The Roles of Human Ribosome Biogenesis Factors in Pre-rRNA Cleavage and p53 Regulation

Graeme Wells

Submitted for Doctor of Philosophy Degree

Final Submission: August 2018

Institute for Cell and Molecular Biosciences

Faculty of Medical Sciences

Newcastle University

Abstract

Ribosome biogenesis is a complex process that is upregulated in many cancers and downregulated during differentiation. Mutations in genes encoding ribosome biogenesis factors are associated with genetic diseases called ribosomopathies. Three of the four ribosomal RNAs (rRNAs), including the 18S rRNA of the small subunit (SSU), released from a single precursor by endonucleolytic cleavages and exonucleolytic processing. Despite extensive research into yeast ribosome biogenesis, the enzymes catalysing several cleavages in 18S rRNA maturation remain unknown, and even less is known in humans. Defective human ribosome biogenesis leads to the activation of the tumour suppressor, p53 via the 5S RNP, a large subunit (LSU) assembly intermediate. Unexpectedly, defective SSU biogenesis also results in p53 induction and multiple proteins implicated in 18S rRNA processing are mutated in cancers. This project aimed to identify endonucleases functioning in human 18S rRNA maturation and investigate the link between SSU biogenesis and p53 regulation.

The data presented here provides evidence that the PIN domain protein UTP24 is the endonuclease responsible for pre-rRNA cleavage at two sites in human 18S rRNA maturation and argues against a direct role for the RNA cyclase-like protein Rcl1/RCL1 in pre-rRNA cleavage in yeast and humans. Additionally, this work shows that an intact UTP23 PIN domain is essential for human pre-rRNA processing, suggesting it may play an enzymatic role in pre-rRNA cleavage. Depletion of UTP23 leads to p53 induction, while a UTP23 mutation associated with colorectal cancer affects pre-rRNA processing *in vivo* but does not result in activation of p53. Multiple ribosome biogenesis factors are frequently mutated in cancers, including the RNA-binding protein RRP5 which plays key roles in 18S rRNA maturation. Depletion of RRP5 results in 5S RNP-dependent p53 activation. Human cell lines stably expressing mutant forms of RRP5, based on mutants used in yeast studies, were generated which will allow for the characterisation of human RRP5.

Declaration and Acknowledgements

I certify that the work presented in this thesis is my own work, except where acknowledged. No part of this work has previously been submitted for any qualification at this or any university.

I would like to thank my supervisor, Dr Claudia Schneider, for her ongoing support, patience and encouragement throughout the course of this project and throughout the thesis-writing process. I would also like to thank my co-supervisor, Dr Nick Watkins for his help in the lab. I am grateful to everyone in both Schneider and Watkins groups, past and present, for their contributions and for making the lab an enjoyable place to work. I would especially like to thank Dr Andria Pelava for her valuable help and the entertainment she provided on a daily basis and Rebecca Cunningham, Justine Lee and Matt Eastham for making our office a nice place to be and for reading this thesis.

I would also like to thank my friends and family for their support over the past few years. In particular, I'd like to thank my Mum for her continuous encouragement and for always being there, and Dr Josh Campbell for distracting me from any stress with his constant nonsense, and for reading this thesis. Finally, I am very grateful to Sinéad Flinders for her support and kindness in the last few months prior to final submission.

Table of Contents

List of Figures

List of Tables

Abbreviations

µg	Microgram
µl	Microlitre
µM	Micromolar
A	Alanine
acp	aminocarboxypropyl
ATP	Adenine triphosphate
BSA	Bovine Serum Albumin
C	Cysteine
cDNA	Complementary DNA
CMV	Cytomegalovirus
CRAC	UV Crosslinking and Analysis of cDNA
CTD	C-terminal domain
C-terminal	Carboxy-terminal
D	Aspartic acid
DAPI	4’,6-Diamidino-2-Phenylindole
DBA	Diamond-Blackfan Anaemia
Del	Deletion
DFC	Dense Fibrillar Component
DMEM	Dulbecco’s Modified Eagle Medium
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleoside triphosphate
DTT	Dithiothreitol
E	Glutamic Acid
E. coli	*Escherichia coli*
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
ES6	Eukaryote-specific expansion sequence 6
EtBr	Ethidium Bromide
ETS	External Transcribed Spacer
FC	Fibrillar Centre
FCS	Fetal Calf Serum
FRT	Flp Recombinase Target
GC	Granular Component

GST	Glutathione S transferase
GTP	Guanosine 5’-triphosphate
H	Histidine
H. sapiens	*Homo sapiens*
HCl	Hydrochloric Acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His-tag	Histidine-tag
HRP	Horseradish Peroxidase
IF	Immunofluorescence
IP	Immunoprecipitation
ITS	Internal Transcribed Spacer
K	Lysine
kb	Kilobase
KCl	Potassium Chloride
kDa	Kilodalton
LB	Lysogeny Broth
LSU	Large Ribosomal Subunit
$MgCl_2$	Magnesium Chloride
ml	Millilitre
mM	Millimolar
mRNA	Messenger RNA
N	Asparagine
nM	Nanomolar
NaCl	Sodium Chloride
NaOAc	Sodium Acetate
NaOH	Sodium Hydroxide
$NaPO_4$	Sodium Phosphate
NMD	Nonsense-Mediated Decay
nt	Nucleotide
NTD	N-terminal domain
N-terminal	Amino-terminal
NTP	Nucleoside Triphosphate
ORF	Open Reading Frame
P	Proline
PAGE	Polyacrylamide Gel Electrophoresis

PARN	Poly(A)-specific Ribonuclease
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
PIN	PilT N terminus
PIPES	Piperazine-1,4′-bis(2-ethanesulfonic acid)
Poly(A)	Polyadenylation
Pol	Polymerase
PNK	Polynucleotide Kinase
Pre-rRNA	Precursor ribosomal RNA
Q	Glutamine
R	Arginine
rDNA	Ribosomal DNA
RNA	Ribonucleic Acid
RNAi	RNA Interference
RNP	Ribonucleoprotein particle
RP	Ribosomal Protein
rRNA	Ribosomal RNA
RRP	Ribosomal RNA-processing protein
RT	Reverse Transcription
S	Svedberg unit
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium Dodecyl Sulphate
siRNA	Small interfering RNA
snoRNA	Small nucleolar RNA
snoRNP	Small nucleolar Ribonucleoprotein
SRP	Signal Recognition Particle
SSC	Saline Sodium Citrate
SSU	Small Ribosomal Subunit
TBE	Tris-Boric Acid-EDTA
TBS	Tris—Buffered Saline
THP	Tris(hydroxypropyl)phosphine
Tm	Melting Temperature
TPR	Tetratricopeptide Repeat
TRAMP complex	Trf4/Air2/Mtr Polyadenylation complex
tRNA	Transfer RNA

Tris	Tris(hydroxymethyl)aminomethane
t-UTP	Transcription-associated Utp
UBF	Upstream-binding Factor
UTP	U Three Protein
v/v	Volume per volume
w/v	Weight per volume
WT	Wild Type
X. laevis	*Xenopus laevis*
Ψ	Pseudouridylation

For clarity, yeast protein names are written in this thesis with the first letter capitalised (e.g. Utp24), with mammalian protein names written with all letters in upper case (e.g. UTP24). When referring to a protein conserved between yeast and mammals, the yeast format has been used (e.g. Utp24).

Chapter One
Introduction

1.1 The Eukaryotic Ribosome

Ribosomes are large molecular machines responsible for protein synthesis via the translation of genetic code of messenger RNA (mRNA). The ribosome is a ribonucleoprotein (RNP) complex composed of ribosomal proteins (RPs) and ribosomal RNAs (rRNAs). A single eukaryotic ribosome has a sedimentation coefficient of 80S and consists of two subunits. The small, 40S subunit (SSU) is the site of transfer (t)RNA binding sites and the decoding function of the ribosome, while the large, 60S subunit (LSU) contains the peptidyltransferase centre (PTC) where peptide bond formation is catalysed (Lafontaine and Tollervey, 2001; Gamalinda and Woolford, 2015). Crystal and cryo-electron microscopy (Cryo EM) structures of eukaryotic ribosomes have enhanced our understanding of ribosome organisation, with RNA molecules at the core of each subunit (Figure 1.1) (Armache *et al.*, 2010a; Armache *et al.*, 2010b; Ben-Shem *et al.*, 2011; Klinge *et al.*, 2011; Rabl *et al.*, 2011; Jenner *et al.*, 2012; Klinge *et al.*, 2012; Anger *et al.*, 2013).

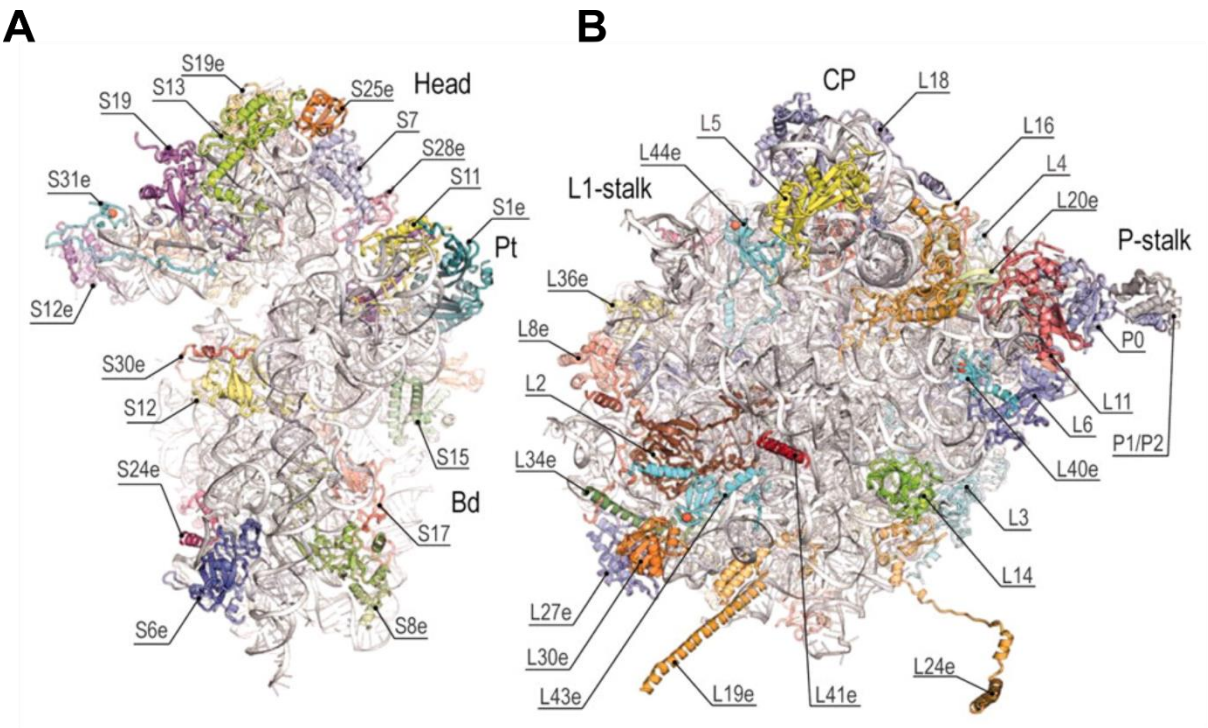


Figure 1.1 Structure of the *Saccharomyces cerevisiae* 40S and 60S ribosomal subunits. (A) Crystal structure of the S. cerevisiae small ribosomal subunit with SSU ribosomal proteins labelled. (B) Crystal structure of the S. cerevisiae large ribosomal subunit with LSU ribosomal proteins labelled. From (Ben-Shem et al., 2011).

Comparison of eukaryotic ribosome structures with structures of prokaryotic and archaeal ribosomes points to a universally conserved core containing the sites of the ribosome's function in protein synthesis (Melnikov *et al.*, 2012). rRNAs, the most abundant noncoding RNAs (ncRNAs) in the cell, are found in both ribosomal subunits, and make up up to 70 % of the ribosome by mass. The SSU contains the 18S rRNA and ~33 RPs, while the LSU contains 28S (25S in yeast), 5.8S and 5S rRNAs and ~46 RPs. RPs are believed to be important for the maintenance of rRNA structure (Lafontaine and Tollervey, 2001)**,** and the number of RPs present in mature ribosomes varies between species (Mager *et al.*, 1997; Planta and Mager, 1998). Three of the four rRNAs, 18S, 5.8S and 28S/25S, are transcribed as a single polycistronic precursor (pre-rRNA; 47S in mammals, 35S in yeast) in the nucleolus by RNA polymerase I (Nazar, 2004). The sequences corresponding to the mature rRNAs are flanked and separated on the primary transcript by external transcribed spacers (5' ETS and 3' ETS) and internal transcribed spacers (ITS1 and ITS2). The 5S rRNA is transcribed separately in the nucleoplasm by RNA polymerase III (Granneman and Baserga, 2004; Orsolic *et al.*, 2016).

1.1.1 Ribosome Production

The synthesis of ribosomes is an extremely complex process and is the major consumer of energy in a cell (Warner, 1999). Production of a single ribosome requires the transcription, processing, folding and modification of ~7,000 nt of pre-rRNA in *Saccharomyces cerevisiae* (budding yeast) and ~13,000 nt in mammals, as well as their assembly with RPs into mature subunits. More than 200 non-ribosomal protein factors assemble onto pre-ribosomal particles containing rRNA precursors and RPs and function in the production of mature ribosome subunits (Henras *et al.*, 2008). In each rapidly growing yeast cell more than 2,000 ribosomes must be assembled per minute, while around 7,500 subunits are produced per minute in a human HeLa cell (Lewis and Tollervey, 2000). As the function of ribosomes dictates the ability of a cell to produce protein, the production of ribosomes is tightly linked to the cell cycle and cellular growth rate (Dez and Tollervey, 2004; Jorgensen *et al.*, 2004; Lempiainen and Shore, 2009). Consistent with this, ribosome production is upregulated in cancer cells, allowing for increased rates of proliferation (Ruggero and Pandolfi, 2003). Defective ribosome biogenesis is associated with multiple diseases, known as ribosomopathies (Freed *et al.*, 2010; Narla and Ebert, 2010).

1.2 Eukaryotic rDNA Transcription

In yeast, around 150 rDNA genes, containing 18S, 5.8S and 25S rRNAs (Figure 1.2A), are found in tandem repeats on chromosome 12 (Kobayashi *et al.*, 1998). These rDNA repeats also contain the 5S rRNA, which is transcribed in the opposite direction. In humans, around 400 tandem rDNA gene repeats, containing the 18S, 5.8S and 28S rRNAs (Figure 1.2B), are found on the short arms of acrocentric chromosomes (chromosome 13, 14, 15, 21 and 22) (Babu and Verma, 1985; Worton *et al.*, 1988; Prieto and McStay, 2005). The human 5S rRNA genes are located separately, on chromosome 1 (Henderson *et al.*, 1972; Steffensen *et al.*, 1974; Sorensen and Frederiksen, 1991). rDNA repeats cluster to form nucleolar organiser regions (NORs), where transcription by RNA polymerase I occurs.

1.2.1 rDNA Transcription

Nucleolar transcription of pre-rRNA transcripts containing the 18S, 5.8S and 25S rRNA sequences by RNA polymerase I represents around 60 % of total yeast transcription, with an elongation rate of 40-60 nt/sec (French *et al.*, 2003; Kos and

Tollervey, 2010). Recruitment of RNA polymerase I to the promoter of an rDNA gene results in the formation of the pre-initiation complex, which requires the recruitment of upstream-binding factor (UBF) and the promoter selectivity factor SL1 complex to the promoter region (Raska *et al.*, 2004; Russell and Zomerdijk, 2005). UBF and SL1 remain at the promoter region as RNA polymerase I begins transcription, allowing for the recruitment of further RNA polymerase I units for successive rounds of transcription (Panov *et al.*, 2001). The 5S rRNA is transcribed in the nucleoplasm by RNA polymerase III (Nazar, 2004).

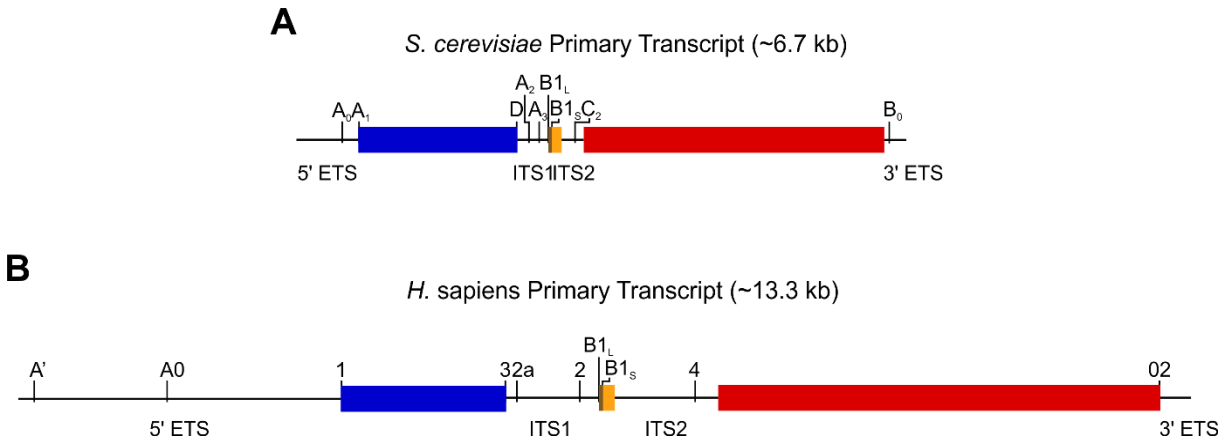


Figure 1.2 Primary pre-rRNA transcripts in S. cerevisiae and H. sapiens. Schematic representation of the primary pre-rRNA transcript in **(A)** yeast and **(B)** humans with names and positions of transcribed spacer regions and endonuclease cleavage sites labelled.

1.2.2 The Nucleolus

The nucleolus is a sub-compartment of the nucleus and is the main site of rRNA transcription, as well as early pre-rRNA modification and processing (Boisvert *et al.*, 2007). While in yeast the crescent-shaped nucleolus contains just two compartments, human nucleoli contain three distinct compartments which appear to reflect the sequential maturation of pre-ribosomes (Thiry and Lafontaine, 2005; Smirnov *et al.*, 2016). The fibrillar centre (FC) contains DNA, including transcription-competent rDNA, as well as RNA polymerase I and transcription factors. The dense fibrillar component (DFC) contains nascent pre-rRNAs and early pre-rRNA processing factors essential for early processing events. The granular component (GC) contains pre-ribosomes. Transcription appears to occur at the boundary between the FC and the DFC, with nascent transcripts extending into the DFC (Cheutin *et al.*, 2002; Huang, 2002;

Koberna *et al.*, 2002; Engel *et al.*, 2013). The first pre-rRNA processing events, endonuclease cleavages within the 5' ETS, at the 5' end of the primary transcript, upstream of the 18S rRNA sequence, occur in the DFC (Derenzini *et al.*, 1990; Granneman *et al.*, 2004). Pre-ribosomes migrate to the GC during ribosome biogenesis, with later processing steps, such as ITS1 processing, downstream of the 18S rRNA sequence, occurring in the GC (Gerbi and Borovjagin, 1997; Lazdins *et al.*, 1997). Pre-ribosomes continue to the nucleoplasm before nuclear export and final maturation in the cytoplasm (Lei and Silver, 2002). The localisation of ribosome biogenesis factors throughout the nucleolus is consistent with their temporal function in pre-rRNA processing. For example, factors required for early cleavages around 18S (in the 5' ETS and ITS1) are generally found in both the DFC and GC (Leary *et al.*, 2004; Prieto and McStay, 2007; Turner *et al.*, 2009).

In addition to ribosome biogenesis, the nucleolus is also important for other functions including the biogenesis of other non-ribosomal RNPs (Boisvert *et al.*, 2007). Indeed, many non-ribosomal proteins are found in the nucleolus, and only around 30% of the nucleolar proteome is involved in ribosome production (Ahmad *et al.*, 2009). One non-ribosomal function of the nucleolus is the biogenesis of the signal recognition particle (SRP), which is an RNP involved in co-translational translocation of specific proteins to the endoplasmic reticulum (ER) membrane (Akopian *et al.*, 2013). The SRP recognises and binds a signal sequence in a nascent peptide and targets the ribosome-peptide complex to the SRP receptor on the ER membrane. Once docked, the SRP is released, the nascent peptide enters the ER membrane and protein synthesis continues, releasing the peptide into the ER. The SRP RNA is transcribed by RNA polymerase III (Dieci *et al.*, 2007) and assembles with the core SRP proteins in the nucleolus to form the pre-SRP (Jacobson and Pederson, 1998; Ciufo and Brown, 2000; Politz *et al.*, 2000; Grosshans *et al.*, 2001; Politz *et al.*, 2002). The pre-SRP is then exported to the cytoplasm and correct 3' processing of the SRP RNA is required for export (Ciufo and Brown, 2000; Grosshans *et al.*, 2001). Maturation of transfer RNA (tRNA) precursors also initialises in the nucleolus (Bertrand *et al.*, 1998; Strub *et al.*, 2007) and small nucleolar RNPs (snoRNPs) also function in alternative pre-mRNA splicing (Damianov *et al.*, 2006; Falaleeva *et al.*, 2016).

1.3 snoRNPs & rRNA Modifications

During ribosome biogenesis, the pre-rRNA undergoes covalent modification in the nucleolus, mainly in mature rRNA sequences that are functionally important within the mature ribosome (Decatur and Fournier, 2002; Sharma and Lafontaine, 2015). Some rRNA modifications are mediated by standalone protein enzymes (Sloan *et al.*, 2017), such as the last step in the modification of the hypermodified uridine, U1191, namely the addition of a aminocarboxypropyl (acp) group, by Tsr3 (Meyer *et al.*, 2016). Many covalent modifications are mediated by snoRNP complexes, which contain a small nucleolar (sno)RNA that guides modification by base pairing to the target sequence on the pre-rRNA, and four core proteins, including the enzyme responsible for catalysing the RNA modification (Kiss-Laszlo *et al.*, 1998; Terns and Terns, 2002; Henras *et al.*, 2004b; Watkins and Bohnsack, 2012). A single snoRNP can guide 2 RNA modification events at different sites, and covalent modifications can occur both post-transcriptionally and co-transcriptionally (Terns and Terns, 2002; Osheim *et al.*, 2004; Kos and Tollervey, 2010). The snoRNPs responsible for rRNA modifications can be separated into two different classes, based on the type of modification they mediate. Box C/D snoRNPs catalyse 2'-*O*-methylation, while box H/ACA snoRNPs catalyse pseudouridylation. A further subset of snoRNPs function in the processing of the pre-rRNA during rRNA maturation (Henras *et al.*, 2008).

1.3.1 Box C/D snoRNPs

Box C/D snoRNPs mediate 2'-*O*-methylation of specific rRNA nucleotides, which stabilises base pairing interactions and can strengthen or alter RNA folding (Helm, 2006). Box C/D snoRNPs consist of a box C/D snoRNA and the four core box C/D proteins, Snu13 (15.5K in humans), Nop56, Nop58 and Nop1 (Fibrillarin in humans) (Figure 1.3) (Watkins *et al.*, 2000). Nop1/Fibrillarin is a methyltransferase that catalyses the transfer of a methyl group to the 2' hydroxyl of the modified rRNA nucleotide (Galardi *et al.*, 2002; Singh *et al.*, 2008). Box C/D snoRNAs contain conserved box C and D motifs, at the 5' and 3' ends respectively, as well as internal box C'/D' motifs. The 5' and 3' ends of the snoRNA base pair, bringing box C and box D close together. Sequences upstream of the D/D' box motifs base pair with the pre-rRNA, mediating methylation by Nop1/Fibrillarin (Kiss-Laszlo *et al.*, 1998; Galardi *et al.*, 2002). Additional base pairing of box C/D snoRNAs adjacent to modification sites

on the rRNA have been described for yeast and humans, which possibly functions to stabilise snoRNA-rRNA interactions and is important for methylation (van Nues *et al.*, 2011). Single copies of both Nop56 and Nop58 associate with RNA guide regions within box C'/D' and box C/D motifs respectively, forming a heterodimer (Lin *et al.*, 2011; Kornprobst *et al.*, 2016). Snu13/15.5K binds to C/D and C'/D' motifs and likely results in conformational changes that permit binding of other core proteins to the snoRNA (Watkins *et al.*, 2002; Szewczak *et al.*, 2005; Qu *et al.*, 2011). Data in yeast indicate that some snoRNAs contain C'/D' motifs differing from the consensus, resulting in formation of different snoRNP structures that allow a single snoRNP to methylate two nucleotides within a target region (van Nues and Watkins, 2017). Rather than catalysing methylation, several box C/D snoRNPs are involved in pre-rRNA processing, including U3, U14 in yeast and higher eukaryotes, and the higher eukaryote-specific U8 and U22 (Henras *et al.*, 2008).

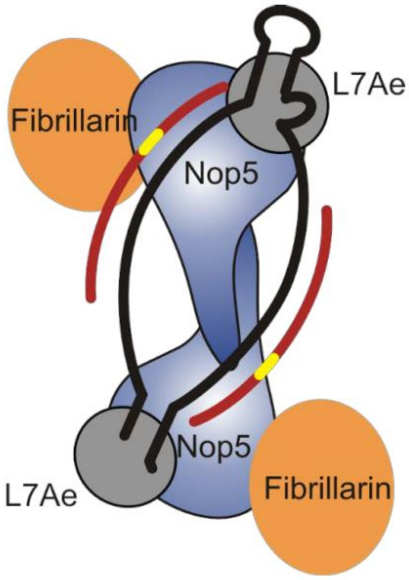


Figure 1.3 Schematic Representation of the archaeal box C/D snoRNP architecture. The box C/D snoRNA is shown in black and the RNA substrate is shown in red, with the modification target nucleotide shown in yellow. Positions of snoRNP proteins Fibrillarin (Nop1 in yeast), Nop5 (Nop56 and Nop58 in eukaryotes) and L7Ae (Snu13 in yeast, 15.5K in humans) are shown. (Watkins and Bohnsack, 2012).

1.3.2 Box H/ACA snoRNPs

Box H/ACA snoRNPs catalyse the conversion of uridine residues to pseudouridine, which is important for stabilising secondary RNA structures (Helm, 2006). Box H/ACA snoRNPs contain the snoRNA component and the four box H/ACA core proteins, Nhp2, Gar1, Nop10 and Cbf5 (Dyskerin in humans) (Figure 1.4) (Henras *et al.*, 1998; Watkins *et al.*, 1998). Cbf5/Dyskerin is a pseudouridine synthase, the catalytic component responsible for the pseudouridylation of the rRNA residue

(Lafontaine *et al.*, 1998). Eukaryotic H/ACA snoRNAs form double stem loop structures, with one copy of each core protein assembled onto each stem (Ganot *et al.*, 1997b; Watkins *et al.*, 1998). Sequences within one or both stem loops base pair with the rRNA, exposing the target uridine residue for pseudouridylation by Cbf5/Dyskerin (Ni *et al.*, 1997; Reichow *et al.*, 2007). In humans, the pseudouridine synthase, Dyskerin, is mutated in the ribosomopathy X-linked dyskeratosis congenita (Narla and Ebert, 2010; Bellodi *et al.*, 2013). As for box C/D snoRNPs, multiple box H/ACA snoRNPs function in pre-rRNA processing instead of pseudouridylation, including snR30/U17 and snR10 (Henras *et al.*, 2008).

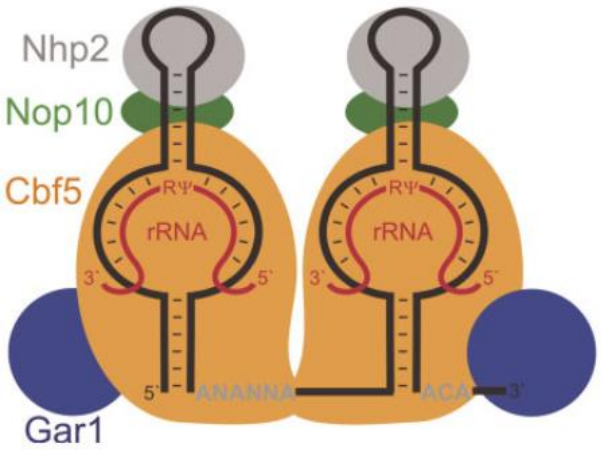


Figure 1.4 Schematic representation of the eukaryotic box H/ACA snoRNP architecture. The box H/ACA snoRNA is shown in black and the RNA substrate is shown in red, with the pseudouridylated nucleotide marked with RΨ. Positions of box H/ACA proteins Cbf5 (Dyskerin in humans), Nop10, Nhp2 and Gar1 are shown. (Watkins and Bohnsack, 2012).

1.4 Pre-rRNA Processing

Ribosomal RNAs are synthesised as precursor molecules called pre-rRNAs. Pre-rRNAs must be processed to release mature rRNAs for assembly into ribosomal subunits. The 18S rRNA of the SSU is transcribed on a polycistronic precursor which also contains the 5.8S and 28S (25S in yeast) rRNAs of the LSU. The mature rRNA sequence on this primary precursor (47S pre-rRNA in mammals, 35S in yeast) are flanked by external transcribed spacer (ETS) sequences (5' ETS and 3' ETS) and separated by internal transcribed spacer (ITS) sequences (ITS1 and ITS2). These spacer regions are removed by a series of endonucleolytic cleavages followed by exonucleolytic processing of cleaved substrates (Venema and Tollervey, 1995). As

with ribosome biogenesis in general, pre-rRNA processing has been most extensively studied in the yeast *S. cerevisiae* (Granneman and Baserga, 2004; Henras *et al.*, 2008; Woolford and Baserga, 2013; Henras *et al.*, 2015; Tomecki *et al.*, 2017). It was originally thought that processing of the yeast pre-rRNA, as well as its covalent modification, always begins with endonucleolytic processing in the 3' ETS of the transcribing pre-rRNA, at a cleavage site called B_0, releasing the 35S pre-rRNA (Venema and Tollervey, 1999). It was later shown, by analysis of nascent pre-rRNA chromatin spreads, that the B_0 cleavage event is not the only co-transcriptional cleavage event, with between 40 % and 80 % of transcripts cleaved within ITS1 during transcription (Osheim *et al.*, 2004). This is consistent with the observed co-transcriptional association of the nascent transcript with RPs, ribosome assembly factors and snoRNPs (Tschochner and Hurt, 2003). More recent biochemical analysis and mathematical modelling estimated that co-transcriptional cleavage at site A_2 occurs in 70-80 % of transcripts (Kos and Tollervey, 2010). Pre-rRNA processing is less well studied in higher eukaryotes but the general processing pathways appear to be largely conserved**.** However, in mammals, the first pre-SSU cleavage, at site A', appears to be the only SSU cleavage to occur co-transcriptionally.

1.4.1 Pre-rRNA processing in Saccharomyces cerevisiae

Pre-rRNA processing in *S. cerevisiae* is well characterised, and the various processing and assembly steps (Figure 1.5) and many of the factors required for each step have been identified (Woolford and Baserga, 2013). The 3' ETS, downstream of the mature 25S rRNA sequence, is cleaved co-transcriptionally at site B_0 by the endonuclease Rnt1 to release the 35S pre-rRNA (Kufel *et al.*, 1999; Delprato *et al.*, 2014) (Henras *et al.*, 2004a). Rnt1 also plays a role in termination of pre-rRNA transcription (Ghazal *et al.*, 2009; Rondon *et al.*, 2009; Nemeth *et al.*, 2013). The ~20 nt long extension 3' of the mature 25S sequence is processed by the exonuclease Rex1 to produce the mature 3' end of 25S (Kempers-Veenstra *et al.*, 1986). The 3' maturation of 25S is coupled to 5.8S 5' maturation, as mutations in the 3' ETS lead to defective cleavage at sites A_3 and B_{1L} (Kufel *et al.*, 1999).

Three early endonuclease cleavages important for maturation of the 18S rRNA of the SSU can occur either post-transcriptionally following B_0 cleavage, or co-transcriptionally on the nascent pre-rRNA (Kos and Tollervey, 2010; Turowski and Tollervey, 2015). The first of these takes place at site A_0 in the 5' ETS, a cleavage

event mediated by an as yet unidentified endonuclease, generating a 33S pre-rRNA intermediate. The B_0 site endonuclease Rnt1 was previously implicated in site A_0 cleavage, but this was later disproven (Elela *et al.*, 1996; Kufel *et al.*, 1999). A second 5' ETS cleavage event, at site A_1 at the 5' end of the mature 18S rRNA sequence, is coupled to A_0 cleavage and these two cleavages occur almost simultaneously. A_0 and A_1 cleavages remove the 5' ETS, generating the mature 5' end of the 18S rRNA and producing a 32S intermediate. The presence of detectable excised 5' ETS spacer fragments spanning from site A_0 to site A_1 indicates that the 5' end of 18S is matured exclusively by endonuclease cleavages, rather than exonucleolytic processing to site A_1 (de la Cruz *et al.*, 1998; Allmang *et al.*, 2000). The PIN domain endonuclease Utp24 was implicated in cleavage at A_1, and recent data argue strongly for a direct role for Utp24 in processing of this site (Bleichert *et al.*, 2006; Tomecki *et al.*, 2015; Wells *et al.*, 2016). The excised fragments of the 5' ETS, generated by A_0 and A_1 cleavages, are degraded by the exosome complex and Rat1 (Petfalski *et al.*, 1998; Allmang *et al.*, 1999; Allmang *et al.*, 2000).

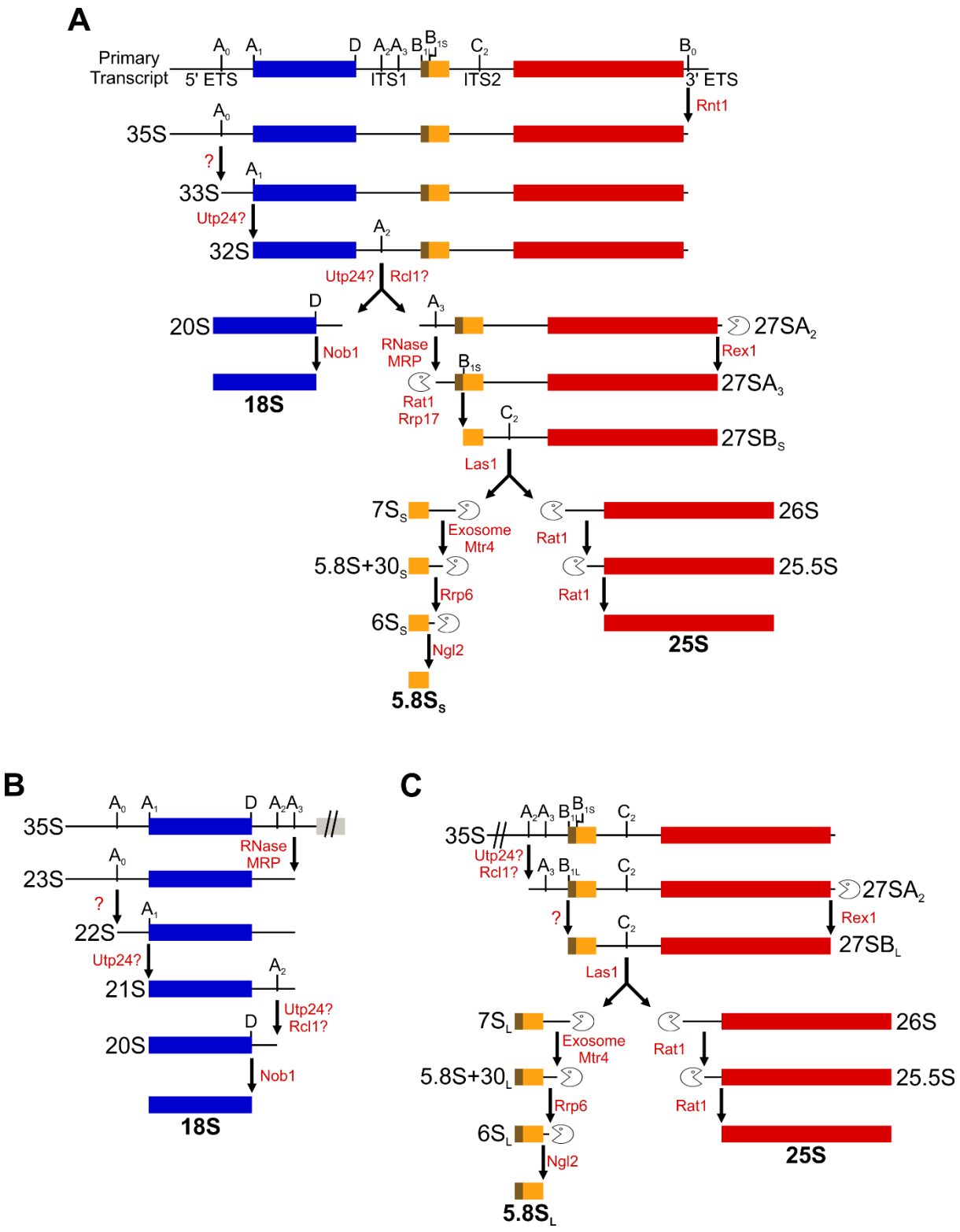


Figure 1.5 Pre-rRNA processing in *S. cerevisiae*. Schematic representation of pre-rRNA processing pathways in yeast. Endonuclease cleavage sites are marked on the pre-rRNA and, where known, the endo- or exo-nuclease catalysing each processing event is shown. **(A)** The major pre-rRNA processing pathway in yeast, involving endonuclease cleavage at sites A_0, A_1 and A_2 prior to A_3 cleavage. This pathway generates the short form of the mature 5.8S rRNA, $5.8S_S$, by exonucleolytic maturation of the 5' end of 5.8S. **(B)** The minor SSU pre-rRNA processing pathway in yeast, involving initial cleavage at site A_3 prior to A_0-A_2 cleavages. **(C)** The minor 5.8S pre-

rRNA processing pathway in yeast, which generates the long form of the mature 5.8S rRNA, $5.8S_L$, by endonucleolytic cleavage at site B_{1L}.

The 18S rRNA precursor is separated from the LSU rRNAs by removal of the ITS1 spacer. This is initiated by endonucleolytic cleavage at site(s) A_2 and/or A_3, generating $27SA_2$ or $27SA_3$ respectively (Kressler *et al.*, 1999). The RNA cyclase-like protein Rcl1 was proposed as the endonuclease responsible for cleaving at the A_2 site (Horn *et al.*, 2011), despite not possessing a known conserved nuclease domain (Tanaka *et al.*, 2011). The role of Rcl1 in A_2 cleavage was recently disputed, and Utp24 was also shown to be responsible for cleavage at this site (Wells *et al.*, 2016). Cleavage at A_2 can occur either co-transcriptionally or post-transcriptionally (Kos and Tollervey, 2010)**,** and co-transcriptional A_2 cleavage requires the exonuclease Rat1 (Axt *et al.*, 2014). The cleavage at A_2 separates the SSU from the LSU rRNA precursors, and the SSU and LSU pre-rRNAs are processed independently following this separation.

SSU and LSU pre-rRNA processing pathways generally require distinct sets of ribosome biogenesis factors, although the RNA binding protein Rrp5 is required for cleavage at both A_2 and A_3, implicating it in both pathways (Venema and Tollervey, 1996). Cleavages at sites A_0, A_1 and A_2 require the U3 snoRNP**,** which base-pairs with 5' ETS and 18S rRNA sequences and is required for correct pre-rRNA folding that permits processing (Hughes and Ares, 1991; Beltrame and Tollervey, 1992; Sharma and Tollervey, 1999; Dutca *et al.*, 2011; Kudla *et al.*, 2011). The U3 snoRNP is a major component of a large RNP complex called the SSU processome, which assembles onto nascent SSU pre-rRNA and is essential for cleavage at sites A_0, A_1 and A_2 (Dragon *et al.*, 2002; Grandi *et al.*, 2002; Bernstein *et al.*, 2004). Around 85 % of pre-rRNA transcripts are cleaved at a second site, A_3, in ITS1 by the RNP RNase MRP (Figure 1.5A) (Schmitt and Clayton, 1993; Chu *et al.*, 1994; Lygerou *et al.*, 1996). A_3 cleavage is followed by exonucleolytic processing of the resulting $27SA_3$ pre-rRNA to the 5' end of the mature 5.8S rRNA (site B_{1S}) by Rat1 and Rrp17, generating the 5' end of the short form of 5.8S ($5.8S_S$) and producing the $27SB_S$ intermediate (Henry *et al.*, 1994; Petfalski *et al.*, 1998; El Hage *et al.*, 2008; Oeffinger *et al.*, 2009). The exonuclease Rat1 is recruited to this region prior to cleavage at the A_3 site (Granneman *et al.*, 2011). The remaining ~15 % of pre-rRNAs not cleaved at A_3 are instead cleaved at an alternative cleavage site, B_{1L}, to mature the 5' end of the long form of 5.8S ($5.8S_L$)

by an unknown endonuclease which produces the $27SB_L$ precursor (Figure 1.5C) (Faber *et al.*, 2006). The 20S pre-rRNA, produced by A_2 cleavage is then exported to the cytoplasm, where it is cleaved at site D by the endonuclease Nob1, forming the mature 3' end of the 18S rRNA (Fatica *et al.*, 2003; Lamanna and Karbstein, 2009; Pertschy *et al.*, 2009). A 23S precursor, cleaved at A_3 but not A_0 (or A_1/A_2) is detectable at low levels in wild type yeast cells, suggesting some variation in the timing of these cleavage events (Dunbar *et al.*, 1997; Gallagher *et al.*, 2004). 22S (cleaved at A_0 and A_3) and 21S (cleaved at A_1 and A_3) precursors have also been shown to accumulate in mutant strains and were thought to be aberrant intermediates and therefore targeted for degradation. However, these "aberrant" precursors were also found to be associated with the pre-rRNA processing machinery in wild type cells, suggesting the possibility of a pathway for their processing into mature 18S rRNA (Figure 1.5B) (Dez *et al.*, 2007; Kuhn *et al.*, 2009). Interestingly, initial ITS1 cleavage at the A_3, rather than A_2, site occurs in response to unfavourable growth conditions, although processing under these conditions does not appear to be productive, and no new ribosomes are produced, suggesting that this is not a *bona fide* pre-rRNA processing pathway (Kos-Braun *et al.*, 2017).

Unlike processing events leading to LSU rRNA maturation in yeast, and in contrast to mammalian ITS1 processing (see below), maturation of yeast 18S rRNA is mediated exclusively by endonuclease cleavages in 5' ETS and ITS1 (Figure 1.5A and 1.5B) (Henras *et al.*, 2015). Cleavage of $27SB_S$ or $27SB_L$ at site C_2 is catalysed by the endonuclease Las1 (Schillewaert *et al.*, 2012; Gasse *et al.*, 2015) and occurs in the nucleus following 5' maturation of 5.8S and 3' maturation of 25S, generating short ($7S_S$) or long ($7S_L$) forms of the 5.8S precursor and the 25S precursor, 26S (Allmang *et al.*, 1999; Michot *et al.*, 1999). The 3' end of 5.8S is matured through a series of exonucleolytic trimming steps. Firstly, the nuclear exosome complex trims to around 30 nt downstream of the 5.8S 3' end, generating a 5.8S+30 intermediate (Mitchell *et al.*, 1996; Allmang *et al.*, 1999). This exosome-mediated processing requires the RNA helicase Mtr4, which likely functions in unwinding of the RNA substrate to permit exosome progression (de la Cruz *et al.*, 1998; Jia *et al.*, 2012). 5.8S+30 is then converted to a 5.8S with an 8 nt 3' extension, called the 6S precursor, by the exosome-associated exonuclease Rrp6 (Briggs *et al.*, 1998). The final 5.8S processing events involve exonucleolytic cleavage by Rex1 and Rex2 (van Hoof *et al.*, 2000) generating 5.8S+5 before final cytoplasmic processing to the 3' end of 5.8S by the exonuclease

Ngl2 (Faber *et al.*, 2002; Thomson and Tollervey, 2010). The 5' end of the 25S rRNA is matured by the exonucleolytic activity of Rat1 (Geerlings *et al.*, 2000; Oeffinger *et al.*, 2009). The exonuclease trimming steps following cleavage at the C_2 site in ITS2 also require the endonuclease Las1 (Schillewaert *et al.*, 2012; Gasse *et al.*, 2015). Many of the exonucleases responsible for processing of cleaved pre-rRNA also function in the degradation of excised spacer fragments. For example, Rat1 is responsible for turnover of excised fragments produced by cleavages at sites A_0 and A_1 as well as sites A_2 and A_3 (Petfalski *et al.*, 1998; Fang *et al.*, 2005).

1.4.2 Mammalian pre-rRNA Processing

Pre-rRNA processing in higher eukaryotes is less well characterised, but experiments indicate a largely conserved pathway between mammals, other vertebrates, plants and the well-studied budding yeast (Hadjiolova *et al.*, 1984a; Hadjiolova *et al.*, 1984c; Hadjiolova *et al.*, 1984b; Lange *et al.*, 2011; Tafforeau *et al.*, 2013; Tomecki *et al.*, 2017). As in *Xenopus* (Borovjagin and Gerbi, 2005) and plants (Delcasso-Tremousaygue *et al.*, 1988), the primary mammalian transcript, the 47S pre-rRNA (Figure 1.6), contains an additional cleavage site upstream of A0 in the 5' ETS, called A' (Bowman *et al.*, 1983; Kass *et al.*, 1987; Hadjiolova *et al.*, 1993; Mullineux and Lafontaine, 2012). Efficient cleavage at the A' site requires a subset of SSU processome components, multiple snoRNPs including U3 and U14, and the exonuclease XRN2 (mammalian homologue of yeast Rat1) (Enright *et al.*, 1996; Sloan *et al.*, 2014; Memet *et al.*, 2017). A' cleavage can be bypassed without affecting production of mature rRNAs, and sites A0 and 1 are cleaved efficiently following impairment of cleavage at A', suggesting that processing of this site could represent a quality control mechanism (Vance *et al.*, 1985; Wang and Pestov, 2011; Sloan *et al.*, 2014). Unlike yeast, where cleavages in both the 5' ETS and in ITS1 can occur co-transcriptionally on the nascent pre-rRNA, A' appears to be the only site cleaved co-transcriptionally in mammals (Lazdins *et al.*, 1997). Cleavage at A' generates a 45S pre-rRNA, and the enzyme responsible for cleaving this site is yet to be identified (Figure 1.6A).

Unlike in yeast, where cleavages in the 5' ETS occur prior to processing of ITS1, in the major human processing pathway, cleavage at site 2 in ITS1 follows A' cleavage, producing the 30S SSU precursor and the 32.5S LSU intermediate (Figure 1.6A). The human orthologue of yeast RNase MRP , which cleaves at yeast site A_3, was initially

thought not to cleave at site 2 (Sloan *et al.*, 2013b), but was later shown to be responsible for this cleavage (Goldfarb and Cech, 2017). The mammalian 5' ETS also contains equivalent cleavage sites to yeast sites A_0 and A_1, called A0 and 1. As in yeast, removal of the 5' ETS appears to be exclusively initiated by endonucleolytic cleavages at sites A0 and 1 (Wang and Pestov, 2011)**,** however recent data suggests a possible alternative processing mechanism when cleavage at 1 is impaired (Tomecki *et al.*, 2015; Wells *et al.*, 2016). Like the equivalent yeast cleavages, cleavages at sites A0 and 1 are coupled and usually occur almost simultaneously. The timing of 5' ETS removal, compared to ITS1 processing, in mammals is variable among different organisms and cell lines, and may also differ in different cell types of a single organism (Bowman *et al.*, 1983; Savino and Gerbi, 1990; Hadjiolova *et al.*, 1993; Borovjagin and Gerbi, 1999). The enzyme responsible for cleavage at site A0 is unknown, while the PIN domain protein UTP24 was proposed as the site 1 endonuclease in human cells (Tomecki *et al.*, 2015; Wells *et al.*, 2016). The excised fragments of the 5' ETS are degraded by the exosome complex and XRN2 (Wang and Pestov, 2011; Sloan *et al.*, 2014). Cleavages at sites A0 and 1 in the 30S intermediate in the major pre-rRNA processing pathway generate the 21S pre-rRNA. Following cleavages at A0, site 1 and site 2, the generated 21S pre-rRNA is processed exonucleolytically from site 2, to site 2a, around 25 nt downstream of the 3' end of the 18S sequence, generating an 18SE precursor (Sloan *et al.*, 2013b). The 18SE pre-rRNA is then exported to the cytoplasm where it is cleaved at site 3 (yeast site D) by the PIN domain endonuclease NOB1, generating the mature 18S rRNA (Preti *et al.*, 2013; Bai *et al.*, 2016).

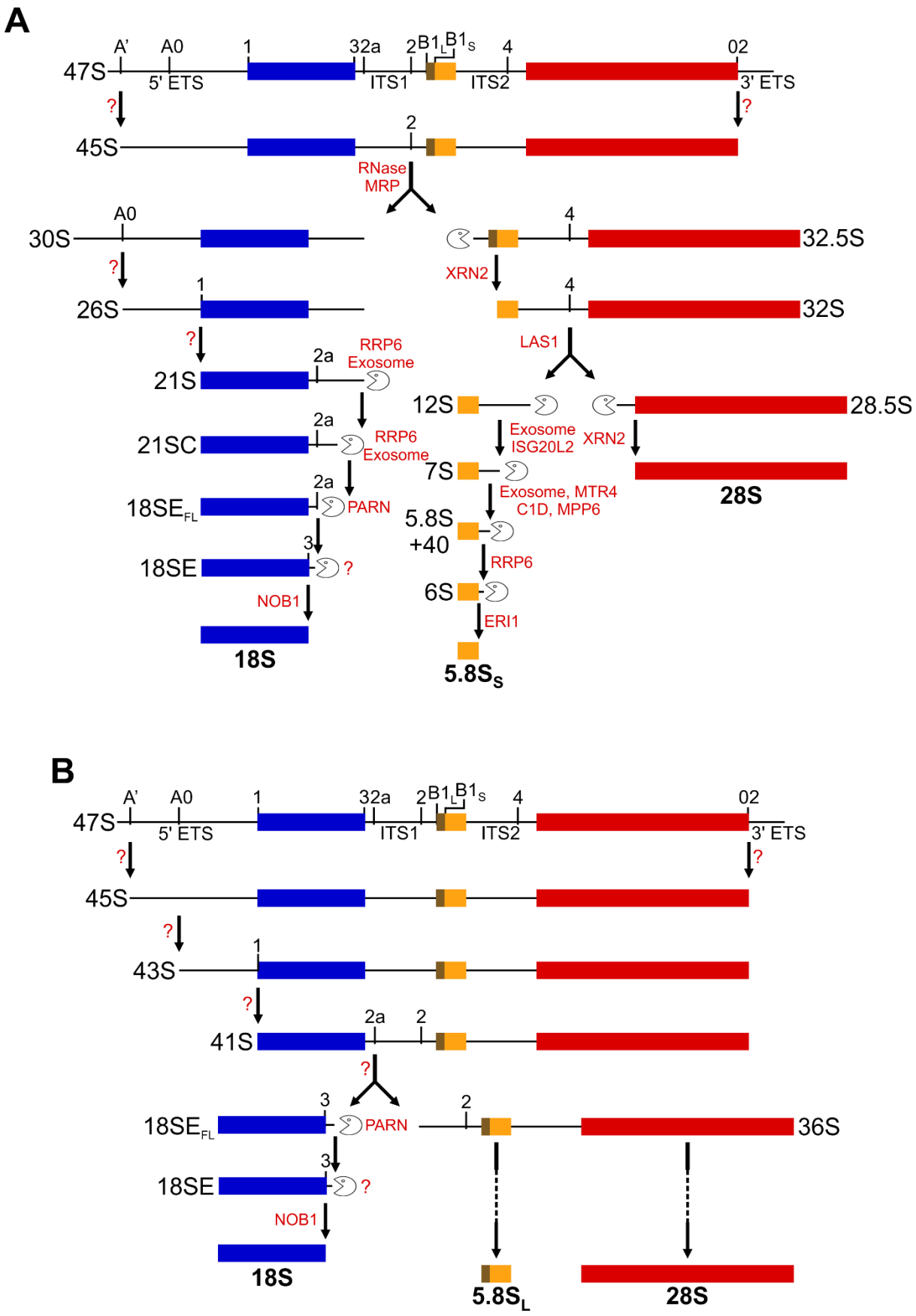


Figure 1.6 Pre-rRNA processing in *H. sapiens*. Schematic representation of pre-rRNA processing pathways in humans. Endonuclease cleavage sites are marked on the pre-rRNA and, where known, the endo- or exo-nuclease catalysing each processing event is shown. (A) The major pre-rRNA processing pathway in humans,

which involves predominant cleavage in ITS1 at site 2 and maturation of the 3' end of 18S via exonucleolytic processing involving the exosome complex without endonucleolytic cleavage at site 2a. (B) The minor pre-rRNA processing pathway in humans, involving cleavage at sites A0, 1 and 2a prior to site 2 cleavage.

As in yeast, mammalian ITS1 contains two endonuclease cleavage sites, called 2a (E) and 2. Unlike in yeast, the 3' most site, site 2, is the major cleavage site in ITS1 (Figure 1.6A)(Hadjiolova *et al.*, 1984a; Hadjiolova *et al.*, 1984c; Hadjiolova *et al.*, 1984b; Hadjiolova *et al.*, 1993; Mullineux and Lafontaine, 2012), suggesting that site 2 may be analogous to the major ITS1 cleavage site in yeast, A_2. However, the efficiency of cleavage at site 2 is not sensitive to impairment of cleavages at A0, site 1 or 2a, and is dependent on the human orthologues of yeast proteins required for cleavage at site A_3 (Lapik *et al.*, 2004; Preti *et al.*, 2013; Sloan *et al.*, 2013b). In contrast, the human 2a site cleavage is dependent on prior cleavages at sites A0 and 1 and is dependent on human orthologues of factors required for yeast A_2 cleavage (O'Donohue *et al.*, 2010; Carron *et al.*, 2011) (Figure 1.6B). These data suggest that site 2a is equivalent to yeast site A_2, while site 2 is equivalent to yeast site A_3. The enzyme responsible for 2a cleavage is unknown, and cleavage at this site is not necessary for production of the mature 18S rRNA under normal conditions. While yeast 18S 3' end processing appears to occur solely via endonucleolytic cleavages within ITS1, both ends of mammalian ITS1 can be processed by both endonucleases and exonucleases (Figure 1.6).

In mammals, there appears to be an alternative, "minor" pre-rRNA processing pathway, which more closely resembles the mechanism observed in yeast. In the minor pathway in human cells, cleavages at A0 and 1 can occur directly after A' cleavage, producing the 41S pre-rRNA. Uncoupling of site A0 and 1 cleavage events in the minor processing pathway results in accumulation of a 43S precursor (cleaved at A0 but not site 1) (Rouquette *et al.*, 2005; O'Donohue *et al.*, 2010). The 18SE precursor can then be produced in alternative manner, by direct endonucleolytic cleavage at site 2a, which also produces the LSU 36S intermediate. The 36S intermediate, which is not usually detectable under normal conditions in human cells, accumulates upon depletion of the 5' – 3' exonuclease XRN2, suggesting that processing can switch to the "minor" pathway in the absence or impairment of site 2 cleavage (Wang and Pestov, 2011; Sloan *et al.*, 2013b). Processing of 18SE then occurs as in the "major" pathway, with a final endonucleolytic cleavage at site 3 by NOB1 (Preti *et al.*, 2013; Bai *et al.*, 2016).

Pre-rRNA processing steps required for the maturation of LSU rRNAs are generally evolutionarily conserved (Lange *et al.*, 2011; Tafforeau *et al.*, 2013; Tomecki *et al.*, 2017). The 3' end maturation of the 28S rRNA occurs early in the processing pathway and may play a role in the termination of RNA polymerase I transcription (Nemeth *et al.*, 2013). After cleavage at site 2 in mammals, XRN2 (Rat1) processes the 5' end of mature 5.8S (Wang and Pestov, 2011; Schillewaert *et al.*, 2012). As in yeast, the 5.8S rRNA exists as a short ($5.8S_S$) and a long ($5.8S_L$) form, but depletion of XRN2, unlike Rat1 depletion in yeast, does not alter the ratio of short and long forms of 5.8S (Schillewaert *et al.*, 2012; Sloan *et al.*, 2013b). A single endonucleolytic cleavage in ITS2, at site 4 (equivalent to yeast site C_2), separates the 5.8S and 28S rRNAs, generating 12S and 28.5S pre-rRNAs. As for the equivalent yeast intermediate, the 12S pre-rRNA undergoes multiple exonuclease processing steps to form the mature 5.8S 3' end. These are mediated by the exosome complex, RRP6 and the ISG20L2 in the nucleus to generate the 6S pre-rRNA (Schilders *et al.*, 2007; Preti *et al.*, 2013; Tafforeau *et al.*, 2013). The mouse 6S intermediate is then exported to the cytoplasm where the 3' end of the 5.8S rRNA is matured by ERI1 (Ansel *et al.*, 2008). XRN2 is responsible for processing the 5' end of the 28.5S pre-rRNA, generating the mature 28S (Wang and Pestov, 2011).

1.4.3 Mammalian-specific aspects of 18S 3' end processing

In yeast, the major ITS1 cleavage occurs at site A_2, separating SSU and LSU pre-rRNA processing pathways. Cleavage at the other ITS1 cleavage site, A_3, is a secondary processing event initiating LSU maturation, but is unimportant for normal SSU maturation. Maturation of the 3' end of 18S in yeast does not appear to involve exonucleases. In mammals however, the major ITS1 processing pathway appears to favour a primary cleavage at site 2 (Mullineux and Lafontaine, 2012), followed by 3' to 5' exonucleolytic processing from site 2 (Figure 1.7A).

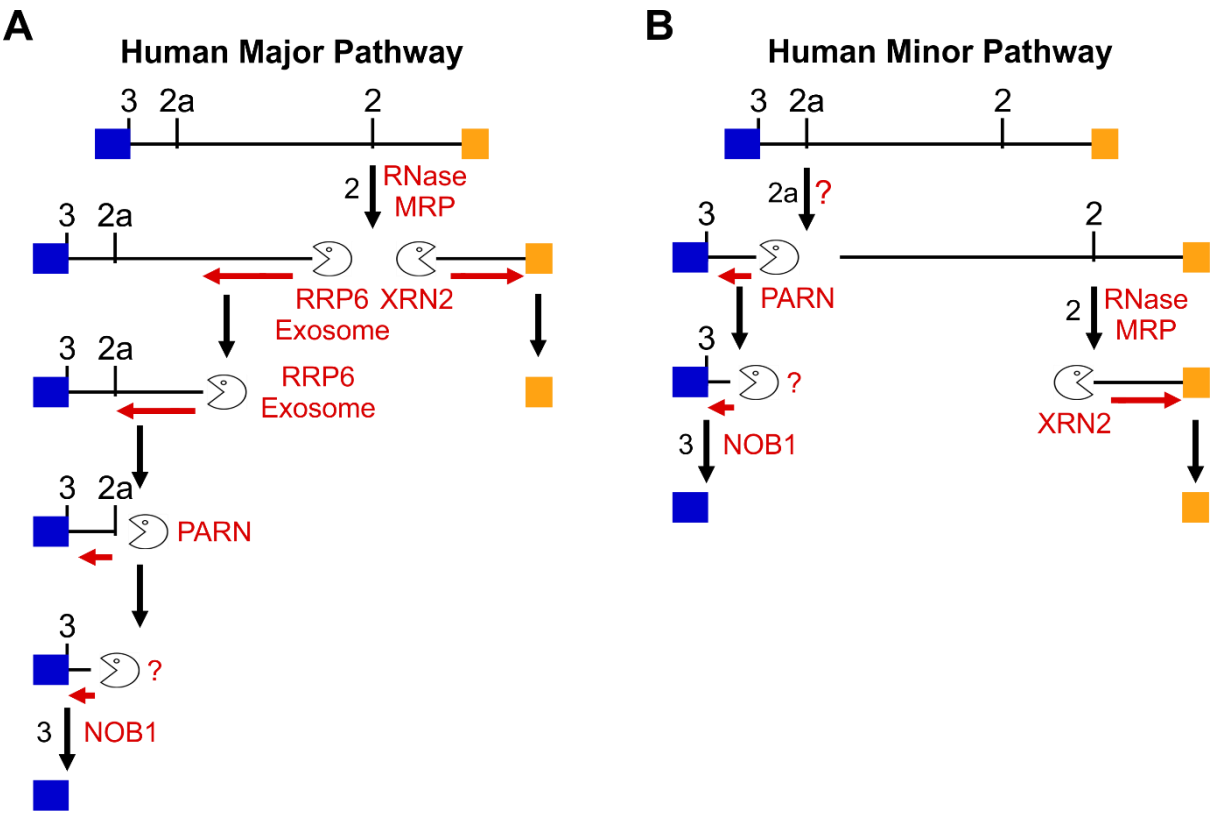


Figure 1.7 Comparison of major and minor ITS1 processing pathways in humans. Schematic representation of major and minor ITS1 processing pathways in humans, showing sequential processing steps and, where known, the enzymes responsible for catalysing processing events. **(A)** The major ITS1 processing pathway in humans, involving initial cleavage at site 2 followed by exonucleolytic processing to the prior to endonuclease cleavage at site 3 by NOB1. **(B)** The minor ITS1 processing pathway, involving initial cleavage at site 2a, followed by 3' to 5' exonucleolytic processing prior to NOB1-mediated site 3 cleavage.

Initial processing continues up to the boundary of the conserved domain C, around 250 nt from site 2, generating the 21SC pre-rRNA (Figure 1.6A; Figure 1.7A) (Carron *et al.*, 2011; Preti *et al.*, 2013; Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013). The 21SC precursor accumulates when the catalytic activity of the exosome-associated exonuclease RRP6 is abolished, suggesting that, in the major ITS1 processing pathway, this precursor is further processed exonucleolytically by RRP6 (a step which also requires the core exosome and exosome cofactors), efficiently producing 18SE (Sloan *et al.*, 2013b). Levels of 18SE were greatly reduced when RRP6 exonuclease activity was abolished or when exosome cofactor MTR4 was depleted and were greatly recovered upon simultaneous depletion of XRN2 and MTR4 (Sloan *et al.*, 2013b). The 18SE pre-rRNA is processed from site 2a, either following RRP6 trimming to this site, or after endonuclease cleavage at site 2a, by the 3' to 5' exonuclease Poly(A)-specific

Ribonuclease (PARN) (Ishikawa *et al.*, 2017; Montellese *et al.*, 2017). In the cytoplasm, the 18SE precursor undergoes further 3' to 5' exonucleolytic trimming, mediated by an unknown exonuclease (Preti *et al.*, 2013; Ishikawa *et al.*, 2017; Montellese *et al.*, 2017). It is unknown whether this cytoplasmic trimming is always followed by endonuclease cleavage at site 3 by NOB1, or if it represents an alternative mechanism for the maturation of the 3' end of 18S. Depletion of XRN2 in human and mouse cells results in accumulation of a 36S intermediate, which is cleaved at 2a but not site 2, indicating a switch from the "major" ("site 2") processing pathway to the "minor" ("2a") pathway (Wang and Pestov, 2011; Preti *et al.*, 2013; Sloan *et al.*, 2013b). XRN2 depletion does not affect levels of the mature 18S rRNA, suggesting that when exonucleolytic production of 18SE is impaired, 18SE can be produced by endonuclease cleavage at site 2a (Wang and Pestov, 2011; Preti *et al.*, 2013; Sloan *et al.*, 2013b) (Figure 1.6B; Figure 1.7B).

Once exported to the cytoplasm, 18SE appears to undergo further 3' exonucleolytic processing, mediated by an as yet unidentified 3' to 5' exonuclease, before endonucleolytic cleavage at the 3' end of 18S (site 3) by NOB1 (Preti *et al.*, 2013; Bai *et al.*, 2016). The exonucleolytic processing of 18SE in the cytoplasm could enable NOB1 to access the pre-rRNA for site 3 cleavage. Alternatively, cytoplasmic exonucleolytic processing of 18SE could continue to the 3' end of the mature 18S, representing an alternative processing mechanism for final maturation of the SSU rRNA. Interestingly, the 18SE pre-rRNA is oligouridylated in the cytoplasm in human cells (Preti *et al.*, 2013). The addition of poly(U) tails to RNAs by poly(U) polymerases has been identified in multiple organisms (Rissland and Norbury, 2008; Wilusz and Wilusz, 2008). A candidate exonuclease for the cytoplasmic processing of oligouridylation is the exosome-independent homologue of yeast Rrp44 (Dis3), DIS3L2, which degrades a wide range of oligouridylated ncRNAs in the cytoplasm (Ustianenko *et al.*, 2013; Pirouz *et al.*, 2016; Ustianenko *et al.*, 2016).

Human cytoplasmic poly(U) polymerases are involved in the degradation of histone mRNA (Mullen and Marzluff, 2008), and so addition of poly(U) tails may target RNA substrates for degradation, similar to short poly(A) tails added to pre-rRNAs in the nucleus by the TRAMP complex, which targets these transcripts for degradation by the exosome complex in yeast (LaCava *et al.*, 2005; Houseley *et al.*, 2006; Holub *et al.*, 2012). The activities of poly(U) polymerases are associated with eukaryotic ribosomes (Wilkie and Smellie, 1968; Hozumi *et al.*, 1975; Milchev and Hadjiolov,

1978; Hayashi and MacFarlane, 1979), and so it is possible that oligouridylation of the 18SE pre-rRNA targets it for degradation by the exosome upon impairment of late SSU processing steps. Alternatively, poly(U) tails may mediate alternative maturation of 18SE to the mature 18S rRNA by recruiting an exonuclease to process this precursor in the absence of NOB1-mediated cleavage at site 3. Interestingly, the proportion of oligouridylated 18SE is increased upon depletion of some SSU RPs and ribosome biogenesis factors (Preti *et al.*, 2013), including RPS10, which is frequently mutated in the ribosomopathy, Diamond-Blackfan Anaemia (DBA) (Doherty *et al.*, 2010).

It is unclear why ITS1 processing should differ between yeast and mammals, but it is likely due to the increased complexity of ribosome biogenesis in higher eukaryotes. The primary pre-rRNA transcript in humans is almost twice as long as in yeast (~13,000 nt vs. ~7,000 nt). The ITS1 spacer is ~350 nt longer in humans compared to yeast and is also considerably more GC-rich. This may make the 2a cleavage site less accessible, hence predominant ITS1 cleavage at site 2. Another key difference between the processing of ITS1 between yeast and mammals is in the timing of important cleavage events. 70-80 % of yeast pre-rRNA transcripts are cleaved at site A_2 in ITS1 co-transcriptionally (Kos and Tollervey, 2010), while human ITS1 appears to be mostly processed after termination of pre-rRNA transcription. Therefore, cleavage at the 2a site in mammalian cells may be less efficient when processing occurs post-transcriptionally, while cleavage at site 2 may act as the major ITS1 cleavage event. This may account for some of the differences observed in ITS1 processing mechanisms. Interestingly, when conditions are unfavourable, pre-rRNA processing in yeast switches to an alternative, post-transcriptional pathway, which more closely resembles the major human pathway with initial cleavage at site A_3 (Kos-Braun *et al.*, 2017).

1.5 snoRNPs in pre-rRNA Processing

While many snoRNPs are essential for ribosome biogenesis through their functions in covalent modification of the pre-rRNA, a subset of snoRNPs play a more direct role in pre-rRNA processing. One of these is RNase MRP, which, as described previously, is responsible for endonucleolytic cleavage at site A_3/2 in ITS1 (Lygerou *et al.*, 1996; Mattijssen *et al.*, 2010; Goldfarb and Cech, 2017). Four other snoRNPs in yeast and humans are known to be important for processing, with functions outside of directly catalysing cleavage events. These are the box C/D snoRNPs, U3 and U14,

and the box H/ACA snoRNPs, snR30/U17 and snR10 (Kiss *et al.*, 2010; Watkins and Bohnsack, 2012). U14 and snR10 appear to have retained their RNA modification functions, playing roles in both covalent modification of RNA and in pre-rRNA processing (Liang and Fournier, 1995; Morrissey and Tollervey, 1997; Liang *et al.*, 2010). U3 and snR30/U17 both lack a known modification target and are important for early endonuclease cleavages in 18S rRNA maturation (Hughes and Ares, 1991; Atzorn *et al.*, 2004; Fayet-Lebaron *et al.*, 2009; Lemay *et al.*, 2011). In higher eukaryotes, additional snoRNPs are also required for processing of both SSU and LSU pre-rRNAs.

1.5.1 U3 snoRNP

U3 was the first snoRNA identified and has been most extensively studied (Hodnett and Busch, 1968). It is evolutionarily conserved and has been identified in all eukaryotes tested (Charette and Gray, 2009; Marz and Stadler, 2009). While U3 is classified as a box C/D snoRNA, it does not direct 2'-*O*-methylation, and instead functions in the direction of pre-rRNA cleavage (Kass *et al.*, 1990; Hughes and Ares, 1991; Borovjagin and Gerbi, 2001). The U3 snoRNP is important for 18S rRNA maturation due to its requirement for cleavages at sites A_0/A0 and A_1/1 in the 5' ETS and at site A_2/2a in ITS1 (Hughes and Ares, 1991; Beltrame and Tollervey, 1992; Sharma and Tollervey, 1999; Langhendries *et al.*, 2016). In higher eukaryotes, U3 is also important for the additional 5' ETS cleavage at A' (Enright *et al.*, 1996; Prieto and McStay, 2005).

The U3 snoRNP contains the U3 snoRNA and the four core box C/D proteins Snu13/15.5K, Nop56, Nop58 and Nop1/fibrillarin, as well as a U3-specific protein, Rrp9/U3-55K (Watkins *et al.*, 2000). The 3' terminal domain of the U3 snoRNA contains unique box B/C motifs, important for the binding of Rrp9/U3-55K (Lubben *et al.*, 1993; Venema *et al.*, 2000; Granneman *et al.*, 2002). Sequence motifs in the 5' domain of the U3 snoRNA are important for base pairing between U3 and multiple regions in the pre-rRNA, both in the 5' ETS and in the mature 18S rRNA sequence (Figure 1.8). These include the box A, A' and GAC motifs, which contain sequences complementary to the mature 18S rRNA (Beltrame and Tollervey, 1995; Samarsky and Fournier, 1998; Antal *et al.*, 2000; Terns and Terns, 2002).

Base pairing between U3 and the pre-rRNA mediates the formation of the central pseudoknot, a conserved structural feature present in the mature 18S rRNA ((Sharma *et al.*, 1999; Granneman *et al.*, 2004; Kudla *et al.*, 2011). The 5' and 3' hinge sequences of U3 base pair with the 5' ETS (Beltrame and Tollervey, 1995; Marmier-Gourrier *et al.*, 2011). The Mpp10 complex, a subcomplex of the SSU processome, also associates with the hinge region of U3 (Wormsley *et al.*, 2001; Gerczei *et al.*, 2009). In *Xenopus*, the U3 hinge sequences base pair close to A' and A0 cleavage sites, potentially ensuring a conformation conducive to endonuclease cleavage of the pre-rRNA (Borovjagin and Gerbi, 2001; Borovjagin and Gerbi, 2004). In all eukaryotes, U3 snoRNA-pre-rRNA base pairing is required for correct folding of the pre-rRNA and likely functions as a chaperone to prevent the premature formation of the central pseudoknot in the 18S rRNA (Sharma and Tollervey, 1999; Dutca *et al.*, 2011). The central pseudoknot is a structural feature that is maintained in the mature ribosome and coordinates the four domains of the 18S RNA to form the decoding centre (Kudla *et al.*, 2011). U3:pre-rRNA base pairing likely also plays a role in the coordination of the A_2/2a cleavage with the 5' ETS cleavages, as it brings the 5' and 3' ends of 18S into close proximity within the SSU processome.

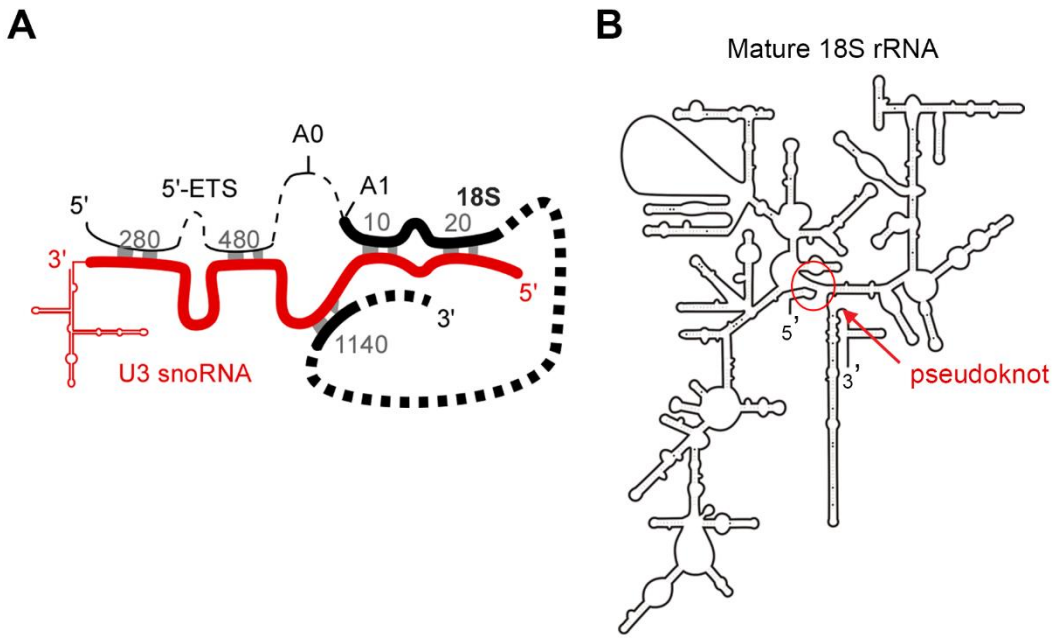


Figure 1.8 Base-pairing interactions between the U3 snoRNA and the pre-rRNA in yeast. (A) Schematic representation of the U3 snoRNA and the pre-rRNA during formation of the 18S central pseudoknot in yeast. The U3 snoRNA is shown in red and the pre-rRNA is shown in black. Labelled nucleotide numbers correspond to the nucleotide of either the 5' ETS or 18S. **(B)** Predicted secondary structure of the mature *S. cerevisiae* 18S rRNA. The position of the central pseudoknot is indicated by a red circle. From (Wells *et al.*, 2017).

The U3 snoRNP is a key subcomplex of the SSU processome, and many SSU processome components were identified through their association with the U3 snoRNA, U3 core proteins and/or Mpp10 (Dragon *et al.*, 2002; Grandi *et al.*, 2002). During assembly of the SSU processome, the U3 snoRNP is recruited to the 5' ETS co-transcriptionally (Chaker-Margot *et al.*, 2015), after the recruitment of the UtpA subcomplex, and around the same time as the UtpB subcomplex (see section 1.7.1) (Zhang *et al.*, 2016b). A recently published structure of the *Chaetomium thermophilum* SSU processome shows the 5' region of the U3 snoRNA penetrating into the core of the SSU processome and stabilising the region of 18S around the A_1 cleavage site (Kornprobst *et al.*, 2016). After early cleavages at sites A_0, A_1 and A_2, the U3 snoRNP (along with other early-acting SSU ribosome biogenesis factors) is released from pre-ribosomal particles (Panse and Johnson, 2010; Strunk *et al.*, 2011). Release of the U3 snoRNA from the pre-rRNA requires the ATP-dependent helicase activity of the SSU processome component Dhr1 (Sardana *et al.*, 2015), which associates with pre-rRNA in the late stages of SSU processome assembly (Chaker-Margot *et al.*, 2015). Controlled release of U3 from base pairing interactions with the pre-rRNA likely permits the spontaneous formation of the central pseudoknot.

1.5.2 snR30/U17

snR30 (U17 in humans) belongs to the box H/ACA snoRNP family, which are responsible for pseudouridylation of rRNA that aid the maintenance of RNA secondary structures (Decatur and Fournier, 2002; Helm, 2006; Kiss *et al.*, 2010). Like canonical box H/ACA snoRNAs, snR30 associates with the four core box H/ACA proteins, Nhp2, Gar1, Nop10 and Cbf5 (Watkins *et al.*, 1998; Liang and Fournier, 2006; Lemay *et al.*, 2011). However, snR30 has no known modification target, and its rRNA recognition sequence occupies a different position within the 3' hairpin compared to those of box H/ACA snoRNAs involved in RNA modification (Atzorn *et al.*, 2004; Fayet-Lebaron *et al.*, 2009). Instead of directing pseudouridylation, snR30/U17 plays an evolutionarily conserved role in 18S rRNA maturation (Morrissey and Tollervey, 1993; Enright *et al.*, 1996; Atzorn *et al.*, 2004). snR30 base pairs to 18S and disruption of this interaction inhibits cleavage at sites A_0 and A_1 in the 5' ETS, and at A_2 in ITS1 (Atzorn *et al.*, 2004; Fayet-Lebaron *et al.*, 2009).

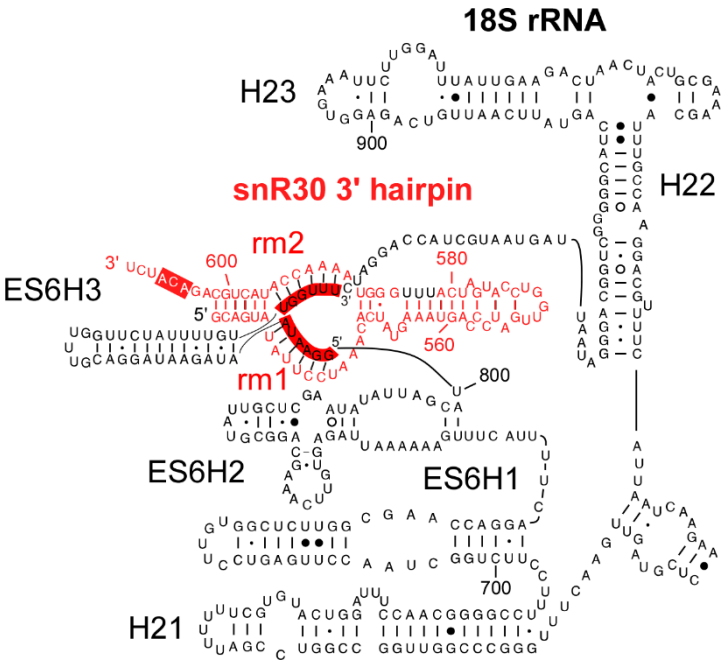


Figure 1.9 Predicted structure of the ES6 region of 18S during snR30 base pairing with the pre-rRNA. The predicted structure of the 18S rRNA ES6 region during base-pairing between 18S and the snR30 snoRNA. The snR30 3' hairpin is shown in red, and the snR30 binding sites, labelled rm1 and rm2, on the 18S structure are highlighted in red. From (Wells *et al.*, 2017).

Recognition elements m1 and m2 within the internal loop of the snR30 3' hairpin base pair to rm1 and rm2 elements in the eukaryote-specific expansion sequence 6 (ES6) region of the 18S rRNA (Figure 1.9) (Atzorn *et al.*, 2004; Fayet-Lebaron *et al.*, 2009). Multiple ribosome biogenesis factors also crosslink to the 18S in this region, including the PIN domain proteins Utp23 (Wells *et al.*, 2017) and Utp24 (Wells *et al.*, 2016), the RNA helicase Rok1 (Martin *et al.*, 2014), the RNA-binding protein Rrp7 (Lin *et al.*, 2013) and the RNA-binding compaction factor Rrp5 (Lebaron *et al.*, 2013). These factors are components of the SSU processome, assembly of which is essential for cleavages at A_0, A_1 and A_2 and snR30 is also required for the assembly of this complex (Lemay *et al.*, 2011). snR30 is also essential for the integration of Utp23 and Kri1 into preribosomes (Hoareau-Aveilla *et al.*, 2012).

Release of snR30 from pre-ribosomes requires the DEAD box RNA helicases Rok1 (Bohnsack *et al.*, 2008; Martin *et al.*, 2014) and Has1 (Liang and Fournier, 2006) and also involves Rrp5 (Khoshnevis *et al.*, 2016). Utp23 is also required for snR30 release (Hoareau-Aveilla *et al.*, 2012) and Utp23 binds to both snR30 and snR30-base pairing regions of 18S (Wells *et al.*, 2017). Therefore, the binding of snR30 and its associated factors may play an important role in coordination of interactions between

other factors and the pre-rRNA during pre-rRNA processing. Much less is known about the human homologue of snR30, the U17 snoRNA. However, human UTP23 has been shown to associate with U17 *in vivo* (Hoareau-Aveilla *et al.*, 2012).

1.5.3 Other snoRNPs in Pre-rRNA Processing

Other snoRNPs with roles in pre-rRNA processing in eukaryotes include the box C/D snoRNP U14, and the H/ACA snoRNP snR10. U14 is highly conserved and essential for cell viability. It base pairs with the 18S rRNA sequence and is required for cleavages at sites A_0, A_1 and A_2 (Lempicki *et al.*, 1990; Li *et al.*, 1990; Liang and Fournier, 1995; Morrissey and Tollervey, 1997). Unlike U3, U14 appears to have retained its rRNA modification function, and therefore plays dual roles in both modification and processing (Torchet and Hermann-Le Denmat, 2002; Atzorn *et al.*, 2004). In higher eukaryotes, U14 is required for the initial 5' ETS cleavage at site A' (Enright *et al.*, 1996; Saez-Vasquez *et al.*, 2004). snR10 is a non-essential H/ACA snoRNP, thought to be specific to yeast which, like U14, has roles in both covalent rRNA modification and pre-rRNA processing (Tollervey, 1987; Liang *et al.*, 2010). The modification role of snR10 involves pseudouridylation of the U2923 residue in the PTC region of the 25S rRNA of the LSU (Ganot *et al.*, 1997a; Ni *et al.*, 1997). The H/ACA box snoRNP snR10 is also important for 18S maturation, as its depletion disrupts cleavage at sites A_0-A_2 (Tollervey, 1987). Although not an essential snoRNA, deletion of snR10 caused defects in growth and pre-rRNA processing (Tollervey, 1987), and it is synthetical lethal upon mutation of the helicase Rok1 or the RNA-binding protein Rrp5 (Venema and Tollervey, 1996; Venema *et al.*, 1997). snR10 interacts with Rrp5 (Lebaron *et al.*, 2013), Rrp7 (Lin *et al.*, 2013) and Rok1 (Martin *et al.*, 2014). There are also multiple higher eukaryote-specific snoRNPs with roles in pre-rRNA processing. These include snoRNPs U8, which is important for maturation of LSU rRNAs 5.8S and 28S in *X. laevis* (Peculis *et al.*, 2001) and humans (Langhendries *et al.*, 2016), and U22 which is required for 5' ETS and ITS1 processing in 18S rRNA maturation in *X. laevis* (Tycowski *et al.*, 1996).

1.6 Pre-ribosomal particles in ribosome biogenesis

As well as snoRNPs, many non-ribosomal protein factors are required for pre-rRNA processing. Like processing snoRNAs, and often through association with snoRNPs, many ribosome biogenesis factors associate with the nascent pre-rRNA

during pre-rRNA transcription (Tschochner and Hurt, 2003). Association of ribosome biogenesis factors, snoRNPs and RPs with the pre-rRNA forms large pre-ribosomal particles, within which pre-rRNA processing takes place. Because the majority of pre-rRNAs are cleaved co-transcriptionally in yeast, at sites A_0, A_1 and A_2, factors required for SSU maturation assemble onto the nascent pre-rRNA early in the pre-rRNA processing pathway. Co-transcriptional pre-rRNA processing can be observed by EM following chromatin spreading (Miller and Beatty, 1969; Mougey *et al.*, 1993).

In these experiments, terminal structures can be observed at the 5' end of nascent pre-rRNA transcripts. These structures were later identified as the SSU processome, which contains the compacted pre-SSU pre-rRNA (Dragon *et al.*, 2002; Osheim *et al.*, 2004; Phipps *et al.*, 2011). Pre-SSU particles are released co-transcriptionally after A_2 cleavage (Osheim *et al.*, 2004). Pre-ribosomal particles are observed on nascent transcripts, but there is currently no evidence of co-transcriptional release of pre-SSU particles in higher eukaryotes, consistent with higher eukaryotic pre-rRNA processing occurring post-transcriptionally (Mougey *et al.*, 1993; Osheim *et al.*, 2004)**.** Many ribosome biogenesis factors have been identified by analysing the effects of their depletion on pre-rRNA processing through detection of pre-rRNA intermediates. Gradient analysis and purification of pre-ribosomal particles have also been useful for identifying factors associating with distinct pre-ribosomal complexes.

1.6.1 Ribosome Biogenesis Factors required for early pre-40S maturation events

Endonuclease cleavage(s) within ITS1 separates pre-SSU processing from the pre-LSU maturation pathway. Following this separation of processing, the two distinct processing pathways occur independently. Therefore, the vast majority of ribosome biogenesis factors function in either SSU or LSU maturation, with very few factors acting in both pathways. The RNA binding protein Rrp5 is an exception to this, as it is involved in processing at both sites within ITS1, at A_2 in 18S rRNA maturation, and A_3 in 5.8S rRNA maturation (Venema and Tollervey, 1996). The ribosome biogenesis factors involved in early steps of 18S maturation are required for cleavage at sites A_0, A_1 and A_2 in the 5' ETS and ITS1. These cleavage events are dependent on the U3 snoRNP, which base pairs with the pre-rRNA in the 5' ETS and with the mature 18S rRNA sequence (Beltrame and Tollervey, 1992; Beltrame and Tollervey, 1995; Borovjagin and Gerbi, 2005; Kudla *et al.*, 2011). The U3 snoRNP is a component of

the SSU processome, the formation of which on the pre-rRNA is required for cleavages at all three early pre-18S cleavages (see section 1.7).

1.6.2 Late pre-40S Particles

Following cleavages at sites A_0, A_1 and A_2, the composition of the yeast SSU pre-ribosomal particle changes dramatically, with many early-acting factors released and late-acting factors being recruited (Zhang *et al.*, 2016b). Late pre-40S complexes in yeast include factors responsible for the final processing step in 18S rRNA maturation, such as the site D endonuclease Nob1 (Fatica *et al.*, 2003; Pertschy *et al.*, 2009; Chaker-Margot *et al.*, 2015). Late-acting factors prevent premature translation initiation by blocking the final, Nob1-mediated, maturation of the 18S rRNA before recruitment of 60S subunits (Lebaron *et al.*, 2012; Strunk *et al.*, 2012). Factors involved in this proofreading step include Tsr1, the kinase Rio2 and the GTPase Fun12 (eIF5B), and the binding of Fun12 to the LSU 25S rRNA is required for Nob1-mediated cleavage of the 20S pre-rRNA (Lebaron *et al.*, 2012; Strunk *et al.*, 2012).

1.6.3 Pre-60S Particles

After cleavage at ITS1, and separation of yeast SSU and LSU pre-rRNA intermediates, many factors required for the maturation of the 5.8S and 25S rRNAs are recruited to the $27SA_2$ precursor (Konikkat and Woolford, 2017) . One of the earliest LSU maturation factors to assemble in pre-60S particles is the compaction factor, Rrp5, which is also important for SSU maturation (Eppens *et al.*, 1999). The majority of $27SA_2$ intermediates generated by A_2 cleavage are also cleaved at site A_3 by RNase MRP to produce a $27SA_3$ precursor (Schmitt and Clayton, 1993; Chu *et al.*, 1994; Lygerou *et al.*, 1996). Multiple ribosome biogenesis factors are required for either cleavage at A_3, or for processing of the $27SA_3$ intermediate, and these are termed A_3 cluster proteins. These include Nop7, Rlp7, Rrp1, Nop12, Nop15, Ebp2, Pwp1, Erb1, Brx1, Nop12, Cic1 and Ytm1 (Pestov *et al.*, 2001; Wu *et al.*, 2001; Gadal *et al.*, 2002; Oeffinger *et al.*, 2002; Oeffinger and Tollervey, 2003; Miles *et al.*, 2005; Granneman *et al.*, 2011; Sahasranaman *et al.*, 2011; Dembowski *et al.*, 2013; Talkish *et al.*, 2014). While important for the removal of the 3' end of ITS1, the A_3 cluster proteins crosslink to a range of sites on the pre-rRNA, including ITS2 and the 5.8S and 25S rRNA sequences (Granneman *et al.*, 2011). Therefore, it is likely that these proteins play structural roles that ensure the correct formation of pre-60S particles that is likely

required for efficient processing of ITS1. Several RNA helicases are involved in early LSU maturation and are released before cleavage at site C_2 in ITS2, likely after functioning in conformational changes in the pre-60S complex (de la Cruz *et al.*, 2004; Woolford and Baserga, 2013).

Many new factors associate with pre-60S complexes during processing of ITS2, while many earlier-acting factors are released (Wu *et al.*, 2016). New factors recruited to pre-ribosomes for cleavage at C_2 include several RNA binding proteins, multiple helicases and a putative methyltransferase (Woolford and Baserga, 2013). After ITS2 removal, the 25S rRNA 5' end is matured by exonucleolytic processing by Rat1 and Rrp17 (Geerlings *et al.*, 2000). Multiple ribosome biogenesis factors are required for the 3' end processing of the 5.8S rRNA, including the AAA ATPase Rea1 (Galani *et al.*, 2004), which may be important for the removal of earlier-acting factors to enable access of exonucleases to the pre-rRNA, as well as structural rearrangements required for export of the pre-60S complex (Bassler *et al.*, 2001). The 5S RNP assembles with the early pre-60S particle (Zhang *et al.*, 2007), before its translocation to the nucleoplasm and subsequent export to the cytoplasm. Final maturation of the 25S rRNA occurs in the cytoplasm and remaining non-ribosomal proteins are released.

1.7 The SSU Processome/90S Pre-ribosome

The SSU processome is a large, ~80S RNP complex that assembles onto the nascent pre-rRNA in the nucleolus (Figure 1.10) (Phipps *et al.*, 2011). It was identified via analysis of proteins co-purifying with the U3 snoRNA, whose role in 18S rRNA maturation was well established, and was predicted to correspond to the terminal knobs observed at the 5' end of nascent pre-rRNAs visualised by electron microscopy (EM) of chromatin spreads of yeast rRNA genes (Dragon *et al.*, 2002; Bernstein *et al.*, 2004). Several components had previously been identified as U3-associated proteins and/or ribosome biogenesis factors, but 17 proteins not previously implicated in pre-rRNA processing were identified and termed U three proteins (Utp1-17). Depletion of the U3 snoRNA or individual Utp proteins resulted in the loss of terminal knobs on the pre-rRNA, suggesting that these knobs correspond to the SSU processome and that individual components are essential for formation of the complex (Dragon *et al.*, 2002). Independently, a large RNP complex containing many overlapping proteins was identified and called the 90S pre-ribosome (Grandi *et al.*, 2002). Both of these complexes exclusively contain SSU RPs and ribosome biogenesis factors required for

SSU maturation, consistent with SSU and LSU processing occurring independently following cleavage in ITS1, and with the ability of independently transcribed pre-SSU and pre-LSU rDNA coding units to be correctly matured into functional ribosomes (Liang and Fournier, 1997). The yeast SSU processome consists of around 70 non-ribosomal protein components as well as multiple snoRNAs.

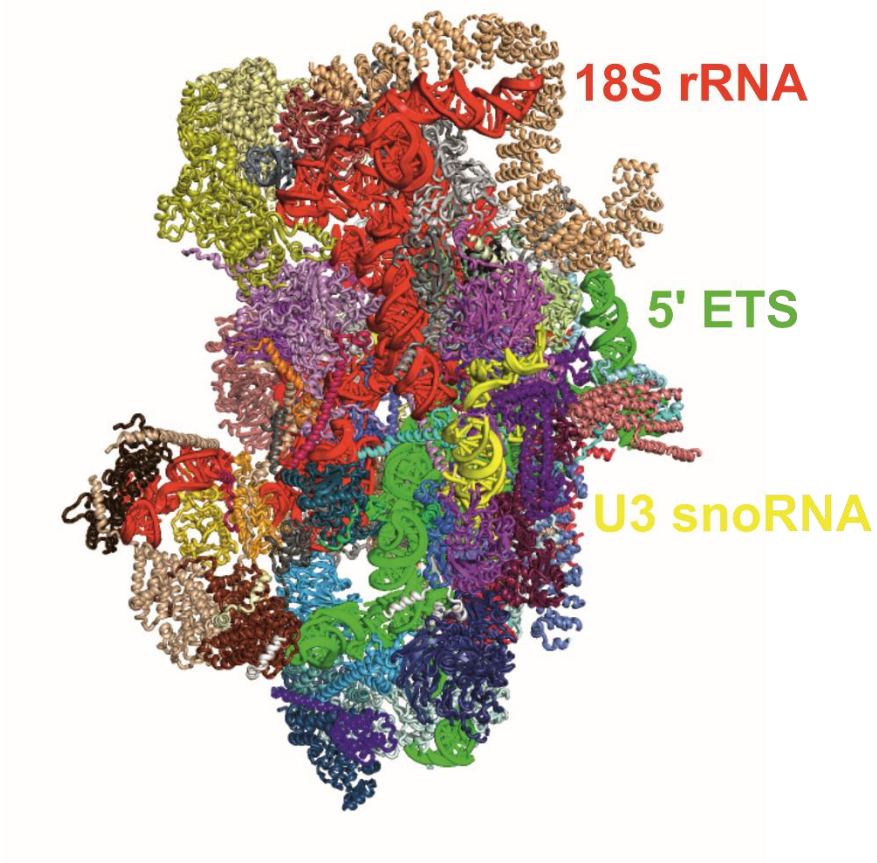


Figure 1.10 Structure of the yeast SSU processome. Cryo-EM structure of the S. cerevisiae SSU processome captured prior to cleavage at site A1. The 18S rRNA is shown in red and the 5' ETS is shown in green. The U3 snoRNA is shown in yellow. Structure from (Barandun et al., 2017).

1.7.1 SSU Processome Assembly

The SSU processome assembles onto nascent pre-rRNA transcripts in a stepwise, hierarchical fashion, beginning with the binding of several subcomplexes to the 5' ETS (Figure 1.11) (Perez-Fernandez *et al.*, 2007; Chaker-Margot *et al.*, 2015; Zhang *et al.*, 2016b). As the pre-rRNA is transcribed, further factors assemble into the complex and the pre-rRNA undergoes processing. After processing events in the 5' ETS and ITS1, the pre-ribosomal particle undergoes several conformational changes,

with many early-acting factors dissociating. The pre-40S complex, containing the 20S pre-rRNA, is then exported to the cytoplasm for final maturation of the 18S rRNA.

The first subcomplex to form during SSU processome assembly is the UtpA, or transcription associated Utp (t-Utp) subcomplex (Krogan *et al.*, 2004), which forms independently of the U3 snoRNP and, unlike other subcomplexes, is required for transcription of pre-rRNA in addition to processing (Bernstein *et al.*, 2004; Gallagher *et al.*, 2004; Prieto and McStay, 2007). All seven yeast UtpA proteins (Utp4, Utp5, Utp8, Utp9, Utp10, Utp15 and Utp17) were confirmed to associate with a fragment of the 5' ETS RNA just 281 nucleotides long, which does not include the first U3 snoRNA-binding site, *in vivo* (Zhang *et al.*, 2016b). This complex is present at the base of the SSU processome in the recently published yeast structure, where it is thought to act as a scaffold and recruit other factors to the 5' ETS (Barandun *et al.*, 2017). The binding of the UtpA subcomplex to the pre-rRNA is required for the recruitment of subsequent subcomplexes and other SSU processome components (Perez-Fernandez *et al.*, 2007; Perez-Fernandez *et al.*, 2011). The requirement of the UtpA subcomplex for pre-rRNA transcription is conserved in human cells and it may be recruited to rDNA by the transcription factor UBF (Prieto and McStay, 2007; Turowski and Tollervey, 2015). However, two of the seven yeast UtpA proteins (Utp8 and Utp9) do not have known human homologs, suggesting there may be as yet unidentified human UtpA subcomplex components (Prieto and McStay, 2007). Indeed, a metazoan-specific UtpA protein, NOL11, has been identified, and is implicated in ribosomopathies through its interaction with the human homolog of Utp4 (Freed *et al.*, 2012; Griffin *et al.*, 2015). When RNA polymerase I-mediated transcription of pre-rRNA is blocked by actinomycin D treatment of human cells, a novel 50S intermediate accumulates, and accumulation of this complex is also observed upon depletion of UtpA proteins (Turner *et al.*, 2009). This 50S SSU processome intermediate does not contain UtpA, UtpB or Mpp10 subcomplex proteins or the pre-rRNA, but does contain the large RNA binding protein Rrp5 and the U3 snoRNP (Turner *et al.*, 2009).

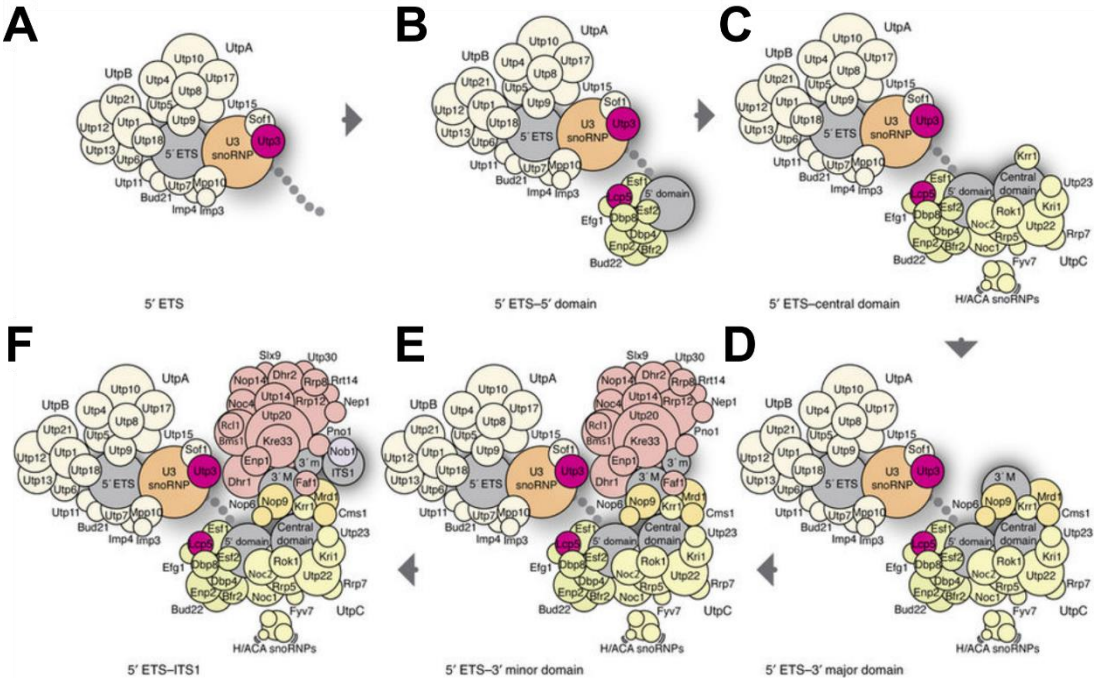


Figure 1.11 Hierarchical assembly of the SSU Processome. Schematic representation of the proposed model for co-transcriptional assembly of the SSU processome. Protein factors hierarchically associating with a pre-rRNA containing sequences up to the end of the 5' ETS (A), the 18S 5' domain (B), the 18S central domain (C), the 3' major domain (D), the 3' minor domain (E), and ITS1 (F) are shown. The pre-rRNA is shown in grey and the U3 snoRNP is shown in orange. From (Chaker-Margot et al., 2015).

The second complex that associates with the 5' ETS during assembly of the yeast SSU processome is the UtpB complex, which consists of Pwp2/Utp1, Utp6, Utp12, Utp13, Utp18 and Utp21 (Krogan *et al.*, 2004). Despite being capable of direct binding to the 5' ETS (Dosil and Bustelo, 2004), prior association of the UtpA subcomplex is required for UtpB recruitment (Perez-Fernandez *et al.*, 2007). Together, the UtpA and UtpB subcomplexes form a "5' ETS particle", which chaperones the pre-rRNA and the U3 snoRNA (Hunziker *et al.*, 2016), and, together, account for ~40 % of the mass of the SSU processome (Sun *et al.*, 2017). Many UtpB proteins contain protein-protein interaction domains (e.g. WD40), and the complex appears to function as the main structural core of the SSU processome (Dosil and Bustelo, 2004; Zhang *et al.*, 2016a; Barandun *et al.*, 2017; Cheng *et al.*, 2017). UtpB proteins associate with the pre-ribosomal complex around the same time as the U3 snoRNP and are important for the incorporation of the U3 snoRNP and the Mpp10 subcomplex (Zhang *et al.*, 2016b). Very little is currently known about the UtpB subcomplex in human cells, although all components are conserved, and UTP13 has been shown to associate with the U3 snoRNA in the human SSU processome (Turner *et al.*, 2009). Mutations within

two UtpB proteins, UTP21 (glaucoma) and UTP6 (neurofibromatosis), are also associated with human disease (Bartelt-Kirbach *et al.*, 2009; Gallenberger *et al.*, 2011).

As described in section 1.5.1, the U3 snoRNP consists of the U3 snoRNA, four core box C/D proteins, Nop1, Nop56, Nop58 and Snu13/15.5K, and a U3-specific protein, Rrp9/U3-55K (Watkins *et al.*, 2000). The U3 snoRNA base pairs with distant sequences in both the 5' ETS and the 18S rRNA, to chaperone RNA folding (Dutca *et al.*, 2011; Kudla *et al.*, 2011; Marmier-Gourrier *et al.*, 2011; Barandun *et al.*, 2017; Cheng *et al.*, 2017; Sun *et al.*, 2017). This base pairing is thought to prevent the premature folding of the pre-rRNA into the central pseudoknot (Hughes, 1996; Sharma and Tollervey, 1999; Kudla *et al.*, 2011). In recent SSU processome EM structures, the U3 snoRNA is positioned at the centre of the pre-ribosomal particle (Kornprobst *et al.*, 2016; Barandun *et al.*, 2017). The U3-specific protein Rrp9 (U3-55K) binds to the box B/C region of the U3 snoRNA and Snu13 (15.5K) (Granneman *et al.*, 2004; Cheng *et al.*, 2017; Sun *et al.*, 2017), and the Rrp9-Snu13 interaction is thought to enhance the binding of Rrp9 to the U3 snoRNA (Zhang *et al.*, 2013). All U3 snoRNP proteins appear to be recruited to the 5' ETS around the same time as the UtpB subcomplex (Zhang *et al.*, 2016b) and incorporation of the U3 snoRNP requires prior UtpB association (Dosil and Bustelo, 2004).

Interestingly, the first U3 binding site in the 5' ETS is not sufficient for the incorporation of the U3 snoRNP (Zhang *et al.*, 2016b), suggesting that protein-protein interactions, in addition to U3-pre-rRNA base pairing, contribute to U3 snoRNP recruitment, however it is unclear if this is the case in the cell. In humans, U3 snoRNA base pairing with the 5' ETS does not appear to be required for U3 snoRNP incorporation into the pre-ribosomal particle**,** and instead, the U3-specific protein U3-55K (Rrp9) is thought to be important for recruitment (Granneman *et al.*, 2004). The binding of U3-55K requires prior binding of 15.5K (Snu13) (Granneman *et al.*, 2002), suggesting the mechanism of U3 snoRNP recruitment may be conserved from yeast to humans.

The final subcomplex that associates with the 5' ETS is the Mpp10 complex, consisting of Mpp10, Imp3 and Imp4 (Lee and Baserga, 1999; Wehner *et al.*, 2002; Granneman *et al.*, 2003). This complex associates with the pre-ribosomal particle around the same time as the UtpB and U3 snoRNP subcomplexes (Zhang *et al.*, 2016b), and is thought to enhance the base pairing interactions between U3 and the

pre-rRNA in both the 5' ETS and the 18S rRNA (Gerczei and Correll, 2004; Gerczei *et al.*, 2009; Shah *et al.*, 2013). MPP10, IMP3 and IMP4 also form a stable complex in human cells and associate with the U3 snoRNA in the SSU processome (Granneman *et al.*, 2003). Another subcomplex, UtpC, associates later in SSU processome assembly. The UtpC complex consists of Rrp7, Utp22 and the four casein kinase II subunits, CkA_1, CkA_2, Ckb1 and Ckb2 (Krogan *et al.*, 2004; Lin *et al.*, 2013). This subcomplex is incorporated into the SSU processome after transcription of the central domain of the 18S rRNA, around the same time as the RNA binding protein Rrp5 (Chaker-Margot *et al.*, 2015; Zhang *et al.*, 2016b). Rrp5 is required for the incorporation of this complex (Vos *et al.*, 2004a; Perez-Fernandez *et al.*, 2007), and the C-terminal, protein-protein interaction, TPR repeats of Rrp5 are positioned close to UtpC in the SSU processome structure (Barandun *et al.*, 2017). This is consistent with the importance of the C-terminal Rrp5 TPR repeats in cleavages at sites A_0, A_1 and A_2 (Torchet *et al.*, 1998; Eppens *et al.*, 1999; Lebaron *et al.*, 2013).

As SSU processome assembly progresses, and pre-rRNA processing is initiated, late-acting factors are recruited to the complex, and early-acting factors dissociate from the SSU processome. Dissociating factors include the PIN domain protein Utp23 and the U14 and snR30 snoRNAs (Zhang *et al.*, 2016b). The U3 snoRNA, in contrast, remains stably associated with the pre-ribosome. One factor found to associate during late assembly is the helicase Dhr1, which is essential for the removal of the U3 snoRNA from the pre-rRNA (Chaker-Margot *et al.*, 2015; Sardana *et al.*, 2015). Within the SSU processome, the four domains of the 18S rRNA, and sequences that form the central pseudoknot (Kudla *et al.*, 2011) in the mature 40S subunit, are separated in different regions, in an open conformation, mediated by base pairing with the U3 snoRNA (Barandun *et al.*, 2017; Cheng *et al.*, 2017). This is entirely different to the compact conformation of these domains in the mature 40S subunit (Ben-Shem *et al.*, 2011). Therefore, the removal of the U3 snoRNA from the pre-rRNA by the helicase activity of Dhr1 likely allows for the final formation of the pseudoknot. Comprehensive analysis of SSU processome assembly, or analysis of SSU processome structure, have not been performed in mammals, and most of our current understanding comes from the study of individual homologues of yeast SSU processome components (Granneman *et al.*, 2002; Granneman *et al.*, 2003; Prieto and McStay, 2007; Turner *et al.*, 2009; Gerus *et al.*, 2010; Wang and Pestov, 2011; Turner

et al., 2012; Sloan *et al.*, 2013b; Sloan *et al.*, 2014; Wang *et al.*, 2014; Sloan *et al.*, 2015; Wang *et al.*, 2016).

1.7.2 SSU Processome Components

In addition to the many structural protein factors that make up the SSU processome, the list of components includes several enzymes, cofactors and regulatory proteins with confirmed, predicted or as yet unknown functions in pre-rRNA processing. This includes candidate endonucleases for pre-rRNA cleavages at sites A_0, A_1 and A_2 (see below) (Billy *et al.*, 2000; Bleichert *et al.*, 2006; Horn *et al.*, 2011; Wells *et al.*, 2016), as well as RNA binding proteins (Sasaki *et al.*, 2000; Dragon *et al.*, 2002; Vos *et al.*, 2004a; Vos *et al.*, 2004b; Hoareau-Aveilla *et al.*, 2012; Lin *et al.*, 2013), RNA helicases and their co-factors (Granneman *et al.*, 2006; Bleichert and Baserga, 2007; Bohnsack *et al.*, 2008; Rodriguez-Galan *et al.*, 2013; Martin *et al.*, 2014), kinases (Vos *et al.*, 2004a) and GTPases (Karbstein *et al.*, 2005; Strunk and Karbstein, 2009).

1.7.2.1 Candidate Endonucleases

While the endonucleases responsible for some early pre-18S cleavages remain elusive or are unconfirmed, it is likely that they are components of the SSU processome, and multiple candidate endonucleases have been identified in the complex. Functional analysis of individual SSU processome components has proved difficult, due to the fact that depletion of SSU processome components always affects the assembly of the complex, and therefore causes an identical pre-rRNA processing phenotype, namely a block in processing at sites A_0/A0, A_1/1 and A_2/2a. Therefore, identification of the specific nuclease that cleaves at an individual site is impossible using simple depletion of a candidate nuclease, and instead relies on mutational analysis.

1.7.2.1.1 Utp24

Utp24 is a candidate endonuclease as it contains a PIN (PilT N-terminal) domain (Figure 1.12). PIN domains share similarity with RNase H and the Fen1 endonuclease, and PIN domain-containing proteins were initially predicted to function as exonucleases (Clissold and Ponting, 2000; Arcus *et al.*, 2004; Levin *et al.*, 2004).

However, PIN domains are found in several endonucleases which function in a diverse range of RNA processing and degradation pathways including nonsense-mediated decay (NMD) (Eberle *et al.*, 2009; Skruzny *et al.*, 2009; Uehata and Akira, 2013). One of the best characterised PIN domain proteins is Nob1, which functions in the final processing step in 18S rRNA maturation, the cytoplasmic endonucleolytic cleavage at site D/3 (Fatica *et al.*, 2003; Lamanna and Karbstein, 2009; Pertschy *et al.*, 2009; Bai *et al.*, 2016). PIN domains exhibiting active endonuclease activity contain three or four conserved acidic residues, which form an active site that coordinates a metal ion (Mn^{2+} or Mg^{2+}) (Arcus *et al.*, 2004; Glavan *et al.*, 2006; Veith *et al.*, 2012). Mutation of just one of the three or four catalytic residues is enough to abolish cleavage activity of PIN domain proteins, as has been shown for the site D endonuclease Nob1 (Pertschy *et al.*, 2009).

While the depletion of Utp24 in yeast disrupts all three endonuclease cleavages in the 5' ETS and ITS1, substitution of an acidic PIN domain active site residue with an uncharged histidine impaired cleavage at sites A_1 and A_2, but permitted normal A_0 cleavage *in vivo* (Bleichert *et al.*, 2006). This led to the suggestion that Utp24 may be the endonuclease responsible for cleavage at sites A_1 and/or A_2, although its endonuclease activity was not confirmed. Utp24 also contains a Zinc finger motif within its PIN domain, predicted to be involved in RNA binding, although the importance of this motif for pre-rRNA processing is unknown. *In vivo* RNA-protein crosslinking (CRAC) experiments determined that yeast Utp24 binds to the pre-rRNA at positions close to the A_1 cleavage site and within the mature 18S rRNA sequence as well as to the U3 snoRNA (Wells *et al.*, 2016). Observed pre-rRNA binding sites included regions of the 18S rRNA and the 5' ETS that are known to base pair with U3 in formation of the central pseudoknot, as well as to the corresponding pre-rRNA base pairing sequences in the U3 snoRNA. These interactions point to a potential role for Utp24 in the coordination or verification of essential pre-rRNA-U3 interactions and pseudoknot formation.

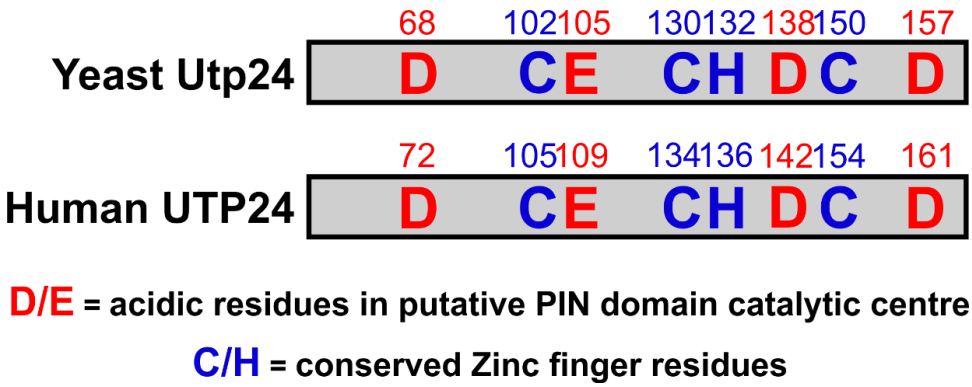


Figure 1.12 Yeast and human Utp24 contain a conserved PIN domain and Zinc finger motif. Cartoon representation of key residues in the PIN domain and Zinc finger motifs of yeast and human Utp24. Acidic residues in the putative PIN domain catalytic centre are shown in red. Conserved Zinc finger motif residues are shown in blue.

CRAC data also provided supporting evidence for a direct enzymatic role for Utp24 in cleavage at the A_1 site at the 5' end of 18S, as a crosslinking site close to this mapped cleavage site on the pre-rRNA was identified (Wells *et al.*, 2016). Consistent with this proposed function, study of the sequential assembly of the SSU processome indicated that Utp24 is first associated with the pre-ribosome after the transcription of the 5' end of the 18S rRNA (Zhang *et al.*, 2016b). Recently published structures of the SSU processome in a post-A_0, but pre-A_1 cleavage state identified Utp24 in a position close to the A_1 cleavage site, further supporting a role for its endonucleolytic cleavage at this site (Barandun *et al.*, 2017; Cheng *et al.*, 2017). In one of these studies, a model was proposed whereby a duplex between the pre-18S RNA and the U3 snoRNA must be unwound to allow Utp24 access to the A_1 cleavage site (Cheng *et al.*, 2017) (Figure 1.13). The endonucleolytic activity of Utp24, via its PIN domain, has also been confirmed *in vitro*, as WT Utp24, but not a mutant containing one or two substituted PIN domain acidic residues cleaved an RNA substrate containing the yeast A_2 site (Wells *et al.*, 2016). Together, these results argue for a catalytic role in endonuclease cleavage at sites A_1 and A_2 for Utp24, although another SSU processome component, Rcl1, was previously proposed as the site A_2 endonuclease (Horn *et al.*, 2011).

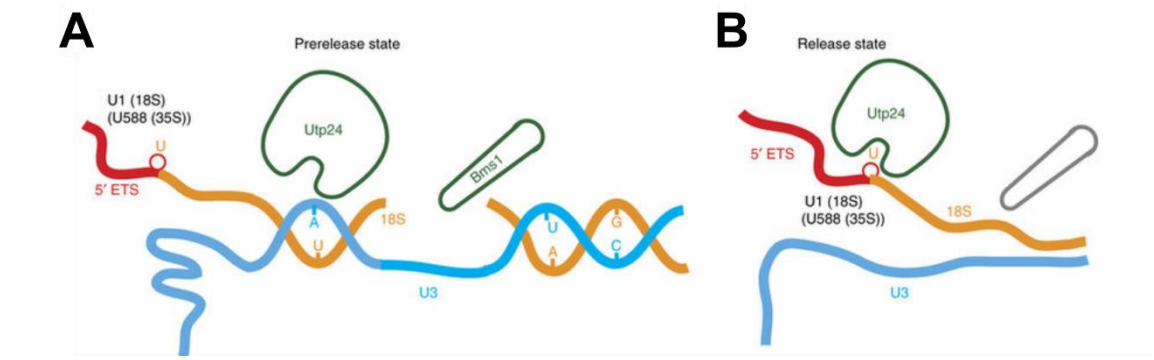


Figure 1.13 Proposed model for Utp24-mediated cleavage at yeast site A_1. (A) Schematic representation of the position of Utp24 in the yeast SSU processome as observed in the Cryo-EM structure in the pre-release state. (B) Proposed model for cleavage at site A_1 by Utp24. From (Cheng et al., 2017).

Compared to yeast, much less is known about UTP24 in mammals. In humans, UTP24 localises to the nucleolus and is a component of the SSU processome (Turner *et al.*, 2012). As such it is required for cleavages at human sites A0, 1 and 2a (Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013)**.** UTP24 is also required for the equivalent three cleavage events in mouse cells (Wang *et al.*, 2014). Like UTP23, UTP24 is not required for the initial 5' ETS cleavage at site A' in either human or mouse cells (Sloan *et al.*, 2014; Wang *et al.*, 2014). In mammals, as in yeast, UTP24 contains conserved putative catalytic residues in its PIN domain (Figure 1.12). A recent study, published during this PhD project, provided strong evidence of a role for UTP24 in cleavage at site 1 in human cells, and showed that an intact UTP24 PIN domain is required for this cleavage *in vivo* (Tomecki *et al.*, 2015).

1.7.2.1.2 Rcl1

A third, surprising, proposed endonuclease in the SSU processome is the RNA cyclase-like protein, Rcl1 (Billy *et al.*, 2000). RNA cyclases are RNA modifying enzymes, which catalyse the conversion of RNA 3' phosphates to 2', 3' cyclic phosphate ends (Genschik *et al.*, 1997). The biological role of RNA cyclases is unknown, but 2', 3' cyclic ends are required for RNA ligation by eukaryotic RNA ligases (Filipowicz and Shatkin, 1983). Cyclic ends are also found in tRNA splicing intermediates (Arn and Abelson, 1996) and the U6 spliceosomal small nuclear (sn)RNA (Lund and Dahlberg, 1992). Despite its similarity to this class of enzymes, Rcl1 does not possess a conserved RNA cyclase catalytic centre (Figure 1.14) (Billy *et al.*, 2000; Tanaka *et al.*, 2011). Rcl1 is essential for yeast growth, associates with

the U3 snoRNP and is required for cleavages at sites A_0, A_1 and A_2 in 18S rRNA maturation (Billy *et al.*, 2000; Horn *et al.*, 2011). Its binding partner, the GTPase Bms1, is also an SSU processome component and is required for the recruitment of Rcl1 to pre-ribosomes (Wegierski *et al.*, 2001; Dragon *et al.*, 2002; Grandi *et al.*, 2002; Karbstein *et al.*, 2005). Analyses of the stepwise co-transcriptional assembly of the SSU processome using truncated pre-rRNA transcripts suggest that Rcl1 assembles into pre-ribosomal particles with Bms1 at a late stage of SSU processome assembly, after the transcription of the 3' minor domain of 18S (Chaker-Margot *et al.*, 2015; Zhang *et al.*, 2016b). Recent SSU processome structures captured prior to A_1 cleavage indicate that Rcl1 is present in the SSU processome as a heterodimer with Bms1 (Barandun *et al.*, 2017).

Rcl1 has been proposed as the endonuclease responsible for cleavage at the A_2 site, as its depletion in yeast affects processing at sites A_0, A_1 and A_2, but particularly A_2, *in vivo* and recombinant Rcl1 was reported to cleave a pre-rRNA mimic containing the A_2 site *in vitro* (Horn *et al.*, 2011). An Rcl1 mutant protein, in which three residues in the C-terminal end of the Rcl1 sequence (R327, D328 and K330) were substituted for alanines, did not show endonuclease activity on this pre-rRNA substrate (Horn *et al.*, 2011). This mutant ("RDK"), which was proposed to be impaired in RNA-binding, as well as a truncation ending at residue 309 (which excludes the "RDK" motif), were not able to rescue the pre-rRNA processing phenotype caused by depletion of Rcl1 *in vivo*. Indeed, the pre-rRNA processing defect observed upon expression of these mutants was strikingly similar to that seen in the absence of Rcl1 (disruption of cleavages at sites A_0, A_1 and A_2, with disruption of A_2 cleavage more pronounced) (Horn *et al.*, 2011). This phenotype is consistent with impaired SSU processome assembly, as all three early cleavage events are affected. One of the residues of the "RDK" motif (R327) was previously shown to be essential for the interaction of Rcl1 with Bms1 which, as mentioned, is essential for the pre-ribosomal association of Rcl1 (Karbstein *et al.*, 2005; Delprato *et al.*, 2014). In addition, the Rcl1 R327A mutation disrupted the nuclear import of Rcl1 (Delprato *et al.*, 2014). These results suggest that, rather than disruption of any RNA substrate-binding or endonuclease activity possessed by Rcl1, the *in vivo* pre-rRNA processing phenotype observed upon expression of the "RDK" mutant is caused by the abolition of Rcl1's incorporation into the SSU processome, and thus impaired assembly of the SSU processome itself.

As previously mentioned, the PIN domain-containing endonuclease Utp24 is a strong candidate for cleavage at the A_2 site (Bleichert *et al.*, 2006; Wells *et al.*, 2016). While this may argue against a role for Rcl1 in cleavage at this site, it remains possible that the two proteins carry out this function in different contexts. For instance, A_2 cleavage can occur both co-transcriptionally or post-transcriptionally, it is therefore possible that both Utp24 and Rcl1 provide this function, depending on the timing of the processing. However, recent work in our lab, published during this PhD project, calls into question the proposal that Rcl1 exhibits endonuclease activity. When *in vitro* RNA cleavage analysis was performed under conditions amenable to cleavage at site A_2 by WT, but not PIN domain mutant forms of Utp24, Rcl1 did not cleave the pre-rRNA at any site (Wells *et al.*, 2016). Further cleavage assays, using conditions under which Rcl1 was reported to cleave RNA (Horn *et al.*, 2011), did indeed show cleavage activity for Rcl1, but also showed equal activity for both WT and catalytically inactive forms of recombinant Utp24 (Wells *et al.*, 2016). This suggests that the RNA cleavage observed under these conditions is not specific. Further experimental evidence of Rcl1's endonuclease activity, and the identification of an inactive mutant disrupting the catalytic site without affecting the interaction with Bms1 or the localisation of Rcl1, are required to categorise Rcl1 as a genuine candidate endonuclease in yeast pre-rRNA processing. As Rcl1 lacks an RNA cyclase catalytic centre (Tanaka *et al.*, 2011), an alternative catalytic centre must be identified within the protein before it can be considered a *bona fide* candidate endonuclease.

In humans, RCL1 is a stable component of the SSU processome and localises to the nucleolus (Turner *et al.*, 2012). Consistent with this, depletion of RCL1 in HeLa cells disrupted processing at cleavage sites A0, 1 and 2a (Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013). As for UTP23 and UTP24, the additional 5' ETS cleavage at site A' does not require RCL1 (Sloan *et al.*, 2014). The well characterised interaction between yeast Rcl1 and the GTPase Bms1 is conserved in both humans and zebrafish (Wang *et al.*, 2016). Despite this, human RCL1 appears to associate with pre-rRNA before A' cleavage and before the predicted recruitment of BMS1 to the pre-rRNA (Turner *et al.*, 2009; Turner *et al.*, 2012). This suggests that the importance of the yeast Rcl1-Bms1 interaction on the nuclear import and pre-ribosomal assembly of Rcl1 is not conserved in human cells. The "RDK" residues important for pre-rRNA processing and yeast Rcl1's interaction with Bms1 are not fully conserved in the human RCL1 protein, and the equivalent residues are instead "RHK" (Figure 1.14). It is unknown whether

these residues are important for the interaction between human RCL1 and BMS1, or for pre-rRNA processing in human cells.

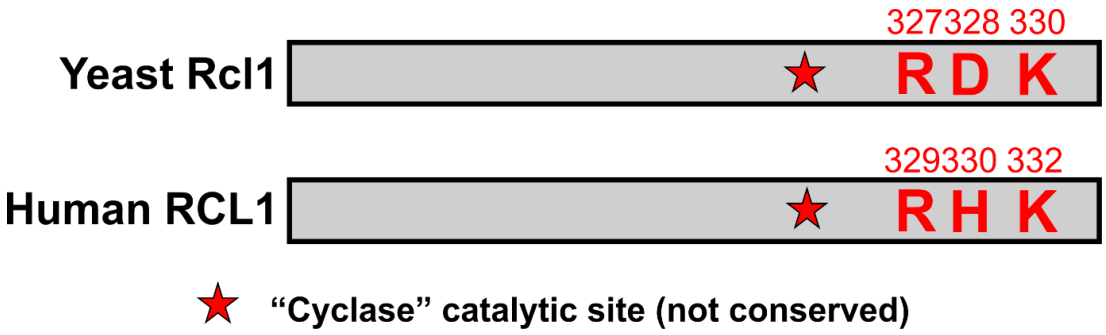
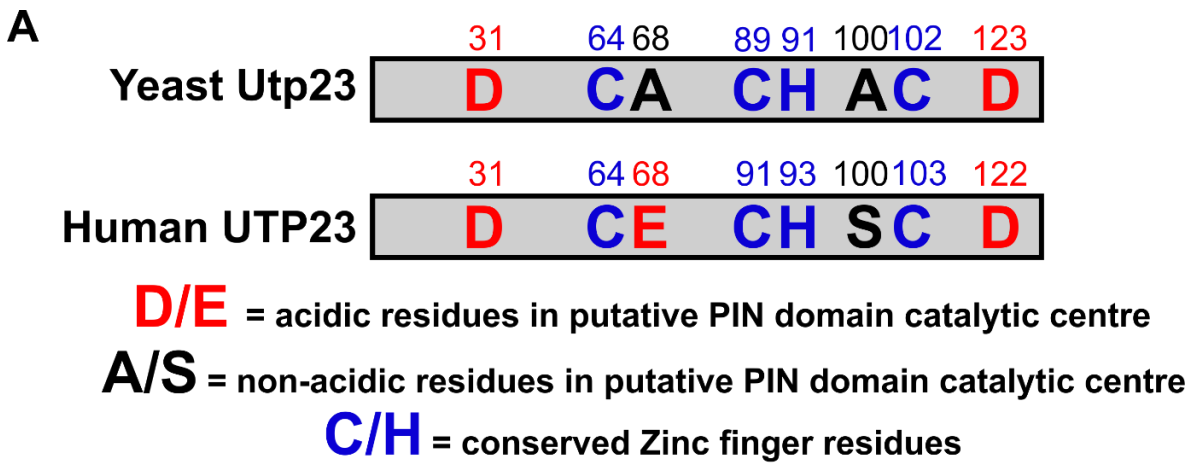


Figure 1.14 Comparison of yeast and human Rcl1 amino acid residues. Cartoon representation of proposed important residues in yeast and human Rcl1. Positions of "RDK" residues revealed to be essential for pre-rRNA processing *in vivo* in yeast (Horn *et al.*, 2011) and the equivalent "RHK" residues in human RCL1 are shown. The position of the non-conserved "cyclase" centre is marked with a star.

1.7.2.1.3 Utp23

Another SSU processome component, Utp23, also contains a PIN domain. Unlike active PIN-domain containing endonucleases, yeast Utp23 has only retained two of the four acidic residues in the active site (Figure 1.15). Accordingly, substitution of the remaining acidic residues in Utp23's PIN domain did not impair pre-rRNA processing *in vivo* and Utp23 mutants containing substitutions in these residues were able to compensate for Utp23 deletion (Bleichert *et al.*, 2006; Lu *et al.*, 2013). Despite not being an endonuclease, as an essential component of the yeast SSU processome, the presence of the Utp23 protein is required for cleavages at sites A_0, A_1 and A_2, suggesting it plays an important, non-enzymatic role in pre-rRNA processing (Bleichert *et al.*, 2006). Within the PIN domain of Utp23, it also contains a conserved Zn finger RNA binding motif (Figure 1.15). Mutation of conserved Zn finger residues severely affected yeast growth, suggesting that the essential function of Utp23 in the SSU processome involves binding to either pre-rRNA or snoRNAs (Lu *et al.*, 2013).



B

Utp23 CLUSTAL O (1.2.4) multiple sequence alignment

```
S. pombe              1 MRQKRAKNYRKLMHTYQLLFGFREPYQVLVDADFLKDLSQQKIDIQA-ALARTVQGAIKP
S. cerevisiae         1 MRQKRAKSYRKQLLVYSHTFKFREPYQVLVDNQLVLECNNSNFNLPS-GLKRTLQADVKV
C. albicans           1 MRQKRAKAYKKQMSVYVHAFKFREPYQIIVDNELITTCQSASFDINK-GFTRTIQAENKP
A. thaliana           1 MRVKRQKKNRRTVRFFTVCYGFRQPYKVLCDGTFVHHLVTNEITPADTAVSELLGGPVKL
D. melanogaster       1 MKISRFKKSHKTLVFFASNFDYREPYQVLIDATFCQAALQQKIGIDE-QIKKYFQCGVKL
X. laevis             1 MKIKRQKHAKKYLTFYRFNFGVREPYQVLLDGTFCQAALKNKIQIKE-QLPKYLMGEVQL
H. sapiens            1 MKITRQKHAKKHLGFFRNNFGVREPYQILLDGTFCQAALRGRIQLRE-QLPRYLMGETQL
M. musculus           1 MKITRQKHAKKHLGFFRNNFGVREPYQILLDGTFCQAALRGRIQLRD-QLPRYLMGETQL
R. norvegicus         1 MKITRQKHAKKHLGFFRNNFGVREPYQILLDGTFCQAALRGRIQLRD-QLPRYLMGETQL

S. pombe             60 MITQCCIRQLYSKSDELKQEIRIAKSFERRRCGHI-DEALSPSECIQSVVNINGRNKHRY
S. cerevisiae        60 MITQCCIQALYETRND--GAINLAKQFERRRCNHSFKDPKSPAECIESVVNISGANKHRY
C. albicans          60 MITQCCIQALYDTKNQ--PAIDIAKSFERRKCNHR--EAIDPSQCIESIVNIKGQNKHRY
A. thaliana          61 FTTRCVIAELEKLGKDFAESLEAAQTLNTATCEHE-EA-KTADECLSEVIG--VQNTEHF
D. melanogaster      60 LTTQCVILESESLGAPLTGATSIVKRFHVHKCGHE-GKPVPASECIKSMTK--DN---RY
X. laevis            60 CTSHCVIKELESLGKELYGAKLIAQRFQVRSCSHF-QDPVSGSACILSLTA--DNNPHHY
H. sapiens           60 CTTRCVLKELETLGKDLYGAKLIAQKCQVRNCPHF-KNAVSGSECLLSMVE--EGNPHHY
M. musculus          60 CTTRCVLKELETLGKELYGAKLIAQKCQVRNCPHF-KSPVSGSECLLSMVD--EGNPHHY
R. norvegicus        60 CTTRCVLKELETLGKELYGAKLIAQKCQVRNCPHF-KSAVSGSECLLSMVD--DGNPHHY

S. pombe            119 VVATQDPELRQALRSVPGVPLIYMKRSVVILEPASRATLLEKHNKESVQMGMSKEEKLLL
S. cerevisiae       118 VVASQDIDLRRKLRTVPGVPLIHLTRSVMVMEPLSTASAKASKI---------TEEQKLY
C. albicans         116 IVASQDLQLRKKLRKIPGVPLIYMNRSVMVMEPISDVSNQYNMN---------YESKKLT
A. thaliana         117 FLGTQDAEFRRKLQQESIVPLVFGLRNILLIDQPSDFQRQSAKDSENKRLTMTDTEKKLL
D. melanogaster     114 VVASQDRLLQESLRKIPGRCLLYLHKATPVLEAPSKASKKWVQRRAKN-LMLGKQV----
X. laevis           117 FIATQDQELATKVKKRAGVPLMFIIQNTIVLDKPSPKSLARVEAVQSNQLVPEHQK----
H. sapiens          117 FVATQDQNLSVKVKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKAVESGQLVSVHEK----
M. musculus         117 FVATQDQNLSVKVKRTPGIPLMFIIQNTIVLDKPSPRTVAFVKAVEAGQLVSVHEK----
R. norvegicus       117 FVATQDQNLSVKVKKNPGIPLMFIIQNTIVLDKPSPRTVAFVKAVEAGQLVSVHEK----
```

Figure 1.15 Conservation of key PIN domain and Zinc finger motif residues in eukaryotic Utp23. (A) Cartoon representation of key residues in the PIN domain and Zinc finger motif of yeast and human Utp23. Conserved acidic residues that form the proposed PIN domain active site are shown in red and non-acidic residues in these positions are shown in black. Conserved Zinc finger motif residues are shown in blue. (B) Alignment of selected eukaryotic protein sequences retrieved by a NCBI protein BLAST search using the S. cerevisiae Utp23 (YOR004W) protein sequence. Retrieved sequences were aligned using ClustalO (http://www.ebi.ac.uk/Tools/msa/clustalo/). Positions of key residues in the PIN domain predicted catalytic centre are highlighted in yellow, with conserved acidic residues shown in red and non-acidic residues shown in black. Positions of conserved Zinc finger motif residues are highlighted in blue.

Pre-rRNA processing at sites A_0, A_1 and A_2 requires the H/ACA snoRNP snR30 (U17 in humans) (Morrissey and Tollervey, 1993). Yeast Utp23 associates with snR30 *in vivo*, and Utp23 is required for snR30 release from pre-ribosomes (Hoareau-Aveilla *et al.*, 2012). snR30 is also required for the stable incorporation of Utp23 into the yeast SSU processome. Utp23 may function in establishing the base pairing interaction between the snR30 snoRNA and the ES6 region of 18S rRNA (Fayet-Lebaron *et al.*, 2009; Hoareau-Aveilla *et al.*, 2012). In the sequential assembly of the SSU processome, Utp23 associates at the same time as snR30, including the box H/ACA core proteins, snR30-interacting factors and ES6-interacting factors (Lin *et al.*, 2013; Chaker-Margot *et al.*, 2015; Zhang *et al.*, 2016b). In recently published structures of the yeast SSU processome, captured in a state following A_0 cleavage but before cleavage at A_1, Utp23 was not detected (Barandun *et al.*, 2017; Cheng *et al.*, 2017). This suggests that Utp23 is released from the pre-ribosomal particle after cleavage at A_0. Our lab has recently published data indicating that yeast Utp23 directly contacts both the 18S rRNA (in the ES6 region) and snR30 *in vivo*, and the binding sites were close to sites of base pairing interaction between the two RNAs (Wells *et al.*, 2017). Utp23 also bound to ES6- and snR30-interacting proteins *in vitro,* including the endonuclease Utp24, suggesting it may play a role in coordinating protein-protein interactions in this region and/or in recruitment of Utp24 to pre-rRNA cleavage sites (Wells *et al.*, 2017).

UTP23 is less well characterised in higher eukaryotes, but it appears to play a conserved role in 18S rRNA maturation. In humans, UTP23 is a component of the SSU processome, localises to the nucleolus, and is essential for cleavage at sites A0, 1 and 2a (Turner *et al.*, 2012; Sloan *et al.*, 2014). UTP23 is not required for the earliest cleavage in the 5' ETS, at site A', emphasising that this cleavage event is independent of, and not required for, subsequent pre-18S rRNA cleavages (Sloan *et al.*, 2014). The interaction between Utp23 and snR30 is also conserved in humans, as UTP23 associates with the human homologue, U17, *in vivo* (Hoareau-Aveilla *et al.*, 2012). This suggests a potentially conserved role for human UTP23 as a non-enzymatic pre-40S ribosome biogenesis factor. However, unlike its yeast counterpart, human UTP23 contains three conserved acidic residues in the active site of its PIN domain (Figure 1.15). The presence of three acidic active site residues has been shown to be sufficient for the endonuclease activity of other PIN domain-containing proteins (Glavan *et al.*, 2006). This raises the exciting possibility that UTP23 may play a distinct, or additional

role as an active endonuclease in human pre-rRNA processing. However, the effect of mutating the active site residues in UTP23's PIN domain has not previously been analysed. Surprisingly, in mouse cells, UTP23 appears to play a different role in 18S rRNA maturation compared to yeast and humans, as depletion of UTP23 disrupted a later processing step, resulting in accumulation of the 18SE pre-rRNA (Wang *et al.*, 2014). Defects in human ribosome biogenesis are linked to disease and cancer predisposition. Human UTP23 was recently found to be mutated in colorectal cancer, although it is unknown if or how this mutation, which is outside of the PIN domain, affects the function of UTP23 in ribosome biogenesis (Timofeeva *et al.*, 2015).

1.7.2.2 Helicases

RNA helicases are enzymes initially defined as proteins capable of ATP-dependent unwinding of RNA duplexes but were later shown to have a broader range of functions, such as displacement of RNA-bound proteins (Jankowsky *et al.*, 2001; Fairman *et al.*, 2004) and annealing of RNA strands (Yang and Jankowsky, 2005). Helicases function in all aspects of RNA metabolism, but the majority have roles in ribosome biogenesis, perhaps unsurprisingly, due to the abundance of RNP complexes involved. RNA helicases are important for the maturation of both SSU and LSU, likely through roles in annealing or unwinding of pre-rRNA-snoRNA base pairing, removal of protein factors from pre-rRNAs or snoRNAs or mediating conformational changes in pre-ribosomal particles (Bleichert and Baserga, 2007; Rodriguez-Galan *et al.*, 2013). Most RNA helicases involved in ribosome biogenesis are DEAD/H box proteins, which contain multiple conserved motifs required for ATP hydrolysis (Linder and Jankowsky, 2011). In yeast, several helicases are required for early cleavages at sites A_0-A_2 in 18S rRNA maturation (Granneman *et al.*, 2006), including Rok1 (Venema *et al.*, 1997), Dhr1 (Colley *et al.*, 2000)**,** Dbp4 (Kos and Tollervey, 2005), Dbp8 (Daugeron and Linder, 2001), Fal1 (Kressler *et al.*, 1997) and Rrp3 (O'Day *et al.*, 1996) and are categorised as SSU processome components (Phipps *et al.*, 2011).

Multiple RNA helicases in the SSU processome are implicated in the removal of snoRNAs from base pairing interactions with the pre-rRNA. These include the unwinding of the U3-pre-18S RNA duplex by the DEAH box helicase Dhr1 for subsequent formation of the central pseudoknot (Colley *et al.*, 2000; Sardana *et al.*, 2015; Zhu *et al.*, 2016), and the requirement for Rok1 in release of snR30 (Bohnsack *et al.*, 2008). Rok1 is a DEAD box helicase that interacts with the RNA binding protein

Rrp5, and has helicase activity *in vitro* (Torchet *et al.*, 1998; Vos *et al.*, 2004a; Garcia *et al.*, 2012).

Rrp5 is essential for the association of Rok1 with pre-ribosomes (Vos *et al.*, 2004a), and they associate with assembling pre-ribosomal particles at around the same time as each other, and as snR30 (Chaker-Margot *et al.*, 2015; Zhang *et al.*, 2016b). Mutations in the conserved motifs of Rok1's helicase domain caused accumulation of snR30 in large pre-ribosomal complexes, suggesting its helicase function is important for snR30 release (Bohnsack *et al.*, 2008) and the interaction with Rrp5 provides substrate specificity *in vitro* (Young *et al.*, 2013). Together, these results suggest that Rok1 is responsible for unwinding the snR30-pre-18S duplex, with Rrp5 acting as a cofactor. However, a recent report indicates that Rok1 is required for the ATP-dependent release of Rrp5 from the pre-ribosome, and disruption of the interaction between Rok1 and Rrp5 causes snR30 accumulation in large particles, despite Rok1 being catalytically active (Khoshnevis *et al.*, 2016). It was therefore suggested that the observed snR30 accumulation upon Rok1 depletion is an indirect effect, caused by the inhibition of Rrp5 release, which was shown to be required for snR30 release (Khoshnevis *et al.*, 2016). It was also shown that Rrp5 interacts with the helicase Has1, which may be responsible for mediating snR30 release (Liang and Fournier, 2006; Khoshnevis *et al.*, 2016). Consistent with this, it has been shown in one study that Rok1 catalyses ATP-independent annealing of RNA duplexes, and it was also suggested that it does not have efficient unwinding activity *in vitro* (Young *et al.*, 2013). However, previous data has shown that Rok1 does indeed possess *in vitro* helicase activity (Garcia *et al.*, 2012).

In humans, there are around twice as many RNA helicases compared to yeast (Umate *et al.*, 2010), and around 40 are thought to be nucleolar (Ahmad *et al.*, 2009), although they are less well characterised than their yeast counterparts. Some helicases involved in ribosome biogenesis are conserved between yeast and humans. These include the human homologue of the SSU processome helicase Fal1 (Kressler *et al.*, 1997), DDX48, which is important for pre-18S cleavages in human cells (Alexandrov *et al.*, 2011). Human DBP4/DDX10, DHR1/DHX37 and HAS1/DDX18, like their yeast homologues, are SSU processome components and associate with U3, while DBP4 is also associated with a 50S SSU processome assembly intermediate when RNA polymerase I transcription is inhibited (Turner *et al.*, 2009). Rok1 (DDX52/ROK1 in humans) is also conserved in humans (Schutz *et al.*, 2010).

1.7.2.3 RNA binding proteins

As expected of a large RNP complex, the SSU processome contains many proteins with RNA binding domains. One such protein is Krr1 (Sasaki *et al.*, 2000), which contains a conserved KH RNA binding domain (Zheng *et al.*, 2014) and associates with the pre-ribosome around the same time as ES6-interacting factors Rok1, Rrp5, and Utp23 as well as snR30 (Chaker-Margot *et al.*, 2015)**.** Rrp5 is a unique SSU processome component, in that it also participates in maturation of the LSU (Venema and Tollervey, 1996; de Boer *et al.*, 2006), and contains multiple S1 RNA binding motifs (Torchet *et al.*, 1998).

1.7.2.3.1 Rrp5

Rrp5 is a large (193 kDa in yeast) multidomain RNA binding protein. It is essential for cleavage at sites A_0, A_1 and A_2 in 18S rRNA maturation, and site A_3 in LSU maturation (Venema and Tollervey, 1996). The protein contains 12 S1 RNA binding domains and seven C-terminal TPR (Tetratricopeptide repeat) protein-protein interaction motifs (Figure 1.16) (Torchet *et al.*, 1998; Young and Karbstein, 2011). Nine S1 domains make up the N-terminal domain (NTD) of Rrp5, which is required for the post-transcriptional A_3 cleavage, while the C-terminal domain (CTD), comprising the remaining three S1 domains and the six C-terminal TPR domains, is required for the mainly co-transcriptional SSU cleavages at sites A_0-A_2 (Figure 1.16A) (Eppens *et al.*, 1999). As well as being functionally distinct, these two halves can be physically separated, and expression of both halves independently is sufficient for the full activity of Rrp5 in pre-rRNA processing (Eppens *et al.*, 1999). The three S1 domains of the CTD (S1 10-12) are specifically required for cleavage at site A_2 (Figure 1.16A) (Vos *et al.*, 2004b). Analysis of *in vivo* RNA-protein crosslinking (CRAC) using N- or C-terminal halves of Rrp5 showed that the CTD binds the pre-rRNA within ITS1, near the A_2 cleavage site (Lebaron *et al.*, 2013). The CTD also crosslinked to all four snoRNAs involved in SSU pre-rRNA processing, U3, U14, snR30 and snR10 and the ES6 region of the 18S rRNA sequence of the pre-rRNA (Lebaron *et al.*, 2013). The NTD, consistent with its requirement for LSU pre-rRNA processing, crosslinked to ITS1 sequences close to the A_3 cleavage site, as well as the RNA component of RNase MRP (Lebaron *et al.*, 2013). Consistent with its known requirement for both ITS1 cleavages, Rrp5 has been proposed as a factor involved in the regulation of the alternative ITS1 cleavages under different growth conditions (Kos-Braun *et al.*, 2017).

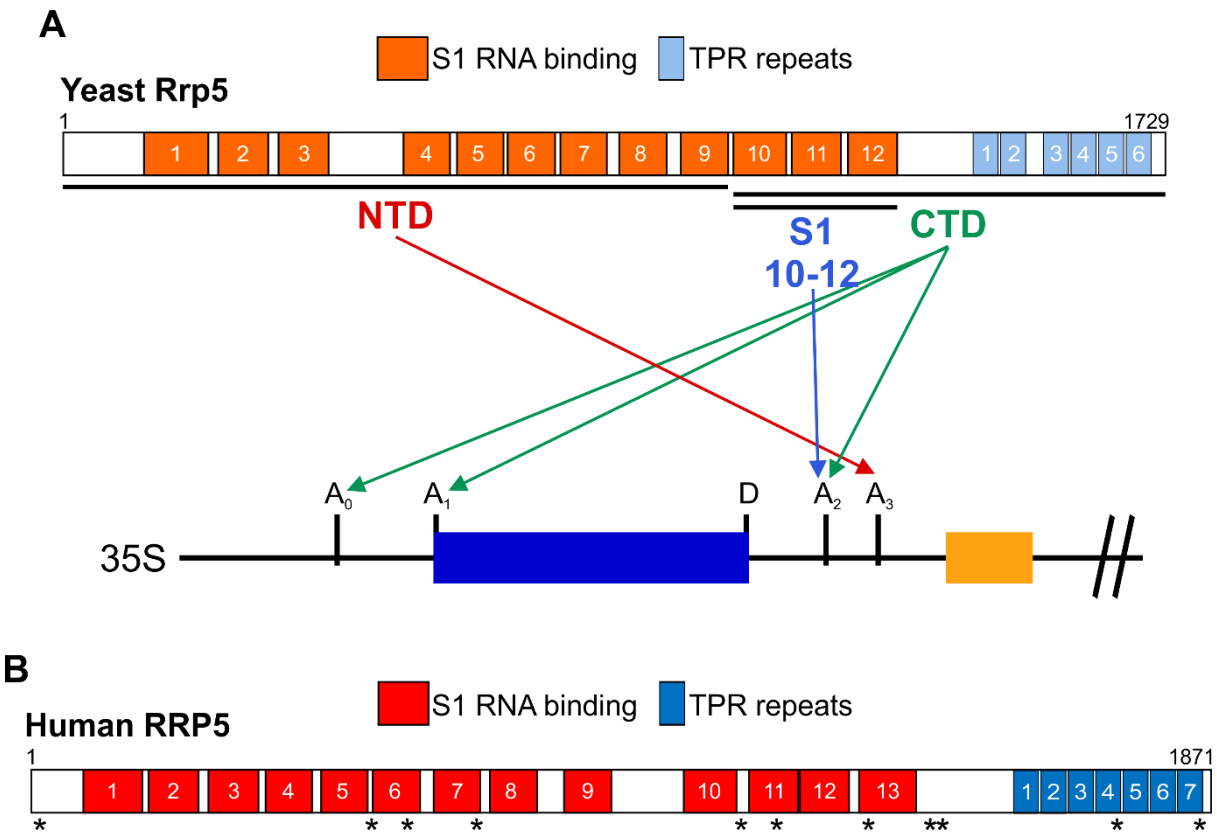


Figure 1.16 The domains of yeast and human Rrp5 proteins. (A) Cartoon depiction of the domain architecture of yeast Rrp5, with S1 RNA binding domains shown in orange and TPR repeats shown in light blue. The functionally independent NTD and CTD domains are marked and arrows indicate the requirement of these domains in processing at different cleavage sites. **(B)** Cartoon representation of human RRP5 domain architecture. S1 RNA binding domains are shown in red and TPR repeats are shown in blue. Positions of cancer-associated mutations (Iorio *et al.*, 2016) are marked with asterisks.

Similarly to depletion of the U3 snoRNA (Osheim *et al.*, 2004)**,** Rrp5 depletion results in the loss of 5' terminal knobs on nascent pre-rRNA upon visualisation of chromatin spreads (Lebaron *et al.*, 2013), indicating that Rrp5 is important for SSU processome assembly. In the sequential assembly of the SSU processome, Rrp5 is required for the association of the UtpC subcomplex (Perez-Fernandez *et al.*, 2007). Consistently, Rrp5 first associates with pre-ribosomes at around the same point as the UtpC proteins, Utp22 and Rrp7, on a pre-rRNA containing the ES6 region in the central domain of the 18S rRNA sequence (Chaker-Margot *et al.*, 2015; Zhang *et al.*, 2016b). In the structure of the yeast SSU processome, the CTD, containing the TPR repeats, of Rrp5 is positioned close to both the UtpC complex and the 18S rRNA central domain (Barandun *et al.*, 2017), consistent with CRAC data showing Rrp5 crosslinking to the ES6 region of 18S (Lebaron *et al.*, 2013). Other factors associating with this region, as

well as associating with the pre-ribosome at around the same time as Rrp5, include the snR30 snoRNA, the PIN domain protein Utp23 (Wells *et al.*, 2017) and the DEAD box helicase Rok1 (Martin *et al.*, 2014)**.** Rok1 is important for the release of snR30 from pre-ribosomes, and Rrp5 is implicated as a co-factor in this process, providing substrate specificity to the activity of Rok1 (Torchet *et al.*, 1998; Bohnsack *et al.*, 2008; Young *et al.*, 2013; Martin *et al.*, 2014)**.** However, a recent study suggests that Rok1 activity is not responsible for the release of snR30, and instead mediates the release of Rrp5 from pre-ribosomes (Khoshnevis *et al.*, 2016). Rrp5 also interacts with another helicase, Has1, so the snR30 accumulation observed upon Rok1 depletion may be caused by disruption of Rrp5's role as a cofactor in Has1-mediated snR30 release (Liang and Fournier, 2006; Khoshnevis *et al.*, 2016).

In addition to functioning in SSU pre-rRNA processing via its CTD, Rrp5 functions in LSU rRNA maturation through the nine S1 domains in its NTD (Eppens *et al.*, 1999; Lebaron *et al.*, 2013). Rrp5 forms a complex with LSU biogenesis factors, Noc1 and Noc2, and its interaction with Noc1 requires the NTD, consistent with the requirement for this region in LSU processing (Hierlmeier *et al.*, 2013). This complex appears to be co-transcriptionally recruited to pre-ribosomes after cleavage at A_2 (Hierlmeier *et al.*, 2013).

Less is known about RRP5 in human cells, but its function in pre-rRNA processing appears to be conserved, as it is required for cleavages at sites A0, 1 and 2 (yeast A_3) (Sweet *et al.*, 2008). As in yeast, RRP5 is also important for cleavage at site 2a (yeast A_2) in the minor pre-rRNA processing pathway of human cells, but not the initial 5' ETS cleavage at site A' (Sloan *et al.*, 2013b). RRP5 is a component of the human SSU processome and associates with the U3 snoRNA in a 50S processome intermediate when rDNA transcription by RNA polymerase I is blocked (Turner *et al.*, 2009). While the general functions of RRP5 in ribosome biogenesis appear to be conserved, it is currently unknown if or how these functions are separated between the 13 S1 RNA binding domains and seven TPR repeats in the 209 kDa protein (Figure 1.16B). RRP5 was found to be frequently mutated in cancer in a recent screen of tumours and cancer cell lines (Iorio *et al.*, 2016).

1.8 Exoribonucleases in Pre-rRNA Processing

While endonucleases play an important role in the processing of the pre-rRNA at multiple stages of ribosome biogenesis, trimming of pre-rRNA intermediates or degradation of spacer sequences or aberrant precursor RNAs via exonucleolytic digestion is also required. Following endonucleolytic cleavage at multiple sites on the yeast and mammalian pre-rRNA, the cleaved transcript is processed exonucleolytically in the 5' or 3' direction. In yeast, processing events involving exonucleases include the maturation of the mature 5.8S rRNA 5' end following cleavage at A_3 in ITS1, the maturation of the 5.8S 3' end and the 25S 5' end after C_2 cleavage in ITS2 (Schillewaert *et al.*, 2012; Pillon *et al.*, 2017), and 25S 3' end maturation after B_0 cleavage in the 3' ETS. Processing of 3' and 5' ends of LSU rRNA sequences is conserved in mammals, with XRN2 (Rat1) mediating maturation of the 5' end of 28S (Wang and Pestov, 2011) and the exosome complex processing the 3' end of 5.8S (Schilders *et al.*, 2007; Preti *et al.*, 2013; Tafforeau *et al.*, 2013).

In humans, exonucleolytic processing is also required for 18S rRNA maturation in the "major" processing pathway of 18SE pre-rRNA production after cleavage at site 2 in ITS1 (Sloan *et al.*, 2013b), and may also play a role in an alternative processing mechanism of 18S 3' end maturation. In addition to processing pre-rRNA in rRNA maturation, exonucleases also function in the degradation of aberrant pre-rRNAs and excised pre-rRNA fragments produced by multiple endonucleolytic cleavages.

1.8.1 Rat1/XRN2

The Rat1 (XRN2 in mammals) 5' to 3' exonuclease functions in processing and degradation of RNA substrates in a broad range of biological processes in the nucleus (Amberg *et al.*, 1992; Kenna *et al.*, 1993). Rat1 shares similar structural features and sequence with another 5' to 3' exonuclease, Xrn1, which performs similar functions in the cytoplasm (Johnson, 1997; Nagarajan *et al.*, 2013). Both Rat1 and Xrn1 possess 5' to 3' exoribonucleolytic activity, which enables processing or degradation of single stranded RNA substrates, which have a monophosphate at the 5' end, and which requires Mg^{2+} or Mn^{2+} ions (Stevens and Poole, 1995; Poole and Stevens, 1997). Xrn1 functions mostly in the degradation of cytoplasmic mRNAs but can also compensate for the loss of Rat1 in some aspects of nuclear RNA processing (Stevens *et al.*, 1991; Petfalski *et al.*, 1998). The N-terminal regions of these two proteins, which contain the

catalytic site, is highly conserved (Page *et al.*, 1998; Solinger *et al.*, 1999), while the C-terminal region of Rat1 is specifically required for its interaction with its cofactor, Rai1 (Xue *et al.*, 2000; Xiang *et al.*, 2009). Rai1 is important for the activation of Rat1 exonuclease activity. Rai1 possesses pyrophosphatase activity and converts RNA 5' triphosphates to 5' monophosphates (Poole and Stevens, 1997; Xiang *et al.*, 2009).

The main role of Rat1 outside of ribosome biogenesis is in the degradation and processing of nuclear RNAs, but it also functions in the termination of transcription of mRNAs by RNA polymerase II (Kim *et al.*, 2004; Luo *et al.*, 2006). Transcription termination is important for recycling of the polymerase machinery for further transcription and preventing aberrant transcription of downstream genes (Rosonina *et al.*, 2006). Cleavage at the 3' end of a gene by the polyadenylation machinery provides a free 5' end for Rat1 to degrade the downstream transcript and trigger transcription termination (Connelly and Manley, 1988; Kim *et al.*, 2004). Rat1 also has a role in termination of RNA polymerase I transcription after cleavage at site B_0 in the 3' ETS by Rnt1 (El Hage *et al.*, 2008; Kawauchi *et al.*, 2008).

In ribosome biogenesis, Rat1, in conjunction with Rai1, participates in the degradation of excised pre-rRNA spacer sequences following endonuclease cleavages in the 5' ETS (at sites A_0 and A_1) and ITS1 (A_2 and A_3) (Petfalski *et al.*, 1998; Fang *et al.*, 2005). Rat1 may also be involved in degradation of the excised fragment between sites D (at the 3' end of 18S) and A_2, in the absence of Xrn1, which is responsible for this degradation (Petfalski *et al.*, 1998). Rat1 and Rai1 also function in the maturation of snoRNA 5' ends, including that of U14 (Petfalski *et al.*, 1998; Lee and Baserga, 1999; Qu *et al.*, 2011). In pre-rRNA processing, Rat1 is responsible for maturation of the 5' end of the 25S rRNA of the LSU following endonucleolytic cleavage at site C_2 in ITS2 by the endonuclease Las1 (Geerlings *et al.*, 2000; Gasse *et al.*, 2015; Fromm *et al.*, 2017). In 5.8S rRNA maturation, Rat1 functions, along with another 5' to 3' exonuclease, Rrp17, in trimming of ITS1 after cleavage at A_3 to site B_{1S} producing the $5.8S_S$ short form of 5.8S (Henry *et al.*, 1994; El Hage *et al.*, 2008; Oeffinger *et al.*, 2009). In the absence of Rat1, cytoplasmic Xrn1 is able to degrade 5' extended forms of 5.8S but is not involved in its maturation (El Hage *et al.*, 2008). In CRAC experiments, Rat1 crosslinked to pre-rRNA sequences in the 5' ETS, ITS1 and ITS2, and is thought to be recruited to the pre-rRNA prior to endonuclease cleavages (Granneman *et al.*, 2011).

In mammals, XRN2 appears to have many roles. Like Rat1 in yeast, human XRN2 is implicated in RNA polymerase II termination (West *et al.*, 2004; Brannan *et al.*, 2012; Fong *et al.*, 2015). While Rat1 activity requires its interaction with Rai1, XRN2 does not bind to the human homolog of Rai1, DOM3Z (Xiang *et al.*, 2009). In *C. elegans* XRN2 forms a complex with an unrelated protein, PAXT-1 via a conserved XRN2-binding domain (XTBD), and this interaction is important for the stability of XRN2 (Miki *et al.*, 2014). In mammals, several proteins have a conserved XTBD, including human C2AIL which, despite a lack of sequence similarity, is able to compensate for loss of *C. elegans* PAXT-1 *in vivo* (Richter *et al.*, 2016). In both human and mouse cells, XRN2 is responsible for the degradation of excised pre-rRNA fragments after endonucleolytic cleavages in the 5' ETS (at sites A0 and 1) and in ITS1 (at sites 2a and 2) (Wang and Pestov, 2011; Preti *et al.*, 2013; Sloan *et al.*, 2013b). XRN2 has also been shown to be important for the maturation of the 5' end of the 5.8S rRNA after site 2 cleavage (Wang and Pestov, 2011; Wang *et al.*, 2015). While the predominant ITS1 cleavage in yeast occurs at site A_2, the major cleavage in mammalian ITS1 processing occurs at site 2, analogous to yeast site A_3 (Wang and Pestov, 2011; Mullineux and Lafontaine, 2012; Sloan *et al.*, 2013b). The 36S precursor, produced by cleavage at 2a prior to site 2 cleavage and not detected under normal conditions, accumulates upon depletion of XRN2, indicating increased use of site 2a as the initial ITS1 cleavage site. Despite the accumulation of this precursor, production of the mature 18S rRNA is not affected, suggesting that the "minor" pathway featuring prominent cleavage at 2a is efficient, and a role for XRN2 in the balance of the two competing pathways (Sloan *et al.*, 2013b).

In higher eukaryotes, an additional endonuclease cleavage occurs in the 5' ETS at site A', and XRN2 is important for this cleavage event as XRN2 depletion results in accumulation of a 5'-extended 30S precursor not cleaved at site A', as well as for degradation of the excised spacer produced by A' cleavage (Wang and Pestov, 2011; Sloan *et al.*, 2014). It was recently shown that human XRN2 is a component of a complex containing NFkB-repressing factor (NKRF) and the human homolog of the RNA helicase Prp43, DHX15, and that this complex is required for efficient A' cleavage (Memet *et al.*, 2017). NKRF contains a conserved XTBD (Richter *et al.*, 2016), and is required for the recruitment of XRN2 to pre-ribosomal particles and for subsequent turnover of excised 5' ETS pre-rRNA fragments (Memet *et al.*, 2017). Under stress conditions, NKRF plays a role in the regulation of XRN2-mediated pre-rRNA

processing and degradation (Coccia *et al.*, 2017). In the proposed model, heat stress induces translocation of NKRF and XRN2 out of the nucleolus into the nucleoplasm, consequently altering pre-rRNA processing (Coccia *et al.*, 2017). XRN2 depletion induces pre-rRNA processing via the minor ITS1 processing pathway (Sloan *et al.*, 2013b), and accumulation of XRN2 in the nucleoplasm likely has a similar effect, which is reminiscent of the observed switch in yeast pre-rRNA processing pathways under stress conditions (Kos-Braun *et al.*, 2017). In a recent high-throughput screen of tumours and cancer cell lines, 88 distinct XRN2 mutations were identified (Figure 1.17A) (Iorio *et al.*, 2016).

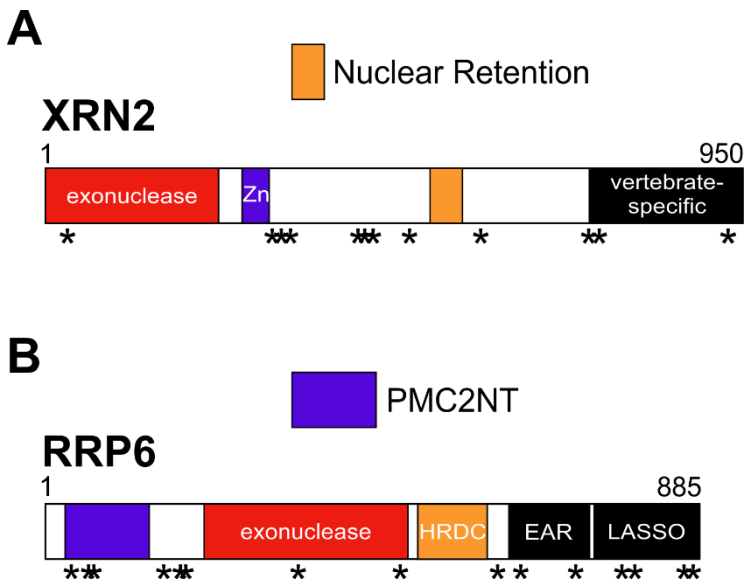


Figure 1.17 The domains of human exoribonucleases XRN2 and RRP6. Cartoon representation of the domain architecture of **(A)** XRN2 and **(B)** RRP6. Positions of cancer-associated mutations (Iorio *et al.*, 2016) are marked with asterisks.

1.8.2 The Exosome Complex and Rrp6

The RNA exosome is a highly conserved multiprotein complex with 3' to 5' exonuclease and endonuclease activities (Mitchell *et al.*, 1997; Lebreton *et al.*, 2008; Schaeffer *et al.*, 2009; Schneider *et al.*, 2009; Januszyk and Lima, 2014). The exosome is responsible for the processing and/or degradation of virtually all classes of RNA, in both the nucleus and the cytoplasm. Cytoplasmic functions of the exosome include the quality control of mRNAs, in mRNA decay pathways such as nonsense-mediated decay (NMD) (Isken and Maquat, 2007). In the nucleus, the exosome is important for 3' end modification of snoRNAs, as well as RNA surveillance, through degradation of

aberrant non-coding RNAs, including snoRNAs, rRNAs or tRNAs that have been misprocessed or misassembled into RNPs (Schneider *et al.*, 2012; Chlebowski *et al.*, 2013). In pre-rRNA processing, the exosome is important for the turnover of excised spacer fragments produced by endonucleolytic cleavage of pre-rRNA intermediates (de la Cruz *et al.*, 1998). In addition, the exonucleolytic activity of the complex has a direct role in LSU rRNA maturation (Mitchell *et al.*, 1996; Allmang *et al.*, 1999).

The eukaryotic core exosome is comprised of nine proteins, with six components (Rrp41, Rrp45, Rrp46, Rrp43, Mtr3 and Rrp42), homologous to RNase PH, forming a hexameric ring, with a channel in the centre, which can accommodate a single stranded RNA substrate (Liu *et al.*, 2006). Three additional proteins (Csl4, Rrp4 and Rrp40) with S1 and/or KH RNA binding domains form the top of the ring (the “cap”), with S1 domains projected towards the channel (Makino *et al.*, 2013). Although homologous to RNase PH, the six ring components lack exonuclease activity, and therefore the catalytic activity of the complex is provided by two additional proteins, Rrp44 (DIS3 in humans) and Rrp6, which associate with the core exosome (Dziembowski *et al.*, 2007; Schneider *et al.*, 2007).

Rrp44 is constitutively associated with both nuclear and cytoplasmic exosome complexes and is a member of the RNase II family of 3’ to 5’ exonucleases (Vincent and Deutscher, 2006). In addition to its RNB domain, which harbours Rrp44’s exoribonuclease activity (Dziembowski *et al.*, 2007; Schneider *et al.*, 2007) and S1 domain important for substrate binding (Schneider *et al.*, 2007), Rrp44 contains a PIN domain, providing endonuclease activity (Lebreton *et al.*, 2008; Schaeffer *et al.*, 2009; Schneider *et al.*, 2009). Structural analysis of the exosome complex showed that Rrp44 is tethered to the bottom of the exosome core via its PIN domain and that an RNA substrate can pass through the central channel of the exosome to reach the exoribonuclease domain of Rrp44 (Wang *et al.*, 2007; Bonneau *et al.*, 2009; Schneider *et al.*, 2009; Makino *et al.*, 2013). In humans, three homologues of Rrp44 exist, but only two of these (DIS3 and DIS3L) associate with the exosome complex (Staals *et al.*, 2010; Tomecki *et al.*, 2010). Like Rrp44, DIS3 has both endo- and exonucleolytic activity, while DIS3L only harbours exonuclease activity (Tomecki *et al.*, 2010). The third human Rrp44 homologue, DIS3L2, is responsible for degradation of oligouridylated ncRNAs in the cytoplasm (Ustianenko *et al.*, 2013; Pirouz *et al.*, 2016; Ustianenko *et al.*, 2016).

The second catalytic component of the exosome complex, Rrp6, is a member of the DEDD superfamily of exonucleases (Steitz and Steitz, 1993; Midtgaard *et al.*, 2006; Wasmuth *et al.*, 2014). Rrp6 associates with the exosome core and Rrp44 to form the nuclear exosome complex in yeast. Yeast and human Rrp6 contain an N-terminal PMC2NT domain, required for its interaction with its co-factor Rrp47 (Stead *et al.*, 2007), an exoribonuclease domain containing four conserved amino acids, which chelate a metal ion for 3' to 5' exonuclease activity, an HRDC domain, which may be responsible for substrate binding (Liu *et al.*, 2006) and a C-terminal domain (CTD) required for the interaction of Rrp6 with the core exosome (Callahan and Butler, 2008; Makino *et al.*, 2013). The CTD of yeast Rrp6 contacts the cap proteins at the top of the core exosome (Makino *et al.*, 2013). Rrp6 interacts with the RNA polymerase II transcription termination factor, Nrd1 (Vasiljeva and Buratowski, 2006) and is required for correct transcription termination (Fox *et al.*, 2015). Rrp44 and Rrp6 have both common and discrete roles in RNA processing and degradation (Gudipati *et al.*, 2012; Schneider *et al.*, 2012). Unlike Rrp44/DIS3, Rrp6 can function, in some processes, independently of the core exosome (Callahan and Butler, 2008; Fox and Mosley, 2016). Human RRP6 (also known as EXOSC10 and PM/SCL-100) contains conserved PMC2NT, HRDC and exonuclease domains, and in addition has a C-terminal domain specific to higher eukaryotes, which is also important for the interaction of RRP6 with the core exosome (Wasmuth and Lima, 2017) (Figure 1.17B). Unlike the yeast protein, human RRP6 is able to degrade structured substrates, suggesting its activity may be regulated differently in human cells (Januszyk *et al.*, 2011). While yeast Rrp6 is exclusively nuclear, RRP6 is concentrated to the nucleolus in human cells, and is also present in the cytoplasm (van Dijk *et al.*, 2007; Tomecki *et al.*, 2010).

The exosome associates with multiple cofactors in its various roles in RNA processing and degradation. The nuclear TRAMP (Trf4/5-Air1/2-Mtr4) polyadenylation complex is an exosome cofactor that assists exosome-mediated degradation by targeting substrates and stimulating exonucleolytic activity (LaCava *et al.*, 2005; Vanacova *et al.*, 2005; Houseley *et al.*, 2006; Callahan and Butler, 2008). Trf4 and Trf5 are poly(A) polymerases that add short poly(A) tails to the 3' end of exosome substrates, targeting them for degradation (Dez and Tollervey, 2004; Wyers *et al.*, 2005). Mtr4 is an RNA helicase that can function as a component of the TRAMP complex or independently, and can unwind RNA duplexes *in vitro* (Bernstein *et al.*, 2006; Wang *et al.*, 2008). The binding of Mtr4 to the core exosome is dependent on

Rrp6 and its cofactor Rrp47 (Schuch *et al.*, 2014). Independent of both the TRAMP complex and the core exosome, Mtr4 enhances the activity of Rrp6 *in vitro* (Callahan and Butler, 2008).

During pre-rRNA processing, the exosome is responsible for degradation of excised 5' ETS fragments, generated by endonuclease cleavages at sites A_0 and A_1 (Lebreton *et al.*, 2008; Schaeffer *et al.*, 2009; Schneider *et al.*, 2009). Upon defective pre-rRNA processing, aberrant precursors are also targeted for degradation by the nuclear exosome (Dez *et al.*, 2006; Wery *et al.*, 2009). In Rrp6 mutants, accumulated pre-rRNAs contain short poly(A) tails, indicating a role for the TRAMP complex in targeting aberrant pre-rRNAs for degradation (Fang *et al.*, 2004; Kuai *et al.*, 2004; Houseley *et al.*, 2006)**.** The exosome also functions in the maturation of the 5.8S 3' end by 3' to 5' exonucleolytic trimming of the 7S precursor following cleavage at site C_2 in ITS2 (Mitchell *et al.*, 1997; Allmang *et al.*, 1999). This processing event also requires the exosome cofactor Mtr4 (de la Cruz *et al.*, 1998; Jia *et al.*, 2012). While the core exosome is responsible for the initial exonucleolytic processing of the 7S pre-rRNA, further processing towards the mature 3' end of the 5.8S rRNA is catalysed by Rrp6, and generates a 6S precursor (Briggs *et al.*, 1998; Allmang *et al.*, 1999). This function of Rrp6 is independent of the core exosome, as Rrp6 lacking the C-terminal region, required for its association with the core, can efficiently process the 7S precursor (Callahan and Butler, 2008).

In humans, depletion of DIS3 leads to accumulation of 3' extended forms of the 5.8S rRNA, suggesting that the role of the exosome in 5.8S maturation is conserved between yeast and humans (Tomecki *et al.*, 2010; Preti *et al.*, 2013). The exosome is also important for degradation of the 5' ETS, indicating that this is a conserved function (Sloan *et al.*, 2014). Similarly, RRP6 depletion also causes accumulation of 3' extended 5.8S precursors (Tomecki *et al.*, 2010; Tafforeau *et al.*, 2013). Surprisingly, however, the exosome is also implicated in maturation of the 3' end of human 18S via processing of ITS1 (Preti *et al.*, 2013; Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013). In contrast to yeast, cleavage at site 2 (yeast site A_3) is the major endonucleolytic processing event in ITS1, and the generated 21S precursor is trimmed from the 3' end to site 2a to produce 18SE (Carron *et al.*, 2011; Preti *et al.*, 2013; Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013). The exonucleolytic activity of RRP6 was shown to be required for the 3' to 5' processing of ITS1 following endonucleolytic cleavage at site 2, and this also required the core exosome protein RRP46 and the exosome cofactor MTR4

(Sloan *et al.*, 2013b). It is unknown whether cleavage at the 2a site occurs after exonucleolytic processing, or if trimming continues to the 2a site. Another 3' to 5' exonuclease, poly(A)-specific ribonuclease (PARN) was recently shown to be required for nuclear processing of the 18SE precursor upstream of site 2a (Ishikawa *et al.*, 2017; Montellese *et al.*, 2017). The observation that 18SE is present in multiple 3' extended forms and is frequently polyuridylated in the cytoplasm, indicates that this precursor is targeted for processing by an unknown cytoplasmic 3' to 5' exonuclease, similar to the addition of short poly(A) tails by the TRAMP complex in the nucleus (Preti *et al.*, 2013). A possible candidate is the cytoplasmic Rrp44 homologue, DIS3L2 (Ustianenko *et al.*, 2013; Pirouz *et al.*, 2016; Ustianenko *et al.*, 2016). A recent high-throughput cancer screen identified 78 cancer-associated mutations in RRP6, with many of these located in the vertebrate-specific CTD (EAR and LASSO domains), suggesting a possible link between the higher eukaryote-specific functions of RRP6 and disease (Figure 1.17B) (Iorio *et al.*, 2016).

1.9 The 5S RNP

Unlike the 18S, 5.8S and 25S/28S rRNAs, the third LSU rRNA, 5S, is transcribed independently in the nucleoplasm by RNA polymerase III (Figure 1.18A) (Granneman and Baserga, 2004). Like RNA polymerase I-transcribed rRNAs in the nucleolus, 5S is transcribed as a precursor (Ciganda and Williams, 2011), which in yeast, undergoes processing by the 3' exoribonucleases Rex1, Rex2 and Rex3 to remove a 3' end extension (van Hoof *et al.*, 2000). In human cells, processing of the 5S rRNA requires the LSU RP RPL5 (Sloan *et al.*, 2013a). Aberrant pre-5S is polyadenylated and targeted for degradation by the exosome complex in yeast (Kuai *et al.*, 2004; Fulnecek and Kovarik, 2007). Free 5S rRNA is unstable and is rapidly degraded and is stabilised by binding of RPL5 immediately after transcription to form the pre-5S RNP complex (Deshmukh *et al.*, 1993; Sloan *et al.*, 2013a). The pre-5S RNP complex is then transported to the nucleolus, where it binds another LSU RP, RPL11, to form the mature 5S RNP (Figure 1.18A) (Sloan *et al.*, 2013a). The association of the yeast 5S RNP with the LSU requires ribosome biogenesis factors Rrs1 and Rpf2 (Zhang *et al.*, 2007). Human homologues of these factors, RRS1 and BXDC1 do not appear to interact with the 5S RNP, but depletion of either protein affects the localisation of RPL5 and RPL11 (Sloan *et al.*, 2013a).

1.10 Ribosomopathies

As a key determinant of the ability of all cells to produce proteins, ribosome biogenesis plays a fundamental role in cell growth and proliferation, and therefore it is unsurprising that defects in a single event in ribosome production can bring about human disorders. Several genetic diseases, known as ribosomopathies, have been characterised by their association with defects in genes relating to ribosome biogenesis (Narla and Ebert, 2010; Danilova and Gazda, 2015). Ribosomopathies are caused by mutations in genes encoding factors required for a range of ribosome assembly events. These include genes encoding RPs of both the SSU (Ebert *et al.*, 2008; Lipton and Ellis, 2010) and LSU (Cmejla *et al.*, 2009), as well as genes encoding non-ribosomal ribosome biogenesis factors (Heiss *et al.*, 1998; Gonzales *et al.*, 2005; Weiner *et al.*, 2012; Angrisani *et al.*, 2014). Despite being caused by mutations in a range of genes, encoding proteins with distinct functions in ribosome biogenesis, many ribosomopathy patients present with similar phenotypes, including anaemia, cranio-facial defects and predisposition to cancer (Yelick and Trainor, 2015).

Several ribosomopathies are caused by defects in SSU processing, and some are caused by mutations in components of the SSU processome (Sondalle and Baserga, 2014). Diamond-Blackfan Anaemia (DBA) is caused by mutations in RPs of both small and large subunits (Delaporta *et al.*, 2014), but the SSU RP RPS19 is the most frequently mutated gene in DBA patients (Flygare *et al.*, 2007; Ellis, 2014). North American Indian childhood cirrhosis (NAIC) is caused by mutation of the SSU processome component UTP4 (Chagnon *et al.*, 2002; Yu *et al.*, 2005). In yeast, Utp4 is required for SSU pre-rRNA processing and rDNA transcription as part of the UtpA subcomplex, while human UTP4 is only required for pre-rRNA processing (Freed and Baserga, 2010). Bowen-Conradi Syndrome (BCS) is a ribosomopathy caused by mutation of the methyltransferase EMG1, and causes symptoms such as undescended testes, and micrognathia (underdeveloped lower jaw) (Bowen and Conradi, 1976) (Armistead *et al.*, 2009). Yeast Emg1 (also called Nep1) is a component of the SSU processome (Liu and Thiele, 2001; Dragon *et al.*, 2002; Bernstein *et al.*, 2004). It is responsible for methylation of the hypermodified pseudouridine at nucleotide 1191 of the 18S rRNA (Wurm *et al.*, 2010). Emg1 interacts with Utp30 in yeast, and this interaction is important for the incorporation of the SSU RP Rps19 into pre-ribosomes (Schilling *et al.*, 2012). As mentioned above, Rps19 is mutated in many DBA patients, but any potential link between these two ribosomopathies remains unknown.

1.11 Ribosome Biogenesis, p53 and Cancer

The majority of ribosomopathy patients have an elevated cancer risk, and many ribosomopathies result in deregulated levels of the transcription factor p53 (Fumagalli and Thomas, 2011). p53 is the major tumour suppressor in human cells and is often mutated in cancers and other diseases. Levels of p53 in the cell are primarily regulated by the E3 ubiquitin ligase MDM2 (Momand *et al.*, 1992). MDM2 binds to and ubiquitinates p53, targeting it for degradation by the proteasome (Figure 1.18) (Wade *et al.*, 2013)**.** Expression of MDM2 is activated by p53, so MDM2-mediated repression of p53 represents a positive feedback loop, ensuring that p53 remains at low levels under normal conditions (Wu and Levine, 1997). Under stress conditions, such as DNA damage, oxidative stress, or oncogenic stress, p53 is stabilised (Brown *et al.*, 2009), often via phosphorylation of either p53 or MDM2, which disrupts their interaction and therefore does not result in MDM2-mediated p53 inhibition (Momand *et al.*, 2000). The activation of p53 leads to a stress response involving either repair, apoptosis, cell cycle arrest or senescence (Meek, 2015).

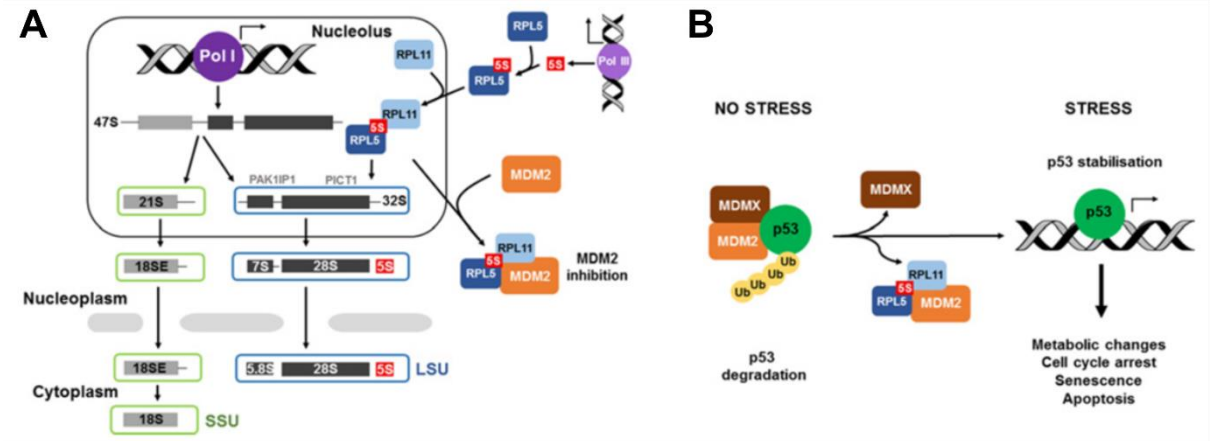


Figure 1.18 Activation of p53 via the 5S RNP-MDM2 pathway. (A) Schematic representation of pre-rRNA processing in human cells and the interaction of the free 5S RNP with MDM2. (B) Schematic representation of the regulation of p53 signalling by MDM2 and the 5S RNP in stressed and unstressed cells. From (Pelava et al., 2016).

1.11.1 Regulation of p53 by the 5S RNP

Many RPs can directly bind to MDM2 and subsequently cause activation of p53 (Figure 1.18B) (Chakraborty *et al.*, 2011). In particular, upon defective ribosome biogenesis, the 5S RNP-associated RPs, RPL5 and RPL11 are essential for activation

of p53 (Sun *et al.*, 2010; Bursac *et al.*, 2012; Sloan *et al.*, 2013a). When not incorporated into the LSU, free 5S RNP in the nucleoplasm binds to and inhibits the function of MDM2, inhibiting p53 ubiquitination and therefore stabilising p53 (Pelava *et al.*, 2016). The complete 5S RNP, containing RPL5 and RPL11 as well as the 5S rRNA, is required for the interaction with MDM2 (Donati *et al.*, 2011; Donati *et al.*, 2013; Sloan *et al.*, 2013a). Defective production of the mature LSU leads to the accumulation of free 5S RNP and subsequent MDM2 inhibition and p53 stabilisation (Horn and Vousden, 2008; Donati *et al.*, 2013; Sloan *et al.*, 2013a). Interestingly, defects in maturation of the SSU are also associated with 5S RNP-dependent p53 activation (Fumagalli *et al.*, 2009; Fumagalli *et al.*, 2012). This is perhaps surprising, as SSU and LSU follow independent maturation pathways, however a mechanism has been proposed whereby SSU defects result in a compensatory upregulation of RP gene transcription, causing an increase in levels of the 5S RNP (Fumagalli *et al.*, 2012). This model suggests that when LSU maturation is defective, upregulation of RPs, in addition to the disruption of 5S RNP incorporation into LSUs, leads to higher levels of free 5S RNP, causing suprainduction of p53 (Fumagalli *et al.*, 2012). A recent screening study showed that depletion of many RPs, especially late-assembling LSU RPs that effect late stages of LSU maturation, affects nucleolar structure and p53 homeostasis (Nicolas *et al.*, 2016).

1.11.2 p53 induction upon depletion of ribosome biogenesis factors

While the effect of RP depletion on p53 levels has been quite comprehensively studied (Fumagalli *et al.*, 2012; Nicolas *et al.*, 2016), the effect on p53 levels upon depletion of non-ribosomal factors is poorly understood. This is an important area of research since several ribosomopathies are caused by mutations in genes encoding ribosome biogenesis factors, and many ribosomopathy phenotypes are dependent on p53 (Fumagalli and Thomas, 2011). Recently, it was shown that depletion of both SSU and LSU ribosome biogenesis factors, functioning at different stages of ribosome production, causes activation of p53, and that this activation is dependent on the 5S RNP (Pelava, 2016). A recent high-throughput screen of tumours and cancer cell lines also identified multiple frequently mutated ribosome biogenesis factors (Iorio *et al.*, 2016).

1.12 Project Aims

While ribosome biogenesis has been extensively studied, particularly in yeast, the enzymes responsible for some key endonucleolytic cleavages during pre-rRNA processing have so far evaded discovery (Henras *et al.*, 2015). These include the endonucleases responsible for cleavage at three early sites required for maturation of the 18S rRNA of the SSU, which requires the formation of the SSU processome. In yeast, the PIN domain of Utp24 was previously shown to be important for cleavage at two sites, A_1 and A_2 (Bleichert *et al.*, 2006), while Rcl1 was controversially proposed as the endonuclease cleaving at site A_2 (Horn *et al.*, 2011). In humans, ribosome biogenesis in general is much less well characterised compared to yeast, and the enzymes responsible for cleavage at all three sites were unknown. While the presence of the human homologues of both Utp24 and Rcl1 is required for all three early cleavages, they had not previously been studied in detail (Tafforeau *et al.*, 2013; Sloan *et al.*, 2014). The importance of the human UTP24 PIN domain had not been investigated, nor had the importance of residues in human RCL1 equivalent to those essential for Rcl1 function in yeast.

A third potential endonuclease in the SSU processome, Utp23 was previously shown to contain a degenerate PIN domain in yeast. While the presence of Utp23 was required for all three early pre-SSU cleavages, mutation of conserved putative catalytic residues in the PIN domain did not affect processing at any site, making Utp23 a highly unlikely candidate for endonuclease cleavage in yeast (Bleichert *et al.*, 2006; Lu *et al.*, 2013). The predicted roles of Utp23 in assisting in the release of snR30 from pre-ribosomes (Hoareau-Aveilla *et al.*, 2012) and in coordinating interactions of ribosome biogenesis factors around the ES6 region of the 18S rRNA were thought to be its primary functions. In humans, however, UTP23 contains a potentially active PIN domain, suggesting it may mediate endonucleolytic cleavage. The importance of the conserved residues in the putative active site of UTP23's PIN domain had not previously been examined. In addition, it was unknown whether roles in snoRNA release or coordination of protein interactions in the pre-ribosome were conserved for human UTP23.

Ribosome biogenesis is tightly linked to cancer, and many ribosomopathies are associated with elevated cancer risks. Ribosome biogenesis is also a key regulator of the tumour suppressor p53, via the 5S RNP-MDM2 pathway. A mutation in UTP23 was

discovered to be associated with colorectal cancer (Timofeeva *et al.*, 2015), although it was unknown what effect this mutation has on the stability of UTP23, on UTP23's function in pre-rRNA processing or whether expression of this mutant causes p53 induction. Other ribosome biogenesis factors were found frequently mutated in a large screen of tumours and cancer cells, including multiple factors important for processing of ITS1 (Iorio *et al.*, 2016). These included the exonucleases XRN2 and RRP6, both of which are important for pre-18S processing (Sloan *et al.*, 2013b; Sloan *et al.*, 2014), among other roles in RNA processing and degradation, and RRP5, which is important for both endonuclease cleavages in ITS1 (Sloan *et al.*, 2013b). It was previously unknown if depletion of these factors results in p53 induction. Interestingly, the pathways of ITS1 processing differs significantly between yeast and humans, with cleavage at an alternative cleavage site favoured in human cells under normal conditions (Sloan *et al.*, 2013b). The fact that multiple factors mediating ITS1 processing in humans are frequently mutated in cancers suggests that defects in ITS1 processing may affect ITS1 processing pathway choice in cancer cells. Yeast Rrp5 can be functionally separated, with the two halves of the protein important for processing at different cleavage sites in ITS1 (Eppens *et al.*, 1999; Lebaron *et al.*, 2013). This suggests that it may influence the ITS1 cleavage site predominantly used. While human RRP5 is important for both cleavages in ITS1, it is unknown if the functional separation of the protein, and its potential impact on ITS1 cleavage site selection, is conserved in humans.

The objectives of this study were to:

- Investigate the roles of putative endonucleases, UTP24 and RCL1 in early cleavages of human 18S rRNA maturation
- Examine the importance of the putative active PIN endonuclease domain of human UTP23, as well as UTP23's association with the U17 snoRNA and protein factors associated with U17 and the ES6 region of the 18S rRNA
- Analyse the effect of a UTP23 colorectal cancer-associated mutation on pre-rRNA processing and p53 induction in human cells
- Study the effect of depleting putative endonucleases and ITS1-associated ribosome biogenesis factors on p53 induction in humans

- Characterise the importance of the different domains of human RRP5 on distinct ITS1 processing events to further understand differences between yeast and human ITS1 processing

Chapter Two
Materials and Methods

The sources of reagents used in specific experiments are stated in the appropriate subsection of this chapter. Chemicals used to make solutions were purchased from Melford Laboratories and cell culture disposables (flasks, plates etc.) were purchased from Sarstedt.

2.1 PCR and Cloning

Open reading frames (ORFs) of ribosome biogenesis factors UTP24, RCL1, UTP23 and RRP5 were cloned into a pcDNA5/FRT/TO vector for transfection into HEK293T or U2OS Flp-In cells to generate stable cell lines. UTP24, RCL1 and RRP5 were cloned into a pcDNA5 vector containing a FLAG-tag sequence to produce proteins with N-terminal FLAG-tags, while the UTP23 ORF was edited to include an HA-tag sequence at the C-terminus before being cloned into a modified pcDNA5 vector without a FLAG-tag sequence. UTP23 WT and D31N pGEX constructs for expression of recombinant protein were already available in the lab. pGEX-UTP23 C103A and P215Q mutant constructs were generated by site-directed mutagenesis (see section 2.1.10).

2.1.1 Design of RNAi-resistant constructs

Constructs were designed to produce mRNAs resistant to the specific siRNA used by altering the DNA sequence of the ORF without altering the coding sequence. RNAi-resistant ORFs were either purchased from IDT or altered by site-directed mutagenesis after pcDNA5 cloning (see section 2.1.10).

2.1.2 Reverse Transcription

Reverse transcription was performed on total RNA extracted from human cells (see section 2.4.1) to produce cDNA for amplification of ORFs. Total RNA was treated with DNase to remove any genomic DNA in the sample with the addition of 2 µl of DNase Turbo (Invitrogen) and 2 µl of DNase Turbo Buffer (Invitrogen) to 8 µl (~16 µg) of total RNA, made up to 20 µl with water. Samples were incubated at 37 °C for 1 hour, followed by incubation at 80 °C for 5 minutes to inactivate the DNase. Approximately

400 ng (1 µl) of DNase-treated total RNA was incubated with 1 µl of oligo-dT primer and 850 µM dNTPs in a 12 µl reaction, before incubation for 5 minutes at 65 °C and 1 minute on ice. DTT was added to a final concentration of 10 mM along with First Strand (FS) Buffer (Invitrogen) at a final concentration of 1x and 1 µl of rRNAsin (Promega). Samples were mixed gently before incubation at 42 °C for 2 minutes and addition of 1 µl of Superscript Reverse Transcriptase III (Invitrogen). The final 20 µl reactions were incubated at 42 °C for 50 minutes, and then 70 °C for 15 minutes. 1 µl of this template was used in PCRs.

2.1.3 ORF amplification by PCR

ORFs were amplified from cDNA, made by reverse transcription of total mRNA (see section 2.1.11), or from purchased gBlock gene fragments (IDT) by Polymerase Chain Reaction (PCR) using Phusion HF DNA polymerase (New England Biolabs). 50 µl PCR reactions were made up of 200 µM dNTPs, 1 µM of each primer (Forward and Reverse) (Table 2.1), ~10 ng of template DNA, high fidelity (HF) Phusion buffer (New England Biolabs) at a final concentration of 1x, and 0.5 µl of Phusion HF DNA polymerase (New England Biolabs). Primers for amplification of genes contained restriction sites for cloning. Conditions used for PCR are shown in Table 2.2.

Table 2.1 Primers used for Gene Amplification

Gene	Forward Primer (5' – 3')	Reverse Primer (5' – 3')
UTP24	CACC**AGATCT**ATGGGGAAGC AAAAGAAAACAAGGAAGTAT GCGACCATGAAGCGAATGCT TAG (**BglII**)	CGCG**CTCGAG**TTAGAATCGAGG GGCTCCATAATCATCTGG (**XhoI**)
RCL1	CACC**GGATCC**ATGGCGACTC AGGCGCACTC (**BamHI**)	CCCC**CTCGAG**TCACTTGAGGGT CTTGCTAAGGTTGG (**XhoI**)
UTP23	CACC**AAGCTT**GGATCCTCCA CCATGAAGATCACAAGGCAG AAACATG (**HindIII**)	CC**GTCGAC**TCATGCATAGTCCG GGACGTCATACGGATAGCCCGC ATAGTCAGG (**SalI**)
RRP5	CACC**AGATCT**ATGGCAAACC TGGAAGAAAGCTTCCC (**BglII**)	GTAC**GTCGAC**CTAGTCCTCTAGC ACTGAGCTCTTGGCC (**SalI**)
RRP5 NTD S1 9	CACC**AGATCT**GCAGCTTTCT CCCTGACCTCTCAC (**BglII**)	GTAC**GTCGAC**CTAGTCCTCTAGC ACTGAGCTCTTGGCC (**SalI**)
RRP5 NTD S1 10	CACC**AGATCT**GCTCTTAACA CTCACTCTGTTAGCC (**BglII**)	GTAC**GTCGAC**CTAGTCCTCTAGC ACTGAGCTCTTGGCC (**SalI**)
RRP5 CTD S1 9	CACC**AGATCT**ATGGCAAACC TGGAAGAAAGCTTCCC (**BglII**)	GTAC**GTCGAC**TCACAGGTGGCC CGTCTCTACCAAG (**SalI**)
RRP5 CTD S1 10	CACC**AGATCT**ATGGCAAACC TGGAAGAAAGCTTCCC (**BglII**)	GTAC**GTCGAC**TCAAGTGTGGCC ATCCTCCAGCTC (**SalI**)
RRP5 Del S1 9	GGGCCACCTGACTGGGAAGC CAGACGTGCTTTCTGC	GCTTCCCAGTCAGGTGGCCCGT CTCTACCAAGGAGGC
RRP5 Del S1 10	GGCCACACTACTGGGAAGCC AGACGTGCTTTCTGC	GCTTCCCAGTAGTGTGGCCATCC TCCAGCTCACTTG

Table 2.2 Phusion HF PCR Conditions

Phusion HF PCR			
Step	**Temperature (°C)**	**Time (sec)**	**Cycles**
Initial Denaturation	98	30	1
Denaturation	98	5	35
Annealing	Tm + 3	10	
Extension	72	15/kb	
Final Extension	72	600	1

2.1.4 Agarose gel electrophoresis

Products of Phusion PCR reactions were diluted in 6x agarose loading dye (30 % glycerol (v/v) and 0.3 % Orange G (w/v)) to a final concentration of 1x. PCR products were then separated on a 1% agarose-1x TBE (90 mM Tris-HCl, 90 mM boric acid, 2 mM EDTA, pH 8.0) gel containing 1x SYBR Safe DNA gel stain (Invitrogen). Gels were run at 130 V in 1x TBE solution (90 mM Tris-HCl, 90 mM boric acid, 2 mM EDTA, pH 8.0) for ~30 minutes and DNA was visualised using a PhosphorImager (Typhoon FLA9000; GE Healthcare).

2.1.5 Extraction and purification of DNA from agarose gels

DNA was extracted from agarose gels using the Wizard SV gel and PCR Clean-Up System Kit (Promega), according to the manufacturer's instructions.

2.1.6 pJET Cloning

Approximately 10 ng (1 µl) of extracted PCR product was added to ~50 ng of pJET1.2 vector (0.5 µl) (Promega), 0.5 µl of T4 DNA ligase (Promega), 5 µl of 2 x Ligase Reaction Buffer (Promega) and 5 µl of water. The pJET ligation reaction was incubated at room temperature for 30 minutes before *E. coli* transformation.

2.1.7 E. coli transformation

Up to 10 µl of ligation reaction (pJET (section 2.1.5)), or conventional cloning reaction (section 2.1.8) or plasmid DNA was added to 100 µl of DH5α *E. coli* competent

cells and incubated on ice for 30 minutes. DH5α *E. coli* competent cells (genotype: F-Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1) were previously prepared by the Inoue method (Sambrook and Russell, 2006). Heat-shock was performed for 1 minute at 42 °C before incubation on ice for 5 minutes. 1 ml of Lysogeny Broth (LB) was added to cells, which were then incubated at 37 °C for 1 hour with shaking at 550 rpm, before centrifugation at 3,000 rpm for 1 minute and removal of 900 µl of media. Transformed cells were resuspended in the remaining ~100 µl media and plated on LB agar plates containing ampicillin for selection of positive transformants. Plates were incubated overnight at 37 °C and a single colony was picked from each plate and added to 3 ml LB medium containing ampicillin (100 µg/ml). Cultures were incubated overnight at 37 °C, before plasmid DNA extraction using the GeneJet Plasmid Miniprep Kit (Thermo Scientific) according to the manufacturer's instructions.

2.1.8 Restriction Digest

DNA inserts were released from the pJET1.2 vector by restriction digest using restriction enzymes (Promega). Approximately 2 µg of plasmid DNA was digested with 10 U of the appropriate restriction enzymes (Promega) and 1x Enzyme Buffer (Promega) in a final volume of 20 µl overnight at 37 °C. Appropriate restriction enzymes and buffers are summarised in Table 2.3. If an appropriate buffer for both restriction enzymes was not available, single digestion was performed with one enzyme, followed by extraction and purification of the cut product using the Wizard SV gel and PCR Clean-Up System Kit (Promega), and a second single digest with the other restriction enzyme. The pcDNA5 target vector was also digested overnight with the appropriate enzymes and buffer, followed by heat inactivation of the restriction enzymes (incubation at 65 °C for 10 minutes) and incubation with 1 µl alkaline phosphatase at 37 °C for 1 hour.

Table 2.3 Restriction enzymes and buffers used for construct digestion

ORF	Enzyme 1	Enzyme 2	Enzyme Buffer
UTP24	BglII	XhoI	D
RCL1	BamHI	XhoI	B
UTP23	HindIII	SalI	B/D
RRP5	BglII	SalI	D

2.1.9 pcDNA5 Cloning

Digested inserts and pcDNA5 target vectors were extracted from 1% agarose gels and purified using the Wizard SV gel and PCR Clean-Up System Kit (Promega). Insert and target vector were then combined at a roughly equimolar ratio with a final DNA concentration of ~10 ng/µl, with 1 µl of 10 x T4 DNA ligase buffer and 1 µl T4 DNA ligase in a 10 µl ligation reaction. Ligation reactions were incubated overnight at 16 °C, and then transformed into 100 µl DH5α competent *E. coli* cells as described above. Colonies were screened by restriction digest and confirmed by DNA sequencing.

2.1.10 DNA Sequencing

The sequences of constructs cloned into the pJET1.2 or pcDNA5 plasmids were confirmed by DNA sequencing performed by either Source Bioscience or GATC Biotech. For smaller ORFs (<2 kB, UTP24, RCL1, UTP23), sequencing was performed with primers corresponding to sequences within the plasmid both upstream and downstream of the ORF. For wild type RRP5 (>5 kB) and RRP5 mutants (2-5 kB), additional sequencing primers were designed and used for sequencing. RRP5 sequencing primers are listed in Table 2.4.

Table 2.4 Primers used for sequencing of RRP5 constructs

Primer	Sequence (5’ – 3’)
RRP5 Seq F1	GCAACGGAGGAGTTGTTAGTC
RRP5 Seq R1	GAGTTTTTTACACTGGCCAGG
RRP5 Seq F2	CTTCCTCAAGCTTGTGAGGAC
RRP5 Seq R2	CGGGGTCAGAGAGTCTAGTCC

2.1.11 Site-directed mutagenesis

Site-directed mutagenesis was used to introduce small changes into pcDNA5 and pGEX constructs, to make point mutations or to make constructs RNAi-resistant. Mutagenesis was carried out using the QuikChange Site Directed Mutagenesis Kit (Stratagene). Approximately 200 ng of pcDNA5 construct DNA was combined with 1x Pfu Turbo Buffer (Stratagene), 100 ng of Forward and Reverse primer, 200 nM dNTPs and 2.5 U Pfu Turbo (Stratagene). Primers used for mutagenesis are shown in Table 2.5. PCR was performed according to manufacturer’s instructions, and PCR conditions are shown in Table 2.6.

Table 2.5 Primers used for site-directed mutagenesis

ORF	Mutation	Forward Primer (5' – 3')	Reverse Primer (5' – 3')
UTP24	D72N	CCACCTTACCACATCCTCGTTAATACCAACTTTATCAACTTTTCC	GGAAAAGTTGATAAAGTTGGTATTAACGAGGATGTGGTAAGGTGG
UTP24	D142N	CCATGTACACACAAAGGAACCTATGCAAATGACTGCTTAGTACAGAGAGTAAC	GTTACTCTCTGTACTAAGCAGTCATTTGCATAGGTTCCTTTGTGTGTACATGG
RCL1	S313	CAGCAGGATGTTTCCTAAGTCCTGCTAGGCCCTC	GAGGGCCTAGCAGGACTTAGGAAACATCCTGCTG
RCL1	RHK-AAA	CGATAGAATTTTTGGCGGCTTTGGCGAGCTTTTTCCAG	CTGGAAAAAGCTCGCCAAAGCCGCCAAAAATTCTATCG
UTP23	D31N	CCGTACCAGATCCTGCTGAACGGCACCTTCTGTCAG	CTGACAGAAGGTGCCGTTCAGCAGGATCTGGTACGG
UTP23	C103A	GCAGTGAGTGGATCAGAAGCTCTGCTTTCCATGGTTGAAGAG	CTCTTCAACCATGGAAAGCAGAGCTTCTGATCCACTCACTGC
UTP23	P215Q	CTTAGTTGTTTGAAGAAAAAGAAAAAGGCACAGGACACACAATCATCTGCTTCTG	CAGAAGCAGATGATTGTGTGTCCTGTGCCTTTTTCTTTTTCTTCAAACAACTAAG
RRP5	RNAi Resistance	CACTGGCCAGGTGGTTAAAGTCGTAGTGCTCAACTGTGAGCCATCCAAAG	CTTTGGATGGCTCACAGTTGAGCACTACGACTTTAACCACCTGGCCAGTG

Table 2.6 PCR conditions used for site-directed mutagenesis

Pfu Turbo PCR			
Step	**Temperature (°C)**	**Time (sec)**	**Cycles**
Initial Denaturation	95	30	1
Denaturation	95	30	17
Annealing	55	60	
Extension	68	60/kb + 60	

PCR products were incubated on ice for 5 minutes and then digested with 1 µl of Dpn1 for 1 hour at 37 °C to digest the methylated, template DNA strand. 8 µl of Dpn1 digested sample was transformed into 100 µl of competent DH5α cells. Plasmid DNA was extracted as described above and sequenced to confirm the mutagenesis and check for secondary mutations.

2.2 Cell Culture and *in vivo* assays

2.2.1 Cell Culture

HEK293T (human embryonic kidney) Flp-In™ T-REx™ cells (Invitrogen), U2OS (osteosarcoma; ATCC) and U2OS Flp-In (provided by Dr Laurence Pelletier, Samuel Lunenfeld Research Institute, Toronto, Canada) cells were cultured with 10 ml Dulbecco's Modified Eagles Medium (DMEM) containing 4500 mg/L glucose, L-Glutamine and sodium bicarbonate (Sigma Aldrich) supplemented with 10% Fetal Calf Serum (FCS) (Sigma Aldrich) and 1% Penicillin/Streptomycin (Sigma Aldrich) in T75 tissue culture flasks (Sarstedt). MCF7 (breast adenocarcinoma; ATCC) and MCF7 Flp-In (provided by Mark Vooijs, Maastricht University Medical Centre, Netherlands) cells were cultured with 10 ml RPMI 1640 (Lonza) medium with L-Glutamine and 25 mM HEPES (4-(2-Hydroxyethyl) piperazine-1-ethanesulfonic acid), supplemented with 10 % FCS (Sigma Aldrich) in T75 tissue culture flasks (Sarstedt). Cells were grown as a monolayer in a humidified incubator at 37 °C with 5 % CO_2. Cells were passaged at ~80 % confluence. Cells were trypsinised using 1 ml of 1x Trypsin-EDTA (Sigma Aldrich) in Phosphate Buffered Saline (PBS) (Sigma Aldrich) for 5-15 minutes at 37°C,

resuspended in 9 ml of the appropriate medium. 8 ml of resuspended cells was discarded, and the remaining cells were re-seeded with the appropriate medium.

Freezing media was prepared for each appropriate medium by addition of 10 % DMSO. Approximately 3x 10^6 cells were resuspended in 1 ml of the appropriate freezing media and resuspended cells were transferred to a cryo-vial, cooled to -80 °C in isopropanol and then stored in liquid nitrogen. Cells were revived from stocks by thawing rapidly at 37 °C, resuspending in 5 ml of the appropriate medium without DMSO, pelleting by centrifugation at 3,000 rpm for 5 minutes and resuspending in 10 ml of the appropriate medium.

2.2.2 siRNA-mediated knockdown

RNAi-mediated depletions were performed using reverse transfection of siRNA duplexes targeting the mRNA sequences of target proteins (Table 2.7). 5 µl of Lipofectamine RNAiMAX transfection reagent (Life Technologies) and 500 µl Opti-MEM (Invitrogen) were incubated together with the siRNA targeting the mRNA of proteins of interest for 15 minutes at room temperature in each well of a 6-well plate. Cells were resuspended in the appropriate medium for the cell line without antibiotics added and counted using a TC20 Automated Cell Counter (BioRad). Approximately 200,000 cells were added to each well and antibiotic-free media was added to the well to bring the total volume to 2 ml. Cells were incubated with the siRNA(s) for 48-72 hours before being harvested by trypsinisation and subsequent resuspension in the appropriate medium.

Table 2.7 siRNAs used for RNAi-mediated knockdown

Target Gene	Sequence (5' – 3')	Source	References
GL2 (control)	CGUACGCGGAAUACUUCGATT	Eurofins MWG	Elbashir *et al*., 2001
UTP24	UCCAAGAUUUGAACGAUUATT	Eurofins MWG	Sloan *et al*., 2013a
RCL1	GAACAUGACUGUAGCGUCCTT	Eurofins MWG	Eurofins MWG
UTP23	GAAAGUAUCAAACAUCUCATT	Eurofins MWG	Sloan *et al*., 2014
XRN2	AAGAGUACAGAUGAUCAUGTT	Eurofins MWG	Sloan *et al*., 2013a
RRP6	UGAGCAGAGUAAUGCAGUATT	Eurogentec	Sloan *et al*., 2013a
RRP5	UGAAGGUUGUCGUAUUGAATT	Eurofins MWG	Sloan *et al*., 2013a
RPL7	CCCAAAAGAUGCUUCGAAATT	Dharmacon Smartpool	
	AAUCAGAGGUAUCAAUGGATT		
	GGUAAAUACGGCAUCAUCUTT		
	GCUCAACAAGGCUUCGAUUTT		
RPL5	GGUUGGCCUGACAAAUUAUTT	Eurofins MWG	Kuroda *et al*., 2011

A siRNA targeting firefly luciferase mRNA (GL2) was used as a negative control in knockdown experiments. siRNAs purchased from MWG Eurofins are supplied pre-annealed and were resuspended in 1x siMAX buffer (MWG Eurofins) to a concentration of 20 µM. RRP6 siRNAs purchased from Eurogentec are supplied as sense and anti-sense strands, which must be annealed into a duplex before use. The two strands are combined at a concentration of 20 µM in 1x Annealing buffer (100 mM KOAc, 2 mM MgOAc, 30 mM HEPES-KOH pH 7.4, heated to 90 °C for 1 minute before incubation at 37 °C for 1 hour.

2.2.3 Flp-In Cells and Creation of Stable Cell Lines

HEK293T, U2OS and MCF7 Flp-In cells contain a stably integrated pFRT/*lac*Zeo target site vector containing a Flp Recombination Target (FRT) site (Figure 2.1). This allows for the stable transfection of the pcDNA5/FRT expression vector containing a gene of interest into the genome by homologous recombination between the FRT sites in both the host cell line and the vector. This is achieved by the co-transfection of the pcDNA5 vector and a pOG44 vector, which encodes the Flp-In recombinase. Flp-In cells also contain a blasticidin S resistance sequence.

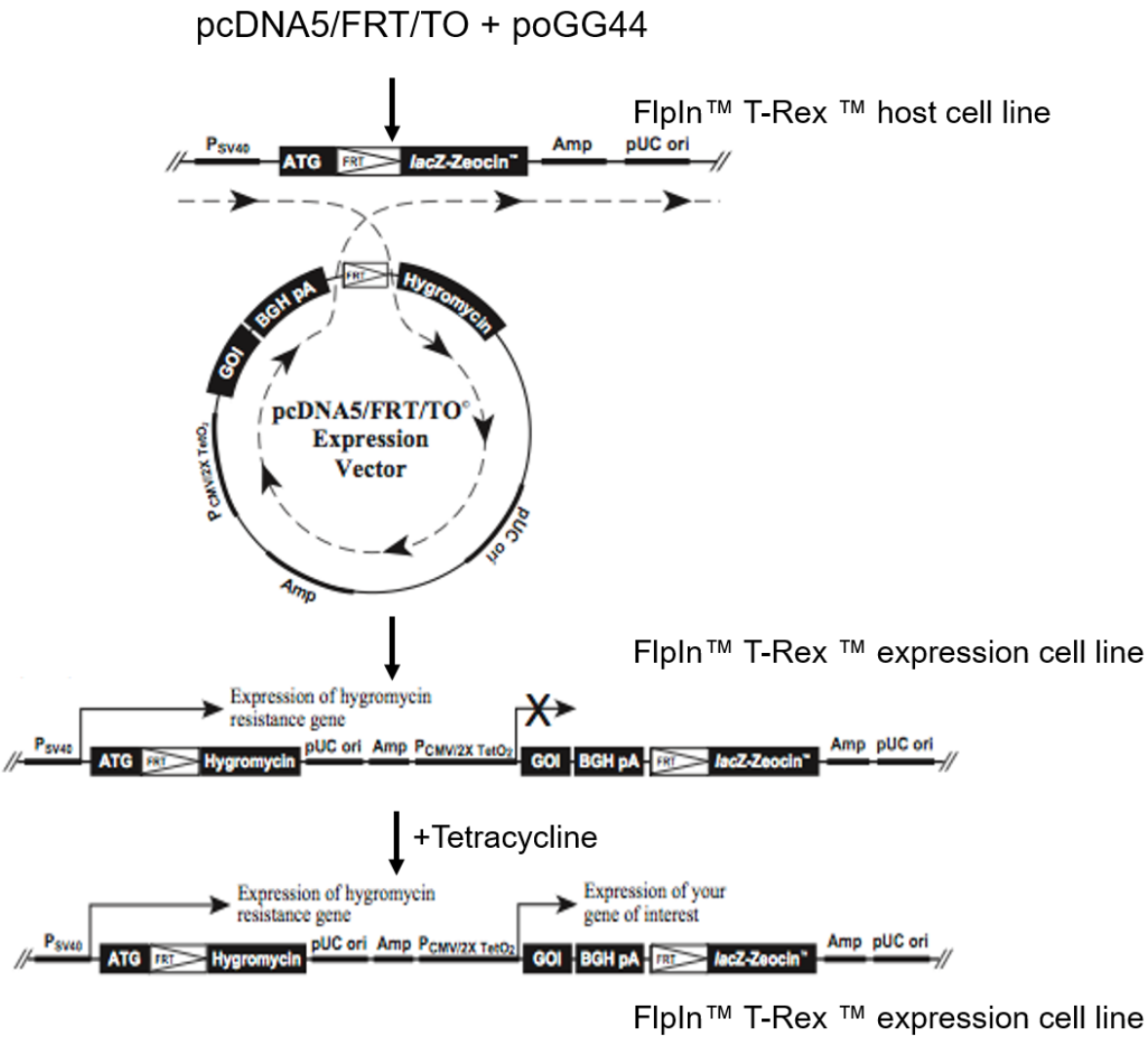


Figure 2.1 Schematic representation of the generation of stable cell lines using the Flp-In Recombination System. Co-transfection of the pcDNA5 expression vector containing a Flp recombination (FRT) site and the pOG44 vector encoding the Flp recombinase induces homologous recombination between the two FRT sites in the Flp-In host cell line and the pcDNA5 vector (shown by the dotted line). This generates a Flp-In expression cell line, containing the hygromycin-resistance gene, ampicillin-resistance gene, pCMV promoter, tetracycline inducible promoter, FLAG-tag sequence and the cloned gene of interest (GOI). (Based on a figure from Invitrogen, taken from (Pelava, 2016)).

The gene of interest was cloned into the pcDNA5 vector where it is under the control of a tetracycline promoter, allowing for induced expression of the gene by addition of tetracyline. pcDNA5 also contains a hygromycin resistance gene which allows for selection of stably transfected cells. The pcDNA5 construct also contains either 2x N-terminal FLAG tags or 2x C-terminal HA tags.

To generate stable cell lines, Flp-In cells were grown in 6-well plates to ~50 % confluency. 9 µl/well of FuGene 6 transfection reagent (Promega) was mixed with 91 µl/well of Opti-MEM and incubated at room temperature for 5 minutes. 1.8 µg of pOG44 was mixed with 0.6 µg of the pcDNA5 construct containing the gene of interest and then added to the transfection reagent mix and incubated at room temperature for 15 minutes. The mix was then added to the well in a dropwise manner and cells were incubated at 37 °C. After ~48 hours cells were trypsinised and transferred to T75 flasks before the initiation of selection with the addition of 100 µg/ml hygromycin B and 10 µg/ml blasticidin S. Stably transfected cells were allowed to grow for around two weeks before resuspension. After generation of stable cell lines, selection was continued with the addition of hygromycin B every passage and blasticidin S every 3rd passage. Expression of tagged proteins was induced by addition of 0-10 µg/ml of tetracycline diluted in 70 % ethanol and protein expression levels were titrated to achieve roughly 1:1 expression of endogenous and tagged proteins. Tetracyline concentrations required to achieve this are listed in Table 2.8.

Table 2.8 Tetracycline concentrations used for induction of protein expression from stable cell lines

Cell Line	Tetracycline concentration (ng/ml)
UTP24 WT (HEK293T)	10,000
UTP24 D72N (HEK293T)	10,000
UTP24 D72N/D142N (HEK293T)	10,000
RCL1 WT (HEK293T)	10
RCL1 S313 (HEK293T)	10
RCL1 RHK (HEK293T)	100
UTP23 WT (HEK293T)	0.5
UTP23 D31N (HEK293T)	100
UTP23 C103A (HEK293T)	100

UTP23 P215Q (HEK293T)	0.5
UTP23 WT (U2OS)	1
UTP23 WT OE (U2OS)	10,000
UTP23 D31N (U2OS)	100
UTP23 P215Q (U2OS)	1
UTP23 P215Q OE (U2OS)	10,000
RRP5 WT (HEK293T)	10
RRP5 NTD S1 9 (HEK293T)	10
RRP5 NTD S1 10 (HEK293T)	10
RRP5 CTD S1 9 (HEK293T)	1
RRP5 CTD S1 10 (HEK293T)	10
RRP5 Del S1 9 (HEK293T)	10
RRP5 Del S1 10 (HEK293T)	10

2.2.4 RNAi rescue

The coding sequence of genes of interest (UTP24, RCL1, UTP23, RRP5) were altered to make the mRNAs resistant to siRNAs used to deplete the endogenous protein. Substitutions were made in the target sequence of the appropriate siRNA within the coding sequence that did not alter the amino acid sequence. After stable transfection of the RNAi-resistant pcDNA5 construct and selection of transformants, expression of RNAi-resistant proteins was induced by addition of tetracycline and cells were incubated with the appropriate siRNA to deplete the endogenous protein. Cells were harvested after 72 hours and protein levels were analysed by western blotting, using the appropriate antibody to detect both endogenous and RNAi-resistant, FLAG/HA-tagged proteins. This confirmed that while levels of endogenous proteins were significantly reduced, the RNAi-resistant, tagged protein levels remained stable.

2.2.5 Immunofluorescence in human cells

Cells were plated onto a single coverslip in one well of a 24-well plate and expression of HA-tagged UTP23 was induced with the addition of tetracycline (see table 2.8). After 72 hours incubation at 37 °C, cells were fixed with 200 µl of 4 % paraformaldehyde in PBS for 20 minutes at room temperature. Paraformaldehyde was washed from the wells with PBS before incubation with 200 µl PBS/0.1% Triton for 15

minutes at room temperature. Cells were washed four times with 500 µl PBS before blocking with PBS/0.1% Triton/10 % fetal calf serum (FCS) for 1-2 hours at room temperature. The cell-coated coverslips were transferred to a clean well and 50 µl of anti-HA or anti-fibrillarin antibody (see Table 2.9) diluted in PBS/0.1% Triton/10 % FCS was added directly to the well and incubated for 1-2 hours. Coverslips were transferred back to the original wells and washed three times quickly with PBS, followed by three 10 minute washes with PBS. Coverslips were then transferred to the second well and 50 µl fluorescent secondary antibody diluted in PBS/0.1% Triton/10 % FCS was added directly to the wells and incubated in the dark for 1-2 hours. Antibodies and antibody dilutions used for immunofluorescence are shown in Table 2.9. Coverslips were again transferred to the original well and washed three times quickly with PBS, followed by one 10-minute PBS wash, one 10-minute incubation with DAPI (4',6'-diamidino-2-phenylindole) (0.1 µg/ml, Sigma) diluted in PBS (1 in 10,000 dilution), and a final 10-minute PBS wash. All washes following incubation with the secondary antibody were performed in the dark. Coverslips were then immersed quickly in water 5 times, then in ethanol 5 times before being left to air-dry in the dark and mounted to a glass microscope slide using Mowiol (Sigma). A Zeiss Axiovert 200M inverted microscope was used to visualise cells and localisation was analysed using Axiovision software and images were processed using Photoshop (Adobe). Anti-HA signal indicated the localisation of HA-tagged UTP23, the anti-fibrillarin signal was used as a marker for the nucleolus and DAPI was used to stain nuclei.

Table 2.9 Antibodies used for Immunofluorescence in human cells

Antibody	Raised In	Source	Dilution
α-HA	Mouse	Babco (MMS-101P)	1:500
α-Fibrillarin	Rabbit	Santa Cruz (sc-25397)	1:200
α-Rabbit IgG (Alexa Fluor 555 conjugate)	Donkey	Invitrogen (A-31572)	1:500
α-Mouse IgG (Alexa Fluor 647 conjugate)	Goat	Invitrogen (A-21235)	1:500

2.2.6 Immunoprecipitation

Approximately 1x 10^7 cells were grown in a T75 flask and expression of HA-tagged protein was induced by addition of tetracycline. Cells were centrifuged at 900 rpm for 5 minutes, and cells were resuspended in 2 ml of Gradient Buffer E (150 mM KCl, 20 mM HEPES pH 8.0, 0.5 mM EDTA, 0.1 mM DTT) and sonicated at 20 % ultrasonic amplitude/0.3 second pulse/1 second off, at 3 x 20 second on, 20 second off intervals using a BANDELIN SONOPULS. Triton-X-100 (to 0.2 % final concentration), glycerol (to 10 %) and $MgCl_2$ (to 1.5 mM) was added to sonicated samples. Samples were centrifuged at 13,000 rpm at 4 °C for 10 minutes to remove insoluble material, and the supernatant was transferred to a new microcentrifuge tube. 10 % of the total sample was kept and stored at -20 °C.

20 µl of anti-HA agarose beads (Sigma) per IP sample and 20 µl of NHS-activated sepharose beads (GE Healthcare) per control sample were prepared by washing in 3 x in cold IP buffer (10% glycerol, 1x G150 (see below), 0.2% Triton-X-100, 1 mM DTT). Cell lysates were divided equally between anti-HA and control beads and samples were gently rotated at 4 °C for 2 hours before centrifugation at 3,000 rpm

for 1 minute. Beads were washed 3 x with 1 ml of cold IP buffer, with centrifugation at 3,000 rpm for 1 minute at 4 °C between each wash. After a final wash with 1x IP buffer, co-precipitated RNA was extracted and analysed.

RNA extraction was performed by adding 180 µl of Homogenisation Buffer (1% SDS, 50 mM Tris, 50 mM NaCl, 0.5 mM EDTA) and 200 µl of Phenol/Chloroform/Isoamylalcohol (Sigma) to the washed beads for each IP sample and 300 µl Homogenisation Buffer and 400 µl Phenol/Chloroform/Isoamylalcohol to the beads for each total sample. Samples were vortexed for 3 minutes and then centrifuged at 13,000 rpm for 5 minutes. After centrifugation, the upper layer was transferred to a clean microcentrifuge tube and 20 µl of 3 M NaOAc (40 µl for totals) and 500 µl of 100 % ethanol (1 ml for totals) was added. 1 µl of GlycoBlue (15 mg/ml stock; Fisher Scientific) was added to IP samples. Samples were incubated at -20 °C overnight, and then centrifuged at 13,000 rpm for 20 minutes at 4 °C and supernatant was removed. The pellet was washed in 500 µl of ice-cold 75 % ethanol and centrifuged at 13,000 rpm for 20 minutes at 4 °C. The supernatant was removed, and the pellet was air-dried before being resuspended in 8 µl of water. RNA was analysed by northern blotting as described in section 2.4.2.

2.2.7 Immunofluorescence in yeast cells

Immunofluorescence was performed on yeast cells using an anti-TAP antibody to assess the localisation of wild type and RDK mutant Rcl1. Plasmids encoding HTP (His_6-TEV-protein A)-tagged Rcl1 wild type or Rcl1 RDK mutant were transformed into yeast strains expressing HA-tagged Rcl1 under control of a repressible *GAL10* promoter (genotype: *MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; RCL1-HTP-URA3*). Cells were grown in minimal media containing glucose for 6 hours to deplete endogenous Rcl1 and were harvested at mid log phase (OD_{600} 0.4 – 0.8). Harvested cells were fixed in 4 % formaldehyde diluted in 1x PBS for 15 minutes at room temperature before washing 3 times with 1x PBS and once in buffer B (0.1 M KPO_4 pH 7.5, 1 M sorbitol, 10 mM DTT). Cells were then incubated for 45 minutes at 30 °C in buffer B containing 50 U/ml of Zymolyase (Zymo Research), before pelleting by centrifugation and resuspension in 100 µl of buffer B per ~$1x10^7$ of cells and transferred to a coverslip in a 24-well plate. Cells were incubated on coverslips for 30 minutes at room temperature and then washed twice in 2 ml of 1x PBS before blocking with 1 ml of PBS containing 5 % milk for 30 minutes at room temperature with shaking. Blocking

solution was removed and 200 µl of primary antibody (anti-TAP or anti-Nop1 as a control), diluted in PBS containing 5 % milk was added to each coverslip and cells were incubated for 1 hour at room temperature. Excess primary antibody was removed with 3 5-minute washes with 2 ml of PBS containing 5 % non-fat dried milk (w/v) (Marvel). Coverslips were incubated with 200 µl of secondary antibody diluted in PBS containing 5 % milk for 1 hour at room temperature in the dark before washing with PBS containing DAPI (4',6'-diamidino-2-phenylindole). Coverslips were mounted onto microscope slides using Vectorshield. A Zeiss Axiovert 200M microscope was used to visualise cells. Images were analysed using Axiovision software and processed using Photoshop (Adobe).

Table 2.10 Antibodies used for Yeast Immunofluorescence

Antibody	**Raised In**	**Source**	**Dilution**
α-TAP (detects linker between TEV and protein A of HTP tag)	Rabbit	Thermo Scientific (CAB1001)	1:1,000
α-Nop1	Mouse	Santa Cruz (sc-57940)	1:2,000
α-Rabbit IgG (Alexa Fluor 555 conjugate)	Donkey	Invitrogen (A-31572)	1:500
α-Mouse IgG (Alexa Fluor 647 conjugate)	Donkey	Invitrogen (A-31570)	1:500

2.3 Protein Analysis

2.3.1 SDS-PAGE

Human cells grown in 6-well plates were trypsinised, transferred to microcentrfuge tubes and then pelleted by centrifugation at 3,000 rpm for 5 minutes. Cell pellets (or protein pellets recovered following TRI reagent RNA extraction, see section 2.4.1) were resuspended in 50-100 µl of 2 x protein loading dye (PLD) (74 mM Tris-HCl pH 6.8, 1.25 mM EDTA, 20 % glycerol, 2.5 % SDS, 0.125 % bromophenol blue, 50 mM DTT). Samples were incubated at 95 °C for 5-10 minutes and briefly centrifuged before loading on a 1.5 mm denaturing SDS polyacrylamide gel (SDS-PAGE) (BioRad) 10 % polyacrylamide resolving gels were used for large (≥80 kDa) proteins and 13 % polyacrylamide resolving gels were used for small (<80 kDa) proteins and a 4 % polyacrylamide stacking gel was used for all samples. Electrophoresis was carried out at 200 V for 50-60 minutes in 1x Protein Running Buffer (25 mM Tris-HCl pH 8.3, 250 mM glycine, 0.1% SDS (w/v)).

2.3.2 Western blotting

After polyacrylamide gel electrophoresis, samples were transferred to nitrocellulose membrane (Protran, GE Healthcare) for 1.5 hours at 65 V in Transfer Buffer (25 mM Tris-HCl pH 8.3, 150 mM glycine, 10 % methanol). Total protein on membranes was detected by staining with Ponceau S solution (0.1% in 5% HAc (w/v)). Membranes were blocked with the appropriate blocking buffer depending on the method of detection for at least 1 hour at room temperature, to limit non-specific binding of antibodies to the membrane. For detection by enhanced chemiluminescence (ECL) (Pierce) and exposure on ECL hyperfilm (GE Healthcare), blocking was performed in 1x PBS, 0.05 % Triton X-100 (v/v), 2 % non-fat dried milk (w/v) (Marvel). For detection of fluorescently-labelled secondary antibody using the Odyssey-LI-COR system (LI-COR), blocking was performed in 1x TBS, 0.1% Tween-20 (v/v), 2 % Marvel skimmed milk powder (w/v).

Primary antibodies (listed in Table 2.11) were diluted in the appropriate blocking buffer to their optimal concentration. Membranes were incubated with primary antibodies at 4 °C overnight before being washed for 3 x 10 minutes with PBS-Triton X-100 (0.05 %, v/v) (ECL detection) or TBS-Tween-20 (0.1%, v/v) (Odyssey-LI-COR). Membranes developed using ECL were incubated with Horseradish Peroxidase

(HRP)-conjugated secondary antibodies. Membranes detected by the Odyssey-LI-COR system were incubated with fluorescently-labelled secondary antibodies. All membranes were incubated with secondary antibodies (listed in Table 2.12) for 1-2 hours at room temperature and then washed 3 x 10 minutes with PBS-Triton X-100 (0.05 %, v/v) (ECL) or TBS-Tween-20 (0.1%, v/v) (Odyssey-LI-COR). Membranes visualised using the Odyssey-LI-COR system were washed a fourth time with 1x TBS without Tween-20. Quantification of western blots was performed using ImageQuant software (GE Healthcare) and quantification data was normalised to the loading control.

Table 2.11 Primary Antibodies used in Western Blotting

Antibody	**Raised In**	**Source**	**Dilution**	**Reference**
α-Karyopherin β1	Rabbit	Santa Cruz (sc-11367)	1:2,000	
α-RCL1	Rabbit	Eurogentec (custom)	1:1,000	Sloan *et al*., 2013a
α-UTP24	Rabbit	Eurogentec (custom)	1:1,000	Sloan *et al*., 2013a
α-UTP23	Rabbit	Aviva (ARP60526_P050)	1:1,000	
α-FLAG	Rabbit	Santa Cruz	1:10,000	Sloan *et al*., 2013a
α-FLAG	Mouse	Santa Cruz	1:10,000	Sloan *et al*., 2013a
α-HA	Mouse	Berkeley Ab Company	1:5,000	Sloan *et al*., 2013a
α-XRN2	Rabbit	Bethyl Laboratories, Inc.	1:2,000	Sloan *et al*., 2013a
α-RRP5	Rabbit	Eurogentec (custom)	1:1,000	Turner *et al*., 2009
α-p53	Mouse	Santa Cruz (sc-126)	1:500	
α-RPL5	Rabbit	SantaCruz (A0912)	1:5,000	
α-RPL7	Rabbit	Abcam (ab72550)	1:2,000	
α-GAPDH	Mouse	Santa Cruz (sc-47724)	1:20,000	

Table 2.12 Secondary Antibodies used in western blotting

Antibody	Raised In	Source/Reference	Dilution
α-Rabbit IgG-HRP	Donkey	Santa Cruz (sc-2313)	1:10,000
α-Mouse IgG-HRP	Donkey	Santa Cruz (sc-2314)	1:10,000
α-Rabbit 800CW	Goat	LI-COR (926-32211)	1:10,000
α-Mouse 800CW	Donkey	LI-COR (926-32212)	1:10,000

2.4 RNA Analysis

2.4.1 RNA extraction

Total RNA from cell pellets was extracted using TRI reagent (Ambion). Cells were harvested from 6-well plates by trypsinisation, transferred to microcentrifuge tubes and pelleted by centrifugation at 3,000 rpm for 5 minutes. 500 µl of TRI reagent (Ambion) was added to whole cell pellets from each well and samples were incubated at room temperature for 5 minutes. 100 µl of chloroform was added before vortexing samples for 15-30 seconds. After 2 minutes incubation at room temperature, samples were centrifuged at 13,000 rpm for 15 minutes to separate RNA (upper phase), DNA (interphase) and protein (lower phase). The upper phase containing RNA was transferred to a clean microcentrifuge tube and 250 µl of isopropanol (Sigma Aldrich) was added. Samples were vortexed briefly, incubated for 5 minutes and then centrifuged at 13,000 rpm for 10 minutes to pellet the RNA. The supernatant was discarded, and the RNA pellet was washed with 500 µl of 75 % ethanol and then centrifuged for 5 minutes at 13,000 rpm. The RNA pellet was dried using a speed-vac for 1-2 minutes and then diluted in 11 µl of water and incubated at 55 °C with shaking at 550 rpm for 10 minutes. RNA was then stored at -20 °C.

The concentration of extracted RNAs was measured using a NanoDrop Spectrophotometer (Thermo Scientific). Around 1 µg of RNA was mixed with 5 volumes of 5 x glyoxal loading dye (GLD) (61.2 % DMSO (v/v), 20.4 % glyoxal (v/v), 12.2 % 1x BPTE buffer (28.7 mM Bis-Tris, 9.9 mM PIPES (piperazine-N,N′-bis(2-ethanesulfonic acid), 1 mM EDTA) (v/v), 4.8 % glycerol (v/v)) with 0.2 mg/ml EtBr and incubated at 55 °C for 1 hour. Glyoxalated RNA samples were loaded on a small 1.2 % agarose BPTE gel and run at 120 V for 20-30 minutes before being visualised on a PhosphorImager (Typhoon FLA9000; GE Healthcare) to confirm the integrity of 28S and 18S rRNAs.

2.4.2 Northern Blotting

2.4.2.1 Gel electrophoresis

For detection of pre-rRNA intermediates, around 2 µg of extracted RNA, estimated based on NanoDrop (Thermo Scientific) data and from small agarose gels (see section 2.4.1), was added to 5 volumes of 5x GLD containing 0.02 mg/ml EtBr and incubated at 55 °C for 1 hour. Samples were separated on a large 1.2% agarose, 1x BPTE gel, using 1x BPTE as a running buffer, at 185 V for 3.5 hours and then mature 28S and 18S rRNAs were visualised using a PhoshphorImager (GE Healthcare). The gel was washed once in 75 mM NaOH at room temperature for 20 minutes, followed by two 15-minute washes in Tris Salt Buffer pH 7.4 (0.5 mM Tris-HCl pH 7.4, 1.5 M NaCl) and a final 20-minute wash in 6x SSC (1 M NaCl, 0.1 M $Na_3C_6H_5O_7$) at room temperature. RNA was then transferred to a Hybond N nylon membrane (Amersham) overnight in 6x SSC at room temperature via the capillary action method. After transfer, RNA was crosslinked to the membrane using a Stratalinker UV Crosslinker (Stratagene) using the “auto” setting twice.

For detection of smaller RNAs (<500 nt), extracted RNA was diluted in 1x RNA loading dye and loaded on an 8 % or 10 % acrylamide/7 M Urea gel. Acrylamide/Urea gels were run for around 1 hour at 220V with 1x TBE (90 mM Tris-HCl, 90 mM boric acid, 2 mM EDTA, pH 8.0) used as running buffer. RNA was then transferred to a Hybond N membrane (Amersham) in 0.5 x TBE solution for 90 minutes at 65 V, before crosslinking of the RNA to the membrane using a Stratalinker (Stratagene) as above. Crosslinked RNA was visualised by staining the membrane with methylene blue.

2.4.2.2 Northern blot hybridisation

For probing with 5’ labelled oligonucleotide probes, crosslinked membranes were pre-hybridised in SES1 buffer (0.5 M sodium phosphate pH 7.2, 7 % SDS (w/v), 1 mM EDTA) for 1 hour at 37 °C. Freshly made 5’ labelled oligonucleotide probes (see section 2.4.2.3) were then added to the SES1 pre-hybridisation buffer, and membranes were hybridised overnight at 37 °C. After removal of the probe, membranes were washed twice for 20 minutes per wash in 1x SSC/ 0.1% SDS before being exposed to a PhosphorImager screen and later visualised using a PhosphorImager (Typhoon FLA9000; GE Healthcare). For random-prime labelled probes, crosslinked membranes were pre-hybridised in Pre-Hyb buffer (25 mM NaPO4 pH 6.5, 6x SSC, 5x Denhardts,

0.5 % SDS (w/v), 50 % deionised formamide, 100 µg/ml denatured salmon sperm DNA) for 2 hours at 42 °C, the probe was added to the Pre-Hyb buffer and hybridisation was performed overnight at 42 °C. After removal of the random-prime probe, membranes were washed in 2 x SSC/0.5 % SDS twice for 5 minutes per wash at 42 °C, followed by two 5-minute washes in 2 x SSC/0.1% SDS at 42 °C and a final 20-minute wash in 2 x SSC/0.1% SDS at 50 °C. Membranes were then exposed to a PhosphorImager screen and visualised using a PhosphorImager (Typhoon FLA9000; GE Healthcare).

2.4.2.3 Northern blotting probes

Oligonucleotide probes were radioactively labelled using ^{32}P γ-ATP (Perkin Elmer). 1 µl (10 µM) of the oligonucleotide primer (Table 2.13) was combined in a 10 µl reaction with 1 µl of T4 Polynucleotide Kinase (PNK) Buffer (New England Biolabs), 2-4 µl of ^{32}P γ-ATP and 1 µl of T4 PNK (New England Biolabs). When using a fresh stock of ^{32}P γ-ATP, 2 µl was used; after more than one half-life 4 µl was used. Samples were incubated at 37 °C for 45 minutes and then made up to 50 µl with water. The sample was passed through a G50 column (Life Technologies) by centrifugation at 3,000 rpm for 2 minutes at room temperature to remove non-incorporated radioactivity. Before being added to SES1 pre-hybridisation buffer, the probe was incubated at 95 °C for 2 minutes.

Table 2.13 Oligonucleotide probes used for detection of pre-rRNAs in Northern blotting

Probe	**Sequence (5’ – 3’)**
18SE (5520)	CCTCGCCCTCCGGGCTCCGTTAATGATC
ITS1 (6121)	AGGGGTCTTTAAACCTCCGCGCCGGAACGCGCTAGGTAC
3’ ITS1 (6603)	CGAGGTCGATTTGGCGAGGGC
U17	TTCCTGCATGGTTTGTCTCC
7SK	GTGTCTGGAGTCTTGGAAGC

To make the random-prime labelled ETS1 probe (Table 2.14), 25-50 ng of PCR product, amplified from a plasmid template, was mixed with water to a volume of 9 µl

and incubated at 95 °C for 5 minutes. 3 µl of random hexamer mix (250 mM Tris pH 7.5, 50 mM MgCl2, 5 mM DTT, 500 µM dATP, 500 µM dGTP, 500 µM TTP), 2 µl of ^{32}P α-CTP (Perkin Elmer) and 1 µl of Klenow polymerase was added and the reaction was incubated at 37 °C for 30 minutes. The reaction was made up to 50 µl with water and passed through a G50 column (Life Technologies). Before being added to the Pre-Hyb buffer, the probe was incubated at 95 °C for 5 minutes.

Table 2.14 Primers used to generate random-prime labelled probe used for northern blotting

Probe	Forward Primer (5' – 3')	Reverse Primer (5' – 3')	Pre-rRNA Region Amplified
ETS1	GCTGACACGCTGTCCT CTGGCGA	CGGACAACCCCGCGGA GACGAGA	1-339

2.4.4 In vitro transcription of RNA substrates

snR30 and U17 snoRNA substrates were transcribed in the presence of ^{32}P α-UTP to produce radiolabelled RNA substrates for *in vitro* RNA binding electromobility shift assays (EMSAs) (section 2.6.1). 0.5 µg of PCR product was added to 4 µl of 5 x transcription buffer (Promega), 1 µl of UTP mix (10 mM ATP, 10 mM GTP, 10 mM CTP, 1 mM UTP) (Promega), 0.5 µl of rRNAsin (Promega), and 1 µl of ^{32}P α-UTP (Perkin Elmer) in a 20 µl reaction. The transcription reaction was incubated at 37 °C for 2 hours, before the addition of 1 µl of RNase-free DNase and a further 30-minute incubation at 37 °C. 30 µl of water was added to the reaction and RNA extraction was performed using 50 µl of Phenol/Chloroform/Isoamylaclohol. After vortexing for 5 minutes and centrifugation at 13,000 rpm for 5 minutes at 4 °C, the upper phase was passed through a G50 column. RNA was precipitated using 1 µl of GlycoBlue (Fisher Scientific), 7.5 µl of NaAc pH 5.2 and 200 µl of 100% ethanol and incubating at -20 °C overnight. RNA samples were then centrifuged at 13,000 rpm for 20 minutes at 4 °C before washing with 80 % ethanol and further centrifugation at 13,000 rpm for 20 minutes at 4 °C. RNA pellets were dried and resuspended in 100 µl of water.

2.4.5 Radiolabelling of RNA oligonucleotides

An RNA oligonucleotide containing the human site A0 (pre-rRNA nts 1623-1648) was radiolabelled for nuclease assays (section 2.6.3) using T4 Polynucleotide Kinase (PNK). Before labelling, the RNA oligonucleotide was gel-purified. 500 pmoles of RNA was mixed with 10 µl of RNA loading dye and incubated at 65 °C for 10 minutes and loaded on a 12 % acrylamide/ 8 M Urea gel. The gel was then stained with SYBR Safe gel stain for 10 minutes at room temperature and de-stained with 1x TBE. The RNA was visualised using a PhosphorImager (Typhoon FLA9000; GE Healthcare) and the band corresponding to the RNA oligonucleotide was cut from the gel and eluted with 500 µl of TNES buffer (20 mM Tris/HCl pH 8, 300 mM NaCl, 5 mM EDTA, 0.1% SDS). RNA was then extracted with 500 µl of Phenol/Chloroform/Isoamylalcohol and precipitated with 1.5 µl of GlycoBlue (Fisher Scientific) and 1 ml of 100 % ethanol overnight at -20 °C. The RNA pellet was washed in 1 ml of 80 % ethanol, dried and resuspended in 50 µl of water. 1 µl of purified RNA was analysed on a 12 % acrylamide/8 M Urea gel before labelling.

1 µl of purified RNA oligonucleotide was mixed with 1 µl of 10x T4 PNK buffer, 2 µl of ^{32}P γ-ATP (Perkin Elmer) and 1 µl of T4 PNK in a 10 µl reaction and incubated at 37 °C for 1 hour. 45 µl of water was added and RNA was extracted with 50 µl of Phenol/Chloroform/Isoamylalcohol by vortexing for 5 minutes and centrifugation at 13,000 rpm for 5 minutes at 4 °C. The upper phase, containing the RNA sample, was passed through a G50 column, and RNA was precipitated by adding 1 µl of glycogen (GlycoBlue (Fisher Scientific)), 7.5 µl of NaAc pH 5.2 and 200 µl of 100 % ethanol and incubating at -20 °C overnight. The sample was centrifuged at 13,000 rpm for 20 minutes at 4 °C and the RNA pellet was washed with 80 % ethanol before centrifugation at 13,000 rpm for 20 minutes at 4 °C. Pellets were dried and then resuspended in 100 µl of water. An aliquot of the labelled RNA was checked on a 12 % acrylamide/ 8 M Urea gel before use in nuclease assays.

2.5 Protein Purification

2.5.1 Expression and purification of GST-tagged proteins

pGEX-6P1 plasmids encoding GST-tagged proteins were transformed into BL21 Codon Plus(DE3) cells (genotype: *E. coli* B^{-} F^{-} *ompT hsdS*(r_{B^-} m_{B^-}) *dcm*$^{+}$ Tet^{r} *gal* λ(DE3) *endA* Hte [*argU ileY leuW* Cam^{r}]) (Agilent Technologies) and cells were grown

on LB agar plates containing 100 µg/ml ampicillin (Amp) and 12.5 µg/ml chloramphenicol (CA) overnight at 37 °C. 25 ml of LB (+CA) was inoculated with a single BL21 Codon Plus colony and incubated for 3 hours at 37 °C before being transferred to a cold 50 ml Falcon tube and incubated on ice for 10 minutes. Cells were pelleted by centrifugation at 2,700 g for 10 minutes at 4 °C and then resuspended in 15 ml of ice-cold 80 mM $MgCl_2$-20mM $CaCl_2$. Cells were pelleted by centrifugation at 2,700 g for 10 minutes at 4 °C and pelleted were resuspended in 1 ml of ice-cold 80 mM $MgCl_2$-20mM $CaCl_2$. 2 µl of pGEX-6P1 plasmid was added to 200 µl of this BL21 cell solution and incubated on ice for 30 minutes. Heat shock was performed at 42 °C for 1 minutes, followed by incubation on ice for 2 minutes. 800 µl of LB media was added and samples were incubated at 37 °C for 45 minutes. Cells were pelleted by centrifugation at 3,000 rpm for 2 minutes and resuspended in ~100 µl of LB media before being pipetted onto LB (+Amp, +CA) plates. Plates were incubated overnight at 37 °C.

A single colony was then used to inoculate 20 ml LB cultures (+ Amp/CA) which were grown for around 16 hours at 37 °C with shaking. 10 ml of overnight culture was used to inoculate 1 L cultures of 2 x YT media (16 g/L Tryptone, 10 g/L Yeast Extract, 5 g/L NaCl) containing 100 µg/ml ampicillin and 12.5 µg/ml chloramphenicol which were grown at 37 °C with shaking until reaching an OD_{600} of between 0.1 and 0.2. Cultures were then cooled to 16 °C and protein expression was induced by adding isopropyl-beta-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Protein expression was confirmed by analysing a 1 ml sample of each culture before and after addition of IPTG by Coomassie staining. Cultures were grown overnight at 16 °C and then harvested by centrifugation at 4,000 rpm at 4 °C for 40 minutes in a Beckman J6-HC centrifuge. Supernatants were discarded, and pellets were resuspended in 10 ml Purification Buffer (10 mM Tris/HCl pH 7.6, 7 5mM NaCl, 4.5 % glycerol, 0.05 % Tween20) containing 100 ng/µl of lysozyme, half an EDTA-free Protease Inhibitor tablet/25 ml and Tris(hydroxypropyl)phosphine (THP) to a final concentration of 0.5 mM.

Cells were lysed by sonication on ice at 90 % amplitude for 3 minutes, with 0.5 second pulse. Samples were then centrifuged at 18,500 rpm at 4 °C for 45 minutes in a JA20 rotor to pellet cell debris, and clarified lysate was used for protein purification. Glutathione sepharose beads, suspended in 50 % ethanol, were prepared by washing three times in Purification Buffer, followed by centrifugation at 3,000 rpm for 1 minute.

After washing, beads were resuspended in 1 ml of Purification Buffer and THP was added to a final concentration of 0.5 mM. Washed beads were added to the clarified lysate and rotated slowly on a wheel at 4 °C for 2 hours. Samples were centrifuged at 4,000 rpm for 3 minutes at 4 °C and supernatant was removed. A 10 µl sample of supernatant was taken for analysis of unbound protein. The beads were then washed twice in 50 ml of Purification Buffer before being resuspended in 1 ml of Purification Buffer containing 50 mM reduced glutathione and 0.5 mM THP and rotated slowly on a wheel at 4 °C for 1 hour. After elution, samples were centrifuged at 4,000 rpm for 3 minutes at 4 °C, and the supernatant, containing the purified protein sample, was recovered and either frozen in liquid nitrogen and stored at -80 °C or desalted immediately (see section 2.5.3). A sample of eluted protein was taken to confirm protein purification by SDS-PAGE.

2.5.2 Expression and purification of His-tagged proteins

pET100 plasmids encoding His-tagged proteins were expressed in BL21 Codon Plus *E. coli* cells as for GST-tagged proteins, except cell pellets were resuspended in 10 ml Purification Buffer containing 20 mM imidazole. Cell pellets were lysed as for GST-tagged proteins and added to Ni-NTA sepharose Superflow agarose beads (QIAGEN), equilibrated in Purification Buffer containing 20 mM imidazole. After rotating on a wheel with the cell lysate for 2 hours at 4 °C, the beads were washed three times for 10 minutes per wash, head-over-tail at 4 °C in Purification Buffer containing 20 mM imidazole. After a further two short washes, multiple elution steps were performed with Purification Buffer containing increasing concentrations of imidazole. Elution steps were performed for 10 minutes, head-over-tail at 4 °C with 600 µl Purification Buffer containing imidazole concentrations between 50 and 1,000 mM. A sample of each eluate was taken for analysis by SDS-PAGE. Eluted protein samples were either frozen in liquid nitrogen and stored at -80 °C or immediately desalted (see section 2.5.3).

2.5.3 Desalting

Desalting was performed to remove glutathione or imidazole from GST- or His-tagged protein samples, respectively. Protein samples were desalted by gravity using illustra NAP-5 Columns according to manufacturer's instructions (GE Healthcare), which were equilibrated with 10 ml of Purification Buffer. Each ~500 µl protein sample

was then eluted into two 500 µl fractions, which were frozen in liquid nitrogen and stored at -80 °C.

2.5.4 Determination of recombinant protein concentration

Protein concentrations were determined using a Bradford assay. 5, 10 and 20 µl of protein sample was added to 795, 790 and 780 µl of water, and 200 µl of Bradford Reagent (BioRad) was added. Samples were incubated at room temperature for 5 minutes and OD_{595} was measured for each sample. 200 µl of Bradford Reagent was also added to 800 µl water for use as a blank control sample. A range of Bovine Serum Albumin (BSA) concentrations (0.2 – 1.2 µg/µl), were used to produce standard OD_{595} measurements for determining protein concentrations.

2.6 *In vitro* Assays

2.6.1 EMSA Band-shift assay

Trace amounts of *in vitro* transcribed, radiolabelled snR30 or U17 snoRNA RNA substrates were used in 10 µl, 15 µl or 20 µl reactions containing 100-3,000 nM of recombinant protein. Reactions containing the RNA substrate in, recombinant protein, 2 mM of DTT, 0.8 units/µl of RNasin, 500 ng/µl of *E. coli* tRNA, were made up to 10 µl, 15 µl or 20 µl with Purification Buffer (10 mM Tris/HCl pH7.6, 75 mM NaCl, 100 ng/µl BSA, 4.5% glycerol, 0.05% Tween-20)**.** Samples were incubated for 10 minutes at 30 °C and then 10 minutes on ice, before adding 2 µl, 3 µl or 4 µl of agarose loading dye (30 % glycerol and 0.3 % Orange G (w/v)). Free RNA and RNP complexes were separated on a 4 % polyacrylamide/1x TBE gel containing 5 % glycerol and visualised using a PhosphorImager (Typhoon FLA9000; GE Healthcare).

2.6.2 Protein-protein Interaction Studies

GST-tagged UTP23 was used as bait protein for protein-protein interaction studies with His-tagged recombinant proteins or *in vitro* translates (TNT, Promega). Approximately 100 pmoles of desalted GST-UTP23 or free GST as a control was coupled to 15 µl of glutathione sepharose beads that had been previously equilibrated in 200 µl of NB Buffer (20 mM Tris/HCl pH 7.6, 150 mM NaCl, 8.7% glycerol, 0.1% Tween-20) containing 10 µg/ml of BSA. Protein and beads were rotated on a wheel for 30 minutes at 4 °C and beads were washed with 1 ml of NB Buffer to remove unbound

protein. His-tagged UTP24, RRP7, or NHP2, or ROK1/DDX52 translated *in vitro* in the presence of ^{35}S methionine was added to the beads, which were rotated on a wheel for 1 hour at 4 °C. Beads were then washed five times in 500 µl of NB Buffer, all buffer was removed and 25 µl of 2 x PLD containing 50mM DTT was added to the beads. Samples were incubated at 95 °C with shaking at 400 rpm for 5 minutes and loaded on a 13 % denaturing SDS-PAGE gel. For samples incubated with His-tagged proteins, 5 % or 10 % of the input material was loaded and protein was transferred to a nitrocellulose membrane. Proteins were analysed by western blotting using antibodies specific to the His-tag or GST-tag. For samples incubated with *in vitro* translated ROK1/DDX52, 1% of the input material was loaded and protein was analysed using a PhosphorImager (Typhoon FLA9000; GE Healthcare).

2.6.3 Nuclease assay

Wild type or PIN domain mutant forms of recombinant GST-tagged UTP23 was used in nuclease assays using a radiolabelled RNA oligonucleotide containing the human A0 site. 20 pmoles of GST-UTP23 or free GST as a control was incubated with 2 mM DTT, 0.8 U/µl of RNasin, 1x BSA, 0.1 µg/µl of *E. coli* tRNA and 5 mM or 10 mM $MnCl_2$ in a 9 µl reaction for 5 minutes at 37 °C and then 5 minutes on ice. 1 µl of 5' end-labelled RNA oligonucleotide was added to the reaction, which was incubated for 1 hour at 37 °C. 10 µl of 2 x RNA loading dye was added to each sample and 2 µl of the 20 µl sample was loaded on a 12% sequencing gel and run at 25 W for 90 minutes. Partial alkaline hydrolysis was performed on the radiolabelled RNA oligonucleotide to produce a ladder in which each nucleotide of the substrate is cleaved. 2 µl of labelled RNA was mixed with 2 µl of 2 x AH (Alkaline Hydrolysis) Buffer (100 mM $NaHCO_3/Na_2CO_2$ pH 9, 2 mM EDTA, 0.5 µg/µl of *E. coli* tRNA) and incubated at 95 °C for 3 minutes, 5 minutes or 10 minutes. The alkaline hydrolysis reaction was stopped by adding 6 µl of AH Stop Buffer (10 M Urea, 0.8 % Bromophenol Blue, 0.8 % Xylene Cyanol) and incubating the samples on ice for 2 minutes before loading on the gel with nuclease assay samples. Gels were then dried and exposed to a PhosphorImager screen before visualising using a PhosphorImager (Typhoon FLA9000; GE Healthcare).

Chapter Three

The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans

3.1 Introduction

The production of the eukaryotic ribosome is a tightly regulated and extremely energy demanding process that requires the action of over 200 non-ribosomal, trans-acting factors (Thomson *et al.*, 2013). Ribosome biogenesis begins with the transcription of three of the four ribosomal (r)RNAs as a single precursor transcript in the nucleolus by RNA polymerase I. This primary transcript is called the 35S pre-rRNA in budding yeast and the 47S pre-rRNA in mammals and contains the 18S rRNA of the small ribosomal subunit (SSU), and the 5.8S and 25S (yeast)/28S (mammals) rRNAs of the large subunit (LSU). The 5S rRNA is processed separately after transcription by RNA pol III in the nucleus. The sequences of the mature rRNAs are separated on the 35S/47S pre-rRNA by internal transcribed spacer (ITS) regions and flanked by external transcribed spacer sequences (ETS) (Henras *et al.*, 2015). Maturation of rRNAs for assembly into functional ribosomes occurs through the removal of these spacer regions by endonuclease cleavages at various cleavage sites, followed by exonuclease trimming of cleaved pre-rRNA substrates. While the sites of endonuclease cleavages on the 35S/47S pre-rRNA have been mapped, the endonucleases responsible for catalysing several cleavage events were unknown at the start of this project.

The majority of our understanding of the pathways involved in ribosome biogenesis comes from research in *Saccharomyces cerevisiae*. The endonucleases that cleave at several sites have been identified but multiple still remained elusive before the start of this project. There was a particular gap in the current knowledge in the identity of the endonucleases responsible for the early nucleolar cleavages required to release the 18S rRNA of the small subunit. 18S rRNA maturation requires several cleavages to remove the 5' ETS and ITS1 sequences (Figure 3.1). In yeast, the 5' ETS is removed by initial cleavage at site A_0 within the 5' ETS followed by a cleavage event at the 5' end of the 18S rRNA sequence at site A_1. Cleavage within ITS1 separates the 18S rRNA from large subunit (5.8S and 25S) rRNAs. In yeast, the major cleavage separating small and large subunit rRNAs occurs at site A_2, and this

cleavage event is coupled to prior cleavages at sites A_0 and A_1. Yeast ITS1 can also be cleaved at a downstream site, called A_3, by RNase MRP (Lygerou *et al.*, 1996). The three early pre-18S cleavages, at sites A_0, A_1 and A_2, occur co-transcriptionally in the nucleolus, and generate a 20S pre-rRNA (Kos and Tollervey, 2010). This final pre-18S intermediate is then exported to the cytoplasm where it is cleaved at site D at the 3' end of the 18S rRNA sequence by the PIN domain endonuclease Nob1 to release the mature 18S rRNA (Fatica *et al.*, 2003; Pertschy *et al.*, 2009). The endonucleases responsible for the three early cleavages were, before the start of this project, unknown or controversial (see later), but all three cleavages are known to require the correct assembly of a large ribonucleoprotein complex called the SSU processome (Dragon *et al.*, 2002).

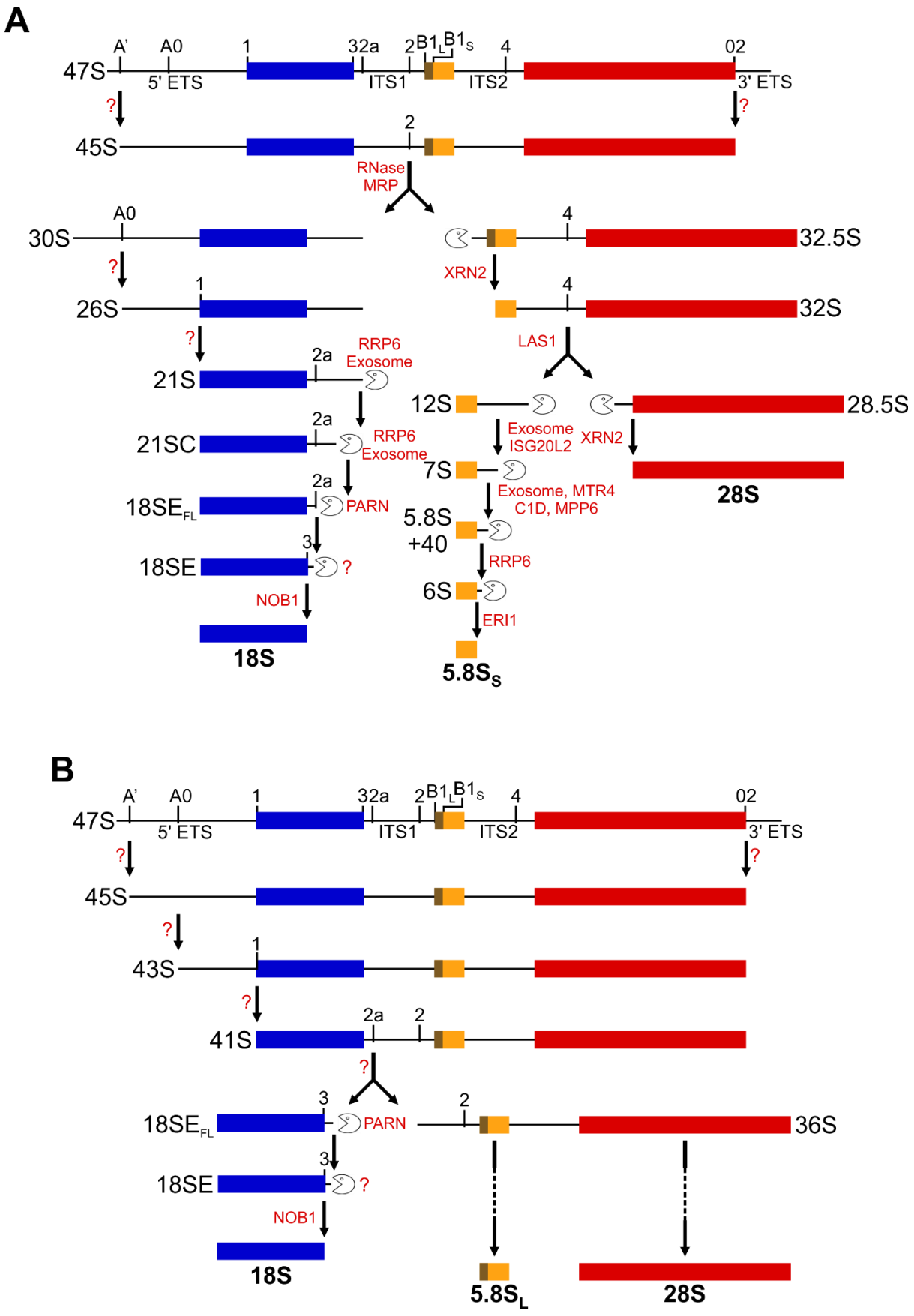


Figure 3.1 Pre-rRNA processing in *H. sapiens*. Schematic representation of pre-rRNA processing in humans. Known pre-rRNA endonuclease cleavage sites are labelled and, where known, the endonuclease or exonuclease(s) responsible for

catalysing each processing event is shown. **(A)** The major pre-rRNA processing pathway in humans, with cleavage at site 2 separating SSU and LSU precursors. **(B)** The minor pre-rRNA processing pathway in humans, involving cleavages at sites A0, 1 and 2a prior to site 2 cleavage.

The SSU processome is a large RNA-protein complex that is recruited to the pre-rRNA co-transcriptionally. The *S. cerevisiae* SSU processome has been purified and the proteins and RNA components of the complex have been identified (Grandi *et al.*, 2002; Bernstein *et al.*, 2004). The SSU processome contains many proteins, in independently assembled subcomplexes, many of which play structural roles. Several RNA-binding and RNA-modifying proteins are also found in the complex and the individual depletion of many SSU processome components inhibits cleavage at sites A_0-A_2, causing the accumulation of a 23S pre-rRNA intermediate. The SSU processome also contains several essential small nucleolar (sno)RNPs, including U3 and snR30. Unlike other snoRNPs in ribosome biogenesis, which function in modification of the pre-rRNA, U3 and snR30 play direct roles in processing at these three early cleavage sites (Hughes and Ares, 1991; Morrissey and Tollervey, 1993). The U3 snoRNA is a key component of the SSU processome which base-pairs with the pre-rRNA within the 18S rRNA sequence and the 5' ETS to ensure correct folding of the pre-rRNA during 18S rRNA maturation (reviewed in (Watkins and Bohnsack, 2012). U3 plays an important role in the formation of a key structural feature of the SSU, the central pseudoknot, and base-pairs to pseudoknot elements within the 18S rRNA (Sharma and Tollervey, 1999; Kudla *et al.*, 2011). The fact that depletion of many components of the SSU processome produces the same pre-rRNA processing phenotype makes it challenging to identify the specific endonucleases catalysing cleavage at these sites. However, the SSU processome contains some proteins with links to RNA processing, including the PIN domain proteins Utp23 and Utp24 and the RNA cyclase-like protein, Rcl1.

While the cleavage sites, processing events and proteins involved in 18S rRNA maturation are largely conserved between yeast and humans, recent studies expanding the knowledge of this process in human cells have highlighted some key differences (Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013). Cleavage sites in the human 5' ETS and ITS1, called A0, 1 and 2a, are analogous to the yeast sites A_0, A_1 and A_2, but the 5' ETS of the human 47S pre-rRNA contains an extra cleavage site, called A' upstream of A0. Cleavage at this site can seemingly be bypassed and is thought to

represent a quality control step (Mullineux and Lafontaine, 2012; Sloan *et al.*, 2014). Cleavage at site A' is the only cleavage event that occurs co-transcriptionally in humans (Lazdins *et al.*, 1997), while the majority of pre-rRNAs are processed co-transcriptionally at these early cleavage sites in yeast (Kos and Tollervey, 2010). Another major difference between the maturation of the 18S rRNA between yeast and humans is the separation of the small and large subunit rRNA sequences. In yeast, this is achieved by a major cleavage at site A_2, while cleavage at the equivalent site in humans, 2a, is part of a 'minor' processing pathway (Sloan *et al.*, 2013b). In humans, the major cleavage event within the ITS1 occurs at site 2 (yeast A3). The final human pre-18S precursor, the 18SE pre-rRNA (equivalent to the yeast 20S pre-rRNA), which is exported to the cytoplasm, is produced by 3'-5' exonucleolytic trimming which requires the exosome complex (Sloan *et al.*, 2013b). The 18SE pre-rRNA, equivalent to yeast 20S, undergoes a final cleavage in the cytoplasm at site 3 (yeast site D) by the PIN domain endonuclease, NOB1 (Bai *et al.*, 2016). Until recently, cleavage at site 2 was not thought to be mediated by human RNase MRP as RNAi-mediated knockdown of protein components of the human complex did not affect pre-rRNA processing in ITS1 (Sloan *et al.*, 2013b). However, a CRISPR study then confirmed that human RNase MRP does indeed cleave human pre-rRNA at site 2, supporting the idea that this site is analogous to yeast site A_3 (Goldfarb and Cech, 2017). Alternatively, 18SE can be produced independently of site 2 cleavage. Endonucleolytic cleavage at 2a, following cleavages at sites A0 and 1, produces the 18SE pre-rRNA, and processing through this pathway can be promoted by depletion of the 5'-3' exoribonuclease, XRN2 (Rat1 in yeast) in both humans (Sloan *et al.*, 2013b) and mouse (Wang and Pestov, 2011). In this 'minor' pathway, the 18SE pre-rRNA is also expected to be cleaved at site 3 by NOB1. Despite these differences, the SSU processome is conserved in humans and is required for the equivalent three early cleavages at sites A0, 1 and 2a. Candidate endonucleases UTP23, UTP24 and RCL1 are essential components of the human SSU processome (Turner *et al.*, 2012).

Utp23 and Utp24 both possess a PIN (PilT N-terminus) domain. The PIN domain is a predicted nuclease domain based on its similarity to 5' exonucleases and RNase H (Arcus *et al.*, 2004). The PIN domain protein Nob1 is an active enzyme responsible for cleavage at site D in yeast (Fatica *et al.*, 2003; Pertschy *et al.*, 2009; Lebaron *et al.*, 2012) and at the equivalent site, site 3, in humans to produce the mature 18S rRNA (Preti *et al.*, 2013; Bai *et al.*, 2016). PIN domains are also found in

endonucleases involved in nonsense-mediated decay (NMD) (Glavan *et al.*, 2006) and RNA degradation and processing via the exosome complex (Lebreton *et al.*, 2008; Schneider *et al.*, 2009). Three or four conserved acidic residues, involved in metal-ion binding, in the catalytic site of the PIN domain are required for the endonuclease activity of PIN domain proteins (Clissold and Ponting, 2000). Yeast Utp24 contains all four conserved acidic residues, making it a promising candidate for enzymatic activity via its PIN domain (Bleichert *et al.*, 2006). As an essential SSU processome component, depletion of Utp24 inhibits cleavages at sites A_0, A_1 and A_2 *in vivo*, while mutation of the conserved acidic residues of the PIN domain disrupts cleavage at sites A_1 and A_2, but not A_0 (Bleichert *et al.*, 2006). This finding led to the suggestion that Utp24, via its PIN domain, could be responsible for cleavage at sites A_1 and A_2 in yeast. In contrast, the PIN domain of yeast Utp23 only possesses two of the four conserved residues, suggesting that it is not an active enzyme.

As for yeast Utp24, human UTP24 also contains a conserved, potentially active, PIN domain and its depletion disrupted cleavage at sites A0, 1 and 2a (equivalent to yeast A_0, A_1 and A_2) in HeLa cells (Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013). A study published while this PhD project was ongoing suggested that UTP24 is responsible for endonuclease cleavage at site 1 in human cells (Tomecki *et al.*, 2015). Interestingly, unlike in yeast, human UTP23 possesses three of the four conserved acidic residues in its PIN domain, raising the possibility that it may play an enzymatic role human pre-rRNA processing.

Rcl1 was suggested as a candidate endonuclease due to its similarity to the RNA-modifying enzymes, RNA cyclases (Billy *et al.*, 2000). RNA cyclases catalyse the synthesis of RNA 2', 3' cyclic phosphate ends from 3' phosphates. Despite this similarity, Rcl1 does not contain a conserved catalytic cyclase domain (Tanaka *et al.*, 2011). Rcl1 has been proposed as the endonuclease responsible for cleavage at the A_2 site in yeast and its depletion disrupts early cleavages at sites A_0, A_1 and A_2 *in vivo*, with particular effect on A_2 cleavage (Horn *et al.*, 2011). In the same study, recombinant Rcl1 was reported to cleave a pre-rRNA mimic at the A_2 cleavage site *in vitro*, and a mutant Rcl1 protein, with three residues (R327, D328, K330) in the C-terminal region of the protein substituted for alanines, was unable to replicate this cleavage. *In vivo*, mutation of these 'RDK' residues produced a pre-rRNA processing phenotype identical to that of Rcl1 depletion in yeast cells, namely inhibition of cleavage at sites A_0, A_1 and A_2. This phenotype is consistent with a defect in SSU processome assembly and/or

function. Consistent with this, the R327 residue in the 'RDK' motif was shown to be essential for the interaction of Rcl1 with its binding partner, the GTPase, Bms1, as well as its subsequent nuclear import and association with pre-ribosomes (Karbstein *et al.*, 2005; Delprato *et al.*, 2014). Depletion of RCL1 in human cells disrupts processing at the equivalent cleavage sites, A0, 1 and 2a, but so far, no mutants have been tested to assess its potential role in endonuclease cleavage at site 2a or any other site.

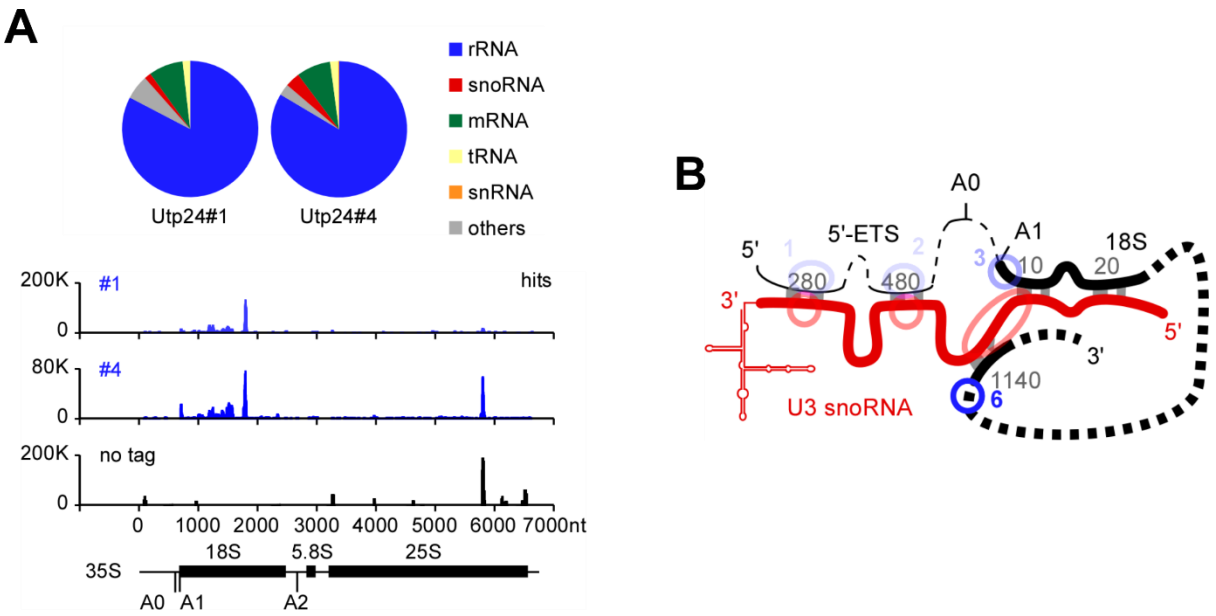


Figure 3.2 Utp24 crosslinking sites on the pre-rRNA and U3 snoRNA. (A) Transcriptome-wide binding profiles of RNAs recovered with purified Utp24 protein in two replicate UV crosslinking and analysis of cDNA (CRAC) experiments. The proportion of sequences mapping to functional RNA classes are shown as pie charts (upper panel), revealing that Utp24 was predominately found crosslinked to rRNA sequences. Sequences aligned to the rDNA sequence encoding the 35S pre-rRNA are shown plotted as reads per million, showing that Utp24 crosslinked to sequences within the 18S rRNA. **(B)** Schematic representation of the pre-rRNA (black) base-paired to the U3 snoRNA (red) during 18S central pseudoknot formation. Positions of Utp24 crosslinking to the pre-rRNA are shown in blue and crosslinking sites on the U3 snoRNA are shown in red. (From (Wells *et al.*, 2016)).

As outlined above, previous studies have provided support for the idea that Utp24 is the enzyme responsible for cleaving pre-rRNA at yeast sites A_1 and A_2, but also that Rcl1 is responsible for cleaving site A_2. A major aim of this PhD project was to characterise the roles of these candidate endonucleases in human cells. Parallel experiments in our lab also attempted to clarify the situation in yeast (Wells *et al.*, 2016). *In vivo* RNA-protein crosslinking (CRAC) was used to determine Utp24 binding

to the pre-rRNA around and within the 18S rRNA as well as to snoRNAs important in ribosome biogenesis (Figure 3.2). This revealed binding sites for Utp24 within the 18S rRNA and in the 5' ETS at regions known to base-pair with the U3 snoRNA and are essential for formation of the central pseudoknot. Utp24 also bound to the corresponding sequences in the U3 snoRNA in regions that base-pair with the pre-rRNA. These results suggest a potential role for Utp24 in verification of interactions between the rRNA and U3. Additionally, a Utp24 binding site very close to the A_1 cleavage site was identified, consistent with a role in cleavage at this site. *In vitro* cleavage analysis using a pre-rRNA mimic containing the yeast A_2 site revealed that Utp24, but not Rcl1, was able to cleave at this site, while a Utp24 mutant containing a substitution in a key acidic residue of its PIN domain did not exhibit endonuclease activity (Figure 3.3). WT Utp24 also cleaved this pre-rRNA mimic at several sites around the correct A_2 site (indicated by asterisks), suggesting that correct folding of the pre-rRNA or the action of other SSU processome components is required for accurate cleavage at this site. These data further supported the idea that yeast Utp24, but not yeast Rcl1 is an active endonuclease responsible for cleavage at two coupled sites in pre-rRNA processing.

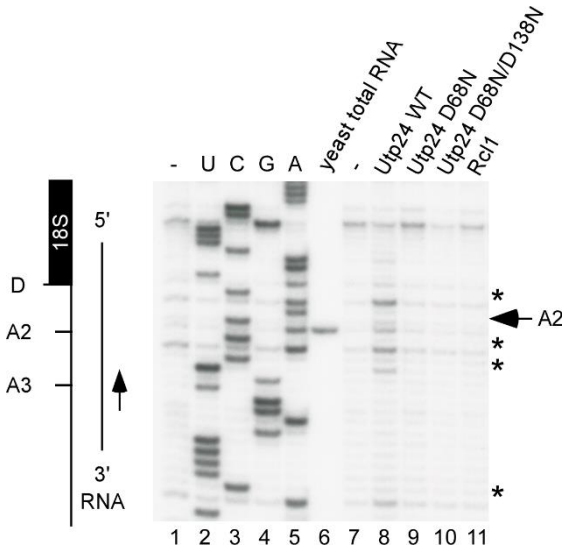


Figure 3.3 Yeast Utp24 cleaves at the A_2 site *in vitro*. An *in vitro* cleavage assay using recombinant yeast Utp24 wild type (WT) or mutant (D68N, D68N/D138N) Utp24 or WT Rcl1. Positions of ITS1 cleavage sites within the pre-rRNA mimic are shown in the left panel. Utp24-mediated cleavage at site A_2 is indicated with an arrow, and cleavages around the cleavage site are marked with asterisks. Adapted from (Wells *et al.*, 2016).

3.2 Results

3.2.1 Human UTP24 is required for 18S rRNA release in three human cell types

Previous studies have shown that UTP24 is important for pre-rRNA processing in HeLa cells (Sloan *et al.*, 2013b) and mice (Wang *et al.*, 2014). To establish the role of UTP24 in a variety of human cell types, RNAi-mediated knockdowns were performed in HEK293T human embryonic kidney cells, U2OS human osteosarcoma cells, and MCF7 human breast cancer cells. A siRNA targeting firefly luciferase (GL2) was used as a control. Depletions were performed by reverse transfection of siRNA duplexes and cells were harvested after 72 hours. Whole cell extract from harvested HEK293T (Figure 3.4A), U2OS (Figure 3.5A) and MCF7 (Figure 3.6A) cells was separated by SDS-PAGE and depletion efficiency was analysed by western blotting using an anti-UTP24 antibody. Karyopherin was used as a loading control. Efficient depletion of UTP24 protein was confirmed in all three cell lines. A representative western blot is shown for each cell line.

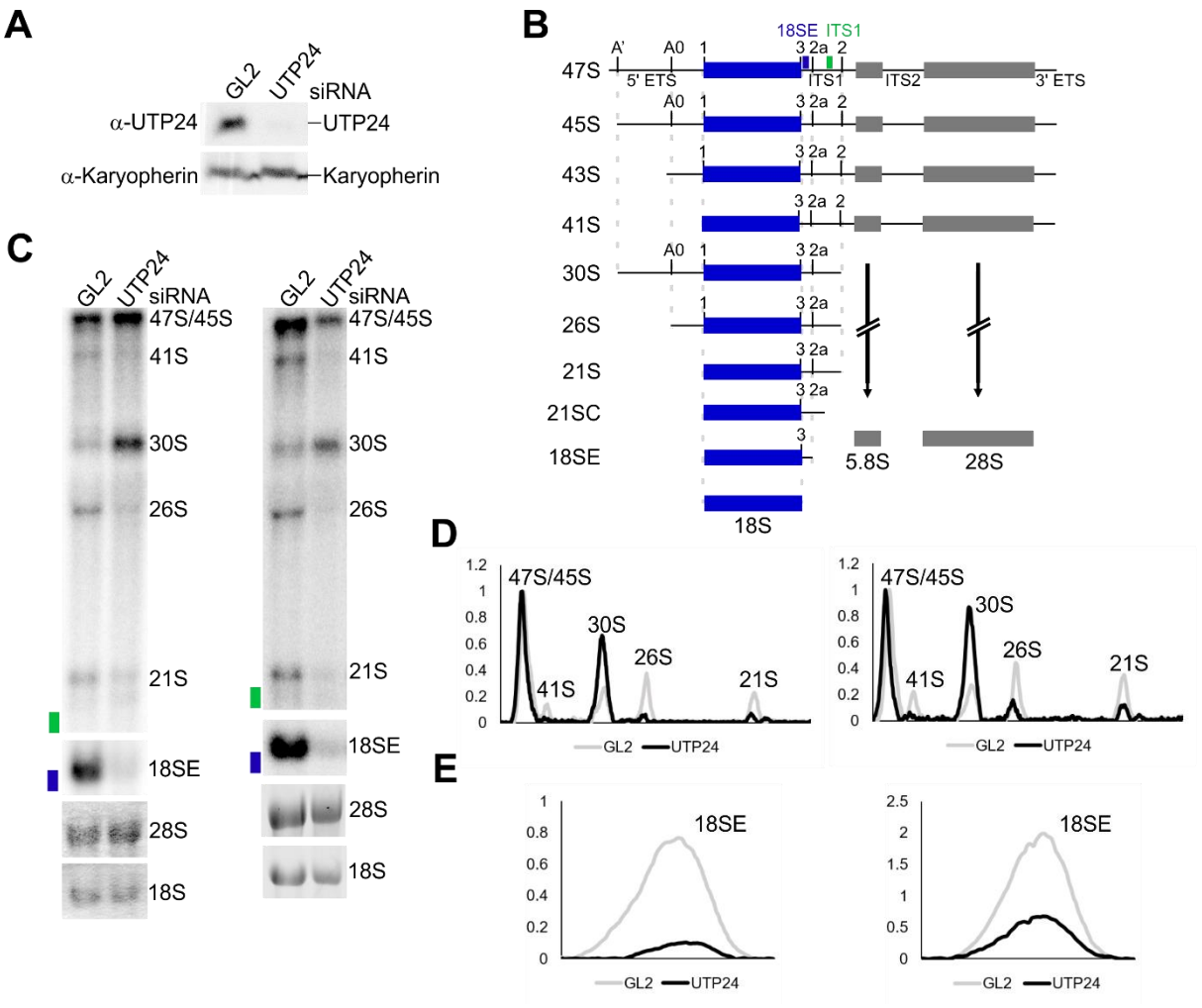


Figure 3.4 UTP24 is required for three early pre-rRNA cleavages in HEK293T cells. (A) Western blot using an anti-UTP24 antibody to confirm depletion of endogenous UTP24 upon RNAi-mediated UTP24 depletion in HEK293T cells. Karyopherin was used as a loading control. **(B)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(C)** RNA from HEK293T cells treated with control siRNA (GL2) or depleted of UTP24 (UTP24) were analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. Two representative experiments are shown. Levels of pre-rRNAs detected by the **(D)** ITS1 or **(E)** 18SE probe from panel C were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

Data from yeast studies have shown that depletion of Utp24 causes accumulation of the 23S pre-rRNA, suggesting a disruption in cleavage at sites A_0, A_1 and A_2 *in vivo* (Bleichert *et al.*, 2006). Previously published data shows that depletion of UTP24 in HeLa cells causes accumulation of the 30S pre-rRNA, suggesting

defective processing at the equivalent sites A0, 1 and 2a (Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013). Depletion of UTP24 (also called FCF1) in mouse caused accumulation of a 34S pre-rRNA, which is equivalent to the human 30S pre-rRNA, and indicates defective processing at the equivalent mouse sites, called A0, 1 and 2b (Wang *et al.*, 2014)**.** This is consistent with SSU processome defects, suggesting that UTP24 is an essential component of this complex.

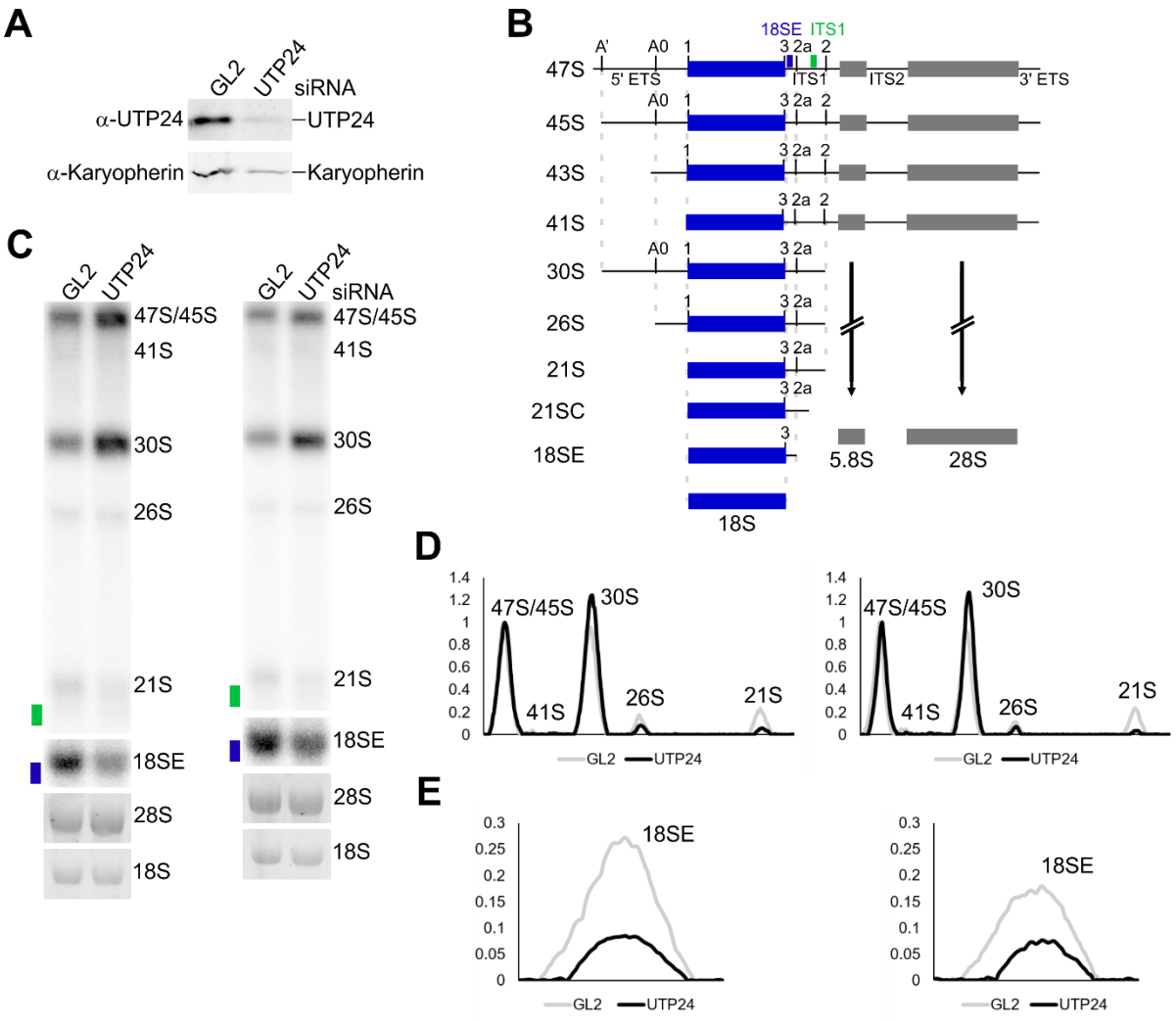


Figure 3.5 UTP24 is required for three early pre-rRNA cleavages in U2OS cells. **(A)** Western blot using an anti-UTP24 antibody to assess depletion of UTP24 upon RNAi-mediated UTP24 knockdown in U2OS cells. Karyopherin was used as a loading control. **(B)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive endonucleolytic or exonucleolytic processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the primary precursor (47S pre-rRNA). **(C)** RNA from U2OS cells treated with control siRNA (GL2) or depleted of UTP24 (UTP24) was analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and the pre-rRNAs detected by ITS1 and

18SE probes are labelled. Two representative experiments are shown. Levels of pre-rRNAs detected by the **(D)** ITS1 or **(E)** 18SE probe from panel C were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted. The pre-rRNA intermediate represented by each peak is labelled.

To assess the effect of UTP24 depletion on pre-rRNA processing, total RNA was extracted from harvested cells following siRNA-mediated knockdown and separated on a glyoxal-agarose gel. Levels of pre-rRNA intermediates were analysed by northern blotting using radiolabelled oligonucleotide probes hybridising to the 5' end of ITS1 ('18SE') and the middle of ITS1 ('ITS1'). Positions of probes on the 47S pre-rRNA are shown in Figure 3.4B, Figure 3.5B and Figure 3.6B. 18SE pre-rRNA levels after UTP24 depletion were quantified separately, normalised to 47S pre-rRNA levels and compared to 18SE levels in control (GL2) samples (HEK293T, Figure 3.4E; U2OS, Figure 3.5E; MCF7, Figure 3.6E). Consistent with previous results in other cell types, depletion of UTP24 in HEK293T (Figure 3.4C), U2OS (Figure 3.5C) and MCF7 (Figure 3.6C) cells caused accumulation of the 30S pre-rRNA and reduction in 18SE levels. The intensity of each band detected with the ITS1 probe was quantified using ImageQuant and normalised to the levels of the 47S primary precursor (HEK293T, Figure 3.4D; U2OS, Figure 3.5D; MCF7, Figure 3.6D). 18SE pre-rRNA levels were quantified and normalised to 47S pre-rRNA levels (Figure 3.4E, Figure 3.5E and Figure 3.6E). Interestingly, the quantification data show a difference in levels of pre-rRNA intermediates in different cell types, even in cells treated with the control siRNA. For example, the ratios of 30S pre-rRNA to 47S pre-rRNA levels vary significantly, with much lower levels of the 30S precursor detected in HEK293T cells compared to U2OS or MCF7. Despite this, 30S pre-rRNA levels were clearly increased compared to control cells upon UTP24 knockdown. These results indicate that the presence of UTP24 is essential for efficient cleavage at sites A0, 1 and 2a in a variety of human cells and that its role in pre-rRNA processing is conserved between yeast, mice and humans, and between multiple human cell lines.

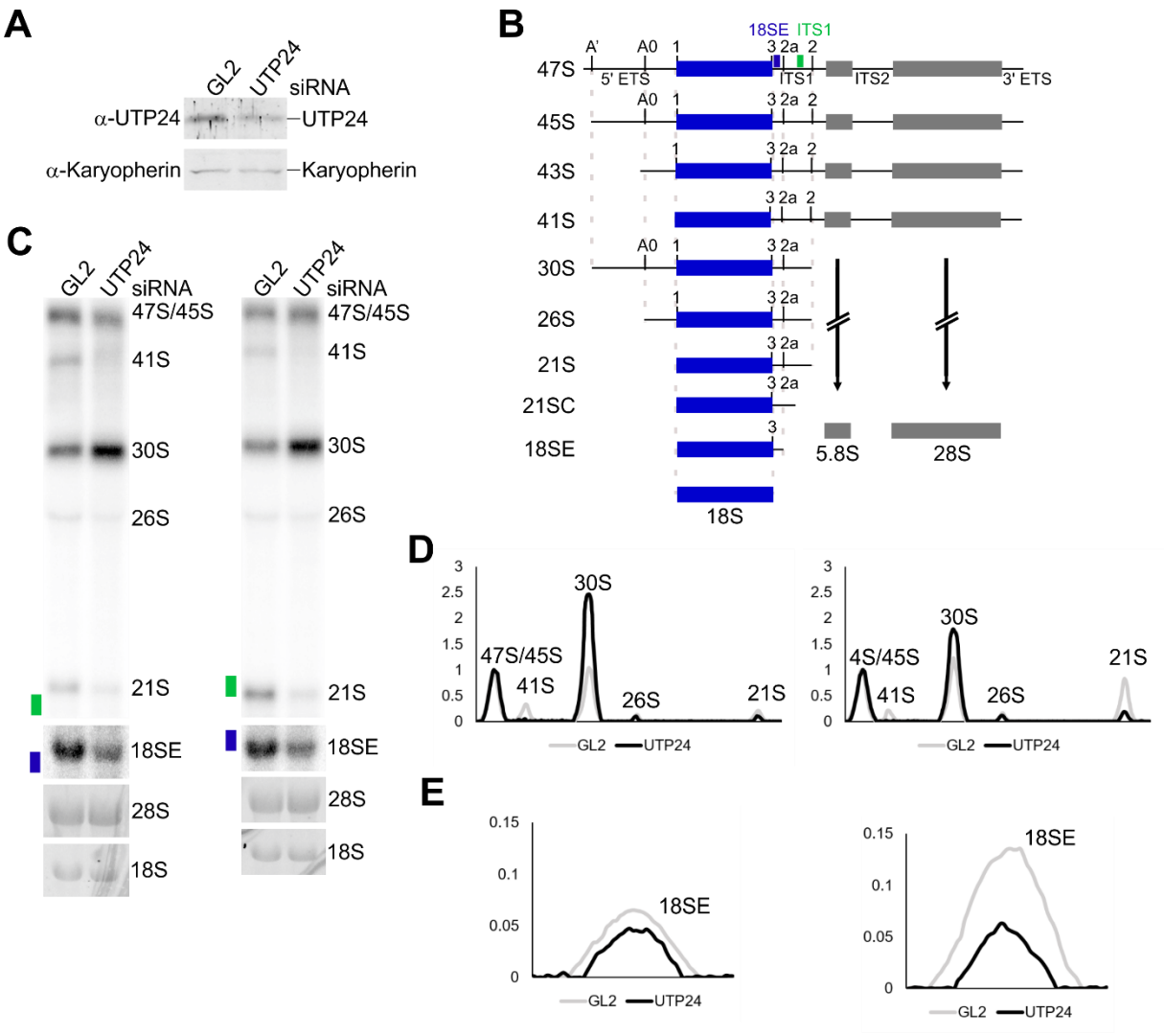


Figure 3.6 UTP24 is required for three early pre-rRNA cleavages in MCF7 cells. **(A)** Western blot using an anti-UTP24 antibody to confirm RNAi-mediated depletion of UTP24 in MCF7 cells. Karyopherin was used as a loading control. **(B)** Schematic of pre-rRNA processing in human cells, showing pre-rRNA intermediates generated from successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(C)** RNA from MCF7 cells treated with control siRNA (GL2) or depleted of UTP24 (UTP24) was analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. Two representative experiments are shown. Levels of pre-rRNA intermediates detected by the **(D)** ITS1 or **(E)** 18SE probe from panel C were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

3.2.2 Creation of stable HEK293T cell lines expressing RNAi-resistant forms of UTP24

While the above results have shown that the presence of UTP24 is required for cleavage at three sites needed for 18S rRNA maturation, the putative catalytic activity of its PIN domain is impossible to investigate by simple RNAi knockdowns. PIN domains contain three or four acidic residues that are important for their catalytic activity (Clissold and Ponting, 2000). In yeast, mutation of putative catalytically-important residues of Utp24 permitted normal cleavage at site A_0, but disrupted cleavages at sites A_1 and A_2 *in vivo* (Bleichert *et al.*, 2006). To test whether mutation of the highly conserved PIN domain of human UTP24 affects pre-rRNA processing, an RNAi-rescue system was established in HEK293T cells. Substitutions were made in the open reading frame (ORF) of UTP24 making it resistant to the siRNA used to deplete UTP24, without altering the amino acid sequence, and then cloned into a modified pcDNA5 plasmid. To produce putative catalytic mutant forms of UTP24, either the first of the four key residues (D72), or both the first and third aspartic acid residues (D72/D142) were substituted for asparagines (D72N and D72N/D142N). Equivalent mutations were shown to disrupt pre-rRNA processing at cleavage sites A_1 and A_2 in yeast *in vivo* (Bleichert *et al.*, 2006) and *in vitro* (Wells *et al.*, 2016). Within the pcDNA5 plasmid, RNAi-resistant UTP24 ORFs were fused to two N-terminal FLAG-tag sequences and were under the control of a tetracycline-inducible promoter. pcDNA5-FLAG-UTP24 constructs were stably transfected into HEK293T cells using the Flp-In recombination system (Life Technologies).

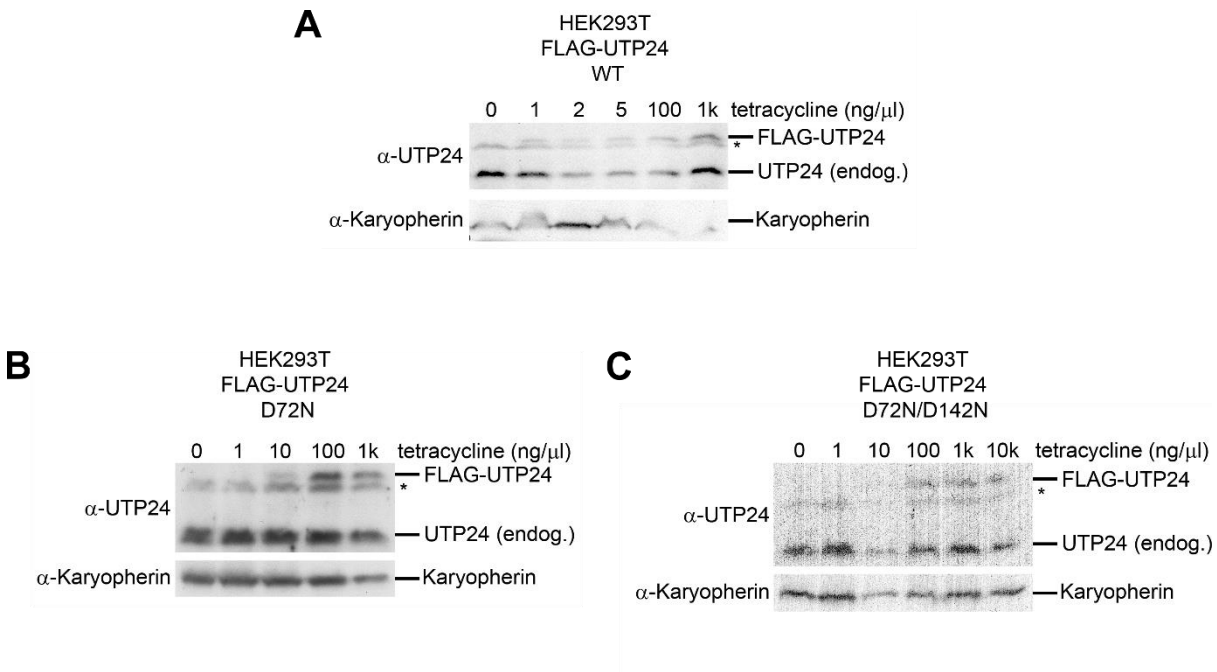


Figure 3.7 Tetracycline induction of FLAG-tagged UTP24 in HEK293T stable cell lines. HEK293T cells stably transfected with wild type (WT; panel A) or mutant (D72N; panel B, D72N/D142N; panel C) UTP24 were incubated with a range of concentrations of tetracycline (0-10,000 ng/µl). Protein was separated by SDS-PAGE and analysed by western blotting using an anti-UTP24 antibody to detect both endogenous and FLAG-tagged UTP24 and an anti-Karyopherin antibody was used as a loading control. Non-specific protein detected by the anti-UTP24 antibody is marked with an asterisk.

After transfection and successful selection of stably transfected cells, expression of FLAG-tagged UTP24 was confirmed by western blotting (Figure 3.7). Cells were incubated with a range of tetracycline concentrations (0-1,000 ng/µl) for 48-72 hours with the aim of identifying the concentration required to express FLAG-tagged UTP24 at similar levels to endogenous UTP24. After tetracycline induction, whole cell extract was separated on an SDS-PAGE gel. UTP24 protein levels were detected using an anti-UTP24 antibody, and karyopherin was used as a loading control. This showed that stably transfected cells efficiently expressed FLAG-tagged UTP24 upon addition of tetracycline, producing an additional band above the endogenous UTP24 band (Figure 3.7A). Expression levels of FLAG-tagged proteins correlated with increasing tetracycline concentration. For both wild type (WT) (Figure 3.7A) and mutant (Figure 3.7B-C) UTP24 however, addition of 1,000 ng/µl tetracycline induced the FLAG-tagged protein at lower levels than endogenous UTP24. Therefore, future experiments used 10,000 ng/µl tetracycline for induction of UTP24.

Once tetracycline-induced expression was confirmed, the RNAi-resistance of the mRNA encoding FLAG-tagged UTP24 was assessed. RNAi-mediated depletion was performed using control (GL2) or UTP24 siRNAs in cells expressing either a FLAG-tag alone (HEK-pcDNA5), a non-resistant FLAG-tagged WT UTP24, or RNAi-resistant FLAG-tagged WT UTP24 construct (Figure 3.8). After 72 hours of RNAi knockdown and tetracycline induction, whole cell extract was separated on an SDS-PAGE gel and protein was analysed by western blotting. The anti-UTP24 antibody was used to detect both endogenous UTP24 and FLAG-tagged UTP24, and karyopherin was used as a loading control. In cells expressing the FLAG-tag alone, UTP24 protein levels were decreased efficiently with the UTP24 siRNA, and no FLAG-tagged protein was produced. In cells expressing a non-resistant UTP24 construct, FLAG-tagged UTP24 was expressed in cells treated with the control (GL2) siRNA, while both endogenous UTP24 and FLAG-tagged UTP24 were efficiently depleted with transfection of the UTP24 siRNA. FLAG-tagged RNAi-resistant UTP24 was expressed efficiently, and levels remained stable with transfection of the UTP24 siRNA, while endogenous UTP24 was efficiently depleted. This showed that RNAi-resistant UTP24 could be expressed while endogenous UTP24 was depleted, allowing for analysis of the effect of UTP24 mutants *in vivo*.

3.2.3 An intact hUTP24 PIN domain is required for cleavage at sites 1 and 2a in HEK293T cells

To assess the potential catalytical role of UTP24 in pre-rRNA processing, an RNAi rescue experiment was set up using cells expressing either a FLAG-tag alone (HEK-pcDNA5), RNAi-resistant WT UTP24, or RNAi-resistant single (D72N) or double (D72N/D142N) UTP24 PIN domain mutants. Cells were incubated with tetracycline and either control (GL2) siRNA or siRNA targeting UTP24, XRN2, or UTP24 and XRN2 together for 72 hours. Depletion of XRN2 was performed to investigate the effect of UTP24 mutations under different processing conditions (see later). Whole cell extract was separated on an SDS-PAGE gel and protein was analysed by western blotting (Figure 3.8). An anti-XRN2 antibody was used to confirm XRN2 depletion, an anti-UTP24 antibody was used to confirm FLAG-tagged UTP24 expression and endogenous UTP24 depletion and karyopherin was used as a loading control. UTP24 and XRN2 were depleted efficiently with UTP24 and XRN2 siRNAs, both alone and together. WT and mutant FLAG-tagged UTP24 were expressed to levels similar to

endogenous UTP24 levels and were not depleted by the UTP24 siRNA. Interestingly, for both UTP24 PIN domain mutant cell lines, depletion of endogenous UTP24 appears to induce increased expression of the FLAG-tagged protein (Figure 3.8).

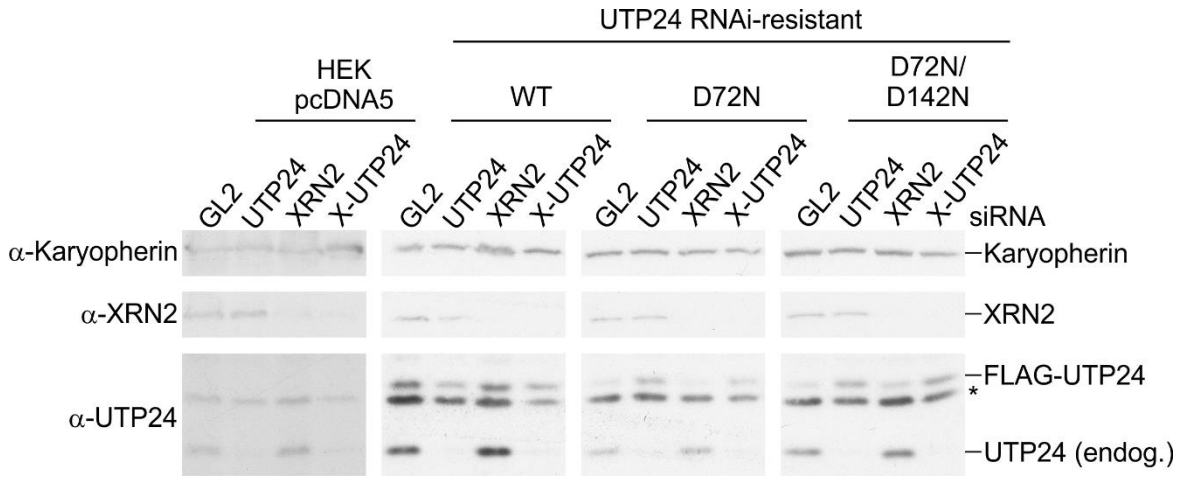


Figure 3.8 Protein analysis of UTP24 HEK293T RNAi-rescue cell lines. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5), or wild type (WT) or mutant (D72N, D72N/D142N) RNAi-resistant UTP24. Following transfection of a control siRNA (GL2) or RNAi-mediated depletion of endogenous UTP24 (UTP24), XRN2 (XRN2) or both (X-UTP24), protein was extracted from cells and separated by SDS-PAGE. Protein was analysed by western blotting using anti-UTP24 and anti-XRN2 antibodies. Anti-Karyopherin was used as a loading control. Non-specific protein detected by the anti-UTP24 antibody is marked with an asterisk.

RNA analysis was performed on RNA extracted from the same cells used in Figure 3.10. Total RNA was extracted from cells following RNAi-mediated knockdown and induction of FLAG-tagged UTP24 and analysed by northern blotting using ‘ITS1’ and ‘18SE’ probes (Figure 3.9A and B). The positions of probes on the 47S pre-rRNA are shown in Figure 3.9A. The intensities of bands detected by the ITS1 probe in Figure 3.8B were quantified, normalised to levels of the 47S pre-rRNA and plotted (Figure 3.9C). In cells expressing the FLAG-tag alone (Figure 3.9B lanes 1 and 2), depletion of UTP24 caused accumulation of the 30S pre-rRNA and reduction of the 41S, 26S, 21S and 18SE pre-rRNAs compared to cells treated with a control (GL2) siRNA. This is the same phenotype observed in non-transfected cells and indicates disrupted cleavage at sites A0, 1 and 2a, consistent with an SSU processome defect. In cells expressing WT RNAi-resistant UTP24 (lanes 3 and 4) expression of this protein was able to rescue the phenotype caused by depletion of endogenous UTP24, as 30S

levels are not increased compared to the GL2 knockdown. In cells expressing either single (D72N, lanes 5 and 6) or double (D72N/D142N, lanes 7 and 8) PIN domain mutant forms of UTP24 the 30S accumulation phenotype seen in control cells is also rescued. Instead, depletion of endogenous UTP24 with expression of these mutants caused significant accumulation of the 26S pre-rRNA, indicating that cleavages at sites 1 and 2a are disrupted but A0 cleavage is unaffected. While the accumulation of the 26S pre-rRNA shows that these cleavages are disrupted, the detection of low levels of 21S suggest that the 5' end of the 18S rRNA is processed by an alternative mechanism when cleavage at site 1 is abolished. It is likely that, in the absence of an endonucleolytic cleavage, the pre-rRNA is trimmed to the 18S rRNA 5' end by a 5' to 3' exonuclease. Levels of a 3'-shortened form of the 21S intermediate (called 21SC) were also accumulated in cells expressing PIN domain mutant forms of UTP24 suggesting that these mutants also disrupt the trimming of ITS1 from site 2 to generate the 18SE pre-rRNA. Based on these results, it seems that an intact UTP24 PIN domain is required for cleavage at sites 1 and 2a, but not A0, while presence of the protein is required for all three early cleavages. This is consistent with yeast data (Bleichert *et al.*, 2006), suggesting that Utp24 may play a conserved role as an endonuclease responsible for cleavage at two sites in yeast and humans.

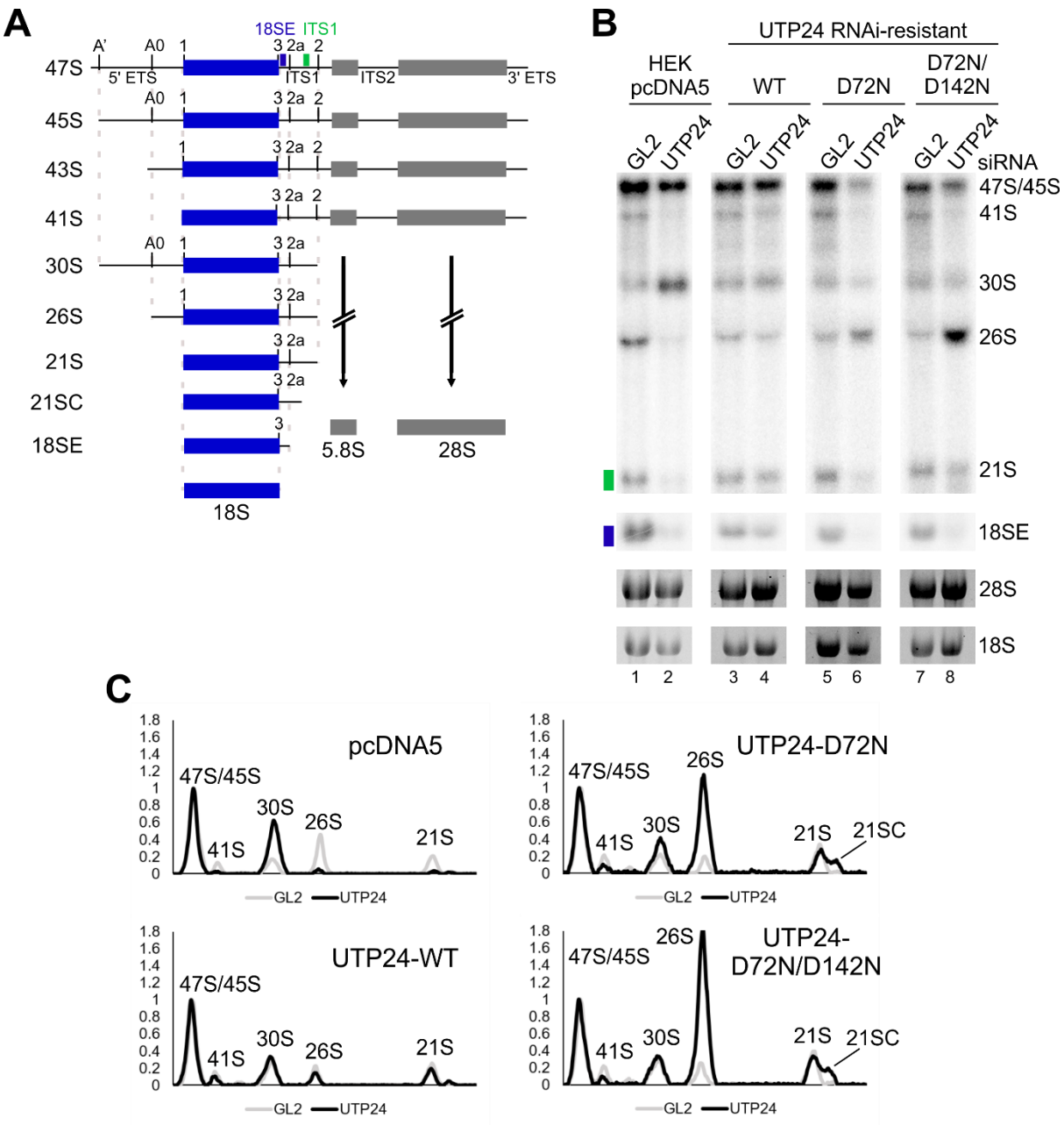


Figure 3.9 An intact PIN domain in UTP24 is required for cleavage at pre-rRNA sites 1 and 2a in HEK293T cells. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5), or wild type (WT) or mutant (D72N, D72N/D142N) RNAi-resistant UTP24. **(A)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive pre-rRNA processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the primary precursor (47S). **(B)** Following treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous UTP24 (UTP24), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane and analysed by northern blotting. Radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle) were used. Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. **(C)** Levels of pre-rRNAs detected by the ITS1 probe from panel B were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

Despite an apparent processing defect at sites 1 and 2a upon expression of a UTP24 PIN domain mutant, the effect on cleavage at the 2a site specifically cannot be readily assessed in human cells under normal conditions. Unlike in yeast, where cleavage at the equivalent site (A_2) is a major processing event, cleavage at the 2a site in human cells is part of a minor processing pathway. In the major human pre-rRNA processing pathway, ITS1 is cleaved at site 2, followed by 3'-5' exonucleolytic trimming to produce the 18SE precursor. Therefore, cleavage at site 2a is not an essential step in normal human cells, meaning that "2a-specific" precursors, the 36S pre-rRNA and the ITS1 fragment (an excised fragment between sites 2a and 2), are barely detectable under normal conditions. However, it is possible to promote pre-rRNA processing through the minor pathway, in which cells are reliant on cleavage at site 2a, by depletion of the 5'-3' exonuclease, XRN2. RNA isolated from cells transfected with either an XRN2 siRNA alone (see Figure 3.8), or XRN2 and UTP24 siRNAs together, was used for northern blot analysis using 'ITS1' and '18SE' probes (Figure 3.10). Depletion of XRN2 alone caused an accumulation of the 36S pre-rRNA and the ITS1 fragment, indicating that cleavage at site 2a takes place in cells used in this experiment (Figure 3.10B lanes 1, 3, 5 and 7).

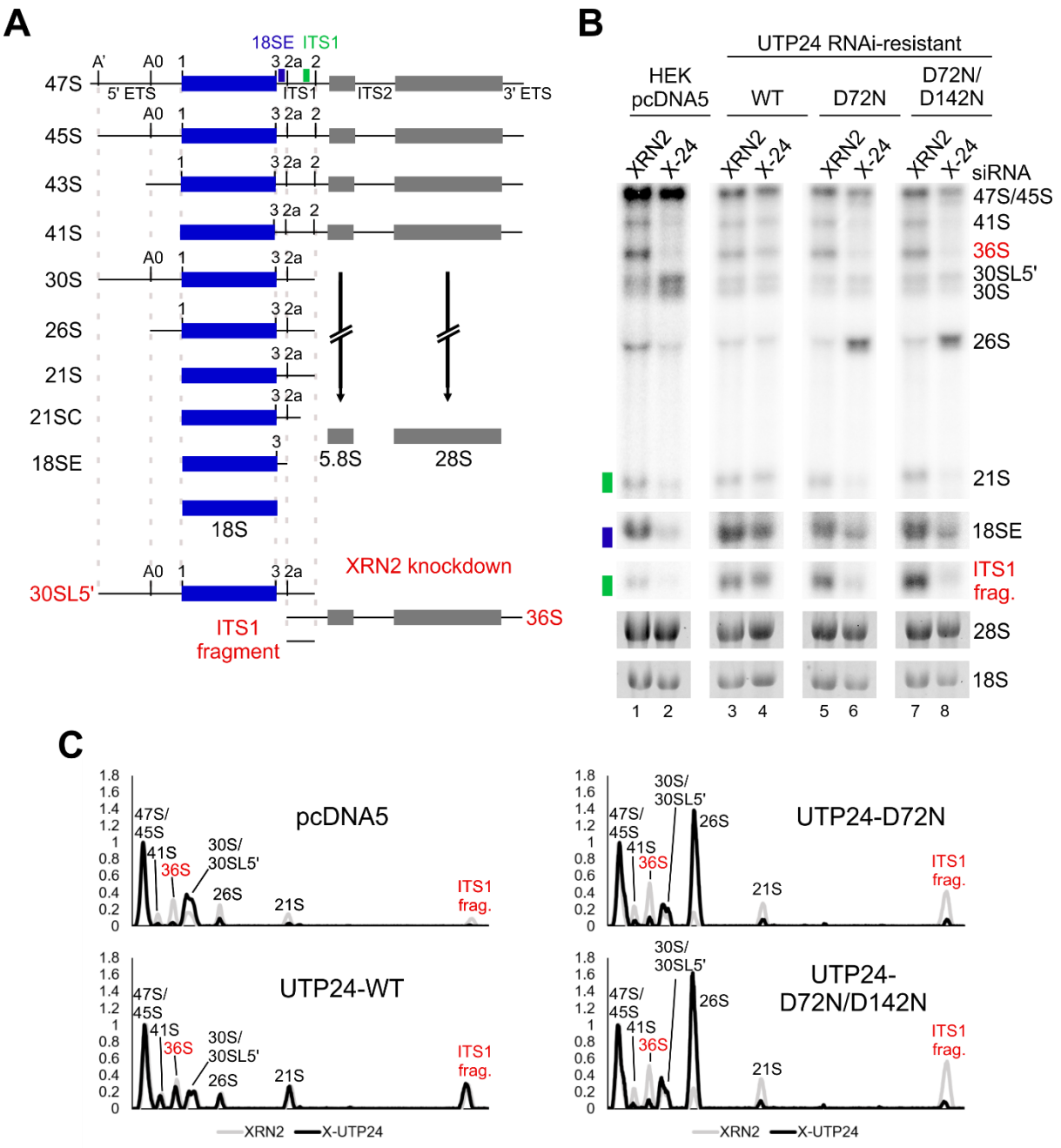


Figure 3.10 An intact PIN domain in UTP24 is required for cleavage at pre-rRNA sites 1 and 2a in HEK293T cells in the absence of XRN2. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5), or wild type (WT) or mutant (D72N, D72N/D142N) RNAi-resistant UTP24. **(A)** Schematic of pre-rRNA processing in human cells, showing pre-rRNA intermediates generated by successive processing events. Pre-rRNA intermediates specific to the minor, 2a-dependent processing pathway, which accumulate upon depletion of XRN2, are shown in red. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(B)** Following treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous UTP24 (UTP24), XRN2 (XRN2) or both (X-24), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane. RNA was analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot

probes are labelled. **(C)** Levels of pre-rRNAs detected by the ITS1 probe from panel B were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted. The pre-rRNA intermediate represented by each peak is labelled.

In cells expressing the FLAG-tag alone, depletion of UTP24 in addition to depletion of XRN2 caused a significant reduction in the 2a-specific precursors, 36S pre-rRNA and ITS1 fragment (Figure 3.10B lane 2), suggesting a disruption in cleavage at site 2a specifically. This is consistent with the role of UTP24 role as a key component of the SSU processome, as this complex is required for cleavage at site 2a. In cells expressing RNAi-resistant WT UTP24, levels of the 36S pre-rRNA and ITS1 fragment are recovered with depletion of both UTP24 and XRN2 (lane 4) compared to cells expressing the FLAG-tag alone, indicating that expression of WT UTP24 was able to rescue the specific "2a" defect. Expression of a single (D72N; lanes 5 and 6) or double (D72N/D142N; lanes 7 and 8) PIN domain mutant UTP24, upon depletion of both endogenous UTP24 and XRN2 (lanes 6 and 8), was unable to rescue the 2a-specific phenotype caused by XRN2 depletion. Levels of 36S and the ITS1 fragment were only detected at very low levels in these cells, suggesting that cleavage at site 2a specifically is disrupted in these cells. This result suggests that the putative catalytic residues in the PIN domain of UTP24 are essential for cleavage at site 2a in human cells. Interestingly, depletion of XRN2 and UTP24 in cells expressing PIN domain mutant forms of UTP24 caused a reduction in levels of the 21S intermediate compared to depletion of UTP24 alone (Figure 3.10), supporting a role for XRN2 in the alternative processing of the 18S rRNA 5' end**.** These *in vivo* data point towards a role for UTP24 as an endonuclease in cleavage at sites 1 and 2a in the maturation of the human 18S rRNA. In addition, *in vitro* data from our lab show that both yeast and human Utp24 were able to cleave a pre-rRNA mimic containing the yeast A_2 cleavage site (Wells *et al.*, 2016).

3.2.4 Human RCL1 is required for 18S rRNA release in HEK293T cells

The likely role of Utp24 as the endonuclease responsible for cleavage at two pre-rRNA sites in yeast and humans, particularly at site A_2/2a, is surprising, as yeast Rcl1 was previously reported to be responsible for cleavage at the yeast A_2 cleavage site (Horn *et al.*, 2011). Rcl1 depletion led to accumulation of the 23S pre-rRNA *in vivo*, caused by disrupted cleavage at sites A_0, A_1 and A_2 (Billy *et al.*, 2000). Rcl1 was also reported to cleave a pre-rRNA mimic at the yeast A_2 site *in vitro*, and this catalytic

activity was abolished by mutation of three residues (R327A, D328A, K330A) predicted to be involved in RNA substrate binding (Horn *et al.*, 2011). Human RCL1 is required for the equivalent three cleavage events, at sites A0, 1 and 2a in HeLa cells (Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013), consistent with a conserved role as a component of the SSU processome in humans. While these data and further data obtained in our lab strongly suggest a role for Utp24 in cleavage at the A_2/2a site, there remains the possibility of redundant activities for the two proteins at this site.

To further investigate the role of RCL1 in human pre-rRNA processing, RNAi-mediated knockdowns were performed in HEK293T cells. A siRNA targeting the mRNA of RCL1 was used to deplete RCL1, and a firefly luciferase siRNA (GL2) was used as a control. After 72 hours incubation with the siRNA, cells were harvested and whole cell extract was loaded on an SDS-PAGE gel and protein levels were analysed by western blotting using an anti-RCL1 antibody (Figure 3.11A). Karyopherin was used as a loading control. Western blotting showed that RCL1 was efficiently depleted in these cells. To analyse the effect of RCL1 depletion on pre-rRNA processing, RNA was isolated from harvested cells following RNAi-mediated knockdown with either the control siRNA or the RCL1 siRNA. RNA was analysed by northern blotting using radiolabelled oligonucleotide probes hybridising to the 5' end of ITS1 ('18SE') or around the middle of ITS1 ('ITS1') (Figure 3.10C). The probes used for northern analysis are marked above the 47S pre-rRNA on the processing scheme (Figure 3.11B).

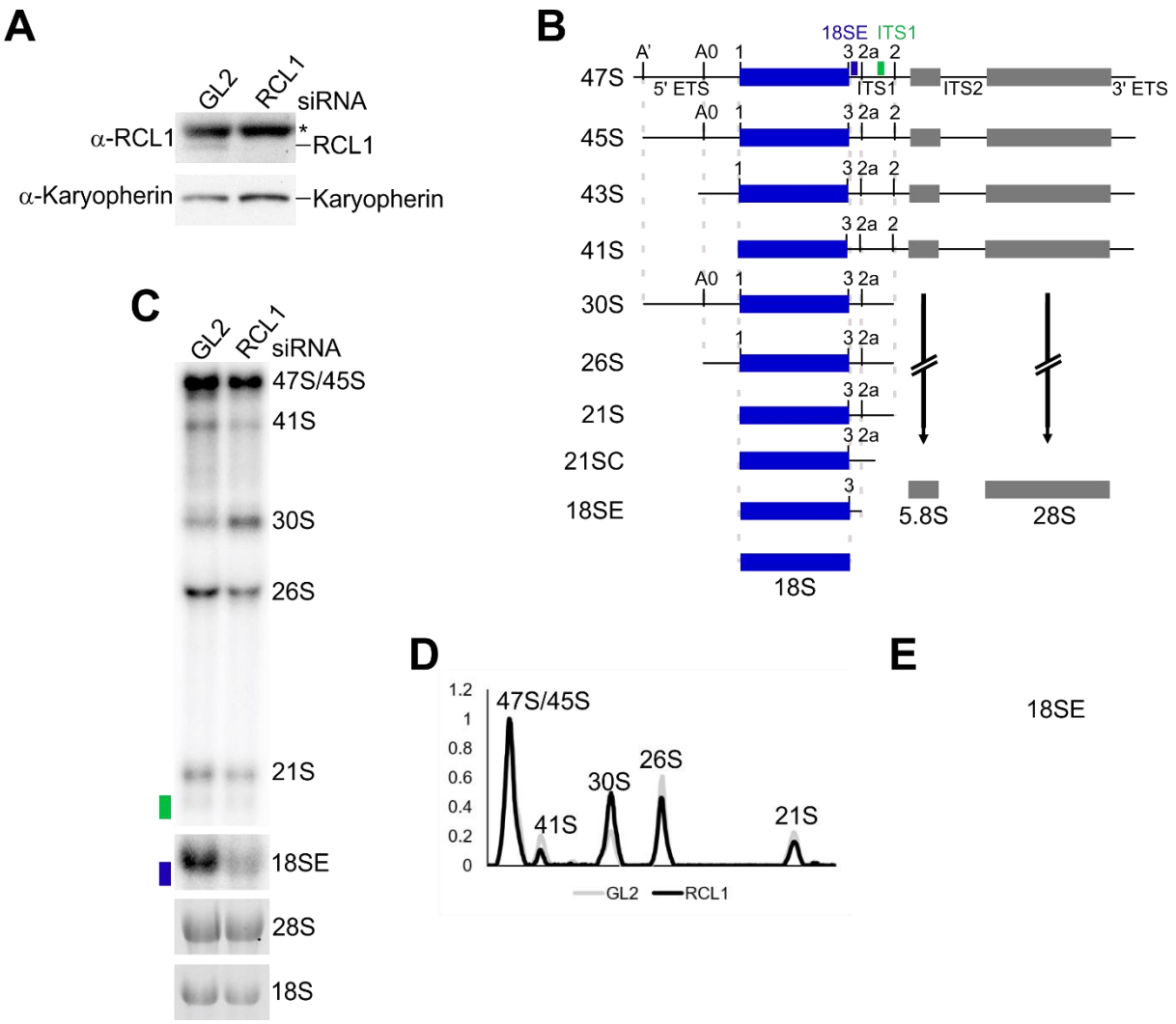


Figure 3.11 RCL1 is required for three early pre-rRNA processing cleavages in HEK293T cells. (A) Western blot using an anti-RCL1 antibody to confirm the RNAi-mediated depletion of endogenous RCL1 in HEK293T cells. An anti-Karyopherin antibody was used as a loading control. **(B)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(C)** RNA from HEK293T cells treated with control siRNA (GL2) or depleted of RCL1 (RCL1) was analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining. Pre-rRNA intermediates detected by ITS1 and 18SE probes are labelled. Levels of pre-rRNAs detected by the **(D)** ITS1 or **(E)** 18SE probe from panel C were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

Depletion of RCL1 in HEK293T cells caused the accumulation of the 30S pre-rRNA and a reduction in levels of the 26S, 21S and 18SE pre-rRNAs, indicating disrupted cleavages at sites A0, 1 and 2a. The intensities of bands detected by the "ITS1" (Figure 3.11D) and "18SE" (Figure 3.11E) probes were quantified, normalised

to the 47S pre-rRNA and plotted on a graph. Attempts to assess the effect of RCL1 depletion on pre-rRNA processing in U2OS and MCF7 cells were unsuccessful (data not shown). The phenotypes observed in these two cell lines were variable and attempts to confirm RNAi knockdown of RCL1 by western blotting failed due to the poor quality of the available antibody. However, the pre-rRNA processing phenotype observed upon RCL1 depletion in HEK293T cells is consistent with yeast data and suggests that the role of Rcl1 in 18S rRNA maturation as part of the SSU processome is conserved.

3.2.5 Creation of stable HEK293T cell lines expressing RNAi-resistant forms of RCL1

In the yeast studies which reported that Rcl1 was able to cleave at the yeast A_2 site *in vitro*, truncation analysis was also performed in an attempt to identify the catalytic site of its putative nuclease activity (Horn *et al.*, 2011). This revealed that residues 309-341 are important for yeast growth, and subsequent mutational analysis within this region identified that mutation of three residues (R327A/D328A/K330A) ("RDK") in the C-terminus caused reduced growth. *In vivo*, these mutants caused pre-rRNA processing defects at sites A_0, A_1 and particularly A_2. The C-terminal Arg-327, Asp-328 and Lys-330 residues were predicted to be involved in pre-rRNA substrate binding. To investigate the potential enzymatic role of human RCL1, stable HEK293T cell lines were generated expressing equivalent mutant forms of RCL1.

Substitutions were made in the ORF of RCL1 to make it resistant to the siRNA used to deplete endogenous RCL1, without altering the amino acid sequence, and cloned into the pcDNA5 plasmid. RCL1 expressed from the pcDNA5 plasmid is under the control of a tetracycline inducible promoter and FLAG-tagged. A pcDNA5-RCL1 mutant construct was produced expressing RCL1 truncated at amino acid 313 (RCL1 S313). This is equivalent to the Rcl1-309 truncation which caused growth and pre-rRNA processing phenotypes in yeast (Horn *et al.*, 2011)**.** The "RDK" residues mutated in the yeast study are not completely conserved in the human protein. A construct expressing RCL1 with the equivalent residues, R329, H330 and K332 ("RHK"), substituted for alanines was produced. WT, S313 and RHK pcDNA5 constructs were stably transfected into HEK293T Flp-In cells using the Flp-In recombination system (Life Technologies).

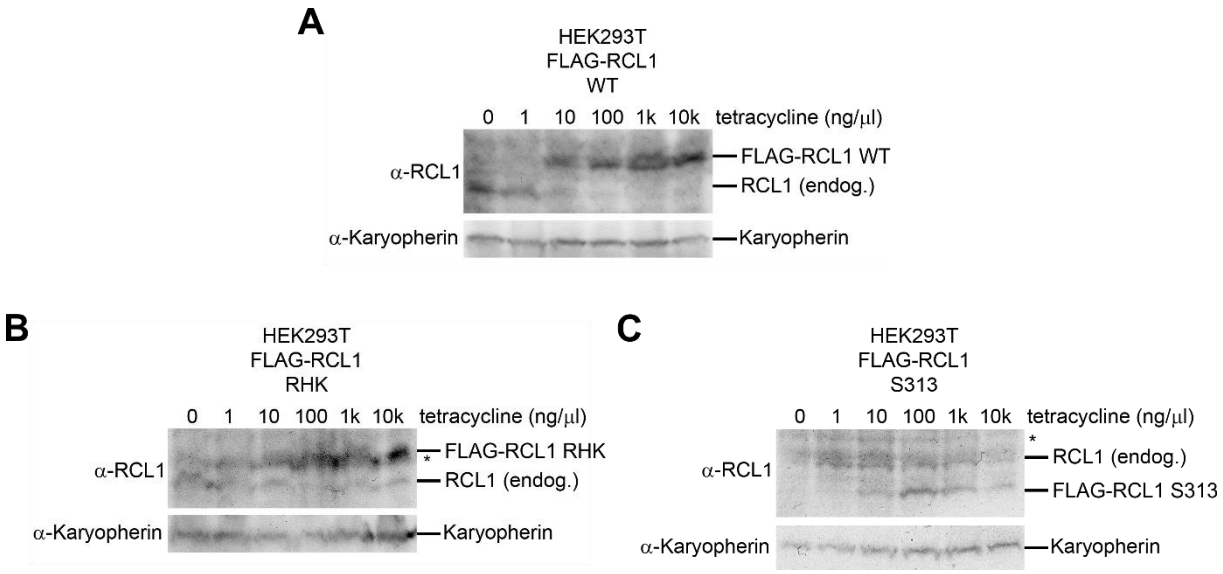


Figure 3.12 Tetracycline induction of FLAG-tagged RCL1 in HEK293T stable cell lines. A range of tetracycline concentrations (0-10,000 ng/µl) were added to HEK293T cells stably transfected with wild type (WT; panel A) or mutant (RHK; panel B, S313; panel C) RCL1. Protein was separated by SDS-PAGE and analysed by western blotting using an anti-RCL1 antibody to detect both endogenous and FLAG-tagged RCL1. An anti-Karyopherin antibody was used as a loading control. Non-specific protein detected by the anti-RCL1 antibody is marked with an asterisk.

Expression of FLAG-tagged RCL1 was induced in stable cell lines using a titration of tetracycline concentrations (0-10,000 ng/µl), aiming to achieve expression of WT and mutant FLAG-tagged RCL1 at levels similar to those of endogenous RCL1. After induction for 48-72 hours, whole cell extract was loaded on an SDS-PAGE gel and protein levels were analysed by western blotting (Figure 3.12). An anti-RCL1 antibody was used to detect levels of both endogenous RCL1 and FLAG-tagged forms of RCL1. Karyopherin was used as loading control. All FLAG-tagged, RNAi-resistant forms of RCL1 were efficiently expressed with addition of tetracycline. Interestingly, expression of FLAG-tagged WT RCL1 (Figure 3.12A), but not RHK (Figure 3.12B) or S313 (Figure 3.12C) mutants, caused depletion of the endogenous protein. The appropriate tetracycline concentration was used for all further experiments. Once expression of FLAG-tagged RCL1 was confirmed, it was necessary to ensure that the expressed proteins were resistant to the RCL1 siRNA. Control cells expressing the FLAG-tag alone (HEK-pcDNA5) or RNAi-resistant RCL1 were incubated with 10 ng/µl tetracycline in conjunction with reverse transfection of either a control (GL2) or RCL1 siRNA for 72 hours. Whole cell extract was loaded on an SDS-PAGE gel and levels of both FLAG-tagged and endogenous RCL1 were analysed by western blotting using an

anti-RCL1 antibody (Figure 3.13). Karyopherin was used as a loading control. Endogenous RCL1 was efficiently depleted with the RCL1 siRNA in control cells. FLAG-tagged RCL1 WT was expressed from the stable cell line at levels similar to that of endogenous RCL1 in control cells and was not depleted by the RCL1 siRNA. This confirmed the RNAi-resistance of FLAG-tagged RCL1 expressed from stable cell lines, allowing me to examine the effects of RCL1 mutants on pre-rRNA processing.

3.2.6 Establishment of an RNAi Rescue system to test the effect of an RCL1 C-terminal truncation and the RCL1 RHK mutant in HEK293T cells

To test the importance of the C-terminal region of RCL1 and the "RHK" residues specifically, on pre-rRNA processing, RNAi rescue systems were set up. Stably-transfected HEK293T cells expressing FLAG-tagged WT RCL1, RCL1 S313 or RCL1 RHK were incubated with the appropriate tetracycline concentration and either a control (GL2) siRNA or a siRNA targeting RCL1, XRN2, or both RCL1 and XRN2 for 72 hours. Depletion of XRN2 was performed to assess effects on processing through the "minor" human pre-rRNA processing pathway (see later). Protein was analysed by western blotting using an anti-RCL1 antibody to detect RCL1 depletion and FLAG-tagged RCL1 expression and an anti-XRN2 antibody was used to confirm XRN2 knockdown (Figure 3.13). Karyopherin or UTP24 was used as a loading control. This confirmed that RCL1 and XRN2 were efficiently depleted with transfection of the respective siRNAs, and WT and mutant forms of RCL1 were expressed at levels similar to endogenous RCL1.

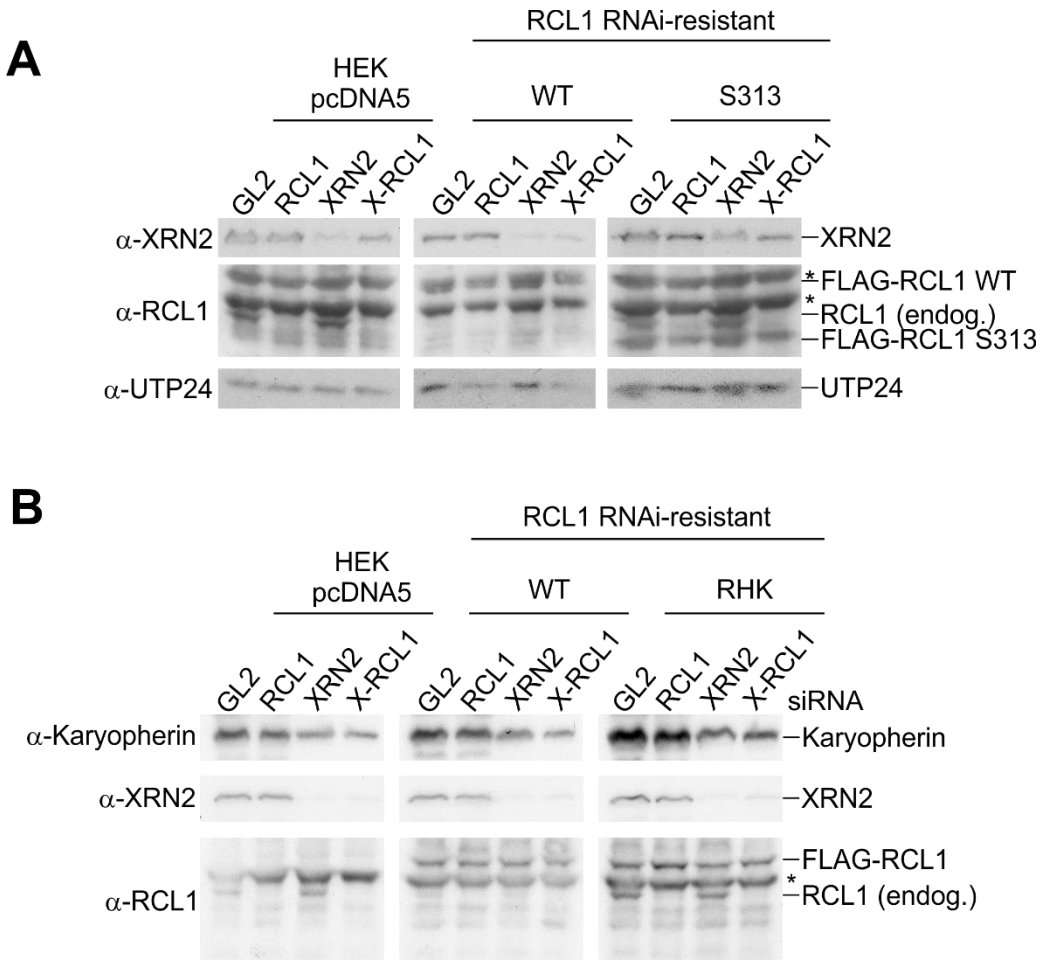


Figure 3.13 Protein analysis of RCL1 HEK293T RNAi-rescue cell lines. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5), or wild type (WT) or mutant **(A)** S313, **(B)** RHK) RNAi-resistant RCL1. Following transfection of a control siRNA (GL2) or RNAi-mediated depletion of endogenous RCL1 (RCL1), XRN2 (XRN2) or both (X-RCL1), protein was extracted from cells and separated by SDS-PAGE. Protein was analysed by western blotting using anti-RCL1 and anti-XRN2 antibodies and Karyopherin was used as a loading control. Non-specific proteins detected by the anti-RCL1 antibody are marked with an asterisk.

3.2.7 A C-terminal truncation of human RCL1 disrupts pre-rRNA processing, likely by affecting SSU processome assembly

To investigate the importance of the C-terminal of RCL1 in pre-rRNA processing total RNA was isolated from the same cells used in Figure 3.16 and northern blotting analysis was performed using 'ITS1' and '18SE' probes (Figure 3.14B). Probes are marked above the 47S pre-rRNA on the pre-rRNA processing pathway scheme (Figure 3.14A). Quantified pre-rRNA intermediate band intensities are shown in Figure 3.17C. In cells expressing RNAi-resistant WT RCL1 (Figure 3.14B, lanes 3 and 4), fully

rescued the 30S pre-rRNA accumulation and 18SE pre-rRNA reduction caused by depletion of RCL1 in control cells expressing the FLAG-tag alone (lanes 1 and 2). Depletion of endogenous RCL1 in cells expressing the RCL1 S313 truncation (lanes 5 and 6) caused accumulation of the 30S pre-rRNA and reduction of 18SE pre-rRNA levels. This indicates a disruption in cleavage at sites A0, 1 and 2a and is the same phenotype observed upon RCL1 depletion in control cells, suggesting that this region of RCL1 is important for RCL1's role in pre-rRNA processing at these sites as part of the SSU processome. It may be that this portion of the protein is important for the incorporation of RCL1 into the SSU processome, although this has not yet been investigated.

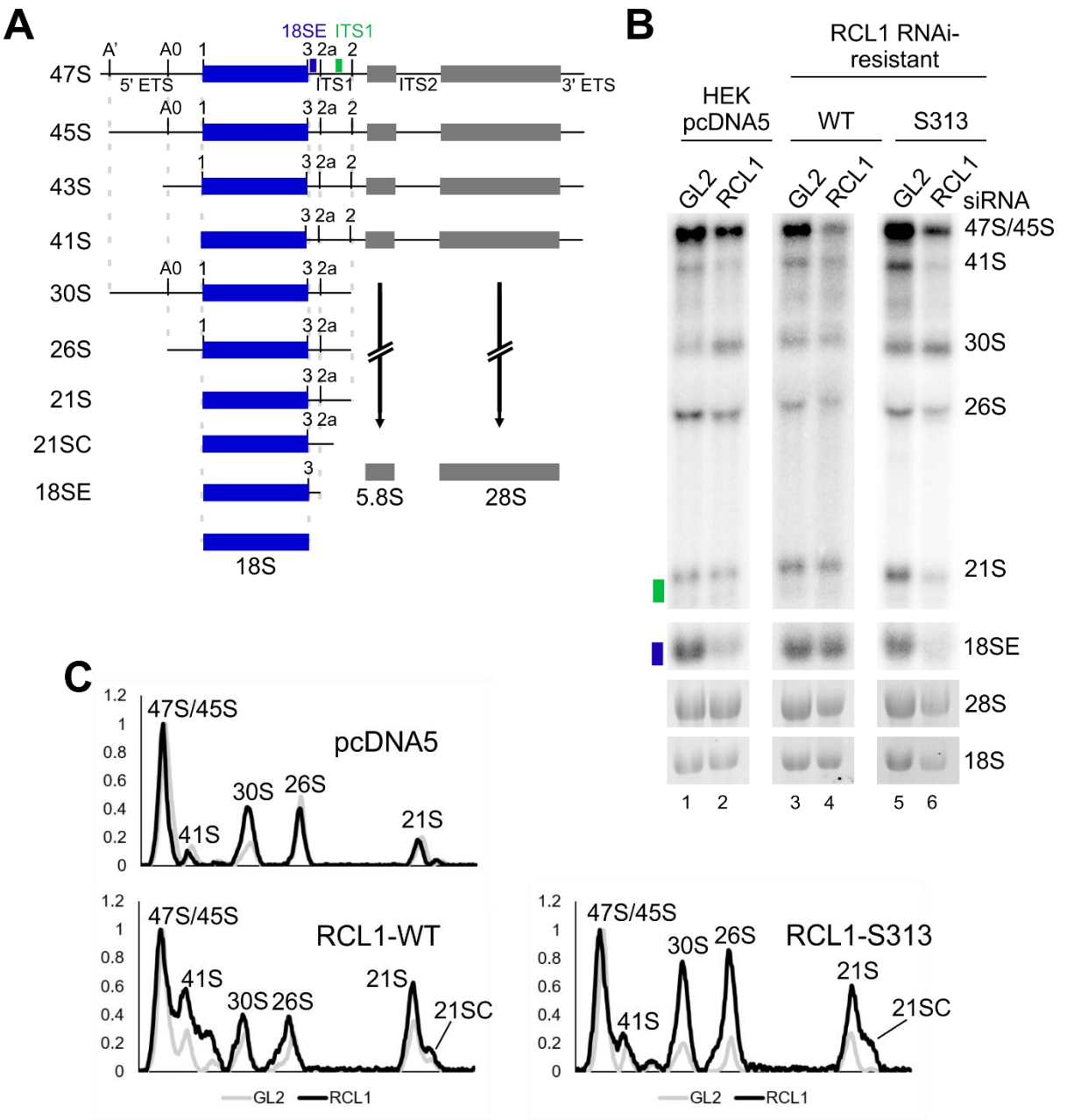


Figure 3. 14 The C-terminus of RCL1 is important for pre-rRNA processing in HEK293T cells. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5), or wild type (WT) or mutant (S313) RNAi-resistant UTP24. **(A)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(B)** Following treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous RCL1 (RCL1), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane. RNA was analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. **(C)** Levels of pre-rRNAs detected by the ITS1 probe from panel B were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

Rcl1 was reported to cleave pre-rRNA at the yeast A_2 site, which is a major event in 18S rRNA maturation in yeast. However, in human cells, the equivalent site (2a) is only cleaved as part of a minor pre-rRNA processing pathway, so pre-rRNA intermediates produced through cleavage at site 2a are barely detectable under normal conditions. To analyse the importance of the C-terminus of RCL1 on site 2a cleavage specifically, pre-rRNA processing through the minor pathway can be promoted by depletion of the 5'-3' exonuclease, XRN2. Northern blotting was therefore performed using RNA from cells depleted of XRN2, both alone and in conjunction with RCL1 knockdown, in cells expressing RCL1 S313 (Figure 3.15B). Depletion of XRN2 alone caused accumulation of the "2a-specific" intermediates, the 36S pre-rRNA and ITS1 fragment, indicating a shift to the 'minor' processing pathway, as cells rely on site 2a cleavage. In control cells depleted of both RCL1 and XRN2, the 36S pre-rRNA and ITS1 fragment levels were reduced to barely detectable levels, as in control cells processed via the 'major' pathway. This indicates that RCL1 plays an important role in 2a cleavage specifically. Expression of RNAi-resistant WT RCL1 rescued this phenotype, with "2a-specific" precursors recovered (lanes 3 and 4).

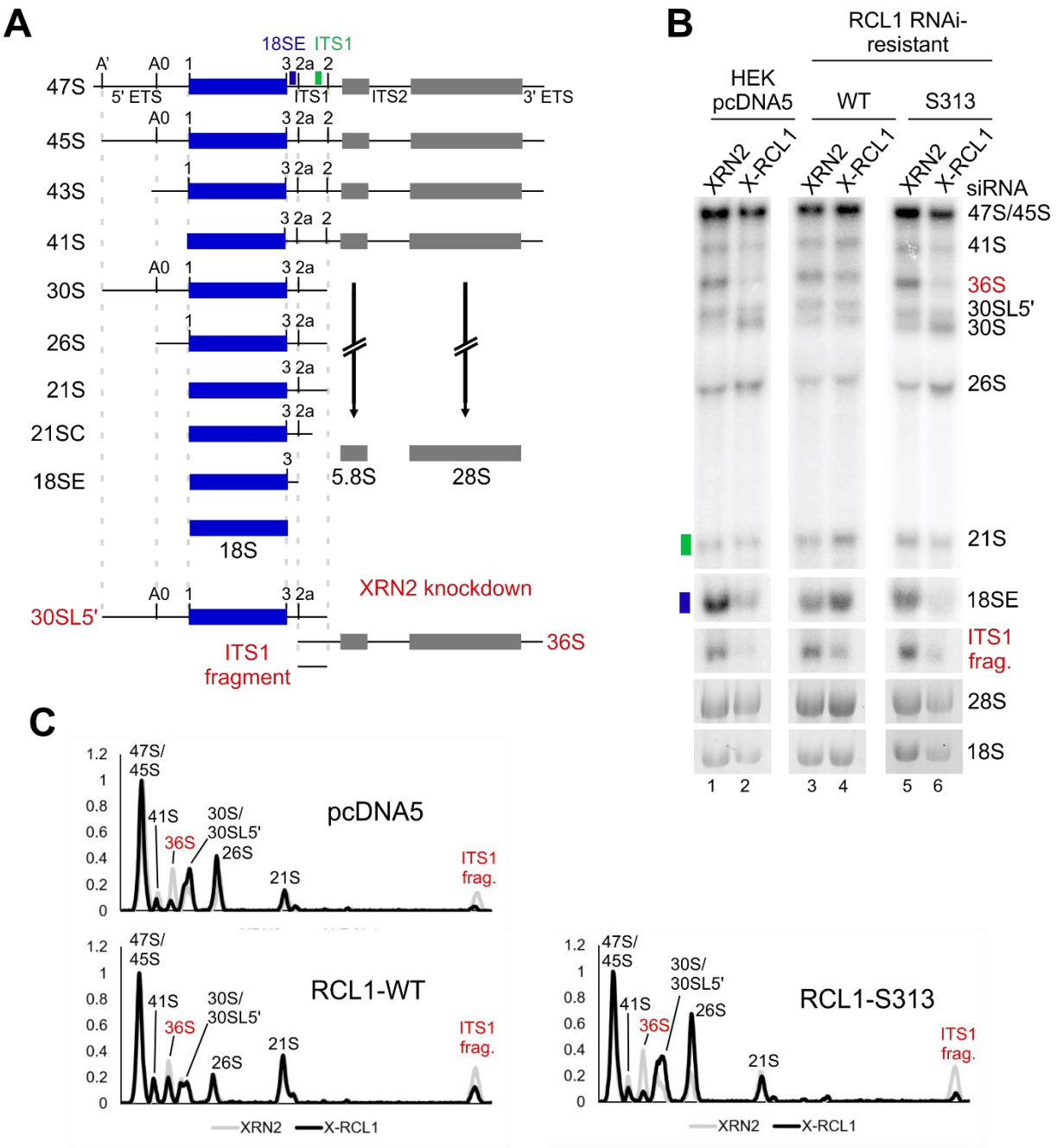


Figure 3.15 The C-terminus of RCL1 is important for pre-rRNA processing in HEK293T cells in the absence of XRN2. Northern analysis of RNA from HEK293T Flp-In cells stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5), or wild type (WT) or mutant (S313) RNAi-resistant RCL1. **(A)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive endonucleolytic or exonucleolytic processing events. Pre-rRNA intermediates specific to the minor, site 2a-dependent processing pathway, which accumulate upon depletion of XRN2, are shown in red. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(B)** Following RNAi-mediated depletion of endogenous XRN2 (XRN2) or both endogenous RCL1 and XRN2 (X-RCL1), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane. RNA was analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot

probes are labelled. **(C)** Levels of pre-rRNAs detected by the ITS1 probe from panel B were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted. The pre-rRNA intermediate represented by each peak in the plot is labelled.

Depletion of both XRN2 and endogenous RCL1 in cells expressing RNAi-resistant RCL1 S313 caused a reduction in levels of both 36S pre-rRNA and the ITS1 fragment, suggesting that the C-terminal region of RCL1 is important for cleavage at the 2a cleavage site. While levels of 2a-specific pre-rRNA intermediates were reduced with RCL1 S313, the reduction was not as significant as observed for the RCL1 knockdown in control cells (pcDNA5), or for UTP24 PIN domain mutants, suggesting that 2a cleavage may not be as severely affected by this mutant. However, this could also be due to a less efficient knockdown of XRN2 in this experiment. Western blotting revealed that XRN2 levels remained higher in the double RCL1-XRN2 knockdown than in the single XRN2 knockdown, potentially explaining the less severe phenotype (Figure 3.13A). The intensities of bands from Figure 3.15B were quantified, normalised to levels of the 47S pre-rRNA and plotted on graphs (Figure 3.15C). These results suggest that the C-terminal portion of RCL1 truncated here are important for the essential role of RCL1 in pre-rRNA processing. Further work is needed to understand this finding, but it likely relates to the incorporation of RCL1 into the SSU processome.

3.2.8 RCL1 RHK residues are not important for pre-rRNA processing in HEK293T cells

While the above results have confirmed that the C-terminus of human RCL1, consistent with previous results in yeast, is important for RCL1's role in pre-rRNA processing at sites A0, 1 and 2a, characterisation as RCL1 as an endonuclease requires the identification of specific sites that may be involved in its putative nuclease activity. While no known catalytic active site in yeast Rcl1 has been identified, three residues within the C-terminal region of the yeast protein were shown to be important for pre-rRNA processing *in vivo* and for Rcl1's reported cleavage activity *in vitro* (Horn *et al.*, 2011). To assess whether the equivalent residues in human RCL1 are similarly important for pre-rRNA processing, an RNAi rescue experiment was set up using cells expressing the equivalent mutant (RCL1 RHK). Total RNA was isolated from cells used in Figure 3.13B and levels of pre-rRNA intermediates were analysed by northern blotting using 'ITS1' and '18SE' probes (Figure 3.16B). Northern blotting probes are marked above the 47S pre-rRNA in Figure 3.16A.

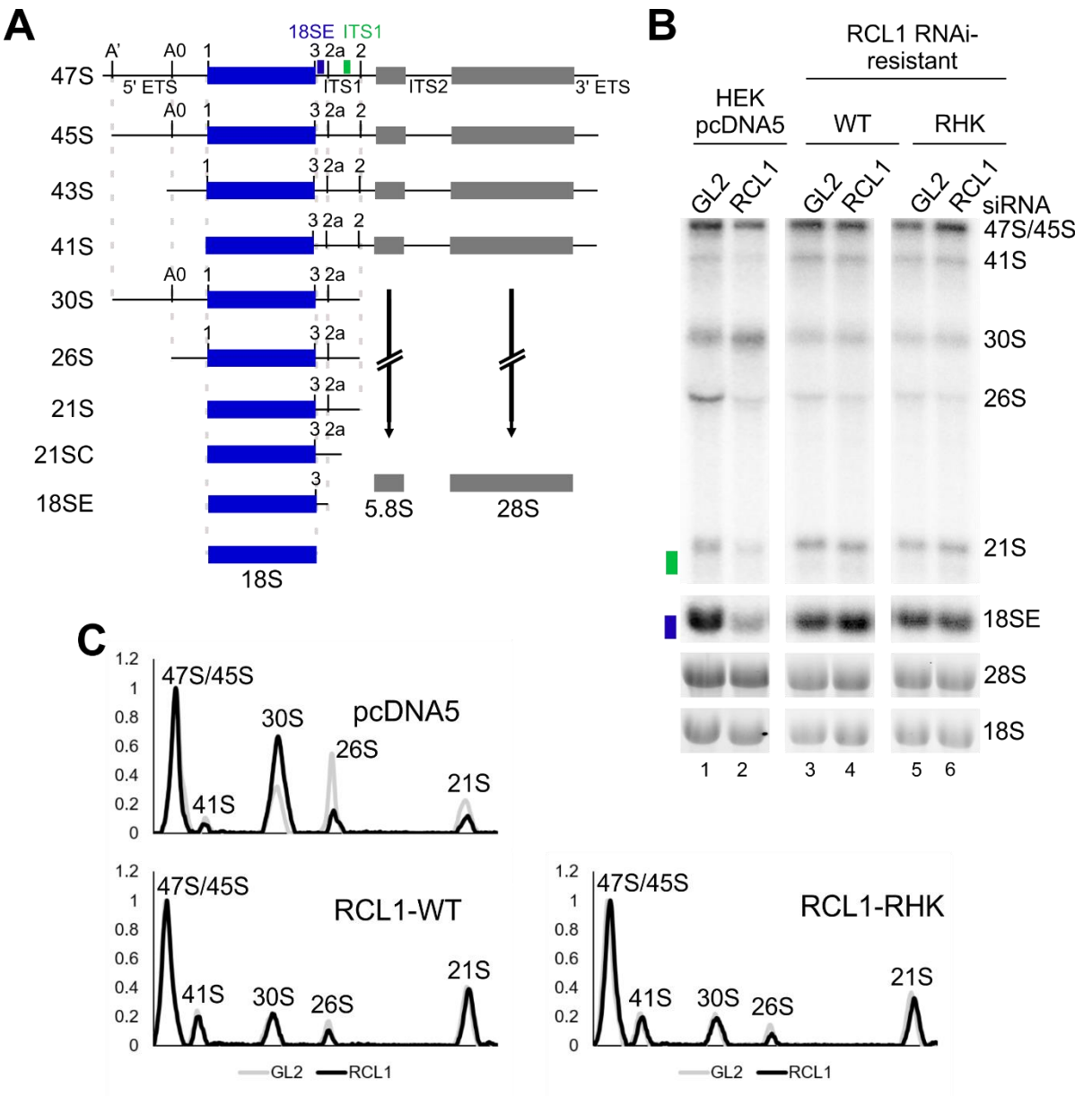


Figure 3.16 RCL1 RHK residues are not important for pre-rRNA processing in HEK293T cells. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5), or wild type (WT) or mutant (RHK) RNAi-resistant RCL1. **(A)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(B)** Following treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous RCL1 (RCL1), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane and analysed by northern blotting. Radiolabelled oligonucleotide probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle) were used. Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. **(C)** Levels of pre-rRNAs detected by the ITS1 probe from panel B were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

In cells expressing the FLAG-tag alone, depletion of endogenous RCL1 caused accumulation of the 30S pre-rRNA and a reduction in levels of the 18SE pre-rRNA (Figure 3.16B, lanes 1 and 2), consistent with its role in pre-rRNA processing at sites A0, 1 and 2a as a component of the SSU processome. In cells expressing FLAG-tagged RNAi-resistant RCL1 RHK, no pre-rRNA processing phenotype was observed upon depletion of endogenous RCL1 (lanes 5 and 6). The intensities of the pre-rRNA bands in Figure 3.16B were quantified and normalised to the intensity of the 47S pre-rRNA band (Figure 3.16C). The lack of 30S pre-rRNA accumulation in cells expressing RCL1 RHK suggests that the R329, H330 and K332 residues of RCL1 are not important for RCL1's role in pre-rRNA processing at sites A0, 1 and 2a.

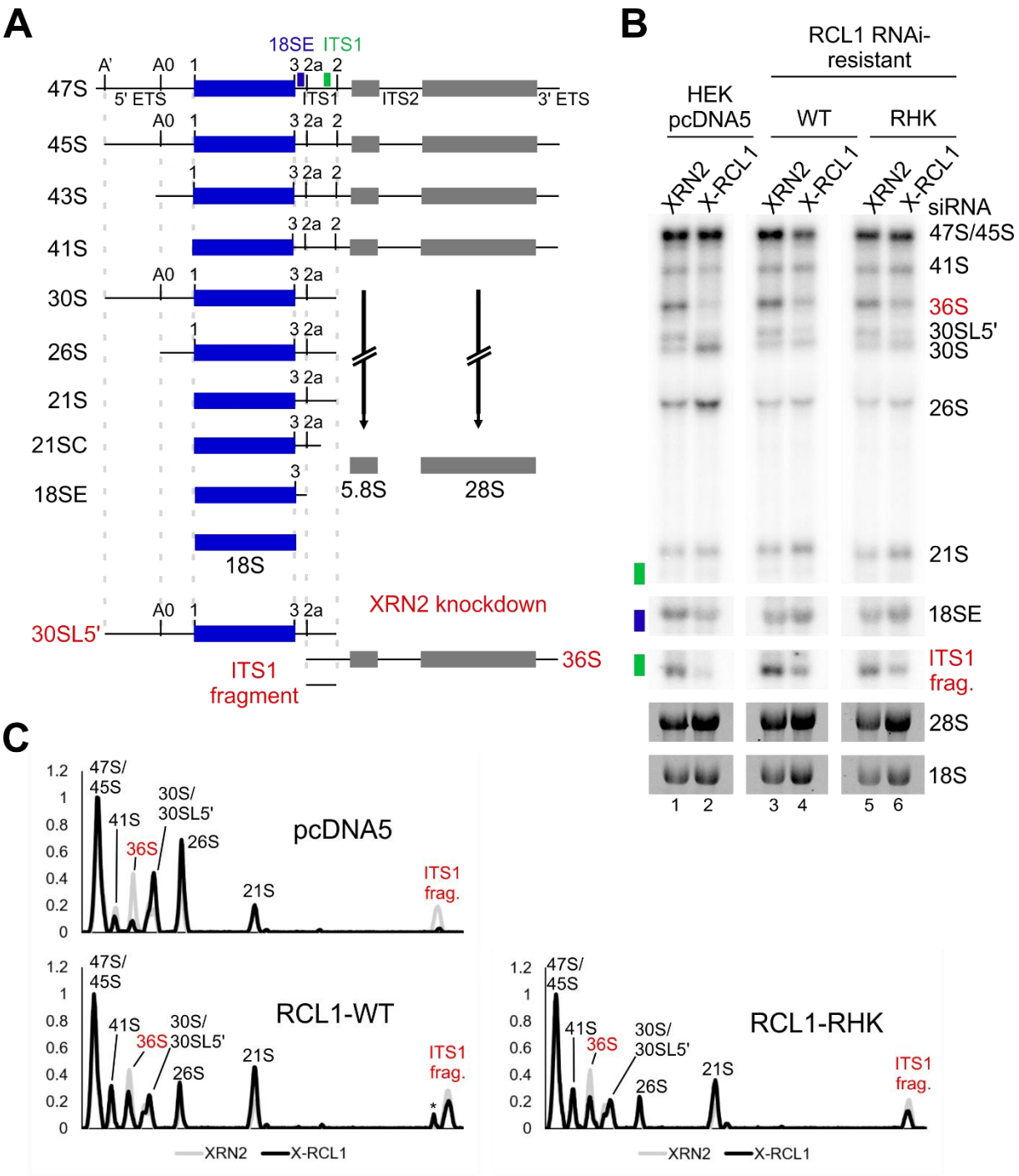


Figure 3.17 RCL1 RHK residues are not important for pre-rRNA processing in HEK293T cells in the absence of XRN2. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5), or wild type (WT) or mutant (RHK) RNAi-resistant RCL1. **(A)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Pre-rRNA intermediates specific to the minor, site 2a-dependent processing pathway, which accumulate upon depletion of XRN2, are shown in red. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(B)** Following RNAi-mediated depletion of XRN2 (XRN2) or both endogenous RCL1 and XRN2 (X-RCL1), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane. RNA was analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S

rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. **(C)** Levels of pre-rRNAs detected by the ITS1 probe from panel B were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

While no pre-rRNA processing phenotype was observed upon expression of RCL1 RHK under normal conditions, it is not possible to assess the effect of this mutant on cleavage at the 2a site specifically under normal conditions. These residues are possibly only important for cleavage at site 2a, which is not an essential event in the 'major' human pre-rRNA processing pathway. RNAi-mediated knockdown of XRN2 was used to shift processing to the "minor" pathway, causing a reliance on cleavage at the 2a site. XRN2 was depleted in cells expressing either the FLAG-tag alone (HEK-pcDNA5), WT RCL1 or RCL1 RHK, either alone or in conjunction with depletion of endogenous RCL1. Northern blotting was performed using 'ITS1' and '18SE' probes (Figure 3.17B) and the intensities of bands in Figure 3.17B were quantified using ImageQuant and normalised to levels of the 47S pre-rRNA (Figure 3.17C). In all cells, depletion of XRN2 caused the accumulation of site 2a-specific precursors, the 36S pre-rRNA and the ITS1 fragment. In cells expressing RNAi-resistant WT RCL1, levels of these 2a-specific precursors are recovered. The 36S pre-rRNA and the ITS1 fragment are also recovered in cells expressing RCL1 RHK, indicating that these residues are not important for the role that RCL1 plays in pre-rRNA processing at site 2a. Therefore, these data suggest that the 'RHK' residues of human RCL1 are not important for pre-rRNA processing at any of the sites important for the maturation of the 18S rRNA. This was surprising, given that the equivalent residues in the yeast Rcl1 protein were reported to be important for pre-rRNA processing at sites A_0, A_1 and A_2 *in vivo* and for Rcl1's observed *in vitro* nuclease activity (Horn *et al.*, 2011).

3.2.9 'RDK' residues are important for the nucleolar localisation of yeast Rcl1

The fact that the human equivalent of the yeast Rcl1 'RDK' mutant supported normal pre-rRNA processing in human cells raises further questions about the role of these residues in Rcl1's putative cleavage activity in yeast. While the 'RDK' mutant disrupted cleavage at the A_2 site in yeast, it also affected cleavage at sites A_0 and A_1, giving a very similar pre-rRNA processing phenotype to that caused by depletion of Rcl1 (Horn *et al.*, 2011). This raises the possibility that this mutant causes a more general defect in Rcl1's likely non-enzymatic role in 18S rRNA maturation as an SSU

processome component. Interestingly, a recent crystal structure showed that mutation of the R327 residue, within the 'RDK' motif, is important for the interaction between Rcl1 and its binding partner, the GTPase Bms1 (Delprato *et al.*, 2014). Deletion of the residues of Bms1 required for its interaction with Rcl1 was also shown to strongly reduce the incorporation of Rcl1 into pre-ribosomes, suggesting that binding by Bms1 is required to recruit Rcl1 to pre-ribosomes (Karbstein *et al.*, 2005). This suggests that this interaction is important for the recruitment of Rcl1 to the SSU processome to perform its essential role in pre-rRNA processing. In yeast cells expressing Rcl1 mutated at the R327 residue, as well as at another residue shown to be important for the Bms1-Rcl1 interaction, this mutant was localised to the nucleus and the cytoplasm, while the WT protein localised correctly to the nucleolus (Delprato *et al.*, 2014). These previous results suggested that the pre-rRNA processing phenotype observed in the Rcl1 'RDK' mutant may be caused by disruption of this protein's localisation or incorporation into the SSU processome, rather than a defect in pre-rRNA substrate binding or catalytic activity.

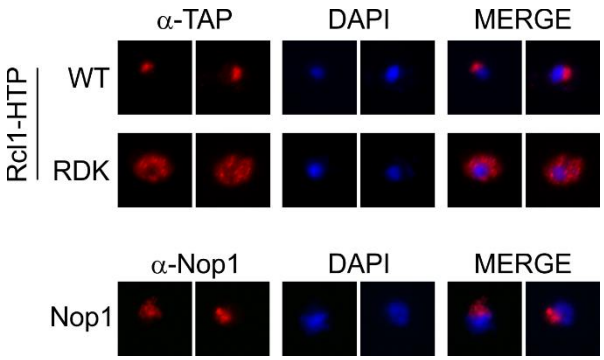


Figure 3.18 Yeast Rcl1 RDK residues are important for the nucleolar localisation of Rcl1. Immunofluorescence was performed on yeast strains expressing plasmid-encoded HTP-tagged yeast Rcl1 wild type (WT; top row) or mutant Rcl1 (RDK; middle row) using an anti-TAP antibody. An anti-Nop1 antibody was used as a nucleolar marker (bottom row). Immunofluorescence signal is shown in red and DAPI staining of the nucleus is shown in blue.

To investigate the localisation of the yeast 'RDK' mutant, immunofluorescence was used with cells expressing HTP-tagged (His6–TEV–protA) WT Rcl1 or Rcl1 RDK (Figure 3.18). An anti-TAP antibody was used to detect the cellular localisation of the tagged Rcl1 protein, and DAPI was used to stain nuclei. The box C/D snoRNP component Nop1 was used as a nucleolar marker. WT Rcl1 localised to the nucleolus, as expected with Rcl1's known role in pre-rRNA processing as part of the SSU

processome. The Rcl1 ‘RDK’ mutant, in contrast, localised mainly in the cytoplasm. This suggests that this mutant is defective in the interaction of Rcl1 with Bms1, and subsequently defective nuclear import of Rcl1, preventing its incorporation into pre-ribosomes and the SSU processome. Together, these results indicate that the pre-rRNA processing phenotype observed in cells expressing the ‘RDK’ mutant is due to these defects, rather than a disruption in pre-rRNA substrate binding. This also explains the fact that the pre-rRNA processing phenotype in cells expressing the ‘RDK’ mutant is very similar to that seen with depletion of Rcl1 (Horn *et al.*, 2011). Further work in our lab showed that the Rcl1 ‘RDK’ mutant does not interact with Bms1 *in vitro*, and does not associate with Nop1, the U3 snoRNA, or 23S, 22/21S and 20S pre-rRNAs (Wells *et al.*, 2016). Together, our recent Rcl1 results cast doubt on its proposed role as an endonuclease responsible for site A_2 cleavage in yeast, and site 2a cleavage in humans.

3.3 Discussion

The identification of the active enzymes responsible for endonucleolytic cleavage of pre-rRNA at several known cleavage sites has proved difficult, despite extensive study into ribosome biogenesis. Particularly in yeast, much work has been done to identify factors involved in processing and the specific processing steps in rRNA maturation. Discovery of the active endonucleases involved in the production of the mature 18S rRNA of the small subunit have been particularly challenging, due to the nature of three early cleavage events at sites around the 18S sequence called A_0, A_1 and A_2 in yeast and equivalent sites called A0, 1 and 2a in humans. In both systems, cleavage at these sites is coupled, and requires the associated action of many factors which assemble into a large ribonucleoprotein complex called the SSU processome. Many enzymatic and non-enzymatic factors make up this complex, many of which are required for the formation or assembly of the complex, and therefore for processing at these sites. This makes the identification of the specific endonuclease responsible for each cleavage event a difficult task.

In yeast, the RNA cyclase-like protein, Rcl1, was controversially reported to have an endonucleolytic function in processing of the A_2 site (Horn *et al.*, 2011). Three residues (RDK) were identified that were reportedly important for this cleavage activity, perhaps through a role in pre-rRNA substrate binding. Casting doubt on this idea, one of these residues was elsewhere shown to be important for the binding of Rcl1 to its

binding partner, the GTPase, Bms1 (Delprato *et al.*, 2014), and this interaction was shown (in the same study and separately) to be important for the nuclear import of Rcl1 and its subsequent incorporation into pre-ribosomes *in vivo* (Karbstein *et al.*, 2005). The putative endonuclease domain (PIN domain) of Utp24 was elsewhere shown to be indispensable for cleavage at sites A_1 and A_2 in yeast, suggesting potential overlapping functions between the two proteins (Bleichert *et al.*, 2006). This part of this PhD project aimed to clarify the situation by examining the human homologues of these two putative endonucleases in human pre-rRNA processing. These data, alongside yeast data from our lab, argues for the role of Utp24, but not Rcl1, as the endonuclease responsible for cleaving pre-rRNA at two coupled sites (A_1/1 and A_2/2a) in both yeast and humans (Wells *et al.*, 2016).

My *in vivo* data, utilising an RNAi rescue system in HEK293T cells, show that the PIN domain of human UTP24 is important for cleavage at sites 1 and 2a. This is consistent with yeast data implicating an active PIN domain in yeast Utp24 in processing at the equivalent sites, A_1 and A_2. Interestingly, induction of UTP24 PIN domain mutant protein expression was increased when endogenous UTP24 was depleted, suggesting that expression of the mutant proteins may be downregulated when the endogenous protein is present. In yeast, expression of Utp24 PIN domain mutants caused a dominant negative pre-rRNA processing defect *in vivo* (Bleichert *et al.*, 2006) which was not observed for human UTP24 in the data presented here. It is possible that the downregulation of mutant UTP24 proteins is responsible for the lack of a dominant negative phenotype in human cells. When endogenous UTP24 is depleted however, the mutant proteins do not appear to be downregulated, potentially reflecting that the presence of an inactive form of UTP24 may be more tolerable to the cell than the total absence of UTP24. Depletion of XRN2 promotes processing of ITS1 through the 2a site, which is normally part of a ‘minor’ processing pathway, rather than at the major ITS1 cleavage at site 2. Under these conditions, the PIN domain of UTP24 was still essential for cleavage at sites 1 and 2a, suggesting that, as in yeast, UTP24 is important for both of these coupled cleavages in human cells. This argues for a conserved role of Utp24, via its PIN domain, in the processing of pre-rRNA at two sites leading to the maturation of the 18S rRNA.

As mentioned above, Rcl1 was proposed to cleave at the A_2 site, and indeed was shown to cleave a pre-rRNA mimic at this site *in vitro* (Horn *et al.*, 2011). In our lab, and under slightly different conditions, we observed cleavage at the A_2 site by WT

yeast Utp24, which was abolished with the mutation of the PIN domain of Utp24 (Wells *et al.*, 2016). Studying RCL1 in *vivo* showed that, like in yeast, the C-terminal of RCL1 is important for the key role of RCL1 in pre-rRNA processing at sites A0, 1 and 2a. A recent Cryo-EM structure of the *S. cerevisiae* SSU processome showed that the C-terminal portion of Rcl1 interacts with Bms1, and so is probably important for the incorporation of Rcl1 in the SSU processome (Barandun *et al.*, 2017). This suggests a conserved role for RCL1 in ribosome biogenesis between yeast and humans. Mutation of the equivalent residues in human RCL1 that were shown to be important for yeast A_2 cleavage had no effect on pre-rRNA processing in human cells. Even when XRN2 was depleted, causing reliance on cleavage at site 2a (equivalent to yeast A_2), the mutation the "RHK" residues had no effect on levels of 2a-specific precursors, suggesting they are not important for 2a cleavage. These results do not point to a role for RCL1 in direct cleavage in human cells.

The implication of both Rcl1 and Utp24 in the processing of pre-rRNA at the yeast A_2 site, but only UTP24 at the human 2a site, raises some interesting possibilities. While the maturation of 18S rRNA is generally conserved between yeast and humans, there are some important differences in the processing pathways, particularly in relation to the processing of ITS1. Firstly, while cleavage at the yeast A_2 site can occur either co-transcriptionally or post-transcriptionally (Henras *et al.*, 2015), all human processing steps are thought to occur post-transcriptionally, with the exception being the initial cleavage at site A' in the 5' ETS (a site not present in the yeast pre-rRNA). This allows for the possibility that two different enzymes (Rcl1 and Utp24) cleave the yeast pre-rRNA at site A_2 in these two distinct contexts, while UTP24 cleaves the human pre-rRNA post-transcriptionally at site 2a. It remains possible that in yeast, Rcl1 cleaves at the A_2 site co-transcriptionally, while Utp24 cleaves here post-transcriptionally.

Despite this possibility, further study in yeast has casted doubt on the role of Rcl1 in cleavage at yeast site A_2 (Wells *et al.*, 2016). The published *in vivo* data from yeast show that mutation of the RDK residues, predicted to be involved in substrate binding, produced a pre-rRNA processing defect strikingly similar to that observed with depletion of Rcl1 (Horn *et al.*, 2011). This is consistent with a disruption of SSU processome assembly. Other work in yeast suggests that mutation of the 'R' residue of the 'RDK' mutant is important for the Bms1-Rcl1 interaction, and hence likely disrupts the nuclear import of Rcl1, subsequently preventing the incorporation of Rcl1

into the SSU processome (Delprato *et al.*, 2014). If the vital role of Rcl1, as a structural but not enzymatic component, occurs within the SSU processome (in the nucleolus), then incorrect localisation of Rcl1 would likely cause a phenotype identical or very similar to that of Rcl1 depletion. To assess this possibility, immunofluorescence was performed in yeast cells expressing either WT Rcl1 or Rcl1 RDK. This showed clearly that while WT Rcl1 was correctly localised to the nucleolus, Rcl1 RDK was retained in the cytoplasm. This, along with data showing that the Rcl1 "RDK" mutant is unable to interact with Bms1 *in vitro* and does not associate with other SSU processome components, the U3 snoRNA, or pre-rRNA *in vivo*, strongly suggests that the phenotype observed for this mutant is caused by incorrect localisation and disruption of SSU processome assembly (Wells *et al.*, 2016).

The lack of a known catalytic active site in Rcl1 (Tanaka *et al.*, 2011), and the fact that the only mutant tested causes defects unrelated to any potential catalytic activity, leaves the reported *in vitro* cleavage of the yeast A_2 site as the only apparent link between Rcl1 and an enzymatic role in pre-rRNA cleavage (Horn *et al.*, 2011). Data from our lab, however, also contests this observation (Wells *et al.*, 2016). Under conditions used in our lab, yeast WT Utp24, but not the Utp24 PIN domain mutant, nor WT Rcl1, was able to cleave the pre-rRNA mimic at the yeast A_2 site. Under the conditions used in Horn *et al*., 2011, Rcl1 did indeed cleave at the A_2 site in our hands, but so did both WT and PIN domain mutant forms of Utp24 (Wells *et al.*, 2016)**.** The fact that cleavage was seen with the inactive form of Utp24 under these conditions suggests that this effect is non-specific and not performed by the recombinant protein used. Further work is needed to assess the ability of Rcl1 to cleave RNA. Namely, a mutant form of Rcl1 incapable of the observed cleavage under the conditions suitable for WT Rcl1 cleavage must be identified. The only mutant used so far for cleavage assays, the "RDK" mutant, has been shown to be important for non-nuclease related functions of Rcl1.

Since mutation of the yeast Rcl1 'RDK' residues had such a significant effect on Rcl1's incorporation into the nucleolus and the SSU processome, it was surprising that mutation of the equivalent residues in human RCL1 ('RHK') had no effect on pre-rRNA processing in human cells. However, looking more closely at the differences in pre-rRNA processing between yeast and humans sheds some light on this. While the nuclear import and SSU processome incorporation of Rcl1 in yeast relies on its interaction with Bms1 (Delprato *et al.*, 2014), and while the interaction between the two

proteins is conserved (Wang *et al.*, 2016), it may not be important for Rcl1 SSU processome incorporation in humans. It has been shown for example, that RCL1 likely enters the SSU processome before BMS1 does (Turner *et al.*, 2012), and therefore the RCL1-BMS1 interaction must not be important for the nuclear import of Rcl1. This, along with further investigation into the importance of the yeast Rcl1 'RDK' residues, explain the difference in *in vivo* pre-rRNA processing phenotype observed with mutation of these residues. Together, these results suggest that Rcl1 is likely not responsible for cleavage of the pre-rRNA in either yeast or humans.

In vitro data from our lab showed that yeast Utp24 can cleave pre-rRNA at the yeast A_2 site. Attempts to confirm *in vitro* cleavage by human UTP24 at the human 2a site have so far been unsuccessful, as the sequence around the 2a site is extremely GC-rich, and therefore difficult to amplify. However, recombinant human UTP24, but not the PIN domain mutant, was able to cleave a pre-rRNA containing the yeast A_2 site and produced the same processing pattern as the yeast Utp24 protein (Figure 3.19) (Wells *et al.*, 2016). This confirms that human UTP24 possesses endonuclease activity and also suggests that the role of Utp24 is conserved between yeast and humans. For both the yeast and human proteins, as well as cleavage at the genuine A_2 site, cleavage was observed at several other sites around the authentic site (Figure 3.18) (Wells *et al.*, 2016). The pattern of cleavage was identical for both yeast and human proteins, suggesting that this activity is specific to the conserved PIN domain in both proteins. It is likely that *in vivo*, within pre-ribosomes, the correct pre-rRNA folding via the U3 snoRNP and/or association of other ribosome biogenesis factors is required for processing solely at the correct A_2 site.

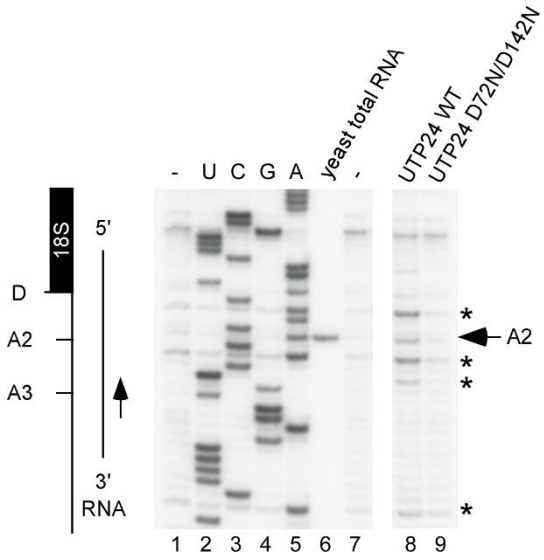


Figure 3.19 Human UTP24 cleaves a pre-rRNA mimic containing the yeast site A_2. An *in vitro* cleavage assay using recombinant human UTP24 wild type (WT) or mutant (D72N/D142N) UTP24. Positions of ITS1 cleavage sites within the pre-rRNA mimic are shown in the left panel. UTP24-mediated cleavage at site A_2 is indicated with an arrow, and cleavages around the cleavage site are marked with asterisks. Adapted from (Wells *et al.*, 2016).

Data from this study, other work from our lab (Wells *et al.*, 2016) as well as data performed elsewhere (Bleichert *et al.*, 2006; Tomecki *et al.*, 2015) strongly links both yeast and human Utp24 to cleavage at site A_1/1 at the 5' of the 18S rRNA. However, attempts to confirm endonuclease activity at site A_1/1 have so far failed (Tomecki *et al.*, 2015; Wells *et al.*, 2016). As mentioned above, it is likely that conditions required for cleavage at this site *in vivo* cannot be replicated in *in vitro* studies. These may include correct folding of the pre-rRNA, and possibly the action of other ribosome biogenesis factors, or even the whole, correctly assembled SSU processome. The recently solved crystal structure of the *S. cerevisiae* SSU processome showed Utp24 positioned close to the un-cleaved A_1 site and associated with the U3 snoRNA and several other protein factors (Barandun *et al.*, 2017). Despite this, primer extension experiments used to analyse the 5' end of the 18S rRNA in two studies showed that expression of a UTP24 PIN domain mutant disrupted accurate cleavage at site 1 in human cells *in vivo* (Tomecki *et al.*, 2015; Wells *et al.*, 2016). Interestingly, both studies observed the same aberrant cleavage 2 nucleotides downstream of site 1 with expression of the UTP24 PIN domain mutant. This explains why some 21S (and 21SC) is still observed in cells expressing PIN domain mutant forms of UTP24 (Figure 3.19). When XRN2 and UTP24 were both depleted in cells expressing UTP24 PIN domain

mutants, levels of the 21S intermediate were significantly reduced, and no 21SC accumulation was observed (Figure 3.10), suggesting a role for the 5' to 3' exonucleolytic activity of XRN2 in the alternative processing mechanism. However, data from our lab and from Tomecki *et al*., 2015 is contrasting in terms of the alternative nuclease responsible for this aberrant processing in the absence of UTP24-mediated cleavage. Our primer extension data showed that with knockdown of the 5'-3' exoribonuclease, XRN2, the aberrant primer extension stop 2 nucleotides downstream of site 1 was significantly reduced (Wells *et al.*, 2016). This implicated XRN2 in exonucleolytic trimming of the 5' ETS towards the 5' end of the 18S rRNA in the absence of endonucleolytic cleavage by UTP24. This is unsurprising, as XRN2 is important for degradation of a fragment of the 5' ETS upstream of the A' site ("ETS1") and the excised fragment resulting from cleavages at sites A0 and 1 ("ETS3") (Sloan *et al.*, 2014). In contrast, Tomecki *et al*. reported that in their hands, depletion of XRN2 did not alter the aberrant processing observed with expression of inactive UTP24. This discrepancy could be related to the difference in the knockdown technique used to deplete UTP24. While our study used simple siRNA-mediated knockdown, Tomecki *et al*., used stable shRNA expression, which had more extreme effects on cell viability and could have caused secondary effects.

Despite years of research into ribosome biogenesis, the endonucleases responsible for early cleavages at three sites in 18S rRNA maturation have remained elusive or controversial. This section of this PhD project has cast doubt on the potential role of Rcl1 in cleavage at the yeast A_2 site and, along with other data from our lab, presented Utp24 as the endonuclease responsible for cleavage at yeast sites A_1 and A_2 and human sites 1 and 2a. The enzymes cleaving the yeast A_0 site and the human A' and A0 sites are yet to be identified.

Chapter Four

Human UTP23 coordinates key interactions in the pre-40S particle and its PIN domain and Zn finger motif are essential for pre-rRNA processing

4.1 Introduction

The SSU processome is a large ribonucleoprotein (RNP) complex required for three coupled early nucleolar cleavages in 18S rRNA maturation, at sites A0, A1 and A2 (A0, 1 and 2a in humans). The *S. cerevisiae* SSU processome contains 72 non-ribosomal proteins and is similar in size to the ribosome itself (Phipps *et al.*, 2011). Of these SSU processome components, many are structural proteins, but the SSU processome also contains RNA helicases, GTPases, kinases and endonucleases. Assembly of the SSU processome occurs in a stage-specific, hierarchical manner, with several factors forming stable subcomplexes that associate with the nascent pre-rRNA (Chaker-Margot *et al.*, 2015; Zhang *et al.*, 2016b). Many protein components of the SSU processome are U three proteins (Utps). The UtpA and UtpB subcomplexes are recruited to the 5' ETS and then chaperone the 5' ETS and the U3 small nucleolar (sno)RNP (Hunziker *et al.*, 2016). The U3 snoRNP, plays a key role in SSU processome assembly through base-pairing to the 5' ETS and within the 18S rRNA sequence, and is important for establishment of the central pseudoknot (Dutca *et al.*, 2011; Marmier-Gourrier *et al.*, 2011).

Many different snoRNPs function in ribosome biogenesis in the covalent modification of the pre-rRNA (Kos and Tollervey, 2010). Box C/D snoRNAs direct the 2'-*O*-methylation while box H/ACA snoRNAs direct pseudouridylation by selecting modification sites on the pre-rRNA (Kiss *et al.*, 2010; Watkins and Bohnsack, 2012). Instead of functioning in methylation of the pre-rRNA, the box C/D snoRNA, U3, is essential for processing of the pre-rRNA at cleavage sites A0, A1 and A2, as described above. Another snoRNA, the box H/ACA snoRNA snR30, which lacks a modification target, is also responsible for pre-rRNA cleavage at these sites, base-pairs to the 18S rRNA and is required for assembly of the SSU processome (Atzorn *et al.*, 2004; Fayet-Lebaron *et al.*, 2009; Lemay *et al.*, 2011). Depletion of many components of the SSU processome disrupts cleavage of the pre-rRNA at sites A0, A1 and A2, leading to the accumulation of a 23S precursor. This makes the identification of enzymes performing

specific functions in pre-rRNA processing, including the endonucleases catalysing cleavage of the pre-rRNA, at these sites difficult.

The Rnt1 endonuclease was at one point thought to cleave at the A0 site, as well as its confirmed function in cleavage in the 3' ETS but was later discarded as a candidate for A0 processing (Elela *et al.*, 1996; Kufel *et al.*, 1999). The RNA cyclase-like protein Rcl1, a component of the SSU processome, was controversially reported to cleave at site A2, despite the lack of a known catalytic site (Horn *et al.*, 2011). Recent research suggested that Rcl1 does not function as an active endonuclease in processing of these sites and identified the PIN domain endonuclease Utp24 as the enzyme responsible for cleaving the pre-rRNA at sites A1 and A2 (Wells *et al.*, 2016). The final pre-18S rRNA cleavage occurs at the 3' end of 18S in the cytoplasm and is mediated by the PIN domain endonuclease Nob1 (Fatica *et al.*, 2003; Pertschy *et al.*, 2009; Lebaron *et al.*, 2012).

Another component of the SSU processome, Utp23, also contains a PIN endonuclease domain. PIN domains are predicted nuclease domains as they show similarity to 5' exonucleases and RNase H (Arcus *et al.*, 2004). A PIN domain is found in endonucleases in many processes, including exosome complex-mediated RNA degradation and RNA processing (Schneider *et al.*, 2009)**,** and nonsense-mediated decay (NMD) (Glavan *et al.*, 2006). The SSU endonucleases cleaving at sites A1, A2 (Utp24) and site D (Nob1) also contain PIN domains which are essential for cleavage at these sites. Catalytically active PIN domains contain four highly conserved acidic amino acids that coordinate a metal ion required for ribonucleolytic activity (Clissold and Ponting, 2000), although research on the NMD endonuclease SMG6 suggests that the conservation of three acidic residues at these positions is sufficient for cleavage activity (Glavan *et al.*, 2006). *S. cerevisiae* Utp23, however, only has two of these four active site amino acids in its PIN domain, suggesting it is not an active endonuclease. Consistent with this, mutation of the two remaining conserved acidic residues in the PIN domain of Utp23 had no effect on pre-rRNA processing *in vivo* (Bleichert *et al.*, 2006) and these mutants were also able to rescue the lethal Utp23 deletion, suggesting these residues are functionally dispensable (Lu *et al.*, 2013). However, depletion of Utp23, as is the case for depletion of many SSU processome components, blocked cleavages at sites A0, A1 and A2, indicating that the presence of Utp23 is important for all three cleavage events. Utp23, like the endonuclease Utp24, also contains a conserved Zn finger motif, which is predicted to be involved in RNA-binding and

mutation of key Zn finger residues were lethal or severely inhibited growth (Lu *et al.*, 2013). These results suggest that Utp23 plays a non-enzymatic role in pre-rRNA processing in yeast, perhaps in mediating correct folding of the pre-rRNA and/or recruitment of other factors to the pre-rRNA or pre-ribosomes.

Despite not being an active endonuclease, Utp23 clearly plays an important role in small subunit pre-rRNA processing. Utp23 has been shown to associate with the H/ACA snoRNP snR30 *in vivo* (Hoareau-Aveilla *et al.*, 2012; Lu *et al.*, 2013), and the processing of the pre-rRNA at sites A0, A1 and A2 requires both Utp23 and snR30 (Morrissey and Tollervey, 1993; Bleichert *et al.*, 2006). The C-terminal Utp23 is required for the release of snR30 from pre-ribosomes, and depletion of snR30 inhibits the incorporation of Utp23 into the SSU processome (Hoareau-Aveilla *et al.*, 2012). The 3' hairpin of snR30 base-pairs with the 18S rRNA in the eukaryotic expansion segment 6 (ES6) (Fayet-Lebaron *et al.*, 2009) and Utp23 may be important in establishment of this interaction (Hoareau-Aveilla *et al.*, 2012). The A1/A2 endonuclease Utp24 also contacts the pre-rRNA in the ES6 region (Wells *et al.*, 2016). A study of the stage-specific SSU processome assembly (Chaker-Margot *et al.*, 2015) on pre-ribosomes showed that Utp23 is incorporated into the complex after assembly of UtpA, UtpB and U3 snoRNP subcomplexes, along with the snR30 snoRNP (including core box H/ACA proteins Gar1, Nhp2, Nop10 and Cbf5), ES6-interacting factors Rrp7 (Lin *et al.*, 2013) (a component of the UtpC subcomplex) and snR30-interacting factors Rok1 (Martin *et al.*, 2014) and Rrp5 (Lebaron *et al.*, 2013). The RNA helicase activity of Rok1 is important for the release of Rrp5 from pre-40S ribosomes, which in turn is required for the release of snR30 (Bohnsack *et al.*, 2008; Martin *et al.*, 2014; Khoshnevis *et al.*, 2016). Utp23 was not present in the SSU processome in two recently published structures captured after cleavage at A0 but before A1 cleavage, suggesting that it is removed from the pre-ribosome shortly after A0 cleavage (Barandun *et al.*, 2017; Cheng *et al.*, 2017).

Recent data from our lab demonstrate that Utp23 makes direct contact with the eukaryotic expansion segment 6 (ES6) of the 18S rRNA and the snR30 snoRNA *in vivo* (Wells *et al.*, 2017). *In vivo* protein-RNA crosslinking (CRAC) experiments showed that Utp23 binding sites on the pre-rRNA were close to, but not overlapping with, the known binding sites of snR30 to the pre-rRNA (Figure 4.1). Direct binding sites of Utp23 on the snR30 were within the 3' portion of the snoRNA, and at sites close to snR30 sequences that base-pair with the 18S rRNA (Wells *et al.*, 2017) (Figure 4.1).

Protein-protein interaction studies further revealed that Utp23 directly interacts with ES6- (Utp24, Rok1, Rrp7) and snR30-interacting (Kri1 (Lemay *et al.*, 2011; Hoareau-Aveilla *et al.*, 2012)) proteins as well as the core box H/ACA protein Nhp2 *in vitro*. The interaction between Utp23 and Utp24 suggests a potential role for Utp23 as a recruitment or chaperone in the Utp24-mediated cleavage events. Utp23 also interacted with itself, suggesting that it may act as a multimer, which has been shown for other PIN domain proteins (Lamanna and Karbstein, 2011). Together, these data suggest that the interaction of Utp23 with snR30 and the ES6 region of 18S rRNA plays a key role in coordinating binding of factors to this region during pre-rRNA processing.

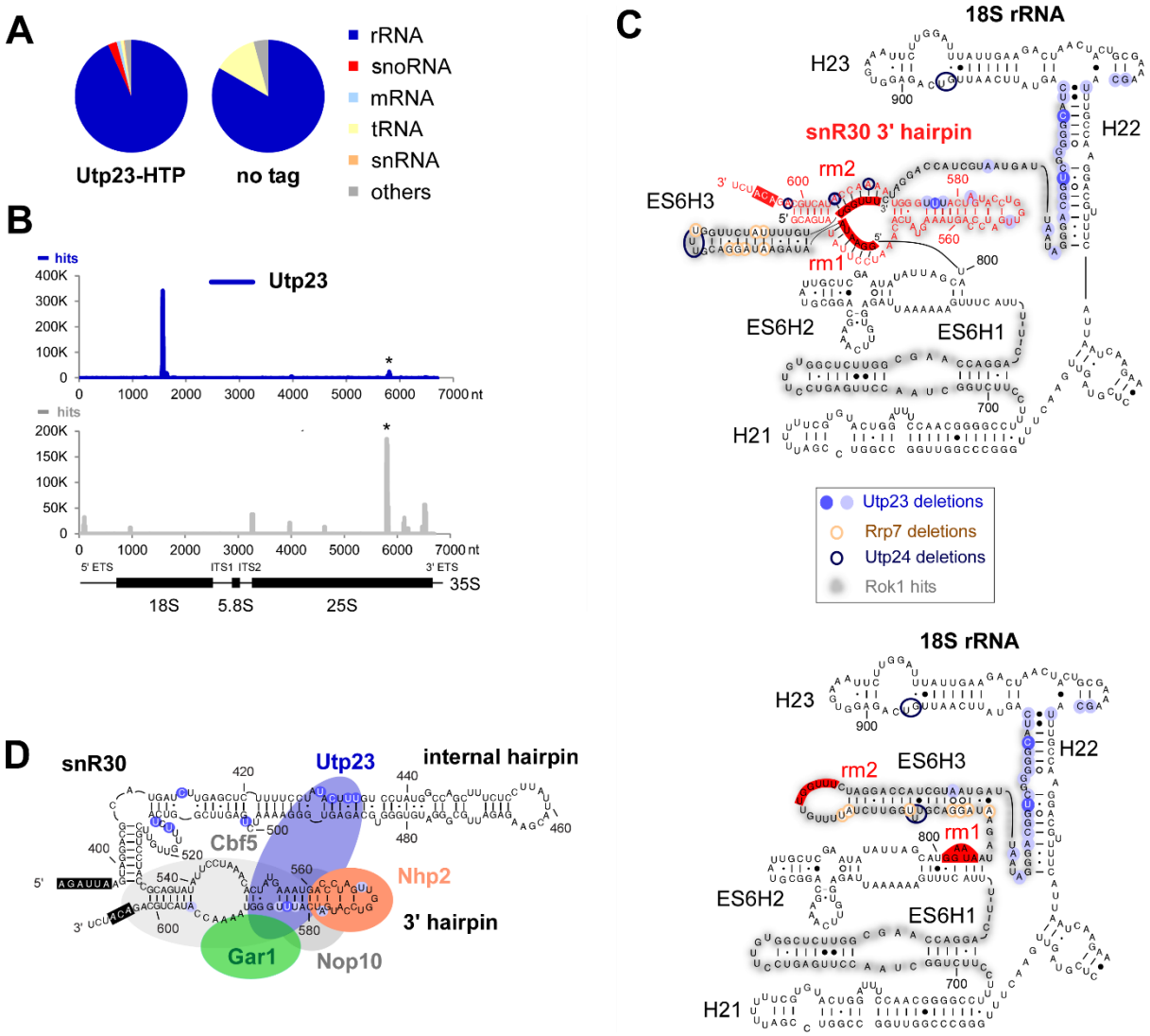


Figure 4.1 Utp23 Crosslinking sites on the pre-rRNA and snR30 snoRNA. (A) Transcriptome-wide binding profiles of RNAs recovered with purified yeast Utp23 protein in a UV crosslinking and analysis of cDNA (CRAC) experiment. The proportion of sequences mapping to functional RNA classes are shown in as pie charts (upper panel), revealing that Utp23 was predominately found crosslinked to rRNA sequences. **(B)** Sequences aligned to the rDNA sequence encoding the 35S pre-rRNA are shown plotted as reads per million, showing a major Utp23 crosslinking site within the 18S rRNA. **(C)** The predicted structure of the ES6 region of the 18S rRNA during base-pairing with snR30 (upper panel) and following snR30 release (lower panel). The 3' hairpin of snR30 is shown in red, with snR30 binding sites on the pre-rRNA highlighted in red and Utp23 crosslinking sites indicated in blue. Sites of Rrp7 (brown), Utp24 (dark blue) and Rok1 (grey) are also shown. **(D)** Schematic depiction of the predicted positions Utp23 and H/ACA core proteins on the snR30 snoRNA. Utp23 crosslinking sites on the snR30 snoRNA are shown as blue circles.

Much less is known about ribosome biogenesis in general, and UTP23 specifically, in humans. Pre-rRNA processing steps in 18S rRNA maturation are generally conserved between yeast and humans, with a few key differences (Sloan *et*

al., 2013b; Tafforeau *et al.*, 2013). The majority of pre-rRNA processing required for yeast 18S rRNA maturation occurs co-transcriptionally (Kos and Tollervey, 2010). In contrast, early cleavages at sites equivalent to yeast sites A0, A1 and A2 (A0, 1 and 2a) occur post-transcriptionally in humans, and only the initial cleavage in the 5' ETS, at a site called A' (a site not present in the yeast pre-rRNA), occurs co-transcriptionally (Lazdins *et al.*, 1997). Cleavage at A' appears to be a quality control step, that can be bypassed, with the initial cleavage in the 5' ETS occurring instead at site A0 (Sloan *et al.*, 2014). While the site 2a cleavage event appears to be functionally analogous to cleavage at yeast site A2, the separation of SSU and LSU rRNAs represents a major distinction in pre-rRNA processing pathways between yeast and humans. In yeast, the major cleavage event in ITS1 occurs at site A2, following cleavages at sites A0 and A1. Cleavage at the equivalent human site, 2a, in contrast, is part of a minor processing pathway, with the major ITS1 cleavage occurring at site 2, equivalent to yeast site A3 (Mullineux and Lafontaine, 2012; Sloan *et al.*, 2013b).

My work, as outlined in chapter 3 implicated human UTP24 in cleavage at sites 1 and 2a via its conserved PIN domain (Wells *et al.*, 2016) and UTP24 was independently implicated in site 1 cleavage (Tomecki *et al.*, 2015). As for the equivalent yeast site, the nuclease responsible for cleaving at site A0 remains elusive. As a conserved component of the SSU processome (Turner *et al.*, 2012), UTP23 is essential for cleavages at sites A0, 1 and 2a in humans (Sloan *et al.*, 2014). Interestingly, UTP23 appears to play a very different role in pre-rRNA processing in mice, where its depletion affects later processing events (Wang *et al.*, 2014). Human UTP23 is also associated with the human homologue of snR30, the U17 snoRNA (Hoareau-Aveilla *et al.*, 2012). Unlike the yeast protein, the PIN domain of human UTP23 contains three of the four key acidic amino acids that are predicted to make up the metal ion-coordinating active site required for endonuclease activity, as well as a conserved Zn finger motif. This suggests that, surprisingly, human UTP23 may be an active endonuclease, and may play a distinct or additional role in human pre-rRNA processing.

Work performed during this PhD project aimed to characterise UTP23 in human cells and investigate its potentially conserved role in coordination of binding of ribosome biogenesis factors at the ES6 region of the 18S rRNA as well as its putative additional role as an active endonuclease in pre-rRNA processing.

4.2 Results

4.2.1 Yeast Utp23 binds the snR30 snoRNA in vitro

The data from *in vivo* UV crosslinking and analysis of cDNA (CRAC) experiments showed significant enrichment of crosslinking of yeast Utp23 to the snoRNA snR30 (Figure 4.1) (Wells *et al.*, 2017). To confirm this interaction, an *in vitro* RNA binding electromobility shift assay (EMSA) was performed using recombinant GST-tagged Utp23 protein and an *in vitro*-transcribed, radiolabelled snR30 RNA (Figure 4.2). GST-Utp23 or free GST (1000 nM) was incubated with the RNA and free RNA and RNA-protein (RNP) complexes were separated on a 4% native polyacrylamide gel and visualised using a PhosphorImager (Figure 4.2A). There was a complete shift from free RNA to RNP complex with GST-Utp23, compared to either the no protein control (-) or GST, suggesting a direct RNA-protein interaction between Utp23 and snR30. To test the strength of this interaction, the EMSA experiment was performed with a range of GST-Utp23 concentrations (0-1000 nM) (Figure 4.2A). The gradual increase in RNP complex compared to free RNA indicates an estimated dissociation constant (K_D) of ~300 nM. This data shows that, consistent with its known association with snR30 *in vivo* (Hoareau-Aveilla *et al.*, 2012), Utp23 directly binds to snR30 *in vitro* and suggests that may play a direct role in its function and release from the pre-ribosome.

Figure 4.2 Recombinant yeast Utp23 binds the snR30 snoRNA *in vitro*. (A) Electromobility shift assay (EMSA) to analyse binding of recombinant yeast GST-Utp23 protein (0-1000 nM) or GST (500 nM) to radiolabelled *in vitro* transcribed snR30 RNA substrate. RNP complexes were separated from free snR30 RNA on a 4% native polyacrylamide gel and visualised using a PhosphorImager. **(B)** Purified GST-Utp23 protein and GST was separated on a 13% polyacrylamide gel and visualised by Coomassie staining.

4.2.2 Human UTP23 binds the U17 snoRNA in vitro

To investigate the conservation of Utp23's role in ribosome biogenesis between yeast and humans, the direct interaction of human UTP23 with the homologue of snR30, the U17 snoRNA, was assessed (Figure 4.3). An EMSA was performed with recombinant GST-tagged UTP23 or free GST and *in vitro* transcribed, radiolabelled U17 RNA (Figure 4.3A). Free RNA and RNP complexes were separated on a 4% native polyacrylamide gel and visualised using a PhosphorImager. The shift from free RNA to RNP complexes observed with increasing concentration of GST-UTP23 (0-600 nM) suggests that UTP23 directly binds the U17 snoRNA *in vitro*, consistent with its observed association with U17 *in vivo* (Hoareau-Aveilla *et al.*, 2012). The *in vitro* binding observed for UTP23 with the U17 snoRNA is considerably weaker than that seen for yeast Utp23 and snR30 (Figure 4.2). Despite this, this data supports a conserved role for Utp23 in the function of snR30 in pre-rRNA processing in both yeast and humans and suggests it may play a similar role in the coordination of factors at the 18S rRNA ES6 region.

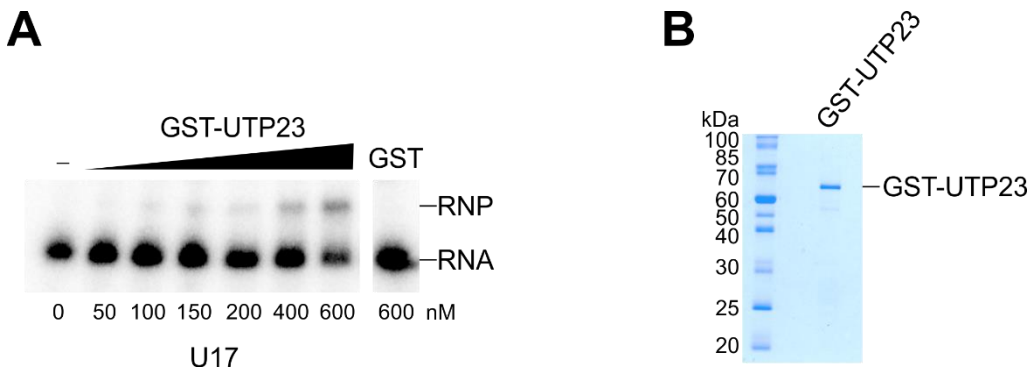


Figure 4.3 Recombinant human UTP23 binds the U17 snoRNA *in vitro*. Electromobility shift assay (EMSA) to analyse binding of recombinant human GST-UTP23 protein (0-600 nM) or GST (600 nM) to radiolabelled *in vitro* transcribed U17 RNA substrate. RNP complexes were separated from free U17 RNA on a 4% native polyacrylamide gel and visualised using a PhosphorImager. **(B)** Purified GST-UTP23 protein used in the EMSA was separated on a 13% polyacrylamide gel and visualised by Coomassie staining.

4.2.3 Human UTP23 interacts with pre-40S processing factors, which associate with the 18S rRNA ES6 region and/or the snR30 snoRNA

Research in yeast suggests a role for Utp23 in the association of protein factors with the ES6 region of the 18S rRNA and to the snR30 snoRNA (Hoareau-Aveilla *et al.*, 2012; Wells *et al.*, 2017). Recent data from our lab showed that Utp23 binds to this region of the pre-rRNA at positions close to base-pairing sites of snR30, and to the internal hairpin of snR30 itself (Wells *et al.*, 2017). A direct interaction between Utp23 and the box H/ACA core protein, Nhp2 was also observed *in vitro* and Utp23 has also been reported to interact with another core protein, Gar1 (Tarassov *et al.*, 2008). The internal hairpin of snR30 is dispensable for snR30 function *in vivo* (Atzorn *et al.*, 2004), so the observed interactions of Utp23 with box H/ACA core proteins likely compensate for the lack of direct RNA-protein binding between Utp23 and snR30 *in vivo*. Recombinant yeast Utp23 also directly interacted with other protein factors binding to the ES6 region or associated with snR30 *in vitro*, including Utp24, Rrp7, and Rok1 (Wells *et al.*, 2017). These results suggest that Utp23 may play a key role in coordinating the interaction of multiple ribosome biogenesis factors with the pre-rRNA during pre-rRNA processing in yeast.

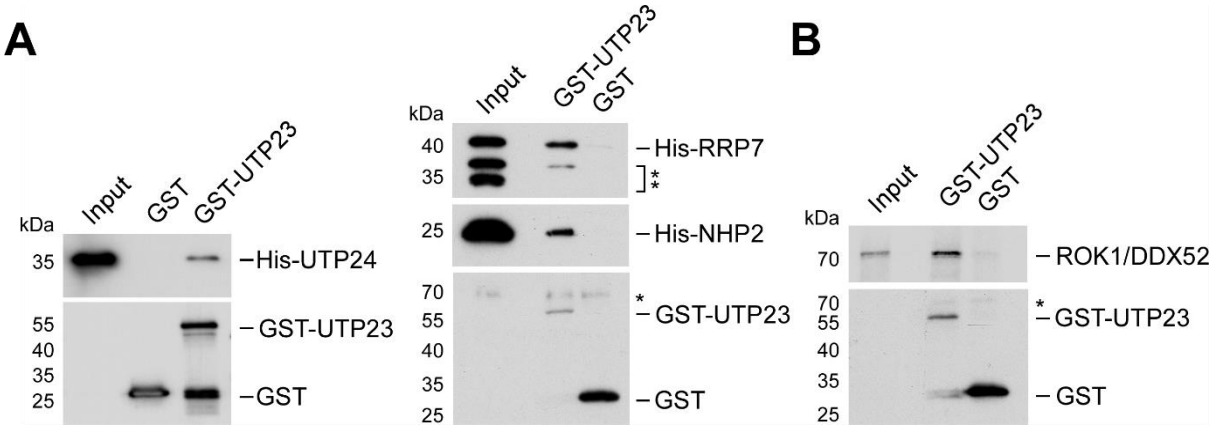


Figure 4.4 Human UTP23 interacts with pre-40S processing factors which associate with the 18S rRNA ES6 region. (A) Protein-protein interaction assay to analyse interactions between recombinant UTP23 and His-tagged ribosome biogenesis factors. After immobilisation on glutathione sepharose, recombinant GST-tagged UTP23 was incubated with His tagged UTP24, NHP2 or RRP7. Eluted material was separated by SDS-PAGE, transferred to nitrocellulose membrane and protein was analysed by western blotting using anti-GST and anti-His antibodies. 10% (UTP24; left panel) or 5% (NHP2/RRP7; right panel) of the input material was also loaded. The double asterisk indicates truncated forms of His-tagged RRP7 and the single asterisk indicates a non-specific protein detected by the anti-GST antibody. **(B)** Protein-protein interaction assay to analyse interactions between recombinant UTP23 and *in vitro*

translated ROK1/DDX52. ROK1/DDX52 was translated in the presence of ^{35}S methionine. After separation by SDS-PAGE, protein was visualised using a PhosphorImager. 1% of the input material was loaded. The asterisk indicates a non-specific protein detected by the anti-GST antibody.

To analyse whether these interactions are conserved in human cells, *in vitro* protein-protein interaction studies were performed using recombinant human UTP23 protein and recombinant or *in vitro* translated human homologues of yeast snR30- and ES6-associated proteins (Figure 4.4). GST-tagged UTP23 or free GST was immobilised on glutathione sepharose and incubated with recombinant His-tagged UTP24, RRP7 or NHP2 or *in vitro* translated ROK1/DDX52. Samples were washed to disrupt non-specific binding, separated by SDS-PAGE and visualised by western blotting using anti-GST and anti-His antibodies (Figure 4.4A) or using a PhosphorImager (Figure 4.4B). This showed that human UTP23 directly interacts with UTP24, RRP7, NHP2 and ROK1/DDX52 *in vitro*, consistent with the interactions observed for yeast Utp23. These results suggest that the key interactions in the ES6 region of the 18S rRNA are evolutionarily conserved and point towards a conserved role for Utp23 at this region in yeast and humans.

4.2.4 Human UTP23 is required for cleavage at sites A0, 1 and 2a in HEK293T cells

In *S. cerevisiae*, Utp23 is required for processing of the 35S pre-rRNA at three sites (A0, A1 and A2) leading to the maturation of the SSU 18S rRNA (Bleichert *et al.*, 2006), presumably through its proposed role in coordinating interactions between the ES6 region of 18S, the snR30 snoRNA and other ribosome biogenesis factors in the SSU processome. Utp24 mediates cleavage at sites A1 and A2 via its PIN endonuclease domain (Wells *et al.*, 2016), but the endonuclease responsible for cleavage at A0 is currently unknown. In humans, UTP23 is also a component of the SSU processome, and was shown to localise to the nucleolus and associate with the U3 snoRNA *in vivo* in HEK293 cells (Turner *et al.*, 2012). Depletion of UTP23 in HeLa cells was shown to disrupt endonucleolytic cleavage at sites A0, 1 and 2a, equivalent to yeast sites A0, A1 and A2, suggesting that its role in pre-rRNA processing is conserved between yeast and humans (Sloan *et al.*, 2014).

To confirm this conservation, depletion of UTP23 via RNAi-mediated knockdown was performed in HEK293T embryonic kidney cells (Figure 4.5). A siRNA targeting firefly luciferase (GL2) was used as a control. After 72 hours depletion, cells were harvested, and protein was analysed by western blotting to confirm the knockdown of UTP23 by the UTP23-targeting siRNA, using an anti-UTP23 antibody (Figure 4.5A). Karyopherin was used as a loading control. RNA was extracted from cells following RNAi knockdown and separated by agarose gel electrophoresis. RNA was transferred to a Hybond N membrane and northern blotting was performed using radiolabelled oligonucleotide probes hybridising to two different sequences within the ITS1 of the pre-rRNA, which detect pre-rRNAs leading to the mature 18S rRNA (Figure 4.5C). The ‘ITS1’ probe hybridises to a sequence between cleavage sites 2a and 2, while the ‘18SE’ probe hybridises between the 3’ of 18S and the 2a cleavage site. The position of probes on the pre-rRNA is shown in Figure 4.5B.

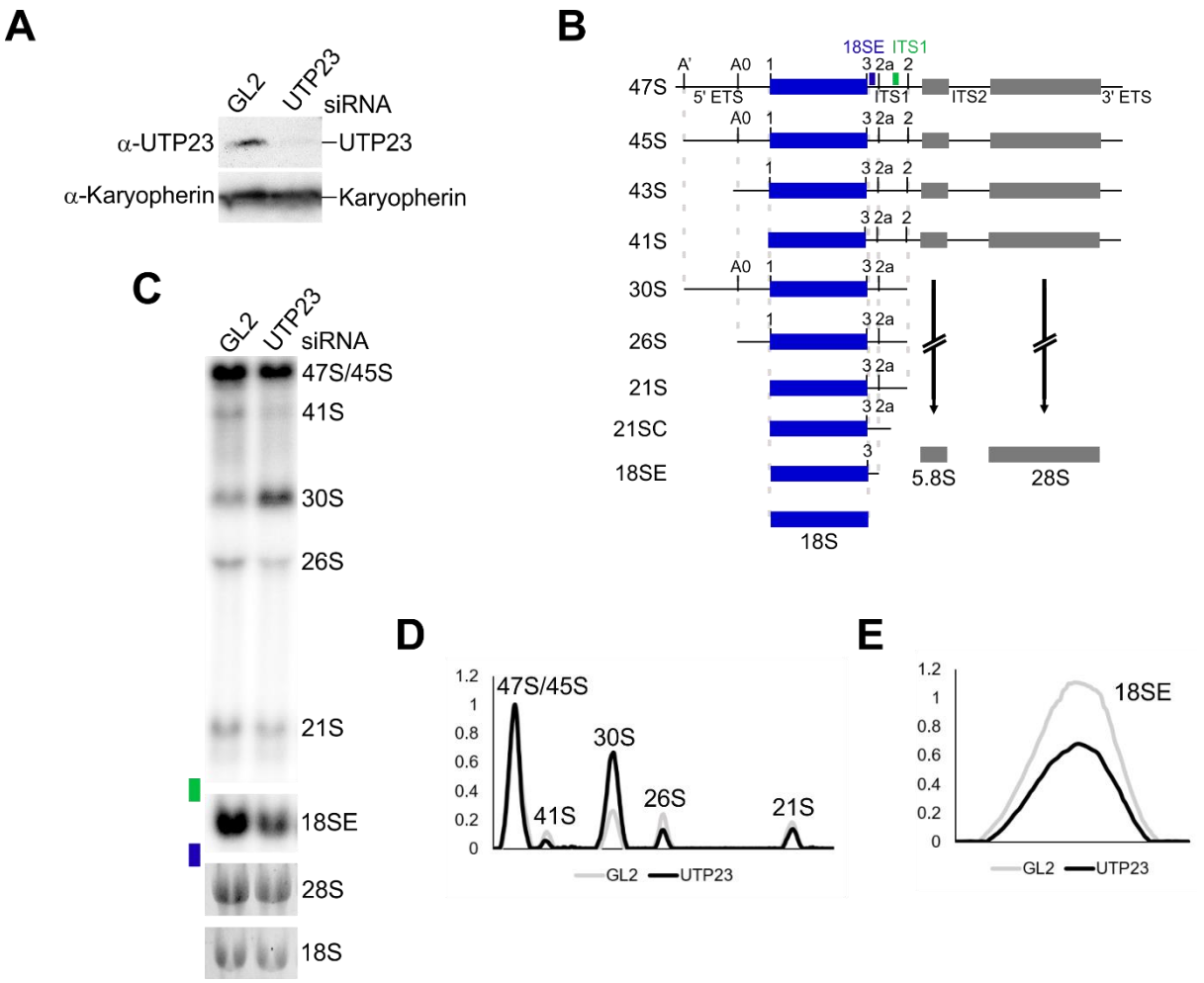


Figure 4.5 UTP23 is required for processing at three early pre-rRNA sites in HEK293T cells. (A) Western blot using an anti-UTP23 antibody to confirm RNAi-mediated depletion of endogenous UTP23 in HEK293T cells. An anti-Karyopherin antibody was used as a loading control. **(B)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(C)** RNA from HEK293T cells treated with control siRNA (GL2) or depleted of UTP23 (UTP23) were analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNA intermediates detected by northern blotting probes are labelled. Levels of pre-rRNA intermediates detected by the **(D)** ITS1 or **(E)** 18SE probe from panel C were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted. The pre-rRNAs represented by each peak are labelled.

Depletion of UTP23 in HEK293T caused an accumulation of the 30S pre-rRNA, which spans from the A' site in the 5' ETS to site 2 in ITS1 (Figure 4.5C). This pre-rRNA phenotype was observed consistently in many independent experiments in these cells. The intensity of bands in Figure 4.5C was quantified, normalised to the intensity

of the 47S/45S band and plotted (Figure 4.5D and E). Accumulation of the 30S pre-rRNA is consistent with normal cleavage at site A' and site 2, but disrupted cleavage at sites A0, 1 and 2a. This is equivalent to the pre-rRNA processing defect observed upon UTP23 depletion in HeLa cells (Sloan *et al.*, 2014), and of Utp23 in yeast cells (Bleichert *et al.*, 2006). These results suggest a conserved role for UTP23 in the processing of pre-rRNA in human cells.

4.2.5 Depletion of UTP23 in U2OS and MCF7 cells results in variable pre-rRNA processing phenotypes

To assess whether the function of UTP23 is conserved across multiple cell types, RNAi-mediated depletion of UTP23 and northern analysis of extracted RNA was performed using U2OS osteocarcinoma and MCF7 breast cancer cells (Figure 4.6 and 4.7). The pre-rRNA processing phenotypes observed upon UTP23 depletion in both cell lines was extremely variable (Figures 4.6 and 4.7). A very mild accumulation of the 30S pre-rRNA was observed in some experiments (Figure 4.6C and D; Figure 4.7C and D), although to a lesser extent than in HEK293T cells. In other experiments, despite confirmation of UTP23 depletion by western blotting (Figure 4.6A; Figure 4.7A), no 30S pre-rRNA accumulation was observed (Figure 4.6C; Figure 4.7C). Interestingly, some experiments appeared to show mild accumulation of the 18SE pre-rRNA (Figure 4.6C and E; Figure 4.7C and E). This is consistent with the pre-rRNA processing phenotype observed upon UTP23 depletion in mouse cells (Wang *et al.*, 2014). The reason for the observed variation in pre-rRNA processing defects in different cell types is unclear.

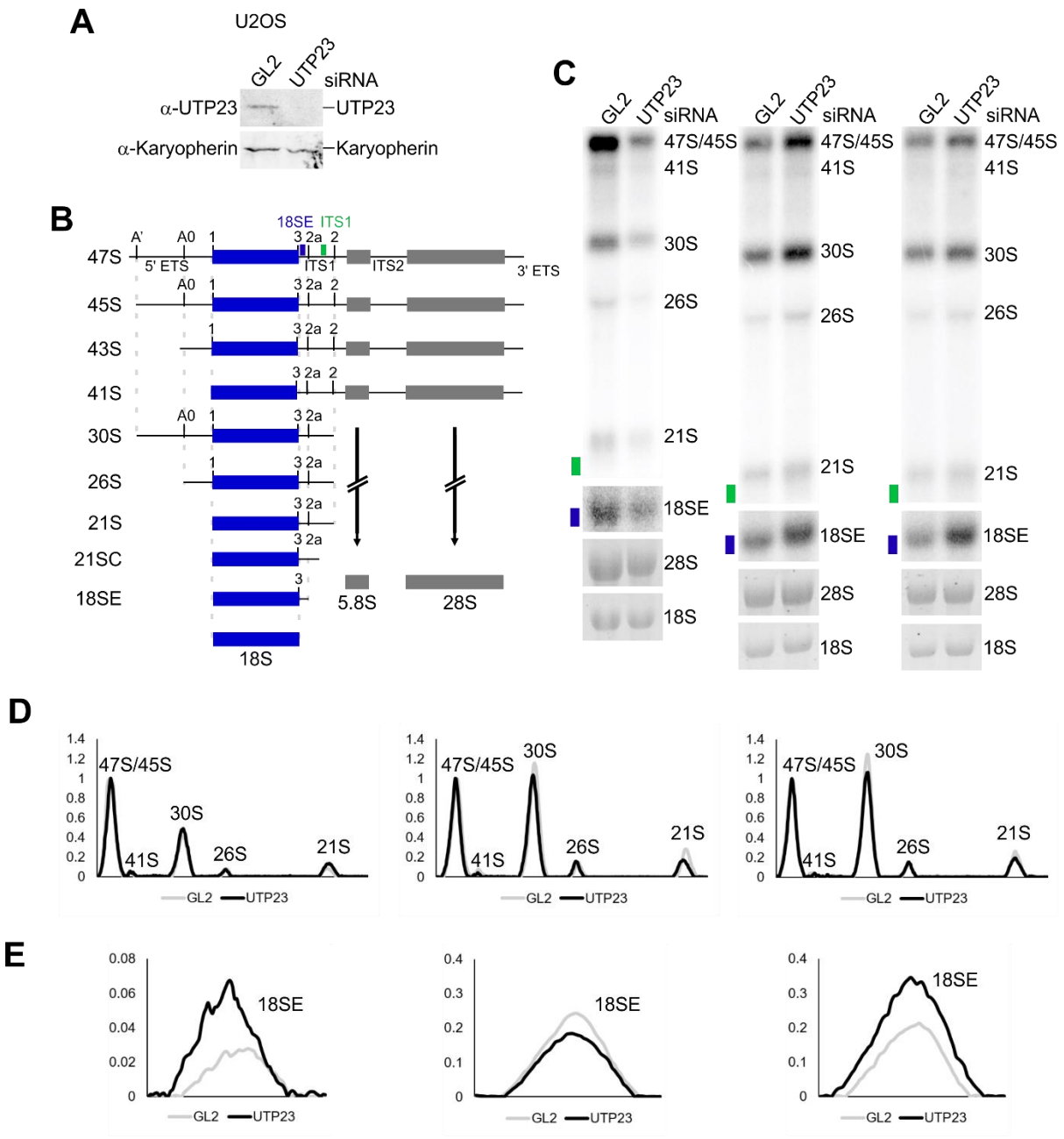


Figure 4.6 Pre-rRNA processing defects upon UTP23 depletion in U2OS cells. (A) Western blot using an anti-UTP23 antibody to confirm depletion of endogenous UTP23 upon RNAi-mediated UTP23 depletion in U2OS cells. Karyopherin was used as a loading control. **(B)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the primary precursor (47S pre-rRNA). **(C)** RNA from U2OS cells treated with control siRNA (GL2) or depleted of UTP23 (UTP23) was analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. Three representative experiments are shown. Levels of pre-rRNAs detected by the **(D)** ITS1 or **(E)** 18SE probe from panel C were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

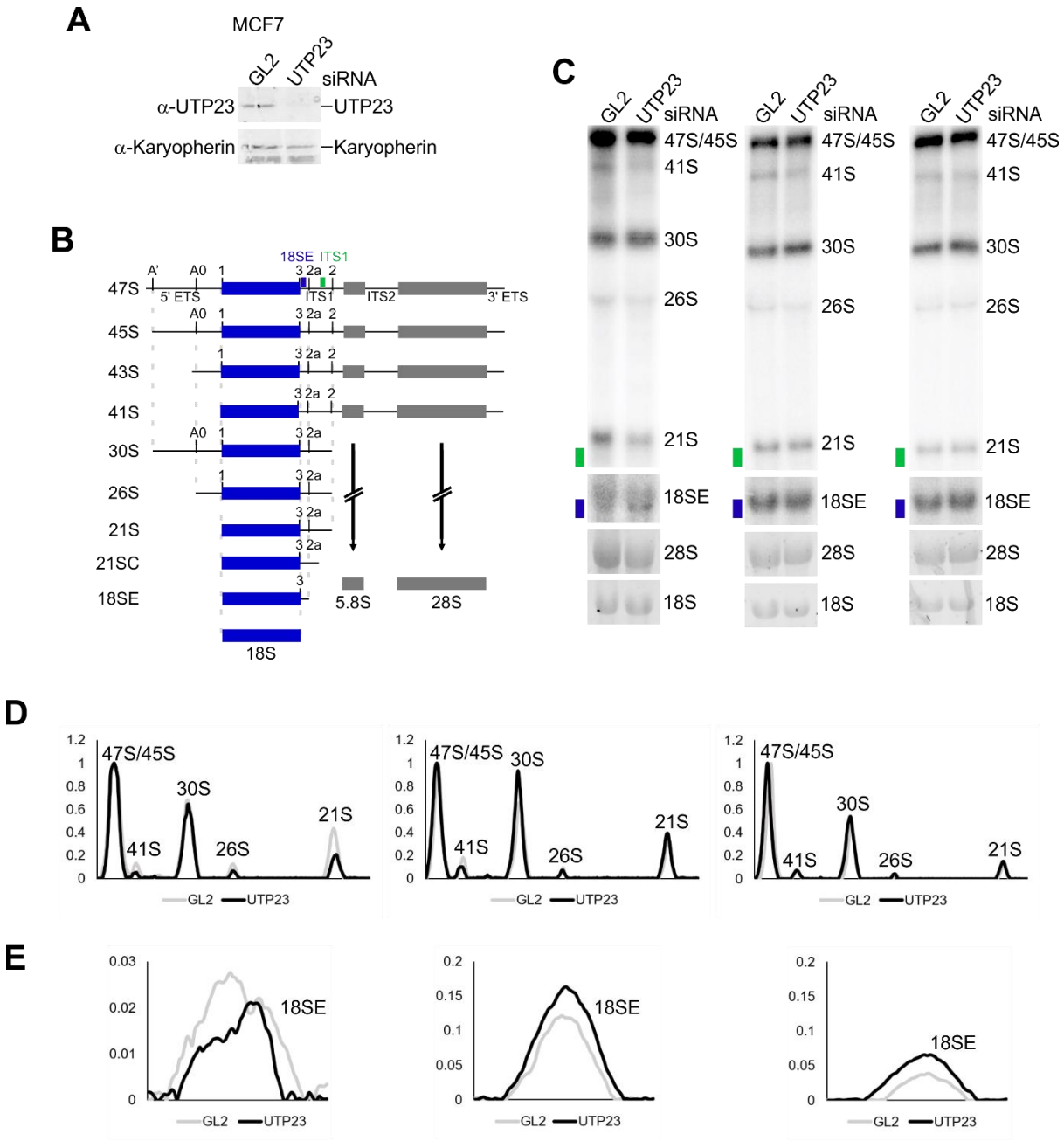


Figure 4.7 Pre-rRNA processing defects upon UTP23 depletion in MCF7 cells. (A) Western blot using an anti-UTP23 antibody to confirm depletion of endogenous UTP23 upon RNAi-mediated UTP23 depletion in MCF7 cells. Karyopherin was used as a loading control. **(B)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(C)** RNA from U2OS cells treated with control siRNA (GL2) or depleted of UTP23 (UTP23) were analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. Three representative experiments are shown. Levels of pre-rRNAs detected by the **(D)** ITS1 or **(E)** 18SE probe from panel C were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

4.2.6 FLAG-tagged WT UTP23 was unable to rescue the pre-rRNA processing defect caused by UTP23 depletion

While pre-rRNA processing phenotypes appear to differ in different cell types, the defect seen with depletion of UTP23 in HEK293T cells, namely accumulation of the 30S intermediate and a reduction in 18SE levels, is consistent with that observed upon UTP23 depletion in HeLa cells and Utp23 depletion in yeast. It is also consistent with the known role of both yeast Utp23 and human UTP23 as a component of the SSU processome. For this reason, further experiments investigating the function of human UTP23 utilised HEK293T cells. The specific function of UTP23 within the SSU processome is unknown. Both yeast and human Utp23 contain a PIN endonuclease domain, but yeast Utp23 contains only two of the four conserved acidic amino acids that make up the catalytic site of PIN domains. Consistent with this, mutation of the two conserved residues did not affect pre-rRNA processing (Bleichert *et al.*, 2006). In contrast, human UTP23 contains conserved acidic residues at three positions in the PIN domain catalytic site, and it has previously been shown for the PIN domain endonuclease, SMG6, which is involved in nonsense mediated decay (NMD), that conservation of three acidic residues are sufficient for endonuclease activity (Glavan *et al.*, 2006). Both yeast and human Utp23 also contain a CCHC Zn finger motif predicted to be involved in RNA binding. Indeed, the Zn finger motif of yeast Utp23 has been shown to be essential for the binding of Utp23 to snR30 *in vitro* (Wells *et al.*, 2017). To examine the importance of the PIN domain and Zn finger motif of human UTP23, an RNAi-rescue system was attempted, as described for UTP24 and RCL1 in chapter three.

The open reading frame (ORF) of UTP23 was first altered to make the mRNA resistant to the siRNA used to deplete endogenous UTP23, without altering the amino acid sequence. This construct was then cloned into the pcDNA5 plasmid which expresses the protein fused with 2x N-terminal FLAG tags under the control of a tetracycline-inducible promoter. pcDNA5-FLAG-UTP23 constructs were stably transfected into HEK293T Flp-In cells using the Flp-In recombination system (Life Technologies). Following successful selection of stably transfected cells, the expression of FLAG-tagged WT UTP23 was analysed using a range of tetracycline concentrations (0-1000 ng/µl), in an attempt to achieve expression to a similar level to that of endogenous UTP23 (Figure 4.8A). Cells were harvested after 72 hours of tetracycline induction and whole cell extract was run on a polyacrylamide gel. Protein

levels were analysed by western blotting using an anti-UTP23 antibody, and karyopherin was used as a loading control (Figure 4.8A). The anti-UTP23 antibody detected both endogenous UTP23 and the larger, FLAG-tagged UTP23 protein, and showed that FLAG-tagged UTP23 was efficiently expressed at higher levels than endogenous UTP23, even with very low tetracycline concentrations.

After confirming expression of FLAG-UTP23 from stable cells, the RNAi-resistance of the FLAG-tagged WT UTP23 construct was assessed (Figure 4.8B). Expression of FLAG-UTP23 was induced with 0.5 ng/µl tetracycline and cells were transfected with either a control siRNA (GL2), or a siRNA targeting UTP23. Protein levels were analysed by western blotting using an anti-UTP23 antibody and karyopherin as a loading control (Figure 4.8B). Treatment with 0.5 ng/µl tetracycline resulted in expression of the FLAG-tagged protein at levels closer to, but still higher than, those of endogenous UTP23. Cells stably transfected with an empty pcDNA5 plasmid, expressing the FLAG-tag alone (HEK-pcDNA5) were used as a control. In these cells, endogenous UTP23 was efficiently depleted with the UTP23 siRNA. In cells expressing RNAi-resistant WT UTP23, endogenous UTP23 was efficiently depleted but levels of the FLAG-tagged protein remained stable. This confirmed that expression of this protein was resistant to the UTP23 siRNA.

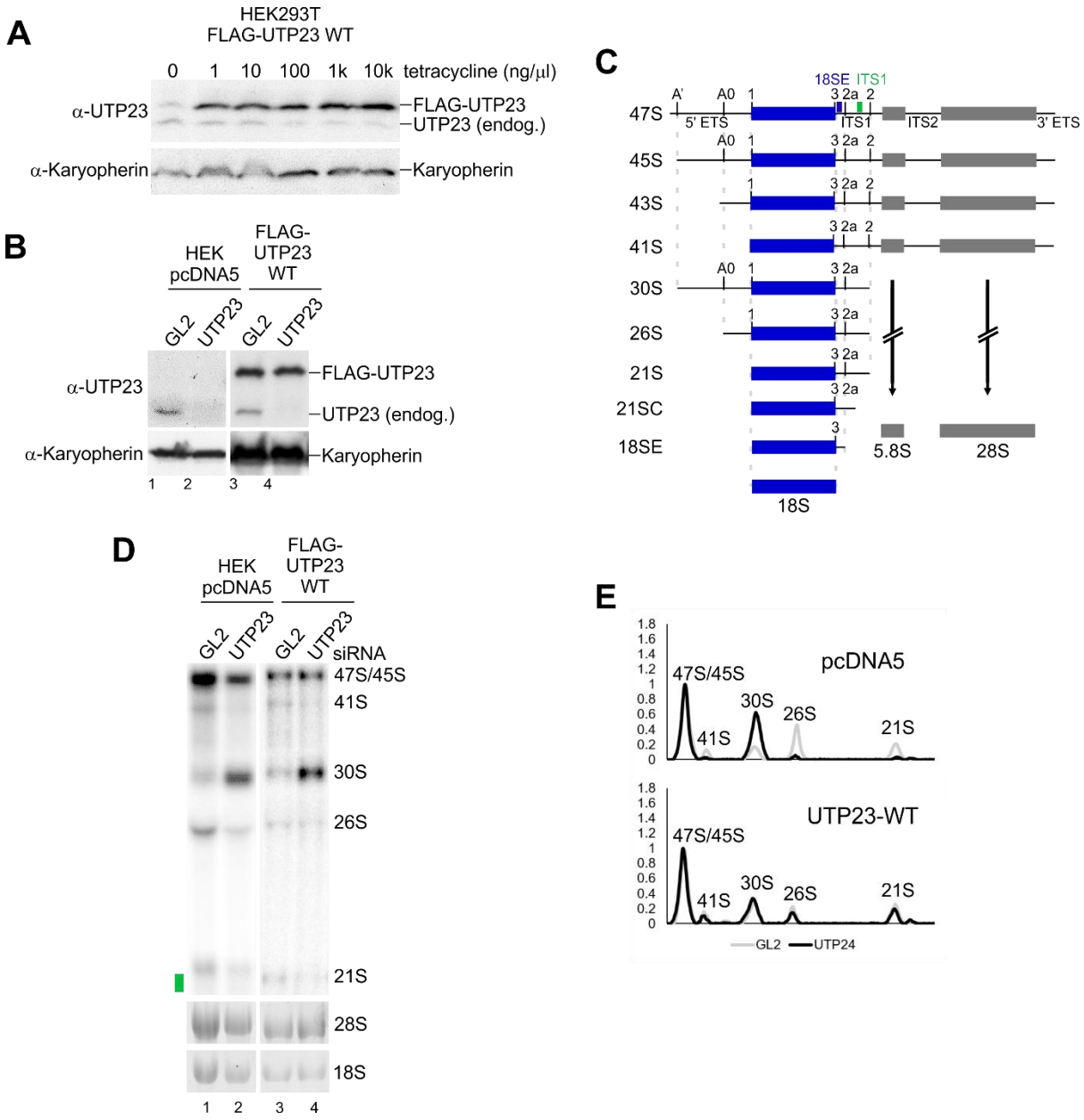


Figure 4.8 Expression of N-terminally FLAG-tagged RNAi-resistant UTP23 does not rescue the pre-rRNA processing phenotype caused by UTP23 depletion in HEK293T cells. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5), or N-terminally FLAG-tagged wild type (WT) RNAi-resistant UTP23. **(A)** HEK293T cells stably transfected with wild type (WT) UTP23 were incubated with a range of tetracycline concentrations (0-10,000 ng/µl) and protein was analysed by western blotting using an anti-UTP23 antibody to detect both endogenous and FLAG-tagged UTP23. An anti-karyopherin antibody was used as a loading control. **(B)** Western blot using an anti-UTP23 antibody to confirm RNAi-mediated depletion of endogenous UTP23 and expression of FLAG-tagged RNAi-resistant UTP23 in stable HEK293T cells. Karyopherin was used as a loading control. **(C)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of the probe used in northern blotting (ITS1; green) is marked on the 47S pre-rRNA. **(D)** Following induction of FLAG-tagged UTP23 expression and treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous UTP23 (UTP23), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane. RNA was

analysed by northern blotting using a radiolabelled probe hybridising to upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. **(E)** Levels of pre-rRNAs detected by the ITS1 probe from panel B were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

To confirm that RNAi-resistant, FLAG-tagged WT UTP23 could rescue the pre-rRNA processing defect observed upon depletion of endogenous UTP23, RNA was extracted from cells expressing either the FLAG-tag alone (HEK-pcDNA5) or FLAG-UTP23. Northern blotting was used to analyse pre-rRNA levels, using 'ITS1' and '18SE' oligonucleotide probes (Figure 4.8D). The position of the probes on the pre-rRNA is marked in Figure 4.8C. In control cells, depletion of UTP23 caused accumulation of the 30S pre-rRNA compared to the control, as seen previously, indicating defective cleavage at sites A0, 1 and 2a. Expression of RNAi-resistant FLAG-tagged UTP23 failed to rescue this pre-rRNA processing phenotype, as depletion of endogenous UTP23 in these cells caused a similar accumulation of 30S pre-rRNA. The fact that this phenotype is observed despite FLAG-UTP23 being expressed at levels similar to or even higher than the endogenous protein (Figure 4.8B) suggest that this protein is non-functional. This is likely due to the size and/or position of the N-terminal FLAG-tag, which may disrupt a vital function of UTP23 in binding of the pre-rRNA or to other protein factors or assembly into the SSU processome *in vivo*.

4.2.7 HA-tagged UTP23 was able to compensate for the depletion of endogenous UTP23

As RNAi-resistant UTP23 containing an N-terminal FLAG-tag was unable to rescue the pre-rRNA processing phenotype caused by UTP23 depletion, an alternative strategy was used to establish the RNAi rescue system. An RNAi-resistant WT UTP23 construct was designed and generated with 2x C-terminal HA-tags instead of the FLAG-tag and cloned into pcDNA5. After generation of stable cell lines, expression of UTP23-HA was induced with tetracycline as for the FLAG-tagged protein above, and the concentration of tetracycline required to express the exogenous protein to similar levels to endogenous UTP23 was determined (see Figure 4.8A). The RNAi-resistance of this construct was analysed by western blotting after RNAi-mediated knockdown using either a control (GL2) siRNA or a siRNA targeting UTP23 mRNA (Figure 4.9A). In control cells, expressing a FLAG-tag alone (pcDNA5), endogenous UTP23 was

efficiently depleted by the UTP23 siRNA. In cells expressing UTP23-HA, endogenous UTP23 was efficiently depleted while RNAi-resistant UTP23-HA levels remained stable.

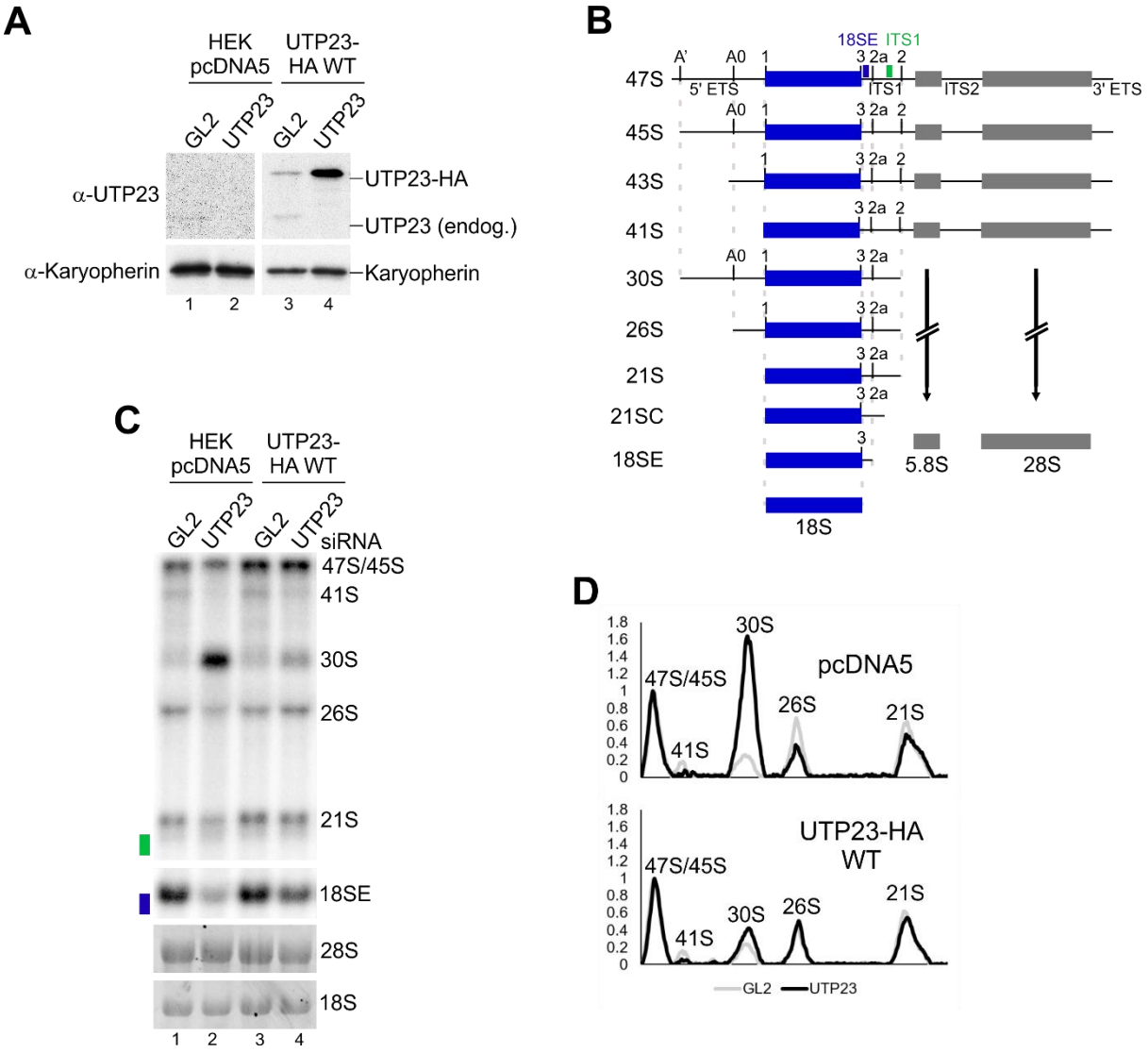


Figure 4.9 Expression of C-terminally HA-tagged RNAi-resistant UTP23 rescues the pre-rRNA processing phenotype caused by UTP23 depletion in HEK293T cells. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag besalone (HEK pcDNA5), or C-terminally HA-tagged wild type (WT) RNAi-resistant UTP23. **(A)** Western blot using an anti-UTP23 antibody to confirm depletion of endogenous UTP23 upon RNAi-mediated UTP23 depletion and expression of HA-tagged RNAi-resistant UTP23 in stable HEK293T cells. An anti-Karyopherin antibody was used as a loading control. **(C)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(D)** Following treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous UTP23 (UTP23), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane. RNA was

analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. **(E)** Levels of pre-rRNAs detected by the ITS1 probe from panel B were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted. The pre-rRNA intermediate represented by each peak is labelled.

Total RNA was extracted from cells analysed in panel A following tetracycline induction and RNAi-mediated knockdown of UTP23. Pre-rRNA levels were analysed by northern blotting using 'ITS1' and '18SE' radiolabelled probes. The positions of the probes on the pre-rRNA are marked in Figure 4.9B. In cells expressing the empty FLAG-tag, depletion of UTP23 caused strong accumulation of the 30S pre-rRNA and a strong reduction in 18SE pre-rRNA levels (Figure 4.9C, lane 2). In cells expressing RNAi-resistant HA-tagged UTP23 these pre-rRNA processing defects were significantly reduced (lane 4). This suggests that, unlike the N-terminally FLAG-tagged protein, C-terminally HA-tagged UTP23 is able to functionally replace endogenous UTP23 *in vivo*.

4.2.8 The conserved PIN domain and Zn finger motif of human UTP23 are essential for pre-rRNA processing

After confirming that HA-UTP23 is functional, substitutions were made in the sequence of the pcDNA5-UTP23-HA construct to generate PIN domain (D31N) and Zn finger motif (C103A) mutants. The first of the four PIN domain acidic residues was substituted for an asparagine (D31N), and a predicted essential cysteine residue of the Zn finger motif was substituted for an alanine (C103A). The human C103A mutant is equivalent to the yeast mutant that abolished *in vitro* binding of Utp23 to snR30 (Wells *et al.*, 2017). These were stably transfected into HEK293T Flp-In cells and a tetracycline induction titration was performed to determine the concentration required to express the mutant proteins to similar levels to endogenous UTP23. After tetracycline induction, proteins were analysed by western blotting using an anti-UTP23 antibody to detect both endogenous and HA-tagged UTP23 (Figure 4.10). Karyopherin was used as a loading control. While WT UTP23-HA (Figure 4.10A) was expressed to levels similar to endogenous UTP23 with 0.5-1 ng/µl tetracycline, PIN domain (50-100 ng/µl) (Figure 4.10B) and Zn finger motif (100 ng/µl) (Figure 4.10C) mutants required higher concentrations of tetracycline, but both mutants were efficiently expressed.

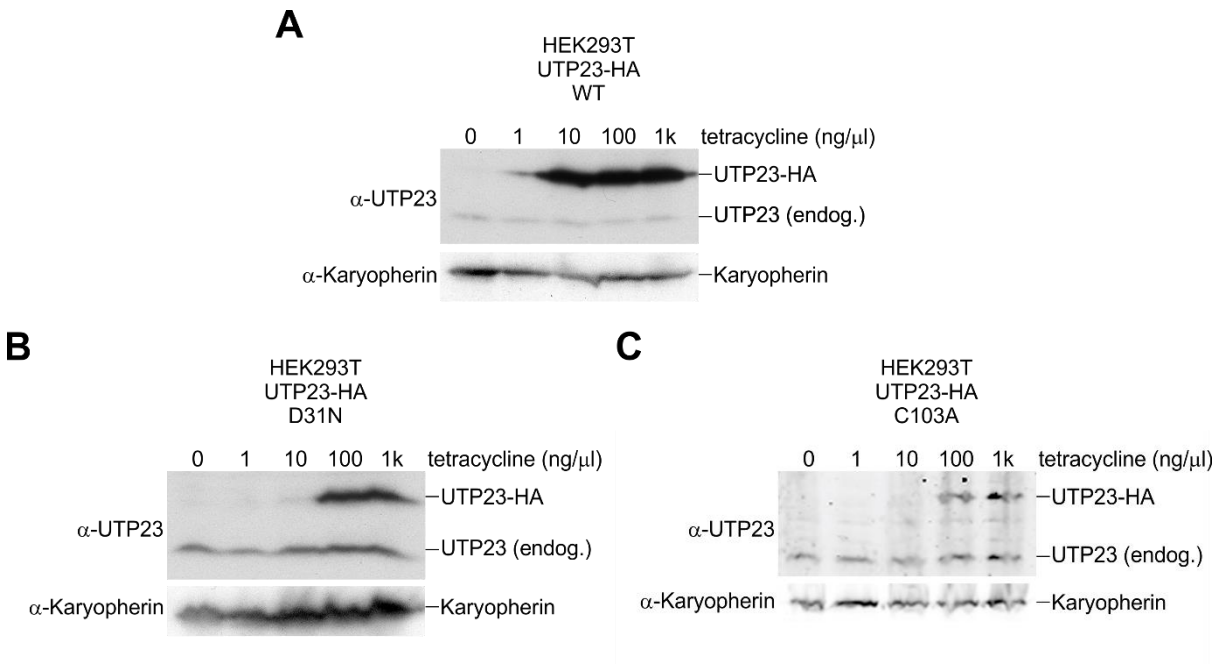


Figure 4.10 Expression of RNAi-resistant UTP23 mutant proteins in stable HEK293T cell lines. HEK293T cells stably transfected with HA-tagged wild type (WT; panel A) or mutant (D31N; panel B, C103A; panel C) UTP23 were incubated with a range of concentrations of tetracycline (0-1,000 ng/μl). Protein was separated by SDS-PAGE and analysed by western blotting using an anti-UTP23 antibody to detect both endogenous and HA-tagged UTP23. Karyopherin was used as a loading control.

An RNAi rescue system was established with cells expressing the HA-tag alone (HEK-pcDNA5), WT UTP23-HA, or PIN domain or Zn finger motif mutant forms of UTP23-HA. Protein levels were assessed by western blotting using an anti-UTP23 antibody, with karyopherin used as a loading control, following expression of UTP23-HA and depletion of endogenous UTP23 (Figure 4.11). Due to the variation in response to tetracycline for the mutant proteins, it was difficult to titrate expression levels of all HA-tagged proteins to similar levels to endogenous UTP23. Therefore, two different WT samples were used for the RNAi rescue experiment, one expressed at levels slightly higher than that of the endogenous protein and comparable to UTP23-D31N levels (WT_1) (lanes 3 and 4) and another at slightly lower levels and comparable to that of the UTP23-C103A mutant (WT_2) (lanes 7 and 8). In all cells, endogenous UTP23 was efficiently depleted by the UTP23 siRNA compared to the control (GL2) siRNA. Levels of WT and mutant RNAi-resistant HA-tagged UTP23 remained stable after treatment with the UTP23 siRNA.

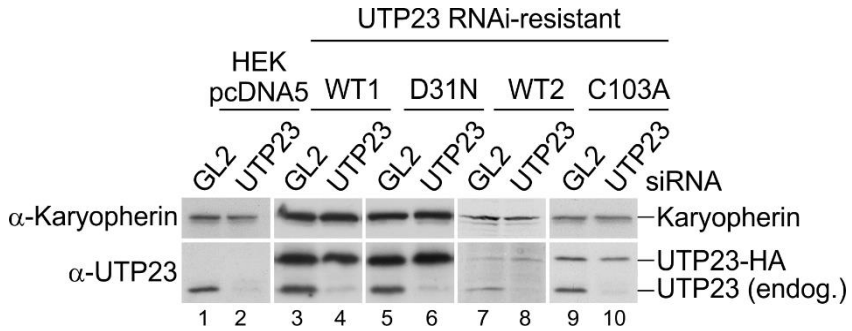


Figure 4.11 Protein analysis of HEK293T UTP23 RNAi rescue. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5; lanes 1 and 2), or wild type (WT1/WT2; lanes 3 and 4/ lanes 7 and 8) or mutant (D31N; lanes 5 and 6, C103A; lanes 9 and 10) HA-tagged RNAi-resistant UTP23. Following transfection of a control siRNA (GL2) or RNAi-mediated depletion of endogenous UTP23 (UTP23), protein was extracted from cells and separated by SDS-PAGE. Protein was analysed by western blotting using an anti-UTP23 antibody. An anti-Karyopherin antibody was used as a loading control.

To investigate the importance of the PIN domain and Zn finger motif of UTP23 on its essential role in pre-rRNA processing, RNA from the RNAi rescue experiment was extracted and separated by agarose gel electrophoresis. Northern blotting was performed using 'ITS1' and '18SE' probes (Figure 4.12A), and the location of these probes on the pre-rRNA are shown in Figure 4.12B. In cells transfected with an empty control vector (pcDNA5), depletion of UTP23 caused accumulation of the 30S pre-rRNA (Figure 4.12A, lanes 1 and 2) and significant reduction of the 18SE pre-rRNA, as seen previously, indicating defective cleavage at sites A0, 1 and 2a. Expression of RNAi-resistant UTP23-HA at levels similar to UTP23-D31N-HA (WT_1) (lanes 3 and 4) or similar to UTP23-C103A-HA (WT_2) (lanes 7 and 8) rescued this phenotype efficiently, restoring pre-rRNA levels back to those observed with the control (GL2) siRNA. To confirm the observed rescue of this phenotype, levels of pre-rRNAs in Figure 4.12A were quantified, normalised to levels of the 47S/45S band and plotted (Figure 4.12C-E). This shows that 30S pre-rRNA levels were similar to those in GL2-treated cells when WT UTP23 was expressed exogenously. Depletion of endogenous UTP23 in cells expressing the RNAi-resistant PIN domain mutant (D31N) form of UTP23 caused a strong accumulation of the 30S pre-rRNA compared to the control (lanes 5 and 6). This suggests that the conserved PIN domain in UTP23 is essential for efficient cleavage of pre-rRNA at sites A0, 1 and 2a. In cells expressing the Zn finger motif mutant (C103A), depletion of endogenous UTP23 also caused

accumulation of the 30S pre-rRNA (lanes 9 and 10), suggesting that the Zn finger motif of UTP23 is also essential for normal cleavage at these sites.

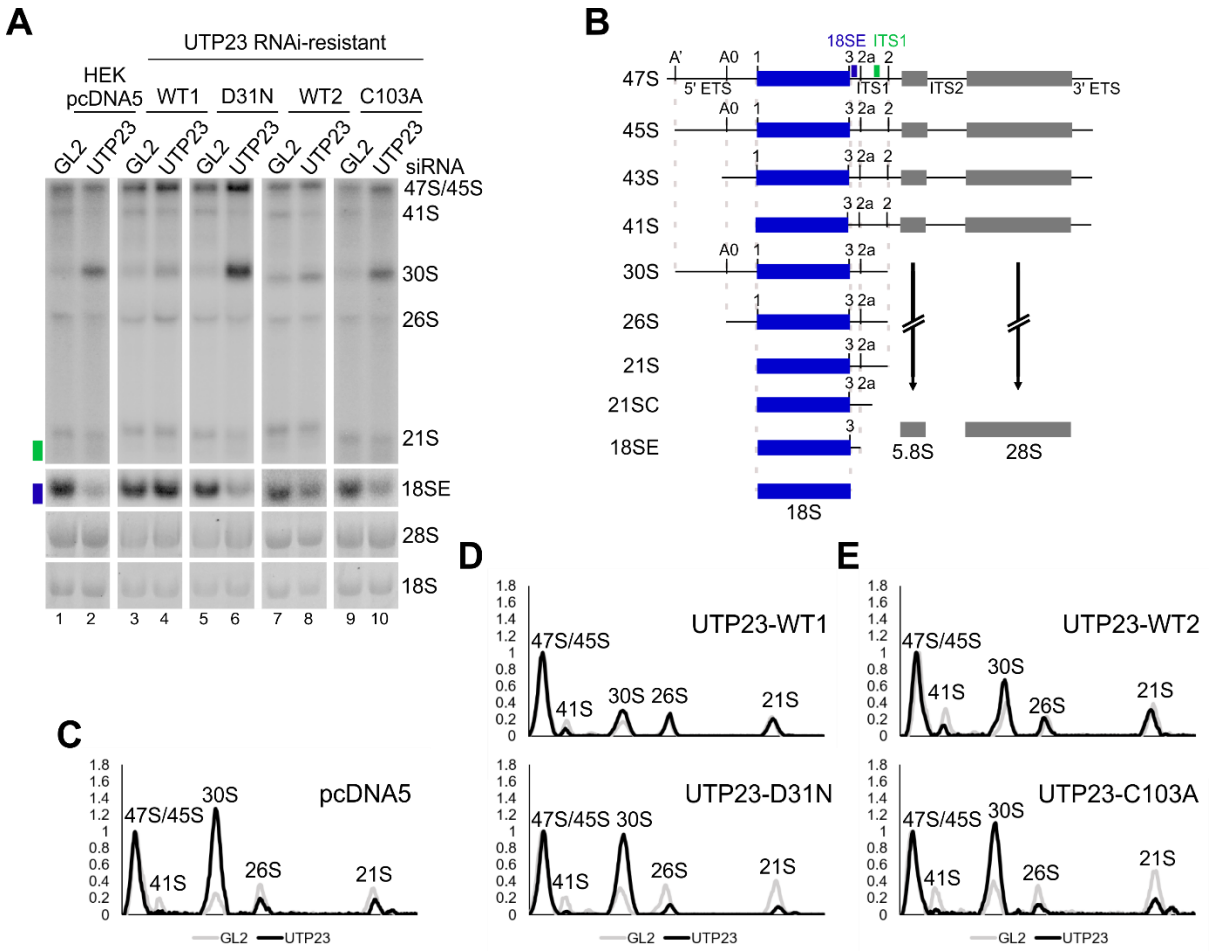


Figure 4.12 Intact PIN domain and Zinc finger in UTP23 are required for pre-rRNA processing at three early cleavage sites in HEK293T cells. Northern analysis of RNA from HEK293T Flp-In cells stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5; panel A lanes 1 and 2), or wild type (WT1/WT2; panel A lanes 3 and 4/ lanes 7 and 8) or mutant (D31N; panel A lanes 5 and 6, C103A; panel A lanes 9 and 10) HA-tagged RNAi-resistant UTP23. **(A)** Following treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous UTP23 (UTP23), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane. RNA was analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. **(B)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(C-E)** Levels of pre-rRNAs detected by the ITS1 probe from panel B were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

The pre-rRNA processing defect observed upon expression of a PIN domain mutant form of UTP23 and the fact that the PIN domain of human UTP23, unlike the yeast protein, contains three conserved acidic residues, suggests that the human UTP23 PIN domain may be active in endonuclease activity and play a direct enzymatic role in pre-rRNA processing. Accumulation of the 30S pre-rRNA is consistent with disrupted cleavage at site A0, as subsequent cleavages at sites 1 and 2a are linked to A0 cleavage, so disruption of cleavage at A0 would also lead to defects in downstream cleavage at sites 1 and 2a. The endonuclease responsible for cleavage at A0 has not been identified, raising the exciting possibility that UTP23 could be responsible for processing at this site. The pre-rRNA processing phenotype seen with mutation of the Zn finger motif of UTP23 suggests that the predicted RNA-binding function of this motif is important for processing of pre-rRNA at these three sites. This is consistent with data indicating that the Zn finger motif of yeast Utp23 is essential for yeast growth (Lu *et al.*, 2013) and is required for its interaction with the snR30 snoRNA *in vitro* (Wells *et al.*, 2017). Given that human UTP23 associates with the human homologue of snR30, the U17 snoRNA *in vivo* (Hoareau-Aveilla *et al.*, 2012) and *in vitro* (Figure 4.4), the disruption of the Zn finger motif may affect the interaction with U17 snoRNA *in vivo*, leading to the observed pre-rRNA processing defect. However, the processing defect observed with these mutants is identical to that observed when UTP23 is depleted in control cells. Therefore, these observations could alternatively be explained by other defects in the mutant proteins that effect the essential function of UTP23 as part of the SSU processome.

4.2.9 UTP23 PIN domain and Zn finger motif mutants localise to the nucleolus

To confirm that the mutations in the PIN domain and Zn finger motif of UTP23 do not cause defects in the essential functions of UTP23 not directly related to pre-rRNA cleavage, the localisation of mutant proteins was first examined by immunofluorescence (Figure 4.13). Stably transfected HEK293T cells were incubated on coverslips with the appropriate tetracycline concentration and immunofluorescence was performed using an anti-HA antibody to detect UTP23-HA. An anti-fibrillarin antibody was used as a nucleolar marker and cell nuclei were stained with DAPI. Fibrillarin localised exclusively to the dense fibrillar nucleolar component, as observed previously (Turner *et al.*, 2012). In control cells (pcDNA5), no signal was detected with the anti-HA antibody. In cells expressing WT UTP23-HA, signal was detected

throughout the nucleolus with the anti-HA antibody, consistent with previous observations and with the predicted role of UTP23 in pre-rRNA processing as part of the SSU processome (Turner *et al.*, 2012). Both PIN domain and Zn finger mutant forms of UTP23 also localised correctly to the nucleolus, indicating that these proteins are not defective in their nuclear or nucleolar import and suggesting they may be efficiently recruited to pre-ribosomes in the nucleolus.

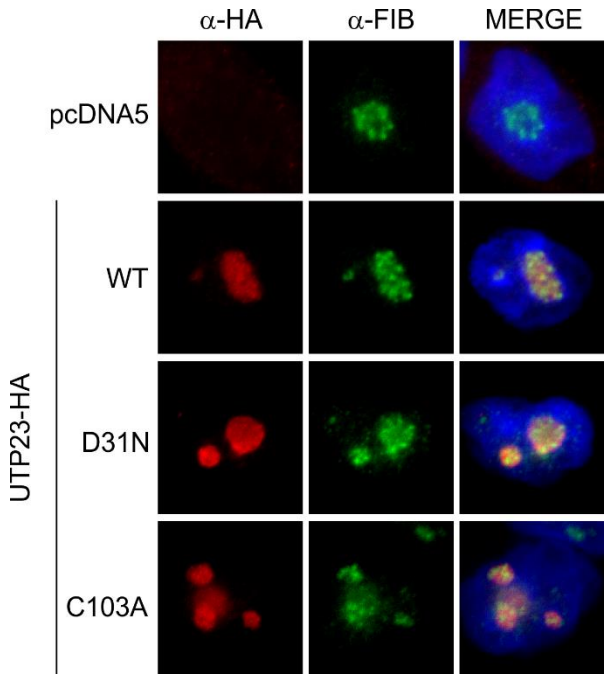


Figure 4.13 UTP23 PIN domain and Zn finger motif mutants localise to the nucleolus in HEK293T cells. Immunofluorescence was performed on HEK293T cells stably expressing a FLAG-tag alone (pcDNA5), HA-tagged UTP23 wild type (WT) or mutant UTP23 (D31N; C103A) using an anti-HA antibody (left column). An anti-Fibrillarin antibody (FIB) was used as a nucleolar marker (middle column). Anti-HA signal is shown in red, anti-Fibrillarin signal is shown in green and DAPI staining of the nucleus is shown in blue.

4.2.11 The UTP23 Zn finger mutant does not associate with the U17 snoRNA in vivo, while the PIN domain mutant shows reduced affinity for U17

To further examine the reason for the pre-RNA processing phenotype observed upon expression of UTP23 mutants, the association of these proteins with the U17 snoRNA was assessed *in vivo* and *in vitro*. To investigate the *in vivo* association of UTP23-HA with the U17 snoRNA, immunoprecipitation was performed using anti-HA antibody-coupled agarose beads with sonicated whole-cell extracts from stably

transfected HEK293T cells expressing a FLAG-tag alone (HEK-pcDNA5), WT UTP23-HA or PIN domain mutant (D31N) or Zn finger motif mutant (C103A) UTP23-HA (Figure 4.14). Co-precipitated RNA was extracted, separated by polyacrylamide gel electrophoresis and analysed by northern blotting using radiolabelled oligonucleotide probes. A probe hybridising to the U17 snoRNA sequence was used to detect U17 co-precipitation, and a probe hybridising to the small nuclear RNA, 7SK, was used as a control to test for non-specific binding.

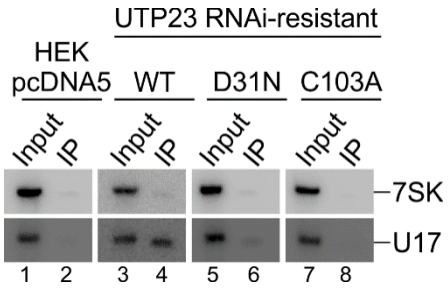


Figure 4.14 The PIN domain and Zinc finger motif of UTP23 are important for its association with the U17 snoRNA *in vivo*. Immunoprecipitation was performed on lysates from cells expressing a FLAG-tag alone (pcDNA5), HA-tagged UTP23 wild type (WT) or mutant UTP23 (D31N; C103A) using an anti-HA antibody. Co-precipitated RNA was extracted and analysed by northern blotting using a probe hybridising to the U17 snoRNA. A probe hybridising to the small nuclear RNA 7SK was used as a loading control. 1.25% of the input material was loaded. RNA was visualised using a PhosphorImager.

In control cells (lanes 1 and 2), no co-precipitation of U17 was observed. WT UTP23-HA efficiently co-precipitated U17, consistent with its known association with U17 *in vivo* (lanes 3 and 4). The UTP23 PIN domain mutant (D31N) (lanes 5 and 6) showed significantly reduced, but still detectable, co-precipitation of the U17 snoRNA, while the Zn finger motif mutant (C103A) (lanes 7 and 8) did not co-precipitate U17 above background levels. These results suggest that both PIN domain and Zn finger motif mutant forms of UTP23 are somewhat defective in association with the U17 snoRNA, with the Zn finger mutation having a more severe effect.

4.2.12 The UTP23 PIN domain mutant binds the U17 snoRNA in vitro while the Zn finger motif mutant shows significantly reduced binding

To assess the ability of mutant forms of UTP23 to bind the U17 snoRNA directly, an *in vitro* RNA binding assay (EMSA) was performed using recombinant GST-tagged

UTP23 proteins and *in vitro*-transcribed, radiolabelled U17 RNA (Figure 4.15). After incubation of RNA and recombinant proteins at a range of concentrations, free RNA and RNP complexes were separated by polyacrylamide gel electrophoresis. Previous data has shown that WT UTP23 directly binds to the U17 snoRNA (Figure 4.4) and this was again confirmed in these experiments (Figure 4.15A**,** upper panel**)**. The PIN domain mutant (D31N) (Figure 4.15A, lower panel) directly bound the U17 snoRNA with similar affinity to the WT protein (upper panel). Quantification of bands from 4.15A are shown in Figure 4.15B. In contrast, the Zn finger motif mutant (C103) (Figure 4.14C, lower panel), showed considerably to U17 compared to the WT protein at the same concentrations (upper panel). Quantification of bands from 4.15C are shown in Figure 4.15D. These results suggest that mutation of the PIN domain of UTP23 does not affect its ability to directly bind the U17 snoRNA, whereas the putative RNA-binding function of the Zn finger motif is important for this interaction.

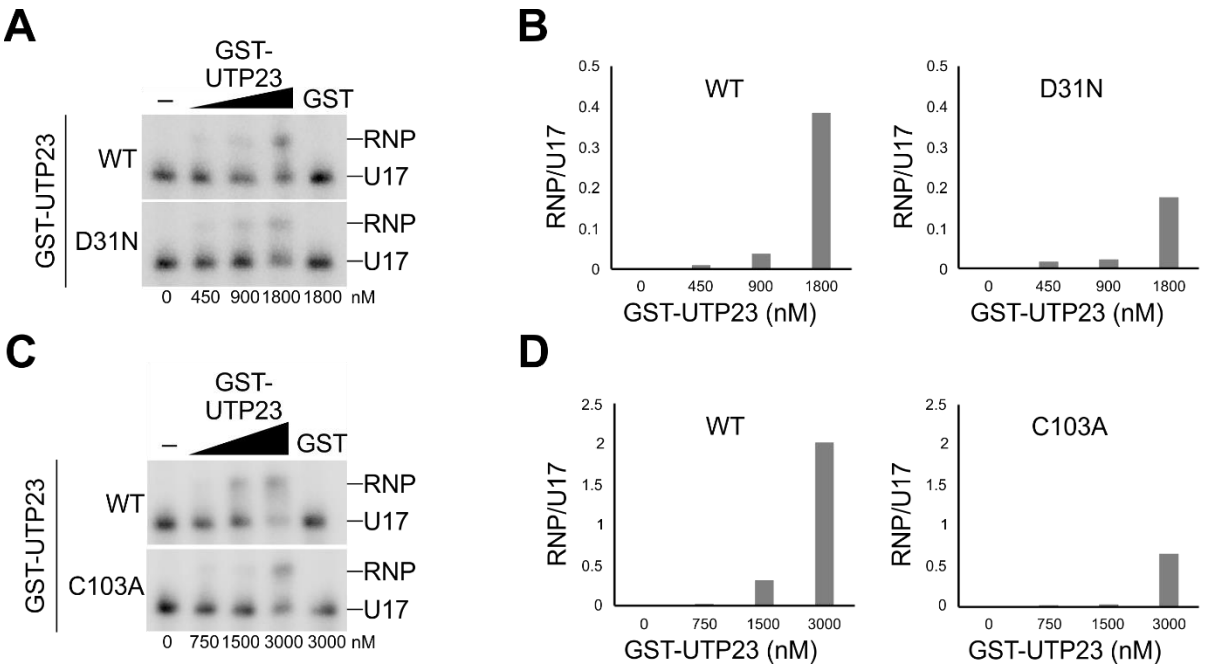


Figure 4.15 An intact Zinc finger motif is important for the direct binding of UTP23 to the U17 snoRNA *in vitro*. Electromobility shift assay (EMSA) to analyse binding of recombinant wild type (WT) and **(A)** PIN domain mutant (D31N) (0-1,000 nM) or **(B)** Zinc finger motif mutant (0-3,000 nM) forms of GST-UTP23 protein or GST (1,800 nM/ 3,000 nM) to radiolabelled *in vitro* transcribed U17 RNA substrate. RNP complexes were separated from free U17 RNA on a 4% native polyacrylamide gel and visualised using a PhosphorImager. **(B and D)** Bands from panels A and C were quantified using ImageQuant and plotted.

4.2.13 Attempts to confirm endonucleolytic activity of UTP23 in vitro

The fact that, unlike the yeast protein, the PIN domain of human UTP23 contains three conserved acidic amino acids in its proposed catalytic site, and that mutation of one of these residues disrupts pre-rRNA processing *in vivo*, suggests that human UTP23 could play an active role as an endonuclease in ribosome biogenesis. Expression of the UTP23 PIN domain mutant (D31N) caused accumulation of the 30S pre-rRNA, indicating defective cleavage at sites A0, 1 and 2a. This is consistent with a block in cleavage at A0, and the endonuclease responsible for cleavage at this site is currently unknown, pointing to a potential role for human UTP23 in processing A0. To assess this possibility, *in vitro* RNA cleavage assays were performed using WT or PIN domain mutant GST-tagged recombinant UTP23 and *in vitro* transcribed RNA substrates. Initial nuclease assays used a pre-rRNA mimic containing the human A0 site (5' ETS nt 1601-1700). After incubation of RNA and protein, RNA was extracted and analysed by primer extension using a radiolabelled oligonucleotide downstream of the A0 cleavage site. The analysis did not show any strong primer extension stops, indicating RNA cleavage, at A0 or any other site after incubation with either WT or PIN domain mutant forms of recombinant UTP23 (data not shown).

Further RNA cleavage assays were performed using 5' end radiolabelled RNA oligonucleotides (Figure 4.16). Recombinant GST-tagged UTP23 (WT or PIN domain mutant, D31N) was incubated with radiolabelled RNA substrate containing the human A0 cleavage site (5' ETS nt 1623-1648) and separated on a 12% PAA/8M urea gel. RNA cleavage assays were performed in the presence of 5 mM or 10mM Mn^{2+}, which has been shown to be important for the *in vitro* nuclease activity of other PIN domain-containing endonucleases (Schneider *et al.*, 2009; Skruzny *et al.*, 2009; Wells *et al.*, 2016). A sequencing ladder was produced for the RNA substrate using partial alkaline hydrolysis, which causes cleavage at every nucleotide. This allowed for detection of the position of the A0 site within the RNA oligonucleotide. Despite not observing any specific cleavage activity at the A0 site, there were several detectable cleavages at other sites in the RNA substrate after incubation with WT GST-UTP23, but not with GST-UTP23-D31N (Figure 4.16).

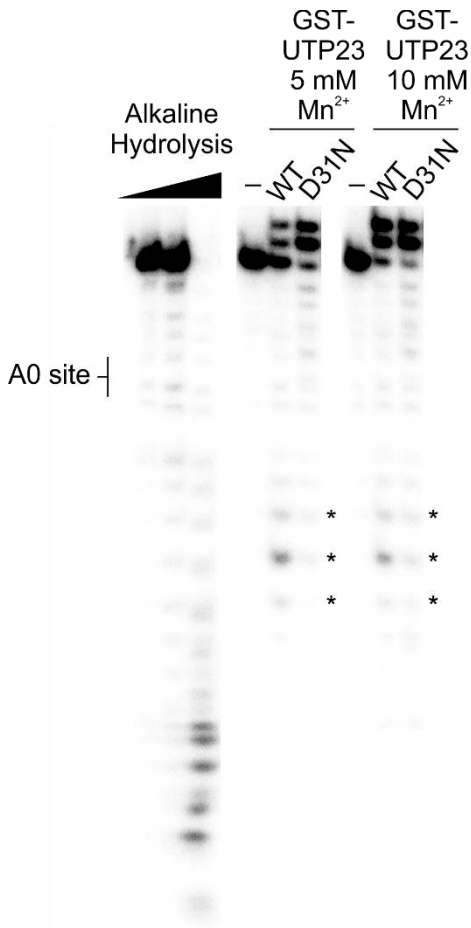


Figure 4.16 Recombinant GST-tagged UTP23 mediates PIN domain-dependent cleavage of a radiolabelled RNA oligonucleotide *in vitro*. An *in vitro* cleavage assay using recombinant human UTP23 wild type (WT) or mutant (D31N) UTP23 and a radiolabelled RNA oligonucleotide containing the human A0 cleavage site. After incubation with recombinant protein RNA was separated on a 12% sequencing gel and visualised using a PhosphorImager. Partial alkaline hydrolysis was performed on the same radiolabelled oligonucleotide to produce a ladder. The position of the A0 cleavage site is indicated and asterisks correspond to UTP23-mediated cleavages.

While no accurate cleavage was observed at the human A0 site for UTP23, the apparent cleavage at other sites on a pre-rRNA substrate with WT, but not PIN domain mutant, GST-tagged UTP23 suggests that UTP23 could have endonuclease activity. This gives support to the potential role of UTP23 as an active endonuclease in ribosome biogenesis. It is possible that accurate cleavage at the correct A0 site requires correct folding of the pre-rRNA substrate, or the presence of other protein factors which either facilitate cleavage or activate the cleavage activity of UTP23. Alternatively, the N-terminal GST-tag on recombinant UTP23 may be disrupting cleavage at the correct site. Further research is needed to confirm the role of human UTP23 as an active endonuclease.

4.3 Discussion

Utp23 is a conserved component of the SSU processome in both yeast and humans (Bleichert *et al.*, 2006; Turner *et al.*, 2009). Correct SSU processome assembly is required for three coupled cleavages at sites A0, A1 and A2 in yeast and the equivalent cleavage events in humans at sites A0, 1 and 2a. PIN domains are found in several known endonucleases, including Nob1/NOB1, which catalyses the final cleavage of the pre-18S RNA at site D/3 (Fatica *et al.*, 2003; Pertschy *et al.*, 2009; Bai *et al.*, 2016) and Utp24/UTP24 which mediates cleavages at sites A1/1 and A2/2a (Tomecki *et al.*, 2015; Wells *et al.*, 2016). The endonuclease activity of PIN domains requires three or four conserved acidic amino acids that coordinate a metal ion, often Mn^{2+}. Yeast Utp23 contains a degenerate PIN domain, containing only two of the four possible residues, meaning it lacks endonuclease activity, and instead expected to play a non-enzymatic, structural role in pre-rRNA processing. Previous research into yeast Utp23 showed that it associates with the snR30 snoRNA, is required for its release from pre-ribosomes and may be important for the base-pairing interactions between snR30 and the ES6 region of 18S rRNA (Hoareau-Aveilla *et al.*, 2012). This suggested a key role for Utp23 in the establishment of interactions around the ES6 region between the pre-rRNA, snR30 and other protein factors functioning in this region. Recent data from our lab provided further evidence for this theory, showing that Utp23 directly interacts with the pre-rRNA and snR30 *in vivo*, and binds to several other ES6- and/or snR30-interacting factors *in vitro* (Wells *et al.*, 2017).

The data presented in this chapter suggests the that role of UTP23 in human cells is largely conserved. The direct binding of yeast Utp23 to the snR30 snoRNA was confirmed *in vitro* (Figure 4.2), as was the fact that this binding is conserved between human UTP23 and the human snR30 homologue, the U17 snoRNA (Figure 4.3). Recombinant UTP23 did not bind to other RNA substrates tested (data not shown), suggesting that the observed binding with U17 is specific. While this interaction is confirmed, the binding of human UTP23 to U17 was considerably weaker than the yeast Utp23-snR30 binding. Yeast Utp23 binds to the internal hairpin (IH) of snR30 (Wells *et al.*, 2017). The human U17 snoRNA contains a much smaller IH, and so the UTP23-U17 interaction is likely to be more dependent on co-factors, which are not present in this *in vitro* assay.

Protein-protein interaction studies also confirmed that GST-tagged human UTP23 interacted with the human homologues of yeast proteins known to interact with the ES6 region of 18S rRNA and with snR30. Protein-protein interactions were confirmed between UTP23 and the endonuclease UTP24, the RNA helicase ROK1/DDX52, the RNA binding protein RRP7 and the box H/ACA snoRNP core protein NHP2 (Figure 4.4). These results suggest that the predicted role of Utp23/UTP23 in coordinating ES6-snR30/U17 interactions in pre-rRNA processing is evolutionarily conserved. No negative control proteins were included in these experiments, which simply aimed to confirm findings observed for yeast Utp23 (Wells *et al.*, 2017). Protein-protein interaction experiments using proteins that are not expected binding partners of human UTP23 would confirm the specificity of the interactions observed here.

The interaction of UTP23 with the endonuclease UTP24 is interesting, as it suggests a potential role for UTP23 in the recruitment of UTP24 for cleavage at site(s) 1 and/or 2a, which may be a conserved function between yeast and humans (Wells *et al.*, 2016; Wells *et al.*, 2017). Alternatively, Utp23/UTP23 may act as a chaperone for the cleavages mediated by Utp24/UTP24 in both systems. In yeast, Utp23 is required for the Rok1-mediated release of snR30 from pre-ribosomes, and this release also requires Rrp5 (Bohnsack *et al.*, 2008; Hoareau-Aveilla *et al.*, 2012; Khoshnevis *et al.*, 2016). Yeast Utp23 interacted with Rok1 *in vitro* (Wells *et al.*, 2017). The *in vitro* interaction between UTP23 and ROK1/DDX52 suggests that the contribution of UTP23 to the release of U17 from pre-ribosomes may be evolutionarily conserved. Rrp7 is present in the yeast 90S pre-ribosome after the release of snR30 and Utp23 (Kornprobst *et al.*, 2016) and binds to the ES6 region of the 18S rRNA (Lin *et al.*, 2013). Rrp7 does not bind snR30 *in vitro* and likely recognises the post-snR30 ES6 structure, suggesting that Utp23 may recruit Rrp7 to this region, facilitating its binding after the release of snR30 and subsequent restructuring of the pre-rRNA (Wells *et al.*, 2017). The interaction between human UTP23 and RRP7 *in vitro* therefore suggests that this proposed function may be conserved in human cells. The reduced size of the IH of the U17 snoRNA compared to snR30 suggests that UTP23 may be more dependent on interaction with box H/ACA core proteins to stabilise the UTP23-U17 interaction in human cells. The observed interaction between UTP23 NHP2 is consistent with this.

The importance of human UTP23 in pre-rRNA processing at three cleavage sites in 18S rRNA maturation has been demonstrated in HeLa cells (Sloan *et al.*, 2014).

Consistently, UTP23 was shown to be a component of the human SSU processome, localising to the nucleolus in HEK293T cells (Turner *et al.*, 2009). The observed accumulation of the 30S pre-rRNA upon depletion of UTP23 in HEK293T cells indicated that UTP23 is essential for cleavages at sites A0, 1 and 2a in these cells (Figure 4.5), suggesting a conserved role for UTP23 in human cells. However, depletion of UTP23 in U2OS and MCF7 cells resulted in variable pre-rRNA processing phenotypes (Figures 4.6 and 4.7), despite efficient RNAi-mediated knockdown of UTP23. In some experiments using these cells, mild accumulation of the 30S intermediate was observed, but other experiments showed no defect or alternative phenotypes. In mouse cells, UTP23 was reported to cause accumulation of the 18SE pre-rRNA (Wang *et al.*, 2014), and this phenotype was also observed in some experiments using U2OS and MCF7 cells, suggesting a role in the later stages of 18S rRNA maturation.

The reasons for the distinct pre-rRNA processing phenotypes observed in different cell types is unclear. More research is required to understand these observations. For example, the localisation and co-precipitation of UTP23 in these cells are required to confirm the presence of UTP23 in the SSU processome. Study of further different cell types is also required to confirm the role of UTP23 in both human and mice cells. It is possible that some cell types have natural defects relating to ribosome biogenesis which alter the pre-rRNA processing phenotype observed upon depletion of UTP23. Indeed, analysis of the sequenced genome of the U2OS cell line shows that these cells contain a missense mutation in the gene encoding the RNA helicase DDX53/ROK1, which has been shown here to interact with UTP23 *in vitro*. The consistent phenotype observed in HEK293T cells upon UTP23 depletion, namely disruption of all three early cleavages, is the same as that reported in HeLa cells (Sloan *et al.*, 2014) and consistent with depletion of Utp23 in yeast (Bleichert *et al.*, 2006). As UTP23, like its yeast counterpart, appears to be a component of the SSU processome (Turner *et al.*, 2012), the pre-rRNA processing defect observed in HEK293T cells is more plausible than the inconsistent phenotypes seen in U2OS and MCF7 cells. Therefore, further experiments used HEK293T cells to explore the function of human UTP23.

To investigate the importance of the conserved PIN endonuclease domain and Zn finger motif of UTP23, an RNAi rescue system was attempted, with HEK293 cells expressing N-terminal FLAG-tagged UTP23. Despite efficient expression and RNAi-

resistance of UTP23, this protein was unable to rescue the pre-rRNA processing defect caused by depletion of endogenous UTP23 (Figure 4.8), suggesting that the N-terminal FLAG tag disrupted the correct folding and/or function of UTP23. N-terminally FLAG-tagged UTP23 was previously shown to localise to the nucleolus (Turner *et al.*, 2012), therefore this tag likely disrupts protein-protein or protein-RNA interactions required for UTP23 function, rather than localisation or nuclear/nucleolar import. To try and prevent this issue, another construct was generated containing a C-terminally HA-tagged UTP23. Expression of this protein was able to rescue this phenotype, suggesting that the position of the tag on the protein does not obstruct its essential function (Figure 4.9).

PIN domain and Zn finger mutant forms of UTP23 were efficiently expressed from stably transfected HEK293 cells upon tetracycline induction, however both mutants showed lower expression levels when compared to WT UTP23 (Figure 4.10). This suggests that these mutants may respond differently to tetracycline compared to the WT protein. Northern analysis of RNA from cells expressing these mutant proteins indicate that the key residues in the PIN domain and Zn finger motif of UTP23 are essential for its role in pre-rRNA processing (Figure 4.12). Depletion of endogenous UTP23 in cells expressing either the D31N PIN domain mutant or the C103A Zn finger motif mutant caused accumulation of the 30S pre-rRNA, consistent with defective processing at sites A0, 1 and 2a. This phenotype suggests that the PIN domain of human UTP23, unlike its yeast counterpart, may have retained some endonuclease activity. It is also consistent with a role for UTP23 in cleavage at site A0 in the human 5’ ETS. The coupling of sites A0, 1 and 2a mean that cleavage at sites 1 and 2a do not occur without prior cleavage at A0, so disruption of A0 cleavage alone would lead to accumulation of the 30S pre-rRNA. The observation of this same pre-rRNA processing defect upon expression of the Zn finger motif mutant suggests that these conserved residues are also important for UTP23 function, presumably due to a role in RNA binding.

While *in vivo* pre-rRNA processing data suggest that the PIN domain and Zn finger are important in 18S rRNA maturation, the fact that the observed processing phenotype is identical to that observed with depletion of UTP23 in control cells suggests that these mutant proteins may be non-functional. While both mutant proteins are expressed at similar levels to WT UTP23 (Figure 4.11), this phenotype could be explained by the disruption of the mutant protein’s localisation, SSU processome

incorporation, or association with the U17 snoRNA. Immunofluorescence experiments indicated that HA-tagged WT UTP23 localised correctly to the nucleolus, as previously observed and consistent with its known role as an SSU processome component (Figure 4.13). Nucleolar localisation was also observed for both PIN domain and Zn finger motif mutant proteins, indicating that mutation of key residues in these proteins does not disrupt the recruitment of UTP23 to the nucleolus.

Immunoprecipitation experiments were used to assess the *in vivo* association of UTP23 mutants with the U17 snoRNA (Figure 4.14). While HA-tagged WT UTP23 efficiently co-precipitated U17, co-precipitation of U17 with the D31N PIN domain mutant was significantly reduced, suggesting that an intact PIN domain is important in some way in the association with UTP23 and the U17 snoRNA. If UTP23 does indeed play a role in cleavage of the pre-rRNA at site A0 via its conserved PIN domain, then disruption of this cleavage activity may affect the subsequent role of UTP23 in binding to and releasing the U17 snoRNA. A0 cleavage may be required for structural rearrangements within the pre-ribosome and for the release of UTP23 from the SSU processome and U17 association. Co-precipitation of the U17 snoRNA by the C103A Zn finger motif mutant was not detectable above background levels, suggesting that the residues making up the Zn finger motif of UTP23 are essential for the UTP23-U17 association *in vivo*.

Finally, the ability of UTP23 PIN domain and Zn finger motif mutants to directly bind the U17 snoRNA *in vitro* was assessed by RNA-protein interaction assays (Figure 4.15). Mutation of a key residue in UTP23's PIN domain did not affect the ability of UTP23 to bind U17, suggesting that the reduced association observed for this mutant *in vivo* is due to a secondary effect, potentially a disruption of UTP23 cleavage activity and subsequent lack of essential structural rearrangements within the pre-ribosome. In contrast, a recombinant UTP23 Zn finger mutant showed reduced binding affinity, suggesting that the key residues forming this motif are important for the direct interaction of UTP23 with the U17 snoRNA. This contrasts with previous data suggesting that, despite mutation of the key Utp23 Zn finger residues being lethal in yeast, the C-terminal tail of yeast Utp23 is important for its binding to snR30 *in vivo* (Lu *et al.*, 2013). However, these results are consistent with recent data that an intact yeast Utp23 Zn finger motif is essential for binding snR30 *in vitro* (Wells *et al.*, 2017). EMSA experiments using different RNA substrates were not performed for UTP23 mutants due to time constraints. These experiments would reveal whether the reduced

association of the Zn finger mutant with U17 is specific for this substrate or if the C103A mutation causes a general reduction in UTP23's RNA binding ability.

The observations that mutation of a key conserved residue within the PIN domain of UTP23 affects pre-rRNA processing but does not alter the cellular localisation, SSU processome incorporation, or *in vitro* binding to the U17 snoRNA, point to the exciting possibility that UTP23 could be an active endonuclease in human ribosome biogenesis. The pre-rRNA processing phenotype observed *in vivo* suggest that UTP23 may be responsible for cleavage at the A0 site. In an attempt to confirm the endonuclease activity of UTP23, *in vitro* RNA nuclease assays were performed using recombinant GST-tagged WT or PIN domain mutant UTP23 and an *in vitro*-transcribed RNA substrate containing the human A0 site (Figure 4.16). Specific cleavage of the pre-rRNA mimic was not observed with either WT or PIN domain mutant UTP23.

These results may suggest that human UTP23, like its yeast counterpart, does not play an enzymatic role in pre-rRNA processing. However, non-specific cleavage was observed at multiple sites around the A0 site with WT GST-UTP23, but not PIN domain mutant (D31N) GST-UTP23, suggesting that the WT protein is capable of RNA cleavage via its PIN domain. It seems likely that its specific endonucleolytic activity has yet to be confirmed, perhaps due to the difference in conditions between the *in vitro* system and that in living cells. Recombinant proteins used in RNA cleavage assays contain an N-terminal GST tag, and RNAi rescue experiments using N-terminally FLAG-tagged UTP23 were unsuccessful, as this tagged protein was unable to functionally replace endogenous UTP23 *in vivo*. The addition of a tag to the N-terminus of UTP23 may therefore disrupt any potential cleavage activity possessed by the recombinant protein. The N-terminal GST-tag did not affect the binding of UTP23 to the U17 snoRNA, or to other protein factors, suggesting that it does not affect the folding of the protein, but it may obstruct the binding of UTP23 to its cleavage substrate, or the cleavage activity itself. Further *in vitro* cleavage assays will use UTP23 fused with alternative affinity tags. Some alternative solutions have been attempted already, although attempts to express recombinant UTP23 containing an N- or C-terminal His tag produced insoluble protein. Attempts to express and purify recombinant UTP23 fused to a C-terminal GST-tag are currently in progress.

An alternative, or additional, explanation for the lack of observable cleavage by UTP23 is that the processing of an RNA substrate at the A0 site may require the presence of other ribosome biogenesis factors. *In vivo*, UTP23 functions as part of the SSU processome, a large complex containing many proteins required for cleavage at multiple sites. It is possible that specific folding of the pre-rRNA is necessary for cleavage at A0 *in vivo* that is not reproduced in the *in vitro* context tested here, and this folding probably requires other SSU processome components. Another SSU processome factor may also act as a co-factor for UTP23-mediated cleavage of pre-rRNA at site A0. Indeed, *in vitro* cleavage of site 3 by the PIN domain endonuclease NOB1 was recently shown to require the co-factor, hCINAP1 (Bai *et al.*, 2016).

Together, the results presented in this chapter suggest an evolutionarily conserved role for human UTP23 in the coordination of interactions with many factors with the pre-rRNA at the ES6 region, as well as a potential additional role in endonucleolytic cleavage at the A0 site.

Chapter Five
The roles of SSU ribosome biogenesis factors in cancer and the p53 signalling pathway

5.1 Introduction

In eukaryotic ribosome biogenesis, three of the four ribosomal (r)RNAs are co-transcribed on a long precursor transcript, which goes through extensive modification and processing steps to release the mature rRNAs. This PhD project has focussed mostly on the pre-rRNA processing steps required to release the mature 18S rRNA from the primary 47S pre-rRNA in human cells, and specifically in identifying endonucleases responsible for cleaving this precursor. The work presented in this chapter focusses on the link between ribosome biogenesis and human health and disease (Pelava *et al.*, 2016).

Defective production of ribosomes is associated with at least 20 genetic diseases, called ribosomopathies (Danilova and Gazda, 2015). Many ribosomopathies are caused by mutations in genes encoding ribosomal proteins. One example is Diamond-Blackfan anaemia (DBA), in which patients have mutations in genes encoding both small subunit (SSU) and large subunit (LSU) ribosomal proteins, and thus exhibit defects in pre-rRNA processing (Ellis, 2014). DBA patients also frequently have physical abnormalities and many of the DBA symptoms have been shown to be dependent on the tumour suppressor p53 (Jaako *et al.*, 2015). Other ribosomopathies are associated with mutations in ribosome biogenesis factors, such as aplasia cutis congenita (ACC), which manifests as localised skin defects. BMS1, a GTPase required for early cleavages essential for maturation of the 18S rRNA is mutated in ACC, and the mutation associated with the disease causes disruption in pre-rRNA processing at these sites (Marneros, 2013). Ribosomopathy patients are often predisposed to cancer, for example DBA patients have a 5-fold higher risk of cancer than the general population (Goudarzi and Lindstrom, 2016).

Ribosome biogenesis is upregulated in cancer and downregulated during cell division and cell differentiation (Gentilella *et al.*, 2015). Many forms of cellular stress, such as DNA damage, also affect ribosome production and ribosome biogenesis is regulated by both oncogenes and tumour suppressors (Orsolic *et al.*, 2016). The link between cancer and defects in ribosome biogenesis is counter-intuitive, as defects in

ribosome biogenesis lead to an increase in p53 levels (Donati *et al.*, 2013; Sloan *et al.*, 2013a). Activation of p53 occurs via the 5S RNP, which is comprised of two LSU ribosomal proteins, RPL5 and RPL11, and the 5S rRNA (Sun *et al.*, 2010; Bursac *et al.*, 2012) (Figure 5.1). Under normal conditions, the 5S RNP, an LSU assembly intermediate, binds to and inhibits MDM2, the inhibitor of p53 (Bursac *et al.*, 2012). Free MDM2 binds to and Ubiquitinates p53, targeting it for proteasomal degradation (Wade *et al.*, 2013).

When ribosome biogenesis is defective, the 5S RNP is accumulated and the 5S RNP-MDM2 interaction is enhanced, leading to stabilisation of p53, and therefore increased p53 levels (Sloan *et al.*, 2013a). Interestingly, this activation is not restricted to LSU defects, as defects in SSU biogenesis also cause p53 activation without noticeably altering LSU production, and p53 induction caused by these defects was dependent on 5S components RPL5 and/or RPL11 (Fumagalli *et al.*, 2012; Pelava, 2016)**.** In addition, both LSU and SSU RPs have been shown to bind to and inhibit MDM2 (Kim *et al.*, 2014; He *et al.*, 2016).

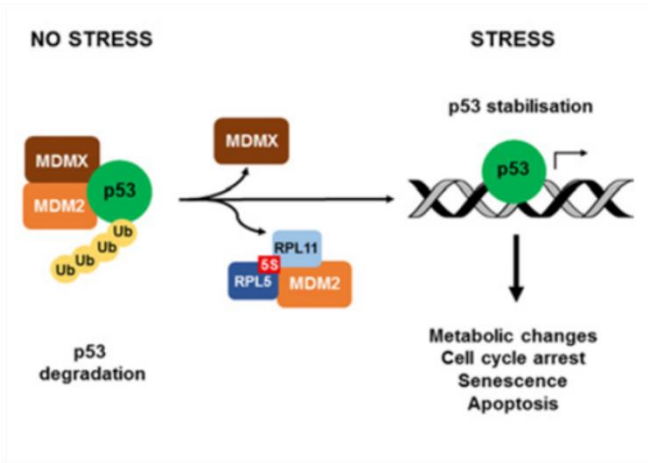


Figure 5.1 Stabilisation of p53 via the 5S RNP. Schematic representation of the regulation of p53 signalling by MDM2 and the 5S RNP in stressed and unstressed cells. Under normal conditions, MDM2 binds to and ubiquitylates p53, targeting it for degradation by the proteasome. Under stress conditions, free 5S RNP binds to and inhibits MDM2, leading to stabilisation of p53 protein levels. From (Pelava, 2016).

A point mutation in the ribosome biogenesis factor UTP23 was recently identified as being associated with colorectal cancer (Timofeeva *et al.*, 2015). UTP23 contains a PIN endonuclease domain and is required for early SSU cleavages at three

sites (A0, 1 and 2a) within the 47S pre-rRNA as part of a large ribonucleoprotein complex called the SSU processome (Turner *et al.*, 2012; Sloan *et al.*, 2014). In yeast, Utp23 associates with, and is required for the release from pre-ribosomes of, the box H/ACA snoRNA snR30, which base-pairs to pre-rRNA and is essential for the three early cleavages required for 18S rRNA release (Hoareau-Aveilla *et al.*, 2012). Less is known about human UTP23 but it has also been shown to associate with the human homolog of snR30, the U17 snoRNA *in vivo* (Hoareau-Aveilla *et al.*, 2012) suggesting that its role in release of U17 from pre-ribosomes may be conserved. Yeast Utp23 contains a degenerate PIN domain, containing only two of the four conserved acidic residues in the putative catalytic site, meaning it lacks endonuclease activity (Bleichert *et al.*, 2006; Hoareau-Aveilla *et al.*, 2012). In contrast, the PIN domain of human UTP23 contains three conserved acidic residues in the putative catalytic centre, and these residues are essential for all three early SSU pre-rRNA cleavages, at sites A0, 1 and 2a, suggesting it may play a role in cleavage at one of these sites (Wells *et al.*, 2017). Another PIN domain protein, Utp24/UTP24 is the likely endonuclease responsible for two of these cleavages, at sites A_1/1 and A_2/2a (Tomecki *et al.*, 2015; Wells *et al.*, 2016), and the RNA cyclase-like protein Rcl1 has been reported to cleave at the yeast A_2 site (Horn *et al.*, 2011). The requirement for an intact UTP23 PIN domain in the three early pre-rRNA cleavages suggests a possible role for UTP23 in cleavage at the A0 site in humans, though its endonuclease activity has not yet been confirmed (see Chapter Four). While many defects in ribosome production can potentially cause an increase in p53 levels via the 5S RNP (Pelava, 2016), it was previously unknown whether depletion and/or mutation of these candidate endonucleases causes p53 induction.

The UTP23 mutation identified in colorectal cancer is a change from a proline at residue 215 to a glutamine (P215Q) (Timofeeva *et al.*, 2015). This residue, in the C-terminal tail of UTP23, is not within the PIN domain of UTP23 or in the RNA-binding Zinc finger motif, so it was unclear whether this substitution would have an effect on the function of UTP23 in pre-rRNA processing or on the stability or expression levels of UTP23. If this mutation affects the stability of the UTP23 protein, it may lead to haploinsufficiency, thereby leading to ribosome biogenesis defects which could contribute to the disease. Alternatively, increased stability of the mutant protein could lead to its overexpression and may exhibit a dominant negative effect.

Many other ribosome biogenesis factors have been found to be mutated in tumours and cancer cell lines (Iorio *et al.*, 2016). Multiple ribosomopathies show defects in processing at the 3' end of the 18S rRNA, within internal transcribed spacer 1 (ITS1), which separates small (18S) and large (28S and 5.8S) rRNA precursors. The processing of ITS1 is used as a clinical diagnostic tool for the ribosomopathy, DBA (Farrar *et al.*, 2014). While much of the ribosome biogenesis production machinery is conserved throughout eukaryotes, the major mechanism of processing of ITS1 is distinct between humans and yeast (Sloan *et al.*, 2013b). In yeast, the major endonuclease cleavage event in ITS1 occurs at site A_2 (equivalent to human site 2a) and is likely mediated by the PIN domain endonuclease Utp24 (Wells *et al.*, 2016). A separate cleavage occurs at site A_3 (equivalent to human site 2) and is mediated by RNase MRP (Lygerou *et al.*, 1996). The 5' end of the 5.8S rRNA is matured by the exonucleolytic activities of Xrn2/Rat1 (Henry *et al.*, 1994) and the 3' end of the 18S rRNA is matured by cleavage at site D by the PIN domain endonuclease Nob1 (Fatica *et al.*, 2003; Pertschy *et al.*, 2009). The RNA-binding protein Rrp5 binds close to, and is required for cleavage of, sites A_2 and A_3 (Venema and Tollervey, 1996; Torchet *et al.*, 1998; Eppens *et al.*, 1999; Eppens *et al.*, 2002; Lebaron *et al.*, 2013).

In humans, the major ITS1 cleavage event occurs at site 2 (equivalent to yeast site 3) by RNase MRP (Goldfarb and Cech, 2017), and the pre-rRNA is trimmed to the 2a site to produce the 18SE pre-rRNA intermediate, which requires the exonucleolytic activity of the exosome-associated exonuclease, RRP6 (Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013). This is surprising, as exosome factors are not required for 18S processing in yeast. In this predominant pre-rRNA processing pathway, cleavage at the 2a site is not essential. As in yeast, the 5' end of the 5.8S rRNA is matured by the exonuclease XRN2 (Preti *et al.*, 2013; Sloan *et al.*, 2013b), and the 3' end of the 18S rRNA is matured via endonucleolytic cleavage by NOB1 (Bai *et al.*, 2016). Human cells also appear to utilise a "minor" pre-rRNA processing pathway, which is similar to the predominant yeast pathway. In this pathway, ITS1 is cleaved at both site 2a (likely by UTP24) and site 2 (Sloan *et al.*, 2013b). A similar situation is observed in yeast upon impaired ribosome biogenesis, with low levels of a precursor cleaved at site A_3 prior to A_2 cleavage detected in wild type cells, and significantly accumulated upon impairment of ribosome biogenesis (Venema and Tollervey, 1999). It is unclear whether this precursor is further processed to generate mature 18S rRNA, representing a "minor" pre-rRNA processing pathway, or if it represents an aberrant precursor that is

degraded. It is unclear why human cells have two distinct pathways for the removal of ITS1, or why the “minor” pathway resembles the predominant yeast pathway, but it is interesting that ribosome biogenesis factors essential for this higher eukaryotic-specific process are often mutated in ribosomopathies.

A recent sequencing study of 11,289 tumours and 1,001 cancer cell lines identified frequent mutations (from here on referred to as “cancer mutations”) in both ribosomal proteins and ribosome biogenesis factors (Iorio *et al.*, 2016). Many mutations were found in factors essential for processing in ITS1 and in particular, in maturation of the 3’ end of 18S rRNA. Two of these factors are the exonucleases RRP6 and XRN2, which play key roles in ITS1 in the maturation of the 3’ end of 18S rRNA and the 5’ end of the 5.8S rRNA respectively. 78 cancer mutations were identified in RRP6, and many of these were in the vertebrate-specific C-terminal domain B (Figure 5.2B). While yeast Rrp6 is important in many aspects of nuclear RNA processing/turnover, human RRP6 is predominantly localised to the nucleolus, suggesting its major role may be in pre-rRNA processing (Tomecki and Dziembowski, 2010). The exosome complex and Rrp6 specifically are required for the exonucleolytic maturation of the 3’ end of the 5.8S rRNA in yeast (Briggs *et al.*, 1998; Thomson and Tollervey, 2010) but human RRP6 seems to play a vertebrate-specific role in the processing of ITS1 in maturation of the 3’ end of the 5.8S rRNA (Sloan *et al.*, 2013b). It was previously unknown whether depletion of RRP6 or disruption of its exonuclease activity results in the activation of p53. In yeast, Xrn2/Rat1 is required for maturation of the 5’ end of the 5.8S rRNA (Henry *et al.*, 1994; Oeffinger *et al.*, 2009) and for co-transcriptional cleavage at site A_2 (Axt *et al.*, 2014). In mammals, XRN2 is also required for 5.8S rRNA 5’ end maturation, plays an additional role in the mammalian-specific co-transcriptional cleavage at site A’ in the 5’ ETS, and its depletion stimulates pre-rRNA processing through the minor pathway, increasing the dependency on cleavage at site 2a by UTP24 (Preti *et al.*, 2013; Sloan *et al.*, 2013b; Wang *et al.*, 2014). Whether depletion of XRN2 activates p53 was previously unknown, but 88 XRN2 cancer mutations were identified, many in the vertebrate-specific C-terminal domain (Iorio *et al.*, 2016) (Figure 5.2A).

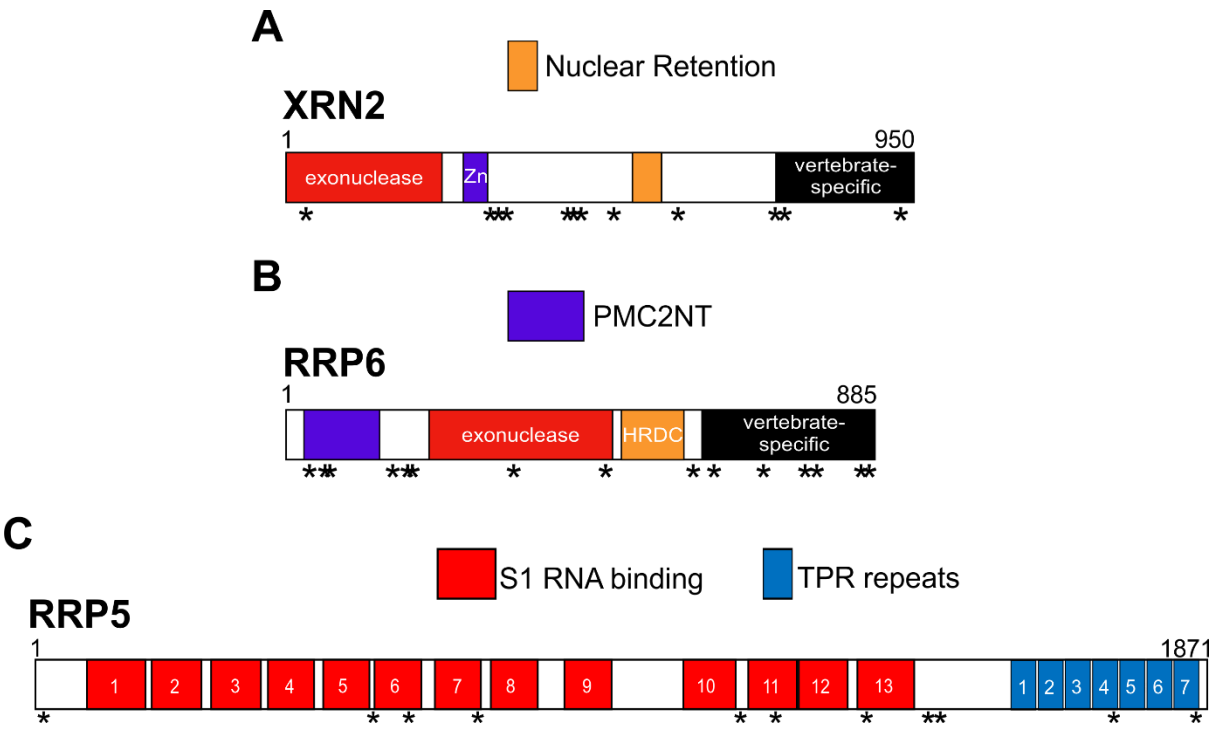


Figure 5.2 The domains of human ribosome biogenesis factors XRN2, RRP6 and RRP5. Cartoon representation of the domain architecture of **(A)** XRN2, **(B)** RRP6 and **(C)** RRP5. Positions of cancer-associated mutations (Iorio *et al.*, 2016) are marked with asterisks.

Another ribosome biogenesis factor important for the processing of ITS1 is the RNA-binding protein RRP5 (Figure 5.2C). In yeast, different parts of the large (193 kDa) multidomain Rrp5 protein are required for distinct processing events. The N-terminal domain (NTD) of Rrp5, containing 9 of its 12 S1 RNA-binding domains, is required for cleavage at site A_3, while the C-terminal domain (CTD), containing the remaining three S1 domains and seven TPR (Tetratricopeptide repeat) protein-protein interaction domains, is important for the three early cleavages at sites A_0, A_1 and A_2 (Venema and Tollervey, 1996; Eppens *et al.*, 1999). UV crosslinking and analysis of cDNA (CRAC) showed that these separate Rrp5 protein halves have distinct binding sites on the pre-rRNA, with the NTD binding adjacent to the A_3 site (human site 2), and the CTD binding adjacent to site A_2 (human site 2a) (Lebaron *et al.*, 2013). This suggests that Rrp5 may play a role in the selection of the cleavage site in ITS1. The last three S1 domains (S1-10 to S1-12) in the C-terminal half of Rrp5 are essential for cleavage at site A_2 (Vos *et al.*, 2004b). Human RRP5 is a component of the SSU processome (Turner *et al.*, 2012), and as such is essential for 18S rRNA production (Sweet *et al.*, 2008). The role of human RRP5 in selection of ITS1 cleavage site usage

is unknown, but it appears to have a conserved role in the processing of both ITS1 cleavages (at sites 2a and 2) in HeLa cells (Sloan *et al.*, 2013b). It is unknown whether these distinct roles are physically separated in relation to the different domains within RRP5, as was shown in yeast, or whether depletion of RRP5 causes induction of p53 via the 5S RNP signalling pathway. 168 cancer mutations were found dispersed throughout the coding region of RRP5 (Iorio *et al.*, 2016).

It is interesting that factors involved in processing of ITS1 are often found in tumours and cancer cell lines (Iorio *et al.*, 2016). While pre-rRNA processing pathways are generally conserved from yeast to humans, processing of ITS1 represents a major distinction between ribosome biogenesis in lower and higher eukaryotes. Frequent mutation of ITS1 processing factors in cancer cells suggest that removal of this region may differ in these cells. It is possible that disease-associated defects in factors involved in ITS1 processing shift cells to the “minor” pre-rRNA processing pathway, causing a reliance on cleavage at 2a. If cancer cells are dependent on processing through this pathway, it could be used as a target to specifically affect cancer cells.

The work presented in this chapter aims to assess whether a UTP23 mutation (P215Q) found in colorectal cancer causes defects in pre-rRNA processing in HEK293T cells, and if its expression leads to p53 induction in U2OS cells. In addition, the effect of depleting candidate endonucleases UTP23, UTP24 and RCL1 on p53 induction in U2OS and MCF7 cells was examined. Finally, this work also aims to assess factors involved in processing of ITS1, the exonucleases RRP6 and XRN2, and the RNA binding protein RRP5, and their importance for p53 induction in U2OS and MCF7 cells and pre-rRNA processing in HEK293T and MCF7 cells.

5.2 Results

5.2.1 Establishment of a stable HEK293T cell line expressing a UTP23 mutant found in a colorectal cancer patient

To establish whether the UTP23 P215Q mutation, observed in a colorectal cancer patient, affects UTP23 stability and UTP23’s role in ribosome biogenesis, a stable HEK293T cell line was generated to analyse this mutant protein *in vivo*. Previous studies have demonstrated the importance of an intact PIN domain and Zinc finger motif in UTP23 (Wells *et al.*, 2017), but the effect of mutations in the C-terminal tail of

UTP23, where the P215 residue is located, has not been investigated. An RNAi rescue system was therefore established to assess the impact of the P215Q mutation on pre-rRNA processing in human cells.

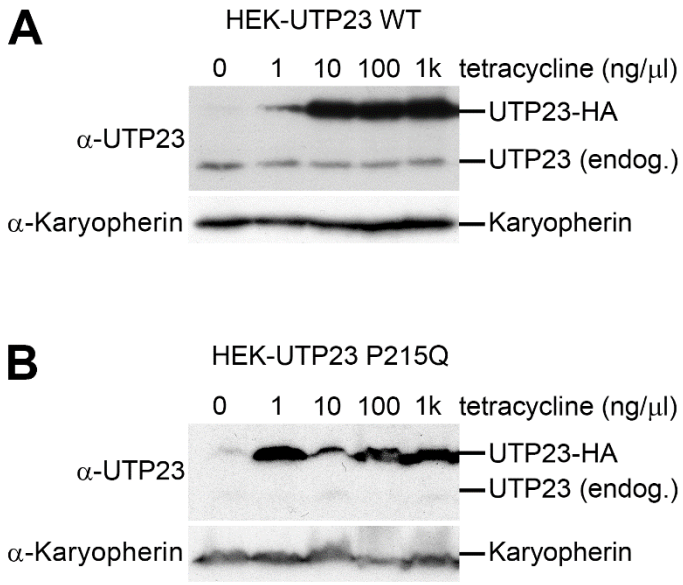


Figure 5.3 Expression of RNAi-resistant UTP23 WT and P215Q mutant proteins in stable HEK293T cell lines. HEK293T cells stably transfected with HA-tagged wild type (WT; panel A) or mutant (P215Q; panel B) UTP23 were incubated with a range of concentrations of tetracycline (0-1,000 ng/μl). Protein was separated by SDS-PAGE and analysed by western blotting using an anti-UTP23 antibody to detect both endogenous and HA-tagged UTP23. An anti-Karyopherin antibody was used as a loading control.

Site-directed mutagenesis of the RNAi-resistant pcDNA5-UTP23-WT-HA construct (see Chapter Four) was performed to substitute the P215 residue for a glutamine (P215Q). The pcDNA5 plasmid allows for regulation of protein expression levels via a tetracycline-inducible promoter. The pcDNA5-UTP23-P215Q-HA construct was then stably transfected into HEK293T Flp-In cells using the Flp-In recombination system (Life Technologies). After selection of stably transfected cells with blasticidin S and hygromycin B, expression of UTP23-P215Q-HA was titrated with a range of tetracycline concentrations (0-1,000 ng/μl) (Figure 5.3). Western blotting with an anti-UTP23 antibody showed that UTP23-P215Q was efficiently expressed upon tetracycline induction, indicating that this mutation does not make the protein unstable, and therefore any potential defects caused by this mutant are unlikely to be due to haploinsufficiency. As for wild type (WT) UTP23, induction of UTP23-P215Q with the

lowest concentration of tetracycline caused expression at levels higher than endogenous UTP23. The P215Q was expressed at higher levels than WT UTP23, suggesting that the mutation may increase the stability of the protein.

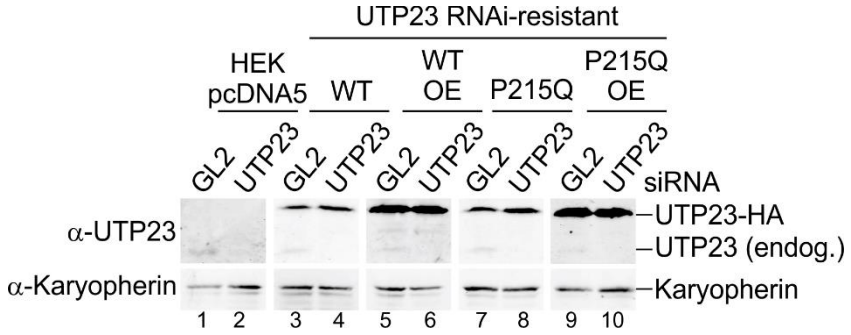


Figure 5.4 Protein analysis of HEK293T UTP23 RNAi rescue. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5; lanes 1 and 2), or wild type (WT/WT OE; lanes 3 and 4/ lanes 5 and 6) or mutant (P215Q/P215Q OE; lanes 7 and 8/ lanes 9 and 10) HA-tagged RNAi-resistant UTP23. Low (1 ng/µl; lanes 3-4 and 7-8) and high (10,000 ng/µl; lanes 5-6 and 9-10) concentrations of tetracycline were used to induce expression of HA-tagged UTP23 at low or high levels. Following transfection of a control siRNA (GL2) or RNAi-mediated depletion of endogenous UTP23 (UTP23), protein was extracted from cells and separated by SDS-PAGE. Protein was analysed by western blotting using an anti-UTP23 antibody and an anti-Karyopherin antibody was used as a loading control.

To confirm that HA-tagged UTP23-P215Q was resistant to RNAi-mediated depletion, whole extract from cells treated with 0.5 ng/µl tetracycline and a siRNA targeting the mRNA of endogenous UTP23 or a control siRNA (GL2) was separated on a polyacrylamide gel. To test the hypothesis that increased stability, and therefore overexpression, of the P215Q mutant protein, may affect pre-rRNA processing, samples from cells expressing UTP23-WT or UTP23-P215Q treated with 10,000 ng/µl (OE) were also analysed. A cell line expressing only a FLAG-tag (pcDNA5) was used as a control. Western blotting was performed using an anti-UTP23 antibody to detect both endogenous and HA-tagged UTP23 (Figure 5.4). Karyopherin was used as a loading control. Endogenous UTP23 was depleted efficiently by the UTP23-targeting siRNA in all cells. WT or P215Q mutant HA-tagged UTP23 was not depleted by the UTP23 siRNA, confirming that this mutant is resistant to RNAi-mediated knockdown.

5.2.2 The UTP23 P215Q mutant affects pre-rRNA processing in HEK293T cells when strongly overexpressed

To determine the effect of the UTP23 P215Q mutant on pre-rRNA processing *in vivo*, RNA was extracted from cells (as in Figure 5.4) expressing WT or P215Q mutant UTP23 after tetracycline induction and RNAi-mediated knockdown of endogenous UTP23. Northern blotting was performed using radiolabelled oligonucleotide probes hybridising to the middle of ITS1 ('ITS1') or to a region between the 3' end of 18S and the 2a cleavage site ('18SE') (Figure 5.5A). The positions of probes on the 47S pre-rRNA are indicated in Figure 5.5B. As previously observed, depletion of UTP23 in control cells (pcDNA5) (Figure 5.5A, lanes 1 and 2) caused accumulation of 30S pre-rRNA and reduction of 18SE pre-rRNA levels compared to cells treated with the control siRNA. Expression of RNAi-resistant WT UTP23-HA after treatment with low (0.5 ng/µl) tetracycline (lanes 3 and 4) rescued this phenotype. Expression of the UTP23-HA P215Q mutant at the same levels also rescued the pre-rRNA processing phenotype (lanes 7 and 8), with no observable accumulation of the 30S pre-rRNA. When WT UTP23-HA expression was induced with high (10,000 ng/µl) tetracycline (OE WT; lanes 5 and 6), levels of pre-rRNA intermediates were unchanged compared to low tetracycline (WT; lanes 3 and 4). In contrast, strong overexpression of the UTP23 P215Q mutant (OE P215Q; lanes 9 and 10), appeared to cause a pre-rRNA processing defect, with observable accumulation of the 30S pre-rRNA, similar to the effect of depleting UTP23 in control cells. Interestingly, 18SE pre-rRNA levels were not strongly reduced in these cells, compared to UTP23 depletion in control cells. This suggests that, rather than abolishing cleavages at sites A0, 1 and 2a, these processing events may occur less efficiently, with the accumulated 30S pre-rRNA eventually processed to 18SE.

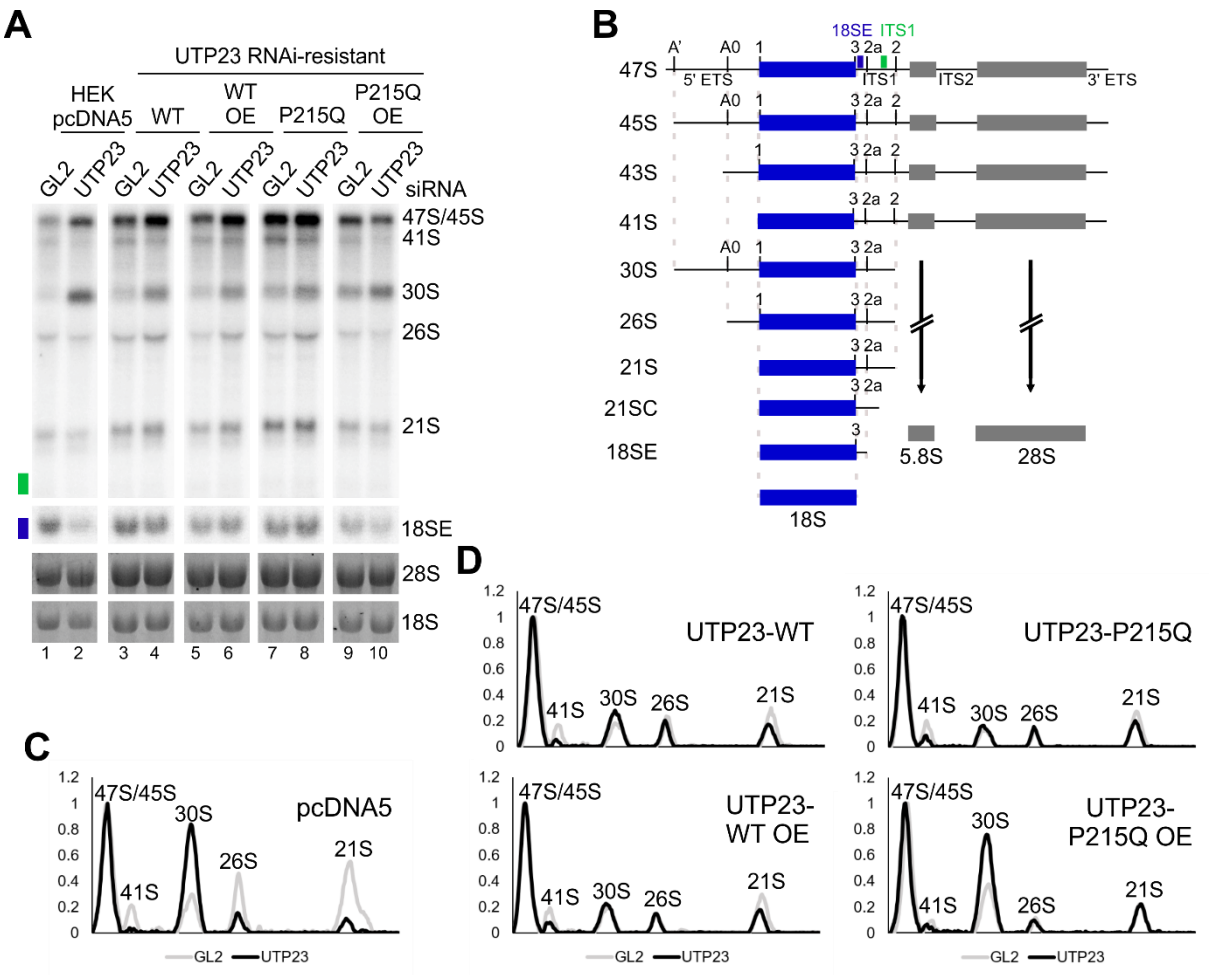


Figure 5.5 Overexpression of the UTP23 P215Q mutant protein disrupts pre-rRNA processing in HEK293T cells. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5; panel A lanes 1 and 2), or wild type (WT/WT OE; panel A lanes 3 and 4/ lanes 5 and 6) or mutant (P215Q/P215Q OE; panel A lanes 7 and 8/ lanes 9 and 10) HA-tagged RNAi-resistant UTP23. **(A)** RNA was extracted following treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous UTP23 (UTP23). RNA was separated by agarose gel electrophoresis and transferred to a nylon membrane and analysed by northern blotting using radiolabelled probes hybridising to the 5' end of ITS1 (18SE; blue rectangle) or upstream of cleavage site 2 (ITS1; green rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. **(B)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of probes used in northern blotting (18SE; blue and ITS1; green) are marked on the 47S pre-rRNA. **(C and D)** Levels of pre-rRNAs detected by the ITS1 probe from panel A were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted. Pre-rRNA intermediates represented by each peak are labelled.

5.2.3 The UTP23 P215Q mutant localises correctly to the nucleolus when overexpressed

In an effort to understand the reason for the pre-rRNA processing phenotype observed upon over-expression of the UTP23 P215Q mutant, immunofluorescence was performed to investigate the localisation of this mutant. As an essential component of the SSU processome, WT UTP23 is localised throughout the nucleolus under normal conditions (Turner *et al.*, 2012; Wells *et al.*, 2017). Expression of a FLAG-tag alone (pcDNA5) or HA-tagged WT or P215Q mutant, with low or high tetracycline concentrations, was induced in HEK293T stable cells, which were grown on coverslips. Immunofluorescence was performed using an anti-HA antibody to detect the localisation of HA-tagged proteins (Figure 5.6). The box C/D snoRNP component Fibrillarin was used as a marker for the dense fibrillar component (DFC) of the nucleolus and nuclei were stained with DAPI. As previously observed, WT UTP23-HA localised throughout the nucleolus in these experiments. The UTP23 P215Q mutant also localised, after induction with either low or high tetracycline, throughout the nucleolus, indicating that this point mutation does not affect the localisation of UTP23.

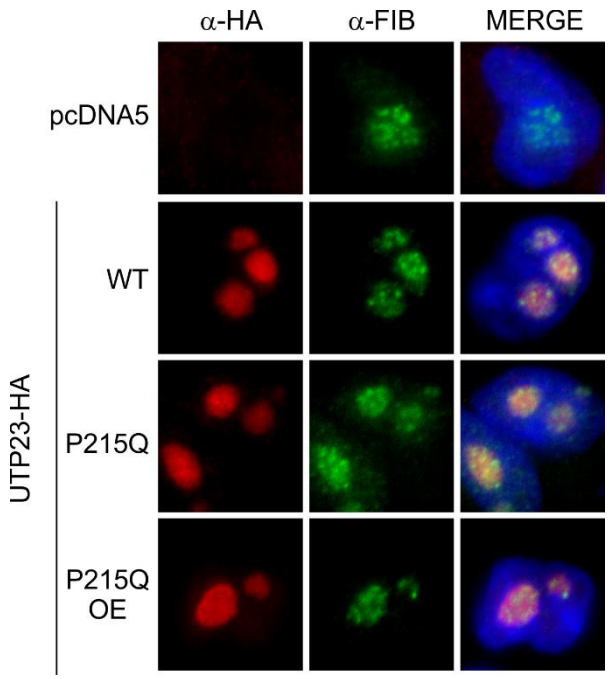


Figure 5.6 UTP23 P215Q mutant localises to the nucleolus in HEK293T cells. Immunofluorescence was performed on HEK293T cells stably expressing a FLAG-tag alone (pcDNA5), HA-tagged UTP23 wild type (WT) or mutant UTP23 expressed at low (P215Q) or high (P215Q OE) levels using an anti-HA antibody (left column). An anti-Fibrillarin antibody (FIB) was used as a nucleolar marker (middle column). Anti-HA signal is shown in red, anti-Fibrillarin signal is shown in green and DAPI staining of the nucleus is shown in blue.

5.2.4 The UTP23 P215Q mutant may associate with the U17 snoRNA more efficiently than WT UTP23 in vivo

While the UTP23 P215Q mutant correctly localises to the nucleolus, the pre-rRNA processing phenotype observed upon over-expression of this mutant suggests that a key role of UTP23 in the SSU processome is impaired by this mutation. UTP23 plays a conserved role in coordinating interactions around the ES6 region of the 18S rRNA sequence in the pre-ribosome, via its association with the U17 snoRNA (snR30 in yeast). While the direct interaction between UTP23 and U17 appears to involve the nucleic acid binding activity of the conserved Zinc finger motif of UTP23 (Wells *et al.*, 2017), it is possible that the P215Q point mutation also affects this interaction or the interaction with other ribosome biogenesis factors in the SSU processome.

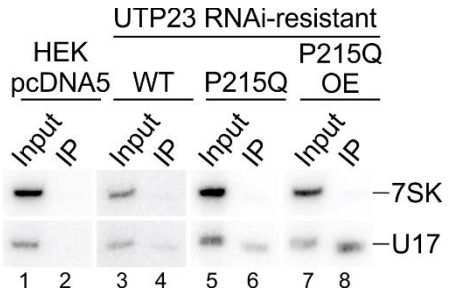


Figure 5.7 The UTP23 P215Q mutant may associate more strongly with the U17 snoRNA than WT UTP23 *in vivo*. Immunoprecipitation was performed on lysates from cells expressing a FLAG-tag alone (pcDNA5), HA-tagged wild type UTP23 (WT) or mutant UTP23 expressed at low (P215Q) or high (P215Q OE) levels using an anti-HA antibody. Co-precipitated RNA was extracted and analysed by northern blotting using a probe hybridising to the U17 snoRNA. A probe hybridising to the small nuclear RNA 7SK was used as a loading control. 1.25% of the input material was loaded. RNA was visualised using a PhosphorImager.

To assess the association of the P215Q mutant with U17 in pre-ribosomes, immunoprecipitation experiments were performed using anti-HA antibody-coupled agarose beads with whole-cell extracts expressing only a FLAG-tag (pcDNA5), or HA-tagged WT or P215Q mutant (treated with low or high tetracycline) UTP23 (Figure 5.7). RNA was extracted from these samples and northern blotting with a probe targeting the U17 snoRNA was performed. A probe hybridising to the small nuclear RNA, 7SK, was used as a loading control. No U17 snoRNA co-precipitation was observed in control samples (lanes 1 and 2). Both WT (lanes 3 and 4) and P215Q mutant (lanes 5 and 6) UTP23-HA co-precipitated U17 efficiently. This suggests that substitution of the P215 residue does not disrupt the ability of UTP23 to bind the U17 snoRNA. The P215Q mutant (lanes 5 and 6) appeared to be more strongly associated with U17 than WT UTP23-HA, and this increased association was more severe upon overexpression of the P215Q mutant (lanes 7 and 8). This preliminary result suggests that this point mutation may increase the affinity of UTP23 for U17, or in some way disrupt the release of U17 from the pre-ribosome, potentially explaining the pre-rRNA processing phenotype observed upon overexpression of the P215Q mutant. Further experiments are needed to assess the effects of this mutation, for example gradient analysis to investigate the association of UTP23-P215Q and U17 with pre-ribosomal particles *in vivo*.

5.2.5 The UTP23 P215Q mutant binds the U17 snoRNA with similar efficiency as the WT protein in vitro

The apparent increase in association of the UTP23 P215Q mutant with the U17 snoRNA suggests that the binding efficiency of the observed interaction of UTP23 to U17 (see Chapter Four) may be increased by this mutation. To test this possibility, an *in vitro* RNA binding electromobility shift assay (EMSA) was performed using recombinant GST-tagged UTP23 WT or P215Q mutant, and a radiolabelled, *in vitro*-transcribed U17 RNA substrate (Figure 5.8). After incubation of GST-UTP23 or free GST with the RNA substrate, free RNA (U17) and RNA-protein complexes (RNP) were separated on a 4% native polyacrylamide gel. For both WT (Figure 5.8A) and P215Q mutant (Figure 5.8B) UTP23, the intensity of the RNP band increased with increasing (0-500 nM) concentrations of GST-UTP23. The shift from free RNA to RNP was similar for WT and the P215Q mutant, indicating that the strength of the interaction between UTP23 and U17 is not affected by the P215Q mutation. This suggests that the increased U17 association of the P215Q mutant observed *in vivo* (Figure 5.7) is not due to any increase in the direct interaction between UTP23 and U17. Therefore, it is possible that this mutation instead increases the association of UTP23 with pre-ribosomal complexes containing U17, such as the SSU processome. Gradient analysis upon expression of the P215Q mutant, with both low and high tetracycline treatment, is required to ascertain the association of the U17 snoRNA and the mutant protein with different pre-ribosomal complexes.

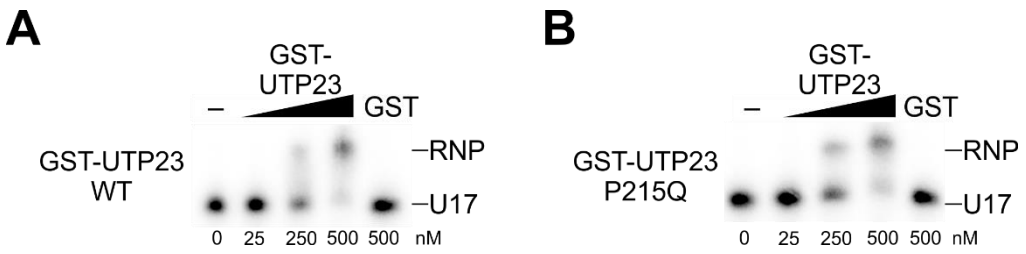


Figure 5.8 The UTP23 P215Q mutation does not affect *in vitro* binding of UTP23 to the U17 snoRNA. Electromobility shift assay (EMSA) to analyse binding of recombinant **(A)** wild type (WT) and **(B)** P215Q mutant (P215Q) forms of GST-tagged UTP23 protein (0-500 nM) or GST (500 nM) to radiolabelled *in vitro* transcribed U17 RNA substrate. RNP complexes and free U17 RNA were separated on a 4% native polyacrylamide gel and RNA was visualised using a PhosphorImager.

5.2.6 Depletion of UTP23, but not UTP24 or RCL1, causes mild induction of p53 in U2OS and MCF7 cells

Defects in ribosome biogenesis are linked to p53 signalling via the 5S RNP-MDM2 pathway, and many ribosomopathies have links to cancer. Upon dysfunction of ribosome biogenesis, free 5S RNP binds to MDM2, preventing MDM2 from inhibiting the tumour suppressor p53 (Pelava *et al.*, 2016). Disruption of ribosome biogenesis therefore causes an increase in p53, which is counterintuitive as many ribosomopathies are associated with increased cancer risk. The identification of a UTP23 point mutation (P215Q) in a colorectal cancer patient prompted me to investigate the link between UTP23, and other putative endonucleases in ribosome biogenesis, and p53 induction. HEK293T cells do not have an active p53 gene, so cannot be used to determine p53 induction. Two other human cell lines, osteocarcinoma-derived cells (U2OS) and breast cancer derived cells (MCF7) do have an active p53 gene and are therefore suitable cell lines to study p53 induction upon depletion of ribosome biogenesis factors. RNAi-mediated knockdown of putative endonucleases UTP23, UTP24 and RCL1 was performed in U2OS and MCF7 cells for 72 hours. Depletion of the LSU RP RPL7 causes significant, strong induction of p53 (Pelava, 2016), so RNAi-mediated knockdown of RPL7 was also performed in U2OS and MCF7 cells as a control for “strong” (up to 4-fold) p53 induction. Cells were harvested, and cell extract was separated by SDS-PAGE. Western blotting was performed using an anti-p53 antibody to detect p53 levels (Figure 5.9). Karyopherin was used as a loading control. Successful depletion of the target protein was also confirmed by western blotting to confirm that RNAi-mediated knockdowns work efficiently in U2OS and MCF7 cells (data not shown; see Chapter Three; see Chapter Four).

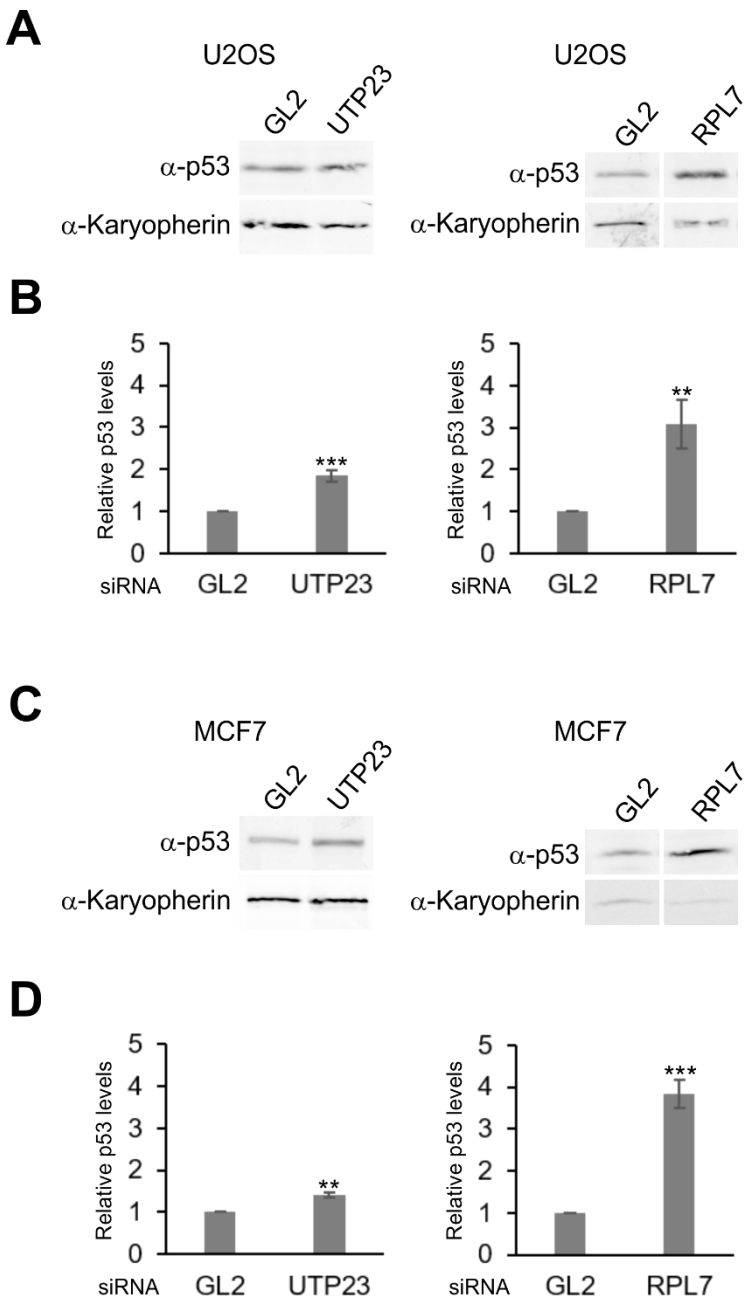


Figure 5.9 Depletion of UTP23 causes p53 induction in U2OS and MCF7 cells. (A) Analysis of protein levels from U2OS cells after RNAi-mediated depletion of UTP23 (left panel) or RPL7 (right panel) by western blotting using an anti-p53 antibody. RPL7 depletion was used as a positive control for strong (~4-fold) p53 induction and an anti-Karyopherin antibody was used as a loading control. Multiple (>3) repeat experiments were performed, and a representative western blot is shown. **(B)** Quantification of p53 protein levels from multiple (>3) western blots following depletion of UTP23 (left panel) or RPL7 (right panel) in U2OS cells. Average intensity of bands detected by the anti-p53 antibody, normalised to levels of Karyopherin are plotted compared to control (GL2) cells. Error bars represent standard error and statistical analysis was performed using an unpaired t-test. **p value <0.01, ***p value <0.0001. **(C)** Analysis of protein levels from MCF7 cells after RNAi-mediated depletion of UTP23 (left panel) or RPL7, as in panel A. **(D)** Quantification of p53 protein levels from multiple (>3) western blots following depletion of UTP23 (left panel) or RPL7 (right panel) in MCF7 cells, as in panel B. **p value <0.01, ***p value <0.001.

A representative western blot for UTP23 knockdown in U2OS cells is shown in Figure 5.9A (left panel) and the quantification of five repeat experiments is shown in Figure 5.9B (left panel). Depletion of UTP23 in U2OS cells caused a mild (~1.75-fold) but significant increase in p53 levels compared to the control (GL2) siRNA. A representative western blot showing induction of p53 upon RPL7 depletion in U2OS cells is shown in Figure 5.9A (right panel) and quantification of p53 levels in six repeats experiments (Figure 5.9B, right panel) shows a significant average p53 induction of around 3-fold**.** A representative western blot for UTP23 depletion in MCF7 cells is shown in Figure 5.9C (left panel) and quantification of p53 levels from three repeat experiments is shown in Figure 5.9D (left panel). Depletion of UTP23 also caused a mild (~1.5-fold), but significant increase in p53 in MCF7 cells, indicating that this effect is conserved between cell types. Depletion of RPL7 in MCF7 cells (Figure 5.9C, right panel) showed an average p53 induction of around 3.8-fold over ten repeat experiments (Figure 5.9D, right panel). RNAi-mediated depletions of candidate endonucleases UTP24 and RCL1 were also performed in both U2OS and MCF7 cells and western blotting was performed to detect changes in p53 levels (Figure 5.10). Depletion of UTP24 or RCL1 in U2OS (Figure 5.10A and B) or MCF7 cells (Figure 5.10C and D) did not cause a significant increase in p53 levels. These results indicate that depletion of UTP24 or RCL1 does not cause p53 induction in either U2OS or MCF7 cells, while depletion of UTP23 causes a mild induction of p53 compared to depletion of RPL7.

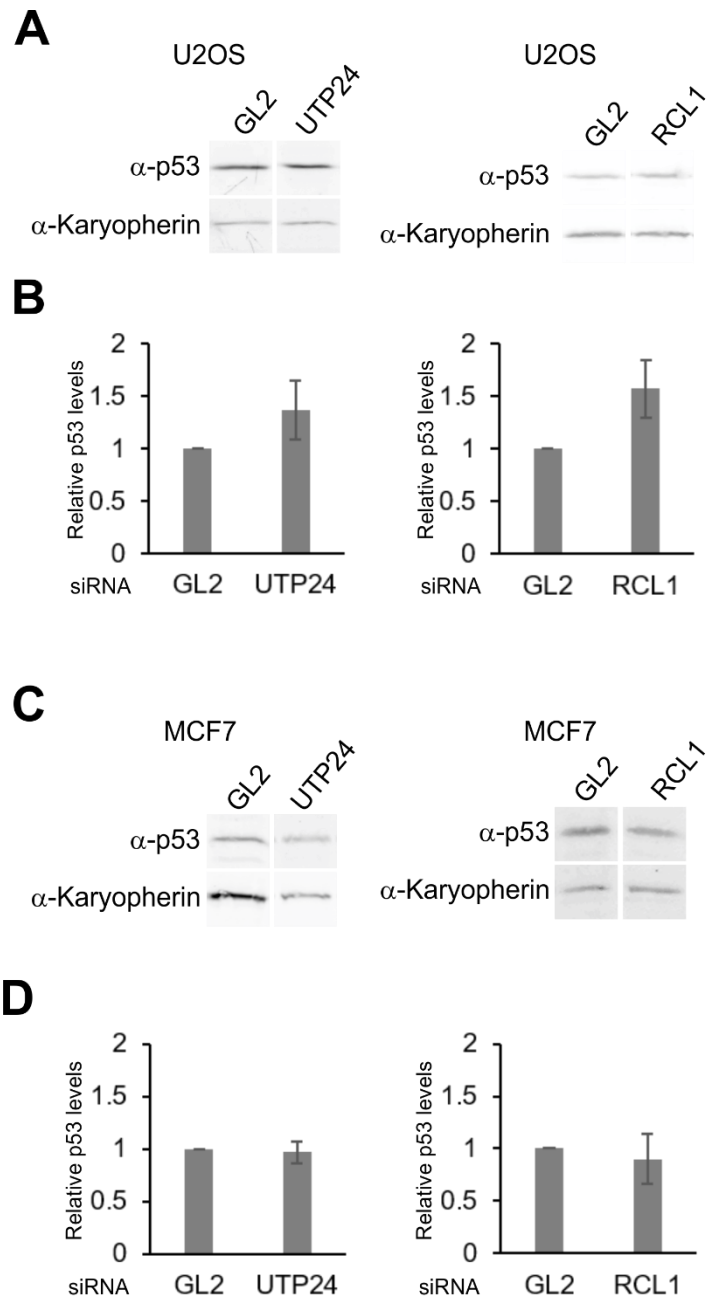


Figure 5.10 Depletion of UTP24 or RCL1 does not cause p53 induction in U2OS or MCF7 cells. (A) Analysis of protein levels from U2OS cells after RNAi-mediated depletion of UTP24 (left panel) or RCL1 (right panel) by western blotting using an anti-p53 antibody. An anti-Karyopherin antibody was used as a loading control. Multiple (>3) repeat experiments were performed, and a representative western blot is shown. **(B)** Quantification of p53 protein levels from multiple (>3) western blots following depletion of UTP24 (left panel) or RCL1 (right panel) in U2OS cells. Average intensity of bands detected by the anti-p53 antibody, normalised to levels of Karyopherin are plotted compared to control (GL2) cells. Error bars represent standard error and statistical analysis was performed using an unpaired t-test. **(C)** Analysis of protein levels from MCF7 cells after RNAi-mediated depletion of UTP24 (left panel) or RCL1 (right panel), as in panel A. **(D)** Quantification of p53 protein levels from multiple (>3) western blots following depletion of UTP24 (left panel) or RCL1 (right panel) in MCF7 cells, as in panel B.

5.2.7 Expression of the UTP23 P215Q colorectal cancer mutant does not causes p53 induction in U2OS cells

The observation that knockdown of UTP23 produced a mild but significant induction of p53 suggests that the pre-rRNA processing defect observed upon overexpression of the P215Q mutation found in colorectal cancer patients may also lead to increased levels of p53. To investigate whether the induction of p53 observed when UTP23 is depleted, is also seen upon expression of the colorectal cancer mutant (P215Q) UTP23, stable UTP23 cell lines were generated using U2OS Flp-In cells. As for the HEK293T stable cell line above, the RNAi-resistant pcDNA5-UTP23-HA constructs were stably transfected into U2OS Flp-In cells using the Flp-In recombination system (Life Technologies). Expression of HA-tagged WT UTP23 and P215Q mutant UTP23 was induced with a range of tetracycline concentrations (0-1,000 ng/µl) to analyse expression levels of UTP23-HA compared to endogenous UTP23. Western blotting was used with an anti-UTP23 antibody to detect both HA-tagged UTP23 and endogenous UTP23, and karyopherin was used as a loading control (Figure 5.11). This showed that WT UTP23 (Figure 5.11A) and P215Q mutant (Figure 5.11B) UTP23 responded to tetracycline in a similar manner. In addition, unlike HEK293T stable cell lines, addition of low concentrations of tetracycline (0.5-10) did not cause expression of UTP23-HA at levels considerably higher than endogenous UTP23.

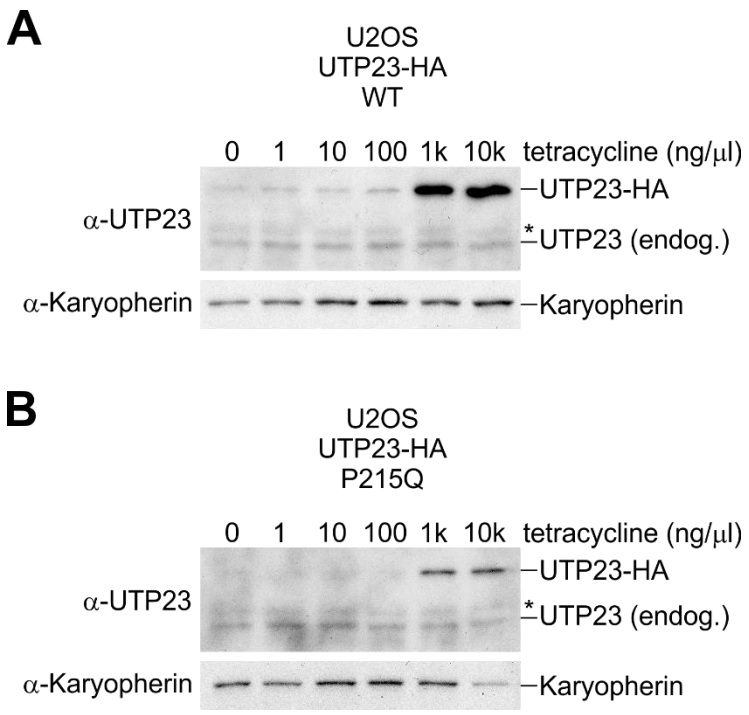


Figure 5.11 Expression of RNAi-resistant UTP23 WT and P215Q mutant proteins in stable U2OS cell lines. U2OS cells stably transfected with HA-tagged wild type (WT; panel A) or mutant (P215Q; panel B) UTP23 were incubated with a range of tetracycline concentrations (0-10,000 ng/μl). Protein was separated by SDS-PAGE and analysed by western blotting using an anti-UTP23 antibody to detect both endogenous and HA-tagged UTP23. An anti-Karyopherin antibody was used as a loading control.

Expression of WT UTP23 or colorectal cancer mutant UTP23 (P215Q) was induced with tetracycline and endogenous UTP23 was depleted using a siRNA targeting UTP23 mRNA. Whole-cell extract was separated on an SDS-PAGE gel and western blotting was performed using an anti-UTP23 antibody to confirm endogenous UTP23 depletion and UTP23-HA expression and an anti-p53 antibody to detect changes in p53 levels (Figure 5.12A). Karyopherin was used as a loading control. The intensity of bands detected by the anti-p53 antibody in three repeat experiments using cells expressing the UTP23 colorectal cancer mutant was quantified and normalised to intensity of the karyopherin band from the same lane (Figure 5.12B). Because the UTP23 P215Q mutant only affected pre-rRNA processing in HEK293T cells when expressed at levels much higher than endogenous UTP23 (Figure 5.4), two concentrations of tetracycline were used to induce expression of both WT UTP23 (1 ng/μl and 10,000 ng/μl) and the P215Q mutant (10 ng/μl and 10,000 ng/μl) (Figure 5.12A lanes 3-6 and 7-10).

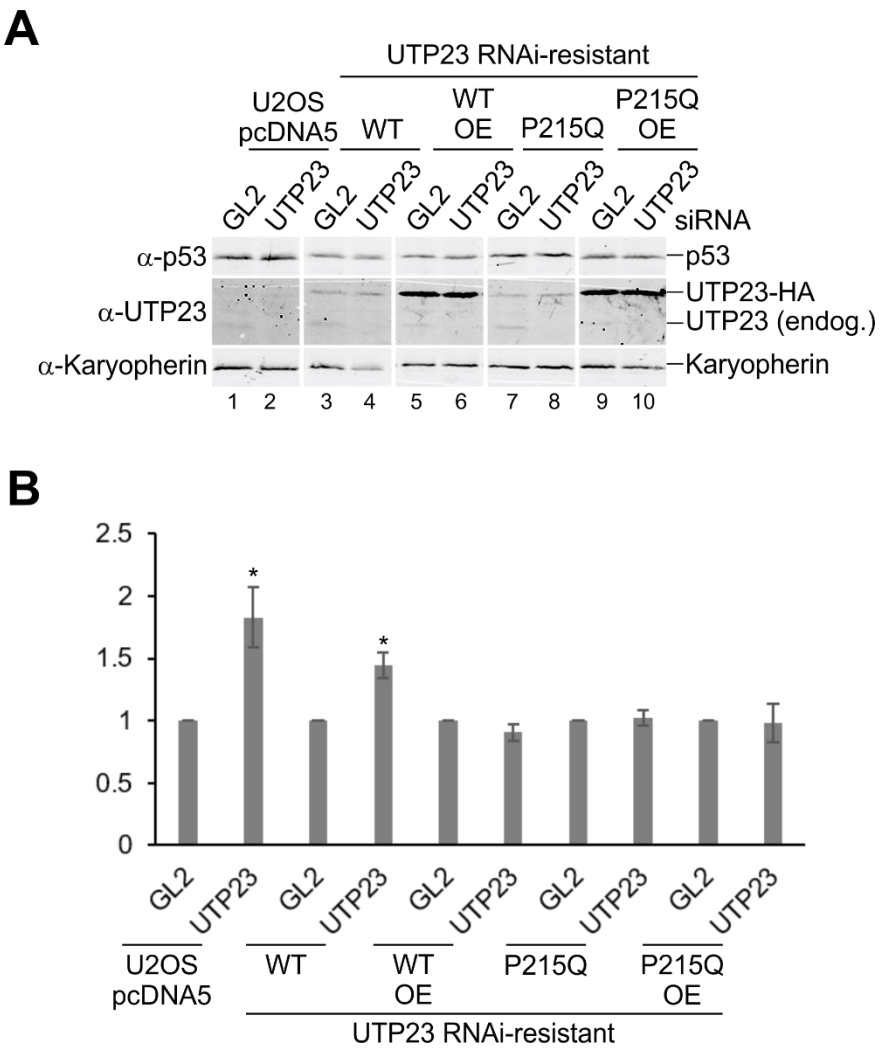


Figure 5.12 Overexpression of the UTP23 P215Q mutant does not cause p53 induction in U2OS cells. (A) U2OS Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (U2OS pcDNA5; lanes 1 and 2), or wild type (WT/WT OE; lanes 3 and 4/ lanes 5 and 6) or mutant (P215Q/P215Q OE; lanes 7 and 8/ lanes 9 and 10) HA-tagged RNAi-resistant UTP23. Following transfection of a control siRNA (GL2) or RNAi-mediated depletion of endogenous UTP23 (UTP23) and induction of HA-tagged UTP23 expression, protein was extracted from cells and separated by SDS-PAGE. Protein was analysed by western blotting using an anti-UTP23 antibody and an anti-p53 antibody. An anti-Karyopherin antibody was used as a loading control. **(B)** Quantification of p53 protein levels from multiple (>3) western blots following depletion of UTP23 in U2OS Flp-In cells stably transfected with plasmids expressing a FLAG-tag alone, or wild type (WT (1 ng/µl tetracycline)/WT OE (10,000 ng/µl tetracycline)) or mutant (P215Q (10 ng/µl tetracycline)/P215Q OE (10,000 ng/µl tetracycline)) HA-tagged RNAi-resistant UTP23. Average intensity of bands detected by the anti-p53 antibody, normalised to levels of Karyopherin are plotted compared to control (GL2) cells for each cell line. Error bars represent standard error and statistical analysis was performed using an unpaired t-test. *p value <0.05.

In control cells, expressing a FLAG-tag alone (pcDNA5; Figure 5.12A lanes 1 and 2), depletion of UTP23 caused a significant (~1.75-fold) increase in p53 levels as observed in plain U2OS and MCF7 cells (Figure 5.8). The expression of WT UTP23-HA at levels similar to endogenous UTP23 (lanes 3 and 4) was unable to rescue the p53 induction caused by depletion of endogenous UTP23, as a significant p53 induction was observed in these cells. This is consistent with the incomplete rescue by WT UTP23-HA, expressed at similar levels, of the pre-rRNA processing defect caused by UTP23 depletion in HEK293T cells (Chapter Four). Expression of WT UTP23-HA at increased levels (lanes 5 and 6), however, did not cause induction of p53, suggesting this protein replaces the endogenous UTP23 protein at higher levels. In cells expressing the colorectal cancer P215Q mutant (lanes 7 and 8), no increase in p53 levels was observed with or without depletion of endogenous UTP23. Levels of p53 did not increase in cells overexpressing the P215Q mutant at levels significantly higher than those of the endogenous protein (lanes 9 and 10), despite the fact that this mutant caused a pre-rRNA processing defect in HEK293T cells when overexpressed. However, there was a clear difference in p53 induction in cells expressing the P215Q compared to UTP23 WT at levels similar to those of endogenous UTP23. While expression of WT UTP23 at similar levels to endogenous UTP23 was unable to rescue the p53 induction caused by endogenous UTP23 depletion, expression of the P215Q mutant at similar levels did not cause p53 induction. This suggests that this mutation may have an effect on the p53 pathway in these cells, although it is unclear why. Further experiments are required to clarify this, such as analysis of pre-rRNA processing in U2OS cells expressing the UTP23 P215Q mutant.

5.2.8 Depletion of the 3'-5' exonuclease RRP6, but not the 5'-3' exonuclease XRN2, causes p53 induction in U2OS and MCF7 cells

Other proteins involved in processing of the human pre-rRNA around the 18S rRNA are also mutated in cancers, particularly factors which are important for processing of the 3' end of 18S within ITS1. This is especially interesting as ITS1 can be processed via different pathways in humans. The major ITS1 processing pathway occurs via a predominant endonucleolytic cleavage at site 2 by RNase MRP (Goldfarb and Cech, 2017), while the minor pathway relies on cleavage at site 2a, similar to the major processing pathway in yeast (Sloan *et al.*, 2013b). The existence of many cancer

mutations in key 3’ 18S processing factors suggest that these mutations may switch processing to the minor, “yeast-like” pathway.

One factor important for 18S 3’ end processing is the exosome-associated protein, RRP6. RRP6 is a 3’ to 5’ exonuclease with roles in a diverse range of RNA processing and degradation processes. It is involved in multiple pre-rRNA processing steps, and the exonuclease activity of RRP6 is required for the trimming of the 21S pre-rRNA to the 18SE pre-rRNA following endonuclease cleavage at site 2 (Sloan *et al.*, 2013b). High-throughput sequencing of cancers identified 78 cancer mutations in RRP6, with several mutations located in the vertebrate-specific C-terminal domain (Iorio *et al.*, 2016). To investigate the potential link between RRP6 and cancer, RNAi-mediated knockdowns were performed in cells U2OS and MCF7 cells, which contain active p53 genes. Western blotting was performed following transfection with a siRNA targeting RRP6, or a control (GL2) siRNA, using an anti-p53 antibody to detect p53 levels (Figure 5.13, left panels). Karyopherin was used as a loading control. RNAi-mediated depletion of RRP6 in both U2OS and MCF7 cells was confirmed by western blotting (data not shown). Quantified p53 levels from three repeat experiments, in both U2OS and MCF7 cells, showed that on average, depletion of RRP6 did not result in significant p53 induction.

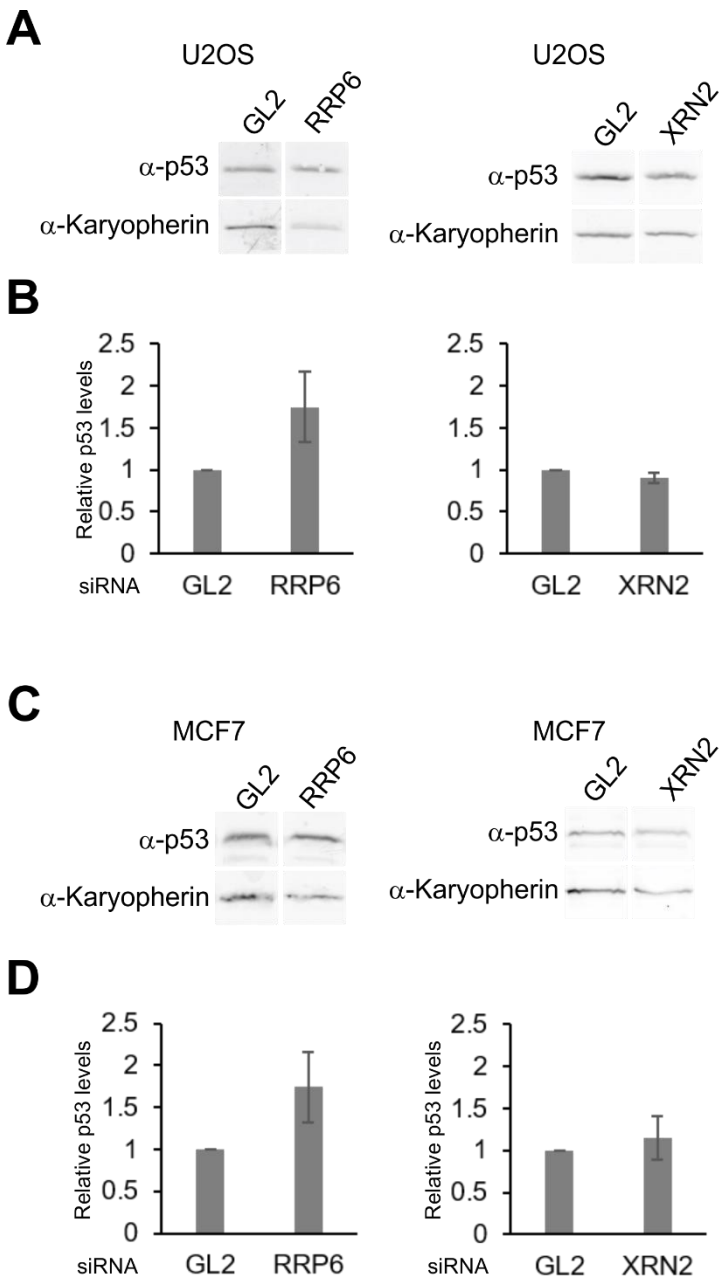


Figure 5.13 Depletion of RRP6 or XRN2 does not cause p53 induction in U2OS or MCF7 cells. (A) Analysis of protein levels from U2OS cells after RNAi-mediated depletion of RRP6 (left panel) or XRN2 (right panel) by western blotting using an anti-p53 antibody. An anti-Karyopherin antibody was used as a loading control. Multiple (>3) repeat experiments were performed, and a representative western blot is shown. **(B)** Quantification of p53 protein levels from multiple (>3) western blots following depletion of RRP6 (left panel) or XRN2 (right panel) in U2OS cells. Average intensity of bands detected by the anti-p53 antibody, normalised to levels of Karyopherin are plotted compared to control (GL2) cells. Error bars represent standard error and statistical analysis was performed using an unpaired t-test. **(C)** Analysis of protein levels from MCF7 cells after RNAi-mediated depletion of RRP6 (left panel) or XRN2 (right panel), as in panel A. **(D)** Quantification of p53 protein levels from multiple (>3) western blots following depletion of RRP6 (left panel) or XRN2 (right panel) in MCF7 cells, as in panel B.

Another ribosome biogenesis factor often mutated in cancers is the 5' to 3' exonuclease, XRN2. Eighty eight XRN2 cancer mutations have been identified (Iorio *et al.*, 2016), and XRN2 also contains a vertebrate-specific C-terminal domain. Like RRP6, XRN2 plays many important roles in RNA processing and degradation, including functions at multiple stages of pre-rRNA processing. In ITS1 specifically, XRN2 processes to the 5' end of 5.8S from site 2, following endonuclease cleavage at this site. Depletion of XRN2 was also shown to shift processing to the minor pathway, inducing cleavage at site 2a, rather than site 2. To test for p53 induction after XRN2 depletion, XRN2 was depleted in both U2OS and MCF7 cells and western blotting was performed using an anti-p53 antibody. Karyopherin was used as a loading control. Efficient depletion of XRN2 in U2OS and MCF7 cells was also confirmed by western blotting (data not shown). p53 levels following XRN2 knockdown did not change compared to the control siRNA (GL2) in multiple experiments with both U2OS and MCF7 cells (Figure 5.13, right panels). This suggests that, despite XRN2 depletion disrupting ribosome biogenesis and altering the ITS1 processing pathway utilised (data not shown), it does not induce p53. These results indicate that, despite their importance in processing of ITS1, a key event in both SSU and LSU biogenesis, and their frequent mutation in cancer, depletion of RRP6 or XRN2 does not induce p53 in U2OS and MCF7 cells.

5.2.9 Depletion of the RNA-binding protein RRP5 in U2OS and MCF7 cells causes a 5S RNP-dependent p53 induction

RRP5 is an RNA-binding protein involved in ribosome biogenesis. It is a component of the SSU processome but is also important for LSU biogenesis. In yeast, the two halves of Rrp5 are able to functionally replace the whole protein when expressed independently. The C-terminal half of Rrp5 binds close to the A_2 cleavage site (site 2a in humans) and is essential for this cleavage event (Venema and Tollervey, 1996; Eppens *et al.*, 1999). The N-terminal half binds near the A_3 site (site 2 in humans) and is essential for cleavage at this site. The exact role of RRP5 in the processing of human 5' ETS and ITS1 is unclear, however it is required for cleavages at sites A0, 1, 2a and 2, suggesting a homologous role to that of the yeast protein (Sweet *et al.*, 2008; Sloan *et al.*, 2013b). It is also a component of the human SSU processome and is part of a ~50S SSU processome assembly intermediate that accumulates upon inhibition of RNA polymerase I transcription (Turner *et al.*, 2009). Like the yeast protein, human

RRP5 is made up of multiple N-terminal S1 RNA-binding domains and C-terminal TPR protein-protein interaction domains. 168 cancer mutations were identified within human RRP5 (Iorio *et al.*, 2016).

Firstly, to assess the effect of RRP5 depletion on p53 levels, western blotting was performed with an anti-p53 antibody following RNAi-mediated knockdown of RRP5 in both U2OS and MCF7 cells (Figure 5.14). Karyopherin was used as a loading control. RNAi-mediated knockdown of RRP5 was also confirmed by western blotting in both U2OS and MCF7 cells (data not shown). A representative western blot with U2OS cells is shown in Figure 5.14A and p53 protein levels from five repeat experiments were quantified and normalised to levels of the loading control (Figure 5.14B, left panel). This showed that depletion of RRP5 resulted in significant increase in p53 levels in U2OS cells, with an average 2.5-fold p53 induction. This was similar to the p53 induction observed upon depletion of RPL7 (Figure 5.14B, right panel). A representative western blot from MCF7 cells is shown in Figure 5.14E and quantified p53 levels from four repeat experiments are shown in Figure 5.14F (left panel). Depletion of RRP5 in MCF7 cells also caused a significant increase in p53 levels, with an average 3-fold induction, which was again similar to the p53 induction observed after RPL7 depletion (Figure 5.14F, right panel).

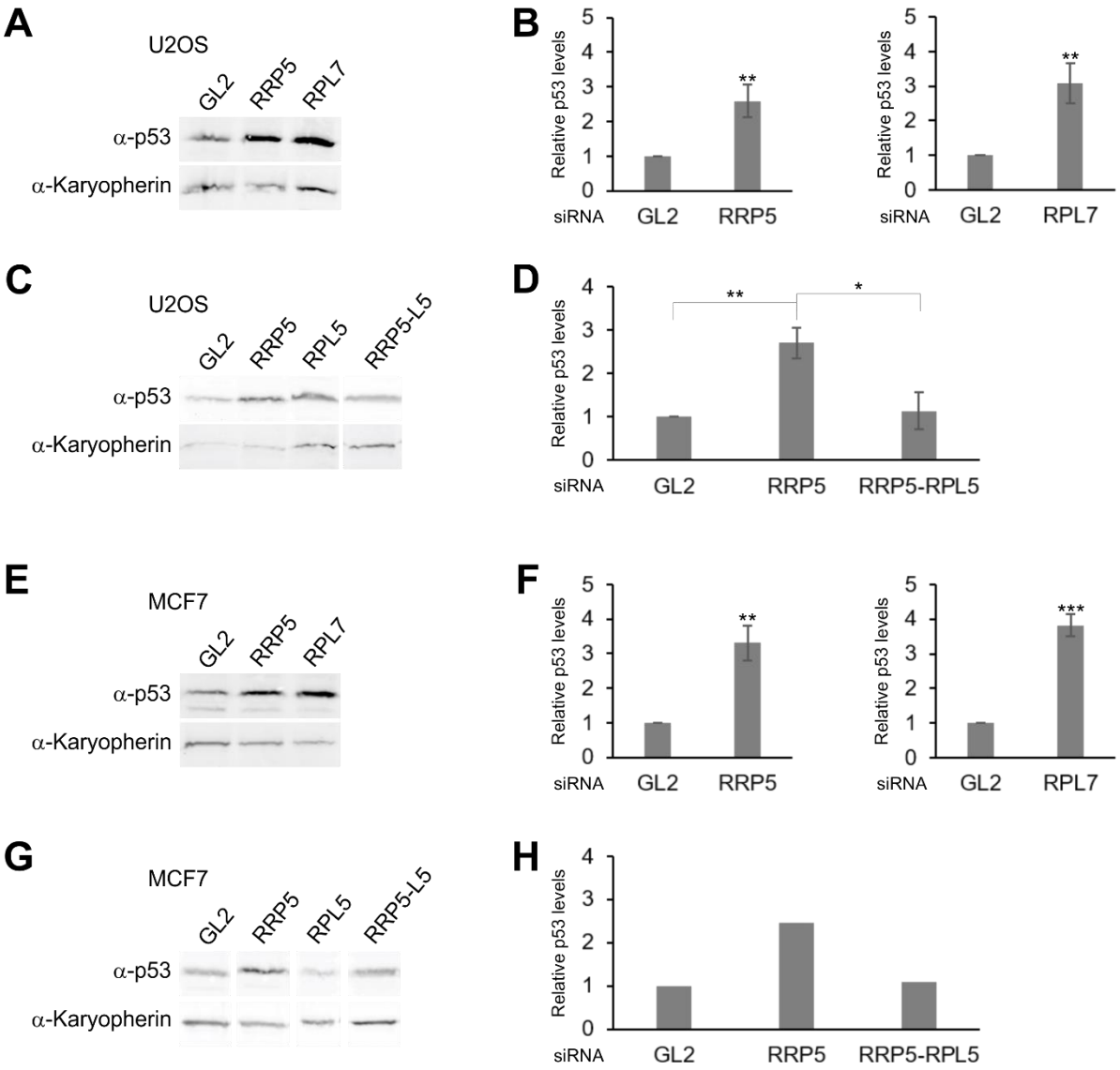


Figure 5.14 Depletion of RRP5 causes strong, 5S RNP-mediated p53 induction in U2OS and MCF7 cells. (A) Analysis of protein levels from U2OS cells after RNAi-mediated depletion of RRP5 or RPL7 by western blotting using an anti-p53 antibody. An anti-Karyopherin antibody was used as a loading control and RPL7 depletion was used as a positive control for strong increase in p53 protein levels. Multiple (>3) repeat experiments were performed, and a representative western blot is shown. **(B)** Quantification of p53 protein levels from multiple (>3) western blots following depletion of RRP5 (left panel) or RPL7 (right panel) in U2OS cells. Average intensity of bands detected by the anti-p53 antibody, normalised to levels of Karyopherin are plotted compared to control (GL2) cells. Error bars represent standard error and statistical analysis was performed using an unpaired t-test. **(C)** Analysis of protein levels from U2OS cells after RNAi-mediated depletion of RRP5, RPL5 or both, as in panel A. **(D)** Quantification of p53 protein levels from multiple (>3) western blots following depletion of RRP5, RPL5 or both in MCF7 cells, as in panel B. **(E)** Analysis of protein levels from MCF7 cells after RNAi-mediated depletion of RRP5 or RPL7, as in panel A. **(F)** Quantification of p53 protein levels from multiple (>3) western blots following depletion of RRP5 (left panel) or RPL7 (right panel) in MCF7, cells as in panel B. **(G)** Analysis of protein levels from MCF7 cells after RNAi-mediated depletion of RRP5, RPL5 or both, as in panel A. **(H)** Quantification of average p53 protein levels from two repeat

experiments following depletion of RRP5 (left panel) or RPL7 (right panel) in MCF7, cells as in panel B. *p value <0.05 **p value <0.01 ***p value <0.001.

To assess whether this p53 induction is dependent on the 5S RNP-MDM2 pathway, RNAi-mediated knockdown of RRP5 was performed in conjunction with knockdown of the LSU ribosomal protein RPL5, a key component of the 5S RNP. A representative western blot from U2OS cells is shown in Figure 5.14C and quantification of p53 levels from three repeat experiments is shown in Figure 5.14D. With depletion of RRP5, p53 levels were significantly increased (average 2.5-fold induction) compared to the control siRNA, while knockdown of both RRP5 and RPL5 caused a significant decrease in p53 levels. This indicates that the p53 induction observed upon depletion of RRP5 is dependent on signalling by the 5S RNP in MDM2 inhibition and subsequent p53 activation in U2OS cells. This effect was also observed in MCF7 cells (Figure 5.14G-H), but only two experiments were performed in these cells, so its significance cannot be determined.

5.2.10 Depletion of RRP5 disrupts pre-rRNA processing in HEK293T and MCF7 cells

As Rrp5 in yeast is required for cleavages at both sites (A_2 and A_3) in ITS1 (Venema and Tollervey, 1996), it is possible that it plays a role in the selection of the pre-rRNA processing pathway used. The function of Rrp5 can be separated with regards to the two ITS1 cleavages. The N-terminal (NTD) portion of the protein binds the pre-rRNA close to, and is important for cleavage at, site A_3, while the C-terminal half (CTD) binds close to site A_2, and is required for cleavages at sites A_0, A_1 and A_2 (Eppens *et al.*, 1999; Lebaron *et al.*, 2013). Human pre-rRNA processing of ITS1 differs from yeast, with the “major” cleavage occurring at site 2 (equivalent to yeast site A_3), rather than 2a (A_2). If RRP5 plays a similar role in humans to the yeast protein within ITS1, it may play a role in recruitment of cleavage enzymes or other factors to these sites, thus mediating selection of a specific pre-rRNA processing pathway. Alternatively, as a large RNA binding protein, its binding to the pre-rRNA may alter structure, making a particular cleavage site more accessible. As defects in multiple factors involved in the 3’ end processing of the 18S rRNA are linked to ribosomopathies and cancers, it is possible that these changes may alter the pre-rRNA processing pathway used in cancer cells.

In yeast, depletion of Rrp5 impairs cleavage at sites A_0, A_1, A_2 and A_3, causing accumulation of a 24S precursor, which is uncleaved at all four sites (Venema and Tollervey, 1996). Depletion of RRP5 in HeLa cells caused an increase in levels of the 45S pre-rRNA intermediate and reduction of 18SE pre-rRNA levels, as well as a mild reduction of 30S pre-rRNA levels (Sloan *et al.*, 2013b). Two novel pre-rRNA intermediates were also observed in the same study, a 3'-extended form of 21S (21SL3') and a 3'-extended form of 30S (30SL3'). Both of these novel intermediates were uncleaved at site 2 and contained the 5.8S rRNA sequence. These results suggested that human RRP5 plays an analogous role to yeast Rrp5 in cleavages at A0 and 1, as well as both cleavage events within ITS1.

To further investigate the role of RRP5 in human pre-rRNA processing, RNAi-mediated knockdown was performed in HEK293T and MCF7 cells. Attempts to extract RNA from U2OS cells were unsuccessful, so pre-rRNA processing was not analysed in these cells. Following RNA isolation, northern blotting was performed with ITS1, 18SE, ETS1 and 3' ITS1 probes (Figure 5.15 and 5.16). Positions of these probes are shown in Figures 5.15A and 5.16A. XRN2 depletion was also performed, both alone (XRN2) and in conjunction with RRP5 knockdown (X-RRP5), to examine the role of RRP5 with respect to cleavage at the 2a site, as XRN2 depletion shifts pre-rRNA processing to the minor, "2a-specific" pathway. In HEK293T cells (Figure 5.15), depletion of RRP5 caused accumulation of an extended 30S intermediate and a 21SL3' intermediate (lanes 1 and 2), as previously observed in HeLa cells (Sloan *et al.*, 2013b). The 30SL3' intermediate cannot be distinguished from the 30SL5' using the ITS1 probe as they are too similar in size. Therefore, two other probes were used to clarify the identity of the band detected by the ITS1 probe. The ETS1 probe (Figure 5.15D detects only the 5' extended precursor, 30SL5', observed upon impairment of A' cleavage, while the 3' ITS1 probe (Figure 5.15E) detects only the 3' extended 30SL3' precursor, which accumulates when cleavage at site 2 is defective.

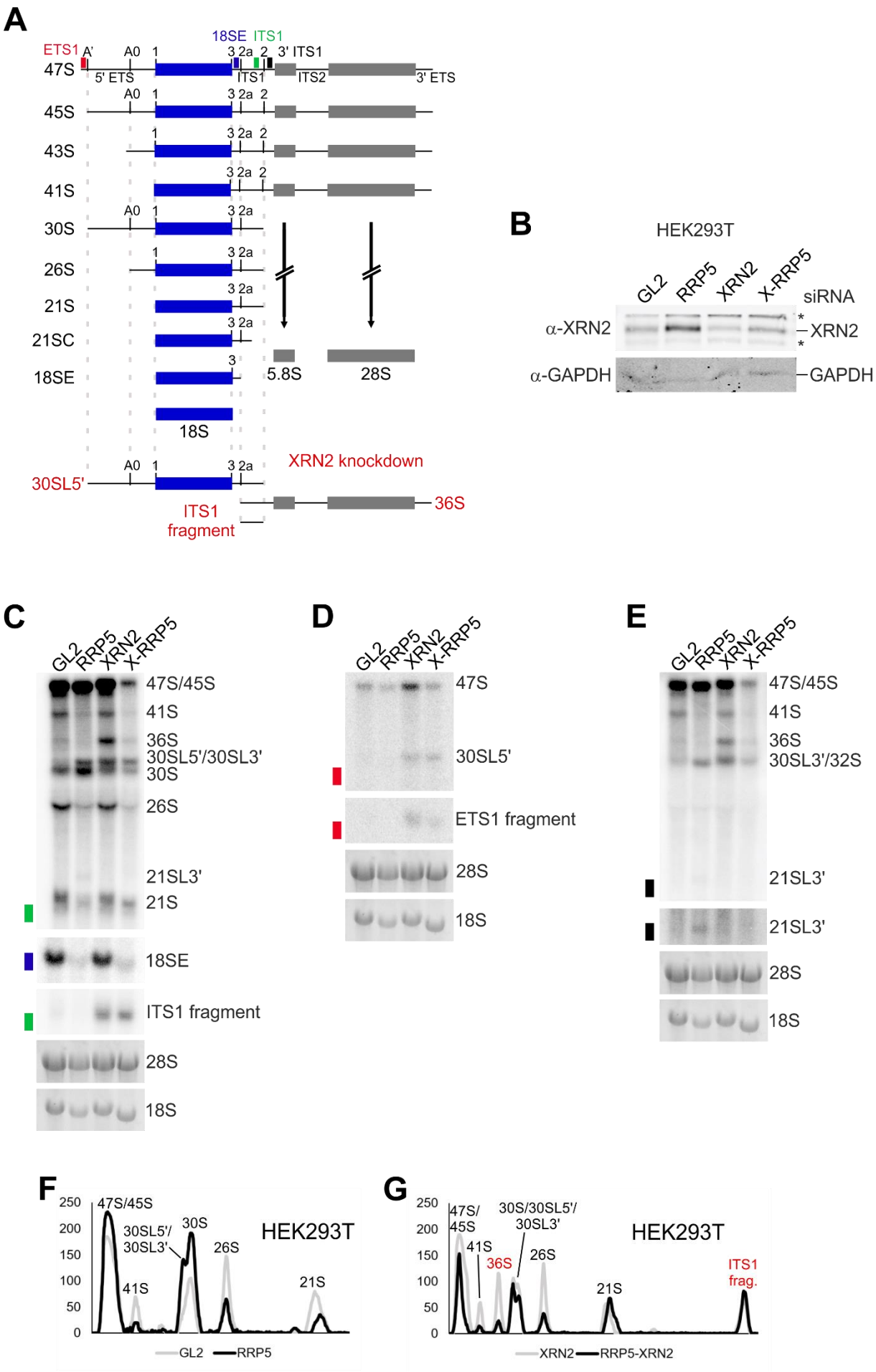

A
18SE ITS1
ETS1 A' A0 1 3 2a 2 3' ITS1
47S 5' ETS ITS1 ITS2 3' ETS
45S
43S
41S
30S
26S
21S
21SC
18SE
5.8S
28S
18S
XRN2 knockdown
30SL5'
36S
ITS1 fragment
B
HEK293T
GL2 RRP5 XRN2 X-RRP5
siRNA
α-XRN2
XRN2
α-GAPDH
GAPDH
C
GL2 RRP5 XRN2 X-RRP5
47S/45S
41S
36S
30SL5'/30SL3'
30S
26S
21SL3'
21S
18SE
ITS1 fragment
28S
18S
D
GL2 RRP5 XRN2 X-RRP5
47S
30SL5'
ETS1 fragment
28S
18S
E
GL2 RRP5 XRN2 X-RRP5
47S/45S
41S
36S
30SL3'/32S
21SL3'
21SL3'
28S
18S
F
250
200
150
100
50
0
47S/45S
30SL5'/
30SL3'
30S
26S
41S
21S
HEK293T
GL2 RRP5
G
250
200
150
100
50
0
47S/
45S
30S/30SL5'/
30SL3'
41S
36S
26S
21S
ITS1
frag.
HEK293T
XRN2 RRP5-XRN2

Figure 5.15 RRP5 is required for cleavages at sites A0, 1, 2a and 2 in HEK293T cells. (A) Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Pre-rRNA intermediates specific to the minor, 2a-dependent processing pathway, which accumulate upon depletion of XRN2, are shown in red. Positions of probes used in northern blotting (18SE; blue, ITS1; green, ETS1; red, 3' ITS1; black) are marked on the primary precursor. **(B)** Western blot using an anti-XRN2 antibody to confirm depletion of endogenous XRN2 upon RNAi-mediated RRP5 depletion in HEK293T cells. Karyopherin was used as a loading control. **(C-E)** Following treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous RRP5 (RRP5), XRN2 (XRN2) or both (X-RRP5), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane. RNA was analysed by northern blotting using radiolabelled probes hybridising to **(C)** the 5' end of ITS1 (18SE; blue rectangle), upstream of cleavage site 2 (ITS1; green rectangle), **(D)** upstream of cleavage site A' (ETS1; red rectangle) or **(E)** downstream of cleavage site 2 (3' ITS1; black rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. Levels of pre-rRNAs detected by the ITS1 probe from panel C after depletion of **(F)** RRP5 alone or **(G)** RRP5 and XRN2 were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

Depletion of RRP5 in HEK293T cells caused accumulation of the 30SL3', detected by the 3' ITS1 probe, consistent with previous results in HeLa cells (Figure 5.15E, lanes 1 and 2). The accumulation of' extended precursors, 21SL3' and 30SL3', as well as reduction of 26S, 21S and 18SE pre-rRNA levels, indicates that all three early SSU cleavages at sites A0, 1 and 2a, as well as the major cleavage in ITS1 at site 2 were disrupted (Figure 5.15C and 5.15E). Quantified northern blotting band intensities detected by the ITS1 probe from these lanes, normalised to levels of the mature 28S rRNA, are shown in Figure 5.15F. Depletion of XRN2 (Figure 5.15C, lane 3) caused accumulation of the "2a-specific" precursors, the 36S pre-rRNA and "ITS1 fragment", as well as the 30SL5', indicating defective A' processing. The band representing the extended form of 30S pre-rRNA upon XRN2 depletion is detected with the ETS1 probe, indicating that this band corresponds to the 30SL5' intermediate (Figure 5.15D). Upon depletion of both XRN2 and RRP5 (Figure 5.15C, lane 4), levels of the 36S pre-rRNA were greatly reduced, suggesting that processing at the 2a site in the "minor" pathway is disrupted with depletion of RRP5. However, levels of the ITS1 fragment remained unchanged, in contrast to results from HeLa cells, suggesting that cleavages at sites 2a and 2 are not fully abolished. The quantified data for these lanes, normalised to levels of the mature 28S rRNA, are shown in Figure 5.15G. Western blotting was performed using an anti-XRN2 antibody to determine XRN2 depletion efficiency (Figure 5.15B).

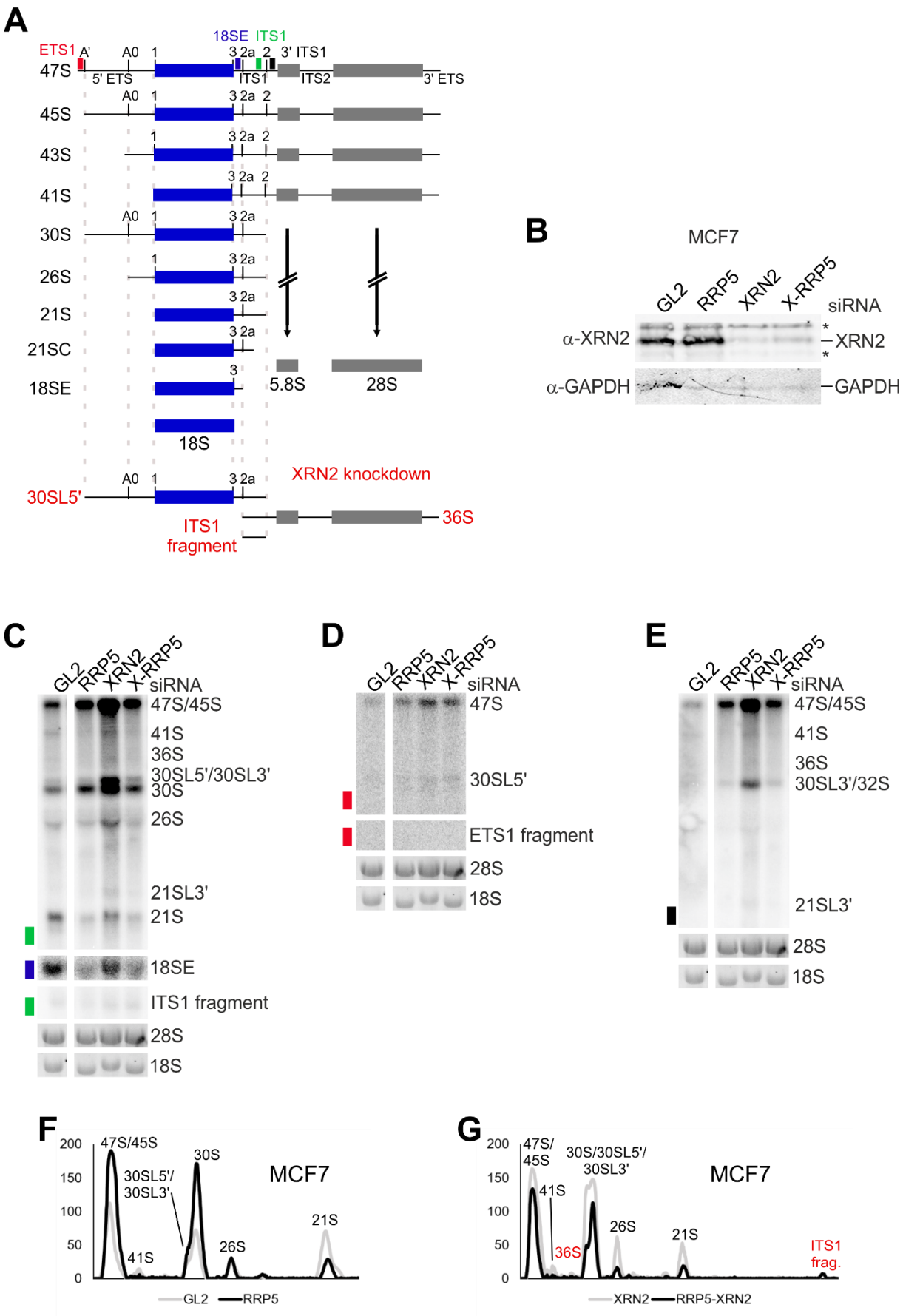


Figure 5.16 Pre-rRNA processing defects observed after RRP5 and XRN2 depletion in MCF7 cells. (A) Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Pre-rRNA

intermediates specific to the minor, 2a-dependent processing pathway, which accumulate upon depletion of XRN2, are shown in red. Positions of probes used in northern blotting (18SE; blue, ITS1; green, ETS1; red, 3' ITS1; black) are marked on the 47S pre-rRNA. **(B)** Western blot using an anti-XRN2 antibody to confirm depletion of endogenous XRN2 upon RNAi-mediated RRP5 depletion in MCF7 cells. Karyopherin was used as a loading control. **(C-E)** Following treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous RRP5 (RRP5), XRN2 (XRN2) or both (X-RRP5), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane. RNA was analysed by northern blotting using radiolabelled probes hybridising to **(C)** the 5' end of ITS1 (18SE; blue rectangle), upstream of cleavage site 2 (ITS1; green rectangle), **(D)** upstream of cleavage site A' (ETS1; red rectangle) or **(E)** downstream of cleavage site 2 (3' ITS1; black rectangle). Mature 28S and 18S rRNAs were visualised using ethidium bromide staining and pre-rRNAs detected by northern blot probes are labelled. Levels of pre-rRNAs detected by the ITS1 probe from panel C after depletion of **(F)** RRP5 alone or **(G)** RRP5 and XRN2 were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted. The pre-rRNA intermediate represented by each peak is labelled.

In MCF7 cells, depletion of RRP5 produced a distinct phenotype compared to those observed in HeLa and HEK293T cells (Figure 5.16). The most prominent phenotype observed with the ITS1 probe was a mild accumulation of the 30S pre-rRNA, and reduction in levels of the 26S, 21S and 18SE precursors, suggesting defects in early cleavages at sites A0, 1 and 2a (Figure 5.16C, lanes 1 and 2). This is consistent with previous analysis of pre-rRNA levels in MCF7 cells**,** which showed that these cells appear to already have higher levels of the 30S intermediate upon treatment with the control siRNA. Unlike in HEK293T cells, only very mild accumulation of the 21SL3' was observed upon RRP5 depletion in MCF7 cells, detected by both ITS1 and 3' ITS1 probes (Figure 5.16C and 5.16E). Quantified pre-rRNA levels detected by the ITS1 probe, normalised to levels of the mature 28S rRNA, are shown in Figure 5.16F. RNAi-mediated knockdown of XRN2 caused the accumulation of the 30SL5' precursor (Figure 5.16C and 5.16D), but levels of the 36S intermediate and the ITS1 fragment were not detectably increased (Figure 5.16C). Western blotting analysis of these samples, (Figure 5.15B) using an anti-XRN2 antibody confirmed that XRN2 was efficiently depleted. Therefore, the lack of accumulation of the "2a-specific" 36S pre-rRNA and ITS1 fragments in these cells may suggest a potential natural defect in cleavage at site 2a in MCF7 cells. Knockdown of both RRP5 and XRN2 caused accumulation of the 30S intermediate and a relative decrease in 30SL5' levels compared to XRN2 knockdown (Figure 5.16C and 5.16D). The quantified pre-rRNA levels from Figure 5.16C (lanes 3 and 4), normalised to mature 28S rRNA levels, are

shown in Figure 5.16G. This pre-rRNA processing phenotype was similar to the single depletion of RRP5, again suggesting a defect in processing through the 2a site in MCF7 cells. The lack of 36S pre-rRNA and ITS1 fragment upon XRN2 depletion suggest that loss of XRN2 does not affect ITS1 processing in MCF7 cells. This suggests that ITS1 processing may occur differently in this cell type compared to HeLa and HEK293T cells, and possibly explains the lack of p53 induction upon XRN2 depletion in these cells.

5.2.11 Creation of stable HEK293T and MCF7 cell lines expressing RNAi-resistant forms of RRP5

Analysis of pre-rRNA intermediate levels upon RRP5 depletion suggests that, at least in HeLa and HEK293T cells, human RRP5 is important for processing of both sites within ITS1, a role conserved from yeast. In yeast, different domains within the large, multidomain Rrp5 protein are important for distinct pre-rRNA processing events (Figure 5.16A**)**. The N-terminal half of Rrp5, comprising 9 of the 12 S1 RNA-binding domains, is important for cleavage at the A_3 site (equivalent to human site 2), while the C-terminal half (three S1 domains and 7 TPR protein-protein domains) is important for cleavage at sites A_0 (human site A0), A_1 (human site 1) and A_2 (human site 2a) (Eppens *et al.*, 1999). Deletion of the last three S1 domains (S1-10 to S1-12) of Rrp5 results in loss of cleavage at the A_2 site (Vos *et al.*, 2004b). When expressed separately, the N- and C-terminal halves of Rrp5 bind the pre-rRNA close to the A_3 and A_2 sites respectively (Lebaron *et al.*, 2013), suggesting Rrp5 may play a role in coordinating endonuclease cleavage at these sites.

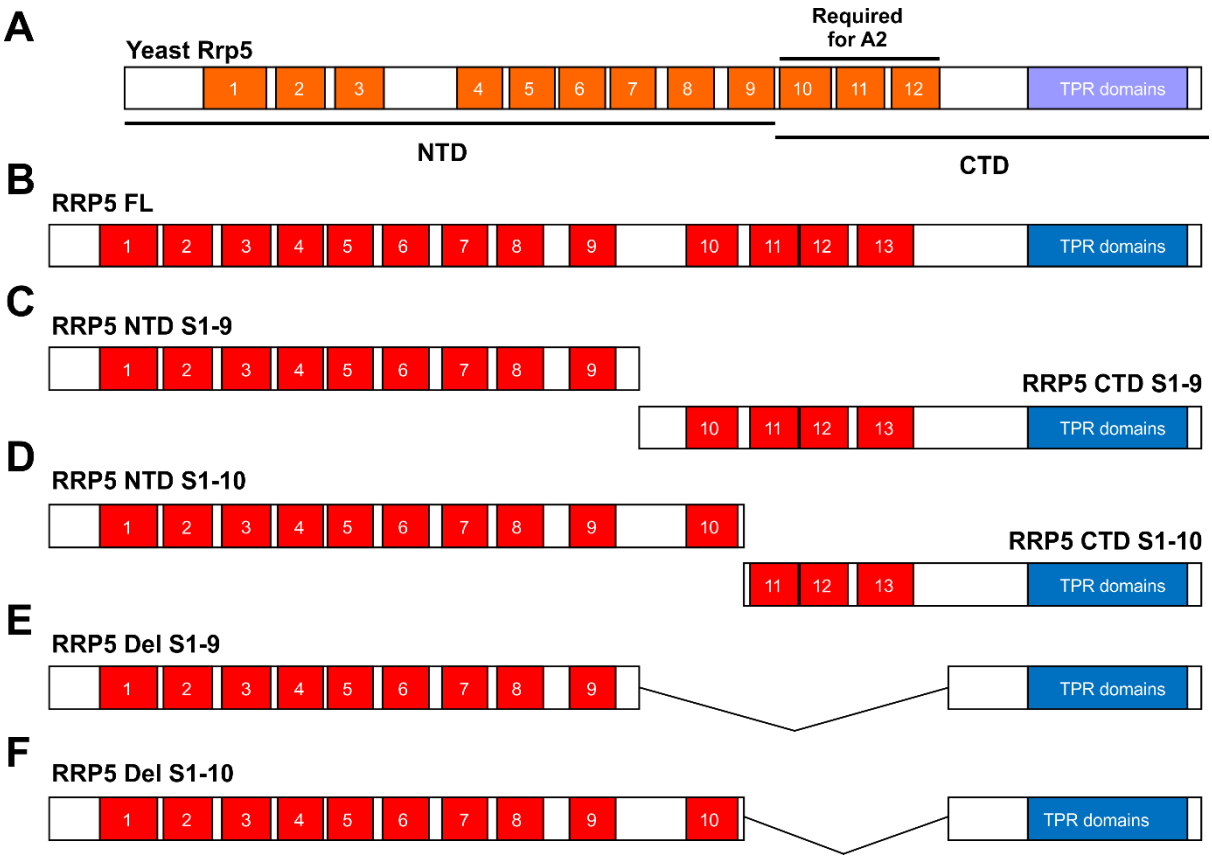


Figure 5.17 The domains of yeast and human Rrp5 proteins and the design of human RRP5 mutants. (A) Cartoon depiction of the domain architecture of yeast Rrp5, with S1 RNA binding domains shown in orange. The functionally independent NTD and CTD domains are marked and arrows indicate the requirement of these domains in processing at different cleavage sites **(B)** Cartoon representation of full-length human RRP5 (RRP5 FL) domain architecture with S1 RNA binding domains shown in red. The domain architecture of human RRP5 mutant proteins, containing **(C)** S1 domains 1 to 9 (RRP5 NTD S1-9), S1 domains 10 to 13 plus the C-terminal TPR domains (RRP5 CTD S1-19), **(D)** S1 domains 1 to 10 (RRP5 NTD S1-10), S1 domains 11 to 13 plus the C-terminal TPR domains (RRP5 CTD S1-10), **(E)** S1 domains 1 to 9 and the C-terminal TPR domains (RRP5 Del S1-9) or **(F)** S1 domains 1 to 10 and the C-terminal TPR domains (RRP5 Del S1-10).

As the processing within human ITS1 is different to yeast, with the major cleavage occurring at site 2 (yeast A_3), and mutations in genes encoding factors required for processing of ITS1 and the 3' end of 18S rRNA, including RRP5, are found in ribosomopathies and cancers, it will be interesting to discover the role of the different domains in human RRP5 in processing of ITS1. Human RRP5 is a >200 kDa protein comprising 13 S1 domains and 7 C-terminal TPR domains (Figure 5.17B). To test the importance of these domains in pre-rRNA processing, N- and C-terminal truncations of RRP5 were designed and generated. As human RRP5 has an extra S1 domain compared to the yeast protein, two different forms of each truncation were produced,

separated after the ninth S1 domain (NTD S1-9 and CTD S1-9) (Figure 5.17C) or separated after the tenth S1 domain (NTD S1-10 and CTD S1-10) (Figure 5.17D). As the three S1 domains in the C-terminal half of Rrp5 (S1-9 to S1-12) were shown to be important for cleavage at site A_2, human RRP5 mutations were also generated deleting these S1 domains (Figure 5.17E and 5.17F). Again, two variants of the deletion mutant were produced, with either S1 domains 10-13 (Del S1-9) (Figure 5.15E) or 11-13 (Del S1-10) (Figure 5.17F) deleted. The ORFs of WT RRP5 and all mutant forms of RRP5 were altered to make them resistant to the siRNA targeting endogenous RRP5. The ORFs of WT RRP5 and all 6 mutant forms were cloned into the pcDNA5 plasmid, where they were fused to two N-terminal FLAG-tags and under the control of a tetracycline-inducible promoter. Stable HEK293T and MCF7 cell lines were generated by stable transfection of pcDNA5-RRP5 constructs into HEK293T or MCF7 Flp-In cells.

5.2.12 Expression of RNAi-resistant WT RRP5 rescues the pre-rRNA processing phenotype caused by depletion of RRP5 in HEK293T cells

After selection of stably transfected cells, expression of FLAG-tagged RRP5 was induced by tetracycline induction at a range of concentrations. Western blotting was used to confirm the expression of FLAG-tagged WT RRP5 (Figure 5.18). As RRP5 is a very large (>200 kDa) protein, the separation between endogenous RRP5 and the slightly larger, FLAG-tagged WT RRP5 protein is not observable with the anti-RRP5 antibody (Figure 5.18A and C). Therefore, an anti-FLAG antibody was also used to confirm expression of the FLAG-tagged form of WT RRP5 (Figure 5.18B and D). The separate blotting of endogenous RRP5 and FLAG-RRP5 prevented the direct comparison of protein levels. To overcome this, a sample from cells expressing RRP5 Del S1-9 (see Figure 5.20) with 10 ng/µl tetracycline (expressed at similar levels to endogenous RRP5) was separated with WT RRP5 and analysed with the anti-FLAG antibody (Figure 5.18B and D). Comparison of FLAG-RRP5 WT to FLAG-RRP5 Del S1-9 therefore allowed for estimation of the tetracycline concentration required to express FLAG-RRP5 WT to similar levels to endogenous RRP5.

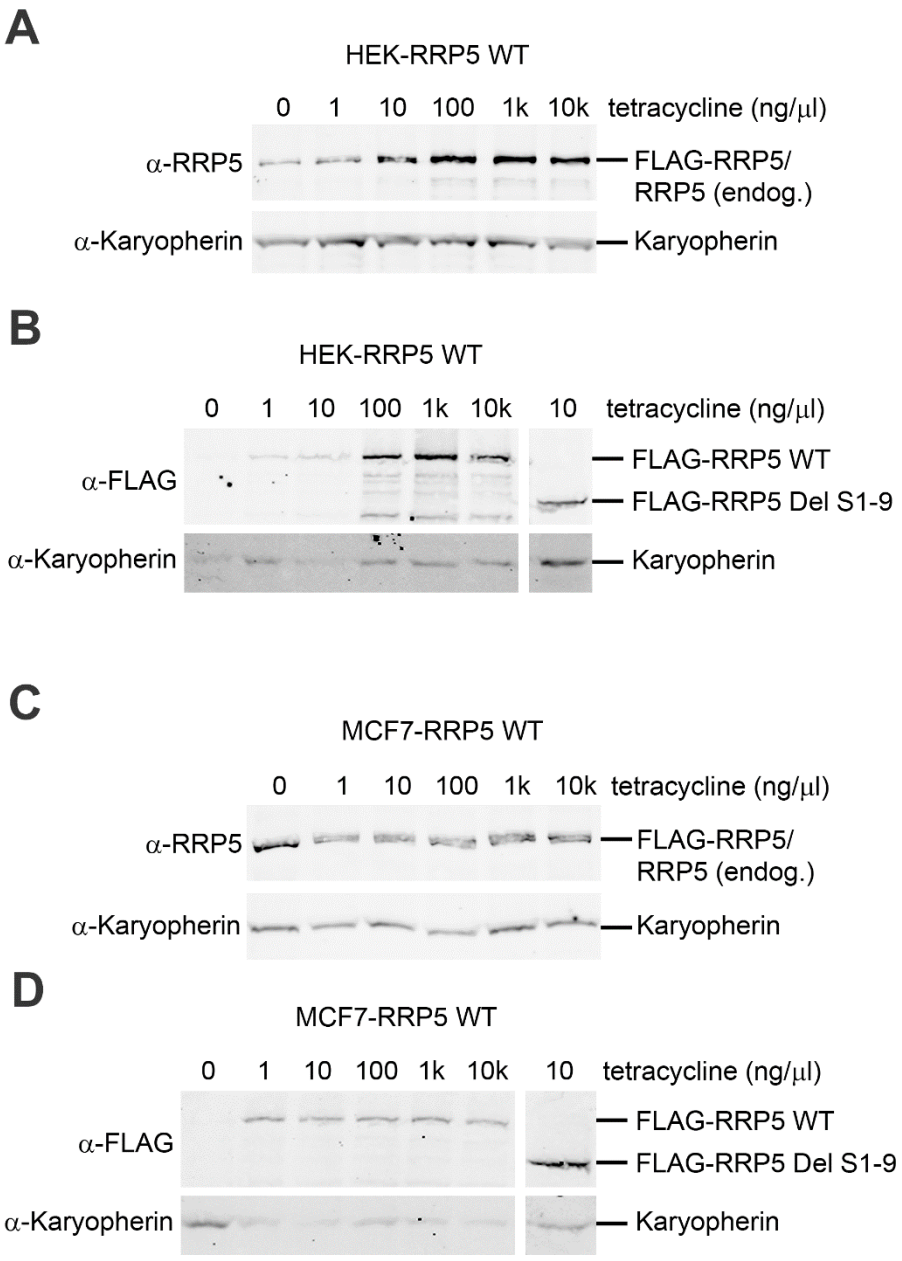


Figure 5.18 Expression of wild type RRP5 in stably transfected HEK293T and MCF7 cells. (A and B) HEK293T Flp-In or **(C and D)** MCF7 Flp-In cells stably transfected with FLAG-tagged wild type RRP5 were incubated with a range of concentrations of tetracycline (0-10,000 ng/μl). Protein was separated by SDS-PAGE and analysed by western blotting using **(A and C)** an anti-RRP5 antibody to detect both endogenous and FLAG-tagged RRP5 or **(B and D)** an anti-FLAG antibody to detect FLAG-tagged RRP5. An anti-Karyopherin antibody was used as a loading control. Panels B and D include a protein sample from cells expressing a FLAG-tagged RRP5 Del S1-9 mutant protein expressed at similar levels to endogenous RRP5 (see Figure 5.19).

In MCF7 cells, expression of FLAG-tagged RRP5 was induced with addition of tetracycline, but it did not appear to be responsive to increased tetracycline concentration, as FLAG-RRP5 levels did not increase with increasing concentration of

tetracycline (Figure 5.18C and 5.18D). In contrast, stably transfected HEK293T cells expressed FLAG-tagged RRP5 at increasing levels with increasing concentration of tetracycline (Figure 5.18A and 5.18B). Therefore, the RNAi-rescue experiments to confirm that expression of WT RRP5 is able to compensate for the depletion of endogenous RRP5 were performed in HEK293T cells.

To confirm that expression of RNAi-resistant WT RRP5 can rescue the observed pre-rRNA processing phenotype caused by depletion of RRP5, northern blotting using ITS1, ETS1 and 18SE probes was performed following induction of FLAG-tagged RRP5 expression and RNAi-mediated knockdown of RRP5 (Figure 5.19D and 5.19E). In control cells, expressing the FLAG-tag alone (pcDNA5) (lanes 1 and 2), depletion of RRP5 caused accumulation of an extended form of the 30S pre-rRNA, as previously observed. This was detected by the ITS1 (Figure 5.19D) and 3' ITS1 probes (Figure 5.19E), indicating that this band corresponds to the 30SL3' intermediate. Mild accumulation of 21SL3', detected by the 3' ITS1 probe (Figure 5.19E), was observed upon depletion of RRP5 in control cells. Levels of 26S, 21S and 18SE pre-rRNAs were also reduced after RRP5 depletion (Figure 5.19D). Quantified bands detected by the ITS1 probe from Figure 5.19D, normalised to levels of the mature 28S rRNA, are shown in Figures 5.19F and 5.19G. Depletion of endogenous RRP5 in cells expressing RNAi-resistant, FLAG-tagged RRP5 (Figure 5.19D, lanes 3 and 4) did not cause accumulation of 30SL3' or 21SL3', indicating that expression of WT RRP5 was able to rescue pre-rRNA processing defects caused by endogenous RRP5 depletion. Quantified northern blotting data show that levels of pre-rRNAs in these cells did not change with RRP5 depletion (Figure 5.19G), and the lack of 30SL3' accumulation is evident. Levels of the 26S, 21S and 18SE precursors were also recovered with expression of RNAi-resistant WT RRP5 (Figure 5.19G).

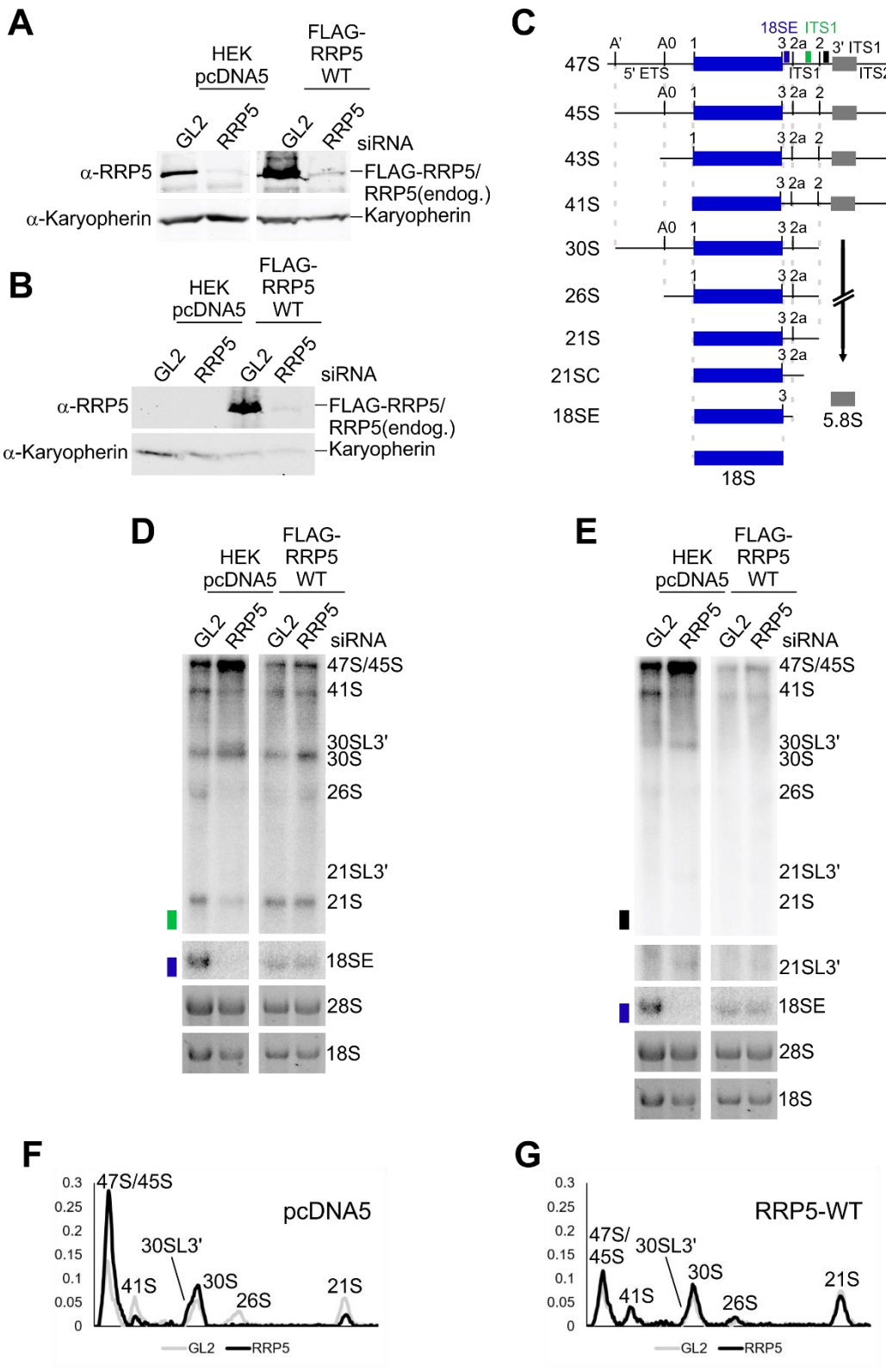


Figure 5.19 Expression of RNAi-resistant RRP5 rescues the pre-rRNA processing defect caused by depletion of endogenous RRP5 in HEK293T cells. HEK293T Flp-In cells were stably transfected with plasmids expressing a FLAG-tag alone (HEK pcDNA5), or N-terminally FLAG-tagged wild type (WT) RNAi-resistant RRP5. **(A)** Western blot using an anti-RRP5 antibody to confirm depletion of endogenous RRP5 upon RNAi-mediated RRP5 depletion in stably transfected HEK293T cells. Karyopherin is used as a loading control. **(B)** Western blot using an anti-FLAG antibody to confirm the expression of FLAG-tagged RRP5 in stably

transfected HEK293T cells. Karyopherin is used as a loading control. **(C)** Schematic of pre-rRNA processing in human cells, showing intermediates generated by successive processing events. Positions of the probes used in northern blotting (ITS1; green, 18SE; blue, 3' ITS1; black) are marked on the 47S pre-rRNA. Following induction of FLAG-tagged RRP5 expression and treatment with a control siRNA (GL2) or RNAi-mediated depletion of endogenous RRP5 (RRP5), RNA was extracted from cells, separated by agarose gel electrophoresis and transferred to a nylon membrane. RNA was analysed by northern blotting using radiolabelled probes hybridising to **(D)** upstream of cleavage site 2 (ITS1; green rectangle), upstream of cleavage site 2a (18SE; blue rectangle) and **(E)** downstream of cleavage site 2. Levels of pre-rRNAs detected by the ITS1 probe from panel D were quantified using ImageQuant, normalised to levels of the 47S/45S pre-rRNA and plotted with the identity of each peak marked.

Western blotting using an anti-RRP5 antibody was used to confirm the depletion of endogenous RRP5 and expression of FLAG-tagged RRP5 (Figure 5.19A). Because of the large size of RRP5, the endogenous and FLAG-tagged forms are indistinguishable by western blotting, but the reduction in the intensity of the band detected by the anti-RRP5 antibody suggests that endogenous RRP5 was efficiently depleted, and also suggests that FLAG-tagged WT RRP5 may be expressed at lower levels than endogenous RRP5. To analyse levels of the FLAG-tagged RRP5 protein, western blotting was also performed using an anti-FLAG antibody (Figure 5.19B). This showed that FLAG-RRP5 was expressed in samples treated with both GL2 and RRP5 siRNAs, but at much lower levels in the sample with endogenous RRP5 depleted (Figure 5.19B, lane 4). However, the pre-rRNA processing phenotype observed in these cells show that this expression is sufficient to compensate for depletion of RRP5. Therefore, stably transfected HEK293T cells were used to express mutant forms of RRP5.

5.2.13 RRP5 N-terminal halves and deletions are efficiently expressed in HEK293T stable cell lines

RRP5 N-terminal (NTD S1-9 and NTD S1-10) and C-terminal (CTD S1-9 and CTD S1-10) and RRP5 deletion (Del S1-9 and Del S1-10) pcDNA5 constructs were transfected into HEK293T Flp-In cells and western blotting was performed to analyse expression of RRP5 mutant proteins (Figure 5.20). As all mutant proteins produced are considerably smaller than full-length RRP5, the anti-RRP5 antibody, raised against a peptide in the N-terminal domain of RRP5, could be used to separately detect levels of both endogenous and mutant FLAG-tagged RRP5. GAPDH was used as a loading

control. RRP5 NTD S1-9 (Figure 5.20A), NTD S1-10 (5.20B), Del S1-9 (5.20C) and Del S1-10 (5.20D) were all expressed efficiently from the respective cell lines upon tetracycline treatment, and protein levels increased with addition of increased tetracycline concentration.

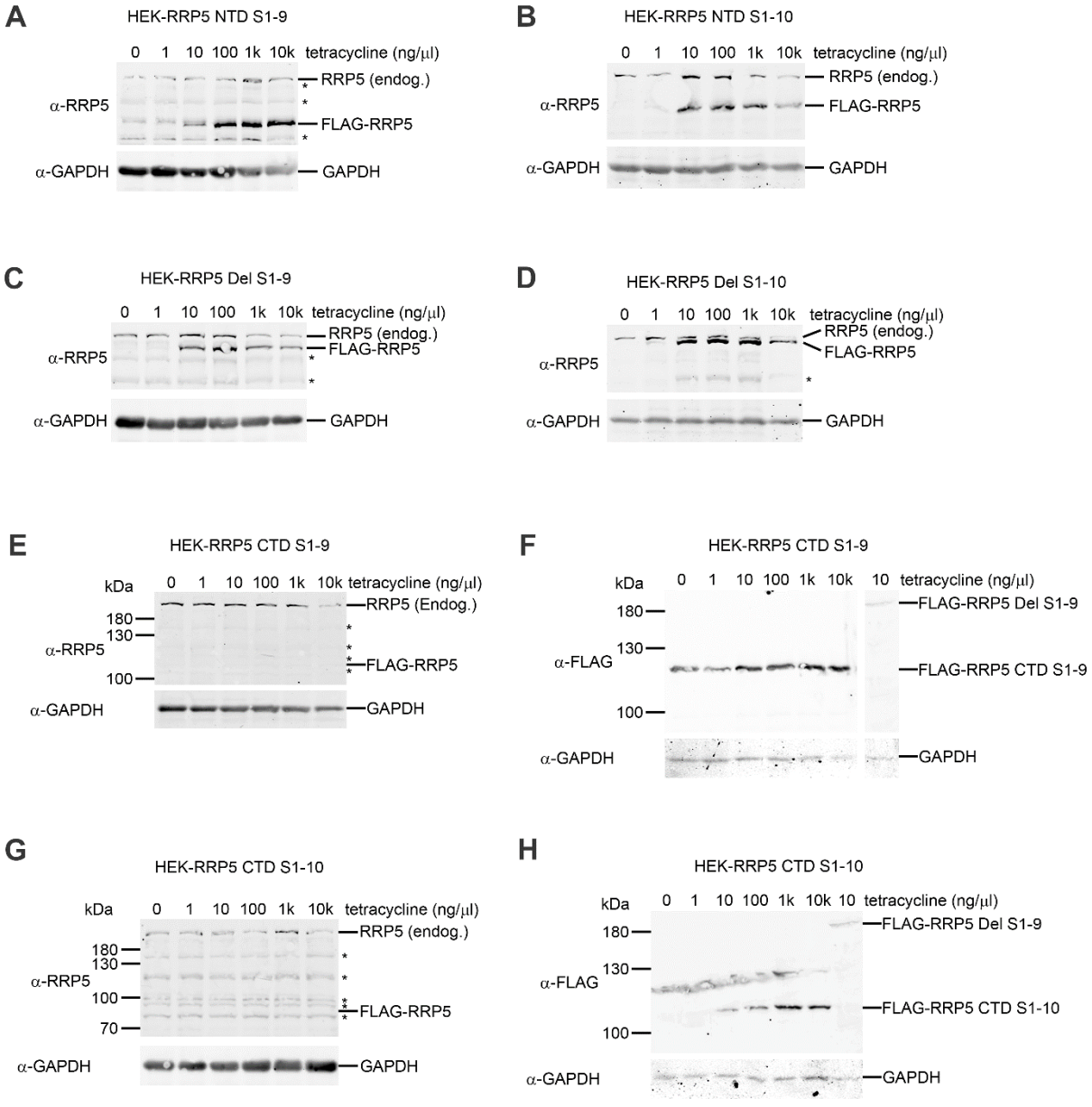


Figure 5.20 Expression of RRP5 mutant proteins in stably transfected HEK293T cells. HEK293T Flp-In cells stably transfected with FLAG-tagged mutant RRP5 proteins were incubated with a range of concentrations of tetracycline (0-10,000 ng/μl). Protein was separated by SDS-PAGE and analysed by western blotting using **(A to E)** an anti-RRP5 antibody to detect both endogenous and FLAG-tagged RRP5. An anti-Karyopherin antibody was used as a loading control. The peptide used to produce the anti-RRP5 antibody is located at the N-terminus of the RRP5 protein, and therefore does not recognise C-terminal halves (RRP5 CTD S1-9 and RRP5 CTD S1-10) (panels

E and G). **(F and H)** An anti-FLAG antibody was used to detect the expression of these mutants. Panels F and H include a protein sample from cells expressing a FLAG-tagged RRP5 Del S1-9 mutant protein expressed at similar levels to endogenous RRP5.

The available anti-RRP5 antibody was raised against a peptide in the N-terminal domain of the protein, so was unable to detect C-terminal truncations, CTD S1-9 and CTD S1-10. However, western blotting using an anti-FLAG antibody shows that these truncations were efficiently expressed and responded to tetracycline (5.20F and 5.20H). An RRP5 Del S1-9 sample treated with 10 ng/µl tetracycline was also separated on the same gel as C-terminal truncations, and blotted with an anti-FLAG antibody, as this mutant was expressed at similar levels to endogenous RRP5. Therefore, comparison of levels of this mutant to the C-terminal truncations allowed me to determine the tetracycline concentration required to express C-terminal truncations at similar levels to endogenous RRP5. The confirmation that RRP5 mutant proteins express efficiently will allow for analysis of these mutants in relation to pre-rRNA processing, and determination of the importance of different domains of RRP5 in different pre-rRNA processing steps.

5.3 Discussion

Ribosome biogenesis is linked to cell growth and proliferation, is up-regulated in most cancers and is defective in at least 20 diseases called ribosomopathies (Pelava *et al.*, 2016). Ribosomopathy patients, who possess defects in either ribosomal proteins or proteins important for ribosome production, are often predisposed to cancer. The main focus of this PhD project has been in the characterisation of candidate endonucleases required for cleavage in the maturation of the 18S rRNA. One candidate endonuclease, the PIN domain protein UTP23, was found to be often mutated in colorectal cancer patients (Timofeeva *et al.*, 2015). UTP23, as part of the SSU processome, is required for early SSU cleavages at sites A0, 1 and 2a and may be an active endonuclease responsible for cleavage at A0 (Wells *et al.*, 2017). The residue mutated in colorectal cancer, P215, is not within the PIN domain or any of the key Zinc finger residues important for RNA binding. Therefore, it was unclear what effect, if any, this substitution has on pre-rRNA processing.

An RNAi-rescue system was established using HEK293T cells, and this showed that the P215Q mutation did not result in a pre-rRNA processing defect when endogenous UTP23 was depleted. However, when this mutant was expressed at levels much higher than endogenous UTP23, accumulation of the 30S pre-rRNA intermediate was observed, indicating disrupted cleavage at sites A0, 1 and 2a. This is the same phenotype observed upon UTP23 depletion in control cells, and also with mutation of key residues in either the PIN domain or Zinc finger motif (Chapter Four). It is therefore difficult to clarify the specific pre-rRNA processing defect caused by the P215Q mutation. Immunofluorescence determined that the P215Q mutant localises correctly to the nucleolus, irrespective of treatment with low or high concentrations of tetracycline, suggesting that the incorporation of UTP23 into the SSU processome is unaffected by this substitution. However, the P215Q mutant showed increased association with the U17 snoRNA compared to WT UTP23 in immunoprecipitation experiments. Yeast Utp23 is essential for the release of snR30 (analogous to U17) from pre-ribosomes (Hoareau-Aveilla *et al.*, 2012). Utp23/UTP23 interacts directly with snR30/U17 and appears to play a key role in snR30/U17-mediated interactions in the ES6 region of the 18S rRNA. Therefore, the apparently increased association of the UTP23 P215Q mutant with the U17 snoRNA in our immunoprecipitation experiments could indicate that U17 release from pre-ribosomes and/or the recruitment or interactions of other ribosome biogenesis factors required for SSU pre-rRNA processing may be impaired. This would explain the pre-rRNA processing phenotype observed when this mutant is present in large quantities, which may contribute to cancer. However, i*n vitro* RNA binding assays indicated that the UTP23 P215Q mutant directly binds to the U17 snoRNA with similar efficiency to WT UTP23. Therefore, if this mutant does associate more strongly with U17 *in vivo*, it is likely due to increased association with pre-ribosomal particles, rather than altered RNA binding. Further experiments are required to clarify the effects observed in these preliminary results, such as gradient analysis to determine the association of this mutant and of the U17 snoRNA with different pre-ribosomal complexes.

The link between cancer and ribosome biogenesis appears to be complex, as tumours often show upregulation of ribosome biogenesis, whereas defects in ribosome biogenesis lead to the activation of the tumour suppressor p53. Indeed, p53 is important for many clinical phenotypes in ribosomopathy patients. It is possible that defects in ribosome biogenesis, and the subsequent p53 induction, lead to

desensitisation of cells to p53 or pressure to disrupt or mutate the p53 pathway. The p53 pathway is linked to ribosome biogenesis via the 5S RNP, comprising the LSU ribosomal proteins RPL5 and RPL11 and the 5S rRNA. This complex binds and inhibits the inhibitor of p53, MDM2, leading to activation of p53 (Bursac *et al.*, 2012; Donati *et al.*, 2013). The interaction between the 5S RNP and MDM2 is enhanced when ribosome production is defective, as these defects lead to an accumulation of the free 5S RNP complex (Sloan *et al.*, 2013a). Ribosome biogenesis defects in both the SSU and LSU can cause activation of p53 via the 5S RNP (Pelava, 2016). This chapter contains preliminary data revealing the effect of depleting SSU ribosome biogenesis factors on p53 induction.

Depletion of UTP23 in U2OS and MCF7 cells caused a mild induction of p53, with an average 1.5- to 1.8-fold increase in p53 levels, compared to the stronger p53 induction (3- to 4-fold increase) upon RPL7 depletion. It is unclear why depletion of RPL7 causes an enhanced p53 induction. It is possible that depletion of LSU factors causes additional accumulation of the 5S RNP, as in addition to ribosome biogenesis defects, disruption of LSU biogenesis may impact the association of the 5S RNP with pre-ribosomes. The mild increase in p53 levels upon UTP23 depletion in U2OS cells suggests that the potential link between UTP23 and cancer may be p53-dependent. It is currently uncertain whether this p53 induction is dependent on the 5S RNP. Further experiments will determine if the p53 induction observed upon UTP23 depletion is rescued by depleting 5S RNP components. The fact that depletion of two other candidate endonucleases, the PIN domain endonuclease UTP24, which cleaves at sites 1 and 2a, and RCL1, did not affect p53 levels suggest that the p53 induction is specific to UTP23. It is unclear why these factors would differ in relation to p53 induction, as they are all required for the same three early pre-SSU cleavages and are all components of the SSU processome. However, in yeast, Utp23 associates with earlier pre-ribosomal complexes, and is not present in complexes after A_0 cleavage but prior to cleavage at site A_1, while Utp24 and Rcl1 are present later (Barandun *et al.*, 2017; Cheng *et al.*, 2017). Therefore, it is possible that earlier impairment of SSU processome assembly induces p53, while later processome formation does not. Further repeat experiments in a range of human cell lines possessing an active p53 gene are required to confirm the effect on p53 regulation of depleting SSU candidate endonucleases. In addition, analysis of the regulation of p53 targets such as p21 and

PUMA will reveal whether p53 activity is altered upon depletion of candidate endonucleases, despite mild or no increase in p53 protein levels.

Surprisingly, expression of the P215Q mutant form of UTP23 found in a colorectal cancer patient did not appear to cause induction of p53 in U2OS cells, even when overexpressed. This suggests that any effects of this mutation in colorectal cancer may not be p53-dependent. However, expression of the P215Q mutant at levels similar to endogenous UTP23 was able to rescue the p53 induction caused by UTP23 depletion, while expression of WT UTP23 at these levels was insufficient. In order to understand the difference between WT UTP23 and the P215Q mutant, pre-rRNA processing needs to be analysed in U2OS cells. Interestingly, the genome of the U2OS cell line contains a point mutation in the helicase ROK1, which interacts with UTP23 *in vitro*. In yeast, Utp23 interacts with Rok1 (Wells *et al.*, 2017) and both factors are required for the release of the snR30 snoRNA from pre-ribosomes (Bohnsack *et al.*, 2008; Hoareau-Aveilla *et al.*, 2012). It is unknown whether this mutation affects ROK1 function or U17 release, but a different cell line may be more appropriate for analysis of UTP23 function. Since HEK293T cells do not contain an active p53 gene, and the MCF7 cell line appears to have other defects in ribosome biogenesis, the human colon carcinoma cell line, HCT 116, may be appropriate for studying pre-rRNA processing defects and p53 induction upon expression of the UTP23 P215Q mutant. Several other UTP23 mutations have been found in tumours and cancer cells, and some of these, like the P215Q mutation, are located in the C-terminal tail of the protein (Iorio *et al.*, 2016). Further work will analyse the effects of these mutations on pre-rRNA processing and p53 induction.

Several other ribosome biogenesis factors important for the maturation of the 18S rRNA have apparent links to cancer. Interestingly, multiple factors involved in the processing of ITS1 and the 3' end of 18S are defective in cancer. While ribosome biogenesis is largely conserved in eukaryotes, the major mechanism of ITS1 processing differs in higher eukaryotes (Sloan *et al.*, 2013b). While the major cleavage event in ITS1 in yeast occurs at site A_2 (equivalent to human 2a), the major human ITS1 cleavage occurs at site 2 (equivalent to yeast A_3), followed by exonucleolytic trimming towards the 2a site. Cleavage at the 2a site in human cells only occurs in a minor processing pathway, and it is possible that defects in factors involved in ITS1 processing may cause a shift to the minor, "yeast-like" pathway. A recent exome sequencing study identified many ribosomal proteins and ribosome biogenesis factors

that are frequently mutated in many cancers, including factors involved in ITS1 processing (Iorio *et al.*, 2016). These include the 3' to 5' exonuclease, RRP6, which is required in ITS1 processing for the trimming towards the 2a site following cleavage at site 2, and the 5' to 3' exonuclease XRN2. Depletion of XRN2 causes a shift to the minor pre-rRNA processing pathway, causing reliance on cleavage at site 2.

The preliminary data presented in this chapter suggests that depletion of neither exonuclease leads to p53 induction in human cells. Depletion of RRP6 appeared to cause a mild induction of p53 in U2OS and MCF7 cells, although this was not significant in either cell line, and showed high variability between experiments. Further experiments, using a range of cell lines, are required to confirm the potential p53 induction upon RRP6 depletion. Surprisingly, depletion of XRN2 did not cause accumulation of p53 in U2OS or MCF7 cells. Depletion of XRN2 in MCF7 cells caused a pre-rRNA processing phenotype unlike that observed in other cells, suggesting these cells may be naturally defective in pre-rRNA processing. Depletion of XRN2 in other cell lines containing an active p53 gene will help to clarify the link between XRN2 and the p53 pathway. XRN2 is frequently mutated in cancer (Iorio *et al.*, 2016), and so while XRN2 depletion does not appear to cause p53 induction, it would be interesting to analyse p53 levels upon expression of XRN2 containing these cancer-associated mutations. As mentioned above for candidate endonucleases, further work would analyse p53 activity upon RRP6 and XRN2 depletion by detection of levels of p53 downstream targets.

Another ribosome biogenesis factor important for ITS1 processing is RRP5. In yeast, Rrp5 is important for both cleavages (at sites A_2 and A_3) in ITS1, and this role seems to be conserved in HeLa cells (Sloan *et al.*, 2013b). Depletion of RRP5 caused a moderate increase in p53 levels in both U2OS and MCF7 cells and this p53 induction was dependent on the 5S RNP. Further characterisation of the role of RRP5 on human pre-rRNA processing showed that RRP5 depletion led to accumulation of 3'-extended forms of the 30S and 21S pre-rRNA intermediates and a decrease in 18SE pre-rRNA levels in HEK293T cells. This is consistent with data from HeLa cells and suggests that RRP5 is important for cleavages at sites 2a and 2 in HEK293T cells. In human cells, the major ITS1 cleavage occurs at site 2, and depletion of XRN2 shifts processing to the "minor" pathway, in which cleavage at site 2a is favoured. This causes accumulation of "2a-specific" precursors, the 36S pre-rRNA and the ITS1 fragment. Depletion of RRP5 in the absence of XRN2 in HEK293T cells, caused a reduction in

levels of the 36S pre-rRNA but, unlike in HeLa cells, levels of the ITS1 fragment were not reduced. One possible explanation for this finding is that RRP5 depletion in HEK293T cells affects cleavage at site 2a more severely than site 2, meaning that the 2a site is not cleaved before site 2 even through the minor processing pathway.

It is not clear why the effect of RRP5 depletion on pre-rRNA processing would differ between HEK293T and HeLa cells, although difference between these two cell lines is much less pronounced than that seen in MCF7 cells. The pre-rRNA processing phenotypes observed in MCF7 cells were strikingly different for both RRP5 and XRN2 depletion. Depletion of RRP5 did not result in observable accumulation of 30SL3' or 21SL3', and instead appeared to cause accumulation of the 30S pre-rRNA, suggesting that cleavages at sites A0, 1 and 2a were disrupted. Depletion of XRN2 did not cause accumulation of the "2a-specific" 36S pre-rRNA or ITS1 fragment, suggesting that cleavage at the 2a site may be defective in these cells. Interestingly, it has previously been observed that the 30SL5' precursor is naturally present in MCF7 cells (Sloan *et al.*, 2014), suggesting that cleavage at the A' site may be impaired in these cells. XRN2 is important for cleavage at A', therefore depletion of XRN2 may not cause further clearly observable pre-rRNA processing defects. This would potentially explain the less severe pre-rRNA processing phenotype and lack of p53 induction upon XRN2 depletion in these cells. Further experiments using a range of different human cell types are required to confirm the role of RRP5 in human pre-rRNA processing.

In yeast, the function of Rrp5 in processing of ITS1 can be physically separated between the multiple RNA-binding and protein-protein interaction domains. The C-terminal half of Rrp5 (comprised of 9 of the 12 S1 RNA-binding domains) is important for processing of sites A_0, A_1 and A_2 (A0, 1 and 2a in humans) while the N-terminal half is important for cleavage at A_3 (Eppens *et al.*, 1999; Lebaron *et al.*, 2013). The final three S1 domains are important for cleavage at site A_2. Therefore, it appears likely that Rrp5 plays a role in the coordination of the predominant cleavage event within ITS1. While RRP5 is required for both human ITS1 cleavages, it is unclear whether the functional distinction of the various domains of the human protein is similar to yeast Rrp5. An RNAi-rescue system was established in order to examine the roles of human RRP5 protein domains, and expression of RNAi-resistant WT RRP5 was able to rescue the pre-rRNA processing phenotype caused by depletion of RRP5, meaning that this system is appropriate for testing pre-rRNA processing upon expression of RRP5 truncations and mutants. Stable HEK293T cell lines were generated expressing N- and

C-terminal halves of RRP5 as well as RRP5 with the last three or four C-terminal S1 domains deleted. Expression of all RRP5 mutant proteins was efficiently induced in stable HEK293T cell lines upon tetracycline treatment. Due to time constraints the effects of pre-rRNA processing upon expression of RRP5 mutants were not tested.

Future work would build on the preliminary data presented here and focus on testing the pre-rRNA processing defects caused by expression of RRP5 mutants from stable cell lines when endogenous RRP5 is depleted. This would illuminate the separate functions of the distinct domains of RRP5 and determine whether the functions confirmed in yeast are conserved in human cells. *In vivo* UV crosslinking would then be performed using these stable cell lines to identify the physical interactions between the different parts of the RRP5 protein and the different regions of ITS1. It will be interesting to see whether the interaction of the different domains with the pre-rRNA is altered upon XRN2 depletion and the shift to the “minor” pre-rRNA processing pathway. If RRP5 is important for the coordination of this switch to cleavage at site 2a, and potentially the recruitment of the 2a endonuclease, UTP24, to this site, it would be expected that RRP5 would crosslink close to this site upon XRN2 depletion. Finally, it would be interesting to see if the recruitment of WT and PIN domain mutant forms of UTP24 is altered in the “minor” pathway, and if this correlates with RRP5 interaction with the pre-rRNA.

Chapter Six
Discussion

6.1 Overview

Ribosome biogenesis is a complex process, which requires a large proportion of the cellular energy and involves many protein and RNP complexes. A key aspect of the production of eukaryotic ribosomes is the processing of the rRNA precursors, which is initiated by endonucleolytic cleavages at multiple pre-rRNA sites. While ribosome biogenesis has been extensively studied in *S. cerevisiae*, the endonucleases responsible for cleavage at multiple sites remain elusive or controversial. In humans, much less is known about the precise roles of many ribosome biogenesis factors in general, and pre-rRNA processing endonucleases specifically. Prior to the start of this project, the enzymes responsible for three early cleavages in the maturation of the 18S rRNA of the SSU in yeast were either unknown or suggested but not confirmed in yeast. In human cells, the enzymes responsible for processing at the equivalent three sites, in addition to the extra SSU cleavage site specific to higher eukaryotes (A'), had not been identified. While the human homologues of the yeast candidate endonucleases have previously been shown to be required for human pre-rRNA processing, detailed study on their specific roles had not been undertaken.

This study aimed to identify the endonucleases responsible for early cleavages in 18S rRNA maturation in human cells, through the analysis of human homologues of yeast candidate endonucleases. Alongside parallel work conducted in our lab using *S. cerevisiae* (Wells *et al.*, 2017), the data from this study provides support for a conserved role for Utp24 in cleavage at sites A_1 and A_2 in yeast, and the equivalent sites 1 and 2a in human cells. Our data also argues against the proposed enzymatic role of Rcl1 in cleavage at yeast site A_2. In addition, this study shows that, unlike in yeast, an intact PIN endonuclease domain in human UTP23 is important for early SSU pre-rRNA cleavages, suggesting a potential role for UTP23 as an endonuclease at human site A0. This possibility highlights a potential major distinction between yeast and human pre-rRNA processing, emphasising the importance of studying ribosome biogenesis in human cells.

The study of human ribosome biogenesis is also significant in regard to the known links between ribosome biogenesis and human health and disease. As a key

determinant of cellular growth rate, ribosome biogenesis is upregulated in cancer cells, and defects in ribosome production are associated with numerous diseases, known as ribosomopathies. Surprisingly, given the upregulation of ribosome biogenesis during tumourigenesis, defective ribosome biogenesis often causes stabilisation and activation of the tumour suppressor p53. The link between ribosome biogenesis and p53 activation is mediated by the LSU biogenesis intermediate, the 5S RNP. However, defects in SSU biogenesis can also lead to 5S RNP-dependent p53 induction. Multiple factors important for SSU biogenesis have been found to be mutated in tumours and/or cancer cell lines, including the putative endonuclease UTP23 (Timofeeva *et al.*, 2015). In this study, it has been shown that depletion of UTP23, but not other candidate endonucleases, UTP24 and RCL1, causes mild induction of p53 in two human cancer cell lines. In addition, overexpression of a cancer-associated UTP23 mutant mildly disrupts pre-rRNA processing, and this mutant may associate more strongly with the U17 snoRNA *in vivo* but does not induce p53 in human cells.

Other SSU biogenesis factors frequently mutated in cancer include the exonucleases XRN2 and RRP6, and the RNA binding protein RRP5 (Iorio *et al.*, 2016). All three of these factors are also important for LSU biogenesis, and XRN2 and RRP6 also participate in RNA degradation and processing more broadly. Surprisingly, this study found that depletion of XRN2 or RRP6 in the two cancer cell lines used did not cause significant induction of p53. In contrast, depletion of RRP5 caused a strong p53 accumulation in both cell lines, and this increase in p53 protein levels was dependent on the 5S RNP. In yeast, the multiple domains of Rrp5 can be functionally separated in regard to their requirement in distinct pre-rRNA processing steps, but while the human protein is broadly conserved, it is unknown if this separation is conserved for human RRP5. Data from this study shows that, consistent with previously published data, RRP5 is important for the three early pre-SSU cleavages, as well as the major ITS1 cleavage in human cells. The pre-rRNA processing phenotype caused by depletion of RRP5 could be efficiently rescued by expression of FLAG-tagged full-length, wild type RRP5, and RRP5 mutants were designed and expressed in human cells, allowing for future analysis of the specific roles of the different RRP5 domains in human ribosome biogenesis.

6.2 Endonucleases responsible for early pre-18S rRNA cleavages

6.2.1 The roles of UTP24 and RCL1 in pre-rRNA processing at sites 1 and 2a

Maturation of the yeast 18S rRNA requires three early cleavage events, which take place at specific sites around the 18S sequence on the nascent pre-rRNA. These three cleavages are coupled by their requirement of the assembly of a large RNP complex called the SSU processome. The composition and assembly of the yeast SSU processome have been well studied (Chaker-Margot *et al.*, 2015; Zhang *et al.*, 2016b; Barandun *et al.*, 2017; Cheng *et al.*, 2017), but the endonucleases within the SSU processome that mediate the three early cleavages were either unknown or their identity had not been confirmed at the start of this PhD project. Two SSU processome components, Utp24 and Rcl1, have previously been implicated in catalysing cleavages at sites A_1 and/or A_2 in yeast. Utp24 contains a PIN endonuclease domain, and the predicted catalytic site residues of its PIN domain were shown to be essential for cleavage at sites A_1 and A_2 (Bleichert *et al.*, 2006). Rcl1 is an RNA cyclase-like protein that was reported to cleave at site A_2 *in vitro*, with three residues in the C-terminal region of the protein shown to be important for all three early cleavages *in vivo* (Horn *et al.*, 2011). These residues were predicted to be important for substrate binding, but no catalytic active site residues have so far been identified in the Rcl1 protein (Horn *et al.*, 2011; Tanaka *et al.*, 2011).

Work performed in yeast in parallel to this PhD project revealed that yeast Utp24 directly contacts the pre-rRNA adjacent to the A_1 cleavage site, supporting a direct role in cleavage at this site (Wells *et al.*, 2016). The position of Utp24 close to site A_1 was recently confirmed in published Cryo EM structures of the SSU processome captured prior to A_1 cleavage (Barandun *et al.*, 2017; Cheng *et al.*, 2017). It was shown that recombinant Utp24 cleaved a pre-rRNA mimic containing the yeast A_2 site *in vitro*, while neither a predicted inactive Utp24 PIN domain mutant nor Rcl1 showed observable cleavage activity (Wells *et al.*, 2016). *In vitro* nuclease assays were also repeated using the same conditions as those used in Horn *et al*., 2011, and these indeed showed the previously reported cleavage activity of Rcl1, but also showed equivalent activity for both WT Utp24 and PIN domain mutants of Utp24, suggesting that this cleavage is non-specific (Wells *et al.*, 2016). These results provided considerable evidence implicating Utp24, via its conserved PIN domain, in catalysis of endonucleolytic cleavage at sites A_1 and A_2 in yeast. In the same study, it was also

shown that mutation of the three residues (RDK) in the C-terminus of Rcl1 that disrupted *in vivo* pre-rRNA processing (Horn *et al.*, 2011), inhibited the interaction between Rcl1 and its binding partner Bms1 *in vitro*, as well as Rcl1's association with pre-rRNA intermediates and the U3 snoRNA (Wells *et al.*, 2016). Consistent with this, these residues directly contact Bms1 in the recently published Cryo-EM structure of the SSU processome (Barandun *et al.*, 2017). Data from this project further showed that, consistent with the reported reliance on the Rcl1-Bms1 interaction for Rcl1 nuclear import (Delprato *et al.*, 2014), the Rcl1 RDK mutant localises to the cytoplasm in yeast cells (Wells *et al.*, 2016).

The yeast data therefore suggest that Utp24 is the enzyme responsible for cleavage at sites A_1 and A_2, while Rcl1 seems unlikely to be an active endonuclease in pre-rRNA processing. However, it remains possible that Rcl1 does indeed cleave at site A_2, and that inhibition of A_2 cleavage upon expression of an inactive Utp24 PIN domain mutant is an indirect consequence of prior disruption of A_1 cleavage. Indeed, no pre-rRNA processing intermediates cleaved at A_2 but not A_1 have so far been identified, suggesting that the order of these two processing events cannot be altered. However, *in vitro* cleavage assays show that Utp24 can cleave at the A_2 site. This leaves open a further alternative possibility whereby both Utp24 and Rcl1 can cleave at this site, perhaps in different *in vivo* contexts. In yeast, early pre-40S cleavages can occur either co-transcriptionally on the nascent pre-rRNA, or post-transcriptionally after release of the primary 35S precursor (Kos and Tollervey, 2010). Therefore, it is possible that cleavage at site A_2 is mediated by different enzymes in these two different contexts. However, the confirmation of Rcl1 as a *bona fide* endonuclease requires further study, particularly in the identification of catalytic site residues required for its putative cleavage activity.

The human homologues of yeast ribosome biogenesis factors are generally well conserved but have been studied in much less detail. Therefore, elucidating the functions of human factors can be helpful in identifying and confirming the roles of homologous factors in other systems. UTP24 is well conserved in higher eukaryotes and contains a conserved PIN domain. Consistent with yeast data, UTP24 is required for early pre-SSU cleavages at three sites, called A0, 1 and 2a (equivalent to yeast sites A_0, A_1 and A_2) in HeLa cells (Sloan *et al.*, 2013b; Tafforeau *et al.*, 2013; Sloan *et al.*, 2014) and mouse cells (Wang *et al.*, 2014). Human RCL1 is also required for these early cleavages in both HeLa and mouse cells (Sloan *et al.*, 2013b; Tafforeau *et al.*,

2013; Sloan *et al.*, 2014; Wang *et al.*, 2014). While the presence of both proteins is required for all three early cleavages, the importance of specific regions and residues of human UTP24 and RCL1 had not previously been investigated.

Data from this project has provided evidence supporting a conserved role for human UTP24 in cleavage, via its PIN endonuclease domain, at sites 1 and 2a, consistent with its proposed equivalent function at yeast sites A_1 and A_2 (Wells *et al.*, 2016). As in yeast, cleavage at both sites requires an intact UTP24 PIN domain *in vivo*. The pre-rRNA processing defect seen upon expression of a predicted inactive UTP24 mutant, in the absence of endogenous UTP24, was observed in cells utilising both major and minor human pre-rRNA processing pathways, showing that the PIN domain of UTP24 is required for these cleavages in either pathway. During this project, UTP24 was independently reported as the enzyme responsible for cleavage at site 1 in humans (Tomecki *et al.*, 2015).

Although data in this study suggests that UTP24 is the enzyme responsible for cleavage at sites 1 and 2a in humans, further work is required to confirm this. Parallel work from our lab tested the nuclease activity of human UTP24 on an RNA substrate containing the yeast site A_2, and the human protein cleaved at the A_2 site and at other sites around A_2, in a pattern identical to the yeast protein (Wells *et al.*, 2016). Cleavage assays with UTP24 using RNA substrates containing human cleavage sites 1 and 2a have so not been possible, due to the high GC content of human ITS1, making its amplification difficult. In addition, cleavage by human UTP24 at the 2a site may require co-factors *in vivo* that are not present in *in vitro* assays. This was recently discovered to be the case for the site 3 endonuclease, NOB1, which required the presence and activity of an ATPase co-factor for accurate cleavage of an 18SE pre-rRNA mimic (Bai *et al.*, 2016).

As in yeast, RCL1 was required for early pre-rRNA processing steps in all human cell lines tested, while expression of an RCL1 C-terminal truncation in HEK293T cells resulted in disruption of all three cleavages, indicating that this region is important for the essential function of Rcl1 in pre-40S processing in yeast and humans (Horn *et al.*, 2011). Mutation of the equivalent three RCL1 residues that are important for the nuclear localisation and SSU processome incorporation of yeast Rcl1 (Wells *et al.*, 2016) did not affect human pre-rRNA processing, suggesting that the

nuclear import of human RCL1, or its association with the SSU processome and/or the pre-rRNA occurs by a distinct mechanism in human cells.

A possible explanation for the dispensability of human RCL1 RHK residues, but not yeast RDK residues, relates to the importance of the interaction of yeast Rcl1 with Bms1 for its nuclear import and association with the SSU processome. In yeast, Rcl1 and Bms1 associate with pre-ribosomes at around the same time during SSU processome assembly (Delprato *et al.*, 2014; Chaker-Margot *et al.*, 2015; Zhang *et al.*, 2016b), and the import of Rcl1 into the nucleus depends on its interaction with Bms1 (Delprato *et al.*, 2014). In humans, while the interaction between the two proteins is conserved (Wang *et al.*, 2016) and likely important for pre-rRNA processing within the SSU processome, RCL1 appears to associate with the SSU processome before BMS1 does (Turner *et al.*, 2012), suggesting that this interaction is not required for the nuclear import of RCL1. The lack of a pre-rRNA processing defect observed upon expression of the human RCL1 RHK mutant in this study therefore suggests that these residues are not important for the RCL1-BMS1 interaction. While early yeast pre-rRNA processing steps predominantly occur co-transcriptionally, human processing appears to be mainly post-transcriptional. Yeast Rcl1 may indeed mediate co-transcriptional cleavage at site A_2, with Utp24 cleaving post-transcriptionally at the same site (Horn *et al.*, 2011). Cleavage at the equivalent human site, 2a, appears to only occur post-transcriptionally, and therefore may not require any potential nuclease activity of human RCL1. Additionally, while cleavage at yeast site A_2 is the predominant ITS1 cleavage, human processing does not rely on cleavage at 2a under normal conditions. Therefore, having multiple potential enzymes that can cleave at this site would probably be unnecessary.

Recent structures of the SSU processome captured prior to cleavage at site A_1 have provided support for the role of Utp24 in direct cleavage at this site (Barandun *et al.*, 2017; Cheng *et al.*, 2017). As disruption of Utp24's PIN domain active site inhibits cleavage at both sites A_1 and A_2 (Bleichert *et al.*, 2006), it is impossible to uncouple these cleavages. Therefore, a yeast SSU processome structure captured after cleavage at site A_1, and prior to A_2 cleavage, will help to confirm the site A_2 nuclease. Structures of the human SSU processome captured both before site 1 cleavage and before site 2a cleavage would also be useful in confirming the endonuclease(s) responsible for cleavage at these sites in human cells.

6.2.2 The role of UTP23 in human ribosome biogenesis

While Utp24/UTP24 appears to be responsible for cleavages at sites A_1/1 and A_2/2a, the enzyme catalysing the first of the three coupled cleavages, at site A_0/A0, is unknown in both yeast and humans. As for A1/1 and A_2/2a cleavages, cleavage at A_0/A0 requires the SSU processome. Aside from Utp24 and Rcl1, there is one other SSU processome component with a known link to endonucleases, the PIN domain protein Utp23. In yeast, Utp23 contains a degenerate PIN domain, with only two of the four putative catalytic acidic residues present, suggesting it does not possess endonuclease activity (Lu *et al.*, 2013). Accordingly, mutation of the remaining two conserved acidic residues had no impact on yeast pre-rRNA processing *in vivo* (Bleichert *et al.*, 2006; Lu *et al.*, 2013). Utp23 was instead proposed to play a non-enzymatic role in ribosome biogenesis, ruling it out as a candidate for cleavage at A_0. Utp23 was shown to associate with the snR30 snoRNA and to be required for the release of snR30 from pre-ribosomes (Hoareau-Aveilla *et al.*, 2012). Human UTP23 also associates with the human homologue of snR30, the U17 snoRNA (Hoareau-Aveilla *et al.*, 2012). Parallel work from our lab has shown that Utp23 directly contacts the pre-rRNA within the 18S rRNA sequence in the ES6 region, which contains snR30 base-pairing sequences (Wells *et al.*, 2017). This work also showed that Utp23 interacts with other factors that bind to this region or to snR30 (Wells *et al.*, 2017). In this PhD study, it has been shown that Utp23 directly binds snR30 *in vitro*, and that this interaction, as well as interactions with ES6- and snR30-interacting factors, is conserved for human UTP23. Together, these results argue for a conserved, non-enzymatic role for Utp23/UTP23, potentially in coordinating the association and release of ribosome biogenesis factors around the ES6 region.

In contrast to the yeast protein, the PIN domain of human UTP23 contains three of the four putative catalytic acidic residues, suggesting it may possess the ability to cleave RNA. UTP23 is a conserved component of the SSU processome in humans and is required for all three early pre-40S cleavages in HeLa cells (Turner *et al.*, 2012; Tafforeau *et al.*, 2013; Sloan *et al.*, 2014). Data from this project showed that UTP23 is also required for these three cleavages in HEK293T cells. However, pre-rRNA processing phenotypes upon UTP23 depletion in U2OS and MCF7 cells were very variable, with some experiments surprisingly showing mild accumulation of the final 18S precursor, 18SE, indicating disruption of a later stage of processing. This is consistent with the phenotype seen upon UTP23 depletion in mouse cells (Wang *et*

al., 2014), and suggests a significantly distinct role compared to that shown for the yeast protein, and to the expected role of an SSU processome component. It is unclear why depletion of UTP23 effects pre-rRNA processing so differently in these cell lines and in mice cells, and further work is required in other cell lines to confirm the overall role of UTP23 in human ribosome biogenesis. The localisation of UTP23 and its interactions with other ribosome biogenesis factors should also be investigated in these cell lines. One possibility for these differences in pre-rRNA processing is that these cell lines may have natural defects in particular stages of ribosome biogenesis. For example, the helicase ROK1/DDX52, a known UTP23- and U17 snoRNA-interacting factor, contains a point mutation in the genome of the U2OS cell line (Tym *et al.*, 2016). Therefore, this could alter the pre-rRNA processing pathway in these cells, compared to that in other cell lines. However, as the phenotype observed upon UTP23 depletion in HEK293T cells was consistent with data from HeLa and yeast, the HEK293T cell line was used for further experiments presented in this study.

While the effect of depleting UTP23 in human cells has been analysed, the importance of the putative catalytic acidic residues in the PIN domain of UTP23 has not previously been investigated. Using an RNAi rescue system in HEK293T cells, this study has shown that an intact PIN domain in UTP23 is important for processing of the pre-rRNA at cleavage sites A0, 1 and 2a. This is in contrast to yeast Utp23 and suggests that the potential enzymatic function of the PIN domain is required for pre-rRNA processing in humans. UTP23 also contains a Zinc finger RNA binding motif, which was also shown to be important for processing at these three sites. The observation that expression of either UTP23 mutant resulted in the same pre-rRNA processing as depletion of UTP23 suggested a potential defect in the general function of the mutant protein as an SSU processome component. However, the nucleolar localisation was unchanged for both mutant proteins.

In yeast, Utp23 is required for the release of snR30 from pre-ribosomes, with Utp23 depletion leading to an increase of snR30 in larger complexes (Hoareau-Aveilla *et al.*, 2012), so it may be expected that disruption of UTP23's RNA binding ability through mutation of the Zinc finger motif would result in a similar increase in the levels of U17 in larger complexes. However, it has also been suggested that yeast Utp23 may be important in establishing initial base pairing between snR30 and the pre-rRNA (Hoareau-Aveilla *et al.*, 2012). If the Zinc finger motif of human UTP23 is indeed important for establishing base pairing between U17 and the pre-rRNA, the motif may

confer the RNA binding ability of UTP23 to either U17 or the pre-rRNA. Indeed, data from this project shows that the UTP23 Zinc finger motif mutant protein does not associate with U17 above background levels *in vivo* and shows reduced binding *in vitro*. This suggests that the Zinc finger motif is at least partly responsible for the binding of UTP23 to the U17 snoRNA. The more severe effect on *in vivo* U17 association compared to *in vitro* U17 binding with mutation of the Zinc finger motif suggests that the direct binding of UTP23 to U17 via this motif is not sufficient for *in vivo* association. It is possible that the Zinc finger motif is also important for UTP23 binding to the pre-rRNA, and this binding may be important for association of UTP23 with U17. Binding of UTP23 to the pre-rRNA was not tested in this study, and future work will aim to elucidate the importance of the Zinc finger motif in the putative pre-rRNA-binding of UTP23.

In contrast, the PIN domain mutant protein showed reduced *in vivo* U17 association, but no reduction in *in vitro* U17 binding. The reduction in U17 co-precipitation with the PIN domain mutant appears to be independent of UTP23 directly binding to the snoRNA, as *in vitro* binding is unaffected. It is unclear why this mutant may associate less efficiently with U17 *in vivo*. One possibility is that mutation of the PIN domain abolishes endonuclease cleavage by UTP23, causing UTP23 to become stalled at the cleavage site, where it is unable to associate with U17.

Disruption of all three early cleavages, as seen upon expression of the UTP23 putative inactive PIN domain mutant, in the absence of endogenous UTP23, is consistent with a block in cleavage at site A0, followed by a subsequent and indirect lack of cleavages at sites 1 and 2a, as these three sites are coupled and occur in a fixed order. The fact that mutation of the PIN domain does not affect expression, localisation or *in vitro* U17 snoRNA binding of UTP23 suggests that the mutant protein is generally functional, and that the pre-rRNA processing defect observed is specific to this point mutation. Therefore, these data raise the exciting possibility that UTP23 may be the endonuclease responsible for cleavage at the A0 site in human cells. Attempts to confirm the endonuclease activity of UTP23 at this site *in vitro*, using recombinant N-terminal GST-tagged UTP23, were unsuccessful. It is possible that the position of the protein tag at the N-terminus disrupts the endonuclease activity, as expression of an N-terminal FLAG-tagged UTP23 was also unable to rescue the *in vivo* pre-rRNA processing phenotype caused by UTP23 depletion. Further experiments will use recombinant UTP23 containing a C-terminal tag to assess this possibility. Alternatively,

UTP23-mediated cleavage *in vivo* may require the presence of other ribosome biogenesis factors that are not present in the *in vitro* system used here. This may include factors that ensure the correct folding of the RNA substrate to permit cleavage, for example the U17 or U3 snoRNPs, or co-factors that activate the endonuclease activity of UTP23. As mentioned previously, the later was discovered to be the case for another PIN domain endonuclease, NOB1, which cleaves at site 3, and requires a co-factor for its *in vitro* endonuclease activity (Bai *et al.*, 2016).

While recombinant GST-UTP23 failed to cleave a pre-rRNA mimic containing the A0 site, data from this study shows some evidence for cleavage activity for this protein. In *in vitro* cleavage assays using a radiolabelled oligonucleotide containing the A0 site, weak cleavage was observed at multiple positions outside of the A0 site with WT GST-UTP23. These cleavages were not observed with incubation of the PIN domain mutant protein, suggesting that this activity is mediated by the PIN domain of UTP23. Further experiments are required to confirm the endonuclease activity of UTP23, but these preliminary *in vitro* results, in addition to the importance of an intact UTP23 PIN domain for pre-rRNA processing *in vivo*, make UTP23 a strong candidate for cleavage at site A0 in human cells. A role for UTP23 in human A0 cleavage was unexpected, considering the non-enzymatic role of yeast Utp23 in ribosome biogenesis. For most other pre-rRNA cleavages, the endonuclease mediating cleavage in yeast appears to play a conserved role in higher eukaryotes. Until recently, the human homologue of RNase MRP, which cleaves site A3 in yeast, was thought not to cleave the equivalent site, site 2, in humans (Sloan *et al.*, 2013b), but was later confirmed as the human site 2 nuclease (Goldfarb and Cech, 2017). Differences between the 5' ETS in yeast and humans may explain the potential distinction in the enzymes that process this spacer. Firstly, while the 5' ETS in yeast is ~700 nt long, the human 5' ETS is >3.5 kb in length. Therefore, while the yeast A_0 site is located around 100 nt from the 5' end of the 18S rRNA sequence, human A0 is around 2 kb from 18S. The human 5' ETS also contains an additional cleavage at site A', upstream of A0. While cleavages at site A_0/A0 in both yeast and humans are functionally equivalent in terms of their reliance on SSU processome formation, yeast A_0 may share similarity with human A', for which the endonuclease also remains unidentified. Indeed, as well as cleavage at these two sites being the initial cleavage event in the 5' ETS in their respective system, they both occur co-transcriptionally (Lazdins *et al.*, 1997; Kos and Tollervey, 2010; Kos-Braun *et al.*, 2017). These differences may explain why

apparently equivalent cleavage sites could be cleaved by different enzymes in the two systems, although further work is required to confirm this.

6.3 The link between SSU biogenesis factors, p53 signalling and cancer

6.3.1 SSU endonucleases in p53 signalling and cancer

As a key determinant of cellular growth rate, ribosome biogenesis is tightly linked to cancer, and is upregulated in tumours (Ruggero and Pandolfi, 2003). It is also linked to regulation of the tumour suppressor p53 via the LSU assembly intermediate, the 5S RNP (Pelava *et al.*, 2016). Defects in ribosome biogenesis lead to accumulation of the 5S RNP, which contains the 5S rRNA and two LSU RPs, RPL5 and RPL11. The 5S RNP binds to and inhibits MDM2, preventing it from targeting p53 for degradation, leading to an accumulation of p53, which results in cell cycle arrest, apoptosis or senescence. While the link between p53 signalling and ribosome biogenesis was initially thought to occur exclusively through the LSU assembly pathway, defects in SSU biogenesis, through depletion of SSU RPs or ribosome biogenesis factors required for SSU maturation, can also lead to p53 induction (Pelava, 2016). SSU processing defects lead to p53 induction via the 5S RNP without altering the levels of the mature LSU rRNAs (Fumagalli *et al.*, 2012; Pelava, 2016), and it is currently unknown how defects in SSU biogenesis lead to increased accumulation of the free 5S RNP. Induction of p53 upon defective ribosome biogenesis is counterintuitive, as ribosome biogenesis factors are mutated in many cancers, which presumably disrupts pre-rRNA processing and should therefore lead to p53 induction and suppression of tumorigenesis. It is possible that cancer cells in which ribosome biogenesis factors are mutated also contain mutations in components of the p53 signalling pathway, or that these cells have adapted and become de-sensitised to p53. This would explain why these cells do not undergo apoptosis or cell cycle arrest.

This study aimed to analyse the effect of SSU candidate endonucleases on p53 signalling, and their potential links to cancer. Preliminary results show that depletion of UTP23, but not UTP24 or RCL1, leads to a mild accumulation of p53 in both U2OS and MCF7 cells. It is unclear why depletion of UTP23 specifically leads to p53 induction, but depletion of UTP24 or RCL1 does not. In HEK293T and HeLa cells, depletion of UTP24, RCL1 or UTP23 disrupts cleavages at sites A0, 1 and 2a, consistent with their importance as part of the SSU processome. However, as shown

in this study, depletion of UTP23 in U2OS or MCF7 cells caused a variable pre-rRNA processing, often resulting in disruption of later cleavage steps. The reason for the distinct pre-rRNA processing upon UTP23 depletion in these cell types is unclear, but it possibly explains the difference in p53 activation compared to depletion of UTP24 and RCL1. It is possible that p53 induction only occurs with disruption of later stages of SSU processing. Further work in different cell types is required to identify the "correct" pre-rRNA processing defect caused by UTP23 depletion. It will be interesting to assess whether p53 induction is observed in p53-active cells in which UTP23 depletion causes the "classical" pre-rRNA processing defect seen with depletion of SSU processome components.

Interestingly, UTP23 is mutated in cancer, and a specific mutation, P215Q, was recently found associated with colorectal cancer (Timofeeva *et al.*, 2015). Whether this particular mutation, which is outside of the PIN domain of UTP23, affects pre-rRNA processing was previously unknown. Many ribosomopathies are associated with increased cancer risk, and defects in ribosome biogenesis are often a result of haploinsufficiency, as the mutant protein is not stably expressed. However, the UTP23 P215Q mutant was efficiently expressed in both HEK293T and U2OS stable cell lines, indicating that this mutation does not affect the stability of the UTP23 protein. Alternatively, ribosome biogenesis defects in ribosomopathies may be caused by overexpression of mutated biogenesis factors, leading to dominant negative effects. Consistent with this idea, overexpression of the UTP23 P215Q mutant in HEK293T cells caused a mild pre-rRNA processing defect, which was not observed with expression of the mutant at lower levels, or with overexpression of wild type UTP23. This defect was similar to that seen upon depletion of UTP23, with cleavages at sites A0, 1 and 2a affected. However, unlike UTP23 depletion, in which disruption of these cleavages leads to significant reduction in 18SE levels, 18SE levels were unchanged in cells overexpressing the P215Q mutant protein. This suggests that these three cleavages are potentially delayed, but not abolished, as processing continues despite accumulation of the 30S precursor.

The P215Q mutant localised correctly to the nucleolus, whether expressed at lower levels or when overexpressed, and bound to the U17 snoRNA at a similar efficiency to WT UTP23 *in vitro*, suggesting that this mutation does not affect the general functions of UTP23 in SSU processome assembly. Preliminary immunoprecipitation experiments suggest that the P215Q mutant may associate more

strongly with U17, either directly or in pre-ribosomal particles, than WT UTP23, which may explain the pre-rRNA processing defect observed when this mutant was expressed at extremely high levels compared to endogenous UTP23. In stable U2OS cell lines, expression of the P215Q mutant, either at levels similar to endogenous UTP23, or overexpressed, did not cause p53 induction. Due to time constraints, pre-rRNA processing was not assessed upon expression of the P215Q mutant in U2OS cells. As depletion of UTP23 caused a distinct pre-rRNA processing defect in U2OS cells compared to HeLa and HEK293T cells, it is possible that overexpression of this mutant would also lead to alternative processing defects. It would also be interesting to analyse p53 levels upon expression of the P215Q mutant in a cell type with a "classical" UTP23 depletion phenotype that also has active p53.

It is unclear whether the UTP23 P215Q mutant protein is overexpressed in colorectal cancer patients, or whether the disease is associated with a potential pre-rRNA processing defect caused by this mutation. However, it is possible that this mutant affects pre-rRNA processing in a way that contributes to the disease, without causing induction of p53, therefore encouraging cancer progression. Other UTP23 mutations are also associated with colorectal cancers (The Cancer Genome Atlas, 2012; Giannakis *et al.*, 2016) as well as other cancer types (Sharpe *et al.*, 2015). Further studies should investigate the impact of these mutations on pre-rRNA processing and the p53 response in order to more fully understand the link between UTP23 and cancer.

6.3.2 ITS1 processing factors in p53 signalling and cancer

One aspect of human ribosome biogenesis that appears to differ significantly from yeast ribosome biogenesis is the removal of ITS1, downstream of the 18S rRNA sequence. Firstly, the major endonuclease cleavage in yeast ITS1 is at site A_2, equivalent to human site 2a, while the major human ITS1 cleavage occurs at site 2, equivalent to yeast site A3. Secondly, while the 3' end of yeast 18S is matured exclusively by endonucleolytic cleavages, human 3' end 18S processing involves 3' to 5' exonucleolytic processing, which requires the exosome-associated exonuclease RRP6 and PARN (Sloan *et al.*, 2013b; Ishikawa *et al.*, 2017; Montellese *et al.*, 2017). In both yeast and human cells, ITS1 processing also involves the 5' to 3' exonuclease Rat1/XRN2 (Henry *et al.*, 1994; Preti *et al.*, 2013; Sloan *et al.*, 2013b) and the RNA binding protein RRP5 (Venema and Tollervey, 1996; Eppens *et al.*, 1999; Sloan *et al.*,

2013b). The reason for such substantial differences between yeast and human ITS1 processing is unclear, but the increased involvement of exonucleases may allow for rapid degradation upon defective processing. This would potentially increase the efficiency of surveillance pathways that may be more vital in human cells, where the production of ribosomes is more complex compared to yeast. While yeast ITS1 is ~350 nt in length, human ITS1 is >1,000 nt long, and this difference in length, and therefore difference between cleavage sites, may also contribute to the differences in processing in this region between the two systems. In addition, unlike yeast ITS1, human ITS1 is very GC-rich, which may also affect how this spacer is processed. These differences are significant clinically, as defects in ITS1 processing are commonly associated with ribosomopathies and for one ribosomopathy, Diamond Blackfan Anaemia (DBA), are even used as diagnostic marker (Farrar *et al.*, 2014). Interestingly, a recent high-throughput screen identified multiple mutations, including nonsense and frameshift mutations that may result in loss of function, in genes encoding RRP6, XRN2 and RRP5 in tumours and cell lines (Iorio *et al.*, 2016). Depletion of XRN2 causes a shift in 18S 3' end processing to the minor pathway, involving cleavage at 2a. Therefore, it is possible that cancer-associated mutations in ITS1 processing factors cause a similar shift to the minor ITS1 processing pathway, which would suggest that cancer cells may rely on processing via the minor pathway. If this is the case, this pathway may be used as a target in cancer treatment.

Surprisingly, depletion of either XRN2 or RRP6 did not cause p53 induction in either U2OS or MCF7 cells. This is particularly surprising as both exonucleases play a broad range of roles in RNA processing and degradation, outside of their roles in ribosome biogenesis. It is unclear why this would be the case. As depletion of XRN2 causes processing to shift to the minor pathway without significantly affecting 18SE or 18S rRNA production, this result suggests that processing through this pathway does not lead to activation of p53. This further implies that the minor ITS1 processing pathway is both productive in 18S rRNA maturation and does not result in cell cycle arrest or apoptosis. The high frequency of XRN2 mutations in cancer cells therefore suggests that cancer cells may exploit the minor processing pathway upon loss of XRN2 function, without activating the p53 pathway. However, XRN2 is also involved in the 5' end processing of the 28S rRNA (Wang and Pestov, 2011), so it is unclear why the LSU processing defect caused by XRN2 depletion would not lead to activation of p53. For RRP6, the explanation for this result is even less clear, as depletion of RRP6

causes a reduction in 18SE pre-rRNA levels and accumulation of 21S and 21SC intermediates, indicating that 18S rRNA maturation is defective in the absence of this protein. One potential explanation is that RRP6 is somehow involved in the MDM2-p53 signalling pathway, and therefore its depletion disrupts the activation of p53, but this possibility is yet to be investigated. Interestingly, data from our lab indicates that depletion of RRP6 in MCF7 and U2OS cells causes a distinct pre-rRNA processing defect, namely accumulation of the 30S pre-rRNA (data not show). Further work will focus on analysing the effect of cancer-associated XRN2 and RRP6 mutations on both pre-rRNA processing and p53 induction.

RRP5 depletion caused strong p53 induction in both U2OS and MCF7 cells, with an increase in p53 protein levels similar to that observed upon depletion of the LSU RP RPL7. This induction was dependent on the presence of the 5S RNP component, RPL5, suggesting that the p53 induction caused by depletion of RPL5 is mediated by inhibition of MDM2 by the 5S RNP. In yeast, Rrp5 is required for both endonuclease cleavages in ITS1 (Venema and Tollervey, 1996; Eppens *et al.*, 1999), and this also appears to be the case in human cells (Sloan *et al.*, 2013b). This suggests that in human cells, depletion of RRP5 would disrupt both major and minor pathways of ITS1 processing, causing defective 18S rRNA production and leading to induction of p53. As depletion of RRP5 also disrupts cleavage at site 2, which is important for LSU maturation, it is impossible to determine whether the p53 induction caused by RRP5 depletion is due to defective SSU or LSU processing, or both. In yeast, the Rrp5 protein can be functionally separated, with the N-terminal region required for site A3 (human site 2) cleavage and the C-terminal region important for A_2 cleavage (human site 2a) (Vos *et al.*, 2004b; Lebaron *et al.*, 2013). In cancer cells containing mutations in the gene encoding RRP5, it is possible that only one of the two ITS1 processing pathways is disrupted, which may not lead to p53 activation. However, it is currently unknown whether human RRP5 can be functionally separated like the yeast protein.

Previous work has revealed that depletion of RRP5 in HeLa cells disrupts processing at sites A0, 1, 2a and 2 (Sloan *et al.*, 2013b), consistent with phenotypes seen upon Rrp5 depletion in yeast (Venema and Tollervey, 1996). Data from this study indicate that RRP5 depletion in HEK293T cells causes a very similar pre-rRNA processing phenotype. When processing through the minor pathway is favoured, by depletion of XRN2, RRP5 depletion caused a reduction in 36S pre-rRNA levels, as previously seen in HeLa cells, indicating disrupted site 2a cleavage, or alternatively,

increased site 2 cleavage relative to site 2a cleavage. In contrast to data from HeLa cells, levels of the ITS1 fragment, the result of cleavage at both site 2a and site 2, were not reduced upon depletion of both XRN2 and RRP5. Together, these results suggest that RRP5 depletion may lead to an increase in cleavage at site 2 relative to site 2a cleavage. Therefore, RRP5 may play a more prominent role in site 2a cleavage when processing occurs through the minor pathway in HEK293T cells.

As with depletion of other ribosome biogenesis factors, depletion of either RRP5 or XRN2 in MCF7 cells produced pre-rRNA processing defects different from those observed in other human cell lines or in yeast. RRP5 depletion did not lead to accumulation of the 30SL3' precursor, instead causing 30S pre-rRNA accumulation, indicating defective cleavage at sites A0, 1 and 2a, but not site 2. XRN2 depletion did not cause strong accumulation of the 36S precursor, or the ITS1 fragment as commonly seen in HeLa and HEK293T cells (Sloan *et al.*, 2014; Wells *et al.*, 2016). This is similar to the distinct pre-rRNA processing defects seen upon depletion of other ribosome biogenesis factors in this study, as well as data from our lab for RRP6 depletion (data not shown). These results suggest that RRP5 is not important for site 2 cleavage in this cell line, and that depletion of XRN2 does not cause a shift to the minor ITS1 processing pathway, suggesting that the MCF7 cell line may be naturally different in pre-rRNA processing. Further work using a range of cell types is required to identify the "normal" pre-rRNA processing phenotype upon depletion of RRP5, although phenotypes seen in HeLa and HEK293T cells are consistent with defects upon depletion of yeast homologues, so HEK293T cells were used for creation of stable RRP5 cell lines.

Creation of stable HEK293T cell lines showed that expression of FLAG-tagged RRP5 is able to rescue the pre-rRNA processing phenotype caused by depletion of RRP5. This allows for further work on understanding the importance of the different regions of the human RRP5 protein in relation to specific cleavages in ribosome biogenesis. HEK293T cells were used as the processing defect seen upon RRP5 depletion was most similar to previous reports in human cells, and consistent with data from depletion of Rrp5 in yeast. Yeast Rrp5 contains a total of 12 S1 RNA binding domains, 9 of which form the N-terminal domain (NTD), required for site A3 cleavage, while the other three, together with the C-terminal TPR protein-protein interaction domains, make up the C-terminal domain (CTD) required for cleavages at sites A_0, A_1 and A_2 (Eppens *et al.*, 1999). The three S1 domains of the CTD are specifically required

for cleavage at site A_2 (Vos *et al.*, 2004b). In contrast, human RRP5 contains 13 S1 domains, so two forms of each RRP5 mutant was designed, splitting the protein after the ninth S1 domain, as in yeast, or after the tenth, leaving three C-terminal S1 domains. All RRP5 mutant proteins were efficiently expressed in transfected cells, suggesting that they are all stable. Further work will determine the ability of each of these mutants to rescue specific pre-rRNA processing defects upon depletion of endogenous RRP5, leading to functional characterisation of human RRP5. In addition, U2OS stable cell lines will be generated and p53 levels will be analysed to establish which, if any, of the RRP5 mutants rescue the p53 induction caused by RRP5 depletion.

6.4 Future Directions

This study has provided strong evidence supporting the role of UTP24 as the endonuclease that cleaves the human pre-rRNA at sites 1 and 2a (Figure 6.1). Building on parallel work performed in *S. cerevisiae*, this data also suggests that this role is conserved between yeast and higher eukaryotes (Wells *et al.*, 2016). UTP24 is now a very strong candidate for these cleavages, but its endonuclease activity at these specific sites has so far not been demonstrated. Parallel work from our lab has shown that human UTP24 can cleave an RNA substrate, and that this cleavage activity is dependent on an intact PIN domain (Wells *et al.*, 2016). However, further work is required to confirm that UTP24 can cleave a substrate containing human cleavage sites 1 and 2a. So far, attempts to confirm this have proved unsuccessful. Specific cleavage of sites 1 and 2a by UTP24 may require other factors that have not so far been present in *in vitro* assays. These factors may be required to induce conformational changes in the RNA substrate that permit UTP24 endonuclease activity, and distinct factors may be required for the two different cleavage sites. UTP24 may also require a co-factor for endonuclease activity at these sites, as was the case for the site 3 endonuclease NOB1 (Bai *et al.*, 2016). Further work will aim to identify any potential cofactors of UTP24. Like UTP23, UTP24 also contains a Zinc finger RNA binding motif, but the importance of this motif has not previously been tested. Future work will establish whether this motif is important for human pre-rRNA processing, and whether it is essential for interactions between UTP24 and the pre-rRNA and/or the U3 snoRNA. Immunofluorescence and gradient sedimentation experiments will determine

the localisation of UTP24 mutants and their association with pre-ribosomal particles, respectively.

This work has also revealed a potential novel role in endonucleolytic cleavage for UTP23, establishing the importance of its PIN domain in three early cleavages in human pre-rRNA processing *in vivo*. The requirement of an intact UTP23 PIN domain in all three early cleavages, and the probable role for UTP24 in cleavage at sites 1 and 2a, leads to the tentative suggestion that UTP23 is the endonuclease responsible for A0 cleavage (Figure 6.1). However, attempts to confirm the endonuclease activity of UTP23 at this site have so far been unsuccessful, potentially due to the presence of the N-terminal GST tag on the recombinant protein used. Future experiments will attempt to rectify this using C-terminally-tagged recombinant UTP23. It is also possible that, as mentioned for UTP24, UTP23-mediated cleavage requires the presence of other ribosome biogenesis factors, so future work will also attempt to identify potential co-factors of UTP23.

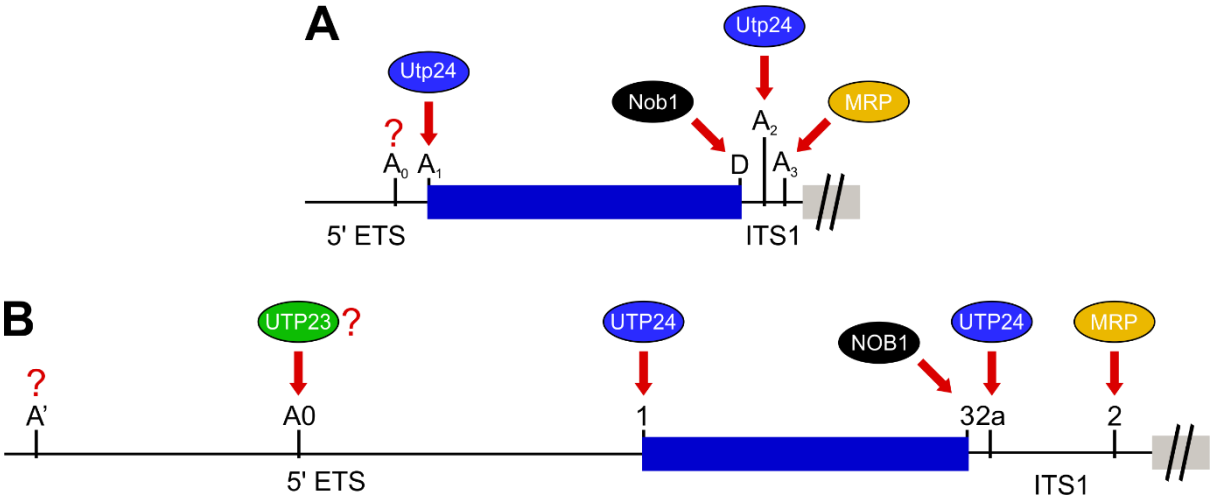


Figure 6.1 Summary of characterised endonucleases in yeast and human 18S rRNA maturation. (A) A region of the yeast primary pre-rRNA transcript containing the 5' ETS, the 18S rRNA and ITS1. Positions of endonuclease cleavage sites are shown, and coloured ovals represent endonucleases, where identified, that cleave at the respective cleavage sites. Red arrows indicate the cleavage site(s) cleaved by each endonuclease. **(B)** A region of the human primary pre-rRNA transcript containing the 5' ETS, the 18S rRNA and ITS1. Endonuclease cleavage sites and endonucleases are shown as for panel A. Red question marks indicate that the enzyme responsible for cleavage at a particular site is unknown, and a red question mark next to a coloured oval represents a potential, but unconfirmed role for cleavage by this factor at this site.

This work has further revealed the importance UTP23's conserved Zinc finger motif in pre-rRNA processing, however the RNA substrate that this motif binds to *in vivo* has not been confirmed. Preliminary data suggests that the Zinc finger motif is important in UTP23's association with the U17 snoRNA, but mutation of this motif only reduces the ability of UTP23 to bind this snoRNA. Therefore, it is likely that other regions of UTP23 are important for this binding, and these remain to be identified. Data from this study has shown that recombinant UTP23 interacts with NHP2, a core box H/ACA protein, *in vitro* and this interaction may contribute to UTP23 association with U17 *in vivo*, thus reducing the importance of the Zinc finger in this association.

Gradient analysis will also reveal the importance of UTP23's PIN domain and Zinc finger motif in the association of both UTP23 and the U17 snoRNA with pre-ribosomal complexes. While depletion of UTP23 in HEK293T cells caused accumulation of the 30S pre-rRNA, consistent with previous work in yeast and HeLa cells, UTP23 depletion in U2OS and MCF7 cells showed variable pre-rRNA processing defects. Therefore, future work requires the study of UTP23 in multiple cell types in order to clarify these distinctions.

While this work has potentially identified the endonucleases responsible for three early cleavages in human ribosome biogenesis, multiple cleavage events still remain to be assigned an endonuclease. In SSU biogenesis, the endonuclease that cleaves the higher-eukaryote specific cleavage at site A' remains to be identified. Unlike the other three SSU cleavages, A' cleavage does not require the whole SSU processome, but only a subset of SSU processome components (Sloan *et al.*, 2014), and therefore it is possible that the A' endonuclease may not be an SSU processome component. As this work has revealed that the yeast A_0 site and the equivalent human site appear to be cleaved by distinct enzymes, further work will also investigate this difference and aim to identify the A_0 nuclease in yeast.

A cancer-associated UTP23 mutation was shown in this study to cause a pre-rRNA processing defect when expressed at extremely high levels relative to endogenous UTP23, and initial results suggest that this mutant protein may associate more strongly with the U17 snoRNA *in vivo*. Further work is required to confirm this increased association, and gradient analysis will assess whether this mutation affects the association of UTP23 or the U17 snoRNA with pre-ribosomal complexes in the absence of endogenous UTP23. These results have also shown that overexpression

of this mutant in cells containing an active p53 gene does not result in p53 induction. Future work will investigate whether pre-rRNA processing is defective upon expression of this mutant in these cells. Other UTP23 mutations are also found in cancers, and their effect on pre-rRNA processing and p53 induction will also be studied.

This work has investigated the link between defective SSU ribosome biogenesis and p53 signalling, and shown that depletion of some, but not all SSU biogenesis factors results in activation of p53. Future work will establish whether the accumulation of p53 upon UTP23 depletion is dependent on the 5S RNP. A further aim is to understand the reason for the observed p53 induction with depletion of some SSU factors, but not others. This will involve testing more SSU biogenesis factors that function in different stages of 18S rRNA maturation and assessing the effect of their depletion upon p53 activation. This will help to understand how SSU biogenesis is connected to the 5S-RNP-MDM2-p53 signalling pathway.

Results from this study have shown that depletion of the RNA binding protein RRP5, which is frequently mutated in cancer, causes activation of p53 in two human cell lines, and that the accumulation of p53 is dependent on the 5S RNP component RPL5. I have established an RNAi rescue system that will allow for investigation of the importance of the different domains of the RRP5 protein in human pre-rRNA processing, as well as how the defects caused by expression of RRP5 mutations affects p53 signalling. Study of the importance of RRP5's domains in different stages of pre-rRNA processing may help to understand the difference in human ITS1 processing compared to yeast. Future work will also look at cancer-associated mutations of RRP5 as well as other ITS1 processing factors including RRP6 and XRN2, in order to determine their impact on pre-rRNA processing, particularly on the ITS1 processing pathway favoured, and p53 induction. Future work will also look at cancer cell lines, which contain mutations in ITS1 processing factors. These are likely to only affect one allele for these essential genes, and may cause changes in protein function, or haploinsufficiency. Therefore, further work will involve determining how these changes or haploinsufficiency lead to pre-rRNA processing defects and/or alter the ITS1 processing pathway used.

6.5 Conclusions

Together, data gathered during this PhD project reveal the potential endonucleases that likely cleave the pre-rRNA at three key sites in human 18S rRNA maturation, revealing a novel distinction, in addition to known differences in ITS1 processing, between yeast and human ribosome biogenesis. This emphasises the importance of studying ribosome biogenesis in human cells, especially given the link between this process and human health. These data also reveal insights into the link between human SSU biogenesis and p53 signalling, as well as the impact of cancer-associated mutations in ribosome biogenesis factors. Further work may lead to novel mechanisms for targeting cancers containing mutations in ribosome biogenesis factors.

References

Ahmad, Y., Boisvert, F.M., Gregor, P., Cobley, A. and Lamond, A.I. (2009) 'NOPdb: Nucleolar Proteome Database--2008 update', *Nucleic Acids Res*, 37(Database issue), pp. D181-4.

Akopian, D., Shen, K., Zhang, X. and Shan, S.O. (2013) 'Signal recognition particle: an essential protein-targeting machine', *Annu Rev Biochem*, 82, pp. 693-721.

Alexandrov, A., Colognori, D. and Steitz, J.A. (2011) 'Human eIF4AIII interacts with an eIF4G-like partner, NOM1, revealing an evolutionarily conserved function outside the exon junction complex', *Genes Dev*, 25(10), pp. 1078-90.

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E. and Tollervey, D. (1999) 'Functions of the exosome in rRNA, snoRNA and snRNA synthesis', *EMBO J*, 18(19), pp. 5399-410.

Allmang, C., Mitchell, P., Petfalski, E. and Tollervey, D. (2000) 'Degradation of ribosomal RNA precursors by the exosome', *Nucleic Acids Res*, 28(8), pp. 1684-91.

Amberg, D.C., Goldstein, A.L. and Cole, C.N. (1992) 'Isolation and characterization of RAT1: an essential gene of Saccharomyces cerevisiae required for the efficient nucleocytoplasmic trafficking of mRNA', *Genes Dev*, 6(7), pp. 1173-89.

Anger, A.M., Armache, J.P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D.N. and Beckmann, R. (2013) 'Structures of the human and Drosophila 80S ribosome', *Nature*, 497(7447), pp. 80-5.

Angrisani, A., Vicidomini, R., Turano, M. and Furia, M. (2014) 'Human dyskerin: beyond telomeres', *Biol Chem*, 395(6), pp. 593-610.

Ansel, K.M., Pastor, W.A., Rath, N., Lapan, A.D., Glasmacher, E., Wolf, C., Smith, L.C., Papadopoulou, N., Lamperti, E.D., Tahiliani, M., Ellwart, J.W., Shi, Y., Kremmer, E., Rao, A. and Heissmeyer, V. (2008) 'Mouse Eri1 interacts with the ribosome and catalyzes 5.8S rRNA processing', *Nat Struct Mol Biol*, 15(5), pp. 523-30.

Antal, M., Mougin, A., Kis, M., Boros, E., Steger, G., Jakab, G., Solymosy, F. and Branlant, C. (2000) 'Molecular characterization at the RNA and gene levels of U3 snoRNA from a unicellular green alga, Chlamydomonas reinhardtii', *Nucleic Acids Res*, 28(15), pp. 2959-68.

Arcus, V.L., Backbro, K., Roos, A., Daniel, E.L. and Baker, E.N. (2004) 'Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease', *J Biol Chem*, 279(16), pp. 16471-8.

Armache, J.P., Jarasch, A., Anger, A.M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S., Marquez, V., Mielke, T., Thomm, M., Berninghausen, O., Beatrix, B., Soding, J., Westhof, E., Wilson, D.N. and Beckmann, R. (2010a) 'Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-A resolution', *Proc Natl Acad Sci U S A*, 107(46), pp. 19748-53.

Armache, J.P., Jarasch, A., Anger, A.M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S., Marquez, V., Mielke, T., Thomm, M., Berninghausen, O., Beatrix, B., Soding, J., Westhof, E., Wilson, D.N. and Beckmann, R. (2010b) 'Localization of eukaryote-specific ribosomal proteins in a 5.5-A cryo-EM map of the 80S eukaryotic ribosome', *Proc Natl Acad Sci U S A*, 107(46), pp. 19754-9.

Armistead, J., Khatkar, S., Meyer, B., Mark, B.L., Patel, N., Coghlan, G., Lamont, R.E., Liu, S., Wiechert, J., Cattini, P.A., Koetter, P., Wrogemann, K., Greenberg, C.R., Entian, K.D., Zelinski, T. and Triggs-Raine, B. (2009) 'Mutation of a gene essential for ribosome biogenesis, EMG1, causes Bowen-Conradi syndrome', *Am J Hum Genet*, 84(6), pp. 728-39.

Arn, E.A. and Abelson, J.N. (1996) 'The 2'-5' RNA ligase of Escherichia coli. Purification, cloning, and genomic disruption', *J Biol Chem*, 271(49), pp. 31145-53.

Atzorn, V., Fragapane, P. and Kiss, T. (2004) 'U17/snR30 Is a Ubiquitous snoRNA with Two Conserved Sequence Motifs Essential for 18S rRNA Production', *Molecular and Cellular Biology*, 24(4), pp. 1769-1778.

Axt, K., French, S.L., Beyer, A.L. and Tollervey, D. (2014) 'Kinetic analysis demonstrates a requirement for the Rat1 exonuclease in cotranscriptional pre-rRNA cleavage', *PLoS One*, 9(2), p. e85703.

Babu, K.A. and Verma, R.S. (1985) 'Structural and functional aspects of nucleolar organizer regions (NORs) of human chromosomes', *Int Rev Cytol*, 94, pp. 151-76.

Bai, D., Zhang, J., Li, T., Hang, R., Liu, Y., Tian, Y., Huang, D., Qu, L., Cao, X., Ji, J. and Zheng, X. (2016) 'The ATPase hCINAP regulates 18S rRNA processing and is essential for embryogenesis and tumour growth', *Nat Commun*, 7, p. 12310.

Barandun, J., Chaker-Margot, M., Hunziker, M., Molloy, K.R., Chait, B.T. and Klinge, S. (2017) 'The complete structure of the small-subunit processome', *Nat Struct Mol Biol*, 24(11), pp. 944-953.

Bartelt-Kirbach, B., Wuepping, M., Dodrimont-Lattke, M. and Kaufmann, D. (2009) 'Expression analysis of genes lying in the NF1 microdeletion interval points to four candidate modifiers for neurofibroma formation', *Neurogenetics*, 10(1), pp. 79-85.

Bassler, J., Grandi, P., Gadal, O., Lessmann, T., Petfalski, E., Tollervey, D., Lechner, J. and Hurt, E. (2001) 'Identification of a 60S preribosomal particle that is closely linked to nuclear export', *Mol Cell*, 8(3), pp. 517-29.

Bellodi, C., McMahon, M., Contreras, A., Juliano, D., Kopmar, N., Nakamura, T., Maltby, D., Burlingame, A., Savage, S.A., Shimamura, A. and Ruggero, D. (2013) 'H/ACA small RNA dysfunctions in disease reveal key roles for noncoding RNA modifications in hematopoietic stem cell differentiation', *Cell Rep*, 3(5), pp. 1493-502.

Beltrame, M. and Tollervey, D. (1992) 'Identification and functional analysis of two U3 binding sites on yeast pre-ribosomal RNA', *EMBO J*, 11(4), pp. 1531-42.

Beltrame, M. and Tollervey, D. (1995) 'Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis', *EMBO J*, 14(17), pp. 4350-6.

Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G. and Yusupov, M. (2011) 'The structure of the eukaryotic ribosome at 3.0 A resolution', *Science*, 334(6062), pp. 1524-9.

Bernstein, K.A., Gallagher, J.E., Mitchell, B.M., Granneman, S. and Baserga, S.J. (2004) 'The small-subunit processome is a ribosome assembly intermediate', *Eukaryot Cell*, 3(6), pp. 1619-26.

Bernstein, K.A., Granneman, S., Lee, A.V., Manickam, S. and Baserga, S.J. (2006) 'Comprehensive mutational analysis of yeast DEXD/H box RNA helicases involved in large ribosomal subunit biogenesis', *Mol Cell Biol*, 26(4), pp. 1195-208.

Bertrand, E., Houser-Scott, F., Kendall, A., Singer, R.H. and Engelke, D.R. (1998) 'Nucleolar localization of early tRNA processing', *Genes Dev*, 12(16), pp. 2463-8.

Billy, E., Wegierski, T., Nasr, F. and 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', *EMBO J*, 19(9), pp. 2115-26.

Bleichert, F. and Baserga, S.J. (2007) 'The long unwinding road of RNA helicases', *Mol Cell*, 27(3), pp. 339-52.

Bleichert, F., Granneman, S., Osheim, Y.N., Beyer, A.L. and Baserga, S.J. (2006) 'The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation', *Proc Natl Acad Sci U S A*, 103(25), pp. 9464-9.

Bohnsack, M.T., Kos, M. and Tollervey, D. (2008) 'Quantitative analysis of snoRNA association with pre-ribosomes and release of snR30 by Rok1 helicase', *EMBO Rep*, 9(12), pp. 1230-6.

Boisvert, F.M., van Koningsbruggen, S., Navascues, J. and Lamond, A.I. (2007) 'The multifunctional nucleolus', *Nat Rev Mol Cell Biol*, 8(7), pp. 574-85.

Bonneau, F., Basquin, J., Ebert, J., Lorentzen, E. and Conti, E. (2009) 'The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation', *Cell*, 139(3), pp. 547-59.

Borovjagin, A.V. and Gerbi, S.A. (1999) 'U3 small nucleolar RNA is essential for cleavage at sites 1, 2 and 3 in pre-rRNA and determines which rRNA processing pathway is taken in Xenopus oocytes', *J Mol Biol*, 286(5), pp. 1347-63.

Borovjagin, A.V. and Gerbi, S.A. (2001) 'Xenopus U3 snoRNA GAC-Box A' and Box A sequences play distinct functional roles in rRNA processing', *Mol Cell Biol*, 21(18), pp. 6210-21.

Borovjagin, A.V. and Gerbi, S.A. (2004) 'Xenopus U3 snoRNA docks on pre-rRNA through a novel base-pairing interaction', *RNA*, 10(6), pp. 942-53.

Borovjagin, A.V. and Gerbi, S.A. (2005) 'An evolutionary intra-molecular shift in the preferred U3 snoRNA binding site on pre-ribosomal RNA', *Nucleic Acids Res*, 33(15), pp. 4995-5005.

Bowen, P. and Conradi, G.J. (1976) 'Syndrome of skeletal and genitourinary anomalies with unusual facies and failure to thrive in Hutterite sibs', *Birth Defects Orig Artic Ser*, 12(6), pp. 101-8.

Bowman, L.H., Goldman, W.E., Goldberg, G.I., Hebert, M.B. and Schlessinger, D. (1983) 'Location of the initial cleavage sites in mouse pre-rRNA', *Mol Cell Biol*, 3(8), pp. 1501-10.

Brannan, K., Kim, H., Erickson, B., Glover-Cutter, K., Kim, S., Fong, N., Kiemele, L., Hansen, K., Davis, R., Lykke-Andersen, J. and Bentley, D.L. (2012) 'mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription', *Mol Cell*, 46(3), pp. 311-24.

Briggs, M.W., Burkard, K.T. and Butler, J.S. (1998) 'Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation', *J Biol Chem*, 273(21), pp. 13255-63.

Brown, C.J., Lain, S., Verma, C.S., Fersht, A.R. and Lane, D.P. (2009) 'Awakening guardian angels: drugging the p53 pathway', *Nat Rev Cancer*, 9(12), pp. 862-73.

Bursac, S., Brdovcak, M.C., Pfannkuchen, M., Orsolic, I., Golomb, L., Zhu, Y., Katz, C., Daftuar, L., Grabusic, K., Vukelic, I., Filic, V., Oren, M., Prives, C. and Volarevic, S. (2012) 'Mutual protection of ribosomal proteins L5 and L11 from degradation is essential for p53 activation upon ribosomal biogenesis stress', *Proc Natl Acad Sci U S A*, 109(50), pp. 20467-72.

Callahan, K.P. and Butler, J.S. (2008) 'Evidence for core exosome independent function of the nuclear exoribonuclease Rrp6p', *Nucleic Acids Res*, 36(21), pp. 6645-55.

Carron, C., O'Donohue, M.F., Choesmel, V., Faubladier, M. and Gleizes, P.E. (2011) 'Analysis of two human pre-ribosomal factors, bystin and hTsr1, highlights differences in evolution of ribosome biogenesis between yeast and mammals', *Nucleic Acids Res*, 39(1), pp. 280-91.

Chagnon, P., Michaud, J., Mitchell, G., Mercier, J., Marion, J.F., Drouin, E., Rasquin-Weber, A., Hudson, T.J. and Richter, A. (2002) 'A missense mutation (R565W) in cirhin (FLJ14728) in North American Indian childhood cirrhosis', *Am J Hum Genet*, 71(6), pp. 1443-9.

Chaker-Margot, M., Hunziker, M., Barandun, J., Dill, B.D. and Klinge, S. (2015) 'Stage-specific assembly events of the 6-MDa small-subunit processome initiate eukaryotic ribosome biogenesis', *Nat Struct Mol Biol*, 22(11), pp. 920-3.

Chakraborty, A., Uechi, T. and Kenmochi, N. (2011) 'Guarding the 'translation apparatus': defective ribosome biogenesis and the p53 signaling pathway', *Wiley Interdiscip Rev RNA*, 2(4), pp. 507-22.

Charette, J.M. and Gray, M.W. (2009) 'U3 snoRNA genes are multi-copy and frequently linked to U5 snRNA genes in Euglena gracilis', *BMC Genomics*, 10, p. 528.

Cheng, J., Kellner, N., Berninghausen, O., Hurt, E. and Beckmann, R. (2017) '3.2-A-resolution structure of the 90S preribosome before A1 pre-rRNA cleavage', *Nat Struct Mol Biol*, 24(11), pp. 954-964.

Cheutin, T., O'Donohue, M.F., Beorchia, A., Vandelaer, M., Kaplan, H., Defever, B., Ploton, D. and Thiry, M. (2002) 'Three-dimensional organization of active rRNA genes within the nucleolus', *J Cell Sci*, 115(Pt 16), pp. 3297-307.

Chlebowski, A., Lubas, M., Jensen, T.H. and Dziembowski, A. (2013) 'RNA decay machines: the exosome', *Biochim Biophys Acta*, 1829(6-7), pp. 552-60.

Chu, S., Archer, R.H., Zengel, J.M. and Lindahl, L. (1994) 'The RNA of RNase MRP is required for normal processing of ribosomal RNA', *Proc Natl Acad Sci U S A*, 91(2), pp. 659-63.

Ciganda, M. and Williams, N. (2011) 'Eukaryotic 5S rRNA biogenesis', *Wiley Interdiscip Rev RNA*, 2(4), pp. 523-33.

Ciufo, L.F. and Brown, J.D. (2000) 'Nuclear export of yeast signal recognition particle lacking srp54p by the Xpo1p/Crm1p NES-dependent pathway', *Curr Biol*, 10(23), p. R882.

Clissold, P.M. and Ponting, C.P. (2000) 'PIN domains in nonsense-mediated mRNA decay and RNAi', *Curr Biol*, 10(24), pp. R888-90.

Cmejla, R., Cmejlova, J., Handrkova, H., Petrak, J., Petrtylova, K., Mihal, V., Stary, J., Cerna, Z., Jabali, Y. and Pospisilova, D. (2009) 'Identification of mutations in the ribosomal protein L5 (RPL5) and ribosomal protein L11 (RPL11) genes in Czech patients with Diamond-Blackfan anemia', *Hum Mutat*, 30(3), pp. 321-7.

Coccia, M., Rossi, A., Riccio, A., Trotta, E. and Santoro, M.G. (2017) 'Human NF-kappaB repressing factor acts as a stress-regulated switch for ribosomal RNA processing and nucleolar homeostasis surveillance', *Proc Natl Acad Sci U S A*, 114(5), pp. 1045-1050.

Colley, A., Beggs, J.D., Tollervey, D. and Lafontaine, D.L. (2000) 'Dhr1p, a putative DEAH-box RNA helicase, is associated with the box C+D snoRNP U3', *Mol Cell Biol*, 20(19), pp. 7238-46.

Connelly, S. and Manley, J.L. (1988) 'A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II', *Genes Dev*, 2(4), pp. 440-52.

Damianov, A., Kann, M., Lane, W.S. and Bindereif, A. (2006) 'Human RBM28 protein is a specific nucleolar component of the spliceosomal snRNPs', *Biol Chem*, 387(10-11), pp. 1455-60.

Danilova, N. and Gazda, H.T. (2015) 'Ribosomopathies: how a common root can cause a tree of pathologies', *Dis Model Mech*, 8(9), pp. 1013-26.
Daugeron, M.C. and Linder, P. (2001) 'Characterization and mutational analysis of yeast Dbp8p, a putative RNA helicase involved in ribosome biogenesis', *Nucleic Acids Res*, 29(5), pp. 1144-55.
de Boer, P., Vos, H.R., Faber, A.W., Vos, J.C. and Raue, H.A. (2006) 'Rrp5p, a trans-acting factor in yeast ribosome biogenesis, is an RNA-binding protein with a pronounced preference for U-rich sequences', *RNA*, 12(2), pp. 263-71.
de la Cruz, J., Kressler, D., Tollervey, D. and 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', *EMBO J*, 17(4), pp. 1128-40.
de la Cruz, J., Lacombe, T., Deloche, O., Linder, P. and Kressler, D. (2004) 'The putative RNA helicase Dbp6p functionally interacts with Rpl3p, Nop8p and the novel trans-acting Factor Rsa3p during biogenesis of 60S ribosomal subunits in Saccharomyces cerevisiae', *Genetics*, 166(4), pp. 1687-99.
Decatur, W.A. and Fournier, M.J. (2002) 'rRNA modifications and ribosome function', *Trends Biochem Sci*, 27(7), pp. 344-51.
Delaporta, P., Sofocleous, C., Stiakaki, E., Polychronopoulou, S., Economou, M., Kossiva, L., Kostaridou, S. and Kattamis, A. (2014) 'Clinical phenotype and genetic analysis of RPS19, RPL5, and RPL11 genes in Greek patients with Diamond Blackfan Anemia', *Pediatr Blood Cancer*, 61(12), pp. 2249-55.
Delcasso-Tremousaygue, D., Grellet, F., Panabieres, F., Ananiev, E.D. and Delseny, M. (1988) 'Structural and transcriptional characterization of the external spacer of a ribosomal RNA nuclear gene from a higher plant', *Eur J Biochem*, 172(3), pp. 767-76.
Delprato, A., Al Kadri, Y., Perebaskine, N., Monfoulet, C., Henry, Y., Henras, A.K. and Fribourg, S. (2014) 'Crucial role of the Rcl1p-Bms1p interaction for yeast pre-ribosomal RNA processing', *Nucleic Acids Res*, 42(15), pp. 10161-72.
Dembowski, J.A., Kuo, B. and Woolford, J.L., Jr. (2013) 'Has1 regulates consecutive maturation and processing steps for assembly of 60S ribosomal subunits', *Nucleic Acids Res*, 41(16), pp. 7889-904.
Derenzini, M., Thiry, M. and Goessens, G. (1990) 'Ultrastructural cytochemistry of the mammalian cell nucleolus', *J Histochem Cytochem*, 38(9), pp. 1237-56.
Deshmukh, M., Tsay, Y.F., Paulovich, A.G. and Woolford, J.L., Jr. (1993) 'Yeast ribosomal protein L1 is required for the stability of newly synthesized 5S rRNA and the assembly of 60S ribosomal subunits', *Mol Cell Biol*, 13(5), pp. 2835-45.
Dez, C., Dlakic, M. and Tollervey, D. (2007) 'Roles of the HEAT repeat proteins Utp10 and Utp20 in 40S ribosome maturation', *RNA*, 13(9), pp. 1516-27.
Dez, C., Houseley, J. and Tollervey, D. (2006) 'Surveillance of nuclear-restricted pre-ribosomes within a subnucleolar region of Saccharomyces cerevisiae', *EMBO J*, 25(7), pp. 1534-46.
Dez, C. and Tollervey, D. (2004) 'Ribosome synthesis meets the cell cycle', *Curr Opin Microbiol*, 7(6), pp. 631-7.
Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M. and Pagano, A. (2007) 'The expanding RNA polymerase III transcriptome', *Trends Genet*, 23(12), pp. 614-22.
Doherty, L., Sheen, M.R., Vlachos, A., Choesmel, V., O'Donohue, M.F., Clinton, C., Schneider, H.E., Sieff, C.A., Newburger, P.E., Ball, S.E., Niewiadomska, E., Matysiak, M., Glader, B., Arceci, R.J., Farrar,

J.E., Atsidaftos, E., Lipton, J.M., Gleizes, P.E. and Gazda, H.T. (2010) 'Ribosomal protein genes RPS10 and RPS26 are commonly mutated in Diamond-Blackfan anemia', *Am J Hum Genet*, 86(2), pp. 222-8.

Donati, G., Brighenti, E., Vici, M., Mazzini, G., Trere, D., Montanaro, L. and Derenzini, M. (2011) 'Selective inhibition of rRNA transcription downregulates E2F-1: a new p53-independent mechanism linking cell growth to cell proliferation', *J Cell Sci*, 124(Pt 17), pp. 3017-28.

Donati, G., Peddigari, S., Mercer, C.A. and Thomas, G. (2013) '5S ribosomal RNA is an essential component of a nascent ribosomal precursor complex that regulates the Hdm2-p53 checkpoint', *Cell Rep*, 4(1), pp. 87-98.

Dosil, M. and Bustelo, X.R. (2004) 'Functional characterization of Pwp2, a WD family protein essential for the assembly of the 90 S pre-ribosomal particle', *J Biol Chem*, 279(36), pp. 37385-97.

Dragon, F., Gallagher, J.E., Compagnone-Post, P.A., Mitchell, B.M., Porwancher, K.A., Wehner, K.A., Wormsley, S., Settlage, R.E., Shabanowitz, J., Osheim, Y., Beyer, A.L., Hunt, D.F. and Baserga, S.J. (2002) 'A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis', *Nature*, 417(6892), pp. 967-70.

Dunbar, D.A., Wormsley, S., Agentis, T.M. and Baserga, S.J. (1997) 'Mpp10p, a U3 small nucleolar ribonucleoprotein component required for pre-18S rRNA processing in yeast', *Mol Cell Biol*, 17(10), pp. 5803-12.

Dutca, L.M., Gallagher, J.E. and Baserga, S.J. (2011) 'The initial U3 snoRNA:pre-rRNA base pairing interaction required for pre-18S rRNA folding revealed by in vivo chemical probing', *Nucleic Acids Res*, 39(12), pp. 5164-80.

Dziembowski, A., Lorentzen, E., Conti, E. and Seraphin, B. (2007) 'A single subunit, Dis3, is essentially responsible for yeast exosome core activity', *Nat Struct Mol Biol*, 14(1), pp. 15-22.

Eberle, A.B., Lykke-Andersen, S., Muhlemann, O. and Jensen, T.H. (2009) 'SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells', *Nat Struct Mol Biol*, 16(1), pp. 49-55.

Ebert, B.L., Pretz, J., Bosco, J., Chang, C.Y., Tamayo, P., Galili, N., Raza, A., Root, D.E., Attar, E., Ellis, S.R. and Golub, T.R. (2008) 'Identification of RPS14 as a 5q- syndrome gene by RNA interference screen', *Nature*, 451(7176), pp. 335-9.

El Hage, A., Koper, M., Kufel, J. and Tollervey, D. (2008) 'Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast', *Genes Dev*, 22(8), pp. 1069-81.

Elela, S.A., Igel, H. and Ares, M., Jr. (1996) 'RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site', *Cell*, 85(1), pp. 115-24.

Ellis, S.R. (2014) 'Nucleolar stress in Diamond Blackfan anemia pathophysiology', *Biochim Biophys Acta*, 1842(6), pp. 765-8.

Engel, C., Sainsbury, S., Cheung, A.C., Kostrewa, D. and Cramer, P. (2013) 'RNA polymerase I structure and transcription regulation', *Nature*, 502(7473), pp. 650-5.

Enright, C.A., Maxwell, E.S., Eliceiri, G.L. and Sollner-Webb, B. (1996) '5'ETS rRNA processing facilitated by four small RNAs: U14, E3, U17, and U3', *RNA*, 2(11), pp. 1094-9.

Eppens, N.A., Faber, A.W., Rondaij, M., Jahangir, R.S., van Hemert, S., Vos, J.C., Venema, J. and Raue, H.A. (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 Res*, 30(19), pp. 4222-31.

Eppens, N.A., Rensen, S., Granneman, S., Raue, H.A. and Venema, J. (1999) 'The roles of Rrp5p in the synthesis of yeast 18S and 5.8S rRNA can be functionally and physically separated', *RNA*, 5(6), pp. 779-93.

Faber, A.W., Van Dijk, M., Raue, H.A. and Vos, J.C. (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*, 8(9), pp. 1095-101.

Faber, A.W., Vos, H.R., Vos, J.C. and Raue, H.A. (2006) '5'-end formation of yeast 5.8SL rRNA is an endonucleolytic event', *Biochem Biophys Res Commun*, 345(2), pp. 796-802.

Fairman, M.E., Maroney, P.A., Wang, W., Bowers, H.A., Gollnick, P., Nilsen, T.W. and Jankowsky, E. (2004) 'Protein displacement by DExH/D "RNA helicases" without duplex unwinding', *Science*, 304(5671), pp. 730-4.

Falaleeva, M., Pages, A., Matuszek, Z., Hidmi, S., Agranat-Tamir, L., Korotkov, K., Nevo, Y., Eyras, E., Sperling, R. and Stamm, S. (2016) 'Dual function of C/D box small nucleolar RNAs in rRNA modification and alternative pre-mRNA splicing', *Proc Natl Acad Sci U S A*, 113(12), pp. E1625-34.

Fang, F., Hoskins, J. and Butler, J.S. (2004) '5-fluorouracil enhances exosome-dependent accumulation of polyadenylated rRNAs', *Mol Cell Biol*, 24(24), pp. 10766-76.

Fang, F., Phillips, S. and Butler, J.S. (2005) 'Rat1p and Rai1p function with the nuclear exosome in the processing and degradation of rRNA precursors', *RNA*, 11(10), pp. 1571-8.

Farrar, J.E., Quarello, P., Fisher, R., O'Brien, K.A., Aspesi, A., Parrella, S., Henson, A.L., Seidel, N.E., Atsidaftos, E., Prakash, S., Bari, S., Garelli, E., Arceci, R.J., Dianzani, I., Ramenghi, U., Vlachos, A., Lipton, J.M., Bodine, D.M. and Ellis, S.R. (2014) 'Exploiting pre-rRNA processing in Diamond Blackfan anemia gene discovery and diagnosis', *Am J Hematol*, 89(10), pp. 985-91.

Fatica, A., Oeffinger, M., Dlakic, M. and Tollervey, D. (2003) 'Nob1p is required for cleavage of the 3' end of 18S rRNA', *Mol Cell Biol*, 23(5), pp. 1798-807.

Fayet-Lebaron, E., Atzorn, V., Henry, Y. and Kiss, T. (2009) '18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences', *EMBO J*, 28(9), pp. 1260-70.

Filipowicz, W. and Shatkin, A.J. (1983) 'Origin of splice junction phosphate in tRNAs processed by HeLa cell extract', *Cell*, 32(2), pp. 547-57.

Flygare, J., Aspesi, A., Bailey, J.C., Miyake, K., Caffrey, J.M., Karlsson, S. and Ellis, S.R. (2007) 'Human RPS19, the gene mutated in Diamond-Blackfan anemia, encodes a ribosomal protein required for the maturation of 40S ribosomal subunits', *Blood*, 109(3), pp. 980-6.

Fong, N., Brannan, K., Erickson, B., Kim, H., Cortazar, M.A., Sheridan, R.M., Nguyen, T., Karp, S. and Bentley, D.L. (2015) 'Effects of Transcription Elongation Rate and Xrn2 Exonuclease Activity on RNA Polymerase II Termination Suggest Widespread Kinetic Competition', *Mol Cell*, 60(2), pp. 256-67.

Fox, M.J., Gao, H., Smith-Kinnaman, W.R., Liu, Y. and Mosley, A.L. (2015) 'The exosome component Rrp6 is required for RNA polymerase II termination at specific targets of the Nrd1-Nab3 pathway', *PLoS Genet*, 11(2), p. e1004999.

Fox, M.J. and Mosley, A.L. (2016) 'Rrp6: Integrated roles in nuclear RNA metabolism and transcription termination', *Wiley Interdiscip Rev RNA*, 7(1), pp. 91-104.

Freed, E.F. and Baserga, S.J. (2010) 'The C-terminus of Utp4, mutated in childhood cirrhosis, is essential for ribosome biogenesis', *Nucleic Acids Res*, 38(14), pp. 4798-806.

Freed, E.F., Bleichert, F., Dutca, L.M. and Baserga, S.J. (2010) 'When ribosomes go bad: diseases of ribosome biogenesis', *Mol Biosyst*, 6(3), pp. 481-93.

Freed, E.F., Prieto, J.L., McCann, K.L., McStay, B. and Baserga, S.J. (2012) 'NOL11, implicated in the pathogenesis of North American Indian childhood cirrhosis, is required for pre-rRNA transcription and processing', *PLoS Genet*, 8(8), p. e1002892.

French, S.L., Osheim, Y.N., Cioci, F., Nomura, M. and Beyer, A.L. (2003) 'In exponentially growing Saccharomyces cerevisiae cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes', *Mol Cell Biol*, 23(5), pp. 1558-68.

Fromm, L., Falk, S., Flemming, D., Schuller, J.M., Thoms, M., Conti, E. and Hurt, E. (2017) 'Reconstitution of the complete pathway of ITS2 processing at the pre-ribosome', *Nat Commun*, 8(1), p. 1787.

Fulnecek, J. and Kovarik, A. (2007) 'Low abundant spacer 5S rRNA transcripts are frequently polyadenylated in Nicotiana', *Mol Genet Genomics*, 278(5), pp. 565-73.

Fumagalli, S., Di Cara, A., Neb-Gulati, A., Natt, F., Schwemberger, S., Hall, J., Babcock, G.F., Bernardi, R., Pandolfi, P.P. and Thomas, G. (2009) 'Absence of nucleolar disruption after impairment of 40S ribosome biogenesis reveals an rpL11-translation-dependent mechanism of p53 induction', *Nat Cell Biol*, 11(4), pp. 501-8.

Fumagalli, S., Ivanenkov, V.V., Teng, T. and Thomas, G. (2012) 'Suprainduction of p53 by disruption of 40S and 60S ribosome biogenesis leads to the activation of a novel G2/M checkpoint', *Genes Dev*, 26(10), pp. 1028-40.

Fumagalli, S. and Thomas, G. (2011) 'The role of p53 in ribosomopathies', *Semin Hematol*, 48(2), pp. 97-105.

Gadal, O., Strauss, D., Petfalski, E., Gleizes, P.E., Gas, N., Tollervey, D. and Hurt, E. (2002) 'Rlp7p is associated with 60S preribosomes, restricted to the granular component of the nucleolus, and required for pre-rRNA processing', *J Cell Biol*, 157(6), pp. 941-51.

Galani, K., Nissan, T.A., Petfalski, E., Tollervey, D. and Hurt, E. (2004) 'Rea1, a dynein-related nuclear AAA-ATPase, is involved in late rRNA processing and nuclear export of 60 S subunits', *J Biol Chem*, 279(53), pp. 55411-8.

Galardi, S., Fatica, A., Bachi, A., Scaloni, A., Presutti, C. and Bozzoni, I. (2002) 'Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA in vitro', *Mol Cell Biol*, 22(19), pp. 6663-8.

Gallagher, J.E., Dunbar, D.A., Granneman, S., Mitchell, B.M., Osheim, Y., Beyer, A.L. and Baserga, S.J. (2004) 'RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components', *Genes Dev*, 18(20), pp. 2506-17.

Gallenberger, M., Meinel, D.M., Kroeber, M., Wegner, M., Milkereit, P., Bosl, M.R. and Tamm, E.R. (2011) 'Lack of WDR36 leads to preimplantation embryonic lethality in mice and delays the formation of small subunit ribosomal RNA in human cells in vitro', *Hum Mol Genet*, 20(3), pp. 422-35.

Gamalinda, M. and Woolford, J.L., Jr. (2015) 'Paradigms of ribosome synthesis: Lessons learned from ribosomal proteins', *Translation (Austin)*, 3(1), p. e975018.

Ganot, P., Bortolin, M.L. and Kiss, T. (1997a) 'Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs', *Cell*, 89(5), pp. 799-809.

Ganot, P., Caizergues-Ferrer, M. and Kiss, T. (1997b) '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 Dev*, 11(7), pp. 941-56.

Garcia, I., Albring, M.J. and Uhlenbeck, O.C. (2012) 'Duplex destabilization by four ribosomal DEAD-box proteins', *Biochemistry*, 51(50), pp. 10109-18.

Gasse, L., Flemming, D. and Hurt, E. (2015) 'Coordinated Ribosomal ITS2 RNA Processing by the Las1 Complex Integrating Endonuclease, Polynucleotide Kinase, and Exonuclease Activities', *Mol Cell*, 60(5), pp. 808-815.

Geerlings, T.H., Vos, J.C. and Raue, H.A. (2000) 'The final step in the formation of 25S rRNA in Saccharomyces cerevisiae is performed by 5'-->3' exonucleases', *RNA*, 6(12), pp. 1698-703.

Genschik, P., Billy, E., Swianiewicz, M. and Filipowicz, W. (1997) 'The human RNA 3'-terminal phosphate cyclase is a member of a new family of proteins conserved in Eucarya, Bacteria and Archaea', *EMBO J*, 16(10), pp. 2955-67.

Gentilella, A., Kozma, S.C. and Thomas, G. (2015) 'A liaison between mTOR signaling, ribosome biogenesis and cancer', *Biochim Biophys Acta*, 1849(7), pp. 812-20.

Gerbi, S.A. and Borovjagin, A. (1997) 'U3 snoRNA may recycle through different compartments of the nucleolus', *Chromosoma*, 105(7-8), pp. 401-6.

Gerczei, T. and Correll, C.C. (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', *Proc Natl Acad Sci U S A*, 101(43), pp. 15301-6.

Gerczei, T., Shah, B.N., Manzo, A.J., Walter, N.G. and Correll, C.C. (2009) 'RNA chaperones stimulate formation and yield of the U3 snoRNA-Pre-rRNA duplexes needed for eukaryotic ribosome biogenesis', *J Mol Biol*, 390(5), pp. 991-1006.

Gerus, M., Bonnart, C., Caizergues-Ferrer, M., Henry, Y. and Henras, A.K. (2010) 'Evolutionarily conserved function of RRP36 in early cleavages of the pre-rRNA and production of the 40S ribosomal subunit', *Mol Cell Biol*, 30(5), pp. 1130-44.

Ghazal, G., Gagnon, J., Jacques, P.E., Landry, J.R., Robert, F. and Elela, S.A. (2009) 'Yeast RNase III triggers polyadenylation-independent transcription termination', *Mol Cell*, 36(1), pp. 99-109.

Giannakis, M., Mu, X.J., Shukla, S.A., Qian, Z.R., Cohen, O., Nishihara, R., Bahl, S., Cao, Y., Amin-Mansour, A., Yamauchi, M., Sukawa, Y., Stewart, C., Rosenberg, M., Mima, K., Inamura, K., Nosho, K., Nowak, J.A., Lawrence, M.S., Giovannucci, E.L., Chan, A.T., Ng, K., Meyerhardt, J.A., Van Allen, E.M., Getz, G., Gabriel, S.B., Lander, E.S., Wu, C.J., Fuchs, C.S., Ogino, S. and Garraway, L.A. (2016) 'Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma', *Cell Rep*, 17(4), p. 1206.

Glavan, F., Behm-Ansmant, I., Izaurralde, E. and Conti, E. (2006) 'Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex', *EMBO J*, 25(21), pp. 5117-25.

Goldfarb, K.C. and Cech, T.R. (2017) 'Targeted CRISPR disruption reveals a role for RNase MRP RNA in human preribosomal RNA processing', *Genes Dev*, 31(1), pp. 59-71.

Gonzales, B., Henning, D., So, R.B., Dixon, J., Dixon, M.J. and Valdez, B.C. (2005) 'The Treacher Collins syndrome (TCOF1) gene product is involved in pre-rRNA methylation', *Hum Mol Genet*, 14(14), pp. 2035-43.

Goudarzi, K.M. and Lindstrom, M.S. (2016) 'Role of ribosomal protein mutations in tumor development (Review)', *Int J Oncol*, 48(4), pp. 1313-24.

Grandi, P., Rybin, V., Bassler, J., Petfalski, E., Strauss, D., Marzioch, M., Schafer, T., Kuster, B., Tschochner, H., Tollervey, D., Gavin, A.C. and 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', *Mol Cell*, 10(1), pp. 105-15.

Granneman, S. and Baserga, S.J. (2004) 'Ribosome biogenesis: of knobs and RNA processing', *Exp Cell Res*, 296(1), pp. 43-50.

Granneman, S., Bernstein, K.A., Bleichert, F. and Baserga, S.J. (2006) 'Comprehensive mutational analysis of yeast DEXD/H box RNA helicases required for small ribosomal subunit synthesis', *Mol Cell Biol*, 26(4), pp. 1183-94.

Granneman, S., Gallagher, J.E., Vogelzangs, J., Horstman, W., van Venrooij, W.J., Baserga, S.J. and Pruijn, G.J. (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 Res*, 31(7), pp. 1877-87.

Granneman, S., Petfalski, E. and Tollervey, D. (2011) 'A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease', *EMBO J*, 30(19), pp. 4006-19.

Granneman, S., Pruijn, G.J., Horstman, W., van Venrooij, W.J., Luhrmann, R. and Watkins, N.J. (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', *J Biol Chem*, 277(50), pp. 48490-500.

Granneman, S., Vogelzangs, J., Luhrmann, R., van Venrooij, W.J., Pruijn, G.J. and Watkins, N.J. (2004) 'Role of pre-rRNA base pairing and 80S complex formation in subnucleolar localization of the U3 snoRNP', *Mol Cell Biol*, 24(19), pp. 8600-10.

Griffin, J.N., Sondalle, S.B., Del Viso, F., Baserga, S.J. and Khokha, M.K. (2015) 'The ribosome biogenesis factor Nol11 is required for optimal rDNA transcription and craniofacial development in Xenopus', *PLoS Genet*, 11(3), p. e1005018.

Grosshans, H., Deinert, K., Hurt, E. and Simos, G. (2001) 'Biogenesis of the signal recognition particle (SRP) involves import of SRP proteins into the nucleolus, assembly with the SRP-RNA, and Xpo1p-mediated export', *J Cell Biol*, 153(4), pp. 745-62.

Gudipati, R.K., Xu, Z., Lebreton, A., Seraphin, B., Steinmetz, L.M., Jacquier, A. and Libri, D. (2012) 'Extensive degradation of RNA precursors by the exosome in wild-type cells', *Mol Cell*, 48(3), pp. 409-21.

Hadjiolova, K.V., Georgiev, O.I. and Hadjiolov, A.A. (1984a) 'Excess 5'-terminal sequences in the rat nucleolar 28S ribosomal RNA', *Exp Cell Res*, 153(1), pp. 266-9.

Hadjiolova, K.V., Georgiev, O.I., Nosikov, V.V. and Hadjiolov, A.A. (1984b) 'Localization and structure of endonuclease cleavage sites involved in the processing of the rat 32S precursor to ribosomal RNA', *Biochem J*, 220(1), pp. 105-16.

Hadjiolova, K.V., Georgiev, O.I., Nosikov, V.V. and Hadjiolov, A.A. (1984c) 'Mapping of the major early endonuclease cleavage site of the rat precursor to rRNA within the internal transcribed spacer sequence of rDNA', *Biochim Biophys Acta*, 782(2), pp. 195-201.

Hadjiolova, K.V., Nicoloso, M., Mazan, S., Hadjiolov, A.A. and Bachellerie, J.P. (1993) 'Alternative pre-rRNA processing pathways in human cells and their alteration by cycloheximide inhibition of protein synthesis', *Eur J Biochem*, 212(1), pp. 211-5.

Hayashi, T.T. and MacFarlane, K. (1979) 'Comparison of endogenous and exogenous RNA primers of poly(U) polymerase in rat hepatic ribosomes', *Biochem J*, 177(3), pp. 895-902.

He, X., Li, Y., Dai, M.S. and Sun, X.X. (2016) 'Ribosomal protein L4 is a novel regulator of the MDM2-p53 loop', *Oncotarget*, 7(13), pp. 16217-26.

Heiss, N.S., Knight, S.W., Vulliamy, T.J., Klauck, S.M., Wiemann, S., Mason, P.J., Poustka, A. and Dokal, I. (1998) 'X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions', *Nat Genet*, 19(1), pp. 32-8.

Helm, M. (2006) 'Post-transcriptional nucleotide modification and alternative folding of RNA', *Nucleic Acids Res*, 34(2), pp. 721-33.

Henderson, A.S., Warburton, D. and Atwood, K.C. (1972) 'Location of ribosomal DNA in the human chromosome complement', *Proc Natl Acad Sci U S A*, 69(11), pp. 3394-8.

Henras, A., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gelugne, J.P. and Caizergues-Ferrer, M. (1998) 'Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs', *EMBO J*, 17(23), pp. 7078-90.

Henras, A.K., Bertrand, E. and Chanfreau, G. (2004a) 'A cotranscriptional model for 3'-end processing of the Saccharomyces cerevisiae pre-ribosomal RNA precursor', *RNA*, 10(10), pp. 1572-85.

Henras, A.K., Dez, C. and Henry, Y. (2004b) 'RNA structure and function in C/D and H/ACA s(no)RNPs', *Curr Opin Struct Biol*, 14(3), pp. 335-43.

Henras, A.K., Plisson-Chastang, C., O'Donohue, M.F., Chakraborty, A. and Gleizes, P.E. (2015) 'An overview of pre-ribosomal RNA processing in eukaryotes', *Wiley Interdiscip Rev RNA*, 6(2), pp. 225-42.

Henras, A.K., Soudet, J., Gerus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A. and Henry, Y. (2008) 'The post-transcriptional steps of eukaryotic ribosome biogenesis', *Cell Mol Life Sci*, 65(15), pp. 2334-59.

Henry, Y., Wood, H., Morrissey, J.P., Petfalski, E., Kearsey, S. and Tollervey, D. (1994) 'The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site', *EMBO J*, 13(10), pp. 2452-63.

Hierlmeier, T., Merl, J., Sauert, M., Perez-Fernandez, J., Schultz, P., Bruckmann, A., Hamperl, S., Ohmayer, U., Rachel, R., Jacob, A., Hergert, K., Deutzmann, R., Griesenbeck, J., Hurt, E., Milkereit, P., Bassler, J. and Tschochner, H. (2013) 'Rrp5p, Noc1p and Noc2p form a protein module which is part of early large ribosomal subunit precursors in S. cerevisiae', *Nucleic Acids Res*, 41(2), pp. 1191-210.

Hoareau-Aveilla, C., Fayet-Lebaron, E., Jady, B.E., Henras, A.K. and Kiss, T. (2012) 'Utp23p is required for dissociation of snR30 small nucleolar RNP from preribosomal particles', *Nucleic Acids Res*, 40(8), pp. 3641-52.

Hodnett, J.L. and Busch, H. (1968) 'Isolation and characterization of uridylic acid-rich 7 S ribonucleic acid of rat liver nuclei', *J Biol Chem*, 243(24), pp. 6334-42.

Holub, P., Lalakova, J., Cerna, H., Pasulka, J., Sarazova, M., Hrazdilova, K., Arce, M.S., Hobor, F., Stefl, R. and Vanacova, S. (2012) 'Air2p is critical for the assembly and RNA-binding of the TRAMP complex and the KOW domain of Mtr4p is crucial for exosome activation', *Nucleic Acids Res*, 40(12), pp. 5679-93.

Horn, D.M., Mason, S.L. and Karbstein, K. (2011) 'Rcl1 protein, a novel nuclease for 18 S ribosomal RNA production', *J Biol Chem*, 286(39), pp. 34082-7.

Horn, H.F. and Vousden, K.H. (2008) 'Cooperation between the ribosomal proteins L5 and L11 in the p53 pathway', *Oncogene*, 27(44), pp. 5774-84.

Houseley, J., LaCava, J. and Tollervey, D. (2006) 'RNA-quality control by the exosome', *Nat Rev Mol Cell Biol*, 7(7), pp. 529-39.

Hozumi, N., Haruna, I., Watanabe, I., Mikoshiba, K. and Tsukada, Y. (1975) 'Poly(U) polymerase in rat brain', *Nature*, 256(5515), pp. 337-9.

Huang, S. (2002) 'Building an efficient factory: where is pre-rRNA synthesized in the nucleolus?', *J Cell Biol*, 157(5), pp. 739-41.

Hughes, J.M. (1996) 'Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA', *J Mol Biol*, 259(4), pp. 645-54.

Hughes, J.M. and Ares, M., Jr. (1991) 'Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA', *EMBO J*, 10(13), pp. 4231-9.

Hunziker, M., Barandun, J., Petfalski, E., Tan, D., Delan-Forino, C., Molloy, K.R., Kim, K.H., Dunn-Davies, H., Shi, Y., Chaker-Margot, M., Chait, B.T., Walz, T., Tollervey, D. and Klinge, S. (2016) 'UtpA and UtpB chaperone nascent pre-ribosomal RNA and U3 snoRNA to initiate eukaryotic ribosome assembly', *Nat Commun*, 7, p. 12090.

Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Goncalves, E., Barthorpe, S., Lightfoot, H., Cokelaer, T., Greninger, P., van Dyk, E., Chang, H., de Silva, H., Heyn, H., Deng, X., Egan, R.K., Liu, Q., Mironenko, T., Mitropoulos, X., Richardson, L., Wang, J., Zhang, T., Moran, S., Sayols, S., Soleimani, M., Tamborero, D., Lopez-Bigas, N., Ross-Macdonald, P., Esteller, M., Gray, N.S., Haber, D.A., Stratton, M.R., Benes, C.H., Wessels, L.F.A., Saez-Rodriguez, J., McDermott, U. and Garnett, M.J. (2016) 'A Landscape of Pharmacogenomic Interactions in Cancer', *Cell*, 166(3), pp. 740-754.

Ishikawa, H., Yoshikawa, H., Izumikawa, K., Miura, Y., Taoka, M., Nobe, Y., Yamauchi, Y., Nakayama, H., Simpson, R.J., Isobe, T. and Takahashi, N. (2017) 'Poly(A)-specific ribonuclease regulates the processing of small-subunit rRNAs in human cells', *Nucleic Acids Res*, 45(6), pp. 3437-3447.

Isken, O. and Maquat, L.E. (2007) 'Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function', *Genes Dev*, 21(15), pp. 1833-56.

Jaako, P., Debnath, S., Olsson, K., Zhang, Y., Flygare, J., Lindstrom, M.S., Bryder, D. and Karlsson, S. (2015) 'Disruption of the 5S RNP-Mdm2 interaction significantly improves the erythroid defect in a mouse model for Diamond-Blackfan anemia', *Leukemia*, 29(11), pp. 2221-9.

Jacobson, M.R. and Pederson, T. (1998) 'Localization of signal recognition particle RNA in the nucleolus of mammalian cells', *Proc Natl Acad Sci U S A*, 95(14), pp. 7981-6.

Jankowsky, E., Gross, C.H., Shuman, S. and Pyle, A.M. (2001) 'Active disruption of an RNA-protein interaction by a DExH/D RNA helicase', *Science*, 291(5501), pp. 121-5.

Januszyk, K. and Lima, C.D. (2014) 'The eukaryotic RNA exosome', *Curr Opin Struct Biol*, 24, pp. 132-40.

Januszyk, K., Liu, Q. and Lima, C.D. (2011) 'Activities of human RRP6 and structure of the human RRP6 catalytic domain', *RNA*, 17(8), pp. 1566-77.

Jenner, L., Melnikov, S., Garreau de Loubresse, N., Ben-Shem, A., Iskakova, M., Urzhumtsev, A., Meskauskas, A., Dinman, J., Yusupova, G. and Yusupov, M. (2012) 'Crystal structure of the 80S yeast ribosome', *Curr Opin Struct Biol*, 22(6), pp. 759-67.

Jia, H., Wang, X., Anderson, J.T. and Jankowsky, E. (2012) 'RNA unwinding by the Trf4/Air2/Mtr4 polyadenylation (TRAMP) complex', *Proc Natl Acad Sci U S A*, 109(19), pp. 7292-7.

Johnson, A.W. (1997) 'Rat1p and Xrn1p are functionally interchangeable exoribonucleases that are restricted to and required in the nucleus and cytoplasm, respectively', *Mol Cell Biol*, 17(10), pp. 6122-30.

Jorgensen, P., Rupes, I., Sharom, J.R., Schneper, L., Broach, J.R. and Tyers, M. (2004) 'A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size', *Genes Dev*, 18(20), pp. 2491-505.

Karbstein, K., Jonas, S. and Doudna, J.A. (2005) 'An essential GTPase promotes assembly of preribosomal RNA processing complexes', *Mol Cell*, 20(4), pp. 633-43.

Kass, S., Craig, N. and Sollner-Webb, B. (1987) 'Primary processing of mammalian rRNA involves two adjacent cleavages and is not species specific', *Mol Cell Biol*, 7(8), pp. 2891-8.

Kass, S., Tyc, K., Steitz, J.A. and Sollner-Webb, B. (1990) 'The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing', *Cell*, 60(6), pp. 897-908.

Kawauchi, J., Mischo, H., Braglia, P., Rondon, A. and Proudfoot, N.J. (2008) 'Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination', *Genes Dev*, 22(8), pp. 1082-92.

Kempers-Veenstra, A.E., Oliemans, J., Offenberg, H., Dekker, A.F., Piper, P.W., Planta, R.J. and Klootwijk, J. (1986) '3'-End formation of transcripts from the yeast rRNA operon', *EMBO J*, 5(10), pp. 2703-10.

Kenna, M., Stevens, A., McCammon, M. and Douglas, M.G. (1993) 'An essential yeast gene with homology to the exonuclease-encoding XRN1/KEM1 gene also encodes a protein with exoribonuclease activity', *Mol Cell Biol*, 13(1), pp. 341-50.

Khoshnevis, S., Askenasy, I., Johnson, M.C., Dattolo, M.D., Young-Erdos, C.L., Stroupe, M.E. and Karbstein, K. (2016) 'The DEAD-box Protein Rok1 Orchestrates 40S and 60S Ribosome Assembly by Promoting the Release of Rrp5 from Pre-40S Ribosomes to Allow for 60S Maturation', *PLoS Biol*, 14(6), p. e1002480.

Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F. and Buratowski, S. (2004) 'The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II', *Nature*, 432(7016), pp. 517-22.

Kim, T.H., Leslie, P. and Zhang, Y. (2014) 'Ribosomal proteins as unrevealed caretakers for cellular stress and genomic instability', *Oncotarget*, 5(4), pp. 860-71.

Kiss-Laszlo, Z., Henry, Y. and Kiss, T. (1998) 'Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA', *EMBO J*, 17(3), pp. 797-807.

Kiss, T., Fayet-Lebaron, E. and Jady, B.E. (2010) 'Box H/ACA small ribonucleoproteins', *Mol Cell*, 37(5), pp. 597-606.

Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S. and Ban, N. (2011) 'Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6', *Science*, 334(6058), pp. 941-8.

Klinge, S., Voigts-Hoffmann, F., Leibundgut, M. and Ban, N. (2012) 'Atomic structures of the eukaryotic ribosome', *Trends Biochem Sci*, 37(5), pp. 189-98.

Kobayashi, T., Heck, D.J., Nomura, M. and 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 Dev*, 12(24), pp. 3821-30.

Koberna, K., Malinsky, J., Pliss, A., Masata, M., Vecerova, J., Fialova, M., Bednar, J. and Raska, I. (2002) 'Ribosomal genes in focus: new transcripts label the dense fibrillar components and form clusters indicative of "Christmas trees" in situ', *J Cell Biol*, 157(5), pp. 743-8.

Konikkat, S. and Woolford, J.L., Jr. (2017) 'Principles of 60S ribosomal subunit assembly emerging from recent studies in yeast', *Biochem J*, 474(2), pp. 195-214.

Kornprobst, M., Turk, M., Kellner, N., Cheng, J., Flemming, D., Kos-Braun, I., Kos, M., Thoms, M., Berninghausen, O., Beckmann, R. and Hurt, E. (2016) 'Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome', *Cell*, 166(2), pp. 380-393.

Kos-Braun, I.C., Jung, I. and Kos, M. (2017) 'Tor1 and CK2 kinases control a switch between alternative ribosome biogenesis pathways in a growth-dependent manner', *PLoS Biol*, 15(3), p. e2000245.

Kos, M. and Tollervey, D. (2005) 'The Putative RNA Helicase Dbp4p Is Required for Release of the U14 snoRNA from Preribosomes in Saccharomyces cerevisiae', *Mol Cell*, 20(1), pp. 53-64.

Kos, M. and Tollervey, D. (2010) 'Yeast pre-rRNA processing and modification occur cotranscriptionally', *Mol Cell*, 37(6), pp. 809-20.

Kressler, D., de la Cruz, J., Rojo, M. and Linder, P. (1997) 'Fal1p is an essential DEAD-box protein involved in 40S-ribosomal-subunit biogenesis in Saccharomyces cerevisiae', *Mol Cell Biol*, 17(12), pp. 7283-94.

Kressler, D., Linder, P. and de La Cruz, J. (1999) 'Protein trans-acting factors involved in ribosome biogenesis in Saccharomyces cerevisiae', *Mol Cell Biol*, 19(12), pp. 7897-912.

Krogan, N.J., Peng, W.T., Cagney, G., Robinson, M.D., Haw, R., Zhong, G., Guo, X., Zhang, X., Canadien, V., Richards, D.P., Beattie, B.K., Lalev, A., Zhang, W., Davierwala, A.P., Mnaimneh, S., Starostine, A., Tikuisis, A.P., Grigull, J., Datta, N., Bray, J.E., Hughes, T.R., Emili, A. and Greenblatt, J.F. (2004) 'High-definition macromolecular composition of yeast RNA-processing complexes', *Mol Cell*, 13(2), pp. 225-39.

Kuai, L., Fang, F., Butler, J.S. and Sherman, F. (2004) 'Polyadenylation of rRNA in Saccharomyces cerevisiae', *Proc Natl Acad Sci U S A*, 101(23), pp. 8581-6.

Kudla, G., Granneman, S., Hahn, D., Beggs, J.D. and Tollervey, D. (2011) 'Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast', *Proc Natl Acad Sci U S A*, 108(24), pp. 10010-5.

Kufel, J., Dichtl, B. and Tollervey, D. (1999) 'Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3' ETS but not the 5' ETS', *RNA*, 5(7), pp. 909-17.

Kuhn, H., Hierlmeier, T., Merl, J., Jakob, S., Aguissa-Toure, A.H., Milkereit, P. and Tschochner, H. (2009) 'The Noc-domain containing C-terminus of Noc4p mediates both formation of the Noc4p-Nop14p submodule and its incorporation into the SSU processome', *PLoS One*, 4(12), p. e8370.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A. and Tollervey, D. (2005) 'RNA degradation by the exosome is promoted by a nuclear polyadenylation complex', *Cell*, 121(5), pp. 713-24.

Lafontaine, D.L., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. and Tollervey, D. (1998) 'The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase', *Genes Dev*, 12(4), pp. 527-37.

Lafontaine, D.L. and Tollervey, D. (2001) 'The function and synthesis of ribosomes', *Nat Rev Mol Cell Biol*, 2(7), pp. 514-20.

Lamanna, A.C. and Karbstein, K. (2009) 'Nob1 binds the single-stranded cleavage site D at the 3'-end of 18S rRNA with its PIN domain', *Proc Natl Acad Sci U S A*, 106(34), pp. 14259-64.

Lamanna, A.C. and Karbstein, K. (2011) 'An RNA conformational switch regulates pre-18S rRNA cleavage', *J Mol Biol*, 405(1), pp. 3-17.

Lange, H., Sement, F.M. and Gagliardi, D. (2011) 'MTR4, a putative RNA helicase and exosome co-factor, is required for proper rRNA biogenesis and development in Arabidopsis thaliana', *Plant J*, 68(1), pp. 51-63.

Langhendries, J.L., Nicolas, E., Doumont, G., Goldman, S. and Lafontaine, D.L. (2016) 'The human box C/D snoRNAs U3 and U8 are required for pre-rRNA processing and tumorigenesis', *Oncotarget*, 7(37), pp. 59519-59534.

Lapik, Y.R., Fernandes, C.J., Lau, L.F. and Pestov, D.G. (2004) 'Physical and functional interaction between Pes1 and Bop1 in mammalian ribosome biogenesis', *Mol Cell*, 15(1), pp. 17-29.

Lazdins, I.B., Delannoy, M. and Sollner-Webb, B. (1997) 'Analysis of nucleolar transcription and processing domains and pre-rRNA movements by in situ hybridization', *Chromosoma*, 105(7-8), pp. 481-95.

Leary, D.J., Terns, M.P. and Huang, S. (2004) 'Components of U3 snoRNA-containing complexes shuttle between nuclei and the cytoplasm and differentially localize in nucleoli: implications for assembly and function', *Mol Biol Cell*, 15(1), pp. 281-93.

Lebaron, S., Schneider, C., van Nues, R.W., Swiatkowska, A., Walsh, D., Bottcher, B., Granneman, S., Watkins, N.J. and Tollervey, D. (2012) 'Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits', *Nat Struct Mol Biol*, 19(8), pp. 744-53.

Lebaron, S., Segerstolpe, A., French, S.L., Dudnakova, T., de Lima Alves, F., Granneman, S., Rappsilber, J., Beyer, A.L., Wieslander, L. and Tollervey, D. (2013) 'Rrp5 binding at multiple sites coordinates pre-rRNA processing and assembly', *Mol Cell*, 52(5), pp. 707-19.

Lebreton, A., Tomecki, R., Dziembowski, A. and Seraphin, B. (2008) 'Endonucleolytic RNA cleavage by a eukaryotic exosome', *Nature*, 456(7224), pp. 993-6.

Lee, S.J. and Baserga, S.J. (1999) 'Imp3p and Imp4p, two specific components of the U3 small nucleolar ribonucleoprotein that are essential for pre-18S rRNA processing', *Mol Cell Biol*, 19(8), pp. 5441-52.

Lei, E.P. and Silver, P.A. (2002) 'Protein and RNA export from the nucleus', *Dev Cell*, 2(3), pp. 261-72.

Lemay, V., Hossain, A., Osheim, Y.N., Beyer, A.L. and Dragon, F. (2011) 'Identification of novel proteins associated with yeast snR30 small nucleolar RNA', *Nucleic Acids Res*, 39(22), pp. 9659-70.

Lempiainen, H. and Shore, D. (2009) 'Growth control and ribosome biogenesis', *Curr Opin Cell Biol*, 21(6), pp. 855-63.

Lempicki, R.A., Jarmolowski, A., Huang, G.Y., Li, H.V. and Fournier, M.J. (1990) 'Mutations in conserved domains of U14 RNA impair 18S ribosomal RNA production in Saccharomyces cerevisiae', *Mol Biol Rep*, 14(2-3), pp. 119-20.

Levin, I., Schwarzenbacher, R., Page, R., Abdubek, P., Ambing, E., Biorac, T., Brinen, L.S., Campbell, J., Canaves, J.M., Chiu, H.J., Dai, X., Deacon, A.M., DiDonato, M., Elsliger, M.A., Floyd, R., Godzik, A., Grittini, C., Grzechnik, S.K., Hampton, E., Jaroszewski, L., Karlak, C., Klock, H.E., Koesema, E., Kovarik, J.S., Kreusch, A., Kuhn, P., Lesley, S.A., McMullan, D., McPhillips, T.M., Miller, M.D., Morse, A., Moy, K., Ouyang, J., Quijano, K., Reyes, R., Rezezadeh, F., Robb, A., Sims, E., Spraggon, G., Stevens, R.C., van den Bedem, H., Velasquez, J., Vincent, J., von Delft, F., Wang, X., West, B., Wolf, G., Xu, Q., Hodgson, K.O., Wooley, J. and Wilson, I.A. (2004) 'Crystal structure of a PIN (PilT N-terminus) domain (AF0591) from Archaeoglobus fulgidus at 1.90 A resolution', *Proteins*, 56(2), pp. 404-8.

Lewis, J.D. and Tollervey, D. (2000) 'Like attracts like: getting RNA processing together in the nucleus', *Science*, 288(5470), pp. 1385-9.

Li, H.D., Zagorski, J. and Fournier, M.J. (1990) 'Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in Saccharomyces cerevisiae', *Mol Cell Biol*, 10(3), pp. 1145-52.

Liang, W.Q. and Fournier, M.J. (1995) 'U14 base-pairs with 18S rRNA: a novel snoRNA interaction required for rRNA processing', *Genes Dev*, 9(19), pp. 2433-43.

Liang, W.Q. and Fournier, M.J. (1997) 'Synthesis of functional eukaryotic ribosomal RNAs in trans: development of a novel in vivo rDNA system for dissecting ribosome biogenesis', *Proc Natl Acad Sci U S A*, 94(7), pp. 2864-8.

Liang, X.H. and Fournier, M.J. (2006) 'The helicase Has1p is required for snoRNA release from pre-rRNA', *Mol Cell Biol*, 26(20), pp. 7437-50.

Liang, X.H., Liu, Q., Liu, Q., King, T.H. and Fournier, M.J. (2010) 'Strong dependence between functional domains in a dual-function snoRNA infers coupling of rRNA processing and modification events', *Nucleic Acids Res*, 38(10), pp. 3376-87.

Lin, J., Lai, S., Jia, R., Xu, A., Zhang, L., Lu, J. and Ye, K. (2011) 'Structural basis for site-specific ribose methylation by box C/D RNA protein complexes', *Nature*, 469(7331), pp. 559-63.

Lin, J., Lu, J., Feng, Y., Sun, M. and Ye, K. (2013) 'An RNA-binding complex involved in ribosome biogenesis contains a protein with homology to tRNA CCA-adding enzyme', *PLoS Biol*, 11(10), p. e1001669.

Linder, P. and Jankowsky, E. (2011) 'From unwinding to clamping - the DEAD box RNA helicase family', *Nat Rev Mol Cell Biol*, 12(8), pp. 505-16.

Lipton, J.M. and Ellis, S.R. (2010) 'Diamond Blackfan anemia 2008-2009: broadening the scope of ribosome biogenesis disorders', *Curr Opin Pediatr*, 22(1), pp. 12-9.

Liu, P.C. and Thiele, D.J. (2001) 'Novel stress-responsive genes EMG1 and NOP14 encode conserved, interacting proteins required for 40S ribosome biogenesis', *Mol Biol Cell*, 12(11), pp. 3644-57.

Liu, Q., Greimann, J.C. and Lima, C.D. (2006) 'Reconstitution, activities, and structure of the eukaryotic RNA exosome', *Cell*, 127(6), pp. 1223-37.

Lu, J., Sun, M. and Ye, K. (2013) 'Structural and functional analysis of Utp23, a yeast ribosome synthesis factor with degenerate PIN domain', *RNA*, 19(12), pp. 1815-24.

Lubben, B., Marshallsay, C., Rottmann, N. and 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 Res*, 21(23), pp. 5377-85.

Lund, E. and Dahlberg, J.E. (1992) 'Cyclic 2',3'-phosphates and nontemplated nucleotides at the 3' end of spliceosomal U6 small nuclear RNA's', *Science*, 255(5042), pp. 327-30.

Luo, W., Johnson, A.W. and Bentley, D.L. (2006) 'The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model', *Genes Dev*, 20(8), pp. 954-65.

Lygerou, Z., Allmang, C., Tollervey, D. and Seraphin, B. (1996) 'Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP in vitro', *Science*, 272(5259), pp. 268-70.

Mager, W.H., Planta, R.J., Ballesta, J.G., Lee, J.C., Mizuta, K., Suzuki, K., Warner, J.R. and Woolford, J. (1997) 'A new nomenclature for the cytoplasmic ribosomal proteins of Saccharomyces cerevisiae', *Nucleic Acids Res*, 25(24), pp. 4872-5.

Makino, D.L., Baumgartner, M. and Conti, E. (2013) 'Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex', *Nature*, 495(7439), pp. 70-5.

Marmier-Gourrier, N., Clery, A., Schlotter, F., Senty-Segault, V. and Branlant, C. (2011) 'A second base pair interaction between U3 small nucleolar RNA and the 5'-ETS region is required for early cleavage of the yeast pre-ribosomal RNA', *Nucleic Acids Res*, 39(22), pp. 9731-45.

Marneros, A.G. (2013) 'BMS1 is mutated in aplasia cutis congenita', *PLoS Genet*, 9(6), p. e1003573.

Martin, R., Hackert, P., Ruprecht, M., Simm, S., Bruning, L., Mirus, O., Sloan, K.E., Kudla, G., Schleiff, E. and Bohnsack, M.T. (2014) 'A pre-ribosomal RNA interaction network involving snoRNAs and the Rok1 helicase', *RNA*, 20(8), pp. 1173-82.

Marz, M. and Stadler, P.F. (2009) 'Comparative analysis of eukaryotic U3 snoRNA', *RNA Biol*, 6(5), pp. 503-7.

Mattijssen, S., Welting, T.J. and Pruijn, G.J. (2010) 'RNase MRP and disease', *Wiley Interdiscip Rev RNA*, 1(1), pp. 102-16.

Meek, D.W. (2015) 'Regulation of the p53 response and its relationship to cancer', *Biochem J*, 469(3), pp. 325-46.

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G. and Yusupov, M. (2012) 'One core, two shells: bacterial and eukaryotic ribosomes', *Nat Struct Mol Biol*, 19(6), pp. 560-7.

Memet, I., Doebele, C., Sloan, K.E. and Bohnsack, M.T. (2017) 'The G-patch protein NF-kappaB-repressing factor mediates the recruitment of the exonuclease XRN2 and activation of the RNA helicase DHX15 in human ribosome biogenesis', *Nucleic Acids Res*, 45(9), pp. 5359-5374.

Meyer, B., Wurm, J.P., Sharma, S., Immer, C., Pogoryelov, D., Kotter, P., Lafontaine, D.L., Wohnert, J. and Entian, K.D. (2016) 'Ribosome biogenesis factor Tsr3 is the aminocarboxypropyl transferase responsible for 18S rRNA hypermodification in yeast and humans', *Nucleic Acids Res*, 44(9), pp. 4304-16.

Michot, B., Joseph, N., Mazan, S. and Bachellerie, J.P. (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 Res*, 27(11), pp. 2271-82.

Midtgaard, S.F., Assenholt, J., Jonstrup, A.T., Van, L.B., Jensen, T.H. and Brodersen, D.E. (2006) 'Structure of the nuclear exosome component Rrp6p reveals an interplay between the active site and the HRDC domain', *Proc Natl Acad Sci U S A*, 103(32), pp. 11898-903.

Miki, T.S., Richter, H., Ruegger, S. and Grosshans, H. (2014) 'PAXT-1 promotes XRN2 activity by stabilizing it through a conserved domain', *Mol Cell*, 53(2), pp. 351-60.

Milchev, G.I. and Hadjiolov, A.A. (1978) 'Association of poly(A) and poly(U) polymerases with cytoplasmic ribosomes', *Eur J Biochem*, 84(1), pp. 113-21.

Miles, T.D., Jakovljevic, J., Horsey, E.W., Harnpicharnchai, P., Tang, L. and Woolford, J.L., Jr. (2005) 'Ytm1, Nop7, and Erb1 form a complex necessary for maturation of yeast 66S preribosomes', *Mol Cell Biol*, 25(23), pp. 10419-32.

Miller, O.L., Jr. and Beatty, B.R. (1969) 'Visualization of nucleolar genes', *Science*, 164(3882), pp. 955-7.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. and Tollervey, D. (1997) 'The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases', *Cell*, 91(4), pp. 457-66.

Mitchell, P., Petfalski, E. and Tollervey, D. (1996) 'The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism', *Genes Dev*, 10(4), pp. 502-13.

Momand, J., Wu, H.H. and Dasgupta, G. (2000) 'MDM2--master regulator of the p53 tumor suppressor protein', *Gene*, 242(1-2), pp. 15-29.

Momand, J., Zambetti, G.P., Olson, D.C., George, D. and Levine, A.J. (1992) 'The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation', *Cell*, 69(7), pp. 1237-45.

Montellese, C., Montel-Lehry, N., Henras, A.K., Kutay, U., Gleizes, P.E. and O'Donohue, M.F. (2017) 'Poly(A)-specific ribonuclease is a nuclear ribosome biogenesis factor involved in human 18S rRNA maturation', *Nucleic Acids Res*, 45(11), pp. 6822-6836.

Morrissey, J.P. and Tollervey, D. (1993) 'Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis', *Mol Cell Biol*, 13(4), pp. 2469-77.

Morrissey, J.P. and Tollervey, D. (1997) 'U14 small nucleolar RNA makes multiple contacts with the pre-ribosomal RNA', *Chromosoma*, 105(7-8), pp. 515-22.

Mougey, E.B., O'Reilly, M., Osheim, Y., Miller, O.L., Jr., Beyer, A. and Sollner-Webb, B. (1993) 'The terminal balls characteristic of eukaryotic rRNA transcription units in chromatin spreads are rRNA processing complexes', *Genes Dev*, 7(8), pp. 1609-19.

Mullen, T.E. and Marzluff, W.F. (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 Dev*, 22(1), pp. 50-65.

Mullineux, S.T. and Lafontaine, D.L. (2012) 'Mapping the cleavage sites on mammalian pre-rRNAs: where do we stand?', *Biochimie*, 94(7), pp. 1521-32.

Nagarajan, V.K., Jones, C.I., Newbury, S.F. and Green, P.J. (2013) 'XRN 5'-->3' exoribonucleases: structure, mechanisms and functions', *Biochim Biophys Acta*, 1829(6-7), pp. 590-603.

Narla, A. and Ebert, B.L. (2010) 'Ribosomopathies: human disorders of ribosome dysfunction', *Blood*, 115(16), pp. 3196-205.

Nazar, R.N. (2004) 'Ribosomal RNA processing and ribosome biogenesis in eukaryotes', *IUBMB Life*, 56(8), pp. 457-65.

Nemeth, A., Perez-Fernandez, J., Merkl, P., Hamperl, S., Gerber, J., Griesenbeck, J. and Tschochner, H. (2013) 'RNA polymerase I termination: Where is the end?', *Biochim Biophys Acta*, 1829(3-4), pp. 306-17.

Ni, J., Tien, A.L. and Fournier, M.J. (1997) 'Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA', *Cell*, 89(4), pp. 565-73.

Nicolas, E., Parisot, P., Pinto-Monteiro, C., de Walque, R., De Vleeschouwer, C. and Lafontaine, D.L. (2016) 'Involvement of human ribosomal proteins in nucleolar structure and p53-dependent nucleolar stress', *Nat Commun*, 7, p. 11390.

O'Day, C.L., Chavanikamannil, F. and Abelson, J. (1996) '18S rRNA processing requires the RNA helicase-like protein Rrp3', *Nucleic Acids Res*, 24(16), pp. 3201-7.

O'Donohue, M.F., Choesmel, V., Faubladier, M., Fichant, G. and Gleizes, P.E. (2010) 'Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits', *J Cell Biol*, 190(5), pp. 853-66.

Oeffinger, M., Leung, A., Lamond, A. and Tollervey, D. (2002) 'Yeast Pescadillo is required for multiple activities during 60S ribosomal subunit synthesis', *RNA*, 8(5), pp. 626-36.

Oeffinger, M. and Tollervey, D. (2003) 'Yeast Nop15p is an RNA-binding protein required for pre-rRNA processing and cytokinesis', *EMBO J*, 22(24), pp. 6573-83.

Oeffinger, M., Zenklusen, D., Ferguson, A., Wei, K.E., El Hage, A., Tollervey, D., Chait, B.T., Singer, R.H. and Rout, M.P. (2009) 'Rrp17p is a eukaryotic exonuclease required for 5' end processing of Pre-60S ribosomal RNA', *Mol Cell*, 36(5), pp. 768-81.

Orsolic, I., Jurada, D., Pullen, N., Oren, M., Eliopoulos, A.G. and Volarevic, S. (2016) 'The relationship between the nucleolus and cancer: Current evidence and emerging paradigms', *Semin Cancer Biol*, 37-38, pp. 36-50.

Osheim, Y.N., French, S.L., Keck, K.M., Champion, E.A., Spasov, K., Dragon, F., Baserga, S.J. and Beyer, A.L. (2004) 'Pre-18S ribosomal RNA is structurally compacted into the SSU processome prior to being cleaved from nascent transcripts in Saccharomyces cerevisiae', *Mol Cell*, 16(6), pp. 943-54.

Page, A.M., Davis, K., Molineux, C., Kolodner, R.D. and Johnson, A.W. (1998) 'Mutational analysis of exoribonuclease I from Saccharomyces cerevisiae', *Nucleic Acids Res*, 26(16), pp. 3707-16.

Panov, K.I., Friedrich, J.K. and Zomerdijk, J.C. (2001) 'A step subsequent to preinitiation complex assembly at the ribosomal RNA gene promoter is rate limiting for human RNA polymerase I-dependent transcription', *Mol Cell Biol*, 21(8), pp. 2641-9.

Panse, V.G. and Johnson, A.W. (2010) 'Maturation of eukaryotic ribosomes: acquisition of functionality', *Trends Biochem Sci*, 35(5), pp. 260-6.

Peculis, B.A., DeGregorio, S. and McDowell, K. (2001) 'The U8 snoRNA gene family: identification and characterization of distinct, functional U8 genes in Xenopus', *Gene*, 274(1-2), pp. 83-92.

Pelava, A. (2016) *Human Ribosome Biogenesis and the Regulation of the Tumour Suppressor p53*. Newcastle University.

Pelava, A., Schneider, C. and Watkins, N.J. (2016) 'The importance of ribosome production, and the 5S RNP-MDM2 pathway, in health and disease', *Biochem Soc Trans*, 44(4), pp. 1086-90.

Perez-Fernandez, J., Martin-Marcos, P. and Dosil, M. (2011) 'Elucidation of the assembly events required for the recruitment of Utp20, Imp4 and Bms1 onto nascent pre-ribosomes', *Nucleic Acids Res*, 39(18), pp. 8105-21.

Perez-Fernandez, J., Roman, A., De Las Rivas, J., Bustelo, X.R. and Dosil, M. (2007) 'The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism', *Mol Cell Biol*, 27(15), pp. 5414-29.

Pertschy, B., Schneider, C., Gnadig, M., Schafer, T., Tollervey, D. and Hurt, E. (2009) 'RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1', *J Biol Chem*, 284(50), pp. 35079-91.

Pestov, D.G., Stockelman, M.G., Strezoska, Z. and Lau, L.F. (2001) 'ERB1, the yeast homolog of mammalian Bop1, is an essential gene required for maturation of the 25S and 5.8S ribosomal RNAs', *Nucleic Acids Res*, 29(17), pp. 3621-30.

Petfalski, E., Dandekar, T., Henry, Y. and Tollervey, D. (1998) 'Processing of the precursors to small nucleolar RNAs and rRNAs requires common components', *Mol Cell Biol*, 18(3), pp. 1181-9.

Phipps, K.R., Charette, J. and Baserga, S.J. (2011) 'The small subunit processome in ribosome biogenesis-progress and prospects', *Wiley Interdiscip Rev RNA*, 2(1), pp. 1-21.

Pillon, M.C., Sobhany, M., Borgnia, M.J., Williams, J.G. and Stanley, R.E. (2017) 'Grc3 programs the essential endoribonuclease Las1 for specific RNA cleavage', *Proc Natl Acad Sci U S A*, 114(28), pp. E5530-E5538.

Pirouz, M., Du, P., Munafo, M. and Gregory, R.I. (2016) 'Dis3l2-Mediated Decay Is a Quality Control Pathway for Noncoding RNAs', *Cell Rep*, 16(7), pp. 1861-73.

Planta, R.J. and Mager, W.H. (1998) 'The list of cytoplasmic ribosomal proteins of Saccharomyces cerevisiae', *Yeast*, 14(5), pp. 471-7.

Politz, J.C., Lewandowski, L.B. and Pederson, T. (2002) 'Signal recognition particle RNA localization within the nucleolus differs from the classical sites of ribosome synthesis', *J Cell Biol*, 159(3), pp. 411-8.

Politz, J.C., Yarovoi, S., Kilroy, S.M., Gowda, K., Zwieb, C. and Pederson, T. (2000) 'Signal recognition particle components in the nucleolus', *Proc Natl Acad Sci U S A*, 97(1), pp. 55-60.

Poole, T.L. and Stevens, A. (1997) 'Structural modifications of RNA influence the 5' exoribonucleolytic hydrolysis by XRN1 and HKE1 of Saccharomyces cerevisiae', *Biochem Biophys Res Commun*, 235(3), pp. 799-805.

Preti, M., O'Donohue, M.F., Montel-Lehry, N., Bortolin-Cavaille, M.L., Choesmel, V. and Gleizes, P.E. (2013) 'Gradual processing of the ITS1 from the nucleolus to the cytoplasm during synthesis of the human 18S rRNA', *Nucleic Acids Res*, 41(8), pp. 4709-23.

Prieto, J.L. and McStay, B. (2005) 'Nucleolar biogenesis: the first small steps', *Biochem Soc Trans*, 33(Pt 6), pp. 1441-3.

Prieto, J.L. and 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 Dev*, 21(16), pp. 2041-54.

Qu, G., van Nues, R.W., Watkins, N.J. and Maxwell, E.S. (2011) 'The spatial-functional coupling of box C/D and C'/D' RNPs is an evolutionarily conserved feature of the eukaryotic box C/D snoRNP nucleotide modification complex', *Mol Cell Biol*, 31(2), pp. 365-74.

Rabl, J., Leibundgut, M., Ataide, S.F., Haag, A. and Ban, N. (2011) 'Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1', *Science*, 331(6018), pp. 730-6.

Raska, I., Koberna, K., Malinsky, J., Fidlerova, H. and Masata, M. (2004) 'The nucleolus and transcription of ribosomal genes', *Biol Cell*, 96(8), pp. 579-94.

Reichow, S.L., Hamma, T., Ferre-D'Amare, A.R. and Varani, G. (2007) 'The structure and function of small nucleolar ribonucleoproteins', *Nucleic Acids Res*, 35(5), pp. 1452-64.

Richter, H., Katic, I., Gut, H. and Grosshans, H. (2016) 'Structural basis and function of XRN2 binding by XTB domains', *Nat Struct Mol Biol*, 23(2), pp. 164-71.
Rissland, O.S. and Norbury, C.J. (2008) 'The Cid1 poly(U) polymerase', *Biochim Biophys Acta*, 1779(4), pp. 286-94.
Rodriguez-Galan, O., Garcia-Gomez, J.J. and de la Cruz, J. (2013) 'Yeast and human RNA helicases involved in ribosome biogenesis: current status and perspectives', *Biochim Biophys Acta*, 1829(8), pp. 775-90.
Rondon, A.G., Mischo, H.E., Kawauchi, J. and Proudfoot, N.J. (2009) 'Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae', *Mol Cell*, 36(1), pp. 88-98.
Rosonina, E., Kaneko, S. and Manley, J.L. (2006) 'Terminating the transcript: breaking up is hard to do', *Genes Dev*, 20(9), pp. 1050-6.
Rouquette, J., Choesmel, V. and Gleizes, P.E. (2005) 'Nuclear export and cytoplasmic processing of precursors to the 40S ribosomal subunits in mammalian cells', *EMBO J*, 24(16), pp. 2862-72.
Ruggero, D. and Pandolfi, P.P. (2003) 'Does the ribosome translate cancer?', *Nat Rev Cancer*, 3(3), pp. 179-92.
Russell, J. and Zomerdijk, J.C. (2005) 'RNA-polymerase-I-directed rDNA transcription, life and works', *Trends Biochem Sci*, 30(2), pp. 87-96.
Saez-Vasquez, J., Caparros-Ruiz, D., Barneche, F. and Echeverria, M. (2004) 'A plant snoRNP complex containing snoRNAs, fibrillarin, and nucleolin-like proteins is competent for both rRNA gene binding and pre-rRNA processing in vitro', *Mol Cell Biol*, 24(16), pp. 7284-97.
Sahasranaman, A., Dembowski, J., Strahler, J., Andrews, P., Maddock, J. and Woolford, J.L., Jr. (2011) 'Assembly of Saccharomyces cerevisiae 60S ribosomal subunits: role of factors required for 27S pre-rRNA processing', *EMBO J*, 30(19), pp. 4020-32.
Samarsky, D.A. and Fournier, M.J. (1998) 'Functional mapping of the U3 small nucleolar RNA from the yeast Saccharomyces cerevisiae', *Mol Cell Biol*, 18(6), pp. 3431-44.
Sambrook, J. and Russell, D.W. (2006) 'The inoue method for preparation and transformation of competent e. Coli: "ultra-competent" cells', *CSH Protoc*, 2006(1).
Sardana, R., Liu, X., Granneman, S., Zhu, J., Gill, M., Papoulas, O., Marcotte, E.M., Tollervey, D., Correll, C.C. and Johnson, A.W. (2015) 'The DEAH-box helicase Dhr1 dissociates U3 from the pre-rRNA to promote formation of the central pseudoknot', *PLoS Biol*, 13(2), p. e1002083.
Sasaki, T., Toh, E.A. and Kikuchi, Y. (2000) 'Yeast Krr1p physically and functionally interacts with a novel essential Kri1p, and both proteins are required for 40S ribosome biogenesis in the nucleolus', *Mol Cell Biol*, 20(21), pp. 7971-9.
Savino, R. and Gerbi, S.A. (1990) 'In vivo disruption of Xenopus U3 snRNA affects ribosomal RNA processing', *EMBO J*, 9(7), pp. 2299-308.
Schaeffer, D., Tsanova, B., Barbas, A., Reis, F.P., Dastidar, E.G., Sanchez-Rotunno, M., Arraiano, C.M. and van Hoof, A. (2009) 'The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities', *Nat Struct Mol Biol*, 16(1), pp. 56-62.
Schilders, G., van Dijk, E. and Pruijn, G.J. (2007) 'C1D and hMtr4p associate with the human exosome subunit PM/Scl-100 and are involved in pre-rRNA processing', *Nucleic Acids Res*, 35(8), pp. 2564-72.
Schillewaert, S., Wacheul, L., Lhomme, F. and Lafontaine, D.L. (2012) 'The evolutionarily conserved protein Las1 is required for pre-rRNA processing at both ends of ITS2', *Mol Cell Biol*, 32(2), pp. 430-44.

Schilling, V., Peifer, C., Buchhaupt, M., Lamberth, S., Lioutikov, A., Rietschel, B., Kotter, P. and Entian, K.D. (2012) 'Genetic interactions of yeast NEP1 (EMG1), encoding an essential factor in ribosome biogenesis', *Yeast*, 29(5), pp. 167-83.

Schmitt, M.E. and Clayton, D.A. (1993) 'Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in Saccharomyces cerevisiae', *Mol Cell Biol*, 13(12), pp. 7935-41.

Schneider, C., Anderson, J.T. and Tollervey, D. (2007) 'The exosome subunit Rrp44 plays a direct role in RNA substrate recognition', *Mol Cell*, 27(2), pp. 324-31.

Schneider, C., Kudla, G., Wlotzka, W., Tuck, A. and Tollervey, D. (2012) 'Transcriptome-wide analysis of exosome targets', *Mol Cell*, 48(3), pp. 422-33.

Schneider, C., Leung, E., Brown, J. and 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 Res*, 37(4), pp. 1127-40.

Schuch, B., Feigenbutz, M., Makino, D.L., Falk, S., Basquin, C., Mitchell, P. and Conti, E. (2014) 'The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase', *EMBO J*, 33(23), pp. 2829-46.

Schutz, P., Karlberg, T., van den Berg, S., Collins, R., Lehtio, L., Hogbom, M., Holmberg-Schiavone, L., Tempel, W., Park, H.W., Hammarstrom, M., Moche, M., Thorsell, A.G. and Schuler, H. (2010) 'Comparative structural analysis of human DEAD-box RNA helicases', *PLoS One*, 5(9).

Shah, B.N., Liu, X. and Correll, C.C. (2013) 'Imp3 unfolds stem structures in pre-rRNA and U3 snoRNA to form a duplex essential for small subunit processing', *RNA*, 19(10), pp. 1372-83.

Sharma, K. and Tollervey, D. (1999) 'Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing', *Mol Cell Biol*, 19(9), pp. 6012-9.

Sharma, K., Venema, J. and Tollervey, D. (1999) 'The 5' end of the 18S rRNA can be positioned from within the mature rRNA', *RNA*, 5(5), pp. 678-86.

Sharma, S. and Lafontaine, D.L. (2015) ''View From A Bridge': A New Perspective on Eukaryotic rRNA Base Modification', *Trends Biochem Sci*, 40(10), pp. 560-75.

Sharpe, H.J., Pau, G., Dijkgraaf, G.J., Basset-Seguin, N., Modrusan, Z., Januario, T., Tsui, V., Durham, A.B., Dlugosz, A.A., Haverty, P.M., Bourgon, R., Tang, J.Y., Sarin, K.Y., Dirix, L., Fisher, D.C., Rudin, C.M., Sofen, H., Migden, M.R., Yauch, R.L. and de Sauvage, F.J. (2015) 'Genomic analysis of smoothened inhibitor resistance in basal cell carcinoma', *Cancer Cell*, 27(3), pp. 327-41.

Singh, S.K., Gurha, P. and Gupta, R. (2008) 'Dynamic guide-target interactions contribute to sequential 2'-O-methylation by a unique archaeal dual guide box C/D sRNP', *RNA*, 14(7), pp. 1411-23.

Skruzny, M., Schneider, C., Racz, A., Weng, J., Tollervey, D. and Hurt, E. (2009) 'An endoribonuclease functionally linked to perinuclear mRNP quality control associates with the nuclear pore complexes', *PLoS Biol*, 7(1), p. e8.

Sloan, K.E., Bohnsack, M.T., Schneider, C. and Watkins, N.J. (2014) 'The roles of SSU processome components and surveillance factors in the initial processing of human ribosomal RNA', *RNA*, 20(4), pp. 540-50.

Sloan, K.E., Bohnsack, M.T. and Watkins, N.J. (2013a) 'The 5S RNP couples p53 homeostasis to ribosome biogenesis and nucleolar stress', *Cell Rep*, 5(1), pp. 237-47.

Sloan, K.E., Leisegang, M.S., Doebele, C., Ramirez, A.S., Simm, S., Safferthal, C., Kretschmer, J., Schorge, T., Markoutsa, S., Haag, S., Karas, M., Ebersberger, I., Schleiff, E., Watkins, N.J. and

Bohnsack, M.T. (2015) 'The association of late-acting snoRNPs with human pre-ribosomal complexes requires the RNA helicase DDX21', *Nucleic Acids Res*, 43(1), pp. 553-64.
Sloan, K.E., Mattijssen, S., Lebaron, S., Tollervey, D., Pruijn, G.J. and Watkins, N.J. (2013b) 'Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing', *J Cell Biol*, 200(5), pp. 577-88.
Sloan, K.E., Warda, A.S., Sharma, S., Entian, K.D., Lafontaine, D.L.J. and Bohnsack, M.T. (2017) 'Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function', *RNA Biol*, 14(9), pp. 1138-1152.
Smirnov, E., Cmarko, D., Mazel, T., Hornacek, M. and Raska, I. (2016) 'Nucleolar DNA: the host and the guests', *Histochem Cell Biol*, 145(4), pp. 359-72.
Solinger, J.A., Pascolini, D. and Heyer, W.D. (1999) 'Active-site mutations in the Xrn1p exoribonuclease of Saccharomyces cerevisiae reveal a specific role in meiosis', *Mol Cell Biol*, 19(9), pp. 5930-42.
Sondalle, S.B. and Baserga, S.J. (2014) 'Human diseases of the SSU processome', *Biochim Biophys Acta*, 1842(6), pp. 758-64.
Sorensen, P.D. and Frederiksen, S. (1991) 'Characterization of human 5S rRNA genes', *Nucleic Acids Res*, 19(15), pp. 4147-51.
Staals, R.H., Bronkhorst, A.W., Schilders, G., Slomovic, S., Schuster, G., Heck, A.J., Raijmakers, R. and Pruijn, G.J. (2010) 'Dis3-like 1: a novel exoribonuclease associated with the human exosome', *EMBO J*, 29(14), pp. 2358-67.
Stead, J.A., Costello, J.L., Livingstone, M.J. and 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 Res*, 35(16), pp. 5556-67.
Steffensen, D.M., Duffey, P. and Prensky, W. (1974) 'Localisation of 5S ribosomal RNA genes on human chromosome 1', *Nature*, 252(5485), pp. 741-3.
Steitz, T.A. and Steitz, J.A. (1993) 'A general two-metal-ion mechanism for catalytic RNA', *Proc Natl Acad Sci U S A*, 90(14), pp. 6498-502.
Stevens, A., Hsu, C.L., Isham, K.R. and Larimer, F.W. (1991) 'Fragments of the internal transcribed spacer 1 of pre-rRNA accumulate in Saccharomyces cerevisiae lacking 5'----3' exoribonuclease 1', *J Bacteriol*, 173(21), pp. 7024-8.
Stevens, A. and Poole, T.L. (1995) '5'-exonuclease-2 of Saccharomyces cerevisiae. Purification and features of ribonuclease activity with comparison to 5'-exonuclease-1', *J Biol Chem*, 270(27), pp. 16063-9.
Strub, B.R., Eswara, M.B., Pierce, J.B. and Mangroo, D. (2007) 'Utp8p is a nucleolar tRNA-binding protein that forms a complex with components of the nuclear tRNA export machinery in Saccharomyces cerevisiae', *Mol Biol Cell*, 18(10), pp. 3845-59.
Strunk, B.S. and Karbstein, K. (2009) 'Powering through ribosome assembly', *RNA*, 15(12), pp. 2083-104.
Strunk, B.S., Loucks, C.R., Su, M., Vashisth, H., Cheng, S., Schilling, J., Brooks, C.L., 3rd, Karbstein, K. and Skiniotis, G. (2011) 'Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates', *Science*, 333(6048), pp. 1449-53.
Strunk, B.S., Novak, M.N., Young, C.L. and Karbstein, K. (2012) 'A translation-like cycle is a quality control checkpoint for maturing 40S ribosome subunits', *Cell*, 150(1), pp. 111-21.

Sun, Q., Zhu, X., Qi, J., An, W., Lan, P., Tan, D., Chen, R., Wang, B., Zheng, S., Zhang, C., Chen, X., Zhang, W., Chen, J., Dong, M.Q. and Ye, K. (2017) 'Molecular architecture of the 90S small subunit pre-ribosome', *Elife*, 6.

Sun, X.X., Wang, Y.G., Xirodimas, D.P. and Dai, M.S. (2010) 'Perturbation of 60 S ribosomal biogenesis results in ribosomal protein L5- and L11-dependent p53 activation', *J Biol Chem*, 285(33), pp. 25812-21.

Sweet, T., Yen, W., Khalili, K. and Amini, S. (2008) 'Evidence for involvement of NFBP in processing of ribosomal RNA', *J Cell Physiol*, 214(2), pp. 381-8.

Szewczak, L.B., Gabrielsen, J.S., Degregorio, S.J., Strobel, S.A. and Steitz, J.A. (2005) 'Molecular basis for RNA kink-turn recognition by the h15.5K small RNP protein', *RNA*, 11(9), pp. 1407-19.

Tafforeau, L., Zorbas, C., Langhendries, J.L., Mullineux, S.T., Stamatopoulou, V., Mullier, R., Wacheul, L. and Lafontaine, D.L. (2013) 'The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors', *Mol Cell*, 51(4), pp. 539-51.

Talkish, J., Campbell, I.W., Sahasranaman, A., Jakovljevic, J. and Woolford, J.L., Jr. (2014) 'Ribosome assembly factors Pwp1 and Nop12 are important for folding of 5.8S rRNA during ribosome biogenesis in Saccharomyces cerevisiae', *Mol Cell Biol*, 34(10), pp. 1863-77.

Tanaka, N., Smith, P. and Shuman, S. (2011) 'Crystal structure of Rcl1, an essential component of the eukaryal pre-rRNA processosome implicated in 18s rRNA biogenesis', *RNA*, 17(4), pp. 595-602.

Tarassov, K., Messier, V., Landry, C.R., Radinovic, S., Serna Molina, M.M., Shames, I., Malitskaya, Y., Vogel, J., Bussey, H. and Michnick, S.W. (2008) 'An in vivo map of the yeast protein interactome', *Science*, 320(5882), pp. 1465-70.

Terns, M.P. and Terns, R.M. (2002) 'Small nucleolar RNAs: versatile trans-acting molecules of ancient evolutionary origin', *Gene Expr*, 10(1-2), pp. 17-39.

The Cancer Genome Atlas, N. (2012) 'Comprehensive molecular portraits of human breast tumours', *Nature*, 490, p. 61.

Thiry, M. and Lafontaine, D.L. (2005) 'Birth of a nucleolus: the evolution of nucleolar compartments', *Trends Cell Biol*, 15(4), pp. 194-9.

Thomson, E., Ferreira-Cerca, S. and Hurt, E. (2013) 'Eukaryotic ribosome biogenesis at a glance', *J Cell Sci*, 126(Pt 21), pp. 4815-21.

Thomson, E. and Tollervey, D. (2010) 'The final step in 5.8S rRNA processing is cytoplasmic in Saccharomyces cerevisiae', *Mol Cell Biol*, 30(4), pp. 976-84.

Timofeeva, M.N., Kinnersley, B., Farrington, S.M., Whiffin, N., Palles, C., Svinti, V., Lloyd, A., Gorman, M., Ooi, L.Y., Hosking, F., Barclay, E., Zgaga, L., Dobbins, S., Martin, L., Theodoratou, E., Broderick, P., Tenesa, A., Smillie, C., Grimes, G., Hayward, C., Campbell, A., Porteous, D., Deary, I.J., Harris, S.E., Northwood, E.L., Barrett, J.H., Smith, G., Wolf, R., Forman, D., Morreau, H., Ruano, D., Tops, C., Wijnen, J., Schrumpf, M., Boot, A., Vasen, H.F., Hes, F.J., van Wezel, T., Franke, A., Lieb, W., Schafmayer, C., Hampe, J., Buch, S., Propping, P., Hemminki, K., Forsti, A., Westers, H., Hofstra, R., Pinheiro, M., Pinto, C., Teixeira, M., Ruiz-Ponte, C., Fernandez-Rozadilla, C., Carracedo, A., Castells, A., Castellvi-Bel, S., Campbell, H., Bishop, D.T., Tomlinson, I.P., Dunlop, M.G. and Houlston, R.S. (2015) 'Recurrent Coding Sequence Variation Explains Only A Small Fraction of the Genetic Architecture of Colorectal Cancer', *Sci Rep*, 5, p. 16286.

Tollervey, D. (1987) 'A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA', *EMBO J*, 6(13), pp. 4169-75.

Tomecki, R. and Dziembowski, A. (2010) 'Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism', *RNA*, 16(9), pp. 1692-724.

Tomecki, R., Kristiansen, M.S., Lykke-Andersen, S., Chlebowski, A., Larsen, K.M., Szczesny, R.J., Drazkowska, K., Pastula, A., Andersen, J.S., Stepien, P.P., Dziembowski, A. and Jensen, T.H. (2010) 'The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L', *EMBO J*, 29(14), pp. 2342-57.

Tomecki, R., Labno, A., Drazkowska, K., Cysewski, D. and Dziembowski, A. (2015) 'hUTP24 is essential for processing of the human rRNA precursor at site A1, but not at site A0', *RNA Biol*, 12(9), pp. 1010-29.

Tomecki, R., Sikorski, P.J. and Zakrzewska-Placzek, M. (2017) 'Comparison of preribosomal RNA processing pathways in yeast, plant and human cells - focus on coordinated action of endo- and exoribonucleases', *FEBS Lett*, 591(13), pp. 1801-1850.

Torchet, C. and 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(1), pp. 70-80.

Torchet, C., Jacq, C. and Hermann-Le Denmat, S. (1998) 'Two mutant forms of the S1/TPR-containing protein Rrp5p affect the 18S rRNA synthesis in Saccharomyces cerevisiae', *RNA*, 4(12), pp. 1636-52.

Tschochner, H. and Hurt, E. (2003) 'Pre-ribosomes on the road from the nucleolus to the cytoplasm', *Trends Cell Biol*, 13(5), pp. 255-63.

Turner, A.J., Knox, A.A., Prieto, J.L., McStay, B. and Watkins, N.J. (2009) 'A novel small-subunit processome assembly intermediate that contains the U3 snoRNP, nucleolin, RRP5, and DBP4', *Mol Cell Biol*, 29(11), pp. 3007-17.

Turner, A.J., Knox, A.A. and Watkins, N.J. (2012) 'Nucleolar disruption leads to the spatial separation of key 18S rRNA processing factors', *RNA Biol*, 9(2), pp. 175-86.

Turowski, T.W. and Tollervey, D. (2015) 'Cotranscriptional events in eukaryotic ribosome synthesis', *Wiley Interdiscip Rev RNA*, 6(1), pp. 129-39.

Tycowski, K.T., Smith, C.M., Shu, M.D. and Steitz, J.A. (1996) 'A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in Xenopus', *Proc Natl Acad Sci U S A*, 93(25), pp. 14480-5.

Tym, J.E., Mitsopoulos, C., Coker, E.A., Razaz, P., Schierz, A.C., Antolin, A.A. and Al-Lazikani, B. (2016) 'canSAR: an updated cancer research and drug discovery knowledgebase', *Nucleic Acids Res*, 44(D1), pp. D938-43.

Uehata, T. and Akira, S. (2013) 'mRNA degradation by the endoribonuclease Regnase-1/ZC3H12a/MCPIP-1', *Biochim Biophys Acta*, 1829(6-7), pp. 708-13.

Umate, P., Tuteja, R. and Tuteja, N. (2010) 'Genome-wide analysis of helicase gene family from rice and Arabidopsis: a comparison with yeast and human', *Plant Mol Biol*, 73(4-5), pp. 449-65.

Ustianenko, D., Hrossova, D., Potesil, D., Chalupnikova, K., Hrazdilova, K., Pachernik, J., Cetkovska, K., Uldrijan, S., Zdrahal, Z. and Vanacova, S. (2013) 'Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs', *RNA*, 19(12), pp. 1632-8.

Ustianenko, D., Pasulka, J., Feketova, Z., Bednarik, L., Zigackova, D., Fortova, A., Zavolan, M. and Vanacova, S. (2016) 'TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs', *EMBO J*, 35(20), pp. 2179-2191.

van Dijk, E.L., Schilders, G. and Pruijn, G.J. (2007) 'Human cell growth requires a functional cytoplasmic exosome, which is involved in various mRNA decay pathways', *RNA*, 13(7), pp. 1027-35.
van Hoof, A., Lennertz, P. and 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', *EMBO J*, 19(6), pp. 1357-65.
van Nues, R.W., Granneman, S., Kudla, G., Sloan, K.E., Chicken, M., Tollervey, D. and Watkins, N.J. (2011) 'Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing', *EMBO J*, 30(12), pp. 2420-30.
van Nues, R.W. and Watkins, N.J. (2017) 'Unusual C/D motifs enable box C/D snoRNPs to modify multiple sites in the same rRNA target region', *Nucleic Acids Res*, 45(4), pp. 2016-2028.
Vanacova, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G. and Keller, W. (2005) 'A new yeast poly(A) polymerase complex involved in RNA quality control', *PLoS Biol*, 3(6), p. e189.
Vance, V.B., Thompson, E.A. and Bowman, L.H. (1985) 'Transfection of mouse ribosomal DNA into rat cells: faithful transcription and processing', *Nucleic Acids Res*, 13(20), pp. 7499-513.
Vasiljeva, L. and Buratowski, S. (2006) 'Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts', *Mol Cell*, 21(2), pp. 239-48.
Veith, T., Martin, R., Wurm, J.P., Weis, B.L., Duchardt-Ferner, E., Safferthal, C., Hennig, R., Mirus, O., Bohnsack, M.T., Wohnert, J. and Schleiff, E. (2012) 'Structural and functional analysis of the archaeal endonuclease Nob1', *Nucleic Acids Res*, 40(7), pp. 3259-74.
Venema, J., Bousquet-Antonelli, C., Gelugne, J.P., Caizergues-Ferrer, M. and Tollervey, D. (1997) 'Rok1p is a putative RNA helicase required for rRNA processing', *Mol Cell Biol*, 17(6), pp. 3398-407.
Venema, J. and Tollervey, D. (1995) 'Processing of pre-ribosomal RNA in Saccharomyces cerevisiae', *Yeast*, 11(16), pp. 1629-50.
Venema, J. and Tollervey, D. (1996) 'RRP5 is required for formation of both 18S and 5.8S rRNA in yeast', *EMBO J*, 15(20), pp. 5701-14.
Venema, J. and Tollervey, D. (1999) 'Ribosome synthesis in Saccharomyces cerevisiae', *Annu Rev Genet*, 33, pp. 261-311.
Venema, J., Vos, H.R., Faber, A.W., van Venrooij, W.J. and Raue, H.A. (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*, 6(11), pp. 1660-71.
Vincent, H.A. and Deutscher, M.P. (2006) 'Substrate recognition and catalysis by the exoribonuclease RNase R', *J Biol Chem*, 281(40), pp. 29769-75.
Vos, H.R., Bax, R., Faber, A.W., Vos, J.C. and Raue, H.A. (2004a) 'U3 snoRNP and Rrp5p associate independently with Saccharomyces cerevisiae 35S pre-rRNA, but Rrp5p is essential for association of Rok1p', *Nucleic Acids Res*, 32(19), pp. 5827-33.
Vos, H.R., Faber, A.W., de Gier, M.D., Vos, J.C. and Raue, H.A. (2004b) 'Deletion of the three distal S1 motifs of Saccharomyces cerevisiae Rrp5p abolishes pre-rRNA processing at site A(2) without reducing the production of functional 40S subunits', *Eukaryot Cell*, 3(6), pp. 1504-12.
Wade, M., Li, Y.C. and Wahl, G.M. (2013) 'MDM2, MDMX and p53 in oncogenesis and cancer therapy', *Nat Rev Cancer*, 13(2), pp. 83-96.

Wang, H.W., Wang, J., Ding, F., Callahan, K., Bratkowski, M.A., Butler, J.S., Nogales, E. and Ke, A. (2007) 'Architecture of the yeast Rrp44 exosome complex suggests routes of RNA recruitment for 3' end processing', *Proc Natl Acad Sci U S A*, 104(43), pp. 16844-9.

Wang, M., Anikin, L. and Pestov, D.G. (2014) 'Two orthogonal cleavages separate subunit RNAs in mouse ribosome biogenesis', *Nucleic Acids Res*, 42(17), pp. 11180-91.

Wang, M., Parshin, A.V., Shcherbik, N. and Pestov, D.G. (2015) 'Reduced expression of the mouse ribosomal protein Rpl17 alters the diversity of mature ribosomes by enhancing production of shortened 5.8S rRNA', *RNA*, 21(7), pp. 1240-8.

Wang, M. and Pestov, D.G. (2011) '5'-end surveillance by Xrn2 acts as a shared mechanism for mammalian pre-rRNA maturation and decay', *Nucleic Acids Res*, 39(5), pp. 1811-22.

Wang, X., Jia, H., Jankowsky, E. and Anderson, J.T. (2008) 'Degradation of hypomodified tRNA(iMet) in vivo involves RNA-dependent ATPase activity of the DExH helicase Mtr4p', *RNA*, 14(1), pp. 107-16.

Wang, Y., Zhu, Q., Huang, L., Zhu, Y., Chen, J., Peng, J. and Lo, L.J. (2016) 'Interaction between Bms1 and Rcl1, two ribosome biogenesis factors, is evolutionally conserved in zebrafish and human', *J Genet Genomics*, 43(7), pp. 467-9.

Warner, J.R. (1999) 'The economics of ribosome biosynthesis in yeast', *Trends Biochem Sci*, 24(11), pp. 437-40.

Wasmuth, E.V., Januszyk, K. and Lima, C.D. (2014) 'Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA', *Nature*, 511(7510), pp. 435-9.

Wasmuth, E.V. and Lima, C.D. (2017) 'The Rrp6 C-terminal domain binds RNA and activates the nuclear RNA exosome', *Nucleic Acids Res*, 45(2), pp. 846-860.

Watkins, N.J. and Bohnsack, M.T. (2012) 'The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA', *Wiley Interdiscip Rev RNA*, 3(3), pp. 397-414.

Watkins, N.J., Dickmanns, A. and 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', *Mol Cell Biol*, 22(23), pp. 8342-52.

Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M. and Luhrmann, R. (1998) 'Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure', *RNA*, 4(12), pp. 1549-68.

Watkins, N.J., Segault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C. and Luhrmann, R. (2000) 'A common core RNP structure shared between the small nucleoar box C/D RNPs and the spliceosomal U4 snRNP', *Cell*, 103(3), pp. 457-66.

Wegierski, T., Billy, E., Nasr, F. and Filipowicz, W. (2001) 'Bms1p, a G-domain-containing protein, associates with Rcl1p and is required for 18S rRNA biogenesis in yeast', *RNA*, 7(9), pp. 1254-67.

Wehner, K.A., Gallagher, J.E. and Baserga, S.J. (2002) 'Components of an interdependent unit within the SSU processome regulate and mediate its activity', *Mol Cell Biol*, 22(20), pp. 7258-67.

Weiner, A.M., Scampoli, N.L. and Calcaterra, N.B. (2012) 'Fishing the molecular bases of Treacher Collins syndrome', *PLoS One*, 7(1), p. e29574.

Wells, G.R., Weichmann, F., Colvin, D., Sloan, K.E., Kudla, G., Tollervey, D., Watkins, N.J. and Schneider, C. (2016) 'The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans', *Nucleic Acids Res*, 44(11), pp. 5399-409.

Wells, G.R., Weichmann, F., Sloan, K.E., Colvin, D., Watkins, N.J. and Schneider, C. (2017) 'The ribosome biogenesis factor yUtp23/hUTP23 coordinates key interactions in the yeast and human pre-40S particle and hUTP23 contains an essential PIN domain', *Nucleic Acids Res*, 45(8), pp. 4796-4809.

Wery, M., Ruidant, S., Schillewaert, S., Lepore, N. and Lafontaine, D.L. (2009) 'The nuclear poly(A) polymerase and Exosome cofactor Trf5 is recruited cotranscriptionally to nucleolar surveillance', *RNA*, 15(3), pp. 406-19.

West, S., Gromak, N. and Proudfoot, N.J. (2004) 'Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites', *Nature*, 432(7016), pp. 522-5.

Wilkie, N.M. and Smellie, R.M. (1968) 'Polyribonucleotide synthesis by subfractions of microsomes from rat liver', *Biochem J*, 109(2), pp. 229-38.

Wilusz, C.J. and Wilusz, J. (2008) 'New ways to meet your (3') end oligouridylation as a step on the path to destruction', *Genes Dev*, 22(1), pp. 1-7.

Woolford, J.L., Jr. and Baserga, S.J. (2013) 'Ribosome biogenesis in the yeast Saccharomyces cerevisiae', *Genetics*, 195(3), pp. 643-81.

Wormsley, S., Samarsky, D.A., Fournier, M.J. and Baserga, S.J. (2001) 'An unexpected, conserved element of the U3 snoRNA is required for Mpp10p association', *RNA*, 7(6), pp. 904-19.

Worton, R.G., Sutherland, J., Sylvester, J.E., Willard, H.F., Bodrug, S., Dube, I., Duff, C., Kean, V., Ray, P.N. and Schmickel, R.D. (1988) 'Human ribosomal RNA genes: orientation of the tandem array and conservation of the 5' end', *Science*, 239(4835), pp. 64-8.

Wu, K., Wu, P. and Aris, J.P. (2001) 'Nucleolar protein Nop12p participates in synthesis of 25S rRNA in Saccharomyces cerevisiae', *Nucleic Acids Res*, 29(14), pp. 2938-49.

Wu, L. and Levine, A.J. (1997) 'Differential regulation of the p21/WAF-1 and mdm2 genes after high-dose UV irradiation: p53-dependent and p53-independent regulation of the mdm2 gene', *Mol Med*, 3(7), pp. 441-51.

Wu, S., Tutuncuoglu, B., Yan, K., Brown, H., Zhang, Y., Tan, D., Gamalinda, M., Yuan, Y., Li, Z., Jakovljevic, J., Ma, C., Lei, J., Dong, M.Q., Woolford, J.L., Jr. and Gao, N. (2016) 'Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes', *Nature*, 534(7605), pp. 133-7.

Wurm, J.P., Meyer, B., Bahr, U., Held, M., Frolow, O., Kotter, P., Engels, J.W., Heckel, A., Karas, M., Entian, K.D. and Wohnert, J. (2010) 'The ribosome assembly factor Nep1 responsible for Bowen-Conradi syndrome is a pseudouridine-N1-specific methyltransferase', *Nucleic Acids Res*, 38(7), pp. 2387-98.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.C., Dufour, M.E., Boulay, J., Regnault, B., Devaux, F., Namane, A., Seraphin, B., Libri, D. and Jacquier, A. (2005) 'Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase', *Cell*, 121(5), pp. 725-37.

Xiang, S., Cooper-Morgan, A., Jiao, X., Kiledjian, M., Manley, J.L. and Tong, L. (2009) 'Structure and function of the 5'-->3' exoribonuclease Rat1 and its activating partner Rai1', *Nature*, 458(7239), pp. 784-8.

Xue, Y., Bai, X., Lee, I., Kallstrom, G., Ho, J., Brown, J., Stevens, A. and Johnson, A.W. (2000) 'Saccharomyces cerevisiae RAI1 (YGL246c) is homologous to human DOM3Z and encodes a protein that binds the nuclear exoribonuclease Rat1p', *Mol Cell Biol*, 20(11), pp. 4006-15.

Yang, Q. and Jankowsky, E. (2005) 'ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1', *Biochemistry*, 44(41), pp. 13591-601.

Yelick, P.C. and Trainor, P.A. (2015) 'Ribosomopathies: Global process, tissue specific defects', *Rare Dis*, 3(1), p. e1025185.

Young, C.L. and Karbstein, K. (2011) 'The roles of S1 RNA-binding domains in Rrp5's interactions with pre-rRNA', *RNA*, 17(3), pp. 512-21.

Young, C.L., Khoshnevis, S. and Karbstein, K. (2013) 'Cofactor-dependent specificity of a DEAD-box protein', *Proc Natl Acad Sci U S A*, 110(29), pp. E2668-76.

Yu, B., Mitchell, G.A. and Richter, A. (2005) 'Nucleolar localization of cirhin, the protein mutated in North American Indian childhood cirrhosis', *Exp Cell Res*, 311(2), pp. 218-28.

Zhang, C., Sun, Q., Chen, R., Chen, X., Lin, J. and Ye, K. (2016a) 'Integrative structural analysis of the UTPB complex, an early assembly factor for eukaryotic small ribosomal subunits', *Nucleic Acids Res*, 44(15), pp. 7475-86.

Zhang, J., Harnpicharnchai, P., Jakovljevic, J., Tang, L., Guo, Y., Oeffinger, M., Rout, M.P., Hiley, S.L., Hughes, T. and Woolford, J.L., Jr. (2007) 'Assembly factors Rpf2 and Rrs1 recruit 5S rRNA and ribosomal proteins rpL5 and rpL11 into nascent ribosomes', *Genes Dev*, 21(20), pp. 2580-92.

Zhang, L., Lin, J. and Ye, K. (2013) 'Structural and functional analysis of the U3 snoRNA binding protein Rrp9', *RNA*, 19(5), pp. 701-11.

Zhang, L., Wu, C., Cai, G., Chen, S. and Ye, K. (2016b) 'Stepwise and dynamic assembly of the earliest precursors of small ribosomal subunits in yeast', *Genes Dev*, 30(6), pp. 718-32.

Zheng, S., Lan, P., Liu, X. and Ye, K. (2014) 'Interaction between ribosome assembly factors Krr1 and Faf1 is essential for formation of small ribosomal subunit in yeast', *J Biol Chem*, 289(33), pp. 22692-703.

Zhu, J., Liu, X., Anjos, M., Correll, C.C. and Johnson, A.W. (2016) 'Utp14 Recruits and Activates the RNA Helicase Dhr1 To Undock U3 snoRNA from the Preribosome', *Mol Cell Biol*, 36(6), pp. 965-78.

Publications and Presentations

Publications

Wells, G.R., Weichmann, F., Colvin, D., Sloan, K.E., Kudla, G., Tollervey, D., Watkins, N.J. and Schneider, C. (2016) 'The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans', *Nucleic Acids Res*, 44(11), pp. 5399-409.

Wells, G.R., Weichmann, F., Sloan, K.E., Colvin, D., Watkins, N.J. and Schneider, C. (2017) 'The ribosome biogenesis factor yUtp23/hUTP23 coordinates key interactions in the yeast and human pre-40S particle and hUTP23 contains an essential PIN domain', *Nucleic Acids Res*, 45(8), pp. 4796-4809.

Oral Presentations

Wells, G.R., (2017) ‘Endonucleases in human small subunit ribosome biogenesis’, ICaMB Away Day, Newcastle University, UK, 18th October.

Wells, G.R., Weichmann, F., Colvin, D., Sloan, K.E., Kudla, G., Tollervey, D., Watkins, N.J. and Schneider, C. (2016) ‘8. Finding the missing endonucleases in yeast and human 18S rRNA release’, RNA UK Conference, Lake District, UK, 29th-31st January.

Poster Presentations

Wells, G.R., Weichmann, F., Sloan, K.E., Colvin, D., Watkins, N.J., Schneider, C. (2018) ‘The ribosome biogenesis factor Utp23 coordinates key interactions in the yeast and human pre-40S particle and human UTP23 contains an essential PIN domain’, RNA UK Conference, Lake District, UK, 26th-28th January.

Wells, G.R., (2017) ‘Endonucleases in the early stages of human ribosome biogenesis’, North East Postgraduate Conference, Newcastle, UK, 30th October.

Wells, G.R., Weichmann, F., Colvin, D., Sloan, K.E., Kudla, G., Tollervey, D., Watkins, N.J. and Schneider, C. (2016) 'The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans’, 21st Annual Meeting of the RNA Society, Kyoto, Japan, 28th June-2nd July.

Nucleic Acids Research Advance Access published March 31, 2016

Nucleic Acids Research, 2016 ***1–11***
doi: 10.1093/nar/gkw213

The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans

Graeme R. Wells[1], Franziska Weichmann[1], David Colvin[1], Katherine E. Sloan[1], Grzegorz Kudla[2,3], David Tollervey[2], Nicholas J. Watkins[1] and Claudia Schneider[1,*]

[1]Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne NE2 4HH, UK, [2]Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, UK and [3]MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK

Received January 08, 2016; Revised March 16, 2016; Accepted March 18, 2016

ABSTRACT

During ribosomal RNA (rRNA) maturation, cleavages at defined sites separate the mature rRNAs from spacer regions, but the identities of several enzymes required for 18S rRNA release remain unknown. PilT N-terminus (PIN) domain proteins are frequently endonucleases and the PIN domain protein Utp24 is essential for early cleavages at three pre-rRNA sites in yeast (A0, A1 and A2) and humans (A0, 1 and 2a). In yeast, A1 is cleaved prior to A2 and both cleavages require base-pairing by the U3 snoRNA to the central pseudoknot elements of the 18S rRNA. We found that yeast Utp24 UV-crosslinked *in vivo* to U3 and the pseudoknot, placing Utp24 close to cleavage at site A1. Yeast and human Utp24 proteins exhibited *in vitro* endonuclease activity on an RNA substrate containing yeast site A2. Moreover, an intact PIN domain in human UTP24 was required for accurate cleavages at sites 1 and 2a *in vivo*, whereas mutation of another potential site 2a endonuclease, RCL1, did not affect 18S production. We propose that Utp24 cleaves sites A1/1 and A2/2a in yeast and human cells.

INTRODUCTION

The eukaryotic rRNAs are processed from the 35S (*Saccharomyces cerevisiae*) or 47S (*Homo sapiens*) rRNA precursors (pre-rRNAs) by endonucleolytic cleavages and exonucleolytic trimming, with concomitant removal of external (5′-ETS, 3′-ETS) and internal (ITS1, ITS2) transcribed spacer sequences (Figure 1 and Supplementary Figure S1) (1).

Early pre-rRNA cleavages, at sites called A0, A1 and A2 in yeast and A0, A1/1 and 2a/E in humans, are important for 18S rRNA maturation. These events require a large ribonucleoprotein complex, the small subunit (SSU) processome (or 90S pre-ribosome) (2). A key component of the SSU processome is the U3 small nucleolar (sno)RNA, which base-pairs with the 5′-ETS and 18S rRNA elements to chaperone the formation of the conserved pseudoknot, a key structural feature of the 40S ribosomal subunit (2,3).

Loss of many different SSU processome components blocks pre-rRNA cleavage at sites A0, A1/1 and A2/2a, making it unclear which factor is the active nuclease. However, two evolutionarily conserved factors feature protein domains linked to RNA processing. Yeast Utp24/Fcf1 and human UTP24 harbor PIN (PilT N-terminus) endonuclease domains (4), whereas yeast Rcl1 and human RCL1 exhibit RNA cyclase-like protein folds (5). A PIN domain is also found in the endonuclease Nob1, which catalyzes the final cytoplasmic maturation step at the 3′-end of the 18S rRNA (site D) in yeast (6–8).

Yeast Utp24 and Rcl1 are essential for growth and conditional depletion of either protein inhibits A0, A1 and A2 cleavage (4,5,9). However, mutation of the PIN domain of Utp24 specifically inhibited cleavage at sites A1 and A2, while cleavage at site A0 was unaffected (4), and the mutant was dominant negative when expressed together with intact Utp24. These observations suggested that the presence of Utp24 is required for SSU processome function in A0-A2 cleavage, whereas the PIN domain harbors the catalytic activity for A1 and A2 cleavage (4), but nuclease activity was not demonstrated. Mutation of Rcl1 also inhibited site A2 cleavage, with less effect at A0 and A1 (5,9,10), and recombinant Rcl1 was reported to cleave pre-rRNA transcripts containing yeast site A2, consistent with direct endonuclease activity (9).

*To whom correspondence should be addressed. Tel: +44 191 208 7708; Fax: +44 191 208 7424; Email: Claudia.Schneider@ncl.ac.uk
Present address:
Franziska Weichmann, Lehrstuhl für Biochemie I, Fakultät für Biologie und Vorklinische Medizin, Universität Regensburg, Regensburg 93053, Germany.
Katherine E. Sloan, Universitätsmedizin Göttingen, Institut für Molekularbiologie, Göttingen 37073, Germany.

In mammals, the equivalent pre-rRNA cleavages also require the presence of UTP24 (FCF1 in mouse) and RCL1 (11–14) and the putative catalytic activity of human UTP24 was very recently linked to site 1 cleavage (15). However, the catalytic roles of UTP24 and RCL1 with respect to 2a cleavage have not been established.

In yeast, mutational and kinetic analyses indicate that processing at sites A1 and A2 is tightly coupled and mainly co-transcriptional (16,17). In vertebrates, pre-rRNA processing appears to be largely post-transcriptional and A′ processing near the 5′-end of the 5′-ETS in humans, a cleavage site not present in yeast, appears to be the only co-transcriptional event (18). Cleavage at site 2a within the human ITS1 (equivalent to yeast A2) is part of a minor pathway in humans (Supplementary Figure S1) (11,12,19). However, no precursors processed at site 2a but still containing 5′-ETS sequences are detected, indicating that A0, 1 and 2a cleavages are coupled within the SSU processome, as in yeast.

Here, we present a combination of *in vivo* and *in vitro* approaches to clarify the roles of Utp24 and Rcl1 in yeast and human ribosome biogenesis. *In vivo* RNA–protein crosslinking studies (CRAC) generated a transcriptome-wide RNA binding profile for yeast Utp24, which provides fresh insights into its function. We also performed *in vitro* assays using recombinant wild-type (WT) and mutant Utp24 and Rcl1 proteins to assess their cleavage activity on pre-ribosomal RNA. Finally, we established RNAi-rescue systems in HEK293 cells to study the effect of presumably catalytically inactive UTP24 and RCL1 mutants on pre-rRNA cleavage in the human system.

MATERIALS AND METHODS

Yeast strains and methods

S. cerevisiae strains (Supplementary Table S1) were constructed by standard methods (20). Cultures were grown at 30°C in medium containing 0.67% nitrogen base (Difco) and 2% glucose or 2% galactose. Strains for crosslinking studies expressed genomically encoded, C-terminal HTP-tagged proteins under the control of their endogenous promoter.

CRAC and data analysis

The CRAC method was performed as previously described (21,22), see Supplementary Figure S2A. To generate RNA–protein crosslinks, actively growing yeast cultures in SD medium (OD_{600} ~0.5) were UV-irradiated in a 1.2 m metal tube for 100 s at 254 nm (22). Illumina sequencing data were aligned to the yeast genome using Novoalign (http://www.novocraft.com). Bioinformatics analyses were performed as described (23). The Illumina sequencing data from this publication have been submitted to the GEO database (http://www.ncbi.nlm.nih.gov/geo/) and assigned the identifier GSE75991.

Cloning and mutagenesis

The open reading frames for Rcl1, Bms1, Utp24 or UTP24 were amplified from yeast genomic DNA or human cDNA adding restriction sites (Supplementary Table S2) and cloned into pET100 vectors (Invitrogen). The constructs were used for *in vitro* translation in the presence of [^{35}S] methionine (TNT, Promega) or subcloned into pGEX-6P1 vectors to express and purify Glutathione *S* - transferase (GST) - tagged recombinant proteins from *E. coli* using standard techniques. The C-terminally HTP-tagged Rcl1 gene was amplified by PCR from yeast genomic DNA (Strain Rcl1-HTP) adding restriction sites to the 5′-end of the gene and the 3′-end of the HTP-tag, respectively, and cloned into pRS316. The coding sequences of UTP24 and RCL1 were altered to make them resistant to the siRNAs (Supplementary Table S3) used to deplete the endogenous mRNAs. These constructs (IDT) were amplified by PCR and cloned into the pcDNA5/FRT/TO vector (Invitrogen) containing 2x N-terminal FLAG tags under the control of a tetracycline-inducible promoter. Point mutations were generated by QuikChange site-directed mutagenesis using overlapping primers (Supplementary Table S2) and verified by sequencing.

Cell culture and RNAi

Constructs were transfected into Flp-In T-Rex HEK293 cells. Stably transfected cells were selected as described by the manufacturor (Invitrogen) and cultured according to standard protocols. Expression of exogenous proteins was induced by addition of tetracycline (UTP24; 1 mg/ml, RCL1; 0.01–0.1 mg/ml). Cells were transfected with siRNA duplexes (Supplementary Table S3) using Lipofectamine RNAiMAX transfection reagent (Invitrogen) and harvested after 72 h depletion.

RNA analysis

RNA was extracted from HEK293 cell pellets using TRI reagent (Sigma-Aldrich). For northern blot analysis, 2 μg of total RNA was separated on 1.2% glyoxal-agarose gels, transferred to nylon membrane and hybridized with 5′-end labeled oligonucleotide probes (Supplementary Table S2). For primer extension analysis, 1 μg of total RNA was converted into cDNA using Superscript III (Invitrogen) and 5′-end labeled oligonucleotide probes and separated on 10% PAA/8 M urea sequencing gels. Results were visualized using a PhosphorImager (Typhoon FLA9000; GE Healthcare). ImageQuant (GE Healthcare) was used to quantify northern blot data, which were normalized to levels of the 47S/45S pre-rRNAs.

In vitro RNA cleavage assay

Recombinant proteins were expressed and purified as described in (24), including 1 mM $MnCl_2$ in all media and buffers. Nuclease assays were performed in 10 mM Tris/HCl pH 7.6, 100 mM NaCl, 2 mM dithiothreitol, 100 μg ml^{-1} bovine serum albumin (BSA), 0.8 unit μl^{-1} RNasin, 4.5% glycerol, 0.05% Tween20, 10 ng μl^{-1} *E. coli* tRNA and 5 mM $MnCl_2$. A total of 10 μl reactions containing ~20 pmol of protein were pre-incubated for 5 min at 30°C. The only exception was for the experiment shown in Supplementary Figure S4, for which detailed conditions

are described in the supplementary figure legend S4. *In vitro* transcribed pre-rRNA substrates (0.125 pmol) containing yeast sites D and A2 (35S + 2301 – 2844) or yeast sites A0 and A1 (35S + 335 – 1146) was added and incubated for 1 h at 30°C. Following proteinase K digestion (30 min at 37°C), the RNA was extracted, precipitated and analyzed by primer extension using probes yD-RT, yA2-RT or yA1-RT, respectively. Products were resolved on 10% PAA/8 M urea sequencing gels and visualized by autoradiography.

Protein–protein interaction studies and purification of Rcl1-containing complexes

GST-bait proteins (~50 pmol) were immobilized on glutathione sepharose and incubated with [^{35}S] *in vitro* translates for 1 h at 4°C in Buffer NB (20 mM Tris/HCl pH 7.6, 150 mM NaCl, 8.7% glycerol and 0.1% Tween20). The beads were washed five times with buffer NB. Retained proteins were separated by SDS-PAGE and visualized by autoradiography. Affinity-purification of Rcl1-HTP complexes (from 500 ml of yeast culture each) was performed essentially as described in (25). Complexes were purified on IgG sepharose at 150 mM NaCl and eluted by TEV protease cleavage. Co-purifying proteins and RNAs were analyzed by immunoblotting or northern blotting, respectively (Supplementary Tables S4 and S2).

Yeast immunofluorescence

Yeast strains expressing HA-tagged Rcl1 under control of the repressible *GAL10* promoter were transformed with plasmids encoding His_6-TEV-protein A (HTP) tagged forms of Rcl1 and $Rcl1_{RDK}$ and grown in minimal medium containing glucose for 6 h to deplete the endogenous protein. Cells at mid-log phase were harvested and fixed in 4% formaldehyde in phosphate buffered saline (PBS) (15 min at RT), washed with PBS and then incubated for 45 min at 30°C in buffer B (0.1 M potassium phosphate pH 7.5, 1 M sorbitol, 10 mM DTT) containing 50 U/ml Zymolase. Cells were washed and incubated in batch with the antibodies diluted in PBS + 5% milk (Supplementary Table S4), washed with PBS containing DAPI (4′,6′-diamidino-2-phenylindole) and mounted onto poly-lysine coated coverslips using Vectorshield. All images were obtained using a Zeiss Axiovert 200 microscope with Plan-Apochromat x100 1.4NA objective, Axiovision software and an Axiocam monochrome camera, and processed in Photoshop (Adobe).

RESULTS

Yeast Utp24 crosslinks in close vicinity to site A1, to the U3 snoRNA and pre-rRNA elements required for pseudoknot formation

To dissect the roles of the two yeast candidate pre-rRNA endonucleases Rcl1 and Utp24, we applied *in vivo* RNA–protein crosslinking (CRAC) to identify their RNA binding sites (21) (Figure 1 and Supplementary Figure S2). C-terminal HTP-tagged (His_6–TEV–protA) Rcl1 and Utp24 were expressed from the chromosomal locus under control of the endogenous promoter. Rcl1-HTP supported WT growth, while the Utp24-HTP strain exhibited a mild growth defect. Actively growing cells were UV-irradiated as described (22) and RNA fragments crosslinked to the purified proteins were isolated and analyzed as outlined in Supplementary Figure S2A. Protein recovery was verified by western analysis (Supplementary Figure S2B). Rcl1 purified well, but crosslinked poorly, and no pre-rRNA target sequence was significantly enriched in every experiment. We were, however, able to reproducibly identify Utp24 RNA crosslinking sites in four independent CRAC experiments. Results from two representative data sets are presented in Figure 1 and Supplementary Figure S2.

Transcriptome-wide RNA binding profiles of Utp24 are shown in Figure 1A. The majority of reads (>83%) identified in both Utp24 data sets were mapped to the (pre-)rRNA, with lower numbers of hits in snoRNAs. Hits were also recovered in mRNAs, but no individual mRNA emerged as a likely Utp24 target. Consistent with its known role in SSU biogenesis, Utp24 was predominately associated with sequences within the 18S rRNA (Figure 1B and C). Importantly, this crosslinking profile was not seen in a CRAC experiment performed with an untagged control strain (Figure 1B). The peak near the 3′-end of the 25S rRNA represents a common CRAC contaminant (21).

In CRAC, microdeletions and/or mutations are often introduced at the site of crosslinking during cDNA preparation and can be used to map precise protein-binding sites (26). Sites of microdeletion (Figure 1C and D) were seen at 18S +3, adjacent to cleavage site A1 (peak 3, see Supplementary Figure S2C for higher resolution), around 18S + 545 (peak 4) and around 18S + 820 and 18S + 895 within expansion segment 6 (ES6) (peak 5), close to the binding site for the snoRNA snR30, which is also required for cleavage at sites A0-A2 (27). The highest peak was located at 18S + 1103 in the vicinity of the 3′-side of the central pseudoknot (peak 6). Sites of microdeletion were also identified in the flanking 5′-ETS (Figure 1E) around 35S + 288 (peak 1) and 35S + 460 (peak 2). Strikingly, these positions in the 5′-ETS correspond to the base-pairing sites for the U3 snoRNA 3′-hinge and 5′-hinge regions, respectively (3).

Analysis of snoRNAs crosslinked to Utp24 (Figure 1F) revealed specific enrichment for the U3 snoRNA (encoded by the genes *SNR17A/B*). Utp24 predominately crosslinked to a U3 region (U3 +15–90), which undergoes multiple interactions with the pre-rRNA. Significant peaks of microdeletions in U3 (Figure 1G and H) were seen at +45 and +52 (peaks b, 5′-hinge) and +68 (peak c, 3′-hinge) that base-pair to 5′-ETS regions at +480 and +280, respectively (2,3). The highest peak of microdeletions (peak a) was at +24, a region of U3 predicted to base-pair with 18S +1140 on the 3′-side of the pseudoknot.

We conclude that Utp24 binds both to the U3 snoRNA and to the corresponding U3-binding sequences in the mature 18S rRNA and the 5′-ETS, interactions that are essential for pseudoknot formation and coupled cleavage of sites A0, A1 and A2. Moreover, the binding of Utp24 at 18S +3 is consistent with a direct role in cleavage at site A1, since the precise location of cleavage is partly defined with respect to this position (28).

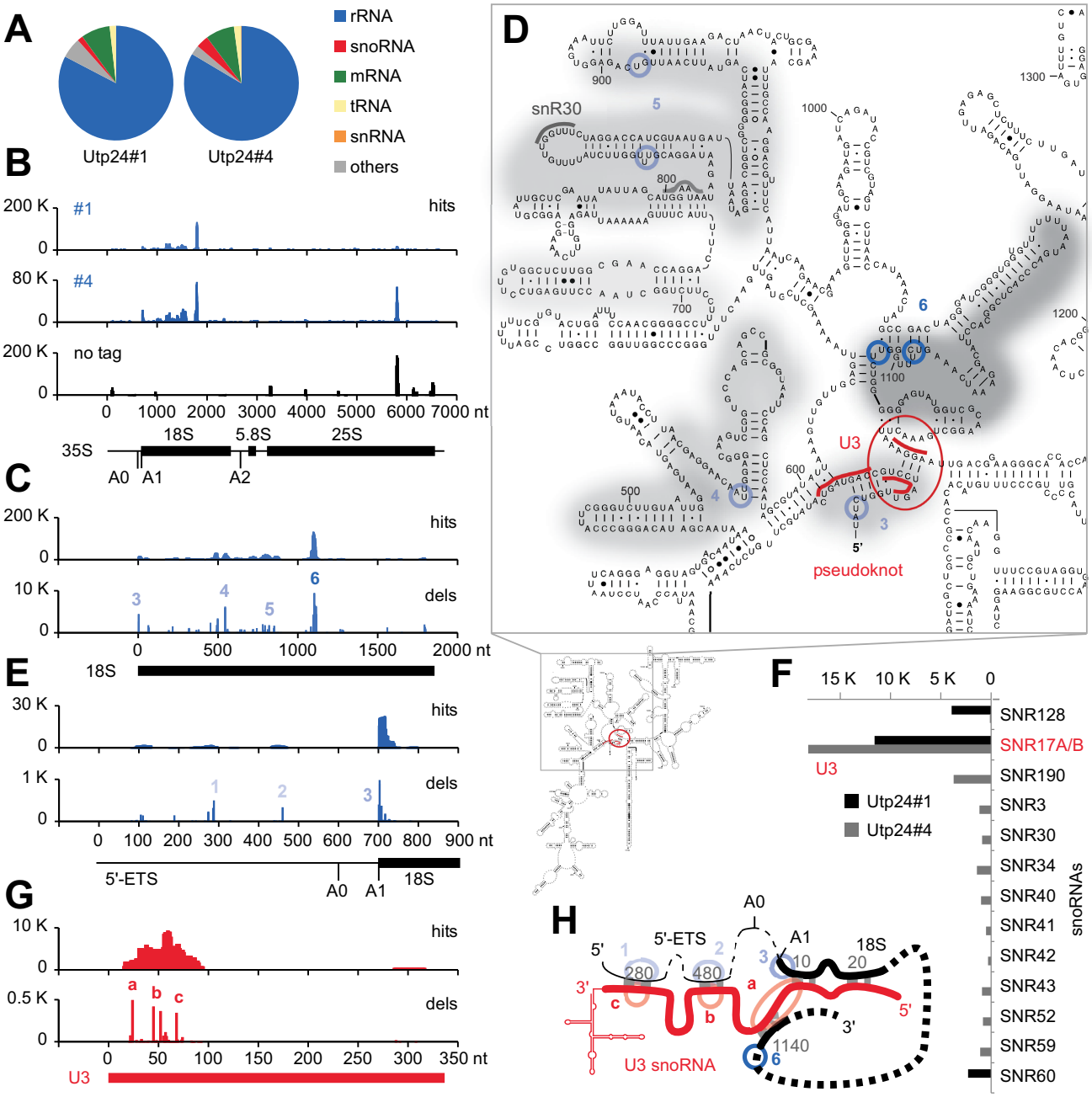


Figure 1. RNA crosslinking sites of yeast Utp24 on (pre-) rRNA and the U3 snoRNA. Illumina sequencing of cDNA libraries generated from crosslinked and subsequently trimmed RNAs recovered with purified Utp24 protein. Normalized data are plotted as reads per million mapped sequences. (**A**) Transcriptome-wide binding profiles. Data of two replicate experiments were mapped to the yeast genome using Novoalign. A total of 979 330 mapped reads were recovered for data set Utp24#1 and 1383819 for data set Utp24#4, respectively. Pie charts illustrate the proportion of all sequences mapped to functional RNA classes (indicated on the right). (**B**) Sequences from data sets Utp24#1 and Utp24#4 and a control experiment with a non-tagged strain (no tag) were aligned with the rDNA (RDN37-1) encoding the 35S pre-rRNA. The frequency of recovery is plotted as hits (total reads) for each individual nucleotide. (**C**, **E** and **G**) RNA binding profiles on (C) the 18S rRNA, (E) the 5′-external transcribed spacer (5′-ETS) and (G) the U3 snoRNA. Hits (upper panels): total reads; dels (lower panels): mutations and microdeletions representing precise binding sites. Microdeletion peaks in the (pre-) rRNA (1-6, blue) and the U3 snoRNA (a–c, red) are labeled. The position of the mature 18S, 5.8S and 25S rRNAs (B, C and E) or the U3 snoRNA (G) are indicated by thick lines. (**D**) Predicted secondary structure of the mature 18S rRNA in *S. cerevisiae*. Utp24 crosslinking sites are marked on the sequence and shades indicate peak height with the highest peak shown in dark grey. Microdeletion peaks (see panel C) are highlighted by shaded blue circles. Binding sites for snoRNAs U3 around the central pseudoknot (red) and snR30 in the ES6 region (grey) are also indicated. (**F**) The distribution of hits mapping to crosslinked snoRNAs is plotted. (**H**) Schematic representation of base-pairing interactions between the pre-rRNA (black) and the U3 snoRNA (red) during pseudoknot formation. The approximate positions of Utp24 microdeletion peaks are indicated in blue (1–6, pre-rRNA, see panels C and E) or red (a–c, U3 snoRNA, see panel G).

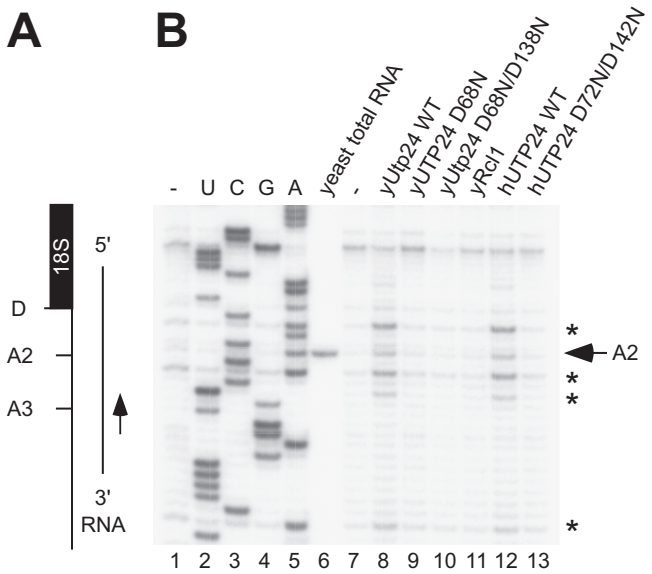


Figure 2. Yeast and human recombinant Utp24 proteins cleave at yeast site A2 *in vitro*. (**A**) Schematic of the RNA substrate mimicking a pre-rRNA fragment before A2 cleavage (yeast 35S +2301-2844). (**B**) *In vitro* transcribed RNA was incubated without recombinant protein (-), wild-type (WT) or mutant Utp24, WT Rcl1 or WT or mutant UTP24 in the presence of 5 mM Mn^{2+} and analyzed by primer extension. The position of the primer is shown in panel A. Non-treated RNA substrate was used to generate a sequencing ladder and endogenous cleavage at site A2 in yeast total RNA is indicated. Recombinant Utp24-mediated cleavages at site A2 and 5′C/A3′ sequences are marked by an arrow and asterisks, respectively.

Utp24 cleaves site A2 in a yeast pre-rRNA substrate *in vitro*

The predicted catalytic center within the Utp24 PIN domain is characterized by four acidic residues (DEDD), which are conserved in all kingdoms of life (29). We therefore assessed the *in vitro* activities of yeast Utp24 and human UTP24 in cleavage assays (Figure 2, Supplementary Figures S3 and S4). WT and PIN domain mutant proteins were expressed in *E. coli* with N-terminal tags containing GST. Double (D72N/D142N) catalytic site mutations in UTP24 were based on yeast Utp24 mutations (D68N, D138N) that cause A1 and A2 cleavage defects *in vivo* (4). The N-terminal GST tag was removed by prescission protease (PP) cleavage following purification (Supplementary Figure S3A).

The proteins were tested in nuclease assays on *in vitro* transcribed RNA containing yeast site A2 (Figure 2A) or A1 (Supplementary Figure S3B). Assays were performed in the presence of 5 mM Mn^{2+}, which is required for the *in vitro* activity of other PIN domain nucleases (7,25,30). The RNA was analyzed by primer extension using a labeled oligonucleotide downstream of the cleavage site.

WT yeast and human Utp24 both cleaved the RNA substrate containing site A2, giving very similar products (Figure 2B, lanes 8 and 12). In contrast, no cleavage activity was observed for the predicted catalytically inactive Utp24 PIN mutants (Figure 2B, lanes 9, 10 and 13), while a different PIN domain endonuclease, Nob1, cleaved site D but not A2 (Supplementary Figure S3D). Utp24-mediated cleavage did not occur exclusively at site A2, but also at several other places with the same sequence context (5′C/A3′, where '/' equals the site of cleavage) around the *bona fide* cleavage site. In contrast, we did not detect clear cleavage at site A1 (Supplementary Figure S3B, see Discussion). An evolutionarily conserved ACAC motif at site A2 was previously identified as a specificity feature for yeast A2 cleavage *in vivo* (31). Since the *in vitro* cleavage pattern was identical for yeast and human Utp24 in our assays, this appears to be a conserved feature of Utp24 PIN nuclease activity.

Yeast Rcl1 was also expressed and purified with a cleavable GST-PP tag (Figure 2 and Supplementary Figure S3A), but no *in vitro* cleavage activity was detected in the presence of 5 mM Mn^{2+} (Figure 2B, lane 11). In an effort to reproduce the reported nuclease activity of Rcl1 (9), we dialyzed the yeast Rcl1 protein, as well as WT and mutant Utp24 proteins, into the buffer optimized for Rcl1 activity (K. Karbstein, personal communication) and performed the nuclease assay in the presence of 10 mM Mg^{2+}. Under these conditions, apparent cleavage at site A2 was observed (Supplementary Figure S4), as reported. However, cleavage was observed with each of the recombinant proteins, including the catalytically inactive Utp24 D68N mutant, indicating that cleavage associated with addition of Utp24 was not specific. We were unable to express and purify a mutant of Rcl1 that is predicted to lack catalytic activity, and were therefore unable to determine whether this is also the case for Rcl1.

An intact PIN domain in human UTP24 is required for accurate cleavages at sites 1 and 2a, but not A0, *in vivo*

Yeast cleavage sites A0–A2 have direct counterparts in humans called sites A0, 1 and 2a (Supplementary Figure S1). RNAi-mediated knockdown of human UTP24 caused 30S pre-rRNA accumulation, indicating that UTP24 is required for these cleavages (11,13). To investigate the putative catalytic role of UTP24, we established an RNAi-rescue system in HEK293 cells. Cells stably expressing the FLAG-tag (pcDNA5) or FLAG-tagged forms of either WT UTP24, or UTP24 with single (D72N) or double (D72N/D142N) catalytic site mutations were generated (Figure 3 and Supplementary Figure S5). These carried silent mutations in the open reading frames rendering them resistant to RNAi-mediated depletion of the endogenous protein. The RNAi-resistant proteins were expressed at endogenous levels using a titratable *TET* promoter. Cells were transfected with either a siRNA specifically targeting endogenous UTP24 or a control siRNA targeting firefly luciferase (GL2) (32). After siRNA treatment for 72 h, the expression of endogenous and FLAG-tagged UTP24 proteins was analyzed by immunoblotting (Supplementary Figure S5A). The UTP24-specific siRNA significantly reduced endogenous protein levels, whereas the RNAi-resistant tagged UTP24 proteins were unaffected.

Total RNA was extracted from RNAi-treated cells and pre-rRNA processing analyzed by northern hybridization using probes complementary to the 5′-end of ITS1 ('h18SE') or downstream of the 2a cleavage site ('hITS1') (Figure 3B and Supplementary Figure S5B). Depletion of UTP24 resulted in 30S accumulation (Figure 3B, lane 2), and this phenotype was almost completely rescued by ex-

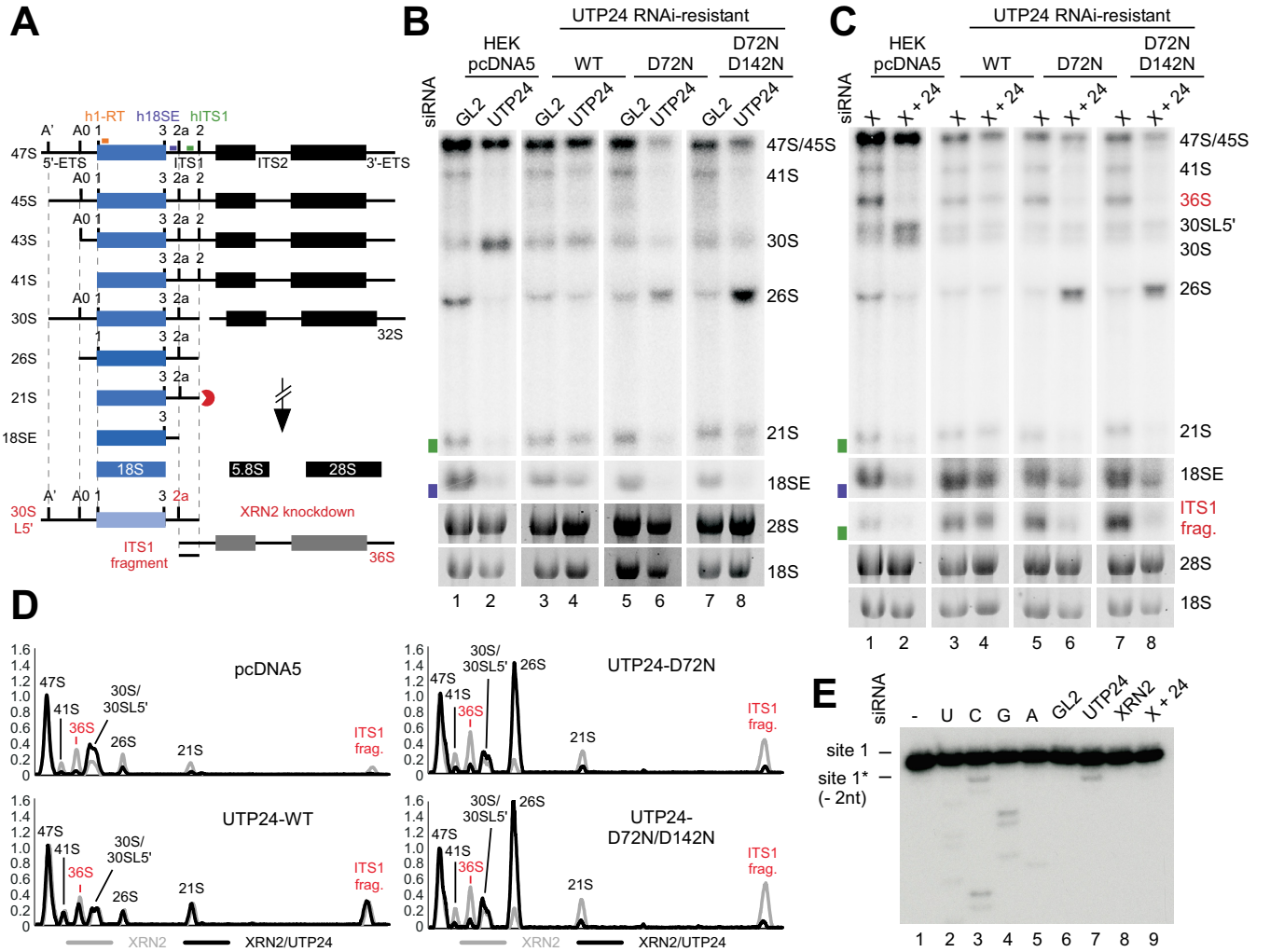


Figure 3. An intact UTP24 PIN domain is required for efficient site 1 and 2a cleavages in the human pre-ribosomal RNA. (**A**) Key steps in human ribosome biogenesis. RNA intermediates of the minor, 2a-dependent pathway accumulating in the absence of XRN2 are shown in red. Radiolabeled probes used for primer extension (h1-RT; orange) or northern blotting (h18SE; purple and hITS1; green) are marked above the 47S precursor. ETS: external transcribed spacer; ITS: internal transcribed spacer. (**B**) HEK293 cells were stably transfected with plasmids expressing the FLAG tag (pcDNA5) or WT or mutant forms of FLAG-UTP24 (D72N, D72N/D142N). RNA from control cells (GL2) or those depleted of endogenous UTP24 (UTP24) by RNAi, was analyzed by northern blotting using probes hybridizing to the 5′-end of ITS1 (h18SE, purple rectangle) or downstream of 2a (hITS1, green rectangle). Pre-rRNAs were detected using a PhosphorImager and total rRNA (28S/18S) was visualized by ethidium bromide staining. RNA species are marked on the right. (**C**) RNA extracted from HEK293 cell lines as in panel B, but depleted of XRN2 alone (XRN2) or XRN2 and UTP24 (X+24) were analyzed as in B. (**D**) RNA levels from panel C were normalized to the 47S/45S pre-rRNAs and plotted for each XRN2-UTP24 double knockdown (black) and the single XRN2 knockdown (grey). The identity of each peak is indicated. Red: RNA species accumulating in the absence of XRN2. (**E**) RNA extracted from HEK293 cell expressing the FLAG-UTP24 D72N mutant, either treated with control siRNA (GL2) or depleted of UTP24 (UTP24), XRN2 (XRN2) or both (X+24), was analyzed by primer extension using probe h1-RT (panel A). Total RNA from control cells (GL2) expressing the FLAG tag alone (pcDNA5) was used to generate a sequencing ladder. Positions of the natural site 1 and 2 nt downstream are indicated on the left.

pression of the WT protein (lane 4). Complementation with UTP24 D72N or D72N/D142N mutants caused accumulation of the 26S pre-rRNA (lanes 6 and 8), indicative of strongly reduced cleavage at sites 1 and 2a. No dominant negative phenotypes were observed upon expression of the PIN mutants in the presence of the endogenous protein (lanes 5 and 7).

The 26S RNA accumulation indicated that the catalytic activity of human UTP24 is required for site 1 and 2a cleavage *in vivo*, while cleavage at site A0 is unaffected. However, 2a cleavage cannot readily be directly assessed because the 36S pre-rRNA and the excised ITS1 fragment, the only intermediates specific to the 'minor' pathway (lighter shades in Figure 3A), are barely detectable in control cells. However, depletion of the 5′–3′-exonuclease XRN2 stimulates processing through this pathway (11,19,33). We, therefore, performed UTP24 knockdown and complementation in combination with XRN2 depletion. Pre-rRNA intermediates were detected by northern hybridization with probe hITS1 (Figure 3C), and quantified using a PhosphorImager. Levels were normalized to the 47S/45S pre-rRNA and plotted to compare the single XRN2 and UTP24/XRN2 double knockdowns (Figure 3D). Depletion of XRN2 alone caused strong accumulation of the 36S pre-rRNA and the ITS1

fragment accompanied by the appearance of the 30SL5' precursor (also referred to as 34S (12) (Figure 3C, lanes 1, 3, 5 and 7), reflecting an A' cleavage defect. Double-knockdown of UTP24 and XRN2 (lane 2) resulted in a strong decrease in the 36S precursor and the ITS1 fragment compared to the single XRN2 knockdown, while expression of the WT UTP24 protein rescued the 2a processing defect (lane 4). Importantly, the 36S pre-rRNA and the excised ITS1 fragment were also severely reduced in double knockdown cells expressing UTP24 D72N (lane 6) or UTP24 D72N/D142N (lane 8), whereas 26S pre-rRNA levels were strongly increased.

Low levels of the 21S precursor were detected in UTP24 PIN mutant cells, suggesting that an alternative mechanism can generate the 18S rRNA 5′-end (site 1) in the absence of endonuclease cleavage. XRN2 degrades the 3′-fragment of the 5′-ETS ('ETS3') (13) and might also digest the 5′-ETS back to site 1 following A0 cleavage. To test this model, RNA was extracted from cells expressing UTP24 D72N, treated with a siRNA against endogenous UTP24, XRN2 or both, and analyzed by primer extension through site 1 (Figure 3E). The major stop at site 1 reflects mature 18S rRNA produced prior to the siRNA treatment. In cells treated with the UTP24 siRNA (lane 7), a weaker primer extension stop was visible two nucleotides downstream of site 1. This was significantly reduced in the UTP24/XRN2 double knockdown (lane 9) and absent when the catalytically inactive UTP24 was not expressed (Supplementary Figure S5C). We speculate that in the absence of site 1 cleavage, XRN2 degrades the 5′-ETS region, but can be blocked by catalytically inactive UTP24 bound at the 5′-end of 18S, leading to the observed truncated product.

Our analyses indicate that point mutations in the PIN domain of UTP24 block cleavage at sites 1 and 2a in human pre-rRNA, while cleavages at sites A0 and A' are unaffected.

Mutation of the proposed pre-rRNA substrate binding site within human RCL1 does not affect 18S production or 2a cleavage

Yeast Rcl1 was also reported to cleave at site A2 (9), suggesting the possibility of redundant activities. To determine whether RCL1 participates in human site 2a cleavage, we established an RCL1 RNAi-rescue system in HEK293 cells, as described for UTP24 (Figure 4 and Supplementary Figure S6). Endogenous and FLAG-tagged RCL1 levels were monitored by immunoblotting (Supplementary Figure S6A). RNAi-mediated depletion of RCL1 caused 30S accumulation (Figure 4A, lane 2), that was rescued by expression of RNAi-resistant FLAG-tagged WT RCL1 (Figure 4A, lane 4).

In yeast Rcl1 an 'RDK-AAA' mutant showed a strong A2 cleavage defect, which was proposed to be due to impaired pre-rRNA substrate binding (9). A stable cell line was constructed expressing RCL1 with the equivalent 'RHK' residues (R329, H330, K332) changed to alanines. Following RCL1 depletion, $RCL1_{RHK\text{-}AAA}$ fully rescued the 30S phenotype, showing that integrity of the proposed RNA substrate-binding pocket is not required for human 18S maturation (Figure 4A, lane 6).

As noted above, XRN2 knockdown increases dependency on 2a cleavage, so the RCL1 knockdowns were repeated in combination with XRN2 depletion (Figure 4B). Following RNAi-mediated knockdown of XRN2 and RCL1, WT RCL1 and the $RCL1_{RHK}$ mutant again rescued the pre-rRNA processing phenotype to a similar extent (Figure 4B, lanes 4 and 6). Quantification of pre-rRNA levels normalized to the 47S/45S pre-rRNAs confirmed that the 2a-dependent pre-rRNA intermediates, the 36S pre-RNA and the ITS1 fragment, are restored to similar levels in cells expressing WT and mutant RCL1, but severely reduced in cells expressing the FLAG-tag only (Figure 4C).

Yeast Rcl1 mutations impair interactions with Bms1, integration into the SSU processome and subcellular localization

The finding that the human equivalent of the yeast $Rcl1_{RDK}$ mutant supports pre-rRNA processing prompted us to re-examine the ribosome synthesis defect in yeast (Figure 5). Recent mutational analysis of Rcl1 based on the crystal structure of the yeast Rcl1-Bms1 dimer showed that mutation of R237, within the RDK motif, disrupts interaction with Bms1 (10). The Rcl1-Bms1 interaction is required for nuclear import of Rcl1 (34) and, consistent with this, an R237A mutation severely impaired the nucleolar localization of yeast Rcl1 (10).

WT Rcl1 and $Rcl1_{RDK}$ were generated by *in vitro* translation in the presence of [^{35}S] methionine. Incubation with recombinant GST-tagged $Bms1^{(aa1\text{-}705)}$ (Figure 5A) revealed a strong and reproducible interaction between the N-terminal region of Bms1 and the WT Rcl1 protein, whereas no binding was detected for the $Rcl1_{RDK}$ mutant.

This finding suggested that $Rcl1_{RDK}$ might not be incorporated into the SSU processome *in vivo*. To test this, we generated a yeast strain expressing HA-tagged Rcl1 under control of the repressible *GAL10* promoter, allowing depletion of the endogenous Rcl1 protein. This strain was then transformed with plasmids encoding His_6-TEV-protein A (HTP) tagged forms of WT Rcl1 and $Rcl1_{RDK}$, or an empty plasmid (pRS316) and grown in minimal medium containing glucose. Rcl1-containing complexes were purified on IgG sepharose followed by TEV cleavage, and associated proteins and RNA were detected by western and northern blotting, respectively (Figure 5). This revealed that WT Rcl1 and $Rcl1_{RDK}$ were expressed and purified to similar levels. However, association of $Rcl1_{RDK}$ with the box C/D snoRNP-specific protein Nop1 (Figure 5B) as well as the U3 snoRNA and several pre-rRNA processing intermediates (23S, 22S/21S, 20S) (Figure 5C) was severely reduced, confirming that the $Rcl1_{RDK}$ mutant fails to associate with the SSU processome. To support this finding, we performed immunofluorescence in cells expressing HTP-tagged WT Rcl1 and $Rcl1_{RDK}$ (Figure 5D). WT Rcl1 localized to the nucleolus as expected, whereas $Rcl1_{RDK}$ was mainly retained in the cytoplasm.

We conclude that pre-rRNA processing defects associated with yeast $Rcl1_{RDK}$ are likely due to impaired nuclear import and SSU processome formation, due to the loss of binding and recruitment via Bms1.

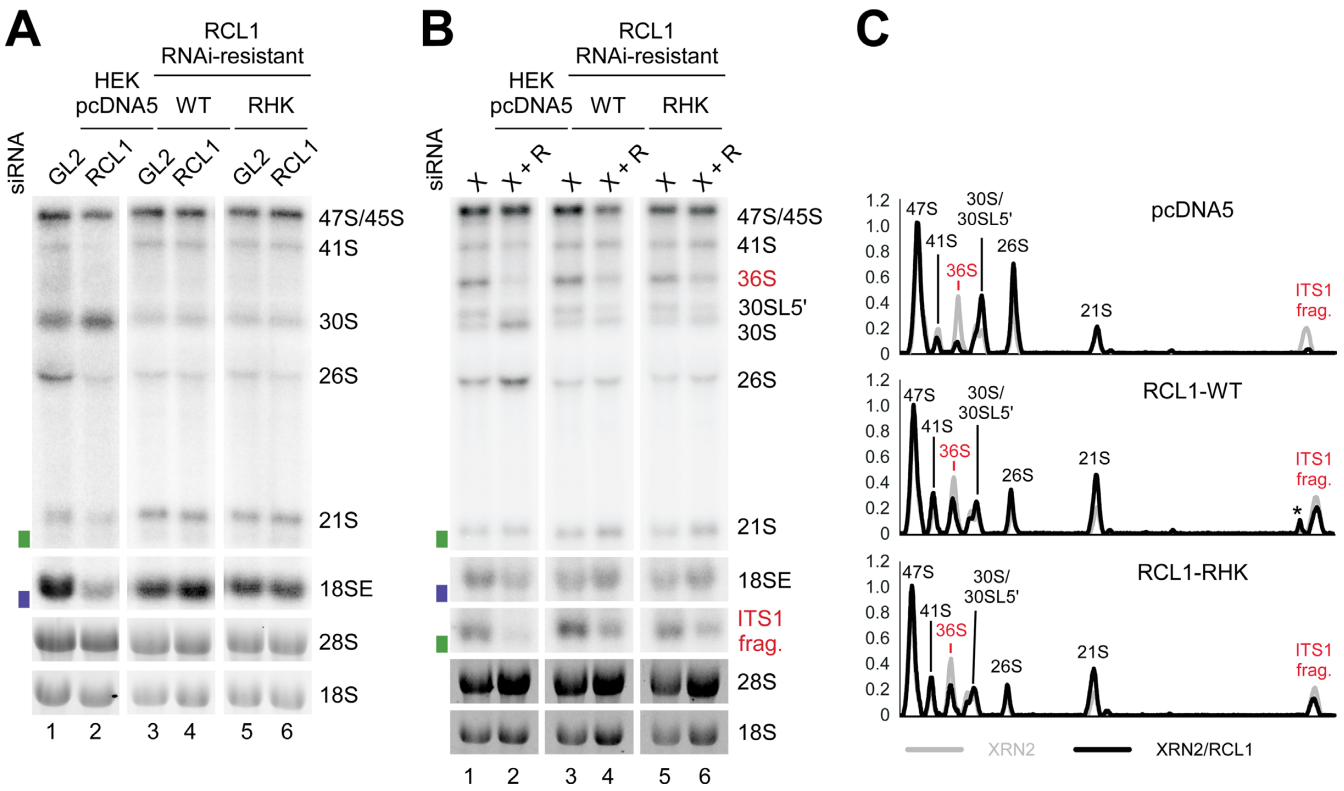


Figure 4. An intact RCL1 RHK domain is not required for human 18S production. (**A**) HEK293 cells were stably transfected with plasmids expressing the FLAG tag alone (pcDNA5) or WT FLAG-RCL1 or the FLAG-RCL1-RHK mutant (RHK). The FLAG-RCL1 mRNAs were rendered resistant to the RCL1 siRNA. RNA was extracted from control cells (GL2), or those depleted of endogenous RCL1 (RCL1) by RNAi and analyzed by northern blotting using a probe hybridizing to the 5′-end of ITS1 (h18SE, purple rectangle) or downstream of site 2a (hITS1, green rectangle). (**B**) RNA extracted from cells as listed in panel A, but depleted of XRN2 alone (XRN2) or XRN2 and RCL1 (X+R), was analyzed as in A. (**C**) The levels of the pre-rRNA intermediates from panel B were normalized to the 47S/45S precursors and plotted for each XRN2-RCL1 double knockdown (black) and the single XRN2 knockdown (grey). The identity of each peak is indicated. RNA intermediates accumulating in the absence of XRN2 are shown in red. Asterisk: non-specific signal.

DISCUSSION

Here, we present data implicating Utp24 as the endonuclease responsible for early pre-rRNA cleavages at sites A1/1 and A2/2a that generate the major precursor to the 18S rRNA.

In vivo UV crosslinking was performed in actively growing cells to identify *bona fide* RNA targets for yeast Utp24 (Figure 1 and Supplementary Figure S2). Utp24 RNA binding was strongest in the region of the 18S rRNA central pseudoknot, a highly conserved structural unit that forms the core of the small ribosomal subunit. Additional crosslinking was observed close to site A1 and in the 5′-ETS pre-rRNA spacer region. Pseudoknot formation is guided by U3 and is required for A0, A1 and A2 cleavages (3). Strikingly, Utp24 was recovered in association with the U3 snoRNA regions that base-pair with the pre-RNA as well as with the U3-binding sites in the 5′-ETS and the 18S elements forming the proximal and distal sides of the pseudoknot (see Figure 1H). These finding indicate a role for U3 in Utp24 recruitment to the A1 cleavage site and, potentially, a role for Utp24 in verification of U3-rRNA interactions.

Utp24 was also crosslinked to the eukaryotic expansion segment 6 (ES6) within the 18S rRNA central domain (Figure 1D). The ES6 region has recently emerged as a binding hub for many ribosome biogenesis factors including the RNA helicase Dhr1, which is required to dissociate U3 from the pre-rRNA following cleavages at sites A0 and A1, and preceding A2 cleavage (35). Utp24 and Dhr1 both crosslink to the same U3 snoRNA regions, which engage in pre-rRNA interactions (35). It is possible that sequential binding of both proteins to U3 might be important to guide structural rearrangements to promote U3 release and/or positioning of Utp24 for A1 and A2 cleavages.

We did not detect clear Utp24 crosslinking around site A2 in the ITS1 sequence. We speculate that Utp24 is recruited to stable docking points within the mature 18S rRNA sequence, and only transiently interacts with the cleavage sites during catalysis. The observed coupling of cleavage at sites A1 and A2 suggests that they are brought into proximity by transient changes in pre-ribosome structure, which remain to be determined.

In vitro endonuclease assays revealed that recombinant Utp24 proteins from yeast and human each show activity on a pre-rRNA substrate containing yeast site A2 (Figure 2). As for other PIN domain proteins (7,25,30), this activity required Mn^{2+} ions and the conserved metal-binding amino acids of the PIN domain. Interestingly, yeast and human Utp24 not only cleaved the authentic site A2 (5′-AAC/ACAC-3′), but also at other positions in a similar sequence context (Figure 2). The cleavage pattern is identical in both proteins and therefore likely specific to the

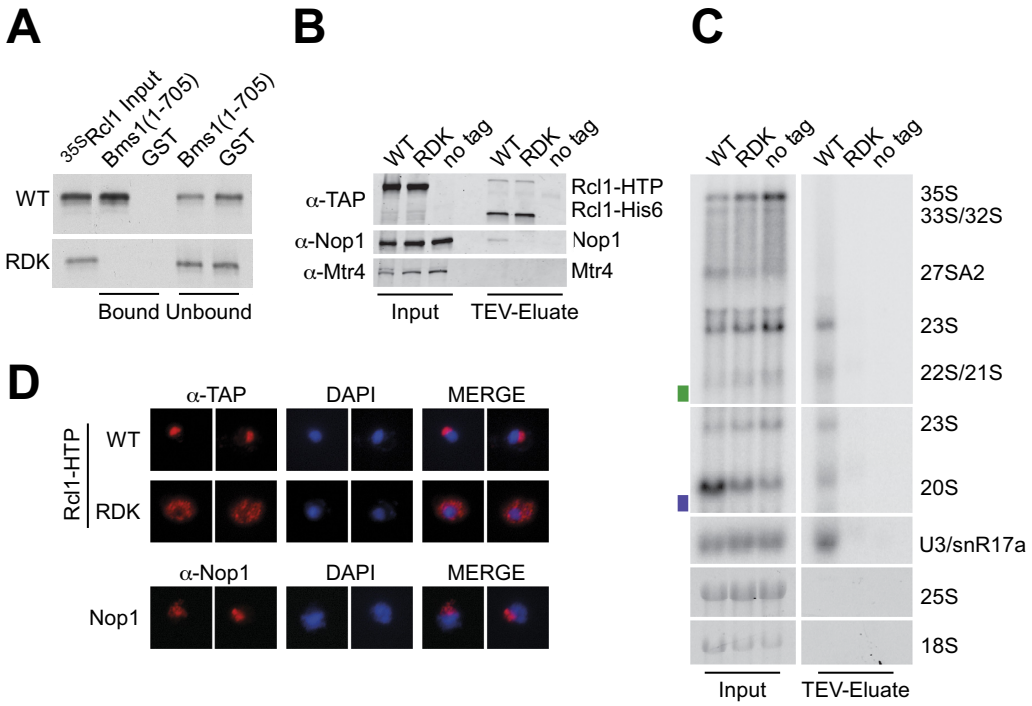


Figure 5. The yeast $Rcl1_{RDK}$ mutant fails to interact with Bms1 *in vitro* and interferes with SSU processome integration and nucleolar localization *in vivo*. (**A**) Recombinant GST-tagged $Bms1^{(aa1\text{-}705)}$ and GST were coupled to glutathione sepharose and incubated with WT and mutant Rcl1 generated by *in vitro* translation in the presence of [^{35}S] methionine. Bound material was eluted under denaturing conditions, separated by SDS-PAGE and analyzed by autoradiography. A total of 10% of the input and unbound material is also loaded. (**B** and **C**) Rcl1-containing complexes were purified on IgG sepharose and eluted by TEV protease cleavage. (B) Co-purifying proteins were analyzed by SDS-PAGE and immunoblotting using antibodies against the C-terminus of the Rcl1-HTP tag after TEV cleavage (anti-TAP), the box C/D snoRNP component Nop1, or Mtr4 as a control. A total of 2.5% of the input material is also loaded. (C) Co-purified RNA was extracted and analyzed by northern blotting using a probe hybridizing to the U3 snoRNA or two regions within the pre-rRNA (yA2-A3, green rectangle or yD-A2, purple rectangle; see Supplementary Figure S1A for location of the probes). Mature 18S and 25S rRNAs were visualized by ethidium bromide staining. A total of 1% of the input material is also loaded. (**D**) Yeast strains expressing plasmid-encoded HTP-tagged WT Rcl1 (top row) or the $Rcl1_{RDK}$ mutant (middle row) were harvested and processed for immunofluorescence miscroscopy using the anti-TAP antibody. Lower row: Cells stained with an antibody against endogenous Nop1 as a nucleolar marker. From left to right: Immunofluorescence signal (red), DAPI (blue) and the merged image.

conserved Utp24 PIN domain. Very low levels of cleavage at these positions around site A2 are also observed *in vivo* (31), but they are underrepresented compared to the *bona fide* A2 site. Bound transacting factors within the SSU processome therefore likely assist the specificity of Utp24-mediated cleavage *in vivo*. The recently mapped 2a site in the human pre-rRNA (5′-C/GAC/GC-3) (19) appears to be very different to yeast A2. However, cleavage takes place at 5′C/purine3′ in both cases. We also tested RNA substrates including yeast site A1, but observed only low levels of cleavage in ACAC-rich regions within the 5′-ETS close to the U3 base-pairing sites (Supplementary Figure S3B and data not shown). Notably, the sequence context of site A1 (5′-GU/UA-3′) does not match yeast A2 or human 2a, highlighting the importance of additional factors and/or structural arrangements within the pre-rRNA that might be needed to facilitate Utp24-mediated cleavage at this site *in vivo*.

Consistent with yeast data, *in vivo* studies on human UTP24 showed that an intact UTP24 PIN domain was required for cleavage at the 5′-end of 18S (site 1) and at site 2a within the ITS1 spacer, but not at A0 or A' in the 5′-ETS (Figure 3). The requirement for UTP24 in site 1 cleavage and the presence of an alternative pathway to process to the 5′-end of 18S in the absence of endonuclease cleavage by UTP24 was recently reported (15). However, in contrast to our results, XRN2 was not implicated in 5′-3′-exonuclease trimming to generate the aberrant 5′-end of 18S. This discrepancy might reflect the use of stable shRNA expression by Tomecki *et al.* (15), which strongly impaired cellular viability and could result in secondary effects. Notably, inspection of published northern analyses reveals phenotypes that link UTP24 to 2a cleavage (Supplementary Figure S10 in (15)).

Both Rcl1 and Utp24 have previously been proposed to cleave site A2 in yeast (4,5,9). A2 cleavage can occur either co- or post-transcriptionally, so two separate enzymes might act in these distinct contexts. The catalytic role of yeast Rcl1 in A2 cleavage was reported to be dependent on an intact RDK motif, which was proposed to mediate RNA substrate binding (9). We found, however, that the $Rcl1_{RDK}$ mutation blocks binding to the SSU processome component Bms1 *in vitro* (Figure 5A), resulting in reduced nuclear import (Figure 5D) and the loss of stable incorporation into the processome (Figure 5B and C) *in vivo*. An intact SSU processome is required for A0-A2 cleavage potentially explaining the pre-rRNA processing defect in strains expressing $Rcl1_{RDK}$. Mutation of the equivalent motif in human

RCL1 ('RHK') did not affect 18S production (Figure 4). However, human SSU processome assembly appears to differ from yeast, since RCL1 associates prior to A' cleavage, while BMS1 largely binds after cleavage at A' (36,37).

This work has implicated Utp24 as the endonuclease responsible for two coupled pre-rRNA cleavages in 18S rRNA maturation. The enzymes responsible for cleavage at yeast and human site A0 and human A' are probably also among the known SSU processome components, but their identities remain an intriguing mystery.

ACCESSION NUMBERS

GEO database (http://www.ncbi.nlm.nih.gov/geo/): identifier GSE75991.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Royal Society [UF100666 to C.S., RG110357 to C.S.]; Wellcome Trust [077248 to D.T., 092076 to N.J.W.]; BBSRC project [BB/F006853/1 to N.J.W]; BBSRC DTP studentship [to D.C.]; ERASMUS internship [to F.W.]. Funding for open access charge: BBSRC [BB/F006853/1 to N.J.W.].

Conflict of interest statement. None declared.

REFERENCES

1. Henras,A.K., Plisson-Chastang,C., O'Donohue,M.F., Chakraborty,A. and Gleizes,P.E. (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA*, **6**, 225–242.
2. Phipps,K.R., Charette,J. and Baserga,S.J. (2011) The small subunit processome in ribosome biogenesis-progress and prospects. *Wiley Interdiscip. Rev. RNA*, **2**, 1–21.
3. Watkins,N.J. and Bohnsack,M.T. (2012) The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA*, **3**, 397–414.
4. Bleichert,F., Granneman,S., Osheim,Y.N., Beyer,A.L. and Baserga,S.J. (2006) The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 9464–9469.
5. Billy,E., Wegierski,T., Nasr,F. and 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. *EMBO J.*, **19**, 2115–2126.
6. Fatica,A., Oeffinger,M., Dlakic,M. and Tollervey,D. (2003) Nob1p is required for cleavage of the 3′ end of 18S rRNA. *Mol. Cell. Biol.*, **23**, 1798–1807.
7. Pertschy,B., Schneider,C., Gnadig,M., Schafer,T., Tollervey,D. and Hurt,E. (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J. Biol. Chem.*, **284**, 35079–35091.
8. Lebaron,S., Schneider,C., van Nues,R.W., Swiatkowska,A., Walsh,D., Bottcher,B., Granneman,S., Watkins,N.J. and Tollervey,D. (2012) Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nat. Struct. Mol. Biol.*, **19**, 744–753.
9. Horn,D.M., Mason,S.L. and Karbstein,K. (2011) Rcl1 protein, a novel nuclease for 18 S ribosomal RNA production. *J. Biol. Chem.*, **286**, 34082–34087.
10. Delprato,A., Al Kadri,Y., Perebaskine,N., Monfoulet,C., Henry,Y., Henras,A.K. and Fribourg,S. (2014) Crucial role of the Rcl1p-Bms1p interaction for yeast pre-ribosomal RNA processing. *Nucleic Acids Res.*, **42**, 10161–10172.
11. Sloan,K.E., Mattijssen,S., Lebaron,S., Tollervey,D., Pruijn,G.J. and Watkins,N.J. (2013) Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. *J. Cell Biol.*, **200**, 577–588.
12. Tafforeau,L., Zorbas,C., Langhendries,J.L., Mullineux,S.T., Stamatopoulou,V., Mullier,R., Wacheul,L. and Lafontaine,D.L. (2013) The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol. Cell*, **51**, 539–551.
13. Sloan,K.E., Bohnsack,M.T., Schneider,C. and Watkins,N.J. (2014) The roles of SSU processome components and surveillance factors in the initial processing of human ribosomal RNA. *RNA*, **20**, 540–550.
14. Wang,M., Anikin,L. and Pestov,D.G. (2014) Two orthogonal cleavages separate subunit RNAs in mouse ribosome biogenesis. *Nucleic Acids Res.*, **42**, 11180–11191.
15. Tomecki,R., Labno,A., Drazkowska,K., Cysewski,D. and Dziembowski,A. (2015) hUTP24 is essential for processing of the human rRNA precursor at site A1, but not at site A0. *RNA Biol.*, **12**, 1010–1029.
16. Sharma,K. and Tollervey,D. (1999) Base pairing between U3 small nucleolar RNA and the 5′ end of 18S rRNA is required for pre-rRNA processing. *Mol. Cell. Biol.*, **19**, 6012–6019.
17. Kos,M. and Tollervey,D. (2010) Yeast pre-rRNA processing and modification occur cotranscriptionally. *Mol. Cell*, **37**, 809–820.
18. Lazdins,I.B., Delannoy,M. and Sollner-Webb,B. (1997) Analysis of nucleolar transcription and processing domains and pre-rRNA movements by in situ hybridization. *Chromosoma*, **105**, 481–495.
19. Preti,M., O'Donohue,M.F., Montel-Lehry,N., Bortolin-Cavaille,M.L., Choesmel,V. and Gleizes,P.E. (2013) Gradual processing of the ITS1 from the nucleolus to the cytoplasm during synthesis of the human 18S rRNA. *Nucleic Acids Res.*, **41**, 4709–4723.
20. Gietz,D., St Jean,A., Woods,R.A. and Schiestl,R.H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.*, **20**, 1425.
21. Granneman,S., Kudla,G., Petfalski,E. and Tollervey,D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 9613–9618.
22. Granneman,S., Petfalski,E. and Tollervey,D. (2011) A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. *EMBO J.*, **30**, 4006–4019.
23. Wlotzka,W., Kudla,G., Granneman,S. and Tollervey,D. (2011) The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J.*, **30**, 1790–1803.
24. Schneider,C., Anderson,J.T. and Tollervey,D. (2007) The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. *Mol. Cell*, **27**, 324–331.
25. Schneider,C., Leung,E., Brown,J. and 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 Res.*, **37**, 1127–1140.
26. Schneider,C., Kudla,G., Wlotzka,W., Tuck,A. and Tollervey,D. (2012) Transcriptome-wide analysis of exosome targets. *Mol. Cell*, **48**, 422–433.
27. Fayet-Lebaron,E., Atzorn,V., Henry,Y. and Kiss,T. (2009) 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. *EMBO J.*, **28**, 1260–1270.
28. Venema,J., Henry,Y. and Tollervey,D. (1995) Two distinct recognition signals define the site of endonucleolytic cleavage at the 5′-end of yeast 18S rRNA. *EMBO J.*, **14**, 4883–4892.
29. Rempola,B., Karkusiewicz,I., Piekarska,I. and Rytka,J. (2006) Fcf1p and Fcf2p are novel nucleolar Saccharomyces cerevisiae proteins involved in pre-rRNA processing. *Biochem. Biophys. Res. Commun.*, **346**, 546–554.
30. Skruzny,M., Schneider,C., Racz,A., Weng,J., Tollervey,D. and Hurt,E. (2009) An endoribonuclease functionally linked to perinuclear mRNP quality control associates with the nuclear pore complexes. *PLoS Biol.*, **7**, e8.

31. Allmang,C., Henry,Y., Wood,H., Morrissey,J.P., Petfalski,E. and Tollervey,D. (1996) Recognition of cleavage site A(2) in the yeast pre-rRNA. *RNA*, **2**, 51–62.
32. Elbashir,S.M., Harborth,J., Weber,K. and Tuschl,T. (2002) Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods*, **26**, 199–213.
33. Wang,M. and Pestov,D.G. (2011) 5′-end surveillance by Xrn2 acts as a shared mechanism for mammalian pre-rRNA maturation and decay. *Nucleic Acids Res.*, **39**, 1811–1822.
34. Karbstein,K., Jonas,S. and Doudna,J.A. (2005) An essential GTPase promotes assembly of preribosomal RNA processing complexes. *Mol. Cell*, **20**, 633–643.
35. Sardana,R., Liu,X., Granneman,S., Zhu,J., Gill,M., Papoulas,O., Marcotte,E.M., Tollervey,D., Correll,C.C. and Johnson,A.W. (2015) The DEAH-box helicase Dhr1 dissociates U3 from the pre-rRNA to promote formation of the central pseudoknot. *PLoS Biol.*, **13**, e1002083.
36. Turner,A.J., Knox,A.A., Prieto,J.L., McStay,B. and Watkins,N.J. (2009) A novel small-subunit processome assembly intermediate that contains the U3 snoRNP, nucleolin, RRP5, and DBP4. *Mol. Cell. Biol.*, **29**, 3007–3017.
37. Turner,A.J., Knox,A.A. and Watkins,N.J. (2012) Nucleolar disruption leads to the spatial separation of key 18S rRNA processing factors. *RNA Biol.*, **9**, 175–186.

Nucleic Acids Research, 2017 ***1–14***
doi: 10.1093/nar/gkw1344

The ribosome biogenesis factor yUtp23/hUTP23 coordinates key interactions in the yeast and human pre-40S particle and hUTP23 contains an essential PIN domain

Graeme R. Wells, Franziska Weichmann, Katherine E. Sloan, David Colvin, Nicholas J. Watkins and Claudia Schneider*

Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK

Received October 12, 2016; Revised December 21, 2016; Editorial Decision December 22, 2016; Accepted December 22, 2016

ABSTRACT

Two proteins with PIN endonuclease domains, yUtp24(Fcf1)/hUTP24 and yUtp23/hUTP23 are essential for early pre-ribosomal (r)RNA cleavages at sites A0, A1/1 and A2/2a in yeast and humans. The yUtp24/hUTP24 PIN endonuclease is proposed to cleave at sites A1/1 and A2/2a, but the enzyme cleaving at site A0 is not known. Yeast yUtp23 contains a degenerate, non-essential PIN domain and functions together with the snR30 snoRNA, while human hUTP23 is associated with U17, the human snR30 counterpart. Using *in vivo* RNA–protein crosslinking and gel shift experiments, we reveal that yUtp23/hUTP23 makes direct contacts with expansion sequence 6 (ES6) in the 18S rRNA sequence and that yUtp23 interacts with the 3′ half of the snR30 snoRNA. Protein–protein interaction studies further demonstrated that yeast yUtp23 and human hUTP23 directly interact with the H/ACA snoRNP protein yNhp2/hNHP2, the RNA helicase yRok1/hROK1(DDX52), the ribosome biogenesis factor yRrp7/hRRP7 and yUtp24/hUTP24. yUtp23/hUTP23 could therefore be central to the coordinated integration and release of ES6 binding factors and likely plays a pivotal role in remodeling this pre-rRNA region in both yeast and humans. Finally, studies using RNAi-rescue systems in human cells revealed that intact PIN domain and Zinc finger motifs in human hUTP23 are essential for 18S rRNA maturation.

INTRODUCTION

Eukaryotic ribosomal (r)RNAs are processed from an initial 35S (*Saccharomyces cerevisiae*) or 47S (*Homo sapiens*) precursor (pre-rRNA) by a series of endonucleolytic cleavages followed by exonucleolytic trimming, which results in the concomitant removal of external (5′ ETS, 3′ ETS) and internal (ITS1, ITS2) transcribed spacer sequences (Figure 1 and Supplementary Figure S1) (1). The early pre-rRNA cleavages, at sites A0, A1 and A2 in yeast and A', A0, A1/1 and 2a/E in humans, are critical for 18S rRNA maturation and require the small subunit (SSU) processome, a large ribonucleoprotein complex (2). The U3 small nucleolar (sno)RNA, a key component of the SSU processome, base-pairs with the 5′ ETS and 18S rRNA sequences to guide the formation of the conserved central pseudoknot, which is an essential feature of the 40S ribosomal subunit (2,3).

Many of the factors involved in ribosome biogenesis are essential for pre-rRNA cleavages, which has made the identification of the nuclease activities responsible for these cleavage events difficult. The PIN (PilT N-terminus) endonucleases yUtp24/Fcf1 (hUTP24 in humans) and yNob1 (hNOB1 in humans) are critical for 18S rRNA processing. Early pre-rRNA cleavages at three sites, A0, A1/1 and A2/2a, require the presence of yUtp24/hUTP24 and the PIN endonuclease domain of the protein is essential for cleavages at A1/1 and A2/2a (4–7). The yeast and human proteins also specifically cleave site A2 *in vitro* and yeast yUtp24 was shown to crosslink to both the U3 snoRNA and close to the A1 cleavage site in the pre-rRNA *in vivo* (7). yNob1/hNOB1 functions later in the 18S rRNA processing pathway and catalyzes the removal of the final part of ITS1 from the 3′ end of the 18S rRNA (site D/3) in the cytoplasm (5,8–11). Only the nucleases that cleave at the A0

*To whom correspondence should be addressed. Tel: +44 191 208 7708; Fax: +44 191 208 7424; Email: Claudia.Schneider@ncl.ac.uk
Present Address:
Franziska Weichmann. Lehrstuhl für Biochemie I, Fakultät für Biologie und Vorklinische Medizin, Universität Regensburg, Regensburg, 93053, Germany.
Katherine E. Sloan. Universitätsmedizin Göttingen, Institut für Molekularbiologie, Göttingen, 37073, Germany.

and the metazoan-specific A' sites have yet to be assigned in the 18S rRNA maturation pathway.

In addition to yUtp24/hUTP24, one other PIN domain protein, yUtp23 (hUTP23 in humans), is essential for early cleavages at A0, A1/1 and A2/2a (4,12,13). However, yeast yUtp23 contains only two of the four conserved amino acids in the catalytic site and the two conserved PIN domain residues are not essential for yUtp23 function (4,14). In addition, the protein contains a conserved CCHC Zinc finger. yUtp23 binds nucleotides 745–859 in the 18S rRNA *in vitro* and specifically associates with the snR30 H/ACA snoRNP *in vivo* (14,15). The snR30 snoRNA is essential for the A0, A1 and A2 cleavages and SSU processome assembly (16,17), and is needed for the integration of yUtp23 into the pre-ribosome (15). Conversely, snR30 does not require yUtp23 for pre-ribosome association, but yUtp23 is essential for snR30 release from the complex (15). The internal loop in the 3′ hairpin of snR30 base-pairs with two elements (rm1 and rm2) in the 18S rRNA expansion sequence 6 (ES6) and is expected to play an important role at this region during pre-SSU maturation (18,19). Interestingly, yUtp23 may be needed to establish base-pairing of snR30 with the 35S pre-rRNA (15). A number of other factors interact with snR30 and/or ES6 including yRok1 (snR30 and ES6 (20)), yRrp7 (ES6 (21)), yUtp24 (ES6 (7)) and yRrp5 (snR30 and ES6 (22)). Release of snR30 from the pre-ribosome requires the RNA helicase yRok1, which directly interacts with yRrp5 (22–24).

Considerably less is known about UTP23 in *Metazoa* but similar to the yeast protein, it is required for pre-rRNA cleavages in the 5′ ETS and ITS1 (at sites A0, 1 and 2a in humans and mice) (12,13) and is also associated with the human homologue of snR30, U17 (15). However, hUTP23 has three of the four key acidic amino acids of the PIN domain, together with the Zinc finger motif, suggesting that it may be an active nuclease in human cells, but why such activity would be required in humans, but not in yeast, remains unclear.

Here, we present a combination of *in vivo* and *in vitro* approaches to determine the role of yUtp23/hUTP23 in both yeast and human ribosome biogenesis. *In vivo* RNA-protein crosslinking studies (CRAC) generated a transcriptome-wide RNA binding profile for yeast yUtp23, which provides new insights into its relationship with snR30 and the ES6 region of the 18S rRNA. We also performed *in vitro* assays using recombinant proteins to determine RNA and protein interaction partners of yeast yUtp23 and human hUTP23 in the SSU processome. Finally, we established RNAi-rescue systems in HEK293 cells to study the effect of mutant hUTP23 on pre-rRNA cleavage in the human system.

MATERIALS AND METHODS

Yeast strains and methods

The *S. cerevisiae* strain expressing genomically encoded, C-terminal HTP-tagged (His_6-TEV-protA) yUtp23 under the control of its endogenous promoter (Supplementary Table S1) was constructed by standard methods. Cultures were grown at 30°C in medium containing 2% glucose and 0.67% nitrogen base.

CRAC and data analysis

Actively growing yeast cultures in SD medium (OD_{600} ~0.5) were UV-irradiated in a 1.2 m metal tube for 100 s at 254 nm to generate RNA–protein crosslinks. The CRAC method (Supplementary Figure S2A) was performed as described in (25,26). Illumina sequencing data were aligned to the yeast genome using Novoalign (http://www.novocraft.com). Sequencing data from this publication were analyzed as previously reported (27) and submitted to the GEO database (http://www.ncbi.nlm.nih.gov/geo/, identifier GSE87238).

Cloning and mutagenesis

The open reading frames of the protein genes listed in Supplementary Tables S2 and S3 were amplified from yeast genomic DNA or human cDNA adding restriction sites and cloned into pET100 vectors (Invitrogen). The constructs were either used directly or sub-cloned into different protein expression vectors to purify affinity-tagged recombinant proteins from *Escherichia coli* using standard techniques or used for *in vitro* translation with [^{35}S] methionine (TNT, Promega).

A human hUTP23 cDNA construct (generated by IDT) containing a C-terminal His_8-PP (PreScission protease recognition site)-2xHA (hemagglutinin) tag was amplified by PCR and cloned into the pcDNA5/FRT/TO vector (Invitrogen) for protein expression under the control of a tetracycline-inducible promoter. The coding sequence of hUTP23 was altered to render it resistant to the siRNAs used to deplete the endogenous mRNA (Supplementary Table S4). Point mutations were generated by site-directed mutagenesis using overlapping primers (Supplementary Table S3) and confirmed by sequencing.

In vitro RNA binding electromobility shift assay (EMSA)

Recombinant proteins were expressed and purified as described previously (28). RNA substrates were transcribed in the presence of [^{32}P]-αUTP, using plasmid constructs or PCR products with T7 promoter sequences as template (Supplementary Tables S2 and S3). RNA binding assays were performed using trace amounts of radiolabeled RNA in 10 mM Tris/HCl pH 7.6, 75 mM NaCl, 2 mM dithiothreitol, 100 ng/μl bovine serum albumin, 0.8 units/μl RNasin, 4.5% glycerol, 0.05% Tween20 and 500 ng/μl *E. coli* tRNA. A total of 10 μl or 15 μl reactions containing 100–5000 nM protein were incubated for 10 min at 30°C, followed by 10 min incubation on ice, and 2 μl or 3 μl native agarose loading dye (30% glycerol and 0.3% Orange G (w/v)) was added. Products were resolved on 4% polyacrylamide/1x TBE (5% glycerol) native gels and visualized using a Typhoon FLA9000 PhosphorImager.

Protein–protein interaction studies

GST-bait proteins were immobilized on glutathione sepharose pre-blocked with bovine serum albumin and incubated with His-tagged recombinant proteins or [^{35}S] *in vitro* translates for 1 h at 4°C in Buffer NB (20 mM Tris/HCl pH 7.6, 150 mM NaCl, 8.7% glycerol and 0.1% Tween20). The beads were washed five times with buffer

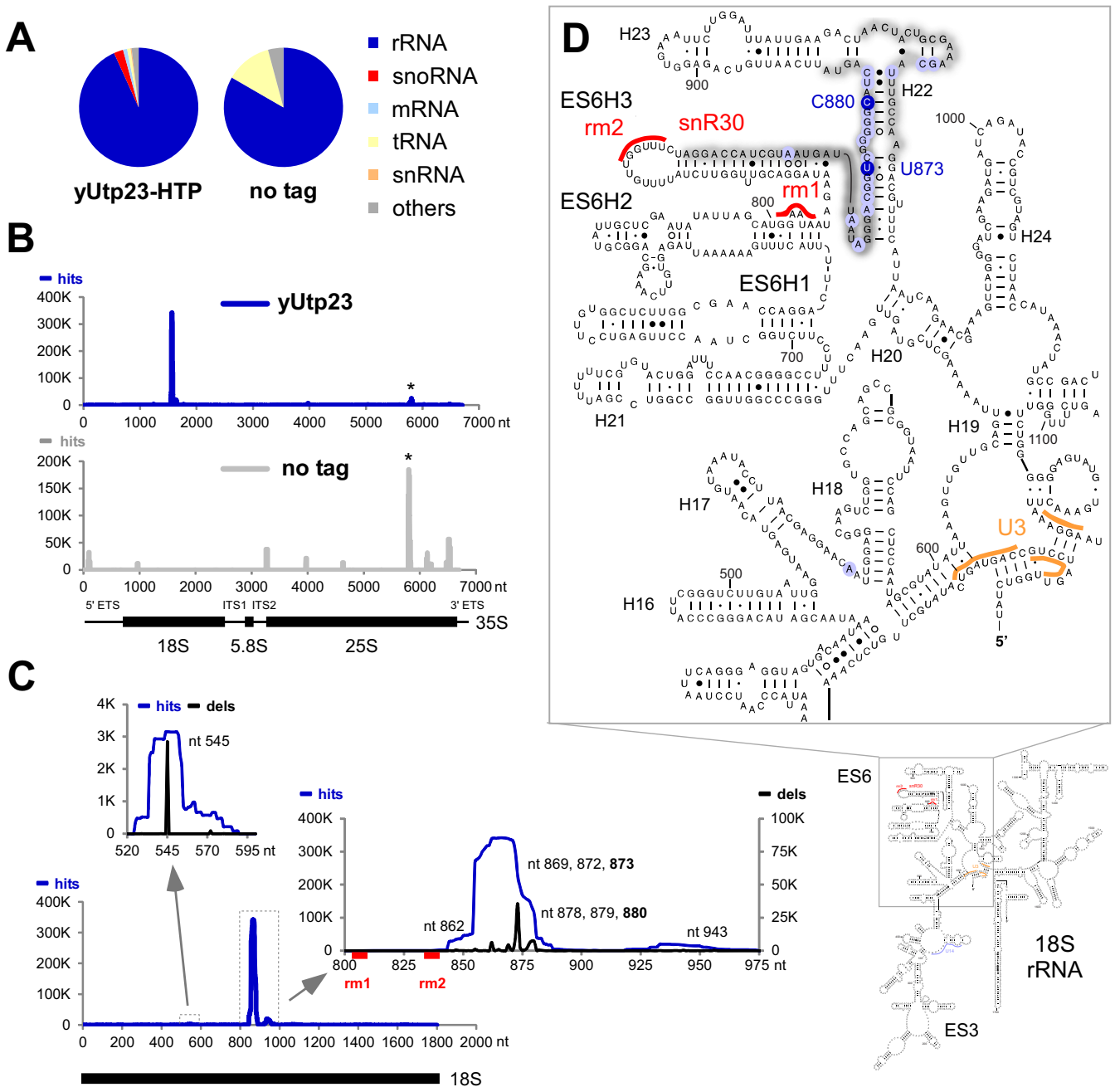


Figure 1. RNA crosslinking sites of yeast yUtp23 on the (pre-) ribosomal RNA. Crosslinked and subsequently trimmed RNAs were purified from yeast strains expressing yUtp23-HTP or a non-tagged control strain and used to generate cDNA libraries. Illumina sequencing data were mapped to the yeast genome using Novoalign. Normalized data for one representative yUtp23-HTP data set and the no tag control are plotted as reads per million mapped sequences ('hits'). (**A**) Transcriptome-wide binding profiles. A total of 2 277 820 mapped reads were recovered for the yUtp23-HTP data set and 7446 reads for the no tag control. Pie charts illustrate the proportion of all reads mapped to functional RNA classes (indicated on the right). (**B**) yUtp23 crosslinking profile on the primary ribosomal RNA transcript. Sequences were aligned with the rDNA (RDN37-1) encoding the 35S pre-ribosomal RNA. The frequency of recovery in the yUtp23 data set (blue) or the no tag control (gray) is plotted as total reads (hits) for each individual nucleotide. The positions of the mature 18S, 5.8S and 25S rRNAs are indicated by thick lines. A common CRAC contaminant at the 3′ end of the 25S rRNA is marked by an asterisk (26). ETS: external transcribed spacer; ITS: internal transcribed spacer. (**C**) yUtp23 crosslinks on the 18S rRNA. Hits (blue): total reads; deletions (black; dels): mutations and microdeletions representing precise binding sites. Prominent microdeletion peaks around nt 545 (left insert) and in the 18S ES6 region (right insert) are labeled. The position of the mature 18S rRNA and the snR30 binding sites in the 18S ES6 region (rm1 and rm2) are indicated by a thick black line or red boxes, respectively. (**D**) Predicted secondary structure of the mature 18S rRNA in *S. cerevisiae*. yUtp23 crosslinking sites are marked on the sequence and gray shades indicate peak height. Microdeletion peaks (see panel C) are highlighted by shaded blue circles. Binding sites for the snoRNAs snR30 in the expansion sequence 6 (ES6) region (rm1 and rm2, red) and U3 around the central pseudoknot (orange) are also indicated.

NB to remove unbound material. Retained proteins were separated by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by Coomassie staining, immunoblotting or by using a PhosphorImager, respectively.

Cell culture and RNAi

hUTP23-His_8-PP-2xHA-pcDNA5 constructs or the empty pcDNA5 plasmid were transfected into Flp-In T-Rex HEK293 cells as described by the manufacturer (Invitrogen) and stably transfected cells were cultured according to standard protocols. Expression of exogenous proteins was induced by addition of tetracycline (WT: 0.1–1 ng/μl; D31N: 50–100 ng/μl; C103A: 100–200 ng/μl). For RNAi-mediated depletion of endogenous proteins, cells were transfected with siRNA duplexes (Supplementary Table S4) using Lipofectamine RNAiMAX transfection reagent (Invitrogen) and harvested after 72 h.

RNA analysis

TRI reagent (Sigma-Aldrich) was used to extract RNA from HEK293 cell pellets. For northern blot analysis, 2 μg of total RNA was separated on 1.2% glyoxal–agarose gels, transferred to nylon membrane and hybridized with [^{32}P] 5′ end radiolabeled oligonucleotides (Supplementary Table S3). Results were visualized with a Typhoon FLA9000 PhosphorImager and quantified using the ImageQuant software (GE Healthcare).

Immunofluorescence

HEK293 cells expressing HA-tagged hUTP23 proteins were grown on coverslips and induced with tetracycline for 72 h before being fixed with phosphate buffered saline (PBS) containing 4% paraformaldehyde. Immunofluorescence analysis was performed as described in (5). Briefly, cells were permeabilized using 0.2% Triton (v/v) and then incubated with primary and secondary antibodies (Supplementary Table S5) diluted in 10% fetal calf serum, PBS, 0.1% Triton (v/v) and washed with PBS, 0.1% Triton (v/v). In the final step, cells were washed with PBS containing DAPI (4′,6′-diamidino-2-phenylindole, 1:10,000) and mounted onto a slide using Mowiol. Images were obtained using a Zeiss Axiovert 200 microscope with Plan-Apochromat x100 1.4NA objective, Axiovision software and an Axiocam monochrome camera and processed in Photoshop (Adobe).

Immunoprecipitation

Immunoprecipitation experiments were performed with sonicated whole-cell extracts as previously described (29), with the exception that anti-HA antibody-coupled agarose beads were used. Co-precipitated RNA was extracted, separated by denaturing polyacrylamide electrophoresis and analyzed by northern blotting using [^{32}P] 5′ end radiolabeled oligonucleotides (Supplementary Table S3).

RESULTS

Yeast yUtp23 crosslinks to the 18S ES6 region and to the snR30 snoRNA

In yeast, yUtp23 is essential for 18S rRNA maturation (4) and required for dissociation of the snR30 small nucleolar RNP from pre-ribosomal particles (15), but its exact role within the SSU processome is unclear. We therefore applied *in vivo* RNA-protein crosslinking (26) to determine yUtp23 RNA binding sites (see Supplementary Figure S2A for overview of the CRAC procedure). We first constructed a yeast strain expressing genomically encoded, C-terminal HTP-tagged (His_6-TEV-protA) yUtp23 protein under the control of its endogenous promoter. The affinity-tagged yUtp23 protein supported wild-type growth and actively growing yeast cultures were UV-irradiated as described (25). Purification of yUtp23 proteins was verified by western analysis (Supplementary Figure S2B) and co-purified crosslinked RNA fragments were isolated and analyzed as outlined in Supplementary Figure S2A. We reproducibly detected yUtp23 RNA crosslinking sites in four independent CRAC experiments. Results from the two largest data sets (>2 million mapped reads each) are presented in Figures 1 and 2 and Supplementary Figure S4.

Figure 1A shows the transcriptome-wide RNA binding profile of yUtp23 compared to a non-tagged control strain. The majority of reads (>93%) in all four yUtp23 data sets were mapped to rRNA sequences, with lower numbers of hits in snoRNAs (on average ∼2%). Importantly, no snoRNAs hits were detected in the non-tagged control data set. Within the primary 35S pre-rRNA transcript, yUtp23 crosslinks were predominately found in the 18S rRNA sequence (Figure 1B). This crosslinking profile is in agreement with the known function of yUtp23 in 18S rRNA maturation and was not seen with the non-tagged control strain.

The main crosslinking peaks in the 18S rRNA sequence (Figure 1C) are located between 18S nt 844–887 and 18S nt 929–953 of the eukaryotic ES6 or directly downstream of this region (helix 22 and 23, respectively) with some reads also found between 18S nt 534–554 (helix 17 and helix 18). To validate the main crosslinking site in the ES6 region, we performed *in vitro* RNA binding studies with recombinant GST-tagged yUtp23 protein expressed in *E. coli* and a radiolabeled 18S rRNA fragment encompassing the crosslinking sites (18S nt 775–963) (Figure 1D and Supplementary Figures S2C and S2D). Consistent with previously published data (14), the yUtp23 protein (but not the GST-only control) bound to this RNA fragment, albeit with very low efficiency. However, the formation of the RNP complex was specifically decreased by addition of a cold competitor RNA of the same sequence, but not an unrelated RNA fragment of similar size (18S nt 1022–1146). This result indicates specific binding of yUtp23 to this region of the rRNA.

During preparation of the cDNA libraries used for sequencing, microdeletions and/or mutations are often introduced at the site of crosslinking and can therefore be exploited to determine precise protein contact sites (30). Major sites of microdeletion within the main 18S rRNA peaks are highlighted in the inserts shown in Figure 1C. Positioning of the predominant yUtp23 reads and microdele-

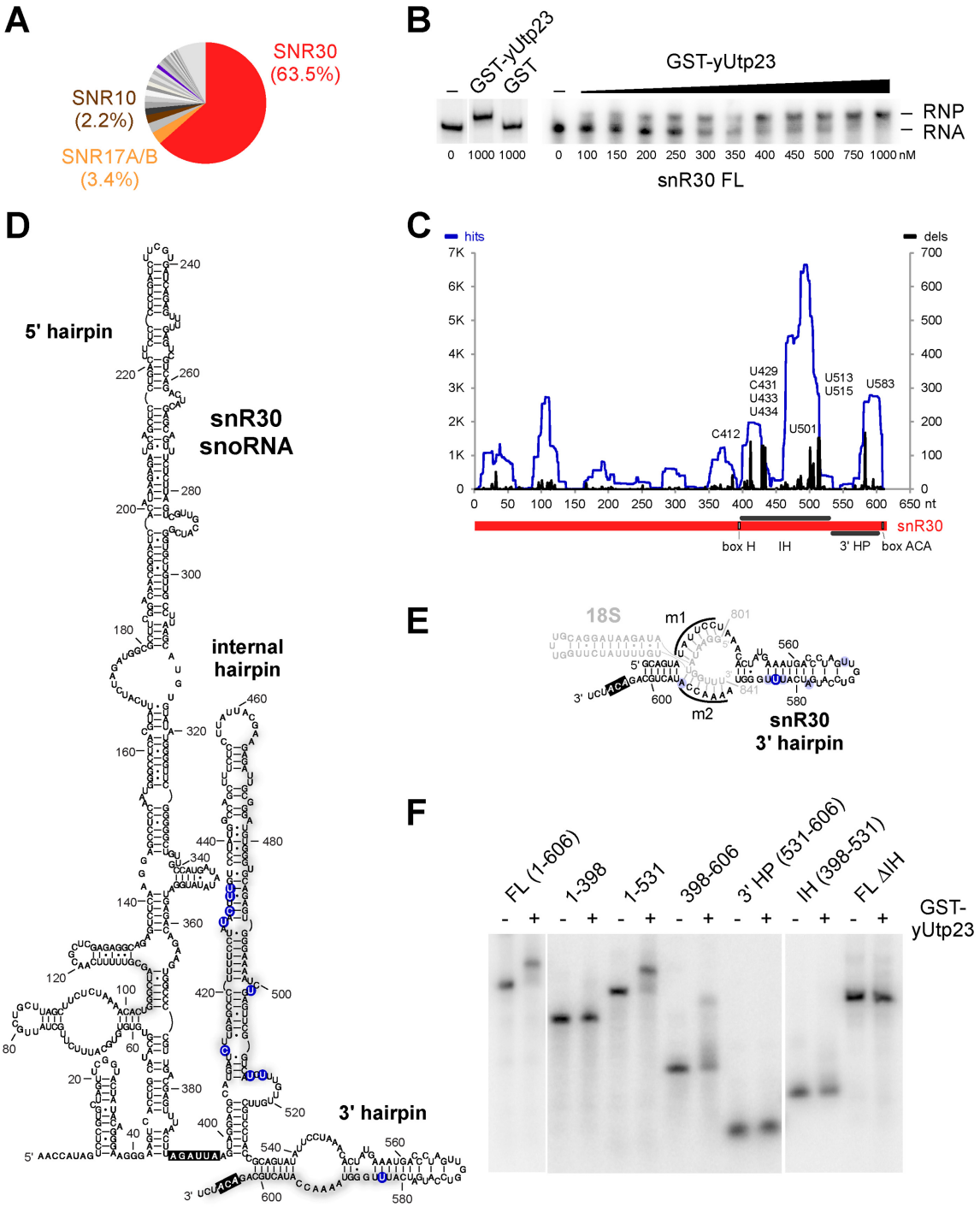


Figure 2. Yeast yUtp23 contacts the internal and 3′ terminal hairpins of the snR30 snoRNA. (**A**) Pie chart of snoRNA reads from yUtp23 crosslinking data sets. The top 20 hits from two combined yUtp23 data sets (see Supplementary Figure S4B for more detail) are represented as percentages of all reads mapped to snoRNAs. Significant and reproducibly enriched RNAs with essential, non-modifying roles in 18S maturation (snR30, U3 and snR10) are highlighted. (**B**) Electromobility shift assay (EMSA) showing the binding of GST-yUtp23 or GST (1000 nM, left) or GST-yUtp23 (100–1000 nM, right) to *in vitro* transcribed radiolabeled snR30 RNA. Free RNA and RNP complexes were separated on 4% native polyacrylamide gels and visualized using a PhoshorImager. (**C**) yUtp23 crosslinking profile on the snR30 snoRNA. Hits (blue): total reads; deletions (black; dels): mutations and microdeletions representing precise binding sites. Prominent microdeletion peaks in the internal hairpin (IH) and 3′ terminal hairpin (3′ HP) are labeled. The positions of the H and ACA boxes within the snR30 RNA are indicated. (**D**) Predicted secondary structure of the snR30 snoRNA in *S. cerevisiae* (adapted from (18)). yUtp23 crosslinking sites are marked on the sequence and grey shades indicate peak height. Predominant microdeletion peaks (see panel C) are highlighted by shaded blue circles. The H and ACA boxes are shown in black. (**E**) A model of the interaction between the 18S ES6 binding sites (m1 and m2) in the 3′ terminal hairpin (3′ HP) of snR30 (black) and the 18S ES6 region (nt 801–841, grey). yUtp23 reads (grey) and microdeletion sites in the 3′ HP (light and dark blue circles) are indicated. Shades of blue represent peak height (see panel C). The ACA box is highlighted in black. (**F**) EMSA showing the binding of GST-yUtp23 (1300 nM) to full-length (FL) snR30 or snR30 RNA fragments. Free RNA and RNP complexes were analyzed as in panel B. IH: internal hairpin; FL ΔIH: full-length snR30 lacking the internal hairpin.

tion sites on the predicted structure of the mature 18S rRNA (Figure 1D) revealed that the main yUtp23 binding sites are in close proximity, but not overlapping with the known snR30 binding sites rm1 (18S nt 801–806) and rm2 (18S nt 836–842) in the ES6 region. Direct yUtp23 protein–RNA contacts within this region were also detected by primer extension analysis on non-digested RNA that was crosslinked to and co-purified with yUtp23 (Supplementary Figure S3A). Notably, similar protein–RNA contacts were also observed between hUTP23 and the human ES6 region when HEK293 cells, expressing an affinity-tagged version of hUTP23, were used for crosslinking and primer extension analysis (Supplementary Figure S3B). This indicates that the interaction between yUtp23/hUTP23 and the 18S ES6 region is evolutionarily conserved.

We next examined the yUtp23 crosslinking sites on snoRNAs (Figure 2 and Supplementary Figure S4). For this, the data for the 20 snoRNAs, to which the maximum number of reads mapped in two independent yUtp23 data sets, were combined and are represented as percentages of reads mapped to all yeast snoRNAs (Figure 2A). Remarkably, 63.5% of snoRNA hits mapped to a single box H/ACA snoRNA, snR30. Significant proportions of reads were also reproducibly mapped to two of the three other snoRNAs involved in 18S rRNA processing (Supplementary Figure S4A), the box C/D snoRNA U3/snR17A/B (3.4%) and the H/ACA snoRNA snR10 (2.2%), whereas reads corresponding to the third, the box C/D snoRNA U14/snR128, were only recovered in one of the analyzed data sets (Supplementary Figure S4B). Hits were also seen for the box C/D snoRNAs snR190 (2.7%) and other modification snoRNAs (0.4–1.8%). The substantial enrichment of snR30 over other snoRNAs suggests that the previously observed relationship between yUtp23 and snR30 involves a direct protein–RNA interaction. Consistent with this, recombinant GST-tagged yUtp23 protein specifically bound to *in vitro* transcribed snR30 RNA, with an estimated dissociation constant (K_D) value of ∼300 nM (Figure 2B). yUtp23 crosslinking hits and microdeletions could also be mapped to the U3/snR17A/B, snR10 and U14/snR128 snoRNA sequences suggesting direct protein–RNA contacts *in vivo* (Supplementary Figure S4C). However, stable interactions with these RNAs were not detected *in vitro* (data not shown) and might therefore require the presence of the pre-rRNA or other factors within the pre-ribosomal particle.

The positions of yUtp23 crosslinking reads and microdeletion sites on the snR30 sequence (Figure 2, panels C and D) suggest that yUtp23 mainly contacts the 3′ portion of the snoRNA. Most reads and deletions mapped to the internal hairpin (IH, nt 398–531) and the 3′ terminal hairpin (3′ HP, nt 532–601) of snR30, whereas fewer reads were found in the 5′ hairpin. The high number of yUtp23 crosslinking sites mapping to the internal hairpin was surprising given that this snR30 region was shown to be dispensable for growth *in vivo* (18). Interestingly, the majority of the precise crosslinking sites on the snR30 3′ HP did not overlap with the snR30 sequences (m1 and m2) that can engage in base-pairing interactions with the 18S rRNA (Figure 2E). Instead, most microdeletions were located within the distal region of the 3′ HP, which was previously proposed to be a putative snR30-specific snoRNP protein binding site (19).

To investigate the importance of the snR30 IH and 3′ HP with respect to yUtp23 association, we designed a series of snR30 fragments and tested them for binding to recombinant yUtp23 proteins *in vitro* (Figure 2F). yUtp23 efficiently bound the full-length (FL) snR30 (nt 1–606). In agreement with the CRAC data suggesting significant contacts to the internal hairpin (IH), binding of yUtp23 was severely reduced in the absence of the IH (FL ΔIH). yUtp23 exhibited no binding to the snR30 5′ hairpin alone (nt 1–398), while binding was restored when the IH (nt 1–531) was included. Moreover, neither the IH (nt 398–531), nor the 3′ HP (nt 531–606) alone were sufficient for binding, whereas combination of both (nt 398–606) enabled binding, but at a reduced level. The data suggest that the internal hairpin and 3′ hairpin in snR30 are both needed for yUtp23 binding *in vitro* and that elements in all three hairpins are necessary for maximally efficient binding.

We conclude that yUtp23 primarily binds to the 3′ portion of the snR30 snoRNA and adjacent to the snR30 base-pairing sites in the mature 18S rRNA sequence. These findings support a direct role for yUtp23 in snR30 function and release.

Yeast yUtp23 and human hUTP23 directly bind proteins that interact with the 18S rRNA ES6 region and snR30/U17

In recent years, several *in vivo* RNA-protein crosslinking studies in yeast have highlighted the 18S ES6 region as a binding platform for early acting SSU synthesis factors, which contact the 18S rRNA in close proximity to the described yUtp23 crosslinking sites (Figure 3A). In the case of the PIN domain endonuclease yUtp24 (7) and the RNA helicase yDhr1 (31), binding to the ES6 region only represented a secondary pre-rRNA contact, while their main crosslinking sites were located within the pseudoknot region. Consistent with this, snoRNA crosslinks for both proteins were mainly mapped to U3, which forms extensive base-pairing interactions with the central pseudoknot. Two other factors, the RNA helicase yRok1 (20) and the RNA-binding protein yRrp7 (21), were almost exclusively crosslinked to the ES6 region. While both proteins are functionally linked to snR30, yRok1 was shown to crosslink to the SSU-associated snoRNAs snR30, U3 and U14, and, to a lesser extent, to snR10, whereas yRrp7 exhibited a strong association with snR10. Lastly, the CTD domain of the pre-ribosomal 'compaction factor' yRrp5 also crosslinked to the ES6 region and analysis of full-length yRrp5 revealed hits to all SSU-associated snoRNAs (U3, snR30, snR10 and U14) (22). Many of these SSU synthesis factors are conserved in the human system and are therefore predicted to engage in similar interactions at the human ES6 region or with the U17 snoRNA, the human snR30 counterpart, but their exact contact sites or interaction partners are not known.

To better understand the role of yUtp23/hUTP23 within this 'binding platform' for ribosome biogenesis factors in yeast and humans, we established an *in vitro* system using recombinant proteins (Supplementary Figures S2 and S5) and [^{35}S] *in vitro* translated proteins to analyze protein–protein

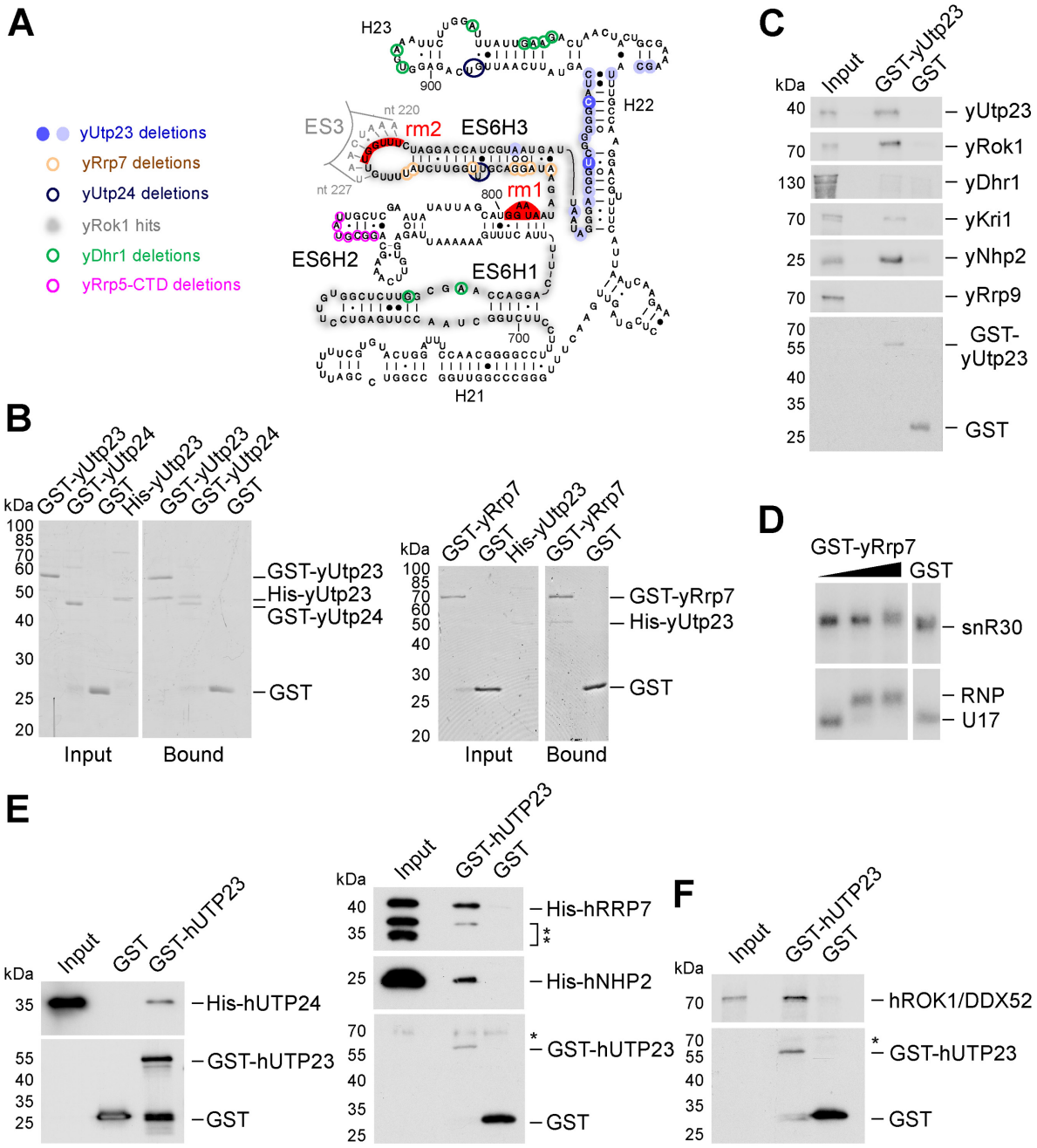


Figure 3. Yeast yUtp23 and human hUTP23 interact with early pre-40S factors contacting the 18S rRNA ES6 region. (**A**) Predicted secondary structure of the mature yeast 18S ES6 region with binding sites of crosslinked ribosome biogenesis factors. Precise crosslinking sites (microdeletions) of yUtp23 (shaded blue circles indicating peak height, see Figure 1C and D) and other published factors (yUtp24 (7), dark blue; yRrp7 (21), brown; yDhr1 (31), green and the yRrp5-CTD (22), pink) are highlighted by circles. yRok1 crosslinking regions (total reads, (20) are shaded in grey. Binding sites for snR30 (rm1 and rm2) are indicated in red. (**B**) Recombinant GST-tagged yeast yUtp23, yUtp24, yRrp7 or free GST were immobilized on glutathione sepharose and incubated with protein-A-His-tagged yUtp23. Bound material was eluted under denaturing conditions, separated by SDS-PAGE and visualized by Coomassie staining. A total of 10% of the input material was loaded. (**C**) Immobilized GST-tagged yeast yUtp23 or free GST was incubated with proteins generated by *in vitro* translation in the presence of [^{35}S] methionine. Bound material was treated as in panel B and analyzed by a PhosphorImager. 2% of the input material was loaded. **D** EMSA showing the binding of GST-yRrp7 (170, 1700 and 3300 nM) or GST (3300 nM) to *in vitro* transcribed radiolabeled snR30 (top) or human U17 (bottom) RNA. RNA and RNP complexes were analyzed as in Figure 2B. (**E**) Recombinant GST-tagged human hUTP23 or free GST was immobilized on glutathione sepharose and incubated with N-terminally His-tagged hUTP24, hRRP7 or hNHP2. Bound material was eluted under denaturing conditions, separated by SDS-PAGE and transferred to nitrocellulose membrane. Proteins were analyzed by immunoblotting using antibodies specific for the His-tag (upper panels) or the GST-tag (lower panel), respectively. Double asterisks: C-terminally truncated forms of the His-tagged hRRP7 protein. The single asterisk denotes a non-specific protein recognized by the anti-GST antibody (lower panel). Left panel (hUTP24): 10% of the input material was loaded. Right panel (hRRP7, hNHP2): 5% of the input material was loaded. (**F**) Immobilized GST-tagged human hUTP23 or free GST was incubated with hROK1/DDX52 protein generated by *in vitro* translation in the presence of [^{35}S] methionine. Bound material was treated as in panel C and analyzed by a PhosphorImager. One percent of the input material was loaded. Asterisk: non-specific protein recognized by the anti-GST antibody.

interactions between ES6-interacting factors (Figure 3, panels B, C, E and F). In these experiments, yeast or human recombinant GST-tagged proteins or free GST were immobilized on glutathione sepharose, incubated with His-tagged recombinant or *in vitro* translated proteins and washed repeatedly to remove non-specifically bound factors. Retained material was eluted under denaturing conditions, separated by SDS-PAGE and visualized by Coomassie staining (panel B), PhosphorImaging (panels C and F) or immunoblotting (using anti-His or anti GST tag antibodies, panel E), respectively.

We firstly found that yeast yUtp23 interacted with itself (panels B and C), suggesting that, similar to other PIN domain proteins (32), it might act as a multimer. Interestingly, we also detected strong interactions between yUtp23 and the PIN domain endonuclease yUtp24 (panel B), which has recently been reported to be directly responsible for the A1 and A2 site cleavages (6,7). This raises the exciting possibility that yUtp23 might act as a recruitment factor or, alternatively, a chaperone for yUtp24, to modulate yUtp24 catalytic activity until it is correctly positioned within the SSU processome. We did not, however, observe interactions between yUtp23 and yDhr1 (panel C) or yRrp5 (not shown).

yUtp23 also specifically interacted with yRrp7 (panel B) and yRok1 (panel C). In yeast, snR30 is required for the stable association of yRrp7 to pre-ribosomes but not vice versa (21). We therefore speculate that yRrp7 contacts the pre-rRNA either during snR30 base-pairing or after snR30 action (see discussion). Consistent with the hypothesis that yRrp7 might associate with the 18S sequence in the absence of snR30, the top snoRNA in the yRrp7 CRAC data set was indeed snR10, and not snR30 (21), and we were unable to observe a direct interaction between recombinant GST-yRrp7 and *in vitro* transcribed snR30 (Figure 3D, upper panel), or, in fact, snR10 (data not shown). Curiously, GST-yRrp7 strongly interacted with U17 (Figure 3D, lower panel), but yRrp7 had also previously been reported to interact with another, functionally unrelated box H/ACA snoRNA, snR5, *in vitro* (21). It is therefore unclear if the observed yRrp7/U17 interaction is biologically relevant. Given that the U17 and snR5 snoRNAs both lack an extended internal hairpin (Supplementary Figure S4D and data not shown), it is also possible that the snR30-specific internal hairpin might interfere with yRrp7 binding to snR30 in our *in vitro* assay.

We further assessed whether yUtp23 directly associates with other known snR30 binding factors. We detected a reproducible interaction between yUtp23 and the nucleolar protein yKri1 (Figure 3, panel C), which, like yUtp23, requires snR30 for pre-ribosomal recruitment (15). Furthermore, we also found that yUtp23 directly interacted with yNhp2, one of the four core box H/ACA proteins, while no interaction was detected with the U3-specific yRrp9 protein. It is likely that the observed protein–protein interaction between yUtp23 and yNhp2 contributes to yUtp23 binding to the snR30 snoRNP. This might explain why the internal hairpin in snR30, which is important for binding of yUtp23 to snR30 *in vitro* (Figure 2F), is not essential for survival *in vivo* (18).

Finally, we also tested human homologues of key ES6-binding factors (hUTP24, hROK1/DDX52 and hRRP7) as well as the core box H/ACA protein hNHP2, for their ability to interact with recombinant GST-tagged human hUTP23 (panels E and F, respectively). Importantly, positive interactions detected in yeast (panels B and C) were also detected in the human system. Human hUTP23 interacted with hUTP24, hRRP7, hNHP2 and hROK1/DDX52. This indicates that key protein–protein interactions at the 18S ES6 rRNA region are evolutionarily conserved.

In conclusion, the observed direct interactions between yUtp23/hUTP23 and factors that also bind to the ES6 region or the snR30/U17 snoRNAs, in combination with the positioning of their crosslink sites on the individual RNAs, provide important insights into their spatial and temporal association within the yeast and human SSU processome (see discussion).

An intact PIN domain and Zinc finger in human hUTP23 are both required for 18S rRNA maturation

In budding yeast, yUtp23 contains only two (D31 and D123) of the four possible acidic residues that make up the characteristic PIN domain endonuclease catalytic center (Figure 4A and Supplementary Figure S6), and mutational analyses of these two residues strongly suggested a non-enzymatic role for yUtp23 within the yeast SSU processome (4,14). The PIN domain of human hUTP23, on the other hand, contains conserved acidic residues at three positions (D31, E68 and D122) and it was shown for the PIN domain of the NMD endonuclease SMG6, that a triad of acidic residues is sufficient for catalytic activity (33). Both human and yeast Utp23 proteins also share a conserved CCHC Zinc finger motif with a predicted role in nucleic acid binding. Single point mutations in the Zinc finger of yeast yUtp23 are lethal (14), but the essential role of this motif has not yet been determined.

RNAi-mediated knockdown of human hUTP23 results in 30S pre-rRNA accumulation, indicating that the presence of hUTP23 is required for the A0, 1 and 2a pre-rRNA cleavages (12) (Figure 4B). The PIN endonuclease hUTP24 is responsible for cleavages at sites 1 and 2a (6,7), but the identity of the enzyme that cleaves at site A0 remains unknown. To investigate a potential catalytic function of human hUTP23 in site A0 cleavage, and to understand the role of its conserved Zinc finger motif, we established an RNAi-rescue system for hUTP23 in HEK293 cells. For this, cells were stably transfected with an empty control vector (pcDNA5) or plasmids encoding C-terminally 2xHA-tagged hUTP23 carrying silent mutations in the hUTP23 open reading frame rendering it resistant to RNAi-knockdown when the endogenous hUTP23 mRNA is targeted. Cells expressing hUTP23 WT or mutant hUTP23 proteins with a single catalytic site mutation in the PIN domain (D31N) or a point mutation in the Zinc finger motif (C103A) were compared. Using a titratable *TET* promoter, conditions were established to express WT and mutant proteins at equivalent levels (Figure 4C). WT and mutant proteins showed some variation in their response to induction by tetracycline and it was therefore difficult to titrate the expression of all HA-tagged proteins to a level equivalent to that of the endogenous protein. To address this, two separate WT samples are presented in Figure 4C,

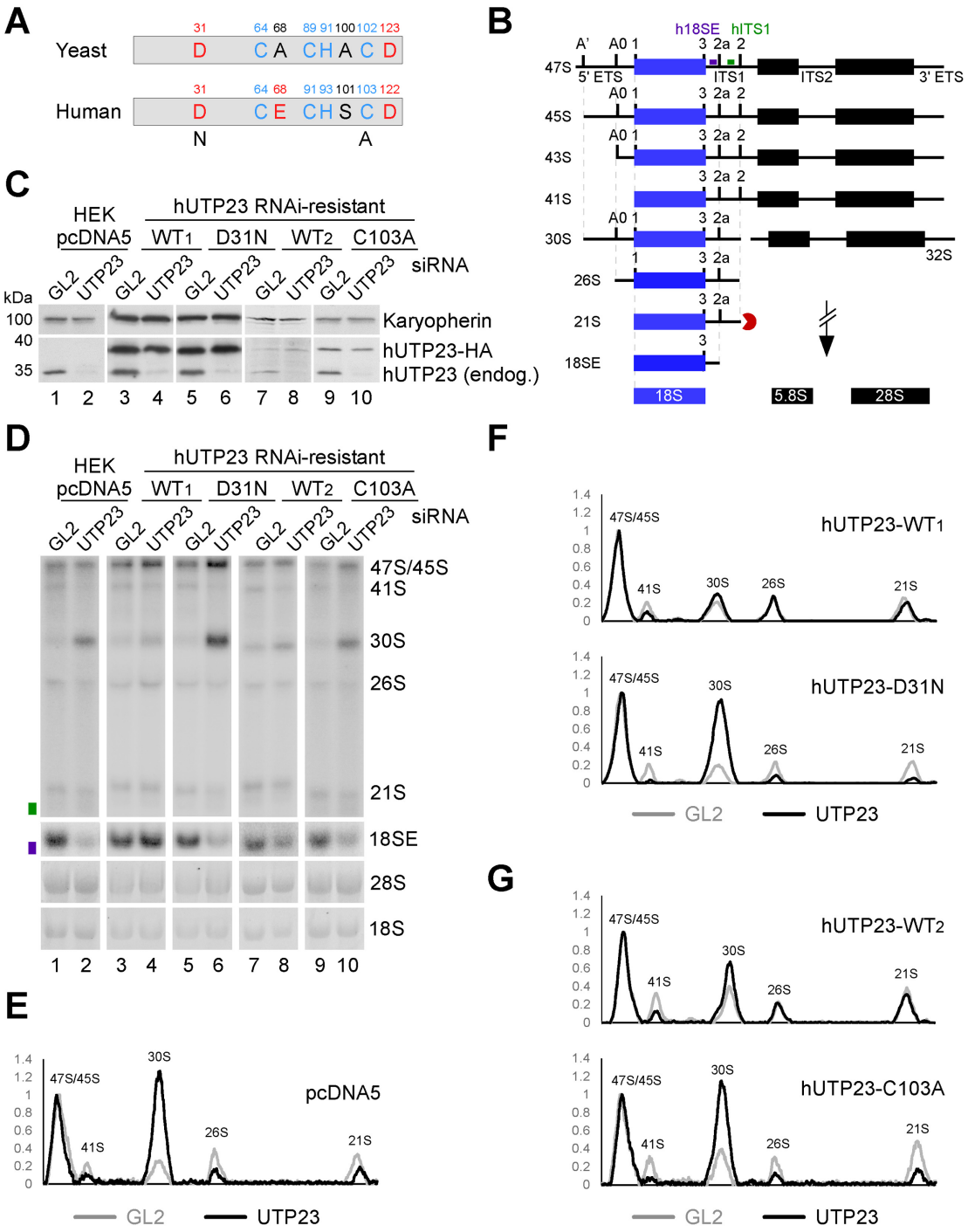


Figure 4. Intact PIN domain and Zinc finger motifs in human hUTP23 are essential for 18S maturation. (**A**) Cartoon depiction of the PIN domain and CCHC Zinc finger motifs in yeast yUtp23 and human hUTP23. Acidic and non-acidic residues in the proposed catalytic centre of the PIN domain are marked in red and black, respectively. Blue: conserved Zinc finger residues. Point mutations generated in the PIN domain (D31N) or the Zinc finger (C103A) of hUTP23 are indicated. (**B**) Schematic representation of the key steps in human ribosome biogenesis. Radiolabeled probes used for northern blotting (h18SE; purple and hITS1; green) are marked above the 47S precursor. ETS: external transcribed spacer; ITS: internal transcribed spacer. (**C**) HEK293 cells were stably transfected with a control plasmid (pcDNA5) or constructs encoding wild type (WT) or mutant forms of HA-tagged hUTP23 (D31N, C103A) and treated with tetracycline to induce protein expression. The hUTP23 coding sequence was modified to render the expressed mRNA resistant to the hUTP23 siRNA. Protein extracted from control cells (GL2), or those depleted of endogenous hUTP23 (UTP23) by RNAi, was separated by SDS-PAGE and transferred to nitrocellulose membrane. Protein levels were analyzed by immunoblotting using antibodies specific for hUTP23 (lower panel) or Karyopherin (upper panel) as loading control. (**D**) RNA from stably transfected and RNAi-treated HEK293 cells as shown in panel C was analyzed by northern blotting using probes hybridizing to the 5′ end of ITS1 (h18SE, purple rectangle) or downstream of 2a (hITS1, green rectangle). Pre-rRNAs were detected using a PhosphorImager and total rRNA (28S/18S) was visualized by ethidium bromide (EtBr) staining. RNA species are labeled on the right. (**E–G**) RNA levels from panel D were normalized to the 47S/45S pre-rRNAs and plotted for each GL2 (gray) or hUTP23 (black) knockdown. The identity of each peak is indicated.

with expression levels that are directly comparable to either the D31N (WT1, compare lanes 3–6) or the C103A (WT2, compare lanes 7–10) mutant, respectively. Cells were transfected with either a siRNA specifically targeting endogenous hUTP23 or a control siRNA targeting firefly luciferase (GL2) (34). After siRNA treatment for 72 h, the expression of endogenous and HA-tagged hUTP23 proteins was analyzed by immunoblotting. The hUTP23-specific siRNA significantly reduced endogenous protein levels, whereas the RNAi-resistant HA-tagged hUTP23 proteins were unaffected (Figure 4C).

Total RNA was extracted from RNAi-treated cells and pre-rRNA processing was analyzed by Northern hybridization using probes complementary to the 5′ end of ITS1 ('h18SE') or downstream of the 2a cleavage site ('hITS1') (Figure 4, panels B and D–G). Depletion of hUTP23 resulted in a significant accumulation of the 30S pre-rRNA, indicative of strongly reduced early cleavages at the A0, 1 and 2a sites, and consequently, a substantial reduction in 18SE levels as previously reported (Figure 4, panels D and E) (12). A 5′-extended 30S precursor (referred to as 30SL5' (12) or 34S (35)), did not accumulate, suggesting that cleavage at A' was not affected (Figure 4D, compare lanes 1 and 2). In cells expressing moderately high levels of wild type HA-tagged hUTP23 (WT1), knockdown of endogenous hUTP23 had no effect on either 18SE or 30S pre-rRNA levels (compare lanes 3 and 4, see panel F for quantification), demonstrating that HA-tagged hUTP23 can functionally replace the endogenous protein in our system. While expression of lower levels of the WT protein (WT2; note that levels are slightly lower than the endogenous protein) did not rescue the processing phenotype to the same extent, a significant reduction of the 30S levels compared to the hUTP23 depletion phenotype was still observed (compare lanes 8 and 2, see panels G and E for quantification).

Surprisingly, expression of either the PIN (D31N, Figure 4D, lane 6) or the Zinc finger (C103A, lane 10) hUTP23 mutants, after hUTP23 knockdown, resulted in strong 30S accumulation compared to their respective WT control (see panels F and G for quantification), a phenotype identical to hUTP23 depletion in the absence of exogenous protein (lane 2). This indicates that the putative active site in the PIN domain of hUTP23 and the integrity of the conserved Zinc finger are both important for hUTP23 function in pre-rRNA processing.

We next investigated the reason(s) why the PIN and Zinc finger mutations impeded hUTP23 function. The significant accumulation of the 30S pre-rRNA caused by expression of the hUTP23 D31N PIN mutant *in vivo* (Figure 4D, lane 6) would be readily explained by a loss of enzymatic activity at site A0. Since site A0, 1 and 2a cleavages are linked, a defect in A0 cleavage would also lead to defects in the cleavages at sites 1 and 2a, which is fitting with the substantial increase in 30S pre-rRNA levels seen upon expression of the hUTP23 D31N PIN mutant. We therefore tested whether recombinant hUTP23 (see Supplementary Figure S5A) would cleave an RNA substrate containing human site A0. However, attempts to demonstrate nuclease activity were so far unsuccessful (data not shown and see discussion).

Immunofluorescence experiments using cells expressing HA-tagged proteins demonstrated that, in contrast to fibrillarin, which is exclusively found in the dense fibrillar component (29), wild-type hUTP23 is localized throughout the nucleolus (Figure 5A). Similar to the wild type protein, the mutant hUTP23 proteins were also found throughout the nucleolus. However, immunoprecipitation experiments revealed that only the WT hUTP23 protein was stably associated with the human homologue of snR30, the U17 snoRNA (Figure 5B, lane 4). The PIN mutant exhibited considerably weaker, but still detectable association with the U17 snoRNA (see lane 6), whereas the C103A Zinc finger mutant did not co-precipitate the U17 snoRNA above background (lane 8). Consistent with our yeast data (Figure 2), recombinant WT hUTP23 directly bound to *in vitro* transcribed U17 snoRNA (Figure 5C and 5D, upper panels). The D31N PIN mutant associated with U17 with similar affinity (Figure 5C, lower panel), whereas the C103A Zinc finger mutant exhibited significantly reduced binding (Figure 5D, lower panel). Interestingly, the same result was observed when we analyzed the binding of WT recombinant yeast yUtp23 and a C102A Zinc finger mutant (Supplementary Figure S5B) to either yeast snR30 or human U17 (Figure 5E). The WT yUtp23 protein efficiently bound both U17 and snR30, but in contrast, the C102A Zinc finger mutant showed significantly reduced binding to both snoRNAs at low protein concentration. In the presence of high levels of the C102A mutant, about half of the U17 substrate was still bound compared to the wild type protein, whereas resolvable complexes of bound snR30 or free RNA could not be detected. This result suggests non-stoichiometric and likely non-specific protein-snR30 interactions when the mutant protein is present at high concentration in the assay. The *in vitro* RNA binding results indicate that the severe growth defect in yeast (14) and the pre-rRNA processing defect in humans (Figure 4), seen upon expression of the yUtp23/hUTP23 Zinc finger mutants, are likely due to a defect in yUtp23/hUTP23 binding to the snR30 and U17 snoRNAs, respectively.

Taken together, our data indicate that an intact PIN domain in hUTP23 and efficient binding of hUTP23 to the U17 snoRNP are both required for 18S maturation in humans.

DISCUSSION

Here, we present data suggesting that the yeast yUtp23-snR30 and human hUTP23-U17 complexes may act as hubs to coordinate the binding and release of factors at the 18S ES6 region. We also show that the PIN domain active site amino acid, D31, is needed for pre-rRNA processing suggesting that hUTP23 may actually be an endonuclease.

Using *in vivo* crosslinking in actively growing yeast cells, we have identified snR30 and the 18S ES6 region as *bona fide* RNA-binding targets for yUtp23 (Figures 1 and 2, and Supplementary Figures S2, S3 and S4). Direct protein-RNA interactions were validated *in vitro* using recombinant proteins and *in vitro* transcribed RNA and we also showed that stable interactions with snR30/U17 required an intact Zinc finger in yUtp23/hUTP23 (Figure 5). The top of the 3′ hairpin in snR30 was predicted to bind a pro-

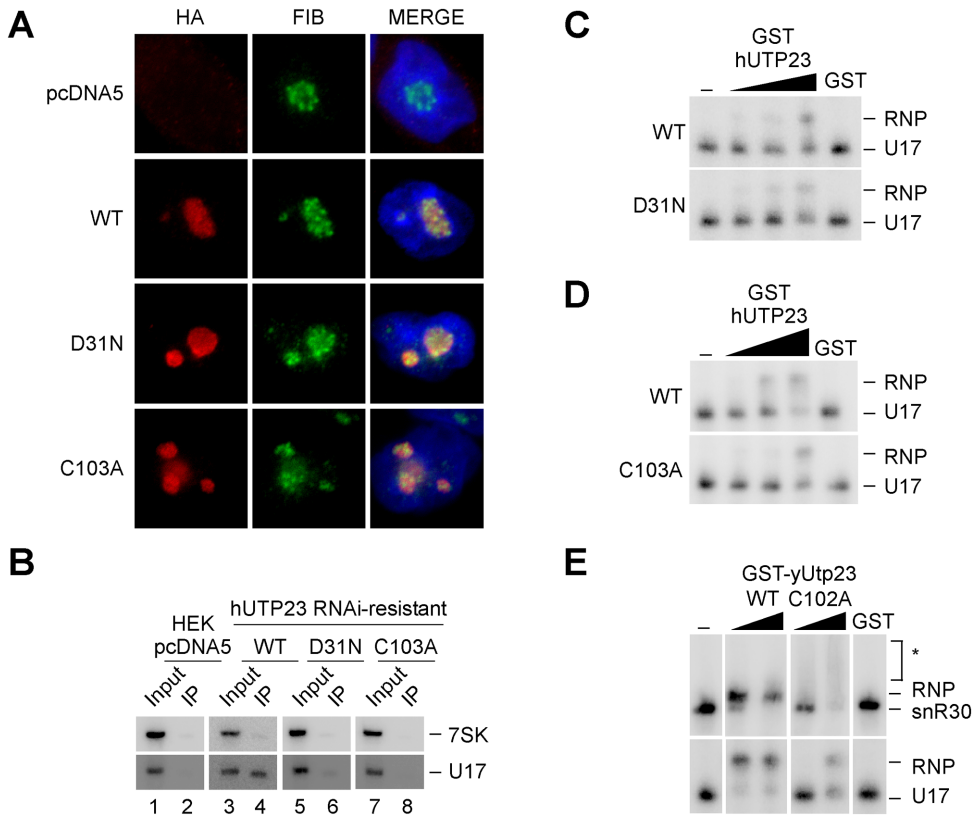


Figure 5. The hUTP23 Zinc finger mutant exhibits typical nucleolar localization, but impaired *in vivo* and *in vitro* binding to U17. (**A**) Cells expressing WT or mutant forms of HA-tagged hUTP23 (D31N, C103A) or no HA-tagged protein (pcDNA5) were harvested and processed for immunofluorescence microscopy using an anti-HA antibody (red). Cells were counterstained with an antibody against endogenous fibrillarin (FIB, green) as a nucleolar marker and DAPI (blue) to highlight the nucleus. From left to right: Immunofluorescence signals for the anti-HA and anti-fibrillarin antibodies and the merged image. (**B**) Soluble lysates from cells, as shown in panel A, were subjected to immunoprecipitation using an anti-HA antibody. Co-precipitated RNA was extracted and separated by denaturing acrylamide electrophoresis, analyzed by northern blotting using radiolabeled probes against the U17 snoRNA and 7SK as a loading control and visualized using a PhosphorImager. A total of 1.25% of the input material was loaded. (**C**) EMSA showing the binding of recombinant wild type (WT, top panel) or D31N PIN mutant (bottom panel) human GST-hUTP23 proteins (450, 900 and 1800 nM) or GST (1800 nM) to *in vitro* transcribed radiolabeled U17 snoRNA. RNA and RNP complexes were analyzed as in Figure 2B. (**D**) Binding of WT (top panel) or the C103A Zn finger mutant (bottom panel) human GST-hUTP23 proteins (750, 1500 and 3000 nM) or free GST (3000 nM) to the U17 snoRNA. Analysis as described in panel C. (**E**) Binding of recombinant WT and Zinc finger mutant (C102A) yeast GST-yUtp23 proteins (250 and 2500 nM) or GST (2500 nM) to snR30 (top) or U17 (bottom) RNAs. Asterisks: unresolved C102A mutant-snR30 complexes, likely due to non-stable protein–RNA interactions. Analysis as described in panel C.

tein essential for snR30 function in 18S rRNA processing (19) and yUtp23 has been identified, alongside yKri1, as a novel non-core snoRNP protein co-purifying with snR30 (15). Interestingly, we detected a direct protein–protein interaction between yUtp23 and yKri1 (Figure 3C). This result was surprising, given that these proteins have been suggested to associate with snR30 in a mutually exclusive manner (15). It was also shown in the same study that yUtp23, but not yKri1, is essential for snR30 release. The observed direct interaction between both proteins may therefore suggest that they are involved in distinct, but possibly consecutive steps of snR30 function, with yUtp23 playing a more active role in snR30 release. Our *in vivo* crosslinking and *in vitro* RNA binding studies have revealed that yUtp23 primarily contacts the top part of the 3′ hairpin and the internal hairpin of snR30 (Figures 2 and 6A). Our data further showed that all three hairpins in snR30 are important for efficient binding of yUtp23 *in vitro*. The finding that the internal hairpin is a major snR30 binding site of the essential yUtp23 protein *in vivo* and *in vitro* is surprising given that this region of the snoRNA has been shown to be dispensable for snR30 function *in vivo* (18). Interestingly, we also discovered an evolutionarily conserved interaction between yUtp23/hUTP23 and the H/ACA snoRNP core component, yNhp2/hNHP2 (Figure 3) and yeast yUtp23 was previously reported to interact with another H/ACA snoRNP component, yGar1 (36). We believe that the interactions between yUtp23 and both yNhp2 and yGar1, which are expected to also bind to the upper part of the snR30 3′ hairpin (Figure 6A), are sufficient to tether yUtp23 on the snR30 snoRNP and thereby compensate for the lack of the internal hairpin *in vivo*. While further experiments are clearly needed to clarify the exact role of yKri1 with respect to snR30 function, our data strongly suggest that yUtp23 is one of the predicted essential snR30-specific protein(s) that bind the 3′ hairpin in the snoRNA.

In vivo protein–RNA crosslinks were also found between yUtp23 and the ES6 region of the 18S rRNA. The speci-

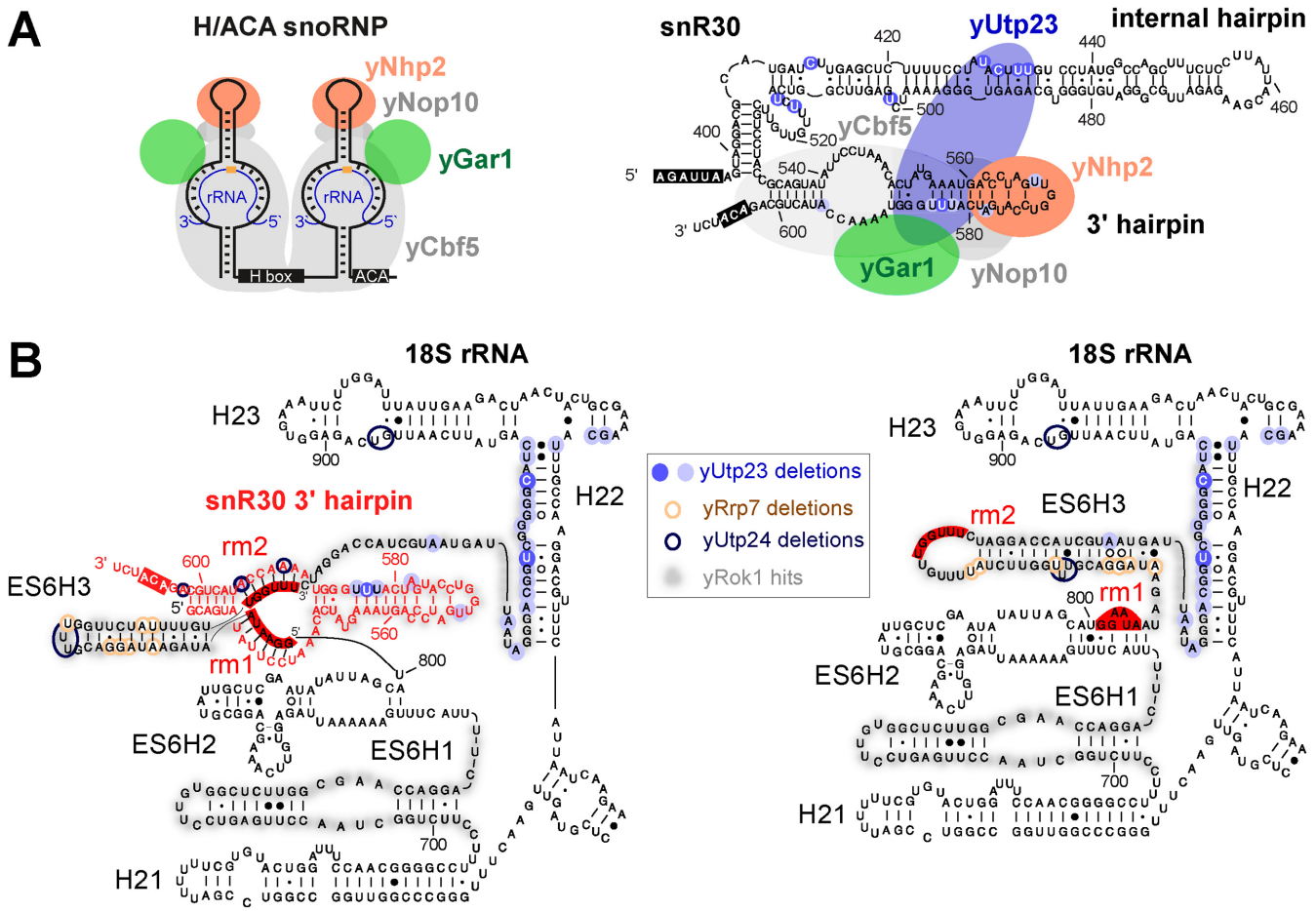


Figure 6. Yeast yUtp23 is central to the integration of ES6 binding factors and snR30 release. (**A**) Left panel: Cartoon of a typical modification H/ACA box yeast snoRNP depicting the positions of the core H/ACA snoRNP proteins yCbf5, yGar1, yNop10 and yNhp2. H and ACA boxes are shown in black. Orange circle: target nucleotide for pseudouridylation in the rRNA (blue). Right panel: Cartoon of the 3′ terminal portion of yeast snR30 with the predicted positions of yUtp23 and core H/ACA snoRNP proteins. yUtp23 microdeletion sites in the internal hairpin and 3′ hairpin are represented as blue circles, shades represent peak height as in Figure 2C. The H and ACA boxes are shown in black. (**B**) Predicted structure of the yeast 18S ES6 region during snR30 base-pairing (left panel) and in the mature rRNA after snR30 release (right panel). Precise crosslinking sites of yUtp23 (blue circles, shades indicate peak height as in Figures 1 and 2), yUtp24 (dark blue) and yRrp7 (brown), and yRok1 crosslinking regions (shaded in grey) on the snR30 3′ hairpin (red) and 18S rRNA sequences (black). Binding sites for snR30 in the ES6 region (rm1 and rm2) and the ACA box are highlighted in red.

ficity of this interaction was validated with an *in vitro* RNA binding experiment (Supplementary Figure S2D). Primer extension analyses of RNAs crosslinked to both yeast yUtp23 and human hUTP23 proteins further revealed that this interaction is evolutionarily conserved (Supplementary Figure S3). Our data therefore show that yUtp23/hUTP23 binds directly to two different RNAs, snR30/U17 and the 18S rRNA ES6 region. Interestingly, the main 18S rRNA interaction site for yUtp23/hUTP23 is in close proximity to the snR30/U17 base-pairing regions in ES6 (rm1 and rm2) and near the known crosslinking sites for yeast yRok1, yRrp7, yDhr1, yUtp24 and yRrp5 (Figure 6B). Furthermore, we could show direct protein–protein interactions between yUtp23 and yRok1, yRrp7 and yUtp24 and their human equivalents hUTP23, hROK1, hRRP7 and hUTP24 (Figure 3). The yUtp23-snR30 and hUTP23-U17 complexes therefore appear to form central hubs within the yeast and human SSU processomes that potentially coordinate the binding of factors to, and their subsequent release from, ES6. Indeed, it has already been shown that yUtp23 is needed for snR30 release from the pre-rRNA (15), a process involving the RNA helicase yRok1 and the compaction factor yRrp5 (23,24).

In vivo crosslinking studies in yeast revealed that several of the protein contact sites in ES6 overlap (Figure 6B), suggesting sequential and coordinated protein binding. Consistent with this idea, some of the ES6 binding proteins, such as yRrp5, yRrp7 and yUtp24, are still present in the 90S pre-ribosome after snR30 and yUtp23 have left the complex (37). Notably, snR30 is required for stable association of yRrp7 with pre-ribosomes, but not vice versa (21). Given the dramatic change in the structure of the yRrp7 binding site in ES6 seen before and after snR30 release (Figure 6B, compare left and right panels), it is unlikely that yRrp7 binds to both structures. yRrp7 is present in 90S pre-ribosomes lacking snR30 and, within this complex, probably still bound to the pre-rRNA after snR30 release. We therefore hypothesize that yRrp7 recognizes the post-snR30 ES6 structure. Supporting this idea, yRrp7 mainly crosslinked to snR10 instead of snR30

(21) and we did not detect stable *in vitro* interactions between yRrp7 and snR30 (Figure 3D). yUtp23 may therefore, together with snR30, function as a platform to recruit yRrp7 to the pre-ribosome through protein–protein interactions. Subsequently, the yRrp7 binding site may then be generated by the yUtp23-dependent release of snR30 and the re-structuring of ES6. Protein–protein interactions between yUtp23 and yRok1 suggest a similar scenario, where yUtp23 could contribute to the role of yRok1 in releasing snR30 from the pre-ribosome. In this context, we were surprised to observe an *in vitro* interaction between yeast yRrp7 and human U17 (Figure 3D). However, future studies in human cells will be needed to clarify whether hRRP7 and U17 interact *in vivo* and whether this is biologically relevant.

Using an RNAi-rescue system, we have also analyzed the function of human hUTP23. Knockdown of hUTP23 inhibits cleavages at sites A0, 1 and 2a (12). These defects were rescued by re-expression of the WT protein, but not hUTP23 containing point mutations in either the Zinc finger or the PIN domain (Figure 4), although WT and both mutant proteins localized to the nucleolus as expected (Figure 5). We observed that the hUTP23 Zinc finger mutant (C103A) showed strongly reduced association with the U17 snoRNA *in vitro* and *in vivo.* The equivalent point mutation in yeast yUtp23 (C102A) also impaired *in vitro* interactions with snR30 (Figure 5) and caused lethality (14). In yeast, snR30 is needed for correct positioning of yUtp23 within the SSU processome to enable yUtp23 binding to the 35S pre-rRNA (15). Our comparative analyses therefore indicate that a direct protein–RNA interaction between yUtp23/hUTP23 and snR30/U17 is evolutionarily conserved and is dependent on the integrity of the Zinc finger.

An intact PIN domain, on the other hand, is not essential for yUtp23 function in yeast (4,14). In humans, however, the PIN domain mutant (D31N) did not rescue hUTP23 function despite correct nucleolar localization. Binding of the recombinant mutant protein to the U17 snoRNA was not affected *in vitro*, indicating that there is no major defect in protein folding caused by the single point mutation in the PIN domain. We did, however, notice significantly decreased co-precipitation of U17 with the hUTP23 PIN mutant *in vivo*. It is therefore possible that the D31N mutation might interfere with other essential hUTP23 interactions within the human pre-40S particle that are needed for U17 association. In yeast and other fungi, two of the four PIN domain active site amino acids are present in yUtp23 (Supplementary Figure S6), however, in metazoan and plant UTP23, three of the key active-site amino acids are conserved. In some PIN-domain proteins, such as SMG6, this is sufficient for nuclease activity (33). Our attempts to detect nuclease activity in human hUTP23, using a variety of human pre-ribosomal RNA substrates containing site A0 have so far been unsuccessful. While the current lack of *in vitro* cleavage data could indicate a non-catalytic role for hUTP23 in human ribosome biogenesis, the strong pre-rRNA processing defects observed upon expression of the PIN mutant *in vivo* make it likely that instead the right *in vitro* substrate has not yet been identified. In addition, other protein factors may be needed for correct RNA folding *in vitro* and/or activation of hUTP23. Indeed, the human PIN domain protein hNOB1 was recently shown to require a co-factor, hCINAP1, for cleavage at site 3 *in vitro* (11). Future experiments will reveal whether hUTP23 is an active enzyme and whether a co-factor(s) is required for its activity.

ACCESSION NUMBERS

GEO database (http://www.ncbi.nlm.nih.gov/geo/): identifier GSE87238.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

The authors thank David Tollervey (Wellcome Trust Centre for Cell Biology, University of Edinburgh) for supporting the early part of this work. The authors also thank Grzegorz Kudla (MRC Institute of Genetics and Molecular Medicine, University of Edinburgh) for help with the CRAC analyses.

FUNDING

Royal Society [UF100666 and RG110357 to C.S.]; Wellcome Trust [092076 to N.J.W.]; BBSRC project [BB/F006853/1 to N.J.W]; ERASMUS internship [to F.W.]. Funding for open access charge: BBSRC [BB/F006853/1 to N.J.W].
Conflict of interest statement. None declared.

REFERENCES

1. Henras,A.K., Plisson-Chastang,C., O'Donohue,M.F., Chakraborty,A. and Gleizes,P.E. (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA*, **6**, 225–242.
2. Phipps,K.R., Charette,J. and Baserga,S.J. (2011) The small subunit processome in ribosome biogenesis-progress and prospects. *Wiley Interdiscip. Rev. RNA*, **2**, 1–21.
3. Watkins,N.J. and Bohnsack,M.T. (2012) The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA*, **3**, 397–414.
4. Bleichert,F., Granneman,S., Osheim,Y.N., Beyer,A.L. and Baserga,S.J. (2006) The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 9464–9469.
5. Sloan,K.E., Mattijssen,S., Lebaron,S., Tollervey,D., Pruijn,G.J. and Watkins,N.J. (2013) Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. *J. Cell Biol.*, **200**, 577–588.
6. Tomecki,R., Labno,A., Drazkowska,K., Cysewski,D. and Dziembowski,A. (2015) hUTP24 is essential for processing of the human rRNA precursor at site A1, but not at site A0. *RNA Biol.*, **12**, 1010–1029.
7. Wells,G.R., Weichmann,F., Colvin,D., Sloan,K.E., Kudla,G., Tollervey,D., Watkins,N.J. and Schneider,C. (2016) The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans. *Nucleic Acids Res.*, **44**, 5399–5409.
8. Fatica,A., Oeffinger,M., Dlakic,M. and Tollervey,D. (2003) Nob1p is required for cleavage of the 3' end of 18S rRNA. *Mol. Cell. Biol.*, **23**, 1798–1807.

9. Pertschy,B., Schneider,C., Gnadig,M., Schafer,T., Tollervey,D. and Hurt,E. (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J. Biol. Chem.*, **284**, 35079–35091.
10. Lebaron,S., Schneider,C., van Nues,R.W., Swiatkowska,A., Walsh,D., Bottcher,B., Granneman,S., Watkins,N.J. and Tollervey,D. (2012) Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nat. Struct. Mol. Biol.*, **19**, 744–753.
11. Bai,D., Zhang,J., Li,T., Hang,R., Liu,Y., Tian,Y., Huang,D., Qu,L., Cao,X., Ji,J. *et al.* (2016) The ATPase hCINAP regulates 18S rRNA processing and is essential for embryogenesis and tumour growth. *Nat. Commun.*, **7**, 12310.
12. Sloan,K.E., Bohnsack,M.T., Schneider,C. and Watkins,N.J. (2014) The roles of SSU processome components and surveillance factors in the initial processing of human ribosomal RNA. *RNA*, **20**, 540–550.
13. Wang,M., Anikin,L. and Pestov,D.G. (2014) Two orthogonal cleavages separate subunit RNAs in mouse ribosome biogenesis. *Nucleic Acids Res.*, **42**, 11180–11191.
14. Lu,J., Sun,M. and Ye,K. (2013) Structural and functional analysis of Utp23, a yeast ribosome synthesis factor with degenerate PIN domain. *RNA*, **19**, 1815–1824.
15. Hoareau-Aveilla,C., Fayet-Lebaron,E., Jady,B.E., Henras,A.K. and Kiss,T. (2012) Utp23p is required for dissociation of snR30 small nucleolar RNP from preribosomal particles. *Nucleic Acids Res.*, **40**, 3641–3652.
16. Morrissey,J.P. and Tollervey,D. (1993) Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol. Cell. Biol.*, **13**, 2469–2477.
17. Lemay,V., Hossain,A., Osheim,Y.N., Beyer,A.L. and Dragon,F. (2011) Identification of novel proteins associated with yeast snR30 small nucleolar RNA. *Nucleic Acids Res.*, **39**, 9659–9670.
18. Atzorn,V., Fragapane,P. and Kiss,T. (2004) U17/snR30 is a ubiquitous snoRNA with two conserved sequence motifs essential for 18S rRNA production. *Mol. Cell. Biol.*, **24**, 1769–1778.
19. Fayet-Lebaron,E., Atzorn,V., Henry,Y. and Kiss,T. (2009) 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. *EMBO J.*, **28**, 1260–1270.
20. Martin,R., Hackert,P., Ruprecht,M., Simm,S., Bruning,L., Mirus,O., Sloan,K.E., Kudla,G., Schleiff,E. and Bohnsack,M.T. (2014) A pre-ribosomal RNA interaction network involving snoRNAs and the Rok1 helicase. *RNA*, **20**, 1173–1182.
21. Lin,J., Lu,J., Feng,Y., Sun,M. and Ye,K. (2013) An RNA-binding complex involved in ribosome biogenesis contains a protein with homology to tRNA CCA-adding enzyme. *PLoS Biol.*, **11**, e1001669.
22. Lebaron,S., Segerstolpe,A., French,S.L., Dudnakova,T., de Lima Alves,F., Granneman,S., Rappsilber,J., Beyer,A.L., Wieslander,L. and Tollervey,D. (2013) Rrp5 binding at multiple sites coordinates pre-rRNA processing and assembly. *Mol. Cell*, **52**, 707–719.
23. Bohnsack,M.T., Kos,M. and Tollervey,D. (2008) Quantitative analysis of snoRNA association with pre-ribosomes and release of snR30 by Rok1 helicase. *EMBO Rep.*, **9**, 1230–1236.
24. Khoshnevis,S., Askenasy,I., Johnson,M.C., Dattolo,M.D., Young-Erdos,C.L., Stroupe,M.E. and Karbstein,K. (2016) The DEAD-box protein Rok1 orchestrates 40S and 60S ribosome assembly by promoting the release of Rrp5 from Pre-40S ribosomes to allow for 60S maturation. *PLoS Biol.*, **14**, e1002480.
25. Granneman,S., Petfalski,E. and Tollervey,D. (2011) A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. *EMBO J.*, **30**, 4006–4019.
26. Granneman,S., Kudla,G., Petfalski,E. and Tollervey,D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 9613–9618.
27. Wlotzka,W., Kudla,G., Granneman,S. and Tollervey,D. (2011) The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J.*, **30**, 1790–1803.
28. Schneider,C., Anderson,J.T. and Tollervey,D. (2007) The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. *Mol. Cell*, **27**, 324–331.
29. Turner,A.J., Knox,A.A. and Watkins,N.J. (2012) Nucleolar disruption leads to the spatial separation of key 18S rRNA processing factors. *RNA Biol.*, **9**, 175–186.
30. Schneider,C., Kudla,G., Wlotzka,W., Tuck,A. and Tollervey,D. (2012) Transcriptome-wide analysis of exosome targets. *Mol. Cell*, **48**, 422–433.
31. Sardana,R., Liu,X., Granneman,S., Zhu,J., Gill,M., Papoulas,O., Marcotte,E.M., Tollervey,D., Correll,C.C. and Johnson,A.W. (2015) The DEAH-box helicase Dhr1 dissociates U3 from the pre-rRNA to promote formation of the central pseudoknot. *PLoS Biol.*, **13**, e1002083.
32. Lamanna,A.C. and Karbstein,K. (2009) Nob1 binds the single-stranded cleavage site D at the 3'-end of 18S rRNA with its PIN domain. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 14259–14264.
33. Glavan,F., Behm-Ansmant,I., Izaurralde,E. and Conti,E. (2006) Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *EMBO J.*, **25**, 5117–5125.
34. Elbashir,S.M., Harborth,J., Weber,K. and Tuschl,T. (2002) Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods*, **26**, 199–213.
35. Tafforeau,L., Zorbas,C., Langhendries,J.L., Mullineux,S.T., Stamatopoulou,V., Mullier,R., Wacheul,L. and Lafontaine,D.L. (2013) The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol. Cell*, **51**, 539–551.
36. Tarassov,K., Messier,V., Landry,C.R., Radinovic,S., Serna Molina,M.M., Shames,I., Malitskaya,Y., Vogel,J., Bussey,H. and Michnick,S.W. (2008) An in vivo map of the yeast protein interactome. *Science*, **320**, 1465–1470.
37. Kornprobst,M., Turk,M., Kellner,N., Cheng,J., Flemming,D., Kos-Braun,I., Kos,M., Thoms,M., Berninghausen,O., Beckmann,R. *et al.* (2016) Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome. *Cell*, **166**, 380–393.