Mutational analysis and function of the 18S rRNA synthesis factor Utp3/Sas10

A thesis submitted to the University of Sheffield for the degree of

Doctor of Philosophy

By

Shler Hussien

Department of Molecular Biology and Biotechnology
University of Sheffield
United Kingdom

April 2017

Abstract

The transcription, modification, processing and assembly of the pre-rRNA into ribosomes requires more than 200 proteins and 79 small RNAs in yeast, and involves most of the transcriptional activity of a cell. A large complex termed the processome is responsible for early cleavage events that lead to 18S rRNA synthesis. The processome subunit Utp3 shares sequence homology with the C1D protein interaction domain of Rrp47. Two hybrid analyses suggest Utp3 interacts with multiple processome components, including the Utp6 and Utp21 components of the Utp-B complex, Mpp10 of the Mpp10 subcomplex and Utp25. Utp3 also interacts with U3 snoRNA. This project aimed to (i) determine the regions within Utp3 that are required for its function *in vivo*; (ii) to map the sites of interaction between Utp3 and the processome components Utp6, Utp21, and Utp25, and (iii) to analyse molecular contacts between Utp3 and the pre-rRNA and/or U3 snoRNA in growing yeast cells. Mutants lacking the central region of Utp3 (which includes the C1D domain) complemented a *GAL::UTP3* but caused dissociation of Utp3 from 90S complexes. In contrast, deletion of either the N- or C-terminal regions did not complement the *GAL::UTP3* and caused depletion of 18S rRNA. Pull-down reactions suggest that interactions between the C1D domain of Utp3 and the processome may be mediated, at least in part, through Utp25. In contrast, interactions with the Utp-B complex are dependent upon both the C1D and CTD domains. Crosslinking and cloning (CRAC) analyses of Utp3/RNA complexes reveals that Utp3 interacts with 35S pre-rRNA and U3 snoRNA directly *in vivo*. These interactions are probably mediated by the CTD domain, since this region of Utp3 can bind nucleic acids *in vitro*. The nature of the critical role of the N-terminal region of Utp3 remains unclear. Cumulatively, the data supports the model that Utp3 is involved in multiple protein-protein and protein-RNA interactions and provides a molecular platform for interactions within the processome complex.

Acknowledgements

Firstly, all praise is to God Almighty for all grants awarded to me. I am especially thankful to my supervisor Dr. Phil Mitchell for giving me a chance to do my PhD in his laboratory and to Prof. Alastair Goldman for making it all possible. Thank you both for your guidance, support and patience, and for always being available to discuss my concerns and ideas. I also extend my thanks to the members of the lab, both past and present, especially to Monika Feigenbutz for her kindness and scientific training, to Will Garland, to Ashley Pridgeon, to Llywelyn Griffith, to Danni Meehan, to Dean and Glyn for technical support, to Mark Johnson for help with IT, to Linda Harris and to all people in the E21 office and everyone else who contributed to this thesis.

I also thank my friend Atiqa, who helped me with table of contents, and Tajdeda and Jess for proof-reading.

The most thanks go to my family: my sisters Halaa, Gulala and my brother Dler for providing me with financial support since 2013. Finally, I would like to thank all my colleagues in the Molecular Biology and Biotechnology department for their welcome and support.

LIST OF ABBREVIATIONS

µg- micrograms

µl- microlitres

ATP- adenosine triphosphate

bp- base pairs

C1D- C1D domain

cDNA- complementary DNA

CPK- central pseudoknot

CTD- C-terminal domain

Da- Daltons

DEPC- diethylpyrocarbonate

DNA- deoxyribonucleic acid

E- eluate

EM- electronmicroscopy

EMSA- electromobility shift assay

ETS- external transcribed region

GFP- green fluorescent protein

GST- Glutathione-S-transferase

GTP- guanosine triphosphate

HCl- hydrochloric acid

His- histidine

HTP- tandem purification tag consisting of the hexahistidinyl tag, the cleavage site of Tobacco etch virus protease, and two copies of the z domain of protein A from Staphylococcus aureus

IP- immunoprecipitation

IPTG- isopropyl-β-D-thiogalactopyranoside

ITS- internal transcribed region

kDa- kilodalton

LSU- large ribosomal subunit

MCR- multiple cloning region

mg- milligram

ml- millilitres

mM- milimolar

mRNA- messenger RNA

rRNA– ribosomal RNA

MCS – multiple cloning site

MW- molecular weight

ng- nanograms

NLS- nuclear localization signal

nt - nucleotide

OD- optical density

Oligos- oligodeoxyribonucleotides

PAGE- polyacrylamide gel electrophoresis.

PCR- polymerase chain reaction

PEG- polyethylenglycol

rDNA- ribosomal DNA

RNA- ribonucleic acid

RNP- ribonucleoprotein particle

SDM site-directed mutagenesis

snoRNA-small nucleolar RNA

snoRNP-small nucleolar ribonucleoprotein

SSU- small subunit

SSUP- Small subunit processome

TBS - Tris-buffered saline with Tween

Table of Contents

List of tables

List of Figures

Chapter 1 : Introduction

1.1: Ribosome biogenesis

Ribosomes are complex ribonucleoprotein (RNP) organelles that translate the genetic code from mRNA and catalyse protein synthesis in all cells. The structure and function of ribosomes is conserved throughout all domains of life.

Ribosomal synthesis is an extraordinarily complex and energy demanding process. It is the main cellular activity in all living organisms. The production of ribosomes is in parallel with the growth and proliferation of cells. In a typical yeast cell, there are about 200,000 ribosomes and approximately 2,000 are synthesised each minute. De novo synthesis of ribosomes in yeast takes about 15 minutes (Osheim et al., 2004; Kos and Tollervey, 2010). This operation requires over 200 non-ribosomal factors, including enzymes such as ATPase's, GTPase's, RNA helicases, kinases, nucleases, RNA modification enzymes and approximately 100 small nucleolar RNAs (snoRNAs) to assemble into pre-ribosomal complexes (Warner, 1989; Woolford and Baserga, 2013).

Pre-ribosomal complexes are assembled by the progressive association of 80 ribosomal proteins with four different ribosomal RNAs. Eukaryotic ribosomes are molecular machines consisting of a large and a small subunit. The large subunit in yeast contains the peptidyl transferase centre (PTC), responsible for the chemical reaction of peptide bond formation, and contains the 5S, 5.8S and 25S rRNAs and 46 ribosomal proteins. The small ribosomal subunit (SSU), which includes the decoding centre, contains one ribosomal rRNA called 18S rRNA and approximately 34 ribosomal proteins (Phipps et al., 2011; Woolford and Baserga, 2013).

The yeast *Saccharomyces cerevisiae* has been studied extensively as a model system for eukaryotic ribosome biogenesis. This is because of the simplicity of its system and the vast amount of genetic information available compared to other eukaryotic cells. Furthermore, approximately all of the proteins involved in ribosome biogenesis in yeast have homologues in higher organisms, making

the yeast ribosome an excellent model for characterizing eukaryotic ribosome structure and function (Venema and Tollervey, 1999). The volume of yeast cell is around 37 μm^3, whereas a human HeLa cell is around 2500 μm^3. In spite of this difference in size, total ribosome synthesis is surprisingly similar in these very different cells (Turowski and Tollervey, 2015).

Deregulation, loss of function or mutation of many assembly factors and ribosomal proteins can cause defects in ribosome biogenesis. The defects in ribosome biogenesis can be lethal in yeast cells, have profound consequences on the health of the organism and can be embryonically lethal or cause disease in higher organisms. Therefore, defects in ribosome biogenesis can be a key to the development of cancer and genetic disease (Ruggero and Pandolfi, 2003; Freed et al., 2010).

1.1.1: Ribosome structure

Ribosomes are large protein-synthesising ribozyme complexes. Eukaryotic ribosomes share generic structural features with their prokaryotic counterparts. Current understanding of ribosome functionality is based on data from both prokaryotic and eukaryotic kingdoms. Ribosomes in both categories of life are universally composed of one large and one small subunit (LSU and SSU, respectively).

There are a number of important differences between prokaryotic and eukaryotic ribosomes. Eukaryotic ribosomes are more structurally complex and larger: a bacterial ribosome is about 2.5 MDa, that of yeast is about 3.3 MDa and those of higher organisms can be up to 4.5 MDa (Melnikov et al., 2012). The assembly of eukaryotic ribosomes and the initiation of translation in eukaryotic cells is subject to additional forms of regulation and quality control, with the degradation of faulty ribosomes (Dinman, 2009). Prokaryotic ribosome maturation requires a relatively small number of non-ribosomal factors, whereas eukaryotic ribosome assembly requires approximately 200 maturation factors (Kressler et al., 2010; Shajani et al., 2011). Twenty five yeast ribosomal proteins do not have bacterial homologues (Ben-Shem et al., 2011). Extra nucleotides present only in eukaryotic rRNAs and absent from prokaryotic

rRNAs are found in specific regions known as expansion segments (ES) (Spahn et al., 2001). These enable attachment of an increased number of ribosomal proteins and appear to consist of distinct localised protein and RNA structures that participate in specific steps of ribosome biogenesis (Spahn et al., 2001; Alkemar and Nygard, 2004; Yokoyama and Suzuki, 2008).

The small 40S subunit of the eukaryotic yeast ribosome contains only 18S (1798 nt) rRNA and 34 ribosomal proteins. The SSU has three distinct secondary structural domains: the head, the beak and the body, this latter part consisting of the platform, the shoulder and the foot. The 3'-end of the 18S rRNA is contained in the head and the 5' end of the 18S rRNA is contained in the shoulder and foot; the platform is in the central rRNA region (Wimberly et al., 2000; Spahn et al., 2001; Spahn et al., 2004; Chandramouli et al., 2008). The SSU platform leads to the decoding centre, which reads the mRNA with the aid of tRNAs. The large 60S subunit of the yeast ribosome is composed of the 25S (3392nt), 5.8S (158 nt) and 5S (121 nt) rRNAs and 46 ribosomal proteins. The LSU contains the peptidyl transferase centre (PTC), responsible for the chemical reaction of peptide bond formation.

The large and small subunits are independently exported to the cytoplasm and do not unite until unwound mRNA is bound by the SSU, which scans in the 5' to 3' direction until the start codon (AUG) that triggers their association is located. The small subunit unites the mRNA with tRNA following Watson and Crick base pairing of the codons and anticodons (Yusupova et al., 2001). In a GTP-dependent step, the large subunit is then recruited to form the complete 80S ribosome. The large subunit contains three binding sites for tRNAs in the platform region: these binding sites are known as A (Acceptor), P (Peptidyl) and E (Exit) (Kapp and Lorsch, 2004). For translation initiation an aminoacylated $tRNA^{Met}$ occupies the P site to provide the carboxyl group for the primary peptidyl transferase reaction (the initiating $tRNA^{Met}$ is the only tRNA that enters directly to the P site and all subsequent tRNAs enter at the A site). The 18S rRNA terminates protein synthesis by binding the release factor eRF1 that recognises a stop codon at the ribosomal A-site (Brown et al., 2015).

1.2: Transcription and production of polycistronic pre-rRNA

The *Saccharomyces cerevisiae* pre-ribosomal RNA synthesis and processing pathway involves many dynamic RNPs and occurs in three separate cellular regions: the nucleolus, the nucleoplasm and the cytoplasm (Woolford and Baserga, 2013; Gerhardy et al., 2014).The starting point of ribosomal biogenesis occurs in a specialised nuclear sub-compartment called the nucleolus that contains around 150-200 tandem clusters of the rDNA repeat sequence, in which all four genes coding for 18S, 25S, 5.8S and 5S rRNAs are found on chromosome XII (Venema and Tollervey, 1999).

Approximately half of the 150 tandem repeats of the rDNA are actively used to create rRNA and the excess of rDNA copies are maintained in a transcriptionally inactive state. Researchers have shown that heavily transcribed genes are toxic to the cells. Therefore, the untranscribed copies function to protect yeast from mutagenic damage that would cause rDNA repeat instability (Woolford and Baserga, 2013).

RNA polymerase I is the most active cellular RNA polymerase transcribing pre-rRNA and is responsible for 60% of the total cellular transcription. To reach the level of ribosome production required by the cell, RNA polymerase I continuously elongates at a rate of 40–60 nt/sec. Studies have shown that changes that affect the elongation activity of RNA Pol I can lead to defects in rRNA maturation (Lebaron et al., 2012).

In yeast, RNA polymerase I (RNA Pol I), transcription starts with the recruitment of an RNA Pol I initiation complex at the rDNA promoter. The Pol I pre-initiation complex requires two basal transcription factor complexes. UAF (UAS-binding upstream activity factor) is associated with the DNA-binding TATA-box binding protein (TBP), common to all RNA polymerases, and binds the upstream promoter element. The core factor (CF) complex composed of Rrn6, Rrn7, and Rrn11 binds the core promoter element. This allows the recruitment of the initiation competent RNA Pol I that is associated with the Pol-I-specific initiation factor Rrn3 (Moss, 2004; Lebaron et al., 2012).

The result of transcription is a single long molecule that encodes three of the four rRNA species as one polycistronic transcript, termed the 35S pre-rRNA. It is approximately 7 Kb in yeast. It contains the 18S, 5.8S and 25S rRNA sequences, named according to their sedimentation properties (Woolford and Baserga, 2013). The fourth rRNA species, 5S rRNA, is transcribed independently of the others as a 3' extended precursor by RNA polymerase III. Transcription of 5S rRNA occurs in the opposite direction to transcription by RNA pol I 5S rRNA is the only ribosomal RNA that is assembled into an independent RNP complex prior to its assembly into the ribosome (Philippsen et al., 1978; Ciganda and Williams, 2011).

The 35S rRNA contains several important non-coding sequences that do not exist as part of the mature ribosome. These are known as external transcribed spacer (3'ETS and 5'ETS) sequences and internal transcribed spacer (ITS1 and ITS2) sequences and are removed during processing of the pre-rRNA (see Fig. 1.2).

1.3: Modifications of the pre-rRNA

Ribosomal RNA molecules, like all classes of RNA, is post-transcriptionally modified at a large number of sites along the molecule (Lebaron et al., 2012). Modifications are mostly directed to highly conserved regions that have functional importance in the ribosome (Lane et al., 1995; Maden, 1986; Kos and Tollervey, 2010).

The initial 35S pre-rRNA transcript undergoes extensive modification of specific nucleotides by small nucleolar RNAs (snoRNAs). The snoRNAs base-pair with their target sequence on the pre-rRNA, flanking the modification site, and thereby guide the appropriate enzymes to their substrates (van Nues et al., 2011; Watkins and Bohnsack, 2012) . This is followed by extensive processing events including both endonucleolytic cleavage and processing from the 5' and 3' ends by exoribonucleases. Some of these processing events do or can occur once transcription of the 35S pre-rRNA is completed but many events occur co-transcriptionally and it has been shown that up to 70% of

nascent rRNA transcripts are cleaved co-transcriptionally in fast growing yeast cells (Bleichert et al., 2006; Kos and Tollervey, 2010).

Eukaryotic RNA modifications include methylation (CH_3) and conversion of uridines into pseudouridines (Ψ) (Brand et al., 1977; Brand et al., 1979). Methylation can occur on either the 2'-O-position of the ribose sugar or on the base of the nucleotide, with the former being more common (Maden, 1990). Ribosomal RNA contains less base modifications than tRNA, but is rich in pseudouridines and ribose 2' methylations (Decatur and Fournier, 2002). All these modifications provide a functional advantage to the ribosome during protein synthesis events through interactions between the rRNA, tRNA and mRNA that confer translational fidelity (Venema and Tollervey, 1999; Baxter-Roshek et al., 2007; Jack et al., 2011). Alternative theories are that nucleotide modifications protect the rRNAs from degradation during assembly events or that they aid folding of the rRNA molecules and increase stability (Hughes, 1996a; Bernstein and Toth, 2012). The latter model is supported by knowledge that pseudouridines have higher potential for hydrogen bonding and are thus presumed to facilitate rRNA folding (Smith and Steitz, 1997; Ni et al., 1997). Similarly, methylation of the 2' hydroxyl group of sugar residues (2'-O-ribose methylation) can impart conformational rigidity to the nucleotide (Kawai et al., 1992).

The number of modified sites found within pre-rRNA correlates with the complexity of the organism. For instance, mammalian pre-rRNA molecules have ~200 modified nucleotides in total (Maden, 1986; Maden, 1990), while yeast rRNAs hold about half of this number of modifications (Ofengand et al., 1995; Smith and Steitz, 1997). Eubacteria have a higher frequency of base modifications and only several sugar methylations and pseudouridylated residues (Smith and Steitz, 1997). Upon methylation snoRNP are removed from pre-rRNA by RNA helicases, it is not known if all snoRNAs detach before early pre-rRNA cleavage steps or if some remain associated for longer. However, Prp43 genetically interacts with proteins required for late assembly steps of the 40S ribosomal subunit that occur in the cytoplasm, causing

speculation that some snoRNAs are not separated until later in assembly(Kiss-László et al., 1996; Pertschy et al., 2009).

1.3.1: snoRNAs involved in modification

In eukaryotes, pre-rRNA modifications are carried out mainly by guide snoRNAs within the nucleolus where the rRNA also assembles with ribosomal proteins to form the pre 40S and 60S subunits (Reichow et al., 2007; Henras et al., 2008). There are over 100 guide snoRNAs, each with one or more target sites of modification (Venema and Tollervey, 1999).

Although the lack of individual or groups of modifications does not affect cell growth or ribosome biogenesis (Kiss-László et al., 1996), ribosomes lacking modifications show reduced translation efficiency and depletion of the essential associated proteins disrupts both ribosome biogenesis and cell growth (Tollervey et al., 1993).

There are two families of guide snoRNAs differing in both structure and function. Box C/D and box H/ACA, snoRNAs are named after characteristic structural features of each class and guide ribose 2'-hydroxy-methylation and pseudouridylation, respectively (Kiss-László et al., 1996; Kiss et al., 2010; Watkins and Bohnsack, 2012).

All snoRNAs become complexed with specific proteins to form small nucleolar Ribonucleoprotein Particles (snoRNPs). Recruitment and association of a core set of snoRNP proteins is required for the maturation, stability and nuclear localisation of snoRNAs. This core includes four protein components of box C/D snoRNPs: fibrillarin/Nop1, Nop56, Nop58, Snu13 and four protein components of box H/ACA snoRNPs: Cbf5, Gar1, Nhp2 and Nop10 (Balakin et al., 1996; Ganot et al., 1997; Lafontaine and Tollervey, 1999; Henras et al., 2004). Nop1 is the methyltransferase that catalyses nucleotide modification in box C/D snoRNPs and Cbf5 is the pseudouridine synthase in box H/ACA snoRNPs (Singh et al., 2008; Zhou et al., 2011). Nop56 and Nop58 assist box C/D snoRNA binding to rRNA (van Nues et al., 2011).

RNA helicases are responsible for unwinding RNA to enable base-pairing duplexes to form between snoRNA and rRNA (Woolford and Baserga, 2013) . (Venema and Tollervey,1999) first suggested that helicases would also be needed to release snoRNAs from pre-rRNA; these duplexes can be up to 21 nucleotides long in the case of those mediated by Prp43 and its co-factor Pfa1 and up to 10 nucleotides long in the case of those mediated by Dbp3, Dbp4, Rok1 and Rrp3 (Woolford and Baserga, 2013).

Depletion of 4 RNA helicases impairs removal of snoRNAs from the 90S processome after modification has occurred. However, depletion of some non-helicase proteins (Utp23, Esf1 and Esf2) has a similar effect and it is proposed that the RNA helicases may have their effect not only through RNA-RNA dissociation activity but also through RNA-protein remodelling within the processome (Woolford and Baserga, 2013).

1.3.2:snoRNAs involved in ribosome biogenesis

While a large number of snoRNAs guide pre rRNA modification, only a few snoRNAs appear to directly guide cleavage of the 35S pre-rRNA to produce 18S pre-rRNA. These are snR30 (box H/ACA), U3 and U14 (box C/D). In contrast to the snoRNAs that function purely to guide pre-rRNA methylation or pseudouridylation, these snoRNAs are essential for cell growth (Li et al., 1990; Hughes and Ares, 1991; Morrissey and Tollervey, 1993; Lemay et al., 2011). Another small snoRNA, snR10, is also involved in 35S cleavage but is not essential for growth (Tollervey, 1987). The RNAase MRP snoRNA is also essential for growth but is not a member of either the box C/D or H/ACA class. RNAase MRP is required for cleavage at site A_3 and contributes to the production of 5.8S rRNA and the LSU (Schmitt and Clayton, 1993).

U14 and snR10 belong to the box C/D class of snoRNAs. They are both implicated in both pre-rRNA processing at sites A_1 and A_2, but not at A_0, as well as having a function in rRNA modification (Liang and Fournier, 1995).Recent studies have found that liberation of U14 snoRNA from pre rRNA requires Dbp4, which is an RNA helicase protein essential for 18S rRNA synthesis. Depletion of Dbp4 impairs the release of U14 snoRNA from pre-rRNA

indicating that Dbp4 unwinds the U14 snoRNA–pre-rRNA duplex (Kos and Tollervey, 2005; Soltanieh et al., 2015).

snR30 is a box H/ACA snRNA that is found throughout eukaryotes. It is essential to rRNA synthesis and is required for pre-rRNA cleavage at A_0, A_1, and A_2 sites (Lemay et al., 2011). Although it is known to bind rRNA, it does not have a known RNA modification target. A protein called Utp23 is required for the removal of snR30 snoRNA from nascent pre-rRNA, since depletion of Utp23 caused accumulation of snR30 RNA in large preribosomal complexes (Hoareau-Aveilla et al., 2012; Lu et al., 2013).

1.3.3: U3 snoRNA involvement in ribosome biogenesis

The most abundant snoRNA in the nucleolus is U3 snoRNA. U3 snoRNA co-purifies with pre-ribosomes and is a non-canonical box C/D-containing snoRNA that has specialised functions in assisting pre-rRNA processing on the pathway of 18S rRNA synthesis and in the maturation and accurate folding of the pre rRNA (Kass et al., 1990; Phipps et al., 2011). U3 snoRNA is evolutionarily conserved and exists in all eukaryotes examined so far, including yeast and humans (Marz and Stadler, 2009; Charette and Gray, 2009). U3 snoRNA from *S. cerevisiae* is over 100 nucleotides longer than U3 from most other eukaryotes, including human and rat, but shares conserved primary and secondary structure elements with them (Hughes et al., 1987).

The secondary structure of U3 snoRNA consists of three domains: a short 5' domain and a longer 3' domain, connected by a hinge region consisting of the 5’ hinge and the 3’ hinge. Each region of U3 snoRNA has specific sequence elements, protein components, and functions (Parker and Steitz, 1987; Segault et al., 1992). The 5' domain contains the conserved sequences GAC, box A and box A’ that base-pair with the pre-rRNA, and both parts of the hinge region. The 3’ domain of U3 snoRNA contains the motifs box B, C and D and is involved in multiple protein interactions. It is associated with the common box C/D snoRNP proteins Nop1, Nop56, Nop58 and Snu13, as well as the U3 snoRNP-specific protein Rrp9 (Venema et al., 2000).

During early 35S pre-rRNA processing, U3 snoRNA is recruited to the pre SSU processome through base-pairing with the 5' ETS and 18S rRNA coding sequence. These interactions are critical for correct folding and processing of the pre-rRNA, including the formation of the central pseudoknot (CPK) within 18S rRNA and preparation for cleavage at A_0, A_1 and A_2 (Fig 1.1). It is thought that base-pairing between 35S pre-rRNA and U3 snoRNA also functions to prevent premature completion of the CPK and pre-rRNA cleavage (Beltrame and Tollervey, 1992).

Initial base-pairing between U3 snoRNA and 35S pre-rRNA occurs within the 5' ETS at so-called site B or ETS 2 (nt 470-479) via the 5' hinge. This interaction is required for the formation of the other two U3 snoRNA/pre-rRNA duplexes (Dutca et al., 2011). The subsequent interactions involve base-pairing between the so called site A or ETS 1 within the 5' ETS (nt 281-291) and the 3' hinge, between the 5' end of 18S rRNA (nt 4-11, 15-22) and the GAC/A' site and between helix 1 at the 3' end of 18S rRNA (nt 1139-1142) and Box A (Hughes, 1996).

The 5' hinge domain of U3 snoRNA is required for the recruitment of the Mpp10 subcomplex, which consists of the proteins Mpp10, Imp3 and Imp4 (Wormsley et al., 2001). Imp3 is important for the hybridisation of pre-RNA and U3 snoRNA, and Imp4 is required for the subsequent release of U3 snoRNA (Shah et al., 2013). Studies have shown that the SSU processome component Utp14 interacts with Dhr1 and activates the helicase activity of Dhr1-like substrates to release U3 snoRNA from pre-rRNA (Sardana et al., 2015; Zhu et al., 2016). Recently, structural studies have revealed the shape of the U3 snoRNP for the first time and have shown it to be located at the core of the 90S pre-ribosomal particle (Kornprobst et al., 2016)

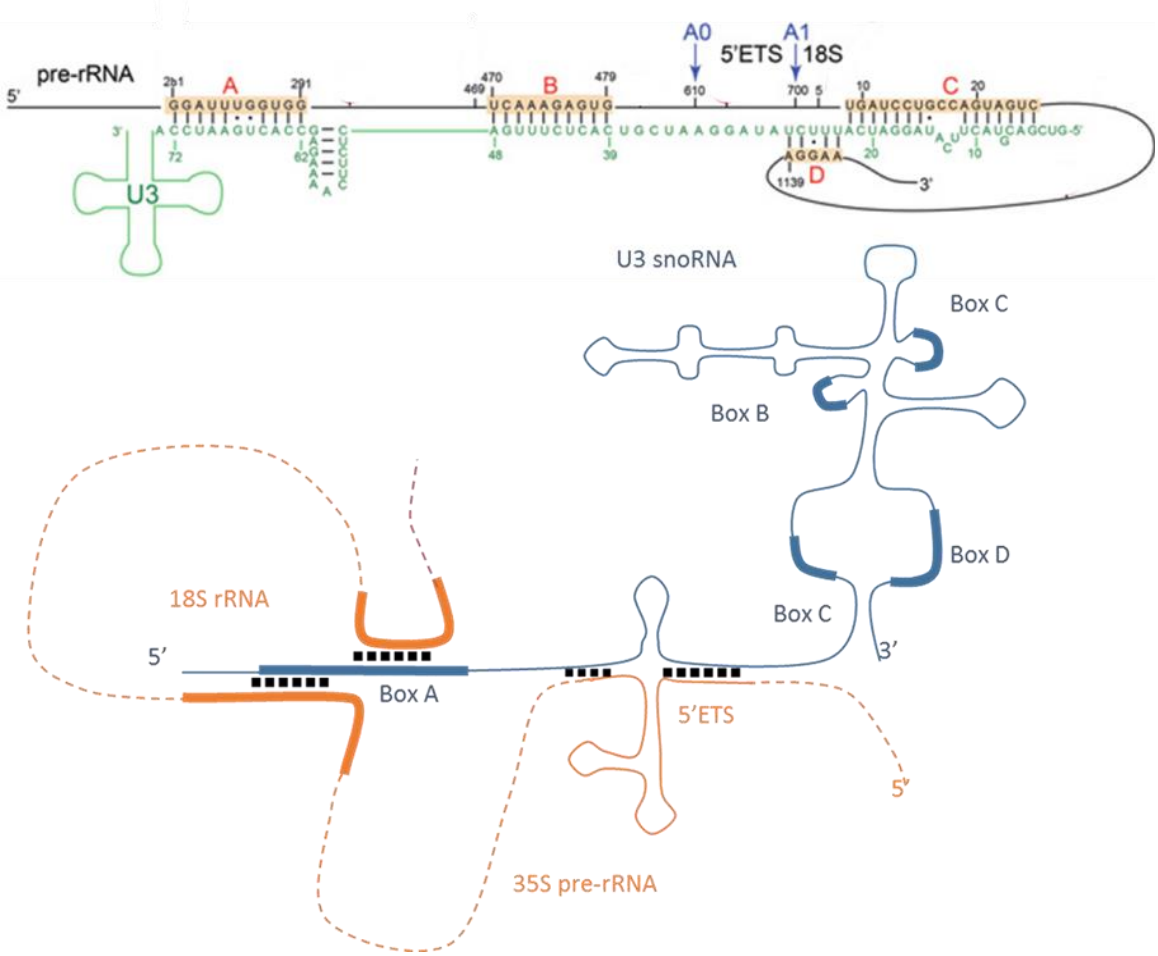

Figure 1.1: Schematic representations of base-pairing interactions between U3 snoRNA and 35S pre-rRNA.

Upper panel: Base-pairing sites are labelled in red, A-D. Specific nucleotide sites on U3 snoRNA and 5'ETS/pre 18S involved in base-pairing are indicated. Cleavage sites A_0 and A_1 on the 5'ETS shown in blue.

Lower panel: U3 snoRNA shown in blue and 35S pre-rRNA shown in orange. Interactions are indicated by black dotted lines. The conserved sequence boxes of U3 snoRNA (Box A, B, C and D) are shown as thickened blue lines. Base-pairing occurs in four places, two between the 3' and 5' hinges of U3 snoRNA and the 5'ETS, and two between the box A and A'/GAC region of U3 snoRNA and complementary nucleotides at conserved locations within the mature 18S rRNA sequence. Orange dotted lines have been used to connect the regions of pre-rRNA that anneal to U3 snoRNA. Adapted from (Dutca et al., 2011; Zhang et al., 2016).

1.4: Processing of the 35S pre rRNA

Within the 90S pre-ribosome (also known as the SSU processome), 35S pre-rRNA undergoes precise, co-ordinated, endonucleolytic cleavage events at sites $A_{0,}$ A_1 and A_2 to remove the 5' external transcribed spacer (5' ETS) and to initiate processing within first internal transcribed spacer (ITS1) (Venema and Tollervey, 1999). These processing reactions are dependent upon base-pairing of U3, U14 and snR30 snoRNAs with the 35S pre-rRNA (Hughes and Ares Jr, 1991; Hughes and Ares, 1991). Depletion of the non-essential snoRNA snR10 also affects pre-rRNA processing (Liang et al., 2010). Several protein factors are additionally required for pre RNA cleavage (Phipps et al., 2011).

Sequential endonucleolytic cleavage of the 35S pre-rRNA at site A_0 and A_1 yields the 33S and 32S precursors, respectively. These RNAs are not stable; the 32S pre-rRNA is quickly processed at site A_2 within the ITS1 region. This splits the pre-rRNA into two sections, a 20S pre-rRNA containing the mature 18S rRNA sequence and the 27S A_2 pre-rRNA that contains the mature 5.8S and 25S rRNA sequences. The 20S pre-rRNA is then exported to the cytoplasm where it is dimethylated by Dim1 and cleaved at site D by Nob1 to form the mature 18S rRNA (Lafontaine et al., 1994; Fatica et al., 2004). The 27S A_2 pre-rRNA is further processed to the mature 5.8S and 25S rRNAs that become a substantial part of the LSU of the ribosome (Fig. 1.2).

A step within the ribosome synthesis pathway can be dependent upon an earlier step, presumably due to prerequisite structural rearrangements of the pre-rRNA during processing (Vos et al., 2004; Ferreira-Cerca et al., 2005; Lamanna and Karbstein, 2011). In other situations, blocking an "early" step does not prevent "later" steps from occurring (Torchet and Hermann-Le Denmat, 2000). Therefore, the order of processing reactions can be dictated by the relative rate at which the processing sites are identified and used, rather than a strict dependence upon completion of a previous step.

The pre-rRNA undergoes endonucleolytic cleavages at precise processing sites within the external and internal transcribed spacers (ETS and ITS,

respectively). Early cleavage events at sites A_0, A_1 and A_2 are coordinated and require the U3 snoRNA, along with a large number of protein components that form a large ribonucleoprotein complex called SSU processome (Armache et al., 2010). Absence of components of the processome can cause loss of cleavage at sites A_0, A_1 and A_2, and result in direct cleavage of the 35S pre-RNA at site A_3. This generates a 27S A_3 pre-rRNA species that can undergo further productive processing to release 5.8S and 25S rRNAs, and a 5'23S pre-rRNA fragment (from the 5' end of the 35S up to site A_3).

The 23S pre-rRNA is degraded in the absence of functional processome components, leading to the depletion of 18S rRNA levels (Fig. 1.2 lower panel). A surveillance mechanism is triggered, which involves polyadenylation of the 23S pre-rRNA by the TRAMP complex and its subsequent degradation by the exosome nuclease complex (Wery et al., 2009). Thus, mutants that affect early processing cleavage events show a loss of 40S ribosomal subunits and normal production of the large 60S ribosomal subunits. Mature rRNAs and pre-rRNA processing intermediates can be distinguished by Northern blot hybridisation, using appropriate oligonucleotide probes. Thus, depletion of ribosome biogenesis factors can cause specific pre-rRNA processing defects that can be visualized by the accumulation and/or depletion of distinct mature rRNAs and pre-rRNA processing intermediates (Phipps et al., 2011).The processome component protein Utp24 appears to be the enzyme responsible for A_1 cleavage. Point mutations at the active site of Utp24, a putative endonuclease PINc domain, inhibit A_1 cleavage (Bleichert et al., 2006), and cross-linking analysis (CRAC) has confirmed Utp24 to cleave at site A_1 (Wells et al., 2016). The processome component Rcl1, a cyclase, has been shown to be involved in cleavage of the pre-rRNA at site A_1 and A_2 (Bleichert et al., 2006; Horn et al., 2011). The nature of the nuclease that makes the initial pre-rRNA cleavage at site A_0 remains unknown. The non-coding spacers, especially the 5' ETS, are thought to be critically important in controlling the sequence of RNA processing. In fact, it has been shown that the 5' ETS is sufficient to nucleate the formation of an approximately 2MDa particle that contains U3 snoRNA and 28 proteins, including several SSU sub-complexes (Chaker-Margot et al., 2015).

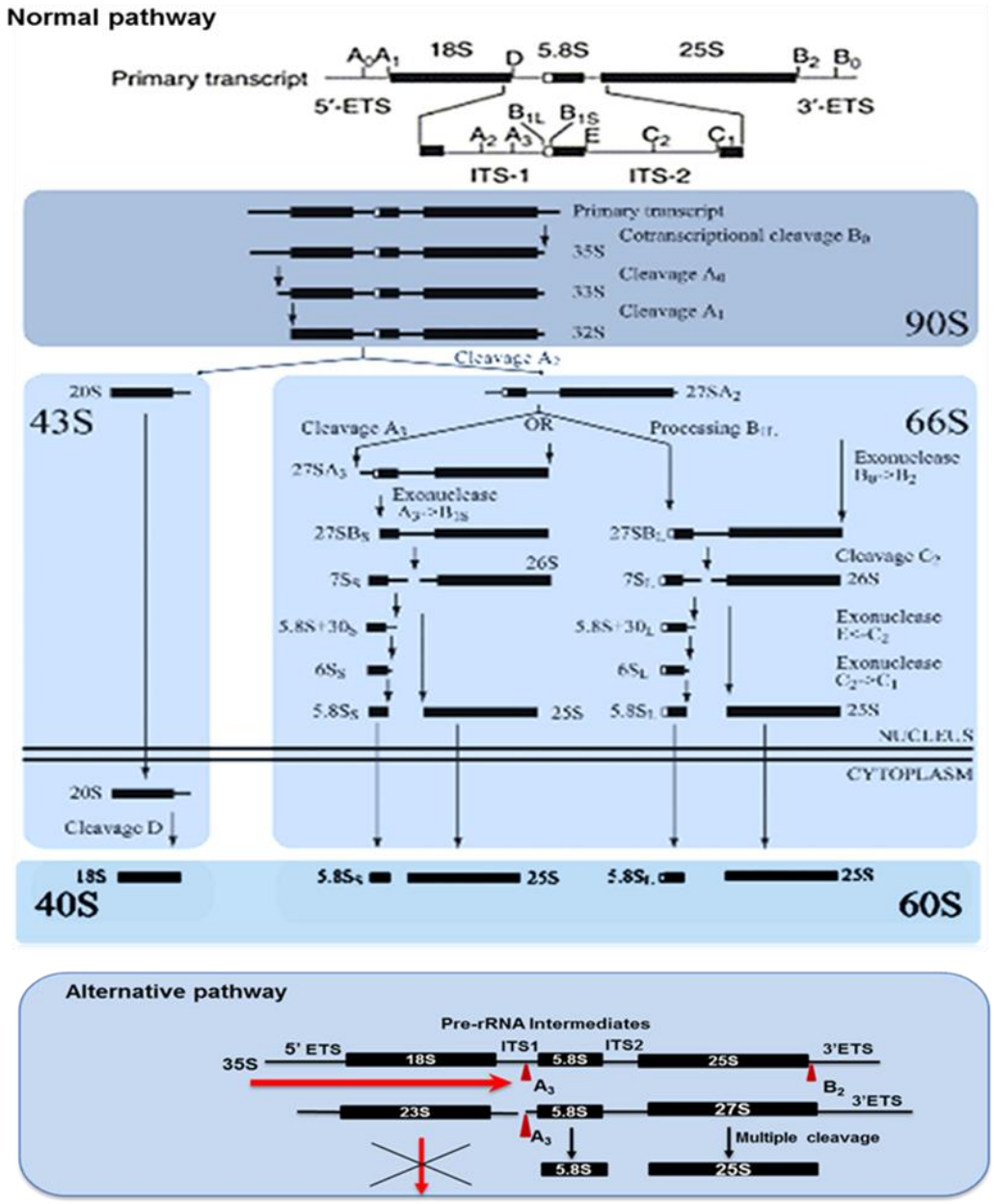

Figure 1.2: Schematic representation of nascent pre-rRNA processing pathway in *S. cerevisiae.*

Upper panel: 35S pre-rRNA with detailed cleavages sites.

Middle panels: Normal 90S processing with endonucleolytic cleavages at $A_{0,}$ A_1 and A_2 followed by exonucleolytic cleavages for the 66S and finally cytoplasmic cleavage at site D of 20S with production of 18S, 5.8S and 25S pre-rRNA intermediates. The pathways essentially act in concert with assembly of the 40S and 60S ribosomal subunits, which are then transported to the cytoplasm.

Lower panel: Loss of 18S potentially due to depletion of one of many SSU assembly factors preventing $A_{0\text{-}2}$ and resulting in accumulation of 23S. Adapted from (Choque et al., 2011).

1.5: The 90S pre-ribosomal particle

The SSU processome or 90S particle that carries out the early processing of the pre-rRNA at sites A_0 - A_2 consists of at least 43 proteins, although some reports suggest as many as 72 proteins (Lim et al., 2011; Zhang et al., 2016) These include a large collection of components known as Utps (U three proteins) (Dragon et al., 2002; Champion et al., 2008). Since pre-rRNA processing occurs together with the assembly of the ribosomal particles in a co-ordinated manner, the processome complex also contains 19 identified ribosomal proteins (Kornprobst et al., 2016).

All SSU processome components examined to date are conserved throughout eukaryotes. Notably, there is strong connection between the function of the SSU processome machinery and progression through the cell cycle; genetic depletion of SSU components results in G1 arrest in yeast (Méreau et al., 1997; Marz and Stadler, 2009; Charette and Gray, 2009).

The rRNA components are transcribed and partially processed in large ribonucleoprotein (RNP) particles before being transported to the cytoplasm, where final processing occurs. Many protein components of the mature ribosome can be found in these intermediate particles, while others join the complex much later in the cytoplasm (Nissan et al., 2002; Lo et al., 2010). The SSU processome is assembled on the 35S pre-rRNA and so contains the sequences found in the mature 25S and 5.8S rRNAs that are found in the large ribosomal subunit. However, proteins involved in LSU biogenesis are not found in the SSU processome and assemble after the LSU rRNAs are released from the polycistronic precursor. This observation strongly suggests that pre-SSU and pre-LSU biogenesis are two different activities (Nissan et al., 2002).

The SSU processome in yeast is a large complex of 2.2 - 6 MDa. It is the largest RNP particle in the cell. It can be visualised by electronmicroscopy (EM) in Miller spreads of chromatin as knobs at the 5’ ends of the nascent pre-rRNA, which have been likened to a Christmas tree in appearance (Dragon et al., 2002) (Fig. 1.3). Studies have shown that early pre-ribosomal particles

change in size and shape, indicating that they are assembled in a step-wise manner (Mougey et al., 1993).

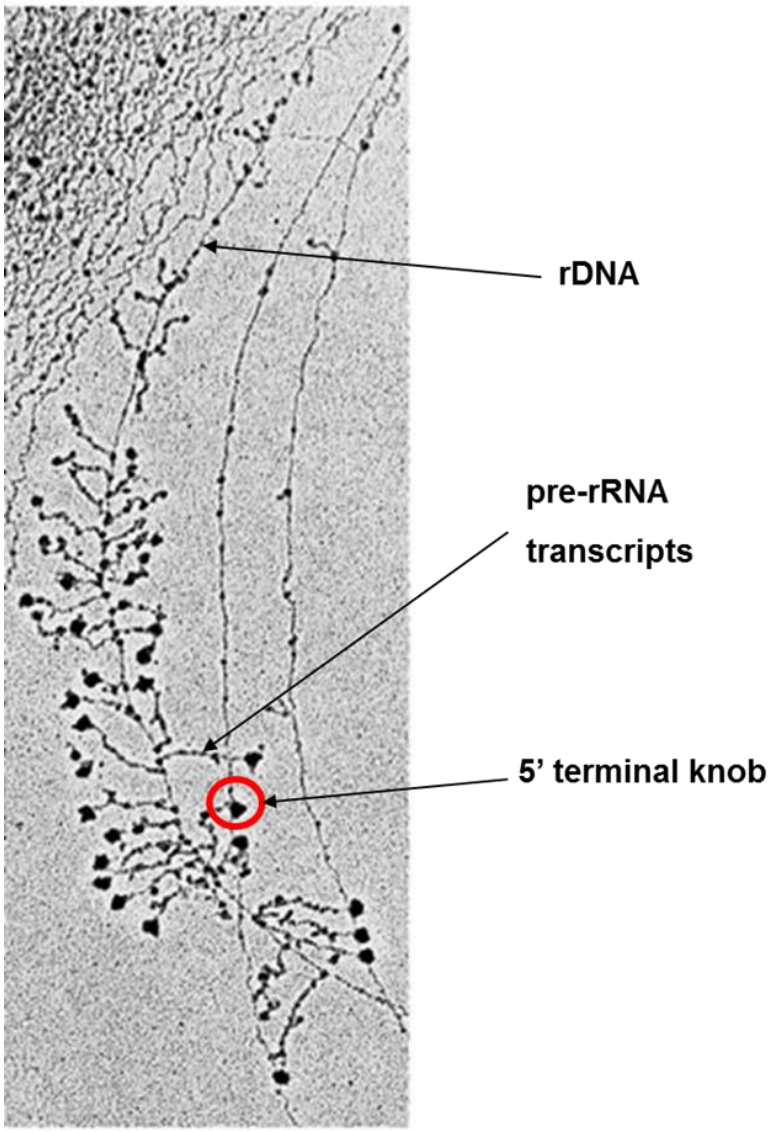

Figure 1.3: A classic electron microscopy (EM) image of Miller spreads showing RNA polymerase I transcription in *S. cerevisiae*

The long filaments represent rDNA chromatin. Pre-rRNA transcripts can be seen as strands branching from the chromatin. Terminal knobs are visible at the 5' end of the nascent pre--rRNA transcripts. Adapted from (Dragon et al., 2002).

At present, the structure and the assembly system of the SSU processome are not fully understood. Despite the fact that a large number of components of 90S pre-ribosomes have been identified and characterised, it is not clear how they co-transcriptionally assemble with the primary transcript to produce a processing-competent particle. Available evidence indicates that eukaryotic cells might reduce the complexity of this operation by pre-assembling 90S

particle components into smaller complexes. These are then assembled on the pre-rRNA in a stepwise and systematic sequence that is dependent on rRNA transcript (Osheim et al., 2004; Bernstein and Baserga, 2004; Perez-Fernandez et al., 2007; Merl et al., 2010; Perez-Fernandez et al., 2011; Zhang et al., 2016).

The 90S processome was first characterised when Baserga's lab purified and identified proteins associated with the box C/D snoRNP component, Nop58, and the U3 snoRNP-specific protein, Mpp10, by tandem affinity purification and mass spectrometry analyses (Dragon et al., 2002).This study identified 17 new U3-associated proteins (called Utp proteins). Further studies subsequently identified roughly 70 SSU components (Chaker-Margot et al., 2015).These studies have demonstrated that a significant fraction of the proteins of the SSU processome (43%), are present within stable subcomplexes. These include the t-Utp/UtpA (transcriptional U three proteins), complex, which links pre-rRNA processing to transcription (Krogan et al., 2004; Gallagher et al., 2004; Granneman et al., 2006), the UtpB complex (Krogan et al., 2004; Champion et al., 2008; Gallagher et al., 2004), the UtpC complex (Krogan et al., 2004), the U3 snoRNP particle (Granneman et al., 2009)and the Mpp10 subcomplex (Dunbar et al., 1997; Charette and Baserga, 2010). Factors that are not present in the above complexes may exist in additional, as yet poorly characterised subcomplexes, as evidenced by interactions between individual components such as Utp7, Utp14 and Sof1 (Karbstein, 2011). Other processing factors are believed to associate with the 90S processome independently (Perez-Fernandez et al., 2007).

Some of these independant factors are RNA helicases, which bind and remodel RNA and RNP complexes in an ATP-dependant manner and participate in virtually all aspects of RNA metabolism in living cells. RNA helicases are thought to enable both the access of and release of RNA binding proteins and snoRNPs to ribosomal RNA. RNA helicase action is modulated by specific co-factors that help locate RNA binding sites and regulate recruitment and enzymatic activity, for instance Prp43 and Pfa1 (Lebaron et al., 2009).

There are 19 RNA helicases involved in ribosome biogenesis in the 90S processome of *S. cerevisiae*. Seven of these (Dbp4, Dbp8, Dhr1, Dhr2, Fal1, Rok1, Rrp3) are required for SSU biogenesis, 10 participate in LSU biogenesis and two (Prp43 and Has1) are required for both (Woolford and Baserga, 2013). All but one are super family 2 (SF2) helicases, characterised by DEAD- or DEAH-box protein motifs (Fairman-Williams et al., 2010). The exception, Mtr4/Dob1, is a Ski 2 family related helicase involved in both rRNA processing and rRNA degradation (Bleichert and Baserga, 2007; Rodríguez-Galán et al., 2013). Depletion of Dbp3, Dbp4, Has1, Rok1, Rrp3 or Prp43 impairs the release of snoRNAs from the processome (Bohnsack et al., 2009). Impairment of RNA helicases Dbp4, Dhr1, Dhr2, Fal1 or Rrp3 causes defects in processing at sites A_0, A_1 and A_2 (Colley et al., 2000; Koš and Tollervey, 2005; Bernstein et al., 2006; Martin et al., 2013; Kos and Tollervey, 2005) .

Prp43 is an RNA helicase (DEAH-box protein) that can be cross-linked to both snoRNAs and their complementary sequences within 25S rRNA, suggesting that it enables snoRNP release (Bohnsack et al., 2009). Prp43 also interacts with Nob1, a PIN domain protein responsible for site D cleavage at the 3' end of 18S rRNA occurring in the cytoplasm (Granneman et al., 2010).Another RNA helicase, Dhr1, associates with both Mpp10 and the U3 snoRNP and has been suggested to trigger structural changes that allow formation of the central pseudoknot(Colley et al., 2000).

Component	**Comments**	**Subcomplex**
t-Utp4	WD40 repeats	**UtpA**
t-Utp5	Utp12 motif, WD40 and GAR repeats	
t-Utp8	No known motifs; also involved in nuclear tRNA export	
t-Utp9	WD40 repeats	
t-Utp10	t-Utp10 domain, HEAT/ARM-type fold and BAP28-like motif	
t-Utp15	WD40 repeats and Utp15 motif	
t-Utp17/Nan1	WD40 repeats	
Pol5	DNA polymerase u; not required for DNA replication; required for rRNA transcription; ARM-type fold	
Utp1	WD40 repeats and Utp12 motif	

Utp6	HAT motifs	**UtpB**
Utp12	WD40 repeats and Utp12 domain	
Utp13	WD40 repeats and Utp13 domain	
Utp18	WD40 repeats	
Utp21	Utp21 domain, and CCs	
Cka1	a-catalytic subunit of CK2; also part of the CURI complex	**UtpC**
Cka2	a9-catalytic subunit of CK2; also part of the CURI complex	
Ckb1	b-regulatory subunit of CK2; also part of the CURI complex	
Ckb2	b9-regulatory subunit of CK2; also part of the CURI complex	
Rrp7	No known motifs; also part of the CURI complex	
Rrp36	No known motifs; contains DUF947	
Utp22	NRAP domain; also part of the CURI complex	
Mpp10	CCs and glutamic acid-rich repeats; associated with the hinge region of U3 snoRNA	**Mpp10**
Imp3	rpS4/S9 RRM	
Imp4	s70-like motif/Brix domain	
Nop1/Lot3	29-O-methyltransferase; GAR and fibrillarin domains	**U3 snRNP**
Rrp9	CCs and WD40 repeats; binds to the B/C motif of U3 snoRNA	
Nop56	Nop5, NOSIC and Nop domains; RNAP I site and KKE/D repeats	
Nop58	Nop5, NOSIC, and Nop domains; CCs, RNAP I site, and KKE/D repeats	
Snu13	rpL7Ae motif; binds to the K-turn motifs of U3 snoRNA; also part of the U4/U6-U5 tri-snRNP	
Bms1	GTPase, stimulated by Rcl1; DUF663 and AARP2CN domains; lysine-rich region and CCs	**Bms1/Rcl1**
Rcl1	RNA-terminal phosphate cyclase-like protein; stimulates Bms1; no cyclase activity detected; also contains a RTC insert domain	
Utp2	Nop14-like domains and CCs; also involved in SSU nuclear export	
Utp3	Utp3 domains, CCs, and a glutamic acid-rich region;disrupts silencing	
Utp7	WD40 repeats and BING4CT; adenylate binding site	
Utp11	Utp11 domain and CCs	
Utp14	Utp14 domain, CCs, and lysine-rich region; ATP/GTP binding site (P-loop)	

Utp16	U3 snoRNA-associated superfamily domain and CCs	
Noc4/Utp19	Noc and CBF domains; also involved in SSU nuclear export	
Utp20	ARM repeats, DRIM motif, and CCs	
Utp23	Lysine-rich region; PINc nuclease domain not required for function	
Utp24	PINc nuclease domain required for function	
Utp25	DEAD-box helicase-like motif; DUF1253 domain (digestive organ expansion factor	
Utp30	rpL1 motif	
Dbp8	DEAD-box RNA helicase; stimulated by Esf2; HELICc domain and CCs	
Dhr1	DEAH-box RNA helicase; HELICc, HA2, and OB Fold/DUF1605 domains	
Dhr2	DEAH-box RNA helicase; HELICc, HA2, and OB Fold/DUF1605 domains	
Nop19/Dhi1	DUF2702	
Emg1	Member of a/b knot fold methyltransferase (SPOUT) superfamily; displays pseudouridine	**Unassigned proteins**
Krr1	KH domain and CCs	
Rok1	DEAD-box RNA helicase; HELICc domain	
Rrp3	DEAD-box RNA helicase; HELICc domain and CCs	
Rrp5	S1 RNA-binding motifs and HAT repeats; binds single-stranded tracts of U's; also participates in A_3 cleavage in 5.8S processing	
Sof1	WD40 repeats and Sof1 domain; similar to b-subunit of trimeric G-proteins	
Dbp4	DEAD-box RNA helicase; HELICc motif and CCs	
Lcp5	Utp3 domain and CCs	
Esf1	Unknown NUC153 domain, CCs, and lysine-rich region	
Esf2	RRM and CCs; binds to RNA and stimulates ATPase activity of Dbp8	
Enp1	ystin domain, glutamic acid-rich region, and CCs; also associates with U14	
Prp43	DEAH-box RNA helicase; also involved in mRNA splicing and LSU biogenesis	

Table1.1:90S processome factors involve in ribosome biogenesis

Mature ribosome complexes and pre-ribosomal processing intermediates are large RNP particles that are well suited to analysis by sucrose density gradient centrifugation. Studies on the sedimentation profiles of distinct of ribosomal particles have provided confirmation of three categories of SSU assembly intermediates in wild-type cells. Particles that sediment at ~90S represent early processing intermediates and contain the full length 35S pre-rRNA (Udem and Warner, 1972; Delprato et al., 2014). Processing within the ITS1 region generates later processing intermediate complexes that sediment at 43S and ~ 66S and which contain the 20S and 27S pre-rRNAs, respectively (Delprato et al., 2014). Co-sedimentation of a protein of interest with the SSU processome or other ribosomal complexes can be readily analysed by ultracentrifugation of lysate on sucrose density gradients. The sedimentation profile of epitope-tagged fusion proteins can detect by Western blotting and compared to the profile observed for ribosomal proteins or specific ribosomal RNA species.

A new technique to isolate and study the composition of intermediate complexes in the assembly of the SSU processome has been reported that uses a dual step, affinity purification approach. This technique depends upon the expression of truncated MS2-tagged pre-rRNA transcripts of known length. It was hypothesized that the expression of truncated pre-rRNAs would restrict the assembly of SSU processome complexes at specific steps along the normal assembly pathway (Perez-Fernandez et al., 2007; Chaker-Margot et al., 2015). More recently, a study using the same technique expanded the available data to provide a framework for the understanding of the biogenesis of the small ribosomal subunit (Zhang et al., 2016). This study resolved the assembly point of 65 individual proteins, along with the U3, U14, and snR30 snoRNA particles to support a stepwise and dynamic view of the assembly of the SSU processome.

1.5.1: 90S processome subcomplexes

The t-Utps/UtpA complex is a heptameric complex composed of Utp4, Utp5, Utp8, Utp9, Utp10, Utp15 and Utp17 (Nan1), and is the first subcomplex to assemble on the 5' end of the nascent transcript. Although many of the SSU processome components are conserved in yeast and humans, Utp8 and Utp9 are specific to yeast, with no known human homologues. The UtpA subcomplex is required for the efficient complete transcription of the 35S pre-rRNA (Perez-Fernandez et al., 2011), and also nucleates the core of SSU processome assembly. It is required for subsequent docking of both the UtpB and UtpC complexes, which occur concurrently but independently of one another (Dosil and Bustelo, 2004; Vos et al., 2004). In yeast, a direct interaction between t-Utp4 and Utp18 (a UtpB component), may mediate UtpB recruitment (Tarassov et al., 2008). More recently, it has been shown that both UtpA and UtpB subcomplexes act as RNA chaperones to initiate eukaryotic ribosome assembly (Hunziker et al., 2016).

The UtpB subcomplex (also known as the Pwp2 subcomplex) is a stable hexameric structure comprised of Utp1/Pwp2, Utp6, Utp12/Dip2, Utp13, Utp18, and Utp21 (Grandi et al., 2002; Krogan et al., 2004). All components of this subcomplex contain repeat motifs: Utp1, Utp12, Utp13, Utp18 and Utp21 consist of WD40 repeats, while Utp6 contains a HAT (half tetratricopeptide) motif. The UtpB subcomplex is a core component of the SSU processome that assembles on the pre-rRNA in concert with the U3 snoRNP particle and is required for the incorporation of the U3 snoRNP particle and the Mpp10 subcomplex into the 90S particle (Perez-Fernandez et al., 2007; Perez-Fernandez et al., 2011; Chaker-Margot et al., 2015).

The UtpC complex is a hexamer containing proteins Utp22, Rrp7, Cka1, Cka2, Ckb1, and Ckb2. The Cka1, Cka2, Ckb1 and Ckb2 proteins constitute the casein kinase 2 complex, which is known to phosphorylate SSU ribosomal proteins (Wojda et al., 2002). This suggests the possibility that the casein kinase 2 complex also phosphorylates components of the SSU processome. Utp C recruitment requires Rrp5, which has not been assigned to a specific subcomplex and appears to be recruited into the SSU processome

independently. Rrp5 is one of the few ribosome processing factors that functions in the biogenesis of both subunits. The C-terminal domain of Rrp5 is essential for early cleavage events at sites A_0–A_2 of the 35S pre-rRNA in the 18S rRNA synthesis pathway, while the N-terminal domain is required for cleavage at site A_3 that initiates processing of the 27S A_2 species in the 5.8S/25S rRNA synthesis pathway (Venema and Tollervey, 1996; Lebaron et al., 2013). Rrp5 contains seven protein interaction domains and interacts with numerous ribosome synthesis factors. It is proposed to form an extremely large complex in this way that contributes to a structural framework for the remodelling of early pre ribosomes (Turowski and Tollervey, 2015).

The Bms1/Rcl1 complex contains two proteins. Bms1 is the only known GTPase in the SSU processome and is not required for the primary assembly steps. Rcl1 is a cyclase found in complex with Bms1 in the particles containing the pre-rRNA substrates for A_2 cleavage. Upon A_2 cleavage, both proteins are incorporated into pre-40S particles. It is proposed that GTP binding or hydrolysis by Bms1 may stimulate conformational rearrangements within the Bms1p-Rcl1p complex that allow the interaction between Rcl1p and its RNA substrate. Additional studies suggest that Rcl1 may be involved in cleavage of the pre-rRNA at site A_1 (Billy et al., 2000; Wegierski et al., 2001; Delprato et al., 2014).

The Mpp10 subcomplex consists of Mpp10, Imp3, and Imp4 (Lee and Baserga, 1999). Imp3 and Imp4 both associate with the U3 snoRNA to stimulate rearrangement of the 18S rRNA structure that is mediated by base-pairing interactions with the U3 snoRNA. Hybridisation of 18S rRNA sequences with U3 snoRNA is dependent upon Imp3, while Imp4 destabilises the U3/18S duplex to dissociate U3 snoRNA (Lee and Baserga, 1999; Gerczei et al., 2009; Shah et al., 2013)

The U3 snoRNP subcomplex includes the Nop1, Nop56, Nop58, Snu13, and Rrp9 proteins. The U3 snoRNP particle is recruited to the 90S processome at the same time as the UtpB subcomplex but these events are independent.

Base-pairing between the U3 snoRNA and the 5' ETS and the 18S rRNA mediate correct pre-rRNA folding (Perez-Fernandez et al., 2011).

With a few notable exceptions, such as Enp1, Dim1, and Nob1, the majority of SSU processome components detach from the 20S pre-rRNA before export from the nucleus to the cytoplasm. Enp1, Dim1 and Nob1 are required for cytoplasmic maturation of the pre-40S small ribosomal subunit precursor complex (Lafontaine et al., 1994). Recently, Rrp12 has been shown to be required for nuclear export of pre-40S particles and both Rrp12 and the exportin Crm1 are required for proper maturation during 40S ribosomal subunit biogenesis (Moriggi et al., 2014). (Fig1.4).

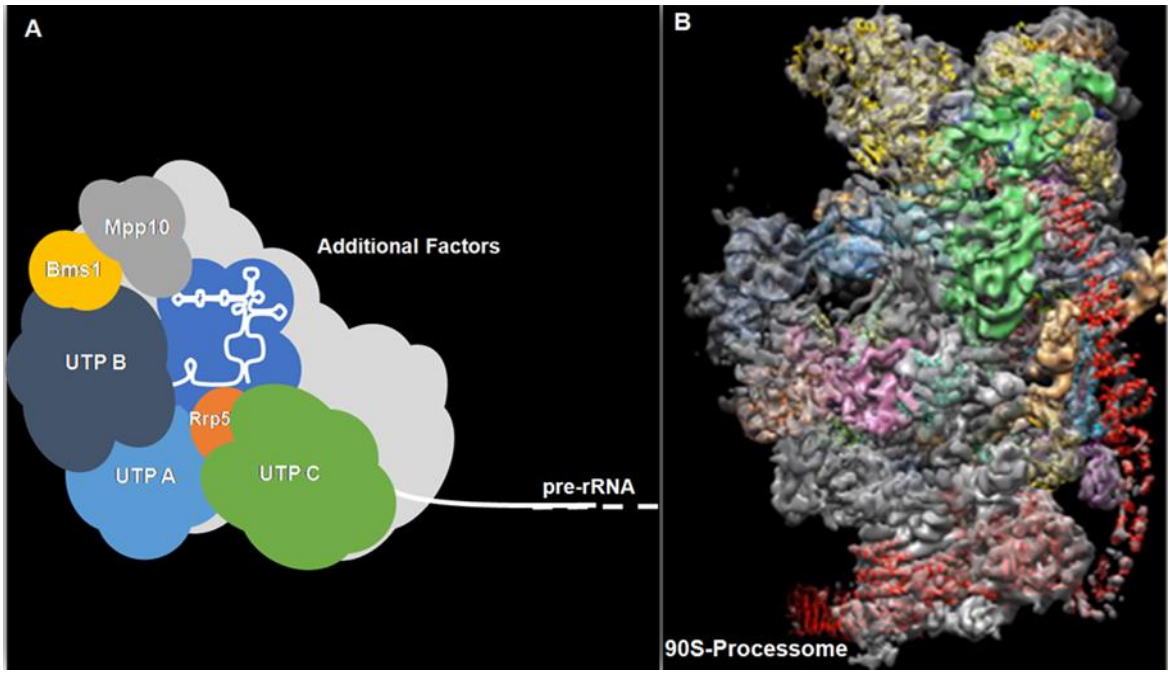

Figure 1.4: Model of the 90S SSU processome based on current knowledge

A: The processome consists of the UtpA, UtpB, UtpC, Mpp10, U3 snoRNP, and Bms1-Rcl1 subcomplexes, in addition to snoRNAs, numerous ribosome biogenesis factors, ribosomal proteins and the nascent pre RNA transcript. Subcomplexes and other components are incorporated into the SSU processome in a sequential and hierarchical fashion as it associates with the nascent pre-rRNA. Adapted from (Perez-Fernandez et al., 2007)

B: The structure of the 90S SSU processome, as determined by cryo-EM. (Kornprobst et al., 2016).

1.6: UtpB subcomplex interaction network

UtpB is a hexameric complex with a molecular weight 525 kDa (Grandi et al., 2002; Krogan et al., 2004). All UtpB components are essential for growth in yeast and are highly conserved across eukaryotic species. Utp1, Utp12, Utp13 and Utp21 each contain two WD40 domains, while Utp18 contains one WD40 domain. The WD domain consists of 40–60 residues arranged in a β-propeller structure and is involved in mediating protein-protein interactions in a variety of cellular activities (Voegtli et al., 2003; Li et al., 2006; Stirnimann et al., 2010). WD domains in the UtpB subcomplex may mediate interactions with other subcomplexes or pre-rRNA processing factors. In addition, WD40 repeats are predicted to interact with unfolded proteins (Tarassov et al., 2008). Utp6 possesses HAT repeats, which also function as a protein-interaction domain (Dosil and Bustelo, 2004; Champion et al., 2008). An interaction network involving the six components of the UtpB subcomplex has been mapped by yeast two-hybrid analyses. However, the particular domain of each protein that is involved in the identified protein-protein interaction is not defined in most cases (Champion et al., 2008; Bartelt-Kirbach et al., 2009).

Structural studies of the UtpB subcomplex indicate a structural core consisting of Utp12, Utp13, Utp21 and Utp1, with a smaller module consisting of Utp6 and Utp18. The N-terminal region of Utp18 attaches to the N-terminal of Utp21 and is required for Utp6 association. The Utp6/Utp18 module appears to be flexible and contains the HAT motif of Utp6 (Zhang et al., 2014; Hunziker et al., 2016)

Utp21 is a central component of the UtpB subcomplex. It is nucleolar protein composed of 939 amino acids with a predicted molecular weight of 105 kDa that is required for 18S ribosomal RNA biogenesis (Dragon et al., 2002; Schafer et al., 2003). Mutations in WDR36, the gene encoding the human homologue of Utp21, have been associated with primary open angle glaucoma (Monemi et al., 2005).

Utp21 interacts with four of the five components of UtpB complex (Utp18, Utp12, Utp6 and Utp1). Studies of the protein-protein interactions involving

Utp21 have demonstrated that residues 274–279 of Utp21 constitute a key binding site with Utp6 and the tandem WD domains of Utp21 interact with residues 100–190 of Utp18 (Champion et al., 2008; Zhang et al., 2014). Utp18 has recently been shown to function as an adaptor protein to recruit the TRAMP and exosome nuclease complexes to the 23S pre-rRNA species to mediate its degradation (Thoms et al., 2015). Utp21 also interacts with the independent SSU biogenesis factors Sas10/Utp3 and Utp25 and these protein-binding sites are most likely conserved on the surface of the Utp21 structure (Zhang et al., 2016).

Utp21 further interacts with Hsp90, a molecular chaperone protein that is required for the folding and maturation of hundreds of cytosolic and nuclear proteins that play key roles in cellular signalling pathways (Zhao et al., 2005; Wu et al., 2012). Additionally, it has been shown that a mutation in Hsp90 causes depletion of Utp21 or accumulation of misfolded Utp21 that cannot perform its essential functions (Tenge et al., 2014).

Utp6 is composed of 440 amino acids with a predicted molecular weight of 52 kDa and is both essential for yeast viability and required for pre-rRNA processing. It has been shown that mutations in the HAT domain of Utp6 cause defects in pre-rRNA processing, accompanied by a corresponding defect in growth (Champion et al., 2008). Additionally, haplo-insufficiency of Utp6 in humans may contribute to the severity of neurofibromatosis type1 (Jenne et al., 2003).

The HAT repeat is a protein-protein interaction motif found only in proteins involved in RNA metabolism. In yeast, the HAT repeat containing proteins Prp6, Prp39, Prp42, Clf1, and Syf1 participate in pre-mRNA splicing (Ben-Yehuda et al., 2000) HAT repeats are also found in Utp6 and Rrp5, both proteins required for pre-rRNA processing (Dragon et al., 2002). Utp6 interacts with Utp18 via its N- terminal and is reported to interact with Utp21 via its HAT domain (Champion et al., 2008). Utp6 may possibly provide an additional contact between the UtpB complex and the Mpp10 complex through its interaction with Mpp10(Charette and Baserga, 2010).

1.7: The role of Utp25 in SSU processome and ribosome biogenesis

Utp25 is a component of the SSU processome that is not a subunit of one of the characteried subcomplexes. The protein is 721 amino acid residues in length and has a predicted molecular weight of 84kDa (Charette and Baserga, 2010). Utp25 is essential for mitotic growth in yeast and is required for 18S rRNA production; depletion of Utp25 causes accumulation of the 35S pre-rRNA and the aberrantly processed 23S pre-rRNA, and depletion of mature 18S rRNA levels. Utp25 contains DUF1235 domain, which are found in several eukaryotic proteins that are localized in the nucleolus in yeast, zebrafish, and humans (Giaever et al., 2002; Hazbun et al., 2003; Huh et al., 2003; Ahmad et al., 2009).

The structure of Utp25 consists of two distinct regions. The N-terminal region, comprising residues 1- 213, contains no known sequence or structural motifs and is poorly conserved. This region is composed of simple, repetitive, negatively charged sequences. Similar repetitive regions are found in other SSU processome components (including Mpp10, Rrp9, Utp3, Utp14, and Utp18) and proteins that function in other aspects of RNA metabolism. In contrast, the C-terminal region of Utp25, consisting of residues 214 – 721, contains the DUF1253 domain that shows limited sequence similarity to the DEAD box helicase domain. Although Utp25 is related in sequence to members of the DEAD-box helicase family, it is itself not a functional helicase (Charette and Baserga, 2010).

Yeast two-hybrid interaction studies suggest that Utp25 interacts with the SSU processome components Utp3 and Mpp10 (Dunbar et al., 1997; Charette and Baserga, 2010; Goldfeder and Oliveira, 2010). Utp25 also co immunoprecipitates with U3 snoRNA, strongly indicating that it is a component of the SSU processome. Additionally, it has been shown that Utp25 copurifies under stringent conditions with Nop19, another SSU processome component required for A_0, A_1 and A_2 cleavages. Nop19 appears to function independently of Utp subcomplexes and to be essential for the incorporation of Utp25 into

pre-ribosomes. Nop19 also interacts preferentially with the RNA helicase Dhr2 (Choque et al., 2011).

1.8: Utp3

The systematic name for the *UTP3* gene is YDL153c. Utp3 was first characterised as a gene product that, when overexpressed, repressed heterochromatin silencing. The protein was named Sas10 (something about silencing 10). Genetic analyses showed that this suppression functions separately from the Sir protein silencing mechanism. Sas10 was shown to be essential for yeast growth and localised completely to the nucleus in all stages of the cell cycle. Utp3 depletion resulted in cell cycle arrest in late S or G2/M phase (Kamakaka and Rine, 1998).

Utp3 is composed of 610 amino acid residues and has a predicted molecular weight of 70,259 kDa. There is no available structural data for Utp3 but secondary structure predictions and sequence alignments suggest the existence of a number of distinct structural features (Fig 1.5). The N-terminal region, like that of Utp25, has no structural motifs and shows regions of simple, repetitive, negatively charged sequences. Residues 219-298 of Utp3 comprise a C1D domain (also called a Sas10/C1D domain). This domain is also found in Lcp5, another pre-rRNA processing factor and SSU processome component. Residues 344-364 are predicted to constitute a coiled-coil structure. The C-terminal domain of Utp3, comprising of residues 532-608, is conserved and is strikingly rich in basic residues. This could in theory contribute to RNA binding (Wilson and Szostak, 1999).

Utp3 is now well established as a component of the SSU processome. N-terminally HA-tagged Utp3 coimmunoprecipitates with U3 snoRNA (Dragon et al., 2002). It has been shown that genetic depletion of Utp3 has the equivalent effect on growth and early pre-18S rRNA processing as depletion of the core SSU processome component, U3 snoRNA (Hughes and Ares Jr, 1991).

Utp3 does not belong to either the UtpB or Mpp10 subcomplexes, but it does interacts with both Utp21 and Utp6 components of the UtpB subcomplex and the Mpp10 subunit of the Mpp10 subcomplex, as demonstrated by yeast two-hybrid experiments (Charette and Baserga, 2010). Utp3 is also able to interact with Utp25, another SSU processome component that is not a component of a characterised subcomplex. Yeast two-hybrid studies have shown that the first 227 residues of Utp3 are sufficient for the interaction with the DUF1253 domain of Utp25 (Charette and Baserga, 2010; Goldfeder and Oliveira, 2010). Utp3 is not required for SSU processome assembly. However, it might provide a link between the UtpB and Mpp10 subcomplexes (Charette and Baserga, 2010)

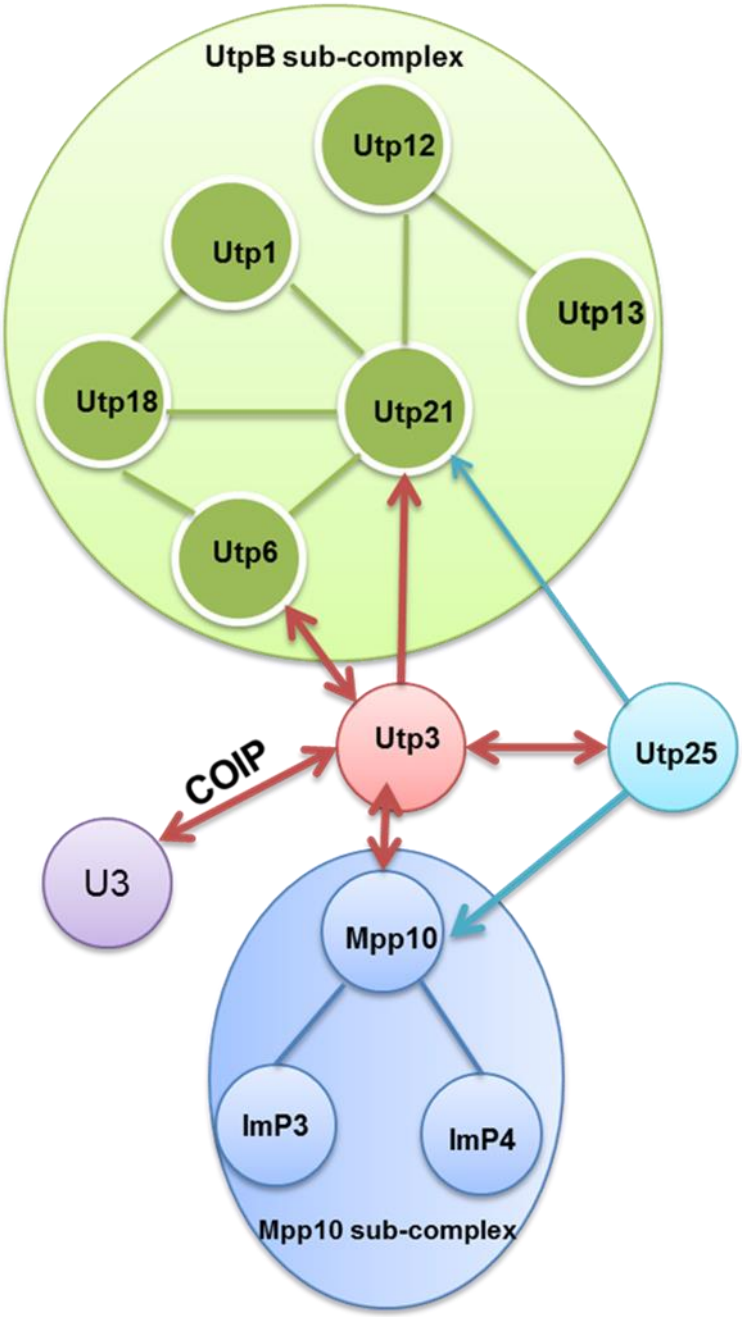

Figure 1.5: Schematic representation of Utp3 interactions.

Schematic showing previously characterised interactions between Utp3, Utp25 and processome components. Yeast two-hybrid interactions between Utp25 and Utp21 or Mpp10 are shown in blue. Yeast two-hybrid interactions between Utp3 and Utp25, Utp6, Utp21 and Mpp10 are shown in red. Interactions between bait and prey proteins are shown as arrows, double-headed arrows indicate interactions observed with proteins both as bait and prey. The coimmunoprecipitation interaction between Utp3 and U3 snoRNA is also indicated (Dragon et al., 2002). Adapted from (Charette and Baserga, 2010).

1.9: The C1D domain family

One feature of the Utp3 protein of particular interest is the C1D domain. The C1D domain consists of approximately 80 amino acids and was initially predicted through bioinformatical analyses (Staub et al., 2004), but has since been established experimentally (Costello et al., 2011). The C1D domain is found in the yeast proteins Utp3, Lcp5 and Rrp47/Lrp1, as well as the metazoan C1D protein (the Rrp47 homologue) and neuroguidin (Jung et al., 2006; Stead et al., 2007; Finn et al., 2010; Mitchell, 2010)(Fig 1.6).

Utp3 and Lcp5 are part of the SSU processome and are essential for the early pre-rRNA cleavage events that lead to the synthesis of the 18S rRNA of the small ribosomal subunit (Wiederkehr et al., 1998; Dragon et al., 2002). Rrp47 is the specific cofactor of RNA exosome complexes containing the exoribonuclease Rrp6 (Mitchell et al., 2003). It has been demonstrated that the C1D domain of yeast Rrp47 forms a heterodimer with the PMC2NT domain in the N-terminal region of Rrp6 and this heterodimer interacts with the RNA helicase Mtr4, thereby providing a physical link between the exosome and TRAMP complexes (Schuch et al., 2014).

The human homologues of both Rrp6 and Rrp47 (PM/Scl-100 and C1D protein, respectively) also interact with each other, with evolutionary conservation of the role in exosome function (Schilders et al., 2007). The C1D protein, like Rrp47, is also involved in DNA repair (Yavuzer et al., 1998). The C1D domain has been suggested to have a coordinating role in maintaining genomic stability (Jackson et al., 2016). Neuroguidin has been shown to interact with the eukaryotic initiation factor eIF4E and to function in cytoplasmic polyadenylation-dependent translational control in neuronal cells. Neuroguidin is proposed to inhibit the expression of specific mRNAs and, like Rrp47, acts as part of a heterodimeric complex (Jung et al., 2006). Recently, it has become clear that neuroguidin is also homologous to Lcp5 and functions in ribosome synthesis (Bammert et al., 2016).

The human Rrp47 homologue C1D protein was initially identified as a DNA-binding protein (Nehls et al., 1998) and yeast Rrp47 also binds RNA and DNA

in vitro with comparable affinity (Stead et al., 2007). However, the C1D domain of Rrp47 is not sufficient for nucleic acid binding, residues at the C-terminus being necessary for this activity *in vitro* (Costello et al., 2011). Similarly, the C1D domains of Lcp5 and Utp3 are not sufficient for DNA binding, nor is there convincing evidence of RNA binding *in vitro.* However, Lcp5 has been shown to interact directly with nucleotides within 18S rRNA *in vivo* (Turner, 2011). Genetic interactions reveal that the C1D domain of Rrp47 is sufficient for protein function *in vivo*. This suggests that the critical function of Rrp47 lies in its ability to interact with Rrp6, rather than its function as a nucleic acid binding protein. The interaction between Rrp47 and Rrp6 is important for their mutual stability (Garland et al., 2013; Feigenbutz et al., 2013). Moreover, the interaction between the Rrp47/Rrp6 heterodimer and Mtr4 allows a physical coupling between the exosome and TRAMp complexes (Schuch et al., 2014). Although the C1D domain of Rrp47 has been demonstrated to form part of the interface of the Rrp47/Rrp6 heterodimer, to be critical for normal Rrp6 expression levels (and therefore to be required for normal RNAase activity of the exosome) and to mediate the interaction between the exosome and TRAMP complexes, the function of the C1D domain in other members of the family has not been addressed. To date, the function of the C1D domains in these proteins has been largely inferred

It has been suggested that the C1D domains of Utp3 and Lcp5 might be responsible for binding to DNA/RNA and other interacting proteins simultaneously, allowing them to target these binding partners to specific nucleic acid substrates (Mitchell, 2010). The C- terminal domains of Utp3, Lcp5 and Rrp47 are all rich in base residues so have an affinity for nucleic acid binding (Wilson and Szostak, 1999). A central hypothesis of this work is that the C1D domain of Utp3 is involved in an interaction with other processome component(s) that is structurally analogous to the Rrp47/Rrp6 heterodimer. Given that both Rrp47 and Lcp5 have been demonstrated to interact with RNA (Costello et al., 2011; Turner, 2011), it is strongly predicted that Utp3 interacts directly with the 35S pre-rRNA.

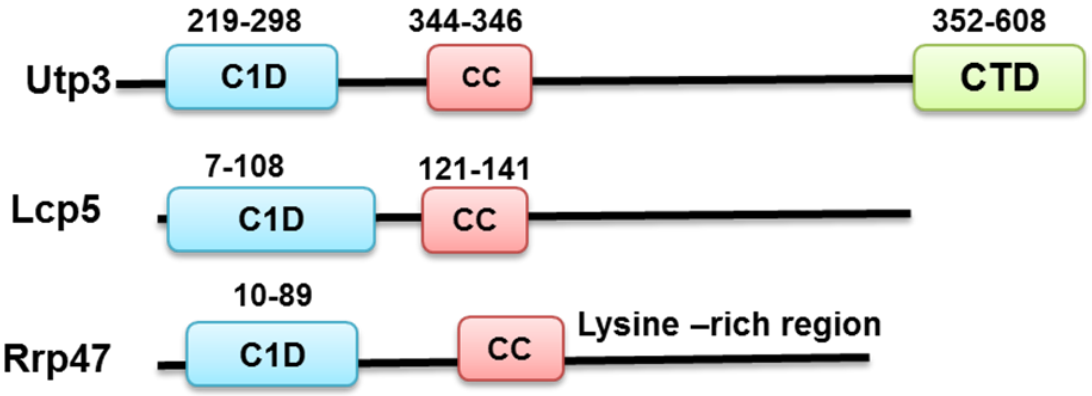

Figure 1.6: Schematic of the domain structure of proteins that share a C1D domain. The domain structure of the yeast proteins Rrp47, Lcp5 and Utp3/Sas10. The amino acid residues that comprise the Sas10/C1D domain of all three proteins, the predicted coiled coil (CC) regions in Utp3 and Lcp5, and the Sas10 CTD domain of Utp3 are indicated. The positions of the domains were taken from the Pfam database (Finn et al., 2016).

1.10:Aims of this study

As illustrated above, the complexity of ribosome biogenesis requires over 200 non-ribosomal factors and approximately 80 small nucleolar RNAs (snoRNAs), many of which assemble into subcomplexes of pre-ribosomal complexes that process and package the pre-rRNA together with the correct ribosomal proteins.

One aim of this project was to address the molecular function of the SSU processome factor Utp3 by generating a series of deletion mutants, based on the known domain structure of the protein, and assaying their ability to complement phenotypes associated with a Utp3 loss of function mutant. Furthermore, in order to define the regions within Utp3 required for interaction with its known binding partners Utp6, Utp12 and Utp25, the Utp3 deletion mutants were also analysed in protein pull-down assays. In addition, CRAC (cross-linking and analysis of cDNA) experiments were performed on Utp3 to determine the RNA binding sites for this protein in growing cells on a genome-wide scale and at nucleotide resolution. Finally, the nuclear localisation of Utp3 was confirmed using a GFP fusion protein and the deletion mutations were assayed to determine the protein domains that are required for correct subcellular localisation. Further detailed information regarding protein-protein and RNA-protein interactions will shed light on the spatial/temporal assembly of the 90S processome and its subsequent disassemble

Chapter 2 : Materials and Methods

2.1: Materials

2.1.1:Plasmids

(Plasmids numbers refer to the laboratory nomenclature reference system)

Plasmid	Description of the plasmid	Supplier Reference
p44	A pRS416 derivative containing the *RRP4* promoter upstream of zz tag cassette for expression of zz N-terminal fusion proteins in yeast.	(Mitchell et al., 1996)
p693	A construct for the expression of an N-terminal zz fusion of Utp3 in yeast. The *UTP3* ORF and downstream region were cloned into p44.	P.J.Mitchell
p694	A derivative of p693 lacking the *EcoR*I site in the *UTP3* 3' UTR. Generated by SDM on p693 using o934 and o935.	This study
p699	A derivative of p694 lacking the *EcoR*I *site* within the *UTP3* ORF. This construct has a unique *EcoR*I site at the 5' end of the *UTP3* ORF. Generated by SDM on p694 using o932 and o933.	This study
p709	A derivative of p699, containing an *EcoR*I *site* 5' of the C1D domain. Generated by SDM using o936 and o937.	This study
p710	A derivative of p699, containing a stop codon at residue 556. Generated by SDM, using primers o938 and o939.	This study
p714	A derivative of p699, containing a stop codon at residue 577. Generated by SDM, using primers o940 and o941.	This study
p718	Expression construct for the *UTP3* ΔN mutant. Generated by *EcoR*I digestion and religation of p709.	This study
p729	A derivative of p699, containing a *Cla*I site and adjacent stop codon at residue 210. Generated by SDM, using o628 and o629. Intermediate to generate	This study

	an internal deletion within the *UTP3* ORF.	
p730	A derivative of p729, containing a *Cla*I site downstream of a stop codon at residue 523 in the *UTP3* ORF. Generated by SDM, using o704 and o705.	This study
p731	Expression construct for the *UTP3* ΔM mutant. Generated by *Cla*I digestion and religation of p730.	This study
p574	Expression construct for the *UTP3* 532X mutant.	P.J.Mitchell
p575	Expression construct of the *UTP3* ΔC1D mutant.	P.J.Mitchell
pRSETb	Bacterial expression vector for N-terminal His-tagged proteins.	Invitrogen
p742	Bacterial expression vector for His-tagged fusion of the N-terminal region of *UTP3*	This study
p469	Bacterial expression vector for His-tagged fusion of the C1D domain of *UTP3*.	P.J.Mitchell
p739	Bacterial expression vector for His-tagged fusion of the C-terminal region of *UTP3*	This study
pGEX6 P-1	Bacterial expression vector for N-terminal GST fusion proteins.	Amersham (Smith and Johnson, 1988)
p741	Bacterial expression vector for the expression of GST-Utp6 fusion protein.	This study
P742	Bacterial expression vector for the expression of GST-Utp21 fusion protein.	This study
P743	Bacterial expression vector for the expression of an N-terminal GST fusion of the DUF1253 domain of Utp25.	This study
p461	Bacterial expression vector for the expression of N-terminal fusion proteins containing the GB1 and hexahistidinyl tag.	(Hautbergue et al., 2008)
p805	A construct for the expression of an N-terminal GFP fusion of *UTP3* in yeast.	This study
p806	A construct for the expression of the *UTP3* 556X mutant as a GFP fusion protein.	This study
p807	A construct for the expression of the *UTP3* 577X	This study

	mutant as a GFP fusion protein.	
p808	A construct for the expression of the *UTP3* ΔN mutant as a GFP fusion protein.	This study
p809	A construct for the expression of the *UTP3* ΔM mutant as a GFP fusion protein.	This study

Table2.1: Plasmids used for construction and expression

2.1.2:Oligonucleotides used in this study

(Oligo numbers refer to the laboratory nomenclature reference system)

Oligo	Sequence	Description/Use
932	gctccactttcaaaagaatttacagaattagctccaaagtttgac g	To remove the *EcoR*I site within the *UTP3* UTR by SDM
933	cgtcaaactttggagctaattctgtaaattcttttgaaagtggagc	
934	gctccactttcaaaagagttcacagaattagctccaaagtttga cg	To remove the *EcoR*I site within the *UTP3* ORF by SDM
935	cgtcaaactttggagctaattctgtgaactcttttgaaagtggag c	
936	aatatggatgacgaagcaagggagaattctttaaggacaatgt ttcctga-	To introduce an *EcoR*I site 5' of the *UTP3* C1D domain.
937	tcaggaaacattgtccttaaagaattctcccttgcttcgtcatcca tatt	
938	ggtctaacgcctaaaagaaactagtataacagaaactctcgtg tc	To introduce a stop codon within the *UTP3* ORF at residue K556.
939	gacacgagagtttctgttatactagtttcttttaggcgttagacc	
940	caaaaggcacaaaagaaactctagacggttcgtgcagtttatt ctggt	To introduce a stop codon within the *UTP3* ORF at residue K577.
941	accagaataaactgcacgaaccgtctagagtttcttttgtgcctt g	
628	aatatggatgacgaagcaatcgattaatatttaaggacaatgttt cc	To introduce a *Cla*I site 5' of the *UTP3* C1D domain.
629	ggaaacattgtccttaaatattaatcgattgcttcgtcatccatatt	
704	ctgttattgctgcaaggtaaatcgatttggcggaattggctg	To introduce a *Cla*I site 5' of the *UTP3* CTD domain.
705	cagccaattccgccaaatcgatttaccttgcagcaataacag	

954	aaaggatcctatggtacgcaaaggct	To amplify the N-terminal region of the *UTP3* gene by PCR.
955	aaaggtacc ttaccttgcttcgtcatcca	
956	aaaggatccattggcggaattggctg	To amplify DNA encoding the C-terminal domain of *UTP3* by PCR.
957	aaaggtaccttaattcttgaatttaacaga	
415	ggaccatggttcctgaatttgct	amplification of the C1D and CTD domains of *UTP3*
1028	aaaaagctt ttaattcttgaatttaacaga	
962	aaacccgggtatgtcagaactctttgga	amplification of the *MPP10* gene
963	aaactcgagtcaaagttttatatttgtgc	
964	aaacccgggtatgtcgaagacaagatac	amplification of the *UTP6* gene
965	aaactcgagttaaagtttgctgataattaa	
966	aaacccgggt atgtctatc gacttgaaaaa	amplification of the *UTP21* gene
967	aaactcgagtcacgcggtggtcac	
968	aaacccgggt agtgacagtgaatcatata	amplification of the *UTP25* DUF1253 domain
969	aaactcgagttatttaaattcataaatttcc	
0447	attcgcctttgctggcc	amplification of the *UTP3* gene
0448	tgacccactcttcttctt	
0449	ctagagataacgtgtgtgt	amplification of the *UTP3* gene
0450	cccatgtcaaattctttgg	

Table 2.2: Oligonucleotides used in this study.

2.1.3:Oligonucleotides for Northern probes

(Oligo numbers refer to the laboratory nomenclature reference system)

Oligo No	Sequence	Description/Use
o236	gcgttgttcatcgatgc	To detect 5.8S rRNA
o405	catggcttaatctttgagac	To detect 18S rRNA
o406	ctccgcttattgatatgc	To detect 25S rRNA
o443	ttcggtttctcactctggggtac	To detect U3 snoRNA
o486	tggtgcgaactctgtg	To detect $tRNA^{Phe}$
o925	ctactcggtcaggctc	To detect 5S rRNA

Table 2.3: Oligonucleotides for Northern probes.

2.1.4: Bacterial strains used in this study

Bacterial strains	Genotype	Supplier Reference
BL21 (DE3) pLysS	*F– dcm ompT hsdS(rb-mb-) GAL λ(DE3) [pLysS Camr]*	Stratagene
DH5α	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	Stratagene

Table 2.4: Bacterial strains used in this study

2.1.5: List of yeast strains

Strain	Genotype	Construction/Source
P851	*MATa ade2-101 his3Δ200 leu2Δ1 lys2-80 trp1-Δ63 ura3-52 KANMX6-GAL::HA3-UTP3*	A kind gift from Susan Baserga (Yale) (Dragon et al. 2002).
P852	*Mata ade2-101 his3Δ200 leu2Δ1 lys2-80 trp1-Δ63 ura3-52*	A kind gift from Susan Baserga
P364	*Mata his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	BY4741. EUROSCARF, Frankfurt
P1584	*Mata his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 HTP-UTP3*	Lab stock strain
P1065	*Mata his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 UTP3-TAP::HIS3*	Lab stock strain
P1066	*Mata his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 UTP3-HTP::HIS3*	Lab stock strain
P1863	*Mata his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 UTP21-GFP::HIS3*	Invitrogen (Hu et al., 2003)
P1864	*Mata his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 UTP25-GFP::HIS3*	Invitrogen (Hu et al., 2003)
P1865	*Mata his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 UTP6-GFP::HIS3*	Invitrogen (Hu et al., 2003)

Table 2.5: yeast strains used in this study

2.1.6:Media used in this work

ercentages given are weight/volume. 2% agar was added for solid media.

Media	Composition
LB (Luria Bertani) broth	1% Peptone, 0.5% Yeast Extract, 1% NaCl
YPD	2% Peptone, 1% Yeast Extract, 2% Glucose
YPGal	2% Peptone, 1% Yeast Extract, 2% Galactose
Minimal medium(SD)	0.17% yeast nitrogen base without amino acids (Melford), 0.5% ammonium sulphate, 2% glucose. Amino acids and bases were supplemented as required from 100 X stocks by adding to the media prior to autoclaving.
Minimal medium (S Gal)	0.17% yeast nitrogen base without amino acids (Melford), 0.5% ammonium sulphate, 2% galactose. Amino acids were added as required and were initially prepared as 100X stocks in water, as described (Table 2.6) and then added to media prior to autoclaving.

Table 2.6: Media used in this study

2.1.7: Weights of amino acid salts (Sigma) used to generate 100X stocks

Amino Acid	Mass per litre (g)
Adenine hemisulfate	2
Arginine monohydrochloride	2
Histidine monohydrochloride	2
Leucine	6
Lysine monohydrochloride	3
Methionine	2
Phenyalanine	5
Threonine	20
Tryptophan	2
Tyrosine	3
Uracil	2

Table 2.7: List of amino acid salts (Sigma) used to generate 100X stocks

2.1.8: Buffers and solutions used in this work

Buffer/Solution	Preparation
TfbI Solution	10mM KAc, 100mM $RbCl_2$, 10mM$CaCl_2$, 50mM $MnCl_2$, 15% v/v glycerol. The pH was adjusted to 5.8 using 0.2M acetic acid.
TfbII Solution	10mM MOPS, 75mM $CaCl_2$, 10mM $RbCl_2$, 15% v/v glycerol. The pH was adjusted to 6.5 using 0.5M KOH
TE(Tris-EDTA) buffer	10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0
LiT Buffer	10mM Tris-HCl pH 7.6, 1mM EDTA pH 8.0 ,100mM lithium acetate pH 7.5
TBE	90mM Tris, 90mM boric acid, 2mM EDTA pH 8.0
TGS(Tris-Glycine SDS running buffer)	25mM Tris, 192mM glycine, 0.1% (w/v) sodium dodecyl sulphate
Western transfer buffer	12.5mM Tris, 96mM glycine, 0.1% SDS, 20% methanol
Trans –blot Turbo transfer systems buffer	1 x transfer buffer (Bio-Rad), 20% ethanol
Western Blot Blocking Buffer	1% w/v dried, skimmed milk powder (Sigma) in 1 X TBS.
TBS (Tris-buffered saline)	10mM Tris-HCl pH 7.4, 150mM NaCl
TBST(Tris-buffered saline + Tween)	10mM Tris-HCl pH 7.4, 150mM NaCl, 0.1% Tween 20
ECL solution 1	2.5mM luminol (Sigma), 400µM p-coumaric acid (Sigma), 100mM Tris-HCl (pH 8.7). The solution was stored at 4°C
ECL solution 2	5.4mM H_2O_2, 100mM Tris-HCl (pH 8.7). The solution was stored at 4°C
Yeast Alkaline Lysis Extraction solution	0.2M NaOH, 0.2% v/v β-mercaptoethanol. The solution was stored at 4°C
10 X MOPS Buffer	200mM MOPS pH 7.0, 50mM sodium acetate, 10mM EDTA. After dissolving, the pH was set to 7 with NaOH.
1X SSPE	150mM NaCl, 9mM sodium dihydrogen phosphate, 1mM EDTA. The buffer was prepared as a 20 X stock

	solution and, after dissolving, the pH was set to 7 with NaOH.
Norther stripping solution	0.1 x SSPE, 0.1% SDS
50X Denhardts solution	2% Ficoll, 2% polyvinyl pyrrolidone, 2% BSA. The solution was stored at -20°C.
Blot neutralisation buffer	0.5M Tris-HCL pH7.4, 1.5M NaCl
DEPC-H_2O	DEPC (diethyl pyrocarbonate) was added to Millipore filitered H_2O to final concentration of 0.1%, mixed, left overnight and then autoclaved.
6X DNA loading dye	0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol, 30% (v/v) glycerol
2X RNA loading dye	95% formamide, 20mM EDTA, 0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol.
2X protein loading buffer	160mM Tris-HCl pH 6.8, 10% β-mercaptoethanol, 2% SDS, 10% glycerol
3X glyoxal RNA loading buffer	50% DMSO, 1M glyoxal, 20mM MOPS pH 7.8, 0.2mg/ml ethidium bromide, 0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol
Pull down cell lysis buffer	50mM HEPES-NaOH pH 7.4, 250mM NaCl, 5mM $MgCl_2$, 10% glycerol.
Pull down wash buffer	50mM HEPES-NaOH pH7.4, 250mM NaCl, 5mM $MgCl_2$, 0.1% NP40, 1mM DTT.
Imidazole (5M stock)	5M Imidazole-HCl, pH 7.6
H150 lysis buffer	50 mM Tris-HCl pH7.6, 150mM NaCl. PMSF (100 mM stock solution in isopropanol) was added to a final concentration of 2 mM
H 150 wash buffer	50mM Tris-HCl pH 7.6, 150mM NaCl, 10mM imidazole
H 150 elution buffer	50mM Tris-HCl pH 7.6, 150mM NaCl, 250mM imidazole
Alkaline lysis solution I	50mM glucose, 25mM Tris-HCl pH8, 10mM EDTA
Alkaline lysis solution II	0.2M NaOH, 1% SDS.
Alkaline lysis solution III	3M potassium acetate, 11.5% glacial acetic acid.

TMN-150 lysis buffer for sucrose gradient	10mM Tris-HCl pH 7.6, 5mM $MgCl_2$, 150mM NaCl
EMSA buffer	100mM Tris-HCl pH 8.0, 200mM KCl, 20mM $MgCl_2$, 1mM EDTA

Table 2.8: Buffers and solutions used in this study

2.1.9: Antibodies

Antibody	Storage (°C)	Dilution	Incubation period	Manufacturer
Rabbi anti glutathione-S-transferase	-20	1:10,000	2 hours	SIGMA
Goat anti-rabbit peroxidase conjugate	-20	1:10,000	1 hour	SIGMA
peroxidase/anti-peroxidase conjugate	-20	1:10,000	1 hour	SIGMA
mouse anti-green fluorescent protein	-20	1:3,000	2 hours	Roche
goat anti-mouse peroxidase conjugate	4	1:5,000	1 hour	Bio-Rad
mouse anti-phosphoglycerate kinase 1	4	1:10,000	2 hours	Invitrogen
mouse anti-penta his	-20	1:10,000	2 hours	QIAGEN
rabbit anti-mouse peroxidase conjugate	-20	1:20,000	1 hour	SIGMA

Table 2.9: Antibodies used in this study

2.2: METHODS

2.2.1: Preparation of *E. coli* competent cells (rubidium chloride method)

E. coli cells were grown on LB agar overnight at 37℃. A few colonies were inoculated into 5ml LB broth overnight. The culture was diluted 100-fold into pre-warmed LB medium and incubated with shaking at 37℃ until the cell density reached OD_{595nm} of 0.48. The cells were then harvested, transferred to 50ml Falcon tubes following centrifugation at 3,200 xg at 4℃ for 5min and the medium was thoroughly drained from the cell pellet. Cell pellets were resuspended in 40ml chilled tfb1buffer, incubated on ice for 10 min and then harvested. The cell pellet was dried again, resuspended and incubated in 5ml of tfb2 buffer. Cells were incubated on ice for 15min and then 100µl aliquots of competent cells were snap-frozen in liquid nitrogen and stored at -80℃ until use.

2.2.2: Transformation of *E. coli* competent cells

20-100ng of plasmid DNA were introduced into 100µl of thawed competent cells. The mixture was kept on ice for 30 minutes, subjected to heat shock at 42°C for 90 seconds and then placed on ice for 2 minutes. A volume of 900µl LB medium was added and the cells were incubated at 37°C for 1 hour to allow transformants to recover. The cells were centrifuged, resuspended in LB broth and spread on agar plates containing the appropriate antibiotic and incubated at 37°C overnight.

2.2.3: Pellet preparation for the Isolation of plasmid DNA from transformed *E. coli*

5ml LB broth containing the appropriate antibiotic was inoculated with a single colony of a transformed *E. coli* strain grown on solid growth medium. The samples were incubated at 37°C with shaking overnight. Cells were harvested via centrifugation at 15,000xg for 1 minute and supernatant was discarded. The plasmid was purified from the cell pellet either by the alkaline lysis method (Maniatis et al., 1982) or using a commercially available spin column kit (E.Z.N.A. Minikit I).

2.2.3.1: Alkaline lysis method

Cell pellets were obtained from 1.5ml saturated overnight cultures. The pellets were resuspended in 100µl alkaline lysis solution I. A volume of 200µl alkaline lysis solution II was added and mixed by inversion; the mixture was then incubated on ice for 5 min. This was then followed by the addition of 150µl of cold alkaline solution III. The microfuge tubes were inverted several times to neutralise the solution and the mixtures were incubated on ice for a further 5 minutes. The mixtures were then centrifuged for 10 minutes at 13,000xg. The resulting supernatant was transferred to a new microfuge tube, mixed with an equal volume of phenol/chloroform, briefly vortexed and centrifuged for 5 minutes. The aqueous phase obtained from this process was then transferred to a fresh tube and DNA precipitation performed by the addition of two volumes of chilled 100% ethanol. The DNA was pelleted by centrifugation for 20 minutes at 13,000xg, followed by washing twice with chilled 70% ethanol. The pellets were air-dried and then resuspended in 19µl H_2O. RNAase A (1µl of a 1mg/ml solution) was added to the resuspended pellets and 1µl of the plasmid miniprep sample was then analysed by agarose gel electrophoresis.

2.2.3.2: Miniprep kit method

Plasmid preparations that were made for lab stocks, subsequent site-directed mutagenesis or for sequencing analysis were purified from cell pellets (as described in section 2.2.3) using commercial spin column kits (E.Z.N.A. Minikit I). Plasmid DNA was isolated according to the manufacturer's instructions.

2.2.4: Purification of genomic DNA from yeast

Genomic DNA was isolated from yeast for PCR amplification of the *UTP3* gene and for confirmation of alleles. The yeast genomic DNA extraction method was adapted from (Hoffman and GARRISON, 1997). 10ml of yeast culture was grown to saturation in appropriate growth medium and the cells were harvested by centrifugation at 3,200xg for 5 minutes. The supernatant was discarded and the cell pellet was washed with 5ml TE buffer. The cell pellet was resuspended in 200µl cell breaking buffer (10 mM Tris-HCl pH 8, 100 mM NaCl, 1 mM EDTA, 1% SDS, 2% Triton X-100). 200 µl glass beads (Sigma) and 200µl phenol/chloroform pH 8 were added and the cell mixture was

vortexed for 5 min. 200 µl TE buffer was then added, the mixture vortexed briefly and then centrifuged for 5 min at 15,000xg. The supernatant was transferred into a new microfuge tube and the DNA was precipitated by the addition of 2 volumes of ice-cold 100% ethanol and incubation at -20°C for one hour. The samples were centrifuged for 20 minutes at 15,000xg and the cell pellet was dissolved in 400µl TE buffer. RNA was digested by the addition of 30µg RNAase A (30µl of a 1mg/ml solution) and incubation of the mixture at 37°C for 30 minutes. The DNA was recovered from the mixture by phenol/chloroform extraction and precipitation at room temperature with an equal volume of isopropanol in the presence of 2.5M ammonium acetate. The DNA was resuspended in TE buffer and subsequently precipitated with isopropanol 3 times. The pellet was washed with 70% ice-cold ethanol, air-dried and resuspended in 50µl TE buffer. One µl of the purified yeast genomic DNA was routinely analysed by agarose gel electrophoresis.

2.2.5: Polymerase Chain Reaction

PCR reactions were done using Taq DNA polymerase (Go Taq®, Promega). PCR reactions were routinely carried out in 50µl volumes of GoTaq Flexi® buffer containing 2.5mM $MgCl_2$, 200µM dNTP, 1µl of genomic DNA (~ 1µg) and 5 pmol of each oligodeoxyribonucleotide primer after the addition of 0.5 units of Taq polymerase. After denaturation for 2 minutes at 95°C, PCR reactions were routinely performed for 30 cycles, involving a 30 second incubation at 95°C, annealing at 45°C for 1 minute and primer extension at 72°C for 3 minutes. An extended annealing step including an additional incubation at 50°C for 1 minute was found to increase to yield of amplicon for some reactions. Upon completion, 1µl of the PCR mixture was analysed by agarose gel electrophoresis.

2.2.5.1: Purification of PCR products

PCR amplicons and restricted DNA fragments were purified from primers and oligonucleotides as required, using a commercially available PCR clean up kit (omega bio-tek). DNA purification was performed according to the protocol provided by the manufacturer.

2.2.6: DNA restriction digestion

Restriction digests with commercial enzymes were conducted according to protocol provided by the manufacturer. Typically, 100-500ng of plasmid was digested in a reaction volume of 20-50µl. For preparative cloning purposes, several micrograms of plasmid were digested. Restriction digestions were routinely analysed by agarose gel electrophoresis.

2.2.6.1: Dephosphorylation of digested DNA

5' phosphate groups were removed from restricted DNA for molecular cloning, using shrimp alkaline phosphatase (New England Biolabs). 1µl (1U) of the enzyme was added to a 50µl DNA restriction digest mixture. Following incubation at 37°C for 60 min, the enzyme was inactivated by heat denaturation at 80°C for 5 min. The dephosphorylated vector fragment was then purified by agarose gel electrophoresis.

2.2.6.2: Klenow fragment treatment of DNA

The Klenow fragment of DNA polymerase I from *E. coli* was used to fill in receded 3' ends of restriction digested DNA for subsequent blunt end ligation. After restriction digestion of DNA (Section 2.2.6), 1µl (10U) of Klenow fragment (Fermentas) and dNTPs (final concentration,200µM) was added and incubated at 37°C for 10 minutes. The enzyme was heat-inactivated at 65°C for 5 minutes and the DNA was then purified by agarose gel electrophoresis.

2.2.6.3: T4 polymerase treatment of DNA

To remove 3' overhangs of restricted DNA for blunt end ligation, T4 DNA polymerase was used. After restriction digestion, 1µl (10U) of T4 polymerase was added and the mixture was incubated at 37°C for a further 10 minutes. dNTPs were then added to a final concentration of 200µM to fill in any recessed 3' ends. The enzyme was then heat-inactivated by incubation at 65°C for 5 minutes and the DNA was purified by agarose gel electrophoresis.

2.2.7: DNA ligation

DNA ligations were conducted using T4 DNA ligase (Promega). Typical ligation reactions comprised 8µl of both the gel-purified, vector and insert DNA fragments, 2µl 10X T4 ligase buffer and 1µl T4 DNA ligase. Reaction mixtures were incubated at 15°C overnight. The following day, a further 1µl of T4 DNA ligase was added and the reaction mixture was incubated for further 4-6 hours. 10µl of the reaction mixture was subsequently transformed into an aliquot of competent *E. coli* DH5α cells.

2.2.8: Agarose gel electrophoresis preparation

1% (w/v) agarose gels containing ethidium bromide (final concentration, 0.1µg/ml) were used to resolve mixtures of DNA fragments by electrophoresis in 0.5 x TBE buffer. Appropriate volumes of DNA were mixed with 6x loading dye and the samples were resolved adjacent to 500ng of a set of DNA molecular weight markers (1kb ladder, Thermo Scientific). Resolved DNA was visualized using a G:Box iChemi XL gel documentation system (Syngene) and digital image files were processed using ImageJ64 (NIH, Bethesda).

2.2.9: Purification of DNA from agarose gels

Following gel electrophoresis, resolved DNA fragments were visualised under low intensity long wavelength UV light and slices containing the DNA were excised from the agarose gel using a scalpel blade. DNA was recovered from the gel slices using a commercially available gel extraction kit (Omega Biotech) according to the manufacturer's protocol. DNA retained on the spin column was eluted in 30µl of the provided elution buffer and 1µl of the purified DNA was routinely analysed by agarose gel electrophoresis (Section 2.2.8).

2.2.10: Site-Directed Mutagenesis (SDM)

Nucleotide substitutions were introduced into plasmids using the Quikchange® Site Directed Mutagenesis Kit (Stratagene). Two complementary oligonucleotides including the desired mutation were designed for each mutant to have a minimal melting temperature of 78°C and a maximal length of 50 nucleotides, as described in the manufacturer's instructions. SDM reactions were carried out essentially according to the manufacturer's instructions

except that PCR reactions were performed for 30 cycles and in 25µl reaction volumes. Specifically, the reaction mixtures contained 1 x reaction buffer, 0.5µl dNTP mix, 1µl plasmid (50-100ng) DNA, 5pmol of each oligodeoxy ribonucleotide primer and 0.5U Pfu DNA polymerase. After an initial heat denaturation at 95°C for 2 minutes, PCR was performed for 30 cycles of denaturation at 95°C for 30 seconds, primer annealing at 55°C for 1 minute and primer extension at 68°C for 12 minutes. A final extension round of 10 minutes was included to ensure that the amplicons were full-length. Subsequently, the reaction mixture was incubated with 1µl Dpn1 (20 units) at 37°C for 1 hour to digest the methylated and hemimethylated parental plasmid. A 15µl aliquot was used to transform competent *E. coli* cells, as described in section 2.2.2.

2.2.11: Yeast transformation

2.2.11.1: Colony transformation

Yeast transformations were performed using lithium acetate (Gietz et al., 1992). A number of small single colonies were suspended in 1ml of TE buffer, followed by centrifugation at 13,000 g for 1 minute. Cell pellets were then resuspended in LiT buffer and 50µl aliquots were added to 1-2 µg of plasmid DNA and 5µl herring sperm DNA (10mg/ml, Roche). After incubation for 30 minutes, 100µl of freshly prepared and sterile-filtered 40% PEG in 1X LiT buffer was added. The mixture was incubated at room temperature for a further 30 minutes. Subsequently, 15µl DMSO (Sigma) was added, and the cells were subjected to heat shock for 15 min at 42°C. The mixtures were centrifuged for 30 secconds and the cell pellets were resuspended in 1ml TE buffer. After recentrifugation, the cell pellet was resuspended in 100µl TE buffer and the cells were plated onto selective solid growth medium.

2.2.11.2: High-efficiency yeast transformation

A high-efficiency yeast transformation procedure (Gietz et al., 1992) was used for transformation of plasmids that had been modified by SDM. Overnight cultures of the appropriate yeast strain were used to inoculate 50ml of pre-warmed (30°C) liquid media to a starting OD_{600nm} of ~0.1. The cells were cultured with constant shaking and the OD monitored until it reached an

OD_{600nm} of 0.5-1. Cells were then harvested by centrifugation at 3,200xg for 5 minutes. Cells were resuspended in 5ml TE buffer, recentrifuged and then taken up in 5ml LiT buffer. The cells were centrifuged as before and then resuspended in 0.5ml LiT buffer. Aliquots of 100µl were used per transformation with ~1µg transforming DNA, 50 µg carrier DNA and 300µl sterilised freshly prepared PEG (40% w/v in LiT buffer). The mixtures were incubated at room temperature for 30 minutes, after which time 50µl DMSO was added and the cells were subject to heat shock at 42°C for 15 minutes. Mixtures were then centrifuged for 30 seconds and the cell pellets were resuspended in 1ml 1X TE buffer. After an additional round of centrifugation, the cell pellets were resuspended in 100µl TE buffer and plated onto appropriate selective media.

2.2.12: Storage of yeast strains

S. cerevisiae strains were grown in selective media at 30°C, unless otherwise stated. Strains were stored for days or weeks on growth medium plates at 4°C, as sealed slants on rich medium at 4°C for several months or in 25% glycerol at -80°C for long-term storage.

2.2.13: Growth analyses of yeast strains

Growth of different yeast strains were compared using a standard spot growth assay. Strains were inoculated into 5ml of selective growth medium and grown to saturation over 1-2 days. The cultures were normalised for their OD_{600nm} values and ten-fold serial dilutions were made in sterile H_2O. 4µl of each dilution were spotted onto appropriate solid medium plates, left to dry and the plates were then incubated at 30°C for 2-4 days.

For complementation analyses of the RNA processing defect of the conditional *GAL::UTP3* strain with plasmids encoding the various Utp*3* deletion mutants, 50ml of pre-warmed (30°C) galactose-based selective growth medium was inoculated with an appropriate overnight culture to a starting OD_{600nm} of approximately 0.1. The culture was incubated in a growth cabinet at 30°C until the OD_{600nm} reached a value of ~0.5. Glucose was added to a final

concentration of 2% and the culture was maintained at an OD_{600nm} between ~0.25 and 0.5 during subsequent growth by dilution with glucose-based selective medium. Cells were harvested upon each doubling for protein and RNA analyses. Cultures were maintained for 24 hours.

2.2.14: Yeast protein lysate preparation

2.2. 14 .1: Alkaline lysis protocol

Expression levels of the Utp3 mutants and fusion proteins were analysed by lysing *cells* under denaturing conditions to minimise artefacts arising from protein degradation (Motley et al., 2012). 10 OD_{600nm} of yeast cells were harvested during mid-log growth (at an OD_{600nm} of approximately 1) in selective medium by centrifugation at 3,200xg for 5 minutes. Cell pellets were either stored at -80°C or used immediately by resuspension in 500µl ice-cold yeast alkaline lysis extraction solution (0.2M NaOH, 0.2% β-mercaptoethanol) and incubated on ice for 10 minutes. Protein was recovered by addition of trichloroacetic acid to a final concentration of 10% (v/v), incubation on ice for 10 minutes and centrifugation at 13,000xg for 5 minutes. The pellets were resuspended in 10µl Tris-HCl pH 9.4 and 90 µl 1 x SDS protein loading buffer and denatured for 5 minutes at 95°C. Samples were centrifuged for 1 min at 15,000xg before analysing 10µl (equivalent to the protein from 1 OD_{600nm} of cells) by SDS-PAGE.

2.2. 14.2: Preparation of native yeast cell extract

Yeast cell extracts were prepared under native conditions for gradient analyses and pull-down experiments. Cells were grown in selective media at 30°C to an OD_{600nm} of 1-1.5. Following centrifugation at 4,200xg for 5 minutes, cell pellets were washed with TE buffer and either used immediately or stored at -80°C. Yeast pellets were weighed and (assuming a density of 1mg/ml) resuspended in an equal volume of extraction buffer. Phenyl methyl sulfonyl fluoride (PMSF, Sigma) was added to the cell suspension to final concentration of 1mM, in addition to 50µl/g cell pellet of a yeast protease inhibitor cocktail (Melford laboratories). An equal volume of glass beads (Sigma) was added and the cells were lysed by vortexing for 5 minutes (10 x 30 second pulses, with a 1 minute incubation on ice between each pulse). Non-lysed cells were

pelleted by centrifugation at 5,000xg for 5 minutes. The lysates were then transferred to clean 1.5ml microfuge tubes and clarified by further centrifugation at 15,000xg for 30 minutes at 4°C.

2.2.14.3: Determination of protein concentration

The concentration of protein was determined spectrophotometrically, either by measuring the direct optical absorbance at 280nm or colorimetrically using a commercial Bradford assay reagent (Bio-Rad) (Bradford, 1976)

2.2.15: Coimunoprecipitation of GFP-tagged proteins

Pull-down experiments were performed on native yeast cell extracts to determine whether the Utp3 mutants were incorporated into complexes containing with components of the processome. Strains were generated that express chromosomally encoded GFP-tagged fusions of Utp6, Utp21 or Utp25 (Life Technologies) (Huh et al., 2003) and that harbour plasmids encoding zz fusions of wild-type or mutant Utp3. Pull-downs were performed using IgG-sepharose beads (GE Healthcare) and strains bearing the pRS416 yeast expression plasmid were used as a control for nonspecific interaction between the GFP fusion proteins and the resin. Native yeast lysates were incubated with 250µl pre-washed IgG-sepharose beads for 2 hours on a rolling platform at 4°C. The supernatant was collected and the beads were washed 3 times with lysis buffer for 5 minutes. Bound proteins were eluted either by addition of an equal volume of 2 x SDS-PAGE loading dye and heating to 90°C for 5 minutes, or by the addition of 0.5M acetic acid. Acetic acid eluates were lyophilised in a SpeedVac™ concentrator (Savant, Thermofisher). The dried pellets were dissolved in 100µl SDS-PAGE loading dye.

2.2.16: Sucrose gradient ultracentrifugation

Cell extracts from strains expressing either a chromosomally encoded HTP-Utp3 fusion or a plasmid-borne zz-tagged Utp*3* construct were subjected to sucrose gradient ultracentrifugation. 12ml 10-50% sucrose gradients in SW41 tubes (Beckman-Coulter) were prepared in TMN150 buffer (10mM Tris-HCl pH 7.6, 150mM NaCl, 5mM $MgCl_2$) by underlaying, using a Beckman 385 Gradient former (Beckman-Coulter). Pairs of gradients were balanced for weight by the

addition of 10% sucrose/TMN 150 buffer to the lighter gradient. 200-500µl of cell lysate was carefully loaded on to the top of the gradient. Ultracentrifugation was performed in a Beckman-Optima LE-80X ultracentrifuge using a SW41 rotor at 36,000xg and a run time of 200 min at 4°C. 18 fractions of 680µl were collected manually from the top to the bottom of the gradients using a Gilson pipette. 50µl aliquots of each gradient fraction were analysed by SDS-PAGE and by Western blotting. The distribution of total cellular protein through each gradient was analysed by resolving aliquots through 10% and 15% SDS-PAGE gels and staining using a colloidal Coomassie blue stain. Western blot analyses of the Utp3 and GFP fusion proteins were performed using appropriate antibodies after electrophoresis of the gradient fractions through 10% SDS-PAGE gels.

2.2.16.1: RNA recovery from sucrose gradient fractions

To remove protein from the sucrose gradient fractions and recover the RNA, sodium acetate buffer pH 5 was added to a final concentration of 300mM and the samples were extracted with one volume of phenol/chloroform by vortexing briefly, followed by centrifugation for 5 min at 15,300xg. The aqueous layer was transferred to fresh microfuge tube and re-extracted with a second volume of phenol/chloroform. The aqueous layer from the second extraction was added to 1 ml ice-cold 100% ethanol and left at -20°C for 2 hours. The precipitated RNA was pelleted by centrifugation at 15,300xg for 20 min, washed with ice-cold 70% (v/v) ethanol and air-dried. The RNA pellets were resuspended in 20µl DEPC-treated water. Aliquots from each gradient fraction were mixed with an equal volume of 2 X RNA loading dye (see Table 2.8), denatured by heating at 65°C for 5 minutes, resolved by PAGE through acrylamide/urea gels (see 2.2.19) and analysed by Northern blot hybridisation.

2.2.17: SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) analysis of proteins

Proteins were analysed according to their electrophoretic mobility through sodium dodecylsulphate polyacrylamide gels (SDS-PAGE) (Shapiro et al., 1967). SDS-PAGE gels were made with stacking gels containing 4% acrylamide and a resolving gel containing acrylamide concentrations ranging

from 8-16% (v/v), depending upon the size of protein under analysis. Protein samples were heat-denatured in protein loading buffer (Table 2.8) at 90°C for 10 min and then resolved by SDS-PAGE in 1 X TGS buffer using the Mini-Protean® Cell electrophoresis system (Bio-Rad). Proteins were resolved through the stacking gel at 100V and then the voltage was increased to 150 volts, once the dye front had reached the resolving gel. Gels were run until the dye-front had reached the bottom of the gel plate. To determine the molecular weight of the resolved proteins, molecular weight markers (Precision Plus Protein All Blue Standard, Bio-Rad) were resolved in parallel. SDS-PAGE gels were either stained to visualise protein content or subjected to Western blot analysis (see below).

2.2.17.1: Coomassie staining of SDS-PAGE gels

To visualize proteins separated by SDS-PAGE, gels were stained in 10ml of a colloidal Coomassie blue stain (Instant Blue, Expedeon) for 30 min with gentle rocking at room temperature. Excess stain was removed by washing with water and the gels were photographed using a G:Box iChemi XL gel documentation system (Syngene). Data was collected using the GeneSnap software (Syngene) and images were generated using ImageJ64 (NIH Bethesda).

2.2.17.2: Western blotting

Following separation according to their size by SDS-PAGE, proteins were transferred to Hybond nitrocellulose membranes (Protran^{TM}, GE Healthcare) by electrophoresis in Western blot transfer buffer. Transfer was performed either overnight at 15 V using a HSI, TE Series Transphor Electrophoresis Unit (HSI, Hoefer Scientific Instruments) or for ~1 hour using a semi-dry blotting apparatus (Trans-Blot Turbo Transfer system, Biorad). Following transfer, the membranes were washed briefly in TBS Buffer and stained with Ponceau S solution (Thermo Scientific) in TBS buffer and photographed. The membranes were washed in TBS buffer to remove the Ponceau S stain and then blocked in 1% skimmed milk powder in TBS buffer for 30 min at room temperature. The blocked membranes were washed 3 times for 5 minutes with TBS buffer before addition of the primary antibody (see Table 2.9 for the antibody dilutions

used and appropriate incubation periods). After incubation, unbound primary antibody was removed by three washes for 5 minutes in TBS buffer. The blot was then incubated in TBS buffer containing an appropriate HRP-conjugated secondary antibody. Unbound secondary antibody was then removed by washing, as described above. Antibody/epitope complexes were detected by ECL, using in-house prepared reagents. Equal volumes of ECL solutions 1 and 2 were mixed and applied to the membrane for 1 minute with gentle agitation. Data were captured using a G: BoxiChemi XL gel documentation system and GeneSnap software (Syngene). Images were generated using the ImageJ64 package (NIH, Bethesda). For reiterative staining with a second antibody, the Western blot was washed extensively with TBST (see table 2.8) and the membranes were subsequently incubated with the next antibody.

2.2.18: RNA extraction from yeast

Total RNA extraction from the yeast cells was performed using the hot phenol method (Maniatis et al., 1982), with additional modifications from another study (Tollervey and Mattaj, 1987). A 50ml culture of the appropriate yeast strain was grown in selective medium to an OD_{600nm} of 0.5 and the cells were harvested by centrifugation at 3,200xg for 10 min at 4°C. The yeast cell pellet was either stored at -80°C until required, or processed directly. 0.5ml GTC mix (50mM Tris-HCl pH 8, 4M GTC, 10mM EDTA pH 8, 2% sarskosyl, 1% β-mercatoethanol) and 0.5ml phenol pH 4.0 were added to the cell pellet and the cells were lysed by vortexing for 5 minutes in the presence of 1 ml of DEPC-treated glass beads (425-600 µm, Sigma). A further 1.5ml of phenol was then added and the contents of the tube were mixed by briefly vortexing. The samples were incubated at 65°C for 10 min. After cooling on ice for 5 minutes, 2ml chloroform and 1ml sodium acetate mix (10mM Tris-HCl pH 8, 100mM NaAc pH 5, 1mM EDTA) were added and the mixtures were vortexed briefly. The organic and aqueous layers were separated by centrifugation at 3,200xg for 5 minutes at 4°C. The aqueous layer was then transferred to a new tube and re-extracted with the addition of 2.5ml of phenol/chloroform pH 4.5, followed by centrifugation for 5 minutes. The aqueous layer containing RNA was transferred into a fresh tube and mixed with 2 volumes of 100% ethanol. The RNA was precipitated at -80°C for 1-16h and pelleted by centrifugation at

3,200xg for 30 minutes. The RNA pellet was washed twice with 70% (v/v) ethanol and air dried for 10-20 min at room temperature. The RNA pellet was resuspended in a final volume of 100µl DEPC-treated H_2O. 5µl of the RNA was diluted to 800µl and the optical absorbance at 260nm was measured. The concentration of RNA in the purified samples was calculated, assuming an A_{260nm} reading of 1.0 is equivalent to 40µg/ml. RNA samples were analysed by electrophoresis under denaturing conditions, using polyacrylamide agarose gels.

2.2.19: Acrylamide gel electrophoresis of RNA

5µg aliquots of total cellular RNA were resolved through 8% polyacrylamide (19:1), 50% (w/v) urea gels containing 0.5 X TBE buffer. RNA samples were diluted to a volume of 8µl in RNA gel loading solution and denatured by heating at 65°C for 5 minutes. Samples were loaded on 20 x 20 cm gels and electrophoresis was performed at 65V in 0.5 X TBE buffer using an EV200 Large Format PAGE Unit Gel Unit (Engineering and Design Plastics Ltd., Cambridge). Electrophoresis was performed until the xylene cyanol dye front had reached the bottom of the gel plate. The gel was stained with ethidium bromide (0.1µg/ml in 0.5 X TBE buffer) for 10-20 min with gentle agitation and the RNA was visualized using a G: Box iChemi XL gel documentation system and Genesnap software (Syngene). The RNA was then transferred to Hybond-N^+ membranes (GE Healthcare) by electrophoresis in 0.5 X TBE buffer at 10 volts overnight using an MWB2 electrotransfer unit (Engineering and Design Plastics Ltd, Cambridge). After transfer, the RNA was cross-linked to the membrane by irradiation with 254nm UV light at a dose of 120 mjoules/cm^2, using a CL-1000 UV crosslinker (UVP, Cambridge). Immediately prior to hybridisation, crosslinked membranes were soaked in hybridisation buffer (6 X SSPE, 5 X Denhardt's solution, 0.2% SDS) and incubated at 37°C for at least 30 minutes.

2.2.20: Agarose gel electrophoresis of RNA

For analysis of large rRNA species, total cellular RNA was fractionated on 25 cm long, 1.2% agarose gels. 10µg of total RNA was lyophilized in a Savant SpeedVac SC100A PlusConcentrator (Thermo Scientific). The RNA pellet was

re-suspended in DEPC-treated water and glyoxal loading dye, and the RNA was denatured for 1 hour at 55°C. After loading the samples, the agarose gel was run in 1 X MOPS buffer until the bromophneol blue dye had travelled ~ 16 cm. After electrophoresis, the gel was soaked for 15 minutes in 60mM NaOH, followed by 15 minutes in neutralization buffer and 15 minutes in 2 X SSPE buffer. The RNA was transferred onto Hybond-N^+ membranes (GE Healthcare) by downward capillary action overnight in 2 X SSPE buffer. After transfer, the RNA was cross-linked to the Hybond membrane as described above. Agarose gel blots were stained with methylene blue (0.2% methylene blue in 0.5M sodium acetate) for 15 minutes and destained by briefly rinsing with Millipore-filtered water. Agarose gel northern blots were pre-hybridised in buffer containing 6 x SSPE, 5 x Denhardt's solution, 0.2% SDS, as described above.

2.2.21: Southern blotting

DNA fragments were separated by agarose gel electrophoresis in 0.5 X TBE buffer. After visualization of the DNA by ethidium bromide staining using a G: BoxiChemi XL gel documentation system, the gels were soaked consecutively for 15 minutes in 0.4M NaOH, neutralization buffer and 10 X SSPE buffer. The DNA was then transferred to Hybond N^+ membranes (GE Healthcare) by downward capillary action. After transfer, DNA was cross-linked to the membrane using a CL-1000 UV crosslinker and pre-hybridised, as described above.

2.2.22: 5' end radiolabelling of oligonucleotides

Oligodeoxyribonucleotides (Eurofins Genomics) were labeled at their 5' hydroxyl group with ^{32}P using T4 polynucleotide kinase and [^{32}P]-ATP (PerkinElmer, Mass, USA) (Maniatis et al., 1982). Typically, 5pmol of DNA oligonucleotide were incubated with 5 units of polynucleotide kinase and ~ 6 pmoles of γ[^{32}P]-ATP (specific activity, 6000Ci/mmol) at 37°C for 30 min. The enzyme was then heat-inactivated by incubation at 65°C for 5 min and the mixture was filtered through a 0.2µm filter (Millipore) directly into the pre-hybridisation mixture.

2.2.23 Hybridisation of Northern and Southern blots

Northern and Southern blots were hybridised with 5' [^{32}P]-labelled oligonucleotide probes in a minimal volume (30-50ml) of hybridisation buffer (6 x SSPE, 5 x Denhardt's solution, 0.2% SDS) overnight at 37°C. After hybridisation, the hybridisation probe was drained from the blot and stored in a Falcon tube for further use, if required. The blot was then rinsed briefly three times with 6 X SSPE buffer at room temperature, followed by a 30 minute incubation in 6 x SSPE buffer at 37°C. Excess moisture was removed from the blots by laying them briefly on paper towels and the dried blots were wrapped in Saranwrap. The wrapped blots were then either subjected to autoradiography using MS film (Kodak) or placed under phosphor storage screens, using X-ray film storage cassettes. Images from nonsaturated exposures were obtained by photographic development of the film or by using a Typhon FL 7000 laser scanner (GE Healthcare), respectively.

To remove hybridisation signals from blots for subsequent hybridisation with other probes, ~ 200 ml stripping buffer (0.1 X SSPE, 0.1% SDS) was heated to boiling and poured directly onto the blot. The blot was then incubated in the stripping buffer on a shaking platform at room temperature until it had cooled, at which point it was either rehybridised or stored after being wrapped in Saranwrap.

2.2.24: Expression of recombinant proteins in *E. coli*, purification and investigation

The *E. coli* strain BL21 (DE3) was used for the expression of recombinant yeast proteins. After transformation (see 2.2.2) of the appropriate plasmid, and selection of transformants on solid growth medium containing ampicillin, several colonies were chosen and inoculated into 10ml LB medium containing both ampicillin and chloramphenicol. The culture was incubated with shaking at 37°C overnight. The overnight culture was diluted 100-fold into pre-warmed growth medium containing ampicillin and chloramphenicol and incubated at 37^0C until the OD_{600nm} was approximately 0.5-0.8. Recombinant protein expression was then typically induced by adding isopropyl-β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5-1mM and the cells

were incubated with shaking for further 3 hours at 37°C. Cells expressing recombinant protein were harvested by centrifugation at 5000xg for 5 min at 4C° and stored at -20 °C. Notably, the GST-Utp6 fusion protein was expressed by induction at 25°C.

2.2.24.1: *E. coli* lysate preparation

Stored pellets from 500ml of *E. coli* culture were thawed on ice and resuspended in 5ml H150 lysis buffer (50 mM Tris-HCl pH 7.6 150 mM NaCl) containing lysozyme. Immediately prior to use, PMSF (Sigma) was added to a final concentration of 2mM. Cells were lysed by sonication with a MSE Soniprep 150 Sonicator (MSE UK Ltd) using 6 cycles of a twenty second sonication at amplitude of 16µm, followed by 1 min intervals on ice. The *E. coli* lysates were clarified by centrifugation at 15,000xg for 30 min and either used immediately or stored at -20°C.

2.2.25: Analysis of recombinant protein-protein interactions in vitro

One ml aliquots of clarified lysate from cells expressing GST-and His6-tagged protein were incubated with 50µl pre-washed glutathione-sepharose beads (GE Healthcare) for 2 hours on a rolling platform. The unbound material in the supernatant was removed from the beads following a brief centrifugation. The beads were then washed three times for 10 minutes using cell lysis buffer and the retained proteins were then eluted by boiling in 2 x SDS Loading Dye. Equivalent fractions of the input, bound and unbound material were then analyzed by SDS-PAGE and Western blotting, using antibodies specific to the GST and hexahistidine tags.

2.2.26: Electrophoretic mobility shift assays

Electrophoretic mobility shift assays (EMSAs) were performed according to Revzin (Revzin, 1989), using recombinant proteins that were purified from cell lysates by Ni-NTA affinity chromatography and dialysed into EMSA buffer (10mM Tris-HCl pH 8, 20mM KCl, 2mM $MgCl_2$, 0.1mM EDTA, 6% glycerol) prior to analysis. Two-fold serial dilutions of purified recombinant protein (initial concentration, 10µM) were made in EMSA buffer and 20µl aliquots were mixed

with 50ng of linearised plasmid DNA, yeast tRNA or poly (A) RNA. The mixtures were incubated on ice for 15 min. 5µl loading buffer was then added and bound and unbound fractions of nucleic acid were resolved by electrophoresis through 6% acrylamide gels using 0.5 X TBE buffer. Nucleic acid was visualised by staining the acrylamide gels with ethidium bromide, as described above (see 2.2.19), or by subsequent electrotransfer and hybridisation with ^{32}P-labelled oligonucleotide probes.

2.2.27: Yeast growth for fluorescence microscopy

For GFP fluorescence microscopy, cells were grown in minimal medium supplemented with appropriate amino acids at 30 °C to an OD_{600nm} of ≤0.5-0.7. 1ml aliquots of cells were collected by centrifugation and resuspended in 100µl media. Living cells were mounted on a slide and observed using an inverted fluorescence microscope (Axiovert 200M, Zeiss) fitted with a high intensity Exfo X-cite 120 excitation light source, band pass filters and a high sensitivity digital CCD camera (Orca ER, Hamamatsu Photonics). Image acquisition was carried out using Volocity image analysis software (Perkin Elmer). Fluorescence images were collected as 0.5µm Z stacks using exposures of up to 300ms, merged into one plane using Openlab software (PerkinElmer) and processed further using the ImageJ64 package.

2.2.28: Bioinformatics

Yeast DNA sequences were obtained from the Saccharomyces Genome Database (SGD) (http://www.yeastgenome.org). Bioinformatics tools and software were used extensively throughout this study and are referenced in the text.

Chapter 3 : Mutational analyses reveal critical roles for the N- and C-terminal regions of Utp3

A number of laboratories have identified and isolated the 90S complex components that are essential for early processing events of 35S rRNA precursors using epitope tagging and mass spectrometry techniques. Recent high resolution cryo-EM studies, combined with X-ray crystallography information, have enabled a dynamic 3D picture of assembly of the processome. The data suggest that the 90S RNP complex forms on the 35S pre- rRNA through an ordered assembly of Utp subcomplexes and the U3 snoRNP particle, which form a mould-like structure around and within the developing 40S ribosomal subunit which facilitate a multitude of processing interactions and structural rearrangements. Sequential processing steps and structural changes are thought to be mediated through the changing availability of RNA-RNA, protein-protein and RNA-protein binding sites (Bassler et al., 2001; Dragon et al., 2002; Grandi et al., 2002; TA et al., 2002; Nissan et al., 2002; Perez-Fernandez et al., 2011; Karbstein, 2011).

Utp3 is an essential protein of the 90S processome particle that is required for early cleavage of 35S pre rRNA precursor at site A_0, A_1 and A_2 that leads to the production of 18S rRNA (Dragon et al., 2002). Utp3 is a large protein of approximately 70 kDa containing an N-terminal region that is composed of simple, repetitive, negatively charged sequences and three distinct protein domains: a central Sas10/Utp3 C1D domain, a coiled-coil domain and a C-terminal domain rich in basic residues (Kamakaka and Rine, 1998; Finn et al., 2016). Depletion of Utp3 has the same effect on cell growth and early pre-rRNA processing events as depletion of the core SSU processome component, U3 snoRNA (Hughes and Ares, 1991). A complete understanding of the function of each component of the SSU is crucial in order to uncover and understand the molecular function of the 90S processome complex. As with many factors found to be essential for SSU processome function, there is limited information currently available regarding the molecular functions of Utp3 in ribosome assembly and pre- rRNA processing. Utp3 is not a component of the characterized UtpB, Mpp10 or U3 snoRNP subcomplexes of

the 90S processome. However, it has been shown copurify with the 90S processome and to interact with components of the UtpB and Mpp10 complexes, as well as Utp25, in yeast two hybrid assays (Charette and Baserga, 2010).

In this chapter, I describe experiments that investigate the effect of specific mutations within Utp3 on cell viability, 18S rRNA synthesis and the incorporation of Utp3 into the SSU processome complex. To address the function of the generated *Utp3* alleles, a conditional mutant strain was used that expresses a Utp3 fusion protein bearing a triple HA tag at its N-terminus under the control of the tightly regulated *GAL* promoter (Dragon et al., 2002).

Mutants of the *UTP3* gene were generated on a plasmid and expressed from a constitutive yeast promoter. Strains expressing an essential gene from the *GAL* promoter are typically viable on galactose-based medium but cease to grow upon shift to glucose-based medium, when transcription from the *GAL* promoter is repressed. Upon transformation of the *GAL::UTP3* strain with a plasmid expressing a functional copy of the *UTP3* gene, the transformants will be able to grow on glucose medium. Conversely, transformants bearing a non-functional, plasmid-borne mutant *UTP3* allele will not be able to grow on glucose medium (Figure 3.1).

GAL::UTP3 transformants harbouring *UTP3* alleles were assayed for complementation of the conditional growth phenotype using plate assays on galactose- and glucose-based medium. RNA was isolated from strains grown in galactose and glucose based medium and analysed for defects in 18S rRNA synthesis by agarose gel electrophoresis and northern blot hybridisation. Furthermore, lysates from wild-type strains expressing epitope-tagged mutant Utp3 proteins were resolved by sucrose gradient ultracentrifugation and subjected to Western blot analyses to determine whether the Utp3 proteins were incorporated into processome-sized complexes. Taken together, these data establish that both the N-terminal region of Utp3 and the Sas10 C-terminal domain of the protein are necessary and sufficient for Utp3 function *in vivo*

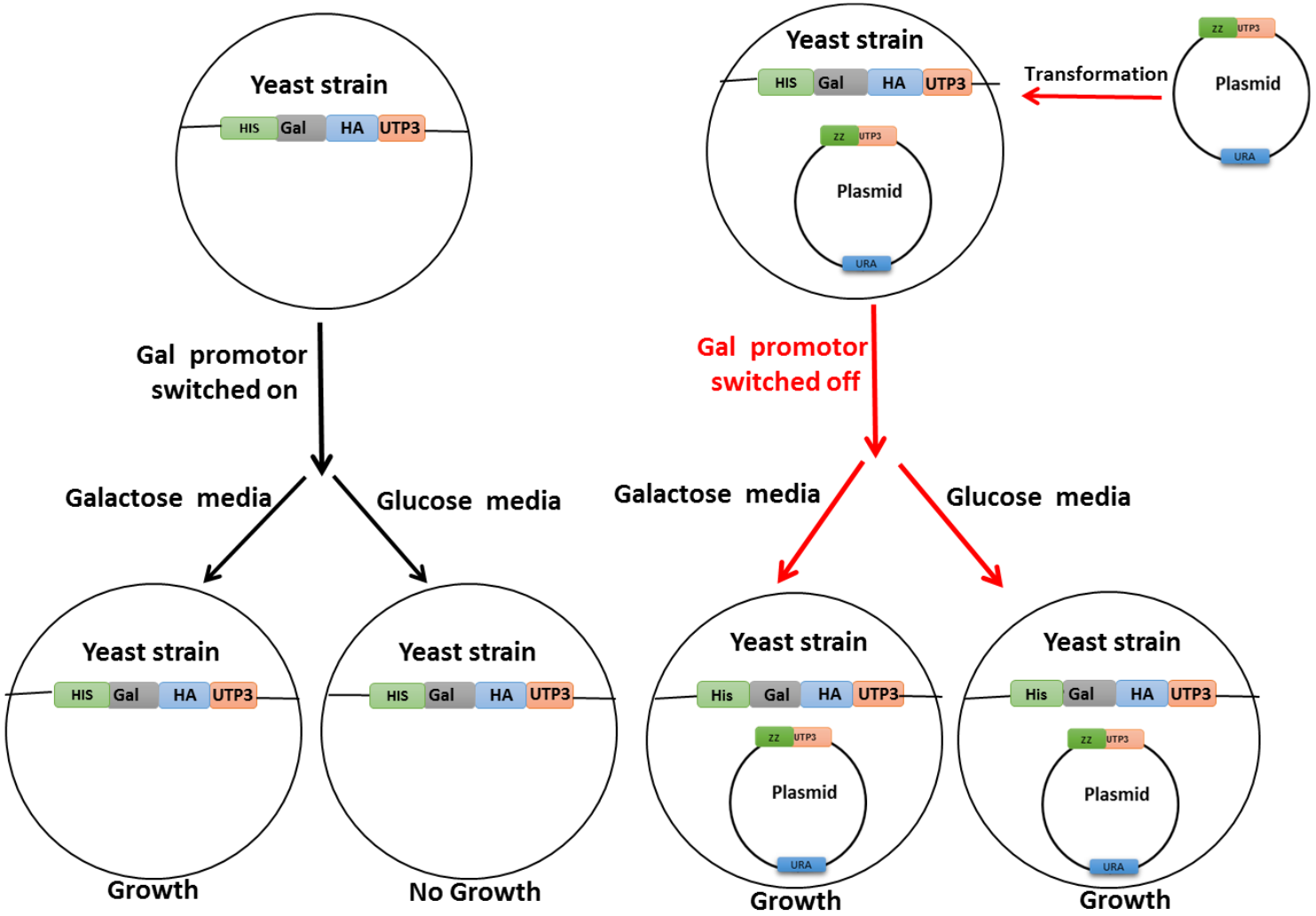

Figure 3.1: Schematic representation of the *GAL::UTP3* growth complementation assay.

The *GAL::UTP3* strain expresses an N-terminal HA-Utp3 fusion protein that is able to grow on galactose medium but is not able to grow on glucose medium. Upon transformation of the *GAL::UTP3* strain with a plasmid expressing a functional copy of the *UTP3* gene, the transformants will be able to grow on glucose medium.

3.1: Analysis of the *GAL::UTP3*

The *GAL::UTP3* yeast strain was kindly provided by the Baserga laboratory (Dragon et al., 2002). Initially, experiments performed to confirm the conditional growth phenotype of the *GAL::UTP3*. Utp3 depletion has been shown to cause a reduction in 18S rRNA levels (Dragon et al., 2002).Since the *GAL::UTP3* strain expresses a triple HA epitope-tagged Utp3 fusion protein, depletion of the protein upon transcriptional repression could be directly assayed by Western blotting.

Initially, growth of the *GAL::UTP3* strain was compared to growth of an isogenic wild-type strain on solid medium containing either galactose or glucose as a carbon source (Fig. 3.2A). The wild-type strain exhibited clear growth on both media within 3 days incubation at 30°C. As expected, growth of the wild type strain was better on glucose medium than on galactose medium. In contrast, the *GAL::UTP3* strain showed growth on galactose medium but had a severe growth defect on glucose medium (Fig. 3.2B). This observation is consistent with the essential nature of the *UTP3* gene (Nehls et al., 1998) also suggests that the expression of the *GAL::UTP3* allele can be tightly repressed in this strain. Therefore, the *GAL::UTP3* strain is potentially suitable for genetic complementation analyses. Comparison of the growth rate of the *GAL::UTP3* with an isogenic wild-type strain during incubation in galactose minimal medium revealed that there was no detectable growth defect observed under permissive conditions (data not shown). Both strains had a doubling time of approximately 6 hours at 30°C. Therefore, expression of Utp3 from the *GAL* promoter under permissive conditions is not noticeably deleterious for the cell growth.

In order to analyse the effect of Utp3 depletion upon cell growth and 18S rRNA synthesis, the *GAL::UTP3* and its isogenic wild-type control strains were grown in galactose minimal media to mid-exponential phase and then shifted to glucose-based media (Fig. 3.2). After addition of glucose, the two strains maintained a similar growth rate for approximately six hours. After six hours, the growth rate of the *GAL::UTP3* was substantially reduced, compared to the wild-type control strain, and effectively ceased after incubation for ~9 hours in

glucose medium. During this time, the OD_{600nm} of the *GAL::UTP3* increased ~6-fold. Conditional mutants that are defective in ribosome synthesis typically show similar growth characteristics, which reflect depletion of ribosome levels below the critical threshold required for progression through the cell cycle (Méreau et al., 1997; Marz and Stadler, 2009; Charette and Gray, 2009). In previously published studies, other SSU processome components such as Utp25, Rrp9, Dhr1 and Dbp4 show a growth defect 6-25 hours after transcriptional repression, with a variable range of phenotypic severity (Venema et al., 2000; Hage and Tollervey, 2004; Charette and Baserga, 2010; Goldfeder and Oliveira, 2010; Choque et al., 2011; Sardana et al., 2015; Soltanieh et al., 2015).

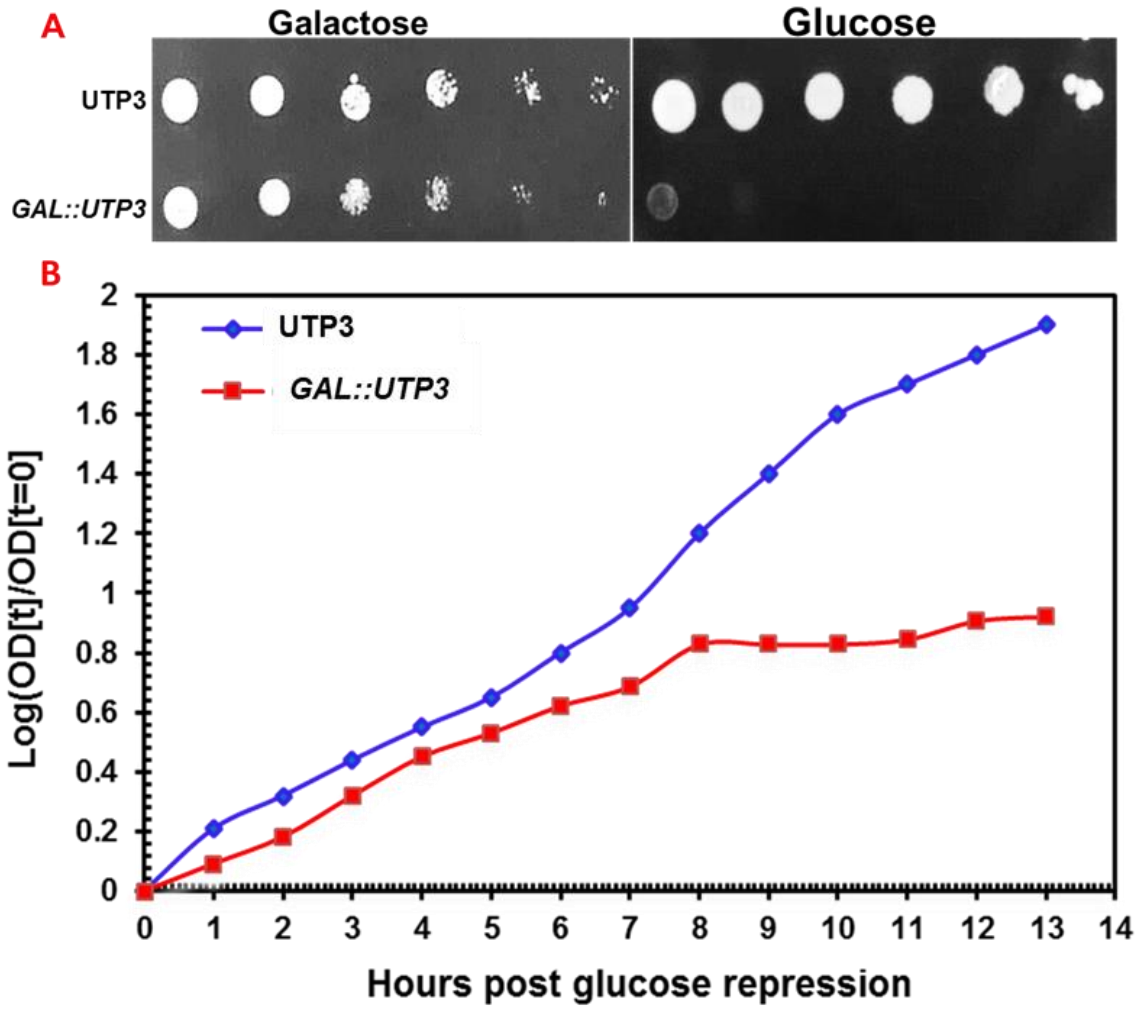

Figure 3.2: Growth of the *GAL::UTP3* strain is inhibited on glucose medium.

A: Spot growth assay of *UTP3* wild-type and *GAL::UTP3* strain. The *GAL::UTP3* (strain 851) and its isogenic wild-type strain (852) were spotted on galactose and glucose-containing solid media. Growth is shown after three days at 30°C.

B: Repression of Utp3 expression severely decreases growth. *GAL::UTP3* and wild-type strains were grown to mid log phase in selective galactose-based medium and then glucose was added to repress the GAL promoter. The increase in OD at $_{600\ nm}$ was measured during incubation at 30°C and the cultures were diluted as necessary to maintain exponential growth. Log_{10} values of the increase in OD (logODt/ODt=0) is plotted against the time after glucose addition.

To follow the depletion of the HA-tagged Utp3 protein upon transcriptional repression, aliquots of the *GAL::UTP3* strain culture were harvested both during growth in galactose medium and after incubation in glucose-based medium for 2, 5, 9 or 24 hours. Total protein was extracted from the harvested cells under alkaline denaturing conditions and subjected to Western blot analysis using an anti-HA antibody. The same blots were also incubated with a Pgk1-specific antibody to indicate the amount of lysate analysed at each time-point. The strong, constitutive Pgk1 promoter is expressed comparably during growth on different fermentable media (Partow et al., 2010). Fig 3.3 shows that the HA-Utp3 signal is severely depleted from the yeast culture after 2 hours and is absent after 5 hours (samples 2 and 3). This indicates a significant loss of Utp3 protein upon repression of the *GAL* promoter by transfer to glucose medium.

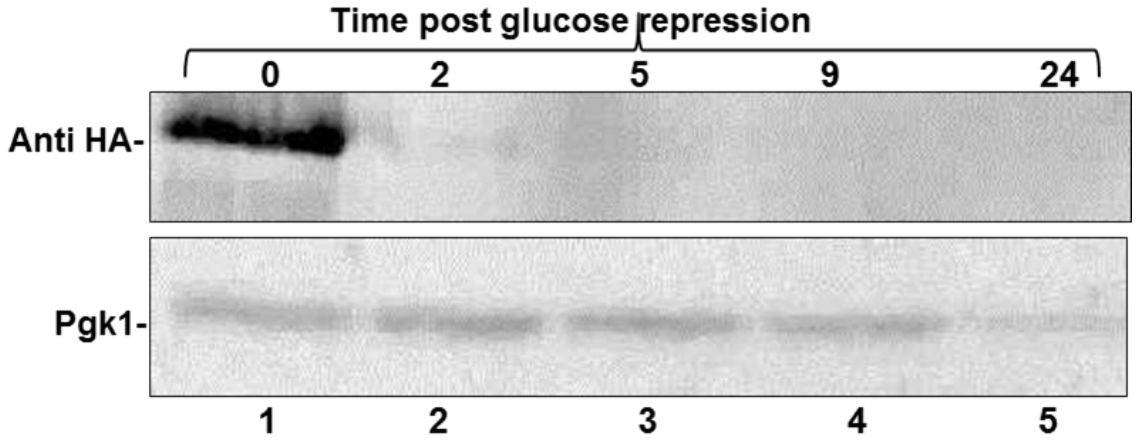

Figure 3.3: Depletion of Utp3 upon transfer of the *GAL::UTP3* to glucose medium.
Total protein was extracted from the *GAL::UTP3* under denaturing conditions after incubation in glucose medium for the times indicated and analysed by Western blot using antibodies against the HA tag. Pgk1 was used as a loading control, as shown in the lower panel.

To analyse the effect of Utp3 depletion on 18S rRNA synthesis, total RNA was isolated from the *GAL::UTP3* during growth in galactose medium and at time-points after transfer to glucose-based medium. RNA was also prepared from the isogenic wild-type control strain during growth in galactose- and glucose-based media. 10µg aliquots of RNA were denatured with glyoxal, resolved by agarose gel electrophoresis and then transferred to Hybond-N^+ nylon

membranes. 18S and 25S rRNAs were detected by methylene blue staining and by subsequent Northern blot hybridisation (Fig. 3.4). The two principal RNA bands detected by methylene bue staining were confirmed to be 18S and 25S rRNA. The relative levels of 18S and 25S rRNA were not altered in the wild-type strain upon transfer from galactose medium to glucose medium (note that the Northern blot hybridisation signal obtained for the 18S rRNA was significantly stronger than that for the 25S rRNA). In contrast, the level of 18S rRNA in the *GAL::UTP3* was depleted upon transfer to glucose-based medium (Fig. 3.4B), whereas 25S rRNA levels were not depleted (the total amount of RNA at the 24-hour time-point is clearly under loaded) (Fig 3.4 B). Quantification of the 18S rRNA:25S rRNA ratio observed in the *GAL::UTP3* strain revealed that 18S rRNA levels were reduced to ~ 60% of the level seen during growth in galactose based medium after 2 hours incubation in the presence of glucose and to ~10% after incubation for 5 hours (Fig. 3.4, panel C). Therefore, 18S rRNA depletion in the *GAL::UTP3* strain occurred after transcription repression and before any detectable effect on cell growth, consistent with the depletion of HA-tagged Utp3 observed by Western analyses (Fig. 3.3).

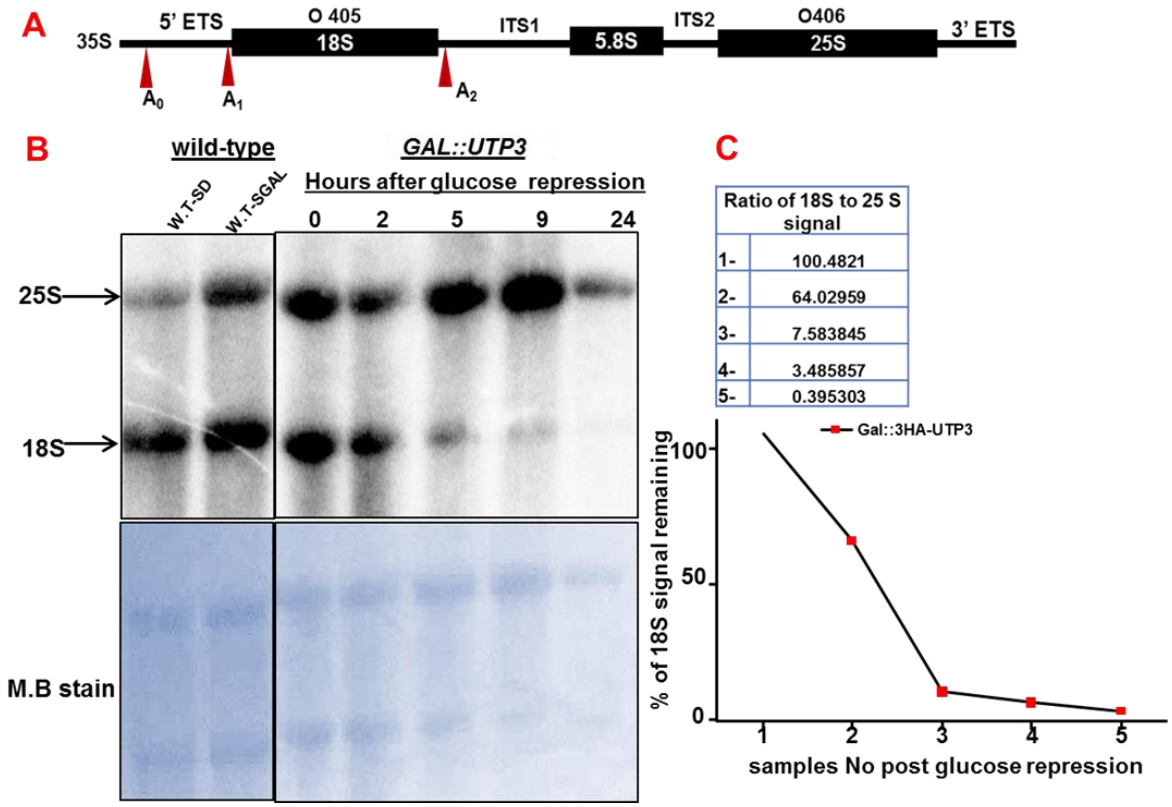

Figure 3.4: Utp3 depletion causes loss of 18S rRNA.

A: Schematic of the pre- rRNA transcript, indicating the relative position of the 18S, 5.8S and 25S rRNA coding regions and the sequences that are complementary to the oligonucleotide probes o405 and o406 that were used for detection of the 18S and 25S rRNAs, respectively.

B: Northern blot analyses of total RNA isolated from the *GAL::UTP3* and wild-type strain during growth in galactose- and glucose-based media. Total RNA was resolved by agarose gel electrophoresis, transferred to a nylon membrane and detected by methylene blue staining (lower panel). 18S and 25S rRNAs were detected by hybridisation with specific oligonucleotide probes (top panel).

C: Plot of the loss of 18S rRNA in the *GAL::UTP3* as a function of time. The 18S rRNA level, relative to 25S rRNA, was plotted for each time-point and is expressed as a percentage of the signal in galactose medium.

3.2: Removal of *EcoR*I site from a plasmid-borne zz-Utp*3* allele

In order to address the dependence upon specific domains within Utp3 for the molecular function of this protein *in vivo*, use was made of an available plasmid encoding an N-terminally epitope-tagged Utp3 fusion protein that contains two tandem copies of the z domain from protein A of *Staphylococcus aureus* (Turner, 2011). Expression of zz-tagged proteins from this construct would enable the expression levels of mutant and wild-type Utp3 proteins to be readily compared. Expression of fusion proteins encoded in this plasmid is driven by constitutive expression from the yeast *RRP4* promoter (Mitchell et al., 1996). The plasmid also harbours a *URA3* selectable marker.

The original zz-Utp3 construct (Turner, 2011) could be used directly to generate C-terminal deletion mutants by site-directed mutagenesis (SDM). The construct contains an *EcoR*I site at the junction between the zz tag and the *UTP3* open reading frame (ORF) that could potentially be exploited to generate epitope-tagged N-terminal deletion mutants of Utp3. However, the plasmid contains two additional *EcoR*I sites within the insert, one within the *UTP3* ORF at position 4088, and another within the 3' UTR of the *UTP3* gene at position 5500 (Fig. 3.5A). In order to use the *EcoR*I site at the 5' end of the *UTP3* ORF, the additional sites were therefore initially removed by site-directed mutagenesis. The *EcoR*I site within the *UTP3* ORF was deleted in a manner that introduced a silent mutation by exchanging the GAA codon for glutamate at residue 226 to GAG.

Site-directed mutagenesis (SDM) was initially performed to remove the *EcoR*I site from the 3' UTR of the *UTP3* gene (see 2.2.10). Plasmid DNA samples from candidate clones were screened for the successful removal of the *EcoR*I site by restriction digestion (Fig. 3.5, panel B). Digestion of the parental plasmid with *EcoR*I released three fragments of ~ 5 kb, 1.4 kb and 700 bp, as expected. Digestion of all three candidate clones demonstrated loss of the 1.4 kb fragment and release of the 700 bp fragment, indicative of correct deletion of the *EcoR*I site within the 3' UTR. Loss of the restriction site at position 5500 was confirmed by sequence analysis of the corresponding region of the plasmid (Fig. 3.5, panel D).

A further round of SDM was performed to remove the *EcoR*I site within the *UTP3* ORF and candidate clones were screened by *EcoR*I digestion of plasmid DNA, as described above. The 700 bp fragment was not observed upon restriction digestion of DNA from the two candidate clones, although the plasmid was linearised, suggesting that the *EcoR*I site at position 4088 had been deleted. Subsequent sequence analysis through the 5' region of the *UTP3* gene confirmed that the site had indeed been deleted (data not shown). The resultant plasmid contains a single *EcoR*I site at the beginning of the *UTP3* ORF and can therefore be used to generate plasmids encoding N-terminal deletion mutants of Utp3 that have an N-terminal zz epitope tag.

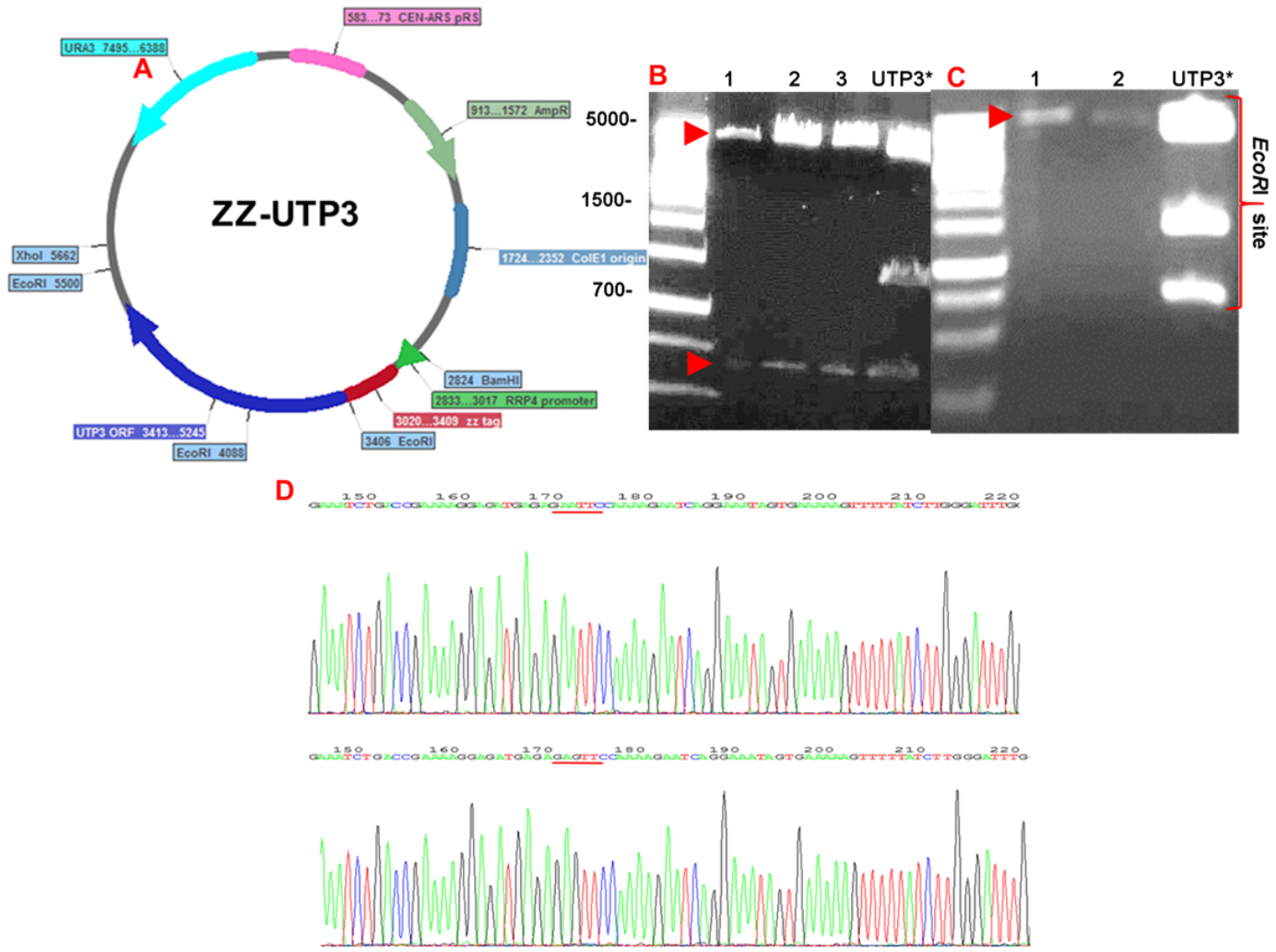

Figure 3.5: Removal of two *EcoR*I sites from the *UTP3* gene.

A: Schematic representation of the zz -UTP3 construct that contains three *EcoR*I sites: one at the junction between the zz tag and the *UTP3* ORF, one within the *UTP3* ORF and one within the *UTP3* 3' UTR.

B, C: Screening of candidates for loss of *EcoR*I sites within the *UTP3* 3' UTR and ORF, respectively. Purified plasmids were digested with *EcoR*I and the incubation mixtures were resolved by agarose gel electrophoresis. Digests of the original construct (denoted *UTP3**) were included as controls. Restriction fragments observed in the positive candidates are indicated with red arrowheads.

D: Sequencing reads of the *UTP3**construct and the derivative after deletion of the *EcoR*I site within the 3' UTR.

3.2.1: Validation of a new zz-Utp3 allele

The construct containing a unique *EcoR*I restriction site at the beginning of the *UTP3* ORF, the original plasmid encoding the genomic *UTP3* sequence and a control vector that lacked DNA derived from the *UTP3* gene were transformed into the conditional *GAL::UTP3* strain (P851) and also an isogenic wild-type strain (P852). The growth phenotypes of the transformants were then tested on media that was either permissive or non-permissive for the *GAL::UTP3*, using a spot growth assay (Fig. 3.6, panel A).

Expression of either the allele containing the genomic *UTP3* sequence (denoted *UTP3**) or the derivative that contains a single *EcoR*I site (denoted *UTP3*) in the *GAL::UTP3* strain allowed comparable growth on glucose-based selective medium (Fig. 3.6, panel A). Transformants of the *GAL::UTP3* harbouring the control vector were not able to grow on glucose-based medium, whereas all transformants showed comparable growth on medium containing galactose. Furthermore, all transformants of the wild-type *UTP3* strain were able to grow on both galactose and glucose medium. These data demonstrate that the plasmid-borne *UTP3* alleles complement the growth phenotype of the *GAL::UTP3*, and that deletion of the two genomically encoded *EcoR*I sites has no obvious detrimental effect on the function of the *UTP3* gene.

To compare the expression levels of the Utp3 proteins under different growth conditions, cell lysates were prepared from transformants that had been grown in galactose and glucose media and the epitope-tagged zz fusion proteins were visualized by SDS-PAGE analysis and Western blotting (Fig. 3.6, panel B). The expression levels of the zz-Utp3 protein were similar to the expression levels of the wild-type *UTP3** allele, suggesting that removal of the *EcoR*I site of *UTP3* gene has no impact on Utp3 protein expression. Strikingly, the expression level of Utp3 was significantly higher in cells grown in glucose medium than those grown in galactose-based medium. This difference may reflect a requirement for higher levels of Utp3 during faster growth in medium containing the more preferable carbon source. Thus, expression of Utp3 may be regulated by cell growth rate.

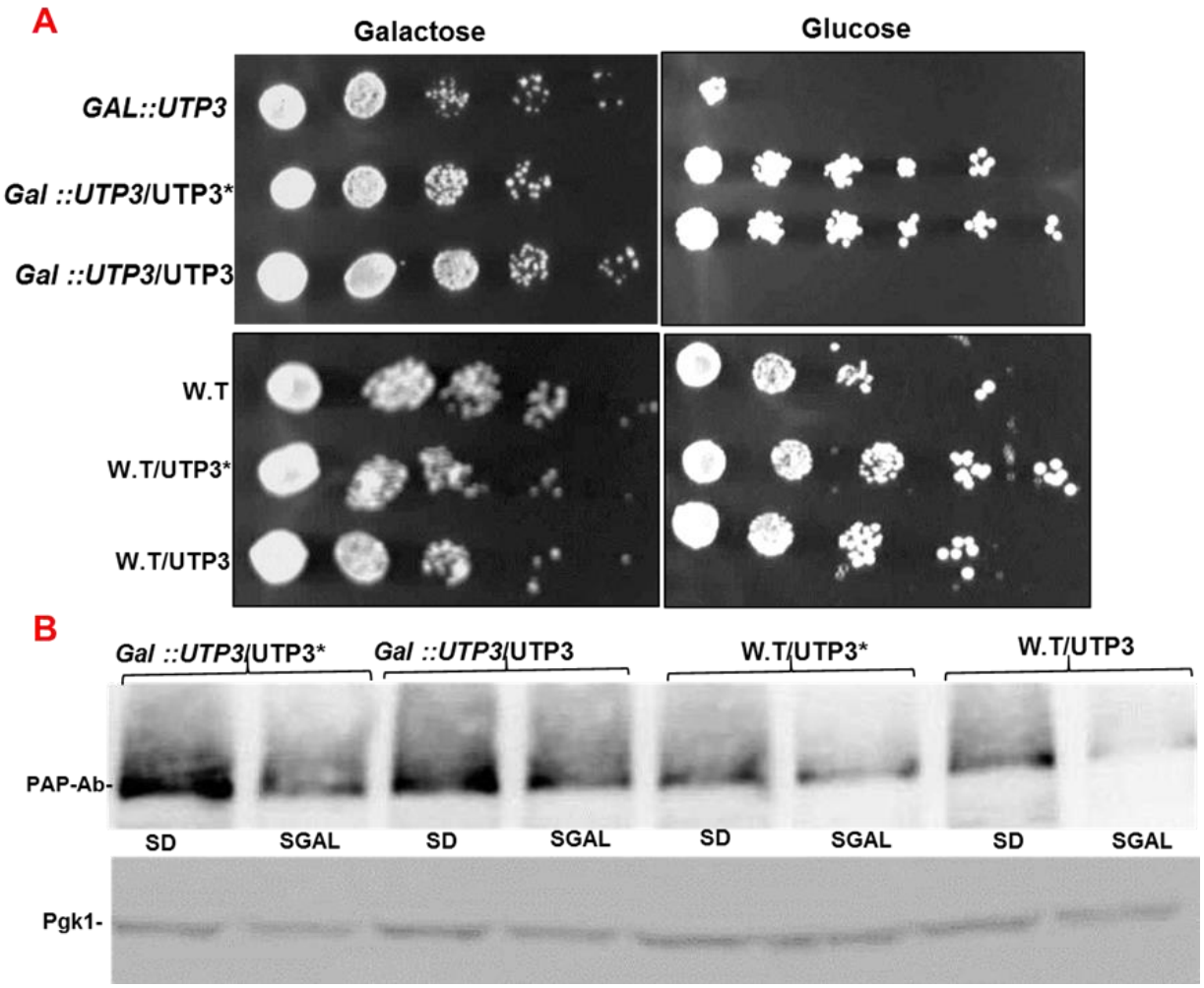

Figure 3.6: Removal of *the EcoR*I sites from UTP3 gene does not affect its ability to complement the *GAL::UTP3* strain or the expression level of Utp3 protein.

A: Ten-fold serial dilutions of the *GAL::UTP3* strain (851) and its isogenic *UTP3* wild type strain (852) harbouring plasmids expressing either the wild-type *UTP3* allele encoding the chromosomal *UTP3* sequence (denoted *UTP3**), the derived allele containing a single *EcoR*I site (denoted *UTP3*), or the cloning vector, were grown on selective minimal media containing either galactose or glucose. Images were taken after growth for two days at 30°C.

B: *GAL::UTP3* and isogenic *UTP3* wild-type strains expressing either the wild-type or mutant zz-Utp3 fusion proteins (denoted Utp3* and Utp3, respectively) were grown in selective galactose- or glucose-based minimal media. Cell lysates were resolved by SDS-PAGE and Utp3 fusion protein expression levels were analysed by Western blotting using the PAP antibody (upper panel). The Pgk1 protein was analysed as a loading control (lower panel).

3.2.1: A comparison of the sedimentation profiles of endogenously expressed HTP-Utp3 and plasmid-encoded zz-Utp3 proteins

90S ribosomal complexes can be resolved from the large majority of cellular protein by sedimentation of cell lysates through sucrose density gradients. The distribution of specific epitope-tagged proteins through sucrose density gradients can be analysed by SDS-PAGE and Western analysis of individual gradient fractions. Sedimentation profiles of Utp3 were compared in strains expressing the protein from its normal chromosomal locus as an N-terminally tagged HTP protein fusion (denoted HTP-Utp3) or as the N-terminally tagged zz protein fusion from a centromeric plasmid (denoted zz-Utp3, see above) (Fig. 3.7). The strain harbouring the *HTP::UTP3* Locus was designed for CRAC analyses of the RNA binding profile of Utp3 and is described in more detail in Chapter 5. Here, the HTP-Utp3 fusion protein was analysed together with the plasmid–encoded protein because its expression is under the control of the homologous *UTP3* promoter and expected to mirror more accurately the expression of the wild-type protein.

Aliquots of each fraction of the sucrose gradients were resolved by SDS-PAGE and the distribution of total cellular protein was visualised by Coomassie blue staining (Fig. 3.7, lower panels). The patterns observed were very similar for lysates from cells expressing either the HTP-Utp3 or the zz-Utp3 proteins, demonstrating reasonable technical reproducibility for the manually prepared gradients. Both profiles reveal prominent heterogeneous protein peak at the top of the gradient, which reflects the majority of cellular protein. Fractions in the middle of the gradient (lanes ~8-13) are resolved into a number of clearly defined bands corresponding to low molecular weight proteins. These fractions contain ribosomal complexes, given their sedimentation behaviour through sucrose gradients and the molecular weight profiles of the proteins.

The sedimentation profiles of HTP-Utp3 and zz-Utp3 fusion proteins were analysed in the same gradients by SDS-PAGE and Western blotting. HTP-Utp3 was observed throughout the lower part of the gradient, broadly reflecting the distribution of the ribosomal complexes, with a small amount in fractions at the top of the gradient. The distribution of the zz-Utp3 protein also reflected the

sedimentation profile of ribosomal complexes, with a clear peak in fractions 9 and 10, but it was predominantly found in fractions towards the top of the gradient. This difference might reflect a difference in expression levels of Utp3 seen from the plasmid-borne *RRP4* promoter, compared with that seen upon expression from the *UTP3* promoter. Expression levels of zz-Utp3 are considerably higher than *HTP::UTP3* (see Fig. 3.10, below). It is possible that expression from the *RRP4* promoter generates a surplus of the protein unable to be incorporated into complexes that function in pre- rRNA processing.

RNA was also recovered from each sucrose density gradient fraction of the zz-Utp3 lysate. Ethidium bromide staining of the acrylamide RNA gel clearly resolved fractions rich in tRNA molecules (lanes 1-5) from fractions containing the 5S and 5.8S ribosomal RNAs (lanes 10-14 and, to a lesser extent, fractions further towards the bottom of the gradient) (Fig. 3.8, panel A). The presence of rRNAs predominantly in lanes 10-14 is consistent with the observation of low molecular, ribosomal proteins observed upon SDS-PAGE analysis (Fig. 3.7) To analyse the distribution of U3 snoRNA, the RNA was transferred from the acrylamide gel to a nylon membrane and subjected to Northern blot hybridisation analyses (Fig. 3.8, panel B). U3 snoRNA was widely distributed between fractions 10-17, consistent with its association with the large 90S SSU processome complex. U3 snoRNA was also observed in fractions 4-5 and 7-8, indicative of its presence in smaller complexes. 5S and 5.8S rRNA were observed upon Northern hybridisation analyses in higher density fractions (lanes 11-18), as expected from the pattern of the ethidium stained gel. Notably, Northern hybridisation using a probe specific to $tRNA^{Phe}$ detected two resolved bands. The larger fragment of $tRNA^{Phe}$ may represent unspliced tRNA or cross-hybridisation with an unknown RNA.

In conclusion, protein and RNA analyses of the sucrose density gradient fractions supports the conclusion that Utp3 cosediments with pre-ribosomal complexes.

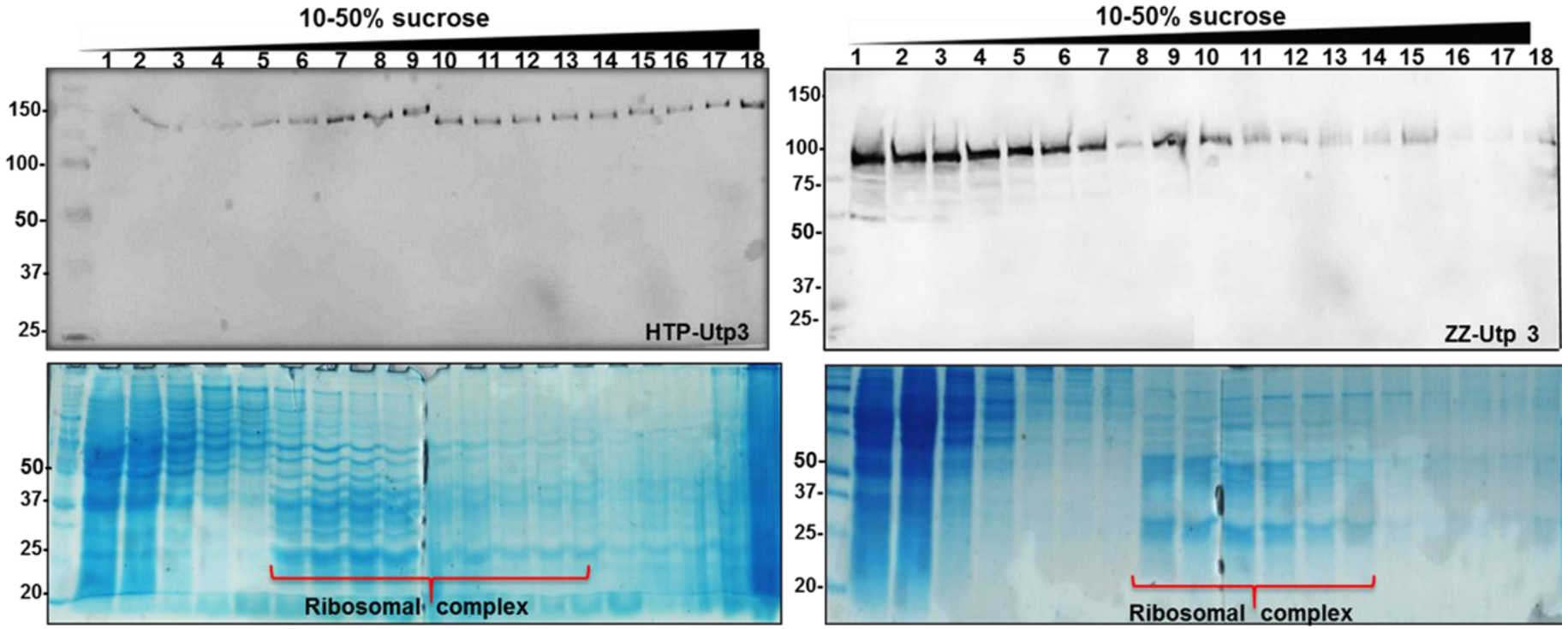

Figure 3.7: Both chromosomally encoded HTP-Utp3 and plasmid-encoded zz-Utp3 cosediment with pre ribosomal particles

Lysates from strains expressing chromosomally encoded Utp3 or plasmid-encoded zz-Utp3 were fractionated through 10-50% sucrose gradients. Fractions were collected from each gradient, aliquots were resolved through SDS-PAGE gels and the proteins visualised by staining with Coomassie blue (lower panels) or transferred to nitrocellulose membranes for Western analyses using the PAP antibody (upper panels). The electrophoretic migration of protein molecular weight markers is indicated (size is indicated in kDa).

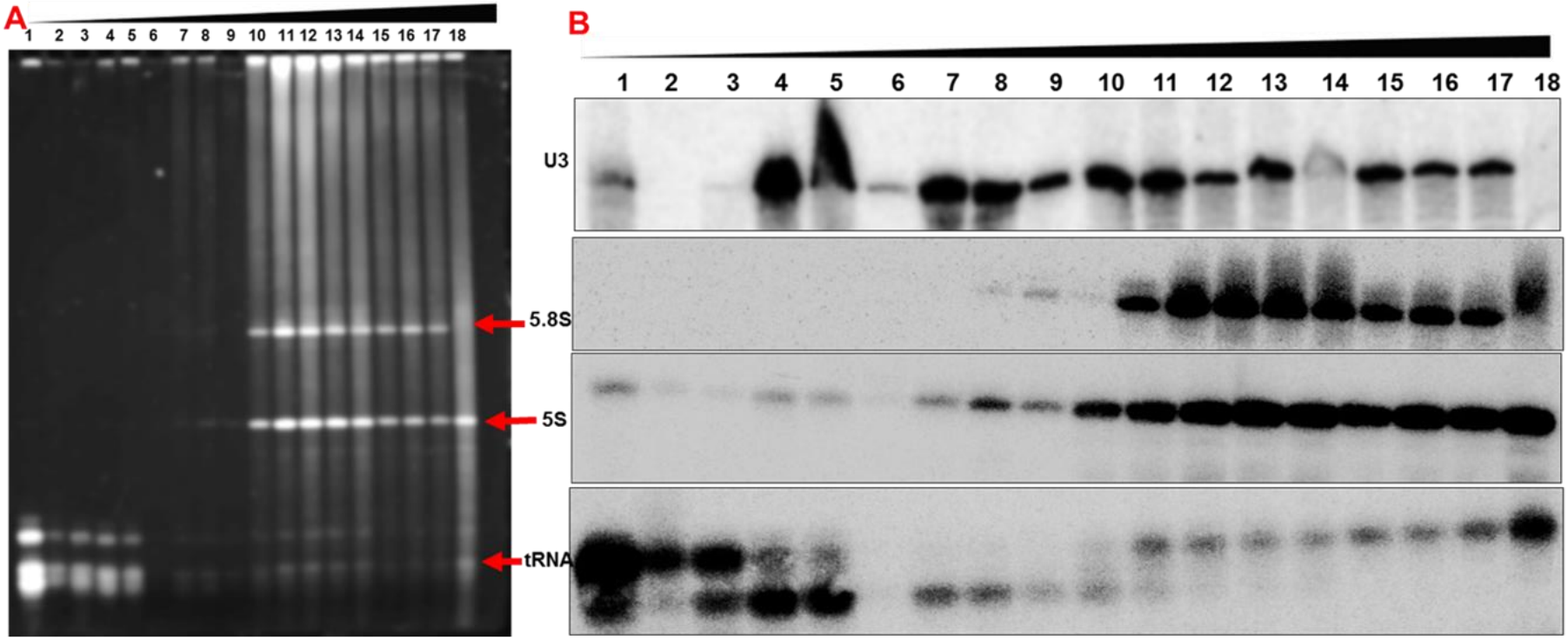

Figure 3.8: RNA analyses of cell lysate from a strain expressing zz-Utp3 after fractionation through a 10-50% sucrose gradient

RNA was recovered from sucrose gradient fractions of cell lysate from a strain expressing zz-Utp3 and resolved through an 8% polyacrylamide gel.

A: Image of the ethidium stained acrylamide gel, showing the differential migration of RNA through the gradient. The position of the major RNA species are indicated.

B: RNA was transferred to nylon membranes and analysed by reiterative Northern blot hybridisation, using oligonucleotide probes complementary to U3 snoRNA, 5S rRNA, 5.8S rRNA and $tRNA^{Phe}$.

3.3: Deletion analysis of Utp3 function

Utp3 is composed of 610 amino acid residues, with a predicted molecular weight of 70,259 Daltons, as inferred from the yeast genome sequence. Utp3 contains an N-terminal region containing repetitive, negatively charged sequences. The adjacent C1D domain encompasses residues 219-298 (Staub et al., 2004). Amino acids 344-364 constitute a coiled-coil domain, while the Sas10/Utp3 C-terminal domain (CTD) comprises residues 532-608.

There is currently little published data available on the characterisation and overall functions of Utp3, beyond that it is involved in 18S processing (Dragon et al., 2002) and an early report that over-expression of the protein has an impact on the expression of heterochromatic loci (Kamakaka and Rine, 1998). Moreover, the importance of one or more protein domains within Utp3 in its molecular function has not been reported. One principal aim of this study is to define the regions of Utp3 that are required for its function. A series of N-terminal, C-terminal and internal deletion mutations within the *UTP3* gene were generated, based on the bioinformatically defined features of the protein (see Fig. 3.9, panel A), and the mutants were analysed for their ability to complement the conditional growth phenoytpe of a *GAL::UTP3* strain (Dragon et al., 2002).

As noted above, the N-terminal region of Utp3 has repetitive sequences that are rich in acidic residues. An N-terminal deletion mutant (plasmid p718, denoted Δ N-terminal) was generated by introducing a second in-frame *EcoR*I site (in addition to the site at the 5' end of the *UTP3* ORF) by site-directed mutagenesis at the 5' end of the C1D domain, restriction with *EcoRI* and religation.

Truncation of the complete CTD of Utp3 (at residue G523) leads to loss of function (Turner, 2011). Two shorter C-terminal deletions were generated by site-directed mutagenesis, in which termination codons replace codons for lysine residues at positions 577 (p714, denoted K577X) or 566 (p710, denoted K556X). These mutations were made in order to remove known potential NLS

sequences contained in the C-terminal domain (Kamakaka and Rine, 1998). The original CTD deletion (denoted ΔCTD) was analysed in parallel (Fig.3.9A).

To determine whether these Utp*3* mutants encode functional proteins, the mutant constructs and a plasmid encoding the wild-type Utp3 fusion protein were transformed into the *GAL::UTP3* strain (Strain 851). Transformations were also performed using the cloning vector as a negative control. The transformants were grown up in minimal galactose-based medium and their growth was compared on glucose and galactose-based minimal medium, using a spot growth assay. This assay provides a good indication of growth by comparing the cell mass obtained from a serial dilution of cells that have been normalised by optical density (OD_{600nm}) (Fig. 3.9 B).

After three days' growth at 30°C, growth of all strains was comparable on galactose-based medium lacking uracil. On glucose-based medium, as previously shown (Fig. 3.6), growth of cells harbouring the original wild type zz-Utp3* construct and the derivative containing a single *EcoR*I site (zz-Utp3) was comparable. Consistent with previous work in our laboratory (Turner, 2011), deletion of the CTD prevented cell growth. The K556X mutation was also unable to complement the growth phenotype of the *GAL::UTP3*. The K577X mutant showed noticeably decreased growth, compared to the wild-type alleles. Deletion of the N-terminal region of the protein caused a clear loss of complementation. These data clearly demonstrate that expression of Utp3 lacking either the N-terminal or the C-terminal region is unable to complement the conditional growth phenotype of the *GAL::UTP3*.

The above data suggests that the NTD and CTD regions are required for Utp3 function. To test whether these two regions alone are sufficient for Utp3 function *in vivo*, an additional *UTP3* allele was generated that harboured an internal deletion of the central region of the protein (Fig. 3.9, panel A). In-frame *Cla*I sites were generated at residues 208/209 and 524/525 by site-directed mutagenesis and the intervening sequence was deleted by restriction digestion and ligation. The resulting mutation (plasmid 731, denoted ΔM) removed both the C1D domain and the coiled-coil region, as well as the central region of the

protein, and resulted in the expression of a protein less than half of the size of the wild-type protein. The ΔM mutant supported growth of the *GAL::UTP3* on glucose-based medium, although the growth was significantly weaker than that seen upon expression of the wild-type protein. This result demonstrates that the N–terminal and CTD domains are indeed sufficient to support cell growth.

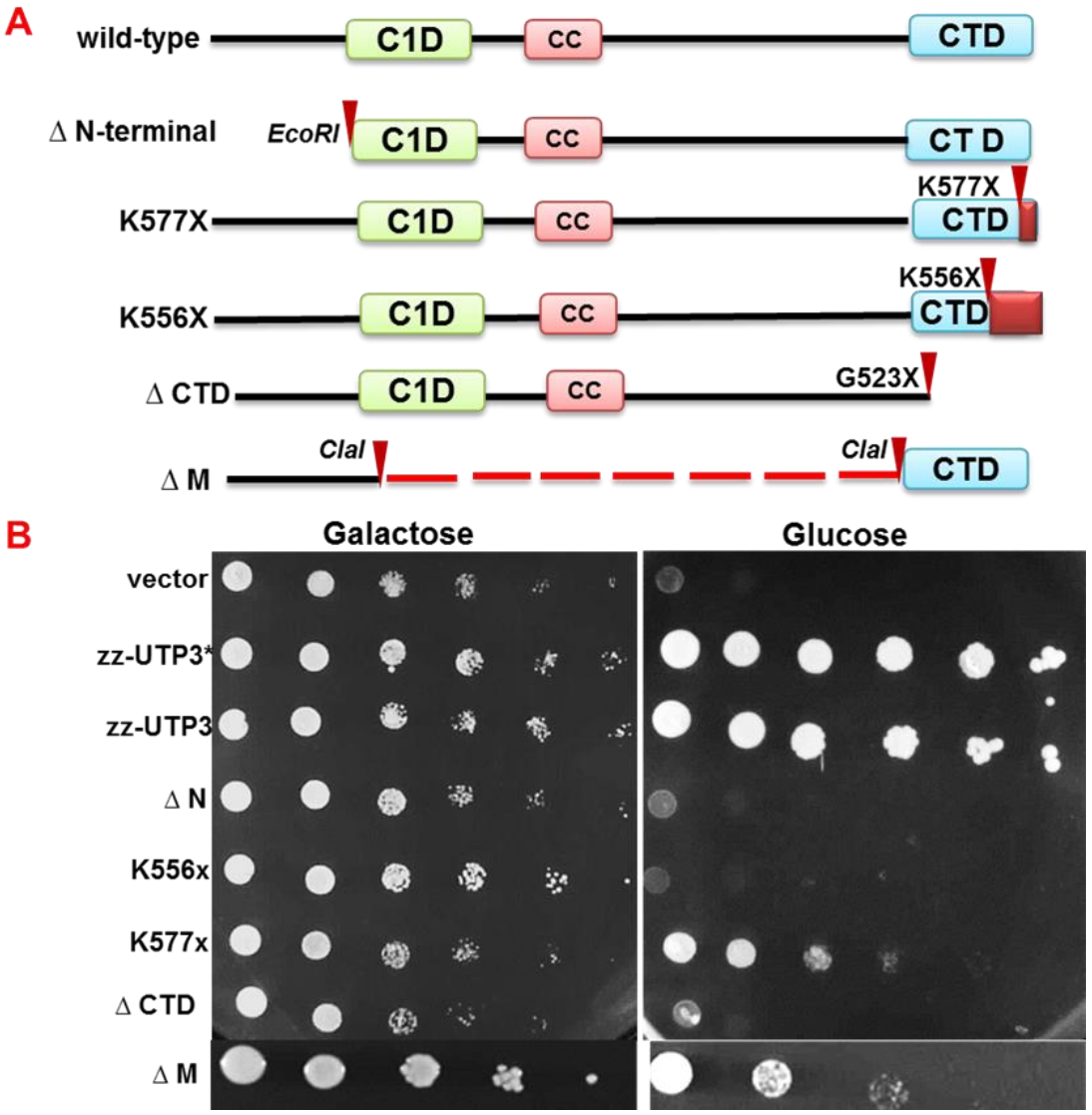

Figure 3.9: Utp3 requires both the CTD and the N-terminal region for function.

A: Schematic representation of the domain organisation of Utp3. The protein contains a C1D domain, a coiled coil region and a C-terminal domain. A series of N-terminal, C-terminal or internal deletion mutants were generated to test the functional requirement of these regions of Utp3. The positions of restriction sites (to generate the N-terminal and internal deletions) or stop codons (to generate the C-terminal deletions) that were introduced by site-directed mutagenesis are indicated by red arrowheads. Portions of the CTD that are not expressed in the deletion mutants are coloured red. The internal deletion is indicated by a broken line.

B: Spot growth assays of *UTP3* mutants. Transformants of the conditional *GAL::UTP3* strain, bearing an empty vector or one of the Utp3 mutants, were grown in galactose minimal medium and normalized by optical density (600nm). 10-fold serial dilutions were spotted on to glucose- and galactose based minimal medium plates. Plates were photographed after incubation at 30°C for 3 days with the exception of the plate assaying the ΔM mutant, which was photographed after 5 days.

3.3.1: Analysis of the expression levels of Utp3 mutants

In order to investigate the relative expression levels of mutant and wild-type Utp3 proteins, yeast strains harbouring each epitope-tagged zz fusion construct were grown in selective minimal medium and cell lysates made under denaturing conditions were analysed by SDS-PAGE and Western blotting. Lysate from a strain expressing epitope-tagged HTP-Utp3 fusion protein (which bears the same zz tag) was analysed in parallel. The *HTP::UTP3* strain expresses Utp3 from the homologous promoter and provides a measure for the level required to support normal cell growth.

A representative Western blot of Utp3 expression levels and a quantitative analysis of biological replicates is shown in Fig. 3.10. The data represents averaged results from two experiments for five mutants, normalized against Pgk1 levels and standardized to the amount of protein in the HTP-Utp3 strain.

It can be seen that the protein expression level of zz-Utp3 is almost twice that of HTP-Utp3. Therefore, expression of Utp3 from the plasmid is considerably higher than from the normal chromosomal locus. This may reflect the relative strengths of the *RRP4* and *UTP3* chromosomal promoters.

The Western blot shows discrete bands for all fusion proteins and the pattern is consistent with all deletion mutants expressing a significant amount of truncated Utp3 of the appropriate molecular weight. The expression levels of the K577X and ΔM mutants were not clearly different from the plasmid-borne wild-type protein, whereas the levels of the K556X, ΔCTD and ΔNTD mutants were consistently reduced. Deletion of the N-terminal region has the greatest effect on protein expression levels. This suggests an important role for the N-terminal region of Utp3 in protein folding/stability.

Interestingly, looking at all six Western blotting results (data not shown), the ΔM mutant showed a consistently increased expression level, compared to the wild-type protein. This increased abundance of the ΔM mutant protein may somehow help to compensate for the lack of the C1D and coiled-coil domains.

Alternatively, it is possible that C1D and coiled-coil regions could be involved in autoregulation of protein levels.

Significantly, all Utp3 mutants are expressed at levels that are higher than the HTP-Utp3 fusion protein. Therefore, the differences in growth phenotypes observed for the *GAL::UTP3* transformants in this study are likely to reflect a decrease in the function of the Utp3 mutants, rather than simply an effect on protein expression levels.

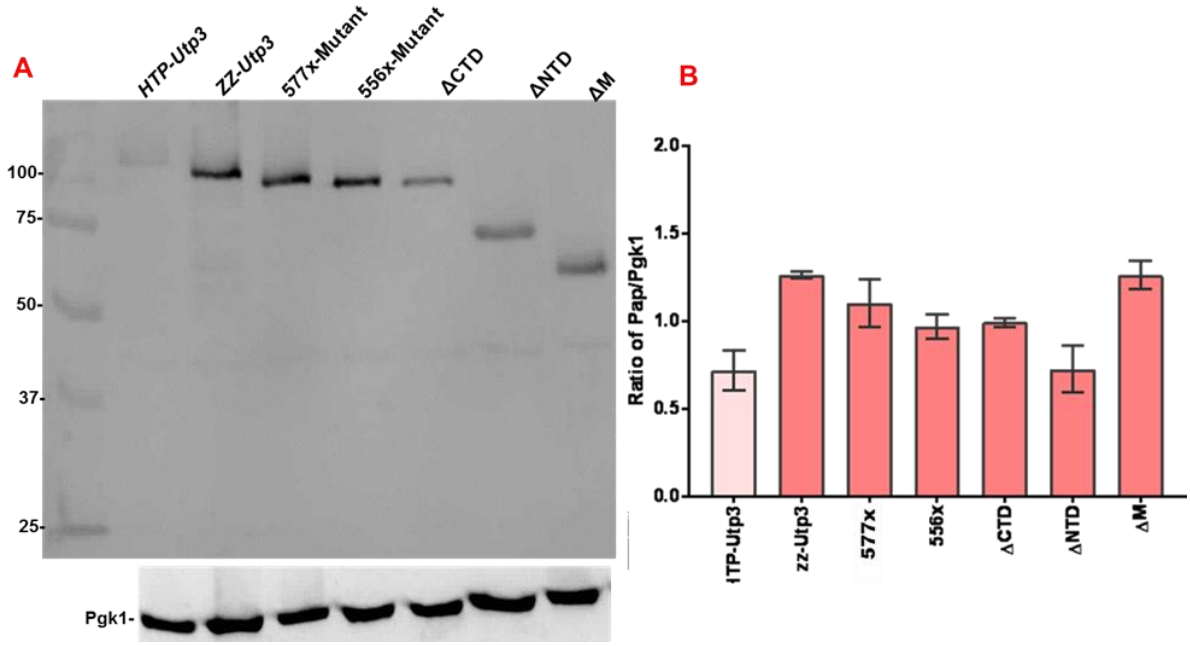

Figure 3.10: zz-Utp3 and all mutants expressed at approximately comparable levels.

A: Cell lysates from a Utp3 wild-type strain expressing either a full-length zz-Utp3 fusion protein or one of the Utp3 mutant derivatives were resolved by 10% SDS-PAGE and analysed by Western blotting using the PAP antibody (upper panel). Lysate from a strain expressing the chromosomally encoded HTP-Utp3 fusion protein under the control of the homologous *UTP3* promoter was also analysed to allow comparison with normal expression levels of Utp3. Blots were subsequently incubated with anti-Pgk1 antibody for normalisation (lower panel).

B: Quantification of the expression levels of Utp3 mutant proteins, relative to the wild-type protein expressed from the *UTP3* promoter. The error bars indicate the range of experimental values for biological duplicate samples.

3.4: Analysis of 18S rRNA synthesis defects in the Utp3 mutants

Utp3 is required for early pre- rRNA processing steps and for 18S rRNA production (Dragon et al., 2002). To address the impact of Utp3 mutants on 18S rRNA synthesis, *GAL::UTP3* transformants expressing the Utp*3* mutants

were grown to mid-log phase in galactose-based medium and then transferred to glucose-based medium. Cells were harvested at various time-points after addition of glucose and total cellular RNA was isolated, resolved through agarose gels and analysed by Northern blotting using probes specific to the major rRNA species (Fig. 3.11). RNA was also isolated and analysed from a wild-type strain during growth in galactose- and glucose-based medium.

The relative levels of 18S and 25S rRNAs in the wild-type strain remained comparable in galactose- and glucose-based minimal medium. In the *GAL::UTP3*, the amount of 18S rRNA was significantly depleted compared to the level of 25S rRNA after approximately 5 hours incubation in glucose-based minimal medium. A similar depletion of 18S rRNA was observed for the ΔN and K556X mutants. However, the 18S rRNA levels were not depleted in the K577X mutant. These observations are consistent with the genetic complementation data described above (Fig. 3.9) and support the conclusion that the K577X mutant protein is functional. In the case of the ΔCTD mutant, a significant depletion of 18S rRNA was observed but the effect appeared to require longer incubation in the presence of glucose and the levels of 18S rRNA appeared to recover by the 24-hour time-point. It could be that in this instance the 24-hour sample of rRNA was accidentally overloaded relative to the other lanes. Alternatively, suppression of RNA processing phenotypes upon prolonged incubation of strains with *GAL* regulated conditional alleles is a common phenomenon. Note that although multiple samples were analysed for each depletion experiment, these analyses are based on a single set of cultures.

Taken together, the genetic complementation analyses and the rRNA analyses indicate that the ΔN and K556X mutants are loss of function alleles that show a block in 18S rRNA production, whereas the K577X mutant allows 18S rRNA synthesis but shows a slow growth phenotype (Figs. 3.9 and 3.11).

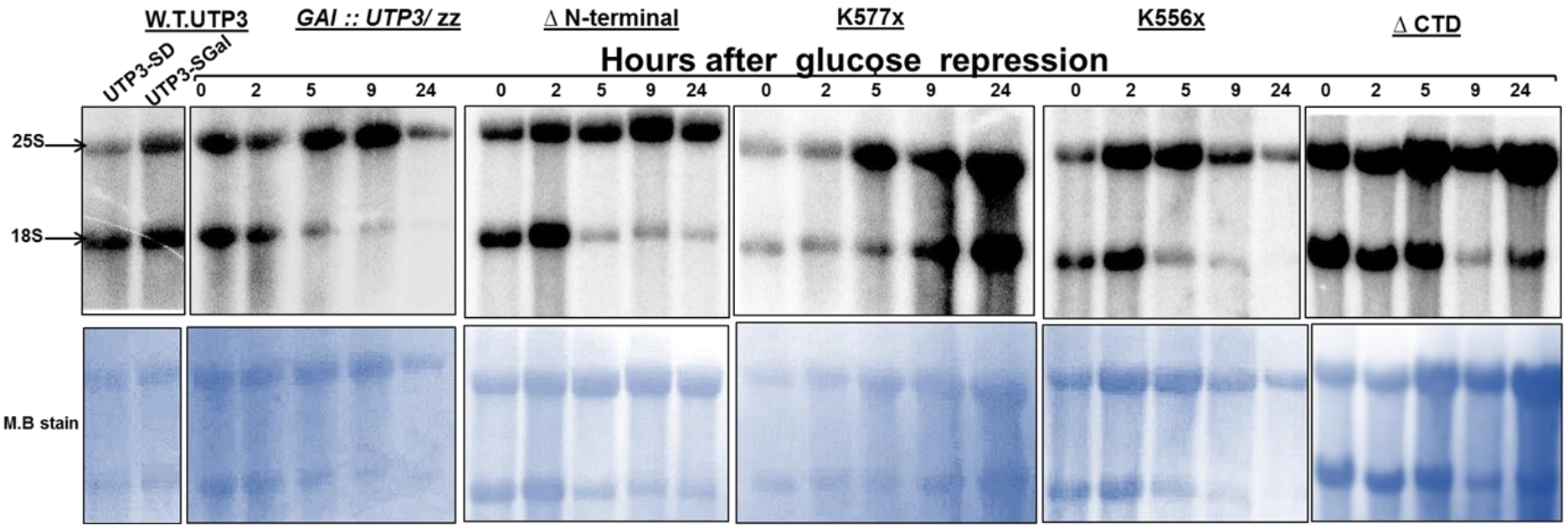

Figure 3.11: Loss of function Utp3 mutants causes depletion of 18S rRNA

Agarose gel Northern blot analysis of rRNA species in the conditional *GAL::UTP3* allele harbouring the ΔN-terminal, K775X, K556X or ΔCTD *utp3* mutants. *GAL::UTP3* harbouring plasmids encoding the wild-type zz-Utp3 or the cloning vector were also included as a positive and negative control, respectively. Strains were grown to mid-log phase in minimal galactose medium. Glucose was then added and samples collected at different time points, as indicated. The *GAL::UTP3* strain complemented with the wild-type *Utp3* gene was harvested after growth in glucose (SD) or galactose-based medium (SGal). RNA was isolated and 10µg aliquots were denatured and resolved through an agarose gel by electrophoresis. RNA was transferred to nylon membranes and analysed by Northern blot hybridizations using labelled oligonucleotides complementary to sequences within 18S and 25S rRNAs. Upper panels show the Northern blot hybridisation data; lower panels show images of blots stained with methylene blue prior to hybridisation.

3.5: Analysis of the sedimentation profile of Utp3 mutants

Data described above (Fig. 3.7) demonstrates that plasmid-expressed Utp3 co-sediments in part with large, ribosome-sized complexes that can be clearly resolved from soluble cytosolic proteins by sucrose density gradient centrifugation.

To determine whether SSU processome assembly was impaired in cells containing either viable or non-viable Utp3 mutants, sucrose gradients were performed on lysates of each mutant and compared to wild type zz-Utp3.

To address whether the Utp3 mutant proteins can be stably incorporated into SSU processome complexes, cell lysates from wild-type strains expressing each epitope-tagged Utp3 fusion protein mutant were resolved through sucrose gradients. Aliquots of the gradient fractions were analysed in parallel by SDS-PAGE (to determine the distribution of total cellular protein) and Western blotting (to address the sedimentation profile of Utp3 proteins) (Fig. 3.12).

Coomassie Blue staining of SDS-PAGE gels showed a consistent distribution pattern of cellular protein throughout the gradients, demonstrating that the gradients were technically reproducible. The majority of cellular protein was observed in fractions 2-5 at the top of the gradients, while the predominantly low molecular weight ribosomal proteins were observed in the middle of the gradiets in fractions 9-13.

Western blotting analyses revealed that wild-type Utp3 showed a similar distribution as before (see Fig. 3.7), with most of the protein found at the top of the gradient. As seen previously, a small discernible peak of Utp3 protein was seen in fractions 9 and 10 and leads further into the gradient. These fractions contain ribosomal complexes, based on the Coomassie Blue staining patterns, consistent with the incorporation of Utp3 into 90S complexes

The ΔCTD and K556X mutants show a similar distribution pattern to the wild-type protein; a considerable fraction of these Utp3 mutants is found in the

fractions at the top of the gradient but there is also a clear distribution through the middle fractions of the gradients (6-12). In the case of the ΔCTD mutant, there is a clearly pronounced peak in fractions 7-12. This suggests the ΔCTD protein is assembled into pre-ribosomal complexes that are nevertheless functionally impaired in 18S rRNA synthesis.

The K577X and ΔM mutants, both of which support cell growth, are distributed predominantly in fractions at the top of the gradient. This is particularly true for the ΔM mutant. In contrast, the distribution of the ΔNTD mutant shows a clear correlation with ribosomal complexes and there is hardly any protein in the fractions at the top of the gradient.

These data show that loss of function deletions within the CTD of Utp3 do not impact on its ability to assemble into larger complexes. Further analysis of the complex observed in the Δ M mutant may provide information regarding important protein interactions between Utp3 and other processome components.

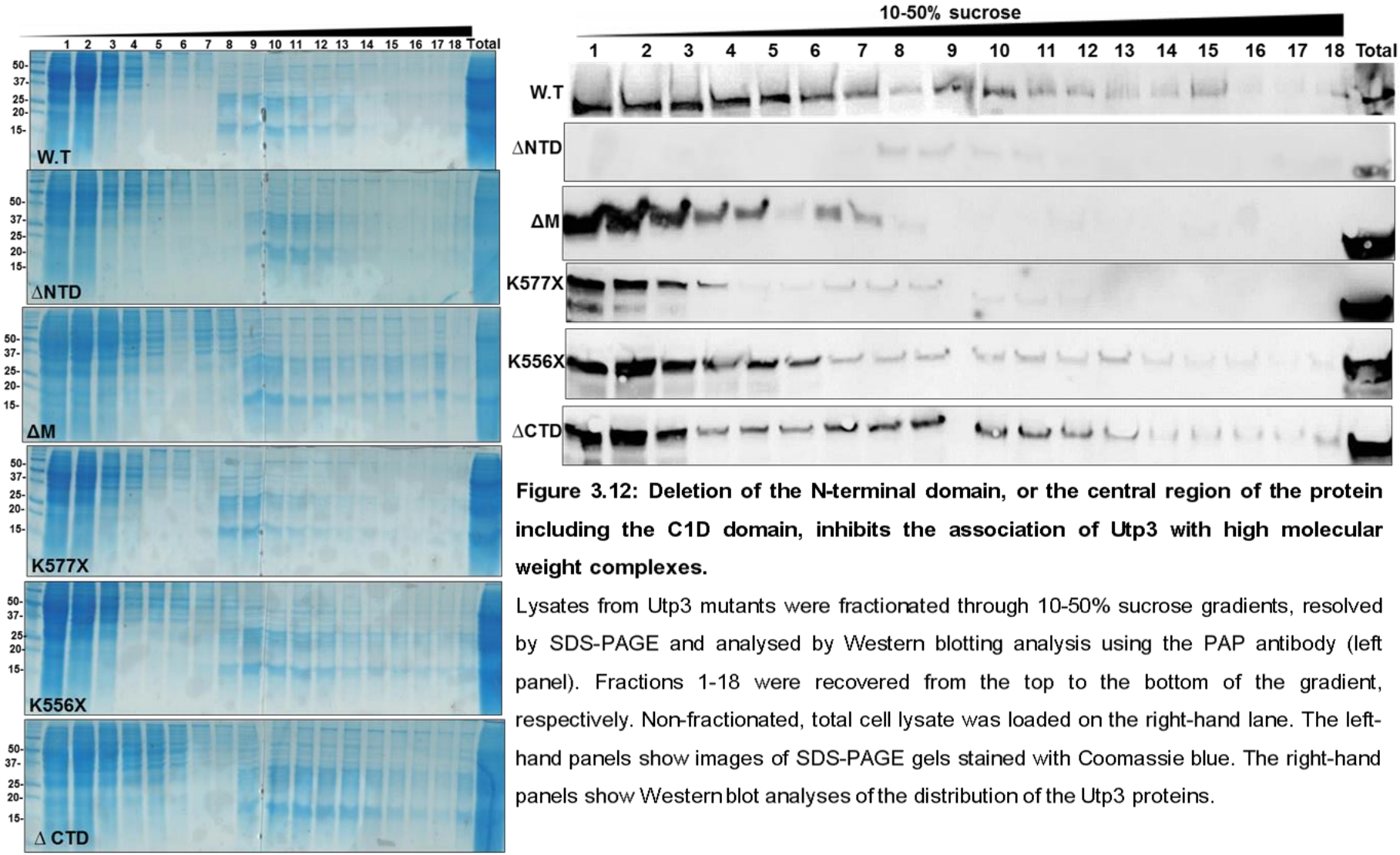

Figure 3.12: Deletion of the N-terminal domain, or the central region of the protein including the C1D domain, inhibits the association of Utp3 with high molecular weight complexes.

Lysates from Utp3 mutants were fractionated through 10-50% sucrose gradients, resolved by SDS-PAGE and analysed by Western blotting analysis using the PAP antibody (left panel). Fractions 1-18 were recovered from the top to the bottom of the gradient, respectively. Non-fractionated, total cell lysate was loaded on the right-hand lane. The left-hand panels show images of SDS-PAGE gels stained with Coomassie blue. The right-hand panels show Western blot analyses of the distribution of the Utp3 proteins.

Chapter 4 : Analysis of protein interactions involving Utp3

Utp3 has previously been shown to interact with the SSU processome components Mpp10, Utp6, Utp21 and Utp25 through two-hybrid assays (Charette and Baserga, 2010). The mutational analyses described in Chapter 3 demonstrate that the N-terminal region and the conserved C-terminal domain of Utp3 are critical for its function *in vivo*. However, the function of specific domains or regions within Utp3 in interactions with other components of the SSU processome has not been addressed. This chapter describes experiments that were aimed to define the domains or regions within Utp3 that mediate characterised interactions with Utp6, Utp21 and Utp25. Protein interactions were tested in pull-down assays peformed on yeast cell lysates that made use of the "zz" epitope-tagged wild-type, mutant Utp3 fusion proteins described in Chapter 3 and commercially available yeast strains that express GFP fusions of Utp6, Utp21 and Utp25 (Huh et al., 2003). In addition, a series of pull-down assays were done between recombinantly expressed C1D and CTD domains of Utp3 and either full-length Utp21 or Utp6 proteins, or the DUF1253 domain of Utp25.

4.1: Validation of strains expressing GFP fusions of the SSU proteins Utp21, Utp25 and Utp6 proteins

Haploid yeast strains expressing endogenous C-terminal, GFP-tagged fusions of Utp21, Utp25, and Utp6 were obtained from a commercially available source (Life Technologies). An *Mpp10-GFP* strain was not available and so the interaction between Utp3 and Mpp10 was not analysed in these studies.

The yeast *UTP6*, *UTP21* and *UTP25* genes are essential for cell growth (Giaever et al., 2002). To verify that the GFP fusion proteins are able to support normal growth, the growth of the *UTP6-GFP*, *UTP21-GFP* and *UTP25-GFP* strains were compared with the isogenic wild-type strain(p364) on minimal medium using a spot growth assay (Fig. 4.1, panel A). There was no difference observed between the growth of the wild-type strain and any of the strains expressing GFP-tagged SSU processome components, indicating that epitope tag did not significantly disrupt the function of Utp21, Utp25 or Utp6.

To confirm that the fusion proteins could be readily detected, cell extracts were prepared from the *UTP21-GFP*, *UTP25-GFP* and *UTP6-GFP* strains by alkaline lysis. Total cellular protein was resolved by SDS-PAGE and the fusion proteins was visualised by Western blot analysis using a GFP-specific antibody. Discrete bands corresponding to the fusion proteins were detected in the lysates of each strain. The observed migration of the fusion proteins was in agreement with the predicted molecular weights of the epitope-tagged variants of Utp21, Utp25 and Utp6 (~130kDa, 110 kDa and 75 kDa respectively) (Fig 4.1, panel B). The Utp21 signal was observed very weak barely detectable. This might be due to either degradation of the protein or protein insolubility.

To verify that the GFP cassette had been integrated into the correct chromosomal locus in these strains, PCR reactions were performed on genomic DNA samples using gene-specific forward primers and a GFP reverse primer. A specific band corresponding to the predicted size of the amplified DNA product was observed in the PCR reaction mixtures using genomic DNA from the *UTP21-GFP, UTP25-GFP* and *UTP6-GFP* strains (3.0, 2.7 and 2.1 Kb respectively). Furthermore, PCR reactions performed on genomic DNA from a wild-type strain did not give any product with any set of primers. These data confirm that the *UTP21-GFP*, *UTP25-GFP* and *UTP6-GFP* strains express functional GFP fusion proteins (Fig. 4.1 C).

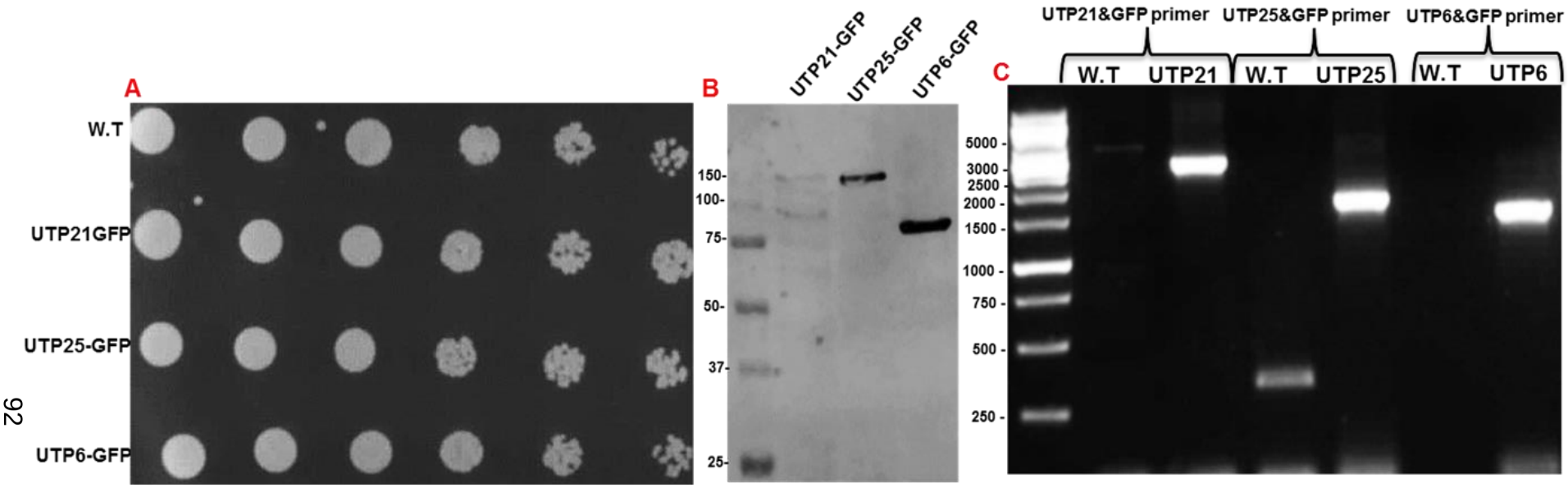

Figure 4.1 : Validation of UTP21-UTP25-UTP6 GFP tagging construct

A: GFP tagged strains show a similar rate of growth compared with wild-type strain. Image of 10- fold serial dilutions of cells grown in glucose based minimal medium at 30°C was taken after 3 days.

B: Total cell lysates were prepared by alkaline lysis from GFP tagged fusion protein Utp21, Utp25, Utp6 and analysed by Western blot using GFP antibody. GFP fusion protein migration corresponds to the estimated molecular weight of Utp21, Utp25, Utp6 (~130kDa, 110 kDa, 75 kDa respectively).

C: PCR reaction performed using wild-type and GFP strain genomic DNA. Wild-type was used as a control with forward primer of each GFP strain and general GFP reverse primer. PCR product of GFP strains corresponds to gene sizes for *UTP21*, *UTP25* and *UTP6* respectively. Using corresponding wild-type genomic DNA did not give any product.

4.2: Utp21, Utp25 and Utp6 show distinct sedimentation profiles

To analyse the size distribution of complexes containing Utp-21, Utp25 and Utp6, cell extracts were prepared from the strains expressing GFP fusion proteins and resolved through sucrose density gradients by ultracentrifugation. The distribution profiles of Utp21, Utp25 and Utp6 fusion proteins in the resulting fractions were analysed by SDS-PAGE and Western blotting (Fig. 4.2).

SDS-PAGE analysis of the distribution of total cellular protein through the gradients revealed patterns similar to those described previously, with most of the cellular protein present in fractions at the top of the gradient and fractions in the middle of the gradient being rich in low molecular weight, ribosomal protein. However, these gradients showed a wider range of variability in the resolution of ribosomal complexes and soluble protein. Utp21 and Utp6 tagged proteins, which are both components of the UtpB subcomplex, were observed in fractions through the middle of the gradients. These observations are consistent with their association with large complexes the size of pre-ribosomes (Fig. 4.2). A relatively small amount of Utp21 and Utp6 was observed in the fractions at the top of the gradients. In contrast, the distribution profile of Utp25-GFP showed the protein detected predominantly in fractions at the top of the sucrose gradient with no significant signal in the fractions containing ribosomal complexes. Utp25 has been shown to interact with Mpp10 and Utp3 by co-immunoprecipitation and two-hybrid analyses (Krogan et al., 2006; Charette and Baserga, 2010), but has not been shown to be physically associated with the 90S SSU processome. Previous studies have found reproducible low co-immunoprecipitation efficiencies of Utp25 with processome components, specifically Mpp10, Utp8, Utp18, Rrp9 and U3 snoRNA in comparison to Utp21 and Utp6 (Bleichert et al., 2006; Charette and Baserga, 2010)' These data suggest that the association of Utp25 with the 90S SSU may be either labile or transient.

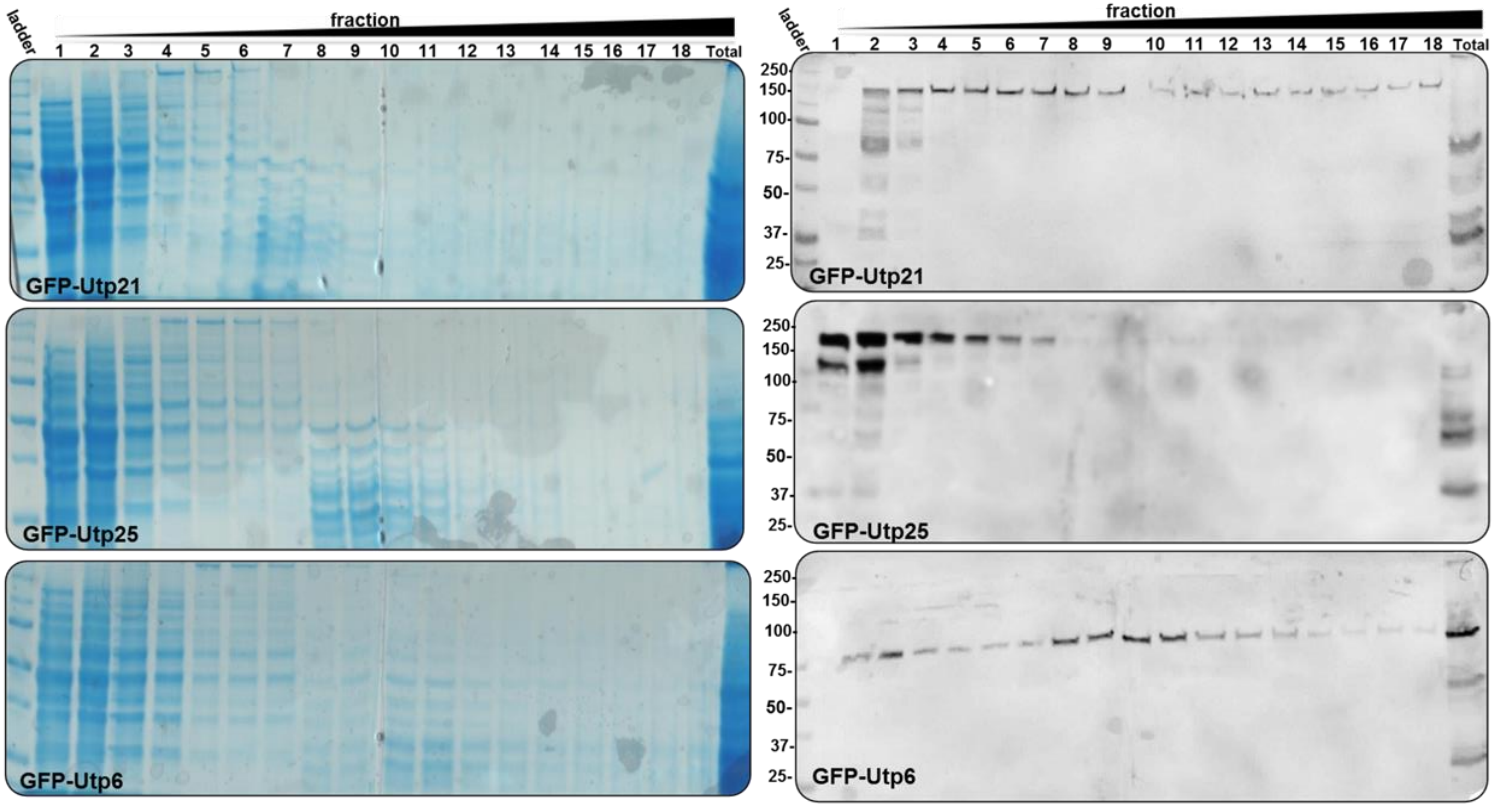

Figure 4.2 : Sedimentation profile Utp21, Utp25 and Utp6 through sucrose gradients

Cell Lysates from strains expressing GFP fusion proteins of Utp6, Utp21 or Utp25 were fractionated through 10-50% sucrose gradients and analysed by SDS-PAGE and Western blot analysis using an anti-GFP antibody. Aliquots from identical fractions were analysed by Coomassie blue staining (left hand panel) and Western lotting (right hand panel). Fractions 1-18 were taken from the top to the bottom of the gradients. Nonfractionated whole cell lysate (Total) was loaded in the right hand side lanes. The electrophoretic mobilities of molecular weight marker proteins (loaded on the left hand side of each panel; sizes are in kDa) are indicated. Note that samples for Coomassie blue staining were resolved through 16% SDS-PAGE gels, whereas western blot analyses were performed on samples that were resolved through % 10 SDS-PAGE gels.

4.3: Deletion of either the C1D or CTD domain of Utp3 impairs its interaction with the UtpB subcomplex

To determine whether specific deletion mutations within Utp3 block its ability to interact with Utp21, Utp25 and Utp6, pull-down experiments were performed on cell lysates from strains expressing zz epitope-tagged fusions of either the wild-type or mutant Utp3 protein and GFP-tagged Utp21, Utp25 and Utp6 fusion proteins. Preliminary Western analyses confirmed the expression of both the tagged Utp3 and tagged Utp21, Utp25 and Utp6 in the transformed yeast strains, and showed that coexpression of the two fusion proteins did not alter their distribution profiles within sucrose density gradients (data not shown).

Initially, pull-down experiments were performed on cell lysates after fractionation through sucrose density gradients in order to address whether the complexes observed were found within ribosome-sized or smaller protein complexes. However, the Western analysis data obtained from these experiments were difficult to interpret, with the proteins consistently appearing as smeared rather than distinct bands. Therefore, pull-downs were performed on non-fractionated cell extracts.

Initial pull-down reactions on Utp21-GFP and Utp6-GFP strains using cell lysis buffers containing 150mM NaCl showed a degree of nonspecific retention of epitope-tagged proteins on the IgG-sepharose beads. Therefore, subsequent assays were performed on cell lysates prepared using buffers containing 250mM KCl or NaCl.

Figure 4.3 shows a dataset from pull-down experiments assaying wild-type and mutant variants of zz-tagged Utp3 for interaction with Utp21-GFP. For each binding assay, equivalent fractions of whole cell lysate (WCL) and flow-through (F) samples, and a 10-fold equivalent of the eluate (E) fractions were analysed. Samples were resolved by SDS-PAGE and the fusion proteins were visualised by incubation of a single Western membrane, firstly with the anti-GFP antibody and subsequently with the PAP antibody complex. For this

reason, the signal for the Utp21-GFP fusion protein is also visible on the Western analysis of the zz-Utp3 proteins. The two fusion proteins were differentiated by comparative analysis of the two Western images obtained for each blot. For clarity, the bands corresponding to full-length Utp3 and Utp21 fusion proteins are indicated by red arrowheads and asterisks, respectively.

Comparision of the Western signals shows that each of the Utp3 proteins (indicated by arrowheads in each of the PAP Western images) was efficiently depleted from the cell lysates. A strong signal was observed for Utp21-GFP in the eluate fraction from lysate containing wild-type Utp3 protein and retention on the IgG sepharose beads was dependent upon Utp3 protein, since no signal was observed in the eluate fraction from the vector control. A strong interaction was also observed between *Utp21-GFP* and the Utp3-ΔNTD mutant that lacks the N-terminal region of the protein. In contrast, only very weak binding was observed for Utp3 proteins lacking the C1D domain, the central region of the protein (Utp3-ΔM) or with deletions at the C-terminus. Notably, deletion of the whole C-terminal domain of Utp3 had a stronger effect on Utp21 binding than the less extensive K577X and K556X deletions. These data clearly show that the interaction between Utp3 and Utp21 is independent of the N-teminal region of Utp3, whereas deletion of either the C1D domain or C-terminal domain of Utp3 has a strong inhibitory effect on this interaction.

A very similar set of observations were made in pull-down assays between Utp3 and another UtpB subcomplex component, Utp6 (Fig. 4.4). The wild-type and mutant Utp3 proteins were efficiently depleted from the cell extracts and Utp6 was detected in the eluates of pull-downs in the presence of wild-type Utp3 but not in the negative control. Moreover, deletion of the N-terminal region of Utp3 had no effect on Utp6 binding but deletion of the C1D domain or C-terminal truncations completely blocked the interaction with Utp3. Efficient binding was observed for the K577X Utp3 mutant, compared to the more extensive K556X and ΔCTD deletions.

These experiments demonstrate that the C1D and CTD domains of Utp3 are required for stable interaction with Utp21 and Utp6. Binding with both Utp21

and Utp6 was more evident with the smaller C-terminal deletion (K577X) than more extensive deletions (K556X and ΔCTD). In contrast, the interaction was independent of the N-terminal region of the protein. Taken together, both groups of pull-down experiments show that a stable interaction between Utp3 and the UtpB subcomplex is dependent upon both the C terminal and the C1D domain of Utp3.

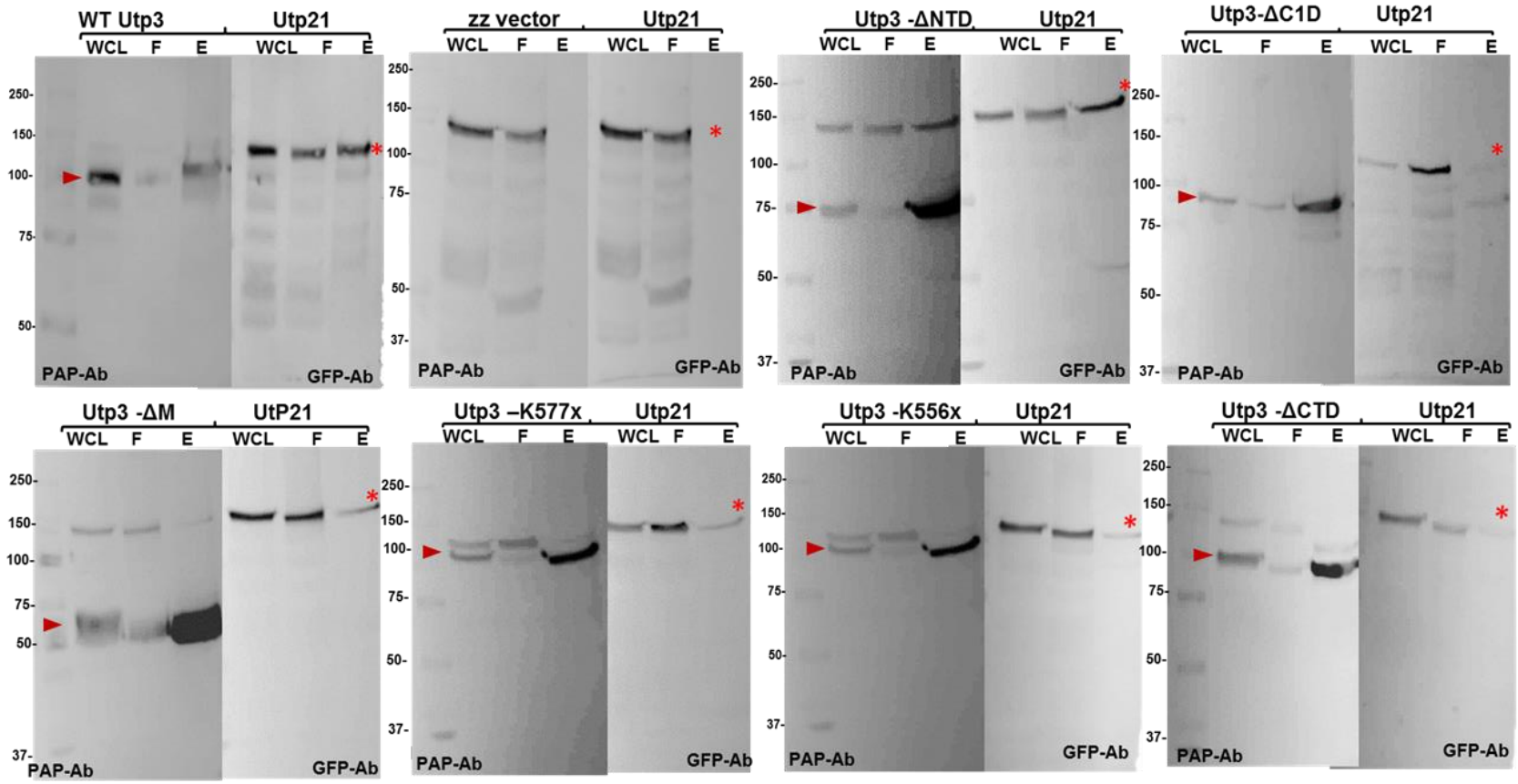

Figure 4.3: Deletion of C1D or CTD domains of Utp3 impacts on its interactions with UtpB subcomplex component Utp21

Cell extracts were prepared from Utp21-GFP strains that express zz-tagged wild-type or mutant *Utp3* alleles, and pull-downs were performed using IgG sepharose beads. Transformants harbouring a plasmid that expresses only the epitope tag (zz vector) were analysed as a negative control. Whole cell lysate (WCL), flow through (F) and elute (E) fractions were resolved by SDS-PAGE and analysed by Western blotting. Individual membranes from each pull-down were reiteratively probed with anti-GFP antibody. Membranes were then stripped and re-probed with anti-PAP antibody. The PAP Westerns contain residual signal from the preceding anti-GFP Westerns. Bands detected with the PAP antibody are indicated with red arrowheads and the GFP signal with asterisks. The eluate samples were 10-fold the equivalent fractions from whole cell lysates and flow-through samples. The elctrophoretic mobilities of molecular weight marker proteins (M, sizes indicated in kDa) are shown.

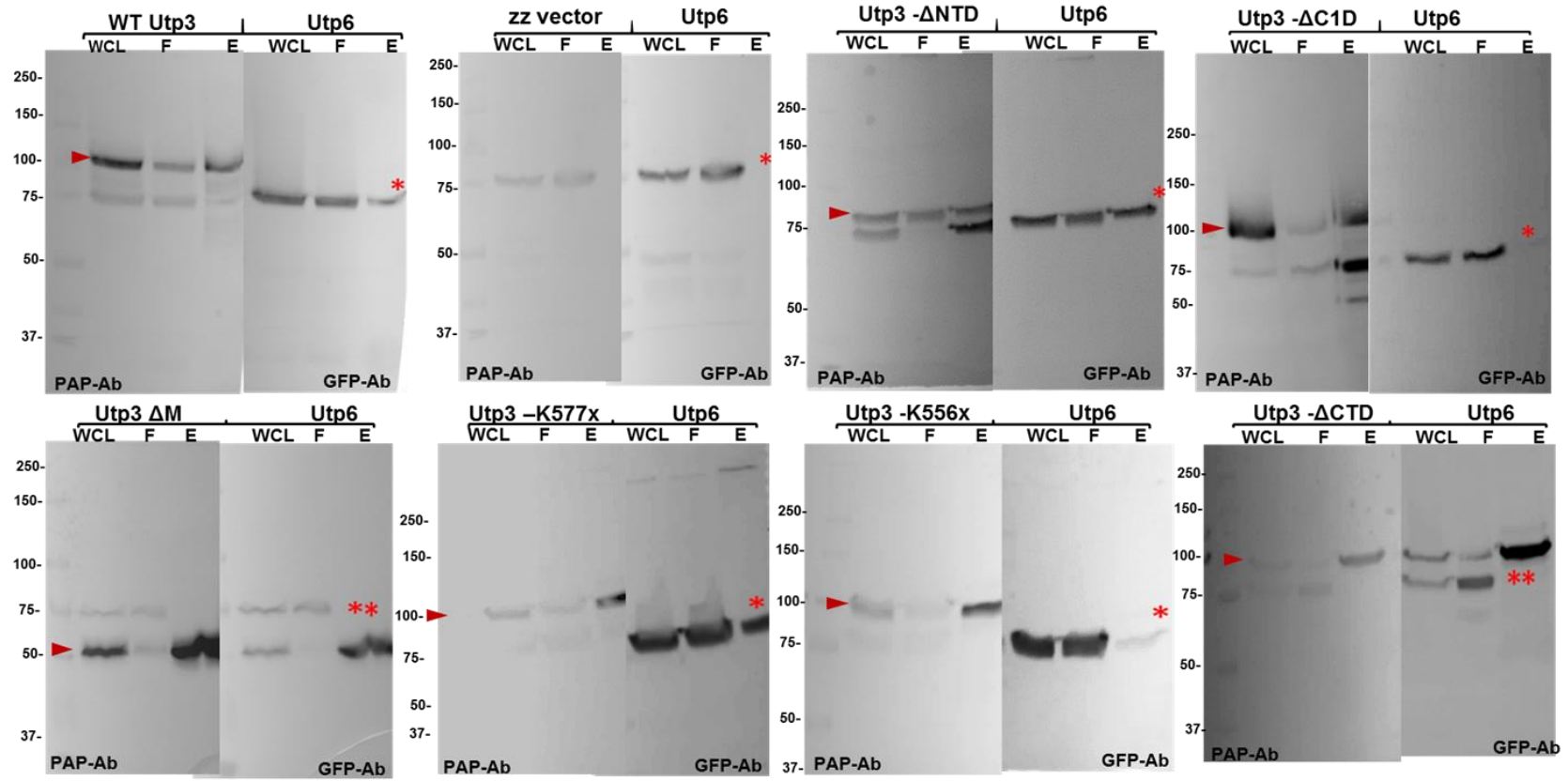

Figure 4. 4: Deletion of C1D or CTD domains of Utp3 impacts on its interactions with UtpB subcomplex component Utp6

Cell extracts were prepared from Utp6-GFP strains that express zz-tagged wild-type or mutant *Utp3* alleles, and pull-downs were performed using IgG sepharose beads. Transformants harbouring a plasmid that expresses only the epitope tag (zz vector) were analysed as a negative control. Whole cell lysate (WCL), flow through (F) and elute (E) fractions were resolved by SDS-PAGE and analysed by Western blotting. Individual membranes from each pull-down were reiteratively probed with anti-GFP antibody. Membranes were then stripped and re-probed with anti-PAP antibody. The PAP Westerns contain residual signal from the preceding anti-GFP Westerns. Bands detected with the PAP antibody are indicated with red arrowheads and the GFP signal with asterisks. The eluate samples were 10-fold the equivalent fractions from whole cell lysates and flow-through samples. The elctrophoretic mobilities of molecular weight marker proteins (M, sizes indicated in kDa) are shown. (**)Cross reaction from secondary antibody.

4.3.1 Utp3 interactions with the UtpB subcomplex are independent of RNA

Our *in vivo* crosslinking studies (described in Chapter 5) show that Utp3 directly interacts with RNA. Furthermore, the CTD domain of Utp3, which is required for the interaction between Utp3 and the UtpB complex (Fig. 4.3 and 4.4), can interact with nucleic acid *in vitro* (Chapter 5). To determine whether the observed interactions between Utp3 and components of the UtpB sub-complex are dependent on RNA, a further series of pull-down experiments were undertaken in the presence or absence of RNAase A.

Utp21 and Utp6 were clearly detected in the eluates of pull-downs with Utp3, but were not detected in the eluates of the vector control samples lacking epitope-tagged Utp3. Importantly, the proportion of Utp6-GFP binding was not obviously different after incubation with RNAase A (Fig. 4.5 panel B, compare the Utp6 signal in the left and centre panels). To demonstrate that any RNA within the pull-down samples would have been degraded, 1μg tRNA was incubated in the presence or absence of the same amount of RNAase A enzyme, under the same conditions used for the pull-down reactions, and the reaction mixtures were resolved through an acrylamide/urea gel. tRNA was not detectable after incubation in the presence of RNAase A (Fig. 4.5, panel C). It can be concluded from this data that the interaction between Utp3 and the UtpB subcomplex might be not not indirectly mediated through protein-RNA interactions. The pull-down data for Utp21 were inconclusive, since the Utp21-GFP and zz-Utp3 proteins comigrated in these analyses (Fig. 4.5, panel A). another possibility is RNA was not digested due to protection from bound protein.

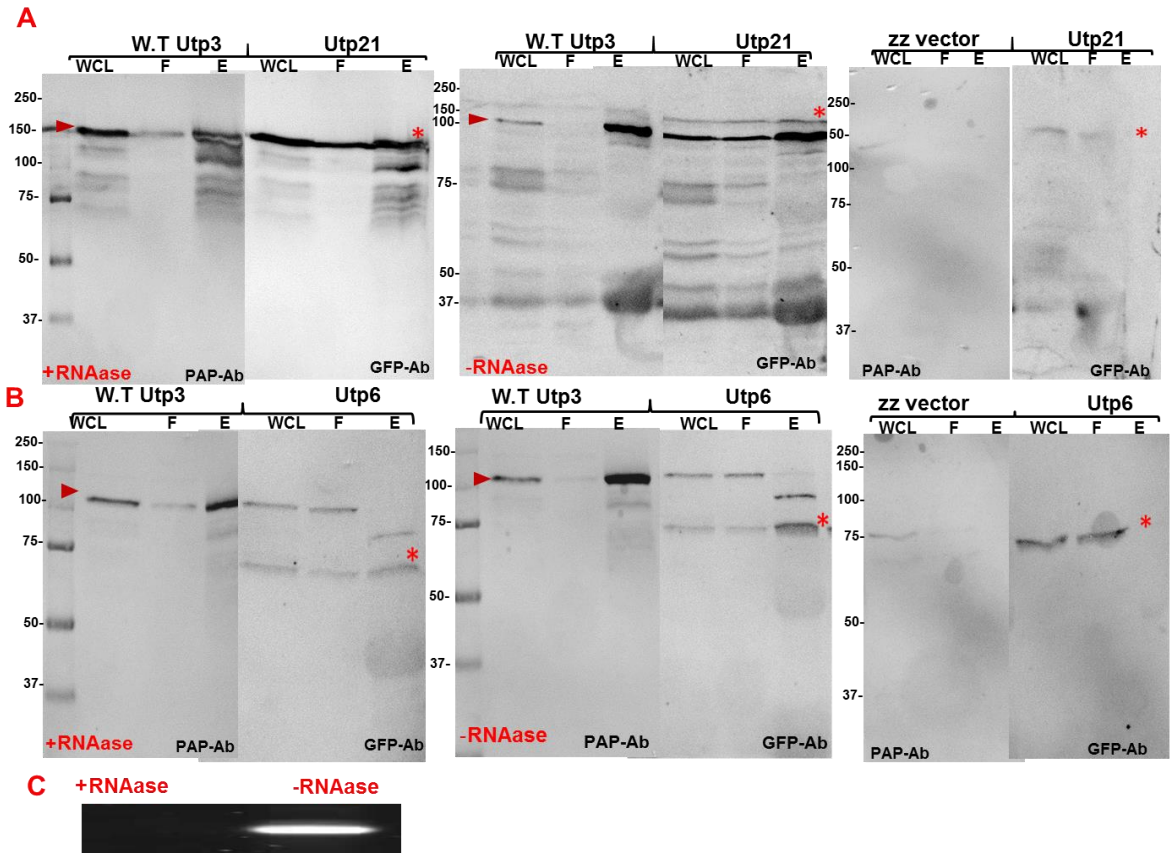

Figure 4.5:**Utp3 interaction with either Utp21 or Utp6 was independent of RNA.**

A, B: Pull-down reactions were performed on lysates from *Utp21-GFP* and *Utp6-GFP* strains expressing epitope-tagged zz-Utp3 or harbouring the vector that expresses only the zz-tag. The electrophoretic migration of molecular weight marker proteins (sizes in kDa), are indicated. Pull-downs were eluted either with or without prior incubation with RNAase A. Reiterative Western blot analyses were performed on whole cell lysate (WCL), flow through (F) and elute (E) samples using GFP antibody. The aliquots of elute fractions analysed were 10-fold the equivalent aliquots from cell lysates and flow-through fractions. Membranes were stripped and re-probed with anti-PAP antibodies. Images of Western blots obtained with the PAP antibody contain residual signal from the anti-GFP antibody. The full length zz-Utp3 fusion protein is indicated with red arrowheads. The GFP signal is indicated with asterisks.

C: Verification of RNAase activity. Yeast tRNA (1µg) was incubated with 1µg RNAase A under the conditions used for the pull-down assays. Reaction products were resolved through an acrylamide/urea gel. RNA was detected by ethidium bromide staining.

4.4: The Utp3 interaction with Utp25 is salt-sensitive

A comparable set of pull-down experiments were also performed to address the regions or domains within Utp3 that are required for its interaction with Utp25. When the pull-down experiments were carried out on lysates prepared using buffers containing 150mM NaCl, Utp25 was detected in the eluate fraction of the lysates containing zz-Utp3 fusion protein, but not in the eluate fractions of lysate from cell harbouring the control vector (Fig. 4.6). These data further support a specific interaction between the Utp3 and Utp25 proteins in yeast cell extracts (Krogan et al., 2006). Further investigations were undertaken to analyse the effect of different Utp3 mutations upon its interaction with Utp25 and thereby to determine the region(s) of Utp3 that are critically required for this interaction. Utp25 was detected in the eluate fraction of pull-downs using the N-terminal deletion of Utp3 (ΔNTD). In contrast, essentially no Utp25 was detected in the eluate fraction of pull-down reactions involving the ΔC1D Utp3 mutant. Weak interactions were observed to varying degrees in the presence of the C-terminal deletion mutants. The ΔM Utp3 mutant also showed a weak interaction. These data are largely consistent with the observations made upon analysis of Utp3 interactions with the UtpB subcomplex components (Figs. 4.3 and 4.4).

When the pull-down experiments were performed in buffers containing 250mM NaCl, Utp25 was observed bound to the full-length Utp3 protein but no clear association was observed with any of the Utp3 deletion mutants (Fig. 4.7). Taken together, these results suggest a similar interaction pattern between Utp3 and either Utp25 or the UtpB complex, with both interactions being independent of the N-terminal region of the protein but dependent upon the C1D domain and C-terminal domain of Utp3. The interaction between Utp3 and Utp25 is more sensitive to increased ionic strength than the interaction with the UtpB subcomplex components.

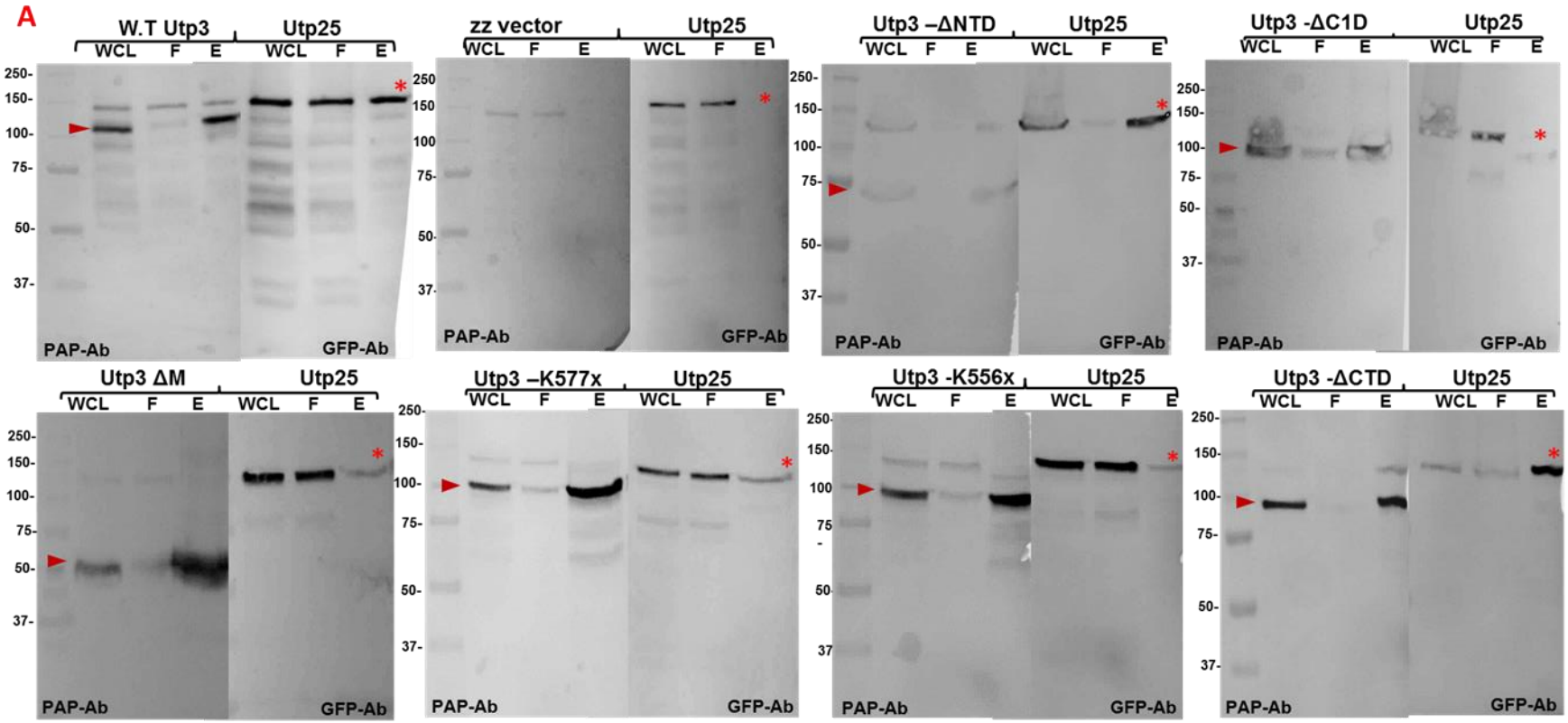

Figure 4.6: The Utp3 / Utp25 interaction is dependant upon the C1D domain and is salt sensitive

A: Deletion of the C1D domain inhibits the Utp3-Utp25 interaction. Cell lysates were prepared from a *Utp25-GFP* strain expressing epitope-tagged wild-type or mutant Utp3 alleles with a buffer containing 150 mM NaCl and pull-downs were performed on IgG sepharose beads. Equvalent fractions of whole cell lysate (WCL) and flow-through (F) fractions were analysed along with 10-fold equivalent of elute (E) fractions. Samples were resolved by SDS-PAGE and analysed by Western blotting, using anti-GFP antibody. Membranes were stripped and then re-probed with anti-PAP antibody. The electrophoretic migration of molecular weight marker proteins (sizes in kDa), are indicated. Bands derived from all PAP signals are indicated with red arrowheads and Utp25-GFP signals are indicated with asterisks.

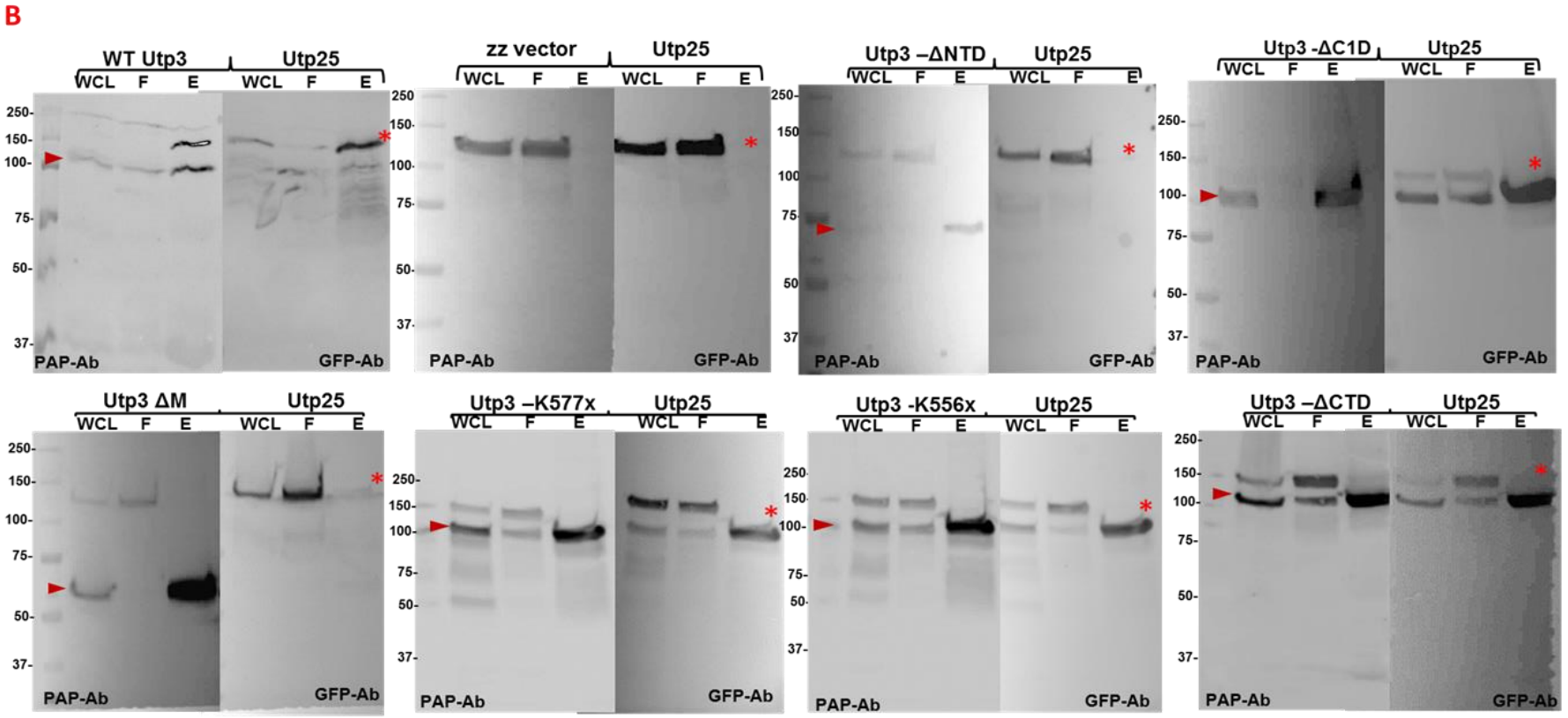

Figure 4.7: The Utp3/Utp25 interaction is dependant on the C1D domain and is salt sensitive

B: All *Utp3* mutants lost interaction with Utp25 in 250 mM NaCl. Pull-downs were performed and analysed as described in **4.6**, but using a cell extraction buffer containing 250 mM NaCl.

4.4.1: Utp3 interactions with Utp25 was independent on RNA

To determine whether the Utp3-Utp25 interaction is RNA-independent pull-down reactions were performed on lysates from the Utp25-GFP cells that either expressed full-length zz-Utp3 protein or harboured the control vector (Fig. 4.8). Utp25 was detected in the eluate fraction of the pull-down on lysate containing the zz-Utp3 protein but not in the eluate fraction fom the control lysate. Treatment of the bound fractions with RNAase A prior to elution did not cause a loss of Utp25 signal. These observations support the conclusion that the interaction between Utp3 and Utp25 is independent of RNA.

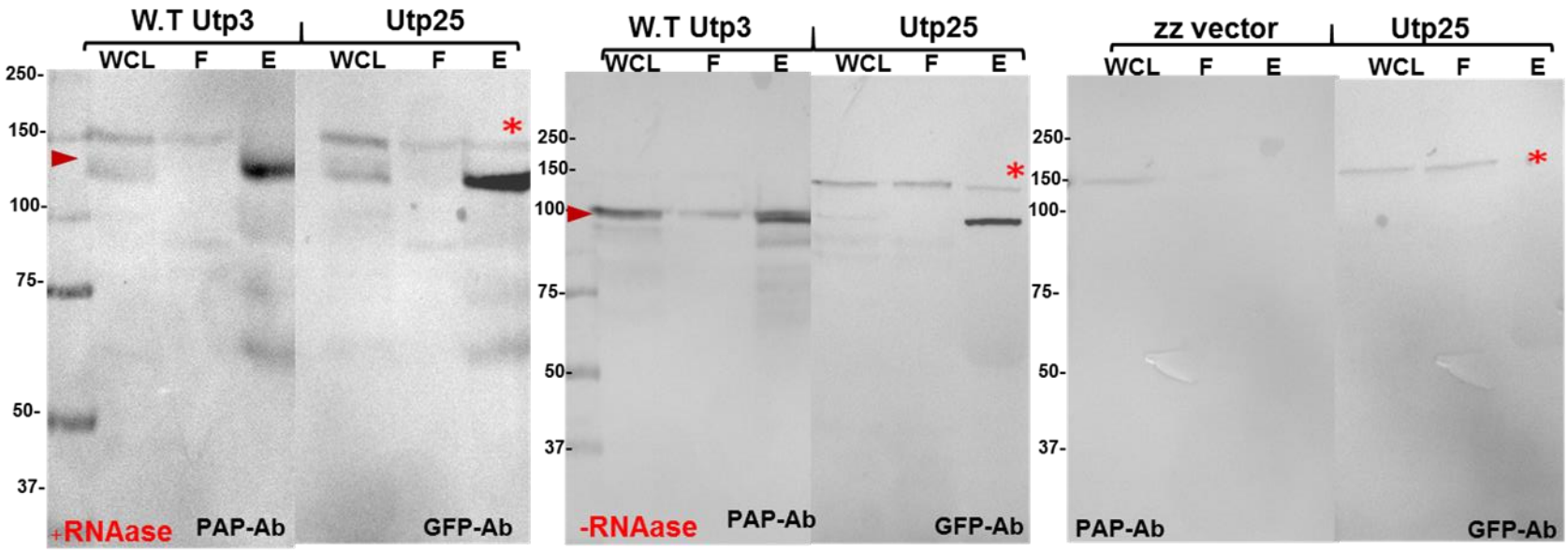

Figure 4.8 : Utp3 interaction with Utp25 is RNA independant.

Pull down reactions were performed on lysate from Utp25-GFP strains expressing epitope-tagged zz-Utp3 or harbouring the vector that expresses only the zz-tag. The electrophoretic migrations of molecular weight marker proteins (sizes in kDa), are indicated. Pull-downs were eluted either with or without prior incubation with RNAase A. Reiterative Western blot analyses were performed on whole cell lysate (WCL), flow-through (F) and elute (E) samples using GFP antibody. The membranes were stripped and reprobed with PAP antibody. The aliquots of elute fractions analysed were 10-fold the equivalent aliquots from cell lysate and flow-through fractions. Images of Western blots obtained with the PAP antibody contain residual signal from the anti-GFP antibody. The full-length Utp3 protein is indicated with red arrowheads . Utp25-GFP signal is indicated with asterisks.

4.5: Construction of bacterial expression constructs for recombinant protein interaction studies

Pull-down assays on yeast cell lysates described above have shown that Utp3 interacts with components of the UtpB subcomplex in a manner that is dependent upon its C1D and CTD domains and independent of the N-terminal region of the protein. The interaction with Utp25 appears to have similar physical requirements but is less stable.

To determine whether Utp3 interacts directly with Utp21, Utp25 or Utp6, constructs were generated to enable the expression and purification of recombinant proteins from *Escherichia coli*. To facilitate protein interaction assays, the N-terminal region, C1D domain and C-terminal domain of Utp3 were expressed as His-tagged proteins, while Utp21, Utp6 and the DUF1253 domain of Utp25 were expressed as GST fusions. A construct for the expression of the Utp3 C1D domain was already available from a previous study (Turner, 2011). All other constructs used in this study were generated as part of this work. Genomic DNA encoding residues 1-219 or 532-608 of the *UTP3* ORF (comprising the N-terminal region and the CTD of Utp3, respectively) were amplified by PCR and cloned by restriction digestion and ligation into the pRSETb vector. The resultant constructs encode N-terminal hexahistidine (6×His) tagged polypeptides, the expression of which are IPTG-inducible. Candidate clones were screened by restriction digestion (Fig. 4.9, panel A) and validated by sequence analysis through the complete length of the amplicon.

Yeast two-hybrid studies suggest that the interaction between Utp25 and Utp3 is mediated via the DUF1253 domain of Utp25 (Charette and Baserga, 2010). The region within Utp21 or Utp6 that interacts with Utp3 is not known. Therefore, constructs were generated for the bacterial expression of full-length Utp21 and Utp6 proteins and the Utp25 DUF1253 domain. The corresponding sequences were amplified from yeast genomic DNA and cloned into the pGEX-6P-1 vector, which encodes N-terminal GST fusion proteins. Candidate clones were screened by restriction digestion (Fig. 4.9, panel B) and verified by sequence analysis through the length of the whole insert.

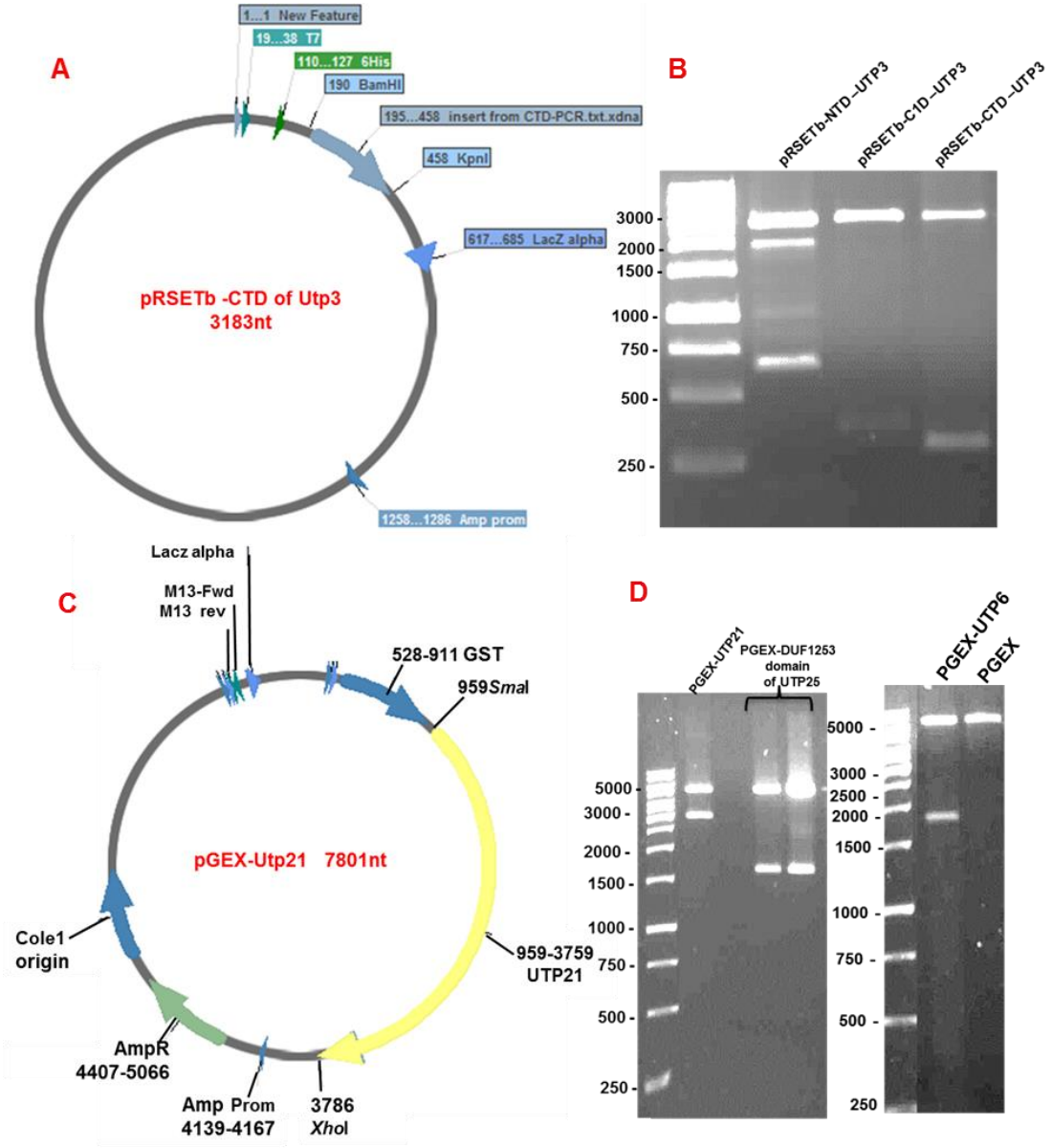

Figure 4.9: Construction of clones encoding His fusions of NTD and CTD of Utp3 and GST fusions of Utp21, Utp6 and the DUF1235 domain of Utp25.

A: The DNA encoding for NTD and CTD domains of Utp3 was cloned as *BamH*I and *Kpn*I fragments into the bacterial expression vector pRSETb. A schematic representation of pRSETb-CTD of Utp3 construct is shown.

B: pRSETb-NTD and pRSET-CTD constructs were digested with *BamH*I and *Kpn*I, then the incubation mixtures were resolved through 1% agarose gel to confirm successful cloning.

C: The full-length ORFs of *UTP6* and *UTP21,* and the nucleotide sequence encoding the DUF1235 domain of UTP25 were cloned as *Sma*I-*Xho*I fragments into the bacterial expression vesctor pGEX-6P1. A schematic representation of the pGEX-Utp21 construct is shown.

D: The constructs in **C** (above) were digested with *Sma*I and *Xho*I and then the incubation mixtures were resolved through 1% agarose gel to confirm successful cloning.

4.5.1: Optimizing the expression of recombinant proteins

Various protocols were tested to maximise the expression of the recombinant protein in a soluble form. Recombinant proteins were induced in bacterial cultures during incubation at temperatures ranging from 20°C to 37°C and by the addition of IPTG to a final concentration of between 0.05 and 1mM. To determine the relative level of soluble and insoluble recombinant protein obtained, native and denatured extracts were prepared from induced cells, resolved by SDS-PAGE and screened by Western analyses.

Optimal results for expression of the Utp3 C1D domain were obtained by inducing expression with 0.2mM IPTG and the best results for the Utp3 CTD domain were obtained by inducing expression with 1mM IPTG, both at 37°C. Expression of the N-terminal domain of Utp3 proved ineffective under all conditions tested.

Utp21 was expressed best when induced with 0.5mM IPTG at 37°C, whereas Utp6 was expressed best when induced with 0.2mM IPTG at 25 °C. None of the methods trialled to obtain expression of DUF1253 domain of Utp25 were effective. Representative gels showing the purification of His-tagged C1D and CTD domains of Utp3 are shown in (Fig 4.10).

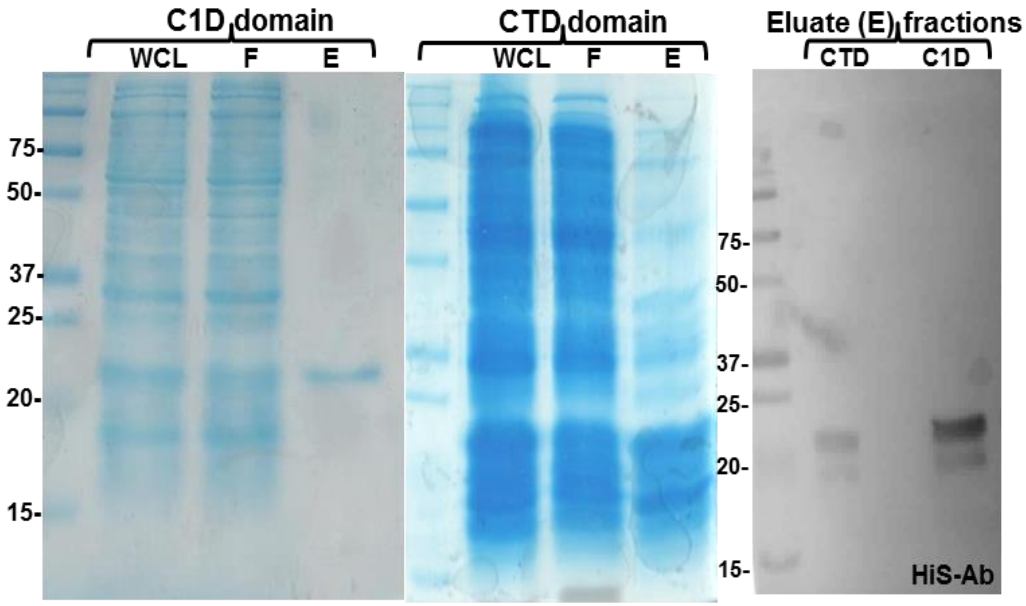

Figure 4.10: Construction of clones encoding His fusions of NTD and CTD of Utp3 and GST fusions of Utp21, Utp6 and the DUF1235 domain of Utp25.

The Hexahistidinyl-tagged C1D and CTD domains of Utp3 were expressed in *E. coli* and purified from cell lysates by affinity chromatography using Ni-NTA beads. Whole cell lysate (WCL), flow through (F) and eluate (E) fractions were resolved by SDS-PAGE through 16% acrylamide gels and transferred to nitrocellulose membranes for Western blotting. Total proteins (left and centre panels) were visualised by staining with colloidal Coomassie blue. Western blot analysis of the purification of the C1D fusion protein (right panel) was performed using an antibody against his tag.

4.6: Mapping the Utp3 interaction with Utp21 and Utp6

Data obtained from the yeast pull-down experiments described above suggest, but do not prove, that Utp3 interacts with Utp21 and Utp6 through the C1D and CTD domains. To investigate whether these interactions reflect direct contacts, recombinantly expressed, His-tagged C1D and CTD domains of Utp3 were tested for interaction with recombinant GST fusions of Utp21 and Utp6.

The recombinant proteins were expressed separately in *E. coli* and pull-down experiments were performed using glutathione beads, since this gave less apparent background binding than the Ni-NTA beads. Lysates from cells expressing either a GST Utp21 or Utp6 fusion protein, or the GST tag alone, were passed over glutathione beads. Lysate from cells expressing either His-tagged C1D or CTD domains from Utp3 were then incubated with the charged glutathione beads and, after extensive washing, retained protein was eluted using gel loading buffer.

When the beads were incubated with lysate from cells expressing the GST tag alone and subsequently with lysate from cells expressing the C1D or CTD domain of Utp3, there was no detectable signal for the His-tagged Utp3 polypeptides in the eluate fractions (Fig. 4.11 panel-A). These data demonstrate that neither the C1D domain nor the CTD domain of Utp3 has an affinity for the GST tag. In contrast, Western blot analysis revealed a signal in the eluate fractions for both the C1D and CTD domains when beads were incubated with GST-Utp21 (Fig. 4.11- Panel-B). Similarly, Western analyses of pull-downs with GST-Utp6 revealed an interaction with both the C1D and CTD domain of Utp3 (Fig. 4.11- Panel-B –lower panel). These data suggest direct interactions between both the C1D domain and CTD domains of Utp3 and Utp21, as well as Utp3 and Utp6. This indicates that multiple contacts exist between Utp3 and the UtpB subcomplex of the 90S processome.

The recombinant protein pull-down assays suggest that the C1D and CTD domains of Utp3 can independently and directly interact with the UtpB subcomplex. To further address these findings, pull-down assays were performed on whole cell lysates from yeast strains expressing Utp21-GFP,

Utp6-GFP or Utp25-GFP fusion proteins using Ni-NTA beads that were charged with the recombinant C1D and CTD domains of Utp3 (Fig. 4.12).

Clear signals were obtained in the eluate fractions for Utp21-GFP and Utp6-GFP when the beads were incubated with either the C1D domain or the CTD domain of Utp3, whereas no signal was detectable in either pull-down reaction for Utp25-GFP (Fig. 4.12). These data provide further support to the conclusion that both the C1D and CTD domains of Utp3 are engaged in interactions with the UtpB subcomplex. Moreover, these data extend the findings described above that the two domains of Utp3 can interact independently with the UtpB subcomplex.

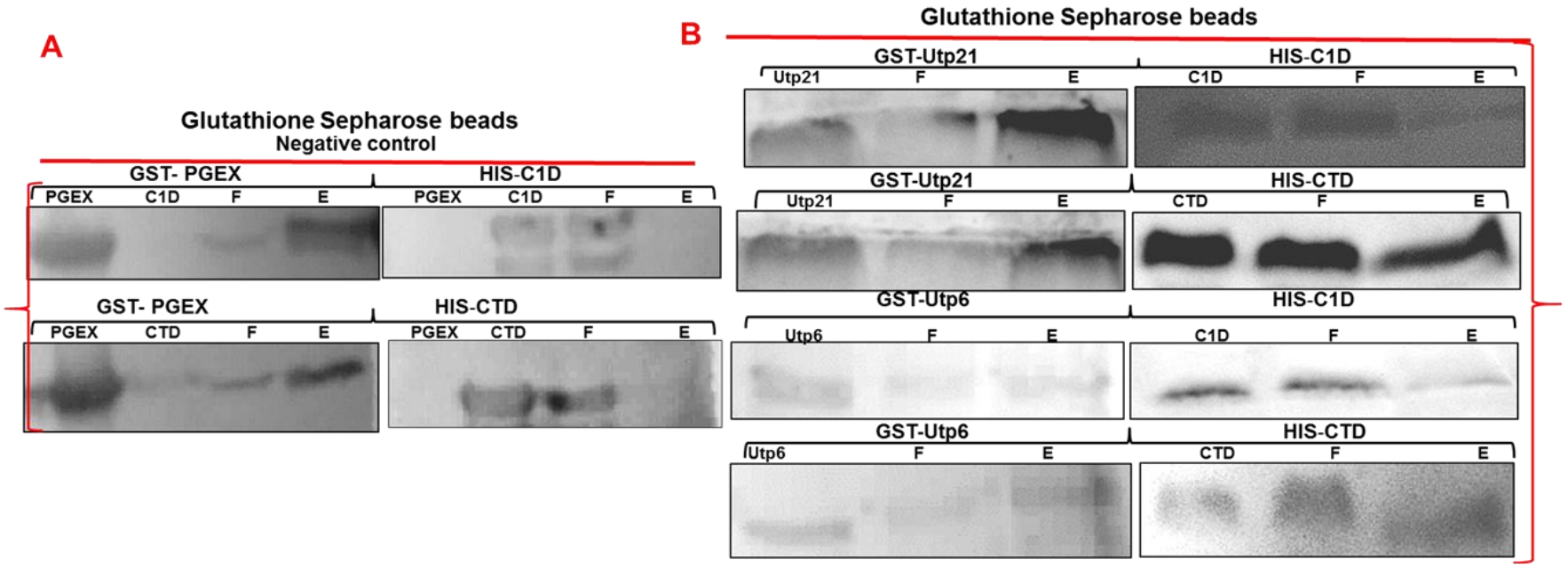

Figure 4.11: C1D and CTD interact directly with UtpB complex via Utp21 and Utp6 in vitro

Cell lysates prepared from C1D and CTD of Utp3 expressed in *E .Coli* along with GST-Utp21, GST-Utp6. Pull-downs were performed on Glutathione Sepharose beads. Input (WCL), flow-through (F) and elute (E) were resolved by SDS-PAGE and analysed by Western blotting using anti-GST and anti His antibody respectively. There is clear binding of C1D and CTD, but not empty vector.

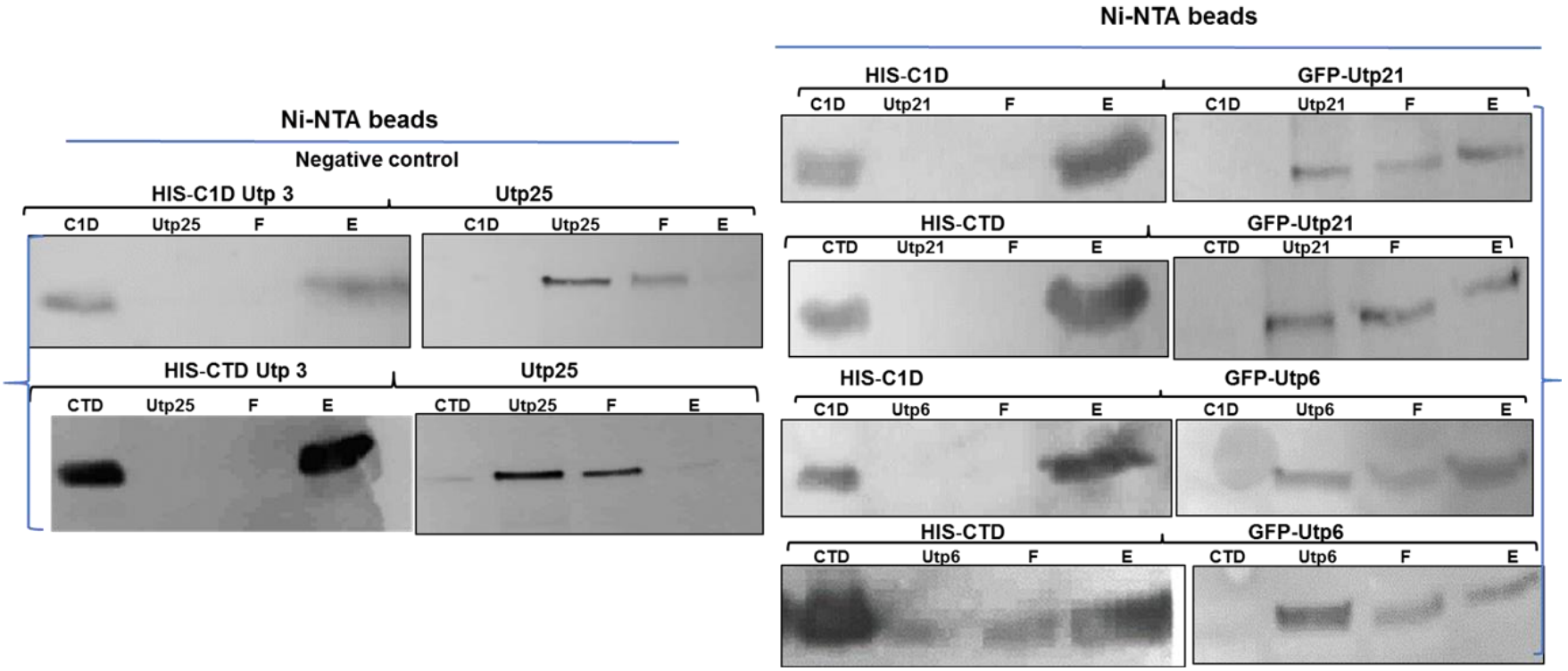

Figure 4.12: C1D and CTD interact directly with UtpB complex via Utp21 and Utp6 in vitro

Cell lysates prepared from C1D and CTD of Utp3 expressed in *E .Coli* were incubated with GFP-Utp21, GFP-Utp6 and GFP-Utp25 yeast strain lysates. Pull-downs were performed on Ni-NTA beads. Input (WCL), flow-through (F) and elute (E) were resolved by SDS-PAGE and analysed by Western blotting using anti-His and anti GFP antibody respectively. There is clear binding of Utp21 and Utp6, but not Utp25. The lack of binding with Utp25 indicates that the interactions are independent of the tag.

Chapter 5 : Analysis of the RNA binding properties of Utp3

Genetic and biochemical studies have shown that some components of the SSU processome seem to be scaffolding proteins harbouring motifs involved in protein-RNA binding or protein-protein interaction (Lin et al., 2013; Lu et al., 2013). For Utp3, there is no published information demonstrating that it is directly involved in RNA-protein interactions. However, since Utp3 contains a C1D domain that is found specifically in proteins that are present in RNA binding complexes, its function in 18S rRNA synthesis as part of the SSU processome may involve direct interaction with RNA.

An increasing number of proteins within the SSU processome have been shown to directly interact with rRNA using a technique known as CRAC (crosslinking and analysis of cDNAs) (Granneman et al., 2009; Granneman et al., 2010; Lebaron et al., 2012; Sardana et al., 2015; Wells et al., 2016; Hunziker et al., 2016). As part of the analysis of Utp3 function and in collaboration with the Tollervey lab, which established the CRAC experimental approach, the sites of interaction between Utp3 and cellular RNA were analysed on a genome-wide scale at nucleotide resolution. In this technique, either growing yeast cells or cell lysates are exposed to UV cross-linking. After purification of the epitope-tagged protein, cross-linked RNAs are trimmed and ligated to linkers followed by reverse transcriptase (RT)-PCR amplification and illumina sequencing. The sequences obtained are aligned with the yeast genome, and identified targeted RNAs sorted into functional categories. The number of mapped reads in each category for a tagged strain can then be compared to results for wild-type strain.

Having demonstrated that Utp3 does indeed directly contact RNA within growing cells, the domain within Utp3 involved in RNA interactions was addressed. In vitro electromobility shift assays (EMSA) (Hellman and Fried, 2007) were performed on DNA and RNA using a recombinantly expressed form of the C-terminal domain of Utp3 (pRSETb–CTD).

5.1: Generation of tagged Utp3 for target RNA identification

Epitope tagging is a routinely used technique in functional studies on *Saccharomyces cerevisiae*. Epitope tags are important tools in protein studies that help to reveal a protein's function through its purification, an analysis of its localisation and expression levels, and by providing a means to identify interaction partners.

To create an epitope-tagged strain, DNA encoding the tag and a selectable marker gene is amplified by PCR and introduced into the yeast genome through homologous recombination (Wach et al., 1994). Epitope tags can be introduced either at the C or the N terminal of proteins. C-terminal tagging of proteins is relatively easy and widely used, compared to N-terminal tagging, which is more complex (Longtine et al., 1998; Knop et al., 1999; Booher and Kaiser, 2008)

The epitope tag used universally for CRAC analysis is the HTP (His$_6$-TEV-protA) tag. The fusion proteins are expressed from the natural chromosomal locus under the control of the endogenous promoter. A C-terminally tagged *UTP3::HTP* yeast strain had been created as part of a previous study (Turner, 2011). However, cells expressing a C terminal Utp3-HTP fusion protein had a slow growth phenotype, compared to the isogenic wild type strain, suggesting that the C-terminal tag interfered with protein function and/or stability. Due to the slow growth phenotype of the *HTP::UTP3* strain, an N terminal tagged *HTP::UTP3* yeast strain that had been specifically developed in this laboratory to be used in CRAC analysis work was used.

The procedure used to make N-terminal tagged Utp3 depends on insertion of a selectable marker cassette, in this case the *URA3* gene and the epitope tag gene (*HTP-UTP3*), under the control of the *GAL* promoter. This is done upstream of the target gene of interest and then the Cre–Lox system of homologous DNA recombination is used to eliminate the *URA3* marker gene and *GAL* promoter (Sauer, 1987) (Fig. 5.1A). The Cre recombinase enzyme recognises a *LoxP* sequence introduced within the cassette to eliminate the *GAL* promoter and the *URA3* marker gene. The strain is then cultured on

medium containing 5-fluoroorotic acid (5-FOA) to select for the HTP-Utp3 recombinant isolates and simultaneously to purge the cells of the plasmid encoding the recombinase.

To characterize the growth phenotype of the resultant isolates, a yeast strain expressing the HTP-Utp3 fusion protein was plated on medium containing 5-FOA, glucose medium lacking uracil, glucose medium lacking histidine, YPGAL and YPD, and its growth was compared to that of the isogenic wild-type and *GAL::HTP-UTP3* strains. Plates were incubated at 30°C for 3 days before photographing. There was a clear growth of the *HTP::UTP3* strain on medium containing 5-FOA, whereas this strain was unable to produce colonies on medium lacking histidine or uracil (Fig. 5.1). These data are consistent with loss of the integrated *URA3* marker in the *GAL:: HTP-UTP3* strain and both *HIS3* and *URA3* markers encoded by the plasmid containing the gene encoding the Cre recombinase. The *HTP::UTP3* strain showed clear growth on media containing either glucose or galactose (YPD and YPGAL), whereas the *GAL::HTP-UTP3* strain grew on YPGAL but not YPD medium (Fig. 5.1 B and C).

To compare the growth of strains expressing the N-terminal tagged HTP-Utp3 fusion protein, the C-terminal-tagged Utp3-HTP protein and a C-terminal tagged TAP- Utp3 protein with wild-type strain, serial dilutions of these strains were prepared and spotted onto glucose-based minimal media (YMM) plates. After incubation for 3 days, growth of the strain expressing the HTP-Utp3 fusion protein was similar to that of the wild type strain (Fig 5.1, panel D). In contrast, it can be seen that the growth of strains expressing the C-terminal tagged Utp3 fusion proteins Utp3-HTP and Utp3-TAP was restricted compared to the wild-type strain.

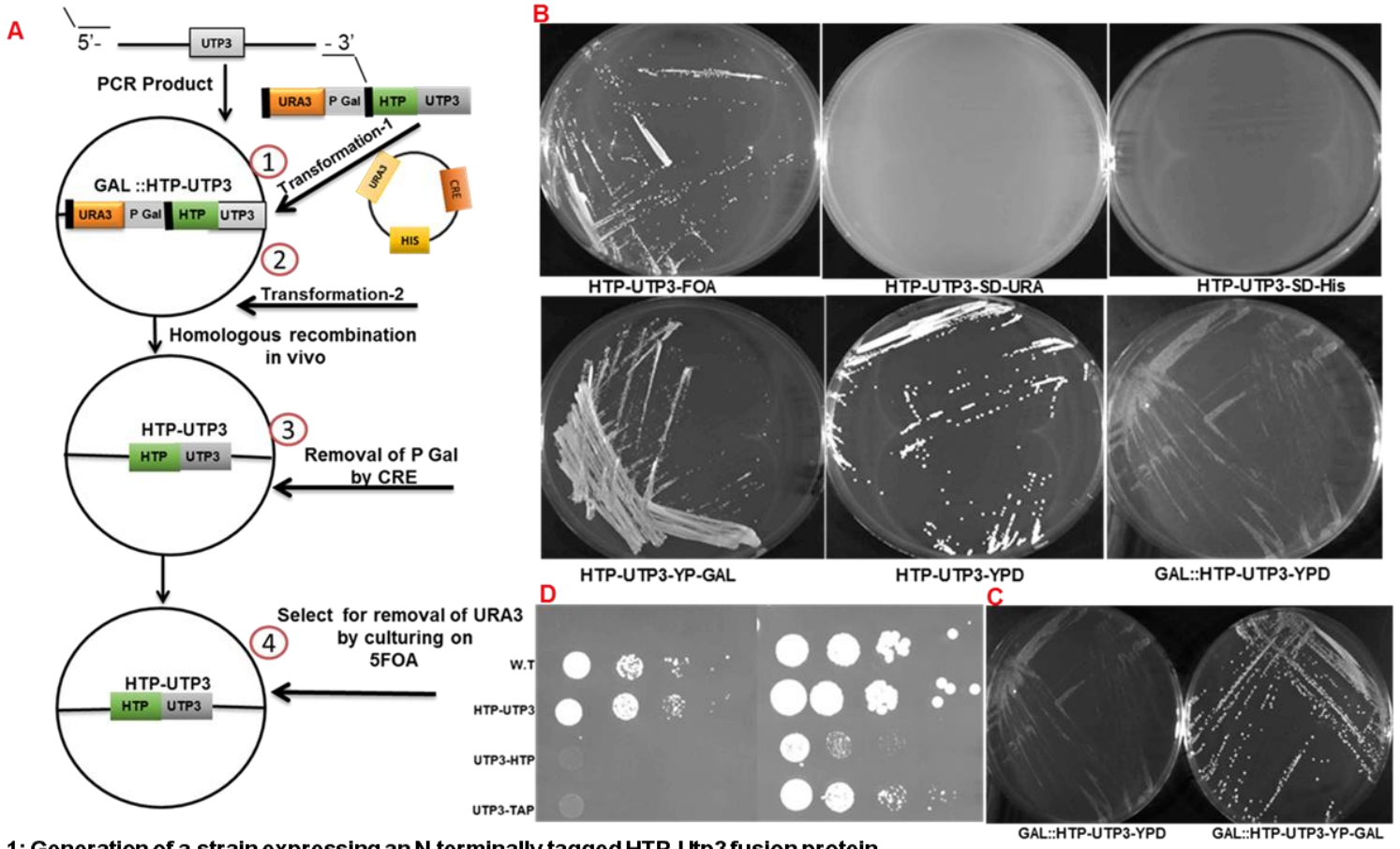

Figure 5.1: Generation of a strain expressing an N-terminally tagged HTP-Utp3 fusion protein.

A: Schematic diagram showing how the HTP-*UTP3* allele was generated. An initial integration event inserted a DNA sequence encoding the HTP tag at the 5' end of the *UTP3* ORF and placed the *UTP3* gene under the control of the *GAL* promoter. After introduction of a plasmid encoding the CRE recombinase, the *GAL* promoter was excised, repositioning the *Utp3* promoter upstream of the HTP-*utp3* ORF. Finally, the strain was purged of the CRE plasmid by growth on medium containing 5-FOA.

B: Growth of the candidate HTP-*utp3* strains was compared on galactose-based (YPGal) or glucose-based (YPD) medium, and on selective growth medium containing 5'FOA or lacking either uracil (SD-URA) or histidine (SD-HIS).

C: The parental strain *GAL::HTP-UTP3* can grows on YP-GAL medium but not on YPD medium

D: A wild-type strain (WT) and strains expressing different Utp3 fusion proteins (HTP-*utp3*, N-terminal HTP tag; *utp3*-HTP, C-terminal HTP tag; *utp3*-TAP, C-terminal TAP tag) were pre-grown in YPD medium and serial dilutions were spotted onto glucose-based minimal medium. Images were taken after 2 and 3 days (left and right plates, respectively).

5.1.1: Validation of HTP-UTP3

To confirm that the HTP cassette had integrated at the *UTP3* locus within the yeast chromosome, genomic DNA preparations were made from the *HTP::UTP3* yeast strain and two wild-type strains. PCR reactions were performed to amplify the *UTP3* locus in each strain. Primers were chosen that are complementary to sequences 500 base pairs (bp) upstream or downstream of the *UTP3* initiation codon. The PCR products were then resolved by electrophoresis through a 1% agarose gel.

As expected, the result of PCR assays showed the product obtained upon amplification of genomic DNA from the *HTP::UTP3* strain was longer than that obtained from the wild type strains (Fig. 5.2, panel A). A product of ~ 1.1kb was obtained upon amplification of the *UTP3* locus from genomic DNA of the wild type strains, consistent with yeast genomic DNA sequence. In contrast, the PCR product obtained upon amplification of genomic DNA from the *HTP::UTP3* strain was closer to ~1.5kb.

To compare the expression level of Utp3 fusion proteins, a wild-type strain and strains expressing the HTP-Utp3, Utp3-HTP or Utp3-TAP fusion proteins were grown in glucose minimal medium and cell lysates were prepared in SDS-loading buffer. The expression levels of the epitope-tagged Utp3 proteins were compared by resolving the lysates through 10% SDS-PAGE gels and performing Western blot analyses using the PAP antibody complex. The levels of Pgk1 in the samples were also compared as a loading control (Fig. 5.2, panel B). All three epitope-tagged Utp3 proteins migrated at approximately 130 kDa. Utp3 is a predicted molecular weight of 70 kDa, the TAP tag is approximately 21kDa and the HTP tag is approximately 18kDa. The Utp3 fusion proteins therefore migrated slower than predicted for proteins of their size. Expression of HTP-Utp3 was better than the other strains, making it the best subject for CRAC analysis.

To compare the expression level of Utp3 fusion proteins, a wild-type strain and strains expressing the HTP-Utp3, Utp3-HTP or Utp3-TAP fusion proteins were grown in glucose minimal medium and cell lysates were prepared in SDS-

loading buffer. The expression levels of the epitope-tagged Utp3 proteins were compared by resolving the lysates through 10% SDS-PAGE gels and performing Western blot analyses using the PAP antibody complex. The levels of Pgk1 in the samples were also compared as a loading control (Fig. 5.2, B). All three epitope-tagged Utp3 proteins migrated at approximately 130 kDa. Utp3 is a predicted molecular weight of 70 kDa, the TAP tag is approximately 21 kDa and the HTP tag is approximately 18kDa. The Utp3 fusion proteins therefore migrated more slowly than predicted for proteins of their size.

The Western data from further confirmatory tests showed that HTP-Utp3 was expressed well in cells grown in either glucose or galactose medium and at a higher level than the C-terminally tagged Utp3-HTP protein (Fig 5.2C). The experiment was repeated many times (data not shown). In all assays there was no defect in the protein expression of the HTP-Utp3 protein.

As *Utp3::HTP* strain had previously not generated any useful CRAC data and the *HTP::UTP3* strain made for this study grew well on both media and expressed the fusion protein consistently well, this N-terminally tagged strain was used for CRAC analysis.

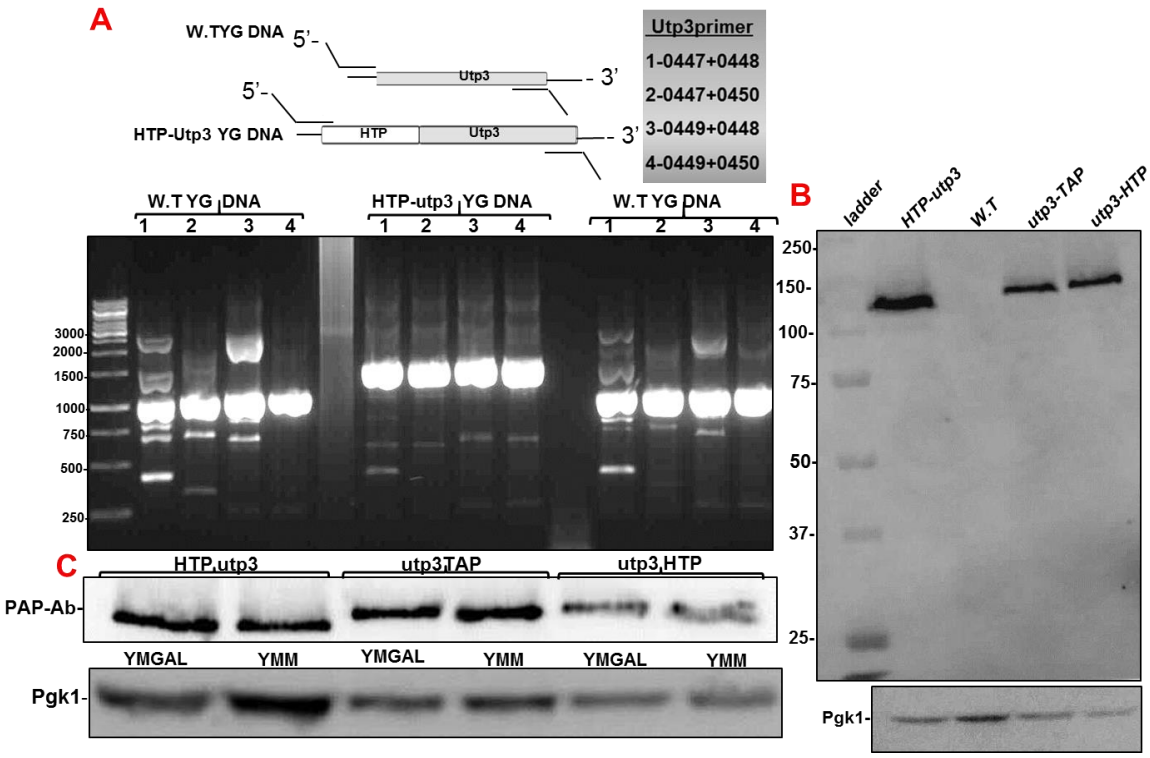

Figure 5.2: Verification of the HTP::*UTP3* strain.

A: PCR was performed on genomic DNA from two wild-type strains and the candidate HTP::*UTP3* strain, using primers complementary to sequences within the *UTP3* promoter region and the 3' end of the ORF. The amplicons were resolved by agarose gel electrophoresis and visualised by ethidium staining.

B: Western blot analysis of the relative expression level of Utp3 fusion proteins bearing epitope tags at either the N- or C-terminus during growth in rich medium. Total cell lysates were prepared under denaturing conditions from a wild-type strain and strains expressing different Utp3 fusion proteins and resolved by SDS-PAGE. Western blot analysis was performed using PAP antibody and an antibody specific to the Pgk1 protein. The electrophoretic mobility of molecular weight markers (in kDa) are indicated on the left.

C: Western analyses of Utp3 fusion proteins during growth in galactose-based (YMGAL) or glucose-based (YMM) minimal media. Total cell lysates were prepared by alkaline lysis and analysed by Western blot using PAP and anti-Pgk1 antibodies, as above.

5.1.2: CRAC analysis of Utp3

In order to investigate the exact RNA binding sites of Utp3 in growing yeast cells, the HTP::*UTP3* strain was subjected to CRAC by David Tollervey's laboratory (Welcome Trust Centre for Cell Biology, University of Edinburgh). The parental strain BY4741 was analysed in parallel as a negative control. The distribution of cDNA reads from the *HTP::UTP3* and wild-type strains throughout the yeast genome were mapped to the *S. cerevisiae* genome using the SGD genome browser tool. An overview of the number of cDNA reads obtained in these analyses are presented in Table 5.1.

It can be seen that the *HTP::UTP3* strain generated more than 10 times the reads than the wild-type strain overall. Clusters of cDNA reads were observed in the *HTP::UTP3* strain throughout the rRNA repeat and at the snoRNA loci, while the reads from the wild-type stain were more evenly distributed throughout the genome. Compared to the wild-type strain, the *HTP::UTP3* strain shows significant enrichment for 18S rRNA (2-fold), 5S rRNA (5-fold), snoRNA (U3 snoRNA 14-fold; other snoRNAs 1.9-fold) and less strikingly, mRNA (1.4-fold).Conversely, there is a lower frequency of hits from 25S rRNA (2-fold), 5.8S rRNA (4.7-fold), tRNA (1.2-fold), antisense reads (1.6-fold) and intergenic transcripts (2.9-fold). (Table 5.2andFig 5.4)

The highest number of cDNA reads in the *HTP::UTP3* strain for a specific gene was seen for the 18S rRNA sequence (more reads were observed for the mRNA transcripts but these reads represent hits to ~ 6,000 protein-coding genes). This is consistent with the known function of Utp3 in 18S rRNA synthesis. The highest enrichment of cDNA reads in the *HTP::UTP3* strain, compared to the wild-type strain, was seen for U3 snoRNA (a 14-fold increase compared to the wild-type strain). U3 snoRNA is a component of the SSU processome and is known to function as a chaperone in the formation of the pseudoknot structure at the 5' end of 18S rRNA (Hughes, 1996). There is also a 1.9-fold enrichment of cDNA hits for other snoRNAs. A significant enrichment of cDNA reads was also seen in the *HTP::UTP3* strain for 5S rRNA (5-fold) and for snRNAs (3-fold). Large numbers of reads were also observed for 25S rRNA and for tRNAs. However, the relative frequency of these reads were

higher in the control strain than the *HTP::UTP3* strain. 25S RNA is a well-recognised contaminant in CRAC analyses (Granneman et al., 2009).

cDNA reads obtained in CRAC analyses typically contain single nucleotide substitutions or deletions that arise as the reverse transcriptase passes through the site of crosslinking. Sequence alignment of the cDNa reads can therefore identify the protein-RNA contacts at nucleotide resolution. The sites of cross-linking can then be mapped onto the known secondary structure of the RNA. Given the known function of Utp3 in 18S rRNA synthesis and the distribution of cDNA reads obtained (Table 5.1), analyses of cross-link sites were focussed on the pre-rRNA transcript and on U3 snoRNA.

Sequencing generated from cross linked samples	*HTP::utp3*	BY4741
Total number of reads	62623	5306
Total number of mapped cDNAs	62623	5306
Number of reads without overlapping features	296	46
Number of overlapping genomic features	120151	11209
Number of sense reads within genomic features	69104	6704
Number of antisense reads within genomic features	50751	4459
37S rRNA	20815	2157
18S rRNA	10648	482
25S rRNA	7611	1459
5S rRNA	1362	28
5.8S rRNA	412	189
Other rRNA	46	13
antisense	1206	179
mRNA	14882	1009
U3 snoRNA	1945	11
Other snoRNAs	3478	177
tRNA	4317	500
Intergenic	1466	401
XUTs	267	43
SUTs	195	31
CUTs	115	13
snRNAs	235	10
ncRNAs	79	0
Pseudogenes	25	2

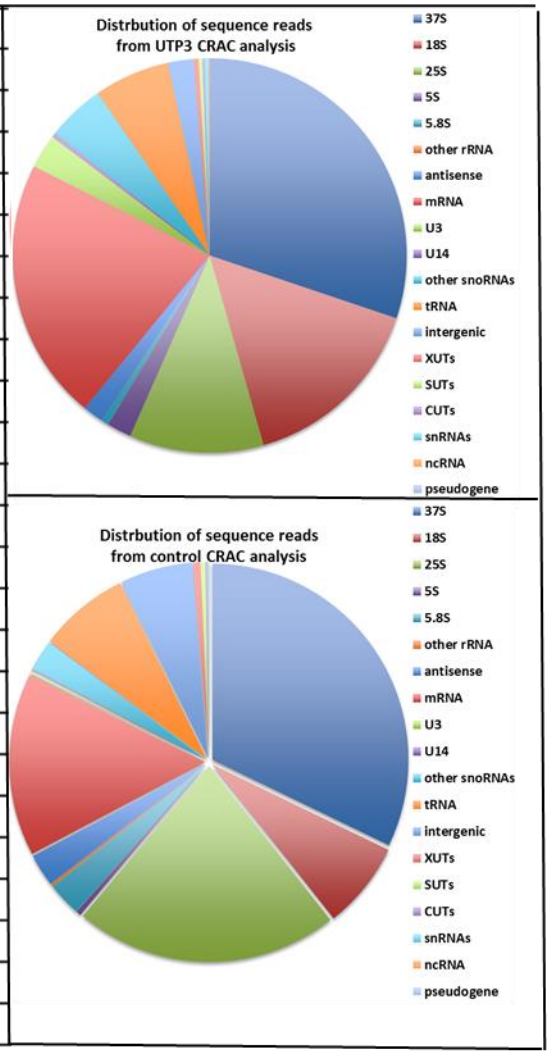

Table 5.1and Figure 5.3: An Overview of the Utp3 CRAC Data.

CRAC analyses were performed on the *HTP::UTP3* strain and the isogenic BY4741 control strain. The number of sense reads within specific classes of transcripts are given. Hits within 18S, 5.8S and 25S rRNA are also ascribed to the 37S rRNA precursor transcript, which (in addition to the mature rRNA sequences) also contains the 5' ETS, 3' ETS, ITS1 and ITS2 sequences. The pie charts show the distribution of mapped reads among the different classes of genomic features

	UTP3 %	control %	fold increase inUtp3 compared to wt
37S	30.1	32.2	
18S	15.4	7.2	2
25S	11	21.8	-2
5S	2	0.4	5
5.8S	0.6	2.8	-4.7
other rRNA		0.2	
antisense	1.7	2.7	-1.6
mRNA	21.5	15.1	1.4
U3	2.8	0.2	14
U14	0.4	0.2	
other snoRNAs	4.7	2.5	1.9
tRNA	6.2	7.5	-1.2
intergenic	2.1	6	-2.9
XUTs	0.4	0.6	-1.5
SUTs	0.3	0.4	-1.3
CUTs	0.2	0.2	
snRNAs	0.3	0.1	3
ncRNA	0.1		
pseudogene			

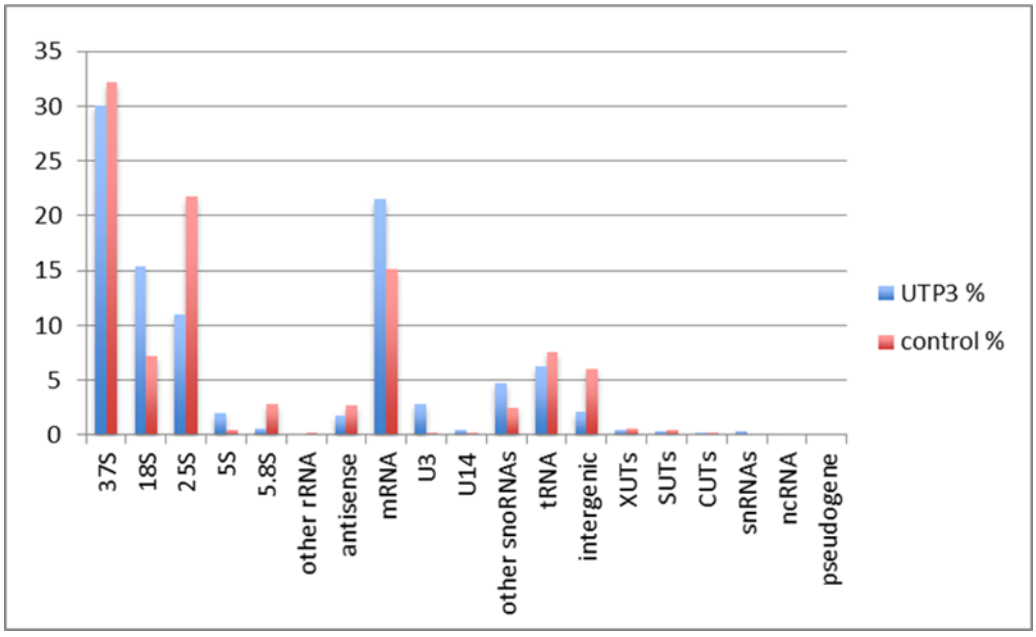

Table 5.2 and Figure 5.4: An average mapped reads comparison between the CRAC data for Utp3 and the negative control

The *HTP::UTP3* strain and a wild-type control were subjected to the CRAC protocol and sequence reads were mapped to the *S. cerevisiae* reference genome. Mapped reads were assigned to characterised genomic features. The histogram shows the distribution of mapped increase read of HTP-UTP3 10 times more than the control sample among the different classes of genomic features.

5.1.3: Analysis of Utp3 cross-link sites within U3 snoRNA

CRAC analysis of the *HTP::UTP3* strain and wild-type control cross-linked to U3 snoRNAs revealed a specific enrichment of Utp3 for the U3 snoRNA, encoded by the *SNR17A* and *SNR17B* genes (Fig. 5.5A upper panel). *SnR17A* is 5-10 fold more abundant than s*nR17B*. Each is 328 nucleotides long and they are 96% identical in the region of the mature RNA (Hughes et al., 1987; Beltrame and Tollervey, 1992). All of the cDNA hits were mapped to the U3 snoRNA exonic sequences. U3 snoRNA functions by base-pairing in four places on the 35S pre-rRNA within the 90S processome (Dutca et al., 2011; Kudla et al., 2011). Two of these interactions are to sites within the 5' external transcribed spacer (5'ETS), via the U3 3' hinge (nt 62-72) and U3 5' hinge (nt 39-48) regions. U3 snoRNA also base-pairs with the 5' end of pre 18S via U3 GAC/A' site (nt 4-11 and 15-22) and a short sequence far away at the 3' end of the 18S pre-rRNA via U3 Box A (nt 23-27) (Fig.1.2).

By mapping cDNA reads from the *HTP::UTP3* and wild-type strains to the spliced U3 snoRNA sequence, plotted as the number of reads at each nucleotide position, it is possible to identify the regions of Utp3/U3 snoRNA interaction. These indicate three general areas of Utp3 linkage to U3 snoRNA (Fig. 5.4A upper chart). Analysis of the sites of nucleotide substitution or deletion within these cDNA reads indicates a number of peaks that are distributed within the same regions of U3 snoRNA, at nt ~30-40, nt ~120 and nt ~270 (Fig. 5.5A lower chart). The majority of single nucleotide deletions were observed within a sequence of six consecutive uridine residues at positions 281-286. These residues map to the tip of Helix 3 within the secondary structure of U3 snoRNA (Fig. 5.5, panel B). This data suggests that Utp3 makes a direct contact with this structural element of U3 snoRNA. The nucleotide substitutions reveal evidence of HTP-Utp3 cross-linking to boxes A, B and C, but not D, of U3 snoRNA (Fig. 5.5B).

Helix 3 is directly adjacent to Box C and very close to Box B of U3 snoRNA in yeast (Mereau et al 1997). Thus, the vast majority of Utp3 binding to U3 snoRNA is in this area. Previous CRAC studies have identified the proximal

region of Helix 3 on U3 snoRNA as a binding site for the Nop1 and Nop56 protein components of the U3 snoRNP (Granneman, 2009). Very recently, a CRAC study has demonstrated clear *in vivo* binding of Utp10 to the Helix 3 region of U3 snoRNA (Hunziker et al., 2016). Utp10 is a very large protein component of the UtpA pre-ribosomal sub complex with a predicted molecular weight of 200 kDa (Krogan et al., 2004). It is suggested to have a role in recruiting the 3' domain of U3 snoRNA during processing of 35S pre-rRNA (Hunziker et al., 2016) and, consistent with this, Utp10 has been shown to form an extended arch that connects the base structure (containing the 5' ETS rRNA) to the head (containing the pre-40S rRNA region) of the 90S processome complex (Kornprobst et al., 2016).

There is some evidence from CRAC data that Utp3 binds to U3 snoRNA at Box A, which base pairs to the 3' end of the 18S rRNA coding sequence. However, Utp3 does not bind at the important sites within the 5' ETS where U3 snoRNA binds 35S pre rRNA at its 3' hinge (called Site A on the 5'ETS; nt 281-291) or, its 5' hinge (called Site B on the 5'ETS; nt 470-479) (Fig. 1.1).

Figure 5.9 shows a representation of the recently published three-dimensional structure of the 90S processome complex from the thermophilic eukaryote *Chaetomium thermophilum* (Kornprobst et al., 2016) . In this structure, helix 2 and helix 4 of U3 snoRNA are well resolved and protrude from the body of the particle while its 5' region penetrates into the core of the 90S complex to base-pair with the 5' ETS RNA. Three identified sites of Utp3 binding to U3 snoRNA in the Box B and Box C region are labelled (A110, U112 and U261). These nucleotides appear close together on the secondary structure of U3 snoRNA (Fig 5.5B) but relatively separated in the three-dimensional model, perhaps due to variability in the 3' domain of U3 snoRNA between the two species (Phipps et al., 2011). Other Utp3 binding sites on helix 3 and Box A of U3 snoRNA are deeper inside the 90S processome base. Notably, the processome subunits Utp6 and Utp21 can be seen close to this region of the 90S structure (Fig. 5.9).

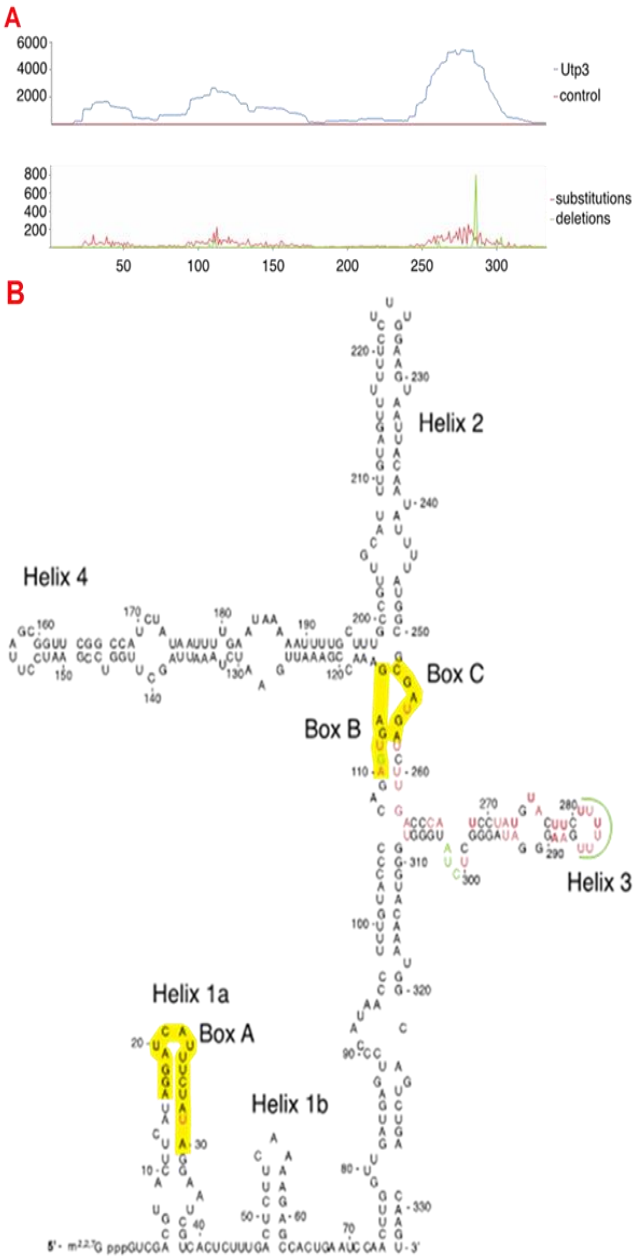

Figure5.5: **Utp3 contacts within U3 (snR17A) snoRNA**

A: Profiles showing the distribution of sequence reads across the spliced RNA sequence (no reads were observed within the U3 snoRNA intron). Upper profiles Comparison of the distribution of reads observed in the *HTP::UTP3* strain (shown in blue), compared to the wild-type control (shown in red). Lower profiles Comparison of the distribution of nucleotide substitutions (shown in red) and deletions (shown in green) observed within HTP-Utp3 reads (shown in red).

B: Secondary structure of U3 snoRNA (Granneman et al., 2009), showing the distribution of nucleotides that are substituted (shown in red) or deleted (shown in green) within the CRAC reads. The predominant single nucleotide deletions are located within the run of uridine residues at the tip of helix 3. The position of the highly conserved sequence boxes A, B and C are indicated.

5.1.4 Analysis of Utp3 cross-link sites within the 35S pre-rRNA sequence

Our CRAC analysis showed significant levels of cross-linking between HTP-Utp3 and sequences within the 35S pre-rRNA, specifically within the 5' ETS region and the 18S rRNA sequence. Panel A of figure 5.6 shows the frequency distribution of cDNA sequence reads along the length of the 35S pre-rRNA in the *HTP::UTP3* and wild-type strains. There is a concentration of cDNA reads from the *HTP::UTP3* strain at many sites along the 35S pre-rRNA sequence, in particular through the 5' ETS and 18S rRNA. The peak towards the 3'-end of the 25S rRNA sequence is a common false positive or contaminant (Granneman et al., 2009) .

Analyses were undertaken to map nucleotide deletions and substitutions within the cDNA reads to the 5' ETS and 18S rRNA sequences. Nucleotide deletions were rare in the cDNA reads of the 5' ETS sequence from the *HTP::UTP3* strain but substitutions at U116, A134, C204, G209 and U438 were the most predominantly observed alterations (Fig. 5.6 B). Similar analyses of cDNAs derived from 18S rRNA sequence revealed prominent substitutions at A156, A518, A544, U582, A1203, U1340 and U1558, and deletions at U132, U493, U1239, U1361, U1491 and U1566 within the HTP-Utp3 reads (Fig. 5.6, panel C).

The predominant sites of cross-linking to Utp3 (U116, A134, A184, C204, G209 and U438) were mapped onto the secondary structure of the 5' ETS (Fig. 5.6). The nucleotides that are cross-linked to Utp3 are clustered within the secondary structure of the 5' ETS. A134 lies on the opposite side of a helix to U116, while A184, C204 and G209 are located towards the terminal loop of another helix. These sites are predominantly towards the 5' end of the 5' ETS. Notably, there is no apparent proximity in the secondary structure between the residues within the 5' ETS that are cross-linked to Utp3 and the residues that interact with U3 snoRNA (nucleotides 279-291 and 469-479, highlighted in yellow in Fig. 5.6) or the A_0 or A_1 cleavage sites (nucleotides 601 and 700).

The interaction between Utp3 and the 5' end of the 5' ETS is in good agreement with the findings from two recent studies on the assembly of the

processome complex, which used mass spectrometry to identify yeast proteins that copurified with pre-rRNA fragments that extend to variable extents from the 5’ end (Chaker-Margot et al., 2015; Zhang et al., 2016). Both groups demonstrated the presence of Utp3 at an early stage during the step-wise recruitment of the UtpA complex, the UtpB complex, the U3 snoRNP and the Mpp10 sub-complexes to the 5’ ETS. The rRNA truncations used by Zhang et al. (Zhang et al., 2016) detected low levels of Utp3 associated with pre-rRNA fragments extended to +309 but not on fragments extended to +281. Our CRAC data suggest that Utp3 binds to the 5’ region of the 5’ ETS, consisting of the first ~ 260 nucleotides. However, Zhang et al. noted that protein binding to sections of the 5’ ETS was initially weak and easy to disrupt but became more stabilised within longer transcripts.

An earlier study using CRAC analysis reported that all seven subunits of the UtpA complex show marked binding to the 5’ proximal region of the 5’ ETS, consistent with a role for the UtpA complex in early pre rRNA processing (Hunziker et al., 2016). Utp10, a large protein component of the UtpA subcomplex, was shown to cross-link with the 5’ ETS with a peak at nucleotide 110. This is very close to the peak found for Utp3 at U116. In addition, Utp5, another UtpA subunit, cross-linked to the 5’ ETS maximally at nucleotide 130, which is close to the Utp3 cross-link site at A134 and is base-paired to U116. These data suggest that Utp3 may be bound to the 5’ ETS in close proximity to components of the UtpA complex. UtpA has been proposed to stimulate the recruitment of the U3 snoRNP to the 5’ ETS via the interaction between Utp10 and the 3’ domain of U3 snoRNA (Hunziker et al., 2016). The CRAC data showed prominent binding of Utp3 to both the 5’ domain of the 5’ ETS and the 3’ domain of U3 snoRNA (Fig. 5.5 and 5.7)., taken together with the mass spectrometry data demonstrating the association of Utp3 with the U3 snoRNP particle and the 5’ ETS region (Chaker-Margot et al., 2015) suggests comparable roles for Utp3 and UtpA in U3 snoRNP recruitment and/or stabilisation.

This study has shown that Utp3 interacts with Utp21 and Utp6, two components of the UtpB subcomplex (Figs. 4.3 and 4.4). The UtpB

subcomplex is a large structure that is proposed to stabilise U3 snoRNA base pairing with the 5’ ETS at nucleotide +280 (site A) via Utp1 (Hunziker et al., 2016). The UtpB subcomplex has also been shown to contact the 5’ ETS via Utp18 at nucleotide +90, adjacent to the UtpA complex binding site. Neither Utp21 nor Utp6 were observed to bind to the 35S rRNA (Hunziker et al., 2016), consistent with earlier findings in this study that the interaction between Utp3 and Utp21 or Utp6 is independent of RNA (Fig. 4.5).

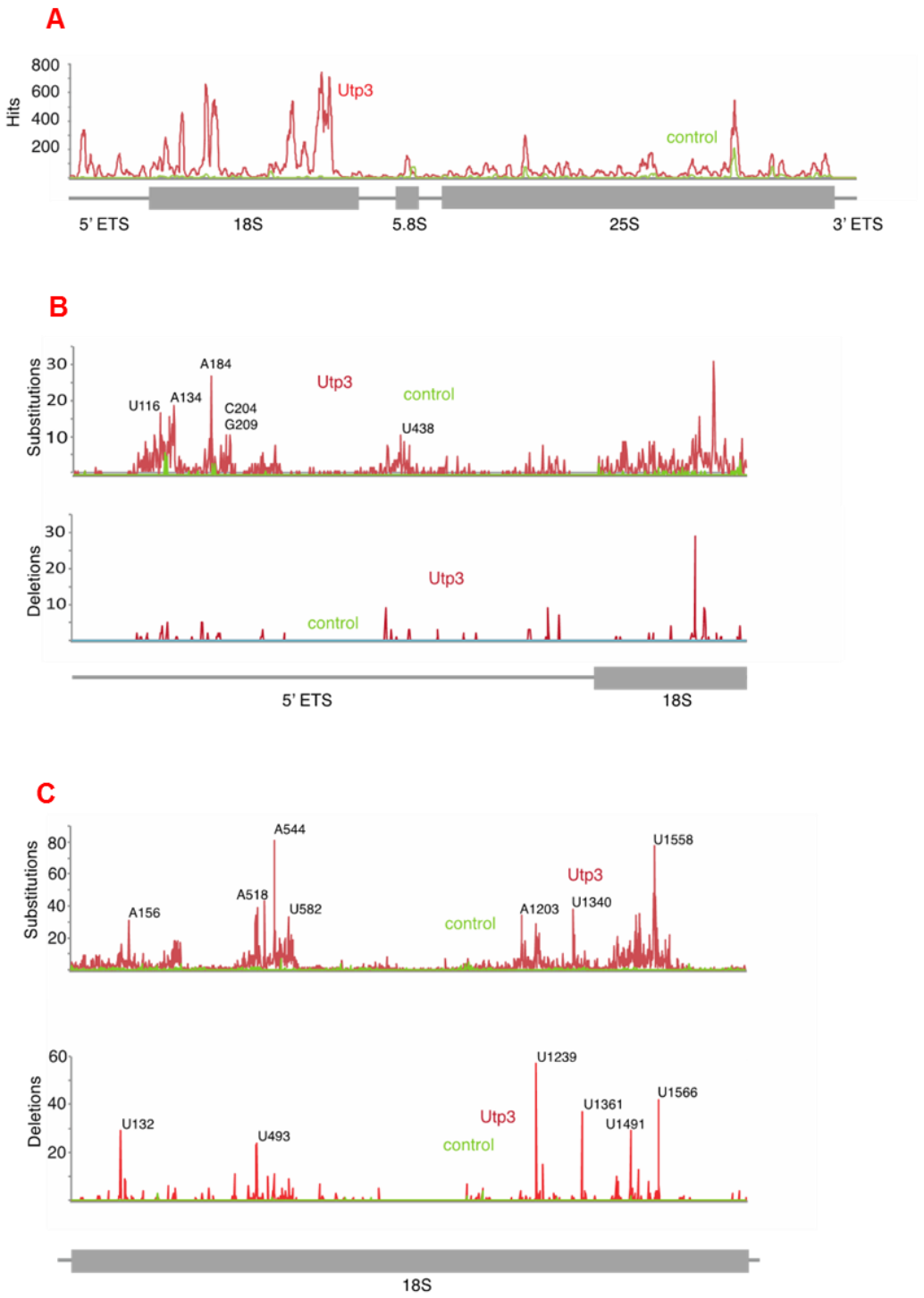

Figure 5.6: Utp3 crosslinks within the 35S rRNA

A: Distribution of reads across the 35S rRNA sequence obtained with the *HTP::UTP3* strain (shown in red) and the wild-type control (shown in green).

B: Distribution of substitutions and deletions across the 5' ETS region. Utp3 reads are shown in red; control reads are shown in green. Prevalent sites are indicated.

C: Distribution of substitutions and deletions across the 18S rRNA sequence. Utp3 reads are shown in red; control reads are shown in green. Major sites of crosslinking to Utp3 are indicated.

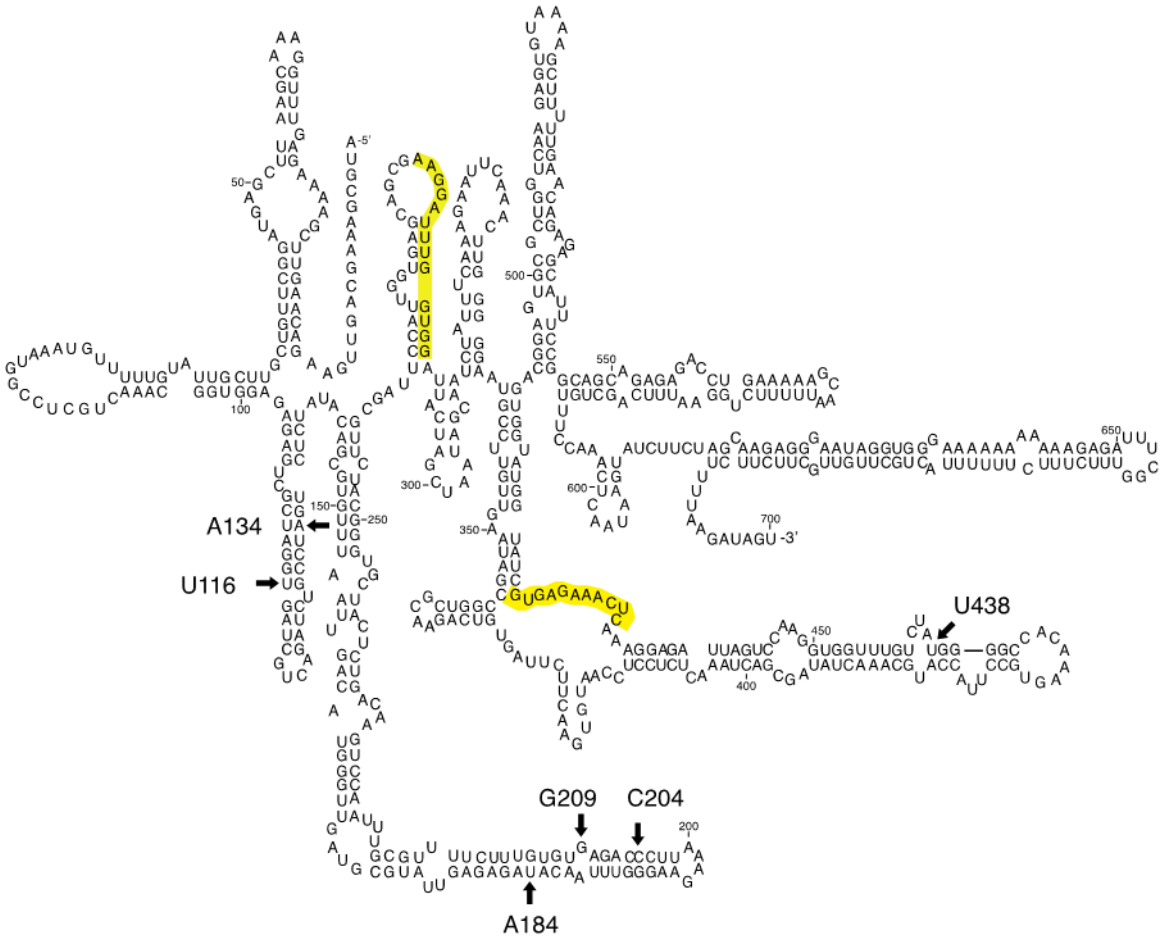

Figure5.7: Location of Utp3 crosslinks within the secondary structure of the 5' ETS

Prevalent nucleotide substitutions observed within sequence reads obtained with the HTP-Utp3 strain, indicative of sites of cross-linking, are mapped to the secondary structure of the 5' ETS region of pre-rRNA from *S. cerevisiae* (adapted from Yeh and Lee, 1992). Nucleotides known to interact with U3 snoRNA (Beltrame and Tollervey, 1992) are indicated in yellow.

.

The most frequent sites of HTP-Utp3 cross-linking to 18S rRNA were mapped onto the secondary structure (Cannone et al., 2002) (Fig. 5.8). In contrast to the contacts within U3 snoRNA or the 5' ETS, Utp3 cross-linked to nucleotides within 18S rRNA that are predominantly located in stem loops and bulges within the secondary structure. Nucleotides that are not involved in base-pairing may be structurally less restrained, which may be significant for contacts with Utp3. Figure 5.8 shows six sites mapped on to the 5' domain of 18S rRNA (between Helix 7 and 8, Helix 8, Helix 16, Helix 17 and Helix 18) and seven sites mapped on to the 3' major domain of 18S rRNA (helix 31, helix 33, helix 41 and two sites in helices 39 and 42).

CRAC analysis did not reveal any significant cross-linking sites for Utp3 in the 3' minor domain, the central domain that contains sites of interaction with U3 snoRNA, or the internal transcribed spacer, ITS1 (see Fig. 5.5, panel A). These are the last three domains of the pre-40S subunit to be completed (de la Cruz et al., 2015).

The study by (Zhang et al., 2016). on the assembly pathway of the 90S particle proposed that completion of transcription of the nucleotides that comprise helix 44 within the mature 18S rRNA (the long penultimate helix in Fig. 5.8) triggers a marked change in composition of the RNP particle, with the release of many proteins and snoRNAs. However, all proteins that are bound to the 5' ETS and U3 snoRNA (including Utp3) remain stably associated with the transcript throughout the assembly of the 90S pre-ribosomal particle, consistent with similar mass spectrometry data showing little change in abundance of these early factors (Chaker-Margot et al 2015). The U3 snoRNA, the 5' ETS region and their associated proteins and subcomplexes are important in stabilising and encapsulating the 35S pre-rRNA while it is modified and folded (Kornprobst et al., 2016).

Four of the Utp3 cross-link sites to nucleotides within the 18S rRNA (U132, A156, U493 and A518) that lie within the 5' domain could be mapped on to the three-dimensional model of the 90S processome complex of *C. thermophilum*. The 18S rRNA coding sequence is not in its mature three-dimensional

structure within the 90S processome, but nevertheless the 5’ domain is close to its conformation in the mature state in the ribosome. The 5’ domain of the 18S rRNA is found in the so-called “head” region of the 90S processome, while the 5’ ETS, the UtpA and UtpB subcomplexes are found in the “base” and “body” of the 90S processome (Kornprobst et al., 2016). Sites corresponding to Utp3 binding to the 3’ major domain are not resolved in the current structure of the 90S processome, but these would most probably lie within the so-called body region that contains Utp21 and Utp6 (Fig. 5.9) (Kornprobst et al., 2016).

The 90S pre-ribosomal particle is subject to biochemical and structural alterations throughout its existence and the mature conformation of the 40S particle is only achieved after A_0, A_1 and A_2 cleavage within the 5’ ETS and ITS1 regions, followed by release of U3 snoRNP by Dhr1 (Sardana et al., 2015) and subsequent cleavage at site D in the cytoplasm. It is currently not possible from the CRAC data alone to say at what point during the assembly of the 90S procesome that specific contacts between Utp3 and nucleotides within the 18S rRNA sequence occur or to speculate on the importance for Utp3 function.

Fig. 5.10 shows the location of nucleotides that cross-link to Utp3 within the three-dimensional structure of the 40S ribosomal subunit. Notably, the relative orientation of the recently published structure of the 90S processome is inverted when compared to the established models of the 40S subnit; the 5’ domain of 18S rRNA, which is positioned in the “head” of the 90S processome structure, is in the lower “body” region of the 40S subunit, while the 3’ major domain of 18S rRNA lies within the “body” of the processome and the “head” of the 40S subunit. Seven of the predominant cross-link sites (U1361, U1566, U1491, U1558, U1340, U1239 and A1203) are located in the 3’ major domain within the head area of the 40S subunit, three contacts (A518, U582 and A544) are seen close to the platform region where codon recognition occurs and two sites within the 5’ domain (U132, A156) are located within the body and foot of the 40S subunit. These sites are broadly distributed through the three-dimensional structure of the ribosome. However, Utp3 is not present in the mature 40S SSU and therefore these sites reflect previous contacts with

Utp3 in the core of the 90S processome. As there are contact sites with Utp3 along the 18S rRNA sequence that must be occurring within the processome, this suggests either that Utp3 is extremely elongated through the processome structure in a similar way to Utp10 or that Utp3 contacts 18S at multiple sites as it is being transcribed. From what we know about Utp3, it is likely that the sites of cross-linking within the 35S pre-rRNA reflect contacts made when the protein is stably associated with the 5' ETS region in the body/base of the processome complex that includes UtpA and UtpB subcomplexes, at early stage during transcription.

In conclusion, CRAC analysis has demonstrated that Utp3 binds both to the U3 snoRNA and to multiple sites on the 5' ETS and 18S region of the 35S pre rRNA. However, the topological information regarding Utp3 interactions does not readily offer an explanation of its function, except that it may have a generalised role in the folding of the pre-rRNA. It could also be speculated from what is known about the role of the C1D domain, the RNA binding capacity of the Utp3 C-terminal domain, and the role of U3 snoRNA, that Utp3 functions as a pre-rRNA chaperone protein to promote the correct folding of the pre-rRNA.

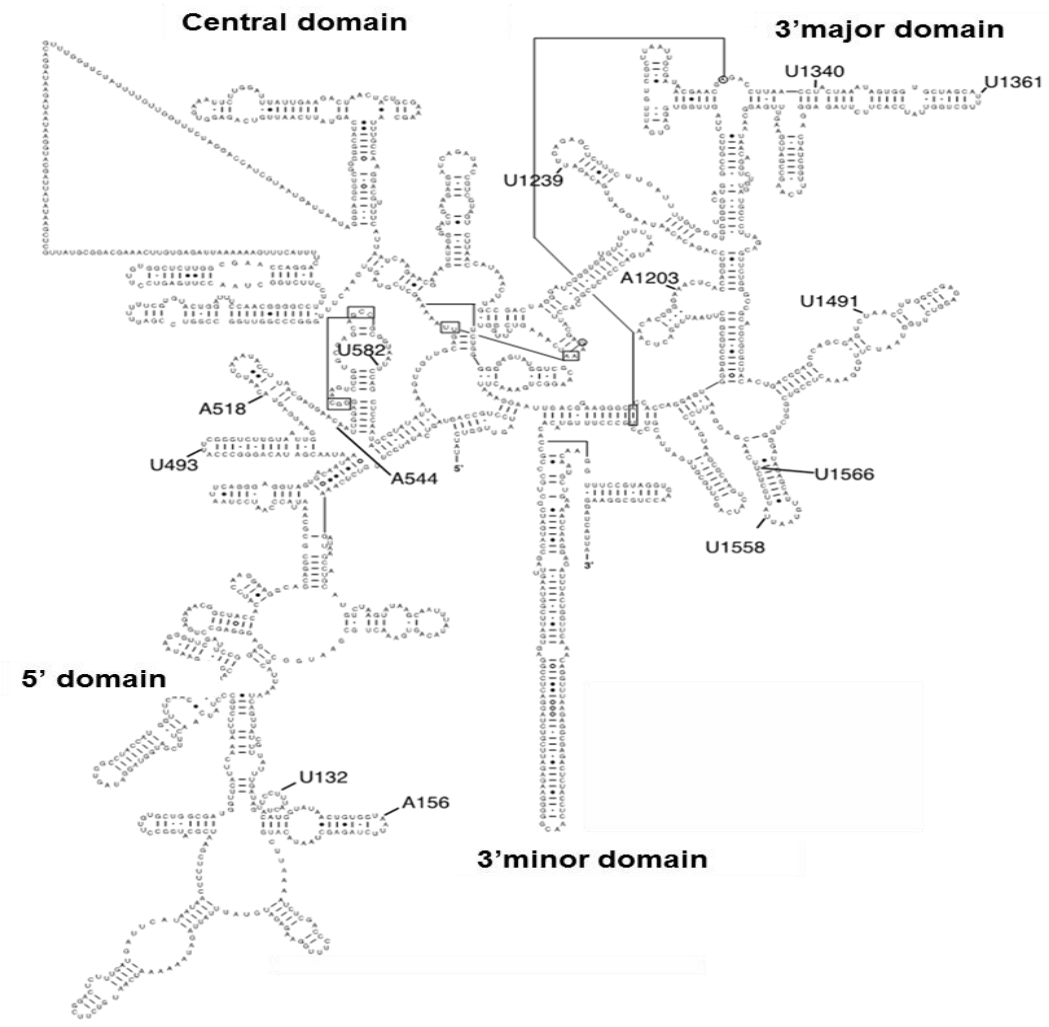

Figure 5.8: Location of Utp3 cross-linking sites within the secondary structure of 18S rRNA

Schematic of the secondary structure of 18S rRNA from *S. cerevisiae* (Cannone et al., 2002). The location of prevalent crosslinks to Utp3 are indicated.

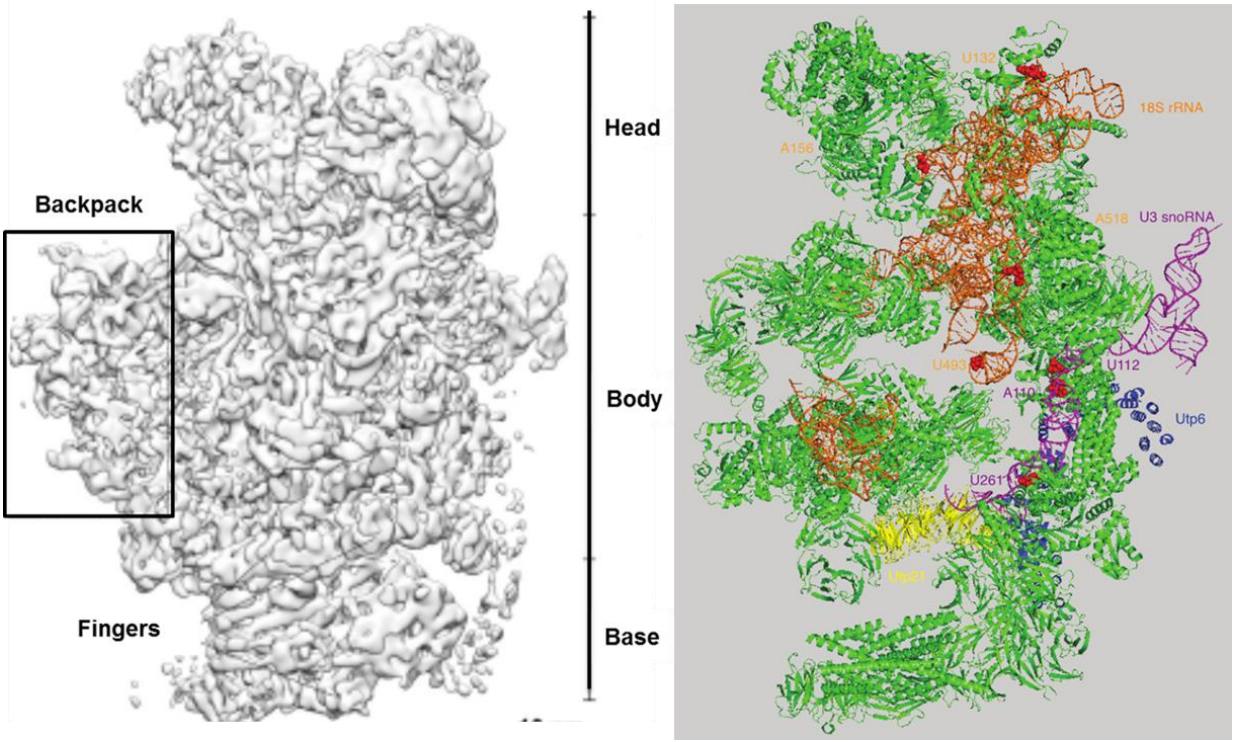

Figure 5.9: Sites of Utp3 crosslinking within the three-dimensional structure of the 90S processome.

The three-dimensional structure of the 90S processome from *Chaetomium thermophilum* (Kornprobst et al., 2016) is shown. Identifiable protein folds that could be fitted to the cryo-electron density map are shown in green, resolved regions of rRNA are indicated in orange and portions of U3 snoRNA that could be fitted to the model are shown in purple. Residues within the *C. thermophilum* 16S RNA sequence that are homologous to nucleotides within 18S rRNA of *S. cerevisiae* that were identified to crosslink with Utp3 are highlighted as red spheres. Resolved regions of proteins Utp21 and Utp6, which are thought to contact Utp3 directly, are shown in yellow and blue, respectively. The image was generated from PDB entry 5JPQ, using MacPyMOL software. The perspective is shown to allow direct comparison with published images (Kornprobst et al., 2016).

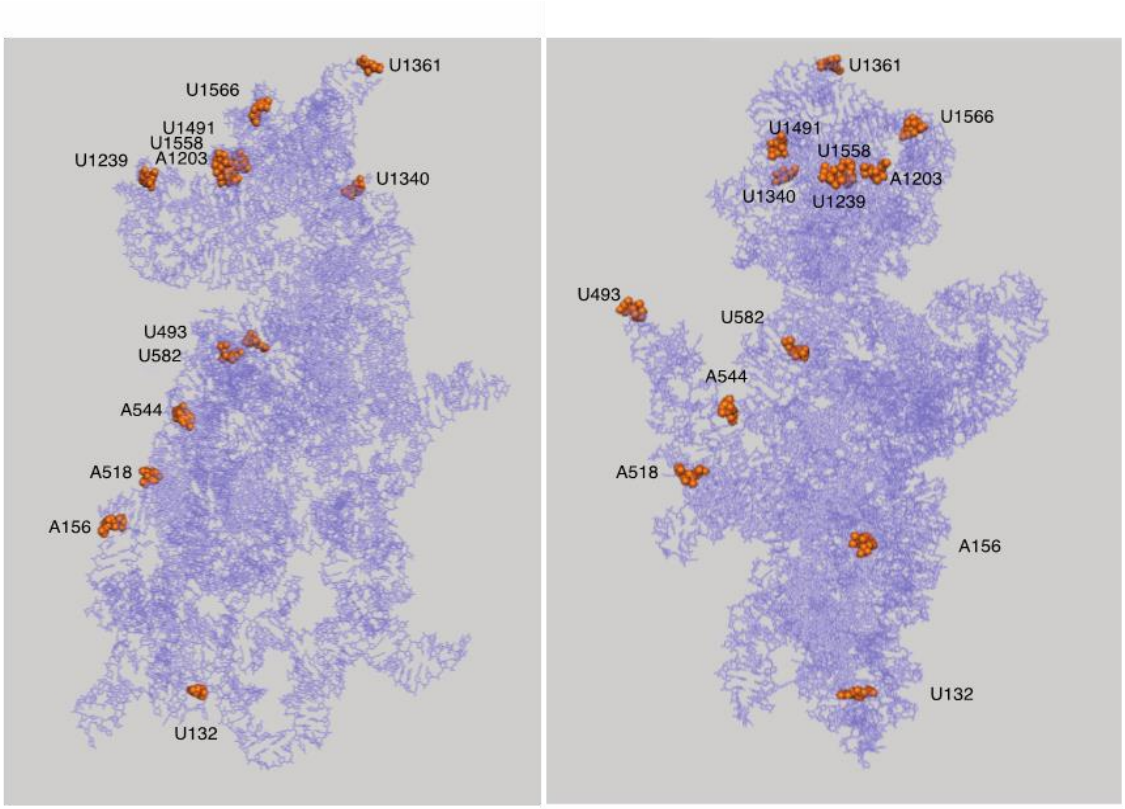

Figure 5.10: Location of Utp3 crosslink sites to 18S rRNA within the three dimensional structure of the 40S small ribosomal subunit.

The three-dimensional structure of 18S rRNA is shown in blue as a line diagram. Utp3 crosslink sites within 18S rRNA are highlighted as spheres and shown in orange. Two distinct perspectives are shown that differ by rotation around the vertical axis. Images were generated with MacPyMOL, using PDB structure 3O2Z (Ben-Shem et al., 2010).

5.2: The CTD domain of Utp3 possesses DNA binding activity

The CRAC analyses demonstrate that Utp3 shows RNA binding to 35S pre-rRNA and U3 snoRNA in vivo. This raises the question as to which region of Utp3 is specifically involved in this interaction.

Several studies have found that Rrp47 and its human homologue (known as C1D) bind both RNA and DNA (Nehls et al., 1998; Stead et al., 2007; Costello et al., 2011). The CTD of Rrp47 is required for its ability to bind RNA and DNA *in vitro*. The stretches of basic residues within the C-terminal region of Utp3 make it a likely candidate for nucleic acid binding. BLAST analysis suggested that the C-terminal sequence of Utp3 has no significant homology to other proteins (Fig. 5.11). Further work was undertaken in this study to investigate the role of the C-terminal domain of Utp3 in RNA/DNA binding. Electromobility band-shift assays (EMSA) were performed on purified recombinant Utp3 CTD proteins expressed in *E. coli*.

The CTD domain of Utp3 was expressed in *E coli* and purified as a 6×His-tagged fusion by affinity chromatography using Ni-NTA beads and ion exchange chromatography using SP sepharose beads (Fig. 5.12, panel A). Recombinant His-tagged Rrp47 (Costello et al., 2011) was purified and assayed in parallel as a positive control, while lysate from vector-transformed *E. coli* was subjected to the same purification procedure and used as a negative control. EMSA gels were analysed either by staining with ethidium bromide or by electrotransfer and hybridisation analyses, using a radiolabelled probe. Consistent with previous studies (Stead et al., 2007), recombinant Rrp47 showed a clear band-shift when assayed with linearised plasmid DNA (Fig. 5.11, panel B). There was a similar clear band-shift of DNA incubated with the CTD domain of Utp3, while no shift was detected for the negative control. The CTD construct of Utp3 therefore appears clearly to bind DNA. The pattern is consistent with binding at multiple sites within the DNA, as there is no shift to a single sized product band (Fig. 5.12).

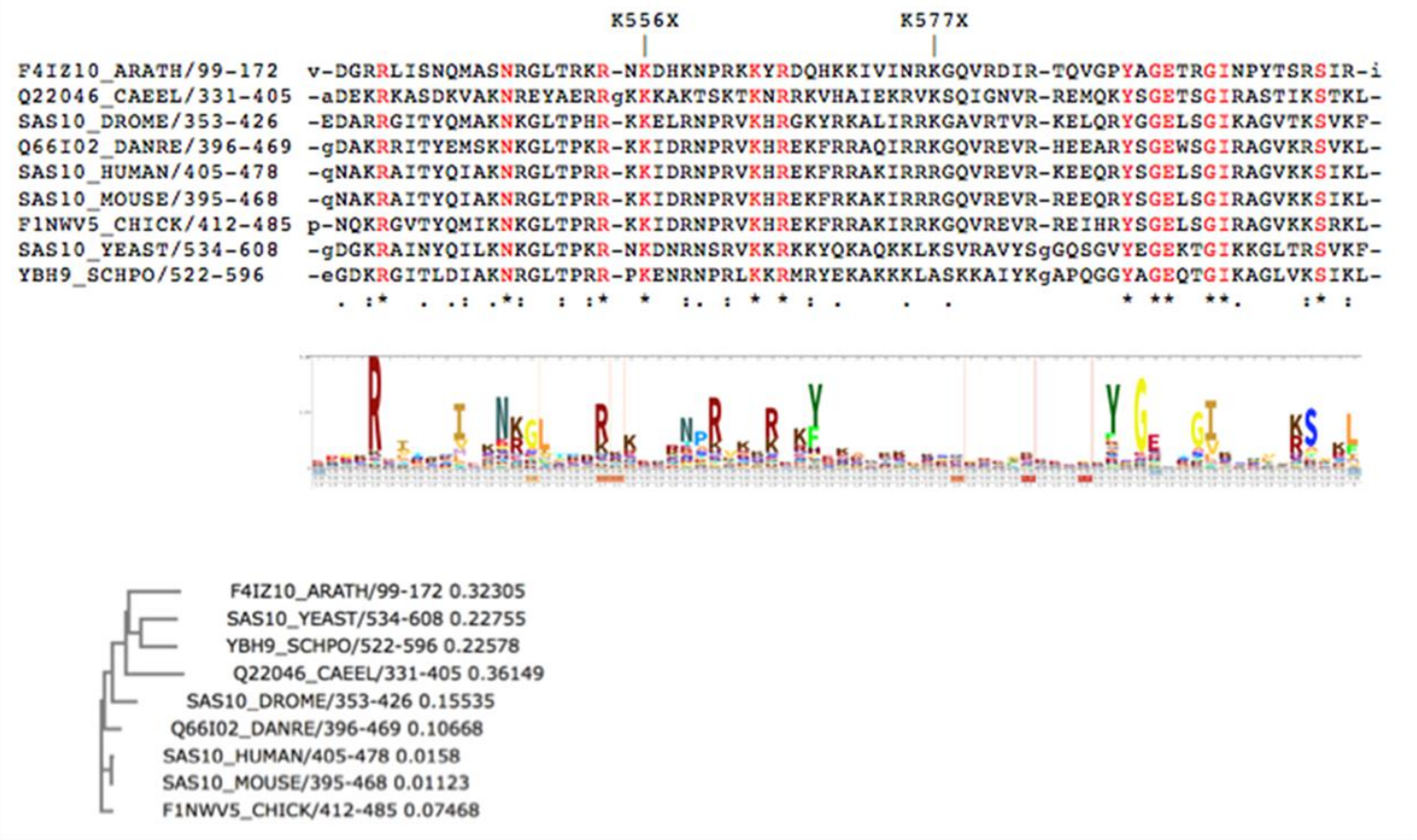

Figure 5.11:. A sequence alignment of the C-terminal domain of Utp3.

Utp3 sequences from representative model organisms were selected from Pfam (PF09368) and aligned using CLUSTAL Omega. Residues that are conserved in all selected sequences are highlighted in red and indicated with an asterisk. Highly conserved and weakly conserved residues are indicated with : and . symbols, respectively. The position of residues that were mutated by site-directed mutagenesis to generate C-terminal truncation mutants (K556X, K577X) are indicated above the alignments. The middle panel shows an HMM logo of the Utp3 CTD (taken from Pfam). The lower panel shows a phylogenetic tree of the sequences (generated in (CLUSTAL).

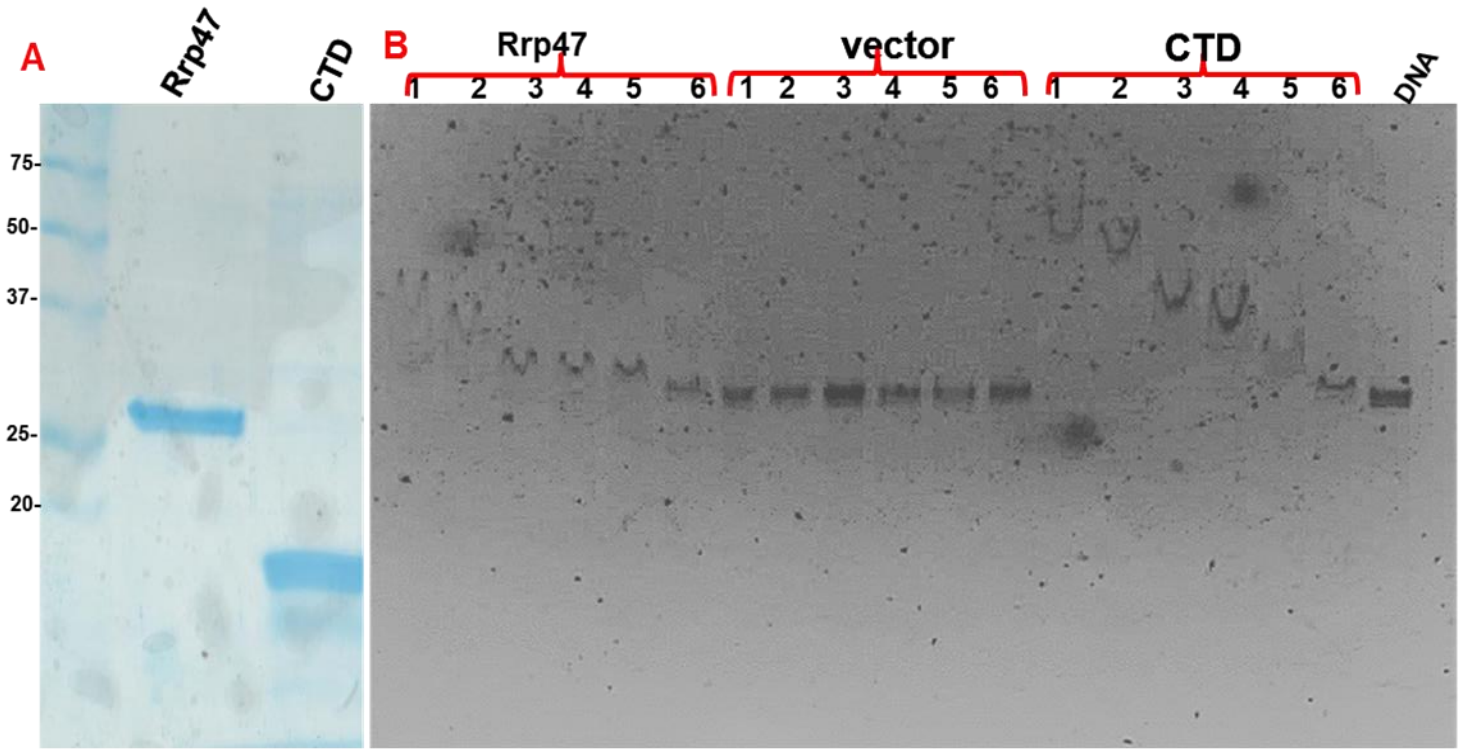

Figure 5.12: The CTD domain of Utp3 has DNA-binding activity in vitro.

A: Purification of recombinant His(6)-tagged fusions of Rrp47 and the CTD of Utp3. Proteins were purified from bacterial cell lysates by affinity chromatography using Ni-NTA beads, ion exchange chromatography using SP sepharose beads and subsequent dialysis. An aliquot of the purified protein was resolved by SDS-PAGE and protein visualised by staining with colloidal Coomassie blue. A control sample from cells transformed with the cloning vector was purified under the same conditions.

B: Electromobility shift assays (EMSA) of purified recombinant Rrp47 and the C1D domain of Utp3, using linearised plasmid DNA. Two-fold serial dilutions were made from purified protein samples, such that the incubation mixtures contained the following protein concentrations: lane 1, 10 uM; lane 2, 5 uM; lane 3, 2.5 uM; lane 4, 1.25 uM; lane 5, 625 nM; lane 6, 312 nM. For the vector control, serial dilutions were made from the neat sample obtained after dialysis. Incubation mixtures were resolved by electrophoresis through 1% agarose gels and the DNA was visualised by staining with ethidium bromide. A reaction mixture in the absence of protein (labelled DNA) is shown on the right hand side.

Chapter 6 : Utp3 localisation to the nucleolus

Utp3 protein must pass through nuclear pores via complexes (NPCs) to reach the nucleus and subsequently be localised in the nucleolus, which is the site of Utp3 activity in all stages of the cell cycle. The C-terminal region of Utp3 has been suggested to carry putative nuclear localisation signals (Kamakaka and Rine, 1998). These short sequences, typically rich in positively charged lysine or arginine residues, bind a protein for subsequent import into the cell nucleus. Loss of an NLS sequence that is critical for nuclear accumulation would result in aberrant cytosolic accumulation of the protein.

Not all proteins that are imported into the nucleus have an NLS. The nuclear import of ribosomal proteins through nuclear pores may be passive in the case of molecules <40kDa (Raices and D'Angelo, 2012) and the import of larger proteins may be facilitated through association with other proteins that contain an NLS (Bange et al., 2013). NLSs are recognised and bound by transport protein receptors, known as importins. For example, the yeast importin-α homologue Srp1 binds to the NLS of its protein substrates and forms a cargo complex with Kap95 (karyopherin-β) to mediate translocation of proteins through the nuclear pore (Mohr et al., 2009; Grunwald et al., 2011).

To address the contribution of specific regions of the Utp3 protein to its correct subcellular localisation, all of the generated Utp3 deletion mutants were sub-cloned into a related plasmid in which the N-terminal zz tag was replaced by a GFP domain. These plasmids carry a *HIS3* marker and express N-terminally tagged GFP fusion proteins under the control of a constitutive *RRP4* promoter. To test whether the mutant GFP-Utp3 fusion proteins were functional, the plasmids were transformed into a wild-type strain (P852) and the isogenic *GAL::UTP3* (strain P851), and spot growth assays were carried out on glucose-based minimal medium. There was no growth defect observed for the *GAL::UTP3* transformant harbouring a plasmid encoding the wild-type *UTP3* sequence downstream of the GFP tag, indicating that the function of Utp3 was not affected by the GFP tag. As observed previously (Fig. 3.9), the ΔCTD truncation mutant failed to complement the conditional growth phenotype of

the *GAL::UTP3* allele. However, some mutations (such as the K556X mutant and the ΔN mutant) that were non-functional in this genetic assay as zz fusion proteins nevertheless supported growth of the *GAL::UTP3* on glucose-based medium. To determine whether these effects may reflect differences in protein expression levels, whole cell lysates were prepared from strains expressing each of the GFP-Utp3 fusions, resolved by SDS-PAGE through 10% acrylamide gels and analysed by Western blotting using an anti GFP antibody. These analyses revealed a clear signal with GFP antibody for the wild-type protein and all of the mutants (Fig. 6.1, panel C). The reason why the genetic complementation by some of the Utp3 mutants is dependent upon the epitope tag is currently not clear.

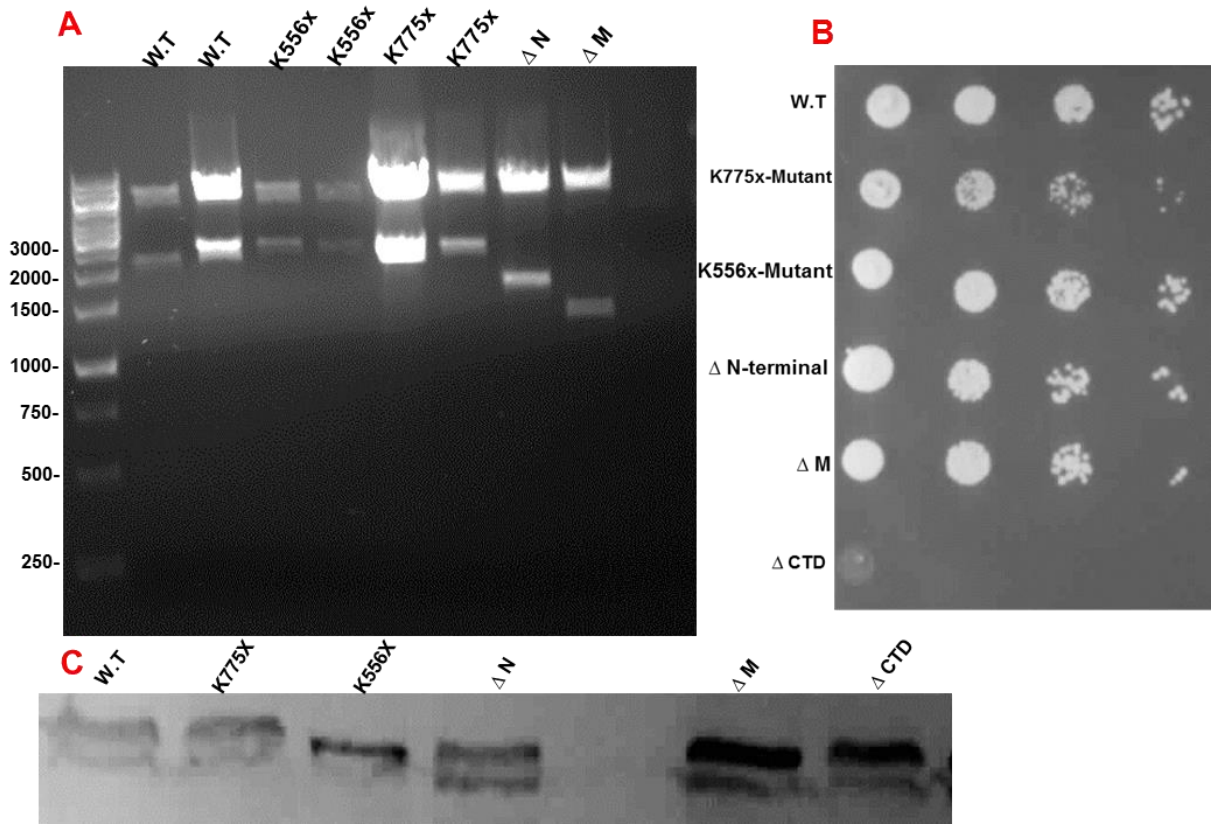

Figure 6.1 : GFP-UTP3 strain was not affected by GFP-tag.

A: Restriction digestion analysis of candidate clones for the expression of *Utp3*-GFP fusion proteins. Plasmids were screened by incubation with *EcoRI* and *AvaI* and resolution of the reaction products was demonstrated by agarose gel electrophoresis.

B: Spot growth complementation assays of the *GAL::UTP3* harbouring plasmids that express either wild-type Utp3 or Utp*3* deletion mutant. Transformants were grown in glactose-based selective minimal medium and ten-fold serial dilutions of the normalised culture were spotted onto glucose-based minimal medium. The image was taken after growth at 30°C for 2 days.

C: Western blot analysis of the expression level of GFP-Utp3 fusion proteins. Cell lysates of a wild-type strain transformed with plasmids that express GFP fusions of either the wild-type Utp3 protein or a deletion mutant were made under denaturing conditions and resolved by SDS-PAGE. The expression levels of the Utp3 proteins were assessed by Western blotting, using a GFP-specific antibody.

6.1: Utp21, Utp6 and Utp25 show specific subcellular localisation

The Utp21 and Utp6 components of the UtpB subcomplex and the Utp25 component of the SSU processome have been shown to localise to the nucleolus (Goldfeder and Oliveira, 2010; Tenge et al., 2014). In an initial step to compare the localisation of Utp3 proteins to known nucleolar markers, we analysed the *UTP21-GFP*, *UTP6-GFP* and *UTP6-GFP* strains by GFP fluorescence microscopy. A vector expressing GFP alone was used as a control (Fig. 6.2). Cell images showed that GFP fusions of Utp21, Utp25 and Utp6 were all localised to a single focus within the cell, consistent with previous reports.

6.2: Utp3 and all mutants are localised to a specific cellular structure

Transformants of a wild-type strain expressing the Utp3 GFP fusions were grown in selective minimal medium and the cellular localisation of the proteins were analysed by GFP fluorescence microscopy, as above. The GFP signal obtained for wild-type GFP-Utp3 and each of the mutants was also observed to be focused in a single subcellular structure, similar to the localisation of Utp21-GFP, Utp25-GFP and Utp6-GFP. This is in agreement with previous findings that localised Utp3 to the nucleolus (Kamakaka and Rine 1998). Furthermore, the data show that no one of the Utp3 deletions or truncations prevented the normal localisation of the protein. These findings demonstrate that although the CTD of Utp3 contains an NLS amino acid sequence, its absence does not impair Utp3 localisation to the nucleus. Moreover, since the ΔN, ΔM and ΔCTD deletions span the complete Utp3 structure, there is no single NLS present within the Utp3 protein that is necessary for nuclear localisation. It is possible that the protein has redundant NLS sequences, or that it is imported whilst associated with another protein that contains an NLS.

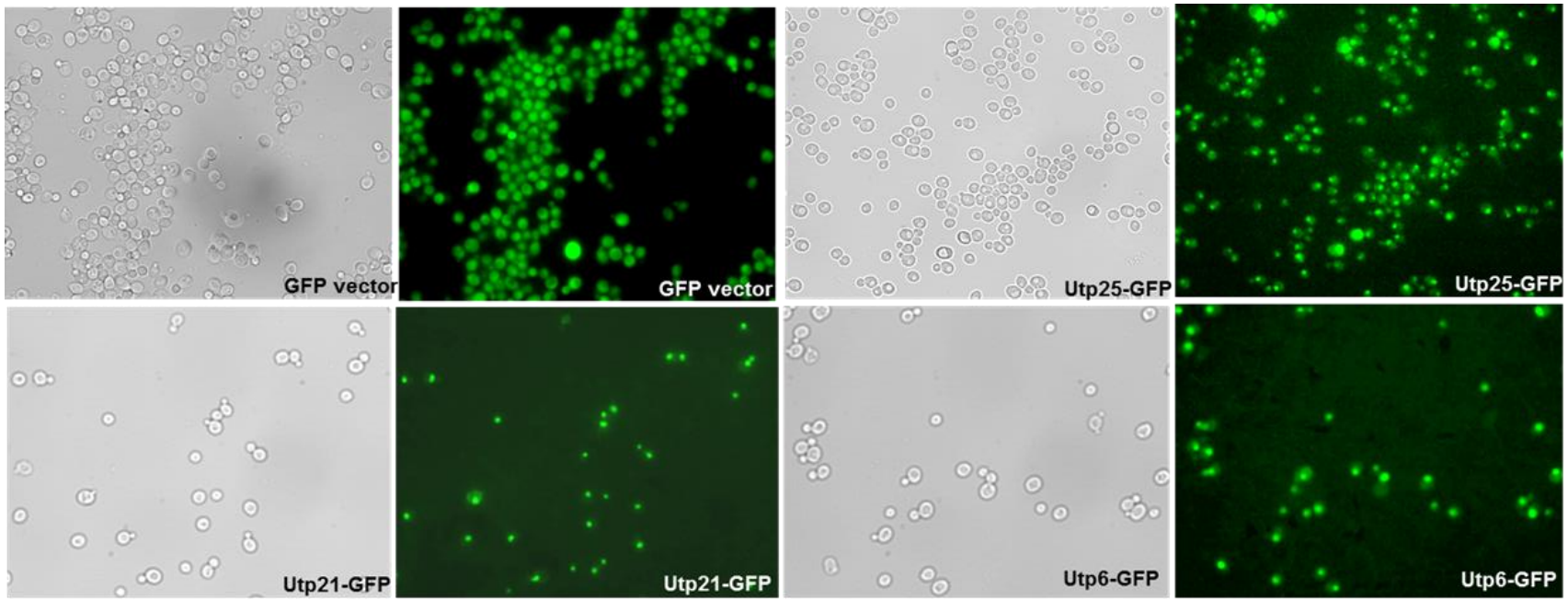

Figure 6.2: Utp21-GFP, Utp25-GFP and Utp6-GFP are clearly localised to the nucleus
Cells expressing Utp21-GFP, Utp25-GFP and Utp6-GFP from the normal chromosomal location were grown to OD_{600}<1 in minimal media. Empty vector was included as a negative control. Florescence microscopy was performed using (Axiovert 200M; Carl Zeiss) microscope. Bright field images and GFP fluorescence signals were captured from living cells.

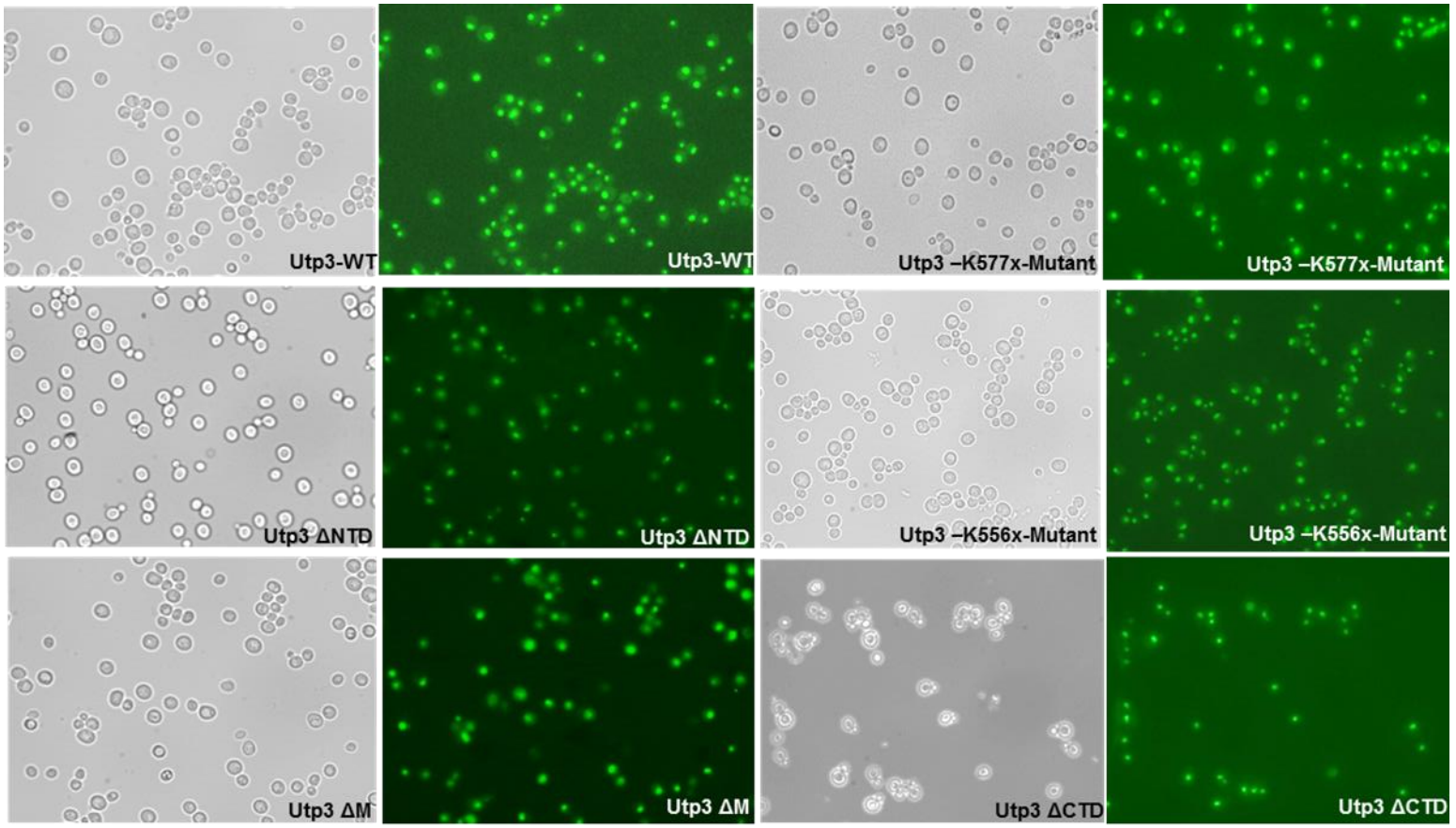

Figure 6.3: GFP-*Utp3* fusion and all mutants show localisation to the nucleolus

GFP-*Utp3* fusion and all *Utp3* mutants expressed in wild type yeast strain (p852). Live cell images were taken using florescence microscopy (Axiovert 200M; Carl Zeiss). Bright field images and GFP fluorescence signals were captured from living cells.

6.3: An RFP-Nop1 fusion provides a suitable nucleolar marker

Nop1 is a core component of box C/D snoRNPs that is implicated in the early, nucleolar phase of the pre-rRNA maturation pathway and is routinely used as a nucleolar localisation marker. To confirm that wild-type and mutant variants of Utp3 co-localise to the nucleolus with Nop1, a plasmid encoding an RFP fusion of Nop1 (Baßler et al., 2010), or the control vector was transformed into a wild-type yeast strain and the subcellular localisation of Nop1 was addressed by fluorescence microscopy (Fig. 6.4). Cell images revealed that the RFP-Nop1 showed a clear localisation to a single point within the yeast cells, indicating the position of the nucleolus. Future studies will address the co-localisation of GFP-tagged Utp3 proteins with the RFP-Nop1 reporter.

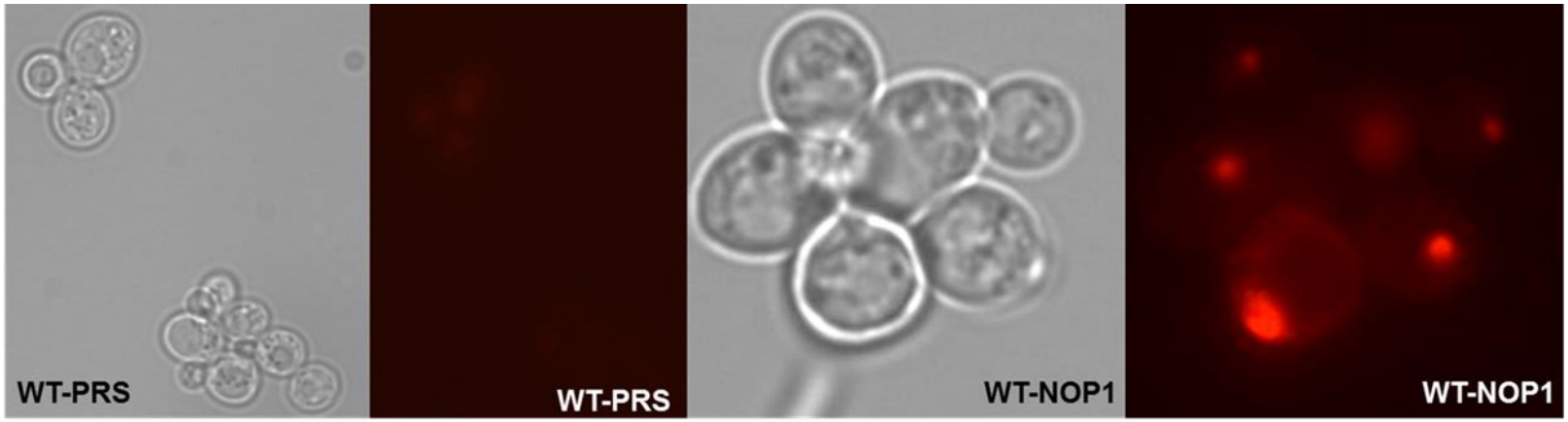

Figure 6.4: Nop1p (mCherry-Nop1p) expressed in the nucleolus
A wild-type strain bearing the *TRP1* auxotrophic marker (P455) was transformed with the plasmid encoding mCherry-Nop1p (p810) or the vector pRS314 and grown in selective growth medium. Imaging was performed on live cells. Data was collected under the bright field and by fluorescence through the RFP channel.

Chapter 7 : Discussion

Most of the more than 70 ribosomal proteins within the small subunit processome are uncharacterized and a significant number remain unassigned to subcomplexes. Utp3 is a large pre-ribosomal protein that is not a stable component of a characterised subcomplex and contains a C1D domain. Depletion of Utp3 is reported not to prevent the assembly of the SSU processome (Charette and Baserga, 2010). However, depletion of Utp3 severely impairs the production of 18S rRNA (Dragon et al., 2002). The specific role of Utp3 remains unclear. All proteins containing a C1D domain are involved in RNA metabolism. The C1D domain of Rrp47 is known to mediate the specific interaction with the PMC2NT domain of Rrp6 within the exosome complex (Stead et al., 2007; Costello et al., 2011; Feigenbutz et al., 2013; Schuch et al., 2014). In addition, the C1D domain of Lcp5 specifically interacts with Bfr2 (Griffith and Mitchell, unpublished data). We therefore decided to study Utp3 in more detail to ascertain how it functions within the SSU processome, using both *in vivo* and *in vitro* approaches.

Yeast 2-hybrid experiments have shown that Utp3 interacts with Utp25, a pre-ribosomal protein that has not been assigned to a subcomplex. Utp3 is also known to interact with the UtpB subcomplex, via both Utp6 and Utp21, and the Mpp10 subcomplex, through Mpp10 (Goldfeder and Oliveira, 2010; Charette and Baserga, 2010; Lim et al., 2011). These interactions with different subcomplexes probably occur concomitantly with the stepwise assembly of the complete SSU complex and most likely involve multiple regions of Utp3.

In this study, we aimed to characterise in more detail the behaviour of Utp3 in the early pre-rRNA cleavage events that generate 18S rRNA. To do this, a series of site-directed mutagenesis experiments were undertaken to generate a set of deletion mutants based on bioinformatically defined domains identified within Utp3. Interactions between Utp3 and several SSU subunit protein components were investigated using pull-down assays. The CTD domain of Utp3 was assayed for nucleic acid binding activity in vitro, while Utp3 RNA

binding in vivo was addressed through CRAC analyses. Finally, I investigated the effect of a set of deletion mutants on the subcellular localization of Utp3.

This study confirmed that depletion of Utp3 leads to a severe growth defect in yeast. As with U3 snoRNA depletion, growth was affected within six hours of transcriptional shut-off. This is faster than that observed for many other pre rRNA processing factors, many of which only show a growth defect after 6-25 hours post inhibition with a varying range of severity (Hage and Tollervey, 2004; Choque et al., 2011; Soltanieh et al., 2015).

This study also confirmed that Utp3 depletion causes a specific depletion in 18S pre-rRNA levels. This is also observed upon depletion of either U3 snoRNA or the U3 associated proteins Nop58, Nop1, Mpp10, Imp3, Imp4, Lcp5, Sof1, Utp23, Utp24, Enp1, Dpb4, Utp25 and Nop19. Depletion of these proteins resulted in an accumulation of the 35S pre-rRNA and the aberrant 23S pre-rRNA. There was also a decline in the levels of 20S pre-rRNA that is subsequently processed into mature 18S rRNA, but with varying degrees of severity (Dunbar et al., 1997; Lee and Baserga, 1997; Wiederkehr et al., 1998a; Tollervey et al., 1991; Chen et al., 2003; Choque et al., 2011; Soltanieh et al., 2015).

Western blot analysis in this study revealed that the level of Utp3 in the *GAL::UTP3* strain decreased dramatically within the first hour and totally vanished after 2 hours post transcriptional repression. This explains the rapid change in the 18S:25S rRNA ratio. Northern blotting shows the vast majority of 18S rRNA was depleted after just five hours and continued to fall. Taken together, this shows that Utp3 plays a key role in 18S synthesis and as such, lower levels of this protein are not tolerated by the cell. This result is in agreement with a previous study that found Utp3 co-precipitated with 35S pre-rRNA, showing that it is present in the complex before the primary cleavage occurs at A_0 (Chaker-Margot et al., 2015). This is supported by an accumulation of the 23S pre-rRNA, rather than 22S or 21S species that accumulate as a result of A_1 and A_2 site cleavage inhibition (Turner, 2011).

The depletion of many SSU processome components, such as Nan1, Pwp2/Utp1 or Imp4, leads to a loss of U3 snoRNA from 90S complexes (Perez-Fernandez et al., 2011). U3 snoRNA is an indicator of SSU assembly. Previous work (Charette and Baserga, 2010) showed that incorporation of U3 snoRNA into the 90S complex was not affected by the absence of Utp3. This suggests a secondary role of Utp3 in SSU processome assembly and the authors postulated that Utp3 either recruits other essential processing proteins or protein sub-complexes like Mpp10 during the tertiary assembly step of the SSU processome, or is an important factor that contributes to pre-rRNA cleavage.

In this study, a comparison of the sedimentation of zz-Utp3 and chromosomally encoded HTP-Utp3 through sucrose gradients showed that both were found in fractions containing ribosomal complexes. These findings were confirmed when the same lysates were fractionated through polysome gradients. This work validated HTP-Utp3 as a chromosomally tagged control for subsequent experiments on zz-Utp3 mutants. Northern blotting analyses of the RNA recovered from all 18 sucrose gradient fractions confirmed that Utp3 sediments with ribosomal complexes. This has not been previously demonstrated, but builds on work using tagged pre-rRNA and mass spectrometry showing that Utp3 is present in the 90S pre-ribosomal complex (Chaker-Margot et al., 2015; Zhang et al., 2016).

A previous study (Turner, 2011) found that deletion of the C1D domain of Utp3 did not affect strain viability, but that deletion of the CTD was lethal. In this study, yeast growth was also significantly impaired by the deletion of the CTD of Utp3. Additional Utp3 mutants were created for this study, including a deletion of the NTD, smaller deletions of the CTD and an internal deletion (ΔM) that comprises only the N- and C-terminal domains. Deletion of the NTD, or extensive truncations of the CTD were lethal, while the ΔM deletion did not clearly affect viability and a small C-terminal deletion (577X) resulted in a slow growth phenotype.

This study demonstrated loss of 18S rRNA production in all non-viable mutants. It seems very probable that this is the cause of the block in cell growth, as 25S rRNA levels remain unchanged in these mutants, indicating that processing of the LSU pre-rRNAs is unaffected. Taken together, these findings support a role of Utp3 through its CTD and NTD in 35S pre-rRNA processing to facilitate 18S rRNA synthesis.

Protein expression levels of the Utp3 mutants could potentially explain differences in their functionality. Quantitative Western blot analyses of the expression levels of zz-tagged wild-type and mutant Utp3 proteins showed that all mutants, with the exception of the N-terminal deletion, were expressed at comparable levels. Therefore, the growth phenotypes of the Utp3 mutants should be ascribed to an impairment in function rather than a deficiency in expression levels.

.

Sucrose gradient analyses and Northern blotting studies in this study demonstrated a widespread distribution of Utp3 throughout the gradient, signifying the involvement of Utp3 in a range of ribosomal complexes of varying size. It has been postulated for a long time that the SSU processome assembles in a hierarchical manner, with pre-assembled subcomplexes that are recruited sequentially (Mougey et al., 1993; Dragon et al., 2002; Osheim et al., 2004). Recent mass spectrometry findings support this and also suggest the presence of Utp3 at the time that early pre-rRNA processing events occur (Chaker-Margot et al., 2015; Zhang et al., 2016). SDS-PAGE analyses of the sucrose gradient fractions showed that all Utp3 mutants were incorporated into ribosomal complexes.

The multi-domain architecture of Utp3 suggests different roles of different regions of the protein. These undoubtedly include protein-protein interactions, as indicated by published yeast 2-hybrid interactions (Goldfeder and Oliveira, 2010; Charette and Baserga, 2010; Lim et al., 2011). The C1D and CTD regions of Utp3 are the strongest candidates to mediate these interactions. The known protein-protein interaction between Rrp6 and the C1D domain of Rrp47 suggests that C1D of Utp3 may have a role in protein-protein

interactions (Stead et al., 2007; Costello et al., 2011; Schuch et al., 2014). The results in this study demonstrate that the C1D domain alone was not sufficient for interaction with the UtpB sub complex, while it was sufficient for interaction with Utp25 in low salt lysate but not high salt lysate.

From earlier yeast 2-hybrid studies, Utp3 is known to interact with Utp21, Utp6, Utp25 and Mpp10 (Charette and Basega 2010). Pull-down experiments in this study support the conclusion that Utp3 is bound to the UtpB subcomplexes containing Utp21 and Utp6. Interaction with the UtpB complex was observed in lysates containing 150mM or 250mM NaCl. Further pull-down assays in this study using Utp3 mutants demonstrated that deletion of either the C1D or CTD domain independently prevented Utp3 association with both Utp21 and Utp6 in cell lysates. Furthermore, pull-down experiments on recombinant *E. coli* proteins and yeast lysates supported the conclusion that Utp3 interacts with the UtpB subcomplex via the C1D and CTD domains. The requirement of both C1D and CTD domains of Utp3 for interaction with Utp21 and Utp6 might either be because both interact independently, or because the C1D site is responsible for protein binding and the CTD stabilises the interaction in some way. The loss of viability seen in the ΔCTD mutant may be due to weakening of the interaction with the UtpB subcomplex. Recent studies have shown that the UtpB subcomplex together with the UtpA subcomplex contribute to a stabilising scaffold for the 35S pre-rRNA (Kornprobst et al., 2016). The interaction between Utp3 and the UtpB subcomplex could either involve protein interactions or be mediated through RNA. Pull-downs of Utp3 with Utp21 or Utp6 GFP fusion proteins were comparable, whether or not the samples were incubated with RNAase. Thus, the interaction between Utp3 and the UtpB complex is not simply because both are able to interact with RNA.

Previous studies using yeast 2-hybrid experiments showed that Utp3 interacts with Utp25 (Goldfeder and Oliveira, 2010; Charette and Baserga, 2010). Pull-down experiments in this study confirmed the Utp3/Utp25 interaction. Furthermore, interaction between Utp3 and Utp25 was direct and independent of RNA. Pull-downs were performed to look at Utp25 interaction with zz-Utp3 and Utp3 mutants using buffers at two salt concentrations and all Utp3

mutants, but not the wild-type zz-Utp3 protein, lost interaction with Utp25 in buffer containing 250mM NaCl. In buffer containing 150mM NaCl, only the ΔC1D mutant of Utp3 lost interaction with Utp25. This suggests that the Utp3/Utp25 interaction is more salt-sensitive (a reduction in signal can be seen even with the wild-type protein at 250mM NaCl) than the interaction between Utp3 and the UtpB subcomplex.

These results disagree with the findings of a previous study that suggested the N-terminal domain of Utp3 was sufficient for its interaction with Utp25 (Goldfeder and Oliveira, 2010). The truncation of Utp3 (effectively a bisection) used by Goldfeder actually includes a small part of the C1D domain of Utp3. Goldfeder also used a yeast 2-hybrid method, known to be more vulnerable to false positives, due to imbalanced expression of bait and prey proteins (Ito et al., 2001). In this study, pull-down experiments were only performed after confirming good expression of both tagged proteins by Western blotting.

This study suggests for the first time that the Utp3-Utp25 interaction involves the C1D domain of Utp3, and that the C1D domain is also involved, together with the CTD domain, in Utp21 and Utp6 interaction. It has been suggested that both Utp3 and Utp25 link the UtpB and Mpp10 subcomplexes (Charette and Baserga, 2010). My data suggest that these interactions are mediated in part by the C1D domain of Utp3. Figure 8.1 shows a schematic of the postulated domain-specific interactions between Utp3 and the UtpB subcomplex, Utp25 and U3 snoRNA/pre ribosomal RNA in the early 90S processome.

This study confirms that Utp3 is localised to a specific subcellular structure within the nucleus, presumably the nucleolus, but did not find any evidence of a nuclear localisation signal (NLS), in the C terminal, or any other region of Utp3, contrary to previous expectations (Kamakaka and Rine, 1998). This could be because Utp3 has more than one (redundant) NLS or because it is delivered to the nucleus through its association with another protein.
The second half of this study examined the RNA-binding ability of Utp3. *In vitro* EMSA studies indicated that Utp3 binds to DNA via its C-terminal domain and

there were weaker indications of RNA binding, also via the C-terminal domain. Previous studies have shown that binding of Rrp47 to structured RNA requires both the CTD and the C1D domain (Costello et al., 2011).

In vivo CRAC analyses identified an unexpectedly high number of rRNA binding sites for Utp3: binding might have been predicted to occur to the 5' ETS and with U3 snoRNA, but contacts were also identified within 18S rRNA. Even the sites identified in the 5' ETS and U3 snoRNA did not readily match known functional sites that are required for pre-rRNA processing. The clustering of Utp3 binding sites within helix 3 and adjacent to the Box B and Box C sites of U3 snoRNA was striking. The proximity of this site to Utp10 binding, together with the location of Utp3 binding very close to the Utp10 binding site within the 5' ETS, suggests a role for both Utp3 and UtpA in recruiting or stabilising the U3 snoRNP within the 90S processome. CRAC data for the 5' ETS showed Utp3 binding at sites close to the extreme 5' end. This supports the hypothesis that Utp3 has a role in very early rRNA processing, when only UtpA is recruited to the 5'ETS and goes against the suggestion that Utp3 (with Utp25) is recruited to the SSU processome after UtpA, UtpB and U3 snoRNA (Charette and Baserga, 2010). Utp3 binding also was demonstrated close to Utp5, another UtpA component, on the 5'ETS.

This study has demonstrated pull-down interactions between Utp3 and Utp21 or Utp6 components of the UtpB subcomplex and shown it to be independent of RNA. Consideration of the structure of the 90S processome (Kornprobst et al., 2016) suggests that the direct interaction between Utp3 and the UtpB complex could contribute to stable association of U3 snoRNA with the processome via its 3' domain, as these features appear in close spatial proximity.

The nascent 35S pre-RNA is incorported into an evolving supporting structure comprising the U3 snoRNP particle and multiple protein subcomplexes of the processome. The location of some 13 binding sites for Utp3 on the 18S rRNA coding sequence, close to areas of likely increased flexibility, could indicae a role for Utp3 in helping the pre-rRNA to fold correctly or maintain its position

within this scaffold at various stages in the dynamic development of the 90S particle.

Given the known role of the C1D domain in RNA processing and repair (Mitchell, 2010; Garland et al., 2013; Feigenbutz et al., 2013), it is also speculatively possible that the multiple Utp3 binding sites to the nascent rRNA represent a function related to stabilisation, surveillance or repair of the pre ribosomal transcript. A role in ensuring the accuracy of ribosome biogenesis is not inconsistent with the suggested role of a rRNA chaperone.

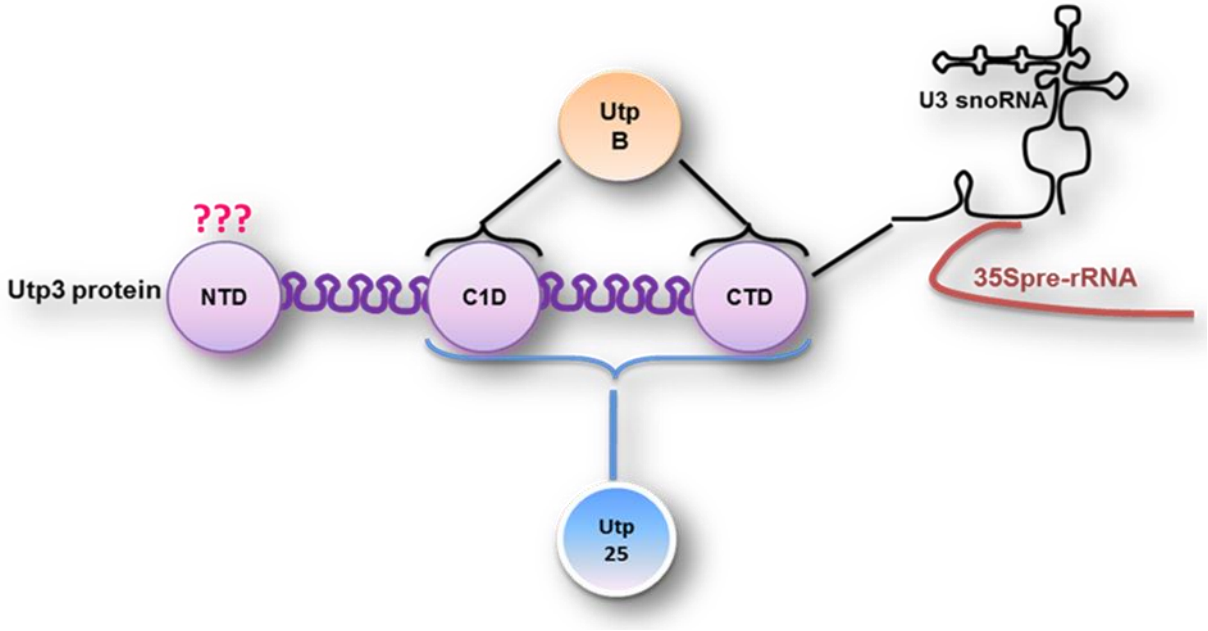

Figure 7.1: Functions of Utp3 domains

Utp3 interaction with UtpB requires either the C1D or CTD domain of Utp3, while Utp3 interaction with Utp25 requires both. Utp3 CTD binds rRNA in the pre-ribosome and ribosomal snoRNA.

.

Chapter 8 : Conclusions and Recommendations

In order to elucidate the role of Utp3 in yeast ribosome biogenesis, I present the data from a combination of in vivo and in vitro genetic manipulations and biochemical analyses. The data in this study complements studies showing yeast 2-hybrid interactions between Utp3 and components of the SSU processome. We demonstrated that the C1D and CTD domains of Utp3 are important for its interaction with Utp21 and Utp6 in cell extracts and using recombinant proteins. The loss of function associated with the ΔCTD mutant may be due to disruption of the interaction between Utp3 and the UtpB subcomplex. Further pull-down investigations of the interaction between Utp3 and the Mpp10 subcomplex would help increase our understanding of the function of specific domains within Utp3 and their role in early pre-rRNA processing. We found that the interaction between Utp3 and Utp25 is not as strong as that found between Utp3 and the UtpB complex. Utp3 interactions with Utp21, Utp6 and Utp25 were independent of RNA. Reciprocal experiments should be performed in order to identify Utp21 and Utp6 interaction domains with Utp3.

I demonstrated that the CTD domain of Utp3 is able to interact directly with DNA in vitro. The CTD domain most probably mediates Utp3 interactions with RNA. Further work developing a CTD deletion construct for CRAC analysis would clarify this. Crosslinking studies (CRAC) revealed a wide RNA binding profile for yeast Utp3, suggesting a role with UtpA in recruiting or stabilising U3 snoRNP, and were consistent with a direct protein-protein interaction with UtpB that from structural studies could also stabilise U3 snoRNP within the 90S processome. I speculate that Utp3 performs a generalised function as an RNA chaperone in the 90S processome. In this study we found that no Utp3 mutation affected its localisation within the cell, although the CTD domain had previously been predicted to contain an NLS.

Finally, the data presented here paints a contrasting picture of Utp3 function. The effect of loss of Utp3 mimics the depletion of the undoubtedly core processome component U3 snoRNA. The SSU processome can assemble in

the absence of Utp3 (Perez-Fernandez et al., 2011), whereas assembly is dependent upon U3 snoRNA. Utp3 function would appear to reflect that of components such as the Mpp10 subcomplex, that associate later but are nevertheless required for core processome activity (Gerczei et al., 2009). Utp3, like Mpp10 components, may help recruit and stabilise U3 snoRNA in relation to the evolving 35S pre-ribosomal rRNA. The mutagenesis and yeast complementation analyses reveal, as illustrated in Figure 8.1, that the C-terminal domain and C1D domain of Utp3 comprise regions that are key to its function in terms of its interactions with the UtpB subcomplex, with Utp25 and, in the case of the CTD domain, in nucleic acid binding.

References

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

Alkemar, G. & Nygard, O. 2004. Secondary structure of two regions in expansion segments ES3 and ES6 with the potential of forming a tertiary interaction in eukaryotic 40S ribosomal subunits. *RNA,* 10(3)**,** pp 403-11.

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. & Beckmann, R. 2010. 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.

Balakin, A. G., Smith, L. & Fournier, M. J. 1996. The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell,* 86(5)**,** pp 823-834.

Bammert, L., Jonas, S., Ungricht, R. & Kutay, U. 2016. Human AATF/Che-1 forms a nucleolar protein complex with NGDN and NOL10 required for 40S ribosomal subunit synthesis. *Nucleic Acids Res*.

Bange, G., Murat, G., Sinning, I., Hurt, E. & Kressler, D. 2013. New twist to nuclear import: When two travel together. *Commun Integr Biol,* 6(4)**,** pp e24792.

Bartelt-Kirbach, B., Wuepping, M., Dodrimont-Lattke, M. & 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. & Hurt, E. 2001. Identification of a 60S preribosomal particle that is closely linked to nuclear export. *Mol Cell,* 8(3)**,** pp 517-29.

Baxter-Roshek, J. L., Petrov, A. N. & Dinman, J. D. 2007. Optimization of ribosome structure and function by rRNA base modification. *PloS one,* 2(1)**,** pp e174.

Baßler, J., Kallas, M., Pertschy, B., Ulbrich, C., Thoms, M. & Hurt, E. 2010. The AAA-ATPase Rea1 drives removal of biogenesis factors during multiple stages of 60S ribosome assembly. *Mol Cell,* 38(5)**,** pp 712-721.

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

Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G. & Yusupov, M. 2011. The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution. *Science,* 334(6062)**,** pp 1524-1529.

Ben-Yehuda, S., Dix, I., Russell, C. S., McGarvey, M., Beggs, J. D. & Kupiec, M. 2000. Genetic and physical interactions between factors involved in both cell cycle progression and pre-mRNA splicing in *Saccharomyces cerevisiae*. *Genetics,* 156(4)**,** pp 1503-17.

Bernstein, J. & Toth, E. A. 2012. Yeast nuclear RNA processing. *World Journal of Biological Chemistry,* 3(1)**,** pp 7-26.

Bernstein, K. A. & Baserga, S. J. 2004. The small subunit processome is required for cell cycle progression at G1. *Mol Cell* 15(11)**,** pp 5038-46.

Bernstein, K. A., Granneman, S., Lee, A. V., Manickam, S. & Baserga, S. J. 2006. Comprehensive mutational analysis of yeast DEXD/H box RNA helicases involved in large ribosomal subunit biogenesis. *Mol Cell,* 26(4)**,** pp 1195-1208.

Billy, E., Wegierski, T., Nasr, F. & Filipowicz, W. 2000. Rcl1p, the yeast protein similar to the RNA 3'-phosphate cyclase, associates with U3 snoRNP and is required for 18S rRNA biogenesis. *EMBO J,* 19(9)**,** pp 2115-26.

Bleichert, F. & Baserga, S. J. 2007. The long unwinding road of RNA helicases. *Mol Cell,* 27(3)**,** pp 339-352.

Bleichert, F., Granneman, S., Osheim, Y. N., Beyer, A. L. & Baserga, S. J. 2006 The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proceedings of the National Academy of Sciences,* 103(25)**,** pp 9464-9469.

for the early cleavage steps in 18S rRNA maturation. *Proc Natl Acad Sci U S A,* 103(25)**,** pp 9464-9.

Bohnsack, M. T., Martin, R., Granneman, S., Ruprecht, M., Schleiff, E. & Tollervey, D. 2009. Prp43 bound at different sites on the pre-rRNA

performs distinct functions in ribosome synthesis. *Mol Cell,* 36(4)**,** pp 583-592.

Booher, K. R. & Kaiser, P. 2008. A PCR-based strategy to generate yeast strains expressing endogenous levels of amino-terminal epitope-tagged proteins. *Biotechnol J,* 3(4)**,** pp 524-9.

Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry,* 72(1)**,** pp 248-254.

Brand, R. C., Klootwijk, J., Sibum, C. P. & Planta, R. J. 1979. Pseudouridylation of yeast ribosomal precursor RNA. *Nucleic acids research,* 7(1)**,** pp 121-134.

Brand, R. C., Klootwijk,, J., Steenbergen, T. J., Kok, A. J. & Planta, R. J. 1977. Secondary methylation of yeast ribosomal precursor RNA. *European Journal of Biochemistry,* 75(1)**,** pp 311-318.

Brown, A., Shao, S., Murray, J., Hegde, R. S. & Ramakrishnan, V. 2015. Structural basis for stop codon recognition in eukaryotes. *Nature,* 524(7566)**,** pp 493-496.

Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Müller, K. M., Pande, N., Shang, Z., Yu, N. & Gutell, R. R. 2002. The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics,* 3(2-2.

Chaker-Margot, M., Hunziker, M., Barandun, J., Dill, B. D. & 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 920-923.

Champion, E. A., Lane, B. H., Jackrel, M. E., Regan, L. & Baserga, S. J. 2008. A direct interaction between the Utp6 half-a-tetratricopeptide repeat domain and a specific peptide in Utp21 is essential for efficient pre-rRNA processing. *Mol Cell Biol,* 28(21)**,** pp 6547-56.

Chandramouli, P., Topf, M., Menetret, J. F., Eswar, N., Cannone, J. J., Gutell, R. R., Sali, A. & Akey, C. W. 2008. Structure of the mammalian 80S ribosome at 8.7 A resolution. *Structure,* 16(4)**,** pp 535-48.

Charette, J. M. & Baserga, S. J. 2010. The DEAD-box RNA helicase-like Utp25 is an SSU processome component. *RNA,* 16(11)**,** pp 2156-69.

Charette, J. M. & Gray, M. W. 2009. U3 snoRNA genes are multi-copy and frequently linked to U5 snRNA genes in Euglena gracilis§. *BMC Genomics,* 10(1)**,** pp 1-19.

Chen, W., Bucaria, J., Band, D. A., Sutton, A. & Sternglanz, R. 2003. Enp1, a yeast protein associated with U3 and U14 snoRNAs, is required for pre-rRNA processing and 40S subunit synthesis. *Nucleic Acids Res,* 31(2)**,** pp 690-9.

Choque, E., Marcellin, M., Burlet-Schiltz, O., Gadal, O. & Dez, C. 2011. The nucleolar protein Nop19p interacts preferentially with Utp25p and Dhr2p and is essential for the production of the 40S ribosomal subunit in *Saccharomyces cerevisiae*. *RNA Biol,* 8(6)**,** pp 1158-72.

Ciganda, M. & Williams, N. 2011. Eukaryotic 5S rRNA biogenesis. *Wiley Interdisciplinary Reviews: RNA,* 2(4)**,** pp 523-533.

Colley, A., Beggs, J. D., Tollervey, D. & Lafontaine, D. L. 2000. Dhr1p, a putative DEAH-box RNA helicase, is associated with the box C+ D snoRNP U3. *Mol Cell,* 20(19)**,** pp 7238-7246.

Costello, J. L., Stead, J. A., Feigenbutz, M., Jones, R. M. & Mitchell, P. 2011. The C-terminal region of the exosome-associated protein Rrp47 is specifically required for box C/D small nucleolar RNA 3'-maturation. *J Biol Chem,* 286(6)**,** pp 4535-43.

de la Cruz, J., Karbstein, K. & Woolford, J. L. 2015. Functions of Ribosomal Proteins in Assembly of Eukaryotic Ribosomes In Vivo. *Annual review of biochemistry,* 84(93-129.

Decatur, W. A. & Fournier, M. J. 2002. rRNA modifications and ribosome function. *Trends in Biochemical Sciences,* 27(7)**,** pp 344-351.

Delprato, A., Al Kadri, Y., Perebaskine, N., Monfoulet, C., Henry, Y., Henras, A. K. & Fribourg, S. 2014. Crucial role of the Rcl1p-Bms1p interaction for yeast pre-ribosomal RNA processing. *Nucleic Acids Res,* 42(15)**,** pp 10161-72.

Dinman, J. D. 2009. The eukaryotic ribosome: current status and challenges. *J Biol Chem,* 284(18)**,** pp 11761-5.

Dosil, M. & 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. G., 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. & Baserga, S. J. 2002. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature,* 417(6892)**,** pp 967-970.

Dunbar, D. A., Wormsley, S., Agentis, T. M. & 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. & 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.

Fairman-Williams, M. E., Guenther, U.-P. & Jankowsky, E. 2010. SF1 and SF2 helicases: family matters. *Current opinion in structural biology,* 20(3)**,** pp 313-324.

Fatica, A., Tollervey, D. & DlakiĆ, M. 2004. PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA. *RNA,* 10(11)**,** pp 1698-1701.

Feigenbutz, M., Garland, W., Turner, M. & Mitchell, P. 2013. The exosome cofactor Rrp47 is critical for the stability and normal expression of its associated exoribonuclease Rrp6 in *Saccharomyces cerevisiae*. *PLoS One,* 8(11)**,** pp e80752.

Ferreira-Cerca, S., Pöll, G., Gleizes, P.-E., Tschochner, H. & Milkereit, P. 2005. Roles of Eukaryotic Ribosomal Proteins in Maturation and Transport of Pre-18S rRNA and Ribosome Function. *Mol Cell,* 20(2)**,** pp 263-275.

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J. & Bateman, A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res,* 44(D1)**,** pp D279-85.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R. & Bateman, A. 2010. The Pfam protein families database. *Nucleic Acids Res,* 38(Database issue)**,** pp D211-22.

Freed, E. F., Bleichert, F., Dutca, L. M. & Baserga, S. J. 2010. When ribosomes go bad: diseases of ribosome biogenesis. *Molecular BioSystems,* 6(3)**,** pp 481-493.

Gallagher, J. E., Dunbar, D. A., Granneman, S., Mitchell, B. M., Osheim, Y., Beyer, A. L. & 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.

Ganot, P., Bortolin, M.-L. & Kiss, T. 1997. Site-Specific Pseudouridine Formation in Preribosomal RNA Is Guided by Small Nucleolar RNAs. *Cell,* 89(5)**,** pp 799-809.

Garland, W., Feigenbutz, M., Turner, M. & Mitchell, P. 2013. Rrp47 functions in RNA surveillance and stable RNA processing when divorced from the exoribonuclease and exosome-binding domains of Rrp6. *RNA,* 19(12)**,** pp 1659-1668.

Gerczei, T., Shah, B. N., Manzo, A. J., Walter, N. G. & 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.

Gerhardy, S., Menet, A. M., Peña, C., Petkowski, J. J. & Panse, V. G. 2014. Assembly and nuclear export of pre-ribosomal particles in budding yeast. *Chromosoma,* 123(4)**,** pp 327-344.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., Arkin, A. P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K. D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Guldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kotter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P.,

Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C. Y., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W. & Johnston, M. 2002. Functional profiling of the Saccharomyces cerevisiae genome. *Nature,* 418(6896)**,** pp 387-91.

Gietz, D., St Jean, A., Woods, R. A. & Schiestl, R. H. 1992. Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res,* 20(6)**,** pp 1425.

Goldfeder, M. B. & Oliveira, C. C. 2010. Utp25p, a nucleolar Saccharomyces cerevisiae protein, interacts with U3 snoRNP subunits and affects processing of the 35S pre-rRNA. *Febs j,* 277(13)**,** pp 2838-52.

Grandi, P., Rybin, V., Bassler, J., Petfalski, E., Strauss, D., Marzioch, M., Schafer, T., Kuster, B., Tschochner, H., Tollervey, D., Gavin, A. C. & 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., Bernstein, K. A., Bleichert, F. & Baserga, S. J. 2006. Comprehensive mutational analysis of yeast DEXD/H box RNA helicases required for small ribosomal subunit synthesis. *Mol Celll,* 26(4)**,** pp 1183-94.

Granneman, S., Kudla, G., Petfalski, E. & Tollervey, D. 2009a. 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(24)**,** pp 9613-8.

Granneman, S., Petfalski, E., Swiatkowska, A. & Tollervey, D. 2010. Cracking pre-40S ribosomal subunit structure by systematic analyses of RNA–protein cross-linking. *The EMBO J,* 29(12)**,** pp 2026-2036.

Grunwald, D., Singer, R. H. & Rout, M. 2011. Nuclear export dynamics of RNA-protein complexes. *Nature,* 475(7356)**,** pp 333-341.

Hage, A. E. & Tollervey, D. 2004. A surfeit of factors: why is ribosome assembly so much more complicated in eukaryotes than bacteria? *RNA Biol,* 1(1)**,** pp 10-5.

Hautbergue, G. M., Hung, M. L., Golovanov, A. P., Lian, L. Y. & Wilson, S. A. 2008. Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proc Natl Acad Sci U S A,* 105(13)**,** pp 5154-9.

Hazbun, T. R., Malmstrom, L., Anderson, S., Graczyk, B. J., Fox, B., Riffle, M., Sundin, B. A., Aranda, J. D., McDonald, W. H., Chiu, C. H., Snydsman, B. E., Bradley, P., Muller, E. G., Fields, S., Baker, D., Yates, J. R., 3rd & Davis, T. N. 2003. Assigning function to yeast proteins by integration of technologies. *Mol Cell,* 12(6)**,** pp 1353-65.

Hellman, L. M. & Fried, M. G. 2007. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc,* 2(8)**,** pp 1849-61.

Henras, A. K., Dez, C. & Henry, Y. 2004. RNA structure and function in C/D and H/ACA s(no)RNPs. *Current Opinion in Structural Biology,* 14(3)**,** pp 335-343.

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

Hoareau-Aveilla, C., Fayet-Lebaron, E., Jady, B. E., Henras, A. K. & Kiss, T. 2012. Utp23p is required for dissociation of snR30 small nucleolar RNP from preribosomal particles. *Nucleic Acids Res,* 40(8)**,** pp 3641-52.

Hoffman, C. & GARRISON, T. R. 1997. Saccharomyces cerevisiae. *Current Protocols in Molecular Biology, FM Ausubel, R. Brent, RE Kingston, DD Moore, JG Seidman, JA Smith, and K. Struhl, eds (New York: John Wiley & Sons),* 13(5-13.7.

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

Hughes, J. & Ares Jr, M. 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. *The EMBO J,* 10(13)**,** pp 4231.

Hughes, J. M. 1996a. Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA. *Journal of molecular biology,* 259(4)**,** pp 645-654.

Hughes, J. M., Konings, D. A. & Cesareni, G. 1987. The yeast homologue of U3 snRNA. *The EMBO J,* 6(7)**,** pp 2145-2155.

Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S. & O'Shea, E. K. 2003. Global analysis of protein localization in budding yeast. *Nature,* 425(6959)**,** pp 686-91.

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. & Klinge, S. 2016. UtpA and UtpB chaperone nascent pre-ribosomal RNA and U3 snoRNA to initiate eukaryotic ribosome assembly. *Nat Commun,* 7(12090.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. & Sakaki, Y. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A,* 98(8)**,** pp 4569-74.

Jack, K., Bellodi, C., Landry, D. M., Niederer, R. O., Meskauskas, A., Musalgaonkar, S., Kopmar, N., Krasnykh, O., Dean, A. M. & Thompson, S. R. 2011. rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Mol Cell,* 44(4)**,** pp 660-666.

Jackson, R. A., Wu, J. S. & Chen, E. S. 2016. C1D family proteins in coordinating RNA processing, chromosome condensation and DNA damage response. *Cell division,* 11(1)**,** pp 1.

Jenne, D. E., Tinschert, S., Dorschner, M. O., Hameister, H., Stephens, K. & Kehrer-Sawatzki, H. 2003. Complete physical map and gene content of the human NF1 tumor suppressor region in human and mouse. *Genes Chromosomes Cancer,* 37(2)**,** pp 111-20.

Jung, M. Y., Lorenz, L. & Richter, J. D. 2006. Translational control by neuroguidin, a eukaryotic initiation factor 4E and CPEB binding protein. *Mol Cell,* 26(11)**,** pp 4277-87.

Kamakaka, R. T. & Rine, J. 1998. Sir- and silencer-independent disruption of silencing in Saccharomyces by Sas10p. *Genetics,* 149(2)**,** pp 903-14.

Kapp, L. D. & Lorsch, J. R. 2004. The molecular mechanics of eukaryotic translation. *Annual review of biochemistry,* 73(1)**,** pp 657-704.

Karbstein, K. 2011. Inside the 40S ribosome assembly machinery. *Curr Opin Chem Biol,* 15(5)**,** pp 657-63.

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

Kawai, G., Yamamoto, Y., Kamimura, T., Masegi, T., Sekine, M., Hata, T., Iimori, T., Watanabe, T., Miyazawa, T. & Yokoyama, S. 1992. Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group. *Biochemistry,* 31(4)**,** pp 1040-1046.

Kiss, T., Fayet-Lebaron, E. & Jády, B. E. 2010. Box H/ACA small ribonucleoproteins. *Mol Cell,* 37(5)**,** pp 597-606.

Kiss-László, Z., Henry, Y., Bachellerie, J.-P., Caizergues-Ferrer, M. & Kiss, T. 1996. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell,* 85(7)**,** pp 1077-1088.

Knop, M., Siegers, K., Pereira, G., Zachariae, W., Winsor, B., Nasmyth, K. & Schiebel, E. 1999. Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. *Yeast,* 15(10B)**,** pp 963-72.

Kornprobst, M., Turk, M., Kellner, N., Cheng, J., Flemming, D., Koš-Braun, I., Koš, M., Thoms, M., Berninghausen, O., Beckmann, R. & 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, M. & 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. & Tollervey, D. 2010. Yeast pre-rRNA processing and modification occur cotranscriptionally. *Mol Cell,* 37(6)**,** pp 809-20.

Kressler, D., Hurt, E. & Bassler, J. 2010. Driving ribosome assembly. *Biochim Biophys Acta,* 1803(6)**,** pp 673-83.

Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P., Punna, T., Peregrín-Alvarez, J. M., Shales, M., Zhang, X., Davey, M., Robinson, M. D., Paccanaro, A., Bray, J. E., Sheung, A., Beattie, B., Richards, D. P., Canadien, V., Lalev, A., Mena, F., Wong, P., Starostine, A., Canete, M. M., Vlasblom,

J., Wu, S., Orsi, C., Collins, S. R., Chandran, S., Haw, R., Rilstone, J. J., Gandi, K., Thompson, N. J., Musso, G., St Onge, P., Ghanny, S., Lam, M. H. Y., Butland, G., Altaf-Ul, A. M., Kanaya, S., Shilatifard, A., O'Shea, E., Weissman, J. S., Ingles, C. J., Hughes, T. R., Parkinson, J., Gerstein, M., Wodak, S. J., Emili, A. & Greenblatt, J. F. 2006. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature,* 440(7084)**,** pp 637-643.

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. & Greenblatt, J. F. 2004. High-definition macromolecular composition of yeast RNA-processing complexes. *Mol Cell,* 13(2)**,** pp 225-39.

Kudla, G., Granneman, S., Hahn, D., Beggs, J. D. & Tollervey, D. 2011. Cross-linking, ligation, and sequencing of hybrids reveals RNA–RNA interactions in yeast. *Proceedings of the National Academy of Sciences,* 108(24)**,** pp 10010-10015.

Lafontaine, D., Delcour, J., Glasser, A.-L., Desgres, J. & Vandenhaute, J. 1994. The DIM1 Gene Responsible for the Conserved m62Am62A Dimethylation in the 3′-Terminal Loop of 18 S rRNA is Essential in Yeast. *Journal of Molecular Biology,* 241(3)**,** pp 492-497.

Lafontaine, D. L. & Tollervey, D. 1999. Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *RNA,* 5(3)**,** pp 455-467.

Lamanna, A. C. & Karbstein, K. 2011. An RNA Conformational Switch Regulates Pre-18S rRNA Cleavage. *Journal of Molecular Biology,* 405(1)**,** pp 3-17.

Lane, B. G., Ofengand, J. & Gray, M. W. 1995. Pseudouridine and O2-methylated nucleosides. Significance of their selective occurrence in rRNA domains that function in ribosome-catalyzed synthesis of the peptide bonds in proteins. *Biochimie,* 77(1–2)**,** pp 7-15.

Lebaron, S., Papin, C., Capeyrou, R., Chen, Y. L., Froment, C., Monsarrat, B., Caizergues-Ferrer, M., Grigoriev, M. & Henry, Y. 2009. The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein

Pfa1p during yeast ribosome biogenesis. *The EMBO J,* 28(24)**,** pp 3808-3819.

Lebaron, S., Schneider, C., van Nues, R. W., Swiatkowska, A., Walsh, D., Bottcher, B., Granneman, S., Watkins, N. J. & Tollervey, D. 2012a. 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. & Tollervey, D. 2013. Rrp5 binding at multiple sites coordinates pre-rRNA processing and assembly. *Mol Cell,* 52(5)**,** pp 707-19.

Lee, S. J. & Baserga, S. J. 1997. Functional separation of pre-rRNA processing steps revealed by truncation of the U3 small nucleolar ribonucleoprotein component, Mpp10. *Proc Natl Acad Sci U S A,* 94(25)**,** pp 13536-41.

Lee, S. J. & 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.

Lemay, V., Hossain, A., Osheim, Y. N., Beyer, A. L. & Dragon, F. 2011. Identification of novel proteins associated with yeast snR30 small nucleolar RNA. *Nucleic acids research,* gkr659.

Li, H., Zagorski, J. & Fournier, M. J. 1990. Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in *Saccharomyces cerevisiae*. *Mol Cell,* 10(3)**,** pp 1145-1152.

Li, T., Chen, X., Garbutt, K. C., Zhou, P. & Zheng, N. 2006. Structure of DDB1 in complex with a paramyxovirus V protein: viral hijack of a propeller cluster in ubiquitin ligase. *Cell,* 124(1)**,** pp 105-17.

Liang, W. Q. & 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, X.-h., Liu, Q., Liu, Q., King, T. H. & Fournier, M. J. 2010. Strong dependence between functional domains in a dual-function snoRNA infers coupling of rRNA processing and modification events. *Nucleic Acids Research,* 38(10)**,** pp 3376-3387.

Lim, Y. H., Charette, J. M. & Baserga, S. J. 2011. Assembling a protein-protein interaction map of the SSU processome from existing datasets. *PLoS One,* 6(3)**,** pp e17701.

Lin, J., Lu, J., Feng, Y., Sun, M. & 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)**,** pp e1001669.

Lo, K. Y., Li, Z., Bussiere, C., Bresson, S., Marcotte, E. M. & Johnson, A. W. 2010. Defining the pathway of cytoplasmic maturation of the 60S ribosomal subunit. *Mol Cell,* 39(2)**,** pp 196-208.

Longtine, M. S., McKenzie, A., 3rd, Demarini, D. J., Shah, N. G., Wach, A., Brachat, A., Philippsen, P. & Pringle, J. R. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in S*accharomyces cerevisiae*. *Yeast,* 14(10)**,** pp 953-61.

Lu, J., Sun, M. & Ye, K. 2013. Structural and functional analysis of Utp23, a yeast ribosome synthesis factor with degenerate PIN domain. *Rna,* 19(12)**,** pp 1815-24.

Maden, B. 1990. The numerous modified nucleotides in eukaryotic ribosomal RNA. *Progress in nucleic acid research and molecular biology,* 39(241-303.

Maden, B. E. H. 1986. Identification of the locations of the methyl groups in 18 S ribosomal RNA from Xenopus laevis and man. *Journal of Molecular Biology,* 189(4)**,** pp 681-699.

Maniatis, T., Fritsch, E. & Sambrook, J. 1982. Cold Spring Harbor Laboratory. Cold Spring Harbor, NY.

Martin, R., Straub, A. U., Doebele, C. & Bohnsack, M. T. 2013. DExD/H-box RNA helicases in ribosome biogenesis. *RNA biology,* 10(1)**,** pp 4-18.

Marz, M. & Stadler, P. F. 2009a. Comparative analysis of eukaryotic U3 snoRNA. *RNA Biology,* 6(5)**,** pp 503-507.

Melnikov, S., Ben-Shem, A., de Loubresse, N. G., Jenner, L., Yusupova, G. & Yusupov, M. 2012. One core, two shells: bacterial and eukaryotic ribosomes. *Nature structural & molecular biology,* 19(6)**,** pp 560-567.

Merl, J., Jakob, S., Ridinger, K., Hierlmeier, T., Deutzmann, R., Milkereit, P. & Tschochner, H. 2010. Analysis of ribosome biogenesis factor-modules

in yeast cells depleted from pre-ribosomes. *Nucleic Acids Res,* 38(9)**,** pp 3068-80.

Mitchell, P. 2010. Rrp47 and the function of the Sas10/C1D domain. *Biochem Soc Trans,* 38(4)**,** pp 1088-92.

Mitchell, P., Petfalski, E., Houalla, R., Podtelejnikov, A., Mann, M. & Tollervey, D. 2003. Rrp47p is an exosome-associated protein required for the 3′ processing of stable RNAs. *Molecular and Cellular Biology,* 23(19)**,** pp 6982-6992.

Mitchell, P., Petfalski, E. & 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.

Mohr, D., Frey, S., Fischer, T., Guttler, T. & Gorlich, D. 2009. Characterisation of the passive permeability barrier of nuclear pore complexes. *EMBO J,* 28(17)**,** pp 2541-53.

Monemi, S., Spaeth, G., DaSilva, A., Popinchalk, S., Ilitchev, E., Liebmann, J., Ritch, R., Heon, E., Crick, R. P., Child, A. & Sarfarazi, M. 2005. Identification of a novel adult-onset primary open-angle glaucoma (POAG) gene on 5q22.1. *Hum Mol Genet,* 14(6)**,** pp 725-33.

Moriggi, G., Nieto, B. & Dosil, M. 2014. Rrp12 and the Exportin Crm1 participate in late assembly events in the nucleolus during 40S ribosomal subunit biogenesis. *PLoS Genet,* 10(12)**,** pp e1004836.

Morrissey, J. P. & Tollervey, D. 1993. Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol Cell,* 13(4)**,** pp 2469-2477.

Moss, T. 2004. At the crossroads of growth control; making ribosomal RNA. *Curr Opin Genet Dev,* 14(2)**,** pp 210-7.

Motley, A. M., Nuttall, J. M. & Hettema, E. H. 2012. Pex3-anchored Atg36 tags peroxisomes for degradation in *Saccharomyces cerevisiae*. *EMBO J,* 31(13)**,** pp 2852-68.

Mougey, E. B., O'Reilly, M., Osheim, Y., Miller, O. L., Jr., Beyer, A. & 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.

Méreau, A., Fournier, R., Grégoire, A., Mougin, A., Fabrizio, P., Lührmann, R. & Branlant, C. 1997. An in vivo and in vitro structure-function analysis of

the *Saccharomyces cerevisiae* U3A snoRNP: protein-RNA contacts and base-pair interaction with the pre-ribosomal RNA1. *Journal of Molecular Biology,* 273(3)**,** pp 552-571.

Nehls, P., Keck, T., Greferath, R., Spiess, E., Glaser, T., Rothbarth, K., Stammer, H. & Werner, D. 1998. cDNA cloning, recombinant expression and characterization of polypetides with exceptional DNA affinity. *Nucleic Acids Res,* 26(5)**,** pp 1160-6.

Ni, J., Tien, A. L. & Fournier, M. J. 1997. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell,* 89(4)**,** pp 565-573.

Nissan, T. A., Bassler, J., Petfalski, E., Tollervey, D. & Hurt, E. 2002. 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. *EMBO J,* 21(20)**,** pp 5539-47.

Ofengand, J., Bakin, A., Wrzesinski, J., Nurse, K. & Lane, B. G. 1995. The pseudouridine residues of ribosomal RNA. *Biochemistry and cell biology,* 73(11-12)**,** pp 915-924.

Osheim, Y. N., French, S. L., Keck, K. M., Champion, E. A., Spasov, K., Dragon, F., Baserga, S. J. & 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*.*Molecular Cell,* 16(6)**,** pp 943-954.

Parker, K. A. & Steitz, J. A. 1987. Structural analysis of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base pairing with pre-rRNA. *Mol Cell,* 7(8)**,** pp 2899-913.

Partow, S., Siewers, V., Bjørn, S., Nielsen, J. & Maury, J. 2010. Characterization of different promoters for designing a new expression vector in Saccharomyces cerevisiae. *Yeast,* 27(11)**,** pp 955-964.

Perez-Fernandez, J., Martin-Marcos, P. & Dosil, M. 2011. Elucidation of the assembly events required for the recruitment of Utp20, Imp4 and Bms1 onto nascent pre-ribosomes. *Nucleic Acids Res,* 39(18)**,** pp 8105-21.

Perez-Fernandez, J., Roman, A., De Las Rivas, J., Bustelo, X. R. & 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. & 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.

Philippsen, P., Thomas, M., Kramer, R. A. & Davis, R. W. 1978. Unique arrangement of coding sequences for 5 S, 5.8 S, 18 S and 25 S ribosomal RNA in *Saccharomyces cerevisiae* as determined by R-loop and hybridization analysis. *Journal of molecular biology,* 123(3)**,** pp 387-404.

Phipps, K. R., Charette, J. & Baserga, S. J. 2011. The small subunit processome in ribosome biogenesis-progress and prospects. *Wiley Interdiscip Rev RNA,* 2(1)**,** pp 1-21.

Raices, M. & D'Angelo, M. A. 2012. Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. *Nat Rev Mol Cell Biol,* 13(11)**,** pp 687-99.

Reichow, S. L., Hamma, T., Ferre-D'Amare, A. R. & Varani, G. 2007. The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res,* 35(5)**,** pp 1452-64.

Revzin, A. 1989. Gel electrophoresis assays for DNA-protein interactions. *Biotechniques,* 7(4)**,** pp 346-55.

Rodríguez-Galán, O., García-Gómez, J. J. & de la Cruz, J. 2013. Yeast and human RNA helicases involved in ribosome biogenesis: current status and perspectives. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms,* 1829(8)**,** pp 775-790.

Ruggero, D. & Pandolfi, P. P. 2003. Does the ribosome translate cancer? *Nat Rev Cancer,* 3(3)**,** pp 179-192.

Sardana, R., Liu, X., Granneman, S., Zhu, J., Gill, M., Papoulas, O., Marcotte, E. M., Tollervey, D., Correll, C. C. & 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)**,** pp e1002083.

Sauer, B. 1987. Functional expression of the cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae. Mol Cell Biol,* 7(6)**,** pp 2087-96.

Schafer, T., Strauss, D., Petfalski, E., Tollervey, D. & Hurt, E. 2003. The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. *EMBO J,* 22(6)**,** pp 1370-80.

Schilders, G., van Dijk, E. & 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.

Schmitt, M. E. & Clayton, D. A. 1993. Nuclear RNAase MRP is required for correct processing of pre-5.8S rRNA in *Saccharomyces cerevisiae*. *Mol Cell Biol,* 13(12)**,** pp 7935-41.

Schuch, B., Feigenbutz, M., Makino, D. L., Falk, S., Basquin, C., Mitchell, P. & 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.

Segault, V., Mougin, A., Gregoire, A., Banroques, J. & Branlant, C. 1992. An experimental study of *Saccharomyces cerevisiae* U3 snRNA conformation in solution. *Nucleic Acids Res,* 20(13)**,** pp 3443-51.

Shah, B. N., Liu, X. & 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.

Shajani, Z., Sykes, M. T. & Williamson, J. R. 2011. Assembly of bacterial ribosomes. *Annu Rev Biochem,* 80(501-26.

Shapiro, A. L., Vinuela, E. and Maizel, J. V., Jr. 1967. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochem Biophys Res Commun,* 28(5)**,** pp 815-20.

Singh, S. K., Gurha, P. & 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.

Smith, C. M. & Steitz, J. A. 1997. Sno storm in the nucleolus: new roles for myriad small RNPs. *Cell,* 89(5)**,** pp 669-672.

Smith, D. B. & Johnson, K. S. 1988. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene,* 67(1)**,** pp 31-40.

Soltanieh, S., Osheim, Y. N., Spasov, K., Trahan, C., Beyer, A. L. & Dragon, F. 2015. DEAD-box RNA helicase Dbp4 is required for small-subunit processome formation and function. *Mol Cell Biol,* 35(5)**,** pp 816-30.

Spahn, C. M., Beckmann, R., Eswar, N., Penczek, P. A., Sali, A., Blobel, G. & Frank, J. 2001. Structure of the 80S ribosome from Saccharomyces cerevisiae--tRNA-ribosome and subunit-subunit interactions. *Cell,* 107(3)**,** pp 373-86.

Spahn, C. M., Gomez-Lorenzo, M. G., Grassucci, R. A., Jorgensen, R., Andersen, G. R., Beckmann, R., Penczek, P. A., Ballesta, J. P. & Frank, J. 2004. Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. *EMBO J,* 23(5)**,** pp 1008-19.

Staub, E., Fiziev, P., Rosenthal, A. & Hinzmann, B. 2004. Insights into the evolution of the nucleolus by an analysis of its protein domain repertoire. *Bioessays,* 26(5)**,** pp 567-81.

Stead, J. A., Costello, J. L., Livingstone, M. J. & 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.

Stirnimann, C. U., Petsalaki, E., Russell, R. B. & Muller, C. W. 2010. WD40 proteins propel cellular networks. *Trends Biochem Sci,* 35(10)**,** pp 565-74.

TA, N., Bassler J Fau - Petfalski, E., Petfalski E Fau - Tollervey, D., Tollervey D Fau - Hurt, E. & E, H. 2002. - 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. *EMBO J J,* 21(20)**,** pp 5539-47.

Tarassov, K., Messier, V., Landry, C. R., Radinovic, S., Serna Molina, M. M., Shames, I., Malitskaya, Y., Vogel, J., Bussey, H. & Michnick, S. W. 2008. An in vivo map of the yeast protein interactome. *Science,* 320(5882)**,** pp 1465-70.

Tenge, V. R., Knowles, J. & Johnson, J. L. 2014. The ribosomal biogenesis protein Utp21 interacts with Hsp90 and has differing requirements for Hsp90-associated proteins. *PLoS One,* 9(3)**,** pp e92569.

Thoms, M., Thomson, E., Bassler, J., Gnadig, M., Griesel, S. & Hurt, E. 2015. The Exosome Is Recruited to RNA Substrates through Specific Adaptor Proteins. *Cell,* 162(5)**,** pp 1029-38.

Tollervey, D. 1987. A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA. *EMBO J,* 6(13)**,** pp 4169-75.

Tollervey, D., Lehtonen, H., Carmo-Fonseca, M. & Hurt, E. C. 1991. The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. *EMBO J,* 10(3)**,** pp 573-83.

Tollervey, D., Lehtonen, H., Jansen, R., Kern, H. & Hurt, E. C. 1993. Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell,* 72(3)**,** pp 443-457.

Tollervey, D. & Mattaj, I. 1987. Fungal small nuclear ribonucleoproteins share properties with plant and vertebrate U-snRNPs. *The EMBO J,* 6(2)**,** pp 469.

Torchet, C. & Hermann-Le Denmat, S. 2000. Bypassing the rRNA processing endonucleolytic cleavage at site A2 in *Saccharomyces cerevisiae. RNA,* 6(11)**,** pp 1498-508.

Turner, M. 2011. *An investigation of the C1D family of proteins in Saccharomyces cerevisiae.* PhD, University of sheffield.

Turowski, T. W. & Tollervey, D. 2015. Cotranscriptional events in eukaryotic ribosome synthesis. *Wiley Interdisciplinary Reviews: RNA,* 6(1)**,** pp 129-139.

Udem, S. A. & Warner, J. R. 1972. Ribosomal RNA synthesis in *Saccharomyces cerevisiae*. *J Mol Biol,* 65(2)**,** pp 227-42.

van Nues, R. W., Granneman, S., Kudla, G., Sloan, K. E., Chicken, M., Tollervey, D. & Watkins, N. J. 2011. Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. *The EMBO J,* 30(12)**,** pp 2420-2430.

Venema, J. & 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. & Tollervey, D. 1999. Ribosome synthesis in *Saccharomyces cerevisiae. Annu Rev Genet,* 33(261-311.

Venema, J., Vos, H. R., Faber, A. W., van Venrooij, W. J. & 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.

Voegtli, W. C., Madrona, A. Y. & Wilson, D. K. 2003. The structure of Aip1p, a WD repeat protein that regulates Cofilin-mediated actin depolymerization. *J Biol Chem,* 278(36)**,** pp 34373-9.

Vos, H. R., Faber, A. W., de Gier, M. D., Vos, J. C. & Raué, H. A. 2004. 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. *Eukaryotic Cell,* 3(6)**,** pp 1504-1512.

Wach, A., Brachat, A., Pohlmann, R. & Philippsen, P. 1994. New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast,* 10(13)**,** pp 1793-808.

Warner, J. R. 1989. Synthesis of ribosomes in *Saccharomyces cerevisiae. Microbiological Reviews,* 53(2)**,** pp 256-271.

Watkins, N. J. & 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 Interdisciplinary Reviews: RNA,* 3(3)**,** pp 397-414.

Wegierski, T., Billy, E., Nasr, F. & Filipowicz, W. 2001. Bms1p, a G-domain-containing protein, associates with Rcl1p and is required for 18S rRNA biogenesis in yeast. *RNA,* 7(9)**,** pp 1254-67.

Wells, G. R., Weichmann, F., Colvin, D., Sloan, K. E., Kudla, G., Tollervey, D., Watkins, N. J. & Schneider, C. 2016. The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans. *Nucleic Acids Res*.

Wery, M., Ruidant, S., Schillewaert, S., Lepore, N. & 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.

Wiederkehr, T., Pretot, R. F. & Minvielle-Sebastia, L. 1998. Synthetic lethal interactions with conditional poly(A) polymerase alleles identify LCP5, a gene involved in 18S rRNA maturation. *RNA,* 4(11)**,** pp 1357-72.

Wilson, D. S. & Szostak, J. W. 1999. In vitro selection of functional nucleic acids. *Annual review of biochemistry,* 68(1)**,** pp 611-647.

Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T. & Ramakrishnan, V. 2000. Structure of the 30S ribosomal subunit. *Nature,* 407(6802)**,** pp 327-339.

Wojda, I., Cytrynska, M., Frajnt, M. & Jakubowicz, T. 2002. Protein kinases CKI and CKII are implicated in modification of ribosomal proteins of the yeast Trichosporon cutaneum. *Acta Biochim Pol,* 49(4)**,** pp 947-57.

Woolford, J. L., Jr. & 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. & Baserga, S. J. 2001. An unexpected, conserved element of the U3 snoRNA is required for Mpp10p association. *RNA,* 7(6)**,** pp 904-19.

Wu, Z., Moghaddas Gholami, A. & Kuster, B. 2012. Systematic identification of the HSP90 candidate regulated proteome. *Mol Cell Proteomics,* 11(6)**,** pp M111 016675.

Yavuzer, U., Smith, G. C., Bliss, T., Werner, D. & Jackson, S. P. 1998. DNA end-independent activation of DNA-PK mediated via association with the DNA-binding protein C1D. *Genes Dev,* 12(14)**,** pp 2188-99.

Yokoyama, T. & Suzuki, T. 2008. Ribosomal RNAs are tolerant toward genetic insertions: evolutionary origin of the expansion segments. *Nucleic Acids Res,* 36(11)**,** pp 3539-51.

Yusupova, G. Z., Yusupov, M. M., Cate, J. H. & Noller, H. F. 2001. The path of messenger RNA through the ribosome. *Cell,* 106(2)**,** pp 233-41.

Zhang, C., Lin, J., Liu, W., Chen, X., Chen, R. & Ye, K. 2014. Structure of Utp21 tandem WD domain provides insight into the organization of the UTPB complex involved in ribosome synthesis. *PLoS One,* 9(1)**,** pp e86540.

Zhang, L., Wu, C., Cai, G., Chen, S. & Ye, K. 2016. Stepwise and dynamic assembly of the earliest precursors of small ribosomal subunits in yeast. *Genes Dev,* 30(6)**,** pp 718-32.

Zhao, R., Davey, M., Hsu, Y. C., Kaplanek, P., Tong, A., Parsons, A. B., Krogan, N., Cagney, G., Mai, D., Greenblatt, J., Boone, C., Emili, A. & Houry, W. A. 2005. Navigating the chaperone network: an integrative

map of physical and genetic interactions mediated by the hsp90 chaperone. *Cell,* 120(5)**,** pp 715-27.

Zhou, J., Liang, B. & Li, H. 2011. Structural and functional evidence of high specificity of Cbf5 for ACA trinucleotide. *RNA,* 17(2)**,** pp 244-50.

Zhu, J., Liu, X., Anjos, M., Correll, C. C. & 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.