CONSTRUCTION AND INITIAL CHARACTERIZATION OF YEAST STRAINS WITH DYSKERATOSIS CONGENITA MUTATIONS

ABEER BINT ABDULLAH BIN AHMAD FARAH AL-OGAILAN
Bachelor of Science, Biochemistry, King Saud bin Abdul-Aziz University in Riyadh, 2009

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

MASTER OF SCIENCE

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

CONSTRUCTION AND INITIAL CHARACTERIZATION OF YEAST STRAINS WITH DYSKERATOSIS CONGENITA MUTATIONS

ABEER BINT ABDULLAH BIN AHMAD FARAH AL-OGAILAN

Date of Defence: August 02, 2017

Dr. Ute Kothe Thesis Supervisor	Associate Professor	Ph.D.
Dr. Marc Roussel Thesis Examination Committee Member	Professor	Ph.D.
Dr. Nehalkumar Thakor Thesis Examination Committee Member	Assistant Professor	Ph.D.
Dr. Ken Vos Chair, Thesis Examination Committee	Associate Professor	Ph.D.

Abstract

The effect that Dyskeratosis congenita mutations have on ribosome biogenesis has not been fully investigated. Here, I am studying the effect of Dyskeratosis congenita mutations using a yeast model system with the goal of identifying impacts on ribosome biogenesis. I constructed six haploid yeast strains with a single mutation in the *CBF5* gene each using site-directed mutagenesis and chromosomal integration. First, I screened for phenotypic effects on growth on solid and in liquid medium at different temperatures; these experiments show normal growth behavior of the mutant strains compared to wild type. Second, Cbf5p expression levels were examined in whole-cell extract via western blotting analysis, which again did not reveal any difference to wild type. Finally, progression through the cell cycle was investigated in synchronized cells. Together, this study indicates that individual Dyskeratosis congenita mutations are well tolerated in the unicellular eukaryote *Saccharomyces cerevisiae*.

Table of Contents

List of Tables

List of Figures

List of Abbreviation

BSA	Bovine serum albumin
DKC1	Dyskerin (human homolog of yeast *CBF5*)
CMC	N-cyclohexyl-N´-(2-morpholinoethyl)-carbodiimide metho-p-toluenesulfonate
FACS	Fluorescence Activated Cell Sorter
FSC	Forward scattering
H/ACA RNA	H/ACA class of guide RNA
H/ACA scaRNA	Small Cajal body specific H/ACA RNA
H/ACA snoRNA	Small nucleolar H/ACA RNA
H/ACA sRNP	H/ACA small Ribonucleoprotein
IRES	internal ribosome entry site
LB medium	Lysogeny broth
LB-Amp	Lysogeny broth with ampicillin
Naf1	nuclear assembly factor 1
nt	nucleotide
OD_{600}	optical density at wavelength of 600 nm
PBS	Phosphate-buffered saline
PCR	polymerase chain reaction
PGK1	Phosphoglycerate kinase
pre-rRNA	precursor ribosomal RNA
Pol II	RNA polymerase II
PTC	peptidyl transferase center
PTCL	peripheral T-cell lymphoma
PUA	pseudouridine synthase and archaeosine transglycosylase domain
PUS	Pseudouridine synthases
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SSC	Side scattering
S-broth	synthetic broth
TERC	Telomerase RNA component
TERT	Telomerase reverse transcriptase
uMSC	human umbilical cord blood derived mesenchymal stem cells
YIp5	Yeast integrating plasmid 5
YEPD	yeast extract peptone dextrose
5 FOA	5 Fluoroorotic Acid

Chapter 1: Introduction

1.1 Dyskeratosis congenita – an inherited bone-marrow failure syndrome

Dyskeratosis congenita as a term was only suggested in 1930, but the disease was firstly described as a syndrome by Professor Zinsser already in 1910; therefore this syndrome is also known as Zinsser-Engman-Cole syndrome (Walne and Dokal 2008). The clinical diagnosis of Dyskeratosis congenita as a disease was very difficult as it is a progressive, multi-systemic disorder with a vast range of severities and a variety of clinical presentations such that it can be easily be mistaken for other syndromes. Practitioners at that time used to focus on one symptom only which they considered as the most prominent clinical feature. That was the case until professor Zinsser diagnosed two brothers where the youngest suffered from ear infection and unusual skin pigmentation, while the oldest suffered from symptoms including white patches in the mouth as well as abnormal nail growth. After thorough examination, it was revealed that both brothers shared the same syndrome, and only then Professor Zinsser linked these three symptoms (abnormal skin pigmentation, oral leukoplacia and nail dystrophy) together and began using them as diagnostic markers for Dyskeratosis congenita (Mason and Bessler 2011). Today, in order for a patient to be diagnosed with Dyskeratosis congenita, the patient must exhibit, in addition to the three previous classical markers, at least two triads that are related to somatic retardation. Since 1910, multiple defects in somatic systems have been documented in Dyskeratosis congenita patients that occur either sequentially or together including alopecia, dental lesion, dysphagia, dysplasia, hyperthyroid, mental retardation, anemia, epiphora, hematological abnormalities etc. In addition, patients are highly susceptible to pulmonary retardation and fibrosis, immune deficiency as well as cancer. In

the majority of Dyskeratosis congenita cases, the disease eventually progresses to reach a point where stem cells fail to regenerate leading to bone barrow failure, and this is the main cause of premature death among patients. Following that, the majority of Dyskeratosis congenita patients' deaths are related to pulmonary fibrosis and lastly malignancy (Alter, Giri et al. 2009, Bessler, Wilson et al. 2010, Dokal 2011, Angrisani, Angrisani et al. 2014).

Originally, Dyskeratosis congenita was conceived to be an X-linked disorder as all the reported cases were male (Vulliamy, Beswick et al. 2008). However, in 1963 Sarrow and Hitch reported the first female patient who had multiple Dyskeratosis congenita symptoms including the classical triads at age of 31 and had carcinoma in the cervix before her death. As no one in her family exhibited any clinical features of Dyskeratosis congenita, practitioners diagnosed her with autosomal recessive Dyskeratosis congenita (Sorrow and Hitch 1963). In 1971, a Dyskeratosis congenita case with autosomal dominant inheritance was reported after discovering another female patient with Dyskeratosis congenita clinical features whose mother was asymptomatic, while her uncle as well as her father, sister and grandmother suffered from Dyskeratosis congenita symptoms, too. In conclusion, Dyskeratosis congenita can be X-linked, autosomal recessive or autosomal dominant (Ruggero, Silvia, Piazza et al. 2003, Vulliamy, Beswick et al. 2008).

Up to this date, eleven genes (DCK1, TERC, TERT, NOP10, NHP2, TINF2, ACD, TCAB1 (WRD79), CTC1 and RTEL1) have been confirmed by genetic screening to contribute to Dyskeratosis congenita. All these genes are encoding for proteins that are involved in chromosome maintenance (Islam, Rafiq et al. 2013, Walne, Collopy et al. 2016). More recently, a research group discovered other genes (USB1, LIG4 and

GRHL2) that are homozygous variants of the Dyskeratosis congenita syndrome, and patients exhibit the classical Dyskeratosis congenita symptoms (Walne, Collopy et al. 2016).

Bone marrow failure is considered one of the classical Dyskeratosis congenita markers, and many Dyskeratosis congenita patients have stem cells with short telomeres. In eukaryotes, chromosome ends contain a DNA/protein structure known as telomere. In most tissues, telomeres get shorter directly right after the first year of life and gradually get reduced in length throughout the lifespan (Bessler, Wilson et al. 2010). This reduction is a cellular checkpoint for tissues to ensure that cells will not go through uncontrolled genetic instability during aging. However, in rapidly dividing tissues such as the bone marrow, telomere length reduction could be rather massive and therefore lethal (Marrone, Walne et al. 2005). Therefore, in these tissues, telomeres retrieve their length by activation of an enzyme known as telomerase. Eventually, rapidly dividing tissues will also have short telomeres despite telomerase function which simply is a result of aging. In general, telomere length has a direct correlation with the aging process; and therefore, a person at a certain age will typically have the same range of telomere length as the aged-matched control. Therefore, in somatic cells for instance, telomere length decreases gradually until cells undergo senescence or apoptosis (Shay and Wright 2004). However, this process is much faster among Dyskeratosis congenita patients which explains why Dyskeratosis congenita patients have abnormally short telomeres in the stem cells (Shay and Wright 2004)

The discovery of telomerase has granted Elizabeth Blackburn, Jack Szostak, and Carol Greider a Nobel Prize in 2009 (Corey 2009). Telomerase is a ribonucleoprotein, which

contains a telomerase RNA component (TERC) and proteins including telomerase reverse transcriptase (TERT, present in all organisms) and other organism-specific proteins such as H/ACA proteins (Cbf5, Nop10, Gar1, Nhp2) in mammals (Calado and Young 2008, Egan and Collins 2012). The RNA component contains a template region that is used by TERT to add a repetitive sequence to the telomeres. This enzymatic activity is crucial to maintain chromosome integrity after many rounds of cell division throughout the lifespan of an organism (Jaskelioff, Muller et al. 2011, Ale-Agha, Dyballa-Rukes et al. 2014).

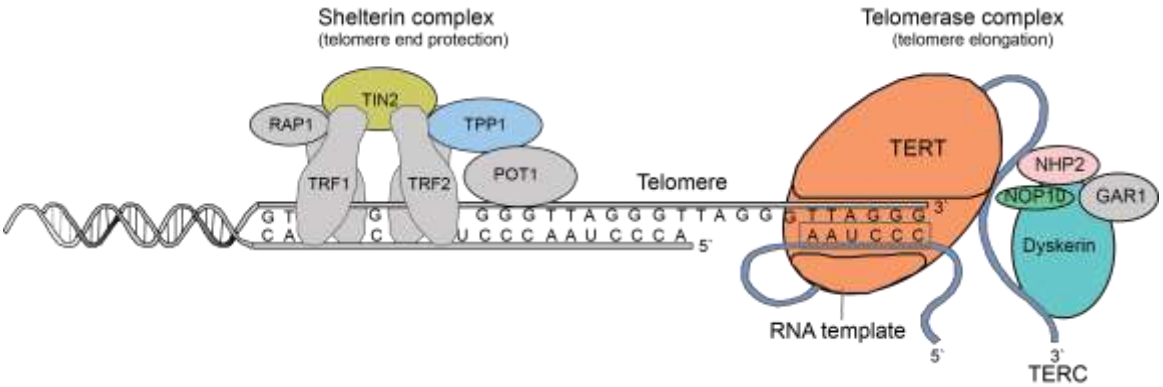


Figure 1: An illustration of complexes that are implicated in telomere maintenance. These consist of the telomerase complex on the right and the shelterin complex on the left. The protein complex of H/ACA proteins dyskerin (Cbf5), Nop10, Nhp2 and Gar1 bound to TERC is depicted on the right.

In general, the majority of Dyskeratosis congenita patients have either very short or relatively short telomeres; therefore, Dyskeratosis congenita is generally believed to be a telomere-linked disorder (Vulliamy, Beswick et al. 2008).

1.1.1 X-linked Dyskeratosis congenita

Most Dyskeratosis congenita cases belong to the X-linked version of the disease (McGrath 1999). In 1986, the Connor lab was able to narrow down the location of the mutated gene that causes X-linked Dyskeratosis congenita to the locus Xq28. In 1998,

Heiss and coworkers identified the gene and called it Dyskeratosis Congenita 1 (*DKC1*) and the corresponding protein dyskerin (Heiss, Knight et al. 1998). Most X-linked Dyskeratosis congenita patients are found to have missense mutation in DKC1. So far, there are about 50 different missense mutations documented in DKC1, in addition to one deletion and one mutation in the promoter region (Connor, Gatherer et al. 1986). Until now, researchers have not linked precisely any of the DKC1 mutation to the short telomere phenotype (Mason and Bessler 2011). The most severe form of Dyskeratosis congenita is X-linked and is known as Hoyeraal-Hreidarsson syndrome. The majority of Hoyeraal-Hreidarsson patients are male, but a female case was also reported which results from a mutation at the TERC gene. Hoyeraal-Hreidarsson patients exhibit aplastic anemia during childhood, as well as growth defects, microcephaly and development defects, and almost all of them die during childhood (Vulliamy, Beswick et al. 2008).

1.1.2 Autosomal recessive Dyskeratosis congenita

The majority of the mutated genes discovered in patients affected by autosomal recessive Dyskeratosis congenita are implicated in telomere maintenance; specifically, the identified genes are telomerase reverse transcriptase (TERT), NOP10, NHP2, TCAB1 (WRAP53), and CTC1 (Cunningham, Moreno et al. 2014, Walne, Collopy et al. 2016). Families with TERT mutations suffer from either mild or severe forms of aplastic anemia. The latter symptom indicates the presence of abnormal telomere length in hematopoietic stem cells. Many studies were performed on three generations with this mutated gene revealing that the first generation only exhibited mild anemia and somewhat shortened telomeres, but not pathogenically enough to express the disease. Other studies on five children with TERT mutations show normal blood count as well as normal telomere

length. However, during their life, their telomeres were getting shorter at a higher rate than in a healthy control (Mason and Bessler 2011) (Dokal 2011, Gu, Fan et al. 2011). Nop10 is a very small protein that contributes to telomerase stability and hence telomere maintenance. It was found to be mutated among families with consanguineous marriage (Vulliamy, Beswick et al. 2008). Both heterozygous and homozygous patients suffer from the classical Dyskeratosis congenita symptoms and also display very short telomeres as well as low TERC levels, but the disease was much more severe for homozygous individuals (Mason and Bessler 2011). Nhp2, similarly to Nop10, was found mutated in patients from consanguineous marriages. It is also a small protein that is part of the H/ACA sRNP, which is involved in telomere integrity. Individuals showed short telomeres as well as low TERC levels (Vulliamy, Beswick et al. 2008, Dokal 2011).

1.1.3 Autosomal dominant Dyskeratosis congenita

This class involves the telomerase RNA component (TERC), TINF1 and RTEL1. Human TERC is a large RNA with three conserved domains, firstly, the pseudoknot which contains the template region with the repetitive sequence, second, the CR4-CR5 domain to which TERT binds, and lastly, toward the 3' end the scaRNA domain where a set of four proteins (dyskerin, Nop10, Gar1, Nhp2) bind and protect it from degradation (Calado and Young 2008, Armanios 2009).

Many studies show that families with TERC mutations start to have short telomeres only at the end of the fourth or the fifth decade of their life, and they are recorded to be asymptomatic with mild anemia. Moreover, it is assumed that their children with more severe symptoms have inherited the mutated gene as well as the already shortened telomeres from their parents. Therefore, Dyskeratosis patients with TERC mutations (as

well as some patients with TERT mutations) may die because of haploinsufficiency of telomerase (Kirwan and Dokal 2009).

1.2 Pseudouridines

RNAs are heavily modified, and so far over 100 types of chemical modifications have been identified (Kiss 2001). Pseudouridine is the most abundant site-specific nucleotide modification that is enzymatically and post-transcriptionally introduced into many RNAs (Figure 2) (Roovers, Hale et al. 2006). Although both pseudouridine and uridine share an identical molecular mass and the same UV spectrum, pseudouridine is able to contribute to RNA stability through an additional hydrogen bond to the remaining RNA nucleotides (Spenkuch, Motorin et al. 2014). In fact, the interaction between pseudouridine and adenine is more stable than the one between U-A within a helix or at the end of RNA helices (Hudson, Bloomingdale et al. 2013).

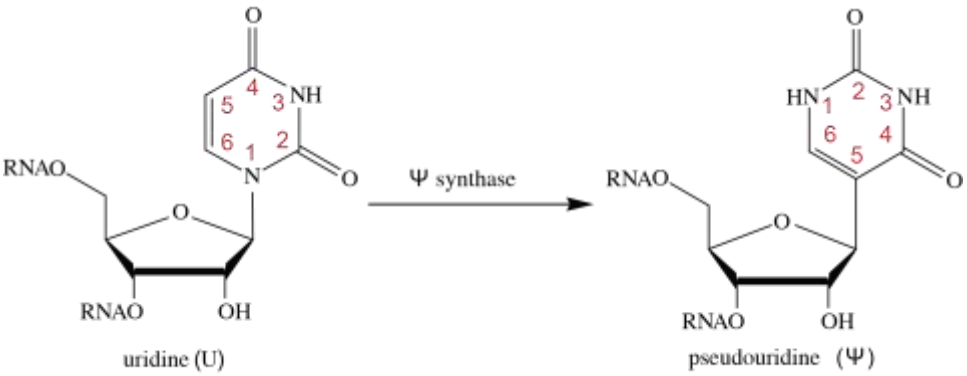


Figure 2: Structural differences between uridine and pseudouridine**.** Uridine is the substrate for pseudouridine synthases, which they isomerize to pseudouridine.

The number of pseudouridines in bacteria is limited, and knocking out any of the pseudouridine synthase genes has no lethal effect on bacteria. However, the number of pseudouridines surges in existence in eukaryotes to reach 100 in human rRNA. So far,

thirteen pseudouridine synthases have been identified in humans (Antonicka, Choquet et al. 2016, Rintala-Dempsey and Kothe 2017).

Scientists are not yet able to fully elucidate the structural or functional role of pseudouridine within the target RNA. However, pseudouridines are clustered in locations that are known to have a vital function within the RNA such as the peptidyltransferase region in rRNA (Dunin-Horkawicz, Czerwoniec et al. 2006). It has been suggested that pseudouridine increases the overall stability of rRNA structure by providing an extra H bond that coordinates a water molecule bridging to the RNA phosphate backbone (Arnez and Steitz 1994). This notion was strengthened by the finding that the tRNA structure is rigidified by the presence of pseudouridines (Yarian, Basti et al. 1999). Specifically, pseudouridines at positions 38, 39 and 55 are found to stabilize tRNA structure (Arnez and Steitz 1994). Moreover, in particular pseudouridines in the anticodon domain of tRNA are enhancing anticodon–codon interactions which indicates that pseudouridines are increasing the accuracy of the translation when compared to the unmodified tRNA (Harrington, Nazarenko et al. 1993, Davis, Veltri et al. 1998).

Pseudouridines are likely integrated to stay; therefore, there are so far no enzymes known in mammals that can reverse pseudouridylation. Pseudouridines leave the organism through urine as a nucleoside. An enzyme has been identified that can dephosphorylate pseudouridines in humans. This enzyme known as pseudouridine-5'-phosphatase might be important to prevent random incorporation of pseudouridines in RNAs during transcription (Goldberg and Rabinowitz 1961, Preumont, Rzem et al. 2010, Spenkuch, Motorin et al. 2014).

Detection of pseudouridine sites in RNAs is a challenging task. Often, site-specific pseudouridylation in rRNA was predicted based on sequence complementarity to a guide

RNA. Detection of site-specific pseudouridylation includes extraction of the total RNAs followed by isolating rRNA as 80% of total RNAs is rRNA. Then the RNA mixture is treated with *N*-cyclohexyl-*N′*-(2-morpholinoethyl)-carbodiimide metho-*p*-toluenesulfonate (CMC) and subsequently with strong alkaline treatment to remove any unspecific association of CMC with uridines and guanines. The next step is a reverse transcriptase assay that results in a premature stop in the presence of CMC-labelled pseudouridine (Bakin and Ofengand 1993). The disadvantage of this procedure is that in order to reach a detectable threshold, the RNA species should be highly expressed, which is not the case for the majority of the RNAs species. Moreover, many RNAs get degraded quickly during the CMC and the alkaline treatments. Initially, this was the only detection method for pseudouridines at a specific site in a specific RNA. Nowadays, there is global pseudouridine detection available by treating all cellular RNAs with CMC which is then followed by next-generation RNA sequencing. With this new detection method, many additional pseudouridine sites have been found in a vast range of noncoding RNAs as well as mRNAs. For example, in addition to rRNA and cytosolic tRNA, many pseudouridines sites are found in mitochondrial tRNA, small nuclear RNAs, H/ACA sRNA, C/D sRNA, telomerase RNA, 7SK snRNA and many other RNA species (Alfonzo 2014, Carlile, Rojas-Duran et al. 2014, Lovejoy, Riordan et al. 2014, Schwartz, Bernstein et al. 2014).

1.2.1 Pseudouridine synthases

Pseudouridine synthases can be subcategorized into two groups: 1) standalone pseudouridine synthases (called the PUS family in eukaryotes) and 2) RNA-guided pseudouridine synthases (H/ACA sRNP). Standalone pseudouridine synthases are enzymes that have been found in all domains of life. All standalone pseudouridine synthase have one or a few known uridine target sites. Standalone pseudouridine synthases are proteins that are conducting pseudouridine formation on target uridines by sequence recognition or/and structural recognition of the substrate RNA. They are targeting mainly tRNA, mitochondrial rRNA and some mRNAs (Rintala-Dempsey and Kothe 2017). The second type of pseudouridine synthases (H/ACA sRNP) requires complex formation of RNA and protein prior to targeting the substrate uridines that are mostly found in rRNAs and small nuclear RNAs, but also some mRNAs.

All pseudouridine synthases share a very similar catalytic domain that possesses a catalytic aspartate residue. In general, the activity of these enzymes does not require any energetic investment such as ATP; pseudouridine synthases simply break the glycosidic bond between the N1 and C1' atoms, then the bound is formed again between two C5-C1' atoms. Recently it has been proven that the catalytic aspartate deprotonates C2' in the ribose ring that leads to breakage of the glycosidic bound and formation of a glycal intermediate (Veerareddygari, Singh et al. 2016). Then the uracil ring is either flipped or rotated and forms a new carbon-carbon glycosidic bound.

A human disorder has been identified that arises as a result of mutations in the PUS1 gene. Pseudouridine target sites in tRNA species (mitochondrial and cytoplasmic tRNAs) of this enzyme were not implemented despite the normal cellular localization of PUS1

(Bykhovskaya, Casas et al. 2004, Fernandez-Vizarra, Berardinelli et al. 2007) (Patton, Bykhovskaya et al. 2005).

1.2.2 H/ACA small Ribonucleoproteins (sRNPs)

RNA-guided pseudouridine synthases have been identified in archaea and eukaryotes. RNAs that guide pseudouridylation belong to the H/ACA class of noncoding RNAs. Upon transcription, these H/ACA guide RNAs are transported inside the cell either to the nucleolus or the Cajal body. Regardless of their final destination, all H/ACA RNAs bind to the same set of proteins and guide formation of pseudouridines in a vast range of RNA species. All H/ACA RNAs share one consensus secondary structure which is the 5'-hairpin-hinge-hairpin-tail-3' structure. In general, H/ACA sRNAs are 100-150 nucleotides (nt) in length, but there are a few H/ACA RNAs that reach around 250 nt and more (snR30). The length of each hairpin ranges from 60 to 75 nt. Both the hinge and the tail regions are stretches of single-stranded sequences. The hinge region contains a conserved motif (AnAnnA) called H box that follows the first hairpin. The tail has also a conserved sequence (Lermontova, Schubert et al. 2007) called an ACA box that follows the 3' hairpin. The two boxes are serving as a signal for the protein complex to bind as well as for subcellular localization (Ueda, Calado et al. 2014). Each hairpin contains two stems, the proximal P1 stem and distal P2 stem, that are separated by a large internal loop that has unpaired sequences; this internal loop is known as the pseudouridylation pocket. Each pseudouridylation pocket has a potential antisense target sequence in noncoding RNAs, mainly cytoplasmic rRNA and snRNAs, but also mRNAs (McMahon, Contreras et al. 2015, Li, Ma et al. 2016).

Substrate RNA can be paired to the sequence in the pseudouridylation pocket of an

H/ACA guide RNA in a way that the target U and the adjacent nucleotide are always unpaired; meanwhile about 3-10 nucleotides that are flanking the target U are paired to sequences in the pseudouridylation pocket. It has been proposed that this precise arrangement is needed for the isomerization of uridine to pseudouridine (McMahon, Contreras et al. 2015).

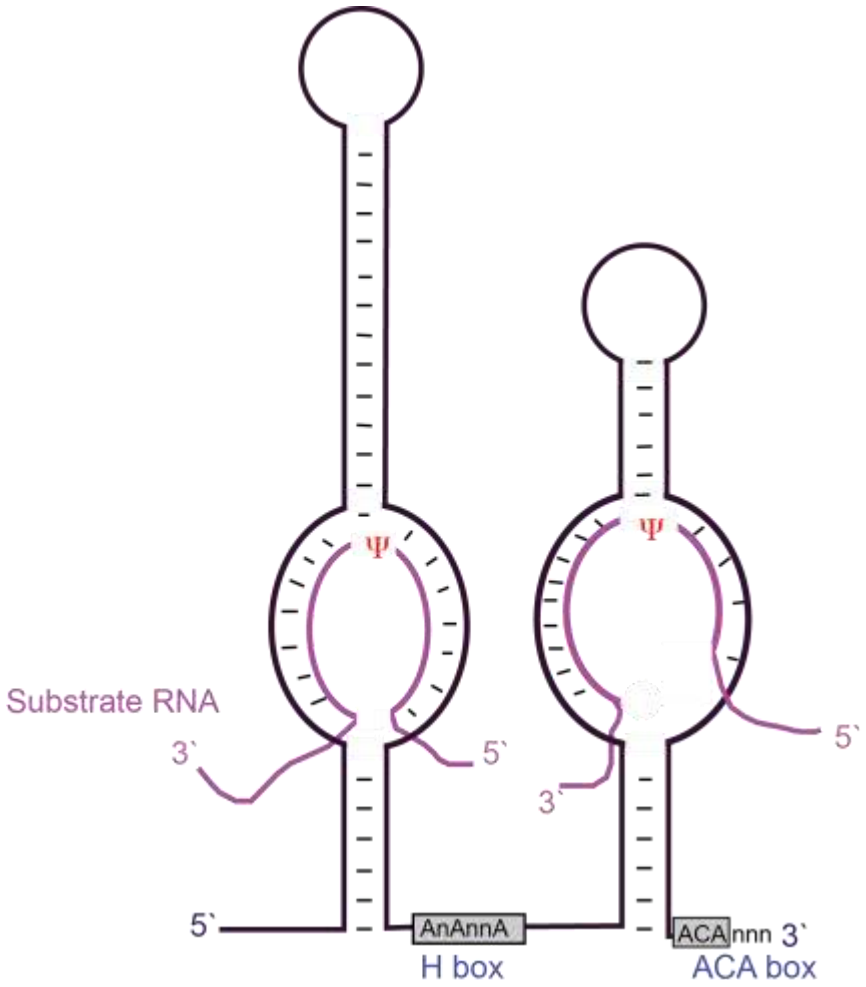


Figure 3: The consensus secondary structure of H/ACA guide RNA. It consists of two hairpins separated by the H box while the second hairpin is followed by the ACA box. Each stem has two pockets that base-pair to the substrate RNA shown in pink. The position of the modified U is shown in red as a psi symbol representing pseudouridine.

Unlike standalone pseudouridine synthases which have one to a few target sites in specific RNAs, H/ACA sRNPs target many uridines in many RNAs species. H/ACA

sRNPs are composed of one guide H/ACA RNA and four proteins, Cbf5 in yeast (Dkc1 in human), Nop10, Nhp2 and Gar1 (Figure 4). H/ACA sRNPs are conserved across archaea and eukaryotes, but are not found in bacteria. Notably, more than 10% of the human genome can be transcribed into one subtype of H/ACA sRNAs known as AluACA RNAs which all bind to the same set of proteins(Ketele, Kiss et al. 2016). The remaining types of H/ACA RNAs also bind to the same four proteins. H/ACA sRNPs perform many vital functions that will be described latter in this thesis.

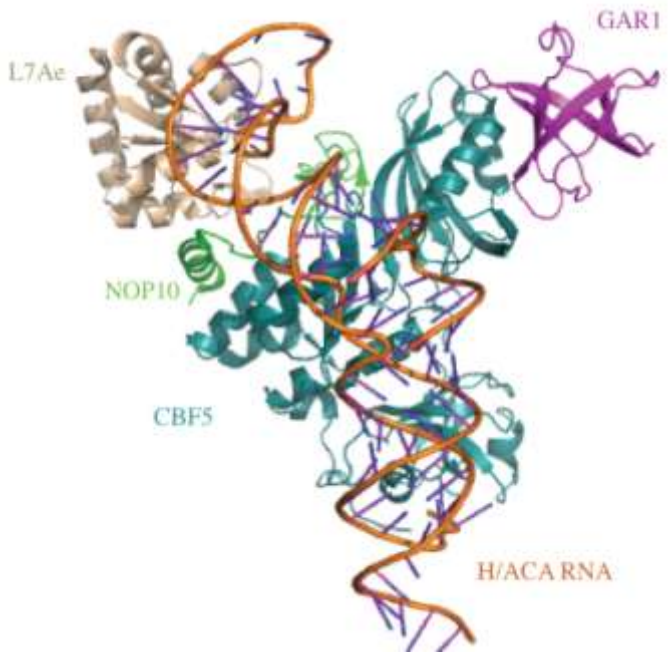


Figure 4: Crystal structure of an H/ACA sRNP obtained from *Pyrococcus furiosus* (PDB ID: 3HAY). The protein complex consists of L7Ae (Nhp2 homolog) in sandy color, bound to Nop10 in light green, Cbf5 in turquoise and Gar1 in pink. The guide RNA bound to L7Ae-Nop1-Cbf5 is shown in orange.

Cbf5 was originally thought to be a centromere binding protein (centromere binding factor 5) giving Cbf5 its name (Jiang, Middleton et al. 1993). This protein is essential in yeast and eukaryotes as it cannot be deleted. It is the only pseudouridine synthase that directs pseudouridylation in yeast rRNA, as substitution of aspartate 95 in yeast Cbf5

with alanine resulted in total absence of pseudouridines in rRNA (Zebarjadian, King et al. 1999). Based on sequence similarities, Cbf5 belongs to the pseudouridine synthase family known as the TruB family. Both enzymes have a similar catalytic domain as well as a conserved RNA-binding domain known as pseudouridine synthase and archaeosine transglycosylase (PUA) domain (Zucchini, Strippoli et al. 2003). Most knowledge of the H/ACA sRNP structure stems from studies of the archaeal H/ACA sRNP complex (Gupta, Gurha et al. 2007, Xue, Liang et al. 2007). Reconstituting the yeast H/ACA sRNP complex is rather difficult, but in 2011, the crystal structure of the yeast trimeric protein complex Cbf5-Nop10-Gar1 was reported (Li, Duan, Li, Yang et al. 2011). *Saccharomyces cerevisiae* Cbf5 contains a catalytic domain compromised of residues 39-292 which are divided into two subdomains D1 and D2. Both subdomains are joined by a structure known as the catalytic cleft. The PUA domain is highly positively charged which explains its affinity for RNA. It consists of residues in the N terminus (1-32) and the C terminus (266 -351) (Li and Ye 2006). The remaining sequence (352-483) is an unstructured tail that contains 10 tandem KKE/D repeats. The PUA domain recognizes the Box H and Box ACA in the H/ACA RNAs (Rashid, Liang et al. 2006, Li, Duan, Li, Yang et al. 2011). The PUA domain is one of the most ancient RNA recognition domains, and it consists of about 64-96 residues forming both β and α secondary structures. It is found in more than 600 catalytic enzymes from different families in all domains of life including archaea and eukaryotes. Many of these enzymes and proteins are known to be associated with tRNA, proline synthesis, translation initiation, rRNA modifications and ribosome biogenesis (Cerrudo, Ghiringhelli et al. 2014). In most cases, the PUA domain is located in the C terminus of the enzyme (Cerrudo, Ghiringhelli et al. 2014). To this date, there is no eukaryotic crystal structure of the complete H/ACA sRNP, but in the

crystal structure obtained from *P. furiosus* H/ACA sRNP, the PUA domain interacts with the double-stranded guide RNA P1 stem and recognizes the ACA (and the AnAnnA) sequence (Pérez‐Arellano, Gallego et al. 2007). The PUA domain in Cbf5 is not part of the catalytic domain, but it helps in positioning the target uridines in the active site (Li and Ye 2006). In yeast, the PUA domain is much larger than in archaeal Cbf5.

Nop10 is highly unstructured and contains few hydrophilic amino acids that contribute to its tertiary structure stability. It has an elongated structure where the C terminus contains an α helix and is connected to the N terminus ribbon domain via a linker. Nop10 binds to Cbf5 at the D1 subdomain, but also touches the D2 surface as well (Li, Duan, Li, Yan et al. 2011).

Nhp2 is the only protein in the complex that cannot associate directly with Cbf5, rather, it is believed to attach to Nop10 as its archaeal L7Ae homolog. Two cases of Dyskeratosis Congenita have been linked to Nhp2 mutations that impair its binding to Nop10 (Trahan, Martel et al. 2010). Unlike L7Ae, the archaeal homolog of yeast Nhp2, the latter does not recognize k-turn motifs in the guide RNA. However, it has been found that yeast Nhp2 can bind different RNAs non-specifically *in vitro* (Henras, Dez et al. 2001). The ternary complex of Nhp2-Nop10-Cbf5 is recruited to the H/ACA sRNA which aids Nhp2 to interact with the guide RNA (Koo, Park et al. 2011).

Yeast Gar1 is named after the two glycine/arginine rich stretches that are flanking a core domain of residues 32-124. Interestingly, Gar1 was the first protein to be purified and identified in the H/ACA sRNP complex (Ganot, Caizergues-Ferrer et al. 1997), and yet it is the last protein to be assembled to the H/ACA sRNP complex as the Gar1 binding site on Cbf5 is initially pre-occupied by the assembly factor Naf1. Gar1's functional

contribution to the H/ACA sRNPs is to facilitate substrate turnover by interacting with the thumb loop in Cbf5 (Wang, Yang et al. 2015).

1.3 H/ACA RNA expression

The maturation process of H/ACA sRNPs is highly coordinated. In yeast, most snoRNAs are transcribed from independent genes; however, a small fraction of snoRNAs is encoded by introns in protein-coding genes. As known so far, the majority of H/ACA sRNAs in humans are transcribed as introns of a mRNA, therefore their transcription is regulated together with the protein-coding gene and they lack independent expression regulation. Most human introns are rapidly degraded upon release from pre-mRNA; therefore, the most important step in H/ACA sRNP biogenesis is facilitating the initial binding of H/ACA sRNAs to the RNP complex thus protecting the nascent H/ACA sRNAs from degradation (Bousquet-Antonelli, Henry et al. 1997, Dez, Henras et al. 2001, Darzacq, Kittur et al. 2006, Hoareau-Aveilla, Bonoli et al. 2006). The H/ACA sRNA is released from introns of pre-mRNAs by exonucleases (Kiss and Filipowicz 1995). All H/ACA sRNAs that have been identified to this date are transcribed within pre-mRNA by RNA-Pol II. Processing of H/ACA sRNA is usually independent from its location relative to the splicing site in the pre-mRNA. In addition, H/ACA proteins can be recruited to the H/ACA sRNA even before it is processed from the host nascent mRNA during transcription or shortly after. Processing of some H/ACA sRNAs such as U64 is governed by the RNA polymerase II machinery (Pol II) as H/ACA RNA cannot be accurately processed when transcribed by Pol I or Pol III rather than Pol II. (Richard, Kiss et al.

2006).

Interestingly, many snoRNAs as well as scaRNAs are transcribed as part of a host mRNA for proteins that are implicated in ribosome biogenesis (Lestrade and Weber 2006). This raises the possibility that H/ACA sRNPs levels can be used as a biological marker to reflect an increase of cellular activity such as growth (McMahon, Contreras et al. 2015). Many protein levels are increased in response to growth including ribosomal proteins and proteins involved in ribosome biogenesis. A few studies show that indeed H/ACA sRNPs levels are reflecting an increased cellular demand on protein production. For example, it was shown that Dkc1 transcript levels are up-regulated in response to the oncogene c-Myc (Alawi and Lee 2007). This oncogene is one of the human proteins that is highly altered in cancer. C-Myc coordinates the activation of multiple genes that are implicated in cellular proliferation including genes for ribosome biogenesis as well as telomerase activation (Felsher, Yetil et al. 2010).

Until now, it is not fully known which factors regulate H/ACA sRNA transcription, and whether the regulation itself is different between the H/ACA sRNAs that guide pseudouridylation and the ones with other cellular functions. It has been shown that the expression levels of distinct sets of H/ACA sRNAs are tissue-specific (Cavaille, Buiting et al. 2000, Castle, Armour et al. 2010, Bellodi, McMahon et al. 2013). Furthermore, the expression levels of a particular subset of H/ACA sRNAs are deregulated in cancer (Bellodi, McMahon et al. 2013).

1.3.1 Assembly and localization of H/ACA sRNPs

Assembly of H/ACA snoRNAs and the H/ACA proteins requires two essential chaperones, Shq1 and the nuclear assembly factor 1 (Naf1) (Grozdanov, Roy et al. 2009). So far, Naf1 and Shq1 are exclusive proteins to eukaryotes, which indicates potential functional specificity (Fatica, Dlakic et al. 2002). In yeast for example, a significant loss of H/ACA snoRNAs was detected upon depletion of Shq1 and Naf1. It has been found that Naf1 stabilizes newly transcribed H/ACA snoRNA, and its recruitment to the newly transcribed RNAs is Dkc1 dependent; interestingly, the Dkc1 level was also dependent on Naf1 (Darzacq, Kittur et al. 2006). One potential explanation is based on a study using chromatin coimmunoprecipitation to show that a complex of Naf1-Cbf5 is recruited to the H/ACA snoRNA genes (Ballarino, Morlando et al. 2005). In fact, an early association of H/ACA snoRNAs to the trimeric complex of Cbf5-Nop10-Nhp2 (but not Gar1) was found at the site of H/ACA sRNAs transcription. It was suggested that Naf1 is involved in this early steps of H/ACA snoRNP biogenesis, but upon maturation of the complex, it gets replaced by Gar1 (Darzacq, Kittur et al. 2006). Interestingly, Naf1 and Shq1 were never detected binding simultaneously to Cbf5 (Grozdanov, Roy et al. 2009). Unlike Naf1, Shq1 is not found at the final destinations (nucleolar and the Cajal body) of the mature H/ACA sRNP complex. Moreover, it was not found near the newly transcribed H/ACA sRNAs genes. In fact, Shq1 binds *in vivo* only to Cbf5 alone and is not associating with any of the H/ACA sRNPs or Naf1 (Grozdanov, Roy et al. 2009). Consequently, Shq1 is involved in much earlier steps than Naf1 in H/ACA sRNPs biogenesis.

1.3.2 H/ACA sRNAs

The class of H/ACA small RNAs was identified almost 50 years ago, but so far the functional role of this particular type of RNA is not completely understood. Deregulation of H/ACA sRNAs is implicated in diseases including leukemia (Valleron, Laprevotte et al. 2012). The H/ACA sRNA class can be divided into two main subgroups according to their cellular location: 1) small nucleolar RNA (H/ACA snoRNAs) and 2) small Cajal body specific RNA (H/ACA scaRNAs).

1.3.3 H/ACA snoRNAs

As was described before, these guide RNAs have a consensus secondary structure of two hairpins and contain characteristic sequence elements with the H box and the ACA box. The majority of H/ACA snoRNAs guide pseudouridylation of many uridines in the nascent pre-rRNAs. However, some H/ACA snoRNAs have no known target substrate and therefore they have been further classified as orphan H/ACA snoRNAs (Vitali, Royo et al. 2003, Fatica and Bozzoni 2014). Many studies show that H/ACA snoRNAs can execute functions other than guiding pseudouridylation. In this context, many H/ACA sRNAs exhibit tissue specificity indicating that they could conduct variety of cellular functions. Interestingly, in some human diseases, H/ACA sRNAs level are distorted. Moreover, a subset of H/ACA sno/caRNAs are altered in human diseases, but it is not known how and why H/ACA sRNA expression is regulated. Nevertheless, H/ACA sRNAs have been used as biological markers for certain cancers such as leukemia (Ronchetti, Mosca et al. 2013, Mannoor, Shen et al. 2014). The majority of snoRNAs, that are deregulated in cancers, do not display genetic alterations in their genes. However,

a group of snoRNAs has been reported with either deletions or nucleotide substitutions which are implicated in human disorders, predominantly cancer (Skreka, Zywicki et al. 2012, Valleron, Laprevotte et al. 2012, Valleron, Ysebaert et al. 2012). Some of the cellular functions that have been designated to H/ACA sRNAs will be discussed below including hematopoietic stem cell differentiations, stem cell biogenesis, chromatin biology and other cellular processes such as cholesterol metabolism.

1.3.4 Hematopoietic stem cell differentiations

The levels of a specific type of H/ACA sno/scaRNAs are exclusively altered in hematopoiesis-related pathologies (Ronchetti, Mosca et al. 2013) including acute myeloblastic leukemia (Kocak, Ballew et al. 2014) (Valleron, Laprevotte et al. 2012), multiple myeloma (Ronchetti, Todoerti et al. 2012) and peripheral T-cell lymphoma (PTCL) (Valleron, Ysebaert et al. 2012). These H/ACA sno/caRNAs are sometimes up-regulated, but in most of the cases they are down-regulated. Interestingly, there are some cases where the down regulation of these H/ACA sno/scaRNAs is independent from the host gene level (Ronchetti, Todoerti et al. 2012). In this context, it is not fully known yet if there is a post-transcriptional regulation that is dynamically involved in H/ACA sno/sca RNA biogenesis in these cells. Nevertheless, the main function of most of these affected H/ACA sno/scaRNAs is guiding pseudouridine formation in rRNAs, and there are some that are involved in spliceosome RNAs modification.

One study examined hematopoietic stem cell differentiation in CD34+ cells with a mutation in the promoter region of the Dkc1 gene. Firstly, the authors tested H/ACA

snoRNA and scaRNA levels as well as TERC abundance and telomere length. While both TERC levels and telomere length were normal, there was a heterogenic reduction in some H/ACA sRNA populations, in particular of two H/ACA snoRNAs that are guiding pseudouridine formation in 18S rRNA (Bellodi, McMahon et al. 2013). Furthermore, the authors tested cells from different Dyskeratosis congenita patients and different tissues and found that H/ACA sRNA levels were different across tissues. Then researchers transfected cells with WT Dkc1, which induced H/ACA sRNA transcription to reach almost normal levels. When they tested pseudouridine levels in rRNA in these cells, they discovered that indeed specific pseudouridines were reduced in rRNA corresponding to the down-regulated guide H/ACA sRNA. Moreover, the ability of hematopoietic cells to differentiate was rescued by introducing WT Dkc1 (Bellodi, McMahon et al. 2013). The observation that H/ACA sRNAs are tissue specific raises the question of whether the function of H/ACA sRNAs in pseudouridine formation is vital or whether these H/ACA sRNAs are involved in other hematopoiesis-related functions in these particular cell types.

1.3.5 Stem cell biogenesis

There are many types of stem cells such as for example uMSC (human umbilical cord blood derived mesenchymal stem cells). This type of pluripotent stem cell can differentiate into many types of cells, which makes it a very good candidate for cell-based therapy. In this particular cell line, one type of H/ACA snoRNA (SNORA7A) is crucial for regeneration of the cells as well as their differentiation abilities. When SNORA7A

expression is reduced using siRNAs, uMSC severely fail to proliferate. And when researchers compromised the structure of SNORA7A by expressing SNOR7A with a mutation that prevents it from binding to the H/ACA proteins, the uMSC cells fail to self-renew and differentiate (Zhang, Xu et al. 2016). Interestingly, Dkc1 seems to be important for the H/ACA sRNA to fulfill its function in these cells (Zhang, Xu et al. 2016).

1.3.6 H/ACA snoRNAs and chromatin

A significant population of snoRNAs such as ACA44 has been found to stably interact with chromatin in human cells and *Drosophila* cells (Schubert, Pusch et al. 2012). These snoRNAs are enriched within chromatin-associated RNAs. Our understanding of the functional role of these particular H/ACA snoRNAs for chromatin is still limited. Another type of noncoding RNA known as C/D snoRNA associates with chromatin binding protein Df31 and promotes the open state of chromatin. Therefore, it is possible that H/ACA snoRNAs are also conducting similar functions as well (Schubert, Pusch et al. 2012, McMahon, Contreras et al. 2015). Surprisingly, the interaction of the H/ACA snoRNAs with chromatin is independent from the H/ACA proteins. Scientists did not detect any of the protein complexes associated with chromatin by mass spectrometry. This study provides new evidence that H/ACA sRNAs could participate in cellular functions independently from their canonical protein complex (Schubert, Pusch et al. 2012). If H/ACA sRNAs are able to perform different cellular functions, they could potentially have other protein partners that enable them to do so.

1.3.7 H/ACA sRNAs and micro RNAs

Interestingly many H/ACA snoRNAs that are involved in rRNA biogenesis can serve as precursors for much smaller RNAs that behave like microRNAs of which many are conserved among eukaryotes including humans (Scott, Avolio et al. 2009, Taft, Glazov et al. 2009). These RNAs are "like" microRNAs in the sense that they get processed mostly from the 3' hairpin of H/ACA sRNAs by Dicer and are associated with Ago1 and Ago2 in humans (Ender, Krek et al. 2008). The latter proteins form complexes with H/ACA scaRNA-like miRNAs and control gene expression at the post-transcriptional level (Ender, Krek et al. 2008). In addition, further investigations at the genomic level of 20 microRNAs in humans revealed that they have evolved from H/ACA snoRNAs. Five members of these microRNAs are found to bind to Dkc1 (Scott, Avolio et al. 2009). It is not clear whether snoRNA-like microRNAs are the byproduct of H/ACA sRNAs degradation or whether they possess additional functions. One proposed mechanism hindering H/ACA RNA-like microRNA transcription is hypermethylation of the gene (Ferreira, Heyn et al. 2012, Lopez-Serra and Esteller 2012).

1.3.8 Small Cajal body-specific RNAs targeting snRNAs

In eukaryotes, the majority of mRNAs are transcribed as pre-mRNAs that undergo maturation by removal of introns either by self-cleavage or by the spliceosome. The latter constitutes a massive RNP structure of almost 100 proteins and five snRNAs (U1, U2, U4, U5 and U6). U2 snRNA is exceptionally important as it contains a region to interact with the pre-mRNA intron. U2 is heavily modified, and the bulk of the nucleotide

modifications are pseudouridines that represent about 60% of the total modifications (Yu, Ge et al. 2011). This heavily-modified region is the branch site recognition region and contains three conserved pseudouridines. These pseudouridines are important for the accurate structure of U2 snRNA as well as for activation of the ATPase activity of one of the proteins that is implicated in the splicing mechanism (Wu, Adachi et al. 2016). Many pseudouridines are also found in regions of U6 and U4 that are important for base paring as well as in the U6 catalytic core (Wu, Yu et al. 2011). Although some of these pseudouridines are introduced by standalone pseudouridine synthases, some pseudouridines, which are conserved and also play fundamental roles in the spliceosome machinery, are generated by H/ACA scaRNP complexes. These guide H/ACA sRNAs reside in Cajal bodies, and they are structurally very similar to H/ACA snoRNAs with only a slight difference as they contain a CAB signal for localization (Marz, Gruber et al. 2011).

1.3.9 Intron-derived noncoding RNAs with snoRNA ends (sno-IncRNAs)

The class of H/ACA sno-IncRNAs is poorly studied. Members of this class are found to be tissue specific, conserved and expressed in humans and monkeys (Zhang, Yin et al. 2014). A group of these sno-IncRNAs are flanked by H box sequences at their 5' end and ACA box sequences at the 3' end. It is not known yet if these RNAs bind to the same H/ACA proteins or to other proteins. The function of these sno-IncRNAs remains unexplored as well (Baker 2011, Zhang, Yin et al. 2014).

1.3.10 Intronic AluACA RNAs

About 10% of the human genome is encoding a subclass of H/ACA sRNA known as AluACA RNA. Despite this massive occupancy in the human genome, the functional role of AluACA RNAs is still unclear. Strikingly, expression of AluACA RNA is extremely low, and it is not fully known if the level of expression is coupled to stress. Mature AluACA RNAs successfully bind to the H/ACA proteins (Jády, Ketele et al. 2012, Ketele, Kiss et al. 2016).

1.3.11 Ribosome biogenesis and the contribution of H/ACA snoRNPs

In the nucleolus, ribosomal RNA is transcribed as a massive precursor (pre-rRNA) that is comprised of the mature rRNA sequences (18S, 5.8S and 25S/28S rRNA) along with the internal (ITS1 and ITS2) and external (5' ETS and 3' ETS) transcribed spacers that have to be excised (called rRNA processing). Many trans-acting factors are involved in pre-rRNA processing and maturation including hundreds of H/ACA snoRNPs that catalyze isomerization of target uridines to pseudouridines. H/ACA snoRNPs are involved in the early stages of ribosome biogenesis in the nucleolus. During transcription or shortly after transcription, 18S rRNA can be processed from its precursor 20S pre-rRNA through cleavage at three distinct sites. One of the factors involved in this process belongs to the class of H/ACA snoRNPs. Specifically, one highly conserved and essential factor for cell viability is an RNA known as U17 (human) or snR30 (yeast) that belongs to H/ACA snoRNA class. This particular H/ACA snoRNA has no known target uridine in all ncRNAs including pre-rRNA, but it guides the H/ACA snoRNP complex to defined

sequences in the pre-35S rRNA (Morrissey and Tollervey 1993, Kiss, Jady et al. 2004). Although it is not clear how snR30 facilitates processing of pre-18SrRNA, it has been found that pre-18S rRNA fails to be processed at sites A0, A1 in the 5' ETS region and A2 sites in the ITS1 region in absence of snR30 or with snR30 mutated at the 3' hairpin that prevents direct base pairing to pre-rRNA. As snR30 is essential for pre-rRNA processing, any defect in this H/ACA snoRNA completely abolishes 18S rRNA accumulation (Fayet-Lebaron, Atzorn et al. 2009). SnR30 is one of the longest members of the H/ACA sRNA family as it contains about 600 nucleotides. It folds into a secondary structure that resembles the consensus 5'hairpin-hinge-hairpin-3'tail in H/ACA sRNA but with an additional middle hairpin (Atzorn, Fragapane et al. 2004). Further structural analysis of snR30 was conducted revealing that the most important regions that are needed for snR30 to fulfill its function are contained within the 3' hairpin (Atzorn, Fragapane et al. 2004).

Maturation of rRNA also involves numerous nucleotide modifications. Over 100 uridines in rRNA sequences are known to be the canonical targets for a number of H/ACA snoRNAs. The latter directly base pair with the target rRNA sequence which allows the H/ACA snoRNP complex to pseudouridylate a very distinct uridine (Ni, Tien et al. 1997, Kiss, Jady et al. 2004). About 8-10% of the total uridines in human 25S rRNA and 18S rRNA are isomerized to pseudouridines (Lestrade and Weber 2006). All these modifications in the target rRNAs are exclusively introduced by H/ACA snoRNPs. Although the target uridine is not necessarily conserved among eukaryotes, interestingly the general locations of pseudouridines in rRNA are conserved as pseudouridines are found in rRNA regions that are critical for ribosomal function. For example, the peptidyl

transferase center (PTC), the A, P and E sites as well as the A- site finger region, the decoding region, and also important helices that form inter-subunit bridges such as Helix 69 are rich in pseudouridines (Liu, Liang et al. 2008, Baudin-Baillieu, Fabret et al. 2009). Ribosome biogenesis is upregulated in cancer while on the other hand alterations in ribosome biogenesis give rise to inherited disorders (Montanaro, Treré et al. 2012, Ruggero and Shimamura 2014).

Bacterial rRNA has relatively few pseudouridines and their absence does not affect bacterial survival as all stand-alone, ribosomal pseudouridine synthases are non-essential in bacteria. However, the number of pseudouridines is drastically increased in highly complex organisms which may aid the ribosome to conduct a broader range of functions (Ruggero, Yoon et al. 2007, Piekna-Przybylska, Przybylski et al. 2008). For example, the Dyskeratosis congenita disorder is characterized by a reduction in pseudouridine levels in the rRNA which is not tolerated in other tested eukaryotic organisms including mouse (He, Navarrete et al. 2002). Ribosomes with reduced levels of pseudouridines are not able to translate a unique set of mRNAs with regulatory elements known as internal ribosome entry sites (IRES) (Blazewicz, Barnard et al.). These types of mRNAs are encoding for factors that initiate crucial cellular events such as apoptosis and tumor suppressors (Yoon, Peng et al. 2006, Jack, Bellodi et al. 2011, Penzo, Rocchi et al. 2015). The fact that the ribosome's functional ability is changed in response to pseudouridine formation makes it tempting to speculate that specific ribosomes could be dynamically involved in gene expression (Xue and Barna 2012). In this respect, it is noteworthy that many H/ACA snoRNAs, that have been involved in pre-rRNA pseudouridylation, are either up- or down-regulated in many types of cancer (Ronchetti, Todoerti et al. 2012, Mannoor, Shen

et al. 2014).

The Fournier lab examined how preventing pseudouridine formation in particular regions of rRNAs affects yeast cells. First, overall elimination of pseudouridine in yeast rRNA by making inactive Cbf5 is not well tolerated, as cells were growing only at 25 °C albeit very slowly (Zebarjadian, King et al. 1999). The group further eliminated pseudouridine formation at particular sites in the rRNA by deleting individual or groups of H/ACA snoRNAs, for example RNAs targeting helix 69 in 25S rRNA containing five modifications (three are in the loop and two are in the stem). When these pseudouridines are deleted in combination, the ribosome structure was altered and the translational read-through of stop codons increased (Liang, Liu et al. 2007).

In addition, examination of cells obtained from X-DC patients shows that a particular H/ACA sRNA (SNORA15) that directs pseudouridylation in 18S rRNA at position U1367 is expressed only at very low levels and thus the modification is also decreased at that position (Bellodi, McMahon et al. 2013). However, there is currently no clear explanation for how deregulation of few a subsets of H/ACA snoRNAs is contributing to the disease Dyskeratosis congenita. On the other hand, some H/ACA sRNAs that guide pseudouridylation in 18S rRNA such as H/ACA snoRNA42 are up-regulated in solid tumors (Bellodi, McMahon et al. 2013).

1.3.12 H/ACA sRNPs acting on messenger RNAs

Discovering that pseudouridines are also found in mRNA has shifted our perspective

about this modification (Carlile, Rojas-Duran et al. 2014, Lovejoy, Riordan et al. 2014, Schwartz, Bernstein et al. 2014). Using novel sequencing techniques described above, over 2000 pseudouridine sites were identified in human mRNA (Li, Zhu et al. 2015). Although some uridines in mRNA are isomerized by H/ACA sRNPs, the majority of pseudouridines seems to be introduced by standalone pseudouridine synthases (Schwartz, Bernstein et al. 2014). Many mRNAs in yeast have been identified that are pseudouridylated in normal conditions, and the number of pseudouridines further increases upon stress such as heat shock (Schwartz, Bernstein et al. 2014, Abidi, Sayari et al. 2016). So far the majority of the pseudouridines that have been detected in the mature mRNA are located within the translated sequences (open reading frame). Moreover, only a few pseudouridines are found in the 3' UTR and only in very rare cases in the 5' UTR (Lovejoy, Riordan et al. 2014). Artificial introduction of pseudouridines at the first position in a stop codon impairs termination of translation in both *in vivo* as well as *in vitro*, and these stop codons with pseudouridines are translated into serine and threonine for ψAA and ψAG, and phenylalanine and tyrosine for ψGA (Karijolich and Yu 2011). This finding suggests not only that pseudouridine can change the genetic code, but it could also regulate gene expression. The crystal structure of the ribosome shows the tRNA anticodon forming non-canonical interactions with ψAG in mRNA (Fernández, Ng et al. 2013). Further investigations are needed to delineate the consequences of pseudouridines in mRNA for gene expression and regulation thereof.

Chapter 2. Objectives

The exact effects of Dyskeratosis congenita mutations on ribosome biogenesis have not been identified yet. Interestingly, the symptoms of Dyskeratosis congenita are varied and the severity of the syndrome is somewhat dependent on the location of the mutation. Here, I wanted to characterize yeast strains that each have a single Dyskeratosis congenita mutation. Yeast is a model organism that is used extensively by many researchers to better understand for instance a common pathway or to investigate the function of a protein in eukaryotes as it is much simpler than higher eukaryotes and much easier to work with. First, I planned to generate six haploid *S. cerevisiae* yeast strains by taking advantage of the homologous recombination mechanism. Three of the selected mutations are located within the catalytic domain of Cbf5, two in the PUA domain and one in the unstructured C-terminal extension of Cbf5. Second, I aimed to assess the growth behavior to screen for phenotypes in these strains. Subsequently, I planned to examine Cbf5p expression levels among mutant strains to see if there is any difference indicating reduced protein stability compared to the wild type. Finally, it was my objective to investigate cell cycle progression in these strains.

Chapter 3 Materials and Methods

3.1 Site-directed mutagenesis

Overlapping primers for each mutation were designed and the annealing temperature for each set of primers was adjusted in each PCR reaction (Table 1). The PCR reaction (20 µL total) contained Q5 DNA polymerase (NEB) and its corresponding buffer, 25 mM $MgCl_2$, 0.4 mM dNTPs, 0.5 µM forward and reverse primers each, 10 ng/µL of DNA template (pUC19-ScCbf5-YIp5). The PCR program was the following, (1) initial denaturation at 98°C for 30 s, (2) denaturation at 98°C for 10 s, (3) annealing & extension at 72°C for 6 min, (4) final extension at 72°C for 6 min and finally hold at 4°C. Steps (2) and (3) were repeated for 30 cycles. The PCR product was digested with DpnI (1 U/µL, Fisher Scientific) overnight at 37°C. Samples were analyzed by 1% agarose gel electrophoresis, stained with ethidium bromide and visualized with UV light (260 nm).

Table 1: Overlapping primers that have been designed to implement the desired mutation together with the corresponding annealing temperatures.

cbf5 mutation	Forward primers (5` to 3`)	Reverse primers (5` to 3`)	Annealing temperature
S91G	GGTCACGGTGGTACATT GGATCCAAAAGTTACAG	CAATGTACCACCGTGAC CAGTCTTCTCACAAC	72°C
R128W	GTATTGTCTGGTTGCAT GATGCTTTGAAGGATG	CATGCAACCAGACAATA CACACATACTCCTTACC	71°C
S250R	GAACGTTACTTGAGATC CATTATTCAACC	GGATCTCAAGTAACGTT CATCTCTTGTATTGTCG	67°C
K284R	GGTGCAAGGTTGATGAT CCCTGGTTTATTGCG	GATCATCAACCTTGCAC CATAACATACTGCATTG AC	72°C
M320T	CGCACAAACGTCAACCG TTGACTTAGCTTCCTG	CGGTTGACGTTTGTGCG ATGGCAACG	72°C
G372E	GGATAAATACGAGCGTG TTAACGAAAACACCCCA G	CACGCTCGTATTTATCCA ATTTGCCGTCTGC	72°C

3.2 Chemical Transformation into *E. coli* DH5α competent cells.

1.8 μL (20-60 ng/μL) of the mutagenesis PCR product or of a ligation reaction was transformed into 20 μL of highly competent *E. coli* DH5α cells (New England Biolabs). After addition of the DNA to the cells, the mixture was incubated on ice for 30 min followed by heat shock at 42°C for 45 s. After incubating on ice for 1 min, 250 μL LB medium was added and the reaction was incubated with shaking at 37°C for 1 hour. The transformation mixture was then plated onto LB agar containing 100 mg/mL ampicillin followed by incubation of the plates for 12-16 hours at 37°C.

3.3 Preparation of Glycerol Stocks

Selected colonies were used to inoculate a 5 mL liquid LB-Amp culture which was incubated 37°C for 16 hours. 750-900 μL of the culture was used to make glycerol stocks by adding the culture to 500 μL autoclaved glycerol followed by vortexing and shock freezing in liquid nitrogen. Glycerol stocks were stored at -80°C.

3.4 Miniprep of plasmid DNA

3-4 mL of an overnight *E. co*li culture were used to plasmid using a Miniprep Kit from Biobasic (following the manufacturer's instructions for high copy plasmid extraction).

3.5 Sequence confirmation

Plasmids were sent to GENEWIZ for sequencing using primers that are located about 120 base pairs upstream or downstream of the mutation site. The resulting sequence was aligned to the wild-type sequence of CBF5 using the online sequence alignment tool

CLUSTALW to identify positive mutants. The sequence of the whole gene was examined to ensure the absence any other mutations.

3.6 Plasmid restriction, gel extraction and dephosphorylation

To isolated the mutated *cbf5* gene, the pUC19-ScCbf5-YIp5 plasmids (1.6 ng/μL) were digested for 4 hours with EcoRI and SalI (2 U/μL, Fisher Scientific) in Orange buffer at 37°C. YIp5 plasmid was digested using the same conditions. The digestion products were separated by 1% agarose gel electrophoresis. The bands that correspond to the length of *cbf5* were excised from the gel, and the DNA was purified using an EZ-10 Spin column DNA Gel Extraction kit from Biobasic according to the manufacturer's instructions. The linear YIp5 vector was then dephosphorylated with shrimp alkaline phosphatase (1 U/μL, Fisher Scientific) in SAP buffer for 1 hour at 37°C. Subsequently, the enzyme was deactivated by heating the reaction at 65°C for 20 min.

3.7 Sticky-end ligation

To ligate the isolated *cbf5* insert to the linear, dephosphorylated YIp5 vector, insert and vector were mixed at a molar ratio of 4:1 insert to vector with T4 DNA ligase (2 U/μl, Fisher Scientific) and its complement buffer in 20 μL total reaction volume. The reaction was incubated at 16°C for 24 hours. The next day, after deactivation of the ligase enzymes by heating at 65°C for 20 min, about 1.8 μL of the ligation product was transformed into *E. coli* DH5α competent cells as described above.

3.8 Restriction of YIp5-ScCbf5

We have induced a single cut in *ScCBF5* using BshTI for *cbf5* K284R, M320T and G372E or HpaI (KspAI) for *cbf5* S91G and R128W and BamHI for S250R. In brief, 1.5 ng/μL of plasmid were incubated with 2 U/μl restriction enzyme (Fisher Scientific) in a total volume of 20 μL for 60 min at 37°C. Subsequently, restrictions enzymes were heat inactivated by incubation for 20 min at 80°C for BashTI similarly for BamHI and at 65°C for HpaI.

3.9 Yeast Transformation

First, *S. cerevisiae* BY4741 cells were made competent as follows. This strain is a laboratory strain derivative of S288C with mating type MATa and genotype his3Δ1 leu2Δ0 met15Δ0 ura3Δ0. A single BY4741 yeast colony was selected for growth in 5 mL YEPD medium with 2% glucose for 5-7 hours at 30 °C and 200 rpm shaking. About $2.5x10^8$ cells of this culture were used to inoculate a 50 mL culture in YPDE + 2% glucose. Growth was monitored by counting using a hemacytometer, and once the culture reached the exponential phase at 2 x 10^6 cells/mL, the cells were harvested by centrifugation at 3000*g* for 5 min. Next, the cells were resuspended in 25 mL distilled water and washed twice by centrifugation. Then the cells were washed once with 10 mL of SORB buffer (100 mM LiOAc, 10 mM Tris-HCL pH 8.0, 1 mM EDTA/NaOH PH 8.0, 1 M sorbitol) and centrifuged gently. The cells were resuspended in 360 μ SORB and 40 μL DNA carrier (10 mg/mL Salmon sperm DNA). Finally, 50 μL aliquots of the competent yeast cells were prepared, shock frozen in liquid nitrogen and stored at -80°C.

The linearized YIp5-Sc*cbf5* plasmid (1-2 ng/μL) was mixed with 50 μL competent BY4741 cells, and 40% PEG solution was added and mixed via vortexing. The reaction was left at room temperature for 30 min, then a total of 10% DMSO was added, and the transformation mixture was subjected to heat shock at 42°C for 20 min. Subsequently, cells were collected through centrifugation and gently resuspended in 200 μL of S-broth. Finally, the transformation mixture was plated onto two Sc-ura+ 2% glucose plates. The plates were then incubated at 30 °C for 48 hours.

3.10 Preparation of *S. cerevisiae* glycerol stocks

Colonies are selected and grown in 5 mL Sc-ura+ glucose medium at 30°C. At the mid of exponential phase of about 0.4 - 0.7 OD_{600}, and 750 μL of the cultures was collected to make a glycerol stock by mixing with 250 μL autoclaved glycerol in a cryo-vial. The cell suspension was shock frozen in liquid nitrogen, and the glycerol stock was stored at -80°C. (Growth of the remaining culture was continued to grow over night to harvest cells for genomic DNA extraction.)

3.11 Genomic DNA extraction and PCR amplification of ScCbf5

Genomic DNA from yeasts cells was extracted by using Geneaid-Presto Mini gDNA Yeast Kit from FroggaBio Scientific Solution following the manufacturer's instructions. The gDNA that was extracted was visualized by 0.75% agarose gel electrophoresis. A total of 25 μL PCR reaction was prepared by adding 10-130 ng/μL of genomic DNA, 10x Pfu DNA polymerase buffer, 0.5 μM each of the forward primer (EcoR1-ScCbf5-up) and the reverse primer (YIp5 downstream SalI reverse), as well as 0.4 mM dNTPs. The

reaction was incubated in the thermocycler using the following program: initial temperature 95°C for 5 min, denaturation at 95°C for 30 s, annealing at 52°C for 30 s and synthesis at 72°C for 2.5 min, final extension at 72°C for 7 min and hold at 4°C. The PCR product was visualized via agarose gel electrophoresis. All PCR products representing successful amplification of *ScCbf5* $^{(+mutation)}$ were analyzed by Sanger sequencing using primers that detected the mutations as well as primers that span the whole *ScCbf5* gene to ensure that there are no other mutations. The primers that were used in the PCR reaction as well as for sequencing are listed in Table 2. The location of the primers is illustrated in Figure 5.

Table 2: Primers that were used in the PCR reaction to amplify *cbf5* or for sequencing.

Primer Name (abbreviation used in Fig. 5)	Sequence (5` to 3`)
EcoR1-ScCbf5-up (EcoR I Fw)	AATTCATGTCAAAGGAGGATTTCGTTAT
SalI-ScCbf5down(Sal I down)	GTCGACGGAATGCAATAATCGGCGATAT
ScCbf5 seq 1038 antisense (1308 rev)	GGACCCAAACCCCATCTTCTTGGG
ScCbf5 seq 562 antisense (562 rev)	GCATATAAGTACCAGCTTCACAGGAAGC
YIp5 downstream SalI reverse (YIp5-Sal I down)	CAGTCATAAGTGCGGCGACG
ScCbf5 seq sense 976 (976 fw)	GCTTCCTGTGATCATGGTGTTGTTGC
Yeast Cbf5 antisense HindIII (anti-Hind III rev)	AAGCTTTCATTTCTTAGATTTCTTAGATTTC
YIp5 integration Cbf5(C) reverse (YIP5 Cbf5(c) rev)	CCTCAGAATCACCGTCTTCAG
YIP5 integrated Amp forward	ATAATACCGCGCCACATAGC

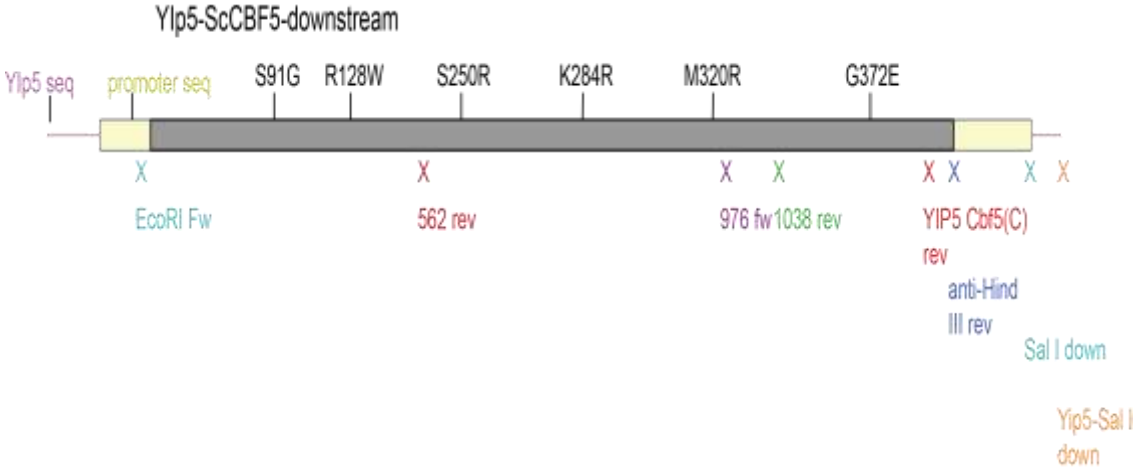


Figure 5: Location of primers used for PCR reaction or used for mutation confirmation. The *CBF5* gene is shown in gray and the promoter sequence in light yellow. The location of the mutations is depicted above the gene while the name of the primers is indicated underneath.

3.12 5-Fluoroorotic Acid (5-FOA) Pop-out procedure

SD+Ura+5-FOA plates were prepared by adding 136 mL S-broth, 0.288 g SC-ura, 2.4 g agar in a total volume of 150 mL distilled water. The mixture was then autoclaved, and while the mixture was still hot, 5.7 mM 5-FOA was added. Right before pouring into plates, 6 mL of 50% w/v glucose and 2 mg/mL uracil were added and mixed gently. Yeast strains were streaked on SD-ura plates, and one colony was selected and incubated in 5 mL YEPD liquid culture for about 7-8 h until the cells reached the exponential phase (0.4 OD_{600}) at 30°C with shaking at 200 rpm. The cells were collected via centrifugation at 3000 x*g* for 5 min, then the cells were resuspended in YEPD to $5x10^8$ cells/mL. Serial dilutions of $1x10^6$ and $1x10^7$ cells/mL were made to be plated on 5-FOA plates. The plates were incubated at 30°C for 48 h. Following the pop-out, genomic DNA was prepared of selected colonies as described above, the Cbf5 gene was amplified using the primers "EcoR1-ScCbf5-up" and "YIp5 downstream SalI reverse", and the PCR products were sequenced.

3.13 Growth analysis

To analyze growth in liquid medium, yeast strains were streaked onto YPDE + 2% glucose plates, and allowed to grow for 48 hours at 30°C. Three colonies from each strain were selected and grown in 5 mL YPDE + 2% glucose medium at 30°C and 200 rpm shaking for 5 hours. For each colony, three technical replicates were generated by preparing three dilutions (from each colony) of $1x10^6$ of cells/130 µL of YPDE + 2% glucose. Each dilution was placed into one well in a 96-well plate; as blank, YPDE+ 2% glucose was used. The microplate reader from BioTek with Gene5 software was used in this assay with shaking at 250 rpm and measurement of the OD_{600} every 15 min for 60 h.

This experiment was conducted at 30°C, 18°C and at 37°C. To analyze growth at 18°C, the plate reader was placed into a refrigerator at 14°C to make sure that the temperature will be stable at 18°C. The data is then exported as an Excel sheet and further analyzed in Excel, and the figures were made using the Prizm software. First, the average for the four technical replicates was calculated. Next, I calculated for each biological replicate (based the average of the technical replicates) the value $\ln(A_0/A_t)$ for each time point t where ln is the natural logarithm, A_0 is the OD_{600} reading at time zero (start time), and A_t is the OD_{600} at time t. Subsequently, $\ln(A_0/A_t)$ was plotted against time (in minutes), the linear region was selected and fit to a linear regression to obtain the slope y. The value of the slope y is used in the equation ln(2)/y to get the doubling time in minutes. Lastly, the average and the standard deviation of the doubling times of the different biological replicates was calculated.

To analyze growth on solid medium, cells were similarly grown in YPDE + 2% glucose until they reached the mid of the exponential phase. The cultures were used to prepare serial dilutions of $1x10^3$, 100, 50, 10 and 5 cells per 5 µL. The cells were placed in 5 µL drops onto YPDE + 2% glucose plates, left to dry, and inverted plates were incubated at 30°C or 18°C or 37°C for 48 - 72 hours. A picture of each was plate was taken with a smart phone camera or a Canon camera.

3.14 Whole-cell extract preparation

Whole-cell extract was prepared from 2 OD_{600} of cells from an exponentially growing culture. After centrifugation of the cell culture at 300 x*g* for 15 min, the cells were resuspended and incubated in 0.4 mM NaOH for 5 min. After a second centrifugation, the cells were resuspended in Tris-Urea (0.1 M Tris and 8 M Urea) pH 8 and 10% SDS,

boiled for 7 min, and 6x SDS loading dye was added before continuing to boil the samples for additional 3 min.

3.15 Western Blotting

To detect and semi-quantify Cbf5 protein levels, whole-cell extract is separated by 12% SDS-PAGE. The SDS-PAGE is subsequently blotted onto nitrocellulose at 100 V for 70 min in cold transfer buffer (48 mM Tris, 38.7 mM glycine, 10% (w/v) SDS, 20% (v/v) methanol). Subsequently, the nitrocellulose membrane is washed twice with 1x TBS buffer (50 mM Tris pH 7.6, 150 mM NaCl), and then incubated for 1 hour in blocking buffer (2% (w/v) BSA in 1x TBS). The primary antibody used in this assay was a custom-made antibody from BioBasic Inc.; it was generated by infecting two rabbits with an antigen of a 15-residue long, synthesized peptide designed from a sequence at the C terminus of Cbf5 (aa 461-475: KKEKKRKSEDGDSEE). A 1:1,000 dilution in 3% BSA in TBS of this primary antibody was made, and the blot was incubated with primary antibody overnight at 4°C. The next day, the membrane was washed one time with TBS buffer followed by three times wash with TBS+TT, then incubated with the secondary antibody (horseradish peroxidase coupled anti-rabbit antibody (Sigma), 1:1,000 dilution in 3% BSA and 1x TBS) for 1 to 2 hours, and finally washed again with TBS-TT three times.

For detection, the blot was incubated with 25 mL luminol solution (100 mM HCl pH 8.5, 11.3 mg luminol dissolved in 0.25 mL DMSO and 91.3 mM p-coumaric acid dissolved in DMSO) and 15 μL of H_2O_2 solution (40 μM H_2O_2 and 100 mM HCl) and developed for 5-8 min. A picture of the chemiluminescence was taken with an Amersham Imager 600. The membrane underwent mild stripping procedure by washing the membrane twice for

10 min with stripping buffer (160 mM glycine pH 2.2, 1 g SDS, 10 mL Tween 20 dissolved in 800 mL distilled water), followed by washing three times with PBS buffer, finally washing twice with TBS-TT for 5 min. The membrane is blocked for one hour with 3% PBS and then incubated with 1:1000 PGK1 primary antibody conjugated with horseradish peroxidase (Abcam) overnight.

3.16 Propidium iodide staining

5 mL yeast cultures starting at 0.3-0.4 OD_{600} in YEPD + 2% glucose were grown in the morning for 1-2 hours at 30°C to reach the early exponential growth phase before adding 15 μg/μL nocodazole (Calbiochem- EMD Millipore Corp). The culture was incubated for additional 2 hours. A second dose of nocodazole (7.5-15 μg/μL) was added, and the culture was incubated for another hour. 3μL of culture were observed in a microscope to ensure that more than 98% of cells were budding and had reached the G2/M phase. Cells were released from the arrest by centrifugation followed by washing twice with pre-warmed (30°C) water. Next, the cells were resuspended into new pre-warmed (30°C) 5mL YEPD + 2% glucose. Cell samples of 200 μL were collected before nocodazole treatment, immediately after nocodazole treatment, 15 min, 1 hour, 2 hours and 3 hours after restarting cell cycle by removing the nocodazole. The cells were collected via centrifugation, then washed twice with distilled water and resuspended in 4 mL 70 % cold ethanol and samples were stored at -20°C. The next day, the cells were centrifuged and washed twice with cold Dulbecco's phosphate buffer saline DPBS buffer (136.9 mM NaCl, 26.8 mM KCl, 8.1 mM Na_2HPO_4, 1.5 mM KH_2PO_4 in 1 L distilled water and adjusted to pH 7.4). Next, cells were treated with 50 mM NaCl and 50 μL of 10 μg/mL

RNase A for 30 min at 30°C. Cells were centrifuged again and washed with DPBS buffer. Subsequently, cells are incubated with 0.5 mL DPBS buffer and 50μg/mL propidium iodide for at least 10 min.

3.17 Flow cytometry

Propidium iodide-stained cells were analyzed in a Fluorescence Activated Cell Sorter (FACS) from BD Biosciences by exciting at 488 nm and detection at 610 ± 20 nm. In addition, the forward (FSC) and side scattering (SSC) was recorded. Cells were injected using a 70 μm nozzle, and for each sample 50,000 cells were measured. The data were analyzed using the Flowing Software to visualize the data on a dot plot (forward and side scattering) and also on a histogram (propidium iodide fluorescence). The dot plot data was used to select the populations of single yeast cells from cell debris and cell aggregates (see Appendix Fig. A2 for gating information). The gating was kept constant for all mutants and the wild-type for all time points.

Chapter 4. Results

4.1 Site-directed mutagenesis

Site-directed mutagenesis was used to create BY4741 mutant strains each containing a single mutation within the genomic *CBF5* gene. All of the mutations that have been selected are within regions that are conserved between yeast *CBF5* and human *DKC1*. Moreover, the mutations correspond to the ones that have been documented in the human disorder, Dyskeratosis congenita. The mutations encode amino acid substitutions within either the catalytic or PUA domains of *CBF5* or in the unstructured C-terminus (Figure 6). The ones that are located in the catalytic domain are S91G, R128W, and S250R which correspond to T121G, R158W and S280R in humans, respectively. The mutations that lie within the PUA domain are K284R, and M329T and the corresponding mutations in human are K214R, and M350T (Figure 6). Lastly, the mutation G372E resides in the unstructured C-terminal region of the Cbf5 protein and corresponds to G402E in humans. Most importantly, these mutations result in disease phenotypes ranging from severe to very severe. The majority of the mutations within the catalytic domain causea more severe disease than the ones within the tail region (Dokal 2000, Marrone, Walne et al. 2005, Dokal 2011). We hypothesized that these mutations are likely to have an effect on the *S. cerevisiae* model organism. In addition, patients with these mutations often develop different types of cancer as the disease progresses, a feature of Dyskeratosis congenita that we would like to study in the future.

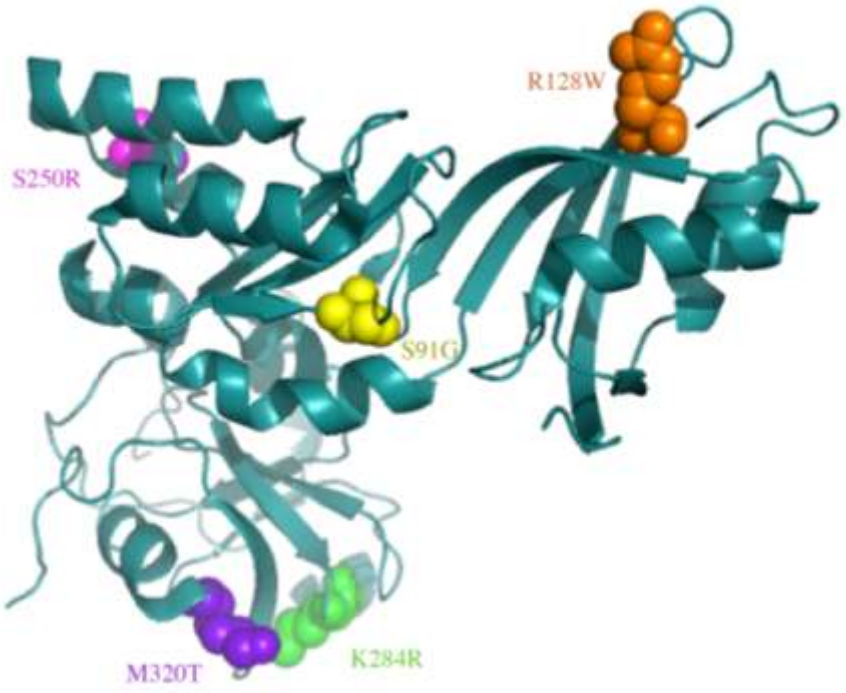


Figure 6: Location of Dyskeratosis congenita substitutions mapped onto the crystal structure of ScCbf5p (PDB:3U28). Within the catalytic domain are displayed in the following colors: S91G in yellow, R128W in orange and S250R in pink. The mutations in the PUA domain are shown as follows: K284R in green and M320T in purple. The last mutation G372E is within the unresolved C-terminal region.

There are two stages of generating yeast strains harboring mutations in the genomic *CBF5* gene: the first one is carried out in *E. coli*, while the second stage is carried out in yeast. Initially, this project started with site-directed mutagenesis of the pUC19-*ScCBF5+downstream* plasmid which contains the entire *CBF5* coding region flanked by 33 nt upstream and 348 nt downstream of the coding region. This was followed by sequence validation to assure the integrity of the *CBF5* gene and also successful mutagenesis (Figure 7). As shown in Figure 7, three samples were collected at different time points: 1) before starting the PCR reaction, 2) after the PCR and 3) after overnight DpnI digestion. Before the initiation of the PCR reaction, the plasmid migrates similarly to the 3500bp ladder band, even though pUC19-*ScCBF5*-downstream is 4100 bp in size;

this is because the plasmid is supercoiled and travels much faster than a linear DNA. After the PCR reaction, the successful mutagenesis product appears as a band much higher than the size of the supercoiled DNA template because the plasmid is linearized and not supercoiled and thus traveled corresponding to its length of 4100 bp. The control reaction does not contain DNA polymerase; therefore the after-PCR band corresponds to the original size of the DNA template. Unsuccessful mutagenesis is evident for the mutations S91G and G372E as no products appear on the agarose gel after the PCR (Figure 7); these mutagenesis reactions were later repeated. Next, PCR products are subjected to DpnI digestion to ensure that no parent DNA template is left in the reaction. DpnI cuts at the GATC sequence only when the A is methylated, and parental DNA (purified from *E. coli*) is methylated unlike the PCR-amplified DNA.

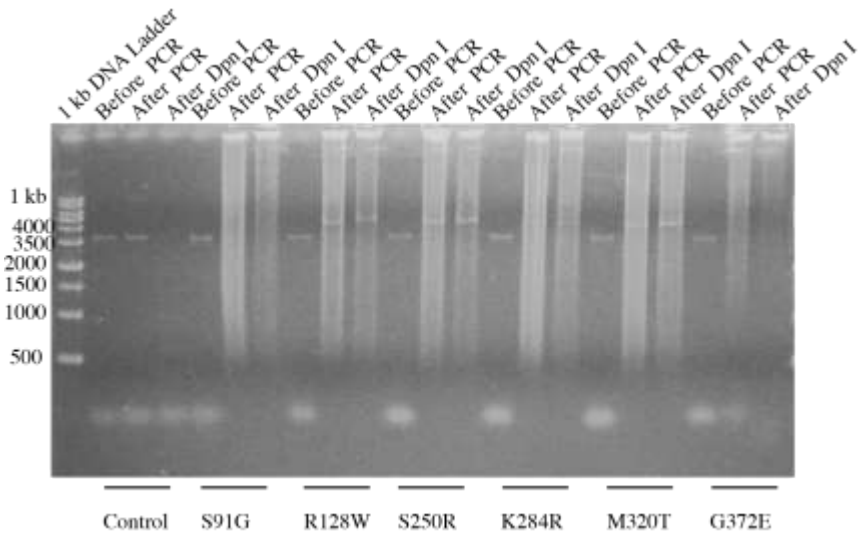


Figure 7: PCR reaction for site-directed mutagenesis of Cbf5 analyzed by agarose gel electrophoresis. Site-directed mutagenesis of pUC19-*ScCBF5*. Mutagenic primers were used in a PCR reaction to encode the amino acid substitutions in *CBF5* indicated at the bottom of the figure. A control PCR was carried out without DNA polymerase. After the PCR, the template plasmid was digested by DpnI. Samples before and after PCR as well as after the DpnI digestion (as indicated on the top) were analyzed by 1% agarose gel electrophoresis.

The presence of a band after the restriction with DpnI is an indication that the site-directed mutagenesis most likely is successful. After site-directed mutagenesis, samples were transformed into DH5α *E. coli* competent cells, followed by colony screening, plasmid extraction and sequencing.

Next, the *cbf5* gene was restricted from pUC19-*Sccbf5* and ligated into YIp5 (yeast integrating plasmid). YIp5 allows for the initiation of the second stage of the experiment, namely the transition from *E. coli* to yeast. First, *ScCbf5*$^{(+mutation)}$ was extracted from pUC19 by double restriction and then ligated to YIp5 that was also subjected to the same double restriction. The ligation reaction was transformed into *E. coli* DH5α competent cells, after which several colonies were selected, the plasmid DNA was isolated, and then plasmids underwent double restriction with EcoRI and SalI (Figure 8).

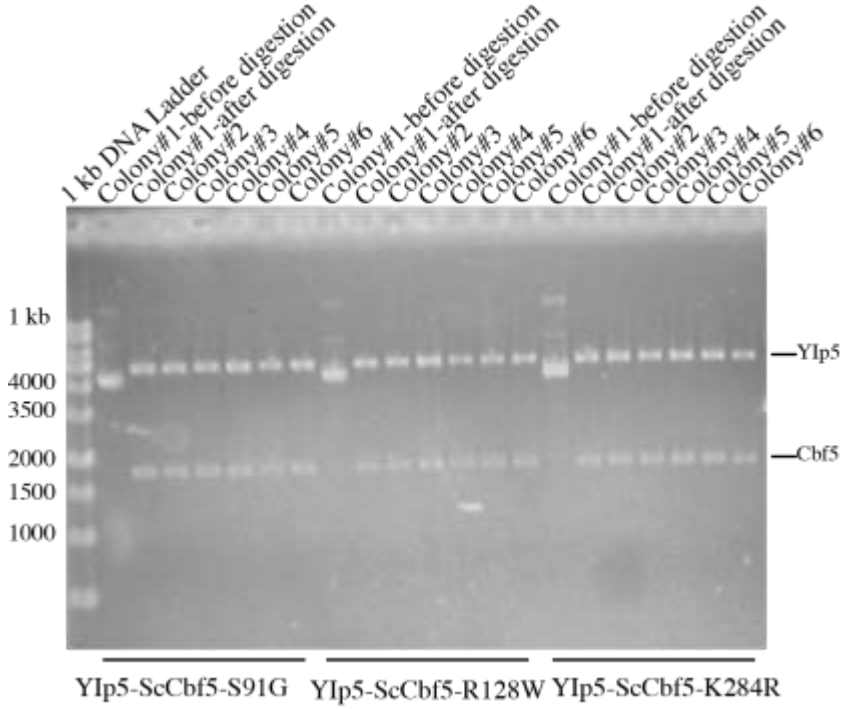


Figure 8: Restriction screening for the successful ligation of representative mutant *Sccbf5* genes into YIp5. After isolating plasmid DNA, the plasmids were double-restricted to release the *Sccbf5* genes from YIp5. Then the restriction products were analyzed by agarose gel electrophoresis. For each *cbf5* mutation (indicated underneath the gel), several colonies were analyzed as shown above the gel.

The restriction products are separated 1% agarose gel which was stained with ethidium bromide and visualized at 260nm. The gel shows clearly that all plasmids that have been isolated contain the *cbf5* gene and the vector YIp5 (Figure 8). Next, a few of them are further selected and verified by sequencing.

A single double-strand cut was introduced into the *cbf5* gene within the YIp5-*Sccbf5+downstream* plasmid before transforming it into BY4741 yeast competent cells. Restriction enzymes that are used for this step are indicated in the method section. The location of the cut site depends on the location of the mutation. For example, for the mutations in the N-terminal region *cbf5* such as S91G, the cut is induced downstream from the site of the mutation. For the mutations that are located at the C terminus such as G372E, a cut was introduced upstream from that location. This ensures that several hundred nt are available on both sites of the cut to pair with the yeast chromosome during homologous recombination. Usually we select a distance of at least 130 nt between the mutation and the restriction site to minimize the likelihood that the mutation is reversed during homologous recombination. The transformation mixture was plated on SD-Ura and grown for 48 hours. At this point, the well-known DNA repair mechanism of homologous recombination was exploited. When the yeast cell encounters a damaged *cbf5* gene, it activates a repair mechanism that joins homologous sequences from native *CBF5* to the *cbf5* gene within the plasmid. Thus, the entire plasmid sequence becomes incorporated into the yeast genome (Figure 10). The *cbf5* sequences were examined by PCR using the genomic DNA as template and primers that target the *cbf5* gene (Figure 9). I used in each PCR either primer set 1) (YIP5 integrated Amp forward & yeast Cbf5antiHindIII) which will generate a band ~2300 bp in size, or primer set 2) (YIP5 integrated Amp forward & YIp5 integration Cbf5(c)reverse) which will produce a band

~2200 bp in size or primer set 3) (ScCbf5 EcoRI up & (YIp5 integration Cbf5(c)reverse) which will result in a band ~1500 bp in size. The primer set used typically depends on the location of the mutation. For example, PCR with either primer set 1) or 2) are used in order to get good sequencing results for mutations near the N terminus (mutations in the catalytic domain) whereas for the mutation at the PUA domain (near the C terminus) primer set 3) was used (or sometimes primer set 4) (ScCbf5 EcoRI up yeast Cbf5antiHindIII) and primer set 5) (ScCbf5 EcoRI up SalI-ScCbf5 downstream) which generates a band at less than 1800 bp).

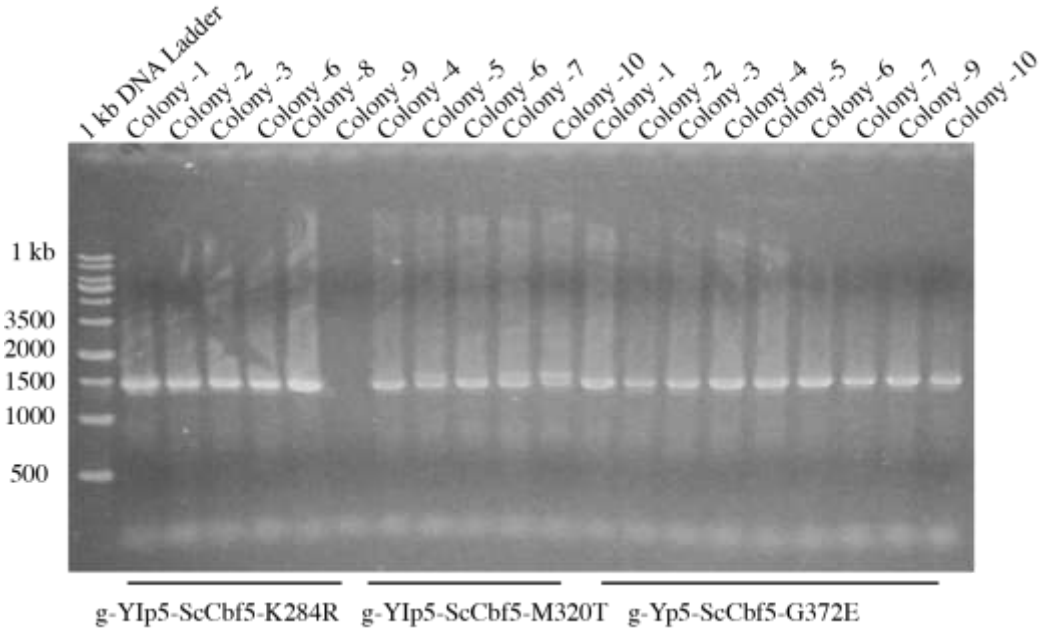


Figure 9: PCR screening of yeast strains after integration of YIp5-Sc*cbf5* plasmids. Genomic DNA was prepared from yeast cells and used as a template in a PCR reaction to amplify the *cbf5* gene. The PCR products were analyzed on a 1% agarose gel with the type of mutation indicated below and the colony number above the gel. Subsequently, PCR products were sequenced for confirmation of the mutation. The primers used in this reaction (ScCbf5 EcoRI up & YIp5 integration Cbf5(c) reverse) are indicated with blue dots in Figure 10.

PCR products that corresponded to the size of *cbf5* were sent for sequencing for validation. Due to the presence of two *cbf5* genes at this stage which were both amplified

by the used primers, sequencing results were of poor quality and mutation sites typically exhibited two nucleotide peaks in the sequencing chromatogram.

Naturally, the yeast cells containing an integrated plasmid may rarely eliminate this sequence by undergoing a second homologous recombination step. Cells that have undergone this step are differentiated from cells that have not by plating yeast cells on SD+Ura+5FOA. All cells that did not eject the plasmid sequence from its genome will still have the URA3 gene in its genome. The URA3 gene encodes for an enzyme that can use 5FOA as precursor to make 5-fluorouracil which is lethal to the cell. In contrast, cells that already underwent the pop out will not be harmed by 5FOA and will survive by consuming the uracil that has been added to the medium. This process is known as 5FOA Pop-out, which is described in the schematic below (Figure 10).

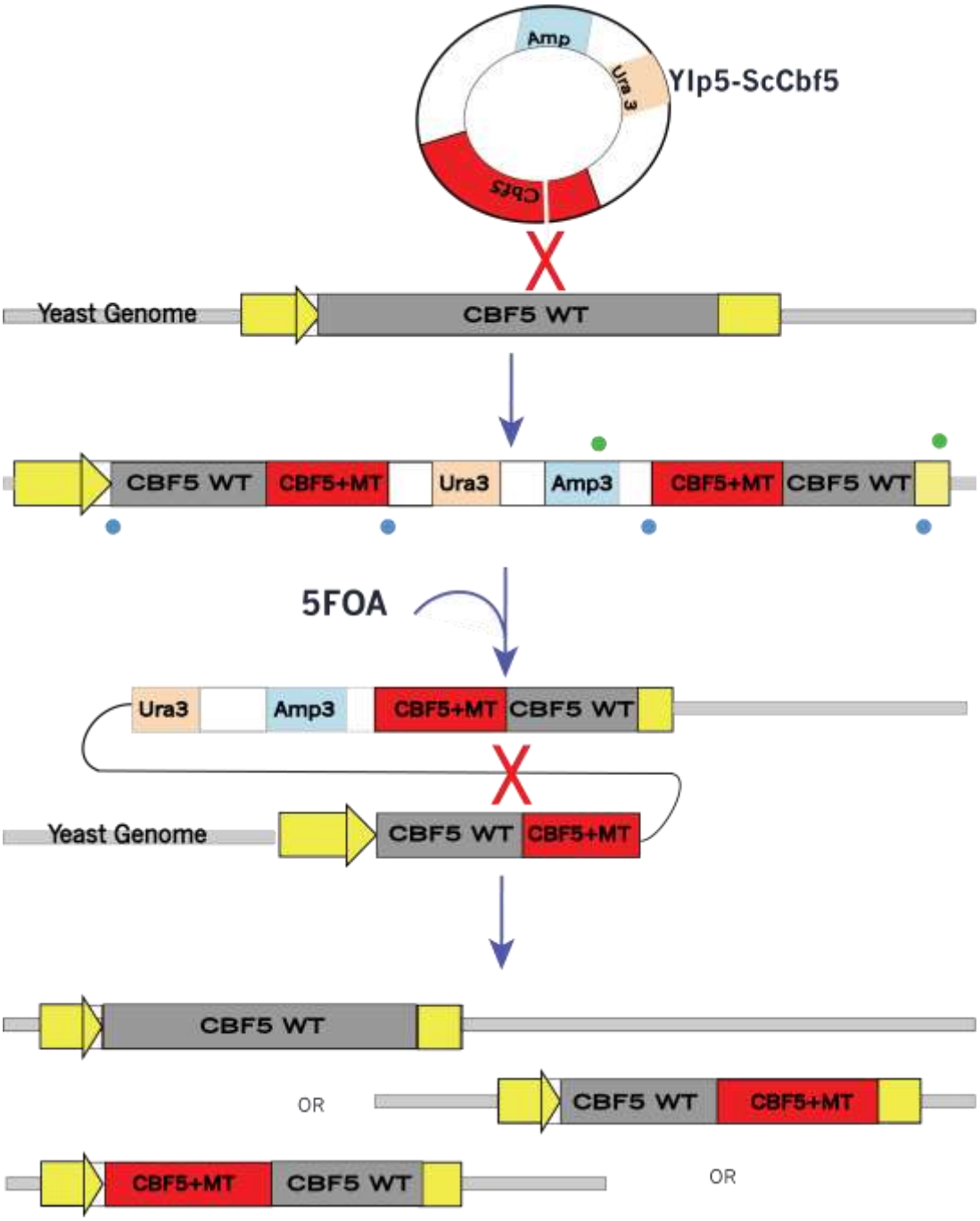


Figure 10: Schematic overview of generating mutations in the chromosomal *CBF5* gene. Upon transformation of the linearized YIp5-Sc*cbf5* plasmid into *S. cerevisiae*, the first homologous recombination leads to full integration of the YIp5-*Sccbf5* plasmid into the yeast genome. Subsequently, we selected cells that underwent the second homologous recombination (pop-out) by plating the strains on SD+Ura+5FOA. The green and blue dots are primers that were used to amplify *cbf5*. The primers that correspond to the green dots are (YIp5 integrated Amp forward & YIp5 integration Cbf5(c)reverse) whereas the primers that correspond to the blue dots are (ScCbf5 EcoRI up & YIp5 integration Cbf5(c) reverse).

The SD+Ura+5FOA plates always exhibited four types of colonies: large colonies, medium-sized colonies, relatively small and extremely small colonies. Ten colonies from each of the first three categories were selected and genomic DNA extraction was performed. A PCR reaction was conducted using primers are indicated in (Figure 11) that target *CBF5,* and the PCR product was visualized on a 1% agarose gel (Figure 11). Next, the PCR products were sent for sequencing for verification. Usually, all the sequencing results of PCRs performed on large colonies yielded a wild-type sequence. However, colonies of medium sized or small colonies almost always contained the mutated *cbf5* gene. Moreover, the sequencing results showed only one peak in the sequencing chromatogram at the site of the mutation.

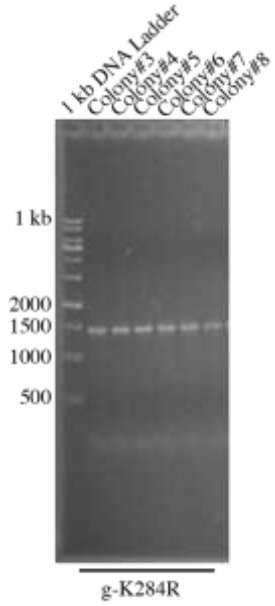


Figure 11: PCR screening of yeast strains potentially containing mutant *cbf5* alleles after the popout of YIp5 vector sequences. Yeast genomic DNA was used as a template in a PCR reaction to amplify the *cbf5* gene including parts of the promoter sequence; the PCR reactions were analyzed by 1% agarose gel electrophoresis. As an example, the screening of the *cbf5* K284R strain is shown here; the different colonies tested are indicated above the gel. Primers used in this reaction are ScCbf5 EcoRI up & YIp5 integration Cbf5(c)reverse.

4.2 Phenotypic screening

Growth analysis studies are one of the most widely used experiments for phenotype detection in yeast. Therefore, a 96-well plate and plate reader system was used to examine yeast cell growth at 18°C, 30°C and 37°C. Three selected colonies were each incubated in 5 mL YEPD culture. Once the cells reached the middle of the exponential growth phase (0.4 OD_{600}), the cells were diluted to 0.1 OD_{600}, placed in a well of the plate and incubated with shaking for 60-72 h. At 30°C, the doubling time for the wild type cells was about 100 min. The doubling time for each mutant was then compared to the wild type to examine any differences in growth rate. None of the mutants grew slower than the wild type cells given the standard deviations. Therefore, it was concluded that there are no notable differences between the wild type and the mutant strains regarding their growth at 30°C in YEPD medium. This experiment was repeated at low and high temperatures to investigate the presence of any phenotypic features at suboptimal conditions. First the same experiment was repeated at 37°C for 72h.

Table 3: Summary of the doubling time for the WT and the *cbf5* mutant strains at 18°C, 30°C and 37°C. Cells were grown in liquid rich medium for 30 -72 h. The growth test was conducted using a plate reader and a 96-well plate.

	18°C		30°C		37°C	
Muta-tion	Doubling time, min	Standard deviation, min	Doubling time, min	Standard deviation, min	Doubling time, min	Standard deviation, min
WT	229	11	110	7	92	5
S91G	245	5	87	3	92	6
R128W	241	10	108	12	92	1
S250R	288	6	90	1	75	7
K284R	247	9	98	6	92	7
M320T	267	5	91	3	75	8
G372E	241	5	105	3	95	1

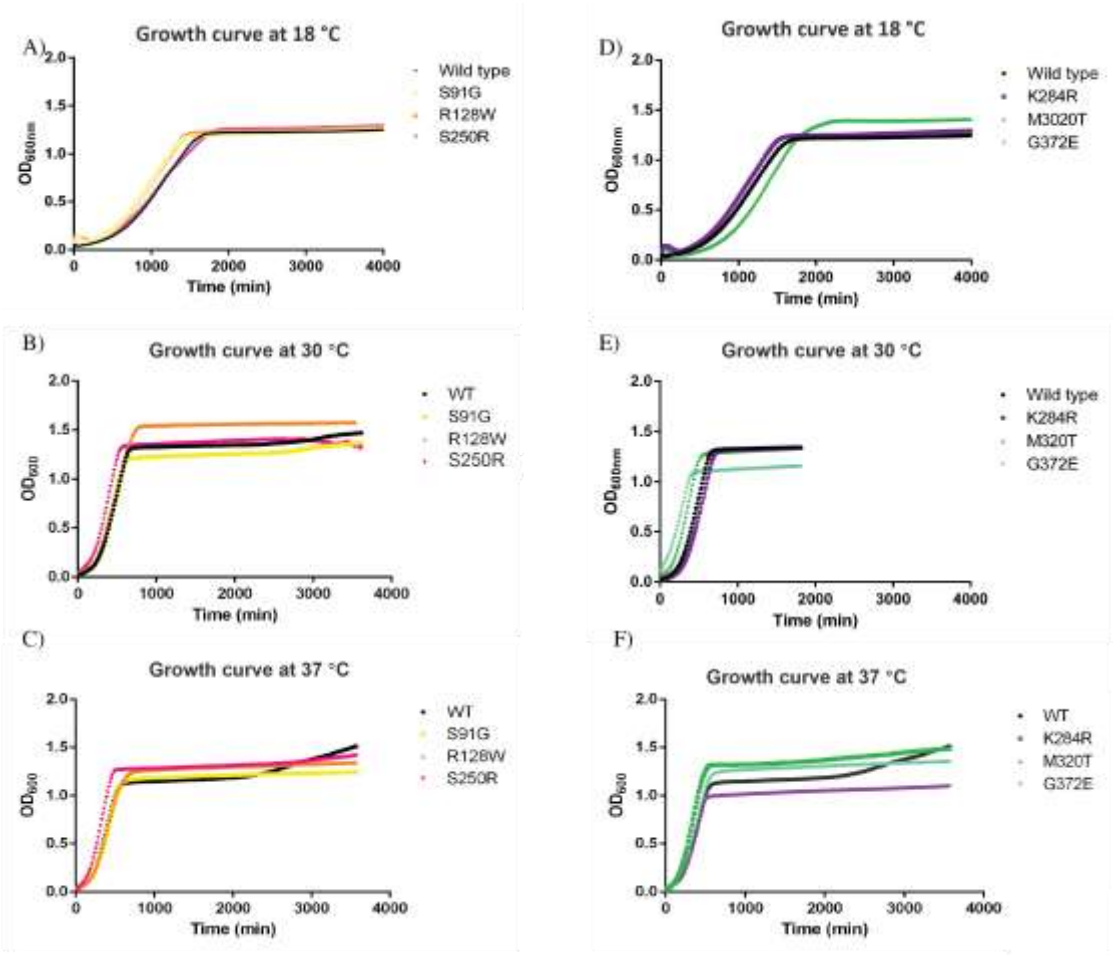


Figure 12: Growth curves of the wild type and the mutant *cbf5* strains at different temperatures in liquid rich medium. The graphs on the left are for the wild type, *cbf5* S91G, R128W and S250R at 18°C in the top, at 30°C in the middle and the 37°C in the bottom. Similarly, on the right are the graphs for the growth rate for the wild type, *cbf5* K284R, M320T and G372E.

The data show that the wild type cells maintain a similar growth rate at 37°C to cells grown at 30°C. All mutant strains were compared to the wild type cells and as was observed at 30°C, all mutants grew at least as fast as the wild type cells and displayed temperature resilience. Some growth curves will need to be reproduced to assess whether some mutations enable a faster growth than wildtype (S250R, K284R). Lastly, we examined the growth rate at 18°C, and for most mutants, no detectable phenotype was observed (Figure 12 & Table 3). However, the mutants S250R and M320T displayed a slight decrease in doubling time suggesting a mild cold sensitivity for these strains.

For the growth analysis on a solid YEPD medium, serial of dilutions are made and plated on YEPD + agar medium. The plates are left to incubate at 30°C, 37°C and 18°C respectively for 48-72 h. The results indicate that all strains were able to grow in a similar manner as the wild type at 30°C. Moreover, the growth behavior did not change at 37°C and 18°C and and remained similar to the wild type (Figure 13).

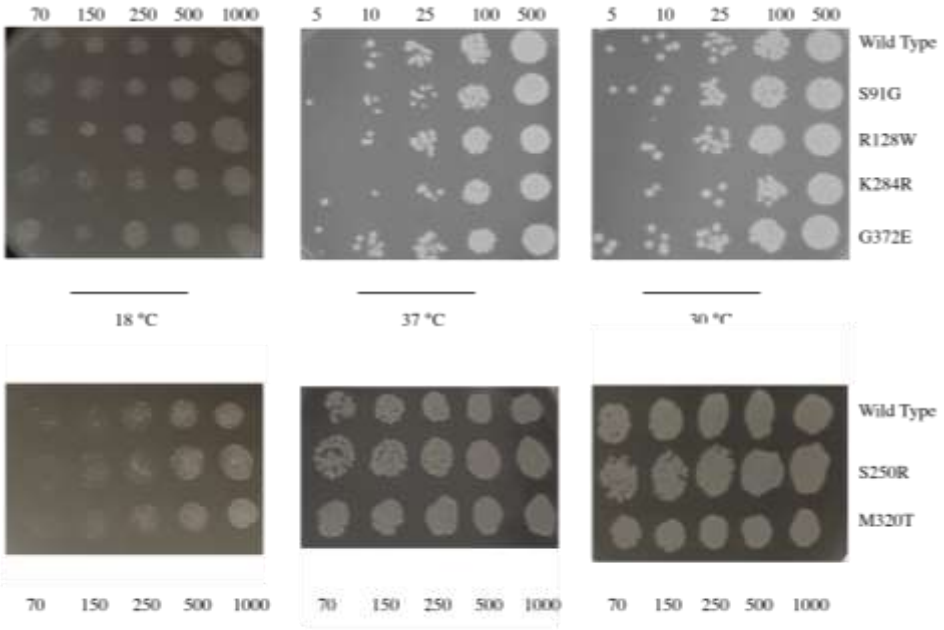


Figure 13: Growth analysis on solid rich medium at different temperatures. The different yeast strains are labelled on the right whereas the number of cells is described either above or below each plate. The temperature of the test is indicated in the middle.

4.3 Quantification of Cbf5 protein levels

Western blotting was used to examine ScCbf5 protein levels in whole-cell extracts of the mutant yeast strains and to compare these to the wild type Cbf5 protein levels. Yeast cells were grown to the mid-point of the exponential growth phase, and a 2 OD_{600} sample was collected. The samples were subjected to alkaline extraction, and total protein from each

yeast strain (wild-type and mutants) was then separated by SDS-PAGE (Figure 14). The proteins were transferred from the SDS gel to a nitrocellulose membrane for Western blotting (Figure 15).

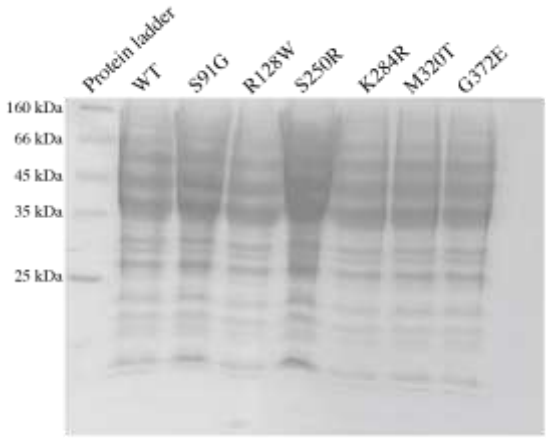


Figure 14: SDS-PAGE of yeast whole-cell extract from wild-type and mutant yeast strains. Whole cell extract was prepared using the alkaline extraction method. Cbf5 is approximately 55 kDa in size.

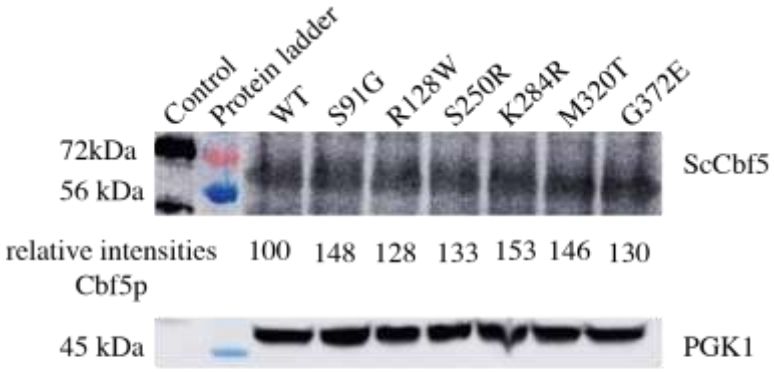


Figure 15: Western blot of Cbf5 from wild-type and mutant yeast strains. For Western blotting, a primary antibody against a small peptide chain in the C-terminus of Cbf5 was generated. The membrane was stripped and incubated with PGK1 primary antibody to show PGK1 as a loading control. Relative intensity of Cbf5p bands is indicated beneath each band. As a control (left lane), ScCbf5 protein was loaded that was expressed in *E. coli* and purified using a hexahistidine tag; as this recombinant Cbf5 protein may not be phosphorylated in contrast to endogenous Cbf5 in yeast, it may migrate a bit slower on the gel.

Several bands were detected including one that corresponded in size to the molecular weight of Cbf5. The membrane was stripped and then incubated with primary antibody to detected phosphoglycerate kinase (PGK1) as a loading control.

As shown in Figure 15, Cbf5 is visible in all samples. To quantify Cbf5 levels in each of the samples Image J software was used to determine band intensity and to normalize the Cbf5 level relative to the loading control PGK1. This analysis revealed that all Cbf5 variants are expressed at levels comparable to wild-type Cbf5. It will be necessary to reproduce this experiment to determine standard deviations for the relative Cbf5 levels in the mutant strains.

4.4 Examination of the cell cycle in yeast strains

The following experiment was used to investigate any unusual delay of cell cycles phases in the mutant strains. Towards this aim, yeast cells were stained with propidium iodide after synchronization with nocodazole and subsequently analyzed by flow cytometry to detect the propidium iodide-stained DNA levels. To examine the distribution between G1 and G2 phases, cells were collected at specific time points after removal of nocodazole. The first sample was collected from cell culture before the addition of nocodazole. Figure 16 shows that the G1 and G2 phases are clearly represented as two independent peaks in propidium iodide fluorescence corresponding to 1*n* and 2*n* DNA content, respectively. The G2 peak is larger than the G1 peak. Next, nocodazole was added to the cultures which were allowed to incubate until 95% of the cells reached the G2/M phase. This was evident under the microscope where all cells exhibited the presence of buds, with daughter cells similar in size to the parent cells. At this particular point, cells were collected for the flow cytometry analysis. Figure 16 shows that all strains, including the

wildtype strain, appear to be stalled at the G2 phase while the G1 phase is represented only by a small peak. After a 3-hour incubation with nocodazole, the cells were pelleted and washed twice with distilled water and then re-incubated with fresh pre-warmed medium. By doing this, the cells were able to resume the cell cycle in a synchronized fashion. Samples after 15-20 min of releasing from the cell cycle are collected: at this time point, the G1 peak is getting larger again indicating that cell cycle is progressing and that the cells stalled in G2 previously have no divided and reached the G1 phase. After 2 h of incubation, cells are also collected, and the data shows that the cell population, that is in G2 phase, is much greater than the cell population in G1 phase. This indicates that the cells, that had just divided at the previous time point (15-20 min) and were in G1 phase have no duplicated their genome reaching the G2 phase. After 4 h, the results show that the G1 peak has increased again slightly, and after 6 h, both a pronounced G1 peak and a clear G2 peak can be observed suggesting that cells are progressing through the cell cycle and that the population is becoming asynchronous again. This trend is consistent across all strains, and no apparent abnormality in the mutant strains compared to the wild-type is observed.

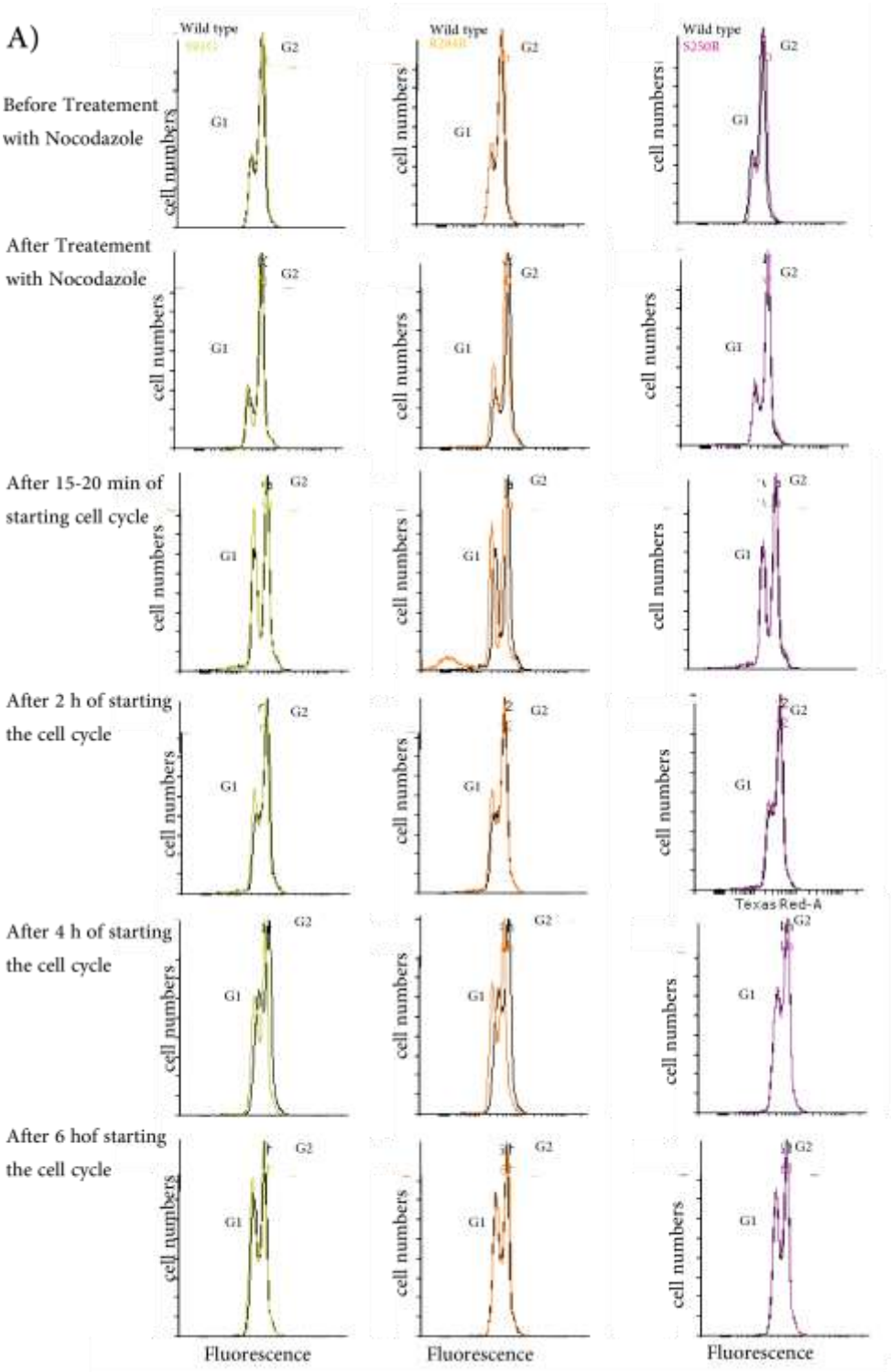


Figure 16 A: Cell cycle analysis by flow cytometry of wild-type and mutant yeast strains. The cell cycle was arrested in the G2 phase by nocodazole and synchronized progression through the cell cycle was monitored after removal of nocodazole. The x-axis shows the fluorescence of propidium iodide. The data set for the wild-type is shown in each panel in black for direct comparison (legend is continued on next page).

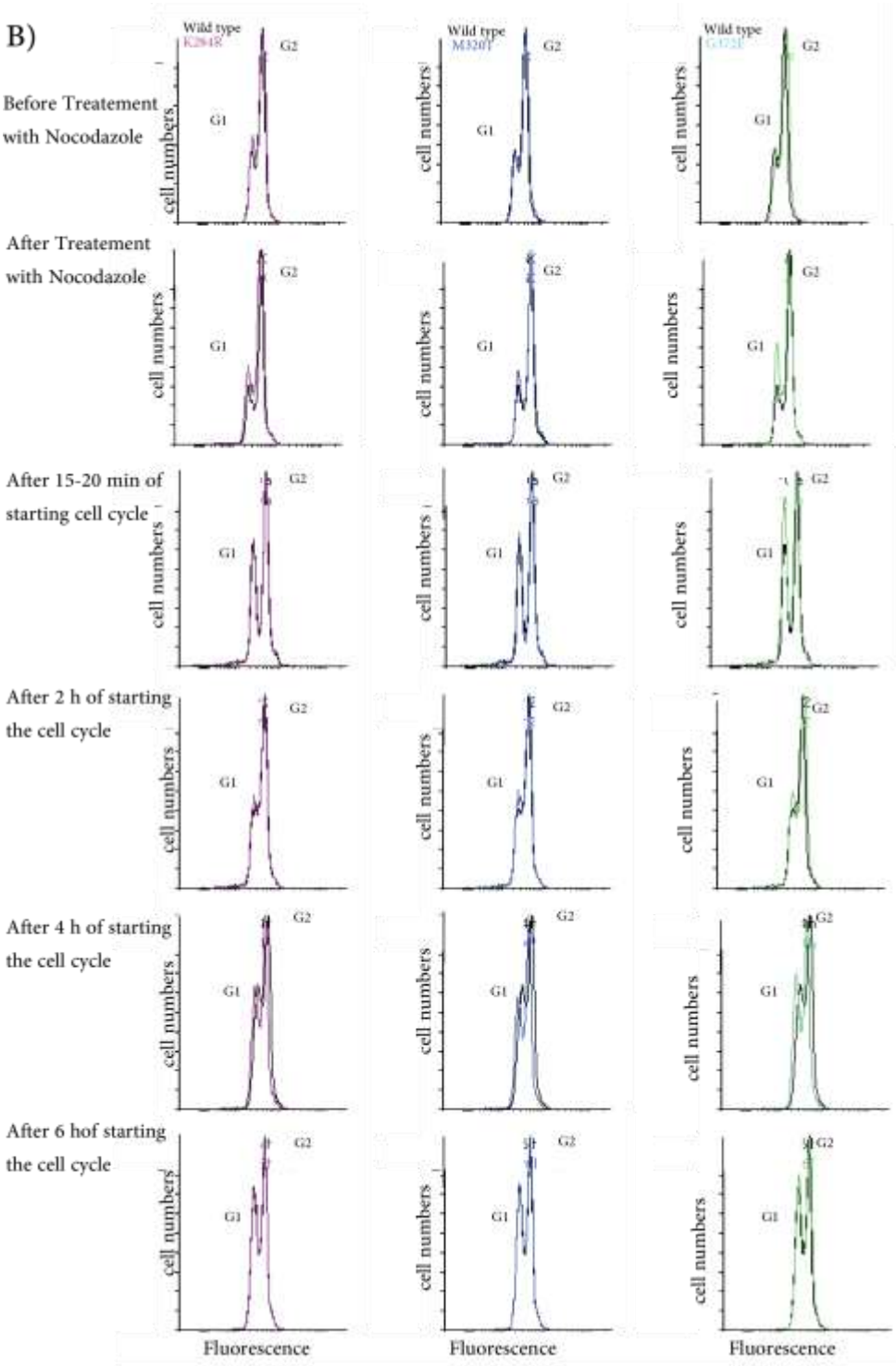


Figure 16 B: continued. Yeast cells were stained with propidium iodide for flow cytometry analysis. Panel A) shows histograms at different time points for the wild-type (in black), *cbf5* S91G (green), *cbf5* R128W (orange) and *cbf5* S250R (purple) yeast strains. Panel B) shows flow cytometry histograms at different time points for the WT overlaid in black with *cbf5* K284R (purple), *cbf5* M320T (blue) and *cbf5* G372E (green) yeast strains.

Chapter 5. Discussion

Here, I constructed six yeast strains which each have a single mutation in the *cbf5* gene and screened the strains for phenotypes. All strains grow similarly to the wild type at the optimal temperature (30°C) in rich medium; similar results were also observed at 18°C and 37°C. In rich solid medium, all mutants also exhibit similar growth behavior as the wild type. The next assay was to examine Cbf5p expression levels across strains via western blotting. This analysis reveals that Cbf5 expression levels are relatively similar among the mutant strains and the wild type strain. Finally, all strains were characterized for progression through the cell cycle following nocodazole treatment using flow cytometry; no difference between the wild type and the constructed strains was observed. The overall conclusion of this study is that single Dyskeratosis congenita mutations in *cbf5* are highly tolerated in yeast even though the counterpart mutations in humans cause severe symptoms. In the following, I will speculate how these results can be explained in the context of current literature on H/ACA sRNPs.

Firstly, the protein constituents of H/ACA sRNPs are bound to less than 100 different H/ACA RNAs in yeast (Lestrade and Weber 2006). The population of H/ACA RNAs in humans is significantly greater than in yeast (Dieci, Preti et al. 2009, Ketele, Kiss et al. 2016). Furthermore, human H/ACA RNAs are the precursors for microRNAs and Piwi-associated RNAs (Scott, Avolio et al. 2009). This difference could explain why there is no obvious phenotype in the constructed yeast strains. Moreover, H/ACA RNA expression levels in human are tissue-dependent and stress-dependent. Deregulation of H/ACA RNA levels can be used as a biological marker for diseases such as many types of

cancer. Therefore, a great range of symptoms can be expressed in humans whereas in a single cell such as yeast, a phenotypic effect manifests typically as slower growth. Secondly, studies conducted on human cell lines with *dkc1* being mutated showed that overall translation fidelity of mRNA by ribosomes was relatively normal (Jack, Bellodi et al. 2011, Penzo, Rocchi et al. 2015). However, the ribosomes failed to translate certain type of mRNAs containing an IRES element. At least 30 IRES mRNAs were found in human tissues such as brain tissues, and some mRNAs that are involved in apoptosis contain an IRES structure as well (Hellen and Sarnow 2001, Penzo, Rocchi et al. 2015). In yeast, there are seven mRNAs containing an IRES element which are could play a role in invasive growth (Gilbert, Zhou et al. 2007). But not much is known about other IRES-containing mRNAs in yeast, and their importance for cellular growth. In addition, it has been shown that yeast cells with pseudouridine-free ribosomes are still viable, but display a severe slow-growth phenotype (Zebarjadian, King et al. 1999), whereas deregulation of one of the pseudouridine-guiding H/ACA RNAs is linked to cancer in humans (Taft, Pang et al. 2010, Esteller 2011). These findings could suggest that unlike yeast, the human biological system is more complex and highly sensitive to slight alterations.

Therefore, to see phenotypic effects in yeast strains, we may have to combine several Dyskeratosis congenita mutations in each yeast strain. In future, we need to be strategic while choosing mutations to be implemented. Based on the H/ACA sRNP crystal structure, there are few possible types of mutations in Cbf5 that we can investigate. One obvious strategy is to target the PUA domain, where we could focus on interfering with guide RNA binding. We could make mutations that make the binding to the guide RNA either stronger or weaker. The PUA domain is highly positively charged and thereby binds to the H/ACA RNA. We could design the PUA domain in in a way that makes it

almost completely positively charged which could possibly further increases the affinity to guide RNA. However, the danger of this strategy is that the PUA domain could become destabilized by a too high density of positive charges on the protein surface. One consequence of mutations that tighten binding of the PUA domain in Cbf5p to the guide RNA could be to make the H/ACA sRNP complex very rigid and thus eliminate intrinsic flexibility in the complex that is needed for the interaction with the substrate RNA. Alternatively, we can weaken the interaction of the PUA domain with guide RNA by introducing negatively charged residues. We could then test the effect of amino acid substitutions in Cbf5 *in vitro* by measuring pseudouridylation of target uridines in a variety of substrate RNAs using a tritium release assay and by determining the affinity for guide RNA and substrate RNA by nitrocellulose filtration (Kamalampeta and Kothe 2012). In addition, it has been speculated that H/ACA sRNPs could be functioning as RNA chaperones (Keffer-Wilkes, Veerareddygari et al. 2016).In this case, one feature that we can test is the rRNA profile to assess processing of rRNA from pre-rRNA by pulse-chase labeling of rRNA with radiolabeled [^{3}H]-methionine in exponentially growing cells. Eventually, pre-rRNA variants will be processed into fully mature rRNA, but it could be much slower than in the wild type. We could also subject ribosomes to sucrose gradient centrifugation to examine the polysome profile and the 40S to 60S ratio (Liang, Liu et al. 2007). In addition, we could examine pseudouridine content compared to the wild type.

Many transacting factors are involved in ribosome biogenesis, and in humans about 100 H/ACA sRNPs are crucial factors in pre-RNA maturation. These H/ACA sRNPs are involved in pre-rRNA biogenesis at different time points. Therefore, we could test the hierarchy of different H/ACA sRNPs interacting with pre-rRNA. We could do that by

integrating a MS2-tagged guide RNA into the yeast genome. Next we could target this MS2-H/ACA RNA with a fusion protein of MS2 protein and a red fluorescent protein. We can next similarly insert another RNA tag such as RNA mango to the 5`end of the pre-rRNA (Dolgosheina, Jeng et al. 2014). We could then detect co-localization of the guide RNA and the pre-rRNA by dual fluorescence microscopy. Thereby, we could observe the lifetime of H/ACA sRNPs pairing and dissociation from pre-rRNA with the help of the fluorescent signal. This could also be particularly interesting regarding the snR30 H/ACA RNA in yeast as we could study when it base-pairs with pre-rRNA (Lemay, Hossain et al. 2011). snR30 is involved in pre-rRNA maturation, but the full mechanism is not fully understood. Understanding the interaction of this particular guide RNA could add significant information to the process of ribosome biogenesis. This fluorescence microscopy assay could also be used to assess the effects of *cbf5* mutations on ribosome biogenesis *in vivo*.

Once we get a yeast strain with one or more Dyskeratosis congenita mutations that has a phenotypic effect, we could test whether the overall pseudouridylation level will decrease or increase in the cell. Few studies have shown that during stress, the pseudouridylation level increases in noncoding RNA as well as mRNA (Li, Zhu et al. 2015). Another study has shown that when PUS1 is knocked out, its target uridine in U2 snRNA still gets pseudouridylated by an H/ACA sRNP (Deryusheva and Gall 2017). We could test this hypothesis by HPLC analysis of total RNA and then compare the level of pseudouridines formation in mutant strains relative to the wild type.

Finally, I propose making more Dyskeratosis congenita mutations in the unstructured C-terminal tail of Cbf5 including a complete deletion of this part of Cbf5. This location has not been investigated in eukaryotes, and this approach could provide information about

the earlier stages of H/ACA sRNP biogenesis, and also about the final location of H/ACA snoRNPs in the nucleolus as the tail could contribute to the gel-like structure of the nucleolus (Thiry and Lafontaine 2005). It is not known if the tail region in Cbf5p gets phosphorylated right after translation or if it gets first covered by Shq1. Deletion of the C terminus could also allow us to discover new protein partners that interact with the complex via the unstructured tail. Interestingly, upon Shq1p binding, the tail region in Cbf5 becomes highly structured (Li, Duan, Li, Ma et al. 2011). Cbf5p is heavily modified post-translationally in the N terminus, the PUA domain and in the C terminus. Notably, Dkc1p in humans and also Dkc1p in rats are modified in the same regions as yeast Cbf5p (Robles, Humphrey et al. 2016) (Swaney, Beltrao et al. 2013).

The vast majority of these post-translation modifications are phosphorylations. This could imply that modification of Cbf5 regulates its activity in the cell which raises the question regarding the functional difference of modified Cbf5 and unmodified Cbf5. Can we classify unmodified Cbf5p as inactive while the modified one is active or vice versa? Could it be that the phosphorylated tail triggers a signal in the cell? Or could it activate undiscovered or uncharacterized functions of Cbf5 in the cell? These and more questions can be answered by making yeast strains where the phosphorylation sites of Cbf5 are mutated to aspartate to mimic phosphorylation, or alanine, or where the tail is deleted completely.

In conclusion, Dyskeratosis congenita seems to be caused by disrupted pathways in the cell that might be specific to humans or mammals, and impaired ribosome biogenesis could be only one of several affected pathways since the H/ACA sRNPs complex is involved in pre-rRNA processing as well as incorporation of pseudouridines in rRNAs. Thus, introducing more or other Dyskeratosis congenita mutations in a simple eukaryote

such as yeast could in the future reveal affected step(s) with respect to ribosome biogenesis.

References

Abidi, K., T. Sayari, M. Jallouli, Y. Hammi, R. G. Louzir and T. Gargah (2016). "Diffuse Mesangial Sclerosis in a Child With Dyskeratosis Congenita Leading to End-stage Renal Disease." Iranian Journal of Kidney Diseases **10**(6): 416-418

Alawi, F. and M. N. Lee (2007). "DKC1 is a direct and conserved transcriptional target of c-MYC." Biochemical and Biophysical Research Communications **362**(4): 893-898.

Ale-Agha, N., N. Dyballa-Rukes, S. Jakob, J. Altschmied and J. Haendeler (2014). "Cellular functions of the dual-targeted catalytic subunit of telomerase, telomerase reverse transcriptase--potential role in senescence and aging." Experimental gerontology **56**: 189-193.

Alfonzo, J. D. (2014). "Post-transcriptional modifications are very important after all." RNA Biology **11**(12): 1481-1482.

Alter, B. P., N. Giri, S. A. Savage and P. S. Rosenberg (2009). "Cancer in dyskeratosis congenita." Blood **113**(26): 6549-6557.

Angrisani, A., A. Angrisani, R. Vicidomini, M. Turano and M. Furia (2014). "Human dyskerin: beyond telomeres." Biological chemistry **395**(6): 593-610.

Antonicka, H., K. Choquet, Z. Y. Lin, A. C. Gingras, C. L. Kleinman and E. A. Shoubridge (2016). "A pseudouridine synthase module is essential for mitochondrial protein synthesis and cell viability." EMBO reports **18**(1): 28-38.

Armanios, M. (2009). "Syndromes of telomere shortening." Annual review of genomics and human genetics **10**(1): 45-61.

Arnez, J. G. and T. A. Steitz (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**(24): 7560-7567.

Atzorn, V., P. Fragapane and T. Kiss (2004). "U17/snR30 Is a Ubiquitous snoRNA with Two Conserved Sequence Motifs Essential for 18S rRNA Production." Molecular and Cellular Biology **24**(4): 1769-1778.

Baker, M. (2011). "Long noncoding RNAs: the search for function." Nature Methods **8**(5): 379-383.

Bakin, A. and J. Ofengand (1993). "Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: Analysis by the application of a new sequencing technique." Biochemistry **32**(37): 9754-9762.

Ballarino, M., M. Morlando, F. Pagano, A. Fatica and I. Bozzoni (2005). "The Cotranscriptional Assembly of snoRNPs Controls the Biosynthesis of H/ACA snoRNAs in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **25**(13): 5396-5403.

Baudin-Baillieu, A., C. Fabret, X.-h. Liang, D. Piekna-Przybylska, M. J. Fournier and J.-P. Rousset (2009). "Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy." Nucleic Acids Research **37**(22): 7665-7677.

Bellodi, C., M. McMahon, A. Contreras, D. Juliano, N. Kopmar, T. Nakamura, D. Maltby, A. Burlingame, S. A. Savage, A. Shimamura and D. Ruggero (2013). "H/ACA small RNA dysfunctions in disease reveal key roles for noncoding RNA modifications in hematopoietic stem cell differentiation." Cell Reports **3**(5): 1493-1502.

Bessler, M., D. B. Wilson and P. J. Mason (2010). "Dyskeratosis congenita." FEBS Letters **584**(17): 3831-3838.

Blazewicz, S. J., R. L. Barnard, R. A. Daly and M. K. Firestone (2013). "Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses." ISME J **7**(11): 2061-2068.

Bousquet-Antonelli, C., Y. Henry, P. G'Elugne J, M. Caizergues-Ferrer and T. Kiss (1997). "A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs." EMBO J **16**(15): 4770-4776.

Bykhovskaya, Y., K. Casas, E. Mengesha, A. Inbal and N. Fischel-Ghodsian (2004). "Missense Mutation in Pseudouridine Synthase 1 (PUS1) Causes Mitochondrial Myopathy and Sideroblastic Anemia (MLASA)." American Journal of Human Genetics **74**(6): 1303-1308.

Calado, R. T. and N. S. Young (2008). "Telomere maintenance and human bone marrow failure." Blood **111**(9): 4446-4455.

Carlile, T. M., M. F. Rojas-Duran, B. Zinshteyn, H. Shin, K. M. Bartoli and W. V. Gilbert (2014). "Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells." Nature **515**(7525): 143-146.

Castle, J. C., C. D. Armour, M. Löwer, D. Haynor, M. Biery, H. Bouzek, R. Chen, S. Jackson, J. M. Johnson, C. A. Rohl and C. K. Raymond (2010). "Digital genome-wide ncRNA expression, including SnoRNAs, across 11 human tissues using polyA-neutral amplification." PloS One **5**(7): e11779.

Cavaille, J., K. Buiting, M. Kiefmann, M. Lalande, C. I. Brannan, B. Horsthemke, J.-P. Bachellerie, J. Brosius and A. Huttenhofer (2000). "Identification of Brain-Specific and Imprinted Small Nucleolar RNA Genes Exhibiting an Unusual Genomic Organization." Proceedings of the National Academy of Sciences of the United States of America **97**(26): 14311-14316.

Cerrudo, C. S., P. D. Ghiringhelli and D. E. Gomez (2014). "Protein universe containing a PUA RNA‐binding domain." FEBS Journal **281**(1): 74-87.

Connor, J. M., D. Gatherer, F. C. Gray, L. A. Pirrit and N. A. Affara (1986). "Assignment of the gene for dyskeratosis congenita to Xq28." Human genetics **72**(4): 348-351.

Corey, D. R. (2009). "Telomeres and Telomerase: From Discovery to Clinical Trials." Chemistry & Biology **16**(12): 1219-1223.

Cunningham, J. T., M. V. Moreno, A. Lodi, S. M. Ronen and D. Ruggero (2014). "Protein and nucleotide biosynthesis are coupled through a single rate limiting enzyme, PRPS2, to drive cancer." Cell **157**(5): 1088-1103.

Darzacq, X., N. Kittur, S. Roy, Y. Shav-Tal, R. H. Singer and U. T. Meier (2006). "Stepwise RNP Assembly at the Site of H/ACA RNA Transcription in Human Cells." The Journal of Cell Biology **173**(2): 207-218.

Davis, D. R., C. A. Veltri and L. Nielsen (1998). "An RNA model system for investigation of pseudouridine stabilization of the codon-anticodon interaction in tRNALys, tRNAHis and tRNATyr." Journal ofBiomolecular Structure & Dynamics **15**(6): 1121-1132.

Deryusheva, S. and J. G. Gall (2017). "Dual nature of pseudouridylation in U2 snRNA: Pus1p-dependent and Pus1p-independent activities in yeasts and higher eukaryotes." RNA **23**(7): 1060-1067.

Dez, C., A. Henras, B. Faucon, D. Lafontaine, M. Caizergues-Ferrer and Y. Henry (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 Research **29**(3): 598-603.

Dieci, G., M. Preti and B. Montanini (2009). "Eukaryotic snoRNAs: A paradigm for gene expression flexibility." Genomics **94**(2): 83-88.

Dokal, I. (2000). "Dyskeratosis congenita in all its forms." British Journal of Haematology **110**(4): 768-779.

Dokal, I. (2011). "Dyskeratosis Congenita." ASH Education Program Book **2011**(1): 480-486.

Dolgosheina, E. V., S. C. Y. Jeng, S. S. S. Panchapakesan, R. Cojocaru, P. S. K. Chen, P. D. Wilson, N. Hawkins, P. A. Wiggins and P. J. Unrau (2014). "RNA Mango Aptamer-Fluorophore: A Bright, High-Affinity Complex for RNA Labeling and Tracking." ACS Chemical Biology **9**(10): 2412-2420.

Dunin-Horkawicz, S., A. Czerwoniec, M. J. Gajda, M. Feder, H. Grosjean and J. M. Bujnicki (2006). "MODOMICS: a database of RNA modification pathways." Nucleic Acids Research **34**(Database issue): D145-D149.

Egan, E. D. and K. Collins (2012). "An Enhanced H/ACA RNP Assembly Mechanism for Human Telomerase RNA." Molecular and Cellular Biology **32**(13): 2428-2439.

Ender, C., A. Krek, M. R. Friedländer, M. Beitzinger, L. Weinmann, W. Chen, S. Pfeffer, N. Rajewsky and G. Meister (2008). "A Human snoRNA with MicroRNA-Like Functions." Molecular Cell **32**(4): 519-528.

Esteller, M. (2011). "Non-coding RNAs in human disease." Nature reviews. Genetics **12**(12): 861-874.

Fatica, A. and I. Bozzoni (2014). "Long non-coding RNAs: new players in cell differentiation and development." Nature Reviews. Genetics **15**(1): 7-21.

Fatica, A., M. Dlakic and D. Tollervey (2002). "Naf1p is a box H/ACA snoRNP assembly factor." RNA **8**(12): 1502-1514.

Fayet-Lebaron, E., V. Atzorn, Y. Henry and T. Kiss (2009). "18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences." The EMBO Journal **28**(9): 1260-1270.

Felsher, D. W., A. Yetil and J. van Riggelen (2010). "MYC as a regulator of ribosome biogenesis and protein synthesis." Nature Reviews Cancer **10**(4): 301-309.

Fernandez-Vizarra, E., A. Berardinelli, L. Valente, V. Tiranti and M. Zeviani (2007). "Nonsense mutation in pseudouridylate synthase 1 (PUS1) in two brothers affected by myopathy, lactic acidosis and sideroblastic anaemia (MLASA)." Journal of Medical Genetics **44**(3): 173-180.

Fernández, I. S., C. L. Ng, A. C. Kelley, G. Wu, Y.-T. Yu and V. Ramakrishnan (2013). "Unusual base pairing during the decoding of a stop codon by the ribosome." Nature **500**(7460): 107-110.

Ferreira, H. J., H. Heyn, C. Moutinho and M. Esteller (2012). "CpG island hypermethylation-associated silencing of small nucleolar RNAs in human cancer." RNA Biology **9**(6): 881-890.

Ganot, P., M. Caizergues-Ferrer and T. Kiss (1997). "The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation." Genes & Development **11**(7): 941-956.

Gilbert, W. V., K. Zhou, T. K. Butler and J. A. Doudna (2007). "Cap-Independent Translation Is Required for Starvation-Induced Differentiation in Yeast." Science **317**(5842): 1224.

Goldberg, I. H. and M. Rabinowitz (1961). "The incorporation of 5-ribosyluracil triphosphate into RNA in nuclear extracts of mammalian cells." Biochemical and Biophysical Research Communications **6**(5): 394-398.

Grozdanov, P. N., S. Roy, N. Kittur and U. T. Meier (2009). "SHQ1 is required prior to NAF1 for assembly of H/ACA small nucleolar and telomerase RNPs." RNA **15**(6): 1188-1197.

Gu, B. W., J. M. Fan, M. Bessler and P. J. Mason (2011). "Accelerated hematopoietic stem cell aging in a mouse model of dyskeratosis congenita responds to antioxidant treatment." Aging Cell **10**(2): 338-348.

Gupta, R., P. Gurha, P. Chaurasia and A. Joardar (2007). "Differential Roles of Archaeal Box H/ACA Proteins in Guide RNA-Dependent and Independent Pseudouridine Formation." RNA Biology **4**(2): 101-109.

Harrington, K. M., I. A. Nazarenko, D. B. Dix, R. C. Thompson and O. C. Uhlenbeck (1993). "*In vitro* analysis of translational rate and accuracy with an unmodified tRNA." Biochemistry **32**(30): 7617-7622.

He, J., S. Navarrete, M. Jasinski, T. Vulliamy, I. Dokal, M. Bessler and P. J. Mason (2002). "Targeted disruption of Dkc1, the gene mutated in X-linked dyskeratosis congenita, causes embryonic lethality in mice." Oncogene **21**(50): 7740-7744.

Heiss, N. S., S. W. Knight, T. J. Vulliamy, S. M. Klauck, S. Wiemann, P. J. Mason, A. Poustka and I. Dokal (1998). "X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions." Nature Genetics **19**(1): 32-38.

Hellen, C. U. T. and P. Sarnow (2001). "Internal ribosome entry sites in eukaryotic mRNA molecules." Genes & Development **15**(13): 1593-1612.

Henras, A., C. Dez, J. Noaillac-Depeyre, Y. Henry and M. Caizergues-Ferrer (2001). "Accumulation of H/ACA snoRNPs depends on the integrity of the conserved central domain of the RNA-binding protein Nhp2p." Nucleic Acids Research **29**(13): 2733-2746.

Hoareau-Aveilla, C., M. Bonoli, M. Caizergues-Ferrer and Y. Henry (2006). "hNaf1 is required for accumulation of human box H/ACA snoRNPs, scaRNPs, and telomerase." RNA **12**(5): 832-840.

Hudson, G. A., R. J. Bloomingdale and B. M. Znosko (2013). "Thermodynamic contribution and nearest-neighbor parameters of pseudouridine-adenosine base pairs in oligoribonucleotides." RNA **19**(11): 1474-1482.

Islam, A., S. Rafiq, M. Kirwan, A. Walne, J. Cavenagh, T. Vulliamy and I. Dokal (2013). "Haematological recovery in dyskeratosis congenita patients treated with danazol." British Journal of Haematology **162**(6): 854-856.

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

Jády, B. E., A. Ketele and T. Kiss (2012). "Human intron-encoded Alu RNAs are processed and packaged into Wdr79-associated nucleoplasmic box H/ACA RNPs." Genes & Development **26**(17): 1897-1910.

Jaskelioff, M., F. L. Muller, J.-H. Paik, E. Thomas, S. Jiang, A. C. Adams, E. Sahin, M. Kost-Alimova, A. Protopopov, J. Cadiñanos, J. W. Horner, E. Maratos-Flier and R. A. DePinho (2011). "Telomerase reactivation reverses tissue degeneration in aged telomerase-deficient mice." Nature **469**(7328): 102-106.

Jiang, W., K. Middleton, H. J. Yoon, C. Fouquet and J. Carbon (1993). "An essential yeast protein, CBF5p, binds *in vitro* to centromeres and microtubules." Molecular and Cellular Biology **13**(8): 4884-4893.

Kamalampeta, R. and U. Kothe (2012). "Archaeal proteins Nop10 and Gar1 increase the catalytic activity of Cbf5 in pseudouridylating tRNA." Sci Rep **2**: 663 doi:10.1038/srep00663.

Karijolich, J. and Y.-T. Yu (2011). "Converting nonsense codons into sense codons by targeted pseudouridylation." Nature **474**(7351): 395-398.

Keffer-Wilkes, L. C., G. R. Veerareddygari and U. Kothe (2016). "RNA modification enzyme TruB is a tRNA chaperone." Proceedings of the National Academy of Sciences of the United States of America **113**(50): 14306-14311.

Ketele, A., T. Kiss and B. E. Jády (2016). "Human intron-encoded AluACA RNAs and telomerase RNA share a common element promoting RNA accumulation." RNA Biology **13**(12): 1274-1285.

Kirwan, M. and I. Dokal (2009). "Dyskeratosis congenita, stem cells and telomeres." Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease **1792**(4): 371-379.

Kiss, A. M., B. E. Jady, E. Bertrand and T. Kiss (2004). "Human box H/ACA pseudouridylation guide RNA machinery." Mol Cell Biol **24**(13): 5797-5807.

Kiss, T. (2001). "Small nucleolar RNA‐guided post‐transcriptional modification of cellular RNAs." The EMBO Journal **20**(14): 3617-3622.

Kiss, T. and W. Filipowicz (1995). "Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns." Genes & Development **9**(11): 1411-1424.

Kocak, H., B. J. Ballew, K. Bisht, R. Eggebeen, B. D. Hicks, S. Suman, A. O'Neil, N. Giri, I. Maillard, B. P. Alter, C. E. Keegan, J. Nandakumar, S. A. Savage, N. D. C. G. R. Laboratory and N. D. C. S. W. Group (2014). "Hoyeraal-Hreidarsson syndrome caused by a germline mutation in the TEL patch of the telomere protein TPP1." Genes & Development **28**(19): 2090-2102.

Koo, B.-K., C.-J. Park, C. F. Fernandez, N. Chim, Y. Ding, G. Chanfreau and J. Feigon (2011). "Structure of H/ACA RNP Protein Nhp2p Reveals Cis/ Trans Isomerization of a Conserved Proline at the RNA and Nop10 Binding Interface." Journal of Molecular Biology **411**(5): 927-942.

Lemay, V., A. Hossain, Y. N. Osheim, A. L. Beyer and F. Dragon (2011). "Identification of novel proteins associated with yeast snR30 small nucleolar RNA." Nucleic Acids Research **39**(22): 9659-9670.

Lermontova, I., V. Schubert, F. Börnke, J. Macas and I. Schubert (2007). "*Arabidopsis* CBF5 interacts with the H/ACA snoRNP assembly factor NAF1." Plant Molecular Biology **65**(5): 615-626.

Lestrade, L. and M. J. Weber (2006). "snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs." Nucleic Acids Research **34**(Database issue): D158-D162.

Li, L. and K. Ye (2006). "Crystal structure of an H/ACA box ribonucleoprotein particle." Nature **443**(7109): 302-307.

Li, S., J. Duan, D. Li, S. Ma and K. Ye (2011). "Structure of the Shq1–Cbf5–Nop10–Gar1 complex and implications for H/ACA RNP biogenesis and dyskeratosis congenita." The EMBO Journal **30**(24): 5010-5020.

Li, S., J. Q. Duan, D. D. Li, B. Yang, M. Q. Dong and K. Q. Ye (2011). "Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase." Genes & Development **25**(22): 2409-2421.

Li, X., S. Ma and C. Yi (2016). "Pseudouridine: the fifth RNA nucleotide with renewed interests." Current opinion in chemical biology **33**: 108-116.

Li, X. Y., P. Zhu, S. Q. Ma, J. H. Song, J. Y. Bai, F. F. Sun and C. Q. Yi (2015). "Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome." Nature Chemical Biology **11**(8): 592-U593.

Liang, X.-h., Q. Liu and M. J. Fournier (2007). "rRNA Modifications in an Intersubunit Bridge of the Ribosome Strongly Affect Both Ribosome Biogenesis and Activity." Mol Cell **28**(6): 965-977.

Liu, B., X.-h. Liang, D. Piekna-Przybylska, M. J. Fournier and Q. Liu (2008). "Mis-targeted methylation in rRNA can severely impair ribosome synthesis and activity." RNA Biology **5**(4): 249-254.

Lopez-Serra, P. and M. Esteller (2012). "DNA methylation-associated silencing of tumor-suppressor microRNAs in cancer." Oncogene **31**(13): 1609-1622.

Lovejoy, A. F., D. P. Riordan and P. O. Brown (2014). "Transcriptome-Wide Mapping of Pseudouridines: Pseudouridine Synthases Modify Specific mRNAs in *S. cerevisiae*." PLoS One **9**(10): e110799.

Mannoor, K., J. Shen, J. Liao, Z. Liu and F. Jiang (2014). "Small nucleolar RNA signatures of lung tumor-initiating cells." Molecular Cancer **13**(1): 104-104.

Marrone, A., A. Walne and I. Dokal (2005). "Dyskeratosis congenita: telomerase, telomeres and anticipation." Current Opinion in Genetics & Development **15**(3): 249-257.

Marz, M., A. R. Gruber, C. Höner zu Siederdissen, F. Amman, S. Badelt, S. Bartschat, S. H. Bernhart, W. Beyer, S. Kehr, R. Lorenz, A. Tanzer, D. Yusuf, H. Tafer, I. L. Hofacker, P. F. Stadler, v. Universitetets särskilda, f. Medicinska, H. Create, M. V. Onkologi och Patologi, Oncology, M. V. Pathology, V. Division, v. Särskilda, C. University Specialised, a. Special, M. Faculty of, L. Department of Clinical Sciences, u. Lunds, V. Sektion, L. Institutionen för kliniska vetenskaper and U. Lund (2011). "Animal snoRNAs and scaRNAs with exceptional structures." RNA Biology **8**(6): 938-946.

Mason, P. J. and M. Bessler (2011). "The genetics of dyskeratosis congenita." Cancer Genetics **204**(12): 635-645.

McGrath, J. A. (1999). "Dyskeratosis congenita: new clinical and molecular insights into ribosome function." The Lancet **353**(9160): 1204-1205.

McMahon, M., A. Contreras and D. Ruggero (2015). "Small RNAs with big implications: New insights into H/ACA snoRNA function and their role in human disease." Wiley Interdisciplinary Reviews. RNA **6**(2): 173-189.

McMahon, M., A. Contreras and D. Ruggero (2015). "Small RNAs with big implications: new insights into H/ACA snoRNA function and their role in human disease." Wiley Interdiscip Rev RNA **6**(2): 173-189.

Montanaro, L., D. Treré and M. Derenzini (2012). "Changes in ribosome biogenesis may induce cancer by down-regulating the cell tumor suppressor potential." Biochimica et Biophysica Acta (BBA) - Reviews on Cancer **1825**(1): 101-110.

Morrissey, J. P. and D. Tollervey (1993). "Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis." Molecular and Cellular Biology **13**(4): 2469-2477.

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

Patton, J. R., Y. Bykhovskaya, E. Mengesha, C. Bertolotto and N. Fischel-Ghodsian (2005). "Mitochondrial myopathy and sideroblastic anemia (MLASA): missense mutation in the pseudouridine synthase 1 (PUS1) gene is associated with the loss of tRNA pseudouridylation." J Biol Chem **280**(20): 19823-19828.

Penzo, M., L. Rocchi, S. Brugiere, D. Carnicelli, C. Onofrillo, Y. Coute, M. Brigotti and L. Montanaro (2015). "Human ribosomes from cells with reduced dyskerin levels are intrinsically altered in translation." FASEB JournaL **29**(8): 3472-3482.

Pérez‐Arellano, I., J. Gallego and J. Cervera (2007). "The PUA domain−a structural and functional overview." FEBS Journal **274**(19): 4972-4984.

Piekna-Przybylska, D., P. Przybylski, A. Baudin-Baillieu, J.-P. Rousset and M. J. Fournier (2008). "Ribosome Performance Is Enhanced by a Rich Cluster of Pseudouridines in the A-site Finger Region of the Large Subunit." Journal of Biological Chemistry **283**(38): 26026-26036.

Preumont, A., R. Rzem, D. Vertommen and E. Van Schaftingen (2010). "HDHD1, which is often deleted in X-linked ichthyosis, encodes a pseudouridine-5'-phosphatase." The Biochemical Journal **431**(2): 237-244.

Rashid, R., B. Liang, D. L. Baker, O. A. Youssef, Y. He, K. Phipps, R. M. Terns, M. P. Terns and H. Li (2006). "Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita." Mol Cell **21**(2): 249-260.

Richard, P., A. M. Kiss, X. Darzacq and T. Kiss (2006). "Cotranscriptional Recognition of Human Intronic Box H/ACA snoRNAs Occurs in a Splicing-Independent Manner." Molecular and Cellular Biology **26**(7): 2540-2549.

Rintala-Dempsey, A. C. and U. Kothe (2017). "Eukaryotic stand-alone pseudouridine synthases – RNA modifying enzymes and emerging regulators of gene expression?" RNA Biology doi: 10.1080/15476286.2016.1276150. [Epub ahead of print].

Robles, M. S., S. J. Humphrey and M. Mann (2016). "Phosphorylation Is a Central Mechanism for Circadian Control of Metabolism and Physiology." Cell Metabolism **25**(1): 118-127.

Ronchetti, D., L. Mosca, G. Cutrona, G. Tuana, M. Gentile, S. Fabris, L. Agnelli, G. Ciceri, S. Matis, C. Massucco, M. Colombo, D. Reverberi, A. G. Recchia, S. Bossio, M. Negrini, P. Tassone, F. Morabito, M. Ferrarini and A. Neri (2013). "Small nucleolar RNAs as new biomarkers in chronic lymphocytic leukemia." BMC Medical Genomics **6**(1): 6-27.

Ronchetti, D., K. Todoerti, G. Tuana, L. Agnelli, L. Mosca, M. Lionetti, S. Fabris, P. Colapietro, M. Miozzo, M. Ferrarini, P. Tassone and A. Neri (2012). "The expression pattern of small nucleolar and small Cajal body-specific RNAs characterizes distinct molecular subtypes of multiple myeloma." Blood Cancer Journal **2**: e96.

Roovers, M., C. Hale, C. Tricot, M. P. Terns, R. M. Terns, H. Grosjean and L. Droogmans (2006). "Formation of the conserved pseudouridine at position 55 in archaeal tRNA." Nucleic Acids Research **34**(15): 4293-4301.

Ruggero, D. and A. Shimamura (2014). "Marrow failure: a window into ribosome biology." Blood Journal **124**(18): 2784-2792.

Ruggero, D., A. Yoon, G. Peng, O. Zollo, R. Adamo, N. Haynes, W. Xu and E. Rego (2007). "130 INVITED Novel insight into the pathogenesis of dyskeratosis congenita: how defective ribosome activity can cause cancer and disease." EJC Supplements **5**(4): 35-36.

Ruggero, D. Silvia, G., F. Piazza, E. Rego and et al. (2003). "Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification." Science **299**(5604): 259-262.

Schubert, T., M. C. Pusch, S. Diermeier, V. Benes, E. Kremmer, A. Imhof and G. Längst (2012). "Df31 protein and snoRNAs maintain accessible higher-order structures of chromatin." Mol Cell **48**(3): 434-444.

Schwartz, S., D. A. Bernstein, M. R. Mumbach, M. Jovanovic, R. H. Herbst, B. X. Leon-Ricardo, J. M. Engreitz, M. Guttman, R. Satija, E. S. Lander, G. Fink and A. Regev (2014). "Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA." Cell **159**(1): 148-162.

Scott, M. S., F. Avolio, M. Ono, A. I. Lamond and G. J. Barton (2009). "Human miRNA precursors with box H/ACA snoRNA features." PLoS Computational Biology **5**(9): e1000507.

Shay, J. W. and W. E. Wright (2004). "Telomeres in dyskeratosis congenita." Nature Genetics **36**(5): 437-438.

Skreka, K., M. Zywicki, M. Karbiener, A. Hüttenhofer, M. Scheideler and M. Rederstorff (2012). "Expression Profiling of a Heterogeneous Population of ncRNAs Employing a Mixed DNA/LNA Microarray." Journal of Nucleic Acids **2012**: 1-10.

Sorrow, J. J. M. and J. M. Hitch (1963). "Dyskeratosis congenita. first report of its occurrence in a female and a review of the literature." Archives of Dermatology **88**: 340-347.

Spenkuch, F., Y. Motorin and M. Helm (2014). "Pseudouridine: still mysterious, but never a fake (uridine)!" RNA Biol **11**(12): 1540-1554.

Swaney, D. L., P. Beltrao, L. Starita, A. Guo, J. Rush, S. Fields, N. J. Krogan and J. Villen (2013). "Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation." Nat Meth **10**(7): 676-682.

Taft, R. J., E. A. Glazov, T. Lassmann, Y. Hayashizaki, P. Carninci and J. S. Mattick (2009). "Small RNAs derived from snoRNAs." RNA (New York, N.Y.) **15**(7): 1233-1240.

Taft, R. J., K. C. Pang, T. R. Mercer, M. Dinger and J. S. Mattick (2010). "Non-coding RNAs: regulators of disease." The Journal of Pathology **220**(2): 126-139.

Thiry, M. and D. L. J. Lafontaine (2005). "Birth of a nucleolus: the evolution of nucleolar compartments." Trends in Cell Biology **15**(4): 194-199.

Trahan, C., C. Martel and F. Dragon (2010). "Effects of dyskeratosis congenita mutations in dyskerin, NHP2 and NOP10 on assembly of H/ACA pre-RNPs." Hum Mol Genet **19**(5): 825-836.

Ueda, Y., R. T. Calado, A. Norberg, S. Kajigaya, G. Roos, E. Hellstrom-Lindberg, N. S. Young (2014). "A mutation in the H/ACA box of telomerase RNA component gene (TERC) in a young patient with myelodysplastic syndrome." BMC Medical Genetics **15**(68): 1-8.

Valleron, W., E. Laprevotte, E. Gautier, C. Quelen, C. Demur, E. Delabesse, X. Agirre, F. Prosper, T. Kiss and P. Brousset (2012). "Specific small nucleolar RNA expression profiles in acute leukemia." Leukemia **26**(9): 2052-2060.

Valleron, W., L. Ysebaert, L. Berquet, V. Fataccioli, C. Quelen, A. Martin, M. Parrens, L. Lamant, L. de Leval, C. Gisselbrecht, P. Gaulard and P. Brousset (2012). "Small nucleolar RNA expression profiling identifies potential prognostic markers in peripheral T-cell lymphoma." Blood **120**(19): 3997-4005.

Veerareddygari, G. R., S. K. Singh and E. G. Mueller (2016). "The Pseudouridine Synthases Proceed through a Glycal Intermediate." Journal of the American Chemical Society **138**(25): 7852-7855.

Vitali, P., H. Royo, H. Seitz, J.-P. Bachellerie, A. Hüttenhofer and J. Cavaillé (2003). "Identification of 13 novel human modification guide RNAs." Nucleic Acids Research **31**(22): 6543-6551.

Vulliamy, T., R. Beswick, M. Kirwan, A. Marrone, M. Digweed, A. Walne and I. Dokal (2008). "Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita." Proceedings of the National Academy of Sciences of the United States of America **105**(23): 8073-8078.

Walne, A. J., L. Collopy, S. Cardoso, A. Ellison, V. Plagnol, C. Albayrak, D. Albayrak, S. S. Kilic, T. Pat roglu, H. Akar, K. Godfrey, T. Carter, M. Marafie, A. Vora, M. Sundin, T. Vulliamy, H. Tummala and I. Dokal (2016). "Marked overlap of four genetic syndromes with dyskeratosis congenita confounds clinical diagnosis." Haematologica **101**(10): 1180-1189.

Walne, A. J. and I. Dokal (2008). "Dyskeratosis Congenita: A historical perspective." Mechanisms of Ageing and Development **129**(1–2): 48-59.

Wang, P., L. Yang, Y. Q. Gao and X. S. Zhao (2015). "Accurate placement of substrate RNA by Gar1 in H/ACA RNA-guided pseudouridylation." Nucleic Acids Res **43**(15): 7207-7216.

Wu, G., H. Adachi, J. Ge, D. Stephenson, C. C. Query and Y. T. Yu (2016). "Pseudouridines in U2 snRNA stimulate the ATPase activity of Prp5 during spliceosome assembly." The EMBO Journal **35**(6): 654-667.

Wu, G., A. T. Yu, A. Kantartzis and Y. T. Yu (2011). "Functions and mechanisms of spliceosomal small nuclear RNA pseudouridylation." Wiley Interdiscip Rev RNA **2**(4): 571-581.

Xue, S. and M. Barna (2012). "Specialized ribosomes: a new frontier in gene regulation and organismal biology." Nature reviews. Molecular Cell Biology **13**(6): 355-369.

Xue, S., B. Liang, M. P. Terns, H. Li and R. M. Terns (2007). "Substrate RNA positioning in the archaeal H/ACA ribonucleoprotein complex." Nature Structural & Molecular Biology **14**(12): 1189-1195.

Yarian, C. S., M. M. Basti, R. J. Cain, G. Ansari, R. H. Guenther, E. Sochacka, G. Czerwinska, A. Malkiewicz and P. F. Agris (1999). "Structural and functional roles of the N1- and N3-protons of psi at tRNA's position 39." Nucleic AcidsResearch **27**(17): 3543-3549.

Yoon, A., G. Peng, Y. Brandenburg, O. Zollo, W. Xu, E. Rego and D. Ruggero (2006). "Impaired Control of IRES-Mediated Translation in X-Linked Dyskeratosis Congenita." Science **312**(5775): 902-906.

Yu, A. T., J. Ge and Y. T. Yu (2011). "Pseudouridines in spliceosomal snRNAs." Protein Cell **2**(9): 712-725.

Zebarjadian, Y., T. King, M. J. Fournier, L. Clarke and J. Carbon (1999). "Point Mutations in Yeast CBF5 Can Abolish In Vivo Pseudouridylation of rRNA." Molecular and Cellular Biology **19**(11): 7461-7472.

Zebarjadian, Y., T. King, M. J. Fournier, L. Clarke and J. Carbon (1999). "Point mutations in yeast CBF5 can abolish in vivo pseudouridylation of rRNA." Mol Cell Biol **19**(11): 7461-7472.

Zhang, X.-O., Q.-F. Yin, H.-B. Wang, Y. Zhang, T. Chen, P. Zheng, X. Lu, L.-L. Chen and L. Yang (2014). "Species-specific alternative splicing leads to unique expression of sno-lncRNAs." BMC genomics **15**(1): 287-287.

Zhang, Y., C. Xu, D. Gu, M. Wu, B. Yan, Z. Xu, Y. Wang and H. Liu (2016). "H/ACA Box Small Nucleolar RNA 7A promotes the Self-Renewal of Human Umbilical Cord Mesenchymal Stem Cells: SNORA7A promotes uMSC self-renewal." Stem cells **35**(1): 222-235.

Zucchini, C., P. Strippoli, A. Biolchi, R. Solmi, L. Lenzi, P. D'Addabbo, P. Carinci and L. Valvassori (2003). "The human TruB family of pseudouridine synthase genes, including the Dyskeratosis Congenita 1 gene and the novel member TRUB1." InternationalJournal ofMolecular Medicine **11**(6): 697-704.

Appendices

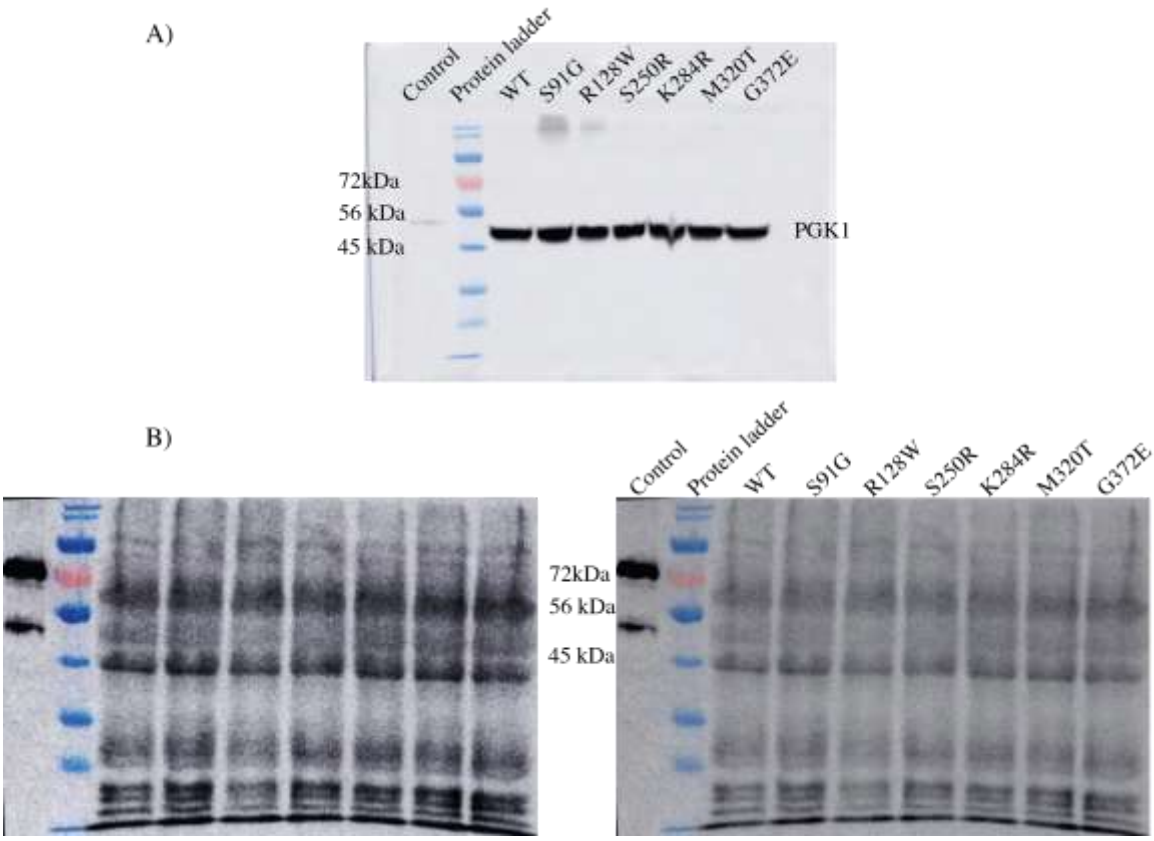
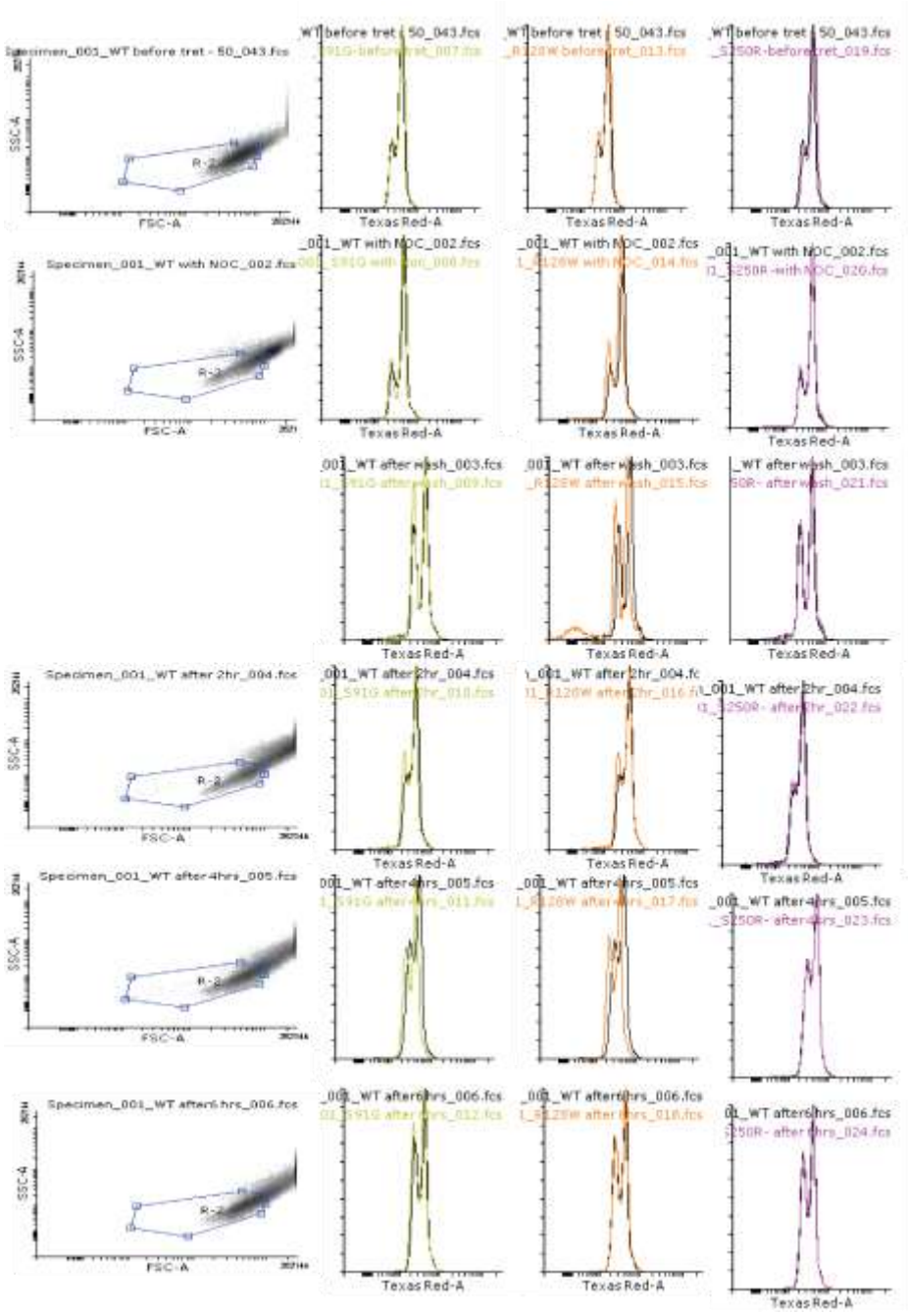


Figure A1: The full-size membranes of the western blot of Cbf5 from wild-type and mutant yeast strains. Panel A) shows the whole membrane with the PGK1 loading control. In panel B), the membrane is shown that was incubated with the primary anti-Cbf5 antibody, and the intensity was slightly adjusted using Irfan View software. On the right of panel B) is the same membrane without any intensity adjustment.


Specimen_001_WT before tret - 50_043.fcs
Specimen_001_WT with NOC_002.fcs
Specimen_001_WT after 2hr_004.fcs
Specimen_001_WT after4hrs_005.fcs
Specimen_001_WT after6 hrs_006.fcs
SSC-A
FSC-A
Texas Red-A

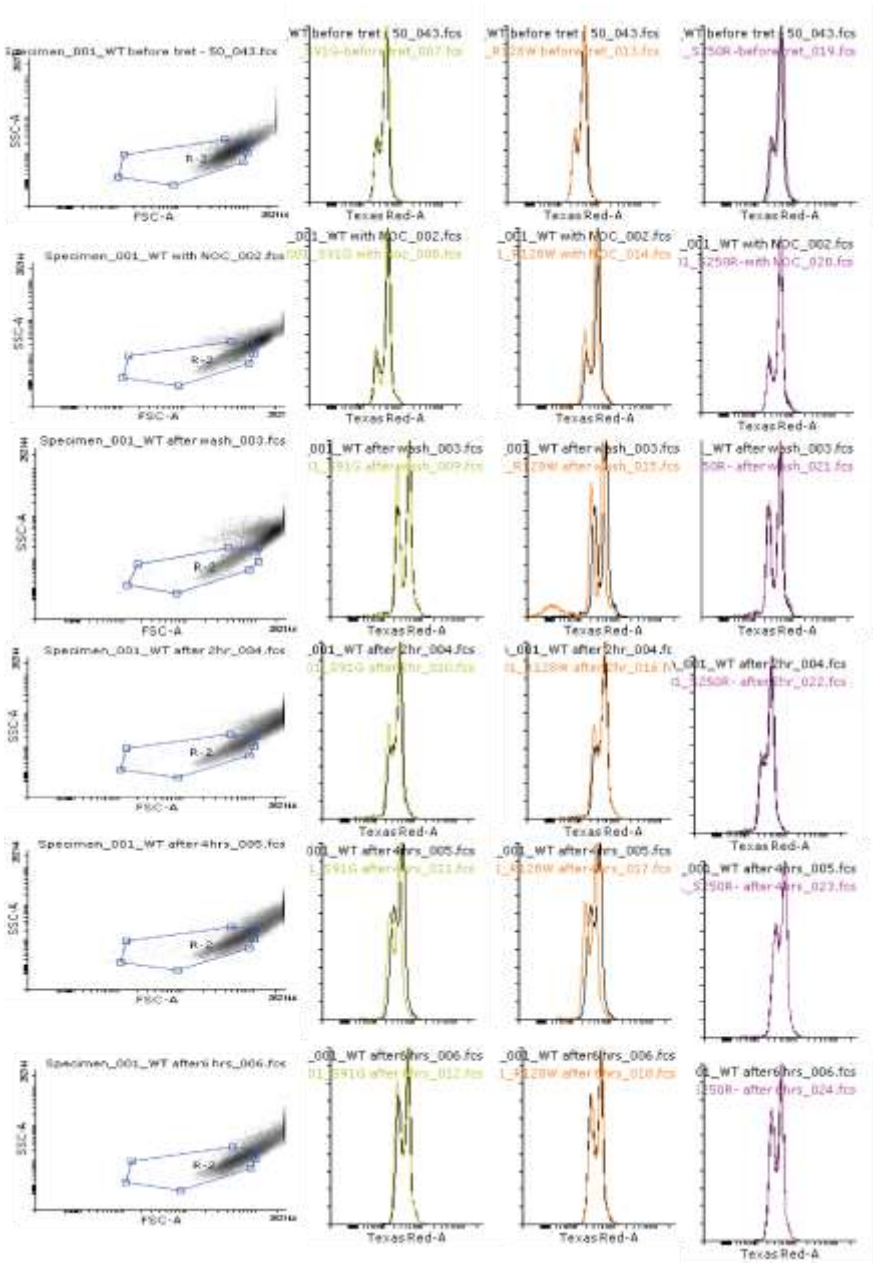


Figure A2: Flow cytometry data including gating information for the mutations in the catalytic domain of Cbf5p as well as the wild-type. The gating for the representative wild-type is shown in the left plots and the resulting histograms for wild-type and mutants are depicted on the right plots). The same gating was consistently used throughout the analysis for all data (time points and strains).

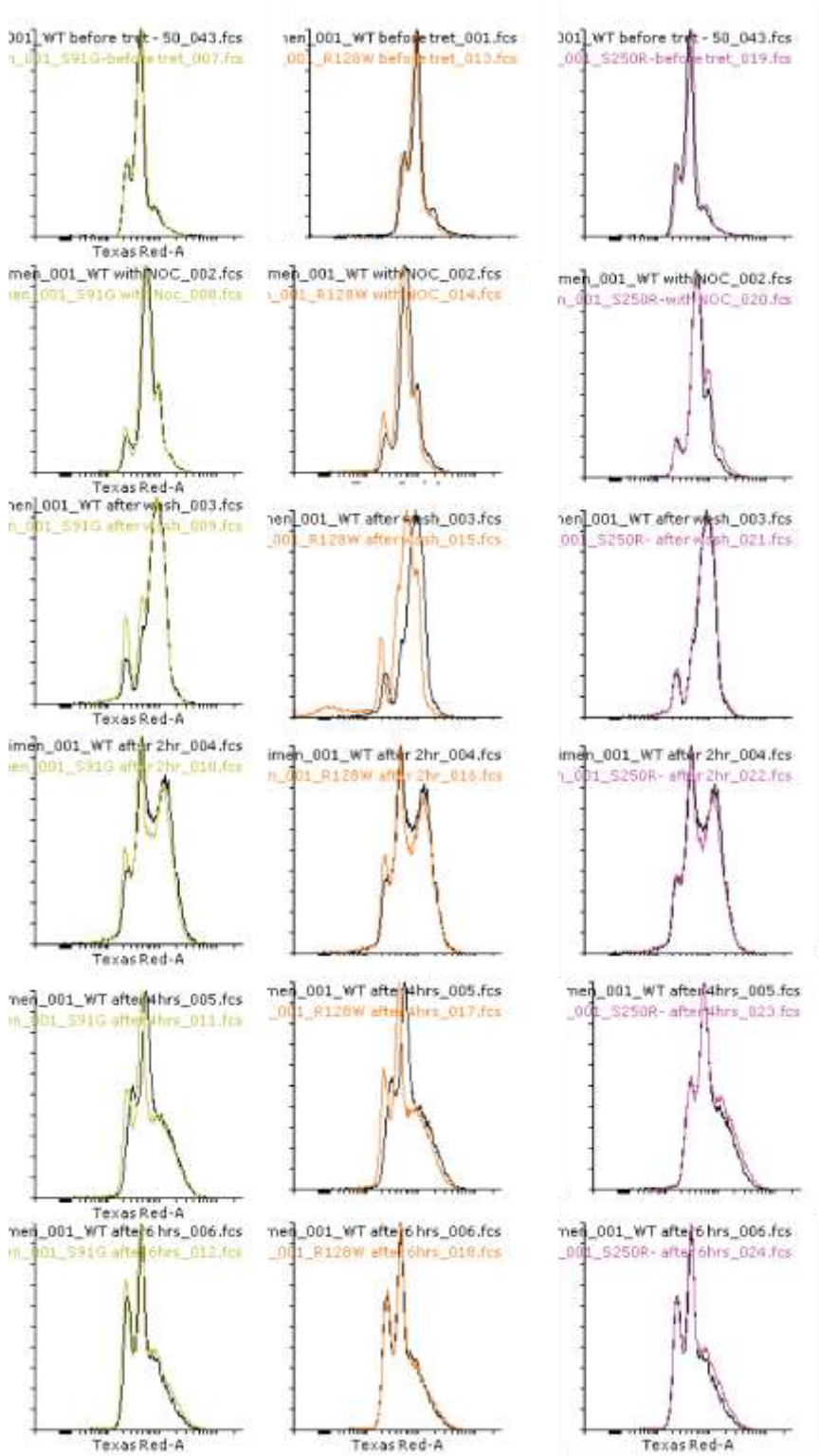


Figure A3: Ungated Flow cytometry data. On this page, histograms are shown for the mutants that are located in the catalytic domain (each overlaid onto wild-type in black).

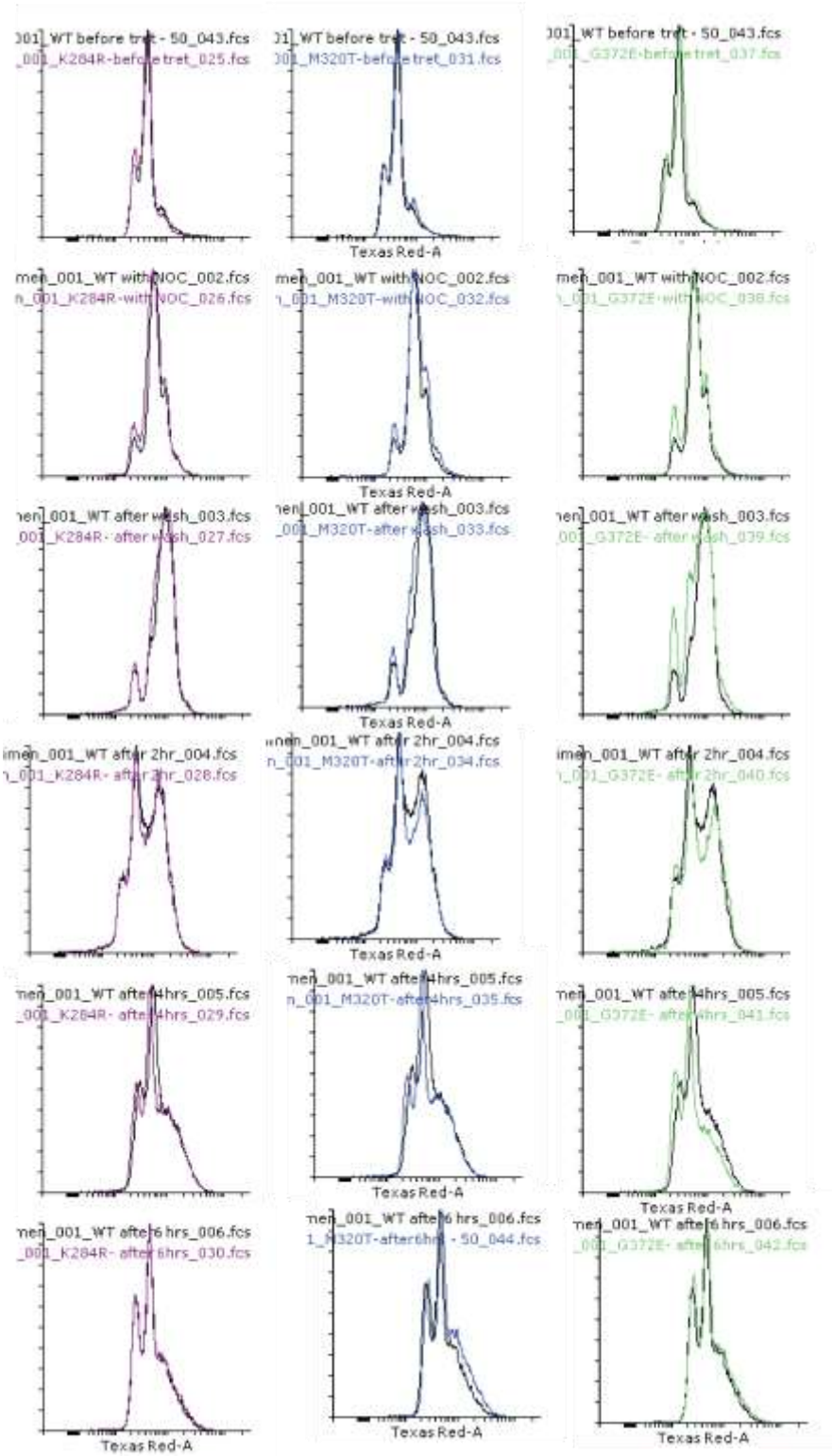


Figure A3: continued (Ungated Flow cytometry data). This second page shows ungated flow cytometry data for the mutants that are located in the PUA domain overlaid onto wild-type in black.