BINDING OF IN VITRO RECONSTITUTED SNR30 RIBONUCLEOPROTEIN TO RIBOSOMAL RNA

TIMOTHY JOHN VOS
Bachelor of Science, University of Lethbridge, 2018

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

In

BIOCHEMISTRY

University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

BINDING OF IN VITRO RECONSTITUTED SNR30 RIBONUCLEOPROTEIN TO RIBOSOMAL RNA

TIMOTHY JOHN VOS

Date of Defence: April 21, 2021

Dr. U. Kothe Thesis Supervisor	Professor	Ph.D.
Dr. H.J. Wieden Thesis Examination Committee Member	Professor	Ph.D.
Dr. M. Roussel Thesis Examination Committee Member	Professor	Ph.D.
Dr. F. Bachand External Examiner Université de Sherbrooke Sherbrooke, Quebec	Associate Professor	Ph.D.
Dr. M. Gerken Chair, Thesis Examination Committee	Professor	Ph.D.

DEDICATION

To my parents, without whom I never would have shown up on time.

ABSTRACT

The creation of eukaryotic ribosomes is a complex process involving hundreds of assembly factors, including the small nucleolar RNA snR30. Through interacting with the unprocessed precursor ribosomal RNA (pre-rRNA), snR30 is essential for the maturation of pre-rRNA forming the small ribosomal subunit. To characterize snR30's interaction network, an *in vitro* approach was used to quantify the affinity of snR30 to the core H/ACA protein complex and the PIN endonuclease Utp23 revealing very tight interactions with affinities in the sub- and low-nanomolar ranges, respectively. Furthermore, the snR30 complex binds tightly to its primary binding site in ribosomal RNA called the C1 site, but only weakly to other predicted interaction sites. Utp23 binds tightly to its predicted interaction site in ribosomal RNA only if adjacent rRNA elements are also present. In conclusion, these findings serve as foundational work to further explore the mechanisms of snR30 during maturation of rRNA.

ACKNOWLEDGEMENTS

First, I would like to thank my lab mates who have somehow put up with me for my duration in the lab. Next, I would like to thank my incredible supervisor Dr. Ute Kothe. Without her continued support and assistance, none of this would have been possible. Furthermore, for my family for agreeing to host me these past six years of my postsecondary education. Their undying love and support have been a tremendous boon in helping me get here. Lastly, I would like to thank my committee and external examiner. The feedback from Dr. Russel and Dr. Roussel has been incredibly valuable in shaping the progress of my thesis and resolving problems. Dr. Wieden filling in for Dr. Russel on short notice is an incredibly kind gesture that I greatly appreciate. Dr. Bachand has been generous enough to agree to attend my defence and provide invaluable feedback from a different perspective. Lastly, Dr. Gerken deserves thanks for agreeing to chair the thesis defence.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Amp	Ampicillin
bp	Base pair
CV	Column Volume
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ES	Expansion segment
ETS	Externally transcribed spacer
GMP	Guanosine Monophosphate
iPPase	Inorganic pyrophosphatase
IPTG	Isopropyl β- d-1-thiogalactopyranoside
ITS	Internally transcribed spacer
ncRNA	Noncoding Ribonucleic acid
nt	Nucleotide
NTP	Ribonucleotide triphosphate
OD	Optical Density
PAGE	Polyacrylamide gel electrophoresis
PMSF	Phenylmethylsulfonyl fluoride
RNP	Ribonucleoprotein
rRNA	Ribosomal ribonucleic acid
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SHAPE	Selective 2' Hydroxyl Acylation analyzed by Primer Extension
snoRNA	Small nucleolar ribonucleic acid
SSU	Small subunit
TE	Tris and Ethylenediaminetetraacetic acid
TBS	Tris buffered saline
Utp	U3 protein

CHAPTER 1: INTRODUCTION[1]

1.1 snR30 Overview

Eukaryotic ribosomes are synthesized in the nucleolus with the help of hundreds of assembly factors. In a highly orchestrated fashion, multiple protein and RNA assembly factors are recruited to nascent pre-ribosomal RNA (pre-rRNA) and depart at defined times, whereas ribosomal proteins also associate in a defined order, eventually forming ribosomal subunits together with the matured rRNAs. Among the ribosome assembly factors, the nucleolar snR30 is an essential H/ACA RNA that is critical for the formation of the small ribosomal subunit, but only limited information is available on the mechanistic role of snR30 during the complex process of ribosome formation. Before discussing the function of snR30 in detail, we will first provide a brief overview of eukaryotic ribosome synthesis, in particular, small subunit (SSU) formation, to place the role of snR30 into context. Unless otherwise indicated, the information presented is based on the *Saccharomyces cerevisiae* model system used widely in ribosome biogenesis research (as reviewed in References [3-5]).

Ribosome biogenesis occurs co-transcriptionally on the pre-rRNAs, and snR30 is also thought to act co-transcriptionally during the processing of 18S rRNA. Three of the four mature rRNAs (18S, 5.8S, and 25S rRNA) are transcribed by RNA polymerase I as a single 35S precursor RNA (pre-rRNA, 47S pre-rRNA in humans), whereas the fourth rRNA, 5S rRNA, is transcribed independently by RNA Polymerase III. In addition to the mature rRNA sequences, the 35S pre-rRNA contains two external and two internal transcribed spacers abbreviated as ETS and ITS respectively, which are removed through several nucleolytic cleavage events such that these

[1] This chapter from section 1.1 to 1.6 inclusive are from Vos and Kothe (2020). U.K. and T.J.V. conceptualized the review; T.J.V. prepared the original draft including figures; U.K. and T.J.V. edited the text and figures. 1. Vos, T.J. and U. Kothe, *snR30/U17 Small Nucleolar Ribonucleoprotein: A Critical Player during Ribosome Biogenesis.* Cells, 2020. **9**(10).

regions are not present in the mature ribosome. The initial sites of pre-rRNA cleavage, which separate the 18S rRNA from the remaining pre-rRNA, are called A sites, specifically A_0, A_1, A_2, and A_3 in yeast and 01, 1, and 2 in humans (Figure 1.1) [6-8]. Notably, snR30 is involved in pre-rRNA cleavage at sites A_0, A_1, and A_2.

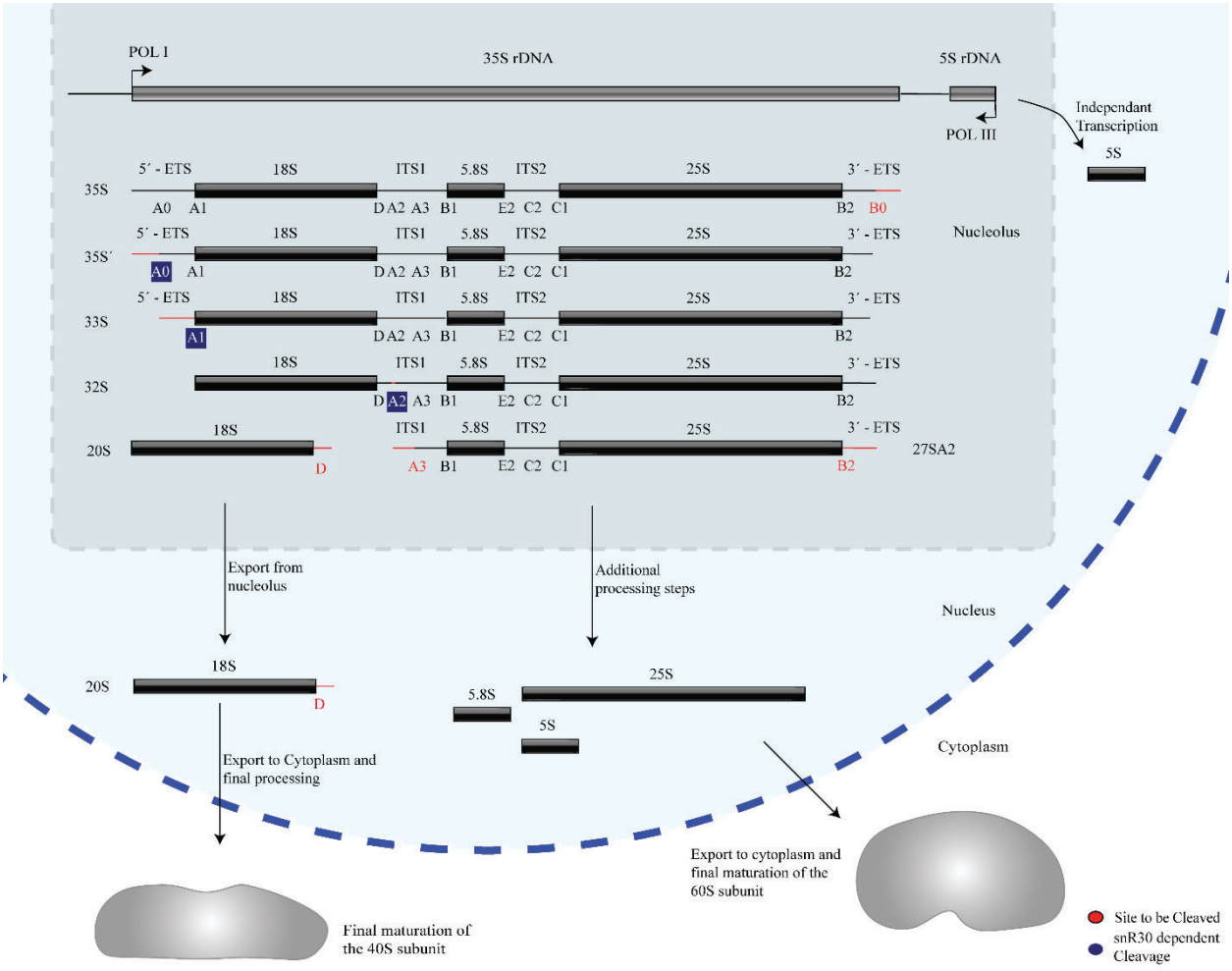

Figure 1.1. Overview of rRNA processing in yeast. The mature 18S, 5.8S, and 25S rRNAs are generated from a 35S pre-rRNA precursor through several nucleolytic cleavage events, shown here generating characteristic pre-rRNA intermediates. For each step, the subsequent site(s) of cleavage is highlighted in red. snR30 is essential for cleavage at sites A_0, A_1, and A_2, which are highlighted in blue.

During ribosome biogenesis, snR30 does not act alone, but rather forms a stable ribonucleoprotein (RNP) that belongs to the family of H/ACA small nucleolar RNPs (snoRNPs).

The snoRNPs represent a large portion of the ribosome biogenesis factors found in the nucleolus, and most snoRNPs chemically modify rRNA, but notably, this is not the case for the snR30 snoRNP. Occurring in the form of either C/D box or H/ACA box guide RNA systems, snoRNPs direct site-specific 2′-O-methylation and pseudouridylation, respectively [9, 10]. Each of these complexes is composed of one guide RNA, also called small nucleolar RNA (snoRNA), and four core proteins. The proteins Nop1 (Fibrillarin in vertebrates), Nop56, Nop58, and Snu13 (15.5K) assemble on C/D box guide RNAs, forming C/D snoRNPs [11-14], whereas the proteins Cbf5 (Dyskerin), Nop10, Gar1, and Nhp2 are components of the H/ACA box snoRNP together with H/ACA box guide RNA [15-17]. During ribosome assembly, the H/ACA snoRNPs introduce many pseudouridines in rRNA that play an important role in the maintenance of ribosome structure, stability, and translational fidelity [18-24].

Interestingly, of the more than one hundred snoRNAs known in yeast, only three guide RNAs are essential, namely U3, U14, and snR30. In addition, deletion of *S. cerevisiae* snR10 results in slow growth, and a cold-sensitive phenotype [25]. Remarkably, the primary function of these essential snoRNAs is not rRNA modification, but they are all directly or indirectly required for rRNA processing during ribosome synthesis. Additionally, snR30 may also play a role in cholesterol trafficking in higher eukaryotes [26-28]. In comparison to snR30, much more is known about the essential U3 snoRNA, a C/D box RNA that is responsible for the correct folding of the central pseudoknot in the 18S rRNA [29]. The central pseudoknot is an SSU rRNA structure conserved from prokaryotes to higher eukaryotes that connects the different domains of the SSU rRNA [30]. To facilitate pseudoknot formation, U3 makes multiple essential base pair interactions to the pre-rRNA in both the 5′ ETS and the 18S rRNA [29]. In particular, the box A and A′ motifs in the U3 snoRNA bind to the 18S rRNA regions that form the central pseudoknot. When U3 is

deleted, biogenesis of the small subunit is halted early due to the U3 snoRNP having a central role in biogenesis. This leads to accumulation of 23S pre-rRNA due to an inability to process at A_0, A_1, or A_2 [31]. Lastly, U14 is another C/D RNA that is essential for 18S rRNA formation [32]. U14 appears to have two distinct roles, as one region, which base pairs to the pre-rRNA and is not essential, guides the introduction of a 2′-O-methylation at C414 in 18S rRNA, whereas another region is essential and binds to the 18S rRNA extensively in the 5′ domain of the 18S rRNA [33-35]. Depletion of U14 leads to the exact same phenotype as depletion of U3 or snR30. Thereby, U14 resembles the H/ACA box RNA snR10 that also has dual functions by directing pseudouridine formation at position 2923 in 25S rRNA and by contributing to 18S rRNA processing. In cases where snR10 is lacking, the phenotype is similar to a knockout of the essential genes, except less pronounced. There is minor 35S pre-rRNA accumulation that leads to an increase in 23S pre-rRNA and a 21S pre-rRNA product. However, at permissive temperatures, this pre-rRNA is still processed, albeit slowly, into mature 18S rRNA [36].

Knowledge of ribosome biogenesis, and specifically small subunit biogenesis, has significantly increased due to recent advances in cryo-electron microscopy (cryo-EM), yielding high-resolution structures of the SSU processome complex [37-39], but no structures of snR30 bound to the SSU processome are available to date. The recent cryo-EM structures provide insight into the interactions within the SSU processome, which has led to advances in both the understanding of protein positioning and timings of large-scale rearrangements in the small ribosomal subunit during ribosome biogenesis. However, many of the early-acting factors, such as the snR30 snoRNP, are not present in the structures solved so far, presumably because the snR30 snoRNP has already dissociated before formation of the stable SSU processome. Therefore, no structural information is currently available on the interaction of snR30 with the SSU processome,

but instead, our knowledge is derived mostly from genetic and biochemical studies. For example, by tagging and purifying a known biogenesis factor, the timing of SSU processome association of this factor can be roughly estimated by identifying other co-purified factors [40, 41]. More recently, different ribosomal precursors have also been studied by expressing and purifying truncated and tagged pre-rRNA variants followed by proteomic analysis, providing additional information on when assembly factors associate and dissociate during pre-rRNA transcription [42-45]. Notably, taking such an approach, snR30 was detected in early ribosomal intermediates that arise before 18S rRNA transcription is complete, but it was no longer detected in later intermediates [42]. These findings suggest that snR30 acts relatively early during ribosome biogenesis when the structure of the SSU processome might still be rather flexible and labile.

The early stages of eukaryotic ribosome synthesis are characterized by a dynamic interaction network of several factors, including snR30. Ribosome biogenesis starts with the binding of the UTP-A complex to the 5′ ETS of the 35S pre-rRNA transcript. Subsequently, the U3 snRNP, UTP-B, and Mpp10 complexes associate [42, 43, 46]. While these complexes contain a large portion of the assembly factors, there are many individual proteins, small complexes, and modification enzymes that also act during this period. snR30 is part of one of these complexes that seems to only be present until the 3′ minor domain of 18S rRNA is transcribed [42]. The primary interaction point of snR30 is within expansion segment 6 (ES6) in the central domain, which is unstructured and therefore not resolved in the cryo-EM structures of the SSU processome [39]. The possible function of snR30 binding for pre-rRNA folding will be further discussed below.

By recruiting all these assembly factors, the pre-rRNA in the SSU processome is held in a state where the mature subunit is recognizable, but not fully formed, as evident in the cryo-EM structures [37-39], and snR30 may contribute to preventing premature interactions as further

discussed below. One of the most striking differences between the SSU processome structure compared to the mature 40S subunit is that the central pseudoknot is unable to form due to the presence of Sas10, Lcp5, the U3 snRNP, and other factors [38]. Furthermore, the SSU processome separates the four domains of the 18S rRNA into independent units which can fold separately [39, 47]. More recent structural information from the thermophilic fungus *Chaetomium thermophilum* has suggested that the 5′, central, 3′ major, and 3′ minor domains of 18S rRNA are not sequentially integrated into the pre-40S subunit, but that the 5′ domain may join the SSU processome last; however, it remains to be investigated whether this finding holds true in other organisms [48]. Thus, the early stages of SSU formation are characterized by dynamic interactions and several conformational changes in rRNA that are facilitated by many factors, including snR30.

Research into ribosome biogenesis is a rapidly evolving field as the concerted functions of hundreds of poorly understood factors and steps including snR30 need to be identified. In this review, we consolidate the information known about snR30 and develop hypotheses about its role during ribosome formation. Thereby, we hope to inspire future research into the function and mechanism of snR30, which will fill a critical gap in our understanding of ribosome formation.

1.2 snR30 Biogenesis, Sequence, and Structure

1.2.1 Transcription and Processing

In unicellular eukaryotes, such as *S. cerevisiae,* snR30 is transcribed as an independent gene by RNA Pol II, the mRNA producing polymerase. Contrary to mRNA, the snoRNA undergoes processing to remove the polyA tail by initially binding the Nrd1-Nab3-Sen1 complex, which presents the snR30 3′ end to the exosome for exonucleolytic processing [49, 50]. On the 5′ end, snR30 possesses a trimethyl guanosine cap like most other H/ACA snoRNAs in yeast [26, 51], which is added by the conserved methyltransferase Tgs1 [52].

In metazoans such as humans, transcription and processing of U17a and U17b, two human homologs of snR30, is significantly different. Both U17a and U17b are encoded by the U17HG gene upstream of the cell cycle regulatory gene RCC1 [53]. Interestingly, the U17HG gene harbors U17a and U17b in two introns; however, the U17HG gene seems not to encode a protein and only the U17 sequences are conserved [54]. However, in other species, the location of the U17 gene varies such as in *X. laevis,* where U17 is transcribed in all six introns of the r-protein S8 gene [55]. Following transcription of the host gene, the excised intron containing U17 is then processed exonucleolytically at both the 5′ and 3′ ends in both Xenopus and HeLa cells, implying a conserved mode of maturation [56, 57]. The 3′ end is processed by the exosome [58], whereas the 5′ end is processed by an unknown endonuclease [56, 59]. Further information on snoRNA biogenesis and processing events have been reviewed by Kufel and Grzechnik [60].

1.2.2 snR30 snoRNP Formation and Nucleolar Localization

In yeast, the core snR30 RNP comprising the RNA and the four canonical H/ACA proteins is formed in the same way as other H/ACA snoRNPs that modify rRNA. This process begins with formation of a protein complex initiated by the binding of the protein Shq1 to Cbf5 in the cytoplasm [61]. Shq1 associates with the RNA-binding domain of Cbf5 mimicking interactions of H/ACA snoRNA, which is known to be strongly bound by Cbf5 [62, 63]. The Shq1-Cbf5 complex is then shuttled into the nucleus where the majority of RNP maturation takes place. During this process, Nop10, Nhp2, and an assembly factor called Naf1 bind, forming a large protein complex [64]. At the site of snR30 transcription, a pair of ATPases catalyzes the release of Cbf5 from Shq1, freeing the enzyme to tightly bind the snoRNA in its place [65]. This interaction is mediated by Naf1 binding to the large subunit of RNA polymerase II [66]. After binding of Cbf5 to the H/ACA guide RNA, the subsequent maturation step involves the competitive binding of Gar1 to Cbf5,

resulting in the displacement of Naf1, which is recycled back to the cytoplasm for further maturation. The now fully assembled H/ACA snoRNP is shuttled from Cajal bodies to the nucleolus by Nopp140 [64, 67]. Localization signals for transport into the nucleolus appear to be located in the H and ACA boxes, as well as the general structure of the guide RNA, at least for vertebrates [68, 69]. For U17, the ACA box is required for assembly in vitro and presumably in vivo, signifying that nucleolar localization is dependent upon formation of the RNP [70]. As only mature snoRNPs are found within the nucleolus, maturation must occur between transcription in the nucleoplasm and integration into the nucleolus [71]. It is possible that some steps of snR30 maturation may occur within Cajal bodies similarly to the maturation of other snoRNAs. Immature C/D box RNPs as well as a subgroup of H/ACA RNAs called H/ACA small Cajal body RNAs (scaRNA) can be readily detected within Cajal bodies, and some evidence suggests that H/ACA snoRNPs may in general also traverse Cajal bodies [72-74]. In conclusion, current evidence suggests that the maturation of the snR30/U17 RNP follows the same steps as canonical H/ACA guide RNA as all H/ACA RNAs including snR30 assemble with the same proteins.

1.3. Conservation of snR30/U17 Structure, Sequence, and Motifs

Typical H/ACA guide RNAs in eukaryotes share a similar secondary structure comprised of two hairpins connected by a hinge region. The two hairpins are followed by two conserved sequence motifs, the H Box (ANANNA) and the ACA Box, respectively. While most known H/ACA RNAs contain two hairpins in eukaryotes, there are instances of H/ACA guide RNAs having one or three hairpins in selected eukaryotic organisms as well as in archaea. snR30 is an unusual H/ACA guide RNA that has two primary hairpins, the 5′ and 3′ hairpins, but also possesses a third internal hairpin as well as a 41-nt leader sequence at its 5′ end (Figure 1.2). Notably, the 5′ hairpin of snR30 is much longer than a standard H/ACA hairpin such that *S. cerevisiae* snR30 has

an unusual length of 606 nt, almost triple the length of the average yeast H/ACA guide RNA (~200 nt). While not found in all H/ACA snoRNAs, the internal hairpin and the 41-nt leader are both features that are also present in some other H/ACA guide RNAs [75]. Interestingly, snR30 lacks an unpaired internal bulge following the first stem of the 5′ hairpin, a feature known as the pseudouridylation pocket in standard H/ACA guide RNAs. In contrast, the 3′ hairpin contains a single-stranded bulge like all other H/ACA guide RNA hairpins, and the top of the bulge is located at a conserved 14–16 nucleotide distance from the base of the hairpin and the ACA box [76]. In H/ACA snoRNAs directing pseudouridylation, this distance is important for properly positioning the guide RNA on the Cbf5-Nop10-Nhp2 binding surface, allowing binding of the target RNA to the pseudouridylation pocket and positioning of the target uridine into the active site of Cbf5 [63, 76]. In the 5′ hairpin of snR30, the only similar bulge occurs too far away from the base of the stem and the H box for correct positioning of Cbf5. Accordingly, no pseudouridine has been suggested to be introduced by the 5′ hairpin of snR30. As is the case for canonical H/ACA snoRNAs, and based on the location of the H and ACA Boxes, snR30 is expected to bind one set of the H/ACA core proteins (Cbf5, Nop10, Gar1, and Nhp2) to each of the 5′ and 3′ hairpins, resulting in a predicted 2:1 stoichiometry between the proteins and the snR30 RNA. Hence, despite its elongated structure, the only unique aspect of snR30 compared to modification H/ACA snoRNAs is that its 5′ hairpin does not have any known RNA targets.

Human U17 RNA (207 nt) is shorter than yeast snR30 and comprises four hairpins (Figure 1.3B), forming a secondary structure consisting of a 5′-variable domain and a 3′-conserved domain [26, 75, 77]. Thus, the comparison of yeast snR30 and human U17 can reveal functionally important regions of this conserved H/ACA snoRNA. During evolution, the 5′ region of snR30/U17 was compacted to a smaller size, effectively reducing transcriptional cost, which is

similar to the general trend of guide RNA shortening between single-cell and complex eukaryotes [78]. In humans, the 5′ end of U17 is composed of two stems of similar size prior to the H box. Whereas the H box of U17 is not required for in vitro RNP formation, the H box of snR30 is critical for accumulation of snR30 in vivo [70, 75]. U17 also contains an internal hairpin, although it is much smaller than that of snR30. Since the 5′ structure of snR30/U17 is not conserved, mutational studies investigated whether the 5′ and internal hairpins of yeast snR30 are critical for cell viability [75]. Indeed, both the 5′ hairpin and the internal hairpin can be individually deleted without affecting cell viability, and cell viability was maintained at a reduced level when replacing both the 5′ and internal hairpins with the 5′ hairpin of another box H/ACA RNA. Therefore, the 5′ hairpin and the internal hairpin of yeast snR30 are not directly responsible for its essential role within the cell. In contrast to the 5′ domain, the 3′ hairpin of U17 is extremely similar to that of snR30 as they both possess a structure identical to a standard H/ACA guide RNA hairpin including an unpaired bulge. As further outlined below, the conserved nature of the 3′ hairpin already indicates that this region in snR30/U17 is functionally most important.

By aligning the sequences of the snR30/U17 species from yeast, *Xenopus*, and humans, two strongly conserved sequence motifs in the 3′ hairpin were discovered and called m1 and m2, which are critical for ribosome formation [75]. The m1 and m2 regions are located in the non-productive, unpaired ‘pseudouridylation pocket’ on the 3′ hairpin of snR30/U17. However, rather than being located on the distal side of the pocket where modification H/ACA guide RNAs bind their targets, they are located on the basal side of the pocket. Two complementary sequences in 18S rRNA were discovered and designated as rm1 and rm2, and mutational studies confirmed Watson-Crick base pairing between the m1/rm1 and m2/rm2 sequences that is necessary for pre-rRNA processing [75]. While there is some variation in the m1 and m2 sequences across eukaryotes, these are always

matched by compensatory mutations in the rm1 and rm2 motifs in 18S rRNA (Table 1.1), underlining the importance of this base-pairing of snR30/U17 with 18S rRNA for ribosome biogenesis.

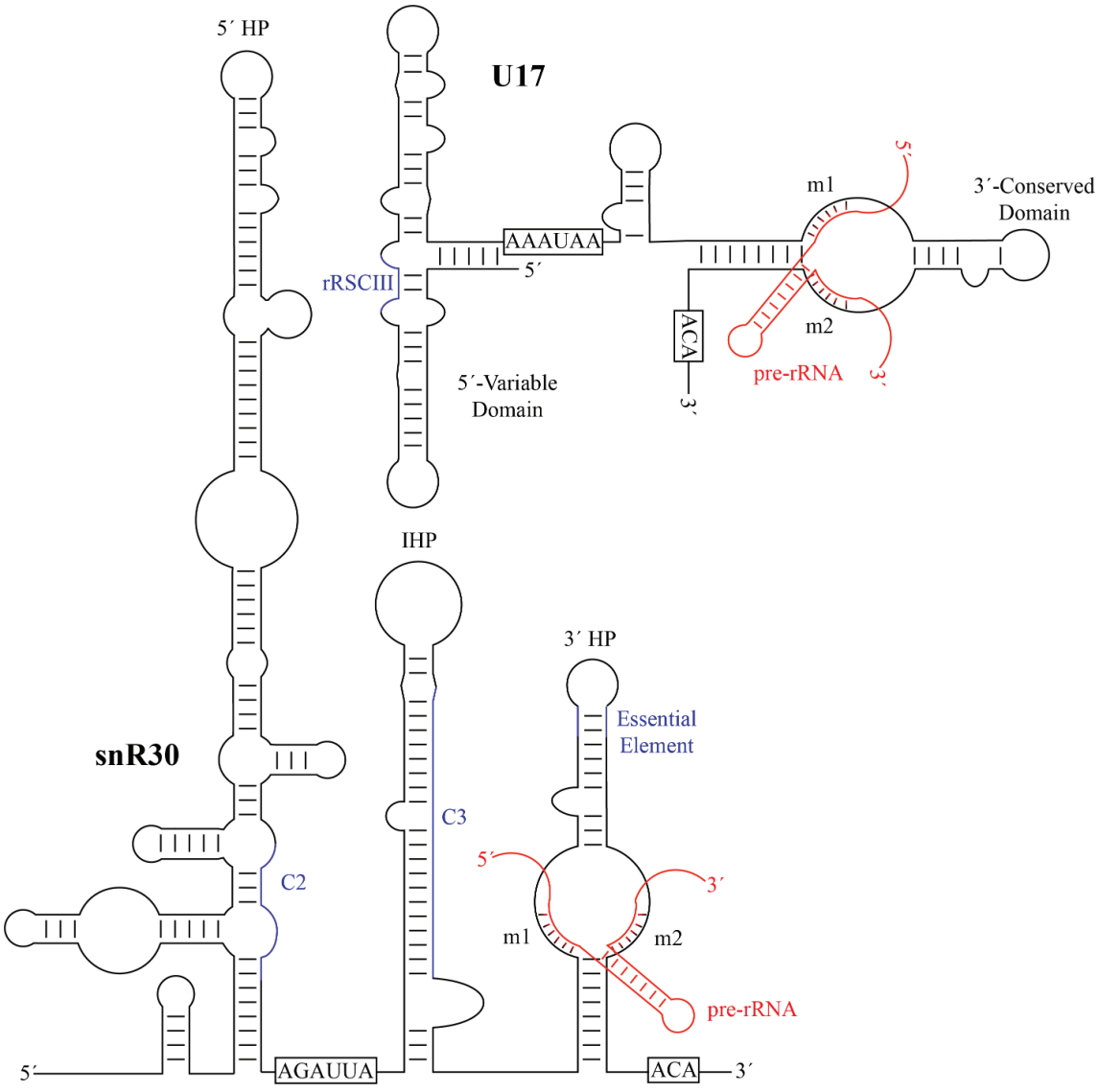

Figure 1.2. Secondary structures of yeast snR30 and human U17 and their modus operandi of pre-RNA binding. The H and ACA sequences, that characterize each H/ACA snoRNA, are boxed. The m1 and m2 motifs are labelled and the base-pairing to target pre-rRNA is shown in red. Additional regions of snR30/U17 predicted to have a function such as forming further interactions with 18S rRNA are depicted in blue (C2 and C3 in *S. cerevisiae* snR30, rRSCIII in human U17). *HP*, hairpin; *IHP*, internal hairpin. Minor bulges and imperfect base-pairing are not represented. snR30 was adapted from Atzorn et al. [75], and U17 was adapted from Ruhl et al. [79].

In addition to the critical and highly conserved m1 and m2 regions, additional elements of snR30 or U17 have been identified that also bind to 18S rRNA, although these secondary interactions are not conserved across all species. In *S. cerevisiae*, crosslinking, ligation, and sequencing of hybrids (CLASH) uncovered additional areas of interaction between snR30 and 18S rRNA [35]. Notably, two of the strongest interaction sites, called C2 and C3, also contain regions of sequence complementarity between snR30 and 18S rRNA and are conserved among fungi (Figure 1.3). First, a region in the 5′ hairpin of snR30 is proposed to interact with helix 1 in expansion segment 6 (ES6) of 18S rRNA in close vicinity to the interaction of the m1 and m2 regions with helix 3 of ES6. However, the importance of this interaction remains to be investigated since the 5′ hairpin of snR30 is dispensable. Second, a 19-nt region within the internal hairpin of snR30 has the potential to base-pair to expansion segment 7 (ES7) in 18S rRNA, which is also conserved in *Schizosaccharomyces pombe*. Whereas these interaction sites are likely specific to fungi, other putative contacts between U17 and 18S rRNA have been reported in vertebrates. In humans, the U17 rRCSIII sequence in stem 2 of the 5′ domain is predicted to base-pair with 18S rRNA at positions 967–976 (Figure 1.2) [80], and this sequence complementarity is conserved not only in mammals, but also birds, fish, amphibians, and reptiles. An additional element in stem 1 of the 5′ domain, called rRCSI, may form 12 base-pairs to the 18S rRNA, but is only conserved in fish and amphibians, not in mammals [80]. It remains unknown whether these predicted contacts between vertebrate U17 and 18S rRNA form in vivo and whether they are of functional importance for ribosome biogenesis. Due to the divergence of snR30/U17 sequence and structure over evolution, different interactions with 18S rRNA may have formed that may serve similar functions in stabilizing binding of this H/ACA snoRNA to the SSU processome.

Table 1.1. Comparison of snR30/U17 RNA across eukaryotes.

Organism	Common Name	Gene Name	Length (nt)	m1	m2	Accession Number	Reference
Apteryx rowi	Kiwi	*E1/U17*	205	AUAUUCCUA	AAACCAU	XR_003255265.1	N/A
Caenorhabditis elegans	Round worm	*CeN96*	221	AUAUUCCU**U**	AAACCAU	AY948718.1	[81]
Caretta caretta	Turtle	*U17*	213	AUAUUCCUA	AAACCAU	AJ306558	[80]
Candida glabrata	Yeast	*snR30*	576	AUAUUCCU**G**	AAACCAU	URS00006E7B8C	[82] [a]
Danio rerio	Zebra fish	*E1/U17*	>140	AUAUUCCUA	AAACCAU	LR812563.1	N/A
Fugu rubripes	Puffer fish	*U17*	218	AUAUUCCUA	AAACCAU	X94942	[83]
Homo sapiens	Humans	*U17*	207	AUAUUCCUA	AAACCAU	L16791	[75]
Kazachstania naganishii	Yeast	*snR30*	589	AUAUUCCUA	AAACCAU	URS0000BE6F45	[82] [a]
Mus musculus	Mouse	*U17/SNORA73*	203	AUAUUCCUA	AAACCAU	XR_003836540.1	N/A
Pelomedusa subrufa	Turtle	*U17*	215	AUAUUCCUA	AAACCAU [c]	AJ306565	[80]
Saccharomyces cerevisiae	Yeast	*snR30*	606	AUAUUCCUA	AAACCAU	NR_132204	[75]
Saccharomyces pombe	Yeast	*U17*	325	AUAUUCCUA	AAACCAU	AJ544685	[75]
Tetrahymena thermophila	Algae	*U17*	~240 [b]	AUAUUCCU**G**	AAACCAU	AJ544686	[75]
Xenopus laevis	Frog	*U17*	222	AUAUUCCUA	AAACCAU	X71081	[55]

Non-conserved nucleotides in the sequences from *C. glabrata* and *T. thermophila* are highlighted in bold. Accession numbers are from GenBank except for *S cerevisiae*, *K. naganishii*, and *C. glabrata*, which are from NCBI or RNA Central. References refer to the publication in which each motif was identified; no reference is provided, indicated as N/A, when the RNA sequence was only identified through a BLAST search at NCBI. [a] The motif was found manually in the sequences retrieved from RNA Central. [b] Accession number provides only partial sequence. Atzorn et al. [75] purified and sequenced the 5′ end of *T. thermophila* snR30, and its length was estimated based on gel size. [c] The *Psu* sequence has two A insertions 5′ of the motif, however, the rRNA motif stays the same, indicating that the insertion does not affect base pairing.

1.4 Role of snR30/U17 in rRNA Processing

1.4.1 Site-Specific Binding to the 18S rRNA

To promote formation of the small ribosomal subunit, snR30/U17 binds to the pre-rRNA and enables cleavages at sites A_0, A_1, and A_2 (Figure 1.2) [27, 84]. However, the molecular mechanism of the snR30/U17 snoRNP, including the details of its interaction with the SSU processome, remain unknown. As mentioned earlier, snR30/U17 base-pairs to expansion segment 6 of the 18S rRNA with the m1 and m2 motifs, which flank both sides of the basal end of the pocket on the 3′ hairpin. This snoRNA–rRNA interaction is markedly different from rRNA recognition by modification H/ACA RNAs, which bind to their target using the distal half of the pseudouridylation pocket. Therefore, it is this inverse binding from snR30/U17 that prevents it from introducing a pseudouridine into the 18S rRNA (Figure 1.3) [75]. It has been speculated that this change in orientation allows for snR30 to stay on the SSU processome longer, until it is specifically removed by an external factor. Notably, the base-pairing interaction between m1/rm1 and m2/rm2 is essential, indicating that the binding of snR30's 5′ hairpin to ES6 of the rRNA is of functional importance [85]. Binding of snR30 to ES6 sequesters the rm1 and rm2 regions of the pre-rRNA and the sequence between rm1 and rm2 is predicted to form a new hairpin structure in ES6 that is not present in the mature 18S rRNA, as observed in the ribosome (Figure 1.3) [85]. At the same time, formation of this new helix would result in the unwinding of ES6 hairpin 3 (ES6H3) and possibly also ES6 hairpin 2 (ES6H2). In addition to the interaction of snR30 with the rm1 and rm2 sequences in 18S rRNA, other interactions between the snR30 snoRNP and the SSU processome will likely occur and stabilize the interaction, but these interactions might differ between organisms. In fungi, snR30 was observed to crosslink to the C2 and C3 sites which are in relatively close proximity to the rm1 and rm2 sequences (also called C1) in the secondary structure

of 18S rRNA (Figure 1.3A) [35]. The ES6 structure is not resolved in the early SSU processome structures when snR30 is expected to bind, but recent structural information on the 90S pre-ribosome to 40S subunit transition in *C. thermophilum* suggest that this region becomes further organized during cleavage of the A1 site [86]. These later intermediates indicate that the C1, C2, and C3 interaction sites of *S. cerevisiae* snR30 with 18S rRNA are located in close three-dimensional proximity on the surface of the early SSU processome, allowing simultaneous interaction with the snR30 snoRNP. snR30 alters the secondary structure of 18S rRNA in ES6, but the functional importance of this conformational change is unknown. By preventing the 18S rRNA from adopting its mature conformation, snR30 may hold the SSU processome in a higher energy state until additional factors cause the release of snR30 from the SSU processome (vide infra). Interestingly, a similar restructuring of the pre-rRNA is also seen upon U3 binding to the 5′-ETS region of the pre-rRNA, promoting formation of the central pseudoknot. Therefore, the essential snoRNAs (snR30, U3, U14, as well as snR10) may generally bind to the transcribing pre-rRNA acting as rRNA chaperones by keeping the rRNA in a particular conformation to avoid misfolding of the rRNA [29]. The change in ES6 conformation induced by snR30 could be a critical signal transmitted through the rRNA to sites A_0, A_1, and A_2, allowing the SSU processome to surpass a checkpoint such that the next step along the maturation pathway, namely pre-rRNA processing, may occur.

Another possibility is that the snR30-induced conformational change in rRNA influences the recruitment of protein assembly factors. Certain factors may only bind to the particular rRNA conformation induced by snR30 and will thus depend on snR30 association with the ribosome. Other proteins may be unable to bind to this snR30-induced rRNA conformation and will therefore

depend on snR30 dissociation. Thus, the interaction of snR30 with pre-rRNA may regulate the timing of protein association with the SSU processome.

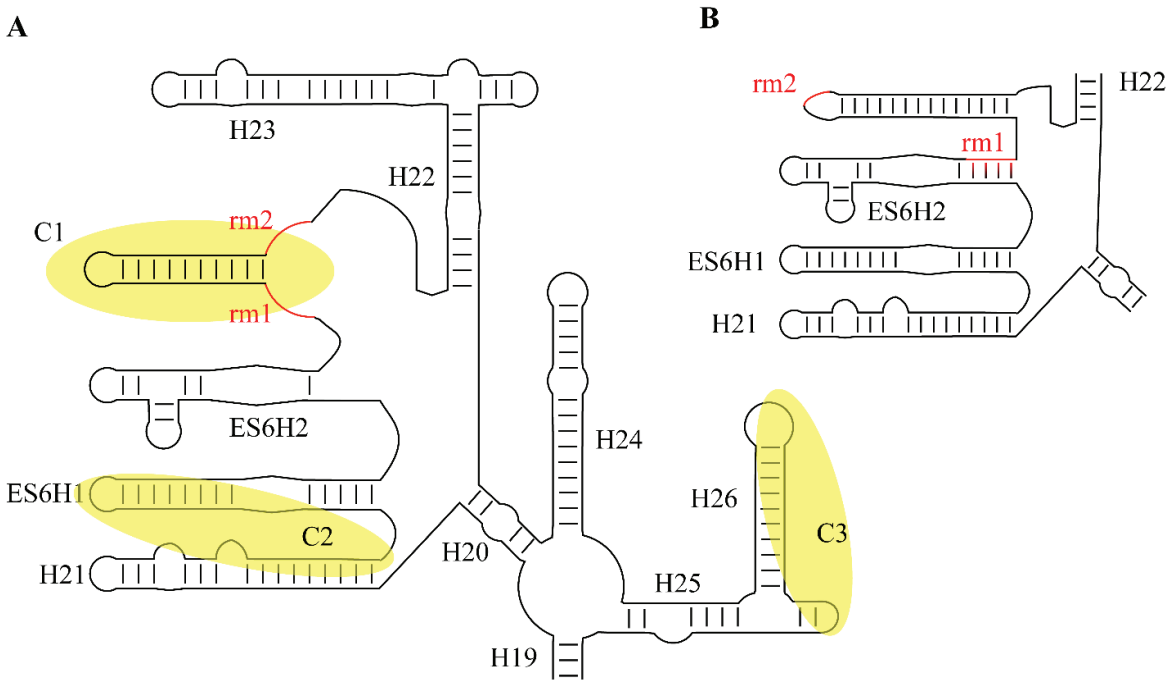

Figure 1.3. Conformational changes within ES6 of the 18S rRNA upon interaction with snR30. (**A**) Predicted rRNA conformation while snR30 base-pairs to the sequence elements rm1 and rm2 (red, crosslinking site C1) of the SSU processome. In addition, the crosslinking sites C2 and C3 are highlighted in yellow that were observed for *S. cerevisiae* snR30 (compare to Figure 1.2 for the corresponding crosslinking sites in snR30). (**B**) rRNA secondary structure of ES6 as visualized in cryo-EM models of the mature ribosome showing the different base-pairing in absence of snR30 [87].

1.4.2 snR30 Protein Recruitment

Since snR30 has no endonuclease activity itself, it must indirectly promote pre-rRNA cleavage events during ribosome biogenesis. To understand the molecular mechanism of snR30, it is therefore mandatory to dissect the interaction network of snR30 not only with rRNA, but also with additional ribosome assembly factors, including putative endonucleases which may be recruited by snR30. Notably, the endonuclease(s) responsible for rRNA cleavage at the A_0, A_1, and A_2 sites have not been unambiguously identified so far. The PIN endonuclease Utp24 has been proposed

to catalyze the cleavages at sites A_1 and A_2 [88, 89], but Rcl1 also has the ability to cleave A_2 in vitro [90].

Several assembly factors have been identified as interaction partners of snR30 through immunoprecipitations (Figure 1.4). Besides the H/ACA core proteins Cbf5, Nop10, Gar1, and Nhp2, snR30 also binds to Nop6 [84, 91], the DEAD-box helicases Has1 [92] and Rok1, the PIN domain endonuclease Utp23, and Kri1 [93, 94]. Additional interactions of snR30 were reported with ribosomal proteins S9 and S18, and the histones H2B and H4, but these may possibly represent unspecific interactions [84, 91].

The interaction of snR30 with Utp23 is of particular interest as snR30 and Utp23 together may serve as an essential assembly platform to facilitate pre-rRNA processing and ribosome formation. The essential 3′ hairpin and the internal hairpin of snR30 strongly crosslink to the protein Utp23, which also directly interacts with the H/ACA core protein Nhp2 [88]. In addition to the core components of the snR30 snoRNP, Utp23 also binds Rok1, Rrp7, and Utp24 in both yeast and human, underlining the role of Utp23 as a critical assembly hub during ribosome biogenesis [88]. Whereas snR30 is required for the incorporation of both Utp23 and Kri1 into the pre-ribosome, interestingly, Utp23 is in turn needed to later release snR30 from the pre-ribosome [94]. Like Utp24, a putative rRNA endonuclease, the snR30-interaction partner Utp23 also contains a PIN domain endonuclease fold. However, the functional importance of this domain remains unknown as yeast Utp23 does not possess the catalytic residues required for endonucleolytic cleavage [95]. In contrast, human Utp23 (hUtp23) contains catalytic residues and these are essential for cell viability, suggesting that Utp23 may possibly play additional roles in human ribosome formation compared to yeast [88]. Obviously, the direct interaction of Utp23 with the catalytically active endonuclease Utp24 may be responsible for mediating the role of snR30 in facilitating pre-rRNA

processing. In this context, it is important to note that Utp23 will not simply recruit Utp24 to the pre-ribosome as Utp24 is known to interact early with the pre-ribosome already while the 5′-ETS is being transcribed [96]. Therefore, the recruitment of the snR30 snoRNP together with Utp23 to the pre-ribosome is likely followed by the subsequent interaction of Utp23 and Utp24 on the pre-ribosome, which could lead to a re-positioning of Utp24 within the pre-ribosome. Since Utp24 strongly crosslinks and directly interacts with the U3 snoRNA [89], further conformational changes in the pre-ribosome could be indirectly induced by snR30 and Utp23. Thus, snR30, Utp23, and Utp24 could be critical in coordinating conformational changes across rRNA domains within the three-dimensional (3D) structure of the pre-ribosome leading to pre-rRNA cleavage at sites A_0, A_1, and A_2 by Utp24, Rcl1, or an unknown endonuclease.

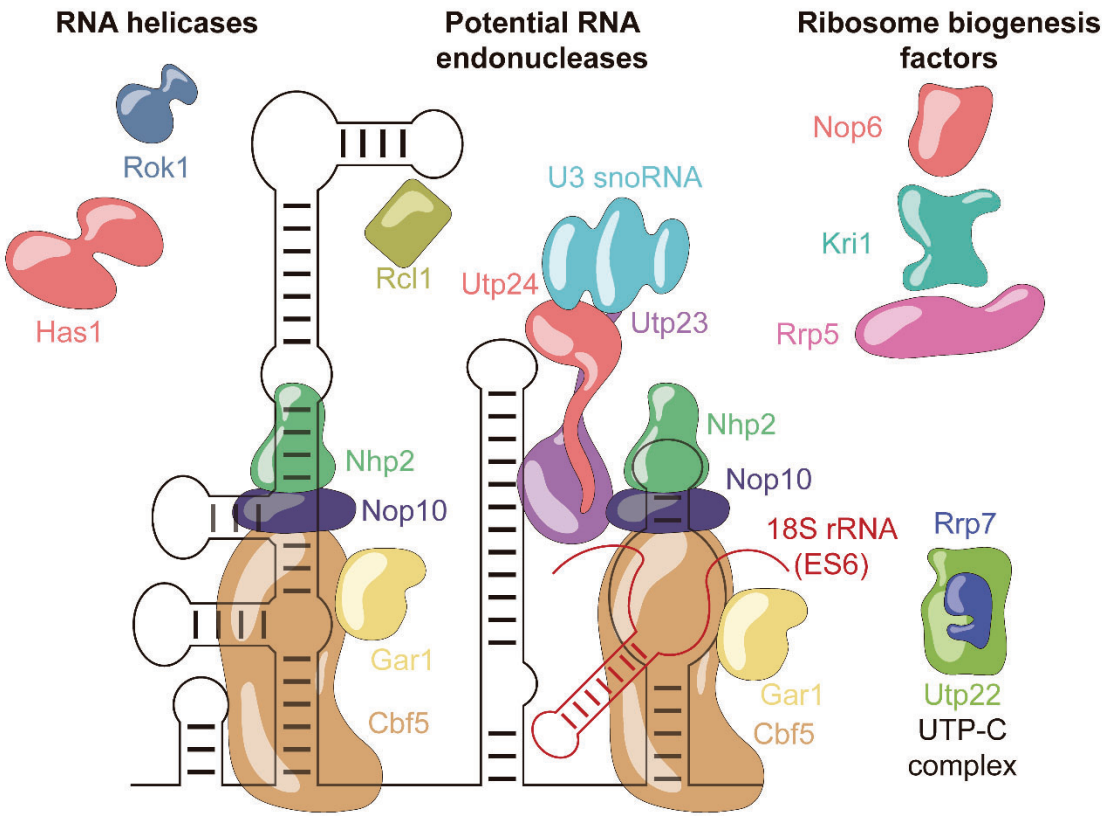

Figure 1.4. Schematic representation of the snR30 protein interaction network. snR30 snoRNA and its binding partners relevant to ribosome biogenesis are sorted and highlighted by different colors according to protein function. snR30 RNA is displayed in black, and pre-rRNA is in red.

1.4.3 Hypothetical Function of snR30 during Ribosome Formation

As outlined above, snR30 acts as a critical assembly hub during ribosome biogenesis by interacting both with 18S rRNA and by binding critical assembly factors. Both through the snR30-rRNA (Figure 1.3) as well as the snR30–protein interactions and networks (Figure 1.4), snR30 can directly or indirectly facilitate processing of pre-rRNA. Accordingly, we propose the following two hypotheses regarding the molecular mechanism of snR30, which are also summarized in Figure 1.5.

1.4.3.1 snR30 Mechanism – Hypothesis 1

Upon binding of snR30 with its 3′ hairpin to the expansion segment 6 (ES6) of 18S rRNA, snR30 induces conformational changes in the pre-rRNA, acting as an rRNA chaperone. These conformational changes caused by snR30 may include both the formation of specific structures such as the new helix in ES6 (Figure 1.3) as well as the unfolding of prematurely formed rRNA structures. The snR30-induced changes in pre-rRNA may be either transmitted directly as rRNA conformational changes through the SSU processome or may indirectly enable the recruitment of additional ribosome assembly factors to the snR30-induced conformation of pre-rRNA, leading to the correct positioning of the pre-rRNA cleavage sites A_0, A_1, and A_2 relative to the responsible endonucleases, such that snR30 enables pre-rRNA processing.

1.4.3.2 snR30 Mechanism – Hypothesis 2

The snR30 snoRNP strongly interacts with Utp23 through direct RNA–protein interaction as well as Nhp2-Utp23 protein–protein interaction, thereby recruiting Utp23 to the SSU processome, where Utp23 coordinates the binding and stabilization of additional ribosome assembly factors such as Rok1 and Kri1. Most importantly, Utp23 will interact with the SSU processome-bound

PIN domain endonuclease Utp24, possibly inducing conformational changes in Utp24 and the SSU processome, enabling processing of pre-rRNA at sites A_0, A_1, and A_2.

Importantly, these two hypotheses are not mutually exclusive, and it can rather be envisioned that both mechanisms occur during ribosome biogenesis, leading to concerted conformational changes in both pre-rRNA and protein assembly factors that enable pre-rRNA processing.

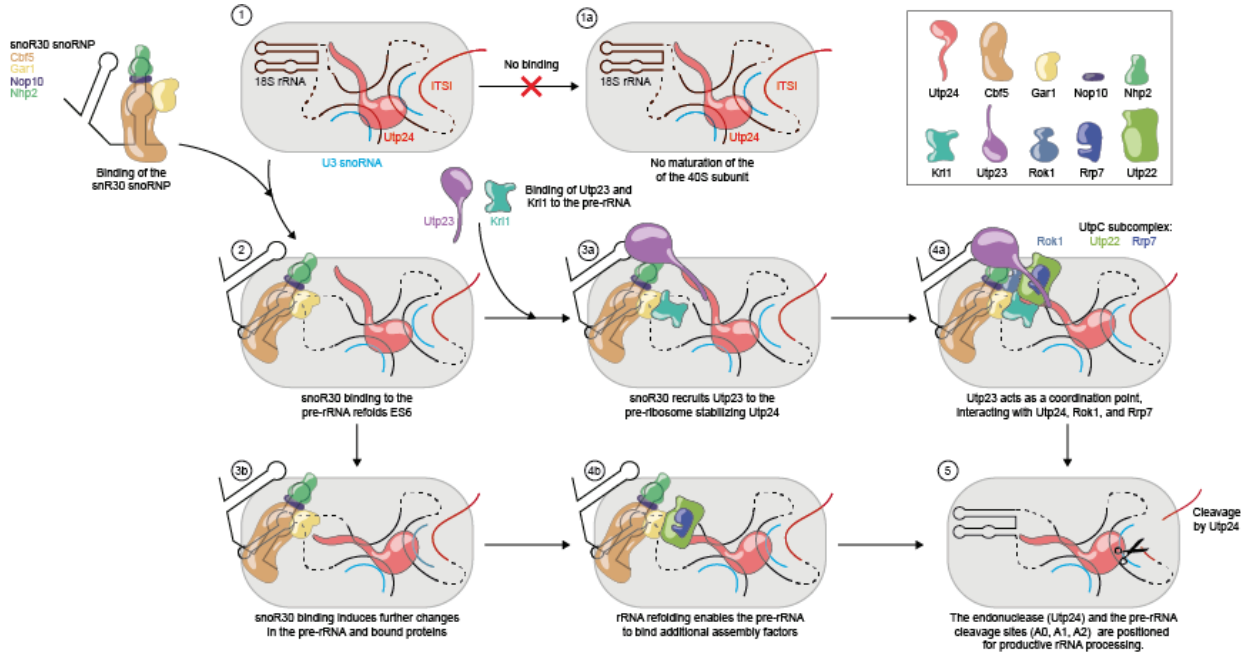

Figure 1.5. Putative molecular mechanisms of snR30 during ribosome biogenesis. A schematic of the SSU processome is shown as a light gray background with Utp24 (cyan) bound to the 18S rRNA and U3 snoRNA (**1**). In the absence of snR30, no processing occurs (**1a**). Base-pairing of snR30 to the rm1 and rm2 sites in expansion segment 6 (ES6) of 18S rRNA causes the two hairpins in ES6 to unfold and to instead refold into a single hairpin (**2**). Once bound to the SSU processome, snR30 likely recruits the proteins Utp23 and Kri1 to the SSU processome (**3a**). Utp23 may act as a coordination point interacting with several other proteins such as Rrp7, Rok1, and the endonuclease Utp24 (**4a**). These interactions could lead to a stabilization and possible reorganization of the SSU processome, allowing processing to occur, e.g., by bringing the endonuclease(s) (Utp24 or Rcl1) to the pre-rRNA cleavage sites (**5**). Alternatively, or in addition, the rearrangement of ES6 upon snR30 binding as well as other folding and unfolding events in rRNA caused by snR30 may induce conformational changes throughout the 18S pre-rRNA (**3b**). This could lead to repositioning of already bound factors like Utp24 and to the recruitment of essential factors like the UtpC subcomplex (**4b**). Ultimately, these conformational changes in the SSU processome correctly position the endonucleases relative to the pre-rRNA cleavage site within the SSU processome (**5**). snR30 and the 18S section of the pre-rRNA is black, the ITS1 of the pre-rRNA is red, and U3 snoRNA is depicted in light blue.

1.5 snR30 Release from the Pre-Ribosomal Particle

1.5.1 Required Factors for snR30 Release

Unlike canonical H/ACA snoRNPs that modify rRNA, the snR30 snoRNP requires RNA helicases to catalyze snR30 release from the pre-rRNA [92]. The dependence of snR30 on a helicase may result from the fact that the rRNA is bound in the inverse orientation to the internal bulge in the 3′ hairpin of snR30 compared to substrate RNA binding by canonical H/ACA snoRNAs. As a consequence of this inverse orientation, no pseudouridine can be formed by the snR30 snoRNP. In canonical H/ACA snoRNAs, pseudouridine formation triggers a conformational change in Gar1 which in turn catalyzes substrate turnover by altering the conformation of the thumb loop in Cbf5 which interacts with the target RNA [97]. Thus, the helicases may replace Gar1's function in facilitating dissociation of the snR30 snoRNP from pre-rRNA.

The DEAD-box helicase Rok1 is a critical player in snR30 release from the SSU processome [93], but the helicase Has1 as well as Utp23 are also essential for this function [92, 94]. Rather than directly removing snR30 from the ribosome, the helicase Rok1 is involved in a complex interplay with another critical ribosome assembly factor, the protein Rrp5, and together, Rok1, Has1, Utp23, and Rrp5 contribute to snR30 dissociation from the pre-40S ribosome. As for snR30, Rok1, Utp23, and Kri1, the assembly factor Rrp5 is recruited to the SSU processome during transcription of the central domain and acts as a compaction factor [98]. Unlike most other ribosome assembly factors, Rrp5 is important for the maturation of both the small subunit and the large subunit. Absence of Rrp5 prevents pre-rRNA cleavages at the A_0, A_1, and A_2 sites, like depletion of snR30, but also abolishes cleavage at the A_3 site by the ribonuclease MRP [99]. Rok1 directly binds to the A_2 site preventing early processing, and after domain 1 of the 25S rRNA is

transcribed, the protein complex Noc1/Noc2 rearranges Rrp5, thereby freeing the A_2 cleavage site. At this time, Rok1 catalyzes the release of Rrp5 from the SSU processome [100, 101]. Interestingly, the release of Rrp5 from the pre-40S ribosome by Rok1 is a prerequisite for snR30 dissociation. Thus, the most likely scenario is that Rok1 first induces conformational changes in Rrp5 or rRNA causing its release, and subsequently, the snR30 snoRNP is actively removed from the SSU processome with the help of the helicase Has1. The removal of snR30 may further be mediated by the direct interaction of Rrp5 with Has1 [101], but the molecular details of the interaction network of Rok1, Rrp5, Has1, and snR30 remain to be elucidated.

In addition to the helicases Rok1 and Has1, the PIN endonuclease Utp23 is also required for snR30 release [94]. Since snR30 is required for Utp23 binding in the first place, it is currently not clear how Utp23 is required for snR30's dissociation [94]. Possibly, Utp23 is not directly required for the snR30 release, but rather for a previous functional step of snR30 such as mediating pre-rRNA processing (vide supra). The cleavage of pre-rRNA may constitute another checkmark that is required to allow snR30 dissociation. Accordingly, the absence of Utp23 may cause snR30 to remain bound to the SSU processome by inhibiting the pre-rRNA processing checkmark.

1.5.2 Timing of the snR30 snoRNP Release

Our knowledge about snR30 is limited by the fact that this snoRNA co-transcriptionally associates and dissociates from pre-rRNA already before the SSU processome is fully assembled. The timing of protein and snoRNA association with pre-rRNA, including snR30, was elegantly determined by analyzing stalled pre-ribosomal complexes assembled on truncated pre-rRNA [42]. Interestingly, snR30 can only be detected in the pre-ribosomal particle when about half of the central domain of 18S rRNA is already transcribed, including 100 nucleotides upstream of the rm2 motif. As expected, at the same time, H/ACA proteins Cbf5, Nop10, Gar1, and Nhp2 can be

detected on the SSU processome, suggesting that snR30 is the H/ACA snoRNA that promotes most stable binding of the H/ACA proteins to the SSU processome. It is noteworthy that after transcription of snR30's main interaction sites within 18S rRNA, namely the rm1 and rm2 sites, additional rRNA needs to be synthesized before snR30 can be stably detected. Possibly, this additional rRNA stretch stabilizes the ES6 where snR30 binds. Furthermore, transcription of a potential base-pairing between snR30 and a region in 18S rRNA downstream of rm2 (C3, vide supra) [35], which also encompasses the position of Utp23 crosslinking to pre-rRNA, might further stabilize snR30 binding to the SSU processome [88].

The snR30 snoRNP remains bound to the SSU processome until the 3′ major domain of 18S rRNA is transcribed [42]. This is a surprising finding as snR30 thus seems to leave the SSU processome before the A_2 site in ITS1 is transcribed, although snR30 is crucial for facilitating A_2 cleavage [27]. Further mechanistic research is required to reconcile these observations. Interestingly, the C/D box U14 RNP dissociates at the same time as snR30, whereas the U3 snoRNP remains present, forming the stable 90S particle observed by cryo-EM. The release of snR30, U14, and fourteen other factors, including proteins interacting with snR30 such as Cbf5, Utp23, Kri1, and Nop6, could be triggered by a conformational rearrangement during transcription of ITS1 [84, 94, 102]. According to these findings, it is possible that in particular snR30 and U14 may form a snoRNA complex within the SSU processome and constitute a functional unit.

1.6 Open Questions

snR30/U17 is an unusual H/ACA guide RNA with a critical function during ribosome biogenesis that requires further mechanistic studies to be fully understood. With respect to ribonucleoprotein formation and synthesis, snR30/U17 behaves like a canonical H/ACA guide RNA. However, the essential function of snR30 differs greatly from the standard H/ACA

modification RNAs as its binding orientation on the pre-rRNA and its effect on the pre-ribosome are entirely different. Thus, one of the most interesting challenges in the ribosome synthesis research field remains to uncover the molecular mechanism of the snR30 snoRNP in promoting pre-rRNA processing.

Research on snR30 is lagging behind other essential snoRNAs like U3 for several reasons, leaving many unanswered questions about its function. As for other essential RNAs, the cellular role of snR30 can only be assessed through transient knock-downs, mutational, and deletion studies, which all contributed to identify the functional elements of snR30 [75, 85]. Additional crosslinking, pull-down, and mass spectrometry studies have helped to provide initial insight into the interactions of snR30 with the SSU processome [35, 42]. However, these experimental approaches provide only limited insight into the functional mechanism of snR30 from ribosome binding to inducing some change in the ribosome to dissociating in a controlled manner, and we are therefore still missing a detailed description of snR30's action on the SSU processome. Such understanding is particularly limited due to the transient interaction of snR30 with the maturing SSU processome as it has been reported to dissociate already before 18S rRNA transcription is complete [42]. This transient binding of snR30 with pre-rRNA has so far prevented us from obtaining high-resolution structural information on the interaction of this essential snoRNA with the SSU processome.

Two major hypotheses are proposed to explain the mechanistic role of snR30 within ribosome biogenesis. First, the protein recruitment hypothesis suggests that snR30 together with its associated proteins including Utp23 is essential to recruit and position additional factors on the SSU processome, ultimately helping Utp24 or another endonuclease to be correctly positioned for cleaving the pre-rRNA at sites A_0, A_1, and A_2 (Figure 1.5). However, open questions remain

regarding the details of the snR30-mediated interaction network and the timing of pre-rRNA processing. For example, we still do not know unambiguously the endonucleases responsible for A_0, A_1, and A_2 cleavage, and it remains unclear how the timing of snR30 dissociation and the cleavage at the A_2 site in the ITS1 are coordinated [37, 42]. Second, the rRNA folding hypothesis indicates that snR30 may act as an RNA chaperone in promoting conformational changes within pre-rRNA that ultimately are propagated to enable positioning of endonucleases at the A_0, A_1, and A_2 cleavage sites (Figure 1.5). Interestingly, in the cryo-EM structures of the SSU processome, expansion segment 6 is flexible and not folded into its mature conformation, such that it cannot be visualized [47]. If this complex represents a SSU processome structure after snR30 dissociation, then the function of snR30 could be to maintain the central domain of 18S rRNA in an unfolded state. Based on the secondary structure predictions of snR30 bound to 18S rRNA, it is equally conceivable that snR30 induces an alternative conformation in ES6. Both the un- or the re-folding of the 18S rRNA is likely a critical function of snR30, but in the absence of structural information, many questions regarding the impact of snR30 on SSU processome conformation remain.

In addition to our lack of understanding as to how snR30 promotes pre-rRNA processing, we also have limited knowledge on the timing, control, and mechanism of how the snR30 RNP is released from the SSU processome. Interestingly, snR30 requires two helicases, Has1 and Rok1, as well as Utp23 to dissociate from the SSU processome [94]. These three proteins may cooperate to signal a checkpoint that snR30 has completed its function before it dissociates. The dissociation of snR30 may be further coupled to progress in pre-rRNA transcription. In general, the finding that many factors are essential for the release of snR30 indicates that snR30 is responsible for an important step in ribosome maturation and that early dissociation of snR30 would abrogate small subunit formation. However, it remains to be uncovered how the SSU processome senses that

snR30's function is fulfilled such that it can be released and subsequent steps in ribosome biogenesis can occur.

In conclusion, snR30 is a highly interesting, still poorly understood snoRNA with critical function in ribosome biogenesis. This RNA, which has been hypothesized to be the ancestor to all H/ACA guide RNAs [103], is often overshadowed by the highly studied U3 RNA, but warrants further investigations to shed light not only on snR30, but also on critical steps during early ribosome synthesis. To answer the many open questions regarding snR30, we will need to apply a combination of traditional and innovative methods. Obviously, it will be most exciting to visualize snR30 bound to the SSU processome by cryo-EM. Significant challenges to purify a stable complex of an early ribosomal intermediate with snR30 will have to be overcome to reach this goal. Complementing such structural studies, now is the time to also utilize the power of in vitro studies to help elucidate the mysteries around snR30. Since the purification of active yeast H/ACA snoRNPs has been established, this route seems feasible to provide quantitative information on snR30's RNA and protein interactions, to generate kinetic information on its impact on pre-rRNA and to ultimately generate mechanistic information [63, 104]. Generally, solving the puzzle of snR30's function will only be possible when we develop a thorough understanding on rRNA folding and transient rRNA–snoRNA interactions during ribosome synthesis, which will require innovative experimental approaches. Based on the rapid and stunning progress in uncovering the mechanism of ribosome formation in the past years, we predict that the next years will yield interesting insights into the role of snR30 mediating pre-rRNA processing and folding.

1.7 Objectives

As discussed previously, the current knowledge about snR30 leaves many questions unanswered regarding its molecular interactions and mechanisms during ribosome biogenesis. Previously, the Kothe lab has published the *in vitro* reconstitution of the *S. cerevisiae* H/ACA snoRNP complex [63]. Building on this work, I reconstituted the snR30 RNP from purified components and analyzed the mechanisms of snR30 complex formation. Towards this goal, the function of the *in vitro* purified factors was first confirmed. Importantly, the PIN endonuclease Utp23 was purified and tested. To confirm that the purified Utp23 functions as previously published, the previously reported interaction between Nhp2 and Utp23 was tested via pulldown assay [88]. Following confirmation of the factors' purity and function, three objectives were addressed.

The first objective of this thesis was to determine the affinity of the guide RNA snR30 to the core H/ACA proteins Cbf5-Nop10-Gar1-Nhp2. This investigation revealed that snR30 behaves similarly to other H/ACA guide RNAs. Subsequently, the finding that Nhp2 is dispensable for binding of snR30 to the core H/ACA proteins was confirmed by separately testing snR30 binding by the core Cbf5-Nop10-Gar1 trimeric complex and Nhp2. Thereby, I confirmed the hypothesis that snR30, as an orphan H/ACA guide RNA, binds H/ACA proteins similarly to the modification guide RNAs. This was coupled to analyzing binding of snR30 with Utp23. The high affinity of this interaction provided evidence that the interaction of Utp23 with the snR30 RNP is predominantly protein-RNA mediated.

After reconstituting and characterizing the formation of the snR30 complex, the second objective was to investigate its interaction with the 18S rRNA. Previously, multiple sites of interaction between snR30 and the 18S rRNA have been reported [35]. By transcribing various

sections of the ES6 and surrounding helices of 18S rRNA, I was able to characterize 18S rRNA's interaction with the snR30 complex. This experiment fully utilizes the capabilities of the purified system to determine an affinity between the snR30 complex and the different sites of the 18S rRNA. Original studies on snR30 determined that only the C1 interaction site located in the basal half of snR30's 3′ hairpin possesses a physiological function; my work has shown that the C2 and C3 sites identified in Martin et al [35] are not contributing to the affinity of snR30 binding to the pre-rRNA, and instead snR30 is predominantly anchored on the pre-rRNA by the C1 site.

Lastly, the affinity of Utp23 to these same sites in the 18S rRNA was measured. Utp23 crosslinking data revealed interaction sites in helix 22, an area immediately adjacent to ES6 [88]. My experiments have revealed the requirements for Utp23 binding to the 18S rRNA, which include both the ES6H3 and H22 regions in pre-rRNA. In summary, my has revealed that Utp23's interaction with the rRNA is specific, and independent of other factors although, it cannot be ruled out that Utp23 binding is further enhanced upon snR30 binding.

In conclusion, I have reconstituted a functional snR30 complex *in vitro* and characterized its interaction network with ribosomal RNA and the assembly factor Utp23. In doing so, I have begun to clarify the questions how snR30 interacts with the ribosomal RNA. Furthermore, these data sets have allowed me to not only test known interaction sites but also elucidate any unknown interactions. Finally, the completion of these experiments has generated insight that the rRNA can be bound solely by the snR30 complex so long as the C1 site is present. Based on snR30's reliance on helicases to be removed from the pre-ribosomal particle, the interaction between snR30 and the rRNA is specific, tight, and independent of other factors.

CHAPTER 2: MATERIALS AND METHODS

2.1 RNA Design, Purification, and Labelling

DNA templates for *in vitro* transcriptions were generated using the Polymerase Chain Reaction (PCR) [105] method with purified *S. cerevisiae* genomic DNA (2.5 ng/µl) as a template. In these reactions, Pfu polymerase from Truin Science (0.02 U/µl) was used along with primers (Table 2.1) at a concentration of 0.5 µM and a dNTP concentration of 200 µM. Standard PCR protocol was used, denaturation at 95°C for 30 seconds, annealing at 54°C for 30 seconds and elongation at 72°C for 3 minutes. This cycle is repeated 30 times to achieve maximum DNA amplification. Product formation was confirmed by use of DNA-PolyAcrylamide Gel Electrophoresis (DNA-PAGE). Two different loci were targeted for amplification: snR30 and the entire 35S rDNA. The oligonucleotides (IDT) to amplify these regions are displayed in Table 2.1. These amplified sections of DNA were blunt-end ligated into SmaI-restricted pUC19 plasmids using T4 DNA Ligase (Thermo Fisher; 0.5 U/µl) and transformed into chemically competent DH5α *E. coli* cells from New England Biolabs. *E. coli* strains were stored in 30% glycerol at -80°C. To create RNA, the plasmids were first subjected to a round of amplification to create the DNA template. This PCR is identical to above except the extension time is only 30 seconds. A total of 12 DNA templates were created from the DNA oligonucleotides that are listed in Table 2.1; the pairs used were: snR30 T7 sense-snR30 antisense, snR30Δ292 T7 sense-snR30 antisense, snR30Δ397 T7 sense-snR30 antisense, H21 T7 sense-H22/23 antisense, H21 T7 sense-ES6H1 antisense, H21 T7 sense-ES6H3 antisense, ES6H1 T7 sense-ES6H3 antisense, ES6H2 T7 sense-ES6H3 antisense, Rm1-Rm2 T7 sense-Rm1-Rm2 antisense, ES6H2 T7 sense-H22/23 antisense, H22/23 T7 sense-H22/23 antisense, and H25/26 T7 sense-H25/26 antisense. Subsequently, the DNA template is purified using EZ10 spin columns from an EZ10 Spin Column Plasmid DNA kit

(Bio Basic). For templates created using a T7 sense primer that start with a 5′-T and not a 5′-GC double nucleotide, a second round of amplification was required. This PCR used a T7 only primer (5′- GCTAATACGACTCACTATAGGG-3′) to add the two 5′-GC nucleotides and had the same cycling conditions as previously except the annealing temperature was lowered to 50°C. The product of this reaction was again purified using the same EZ10 Spin Column Plasmid DNA kit (Bio Basic).

Table 2.1: List of all Oligonucleotides and their purpose.

Primer Name	Sequence (5′-3′)	Purpose
snR30 T7 sense	GCTAATACGACTCACTATAGGGAACCATAGTCTCGTGCTAGTTC GGTACTATACAGGG	Plasmid Creation Template Creation
snR30Δ292 T7 sense	GCTAATACGACTCACTATAGGGTGCTGTTGCCTTAACGATGTGT ATATGGGG	Template Creation
snR30Δ397 T7 sense	GCTAATACGACTCACTATAGGTAGGACGCATGATCTTGAGCTCT TTTCCTATACTTTG	Template Creation
snR30 antisense	mAmGATGTCTGCAGTATGGTTTTACCCAAATGATCATGGACC	Plasmid Creation Template Creation
35S rDNA sense	ATGCGAAAGCAGTTGAAGACAAGTTCG	Plasmid Creation
35S rDNA antisense	CAAATCCTTTCACGCTCGGGAAGC	Plasmid Creation
H21 T7 sense	TAATACGACTCACTATAGGGATTTTTTCGTGTACTGGATTTCCA ACGGG	Template Creation
ES6H1 T7 sense	TAATACGACTCACTATAGGGTTCTGGCTAACCTTGAGTCCTTG	Template Creation
ES6H2 T7 sense	TAATACGACTCACTATAGGGTTACTTTGAAAAAATTAGAGTGTT CAAAGCAGGCG	Template Creation
Rm1-Rm2 T7 sense	GCTAATACGACTCACTATAGGGCATGGAATAATAGAATAGGAC GTTTGGTTC	Template Creation
H22/23 T7 sense	TAATACGACTCACTATAGGGTTAATAGGGACGGTCGGGGG	Template Creation
ES6H1 antisense	TCCTGGTTCGCCAAGAGCC	Template Creation

Rm1-Rm2 antisense	CCTAGAAACCAACAAAATAGAACCAAACGTCCTATTCTATTATTCC	Template Creation
ES6H3 antisense	mTmCATTACGATGGTCCTAGAAACCAAC	Template Creation
H22/23 antisense	GAAAACGTCCTTGGCAAATGCTTTCG	Template Creation
H25/26 T7 sense	TAATACGACTCACTATAGGGCCGACTAGGGATCGGGTGG	Template Creation
H25/26 antisense	ACCCAAAGACTTTGATTTCTCGTAAGGTGC	Template Creation

In vitro transcriptions were carried out as described in Wright et al [106]. To summarize, the reaction was incubated at 37^0C for a variable amount of time (1 hour for test, 4 hours for large scale, overnight for radioactive *in vitro* transcriptions). The reaction buffer contained 40 mM Tris-HCl pH 7.5, 15 mM $MgCl_2$, 2 mM spermidine, and 10 mM NaCl. Additional components in the reaction were: 10 mM dithiothreitol (DTT), 3 mM nucleoside triphosphate (NTP) mixture containing adenosine triphosphate (ATP), cytidine triphosphate (CTP), guanosine triphosphate (GTP), and uridine triphosphate (UTP), 5 mM guanosine monophosphate (GMP), 0.01 U/µl inorganic pyrophosphatase (iPPase), 0.3 µM T7 RNA polymerase, and 0.12 U/µl RiboLock RNase inhibitor (Thermo Fisher). For radioactive *in vitro* transcriptions, [C5-^{3}H]UTP (Moravek) was combined with cold UTP to create a mixture at a final concentration of 0.1 mM or 1 mM in the *in vitro* transcription. Reactions with 0.1 mM [C5-^{3}H]UTP were attempted for all RNA, any that failed were repeated using 1 mM [C5-^{3}H]UTP. Following incubation, DNase (Thermo Fisher Scientific) was added to a final concentration of 0.002 U/µl, and the reaction was incubated for an additional hour at 37°C. The reaction was quenched by adding 3 M NaOAc to a final concentration of 0.3 M. Visualization of the product was carried out by Urea-PolyAcrylamide Gel

Electrophoresis (Urea-PAGE), staining with Ethidium Bromide (EtBr) and imaging on a UV transilluminator.

Non-radioactive *in vitro* transcription reactions were carried out in a total volume of 1 ml. These RNAs were purified using a Superdex 200 column (GE Healthcare) on a Biorad Duoflow chromatography system. Fractions corresponding to the RNA based on A_{260} absorbance were collected and precipitated. The precipitation was done by adding 1/3 vol 7.5 M LiCl, 50 mM EDTA, pH 6 plus 1/2 vol. isopropanol. The precipitation was incubated at -20^{o}C overnight before centrifugation at 5,000 x g for 30 minutes at 4^{o}C. The supernatant was discarded, and the RNA pellet was washed once with 70% ethanol. To prevent RNA loss, the pellet was again collected by centrifugation at 5,000 x g for 5 minutes. Following the wash, the RNA pellet was resuspended in RNase-free water. The concentration of the RNA was calculated using its A_{260} extinction coefficient. All coefficients were determined by the sum of the individual nucleotides' absorbance (http://www.fechem.uzh.ch/MT/links/ext.html).

2.1.1 Radioactive End-labelling

First, the 5′ phosphate of 200 pmol of snR30 (nonradioactive) was removed in a dephosphorylation reaction using Calf Intestinal Alkaline Phosphatase (New England Biolabs) at a concentration of 0.1 U/µl. Second, rephosphorylation was carried out using T4 polynucleotide kinase (0.5 U/µl) in the presence of [γ-^{32}P] ATP (10 µCi). The RNA was separated from excess nucleotide and enzymes by an RNA EZ10 clean-up kit (Bio Basic).

2.1.2 Purification of Radioactive In Vitro Transcriptions

All RNAs containing a section of the 18S rRNA were transcribed in the presence of [C5-^{3}H]UTP. These reactions were analyzed by Urea-PAGE for successful transcription, followed by

staining with SYBR Gold (Invitrogen) and imaging on an Amersham Typhoon 5 (GE Healthcare). Subsequently, the RNA was purified using Nucleobond Xtra Midi columns (Machery Nagel). To the RNA reaction, a mixture of buffer R0 (100 mM Tris/Acetate, 10 mM $MgCl_2$, 15% Ethanol (EtOH), pH 6.3) and R3 (R0, 1150 mM KCl) were added together to create a final concentration of 0.2 M KCl. This reaction was added to a column pre-equilibrated with buffer R0. The RNA was washed with 5 ml of buffer R1 (R0, 300 mM KCl) and then eluted in 2 ml of buffer R3. To this elution, 1/2 volume of isopropanol was added, and the mixture was allowed to precipitate at 4°C overnight. Following precipitation, the RNA was pelleted by centrifugation at 4°C and then washed once with 70% ethanol. The pellet was then dried and resuspended in pure water. Concentration of the RNA was determined by absorbance at 260 nm as described for nonradioactive RNA, and specific activity was calculated by scintillation counting. Size and details of all RNAs are presented in Table 2.2.

Table 2.2: Parameters of the snR30 and deletions thereof as well as 18S rRNA fragments

RNA	Length (nt)	Extinction coefficient (ε_{260}) μM^{-1} cm^{-1}*
snR30 full-length	612	6.788
snR30Δ292	317	3.551
snR30Δ397	213	2.364
H21-H22/23	302	3.427
H21-ES6H1	100	1.046
H21-ES6H3	199	2.237
ES6H1-ES6H3	162	1.850
ES6H2-H3	116	1.363
Rm1-Rm2	52	0.601
ES6H2-H22/23	219	2.554
H22/23	103	1.191
H25/26	72	0.801

*Coefficients calculated by use of: http://www.fechem.uzh.ch/MT/links/ext.html

2.2 Purification of Utp23

The pGEX5x-3 plasmid constructs containing the sequence of *S. cerevisiae* Utp32 codon-optimized for expression were synthesized and cloned by Genewiz. The plasmid was transformed into chemically competent BL21 (DE3) *E. coli* cells (New England Biolabs). Using these transformants, four colonies were chosen to screen for optimal protein production.

Utp23-expressing *E. coli* was grown in standard LB media containing 0.05 mg/ml Ampicillin. A total of eight 500 ml flasks were inoculated from two 50 ml precultures. These flasks were incubated at 37°C until an OD_{600} of 0.6-0.9. After inducing with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), the cells were transferred to 18°C and grown overnight. The cells were pelleted by centrifugation at 5,000 x g for 20 minutes. Collected cell pellets weighed between 5 and 10 g. Pellets were shock frozen in liquid nitrogen and stored at -80°C until protein purification.

Utp23 was purified based on the protocol by Wells et al [88] by resuspending the cell pellet in 5 ml of buffer NBA (20 mM Tris pH 7.6, 150 mM NaCl, 10% Glycerol, 0.2 mM DTT, 0.1% NP40, 5 mM $MgCl_2$, 0.5 mM phenylmethylsulfonyl fluoride (PMSF)) per 1 g of pellet. After thawing, lysozyme (Biobasic) was added to a final concentration of 1 mg/ml, and the suspension was allowed to incubate for half an hour on ice with slow stirring. Finally, 12.5 mg/g of cells of sodium deoxycholate was added to the mixture, and the cells were sonicated using a Branson Sonifier 450 in five rounds of 2 minutes at duty cycle 60%. Cell debris was separated by centrifugation at 30,000 x g for 45 minutes. The cleared lysate was then loaded on 2 ml of Glutathione Sepharose Fast Flow resin in a gravity flow column. The resin was prepared by washing with 20 column volumes (CV) of buffer NBA. The lysate was incubated with the resin for 30 minutes at 4°C. After binding, 20 CV of buffer NBA were used to wash the resin and remove any contaminants. Lastly, eight 1-ml elutions were collected using buffer NBA lacking PMSF and

containing 20 mM reduced glutathione. Elutions were analyzed on a 12% sodium dodecyl sulfate (SDS)-PAGE to determine which elutions to combine. The protein was then aliquoted into 50 or 100 µl fractions and shock frozen in liquid nitrogen. The protein was stored at −80^{0}C until use.

2.2.1 Determination of Utp23 Concentration

The concentration of Utp23 was determined by SDS-PAGE. A BSA standard curve was generated on an SDS-PAGE using masses between 10 µg and 200 µg. At least three different concentrations of Utp23 were analyzed alongside the standard curve. These samples were separated on an SDS-PAGE and stained with Coomasie brilliant blue. After scanning the gel, ImageJ software was used to quantify the intensity of the bands. Any protein point that was outside of the standard curve was discarded. Using the standard curve, the concentration of each protein sample was determined, and the average was calculated.

2.3 Nhp2-Utp23 Pull-down Assay

To determine if Nhp2 and Utp23 interact with each other, the proteins were previously purified with affinity tags fused onto their N-termini. His-Nhp2 binds specifically to Nickel Sepharose and GST-Utp23 binds specifically to Glutathione Sepharose resin. 10 pmols of each protein were incubated together at 30°C for 10 minutes in buffer NBA (with 0.1% v/v Tween 20 substituting for NP40). In a microcentrifuge tube, 100 µl of either Ni-Sepharose Fast Flow or Glutathione Sepharose Fast Flow (GE Healthcare) resin was prepared by three washes of buffer NBA each of 3 CV. The protein mixture was loaded onto the resin and an input sample was taken equal to 10% of the resin. Following three washes with buffer NBA, each 3 CV, another 10% of the resin was taken as the bound sample. When Ni-Sepharose Fast Flow resin was used, neither EDTA nor DTT were in the buffer due to their chelating effect on the nickel ions. The

input, final wash, and bound resin were all resolved on a 15% SDS-PAGE for imaging Nhp2 (12% for Utp23) followed by western blot against the protein that cannot bind to the resin. The SDS-PAGE containing samples of the protein was dry-blotted on a Trans-Blot Turbo (Biorad) for 20 min at 30 V. Transfer was judged based on visibility of the pre-stained ladder on the nitrocellulose membrane (Pall Corporation). The membrane was blocked by incubation in a 5% BSA TBS buffer (20 mM Tris pH 7.5, 150 mM NaCl) for one hour. Following blocking, the membrane was transferred into a 1:5,000 dilution of antibody containing 3% BSA in TBS and incubated overnight with gentle rocking at room temperature. The two antibodies were anti-His Horseradish Peroxidase (HRP; Genscript) and anti-GST HRP (Genscript) for imaging his-tagged Nhp2 and GST-tagged Utp23, respectively. Following incubation with the antibody, the membrane was washed twice for ten minutes each in TBS+TT buffer (TBS, 0.05% Triton-X100, 0.2% Tween20). To detect the antibody, the membrane was lastly transferred into a mixture containing 0.375 mg/ml luminol, 0.5 mg/ml p-coumaric acid, 0.1% H_2O_2 in 100 mM Tris-HCl, pH 8.5. After five minutes of incubating with gentle shaking at room temperature, the luminescence was imaged using an Amersham Imager 600 (GE Healthcare).

2.4 Nitrocellulose Filter Binding

2.4.1 Nitrocellulose Filter Binding of Cbf5-Gar1-Nop10-Nhp2 to RNA and Utp23 to RNA

To determine the affinity of a singular RNA to a protein or protein complex, the radiolabelled RNA is diluted in buffer RB (20 mM HEPES pH 7.4, 150 mM NaCl, 1.5 mM $MgCl_2$, 10% glycerol) and heated to 80°C for 1 minute before being allowed to cool to room temperature for 10 minutes in order to allow the RNA to fold. Afterwards, the RNA is incubated with either Cbf5-Gar-1-Nop10-Nhp2, Cbf5-Nop10-Gar1, Nhp2, or Utp23 at 30°C for 10 minutes in buffer RB over a minimum of 6 concentrations in the respective concentration ranges 0.2 nM

– 10 nM; 0.2 nM – 10 nM; 50 nM – 2000 nM; and 4 nM – 150 nM. The reaction solution is applied to a nitrocellulose membrane (GE Healthcare Life Sciences) and washed with 1 ml of ice-cold buffer RB. Following washing, the membrane is dissolved in 10 ml of scintillation cocktail (Ecolite (+), MP Biomedical). Lastly, 2 ml of the cocktail is added to a sample of unfiltered radioactive RNA to determine specific activity. The fraction of RNA that remained bound to the filter was determined by scintillation counting (Tri-Carb 2810TR) the filter and a control sample. The decays per minute (dpm) of the radioactivity retained on the membrane are divided by the dpm of the total reaction to yield the fraction of RNA bound which is converted into a percentage. Using GraphPad Prism, the percentage RNA binding is plotted versus the protein concentration and fitted with the following equation 2.1.

$$Y = B_{max} \times [S] / (K_D + [S]) \quad (2.1)$$

where Y is the percent bound, [S] is the concentration of RNA in the reaction, K_D is the dissociation constant, and B_{max} is the maximum value the fit reaches on the y-axis.

Using equation 2.1, the dissociation constant (K_D) for the interaction was determined. This was repeated in triplicate and two approaches were taken. First, the K_Ds were averaged, and the highest standard deviation was used which are shown in the appendix. Second, the data points were averaged to create a single average curve with its own K_D and standard deviation seen in chapter 3. This process was used to determine the affinities for all interactions involving only a single RNA.

2.4.2 snR30 Complex Nitrocellulose Filter Binding with 18S rRNA Fragments

First, the H/ACA protein complex is assembled on the snR30 RNA by adding the Cbf5-Nop10-Gar1 trimeric complex to the snR30 RNA and incubating it at 30°C for 5 minutes in

buffer RB. Following this, Nhp2 was added in a 1:1 ratio to the other H/ACA proteins, and the assembly reaction was further incubated at 30°C for 5 minutes. The final ratio of protein to RNA was 2:1 for full-length snR30 and 1:1 for snR30 truncations. This complex was then incubated at a concentration of 5 nM with an 18S rRNA fragment for 10 minutes at 30°C in buffer RB before filtration and scintillation counting as described above. Treatment and analysis of the data was identical to single RNA filter binding.

CHAPTER 3: RESULTS

3.1 Purification of RNA

3.1.1 Purification of snR30

To study how snR30 functions *in vitro*, purifying the RNA is a crucial first step. To gain insight into the structurally relevant portions of snR30, two truncations were conceived in addition to the full length snR30. The first truncation reduces the length of snR30 by 292 nucleotides, therefore, it is named snR30 Δ292. This truncation removes the majority of the 5′ hairpin but leaves the C2 interaction site intact, allowing for binding to the complementary site within helices 21 and ES6H1 in the 18S rRNA. Furthermore, the snR30 Δ292 starts at a sequence of three guanine nucleotides. This sequence motif aids in transcription start for the T7 polymerase and therefore does not require the addition of extra nucleotides to the RNA [107]. The second truncation deletes the entire 5′ hairpin and the H box leaving only the internal hairpin and 3′ hairpin. This truncation removes the C2 interaction site but leaves the C1 and C3 interactions intact. Since it removes the 5′ 397 nucleotides it is called snR30 Δ397. Together with the full-length snR30, these three variants of the snR30 RNA were generated. The locations of the truncations in the secondary structure are displayed in Figure 3.1 A.

In order to assess the efficiency of transcribing the very long snR30 RNA, I conducted small test *in vitro* transcriptions. These tests demonstrate that despite the length of the RNA, the *in vitro* transcription generates sufficient amounts of RNA for purification and most experimental assays. Figure 3.1 B shows a test transcription of full-length snR30, in which one can observe high levels of RNA production over a short time frame. Following test transcriptions, all three variants of snR30 were transcribed using large-scale reactions. After

confirming the RNA transcriptions were successful, the RNA was purified by size exclusion chromatography (SEC). As is evident in the chromatograms, each RNA (first peak) was well separated from the free nucleotides (second peak; Figure 3.1 C-E). Both snR30 and Δ292 snR30 eluted as one sharp peak. Δ397 snR30, however, had a shoulder indicating the presence of a larger specimen in the purification. Upon visualizing the shoulder and peak on a denaturing gel, both resolved at the same size (data not shown). As the shoulder contains RNA of the same size as the peak, the faster migrating shoulder is speculated to be mis-folded RNA. To finish the purification, the fractions corresponding to the RNA were collected and pooled. After precipitation and resuspension, the RNAs designated to be radiolabeled were dephosphorylated. Final confirmation of the RNAs' presence was assessed on a urea-PAGE following dephosphorylation (Figure 3.1 F). From the resultant bands, I can conclude that five of the six samples had RNA following the dephosphorylation. For the snR30 Δ292 sample 2, where no band is visible, the RNA was most likely lost during the procedure required to remove the proteins. Probably, the concentration of RNA was too low following phenol-chloroform extraction for successful precipitation to occur. Furthermore, at this stage, the smearing in the RNA samples is notably greater than it was following transcription (compare panels B and F in Figure 1). The increase in smearing is likely due to RNA degradation over time from freeze-thaw cycles or heating. Interestingly, this potential degradation appears to affect the full-length and snR30 Δ292 more than the snR30 Δ397. It may be that the shorter length of the snR30 Δ397 or a missing structural component helps protect from degradation.

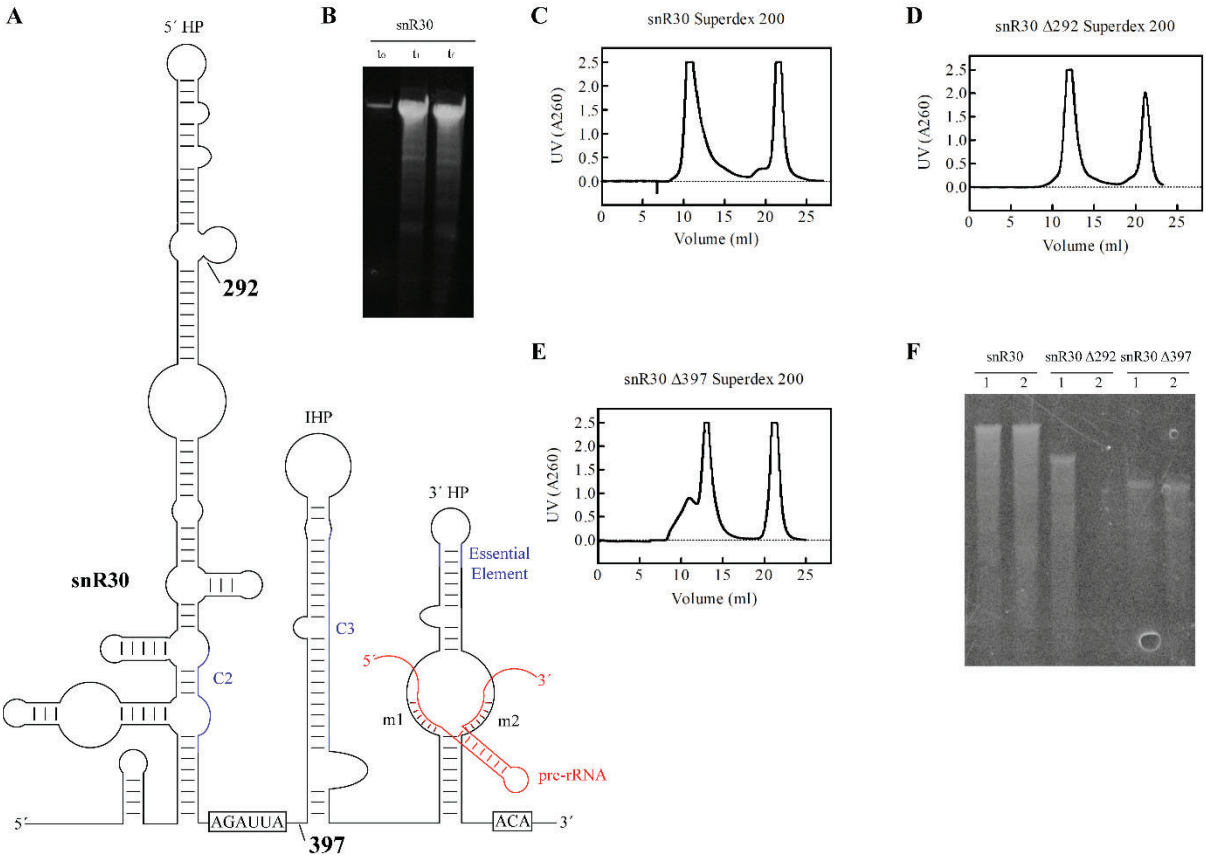

Figure 3.1. Purification of snR30 RNA and its truncations. (**A**) Secondary structure of full-length snR30. Location of truncations at nucleotides 292 and 397 are marked. Main binding sites m1 and m2 are shown interacting with 18S rRNA (red). The secondary sites of interaction at C2 and C3 are highlighted in blue as well as an essential element in the 5′ hairpin described in Fayet-Lebaron et al [85]. (**B**) Test *in vitro* transcription of full-length snR30. The following samples were analyzed: 0 minutes, t_0; 60 minutes, t_1; post 1-hour DNase, t_f. (**C**) Chromatogram of Superdex 200 purification of snR30 full-length. (**D**) Chromatogram of Superdex 200 purification of snR30 Δ292. (**E**) Chromatogram of Superdex 200 purification of snR30 Δ397. In all three purifications of the snR30 variants, the first peak is the RNA, the second is free NTPs. Absorbance at 260 nm is plotted against volume (ml). (**F**) Visualization of dephosphorylated snR30 RNA on an 8M 8% Urea-PAGE.

3.1.2 Purification of 18S rRNA Interaction Sites

Sections of the 18S rRNA that may contain binding sites for the snR30 complex were identified based on publications by Martin et al [35] and Wells et al [88]. These studies discovered novel locations in 18S rRNA of snR30 binding and Utp23 crosslinking sites, respectively. Many of these interactions between snR30 and 18S rRNA, including the primary

interaction sites rm1 and rm2, are located within expansion segment (ES) 6 and surrounding helices of the 18S rRNA (see Figure 3.2 A). Martin et al [35] identified three primary interaction sites and called them C1 (composed of rm1 and rm2 in ES6H2 and ES6H3), C2 (helices H21 and ES6H1), and C3 (helices H25 and H26). The Utp23 crosslinking site to the 18S rRNA is primarily located at the base of helix 22 [88]. Utilizing this knowledge, I separated this section of the 18S rRNA into five almost equal length sections, each comprised of a single helix (except helices 22/23). From these five sections, I created eight different RNA constructs shown in Figure 3.2 B. Lastly, to probe the interaction of C3 with the snR30 complex, I created a ninth RNA construct comprising helices 25 and 26 (Figure 3.2 C).

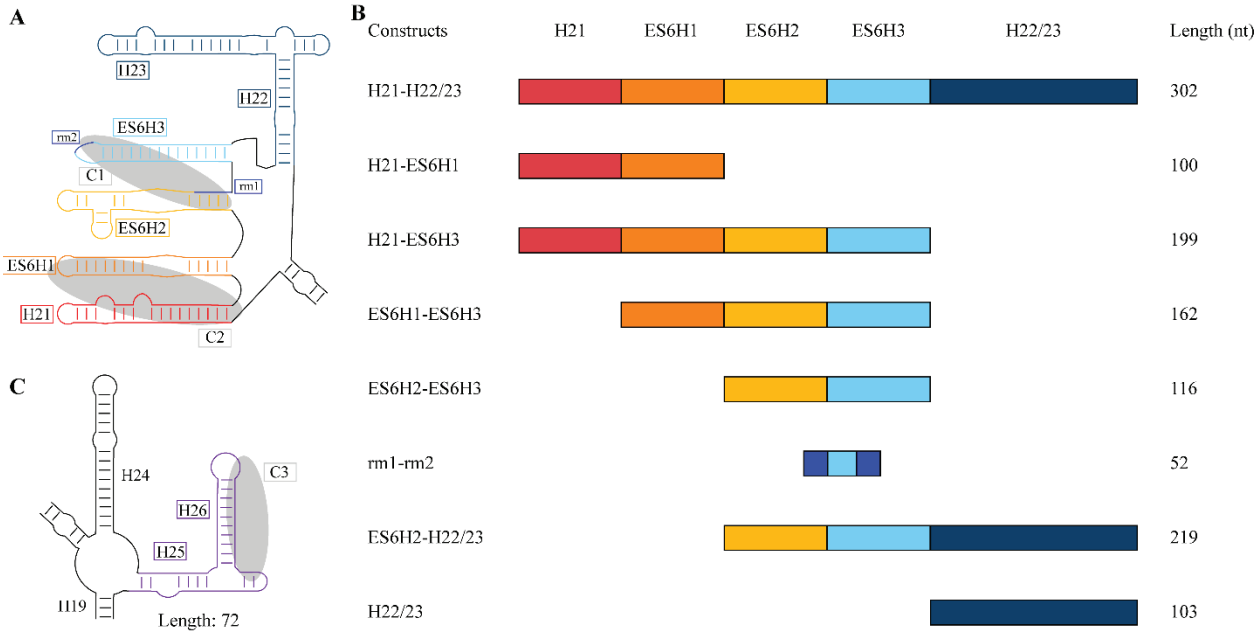

Figure 3.2. Design of RNA elements comprising different sections of 18S rRNA. (**A**) Secondary structure of ES6 and surrounding helices. The locations of cross-linking sites C1 and C2 are highlighted in gray. (**B**) Schematic representation of the different constructs derived from ES6. From left to right: name of construct, illustration of structure elements contained in construct, length of construct in nucleotides. Colors of helices in the schematic correlate to the colors in (A). (**C**) Secondary structure of helices 25 and 26 of 18S rRNA. The C3 site is highlighted in gray, the portion in purple represents the 72-nucleotide transcript used in this thesis.

The 18S rRNA fragments were purified differently than the snR30 guide RNA, as they are radiolabeled with a different isotope. These RNAs were transcribed in the presence of [^{3}H]UTP, an isotope with a longer half-life. This process allowed the generation of radiolabeled RNA in a single step, also possessing a suitable specific activity for experimental assays. Test transcriptions for the nine RNAs were completed on a small scale, and all produced RNA (data not shown). Following test *in vitro* transcriptions, the RNA was produced using a mixture containing both non-radioactive UTP and [^{3}H]UTP for a final concentration of 0.1 mM UTP. At this final concentration of UTP, well below the test conditions of 3 mM UTP, only some RNAs were successfully transcribed (Figure 3.3 A).

To determine the reason for failure of radioactive transcription, I quantified the prevalence of uridine in the transcribed regions. Of the 302 nucleotides composing the H21-H22/23 construct, 114 are uridines. Due to the high percentage of uridine in the transcript, I hypothesized that the low concentration of UTP in the radioactive transcription is insufficient to efficiently generate full-length transcripts. To confirm this, I performed test transcriptions at a variety of UTP concentrations (Figure 3.3 B) and observed reduced band intensity in the urea-PAGE analysis of the *in vitro* transcriptions at 0.5 mM UTP. This suggests that at UTP concentrations under 1 mM, the transcription of longer RNAs, that contain more uridines, is inefficient. Notably, these transcriptions are not failing, as there is no evidence of smearing that would result from short, abortive transcripts. Therefore, I repeated the radioactive transcriptions with a final concentration of 1 mM UTP to maximize efficiency of transcription while taking a reduced final specific activity of the radiolabeled RNA into account.

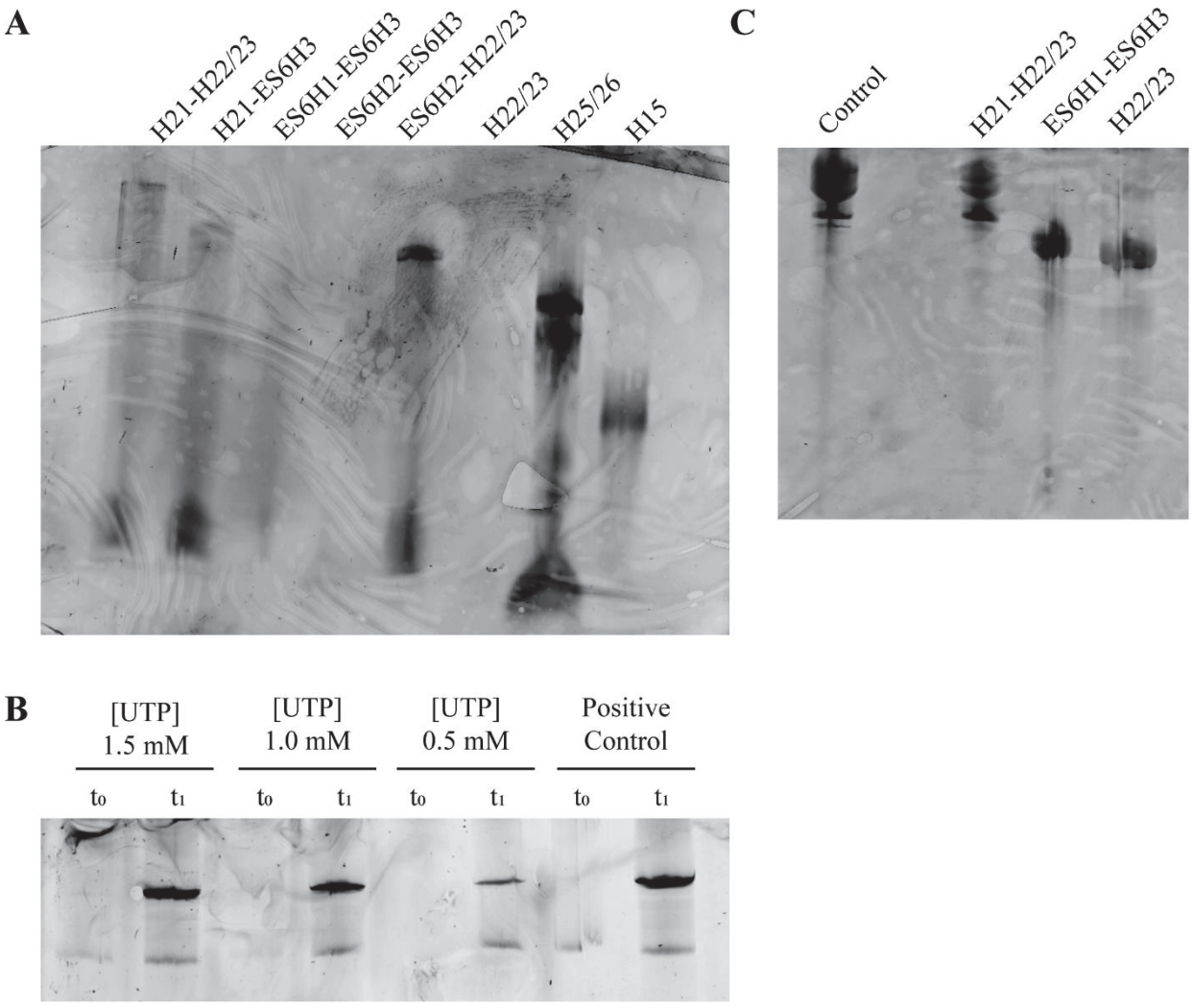

Figure 3.3. Representative radioactive *in vitro* transcriptions of 18S rRNA fragments. (**A**) Initial *in vitro* transcriptions with 0.3 mM UTP. (**B**) Test *in vitro* transcriptions with UTP concentration titration illustrating transcription levels of RNA at different concentrations. Positive control represents non-radioactive transcription conditions with 3 mM UTP. (**C**) Radioactive *in vitro* transcriptions using 1 mM UTP. Control lane contains non-radioactive H21-H22/23 RNA for size comparison.

The radioactive *in vitro* transcriptions were purified independently by anion exchange gravity flow chromatography. The final concentration and specific activity of the RNA was determined by spectrophotometer and scintillation counting.

3.2 Purification of Utp23

The method of purifying the *S. cerevisiae* nucleolar protein Utp23 was developed based on previously published methods [88]. Initial attempts to express the GST fusion tagged Utp23 in *E. coli* demonstrated only low levels of expression resulting in practically undetectable protein quantities in elution samples after affinity chromatography purification (data not shown). The expression levels were increased by growing *E. coli* overnight at 18°C after induction by IPTG. To further improve the purification, I added dithiothreitol (DTT) and phenylmethylsulfonylfluoride (PMSF) into the cell opening and wash buffers. Furthermore, the glycerol concentration was increased from 8.7% to 10% v/v.

Purification of Utp23 following this method yielded protein that was successfully purified using Glutathione Sepharose resin as evident from a band at the expected size in the elution samples (Figure 3.4 A). However, the elution samples displayed numerous other bands besides Utp23. These other bands indicate that Utp23 co-purifies with significant amounts of contaminants: quantification of elutions 5 and 8 generated an average purity of only 63 ±3%. Therefore, concentration determination had to be performed by comparative gel intensity quantification. Elutions 1-7 were used for SEC while Elution 8 (Figure 3.4A) was kept for experimental assays and its concentration was determined to be 3 ± 0.3 μM. Mass spectrophotometry experiments determined that the major contaminants were the protein chaperones DnaK and GroL (data not shown). These chaperones could not be resolved from Utp23 using size exclusion chromatography as only a single peak eluted (Figure 3.4 B).

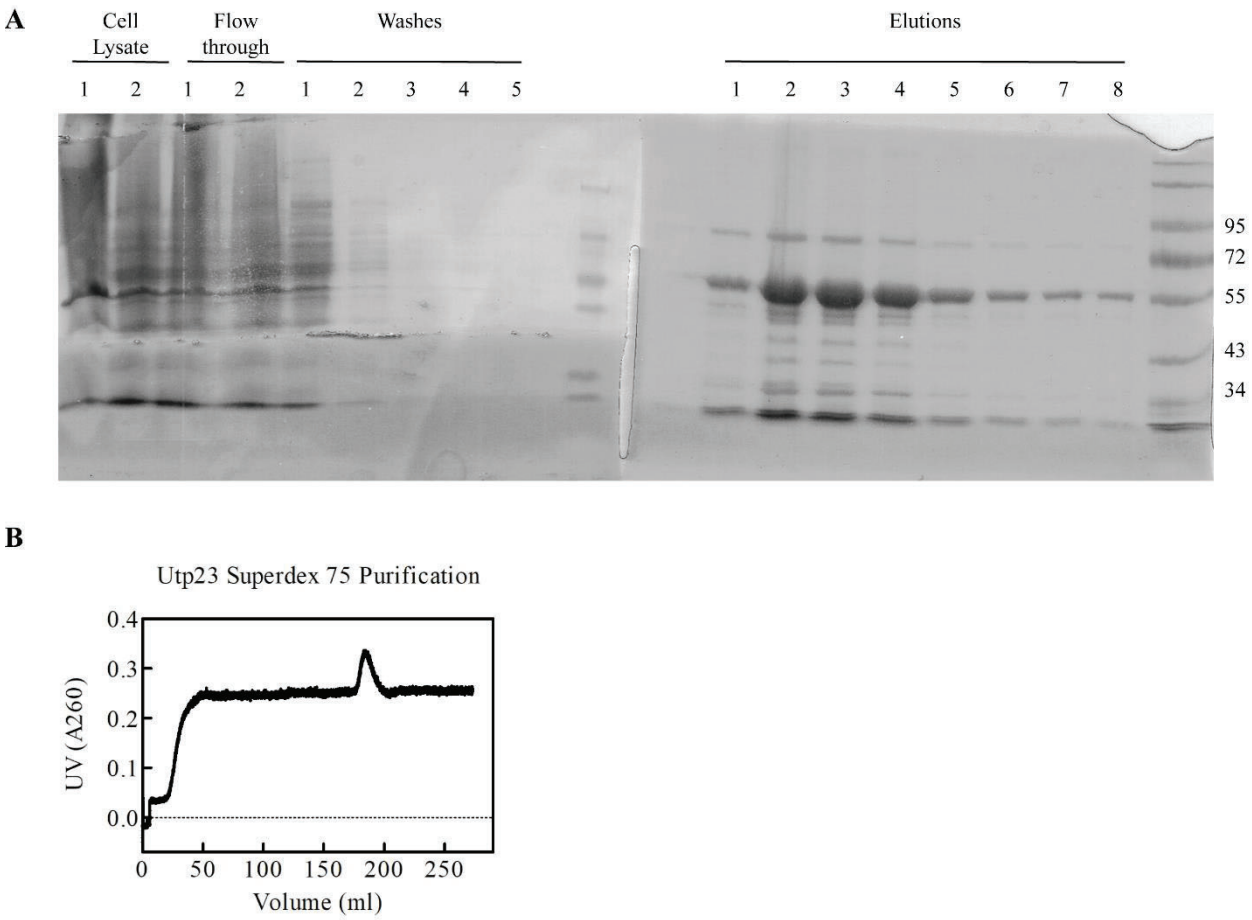

Figure 3.4. Purification of Utp23. (**A**) Glutathione-Sepharose purification of Utp23. The progress of the purification from cell opening to elution was analyzed by 10% SDS-PAGE. (**B**) Superdex 75 chromatogram of Utp23 purification.

3.3 Purified Nhp2 and Utp23 Interact in a Pulldown Assay

To confirm whether the purified protein components are functional, I tested for the interaction between Nhp2 and Utp23 that was previously reported [88]. First, a pulldown assay utilizing the GST-tag on the purified Utp23 was completed. Following incubation of GST-tagged Utp23 with hexahistidine-tagged Nhp2, the protein sample was incubated with the Glutathione Sepharose resin. Following the elutions from the resin, samples were analyzed by Western Blot against the His-tag on Nhp2 in order to determine if Nhp2 had been retained on the resin by

interacting with Utp23. The input and bound samples of glutathione resin, before and after the washes respectively, show increased levels of Nhp2 compared to the washes and elutions (Figure 3.5 A). As the post-bound lane (resin taken after the elutions) has an increased level of Nhp2 compared to the elutions, Nhp2 was retained on the resin through the elutions. The retention of Nhp2 indicates that potentially the concentration of reduced glutathione used is not sufficient for complete removal of GST-Utp23 from the resin. As oxidized glutathione will not bind to the GST-tag causing release from the resin, it is possible that too much glutathione became oxidized between buffer preparation and carrying out the assay. Alternatively, the Nhp2 could have precipitated leading to high retention on the resin. Nhp2, however, purifies at high concentrations indicating high solubility and produced no visible evidence of precipitation (white solid) during the assay. Furthermore, Nhp2 is retained on the resin after the washing steps, suggesting Nhp2 has a specific interaction to GST, Utp23, or the affinity resin. It is highly improbable that Nhp2 is facilitating its own interaction to an affinity resin without the requisite tag; therefore, I conclude that Nhp2 is likely being specifically bound by GST-Utp23. Unfortunately, this pulldown assay cannot explicitly rule out the possibility that the GST tag has an interaction surface for Nhp2 to bind. To rule out this possibility, another pulldown needs to be done with just GST in addition to the GST-Utp23 fusion protein.

To confirm that Nhp2 was being retained on the resin through an interaction with Utp23, a second pulldown assay was performed utilizing Ni-Sepharose resin to bind Nhp2 (Figure 3.5 B). The Western Blot was developed using an antibody specific for the GST tag fused to Utp23. In this assay, a large portion of the GST-Utp23 was found in the flowthrough and wash steps. This is in contrast to the pulldown assay with glutathione resin, which had relatively little Nhp2 in either sample. This difference could simply be due to a difference in protein amounts or minor

fluctuations in the surrounding environment reducing the affinity of the interaction. Nonetheless, the pulldown assay revealed high levels of GST-Utp23 present in the elutions. Comparing the wash 3 sample to elution 2, there is an apparent increase in the concentration of GST-Utp23, which is indicative of the protein being bound to the resin by a specific interaction. As Ni-Sepharose is an affinity resin highly specific to hexa-histidine tagged proteins, it is unlikely that the GST-Utp23 is binding the Ni-Sepharose on its own, but this could be tested independently in the future. Consequently, GST-Utp23 is likely forming an interaction with Nhp2 in order to facilitate its binding to the resin and eluting specifically in high concentrations of imidazole. Considering the pulldown assays together, I can conclude that GST-Utp23 is able to bind to Nhp2 in a manner that is not reliant upon the presence of other factors; however it remains to be confirmed whether this interaction is specifically mediated by Utp23 as hypothesized or whether the GST tag contributed to the pulldown.

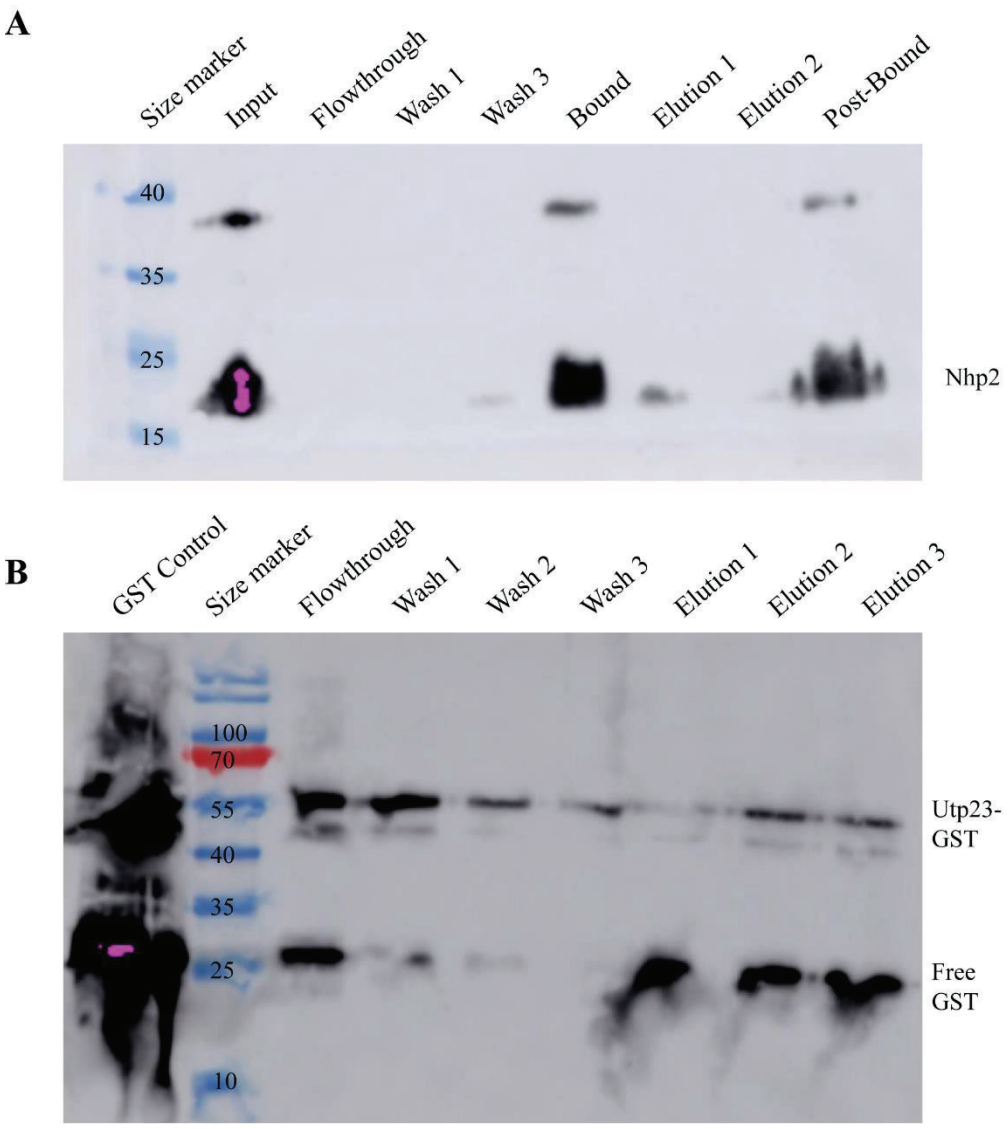

Figure 3.5. Pulldown assays to validate the interaction between Nhp2 and Utp23. (**A**) Pulldown assay using Glutathione Sepharose resin to capture GST-tagged Utp23. 10% of the resin was taken following protein incubation as the input sample. Following washing, another 10% was removed as the bound sample. After elutions, a final 10% of the resin was taken as the post-bound sample. Samples were analyzed by Western blot with anti-His antibody to visualize Nhp2. (**B**) Pulldown assay using Ni-Sepharose resin binding hexa-histidine-tagged Nhp2 followed by a Western blot with anti-GST antibody to image Utp23.

To summarize, both resins used in the pulldown assays indicate (Figure 3.5 A, B) that Nhp2 and GST-Utp23 are interacting with each other. These assays establish that the binding between Nhp2 and GST-Utp23 is reciprocal and independent of the resins. By confirming the

pulldown with two resins, I have demonstrated that Nhp2 and GST-Utp23 form a stable complex. Interestingly, an unknown band of ~38 kDa appears in the anti-His blots for imaging Nhp2 (Figure 3.5 A). Since this higher band in the Nhp2 blots shows up specifically with high concentrations of Nhp2, it is possible that the band is a dimer of Nhp2 that has been previously reported [108]. The sample is visualized by gel electrophoresis through an acrylamide matrix containing sodium-dodecyl-sulfate (SDS). SDS is a protein surfactant that linearizes the protein masking its charge and allows for separation of proteins based solely upon their size. It is possible however, that the amount of SDS added to the sample was inadequate to cover every protein therein, and therefore is insufficient to prevent the formation of Nhp2 dimers. Alternatively, Nhp2 has a single cysteine residue at position 94 that may form a disulfide bridge that is not reduced by the beta-mercaptoethanol in the reaction. Surprisingly, there is a high prevalence of free GST eluting with Nhp2. One possible explanation of this phenomenon is that the high ratio of free-GST to GST-Utp23 in the elutions results from GST's ability to dimerize [108]. While Utp23 is interacting with Nhp2, the GST tag on Utp23 may bind to another free GST monomer. In this scenario, the ratio of free GST to GST-Utp23 would be one to one; however, the amount of free GST exceeds the level of detected GST-Utp23 such that additional factors must contribute to the retention of free GST on the resin. Another factor that may have a large influence on the level of Utp23 detected in the sample is proteolysis. During sample preparation, the proteins are boiled at 95°C for 5 minutes. Throughout this process, it is possible that the heat induces cleavage in the linker between GST and Utp23, and the antibody only detects the GST and not the free Utp23. Therefore, it is possible that the prevalence of free GST in the sample is higher than expected based on stoichiometry. The heat-induced proteolysis could also be the reason why the GST-Utp23 often displays double banding in addition to minor

smearing. In summary, I have verified that the purified Nhp2 and GST-Utp23 proteins are capable of binding to one other. This interaction indicates that the purified proteins are functional and active, but further validation including free GST controls are required.

3.4 snR30's Affinity to the Core H/ACA Proteins and Utp23

3.4.1 The core H/ACA Protein Complex Cbf5-Gar1-Nop10-Nhp2 Interacts Tightly with snR30

Utilizing the trimeric complex of Cbf5-Gar1-Nop10 as well as Nhp2, both purified according to Caton et al. [63], I determined the affinity of snR30 to the core H/ACA protein complex via nitrocellulose filter binding. In accordance with the core H/ACA proteins' (Cbf5-Nop10-Gar1-Nhp2) interaction with other modification H/ACA guide RNAs, I expected an affinity around 1 nM [63]. Despite snR30 being an orphan H/ACA RNA that does not induce pseudouridylation, I anticipated that snR30 would form a similarly tight complex with the core H/ACA proteins since the snR30 RNP undergoes the same biogenesis pathway in yeast cells as modification H/ACA snoRNPs [1].

To verify my hypothesis, the affinity of each of the three variants of snR30 to the four-protein complex was determined in triplicate. The percent binding of the RNA was averaged across the three replicates and plotted against the protein concentration (Figure 3.6 A-C). The affinity of the interaction was determined by fitting to a hyperbolic function, and the dissociation constants are summarized in Table 3.1. As predicted, the binding of the snR30 and its truncations to the core H/ACA proteins is extremely tight. The weakest interaction observed is the binding of full-length snR30 to the Cbf5-Nop10-Gar1-Nhp2 complex with a K_D of 0.9 ± 0.2 nM, and the tightest is binding of snR30 Δ397 with a K_D of 0.4 ± 0.1 nM.

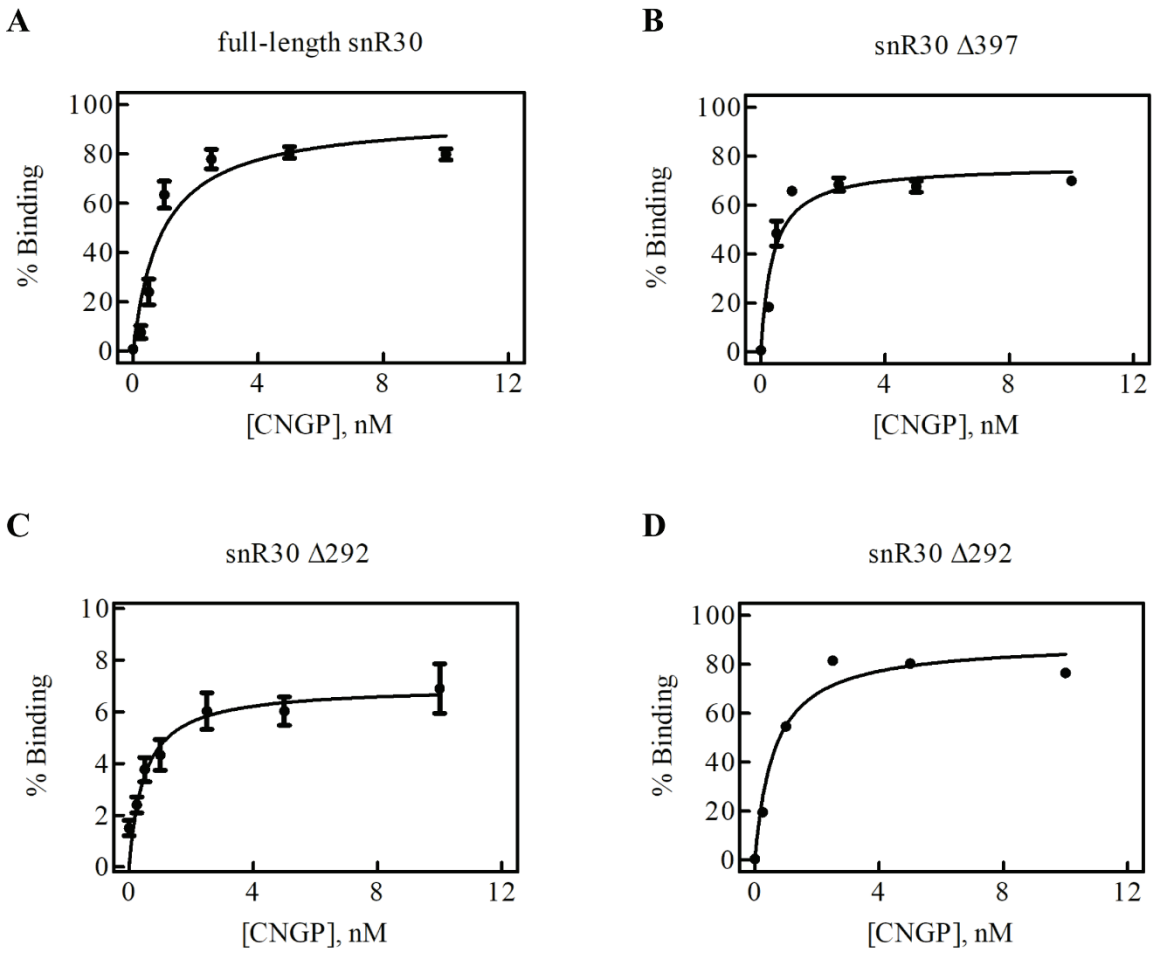

Figure 3.6. Average binding of the core H/ACA proteins Cbf5-Nop10-Gar1-Nhp2 with snR30. Percent binding of snR30 and its truncations to the Cbf5-Nop10-Gar1-Nhp2 (CNGP) core H/ACA protein complex. (**A-C**) Average binding of the respective RNAs, n=3. The smooth lines represent fits using a hyperbolic function (equation 2.1). (**D**) Singular experiment with snR30 Δ292. The dissociation constants are listed in Table 3.1.

It is worth noticing that despite the similar affinity of all three snR30 variants, snR30 Δ292 displayed a markedly reduced amplitude when binding to the H/ACA protein complex. Therefore, I repeated the experiment to characterize the binding of snR30 Δ292 with Cbf5-Nop10-Gar1-Nhp2 on a different day (Figure 3.6 D). Based on the increased amplitude reaching expected levels with no significant change in the affinity (0.5 ± 0.2 to 0.6 ± 0.2 nM), it was concluded that the low amplitude was an artifact observed on a particular day that may have

arisen through improper refolding of snR30 resulting in a conformation that is unable to bind the protein complex.

Table 3.1. Affinity of snR30 variants binding to the Cbf5-Nop10-Gar1-Nhp2 protein complex

snR30 Variant	Dissociation constant (nM)	Binding Amplitude (% Binding)
Full-length	0.9 ± 0.2	95 ± 6
Δ397	0.4 ± 0.1	76 ± 4
Δ292	0.5 ± 0.2	7 ± 0.5
Δ292*	0.6 ± 0.2	89 ± 7

*Corresponds to the singular trial in Figure 3.6 D. Dissociation constant is not included in the average for snR30 Δ292.

3.4.2 Contribution of Nhp2 to H/ACA Complex Binding

The archaeal Nhp2 homolog is L7Ae, an essential protein that specifically binds to a kink-turn motif; however, the eukaryotic H/ACA RNAs do not possess a kink-turn motif. Thus, it becomes interesting to consider two questions: does Nhp2 confer the sub-nanomolar affinity of the H/ACA proteins for snR30 despite the absence of a kink-turn and if not, is Nhp2 dispensable for binding *in vitro*? These questions were addressed for the *S. cerevisiae* H/ACA guide RNA snR34 in Caton et al [63]. When binding snR34 on its own, Nhp2 interacts with an affinity of 600 ± 200 nM while the Cbf5-Nop10-Gar1 trimeric complex has a K_D of 0.3 ± 0.1 nM for snR34. Therefore, for the modification guide RNA snR34, Nhp2 does not convey tight binding nor is it essential for protein binding *in vitro*. To demonstrate that snR30 behaves similarly to the modification guide RNAs, binding of the full-length snR30 and snR30 Δ397 to the trimeric complex of Cbf5-Nop10-Gar1 were tested in single replicates as a proof of concept (Figure 3.7 A, B).

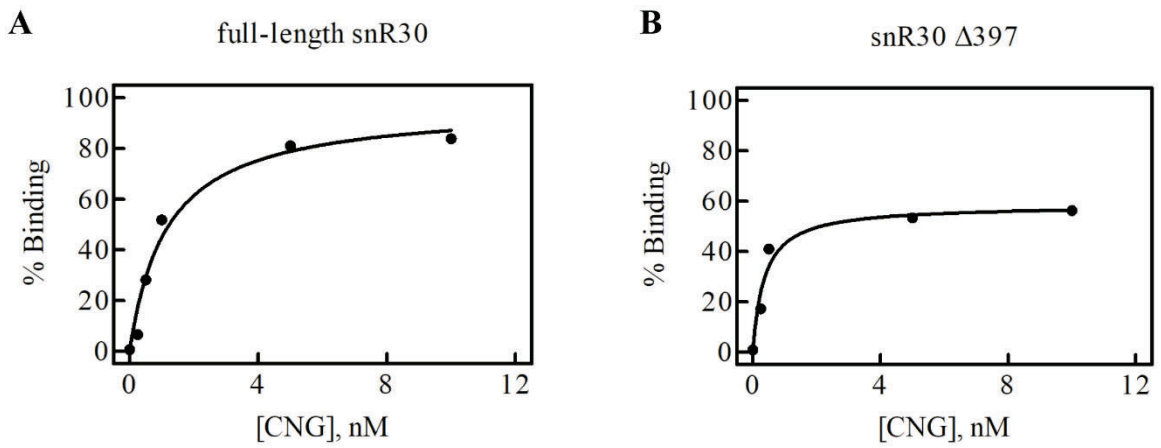

Figure 3.7. Nitrocellulose membrane filter binding of snR30 and snR30 Δ397 to the trimeric protein complex of Cbf5-Nop10-Gar1 (CNG). (**A, B**) Singular trial of the respective RNAs. The binding curves were fitted using equation 2.1. Dissociation constants are listed in Table 3.2.

The K_D of full-length snR30 binding to the trimeric protein complex of Cbf5-Nop10-Gar1 is 1.2 ± 0.3 nM (Table 3.2). This affinity is within the same range as the affinity of full-length snR30 to the entire core complex supplemented with Nhp2 (0.9 ± 0.2 nM). The same is true for snR30 Δ397, where the absence of Nhp2 has no measurable effect on the K_D (0.4 ± 0.1 nM; Tables 3.1 &3.2). Accordingly, I can conclude that the protein core of Cbf5-Nop10-Gar1 has extremely tight binding to snR30, even in the absence of Nhp2. This data corroborates the published findings that Nhp2 is dispensable for binding of the core H/ACA proteins to the guide RNA *in vitro* [63].

Table 3.2. Affinity between snR30 variants and the Cbf5-Nop10-Gar1 trimeric protein complex

snR30 Variant	Dissociation constant (nM)	Binding Amplitude (% Binding)
Full-length	1.2 ± 0.3	97 ± 8
Δ397	0.4 ± 0.1	58 ± 5

Having shown that Nhp2 is not required for tight binding of snR30 to the Cbf5-Nop10-Gar1 complex, it is still interesting to determine whether Nhp2 can bind to the snR30 RNA on its

own as Nhp2 was previously reported to interact with snR34 with a K_D of 600 ± 200 nM [63]. Comparing the K_Ds of snR30 for the trimeric protein complex of Cbf5-Nop10-Gar1 and Nhp2 (Fig. 3.8, Table 3.3), I observe that Nhp2's affinity for snR30 is more than 10-fold lower than the affinity of the trimeric complex for snR30. In conclusion, snR30, like other H/ACA guide RNAs, binds extremely tightly to the trimeric complex of Cbf5, Nop10, and Gar1. Nhp2 alone, meanwhile, interacts with a much lower affinity with snR30. Unlike the trimeric complex, however, Nhp2 on its own has a large change in affinity between the full-length snR30 and the snR30 Δ397 truncation. The resultant dissociation constants differ by almost an order of magnitude from 27 ± 3 nM for full length snR30 to 240 ± 55 nM for snR30 Δ397. As a member of the L7Ae class of proteins, Nhp2 retains the motif that allows for binding to kink-turn motifs [109]; however, there is no clear kink-turn in the first 397 nucleotides of snR30 that satisfy the requirements of three bulged nucleotides followed by G-A and A-G base pairs [110]. Therefore, the reason behind the change in affinity from the previously reported 600 ± 200 nM between Nhp2 and snR34, to 240 ± 55 nM for snR30 Δ397 and 27 ± 3 nM for the full-length snR30 remains unclear.

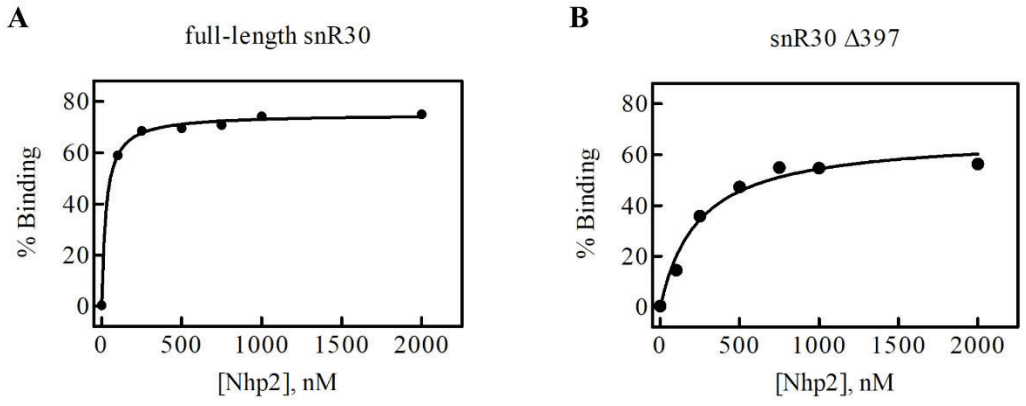

Figure 3.8. Nitrocellulose membrane filter binding of snR30 and snR30 Δ397 interacting with the core H/ACA protein Nhp2. (**A, B**) Singular trial of the respective RNA, fit using equation 2.1. Dissociation constants are listed in Table 3.3.

Table 3.3. Nitrocellulose filter binding parameters for the interaction of snR30 variants and Nhp2.

snR30 Variant	Dissociation constant (nM)	Binding Amplitude (% Binding)
Full-length	27 ± 3	75 ± 1
Δ397	240 ± 55	67 ± 4

In conclusion, the H/ACA proteins Cbf5, Nop10, Gar1, and Nhp2, which are ubiquitous partners to all H/ACA guide RNAs, can be divided into two groups: the Cfb5-Nop10-Gar1 co-purified complex and Nhp2. Distinguishing between these two groups allows for determination of the major contributor to the interaction affinity for snR30. It is clear from the difference in affinity between the proteins to the full-length snR30 (1.2 ± 0.3 nM, Cbf5-Nop10-Gar1; 27 ± 3 nM, Nhp2) that the trimeric complex provides the very high RNA affinity. By testing these interactions separately, I was able to determine that snR30 behaves very similarly to modification H/ACA guide RNAs with respect to binding the core H/ACA proteins.

3.4.3 snR30 Interacts Tightly with Utp23 in Absence of the H/ACA Proteins

During ribosome biogenesis, Utp23 is recruited to the pre-ribosomal particle by the snR30 complex [94]. Based on this and the crosslinking data from Wells et al [88], I hypothesized that Utp23 interacts tightly with snR30. Since all crosslinking sites identified between Utp23 and snR30 are within the 3′ 200 nucleotides that are included in all three variants of snR30, I proposed that all snR30 variants should bind with similar affinities to Utp23. Upon nitrocellulose filter binding of Utp23 interacting with the three variants of snR30 in triplicate, my initial hypothesis was confirmed (Figure 3.9 A-C). I observe tight binding with dissociation constants in the range of 5-9 nM that are not significantly different across the three snR30 variants (Table 3.4). Despite the consistency in binding affinities, there are distinct changes in

amplitude among the RNA variants. As the length of the RNA decreases from the full-length transcript to snR30 Δ397, the maximum amplitude of binding decreases. It is unclear if this effect is due to the length of the RNA or due to an unknown function of the 5′ hairpin of snR30, as the truncated versions are both missing the 5′ hairpin. Regardless, Utp23 is shown to bind snR30 with a tight affinity, confirming my hypothesis that the protein would interact with the RNA, even in the absence of any other cellular factors.

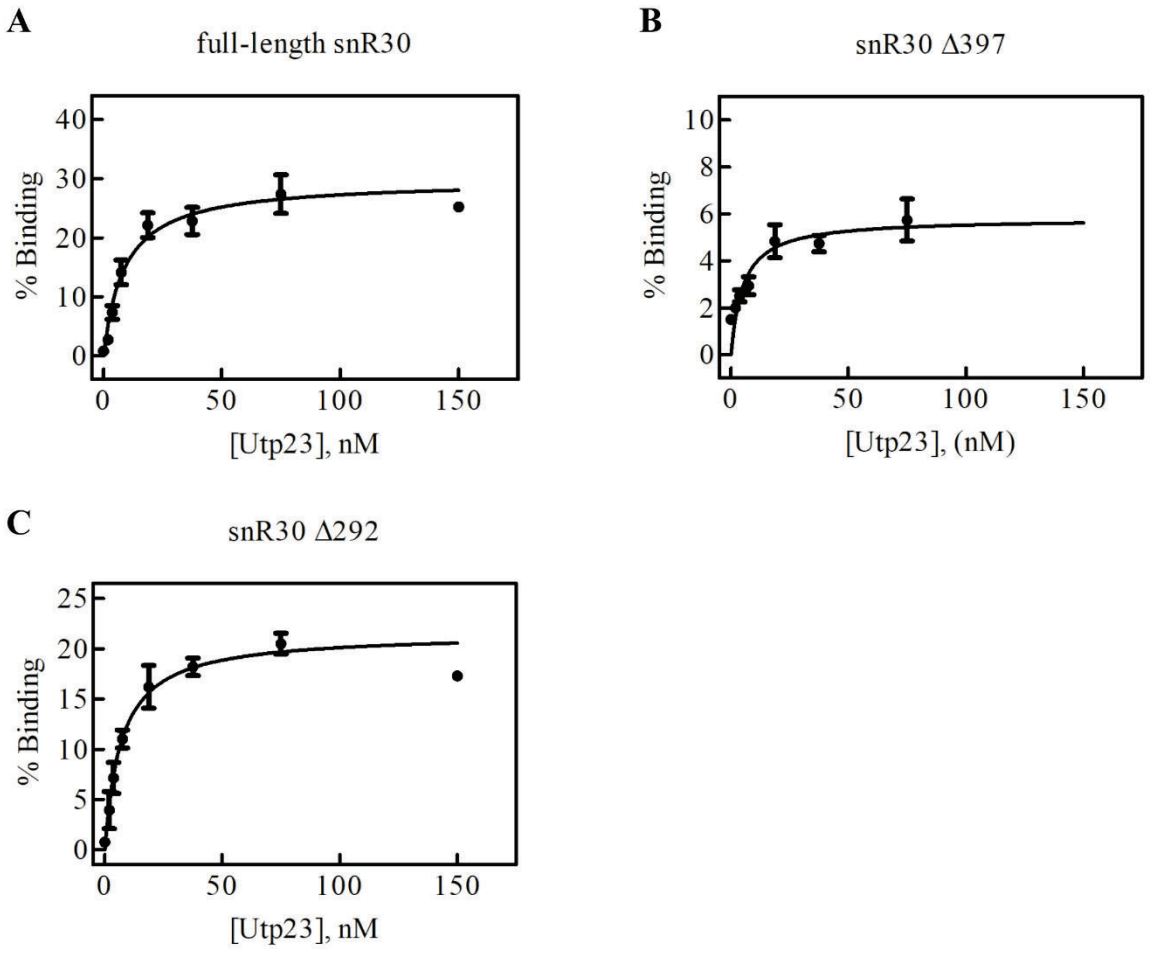

Figure 3.9. Binding of Utp23 to snR30 and its truncations. (**A-C**) Average binding of the respective RNA n=3, fit using equation 2.1. Dissociation constants are displayed in Table 3.4.

Despite the high affinity of Utp23 for snR30, it is worth noting that the protein does not bind the majority of the RNA present in the experiments. The highest amplitude reached is only

30% binding of the full-length snR30. This is in direct contrast to experiments of Cbf5-Nop10-Gar1-Nhp2 binding to snR30, in which I observe nearly 100% binding. Even Nhp2 alone, despite its lower affinity, reached an amplitude of 75% binding to the full-length snR30. This unexpected observation could be explained in two ways: either a large percentage of the guide RNA is in a state where Utp23 is unable to bind it, or the protein itself has a high dissociation rate such that a large portion of the bound RNA is removed during the wash step of the nitrocellulose filter binding experiment. Currently, it is impossible to distinguish between these explanations of the different, low binding amplitudes. The cause can be determined by determining the k_{off} rate of Utp23 dissociating from snR30, which can be done using chase filter binding, a variation of nitrocellulose filter binding. Either way, the consistent dissociation constants of Utp23 for the three snR30 truncations validate the crosslinking data and suggest that Utp23 has a specific and tight binding site within the internal hairpin of snR30.

Table 3.4. Dissociation constants for the binding between snR30 variants and Utp23

snR30 Variant	Dissociation constant (nM)	Binding Amplitude (% Binding)
Full-length	9 ± 2	30 ± 2
Δ397	5 ± 2	6 ± 0.5
Δ292	7 ± 1	22 ± 1

3.5 snR30 Complex Interaction with 18S rRNA

3.5.1 The snR30 Complex Binds Tightly to the C1 Site

Initial studies on snR30 by Atzorn et al [75] proved that Watson-Crick base pairing between the m1 and m2 motifs on snR30 and the rm1 and rm2 motifs on the 18S rRNA are essential for snR30's function. Subsequent work on snR30 discovered the interaction with 18S

rRNA is not limited to the m1 and m2 motifs; besides the C1 site (previously referred to as m1 and m2), two more interaction sites called C2 and C3 were described [35]. Of the three interaction sites, I expected the C1 site to have the highest affinity based on its location within a single-stranded loop of the 5′ hairpin of snR30, as well as its physiological relevance. The m1 and m2 sequences on snR30 have been shown to be essential for snR30 function, whereas the locations in snR30 where the C2 and C3 sites of 18S rRNA are proposed to bind are not essential [75]. Taken together, I postulated that C1 is the primary interaction site whereas the C2 and C3 sites are only complementary. Therefore, the affinity of the snR30 RNP complex for the C2 and C3 sites may be lower than for the C1 site, where the primary interaction should be occurring.

The snR30 RNP binds the C2 site contained within the H21-ES6H1 construct with a dissociation constant of 405 ± 191 nM compared to the dissociation constant for the rm1-rm2 RNA's of 106 ± 51 nM (Table 3.5; Figure 3.10 B, E). No conclusion can be drawn for the interaction between H25/26 (the C3 site) and the snR30 RNP for which the dissociation constant is 225 ± 227 nM as all three trials behaved differently (Figure 3.10 H; Figure A.1 H). Every construct comprising the C1 motif (rm1-rm2) binds to the snR30 RNP with an affinity equal to or higher than observed for the rm1-rm2 construct (H21-H22/23, 84 ± 16 nM; ES6H1-ES6H3, 54 ± 19 nM; ES6H2-ES6H3, 77 ± 21 nM; ES6H2-H22/23, 19 ± 10 nM; Figure 3.10 A,C,D,F). These specific and tight interactions are contrasted by the relatively weak binding the snR30 RNP has for the negative control, mitochondrial $tRNA^{Met}$ (746 ± 405 nM; Figure 3.10 I). Based on the different subcellular localization, there cannot be any *in vivo* interactions between the nucleolar snR30 RNP and the mitochondrial $tRNA^{Met}$. Therefore, I can conclude that the snR30 RNP has an unspecific interaction with RNA with a dissociation constant around 700 nM. Furthermore, I observe that every single construct that does not contain the C1 interaction site

located in helices ES6H2 and ES6H3 has an affinity more than two times lower than the affinity of the short rm1-rm2 RNA (minimal C1 site) for the snR30 RNP. Therefore, the experimental data generated with the full-length snR30 complex supports my hypothesis that the main region required for proper function of snR30 *in vivo* is the C1 motif.

Unexpectedly, the amplitudes of the binding curves vary widely (Table 3.5). Only two constructs, rm1-rm2 (48 ± 10 %) and ES6H1-ES6H3 (70 ± 7 %), have an amplitude less than 100%. Notably, the amplitudes of binding reach unexpectedly high levels, e.g., 3150 ± 950 % with the mitochondrial $tRNA^{Met}$. Since the radiolabeled RNA is being titrated in these experiments, 100% binding indicates that every snR30 RNP complex has bound one radiolabeled RNA molecule. Theoretically, it is possible for a single molecule of the snR30 RNP complex to interact with more than one RNA molecule, reaching levels beyond 100% binding. Due to the presence of two sets of H/ACA proteins per guide RNA, any direct protein-RNA interaction, that is not mediated by specific snR30 guide RNA base-pairing, can exceed 100%. When accounting for the potential of protein-RNA interactions to increase the amplitude above 100%, a trend arises between the binding affinity and amplitudes. Only RNA with a dissociation constant above 200 nM exhibit an amplitude beyond 200% binding (Table 3.5). Therefore, the data suggests that a lower amplitude indicates a specific mode of interaction that is dependent upon the guide RNA. This mode of interaction can be distinguished with the help of the negative control RNA, mitochondrial $tRNA^{Met}$, which exemplifies how high the amplitude can be during unspecific binding of the RNA to the snR30 RNP complex. The experiments with mitochondrial $tRNA^{Met}$ establish a baseline that can be used to quantify which RNAs interact with the full-length snR30 complex in an unspecific manner. Thus, I conclude from the data for the different 18S rRNA fragments, that any RNA element with a dissociation constant above 200 nM and an amplitude

above 400% is binding in an unspecific manner to the full-length snR30 RNP. Using these criteria, H21-ES6H1 (C2 site), H22/23, and H25/26 (C3 site) bind non-specifically, whereas H21-H22/23, ES6H1-ES6H3, ES6H2-ES6H3, rm1-rm2, and ES6H2-H22/23 bind in a specific manner to the snR30 RNP. As all the RNA interacting in a specific manner contain the C1 motif, it is likely that they bind to the snR30 RNP through Watson-Crick base-pairing with the snR30 RNA whereas the unspecific RNA interactions may be mediated by the H/ACA proteins.

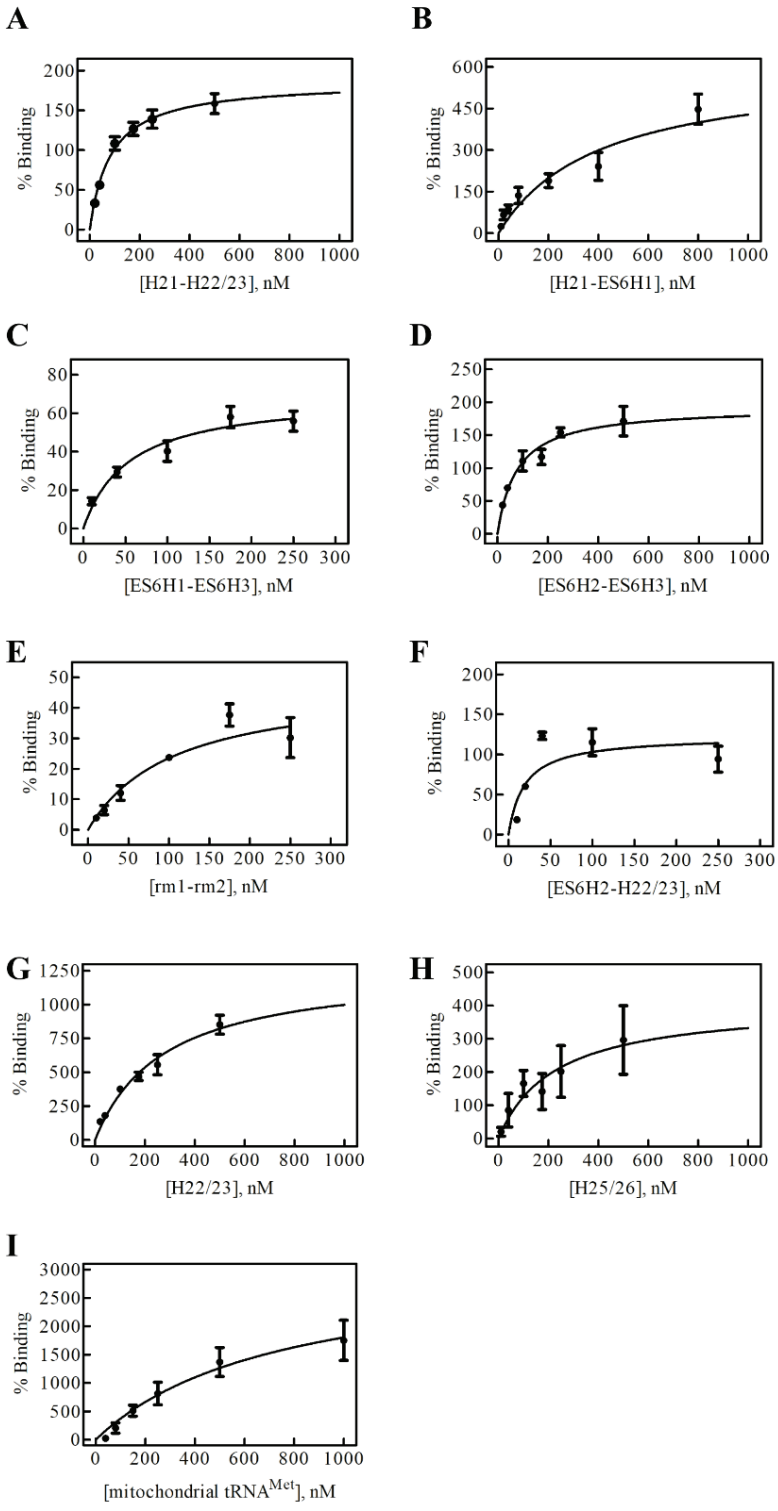

Figure 3.10. Full-length snR30 complexed with Cbf5-Nop10-Gar1-Nhp2 Nitrocellulose Filter Binding to 18S rRNA fragments. (**A-I**) Binding of the respective RNA, n=3 fit using equation 2.1. Dissociation constants are displayed in Table 3.5. Individual trials shown in Figure A.1.

In conclusion, the nitrocellulose filter binding experiments support the hypothesis that snR30 is primarily anchored to the 18S rRNA via the essential motifs m1 and m2. Despite the more recent work discovering additional interaction sites of snR30 in pre-rRNA, neither of them appears to strongly contribute to the affinity of snR30 binding to 18S rRNA. Since the H21-H22/23 construct has a remarkably similar affinity for snR30 as the rm1-rm2 construct, the addition of extra structural elements within the longer 18S rRNA fragment may not significantly enhance binding to the snR30 RNP.

Table 3.5. Interaction affinity between 18S rRNA constructs and the snR30 complex.

snR30 Variant	18S rRNA Construct	Dissociation constant (nM)	Binding Amplitude (% Binding)
Full-length	H21-H22/23	84 ± 16	187 ± 11
	H21-ES6H1	405 ± 191	600 ± 140
	ES6H1-ES6H3	54 ± 19	70 ± 7
	ES6H2-ES6H3	77 ± 21	193 ± 16
	rm1-rm2	106 ± 51	48 ± 10
	ES6H2-H22/23	19 ± 10	123 ± 18
	H22/23	276 ± 72	1280 ± 170
	H25/26	225 ± 227	410 ± 190
	Mitochondrial $tRNA^{Met}$	746 ± 405	3150 ± 950
Δ397	H21-H22/23	40 ± 15	102 ± 10
	H21-ES6H1	90 ± 27	154 ± 21
	ES6H1-ES6H3	ND	ND
	ES6H2-ES6H3	32 ± 18	57 ± 8
	rm1-rm2	50 ± 11	24 ± 2
	ES6H2-H22/23	ND	ND
	H22/23	84 ± 27	430 ± 46
	H25/26	92 ± 39	640 ± 93
	Mitochondrial $tRNA^{Met}$	NC	NC

ND means the affinity was not determined for the RNA. NC means the affinity could not be calculated.

I investigated the interactions between the same 18S rRNA constructs, that were used for full-length snR30, with the snR30 Δ397 truncation. As aforementioned, this snR30 truncation lacks the 5′ leader and the entire 5′ hairpin, including the H box. The deletion removes the predicted C2 interaction site within the 5′ hairpin of snR30 that should interact with helices 21 and ES6H1 of the 18S rRNA. As such, the snR30 Δ397 RNP should not be able to bind to the H21-ES6H1 construct that relies upon this interaction. Disagreeing with this, the low affinity of the full-length snR30 to H21-ES6H1 of 405 ± 191 nM (Table 3.5; Figure 3.11 B) increases to 90 ± 27 nM for snR30 Δ397 (Table 3.5; Figure 3.11 B). Moreover, the H21-ES61 construct has an amplitude of 154 ± 21 % suggesting a specific mode of interaction with the snR30 Δ397 RNP (Figure 3.11 B). All three 18S rRNA constructs, which contain the sequence from rm1 to rm2, display nearly identical dissociation constants of 40 ± 15, 32 ± 18, and 50 ± 11 nM for H21-H22/23, ES6H2-ES6H3, and rm1-rm2 for binding to Δ397 snR30 , respectively (Figure 3.11 A,C,D).

When comparing the affinities of snR30 Δ397 to the full-length snR30, one of the primary differences is a potential change in affinity for H25/26. This construct composing the C3 interaction site has an affinity of 225 ± 227 nM to the full-length snR30 RNP. With such a high error in the measurement, the main metric that suggests unspecific binding of H/25/26 to the full-length snR30 RNP is the high amplitude. However, the binding of the H25/26 construct to the snR30 Δ397 complex is characterized by an affinity of 92 ± 39 nM suggesting a specific interaction (Figure 3.11 F). Despite the confirmation of tight binding between the H25/26 construct and the snR30 Δ397 complex, the resultant amplitude was again high (410 ± 190 %, full-length; 640 ± 93 %, Δ397). The observed tight binding of H25/26 to snR30 with a high

amplitude calls into question whether indeed all interactions with amplitudes above 400% should be classified as unspecific binding.

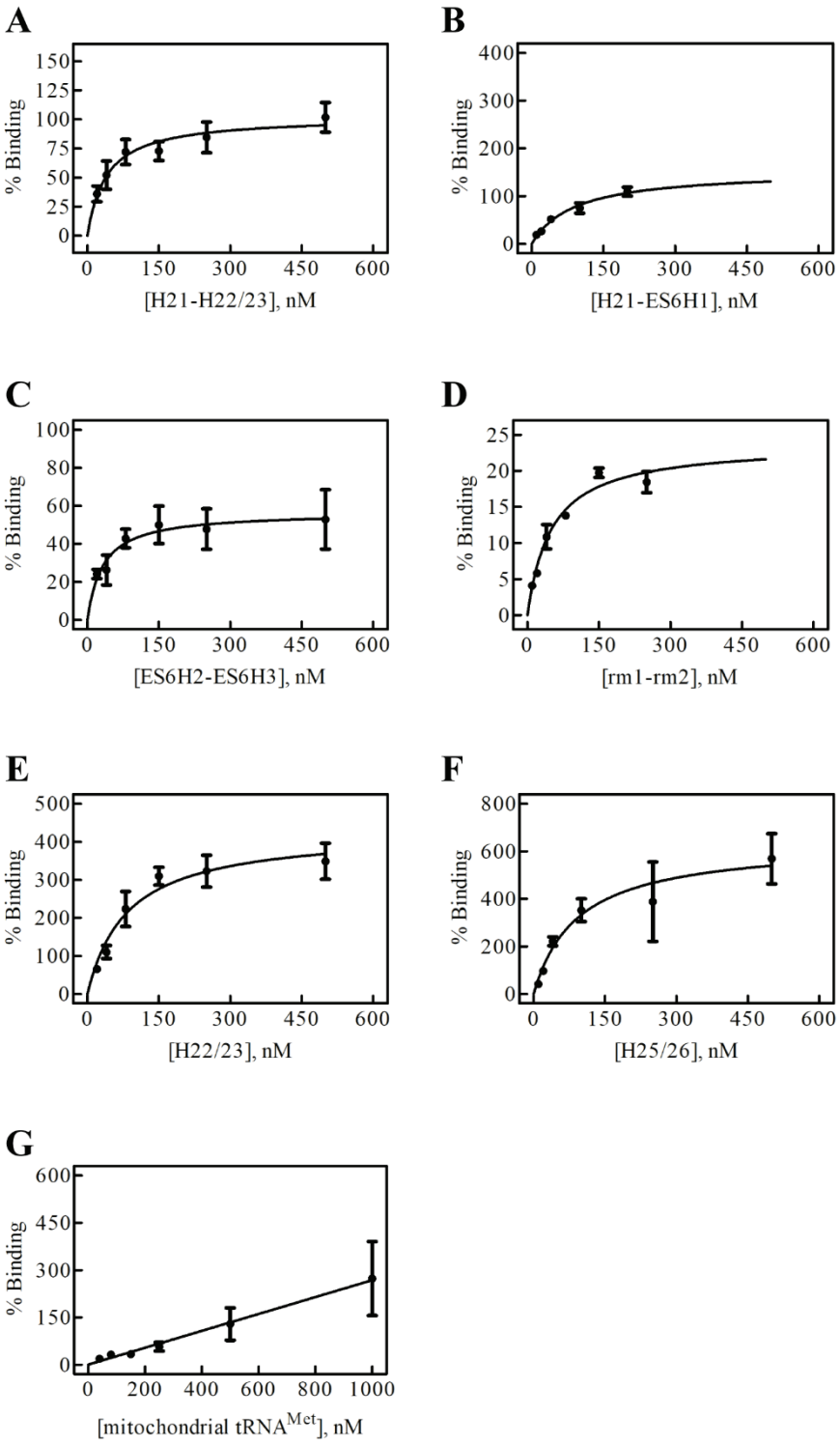

Figure 3.11. Nitrocellulose filter binding of 18S rRNA constructs interacting with snR30 Δ397 complexed with Cbf5-Nop10-Gar1-Nhp2. (**A-G**) Binding of the respective RNA, n=3 fit using equation 2.1. Dissociation constants are displayed in Table 3.5. Individual trials shown in Figure A.2.

The final notable change between the full-length snR30 and the snR30 Δ397 complexes is the difference in binding to H22/23. The full-length snR30 has a K_D of 276 ± 72 nM to H22/23 compared to snR30 Δ397 which has a K_D of 84 ± 27 nM. The two-fold change in affinity between the full-length and the snR30 Δ397 RNP when binding the H22/23 construct suggests a change from unspecific to specific binding. However, the amplitude exhibits little change from the full-length (1280 ± 170%, Table 3.5) to the snR30 Δ397 RNP (430 ± 46%, Table 3.5). The H22/23 amplitude is on the cut-off determined for the full-length snR30 RNP of 400% as one of the determinants for unspecific binding. Such a high amplitude indicates unspecific protein-RNA interactions between the core H/ACA proteins to the 18S rRNA fragment that are not mediated by the snR30 guide. Therefore, the mechanism of binding between the Δ397 snR30 RNP and the H22/23 construct is still undetermined. Additional experiments can hopefully quantify the change in affinity between snR30 variants and investigate whether the interaction is guide RNA mediated.

It is worth mentioning that the difference in binding at C2, C3, and H22 between the full-length snR30 RNP and the snR30 Δ397 RNP, which form only weak interaction to the full-length snR30 complex, may be an indirect result of the ratio of guide RNA to protein. In the assays with full-length snR30, the ratio of protein to RNA is 2:1, whereas with Δ397 snR30 the protein:RNA ratio is 1:1 to adjust to the removal of the H box. Therefore, even though the concentration of the complex within the reaction is the same, the total concentration of protein is half for experiments with snR30 Δ397 compared to reactions with the full-length snR30. This reduction of total protein may influence the stoichiometry of RNA in the reaction providing a larger binding surface for RNA-RNA interactions. It remains unclear whether this difference between the full-length and snR30 Δ397 complexes during filter binding changes the affinity to

RNA that is suspected to be binding in an unspecific manner like H22/23. Either way, more work is necessary to make conclusive deductions regarding the formation and binding of the snR30 RNP to the 18S rRNA fragments.

In conclusion, while the data presented here support the importance of the interaction of snR30 with the C1 site, the mechanism through which the snR30 RNP may interact with C2, C3, and H22 in the 18S rRNA remains unclear. The amplitudes exhibited by the H22/23 and H25/26 constructs are reminiscent of unspecific binding (430 ± 46 %, 640 ± 93 %; Table 3.5) when binding to the snR30 Δ397 RNP. Contrary to this high amplitude, the binding curves for H22/23 and H25/26 are reproducible and show a tight interaction that is not present in the full-length complex. Further data is required to clarify whether the high amplitude indicates an interaction mediated through protein-RNA binding or if the snR30 Δ397 RNP can make RNA-RNA interactions to H22/23 and H25/26.

3.5.2 Utp23 Forms a Tight, Specific Interaction to ES6H2-H22/23

The PIN (PilT N-terminus) endonuclease Utp23 is essential in *S. cerevisiae* for formation of the SSU processome, and by extension, the 40S small subunit. The Schneider lab discovered Utp23's interaction sites with the 18S rRNA, which are primarily located in helix 22 [88]. To build on this, I investigated Utp23's affinity to the various regions of ES6 of the 18S rRNA and surrounding regions by using nitrocellulose filter binding (Figure 3.12). Most surprising is my observation of how Utp23 interacted with H21-H22/23 and H22/23. Contrary to one-site binding curves where the percentage of the RNA bound increases in a hyperbolic manner, the binding peaks between 10-20 nM Utp23 and subsequently decreases at higher Utp23 concentrations preventing the determination of a K_D (Table 3.6). Interestingly, ES6H2-H22/23 is binding to Utp23 as expected following a hyperbolic binding curve. Therefore, despite ES6H2-H22/23

containing the H22/23 binding site similar to H21-H22/23 and H22/23, it did not follow the same trend of reaching a maximum amplitude around 20 nM Utp23 before decreasing. Notably, this trend of reaching a maximum RNA binding around 10-20 nM Utp23 followed by a decrease is reproducible and RNA specific. Other RNAs such as H21-ES6H3 and ES6H1-ES6H3 both display a phenomenon where the final data point is lower than the preceding one. It is possible that this reduction, while far smaller than the binding decrease witnessed in H21-H22/23, is a result of the same property of Utp23. If so, the affinities reported for H21-ES6H3 and ES6H1-ES6H3 in Table 3.6 may not be accurate.

In all binding reactions with Utp23, only two RNAs reached an amplitude over 10%: H21-H22/23 and ES6H2-H22/23 (Table 3.6). Of these, ES6H2-H22/23 had the highest affinity; it yielded a dissociation constant lower than 2 nM, which is below the RNA concentration in the reaction. This high affinity indicates tighter binding between Utp23 and the 18S rRNA than Utp23 binding to snR30 (see section 3.4.3) and demonstrates that Utp23 has the ability to form a complex with the 18S rRNA. Nevertheless, no RNA tested by nitrocellulose filter binding with Utp23 generated an amplitude over 30%. These low amplitudes, sometimes even less than 5%, reveal an inability of the protein to bind most of the RNA in the reaction. The most likely explanation is that every RNA that does not reach at least 10% binding potentially represents an unspecific interaction with Utp23. Similarly, those RNAs displaying less than 5% binding (H21-ES6H1, H22/23, H25/25, mitochondrial $tRNA^{Met}$) are extremely likely to be nonspecifically binding to Utp23. The 5% cut-off is influenced by previous filter binding experiments in which the snR30 Δ397 bind to Utp23 with less than 10% amplitude but still maintains a similar affinity compared to the full-length and snR30 Δ292 variants (Table 3.4). However, when the amplitude is under 5% binding, the noise of the measurements is too large compared with the weak signal

leading to high error as demonstrated by the 19 ± 28 nM K_D between Utp23 and the control RNA, mitochondrial $tRNA^{Met}$. As such, there is no evidence for binding between Utp23 and the mitochondrial $tRNA^{Met}$. For this reason, the primary conclusion of this series of experiments is that Utp23 is capable of binding to both the H21-H22/23 and ES6H2-H22/23 constructs. Both H21-H22/23 and ES6H2-H22/23 contain the H22/23 interaction site in addition to extra structural elements, notably ES6H3. Previous studies indicate that the main Utp23 interaction site is within H22, but the nearby secondary structure of ES6H3 may be necessary for stabilization of Utp23 to the 18S rRNA [88]. Utp23 may also specifically bind H21-ES6H3, ES6H1-ES6H3, ES6H2-ES6H3, and rm1-rm2. All four of these RNAs contain ES6H3 supporting the hypothesis that ES6H3 is important in stabilizing Utp23 on the 18S rRNA. The binding of Utp23 to H21-ES6H1, H22/23, H25/26, and mitochondrial $tRNA^{Met}$ is likely to be unspecific or require another factor to stabilize the binding interaction that is not present *in vitro*.

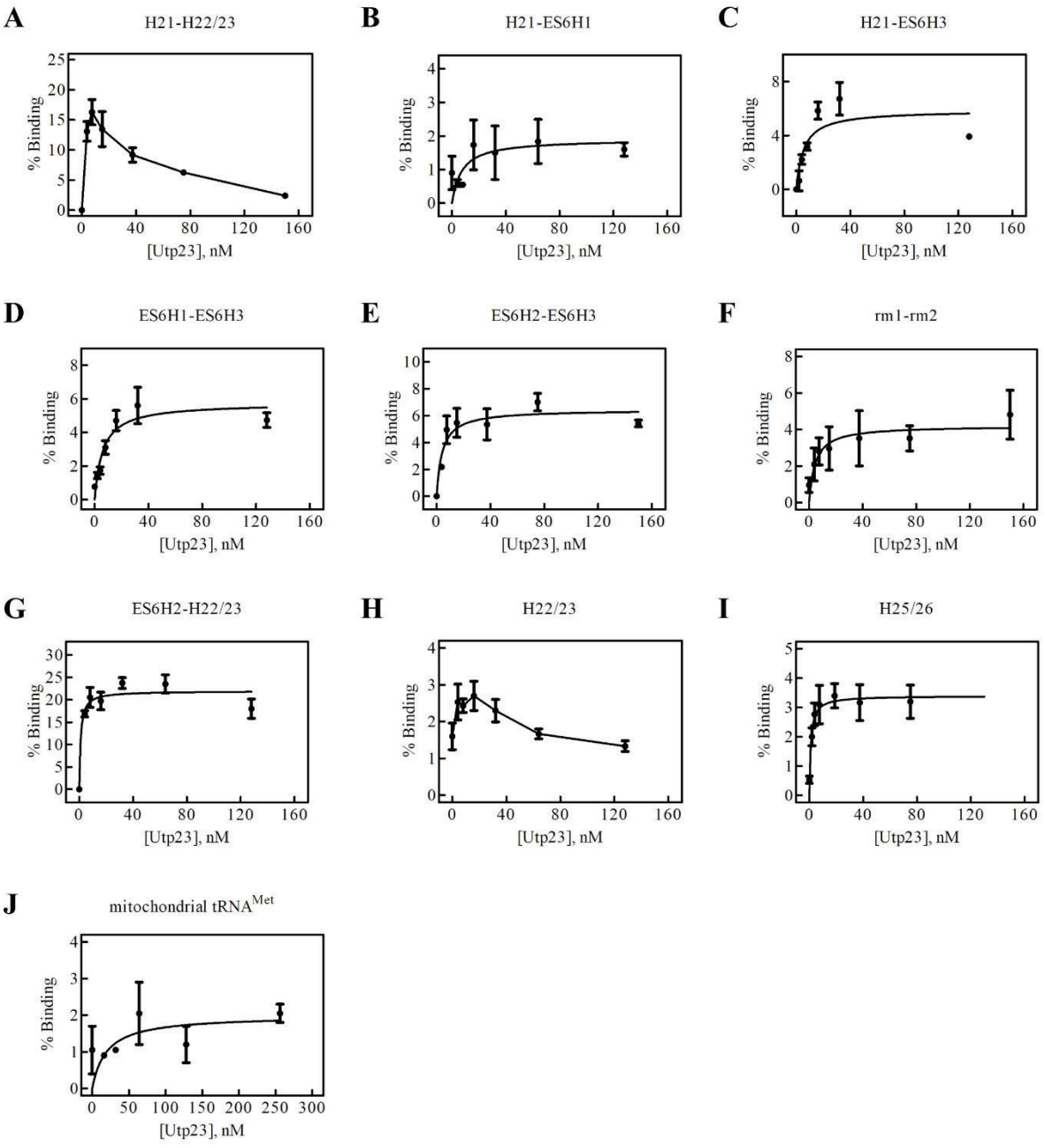

Figure 3.12. Utp23 nitrocellulose filter binding to different 18S rRNA fragments. (**A, H**) Average binding to H21-H22/23 and H22/23 (n=3), data points are connected by straight lines as a visual guide. (**B-G, I**) Average binding to the respective RNAs (n=3); data is fit with equation 2.1. (**J**) Average binding to mitochondrial $tRNA^{Met}$ (n=2), data is fit with equation 2.1. Dissociation constants are displayed in Table 3.6. Individual trials shown in Figure A.3.

To summarize, Utp23 seems to bind to each of the 18S RNA constructs tested. However, while there is an observable interaction between Utp23 and the 18S rRNA fragments, Utp23 binds an exceptionally low percentage of the RNA in most cases. This binding is hypothesized to be unspecific based on the low amplitude which is contrasted by the tight and higher amplitude binding of the RNAs (ES6H2-H22/23, full-length snR30) that are hypothesized to be specific. The caveat to this conclusion is the observation about binding decreasing at higher Utp23 concentrations for both H21-H22/23 and H22/23.

Table 3.6. Affinity between 18S rRNA fragments and Utp23

18S rRNA Constructs	Dissociation constant (nM)	Binding Amplitude (%)
H21-H22/23	NC	NC
H21-ES6H1	7 ± 8	2 ± 1
H21-ES6H3	5 ± 3	6 ± 1
ES6H1-ES6H3	6 ± 2	6 ± 1
ES6H2-ES6H3	4 ± 2	6 ± 1
rm1-rm2	4 ± 4	4 ± 1
ES6H2-H22/23	< 2 ± 1*	22 ± 1.1
H22/23	NC	NC
H25/26	< 2 ± 1*	3 ± 1
mitochondrial $tRNA^{Met}$	19 ± 28	2 ± 1

NC means the affinity could not be determined. * Affinity is below the concentration of RNA used in the reaction.

CHAPTER 4: DISCUSSION

4.1 Optimizations, Benefits, and Disadvantages of an *In Vitro* System

The prerequisite for my investigations was the successful *in vitro* reconstitution of the snR30 RNP from purified components. While purifying the numerous factors required for *in vitro* studies of the snR30 RNP, many improvements were made to allow for the consistent generation of these RNAs and proteins. Despite this, I think that there are further optimizations possible to generate additional factors expanding the capabilities of the snR30 RNP system. Currently, I have produced and optimized the production of snR30 and its truncation, fragments of the 18S rRNA, and the essential nucleolar protein Utp23.

Firstly, the preparation of snR30 has been achieved. Transcription of snR30 is challenging due to its length, and its tendency to degrade. Most likely, this degradation is enhanced by the numerous freeze-thaw cycles that the RNA may undergo. By streamlining the purification workflow, the degradation of the RNA should be reduced. An alternative solution is to instead work with the snR30 Δ397, a far shorter RNA that is more stable. Furthermore, this truncation contains all essential elements of snR30 except for the H box, whose role most likely deals with biogenesis and nucleolar import of the ribonucleoprotein rather than ribosome biogenesis [1]. As such, the snR30 Δ397 should be comparable to full-length snR30 in making essential interactions with both proteins and the 18S rRNA. For example, the truncation can be used when a higher concentration of snR30 is required than is currently possible with the full-length construct. A high concentration of the RNA becomes crucial when investigating the rate of dissociation between snR30 and a protein in chase-filter binding experiments. Therein, the complex of protein and radiolabeled RNA is incubated with a large excess of non-radiolabeled RNA, ensuring that upon complex dissociation, the protein binds to non-radiolabeled RNA.

Chase-filter binding creates a time-dependent exponential decay signal allowing the determination of the dissociation rate constant, k_{off}. As snR30 Δ397 can be purified at a higher concentration than full-length snR30, and contains the essential elements of full length snR30, experiments such as chase filter binding can be simplified through the use of the truncation.

Next, the production of RNAs based on the 18S rRNA was greatly improved. Initial *in vitro* transcriptions generated extremely low amounts of RNA (data not shown). Through rigorous testing, it was discovered that the presence of a 5′ GC double nucleotide preceding the T7 promoter sequence vastly improves transcription efficiency. Most likely, these nucleotides are important for providing additional stability to the DNA double helix, as well as closing off the end of the transcription bubble that the T7 RNA polymerase creates. Upon addition of the 5′ GC nucleotides, the new templates generate consistent and high levels of RNA in *in vitro* transcription reactions. The last optimization required to produce these RNAs radioactively was adjusting the concentration of UTP in the reactions. Since the region from helix 21 to helix 23 of the 18S rRNA is comprised of more than one-third uridine, the standard radiolabeling procedure was unsuccessful in producing many regions of the 18S rRNA [106]. Since [^{3}H]UTP is suspended in ethanol, a chemical that hinders *in vitro* transcriptions, the concentration of [^{3}H]UTP cannot be increased. However, by adding more non-radioactive UTP, the total concentration of UTP can be increased. This approach reduces the specific activity in the resultant RNA but enhances the efficiency of the *in vitro* transcription reaction. Through these optimizations, the production of new 18S rRNA fragments has been refined allowing us in the future to produce other regions of the 18S rRNA that are predicted to interact with snR30, such as helix 15 [35].

Lastly, the purification of Utp23 required a lot of optimization, as this PIN endonuclease protein displayed minimal expression under the previously published protocol [88]. The first optimization to enhance production of Utp23 was changing the incubation temperature during expression from 23°C for three hours to 18°C overnight. By lowering the expression temperature, more Utp23 is expressed as a soluble protein as has been shown for fusion proteins previously [111]. More Utp23 may be expressed at lower temperature because the protein has more time during slower protein translation to fold correctly [112]. Following successful expression, the purification of Utp23 needed to be improved. To do this, PMSF and DTT were added to the purification buffer to prevent protease activity, and to ensure the protein was maintained in a reduced environment, respectively. Reducing conditions may be particularly important when expressing Utp23, as it is localized to the nucleolus in yeast, an organelle that is maintained in a reduced state and only becomes oxidized when responding to stress [113]. Following purification of Utp23, visualization of the protein revealed a high abundance of contamination in the form of the *E. coli* chaperones DnaK and GroEL totaling over 5% of the protein content. SEC revealed that the affinity of the chaperones to Utp23 is high enough to cause the proteins to co-elute as a single peak. The addition of ATP to the protein sample before loading onto the column may aid in separation of Utp23 and the chaperones, by inducing chaperone turnover and release of Utp23 from the complex [114]. According to previously published studies, bacterial chaperones are common contaminants in the purification of fusion proteins; however, they can often be removed by washing the protein when it is bound to the (affinity) resin with buffer supplemented with magnesium ions, ATP, and denatured *E. coli* proteins [114, 115]. Another step that may aid in enhancing protein purity is a wash step when Utp23 is bound to the glutathione resin at high ionic strength. By increasing salt concentration in

the buffer, weak ionic interactions of other contaminants (besides the chaperones) to either the resin or GST-Utp23 will be disrupted. By applying high salt wash steps in parallel to the addition of ATP during the Glutathione Sepharose chromatography, it should be possible to generate Utp23 at a higher purity than currently obtained. Ideally, SEC is utilized following affinity purification to increase protein purity. SEC dilutes the protein to a concentration lower than the minimum concentration needed for most assays, requiring the fractions containing protein to be concentrated, e.g., by ultrafiltration. This process is impossible with NP40, as the detergent forms micelles larger than GST-UTP, preventing the detergent from flowing through the cut-off filter resulting in NP40 being concentrated together with the protein. As such, one essential change in the purification of Utp23 in the future is to alter the detergent in the elution buffer from NP40 to Tween 20. Tween 20 is the detergent of choice as it does not form micelles, thus allowing for protein concentration by ultrafiltration. Furthermore, Tween 20 was used in previous studies of Utp23 during protein-protein interaction assays [88]. Alternative approaches to Utp23 purification include purifying the protein from yeast cells. This approach has the advantage that yeast cells can properly fold and localize Utp23 to the nucleolus unlike *E. coli*. Additionally, Utp23 possesses a phosphoserine at position 182 which may aid in stability and function and which will be lacking when expressing Utp23 in *E. coli* [116]. The last possibility is co-expression of Utp23 with a stabilizing factor such as Utp24. As the proteins are known to interact, the protein-protein interaction may yield a higher percentage of stable, functional protein [88].

In conclusion, the purification of the *in vitro* factors has been optimized. Thereby, I ensure future expansions and investigations of the snR30 *in vitro* reconstitution system are streamlined and efficient. Notably, my optimizations do not include the work required to

reconstitute the H/ACA core proteins which have already been published [63]. Based on my reconstitution of a purified snR30 RNP, the potential of this *in vitro* system to discover novel information surrounding thermodynamics and kinetics is unmatched. In particular, there are three main benefits to working in an *in vitro* system with purified components: isolation from external factors, quantitative data on the mechanism, and the capability for increasing system complexity. First, an *in vitro* system removes the influence from other factors. By utilizing a system comprised solely of purified components, each interaction can be studied in isolation to yield quantitative insight into the architecture of the system. For snR30, this is especially important, as its role during ribosome biogenesis is to bind co-transcriptionally and then be released prior to completion of the 35S pre-rRNA transcript. The short time that snR30 is bound to rRNA for has limited studies about the structure, as well as the function of snR30. An *in vitro* system eliminates the helicases responsible for removing snR30 from the pre-ribosomal particles, enabling studies of how snR30 binds and interacts with the 18S rRNA. The second benefit of *in vitro* systems is the ability to generate quantitative thermodynamic data for the interaction of different components. By measuring the affinity of proteins and RNAs for each other, it is possible to tease out differences as small as a two-fold change that would be invisible to qualitative studies such as the yeast two hybrid approach. These small differences in affinity, or lack thereof, provide insight into which elements are required to stabilize an interaction. By informing which factors are required for stabilization, the *in vitro* system can generate novel findings utilizing quantitative data. Finally, the *in vitro* system created here can be made increasingly more complex as additional ribosome assembly factors such as Rok1, Has1, and Utp24 are purified and added. Increasing the intricacy of the system by a single factor at a time allows studying the exact contribution of each factor for complex stability and its interaction

with snR30. This benefit is highlighted in the discovery that Nhp2 is dispensable for tight binding of the Cbf5-Nop10-Gar1 trimeric complex to snR30.

Using a highly purified *in vitro* experimental system also has significant disadvantages. As previously highlighted, purification of the individual components is a major bottleneck in generating a functional system. Thus, the versatility of the system is limited by the number of factors one can purify. Furthermore, the *in vitro* system can only address known interactions, and any interactions involving unknown factors will be missed. In the scenario that an interaction is missing, the data generated from the *in vitro* system will give incomplete insight into the system's mechanism. Another way that interactions can be missed is in lacking post-translational modifications. For the human homologs of Nhp2 and Cbf5, both are SUMOylated leading to increased binding to snoRNAs [117]. However, in my system, the purified proteins from *E. coli* are lacking all modifications. Lastly, by purifying the components, the interactions between them will not be affected by any competing factors or the cellular environment. Therefore, the interactions measured *in vitro* may be different from the *in vivo* situation if there is a missing factor or if the reaction conditions are significantly different from the normal cellular environment. During the purification of snR30, this possibility must be considered as the long RNA may adopt many different confirmations. To determine how snR30 folds *in vitro,* Selective 2' Hydroxyl Acylation analyzed by Primer Extension (SHAPE) could be used to assess whether the RNA folds into the *in vivo* conformation. Due to this, it is essential that all results obtained with an *in vitro* system are carefully considered to make sure they can be understood in the context of the living cell and the role the factors are expected to play therein.

4.2 Mechanism and Function of snR30 during Ribosome Biogenesis

4.2.1 Formation of snR30 snoRNP in Comparison to Modification H/ACA snoRNPs

Like the modification H/ACA guide RNAs that direct pseudouridine formation, I have demonstrated that snR30 also forms a tight, sub-nanomolar interaction with the core H/ACA proteins Cbf5-Nop10-Gar1-Nhp2 [63]. By confirming that snR30 interacts with H/ACA core proteins in a similar manner as modification guide RNAs, I have shown that all types of H/ACA guide RNAs follow a conserved mode of biogenesis despite playing different roles in the cell. The similar affinity of all H/ACA RNAs for H/ACA proteins also implies that snR30 likely forms an active complex upon binding to the core H/ACA proteins; however, this confirmed activity of modification guide RNAs cannot be proven for snR30, as it does not have a target for pseudouridylation [63]. As an orphan RNA with no pseudouridylation target, snR30 differs from snR10, which has two active hairpins: the 5′ hairpin, which binds the 18S rRNA but does not modify it, and the 3′ hairpin, which introduces a pseudouridine at position 2923 of the 25S rRNA [10]. While the pseudouridine introduced by snR10 is dispensable with no phenotype, yeast lacking the binding region from the 5′ hairpin of snR10 develop a cold-sensitive phenotype [36]. Since all characterized H/ACA snoRNAs (snR30, snR5, snR34) bind to the core H/ACA proteins Cbf5-Nop10-Gar1-Nhp2 with a similar affinity, it is reasonable to hypothesize that all H/ACA guide RNAs, including snR10, will interact in the same tight manner with the H/ACA core proteins, particularly the Cbf5-Nop10-Gar1 core trimeric complex [63].

Nhp2's affinity to snR30 is different than expected. Caton et al [63] determined the affinity of Nhp2 to the H/ACA guide RNA snR34 to be 600 ± 200 nM. My data indicates that the dissociation constant between snR30 and Nhp2 is 27 ± 3 nM for the full-length RNA. Notably, Nhp2's affinity for snR30 is far lower than the sub-nanomolar affinity of Cbf5-Nop10-Gar1 for snR30. *In vivo,* the absence of Nhp2 is lethal and prevents H/ACA sRNP-mediated pseudouridylation; pre-rRNA processing also fails in the absence of Nhp2 indicating that the

snR30 RNP cannot function either [17, 118]. Based on this functional requirement for Nhp2, the large change in affinity between a modification H/ACA guide RNA and snR30 suggests Nhp2 may have a unique binding site within snR30. The tighter binding of Nhp2 to snR30 may involve the essential element located near the top of snR30's 3′ hairpin [85]. This hypothesis is based on both the observation that this element is essential for snR30 function as well as that the region is very close to where Nhp2 would bind in a modification guide RNA complex [119]. Furthermore, my data confirms that Nhp2 and GST-Utp23 interact with each other. Since Utp23 requires snR30 for its recruitment to the pre-ribosomal particle, Utp23 may rely on Nhp2 to remain bound to the snR30 RNP [94]. This role seems likely as the internal hairpin of snR30 is dispensable for activity *in vivo* [75]; therefore, Nhp2 is likely critical for maintaining Utp23's interaction to the snR30 RNP. Thus, Nhp2's higher affinity for snR30 than the modification H/ACA guide RNAs may be related to Nhp2's unique role in recruiting Utp23 to the snR30 RNP. By interacting with Nhp2, Utp23 is most likely held in a conformation where it is also able to bind to the internal hairpin of snR30.

To summarize, snR30 is an orphan H/ACA guide RNA that behaves similarly to the modification guide RNAs in binding H/ACA proteins during biogenesis of the RNP. Improvements in cryo-EM microscopy may hold the key to determining both how snR30 folds and how the proteins composing its RNP interact. With the recent advancements in cryo-EM data gathering, collection, and processing, resolving a high-resolution structure may be possible in the future (reviewed in [120]). To do so, the RNP components need to be pure and at a high enough concentration for imaging which will be facilitated by my work reported in this thesis to reconstitute a pure snR30 RNP.

4.2.2 Insight into snR30's Role in Ribosome Biogenesis

According to the nitrocellulose filter binding data presented in this thesis, the affinity between the C2 and C3 sites and the snR30 RNP is lower than the interaction of the snR30 RNP with its primary C1 interaction (>200 nM compared to <100 nM). Potentially, this weak binding of snR30 to C2 and C3 is a result of extended secondary structure in the full-length snR30 RNP that prevents base-pairing to the secondary sites in the 18S rRNA fragments. Without the entire pre-ribosome and assembly factors present, the C2 and C3 sites may not be able on their own to facilitate binding between the snR30 RNP and the 18S rRNA. Unfortunately, high error in the dissociation constant for binding of snR30 RNP to the C3 site comprised in the H25/26 fragment (225 ± 227 nM) prevents drawing clear conclusions. The binding of snR30 Δ397 RNP to the H25/26 construct has an affinity of 92 ± 39 nM. This more accurate data suggests that the C3 site may be able to interact independently of exterior factors, but further experiments are required to confirm this. Most importantly, my data shows that as long as the C1 site is present within the 18S rRNA , adding the C2 or C3 site in the target RNA does not increase the affinity for the snR30 RNP. When the C1 interaction site is present, a dissociation constant of approximately 80 nM is observed for the H21-H22/23, ES6H1-ES6H3, ES2H3-ES6H3, rm1-rm2, and ES6H2-H22/23 RNAs. Thus, my findings suggest that the C1 site stabilizes the snR30 RNP to the pre-ribosome while the presence of the C2 and C3 sites do not contribute to stabilizing the interaction between the 18S rRNA and the snR30 RNP. However, while there is little to no measurable effect to having the C2 and C3 interaction sites *in vitro*, they may possess a role *in vivo* that is not stabilization of the snR30 RNP to the pre-ribosome.

As mentioned previously, Utp23's main binding site within snR30 is located within the internal hairpin, and therefore may allow Utp23 to unwind the internal hairpin allowing access to the C3 site. I hypothesize that repetition of the nitrocellulose filter binding experiments with

Utp23 present will increase the affinity of the snR30 RNP specifically for 18S rRNA constructs containing the C3 interaction site. This possibility may extend to the C2 site but cannot be accurately predicted at this time. Previous data demonstrates that the C1 site is the only one of the three binding sites in 18S rRNA that is essential for snR30 function. Binding of snR30 to the rm1 and rm2 motifs (the C1 site) may be essential because it induces a refolding of the 18S rRNA that may be required for subsequent steps in the process of ribosome biogenesis [75, 88]. Even if the C2 and C3 interaction sites aid in the stabilization of either snR30 to the 18S rRNA or of snR30 to the hypothesized snoRNA network, the C2 and C3 sites are redundant and not essential for snR30 function, but their presence may still provide some benefit [35]. In Atzorn et al [75], the authors replaced the 5′ hairpin of snR30 with the snR5 5′ hairpin with no noticeable phenotype. Interestingly, when both the 5′ and internal hairpin of snR30 were replaced by the snR5 5′ hairpin, the growth of the yeast was reduced by approximately 40%. As such, the C2 and C3 sites within these hairpins are likely relevant in aiding snR30 in completing its function in ribosome biogenesis.

To conclude, the snR30 RNP interacts tightly with the C1 site composed of the rm1 and rm2 motifs. Addition of other proposed 18S rRNA interaction sites did not stabilize the interaction further. My data demonstrates that this tight interaction between the C1 site and the snR30 RNP is likely maintained during the entire time when snR30 is associated with the pre-ribosomal particle. As such, during ribosome biogenesis, the C1 site anchors the snR30 complex to the pre-ribosome. Unlike the C1 site, the contacts of the C2 and C3 sites with the pre-ribosome during biogenesis may not be preserved throughout all remodeling steps that the pre-ribosome undergoes. From these findings, I have determined that the C2 and C3 interaction sites are not necessary for snR30 stabilization to the pre-ribosome. The function of C2 and C3 may be

to help chaperoning the pre-rRNA or orientating the snR30 RNP correctly to form the snoRNA interaction network hypothesized in Martin et al. [35]. This snoRNA interaction network hypothesis suggests that snR30 may contribute to folding many of the expansion segments of the 18S rRNA by interacting with the snR10, U14, and U3 snoRNAs [35]. The interaction network is complemented by the helicase Rok1 and likely additional snR30-specific proteins like Utp23 and Kri1. The hypothesis that the snoRNAs form an interaction network is further supported by the fact that the majority of the snoRNA interaction sites occur in expansion segments of the rRNA; thus, the snoRNAs may refold ribosomal elements that do not exist in the bacterial ribosome – the evolutionary precursor to the eukaryotic ribosome [35]. Since the snoRNAs themselves are a feature specific to eukaryotes and archaea, this raises the tantalizing possibility that the essential guide RNAs U3, U14, snR10, and snR30 were among the first of their class to evolve in tandem with the expansion segments not found in the bacterial ribosome. While this potential for co-evolution between the ribosome and snoRNAs is interesting, it does not clarify snR30's role any further.

A large challenge in discovering more about snR30 is the transient nature of snR30's interaction with pre-rRNA during ribosome biogenesis. This is highlighted by the fact that in every pre-ribosomal structure published so far snR30 is missing, likely because it has already dissociated from the pre-ribosome [42]. Therefore, the exact interaction network between snR30 and the other ribosome biogenesis factors remains inconclusive. To expand on my findings reported here, there are numerous further avenues to pursue. First, it will be interesting to repeat the nitrocellulose filter binding experiments with 18S rRNA fragments by supplementing the snR30 complex with Utp23. Thereby, it will be possible to quantify any change Utp23 induces in binding of the snR30 RNP to pre-rRNA. In particular, I hypothesize that the largest change in

affinity will occur for the H25/26 rRNA fragment. This experiment will determine if Utp23 can melt the internal hairpin of snR30 and enable specific binding of the H25/26 region to the internal hairpin. To complement this study, another RNA construct comprising helices 22 to 26 could be generated. In this H22-H26 construct, Utp23's binding site within 18S rRNA is present. I hypothesize that adding Utp23 to the system will increase the affinity of the snR30 RNP to the H22-H26 RNA. The region from helices 22 to 26 contains not only Utp23's binding site, but there are also two binding sites for Utp23 on the snR30 RNP through interactions with Nhp2 and the internal hairpin. Therefore, there are in total three different interaction sites for Utp23 enabling it to stabilize binding of the H22-H26 RNA to the snR30 RNP. Furthermore, the proteins used in snR30 RNP binding assays should be expanded to include the helicases Rok1 and Has1, the essential nucleolar protein Kri1, and the endonuclease Utp24.

4.3 Utp23's Role and Interactions

In the nitrocellulose filter binding experiments with snR30, Utp23 behaves as hypothesized; it binds to full-length, Δ292, and Δ397 snR30 specifically with low-nanomolar affinities (<10 nM). Notably, Utp23's nanomolar affinity to snR30 is an order of magnitude lower than the subnanomolar affinity of snR30 to the core H/ACA proteins, Cbf5-Nop10-Gar1-Nhp2 (<1 nM). As an accessory protein to the snR30 RNP, Utp23 presumably binds to snR30 after assembly of the Cbf5-Nop10-Gar1-Nhp2 proteins. Most likely, Utp23 binds the snR30 RNP during ribosome biogenesis and dissociates just after snR30's release from the pre-ribosomal particle [42, 94]. Thus, the interaction between Utp23 and snR30 is expected to be weaker than the interaction between snR30 and the core H/ACA proteins.

However, there is a low amplitude when snR30 is binding Utp23 in the nitrocellulose filter binding experiments, which is unexpected. As discussed earlier, this amplitude change can

be based on a number of factors such as how quickly the protein releases the RNA, how stringent the washing during nitrocellulose filtration is, and whether a large amount of RNA is in a conformation that cannot be bound. Of these three possibilities, the probability of Utp23 quickly releasing the RNA is unlikely. The amount of RNA that is washed away is dependent upon the protein's rate of dissociation. Potentially, Utp23 has a very high association rate constant k_{on} rate that is offset by a high dissociation rate constant k_{off}. As a ribosome assembly factor however, Utp23 needs to able to stably interact with the pre-ribosome. As such, Utp23 should not dissociate from snR30 once it finds the correct binding location. Likewise, the nitrocellulose filtration washing step is unlikely to be too stringent as the same washing conditions can yield amplitudes near 100%, albeit for different protein. Lastly, the possibility that the amplitude is limited by how much RNA is in a conformation that can be bound to the complex is harder to discard, as Utp23's requirement for binding snR30 is unknown. As mentioned, *in vivo* Utp23 binds to the snR30 RNP, not snR30. Therefore, *in vitro* it is unknown how the snR30 RNA will fold. The amplitude of binding decreases from the full-length snR30 to snR30 Δ397 with Utp23. This reduction could indicate that the increased length of the full-length snR30 aids in proper folding, and the snR30 Δ397 may sample more conformations that cannot bind Utp23. Assuming Utp23's main binding site to snR30 is in the internal hairpin as determined by Wells et al., then proper folding of the internal hairpin may be essential for Utp23 to bind [88]. In the gathered data, the amplitude decreased from 30 ± 2 % for full length snR30 to 6 ± 0.5 % for snR30 Δ397. Currently, it is unknown why this decrease is happening. Future experimentation is required to determine a mechanism or cause. As such, my data indicates that Utp23 may be unable to interact with a large percentage of snR30 in absence of the core H/ACA proteins, as these proteins should stabilize snR30 in its mature conformation.

Next, I quantified Utp23's interaction with the 18S rRNA fragments. These experiments do not mimic *in vivo* conditions where Utp23 requires snR30 to be recruited to the pre-ribosome. I observed tight binding following a hyperbolic binding curve for only one 18S rRNA fragment, ES6H2-H22/23. The ES6H2-H22/23 RNA contains both the ES6H3 and H22 helices of the 18S rRNA that Utp23 has been previously characterized to interact with *in vivo* [88]. Of the other 18S rRNA constructs, Utp23 is unable to bind more than 5% of the H21-ES6H1, rm1-rm2, H22/23, H25/26, and mitochondrial $tRNA^{Met}$. According to the fitting with a hyperbolic equation, Utp23 appears to interact with low nanomolar affinities with all four constructs. During ribosome biogenesis, there are many RNA sites for Utp23 to bind. As such, I speculate that this consistent binding with extremely low amplitudes indicates that Utp23 is able to nonspecifically interact with RNA to sample it for the correct binding site before dissociation. These extremely low amplitudes are contrasted by the tightest bound RNA, ES6H2-H22/23, which reaches an amplitude above 20%. However, as a ribosome biogenesis protein, Utp23 should be able to bind more than 20% of the RNA. Therefore, as the snR30 RNP is required to recruit Utp23 to the pre-ribosome *in vivo*, the snR30 RNP is likely required to stabilize Utp23 on the pre-rRNA [88].

When binding to the longest RNA construct, H21-H22/23, Utp23 exhibits extremely strange behavior: after reaching a binding amplitude of approximately 22%, the percentage of RNA being bound begins to decrease upon increasing the Utp23 concentration (Figure 3.12 A). This phenomenon was observed to various extents for several of the different RNAs tested, but was notably absent during snR30 binding, implying that this is happening in an RNA-dependent fashion that is sensitive to Utp23 concentration. One possible explanation for this pattern is that Utp23 may be dimerizing [88]. This dimerization may not be physiologically relevant, as the protein Utp23 is expected to bind to its paralog Utp24 *in vivo* whose nuclease activity is essential

for cleavage at the A_1 and A_2 sites of the pre-rRNA [89]. Therefore, when two monomers of Utp23 bind to each other, the dimer may be unable to bind the RNA construct. Unfortunately, the current data does not provide enough information to speculate further; however, a series of further experiments should help elucidate the reason less RNA is bound at higher Utp23 concentrations. The most promising experiment will be to repeat the binding of Utp23 to the H21-H22/23 construct but switch the titrant to the RNA. In this set-up, the concentration of Utp23 is held constant and replicates can be generated at a variety of RNA concentrations. Based on my data, where binding of Utp23 to H21-H22/23 reaches a maximum around 20 nM of Utp23, the potential dimerization of Utp23 may occur with a K_D of about 20 nM. At Utp23 concentrations below 20 nM, Utp23 may be in its monomeric state and able to successfully bind the RNA construct. Thus, I propose that at Utp23 concentrations below 20 nM, the binding curve will be hyperbolic and reach high levels of binding at high RNA concentrations with a consistent dissociation constant. However, when the Utp23 concentration is closer to 100 nM, it is likely that a low constant percentage of RNA binding will be observed independent of the RNA concentration. This binding will reflect the low concentration of monomeric Utp23 capable of binding RNA in contrast to the larger fraction of dimeric Utp23 that is unable to bind RNA. If this is the behavior shown in the experiment, it is conclusive evidence that Utp23 undergoes a change in its ability to bind RNA, likely through dimerization, at a concentration around 20 nM.

In vivo, Utp23 may not be able to dimerize due to its interaction with Utp24 or Nhp2. To study whether the interaction between Utp23 and Nhp2 prevents Utp23 dimerization, SEC can be used. The GST-Utp23 dimer will have a mass of around 106 kDa. If Nhp2 disrupts dimerization, the elution volume should increase as the GST-UTP23-Nhp2 heterodimer has a size of 72 kDa. However, if the dimerized Utp23 can bind Nhp2, the resultant trimeric complex will be

characterized by a larger size than the Utp23 homodimer. To gain quantitative analysis, SEC coupled to multi-angle light scattering (SEC-MALS) can be employed. This technique has to the ability to quantify the size of the eluting particle confirming the stoichiometry of the complex [121].

To summarize, Utp23 is an extremely interesting ribosome biogenesis factor, and there is more work needed to fully elucidate its function. Notably, Utp23 binds tightly to the ES6H2-H22/23 RNA construct. With an affinity of $<2 \pm 1$ nM and an amplitude of 22 ± 1.1 %, only this construct of all the 18S rRNA fragments tested displays both an amplitude over 20% as well as a hyperbolic fit. As indicated by Wells et al., Utp23 binds to the pre-rRNA at ES6H3 and H22 [88]. My data corroborates this finding that helix 22 is not enough on its own to bind Utp23, and that Utp23 also relies upon ES6H3 for stabilization to the pre-rRNA. Since the snR30 RNP is required for Utp23 recruitment to the pre-ribosome, the presence of the snR30 RNP may enhance Utp23's binding to the pre-rRNA. Utp23 in turn may also aid in the stabilization of the snR30 RNP on the pre-ribosome. With future studies, it could be possible to unravel exactly how Utp23 interacts with both snR30 and the 18S rRNA, in addition to expanding the purified factors to elucidate if there is another protein which is required for proper Utp23 binding to 18S rRNA.

4.4 Future Directions

The work in this thesis represents an initial quantitative characterization of the snR30 RNP and its interaction with the pre-rRNA and the Utp23 assembly factor during ribosome biogenesis. Based on my findings and given the complex structure and function of snR30, there are many questions left to answer surrounding snR30. While my work has clarified that the C1 site comprised of the rm1 and rm2 motifs is the main anchor for the snR30 RNP on the pre-ribosome, it remains to be clarified how much the other interactions contribute to the stability of

the interaction *in vivo*. Also, while it is known that neither the C2 nor the C3 sites are essential, they appear to convey some fitness advantage [75]. Why is that advantage not seen in the *in vitro* system? Furthermore, is that fitness advantage related to the snoRNA interaction network that snR30 is a part of [35]? Through expansion of the *in vitro* system, insights into all of these questions can be gained. Systematic addition of both proteins and other sections of the 18S rRNA will elucidate which protein factors and/or sections of the 18S rRNA maximize the interaction affinity. Likewise, by continuing to utilize the full-length and snR30 Δ397 constructs in addition to including others such as a deletion of the internal hairpin, exact knowledge about the sections required for tight binding to the pre-ribosome will be determined.

As demonstrated by the Nhp2-Utp23 pulldowns, probing additional protein-protein interactions by pulldown assays is critical and feasible. Pulldown assays will help to clarify the exact interactions between the proteins composing the snR30 RNP. To supplement the reported binding of Nhp2 to Utp23, it will be important to show if Utp23 binds to the protein core of Cbf5-Nop10-Gar1 to establish how extensive the interactions between Utp23 and the core H/ACA protein complex are [88]. Previous yeast genetic studies indicate an interaction between Gar1 and Utp23 which suggests that Utp23 interacts with the H/ACA proteins beyond Nhp2 [122]. Further protein interactions to test include Utp23 – Rok1, Utp23 – Has1, Nhp2 – Rok1, Nhp2 – Has1, Cbf5-Nop10-Gar1 – Rok1, Cbf5-Nop10-Gar1 – Has1, and Utp23 – Utp24. Importantly, Rok1 and Has1 are RNA helicases involved in snR30 release from the pre-ribosomal particles. Pull-down assays will clarify which snR30-associated proteins interact with Rok1 and Has1 providing insight into how these helicases recognize the snR30 RNP complex. Furthermore, investigating the interaction of Rok1 and Has1 with the snR30 RNP and 18S rRNA will aid in identifying which helicase can unwind the base-pairing between snR30 and the 18S

rRNA. By expanding the range of protein factors, the various protein interactions facilitating snR30 binding to and release from the pre-ribosome can be resolved.

Perhaps the most tantalizing future direction is the possibility of resolving a structure of the snR30 RNP, for example using cryo-electron microscopy. Towards this goal, the snR30/U17 RNP could either be purified from yeast or human cells, or reconstituted from purified components. While the purification from cells has a greater chance of producing a biologically relevant complex, the reconstitution may be the more feasible method as most of the components have now been purified. Also, as it is likely that many of the snR30-associated proteins form transient interactions, purification of the entire complex from cells using tagged snR30 may lead to high complex heterogeneity. This problem can be avoided using purified components followed by SEC to verify whether the complex elutes in a homogeneous fashion. As generating cryo-EM data requires expensive equipment and expertise in cryo-electron microscopy; confirmation of the complex homogeneity prior to data collection is essential [120].

Ultimately, insight into snR30's role during ribosome biogenesis is incomplete without knowledge of the molecular interactions the snR30 RNP forms upon binding to the pre-rRNA. To resolve the structure of the snR30 RNP docked on the pre-ribosome, the snR30 complex must be stabilized on the pre-ribosome. Interestingly, it has already been shown that the RNA helicase Rok1 is required for snR30 dissociation [93]. Therefore, in a Rok1 knock-down yeast strain, any purified pre-ribosomes should have snR30 still attached to them. A structure of this complex would give the best insight into snR30's exact role during ribosome biogenesis.

In conclusion, this thesis provides the basis for many future avenues of research. One of the most crucial next steps is the expansion of available protein factors to study. Utp24 is the most interesting based on its role as the endonuclease catalyzing the cleavages at site A_1 and A_2

of the pre-rRNA [89]. While the majority of Utp24 crosslinking sites are to the U3 snoRNA, Utp24 is of particular interest due to its role as an endonuclease in ribosome biogenesis [89]. The *in vitro* system purified in this thesis has the capacity to quantify Utp24's affinity to the 18S rRNA, especially the crosslinking sites detected in ES6H3, H22, and H26 [89].

4.5 Conclusions

In this thesis, I report the quantitative characterization of critical interactions formed by snR30 and its associated proteins during ribosome biogenesis. First, the binding of both the core H/ACA proteins Cbf5-Nop10-Gar1-Nhp2 and Utp23 to snR30 was characterized through nitrocellulose filter binding assays. These assays verified that snR30 is bound tightly to the core H/ACA proteins similarly to modification H/ACA RNAs, and the accessory protein Utp23 binds to snR30 in a specific manner with high affinity. Furthermore, the affinity between the snR30 RNP and the 18S rRNA constructs was determined by nitrocellulose filter binding. From this data, I have gained novel insights into the affinity of snR30 for the three proposed interaction sites C1, C2, and C3 [35]. As hypothesized, the full-length snR30 complex interacts more strongly with the rm1 and rm2 motifs comprising the C1 interaction site in expansion segment 6 of the 18S rRNA than with the secondary sites of C2, C3, and H22/23. Finally, the PIN endonuclease Utp23 was shown to interact tightly only to ES6H2-H22/23 suggesting that Utp23 likely requires multiple contact points in 18S rRNA, but also the snR30 RNP, to stably associate with the pre-ribosome.

The work in this thesis provides a first quantification of the affinity between snR30 and the pre-rRNA and as such, has generated three conclusions about ribosome biogenesis. First, the tight binding to snR30 stabilizes Utp23 on the pre-ribosome allowing it to fulfill its function. Next, the C1 site is the most important of the three RNA-RNA interactions between the snR30

RNP and the pre-ribosome, anchoring snR30 on the pre-ribosome during remodeling of the pre-ribosome and pre-rRNA processing. Lastly, after recruitment to the pre-ribosome by the snR30 RNP, Utp23 forms further interactions to H22 and ES6H3 in the pre-rRNA thereby contributing to the stabilization of snR30 RNP on the pre-ribosome.

To summarize, the results presented in this thesis are of foundational importance to establishing an exact modus operandi for snR30 function. My *in vitro* system has gathered the first quantitative data about the interactions formed by the unique snoRNA, snR30. This work sets the groundwork for many potential advances in related fields. Of particular interest is snR30's human homolog, U17, which has the potential to serve as an anti-cancer target. Since snR30/U17 is required for ribosome biogenesis, a process that is upregulated in actively dividing cells such as cancers, understanding U17's interactions within the pre-ribosomal particle is crucial in designing an antisense oligonucleotide that can target and interfere with its function [123]. Notably, the m1 and m2 motifs in snR30 are conserved with no base changes between snR30 and U17. The high conservation of these motifs combined with my data confirm that these motifs which compose the C1 interaction site are responsible for anchoring the snR30/U17 RNP on the pre-ribosome. Therefore, my findings may serve as a foundation for designing an oligonucleotide to target snR30/U17 function similarly to previous reports targeting the essential snoRNAs U3 and U8 [124].

References

1. Vos, T.J. and U. Kothe, *snR30/U17 Small Nucleolar Ribonucleoprotein: A Critical Player during Ribosome Biogenesis.* Cells, 2020. **9**(10).
2. Eliceiri, G.L., *The vertebrate E1/U17 small nucleolar ribonucleoprotein particle.* J Cell Biochem, 2006. **98**(3): p. 486-495.
3. Thomson, E., S. Ferreira-Cerca, and E. Hurt, *Eukaryotic ribosome biogenesis at a glance.* J Cell Sci, 2013. **126**(Pt 21): p. 4815-4821.
4. Klinge, S. and J.L. Woolford, Jr., *Ribosome assembly coming into focus.* Nat Rev Mol Cell Biol, 2019. **20**(2): p. 116-131.
5. Phipps, K.R., J. Charette, and S.J. Baserga, *The small subunit processome in ribosome biogenesis-progress and prospects.* Wiley Interdiscip Rev RNA, 2011. **2**(1): p. 1-21.
6. Kos, M. and D. Tollervey, *Yeast pre-rRNA processing and modification occur cotranscriptionally.* Mol Cell, 2010. **37**(6): p. 809-820.
7. Kister, K.P., B. Muller, and W.A. Eckert, *Complex endonucleolytic cleavage pattern during early events in the processing of pre-rRNA in the lower eukaryote, Tetrahymena thermophila.* Nucleic Acids Res, 1983. **11**(11): p. 3487-3502.
8. Hadjiolova, K.V., et al., *Alternative pre-rRNA processing pathways in human cells and their alteration by cycloheximide inhibition of protein synthesis.* Eur J Biochem, 1993. **212**(1): p. 211-215.
9. Kiss-Laszlo, Z., et al., *Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs.* Cell, 1996. **85**(7): p. 1077-1088.
10. Ni, J., A.L. Tien, and M.J. Fournier, *Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA.* Cell, 1997. **89**(4): p. 565-573.
11. Tollervey, D. and T. Kiss, *Function and synthesis of small nucleolar RNAs.* Curr Opin Cell Biol, 1997. **9**(3): p. 337-342.
12. Lafontaine, D.L. and D. Tollervey, *Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability.* RNA, 1999. **5**(3): p. 455-467.
13. Lafontaine, D.L. and D. Tollervey, *Synthesis and assembly of the box C+D small nucleolar RNPs.* Mol Cell Biol, 2000. **20**(8): p. 2650-2659.
14. Watkins, N.J., et al., *A common core RNP structure shared between the small nucleoar box C/D RNPs and the spliceosomal U4 snRNP.* Cell, 2000. **103**(3): p. 457-466.
15. Girard, J.P., et al., *Gar1 Is an Essential Small Nucleolar Rnp Protein Required for Pre-Ribosomal-Rna Processing in Yeast.* Embo Journal, 1992. **11**(2): p. 673-682.
16. Lafontaine, D.L.J., et al., *The box H+ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase.* Genes & Development, 1998. **12**(4): p. 527-537.
17. Henras, A., et al., *Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs.* EMBO J, 1998. **17**(23): p. 7078-7090.
18. Zebarjadian, Y., et al., *Point mutations in yeast CBF5 can abolish in vivo pseudouridylation of rRNA.* Mol Cell Biol, 1999. **19**(11): p. 7461-7472.
19. Jack, K., et al., *rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells.* Mol Cell, 2011. **44**(4): p. 660-666.
20. Liang, X.H., Q. Liu, and M.J. Fournier, *Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing.* RNA, 2009. **15**(9): p. 1716-1728.

21. Baudin-Baillieu, A., et al., *Nucleotide modifications in three functionally important regions of the Saccharomyces cerevisiae ribosome affect translation accuracy.* Nucleic Acids Res, 2009. **37**(22): p. 7665-7677.
22. Piekna-Przybylska, D., et al., *Ribosome performance is enhanced by a rich cluster of pseudouridines in the A-site finger region of the large subunit.* J Biol Chem, 2008. **283**(38): p. 26026-26036.
23. Liang, X.H., Q. Liu, and M.J. Fournier, *rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity.* Mol Cell, 2007. **28**(6): p. 965-977.
24. King, T.H., et al., *Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center.* Mol Cell, 2003. **11**(2): p. 425-435.
25. Tollervey, D. and C. Guthrie, *Deletion of a yeast small nuclear RNA gene impairs growth.* EMBO J, 1985. **4**(13B): p. 3873-3878.
26. Bally, M., J. Hughes, and G. Cesareni, *Snr30 - a New, Essential Small Nuclear-Rna from Saccharomyces-Cerevisiae.* Nucleic Acids Research, 1988. **16**(12): p. 5291-5303.
27. Morrissey, J.P. and D. Tollervey, *Yeast Snr30 Is a Small Nucleolar Rna Required for 18s Ribosomal-Rna Synthesis.* Molecular and Cellular Biology, 1993. **13**(4): p. 2469-2477.
28. Jinn, S., et al., *snoRNA U17 regulates cellular cholesterol trafficking.* Cell Metab, 2015. **21**(6): p. 855-867.
29. Dutca, L.M., J.E. Gallagher, and S.J. Baserga, *The initial U3 snoRNA:pre-rRNA base pairing interaction required for pre-18S rRNA folding revealed by in vivo chemical probing.* Nucleic Acids Res, 2011. **39**(12): p. 5164-5180.
30. Brink, M.F., M.P. Verbeet, and H.A. de Boer, *Formation of the central pseudoknot in 16S rRNA is essential for initiation of translation.* EMBO J, 1993. **12**(10): p. 3987-3996.
31. Hughes, J.M. and M. Ares, Jr., *Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA.* EMBO J, 1991. **10**(13): p. 4231-4239.
32. Li, H.D., J. Zagorski, and M.J. Fournier, *Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in Saccharomyces cerevisiae.* Mol Cell Biol, 1990. **10**(3): p. 1145-1152.
33. Liang, W.Q. and M.J. Fournier, *U14 base-pairs with 18S rRNA: a novel snoRNA interaction required for rRNA processing.* Genes Dev, 1995. **9**(19): p. 2433-2443.
34. Morrissey, J.P. and D. Tollervey, *U14 small nucleolar RNA makes multiple contacts with the pre-ribosomal RNA.* Chromosoma, 1997. **105**(7-8): p. 515-522.
35. Martin, R., et al., *A pre-ribosomal RNA interaction network involving snoRNAs and the Rok1 helicase.* RNA, 2014. **20**(8): p. 1173-1182.
36. Tollervey, D., *A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA.* EMBO J, 1987. **6**(13): p. 4169-4175.
37. Cheng, J., et al., *3.2-A-resolution structure of the 90S preribosome before A1 pre-rRNA cleavage.* Nat Struct Mol Biol, 2017. **24**(11): p. 954-964.
38. Barandun, J., et al., *The complete structure of the small-subunit processome.* Nat Struct Mol Biol, 2017. **24**(11): p. 944-953.
39. Chaker-Margot, M., et al., *Architecture of the yeast small subunit processome.* Science, 2017. **355**(6321).

40. Dragon, F., et al., *A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis.* Nature, 2002. **417**(6892): p. 967-970.
41. Grandi, P., et al., *90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors.* Mol Cell, 2002. **10**(1): p. 105-115.
42. Zhang, L.M., et al., *Stepwise and dynamic assembly of the earliest precursors of small ribosomal subunits in yeast.* Genes & Development, 2016. **30**(6): p. 718-732.
43. Chaker-Margot, M., et al., *Stage-specific assembly events of the 6-MDa small-subunit processome initiate eukaryotic ribosome biogenesis.* Nat Struct Mol Biol, 2015. **22**(11): p. 920-923.
44. Chaker-Margot, M. and S. Klinge, *Assembly and early maturation of large subunit precursors.* RNA, 2019. **25**(4): p. 465-471.
45. Chen, W., et al., *Stepwise assembly of the earliest precursors of large ribosomal subunits in yeast.* Nucleic Acids Res, 2017. **45**(11): p. 6837-6847.
46. Perez-Fernandez, J., et al., *The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism.* Mol Cell Biol, 2007. **27**(15): p. 5414-5429.
47. Kornprobst, M., et al., *Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome.* Cell, 2016. **166**(2): p. 380-393.
48. Cheng, J., et al., *Thermophile 90S Pre-ribosome Structures Reveal the Reverse Order of Co-transcriptional 18S rRNA Subdomain Integration.* Mol Cell, 2019. **75**(6): p. 1256-1269 e7.
49. Ursic, D., et al., *The yeast SEN1 gene is required for the processing of diverse RNA classes.* Nucleic Acids Res, 1997. **25**(23): p. 4778-4785.
50. Steinmetz, E.J. and D.A. Brow, *Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1.* Mol Cell Biol, 1996. **16**(12): p. 6993-7003.
51. Chanfreau, G., P. Legrain, and A. Jacquier, *Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism.* J Mol Biol, 1998. **284**(4): p. 975-988.
52. Mouaikel, J., et al., *Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus.* Mol Cell, 2002. **9**(4): p. 891-901.
53. Kiss, T. and W. Filipowicz, *Small nucleolar RNAs encoded by introns of the human cell cycle regulatory gene RCC1.* EMBO J, 1993. **12**(7): p. 2913-2920.
54. Pelczar, P. and W. Filipowicz, *The host gene for intronic U17 small nucleolar RNAs in mammals has no protein-coding potential and is a member of the 5'-terminal oligopyrimidine gene family.* Mol Cell Biol, 1998. **18**(8): p. 4509-4518.
55. Cecconi, F., et al., *U17XS8, a small nucleolar RNA with a 12 nt complementarity to 18S rRNA and coded by a sequence repeated in the six introns of Xenopus laevis ribosomal protein S8 gene.* Nucleic Acids Res, 1994. **22**(5): p. 732-741.
56. Kiss, T. and W. Filipowicz, *Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns.* Genes Dev, 1995. **9**(11): p. 1411-1424.
57. Cecconi, F., P. Mariottini, and F. Amaldi, *The Xenopus intron-encoded U17 snoRNA is produced by exonucleolytic processing of its precursor in oocytes.* Nucleic Acids Res, 1995. **23**(22): p. 4670-4676.

58. Szczepinska, T., et al., *DIS3 shapes the RNA polymerase II transcriptome in humans by degrading a variety of unwanted transcripts.* Genome Res, 2015. **25**(11): p. 1622-1633.
59. Caffarelli, E., et al., *Processing of the intron-encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA.* EMBO J, 1996. **15**(5): p. 1121-1131.
60. Kufel, J. and P. Grzechnik, *Small Nucleolar RNAs Tell a Different Tale.* Trends Genet, 2019. **35**(2): p. 104-117.
61. Yang, P.K., et al., *The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis.* J Biol Chem, 2002. **277**(47): p. 45235-45242.
62. Walbott, H., et al., *The H/ACA RNP assembly factor SHQ1 functions as an RNA mimic.* Genes Dev, 2011. **25**(22): p. 2398-2408.
63. Caton, E.A., et al., *Efficient RNA pseudouridylation by eukaryotic H/ACA ribonucleoproteins requires high affinity binding and correct positioning of guide RNA.* Nucleic Acids Res, 2018. **46**(2): p. 905-916.
64. Massenet, S., E. Bertrand, and C. Verheggen, *Assembly and trafficking of box C/D and H/ACA snoRNPs.* RNA Biol, 2017. **14**(6): p. 680-692.
65. Machado-Pinilla, R., et al., *Mechanism of the AAA+ ATPases pontin and reptin in the biogenesis of H/ACA RNPs.* RNA, 2012. **18**(10): p. 1833-1845.
66. Darzacq, X., et al., *Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells.* J Cell Biol, 2006. **173**(2): p. 207-218.
67. Grozdanov, P.N., et al., *SHQ1 is required prior to NAF1 for assembly of H/ACA small nucleolar and telomerase RNPs.* RNA, 2009. **15**(6): p. 1188-1197.
68. Narayanan, A., et al., *Nucleolar localization signals of box H/ACA small nucleolar RNAs.* EMBO J, 1999. **18**(18): p. 5120-5130.
69. Lange, T.S., et al., *Box H and box ACA are nucleolar localization elements of U17 small nucleolar RNA.* Mol Biol Cell, 1999. **10**(11): p. 3877-3890.
70. Dragon, F., V. Pogacic, and W. Filipowicz, *In vitro assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs.* Mol Cell Biol, 2000. **20**(9): p. 3037-3048.
71. Samarsky, D.A., et al., *The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization.* EMBO J, 1998. **17**(13): p. 3747-3757.
72. Narayanan, A., et al., *Role of the box C/D motif in localization of small nucleolar RNAs to coiled bodies and nucleoli.* Mol Biol Cell, 1999. **10**(7): p. 2131-2147.
73. Kiss, T., et al., *Biogenesis and intranuclear trafficking of human box C/D and H/ACA RNPs.* Cold Spring Harb Symp Quant Biol, 2006. **71**: p. 407-417.
74. Meier, U.T., *RNA modification in Cajal bodies.* RNA Biol, 2017. **14**(6): p. 693-700.
75. Atzorn, V., P. Fragapane, and T. Kiss, *U17/snR30 is a ubiquitous snoRNA with two conserved sequence motifs essential for 18S rRNA production.* Mol Cell Biol, 2004. **24**(4): p. 1769-1778.
76. Ganot, P., M.L. Bortolin, and T. Kiss, *Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs.* Cell, 1997. **89**(5): p. 799-809.
77. Cervelli, M., et al., *Structural and sequence evolution of U17 small nucleolar RNA (snoRNA) and its phylogenetic congruence in chelonians.* J Mol Evol, 2003. **57**(1): p. 73-84.
78. Hughes, J.M., D.A. Konings, and G. Cesareni, *The yeast homologue of U3 snRNA.* EMBO J, 1987. **6**(7): p. 2145-2155.

79. Ruhl, D.D., M.E. Pusateri, and G.L. Eliceiri, *Multiple conserved segments of E1 small nucleolar RNA are involved in the formation of a ribonucleoprotein particle in frog oocytes.* Biochem J, 2000. **348 Pt 3**: p. 517-524.
80. Cervelli, M., et al., *Comparative structure analysis of vertebrate U17 small nucleolar RNA (snoRNA).* J Mol Evol, 2002. **54**(2): p. 166-179.
81. Deng, W., et al., *Organization of the Caenorhabditis elegans small non-coding transcriptome: genomic features, biogenesis, and expression.* Genome Res, 2006. **16**(1): p. 20-29.
82. Nawrocki, E.P., et al., *Rfam 12.0: updates to the RNA families database.* Nucleic Acids Res, 2015. **43**(Database issue): p. D130-137.
83. Cecconi, F., et al., *A functional role for some Fugu introns larger than the typical short ones: the example of the gene coding for ribosomal protein S7 and snoRNA U17.* Nucleic Acids Res, 1996. **24**(16): p. 3167-3172.
84. Lemay, V., et al., *Identification of novel proteins associated with yeast snR30 small nucleolar RNA.* Nucleic Acids Research, 2011. **39**(22): p. 9659-9670.
85. Fayet-Lebaron, E., et al., *18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences.* Embo Journal, 2009. **28**(9): p. 1260-1270.
86. Cheng, J., et al., *90S pre-ribosome transformation into the primordial 40S subunit.* Science, 2020. **369**(6510): p. 1470-1476.
87. Spahn, C.M., et al., *Structure of the 80S ribosome from Saccharomyces cerevisiae--tRNA-ribosome and subunit-subunit interactions.* Cell, 2001. **107**(3): p. 373-386.
88. Wells, G.R., et al., *The ribosome biogenesis factor yUtp23/hUTP23 coordinates key interactions in the yeast and human pre-40S particle and hUTP23 contains an essential PIN domain.* Nucleic Acids Research, 2017. **45**(8): p. 4796-4809.
89. Wells, G.R., et al., *The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans.* Nucleic Acids Res, 2016. **44**(11): p. 5399-5409.
90. Horn, D.M., S.L. Mason, and K. Karbstein, *Rcl1 protein, a novel nuclease for 18 S ribosomal RNA production.* J Biol Chem, 2011. **286**(39): p. 34082-34087.
91. Lubben, B., et al., *Isolation and Characterization of the Small Nucleolar Ribonucleoprotein Particle Snr30 from Saccharomyces-Cerevisiae.* Journal of Biological Chemistry, 1995. **270**(19): p. 11549-11554.
92. Liang, X.H. and M.J. Fournier, *The helicase Has1p is required for snoRNA release from pre-rRNA.* Molecular and Cellular Biology, 2006. **26**(20): p. 7437-7450.
93. Bohnsack, M.T., M. Kos, and D. Tollervey, *Quantitative analysis of snoRNA association with pre-ribosomes and release of snR30 by Rok1 helicase.* Embo Reports, 2008. **9**(12): p. 1230-1236.
94. Hoareau-Aveilla, C., et al., *Utp23p is required for dissociation of snR30 small nucleolar RNP from preribosomal particles.* Nucleic Acids Research, 2012. **40**(8): p. 3641-3652.
95. Lu, J., M.Y. Sun, and K.Q. Ye, *Structural and functional analysis of Utp23, a yeast ribosome synthesis factor with degenerate PIN domain.* Rna-a Publication of the Rna Society, 2013. **19**(12): p. 1815-1824.
96. Barandun, J., M. Hunziker, and S. Klinge, *Assembly and structure of the SSU processome-a nucleolar precursor of the small ribosomal subunit.* Curr Opin Struct Biol, 2018. **49**: p. 85-93.
97. Duan, J., et al., *Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase.* Mol Cell, 2009. **34**(4): p. 427-439.

98. Lebaron, S., et al., *Rrp5 Binding at Multiple Sites Coordinates Pre-rRNA Processing and Assembly.* Molecular Cell, 2013. **52**(5): p. 707-719.
99. Venema, J. and D. Tollervey, *RRP5 is required for formation of both 18S and 5.85 rRNA in yeast.* Embo Journal, 1996. **15**(20): p. 5701-5714.
100. Khoshnevis, S., et al., *Rrp5 establishes a checkpoint for 60S assembly during 40S maturation.* RNA, 2019. **25**(9): p. 1164-1176.
101. Khoshnevis, S., et al., *The DEAD-box Protein Rok1 Orchestrates 40S and 60S Ribosome Assembly by Promoting the Release of Rrp5 from Pre-40S Ribosomes to Allow for 60S Maturation.* Plos Biology, 2016. **14**(6).
102. Watkins, N.J., et al., *Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure.* RNA, 1998. **4**(12): p. 1549-1568.
103. Lafontaine, D.L. and D. Tollervey, *Birth of the snoRNPs: the evolution of the modification-guide snoRNAs.* Trends Biochem Sci, 1998. **23**(10): p. 383-388.
104. Li, S., et al., *Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase.* Genes Dev, 2011. **25**(22): p. 2409-2421.
105. Mullis, K., et al., *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.* Cold Spring Harb Symp Quant Biol, 1986. **51 Pt 1**: p. 263-273.
106. Wright, J.R., et al., *Pre-steady-state kinetic analysis of the three Escherichia coli pseudouridine synthases TruB, TruA, and RluA reveals uniformly slow catalysis.* RNA, 2011. **17**(12): p. 2074-2084.
107. Conrad, T., et al., *Maximizing transcription of nucleic acids with efficient T7 promoters.* Commun Biol, 2020. **3**(1): p. 439.
108. Koo, B.K., et al., *Structure of H/ACA RNP protein Nhp2p reveals cis/trans isomerization of a conserved proline at the RNA and Nop10 binding interface.* J Mol Biol, 2011. **411**(5): p. 927-942.
109. Koonin, E.V., P. Bork, and C. Sander, *A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme?* Nucleic Acids Res, 1994. **22**(11): p. 2166-2167.
110. Huang, L. and D.M.J. Lilley, *The Kink Turn, a Key Architectural Element in RNA Structure.* J Mol Biol, 2016. **428**(5 Pt A): p. 790-801.
111. Vera, A., et al., *The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures.* Biotechnol Bioeng, 2007. **96**(6): p. 1101-1106.
112. Rosano, G.L. and E.A. Ceccarelli, *Recombinant protein expression in Escherichia coli: advances and challenges.* Front Microbiol, 2014. **5**: p. 172.
113. Yang, K., et al., *A redox mechanism underlying nucleolar stress sensing by nucleophosmin.* Nat Commun, 2016. **7**: p. 13599.
114. Rial, D.V. and E.A. Ceccarelli, *Removal of DnaK contamination during fusion protein purifications.* Protein Expr Purif, 2002. **25**(3): p. 503-507.
115. Rohman, M. and K.J. Harrison-Lavoie, *Separation of copurifying GroEL from glutathione-S-transferase fusion proteins.* Protein Expr Purif, 2000. **20**(1): p. 45-47.
116. Albuquerque, C.P., et al., *A multidimensional chromatography technology for in-depth phosphoproteome analysis.* Mol Cell Proteomics, 2008. **7**(7): p. 1389-1396.
117. Ryu, H., et al., *The deubiquitinase USP36 promotes snoRNP group SUMOylation and is essential for ribosome biogenesis.* EMBO Rep, 2021: p. e50684.

118. Kolodrubetz, D. and A. Burgum, *Sequence and genetic analysis of NHP2: a moderately abundant high mobility group-like nuclear protein with an essential function in Saccharomyces cerevisiae.* Yeast, 1991. **7**(2): p. 79-90.
119. Li, L. and K. Ye, *Crystal structure of an H/ACA box ribonucleoprotein particle.* Nature, 2006. **443**(7109): p. 302-307.
120. Cheng, Y., *Single-Particle Cryo-EM at Crystallographic Resolution.* Cell, 2015. **161**(3): p. 450-457.
121. Some, D., et al., *Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS).* J Vis Exp, 2019(148).
122. Tarassov, K., et al., *An in vivo map of the yeast protein interactome.* Science, 2008. **320**(5882): p. 1465-1470.
123. Pelletier, J., G. Thomas, and S. Volarevic, *Ribosome biogenesis in cancer: new players and therapeutic avenues.* Nat Rev Cancer, 2018. **18**(1): p. 51-63.
124. Langhendries, J.L., et al., *The human box C/D snoRNAs U3 and U8 are required for pre-rRNA processing and tumorigenesis.* Oncotarget, 2016. **7**(37): p. 59519-59534.

Appendix 1

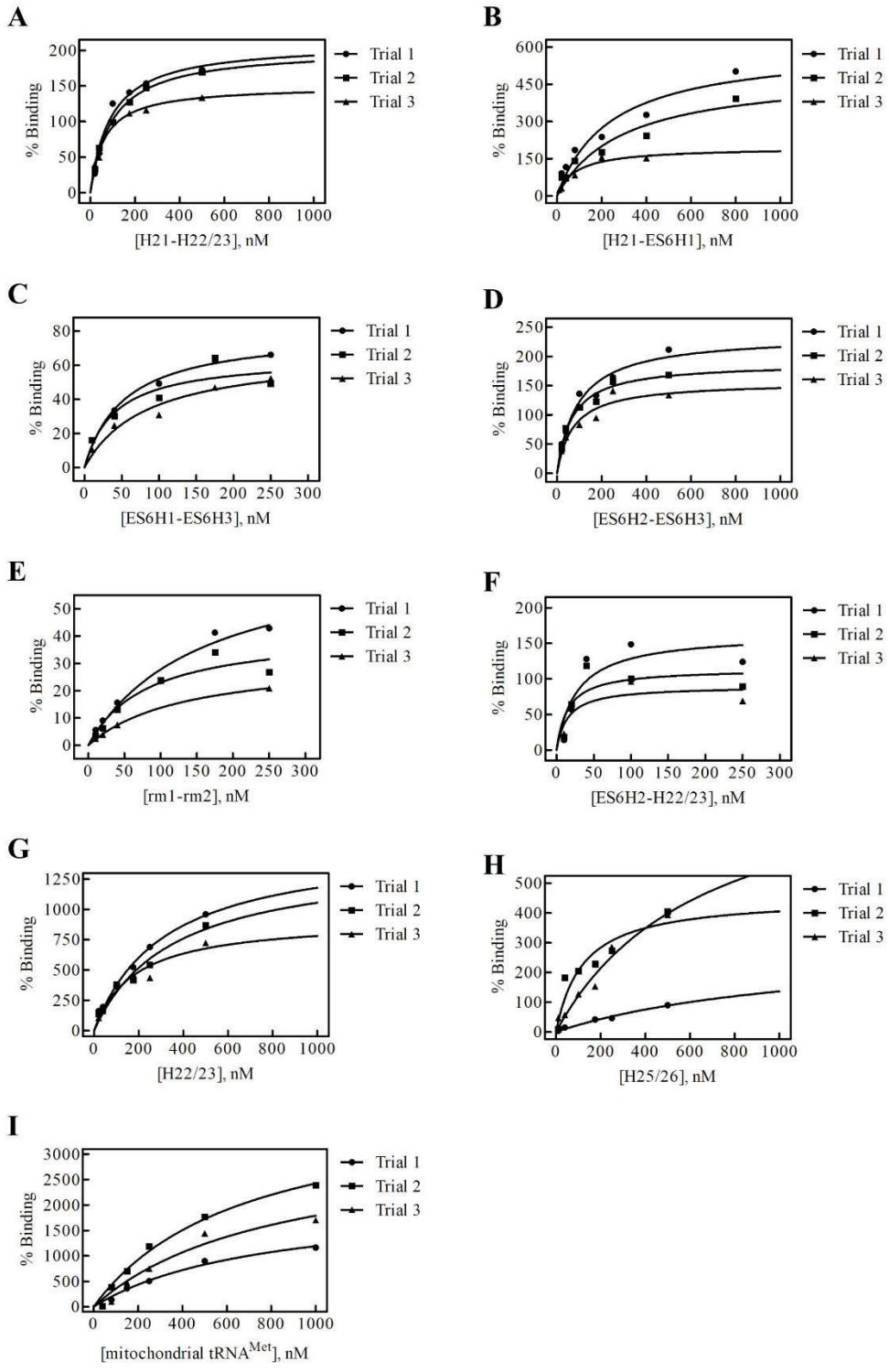

Figure A.1. All replicates of full-length snR30 complexed with Cbf5-Nop10-Gar1-Nhp2 Binding to 18S rRNA Constructs. (**A-I**) Binding of the respective RNA (n=3), data is fit with equation 2.1. Dissociation constants are displayed in Table A.1.

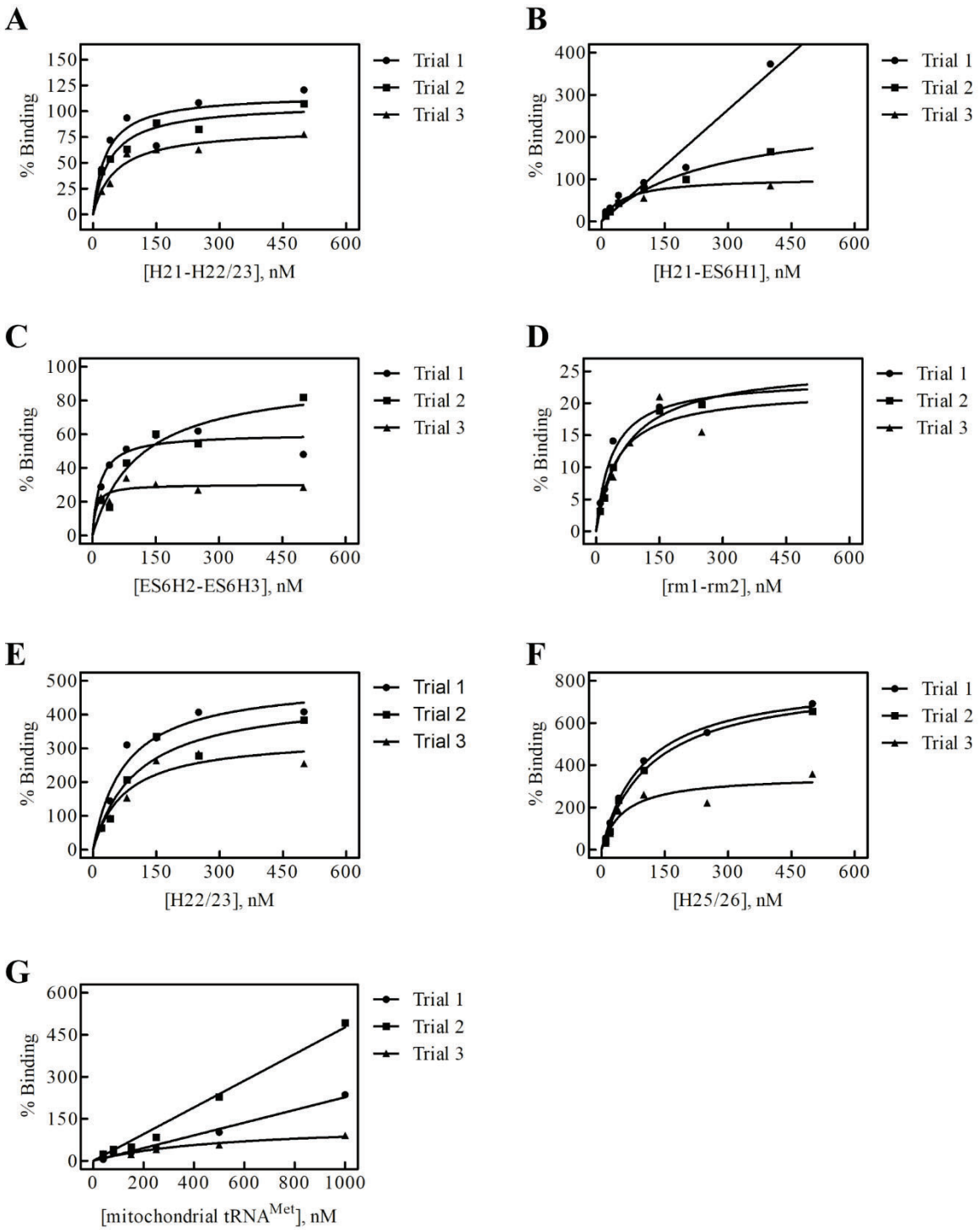

Figure A.2. Δ397 snR30 complexed with Cbf5-Nop10-Gar1-Nhp2 binding to 18S rRNA Constructs. (**A-G**) Binding of the respective RNA (n=3), data is fit with equation 2.1. Dissociation constants are displayed in Table A.1.

Table A.1. Averaged interaction affinity of the replicates between 18S rRNA constructs and the snR30 complex.

snR30 Variant	18S rRNA Construct	Dissociation constant (nM)	Binding Amplitude (% Binding)
Full-length	H21-H22/23	84 ± 21	187 ± 16
	H21-ES6H1	207 ± 149	431 ± 107
	ES6H1-ES6H3	60 ± 36	71 ± 12
	ES6H2-ES6H3	77 ± 27	193 ± 23
	rm1-rm2	125 ± 59	49 ± 13
	ES6H2-H22/23	19 ± 21	123 ± 38
	H22/23	282 ± 141	1300 ± 314
	H25/26	652 ± 762	560 ± 343
	Mitochondrial $tRNA^{Met}$	758 ± 407	3147 ± 925
Δ397	H21-H22/23	42 ± 17	102 ± 15
	H21-ES6H1	156 ± 84*	184 ± 44*
	ES6H1-ES6H3	ND	ND
	ES6H2-ES6H3	46 ± 48	62 ± 15
	rm1-rm2	51 ± 27	24 ± 4
	ES6H2-H22/23	ND	ND
	H22/23	85 ± 45	432 ± 72
	H25/26	90 ± 26	661 ± 55
	Mitochondrial $tRNA^{Met}$	549 ± 325**	133 ± 40**

ND means the affinity was not determined for the RNA. NC means the affinity could not be calculated. *Only two trials had data that could be averaged. **Only one trial had a determinable K_D.

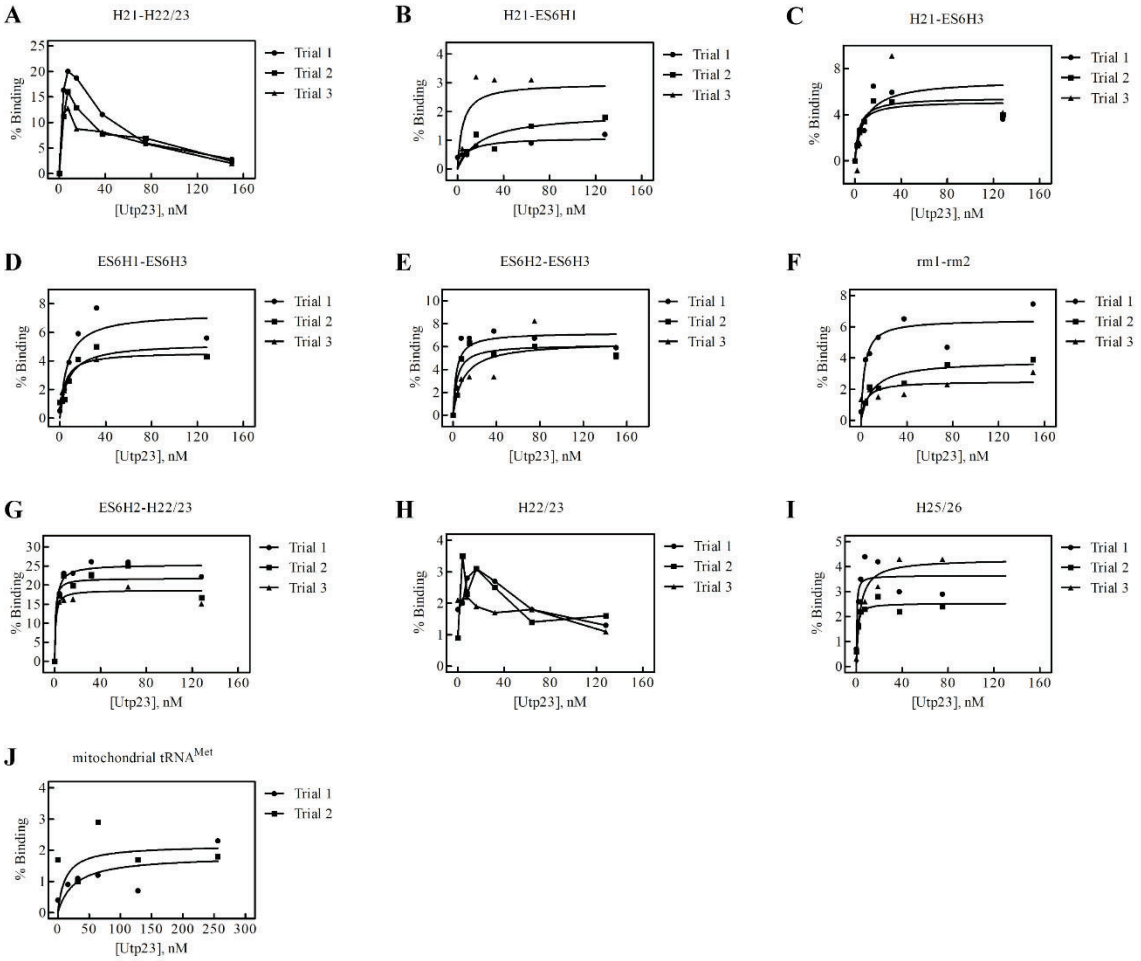

Figure A.3. Utp23 nitrocellulose filter binding to 18S rRNA constructs. (**A, H**) Binding to H21-H22/23 and H22/23 n=3, data points are connected without fitting. (**B-G, I**) Binding to the respective RNAs n=3, data is fit with equation 2.1. (**J**) Binding to mitochondrial $tRNA^{Met}$ (n=2), data is fit with equation 2.1. Dissociation constants are displayed in Table A.2

Table A.2. Averaged affinity of the replicates between 18S rRNA fragments and Utp23.

18S rRNA Construct	Dissociation constant (nM)	Binding Amplitude (%)
H21-H22/23	NC	NC
H21-ES6H1	11 ± 16	2 ± 1
H21-ES6H3	6 ± 11	6 ± 3
ES6H1-ES6H3	6 ± 3	6 ± 1
ES6H2-ES6H3	6 ± 8	6 ± 1
rm1-rm2	6 ± 5	6 ± 1
ES6H2-H22/23	< 2 ± 1*	22 ± 2
H22/23	NC	NC
H25/26	< 2 ± 1*	3 ± 1
mitochondrial $tRNA^{Met}$	19 ± 49**	2 ± 1**

NC means the affinity could not be calculated. *Affinity is below the concentration of RNA used in the reaction. **Only two replicates were done.