KINETIC ANALYSIS OF PSEUDOURIDINE FORMATION IN TRNA BY ARCHAEAL CBF5 AND PUS10

RAJASHEKHAR KAMALAMPETA
M.Sc. University of Skövde, 2008

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

DOCTOR OF PHILOSOPHY

IN

BIOMOLECULAR SCIENCE

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

Dedicated to

My parents for their unconditional love and support

&

Those passionate teachers for instilling my interest in science

Abstract

Pseudouridines (ψ), the most common modifications in RNA, are formed by stand-alone ψ synthases in all organisms. In addition, archaea and eukaryotes use H/ACA small ribonucleoproteins for pseudouridylation. Cbf5, the catalytic component of these complexes, can also introduce ψ55 in archaeal tRNAs in a guide RNA-independent manner. Here, kinetic and thermodynamic analyses revealed that both *Pyrococcus furiosus* Nop10 and Gar1 proteins enhance the catalytic ability of Cbf5 and increase its affinity for tRNA. Pus10, representing a novel ψ synthase family, is the *in vivo* archaeal tRNA ψ55 synthase. Characterization of several Pus10 variants demonstrated the importance of the thumb loop for catalysis, a potential role of the THUMP domain in tRNA binding and a new catalytic arginine which may flip the target uridine into Pus10's active site. The quantitative characterization of the archaeal pseudouridine synthases Cbf5 and Pus10 reported here sheds light on their cellular roles in RNA modification.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to Dr. Ute Kothe, my supervisor, for her invaluable guidance, mentorship, encouragement, and constructive suggestions throughout my Ph.D.

I wish to thank Dr. Marc Roussel, my co-supervisor, for his helpful advice and constant support. I would also like to sincerely thank my supervisory committee members Drs. Steven Mosimann and Elizabeth Schultz for their valuable suggestions and critical feedback.

I would like to extend special thanks to Dr. Hans-Joachim Wieden for his supervision in the absence of my supervisor and for his critical advice on my projects. I would also like to extend my thanks to Dr. George Owttrim for accepting to be the external examiner, and to Dr. Paul Hayes for chairing the examination committee.

I am grateful to all Kothe and Wieden lab members for making this journey a great learning experience. Special appreciation goes to Laura, Jenna, Evan Caton, Jessica, Jaden, Roxanne, Lisza, Alvin, Evan, Evelina, Jeff, Marion, Tobias, Fan, Kirsten, Mike, Lindsay and many other wonderful people for their critical discussions and continuous help during my thesis. I must also thank Andy for all his helpful conversations.

I would also like to acknowledge all those brilliant students, Aruna, Chaitra, Christian, Dominik, Justin, Jenna, and Vanessa for assisting me in the preparatory work and for giving me a great opportunity to supervise them. Also thanks are due to the previous Kothe lab members who contributed to this project.

Special thanks goes to Susan and many people from the School of Graduate Studies for their help.

Finally, I wish to thank my family and friends for their support and encouragement. I must not forget to thank my wife Malini for being there during all ups and downs and for contributing immensely to my success.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ψ	Pseudouridine
ATP	Adenosine-5'-triphosphate
BSA	Bovine serum albumin
Cbf5	Centromere binding factor 5
CTP	Cytidine-5'-triphosphate
DEAE	Diethylaminoethyl
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
Gar1	Glycine and arginine rich protein 1
GTP	Guanosine-5'-triphosphate
His-tag	Histidine-tag
HEPES	N-2-Hydroxyethylpiperazine-N'-2-Ethanesulfonic acid
iPPase	inorganic pyrophosphatase
IPTG	Isopropyl-β-D-galactopyranoside
Kan	Kanamycin
LB	Luria-Bertani
miRNA	Micro RNA
mRNA	Messenger RNA
MWCO	Molecular weight cutoff
Nop10	Nucleolar protein 10
NTP	Nucleotide-5'-triphosphate
OD_{600}	Optical Density at 600 nm
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB ID	Protein data bank identification code
PMSF	Phenylmethylsulfonylfluoride
PTC	Peptidyl transferase centre
PUA domain	PseudoUridine synthase and Archaeosine-transglycosylase domain
Pus1-10	Pseudouridine synthases 1-10
RF	Release factor
RluA	Ribosomal large subunit pseudouridine synthase A
RNA	Ribonucleic acid
RNase	Ribonuclease
rRNA	Ribosomal RNA
RsuA	Ribosomal small subunit pseudouridine synthase A
SDS	Sodium dodecyl sulphate
siRNA	Small interfering RNA
snoRNA	Small nucleolar RNA
snoRNP	Small nucleolar ribonucleoprotein

List of Abbreviations (continued)

snRNA	Small nuclear RNA
sRNP	Small ribonucleoprotein
TERC	Telomerase RNA component
TERT	Telomerase Reverse Transcriptase
THUMP	Thiouridine synthases, Methylases, and Pseudouridine synthases
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
TruA-D	tRNA pseudouridine synthases A-D
U1-U6 snRNAs	Uridyl-rich snRNAs
UTP	Uridine-5'-triphosphate

Chapter 1: Introduction

1.1 Ribosome biogenesis

Proteins are indispensable for the survival of living cells and are synthesized by macromolecular complexes known as ribosomes. Ribosomes are composed of both RNA (~ 60%) and protein (~ 40%). Formation of these macromolecular complexes, a process known as ribosome biogenesis, is highly complex and involves synthesis, processing, modification of ribosomal RNA (rRNA) and its assembly with proteins. Ribosomes in all domains of life are composed of two subunits, a large and a small subunit. Although the function of ribosomes in prokaryotes and eukaryotes is conserved, there exist differences in their size and complexity. While prokaryotes contain 70S ribosomes composed of a 30S small subunit and a 50S large subunit, eukaryotes carry 80S ribosomes made up of 40S and 60S subunits, respectively. Also both the rRNA and protein composition of the subunits is slightly different – the prokaryotic 30S subunits contain 16S rRNA and about 20 ribosomal proteins, while the eukaryotic 40S subunit consists of 18S rRNA and ~ 30 ribosomal proteins. The prokaryotic 50S subunit comprises two rRNA components (5S and 23S) and ~ 30 ribosomal proteins, and the eukaryotic 60S subunits has three rRNAs (5S, 5.8S and 25S/28S) along with ~ 45 ribosomal proteins (Wool 1979).

Although there are differences in the primary structure of rRNAs, the secondary structures display significant conservation between prokaryotes and eukaryotes (Lee and Gutell 2012). For obvious reasons, all parts of prokaryotic ribosome assembly takes place in the same cellular compartment, whereas in eukaryotes, this process occurs in the nucleolus for the most part, but also in the cytoplasm (Hadjiolov 1985).

1.1.1 Organization of genes encoding ribosomal RNA and proteins

In bacteria, genes encoding rRNA species are organized into operons (referred to as *rrn*). Seven such transcriptional units are found in *Escherichia coli* (Kiss *et al.* 1977). Co-transcription of these operons results in 30S pre-transcripts consisting of the 16S, 23S, and 5S rRNAs (Schlessinger 1980). Spacer regions located in these operons carry genes coding for a few transfer RNA (tRNA) molecules. In contrast, eukaryotic rRNA genes are arranged into arrays of several tandem repeats (150 – 200 in yeast) each coding for a 35S pre-transcript containing 5.8S, 18S, and 25S/28S rRNA. Between 150 and 200 repeats are found on chromosome XII in yeast, accounting for about 10% of the yeast genome (Merz *et al.* 2008). Humans carry about 400 rRNA genes distributed into tens of tandem repeats located on chromosomes 13, 14, 15, 21, and 22 (Raska *et al.* 2004). In most eukaryotes, ribosomal genes for 5S rRNA are located separately in large clusters in the nucleus (Nazar 2004). A high copy number of rRNA genes ensures that the synthesis of ribosomes meets the demand for intensive protein synthesis in dividing cells.

In addition to the internal spacer sequences separating each rRNA gene, eukaryotes also contain both external transcribed and non-transcribed spacer elements (Sylvester *et al.* 2004). Variations in the transcribed spacer sequences contribute to the size and composition differences observed in eukaryotic pre-RNA sequences. The secondary structures formed by internal spacer sequences appear to be important for interaction with the trans-acting factors and in defining the accuracy, efficiency and the order of processing steps (Nazar 2004).

In prokaryotes, genes coding for ribosomal proteins are found in genomic clusters that include several partially conserved operons (Wang *et al.* 2009). While in eukaryotes, ribosomal protein genes are scattered throughout the genome (Kenmochi *et al.* 1998), they are found to be coordinately expressed (Li *et al.* 2005). Bioinformatic analysis revealed that there are 34 ribosomal proteins that are conserved in all domains of life, while 33 are conserved only in archaea and eukaryotes; besides these, 1 ribosomal protein specific to archaea, 23 specific to bacteria and 11 specific to eukaryotes are observed (Lecompte *et al.* 2002).

1.1.2. Eukaryotic rRNA processing and ribosome assembly

In eukaryotes, ribosome biogenesis begins with the transcription of the long pre-rRNA by RNA polymerase I in the nucleolus, except for 5S rRNA which is transcribed separately by RNA polymerase III in another area of the nucleus (Figure 1.1). Upon synthesis, 5S rRNA is transported to the nucleolus as a ribonucleoprotein complex (Nazar 2004). With the involvement of RNA polymerase II which transcribes the messenger RNAs (mRNAs) required for the synthesis of ribosomal proteins and accessory factors, a significant portion of the transcription machinery is dedicated to the synthesis of ribosomes in rapidly growing cells (Warner 1999).

Either co-transcriptionally or immediately after transcription, pre-rRNA molecules undergo extensive modifications guided by non-coding guide RNAs called small nucleolar RNA (snoRNA). The majority of the modifications include 2'-O methylations of ribose and isomerization of uridine to pseudouridine guided by C/D box and H/ACA box guide RNAs, respectively. Over a hundred guide RNA molecules direct the small

nucleolar ribonucleoprotein (snoRNP) complexes to the specific sites in the pre-rRNA for modification (for more details refer to section 1.3.2). Initially, a 90S pre-ribosome complex is formed in eukaryotes from precursor rRNA transcripts and several ribosomal and non-ribosomal proteins (Henras *et al.* 2008). Next, a number of site-specific pre-rRNA cleavage steps are carried out by the processing snoRNPs guided by C/D snoRNAs U3, U8, U14 and U22, and H/ACA snoRNAs snR10 and snR30 (U17 in mammals) (Tollervey and Kiss 1997). Following these steps, further processing occurs by the action of exo- and endo-nucleases, which remove the spacer sequences and release the 18S rRNA from the pre-transcript. Upon multiple cleavages, these steps result in pre-40S and pre-60S ribosomal subunits. Several other intermediates are formed during the processing, each with a different subset of proteins.

Together with nucleases and snoRNPs, there are about 200 accessory factors participating in ribosome assembly, which comprise putative ATP-dependent RNA helicases, chaperones, GTPases, and export factors (Henras *et al.* 2008). RNA helicases are predicted to be involved in restructuring the RNA, facilitating the ribosomal protein-RNA interaction and mediating the remodeling of pre-ribosomal complexes (Martin *et al.* 2013). Pre-40S and pre-60S complexes are transported from the nucleus to the cytoplasm for the final maturation steps (Figure 1.1). This unidirectional transport occurs through the nuclear pore complexes with the help of GTPases (Ran or Gsp1p in yeast) and several export receptors (Pemberton and Paschal 2005). Finally, further processing of 18S rRNA and the replacement of non-ribosomal transport proteins with ribosomal proteins leads to the formation of mature and translation-competent ribosomes in the cytoplasm (Henras *et al.* 2008).

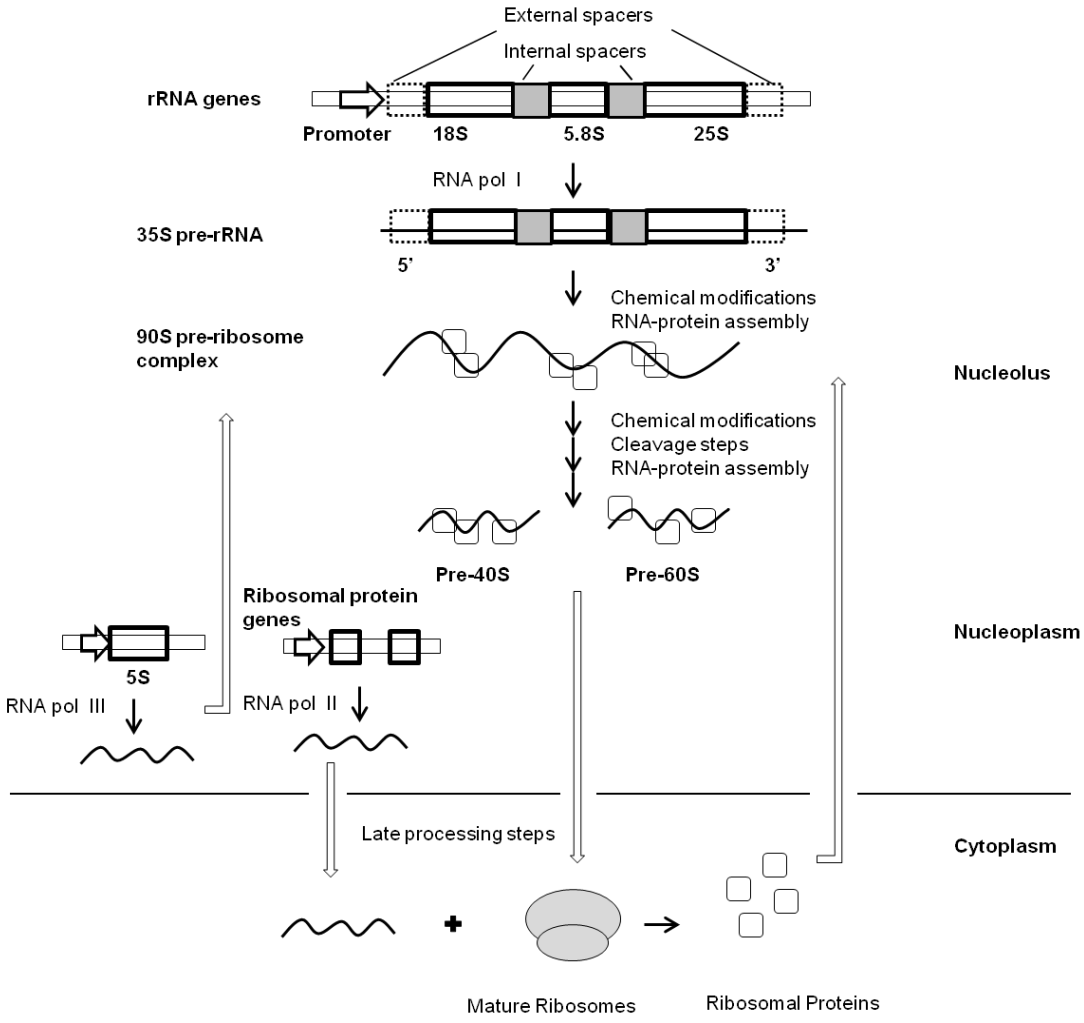


Figure 1.1. Summary of ribosome biogenesis in eukaryotes

The long pre-rRNA transcript containing all the rRNAs, except 5S rRNA, is transcribed from rRNA gene arrays in the nucleolus. 5S rRNA transcription occurs separately in the nucleus. mRNA transcribed from the ribosomal genes in the nucleus is transported to the cytoplasm, resulting in the synthesis of ribosomal proteins which are then transported back into the nucleolus. Together, all four rRNA species and ribosomal proteins form an initial 90S pre-ribosome complex. Several RNA modifications, cleavage events, and protein assembly steps take place to form the pre-40S and pre-60S ribosome complexes. Following the remodelling of pre-ribosomal complexes, they are transported to the cytoplasm for final maturation steps which lead to the formation of mature ribosomes.

Ribosome biogenesis is a highly complicated process with several processing steps facilitated by a number of factors, while many proteins responsible for certain processing

steps are unknown. Given the dynamic and intricate nature of the steps involved in ribosome synthesis, extensive coordination and tight regulation is required between rRNA processing and ribosome assembly for maintaining the accuracy in ribosome formation which otherwise leads to undesired consequences such as cancer (Montanaro *et al.* 2008). In addition, to efficiently carry out this energy demanding task, eukaryotes also exhibit spatio-temporal ordering of the various events involved in ribosome synthesis in the nucleolus, nucleoplasm and cytoplasm (Fromont-Racine *et al.* 2003). Faulty ribosomes are found to be eliminated by polyadenylation followed by selective degradation by the exosome (Dez *et al.* 2006).

1.2 RNA modifications

1.2.1 Brief introduction to RNA

RNA, besides DNA and proteins, is an essential biomolecule found in all living cells. It replaces DNA in several viruses, acting as the genetic material. Unlike DNA, RNA is typically single-stranded and composed of ribonucleotides containing adenine, guanine, and cytosine as in DNA, but uracil instead of thymine. Although single stranded in nature, RNA can assume various secondary structures through intra-strand base pairing and forms complex tertiary structures, which are important for its function (Hermann and Patel 1999). In the cell, there are several different types of RNA, most importantly rRNA, mRNA, and tRNA. As described in the previous section, rRNA is the major and an essential component of the ribosome, which functions as the protein synthesizing machinery. mRNA is involved in transferring the genetic information from DNA to

proteins, while tRNA works as a molecular adaptor that binds to mRNA, delivering amino acids to the site of protein synthesis.

Human genome sequencing revealed that less than 2% of the cellular DNA is protein-coding. However, most of the remaining DNA is predicted to actively undergo transcription, generating non-coding RNAs (Mattick 2001; Alexander *et al.* 2010). In addition to tRNA and rRNA, several non-coding RNAs were discovered with diverse functional abilities. Some of the well studied non-coding RNAs include microRNA (miRNA) involved in the regulation of gene expression, small nuclear RNA (snRNA) in splicing, small interfering RNA (siRNA) in gene silencing, and snoRNA in RNA processing and maturation (Mattick and Makunin 2006). Several other non-coding RNAs are emerging, expanding their repertoire of biological functions. Bacteria and archaea also contain non-coding RNAs, but they account only for a small fraction of the genomes, which correlates with the low complexity of these organisms (Mattick 2001, Costa 2008).

1.2.2 Modifications in RNA and their significance

'RNA modification' refers to a chemical change applied to the existing nucleotide in RNA and does not include variations that lead to the alteration of genetic meaning (Grosjean 2005). In the literature, sometimes, 'RNA editing', which describes alterations such as base substitutions and deletions, is synonymously used with 'RNA modification'. Following transcription, most non-coding RNAs undergo chemical modifications. These chemical modifications together with RNA editing and other more complex processing steps lead to the maturation of RNA, rendering it functional by attributing to it new structural and chemical properties (Grosjean 2005). The RNA modification database

reports 109 different modified nucleotides in RNA with known chemical structures (http://rna-mdb.cas.albany.edu/RNAmods/rnaover.htm). tRNA displays the most diverse modifications among any known RNA species with 93 different modifications, rRNA being the second most commonly modified RNA containing 31 modifications, while mRNA provides 13 additional cases and 14 modifications were found in other RNA classes such as snRNAs (Cantara *et al.* 2011). Besides the most prevalent modifications (2'-O-methylations and uridine to pseudouridine conversions), RNA contains other modifications such as base methylation, base thiolation, base reduction, certain hypermodifications (eg. s^2m^5U), etc. Both the abundant ribose-methylations and pseudouridine formations are introduced by either site-specific enzymes (specific to one or more sites) called methylases and pseudouridine synthases, respectively, or by RNA-guided snoRNP (or just small RNP (sRNP) in archaea) complexes. While only stand-alone enzymes are found in bacteria, both archaea and eukarya in addition use C/D box RNAs to site-specifically guide methylation and H/ACA box RNAs for pseudouridine formation.

RNA modification is a universal phenomenon displayed in all known RNAs (Lapeyre 2005). Interestingly, there seems to be a correlation between the number of modifications and the complexity of the organism. For instance, yeast mitochondrial ribosomes, which synthesize about 10 proteins, contain 3 modifications in their mitochondrial rRNA, compared to over 200 modifications found in rRNA of metazoan ribosomes that synthesize thousands of proteins (Lapeyre 2005). The observation that several modifications and their sequence locations are conserved, suggests important roles for the RNA modifications in the cell. Moreover, modifications mapped onto the three-

dimensional structure of the ribosome indicated their selective occurrence in functionally important regions (Decatur and Fournier 2002; Omer *et al.* 2003), further supporting their importance. Surprisingly, preventing the formation of individual modifications in functionally important regions either by knocking out the responsible guide RNA or by mutating the enzyme involved in the modification resulted in no or only slight defects. But, the absence of these modifications in combination with the removal of modifications at other sites displayed effects on growth rate, ribosome formation and translation efficiency (King *et al.* 2003; Liang *et al.* 2007; Liang *et al.* 2009). These observations led to the hypothesis that the modifications may act in a cooperative manner.

tRNA displays a clover-leaf secondary structure composed of acceptor stem and T-, anticodon-, and D-arms that in turn forms the L-shaped tertiary structure (Figure 1.2). The names D- and T-arms in fact originate from the preserved modifications in their loop regions, dihydrouridine (D) and ribothymidine (T), respectively. The T-arm is also commonly known as the TψC-arm due to the presence of a universally conserved pseudouridine (ψ) at position 55 followed by cytidine in the corresponding loop region. Nucleotide modifications found in tRNA stabilize its tertiary structure and prevent it from premature degradation (Motorin and Helm 2010). Especially, modifications found in the anticodon loop at position 34 and the nearby position 37 are essential for the decoding function of tRNA (Satoh *et al.* 2000).

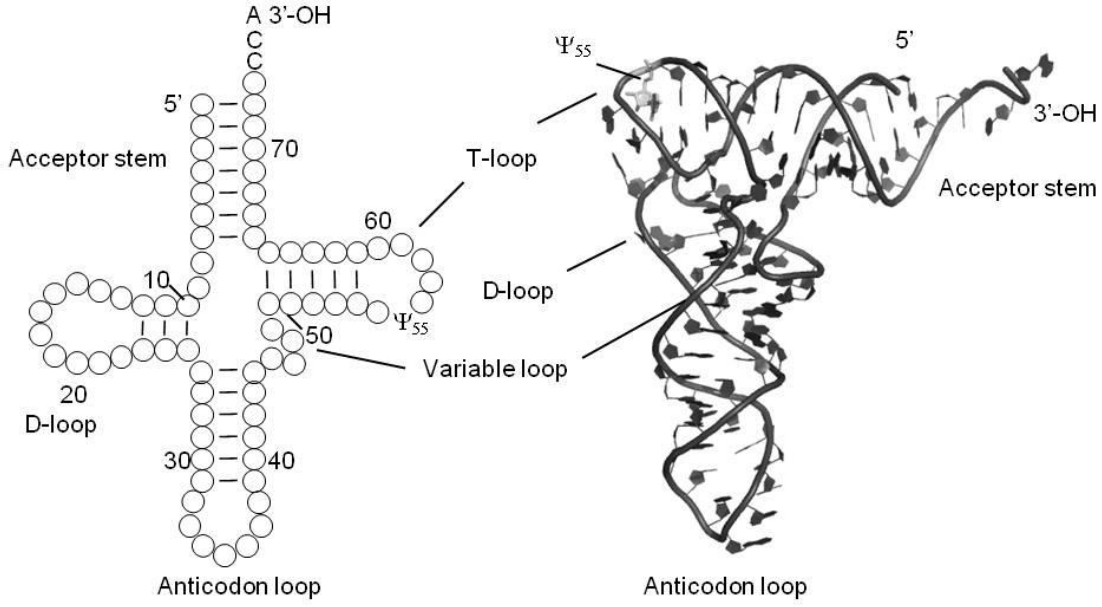


Figure 1.2. Structure of tRNA

Schematic representation of tRNA secondary structure (left), cartoon representation of tertiary structure of yeast $tRNA^{Phe}$ (PDB ID: 4TRA) (right). Key structural elements are labeled. Every tenth nucleotide is indicated and the universally conserved ψ55 is highlighted. Single stranded tRNA molecule base-pairs through complementary regions forming a cloverleaf secondary structure. The secondary structure shows three stem-loops, and a stem formed from the base pairing of 3'- and 5'- ends. In addition, tRNA also displays a small variable loop. Through coaxial stacking of helices, tRNA forms a three dimensional L-shaped structure.

The 2'-O-methylation of ribose stabilizes the C3'-endo form of ribose resulting in conformational rigidity (Kawai *et al.* 1992), contributing to the tRNA function in codon recognition and protecting the RNA from unspecific degradation (Motorin and Helm 2010). Methylations are linked to the thermal stability of tRNA and rRNA in hyperthermophiles (Kowalak *et al.* 1994; Noon *et al.* 1998), and to the conferral of resistance against antibiotics (Douthwaite *et al.* 2005). It was also proposed that the methylations clustered on the exit tunnel of the ribosome provide a hydrophobic environment that may prevent the nascent polypeptides from sticking to the tunnel

(Nissen *et al.* 2000). Together these observations underline both the structural and functional importance of modifications in RNA.

1.2.3 Pseudouridine

1.2.3.1 Structure and important properties of pseudouridine

Initially, RNA was believed to be composed of four canonical nucleotides. Research conducted in the early 1950s using ribonuclease-digested yeast RNA extract led to the isolation of a minor RNA constituent that was referred to as the 'fifth nucleotide' (Davis and Allen 1957). As this newly identified nucleotide (5-ribosyluracil) exhibited similar physical and chemical properties to the canonical uridine (1-ribosyluracil), it was named 'pseudouridine', and abbreviated 'ψ' (Cohn 1959). Further analysis of this compound revealed that it is a uridine derivative with an unusual C-C glycosidic bond, instead of the typical N-C glycosidic bond (Figure 1.3) (Cohn 1960). It has been proposed that the isomerization of uridine to ψ involves breakage at N_1-C_1', followed by a rotation around the N_3-C_6 axis, with the concomitant formation of a new glycosidic bond between C_5 and C_1' (Goldwasser and Heiniukson 1966).

Despite similarities to its parent uridine nucleotide, ψ displays a few distinct properties attributing to it certain advantages over a uridine in RNA (Figure 1.3). Its unique C-C glycosidic linkage is predicted to provide enhanced conformational flexibility over the N-C glycosidic bond through increased rotational freedom about the C-C link (Lane *et al.* 1995). With this increased rotational flexibility, ψ was thought to have a preference for the *syn* glycosyl conformation over the *anti* conformation found in uridine, suggesting a role for ψ as a 'conformational switch' in RNA (Neumann *et al.* 1980).

While this was shown to be the case with the free nucleoside, in RNA however, ψ has been found to always exist in the *anti* conformation (Yarian *et al.* 1999). In the *anti* conformation, ψ maintains the ability to base pair with an adenine nucleotide like uridine.

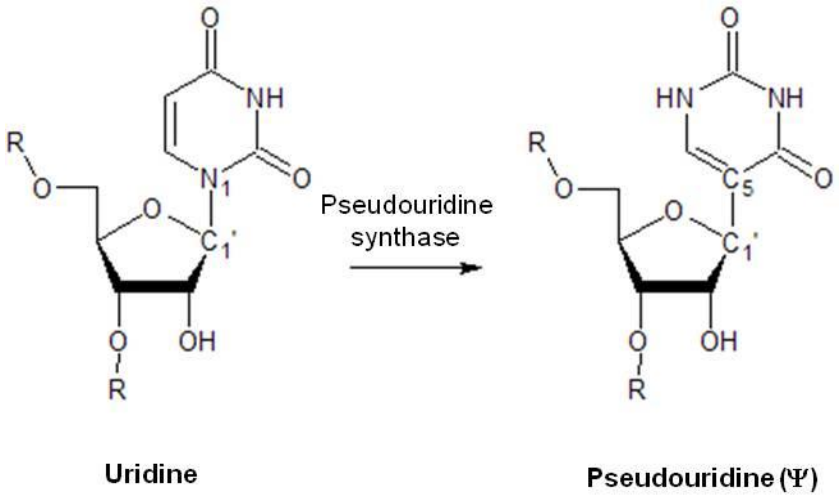


Figure 1.3. Isomerization of uridine to pseudouridine.

Pseudouridine is introduced at specific sites in RNA by pseudouridine synthase. The two important differences found in pseudouridine compared to uridine are the C-C glycosidic bond and an extra -NH group that can participate in a new hydrogen bond.

The other important feature is the extra N_1H imino group in ψ which has the potential to participate in an additional hydrogen bond, establishing novel interactions in RNA (Figure 1.3). tRNA co-crystallized with a tRNA synthetase revealed a water molecule linking the ψ to the phosphate backbone through hydrogen bonding with N_1H of ψ and the 5' phosphate of the same residue (Arnez and Steitz 1994). Through this novel hydrogen bond, ψ stabilizes the local structure of RNA resulting in increased rigidity (Yarian *et al.* 1999). In addition to the structural stabilization by extra hydrogen bonding, studies conducted using model RNA oligonucleotides containing ψ demonstrated that ψ increases local RNA stacking by promoting a 3'-endo conformation of ribose (Davis

1995). This also increases the base stacking in neighboring nucleotides in a cooperative manner contributing to increased stability of RNA structure. Consistent with these observations, ψ has been shown to enhance the thermal stability of RNA (Arnez and Steitz 1994).

1.2.3.2 Distribution of Ψ in RNA and its biological significance

ψs are the most abundant individual modifications observed in RNA. They are distributed in all three domains of life in many non-coding RNAs, such as tRNA, rRNA, snRNA and snoRNA (Charette and Gray 2000). They are very common in tRNA; especially, ψ at position 55 is found in all tRNAs with the exception of few initiator tRNAs (Samuelsson *et al.* 1987). ψs are also commonly encountered in the anticodon stem-loop (ψ38, ψ39) and the D-arm (ψ13) of tRNA. Although less common, they are also found at other sites in tRNA in a phylogenetic domain specific manner (Auffinger and Westhof 1998). Most of the conserved ψs tend to be located at the stem and loop junctions, owing to the stabilizing effect that ψ has on these structural motifs (Charette and Gray 2000). Grosjean and colleagues have demonstrated that ψ38 and ψ39 in tRNAs play a role in enhancing stop codon readthrough and +1 frame shifting in yeast (Lecointe *et al.* 2002). ψs introduced at multiple sites in eukaryotic tRNA by pseudouridine synthase 1 (Pus1) are predicted to play an important role in nuclear export of tRNA (Grosshans *et al.* 2001). Notably, it has been shown that ψ55 regulates the modifications at the other sites in tRNA, in particular Gm18, m^5s^2U54, and m^1A58 for low-temperature adaptation in *Thermus thermophilus* (Ishida *et al.* 2011). Together with other

modifications in tRNA, ψs take part in fine-tuning the tRNA structure that in turn influences codon reading and the accuracy of protein synthesis (Harrington *et al.* 1993).

As discussed in section 1.1, rRNA undergoes extensive chemical modifications during rRNA processing and its assembly into the ribosome. ψs are ubiquitous in both small and large subunit rRNAs. It appears that the number of ψs increases with the complexity of the organism. While *E. coli* rRNA contains only 11 such modifications, rRNA from yeast and humans show around 50 and 100 ψs, respectively (Ofengand 2002). ψs are found to cluster in functional centers of the ribosome, corresponding to the peptidyl transferase center, the decoding center (site for interaction of tRNA with mRNA), the peptide exit tunnel and also the interface of large and small subunits of the ribosome (Charette and Gray 2000; Decatur and Fournier 2002). Based on their prevalence in functional sites, ψs in rRNA are predicted to stabilize local RNA structure through RNA-RNA and RNA-protein interactions and to play an important role in ribosome biogenesis and ribosome function (Charette and Gray 2000). In accordance with this prediction, blocking the formation of a conserved ψ (position 2920) in the A-site of the peptidyl transferase center in yeast rRNA resulted in a reduced translation rate and impaired polysome formation (King *et al.* 2003). In *E. coli*, three ψs at positions 1911, 1915, and 1917 are found in helix 69 of the 23S rRNA, which forms a bridge between the large and small ribosomal subunits. Loss of these modifications in bacteria by deleting RluD, the pseudouridine synthase responsible for their formation, revealed that these ψs play an important role in translation termination and peptide release by influencing the activity of releasing factor 2 (Kipper *et al.* 2011). Extensive analysis of modifications in yeast rRNA by the Fournier lab further demonstrated that conserved ψs in helix 69 of eukaryotic ribosomes play

broader roles, influencing ribosome formation, stability, and function (Liang *et al.* 2007). Depletion of ψs in combination with methylations in the decoding center of the yeast ribosome resulted in reduced translational activity and defects in synthesis of the ribosomal small subunit (Liang *et al.* 2009). A common observation made in all these studies is that a minimal or null effect is observed upon removal of a single modification, while deletion of modifications at 3 or more sites from the same region showed significant effects indicating that the modifications act cooperatively.

In eukaryotes, small nuclear RNA (snRNA) combines with proteins to form 'small nuclear ribonucleoprotein complexes' (snRNPs) referred to as the 'spliceosomal machinery'. These complexes are responsible for splicing non-protein coding introns from eukaryotic pre-mRNA resulting in mature mRNA containing protein-coding exons. All major spliceosomal RNAs (U1, U2, U4, U5 and U6) display extensive pseudouridylation. Notably, *Xenopus* U2 snRNA contains 13 ψs, accounting for 12% of its nucleotides and 60% of the total RNA modifications found in this RNA (Reddy and Busch 1988). In contrast, few ψs are found in minor spliceosomal RNAs. As in rRNA and tRNA, many of these ψs in spliceosomal RNAs are conserved across species and are found in functionally important regions (Karjolich and Yu 2010). The significance of ψs in snRNAs is not well understood except for their importance in U2 snRNA. Using a reconstituted system from *Xenopus* oocytes, Zhao and Yu (2004) demonstrated the essential role of ψs in branch site recognition region by U2 snRNA in the assembly of functional spliceosome and pre-mRNA splicing. Loss of ψ35 in the branch site sequence from U2 snRNA when coupled with either a substitution of uridine to guanine at position 40 or deletion of the uridine, leads to the accumulation of pre-mRNA and results in a

temperature-sensitive growth phenotype in yeast (Yang *et al.* 2005). ψ7 found in U5 snRNA has been shown to participate in binding the pre-mRNA, suggesting a role for this residue in stabilizing the interaction between U5 snRNA and pre-mRNA (Karjolich and Yu 2010). Interestingly, recent studies performed with yeast subjected to nutrient deprivation and heat shock, revealed the induced formation of ψs at novel sites in U2 snRNA, indicating a probable regulatory role for ψs under stress conditions (Wu *et al.* 2011).

ψs have not yet been reported to naturally exist in mRNAs, owing to the low abundance of mRNAs compared to rRNAs and tRNAs, which makes it difficult to analyze their modifications. However, several studies point towards the possibility of widespread ψs in mRNA (Ge and Yu 2013). Studies using artificial H/ACA RNAs indicated that mRNA can in principle be pseudouridylated at target sites (Chen *et al.* 2010). Interestingly, when uridine present in nonsense codons is subjected to pseudouridylation, it suppresses translation termination (Karjolich and Yu 2011). Further, Kariko and colleagues have shown that pseudouridylation of *in vitro* transcribed mRNA increases translation by decreasing the activation of RNA-dependent protein kinase compared to uridine containing mRNA (Anderson *et al.* 2010). Increased translational efficiency combined with reduced immunogenicity observed due to ψ incorporation makes mRNA containing ψs a potential choice in RNA-based therapeutic applications (Anderson *et al.* 2010). Together, all these findings suggest that ψs are selected in evolution to play important roles in diverse RNA species, particularly by stabilizing the structural motifs present in functional centers, thereby influencing several key cellular activities in all domains of life.

ψ was shown to be metabolized in *E. coli* using a pseudouridine kinase (phosphorylates ribose) and pseudouridine-5'-phosphate glycosidase (hydrolyzes the unique C-C glycosidic bond in ψ), releasing uracil and ribose-5'-phosphate (Preumont *et al.* 2008). Like other nucleotides, this study indicated that ψ can be recycled for efficient use of cellular components. Notably, homologs of these enzymes are found in most eukaryotes, but not in mammals.

1.3 Ψ synthases

As discussed above, ψs are extensively distributed in almost all RNA classes and are important for several cellular activities. In all domains of life, specific uridines in RNA are isomerized to ψs by a group of enzymes known as 'ψ synthases'. Although bacteria use protein-only ψ synthases (hence referred to as stand-alone ψ synthases), archaea and eukaryotes employ RNA-dependent protein complexes to accomplish this task in addition to stand-alone enzymes.

1.3.1 Stand-alone ψ synthases

Johnson and Söll (1970) at Yale University were first to demonstrate pseudouridine synthase activity in *E. coli* extracts using [^{14}C] labeled *in vitro* transcribed tRNA. Their experiments revealed that ψs are specifically derived from existing uridines in RNA by enzyme(s) present in an *E. coli* cell extract. Studies based on observations that ψs are absent from the anticodon loop of $tRNA^{His}$ in bacterial *hisT* (a gene found in the histidine operon) mutant strains led to the discovery and characterization of the first bacterial pseudouridine synthase (pseudouridine synthase I), the product of the *hisT* gene (Cortese

et al. 1974; Marvel *et al.* 1985). A decade later, the Ofengand and Lane labs together purified and characterized bacterial ψ55 synthase, which they named TruB (for **tR**NA pseudo **U** modification), renaming the previously identified pseudouridine synthase I as TruA (Nurse *et al.* 1995). Continued efforts to identify the other ψ synthases using bioinformatic and various biochemical approaches resulted in the discovery of several ψ synthases in bacteria and eukarya. Knowledge of ψ synthases in archaea lags behind with only a handful of them having been characterized so far (Grosjean *et al.* 2008; Blaby *et al.* 2011).

1.3.1.1 Ψ synthase families and their structural organization

On the basis of sequence similarity, ψ synthases are grouped into six families (Table 1.1) (Hamma and Ferré-D'Amaré 2006; McCleverty *et al.* 2007). Each of these families is named after its *E. coli* representative. The only exception to this is Pus10, a recent addition to the ψ synthase families with no known homologs in bacteria, which is named after the archaeal enzyme (Roovers *et al.* 2006). Representative crystal structures of all six ψ synthases families have been solved (Foster *et al.* 2000; Hoang and Ferré-D'Amaré 2001; Sivaraman *et al.* 2002; Hoang and Ferré-D'Amaré 2004; Phannachet *et al.* 2005; Hoang *et al.* 2006; McCleverty *et al.* 2007). Comparative analysis of these structures revealed that despite the lack of sequence similarity, all ψ synthases display a common catalytic fold and a conserved active-site structure (Figure 1.4). The core architecture unique to ψ synthases consists of an extended platform of eight continuous mixed β-sheets with several helices and loops surrounding them. The center of the platform acts as an active-site cleft flanked by a conserved loop-helix structure on one side and a long

loop on the other side. This long loop contains a strictly conserved aspartate residue, referred to as the catalytic aspartate, found in all ψ synthases.

Table 1.1. ψ synthase families with representative enzymes

Family Name	**Accessory domain**	**Example: substrate specificity of enzyme naming the family**		
		Substrate	**Modification sites**	**Organism**
TruA	None	tRNA	38, 39, 40	*E. coli*
TruB	PUA (C-terminal)	tRNA	55	*E. coli*
RluA	S4-like domain* (N-terminal)	tRNA	32	*E. coli*
		23S rRNA	746	*E. coli*
RsuA	S4-like domain (N-terminal)	16S rRNA	516	*E. coli*
TruD	TruD	tRNA	13	*E. coli*
Pus10	THUMP (N-terminal)	tRNA	55	*P. furiosus*

* The accessory domain is absent in the enzyme RluA but found in other members of the family

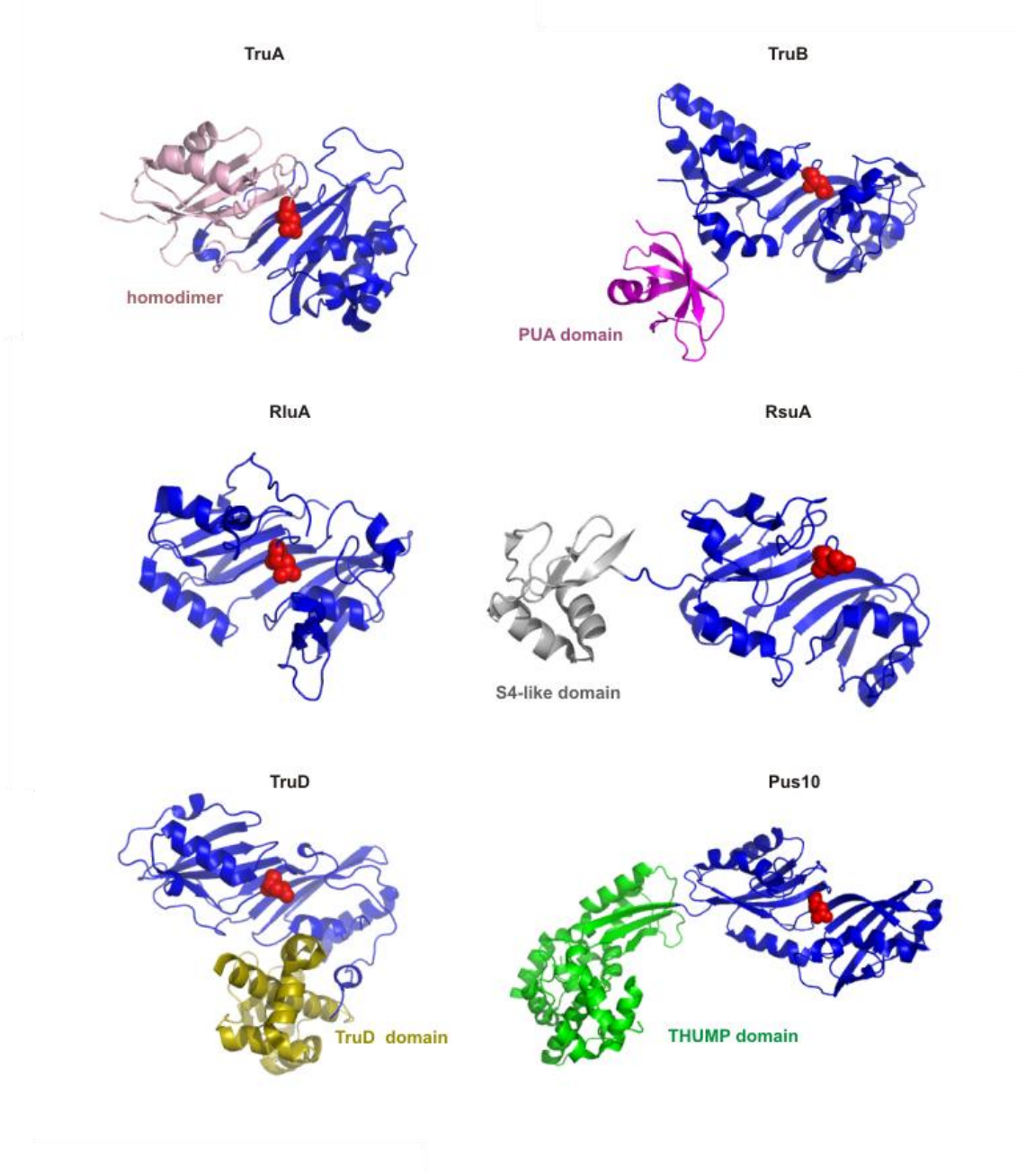


Figure 1.4. Structures of representative Ψ synthases.

Cartoon representation of a representative structure from each of the six families of ψ synthases. *E. coli* TruA (1DJ0), TruB (1K8W), RluA (2I82), RsuA (1KSK), TruD (1SZW), and human Pus10 (2V9K). The catalytic domain of each protein is shown in the same orientation and in a single color (blue), while the accessory domains are colored differently (TruA additional subunit: pink, PUA: magenta, S4: grey, TruD: olive, THUMP: green) and labeled. The catalytic aspartate in the catalytic pocket is highlighted in red.

ψ synthases show diverse secondary structural insertions adjacent to the catalytic site. While TruB has a thumb-like insertion on one side of the catalytic pocket, RluA, in addition to the thumb loop, contains a forefinger loop on the other side of the catalytic pocket. These loops are believed to play an important role in interaction with RNA (Hamma and Ferré-D'Amaré 2006). Besides the catalytic domain, most of the pseudouridine synthases have distinct accessory domains appended either at the N- or C-terminus of the protein (Table 1.1). While the TruB family exhibits a C-terminal PUA domain (named after its presence in some **P**seudo**U**ridine synthases and **A**rchaeosine-transglycosylases) (Hoang and Ferré-D'Amaré 2001), some members of the RluA (referring to **R**ibosomal **l**arge sub**u**nit) and RsuA (**R**ibosomal **s**mall sub**u**nit) families contain an N-terminal domain similar to the one in ribosomal protein S4 (Hamma and Ferré-D'Amaré 2006). These accessory domains are predicted to facilitate RNA binding. An interesting observation was made from a crystal structure of *E. coli* TruD, wherein the catalytic domain of this enzyme displays a circular permutation of the secondary structural elements (Hoang and Ferré-D'Amaré 2004), besides an additional TruD specific domain inserted into its catalytic domain (Figure 1.4). In contrast to all other ψ synthases, only TruA functions as a dimer (Hur and Stroud 2007).

Based on the conserved catalytic core found in all ψ synthases, these enzymes are predicted to be evolved from a common ancestor through divergent evolution (Mueller 2002; Hur *et al.* 2006). Among all ψ synthase families, TruD shows the lowest sequence similarity to other families and has homologs in all domains of life indicating that it could have diverged first from other ψ synthases, followed by TruA (Hamma and Ferré-

D'Amaré 2006). RluA and RsuA are relatively similar sharing three conserved motifs (Motifs I, II, and III), along with TruB, which shares motif I and II (Koonin 1996).

1.3.1.2 Substrate recognition and catalysis

RNA targets of ψ synthases vary from simple RNA stem-loops to long rRNA molecules with complex three-dimensional structures. ψ synthases display diverse specificity with respect to the number of modification sites and types of RNAs modified (Table 1.1). Some of them modify only a specific site in a single class of RNA; for example, so far RsuA has only been shown to introduce ψ at position 516 in 16S rRNA (Wrzesinski *et al.* 1995). Others modify several sites in structurally similar regions in multiple RNAs of the same type. TruB, also known as ψ55 synthase, modifies uridine at position 55 in almost all bacterial tRNAs (Nurse *et al.* 1995). TruA on the other hand is responsible for modifying three different nearby sites (38, 39, and 40) in the same tRNA (Hur and Stroud 2007). Furthermore, some enzymes are able to modify different positions in entirely different classes of RNA. For instance, RluA pseudouridylates a single uridine in tRNA (U32) as well as in rRNA (U746 in 23S rRNA) (Wrzesinski *et al.* 1995).

ψ synthases employ different approaches to recognize the target uridine in RNA, either in the context of structure or sequence. In the case of TruB, RNA recognition takes place mainly by shape complementarity, wherein the enzyme recognizes the native structure of the T-loop containing the modification site (Hoang and Ferré-D'Amaré 2001). Further, to enhance specificity, TruB also makes contact with a few conserved nucleotides in the substrate, such as the C56 located close to the target site (Hamma and

Ferré-D'Amaré 2006). RluA on the other hand induces a structural reorganization in the anticodon loop containing the target U32 nucleotide, subsequently recognizing this newly formed structure through shape complementarity (Hoang *et al.* 2006). In contrast to the two strategies used by TruB and RluA, TruA establishes an interaction with two conserved tRNA structural elements (elbow and the D stem) with the help of an additional subunit forming the homodimer, thereby recognizing the dynamic anticodon loop containing the three target sites located in a row (Hur and Stroud 2007). While RluF, a member of RsuA family has been shown to follow a similar approach to RluA in identifying the target uridine (Alian *et al.* 2009), substrate interactions of TruD and Pus10 is not well understood. In addition to all these mechanisms used in substrate recognition, Cbf5, a TruB family ψ synthase, uses a completely different strategy by employing a guide RNA to recognize the target substrate (see section 1.3.2 for details). In all of these modification events, in order to gain access to the target site, usually the target uridine is flipped out from the RNA into the active site of the enzyme, with up to two additional bases undergoing base-flipping (Hamma and Ferré-D'Amaré 2006).

Bioinformatic and structural studies revealed five signature motifs present in the catalytic domains of ψ synthases, motif I, II, IIa, III and IIIa (Koonin 1996; del Campo *et al.* 2004; McCleverty *et al.* 2007). The catalytic pocket of ψ synthases is largely hydrophobic in nature and harbours three conserved active site residues. Besides the invariant aspartate residue located in motif II, a tyrosine (replaced by phenylalanine in TruD) in motif IIa, and a basic residue, either arginine or lysine in motif III, are found in the active site of ψ synthases (Hamma and Ferré-D'Amaré 2006). Structural and biochemical analyses have confirmed the critical role of the aspartate residue in catalysis

(Ramamurthy *et al.* 1999; Hoang *et al.* 2005). The conserved tyrosine/phenylalanine (Y179 in TruB) is involved in a stacking interaction with the target base and is proposed to act as a general base in a later step of the catalytic reaction (Phannachet *et al.* 2005). The arginine or lysine (R181 in TruB) makes a salt bridge with the aspartate and it might play a role in positioning the aspartate for catalysis (Pan *et al.* 2003). Besides these conserved active site residues, except TruB and TruD, the other four ψ synthase families contain an arginine residue two nucleotides prior to the catalytic aspartate. TruB, on the other hand has a histidine five nucleotides prior to the catalytic aspartate. Cocrystal structures of TruB and RluA have shown that these histidine and arginine residues occupy the location vacated by the flipped-out target nucleotide in tRNA, respectively (Hoang and Ferré-D'Amaré 2001; Hoang *et al.* 2006).

Despite significant knowledge of the structure and substrate recognition by ψ synthases, the actual chemical mechanism of ψ formation is unknown. The same catalytic fold and the conserved active-site structure of ψ synthase families suggest that all ψ synthases likely employ a common catalytic mechanism. In support of this hypothesis, kinetic analysis of *E. coli* TruB, TruA, and RluA by our lab revealed that all three studied enzymes have very similar rate constants with a uniformly slow catalytic step (Wright *et al.* 2011). Some mechanistic insight has been obtained using a 5-fluorouridine (5-FU) substituted RNA to probe the chemical steps occurring during ψ formation (Gu *et al.* 1999; Hoang and Ferré-D'Amaré 2001). At minimum, ψ formation involves three chemical steps; breakage of the N_1-C_1' glycosidic bond, rotation of the detached base, and reattachment of the base to form a new glycosidic bond between C_5 and C_1' (Hamma and Ferré-D'Amaré 2006). Three catalytic mechanisms have been

proposed based on different roles of the catalytic aspartate. According to the first mechanism (Gu *et al.* 1999), the aspartate residue attacks the C6 of the uracil base in a Michael addition forming an ester intermediate (Michael adduct) (Figure 1.5). Base rotation takes place around the ester bond while the uracil is attached to the enzyme, followed by hydrolysis of the ester linkage resulting in ψ. According to the second mechanism (Huang *et al.* 1998), the aspartate carries out a nucleophilic attack onto C1' of ribose generating an acylal intermediate (Figure 1.5). The liberated base then undergoes a rotation allowing the new C-C glycosidic bond formation.

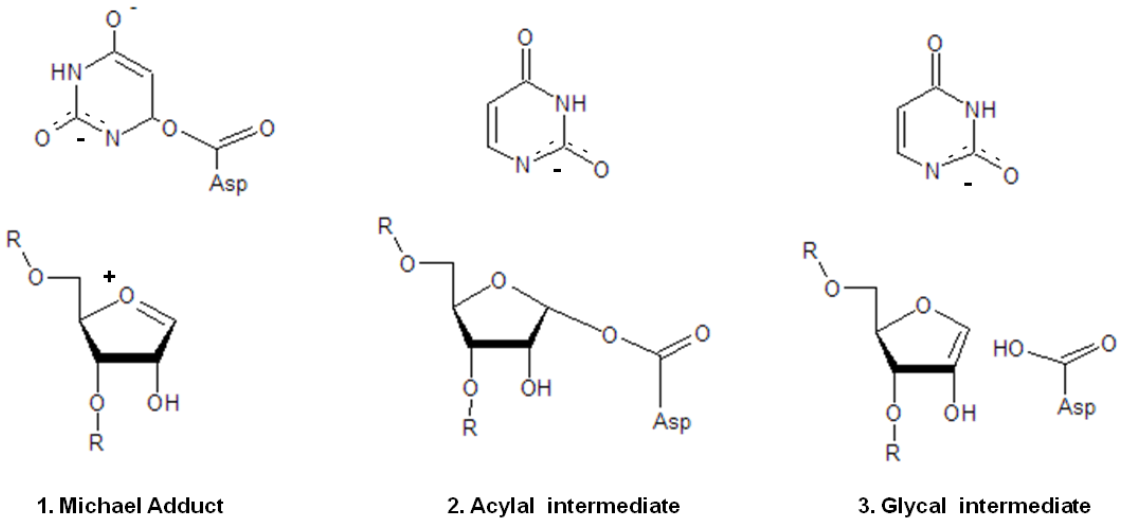


Figure 1.5. Potential intermediates formed during ψ formation

The three mechanisms proposed to explain ψ formation proceed through different intermediates. 1. Michael adduct is generated during the nucleophilic attack at C6 of uracil by the catalytic aspartate. 2. Acylal intermediated forms when the nucleophilic attack occurs at C1' of ribose 3. Proton abstraction from C2' of ribose produces a glycal intermediate.

More recently, a third mechanism has been proposed by the Mueller group based on the observation of two isomeric hydrated products formed when a 5-FU containing tRNA is incubated with TruB (Miracco and Mueller 2011). According to this mechanism, the

reaction proceeds through a glycal intermediate produced as a result of the abstraction of a proton by the aspartate from C2' of ribose. This study rules out the Michael addition mechanism, but is also in accordance with the mechanism containing an acylal intermediate. Further studies are required to clarify the exact mechanism involved in ψ formation by ψ synthases.

1.3.2 H/ACA small ribonucleoprotein complex

As mentioned earlier, isomerization of uridine to ψ is also performed by H/ACA sRNP or snoRNP complexes in archaea and eukaryotes, respectively. H/ACA RNAs were first reported by the Fournier group as a new major class of RNA besides C/D box RNAs in the nucleolus of yeast and vertebrates (Balakin *et al.* 1996). Whereas the C/D box RNAs function in guiding the site-specific methylation of ribose, H/ACA box RNAs are involved in the pseudouridylation process (Ni *et al.* 1997). Although they mainly guide rRNA and snRNA pseudouridylations, there is one case reported in archaea where they could guide tRNA modification (Muller *et al.* 2009). H/ACA RNAs involved in guiding snRNA modifications are known as small Cajal body specific RNAs (scaRNAs) owing to their localization in Cajal bodies (Darzacq *et al.* 2002). Cajal bodies are dynamic nuclear regions found in eukaryotes and are the centers for the biogenesis of snRNPs involved in splicing. To perform site-specific pseudouridylation, H/ACA RNA forms a complex with a set of four evolutionarily conserved proteins consisting of a catalytic component, Cbf5 (dyskerin in humans) and three accessory proteins, Nop10, L7Ae (Nhp2 in eukaryotes) and Gar1.

1.3.2.1 H/ACA snoRNAs

H/ACA RNAs are typically 70 - 250 nucleotides long and adopt a 5'-hairpin-hinge-hairpin-tail-3' secondary structure (Figure 1.6A) (Balakin *et al.* 1996). The eukaryotic snoRNAs usually have a two-stem H/ACA RNA with a hinge region connecting them. Although single hairpins are common in archaea, they also have two and three hairpin structures (Omer *et al.* 2003). In eukaryotes, a set of core proteins are proposed to bind each hairpin region and may direct the pseudouridylation of two distantly spaced uridines in rRNA (Watkins *et al.* 1998; Kiss *et al.* 2010). All H/ACA RNAs show two characteristic conserved sequences, box 'H' referring to the sequence in the hinge region (ANANNA, where N is any nucleotide) and 'ACA' (ANA in archaea) referring to the sequence found in the tail region, exactly three nucleotides upstream of the 3'-end of the RNA. In each hairpin structure at the junction of two stem regions, there is an internal loop called the 'pseudouridylation pocket' with 9-13 nucleotides on each strand (Reichow *et al.* 2007). The pseudouridylation pocket displays complementarity to a specific sequence in the substrate RNA, allowing base pairing to position the unpaired target uridine at the base of the upper stem for modification. The distance between the target uridine and the conserved H or ACA box elements is typically between 14 and 17 nucleotides (Ni *et al.* 1997). A kink-turn, a common structural motif found in RNA, is formed by a 3-nucleotide bulge flanked with non-canonical GA base pairs on one side and canonical GC base pairs on the other side (Klein *et al.* 2001). Archaeal, but not eukaryotic H/ACA RNAs, contain a kink-loop (a variant of kink-turn) in the upper half of the stem, where the accessory protein L7Ae binds (Rozhdestvensky *et al.* 2003). A kink-loop is similar to a kink turn, but lacks canonical stem.

1.3.2.2 Structure and function of H/ACA sRNP complex

In vitro reconstitution of H/ACA sRNPs from archaea (Baker *et al.* 2005; Charpentier *et al.* 2005) and the crystal structures of various subcomplexes as well as complete H/ACA sRNP with and without substrate (Hamma *et al.* 2005; Li and Ye 2006; Rashid *et al.* 2006; Duan *et al.* 2009; Liang *et al.* 2009) provided a deeper insight into their structural organization and function. The catalytic component, Cbf5, is a TruB family ψ synthase and contains an N-terminal catalytic domain and a large C-terminal PUA domain with N- and C-terminal extensions compared to TruB. Cbf5 specifically recognizes the conserved 3'-ACA sequence of H/ACA RNA and interacts with the lower stem of the hairpin through the PUA domain (Figure 1.6B). Nop10 forms extensive contacts with Cbf5 and also binds the upper stem of the H/ACA RNA hairpin, whereas Gar1 exclusively binds to Cbf5 at a distant region, especially to Cbf5's thumb loop which is involved in substrate RNA binding (see below). In contrast, L7Ae interacts with the kink-loop in the upper stem region of archaeal H/ACA sRNA and also makes few contacts with Nop10. Substrate RNA is recruited into the pseudouridylation pocket through base pairing to one side of the guide RNA forming a U-shaped substrate structure. Substrate binding induces a conformational reorganization of the pseudouridylation pocket and influences the two flanking stems of the guide RNA thereby positioning the target uridine in the active site of Cbf5. In the absence of substrate, the thumb loop of Cbf5 assumes an open conformation by binding to Gar1, but upon substrate recruitment the thumb loop establishes an interaction with the substrate RNA, stabilizing it in the active site (closed conformation) (Duan *et al.* 2009). The interaction of Gar1 with the thumb loop was suggested to play a role in sensing the

successful modification allowing the concomitant release of the product. The yeast H/ACA snoRNP complex has also been successfully reconstituted and its crystal structure has been determined, which revealed a similar structural organization of the yeast H/ACA snoRNP complex and the archaeal complex (Li *et al.* 2011).

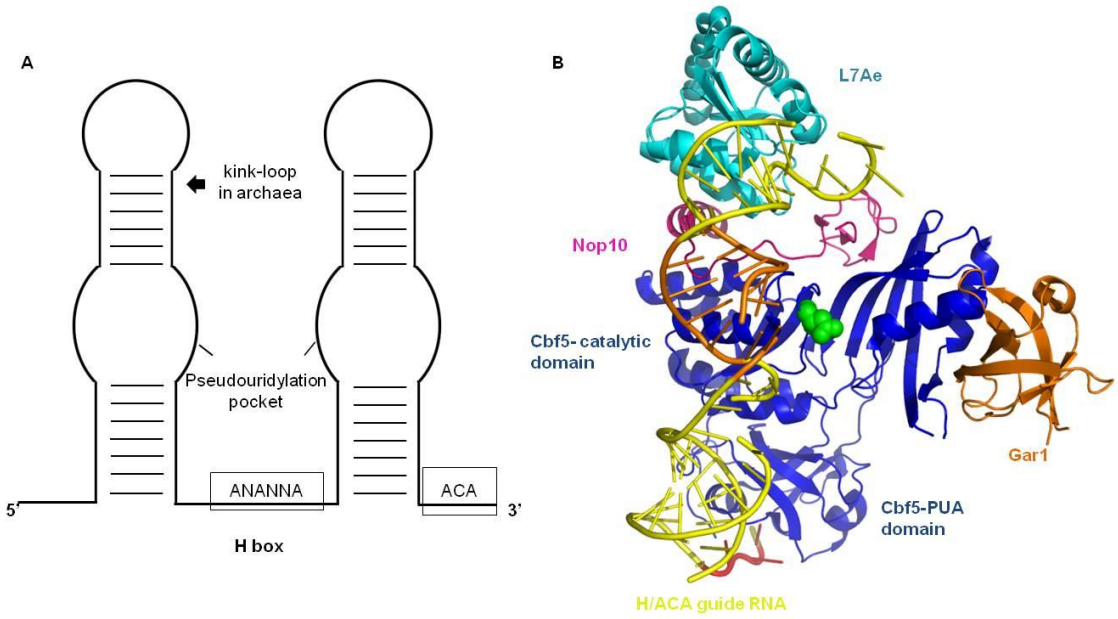


Figure 1.6. Structure of H/ACA guide RNA and snoRNP complex

A. Schematic diagram of H/ACA guide RNA showing a two-hairpin H/ACA guide RNA with the conserved box H and ACA motifs. The pseudouridylation pocket is the site for substrate binding, and the kink-loop is found in the upper stem of archaeal guide RNAs where L7Ae binds. B. Cartoon representation of the structure of a single-hairpin H/ACA guide RNA (yellow, pseudouridylation pocket in orange and 3'-ACA in red) from *P. furiosus* bound to all four H/ACA proteins (Cbf5 in blue, Nop10 in pink, Gar1 in orange, and L7Ae in cyan) in the absence of substrate RNA (PDB ID 2HVY). H and ACA motifs in RNA and the catalytic aspartate in Cbf5 are highlighted in red and green, respectively.

Recently, Yang and colleagues reported the kinetic and thermodynamic analysis of the archaeal H/ACA sRNP complex (Yang *et al.* 2012). Based on the data obtained using fluorescence correlation spectroscopy, they proposed a two-step sequential model for

substrate binding and product release. The substrate initially binds the guide RNA through base pairing, and is then recruited into the active site by interacting with the thumb loop of Cbf5. Following ψ formation, the thumb loop undergoes a conformational change, thereby releasing the product RNA. In this work, the authors also reported a dual role for Gar1, both in catalysis and in product release. Further evaluation is required to test if the binding and kinetic parameters reported in this work are relevant since experiments were conducted at 27 ˚C as opposed to the reported optimal temperature (70 ˚C) for *P. furiosus* Cbf5.

In addition to the major function of the H/ACA sRNP complex in ψ formation in multiple substrates, they also play other important roles in eukaryotes. Some H/ACA snoRNAs are involved in rRNA processing (see section 1.1). For instance, snR30 from yeast base pairs to short internal sequences in 18S rRNA and is believed to guide pre-rRNA processing factors essential for the nucleolytic cleavage of 18S rRNA from 35S pre-rRNA (Tollervey and Kiss 1997)

Telomerase is an essential eukaryotic enzyme that adds nucleotide repeats to the telomere, maintaining its length. Telomerase is composed of telomerase reverse transcriptase (TERT) and an RNA component (TERC) that acts as a template for the DNA extension. Notably, mammalian TERC contains an H/ACA domain that is essential for telomerase function and localization (Collins 2006). Similar to the H/ACA snoRNP complexes guiding pseudouridylation, the telomerase RNA H/ACA domain also forms a complex with the same four core proteins (Dez *et al.* 2001), but there are no known pseudouridylation targets of this complex.

Dyskeratosis congenita (DC) is an inherited rare bone marrow failure syndrome with the affected patients showing defects in telomerase function. DC is genetically heterogenous with the causative mutations found in human Cbf5 (dyskerin), NOP10, NHP2, TERC and TERT and in a telomere protecting protein (Tin2) (Kirwan and Dokal 2009). One of the most common and severe X-linked forms of this syndrome is caused by mutations in the PUA domain at the binding interface of H/ACA snoRNA and dyskerin (Rashid *et al.* 2006), implying that RNA binding is impaired. DC mutations in dyskerin, NHP2 or NOP10 are reported to affect the assembly of H/ACA snoRNP complexes (Trahan *et al.* 2010). Although initial studies from animal models suggested that defects in ribosome biogenesis and function due to mutations in dyskerin are likely responsible for DC, current studies favour the effects on telomerase as the principal factor leading to DC (Kirwan and Dokal 2009).

1.3.2.3 Eukaryotic H/ACA snoRNP biogenesis

snoRNA genes show variation in organizational structure. In archaea and yeast, they are mostly encoded by independent genes organized into mono- or polycistronic transcriptional units. In contrast, in higher eukaryotes, most of them are encoded in intronic regions of protein-coding or non-coding genes (Kiss *et al.* 2010). The proteins expressed from these host genes are usually involved in ribosome biogenesis and function, indicating the coordination between the synthesis of snoRNAs and ribosome-associated factors (Filipowicz and Pogacic 2002). snoRNAs transcribed from polycistronic genes are processed by endonucleases to liberate individual snoRNAs. Intronic snoRNAs are either generated by canonical splicing or by the action of specific

endonucleases that remove non-snoRNA regions (Filipowicz and Pogacic 2002). The 5' and 3' ends of the snoRNAs are then trimmed by exonucleases.

The H/ACA snoRNP precursor assembly process begins co-transcriptionally in the nucleoplasm. In addition to the nucleoplasm, snoRNP biogenesis takes place in Cajal bodies with the accumulation of final mature snoRNPs in the nucleolus (Watkins and Bohnsack 2012). The conserved box elements and the stem structures of H/ACA RNA play an essential role in the snoRNP assembly and are thus important for the snoRNA stability and localization (Filipowicz and Pogacic 2002). Assembly of H/ACA snoRNP complex starts with the assembly factor Shq1 binding to Cbf5, which may assist the recruitment of Nop10 and Nhp2, forming a ternary complex of Cbf5-Nop10-Nhp2 (Grozdanov *et al.* 2009). Naf1, another recruiting factor binds Cbf5 in the ternary complex (Ballarino *et al.* 2005; Kiss *et al.* 2010). Interaction of Shq1 and Naf1 is thought to stabilize Cbf5 and to ensure its correct subcellular localization (Darzacq *et al.* 2006). By interacting with the C-terminal domain of RNA polymerase II, Naf1is suggested to guide the Cbf5-Nop10-Nhp2 complex to the nascent H/ACA snoRNA (Yang *et al.* 2005). In the final step of the assembly, Naf1, which has a domain similar to Gar1, is replaced by Gar1 resulting in the final mature complex composed of H/ACA snoRNA and Cbf5-Nop10-Nhp2-Gar1 complex. In addition to the specific assembly factors discussed, the survival of motor neuron (SMN) complex, general assembly factors such as Hsp90 and several other unknown trans-acting factors are believed to assist in this complex and dynamic H/ACA snoRNP biogenesis process (Kiss *et al.* 2010).

1.4 Objectives

As summarized in this chapter, ψs are universally distributed in non-coding RNAs from all domains of life and are important in several essential processes in the cell. Although significant insight has been obtained into this most common RNA modification, several questions remain to be answered, such as: How do ψ synthases that specifically recognize one or two target sites function compared to the ψ synthases with more targets? Is there any order in the occurrence of ψ modifications in the cell? If so, how does the cell regulate ψ synthases to coordinate the modification at multiple target sites in a prioritized manner? Even more challenging questions to tackle are, what is the chemical mechanism involved in ψ formation and what is the biological significance of ψs in RNA?

To address these complex problems, smaller questions that are an integral part of these challenges need to be identified and answered first, which will ultimately guide us towards a comprehensive understanding of ψ formation. With this approach in mind, in this thesis questions have been asked regarding Cbf5, a complex enzyme acting as part of the H/ACA sRNP complex, and Pus10, a new stand-alone ψ synthase. Both of these enzymes have been shown to be involved in the pseudouridylation of U55 in archaeal tRNAs (Roovers *et al.* 2006).

Cbf5, which is typically involved in modifying several target sites in a guide RNA-dependent manner, and can also modify archaeal tRNAs at position 55 with the help of two accessory proteins, but without a guide RNA (Gurha *et al.* 2007). Chapter 2 describes the studies on this unexpected guide RNA-independent function of Cbf5 in

tRNA modification, with the aim to understand the contribution of the two accessory proteins, Nop10 and Gar1. To gain insight into this unexpected function of Cbf5, I have analyzed the steady-state kinetics of this reaction along with thermodynamic binding studies using purified components from an archaeal model, *Pyrococcus furiosus*.

In addition to Cbf5, surprisingly, archaea evolved a novel ψ55 synthase, Pus10, with poor sequence similarity to the TruB family of ψ55 synthases (TruB in bacteria, Pus4 and Cbf5 in eukaryotes) (Watanabe *et al.* 2000). In Chapter 3, as a first step towards understanding the newly discovered ψ synthase Pus10, the biochemical analysis of wild-type *P. furiosus* Pus10 and of several variants constructed using site-directed mutagenesis is described. The aim of this work is to determine if Pus10 evolved to be a more efficient enzyme than Cbf5, which would explain its *in vivo* role in ψ55 formation, and to learn how Pus10 functions in comparison to other known ψ synthases.

Together, it was the objective of these studies on Cbf5 and Pus10 (EC: 5.4.99.25) to provide comparative insight into ψ55 incorporation by these two different enzymes. In Chapter 4, I have summarized the key findings from this thesis and the significance of this work, and discussed the potential future directions in the research of ψ modification that will eventually guide us towards a better understanding of this significant constituent of non-coding RNA.

Chapter 2: tRNA pseudouridylation by archaeal Cbf5 and the contribution of Nop10 and Gar1

Reprinted from

Archaeal proteins Nop10 and Gar1 increase the catalytic activity of Cbf5 in pseudouridylating tRNA

Rajashekhar Kamalampeta and Ute Kothe

Scientific Reports 2012; 2: 663

DOI: 10.1038/srep00663 (2012)

Contributions:

UK designed the research, RK carried out all experiments, and UK wrote the manuscript. All authors reviewed the manuscript.

Changes incorporated:

An experiment performed to determine the optimal reaction conditions required for the pseudouridylation has been added, and some wording has been changed to maintain consistency with the other chapters.

2.1 Introduction

ψ synthases are found in all domains of life as they catalyze the formation of the most abundant RNA modification, the site-specific conversion of uridines to ψs. Based on their structure and sequence similarities, ψ synthases are classified into six families named by their bacterial representatives, TruA, TruB, TruD, RsuA and RluA (Hamma and Ferré-D'Amaré 2006) as well as the unrelated ψ synthase Pus10 found in archaea and some eukaryotes (Watanabe and Gray 2000). The diversity of ψ synthases allows them to

site-specifically target cellular RNAs such as ribosomal RNA, tRNA as well as small nuclear and small nucleolar RNA in eukaryotes. Typically, stand-alone ψ synthases functioning as a single protein recognize one or a small number of related substrate RNAs based on structure and/or sequence (Hamma and Ferré-D'Amaré 2006). In addition, archaea and eukaryotes harbor H/ACA small ribonucleoproteins comprised of the ψ synthase Cbf5 (dyskerin in humans), the accessory proteins Nop10, Gar1 and archaeal L7Ae or eukaryotic Nhp2 as well as an H/ACA guide RNA (Gannot *et al.* 1997; Ni *et al.* 1997). Here, the different H/ACA guide RNAs are responsible for recruiting the target RNAs through specific base-pairing interactions while Cbf5 remains the catalytic component of the complex (Lafontaine *et al.* 1998). However, the roles of the accessory proteins Nop10, Gar1 and L7Ae have not yet been fully established.

Compared to uridines, ψs are characterized by a C-C glycosidic bond and extra second imino group in the base which can participate in additional hydrogen bonds. Presumably, these types of additional interactions confer increased stability to RNA containing ψs (Charette and Gray 2000). Furthermore, ψs near the active centers of the spliceosome and the ribosome have been implicated in the function of these molecular machines (Yang *et al.* 2005; Liang *et al.* 2007). While the exact details of the catalytic mechanism are still under investigation, it is very likely that all ψ synthases employ the same mechanism for pseudouridylation (McDonald *et al.* 2011) since all ψ synthases share a structurally very similar catalytic domain including a strictly conserved aspartate residue which may form a covalent bond to the ribose (Miracco and Mueller 2011). In addition, the active sites of ψ synthases are composed of a positively charged residue that interacts with the catalytic aspartate, and an aromatic residue that forms stacking interactions with the uracil ring

(Hamma and Ferré-D'Amaré 2006). In agreement with the suggested common catalytic mechanism, our lab has recently shown that three families of bacterial ψ synthases (TruA, TruB, and RluA) are characterized by a uniformly slow catalytic step (Wright *et al.* 2011).

Cbf5 is the most complex ψ synthase as it is acting in conjunction with a guide RNA and proteins Nop10, Gar1, and archaeal L7Ae. The structure of the *Pyrococcus furiosus* H/ACA small ribonucleoprotein in the presence and absence of substrate RNA provided insight into the molecular architecture of the complex and suggested possible functions for its components (Li and Ye 2006; Duan *et al.* 2009). As mentioned, Cbf5 is the catalytic unit and interacts extensively with the guide RNA. Nop10 binds to Cbf5 close to the active site and has been proposed to stabilize it (Hamma *et al.* 2005), but it also forms some contacts to the guide RNA. Without Nop10, the H/ACA small ribonucleoprotein is inactive in modifying target RNA (Baker *et al.* 2005; Charpentier *et al.* 2005). Gar1 is the only protein not directly interacting with RNA; instead it can bind to the thumb loop of Cbf5, stabilizing it in an open conformation. Omission of Gar1 limits the guide RNA-dependent pseudouridylation activity of H/ACA small ribonucleoproteins to a single round (Duan *et al.* 2009), presumably because product release is impaired when Gar1 is not inducing an open conformation of the thumb loop. Lastly, L7Ae binds to the kink-turn motif in archaeal H/ACA guide RNA thereby helping to position the guide RNA and the substrate RNA within the complex (Liang *et al.* 2007). Interestingly, archaeal Cbf5 is also able to act in a guide RNA-independent manner as it can on its own introduce ψs at position 55 in the T arm of tRNAs like its bacterial homologue TruB (Roovers *et al.* 2006). This activity is greatly enhanced by the addition of Nop10 and Gar1 (Gurha *et al.*

2007; Muller *et al.* 2007). The significance of this guide-independent activity of Cbf5 is not clear as it has been demonstrated that tRNAs are pseudouridylated *in vivo* by another archaeal enzyme, Pus10 (Blaby *et al.* 2011).

Here, we ask the question why Cbf5 requires additional proteins, in particular Nop10 and Gar1, for its optimal function. Answering this question is not possible by investigating the guide RNA-dependent reaction of the H/ACA small ribonucleoprotein complex, as it loses its complete guide-dependent activity without Nop10. Therefore, we dissected the role of Nop10 and Gar1 for the guide-independent tRNA modification by Cbf5. Our results clearly show that Nop10 and Gar1 not only increase Cbf5's affinity to tRNA, but that they also enhance its catalytic activity.

2.2 Methods

2.2.1 Buffers and reagents

Reaction buffer: 20 mM HEPES-KOH pH 7.0, 1.5 mM $MgCl_2$, 150 mM KCl, 0.1 mM EDTA unless otherwise stated. Nucleotide triphosphates and guanine monophosphate for *in vitro* transcription, and inorganic pyrophosphatase were from Sigma; all other enzymes were from Fermentas. Chemicals were purchased from VWR, DNA oligos were obtained from IDT and [C5-^{3}H] UTP (MT 553) was from Moravek.

2.2.2 Molecular cloning and mutagenesis

The genes encoding the proteins Cbf5, Nop10, and Gar1 were amplified from *P. furiosus* genomic DNA (ATCC, 43587D-5) using the following primers (restriction site in italics, Table 2.1)

Table 2.1. Primers used in amplifying *P. furiosus* Cbf5, Nop10, and Gar1 genes

Primer	Sequence (5' to 3')
Cbf5 sense (BamHI)	*GGATCC*GGCGAGAGACGAGGTAAGAAG
Cbf5 antisense (SalI)	*GTCGAC*TTAGCTTCTATCTCTTTTTTCCC
Nop10 sense (BglII)	CCCG*AGATCTC*AGGTTTAGGATAAGGAAGTGTC
Nop10 antisense (XhoI)	CATT*CTCGAG*TCATTTTTCCTTCCTCCCTA
Gar1 sense (his-tag, NheI)	ATG*GCTAGC*GAAAAACAGGGTGAAAAAATG
Gar1 sense (no his, NcoI)	GCG*CCATGGGC*GAAAAACAGGGTGAAAAAATG
Gar1 antisense (BamHI)	TTC*GGATCC*TCATCTATTCAGCCTTTTCTTC

Subsequently, the genes were inserted by blunt-end ligation into SmaI restricted pUC19 plasmid. Using restriction sites added through the primers, the genes were removed from the pUC19 plasmid and inserted into an expression vector which had been double-restricted with the appropriate enzymes and gel purified. This generated the following plasmids: pETDuet1-PfCbf5 (gene in multiple cloning site I including an N-terminal hexahistidine tag), pETDuet1-PfNop10(nohis) (gene in multiple cloning site II without tag), pET28a-PfGar1 (including N-terminal hexahistidine tag), and pET28a-PfGar1(nohis) (without tag used for purification in complex with Cbf5). To generate a catalytically inactive variant of Cbf5, QuikChange^{TM} mutagenesis was applied to change the catalytic aspartate to asparagine generating plasmid pETDuet1-PfCbf5D85N. All plasmids were verified by sequencing (Macrogen).

2.2.3 Protein expression and purification

For protein expression, plasmids were individually transformed into Rosetta 2(DE3) competent *E. coli* cells (EMD Bioscience) which provide the codons rarely used in *E. coli* and allow the induction of protein expression using IPTG. To express Cbf5 and Nop10, cells were grown at 37 °C in LB medium supplemented with 100 μg/mL ampicillin; for Gar1 and Pus10 expression, LB medium contained 50 μg/mL kanamycin. At an OD_{600} of ~0.6, protein expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. In case of Gar1, cells were transferred to 30 °C prior to the induction. Cells were harvested three hours after induction by centrifugation at 5,000×*g* for 15 min, flash frozen and stored at -80 °C.

Cbf5 and Gar1 were individually purified. For purification of Cbf5-Nop10 and Cbf5-Nop10-Gar1 complexes, cells were mixed to allow formation of the protein complex during cell opening as in previous reports (Li and Ye 2006). In all cases, cells were resuspended in 5 mL/g buffer A1 for purification of Cbf5-Nop10 and Cbf5-Nop10-Gar1 (25 mM sodium phosphate buffer (pH 7.6), 1M NaCl, 5% (v/v) glycerol, 30 mM imidazole and 0.1 mM phenylmethylsulfonyl fluoride (PMSF)) or buffer A2 for purification of Cbf5 and Gar1 alone (20 mM Tris-HCl pH 8.0, 300 mM KCl, 5 mM β-mercaptoethanol, 5% (v/v) glycerol, 30 mM imidazole, 0.1 mM PMSF). Cells were lysed for 30 min on ice by adding 1 mg/mL lysozyme followed by addition of sodium deoxycholate (12.5 mg/g cells) and further incubation for 15 min on ice. The solution was sonicated five times for 1 min each (intensity level 6, duty cycle 60%, Branson Sonifier) and centrifuged for 45 min at 30,000×*g*, 4 °C. The lysate was then subjected to heat denaturation at 75 °C for 15 min followed by centrifugation for 30 min at 30,000×*g*,

4 °C. For purification of Cbf5 alone, the heat denaturation step was omitted since we observed that this step rendered the protein inactive. The cleared lysate was loaded onto a 5 mL Ni^{2+} Sepharose column (GE Healthcare) using a BioLogic LP chromatography system (BioRad) and washed extensively with buffer A1. The protein was subsequently eluted with a linear gradient (50 mL) to buffer B (same as A except for 500 mM imidazole and no PMSF). For purification of Cbf5, glycerol was immediately added to fractions to a final concentration of 20% (v/v). Peak fractions were analyzed by 15% SDS-PAGE, pooled and concentrated by ultrafiltration (Vivaspin MWCO 30,000 or 10,000). Next, the protein was re-buffered either by ultrafiltration or by size exclusion chromatography using a Superdex 75 column (XK26/100 column, GE Healthcare) in buffer C (20 mM HEPES-KOH pH 7.5, 600 mM KCl, 1 mM EDTA, 20% (v/v) glycerol) at a flow rate of 1 mL/min (BioLogic DuoFlow chromatography system). Peak fractions were concentrated as before, flash frozen and stored in aliquots at -80 °C.

Protein concentration was determined photometrically at 280 nm using a molar extinction coefficient of 46,410 M^{-1} cm^{-1} for Cbf5, 9,970 M^{-1} cm^{-1} for Nop10, and 11,460 M^{-1} cm^{-1} for Gar1 (calculated using ProtParam (Gill and von Hippel 1989)), while concentration of Gar1 alone was determined at 210 nm using the extinction coefficient of 20.5 ml mg^{-1} cm^{-1}. The catalytically inactive mutant (D85N) of Cbf5 was purified either alone or in combination with Gar1, with Nop10, or both, essentially in the same way as explained above.

2.2.4 *In vitro* transcription and purification of tRNA and H/ACA guide RNA

A plasmid, called pIDT-Smart-PftRNA^{Asp}, encoding a T7 promoter followed by the gene for *P. furiosus* tRNA^{Asp} was purchased from Integrated DNA Technology. To generate H/ACA guide RNA, the sequence for *P. furiosus* H/ACA guide RNA Pf4 (5'-AAUGCCCCUCCCCUCUCACACCCCCGUGAGAAGUGAGCGGGGGGCGGUCGG GGAGGGGACAUCA-3') (Klein *et al.* 2002) as well as a T7 promoter was assembled using overlapping oligos and cloned into a pUC19 vector. The template for the *in vitro* transcription of tRNA^{Asp} and Pf4 guide RNA was generated by PCR amplification from the corresponding plasmids using methylated reverse primers to precisely terminate transcription (Sherlin *et al.* 2001). The *in vitro* transcription was performed using the PCR template (10% (v/v)) in transcription buffer (40 mM Tris-HCl pH 7.5, 15 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT) with 3 mM each ATP, CTP and GTP, and 0.1 mM [C5-^{3}H]UTP (23.9 Ci/mmole), 5 mM GMP, 0.01 U/μL inorganic pyrophosphatase, 0.3 μM T7 RNA polymerase and 0.12 U/μL RNase inhibitor at 30 °C for 4 h. Following the *in vitro* transcription, the template was digested with 2 U/mL DNaseI (Fermentas) for 1 h at 37 °C, and the RNA was purified with a Nucleobond AX100 column (Macherey-Nagel) using equilibration buffer R0 (100 mM Tris-acetate pH 6.3, 10 mM $MgCl_2$, 15% (v/v) ethanol), washing buffer R1 (R0 with 300 mM KCl) and elution buffer R3 (R0 with 1150 mM KCl). The RNA was concentrated by isopropanol precipitation and dissolved in H_2O. The tRNA concentration was determined photometrically at 260 nm using the extinction coefficient 5×10^5 M^{-1} cm^{-1}. The specific activity of the purified [^{3}H]tRNA^{Asp} and [^{3}H] guide RNA was determined by scintillation counting.

2.2.5 Nitrocellulose Filtration

Prior to all experiments, [^{3}H]tRNA^{Asp} and proteins were pre-incubated at 70 °C for 5 min. To allow the tRNA to bind to protein, 5 or 10 nM [^{3}H]tRNA^{Asp} was incubated with 0 - 700 nM protein or protein complex in reaction buffer for 10 min at 70 °C. The complete 50 μL reaction mixture was then filtered through a nitrocellulose membrane followed by washing of the membrane with 1 mL cold reaction buffer. Membranes were dissolved for 30 min in 10 mL EcoLite scintillation cocktail (EcoLite (+), MP Biomedical), and the amount of tRNA bound to the protein retained on the membrane was determined by scintillation counting (Perkin-Elmer Tri-Carb 2800TR liquid scintillation analyzer). In order to obtain the dissociation constant (K_D), the increase in the fraction of bound tRNA as a function of the protein concentration was analyzed by fitting to a quadratic equation (Wright *et al.* 2011) with [RNA] = 5 or 10 nM:

$$P_{bound} = Amp \times [(K_D + [R] + [P]) / 2 - \{(K_D + [R] + [P])^2 / 4 - [P] \times [R]\}^{0.5}]$$

Where P_{bound} is the percentage of bound tRNA, P and R are concentrations of protein and RNA, respectively, and Amp is the amplitude or final level of bound tRNA. Each titration was repeated at least three times; the K_D and its standard deviation were determined for each titration by fitting in Graphpad Prism. The average K_D including the largest standard deviation of individual titrations (which is larger than the standard deviation between the K_Ds of individual titrations) is reported in Table 2.2.

2.2.6 Tritium Release Assay

To identify the optimal conditions required for the *in vitro* pseudouridylation experiments, reactions were performed under multiple turnover conditions using 50 nM

Cbf5-Nop10 complex and 600 nM [^{3}H]tRNA^{Asp}. Beginning with the reaction buffer (section 2.2.1), the buffer was changed to determine the optimal concentrations of KCl, and $MgCl_2$, and the optimal pH. Besides these conditions, the effect of 50 ^oC or 70 ^oC temperatures was assessed. For Michaelis-Menten titrations, different concentrations of [^{3}H]tRNA^{Asp} (100 - 3000 nM) were incubated with 10 nM enzyme in reaction buffer plus 0.2% (w/v) bovine serum albumin at 70 °C. For single-turnover experiments, 600 nM [^{3}H]tRNA^{Asp} were incubated with 5 μM enzyme at 70 °C. Samples were removed at the desired time points and added to 5 % (w/v) activated charcoal (Norit A, EMD, CX0655-1) in 0.1 M HCl. Following centrifugation at 10,000×*g* for 2 min, the supernatant was added to 0.5 mL fresh 5 % (w/v) activated charcoal in 0.1 M HCl, mixed and centrifuged again. The supernatant was filtered through a glass wool plug in a 1 mL micropipet tip, and 0.8 mL of the resulting filtrate was then used for scintillation counting in 4 mL EcoLite scintillation cocktail. Each time course was repeated at least three times to determine the initial velocity, v_0, by linear fitting. The dependence of the initial rates v_0 on the tRNA concentration was analyzed by fitting the data in GraphPad Prism using the Michaelis-Menten equation $v_0 = v_{max}[S]/(K_M + [S])$, and the catalytic constant, k_{cat}, was determined by dividing v_{max} by the enzyme concentration (10 nM). The single-turnover experiments were analyzed by fitting the data to a single-exponential equation (Wright *et al.* 2011):

Pseudouridine = Amp – Amp x exp (-k_ψ x t)

Where Amp represents the amplitude and k_ψ is the single-turnover rate constant of ψ formation.

2.3 Results

2.3.1 Optimal conditions required for pseudouridylation reactions

In order to understand the contribution of proteins Nop10 and Gar1 on the pseudouridylation activity of Cbf5, a highly-purified *Pyrococcus furiosus* model system was used, similarly to previous studies (Baker *et al.* 2005; Charpentier *et al.* 2005; Li and Ye 2006). Each protein was individually expressed in *Escherichia coli*, and cells expressing the respective proteins were combined during cell opening to allow for formation of protein complexes. Subsequently, the individual proteins (Cbf5, Gar1) or protein complexes (Cbf5-Nop10, Cbf5-Nop10-Gar1) were purified by affinity and size-exclusion chromatography utilizing the hexa-histidine tag engineered onto the N-terminus of Cbf5 or Gar1. This purification strategy is essentially identical to previously published methods (Li and Ye 2006). All proteins were more than 95% pure as judged by SDS-PAGE. To study Cbf5's activity in modifying archaeal tRNA, *P. furiosus* $tRNA^{Asp}$, a substrate of Cbf5 (Roovers *et al.* 2006), was generated by *in vitro* transcription using [C5-^{3}H]UTP and subsequently purified by anion exchange chromatography.

Diverse buffer conditions have been used in pseudouridylation assays with Cbf5 (Baker *et al.* 2005; Charpentier *et al.* 2005; Roovers *et al.* 2006; Gurha *et al.* 2007). By identifying the common buffer components used in these studies (HEPES, KCl, MgCl2, and EDTA) a systematic characterization was performed to determine the optimal concentrations of the selected buffer components required for ψ formation by Cbf5 in *in vitro* reactions. The extent of pseudouridylation was determined using a well-established tritium release assay detecting liberation of tritium from the C5 of the uracil base upon

formation of the new C-C glycosidic bond (Cortese *et al.* 1974). Multiple turnover experiments using lower enzyme (50 nM Cbf5-Nop10) than substrate concentrations (600 nM [^{3}H]tRNA) were performed in the reaction buffer , while changing the concentration of a single component at a time. Under all conditions, the time courses displayed a gradual increase in ψ formation over time (Figure 2.1). Initial velocities (v_0) were determined from linear fits of the time courses in order to compare the effect of the varied component on the reaction rate. The wide range of KCl concentrations (0 mM - 1000 mM) tested in these reactions revealed that Cbf5-Nop10 can introduce ψ into tRNA even in the absence of KCl. The rate of reaction increases with addition of KCl up to a concentration at 150 mM, reaching about 1 nM min^{-1} under our assays conditions (Figure 2.1A). A further increase in the KCl concentration decreases the v_0 with no ψ formation observed at 600 mM or higher concentrations of KCl. Based on these observations, 150 mM KCl has been determined to be the optimal concentration for ψ formation by Cbf5-Nop10. Next, a titration with $MgCl_2$ showed only small amounts of ψ accumulation when no $MgCl_2$ was used, while an increase in $MgCl_2$ concentration also increases the v_0 values, which stay relatively constant between 1.5 mM to 10 mM $MgCl_2$ (Figure 2.1B).

Subsequently, experiments performed at different pH values (6 - 8) indicated that a neutral pH is optimal for these reactions with lower v_0 values recorded upon decreasing or increasing the pH (Figure 2.1C). Further, tritium release experiments were carried out at 50 ^{o}C and 70 ^{o}C to compare the effect of temperature on ψ formation. At 50 ^{o}C, the reaction progresses very slowly, resulting in less than 30% modification in 2 hours (Figure 2.1 A-C), when compared to the reaction at 70 ^{o}C that reached the same extent of ψ formation in less than 10 minutes (Figure 2.1 D).

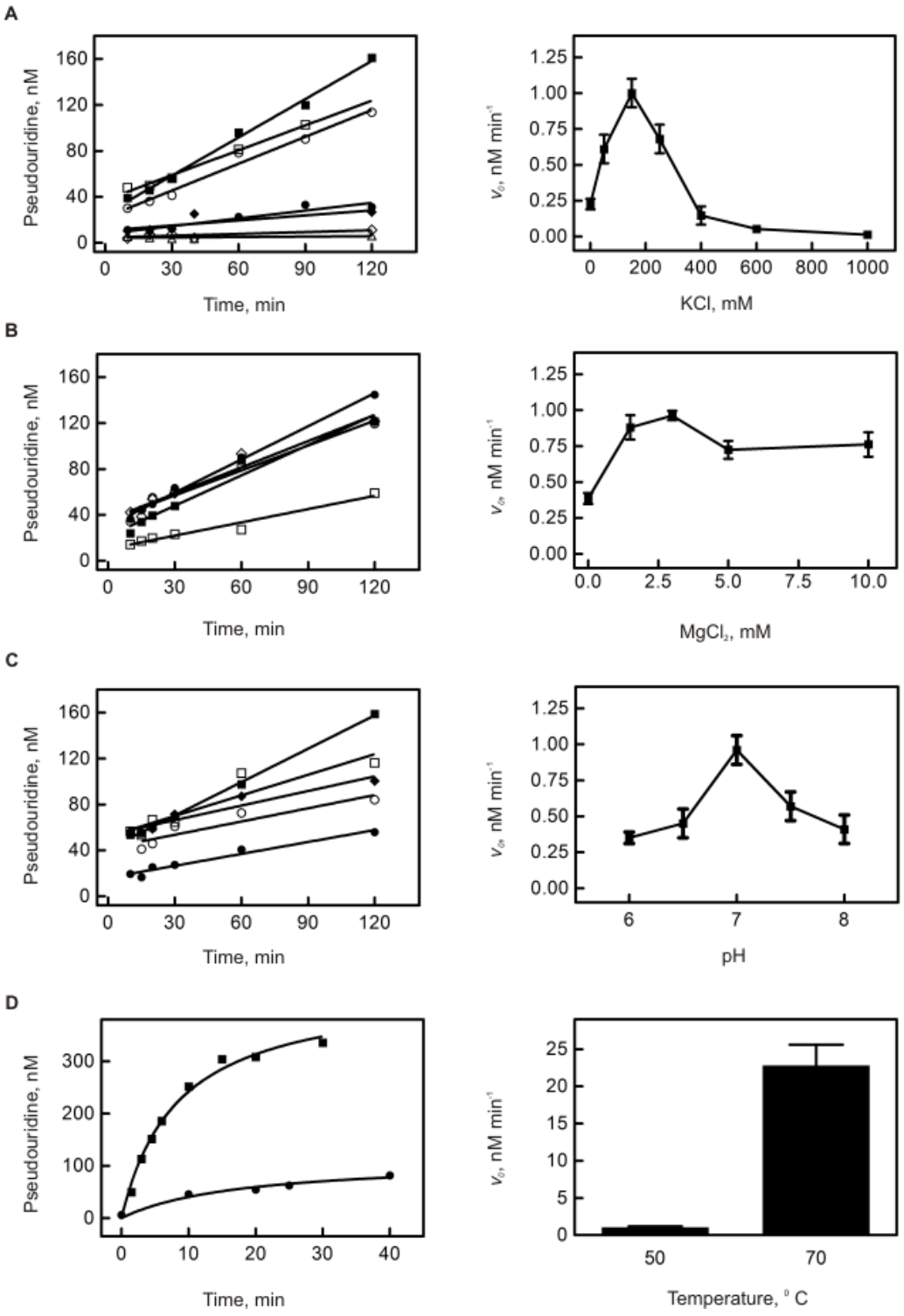


Figure 2.1. Optimization of buffer conditions for ψ formation using the Cbf5-Nop10 complex.

Tritium release assays were performed with 600 nM [^{3}H]tRNA and 50 nM Cbf5-Nop10 in the reaction buffer containing 20 mM HEPES (pH 7.0), 1.5 mM $MgCl_2$, 150 mM KCl

and 0.1 mM EDTA at 50 ºC, unless otherwise specified. ψ formation at different reaction conditions was studied. A) KCl titration: 0 mM (filled circles), 50 mM (open circles), 150 mM (filled squares), 250 mM (open squares), 400 mM (filled diamonds), 600 mM (open diamonds), 1000 mM (open triangles). B) $MgCl_2$ titration: 0 mM (open squares), 1.5 mM (filled squares), 3 mM (filled circles), 5 mM (open circles), 10 mM (open diamonds). C) pH titration: 6.0 (filled circles), 6.5 (open circles), 7.0 (filled squares), 7.5 (open squares), 8.0 (filled diamonds). D) ψ formation compared at different reaction temperatures: 50 ºC (circles) and 70 ºC (squares). Left side panels in the figure show the time courses and the right side panels show the reaction rates determined by fitting the time course to a linear equation. Smooth curves in left panel of figure D were obtained by fitting the time courses to a hyperbolic function.

This corresponds to a significant enhancement in v_0 (over twenty-fold) for increasing the temperature by 20 ºC. In accordance with our findings, 70 ºC has previously been shown to be the optimal temperature for *P. furiosus* Cbf5 (Roovers et al. 2006). In summary, these experiments identified the optimal conditions required for the *in vitro* pseudouridylation reactions by Cbf5 to be 150 mM KCl, pH 7, 1.5 to 10 mM $MgCl_2$ and 70 ºC.

2.3.2 Multiple-turnover catalysis of tRNA modification by Cbf5 in absence and presence of Nop10 and Gar1

To verify the activity of the *in vitro* reconstituted complexes, time courses of ψ formation were recorded at 70°C, and under multiple turnover conditions (Figure 2.2). No tritium was released from the tRNA under these conditions in the absence of proteins. Notably, more than 80% ψ formation was observed after 60 min for the Cbf5-Nop10-Gar1 complex as well as the Cbf5-Gar1 complex while the Cbf5-Nop10 complex yielded 60% ψ formation in 60 min. Interestingly, Cbf5 alone only reached about 20% pseudouridylation after 60 min of incubation. This might be attributed to the general tendency of free Cbf5 to precipitate as observed during purification. It is difficult to

visualize the precipitation in the smaller volumes used in the reaction, but it is likely that the unstable nature of Cbf5 on its own may explain its inability to complete the reaction under *in vitro* conditions, although other reasons cannot be excluded. In general, our findings are consistent with previous studies which showed an increasing activity of Cbf5 upon addition of Nop10 alone, Gar1 alone, or both Nop10 and Gar1, the latter Cbf5-Nop10-Gar1 complex representing the most active complex (Roovers *et al.* 2006; Gurha *et al.* 2007; Muller *et al.* 2007). However all these studies were conducted under single-turnover conditions using a large excess of enzyme over substrate. Therefore, our findings demonstrate for the first time that Cbf5 and its complexes with Nop10 and Gar1 are able to catalyze tRNA modification in a multiple-turnover fashion.

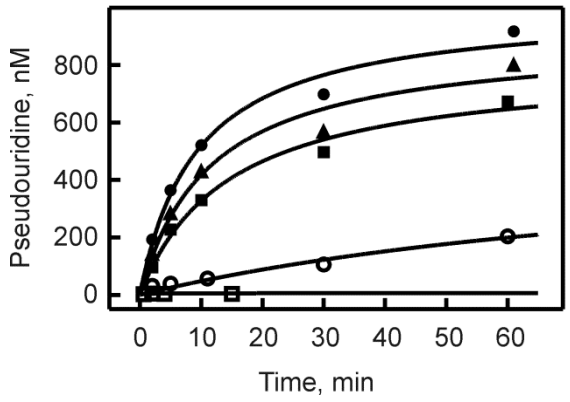


Figure 2.2. Time courses of ψ formation by Cbf5 alone and in the presence of Nop10 and Gar1.

1000 nM [^{3}H]tRNA was incubated at 70 °C with 10 nM Cbf5-Nop10-Gar1 (filled circles), Cbf5-Gar1 (filled triangles), Cbf5-Nop10 (filled squares) or Cbf5 alone (open circles). As a control, 1000 nM [^{3}H]tRNA was incubated in reaction buffer alone (open squares). The extent of ψ formation was quantified using the tritium release assay.

They furthermore reveal that all of the analyzed complexes are capable of efficient product release in contrast to the guide RNA-dependent function of Cbf5 where the reaction is limited to a single round of catalysis when Gar1 is absent, presumably since

the product RNA cannot dissociate from the H/ACA small ribonucleoprotein (Duan *et al.* 2009). In summary, we are using a highly-active, purified reconstituted *in vitro* system capable of multiple-turnover catalysis for studying pseudouridylation by Cbf5 in the presence and absence of Nop10 and Gar1.

2.3.3 Steady-state kinetic analysis of tRNA modification by Cbf5

In order to identify the role of Nop10 and Gar1 for tRNA modification by Cbf5, we have conducted steady-state kinetic experiments utilizing the fully active Cbf5-Nop10-Gar1 complex as well as complexes lacking either Nop10 or Gar1. We did not analyze the Cbf5 enzyme alone due to its limited activity (Figure 2.2). Based on these experiments we have determined the catalytic constants (k_{cat}) as well as the Michaelis constants (K_M), which respectively provide insights into catalysis and interaction with the substrate RNA. Based on the initial, linear phase of product formation using 10 nM enzyme, v_0 of the reaction could be determined by linear fitting (Figure 2.3A). The respective experiments were conducted at different tRNA concentrations ranging from 150 to 3000 nM to determine the dependence of the initial velocity on the substrate concentration (Figure 2.3B-D). Fitting to a Michaelis-Menten equation provided the steady-state kinetic parameters k_{cat} and K_M summarized in Table 2.2. Interestingly, all three analyzed complexes exhibited very similar behavior at low tRNA concentrations (< 300 nM tRNA). However, at higher tRNA concentration, the initial velocity of the Cbf5-Nop10 as well as the Cbf5-Gar1 catalyzed reaction did increase only very slightly (Figure 2.3B and C). In contrast, the Cbf5-Nop10-Gar1 complex showed a strong increase in initial velocity with higher substrate tRNA concentrations up to a velocity of 180 nM min^{-1} at 3000 nM tRNA without reaching saturation (Figure 2.3D). Thus, both Nop10 and

Gar1 contribute significantly to Cbf5's activity in particular at high substrate concentrations. This trend is confirmed by the quantitative analysis of the titrations, as the k_{cat} of the Cbf5-Nop10-Gar1 complex is 0.7 s^{-1}, more than three-fold higher than that of Cbf5-Gar1 (0.2 s^{-1}) and about six-fold larger than the k_{cat} of the Cbf5-Nop10 complex (0.11 s^{-1}).

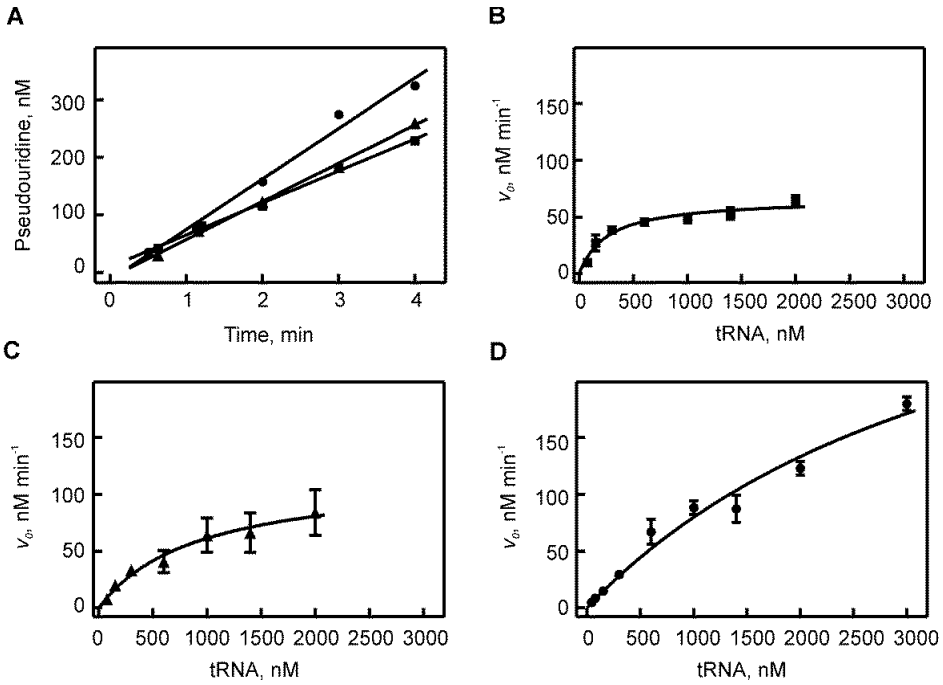


Figure 2.3. Steady-state kinetic analysis of pseudouridylation by the Cbf5-Nop10-Gar1 complex and subcomplexes thereof.

A. Short time courses using 1000 nM [^{3}H]tRNA and 10 nM enzyme to determine the initial velocity (v_0) of ψ formation by linear fitting. Cbf5-Nop10-Gar1 (circles), Cbf5-Gar1 (triangles), Cbf5-Nop10 (squares). Similar time courses were recorded at different tRNA concentrations, and the obtained initial velocities were plotted against the substrate concentration (B-D). Different complexes of Cbf5 were used as enzymes: Cbf5-Nop10 (B), Cbf5-Gar1 (C), and Cbf5-Nop10-Gar1 (D). Fitting to the Michaelis-Menten equation (smooth lines) yielded values for k_{cat} and K_M (see Table 2.2).

Interestingly, the effect of Nop10 and Gar1 on the Michaelis-Menten constant, K_M, is different than on k_{cat}. The K_M for tRNA decreases from about 4000 nM for the Cbf5-Nop10-Gar1 complex to 920 nM for Cbf5-Gar1 and 260 nM for Cbf5-Nop10. This is

surprising as at first view, this would suggest that the Cbf5-Nop10-Gar1 complex is less efficient in interacting with substrate tRNA than the partially assembled complexes.

Table 2.2. Kinetic parameters for tRNA modification by different Cbf5 complexes[a]

	K_M, nM	k_{cat}, s^{-1}	k_Ψ, s^{-1}
Cbf5-Nop10-Gar1	4000 ± 1700	0.7 ± 0.2	> 0.2
Cbf5-Gar1	920 ± 240	0.20 ± 0.03	0.06 ± 0.02
Cbf5-Nop10	260 ± 70	0.11 ± 0.01	0.07 ± 0.02
Cbf5	n.d.	n.d.	0.04 ± 0.01

n.d. – not determined

[a] Each value for K_M, k_{cat} and k_Ψ is reported together with the standard deviation obtained from data fitting as described in the Methods.

2.3.4 Substrate binding by Cbf5 alone and in complex with Nop10 and/or Gar1

In order to shed more light on the mechanism of substrate binding by Cbf5-Nop10-Gar1 and subcomplexes thereof, nitrocellulose filtration assays were performed to determine the dissociation constants (K_D) for $tRNA^{Asp}$ binding. To prevent modification of the bound tRNA, we have constructed a catalytically inactive Cbf5 variant by mutating the catalytic aspartate 85 to asparagine (D85N). This renders the protein completely inactive in pseudouridylation (Zebarjadian *et al.* 1999) (data not shown), while retaining its RNA binding abilities (see below). Subsequent to a 10 minute incubation of 10 nM [^{3}H]tRNA in the presence of excess protein, the reaction mixture was filtered through a nitrocellulose membrane that retains protein and protein-bound tRNA. After washing of

the membrane with reaction buffer, the amount of retained and therefore bound tRNA was determined by scintillation counting of the nitrocellulose filters. In order to assess the role of Nop10 and Gar1 for tRNA binding, we analyzed not only the Cbf5D85N-Nop10-Gar1, the Cbf5D85N-Nop10 and the Cbf5D85N-Gar1 complex, but also Cbf5D85N and Gar1 alone as Gar1 has been shown to bind RNA (Bagni and Lapeyre 1998). For all proteins and protein complexes tested, about 80% of the tRNA was bound to protein at high protein concentrations (Figure 2.4). Gar1 bound tRNA comparatively weakly (K_D of 750 nM, Table 2.3), and Cbf5D85N alone bound tRNA with an intermediate affinity (K_D = 235 nM). However, all other complexes of Cbf5D85N with Nop10 and/or Gar1 displayed a high affinity for tRNA ranging from 45 – 80 nM (Table 2.3). In comparison to Cbf5D85N alone, these results clearly show that both Nop10 and Gar1 enhance Cbf5's ability to bind tRNA to a similar extent. Furthermore, these experiments demonstrate that the complete Cbf5-Nop10-Gar1 complex is fully capable of tight binding to the substrate tRNA despite its high K_M (Figure 2.3 and Table 2.2).

The large difference between the K_D and the K_M can be explained with the high catalytic activity of the Cbf5-Nop10-Gar1 complex. For a relatively simple mechanism assuming single-step binding (k_1, k_{-1}) followed by catalysis (k_2), Briggs-Haldane showed that the K_M is (k_{-1} + k_2) / k_1 (Fersht 1998). Only if the catalytic rate constant (k_2) is low compared to dissociation of substrate (k_{-1}) is the K_D (= k_{-1}/k_1) equal to the K_M, (Fersht 1998). This is not the case for most enzymes including Cbf5-Nop10-Gar1 where K_M is influenced not only by the rate constants for substrate binding (k_1, k_{-1}), but also by the rate constant of catalysis (k_2) or other subsequent steps. While the exact kinetic mechanism of Cbf5-Nop10-Gar1 in modifying tRNA is not known, our data are

consistent with the Briggs-Haldane description of K_M. As the Cbf5-Nop10-Gar1 complex has a relatively high catalytic constant, k_{cat}, which most likely reflects the rate constant of catalysis, k_2, it is expected that K_M increases with k_2. In contrast, the catalytic constant, k_{cat}, is rather low for the Cbf5-Nop10 complex, and hence its K_M value is of a similar order of magnitude as the K_D, *i.e.* the Cbf5-Nop10 complex might follow the Michaelis-Menten mechanism. In conclusion, the high K_M of the Cbf5-Nop10-Gar1 complex seems to be a result of its high k_{cat} value, or in other words the high catalytic activity of the Cbf5-Nop10-Gar1 complex is achieved by "sacrificing" the K_M value for tRNA. Notably, this property does not necessarily have to apply to the complete H/ACA small ribonucleoprotein as it employs a different mechanism for substrate RNA binding based on guide RNA.

Table 2.3. Affinity of Cbf5 complexes to substrate and product tRNA[a]

	K_D, nM (substrate tRNA)	**K_D, nM (product tRNA)**
Cbf5	235 ± 65	n.d.
Cbf5-Nop10-Gar1	45 ± 20	27 ± 10
Cbf5-Gar1	80 ± 25	105 ± 25
Cbf5-Nop10	50 ± 15	60 ± 20
Gar1	750 ± 300	n.d.

n.d. – not determined[a] Each K_D is the average of at least three different nitrocellulose filtration experiments titrating protein against tRNA. Each average K_D value is reported together with the largest standard deviation from individual filtration experiments.

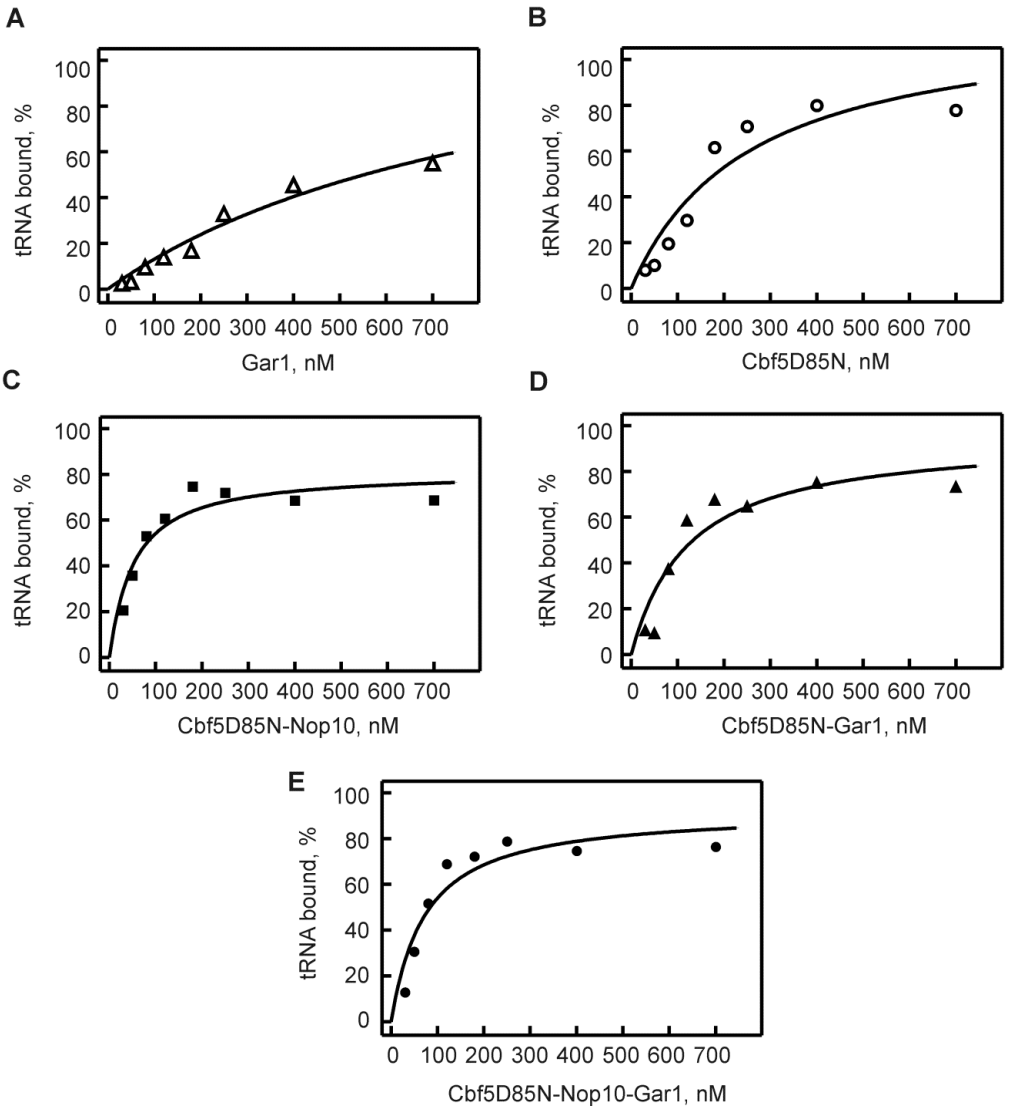


Figure 2.4. Substrate tRNA binding by Cbf5 in the presence and absence of Nop10 and Gar1.

To determine the affinity of Cbf5 and Cbf5 complexes for unmodified substrate tRNA, [^{3}H]tRNA was incubated with increasing concentrations of the catalytically inactive Cbf5 D85N variant and associated proteins. The percentage of bound tRNA was recorded by nitrocellulose filtration and scintillation counting. The experiment was carried out with Gar1 alone (A), Cbf5D85N alone (B), Cbf5D85N-Nop10 (C), Cbf5D85N-Gar1 (D), and Cbf5D85N-Nop10-Gar1 (E). Fitting to a quadratic function (Materials and Methods, smooth lines) provided the dissociation constant, K_D, for the interaction of Cbf5 and its complexes with substrate tRNA (see Table 2.3). Here, individual titrations are shown, but each experiment was repeated at least three times to determine the dissociation constants, K_D, reported in Table 2.3.

2.3.5 Single-turnover tRNA modification by Cbf5 and Cbf5 complexes

Next, we asked whether Gar1 could influence the release of product tRNA as it has been implicated in product release during the guide-dependent activity of the H/ACA small ribonucleoprotein (Duan *et al.* 2009). For this purpose, pseudouridylation assays were conducted under single-turnover conditions, *i.e.* with an excess of enzyme (5 μM) over [^{3}H]tRNA (0.6 μM). Under these conditions, the tritium release assays detect the appearance of the enzyme-product complex as the active site is accessible to water and the released tritium can easily escape the active site. Therefore, the measured rate constant is independent of product release in contrast to the k_{cat} measured under multiple round conditions. If product release is rate-limiting under multiple turnover conditions, for example upon omission of Gar1, the k_{cat} would be lower than the single-round rate constant of ψ formation (k_Ψ). It is therefore the aim of these single-round experiments to assess whether product release is limiting by comparing k_Ψ and k_{cat}. For these experiments, very short time courses have to be measured as the reaction is expected to be rather fast. Usually we would achieve this by using the rapid-mixing quench flow apparatus; however, this is not feasible at 70 ºC. Therefore, the experiments were performed by hand allowing at least a rough estimation of single-round pseudouridylation rate constants (k_Ψ). Again, all Cbf5 complexes with Nop10 and/or Gar1 achieved 80% or more product formation in a short time (Figure 2.5). Interestingly, Cbf5 alone was able to form ψs with a rate of 0.04 s^{-1} under these conditions, but failed to convert more than 30% of all tRNAs, which again might be explained by an instability of Cbf5 during the course of the experiment. As expected based on the k_{cat}, the complete Cbf5-Nop10-Gar1 complex had converted all substrate to product within the first 10 seconds, thus indicating

that the single-round rate constant is at least 0.2 s^{-1} or larger. Interestingly, upon omitting Gar1, the Cbf5-Nop10 complex displayed a single-round rate constant of 0.07 s^{-1} which is very similar to the k_{cat} value of 0.11 s^{-1} for this complex given the precision of the measurements (Table 2.2). This clearly demonstrates that product release is fast for the Cbf5-Nop10 complex. Therefore, Gar1 is not involved in tRNA product release in contrast to its function in the guide-dependent reaction (Duan *et al.* 2009). For the Cbf5-Gar1 complex, the single-round rate constant is 0.06 s^{-1} and therefore also in a comparable magnitude to the k_{cat} (Table 2.2). This indicates that Nop10 is also not involved in product release.

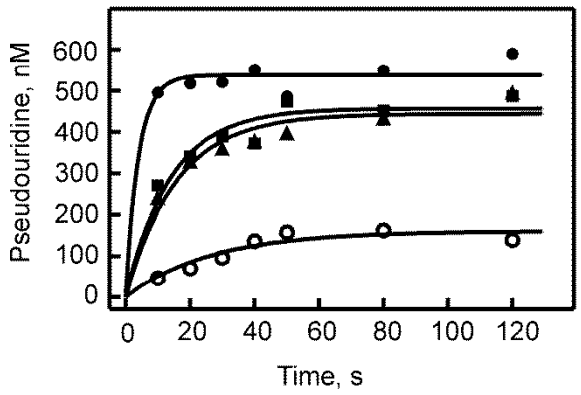


Figure 2.5. Effect of Nop10 and Gar1 on a single-round of ψ formation by Cbf5.

0.6 μM of [^{3}H]tRNA was incubated with 5 μM of Cbf5 and accessory proteins at 70°C, and ψ formation was determined using the tritium release assay. Under these conditions, each Cbf5 can only modify a single tRNA. The tRNA was reacted with Cbf5-Nop10-Gar1 (filled circles), Cbf5-Gar1 (filled triangles), Cbf5-Nop10 (filled squares), or Cbf5 alone (open circles). The time courses were fit to a single-exponential equation to estimate the single-turnover rate constant of ψ formation, k_ψ (Table 2.2).

2.3.6 Interaction of Cbf5 complexes with modified product tRNA and H/ACA guide RNA

Based on previous studies reporting that ψ synthases can bind modified product tRNA (Ramamurthy *et al.* 1999), we next examined whether this is also the case for Cbf5. To this end, the nitrocellulose filtration assays with [^{3}H]tRNA were repeated in the presence of active, wild-type Cbf5 in complex with Nop10 and/or Gar1. As shown in the single-turnover pseudouridylation assay (Figure 2.5), all uridines should be converted to ψs by the Cbf5-Nop10, Cbf5-Gar1 and Cbf5-Nop10-Gar1 complexes during the 10 minute incubation period at 70 ºC allowing the measurement of modified product tRNA binding. Interestingly, all Cbf5 complexes again displayed relatively tight tRNA binding reaching maximal binding at protein concentrations between 100 - 200 nM (Figure 2.6). Fitting of the data revealed the dissociation constants (K_D) as summarized in Table 2.3. The comparison to the respective affinities for unmodified substrate tRNA reveals that Cbf5 complexes with Nop10 and/or Gar1 bind with similar affinities to substrate and product tRNA. Notably, tight binding of the product tRNA does not exclude rapid product release; instead it is likely that product binding is a dynamic equilibrium with rapid dissociation and re-association of the tRNA.

Lastly, we asked how the interaction of the Cbf5-Nop10-Gar1 complex with tRNA compares to its interaction with H/ACA guide RNA as it occurs in the archaeal cell. Therefore, [^{3}H]-labeled H/ACA guide RNA Pf4 (Klein *et al.* 2002) was prepared and used in nitrocellulose filtration assays with Cbf5-Nop10-Gar1. The titration revealed that H/ACA guide RNA binds tightly to Cbf5-Nop10-Gar1 reaching the end level already at 100 nM of protein. The dissociation constant for the interaction of Cbf5-Nop10-Gar1

with H/ACA guide RNA Pf4 is 21 ± 8 nM as determined in three independent experiments. Hence, Cbf5-Nop10-Gar1 binds guide RNA as tightly as tRNA (see Table 2.3).

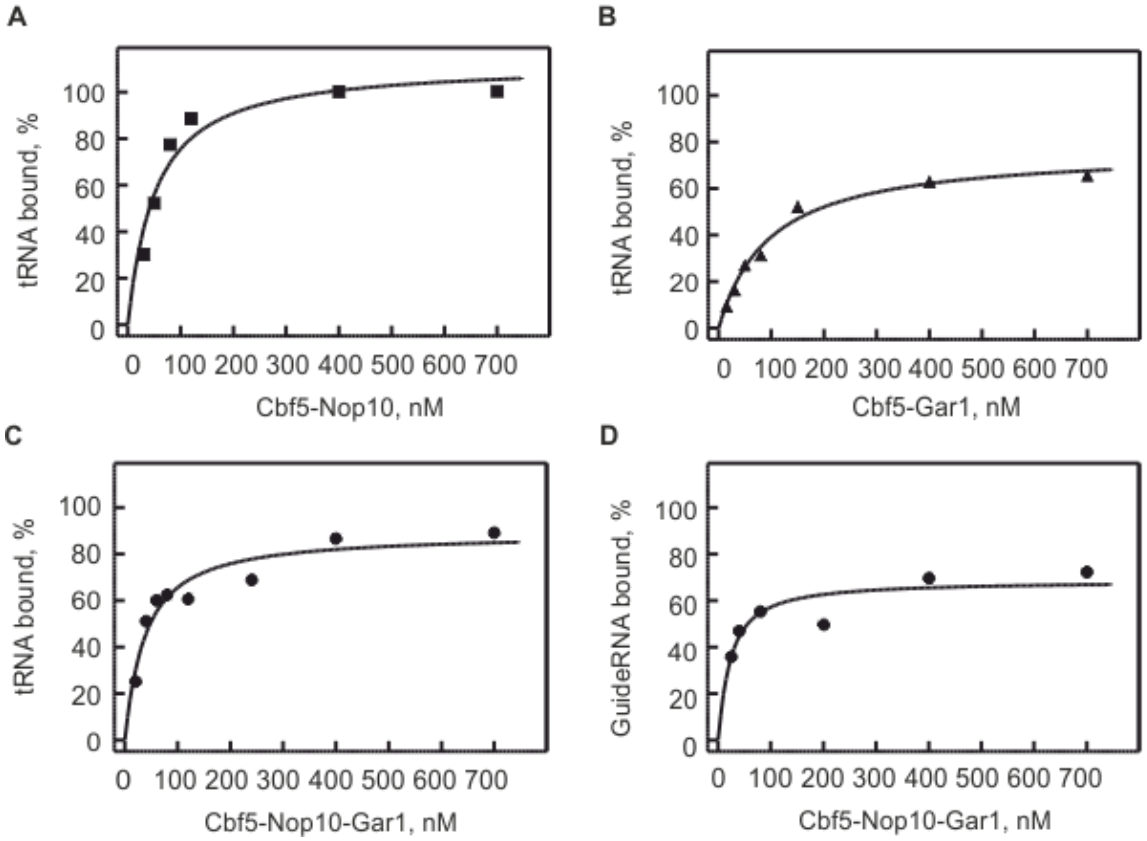


Figure 2.6. Binding of modified product tRNA in comparison to H/ACA guide RNA by Cbf5 in presence and absence of Nop10 and Gar1.

[^{3}H]tRNA or [^{3}H] H/ACA guide RNA was incubated at 70 °C for 10 minutes with increasing concentrations of wild-type, active Cbf5-Nop10-Gar1 as well as complexes missing Nop10 or Gar1 followed by nitrocellulose filtration and scintillation counting to determine the percentage of bound product tRNA. The tRNA experiment was performed with Cbf5-Nop10 (A), Cbf5-Gar1 (B), and Cbf5-Nop10-Gar1 (C). The H/ACA guide RNA was bound to Cbf5-Nop10-Gar1 (D). Smooth lines are the result of fitting to a quadratic function yielding the dissociation constants, K_D, for product tRNA binding (see Table 2.3) and H/ACA guide RNA binding (21 ± 8 nM). Again individual, representative titrations are shown.

2.4 Discussion

Here, we present the first quantitative analysis of guide-independent ψ formation by archaeal Cbf5 in the presence and absence of its accessory proteins Nop10 and Gar1. Our findings demonstrate that both Nop10 and Gar1 enhance Cbf5's catalytic activity. Furthermore, they improve Cbf5's interaction with its substrate tRNA. In contrast to the guide-dependent reaction, Gar1 does not affect product release by Cbf5. All Cbf5 complexes are capable of tight binding to both the substrate and the product tRNA. These quantitative findings allow for the first time a detailed insight into the role of the accessory proteins Nop10 and Gar1.

Our results unambiguously show that lack of either Nop10 or Gar1 from the full Cbf5-Nop10-Gar1 complex reduces the catalytic constant, k_{cat}, revealing a role of Nop10 and Gar1 in enhancing the catalytic ability of Cbf5. In general, the active site of ψ synthases contains three residues that have been implicated in catalysis: an aspartate that is essential for catalysis as well as a tyrosine (phenylalanine in TruD) and an arginine or lysine interacting with the catalytic aspartate (Hamma and Ferré-D'Amaré 2006). Our findings raise the question of how Nop10 and Gar1 can influence the active site of Cbf5.

The different crystal structures of Cbf5-Nop10-Gar1 support the hypothesis that Nop10 and Gar1 may influence all three active site residues of Cbf5 and may contribute to positioning of the substrate tRNA. As seen in the crystal structures of Cbf5-Nop10-Gar1, Nop10 binds in the vicinity of Cbf5's active site whereas Gar1 can interact with Cbf5's thumb loop, but is not close to the active site of Cbf5 (Figure 2.7) (Li and Ye 2006; Rashid *et al.* 2006). Based on these structural constraints, it is highly unlikely that

either Nop10 or Gar1 contribute a residue directly to the active site which is also in accordance with the observation that significant catalytic activity is retained upon loss of Nop10 or Gar1. Instead, we hypothesize that Nop10 and Gar1 are indirectly influencing Cbf5's activity. For Nop10, it has already been proposed based on the crystal structures that it stabilizes the active site of Cbf5 (Hamma *et al.* 2005). Nop10's linker region directly interacts through a so-called proline spine with the conserved motif I in Cbf5 which is located next to the active site and contacts the catalytic aspartate residue (Hamma and Ferré-D'Amaré 2010). Additionally, the N-terminal domain, specifically the conserved tyrosine 14 of Nop10 (*P. furiosus* numbering as in Duan *et al.* 2009), contacts the conserved valine 114 in β4 of Cbf5 which is next to the conserved tyrosine 113 residue that has been implicated in catalysis (Hamma *et al.* 2005). Hence the effect of Nop10 on Cbf5's catalytic ability might result from a stabilization of motif I in Cbf5 and β4 thereby correctly positioning the active site residues aspartate 85 and tyrosine 113 (Figure 2.7). Gar1 contacts the C-terminus of Cbf5's helix 5 which contains arginine 184 at its N-terminus, the third of the active site residues. Additionally, Gar1 can interact with Cbf5's thumb loop in the so-called open conformation (Li and Ye 2006) and maintains interactions with Cbf5's strand β7 preceding the thumb loop in the closed conformation (Duan *et al.* 2009). As the thumb loop interacts with substrate RNA in presence of guide RNA (Li and Ye 2006), it can be envisioned that Gar1's interaction with β7 could also help to correctly position tRNA in Cbf5's active site (Figure 2.7). Thus, Gar1 could influence the active site geometry of Cbf5 by correctly positioning helix 5 of Cbf5 and thereby the catalytic arginine, and it could indirectly enhance catalysis by substrate positioning with the help of the thumb loop.

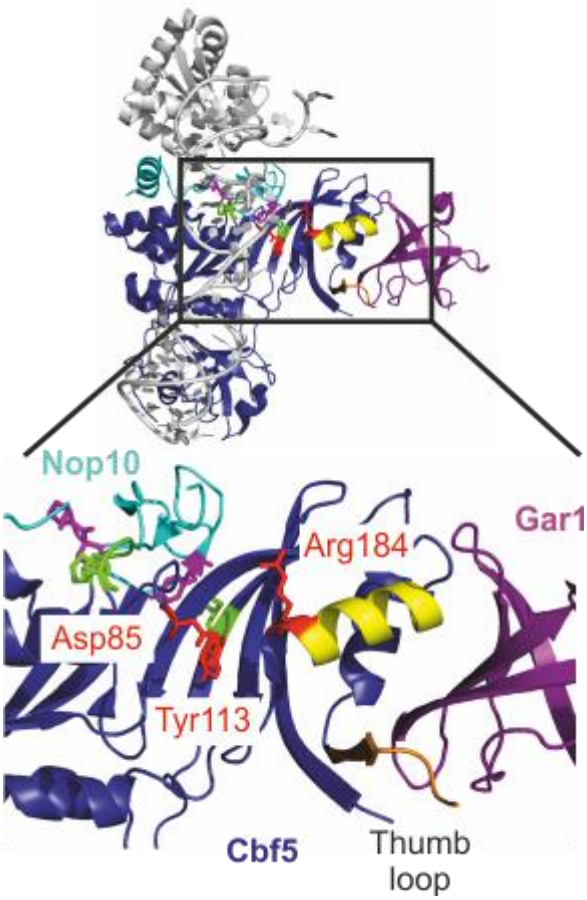


Figure 2.7. Contacts between Nop10 and Gar1 and the active site of Cbf5.

The upper panel shows the structure of the complete H/ACA small ribonucleoprotein with guide RNA and L7Ae depicted in grey (PDB ID: 2HVY (Li and Ye 2006)); the active site of Cbf5 as shown below is indicated by the boxed area. The active site residues of Cbf5 (Asp85, Tyr113, and Arg184) are shown in red. Nop10 is depicted in cyan and Gar1 in purple. Residues of Nop10 indicated in pink are in contact with Cbf5 residues shown in green that are in the direct neighborhood of the active site residues Asp85 and Tyr113. Gar1 contacts helix 5 of Cbf5 (yellow) that contains the active site Arg184 at its N-terminus; furthermore Gar1 interacts with Cbf5's β7 strand preceding the thumb loop (orange) which can interact with the substrate RNA. These contacts can potentially contribute to the stabilization of Cbf5's active site by Nop10 and Gar1. The figure was prepared using PyMol (DeLano Scientific LLC, 2006).

The finding that Nop10 and Gar1 enhance Cbf5's catalytic activity during tRNA modification likely also applies to the guide-dependent pseudouridylation by Cbf5. Both the guide-dependent and the guide-independent reaction analyzed here are taking place in

the same active site of Cbf5; and Nop10 and Gar1 interact in the same way with Cbf5 in the absence and presence of guide RNA as is evident upon comparing the isolated Cbf5-Nop10-Gar1 structure and the full H/ACA small ribonucleoprotein (Li and Ye 2006; Rashid *et al.* 2006), Therefore, we hypothesize that the roles of Nop10 and Gar1 in stabilizing Cbf5's active site during catalysis also hold true for the guide-dependent reaction. Notably, it would not have been possible to identify these functions of Nop10 and Gar1 by studying the guide-dependent reaction as lack of Nop10 completely inhibits ψ formation (Baker *et al.* 2005; Charpentier *et al.* 2005) and lack of Gar1 limits the reaction to a single round (Duan *et al.* 2009). Our findings do not exclude other roles of Nop10 and Gar1 in the guide-dependent reaction in particular for substrate RNA binding and product release which might be substantially different from Cbf5's interactions with tRNA.

Notably, Cbf5-Nop10-Gar1 displays a similar catalytic constant of about 0.7 s^{-1} to the rate constants of pseudouridylation by bacterial ψ synthases TruB, TruA and RluA (0.35 to 0.7 s^{-1}) (Wright *et al.* 2011). It has been previously discussed that this relatively low catalytic rate constant will most likely apply to all bacterial stand-alone ψ synthases. The findings for Cbf5-Nop10-Gar1 now suggest that uniform slow catalysis is a general feature of ψ synthases that holds true also for complex ψ synthases such as Cbf5. Possibly, such a slow rate of catalysis is a result of the chemical mechanism required for ψ formation. Pseudouridylation consists of at least cleavage of the glycosidic bond, rotation of the uracil base and formation of the new C-C glycosidic bond, and this reaction is presumably catalyzed by the same mechanism in all ψ synthases sharing a conserved catalytic domain and conserved active site residues (Hamma and Ferré-

D'Amaré 2006). As this is a chemically complex reaction, it might not be possible to enhance ψ formation to more than 0.35 – 0.7 s^{-1}.

The nitrocellulose filtration assays reveal high-affinity equilibrium binding constants (K_D) in the low nanomolar range for both substrate and product tRNA and all Cbf5 complexes. Both Nop10 and Gar1 are able to enhance Cbf5's ability to bind tRNA as the K_Ds for the protein complexes are between 27 and 105 nM while Cbf5 alone binds tRNA with a K_D of 235 nM (Table 2.3). Again, this improved tRNA binding might be a result of the overall stabilization of Cbf5 by Nop10 and Gar1. Furthermore, Nop10 could directly contribute to tRNA binding as it also forms contacts to the guide RNA in the H/ACA small ribonucleoprotein (Li and Ye 2006). Gar1 does not contact the guide RNA and might not be directly involved in tRNA binding; accordingly, at least for product binding, Gar1's effect on the affinity of Cbf5 for tRNA seems to be smaller than the effect of Nop10 (Table 2.3). Overall, our findings suggest that, although tRNA is not the *in vivo* substrate of Cbf5, Cbf5 has retained the ability to interact with tRNA similarly as its bacterial homologue TruB. Interestingly, Cbf5 alone and its complexes display an even higher affinity for tRNA than TruB (Ramamurthy *et al.* 1999; Wright *et al.* 2011). Based on a structural comparison of Cbf5 with TruB and its interaction with tRNA (Hoang and Ferré-D'Amaré 2001; Li and Ye 2006), it is likely that the tRNA binding site on Cbf5 overlaps with the guide RNA binding site. As this guide RNA binding site of Cbf5 is rather large, the high tRNA affinity of Cbf5 might reflect the ability of Cbf5 to tightly bind H/ACA guide RNA which has an equally high affinity (Figure 2.6D). Furthermore, it is not surprising that binding of modified product tRNA to the Cbf5 complex is very similar to binding of unmodified substrate tRNA as the introduction of ψ represents a

relatively minor change to the overall tRNA structure. Similarly, binding of product tRNA has previously been observed for TruB (Ramamurthy *et al.* 1999).

In contrast to Gar1's function during the guide-dependent reaction, Gar1 is not involved in product release from Cbf5 for the guide-independent modification of tRNA (Duan *et al.* 2009). Our data clearly show that multiple rounds of catalysis can occur rapidly in the absence of Gar1, *i.e.* for the Cbf5-Nop10 complex. Also, the single-round rate constant of catalysis, k_{ψ}, is similar to the multiple round catalytic constant, k_{cat}, for the Cbf5-Nop10 complex indicating that product release is not rate-limiting. In fact, product release is also rapid for Cbf5-Nop10-Gar1 and Cbf5-Nop10, *i.e.* rapid tRNA release seems to be a general feature of the guide-independent reaction. This differential function of Gar1 for product release in the guide-dependent and -independent RNA modification can best be explained by a different mode of substrate binding. In the presence of a guide RNA, the substrate RNA is held in place through several base-pairs. In contrast, the tRNA directly interacts with the proteins, predominantly Cbf5 and maybe Nop10, and these contacts might be easier to break during release of the product tRNA.

In summary, the first quantitative analysis of ψ formation by Cbf5-Nop10-Gar1 reported here reveals that both Nop10 and Gar1 can stabilize the active site of Cbf5 thereby enhancing its catalytic activity. We hypothesize that this is a general feature of Nop10 and Gar1 which could also indirectly contribute to catalysis during the guide-dependent reaction. Furthermore, we demonstrate for the first time that Cbf5-Nop10-Gar1 complexes have very high affinities for tRNA in the low nanomolar range, but are capable of rapidly releasing modified product tRNA. As Cbf5-Nop10-Gar1 displays an equally high affinity to H/ACA guide RNA as to tRNA, we suggest that Cbf5-Nop10-

Gar1 might be mostly found bound to guide RNA in the archaeal cell and might therefore not be available for modifying tRNA which is instead catalyzed by Pus10 *in vivo*. This quantitative characterization of the complex archaeal ψ synthase Cbf5 in tRNA modification paves the way for further studies into the mechanism of guide-RNA dependent ψ formation by the H/ACA small ribonucleoprotein complex.

Chapter 3: Structure-guided functional analysis of archaeal Pus10 in pseudouridine 55 formation in tRNA

3.1 Introduction

tRNAs in all domains of life contain ψ at position 55. ψ55 in tRNA appears to be highly conserved due to its structural role in allowing the tertiary base-pairing with the conserved G18, which in turn facilitates the stacking of neighboring m^5U54 and conserved purine at position 57 (Romby *et al.* 1987). In addition to this well established structural role, the Björk group has demonstrated that preventing the modification of ψ55 reduces expression of several virulence genes in the human pathogen, *Shigella flexneri* (Urbonavicius *et al.* 2002). Their work also shows reduced growth rates and defects in translation of certain codons in *E. coli* cells when ψ55 was excluded from tRNA in combination with two other modifications (Gm18, m^5U54). In bacteria, TruB is responsible for converting uridine to ψ at this position (Nurse *et al.* 1995), while Pus4, a TruB homolog, performs this modification in eukaryotes (Becker *et al.* 1997). Conversely, Cbf5, another homolog of TruB found in archaea and eukaryotes, is shown to be involved in modifying rRNA as part of the H/ACA small ribonucleoprotein complex, which contains three accessory proteins (Nop10, Gar1, and L7Ae) and a guide RNA (Baker *et al.* 2005; Charpentier *et al.* 2005). Notably, archaeal Cbf5 also has the ability to perform ψ55 modification in tRNA under *in vitro* conditions (Roovers *et al.* 2006). This function of tRNA modification by archaeal Cbf5 is guide-RNA-independent, but requires Nop10 and Gar1 for optimal activity (Roovers *et al.* 2006; Gurha *et al.* 2007).

Besides Cbf5, another protein called PsuX (later named Pus10) was identified in archaeal organisms while searching for Cbf5 homologs (Watanabe and Gray 2000). This protein was also demonstrated to modify ψ55 in archaeal tRNAs (Roovers *et al.* 2006). Pus10 is reported to have the ability to introduce ψ54 as well in tRNA, although with varied efficiencies and in a salt-dependent manner (Gurha and Gupta 2008). Interestingly, Pus10 homologs are also found in a few eukaryotes such as *Caenorhabditis elegans*, *Drosophila Melanogaster*, mouse, plants and humans, but not in yeast (Watanabe and Gray 2000; McCleverty *et al.* 2007). It might be that the Pus10 gene has been acquired in these eukaryotes through horizontal gene transfer. No Pus10-related sequences are observed in bacteria (Watanabe and Gray 2000). The discovery of Pus10 revealed a redundant system in archaea for tRNA modification at position 55. It was found that Pus10 can efficiently complement for *truB* in the null strain of *E. coli* supporting a likely role of Pus10 in ψ55 formation *in vivo* (Roovers *et al.* 2006). To provide direct evidence for the involvement of Pus10 in ψ55 formation in the cell, Dr. Gupta's group in collaboration with others conducted *in vivo* studies using a *Haloferax volcanii* Cbf5-deletion strain (Blaby *et al.* 2011). Based on the observation of ψ55 modification in the Cbf5 deletion strain, they concluded that it is Pus10, not Cbf5 that introduces ψ55 in archaea. However, it is not clear what prevents Cbf5 from performing this modification in the cell. It is important to note that in this work, the authors showed that Pus10 is essential for this halophile, as the strain with a chromosomal deletion of Pus10 could only grow in the presence of plasmid-encoded Pus10 expressed under the regulation of an inducible promoter.

Due to the lack of sequence similarity to other known ψ synthases, Pus10 was placed in its own new family *i.e.* the 6th family of ψ synthases (Watanabe and Gray 2000). The crystal structure of human Pus10 was solved in the absence of the substrate (McCleverty *et al.* 2007). Pus10 is a crescent-shaped molecule with a C-terminal catalytic domain and an accessory domain at its N-terminus (Figure 3.1). Many ψ synthases contain an accessory domain that is implicated in binding the substrate RNA. It was proposed that the accessory domain in archaeal Pus10, the so-called THUMP domain (named after its existence in **Th**io**u**ridine synthases, **M**ethylases, and **P**seudouridine synthases), may also play a similar role in tRNA binding (Aravind and Koonin 2001). The crystal structure revealed that the N-terminal domain of Pus10 is composed of nine α-helices arranged in two bundles of three helices each, a mixed three-stranded β-sheet packing against two of the remaining α-helices, and several loops connecting these secondary structures (Figure 3.1). Although this domain exhibits high sequence variation, there are two conserved features identified. First, this domain was observed to bind a zinc ion through tetrahedral coordination by four highly conserved cysteines, at positions 21, 24, 109, and 112 in human Pus10 (corresponding to 16, 19, 57, and 60 in *P. furiosus*). This motif is conserved in all Pus10 sequences, displaying a $CX_2CX_{84}CX_2C$ sequence in human Pus10, with standard spacing within each cysteine pair, while the spacing between the pairs appears to be variable (e.g. in *P. furiosus* it is $CX_2CX_{37}CX_2C$). Pus1, another ψ synthase from yeast, was also shown to contain a zinc-binding motif, in addition to archaeal Nop10, an accessory protein found in the H/ACA sRNP complex. Second, the N-terminal domain of Pus10 possesses several conserved positively charged residues forming a basic cleft that is placed opposite to the putative active site in the C-terminal domain. It was

predicted that this conserved patch of basic residues may be involved in mediating the interaction with tRNA (McCleverty *et al.* 2007). Given the unique presence of the THUMP domain in Pus10 among all ψ synthases, it is not clear if this domain is involved in binding the tRNA. In particular the role of the conserved basic residues and the function of the zinc-binding motif are unknown. It is also not clear if this domain is essential for the function of Pus10.

Despite poor sequence similarity of Pus10 to other ψ synthases, its catalytic domain displays all the conserved motifs found in ψ synthases, including conserved active site residues (McCleverty *et al.* 2007). The C-terminal domain displays an extended platform created by a set of antiparallel, one four-stranded and one five-stranded, β-sheets connected through parallel strands, flanked by two α-helices (Figure 3.1). Four more helices were also found packing against this platform. The crystal structure shows the active site located in a deep basic pocket lined by the conserved sequence motifs containing the catalytic aspartate (210, according to *P. furiosus* numbering), and a highly conserved arginine (208) from motif II, tyrosine (274) from motif IIa, a basic residue (348) from motif III, and a leucine (375) from motif IIIa (Figure A1). Like other ψ synthases, Pus10 also contains two RNA-binding loops, called the forefinger loop and thumb loop. As in several ψ synthases, these loops are located adjacent to the active site of the enzyme and are involved in stabilizing the interaction of substrate with the enzyme (Hamma and Ferré-D'Amaré 2006). In *P. furiosus,* the thumb loop of Pus10 contains several conserved positively charged residues adding to the highly basic nature of the catalytic pocket. In contrast, the forefinger loop appears to have low sequence conservation (Figure A1) with only a single basic residue conserved in eukaryotic

sequences. Based on a model created by docking tRNA onto Pus10, it was proposed that the TψC loop carrying the target uridine is placed into the basic cleft of the catalytic domain by accommodating the tRNA acceptor stem in another basic cleft between the C- and N-terminal domains, while the 3' end is engaged by the positively charged amino acids in the THUMP domain (McCleverty *et al.* 2007). Except for the predictions made using this model, our knowledge on the interaction of tRNA with Pus10 is limited.

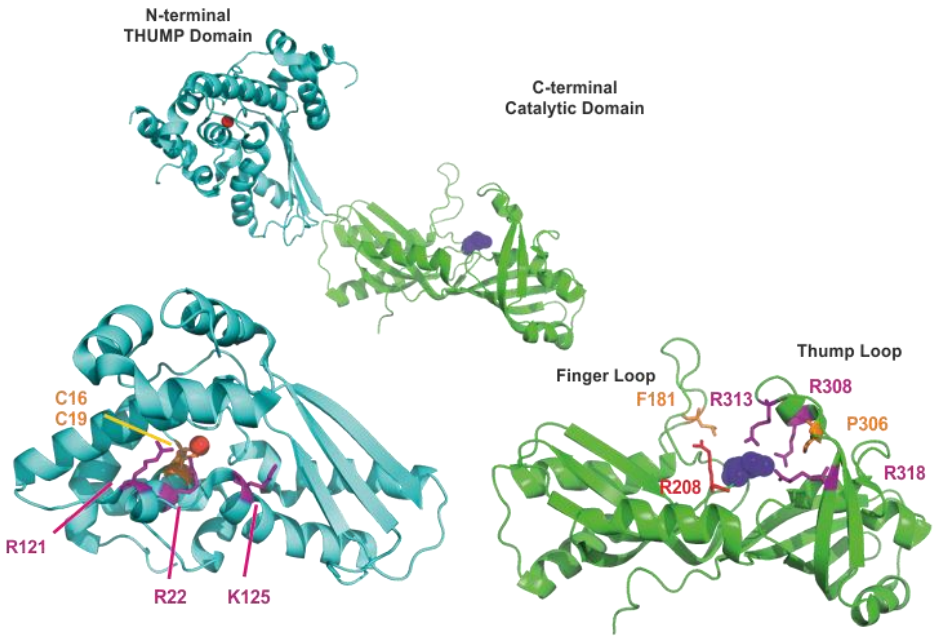


Figure 3.1. Structure of human Pus10 showing the amino acid residues selected for the current study.

Cartoon representation of human Pus10 (full-length Pus10 (top) (PDB ID: 2V9K), N-terminal domain (left, in cyan), and C-terminal domain (right, in green)) showing the amino acid residues (in stick form) targeted for site-directed mutagenesis. Residues are labeled with single letter codes and numbered according to *P. furiosus* Pus10. The zinc ion and the catalytic aspartate are shown as red and blue spheres, respectively. The corresponding numbering of amino acid residues in human Pus10 is as follows, in parentheses: C16 (C21), C19 (C24), R22 (R27), R121 (R176), K125 (K183), F181 (L314), R208 (R342), D210 (D344), P306 (P445), R308 (R447), R313 (R452), and R318 (R457). The figure was prepared using PyMOL (DeLano Scientific LLC, 2006).

To gain insight into the function of Pus10 in pseudouridylation of tRNA, site-directed mutagenesis was performed to create Pus10 variants with selected single amino acid substitutions in the thumb loop and the THUMP domain, and to generate a variant with two point mutations to eliminate zinc binding. In addition, a deletion variant containing only the Pus10 C-terminal catalytic domain was designed. Using the Pus10 wild-type and the Pus10 variants, the first quantitative biochemical characterization was performed to assess the catalytic efficiency of Pus10, and to explore the role of the N-terminal THUMP domain and the contribution of the thumb loop to the catalytic activity of Pus10. Our results demonstrate that the THUMP domain plays an important role in Pus10 interaction with tRNA, and that the thumb loop is critical for efficient catalysis by Pus10. These studies also show the essential catalytic role of arginine 208 and provide insight into the differences in efficiency of Pus10 compared to the Cbf5-Nop10-Gar1 complex in pseudouridylating tRNA.

3.2 Materials and Methods

3.2.1 Buffers and reagents

Buffers and reagents are described in Chapter 2 unless otherwise specified.

3.2.2 Molecular cloning and mutagenesis

The open reading frame of Pus10 was amplified from *P. furiosus* genomic DNA (ATCC, 43587D-5) using the following primers (restriction sites are in italics):

Pus10 sense (NheI) 5’-*GCTAGC*ATACTTGAAAAAGCCAGAGAGATATTGGAG-3’

Pus10 antisense (XhoI) 5’-*CTCGAG*TCAATTATCTCCCTCAACATCGTCC-3’

The amplified Pus10 gene was blunt-end ligated into a SmaI restricted pUC19 plasmid to create pUC19-PfPus10. Subsequently, a double-digestion of pUC19-PfPus10 was performed using XhoI and NheI and the Pus10 gene fragment was purified by gel extraction. The resulting Pus10 gene was inserted into pET28a that was also double-digested by the same set of restriction enzymes to create pET28a-PfPus10. This construct allowed the expression of Pus10 carrying an N-terminal hexahistidine tag.

An N-terminal deletion mutant (Pus10 ΔN) was created in a PCR reaction using pET28a-PfPus10 and the primers given below by amplifying the entire plasmid excluding the N-terminal domain of Pus10 (1-159 residues):

ΔN sense	5'-pCCAATATATGTAGCTGGGAGGTATAGAAAGCTC-3'
ΔN antisense	5'-GCTAGCCATATGGCTGCCGC-3'

Template DNA was removed by overnight DpnI digestion at 37 °C. Next, the linear product obtained in the PCR reaction was ligated by T4 DNA ligase to re-circularize the plasmid which was then transformed into high efficiency competent *E. coli* DH5α cells (New England Biolabs). Colonies were selected for resistance to kanamycin and screened by restriction analysis. Plasmids purified from the selected colonies were verified by sequencing (GENEWIZ).

Selected point mutations were introduced into the Pus10 gene by QuikChange^{TM} site-directed mutagenesis (Stratagene) employing the sense primers listed in Table 3.1. In addition to the point mutations, a double mutant was constructed, wherein two of the four conserved cysteine residues were mutated to alanine. Plasmids positive for the selected

point mutations were identified by restriction analysis utilizing either the restriction site that is removed or newly introduced by the primers and confirmed by sequencing.

Table 3.1. Primers used for site-directed mutagenesis of Pus10

Primer	Sequence (5' to 3')
C16/19A	GGAGCATCAAC***TGGCCA***ACCAC**GC**CTTGGGTAGGTTATTTGG
R22A	CCACTGCTTGGGT**GC**GTTATTTGGA***AAGCTT***GGGAAGGGAAC
R121A	GATTACAAAAGAGTTCAAC**GC**G***GAGCTC***GGGAAAGTTATTGCAG
K125A	GAGTTCAACAGG***GAGCTC***GGG**GC**AGTTATTGCAGTTAGATATGG
F181A	CATTAGAG***GAATTC***CCCCAAACTCCAGCCCCTGGG**GC**TAAGGAGAGC
R208A	CAAAGGAGCTGGG**GC**AGAAGACGT**G**GACGTTAGAATGC
D210N	GGAGCTGGGAGAGAA**A**ACGT**G**GACGTTAGAATGCTG
P306G	CGGAAATTAAACAGAGA***ACCGGT***AGGAGAGTACTCAATAGTAGAGC
R308A	GAACCCCCAGG**GC**AGTGCTCAATAGTAGAGC***AGATCT***AGTTAGAGT TAG
R313A	CCAGGAGAGT**GC**TCAATAGTGCAGC***AGATCT***AGTTAGAGTTAGAAA GG
R318A	AGGAGAGTGCTCAATAGTAGAGC***AGATCT***AGTT**GC**AGTTAGAAAGG TTTACG

Sequences of sense primers are given; antisense primers are exact complements of sense primers. The mutated nucleotides are highlighted in bold and newly introduced restriction sites in italics.

3.2.3 Protein expression and purification

Protein expression and purification of Pus10 wild-type and those carrying the desired mutations were performed essentially as described for Cbf5-Nop10 in Chapter 2. In brief, either wild-type or mutant plasmid was transformed into Rosetta 2(DE3) competent *E. coli* cells; cells were grown at 37 °C in LB medium containing 50 µg/mL kanamycin followed by induction with 1 mM IPTG at an OD_{600} of ~0.6. The cell pellet obtained three hours post-induction was lysed in buffer A1 using 1 mg/mL lysozyme and subsequent sonication. Following heat denaturation of the cell lysate at 75 °C for 15 min, protein purification was carried out using Ni^{2+} Sepharose affinity chromatography. Elutions were performed using buffer B with 20% glycerol. The eluted protein was further purified by size exclusion chromatography using a Superdex 75 column (GE Healthcare) in Buffer C. Protein quantification was performed by SDS-PAGE followed by densitometry using ImageJ (Schneider *et al.* 2012), analyzing various amounts of each purified protein on a 15% SDS-PAGE stained with Coomassie brilliant blue, against the standard curve of purified *P. furiosus* Cbf5. An average concentration was calculated from at least two independent measurements each with two to three different samples. The purity of the protein preparations was estimated to be between 80-90%.

3.2.4 *In vitro* transcription and purification of tRNA

In vitro transcription and purification of *P. furiosus* $tRNA^{Asp}$ was carried out as described in Chapter 2. In brief, $[^3H]tRNA^{Asp}$ was *in vitro* transcribed by T7 RNA polymerase at 30 °C for 4 h using template DNA prepared by PCR from pIDT-Smart-$PftRNA^{Asp}$. Following the removal of template DNA by DNaseI digestion, tRNA was

purified by anion-exchange chromatography using a Nucleobond AX100 column (Machery-Nagel). tRNA was concentrated by isopropanol precipitation and a subsequent precipitation with ethanol to remove salt. The concentration of tRNA dissolved in water was measured photometrically, and the specific activity was determined by scintillation counting.

3.2.5 Nitrocellulose Filtration

Nitrocellulose filtration experiments were conducted as summarized in section 2.2.5. In short, 5 nM [^{3}H]tRNA^{Asp} was incubated with 0 – 700 nM wild-type or 0 – 1200 nM Pus10 variants at 70 °C for 10 min. In the cases of the P306G, R313A, R318A, and ΔN variants, the incubation of tRNA with protein was extended to 30 min to allow the completion of product formation. The amount of tRNA retained on the filter was quantified by scintillation counting, and the K_D was determined by fitting the data to the quadratic equation (as in section 2.2.5) with [RNA] = 5 nM.

3.2.6 Tritium Release Assays

For initial characterization of proteins, tritium release assays were performed essentially as described in Chapter 2 using 5 nM Pus10 wild-type or variant with 600 nM [^{3}H]tRNA^{Asp} at 70 °C. For Michaelis-Menten titrations, either 5 nM Pus10 wild-type or R121A, 10 nM Pus10 ΔN, 15 nM Pus10 P306G, or 25 nM Pus10 R313A was titrated with increasing concentrations of [^{3}H]tRNA^{Asp} (50 nM – 1500 nM). Initial velocities (v_0) were determined from the time courses and v_0/[E] were calculated for each protein. The data obtained by plotting v_0/[E] against tRNA concentration was fit to the Michaelis-Menten equation to determine Michaelis-Menten parameters. Further, to obtain the

single-turnover rate constants, tritium release assays were performed with 5 μM of the selected protein and 600 nM [^{3}H]tRNA^{Asp}, and the data was fit to the single-exponential equation (as in section 2.2.6).

3.3 Results

3.3.1 Selection of amino acids for Pus10 mutational analysis

To select amino acid residues potentially involved in tRNA binding for site-directed mutagenesis, sequence alignments were performed using Pus10 sequences from ten archaeal species representing different classes of archaea and the human Pus10 sequence (Figure A1). *P. furiosus* Pus10 (388 amino acids) is much shorter in length than its human homolog (528 amino acids), with the majority of the amino acid residues (~90%) missing in *P. furiosus* Pus10 corresponding to the N-terminal domain. Therefore, all the aligned archaeal sequences show a shorter N-terminal domain compared to human Pus10. As described in section 3.1, the N-terminal domain shows a conserved zinc-binding motif and several conserved positively charged residues. To study the role of the zinc-binding motif, two of the four conserved cysteine residues (C16 and C19) were changed to alanine, which abolishes their ability to coordinate zinc. To understand the contribution of the conserved basic residues in binding the tRNA, three basic residues (R22, R121, K125) were individually substituted to alanine. Further, to examine the role of the N-terminal domain for the function of Pus10, a deletion variant missing the entire N-terminal domain was constructed (Pus10 ΔN). The forefinger loop in *P. furiosus* is also very short compared to the forefinger loop of human Pus10 with very low sequence conservation (Figure A1). However, the alignment shows a conserved aromatic residue in

the forefinger loop region in archaeal sequences, while the human Pus10 contains a lysine at this position. The phenylalanine present at this location in *P. furiosus* was replaced with alanine (F181A) to test its potential role in interacting with tRNA through base stacking.

Unlike the N-terminal domain and the forefinger loop, the thumb loop is very well conserved and contains several basic residues. In order to understand the role of these conserved basic residues in the interaction with tRNA, arginines at positions 308, 313, and 318 were substituted with alanine. Also, a conserved proline residue located at the start of the thumb loop was replaced with glycine (P306G) to test if it acts as a hinge in accommodating the tRNA in the active site. In addition to the above discussed changes made in the catalytic domain, the absolutely conserved catalytic aspartate (D210) was substituted with asparagine, and arginine 208, located two residues prior to the catalytic aspartate, was substituted with alanine (R208A). The arginine at this position is conserved in all the known ψ synthase families except TruA and TruB, and is proposed to play a role in flipping the target uridine into the active site in RluA (Hoang *et al.* 2006).

3.3.2 Purification of Pus10 and its variants

All the *P. furiosus pus10* mutants carrying the desired nucleotide changes were successfully constructed using QuikChange^{TM} site-directed mutagenesis and by employing PCR to obtain the N-terminal deletion mutant. The resulting constructs were used for over-expression of proteins in *E. coli,* and all proteins were purified by Ni^{2+} Sepharose affinity chromatography followed by size-exclusion chromatography. SDS-

PAGE analysis of the purified proteins showed an intense band corresponding to the full-length Pus10 (46.5 kDa) and a few less intense low molecular weight bands which were retained upon purification by size-exclusion chromatography (Figure 3.2), suggesting them to be degradation products of the full-length protein. Densitometry analysis carried out using ImageJ showed full-length Pus10 constituted more than 80% of the purified protein. Hence, the concentration of the protein corresponding to only the full-length protein was determined using SDS-PAGE and densitometry.

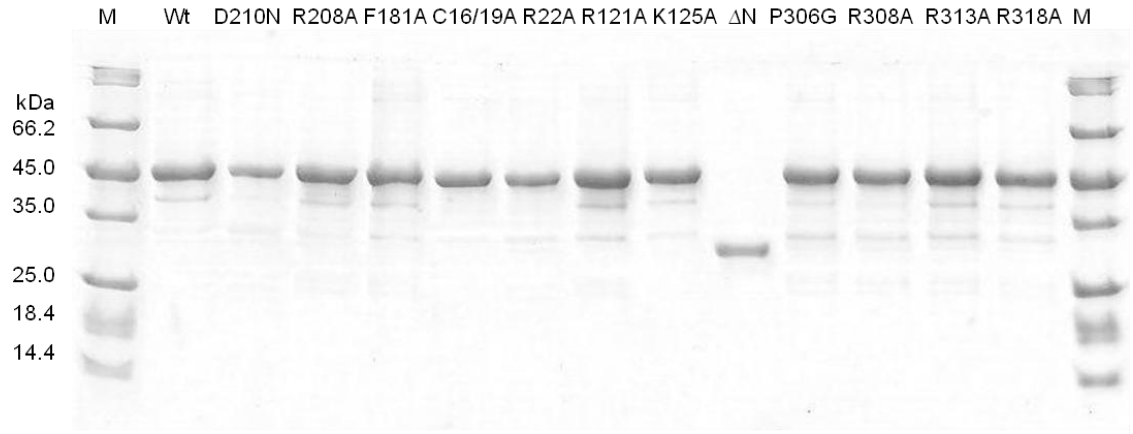


Figure 3.2. Gel analysis of purified Pus10 wild-type and variants.

50 pmol of each purified Pus10 protein was analyzed on 12% tris-tricine-PAGE and visualized by staining with Coomassie blue. M: Protein molecular weight marker. Each lane is identified with the corresponding protein sample, Pus10 wild-type (Wt) and its variants. Molecular weights of the proteins from the marker are labeled.

3.3.3 Optimal KCl concentration for ψ55 formation by Pus10

Previously, optimal buffer conditions required for tRNA modification by Cbf5 were determined (refer to section 2.3.1). These experiments have demonstrated that a KCl concentration of 150 mM is optimal for Cbf5's *in vitro* ψ synthase activity in introducing

ψ55 into tRNA. Surprisingly, Gurha and Gupta (2008) reported that *P. furiosus* Pus10 forms ψs at non-specific sites along with position 54 and 55, when they conducted the experiments at 150 mM NaCl. To test if Pus10 modifies more than one site in our system, tritium release assays were performed at four different KCl concentrations (75 mM – 600 mM) using 5 nM Pus10 wild-type and 600 nM [^{3}H]tRNA (Figure 3.3). At the lowest KCl concentration used in the experiment (*i.e.* 75 mM), about 60% ψ forms in 50 min; and at the highest KCl concentration (*i.e.* 600 mM) tested, about 40% ψ forms in 50 min. In reactions performed at 150 mM KCl and 300 mM KCl, 90% and 70% ψ was detected, respectively.

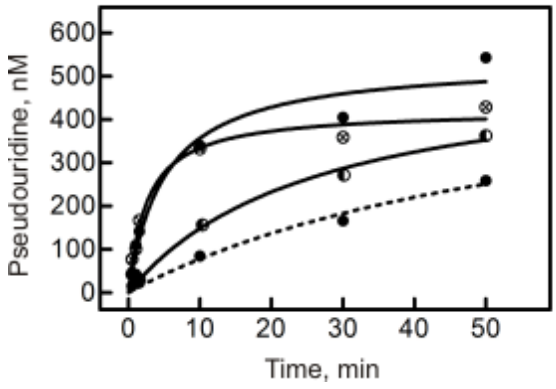


Figure 3.3. KCl titration of pseudouridine formation in tRNA by Pus10.

Tritium release assays were performed with 5 nM Pus10 wild-type and 600 nM [^{3}H]-tRNA, in the presence of 75 mM KCl (half-filled circles), 150 mM KCl (filled circles), 300 mM KCl (crossed circles), and 600 mM KCl (filled circles, dotted lines). Smooth curves were obtained by fitting the data to a hyperbolic equation.

Thus the highest amount of ψ was observed when 150 mM KCl was used in the reaction, and under any tested KCl concentration the percentage of modification is not more than 100%. Moreover, as observed from the initial linear phase, ψ is formed at a much a higher rate in 150 mM and 300 mM KCl, compared to the reaction at 75 mM and

600 mM KCl. Together these data indicate that the optimal KCl concentration for tRNA modification by Pus10 is between 150 mM – 300 mM, and in this range there is only one ψ modified per tRNA. Based on this observation, all further experiments were performed with the buffer containing 150 mM KCl.

3.3.4 Pseudouridylation activity of Pus10 wild-type and its variants

In order to analyze the pseudouridylation ability of Pus10 wild-type, a tritium release assay was performed under multiple-turnover conditions using 5 nM protein and 600 nM [^{3}H]tRNA. To monitor the course of ψ formation, samples were collected at various time points after initiating the reaction, and the amount of ψ was determined in each sample. Pus10 wild-type shows 100% conversion of uridine to ψ in 30 min, while 90% of the modification takes place within 15 min under these conditions (Figure 3.4A). From the linear phase, the initial rate catalyzed by wild-type Pus10 was determined to be 160 ± 20 nM min^{-1} (Table 3.2). This analysis was also performed with all the constructed variants in order to study the effects of amino acid substitutions on Pus10's ability to pseudouridylate tRNA. Pus10 F181A, containing the only substitution made in the forefinger loop, yields a time course that is very similar to the wild-type in terms of both end level and initial rate (140 ± 5 nM min^{-1}) (Figure 3.4A, Table 3.2). As shown in Chapter 2 and in other published works (Ramamurthy *et al.* 1999a; Wright *et al.* 2011), when the absolutely conserved aspartate residue from the catalytic pocket of ψ synthase is substituted with asparagine, it abolished the catalytic activity of the enzyme in modifying tRNA. As predicted, substitution of the aspartate residue at position 210 of Pus10 with asparagine also abolishes Pus10's ability to modify tRNA (Figure 3.4A,

Table 3.2). Interestingly, also the arginine to alanine replacement at position 208 renders Pus10 inactive, similarly to the catalytically inactive Pus10 D210N variant.

Next, the Pus10 variants with single amino acid changes in the thumb loop were tested. Pus10 P306G and Pus10 R318A variants show comparable time courses, reaching ~ 50% ψ formation in 30 min with similar initial rates (Figure 3.4B, Table 3.2). Pus10 R313A results in slower ψ formation, *i.e.* 25% modification in 30 min, with an initial rate of 7 ± 1 nM min^{-1}. Another thumb loop variant constructed in this study, Pus10 R308A shows almost no ψ formation in 60 min (Figure 3.4B, Table 3.2). In summary, all the thumb loop substitutions have exhibited a significant effect on their pseudouridylation activity with the R308A substitution having the most severe effect.

Further analysis of the variants with substitutions in the THUMP domain was carried out. All these variants, excluding the N-terminal deletion variant, result in very similar time courses as that of Pus10 wild-type (Figure 3.4C). Hence, contrary to the thumb loop variants, there is no effect on the pseudouridylation activity of Pus10 due to the substitution of the selected basic residues in the THUMP domain, as well as the removal of the zinc-binding site (Table 3.2). However, the Pus10 variant lacking the N-terminal domain shows a significant decrease in the amount of ψ formed (only 40% modification in 60 min) and a about 36-fold decrease in the initial rate (Table 3.2). Altogether, these results demonstrate the importance of the residues in the thumb loop besides the catalytic pocket and the role of the THUMP domain as a whole for the catalytic function of Pus10 in pseudouridylation.

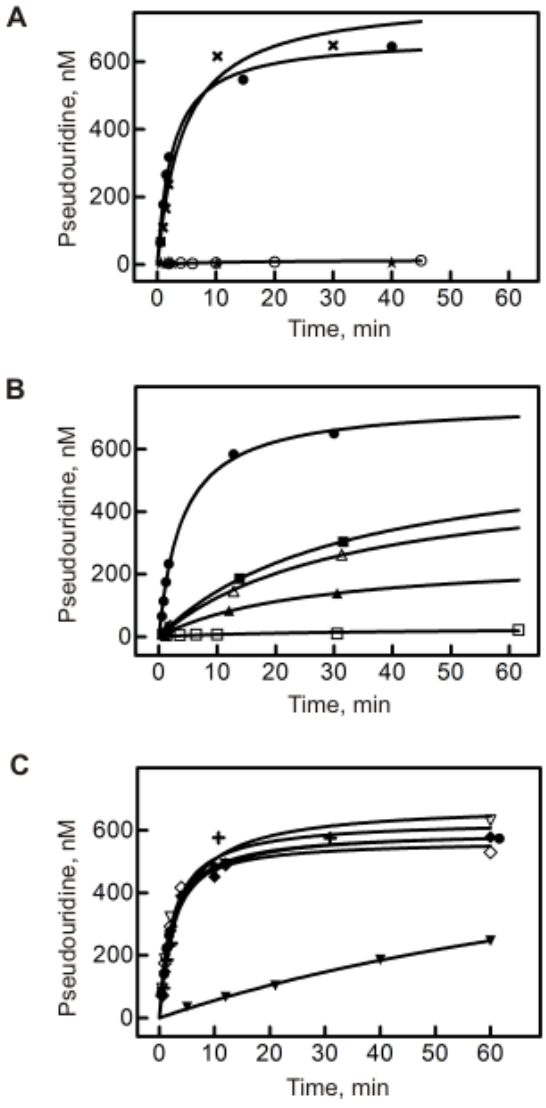


Figure 3.4. Time courses of ψ55 formation in tRNA by Pus10 wild-type and variants.

Tritium release assays were conducted with 5 nM protein and 600 nM [^{3}H]tRNA at 70°C. The time courses were fit to a hyperbolic equation (smooth curves) and initial rates were calculated from the linear fits (not shown) of the initial linear phase. A) Time courses of Pus10 wild-type (filled circles), Pus10 F181A (cross), Pus10 D210N (open circles), and Pus10 R208A (star). B) Time courses of Pus10 wild-type (filled circles) compared to the variants carrying substitutions in the thumb loop; Pus10 P306G (filled squares), Pus10 R308A (open squares), Pus10 R313A (filled triangles), and Pus10 R318A (open triangles). C) Time courses of Pus10 wild-type (filled circles) compared to the variants with substitutions in the THUMP domain: Pus10 C16AC19A (plus), Pus10 R22A (open diamonds), Pus10 R121A (filled diamonds), Pus10 K125A (open inverted triangles), and Pus10 ΔN (filled inverted triangles).

Table 3.2. Initial rates (v_0) for pseudouridine formation by Pus10 wild-type and variants

Region of substitution	**Pus10 variant**	**v_0, nM min^{-1}**
	Pus10 Wild-type	160 ± 20
Catalytic pocket	Pus10 D210N	nd[a]
	Pus10 R208A	nd[a]
Forefinger loop	Pus10 F181A	140 ± 5
Thumb loop	Pus10 P306G	18 ± 4
	Pus10 R308A	0.40 ± 0.05
	Pus10 R313A	7 ± 1
	Pus10 R318A	20 ± 2
Zinc-binding site	Pus10 C16AC19A	150 ± 10
N-terminal THUMP domain	Pus10 R22A	140 ± 20
	Pus10 R121A	135 ± 5
	Pus10 K125A	150 ± 15
	Pus10 ΔN	4.3 ± 0.1

[a] nd – not determined since Pus10 variant is inactive

3.3.5 ψ formation by Pus10 under single-turnover conditions

As noticed in the tritium release assays conducted under multiple-turnover conditions, Pus10 R208A shows no activity, indicating that the point mutation introduced at this site severely affected the catalytic ability of Pus10. Under the conditions used in multiple-turnover experiments, it is difficult to know if the Pus10 R208A variant is limited to only a single round of catalysis given that a single round corresponds to 5 nM tRNA modification (*i.e.* less than 1%). Therefore to answer this question, tritium release assays were carried out using 5 μM protein and 600 nM [^{3}H]tRNA, wherein the protein only has to perform one round of modification corresponding to a maximum of 600 nM tRNA pseudouridylation. Pus10 wild-type converts about 90% of the uridine into ψ in 4 min, reaching about 80% ψ already in one minute, at a specific rate of 0.06 s^{-1} (Figure 3.5). As expected, the catalytically inactive variant, Pus10 D210N showed no ψ formation, and the same was true for the Pus10 R208A variant. The inability of the Pus10 R208A variant to form ψ even in the presence of excess protein confirms that the Pus10 R208A variant is catalytically inactive. In addition to the variants carrying changes in the catalytic pocket, Pus10 R308A, the other variant that also showed undetectable ψ formation in multiple-turnover experiments, was tested under single-turnover conditions. This variant with a mutation in the thumb loop, yielded a moderate level of ψ accumulation by converting 25% of uridine to ψ upon 4 min incubation (Figure 3.5). In summary, Pus10 R308A differed from Pus10 D210N and R208A variants by showing some ψ formation while the two catalytic pocket variants exhibited no pseudouridylation activity. This

indicates that the pseudouridylation ability of Pus10 R308A is not completely abolished, although it is severely impaired.

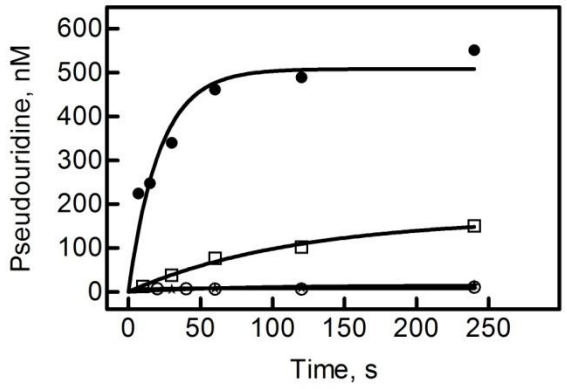


Figure 3.5. Time courses of pseudouridine formation by Pus10 under single-turnover conditions.

Tritium release assays were performed using 5 μM protein and 600 nM tRNA. Pus10 wild-type (filled circles), Pus10 D210N (open circles), Pus10 R208A (stars), and Pus10 R308A (open squares). Time courses were fit to a single exponential equation to obtain the rate of pseudouridine formation (k_{ψ}).

3.3.6 ψ formation by Pus10 R308A under multiple-turnover conditions

From the previous experiments it is not clear if Pus10 R308A is capable of performing multiple rounds of catalysis, which requires the ability to release the product tRNA upon modification. In order to determine this, tritium release assays were conducted with 200 nM protein and 600 nM tRNA. Although Pus10 R308A could not introduce ψ into all the tRNA molecules, it modified about 65% of the tRNA in 60 minutes *i.e.* it performed close to 2 rounds of catalysis (Figure 3.6). This result demonstrates that Pus10 R308A is not limited in product release as it can perform more than a single round of catalysis, but the slower rate of ψ formation under both multiple and single-turnover conditions

confirms that the R308A substitution in the thumb loop severely affects Pus10's catalytic ability.

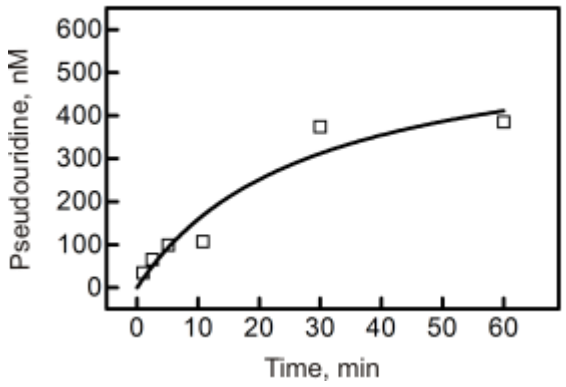


Figure 3.6. Pseudouridylation of tRNA by Pus10 R308A under multiple-turnover conditions.

Tritium release assays were performed with 200 nM Pus10 R308A and 600 nM tRNA. The smooth curve was obtained by fitting the time course to a hyperbolic equation.

3.3.7 Investigating the effect of amino acid changes on tRNA binding by Pus10

One of the questions we would like to answer in this work is what regions of Pus10 mediate the interaction with tRNA. To address this question, analysis of all the constructed variants in comparison to the Pus10 wild-type was conducted by nitrocellulose filtration experiments that allowed us to determine the dissociation constants (K_Ds) for binding of Pus10 to tRNA. [^{3}H]tRNA was incubated with increasing concentrations of either Pus10 wild-type or variant at 70 °C for 10 min. The minimum protein concentration employed in the experiment was at least three-fold higher than the concentration of the tRNA. Under these conditions, most of the tRNA molecules are pseudouridylated, therefore the K_D measured here reflects the K_D for the product tRNA. In the cases of the thumb loop and ΔN variants that showed slower rates of ψ formation,

the incubation step was extended to 30 min to allow the completion of the tRNA modification. The Pus10 variants, R208A and D210N were not subjected to this prolonged incubation as they are inactive. When tRNA was titrated with Pus10 wild-type, about 45% of tRNA was retained on the filter at saturating protein concentrations (Figure 3.7). For most Pus10 variants, the plateau was observed at about 30% of tRNA binding. These low retention efficiencies of RNA in nitrocellulose filtration experiments are not uncommon, and might result from variations in RNA refolding, from the washing step performed to remove unspecifically bound RNA also removing some RNA from the active site, and variations in different RNA preparations (Hall and Kranz 1999; Ramamurthy *et al.* 1999a). However, it is important to note that the end levels remained constant in the repetitions, and the K_Ds were reproducible.

Fitting the data to the quadratic equation function, yielded a K_D of 30 nM for Pus10 wild-type binding to tRNA (Table 3.3). To allow us to determine the K_D for the Pus10 interaction with the substrate tRNA, the catalytically inactive variant, Pus10 D210N was tested and resulted in an identical K_D (29 nM) to that of the product tRNA. Next, the analysis was performed with the constructs carrying the substitutions in the catalytic pocket, forefinger loop, thumb loop, and the THUMP domain. As summarized in Table 3.3, both Pus10 R208A carrying a substitution in the catalytic pocket and Pus10 F181A with a substitution in the forefinger loop exhibit very similar K_Ds to the wild-type. All the variants with changes in the thumb loop show slightly higher K_Ds (two- to five-fold). Interestingly, the variants carrying changes in the N-terminal domain result in K_Ds that are five to seven-fold higher than the K_D obtained for Pus10 wild-type, with the exception of the Pus10 C16AC19A variant, which shows just over a three-fold increase.

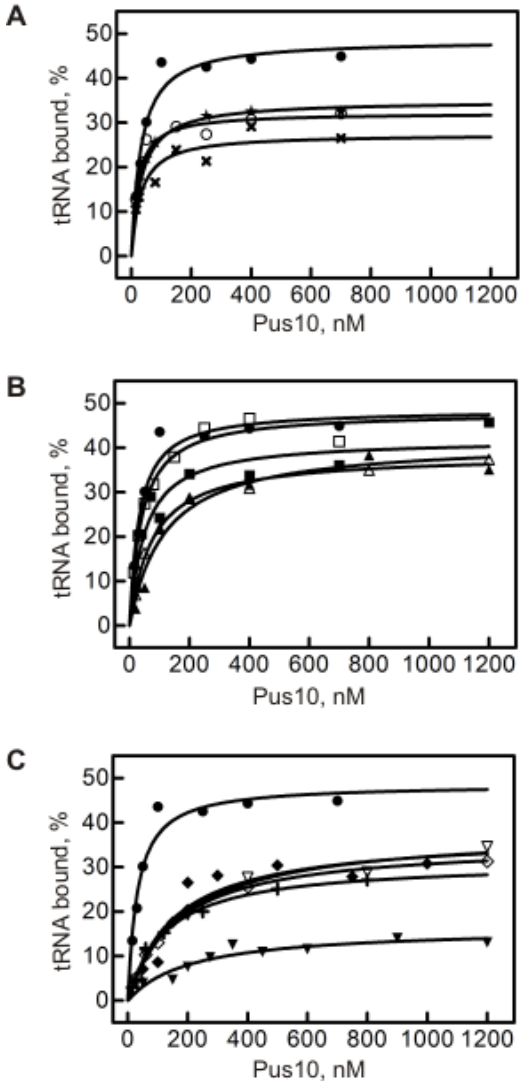


Figure 3.7. tRNA titration with Pus10 wild-type and Pus10 variants to determine the affinities between Pus10 and tRNA.

Nitrocellulose filtration experiments were conducted with 5 nM [^{3}H]tRNA and increasing concentrations of protein. Reaction mixtures were incubated for at least 10 min at 70 °C prior to the filtration. Smooth curves were obtained by fitting the data to a quadratic equation. A) Binding curves of Pus10 wild-type (filled circles), Pus10 F181A (cross), Pus10 D210N (open circles), and Pus10 R208A (star). B) Binding curves of Pus10 wild-type (filled circles) compared to the variants carrying selected substitution in the thumb loop; Pus10 P306G (filled squares), Pus10 R308A (open squares), Pus10 R313A (filled triangles), and Pus10 R318A (open triangles). C) Binding curves of Pus10 wild-type (filled circles) compared to the variants with selected substitution in the THUMP domain: Pus10 C16AC19A (plus), Pus10 R22A (open diamonds), Pus10 R121A (filled diamonds), Pus10 K125A (open inverted triangles), and Pus10 ΔN (filled inverted triangles).

Table 3.3. Equilibrium dissociation constants for the interaction between Pus10 and tRNA[a]

Region of substitution	Pus10 variant	K_D, nM
	Pus10 Wild-type	30 ± 10
Catalytic pocket	Pus10 D210N	29 ± 15
	Pus10 R208A	50 ± 25
Forefinger loop	Pus10 F181A	33 ± 20
Thumb loop	Pus10 P306G	70 ± 30
	Pus10 R308A	50 ± 25
	Pus10 R313A	140 ± 55
	Pus10 R318A	85 ± 25
Zinc-binding site	Pus10 C16AC19A	110 ± 30
N-terminal THUMP domain	Pus10 R22A	170 ± 45
	Pus10 R121A	160 ± 60
	Pus10 K125A	200 ± 50
	Pus10 ΔN	190 ± 65

[a] K_Ds provided here are averages from at least two independent nitrocellulose filtration experiments reported with the highest standard deviation observed.

Notably, deletion of the N-terminal domain had a similar effect on the K_D as the three single-residue substitutions (R22A, R121A, K125A) studied in the N-terminal domain (Table 3.3). In summary, the nitrocellulose filtration experiments revealed a tight interaction between Pus10 and tRNA with a low nanomolar K_D. Pus10 binds both substrate tRNA and product tRNA with the same affinity. While the selected substitutions in the thumb loop have relatively moderate effects on Pus10's interaction with the tRNA, changes in the N-terminal domain had a significantly larger influence.

3.3.8 Steady-state kinetic analysis of Pus10 wild-type and selected Pus10 variants

To gain insight into the catalytic efficiency of Pus10 wild-type and the interaction of Pus10 with tRNA, steady-state kinetic analysis was performed. Further, to probe the effect of the single residue changes introduced into Pus10, selected variants carrying substitutions in the thumb loop (P306G and R313A) and the THUMP domain (R121A) along with the N-terminal deletion variant were also analyzed by steady-state kinetic experiments. Low nanomolar concentrations of Pus10 wild-type or the Pus10 variants were incubated with increasing concentrations of tRNA in tritium release assays and initial rates at each tRNA concentration were determined (Figure 3.8A). To provide an expression that is independent of the enzyme concentration used in the experiment, initial rates over the enzyme concentration (v_0/[E]) was calculated. Data from v_0/[E] plotted against the respective tRNA concentration were fit to the Michaelis-Menten equation to determine the k_{cat} and K_M values (Figure 3.8B and C). All tested proteins have shown a gradual increase in v_0/[E] with increasing tRNA concentration, which saturates at tRNA concentrations higher than 1 μM, with the exception of Pus10 ΔN. This analysis yielded a k_{cat} of 0.9 s^{-1} and a K_M of 400 nM for Pus10 wild-type.

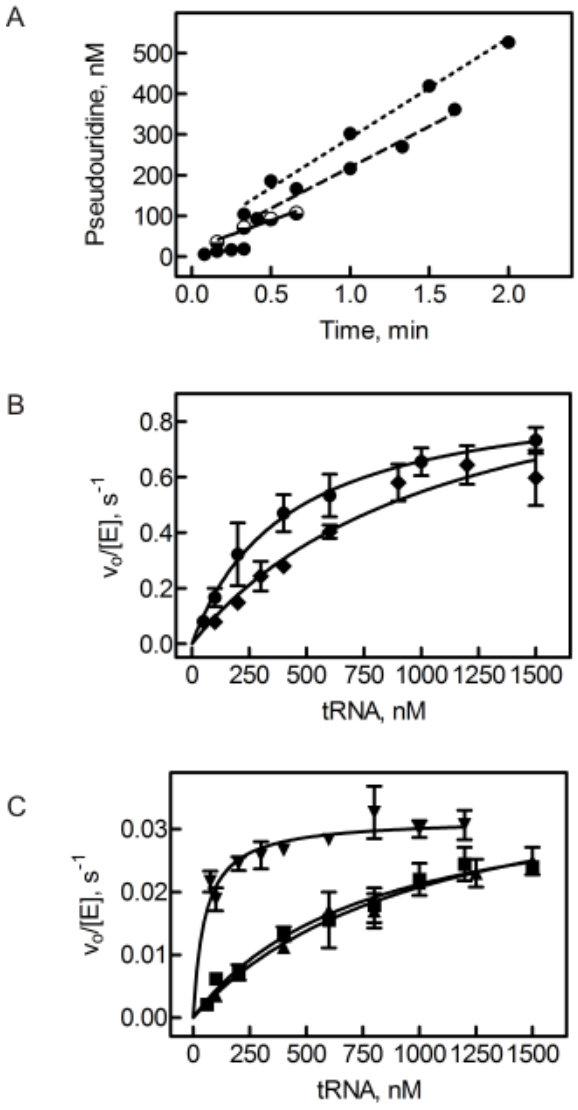


Figure 3.8. Steady-state kinetics of pseudouridylation by Pus10 wild-type and selected Pus10 variants.

Tritium release assays were performed with low nanomolar concentrations of protein (refer to the methods for details) and increasing concentrations of [^{3}H]tRNA at 70 °C. A) Representative short time courses of Pus10 wild-type that were used to calculate the initial velocities. 5 nM Pus10 wild-type was incubated with various concentrations of tRNA, 100 nM (filled circles), 400 nM (half filled circles), 1000 nM (filled circles, dashed lines), and 1500 nM (filled circles, dotted lines). Data were fit to a linear equation. B) Michaelis-Menten titration of Pus10 wild-type (filled circles) and Pus10 R121A (filled diamonds). The v_0/[E] values were calculated at each tested tRNA concentration by dividing the initial velocities by the concentration of enzyme used in the experiment. C) Michaelis-Menten titration of Pus10 P306G (filled squares), Pus10 R313A (filled triangles), and Pus10 ΔN (filled inverted triangles).

The Pus10 R121A variant with the substitution in the THUMP domain exhibits a very similar k_{cat} as the wild-type, but about a three-fold higher K_M (Table 3.4).

In contrast, both analyzed thumb loop variants (Pus10 P306G and Pus10 R313A) showed more than a twenty-fold reduction in the k_{cat}, but only a two-fold increase in the K_M (Figure 3.8C; Table 3.4). Interestingly, incubation of the N-terminal deletion variant with increasing concentrations of tRNA displays only a slight increase in v_0/[E] at low tRNA concentrations, with no further enhancement in v_0/[E] at higher tRNA concentrations. Owing to this behaviour observed for Pus10 ΔN, the K_M could only be estimated for this variant. Overall, the N-terminal deletion variant results in a k_{cat} that is similar to Pus10 P306G and Pus10 R313A, but thirty-fold smaller than the wild-type k_{cat} (Table 3.4), and a K_M that is at least six-fold smaller than the wild-type K_M.

Table 3.4. Summary of steady-state kinetic parameters of Pus10 and its variants

Region of substitution	**Pus10 variant**	**K_M, nM**	**k_{cat}, s^{-1}**
	Pus10 Wild-type	400 ± 40	0.90 ± 0.03
Thumb loop	Pus10 P306G	800 ± 170	0.040 ± 0.004
	Pus10 R313A	1000 ± 190	0.040 ± 0.004
N-terminal domain	Pus10 R121A	1100 ± 350	1.2 ± 0.2
	Pus10 ΔN	< 65	0.030 ± 0.001

3.4 Discussion

Most of our current knowledge on Pus10 is limited to the structural information available through the crystal structure of human Pus10 determined in the absence of the tRNA (McCleverty *et al.* 2007) and to the target site of archaeal Pus10 *i.e.* ψ55 in tRNA both *in vitro* and *in vivo* (Roovers *et al.* 2006; Gurha and Gupta 2008; Blaby *et al.* 2011). To understand the structure-function relationship of Pus10, with an emphasis on Pus10 interaction with tRNA, a kinetic and thermodynamic analysis of several Pus10 variants carrying amino acid residue changes in key structural regions of Pus10 was conducted. Results from these studies demonstrate that Pus10 efficiently modifies tRNA using its N-terminal THUMP domain and conserved thumb loop for binding and catalysis. This study presents the first quantitative analysis of tRNA modification by Pus10.

In light of the observation made by Gurha and Gupta (2008), which reported Pus10's ability to target both uridines at positions 54 and 55 in *Haloferax volcanii* tRNA with additional non-specific uridine conversion when buffer containing 150 mM NaCl was used, here the ψ formation by *P. furiosus* Pus10 at various KCl concentrations was tested. The results from these experiments show that 150 mM to 300 mM KCl is optimal for this reaction, and there is only one site modified under the reaction conditions used. The contrasting results obtained by Gurha and Gupta (2008) could be attributed to the use of a heterologous system containing *P. furiosus* Pus10 and *H. volcanii* $tRNA^{Trp}$ in their studies. Here both Pus10 and tRNA substrate from *P. furiosus* were used, which may explain the higher specificity observed in these experiments. The observation that *P. furiosus* tRNA contains no ψ at position 54, but another modification of uridine, S^2m^5U (Kowalak *et al.* 1994), and the detection of ψs at position 55 (but not at 54), when *in vitro*

transcribed yeast and *H. volcanii* tRNAs were incubated with *P. furiosus* cell-free extract (Constantinesco *et al.* 1999), both suggests it is likely that U55 is modified.

Time courses conducted with a large excess of substrate over enzyme revealed for the first time that Pus10 is a multiple-turnover enzyme. As predicted, substitution of the aspartate residue in the catalytic pocket with asparagine confirmed that D210 is the catalytic residue since Pus10 becomes inactive upon its substitution. Nitrocellulose filtration experiments demonstrated that despite the lack of pseudouridylation ability, Pus10 D210N can tightly bind tRNA with a low nanomolar affinity. Pus10 binds both the substrate and the modified product tRNA with similar affinity, as seen with TruB and Cbf5-Nop10-Gar1 (Wright *et al.* 2011; Kamalampeta and Kothe 2012). The catalytic efficiencies (k_{cat}/K_M) calculated from the values obtained by steady-state kinetic analysis offer insight into the modification of tRNA at position 55 by Pus10 compared to Cbf5-Nop10-Gar1. Pus10 displays a catalytic efficiency of 2.2×10^6 M^{-1} s^{-1} that is about thirteen-fold higher than the catalytic efficiencies calculated for *P. furiosus* Cbf5-Nop10-Gar1 (1.75×10^5 M^{-1} s^{-1}) using the data obtained in our previous work (Kamalampeta and Kothe 2012). As both Pus10 and Cbf5-Nop10-Gar1 exhibit similar k_{cat} values (0.9 s^{-1} and 0.7 s^{-1}, respectively), the difference in their efficiency stems from the ten-fold difference in their K_M values. These data provide the reason for Pus10 being the ψ55 synthase *in vivo* but not the Cbf5-Nop10-Gar1 complex, which is usually involved in rRNA modification in a guide-RNA dependent manner. It is important to note that the k_{cat} value for Pus10 is similar to those reported for other stand-alone ψ synthases from bacteria, which indicates that Pus10 might also display a slow catalytic step as described for TruB, TruA, and RluA (Wright *et al.* 2011), suggesting a similar catalytic mechanism employed

by Pus10 in ψ formation. Surprisingly, the single-turnover rate of ψ formation (k_{ψ}) measured for Pus10 is about fifteen-fold lower than the k_{cat}. k_{ψ} is the rate of ψ formation in a single round whereas k_{cat} is the rate of a multiple-turnover reaction, which consists of all the steps as in single round and in addition a product release step. So, fundamentally k_{ψ} cannot be lower than the k_{cat}. The reason for such a low rate in presence of high enzyme concentrations is not known. Further investigation is needed to identify the problem in these experiments.

A sequence alignment of bacterial ψ synthases along with Pus10 (McCleverty *et al.* 2007) revealed an arginine residue, two amino acids N-terminal to the catalytic aspartate, that is conserved in all ψ synthases except in TruB and TruD. An arginine to alanine substitution at this position (Pus10 R208A) has no effect on Pus10's affinity for tRNA (Table 3.3), but abolishes the catalytic activity, demonstrating an essential role of R208 for the catalytic function of Pus10. It is likely that this arginine residue is involved in flipping the target uridine into the active site during the modification, similarly to the role suggested for the arginine found in RluA in a structurally equivalent position (Hoang *et al.* 2006). In this work, the authors showed that neither methionine, nor lysine could be substituted for the arginine without abolishing pseudouridylation activity, indicating the importance of arginine at this position. Although the bacterial ψ55 synthase, TruB, does not contain arginine at this position, histidine 43 was proposed to play a similar role in promoting the base flipping during the modification of U55 (Hoang and Ferré-D'Amaré 2001). In summary, besides the other three catalytic residues (D210, Y274, K348 in Pus10) that are conserved in ψ synthases, R208 is also essential for the Pus10 catalyzed

tRNA modification. Further studies are required to prove the role of R208 in placing the target uridine in the active site.

The N-terminal THUMP domain in Pus10 is only found in this ψ synthase family similar to other accessory, family-specific domains found in other ψ synthases. Since the accessory domains in the other families of ψ synthases are involved in RNA binding and since the THUMP domain is present in several RNA modifying enzymes, it has been suggested that Pus10's THUMP domain is involved in facilitating RNA binding (Aravind and Koonin 2001; McCleverty *et al.* 2007). Besides this hypothesis, a model created by manually docking tRNA onto Pus10 suggested that the conserved basic residues in the THUMP domain are likely involved in accommodating tRNA (McCleverty *et al.* 2007). Three Pus10 constructs created in this work by individually substituting conserved basic residues in the THUMP domain all show a significantly reduced affinity for tRNA compared to wild-type Pus10 (Table 3.3), confirming the role of the N-terminal domain in interacting with tRNA. None of these Pus10 variants have any effect on the pseudouridylation activity of Pus10 due to the higher substrate concentrations (about three-fold over their K_D) used in these experiments (Figure 3.4C). In addition to these results, the steady-state parameters determined for Pus10 R121A revealed a similar k_{cat} to wild-type, but a two-fold higher Michaelis-Menten constant (Table 3.4), which further supported the effect of these substitutions on tRNA binding but not on the catalytic activity. Interestingly, the N-terminal deletion variant also displayed a similar effect on affinity for tRNA as observed with single amino acid substitutions of the basic residues from this domain. This implies that the conserved basic residues in the THUMP domain are the main contacts for the interaction with tRNA. Furthermore, the deletion of the

entire THUMP domain severely affected Pus10's catalytic activity, yielding a thirty-fold lower k_{cat} compared to wild-type Pus10 (Table 3.4). Such a significant consequence of the N-terminal deletion on the catalytic ability of Pus10 further suggests an important role for the THUMP domain in augmenting the functional ability of Pus10 in pseudouridylating tRNA. This role of the THUMP domain may be accomplished through proper positioning of tRNA into the catalytic pocket. The presence of a much longer THUMP sequence found in higher eukaryotes, together with some remarkably identical regions identified in *Drosophila melanogaster* and humans (McCleverty *et al.* 2007) may point towards a role of the THUMP domain in other functions of Pus10 observed in humans such as a function in TRAIL induced apoptosis (Park *et al.* 2009).

Analysis of the double variant (Pus10 C16AC19A), created to eliminate zinc binding by Pus10, revealed only a small effect (three-fold reduction) on the affinity of Pus10 for tRNA, with no effect on the catalytic activity of Pus10, indicating that the zinc-binding motif is not required for the catalytic function of Pus10. This finding is in contrast to the studies on a yeast ψ synthase, Pus1, where the zinc ion is essential for Pus1 binding to tRNA *in vitro* and hence for the catalytic function of the enzyme (Arluison *et al.* 1998). Nevertheless, it is possible that the zinc-binding motif of Pus10 may have a role in the cell that is not essential for the pseudouridylation activity in the purified *in vitro* system.

Lastly, the four thumb loop variants investigated in this work, each with a single amino acid substitution, all display minor effects on the affinity for tRNA. But, they all show significantly reduced initial rates in pseudouridylation with the most severe effect resulting from the R308A substitution. Time courses performed with Pus10 R308A in the presence of high concentrations of enzyme revealed that its catalytic activity is severely

impaired. The multiple-turnover assay indicated that this variant is not limited in product release. Together, these results confirmed that R308 is very important for the catalytic activity of Pus10. The steady-state kinetic data obtained for Pus10 P306G and Pus10 R313A showed an over twenty-fold reduction in k_{cat}, further supporting the effect on the catalytic activity due to substitutions in the thumb loop. A slight increase in K_M values (about two-fold) observed for these two variants also correspond well with the similar effects seen on the K_Ds. In summary, the results obtained with the thumb loop variants confirm that the conserved basic residues in this loop play a critical role in the catalytic function of Pus10. It is not clear from these experiments whether the positive charge of the arginine or the specific arginine side chain is important for its role. Future experiments with Pus10 variants carrying substitutions of the arginines with methionine and lysine could clarify this question. As proposed for TruB (Hoang *et al.* 2005), the thumb loop may contribute to the stabilization of the interaction of tRNA with Pus10.

In contrast to the important role of the thumb loop, the single phenylalanine (F181) residue we studied from the forefinger loop did not show any contribution towards either tRNA binding or catalysis. This finding can be attributed to the poor conservation observed in this structural element, which may indicate a less significant role of this loop in Pus10 function.

In conclusion, the first quantitative studies reported here on tRNA modification by Pus10 provide evidence for the important role of the THUMP domain and the thumb loop in this reaction. This work also establishes the critical contribution of arginine 208 in Pus10 catalytic activity. Our studies also provide a reason why Pus10 is the archaeal *in vivo* ψ55 synthase, despite the existence of Cbf5 in the cell. Further, we propose that

Pus10 follows an induced-fit mechanism in binding the tRNA, similar to TruB (Pan *et al.* 2003). In this work, by comparing the structures of TruB alone and with that of a TruB bound to RNA, the authors suggested that TruB uses an induced fit mechanism wherein it undergoes a conformational change upon initial binding to enhance the interaction with tRNA. In a similar manner, in the case of Pus10, initial docking of tRNA may take place with the help of the THUMP domain followed by the interaction of the thumb loop which undergoes a conformational change. Then both the THUMP domain and the thumb loop facilitate the correct positioning of tRNA in the active site followed by a conformational change in the active site, wherein arginine 208 probably promotes the flipping of the target nucleotide for catalysis.

Chapter 4: Conclusions and future perspectives

Even though pseudouridine was identified over 50 years ago, its actual function in the cell and the molecular mechanism underlying the modification are yet to be known. Undoubtedly, significant progress has been made in the last one and a half decade which is mainly dominated by the knowledge obtained from the structural studies and biochemical analysis of ψ synthases (Hamma and Ferré-D'Amaré 2006; Grosjean and Ferré-D'Amaré 2009). Besides gaining more insight into the structure and function of ψ synthases, these studies also raised new questions. One such case is the ψ synthase Cbf5 that evolved to function together with H/ACA RNA to create the relatively large number of modifications found in archaea and eukaryotes. Interestingly, being a homolog to TruB, a stand-alone ψ55 synthase, Cbf5 also retained its ability to introduce the ψ55 in archaeal tRNAs (Roovers *et al.* 2006). In this thesis, studies on the contribution of two accessory proteins, Nop10 and Gar1, revealed that they exert a two-fold activating effect on guide RNA-independent activity of Cbf5 by increasing the affinity of Cbf5 for tRNA and by enhancing its catalytic ability. The catalytic rates determined for the guide-RNA-independent modification by Cbf5 indicated that by forming a ternary complex with Nop10 and Gar1 proteins, Cbf5 attains a similar efficiency as other stand-alone ψ synthases. Some of these findings are likely applicable to the guide-RNA-dependent activity of H/ACA sRNPs (see section 2.4 for detailed discussion). Although results from this work establish the basis for the ability of Cbf5 to modify tRNA under *in vitro* conditions, it is not clear if the cell has any advantage for retaining this ability of Cbf5 as Cbf5 appears to be mainly involved in guide-RNA-dependent pseudouridine formation in the cell (Blaby *et al.* 2011). In general, observations in this work indicate how an

enzyme's dependency on partner proteins for attaining the catalytic efficiency can act as a means of regulation. It can be envisioned that inherently unstable Cbf5 in the cell becomes fully active only when its partner proteins are present, which in turn may be involved in the coordination of other related activities in the cell.

I also made attempts to purify the functional H/ACA sRNP complex from *P. furiosus*, but only a partially active complex has been obtained that could not be used for kinetic studies of the guide RNA-dependent function of Cbf5. Very recently, Ye and colleagues published the kinetic analysis of a short synthetic substrate RNA modification by *P. furiosus* H/ACA sRNP complex (Yang *et al.* 2012). However, the experiments were performed at 37 ˚C while the optimal temperature for Cbf5-catalyzed tRNA modification under *in vitro* conditions is 70 ˚C (Roovers *et al.* 2006), which is similar to the physiological temperature of *P. furiosus* (~100 ˚C). In future, it will be useful to carry out the steady-state kinetic analysis at this reported optimal temperature for Cbf5 to obtain a relevant k_{cat} which would allow a comparison of the catalytic efficiencies of Cbf5 in guide RNA-independent and -dependent functions. Further, to confirm the involvement of the residues proposed in mediating the roles of Nop10 and Gar1 (Figure 2.7), individual Cbf5, Nop10, and Gar1 variants carrying substitutions in these amino acids can be constructed and analyzed in both guide-RNA-dependent and -independent reactions.

Initially it was not known if Cbf5 or Pus10 is responsible for the ψ55 formation in archaeal tRNAs. During the course of my thesis, however, it has been shown that Pus10 performs the *in vivo* modification of ψ55 in tRNAs in archaea (Blaby *et al.* 2011). In this thesis, thermodynamic and kinetic studies of Pus10 variants carrying selected amino acid

substitutions in various structural regions of the protein allowed me to identify several key residues and their contributions to tRNA binding and catalysis. Most interestingly, this work demonstrated the important role of conserved amino acids in the thumb loop of Pus10 for catalysis and potential contributions of the unique THUMP domain to tRNA binding. Another key finding from this work is the identification of a new catalytic arginine residue located close to the catalytic aspartate. Pus10 is one of the two known ψ55 synthases to contain a zinc ion. In contrast to the essential role of zinc in the case of the yeast pseudouridine synthase Pus1p (TruA family) (Arluison *et al.* 1998), studies from this thesis indicate no obvious need for zinc in *in vitro* modification reaction by Pus10. It is possible that zinc plays a structural role in Pus10.

Consistent with Pus10's function in tRNA modification *in vivo*, Michaelis-Menten parameters determined in this work indicated a higher catalytic efficiency of Pus10 compared to the Cbf5-Nop10-Gar1 complex. Further, the finding that Pus10 displays a similar k_{cat} as other known ψ synthases also strengthens the argument that pseudouridine synthases employ the same catalytic mechanism. In summary, besides the insight into the unexpected function of Cbf5 in tRNA modification, this work provides the first quantitative biochemical analysis of Pus10 and opens the way for further findings that will enhance our understanding of this new ψ synthase family. This work also extends our understanding of enzymes involved in tRNA modification, in general.

In future, it will be interesting to confirm the suggested role of the newly identified catalytic arginine residue in base-flipping. One way of confirming this is to use tRNA with a fluorescent label in the TψC arm to perform a titration with wild-type Pus10 and

with the Pus10 variant carrying a substitution for arginine 208. A change in fluorescence reflecting base-flipping in the case of wild-type Pus10 that is absent in the arginine 208 variant would provide evidence for its role in base-flipping. Further, to support the contribution of the THUMP domain in tRNA binding, a cocrystal structure of Pus10 and tRNA can be determined which will, in addition, provide insight into the interaction of Pus10 and tRNA. As noticed there is a discrepancy in the literature regarding the site of modification of tRNA by Pus10 which can modify either position 55 or uridine 54 as reported in a halophile (Gurha and Gupta 2008). To confirm if *P. furiosus* Pus10 only introduces ψ55 under the tested conditions, a control tRNA containing a substitution at position 55 (Roovers *et al.* 2006) can be used in tritium release assays to clarify this confusion. Besides these proposed studies, solving the kinetic mechanism of Pus10-catalyzed tRNA modification will further extend our knowledge of this novel ψ synthase.

Towards solving the chemical mechanism of ψ formation, there is a need for a comprehensive understanding of the substrate recognition by ψ synthases. To this end, solving cocrystal structures for remaining representative ψ synthases may prove to be useful. Also, biochemical studies involving mutational analysis of substrates to identify the key determinants of protein-RNA interactions and the minimal substrate requirements will be informative. Experimental techniques to monitor the conformational changes in real time such as stopped-flow may be helpful in gaining further understanding into the enzyme and substrate interactions. Together, a careful integration of knowledge obtained using structural studies combined with biochemical analysis may eventually reveal the chemical mechanism of this deceptively simple modification. Finally, to understand the actual role of ψs there is a need to distinguish the other biological roles of ψ synthases

from that of the modification, which requires considerable further efforts to design both *in vivo* and *in vitro* studies. Given the extent of progress achieved in last one and a half decade one can be optimistic about solving these puzzles in the very near future.

References

Alexander, R. P., Fang, G., Rozowsky, J., Snyder, M. and Gerstein, M. B. (2010). Annotating non-coding regions of the genome. *Nat Rev Genet* 11: 559-571.

Alian, A., DeGiovanni, A., Griner, S. L., Finer-Moore, J. S. and Stroud, R. M. (2009). Crystal structure of an RluF-RNA complex: a base-pair rearrangement is the key to selectivity of RluF for U2604 of the ribosome. *J Mol Biol* 388: 785-800.

Anderson, B. R., Muramatsu, H., Nallagatla, S. R., Bevilacqua, P. C., Sansing, L. H., Weissman, D. and Kariko, K. (2010). Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. *Nucleic Acids Res* 38: 5884-5892.

Aravind, L. and Koonin, E. V. (2001). THUMP- a predicted RNA-binding domain shared by 4-thiouridine, pseudouridine synthases and RNA methylases. *Trends Biochem Sci* 26: 215-217.

Arluison, V., Hountondji, C., Robert, B. and Grosjean, H. (1998). Transfer RNA-pseudouridine synthetase Pus1 of *Saccharomyces cerevisiae* contains one atom of zinc essential for its native conformation and tRNA recognition. *Biochemistry* 37: 7268-7276.

Arnez, J. G. and Steitz, T. A. (1994). Crystal structure of unmodified tRNA(Gln) complexed with glutaminyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry* 33: 7560-7567.

Auffinger, P. and Westhof, E. (1998). Effects of pseudouridylation on tRNA hydration and dynamics: A theoretical approach. *Modification and editing of RNA*. H. Grosjean and R. Benne(Ed). Washington, DC, ASM Press: 103-112.

Bagni, C. and Lapeyre, B. (1998). Gar1p binds to the small nucleolar RNAs snR10 and snR30 *in vitro* through a nontypical RNA binding element. *J Biol Chem* 273: 10868-10873.

Baker, D. L., Youssef, O. A., Chastkofsky, M. I., Dy, D. A., Terns, R. M. and Terns, M. P. (2005). RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. *Genes Dev* 19: 1238-1248.

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

Ballarino, M., Morlando, M., Pagano, F., Fatica, A. and Bozzoni, I. (2005). The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol* 25: 5396-5403.

Becker, H. F., Motorin, Y., Planta, R. J. and Grosjean, H. (1997). The yeast gene YNL292w encodes a pseudouridine synthase (Pus4) catalyzing the formation of pseudouridine 55 in both mitochondrial and cytoplasmic tRNAs. *Nucleic Acids Res* 25: 4493-4499.

Montanaro, L., Treré, D. and Derenzini, M. (2008). Nucleolus, ribosomes, and cancer. *Am J Pathol* 173: 301-310.

Blaby, I. K., Majumder, M., Chatterjee, K., Jana, S., Grosjean, H., de Crecy-Lagard, V. and Gupta, R. (2011). Pseudouridine formation in archaeal RNAs: The case of *Haloferax volcanii*. *RNA* 17: 1367-1380.

Cantara, W. A., Crain, P. F., Rozenski, J., McCloskey, J. A., Harris, K. A., Zhang, X., Vendeix, F. A., Fabris, D. and Agris, P. F. (2011). The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res* 39(Database issue): D195-201.

Charette, M. and Gray, W. (2000). Pseudouridines in RNA: What, where, how, and why. *IUBMB Life* 49: 341 - 351.
Charpentier, B., Muller, S. and Branlant, C. (2005). Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res* 33: 3133-3144.
Chen, C., Zhao, X., Kierzek, R. and Yu, Y. T. (2010). A flexible RNA backbone within the polypyrimidine tract is required for U2AF65 binding and pre-mRNA splicing *in vivo*. *Mol Cell Biol* 30: 4108-4119.
Cohn, W. E. (1959). 5-Ribosyl uracil, a carbon-carbon ribofuranosyl nucleoside in ribonucleic acids. *Biochim Biophys Acta* 32: 569-571.
Cohn, W. E. (1960). Pseudouridine, a carbon-carbon linked ribonucleoside in ribonucleic acids: isolation, structure, and chemical characteristics. *J Biol Chem* 235: 1488-1498.
Collins, K. (2006). The biogenesis and regulation of telomerase holoenzymes. *Nat Rev Mol Cell Biol* 7: 484-494.
Constantinesco, F., Motorin, Y. and Grosjean, H. (1999). Transfer RNA modification enzymes from *Pyrococcus furiosus*: detection of the enzymatic activities *in vitro*. *Nucleic Acids Res* 27: 1308-1315.
Cortese, R., Kammen, H. O., Spengler, S. J. and Ames, B. N. (1974). Biosynthesis of pseudouridine in transfer ribonucleic acid. *J Biol Chem* 249: 1103-1108.
Costa, F. F. (2008). Non-coding RNAs, epigentics and complexity. *Gene* 410: 9-17.
Darzacq, X., Jady, B. E., Verheggen, C., Kiss, A. M., Bertrand, E. and Kiss, T. (2002). Cajal body-specific small nuclear RNAs: a novel class of 2′-O-methylation and pseudouridylation guide RNAs. *EMBO J* 21: 2746-2756.
Darzacq, X., Kittur, N., Roy, S., Shav-Tal, Y., Singer, R. H. and Meier, U. T. (2006). Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J Cell Biol* 173: 207-218.
Davis, D. R. (1995). Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res* 23: 5020 - 5026.
Davis, F. F. and Allen, F. W. (1957). Ribonucleic acids from yeast which contain a fifth nucleotide. *J Biol Chem* 227: 907-915.
Decatur, W. A. and Fournier, M. J. (2002). rRNA modifications and ribosome function. *Trends Biochem Sci* 27: 344-351.
del Campo, M., Ofengand, J. and Malhotra, A. (2004). Crystal structure of the catalytic domain of RluD, the only rRNA pseudouridine synthase required for normal growth of *Escherichia coli*. *RNA* 10: 231-239.
Dez, C., Henras, A., Faucon, B., Lafontaine, D., Caizergues-Ferrer, M. and Henry, Y. (2001). Stable expression in yeast of the mature form of human telomerase RNA depends on its association with the box H/ACA small nucleolar RNP proteins Cbf5p, Nhp2p and Nop10p. *Nucleic Acids Res* 29: 598-603.
Dez, C., Houseley, J. and Tollervey, D. (2006). Surveillance of nuclear-restricted pre-ribosomes within a subnucleolar region of *Saccharomyces cerevisiae*. *EMBO J* 25: 1534-1546.
Douthwaite, S., Formy, D. and Yoshizawa, S. (2005). Nucleotide methylations in rRNA that confer resistance to ribosome-targeting antibiotics. *Fine-tuning of RNA functions by modification and editing*. H. Grosjean(Ed). Heidelberg, Springer-Verlag.

Duan, J., Li, L., Lu, J., Wang, W. and Ye, K. (2009). Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. *Mol Cell* 34: 427-439.
Fersht, A. (1998). *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding*, W. H. Freeman and Company.
Filipowicz, W. and Pogacic, V. (2002). Biogenesis of small nucleolar ribonucleoproteins. *Curr Opin Cell Biol* 14: 319-327.
Foster, P. G., Huang, L., Santi, D. V. and Stroud, R. M. (2000). The structural basis for tRNA recognition and pseudouridine formation by pseudouridine synthase I. *Nat Struct Mol Biol* 7: 23-27.
Fromont-Racine, M., Senger, B., Saveanu, C. and Fasiolo, F. (2003). Ribosome assembly in eukaryotes. *Gene* 313: 17-42.
Gannot, P., Bortolin, M. L. and Kiss, T. (1997). Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* 89: 799-809
Ge, J. and Yu, Y.-T. (2013). RNA pseudouridylation: new insights into an old modification. *Trends Biochem Sci* (published online ahead of print February 4, 2013): dx.doi.org/10.1016/j.tibs.2013.1001.1002.
Gill, S. C. and von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182: 319-326.
Goldwasser, E. and Heiniukson, R. L. (1966). The Biochemistry of Pseudouridine. *Progress in Nucleic Acid Res and Mol Biol*. J. N. Davidson and W. E. Cohn(Ed), Elsevier. 5: 399-416.
Grosjean, H. (2005). Modification and editing of RNA: historical overview and important facts to remember. *Fine-tuning of RNA functions by modification and editing*. H. Grosjean(Ed). Heidelberg, Springer-Verlag: 1-22.
Grosjean, H. and Ferré-D'Amaré, A. R. (2009). Pseudouridine formation, the most Common transglycosylation in RNA. *DNA and RNA modification enzymes: Structure, mechanism, function and evolution*. H. Grosjean(Ed). Austin, TX, Landes Bioscience: 363-376.
Grosjean, H., Gaspin, C., Marck, C., Decatur, W. A. and de Crecy-Lagard, V. (2008). RNomics and Modomics in the halophilic archaea *Haloferax volcanii*: identification of RNA modification genes. *BMC Genomics* 9: 470.
Grosshans, H., Lecointe, F., Grosjean, H., Hurt, E. and Simos, G. (2001). Pus1p-dependent tRNA pseudouridinylation becomes essential when tRNA biogenesis is compromised in yeast. *J Biol Chem* 276: 46333-46339.
Grozdanov, P. N., Roy, S., Kittur, N. and Meier, U. T. (2009). SHQ1 is required prior to NAF1 for assembly of H/ACA small nucleolar and telomerase RNPs. *RNA* 15: 1188-1197.
Gu, X., Liu, Y. and Santi, D. V. (1999). The mechanism of pseudouridine synthase I as deduced from its interaction with 5-fluorouracil-tRNA. *Proc Natl Acad Sci* 96: 14270-14275.
Gurha, P. and Gupta, R. (2008). Archaeal Pus10 proteins can produce both pseudouridine 54 and 55 in tRNA. *RNA* 14: 2521-2527.
Gurha, P., Joardar, A., Chaurasia, P. and Gupta, R. (2007). Differential roles of archaeal Box H/ACA proteins in guideRNA-dependent and independent pseudouridine formation. *RNA Biol* 4: 101-109.

Hadjiolov, A. A. (1985). The nucleolus and ribosome biogenesis. *Cell Biology Monographs*. A. M. Beerman, L. Goldstein, K. R. Portrer and P. Sitte(Ed). New York, Springer-Verlag: 1-263.
Hall, K. B. and Kranz, J. K. (1999). Nitrocellulose filter binding for determination of dissociation constants. *RNA-protein interaction protocols*. S. R. Haynes(Ed). Totowa, NJ, Humana Press. 118: 105-114.
Hamma, T. and Ferré-D'Amaré, A. R. (2006). Pseudouridine synthases. *Chem Biol* 13: 1125-1135.
Hamma, T. and Ferré-D'Amaré, A. R. (2010). The box H/ACA ribonucleoprotein complex: interplay of RNA and protein structures in post-transcriptional RNA modification. *J Biol Chem* 285: 805-809.
Hamma, T., Reichow, S. L., Varani, G. and Ferré-D'Amaré, A. R. (2005). The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat Struct Mol Biol* 12: 1101-1107.
Harrington, K. M., Nazarenko, I. A., Dix, D. B., Thompson, R. C. and Uhlenbeck, O. C. (1993). In vitro analysis of translational rate and accuracy with an unmodified tRNA. *Biochemistry* 32: 7617-7622.
Henras, A. K., Soudet, J., Gerus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A. and Henry, Y. (2008). The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci* 65: 2334-2359.
Hermann, T. and Patel, D. J. (1999). Stitching together RNA tertiary structure. *J Mol Biol* 294: 829-849.
Hoang, C., Chen, J., Vizthum, C. A., Kandel, J. M., Hamilton, C. S., Mueller, E. G. and Ferré-D'Amaré, A. R. (2006). Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure. *Mol Cell* 24: 535-545.
Hoang, C. and Ferré-D'Amaré, A. R. (2001). Cocrystal structure of a tRNA psi55 pseudouridine synthase: Nucleotide flipping by an RNA-modifying enzyme. *Cell* 107: 929-939.
Hoang, C. and Ferré-D'Amaré, A. R. (2004). Crystal structure of the highly divergent pseudouridine synthase TruD reveals a circular permutation of a conserved fold. *RNA* 10: 1026-1033.
Hoang, C., Hamilton, C. S., Mueller, E. G. and Ferré-D'Amaré, A. R. (2005). Precursor complex structure of pseudouridine synthase TruB suggests coupling of active site perturbations to an RNA-sequestering peripheral protein domain. *Protein Sci* 14: 2201-2206.
Huang, L., Pookanjanatavip, M., Gu, X. and Santi, D. V. (1998). A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst. *Biochemistry* 37: 344-351.
Hur, S. and Stroud, R. M. (2007). How U38, 39, and 40 of many tRNAs become the targets for pseudouridylation by TruA. *Mol Cell* 26: 189-203.
Hur, S., Stroud, R. M. and Finer-Moore, J. (2006). Substrate recognition by RNA 5-methyluridine methyltransferases and pseudouridine synthases: a structural perspective. *J Biol Chem* 281: 38969-38973.
Ishida, K., Kunibayashi, T., Tomikawa, C., Ochi, A., Kanai, T., Hirata, A., Iwashita, C. and Hori, H. (2011). Pseudouridine at position 55 in tRNA controls the contents of other

modified nucleotides for low-temperature adaptation in the extreme-thermophilic eubacterium *Thermus thermophilus*. *Nucleic Acids Res* 39: 2304-2318.
Ishitani, R., Yokoyama, S. and Nureki, O. (2008). Structure, dynamics, and function of RNA modification enzymes. *Curr Opin Struct Biol* 18: 330-339.
Johnson, L. and Söll, D. (1970). In vitro biosynthesis of pseudouridine at the polynucleotide level by an enzyme extract from *Escherichia coli*. *Proc Natl Acad Sci* 67: 943-950.
Kamalampeta, R. and Kothe, U. (2012). Archaeal proteins Nop10 and Gar1 increase the catalytic activity of Cbf5 in pseudouridylating tRNA. *Sci Rep* 2: 663.
Karjolich, J. and Yu, Y. T. (2010). Spliceosomal snRNA modifications and their function. *RNA Biol* 7: 192-204.
Karjolich, J. and Yu, Y. T. (2011). Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* 474: 395-398.
Kawai, G., Yamamoto, Y., Kamimura, T., Masegi, T., Sekine, M., Hata, T., Limori, T., Watanabe, T., Miyazawa, T. and Yokoyama, S. (1992). Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group. *Biochemistry* 31: 1040-1046.
Kenmochi, N., Kawaguchi, T., Rozen, S., Davis, E., Goodman, N., Hudson, T. J., Tanaka, T. and Page, P. G. (1998). A map of 75 human ribosomal protein genes. *Genomic Res* 8: 509-523.
King, T. H., Liu, B., R., M. R. and Fournier, M. J. (2003). Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol Cell* 11: 425 - 436.
Kipper, K., Sild, S., Hetenyi, C., Remme, J. and Liiv, A. (2011). Pseudouridylation of 23S rRNA helix 69 promotes peptide release by release factor RF2 but not by release factor RF1. *Biochimie* 93: 834-844.
Kirwan, M. and Dokal, I. (2009). Dyskeratosis congenita, stem cells and telomeres. *Biochim Biophys Acta* 1792: 371-379.
Kiss, A., Sain, B. and Venetianer, P. (1977). The number of rRNA genes in *Escherichia coli*. *FEBS Lett* 79: 77-79.
Kiss, T., Fayet-Lebaron, E. and Jady, B. E. (2010). Box H/ACA small ribonucleoproteins. *Mol Cell* 37: 597-606.
Klein, D. J., Schmeing, T. M., Moore, P. B. and Steitz, T. A. (2001). The kink-turn: a new RNA secondary structure motif. *EMBO J* 20: 4214-4221.
Klein, R. J., Misulovin, Z. and Eddy, S. R. (2002). Noncoding RNA genes identified in AT-rich hyperthermophiles. *Proc Natl Acad Sci* 99: 7542-7547.
Koonin, E. V. (1996). Pseudouridine synthases: Four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. *Nucleic Acids Res* 24: 2411-2415.
Kowalak, J. A., Dalluge, J. J., McCloskey, J. A. and Stetter, K. O. (1994). The role of posttranscriptional modification in stabilization of transfer RNA from hyperthemophiles. *Biochemistry* 33: 7869-7876.
Lafontaine, D. L., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. and Tollervey, D. (1998). The box H/ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev* 12: 527-537.

Lane, B. G., Ofengand, J. and Gray, M. W. (1995). Pseudouridine and O2'-methylated nucleosides. Significance of their selective occurrence in rRNA domains that function in ribosome-catalyzed synthesis of the peptide bonds in proteins. *Biochimie* 77: 7-15.
Lapeyre, B. (2005). Conserved ribosomal RNA modification and their putative roles in ribosome biogenesis and translation. *Fine-tuning of RNA functions by modification and editing*. H. Grosjean(Ed). Heidelberg, Springer-Verlag: 263-284.
Lecointe, F., Namy, O., Hatin, I., Simos, G., Rousset, J. P. and Grosjean, H. (2002). Lack of pseudouridine 38/39 in the anticodon arm of yeast cytoplasmic tRNA decreases *in vivo* recoding efficiency. *J Biol Chem* 277: 30445-30453.
Lecompte, O., Ripp, R., Thierry, J. C., Moras, D. and Poch, O. (2002). Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. *Nucleic Acids Res* 30: 5382-5390.
Lee, J. C. and Gutell, R. R. (2012). A comparison of the crystal structures of eukaryotic and bacterial SSU ribosomal RNAs reveals common structural features in the hypervariable regions. *PLoS One* 7: e38203.
Li, L. and Ye, K. (2006). Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* 443(7109): 302-307.
Li, S., Duan, J., Li, D., Yang, B., Dong, M. and Ye, K. (2011). Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase. *Genes Dev* 25(22): 2409-2421.
Li, X., Zhong, S. and Wong, W. H. (2005). Reliable prediction of transcription factor binding sites by phylogenetic verification. *Proc Natl Acad Sci* 102: 16945-16950.
Liang, B., Xue, S., Terns, R. M., Terns, M. P. and Li, H. (2007). Substrate RNA positioning in the archaeal H/ACA ribonucleoprotein complex. *Nature Struct Mol Biol* 14: 1189-1195.
Liang, B., Zhou, J., Kahen, E., Terns, R. M., Terns, M. P. and Li, H. (2009). Structure of a functional ribonucleoprotein pseudouridine synthase bound to a substrate RNA. *Nat Struct Mol Biol* 16: 740-746.
Liang, X. H., Liu, Q. and Fournier, M. J. (2007). rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Mol Cell* 28: 965-977.
Liang, X. H., Liu, Q. and Fournier, M. J. (2009). Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA* 15: 1716-1728.
Martin, R., Straub, A. U., Doebele, C. and Bohnsack , M. T. (2013). DExD/H-box RNA helicases in ribosome biogenesis. *RNA Biol* 10: 3-17.
Marvel, C. C., Arps, P. J., B.C., R., Kammen, H. O., Penhoet, E. E. and Winkler, M. E. (1985). hisT is part of a multigene operon in *Escherichia coli* K-12. *J Bacteriol* 161: 60-71.
Mattick, J. S. (2001). Non-coding RNAs: the architects of eukaryotic complexity. *EMBO Rep* 2(11): 986-991.
Mattick, J. S. and Makunin, I. V. (2006). Non-coding RNA. *Hum Mol Genet* 15: 17-29.
McCleverty, C. J., Hornsby, M., Spraggon, G. and Kreusch, A. (2007). Crystal structure of human Pus10, a novel pseudouridine synthase. *J Mol Biol* 373: 1243-1254.
McDonald, M. K., Miracco, E. J., Chen, J., Xie, Y. and Mueller, E. G. (2011). The handling of the mechanistic probe 5-fluorouridine by the pseudouridine synthase TruA

and its consistency with the handling of the same probe by the pseudouridine synthases TruB and RluA. *Biochemistry* 50: 426-436.
Merz, K., Hondele, M., Goetze, H., Gmelch, K., Stoeckl, U. and Griesenbeck, J. (2008). Actively transcribed rRNA genes in *S. cerevisiae* are organized in a specialized chromatin associated with the high-mobility group protein Hmo1 and are largely devoid of histone molecules. *Genes Dev* 22: 1190-1204.
Miracco, E. J. and Mueller, E. G. (2011). The products of 5-fluorouridine by the action of the pseudouridine synthase TruB disfavor one mechanism and suggest another. *J Am Chem Soc* 133: 11826-11829.
Motorin, Y. and Helm, M. (2010). tRNA stabilization by modified nucleotides. *Biochemistry* 49: 4934-4944.
Mueller, E. G. (2002). Chips off the old block. *Nature Struct Mol Biol* 9: 320-322.
Muller, S., Fourmann, J. B., Loegler, C., Charpentier, B. and Branlant, C. (2007). Identification of determinants in the protein partners aCBF5 and aNOP10 necessary for the tRNA:Psi55-synthase and RNA-guided RNA:Psi-synthase activities. *Nucleic Acids Res* 35: 5610-5624.
Muller, S., Urban, A., Hecker, A., Leclerc, F., Branlant, C. and Motorin, Y. (2009). Deficiency of the tRNATyr:Psi 35-synthase aPus7 in Archaea of the Sulfolobales order might be rescued by the H/ACA sRNA-guided machinery. *Nucleic Acids Res* 37: 1308-1322.
Nazar, R. N. (2004). Ribosomal RNA processing and ribosome biogenesis in eukaryotes. *IUBMB Life* 56: 457-465.
Neumann, J. M., Bernassau, J. M., Gueron, M. and Tran-Dinh, S. (1980). Comparative conformations of uridine and pseudouridine and their derivatives. *Eur J Biochem* 108: 457-463.
Ni, J., Tien, A. L. and Fournier, M. J. (1997). Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* 89: 565-573.
Nissen, P., Hansen, J., Ban, N., Moore, P. B. and Steitz, T. A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science* 289: 920-930.
Noon, K. R., Bruenger, E. and McCloskey, J. A. (1998). Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hyperthermophile *Sulfolobus solfataricus. J Bacteriol* 180: 2883-2888.
Nurse, K., Wrzesinski, J., Bakin, A., Lane, B. G. and Ofengand, J. (1995). Purification, cloning, and properties of the tRNA psi 55 synthase from *Escherichia coli*. *RNA* 1: 101-112.
Ofengand, J. (2002). Ribosomal RNA pseudouridines and pseudouridine synathses. *FEBS Lett* 514: 17-25.
Omer, A. D., Ziesche, S., Decatur, W. A., Fournier, M. J. and Dennis, P. P. (2003). RNA-modifying machines in Archaea. *Mol Microbiol* 48: 617-629.
Pan, H., Agarwalla, S., Moustakas, D. T., Finer-Moore, J. and Stroud, R. M. (2003). Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through a combination of rigid docking and induced fit. *Proc Natl Acad Sci U S A* 100: 12648-12653.
Park, S. Y., Shin, J. N., Woo, H. N., Piya, S., Moon, A. R., Seo, Y. W., Seol, D. W. and Kim, T. H. (2009). DOBI is cleaved by caspases during TRAIL-induced apoptotic cell death. *BMB Rep* 42: 511-515.

Pemberton, L. F. and Paschal, B. M. (2005). Mechanisms of receptor-mediated nuclear import and nuclear export. *Traffic* 6: 187-198.
Phannachet, K., Youssef, E. and H, H. (2005). Dissecting the roles of a strictly conserved tyrosine in substrate recognition and catalysis by pseudouridine 55 synthase. *Biochemistry* 44: 15488-15494.
Preumont, A., Snoussi, K., Stroobant, V., Collet, J. F. and Van Schaftingen, E. (2008). Molecular identification of pseudouridine-metabolizing enzymes. *J Biol Chem* 283: 25238-25246.
Ramamurthy, V., Swann, S. L., Paulson, J. L., Spedaliere, C. J. and Mueller, E. G. (1999). Critical aspartate residues in Pseudouridine synthases. *J Biol Chem* 274: 22225-22230.
Ramamurthy, V., Swann, S. L., Spedaliere, C. J. and Mueller, E. G. (1999). Role of cysteine residues in pseudouridine synthases of different families. *Biochemistry* 38: 13106-13111.
Rashid, R., Liang, B., Baker, D. L., Youssef, O. A., He, Y., Phipps, K., Terns, R. M., Terns, M. P. and Li, H. (2006). Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol Cell* 21: 249-260.
Raska, I., Koberna, K., Malinsky, J., Fidlerova, H. and Masata, M. (2004). The nucleolus and transcription of ribosomal genes. *Biol Cell* 96: 579-594.
Reddy, R. and Busch, H. (1988). Small nuclear RNAs: RNA sequences, structure, and modifications. *Structure and function of major and minor small nuclear ribonucleoprotein particles*. M. L. Birnsteil(Ed). Heidelberg, Springer-Verlag 1-37.
Reichow, S. L., Hamma, T., Ferré-D'Amaré, A. R. and Varani, G. (2007). The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res* 35: 1452-1464.
Romby, P., Carbon, P., Westhof, E., Ehresmann, C., Ebel, J. P., Ehresmann, B. and Giege, R. (1987). Importance of conserved residues for the conformation of the T-loop in tRNAs. *J Biomol Struct Dyn* 5: 669-687.
Roovers, M., Hale, C., Tricot, C., Terns, M. P., Terns, R. M., Grosjean, H. and Droogmans, L. (2006). Formation of the conserved pseudouridine at position 55 in archaeal tRNA. *Nucleic Acids Res* 34: 4293-4301.
Rozhdestvensky, T. S., Tang, T. H., Tchirkova, I. V., Brosius, J., Bachellerie, J. P. and Huttenhofer, A. (2003). Binding of L7Ae protein to the K-turn of archaeal snoRNAs: a shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. *Nucleic Acids Res* 31: 869-877.
Samuelsson, T., Guindy, Y. S., Lustig, F., Boren, T. and Lagerkvist, U. (1987). Apparent lack of discrimination in the reading of certain codons in *Mycoplasma mycoides*. *Proc Natl Acad Sci* 84: 3166-3170.
Satoh, A., Takai, K., Ouchi, R., Yokoyama, S. and Takaku, H. (2000). Effects of anticodon 2'-O-methylations on tRNA codon recognition in an Escherichia coli cell-free translation. *RNA* 6: 680-686.
Schlessinger, D. (1980). In Ribosomes G. Chambliss, G. R. Craven, J. Davies, K., Kahan, L. and Nomura, M. (Ed). Baltimore, University Park Press: 767-780.
Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9, 671-675.

Sherlin, L. D., Bullock, T. L., Nissan, T. A., Perona, J. J., LaRiviere, F., Uhlenbeck, O. C. and Scaringe, S. (2001). Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. *RNA* 7: 1671-1678.
Sivaraman, J., Sauve, V., Larocque, R., Stura, E. A., Schrag, J. D., Cygler, M. and Matte, A. (2002). Structure of the 16S rRNA pseudouridine synthase RsuA bound to uracil and UMP. *Nature Struct Mol Biol* 9: 353-358.
Sylvester, J. E., Gonzalez, I. L. and Mougey, E. B. (2004). Structure and organization of vertebrate ribosomal DNA. *The Nucleolus*. M. O. J. Olson(Ed), Kluwer Academic / Plenum Publishers: 58-73.
Tollervey, D. and Kiss, T. (1997). Function and synthesis of small nucleolar RNAs. *Curr Opin Cell Biol* 9: 337-342.
Trahan, C., Martel, C. and Dragon, F. (2010). Effects of dyskeratosis congenita mutations in dyskerin, NHP2 and NOP10 on assembly of H/ACA pre-RNPs. *Hum Mol Genet* 19: 825-836.
Urbonavicius, J., Durand, J. M. B. and Bjork, G. R. (2002). Three Modifications in the D and T Arms of tRNA Influence Translation in Escherichia coli and Expression of Virulence Genes in *Shigella flexneri*. *J Bacteriol* 184: 5348-5357.
Wang, J., Dasgupta, I. and Fox, G. E. (2009). Many nonuniversal archaeal ribosomal proteins are found in conserved gene clusters. *Archaea* 2: 241-251.
Warner, J. R. (1999). The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci* 24: 437-440.
Watanabe, Y. and Gray, M. W. (2000). Evolutionary appearance of genes encoding proteins associated with box H/ACA snoRNAs: cbf5p in *Euglena gracilis*, an early diverging eukaryote, and candidate Gar1p and Nop10p homologs in archaebacteria. *Nucleic Acids Res* 28: 2342-2352.
Watkins, N. J. and Bohnsack, M. T. (2012). The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip Rev RNA* 3: 397-414.
Watkins, N. J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M. and Lührmann, R. (1998). Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all box H/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA* 4: 1549-1568.
Wool , I. G. (1979). The structure and function of eukaryotic ribosomes. *Annual Rev Biochem* 48: 719-754.
Wright, J. R., Keffer-Wilkes, L. C., Dobing, S. R. and Kothe, U. (2011). Pre-steady-state kinetic analysis of the three *Escherichia coli* pseudouridine synthases TruB, TruA, and RluA reveals uniformly slow catalysis. *RNA* 17: 2074-2084.
Wrzesinski, J., Bakin, A., Nurse, K., Lane, B. G. and Ofengand, J. (1995). Purification, cloning, and properties of the 16S rRNA pseudouridine 516 synthase from *Escherichia coli*. *Biochemistry* 34: 8904-8913.
Wrzesinski, J., Nurse, K., Bakin, A., Lane, B. G. and Ofengand, J. (1995). A dual-specificity pseudouridine synthase: an *Escherichia coli* synthase purified and cloned on the basis of its specificity for psi 746 in 23S RNA is also specific for psi 32 in tRNA(phe). *RNA* 1: 437-448.
Wu, G., Xiao, M., Yang, C. and Yu, Y. T. (2011). U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *EMBO J* 30(1): 79-89.

Yang, C., McPheeters, D. S. and Yu, Y. T. (2005). Psi35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in Saccharomyces cerevisiae. *J Biol Chem* 280: 6655-6662.

Yang, P. K., Hoareau, C., Froment, C., Monsarrat, B., Henry, Y. and Chanfreau, G. (2005). Cotranscriptional recruitment of the pseudouridylsynthetase Cbf5p and of the RNA binding protein Naf1p during H/ACA snoRNP assembly. *Mol Cell Biol* 25: 3295-3304.

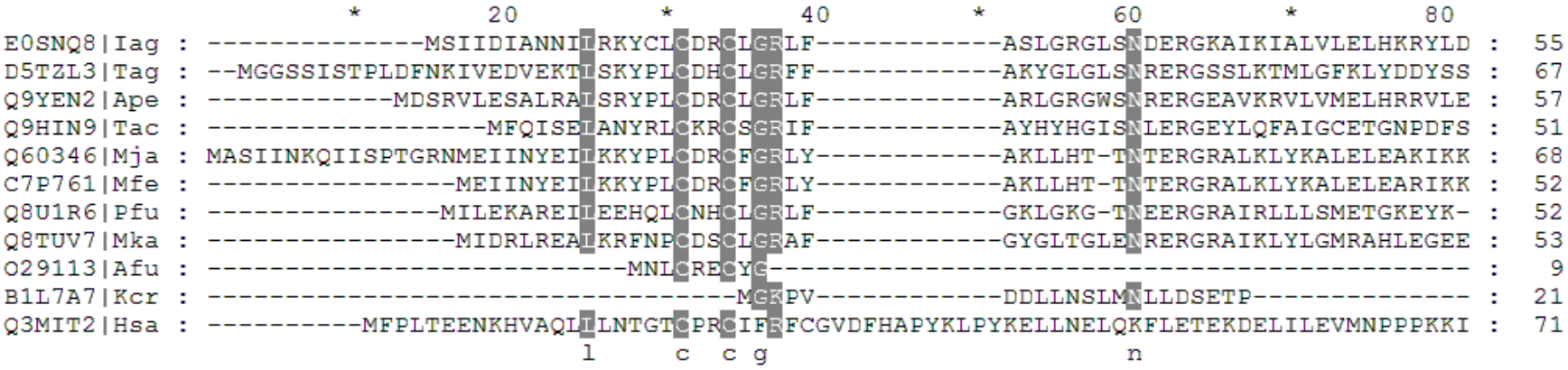
Yang, X., Duan, J., Li, S., Wang, P., Ma, S., Ye, K. and Zhao, X. S. (2012). Kinetic and thermodynamic characterization of the reaction pathway of box H/ACA RNA-guided pseudouridine formation. *Nucleic Acids Res* 40: 10925-10936.

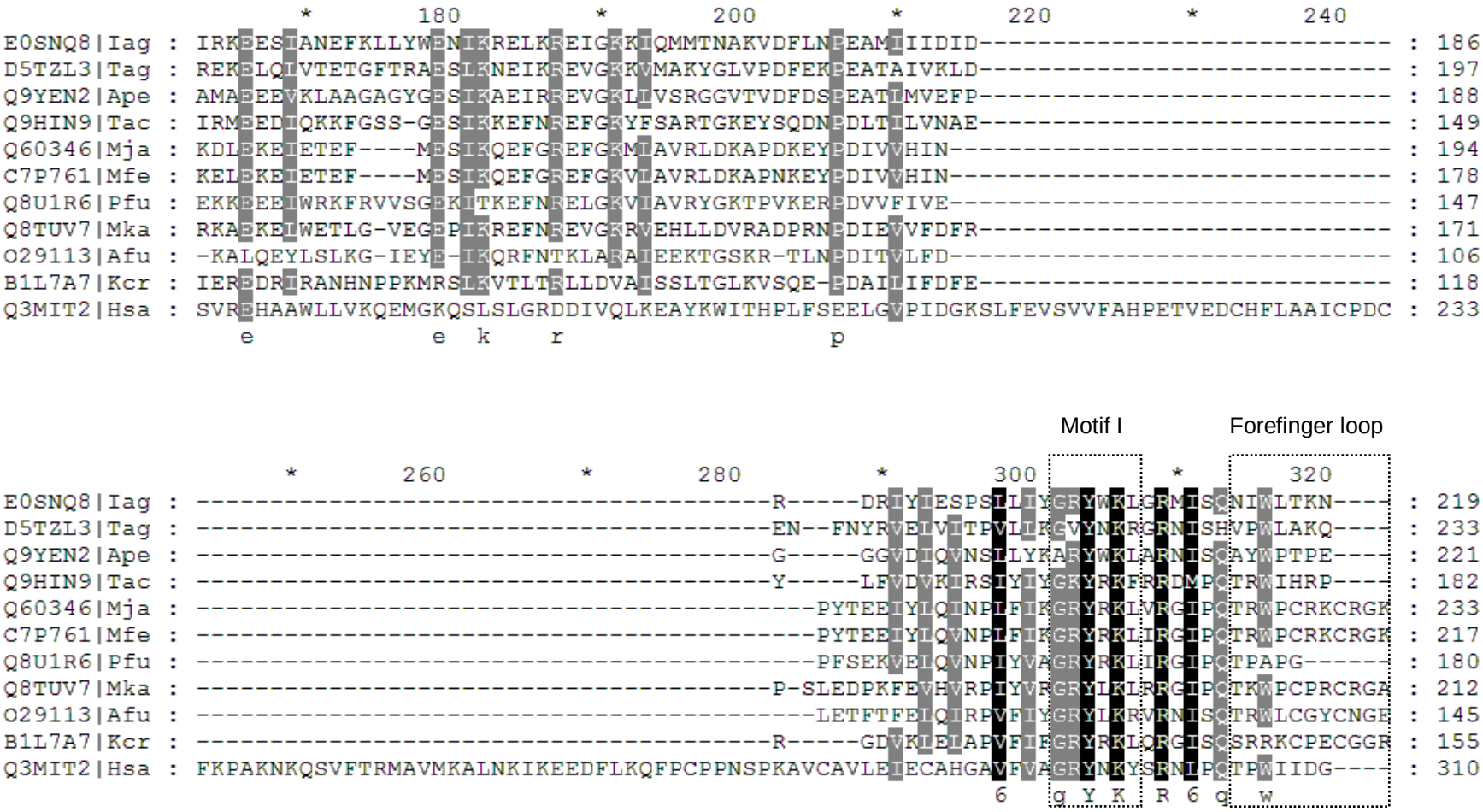
Yarian, C. S., Basti, M. M., Cain, R. J., Ansari, G., Guenther, R. H., Sochacka, E., Czerwinska, G., Malkiewics, A. and Agris, P. F. (1999). Structural and functional roles of the N1- and N3-protons of psi at tRNA's position 39. *Nucleic Acids Res* 27: 3543-3549.

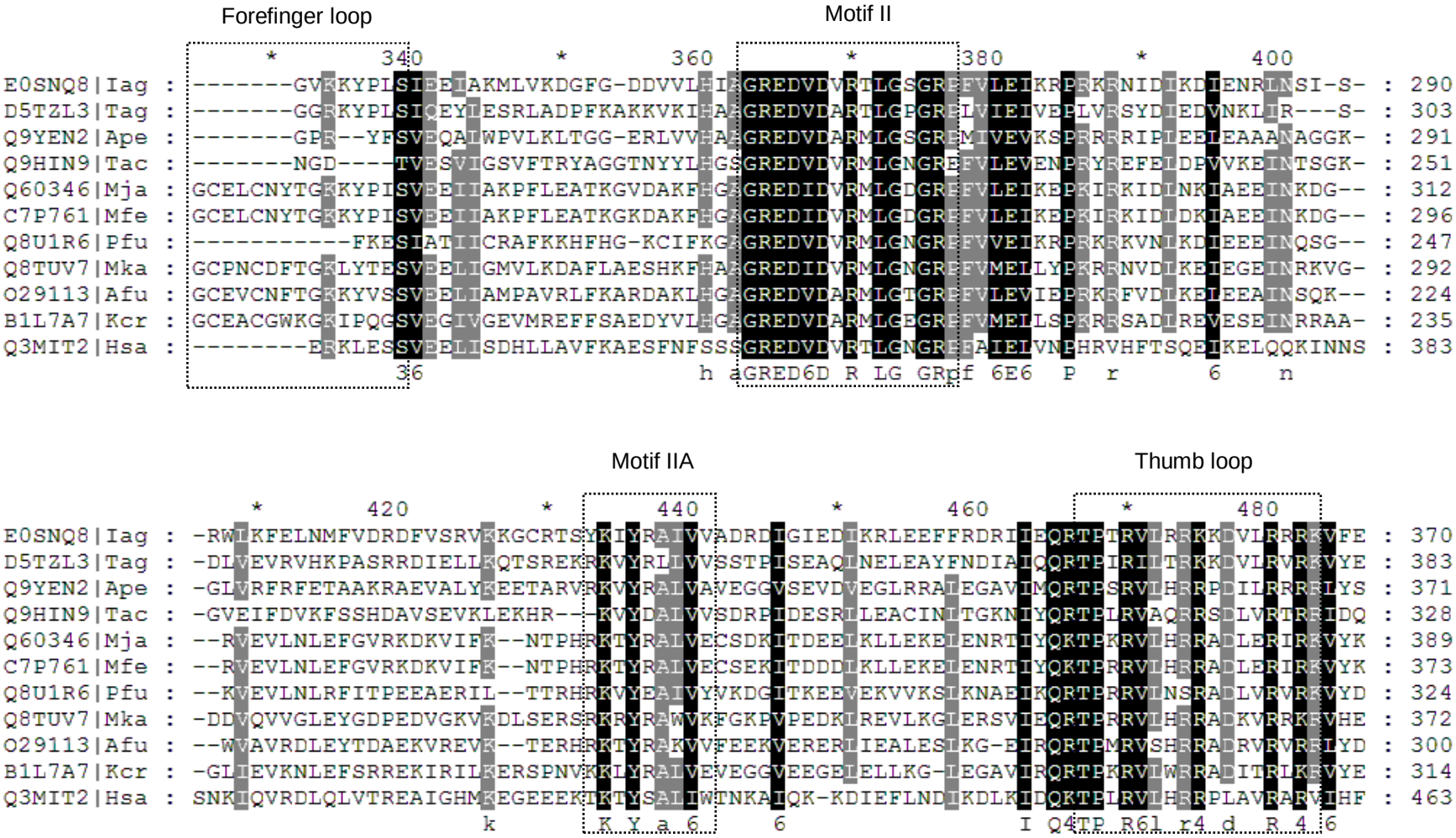
Zebarjadian, Y., King, T., Fournier, M. J., Clarke, L. and Carbon, J. (1999). Point mutations in yeast CBF5 can abolish *in vivo* pseudouridylation of rRNA. *Mol Cell Biol* 19: 7461-7472.

Zhao, X. and Yu, Y. T. (2004). Pseudouridines in and near the branch site recognition region of U2 snRNA are required for snRNP biogenesis and pre-mRNA splicing in Xenopus oocytes. *RNA* 10: 681-690.

Appendix


E0SNQ8|Iag : --------------MSIIDIANNILRKYCLCDRCLGRLF------------ASLGRGLSNDERGKAIKIALVLELHKRYLD : 55
D5TZL3|Tag : --MGGSSISTPLDFNKIVEDVEKTLSKYPLCDHCLGRFF------------AKYGLGLSNRERGSSLKTMLGFKLYDDYSS : 67
Q9YEN2|Ape : ------------MDSRVLESALRALSRYPLCDRCLGRLF------------ARLGRGWSNRERGEAVKRVLVMELHRRVLE : 57
Q9HIN9|Tac : ------------------MFQISELANYRLCKRCSGRIF------------AYHYHGISNLERGEYLQFAIGCETGNPDFS : 51
Q60346|Mja : MASIINKQIISPTGRNMEIINYEILKKYPLCDRCFGRLY------------AKLLHT-TNTERGRALKLYKALELEAKIKK : 68
C7P761|Mfe : ----------------MEIINYEILKKYPLCDRCFGRLY------------AKLLHT-TNTERGRALKLYKALELEARIKK : 52
Q8U1R6|Pfu : ---------------MILEKAREILEEHQLCNHCLGRLF------------GKLGKG-TNEERGRAIRLLLSMETGKEYK- : 52
Q8TUV7|Mka : ----------------MIDRLREALKRFNPCDSCLGRAF------------GYGLTGLENRERGRAIKLYLGMRAHLEGEE : 53
O29113|Afu : ----------------------------MNLCRECYG-- : 9
B1L7A7|Kcr : ----------------------------------MGKPV------------DDLLNSLMNLLDSETP-------------- : 21
Q3MIT2|Hsa : ----------MFPLTEENKHVAQLILNTGTCPRCIFRFCGVDFHAPYKLPYKELLNELQKFLETEKDELILEVMNPPPKKI : 71
l c c g n
E0SNQ8|Iag : GDKDSLKQLVELSPNIGPIAVNLIKNLGIEIEYTPRTCFICDNKIN-DIIDTYSQRIADIINER--HINSFVLGIRGVTSY : 133
D5TZL3|Tag : -KRIGRDSLLALAENAGEPLTRMVEKL-FAENVNPKTCFICGGRISDEWIDSVAEAVYEKIKEA--SVTRFVVGVKLGRNI : 144
Q9YEN2|Ape : GDEAALKTLVSAAPNIGEVARDVVEHLSPGSYREGGPCAVCGGRLE-SVIASAVEEGYRLIRAY--DIERFVVGVRLERGV : 135
Q9HIN9|Tac : ------------------------------FVDSKNCDICHGIFD--RFDDIYDLVAKKVGDA--QYSTFLVGSVFPQDT : 97
Q60346|Mja : AKEKGINYEEELELLKALAKSGVDEIRLEDIEIEKENCPWCRGIFNKQKMEKLLNKAIELIKEY--DFDTFLIGTHIPEEI : 147
C7P761|Mfe : AKEKEIDYKDELELLKALAKSGVDEIRLEEANAEKESCPWCRGIFSKQKMEKLLNKAIELIKEY--DFETFLVGTHLPEEI : 131
Q8U1R6|Pfu : --------------------------------EPEKCELCGGVFN--NIDKFAELCIKAAEGI--EFETFWVGSRFPEEI : 96
Q8TUV7|Mka : ETLELLARSGLEEAAAVLDDP----------PEPEPCGVCRGVLD--KVDEFAEVVACEIKDL--EFRGFVVGSRWPEEI : 119
O29113|Afu : ----------------IIGEG----------VVAEKCEACCNAFD--RVEEFAEEIVKKMSEY--EFETFNVGSRVWGSL : 59
B1L7A7|Kcr : --------------------------------SGGGCHLCDGLLS-REIGKMLRKALSEIEGV--ELSSVKSGVSLPAIL : 66
Q3MIT2|Hsa : RLQELEDSIDNLSQNGEGRISVSHVGSTASKNSNLNVCNVCLGILQEFCEKDFIKKVCQKVEASGFEFTSLVFSVSFPPQL : 152
C C g f g


Motif I
Forefinger loop


Forefinger loop
Motif II
Motif IIA
Thumb loop

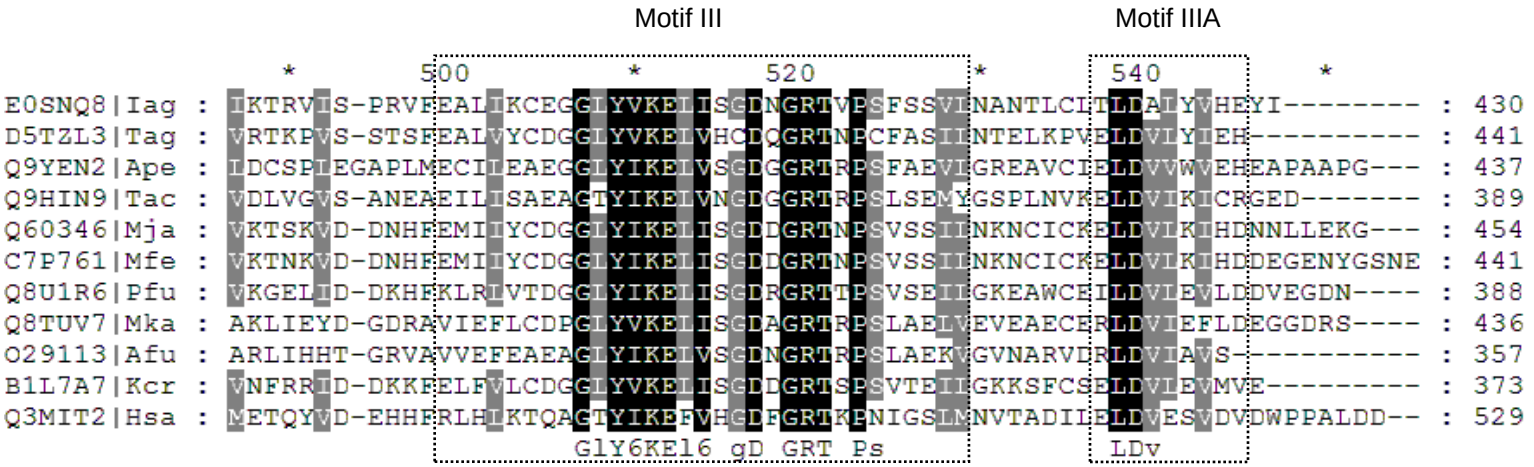


Figure A1: Alignment of Pus10 protein sequences from human and selected archaeal species.

The alignment was performed using ClustalW with Pus10 sequences representing different classes of archaea and human. Each sequence is labeled with the respective UniProt identification number. Conserved motifs, thumb loop and forefinger loop sequences in the catalytic domain are highlighted. Conserved residues are highlighted (black: 100%, grey: >75%). Iag*: Ignisphaera aggregans,* Tag*: Thermosphaera aggregans,* Ape: *Aeropyrum pernix,* Tac*: Thermoplasma acidophilum,* Mja: *Methanococcus jannaschii,* Mfe: *Methanocaldococcus fervens,* Pfu: *Pyrococcus furiosus,* Mka*: Methanopyrus kandleri,* Afu: *Archaeoglobus fulgidus,* Kcr: *Korarchaeum cryptofilum,* and Hsa: *Homo sapiens*.