T7 primer (5' TAATACGACTCACTATAGGG 3'). Reactions were incubated at 37 ^{o}C for 30 min and slowly cooled to room temperature to anneal the labelled primer. Primer extension using limiting amount of dNTPS (75 nM of dATP, dTTP, dCTP and dGTP), 6 mM DTT and 2 U Sequenase Polymerase (USB) was carried out for 4 min at room temperature. Termination Mixes were prepared to contain 80 µM non-terminating dNTP and 8 µM terminal ddNTP and 50 mM NaCl for each of the four nucleotides. The 2', 3' – dideoxynucleoside-5'-triphosphates lack the 3'-OH group necessary for DNA chain elongation. Tubes were set up containing 3.5 µl of one of the four Termination Mixes and the labelling reaction was divided equally between the tubes (2.5 µl each). Reactions were continued at 37 ^{o}C for 5 min during which DNA synthesis continues until a ddNTP is incorporated so the final reaction contains a mixed population of chains terminated at a specific base along the sequence. Reactions were stopped after 5 minutes by addition of RNA Loading Dye (80 % formamide, 10 mM EDTA pH 8.0,

1 mg/ml xylene cyanol FF,1 mg/ml bromophenol blue) and separated on a denaturing 6 % acrylamide/7 M urea sequencing gel. The gel was fixed in 50 % methanol, 10 % acetic acid for 10 min, dried and results visualised by X-ray film exposure overnight. Once the U3 precursor sequence had been determined, this method was then refined to track only the Ts in the sequence and the pattern used to extrapolate the preU3 sequence in the remaining 80 pGEMTeasy clones.

The analysis of the 3' end of U3 precursors in cells in which exosome components have been depleted by RNAi was carried out using the same methodology and pGEMTeasy clones were sequenced from positive colonies by GATC Biotech.

2.5 *In vitro* assays

2.5.1 Protein-Protein Interaction

2.5.1.1 Glutathione sepharose-GST precipitation

Equal amounts (approximately 50 ng) of each of the proteins of interest were mixed in a final volume of 200 µl of the buffer in which they were purified (Exosome or Purification Buffer). Samples were mixed at 4 ^{o}C on a rotary incubator for 1 h to allow formation of protein complexes. 40 µl of Glutathione Sepharose 4 Fast Flow Beads (GE Healthcare) (50 % in ethanol) were washed three times in the buffer pelleting in between washes by centrifugation at 3000 rpm in a benchtop centrifuge. Beads were added to each reaction and reactions were mixed for a further 2 h on a rotary incubator to allow GST tagged proteins and their complexes to bind. The beads were then pelletted by centrifuging at 3000 rpm for 1 min in a benchtop microcentrifuge. The beads were then washed three times in buffer as above and an additional wash step in which the beads were transferred to a clean microcentrifuge tube was included. Beads with bound proteins were then resuspended in 50 µl Protein Loading Buffer (74 mM Tris-HCl pH6.8, 1.25 mM EDTA, 20 % glycerol, 2.5 % SDS, 0.125 % bromophenol blue, 50 mM DTT) before analysis by SDS-PAGE and/or Western blotting.

2.5.1.2 Immunoprecipitation

A method similar to that described in section 2.5.1.1 was used for investigating protein-protein interactions between the thioredoxin tagged proteins, NOP56 and NOP58 and HIS- or GST-tagged exosome proteins. Protein A Sepharose CL-4B beads (GE Healthcare) were washed three times in Exosome Buffer. Anti-thioredoxin antibody was added (1:80 dilution) to the beads and mixed on a rotary incubator at 4 ^{o}C overnight. The wash steps were then repeated to remove any unbound antibody. Proteins were mixed and interactions determined as for pull-downs with Glutathione Sepharose 4 Fast Flow Beads.

2.5.2 Protein-RNA interactions

Regions of 18S rRNA and ITS1 that have been shown to bind pre-40S complex proteins by CRAC in yeast (Granneman et al, 2010) are predicted to bind human proteins *in vitro*. The sequence of 18S and ITS1 is highly conserved between mouse and human and these fragments were amplified by PCR using HotStar Taq polymerase using the primers shown in Table 2.9. Figure 2.2 shows where each of these primers binds on a secondary structural model of the 18S rRNA. For each substrate the forward primer in the PCR reaction contains a T7 promoter to enable the product to be radioactively labelled by *in vitro* transcription.

Table 2.9 Primers for amplification of 18S rRNA substrates

Primer	Direction	Sequence
3761_T7	Forward	GCCGAATTCTAATACGACTCACTATAGGGTCAGTTATGGTTCCTTTGGTC
4980_T7	Forward	GCCGAATTCTAATACGACTCACTATAGGGTGGTGGTGCATGGCCGTTC
4882_T7	Forward	GCCGAATTCTAATACGACTCACTATAGGGTGGAGCCTGCGGCTTAATTTG
5318_T7	Forward	GCCGAATTCTAATACGACTCACTATAGGGTCATAAGCTTGCGTTGATTAAGTCC
5357_T7	Forward	GCCGAATTCTAATACGACTCACTATAGGGACCGCCCGTCGCTACTACCGA
4012	Reverse	GCCGGATCCGCAGACGTTCGAATGGGTCG
5176	Reverse	GCCGGATCCCAGCCCCGGACATCTAAGGGC
5188	Reverse	GACGGATCCGTAGCGCGCGTGCAGCCCCG
5527	Reverse	GCCGGATCCTAATGATCCTTCCGCAGGTTC
5927	Reverse	ACACACAAGACGGGGAGAGCG

Transcription reactions were carried out by combining 50 ng of gel recovered PCR product with 1 x Transcription Buffer (40 mM Tris-HCl pH 7.9, 6 mM $MgCl_2$, 10 mM NaCl, 10 mM DTT, 2 mM Spermidine), 1 mM ATP, 1 mM GTP, 1 mM CTP, 0.1 mM UTP, 0.5 µl RNasin (Ambion), 1 µl T7 RNA Polymerase (kindly provided by N. Zenkin) and 2 µl α-^{32}P UTP (1.32 µM, 1.48 MBq, PerkinElmer) in a total volume of 10 µl. Reactions were incubated at 37 ^{o}C for 2 h before addition of 1 µl Turbo DNase (Ambion) for a further 30 min at 37 ^{o}C to degrade the template DNA. The reaction volume was increased to 50 µl and centrifuged serially through two G50 Spin Columns (GE Healthcare) to remove any unincorporated nucleotides. The activity of each RNA transcripts was determined using a scintillation counter and samples were diluted to 5000 cpm/µl. Approximately 50 ng of GST-tagged protein (or combination of GST- and His-tagged proteins) was diluted in Binding Buffer (20 mM HEPES pH 7.4, 150 mM KCl, 1.5 mM $MgCl_2$, 0.1 % Triton X-100 (v/v), 10 % glycerol (v/v) to a final reaction volume of 200 µl. When two proteins were mixed, reactions were incubated at 4 ^{o}C for

1 h on a rotary wheel to enable complex formation. 20 µl per reaction of Glutathione Sepharose 4 Fast Flow (GE Healthcare) bead slurry in ethanol was washed three times in Binding Buffer precipitating beads in between washes at 3000 rpm in a microfuge. Beads were then added to each reaction and proteins bound to the beads for 1 h at 4 ^{o}C for 1 h on a rotary incubator. 1 µl (5000 cpm) of each labelled RNA transcript was then added to the sample and reactions incubated at 4 ^{o}C for a further 30 min. Beads with bound protein and RNA were then precipitated by centrifugation at 3000 rpm in a microfuge and washed 4 x in Binding Buffer with the beads being transferred to a clean microfuge tube on the final wash step. RNA was eluted from the beads by addition of 160 µl of RNA Extraction Buffer (1 % SDS (w/v), 50 mM Tris pH 7.5, 50 mM NaCl, 0.5 mM EDTA). To this 200 µl phenol:chloroform:isoamylalcohol (25:24:1) was added and shaken vigorously for 3 min. Samples were then centrifuged at 13,000 rpm in a microfuge for 5 min. The upper aqueous layer containing the RNA was removed and 150 mM NaOAc, tRNA (50 µg/ml) and 600 µl ethanol added. Reactions were cooled at -20 ^{o}C for 2-16 h then centrifuged at 13,000 rpm in a benchtop centrifuge for 20 min to pellet RNA. The supernatant was removed and the pellets dried in a speed vac. Samples were resuspended in 10 µl RNA Loading Dye (40 % formamide, 0.5 mM EDTA, 50 µg/ml bromophenol blue, 50 µg/ml xylene cyanol). Samples were denatured at 95 ^{o}C for 2 min, cooled on ice and separated on an 8 % polyacrylamide, 7 M urea gel. Gels were fixed for 10 min in 50 % methanol, 10 % acetic acid and dried onto Whatman Paper (VWR International). Results were visualised by exposure to a phosphorimager screen and detected using a Typhoon imager and then analysed with ImageQuant software (GE HealthCare).

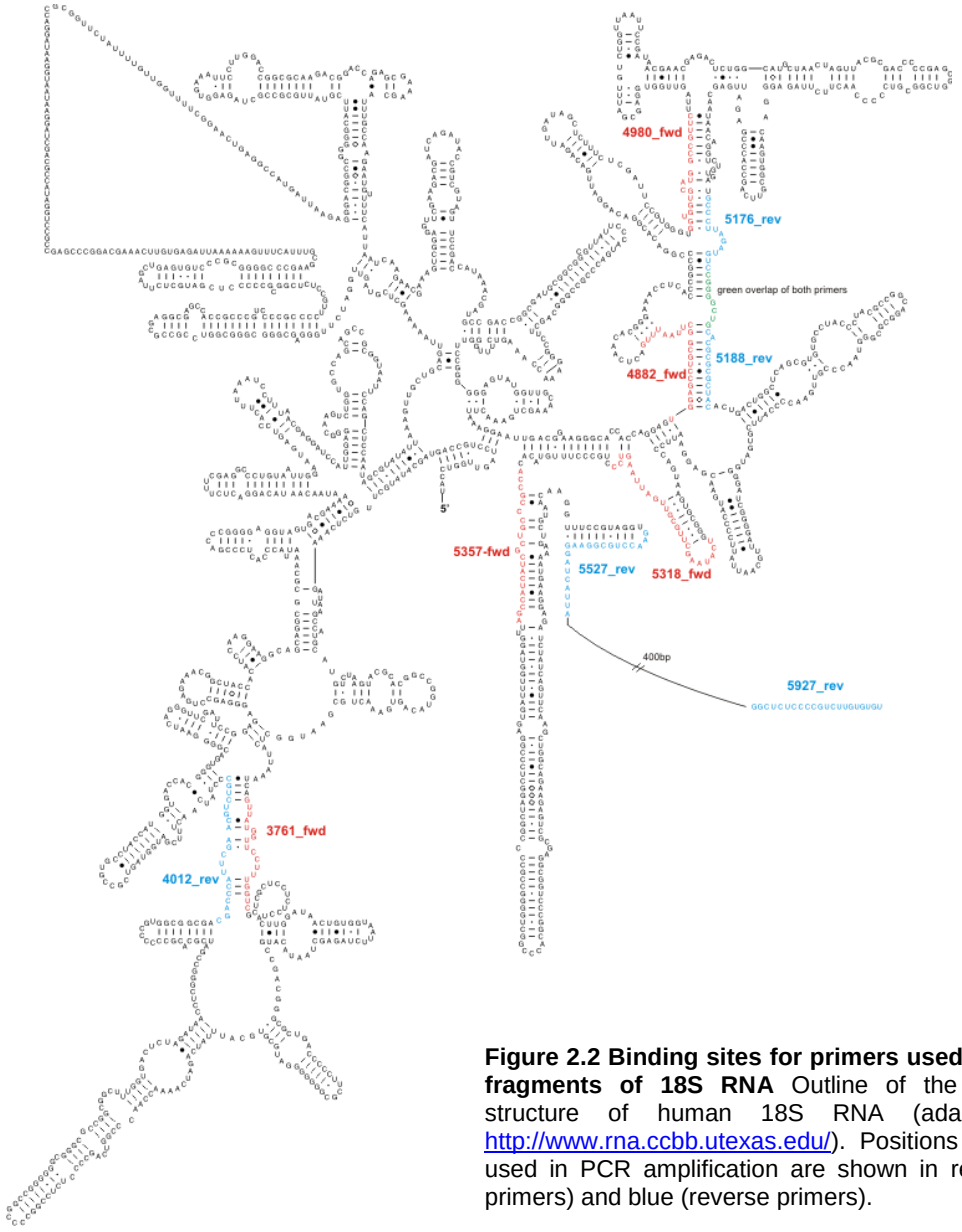


Figure 2.2 Binding sites for primers used to amplify fragments of 18S RNA Outline of the secondary structure of human 18S RNA (adapted from http://www.rna.ccbb.utexas.edu/). Positions of primers used in PCR amplification are shown in red (forward primers) and blue (reverse primers).

2.5.3 Kinase/phosphorylation assay

RIO2 is a protein kinase and an assay was developed to assess both its ability to autophosphorylate and to detect phosphorylated substrates. 10 ng GST-RIO2 or GST-RIO2kd was incubated either alone (autophosphorylation) or with an equivalent amount of potential substrate proteins in a 10µl reaction in Purification Buffer (20 mM Tris-HCl pH 8.0, 300 mM KCl, 5 mM $MgCl_2$, 0.1 % Tween 20 (v/v) and 10 % glycerol (v/v). A further reaction was set up in which HIS-NOB1 and GST-PNO1 had been pre-incubated at 4 ^{o}C on a rotary wheel for 1 h to allow interaction between these proteins before addition of GST-RIO2 or GST-RIO2kd. ^{32}P-ATP was added to a final activity of 1 µCi in both the presence and absence of unlabelled ATP. Unlabelled ATP was added to provide sufficient phosphate to drive the kinase reaction (autphosphorylation

reactions 0-1 mM, substrate phosphorylation 0.01 mM). Reactions were incubated at 37 ^{o}C for 30 min and stopped by addition of Protein Loading Dye (74 mM Tris-HCl pH 6.8, 1.25 mM EDTA, 20 % glycerol, 2.5 % SDS, 0.125 % bromophenol blue, 50 mM DTT) to denature the proteins. Samples were heated to 95 ^{o}C for 2 min and loaded onto a 10 % denaturing polyacrylamide gel and separated by elecrophoresis. The gel was fixed for 10 min in 50 % methanol, 10% acetic acid before being dried onto Whatman paper (VWR International). Phosphoimager screen detection revealed proteins which had been phosphorylated as they had become labelled by incorporation of γ-P^{32}-ATP.

2.5.4 Ribonuclease assays

Ribonuclease assays were used to assess the 3'-5' exonuclease activity of various proteins (human RRP6, C1D, MPP6 and ENP1 and yeast RRP6). Two different substrates were used: a poly(A) substrate 30 nucleotides long (kindly provided by Dr Claudia Schneider, University of Edinburgh) and a fragment of the U14 snoRNA which forms a stem-loop structure, 5' UCGCUGUGAUGAUUUAUUCUGAGCGAA 3' (Eurgoentec). 10 pmol of each substrate was radio-labelled at the 5' end using ^{32}P-ATP which generates a monophosphate end. Reactions were carried out as for the labelling of oligo-probes using T4 PNK. To purify substrates, RNA extraction buffer was added to a volume of 50 µl and an equal volume of phenol:chloroform;isoamylalcohol (25:24:1) added. Reactions were briefly vortexed and centrifuged at 13,000 rpm in microcentrifuge. The upper aqueous layer was removed and passed through a G50 spin column to removed unincorporated nucleotides. To this, NaCl was added to a final concentration of 250 mM along with 1 µl glycerol and 150 µl ethanol to precipitate RNA. Samples were incubated at -20 ^{o}C for > 2 h then centrifuged at 13,000 rpm to pellet the RNA which was then washed in 70 % ethanol. The pellet was resuspended in water to a concentration of 100 fmol/µl. Concentrations of proteins were determined by nanodrop as described above and 200 fmol was used per 10 µl reaction. Proteins were diluted in ExoMg Buffer (10 mM Tris–HCl pH 7.6, 75 mM NaCl, 2 mM DTT, 100 mg/ml BSA, 0.8 U/ml RNasin, 4.5 % glycerol, 0.05 % Triton X-100, 0.5 mM $MgCl_2$) and incubated at 37 ^{o}C for 10 min. The each 10 µl reaction 10 fmol of labelled RNA substrate was added. Reactions were incubated at 37 ^{o}C for the desired duration and reactions stopped by addition of one volume of RNA Loading Dye (80 % formamide, 10 mM EDTA pH 8.0, 1 mg/ml xylene cyanol FF, 1 mg/ml bromophenol blue). Samples were denatured either for 2 min at 95 ^{o}C for unstructured substrates or at 65 ^{o}C for stem loop RNAs before separation on a denaturing 12 % acrylamide, 8 M urea sequencing gel. Products were visualised using a phosphorimager.

2.6 Cell Culture and *in vivo* assays

2.6.1 Cell culture

Human HeLa SS6 (cervical carcinoma) cells were routinely cultured in Dulbecco's Modified Eagles Medium (DMEM) containing 4500 mg/L glucose, L-glutamine and sodium bicarbonate, without pyruvate (Sigma Aldrich) supplemented with 10 % (v/v) foetal bovine serum (FBS), 100 units/m penicillin and 100 µg/ml streptomycin. Cells were grown in a monolayer in a 5 % CO_2 humidified incubator at 37 ^{o}C. When at 80 % confluent, cells were released by treatment with 1 x trypsin EDTA (Sigma Aldrich) in sterile phosphate buffered saline (PBS, Sigma Aldrich) for 5 min at 37 ^{o}C, resuspended in DMEM and reseeding of an appropriate percentage.

A strain of HEK293 (human embryonic kidney) cells, Flp-In™ T-REX™-293 (Invitrogen) were cultured using the same methods as for HeLa cells. These HEK293 cells contain a stably integrated pFRT/*lac*Zeo vector derived from *S. cerevisiae* to provide a specific Flp Recombinase site. They also contain an integrated pcDNA6/TR vector allowing constitutive expression of the Tet-repressor under the control of the human CMV promoter and a sequence conferring blasticidin S resistance. This enabled these cells to be used for generating tetracycline-inducible stable cell lines. To maintain selection of cells containing plasmids for expression of FLAG-tagged proteins of interest, cells were cultured in the presence of 100 µg/ml hygromycin and treated with 10 µg/ml blasticidin at each third passage.

TC7 cells are a derivative of CaCo-2 (colon cancer) cells that are well characterised as a model for differentiation. TC7 cells were cultured in DMEM supplemented with non-essential amino acids (glycine 7.5 mg/L, L-alanine 8.9 mg/L, L-asparagine 1.32 mg/L, L-aspartic acid 1.33.0 mg/L, L-glutamic acid 1.47 mg/L, L-proline 1.15 mg/L, L-serine 1.05 mg/L and 0.1 mM glutamine). When passaging cells, they were grown to 80 % confluency, trypsinised using 4 ml 1 x trypsin in sterile PBS at 37 ^{o}C for 10-15 min. Cells often formed clumps which were allowed to settle prior to reseeding the flask. To achieve differentiation, cells were seeded into a flask as normal and allowed to grow for 22 days during which time the culture media was regularly changed. After this time, differentiated cells were harvested as normal.

HeLa, HEK293 and TC7 cells were all harvested by pelleting cells using a swing bucket centrifuge at 800 rpm for 5 min. Media were removed and pellets used directly or snap frozen in liquid nitrogen and stored at -80 ^{o}C. Stocks of all cell lines were prepared by resuspending approximately 5 x 10^6 cells in 1 ml Freezing Medium (DMEM, 20 % FBS, 10 % DMSO) and slowly cooling to -80 ^{o}C before storing in liquid

nitrogen. Cell stocks were revived by rapid thawing at 37 ^{o}C, washing the cells in 2 x 10 ml culture medium followed by normal growth in a T175 flask.

2.6.2 RNA interference (RNAi)

2.6.2.1 RNAi in HeLa cells

Specific target proteins in human cells were selectively depleted using siRNA targeting specific mRNA sequences (Table 2.10). A control siRNA that targeted firefly luciferase mRNA (GL2) was used in each experiment. This gene is not present in HeLa cells and has previously been demonstrated to have no effect on cell growth or RNA levels in HeLa SS6 cells (Elbashir et al, 2002). siRNAs purchased from Eurogentec need to be pre-annealed into a duplex prior to use. Sense and anti-sense strands were combined to a final concentration of 20 µM in 1 x Annealing Buffer (100 mM KOAc, 2 mM MgOAc, 30 mM HEPES-KOH (4-(2-Hydroxyethyl) piperazine-1-ethanesulfonic acid pH 7.4), heated to 90 °C for 1 min and incubated at 37 °C for 1 h before use. siRNAs purchased from MWG Eurofins are pre-annealed and were resuspended in 1 x Universal siRNA Buffer (MWG Eurofins) to a final concentration of 20 µM. SMARTPool siRNAs (Dharmacon) were also resuspended to 20 µM in 1 x Annealing Buffer prior to use.

Transfection of siRNA duplexes into HeLa cells was done by chemical reverse transfection using Lipofectamine™ RNAiMAX reagent (Invitrogen). Knock-downs were generally carried out in six-well plates but could be scaled up or down depending on the amount of material required; the following protocol is for a six-well plate. Complexes were prepared in the well by combining 150 pmol siRNA(s) with 500 µl Opti-MEM I Reduced serum medium with L-glutamine (Invitrogen) and addition of 4 µl RNAiMAX reagent for 15-20 min at room temperature. HeLa cells were harvested by trypsinising as previously described and resuspending in DMEM media containing no antibiotics. Cells were counted using a CASY cell counter (Innovatis AG) and diluted in media without antibiotics to a concentration of 7.5 x 10^4 cells/ml. 2 ml of cell solution was then added to each well containing siRNAs. Penicillin and Strepotmycin were added to each well 6 h after transfection to the concentrations normally used for growth of HeLa cells. Cells were grown for 60-72 hours in the presence of the siRNA and then harvested from the wells by trypsinisation as described above. Cells were split into pellets containing approx 7.5 x 10^5 cells and either snap frozen in liquid nitrogen for storage at -80 °C or disrupted directly for RNA extraction.

Table 2.11 siRNA duplexes used to deplete specific mRNAs in human cells

Target Gene	Target sequence	Source	Reference
GL2	CGTACGCGGAATACTTCGATT	Eurogentec	Elbashir et al., 2001
RRP46	GGAGCUCACAUUUAUGGAATT	Eurogentec	Watkins et al., 2004
	ACAAGGCCACACUCGAAGUTT		
	ACAUUCAAAGGAGCUCACATT		
RRP6	GAAGGCAGCUGAGCAAACATT	Eurogentec	N/A
	UGAGCAGAGUAAUGCAGUATT		
	AGAUGAAAGUUACGGAUAUTT		
C1D	GUUGGAUCCACUUGAACAATT	MWG	Schilders et al., 2007
MPP6	GAGCACUGGUACUUGGAUUTT	MWG	Schilders et al., 2005
	CAGUAGAGCUUGAUGUGUCTT		
	GAUAUGAGACCUUGGUGGGTT		
MTR4	GACAGCAGCUUGCCAAAUUTT	Eurogentec	N/A
	CUGGUGAUGUUACUAUUAATT		
	GGCUUUACAUUCCUAAAGATT		
DIS3	AGGUAGAGUUGUAGGAAUATT	MWG	Staals et al., 2010
XRN2	AAGAGUACAGAUGAUCAUGTT	MWG	West et al., 2004
	GGGAAGAAAUAUUGGCAAATT		
NOP58	CAAGCATGCAGCTTCTACCGTT	MWG	Watkins et al., 2004
NOP56	CAAUAUGAUCAUCCAGUCCATT	Eurogentec	Watkins et al., 2004
Fibrillarin	CAGTCGAGTTCTCCCACCGCTT	MWG	Watkins et al., 2004
	GATGTGTGTTGATACTGTTGCAC		
RPS19	GAUGGCGGCCGCAAACUGATT	MWG	
POP1	GAAUUUAACCGUAGACAAATT	MWG	Mattijssen et al., 2010
RPP38	GCUAUUGGACUUCAGAAGATT	MWG	Cohen A, et al., 2003
RPP40	CCUUAAACUUGGAUUCUAATT	MWG	Mattijssen et al., 2010
BOP1	AUGGCAUGGUGUACAAUGATT	MWG	Rohrmoser et al., 2007
RBM28	CGAGAAGGCUUGAUUCGUG	Dharmacon SMARTPOOL	N/A
	GAAAUUACCCUCUGAUGUG		
	ACUCAAAUAUGUCCGCAUU		
	CAAAGGCUGUAGAUGACAA		
RRP5	UGAAGGUUGUCGUAUUGAATT	MWG	Sweet et al., 2007
UTP24	UGAAAAGAAGAAUCCGUAA	Dharmacon SMARTPOOL	N/A
	CAAGGAAGUAUGCGACCAU		
	GAAGGAUCCCAGCGCAUUA		
	UCCAAGAUUUGAACGAUUA		
RCL1	ACGAAUGGUUCUCGAAUUG	Dharmacon SMARTPOOL	N/A
	CUACAGGGGUGGAUGCGUA		
	CUGGGAAGUCUCCGGGCUU		
	GAACAUGACUGUAGCGUCC		
BMS1	UCAAAGAGGAAGAAGAUUA	Dharmacon SMARTPOOL	N/A
	CAAAAGAAGCGCUGUUUAA		
	GGGAUUUAGAGGAGGUUAU		
	GGGCAUCAAACGACGGAAA		
RIO2	GGAUCUUGGAUAUGUUUAATT	MWG	Zemp et al., 2009
ENP1	CGAAAUCAGGCGUGAGCUU	MWG	Carron et al., 2010
	GAUGUUCAUGAACAAGAATT		
DIM1	GAUUCAACGCAGAAGGUAUUU	Dharmacon SMARTPOOL	N/A
	GAUGGUCUAGUAAGGAUAAUU		
	CUUAAGACCAACUGAUGUAUU		
	ACCAGAAGAUUUCAGCAUAUU		

Table 2.11 siRNA duplexes used to deplete specific mRNAs in human cells - continued

Target Gene	Target sequence	Source	Reference
NOB1	UCACCGAGGAUCAGCGCUUUU	Dharmacon SMARTPOOL	N/A
	GCAAAUUGGAUGGGCGUCUUU		
	ACAGAAAAGAUGACAGCGAUU		
	AGAAAGUGUCCGUGACCGUUU		
PNO1	GCUAACAGAUACACACCAUUU	Dharmacon SMARTPOOL	N/A
	GGACUUCAGAUACGCUUUAUU		
	AGAGAAUGUGACACGGACAUU		
	CUUAAAAGGGCUCCGAACAUU		
PRP43	UGAAAGUAGUUCUAUAUAUTT	MWG	N/A
	UAAGACUAAAGUAUUUAAUTT		

2.6.2.2 RNAi depletion and rescue in HEK293 cells

HEK293 stable cell lines for RNAi resistant expression of tagged proteins of interest were induced using tetracycline at the appropriate concentration (Table 2.11) to obtain 1:1 expression of tagged:endogenous proteins for 24 h. Transfection of siRNA duplexes to deplete the endogenous protein of interest in the induced HEK293 cells was carried out as for HeLa cells but using 5µl RNAiMAX reagent. Transfected cells continued to be induced with tetracycline and were harvested 48 h after siRNA treatment as described above.

2.6.3 Creation of HEK293 Stable Cell Lines

Genes of interest were cloned into a modified pcDNA5/FRT/TO vector containing 2x FLAG tags, a tetracycline-regulated hybrid human CMV/TetO2 promoter, and a hygromycin B resistance gene with a FRT site in the 5' coding region. When Flp-In T-Rex HEK293 cells are transfected with both this pcDNA5 vector and pOG44, a plasmid constitutively expressing Flp recombinase, homologous recombination between the FRT sites is enabled. This brings about integration of the tagged gene of interest and the hygromycin resistance gene in frame with their respective promoters. Cells in which the gene of interest has been integrated can then be selected for by culturing in the presence of hygromycin B (Figure 2.3).

To generate such stable cell lines, Flp In T-Rex HEK293 cells were grown in six-well plates to 60 % confluency. 9 µl/well of FuGene HD (Roche Applied Sciences) was mixed with 91 µl of Opti-MEM I (Invitrogen) and incubated at room temperature for 5 min. To this, 1.8 µg pOG44 and 0.6 µg pcDNA5 per well were added and allowed to

incubate for a further 15 min before the mixture was added to the cells in a drop wise manner.

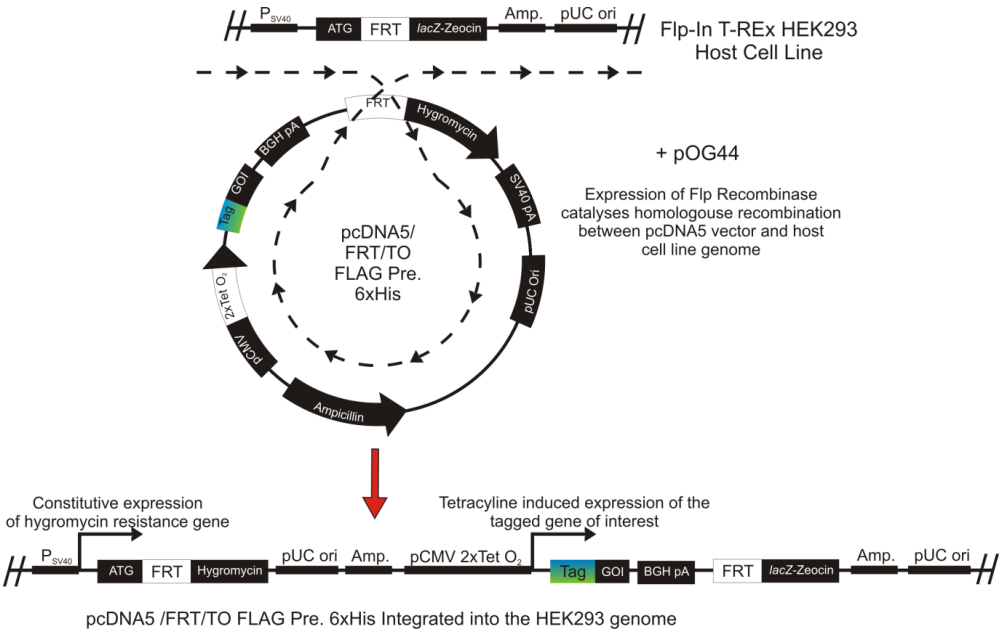


Figure 2.3 Schematic representation of pcDNA5 homologous recombination into Flp-In T-Rex HEK293 cells Homologous recombination (dotted line) is shown between the pcDNA5-FLAG vector and the Flp Recombination sites (FRT) in the modified HEK293 cell genome. This is aided by the Flp Recombinase encoded within pOG44. Shown are; SV40 promoter (P_{SV40}); start codon (ATG); Flp Recombination sites (FRT); hygromycin resistance gene (hygromycin); pUC origin (for propagation in *E. coli*); Ampicillin resistance gene (Amp.); CMV promoter (pCMV); Tetracycline operator sequences (2 x Tet O_2); 2 x FLAG Precission Protease 6 x His Tag (Tag); Gene of interest (GOI); BGH polyadenylation signal (BGH pA, required for efficient termination and polyadenylation of mRNA); and *lacZ* fusion with the zeocin resistance gene (*lacZ*-Zeocin). Based on a figure from Invitrogen.

Chemically transfected cells were incubated for 48 h under normal culture conditions. Selection of positive integrants was then initiated by addition of 100 µg/ml hygromycin B and 10 µg/ml blasticidin S. Colonies of selected cells were allowed to form for up to two weeks and resupended. Expression of tagged protein was induced by addition of 0-1 mg/ml tetracycline for at least 36 h. The concentrations of tetracycline required to induce 1:1 expression of endogenous and tagged-protein in each cell line used in this study this given in Table 2.11.

Table 2.11 Tetracycline concentrations used to induce stably transfected HEK293 cells

HEK293 cell line	Concentration of tetracycline (ng/ml)
RRP6	100
RRP6exo	100
DIM1	5
NOB1	1
PNO1	2
RIO2	1
RIO2kd	1

2.6.4 Metabolic RNA labelling

HeLa cells were cultured in six-well plates and RNAi depletion of factors of interest carried out as required. 48 h after transfection of siRNA duplexes, the culture media were changed for phosphate-free DMEM. This media was prepared according to a Sigma high-glucose DMEM (D5671) formulation without $NaHPO_3$ supplemented with 10 % dialysed FCS. Cells were depleted of phosphate for 1 h. Culture media was then exchanged for phosphate-free DMEM containing 15 µCi/ml ^{32}P labelled inorganic phosphate at 37 ^{0}C for 1 h. Labelled media was then removed and cells were washed with 2 ml normal DMEM supplemented with 10 % FCS. Wells were harvested as required (time points, 0, 15, 30, 60, 120 min) by trypsinisation. RNA was extracted from labelled cells using TRI-reagent and analysed by agarose-glyoxal gel electrophoresis or PAGE as appropriate. PAGE gels were dried and RNAs visualised using a phosphorimager. RNA separated on agarose-glyoxal gels was transferred to Hybond-N membranes by capillary action and RNAs visualised as for PAGE.

2.6.5 Treatment of HeLa cells with leptomycin B

Human cells were treated with 30 nM leptomycin B for 2 h (Muro et al, 2008; Rouquette et al, 2005) before being harvested for RNA extraction or being fixed for immunofluorescence.

2.6.6 Immunofluorescence

HEK293 cells were grown on 10 mm round coverslips in 24-well plates in 500 µl media to 80 % confluence in the presence of tetracycline if required. Cover-slips were washed 3 x in PBS and fixed using 4 % paraformaldehyde (w/v) in PBS pH 7.4 for 20 min. Cover-slips were then washed 3 x in PBS, incubated in PBS, 0.1 % Triton-X100 for 15 min at room temperature, washed 3 x in PBS and blocked for 1 h at room

temperature in PBS, 10 % FBS (v/v), 0.1 % Triton-X100. Primary antibodies (Table 2.12) were then applied using an appropriate dilution of antibody in 50 µl of 10 %FBS (v/v), PBS pH7.4 for 1 h at room temperature. Cover-slips were then washed 3 x briefly and 3 x 10 minutes in PBS before applying 50 µl of appropriate secondary antibody in 10 % FBS (v/v), PBS pH7.4 for 1 h at room temperature. Cover-slips were again washed 3 x briefly and 3 x 10 minutes in PBS with the addition of DAPI (0.1 µg/ml, Sigma Aldrich) in the penultimate wash. Finally, coverslips were briefly immersed in water and ethanol, dried and mounted, inverted onto glass slides using 3.5 µl Moviol. Cells were then examined using a Zeiss Axiovert 200M inverted microscope with a Plan-Apochromat, 100 x / 1.40 oil, ∞ / 0.17 objective (Zeiss). The Zeiss filter sets used were: 02 (DAPI); 20 (Cy3); 26 (Cy5).

Table 2.12 Antibodies used in immunofluorescence

Antibody	Raised in	Source	Dilution
FLAG (polyclonal)	Rabbit	Sigma F7425	1:500
Fibrillarin 72B9	Mouse	Michael Pollard/Ger Pruijn	1:500
Mouse IgG Alexa Fluor 555 conjugate	Donkey	Invitrogen A31570	1:500
Rabbit IgG Alexa Fluor 647 conjugate	Donkey	Invitrogen A31573	1:500

2.6.7 Preparation of whole cell extract

Human cells were cultured as described above, harvested and snap frozen in pellets containing approximately 5 x 10^6 cells. To prepare extracts for gradient analysis or immunoprecipitation a pellet was thawed on ice and resuspended in 500 µl Gradient Buffer (10 % glycerol, 20 mM HEPES-NaOH pH7.9, 150 mM KCl, 0.5 mM EDTA, 1 mM DTT) and disrupted by sonication using a Bandelin Sonopuls HD2070 ultrasonic homogeniser with a 2mm MS72 titanium microtip for 15 s at minimum power. Triton-X100 (0.2 % (v/v)) and 1.5 mM $MgCl_2$ (in the case of immunoprecipitations) were added and insoluble material pelleted by centrifugation at 13,000 rpm in a microfuge at 4 ^oC for 10 min. The cleared supernatant could then be used as required.

2.6.8 Immunoprecipitation

10 µl of Protein G Sepharose beads (GE Healthcare) per immunoprecipitation (IP) were washed in 3 x 1 ml PBS+ 0.1 % Triton-X100 with centrifugation at 3,000 rpm in a microfuge for 1 min at 4 ^oC between steps. After the final wash step, beads were resuspended in 1 ml IP buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 3 mM $MgCl_2$, 0.5 mM DTT, 10 % glycerol (v/v), 0.1 % Triton-X100 (v/v)) with 5 µl of mouse α-FLAG

antibody per IP reaction or no antibody as a negative control, and incubated on a rotary incubator overnight at 4 ^{o}C. Sepharose beads were then washed 3 x in IP Buffer as described above and aliquoted per IP reaction. Cell extract from either HEK293 or HeLa cells was prepared as described in Section 2.6.6 and this was added to the prepared sepharose beads which were rotated for 2 h at 4 ^{o}C. 10 % of the extract was retained for downstream RNA and protein extraction to give a relative measure of input. Immunoprecipitation reactions were washed 3 x in IP buffer as previously, resuspended in 1 ml IP buffer, transferred to a new microcentrifuge tube and pelleted at 3,000 rpm for 1 min. RNA and protein were then extracted from samples as described in Section 2.4.1.1 for analysis by Northern and Western blotting, respectively.

Chapter Three

The roles of human exonucleases in RNA processing and degradation in the nucleus

3.1 Introduction

Exoribonucleases play major roles in both prokaryotic and eukaryotic cells in the processing and degradation of a variety of RNAs. Cells produce large quantities of both coding (mRNAs) and non-coding RNAs (rRNAs, snoRNAs and snRNAs) (Warner, 1999). The majority of eukaryotic transcripts are produced as precursors and undergo controlled processing to form mature, functional RNAs. RNA degradation is also important both for maintaining steady-state RNA levels within the cells and as part of quality control mechanisms that target aberrant RNAs for degradation.

RNA can be processed and degraded from either the 5' or 3' end. XRN2 (Rat1 in yeast) and XRN1 provide the majority of 5'-3' exonuclease activity in the nucleus and cytoplasm, respectively, but other 5'-3' exonucleases such as NOL12 (Rrp17 in yeast) have also been identified (Johnson, 2001; Long & McNally, 2003; Oeffinger et al, 2009). The exosome provides the majority of the 3'-5' exonuclease activity in the cell although other 3'-5' exonucleases, such as the REX proteins also play important roles (Allmang et al, 1999b; Mitchell et al, 1997).

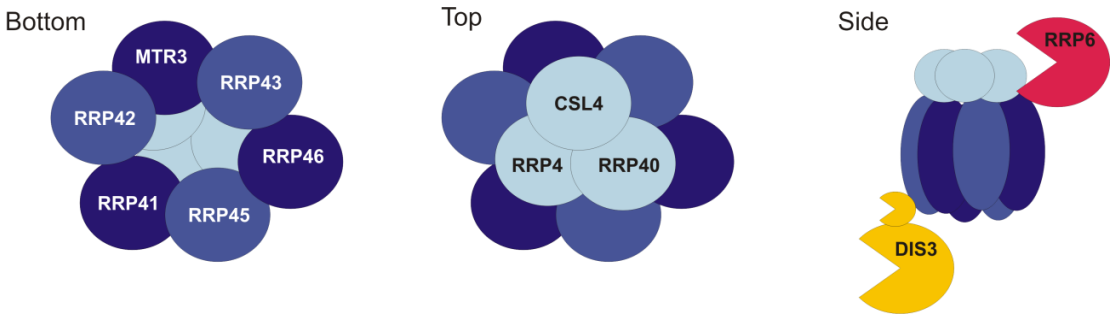


Figure 3.1 Structure of the eukaryotic exosome Schematic representations of the core eukaryotic exosome and active subunits shown for the top, bottom and side. Dark blue and mid-blue ovals represent PH domain proteins homologous to archaeal RRP41 and RRP42 respectively. Pale blue circles correspond to the trimeric cap formed by RNA-binding domain proteins which containing KH and S1 domains which stabilises the interactions of the hexameric ring. RNase D-like RRP6, shown in red, is probably located in the proximity of the trimeric cap while the exo-/endonuclease, DIS3 (yellow) is anchored to the lower face of the core exosome via its PIN domain (small yellow shape).

The exosome is an evolutionarily conserved multi-protein complex which in eukaryotes is formed around an inactive core that consists of nine individual subunits arranged as a hexamaeric ring of PH domain proteins, RRP41, RRP45, RRP42, MTR3, RRP43 and RRP46, stabilised by a trimeric cap composed of KH/S1 domain containing RNA binding proteins, RRP4, RRP40 and CLS4 (Figure 3.1) (Liu et al, 2006). In yeast, the exosome is constitutively associated with Rrp44, a protein which possesses both exonuclease and endonuclease activity (Lebreton et al, 2008; Mitchell et al, 1997; Schneider et al, 2007; Schneider et al, 2009). A second exonuclease, Rrp6, is also a component of the nuclear exosome (Mitchell et al, 2003). In higher eukaryotes, two exosome-associated Rrp44 homologues have been identified, DIS3 (in the nucleus) and DIS3L (in the cytoplasm) although it is not clear if these are as closely associated with the exosome as their yeast counterpart is (Staals et al, 2010; Tomecki et al, 2010b). The human exosome is also associated with the RNase D-like exonuclease, RRP6, predominantly in the nucleolus but also in the nucleus and possibly in the cytoplasm (Tomecki et al, 2010b). Rrp44 is anchored to the lower face (opposite the cap) of the core exosome through its PIN domain and it is proposed that unstructured RNA substrates recognised by the RNA binding proteins in the cap pass through the central channel of the hexameric ring and are degraded by Rrp44 as they exit (Bonneau et al, 2009; Lorentzen et al, 2008; Schneider et al, 2007; Wang et al, 2007). Structured RNA substrates are, however, thought to be directly cleaved by the endonuclease activity of the PIN domain of Rrp44 and are degraded without passing though the core exosome. It is less certain how Rrp6 interacts with the core exosome and different models have been proposed that place Rrp6 either close to the trimeric cap or on the opposite face (Cristodero et al, 2008; Lehner & Sanderson, 2004). The activity of the exosome in both yeast and humans is regulated by a range of different cofactors and in the nucleus, RRP6 interacts with two cofactors, C1D (Rrp47 in yeast) and MPP6, which are thought to recruit the exosome to particular substrates (Milligan et al, 2008; Mitchell et al, 2003; Schilders et al, 2005; Schilders et al, 2007). Another mechanism of exosome recruitment in the nucleus is through the polyadenylation of target substrates by the TRAMP complex which makes them more favourable substrates (Houseley et al, 2006; Kadaba et al, 2004; Kadaba et al, 2006; LaCava et al, 2005; Vanacova et al, 2005). The TRAMP helicase, MTR4 is, however, particularly associated with the nuclear exosome and acts as a TRAMP-independent exosome cofactor in functions such as processing of 5.8S rRNA (Lubas et al, 2011; Schilders et al, 2007).

Many substrates of the yeast exosome in both the nucleus and cytoplasm have been identified and some of the nuclear substrates are shown in Figure 3.2 (Houseley

et al, 2006; Kiss & Andrulis, 2011; Lykke-Andersen et al, 2009; Schmid & Jensen, 2008a; Tomecki et al, 2010a; Vanacova & Stefl, 2007). The yeast exosome is involved in processing of precursors of sn/snoRNAs and was initially characterised through its role in the formation of the mature 3' end of 5.8S rRNA (Allmang et al, 1999a; Mitchell et al, 1997; Mitchell et al, 1996). Degradation functions of the exosome include turnover of fragments of pre-rRNAs released during processing and unstable transcripts such as CUTs and PROMPTs in the nucleus (Chekanova et al, 2007; Milligan et al, 2008; Neil et al, 2009; Preker et al, 2008). Exosome functions in RNA surveillance came to light through its participation in the degradation of defective non-coding RNAs (sn/sno/t/rRNAs) and incorrectly spliced mRNAs in the nucleus. Cofactors of the exosome which also participate in specific exosome functions have been identified. It is not yet clear how these exonucleases determine whether a substrate should be targeted for degradation or whether only processing to generate a specific mature end is required.

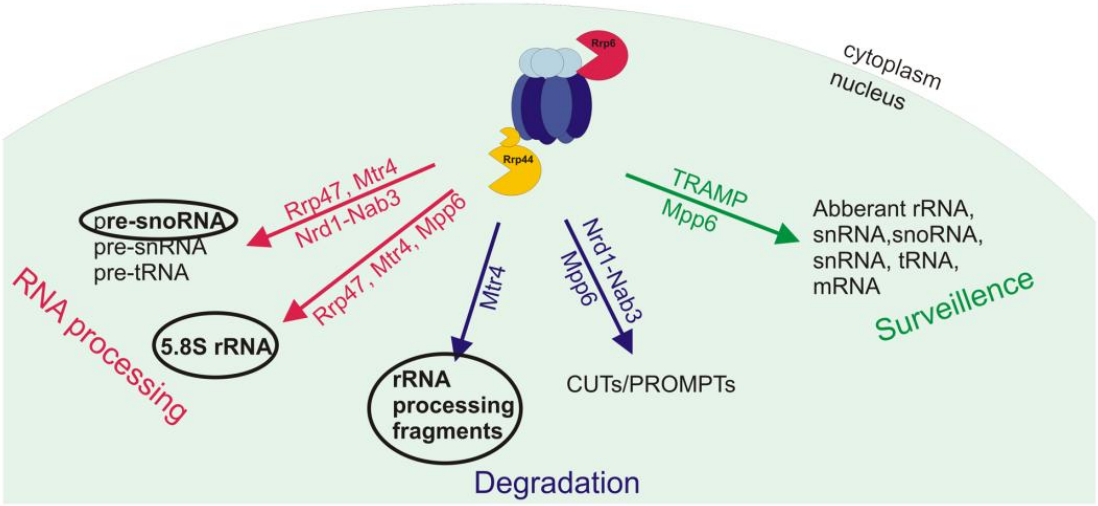


Figure 3.2 Diverse functions of the exosome in the nucleus Schematic representation of the core yeast exosome and some of its nuclear substrates. The cofactors linked to each exosome function are given along the relevant arrow. The role of the human exosome in processing and degradation of the substrates highlighted with black circles will be investigated in this study.

Much less is known about the composition and roles of the human exosome than the yeast exosome. The only example of the human exosome stably associated with a substrate is with U3 and U8 pre-snoRNAs making snoRNA biogenesis an ideal system for investigating human exosome processing in the nucleus (Watkins et al, 2004; Watkins et al, 2007). In yeast, the 3' ends of snoRNAs are processed by the exosome (Allmang et al, 1999a; Mitchell et al, 2003). snoRNPs are required for ribosome biogenesis as they not only carry out important post-transcriptional modifications but because some, such as U3, U14 and U8 are also essential for

specific cleavage steps in pre-rRNA processing (Henras et al, 2008; Watkins & Bohnsack, 2011). Many snoRNAs in vertebrates are encoded within introns of protein coding genes and are released from the pre-mRNA either by splicing of the intron lariat or by endonucleolytic cleavages in regions flanking the snoRNA (Dieci et al, 2009). This is followed by 5' and 3' end processing to produce the mature snoRNA (Allmang et al, 1999a; Petfalski et al, 1998).

Other snoRNAs, such as U3 and U8, are independently transcribed by RNA polymerase II and precursors contain a 5' m^7G cap and a 3' extension. The mature snoRNA is produced by hypermethylation of the cap and removal of the 3' extension by exonuclease processing (Filipowicz & Pogacic, 2002; Mouaikel et al, 2002; Weinstein & Steitz, 1999). In yeast, most snoRNAs are independently transcribed and contain both 5' and 3' extended sequences. It has been shown that the 5' extensions are removed by Rat1 and Xrn1, whereas the 3' extensions are first cleaved by the endonuclease, Rnt1 (Chanfreau et al, 1998b; Kufel et al, 2003) and are then processed by the exosome (Allmang et al, 1999a). snoRNA processing occurs concurrently with the assembly of proteins to form the snoRNP complex. Mature box C/D snoRNPs contain four core proteins which are assembled hierarchically; 15.5K first binds a conserved stem in the box C/D motif and this in turn recruits NOP56, NOP58 and fibrillarin (Watkins et al, 2002; Watkins et al, 2000). In human cells, biogenesis of the U3 snoRNP complex is achieved by a large multi-protein pre-snoRNP complex (Watkins et al, 2004). Proteins linked to snoRNP assembly include biogenesis factors NUFIP, BCD1, NOP17 and the ATPases, TIP48 and TIP49 (McKeegan et al, 2007).

The yeast exosome not only participates in the 3' end formation of 5.8S, but is also involved in other stages of ribosome biogenesis. Three of the four ribosomal RNAs, 18S, 5.8S and 28S are co-transcribed as a single precursor by RNA polymerase I. This precursor includes external transcribed spacer regions (5'ETS and 3'ETS) and two internal transcribed regions (ITS1 and ITS2) which are removed by an ordered series of endonucleolytic cleavages and exonucleolytic processing to release the mature rRNAs (Henras et al, 2008). The 5' external spacer (5'ETS) is removed in a step-wise manner by endonucleolytic cleavages at A_0 and A_1 in yeast and in higher eukaryotes, at an additional cleavage site, A'. In mouse cells, both XRN2 and the exosome are required for turnover of specific fragments released by these cleavages (Kent et al, 2009; Wang & Pestov, 2011). In yeast, both 5'-3' exonucleases including Rat1 (XRN2) and the 3'-5' exonuclease complex, the exosome, are involved in the formation of the mature ends of both 5.8S and 28S rRNAs (Mitchell et al, 1996). The role of the exosome and some of its cofactors in the formation of the 3' end of 5.8S is conserved in human cells (Schilders et al, 2005; Schilders et al, 2007). In both plant

and mouse cells, a role for XRN2 in pre-rRNA processing which does not appear to be directly related to its exonuclease activity has also been demonstrated (Wang & Pestov, 2011; Zakrzewska-Placzek et al, 2010). Further, in yeast the exosome plays an important role in quality control pathways which ensure the production of functional ribosomes by degradation of defective pre-rRNAs (Houseley et al, 2006). These aberrant precursors are polyadenylated by the TRAMP complex targeting them for degradation by the exosome. Polyadenylated pre-rRNAs have been identified in higher eukaryotes implying that this mechanism may be conserved.

The majority of our knowledge about the many functions of exonucleases, particularly the exosome, is derived from studies carried out in *S. cerevisiae*. Although evidence suggests that some of these functions are conserved in human cells, it is likely that exonucleases also have additional or different roles in RNA processing in higher eukaryotes. We therefore aimed to characterise the interactions and activity of the human nuclear exosome *in vitro* and study its function on two substrates with which the human exosome has been shown to be associated, pre-snoRNAs and ribosomal RNAs.

3.2 Results

3.2.1 Purification and interactions of recombinant exonuclease proteins and their cofactors

A crystal structure of the core human exosome composed of nine subunits has been published (schematic in Figure 3.1) but it was not clear how the active subunits and cofactors interacted with this core (Liu et al, 2006). In yeast, Rrp44 provides the majority of the activity of the exosome but DIS3, the nuclear homologue in human cells, is less stably associated with core exosome (Staals et al, 2010; Tomecki et al, 2010b). This suggests that human RRP6 and its cofactors may play a more prevalent role in processing of exosome substrates in human cells. It was therefore decided to investigate how human RRP6 and its cofactors, C1D and MPP6, interact with proteins of the nine-membered core exosome.

His-tagged forms of the nine core exosome proteins; RRP4, RRP40, CSL4, RRP43, RRP46, RRP41, RRP45, RRP42 and MTR3 were over-expressed in *E.coli* and purified by nickel-affinity. GST-tagged RRP6, C1D, MPP6 and GST-alone were similarly over-expressed in *E.coli* and purified using their GST-tags (Figure 3.3A).

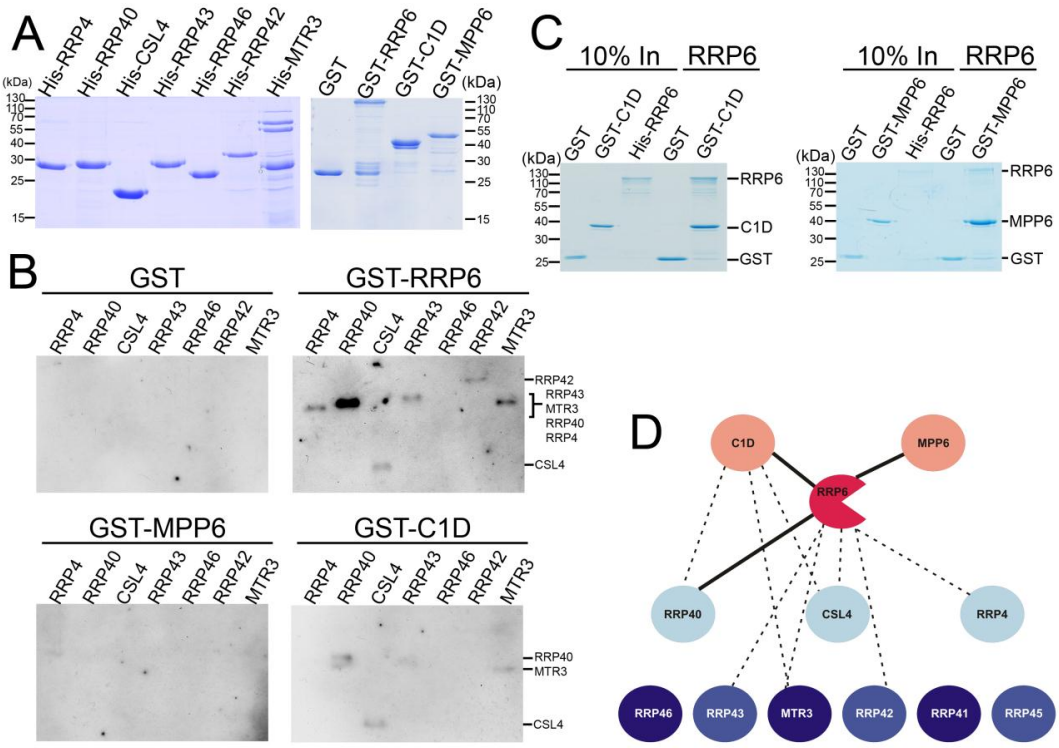


Figure 3.3 Protein-protein interactions between the core exosome, RRP6 and its cofactors **A)** Proteins of interest were expressed in *E. coli*, purified by either their His- or GST-tags, separated by SDS-PAGE and visualised by Coomassie staining. Positions of marker proteins are indicated to the left of the panel. **B)** Recombinant GST or GST-tagged RRP6, C1D7 or MPP6 was bound to glutathione sepharose and incubated with each of the His-tagged core exosome proteins. Co-purified complexes were separated by SDS-PAGE and analysed by Western blotting using an α-His antibody. A) represents 10X input of each protein. Co-purified proteins are indicated to the right of relevant panels **C)** Reactions were carried out as in B) using glutathione sepharose bound to either GST-alone, GST-C1D or GST-MPP6 and untagged-RRP6 was added as indicated above the panel. In each panel, the first three lanes show 10% input (10% In). Co-purified complexes were separated by SDS-PAGE and visualised by Coomassie staining. **D)** Schematic diagram of protein-protein interactions between the core human exosome, RRP6 and it's cofactors. PH domain proteins which form the hexameric ring are shown in dark and mid-blue. RNA binding (KH/S1) domain proteins of the cap are shown in pale blue. RRP6 and its cofactors are shown in red and pink. Strong interactions are indicated by thick black lines and weak interactions are indicated with dashed lines.

His-RRP45 was very unstable and did not remain in solution at usable concentrations so this protein was excluded from subsequent interaction studies. In preliminary experiments, His-RRP41 co-purified with a large number of other proteins (exosome cofactors and snoRNP proteins). The data generated were not reproducible and the large number of potential interactions observed suggested that these interactions were non-specific. For this reason RRP41 was also not included in further protein-protein interaction studies.

Recombinant His-tagged core exosome proteins were incubated with either GST-alone, GST-RRP6, GST-C1D or GST-MPP6 and any complexes formed were purified using glutathione sepharose. Bound proteins were then separated by SDS-PAGE, followed by Western blotting using an α-His antibody. This detected core exosome proteins that had been co-purified with the sepharose bound GST, GST-

RRP6, GST-C1D or GST-MPP6 and are therefore able to interact with the GST-tagged proteins (Figure 3.3B). None of the core exosome proteins interacted with the GST-tag alone showing that interactions observed with exosome proteins were specific to the coding regions of the proteins. RRP6 was able to co-purify RRP40 efficiently suggesting a strong interaction, while weaker interactions were also observed with RRP4, CSL4, MTR3, RRP43 and RRP42. As RRP6 interacts with all the KH/S1 domain proteins, this suggests that RRP6 is associated with the "cap" face of the core exosome. Weak protein-protein interactions were also observed between C1D and RRP40, CSL4 and MTR3. As C1D is a cofactor of RRP6 it is logical that interactions are seen with RRP6-interacting proteins and suggests C1D also interacts with the cap-side of the exosome. All the interactions of C1D are weaker than those observed for RRP6 implying that C1D is less closely associated with the core exosome than RRP6 is. No protein-protein interactions were observed between the core exosome and MPP6.

It was next decided to determine if human RRP6 interacts with C1D and MPP6 as the yeast proteins do. GST-RRP6 was expressed from the pGEX6P1 plasmid, which encodes a PreScission protease cleavage site between the protein coding region and the GST-tag. Purified GST-RRP6 was treated with PreScission protease to produce untagged protein. GST-alone and GST-C1D were incubated with untagged RRP6 and complexes formed were co-purified using glutathione sepharose. Proteins were then analysed by SDS-PAGE and proteins visualised by Coomassie staining (Figure 3.3C, left panel). RRP6 did not co-purify with the GST-alone showing no interaction had taken place, whereas a robust interaction was detected between RRP6 and C1D. 10% protein inputs were also separated by SDS-PAGE and by comparison >10% of RRP6 is co-precipitated by C1D. A similar experiment was carried out using GST-MPP6 and an equally robust interaction with RRP6 was observed (Figure 3.3C, right panel). As MPP6 was not observed to interact the with core exosome proteins, this suggests that RRP6 mediates association of these proteins with the core exosome *in vivo*. Based on these data, a protein-protein interaction network has been outlined (Figure 3.3D).

3.2.2 Ribonuclease activity of the exosome

In yeast, RRP6 has been characterised as having distributive 3'-5' exonuclease activity but, at the time of carrying out these experiments, the exonuclease activity of the human protein had not been investigated. Characterising the *in vitro* activity of

RRP6 is important as this sheds light on which substrates the protein may be able to process *in vivo* and how this processing might take place.

3.2.2.1 Recombinant human RRP6 has distributive 3'-5' exonuclease activity *in vitro*

In vitro exonuclease assays were carried out using purified recombinant human RRP6 (hRRP6) and *S. cerevisae* RRP6 (yRRP6) proteins (Figure 3.4A). RNase D-like nucleases, such as RRP6, contain conserved DEDD motifs in their catalytic sites. These four key amino acid residues are essential for the exonuclease activity of the proteins (Januszyk et al, 2011). A sequence alignment of yeast and mammalian RRP6 identified the human RRP6 DEDD motif and the first amino acid in this as aspartate 313. For the expression of a catalytically inactive enzyme, hRRP6exo, a point mutation was introduced into the plasmid from which the protein was expressed so that D313 would be substituted by alanine. Time course nuclease assays using a 5'-end labelled poly (A) RNA substrate were performed, with samples taken at various time points over two hours. Both human and yeast RRP6 exhibited 3'-5' exonuclease activity under these conditions (Figure 3.4B). In contrast, the hRRP6exo protein showed no nuclease activity confirming that the activity observed for hRRP6 and yRRP6 can be attributed to the protein itself and that D313 is a key amino acid for the catalytic activity of hRRP6 (Figure 3.4B). When hRRP6 was included, the full length labelled RNA was totally converted to shorter products within 10 min. Accumulation of progressively shorter fragments increased during the time course and after 120 min all the labelled RNA had been converted to fragments of 10 nt or less. The total amount of labelled RNA present after 120 minutes was significantly lower than the input suggesting that much of the RNA had been completely degraded to single nucleotides. Similarly, yRRP6 converted the full length RNA substrate to short products within the first 10 min of the assay. The range of fragments detected at each point during the time course was slightly longer than those detected for hRRP6 by approximately 5 nt. After 120 min in the presence of yRRP6 the labelled RNA had all been converted to products 10-17 nt in length. Taken together, the data suggest that the human recombinant protein purified for this experiment had slightly higher 3'-5' exonuclease activity than the yeast protein.

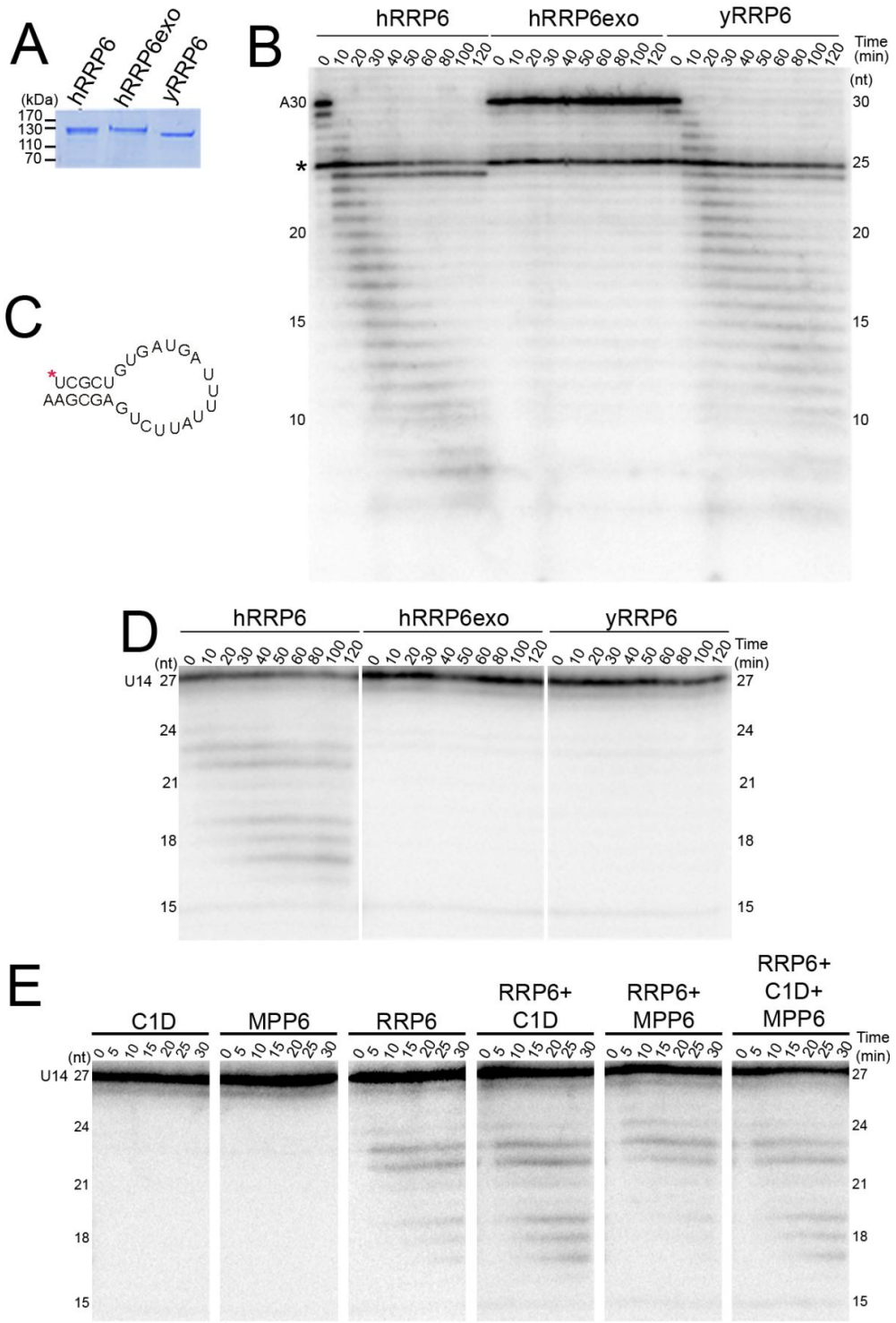


Figure 3.4 ***In vitro*** **exonuclease activity of human RRP6 A)** Human RRP6 (hRRP6), human RRP6 D313A (hRRP6exo) and yeast RRP6 (yRRP6) were expressed in *E. coli* and purified by their GST-tags. Purified proteins were then separated by SDS-PAGE and visualised by Coomassie staining. The positions of the marker proteins are indicated to the left of the panel. **B)** 200 fmol of recombinant protein (hRRP6, RRP6exo or yRRP6) and 10 fmol of 5' end labelled 30-mer-polyadenosine (A30) RNA substrate (shown in **(C)**) were incubated for 120 min at 30 ^{o}C with samples of equal volume being taken at the time points shown above the panel. Reactions were separated on a 12 % polyacrylamide/8 M urea gel and visualised using a phosphorimager. The asterisk marks a labelled contaminant in the substrate which should be disregarded. **D)** Schematic outline of the U14 stem loop substrate. A red asterisk marks the labelled 5' end. **E)** Exonuclease assays using the 5' end U14 RNA substrate and hRRP6, hRRP6exo and yRRP6 were carried out and results visualised as in A). **F)** Exonuclease assays using human RRP6 were carried out in the presence and absence of the exosome cofactors, GST-C1D and GST-MPP6 (see previous figure).

3.2.2.2 Human RRP6 is able to degrade RNA with secondary structure

The poly (A) substrate used in the previous section is not representative of RNA substrates that the exosome is likely to encounter in the cell as it does not form any secondary structures. We therefore investigated the exonuclease activity of both yeast and human RRP6 in the degradation of an RNA substrate containing a stem-loop. A 27 nt fragment of the U14 snoRNA which forms a stem-loop structure with 10 nt involved in base-pairing interactions, 17 nt forming the loop and a single nucleotide overhang at the 3' end was radio-labelled at the 5' end (Figure 3.4C). Neither hRRP6exo or yRRP6 were able to degrade this substrate *in vitro* (Figure 3.4D). Purified hRRP6 did, however, degrade the stem-loop RNA with a ladder of shorter RNAs becoming detectable over time (Figure 3.4D). The activity of hRRP6 on this structured substrate is notably lower than on poly (A) RNA with only approximately 10 nt being removed within two hours and a significant proportion of full length RNA remaining intact for the duration of the time course.

3.2.2.3 The 3'-5' exonuclease activity of RRP6 is stimulated by C1D and inhibited by MPP6

In the nucleus, RRP6 is associated with cofactor proteins, C1D and MPP6, and although these cofactors are required for many of the functions carried out by the RRP6-exosome, for example in 5.8S rRNA processing, it is not clear what role they play (Schilders et al, 2005; Schilders et al, 2007). We have been able to demonstrate protein-protein interactions between RRP6 and its cofactors, C1D and MPP6, so next investigated whether these cofactors influenced the exonuclease activity of RRP6.

Recombinant C1D, MPP6 and RRP6 were incubated individually or in various combinations with the 5' end-labelled U14 stem-loop labelled RNA substrate and samples taken over a time course. No exonuclease activity was observed for either C1D or MPP6 alone. RRP6 displayed 3'-5' exonuclease activity with the appearance of a ladder of truncated 5' labelled substrates after 5 min (Figure 3.4E). When RRP6 was combined with C1D, the exonuclease activity was enhanced as a more significant increase in the accumulation of short RNA species of 16-18 nt was detected (Figure 3.4E). Even after 30 min incubation, these short fragments were barely detectable when RRP6 was mixed with MPP6, indicating that the exonuclease activity of RRP6 was inhibited (Figure 3.4E). Further, the RNA substrate was rapidly degraded, in 5 min to a 22 nt species by RRP6 alone, whereas when MPP6 was also present this fragment was only detectable to a comparable level after 15 minutes. When RRP6 was combined with both C1D and MPP6, the exonuclease activity observed was

comparable to that of RRP6 alone. Taken together, these data suggest that C1D promotes the *in vitro* exonuclease activity of RRP6 while MPP6 regulates or inhibits this activity.

3.2.3 Protein-protein interactions between exosome and snoRNP proteins/biogenesis factors

Association of the human exosome with pre-snoRNAs has been demonstrated *in vivo* by immunoprecipitation of pre-U3 and pre-U8 with antbodies raised against the exosome component, RRP46 (Watkins et al, 2004; Watkins et al, 2007). It was, therefore, decided to investigate if direct protein-protein interactions between exosome subunits and both core snoRNP proteins and snoRNP biogenesis factors could also be detected *in vitro*. Core snoRNP proteins fibrillarin and 15.5K and the snoRNP biogenesis factors TIP48, TIP49, BCD1, NUFIP and NOP17 were expressed with N-terminal GST-tags in *E. coli* and purified using their tags (Figure 3.5A, left panel). The core snoRNP proteins, NOP56 (residues 1-458) and NOP58 (residues 1-435) were expressed with N-terminal thrioredoxin (TRX) and C-terminal His tags in *E. coli* and were purified using their His-tags (Figure 3.5A, right panel). These truncated versions of NOP56 and NOP58 have been shown to function normally in snoRNP complexes (Gautier et al, 1997; Lafontaine & Tollervey, 2000). These recombinant snoRNP proteins were incubated with the His-tagged core exosome proteins described in section 3.2.1 (Figure 3.5B) and any complexes formed were purified using glutathione sepharose (GST-tagged proteins) or anti-thioredoxin bound protein G sepharose (NOP56 and NOP58). Reactions containing the ATPases, TIP48 and TIP49, were supplemented with 1 mM ATP to reflect the ATP-bound state of these proteins during snoRNP biogenesis. Complexes formed were analysed by SDS-PAGE followed by Western blotting to detect the His-tagged exosome proteins (Figure 3.5C). None of the exosome proteins were observed to co-purify with the GST or TRX-tags implying that no interactions were taking place and verifying that any exosome proteins co-purified with the snoRNP proteins were detected due to interactions with the proteins themselves. No interactions were detected between exosome proteins and the mature snoRNP proteins, fibrillarin, 15.5K, NOP56 or NOP58. A robust interaction was detected between the snoRNP biogenesis factor, BCD1, and the exosome protein, MTR3. Weaker, but repeatable, interactions were detected between the ATPases, TIP48 and TIP49, and the S1/KH domain exosome proteins, RRP40 and CSL4. No interactions were detected between the exosome proteins and proteins found in the mature snoRNP. These data imply that the exosome is associated with pre-snoRNPs but may suggest that it dissociates when the mature snoRNP is formed.

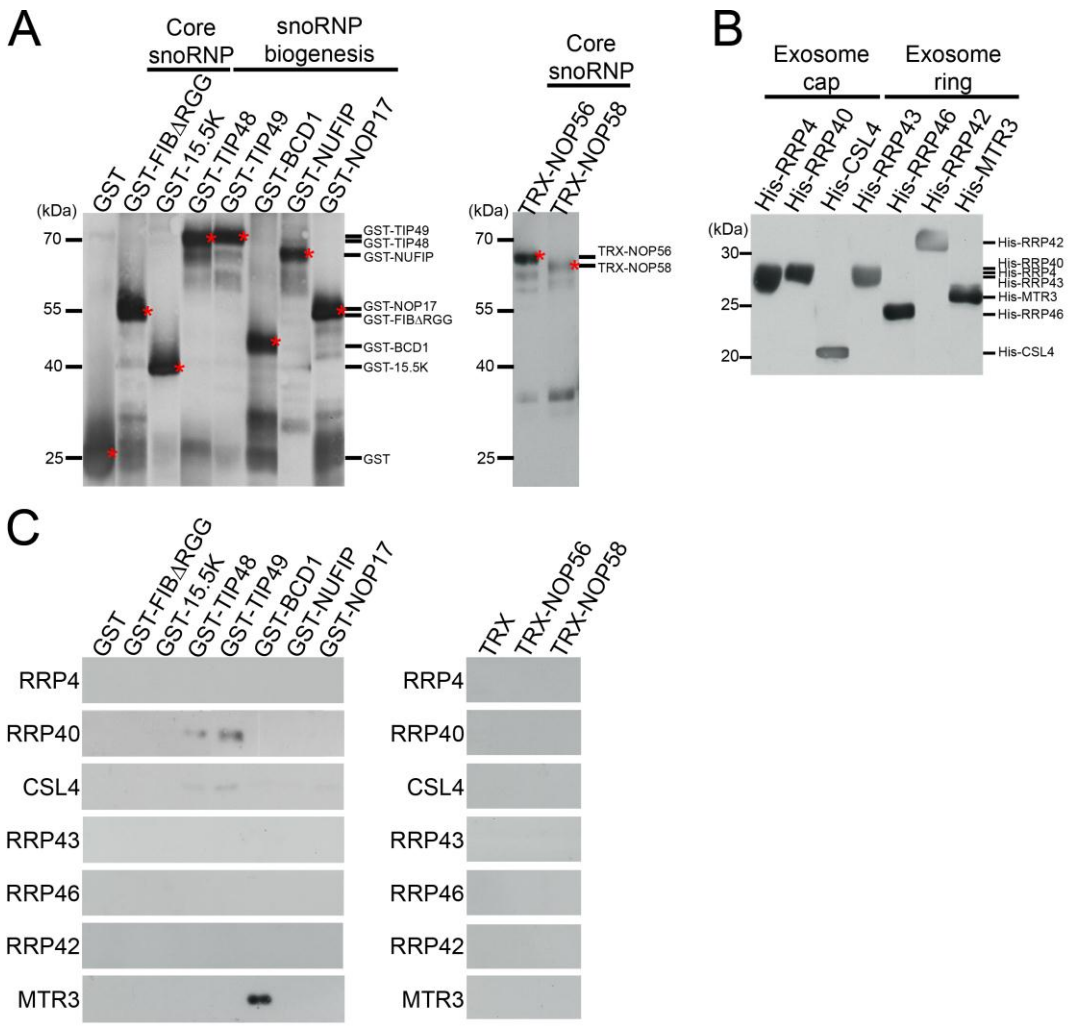


Figure 3.5 Protein-protein interactions between exosome proteins and snoRNP proteins A) GST- or His/Thioredoxin (TRX)-tagged core snoRNP proteins and GST-tagged snoRNP biogenesis factors were over-expressed in E. coli and purified through their GST or His tags. Purified proteins were separated by SDS-PAGE and transferred for Western blotting using antibodies to detect the GST tag (left panel) or the His tag (right panel). Red asterisks mark the full length proteins. Protein loading represents 10% input for interaction experiments (C) except for NOP58 where the amount of protein was increased two fold. **B)** His-tagged core exosome proteins were over-expressed in E. coli and purified using their His-tags. Purified proteins were separated by SDS-PAGE and analysed by Western blotting using an antibody which detects the His-tags. Protein loading represents 10% input for the interaction experiments shown in (C). **C)** Recombinant GST, GST-tagged snoRNP proteins, TRX or His/TRX snoRNP proteins were bound to either glutathione sepharose or anti-thioredoxin bound protein G sepharose and incubated with each of the His-tagged core exosome proteins. Co-purified complexes were separated by SDS-PAGE and analysed by Western blotting using an antibody which detects the His-tags.

3.2.4 Sequencing of precursor snoRNAs

To understand how the exosome may be involved in snoRNA processing we first wanted to investigate the nature of box C/D snoRNA precursor sequences and the pathways by which they are removed in normal cells. The human exosome has been shown to be associated with two box C/D pre-snoRNAs, U3 and U8 making these pre-snoRNAs appropriate candidates for investigation (Watkins et al, 2004; Watkins et al, 2007). The mechanisms by which these two pre-snoRNAs are thought to be processed are, however, different. U8 snoRNA has a 25 nt precursor sequence that is removed in a multi-step pathway with four distinct precursors identifiable by Northern blotting (Watkins et al, 2007), while the U3 snoRNA has a shorter precursor that, it is proposed, is removed in a single processing step. Although human pre-U3 processing has not been thoroughly investigated, in rat, two different U3 precursors have been detected (Stroke & Weiner, 1985; Watkins et al, 2004). We therefore, determined the distribution of 3' end of pre-U3 in wild-type human cells.

Total RNA was extracted from HeLa cells and an oligonucleotide adaptor ligated to the 3' ends of all RNA molecules. Reverse transcription from a primer annealing to the adaptor oligonucleotide generated cDNAs representing all the RNAs to which the adaptor had been ligated. PCR amplifications were then carried out using a forward primer specific for the 3' end of U3. As pre-U3 is much less abundant than mature U3, this primer was designed to extend from the 3' end of the mature sequence 2 nt into the precursor sequence so that pre-snoRNAs would be specifically amplified (Figure 3.6A). PCR products were separated on a non-denaturing polyacrylamide gel (Figure 3.6B) and the band corresponding to pre-U3 excised. PCR products were extracted from the gel and cloned into pGEMTeasy. DNA was extracted from 85 colonies and sequenced. 15 plasmids were sequenced using reactions terminating at each of the four bases and the remainder were sequenced using only reactions terminating at thymine bases (Figure 3.6C). A graphical representation of the number of U3 precursors found terminating at each nucleotide in the precursor sequence is given in Figure 3.6D.

Sequencing data showed that the longest precursor sequences were 13 nt long and could be defined as 5'-TTTTCCTTCTTGA-3'. This sequence differs from the published precursor sequence of U3 (5'-TTTTCCTTCATT-3') which had been defined based on data from Northern blot analysis and comparison to database sequences (Watkins et al, 2004). The approach described here, however, assesses the 3' ends of human pre-U3 directly.

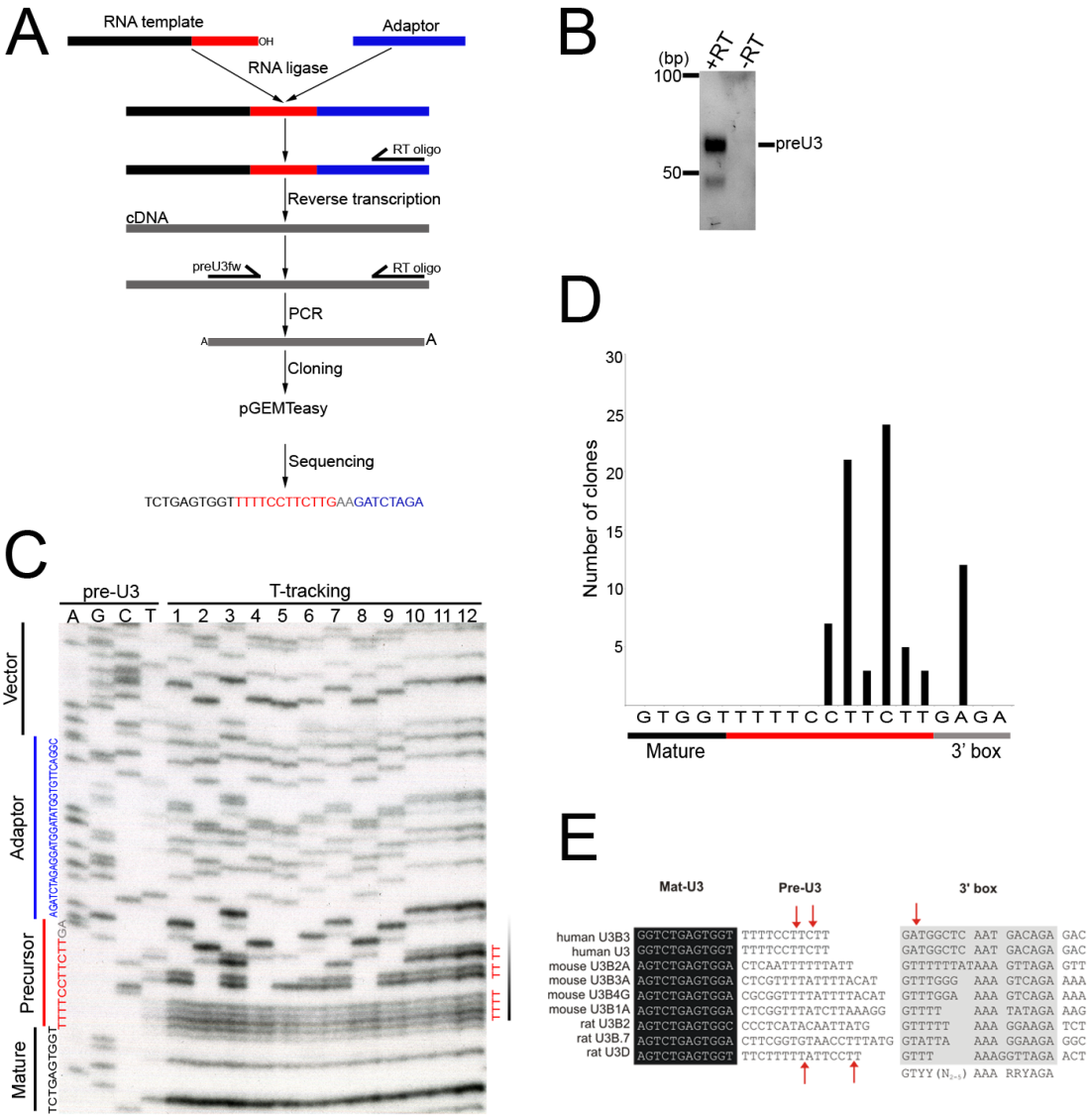


Figure 3.6 U3 precursor sequencing A) Outline of the method used for pre-snoRNA sequencing. Black lines indicate mature snoRNA sequence, red indicates precursor sequence, blue indicates adaptor oilgonucleotide and grey indicates cDNA/PCR products. Reaction steps are named adjacent to flow diagram arrows. **B)** Total RNA was extracted from HeLa cells and an adaptor oligonucleotide ligated to 3' ends from which reverse transcription (+ or- enzyme as indicated) was followed by PCR using a preU3-specific primer. PCR products were separated on a 6 % non-denaturing polyacrylamide gel and visualised with UV. The positions of DNA markers is shown to the left. The pre-U3 specific band which was excised is indicated. **C)** PCR products shown in B were extracted from the acrylamide gel and cloned into pGEMTeasy vector. DNA was extracted from positive clones and sequenced. A full-length precursor sequence terminating at each of the bases is shown to the left. Samples were separated on a denaturing 12 % polyacrylamide/7 M urea gel and products visualised by autoradiography. **D)** Bar chart showing number cloned pre-U3 sequences terminating at each nucleotide in the precursor sequence. **E)** Alignment of U3 precursors from human, mouse and rat adapted from Watkins et al. Cell 2004. Sequences are aligned with respect to the mature snoRNA sequence and the position of the 3' box (Neuman de Vegvar et al., 1986; Hernandez and Weiner, 1986; Neuman de Vagvar and Dahlberg, 1990). 3' ends of intermediate precursors processing steps are indicated with red arrows for rat (Stroke and Weiner, 1985), below and for human sequences, above.

Our data demonstrate that the majority of U3 precursors terminate at one of three different nucleotides indicated by the upper red arrows in Figure 3.6E. This suggests that, like U8 precursors (Watkins et al, 2007), human U3 precursors are heterogeneous and are also processed in a step-wise manner. As the U3 precursors only differ in length by 2-3 nt they can only be distinguished by direct sequencing methods as applied here. Previously published data (Stroke & Weiner, 1985) have shown that pre-U3 in rat is processed via two intermediates also indicated in Figure 3.6E. As we find that both U3 and U8 pre-snoRNAs are processed using a similar multistep mechanism, this may indicate a common pathway for pre-snoRNA processing.

A similar approach was used to try to define the precursor sequence of U8. Although the majority of plasmids from the colonies screened did not contain U8 sequences, sequencing data from the U8-containing clones obtained, identified several pre-snoRNAs terminated 4, 13, 17 and 23 nt from the 3' end of the mature sequence. These correspond with the published precursor intermediates (Watkins et al, 2007). The number of pre-U8 clones was, however, insufficient for any conclusions to be drawn about the relative ratios of each of these precursors.

3.2.5 RNAi depletion of key exosome subunits does not significantly affect pre-snoRNA or mature snoRNA levels

To determine if the exosome is involved in the 3' processing of snoRNAs, the levels of pre-snoRNAs in cells depleted of exosome subunits and cofactors were investigated. If the activity of the exosome is required for converting pre-snoRNAs into mature snoRNAs, decreasing exosome levels is likely to inhibit processing causing pre-snoRNAs to accumulate with a concomitant decrease in mature snoRNA levels. Consistent with this hypothesis, deletion of either Rrp6 or Rrp47 from yeast cells causes the accumulation of 3' extended snoRNAs (e.g. snR38, snR50, snR52) and snRNAs such as U6 (Mitchell et al, 2003).

siRNA duplexes targeting the nuclear exosome subunits, RRP46, RRP6 and DIS3 and the cofactors, C1D, MPP6 and MTR4, were chemically transfected into HeLa cells. A siRNA duplex targeting the mRNA of firefly luciferase, a gene not expressed in HeLa cells, was also used as a negative control. 60 hours after transfection, cells were harvested. Western blot analysis of proteins from RNAi treated cells using antibodies specifically detecting each protein of interest showed that each of the exosome proteins had been depleted to <10% of normal levels while the levels of the cytoskeletal protein, actin, were unaffected (Figure 3.7A). Each siRNA duplex

specifically depleted its target protein and none were observed to affect the levels of any other exosome proteins.

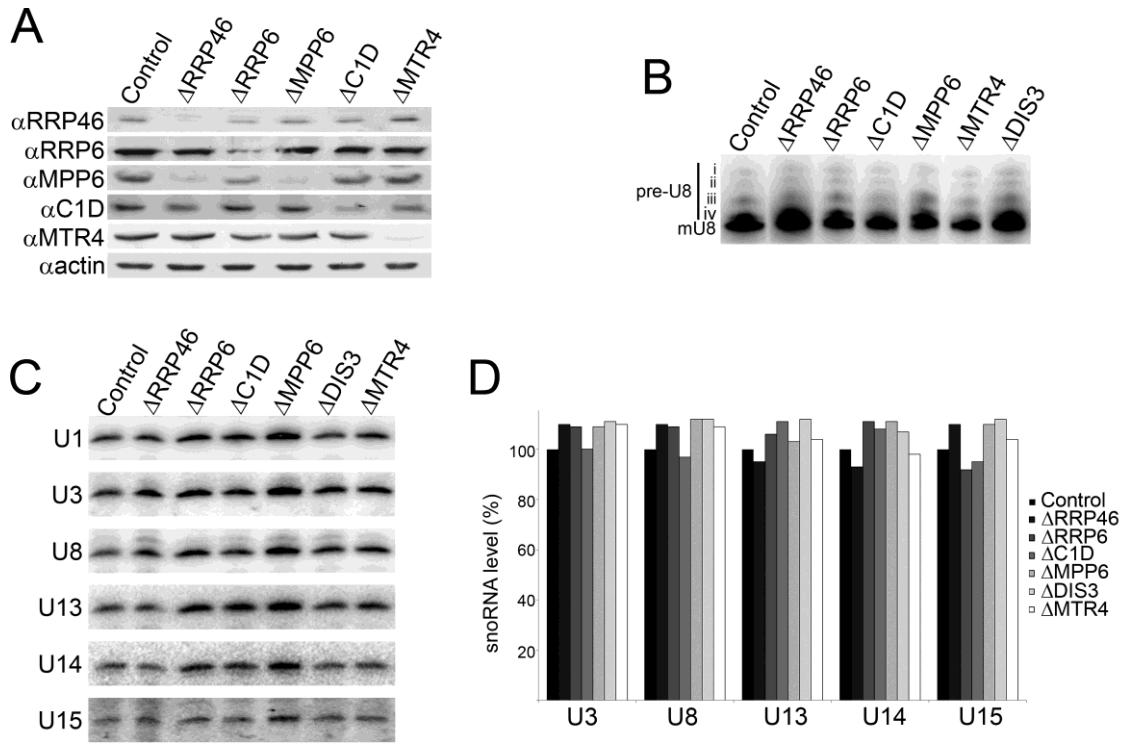


Figure 3.7 Depletion of exosome subunits by RNAi does not lead to changes in precursor or mature snoRNA levels A) HeLa cells were chemically transfected with siRNA duplexes to deplete firefly luciferase (control), RRP46, RRP6, C1D, MPP6, MTR4 or DIS3. Protein was extracted from RNAi depleted cells after 60 hours and analysed by SDS-PAGE followed by Western blotting. Proteins of interest were detected by antibodies specific to each protein, indicated to the left of each panel **B)** RNA was extracted from siRNA treated cells, separated on an 8% acrylamide/7M urea gel, transferred to a nylon membrane and hybridised with an oligonucleotide probe targeting mature U8 snoRNA. **C)** RNA was also analysed by Northern blotting using probes hybridising to the sn/snoRNAs; U1, U3, U8, U13, U14 and U15. **D)** Quantification of sn/snoRNA levels shown in C) were carried out using ImageQuant software. snoRNA levels were normalised to U1 and are given as a percentage compared to snoRNA levels in cells treated with the control siRNA.

To determine if exosome depletion affected pre-snoRNA levels, RNA was extracted from RNAi treated cells and analysed on a denaturing polyacrylamide gel followed by Northern blotting using a probe hybridising to U8. This snoRNA was chosen as it has a long precursor sequence meaning pre-snoRNAs are readily distinguishable from mature snoRNAs using this method. This showed that no significant increase in the levels of any of the four U8 precursors could be observed after depletion of any of the exosome subunits (Figure 3.7B). It was not possible to perform a similar analysis of U3 precursors in cells depleted of exosome proteins using Northern blotting as the U3 precursor sequence is short and the mature U3 snoRNA is relatively so abundant that the precursor cannot easily be distinguished by Northern blotting.

We further examined the levels of a range of different mature snoRNAs in cells depleted of specific exosome proteins. HeLa cells were transfected with siRNAs targeting exosome proteins and cofactors as above. RNA was analysed by PAGE followed by Northern blotting using probes hybridising to the independently transcribed snoRNAs (U3, U8 and U13), snoRNAs derived from introns (U14 and U15) and the snRNA, U1. RNAs were visualised using a phosphorimager and quantified. U1 was used as a loading control and the levels of all the snoRNAs were normalised to U1 levels (Figure 3.6C, D). No significant differences in the levels of any of these snoRNAs could be observed by Northern blot analysis. Quantitation showed that the levels of all these snoRNAs in each of the exosome knockdowns varied by <10 % compared to control cells and no consistent trends were observed for any particular exosome protein on either class of snoRNA. It was concluded that RNAi depletion of exosome proteins and cofactors did not significantly affect snoRNA levels.

3.2.6 Exonucleases are required for turnover of snoRNAs

Another possible function of the exosome is the degradation, rather than the processing, of snoRNAs. As part of other work presented in Chapter 4, POP1, a core protein of the RNase MRP snoRNP, was depleted both individually and in combination with the 5'-3' exonuclease, XRN2. Western blot analysis of proteins from siRNA treated cells showed that the targeted proteins were significantly depleted (Figure 3.8A). RNA from siRNA treated HeLa cells was analysed by PAGE followed by Northern blotting using a probe hybridising to the RNA component of RNase MRP. This revealed that depletion of POP1 caused the levels of MRP RNA to decrease to 37 % of those in cells treated with a control siRNA (Figure 3.8B). Interestingly, depletion of XRN2 caused MRP RNA levels to increase almost two-fold and when XRN2 was co-depleted with POP1, MRP RNA levels were 90 % of those in control cells (Figure 3.8B). This increase in the levels of a mature RNA following depletion of an exonuclease indicated that XRN2 was important for the turnover of MRP RNA and raised the possibility that this method could be used to identify whether the exosome also played a role in degradation of either MRP RNA or box C/D snoRNAs.

Co-depletion of either of the active subunits of the exosome, RRP6 or DIS3, with POP1 did not rescue MRP RNA levels as XRN2 did, indicating that the exosome is not required for turnover of MRP RNA (Figure 3.8C). Depletion of any of the core box C/D snoRNP proteins causes a decrease in snoRNA levels (Newton et al, 2003) and (Watkins lab, unpublished data). HeLa cells were transfected with siRNA duplexes targeting NOP58, RRP6, DIS3, XRN2 or combinations of NOP58 with one of the

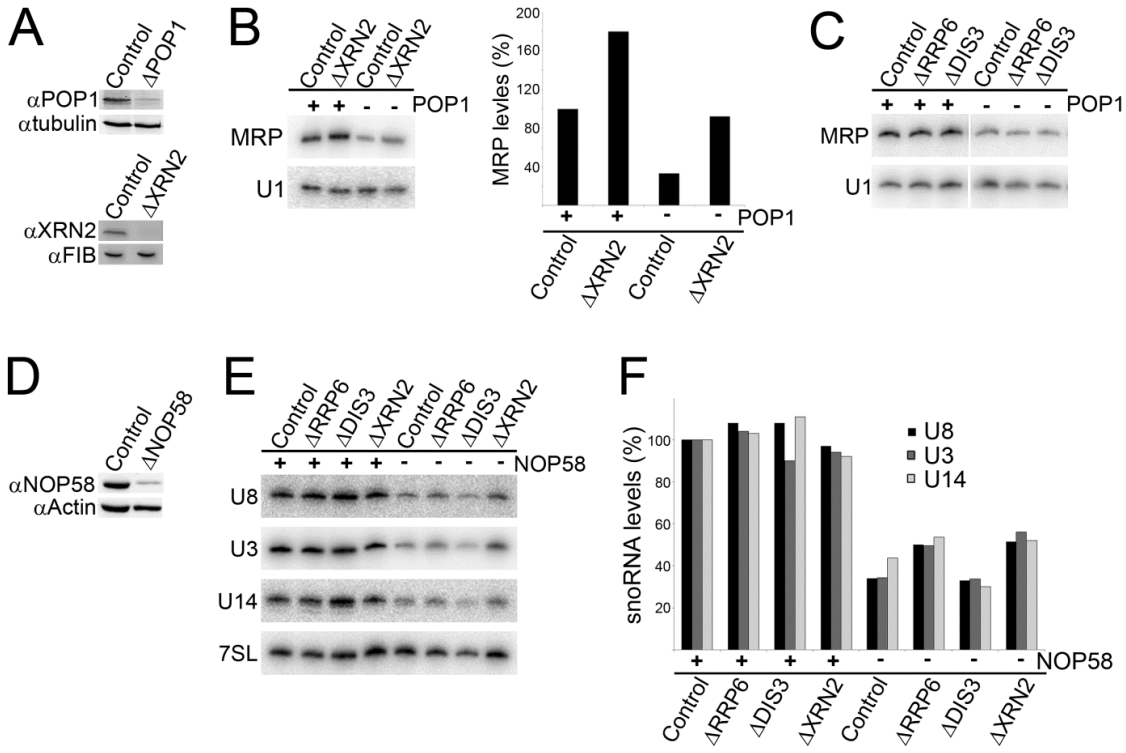


Figure 3.8 Exonucleases are important for turnover of snoRNAs A) HeLa cells were transfected with a control siRNA or siRNA duplexes targeting POP1, XRN2 or both proteins. After 60 hours cells were harvested and protein levels were analysed by Western blotting. Due to lack of anti-POP1 antibody, the membrane carrying control and POP1 samples was probed in the laboratory of G. Pruijn. **B)** RNA was separated on an 8 %polyacrylamide/7M urea gel by electrophoresis and transferred to a nylon membrane. Northern blot hybridisation was carried out using probes recognising MRP or U1 snRNA. + indicates that POP1 is present and – represents depletion of POP1. RNA was visualised using a phosphorimager and quantified using ImageQuant software. **C)** A similar experiment to that described in (A) was performed but instead of XRN2, cells were depleted of the active exosome subunits, RRP6 or DIS3. **D and E)** HeLa cells were depleted of the exonuclease proteins shown above the panel individually (+) or in combination with NOP58 (-). Western blot analysis of proteins from cells in which NOP58 had been depeleted was carried out. Northern blot hybridisation was performed using probes to detect U3, U8 and U14 snoRNAs and the RNA component of the signal recognition particle, 7SL. **F)** Results shown in (E) were quantified using ImageQuant software and loading was normalised to 7SL levels. Relative levels of each snoRNA were then calculated by comparison to their levels in control cells.

exonucleases. Western blot analysis of proteins from siRNA treated cells showed that NOP58 was significantly depleted (Figure 3.8D). RNA was analysed by PAGE followed by Northern blotting using probes hybridising to the independently transcribed snoRNAs, U3 and U8, the intronic snoRNA, U14, and the RNA component of the signal recognition particle (SRP), 7SL. The levels of 7SL were used as a loading control and the levels of all the snoRNAs were normalised to this. Depletion of NOP58 caused the levels of U3, U8 and U14 to be decreased approximately three fold (Figure 3.8E, F). Co-depletion of DIS3 with NOP58 did not significantly restore the levels of any of the snoRNAs suggesting that DIS3 is not involved in snoRNA turnover (Figure 3.8E, F). In contrast, depletion of either RRP6 or XRN2 with NOP58 rescued U8 levels by 15%, U14 levels by approximately 10% and U3 levels by 15 and 22% respectively (Figure 3.8E, F). These results were reproducible and from this we concluded that both RRP6

and XRN2 are involved in the turnover of both independently transcribed and intronic snoRNAs.

3.2.7 Formation of the 3' end of 5.8S rRNA involves the exosome in human cells

We next decided to investigate roles of the exosome in pre-rRNA processing in human cells. One of the best characterised roles of the yeast exosome is in 3' processing of 5.8S pre-rRNA and this function has been shown to be conserved in human cells (Mitchell et al, 1996; Schilders et al, 2005; Schilders et al, 2007; Tomecki et al, 2010b). HeLa cells were transfected with siRNA duplexes targeting RRP46, RRP6, C1D, MPP6, MTR4 or DIS3 and cells were harvested after 60 h. RNA was extracted from RNAi treated cells and analysed by PAGE followed by Northern blotting and hybridisation using probes targeting the mature 5.8S sequence and the 5' end of ITS2. Following depletion of each of the exosome proteins and cofactors two extended precursor forms of 5.8S (Pre-I and Pre-II), which were not found in control cells, were detected (Figure 3.9). These precursors contained sequences at 5' end of ITS2 demonstrating that they are 3' extended precursors of 5.8S. Depletion of different exosome subunits leads to variation in the relative extent to which each of the precursors accumulates. When RRP46 or the helicase, MTR4, were depleted predominantly the longer, Pre-II form was detected. Conversely, in the absence of MPP6, Pre-I is seen to accumulate to a much greater extent than Pre-II. Pre-I and Pre-II accumulate to a similar extent upon depletion of either RRP6, its cofactor, C1D, or DIS3.

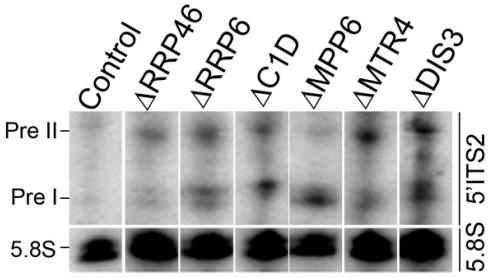


Figure 3.9 The exosome and its cofactors are required for 3' processing of 5.8S rRNA HeLa cells were transfected with a control siRNA or siRNA duplexes targeting the core exosome (RRP46), the active exosome subunits, RRP6 or DIS3, or the nuclear cofactors, C1D, MPP6 or MTR4. Cells were harvested after 60 hours and RNA was extracted. RNA was analysed on an 8% polyacrylamide/7M urea gel followed by Northern blotting using probes hybridising to the mature 5.8S pre-rRNA or the 5' end of ITS2.

3.2.8 Exonucleases are required for the degradation of pre-rRNA fragments

In yeast, another key role of exonucleases in the nucleus is the degradation of pre-rRNA fragments that have been released by endonucleolytic cleavages during processing (Lebreton et al, 2008; Petfalski et al, 1998; Schaeffer et al, 2009; Schneider et al, 2009). This has subsequently also been shown to be the case in mouse cells

(Wang & Pestov, 2011). The 5'ETS is removed from the pre-rRNA transcript in a stepwise manner via three cleavage sites, A', A_0 and A_1, in higher eukaryotes. It was, therefore, of interest to see if any 5'ETS fragments accumulated in the absence of either RRP6 or XRN2 in HeLa cells.

RNAi was used to deplete RRP6 or XRN2 from HeLa cells as previously described. RNA was extracted and analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using probes hybridising to regions of the 5'ETS: 5' to A' (ETS1), between A' and A_0 (ETS2) or between A_0 and the 5' end of 18S (ETS3) (Figure 3.10A). In control cells, none of these 5'ETS fragments were detected but after depletion of XRN2, the ETS1 and ETS3 fragments were seen to accumulate suggesting that XRN2 participates in their turnover (Figure 3.10B). The probe hybridising to the ETS3 region shows significant cross-hybridisation with 18S but the ETS3 fragment, which is approximately 200 nt longer, can be seen to accumulate above 18S after XRN2 depletion. Significantly less of the ETS3 fragment was seen to accumulate than either ETS1 or ETS2. Depletion of RRP6 did not lead to accumulation of either of these fragments but did lead to accumulation of the ETS2 fragment (Figure 3.10B). Two differently sized ETS2 fragments were observed suggesting that turnover of this fragment occurs in multiple stages.

When detected by Northern blotting following agarose-glyoxal gel electrophoresis, the ETS1 fragment appeared as a diffuse band suggesting it is composed of a heterogeneous population of RNAs. To investigate this, RNA from cells depleted of XRN2 was analysed using a 4 % polyacrylamide gel followed by Northern blotting using the probe hybridising to ETS1. The ETS1 fragment accumulated after XRN2 depletion migrated as two bands and comparison to a labelled size markers shows that this fragment is approximately 400 nt in length (Figure 3.10C). These observations are consistent with this fragment being generated by endonucleolytic cleavage at the A' site and XRN2 being the primary exonuclease involved in its degradation.

When RRP6 was depleted, accumulation of a novel intermediate, 37S*, was also detected using both the ETS2 and ETS3 probes but not the ETS1 probe (Figure 3.10B). This indicates that the transcripts from which 37S* are derived have undergone

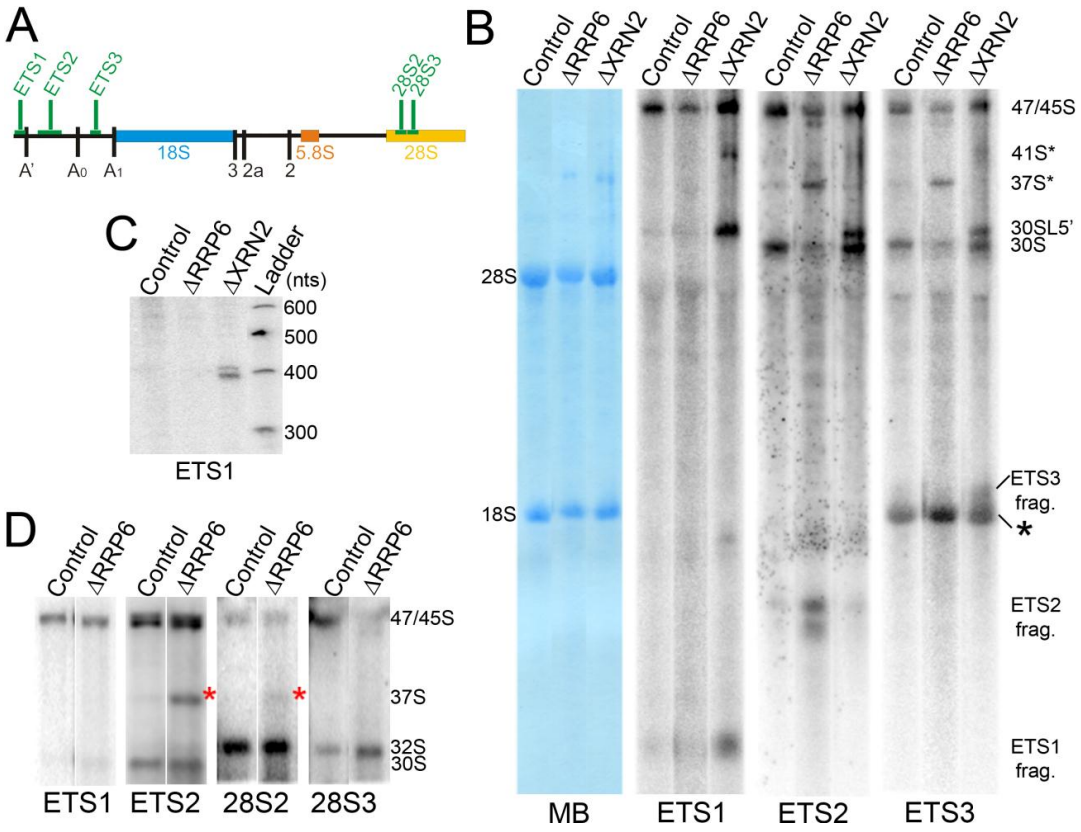


Figure 3.10 Exonucleases in degradation of pre-rRNA fragments A) Schematic outline of human pre-rRNA showing the relative positions of important cleavages and probes using in Northern blot hybridisation. **B)** RNAi using siRNAs targeting either firefly luciferase (Control), RRP6 or XRN2 mRNAs was carried out in HeLa cells. Cells were harvested and RNA extracted 60 hours after transfection. RNA was analysed on a 1.2 % agarose-glyoxal gel which was then transferred to a nylon membrane. The membrane was stained using methylene blue to visualise 28S and 18S mature rRNAs. Northern blot hybridisation was carried out using probes targeting either the 5'-A' (ETS1), A'-A_0 (ETS2) or A_0-A1 (ETS3) fragments of pre-rRNA shown in A. RNAs were visualised using a phosphorimager. Pre-rRNA intermediates detected by each probe are shown to the right of each panel and the asterisk marks cross-reactivity of the ETS3 probe with 18S. **C)** RNA was also separated on a 4 % polyacrylamide/7 M urea gel, transferred to a nylon membrane and hybridised using the ETS1 probe. A 5' end labelled 100 bp DNA ladder was run alongside RNA samples. Results were visualised using a phosphorimager. **D)** RNA from HeLa cells treated with either a control siRNA or depleted of RRP6 was analysed on a 1.2 % agarose-glyoxal gel and transferred to a nylon membrane. Northern blotting was carried out using probes hybridising to ETS1, ETS2 and two regions of 28S. RNA was visualised using a phosphorimager. Red asterisk marks accumulation of 37S*.

A' cleavage but not downstream processing. Assuming that the 5' end of 37S* maps to the A' cleavage site, the size of this intermediate implied that the 3' end would terminate either at the 3' end of ITS2 or within 28S. Northern blot mapping of 37S* in cells depleted of RRP6, using probes spanning this region, was carried out to address this question. 37S* was detected using probes hybridising 1300 nt into 28S but not when using a probe hybridising 1700 nt from the 5' end of 28S (Figure 3.10D). This demonstrated that the 3' end of 37S* lies between 1300-1700 nt downstream of the 5' end of 28S. This suggests it is not an rRNA processing intermediate but that instead, it is a degradation product normally turned over by the exosome. A very weak signal was

detected for 37S* in control cells suggesting this is a naturally occurring fragment that accumulates much more significantly when the exosome is depleted.

Similarly, depletion of XRN2 lead to the accumulation of a novel intermediate, 41S*, that can be detected with all of the 5'ETS probes (Figure 3.10B). 47S is at least 5 fold less abundant than 45S and following XRN2 depletion, 41S* accumulated to a level comparable with 47S suggesting that it is a minor product. Based on the size of this intermediate, the 3' end of 41S* is predicted to lie within the mature 28S sequence indicating that it is a degradation intermediate that is normally turned over from the 5' end by XRN2.

In addition, after RNAi depletion of XRN2, Northern blot hybridisation with the ETS1 probe revealed accumulation of another aberrant processing intermediate, 30SL5', (Figure 3.10B) which will be discussed in more detail in Section 3.2.11.

3.2.9 The exonuclease activity of RRP6 is required for degradation of the ETS2 fragment

By depletion of RRP6 we have shown that this exonuclease is required for degradation of the ETS2 fragment. This does not, however, demonstrate that the enzymatic activity of RRP6 is responsible for this. To confirm whether the exonuclease activity of RRP6 is required for turnover of the ETS2 fragment, stably transfected HEK293 cell lines were generated in which endogenous RRP6 could be depleted by RNAi and expression of FLAG- tagged wild-type, inactive protein (RRP6exo) or the FLAG-tag alone (pcDNA5) could be induced by addition of tetracycline.

The RRP6exo expressed in these cells incorporates the D313A modification shown to inhibit the *in vitro* activity of RRP6 (Figure 3.4). Expression of tagged proteins in these cell lines utilises the normal cellular expression mechanisms and will be targeted by RNAi pathways as for endogenous proteins. In order to express the tagged protein during knockdowns, five individual base changes within the target sequence of each siRNA that did not alter the amino acid sequence of the open reading frame but would prevent siRNA annealing were introduced (Figure 3.11A).

To enable the tagged proteins to be expressed to the same level as the endogenous protein, the concentration of tetracycline required to induce 1:1 expression was determined. Expression of FLAG-tagged protein was tested over a range of tetracycline concentrations (0-1 µg/ml) by treating cells for 48 hours before harvesting and analysis by Western blotting (Figure 3.11B). In each cell line, increasing the tetracycline concentration led to increased expression of FLAG-tagged proteins

corresponding to the correct sizes while levels of the SSU processome protein, PNO1, did not vary. Western blotting using an antibody against RRP6 showed the presence of both the endogenous and tagged forms of each protein. It was determined that the concentration of tetracycline required to mimic endogenous protein levels was 100 ng/ml for RRP6. It should be noted that it was not possible to over-express RRP6 even at the maximum possible tetracycline concentration of 1 µg/ml. Further, increasing expression of the tagged form of RRP6 suppressed endogenous protein.

The sub-cellular localisations of the FLAG-tagged, mutant forms of RRP6 were determined and compared to that of the wild-type FLAG-tagged protein by immunofluorescence since such modifications can lead to protein mislocalisation preventing induced proteins from carrying out the functions of the endogenous protein (Watkins lab, unpublished data). Cells expressing the FLAG-tag only, FLAG-tagged wild-type RRP6 or the RNAi resistant forms of FLAG-RRP6 or RRP6exo were induced with the optimial concentration of tetracycline on cover-slips for 48 hours prior to fixing. Antibodies specifically recognising the FLAG-tag or the core snoRNP protein, fibrillarin, were used in immunofluorescence and DAPI was used to visualise nuclear material. DAPI binds most strongly to DNA in the nucleoplasm and nucleoli appear as “holes” within the nucleus. Immunofluorescence showed that in all cell lines fibrillarin was localised exclusively to the nucleolus and is found both diffused and in concentrated speckles as expected (Puvion-Dutilleul et al, 1991) **(**Figure 3.11C). Consistent with previous immunofluorescence studies in which RRP6-specific antibodies were used (Tomecki et al, 2010b), here, RRP6 was found throughout the nucleus but significantly concentrated in the nucleolus and some weak RRP6-staining could be observed in the cytoplasm (Figure 3.11C).

Cells expressing the FLAG-tag (pcDNA5) or RNAi-resistent FLAG-RRP6 or -RRP6exo were transfected with siRNA duplexes targeting the mRNAs of either firefly luciferase (C) or RRP6 (Δ6) and after 48 hours, the expression of RRP6 was analysed by Western blotting. Following the control knockdown in pcDNA5 cells, the endogenous protein was still present whereas in cells expressing either FLAG-RRP6 or -RRP6exo treated with RRP6 siRNAs, only the larger, FLAG-tagged forms of RRP6 could be detected (Figure 3.11 D). RNA extracted from these cells was analysed by Northern blotting using a probe hybridising to ETS2. The levels of 47/45S and 30S were unaffected and the ETS2 fragment did not accumulate in either control cells or in those where the endogenous protein had been replaced by FLAG-tagged RRP6 (Figure 3.11E). When the endogenous protein was replaced with FLAG-RRP6exo, however, significant accumulation of the ETS2 fragment was observed demonstrating that the activity of RRP6 is indeed required for the degradation of this fragment (Figure 3.11E).

It is interesting to note that 37S* was not detected in HEK293 cells during RNAi-depletion and rescue experiments suggesting that this intermediate may be specific to HeLa cells.

A similar approach was attempted to determine if the exonuclease activity of XRN2 was required for degradation of the ETS1 and ETS3 fragments but this was not successful. HEK293 cell lines for the expression of RNAi-resistant FLAG-tagged XRN2 and inactive XRN2 were produced but it was not possible to induce expression of the tagged proteins to the same level as the endogenous protein.

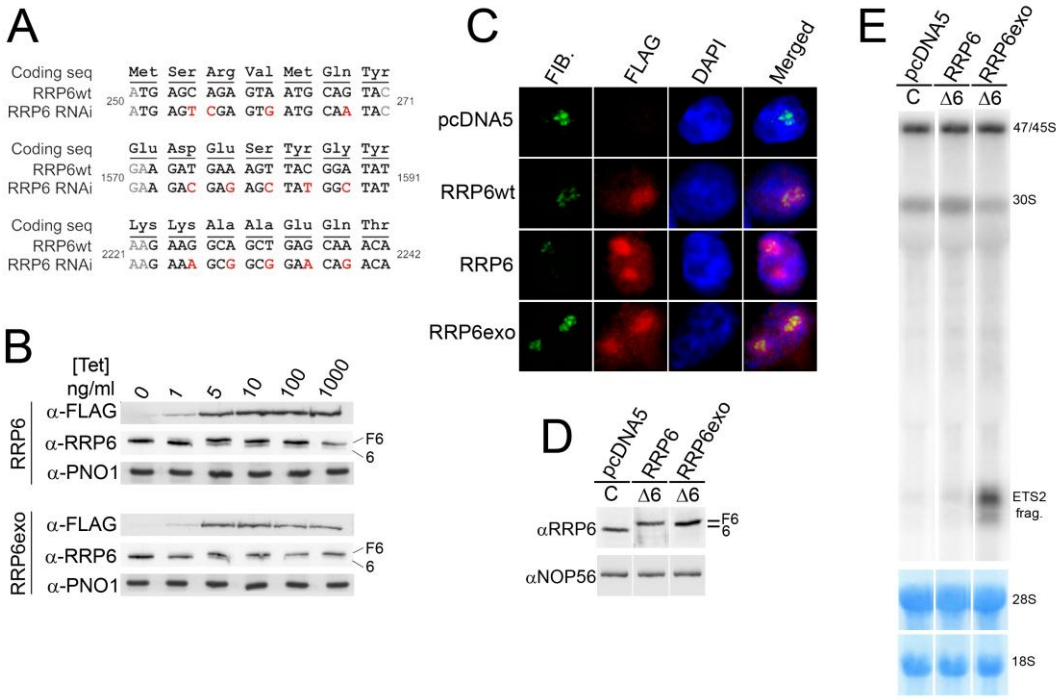


Figure 3.11 The exonuclease activity of RRP6 is required for degradation of the ETS2 fragment A) Outline of siRNA target sites in RRP6. Point mutations are shown in red. **B)** Cells from RNAi resistant FLAG-RRP6 (RRP6) or FLAG-RRP6exo (RRP6exo) HEK293 stable cell lines were cultured in the presence of tetracycline at different concentrations (given above the panel) for 36 hours before harvesting. Proteins were extracted and analysed by SDS-PAGE followed by Western blotting using the antibodies shown to the left of the panel. Tagged (F6) and endogenous (6) proteins are identified. **C)** Cells expressing only the FLAG tag (pcDNA5), FLAG-RRP6 (RRP6wt), RNAi resistant FLAG-RRP6 (RRP6) and RNAi resistant FLAG-RRP6exo (RRP6exo) were grown on coverslips and induced with 100 ng tetracycline for 36 hours, fixed and assayed by immunofluorescence using α-Fibrillarin (green), α-FLAG (red) and DAPI (blue). **D)** pcDNA5, RRP6 and RRP6exo cells were induced with tetracycline for 24hours then treated with siRNAs targeting either firefly luciferase (C) or RRP6 (Δ6) for a further 48 hours before harvesting. Proteins were extracted and analysed by SDS-PAGE followed by Western blotting using antibodies against RRP6 or NOP56. FLAG-tagged and endogenous proteins are identified to the right of the panel. **E)** RNA extracted from these cells was analysed by agarose-glyoxal gel electrophoresis followed by transfer to a nylon membrane. Methylene blue staining shows the levels of mature 28S and 18S rRNAs. Northern blotting using a probe hybridising to ETS2 detected pre-rRNA species identified to the right of the panel.

3.2.10 The role of DIS3 and exosome cofactors in degradation of the ETS2 fragment and 37S*

We have established that RRP6 is required for degradation of both the ETS2 fragment and 37S* in HeLa cells. To investigate if the other active subunit of the exosome in the nucleus, DIS3, or any of the nuclear cofactors of the exosome were also required for these functions, HeLa cells were transfected with siRNAs targeting RRP46, RRP6, C1D, MPP6, MTR4 or DIS3 and pre-rRNA was analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using the ETS2 probe. 37S* was seen to accumulate following depletion of the core exosome and each of the cofactors tested, demonstrating that they are all required for degradation of this intermediate (Figure 3.12). However, 37S* was not accumulated above background levels following depletion of DIS3, implying that this protein did not participate in degradation of this aberrant pre-rRNA. The ETS2 fragments were accumulated after depletion of RRP46, RRP6, MPP6 or MTR4, but not when C1D or DIS3 were depleted (Figure 3.12). Interestingly, these data imply that the exosome requires specific activities and cofactors depending on which RNA substrate is degraded, with C1D being important for degradation of 37S* but not the ETS2 fragment and RRP6, rather than DIS3, required for both of these functions. Further, the longer and shorter forms of the ETS2 fragment accumulated to different extents depending on which exosome component was depleted, although this varied between experiments. Here, depletion of RRP46 lead to significant accumulation of comparable amounts of both forms of the ETS2 fragment whereas depletion of RRP6, MPP6 or MTR4 cause the longer ETS2 fragment to be accumulated to a much higher level than the shorter form. This may suggest that different exosome cofactors are required at different stages of the turnover of this fragment.

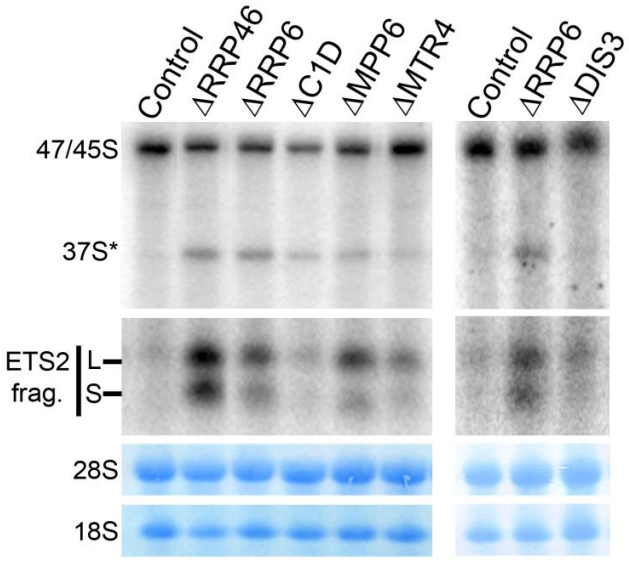


Figure 3.12 The role of DIS3 and the exosome cofactors in the degradation of the ETS2 fragment and 37S* HeLa cells were transfected with siRNA duplexes targeting the core exosome (RRP46), RRP6, C1D, MPP6, MTR4, DIS3 or a control siRNA. After 60 hours, cells were harvested and RNA was extracted. RNA was analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using a probe hybridising between the A' and A_0 sites (ETS2). 28S and 18S rRNA was visualised by methylene blue staining.

3.2.11 XRN2 and MTR4 are required for A' cleavage in HeLa cells

Depletion of XRN2 from HeLa cells caused accumulation of 30SL5' (Figure 3.10A, 3.13A). This intermediate is detectable with the ETS1 probe indicating that it is a 5' extended form of 30S in which A' cleavage has not taken place (Figure 3.13A). This implies that XRN2 is required for A' cleavage in HeLa cells as its homologues are in mouse and plant cells (Wang & Pestov, 2011; Zakrzewska-Placzek et al, 2010). Depletion of the exosome and TRAMP helicase, MTR4, also lead to accumulation of 30SL5' suggesting that this protein plays a hitherto unidentified role in A' cleavage in HeLa cells (Figure 3.13B). Consistent with roles in A' cleavage, depletion of either of these proteins caused accumulation of 47S. Hybridisation with the ETS2 probe which also detects the normal, 30S pre-rRNA processing intermediate, showed that 30SL5' does not accumulate as significantly following MTR4 depletion as after depletion of XRN2 although both proteins were depleted to similar extents, suggesting that XRN2 plays a more important role in this cleavage step (Figure 3.13A, B).

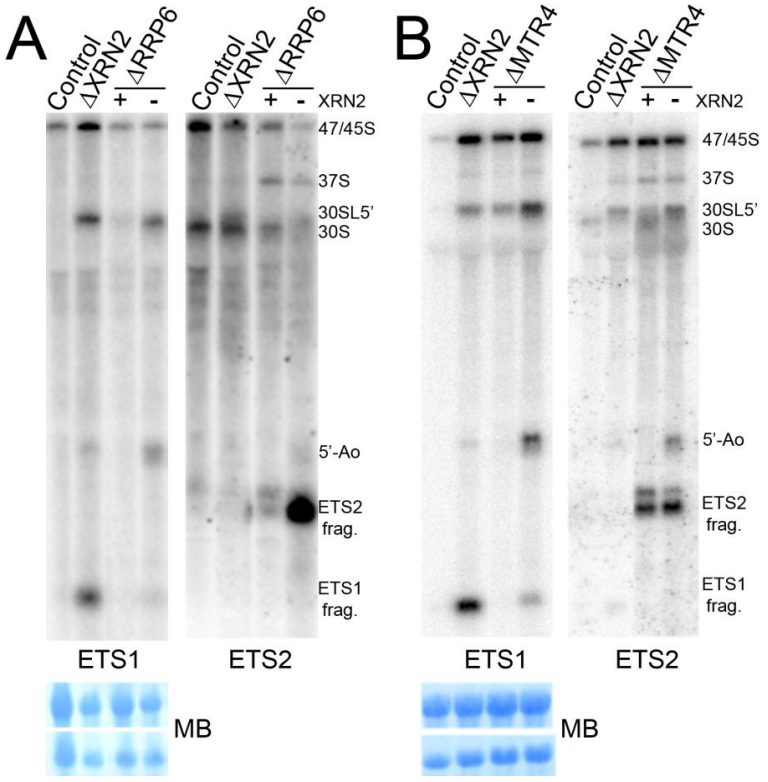


Figure 3.13 XRN2 and MTR4 are required for A' cleavage A) HeLa cells were transfected with siRNAs to deplete either XRN2, RRP6 or combinations of these proteins or a control siRNA. + indicates cells in which XRN2 levels are normal and – denotes cells in which XRN2 has been depelted 60 hours later, RNA was extracted, separated on a 1.2% agarose-glyoxal gel and transferred to a nylon membrane which was stained with methylene blue to visualise 28S and 18S rRNAs. Northern blotting using probes hybridising to ETS1 and ETS2 was carried out and RNAs visualised using a phosphoimager. Pre-rRNA intermediates identified are given to the right of the panel. **B)** Experiments were carried out as in A) using siRNAs to depleted XRN2, MTR4, both these proteins or a control siRNA.

Northern blotting of RNA from cells depleted of XRN2 using the ETS1 probe detected both the ETS1 fragment and also an additional, longer 5'ETS fragment extending from the 5' end of the transcript to the A_0 cleavage site (5'-A_0) (Figure 3.13A). This implies that when XRN2 is depleted, A' cleavage can be bypassed and removal of the 5'ETS can occur through the A_0 site instead. Indeed, when XRN2 and RRP6 were co-depleted, the ETS1 fragment was barely detectable and the 5'-A_0 fragment was observed to accumulate more significantly than when only XRN2 was depleted (Figure 3.13A). Our data would suggest, therefore, that the 3' end of the majority of the ETS1 fragment accumulated after XRN2 depletion is not generated by A' cleavage but instead is produced by 3'-5' exonucleolytic processing by RRP6. This is consistent with the fact that RRP6 is normally responsible for the degradation of this pre-rRNA sequence as part of the ETS2 fragment. Northern blot analysis using the ETS2 probe also revealed that when RRP6 and XRN2 are co-depleted, the ETS2 fragments are accumulated to levels far exceeding those observed when only RRP6 is depleted (Figure 3.13A). In different experiments, the accumulation of the ETS2 fragment was not always as strong as presented in this figure but always exceeded the levels observed when RRP6 was depleted individually. This suggests that although the ETS2 fragment is not accumulated when XRN2 is depleted individually, that XRN2 does participate in its turnover. The significant accumulation of the ETS2 fragment when both RRP6 and XRN2 are depleted may also imply that, as above, the 5'-A_0 arising from bypassing A' cleavage, is processed to the 5' side of the A' cleavage site by 5'-3' exonucleases other than XRN2. Depletion of MTR4 did not cause accumulation of the 5'-A_0 fragment (Figure 3.13B). When MTR4 and XRN2 are co-depleted this fragment accumulates much more significantly than when only XRN2 is depleted and accumulation of the ETS1 fragment was notably reduced (Figure 3.13B). It may be interesting to note that the ratio of the longer and short ETS2 fragments is shifted towards the shorter product by co-depletion of XRN2 with either of the exosome proteins, MTR4 or RRP6.

3.2.12 Core box C/D snoRNP proteins are essential for complete removal of the 5'ETS but not for A' cleavage

To assess how significant the roles of XRN2 and MTR4 in A' cleavage are, it was decided to deplete other factors that are known to be required for this processing step and compare the severity of defects observed. The U3 snoRNP is required for efficient A' cleavage in both yeast and humans and depletion of the U3-specific protein, hU3-55K, causes some accumulation of 30SL5' (Prieto & McStay, 2007). The U3 snoRNP, and other snoRNPs involved in pre-rRNA processing also contain the

common core snoRNP proteins, NOP56, NOP58 and fibrillarin. The effect of depleting these proteins on pre-rRNA processing, particularly on A' cleavage, have not been investigated in human cells.

Each of these U3 snoRNP proteins was, therefore, depleted from HeLa cells using RNAi. Western blot analysis of proteins from siRNA treated cells showed that the levels of NOP56, NOP58, fibrillarin and hU3-55K were all depleted to <10% of that in control cells while the levels of the cytoskeletal protein, actin, or the core exosome protein, CSL4, were unaffected (Figure 3.14A).

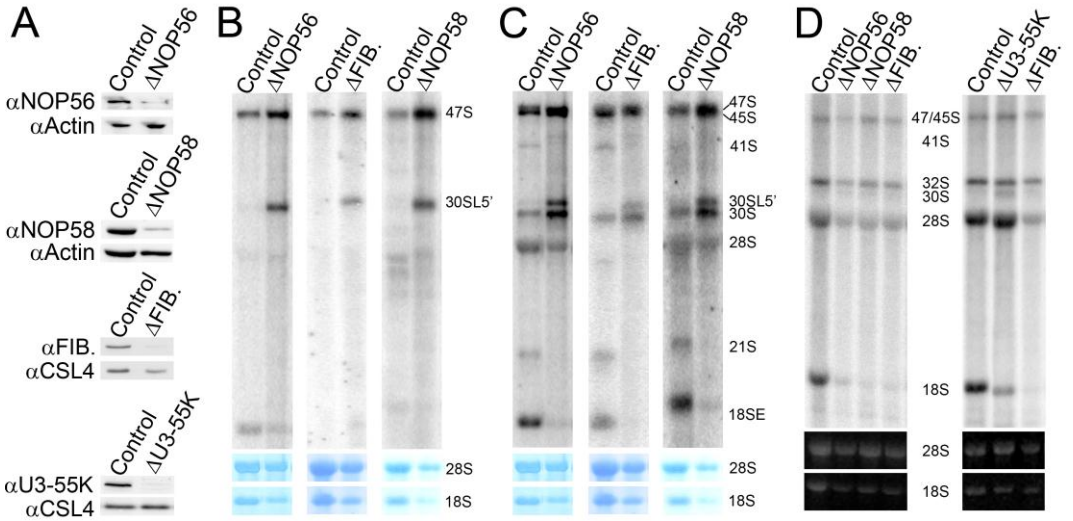


Figure 3.14 Depletion of snoRNP proteins affects A' cleavage but completely inhibits 5'ETS removal HeLa cells were transfected with siRNA duplexes targeting the core snoRNP proteins, NOP56, NOP58 and fibrillarin, the U3-specific protein, hU3-55K, or control siRNAs. **A)** 60 hours after transfection, cells were harvested and both RNA and proteins extracted. Proteins were separated by SDS-PAGE and analysed by Western blotting using antibodies against the targeted proteins and either actin or CSL4 as loading controls. **B and C)** RNA was analysed using agarose-glyoxal gel electrophoresis followed by Northern blotting using a probe hybridising near the 5' end of the pre-rRNA transcript (ETS1) (B) or the 5' end of ITS1 (C). Mature 28S and 18S rRNAs were visualised by methylene blue staining. The positions of the pre- and mature rRNAs are given to the right of the panel. **D)** 48 hours after siRNA transfection, cells were depleted of phosphate, pulse-labelled with ^{32}P orthophosphate and grown in normal media for 3 hours before harvesting. RNA was extracted and analysed by agarose-glyoxal gel electrophoresis and visualised using a phosphorimager. Total RNA was visualised using ethidium bromide staining (UV).

Pre-rRNA from RNAi treated cells was analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using probes hybridising to the 5' end of the pre-rRNA transcript (ETS1) or the 5' end of ITS1. Depletion of NOP56, NOP58 or fibrillarin caused an increase in 47S levels indicating that A' cleavage is impaired in these cells (Figure 3.14B). Also observed when levels of the snoRNP proteins were decreased was the accumulation of 30SL5', a 5' extended form of 30S in which A' cleavage has not occurred (Figure 3.14B, C). This implies that when any of the core snoRNP proteins are depleted, A' cleavage is inefficient. However, after snoRNP

protein depletion, both 45S and 30S were still detected, with 30S being more abundant than 30SL5', implying that in the majority of transcripts, A' cleavage occurs normally. In contrast, accumulation of both 21S and 18SE, the downstream processing products of 30S were dramatically decreased when the snoRNP proteins were depleted (Figure 3.14C). This indicates that while the core snoRNP proteins are important for efficient A', they are essential for complete removal of the 5'ETS. The requirement for the core snoRNP proteins for A' cleavage appears comparable to that for MTR4 or XRN2 as 47S and 30SL5' are observed to accumulate to similar extents, although it must be noted that the proteins were depleted to slightly different extents.

In addition to U3, other snoRNPs are required for pre-rRNA processing and of particular interest is U8 as this snoRNP is only found in higher eukaryotes. In *Xenopus laevis*, U8 is required for biogenesis of the large subunit rRNAs so it was decided to investigate the role of NOP56, NOP58, fibrillain and hU3-55K in biogenesis of both 18S and 28S rRNA using metabolic labelling experiments. HeLa cells were treated with siRNAs targeting the mRNAs of each of the snoRNP proteins. Cells were then metabolically labelled using ^{32}P orthophosphate which was chased with media containing unlabelled phosphate. Cells were harvested after 3 h and pre-rRNA was separated by agarose-glyoxal gel electrophoresis, transferred to a nylon membrane and analysed using a phosphorimager. In cells transfected with a control siRNA, 47/45S, 28S, 18S and the 28S precursor, 32S, were all detected (Figure 3.14D). Depletion of NOP56, NOP58 or fibrillarin caused significant decreases in the levels of 28S rRNA and almost abolished production of 18S rRNA (Figure 3.14D). In contrast, depletion of the U3-specific protein, hU3-55K, lead only to a decrease in 18S levels and did not affect 28S accumulation (Figure 3.14D). Taken together, this indicates that the U3 snoRNP is only required for cleavages affecting small subunit rRNA biogenesis while the proteins found in all box C/D snoRNPs are also necessary for biogenesis of the large subunit rRNAs, potentially as part of the U8 snoRNP.

3.3 Discussion

We have investigated the roles of the human exosome in processing of two different substrates, pre-rRNA and snoRNAs. Extending previous observations showing that the human exosome is associated with pre-snoRNAs (Watkins et al, 2004; Watkins et al, 2007), we have demonstrated *in vitro* interactions between snoRNP biogenesis factors and exosome proteins. Our data question whether the exosome is directly involved in the processing of snoRNA precursors but demonstrate a role for the exosome in turnover and quality control of snoRNAs. We have also

shown that in HeLa cells the exosome and XRN2 have important degradation functions in the recycling of excised pre-rRNA spacer fragments and in the removal of aberrant or excess rRNA precursors. XRN2 and the TRAMP helicase, MTR4, have additional roles in processing of the 5'ETS as they are both required for A' cleavage and XRN2 is clearly important for this step as loss of XRN2 can lead to A' cleavage being bypassed. A' cleavage is similarly inefficient in cells depleted of the core box C/D snoRNP proteins but surprisingly, the major role of these proteins is in 5'ETS processing steps downstream of A' cleavage.

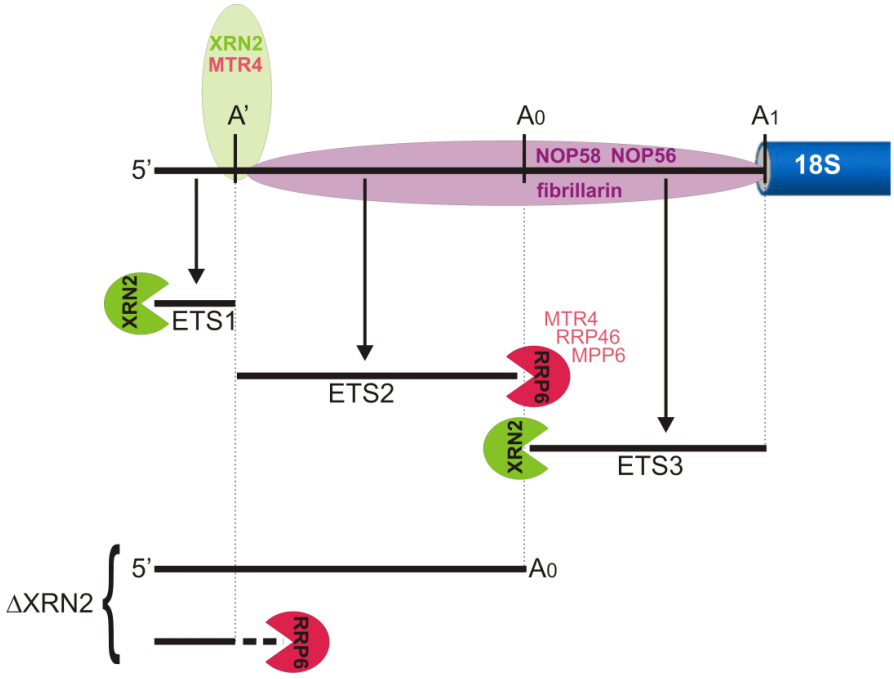


Figure 3.15 Factors required for removal of the 5'ETS in HeLa cells Schematic representation of the 5' external transcribed spacer region (black line) of the pre-rRNA transcript from human cells. Endonuclease cleavage sites are marked above the transcript and the fragments released by these cleavages are shown below. The proteins activities required for degradation of these fragments; exosome (red) and XRN2 (green) are shown with the direction of processing indicated.

The role of exonucleases in degradation of cleaved 5' spacer fragments is conserved from mouse (Wang & Pestov, 2011) to human cells. XRN2 is responsible for the 5'-3' degradation of the fragments extending from the 5' end of the transcript to the A' cleavage site (ETS1) and also from the A_0 to A_1 sites (ETS3). The fragment between the A' and A_0 (ETS2) is, by contrast, primarily degraded from the 3' end by the exosome (Figure 3.15). The main exonuclease responsible for turnover of this fragment is RRP6 and DIS3 does not seem to be involved. Interestingly, the RRP6 cofactors, MPP6 and MTR4 also participate in this processing but C1D does not. Taken together, these data imply that while the A_0 provides an entry point for exonucleolytic processing, that exonucleases are impeded from initiating degradation of these fragments from the site of A' cleavage. Many proteins are required for A' cleavage and the U3 snoRNP

interacts with this region of the pre-rRNA (Borovjagin & Gerbi, 2001; Gerbi et al, 2003) so it is possible that this processing machinery remains associated with either the 3' end of the ETS1 fragment or the 5' end of the ETS2 fragment after cleavage, thereby protecting these ends from the action of the exonucleases. Indeed, in yeast, it has been shown that both the U3 snoRNP and several U3-associated proteins (UTPs) proteins are still associated with the 5'ETS fragments after the 5'ETS cleavages have occurred (Billy et al, 2000; Granneman et al, 2003; Hoang et al, 2005; Krogan et al, 2004; Peng et al, 2004; Wegierski et al, 2001). The two 5'ETS fragments which accumulated when XRN2 was depleted (ETS1 and ETS3) appear to accumulate to different extents. It is not possible to predict how much of these fragments is generated by pre-rRNA processing as they are not normally visible but when both XRN2 and RRP6 were co-depleted, the accumulation of the ETS2 fragment far exceeded the amount of any fragment accumulated when a single exonuclease was depleted. This implies that the accumulation of 5'ETS fragments detected when either XRN2 or RRP6 was depleted individually only represents a small fraction of the total amount of these fragments produced. This suggests that there is redundancy between the exonucleases involved with the Rrp17 homologue, NOL12, being a likely candidate for carrying out the same functions as XRN2 and the REX proteins potentially involved in ETS2 degradation.

The ETS2 and ETS1 fragments detected when the exonucleases RRP6 or XRN2 are depleted both appear as doublets when resolved on agarose-glyoxal and acrylamide gels, respectively, suggesting that they are degraded in multistep pathways. This is supported by the observation that depletion of different subunits and cofactors of the exosome caused variation in the ratio of the longer and shorter forms of the ETS2 fragment detected. Furthermore, co-depletion of either RRP6 or MTR4 with XRN2 also increased accumulation of the shorter form of the ETS2 fragment relative to the longer form, suggesting that different processing activities are required at different stages of turnover of these fragments.

XRN2 is required for A' cleavage in HeLa cells as it is in both mouse and plant cells (Wang & Pestov, 2011; Zakrzewska-Placzek et al, 2010). It has been suggested that XRN2 is not directly involved in this cleavage step but instead, that failure to degrade the ETS1 fragment which also requires XRN2 prevents other factors required for A' cleavage from being recycled (Wang & Pestov, 2011). In *Arabadopsis thalina,* XRN2 has been shown to degrade the 5' end of the full length transcript which exposes the A' site enabling cleavage to take place (Zakrzewska-Placzek et al, 2010). In HeLa cells, depletion of XRN2 caused the accumulation of 30SL5', an aberrant form of the 30S precursor which accumulates if A' cleavage is affected. However, both 30S and

the ETS1 fragment were also detected showing that XRN2 is not essential for this cleavage. It is interesting to note that depletion of XRN2 decreased 45S production significantly while 30S levels were only marginally affected. This suggests that when XRN2 is depleted, A' cleavage is not always the first processing step to occur on the 47S pre-rRNA but can occur after cleavage in ITS1 and convert 30SL5' into 30S. Also observed following XRN2 depletion was a weak accumulation of a fragment extending from the 5' end of the transcript to A_0. This indicates that depletion of XRN2 can result in A' cleavage being bypassed. Co-depletion of XRN2 with RRP6 caused a much more significant accumulation of this fragment and loss of the ETS1 fragment normally detected when XRN2 is depleted. This may suggest that a proportion of the ETS1 fragment detected when XRN2 was depleted individually could actually be derived from the 5'-A_0 fragment undergoing 3'-5' processing by the exosome to a site upstream of A'. Although MTR4 is also required for A' cleavage, the 5'-A_0 fragment was not detected perhaps suggesting that MTR4 is less critical for A' cleavage than XRN2. Taken together these observations suggest that depletion of XRN2 changes the order of cleavage steps required to produce 18S rRNA and promotes cleavage at the A_0 site. It is not, however, clear what role XRN2 or MTR4 perform in A' cleavage. Both XRN2 and MTR4 are involved in the quality control of pre-rRNA processing and in yeast, XRN2 is associated with the rDNA. It is possible therefore, that as part of the surveillance machinery, these proteins function to ensure the correct order of cleavage events. Depletion of the core exosome or RRP6 does not affect A' cleavage so the role of MTR4 in A' processing seems to be an exosome-independent function. It is also possible that the helicase activity of MTR4 may be required to modify the RNA secondary structure of the A' cleavage site to enable the endonuclease to gain access. This would be comparable to how the helicase, Prp43, is proposed to facilitate access of Nob1 to its cleavage site at the 3' end of 18S in yeast (Pertschy et al, 2009).

Association of the U3 snoRNP with pre-rRNA at the A' cleavage site is a key step in formation of pre-ribosomal complexes so it was assumed that the core proteins of the snoRNP would be essential for A' cleavage. While depletion of hU3-55K (Prieto & McStay, 2007) fibrillarin, NOP56 or NOP58 did cause defects in A' cleavage, this processing step was still able to occur, albeit inefficiently. However, cleavages at A_0 or A_1 were completely blocked by depletion of the core snoRNP proteins implying that although the U3 snoRNP binds near the A' cleavage site, that its major function is in fact in the downstream removal of the 5'ETS.

Depletion of exonucleases from HeLa cells also revealed roles for the exosome and XRN2 in pre-rRNA quality control. 37S* and 41S* both appear to be aberrant intermediates normally degraded by the exosome and XRN2, respectively. It is

possible that 37S* is produced by premature termination of transcription within 28S with the partial transcripts then targeted for degradation by the exosome. Alternatively and more likely, the exosome is involved in degradation of excess or aberrant full-length transcripts that need to be removed as part of the cell's RNA surveillance mechanisms. 37S* has undergone A' cleavage which suggests that the normal pre-rRNA processing pathway is initiated before the transcript is targeted for degradation. 37S* is detectable at very low levels in control cells suggesting that the 3' end of this fragment may represent a natural stalling point of the exosome. This could be due to a particular sequence or structure that it is difficult for the exosome to degrade. Since 37S* is probably derived from a full length transcript, other exonucleases such as the REX proteins must be able to initiate degradation but are unable to proceed beyond the 3' end of 37S*. Interestingly, the RRP6 cofactor, C1D, is needed for degradation of 37S* while it is not required for turnover of the ETS2 fragment. This possibly indicates that when the exosome is carrying out recycling functions it does not need C1D but this protein is more important for targeting the exosome to aberrant substrates which are to be degraded as part of the quality control pathways. However, DIS3 does not participate in turnover of 37S* as RRP6 does, suggesting that this fragment is degraded in the nucleolus from which DIS3 is excluded in human cells. However, DIS3 along with RRP6 does function in processing the 3' end of 5.8S pre-rRNA which is likely to occur in the nucleoplasm.

In yeast, deletion of key proteins of the exosome, Rrp6 and Rrp47, causes accumulation of 3' extended precursors of both snoRNAs and snRNAs implying that the exosome is required for the maturation of these small RNAs (Allmang et al, 1999a; Costello et al, 2011). The human exosome is present within the same large complexes as the U3 and U8 pre-snoRNPs (Watkins et al, 2004; Watkins et al, 2007) so it was postulated that the function of the exosome in pre-snoRNA processing is conserved in higher eukaryotes. Protein-protein interactions were detected *in vitro* between exosome proteins (MTR3, CSL4 and RRP40) and snoRNP biogenesis factors (BCD1, TIP48 and TIP49) while no interactions were detected with proteins found in mature snoRNPs. The significance of the individual interactions detected has not been investigated further but they support the observation that the human exosome is stably associated with pre-snoRNP complexes. It is possible that more interactions would be detected between exosome proteins and snoRNP biogenesis factors when these proteins are assembled into the exosome and the pre-snoRNP respectively. Depletion of key activities of the human exosome by RNAi did not, however, cause a significant accumulation of precursors of the U8 snoRNA raising doubts about whether the human exosome is required for pre-snoRNA processing as its yeast counterpart is. It was

difficult to determine whether this was also the case for precursors of the U3 snoRNA since the 3' extension is small and precursors cannot readily be distinguished from the mature snoRNA. The levels of mature snoRNAs transcribed independently (U3, U8) or those that are released from introns by splicing (U14, U15) were also not significantly affected by depletion of exosome factors. It is, however, possible that the human exosome plays a non-essential role in processing of pre-snoRNAs with other 3'-5' exonucleases such as the REX proteins potentially also able to perform this function. However, the exosome subunit, RRP6, and XRN2 are both clearly required for the turnover of other non-coding RNAs, box C/D snoRNAs and the unique snoRNA, RNase MRP RNA. Depletion of NOP58 causes a significant decrease in snoRNA levels because pre-snoRNPs which are not associated with this core protein are detected as aberrant and are therefore degraded. Co-depletion of either RRP6 or XRN2 with NOP58 partially restores snoRNA levels implying that the exosome and XRN2 are involved in this quality controlled degradation. It is likely that this also reflects an involvement of RRP6 and XRN2 in the normal turnover of snoRNAs. The U3 snoRNA is produced in excess (Knox et al, 2011) and it is probable that this is also the case for other snoRNAs meaning unnecessary snoRNAs may need to recycled. DIS3 was not found to participate in snoRNA degradation possibly because in human cells, DIS3 is found predominantly in the nucleoplasm unlike RRP6, which is concentrated in nucleoli where snoRNA turnover occurs.

A model of RNA processing by the yeast exosome is proposed in which the activities of Rrp44 and Rrp6 are coordinated such that the dual endonuclease and processive exonuclease activity of Rrp44 is responsible for removing long, structured RNA sequences and this is followed by fine trimming to a defined specific 3' end using the distributive activity of Rrp6. At the time of carrying out our experiments the *in vitro* activity of human RRP6 had not been investigated although data on this have recently been published (Januszyk et al, 2011). Human and yeast RRP6 both possess similar distributive 3'-5' exonuclease activity. We conclude that the activity was distributive rather than processive as a ladder of partially processed fragments was detected over the course of 2 h whereas previously published data for the processive exosome enzyme, Rrp44, show complete degradation of the same substrate in 60 min (Schneider et al, 2009). It appears that human RRP6 is able to degrade poly (A) RNA at a slightly faster rate than yeast Rrp6 but the difference is subtle.

Specific structured RNA substrates accumulate when yeast Rrp6 is mutated (Allmang et al, 1999a) and concordantly, our data showed that yeast Rrp6 is unable to process RNA with secondary structure *in vitro*. In contrast, human RRP6 is able to degrade structured RNA substrates *in vitro*. An explanation for this difference between

the activities of yeast and human RRP6 can be derived from the recently published crystal structure of a fragment of human RRP6 (Januszyk et al, 2011). This showed that the orientation of the catalytic amino acids of RRP6 (the DEDD box) is well conserved between yeast and humans but that a linker between the exonuclease domain and the HRDC domain, which forms one face of the active cleft, is in a more open conformation in the human protein, thereby allowing double stranded RNAs access to the catalytic site. This raises the possibility that human RRP6 is able to degrade RNAs that the yeast protein cannot and therefore may have additional functions *in vivo*. It is proposed that the decision between degradation and site specific processing could be determined by RNA secondary structures causing exonucleases to stall. It is possible that these would not be sufficient to stop the action of human RRP6 implying that alternative mechanisms, such as proteins bound to RNA 3' ends, may be required to halt the activity of RRP6.

A crystal structure of the nine-membered core human exosome is available but little is known about how RRP6 or nuclear cofactors associate with this core. Our data show that RRP6 interacts with all three RNA binding proteins of the exosome cap suggesting that RRP6 binds to the upper face of the exosome possibly at the opening to the central channel through which RNA substrates are proposed to pass. This is consistent with cryo-electron microscopy studies of the *Leishmania tarentole* exosome that show RRP6 binding to this side of the exosome (Cristodero et al, 2008). In contrast, yeast two-hybrid data demonstrate that human RRP6 interacts with proteins of the hexameric ring (Lehner & Sanderson, 2004) and here weak *in vitro* interactions were also detected between RRP6 and RRP43, RRP42 and MTR3. In yeast, Rrp6 and Rrp44 are both associated with nuclear exosomes and since they are both large proteins, it is possible that the anchoring of Rrp44 to the lower face of the exosome by its PIN domain (Schneider et al, 2009), dictates that RRP6 interacts with the cap-face of the complex. In human cells, RRP6 is predominantly found in the nucleolus while DIS3 is localised to the nucleoplasm (Tomecki et al, 2010b) suggesting that complexes containing RRP6 but not DIS3 are formed and it is possible that in these cases, RRP6 does not interact with the exosome in the same way. However, *in vivo* each of the core exosome proteins, used individually in this assay, would be part of the multi-protein complex so we cannot rule out the possibility that the interactions identified here are not fully representative of all the contacts RRP6 may make with the assembled core exosome.

RRP6 appeared to bridge interactions between the core exosome and the cofactors, C1D and MPP6, since neither of these proteins was seen to interact significantly with the core exosome but both displayed robust interactions with RRP6.

C1D and MPP6 are RNA binding proteins and their yeast homologues are important for recognition of particular RNA features, such as double stranded RNA or pyrimidine rich sequences, so could target the exosome to particular substrates. This may suggest that these factors could stimulate processing by the exosome by recruiting the complex to its substrates (Milligan et al, 2008; Schilders et al, 2005; Stead et al, 2007). However, we have shown that their human counterparts are also able to directly modulate the activity of RRP6 *in vitro*. C1D stimulated the exonuclease activity of RRP6 during the degradation of a structured RNA substrate while MPP6 inhibited it, suggesting that these different cofactors may associate at different times during processing of a substrate, altering the activity of RRP6 as appropriate. It is interesting to note that RRP6, but not the processive DIS3 exonuclease, is required for degradation of several long substrates such as 37S* and the ETS2 fragment but the i*n vitro* activity of RRP6 remains distributive even when stimulated by C1D. This implies that in the context of the cell, the exonuclease activity of RRP6 is significantly enhanced compared to recombinant protein *in vitro*. Further, DIS3 is required for processing of 5.8S precursors, but not degradation of 37S* or turnover of the ETS2 fragment and C1D only participates in 3' end processing of 5.8S. Degradation of 37S* and the ETS2 fragments is likely to occur in the nucleolus, from which DIS3 is excluded, but although C1D localises predominantly to nucleoli, this exosome protein is also not required for their turnover. It is clear therefore, that specific exosome activities and cofactors are required for degradation of different substrates.

Chapter Four

Major and minor pre-rRNA processing pathways for internal transcribed spacer 1 in humans

4.1 Introduction

Three of the four ribosomal RNAs (rRNAs), 18S, 5.8S and 28S (25S in yeast), are co-transcribed as a single precursor (pre-rRNA) by RNA Polymerase I (RNA pol I) in the nucleolus (Venema & Tollervey, 1995). This primary transcript contains the mature rRNA sequences along with two external transcribed spacer regions (5'ETS and 3'ETS) as well as two internal transcribed spacer regions (ITS1 and ITS2) (Figure 4.1A). This precursor undergoes both modification and a series of endonucleolytic cleavages and exonucleolytic processing to remove the spacer regions and release the mature rRNAs which are assembled into the large (LSU) and small (SSU) ribosomal subunits (Figure 4.1B) (Henras et al, 2008).

S. cerevisiae has been extensively used as a model for studying pre-rRNA processing and almost all our understanding of the pathways is derived from these studies. Many of the cleavage sites have been mapped and several of the nucleases responsible for particular steps have been identified but others remain elusive (Henras et al, 2008). One of the key steps in pre-rRNA processing is the separation of the small (18S) and large (5.8S and 28S) subunit rRNAs which is achieved by cleavages at A_2 and A_3 in ITS1. Cleavage at A_2 is linked to removal of the 5'ETS and is reported to be carried out by Rcl1 (Horn et al, 2011). These steps are mediated by the small subunit processome complex which is recruited to the pre-rRNA co-transcriptionally and is primarily composed of the subcomplexes: t-UTP, b-UTP, c-UTP, MPP10 and U3 snoRNP (Kos & Tollervey, 2010; Phipps et al, 2011). Following A_2 cleavage, the mature 3' end of 18S is generated by another endonucleolytic cleavage at site D carried out by Nob1 (Lamanna & Karbstein, 2009; Pertschy et al, 2009). A protein of the SSU processome, Rrp5, provides a link between the two ITS1 cleavages as it is required for processing at both A_2 and A_3 (Venema & Tollervey, 1996). Endonucleolytic cleavage at the A_3 site is mediated by the RNA-protein complex, RNase MRP (Lygerou et al, 1996). A_3 cleavage also requires Nop4 (RBM28 in humans) and is followed by 5'-3' exonucleolytic processing by Rat1 (XRN2 in humans) and Rrp17 to generate the mature 5' end of 5.8S rRNA (Berges et al, 1994; Oeffinger et al, 2009; Petfalski et al, 1998; Sun & Woolford, 1994). A number of proteins, termed the A_3 cluster proteins, which includes Erb1 (BOP1 in humans), are required for this processing after A_3

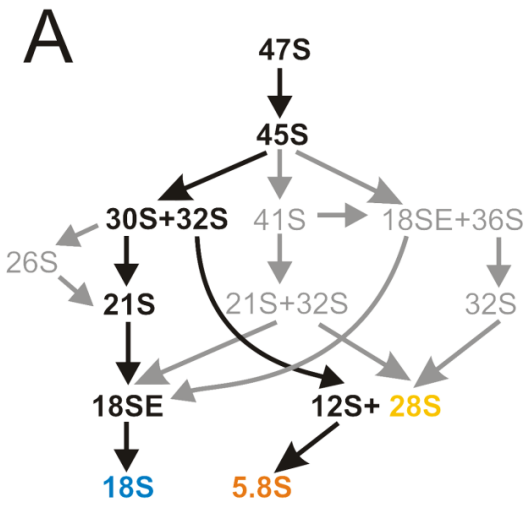
cleavage (Granneman et al, 2011; Rohrmoser et al, 2007; Sahasranaman et al, 2011; Strezoska et al, 2000).

Relatively little is known about pre-rRNA processing in human cells and none of the cleavage sites have been mapped nor have any of the nucleases responsible for mediating these processing steps been identified. Homologues of many of the trans-acting ribosome biogenesis factors identified in yeast, including the nucleases, can be found in human cells although it is not yet clear whether their functions are conserved. Site 2 appears to be the initial cleavage in mammalian ITS1 that separates the SSU and LSU rRNA(s). In higher eukaryotes, pre-rRNA processing appears to be more complex than in yeast and includes both an extra cleavage in the 5'ETS at the A' site and an additional ITS1 processing step that converts 21S into 18SE (Rouquette et al, 2005) (Figure 4.1A, B). It is not understood how this additional ITS1 processing step occurs and two different models have been proposed (Figure 4.1A). One model is based on data showing that the depletion of a late-acting SSU component, ENP1, causes accumulation of a partially processed intermediate smaller than 21S. The heterogeneity of the 3' end of this intermediate led to the suggestion that site 2 cleavage is followed by 3'-5' exonucleolytic processing (Carron et al, 2011). In contrast to this, data from mouse cells provide evidence of an endonucleolytic cleavage at the 2a site near the 3' end of 18S which is stimulated by depletion of the 5'-3' exonuclease, XRN2 (Rat1 in yeast) (Wang & Pestov, 2011). Although in yeast, A_2 is the primary cleavage that separates the LSU and SSU rRNAs, it has been speculated that cleavage at site 2a which generates 18SE, rather than the major site 2 in human pre-rRNA, is equivalent to A_2. This raises the possibility that site 2 is therefore analogous to the yeast A_3 site, which is cleaved by the RNA-protein complex, RNase MRP, but it is not yet clear if this is indeed the case. Depletion of Dicer, the endonuclease linked to miRNA and siRNA production, appears to influence cleavages close to the predicted site 2 cleavage as various differently 5' extended forms of 5.8S rRNA were weakly accumulated (Liang & Crooke, 2011). The significance of this observation is not fully understood and the defects observed were not severe enough to suggest that Dicer is the main endonuclease involved in site 2 processing.

Key questions remain, therefore, about the pathway of mammalian ITS1 processing. Although, similar to yeast, two endonuclease cleavage sites have been identified, it is not clear whether these sites are analogous to the yeast ITS1 cleavages both in terms of the nucleases involved and the manner in which they contribute to processing of ITS1. How ITS1 is processed is of particular interest as this separates the large subunit rRNAs from the small subunit rRNA. As there is an additional ITS1 processing step in higher eukaryotes that is not conserved from yeast, understanding

human ITS1 processing may highlight important differences between the processing pathways used in these different organisms. This is of particular interest as several genetic diseases, such as the ANE and CHH syndromes and Diamond Blackfan anaemia are linked to defects in ribosome production (Freed et al, 2010; Mattijssen et al, 2010b; Narla et al, 2010; Nousbeck et al, 2008). Furthermore, ribosome production is up-regulated in cancer and several ribosome biogenesis factors, such as BOP1 and ENP1, are associated with specific cancers (Chung et al, 2011; Killian et al, 2006; Wang et al, 2009).

We therefore performed a systematic analysis of human homologues of yeast factors linked to ITS1 processing in human cells and characterized the order and nature of the processing events involved.



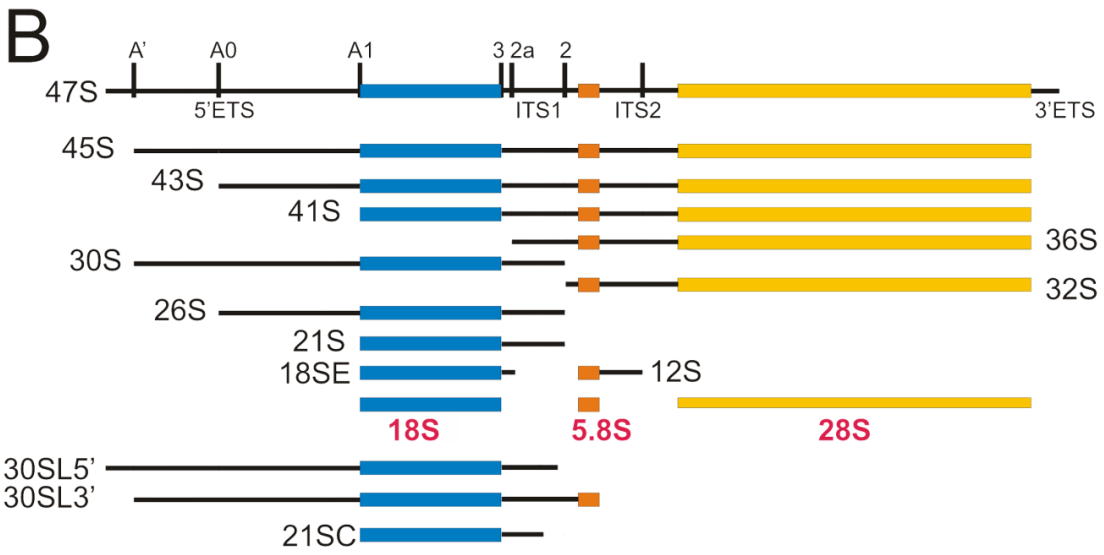


Figure 4.1 pre-rRNA processing pathways in human cells A) Alternative pre-rRNA processing pathways in human cells are shown with the numbers corresponding to the intermediates shown in (B) and arrows indicating the direction of processing. The major pathway is shown in black and minor pathways are shown in grey. **B)** Schematic representations of the initial pre-rRNA transcript containing the mature rRNAs, 18S (blue), 5.8S (orange) and 28S (yellow), the internal transcribed spacer regions (ITS1 and ITS2) and the external transcribed spacers (5'ETS and 3'ETS). The major cleavage sites are marked above the transcript. Normal processing intermediates corresponding to (A) are shown below 47S and three aberrant pre-rRNA fragments described in this chapter are shown at the bottom of this panel.

4.2 Results

4.2.1 Northern blot mapping of ITS1 cleavage site in human pre-rRNA

The two alternative models of how the additional ITS1 processing step is mediated (Carron et al, 2011; Rouquette et al, 2005; Wang & Pestov, 2011) raise questions about the processing sites that exist in human ITS1 and how the ends of the pre-rRNA intermediates resulting from ITS1 cleavages are generated. A Northern blot mapping approach was, therefore, undertaken to determine which ITS1 regions were present in different pre-rRNAs and to define the cleavage sites. RNA was extracted from HeLa cells and separated by agarose-glyoxal gel electrophoresis, transferred to a nylon membrane and analysed by Northern blotting. A series of probes spanning the internal transcribed spacer regions (ITS1 and ITS2) were used for hybridisation. The sequences chosen included the extreme 5' and 3' end of each spacer region and the positions of the internal probes were often dictated by the extreme GC content of human ITS1 sequences. Similar experiments have previously been undertaken by David Tollervey, University of Edinburgh (unpublished data) and the results of these investigations are combined in this dataset (Figure 4.2A, B). The abundance of short pre-rRNAs (30S, 32S, 21S and 12S) relative to 47/45S was calculated as a ratio for each probe and results are given as a bar chart in Figure 4.2B. Several of the probes used significantly cross-hybridised to either 28S or 18S mature rRNAs and in particular instances, this affected the signals detected for 47/45S, 32S and 30S. This only influenced the pre-rRNA of interest in the case of the 6773 probe and here, the fraction that this bias represented was calculated by comparison to Northern blot data generated using a probe directly targeting 28S and was subtracted.

All of the probes used recognised 47/45S but three probes, 6396, 6448 and 6508, only detected 47/45S pre-rRNA and none of the smaller pre-rRNAs, indicating the region spanned by these probes is not associated with either the small or large subunit pre-rRNAs (Figure 4.2A, B). This implies that the cleavage which separates the LSU and SSU pre-rRNAs (site 2) occurs within this region but does not rule out the possibility that more than one cleavage site is used. In addition to 47/45S, the 18S precursors, 30S and 21S, were detected by probes hybridising in the region from the 5' end of ITS1 (5520) to position 6318. Compared to probes hybridising to the 5' end of ITS1, progressively weaker 30S and 21S signals were detected by probes between positions 6121 and 6318. This suggests that the 3' ends of these pre-rRNAs are not homogeneous but instead range between positions 6121 and 6318 implying that following site 2 cleavage, the 3' ends of 30S and 21S are generated by exonucleolytic processing. The large subunit precursors, 32S and 12S, were readily detectable using probes hybridising to ITS2 sequences but a probe at the 3' end of ITS1 only weakly

detected 32S and 12S was not recognised, indicating that all of 12S and the majority of 32S does not contain ITS1 sequences (Figure 4.2A, B). Depletion of the 5'-3' exonuclease, XRN2, increased the amount of 32S containing 3'ITS1 sequences implying that XRN2 is involved in exonucleolytic processing to remove ITS1 sequences from 32S and potentially also 12S (data not shown).

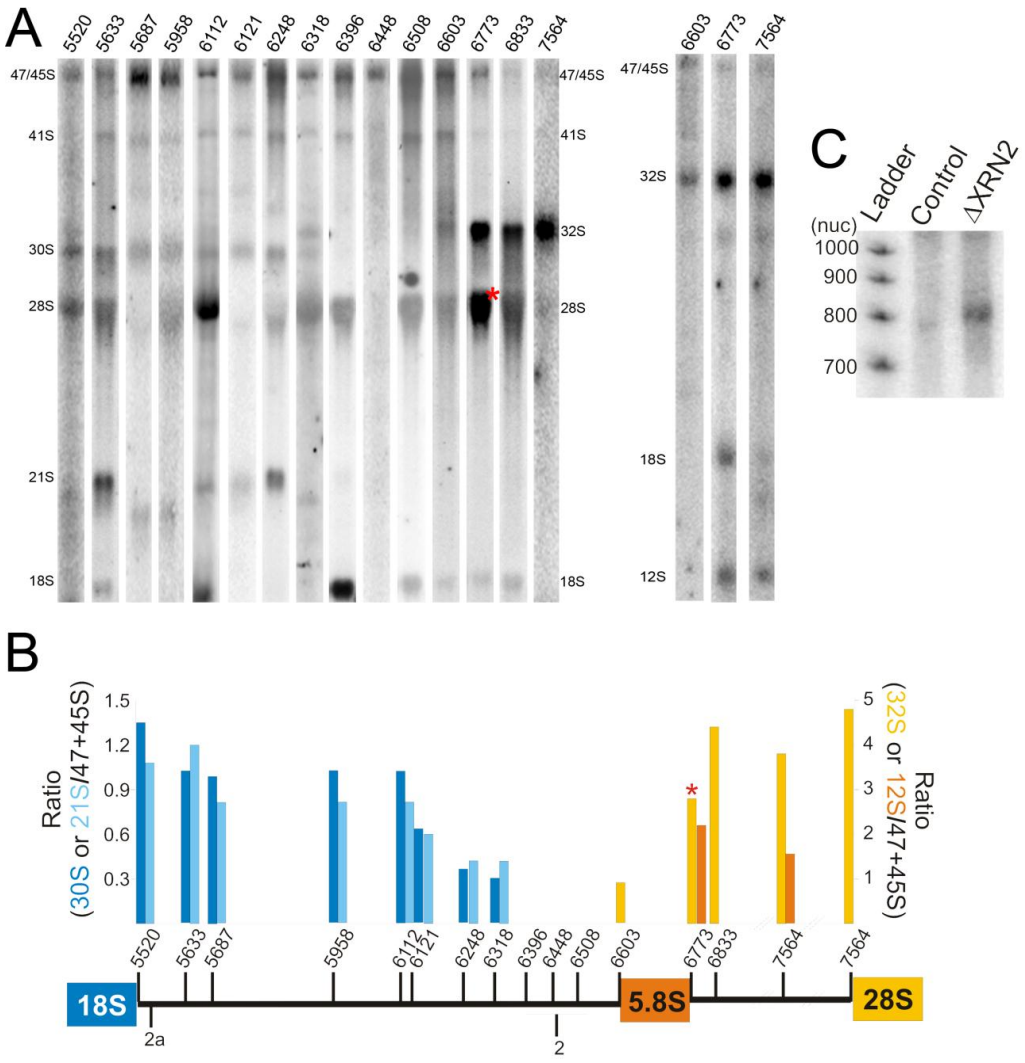


Figure 4.2 Mapping ITS1 cleavages in HeLa cells A) Total RNA was extracted from HeLa cells and separated by agarose-glyoxal gel electrophoresis. RNA was transferred to a nylon membrane and Northern blotting carried out using probes hybridising across ITS1 and ITS2. The number given represents the 5' end nucleotide of the target of each probe within the pre-rRNA sequence. Pre-rRNA intermediates detected are indicated to the left and right of the panel. Red asterisk marks significant probe cross-hybridisation to 28S.). Northern blot data was generated by Katherine Sloan (5520, 5687, 5958, 6112, 6121, 6318, 6448, 7564, 6603, 6773 and 7564) or by David Tollervey (5633, 6248, 6396, 6508, 6603, 6773, 6833). **B)** Levels of 47/45S, 30S, 32S, 21S and 12S pre-rRNA were quantified using Image Quant software and ratios of 30S or 21S and 32S or 12S to 47/45S were calculated. These ratios are given as a bar chart with a schematic outline of ITS1 and ITS2 showing the relative positions of each of the probes. **C)** HeLa cells were transfected with siRNA duplexes to deplete XRN2 or a control siRNA. After 60 h, cells were harvested, RNA was extracted and separated on a 4% acrylamide/7M urea gel alongside a 5'end radiolabelled ladder. RNA was transferred to a nylon membrane and Northern blot hybridisation performed using a probe hybridising to the middle of ITS1 (6121). RNA was visualised using a phosphorimager.

In yeast, fragments of pre-rRNA spacer regions that are released by sequential endonucleolytic cleavages are commonly degraded by exonucleases including the exosome and Rat1 (XRN2). To determine if multiple endonucleolytic cleavages occur in human ITS1, we investigated if fragments of ITS1 are released and accumulate when exonucleases are depleted. HeLa cells were transfected with siRNA duplexes targeting the mRNAs of either XRN2 or the exosome subunit, RRP6. As a negative control, a siRNA duplex targeting firefly luciferase mRNA, which is not expressed in HeLa cells, was also used. RNA was extracted 60 h after siRNA transfection and was separated on a 4% acrylamide denaturing polyacrylamide gel, followed by Northern blotting using probes hybridising to ITS1. Depletion of XRN2 caused accumulation of a fragment of ITS1 that was approximately 800 nt in length (Figure 4.2C). This fragment was not detected in cells transfected with the control siRNA, with those targeting RRP6 or when using a probe hybridising to the very 5' end of ITS1. It was, therefore, concluded that this fragment is generated by endonucleolytic cleavages at site 2 between nucleotides 6396 and 6508, and at site 2a. No other fragments of ITS1 were detected when either XRN2 or RRP6 was depleted suggesting that there is only a single cleavage between nucleotides 6396 and 6508.

Our data indicate that site 2 cleavage is rapidly followed by 5'-3' exonucleolytic processing to generate the 5' end of 5.8S, analogous to the situation in yeast. As the 3' ends of the 18S precursors, 21S and 30S, appear to be heterogeneous, we conclude that these precursors are similarly generated by 3'-5' exonucleolytic processing following site 2 cleavage. This may explain why neither LSU nor SSU precursors could be detected in a region of approximately 160 nt although our data imply that there is only a single cleavage in this region. It appears, therefore, that dual endonuclease cleavages at sites 2 and 2a can occur releasing a fragment of ITS1 which is degraded by XRN2.

4.2.2 RNAi depletion of RRP6, XRN2 or ENP1 leads to altered 18S RNA processing

Detection of the 800 nt ITS1 fragment in HeLa cells after depletion of XRN2 indicates that the 18SE precursor can be generated by an endonucleolytic cleavage at site 2a. Previously published data, however, shows that depletion of the ribosome biogenesis factor, ENP1, causes accumulation of 21SC, suggesting that 18SE could instead be generated by exonucleolytic processing of 21S (Carron et al, 2011). The majority of the 3'-5' exonuclease activity in the nucleolus is derived from RRP6, one of the active subunits of the exosome, making this a likely candidate for carrying out this

processing step. We therefore used RNAi to deplete XRN2, ENP1 or RRP6 from HeLa cells. Proteins were extracted from siRNA treated cells and Western blot analysis showed that each protein was specifically and significantly depleted by the siRNA treatment (Figure 4.3A). The levels of rRNA processing intermediates in these cells were then analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using probes hybridising to the 5' end (5520) or the middle of ITS1 (6121) (Figure 4.3C, D).

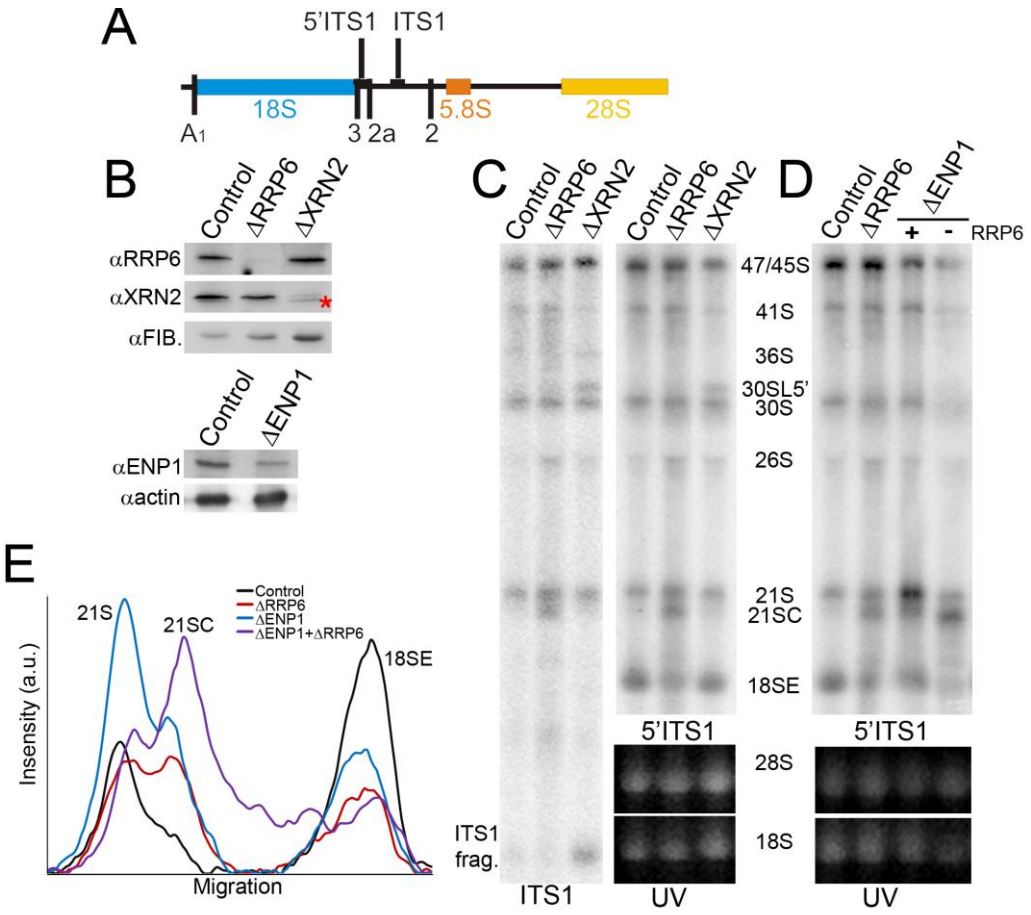


Figure 4.3 RRP6 and ENP1 are required for ITS1 processing A) Schematic outline of a section of the human pre-rRNA transcript with the positions of the probes used in Northern blot hybridisation, 5'ITS1 (5520) and ITS1 (6121) indicated relative to the mature rRNAs. **B)** HeLa cells were transfected with siRNA duplexes to deplete RRP6, XRN2, ENP1 or with a control siRNA targeting firefly luciferase mRNA. Cells were harvested after 60 h and proteins were extracted, separated by SDS-PAGE, followed by Western blotting using antibodies raised against the target proteins and the core snoRNP protein, fibrillarin, or the cytoskeletal protein, actin. The red asterisk indicates residual signal from previous probing of the membrane with an antibody against RRP6, which migrates slightly faster than XRN2. **C)** and **D)** RNA was extracted from siRNA treated cells and analysed by agarose-glyoxal gel electrophoresis where 28S and 18S mature rRNAs were visualised by ethidium bromide staining (UV). RNA was transferred to a nylon membrane and Northern blotting using indicated in (A). Pre-rRNAs were visualised using a phosphorimager and the positions of various pre-rRNA intermediates are indicated between the panels. + indicates RRP6 is present at normal levels while – denotes depletion of RRP6. **E)** The relative levels of pre-rRNAs from the lower portion of panel (D) were quantified using ImageQuant software and are shown as line graphs depicting intensity in arbitrary units (a.u.) compared to the distance migrated through the gel. Levels were normalised to 28S rRNA. The identity of each peak is indicated.

Compared to control cells, depletion of XRN2 led to an increase in 47S levels and the appearance of 30SL5', a 5' extended form of 30S, both of which indicate that XRN2 is required for efficient A' cleavage as discussed in section 3.11 (Figure 4.3C). Depletion of XRN2 also caused an increase in the levels of 26S, only detected in some experiments (see Figure 5.9), implying that the order of cleavage steps during 5'ETS removal is altered by XRN2 depletion as suggested in the previous chapter. Also observed after XRN2 depletion were significant accumulations of 36S and the 800nt ITS1 fragment, implying that an increased proportion of pre-rRNA is cleaved at site 2a prior to site 2 processing. Interestingly, 36S is barely detectable in control cells suggesting endonucleolytic processing at site 2a does not normally occur before site 2 cleavage. However, despite variations in the levels of several pre-rRNA species, the level of the final 18S precursor, 18SE, was not affected by XRN2 depletion suggesting that the order of cleavages is altered without changing the overall kinetics of processing.

Consistent with published data (Carron et al, 2011), depletion of ENP1 did not affect the levels of longer pre-rRNAs (47S, 45S or 30S) but led to a large increase in 21S levels, the appearance of 21SC, a 3' shortened form of 21S and a decrease in 18SE levels (Figure 4.3D, E). Similarly, depletion of RRP6 did not affect the early pre-rRNA processing steps but did cause an even more significant decrease in 18SE levels and the accumulation of 21SC. The appearance of this partially processed form of 21S was accompanied by a range of intermediate products seen as a smear between 21S and 18SE. This implies that loss of RRP6 causes inefficient exonucleolytic processing of 21S. Co-depletion of both ENP1 and RRP6 caused cumulative defects with a stronger accumulation of 21SC, a greater reduction of 21S levels and a significantly stronger decrease in the amount of 18SE than observed in either of the single knockdowns (Figure 4.3C).

Taken together, these data support the notion that 21S is converted to 18SE by exonucleolytic processing and show that the exosome component, RRP6, is required for this step. This pathway is a major route for 18SE production as depletion of RRP6 decreases 18SE levels significantly. Although depletion of XRN2 appears to promote an alternative endonucleolytic processing pathway, this does not change the overall rate of 18SE production, implying that the cell does not normally rely on this alternative endonculeolytic pathway to remove ITS1 sequence.

4.2.3 Metabolic labelling of rRNA shows that RRP6 is required for 18S production

Northern blot analysis (Figure 4.3C, D) clearly demonstrates that depletion of RRP6 inhibits formation of 18SE while although depletion of XRN2 changes the balance of cleavages used to generate 18SE, it does not affect the overall levels of 18SE. Our data therefore suggest that a major pathway by which 18S rRNA is produced requires RRP6. To determine if this is the case, metabolic labelling experiments were carried out. Metabolic labelling of cells following RNAi mediated depletion of a protein of interest enables RNA formed *de novo* to be analysed specifically making any defects caused by loss of the protein more readily detectable than by Northern blotting. HeLa cells were transfected with either a control siRNA duplex or those targeting RRP6 or XRN2 mRNAs. Cells were metabolically labelled using ^{32}P orthophosphate then grown in media containing unlabelled phosphate and harvested over a timecourse of 4 h following removal of the labelled media. Total RNA was separated by agarose-glyoxal or by denaturing polyacrylamide gel electrophoresis, transferred to a nylon membrane and RNA visualised using a phosphorimager (Figure 4.4).

In control cells, at the start of the chase (time 0) only 47/45S was detectable with the first significant appearance of the 32S intermediate after 60 min. The 28S and 18S mature rRNAs then accumulated at later time points. 41S, 30S and 21S were just detectable but due to their relatively low abundance were difficult to follow in these experiments. In cells depleted of either RRP6 or XRN2, after 3 h incubation with unlabelled media, 32S and 28S were accumulated to approximately 80% of the level detected in control cells. In cells depleted of XRN2 the levels of 18S rRNA was also approximately 80% of those found in control cells but in contrast, 18S rRNA levels in cells depleted of RRP6 were decreased to approximately 30% of normal levels (Figure 4.4). This suggests that the levels of 18S and 28S mature rRNAs is affected by depletion of RRP6 or XRN2 (see below for explanation) but that that RRP6 is specifically required for the efficient production of 18S in human cells. We conclude that the majority of the SSU rRNA is produced by exonucleolytic processing of ITS1 following site 2 cleavage.

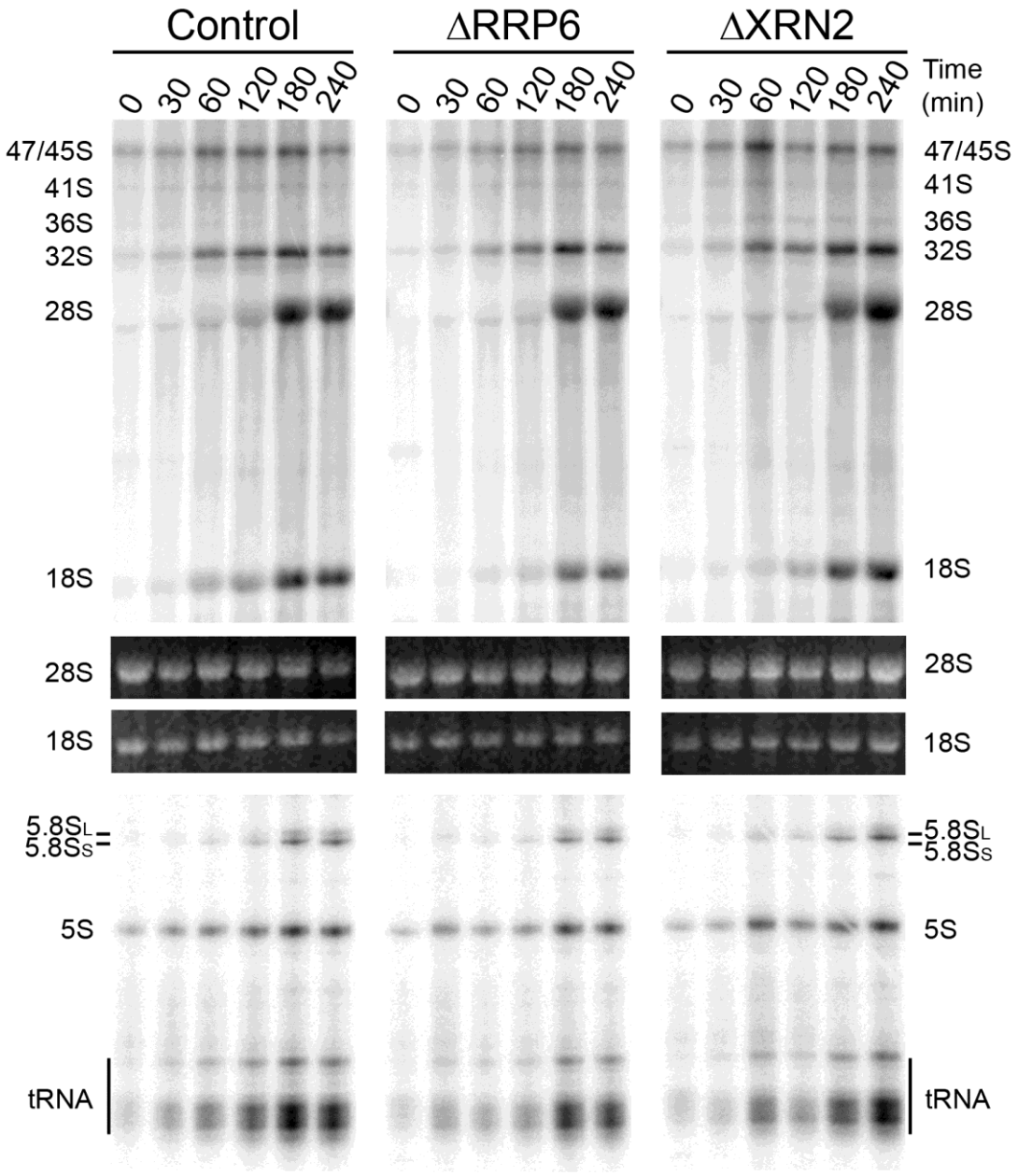


Figure 4.4 RRP6 is required for 18S production HeLa cells were transfected with siRNA duplexes to deplete RRP6 or XRN2 or with control siRNAs. 48 h after transfection, the cells were depleted of phosphate and pulse labelled with ^{32}P orthophosphate. The labelled media was then replaced with normal media for varying lengths of time (indicated above each lane). Cells were harvested and RNA was extracted. This was separated using agarose-gloxal gel electrophoresis (upper panel) or on an 8 % acrylamide/7 M urea gel (lower panel) and RNAs were visualised using a phosphorimager. The total RNA loaded on the agarose-glyoxal gel was visualised by ethidium bromide staining (UV) and the acrylamide gel loaded correspondingly.

Depletion of XRN2 caused a parallel decrease in the levels of both 18S and 28S mature rRNAs and depleting RRP6 affected 28S levels to a similar extent. While it is possible that this represents an equal requirement for these exonucleases in 28S and 18S production, it also raises the possibility that depletion of any essential protein, not only those directly involved in ribosome biogenesis, would affect the growth rate of the HeLa cells and therefore slow the accumulation of mature rRNAs. To test this, RNAi was used to deplete a protein which is known to be essential for cell viability but

that does not participate in formation of ribosomes. The signal recognition particle (SRP) is an RNA-protein complex which is important during translation but is not associated with pre-ribosomes (Grudnik et al, 2009). One of the core subunits of this complex, SRP14, was depleted from HeLa cells (Figure 4.5A) before metabolic labelling and RNA analysis as described above. After depletion of SRP14, the mature rRNAs accumulated at a slower rate than in control cells and after 3 h, 28S and 18S levels were ~80% of those observed in control cells (Figure 4.5B). This implies that depletion of non-ribosome biogenesis factors which have other essential functions in the cell, decreases the rate of pre-rRNA processing by approximately 20%. For a protein factor to be considered critical for the biogenesis of either subunit, depletion of the protein must therefore, cause a greater than 20% decrease in mature rRNA levels or alter the ratio between 28S and 18S levels as is the case when RRP6 is depleted. As the SRP complex is involved in translation, it is possible that the decreased rate of ribosome biogenesis reflects negative feedback from translation inhibition, but this is unlikely.

In these experiments it also appeared that 18S accumulated earlier than 28S with 18S first detectable after 60 min incubation with unlabelled media and 28S not detectable until 3 h after the change of media. This was surprising as the rate of production of both subunits was expected to be the same. The immediate precursor of 18S, 18SE, only extends beyond the 3' end of 18S by approximately 25 nt and cannot be resolved separately from 18S on the agarose glyoxal gel used. It was therefore likely that the 18S signal detected after 60 and 120 min incubation with unlabelled media actually represented 18SE precursor rRNA. To test this possibility, an antisense oligonucleotide primer recognising the sequence 200 nt upstream of the 3' end of 18S was annealed to labelled RNA from control cells and samples were treated with RNaseH to cleave the pre-rRNA specifically at the site of the annealed primer. This generated an approximately 200 nt fragment corresponding to the 3' end of 18S and an ~225 nt RNA derived from the 18SE precursor. RNA was separated on an 8% denaturing acylamide gel that was dried and labelled RNA was visualised using a phosphorimager (Figure 4.5C). 30 and 60 min after addition of unlabelled media, the longer fragment corresponding to18SE was detectable at 2-3 fold higher levels than the short fragments derived from 18S, but after 120 min, levels of the shorter fragment exceeded those of the longer. This shows that, when analysing full length pre-rRNAs by agarose-glyoxal gel electrophoresis, the majority of the signal detected 30 and 60 min after incubating with unlabelled media is 18SE, but at later time points (120-240 min) the majority of the signal is derived from accumulated 18S (Figure 4.6D).

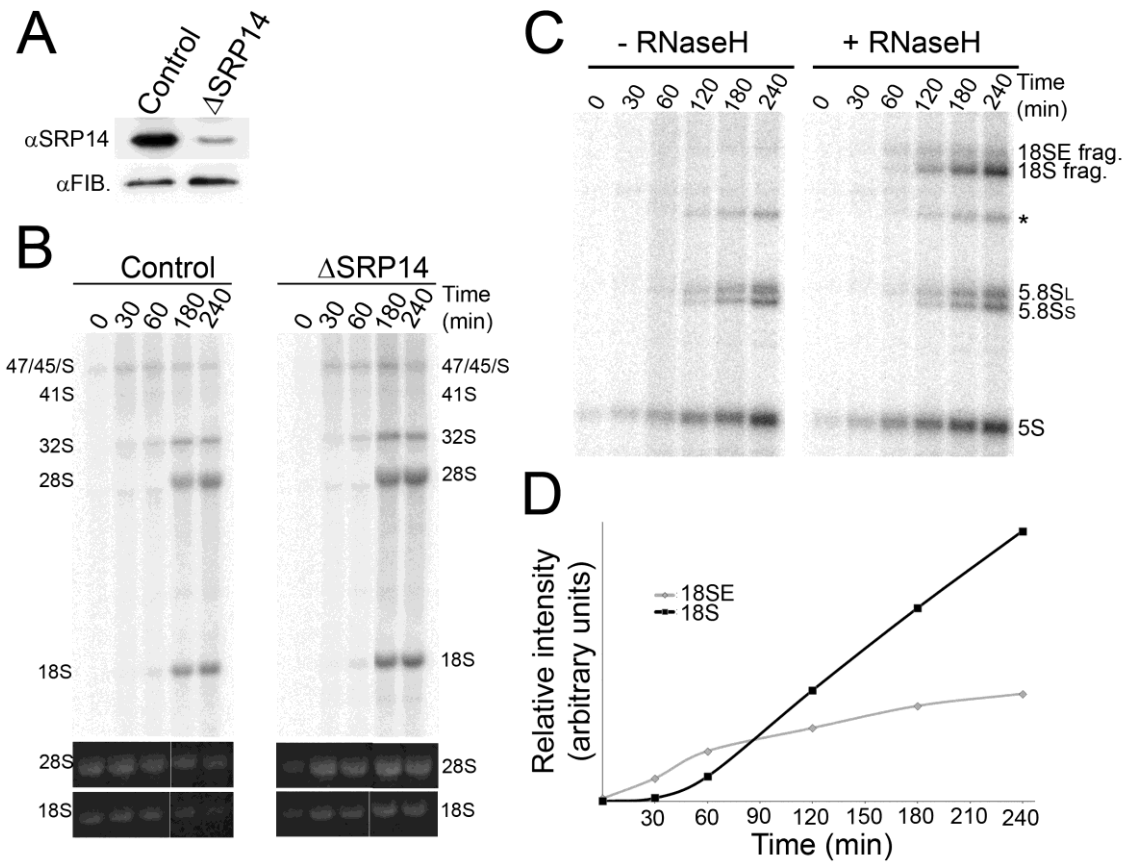


Figure 4.5 Depletion of essential proteins not involved in ribosome biogenesis decreases the rate of mature rRNA synthesis A) HeLa cells were transfected with siRNA duplexes targeting the mRNA of the SRP complex protein, SRP14, or with control siRNAs. After 48 hours cells were harvested and proteins analysed by SDS-PAGE followed by Western blotting using antibodies detecting SRP14 or the snoRNP protein, fibrillarin (FIB.). **B)** Cells were depleted of SRP14 as in (A) and after 48 h, cells were depleted of phosphate, labelled with ^{32}P orthophosphate and then grown in normal media. Cells were harvested at the time points shown above the panel Total RNA was extracted and analysed by agarose-glyoxal gel electrophoresis, transferred to a nylon membrane and visualised using a phosphorimager. **C and D)** HeLa cells were pulse labelled as in (B) but cells were harvested at various time points after addition of normal media (shown above the panel). RNA was extracted and a primer annealing approximately 200 nt upstream of the 3' end of 18S was annealed. Equal amounts of RNA were treated with RNase H and separated on an 8 % polyacrylamide/7 M urea gel. RNA was visualised using a phosphorimager and the levels of the fragments corresponding to 18S and 18SE in RNase H treated samples were quantified using ImageQuant software.

4.2.4 The exonuclease activity of RRP6 is required for processing of 21S to 18SE

Since RRP6 is important for processing to form 18SE and overall 18S production we next wanted to verify whether the exonuclease activity of RRP6 was responsible for the conversion of 21S into 18SE. To determine if the exonuclease activity of RRP6 was required for 18SE production, stably transfected HEK293 cell lines in which endogenous RRP6 could be replaced with either FLAG-tagged wild-type RRP6 or an inactive form of RRP6 (RRP6exo) were used as described in Section 3.2.9, Figure 3.11. Mutation of an amino acid, D313 in the catalytic site abolishes the exonuclease activity of RRP6. Additional mutations were introduced into the open

reading frame (ORF) of both the wild-type and the inactive proteins to enable RNAi resistant expression of these proteins without changing the coding potential

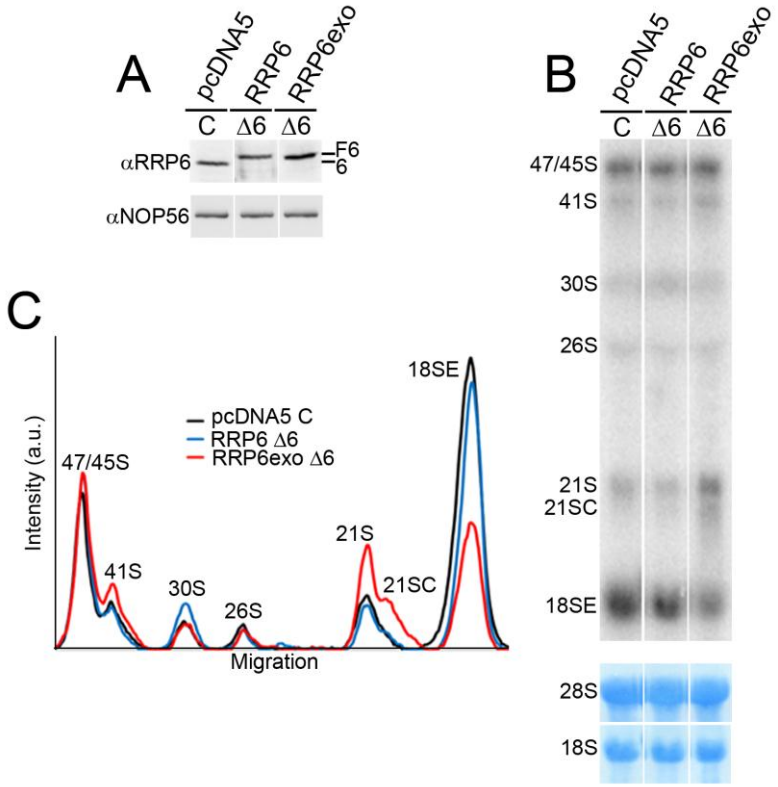


Figure 4.6 The exonuclease activity of RRP6 is required for 18SE production A) HEK293 cells stably transfected with plasmids expressing FLAG-tagged RRP6, RRP6 with a mutation changing D313A in its active site (RRP6exo) or just the FLAG-tag were transfected with either a control siRNA duplex (C) or siRNAs to deplete the endogenous RRP6. The open reading frames of RRP6 in the plasmids had been altered so expression of the tagged forms of RRP6 was resistant to the siRNA treatment. The levels of endogenous and FLAG-tagged proteins were monitored by Western blotting using antibodies against RRP6 or the core snoRNP protein NOP56 as a loading control. **B)** Pre-rRNA was analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using a probe hybridising to the 5' end of ITS1. The levels of mature rRNAs, 28S and 18S were visualised by methylene blue staining (lower panels). The positions of intermediates detected are given to the left of the panel. **C)** The levels of pre-rRNAs from (B) were quantified using ImageQuant software and are given as a line graph showing the intensity in arbitrary units (a.u.) relative to migration through the gel. The identity of each peak is indicated.

Cells expressing FLAG-RRP6, FLAG-RRP6exo or just the FLAG-tag (pcDNA5) were chemically transfected with siRNA duplexes to deplete RRP6 (Δ6) or a control siRNA targeting firefly luciferase mRNA (C). After 48 h, cells were harvested and protein levels were analysed by Western blotting. In cells expressing FLAG-tagged forms of RRP6, the endogenous protein was no longer detectable after siRNA treatment while the longer, tagged forms of the protein were present at the same levels as the endogenous protein in the control cells (Figure 4.6A). RNA was extracted from these cells and was analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using a probe hybridising to the 5' end of ITS1 (Figure 4.6B). In control cells and those expressing FLAG-tagged wild-type RRP6, the levels of all the

pre-rRNA intermediates detected were comparable to those in control cells. However, when endogenous RRP6 was replaced with inactive RRP6, both 21S and 21SC accumulated and the level of 18SE was significantly reduced compared to control cells. These findings clearly demonstrate that the exonuclease activity of RRP6 is the major activity required for processing of 21S to produce 18SE.

4.2.5 The core exosome and exosome cofactors are required for 18SE production

RRP6 usually functions as a component of a multi-protein complex, the exosome. We therefore investigated the roles of the core exosome component, RRP46, the other active subunit of the nuclear exosome, DIS3, and the RRP6 cofactors, C1D, MPP6 and MTR4, in the production of 18SE. Each protein was depleted from HeLa cells using RNAi as described in the previous chapter (Figure 3.7). The effects of these depletions on pre-rRNA processing were analysed by agarose-glyoxal gel electrophoresis, followed by Northern blotting using a probe hybridising to the 5' end of ITS1. Here we focus on the levels of 21S, 21SC and 18SE, as the only significant effect on the longer pre-rRNAs was an accumulation of 30SL5' following MTR4 depletion, which has been discussed in the previous chapter (section 3.11).

Depletion of the core exosome subunit, RRP46, or the TRAMP component, MTR4, led to a decrease in 18SE levels and a strong smear of intermediate products between 21S and 18SE (Figure 4.7A, B). The intensity of this smear of intermediates was significantly higher in these knockdowns than when RRP6 was depleted, but 21SC did not specifically accumulate above this background. Depletion of the RRP6 cofactor, MPP6, led to increased 21S levels and a weak accumulation of 21SC but no overall decrease in 18SE levels suggesting that while MPP6 is important for efficient 18SE production, it is not absolutely required (Figure 4.7A, B). Depletion of C1D or DIS3 had no noticeable effect on 21S, 21SC or 18SE levels implying that these proteins are not required for this step (Figure 4.7A, B). Conversion of 21S into 18SE is likely to occur in the nucleolus and since DIS3 is excluded from this compartment in human cells, it is not surprising that this nuclease does not participate in the production of 18SE. In the previous chapter we demonstrated that both DIS3 and C1D are required for 5.8S processing and C1D functions in the degradation of 37S* confirming that the decreases in protein levels caused by RNAi treatment are sufficient to induce pre-rRNA processing defects. These data therefore show that the core exosome, RRP6, MTR4 and to a lesser extent, MPP6, but not DIS3 or C1D are important for 18SE production.

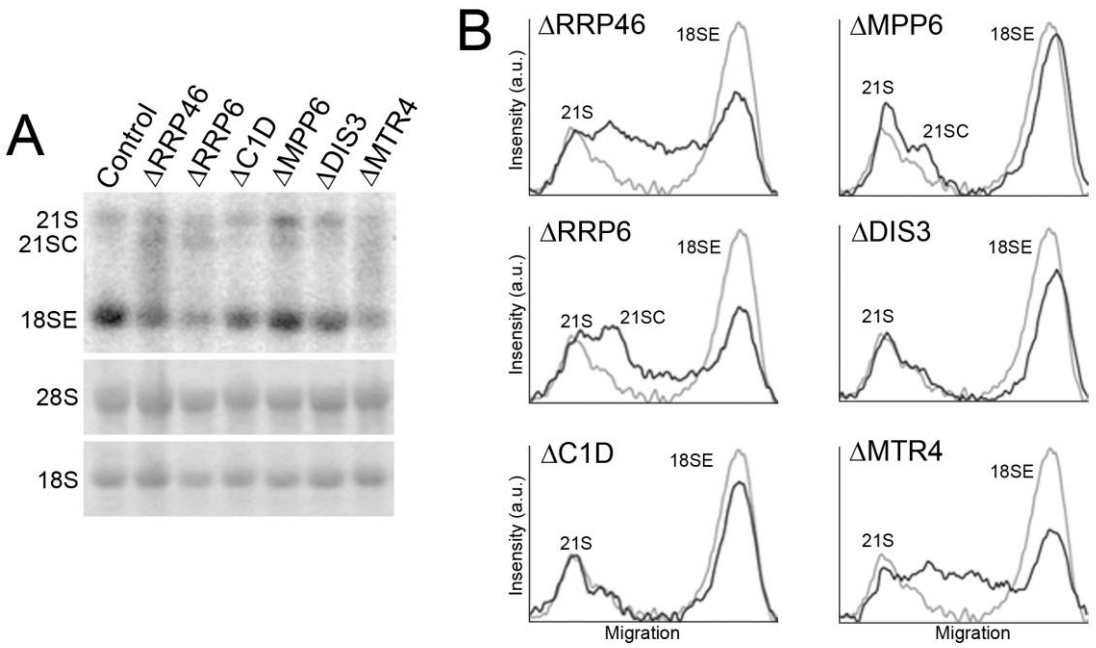


Figure 4.7 The core exosome, MTR4 and MPP6 are required for 18SE production A) HeLa cells were transfected with siRNA duplexes targeting the mRNAs of the core exosome subunit, RRP46, the active nuclear exosome subunits, RRP6 or DIS3 and the cofactors, MPP6, C1D or MTR4. After 60 h, cells were harvested and RNA was extracted. RNA was analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using a probe hybridising to the 5' end of ITS1. Mature rRNAs were visualised by methylene blue staining and are indicated to the left of the panels. **B)** The levels of the pre-rRNA intermediates shown in (A) were quantified using ImageQuant software and were normalised to 28S mature rRNA levels. For each protein depletion, this is shown as line diagrams depicting intensity in arbitrary units (a.u.) versus the distance migrated through the gel. Black lines correspond to pre-rRNAs from control cells and grey lines indicate the levels of pre-rRNAs in cells depleted of the named exosome protein.

4.2.6 The 3' end of 21SC maps to the base of a long stem-loop structure

The heterogeneous 3' ends of 21SC accumulated in cells depleted of ENP1 have been mapped by 3' RACE and terminate between nucleotides 6162-6177 of the human pre-rRNA sequence (Carron et al, 2011). When RRP6 was depleted, the partially processed intermediate, 21SC, accumulated and Northern blot mapping of the 3' end of this intermediate corresponded with the 3' end mapped by Carron et al. (2010) following ENP1 depletion (data not shown). The appearance of a partially processed intermediate when RRP6 is depleted may imply that other 3'-5' exonucleases are also able to initiate ITS1 processing following site 2 cleavage but the decrease in 18SE formation suggests that only the exosome is able to process efficiently beyond the 3' end of 21SC. As 21SC can also be detected at very low levels in control cells, this suggests that the 3' end of 21SC may represent a natural stalling point for the exosome. To investigate what may interrupt exonucleolytic processing at this particular site, the secondary structure of ITS1 was modelled using software available at http://swissmodel.expasy.org and is shown in Figure 4.8. In this model, the

3’ end of 21SC is found at the bottom of a long stem-loop structure. Although we have shown that human RRP6 is able to degrade an RNA with a 5 bp terminal stem *in vitro* (section 3.2.2.2), a stem-loop of this magnitude is still likely to represent a considerable challenge for the exosome. The secondary structure of ITS1 in this region is, therefore, probably responsible for interrupting exonucleolytic processing leading to formation of 21SC.

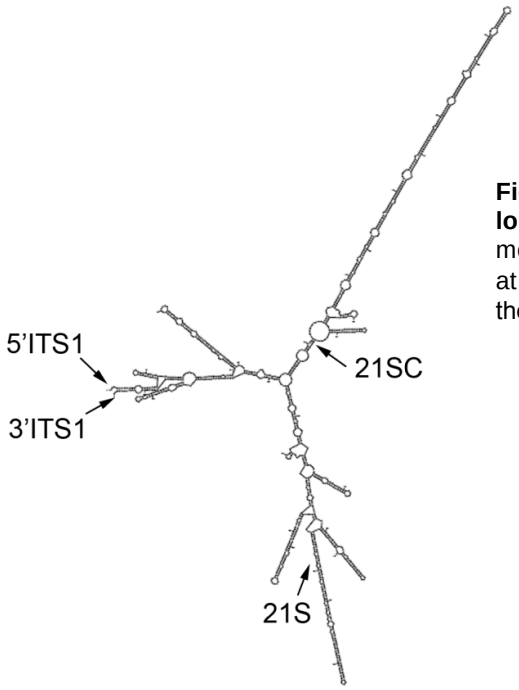


Figure 4.8 The 3’ end of 21SC is a the base of the long stem structure in ITS1 A secondary structure model of ITS1 was generated using software available at http://swissmodel.expasy.org and the positions of the 3’ ends of ITS1, 21S and 21SC are highlighted.

4.2.7 Endonucleolytic cleavage at site 2a provides an exosome independent mechanism for 18SE production

The major pathway for 18SE production involves exonucleolytic processing of 21S by the exosome and RRP6. Depletion of XRN2, however, revealed that 18SE can also be produced by an endonucleolytic cleavage at 2a. In cells depleted of different exosome components, we did not see increased evidence of endonucleolytic cleavage at 2a, characterised by accumulation of 36S and the ITS1 fragment, implying that these pathways are not naturally redundant (Figure 4.9A). However, we investigated whether stimulating endonucleolytic cleavage at 2a by depleting XRN2 could bypass the requirement for the exosome and MTR4 in 18SE production. The effect of depleting RRP46 or MTR4 individually or in combination with XRN2 on 18SE levels was assessed using Northern blotting (Figure 4.9A). The accumulation of 30SL5’ following depletion of MTR4 has been discussed in the previous chapter. As described above, 18SE levels were not affected by XRN2 depletion whereas depletion of either RRP46 or MTR4 led to a significant decrease in 18SE levels. However, 18SE levels in cells

depleted of either RRP46 or MTR4 could be restored to within 10 % of normal levels by co-depletion of XRN2 (Figure 4.9B). This underlines the observation that the majority of 18SE is normally produced by exonucleolytic processing by the exosome but this dependence on the exosome can be bypassed by depletion of XRN2 which stimulates an alternative pathway through endonucleolytic cleavage at 2a.

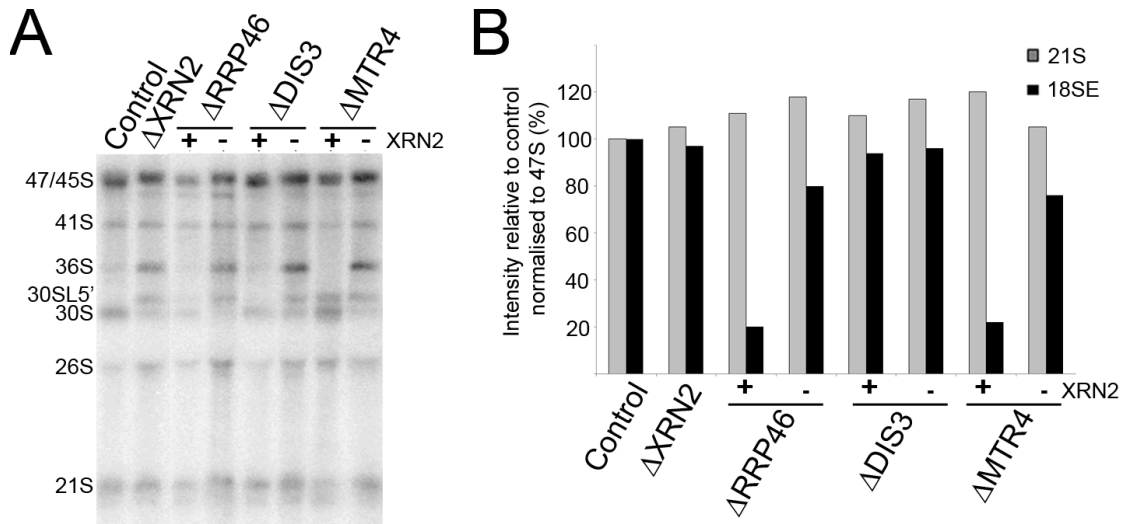


Figure 4.9 Exosome independent production of 18SE A) HeLa cells were transfected with siRNAs duplexes to deplete RRP46, DIS3 or MTR4 either alone or together with XRN2 or with control siRNAs. + denotes pre-rRNA from cells in which XRN2 is present at normal levels and – indicates depletion of XRN2. After 60 h, RNA was extracted and analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using probes hybridising to the middle of ITS1 (top panel) or the 5' end of ITS1 (middle panel) and mature rRNAs (28S and 18S) were visualised using methylene blue staining. Mature and pre-rRNA intermediates detected are indicated to the left of the panel. **B)** 21S and 18SE levels detected with the 5'ITS1 probe were quantified using ImageQuant software. Levels were normalised to mature 28S levels and are shown relative to levels in control cells as a bar chart.

4.2.8 Site 2a in human pre-rRNA is analogous to A_2 in yeast

Our data identify two endonuclease cleavage sites in ITS1; a primary cleavage site at 2 which separates the LSU and SSU rRNAs and a secondary cleavage site at 2a. In yeast, cleavages at A_2 and A_3 separate the LSU and SSU rRNAs but A_3 cleavage is a non-essential step. It has, however, been proposed that site 2a in human pre-rRNA is analogous to the yeast A_2 site. This raises the possibility that this cleavage, equivalent to yeast A_2, only plays a minor or possibly redundant role in human pre-rRNA processing. In yeast, cleavage at A_2 is dependent on many components of the SSU processome and has been suggested to be carried out by the Rcl1 endonuclease

(Horn et al, 2011). We therefore used RNAi to deplete key SSU processome proteins and 18S processing factors, fibrillarin (core snoRNP), UTP24 (putative endonuclease), RCL1 (A_2 endonuclease in yeast), BMS1 (GTPase), RRP5 (required for A_2 and A_3 in yeast), NOB1 (site D endonuclease in yeast), ENP1 and the r-protein, RPS19 (involved in ITS1 exonucleolytic processing), from HeLa cells. This was followed by Northern blotting or metabolic labelling to determine whether these factors were required for either of the cleavages in ITS1. It is difficult to identify factors required for the 2a endonucleolytic cleavage since 36S is the only intermediate specifically generated by this cleavage and it is barely detectable in human cells. However, depletion of XRN2 stimulates the 2a endonucleolytic pathway so each knockdown was carried out both individually and in combination with XRN2 to enable cleavage at 2a to be monitored.

Western blot analysis of proteins from siRNA treated cells showed that each protein of interest had been depleted significantly compared to control cells while levels of other proteins (CSL4, RIO2 or actin) were unaffected by the RNAi treatment (Figure 4.10A). RNA was extracted from siRNA treated cells and analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using probes hybridising to either the 5' end of ITS1 or the middle of ITS1. Depletion of fibrillarin, UTP24 or BMS1 led to increased accumulation of 30S compared to control cells and depletion of fibrillarin also caused accumulation of 30SL5', a 5' extended form of 30S that occurs when A' cleavage is impaired (as discussed in Chapter 3). In all of these knockdowns, 41S, 21S and 18SE precursors could not be detected. This indicates that these core SSU processome proteins are all required for complete removal of the 5'ETS, but since 30S was detected, implies that they are not required for ITS1 cleavage at site 2. Co-depletion of XRN2 with either fibrillarin, UTP24 or BMS1 resulted in the complete absence of 36S and the ITS1 fragment indicating that these proteins are essential for endonucleolytic cleavage at site 2a (Figure 4.10B). Depeletion of RCL1 resulted in increased levels of 26S and 30S implying that RCL1 is required for cleavage at site A_1 in the 5'ETS but is not needed for site 2 cleavage. However, depletion of RCL1 caused a significant decrease in 18SE levels suggesting that it is important for exonucleolytic processing following site 2 cleavage (Figure 4.10B). Similarly, depletion of either ENP1 or RPS19 caused a significant decrease in 18SE levels and in these cases, the levels of 21S and 21SC were dramatically increased. Following depletion of RPS19, 30SL5' was detectable suggesting this protein is also required for A' cleavage (Figure 4.10B and Figure 4.11). Co-depletion of RPS19 with XRN2 abolished accumulation of both 36S and the ITS1 fragment whereas co-depletion of either ENP1 or RCL1 with XRN2 merely decreased the levels of these pre-rRNAs (Figure 4.10B) suggesting all three of these proteins are required for cleavage at site 2a. It is not valid to compare the relative

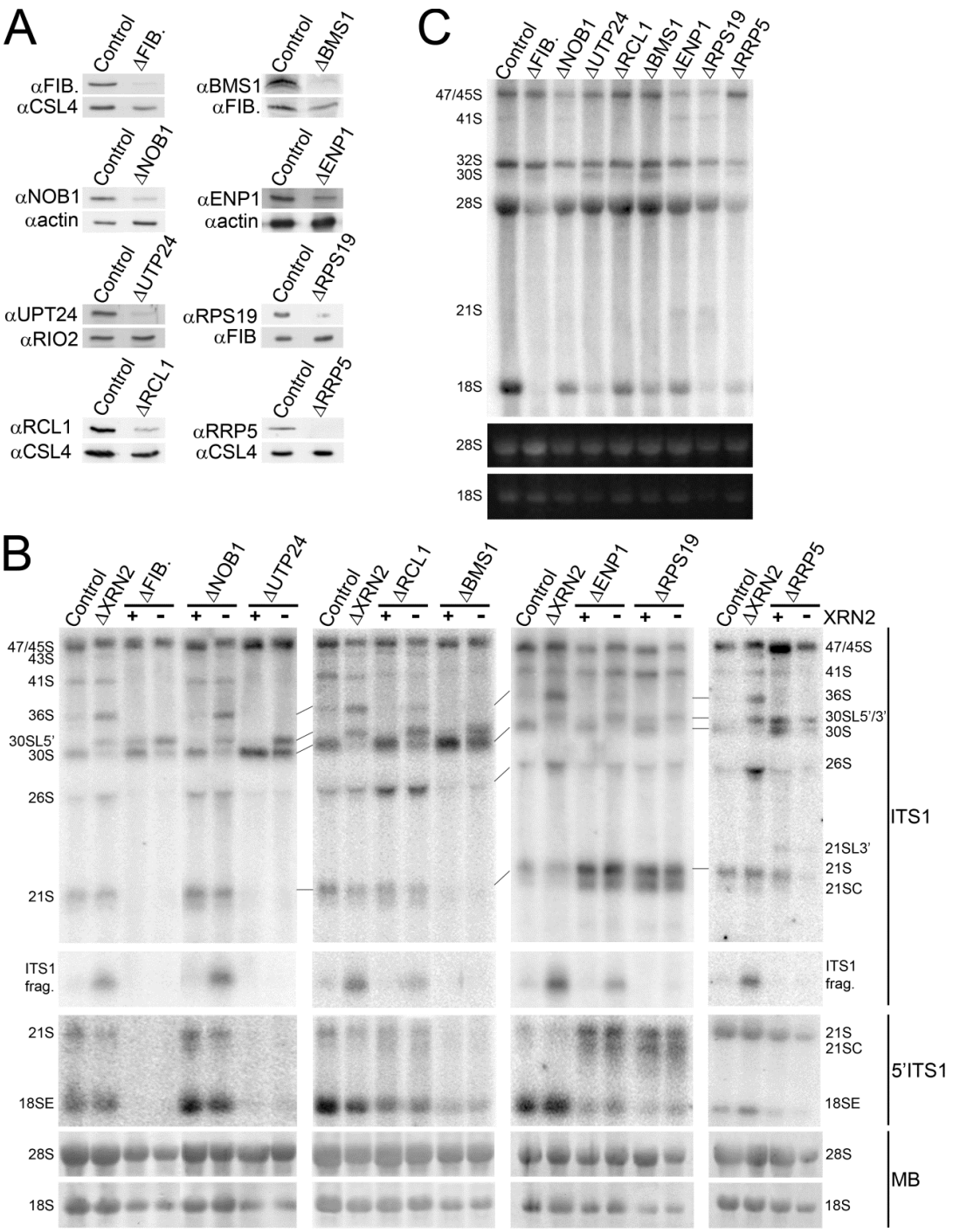


Figure 4.10 SSU processome components are required for endonuclease cleavage at site 2a **A)** and **B)** HeLa cells were transfected with siRNA duplexes targeting the mRNAs of various ribosome biogenesis factors (indicated above the panels) either individually or in combination with those targeting XRN2. + denotes cells in which XRN2 is present at normal levels and – indicates cells in which XRN2 has been depleted. After 60 h cells were harvested and RNA and proteins were extracted. **A)** Protein samples from individual knockdowns were analysed by Western blotting using antibodies against the targeted proteins and CSL4, RIO2, fibrillarin or actin as loading controls. **B)** RNA was separated by agarose-glyoxal electrophoresis followed by Northern blotting using probes hybridising to the middle of ITS1 (top and second panel), the 5' end of ITS1 (third panel). Mature rRNAs 28S and 18S were visualised using methylene blue staining. Mature and pre-rRNAs are indicated to the right and left of the panel and fine lines link the identical pre-rRNAs on the middle sections. **C)** HeLa cells depleted of ribosome biogenesis factors for 48 h were depleted of phosphate and pulse labelled with ^{32}P orthophosphate. Cells were then incubated with normal media for 3 h before harvesting. RNA was extracted and separated using agarose-glyoxal gel electrophoresis and visualised using a phosphorimager. The total RNA was visualised by ehtidium bromide staining (UV)

dependence of 2a on each of these proteins as they were depleted to different extents. Depletion of NOB1 caused no changes to the levels of longer pre-rRNAs either individually or when co-depleted with XRN2 but did result in increased accumulation of 18SE. This suggest that its role as the endonuclease responsible for cleavage at the 3' end of 18S is conserved from yeast to humans and that it is not required for cleavage at either 2 or 2a (Figure 4.10B). Consistent with these data, metabolic labelling experiments showed that depletion of fibrillarin, UTP24, NOB1, BMS1, RCL1, ENP1 or RPS19 resulted in a decrease or the complete loss of 18S rRNA. The increases in 30S levels following depletion of fibrillarin, UTP24, RCL1 and BMS1 and those of 21S following depletion of either ENP1 or RPS19 observed by Northern blotting were also detectable (**Figure 4.10C**). Depletion of fibrillarin also caused a decrease in 28S levels as previously discussed (**Section 3.2.12**).

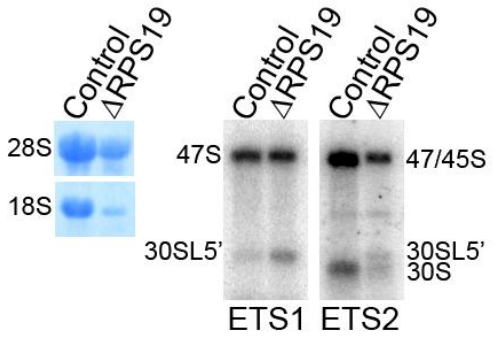


Figure 4.11 RPS19 is required for efficient A' cleavage RNA from cells depleted of RPS19 and control cells was analysed using agarose-glyoxal gel electrophoresis followed by Northern blotting using probes hybridising upstream (ETS1) and downstream (ETS2) of the A' cleavage site in the 5'ETS. Methylene blue staining was used to visualise mature 28S and 18S rRNAs and the positions of both mature rRNAs and pre-rRNAs are indicated to the left and right of the panels.

In yeast, deletion of Rrp5 impairs pre-rRNA processing at sites A_0, A_1, A_2 and A_3 implicating Rrp5 in rRNA processing steps required for formation of both the small and large subunit rRNAs (Venema et al. 1996). Metabolic labelling showed that when cells were depleted of RRP5, 45S levels were radically increased and that production of both 18S and 28S mature rRNAs was significantly inhibited (Figure 4.10C). Northern blot analysis of RNA from cells depleted of RRP5 also showed a major increase in 45S levels and a decrease in the accumulation of 30S and 18SE (Figure 4.10A and Figure 4.12). The decrease in 18SE levels observed was not mirrored by a decrease in 21S levels implying that RRP5 functions in exonucleolytic processing from site 2 to 2a. Northern blotting analysis also revealed accumulation of novel fragments above 30S and 21S which have been termed 30SL3' and 21SL3' respectively. Using agarose-glyoxal gel electrophoresis, the 30SL3' intermediate appears the same size as both 30SL5', the 5' extended form of 30S which accumulates when A' cleavage is affected, and the LSU rRNA precursor, 32S. 30SL3' does not, however, contain the 5' end of the 5'ETS implying that it is a 3' extended form of 30S but this could not be confirmed by Northern blotting because of the overlap with 32S. 21SL3' was also detectable using probes hybridising to the 3' end of ITS1 and 5.8S but not the 5' end of ITS2 (Figure

4.12). By comparison, we infer that the 3' end of 30SL3' was also likely to contain 3'ITS1 and 5.8S sequence. It was concluded that depletion of RRP5 caused accumulation of some pre-rRNAs in which site 2 cleavage had been bypassed and processing instead occurred in ITS2. It is not clear where the cleavage site in ITS2 is and it is likely that it is followed by 3'-5' exonuclease processing to degrade the aberrant precursors. Co-depletion of RRP5 with XRN2 inhibited production of both 36S and the ITS1 fragment indicating RRP5 is also required for 2a cleavage.

Taken together these data imply that the 2a site in human pre-rRNA is analogous to yeast A_2 since cleavage at this site is dependent on several proteins of the SSU processome, including the yeast A_2 endonuclease, Rcl1. Similar to yeast, where Rrp5 is required for cleavage at A_2 and A_3, depletion of RRP5 inhibited both ITS1 cleavages (2 and 2a). This raises the possibility that site 2 in human pre-rRNA may be equivalent to the A_3 cleavage in yeast.

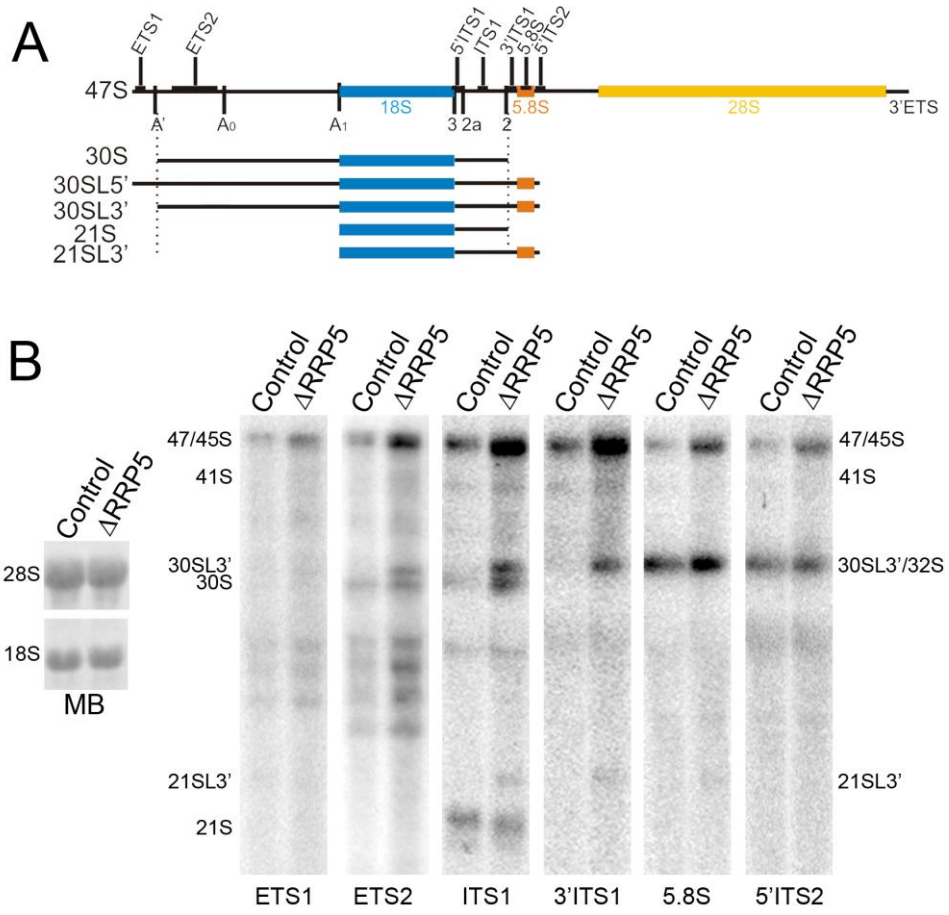


Figure 4.12 Mapping 30SL3' and 21SL3' accumulated after RRP5 depletion A) Schematic outline of partial pre-rRNA transcript showing the relative positions of the probes used in Northern blot hybridisation in green. **B)** RNA from control cells and those depleted of RRP5 was separated using agarose-glyoxal gel electrophoresis. Northern blotting was carried out using probes hybridising to the pre-rRNA regions shown in (A). pre-rRNA intermediates detected are indicated to both the left and right of the panel. Northern blot membranes were stained with methylene blue (MB) to visual mature 28S and 18S rRNAs.

4.2.9 Depletion of BOP1 and to some extent RBM28 impairs site 2 cleavage

In yeast, Nop4, which is homologous to human RBM28, is essential for efficient A_3 cleavage. Following A_3 cleavage in yeast, exonucleolytic processing is used to generate the mature 5' end of 5.8S processing and this requires a cluster of A_3 factors including Erb1 (BOP1 in higher eukaryotes). BOP1 and RBM28 were depleted from HeLa cells, either individually or in combination with XRN2, and the effects of pre-rRNA processing were examined using Northern blotting (Figure 4.13A-C) and metabolic labelling (Figure 4.13E, F). Depletion of BOP1 resulted in a significant decrease in 28S levels and accumulation of 47/45S and 41S without affecting 18S levels (Figure 4.13F). Also observed was a defect in 5.8S accumulation with loss of the normally major, short form of $5.8S_S$ and an increased accumulation of $5.8S_L$ (Figure 4.13E) implying that the pathways used to produce 5.8S following site 2 cleavage are analogous to those used in yeast after A_3 cleavage. Northern blotting of RNA from cells depleted of BOP1 showed increased levels of 47/45S and 41S as well as a significant accumulation 36S and a novel 5' shortened form of 36S, 36SC (Figure 4.13A, B and Figure 4.14). The levels of 32S, 30S, 21S and 12S were notably reduced while 18SE levels remained largely unaffected (Figure 4.14A, B). These data imply that loss of BOP1 impairs site 2 cleavage and co-depletion of BOP1 with XRN2 decreases the amount of the ITS1 fragment accumulated, reflecting this block in site 2 cleavage (Figure 4.13 A, C). Although depletion of BOP1 impairs site 2 cleavage, the levels of 18SE were not altered because an increased proportion of pre-rRNA was cleaved at 2a in the endonucleolytic processing pathway. The appearance of 36SC implies that pre-rRNA cleaved at 2a rather than site 2, can undergo 5'-3' exonucleolytic processing to generate 32S. Co-depletion of BOP1 with XRN2 causes a greater accumulation of 36S but does not change the level of 36SC implying that XRN2 is one of the 5'-3' exonucleases involved in generating this intermediate.

In contrast, depletion of RBM28 caused mild defects in the levels of 28S and 5.8S but the ratio of $5.8S_S$ to $5.8S_L$ and 18S levels were not altered. Metabolic labelling experiments revealed a significant increase in 47/45S levels which was also detectable using Northern blotting implying that all the pre-rRNA processing steps downstream of A' cleavage are slowed when RBM28 is depleted. Also observed using Northern blotting were increases in 41S, 36S and 36SC levels while 32S and 12S levels were slightly decreased. When RBM28 was co-depleted with XRN2, 36S was accumulated to a much greater extent than when XRN2 alone was depleted. These data show that when RBM28, the homologue of Nop4 which is essential for A_3 cleavage, is depleted, ITS1 processing is slowed and an increased proportion of pre-rRNA is cleaved at 2a before cleavage site 2. We therefore conclude that site 2 in human ITS1 is analogous

to yeast A_3 cleavage. This cleavage, however, plays a more important role in pre-rRNA processing in humans than in yeast as it is the endonuclease cleavage that separates the large and small subunit rRNA and our data also suggest that the relative importance of protein factors such as RBM28 (Nop4) in site 2 and A_3 cleavages may differ.

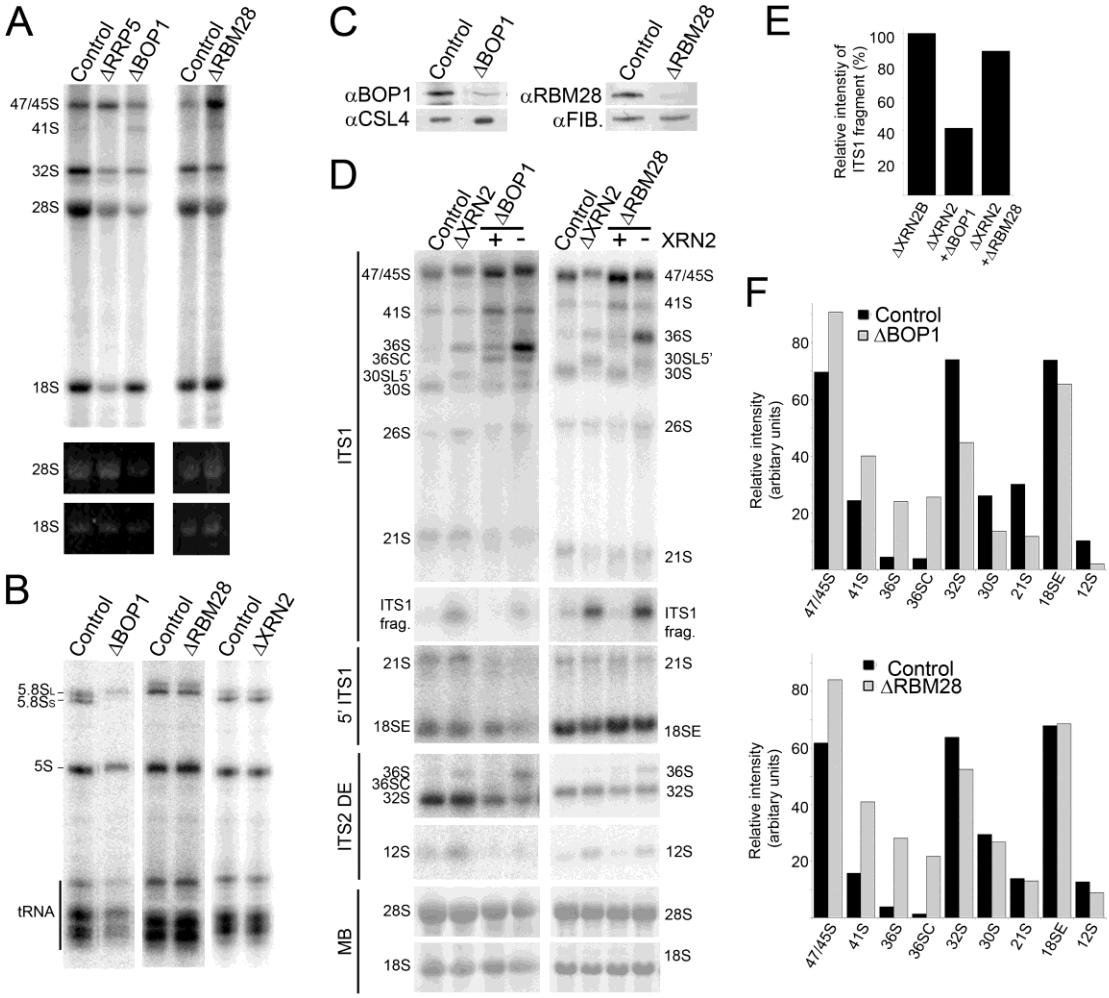


Figure 4.13 BOP1 and RBM28 are required for site 2 cleavage in ITS1 A and B) HeLa cells were depleted of RRP5, BOP1, RBM28 or XRN2 and after 48 h the cells were pulse labelled with ^{32}P orthophosphate then grown in normal media for 3 h. RNA was extracted and separated using agarose glyoxal (A) or 8 % polyacrylamide/7 M urea gel electrophoresis (B) and was visualised using a phosphorimager. Total RNA was visualised using ethidium bromide staining (UV). **C-F)** HeLa cells were transfected with siRNA duplexes to depleted BOP1 or RBM28 either individually or in combination with XRN2 or with control siRNAs. – indicates depletion of XRN2 and + indicates normal XRN2 levels. After 60 h RNA and proteins were extracted. **C)** Proteins were separated by SDS-PAGE and analysed by Western blotting using antibodies raised against mouse BOP1, RBM28 or those recognising CSL4 or fibrillarin (FIB.) as loading controls **D)** RNA was analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using probes hybridising to the 5' end of ITS1, middle of ITS1 or ITS2, as indicated to the left of the panel. Mature 28S and 18S rRNAs were visualised by methylene blue staining of membranes. **E-F)** Levels of pre-rRNA intermediates were determined using ImageQuant software, normalised to mature 18S levels, compared to either control cells (F) or those depleted of only XRN2 (E) and are shown as bar charts.

4.2.10 Human RNase MRP is not required for site 2 cleavage

In yeast, endonucleolytic cleavage at A_3 is a non-essential step that is carried out by the RNP, RNase MRP. The only pre-rRNA processing defect detected when RNase MRP, the enzyme responsible for carrying out this cleavage, is deleted, is a change in the ratio of the long and short forms of 5.8S rRNA. Since the A_3-like cleavage at site 2 in human pre-rRNA is the primary ITS1 cleavage which separates the LSU and SSU rRNAs, RNase MRP would be predicted to be critical for formation of 21S and 32S pre-rRNAs. RNase MRP is a snoRNP consisting of a single RNA and 7-10 proteins including the core components, POP1 and RPP40. In yeast, Pop1 is essential for RNase MRP accumulation and function (Lygerou et al, 1994). Using RNAi, POP1 or RPP40 was depleted from HeLa cells and POP1 was co-depleted with XRN2. Proteins were analysed by Western blotting (probing was carried out in the laboratory of Dr Ger Pruijn due to a lack of available antibodies) which showed that both POP1 and RPP40 had been specifically depleted by the siRNA treatment (Figure 4.14A). Depletion of these proteins using identical siRNA duplexes has previously been shown to inhibit RNase MRP cleavage of the viperin mRNA leading to increased viperin protein levels (Mattijssen et al, 2010a). RNA was analysed on a denaturing polyacrylamide gel followed by Northern blotting using probes hybridising to the RNA component of RNase MRP and the SRP RNA component, 7SL. This showed that in cells treated with each of these siRNAs, the levels of RNase MRP RNA were decreased between four and six fold (Figure 4.14B).

Pre-rRNAs from cells depleted of either POP1 or RPP40 were also analysed by agarose-glyoxal electrophoresis followed by Northern blotting using different probes spanning ITS1, 5.8S and ITS2. No rRNA processing defects were observed in these cells compared to rRNA extracted from control cells (Figure 4.14C, D). The levels of 21S, 30S and 32S, which would be expected to decrease if RNase MRP was required for site 2 cleavage, were not altered, suggesting that although site 2 is to some extent analogous to yeast A_3, RNase MRP does not appear to be the endonuclease responsible for this cleavage. Depletion of these proteins in combination also did not cause any pre-rRNA processing defects. This result was also obtained by an independent group of collaborators (Dr Sandy Mattijssen and Dr Ger Pruijn, unpublished data). POP1 was also co-depleted with XRN2 to determine whether RNase MRP plays any role in the alternative ITS1 cleavage at 2a. However, the levels of 36S and the ITS1 fragment detected when these proteins were co-depleted were the same as those observed when XRN2 was depleted alone ruling out his possibility (Figure 4.14C).

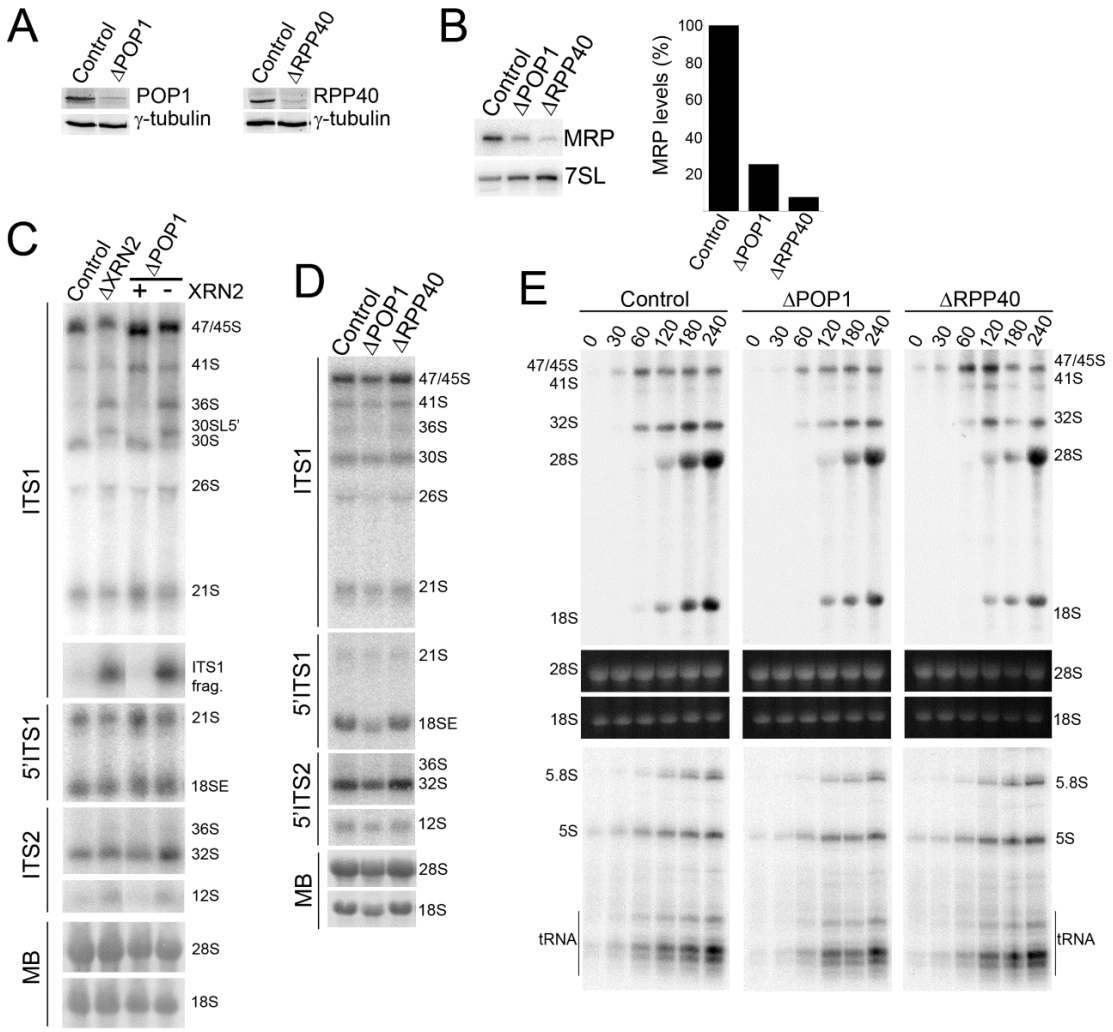


Figure 4.14 Site 2 cleavage does not depend on RNase MRP A) HeLa cells were transfected with siRNA duplexes to deplete the RNase MRP subunits; POP1 or RRP40 or with control siRNAs. After 60 hours RNA and proteins were extracted. Proteins were separated by SDS-PAGE and analysed by Western blotting using antibodies raised against the RNAi targeted proteins or γ tubulin. **B)** RNA from siRNA treated cells was separated on an 8% polyacrylamide/7M urea gel and transferred for Northern blotting using probes hybridising to either the RNA component of RNase MRP or the RNA component of the signal recognition particle (7SL). RNA levels were quantified using ImageQuant software and normalised to 7SL, then compared to levels in control cells. **C)** RNA from cells depleted of POP1, XRN2 or both proteins was analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using probes hybridising to the 5' end of ITS1, the middle of ITS1 or the middle of ITS2. Mature rRNAs were visualised by methylene blue staining. **D)** RNA from cells depleted of RPP40 or POP1 was analysed as in (C). **E)** HeLa cells were transfected with siRNAs to deplete POP1, RPP40 or control siRNAs. After 48 hours cells were pulse labelled with ^{32}P orthophosphate, then grown in normal media with cells harvested at the time points indicated above the panel. RNA was extracted and analysed by both agarose-glyoxal gel electrophoresis and on an 8% denaturing polyacrylamide gel. RNA was then visualised using a phosphoimage. The RNA species detected are indicated to either side of the panel.

To confirm whether any role in ribosome biogenesis for RNase MRP/P could be detected, metabolic labelling experiments were carried out on cells depleted of either POP1 or RPP40. Total RNA was analysed using both agarose-glyoxal gel electrophoresis and polyacrylamide gel electrophoresis (Figure 4.14E). The ratio of 28S to18S rRNAs was not altered compared to control cells and although a slight decrease in the rate of rRNA processing was detected, this was no more significant

than was observed when the essential, non-ribosome biogenesis factor, SRP14, was depleted. No defects in the ratio of long to short forms of 5.8S rRNA were detected as were observed in yeast. Taken together these data strongly suggest that although RNase MRP has other important functions in human cells, that it is not essential for ribosome biogenesis and questions whether site 2 is completely analogous to yeast A_3.

4.2.11 pre-rRNA intermediates in human cell lines and after differentiation

We have identified that two alternative mechanisms exist for the production of 18SE; a major pathway involving exosome processing following site 2 cleavage and a minor pathway involving endonucleolytic cleavage of site 2a. It appears that blocking site 2 cleavage by depleting BOP1 or RBM28 stimulates the alternative cleavage at 2a but inhibiting exonucleolytic processing following site 2 does not, implying these pathways are not redundant. It was, therefore, of interest to investigate whether there is a physiological relevance of conserving two alternative pathways for 18SE production and to determine if there are circumstances when the pathway involving endonucleolytic cleavage at 2a plays a more dominant role. Ribosome biogenesis has been shown to be down-regulated during differentiation when the growth rate of cells is decreased. We therefore speculated that the exonucleolytic processing pathway could represent the most efficient method of 18SE production for rapidly growing cells but when the growth rate is decreased during differentiation, the alternative pathway may be used more. Further, much of the published data describing 36S as the product of endonucleolytic cleavage at 2a was derived from studies carried out in mouse cells and the relative accumulation of 36S to other pre-rRNAs seems greater than in HeLa cells. This raises the possibility that the relative use of these alternative pathways is dependent on cell type. To investigate this hypothesis, the levels of key pre-rRNA intermediates and proteins playing important roles in ITS1 processing were determined in three different cell lines, HeLa, HEK293 and TC7. The TC7 cell line is a derivative of the CaCo-2 cell line (colon cancer cells) and has been used as a model for studying differentiation (Meinl et al, 2008). Cells are grown to confluence and then maintained for a further 22 days during which they undergo differentiation. Pre-rRNA levels were determined in both undifferentiated and differentiated TC7 cells to characterise changes that may occur as ribosome biogenesis is slowed.

HeLa, HEK293 and TC7 cells were grown to 70% confluence and TC7 cells were allowed to differentiate before harvesting. When CaCo-2 undergo differentiation, the levels of the core snoRNP protein, fibrillarin, are unchanged but the levels of many

SSU processome factors including PNO1 are decreased (Knox et al, 2011). To verify that the TC7 cells had undergone differentiation, Western blot analysis was carried out

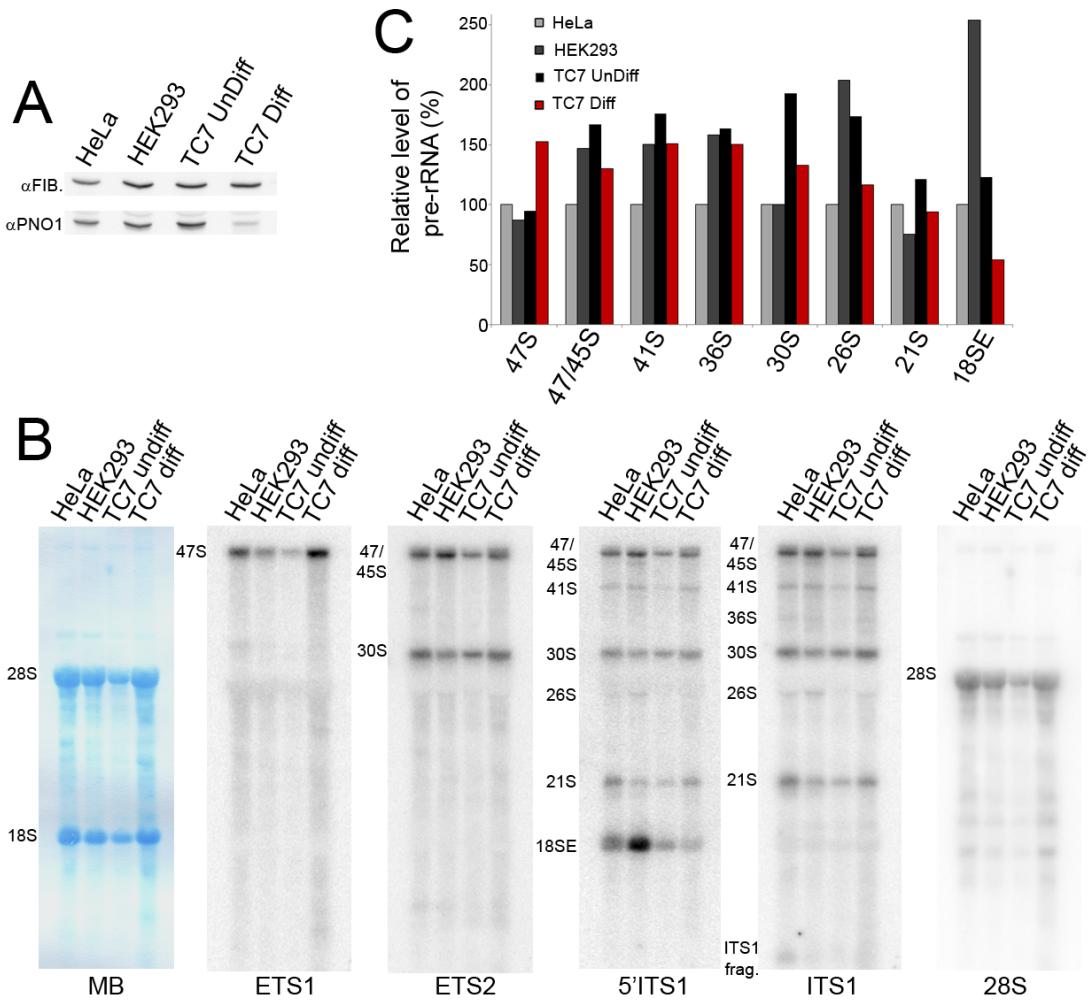


Figure 4.15 The use of alternative ITS1 processing pathways is not altered during differentiation A) HeLa, HEK293 and TC7 cells were cultured. RNA was extracted from HeLa, HEK293 and TC7 cells both before (UnDiff) and after differentiation (Diff). Protein was extracted from cells and separated by SDS-PAGE followed by Western blotting using the antibodies indicated to the left of the panel. **B)** RNA was separated by agarose-glyoxal gel electrophoresis followed by Northern blotting. Membranes were stained with methylene blue (MB) to visualise mature rRNAs. Hybridisation was carried out using radiolabelled probes shown below the panel. Pre-rRNA intermediates observed are indicated using each probe is indicated to the left of each panel. **C)** Quantification was performed using ImageQuant TL software and pre-rRNA levels were normalised to 28S rRNA. Results are summarised in the bar graph. When a pre-rRNA is detected using more than one probe the data shown in the bar graph is as follows: 47S (ETS1), 47/45S (ETS2), 41S (ITS1), 36S (ITS1), 30S (ETS2), 26S (ITS1), 21S (ITS1) and 18SE (18SE) although numbers are representative of data generated with all probes.

using antibodies raised against these two proteins. Fibrillarin levels were observed to be uniform across all cell lines and after differentiation, indicating even loading across the gel, whereas a significant decrease in PNO1 levels was observed in the TC7 cells grown for 22 days, confirming that these TC7 cells had differentiated (Figure 4.15A). RNA was extracted from all these cells and analysed by agarose-glyoxal gel electrophoresis. Northern blotting was carried out using probes hybridising upstream

(ETS1) and downstream (ETS2) of the A' cleavage site in the 5'ETS, at the 5' end of ITS1 and in the middle of ITS1 (Figure 4.15B). The levels of pre-rRNA intermediates detected by each of the probes were determined and compared to levels found in HeLa cells. The results are shown as a bar chart in Figure 4.15C. The levels of 36S, which is generated by endonucleolytic cleavage at the 2a site, were not significantly varied across the cell lines used or after differentiation. It would, therefore, appear that the balance between the alternative processing pathways to produce 18SE is not affected by long-term changes in cell growth rate. It is interesting to note, however, that after differentiation, the levels of 47S are significantly increased compared to those in undifferentiated cells, while the levels of all other pre-rRNA intermediates are decreased after differentiation (Figure 4.15B, C). This demonstrates that as cellular growth rate is reduced and the rate of ribosome biogenesis slows, the rate of rRNA processing is also decreased. Another notable difference is that the level of 18SE is considerably higher in HEK293 cells than in either HeLa or TC7 cells (Figure 4.15B, C).

4.3 Discussion

Here we have characterised two alternative processing pathways for the removal of ITS1 in human cells. The major pathway involves a single endonucleolytic cleavage at site 2, which is to some extent analogous to yeast A_3. As in yeast, this is followed by 5' exonucleolytic processing to produce the mature 5' end of 18S but in contrast to yeast, is also followed by 3' exonucleolytic processing by the exosome to site 2a. This generates 18SE, a precursor of 18S, which extends approximately 25 nt beyond the 3' end of 18S. We also provide evidence that 18SE can be produced by an endonucleolytic cleavage at site 2a, which is equivalent to yeast A_2, but in humans, this cleavage is inefficient and plays a minor role in pre-rRNA processing. Although inhibiting site 2 cleavage stimulates 2a cleavage, this minor pathway cannot compensate for a block in exonucleolytic processing following site 2 cleavage to produce 18SE unless stimulated to do so by co-depletion of XRN2.

Following site 2 cleavage, RRP6, the 3'-5' exonuclease component of the nucleolar exosome, processes to site 2a to remove the majority of ITS1 sequence from the 3' end of 18S. The core exosome, the TRAMP helicase, MTR4, and to a lesser extent, the RRP6 cofactor, MPP6, are required for this processing. Depletion of RRP6 caused a significant decrease, but not an absolute block, in 18SE accumulation and the appearance of a partially processed intermediate, 21SC, suggesting other proteins, such as the REX exonucleases, are also able to inititate and in some cases complete 21S processing to 2a. In contrast, depletion of either the core exosome or MTR4

caused a smear of intermediates perhaps implying that these proteins are required to recruit the exosome to the pre-rRNA or to enhance the activity of RRP6 to enable it to degrade this substrate. The region of ITS1 that is removed is approximately 800 nt long, 82% GC rich and is highly structured, possibly explaining the requirement for the helicase, MTR4, and making it likely that the activity of human RRP6 is significantly increased in the context of the cell compared to that observed *in vitro* (see section 3.2.2). DIS3 was not required for this exonucleolytic processing step and as DIS3 is excluded from nucleoli in human cells, it is likely that ITS1 processing by the exosome predominantly occurs in the nucleolus. The nucleolar RRP6 cofactor, C1D, was also not required for ITS1 processing although we have previously shown that it does function in the turnover of the 37S* and in 5.8S processing (Figure 3.9 and Figure 3.12). We further observed that the SSU biogenesis factors, RPS19, ENP1, RRP5 and RCL1, were required for this processing linking the exosome to the SSU processome.

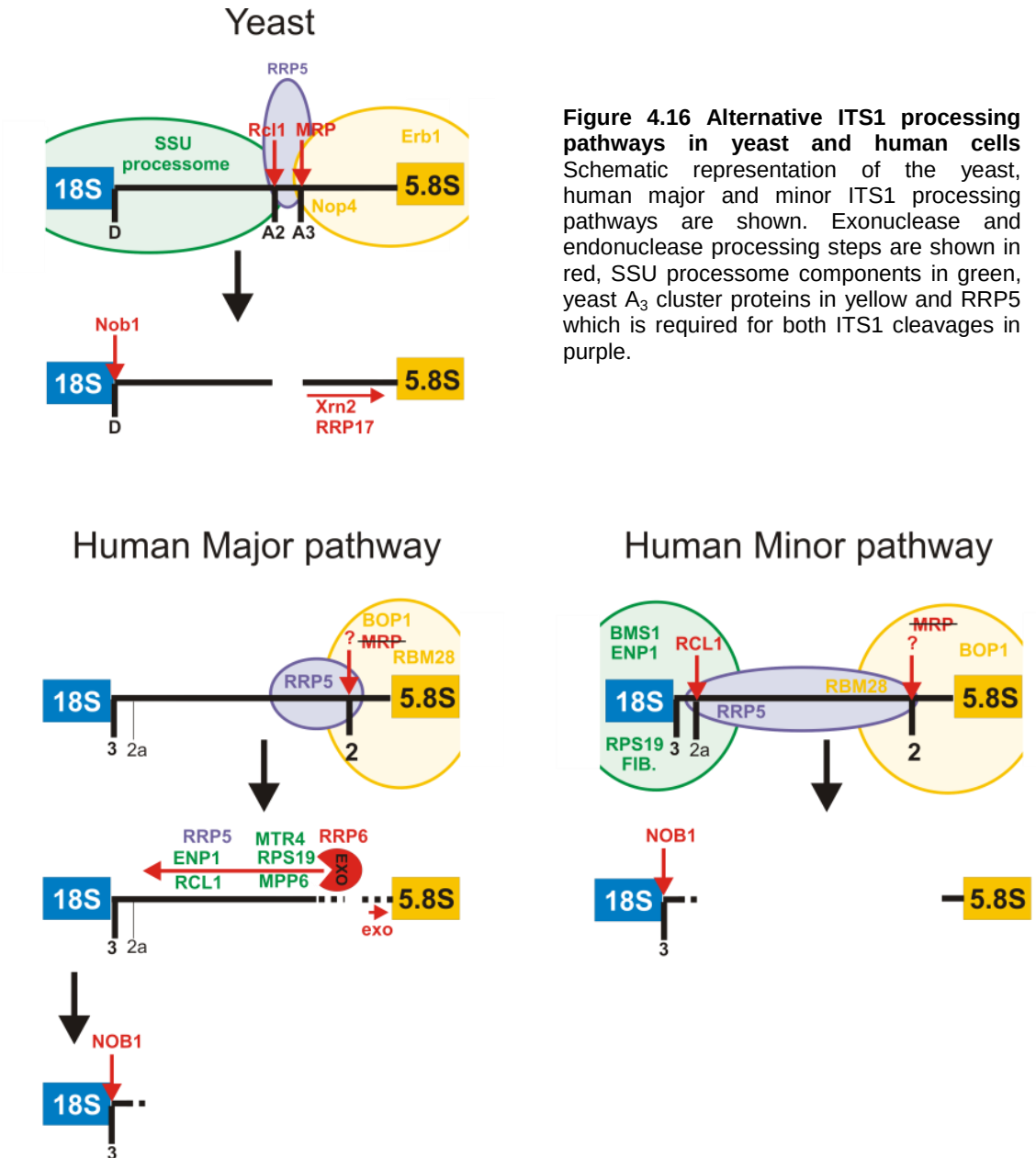


Figure 4.16 Alternative ITS1 processing pathways in yeast and human cells Schematic representation of the yeast, human major and minor ITS1 processing pathways are shown. Exonuclease and endonuclease processing steps are shown in red, SSU processome components in green, yeast A_3 cluster proteins in yellow and RRP5 which is required for both ITS1 cleavages in purple.

Although we have identified alternative ITS1 processing pathways, exonucleolytic processing following site 2 cleavage is clearly the major route since we see little evidence of pre-rRNAs cleaved at 2a prior to site 2 and inhibiting exosome processing of ITS1 severely decreases both 18SE and 18S levels. This is supported by the observation that stimulating 2a cleavage by depletion of XRN2 is able to rescue 18SE production in the absence of the core exosome or MTR4. Exonucleolytic processing between sites 2 and 2a in human cells is somewhat comparable to yeast where modifications in the A_2 site, which block cleavage, result in the exosome processing back to A_2 following A_3 cleavage to enable normal 18S production (Allmang et al, 2000). However, this alternative pathway has evolved to be the major processing mechanism in humans.

We found that multiple components of the SSU processome are required for 2a cleavage, making this cleavage analogous to A_2 cleavage in yeast. In contrast to yeast, cleavage at site 2a is part of a minor pre-rRNA processing pathway in human cells. The only precursor which is specific to this pathway, 36S, is barely detectable in a variety of human cells lines (HeLa, HEK293, TC7) implying that only a small fraction of pre-rRNA is normally processed via this pathway but it is not possible to quantifiy this. When cleavage at site 2a is stimulated, either by blocking site 2 cleavage or by depletion of XRN2, no precursors were detected that had been processed at site 2a but still contained any 5'ETS sequence. This implies that, as in yeast (Venema & Tollervey, 1995), cleavage at 2a (A_2) occurs simultaneously with, or following, removal of the 5'ETS.

There are several reasons why 2a cleavage may be inefficient and these include, the proximity of the cleavage site to the 3' end of 18S, possible inhibition by protein factors or, if removal of the 5'ETS precedes 2a cleavage, that this causes dissociation of the necessary 2a processing machinery. In yeast, early pre-rRNA processing steps, including A_2 cleavage, can occur co-transcriptionally whereas in humans, A' cleavage is the only step thought to occur during transcription. It is possible, therefore, that the efficiency of 2a (A_2) is determined by the timing of processing relative to transcription. Alternatively, depletion of proteins required for site 2 cleavage stimulated 2a processing, possibly indicating that they also inhibit 2a cleavage. This may represent a regulatory mechanism by which the optimal pre-rRNA pathway is favoured while maintaining an alternative method of processing. It is not entirely clear why depletion of XRN2 stimulates 2a cleavage but as mentioned in the previous chapter, XRN2 appears to coordinate the optimal order of pre-rRNA cleavages. It is not, however, entirely clear why there are two alternative pathways for 18SE production that are not fully redundant. We speculated that when cells underwent differentiation, the

decreased growth rate may alter the balance between these pathways but this was not the case. However, similar alternative pathways have also been described in *X. laevis* oocytes suggesting that this is a common feature of pre-rRNA processing in higher eukaryotes (Savino & Gerbi, 1990).

In yeast, a cluster of seven proteins, including Erb1, Nop7 and Ytm1, are co-recruited to the A_3 cleavage site and these proteins are required for processing to form the mature 5' end of $5.8S_S$, recruitment of large subunit ribosomal proteins and regulating pre-rRNA folding in ITS2 (Granneman et al, 2011; Sahasranaman et al, 2011). The homologue of Erb1, BOP1, is also found in a complex with homologues of Nop7, PES1, and Ytm1, WDR12, (Holzel et al, 2005; Lapik et al, 2004) making it likely that this complex is functionally homologous to the yeast A_3 cluster. In yeast, an additional protein associated with the A_3 cluster is Nop4, but unlike the others, this protein is essential for A_3 cleavage. Our data show that cleavage at site 2 in human pre-rRNA is inhibited by depletion of RRP5 or BOP1 but loss of RBM28 (Nop4) only slowed the rate of cleavage at this site. The involvement of these factors in site 2 cleavage implies that it is analogous to yeast A_3 cleavage but the precise roles of the homolgous proteins do not seem to be conserved. Depletion of RRP5, BOP1 or RBM28 caused a delay in pre-rRNA processing following A' cleavage implying that both ITS1 cleavages were slowed. Depletion of RRP5 specifically blocked cleavage at 2a and this effect was sufficient to cause ITS1 cleavages to be bypassed, forcing processing to occur in ITS2. In yeast, blocking A_3 cleavage or the downstream processing inhibits formation of the major, short form of 5.8S, $5.8S_{S,}$ and causes accumulation of $5.8S_L$. Depletion of BOP1 from HeLa cells caused a similar phenotype implying that this step of processing is conserved from yeast to humans.

Although site 2 appeared to be analogous to yeast A_3, depletion of RNase MRP proteins did not affect site 2 cleavage or the subsequent processing to produce 5.8S. RNAi depletion of the RNase MRP/RNase P proteins using identical siRNAs has, however, been shown to inhibit cleavage of the viperin mRNA (Mattijssen et al, 2011). Consistent with this, knockdown of the RNase MRP/P subunit, RPP38, was previously shown to block tRNA processing but only have a minimal effect on 5.8S rRNA processing (Cohen et al, 2003). It therefore, appears that RNase MRP is not the primary endonuclease responsible for site 2 cleavage in humans, although we cannot rule out the possibility that it plays a redundant role in this processing. This suggests that an additional endonuclease, possibly one specific to higher eukaryotes, is required for ITS1 cleavage to separate the LSU and SSU rRNAs.

Characterisation of the specific functions of RBM28, RNase MRP and BOP1 in

human pre-rRNA processing is important as these proteins are linked to ANE and CHH syndromes and colorectal cancer, respectively (Chung et al, 2011; Glazov et al, 2011; Killian et al, 2006; Nousbeck et al, 2008; Ridanpaa et al, 2001; Spiegel et al, 2010). This will be discussed in more detail in section 6.4.

Chapter Five

Analysis of factors involved in formation of the 3' end of 18S rRNA

5.1 Introduction

The small subunit rRNA, 18S, is co-transcribed with the large subunit rRNAs, 5.8S and 28S (25S in yeast). 18S rRNA is separated from the large subunit rRNAs by cleavage in ITS1 at site 2 in human pre-rRNA or at sites A_2 and A_3 in yeast (Figure 5.1). Many proteins required for formation of the small subunit form a multiprotein complex called the SSU processome and five major sub-complexes, U3 snoRNP, MPP10 complex, tUTP, bUTP and cUTP, have been identified (Henras et al, 2008). The SSU processome is recruited to the pre-rRNA co-transcriptionally and is responsible for many steps in the biogenesis of the small subunit including removal of the 5'ETS, 18S pre-rRNA processing and assembly of early pre-ribosomal complexes.

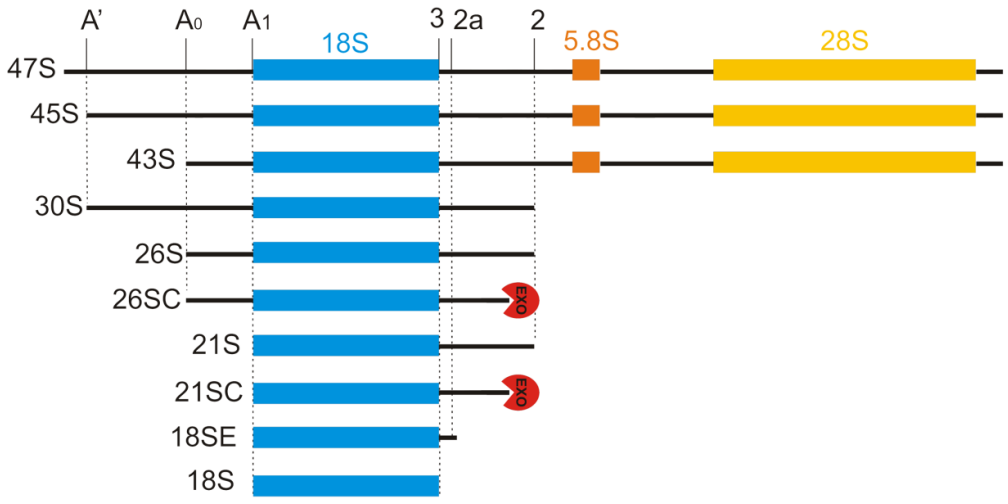


Figure 5.1 pre-rRNA processing to produce 18S rRNA Schematic outlines of the pre-rRNA intermediates in the pathway for 18S production in human cells. Cleavage sites are shown above the full length transcript and 18S (blue), 5.8S (orange) and 28S (yellow) mature rRNAs are shown. The lengths of intermediates is given to the left in Svedburg units (S) and C denotes a 3' shortened form of an intermediate and EXO (red) indicates this is processed by the exosome.

In yeast, a number of additional proteins are required for the final stages of formation of the mature 3' end of 18S which occur as the pre-40S complex is exported from the nucleus into the cytoplasm. Some of these factors, such as Dim1, Nob1, Pno1, Prp43, Enp1, Ltv1 and Tsr1, become associated with early complexes but perform their main functions during or after export into the cytoplasm. Others, such as Rio2, are predominantly localised in the cytoplasm and are only associated with pre-

40S complexes during the final maturation steps. This set of factors includes some proteins whose functions in yeast are known, while the roles of others remain elusive. Nob1 is the endonuclease responsible for cleavage at site D to remove all ITS1 sequence from the 3' end of 18S and Pno1 is proposed to be a cofactor of Nob1 (Lamanna & Karbstein, 2011; Vanrobays et al, 2004). Prp43, is also necessary for this processing step (Pertschy et al, 2009). Since Prp43 is a helicase, it is suggested that structural rearrangement of the pre-rRNA around site D is required for Nob1 cleavage. The dimethyltransferase, Dim1, catalyses the methylation of two conserved adenosine residues at the 3' end of 18S and failure to do this causes formation of non-functional ribosomes (Lafontaine et al, 1998b). Enp1 is a protein which, along with Ltv1, is released from pre-40S complexes when it is phosphorylated by Hrr25 causing a major structural rearrangement of the beak region of the small subunit. Regulation of this step is critical for enabling export of the pre-40S complex into the cytoplasm (Schafer et al, 2006). Much less is known about SSU biogenesis factors in human cells but late pre-40S complexes have been purified through RIO2 and all these proteins, except DIM1, were co-purified, suggesting that the composition of these complexes is relatively well conserved (Zemp et al, 2009).

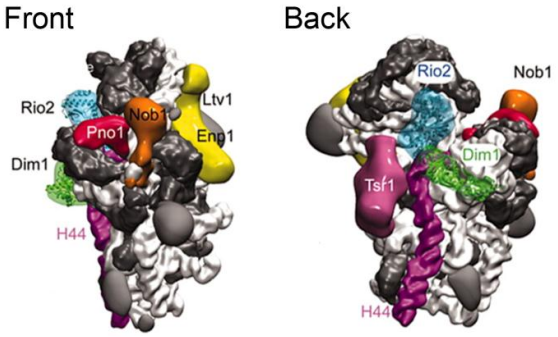


Figure 5.2 Structure of a late pre-40S complex from *S. cerevisiae* Model of late pre-40S complex from *S. cerevisiae* based on a cryo electron microscopy showing the positions of late small subunit transacting factors modelled onto a crystal structure of the *T. thermophilia* small ribosomal subunit. Adapted from Strunk et al. 2011

In yeast, it has been demonstrated, using an *in vivo* cross linking method (CRAC), that many of these late-acting SSU biogenesis factors are directly associated with 18S rRNA and binding sites near the 3' end of 18S have been defined (Granneman et al, 2010). A crystal structure of the eukaryotic 40S small subunit in complex with Initiation factor I from *Tetrahymena thermophilia* (Rabl et al, 2010) shows the three dimensional structure of 18S and the relative positions of the proteins found in the mature ribosome. A structure of a pre-40S complex from *S. cerevisiae* containing most of the late-acting SSU biogenesis factors discussed here has also been generated using cryo-electron microscopy (Figure 5.2) (Strunk et al, 2011). Taken together, these studies enable a clearer understanding of the final structures of the 3' end of 18S and the possible interactions, rearrangements and regulation of the late acting biogenesis factors.

In the previous chapter, we have shown that the pre-rRNA processing pathway used to form the 3' end of 18S in human cells is considerably different to that used in yeast. In yeast, removal of ITS1 occurs by sequential endonucleolytic cleavages at A_2 or A_3 then at site D. In contrast, in human cells, cleavage at site 2, which is in some ways analogous to the yeast A_3 cleavage, is followed by exonucleolytic processing by the exosome to the 2a site which removes the majority of ITS1 sequence (Figure 5.1). Since this additional ITS1 processing step is not found in yeast but the composition of late pre-40S particles seems well conserved (Zemp et al, 2009), it is likely that factors involved in formation of the 3' end of 18S have evolved additional or different functions in human cells to enable this increased complexity in processing to occur. Formation of the mature end of 18S appears to be carried out by NOB1 endonucleolytic cleavage as in yeast (section 4.2.8). The immediate precursor of 18S, 18SE, is only extended from the 3' end of 18S by approximately 25 nt and is predominantly generated by exonucleolytic processing rather than endonucleolytic cleavage. This raises several questions, such as why and how exonucleolytic processing is halted close to, but not directly at, the mature end of 18S. Exonucleolytic processing of ITS1 by the exosome also involves the small subunit protein, RPS19, the SSU biogenesis factors, ENP1 and to some extent RCL1. This implies that ITS1 processing is coordinated with other important events in formation of the mature 3' end of 18S and raises the possibility that the proteins involved in these other steps are required for exonucleolytic processing of ITS1 to 2a. Since these late-acting SSU biogenesis factors bind to the 3' end of 18S in yeast, it is possible that they are responsible for terminating exonucleolytic processing at 2a. It is further possible that other factors involved in formation of the 3' end of 18S are involved in mechanisms by which ITS1 processing is regulated. As phosphorylation often provides a control mechanism, the kinase activity of RIO2 is of particular interest. Although all the proteins identified in yeast as performing the final 18S maturation steps have human homologues, little is known about whether their functions are conserved and what additional roles they may play.

5.2 Results

5.2.1 RNAi depletion of late-acting SSU biogenesis factors

In the previous chapter, we showed that the immediate precursor of 18S, 18SE, is produced by exonucleolytic processing by the exosome. This processing requires the small subunit protein, RPS19, the late acting small subunit (SSU) biogenesis factor, ENP1 and to some extent the SSU processome component, RCL1. We aimed, therefore, to identify other proteins that may be required for efficient exosome

processing of ITS1 to produce 18SE. In yeast, several proteins have been shown to be important for formation of the 3' end of 18S with defects in conversion of 20S into 18S being characteristic of depleting these proteins, making the human homologues likely candidates for involvement in exosome processing of ITS1.

HeLa cells were chemically transfected with siRNAs targeting the mRNAs of ENP1, NOB1, PNO1, RIO2, DIM1 or PRP43. A duplex targeting the firefly luciferase mRNA, which is not expressed in HeLa cells, was used as a negative control. 60 h after transfection, cells were harvested. Proteins were extracted and analysed by SDS-PAGE followed by Western blotting. Antibodies against ENP1, NOB1, PNO1, RIO2 and PRP43 were used and showed that the levels of each of the targeted proteins were specifically decreased compared to control cells while levels of the cytoskeletal protein, actin, or the core snoRNP protein, fibrillarin, were not affected (Figure 5.3A). For DIM1, a usable antibody against the endogenous protein could not be sourced, so it was not possible to confirm whether protein levels had been successfully decreased in HeLa cells. However, to confirm that the siRNA duplexes used were able to specifically decrease DIM1 levels, their effectiveness was tested in stably transfected HEK293 cells expressing FLAG-tagged DIM1. Such cells were transfected with either control siRNAs or those targeting DIM1 and DIM1 levels were monitored by Western blotting using an antibody which detects the FLAG-tag. This showed that FLAG-DIM1 levels were significantly reduced by the siRNA treatment while actin levels remained unchanged (Figure 5.3A).

RNA was also extracted from siRNA treated cells and pre-rRNAs analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using a probe hybridising to the 5' end of ITS1 (Figure 5.3B). Results were visualised using a phosphorimager and quantification of the levels of each pre-rRNA intermediate compared to control cells is given as a line diagram (Figure 5.3C). As described in the previous chapter, depletion of ENP1 led to an increase in 21S, the appearance of 21SC and a decrease in 18SE levels. Depletion of PNO1 did not affect the levels of 47/45S or 30S but caused a significant increase in the 26S pre-rRNA intermediate. 26S corresponds to a precursor extending from the A_0 cleavage site in the 5'ETS to site 2 in ITS1 and is generated when cleavage at A_1 is impaired. An increase in 21S levels and a decrease in 18SE levels were also observed and smears of 26SC and 21SC were detectable below 26S and 21S respectively implying PNO1 may play a role in processing to form 18SE (Figure 5.3 B, C). This will be discussed in more detail in Section 5.2.5. When either RIO2 or NOB1 were depleted from HeLa cells, a significant increase in 18SE levels was observed while the levels of the other pre-rRNA species did not vary relative to cells transfected with the control siRNA (Figure 5.3 B, C). This

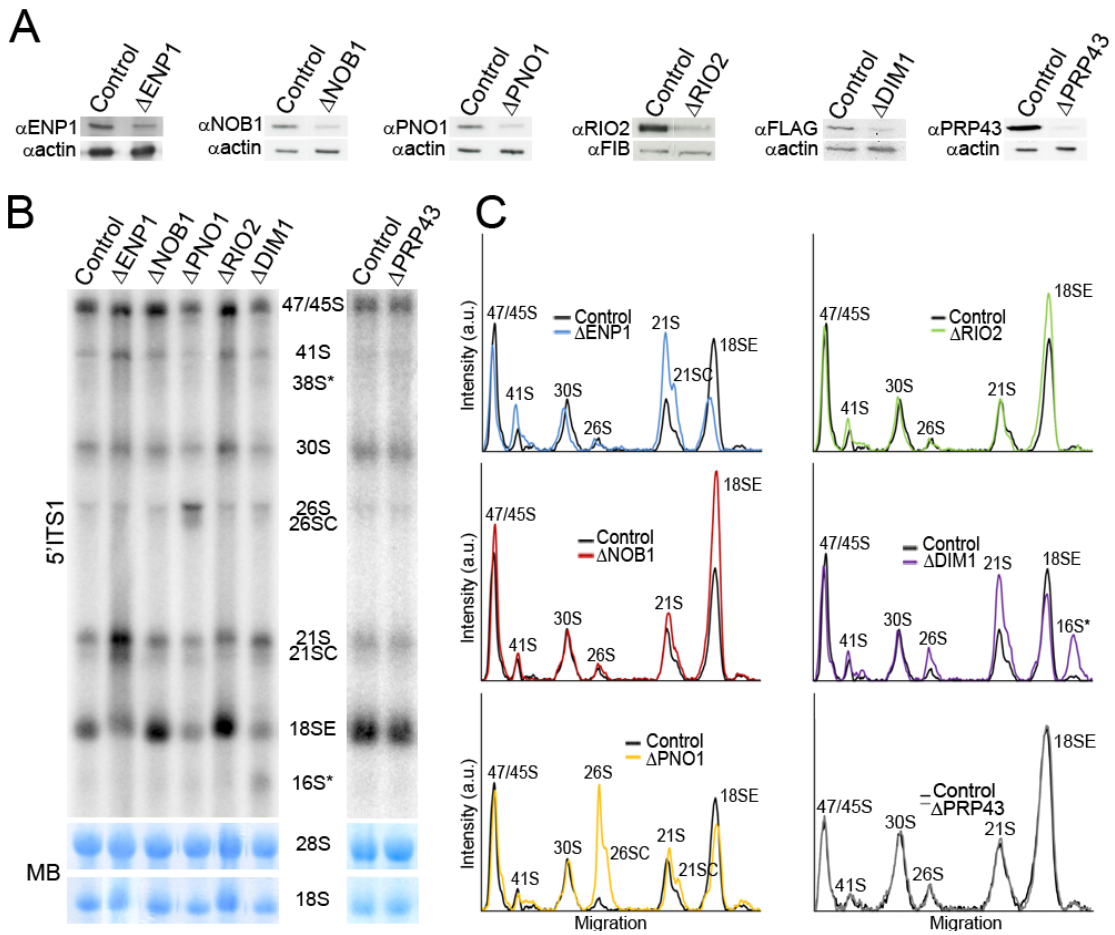


Figure 5.3 Late-acting pre-40S biogenesis factors are required at different stages of pre-rRNA processing A) HeLa cells were chemically transfected with siRNA duplexes specifically targeting firefly luciferase (control), ENP1, NOB1, PNO1, RIO2, DIM1 or PRP43. Protein was extracted from RNAi depleted cells after 60 h and analysed by SDS-PAGE followed by Western blotting. Proteins of interest (except DIM1) were detected by antibodies specific to each protein, indicated to the left of each panel. HEK293 cells expressing FLAG-DIM1 at the same level as the endogenous protein were also transfected with siRNA duplexes to deplete DIM1 or control siRNAs. Proteins were analysed by SDS-PAGE followed by Western blotting using an antibody which detected the FLAG-tag. **B)** RNA was extracted from siRNA treated cells, separated by electrophoresis on a 1.2 % agarose-glyoxal gel and transferred to a nylon membrane. Methylene blue (MB) staining was used to visualise 28S and 18S mature rRNAs and Northern blot hybridisation was performed using a probe detecting the 5' end of ITS1 **C)** The levels of pre-rRNA intermediate detected by Northern blotting in cells depleted of each protein are shown as lane profiles graphs, generated using ImageQuant software. Levels were normalised to 28S levels and each graph depicts intensity in arbitrary units (a.u.) relative to the distance migrated though the gel. pre-rRNA levels in cells depleted of a protein of interest (colour) relative to levels detected in cells transfected with the control siRNA (black) are shown.

implies that, as in yeast, these proteins are only required for the final maturation step in 18S production.

Depletion of DIM1 leads to a small increase in the amounts of 26S and 21S accumulated and a subtle decrease in 18SE levels suggesting that DIM1 may be indirectly involved in exonucleolytic processing of 21S. More strikingly, novel pre-rRNA intermediates smaller than 18S and 41S, here termed, 16S* and 38S*, respectively, were also detected (Figure 5.3B, C, Figure 5.4A, B). 38S* was not always detectable and was only accumulated to significant levels in some knockdown experiments. The sizes of these fragments imply that they are not generated by an alternative processing

pathway, but instead suggest that they are aberrant fragments of pre-RNA undergoing degradation. To map these fragments, RNA extracted from DIM1 depleted cells was analysed by Northern blotting using probes hybridising to the 3' end of the 5'ETS (ETS3) and the middle of ITS1 (ITS1). Both fragments were detected using the ITS1 probe but neither was detected by the ETS3 probe (Figure 5.4A). It was, therefore, concluded that 16S* and 38* are generated from 21S or 18SE and 41S being degraded from the 5' end. This implies that non-DIM1 associated pre-rRNAs are detected as aberrant and are targeted for degradation by 5'-3' exonucleases (Figure 5.4B).

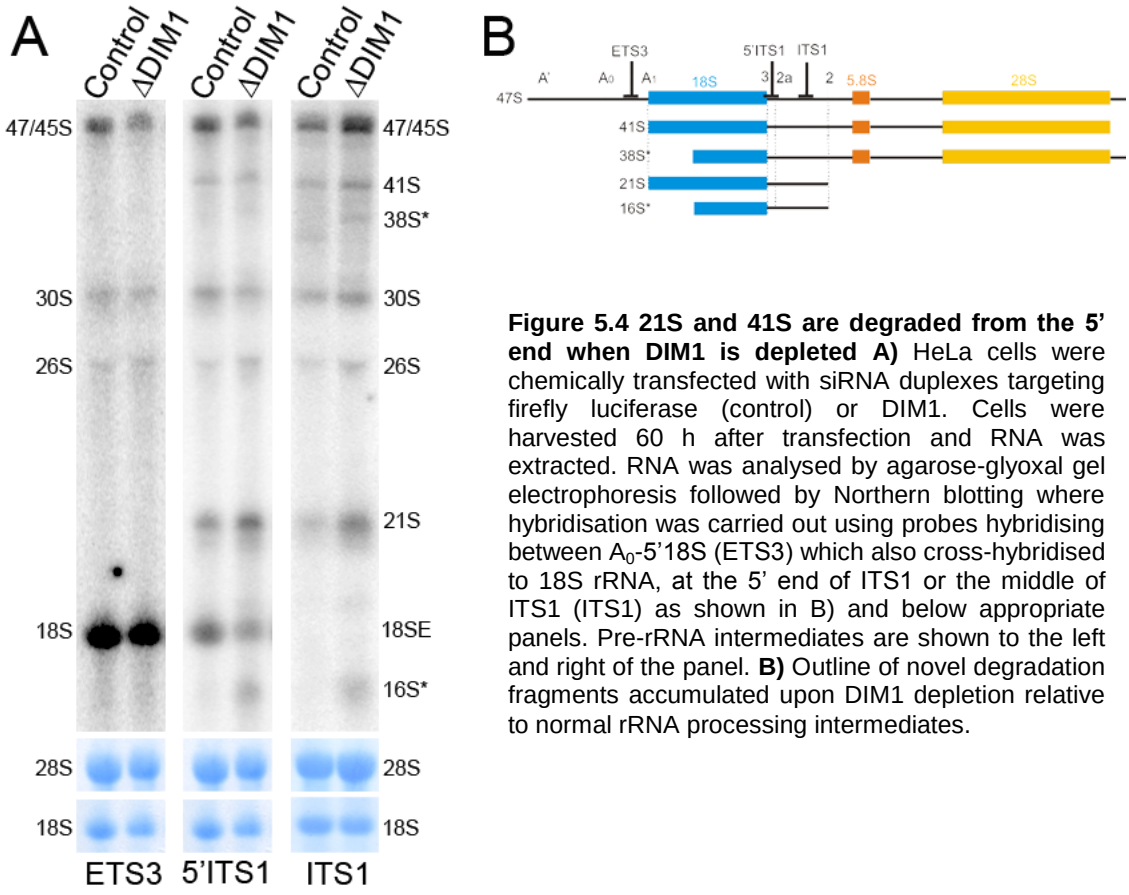


Figure 5.4 21S and 41S are degraded from the 5' end when DIM1 is depleted A) HeLa cells were chemically transfected with siRNA duplexes targeting firefly luciferase (control) or DIM1. Cells were harvested 60 h after transfection and RNA was extracted. RNA was analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting where hybridisation was carried out using probes hybridising between A_0-5'18S (ETS3) which also cross-hybridised to 18S rRNA, at the 5' end of ITS1 or the middle of ITS1 (ITS1) as shown in B) and below appropriate panels. Pre-rRNA intermediates are shown to the left and right of the panel. **B)** Outline of novel degradation fragments accumulated upon DIM1 depletion relative to normal rRNA processing intermediates.

5.2.2 PRP43 is not required for 18S production in HeLa cells

It was expected that the human homologue of yeast Prp43 would have a similar role in ribosome biogenesis to its yeast counterpart and cause an accumulation of the immediate precursor of 18S, but depletion of PRP43 from HeLa cells did not alter the levels of any of the precursors of 18S rRNA (Figure 5.3B, C). This suggests that the function of this protein in cleavage at the 3' end of 18S may not be conserved from yeast to humans. It was, however, possible that the kinetics of 18S production were affected by depletion of PRP43 and Northern blotting may not be a sensitive enough

method to detect this. We therefore investigated the effect of depleting PRP43 on the kinetics of pre-rRNA processing using metabolic labelling.

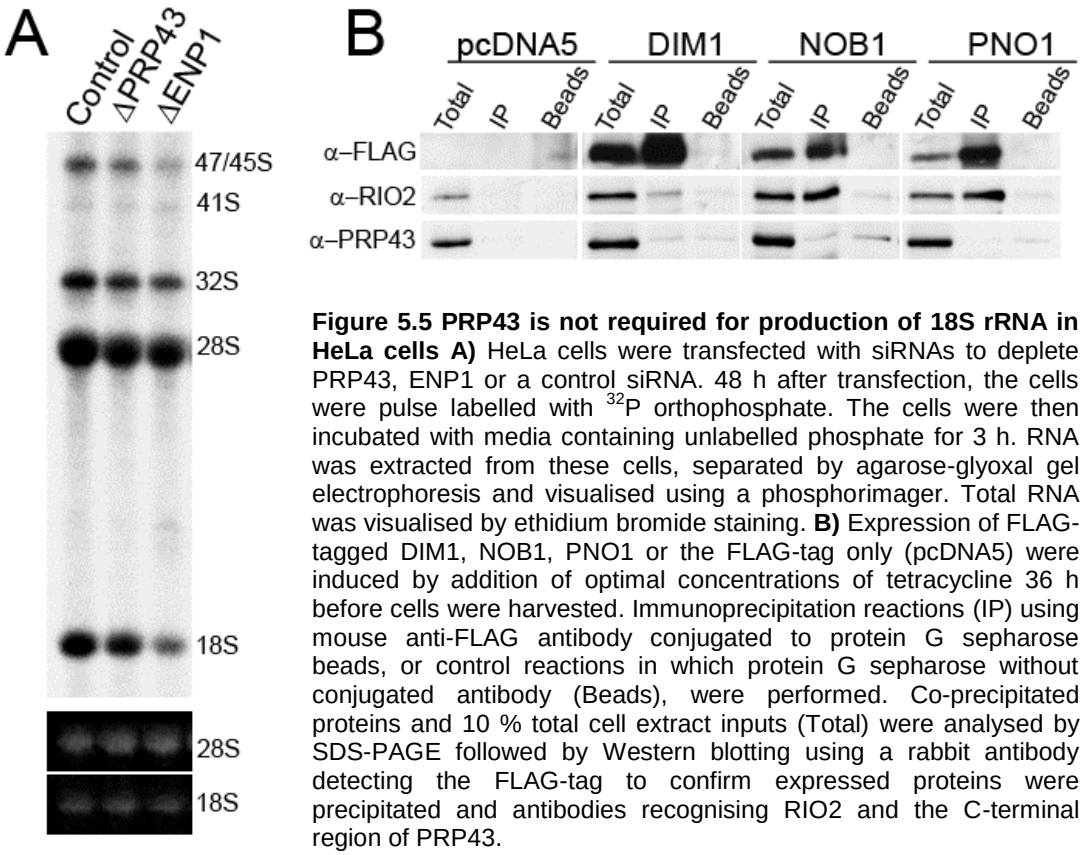


Figure 5.5 PRP43 is not required for production of 18S rRNA in HeLa cells A) HeLa cells were transfected with siRNAs to deplete PRP43, ENP1 or a control siRNA. 48 h after transfection, the cells were pulse labelled with ^{32}P orthophosphate. The cells were then incubated with media containing unlabelled phosphate for 3 h. RNA was extracted from these cells, separated by agarose-glyoxal gel electrophoresis and visualised using a phosphorimager. Total RNA was visualised by ethidium bromide staining. **B)** Expression of FLAG-tagged DIM1, NOB1, PNO1 or the FLAG-tag only (pcDNA5) were induced by addition of optimal concentrations of tetracycline 36 h before cells were harvested. Immunoprecipitation reactions (IP) using mouse anti-FLAG antibody conjugated to protein G sepharose beads, or control reactions in which protein G sepharose without conjugated antibody (Beads), were performed. Co-precipitated proteins and 10 % total cell extract inputs (Total) were analysed by SDS-PAGE followed by Western blotting using a rabbit antibody detecting the FLAG-tag to confirm expressed proteins were precipitated and antibodies recognising RIO2 and the C-terminal region of PRP43.

HeLa cells were depleted of either PRP43 or ENP1, a protein that we have clearly demonstrated is required specifically for 18S production (Figure 4.3 and Figure 4.10) and then metabolically labelled using ^{32}P orthophosphate. Following this, cells were grown in media containing unlabelled phosphate for three hours and harvested. RNA was extracted and analysed by agarose-glyoxal gel electrophoresis then visualised by autoradiography. The ratio of 18S:28S was calculated for both PRP43 and ENP1 depletion and while depletion of ENP1 led to a two fold decrease in 18S levels, depletion of PRP43 did not alter the ratio of the mature rRNAs compared to control cells (Figure 5.5A). It is important to note however, that depletion of ENP1 or PRP43 led to an approximately 20 % decrease in the total amount of labelled rRNA. As discussed in section 4.2.3, depletion of any essential protein, even those not involved in ribosome biogenesis causes a comparable decrease in the amount of labelled rRNA. This is probably due to a decrease in cell growth rate and is not thought to reflect an involvement of the given protein in pre-rRNA processing.

In yeast, many of the proteins involved in formation of the 3' end of 18S have been shown to interact with one another, both *in vitro* and *in vivo*, due to their mutual associations with 18S rRNA and as components of pre-40S complexes. We therefore

investigated if PRP43 was stably associated with complexes containing DIM1, NOB1 or PNO1 using immunoprecipitation. HEK293 cell lines stably transfected with plasmids enabling the inducible expression of FLAG-tagged DIM1, PNO1 and NOB1 to the same level as the endogenous protein or just the FLAG-tag (pcDNA5) were used to make whole cell extracts. Immunoprecipitation reactions were carried out using an antibody against the FLAG-tag conjugated to protein G sepharose to precipitate the FLAG-tagged proteins and any proteins that interact with them. Co-precipitated proteins were separated by SDS-PAGE and analysed by Western blotting using antibodies to detect the FLAG-tag, PRP43 or RIO2. Probing with the antibody detecting the FLAG-tag showed that DIM1, NOB1 and PNO1 were each efficiently precipitated from their cell extracts (Figure 5.5B). RIO2 was also efficiently co-precipitated by DIM1, NOB1 and PNO1 with >20 % being precipitated by NOB1 and PNO1 and approximately 3 % by DIM1 suggesting that RIO2 is a core component of the processing complex formed at the 3' end of 18S rRNA (Figure 5.5B). Different splice variants of PRP43 have been identified so to investigate if PRP43 is associated with pre-40S complexes, antibodies that recognise either the N- or C-terminal of the protein were used in Western blotting. Both antibodies detected only a single protein in total cell extracts and neither showed, PRP43 being co-precipitated by DIM1, NOB1 or PNO1 above the background levels observed in the beads only control lanes. This suggests that PRP43 does not interact either directly or indirectly with these proteins although it is possible that an alternative isoform not recognised by either of these antibodies could be involved (Figure 5.5B). Taken together our data suggest that the role of PRP43 in formation of the 3' end of 18S is not conserved from yeast to humans. PRP43 has therefore, been excluded from further work.

5.2.3 Interactions of proteins involved in 3' end formation of 18S

The immediate precursor of 18S, 18SE, which is produced by exonucleolytic processing only extends beyond the 3' end of 18S by approximately 25-30 nt. It is not clear how the exonuclease activity of the exosome is arrested at the 3' end of this precursor. It is unlikely that RNA secondary structure is responsible for stalling the exosome since the ITS1 region processed by the exosome is highly structured and 25-30 nt is insufficient to form any significant structures. A more likely explanation is that protein(s) bound to either the 3' end of 18S or to the 5' end of ITS1 may physically block the progress of the exosome, preventing it from degrading beyond 18SE. In yeast, CRAC data have identified binding sites for several of the late acting SSU biogenesis factors (NOB1, DIM1, RIO2 and ENP1) at the 3' end of 18S (Granneman et

al, 2010). We decided to investigate the interactions of the human homologues of the late pre-40S complex proteins both with each other and with 18S-ITS1 rRNA.

5.2.3.1 Protein-protein interactions between SSU biogenesis factors *in vitro*

Recombinant N-terminally GST-tagged DIM1, NOB1, PNO1, RIO2 and ENP1 and His-tagged DIM1 and NOB1 were over-expressed in *E. coli* and purified using their tags. Purified proteins were analysed by SDS-PAGE followed by Coomassie staining and in each case, the tagged protein represented the majority of protein in the sample, except for GST-ENP1 in which a significant contaminant of 37 kDa was also observed (Figure 5.6 A). This contaminant contained the GST-tag (data not shown) so is likely to be formed by degradation of full length GST-ENP1.

Recombinant GST or GST-tagged DIM1, NOB1 and PNO1 were bound to glutathione sepharose and incubated with equal amounts of His-tagged DIM1 or NOB1. Complexes formed were then purified and analysed by SDS-PAGE followed by Western blotting. An antibody recognising the His-tag was used to detect His-DIM1 or His-NOB1 that had been co-purified with and therefore interacted with the GST-tagged proteins. Neither His-DIM1 nor His-NOB1 were co-purified by the GST-tag alone confirming that neither interacted with the tags of the GST-proteins tested. Interactions were observed between DIM1 and NOB1 in both reciprocal reactions (His-DIM1+ GST-NOB1 and GST-DIM1+ His NOB1) (Figure 5.6B). DIM1 did not, however, interact with PNO1. A robust interaction was detected between PNO1 and NOB1 with approximately 50 % of the His-NOB1 protein input being co-purified (Figure 5.6B). Interestingly, an interaction was detected between GST- and His-tagged NOB1 implying that NOB1 is able to interact with itself (Figure 5.6B). NOB1 is a PIN domain protein and other proteins containing this domain function as tetramers so these data would suggest that human NOB1 is able to form multimeric complexes. No such interaction was seen between the two tagged forms of DIM1.

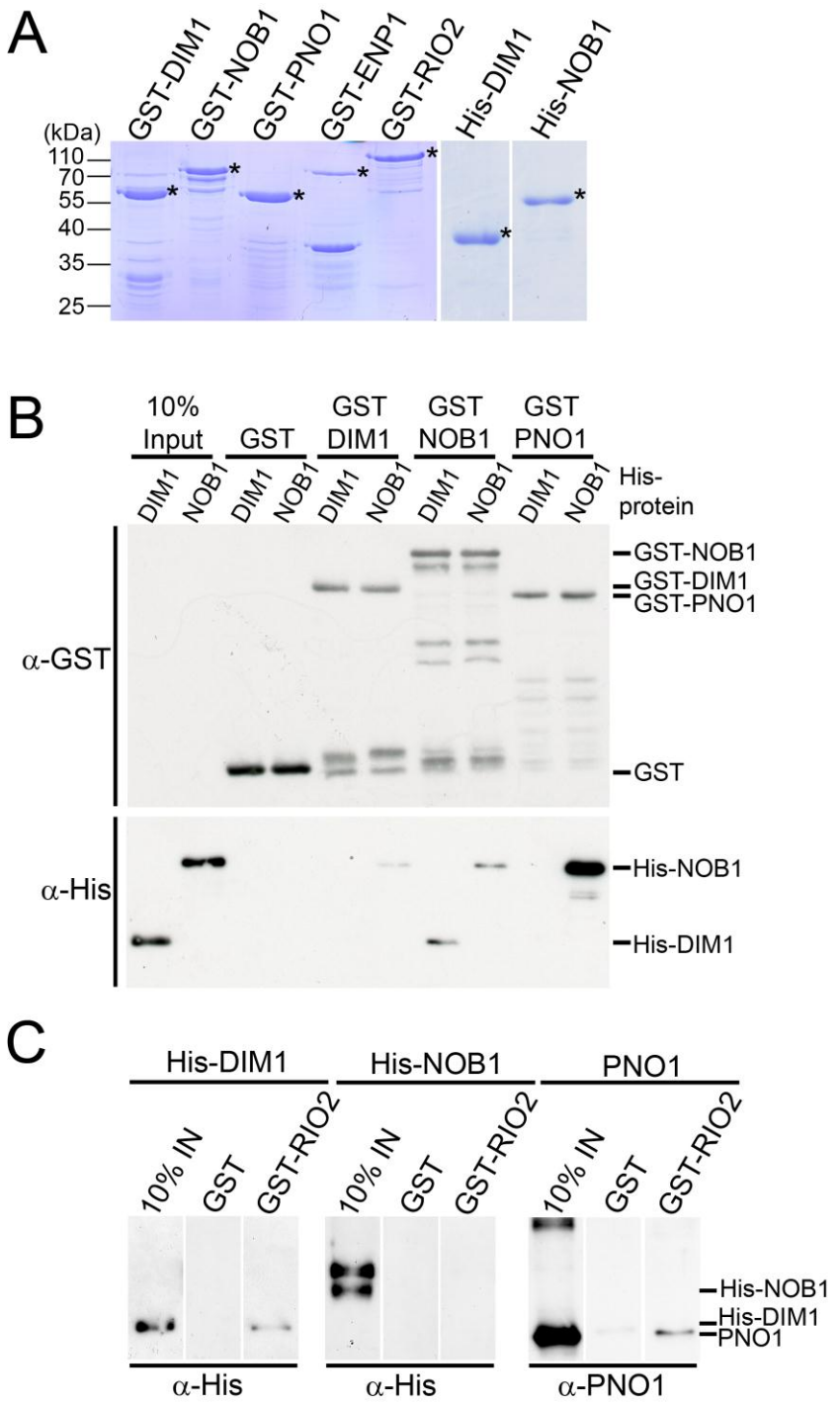


Figure 5.6 Protein purification and protein-protein interactions of late pre-40S complex factors A) Proteins of interest were expressed in *E. coli* and purified by either their GST- or His-tags as indicated with the protein names above the panel. Purified proteins were then separated by SDS-PAGE and visualised by Coomassie staining. Positions of marker proteins are indicated to the left of the panel. **B)** Recombinant GST or GST-tagged DIM1, PNO1 or NOB1 was bound to glutathione sepharose and incubated with an equal amount of His-tagged DIM1 or NOB1. Co-purified complexes were separated by SDS-PAGE and analysed by Western blotting using anti-His and anti-GST antibodies. Proteins detected are indicated to the right of the panel **C)** Equal amounts of recombinant GST and GST-RIO2 were bound to glutathione sepharose and incubated with His-DIM1, His-NOB1 or untagged PNO1. Complexes formed were separated by SDS-PAGE followed by Western blotting using either anti-His antibody (His-DIM1 and His-NOB1) or anti-PNO1 (untagged PNO1). 10 % IN refers to the non-GST tagged protein in each reaction. Proteins detected are identified to the right of the panel.

We next wanted to test the ability of human RIO2 to interact with DIM1, NOB1 and PNO1. Recombinant untagged PNO1 was produced by Prescission protease cleavage to remove the GST-tag from GST-PNO1. Untagged PNO1, His-DIM1 and His-NOB1 were incubated with GST-RIO2 and complexes formed were purified using glutathione sepharose beads. DIM1 and PNO1 both co-purified with RIO2 implying interactions between these proteins whereas NOB1 was not co-purified (Figure 5.6C). A similar experiment was performed using GST-ENP1 in place of GST-RIO2 but neither DIM1, NOB1 nor PNO1 were co-purified suggesting ENP1 does not directly interact with any of these proteins (data not shown). It was not possible to test possible interactions between RIO2 and ENP1 as both proteins were only stable as GST-tagged fusion proteins.

5.2.3.2 Recombinant SSU biogenesis factors bind to 18S rRNA

The binding sites of all these proteins, except PNO1, in yeast 18S rRNA have been determined by CRAC (Granneman et al, 2010). Based on this, RNA substrates containing the equivalent mouse 18S sites and terminating at either the 3' end of 18S or extending in ITS1 were designed (Figure 5.7A, B). The mouse sequence was used as human ITS1 is >80 % GC rich and therefore very difficult to amplify from, but since the sequence of mouse 18S is 99 % identical to that of human 18S, it was expected that interaction with human recombinant proteins would be possible. Substrate A contains helices 34, 40 and 39 including the putative NOB1 binding site. Substrate B is an extended form of A designed to include helices 30, 31, 32 and 33 which are proposed binding sites for yeast Rio2 and Enp1. Substrates C, D, E and F all contain the 3' helices, 44 and 45, and while C and D terminate at the 3' end of 18S sequence, E and F include 40 nt of 5'ITS sequence. C and E are 5' extended versions of D and F, respectively, which also including sequence from helix 28 where yeast Dim1 has been shown to bind (Figure 5.7A, B).

DNA templates for each of these substrates was amplified by PCR from a plasmid containing a mouse rDNA repeat. Forward PCR primers included a T7 promoter enabling each of these PCR products to be transcribed *in vitro* using T7 RNA polymerase, incorporating ^{32}P-labelled uracil (Figure 5.7C). Recombinant GST-tagged DIM1, NOB1, PNO1, ENP1 and RIO2 were prepared as previously described (Figure 5.6A) and GST-TIP48 (Figure 3.5A) was included as a negative control. TIP48 is a snoRNP biogenesis factor (McKeegan et al, 2007) and was not expected to interact with 18S rRNA. Recombinant proteins were bound to glutathione sepharose and RNA transcripts were

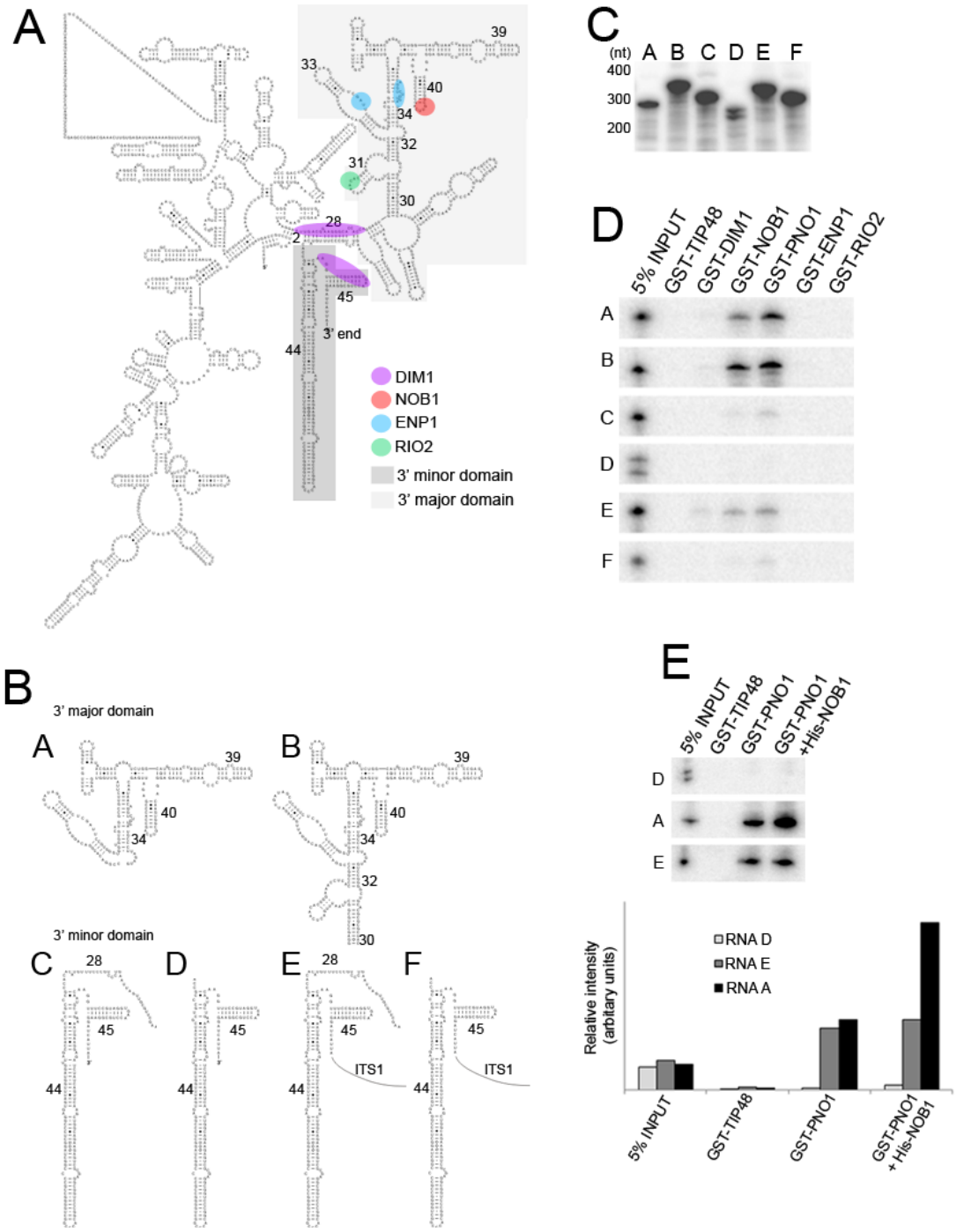


Figure 5.7 Recombinant DIM1, NOB1 and PNO1 bind 18S rRNA *in vitro* A) Outline of human 18S rRNA secondary structure (adapted from http://www.rna.ccbb.utexas.edu/). The 3' minor domain is shown on a dark grey background, the 3'major domain on a light grey background and the central and 5' domains on a white backgorund. Binding sites for proteins of interest in yeast (Granneman *et al.* 2010) are shown and colour coding is explained in the key **B)** Outline of each of the RNA substrates produced with helix numbers indicated. **C)** Specific regions of 18S were amplified by PCR and transcribed using T7 RNA polymerase to incorporate ^{32}P UTP. Transcripts were analysed using an 8 % acyrylamide/7 M urea gel and visualised using a phosphorimager. **D)** Recombinant GST-tagged TIP48, DIM1, NOB1, PNO1, ENP1 and RIO2 were bound to glutathione sepharose and incubated with RNA transcripts (5000 cpm). Complexes formed were separated on an 8 % denaturing PAGE along side 5 % RNA substrate inputs. Results were visualised using a phosphorimager. **E)** RNA-protein binding assay was carried out as in (D) using RNA substrates and proteins shown to the left and above the panel. Bound RNA was quantified using ImageQuant software and the results a shown as a bar chart.

added. Complexes formed were purified and co-purified RNA was extracted, separated by PAGE and detected using a phosphorimager (Figure 5.7D).

None of the RNA substrates were co-purified with GST-TIP48 showing that none of them interacted with this protein and implying that other interactions detected were protein-specific. GST-DIM1 co-purified the substrate E which contained 5'ITS1 sequence and the 3' end of 18S containing part of helix 28. When either the ITS1 sequence or the fragment of helix 28 was not included in the RNA substrate, DIM1 was no longer able to interact, implying that both these regions are important for DIM1 binding to pre-rRNA. Both NOB1 and PNO1 appeared to bind strongly to both substrates A and B which contain the 3' major domain of 18S. This suggests the binding site for human NOB1 is the same as that of the yeast protein and implies that the NOB1 cofactor, PNO1, also binds the same region. It is important to note that, like yeast Nob1, the primary binding site for human NOB1 does not appear to be at the site of its endonucleolytic cleavage at the 3' end of 18S. Weaker interactions were, however, observed between both NOB1 and substrates C and E and between PNO1 and substrates C, E and F suggesting that these proteins are also able to bind to the 3' minor domain of 18S rRNA in the context of the 5' end of ITS1. The most robust interactions with the 3' minor domain were observed when part of helix 28 and 5'ITS1 sequence were present. Neither RIO2 nor ENP1 bound to any of the RNA substrates tested. While it may be the case that these proteins do not directly interact with 18S rRNA in human cells, it is more likely that that the lack of interaction detected is caused by the limitations of this *in vitro* assay (see section 5.3).

DIM1, NOB1 and PNO1 all formed interactions with the 3' end of 18S and the 5' end of ITS1 making it quite conceivable that their binding to the pre-rRNA could protect the very 5' end of ITS1 and prevent the exosome from degrading ITS1 beyond the 2a site.

5.2.3.3 NOB1 and PNO1 bind cooperatively to the 3' major domain of 18S

Both NOB1 and PNO1 interacted with the 3' major domain (substrate A) and 3' minor domain (substrate E) of 18S. Strong interactions between these proteins have been demonstrated using recombinant proteins (Figure 5.7B) and also in whole cell extracts (data not shown). This raises the possibility that the RNA binding affinity of these proteins may be increased when they are in complex with each other. To test this hypothesis, the binding of PNO1 alone and PNO1 in complex with NOB1, to substrates A and E and an 18S substrate that neither NOB1 nor PNO1 bound to individually, D, was tested. RNA binding assays were performed as described above, but where

required, PNO1 and NOB1 were pre-incubated to form a complex prior to addition of the RNA substrates. As previously, GST-TIP48 was not observed to interact with any of the RNA substrates. Equivalent amounts of substrate E were co-purified with PNO1 alone and PNO1 in complex with NOB1 (Figure 5.7E). However, PNO1 was again observed to bind to substrate A but when PNO1 was in complex with NOB1, the amount of RNA co-purified increased more than two fold (Figure 5.7E). This suggests that NOB1 and PNO1 bind cooperatively to the 3' major domain of 18S but not to the 3' minor domain.

5.2.4 ENP1 stimulates the exonuclease activity of RRP6 but is not itself an exonuclease

The pre-rRNA processing defects caused by depletion of either ENP1 or PNO1 are similar to those caused by depletion of RRP6. It was, therefore, decided to investigate in more detail, the particular functions of ENP1 and PNO1 to try to determine what their roles in ITS1 processing may be and how these functions could be coordinated with exosome processing of ITS1.

Depletion of either ENP1 or the exosome component, RRP6, causes accumulation of 21SC and a decrease in 18SE levels. Co-depletion of these two proteins caused significant inhibition of exonucleolytic processing with a large accumulation of 21SC observed and very little of the downstream product, 18SE, formed (Figure 4.3). It is possible, therefore, that ENP1 functions as a cofactor of RRP6 and alters its exonuclease activity such that a greater proportion of processing can proceed beyond the natural stalling point of 21SC when ENP1 is present. To investigate this possibility, we examined the exonuclease activity of RRP6 both with and without ENP1 present.

In vitro exonuclease assays were carried out using a 30 nt 5' end-labelled poly (A) RNA substrate and recombinant RRP6, ENP1 and RRP6exo, a catalytically inactive form of RRP6 described in section 3.2.2 (Figure 5.8A). RNA was separated on a 12 % denaturing polyacrylamide gel and visualised using a phosphorimager. In the presence of RRP6, the poly (A) substrate was degraded with a significant decrease in the levels of the full length RNA and a ladder of shorter products detectable, demonstrating that RRP6 has 3'-5' exonuclease activity (Figure 5.8B). No RNA degradation was observed when RRP6exo was used, confirming that the D313A mutation inhibits the activity of RRP6. Similarly, no nucleolytic activity was detected when only ENP1 was included suggesting that ENP1 is not itself an exonuclease that directly processes ITS1 exonucleolytically (Figure 5.8B). When RRP6 and ENP1 were both present, the extent

of exonucleolytic processing by RRP6 subtly exceeded that when only RRP6 was used (Figure 5.8B, C). In the presence of RRP6 and ENP1, the full length RNA substrate was not detectable after 45 minutes and the longer 28 and 29 nt fragments were considerably weaker than when the substrate was incubated with RRP6 alone. In contrast, the shorter 26 and 27nt fragments were accumulated to a greater extent when both proteins were present than when only RRP6 was used. We concluded, therefore, that ENP1 is able to slightly stimulate the exonuclease activity of RRP6 *in vitro*, suggesting that this may be part of the function of ENP1 in the exonucleolytic processing of ITS1 *in vivo*.

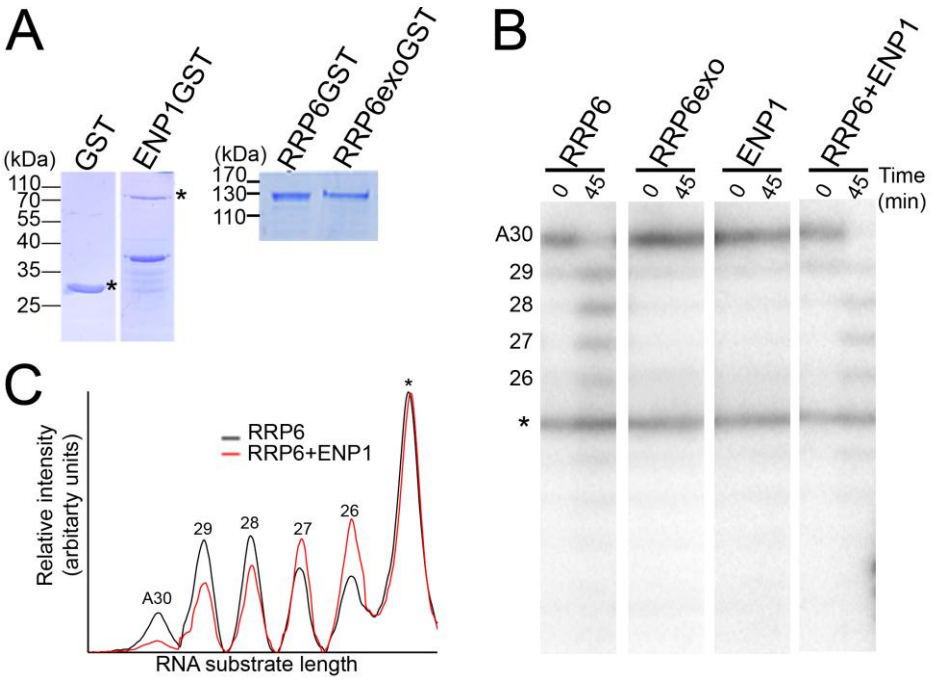


Figure 5.8 ENP1 stimulates the exonuclease activity of RRP6 *in vitro* A) GST and GST-tagged ENP1, RRP6 and RRP6exo were expressed in *E. coli* and purified. Purified proteins were separated by SDS-PAGE and visualised by Coomassie staining. The position of marker proteins is indicated to the left of the panel and full length tagged proteins are indicated by asterisks. **B)** 200 fmol of recombinant protein(s) and 10 fmol of 5' end labelled poly (A) RNA substrate were incubated for 45 min at 30 ^{o}C in Exosome buffer with samples of equal volume being taken at the beginning and end of the experiment. Reactions were separated on a 12 % polyacrylamide/8 M urea gel and visualised using a phosphorimager. To the left of the panel, number mark the length of the RNA species detected in nucleotides with A30 representing the full length substrate. **C)** Levels of processed forms of A30 substrate were determined using ImageQuant software and are shown graphically.

5.2.5 PNO1 may provide a link between ITS1 processing and nuclear export of pre-40S complexes

When PNO1 was depleted from HeLa cells, 21SC was often seen to accumulate and a slight decrease in 18SE levels was detected (Figure 5.3B, C and Figure 5.9A, D). These defects indicate a role for PNO1 in exosome processing of ITS1 following site 2 cleavage, but since the accumulation of 21SC after PNO1 depletion is

much weaker than following RRP6 or ENP1 depletion, PNO1 may only play an indirect role in this step. Depletion of PNO1 did, however, cause a striking accumulation of 26S and a weak accumulation of 43S, indicating that A_1 processing to complete removal of the 5'ETS, requires PNO1 (Figure 5.9A, D). Previously published data shows that treating cells with an inhibitor of CRM1, leptomycin B (LMB), also causes a significant accumulation of 26S (Rouquette et al, 2005). PNO1 is able to interact with CRM1 (Zemp et al, 2009) suggesting that PNO1 and CRM1 may function together in processing downstream of 26S production. Another member of the pre-40S complex that has been shown to interact with CRM1 in a Ran-GTP dependant manner is RIO2. The interaction of RIO2 with CRM1 is mediated through a leucine-rich nuclear export signal (NES) and RIO2 functions as an export-adaptor for pre-40S complexes into the cytoplasm (Zemp et al, 2009). RIO2 is not essential for this function implying that other proteins found in the pre-40S complex are also involved in nuclear export. Both PNO1 and ENP1 also contain highly conserved, but putative leucine-rich NES, (Figure 5.9B) and in yeast, deletion of Pno1 also causes nuclear retention of pre-40S complexes, making these possible candidates for additional pre-40S export adaptor proteins. However, PNO1, but not ENP1, was shown to bind to CRM1 *in vitro* and mutation of three of the key amino acids residues in the putative NES of PNO1 did not affect its CRM1 binding raising questions about whether these proteins play a role in export (Zemp et al, 2009).

Since the pre-rRNA processing defects observed following PNO1 depletion are similar to those published for cells treated with LMB, it was decided to verify whether these parallels could also be drawn when CRM1 was inhibited by LMB in our hands. Further, since depletion of PNO1 caused a slight accumulation of 21SC, suggesting that it is important for ITS1 processing, we wanted to investigate whether this was also the case when CRM1 was inhibited. HeLa cells were treated with LMB and the effects on pre-rRNA processing were examined using agarose-glyoxal gel electrophoresis followed by Northern blotting using a probe hybridising to the 5' end of ITS1. Inhibition of CRM1 with LMB caused an increase in the levels of 43S, 26S and 21SC and a decrease in 18SE levels (Figure 5.9C, D). 21SC accumulated to a similar extent as when RRP6 was depleted by RNAi, suggesting that CRM1 may be an important regulator of ITS1 processing.

Also observed when cells were treated with LMB or depleted of PNO1, was a novel intermediate, 26SC, which resembles a smear below 26S and is likely to be generated by degradation of 26S from the 3' end by exonucleolytic processing. Depletion of other factors required for removal of the 5'ETS, such as XRN2, also causes accumulation of 26S to a similar extent as depletion of PNO1 or treatment with

LMB, but in these cases, 26SC was not detected (Figure 5.8A, D and Figure 4.10). This implies that exonucleolytic processing of ITS1 normally takes place after the 5'ETS has been removed and the 26SC detected after PNO1 depletion or CRM1 inhibition arises due to premature processing of ITS1. This suggests that CRM1 and PNO1 may be required for regulating the timing of exonucleolytic processing of ITS1.

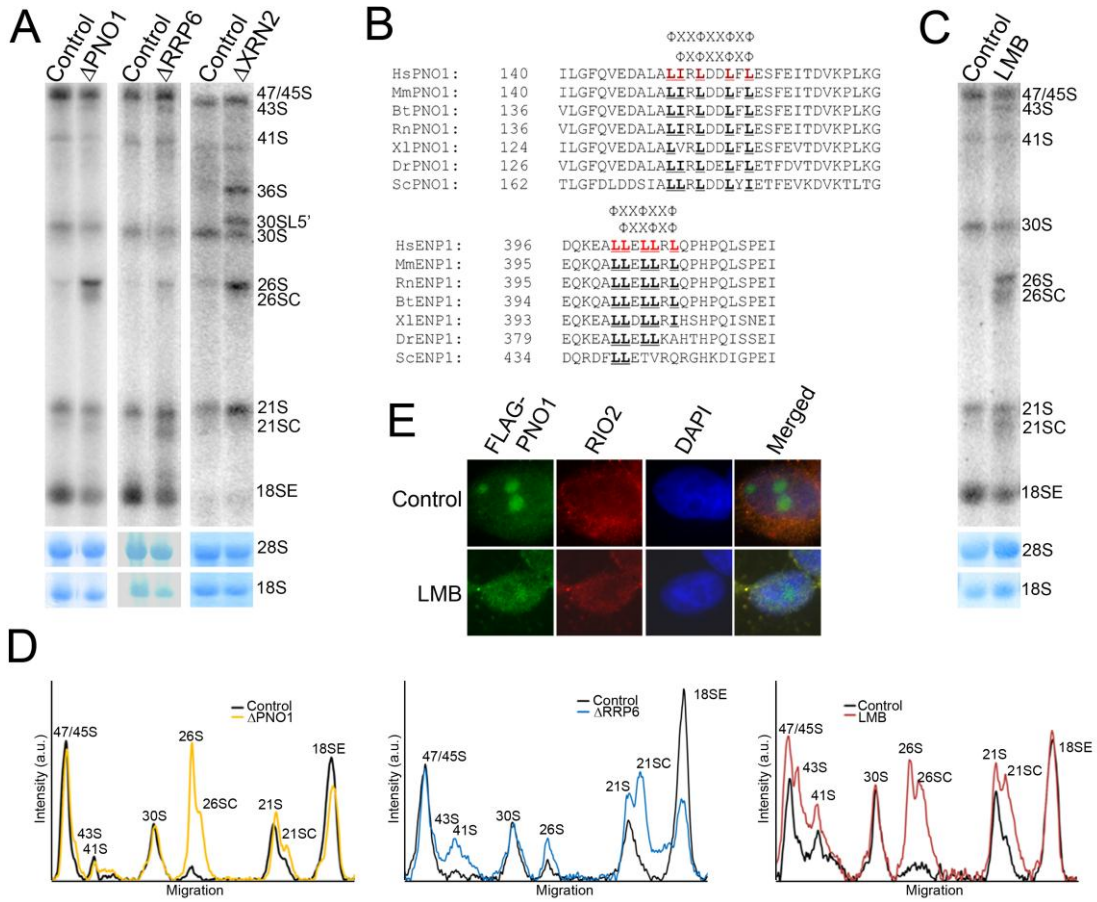


Figure 5.9 CRM1 and PNO1 may regulate the timing of ITS1 processing by the exosome A) RNA from HeLa cells depleted of PNO1, RRP6 or XRN2 was analysed by agarose-glyoxal gel electrophoresis and Northern blotting using probes hybridising to the 5' end of ITS1 (PNO1/RRP6) and the middle of ITS1 (XRN2). Both of these probes hybridise within the 3' end of 21SC/26SC. Pre-rRNA intermediates detected are indicated to the right of the panel. Mature rRNAs were visualised by methylene blue staining. **B)** Alignment of possible nuclear export sequences (NES) from human PNO1 or ENP1. Hs: *Homo sapiens*; Mm: *Mus musculus*; Bt: *Bos Taurus*; Rn: *Rattus norvegicus*; Xl: *Xenopus laevis*; Dr: *Danio rerio*; Sc: *Saccharomyces cerevisiae* were aligned using Clustral W. Φ indicates large hydrophobic amino acids and X represents any amino acid. Residues proposed to form the NES are underlined and shown in red on the human sequence. **C)** HeLa cells were treated with 30 nmol LMB for 2 h and RNA was extracted and analysed as in (A). **D)** The levels of pre-rRNAs in panels (A) and (C) were quantified using ImageQuant software and are given as a line diagram showing intensity in arbitrary units (a.u.) relative to migration through the gel. **E)** HEK293 cells stably transfected with plasmids expressing FLAG-PNO1 were grown on coverslips for 36 h with tetracycline to induce 1:1 expression of tagged and endogenous PNO1. Cells were then either treated with 30 nM LMB for 2 h of left untreated before fixing. Immunfluorescence was carried out using α-FLAG antibody to detect PNO1 (green), α-RIO2 (red) and counterstained with DAPI to visualise nuclear material (blue).

Since PNO1 is putatively linked to nuclear export of pre-40S complexes, we wanted to try to clarify whether PNO1 does function in nuclear export to determine whether the CRM1 regulation of exonucleolytic processing could be due to coordination

of pre-rRNA processing with nuclear export. As a result of its function in nuclear export, RIO2 shuttles between the cytoplasm and nucleus; treating cells with LMB blocks this, causing nuclear accumulation of RIO2. We investigated, therefore, whether the localisation of PNO1 was affected by LMB treatment. HEK293 cells expressing FLAG-tagged PNO1 were treated with LMB and the localisation of FLAG-PNO1 and endogenous RIO2 were monitored by immunofluorescence using antibodies that detected the FLAG-tag and RIO2, respectively. Cells were counter-stained with DAPI to visualise the nucleus (blue) and nucleoli were identified as “holes” within the nucleus (Figure 5.9E). In cells not treated with LMB, PNO1 localised throughout the cell, with high concentrations observed in the nucleolus, whereas RIO2 was predominantly found in the cytoplasm although some protein was also present in the nucleus. After LMB treatment both RIO2 and PNO1 were localised exclusively to the nucleus (Figure 5.9E). This is consistent with previous reports showing that RIO2 shuttling is blocked by CRM1 inhibition (Zemp et al, 2009) and demonstrates that PNO1 also shuttles in a CRM1 dependant manner. Further work is, however, required to determine whether this reflects an active role for PNO1 in nuclear export (see section 5.3).

5.2.6 The kinase activity of RIO2 is required for 18SE processing

RIO2 is a member of a family of atypical protein kinases which is found stably associated with late pre-40S complexes. The kinase activity of RIO2 is suggested to be required for conversion of 18SE into 18S and two amino acids, K123 and D246 in the RIO2 active site, have been identified as important for RIO2 activity (Zemp et al, 2009). It is possible that phosphorylation by RIO2 may be a key regulatory step in coordinating the final steps of 18S maturation both in terms of dissociation of biogenesis factors (Zemp et al, 2009) and pre-rRNA processing (exonucleolytic processing and site 3 cleavage). Direct substrates of RIO2 kinase activity have not been identified. We aimed, therefore, to confirm the requirement for the kinase activity of RIO2 in processing of 18SE and try to identify possible targets of this activity.

To verify the importance of the kinase activity of RIO2 in 18S production, HEK293 stable cell lines were constructed in which endogenous RIO2 could be replaced with either FLAG-tagged wild-type or kinase-inactive RIO2. The kinase activity of RIO2 was inhibited by modification of the two previously identified amino acids in the catalytic site, lysine 123 and aspartic acid 246, to alanine to generate RIO2kd (kinase dead). The siRNA duplex used to deplete endogenous RIO2 targets the 3’ untranslated region (UTR) of the mRNA enabling the tagged forms of RIO2 to be expressed irrespective of siRNA treatment. To determine the concentration of tetracycline required

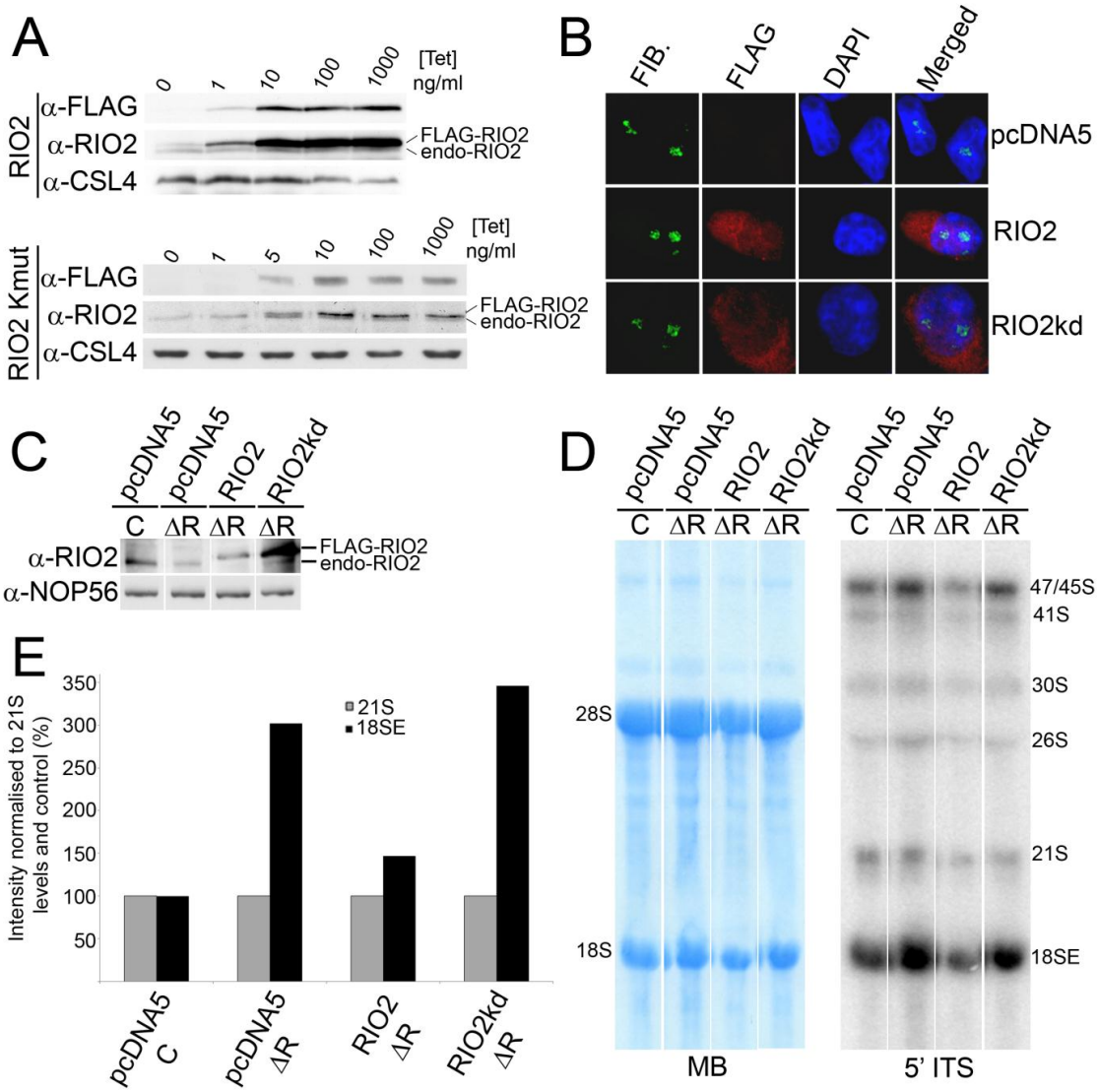


Figure 5.10 The kinase activity of RIO2 is required for conversion of 18SE to 18S A) HEK293 cells stably transfected with plasmids expressing FLAG-RIO2 or FLAG-RIO2kd were cultured in the presence of tetracycline at different concentrations for 36 h before harvesting. Proteins were extracted and analysed by SDS-PAGE followed by Western blotting using the antibodies shown to the left of the panel. Tagged and endogenous proteins are identified **B)** pcDNA5, RIO2 and RIO2kd cells were grown on cover-slips and induced with 1 ng tetracycline for 36 h, fixed and assayed by immunofluorescence using antibodies, α-fibrillarin (FIB.) (green), α-FLAG (red) and DAPI (blue). **C)** HEK293 cells expressing just the FLAG-tag (pcDNA5), FLAG-RIO2 or kinase inactive FLAG-RIO2kd were transfected with a control siRNA (C) or a siRNA duplex targeting the 3'UTR of RIO2 mRNA (ΔR). The levels of endogenous and FLAG-tagged RIO2 were monitored by Western blotting using antibodies against RIO2 and NOP56 as a loading control. **D)** RNA from these cells was analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using a probe hybridising to the 5' end of ITS1. **E)** 21S and 18SE levels were quantified using ImageQuant software. 21S levels were normalised and both 21S and 18SE levels are shown as a bar chart.

to induce expression of the tagged proteins to the same level as the endogenous protein, expression of both FLAG-RIO2 and FLAG-RIO2kd was induced over a range of tetracycline concentrations. Protein expression was monitored by Western blotting using antibodies to detect the FLAG-tag, RIO2 or the exosome protein, CSL4. Expression of FLAG-RIO2 exceeded levels of the endogenous protein when 1 ng/ml tetracycline was used and expression of FLAG-RIO2kd equalled endogenous protein

levels at 5 ng/ml tetracycline (Figure 5.10A). Expression of endogenous RIO2 was suppressed by expression of the tagged forms of the protein as observed when tagged RRP6 was expressed (Figure 3.11).

RIO2 is localised predominantly to the cytoplasm but is also detected in the nucleus (Figure 5.9E). We wanted to confirm that expression of FLAG-tagged or inactive forms of RIO2 did not affect localisation of the protein as this could prevent it from carrying out its normal functions. Cells were grown on cover-slips and expression of FLAG-RIO2 or FLAG-RIO2kd was induced using the optimal concentration of tetracycline for 48 h before fixing the cells. Immunofluorescence was performed using antibodies to detect the FLAG-tag and the snoRNP protein, fibrillarin which is a nucleolar marker, and the nucleoplams was visualised using DAPI. This showed that the FLAG-tagged forms of RIO2 were found mostly in the cytoplasm but were also present in the nucleus, as the endogenous protein is (Figure 5.10B).

Cells expressing FLAG-RIO2 and FLAG-RIO2kd were transfected with siRNAs targeting RIO2 and after 48 h, the expression of RIO2 was analysed by Western blotting. In parallel, control cells expressing only the FLAG-tag (pcDNA5) were transfected with either a control siRNA or one targeting RIO2. In control cells this showed that endogenous RIO2 was successfully depleted from HEK293 cells by the siRNA treatment. In FLAG-RIO2 and FLAG-RIO2kd cells, only the tagged forms of RIO2 were detected at, or in the case of RIO2kd, slightly above, the level of the endogenous protein (Figure 5.10C). RNA was extracted from these HEK293 cells and analysed by agarose-glyoxal gel electrophoresis followed by Northern blotting using a probe hybridising to the 5' end of ITS1. As seen before (Figure 5.3) in control cells, depletion of RIO2 did not affect the levels of the longer pre-rRNAs relative to cells transfected with the control siRNA but 18SE levels were significantly increased. In cells where the endogenous protein had been replaced by FLAG-RIO2, the levels of 18SE were restored almost to the level observed in control cells showing that the defect caused by depletion of RIO2 could be rescued by expression of the tagged wild-type protein (Figure 5.10D, E). However, in cells where RIO2 had been replaced by FLAG-RIO2kd, 18SE levels were three and a half fold higher than those of control cells (Figure 5.10D, E). This supports the previously published data (Zemp et al, 2009) and clearly demonstrates that the kinase activity of RIO2 is required for the conversion of 18SE to 18S. Our data show a three fold increase in 18SE levels when RIO2 is depleted and expression of FLAG-RIO2 restores 18SE levels to only 110% of those in control cells (Figure 5.10E). This suggests that stably expressed FLAG-tagged RIO2 is better able to replace the endogenous protein than the transiently transfect EGFP-tagged RIO2 used in previously published work (Zemp et al, 2009) underlining the

importance of the kinase activity of RIO2 in conversion of 18SE into 18S. The fact that a stronger accumulation of 18SE is detected in cells expressing FLAG-RIO2kd than in control cells depleted of RIO2 may be explained by the observation that expression of the tagged form of the protein suppressed expression of the endogenous protein enhancing the siRNA depletion of the endogenous protein.

5.2.7 RIO2 phosphorylates itself and DIM1 *in vitro*

We next aimed to identify proteins that may be targets of RIO2 phosphorylation. In *S. cerevisiae*, RIO2 is capable of serine phosphorylation *in vitro* (Angermayr & Bandlow, 2002; Geerlings et al, 2003; Vanrobays et al, 2003; Vanrobays et al, 2001). In *A. fulgidus*, RIO1, another member of the RIO family of kinases, which is also needed for 18S production but is not stably associated with pre-40S complexes, has been shown to autophosphorylate at serine 108 in a flexible loop region (LaRonde-LeBlanc & Wlodawer, 2005b). The *in vitro* kinase activity of human RIO2 has not been investigated and it is not known whether this protein is also capable of autophosphorylation and whether other targets of its kinase activity exist. It is possible that autophosphorylation of RIO2 may be required to stimulate its own kinase activity sufficiently to phosphorylate another substrate. RIO2 is able to interact with several of the late-acting SSU biogenesis factors, DIM1, PNO1, NOB1 (Figure 5.5 and Figure 5.6) and the binding sites for yeast RIO2 and ENP1 on 18S rRNA are very close to each other making all of these proteins likely targets of RIO2 phosphorylation. Large scale proteomic analysis of phosphorylated proteins indicates that DIM1, NOB1, PNO1 and ENP1 are all phosphorylated proteins (www.uniprot.org). If any of these proteins are substrates of RIO2 kinase activity, then phosphorylation by RIO2 may be a key regulatory step in formation of the 3' end of 18S. We therefore tested the *in vitro* kinase activity of recombinant human RIO2 both for its ability to phosphorylate itself and also recombinant DIM1, NOB1, PNO1 and ENP1. Since we have demonstrated that NOB1 and PNO1 form a stable complex *in vitro* (Figure 5.6B) and that this complex forms a more robust interaction with 18S rRNA than either protein alone (Figure 5.7E), we also tested the ability of RIO2 to phosphorylate these proteins in complex as well as individually.

GST-RIO2 and GST-RIO2kd which contained the K123A-D246A mutation discussed in the previous section were over-expressed in *E. coli* and purified using their tags. Purification of DIM1, PNO1, NOB1 and ENP1 was carried out as previously (Figure 5.11). Kinase assays were performed by combining equal amounts of either RIO2 or RIO2kd with each of the potential substrates in the presence of ^{32}P labelled

phosphate and 0.01 mM unlabelled ATP to provide sufficient phosphate to enable the reaction to proceed and drive incorporation of the labelled phosphate. After 30 min incubation, proteins were separated by SDS-PAGE and phosphorylated proteins were detected using a phosphorimager. A ten fold excess of each protein input into the assay was also analysed by SDS-PAGE followed by Coomassie staining to enable easy identification of the phosphorylated proteins. ^{32}P labelled RIO2 was strongly detectable in each reaction whereas significantly less ^{32}P labelled RIO2kd was detected, showing that RIO2 is able to phosphorylate itself and that this kinase activity is significantly impaired by mutation of K123 and D246 to alanine (Figure 5.11). No ^{32}P labelled NOB1, PNO1, or ENP1 was detected after incubation of these proteins with either RIO2kd or RIO2 implying that these proteins are not substrates of RIO2 phosphorylation *in vitro*. In contrast, ^{32}P-labelled DIM1 was detected after incubation with RIO2 but not with RIO2kd suggesting that DIM1 is phosphorylated by RIO2 (Figure 5.11). A non-specific product (X) was also observed to incorporate ^{32}P but as this did not correspond to the sizes of any of the substrate proteins, it was disregarded.

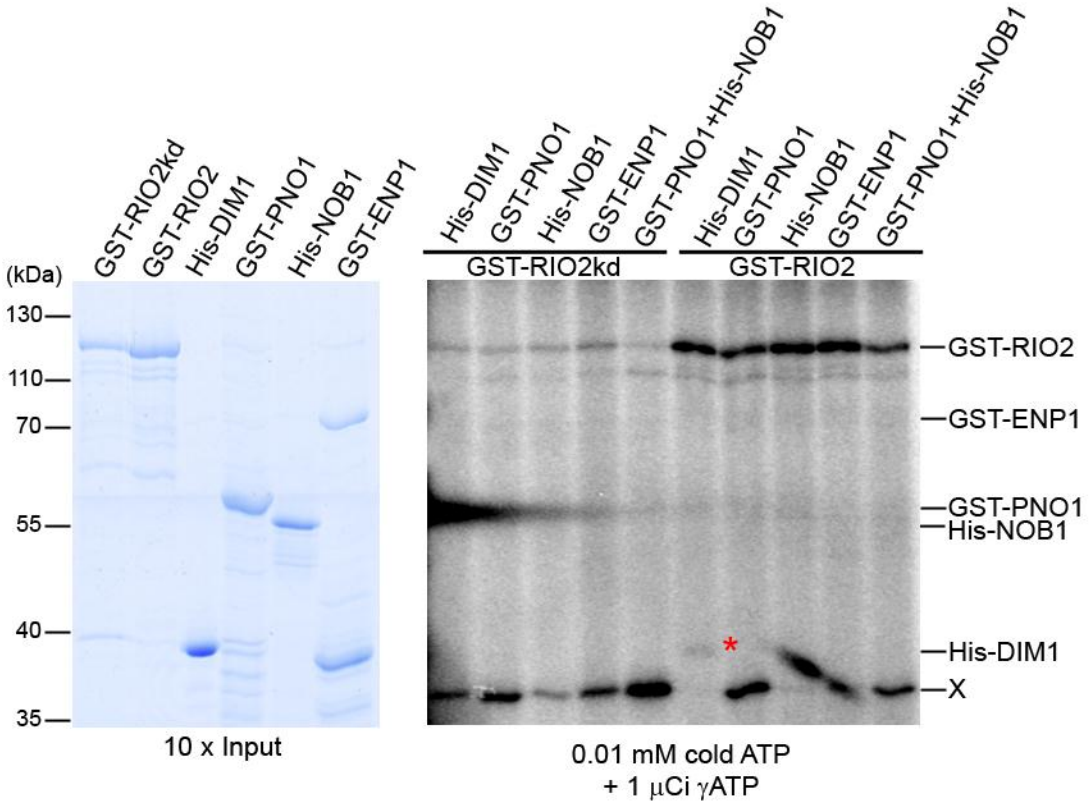
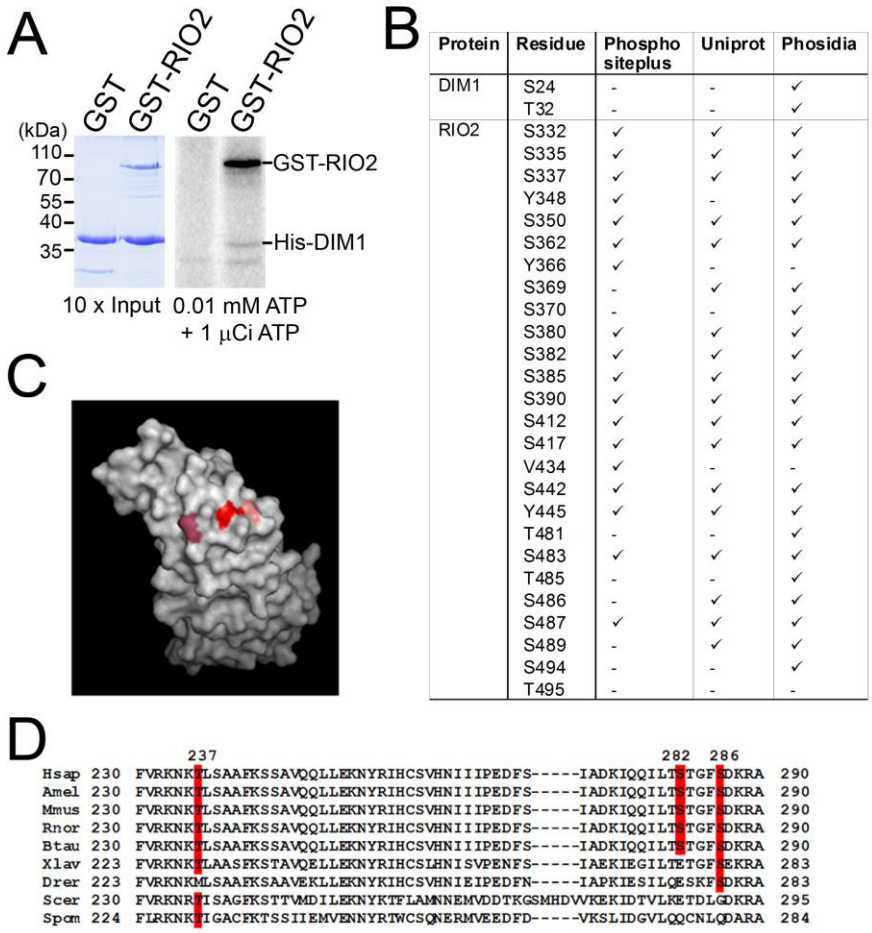


Figure 5.11 RIO2 kinase phosphorylates itself and DIM1 *in vitro* Recombinant proteins of interest were expressed in *E. coli* and purified using either their GST- or His-tags. Phosphorylation assays were carried out by combining RIO2 or RIO2kd with either DIM1, NOB1, PNO1, ENP1 or NOB1 and PNO1 in complex as indicated above the central and right panels.). Reactions were carried out at 30 ^{o}C for 30 min in Purification Buffer containing 0.01 mM ATP and 1 µCi ^{32}PγATP. Samples were analysed by SDS-PAGE and phosphorylated proteins were visualised using a phosphoimager. The positions of each of the potential substrates are given to the right of the panel. 10 x protein inputs were separated by SDS-PAGE and visualised by Coomassie staining (left panel) and the positions of marker proteins are indicated to the left. X indicates a contaminant phosphorylated band which was disregarded. Red asterisk indicates phosphorylated His-DIM1.

5.2.8 Identification of phosphorylation sites in RIO2 and DIM1

As we have demonstrated that human RIO2 is capable of phosphorylating both itself and DIM1 *in vitro* (Figure 5.12A), we next wanted to identify the amino acid residues being phosphorylated. Information about the phosphorylation states of proteins is available in databases including Phosidia (http://www.phosida.com/), Phosphositeplus (http://www.phosphosite.org/homeAction.do) and Uniprot (http://www.uniprot.org/). These data are predominantly collected from large scale proteomic screens and may not represent a full or accurate picture of the important phosphorylated amino acid residues in these proteins. A list of the potential phosphorylation sites in both RIO2 and DIM1 is given in Figure 5.12B.

To identify which, if any, of these sites are the targets of RIO2 phosphorylation, proteomic analysis was carried out. Kinase reactions were performed in which RIO2 and DIM1 were combined both with and without 1 mM ATP. After incubation, proteins were separated on pre-cast NuPAGE protein gels to minimise any contamination. Gels were stained with Coomassie and proteomic analysis carried out in the laboratory of Dr Henning Urlab (Max-Planck Institute, Gottingen, Germany). The results obtained identified thirteen sites in DIM1 as being phosphorylated but nine of these were also identified in reactions carried out in the absence of ATP. This would suggest that these phosphorylations are non-specific or occurred in *E. coli* during preparation of the recombinant protein and they were discounted. The four residues identified as being specifically phosphorylated in reactions containing RIO2 and ATP were: S286, T237, T179 and S282. DIM1 is highly conserved and a crystal structure of Dim1 from the archaeon, *Methanocaldococcus Jannischi*, has been published (Pulicherla et al, 2009). Human DIM1 was modelled on this structure using software available at http://swissmodel.expasy.org and this structure is shown as a surface view model in Figure 5.12C. The four putative RIO2 phosphorylation site residues were identified and are shown marked in shades of red. S286, T237 and S282 were found to be clustered on the surface of DIM1 suggesting that this face of the molecule may represent a target of RIO2 phosphorylation. T179 was, however, buried within the protein molecule. It is possible that the recombinant DIM1 used in these experiments is only partially folded making phosphorylation of this residue possible *in vitro* while it is unlikely to represent a site of RIO2 phosphorylation *in vivo*. Multiple sequence alignment of DIM1 proteins from a range of species showed that the three surface amino acid residues identified are conserved in mammals while only T237 is also conserved in yeast. S286 is conserved in the vertebrates, *X. Laevis* and *Danio rerio*, while S282 is not (Figure 5.12D).



Protein	Residue	Phospho siteplus	Uniprot	Phosidia
DIM1	S24	-	-	✓
	T32	-	-	✓
RIO2	S332	✓	✓	✓
	S335	✓	✓	✓
	S337	✓	✓	✓
	Y348	✓	-	✓
	S350	✓	✓	✓
	S362	✓	✓	✓
	Y366	✓	-	-
	S369	-	✓	✓
	S370	-	-	✓
	S380	✓	✓	✓
	S382	✓	✓	✓
	S385	✓	✓	✓
	S390	✓	✓	✓
	S412	✓	✓	✓
	S417	✓	✓	✓
	V434	✓	-	-
	S442	✓	✓	✓
	Y445	✓	✓	✓
	T481	-	-	✓
	S483	✓	✓	✓
	T485	-	-	✓
	S486	-	✓	✓
	S487	✓	✓	✓
	S489	-	✓	✓
	S494	-	-	✓
	T495	-	-	-

Figure 5.12 Sites of RIO2 phosphorylation in DIM1 A) His-DIM1 was mixed with either GST-alone or GST-RIO2 and 0.01 mM ATP/1 µCi ^{32}P ATP was added. Reactions were incubated at 30 ^{o}C for 30 min. Proteins were separated by SDS-PAGE alongside 10 x protein inputs. Inputs were visualised by Coomassie staining and phosphorylated proteins were detected using a phosphorimager. **B)** Phosphorylation site in DIM1 identified in different databases **C)** Structure of human DIM1 modelled on crystal structure of *Methanocaldococcus Jannischi* DIM1 (http://swissmodel.expasy.org) visualised using PyMol software as a surface view model. Amino acid residues identified as RIO2 kinase phosphorylation sites are shown in shades of red. **D)** Sequences of homologues of human DIM1 were retrieved using Blastp and aligned using ClustralW. Phosphorylation sites in human DIM1 are shown in red and are also marked in red in species where they are conserved. Hsap: *Homo sapiens*; Amel: *Ailuropoda melanoleuca;* Mmus: *Mus musculus*; Rnor: *Rattus norvegicus;* Btau: *Bos Taurus*; Xlav: *Xenopus laevis*; Drer: *Danio rerio*; Scer: *Saccharomyces cerevisiae*; Spom: *Schizosaccharomyces pombe.*

A similar approach was undertaken to try to identify the site(s) of autophosphorylation in RIO2. In this case eleven sites were identified as being phosphorylated in our samples but of these, four were specific to reactions in which ATP was present. The four potential autophosphorylation sites were: S107, S369, S382 and S457. It was not possible to generate a complete model structure of human RIO2 onto which these residues could be mapped as the only available crystal structure of RIO2 is from the archaea, *A. fulgidus*, (LaRonde-LeBlanc & Wlodawer, 2004) and this protein is missing a large C-terminal domain found in the human protein. It was interesting to note however, that the site identified in *A. fulgidus* RIO2 as being

autophosphorylated, serine 108, was not found to be phosphorylated in the human protein.

5.3 Discussion

In the previous chapter, we showed that removal of ITS1 sequence from the 3' end of 18S is achieved by exonucleolytic processing by the exosome to within 25 nt of the mature 3' end. We therefore, screened late-acting SSU biogenesis factors for possible cofactors and regulators of ITS1 processing. We have investigated further the role of ENP1 as a cofactor of RRP6 in ITS1 processing and have found that PNO1 and CRM1 are both important for ITS1 processing and in particular regulating the timing of this processing. DIM1, NOB1 and PNO1 were all shown to bind to RNA substrates containing the 3' end of 18S and 5' end of ITS1 suggesting that these proteins may be able to terminate exosome processing of ITS1 at the 3' end of 18SE. DIM1 has been identified as a substrate of RIO2 kinase activity *in vitro*, establishing a platform from which the effect of this phosphorylation can be further investigated in the future. Further, it appears that the function of NOB1 in site 3 cleavage is conserved from yeast to humans but in contrast to yeast, we find no role for the human homologue of Prp43 in 18S biogenesis (Figure 5.13).

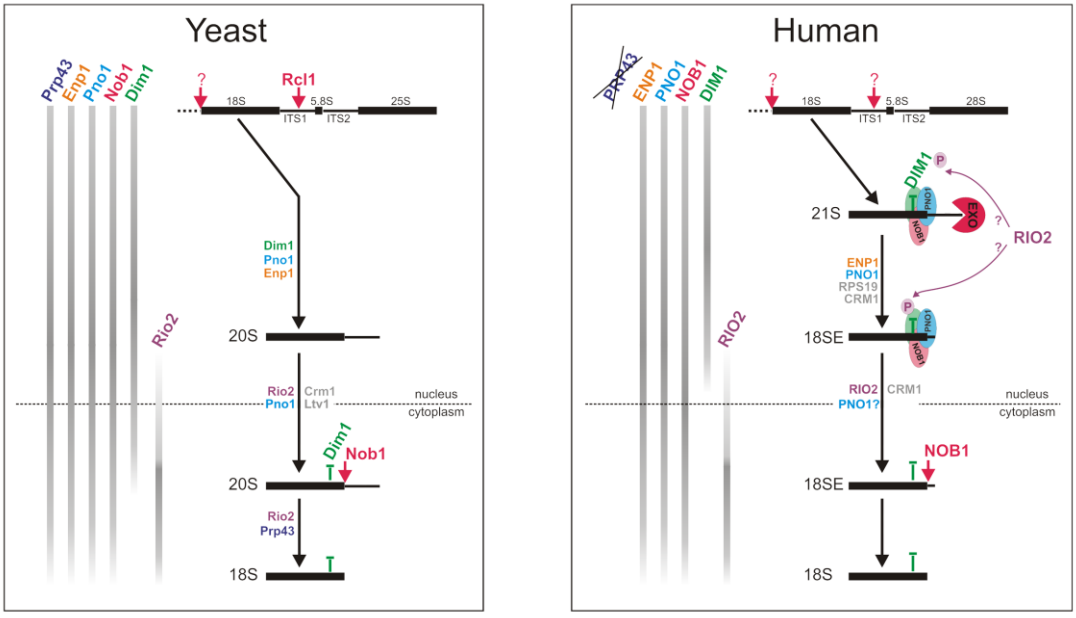


Figure 5.13 Formation of the 3' end of 18S Schematic outline of the association and roles of factors involved in formation of the 3' end of 18S in yeast and humans. pre-rRNA intermediates are shown in black with black arrows indicating processing. Nucleases (Rcl1, NOB1 and the exosome) are shown in red and red arrows indicated endonucleolytic cleavages. Purple "P" and arrows indicate phosphorylation of DIM1 by RIO2 kinase. Green "T" denotes methylation by DIM1. Grey bars to the left and right of the panel indicate association and dissociation of the proteins named above them with pre-40S complexes.

The SSU biogenesis factor, ENP1, is important for ITS1 processing from site 2 to 2a and depletion of ENP1 causes accumulation of 21SC and a decrease in 18SE levels. Since exonucleolytic processing of ITS1 is specific to higher eukaryotes, we can conclude that ENP1 has a very different role in human pre-rRNA processing than in yeast. Indeed, in yeast, Enp1 is required for removal of the 5'ETS (Chen et al, 2003) but our results support those of Carron et al. (2010), showing that depletion of ENP1 from human cells does not affect cleavages at A', A_0 or A_1. The N-terminal region of ENP1 is poorly conserved between yeast and humans so it is possible that this N-terminal is required for the novel functions of ENP1 in human pre-rRNA processing. Although Enp1 in yeast has been shown to cross-link to the 3' end of 18S *in vivo* (Granneman et al, 2010), human ENP1 was not found to bind to any of the regions of 18S tested *in vitro*. While our assay may not reflect all the interactions of ENP1 in the cell, it is possible that the different functions of ENP1 in human cells means that human ENP1 does not bind to 18S rRNA. No specific molecular function has been attributed to ENP1 in yeast or in human cells. Here we have shown that ENP1 is able to subtly stimulate the exonucleolytic activity of RRP6 *in vitro*. This suggests that one of the roles of ENP1 in exonucleolytic processing of ITS1 may be to regulate the processing activity of the exosome. It is also possible that ENP1 is involved in recruiting the exosome to the pre-rRNA since depleting ENP1 causes a processing defect comparable to that of depleting RRP6. In the previous chapter, depletion of RPS19 was also shown to cause accumulation of 21SC. In yeast, Rps19 is required for the association of Enp1 with 35S (Leger-Silvestre et al, 2005), so it could be speculated that, in human cells, these proteins are assembled onto the pre-rRNA in the same manner and this in turn recruits the exosome.

Factors associated with late-pre-40S complexes may also play important roles in regulating ITS1 processing by the exosome. Depletion of PNO1 or, more significantly, inhibition of CRM1, affected exonucleolytic processing of ITS1 with increased levels of 21SC observed in both circumstances. Exonucleolytic processing of ITS1 normally occurs once the 5'ETS has been removed. Depletion of PNO1 or inhibition of CRM1 inhibits cleavage at A_1, indicated by a significant increase in 26S levels. Unlike depletion of other proteins required for complete removal of the 5'ETS, depletion of PNO1 or treatment with LMB causes 26S to undergo premature processing forming 26SC. This suggests that PNO1 and CRM1 may function to regulate the timing of ITS1 processing by the exosome.

As yet, it is not entirely clear how the functions of PNO1 and CRM1 are coordinated so it is only possible to speculate on how CRM1/PNO1 regulation of ITS1 processing may be linked to other events in the formation of the 3' end of 18S. The

major function of CRM1 in ribosome biogenesis is mediating the export of pre-40S from the nucleus into the cytoplasm, raising the possibility that ITS1 processing is linked to nuclear export. The only protein which has been clearly demonstrated to function as an export adaptor for pre-40S complexes, through its interaction with CRM1, is RIO2 (Zemp et al, 2009). It is important to note that unlike treating cells with LMB, the only pre-rRNA processing defect caused by inhibiting export by depleting RIO2 was an accumulation of 18SE, implying that the parallel defects observed after LMB treatment and depletion of PNO1, may arise from an export-independent function of CRM1. There is a precedent for this as other functions of CRM1 have been proposed including the nucleolar localisation of snoRNP proteins (Boulon et al, 2004). RIO2 is not essential for pre-40S complex export and other proteins in this complex are also proposed to have roles in export. Like RIO2, PNO1 directly interacts with CRM1 and contains a leucine rich NES. However, mutation of thee amino acids of this motif in the human recombinant protein did not inhibit interaction with CRM1 raising confusion as to the role of PNO1 as an export-adaptor for pre-40S complexes (Zemp et al, 2009). Deletion of the comparable NES region from the yeast protein leads to retention of pre-40S complexes in the nucleus (Vanrobays et al, 2008). This may suggest that other amino acids in this leucine-rich sequence in PNO1 are sufficient to enable CRM1 binding and that the correct motif for mediating this interaction was not fully identified. To try to clarify whether PNO1 does function as an export adaptor, in the future it is intended to express a series of truncated forms of PNO1, mutate the extended NES identified and investigate the ability of these protein fragments to interact with CRM1. It is hoped that the sequence required for the interaction between PNO1 and CRM1 could be identified, enabling the effect of blocking this interaction on the export of pre-40S complexes to be determined *in vivo*.

Recently, a broad screen of chemotherapeutic agents which inhibit various stages of ribosome biogenesis was published (Burger et al, 2010). This identified several compounds which affect late processing steps and significantly decrease 18S levels: homoharringtonine and cyclohexamide which are both classified as translation inhibitors and 5-fluorouracil which is an antimetabolic agent. It may prove informative to investigate the pre-rRNA processing defects arising in cells treated with these chemicals. The exosome component, RRP6, is a target of 5-fluorouracil (Fang et al, 2004; Silverstein et al, 2011), potentially indicating that this chemotherapeutic drug functions by inhibiting ITS1 processing. Since homoharringtonine and cyclohexamide affect late processing steps and are inhibitors of translation, investigating cells treated with these compounds may give insights into how completing 18S processing is linked to subunit joining and initiation of translation.

The majority of ITS1 sequence is removed from the 3' end of 18S by exonucleolytic processing in human cells, rather than by a series of endonucleolytic cleavages as it is in yeast. This raised the questions of how the exosome is arrested near to the 3' end of 18S and why the final 20-30 nt of ITS1 are not removed. Here we have shown that human DIM1, NOB1 and PNO1 are all able to bind to sequences at the 3' end of 18S that include the 5' end of ITS1. The presence of ITS1 sequence significantly increased interactions between these proteins and the RNA substrates *in vitro* implying that these proteins either bind to the 5' end of ITS1 or that this region is important for stabilising their interactions with the 3' end of 18S. It is possible, therefore, that the presence of these proteins in pre-ribosomal complexes may prevent the exosome accessing the final 20-30 nt of ITS1, thereby protecting them from exonucleolytic degradation (Figure 5.13). Neither RIO2 nor ENP1 bound to 3' 18S fragments *in vitro* but the position of yeast Rio2 in the recently published cryo-EM structure of an *S. cerevisiae* pre-40S complex (Strunk et al, 2011) would suggest that this protein could also participate in blocking exosome access to the 5' end of ITS1. While it is possible that neither protein interacts with 18S in humans, this is unlikely. The lack of binding is more probably caused by a low affinity of RNA binding or could reflect the fact that these proteins function as part of a complex and may, therefore, not be able to form the same interaction individually. Indeed, we have shown that recombinant human RIO2 was able to interact with NOB1, PNO1 and DIM1, suggesting RIO2 forms contacts with several proteins within the pre-40S complex and it is possible that these interactions are necessary for RIO2 binding to 18S rRNA. Yeast Rio2 cross linked to helix 31 and it has been suggested that modification of U1191 in this helix is important for RIO2 binding to this helix (Granneman et al, 2010). In this experiment, RNA substrates were transcribed *in vitro* and did not contain such modifications, which may explain why RIO2 was not observed to interact with this helix.

One of the aims of this work was to identify ways in which ITS1 processing may be controlled so we investigated the kinase activity of RIO2 as phosphorylation is a common regulatory mechanism. It was of particular interest to determine if RIO2 was able to phosphorylate either of the SSU biogenesis factors shown to be involved in ITS1 processing, ENP1 or PNO1, but neither of these proteins were substrates of RIO2 kinase *in vitro*. We are able to reinforce previously published data showing that the kinase activity of RIO2 is required for the processing of 18SE to 18S (Zemp et al, 2009). This suggests that RIO2 kinase functions downstream of exonucleolytic processing to form 18SE. The kinase activity of RIO2 is also required for the dissociation of NOB1, PNO1 and LTV1 from the pre-40S complex in the cytoplasm

(Zemp et al, 2009) implying that the kinase activity of RIO2 is needed both before and after site 3 cleavage.

RIO2 is a highly phosphorylated protein and we have shown that it is capable of autophosphorylation. It is possible that this represents a feedback system by which RIO2 is able to regulate its own kinase activity as is suggest for *E. coli* Rio1 (Angermayr & Bandlow, 2002). Since RIO2 shuttles between the nucleus and cytoplasm, it is possible that the phosphorylation state of RIO2 is altered depending on its sub-cellular localisation. Our data show that RIO2 can phosphorylate DIM1 *in vitro* and potential sites of this phosphorylation have been identified. Although the function of DIM1 as the dimethyltransferase responsible for modification of two adenosine residues at the 3' end of 18S is conserved in human cells, it appears that human DIM1 may perform this function much earlier in the processing pathway than its yeast counterpart (Watkins lab, unpublished). Yeast Dim1 associates with early pre-ribosomal complexes although methylation does not occur until the pre-40S complex reaches the cytoplasm (Lafontaine et al, 1998b). However, human DIM1 is not localised to the cytoplasm (Watkins lab, unpublished) implying that methylation occurs in the nucleus and then DIM1 dissociates from pre-40S complexes. Our data underline the importance of methylation of these two residues, as depletion of DIM1 causes accumulation of partially degraded pre-rRNAs. These degradation intermediates are derived from early precursors, 43S or 41S, making it likely that DIM1 associates with early pre-ribosomal complexes. This suggests that pre-rRNAs that are either not associated with DIM1 or that have not undergone methylation are rapidly detected as aberrant and are targeted for degradation. FLAG-DIM1 is able to precipitate RIO2 from cell extracts confirming that these proteins are found together in pre-40S complexes and recombinant RIO2 and DIM1 also interact *in vitro*, implying that these proteins form direct contacts. Further work is required to determine whether RIO2 phosphorylates DIM1 *in vivo* but these data imply that DIM1 and RIO2 are associated with pre-40S complexes in the nucleus and raise the possibility that after DIM1 methylation, phosphorylation by RIO2 causes DIM1 to dissociate from the pre-40S complex. It is interesting to note that a greater amount of RIO2 was precipitated from cell extracts by FLAG-PNO1 or FLAG-NOB1 than by FLAG-DIM1 possibly indicating that pre-40S complexes containing both DIM1 and RIO2 are transitory. Dissociation of Dim1 from pre-ribosomes in yeast is an important step since the binding site in 18S rRNA identified by CRAC is at the mRNA decoding centre and overlaps with sequences important for tRNA interactions meaning that Dim1 must dissociate before translation could be initiated (Granneman et al, 2010). However, since human DIM1 appears to dissociate from pre-40S complexes in the nucleus rather than cytoplasm, it is possible

that release of DIM1 is important for forming pre-40S complexes that can be exported. The kinase activity of RIO2 is only required for the conversion of 18SE into 18S and if it is the case that RIO2 phosphorylation of DIM1 causes it to dissociate, it is possible that the presence of human DIM1 in late-pre-40S complexes either directly or indirectly prevents cleavage at site 3. To test these ideas it is intended to mutate the identified DIM1 phosphorylation sites both to inhibit phosphorylation and mimic constitutive phosphorylation. We speculate that RIO2 phosphorylation may result in DIM1 dissociation from pre-40S complexes and as we have demonstrated binding of recombinant DIM1 to the 3' end of 18S, we aim to determine the effect of these mutations of the RNA binding ability of DIM1 *in vitro*. It is also intended to investigate the effects of this phosphorylation *in vivo* by modifying the HEK293 stable cells expressing FLAG-RIO2 and kinase-dead FLAG-RIO2 used in this study to enable the association of DIM1 with pre-40S complexes to be assessed. We and others have shown that the kinase activity of RIO2 is required for the conversion of 18SE into 18S so we would like to determine if this is a downstream consequence of DIM1 phosphorylation or whether other substrates of RIO2 kinase that are important for this step could also be identified. It would be of interest to investigate possible pre-40S complex substrates of the other RIO kinase required for 18S production, RIO1, and determine if phosphorylation of DIM1 is specifically achieved by RIO2. We have shown that RIO2 is capable of autophosphorylation and putative phosphorylation sites have been identified. We hope to determine if this represents a mechanism by which the kinase activity of RIO2 is regulated.

Examining the pre-rRNA processing defects caused by depletion of several late-acting small subunit biogenesis factors suggests that the functions of some, such as NOB1, are conserved from yeast while others, including PRP43, are significantly different. Depletion of Prp43, which in yeast is necessary for NOB1 cleavage, possibly by causing structural rearrangements around site D to enable cleavage (Pertschy et al, 2009), did not cause any noticeable defects in pre-rRNA processing suggesting that this function is not conserved. The immediate precursor of 18S in yeast, 20S, is 217 nt longer than 18S while 18SE only extends approximately 25 nt beyond the 3' end of 18S. It is possible that ITS1 sequence in yeast blocks site D and the helicase activity of Prp43 is required for Nob1 to gain access to the cleavage site. Since the ITS1 extension is so much shorter in human cells, this rearrangement by PRP43 may not be required, bypassing the need for PRP43.

Depletion of NOB1 causes an increase in 18SE levels and a decrease in 18S levels. Other data from the Watkins lab (David Colvin) confirms that recombinant human NOB1 is an endonuclease capable of cleaving a substrate containing the 3' end

of 18S *in vitro* and taken together, these data imply that NOB1 is the endonuclease responsible for cleavage at site 3. The primary binding site of both human NOB1 and PNO1 in 18S rRNA appears to be on helices 39 and 40 in the 3' major domain, similar to that of yeast Nob1. We further showed that PNO1 and NOB1 bind cooperatively to this region, perhaps suggesting that a role of PNO1 is recruiting or anchoring NOB1 to 18S. However, although NOB1 and PNO1 also interacted with substrates containing the NOB1 cleavage site at the 3' end of 18S, they did not bind co-operatively to these substrates. Preliminary data also suggest that PNO1 does not increase the endonuclease activity of NOB1 *in vitro* (Watkins lab, unpublished data). NOB1 associates with early pre-40S complexes but does not cleave at site D/3 until the complex has been exported to the cytoplasm. In yeast, it is proposed that the binding of Nob1 to a site in 18S that is distinct from its cleavage site is a regulation mechanism; although Nob1 associates with early-pre-ribosomal complexes, it is unable to cleave at site D until structural rearrangements in the late-pre-40S complex bring it into close proximity with its cleavage site. This may also be the case for the human protein, but crystal structures of the small subunit (Ben-Shem et al, 2011; Ben-Shem et al, 2010) show that helices 39 and 40 are actually relatively close to the Nob1 cleavage site. However, recombinant human NOB1 is able to interact with itself and other work from the Watkins lab has shown that in the cytoplasm, NOB1 is present as a multimer, whereas in the nucleus it is monomeric, so it is possible that multimerisation may represent an alternative mechanism of NOB1 regulation.

Chapter Six

Discussion

6.1 Overview

The exosome has been extensively characterised in *S. cerevisiae* where the activities of the exosome proteins and the mechanisms of RNA processing of a range of different substrates are well understood. At the beginning of this project, a crystal structure of the core human exosome had been published (Liu et al, 2006) but little was known about the active subunits or cofactors and how they interact with the core exosome. It was not clear whether exosome substrates identified in yeast were also processed by the human exosome and if the human exosome has additional functions. Here, we have expanded on previous data showing that the exosome is stably associated with pre-snoRNP complexes (Watkins et al, 2004; Watkins et al, 2007) by demonstrating *in vitro* protein-protein interactions between snoRNP biogenesis factors and exosome proteins. Our data question whether the human exosome is essential for pre-snoRNA processing, but demonstrate that the both the exosome subunit, RRP6, and the 5'-3' exonuclease, XRN2 (Rat1), are important for the degradation and quality control of snoRNAs.

Compared to yeast, removal of the 5'ETS of pre-rRNA in higher eukaryotes involves an additional processing step, A'. We have shown that, as in yeast, the exonucleases, XRN2 and the exosome, participate in degradation of the 5'ETS fragments generated by A' and A_0 cleavages and that XRN2 (Rat1) is required for efficient A' cleavage in HeLa cells. We have also identified an exosome-independent role for the TRAMP helicase, MTR4, in A' cleavage. Further, although the core U3 snoRNP proteins are required for efficient A' cleavage, we have shown that the major role of these proteins is in pre-rRNA processing downstream of A' cleavage. We also provide evidence that the core exosome, RRP6 and MTR4 function in pre-rRNA quality control by degrading partially processed intermediates, such as 37S*.

Processing of ITS1 is a particularly important step in ribosome biogenesis as this is how the rRNAs destined for the SSU and LSU are separated. In higher eukaryotes, ITS1 processing involves an additional step compared to yeast implying the mechanism by which this is achieved may be significantly different (Rouquette et al, 2005). Previously, two alternative models of ITS1 processing following site 2 cleavage were suggested: an endonuclease at 2a involving XRN2 (Rat1) (Wang & Pestov, 2011) or direct 3'-5' exonucleolytic processing from 2 to 2a (Carron et al, 2011). We have

been able to draw parallels between the two human ITS1 cleavage sites (2a and 2) and the yeast ITS1 cleavage sites (A_2 and A_3, respectively). However, we have shown that, unlike in yeast, where A_2 is the primary cleavage site, that the A_3-like site 2 is the major cleavage site in human ITS1. In yeast, the 3' end of 18S is then directly generated by another endonucleolytic cleavage whereas the majority of human ITS1 sequence is removed by exonucleolytic processing by the exosome before an endonucleolytic cleavage at the 3' end of 18S. This exosome-dependant pathway is the major route for production of 18S rRNA but we also provide evidence of an alternative pathway for 18SE production involving an endonucleolytic cleavage at site 2a.

The late-acting 18S biogenesis factor, ENP1, and the small subunit ribosomal protein, RPS19, are required for efficient exonucleolytic processing of ITS1. We also identify CRM1 as a possible regulator of this exonucleolytic processing and speculate whether PNO1 may coordinate the timing of ITS1 processing with nuclear export of pre-40S complexes. We provide evidence that the late-acting SSU biogenesis factors, DIM1, NOB1 and PNO1, interact both with each other and with the 3' end of 18S and 5' end of ITS1 *in vitro*. We suggest, therefore, that the binding of these factors to 18S could be responsible for stalling exonucleolytic processing of ITS1 near the 2a cleavage site. We investigated the roles of other late pre-40S complex factors in human cells and while our data suggest that the role of NOB1 as the 3' 18S endonuclease is conserved from yeast, PRP43 does not appear essential for pre-rRNA processing in human cells. Finally, we have identified DIM1 as a target of RIO2 kinase *in vitro*, but it is not yet clear what the function of this phosphorylation is.

6.2 The human exosome

In contrast to the yeast exosome which associates with two ribonuclease activities, Rrp44 and Rrp6, the human exosome has been shown to associate with three distinct activities, DIS3, DIS3L and RRP6 (Lubas et al, 2011; Staals et al, 2010; Tomecki et al, 2010b). In yeast, Rrp44 is constitutively associated with the exosome, whereas in humans neither DIS3 nor DIS3L are thought to be robustly associated with the core exosome. Unlike yeast Rrp44, which is found throughout the cell, DIS3 is localised to the nucleus but excluded from nucleoli in human cells and DIS3L is only found in the cytoplasm (Tomecki et al, 2010b). Further, RRP6, which in yeast is restricted to the nucleus, is considerably concentrated in the nucleoli and also found in the cytoplasm in human cells. In yeast, exosome complexes exist in two forms; those associated with only Rrp44 and those coupled with both Rrp44 and Rrp6. The localisation of the human homologues implies more variation in the combination of

activities associated with exosome complexes in different compartments of human cells. Perhaps most significantly, in the nucleolus, RRP6 is the major active subunit of the exosome implying that this protein is able to carry out all the RNA processing functions of the exosome which occur there. This includes several substrates which, in yeast, are thought to be primarily processed by the more processive enzyme, Rrp44.

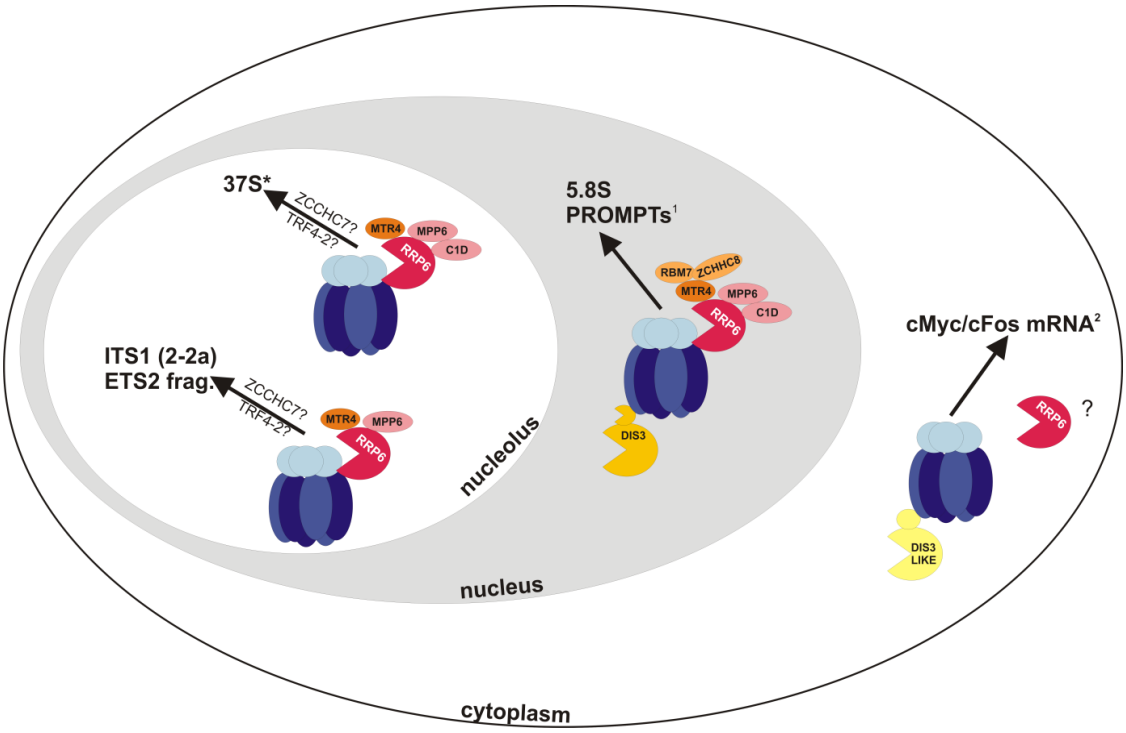


Figure 6.1 Functions and cofactors of the human exosome In the nucleolus, the major active component of the human exosome is RRP6. MTR4 and MPP6 appear to be associated with all exosome complexes but C1D is only required for processing a subset of nucleolar exosome substrates. A newly identified TRAMP-like complex containing ZCCHC7 and TRF4-2 also exists[1] but it is not yet known whether this is involved in exosome processing of the substrates identified in this study. In the nucleus, exosome complexes can associate with both DIS3 and RRP6. Both these activities and the cofactors MTR4, MPP6 and C1D are required for 5.8S processing. The NEXT complex (RBM7 and ZCCHC8), which interacts with the exosome through MTR4 has also been identified and functions in both 5.8S processing and degradation of PROMPTS[1]. In the cytoplasm, the exosome is associated with an alternative form of DIS3, DIS3-LIKE, and functions in mRNA degradation[2]. RRP6 is also found in the cytoplasm of human cells but it is not clear if this interacts with the exosome or whether it is involved in RNA metabolism in this compartment.

[1] Lubas et al. 2011, [2] Tomecki et al. 2010

Our data and other reports show that DIS3 is required for 3' end formation of 5.8S and degradation of PROMPTs (Lubas et al, 2011) but does not participate in ITS1 processing to produce 18SE, recycling of the 5'ETS fragment released by A' and A_0 cleavages or degradation of the aberrant precursor, 37S* (Figure 6.1). In contrast, the core exosome, RRP6 and its cofactors, MPP6 and MTR4, appear to be the major activities required for processing of these RNAs (Figure 6.1). These substrates are long and highly structured (ITS1 ~800 nt and 82 % GC, ETS2 ~1.2 kB and 37S* ~10 kB) making it unlikely that the distributive exonuclease activity of RRP6 observed *in vitro*

would be sufficient to degrade these sequences. Although human RRP6 is able to degrade more structured RNAs than its yeast counterpart, and its cofactors, C1D, MPP6 and ENP1, are able to subtly modulate its activity, our data imply that the activity of RRP6 in the nucleolus must be significantly enhanced to make it more processive. It is possible that the interaction of RRP6 with the core exosome dramatically increases its activity. In yeast, truncated but active Rrp6, which is not able to interact with the core exosome, functions in the 3' processing of 5.8S but cannot degrade long polyadenylated substrates, perhaps suggesting that in isolation Rrp6 is able to carry out functions that involve fine trimming of a minimal number of nucleotides, but contact with the core exosome is necessary for Rrp6 to process longer substrates (Callahan & Butler, 2008). In yeast, the TRAMP complex functions as an exosome cofactor in two different ways; through recruiting the exosome to its substrates by addition of short poly (A) tails and also, independently of its polyadenylation or helicase activities, by increasing the activity of the Rrp6-associated exosome more than 10 fold (Callahan & Butler, 2010). It is not yet entirely clear whether the TRAMP complex and its functions are conserved from yeast but recently, human homologues of the yeast Trf4 polymerase, hTRF4-2, and yeast Air2, hZCCHC7, have been identified in human cells (Fasken et al, 2011; Lubas et al, 2011). These proteins are localised to the nucleolus in complexes the same size as the RRP6-exosome and participate in adenylation of abortive pre-rRNA fragments. Our data also show that MTR4 is involved in ITS1 processing and degradation of both the ETS2 fragment and the novel 37S* intermediate. These events probably occur in the nucleolus, but MTR4 is also required for 5.8S processing which appears to occur in the nucleus. Lubas at al., (2011) have shown that human MTR4 along with RBM7 and ZCCHC8 form another complex, the NEXT complex, which is involved in exosome processing of the 5.8S precursor. The specific roles of MTR4 and these newly identified nucleolar TRAMP homologues in ITS1 processing from site 2 to 2a merits further investigation to determine if these proteins stimulate the exonuclease activity of RRP6 enabling it to degrade this long, structured substrate.

Our data also suggest that, within the nucleolus, the requirement for the RRP6 cofactors, C1D and MPP6, is different depending on which substrate is targeted (Figure 6.1). In yeast, the effect of depleting either Rrp47 (C1D) or Mpp6 parallels the depletion of Rrp6 for the majority of substrates implying they are common cofactors and this also appears to be the case for human MPP6. In contrast, while C1D is required for degradation of 37S* and 5.8S processing, it does not function in ITS1 processing or turnover of the ETS2 fragment, implying another levels of complexity between RRP6-exosomes. As both C1D and MPP6 contain RNA binding motifs, they

have been proposed to function in substrate recognition possibly indicating each protein targets the RRP6-exosome to specific substrates.

6.3 Quality control of ribosome biogenesis

The rate of ribosome production in cells is very high and has been estimated at 7500 ribosomes/min in proliferating HeLa cells (Mayer et al, 2006). The amount of aberrant pre-rRNA intermediates detected in cells depleted of key biogenesis factors is, however, relatively low. For example, metabolic labelling of cells depleted of RRP5 causes a major accumulation of 45S and a significant block in production of both 28S and 18S rRNA, but the aberrant intermediates, 30SL3' and 21SL3', which result from blocking site 2 cleavage are only just detectable by Northern blotting. This implies that these fragments are rapidly and efficiently targeted for degradation and raises the question of how aberrant pre-rRNAs are distinguished from correctly processed intermediates. The potential for incorrectly assembled pre-ribosomal complexes is high since many biogenesis factors contain non-specific RNA binding motifs and common sequences/structural elements are found throughout the 13kB pre-rRNA transcript. These near-cognate interactions would not be significantly unfavourable, implying that RNA surveillance mechanisms must be dynamic rather than passive. It is, therefore, proposed that the fidelity of ribosome production is achieved through a series of irreversible steps, such as nucleolytic cleavages and processing (e.g. NOB1, XRN2, exosome) and steps driven by nucleotide hydrolysis possibly mediated by the multiple GTPases (e.g. BMS1 and TSR1) and ATP-dependent RNA helicases (e.g. MTR4) which function in ribosome biogenesis (Houseley & Tollervey, 2009). The complexity of ribosome assembly suggests that many individual quality control steps are serially required throughout pre-rRNA processing to enhance the overall accuracy of this proof-reading. Potential surveillance steps identified in human ribosome biogenesis include: A' cleavage, incorporation of DIM1 into pre-ribosomal complexes and exonucleolytic processing of pre-rRNA.

It would be possible for the mature rRNAs to be produced from the precursor transcript specifically by endonucleolytic cleavages raising the question of why pre-rRNA processing has evolved to use both exonuclease and endonuclease activities to achieve this. It is possible that recruiting exonucleases to newly formed pre-rRNA ends is a quality control mechanism for complex assembly. It was recently demonstrated that if A_3 cluster proteins are not correctly recruited to the pre-rRNA, the exonuclease, Rat1 (XRN2), switches from processing the mature 5' end of 5.8S to degrading the misassembled complex (Granneman et al, 2011). Comparably, we propose that

proteins, including DIM1, NOB1 and PNO1, bound to the 3' end of 18S arrest the 3'-5' exonucleolytic processing of ITS1 at the 2a site, raising the possibility that if these proteins are not properly assembled, the exosome may instead degrade the aberrant pre-rRNA. Use of exonucleases, which can degrade aberrant transcripts, in key processing steps may, therefore, represent the most efficient surveillance pathway. Similarly, XRN2 (Rat1) is thought to be recruited to the pre-rRNA transcript co-transcriptionally and in higher eukaryotes, is required for both A' cleavage and degradation of abortive pre-rRNA fragments. A' cleavage could, therefore, represent an additional checkpoint in pre-rRNA processing by coupling the docking of the U3 snoRNP complex and SSU processome, which are essential for downstream 5'ETS processing, with initiating processing. If the SSU processome/U3 snoRNP is not correctly associated, XRN2 would be positioned to degrade the transcript from the free 5' end. MTR4, which is also required for A' cleavage, and as part of TRAMP complex participates in RNA surveillance pathways, may also participate in this quality control. Furthermore, exonucleases may be important for directing the optimal pathway of processing by eliminating intermediates with abnormal cleavage patterns. Depletion of XRN2 affects the order of pre-rRNA cleavages such that A' cleavage can be preceded by A_0 or site 2 cleavage and an increased proportion of pre-rRNAs are cleaved at 2a prior to site 2. Consistent with this, in mouse cells, XRN2 has been shown to degrade pre-rRNAs which have bypassed early cleavage steps. Similar to yeast, where depletion of the exosome causes accumulation of partially processed but aberrant pre-rRNA fragments (Allmang et al, 2000), in human cells, depletion of the exosome causes accumulation of the 37S* fragment in which processing has been aborted following A' cleavage.

6.4 Regulation of ribosome biogenesis and disease

We have shown that alternative, but not redundant, processing pathways can be used to generate 18S rRNA in human cells. The major and minor ITS1 processing pathways in human cells are, however, linked because inhibiting the major site 2 cleavage stimulates processing at the minor 2a site. When site 2 cleavage occurs normally, however, the cell is dependent on the exosome to produce 18S rRNA. Similar alternative pathways have also been suggested for ITS1 processing in *X. laevis* (Borovjagin & Gerbi, 2001) implying this is likely to be a conserved feature of pre-rRNA processing in higher eukaryotes. In yeast, alternative non-essential pre-rRNA processing pathways are used to produce long and short forms of 5.8S, which seem functionally identical, and our data suggest this is also the case in human cells. It is not

yet clear what the purpose of these alternative pathways for production of the mature rRNAs is. Altering the long-term growth rate of cells did not affect whether exonucleolytic or endonucleolytic processing was used to produce 18SE in HeLa cells. It is possible, however, that the ratio of $5.8S_S$ and $5.8S_L$ expressed in yeast or human cells may be changed when cells are grown under different conditions. There is a growing body of evidence suggesting that different populations of ribosomes exist within cells and that formation of specialised ribosomes is influenced by changes in the external environment or cellular growth rate (Gilbert, 2011). In yeast and plants, duplication of ribosomal protein genes leads to expression of subtly different isoforms and expression levels of particular proteins are altered by external stimuli. Further, in human cells, RPS19, the protein linked to Diamond Blackfan anaemia, is expressed from different genes with variable 3'UTRs and the length of the UTR correlates with decreased protein expression and increased severity of the disease. Similarly, rRNAs are extensively modified by snoRNPs but it is reported that expression of snoRNAs is divergently regulated under some conditions. It is further suggested that a subset of snoRNAs regulate changes in ribosome function by alternative rRNA modifications states to optimise ribosomes in response to cellular stresses. We speculate that the role of the alternative processing pre-rRNA processing pathways could be to generate subtly different ribosomes which may be optimal for slightly different functions.

Ribosome biogenesis and pre-rRNA processing are regulated globally by pathways involved in transcription regulation, stress response and cell cycle progression. RRP5, which is required for both ITS1 cleavages and recruitment of components of the SSU processome, interacts with the transcription factor, NFkappa B (Sweet et al, 2008), suggesting ribosome biogenesis may be coordinated with expression of particular genes. Further, in yeast, the localisation of Pno1, a cofactor of Nob1, is regulated by the mTOR pathway which coordinates cellular responses to nutritional and oxidative stress (Vanrobays et al, 2008). Recently, inhibition of the mTOR pathway was also shown to block specific steps of pre-rRNA processing required for LSU and SSU biogenesis in HeLa cells, implying mTOR also regulates ribosome biogenesis in higher eukaryotes (Iadevaia et al, 2011). Several of the ribosome biogenesis factors we have investigated here, such as BOP1 and ENP1, have been linked to cell cycle progression (Holzel et al, 2005; Pestov et al, 2001b; Strezoska et al, 2002; Wang et al, 2009). Two key steps in pre-rRNA processing; site 2 cleavage which separates the LSU and SSU rRNAs and exonucleolytic processing to produce 18S, involve BOP1 and ENP1, respectively, suggesting that coupling cell cycle progression with these important pre-rRNA processing events represents a way in which the rate of ribosome biogenesis can be rapidly modulated to match the cellular

growth rate. Interestingly, both BOP1 and ENP1 have been shown to be upregulated in particular cancers, colorectal and hepatocellular carcinoma, respectively (Chung et al, 2011; Killian et al, 2006; Wang et al, 2009). In yeast, the exosome and TRAMP components, Rrp44, Rrp4, Mtr3 and Mtr4, are specifically linked to cell cycle progression as they are important for microtubule formation, implying a possible mechanism linking RNA processing and cell cycle progression (Smith et al, 2011).

Key steps of pre-rRNA processing and assembly of ribosomal proteins need to be carefully regulated to ensure that events occur in the optimal order and at the correct time. Protein phosphorylation is also important for regulating ribosome biogenesis both in terms of pre-rRNA processing and protein association. Our data confirm previously published data showing that the kinase activity of RIO2 is required for conversion of 18SE into 18S (Zemp et al, 2009) and we further show that DIM1 is a substrate of RIO2 phosphorylation *in vitro*. The mechanisms controlling ITS1 processing are not fully understood, but our data suggest that CRM1 may be important for delaying exonuclease processing until the 5'ETS is removed, potentially coordinating processing with nuclear export of pre-40S complexes into the cytoplasm. Exosome processing of ITS1 appears to be regulated by specific exosome cofactors, such as ENP1, MPP6, C1D and possibly the TRAMP complex, which both modulate the activity of the exosome and probably function in recruiting the exosome to the pre-rRNA following site 2 cleavage. It is possible that inhibiting exonucleolytic processing of ITS1 by RRP6 and therefore, 18S production, could be the mechanism by which the chemotherapeutic agent, 5-fluorouracil (5FU), functions. This drug was initially selected for its potential ability to inhibit DNA synthesis by thymidine starvation but in yeast, Rrp6 is targeted by 5FU, raising the possibility that the toxic effect of this drug is cause by blocking ribosome production (Fang et al, 2004; Hoskins & Scott Butler, 2007; Silverstein et al, 2011). In human cells, 5FU has been shown to significantly decrease 18S levels (Burger et al, 2010).

Several of the proteins identified here as playing important roles in pre-rRNA processing have been linked to genetic diseases implying that defects in ribosome biogenesis may be the underlying causes of the diseases (Table 6.1). RPS19 is important for exonucleolytic processing of ITS1 and mutations in this gene cause ribosomal haploinsufficiency and Diamond Blackfan anaemia which is characterised by developmental abnormalities and erythroid deficiency. Mutation of RPS19 has recently been shown to inhibit IRES-mediated translation of mRNAs specifically expressed in erythroid progenitors (Horos et al, 2011). Another genetic disorder, ANE syndrome, is caused by a homozygous point mutation in an RNA binding motif of RBM28 which destabilises the protein causing a significant decrease in RBM28 levels (Nousbeck et

al, 2008; Spiegel et al, 2010). However, RNAi mediated depletion of this protein to <10 % normal levels in HeLa cells slightly slowed ITS1 cleavages, but did not have a

Table 6.1 Ribosome biogenesis factors and disease

Protein	Pre-rRNA step - yeast	Pre-rRNA step - human	Disease	References
Exosome and RRP6	3' 5.8S	3' end 5.8S and ITS1 (2-2a)	Polymyositis-scleroderma	Brouwer et al., 2001, Schilders et al., 2007, Staals and Pruijn 2011
RRP46	3' 5.8S	3' end 5.8S and ITS1 (2-2a)	Chronic myeloid leukaemia	Yang et al, 2002
RPS19	20S accumulation	Exonuclease ITS1 processing and site 2a	Diamond Blackfan anaemia	Idol et al., 2007
ENP1	Structural re-arrangement of pre-40S complex	Exonuclease ITS1 processing and site 2a	Hepatocellular carcinoma	Wang et al., 2009 Fukuda et al., 2008
BOP1	ITS1 and ITS2 cleavages	Site 2 (strong), site 2a (slow), ITS2	Colorectal cancer, hepatocellular carcinoma	Killian et al., 2006, Chung et a., 2011, Holzel et al., 2005, Pestov et al., 2001
RBM28	A_3	Site 2 (mild)	ANE syndrome	Nousbeck et al., 2008, Spiegel et al., 2010
RNase MRP	A_3	No defect	CHH syndrome, scleroderma	Glazov et al., 2011, Ridanpaa et al., 2001, Mattijssen et al., 2010
NOB1	Site D	Site 3	Ovarian cancer, marker of chronic myeloid leukaemia, upregulated in noise injured cochlea	Oehler et al., 2009, Lin et al., 2011, Han et al., 2011
snoRNP	5'ETS and A_2	A' (mild) , A_0, A_1, 2a	Dyskeratosis congenita	Rashid et al., 2006 Gupta and Kumar 2010

major effect on production of the mature rRNAs. Although the number of ribosomes in cells from patients with ANE is decreased, our data question whether this disease is caused by a defect in pre-rRNA processing. However, the RBM28 homologue in yeast, Nop4, is suggested to affect ribosome biogenesis by mediating posttranscriptional methylation of the pre-rRNA transcript so it is possible that the lack of these modifications in human pre-rRNA could be the cause of ANE syndrome (Berges et al, 1994). Alternatively, ANE could be independent of ribosome biogenesis as RBM28 in *C.elegans* regulates expression of the miRNA, lin4, which is linked to development (Bracht et al, 2010). Similarly, mutations in the RNase MRP RNA or promoter which affect either RNase MRP structure or expression levels have been shown to cause CHH syndrome (Mattijssen et al, 2010b; Ridanpaa et al, 2001). As RNase MRP is the endonuclease which cleaves site A_3 in yeast ITS1, this disease has been classified as

a ribosomopathy. Our data suggest, however, that the role of RNase MRP in ITS1 processing is not conserved in human cells. Depletion of RNase MRP subunits which dramatically reduces RNase MRP RNA levels, mimicking those observed in CHH syndrome patients, does not affect pre-rRNA processing, implying that CHH syndrome is not caused by defective ribosome biogenesis. Instead, CHH syndrome is likely to result from the inability of RNase MRP to cleave another of its substrates, such as a particular mRNA. Both RNase MRP and the exosome are also associated with the autoimmune, connective tissue diseases, scleroderma, polymyositis and the PM/Scl overlap syndrome. The exosome proteins, RRP6, RRP45 and C1D, were each first identified as autoantigens and the RNA component of RNase MRP gives rise to two autoantigens, anti-Th and anti-To. It is not yet clear why proteins of these ribonuclease complexes trigger an autoimmune response which leads to these diseases. Further, the core exosome subunit, RRP46, is significantly upregulated in patients with chronic myeloid leukaemia and it is suggested that RRP46 may be a tumour-related antigen and stimulate autoimmune responses in cancer (Staals & Pruijn, 2011; Yang et al, 2002).

6.5 Future work

We have shown that RRP6 is the major activity required for production of 18SE following site 2 cleavage but depletion of RRP6 causes accumulation of a partially processed intermediate and does not completely abolish 18SE production. This implies that other exonucleases may also participate in ITS1 processing. Recently, the human homologue of Rex4, PCM2, has been shown to possess exonuclease activity *in vitro* and homologues of Rex1, Rex2 and Rex3 can also be identified in human cells making these proteins likely candidates to be involved. Depletion of ENP1 or RPS19 caused a pre-rRNA processing defect similar to the lack of RRP6, suggesting these proteins may recruit the exosome to the pre-rRNA. In yeast, Rps19 has been shown to recruit Enp1 to pre-40S complexes supporting this idea. It is, therefore, intended to investigate possible interactions between ENP1, RPS19, the exosome and ITS1 to try to elucidate a mechanism of exosome recruitment.

We have identified several roles for RRP6 in pre-rRNA processing that would require RRP6 to possess more processive exonuclease activity than has been demonstrated *in vitro*. The TRAMP complex has been shown to dramatically increase the activity of the Rrp6-exosome in yeast and components of the TRAMP complex have recently been identified in human cells. It would, therefore, be interesting to investigate the role of these proteins (ZCCHC7, hTRF4-2) and their activities in both

the novel processing (ITS1 and A') and degradation functions (ETS2 and 37S*) of the human exosome and MTR4 identified in this study. Other factors, ZFC3H1, ZC3H18 and ARS2, were recently found to be associated with the human exosome but it remains to be determined if these are in fact exosome cofactors and what their specific functions are.

Our data suggest that CRM1 is a regulator of ITS1 processing and in particular, in determining the correct timing of this processing. The similar pre-rRNA processing defects caused by depletion of PNO1 or RRP6 and inhibition of CRM1, coupled with the role of CRM1 in nuclear export, suggests exonucleolytic processing of ITS1 may be linked to export of pre-40S complexes into the cytoplasm. To try to clarify whether PNO1 does function as an export adaptor, it is intended to express a series of truncated forms of PNO1 and mutate the extended NES identified, then investigate the ability of these protein fragments to interact with CRM1. It is hoped that the sequence required for the interaction between PNO1 and CRM1 could be identified, enabling the effect of blocking this interaction on the export of pre-40S complexes to be determined *in vivo*.

An endonucleolytic cleavage at 2a can produce 18SE and this requires the SSU processome including the putative endonucleases, UTP24 and RCL1. In yeast, it has recently been suggested that Rcl1 is the endonuclease responsible for the analogous A_2 cleavage but interestingly, in the Watkins lab (David Colvin), human UTP24 has been confirmed to have endonuclease activity *in vitro* and has been shown to cleave a substrate containing the 2a site in a specific manner. Another PIN domain protein found in the SSU processome, UTP23, was also observed to cleave this substrate. It is, therefore, unclear which endonuclease is responsible for cleavage at the 2a site and it is intended to investigate the endonuclease activity of all three of these proteins, both *in vitro* and in human cells, to determine which pre-rRNA cleavage steps each is required for. The 2a cleavage site has not been accurately mapped. Based on sequence similarity with the yeast A_2 cleavage site, a potential 2a site has been identified but it remains to be determined if this is indeed the cleavage site.

18SE can be produced by either exonucleolytic processing or an endonucleolytic cleavage at site 2a but exonucleolytic processing is the primary mechanism. 18SE appears to be heterogeneous, extending approximately 30 nt beyond the 3' end of 18S and this sequence is removed by an endonucleolytic cleavage at site 3 by NOB1. We have attempted to investigate differences between 18SE generated by exonucleolytic and endonucleolytic processing. Metabolically labelled RNA from cells in which exonucleolytic processing is slowed by depletion of

RRP6, or those depleted of either BOP1 or XRN2, which rely on endonucleolytic processing, was cleaved using RNase H to generate small fragments in which subtle differences in 18SE length could potentially be resolved by PAGE. However, since the levels the 18SE are significantly decreased in cells depleted of RRP6 and the product is so heterogeneous, with a limited amount of labelled RNA, it was barely possible to detect the 18SE signal. Unlike an ordered series of endonucleolytic cleavages, exonucleolytic processing of ITS1 raises the question of how much of ITS1 sequence must be removed before NOB1 is able to cleave to produce 18S. It is, therefore, intended to try to determine if there is a threshold length of ITS1 that must be processed by the exosome before site 3 can be cleaved.

It has only been possible to map site 2 cleavage to a region of ITS1 approximately 160 nucleotides long since this cleavage appears to be followed by rapid exonucleolytic processing in both directions. To map this cleavage site it would, therefore, be necessary to deplete cells of all the exonucleases involved in ITS1 processing, which in additional to the exosome and XRN2, may include PCM2 (Rex4) and NOL12 (Rrp17) and to identify a protein factor specifically required for site 2 cleavage. We have identified several factors that are required for site 2 cleavage (RRP5, BOP1 and RBM28) but all of these proteins are also required for other steps in pre-rRNA processing. The yeast homologues of BOP1 and RBM28 function as part of a multi-protein cluster also containing Nop15, Cic1, Ytm1, Nop7 and Nop12, so it would be of interest to determine if this complex is fully conserved from yeast to humans. Subtle differences between the involvement of BOP1 and RBM28 in A_3/site 2 cleavage in yeast and humans, respectively, have been identified, making the roles of other proteins of this complex of particular interest. Like BOP1 and RBM28, which have been linked to colorectal cancer and ANE syndrome respectively, the human homologues of other A_3-cluster proteins, such as PES (Ytm1), WDR12 (Nop7), are implicated in cell cycle control. We have also identified MKi67 as the homologue of Nop15 and this protein interacts with Ki67, a major antigen which regulates mitosis and cell cycle progression. Our data suggest that, despite the parallels between yeast A_3 cleavage and human site 2, RNase MRP is not the major activity required for this cleavage. The endonuclease responsible for separating the LSU and SSU rRNAs, therefore, remains elusive. Dicer, an endonuclease involved in the RNAi pathway, which is only found in higher eukaryotes influences cleavages round site 2 but the defects caused by depletion of this protein are relatively minimal implying that if Dicer does participate in site 2 cleavage, it is unlikely to be the major activity involved. Another nucleolar protein which has been shown to possess endonucleolytic activity is nucleophosmin (also known as B23) and this protein is putatively linked to ITS2 cleavage (Herrera et al,

1995; Savkur & Olson, 1998). Nucleophosmin has, however, been shown to interact with PES1, an interacting partner of BOP1 and therefore likely required for site 2 cleavage (Zhang et al, 2009) and with Ebp2 another protein needed for both large and small subunit biogenesis possibly suggesting it could be involved in site 2 cleavage in ITS1.

6.6 Conclusion

Together these data reveal key differences between ribosomal RNA processing and ribosome assembly in yeast and human cells. These findings demonstrate the importance of building on yeast models by also studying ribosome biogenesis in human cells. We are able to provide insights into the proteins required for pre-rRNA processing steps that are specific to higher eukaryotes and we define a model for ITS1 processing in human cells. By studying the roles of different exosome subunits and cofactors in the processing and degradation of several RNA substrates we have also observed some important differences in the sequences and substrates processed by the exosome subunits, RRP6 and DIS3, in human cells compared to yeast. Several proteins that have been shown here to play important roles in pre-rRNA processing and ribosome biogenesis in human cells are also implicated in genetic diseases, cell cycle progression and cancer, underlining the importance of further work to determine the molecular functions of these proteins in human cells.

References

Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D (1999a) Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *The EMBO journal* 18: 5399-5410

Allmang C, Mitchell P, Petfalski E, Tollervey D (2000) Degradation of ribosomal RNA precursors by the exosome. *Nucleic acids research* 28: 1684-1691

Allmang C, Petfalski E, Podtelejnikov A, Mann M, Tollervey D, Mitchell P (1999b) The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases. *Genes & development* 13: 2148-2158

Angermayr M, Bandlow W (2002) RIO1, an extraordinary novel protein kinase. *FEBS Letters* 524: 31-36

Aoki R, Suzuki N, Paria BC, Sugihara K, Akama TO, Raab G, Miyoshi M, Nadano D, Fukuda MN (2006) The Bysl gene product, bystin, is essential for survival of mouse embryos. *FEBS Lett* 580: 6062-6068

Arcus VL, Backbro K, Roos A, Daniel EL, Baker EN (2004) Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *The Journal of biological chemistry* 279: 16471-16478

Arenas JE, Abelson JN (1997) Prp43: An RNA helicase-like factor involved in spliceosome disassembly. *Proceedings of the National Academy of Sciences of the United States of America* 94: 11798-11802

Arigo JT, Eyler DE, Carroll KL, Corden JL (2006) Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Molecular cell* 23: 841-851

Atzorn V, Fragapane P, Kiss T (2004) U17/snR30 is a ubiquitous snoRNA with two conserved sequence motifs essential for 18S rRNA production. *Molecular and cellular biology* 24: 1769-1778

Azzouz N, Panasenko OO, Colau G, Collart MA (2009) The CCR4-NOT complex physically and functionally interacts with TRAMP and the nuclear exosome. *PloS one* 4: e6760

Baillat D, Hakimi MA, Naar AM, Shilatifard A, Cooch N, Shiekhattar R (2005) Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* 123: 265-276

Baserga SJ, Yang XD, Steitz JA (1991) An intact Box C sequence in the U3 snRNA is required for binding of fibrillarin, the protein common to the major family of nucleolar snRNPs. *The EMBO journal* 10: 2645-2651

Bassler J, Kallas M, Pertschy B, Ulbrich C, Thoms M, Hurt E (2010) The AAA-ATPase Rea1 drives removal of biogenesis factors during multiple stages of 60S ribosome assembly. *Molecular cell* 38: 712-721

Baudin-Baillieu A, Fabret C, Liang XH, Piekna-Przybylska D, Fournier MJ, Rousset JP (2009) Nucleotide modifications in three functionally important regions of the Saccharomyces cerevisiae ribosome affect translation accuracy. *Nucleic acids research* 37: 7665-7677

Baxter-Roshek JL, Petrov AN, Dinman JD (2007) Optimization of ribosome structure and function by rRNA base modification. *PloS one* 2: e174

Ben-Shem A, Garreau de Loubresse N, Melnikov S, Jenner L, Yusupova G, Yusupov M (2011) The Structure of the Eukaryotic Ribosome at 3.0 A Resolution. *Science*

Ben-Shem A, Jenner L, Yusupova G, Yusupov M (2010) Crystal structure of the eukaryotic ribosome. *Science* 330: 1203-1209

Berges T, Petfalski E, Tollervey D, Hurt EC (1994) Synthetic lethality with fibrillarin identifies NOP77p, a nucleolar protein required for pre-rRNA processing and modification. *The EMBO journal* 13: 3136-3148

Bernstein KA, Gallagher JE, Mitchell BM, Granneman S, Baserga SJ (2004) The small-subunit processome is a ribosome assembly intermediate. *Eukaryotic cell* 3: 1619-1626

Billy E, Wegierski T, Nasr F, Filipowicz W (2000) Rcl1p, the yeast protein similar to the RNA 3'-phosphate cyclase, associates with U3 snoRNP and is required for 18S rRNA biogenesis. *The EMBO journal* 19: 2115-2126

Bleichert F, Gagnon KT, Brown BA, 2nd, Maxwell ES, Leschziner AE, Unger VM, Baserga SJ (2009) A dimeric structure for archaeal box C/D small ribonucleoproteins. *Science* 325: 1384-1387

Bleichert F, Granneman S, Osheim YN, Beyer AL, Baserga SJ (2006) The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proceedings of the National Academy of Sciences of the United States of America* 103: 9464-9469

Bohnsack MT, Czaplinski K, Gorlich D (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA (New York, NY* 10: 185-191

Bohnsack MT, Martin R, Granneman S, Ruprecht M, Schleiff E, Tollervey D (2009) Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Molecular cell* 36: 583-592

Bonneau F, Basquin J, Ebert J, Lorentzen E, Conti E (2009) The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. *Cell* 139: 547-559

Borovjagin AV, Gerbi SA (2001) Xenopus U3 snoRNA GAC-Box A' and Box A sequences play distinct functional roles in rRNA processing. *Molecular and cellular biology* 21: 6210-6221

Borovjagin AV, Gerbi SA (2005) An evolutionary intra-molecular shift in the preferred U3 snoRNA binding site on pre-ribosomal RNA. *Nucleic acids research* 33: 4995-5005

Boulon S, Marmier-Gourrier N, Pradet-Balade B, Wurth L, Verheggen C, Jady BE, Rothe B, Pescia C, Robert MC, Kiss T, Bardoni B, Krol A, Branlant C, Allmang C, Bertrand E, Charpentier B (2008) The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *The Journal of cell biology* 180: 579-595

Boulon S, Verheggen C, Jady BE, Girard C, Pescia C, Paul C, Ospina JK, Kiss T, Matera AG, Bordonne R, Bertrand E (2004) PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Molecular cell* 16: 777-787

Bowman LH, Emerson CP, Jr. (1977) Post-transcriptional regulation of ribosome accumulation during myoblast differentiation. *Cell* 10: 587-596

Bracht JR, Van Wynsberghe PM, Mondol V, Pasquinelli AE (2010) Regulation of lin-4 miRNA expression, organismal growth and development by a conserved RNA binding protein in C. elegans. *Dev Biol* 348: 210-221

Bradatsch B, Katahira J, Kowalinski E, Bange G, Yao W, Sekimoto T, Baumgartel V, Boese G, Bassler J, Wild K, Peters R, Yoneda Y, Sinning I, Hurt E (2007) Arx1 functions as an unorthodox nuclear export receptor for the 60S preribosomal subunit. *Molecular cell* 27: 767-779

Brand RC, Klootwijk J, Van Steenbergen TJ, De Kok AJ, Planta RJ (1977) Secondary methylation of yeast ribosomal precursor RNA. *European journal of biochemistry / FEBS* 75: 311-318

Bratkovic T, Rogelj B (2010) Biology and applications of small nucleolar RNAs. *Cell Mol Life Sci* 68: 3843-3851

Brouwer R, Pruijn GJ, van Venrooij WJ (2001) The human exosome: an autoantigenic complex of exoribonucleases in myositis and scleroderma. *Arthritis research* 3: 102-106

Burger K, Muhl B, Harasim T, Rohrmoser M, Malamoussi A, Orban M, Kellner M, Gruber-Eber A, Kremmer E, Holzel M, Eick D (2010) Chemotherapeutic drugs inhibit ribosome biogenesis at various levels. *The Journal of biological chemistry* 285: 12416-12425

Burkard KT, Butler JS (2000) A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. ***Molecular and cellular biology*** **20: 604-616**

Burwick N, Shimamura A, Liu JM (2011) Non-Diamond Blackfan anemia disorders of ribosome function: Shwachman Diamond syndrome and 5q- syndrome. ***Seminars in hematology*** **48: 136-143**

Buttner K, Wenig K, Hopfner KP (2005) Structural framework for the mechanism of archaeal exosomes in RNA processing. ***Molecular cell*** **20: 461-471**

Cahill NM, Friend K, Speckmann W, Li ZH, Terns RM, Terns MP, Steitz JA (2002) Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP. ***The EMBO journal*** **21: 3816-3828**

Callahan KP, Butler JS (2008) Evidence for core exosome independent function of the nuclear exoribonuclease Rrp6p. ***Nucleic acids research***

Callahan KP, Butler JS (2010) TRAMP complex enhances RNA degradation by the nuclear exosome component Rrp6. ***The Journal of biological chemistry*** **285: 3540-3547**

Carpousis AJ (2002) The Escherichia coli RNA degradosome: structure, function and relationship in other ribonucleolytic multienzyme complexes. ***Biochemical Society transactions*** **30: 150-155**

Carron C, O'Donohue MF, Choesmel V, Faubladier M, Gleizes PE (2011) Analysis of two human pre-ribosomal factors, bystin and hTsr1, highlights differences in evolution of ribosome biogenesis between yeast and mammals. ***Nucleic acids research*** **39: 280-291**

Chanfreau G (2003) Conservation of RNase III processing pathways and specificity in hemiascomycetes. ***Eukaryotic cell*** **2: 901-909**

Chanfreau G, Legrain P, Jacquier A (1998a) Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. ***Journal of molecular biology*** **284: 975-988**

Chanfreau G, Rotondo G, Legrain P, Jacquier A (1998b) Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. ***The EMBO journal*** **17: 3726-3737**

Chang DD, Clayton DA (1987) A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. ***The EMBO journal*** **6: 409-417**

Chekanova JA, Gregory BD, Reverdatto SV, Chen H, Kumar R, Hooker T, Yazaki J, Li P, Skiba N, Peng Q, Alonso J, Brukhin V, Grossniklaus U, Ecker JR, Belostotsky DA (2007) Genome-wide

high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. *Cell* 131: 1340-1353

Chen CY, Gherzi R, Ong SE, Chan EL, Raijmakers R, Pruijn GJ, Stoecklin G, Moroni C, Mann M, Karin M (2001) AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell* 107: 451-464

Chen W, Bucaria J, Band DA, Sutton A, Sternglanz R (2003) Enp1, a yeast protein associated with U3 and U14 snoRNAs, is required for pre-rRNA processing and 40S subunit synthesis. *Nucleic acids research* 31: 690-699

Chung KY, Cheng IK, Ching AK, Chu JH, Lai PB, Wong N (2011) Block of proliferation 1 (BOP1) plays an oncogenic role in hepatocellular carcinoma by promoting epithelial-to-mesenchymal transition. *Hepatology* 54: 307-318

Cohen A, Reiner R, Jarrous N (2003) Alterations in the intracellular level of a protein subunit of human RNase P affect processing of tRNA precursors. *Nucleic acids research* 31: 4836-4846

Combs DJ, Nagel RJ, Ares M, Jr., Stevens SW (2006) Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis. *Molecular and cellular biology* 26: 523-534

Comella P, Pontvianne F, Lahmy S, Vignols F, Barbezier N, Debures A, Jobet E, Brugidou E, Echeverria M, Saez-Vasquez J (2008) Characterization of a ribonuclease III-like protein required for cleavage of the pre-rRNA in the 3'ETS in Arabidopsis. *Nucleic acids research* 36: 1163-1175

Costello JL, Stead JA, Feigenbutz M, Jones RM, Mitchell P (2011) The C-terminal region of the exosome-associated protein Rrp47 is specifically required for box C/D small nucleolar RNA 3'-maturation. *The Journal of biological chemistry* 286: 4535-4543

Cristodero M, Bottcher B, Diepholz M, Scheffzek K, Clayton C (2008) The Leishmania tarentolae exosome: purification and structural analysis by electron microscopy. *Mol Biochem Parasitol* 159: 24-29

Dauwerse JG, Dixon J, Seland S, Ruivenkamp CA, van Haeringen A, Hoefsloot LH, Peters DJ, Boers AC, Daumer-Haas C, Maiwald R, Zweier C, Kerr B, Cobo AM, Toral JF, Hoogeboom AJ, Lohmann DR, Hehr U, Dixon MJ, Breuning MH, Wieczorek D (2011) Mutations in genes encoding subunits of RNA polymerases I and III cause Treacher Collins syndrome. *Nature genetics* 43: 20-22

de la Cruz J, Kressler D, Tollervey D, Linder P (1998) Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae. *The EMBO journal* 17: 1128-1140

Decatur WA, Fournier MJ (2002) rRNA modifications and ribosome function. *Trends in biochemical sciences* 27: 344-351

Decatur WA, Fournier MJ (2003) RNA-guided nucleotide modification of ribosomal and other RNAs. *The Journal of biological chemistry* 278: 695-698

Dechampesme AM, Koroleva O, Leger-Silvestre I, Gas N, Camier S (1999) Assembly of 5S ribosomal RNA is required at a specific step of the pre-rRNA processing pathway. *The Journal of cell biology* 145: 1369-1380

Dieci G, Preti M, Montanini B (2009) Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics* 94: 83-88

Dixon J, Trainor P, Dixon MJ (2007) Treacher Collins syndrome. *Orthod Craniofac Res* 10: 88-95

Dosil M, Bustelo XR (2004) Functional characterization of Pwp2, a WD family protein essential for the assembly of the 90 S pre-ribosomal particle. *The Journal of biological chemistry* 279: 37385-37397

Douet J, Tourmente S (2007) Transcription of the 5S rRNA heterochromatic genes is epigenetically controlled in Arabidopsis thaliana and Xenopus laevis. *Heredity* 99: 5-13

Dragon F, Gallagher JE, Compagnone-Post PA, Mitchell BM, Porwancher KA, Wehner KA, Wormsley S, Settlage RE, Shabanowitz J, Osheim Y, Beyer AL, Hunt DF, Baserga SJ (2002) A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* 417: 967-970

Dutt S, Narla A, Lin K, Mullally A, Abayasekara N, Megerdichian C, Wilson FH, Currie T, Khanna-Gupta A, Berliner N, Kutok JL, Ebert BL (2011) Haploinsufficiency for ribosomal protein genes causes selective activation of p53 in human erythroid progenitor cells. *Blood* 117: 2567-2576

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

Ebert BL, Pretz J, Bosco J, Chang CY, Tamayo P, Galili N, Raza A, Root DE, Attar E, Ellis SR, Golub TR (2008) Identification of RPS14 as a 5q- syndrome gene by RNA interference screen. *Nature* 451: 335-339

El Hage A, Koper M, Kufel J, Tollervey D (2008) Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. *Genes & development* 22: 1069-1081

Elbashir SM, Harborth J, Weber K, Tuschl T (2002) Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* 26: 199-213

Elela SA, Igel H, Ares M, Jr. (1996) RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. *Cell* 85: 115-124

Ellis SR, Gleizes PE (2011) Diamond Blackfan anemia: ribosomal proteins going rogue. *Seminars in hematology* 48: 89-96

Enright CA, Maxwell ES, Eliceiri GL, Sollner-Webb B (1996) 5'ETS rRNA processing facilitated by four small RNAs: U14, E3, U17, and U3. *RNA (New York, NY* 2: 1094-1099

Eppens NA, Faber AW, Rondaij M, Jahangir RS, van Hemert S, Vos JC, Venema J, Raue HA (2002) Deletions in the S1 domain of Rrp5p cause processing at a novel site in ITS1 of yeast pre-rRNA that depends on Rex4p. *Nucleic acids research* 30: 4222-4231

Esguerra J, Warringer J, Blomberg A (2008) Functional importance of individual rRNA 2'-O-ribose methylations revealed by high-resolution phenotyping. *RNA (New York, NY* 14: 649-656

Faber AW, Van Dijk M, Raue HA, Vos JC (2002) Ngl2p is a Ccr4p-like RNA nuclease essential for the final step in 3'-end processing of 5.8S rRNA in Saccharomyces cerevisiae. *RNA (New York, NY* 8: 1095-1101

Fang F, Hoskins J, Butler JS (2004) 5-fluorouracil enhances exosome-dependent accumulation of polyadenylated rRNAs. *Molecular and cellular biology* 24: 10766-10776

Fasken MB, Leung SW, Banerjee A, Kodani MO, Chavez R, Bowman EA, Purohit MK, Rubinson ME, Rubinson EH, Corbett AH (2011) Air1 zinc knuckles 4 and 5 and a conserved IWRxY motif are critical for the function and integrity of the TRAMP RNA quality control complex. *The Journal of biological chemistry*

Fatica A, Oeffinger M, Dlakic M, Tollervey D (2003a) Nob1p is required for cleavage of the 3' end of 18S rRNA. *Molecular and cellular biology* 23: 1798-1807

Fatica A, Oeffinger M, Tollervey D, Bozzoni I (2003b) Cic1p/Nsa3p is required for synthesis and nuclear export of 60S ribosomal subunits. *RNA (New York, NY* 9: 1431-1436

Fatica A, Tollervey D, Dlakic M (2004) PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA. *RNA (New York, NY* 10: 1698-1701

Filipowicz W, Pogacic V (2002) Biogenesis of small nucleolar ribonucleoproteins. *Current opinion in cell biology* 14: 319-327

Freed EF, Bleichert F, Dutca LM, Baserga SJ (2010) When ribosomes go bad: diseases of ribosome biogenesis. *Mol Biosyst* 6: 481-493

Fromont-Racine M, Senger B, Saveanu C, Fasiolo F (2003) Ribosome assembly in eukaryotes. *Gene* 313: 17-42

Fujiwara T, Suzuki S, Kanno M, Sugiyama H, Takahashi H, Tanaka J (2006) Mapping a nucleolar targeting sequence of an RNA binding nucleolar protein, Nop25. *Exp Cell Res* 312: 1703-1712

Fukuda M, Asano S, Nakamura T, Adachi M, Yoshida M, Yanagida M, Nishida E (1997) CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* 390: 308-311

Fukuda MN, Miyoshi M, Nadano D (2008) The role of bystin in embryo implantation and in ribosomal biogenesis. *Cell Mol Life Sci* 65: 92-99

Galani K, Nissan TA, Petfalski E, Tollervey D, Hurt E (2004) Rea1, a dynein-related nuclear AAA-ATPase, is involved in late rRNA processing and nuclear export of 60 S subunits. *The Journal of biological chemistry* 279: 55411-55418

Galardi S, Fatica A, Bachi A, Scaloni A, Presutti C, Bozzoni I (2002) Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA in vitro. *Molecular and cellular biology* 22: 6663-6668

Gallagher JE, Dunbar DA, Granneman S, Mitchell BM, Osheim Y, Beyer AL, Baserga SJ (2004) RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. *Genes & development* 18: 2506-2517

Ganot P, Caizergues-Ferrer M, Kiss T (1997) The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes & development* 11: 941-956

Gautier T, Berges T, Tollervey D, Hurt E (1997) Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. *Molecular and cellular biology* 17: 7088-7098

Geerlings TH, Faber AW, Bister MD, Vos JC, Raue HA (2003) Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20 S Pre-rRNA in Saccharomyces cerevisiae. *The Journal of biological chemistry* 278: 22537-22545

Geerlings TH, Vos JC, Raue HA (2000) The final step in the formation of 25S rRNA in Saccharomyces cerevisiae is performed by 5'-->3' exonucleases. *RNA (New York, NY* 6: 1698-1703

Gelpi C, Alguero A, Angeles Martinez M, Vidal S, Juarez C, Rodriguez-Sanchez JL (1990) Identification of protein components reactive with anti-PM/Scl autoantibodies. *Clinical and experimental immunology* 81: 59-64

Gerbi SA, Borovjagin AV, Lange TS (2003) The nucleolus: a site of ribonucleoprotein maturation. *Current opinion in cell biology* 15: 318-325

Gerczei T, Correll CC (2004) Imp3p and Imp4p mediate formation of essential U3-precursor rRNA (pre-rRNA) duplexes, possibly to recruit the small subunit processome to the pre-rRNA. *Proceedings of the National Academy of Sciences of the United States of America* 101: 15301-15306

Gerczei T, Shah BN, Manzo AJ, Walter NG, Correll CC (2009) RNA chaperones stimulate formation and yield of the U3 snoRNA-Pre-rRNA duplexes needed for eukaryotic ribosome biogenesis. *Journal of molecular biology* 390: 991-1006

Gerus M, Bonnart C, Caizergues-Ferrer M, Henry Y, Henras AK (2010) Evolutionarily conserved function of RRP36 in early cleavages of the pre-rRNA and production of the 40S ribosomal subunit. *Molecular and cellular biology* 30: 1130-1144

Gherzi R, Lee KY, Briata P, Wegmuller D, Moroni C, Karin M, Chen CY (2004) A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. *Molecular cell* 14: 571-583

Gilbert WV (2011) Functional specialization of ribosomes? *Trends in biochemical sciences* 36: 127-132

Gill T, Cai T, Aulds J, Wierzbicki S, Schmitt ME (2004) RNase MRP cleaves the CLB2 mRNA to promote cell cycle progression: novel method of mRNA degradation. *Molecular and cellular biology* 24: 945-953

Glazov EA, Zankl A, Donskoi M, Kenna TJ, Thomas GP, Clark GR, Duncan EL, Brown MA (2011) Whole-exome re-sequencing in a family quartet identifies POP1 mutations as the cause of a novel skeletal dysplasia. *PLoS genetics* 7: e1002027

Grandi P, Rybin V, Bassler J, Petfalski E, Strauss D, Marzioch M, Schafer T, Kuster B, Tschochner H, Tollervey D, Gavin AC, Hurt E (2002) 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Molecular cell* 10: 105-115

Granneman S, Baserga SJ (2003) Probing the yeast proteome for RNA-processing factors. *Genome Biol* 4: 229

Granneman S, Baserga SJ (2004) Ribosome biogenesis: of knobs and RNA processing. ***Exp Cell Res*** **296: 43-50**

Granneman S, Bernstein KA, Bleichert F, Baserga SJ (2006) Comprehensive mutational analysis of yeast DEXD/H box RNA helicases required for small ribosomal subunit synthesis. ***Molecular and cellular biology*** **26: 1183-1194**

Granneman S, Gallagher JE, Vogelzangs J, Horstman W, van Venrooij WJ, Baserga SJ, Pruijn GJ (2003) The human Imp3 and Imp4 proteins form a ternary complex with hMpp10, which only interacts with the U3 snoRNA in 60-80S ribonucleoprotein complexes. ***Nucleic acids research*** **31: 1877-1887**

Granneman S, Petfalski E, Swiatkowska A, Tollervey D (2010) Cracking pre-40S ribosomal subunit structure by systematic analyses of RNA-protein cross-linking. ***The EMBO journal*** **29: 2026-2036**

Granneman S, Petfalski E, Tollervey D (2011) A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. ***The EMBO journal*** **30: 4006-4019**

Granneman S, Pruijn GJ, Horstman W, van Venrooij WJ, Luhrmann R, Watkins NJ (2002) The hU3-55K protein requires 15.5K binding to the box B/C motif as well as flanking RNA elements for its association with the U3 small nucleolar RNA in Vitro. ***The Journal of biological chemistry*** **277: 48490-48500**

Granneman S, Vogelzangs J, Luhrmann R, van Venrooij WJ, Pruijn GJ, Watkins NJ (2004) Role of pre-rRNA base pairing and 80S complex formation in subnucleolar localization of the U3 snoRNP. ***Molecular and cellular biology*** **24: 8600-8610**

Grudnik P, Bange G, Sinning I (2009) Protein targeting by the signal recognition particle. ***Biological chemistry*** **390: 775-782**

Guo X, Ma J, Sun J, Gao G (2007) The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. ***Proceedings of the National Academy of Sciences of the United States of America*** **104: 151-156**

Gupta V, Kumar A (2010) Dyskeratosis congenita. ***Advances in experimental medicine and biology*** **685: 215-219**

Guttler T, Gorlich D (2011) Ran-dependent nuclear export mediators: a structural perspective. ***The EMBO journal*** **30: 3457-3474**

Guttler T, Madl T, Neumann P, Deichsel D, Corsini L, Monecke T, Ficner R, Sattler M, Gorlich D (2010) NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. ***Nature structural & molecular biology*** **17: 1367-1376**

Hadjiolova KV, Georgiev OI, Nosikov VV, Hadjiolov AA (1984) Mapping of the major early endonuclease cleavage site of the rat precursor to rRNA within the internal transcribed spacer sequence of rDNA. *Biochimica et biophysica acta* 782: 195-201

Hadjiolova KV, Nicoloso M, Mazan S, Hadjiolov AA, Bachellerie JP (1993) Alternative pre-rRNA processing pathways in human cells and their alteration by cycloheximide inhibition of protein synthesis. *Eur J Biochem* 212: 211-215

Hamill S, Wolin SL, Reinisch KM (2010) Structure and function of the polymerase core of TRAMP, a RNA surveillance complex. *Proceedings of the National Academy of Sciences of the United States of America* 107: 15045-15050

Hands-Taylor KL, Martino L, Tata R, Babon JJ, Bui TT, Drake AF, Beavil RL, Pruijn GJ, Brown PR, Conte MR (2010) Heterodimerization of the human RNase P/MRP subunits Rpp20 and Rpp25 is a prerequisite for interaction with the P3 arm of RNase MRP RNA. *Nucleic acids research* 38: 4052-4066

Hartung S, Hopfner KP (2009) Lessons from structural and biochemical studies on the archaeal exosome. *Biochemical Society transactions* 37: 83-87

Helm M (2006) Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic acids research* 34: 721-733

Henras A, Henry Y, Bousquet-Antonelli C, Noaillac-Depeyre J, Gelugne JP, Caizergues-Ferrer M (1998) Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *The EMBO journal* 17: 7078-7090

Henras AK, Capeyrou R, Henry Y, Caizergues-Ferrer M (2004) Cbf5p, the putative pseudouridine synthase of H/ACA-type snoRNPs, can form a complex with Gar1p and Nop10p in absence of Nhp2p and box H/ACA snoRNAs. *RNA (New York, NY* 10: 1704-1712

Henras AK, Soudet J, Gerus M, Lebaron S, Caizergues-Ferrer M, Mougin A, Henry Y (2008) The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci* 65: 2334-2359

Henry Y, Wood H, Morrissey JP, Petfalski E, Kearsey S, Tollervey D (1994) The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. *The EMBO journal* 13: 2452-2463

Hernandez-Verdun D, Roussel P, Thiry M, Sirri V, Lafontaine DL (2010) The nucleolus: structure/function relationship in RNA metabolism. *Wiley interdisciplinary reviews* 1: 415-431

Herrera JE, Savkur R, Olson MO (1995) The ribonuclease activity of nucleolar protein B23. *Nucleic acids research* 23: 3974-3979

Ho JH, Kallstrom G, Johnson AW (2000) Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. *The Journal of cell biology* 151: 1057-1066

Hoang T, Peng WT, Vanrobays E, Krogan N, Hiley S, Beyer AL, Osheim YN, Greenblatt J, Hughes TR, Lafontaine DL (2005) Esf2p, a U3-associated factor required for small-subunit processome assembly and compaction. *Molecular and cellular biology* 25: 5523-5534

Holzel M, Burger K, Muhl B, Orban M, Kellner M, Eick D (2010) The tumor suppressor p53 connects ribosome biogenesis to cell cycle control: a double-edged sword. *Oncotarget* 1: 43-47

Holzel M, Rohrmoser M, Schlee M, Grimm T, Harasim T, Malamoussi A, Gruber-Eber A, Kremmer E, Hiddemann W, Bornkamm GW, Eick D (2005) Mammalian WDR12 is a novel member of the Pes1-Bop1 complex and is required for ribosome biogenesis and cell proliferation. *The Journal of cell biology* 170: 367-378

Horn DM, Mason SL, Karbstein K (2011) RCL1, a novel nuclease for 18S rRNA production. *The Journal of biological chemistry*

Horos R, Ijspeert H, Pospisilova D, Sendtner R, Andrieu-Soler C, Taskesen E, Nieradka A, Cmejla R, Sendtner M, Touw IP, von Lindern M (2011) Ribosomal deficiencies in Diamond-Blackfan anemia impair translation of transcripts essential for differentiation of murine and human erythroblasts. *Blood*

Hoskins J, Scott Butler J (2007) Evidence for distinct DNA- and RNA-based mechanisms of 5-fluorouracil cytotoxicity in Saccharomyces cerevisiae. *Yeast* 24: 861-870

Houseley J, LaCava J, Tollervey D (2006) RNA-quality control by the exosome. *Nat Rev Mol Cell Biol* 7: 529-539

Houseley J, Tollervey D (2009) The many pathways of RNA degradation. *Cell* 136: 763-776

Hughes JM (1996) Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA. *Journal of molecular biology* 259: 645-654

Hung NJ, Lo KY, Patel SS, Helmke K, Johnson AW (2008) Arx1 is a nuclear export receptor for the 60S ribosomal subunit in yeast. *Molecular biology of the cell* 19: 735-744

Huntzinger E, Kashima I, Fauser M, Sauliere J, Izaurralde E (2008) SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA (New York, NY* 14: 2609-2617

Iadevaia V, Zhang Z, Jan E, Proud CG (2011) mTOR signaling regulates the processing of pre-rRNA in human cells. ***Nucleic acids research***

Idol RA, Robledo S, Du HY, Crimmins DL, Wilson DB, Ladenson JH, Bessler M, Mason PJ (2007) Cells depleted for RPS19, a protein associated with Diamond Blackfan Anemia, show defects in 18S ribosomal RNA synthesis and small ribosomal subunit production. ***Blood Cells Mol Dis*** **39: 35-43**

Isken O, Maquat LE (2007) Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. ***Genes & development*** **21: 1833-1856**

Jackson RN, Klauer AA, Hintze BJ, Robinson H, van Hoof A, Johnson SJ (2010) The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing. ***The EMBO journal*** **29: 2205-2216**

Jacobson MR, Cao LG, Wang YL, Pederson T (1995) Dynamic localization of RNase MRP RNA in the nucleolus observed by fluorescent RNA cytochemistry in living cells. ***The Journal of cell biology*** **131: 1649-1658**

Januszyk K, Liu Q, Lima CD (2011) Activities of human RRP6 and structure of the human RRP6 catalytic domain. ***RNA (New York, NY*** **17: 1566-1577**

Jia H, Wang X, Liu F, Guenther UP, Srinivasan S, Anderson JT, Jankowsky E (2011) The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex. ***Cell*** **145: 890-901**

Jiao X, Xiang S, Oh C, Martin CE, Tong L, Kiledjian M (2010) Identification of a quality-control mechanism for mRNA 5'-end capping. ***Nature*** **467: 608-611**

Johnson AW (2001) Rat1p nuclease. ***Methods in enzymology*** **342: 260-268**

Kadaba S, Krueger A, Trice T, Krecic AM, Hinnebusch AG, Anderson J (2004) Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae. ***Genes & development*** **18: 1227-1240**

Kadaba S, Wang X, Anderson JT (2006) Nuclear RNA surveillance in Saccharomyces cerevisiae: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5S rRNA. ***RNA (New York, NY*** **12: 508-521**

Karbstein K, Doudna JA (2006) GTP-dependent formation of a ribonucleoprotein subcomplex required for ribosome biogenesis. ***Journal of molecular biology*** **356: 432-443**

Karbstein K, Jonas S, Doudna JA (2005) An essential GTPase promotes assembly of preribosomal RNA processing complexes. ***Molecular cell*** **20: 633-643**

Kass S, Tyc K, Steitz JA, Sollner-Webb B (1990) The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. *Cell* 60: 897-908

Kent T, Lapik YR, Pestov DG (2009) The 5' external transcribed spacer in mouse ribosomal RNA contains two cleavage sites. *RNA (New York, NY* 15: 14-20

Killian A, Sarafan-Vasseur N, Sesboue R, Le Pessot F, Blanchard F, Lamy A, Laurent M, Flaman JM, Frebourg T (2006) Contribution of the BOP1 gene, located on 8q24, to colorectal tumorigenesis. *Genes, chromosomes & cancer* 45: 874-881

Kim M, Krogan NJ, Vasiljeva L, Rando OJ, Nedea E, Greenblatt JF, Buratowski S (2004) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* 432: 517-522

Kiss-Laszlo Z, Henry Y, Bachellerie JP, Caizergues-Ferrer M, Kiss T (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* 85: 1077-1088

Kiss-Laszlo Z, Henry Y, Kiss T (1998) Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *The EMBO journal* 17: 797-807

Kiss DL, Andrulis ED (2011) The exozyme model: a continuum of functionally distinct complexes. *RNA (New York, NY* 17: 1-13

Kiss T, Fayet E, Jady BE, Richard P, Weber M (2006) Biogenesis and intranuclear trafficking of human box C/D and H/ACA RNPs. *Cold Spring Harb Symp Quant Biol* 71: 407-417

Klinge S, Voigts-Hoffmann F, Leibundgut M, Arpagaus S, Ban N (2011) Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. *Science* 334: 941-948

Knox AA, McKeegan KS, Debieux CM, Traynor A, Richardson H, Watkins NJ (2011) A weak C' box renders U3 snoRNA levels dependent on hU3-55K binding. *Molecular and cellular biology* 31: 2404-2412

Kobayashi T, Heck DJ, Nomura M, Horiuchi T (1998) Expansion and contraction of ribosomal DNA repeats in Saccharomyces cerevisiae: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. *Genes & development* 12: 3821-3830

Kos M, Tollervey D (2010) Yeast pre-rRNA processing and modification occur cotranscriptionally. *Molecular cell* 37: 809-820

Kressler D, Linder P, de La Cruz J (1999) Protein trans-acting factors involved in ribosome biogenesis in Saccharomyces cerevisiae. *Molecular and cellular biology* 19: 7897-7912

Krishnamurthy N, Ngam CR, Berdis AJ, Montano MM (2011) The exonuclease activity of hPMC2 is required for transcriptional regulation of the QR gene and repair of estrogen-induced abasic sites. *Oncogene*

Krogan NJ, Peng WT, Cagney G, Robinson MD, Haw R, Zhong G, Guo X, Zhang X, Canadien V, Richards DP, Beattie BK, Lalev A, Zhang W, Davierwala AP, Mnaimneh S, Starostine A, Tikuisis AP, Grigull J, Datta N, Bray JE, Hughes TR, Emili A, Greenblatt JF (2004) High-definition macromolecular composition of yeast RNA-processing complexes. *Molecular cell* 13: 225-239

Kufel J, Allmang C, Verdone L, Beggs J, Tollervey D (2003) A complex pathway for 3' processing of the yeast U3 snoRNA. *Nucleic acids research* 31: 6788-6797

Kufel J, Dichtl B, Tollervey D (1999) Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3' ETS but not the 5' ETS. *RNA (New York, NY* 5: 909-917

LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* 121: 713-724

Lafontaine D, Vandenhaute J, Tollervey D (1995) The 18S rRNA dimethylase Dim1p is required for pre-ribosomal RNA processing in yeast. *Genes & development* 9: 2470-2481

Lafontaine DL, Bousquet-Antonelli C, Henry Y, Caizergues-Ferrer M, Tollervey D (1998a) The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes & development* 12: 527-537

Lafontaine DL, Preiss T, Tollervey D (1998b) Yeast 18S rRNA dimethylase Dim1p: a quality control mechanism in ribosome synthesis? *Molecular and cellular biology* 18: 2360-2370

Lafontaine DL, Tollervey D (1999) Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *RNA (New York, NY* 5: 455-467

Lafontaine DL, Tollervey D (2000) Synthesis and assembly of the box C+D small nucleolar RNPs. *Molecular and cellular biology* 20: 2650-2659

Lafontaine DL, Tollervey D (2001) The function and synthesis of ribosomes. *Nat Rev Mol Cell Biol* 2: 514-520

Lamanna AC, Karbstein K (2009) Nob1 binds the single-stranded cleavage site D at the 3'-end of 18S rRNA with its PIN domain. *Proceedings of the National Academy of Sciences of the United States of America* 106: 14259-14264

Lamanna AC, Karbstein K (2011) An RNA conformational switch regulates pre-18S rRNA cleavage. *Journal of molecular biology* 405: 3-17

Lange H, Sement FM, Gagliardi D (2011) MTR4, a putative RNA helicase and exosome co-factor, is required for proper rRNA biogenesis and development in Arabidopsis thaliana. *The Plant journal : for cell and molecular biology* 68: 51-63

Lapik YR, Fernandes CJ, Lau LF, Pestov DG (2004) Physical and functional interaction between Pes1 and Bop1 in mammalian ribosome biogenesis. *Molecular cell* 15: 17-29

LaRiviere FJ, Cole SE, Ferullo DJ, Moore MJ (2006) A late-acting quality control process for mature eukaryotic rRNAs. *Molecular cell* 24: 619-626

LaRonde-LeBlanc N, Guszczynski T, Copeland T, Wlodawer A (2005) Autophosphorylation of Archaeoglobus fulgidus Rio2 and crystal structures of its nucleotide-metal ion complexes. *FEBS J* 272: 2800-2810

LaRonde-LeBlanc N, Wlodawer A (2004) Crystal structure of A. fulgidus Rio2 defines a new family of serine protein kinases. *Structure* 12: 1585-1594

LaRonde-LeBlanc N, Wlodawer A (2005a) A family portrait of the RIO kinases. *The Journal of biological chemistry* 280: 37297-37300

LaRonde-LeBlanc N, Wlodawer A (2005b) The RIO kinases: an atypical protein kinase family required for ribosome biogenesis and cell cycle progression. *Biochimica et biophysica acta* 1754: 14-24

Lebaron S, Froment C, Fromont-Racine M, Rain JC, Monsarrat B, Caizergues-Ferrer M, Henry Y (2005) The splicing ATPase prp43p is a component of multiple preribosomal particles. *Molecular and cellular biology* 25: 9269-9282

Lebaron S, Papin C, Capeyrou R, Chen YL, Froment C, Monsarrat B, Caizergues-Ferrer M, Grigoriev M, Henry Y (2009) The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis. *The EMBO journal* 28: 3808-3819

Lebreton A, Tomecki R, Dziembowski A, Seraphin B (2008) Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature* 456: 993-996

Lee SJ, Baserga SJ (1999) Imp3p and Imp4p, two specific components of the U3 small nucleolar ribonucleoprotein that are essential for pre-18S rRNA processing. *Molecular and cellular biology* 19: 5441-5452

Leger-Silvestre I, Caffrey JM, Dawaliby R, Alvarez-Arias DA, Gas N, Bertolone SJ, Gleizes PE, Ellis SR (2005) Specific Role for Yeast Homologs of the Diamond Blackfan Anemia-associated Rps19 Protein in Ribosome Synthesis. *The Journal of biological chemistry* 280: 38177-38185

Leger-Silvestre I, Milkereit P, Ferreira-Cerca S, Saveanu C, Rousselle JC, Choesmel V, Guinefoleau C, Gas N, Gleizes PE (2004) The ribosomal protein Rps15p is required for nuclear exit of the 40S subunit precursors in yeast. *The EMBO journal* 23: 2336-2347

Lehner B, Sanderson CM (2004) A protein interaction framework for human mRNA degradation. *Genome Res* 14: 1315-1323

Lemay V, Hossain A, Osheim YN, Beyer AL, Dragon F (2011) Identification of novel proteins associated with yeast snR30 small nucleolar RNA. *Nucleic acids research* 39: 9659-9670

Lempicki RA, Jarmolowski A, Huang GY, Li HV, Fournier MJ (1990) Mutations in conserved domains of U14 RNA impair 18S ribosomal RNA production in Saccharomyces cerevisiae. *Molecular biology reports* 14: 119-120

Li HD, Zagorski J, Fournier MJ (1990) Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in Saccharomyces cerevisiae. *Molecular and cellular biology* 10: 1145-1152

Li K, Smagula CS, Parsons WJ, Richardson JA, Gonzalez M, Hagler HK, Williams RS (1994) Subcellular partitioning of MRP RNA assessed by ultrastructural and biochemical analysis. *The Journal of cell biology* 124: 871-882

Liang XH, Crooke ST (2011) Depletion of key protein components of the RISC pathway impairs pre-ribosomal RNA processing. *Nucleic acids research* 39: 4875-4889

Liang XH, Liu Q, Fournier MJ (2007) rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Molecular cell* 28: 965-977

Lin J, Lai S, Jia R, Xu A, Zhang L, Lu J, Ye K (2011) Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature* 469: 559-563

Lindahl L, Bommankanti A, Li X, Hayden L, Jones A, Khan M, Oni T, Zengel JM (2009) RNase MRP is required for entry of 35S precursor rRNA into the canonical processing pathway. *RNA (New York, NY* 15: 1407-1416

Liu Q, Greimann JC, Lima CD (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* 127: 1223-1237

Long RM, McNally MT (2003) mRNA decay: x (XRN1) marks the spot. *Molecular cell* 11: 1126-1128

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. *Molecular cell* 29: 717-728

Lorentzen E, Walter P, Fribourg S, Evguenieva-Hackenberg E, Klug G, Conti E (2005) The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nat Struct Mol Biol* 12: 575-581

Lubas M, Christensen MS, Kristiansen MS, Domanski M, Falkenby LG, Lykke-Andersen S, Andersen JS, Dziembowski A, Jensen TH (2011) Interaction profiling identifies the human nuclear exosome targeting complex. *Molecular cell* 43: 624-637

Lubben B, Marshallsay C, Rottmann N, Luhrmann R (1993) Isolation of U3 snoRNP from CHO cells: a novel 55 kDa protein binds to the central part of U3 snoRNA. *Nucleic acids research* 21: 5377-5385

Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U (2004) Nuclear export of microRNA precursors. *Science* 303: 95-98

Lygerou Z, Allmang C, Tollervey D, Seraphin B (1996) Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP in vitro. *Science* 272: 268-270

Lygerou Z, Mitchell P, Petfalski E, Seraphin B, Tollervey D (1994) The POP1 gene encodes a protein component common to the RNase MRP and RNase P ribonucleoproteins. *Genes & development* 8: 1423-1433

Lykke-Andersen S, Brodersen DE, Jensen TH (2009) Origins and activities of the eukaryotic exosome. *Journal of cell science* 122: 1487-1494

Lyman SK, Gerace L, Baserga SJ (1999) Human Nop5/Nop58 is a component common to the box C/D small nucleolar ribonucleoproteins. *RNA (New York, NY* 5: 1597-1604

MacCallum DE, Hall PA (2000) The location of pKi67 in the outer dense fibrillary compartment of the nucleolus points to a role in ribosome biogenesis during the cell division cycle. *J Pathol* 190: 537-544

Maida Y, Yasukawa M, Furuuchi M, Lassmann T, Possemato R, Okamoto N, Kasim V, Hayashizaki Y, Hahn WC, Masutomi K (2009) An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. *Nature* 461: 230-235

Mattaj IW, Muller CW (2010) Solving the NES problem. *Nat Struct Mol Biol* 17: 1288-1289

Mattijssen S, Hinson ER, Onnekink C, Hermanns P, Zabel B, Cresswell P, Pruijn GJ (2010a) Viperin mRNA is a novel target for the human RNase MRP/RNase P endoribonuclease. *Cell Mol Life Sci* 68: 2469-2480

Mattijssen S, Hinson ER, Onnekink C, Hermanns P, Zabel B, Cresswell P, Pruijn GJ (2011) Viperin mRNA is a novel target for the human RNase MRP/RNase P endoribonuclease. *Cell Mol Life Sci* 68: 2469-2480

Mattijssen S, Welting TJ, Pruijn GJ (2010b) RNase MRP and disease. *Wiley interdisciplinary reviews* 1: 102-116

Mayer C, Schmitz KM, Li J, Grummt I, Santoro R (2006) Intergenic transcripts regulate the epigenetic state of rRNA genes. *Molecular cell* 22: 351-361

McKeegan KS, Debieux CM, Boulon S, Bertrand E, Watkins NJ (2007) A dynamic scaffold of pre-snoRNP factors facilitates human box C/D snoRNP assembly. *Molecular and cellular biology* 27: 6782-6793

McKeegan KS, Debieux CM, Watkins NJ (2009) Evidence that the AAA+ proteins TIP48 and TIP49 bridge interactions between 15.5K and the related NOP56 and NOP58 proteins during box C/D snoRNP biogenesis. *Molecular and cellular biology* 29: 4971-4981

Meinl W, Ebert B, Glatt H, Lampen A (2008) Sulfotransferase forms expressed in human intestinal Caco-2 and TC7 cells at varying stages of differentiation and role in benzo[a]pyrene metabolism. *Drug metabolism and disposition: the biological fate of chemicals* 36: 276-283

Merl J, Jakob S, Ridinger K, Hierlmeier T, Deutzmann R, Milkereit P, Tschochner H (2010) Analysis of ribosome biogenesis factor-modules in yeast cells depleted from pre-ribosomes. *Nucleic acids research* 38: 3068-3080

Michot B, Joseph N, Mazan S, Bachellerie JP (1999) Evolutionarily conserved structural features in the ITS2 of mammalian pre-rRNAs and potential interactions with the snoRNA U8 detected by comparative analysis of new mouse sequences. *Nucleic acids research* 27: 2271-2282

Midtgaard SF, Assenholt J, Jonstrup AT, Van LB, Jensen TH, Brodersen DE (2006) Structure of the nuclear exosome component Rrp6p reveals an interplay between the active site and the HRDC domain. *Proceedings of the National Academy of Sciences of the United States of America* 103: 11898-11903

Miles TD, Jakovljevic J, Horsey EW, Harnpicharnchai P, Tang L, Woolford JL, Jr. (2005) Ytm1, Nop7, and Erb1 form a complex necessary for maturation of yeast 66S preribosomes. *Molecular and cellular biology* 25: 10419-10432

Milkereit P, Gadal O, Podtelejnikov A, Trumtel S, Gas N, Petfalski E, Tollervey D, Mann M, Hurt E, Tschochner H (2001) Maturation and intranuclear transport of pre-ribosomes requires Noc proteins. ***Cell*** **105: 499-509**

Milkereit P, Strauss D, Bassler J, Gadal O, Kuhn H, Schutz S, Gas N, Lechner J, Hurt E, Tschochner H (2003) A Noc complex specifically involved in the formation and nuclear export of ribosomal 40 S subunits. ***The Journal of biological chemistry*** **278: 4072-4081**

Miller OL, Jr., Beatty BR (1969) Visualization of nucleolar genes. ***Science (New York, NY*** **164: 955-957**

Milligan L, Decourty L, Saveanu C, Rappsilber J, Ceulemans H, Jacquier A, Tollervey D (2008) A yeast exosome cofactor, Mpp6, functions in RNA surveillance and in the degradation of noncoding RNA transcripts. ***Molecular and cellular biology*** **28: 5446-5457**

Mitchell P, Petfalski E, Houalla R, Podtelejnikov A, Mann M, Tollervey D (2003) Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs. ***Molecular and cellular biology*** **23: 6982-6992**

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: 457-466**

Mitchell P, Petfalski E, Tollervey D (1996) The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism. ***Genes & development*** **10: 502-513**

Miyoshi M, Okajima T, Matsuda T, Fukuda MN, Nadano D (2007) Bystin in human cancer cells: intracellular localization and function in ribosome biogenesis. ***Biochem J*** **404: 373-381**

Monecke T, Guttler T, Neumann P, Dickmanns A, Gorlich D, Ficner R (2009) Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. ***Science (New York, NY*** **324: 1087-1091**

Moritz M, Paulovich AG, Tsay YF, Woolford JL, Jr. (1990) Depletion of yeast ribosomal proteins L16 or rp59 disrupts ribosome assembly. ***The Journal of cell biology*** **111: 2261-2274**

Moritz M, Pulaski BA, Woolford JL, Jr. (1991) Assembly of 60S ribosomal subunits is perturbed in temperature-sensitive yeast mutants defective in ribosomal protein L16. ***Molecular and cellular biology*** **11: 5681-5692**

Morrissey JP, Tollervey D (1993) Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. ***Molecular and cellular biology*** **13: 2469-2477**

Moser MJ, Holley WR, Chatterjee A, Mian IS (1997) The proofreading domain of Escherichia coli DNA polymerase I and other DNA and/or RNA exonuclease domains. *Nucleic acids research* 25: 5110-5118

Mouaikel J, Verheggen C, Bertrand E, Tazi J, Bordonne R (2002) Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Molecular cell* 9: 891-901

Mougey EB, O'Reilly M, Osheim Y, Miller OL, Jr., Beyer A, Sollner-Webb B (1993) The terminal balls characteristic of eukaryotic rRNA transcription units in chromatin spreads are rRNA processing complexes. *Genes & development* 7: 1609-1619

Mukherjee D, Gao M, O'Connor JP, Raijmakers R, Pruijn G, Lutz CS, Wilusz J (2002) The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements. *The EMBO journal* 21: 165-174

Mullen TE, Marzluff WF (2008) Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes & development* 22: 50-65

Muro E, Hoang TQ, Jobart-Malfait A, Hernandez-Verdun D (2008) In nucleoli, the steady state of nucleolar proteins is leptomycin B-sensitive. *Biology of the cell / under the auspices of the European Cell Biology Organization* 100: 303-313

Narla A, Ebert BL (2010) Ribosomopathies: human disorders of ribosome dysfunction. *Blood* 115: 3196-3205

Narla A, Hurst SN, Ebert BL (2010) Ribosome defects in disorders of erythropoiesis. *International journal of hematology* 93: 144-149

Nazar RN (2004) Ribosomal RNA processing and ribosome biogenesis in eukaryotes. *IUBMB life* 56: 457-465

Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* 457: 1038-1042

Newton K, Petfalski E, Tollervey D, Caceres JF (2003) Fibrillarin is essential for early development and required for accumulation of an intron-encoded small nucleolar RNA in the mouse. *Molecular and cellular biology* 23: 8519-8527

Ni J, Tien AL, Fournier MJ (1997) Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* 89: 565-573

Nousbeck J, Spiegel R, Ishida-Yamamoto A, Indelman M, Shani-Adir A, Adir N, Lipkin E, Bercovici S, Geiger D, van Steensel MA, Steijlen PM, Bergman R, Bindereif A, Choder M, Shalev S, Sprecher E (2008) Alopecia, neurological defects, and endocrinopathy syndrome caused by decreased expression of RBM28, a nucleolar protein associated with ribosome biogenesis. ***Am J Hum Genet*** **82: 1114-1121**

O'Donohue MF, Choesmel V, Faubladier M, Fichant G, Gleizes PE (2010) Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. ***The Journal of cell biology*** **190: 853-866**

Oeffinger M, Dlakic M, Tollervey D (2004) A pre-ribosome-associated HEAT-repeat protein is required for export of both ribosomal subunits. ***Genes & development*** **18: 196-209**

Oeffinger M, Leung A, Lamond A, Tollervey D (2002) Yeast Pescadillo is required for multiple activities during 60S ribosomal subunit synthesis. ***RNA (New York, NY*** **8: 626-636**

Oeffinger M, Tollervey D (2003) Yeast Nop15p is an RNA-binding protein required for pre-rRNA processing and cytokinesis. ***The EMBO journal*** **22: 6573-6583**

Oeffinger M, Zenklusen D, Ferguson A, Wei KE, El Hage A, Tollervey D, Chait BT, Singer RH, Rout MP (2009) Rrp17p is a eukaryotic exonuclease required for 5' end processing of Pre-60S ribosomal RNA. ***Molecular cell*** **36: 768-781**

Orban TI, Izaurralde E (2005) Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. ***RNA (New York, NY*** **11: 459-469**

Osheim YN, French SL, Keck KM, Champion EA, Spasov K, Dragon F, Baserga SJ, Beyer AL (2004) Pre-18S ribosomal RNA is structurally compacted into the SSU processome prior to being cleaved from nascent transcripts in Saccharomyces cerevisiae. ***Molecular cell*** **16: 943-954**

Oskarsson T, Trumpp A (2005) The Myc trilogy: lord of RNA polymerases. ***Nat Cell Biol*** **7: 215-217**

Panse VG, Johnson AW (2010) Maturation of eukaryotic ribosomes: acquisition of functionality. ***Trends in biochemical sciences*** **35: 260-266**

Peculis BA (1997) The sequence of the 5' end of the U8 small nucleolar RNA is critical for 5.8S and 28S rRNA maturation. ***Molecular and cellular biology*** **17: 3702-3713**

Peculis BA, Steitz JA (1993) Disruption of U8 nucleolar snRNA inhibits 5.8S and 28S rRNA processing in the Xenopus oocyte. ***Cell*** **73: 1233-1245**

Peculis BA, Steitz JA (1994) Sequence and structural elements critical for U8 snRNP function in Xenopus oocytes are evolutionarily conserved. ***Genes & development*** **8: 2241-2255**

Peng WT, Krogan NJ, Richards DP, Greenblatt JF, Hughes TR (2004) ESF1 is required for 18S rRNA synthesis in Saccharomyces cerevisiae. ***Nucleic acids research*** **32: 1993-1999**

Perez-Fernandez J, Martin-Marcos P, Dosil M (2011) Elucidation of the assembly events required for the recruitment of Utp20, Imp4 and Bms1 onto nascent pre-ribosomes. ***Nucleic acids research*** **39: 8105-8121**

Perez-Fernandez J, Roman A, De Las Rivas J, Bustelo XR, Dosil M (2007) The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. ***Molecular and cellular biology*** **27: 5414-5429**

Pertschy B, Schneider C, Gnadig M, Schafer T, Tollervey D, Hurt E (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. ***The Journal of biological chemistry*** **284: 35079-35091**

Pestov DG, Stockelman MG, Strezoska Z, Lau LF (2001a) ERB1, the yeast homolog of mammalian Bop1, is an essential gene required for maturation of the 25S and 5.8S ribosomal RNAs. ***Nucleic acids research*** **29: 3621-3630**

Pestov DG, Strezoska Z, Lau LF (2001b) Evidence of p53-dependent cross-talk between ribosome biogenesis and the cell cycle: effects of nucleolar protein Bop1 on G(1)/S transition. ***Molecular and cellular biology*** **21: 4246-4255**

Petfalski E, Dandekar T, Henry Y, Tollervey D (1998) Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. ***Molecular and cellular biology*** **18: 1181-1189**

Phillips S, Butler JS (2003) Contribution of domain structure to the RNA 3' end processing and degradation functions of the nuclear exosome subunit Rrp6p. ***RNA (New York, NY*** **9: 1098-1107**

Phipps KR, Charette JM, Baserga SJ (2011) The SSU Processome in Ribosome Biogenesis - Progress and Prospects. ***WIREs RNA*** **2: 1-21**

Piekna-Przybylska D, Przybylski P, Baudin-Baillieu A, Rousset JP, Fournier MJ (2008) Ribosome performance is enhanced by a rich cluster of pseudouridines in the A-site finger region of the large subunit. ***The Journal of biological chemistry*** **283: 26026-26036**

Pluk H, Soffner J, Luhrmann R, van Venrooij WJ (1998) cDNA cloning and characterization of the human U3 small nucleolar ribonucleoprotein complex-associated 55-kilodalton protein. ***Molecular and cellular biology*** **18: 488-498**

Pluk H, van Eenennaam H, Rutjes SA, Pruijn GJ, van Venrooij WJ (1999) RNA-protein interactions in the human RNase MRP ribonucleoprotein complex. *RNA (New York, NY* 5: 512-524

Poole TL, Stevens A (1997) Structural modifications of RNA influence the 5' exoribonucleolytic hydrolysis by XRN1 and HKE1 of Saccharomyces cerevisiae. *Biochem Biophys Res Commun* 235: 799-805

Preker P, Nielsen J, Kammler S, Lykke-Andersen S, Christensen MS, Mapendano CK, Schierup MH, Jensen TH (2008) RNA exosome depletion reveals transcription upstream of active human promoters. *Science* 322: 1851-1854

Prieto JL, McStay B (2005) Nucleolar biogenesis: the first small steps. *Biochemical Society transactions* 33: 1441-1443

Prieto JL, McStay B (2007) Recruitment of factors linking transcription and processing of pre-rRNA to NOR chromatin is UBF-dependent and occurs independent of transcription in human cells. *Genes & development* 21: 2041-2054

Pulicherla N, Pogorzala LA, Xu Z, HC OF, Musayev FN, Scarsdale JN, Sia EA, Culver GM, Rife JP (2009) Structural and functional divergence within the Dim1/KsgA family of rRNA methyltransferases. *Journal of molecular biology* 391: 884-893

Puvion-Dutilleul F, Mazan S, Nicoloso M, Christensen ME, Bachellerie JP (1991) Localization of U3 RNA molecules in nucleoli of HeLa and mouse 3T3 cells by high resolution in situ hybridization. *European journal of cell biology* 56: 178-186

Rabl J, Leibundgut M, Ataide SF, Haag A, Ban N (2010) Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* 331: 730-736

Reichow SL, Hamma T, Ferre-D'Amare AR, Varani G (2007) The structure and function of small nucleolar ribonucleoproteins. *Nucleic acids research* 35: 1452-1464

Rempola B, Karkusiewicz I, Piekarska I, Rytka J (2006) Fcf1p and Fcf2p are novel nucleolar Saccharomyces cerevisiae proteins involved in pre-rRNA processing. *Biochem Biophys Res Commun* 346: 546-554

Ridanpaa M, van Eenennaam H, Pelin K, Chadwick R, Johnson C, Yuan B, vanVenrooij W, Pruijn G, Salmela R, Rockas S, Makitie O, Kaitila I, de la Chapelle A (2001) Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. *Cell* 104: 195-203

Robledo S, Idol RA, Crimmins DL, Ladenson JH, Mason PJ, Bessler M (2008) The role of human ribosomal proteins in the maturation of rRNA and ribosome production. *RNA (New York, NY* 14: 1918-1929

Rohrmoser M, Holzel M, Grimm T, Malamoussi A, Harasim T, Orban M, Pfisterer I, Gruber-Eber A, Kremmer E, Eick D (2007) Interdependence of Pes1, Bop1, and WDR12 controls nucleolar localization and assembly of the PeBoW complex required for maturation of the 60S ribosomal subunit. *Molecular and cellular biology* 27: 3682-3694

Rotenberg MO, Moritz M, Woolford JL, Jr. (1988) Depletion of Saccharomyces cerevisiae ribosomal protein L16 causes a decrease in 60S ribosomal subunits and formation of half-mer polyribosomes. *Genes & development* 2: 160-172

Rougemaille M, Gudipati RK, Olesen JR, Thomsen R, Seraphin B, Libri D, Jensen TH (2007) Dissecting mechanisms of nuclear mRNA surveillance in THO/sub2 complex mutants. *The EMBO journal* 26: 2317-2326

Rouquette J, Choesmel V, Gleizes PE (2005) Nuclear export and cytoplasmic processing of precursors to the 40S ribosomal subunits in mammalian cells. *The EMBO journal* 24: 2862-2872

Ruggero D, Pandolfi PP (2003) Does the ribosome translate cancer? *Nat Rev Cancer* 3: 179-192

Sahasranaman A, Dembowski J, Strahler J, Andrews P, Maddock J, Woolford JL, Jr. (2011) Assembly of Saccharomyces cerevisiae 60S ribosomal subunits: role of factors required for 27S pre-rRNA processing. *The EMBO journal* 30: 4020-4032

Salinas K, Wierzbicki S, Zhou L, Schmitt ME (2005) Characterization and purification of Saccharomyces cerevisiae RNase MRP reveals a new unique protein component. *The Journal of biological chemistry* 280: 11352-11360

San Paolo S, Vanacova S, Schenk L, Scherrer T, Blank D, Keller W, Gerber AP (2009) Distinct roles of non-canonical poly(A) polymerases in RNA metabolism. *PLoS genetics* 5: e1000555

Sasaki M, Kawahara K, Nishio M, Mimori K, Kogo R, Hamada K, Itoh B, Wang J, Komatsu Y, Yang YR, Hikasa H, Horie Y, Yamashita T, Kamijo T, Zhang Y, Zhu Y, Prives C, Nakano T, Mak TW, Sasaki T, Maehama T, Mori M, Suzuki A (2011) Regulation of the MDM2-P53 pathway and tumor growth by PICT1 via nucleolar RPL11. *Nature medicine* 17: 944-951

Savino R, Gerbi SA (1990) In vivo disruption of Xenopus U3 snRNA affects ribosomal RNA processing. *The EMBO journal* 9: 2299-2308

Savkur RS, Olson MO (1998) Preferential cleavage in pre-ribosomal RNA byprotein B23 endoribonuclease. *Nucleic acids research* 26: 4508-4515

Schaeffer D, Tsanova B, Barbas A, Reis FP, Dastidar EG, Sanchez-Rotunno M, Arraiano CM, van Hoof A (2009) The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. ***Nat Struct Mol Biol*** **16: 56-62**

Schafer T, Maco B, Petfalski E, Tollervey D, Bottcher B, Aebi U, Hurt E (2006) Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit. ***Nature*** **441: 651-655**

Schafer T, Strauss D, Petfalski E, Tollervey D, Hurt E (2003) The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. ***The EMBO journal*** **22: 1370-1380**

Schilders G, Raijmakers R, Raats JM, Pruijn GJ (2005) MPP6 is an exosome-associated RNA-binding protein involved in 5.8S rRNA maturation. ***Nucleic acids research*** **33: 6795-6804**

Schilders G, van Dijk E, Pruijn GJ (2007) C1D and hMtr4p associate with the human exosome subunit PM/Scl-100 and are involved in pre-rRNA processing. ***Nucleic acids research*** **35: 2564-2572**

Schillewaert S, Wacheul L, Lhomme F, Lafontaine DL (2011) The evolutionarily conserved protein LAS1 is required for pre-rRNA processing at both ends of ITS2. ***Molecular and cellular biology***

Schmid M, Jensen TH (2008a) The exosome: a multipurpose RNA-decay machine. ***Trends in biochemical sciences***

Schmid M, Jensen TH (2008b) The exosome: a multipurpose RNA-decay machine. ***Trends in biochemical sciences*** **33: 501-510**

Schmitt ME, Clayton DA (1992) Yeast site-specific ribonucleoprotein endoribonuclease MRP contains an RNA component homologous to mammalian RNase MRP RNA and essential for cell viability. ***Genes & development*** **6: 1975-1985**

Schmitt ME, Clayton DA (1993) Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in Saccharomyces cerevisiae. ***Molecular and cellular biology*** **13: 7935-7941**

Schmitt ME, Clayton DA (1994) Characterization of a unique protein component of yeast RNase MRP: an RNA-binding protein with a zinc-cluster domain. ***Genes & development*** **8: 2617-2628**

Schneider C, Anderson JT, Tollervey D (2007) The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. ***Molecular cell*** **27: 324-331**

Schneider C, Leung E, Brown J, Tollervey D (2009) The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic acids research* 37: 1127-1140

Seiser RM, Sundberg AE, Wollam BJ, Zobel-Thropp P, Baldwin K, Spector MD, Lycan DE (2006) Ltv1 is required for efficient nuclear export of the ribosomal small subunit in Saccharomyces cerevisiae. *Genetics* 174: 679-691

Sharma K, Tollervey D (1999) Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing. *Molecular and cellular biology* 19: 6012-6019

Sharma K, Venema J, Tollervey D (1999) The 5' end of the 18S rRNA can be positioned from within the mature rRNA. *RNA (New York, NY* 5: 678-686

Silverstein RA, Gonzalez de Valdivia E, Visa N (2011) The incorporation of 5-fluorouracil into RNA affects the ribonucleolytic activity of the exosome subunit Rrp6. *Molecular cancer research : MCR* 9: 332-340

Smith SB, Kiss DL, Turk E, Tartakoff AM, Andrulis ED (2011) Pronounced and extensive microtubule defects in a Saccharomyces cerevisiae DIS3 mutant. *Yeast* 28: 755-769

Spiegel R, Shalev SA, Adawi A, Sprecher E, Tenenbaum-Rakover Y (2010) ANE syndrome caused by mutated RBM28 gene: a novel etiology of combined pituitary hormone deficiency. *Eur J Endocrinol* 162: 1021-1025

Srivastava L, Lapik YR, Wang M, Pestov DG (2010) Mammalian DEAD box protein Ddx51 acts in 3' end maturation of 28S rRNA by promoting the release of U8 snoRNA. *Molecular and cellular biology* 30: 2947-2956

Staals RH, Bronkhorst AW, Schilders G, Slomovic S, Schuster G, Heck AJ, Raijmakers R, Pruijn GJ (2010) Dis3-like 1: a novel exoribonuclease associated with the human exosome. *The EMBO journal* 29: 2358-2367

Staals RH, Pruijn GJ (2011) The human exosome and disease. *Advances in experimental medicine and biology* 702: 132-142

Stead JA, Costello JL, Livingstone MJ, Mitchell P (2007) The PMC2NT domain of the catalytic exosome subunit Rrp6p provides the interface for binding with its cofactor Rrp47p, a nucleic acid-binding protein. *Nucleic acids research* 35: 5556-5567

Steitz TA, Steitz JA (1993) A general two-metal-ion mechanism for catalytic RNA. *Proceedings of the National Academy of Sciences of the United States of America* 90: 6498-6502

Strezoska Z, Pestov DG, Lau LF (2000) Bop1 is a mouse WD40 repeat nucleolar protein involved in 28S and 5. 8S RRNA processing and 60S ribosome biogenesis. *Molecular and cellular biology* 20: 5516-5528

Strezoska Z, Pestov DG, Lau LF (2002) Functional inactivation of the mouse nucleolar protein Bop1 inhibits multiple steps in pre-rRNA processing and blocks cell cycle progression. *The Journal of biological chemistry* 277: 29617-29625

Stroke IL, Weiner AM (1985) Genes and pseudogenes for rat U3A and U3B small nuclear RNA. *Journal of molecular biology* 184: 183-193

Strunk BS, Loucks CR, Su M, Vashisth H, Cheng S, Schilling J, Brooks CL, 3rd, Karbstein K, Skiniotis G (2011) Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. *Science* 333: 1449-1453

Su JY, Maller JL (1995) Cloning and expression of a Xenopus gene that prevents mitotic catastrophe in fission yeast. *Mol Gen Genet* 246: 387-396

Sun C, Woolford JL, Jr. (1994) The yeast NOP4 gene product is an essential nucleolar protein required for pre-rRNA processing and accumulation of 60S ribosomal subunits. *The EMBO journal* 13: 3127-3135

Sun C, Woolford JL, Jr. (1997) The yeast nucleolar protein Nop4p contains four RNA recognition motifs necessary for ribosome biogenesis. *The Journal of biological chemistry* 272: 25345-25352

Suzuki N, Zara J, Sato T, Ong E, Bakhiet N, Oshima RG, Watson KL, Fukuda MN (1998) A cytoplasmic protein, bystin, interacts with trophinin, tastin, and cytokeratin and may be involved in trophinin-mediated cell adhesion between trophoblast and endometrial epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America* 95: 5027-5032

Suzuki S, Fujiwara T, Kanno M (2007) Nucleolar protein Nop25 is involved in nucleolar architecture. *Biochem Biophys Res Commun* 358: 1114-1119

Sweet T, Khalili K, Sawaya BE, Amini S (2003) Identification of a novel protein from glial cells based on its ability to interact with NF-kappaB subunits. *J Cell Biochem* 90: 884-891

Sweet T, Yen W, Khalili K, Amini S (2008) Evidence for involvement of NFBP in processing of ribosomal RNA. *J Cell Physiol* 214: 381-388

Symmons MF, Jones GH, Luisi BF (2000) A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure* 8: 1215-1226

Szewczak LB, DeGregorio SJ, Strobel SA, Steitz JA (2002) Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP. *Chem Biol* 9: 1095-1107

Tanaka N, Smith P, Shuman S Crystal structure of Rcl1, an essential component of the eukaryal pre-rRNA processosome implicated in 18s rRNA biogenesis. *RNA (New York, NY* 17: 595-602

Tang L, Sahasranaman A, Jakovljevic J, Schleifman E, Woolford JL, Jr. (2008) Interactions among Ytm1, Erb1, and Nop7 required for assembly of the Nop7-subcomplex in yeast preribosomes. *Molecular biology of the cell* 19: 2844-2856

Terns MP, Terns RM (2002) Small nucleolar RNAs: versatile trans-acting molecules of ancient evolutionary origin. *Gene Expr* 10: 17-39

Thiebaut M, Kisseleva-Romanova E, Rougemaille M, Boulay J, Libri D (2006) Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance. *Molecular cell* 23: 853-864

Thomson E, Tollervey D (2010) The final step in 5.8S rRNA processing is cytoplasmic in Saccharomyces cerevisiae. *Molecular and cellular biology* 30: 976-984

Tollervey D (2004) Molecular biology: termination by torpedo. *Nature* 432: 456-457

Tomecki R, Drazkowska K, Dziembowski A (2010a) Mechanisms of RNA degradation by the eukaryotic exosome. *Chembiochem : a European journal of chemical biology* 11: 938-945

Tomecki R, Kristiansen MS, Lykke-Andersen S, Chlebowski A, Larsen KM, Szczesny RJ, Drazkowska K, Pastula A, Andersen JS, Stepien PP, Dziembowski A, Jensen TH (2010b) The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *The EMBO journal* 29: 2342-2357

Tone Y, Toh EA (2002) Nob1p is required for biogenesis of the 26S proteasome and degraded upon its maturation in Saccharomyces cerevisiae. *Genes & development* 16: 3142-3157

Torchet C, Hermann-Le Denmat S (2002) High dosage of the small nucleolar RNA snR10 specifically suppresses defects of a yeast rrp5 mutant. *Mol Genet Genomics* 268: 70-80

Tschochner H, Hurt E (2003) Pre-ribosomes on the road from the nucleolus to the cytoplasm. *Trends in cell biology* 13: 255-263

Turner AJ, Knox AA, Prieto JL, McStay B, Watkins NJ (2009) A novel small-subunit processome assembly intermediate that contains the U3 snoRNP, nucleolin, RRP5, and DBP4. *Molecular and cellular biology* 29: 3007-3017

Turner AJ, Knox AA, Watkins NJ (2012) Nucleolar disruption leads to the spatial separation of key 18S rRNA processing factors. *RNA biology* In press

Tyc K, Steitz JA (1989) U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus. *The EMBO journal* 8: 3113-3119

Tycowski KT, Shu MD, Steitz JA (1994) Requirement for intron-encoded U22 small nucleolar RNA in 18S ribosomal RNA maturation. *Science* 266: 1558-1561

Tycowski KT, Smith CM, Shu MD, Steitz JA (1996) A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in Xenopus. *Proceedings of the National Academy of Sciences of the United States of America* 93: 14480-14485

Valdez BC, Henning D, So RB, Dixon J, Dixon MJ (2004) The Treacher Collins syndrome (TCOF1) gene product is involved in ribosomal DNA gene transcription by interacting with upstream binding factor. *Proceedings of the National Academy of Sciences of the United States of America* 101: 10709-10714

van Hoof A, Lennertz P, Parker R (2000) Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5, RNase MRP and RNase P RNAs in yeast. *The EMBO journal* 19: 1357-1365

van Nues RW, Granneman S, Kudla G, Sloan KE, Chicken M, Tollervey D, Watkins NJ (2011) Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. *The EMBO journal* 30: 2420-2430

Vanacova S, Stefl R (2007) The exosome and RNA quality control in the nucleus. *EMBO reports* 8: 651-657

Vanacova S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, Keller W (2005) A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS biology* 3: e189

Vanrobays E, Gelugne JP, Caizergues-Ferrer M, Lafontaine DL (2004) Dim2p, a KH-domain protein required for small ribosomal subunit synthesis. *RNA (New York, NY* 10: 645-656

Vanrobays E, Gelugne JP, Gleizes PE, Caizergues-Ferrer M (2003) Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in Saccharomyces cerevisiae. *Molecular and cellular biology* 23: 2083-2095

Vanrobays E, Gleizes PE, Bousquet-Antonelli C, Noaillac-Depeyre J, Caizergues-Ferrer M, Gelugne JP (2001) Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rrp10p, an essential non-ribosomal cytoplasmic protein. ***The EMBO journal*** **20: 4204-4213**

Vanrobays E, Leplus A, Osheim YN, Beyer AL, Wacheul L, Lafontaine DL (2008) TOR regulates the subcellular distribution of DIM2, a KH domain protein required for cotranscriptional ribosome assembly and pre-40S ribosome export. ***RNA (New York, NY*** **14: 2061-2073**

Veith T, Martin R, Wurm JP, Weis BL, Duchardt-Ferner E, Safferthal C, Hennig R, Mirus O, Bohnsack MT, Wohnert J, Schleiff E (2011) Structural and functional analysis of the archaeal endonuclease Nob1. ***Nucleic acids research***

Venema J, Tollervey D (1995) Processing of pre-ribosomal RNA in Saccharomyces cerevisiae. ***Yeast*** **11: 1629-1650**

Venema J, Tollervey D (1996) RRP5 is required for formation of both 18S and 5.8S rRNA in yeast. ***The EMBO journal*** **15: 5701-5714**

Venema J, Vos HR, Faber AW, van Venrooij WJ, Raue HA (2000) Yeast Rrp9p is an evolutionarily conserved U3 snoRNP protein essential for early pre-rRNA processing cleavages and requires box C for its association. ***RNA (New York, NY*** **6: 1660-1671**

Verheggen C, Lafontaine DL, Samarsky D, Mouaikel J, Blanchard JM, Bordonne R, Bertrand E (2002) Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. ***The EMBO journal*** **21: 2736-2745**

Vincent HA, Deutscher MP (2006) Substrate recognition and catalysis by the exoribonuclease RNase R. ***The Journal of biological chemistry*** **281: 29769-29775**

Vos HR, Bax R, Faber AW, Vos JC, Raue HA (2004) U3 snoRNP and Rrp5p associate independently with Saccharomyces cerevisiae 35S pre-rRNA, but Rrp5p is essential for association of Rok1p. ***Nucleic acids research*** **32: 5827-5833**

Wang H, Xiao W, Zhou Q, Chen Y, Yang S, Sheng J, Yin Y, Fan J, Zhou J (2009) Bystin-like protein is upregulated in hepatocellular carcinoma and required for nucleologenesis in cancer cell proliferation. ***Cell Res*** **19: 1150-1164**

Wang HW, Wang J, Ding F, Callahan K, Bratkowski MA, Butler JS, Nogales E, Ke A (2007) Architecture of the yeast Rrp44 exosome complex suggests routes of RNA recruitment for 3' end processing. ***Proceedings of the National Academy of Sciences of the United States of America*** **104: 16844-16849**

Wang M, Pestov DG (2011) 5'-end surveillance by Xrn2 acts as a shared mechanism for mammalian pre-rRNA maturation and decay. ***Nucleic acids research*** **39: 1811-1822**

Wang X, Jia H, Jankowsky E, Anderson JT (2008) Degradation of hypomodified tRNA(iMet) in vivo involves RNA-dependent ATPase activity of the DExH helicase Mtr4p. *RNA (New York, NY* 14: 107-116

Warner JR (1999) The economics of ribosome biosynthesis in yeast. *Trends in biochemical sciences* 24: 437-440

Watkins NJ, Bohnsack MT (2011) The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley interdisciplinary reviews*

Watkins NJ, Dickmanns A, Luhrmann R (2002) Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Molecular and cellular biology* 22: 8342-8352

Watkins NJ, Lemm I, Ingelfinger D, Schneider C, Hossbach M, Urlaub H, Luhrmann R (2004) Assembly and maturation of the U3 snoRNP in the nucleoplasm in a large dynamic multiprotein complex. *Molecular cell* 16: 789-798

Watkins NJ, Lemm I, Luhrmann R (2007) Involvement of nuclear import and export factors in U8 box C/D snoRNP biogenesis. *Molecular and cellular biology* 27: 7018-7027

Watkins NJ, Segault V, Charpentier B, Nottrott S, Fabrizio P, Bachi A, Wilm M, Rosbash M, Branlant C, Luhrmann R (2000) A common core RNP structure shared between the small nucleoar box C/D RNPs and the spliceosomal U4 snRNP. *Cell* 103: 457-466

Wegierski T, Billy E, Nasr F, Filipowicz W (2001) Bms1p, a G-domain-containing protein, associates with Rcl1p and is required for 18S rRNA biogenesis in yeast. *RNA (New York, NY* 7: 1254-1267

Weinstein LB, Steitz JA (1999) Guided tours: from precursor snoRNA to functional snoRNP. *Current opinion in cell biology* 11: 378-384

Weir JR, Bonneau F, Hentschel J, Conti E (2010) Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance. *Proceedings of the National Academy of Sciences of the United States of America* 107: 12139-12144

Welting TJ, Kikkert BJ, van Venrooij WJ, Pruijn GJ (2006) Differential association of protein subunits with the human RNase MRP and RNase P complexes. *RNA (New York, NY* 12: 1373-1382

Welting TJ, Peters FM, Hensen SM, van Doorn NL, Kikkert BJ, Raats JM, van Venrooij WJ, Pruijn GJ (2007) Heterodimerization regulates RNase MRP/RNase P association, localization, and expression of Rpp20 and Rpp25. ***RNA (New York, NY*** **13: 65-75**

Welting TJ, van Venrooij WJ, Pruijn GJ (2004) Mutual interactions between subunits of the human RNase MRP ribonucleoprotein complex. ***Nucleic acids research*** **32: 2138-2146**

West S, Gromak N, Proudfoot NJ (2004) Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. ***Nature*** **432: 522-525**

Widmann B, Wandrey F, Badertscher L, Wyler E, Pfannstiel J, Zemp I, Kutay U (2011) The kinase activity of human Rio1 is required for final steps of cytoplasmic maturation of 40S subunits. ***Molecular biology of the cell***

Wild T, Horvath P, Wyler E, Widmann B, Badertscher L, Zemp I, Kozak K, Csucs G, Lund E, Kutay U (2010) A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. ***PLoS biology*** **8: e1000522**

Wlotzka W, Kudla G, Granneman S, Tollervey D (2011) The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. ***The EMBO journal*** **30: 2982**

Wojda I, Cytrynska M, Frajnt M, Jakubowicz T (2002) Protein kinases CKI and CKII are implicated in modification of ribosomal proteins of the yeast Trichosporon cutaneum. ***Acta Biochim Pol*** **49: 947-957**

Woolls HA, Lamanna AC, Karbstein K (2011) Roles of Dim2 in ribosome assembly. ***The Journal of biological chemistry*** **286: 2578-2586**

Wu K, Wu P, Aris JP (2001) Nucleolar protein Nop12p participates in synthesis of 25S rRNA in Saccharomyces cerevisiae. ***Nucleic acids research*** **29: 2938-2949**

Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, 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: 725-737**

Xiang S, Cooper-Morgan A, Jiao X, Kiledjian M, Manley JL, Tong L (2009) Structure and function of the 5'-->3' exoribonuclease Rat1 and its activating partner Rai1. ***Nature*** **458: 784-788**

Xiao S, Hsieh J, Nugent RL, Coughlin DJ, Fierke CA, Engelke DR (2006) Functional characterization of the conserved amino acids in Pop1p, the largest common protein subunit of yeast RNases P and MRP. ***RNA (New York, NY*** **12: 1023-1037**

Xue S, Wang R, Yang F, Terns RM, Terns MP, Zhang X, Maxwell ES, Li H (2010) Structural basis for substrate placement by an archaeal box C/D ribonucleoprotein particle. *Molecular cell* 39: 939-949

Xue Y, Bai X, Lee I, Kallstrom G, Ho J, Brown J, Stevens A, Johnson AW (2000) Saccharomyces cerevisiae RAI1 (YGL246c) is homologous to human DOM3Z and encodes a protein that binds the nuclear exoribonuclease Rat1p. *Molecular and cellular biology* 20: 4006-4015

Yang XF, Wu CJ, Chen L, Alyea EP, Canning C, Kantoff P, Soiffer RJ, Dranoff G, Ritz J (2002) CML28 is a broadly immunogenic antigen, which is overexpressed in tumor cells. *Cancer research* 62: 5517-5522

Yao W, Roser D, Kohler A, Bradatsch B, Bassler J, Hurt E (2007) Nuclear export of ribosomal 60S subunits by the general mRNA export receptor Mex67-Mtr2. *Molecular cell* 26: 51-62

Yao Y, Demoinet E, Saveanu C, Lenormand P, Jacquier A, Fromont-Racine M (2010) Ecm1 is a new pre-ribosomal factor involved in pre-60S particle export. *RNA (New York, NY* 16: 1007-1017

Ye K, Jia R, Lin J, Ju M, Peng J, Xu A, Zhang L (2009) Structural organization of box C/D RNA-guided RNA methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* 106: 13808-13813

Young CL, Karbstein K (2011) The roles of S1 RNA-binding domains in Rrp5's interactions with pre-rRNA. *RNA (New York, NY* 17: 512-521

Zakrzewska-Placzek M, Souret FF, Sobczyk GJ, Green PJ, Kufel J (2010) Arabidopsis thaliana XRN2 is required for primary cleavage in the pre-ribosomal RNA. *Nucleic acids research* 38: 4487-4502

Zemp I, Kutay U (2007) Nuclear export and cytoplasmic maturation of ribosomal subunits. *FEBS Lett* 581: 2783-2793

Zemp I, Wild T, O'Donohue MF, Wandrey F, Widmann B, Gleizes PE, Kutay U (2009) Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. *The Journal of cell biology* 185: 1167-1180

Zhang J, Harnpicharnchai P, Jakovljevic J, Tang L, Guo Y, Oeffinger M, Rout MP, Hiley SL, Hughes T, Woolford JL, Jr. (2007) Assembly factors Rpf2 and Rrs1 recruit 5S rRNA and ribosomal proteins rpL5 and rpL11 into nascent ribosomes. *Genes & development* 21: 2580-2592

Zhang J, Yang Y, Wu J (2009) B23 interacts with PES1 and is involved in nucleolar localization of PES1. *Acta biochimica et biophysica Sinica* 41: 991-997

Zhu B, Mandal SS, Pham AD, Zheng Y, Erdjument-Bromage H, Batra SK, Tempst P, Reinberg D (2005) The human PAF complex coordinates transcription with events downstream of RNA synthesis. ***Genes & development*** **19: 1668-1673**

Publications and Presentations

Publications

Sloan, K., Mattijssen, S., Tollervery, D., Pruijn, G., Watkins, N. (2012) Early Human Pre-rRNA Processing in ITS1 uses Different Processing Pathways and Nucleases from the Established Yeast Model (Submitted)

Sloan, K., Schneider, C., Watkins, N., (2012) Comparison of the yeast and human nuclear exosome complexes. Biochem Soc Trans (submitted)

van Nues, R., Granneman, S., Kudla, G., Sloan, K., Chicken, M., Tollervey, D., Watkins, N., (2011) Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. EMBO J, 30, 2420-2430.

Small, E., Marrington, R., Roger, A., Scott, DJ., Sloan, K., Roper, D., Daffron, TR., Addinall, SG., (2007) FtsZ polymer-bundling by the Escherichia coli ZapA orthologue, YfgE, involves a conformational change in bound GTP. J. Mol Biol, 369, 210-221

Addinall, SG., Johnson, KA., Daffron, T., Smith, C., Roger, A., Gomex, RP., Sloan, K., Blewett, A., Scott, DJ., Roper, DI., (2005) Expression, purification and crystallization of the cell-division protein YgfE from Escherichia coli. Acta Crystallogr, 61, 305-307

Presentations

Oral presentation

Sloan, K., Mattijssen, S., Tollervery, D., Pruijn, G., Watkins, N (2012) Key differences between yeast and mammalian rRNA processing revealed from the analysis of ITS1 processing in human cells. RNA UK Conference, Lake District, UK, January 20th-22nd.

Poster presentations

Sloan, K., Watkins, N., (2011) Unraveling the major and minor pathways of human 18S rRNA 3' end processing. 16th Annual Meeting of the RNA Society, Kyoto, Japan, June 14th-18th.

Knox, A., Sloan, K., Colvin D., Watkins, N., (2011) Regulating NOB1 multimerisation as a mechanism to control the timing of the 3' processing of 18S rRNA. 16th Annual Meeting of the RNA Society, Kyoto, Japan, June 14th-18th.

Sloan, K., Watkins, N., (2010) The role of the human exosome in pre-rRNA processing and degradation. 15th Annual Meeting of the RNA Society, Seattle, Washington, USA, June 22nd-26th.

Knox, A., Sloan, K., Watkins, N., (2010) The 3' maturation of human 18S rRNA. 15th Annual Meeting of the RNA Society, Seattle, Washington, USA, June 22nd-26th.