

**Functional Characterisation of Ribosome
Biogenesis Cofactors in *Saccharomyces
cerevisiae***

Dissertation

for the award of the degree "Doctor rerum naturalium" (Dr.rer.nat.)

of the Georg-August-Universität Göttingen

within the doctoral program: Grundprogramm Biologie
of the Georg-August University School of Science (GAUSS)

submitted by

Roman Martin

from Schwabmuenchen

Goettingen, 2014

Thesis committee

Prof. Dr. Jörg Stülke Department of General Microbiology, Georg-August-University
Goettingen

Prof. Dr. Markus T. Bohnsack Department of Molecular Biology, University Medical Center
Goettingen

Members of the examination board

Reviewer:

Prof. Dr. Jörg Stülke Department of General Microbiology, Georg-August-University
Goettingen

Second reviewer:

Prof. Dr. Markus T. Bohnsack Department of Molecular Biology, University Medical Center
Goettingen

Additional reviewer:

Prof. Dr. Dagmar Klostermeier Institute for Physical Chemistry, University of Muenster

Further members of the examination board:

Prof. Dr. Gerhard Braus Institute for Microbiology and Genetics
Dept. of Molecular Microbiology and Genetics,
Georg-August-University Goettingen

Prof. Dr. Ralf Ficner Institute for Microbiology and Genetics
Dept. of Molecular Structural Biology,
Georg-August-University Goettingen

Prof. Dr. Heinz Neumann Molecular Structural Biology
Dept. of Applied Synthetic Biology,
Georg August University Göttingen

Prof. Dr. Thomas Pieler Center for Biochemistry and Molecular Biology
Dept. of Developmental Biochemistry,
Georg-August-University Goettingen

Date of the oral examination: 23.01.2015

Table of contents

Table of contents	II
Figures.....	V
Tables.....	VI
Abbreviations	VII
Summary	VIII
1 Introduction.....	1
1.1 The structure and function of eukaryotic ribosomes.....	1
1.2 Ribosome biogenesis in <i>Saccharomyces cerevisiae</i>	3
1.3 Transcription of rRNA.....	3
1.4 Processing of precursor rRNA.....	4
1.5 Assembly of pre-ribosomal particles	6
1.6 The structure and function of snoRNPs	9
1.6.1 Box C/D snoRNPs	10
1.6.2 Box H/ACA snoRNPs	10
1.6.3 SnoRNPs involved in pre-rRNA processing.....	11
1.7 RNA helicases.....	13
1.7.1 RNA helicases in ribosome biogenesis.....	15
1.7.2 The Rok1 DEAD box RNA helicase is required for biogenesis of the SSU.....	15
1.7.3 The DEAH box helicase Prp43 is implicated in the maturation of both ribosomal subunits	16
1.8 Aims and objectives	17
2 Material and methods	18
2.1 Standard molecular biological and biochemical methods.....	18
2.1.1 Molecular cloning.....	18
2.1.2 Purification of nucleic acids	20

2.1.3	Gel electrophoresis.....	21
2.1.4	Northern and Western blotting.....	21
2.2	Cultivation and generation of yeast strains.....	23
2.2.1	Transformation of <i>Saccharomyces cerevisiae</i>	23
2.2.2	Depletion and growth analysis of yeast cultures.....	24
2.3	Crosslinking and analysis of cDNA (CRAC).....	24
2.3.1	Yeast strains used for crosslinking experiments.....	25
2.3.2	Crosslinking methods	25
2.3.3	Extract preparation and affinity purification.....	25
2.3.4	Bioinformatics and data analysis	29
2.4	Identification of chimeric reads by CLASH and prediction of snoRNA basepairing sites.....	29
2.5	Purification of pre-ribosomal complexes.....	30
2.6	<i>In vivo</i> structure probing of pre-rRNA.....	30
2.7	qPCR based analysis of snoRNA levels in pre-ribosomal complexes.....	31
2.8	Nob1 <i>in vitro</i> cleavage assay	36
3	Results	37
3.1	Functional analysis of ribosome biogenesis cofactors by <i>in vivo</i> UV crosslinking.....	37
3.2	Crosslinking experiments of Rok1 reveal high affinity for rRNA	37
3.3	Rok1 interacts with the 18S rRNA sequence	39
3.4	Structure probing confirms interaction of Rok1 with ES6H3.....	44
3.5	Rok1 interacts with snR30 and its basepairing sites on pre-rRNA	45
3.6	Identification of additional snR30 basepairing sites on 18S rRNA.....	48
3.7	Identification of basepairing sites for snoRNAs involved in pre-rRNA processing.....	50
3.8	Genome wide mapping shows high affinity of Prp43 to rRNA and box C/D snoRNAs	54
3.9	Prp43 is required for pre-ribosomal release of box C/D snoRNAs involved in LSU maturation..	56
3.10	Prp43 crosslinks 25S rRNA sequences near snoRNA guided modification sites.....	59
3.11	Prp43 crosslinks to the affected snoRNAs and copurifies with snoRNA-rRNA hybrids.....	63

3.12 Prp43 crosslinks close to the 3'-end of the 18S rRNA sequence	67
3.13 The amino acid sequence of Nob1 is conserved from archaea to eukaryotes	68
3.14 The function of Nob1 is conserved in archaea and eukaryotes	71
3.15 Nob1 is structurally and functionally conserved from archaea to eukaryotes	76
4 Discussion	78
4.1 Characterising RNA helicase function using CRAC analysis	78
4.2 A pre-ribosomal interaction network regulated by Rok1	78
4.2.1 Rok1 cooperates with protein cofactors of ribosome biogenesis	79
4.2.2 Rok1 is associated with snoRNAs involved in pre-rRNA processing	81
4.3 The role of RNA helicases in releasing snoRNAs from pre-ribosomes	83
4.4 Structural remodelling of RNP complexes by RNA helicases	85
4.5 RNA helicases are regulated by cofactors	89
4.6 Conclusions	90
Literature	91
Supplementary figure	107
Acknowledgements	i
Publications associated with this dissertation	ii
Promovierenden-Erklärung der Georg-August-Universität Göttingen	iii
Curriculum vitae	iv

Figures

Figure 1: The secondary and tertiary structures of the <i>S. cerevisiae</i> rRNAs and ribosomal subunits.....	2
Figure 2: Genome organisation of the rDNA locus and schematic overview of pre-rRNA processing in <i>S. cerevisiae</i>	5
Figure 3: Model of ribosomal subunit assembly in <i>S. cerevisiae</i>	8
Figure 4: The composition and structure of eukaryotic box C/D and box H/ACA snoRNPs.....	12
Figure 5: The CRAC method.....	27
Figure 6: Preparation of RNA for determination of snoRNA levels by qPCR.....	32
Figure 7: Plasmid encoded Rok1-C-HTP complements the growth phenotype of endogenous depletion.....	38
Figure 8: Rok1 crosslinks predominantly to ribosomal RNA.....	39
Figure 9: Rok1 crosslinks mainly to the 18S rRNA sequence.....	40
Figure 10: The major Rok1 crosslinking site on 18S rRNA is located within ES6.....	42
Figure 11: The Rok1 crosslinking sites cluster on the 3D structure of the mature 40S subunit.....	43
Figure 12: DMS structure probing confirms the Rok1 interaction with ES6H3.....	45
Figure 13: Rok1 crosslinks to snR30 within pre-ribosomal complexes.....	47
Figure 14: Rok1 crosslinking sites include new pre-rRNA basepairing sites of the snR30 snoRNA.....	49
Figure 15: SnR10 interacts with 18S rRNA sequences at Rok1 crosslinking sites.....	51
Figure 16: U3 interacts with 18S rRNA sequences at Rok1 crosslinking sites.....	52
Figure 17: U14 interacts with 18S rRNA sequences at Rok1 crosslinking sites.....	54
Figure 18: Prp43 depletion alters the accumulation of several box C/D snoRNAs on pre-ribosomal particles.....	57
Figure 19: Prp43 depletion does not affect total cellular levels of box C/D snoRNAs.....	58
Figure 20: Prp43 crosslinks to rRNA sequences of both ribosomal subunits.....	60
Figure 21: Prp43 crosslinks to sequences of the 25S rRNA in close vicinity to snoRNA modification sites.....	61
Figure 22: Prp43 crosslinking sites cluster mainly in the centre of the 3D LSU structure.....	63

Figure 23: Prp43 crosslinks to box C/D snoRNAs whose pre-ribosome association is affected by depletion of the RNA helicase.	64
Figure 24: Chimeric reads within the CRAC data validate simultaneous interaction of Prp43 with the affected snoRNAs and pre-ribosomal RNA.	66
Figure 25: Prp43 crosslinks mainly to helix 44 in the 18S rRNA sequence.	68
Figure 26: The amino acid sequence of Nob1 is conserved from archaea to eukaryotes.	70
Figure 27: Recombinant Nob1 from <i>Pyrococcus horikoshii</i> specifically cleaves pre-rRNA at the processing site D <i>in vitro</i>	72
Figure 28: The cleavage mechanism of PhNob1 is conserved from archaea to eukaryotes.	74
Figure 29: Structural analysis of PhNob1 by NMR identified critical residues for its endonuclease function.	75
Figure 30: <i>A. thaliana</i> Nob1 is able to perform a specific cleavage reaction at site D.	76
Figure 31: Interaction network of 40S pre-ribosomal maturation cofactors.	80
Figure 32: 3D Model of Nob1 binding to the 40S ribosome structure.	88
Figure S: CRAC control experiments.	107

Tables

Table 1: Plasmids.	19
Table 2: Hybridisation probes.	22
Table 3: Yeast strains.	24
Table 4: Oligonucleotides used for CRAC analysis.	28
Table 5: qPCR primer.	34
Table 6: Genome wide mapping of Rok1 sequencing reads.	39
Table 7: Prp43 crosslinks to snoRNAs and pre-rRNA.	56

Abbreviations

ATP	Adenosine triphosphate
cDNA	Complementary DNA
CLASH	Crosslinking, Ligation And Sequencing of Hybrids
CRAC	Crosslinking and analysis of cDNA
cryo-EM	Cryo-electron microscopy
Ct	Cycle threshold
DMS	Dimethyl sulphate
dNTPs	Deoxyribonucleotides
ES	Eukaryotic expansion segment
ETS	External transcribed spacer
ITS	Internal transcribed spacer
iTRIC	<i>In culturo</i> temperature regulated interaction crosslinker
LSU	Large ribosomal subunit
mRNA	Messenger RNA
NMR	Nuclear magnetic resonance spectroscopy
NTP	Nucleoside triphosphate
NTS	Non-transcribed spacer sequence
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
Pol I / II / III	RNA polymerase I / II / III
pre-rRNA	Pre-ribosomal RNA
PVDF	Polyvinylidene difluoride
qPCR	Quantitative polymerase chain reaction
rDNA	Ribosomal DNA
RPs	Ribosomal proteins
rRNA	Ribosomal RNA
RT	Reverse transcription
S	Svedberg units
SDS	Sodium dodecyl sulfate
snoRNA	Small nucleolar RNA
snoRNP	Small nucleolar ribonucleoprotein complex
snRNA	Small nuclear RNA
SSU	Small ribosomal subunit
TEV	Tobacco Etch Virus
tRNA	Transfer RNA
UTR	Untranslated regions
UV	Ultra violet

Summary

The biogenesis of eukaryotic ribosomes is a highly complex and dynamic process, which requires a multitude of cofactors. In *Saccharomyces cerevisiae*, approximately 200 proteins have been identified to function in this pathway, coordinating ribosomal protein assembly and processing of ribosomal RNA (rRNA) precursors. Furthermore, small nucleolar ribonucleoprotein complexes (snoRNPs) introduce numerous nucleotide modifications within the rRNAs, which are guided by basepairing of their RNA component (snoRNA) with the designated target pre-rRNA. SnoRNA binding and pre-rRNA processing involve structural reorganisations of the pre-ribosomal complexes during their maturation. RNA helicases are thought to be key regulators of these events. In yeast, 19 RNA helicases are known to be involved in ribosome biogenesis. However, a detailed characterisation of their molecular functions has been impeded by a lack of information about their target sites.

In this study, the recently developed crosslinking technique (CRAC) was used to identify the substrates and binding sites of two RNA helicases, Prp43 and Rok1. All pre-rRNA crosslinking sites of the DEAD box helicase Rok1 were found to cluster on one face of the small ribosomal subunit (SSU). The main binding site, which was confirmed by chemical probing, was found in the eukaryotic expansion segment 6 (ES6), where Rok1 is required for release of the snoRNA snR30. Besides snR30, Rok1 was found to interact also with other snoRNAs involved in pre-rRNA processing. In addition, a search for chimeric sequence reads in the crosslinking data set (CLASH) allowed the identification of several novel basepairing sites of these snoRNAs in the 18S rRNA sequence, close to the Rok1 target sites. New snoRNA interaction sites were discovered mainly in eukaryotic expansion segments of the rRNA sequence, suggesting that these snoRNAs might play important roles in bridging long-range interactions and are part of an extensive interaction network that is important for effective pre-rRNA folding and the coordination of the early processing events during the SSU synthesis.

CRAC analysis for the DEAH box helicase Prp43 revealed several crosslinking sites in the pre-rRNA and on box C/D snoRNAs. Interestingly, several box C/D snoRNAs, which guide a cluster of nucleotide modifications in 25S rRNA, were found to accumulate on pre-ribosomal particles after Prp43 depletion. Mapping of the Prp43

crosslinking sites on the LSU structure revealed that they are in close proximity to the modification sites of the affected snoRNAs. CLASH analysis uncovered hybrids containing sequences of the affected snoRNAs and their target sites on 25S rRNA, strongly supporting a direct function of Prp43 in unwinding the basepairing interactions of these snoRNAs. Additionally, Prp43 crosslinks 25S rRNA sequences close to the basepairing sites of two snoRNAs, whose association with pre-ribosomes was decreased after depletion of the helicase. This finding suggests that Prp43 may function in remodelling of pre-ribosomal complexes, enabling these snoRNAs to gain access to their target sites. Similarly, Prp43 was found to crosslink near the 3'-end of the 18S rRNA, where it has been proposed to function in restructuring events that facilitate the final maturation of the 18S rRNA by the endonuclease Nob1.

Homologues of yeast Nob1 were identified in higher eukaryotes and archaea. *In vitro* cleavage assays using Nob1 from *Pyrococcus horikoshii* or *Arabidopsis thaliana* demonstrated that the homologs are also able to perform a manganese dependent endonucleolytic cleavage reaction at the cleavage site D of 16S and 18S rRNA, respectively. The availability of the *P. horikoshii* Nob1 structure allowed the identification and functional analysis of residues important for substrate binding and cleavage within the PIN domain, while the zinc ribbon domain likely functions in anchoring and positioning of the enzyme on the pre-ribosomal complexes near the processing site.

1 Introduction

Ribosomes (from ribonucleic acid and Greek soma, meaning body) are macromolecular machines, which consist of ribonucleic acid (RNA) and proteins. Ribosomes are found in all kingdoms of life and are responsible for the translation of all cellular proteins.

1.1 The structure and function of eukaryotic ribosomes

The structure and function of eukaryotic ribosomes is best studied in *Saccharomyces cerevisiae*. The yeast ribosome (80S) consists of two subunits (40S and 60S) that are named according to their sedimentation properties during sucrose density centrifugation counted in Svedberg units (S; Deley, 1964). Ribosomes are composed of 79 ribosomal proteins (RPs) and four ribosomal RNAs (rRNAs; Wilson and Doudna Cate, 2012). The small 40S subunit (SSU) incorporates the 18S rRNA whereas the large 60S subunit (LSU) contains the 25S, 5.8S and 5S rRNAs. Recently, cryo-electron microscopy (cryo-EM) and atomic resolution crystallography of yeast and *Tetrahymena* ribosomes revealed the structure of the ribosomal subunits, where the rRNA forms a scaffold onto which the RPs are assembled (Armache et al., 2010; Ben-Shem et al., 2011; Klinge et al., 2011; Klinge et al., 2012; Rabl et al., 2011). The rRNA forms phylogenetically conserved secondary structures, which form the basic subdomains of the ribosomal subunits (Petrov et al., 2013; Petrov et al., 2014a). The 18S rRNA folds into the 5', the central, the 3' major and the 3' minor domain (Figure 1 A). Together with the RPs these domains form the characteristic tertiary structures (body, shoulder, platform, head and beak) of the SSU (Figure 1 B). The LSU is subdivided into six secondary structure domains (I-VI; Figure 1 C). Characteristic tertiary features of the 60S ribosomes are the central protuberance, which contains the basepairing interaction of 25S and 5S rRNA, the L1 stalk, and the acidic stalk (Figure 1 D). The comparison of eukaryotic, prokaryotic and archaeal ribosome structures revealed that the functional core is largely conserved (Petrov et al., 2014b). However, eukaryotic ribosomes are considerably larger than prokaryotic ribosomes due to additional RP content (Wilson and Doudna Cate, 2012) and extra rRNA sequences that are termed eukaryotic

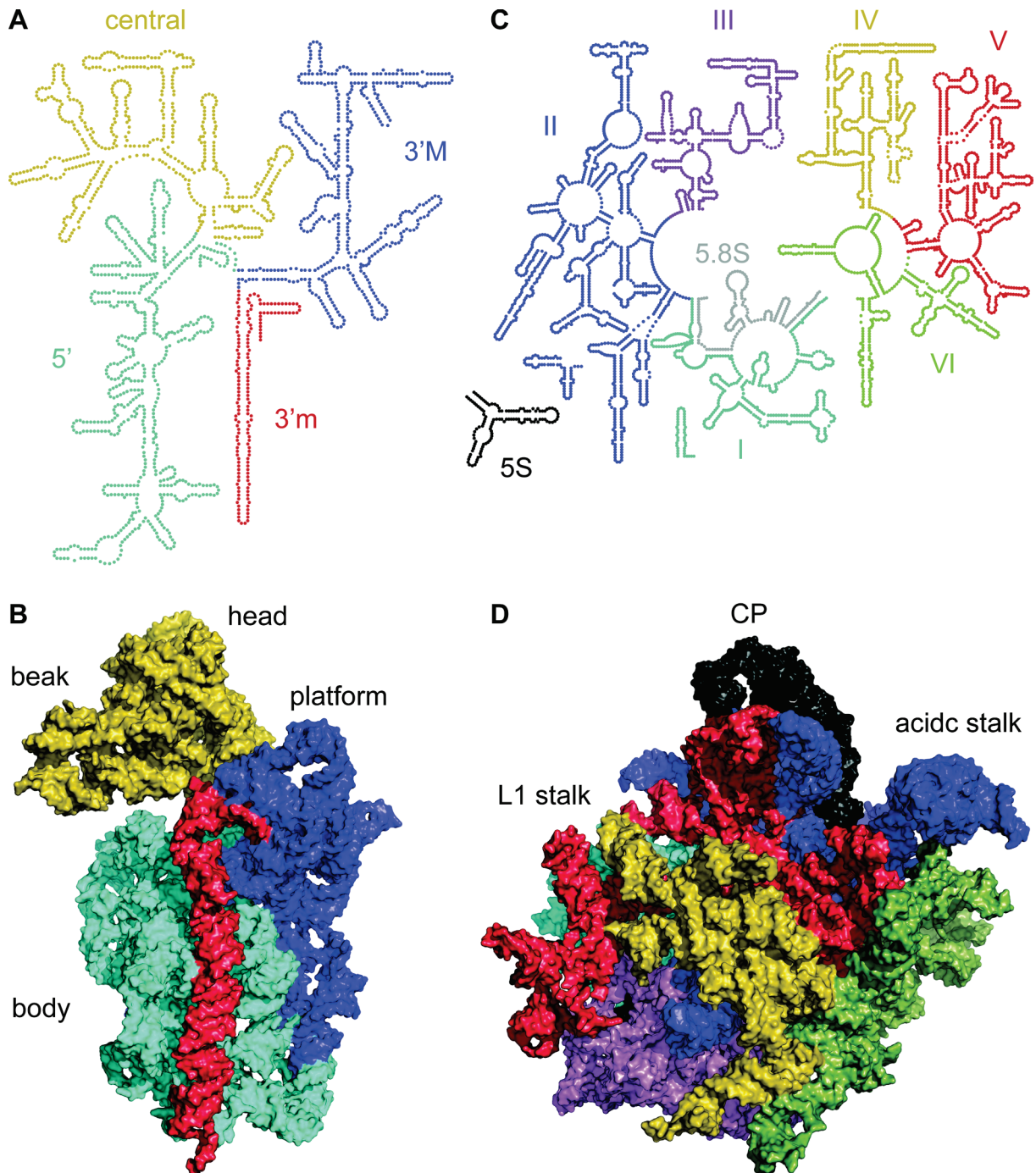


Figure 1: The secondary and tertiary structures of the *S. cerevisiae* rRNAs and ribosomal subunits. (A) Folding of the 18S rRNA is shown as a secondary structure map (Petrov et al., 2013). Individual structural domains are differentially coloured (5': 5'-domain – cyan, central: central domain – yellow, 3'M: 3'-major domain – blue, 3'm: 3'-minor domain – red). (B) The crystal structure of the 40S subunit is shown (Ben-Shem et al., 2011). The secondary structure domains are indicated with the same colours as in A. The characteristic and conserved 3D structures are labelled (beak, head, platform, body). (C) Folding of the 25S, 5.8S and 5S rRNAs shown as a secondary structure map (Petrov et al., 2013). The six secondary structure domains are indicated by differential colouring and labelled from I to VI. (I: domain I – cyan, II: domain II – blue, III: domain III – magenta, IV: domain IV – yellow, V: domain V – red, VI: domain VI – green, 5.8S: 5.8S rRNA – grey, 5S: 5S rRNA – black). (D) The crystal structure of the 60S ribosome is shown (Ben-Shem et al., 2011). The secondary structure domains are coloured as specified in C. The characteristic 3D structures of the LSU (L1 stalk, CP: central protuberance, acidic stalk) are indicated.

expansion segments (ES; Gerbi, 1986). During translation the ribosome migrates along the messenger RNA (mRNA) and the SSU mediates the decoding by codon-anticodon basepairing between transfer RNAs (tRNA) and the mRNA molecules (reviewed by Dever and Green, 2012; Rodnina et al., 2002; Rodnina et al., 2007). In contrast the LSU harbours the catalytic centre of the ribosome; its peptidyltransferase activity generates the peptide bonds between the amino acids of the emerging protein chains. Translational fidelity and efficiency of ribosomes is essential for the viability and the growth of cells and is therefore tightly regulated (Pichon et al., 2012; Van Der Kelen et al., 2009) Furthermore, ribosome biogenesis and activity is linked to cell cycle progression and various key signalling pathways (Proud, 2007).

1.2 Ribosome biogenesis in *Saccharomyces cerevisiae*

The synthesis of ribosomes is especially important in fast proliferating tissues or exponentially growing microorganism cultures. To maintain the steady state level of ribosomes during mitosis single yeast cells produce approximately 2000 ribosomes per minute (Warner, 1999). Moreover, rRNA accounts for 80% of the total cellular RNA and 50% of the RNA polymerase II (Pol II) transcribed mRNA encodes ribosomal proteins (Li et al., 1999; Warner, 1999). Therefore, ribosome biogenesis is one of the most energy consuming processes in actively dividing cells.

The biogenesis of ribosomes requires the coordinated action of all three RNA polymerases as well as approximately 200 non-ribosomal cofactors, including nucleases, RNA helicases and 75 small nuclear ribonucleoprotein complexes (snoRNPs; reviewed by Fromont-Racine et al., 2003; Henras et al., 2008; Thomson et al., 2013; Venema and Tollervey, 1999; Woolford and Baserga, 2013).

1.3 Transcription of rRNA

The synthesis of ribosomes is initiated by the transcription of the rRNA from the RDN locus, which consists of 100-150 copies of rRNA genes arranged in a tandem array on chromosome XII (Figure 2; reviewed in Long and Dawid, 1980). Each repeat contains the ribosomal DNA (rDNA) of 5S rRNA, which is transcribed by RNA polymerase III (Pol III) and the primary 35S transcript that is synthesised by RNA polymerase I (Pol I). This polycistronic rRNA sequence includes the mature 18S, 25S and 5.8S rRNAs and spacer elements, which are excised during maturation. The

internal transcribed spacers (ITS1 and ITS2) separate the mature rRNAs whereas the external transcribed spacers (5'-ETS and 3'-ETS) flank the ends of the 35S transcript. Non-transcribed spacer sequences (NTS1 and NTS2) separate the 5S (RDN5) from the 35S (RDN37) rRNA genes. Chromatin spreads (Miller Spreads) revealed that multiple transcripts can be produced simultaneously from a single active rDNA gene (Miller and Beatty, 1969). rRNA transcription occurs within the nucleolus, a non-membrane delimited sub-compartment of the nucleus, which forms around these highly compacted transcription loci of rDNA (Thiry and Lafontaine, 2005).

1.4 Processing of precursor rRNA

Processing of the 35S pre-ribosomal RNA (pre-rRNA) involves multiple sequential endonucleolytic cleavage and exonucleolytic processing events, which are crucial for the excision of the mature rRNA sequences (Figure 2; reviewed in Henras et al., 2014; Mullineux and Lafontaine, 2012). In *S. cerevisiae* the processing of the 35S pre-rRNA is initiated by cleavage events at sites A_0 and A_1 in the 5'-ETS, which generate the 33S and 32S pre-rRNA intermediates and give rise to the mature 5'-end of the future 18S rRNA (Beltrame et al., 1994; Hughes and Ares, 1991). Processing at site A_2 in ITS1 separates the precursors of the SSU (20S pre-rRNA) from the LSU (27SA₂ pre-rRNA) and results in the independent maturation of the pre-ribosomal subunits (Udem and Warner, 1972). It has been shown that the steady state levels of primary 35S transcript represents only one third of the total amount of the newly synthesised pre-rRNA transcripts since 70% of the Pol I transcripts are already cleaved cotranscriptionally at sites A_0 , A_1 and A_2 (Kos and Tollervey, 2010). After A_0 - A_2 cleavages, the pre-SSU complexes are rapidly exported to the cytoplasm where the 20S pre-rRNA is processed at site D by the endonuclease Nob1 to produce the mature 3'-end of 18S rRNA (Fatica et al., 2003; Pertschy et al., 2009). In contrast to the 20S pre-rRNA, the 27SA₂ pre-rRNA, which is a component of early pre-LSU particles, is further processed in the nucleus by two alternative cleavage pathways. Approximately 85% of the 27SA₂ pre-rRNA undergoes endonucleolytic cleavage catalysed by RNase MRP at site A_3 within ITS1 (Chu et al., 1994; Lindahl et al., 2009; Lygerou et al., 1996; Schmitt and Clayton, 1993; Shuai and Warner, 1991). Subsequently, the cleavage product (27SA₃) is rapidly trimmed to the B1_S site by the

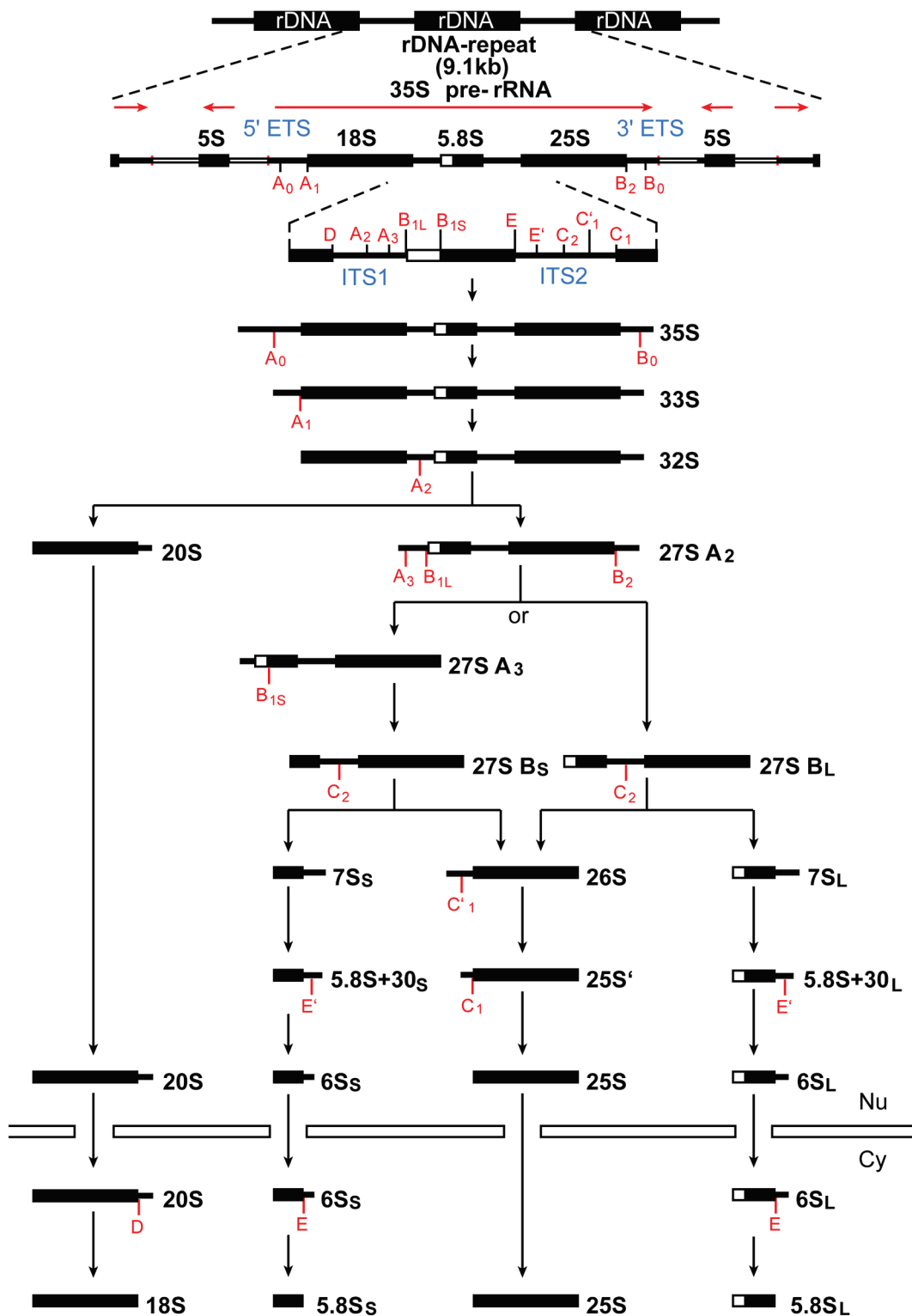


Figure 2: Genome organisation of the rDNA locus and schematic overview of pre-rRNA processing in *S. cerevisiae*. The organisation of an rDNA locus is illustrated at the top. Black bars represent transcribed regions; thick black bars indicate the mature rRNA sequences. The direction of the rDNA transcription is indicated by red horizontal arrows. Transcribed spacer regions are labelled in blue. The location of the known processing sites are indicated in red and referenced along the processing path. The names of individual processing intermediates are specified on the right. The nuclear envelope is represented by horizontal rectangles and the nuclear (Nu) and cytoplasmic (Cy) compartments are indicated.

5'-3' exonucleases Rat1 and Rrp17 to produce the 5'-end of 5.8S rRNA (5.8S_S; Henry et al., 1994; Oeffinger et al., 2009). Alternatively, about 15% of 27SA₂ is directly cleaved at processing site B1_L producing an alternative 5'-end of 5.8S rRNA (5.8S_L). Both variants of the 5.8S rRNA are later found in functional ribosomes and were suggested to provide specificity for certain mRNA subsets during translation (Schmitt and Clayton, 1993). The 27SB_{L/S} pre-rRNAs are subsequently processed in a uniform manner, their 5'-ends are generated after cleavage in ITS2 at site C₂ by an yet unidentified endonuclease and subsequent 5'-3' trimming of the 26S pre-rRNA to processing sites C'₁ and C₁ by the exoribonucleases Rat1 and Xrn1 (Geerlings et al., 2000).

The 7S_{L/S} pre-rRNAs are separated from the pre-26S rRNA after the cleavage at site C₂ and is further processed to sites E' and E by the exosome associated 3'-5' exonuclease Rrp6 producing a 5.8S precursor with a 30-nucleotide 3'-terminal extension (5.8S+30) and the 6S rRNA (Allmang et al., 1999; Briggs et al., 1998; Henry et al., 1994; Mitchell et al., 1996; Mitchell et al., 1997), which is finally trimmed to the mature 5.8S rRNA in the cytoplasm (Faber et al., 2002; van Hoof et al., 2000).

Besides the conventional pre-rRNA processing intermediates, aberrant cleavage products often accumulate upon depletion or deletion of specific maturation cofactors. Deprivation of factors required for cleavages at sites A₀-A₂ cause accumulation of 23S pre-rRNA, which arises after cleavage of the 35S rRNA at site A₃, and mainly affects the synthesis of the SSU, whereas maturation of the LSU remains unaffected (Torchet and Hermann-Le Denmat, 2000). In contrast, defects of LSU maturation cofactors largely cause a significant delay in the A₀ to A₂ cleavage events and affect also the synthesis of 40S ribosomes (Fromont-Racine et al., 2003).

1.5 Assembly of pre-ribosomal particles

In vitro reconstitution of functional ribosomal subunits from *E. coli* revealed that ribosomes are assembled in a hierarchical manner (Held et al., 1973; Nierhaus and Dohme, 1974). The maturation process of eukaryotic ribosomes is much more complex but the feature of sequential recruitment of RPs and biogenesis cofactors is conserved (Kruiswijk et al., 1978; Perez-Fernandez et al., 2007). Extensive pulldown analyses have revealed an inventory of biogenesis cofactors and provided insights into the timing of their association with pre-ribosomal complexes (Kressler et al.,

2010; Krogan et al., 2004; Nissan et al., 2002; Schafer et al., 2003). Ribosome biogenesis requires approximately 200 protein cofactors, which comprise nucleases, ATPases or GTPases, RNA helicases and non-catalytic proteins with structural functions (Fromont-Racine et al., 2003; Venema and Tollervey, 1999; Woolford and Baserga, 2013).

Initial protein cofactors are recruited co-transcriptionally to the nascent transcript leading to the formation of ball-like structures (terminal-knobs) that can be visualised by chromatin spreads (Figure 3; Miller and Beatty, 1969). These particles correspond to emerging 90S pre-ribosomal complexes, which are the first stable intermediates of the ribosome maturation process (Dragon et al., 2002; Grandi et al., 2002; Perez-Fernandez et al., 2007; Perez-Fernandez et al., 2011). They were found to contain mainly RPs and cofactors required for the early processing events at sites A_0 - A_2 , which are essential for the production of 40S ribosomes and accordingly these biogenesis factors are collectively termed SSU processome (Dragon et al., 2002; Grandi et al., 2002). Many of the SSU processome components were found to assemble stepwise as preformed subcomplexes (Krogan et al., 2004; Perez-Fernandez et al., 2007; Perez-Fernandez et al., 2011). Initially the UTP-A or t-UTP (transcription U3 protein) complex, which has a dual role in rRNA processing and transcription, is recruited to the nascent transcript of the 35S pre-rRNA (Gallagher et al., 2004; Granneman and Baserga, 2005; Perez-Fernandez et al., 2007). Subsequently, the UTP-B complex assembles on the pre-rRNA together with the U3 snoRNP and the Mpp10 subcomplex (Lee and Baserga, 1999). The UTP-C subcomplex is recruited to the pre-ribosome together with the biogenesis cofactor Rrp5 on an independent assembly pathway (Perez-Fernandez et al., 2007; Perez-Fernandez et al., 2011).

Extensive remodelling of the pre-40S complexes after A_2 cleavage results in the dissociation of most early-acting maturation factors and the association of several late-acting cofactors. The precursors of the SSU are rapidly exported to the cytoplasm where the final remodelling and maturation steps occur (Lamanna and Karbstein, 2011; Loc'h et al., 2014; Pertschy et al., 2009; Schafer et al., 2003). A surveillance step, which includes a pre-mature subunit joining mechanism, verifies

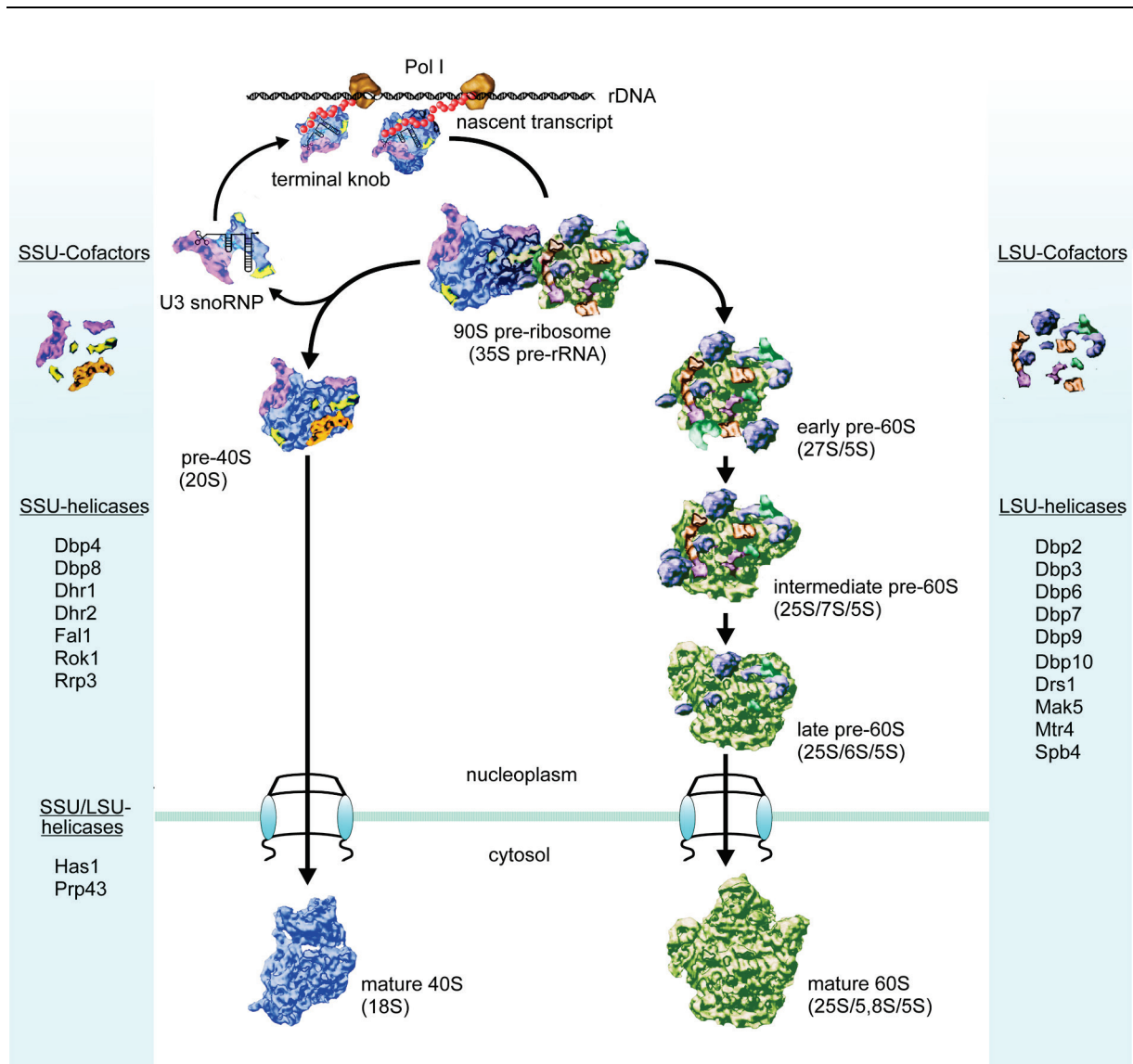


Figure 3: Model of ribosomal subunit assembly in *S. cerevisiae*. The schematic model shows the major pre-ribosomal particles that form during the biogenesis of the 40S (blue) and the 60S (green) ribosomal subunits. The cotranscriptional association and formation of the terminal knobs and the 90S pre-ribosomal particles are illustrated at the top. The separate maturation paths of the individual subunits, following the cleavage at site A₂, is depicted below. The left panel lists the RNA helicases involved in maturation of the SSU (top) and the RNA helicases implicated in the biogenesis of both subunits (bottom). The right panel lists the RNA helicases participating in the synthesis of the LSU.

the integrity of the newly synthesised 40S subunits (Lebaron et al., 2012; Strunk et al., 2011) and triggers cleavage of the 20S pre-rRNA at cleavage site D by the endonuclease Nob1 to produce the mature 3'-end of the 18S rRNA (Fatica et al., 2003). In contrast to the pre-40S particles, the maturation of the pre-60S pre-ribosomal particles continues as they are transported from the nucleolus across the nucleoplasm to the cytoplasm (Nissan et al., 2002). Three main nuclear pre-ribosomal intermediates have been described so far: early pre-60S particles, which include the 27S pre-rRNA and the 5S rRNA, intermediate pre-60S particles, in which the 25S and 7S pre-rRNA are present, and late pre-60S complexes, which contain

the fully processed 25S and the 6S pre-rRNA. The 5S rRNA assembles into early pre-60S particles as a preformed subcomplex together with the ribosomal proteins Rpl5 and Rpl11 (Ciganda and Williams, 2011; Kressler et al., 2012).

So far more than 80 cofactors have been implicated in the maturation of the LSU complexes (recently reviewed by Thomson et al., 2013; Woolford and Baserga, 2013). However, the functions of most maturation factors are still unknown. This applies also for the cofactors involved in maturation of the SSU, which promote the early cleavage events A_0 - A_2 and coordinate structural rearrangements within the SSU processome. It is assumed that most biogenesis factors might organise structural transformations, contribute to the stabilisation of maturation intermediates or mutually control their association and dissociation to the pre-ribosomal complexes.

1.6 The structure and function of snoRNPs

SnoRNPs are small nucleolar ribonucleoprotein complexes, which function in modification of pre-rRNA (Watkins and Bohnsack, 2012). Ribosome biogenesis in *S. cerevisiae* involves 75 snoRNPs, which guide more than 100 rRNA modifications, primarily in regions that are important for the catalytic activity of the ribosome (Brimacombe et al., 1993; Decatur and Fournier, 2002; Piekna-Przybylska et al., 2007). Many individual nucleotide modifications do not seem to be essential for ribosome production and fidelity. However, several clusters of modifications have been shown to be important for the stability and functional integrity of ribosomes (Baudin-Baillieu et al., 2009; Gigova et al., 2014).

In general, snoRNPs consist of a small nucleolar RNA (snoRNA) and associated proteins. Most snoRNAs guide modification of a specific nucleotide by basepairing with a designated target sequence on pre-rRNA. The associated protein cofactors perform structural functions and are, in particular, essential for the catalytic activity of the snoRNP complexes (Reichow et al., 2007; Watkins and Bohnsack, 2012). SnoRNAs can be classified into box C/D and box H/ACA snoRNAs according to the modifications they guide and conserved motifs within their RNA sequence (Balakin et al., 1996; Kiss-Laszlo et al., 1998; Ni et al., 1997). The two classes also differ in their RNA structure and in their protein composition (Reichow et al., 2007; Watkins and Bohnsack, 2012).

1.6.1 Box C/D snoRNPs

Box C/D snoRNAs guide the covalent attachment of a methyl group at the 2'-OH position of the sugar moiety of ribonucleosides. These 2'-O-methylations (Nm) occur at about 55 different positions of rRNA (Kiss-Laszlo et al., 1996). They influence the folding characteristics of rRNA sequences and additionally increase stability against hydrolysis by bases and nucleases (Helm, 2006).

Typical box C/D snoRNAs of *S. cerevisiae* contain a box C/D motif at the 3'- and 5'-ends of the RNA as well as a degenerate, internal box C'/D' motif that separates two potential modification guide sequences (Figure 4 A; Kiss-Laszlo et al., 1998). Basepairing of both the C/D and C'/D' motifs results in the formation of two characteristic secondary structures, which are known as the kink-stem (K-stem) and kink-loop (K-loop; Klein et al., 2001). Each of these secondary structure motifs constitutes a binding platform for the four snoRNP core proteins Nop1, Nop56, Nop58 and Snu13 (Reichow et al., 2007; Watkins and Bohnsack, 2012). The core of the snoRNP is formed by hetero-dimerization of Nop58 and Nop56, which directly bind the box C/D and box C'/D' motifs, respectively (Cahill et al., 2002). Additionally, both proteins were suggested to contribute to substrate binding and recognition since both contact the guide-substrate RNA duplex. Snu13 was initially thought to exclusively bind to the box C/D motif (Cahill et al., 2002; Szewczak et al., 2002). However, recent findings also suggest an interaction with the box C'/D' motif (Qu et al., 2011). The catalytic activity of these snoRNPs is provided by the methyltransferase Nop1, which binds to the box C/D and the box C'/D' motifs and also directly interacts with the guide-substrate duplex (Cahill et al., 2002). The basepairing interaction between the guide sequence and pre-rRNA comprises at least 10-12 nucleotides (Kiss-Laszlo et al., 1996). A recent study revealed that the primary guide-substrate interaction is often stabilised by additional short basepairing contacts, which form in close proximity of the main target site (van Nues et al., 2011).

1.6.2 Box H/ACA snoRNPs

Box H/ACA snoRNPs guide the isomerisation of uridine to pseudouridine (Ψ) at about 45 locations in pre-rRNA (Ganot et al., 1997a; Ni et al., 1997; Piekna-Przybylska et al., 2007). Ψ stabilises base stacking (Davis, 1995) and affects the

folding properties of the surrounding secondary structure by conferring enhanced rigidity of the rRNA (Arnez and Steitz, 1994; Charette and Gray, 2000; Helm, 2006).

Eukaryotic box H/ACA snoRNAs contain two stem-loop structures that are connected by a single stranded hinge region, which includes the box H sequence (Figure 4 B). The second hairpin is followed by a single stranded sequence, which contains the box ACA, exactly three nucleotides upstream of the 3'-end of the snoRNA (Ganot et al., 1997b). The pseudouridylation pocket is formed by short sequence elements located within the loop regions of the two hairpin structures. The cooperative binding of both sequences forces the substrate RNA into a U-like conformation (Wu and Feigon, 2007). The modified uridine residue is unpaired and therefore accessible for the snoRNP pseudouridine synthetase Cbf5 (Lafontaine et al., 1998).

A current model suggests that each hairpin is bound by one copy of the snoRNP core proteins, respectively (Watkins et al., 1998). Besides its catalytic activity, Cbf5 has structural functions in organising each stem of the hairpin (P1 and P2) into a nearly linear conformation (Li and Ye, 2006). The snoRNP core proteins Nop10 and Nhp2 bind the upper stems (P2) of the hairpin structures. Whether Gar1 interacts with the snoRNA in eukaryotes is still unclear (Bagni and Lapeyre, 1998), but a crystal structure of an archaeal box H/ACA snoRNA suggests that Gar1 is attached to the complex via an interaction with Cbf5 (Li and Ye, 2006).

1.6.3 SnoRNPs involved in pre-rRNA processing

Besides their function in guiding nucleotide modifications, snoRNA-pre-rRNA interactions also contribute to pre-rRNA folding and processing (Bachellerie et al., 2002; Watkins and Bohnsack, 2012). The snoRNPs snR10, snR30, U3 and U14 are required for the early pre-rRNA processing events A_0 - A_2 (Hughes and Ares, 1991; Liang and Fournier, 1995; Morrissey and Tollervey, 1993; Tollervey, 1987).

The box C/D snoRNP U3 is a fundamental component of the SSU processome (Dragon et al., 2002) and is essential for early pre-rRNA cleavage events A_0 - A_2 but does not guide nucleotide modifications (Hughes and Ares, 1991). The 3'-part of U3 is complementary to sequences in the central pseudoknot structure of the 18S rRNA (Hughes, 1996) and in the 5'-ETS of 35S pre-rRNA near the cleavage sites A_0 and A_1 (Beltrame and Tollervey, 1992; Beltrame and Tollervey, 1995; Dutca et al., 2011).

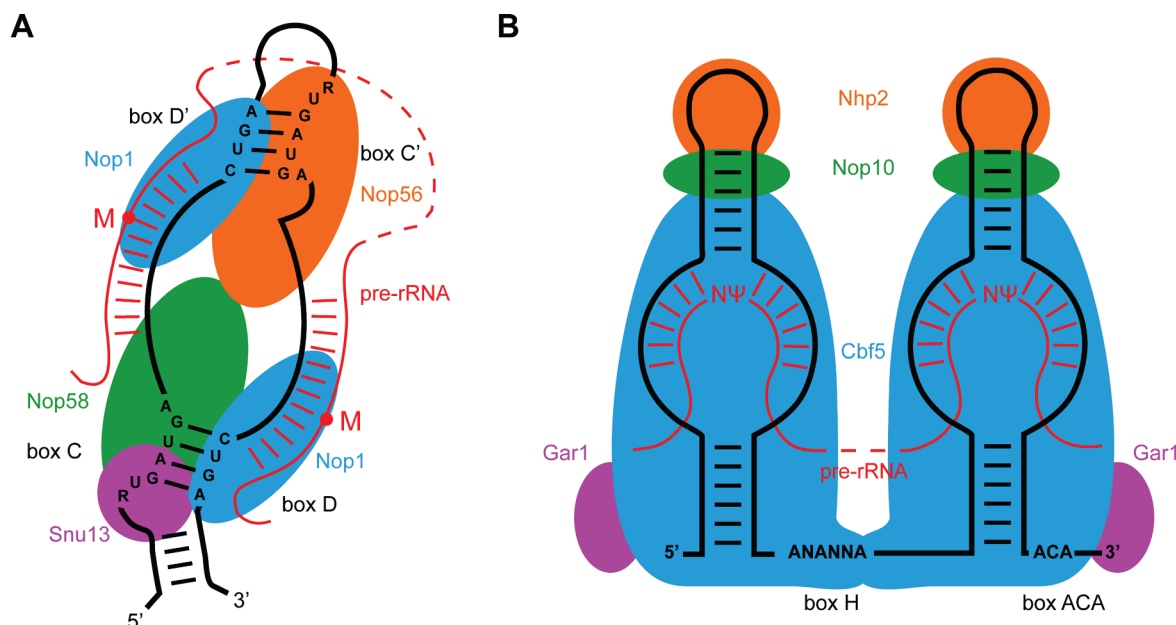


Figure 4: The composition and structure of eukaryotic box C/D and box H/ACA snoRNPs. (A) The secondary structure of box C/D snoRNAs is shown as a black line. The consensus sequence and the position of the conserved motifs (boxes C/D and C'/D') are indicated. The basepairing to pre-rRNA (red) target sequences is illustrated and the methylated residue is indicated (M). The relative positions of the snoRNP core proteins are demonstrated as coloured spheres (Nop1: blue, Nop56: orange, Nop58: green, Snu13: magenta). **(B)** The secondary structure of a box H/ACA snoRNA, containing two characteristic hairpins is shown as a black line. The consensus sequence and the position of the conserved motifs (boxes H and ACA) are indicated. The cooperative basepairing of the substrate pre-rRNA is illustrated in red. The unpaired pseudouridylated nucleotide is labelled (Ψ). Coloured spheres illustrate the relative binding position of the snoRNP core proteins (Cbf5: blue; Nop10: green; Nhp2: orange; Gar1 magenta). The snoRNP models shown were adapted from Watkins and Bohnsack (2012).

Recently two additional interactions of the U3 3'-domain with 18S rRNA were discovered suggesting that the basepairing interactions of U3 and pre-rRNA are even more complex than initially assumed (Kudla et al., 2011). The multiple basepairing interactions of U3 at locations critical for pre-rRNA processing play a crucial role in folding and structural organisation of the 18S pre-rRNA (Dutca et al., 2011). However, U3 could also contribute to the processing of pre-rRNA by delivering associated protein cofactors to the early pre-ribosomal complexes (Dragon et al., 2002), such as Rrp9, which is a U3 specific snoRNP core protein (Granneman et al., 2009). Moreover, the cleavage of the pre-rRNA occurs within the SSU processome, whose assembly depends on U3 snoRNP association and subsequent recruitment of additional cofactors (Dragon et al., 2002; Perez-Fernandez et al., 2007).

Similar to U3, the box H/ACA snoRNP snR30 is essential for the early processing events within the 90S particles (Bally et al., 1988; Lemay et al., 2011; Morrissey and Tollervey, 1993). The processing function of snR30 requires basepairing of two short

conserved sequences in its 3'-hairpin (m1 and m2) with 18S pre-rRNA in expansion segment 6 (ES6; Atzorn et al., 2004; Fayet-Lebaron et al., 2009). Although snR30 does not perform nucleotide modifications, the interaction with ES6 requires the cooperative basepairing of both short sequences within the loop structure of the 3'-hairpin, which is similar to the interaction mode of a classical pseudouridylation pocket. The remarkably long nucleotide sequence of snR30 is a potential binding platform for additional protein cofactors that might contribute to the processing of pre-rRNA. Association to snR30 was proposed for Utp23, Kri1 and Nop6 (Hoareau-Aveilla et al., 2012; Lemay et al., 2011; Lu et al., 2013). However, if these proteins are part of the snoRNP core complex or only transiently associated in the context of pre-ribosomal complexes remains unclear.

The box C/D snoRNA U14 and the box H/ACA snoRNA snR10 possess dual functionality in both pre-rRNA processing and nucleotide modification. U14 guides methylation of residue Cm414 in 18S rRNA (Dunbar and Baserga, 1998) and has an essential role during the early processing events A_0 - A_2 (Jarmolowski et al., 1990; Liang and Fournier, 1995), which depends on a base pairing interaction with the 5'-domain of the 18S rRNA sequence (Liang and Fournier, 1995; Morrissey and Tollervey, 1997). SnR10 is responsible for the pseudouridine formation at position U2923 in the 25S rRNA sequence (Ni et al., 1997) and its deletion causes temperature-dependent cleavage defects (Tollervey and Guthrie, 1985; Tollervey, 1987), which are mediated by a short 7-nucleotide long sequence element in its 5'-hairpin (Liang et al., 2010).

It is commonly accepted that snoRNAs contribute to the structural organisation of pre-rRNA. However, how the diverse basepairing interactions of the snoRNAs with the pre-rRNA regulate the maturation of ribosomes is not yet fully understood.

1.7 RNA helicases

RNA helicases were initially defined as enzymes that utilise the energy of nucleotide triphosphate (NTP) binding and hydrolysis to unwind RNA duplexes. However, extensive research during the last decade has expanded their range of functions and showed that RNA helicases are a large group of ubiquitous enzymes, which are involved in various aspects of RNA remodelling.

Sequence comparison enabled the classification of helicases into six superfamilies (SF1-6; Gorbalenya and Koonin, 1993; Singleton et al., 2007). The largest superfamily (SF2) contains the two major families of RNA specific helicases: DEAD and DEAH/RHA RNA helicases, which are named according to a consensus amino acid sequence motif (Fairman-Williams et al., 2010; Gorbalenya and Koonin, 1993). RNA helicases possess up to 12 conserved sequence motifs within the two RecA domains of the helicase core (Story et al., 1992; Ye et al., 2004). The individual functions of the conserved motifs, comprising NTP binding, NTP hydrolysis, RNA binding and regulatory tasks, were assigned during structural and biochemical characterisations of various RNA helicases (Bleichert and Baserga, 2007; Fairman-Williams et al., 2010).

In addition to the structural differences DEAH and DEAD box helicases vary significantly in their operation mode. DEAH helicases are thought to act in a processive manner similar to many DNA helicases (Lohman and Bjornson, 1996). Effective access to the substrate commonly requires a single stranded 3'-overhang on the substrate duplex. Unwinding then occurs by consecutive NTP consumption, which provides the energy for a directional translocation of the helicase along the backbone of the loading strand and simultaneous removal of the complementary top strand (for a detailed review see: Pyle, 2008). In contrast, DEAD box helicases directly target the substrate duplexes and perform the strand separation by a local unwinding mechanism, which enables the enzyme to dissolve interactions up to 10 – 12 basepairs (Jankowsky, 2011; Linder and Jankowsky, 2011; Yang et al., 2007). In most cases, strand separation only requires NTP binding (Chen et al., 2008; Henn et al., 2010). The hydrolysis of NTP is essential for the release of the unwound RNA strands and for the recycling of the enzyme (Liu et al., 2008). Although the unwinding occurs in a non-processive manner, interaction of additional helicase domains with single stranded or structured extensions on the RNA substrate were found to significantly stimulate the activity of DEAD box helicases (Garcia et al., 2012; Hilbert et al., 2009). These additional domains are also thought to confer substrate specificity and enable the interaction with protein cofactors (Jankowsky, 2011; Jarmoskaite and Russell, 2014).

Besides the classical unwinding of RNA duplexes, RNA helicases have been shown to anneal RNA strands, function as clamps or displace proteins from RNA containing

complexes (Jarmoskaite and Russell, 2014; Linder and Jankowsky, 2011). This broad range of functions explains why RNA helicases are ubiquitous and often essential for viability. RNA helicases play key roles in many pathways of RNA metabolism but the majority is implicated in pre-mRNA splicing and in ribosome biogenesis (Bleichert and Baserga, 2007; Jarmoskaite and Russell, 2014).

1.7.1 RNA helicases in ribosome biogenesis

The genome of *S. cerevisiae* encodes 39 putative RNA helicases, of which 19 are suggested to participate in the maturation of ribosomes (Martin et al., 2013). 16 of these RNA helicases are essential for normal cell proliferation indicating the importance of their functions. Protein depletion followed by analysis of pre-RNA processing intermediates and co-immunoprecipitation studies have revealed that seven RNA helicases are required for the early maturation events of the 40S subunit, while ten RNA helicases were found to act in maturation of the 60S particles and two RNA helicases (Prp43 and Has1) are required for the maturation of both subunits (Figure 3). Despite this preliminary classification, a precise functional characterisation is still missing for most RNA helicases. However, it is thought that they are involved in various remodelling processes of pre-rRNA and the removal of protein cofactors or snoRNPs from pre-ribosomal complexes. Initial evidence for a role of RNA helicases in snoRNA release was provided by several studies, which showed accumulation of snoRNAs on pre-ribosomal particles after depletion of the RNA helicases Has1, Dbp4 and Rok1 (Bohnsack et al., 2008; Kos and Tollervey, 2005; Liang and Fournier, 2006). However, direct release of a snoRNA by an RNA helicase has not been demonstrated so far, mainly because precise target sites of RNA helicases on pre-rRNA are still elusive.

1.7.2 The Rok1 DEAD box RNA helicase is required for biogenesis of the SSU

The DEAD box helicase Rok1 is an essential nuclear protein, which is involved in pheromone induced gene expression (Kim et al., 2004) and required for the early processing events of the primary rRNA transcript. The depletion of Rok1 leads to accumulation of 35S and the aberrant 23S pre-rRNA (Song et al., 1995; Venema et al., 1997). Physical and genetic interactions with early 40S ribosome biogenesis maturation factors, such as Rrp5 and the snR10 snoRNP, showed that Rok1 acts in context of the SSU processome (Lebaron et al., 2013; Venema et al., 1997; Vos et

al., 2004). Consistent with this, the enzymatic activity of Rok1 is required for the release of snR30 from pre-ribosomal complexes (Bohnsack et al., 2008; Oh and Kim, 1999). However, the details of Rok1 mediated snoRNA release and possible extra functions of the RNA helicase are still elusive. The low unwinding rates of Rok1 measured *in vitro* (Garcia et al., 2012), suggest that the helicase might be regulated by additional SSU components or is stimulated by specific RNA substrates. Moreover, a recent *in vitro* study, which proposed an Rrp5 dependent pre-rRNA annealing activity for Rok1 (Young et al., 2013), indicated that the tasks administrated by the RNA helicase could even be more diverse than initially assumed.

1.7.3 The DEAH box helicase Prp43 is implicated in the maturation of both ribosomal subunits

The DEAH box helicase Prp43 is a multifunctional RNA helicase that was first described as a component of late spliceosomal particles (Arenas and Abelson, 1997). Prp43 is responsible for the release of the intron lariat and contributes to the disassembly of the spliceosome (Fourmann et al., 2013; Koodathingal et al., 2010; Martin et al., 2002; Mayas et al., 2010). In addition, Prp43 was found to contribute in the maturation of both ribosomal subunits since depletion of the helicase causes defects in the production of the 20S, 27S and 7S pre-rRNAs and co-immunoprecipitation analyses revealed physical interactions with pre-rRNAs and with box C/D snoRNAs (Combs et al., 2006; Lebaron et al., 2005; Leeds et al., 2006).

In both, splicing and ribosome biogenesis, Prp43 function is regulated by different protein cofactors, which contain a domain with several conserved glycine residues (G-patch; Aravind and Koonin, 1999). During assembly of the 40S ribosome Prp43 was shown to cooperate with the G-patch protein Pfa1 (Lebaron et al., 2005; Pertschy et al., 2009), whereas the G-patch protein Gno1 is a potential cofactor of Prp43 within 90S particles or early intermediates of the LSU (Chen et al., 2014; Guglielmi and Werner, 2002). Although the functional implications of Prp43 during splicing are well defined, the roles of the helicase in ribosome biogenesis are largely elusive.

1.8 Aims and objectives

Ribosome biogenesis is a highly complex and dynamic process that requires extensive structural reorganisations of pre-rRNAs as well as the release of protein cofactors or snoRNPs from precursor particles. ATP-dependent RNA helicases are thought to be the main mediators of such remodelling processes. However, the details of RNA helicase functions in maturation of ribosomal subunits could not be explored so far, because their target sites in pre-ribosomal particles have remained elusive.

The objectives of this work were to identify interacting RNAs and RNA binding sites for the RNA helicases Prp43 and Rok1 in the yeast *S. cerevisiae* using the UV crosslinking and analysis of cDNA (CRAC) approach. The functions of the two RNA helicases were then analysed.

This study therefore aimed to:

- Identify crosslinking sites for the RNA helicase Rok1 on pre-rRNAs.
- Confirm Rok1 binding to 18S rRNA precursors by structure probing.
- Analyse the snoRNA-pre-rRNA interactions in pre-ribosomal particles based on the CRAC data obtained for Rok1.
- Identify crosslinking sites for the RNA helicase Prp43 on pre-rRNAs.
- Analyse the function of Prp43 in the biogenesis of the large ribosomal subunit.
- Study the function and conservation of the pre-ribosomal endonuclease Nob1, which has been shown to genetically interact with Prp43.

2 Material and methods

Chemicals used in this study were purchased from Agilent Technologies (Santa Clara, US), Life Technologies (Darmstadt, GER), Merck (Darmstadt, GER), Neolab (Heidelberg, GER), Perkin Elmer (Rodgau, GER), Roche (Mannheim, GER), Roth (Karlsruhe, GER), Sigma Aldrich (Mannheim, GER), and VWR (Darmstadt, GER). Antibodies were ordered either from Abcam (Cambridge, UK) or Sigma Aldrich. Kits and enzymes used for molecular cloning, and standard RNA or DNA techniques, were derived from Affymetrix (Santa Clara, US), New England Biolabs (Frankfurt am Main, GER), Macherey-Nagel (Dueren, GER), Qiagen (Hilden, GER), Sigma Aldrich and Thermo Scientific (St. Leon-Roth, GER). Standard oligonucleotides were purchased from Sigma Aldrich, adapters used for CRAC analysis were synthesised by Integrated DNA Technologies (Coralville, US), Radiochemicals were ordered from Perkin Elmer.

2.1 Standard molecular biological and biochemical methods

Current methods of molecular cell biology were in general performed according to Sambrook and Russell (2001) and frequently used techniques in this study are described in more detail as follows.

2.1.1 Molecular cloning

The construction of the plasmids used in this study (Table 1) was performed according to standard techniques described in Sambrook and Russell (2001). In short, for expression of tagged Rok1 protein in *S. cerevisiae* the coding sequence of Rok1 was amplified by polymerase chain reaction (PCR) from genomic yeast DNA and cloned (using BamHI and BglII restriction sites) into pRS415-C-HTP and pRS415-N-HTP, derivatives of the pRS415 plasmid (Sikorski and Hieter, 1989). Constructs for recombinant expression of PhNob1 were generated, by cloning (using restriction sites NdeI and BamHI) the synthetic PhNob1 gene (Entelechon), which was codon optimized for expression in *E. coli*, into the pET11a plasmid (Novagen). The coding sequence of AtNob1 was cloned into plasmids A21 (using BamHI and XmaI restriction sites) and A46 (NcoI and XmaI), two derivatives of pET11a. Mutants were generated by site-directed mutagenesis (Zheng et al., 2004).

RNA used for the *in vitro* cleavage experiments was transcribed from constructs, which were generated by annealing of two oligonucleotides containing the desired rRNA sequences and cloning of the fragments into the pGEM4Z (Promega; using restriction sites HindIII and EcoRI) or pBluescript SK (Stratagene; using SacI and EcoRI restriction sites) plasmids. The constructs for transcription of the stable RNA hairpins used for *in vitro* analysis were designed according to Schneider et al. (2009) and Pertschy et al. (2009). The *in vitro* transcripts of PhNob1 pre-rRNA contained the pre-rRNA sequence 5'-GGGAGACAAGCUUAAGUCGUAACAAGGUAGCCGUAGG GGAACCUACGGCUCGAUCACCUCCUAUCGCCGGAAACCCCGUCCGGGGGAAU U-3'. The hairpin constructs included the pre-rRNA sequences 5'-AUCACCUCCUAU CGCC-3' and 5'-CGAUCACCUCCUAUCGCC-3'.

Table 1: Plasmids

ID	Name	Reference
S05-045	pMK140	D. Tollervey (unpublished)
S05-137	pRS415-Rok1-C-HTP	Martin et al. (2014)
S05-139	pRS415-N-HTP-Rok1	Martin et al. (2014)
S05-188	pET11a-PhNob1	Veith et al. (2012a)
S05-270	pGEM4Z-PhRRNA-site-D	Veith et al. (2012a)
S05-376	pET11a-PhNob1-D12N	Veith et al. (2012a)
S05-377	pET11a-PhNob1-S79A	Veith et al. (2012a)
S05-378	pET11a-PhNob1-R115A	Veith et al. (2012a)
S05-379	pET11a-PhNob1-D100N	Veith et al. (2012a)
S05-381	pGEM4Z-PhRRNA-hairpin-S	Veith et al. (2012a)
S05-382	pGEM4Z-PhRRNA-hairpin-L	Veith et al. (2012a)
S05-405	A21-AtNob1-D50N	Missbach et al. (2013)
S05-439	pBS1539-His3-HTP	Martin et al. (2014)
S05-467	A46-AtNob1	Missbach et al. (2013)
S05-481	pBluescript SKII(+)	Stratagene
S05-490	pBluescript SKII(+)-AtRRNA-site-D	Missbach et al. (2013)
S05-1014	pAV162	K. D. Entian (unpublished)

2.1.2 Purification of nucleic acids

Small-scale plasmid DNA purification from *E. coli* was performed by alkaline extraction (Birnboim and Doly, 1979). Large amounts of plasmid DNA were prepared using the NucleoBond plasmid DNA purification kit (Macherey-Nagel), according to manufacturers instructions.

For purification of genomic yeast DNA, exponentially growing cells (OD₆₀₀ 0.5) were harvested and resuspended in 250 µl Lysis Buffer (2% (v/v) TritonX 100, 1% (w/v) SDS, 100 mM NaCl, 10 mM Tris/HCl pH 8.0 and 1 mM EDTA). The cells were lysed by vortexing the suspension together with glass beads and 250 µl phenol-chloroform-isoamylalcohol (25:24:1) at 4 °C. The nucleic acids were precipitated by addition of 25 µl of 3 M sodium acetate (pH 5.2) and 750 µl 100% ethanol followed by centrifugation (20,000 g, 5 min, RT).

Purification of RNA from yeast cells was performed by hot phenol chloroform extraction described in Rio et al. (2011). In short, 25 ml of an exponentially growing yeast culture (OD₆₀₀ 0.5-0.6) were harvested by centrifugation (6,000 g, 5 min, 4 °C) and resuspended in 2 ml of Lysis Buffer (50 mM sodium acetate, 10 mM EDTA). Cells were lysed by addition of 250 µl of 10% (w/v) sodium dodecyl sulphate (SDS) solution and 2.5 ml of Phenol A Solution containing 50% (v/v) Lysis Buffer and 50% (v/v) phenol by vortexing for 1 min at 65 °C. Subsequently the samples were cooled on ice for 5 min and centrifuged (20,000 g, 5 min, 4 °C). The aqueous phase was recovered and a second RNA extraction step was performed by addition of 2.5 ml of Phenol B Solution containing 25% (v/v) phenol, 25% (v/v) chloroform and 50% (v/v) Extraction Buffer (10 mM sodium acetate, 100 mM NaCl, 1 mM EDTA). The aqueous phase was recovered after centrifugation (20,000 g, 5 min, 4 °C) and residual phenol was removed by addition of 2 ml of chloroform-isoamyl mixture (24:1). The aqueous supernatant was recovered by centrifugation (20,000 g, 5 min, 4 °C) and precipitated with 350 µl sodium acetate (pH 5.2) and 10 ml 100% ethanol. The RNA pellet was washed with 70% ethanol and resuspended in 50 µl water.

Purification of RNA from cell lysates or sucrose gradient fractions was performed according to a short phenol chloroform extraction procedure described in Chomczynski and Sacchi (1987) and Chomczynski and Sacchi (2006). In short, 530 µl of cell extract was mixed with 400 µl of GTC mix (2 M guanidinium thiocyanate,

25 mM Tris/HCl pH 8.0, 5 mM EDTA, 1% (v/v) N-Lauroylsarcosine 0.07 M β -mercaptoethanol) by vortexing. Extraction of RNA was performed with 800 μ l phenol followed by vortexing. After addition of 210 μ l chloroform and 40 μ l of 3 M sodium acetate (pH 5.2) samples were centrifuged (20,000 g, 5 min, 4 °C). The aqueous phase was collected and precipitated with 100% ethanol. The RNA pellet was washed with 70% ethanol and resolved in 30 μ l water.

The concentrations and quality of all RNA or DNA preparations were determined using the NanoDrop 2000 spectrophotometer (Thermo Scientific).

2.1.3 Gel electrophoresis

Proteins and nucleic acids were separated according to their size and charge in an electric field using gel electrophoresis. For denatured proteins SDS polyacrylamide gel electrophoresis (PAGE) according to Laemmli (1970) was used.

DNA and large RNA molecules were analysed by agarose gel electrophoresis. DNA was resolved in 1.5-2% (w/v) agarose gels in TAE buffer (40 mM Tris, 20 mM acetic acid and 1 mM EDTA). Large RNA molecules were denatured at 55 °C for 1 h in Glyoxal Loading Buffer (60% (v/v) DMSO, 20% (v/v) deionized glyoxal, BPTe Buffer, 5% (v/v) glycerol, 2 mg ethidium bromide) and separated in 1.2% (w/v) agarose gels in BPTe buffer (10 mM Pipes, 30 mM Bis-Tris, 1 mM EDTA).

Small RNA and cDNA fragments were prepared for electrophoresis in Formamide Loading Buffer consisting of 95% (v/v) formamide, 0.025% (w/v) bromophenolblue, 0.025% (w/v) xylene cyanole, 0.025% (w/v) SDS and 5 mM EDTA for 5 min at 85 °C. The denatured RNA samples were resolved by polyacrylamide gel electrophoresis (PAA) in TBE buffer (89 mM Tris, 89 mM boric acid, 1 mM EDTA). For single nucleotide resolution PAA gels were run in the S2 sequencing gel apparatus (Biometra).

2.1.4 Northern and Western blotting

Proteins separated by SDS PAGE were transferred on polyvinylidene difluoride (PVDF) membranes using a semi-dry electrophoretic blotting method (Towbin et al., 1979). The proteins were visualised by immunodetection and the Western Lightning enhanced chemiluminescence reagents (Perkin Elmer) or with Alexa Fluor fluorescent

secondary antibodies (Life Technologies) and the Odyssey Sa infrared imaging system (Ly-Cor).

Small RNA molecules separated on PAA gels were transferred on positively charged Hybond N+ nylon membranes (GE Healthcare) by wet electrophoretic blotting in TBE buffer (2.1.3).

Large RNA molecules resolved by agarose gel electrophoresis were fragmented by a 20 min treatment of the gel in 0.1 M sodium hydroxide solution. The fragmentation was quenched with Neutralisation Buffer (0.5 M Tris/HCl pH 7.5, 1.5 M NaCl). The transfer of the RNA on Hybond N+ membranes (GE Healthcare) was performed in 6x SSC buffer (150 mM NaCl, 15 mM sodium citrate) using the VacuGene XL vacuum blotting system (GE Healthcare).

RNA was visualised using methylene blue staining (0.02% (w/v) methylene blue, 300 mM sodium acetate pH 5.5) or by hybridisation with radioactively labelled oligonucleotides (Table 2) and a phosphorimager. Transfer of the radioactive γ -phosphate from 25 μ Ci [γ - 32 P] adenosine triphosphate (ATP) to the 5'-end of the oligonucleotides was performed using T4 polynucleotide kinase (PNK, Thermo Scientific), according to manufacturers instructions. The hybridisation was performed at 37 °C for 6 hours in Hybridisation Buffer (6x SSC buffer, 1% (w/v) SDS, 5x Denhard's reagent (Sambrook and Russell, 2001), 100 μ g/ml salmon sperm DNA and 1-5 μ g/ml radioactively labelled oligonucleotide).

Table 2: Hybridisation probes

Target	Sequence (5'-3')
U14	TCACTCAGACATCCTAGG
snR59	GGTGATTAACGACAGCATTGTCAAAGACTAGTCGA
snR50	GCTGCAAATTGCTACCTCTTTCA
U3A	UUAUGGGACUUGUU
snR39	GGTGATAAGTTACGACAGC
snR41	GGGTTGTCGACATGTAGTTA
snR39b	TGCTAGTCACTTTTTGGAATGCC
snR60	CAATCAGTTGAACTATGCATCTTTGG
snR72	GATCAGACTGACGTGCAAATCATT
ITS1	CGGTTTTAATTGTCCTA
snR30	CTAAGTTAAACTCGTCAACG

2.2 Cultivation and generation of yeast strains

The yeast strains used in this work (Table 3) are all derivatives of *S. cerevisiae* BY4741 (Brachmann et al., 1998). Cultures were grown in sterile YPD medium (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) D-glucose) or synthetic dropout medium (Formedium) and kept in mid log-phase until harvesting of cells.

2.2.1 Transformation of *Saccharomyces cerevisiae*

Transformation of yeast strains and genomic integration of the tetracycline repressible promoter and protein tags were performed according to Gietz et al. (1992) and Longtine et al. (1998). In short, a 15 ml culture of exponentially growing yeast cells with an optical density (OD₆₀₀) of 0.5-0.7 (~1x10⁸ cells) were harvested, washed with sterile water and resuspended in Transformation Mixture (10 mM Tris/HCl pH 7.5, 1 mM EDTA, 100 mM lithium acetate, 50 µg salmon sperm carrier DNA, 40% (v/v) polyethyleneglycole (PEG) 4000, 5 µg DNA). The transformation was performed using either plasmid DNA or precipitated PCR product amplified from plasmids pMK140 (Table 1) for integration of the TetO₇ promoter or pBS1539-His3-HTP (Table 1) for genomic His₆-Tev-Protein A (HTP) tagging. Homologous recombination of the PCR products was achieved by target specific sequences incorporated in the forward and reverse primers. The transformation was performed by 30 min incubation at 30 °C, followed by a “heat shock” at 42 °C for 15 min. Selection of positive transformants was performed by plating on synthetic dropout medium. Alternatively, selection using antibiotic marker genes was performed after 16 h pre-incubation on YPD plates followed by replicate plating on selective medium. To verify successful genomic integration of HTP or HA tags, 2 ml of cells from saturated culture were pelleted, resuspended in 100 ml water and lysed by vortexing together with glass beads. The extracted proteins were precipitated with 20% (v/v) trichloroacetic acid (TCA), the pellets were washed with ice-cold 100% acetone and subsequently analysed by SDS PAGE (2.1.3) and Western blotting (2.1.4) using anti-HA (Sigma-Aldrich) and anti-PAP (Sigma-Aldrich) antibodies.

Table 3: Yeast strains

ID	Strain Name	Genotype	Reference
YMB006	BY4741	MATa, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Brachmann et al. (1998)
YMB146	TetO ₇ -Rok1	YMB279, P _{tet} -3HA-rok1-NatMX6	Bohnsack et al. (2008)
YMB147	TetO ₇ -Rok1	YMB279, P _{tet} -3HA-rok1-NatMX6	Bohnsack et al. (2008)
YMB279	BY4741 (TetO ₇ parent)	MATa, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2::tTA, leu2::P _{cmv} tetR'SSN6 URA3-K.I	Alexander et al. (2010)
YMB348	TetO ₇ -Rok1 (empty plasmid)	YMB146 pRS415-Leu2	Bohnsack et al. (2008)
YMB350	Rok1 (WT)	YMB146, pRS415-rok1(wt)-Leu2	Bohnsack et al. (2008)
YMB455	Prp43-C-HTP	YMB006, prp43-C-HTP-URA3	Bohnsack et al. (2009)
YMB458	TetO ₇ -Prp43	YMB279, P _{tet} -3HA-prp43-NatMX6	Bohnsack et al. (2009)
YMB536	Rok1-C-HTP	YMB147, pRS415-rok1-C-HTP-Leu2	Martin et al. (2014)
YMB720	Rok1-N-HTP	YMB146, pRS415-NHTP-rok1-Leu2	Martin et al. (2014)
YMB870	TetO ₇ -Rok1, Enp1-C-HTP	YMB146, enp1-C-HTP-HIS3	Martin et al. (2014)

2.2.2 Depletion and growth analysis of yeast cultures

Growth analysis was performed with pre-grown cultures that were kept in mid-log phase for at least 24 hours. Depletion of proteins with a tetracycline repressible promoter was induced by addition of 5 µg/ml doxycycline (Sigma Aldrich) in YPD medium or 20 µg/ml in synthetic dropout medium (Alexander et al., 2010). The cell proliferation rates were monitored by optical density measurement until growth defects were observed.

2.3 Crosslinking and analysis of cDNA (CRAC)

The CRAC (Crosslinking and analysis of cDNA) analysis was performed in collaboration with Maike Ruprecht (working group of Prof. Dr. Enrico Schleiff, Goethe-University, Frankfurt am Main) and as previously described by Bohnsack et al. (2012) and Granneman et al. (2009). A schematic of the CRAC procedure is shown in Figure 5.

2.3.1 Yeast strains used for crosslinking experiments

CRAC of Prp43 was performed with yeast strain YMB455 (Table 3), which carries a C-terminal His₆-Tev-Protein A (HTP) tag at the genomic Prp43 locus. Rok1 crosslinking was performed with strains YMB536 and YMB720 (Table 3), which were constructed by transfection of plasmids S05-137 and S05-139 (Table 1) into the Rok1 depletion strains YMB146 and YMB147, enabling expression of C- or N-terminal HTP tagged Rok1, respectively (Table 3). Untagged genomic Rok1 was depleted by addition of 20 µg/ml doxycycline for 9 hours before UV irradiation.

2.3.2 Crosslinking methods

Crosslinking was induced by application of UV light (254 nm) to living cells in culture medium (*in culturo*) or after pelleting and resuspension in a small volume (*in vivo*) or after isolation of Prp43 containing complexes on IgG sepharose (*in vitro*).

In culturo crosslinking was performed in a custom-build iTRIC (*in culturo* temperature regulated interaction crosslinker) device, allowing the UV irradiation of the culture, while active water-cooling maintains a constant temperature (Bohnsack et al., 2012). The crosslinking was performed in minimal medium lacking the aromatic amino acids tryptophan and phenylalanine. In total 1.6 J/cm² of UV light was applied, corresponding to 5-30 min of irradiation.

In vivo crosslinking was performed with intact cells, resuspended in ice-cold phosphate buffered saline (PBS) buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄) and spread in Petri dishes. A Stratalinker 1800 (Agilent Technologies) was used to apply 1.6 J/cm² of UV light.

In vitro crosslinking was performed after elution of the RNA-protein complexes from the IgG sepharose beads with a Stratalinker 1800 (Agilent Technologies). The samples were spread on a Petri dish and kept on ice during irradiation with 0.4 J/cm² UV light.

2.3.3 Extract preparation and affinity purification

CRAC analysis was performed with 0.5-1 l of exponentially growing yeast culture with an optical density (OD₆₀₀) of 0.5-0.8. Extraction of the cells was performed in 1.5 ml Lysis Buffer (50 mM Tris/HCl pH 7.6, 150 mM NaCl, 1.5 mM MgCl₂, 0.1% (v/v) Nonidet P-40, 5 mM β-mercaptoethanol) by vortexing 5 min at 4 °C with 3 ml Zirkonia

beads (0.5 mm; Thistle Scientific). The lysate was filled-up with 3 ml Lysis Buffer per gram of yeast cells and cleared by two centrifugation steps (4,500 g, 20 min, 4 °C, and 20,000 g, 20 min, 4 °C). Complexes containing tagged proteins were purified by incubation on IgG sepharose beads for at least 30 min at 4 °C. Beads were washed three times with 10 ml Lysis Buffer and bound complexes were eluted in 500 µl Lysis Buffer with GST-tagged tobacco etch virus (TEV) protease for 2 h at 18 °C. RNase trimming of crosslinked RNA was performed with 1 U RNase-IT RNase A and RNase T1 mixture (Agilent Technologies) for 5 min at 37 °C. RNase digestion was terminated by addition of 0.4 g guanidine hydrochloride (final concentration of 6 M), 300 mM sodium chloride and 10 mM imidazole. For a second purification step under denaturing conditions, samples were incubated over night at 4 °C with 50 µl of pre-equilibrated nickel-NTA beads (Qiagen). Beads were washed twice with 500 µl Wash Buffer (50 mM Tris/HCl pH 7.6, 300 mM NaCl, 10 mM imidazole, 6 M guanidinium-HCl, 0.1% (v/v) NP-40, 5 mM β-mercaptoethanol) and three times with PNK Buffer (50 mM Tris/HCl pH 7.6, 10 mM MgCl₂, 0.5% (v/v) Nonidet P-40, 10 mM β-mercaptoethanol). To remove the terminal 3'-phosphate, a product of the RNase trimming step, the beads were resuspended in 80 µl of PNK Buffer and incubated for 30 min at 37 °C with 8 U of thermo sensitive alkaline phosphatase (TSAP, Promega) and 80 units of RNasin (Promega). The reaction was stopped by washing with 500 µl wash buffer and three times with 500 µl PNK Buffer. For 3'-adapter (Table 4) ligation, beads were resuspended in 80 µl of PNK Buffer containing (80 pmol) of 5' adenylated (App) and 3' blocked (ddC) DNA oligonucleotides, 20% (v/v) polyethylene glycol (PEG) 4000, 80 U RNasin, 40 U T4 RNA ligase 2 (K227Q, New England Biolabs) and incubated for at least 5 h at 25 °C. The reaction was terminated by washing twice with 500 µl Wash Buffer followed by three washing steps with 500 µl PNK Buffer. The crosslinked RNA was 5' radiolabelled using 40 µCi [γ -³²P] ATP 80 U of RNasin and 20 U T4 polynucleotide kinase (PNK, Sigma Aldrich) in 80ml PNK Buffer for 40 min at 37 °C. The reaction was allowed to proceed for 20 min after addition of cold ATP (final concentration: 1 mM). The reaction was stopped by washing with 500 µl Wash Buffer and three times with 500 µl PNK Buffer. Ligation of the 5'-adapter (Table 4) was performed in 80 µl PNK Buffer containing 80 U of RNasin, 40 U T4 RNA ligase, 100 pmol 5'-adapter and 1 mM ATP at 16 °C over night. The reaction was terminated by washing with 500 µl Wash Buffer followed by

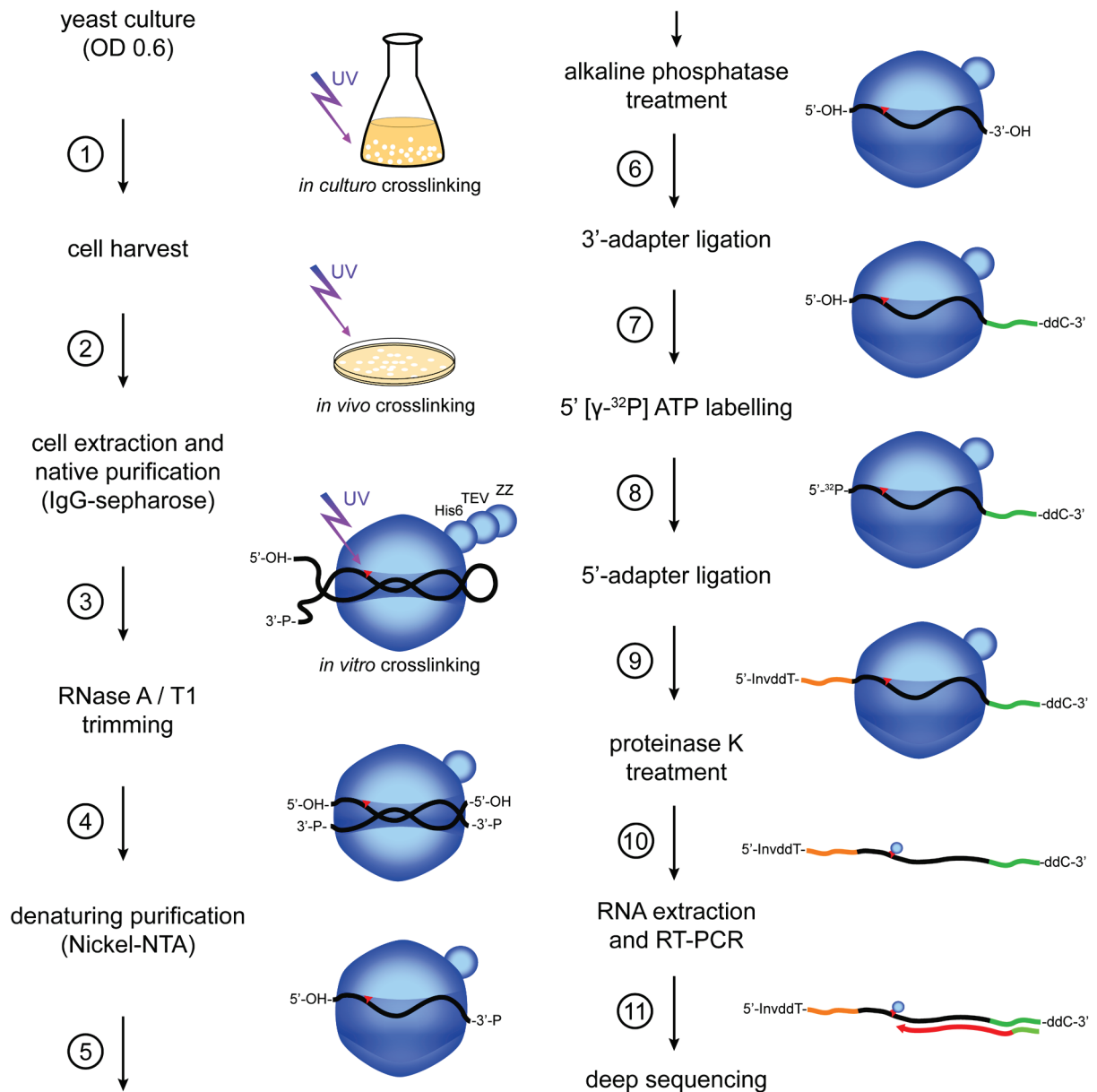


Figure 5: The CRAC method. The scheme shows the main steps of the CRAC technique. The protein of interest carries a His₆-Tev-Protein A (HTP) tag and is depicted in blue. The bound RNA, shown in black, is crosslinked either by UV irradiation of cultures (*in culturo*), on pelleted and resuspended cells (*in vivo*) or after elution from IgG sepharose by TEV protease (*in vitro*). The crosslink, indicated in red, prevents dissociation of the RNA from the bait protein during the second, denaturing purification step on Nickel-NTA beads. The 5' adenylated (App) and 3' blocked (ddC) 3'-adapter, illustrated in green is ligated after alkaline phosphatase treatment. The 5' protected (InvddT) 5'-adapter is coloured orange and is ligated before digestion with proteinase K. A primer depicted in light green, is used for synthesis of cDNA (red) by reverse transcription.

three washing steps with 500 μ l PNK Buffer. Radiolabelled protein-RNA complexes were eluted from the Nickel-NTA matrix with Elution Buffer (50 mM Tris/HCl pH 7.6, 50 mM NaCl, 200 mM imidazole, 0.1% (v/v) Nonidet P-40, 5 mM β -mercaptoethanol), subsequently precipitated with 20% (w/v) TCA, washed with 100% acetone and dried. The samples were resolved on Bis-Tris NuPAGE gels (Invitrogen), transferred to a Hybond C nitrocellulose membrane (GE Healthcare) by electrophoretic western blotting and visualised by autoradiography. Blot areas containing signals, which corresponded to the size of the protein of interest were excised and associated RNA was eluted from the membrane slices with proteinase K solution (100 μ g proteinase K, 50 mM Tris/HCl pH 7.6, 50 mM NaCl, 10 mM imidazole, 0.1% (v/v) Nonidet P-40, 5 mM β -mercaptoethanol 1% (w/v) SDS and 5 mM EDTA). The RNA was further purified with phenol-chloroform and precipitated with ethanol. Reverse transcription (RT) of the RNA fragments was performed using an oligonucleotide primer complementary to the 3'-adapter sequence (Table 4) and the Super Script III reverse transcriptase system (Invitrogen), which was used according to manufacturers instructions. For amplification of the cDNA a PCR reaction containing 1 μ l of cDNA 2.5 U LA Taq Polymerase (TaKaRa), 1x LA TaqPCR buffer, 0.2 μ M of each PCR primer (Table 4), and 0.25 mM deoxyribonucleotides (dNTPs). The PCR products were separated on a 3% (w/v) Methaphor agarose gels (Lonza), fragments of 10-50 base pairs length were excised from the gel and slices were extracted using the MiniElute reaction cleanup kit (Qiagen). Elutes were directly send for single-end deep sequencing (Illumina).

Table 4: Oligonucleotides used for CRAC analysis

Target	Sequence (5'-3')
CRAC 3'-Linker	rAppTCGTATGCCGTCTTCTGCTTGT/ddC/
CRAC 5'-Linker	InvddT/GTTCAGAGUUCUACAGUCCGACGAUC
CRAC RT	CAAGCAGAAGACGGCATAAC
CRAC Forw	AATGATACTGCGACCACCGACAGGTTTCAGAGTTCTACAGTCCGA
CRAC Rev	CAAGCAGAAGACGGCATAACGA

2.3.4 Bioinformatics and data analysis

Genome wide mapping of sequence reads was performed in collaboration with the group of Prof. Dr. Enrico Schleiff (Goethe-University Frankfurt am Main). The CRAC-Pipe software used is described in detail by Simm (2013).

Graphical presentation of mapped sequencing reads on pre-rRNA and snoRNAs was performed with custom-built scripts written in python. Crosslinking profiles were generated using the python (version 2.7) and the python 2D plotting library matplotlib (Hunter, 2007). Mapping of the crosslinking reads on the 2D structures of mature rRNA (Petrov et al., 2013) was performed using python (version 2.7) and vector graphics downloaded from: <http://apollo.chemistry.gatech.edu/RibosomeGallery>. Mapping on the mature 3D rRNA structure (Ben-Shem et al., 2010; Ben-Shem et al., 2011) was performed using PDB entries 3O2Z, 3O58, 3U5B, 3U5C, 3U5D and 3U5E, python (version 2.7) scripts and pymol (version 1.6).

2.4 Identification of chimeric reads by CLASH and prediction of snoRNA basepairing sites

Identification of chimeric sequencing reads was performed by CLASH (Crosslinking, Ligation And Sequencing of Hybrids) as described by Kudla et al. (2011). In short, the deep sequencing reads obtained by CRAC were analysed with the Hyb-pipeline (Travis et al., 2014) using Flexbar (version 2.4) for adapter removal (arguments: trim=0, filter=0, minimum=4, length=17) and Blastall (version 2.2.26) for mapping of the sequences (arguments: word=11 eval=0.01).

Computational prediction of snoRNA-rRNA interactions was performed with a custom-built python (version 2.7) script, which enabled the screening for basepairing interactions, by splitting the RNA sequences into 20 nucleotides long consecutive fragments with 15 nucleotides overlap. The individual fragments were mutually tested for their basepairing properties using the RNA duplex algorithm (Vienna RNA package; Lorenz et al. 2011). The computed Gibbs energy (ΔG) was used to detect highly stable basepairs, with a minimum free energy of -20 kcal/mol. The clustering of chimeric sequence reads obtained by CLASH and computational predictions was performed with python (version 2.7) and the python 2D plotting library matplotlib (Hunter, 2007).

2.5 Purification of pre-ribosomal complexes

For purification of pre-ribosomal complexes in presence and absence of Rok1 a yeast strain, with a tetracycline repressible promoter in front of the Rok1 gene locus and C-terminal HTP tagged Enp1 was used (YMB870; Table 3). 1 l of cells expressing C-terminally HTP tagged Enp1 were harvested in mid-log phase (OD_{600} 0.6) washed with ice cold water and resuspended in 1 ml Lysis Buffer (20 mM Hepes pH 8.0, 75 mM KCl, 1.5 mM $MgCl_2$, 5 mM vanadyl ribonucleoside complexes (New England Biolabs), and 1 tablet Complete Mini protease inhibitor mix (Roche) per 10 ml buffer volume) per gram of yeast cells. Lysing of cells was accomplished by grinding in liquid nitrogen. 1 ml Lysis Buffer per gram of yeast cells was added to the lysates before pre-clearance of the suspension by centrifugation (12 min, 20,000 g). Enp1 containing pre-ribosomal complexes were purified for 2 h with 150 μ l of IgG sepharose in a single step pulldown. Beads were washed 6 times in 5 ml of Washing Buffer (20 mM Hepes pH 8.0, 75 mM KCl, 1.5 mM $MgCl_2$, 5 mM, 0.2 mM EDTA pH 8.0). Copurified pre-rRNA was eluted from the IgG sepharose beads with hot phenol chloroform extraction (2.1.2) according to Rio et al. (2011) and subsequently analysed by Northern blotting (2.1.4) using methylene blue staining (Sambrook and Russell, 2001) for visualisation of mature rRNA. Various low abundant pre-rRNA intermediates were detected by Northern blotting (2.1.4) using a radioactively 5'-terminal labelled hybridisation probe complementary to the 5'-end of ITS1 (Table 2). The protein composition of the purified pre-ribosomal complexes was analysed by SDS gel electrophoresis (2.1.3) followed by Western blotting (2.1.4) using anti-HA (Rok1; Sigma-Aldrich), anti- PAP (Enp1; Sigma-Aldrich), and anti-GAPDH (Abcam) antibodies.

2.6 *In vivo* structure probing of pre-rRNA

The dimethyl sulphate (DMS) *in vivo* structure probing was performed as previously described by Swiatkowska et al. (2012). In short, 1-2 l YMB870 (Table 3) yeast culture expressing C-terminal HTP tagged Enp1 but depleted of Rok1 was pelleted and resuspended in 25 ml of YPD medium. Chemical modification of RNA was performed at 30 °C for 4 min by addition of 1% (v/v). DMS the reaction was quenched by addition of 11 ml of 2 M β -mercaptoethanol and 12.5 ml 100% isoamylalcohol. A second quenching step was performed with 20 ml of 0.7 M

β -mercaptoethanol followed by washing with 20 ml of ice cold water. The purification of pre-ribosomal complexes and extraction of pre-RNA intermediates was performed as described above (2.5). DMS induced modifications in the RNA were detected by primer extension using a radioactively 5'-terminal labelled primer (5'-CGTCCTTGGCAAATGC-3'), complementary to nucleotides 942-957 of the 18S rRNA and Superscript III reverse transcriptase (Invitrogen). The labelling reaction of the oligonucleotide primer (5 μ M) was performed using, 50 μ Ci of [γ - 32 P] ATP and T4 polynucleotide kinase (Thermo Scientific) in a total reaction volume of 40 μ l. The labelled primer was separated from free [γ - 32 P] ATP by gel filtration in a micro centrifuge (1,000 g, 4 min, room temperature), using mini Quick Spin DNA Columns (Roche). 0.125 μ M of the radiolabelled oligonucleotide primer was annealed to 1 μ g of template pre-RNA in a total volume of 10 μ l by a stepwise reduction of the incubation temperature (5 min 65 $^{\circ}$ C, 10 min 37 $^{\circ}$ C and 10 min 4 $^{\circ}$ C). The reverse transcription reaction mix contained 10 μ l Annealing Mix, 1 μ l RNasin (Promega), 100 U Superscript III (Invitrogen), 1x Superscript III reaction buffer, 0.5 mM deoxynucleoside triphosphates (dNTPs), 5 mM dithiothreitol (DTT). The cDNA obtained by the primer extension reaction was precipitated by phenol-chloroform extraction (2.1.2) and resolved on a 10% (v/v) denaturing polyacrylamide, sequencing gel (2.1.3). DMS induced stops were visualised using a phosphorimager. A sequencing ladder was produced with the Sequenase Version 2.0 DNA sequencing kit (Affymetrix) using a plasmid-encoded copy of the RDN37 locus (pAV162; Table 1) as template and the radiolabelled oligonucleotide primer used for the primer extension reaction.

2.7 qPCR based analysis of snoRNA levels in pre-ribosomal complexes

The analysis of snoRNA levels bound to pre-ribosomal particles was performed with Prp43 depletion strain YMB458, carrying a tetracycline repressible promoter in front of the Prp43 gene locus. The YMB279 strain, which only contains the tetracycline repressor gene of the TetO₇ system, served as control (Table 3). The experimental procedure was previously described in Bohnsack et al. (2008). Briefly, cells were depleted of Prp43 protein for 6 hours, harvested by centrifugation. The pellet was resuspended in 1 ml Lysis Buffer (50 mM Tris/HCl pH 7.5, 100 mM NaCl, 5 mM MgCl₂, 1 mM DTT, 5 mM vanadyl ribonucleoside complexes (VRC; New England

Biolabs), and 1 tablet Complete Mini protease inhibitor mix (Roche) per 10 ml buffer volume) per gram yeast cells and lysed by grinding in liquid nitrogen (Kos and Tollervey, 2005). The cell powder was thawed, 1 ml Lysis Buffer per gram of cell pellet was added and the lysate was pre-cleared by centrifugation (20,000 g 12 min, 4 °C). The supernatant was loaded on top of a 12 ml sucrose gradient (10-45% (w/v) sucrose, 50 mM Tris/HCl pH 7.5, 100 mM NaCl, 5 mM MgCl₂, 1 mM DTT) and centrifuged (23,500 rpm, 16 h, 4 °C) in a TST41.14 rotor (Sorvall). The gradients were produced using a Gradient Master apparatus (Biocomp). Gradients were fractionated into 530 µl aliquots and RNA was extracted as described in chapter 2.1.2. Fractions containing free and ribosome bound RNA were pooled, and polyadenylated for 2 h at 37 °C (Figure 6). The reaction mixture contained 5-10 µg total RNA, 8 U poly(A) polymerase (Life Technologies), 2.5 mM MnCl₂ and 5 mM (free pool) or 2.5 mM (ribosome-bound pool) ATP (Roche). Small polyadenylated RNAs were purified using the MirVana miRNA isolation kit (Life Technologies) according to manufacturers instructions. The purified RNA was eluted in 40 µl water

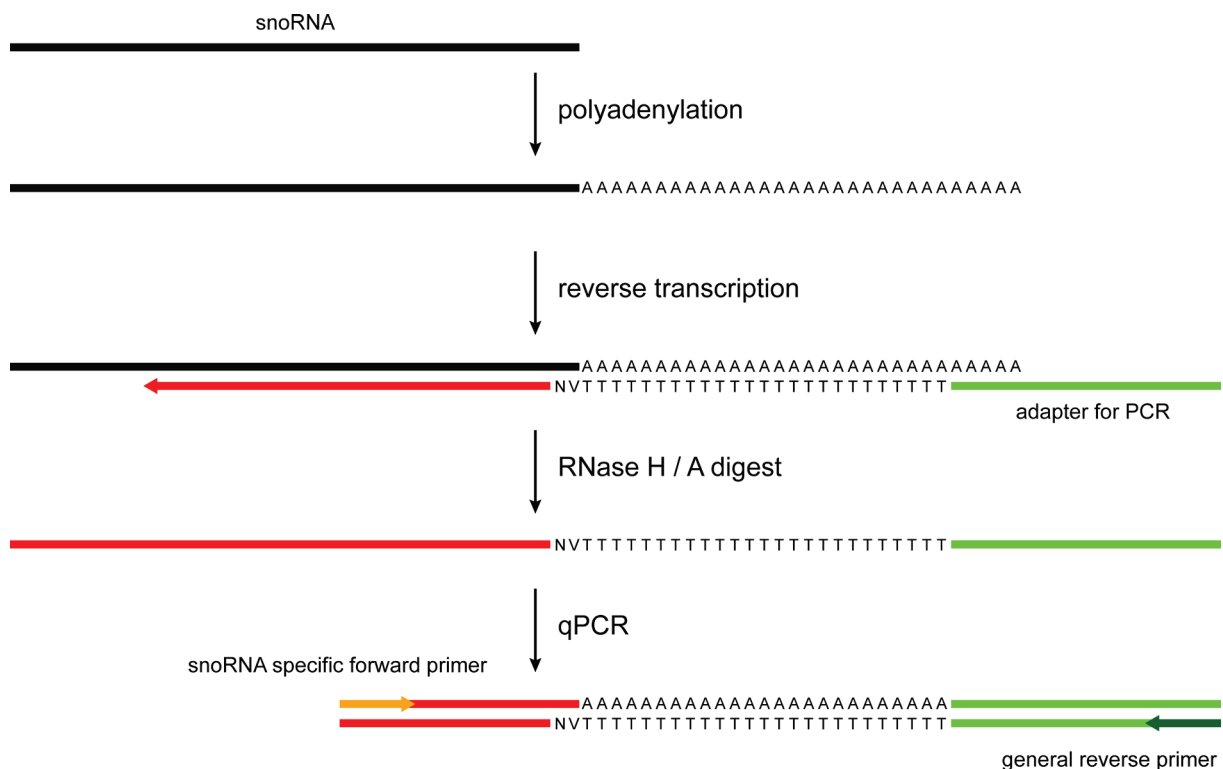


Figure 6: Preparation of RNA for determination of snoRNA levels by qPCR. The scheme was adapted from (Bohnsack et al., 2008). SnoRNAs shown as black bars were polyadenylated and reverse transcribed using an oligo(dT) adapter (Table 5) coloured in green. The RNA was digested with RNase H and RNase A and levels of snoRNAs were determined by qPCR using a snoRNA specific forward primer (orange) and a general reverse primer (dark green), complementary to the oligo(dT) adapter sequence.

and subsequently reverse transcribed, using a standard anchored oligo(dT) linker (Table 5). The Annealing Mix contained 12.5 µg (free pool) or 7.5 µg (ribosome bound-pool) of the linker and the 40 µl of the polyadenylated and purified RNA samples. The annealing was performed by incubation for 10 min at 65 °C followed by 5 min on ice. The reverse transcription reaction was performed for 1 h at 50 °C and contained the Annealing Mix, 2.5 µl RNasin (Promega), 1000 U Superscript III reverse transcriptase (Invitrogen), 1x Superscript III reaction buffer, 1 mM deoxynucleoside triphosphates (dNTPs) 5 mM dithiothreitol (DTT) in a total volume of 100 µl. The reaction was terminated by incubation at 70 °C for 15 min. The RNA was digested at 37 °C for 2 h with 25 U RNase H (New England Biolabs) and 10 U RNase A. The reaction was stopped by incubation at 65 °C for 20 min. The cDNA was purified using the QIAquick PCR purification kit (Qiagen), according to manufacturers instructions and total cDNA concentration of the samples processed in parallel was adapted. Relative levels of all 75 snoRNAs from *S. cerevisiae* were determined using quantitative polymerase chain reaction (qPCR) performed with the Mx3000P qPCR system (Agilent Technologies; Waldbronn, GER), a universal reverse primer and snoRNA specific forward primers (Table 5). Data were analysed using the ddCt-method as described by Bohnsack et al. (2008), and python (version 2.7) scripts. In short, individual snoRNA cycle threshold (Ct) values of both pools were normalised to the median of all 75 measured snoRNA Ct values and to the Ct value of undepleted wild type sample (YMB279) processed in parallel. For data presentation the ratio of ribosome bound to free snoRNA was calculated and plotted. Effects of Prp43 depletion on the total cellular levels of selected snoRNAs were tested by PAA gel electrophoresis followed by Northern blotting (2.1.4), using probes complementary to specific snoRNAs (Table 2).

Table 5: qPCR primer

Target	Sequence (5'-3')
RTQ primer	CGAATTCTAGAGCTCGAGGCAGGCGACATGGCTGGCTAGTTAA GCTTGGTACCGAGCTCGGATCCACTAGTCCTTTTTTTTTTTTTTTTT TTTTTTTTTVN
General reverse primer	CGAATTCTAGAGCTCGAGGCAGG
U18	TGACAAAAGAGATGTGGTTGAC
U24	GAGACATACCAATTATCACCAAG
snR3	GTTTTGATTAGCTGAATGAGAC
snR4	CCTTTATAGCGGTGCTTTAAC
snR5	ATTGGTTCGCTCTAGGTGTAC
snR8	ATCGGTACTGCGGAGTGAG
snR9	ACCTATCATTAGTCCTTCAGAC
snR102	TCTGTCGTCTGTTTTTAGCAG
snR112	AAGAAAGTGAGTGGATCTTCCC
snR13	GTGTGGAAAACTCAAGCTAC
snR161	TCAGGCTGTATTTTCATAACACTAC
snR17ab	GGTACAAATGGCAGTCTGAC
snR30	CATTTGGGTAACCATACTG
snR31	TAAACACCTGATACAGTTGGTC
snR32	GAAATGAGATATTGGGAATCAG
snR33	AATTGATATAGAAGTGTGTGGAC
snR34	TGTCTCAAACGAGGCGATAG
snR35	CAAGGGCTGGTAGGACAGAC
snR36	GCTATTTTTATCTCACGGTATC
snR37	CCTAAGCGACTCTTCTTCATG
snR382	CTGAATGGGTAATAATAGGTAACC
snR39b	ATGTTGTCAACTTAAATTACACC
snR39	GCTGTCGTAATTATCACCATC
snR401	AAGTTTCAGCCTTGTATGAG
snR412	CCTTTTTTCGTTAAGTTTCAG
snR42	GTTAAGCGACCCATGAAATG
snR43	AATCTCTGGGTTGTTTAGATG
snR44	TCCATTACCGTTTACTTTTCC
snR45	GCAACCCATTGATCTTGTTAC
snR46	ATAAAGTTGTGCTATTTCCATG
snR471	ACAATAGCTTTTTAACACTG
snR482	ATGTTAGGATGTGAAGTTAAGTAC
snR492	AAGATTTATCTTTTTGTCCATC
snR50	CTTTACAGAACCGCTACACTG

snR512	TAGATTGGTCTCTTTAACGAAGG
snR52	GACATTAGCGTGAACAATCTC
snR53	TGATTAAAATTGTTGTTTACGC
snR54	CGATCTTGTAGAGAACTTTTACTC
snR55	CACAATCGTCTTTTTTTTTATCC
snR562	ACACAGACCTGTAAGAACTTTTC
snR57	ATTTTTCTGAGGAAGTATATGC
snR58	CTCCTATGGAAGAGAACTC
snR593	AATCACCATCTTTCGGCTGAC
snR60	CAACTGATTGAACATACTATCG
snR61	AAGATAACCAATTTTACCAAAG
snR62	ATGGAAGATATACGACTATCAAC
snR632	GAGTCTTTTAATGTGATGAGTGG
snR64	TAAAGCCCAGTTTTTAGTAGAG
snR65	AACTTTATGATTACAGTGTTTTTC
snR66	GAGATTGCTTTTTTATTACTGAC
snR67	GATTTTACAAACAACAACACTG
snR682	CGAGGAAATTGACTCTTAACAG
snR69	AAAGGGAGAAGATTTTTTTGTC
snR70	TGATTGGTCACAAGACATCTG
snR71	ATTTCTATCCAACATTCATC
snR72	GAGAACATCAATGAAGAAAACG
snR73	TGTGACAGGGCGTGGTACTG
snR74	GAAACAAATTACTCAAATAGACAAG
snR75	AACTATTTAAATTACCATTTCATGC
snR76	TCCTTTCAAATGAGTGACAATG
snR77	GAGTATATGTTGATACGTTTTTGC
snR78	AGTTTCTGAATCTTTTGTGATTAG
snR791	CAAGACTACAACGGTATCTG
snR80	GGTTATATTAGTCCATTTTCATAGC
snR81	ACGCTTTTTCACATCTTCTTG
snR82	TATAGTTTGATAGTAGATGGGCG
snR83	CCCGATTTGTATTTTATTTTTTC
snR84	AGGAACATGACTCAAAGAGACAC
snR85	TACCATATAGAGGTGTCAAGTACAC
snR86	AGAATGTAGTTTCATACCCG
snR87	TGTTCTATATGGGTGATTAGCG
snR1891	ACTTTCAAGTACTTCACACG
snR1902	GAAAAGATGTTGCTTCTGTGAC
snR191	TATGTTCGTAGTAAAGATCCTCAC

2.8 Nob1 *in vitro* cleavage assay

In vitro endonuclease assays were performed as previously described by Pertschy et al. (2009). RNA substrates were transcribed *in vitro* using T7 RNA polymerase (Thermo Scientific) and linearised plasmid DNA as a template (Table 1). The reaction mixture contained 100 U T7 RNA polymerase (Thermo Scientific), 1x T7 RNA polymerase reaction buffer, 5 µg of linearised plasmid DNA (EcoRI), 0.5 mM of each ribonucleoside triphosphate (Roche), 40 U RNase inhibitor (Thermo Scientific), 5 mM dithioereithol (DTT) and 2 mg/ml bovine serum albumin (BSA) in a total volume of 50 µl. The reaction was performed for 2 h at 37 °C and terminated by phenol-chloroform extraction (2.1.2). The transcripts were radiolabelled either internally by addition of 25 µCi [α -³²P] GTP nucleotides during transcription or at the 5'-end using 25 µCi of [γ -³²P] ATP and T4 polynucleotide kinase (Thermo Scientific) according to manufacturers instructions.

The cleavage reaction was performed with PhNob1 and AtNob1 protein. Both were purified as described in Veith et al. (2012a) and Missbach et al. (2013) in the group of Enrico Schleiff (Goethe University, Frankfurt am Main). The cleavage reaction was performed for 1 h at 30 °C in a total volume of 10 µl containing 0.5 µl of substrate RNA and Reaction Buffer (25 mM Tris/HCl pH 7.6, 75 mM NaCl, 5 mM MnCl₂, 8 U RNAsin (Thermo Scientific), 2 mM DTT, 4.5% (v/v) glycerol, 100 µg/ml BSA (Thermo Scientific), 0.05% (v/v) Nonidet P-40, 3.5 µM *E. coli* tRNA) The Nob1 proteins were used in concentrations as specified (Figure 27; Figure 28; Figure 29; Figure 30). The reaction was terminated for 30 min at 37 °C with proteinase K solution (100 µg proteinase K, 50 mM Tris/HCl pH 7.5, 0.1% (w/v) SDS). The RNA was extracted using phenol-chloroform extraction (2.1.2). The RNA pellets were dissolved in 4 µl water and 6 µl Formamide Loading Buffer (2.1.3) and resolved on a 8% (v/v) denaturing polyacrylamide, sequencing gel and visualised using a phosphorimager. The ends of cleavage products were determined by comparison to a sequencing ladder, which was produced with the T7 sequencing kit (Affymetrix) according to manufacturers instructions. The sequencing reaction was performed with the plasmid-encoded template DNA (Table 1) that was used for *in vitro* transcription, and an oligonucleotide complementary to the T7 promoter (5'-CTATAGGGAGA CAAGCTTAAGTC-3'). *E. coli* cell lysates applied for degradation controls were prepared as, described in Veith et al. (2012a).

3 Results

3.1 Functional analysis of ribosome biogenesis cofactors by *in vivo* UV crosslinking

In *S. cerevisiae* 19 RNA helicases are involved in the maturation of the ribosomal subunits. Preliminary studies have allowed a basic functional classification of most of these helicases by analysis of pre-rRNA processing defects induced upon their deletion or depletion (reviewed in Martin et al., 2013; Rodriguez-Galan et al., 2013). This has led to the identification of several RNA helicases that function during the maturation of the LSU, others that participate in biogenesis of the SSU and two that are implicated in both pathways. However, the target sites on the rRNA are still elusive for most RNA helicases involved in ribosome biogenesis and a detailed functional characterisation of these proteins was therefore hardly feasible. To address this and gain a more detailed understanding of RNA helicase functions in the maturation of ribosomal particles, the DEAD box RNA helicase Rok1, which is involved in the early steps of SSU maturation and the DEAH box RNA helicase Prp43, which participates in maturation of both subunits, were subjected to the recently developed CRAC (Crosslinking and analysis of cDNA) analysis (Bohnsack et al., 2012). This method is based on UV induced crosslinking, subsequent purification of the protein-RNA-complexes under denaturing conditions, followed by reverse transcription of the associated RNA fragments and analysis of the cDNA-library by Illumina deep sequencing (2.3).

3.2 Crosslinking experiments of Rok1 reveal high affinity for rRNA

An essential prerequisite for the identification of protein crosslinking sites on RNAs by CRAC is the fusion of a His6-TEV-Protein A-tag (HTP-tag) to the coding sequence of the protein of interest (2.3.1). The gene sequence of the Rok1 RNA helicase was cloned into a derivative of the pRS415 vector also encoding a C-terminal HTP-tag. The plasmid was transformed into yeast strains, in which expression of the endogenous Rok1 protein is controllable by a tetracycline repressible promoter. Depletion of endogenous Rok1 and the constitutive expression of plasmid derived Rok1 results in the presence of predominantly HTP-tagged Rok1 protein within the cells. It is important to note that the HTP-tag did not interfere with the essential Rok1

function since the plasmid encoded Rok1 copy was able to rescue the phenotype of the endogenous Rok1 depletion in a growth assay (Figure 7). In contrast, empty plasmid was not able to complement the loss of Rok1 and resulted in a significantly decreased growth rate of the culture after approximately 10 h.

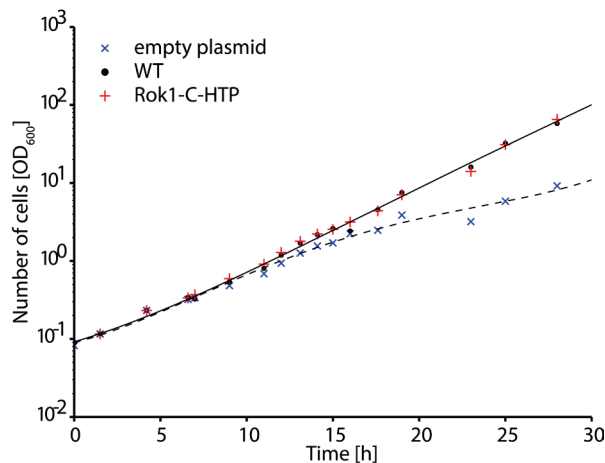


Figure 7: Plasmid encoded Rok1-C-HTP complements the growth phenotype of endogenous depletion. Growth analysis was performed for strains carrying Rok1-C-HTP or empty plasmid (pRS415) after depletion of Rok1 by a tetracycline repressible promoter and compared to the growth rate of wild type cultures. The growth rate was monitored by optical density measurement of liquid cultures.

Crosslinking by UV irradiation was performed after depletion of endogenous Rok1 initially on pelleted living cells *in vivo* for 5 min (2.3.2). To isolate Rok1 containing complexes a pulldown on IgG sepharose was performed (2.3.3). Proteolytic cleavage at the TEV site within the tag enabled the elution of the complexes from the IgG sepharose and allowed a subsequent purification step on a Nickel-NTA matrix. This procedure was performed under denaturing conditions, which ensured that only RNA covalently linked to the bait protein was copurified. RNase trimming of the copurified RNA and ligation of 5'- and 3'-adapters allowed the construction of a cDNA library that was sent for Illumina deep sequencing. Prior to alignment of the obtained sequence reads on the yeast genome the linker sequences were removed *in silico*. The bioinformatic data processing (2.3.4) was performed in collaboration with the group of Prof. Dr. Enrico Schleiff (Simm, 2013). Mapping of aligned sequence reads on the annotated yeast genome allowed the classification of the crosslinked RNAs into functional groups. The percentage of Rok1 crosslinks to individual transcripts, normalised on the number of total mapped reads, is presented in Figure 8 and Table 6. The crosslinking profile observed for Rok1 to total cellular RNA differed significantly in the portion of copurified RNA sequences when compared to the *in vivo* control sample. In particular, and consistent with the known role of Rok1 in

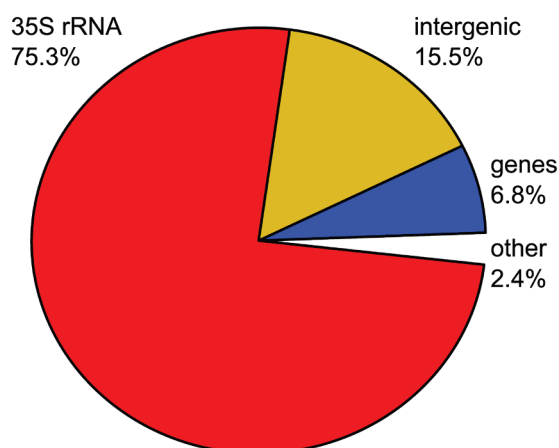


Figure 8: Rok1 crosslinks predominantly to ribosomal RNA. Cells expressing plasmid encoded C-HTP-tagged Rok1 were depleted of endogenous Rok1 and UV crosslinked *in vivo*. RNA copurifying with Rok1 containing complexes, was trimmed and ligation of linkers enabled the construction of a cDNA library, which was analysed by Illumina deep sequencing. Obtained reads were mapped on the *S. cerevisiae* genome and grouped into transcript classes. Relative distribution of Rok1 crosslinking hits for the three most abundant genome feature categories (> 5%) are shown.

Table 6: Genome wide mapping of Rok1 sequencing reads. Upper panel: Mapping statistics for Rok1-C-HTP and control sample. Lower Panel: Counts of crosslinking hits shown for individual transcript classes of both samples were normalised to the total number of mapped reads (per million).

Protein	Rok1 <i>vivo</i>	Cont <i>vivo</i>
mapped reads	27,477,553	4,106,088
discarded reads	6,635,141	2,761,161
unique mapped reads	19,634,843	1,082,670

Features	Hits per million	
35S rRNA	754,162	448,956
Intergenic	155,484	296,175
Genes	67,692	218,662
tRNA	14,578	26,232
5S rRNA	3,962	1,882
snRNA	1,648	3,109
snoRNA (H/ACA)	1,633	3,109
snoRNA (C/D)	756	742
Rep origin	460	492
SRP RNA	350	1,132
Repeat region	299	993
LTR	97	65
RNase P RNA	71	339
Misc feature	42	112
RNase MRP RNA	35	54
Telomerase RNA	20	143
ncRNA	20	354

ribosome biogenesis (Venema et al., 1997), the number of sequences mapping to the 35S transcript was significantly increased.

3.3 Rok1 interacts with the 18S rRNA sequence

To identify the precise Rok1 crosslinking locations within pre-rRNA, the number of crosslinked sequence reads was determined for each individual nucleotide and normalised to the overall amount of mapped reads in the experiment. The resulting hit number was plotted on the linear sequence of the rDNA locus RDN37 (Figure 9 A). Peaks represent clusters of reads and indicate potential Rok1 binding sites. Peaks also present in control samples (Figure S) were disregarded. The most prominent crosslinking sites of Rok1 were found between positions 800 and 1600 of the 35S transcript (Figure 9 A, inset) in the 18S rRNA sequence. To validate the Rok1 specific peaks on 35S rRNA, two additional experiments using a strain carrying an N-terminal HTP-tagged Rok1 copy were performed. In addition to the described *in vivo* crosslinking approach, this strain was also subjected to *in culturo* CRAC analysis (2.3.2). This technique has the advantage of allowing the application of UV light to

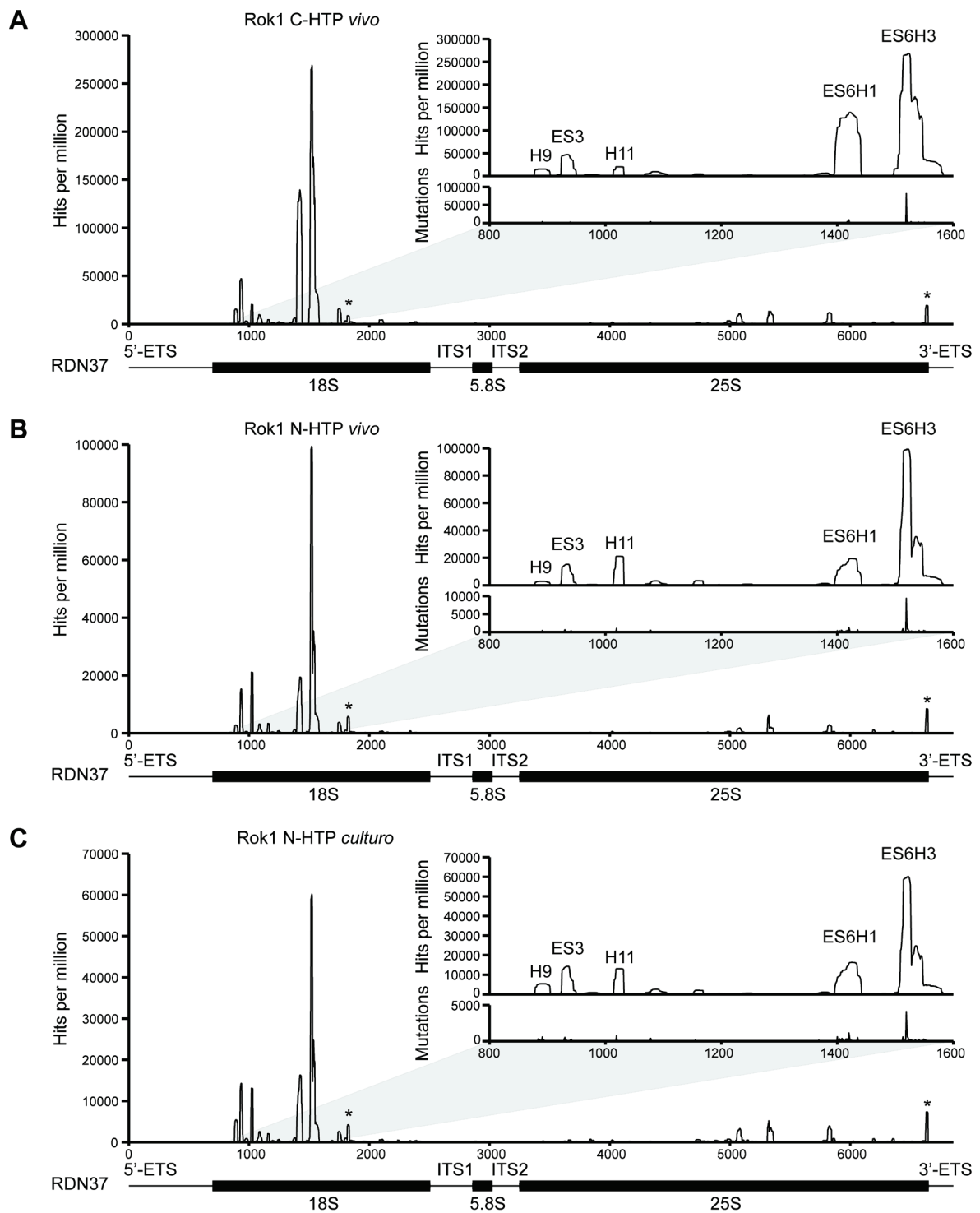


Figure 9: Rok1 crosslinks mainly to the 18S rRNA sequence. Sequence reads obtained by crosslinking of (A) *in vivo* Rok1-C-HTP, (B) *in vivo* Rok1-N-HTP and (C) *in vitro* Rok1-N-HTP were mapped on the linear RDN37 sequence encoding the 35S pre-rRNA. The number of hits was normalised to total mapped reads per million for each nucleotide. Relative positions of the mature 18S, 5.8S and 25S rRNA sequences are indicated below the graphs. Magnifications of the major Rok1 crosslinking sites are provided and their relative positions on the 18S rRNA sequence are indicated (H: helix; ES: expansion segment). The lower panels of the zoom show the position of mutations resulting from errors during reverse transcription and indicate direct crosslinks between nucleotides and amino acids. Peaks also present in the corresponding control experiments (Figure S) are labelled with asterisks.

actively growing yeast cultures rather than pelleted yeast cells. Mapping of the obtained sequence reads on the 35S rRNA locus revealed a nearly identical crosslinking pattern for all three independent experiments (Figure 9 A-C). This suggests that neither the position of the HTP-tag nor the method of UV irradiation significantly influences the crosslinking pattern of Rok1 on pre-rRNA. The analysis of reads obtained by CRAC also enables the detection of single nucleotide substitutions or deletions within the aligned sequences. These mutations arise during reverse transcription and are caused by UV induced covalent attachment of amino acids to the crosslinked RNA (Bohnsack et al., 2012). They indicate positions of direct protein RNA interactions. The accumulation of such mutations at specific sites within the main crosslinking peaks provide additional support to these as specific Rok1 interaction sites (Figure 9, lower panel of zoom).

The analysis of crosslinking sites on the linear 35S rRNA transcript showed that the major crosslinking peaks of Rok1 are located within the 18S rRNA sequence. In order to find potential functional relevance of the crosslinking sites and to visualise the potential interaction sites in a structural context, hit numbers were assigned to the mature secondary structure of 18S rRNA. The main crosslinking sites of Rok1 were located in the central domain of the 18S rRNA sequence within the eukaryotic expansion segment 6 (ES6) on helix 1 (ES6H1) and helix 3 (ES6H3; Figure 10). Further putative Rok1 binding sites reside within in the 5'-domain of 18S, where Rok1 crosslinks to sequences of helix 11 (H11), the eukaryotic expansion segment 3 (ES3) and to helix 9 (H9), which is in the immediate vicinity of ES3. Additionally, weak Rok1 crosslinking in the eukaryotic expansion segment 7 (ES7) was observed.

Although some crosslinking sites of Rok1 are in close proximity, in general, the potential target sites of the helicase are distributed over the 5'- and central domain of the secondary rRNA structure. However, the ability of RNA to form unusual basepairs provides the basis for complex tertiary structures of long RNA molecules (Westhof and Fritsch, 2000) that may allow the clustering of distant sequences in 3D. Since structures of pre-40S complexes are not available, the positions of the Rok1 crosslinking sites were mapped on the mature 3D structure of the 40S ribosome (Ben-Shem et al., 2011). Remarkably, the interaction sites of Rok1 in ES3, ES6, H9 and H11 cluster on one face of the SSU foot region (Figure 11 A). The Rok1 crosslinking site in ES7 is located on the same face but above this main interaction

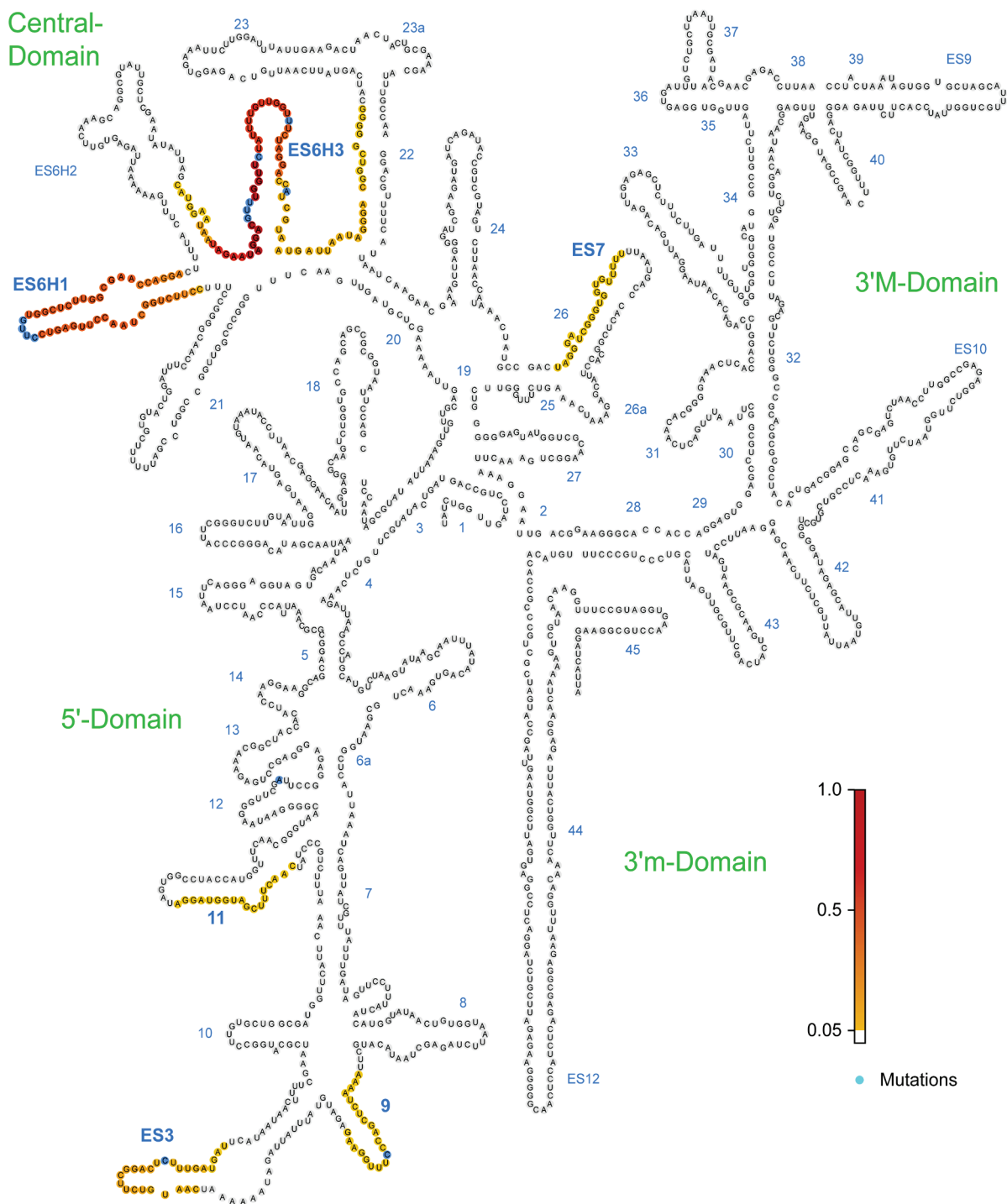


Figure 10: The major Rok1 crosslinking site on 18S rRNA is located within ES6. Rok1-C-HTP crosslinking hits above the threshold of 5% of the highest peak within 35S pre-rRNA were mapped on a secondary structure model of mature 18S rRNA (Petrov et al., 2013). Rok1 crosslinking sites are indicated as coloured circles. The fill colour of the circles indicates the relative hit values as colour gradient from red (100%) to yellow (5%). Blue circles mark positions of mutations that arise from crosslinking induced reverse transcription errors. Helices of the 18S structure are indicated in blue and domains (3'M: 3'-major, 3'm: 3'-minor) are shown in green.

cluster. Interestingly, ES7 is positioned between the ribosomal proteins Rpl13 and Rpl14, which have been reported to be required for Rok1 association with pre-40S complexes (Jakob et al., 2012). In addition, it is noteworthy that folding of the mature 40S structure allows a basepairing interaction of the Rok1 crosslinking sites in ES6H3 and ES3 (Figure 11 B; Alkemar and Nygard, 2003). Taken together, the CRAC analysis of Rok1 revealed potential interactions with rRNA in the 18S sequence. These results are consistent with a role of Rok1 in the process of SSU maturation. Clustering of the crosslinking sites in 3D suggests that the individual crosslinking sites might be the result of a single Rok1 binding site in pre-40S complexes.

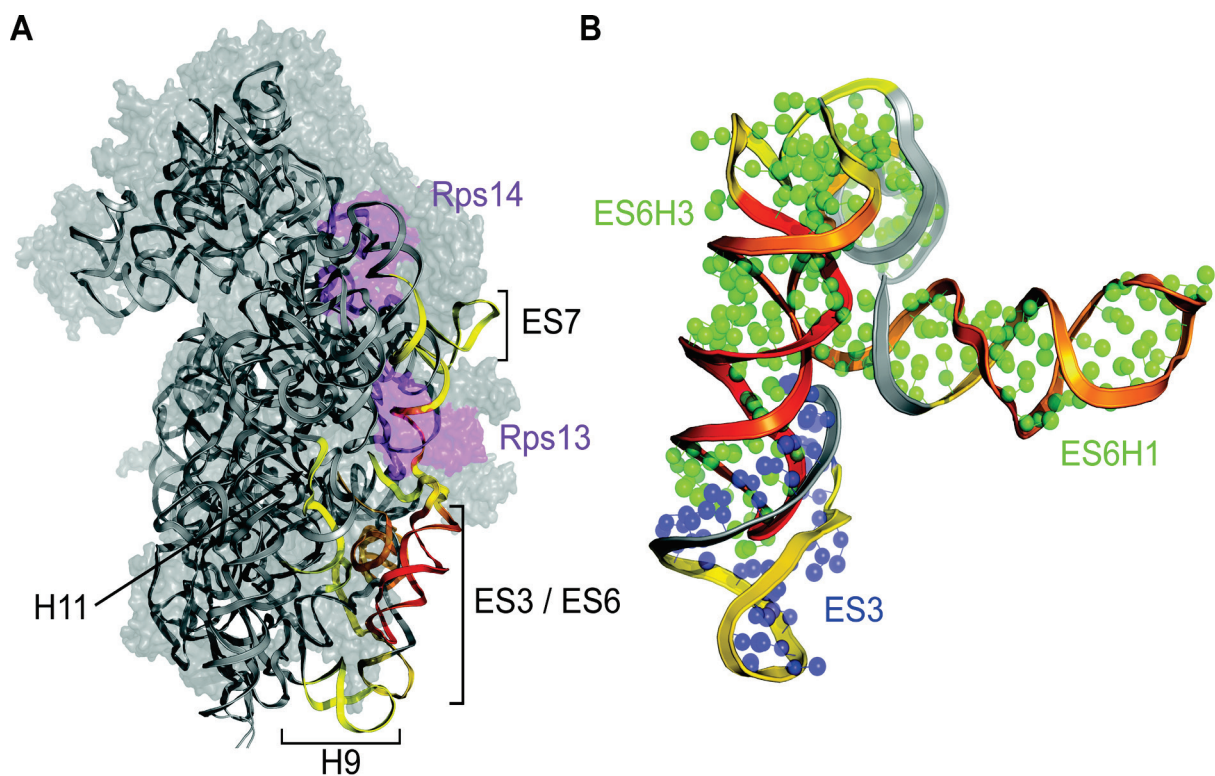


Figure 11: The Rok1 crosslinking sites cluster on the 3D structure of the mature 40S subunit. Rok1-C-HTP crosslinking sites were mapped on the 3D structure of the mature 40S ribosomal subunit (Ben-Shem et al., 2011). **(A)** Rok1 crosslinking sites are visualised as a colour gradient from red (100%) to yellow (5%) on the RNA backbone. Locations of Rok1 crosslinking sites are indicated (H: helix, ES: expansion segment). The surfaces of ribosomal proteins included in the crystal structure are shown in light grey. Ribosomal proteins Rpl13 and Rpl14 are shown in purple. **(B)** Magnified view of the Rok1 crosslinking site in ES6 and ES3. The nucleotide bases of the expansion segments are indicated as coloured spheres ES3 (blue) and ES6 (green). The zoom is presented in side perspective. Rok1 crosslinking sites are presented as in A.

3.4 Structure probing confirms interaction of Rok1 with ES6H3

In order to validate the most prominent crosslinking site of Rok1 on pre-rRNA, an *in vivo* dimethyl sulfate (DMS) structure probing analysis (2.6; Wells et al., 2000) was performed. Application of DMS on actively growing cells induces the transfer of a methyl group from the highly reactive DMS predominantly to the N1 of adenine residues depending on their chemical and steric accessibility within RNA. The resulting DMS modifications are detectable as stops in primer extension reactions. The presence or absence of RNA binding proteins can therefore induce changes in the DMS modification pattern of the RNA sequence.

Since Rok1 crosslinks to sequences that are part of the mature 18S rRNA the purification of Rok1-containing pre-ribosomal complexes after DMS treatment was necessary (2.5). Therefore, the SSU maturation factor Enp1, which is present in early and late pre-ribosomal complexes, was selected for the purification of Rok1 complexes. An initial pulldown experiment with genomically HTP-tagged Enp1 and subsequent analysis of the sample fractions by Western and Northern blotting was performed to confirm the presence of Rok1 and pre-rRNAs, such as 35S, 23S and 20S, in the isolated complexes (Figure 12 A and B). Moreover, it was demonstrated that the elution fractions were substantially depleted of mature rRNA. To detect Rok1 induced alterations in DMS accessibility, chemical probing was performed for cultures expressing Rok1 or depleted of the protein. Rok1 depletion was achieved by addition of tetracycline and resulted in the presence of multiple DMS induced primer extension stops in ES6H3 of Enp1 associated pre-18S rRNA (Figure 12 C). In contrast, presence of Rok1 resulted in a significant decrease in the signal for residues A818, G823, A829 and U832. These four nucleotides are located directly within the centre of the Rok1 crosslinking site. However, residues A803, A809, A811 and A850, which are outside of the main crosslinking site in ES6H3 remained unaffected. Therefore, the structure probing approach confirmed the interaction of Rok1 and ES6H3 identified by CRAC analysis as a bona fide Rok1 binding site.

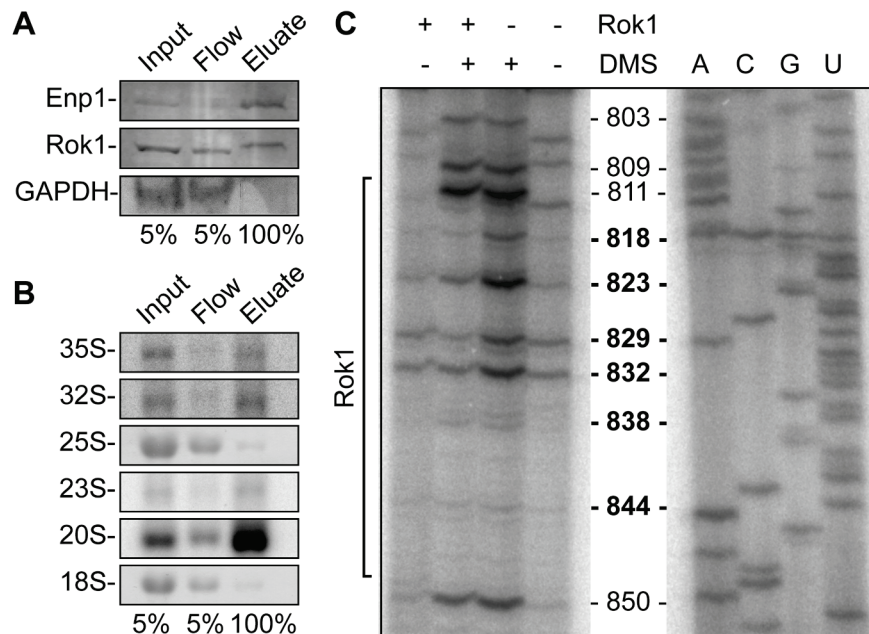


Figure 12: DMS structure probing confirms the Rok1 interaction with ES6H3. Extracts of cells that express HA-tagged Rok1 under the control of a tetracycline repressible promoter and genomically HTP-tagged Enp1 for complex purification, were subjected to co-immunoprecipitation on IgG sepharose. **(A)** Proteins in the eluates were analysed by Western blotting. Antibodies for detection of the HA-tag (Rok1), HTP-tag (Enp1) and endogenous GAPDH were used. **(B)** RNA in the eluate was extracted and analysed by Northern blotting. Pre-rRNAs (35S, 32S, 23S and 20S) were detected with a probe complementary to the 5'-end of ITS1 and visualised with a phosphorimager. The mature rRNAs (18S and 25S) were detected by methylene blue staining. **(C)** Cells expressing or depleted of Rok1 were treated with DMS *in vivo*. RNA that copurifies with HTP-tagged Enp1 was extracted and subjected to primer extension analysis using an oligonucleotide annealed downstream of ES6H3. The reaction products were separated using denaturing polyacrylamide gel electrophoresis and the transcription stops caused by DMS modifications were monitored with a phosphorimager. The indicated nucleotide positions of DMS modifications were determined by comparison to a sequencing ladder shown on the right. Bold numbers depict residues with changed DMS reactivity upon Rok1 depletion. The position of the Rok1 crosslinking site is indicated on the left.

3.5 Rok1 interacts with snR30 and its basepairing sites on pre-rRNA

The RNA helicase Rok1 was previously shown to be required for the release of the H/ACA snoRNA snR30 from pre-ribosomal complexes (Bohnsack et al., 2008). SnR30 is essential for early processing events of pre-rRNA (Bally et al., 1988) and harbours two conserved motifs (m1 and m2), which were shown to establish basepairing contacts with nucleotides in ES6 of the 18S rRNA sequence (Fayet-Lebaron et al., 2009). The corresponding sequences of the rRNA were accordingly named rm1 and rm2. Interestingly, rm2 is located within ES6H3 and lies within the Rok1 binding site identified by the CRAC analysis. The function of Rok1 in snR30 release from pre-ribosomal complexes (Bohnsack et al., 2008) provides additional,

independent evidence for the binding of the helicase to ES6H3. This could suggest a direct role of Rok1 in regulating snR30 basepairing to sequences of ES6 in pre-18S rRNA.

The release of snR30 by Rok1 would require a direct interaction of the helicase with the snoRNA-rRNA duplex. To analyse this further, crosslinking of Rok1 to the snoRNAs involved in maturation of the 40S subunit was analysed (Figure 13 A). As expected, snR30 was highly enriched in all three CRAC experiments. The sequence read distribution on the linear snR30 transcript showed that Rok1 crosslinking signals are distributed over the total length of the snoRNA but accumulate in the 3'-region of the snoRNA (Figure 13 B). The peaks were validated by the occurrence of mutations that arise from nucleotide-amino acid crosslinking events that cause errors during reverse transcription (Figure 13 B, lower panel). Mapping of the crosslinking sites onto the secondary structure of an snR30-ES6 basepairing model (Fayet-Lebaron et al., 2009) revealed that Rok1 crosslinks predominantly to a structurally important internal hairpin and also to the 3'-hairpin of snR30 (Figure 13 C). Interestingly, the 3'-hairpin has been shown to be essential for snR30 function in rRNA processing and harbours the conserved motifs m1 and m2 (Atzorn et al., 2004). Rok1 crosslinks significantly to the m2 motif, which basepairs with the rm2 motif in ES6H3, another Rok1 binding site. These crosslinking data suggest a direct interaction of Rok1 with both the snoRNA and pre-rRNA sequences involved in this basepairing interaction and therefore strongly support a model of direct duplex unwinding by the helicase (Bohnsack et al., 2008). Dissociation of snR30 from ES6H3 frees rm2 and might thereby allow the basepairing of this sequence with ES3, an interaction found in the mature ribosome (Figure 11 B; Alkemar and Nygard, 2003). Possibly, Rok1 is involved in the establishment of this interaction since the helicase crosslinks also to the critical residues for basepair formation in ES3 (Figure 10; Figure 11).

Taken together, the identified crosslinking sites of Rok1 on pre-rRNA and snR30 suggest a simultaneous interaction of the helicase with both substrate RNAs and support a model of Rok1 dependent snR30 release from pre-ribosomal complexes by local strand unwinding. In addition, crosslinking of Rok1 to ES3, which basepairs with ES6H3 in mature ribosomes, could suggest that Rok1 might regulate structural remodelling events during ribosome biogenesis.

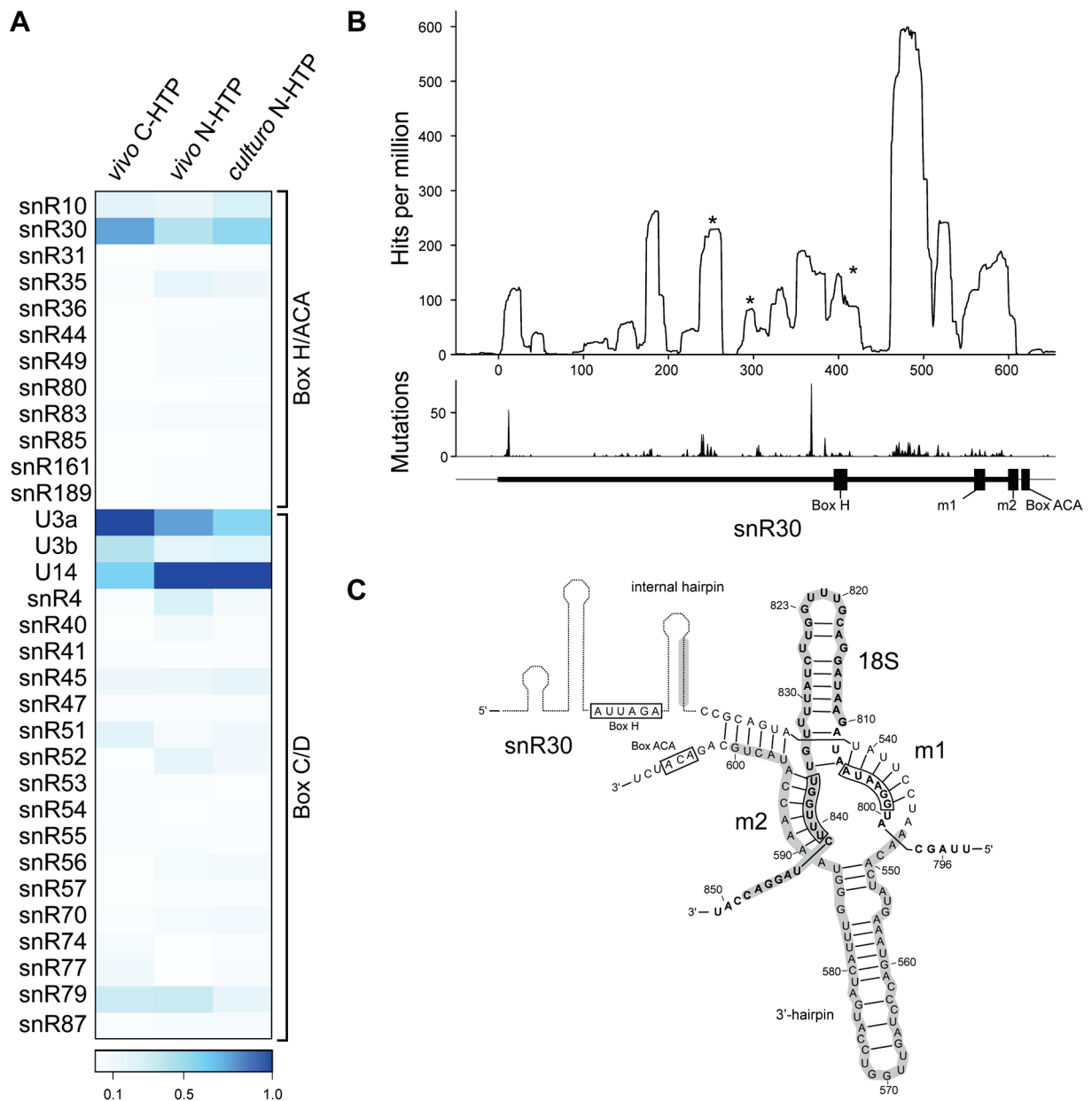


Figure 13: Rok1 crosslinks to snR30 within pre-ribosomal complexes. (A) Total Rok1 reads that map to the individual snoRNAs involved in maturation and modification of the 40S subunit were normalised to the overall amount of snoRNA hits and are presented as colour map. Reference colours for the percentage values are provided as colour bar at the bottom. SnoRNAs were sorted according to their classification (box C/D and box H/ACA) as indicated on the right. **(B)** Number of Rok1 crosslinking hits on the linear genomic snR30 sequence, normalised to total mapped reads per million, is shown. Asterisks mark peaks that were also present in control samples. The position of mutations are presented in the lower panel. The relative positions of the mature snR30 sequence and the location of conserved motifs (Box H, Box ACA, m1 and m2) are indicated by the bar at the bottom. **(C)** Secondary structure model illustrating the potential snR30-ES6 interaction (Fayet-Lebaron et al., 2009). The Rok1 crosslinking sites on both RNAs are shown in grey. Nucleotides of the 18S rRNA sequences are highlighted as bold letters and the residues, which form basepairing interactions, are boxed. Nucleotide positions and conserved motifs of snR30 are indicated.

3.6 Identification of additional snR30 basepairing sites on 18S rRNA

Recent studies using newly developed experimental and computational methods have revealed that the number of basepairing interactions between snoRNA and pre-rRNA is much higher than previously expected (Kudla et al., 2011). Apart from the 3'-hairpin of snR30 a prominent crosslinking peak for Rok1 was found within the internal hairpin of the snoRNA (Figure 13 B). The internal hairpin seems to be structurally important but a specific function for this part of snR30 has not been described so far (Atzorn et al., 2004). However, the strong crosslinking signal of Rok1 at this position raised the possibility that the 5'-hairpin might include additional sequences that can basepair with pre-rRNA.

To test this idea, two independent approaches were used. First, a computational prediction of potential snoRNA-rRNA interactions between snR30 and 18S rRNA was performed. The sequences of the 18S rRNA and snR30 were split into short overlapping fragments and screened for duplex formation using the RNAduplex algorithm from the Vienna RNA Package (Lorenz et al., 2011). The computed Gibbs energy (ΔG), which is a measure for the basepairing stability, was used to identify sequences that form highly stable basepairs (2.4). This sensitive and unbiased approach was complemented by a search for chimeric sequence reads within the experimentally curated Rok1 crosslinking data. For this, the recently published CLASH (Crosslinking, Ligation And Sequencing of Hybrids) analysis was used (Kudla et al., 2011). Such chimeric reads are thought to arise from ligation of two short sequence fragments that copurify with the bait protein and thus form hetero-concatemers or by strand switching of the reverse transcriptase during cDNA synthesis of crosslinked RNA duplexes (Houseley and Tollervey, 2010). The RNA hybrids identified in the CLASH analysis provide strong experimental evidence for snoRNA-rRNA interactions within the binding sites of Rok1 but do not necessarily include the sequences that mediate basepairing. To locate novel basepairing sites of snR30 on 18S-rRNA, the experimentally obtained CLASH data were compared to the predicted interactions, which provide likely basepairing locations of the two RNA molecules. Colocalisation analysis was performed considering the sequence position of interacting sequences within both RNA molecules. Relative positions of predicted basepairing sites and hybrid fragments found in the CLASH analysis were correlated, graphically overlaid and are shown as scatterplot (Figure 14 A). This revealed three

interaction clusters (C1-3) that represent overlaps of hybrids and predictions. The identification of two CLASH hybrids (C1), which are located in close proximity to the known snR30-18S basepairing site in ES6H3, confirmed the effective detection of basepairing sites by the CLASH analysis (Figure 14 B). It should be noted that computational prediction of this known basepairing site was not possible since it is mediated cooperatively by the short m1-rm1 and m2-rm2 duplexes (Fayet-Lebaron et al., 2009), whose stability is below the threshold set in the analysis.

Besides, validation of the known snR30 basepairing site in ES6H3 (C1), two additional putative interaction sites were identified (C2 and C3), both composed of overlapping hybrids and predictions. The potential interactions found in C2 suggest basepairing of ES6H1 to nucleotides at the base of the snR30 5'-hairpin structure and near the conserved box H of the snoRNA (Figure 13 B and C; Figure 14 B). Additionally, both duplex prediction and CLASH analysis identified a highly stable interaction of the main Rok1 crosslinking site on snR30 within the internal hairpin, to ES7 of the 18S rRNA sequence (C3). A basepairing interaction of snR30 and

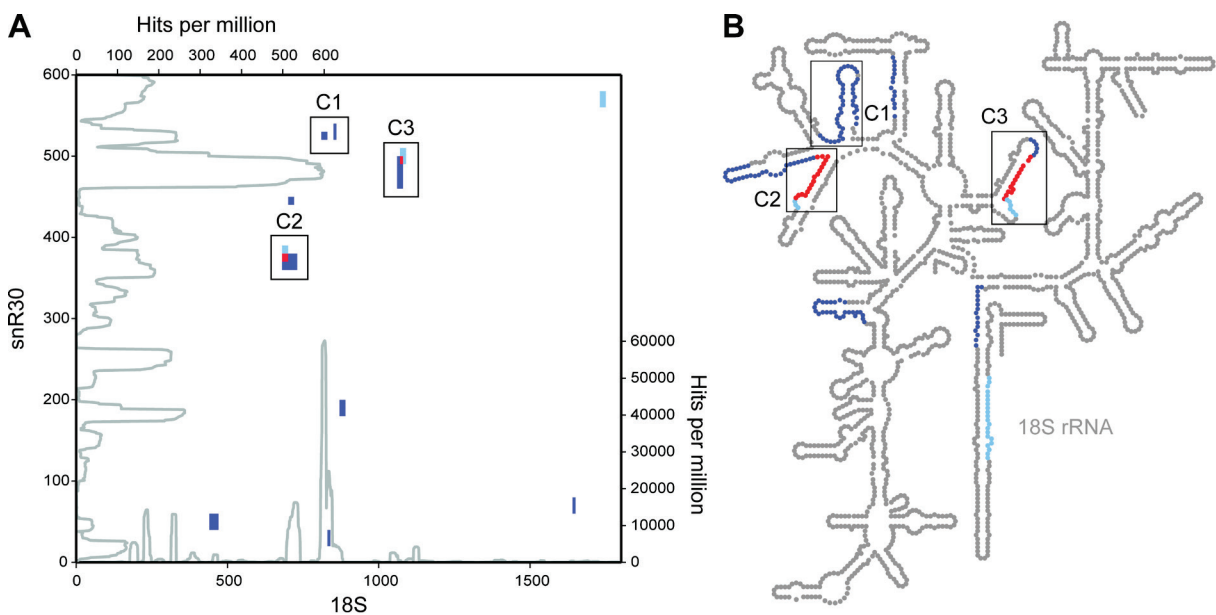


Figure 14: Rok1 crosslinking sites include new pre-rRNA basepairing sites of the snR30 snoRNA. (A) Predicted basepairing sites (light blue) and chimeric reads obtained by CLASH analysis (dark blue) between snR30 (vertical) and 18S (horizontal) were correlated according to their positions on the linear RNA sequences. Overlapping positions of the two datasets on both molecules are highlighted in red. Boxes mark clustering of interactions (C1-C3). The distribution of Rok1 crosslinking reads on both sequences is indicated (grey) according to Figure 9 and Figure 13 B. (B) Potential interaction sites from predictions and identified hybrids were mapped on the secondary structure model of mature 18S rRNA. Colouring and interaction clusters (C1-3) are indicated as in A.

ES7 mediated by sequences of the internal hairpin, might explain the strong crosslinking signal of Rok1 in this region of the snoRNA.

In summary, the application of advanced computational tools such as duplex prediction and search for sequence hybrids within experimentally curated data verified the known snR30 binding site in ES6H3 on the 18S rRNA sequence and identified potential new interactions between the snoRNA and ES6H1 as well as ES7 of the 18S rRNA sequence.

3.7 Identification of basepairing sites for snoRNAs involved in pre-rRNA processing

In addition to snR30, the box H/ACA snoRNA snR10 and the box C/D snoRNAs U3 and U14 participate in early pre-rRNA processing events (Tollervey and Guthrie, 1985) and were significantly crosslinked by Rok1 (Figure 13 A).

In contrast to single modification events, the function of these snoRNAs in pre-rRNA processing requires the formation of various long-range interactions with pre-rRNA (Beltrame and Tollervey, 1992; Beltrame and Tollervey, 1995; Dutca et al., 2011; Kudla et al., 2011; Liang et al., 2010; Morrissey and Tollervey, 1993). These snoRNA-rRNA contacts are mostly quite short, impeding their identification by computational prediction analysis. Rok1 crosslinks to these snoRNAs but its crosslinking sites on pre-rRNA do not significantly overlap with known basepairing sites of these snoRNAs. This suggests that the processing snoRNAs may form additional interactions with pre-rRNA that are located in close proximity to the Rok1 crosslinking sites. To test this hypothesis, a search for chimeric reads within the Rok1 crosslinking data was performed by CLASH analysis.

The H/ACA snoRNA snR10 was shown to guide pseudouridine formation within the 25S rRNA sequence (Ganot et al., 1997a) and loss of the snoRNA causes processing defects (Tollervey and Guthrie, 1985). Additionally, snR10 was found to interact with sequences of the 5'-ETS in pre-rRNA. This basepairing is mediated by a short sequence element (7-nucleotide processing element) in the 5'-part of the snoRNA (Liang et al., 2010). In line with previously known genetic interactions (Venema et al., 1997), a recent report proposed an additional interaction site of snR10 within the 5'- and central domains of the 18S rRNA (Lebaron et al., 2013), which harbour several Rok1 crosslinking sites. Analysis of hit distribution along the

linear snR10 sequence revealed extensive Rok1 crosslinking to the 3'-part of the snoRNA, which includes the conserved box ACA motif (Figure 15 A). This finding is supported by the accumulation of mutations within the sequences mapping to the 3'-part of the snoRNA (Figure 15 A, lower panel). Moreover, a crosslinking site, which is located in the 5'-part of snR10 and overlaps with the 7-nucleotide processing element, was detected. CLASH analysis of snR10 revealed local accumulations of chimeric reads after correlation of the sequence positions on both interacting RNA molecules (Figure 15 A). Mapping the sequence position of the chimeric reads on the secondary structure model of the mature 18S rRNA (Figure 15 B) revealed potential interactions of the snR10 3'-part and ES6. Additional hybrids consisting of sequences, which map to the 5'-part of snR10 and form potential interactions with ES6 and accordingly H9, were found.

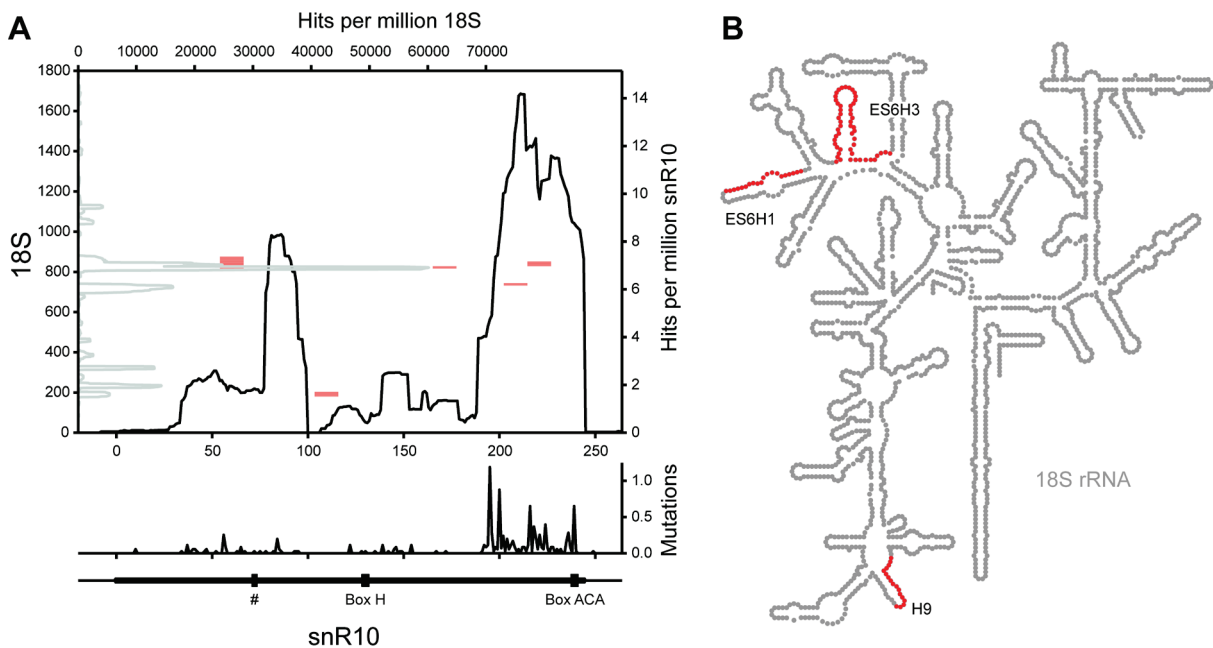


Figure 15: SnR10 interacts with 18S rRNA sequences at Rok1 crosslinking sites. (A) Distribution of mapped reads (upper panel) and mutations arising from reverse transcription stops at Rok1 crosslinking sites (lower panel) over the snR10 locus (horizontal) is shown. Relative position of the mature snR10 sequence within the genomic locus of the snoRNA and conserved motifs are indicated as bars at the bottom. The hash indicates the position of the 7-nucleotide processing element. Rok1 crosslinking peaks on the 18S rRNA sequence (vertical) are shown in grey. Chimeric reads containing sequences of snR10 and 18S rRNA are illustrated as red rectangles and their position on the plot corresponds to the relative location of the sequence fragments within the respective RNA molecule. **(B)** The sequence fragments of the identified hybrid reads were mapped on the 2D structure model of mature 18S rRNA (red circles; Petrov et al., 2013). Numbers of helices containing sequences identified by CLASH are specified.

The box C/D snoRNA U3 performs its essential function in coordinating pre-rRNA cleavage events through its multiple basepairing sites within the 5'-ETS and the 18S rRNA sequence (Beltrame and Tollervey, 1992; Beltrame and Tollervey, 1995; Dutca et al., 2011; Hughes and Ares, 1991). Interestingly, an additional basepairing site of the snoRNA in ES7 was recently proposed (Kudla et al., 2011). Analysis of hit distribution and mutations on the linear sequence of U3 showed strong crosslinking of Rok1 to the 3'-terminal part of the snoRNA, which contains the two conserved elements box C and box D (Figure 16 A). It is unlikely that the rectangular shaped peak at the intron-exon border constitutes a Rok1 interaction site since all reads mapping to this location are short and identical and probably represent a PCR or mapping artefact (van Dijk et al., 2014). The CLASH analysis of Rok1 crosslinking data and correlation of the obtained hybrids revealed several chimeric reads (Figure 16). Hybrids flanking the conserved U3 box D' motif indicate an interaction with ES6H3 of the 18S rRNA sequence (Figure 16 A). Chimeric reads, which map to

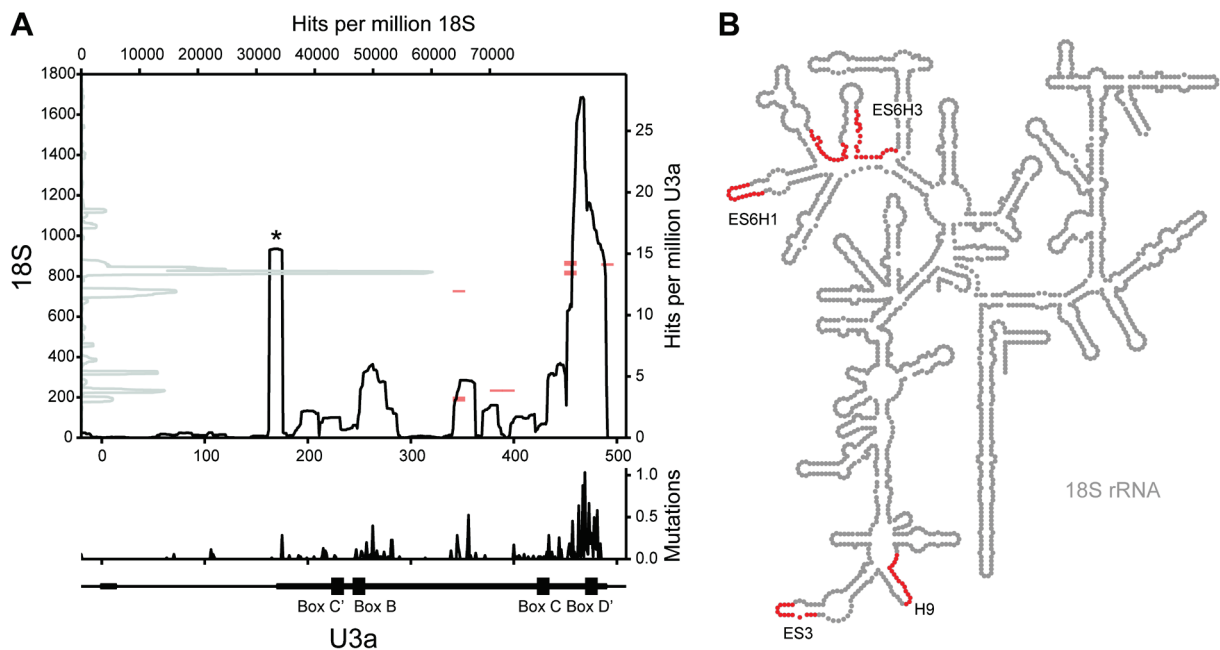


Figure 16: U3 interacts with 18S rRNA sequences at Rok1 crosslinking sites. (A) Distribution of mapped reads (upper panel) and mutations arising from reverse transcription stops at Rok1 crosslinking sites (lower panel) over the U3 locus (horizontal) is shown. The asterisk marks a crosslinking peak that should be disregarded. Bars at the bottom indicate the position of the mature U3 sequence within the genomic locus of the snoRNA, the position of the intron and conserved motifs. Rok1 crosslinking peaks on the 18S sequence (vertical) are shown in grey. Chimeric reads containing sequences of U3 and 18S rRNA are illustrated as red rectangles and their position on the plot corresponds to the relative location of the sequence fragments within the respective RNA molecule. **(B)** The sequence fragments of the identified hybrid reads were mapped on the 2D structure model of mature 18S rRNA (red circles; Petrov et al., 2013). Numbers of helices containing sequences identified by CLASH are specified.

sequences preceding the conserved box C motif, suggest possible interactions with ES3 and H9 in the 3'-part of the 18S rRNA sequence. Until recently, basepairing to pre-rRNA was only shown for the 5'-domain of the U3 snoRNA and the 3'-domain was thought to serve as binding platform for U3 specific proteins (Mereau et al., 1997; Samarsky and Fournier, 1998). The chimeric reads found here together with the lately identified potential interaction sites between the U3 3'-domain and ES7 as well as H28 of pre-18S rRNA (Kudla et al., 2011) suggest that the 3'-domain of U3 also contributes to the anchoring of the snoRNA on pre-rRNA by formation of additional basepairing interactions. The existence of multiple basepairing and potential interaction sites in the 3'-domain of the snoRNA suggests that U3 serves as a scaffold within the pre-ribosomal complexes interlinking the distinct eukaryotic expansion segments ES3, ES6 and ES7. The formation of a specific structural environment within the pre-ribosomal complexes might be the key function of U3 and explains its importance for the processing of pre-rRNA (Dutca et al., 2011).

The box C/D snoRNA U14 is another snoRNA involved in early processing events of the pre-rRNA. Its essential function in pre-rRNA processing is mediated by a sequence (domain A) that basepairs with the 5'-region of the 18S rRNA sequence (Jarmolowski et al., 1990; Liang and Fournier, 1995; Morrissey and Tollervey, 1997). Additionally, U14 guides a 2'-O-methylation reaction at residue Cm414 in 18S rRNA. This requires the basepairing of U14 domain B with pre-18S rRNA at the modification site (Dunbar and Baserga, 1998; Kiss-Laszlo et al., 1996; Trinh-Rohlik and Maxwell, 1988). A detailed analysis of the Rok1 crosslinking pattern on the linear sequence of U14 revealed an accumulation of sequence reads and mutations at the 3'-end of the snoRNA, which contains the conserved box D and domain B (Figure 17 A; Liang and Fournier, 1995; Morrissey and Tollervey, 1997). SnoRNA-rRNA interactions identified by CLASH analysis were found between the 3'-part of U14 and the 5'-domain of the 18S rRNA sequence (Figure 17 B). In addition to confirming the previously reported basepairing site of U14 (Dunbar and Baserga, 1998; Kiss-Laszlo et al., 1996) chimeric reads, which include sequences of the Rok1 crosslinking sites in H11 and H9, were identified suggesting extra interactions of U14 with these structures in pre-ribosomes.

In summary, CRAC and structure probing analyses revealed a Rok1 binding site in ES6H3 on precursors of the 18S rRNA and support a direct release of snR30 by the

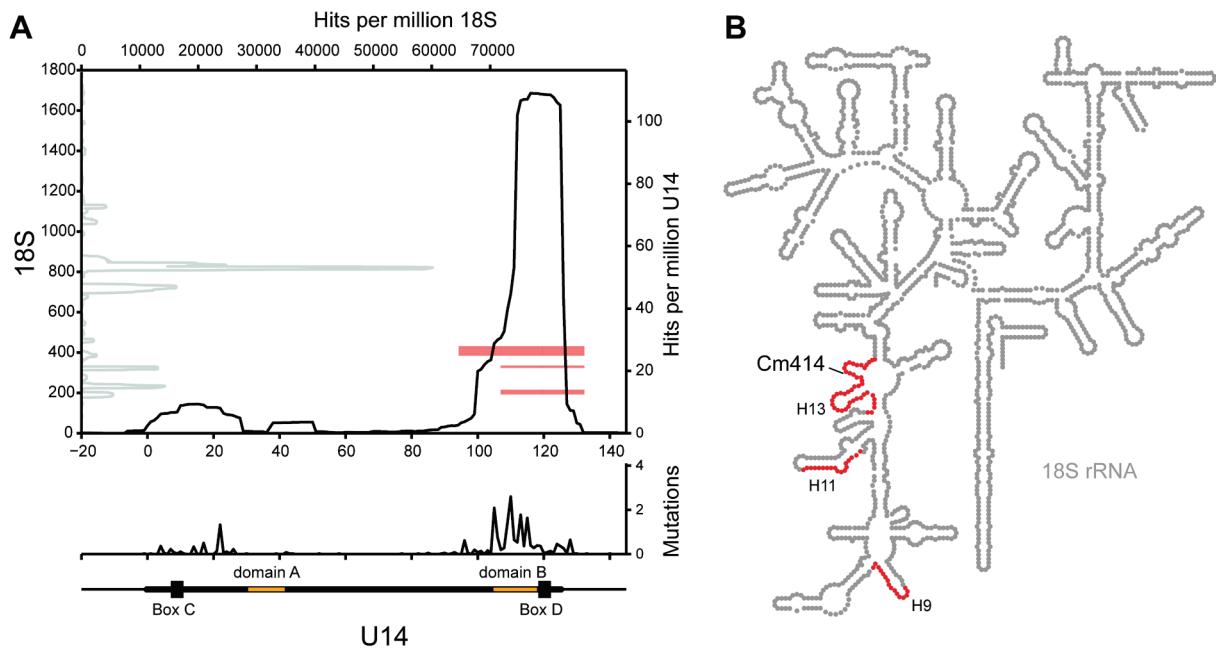


Figure 17: U14 interacts with 18S rRNA sequences at Rok1 crosslinking sites. (A) The distribution of mapped reads (upper panel) and mutations arising from reverse transcription stops at Rok1 crosslinking sites (lower panel) over the U14 locus (horizontal) is shown. Bars at the bottom indicate the position of the mature U14 sequence within the genomic locus of the snoRNA and positions of conserved motifs and the basepairing sequences domain A and domain B (yellow). Rok1 crosslinking peaks on the 18S rRNA sequence (vertical) are shown in grey. Chimeric reads containing sequences of U14 and 18S rRNA are illustrated as red rectangles, their position on the plot corresponds to the relative location of the sequence fragments within the respective RNA molecule. **(B)** The sequence fragments of the identified hybrid reads were mapped on the 2D structure model of mature 18S rRNA (red circles; Petrov et al., 2013). Numbers of helices containing sequences identified by CLASH are specified. The 2'-O-methylation at site Cm414 guided by U14, is indicated.

RNA helicase from this region. In addition, crosslinking patterns of Rok1 to snoRNAs involved in pre-rRNA processing events combined with CLASH analysis validated the previously reported methylation site of U14 snoRNA at position Cm414 and identified new potential interactions between these snoRNAs and pre-rRNAs. Together, these findings indicate the existence of an extensive interaction network between snoRNAs involved in the early processing events and the Rok1 interaction sites in ES3, ES6 and ES7 within early pre-ribosomal complexes.

3.8 Genome wide mapping shows high affinity of Prp43 to rRNA and box C/D snoRNAs

The DEAH box RNA helicase Prp43 was initially identified as cofactor of the spliceosome (Arenas and Abelson, 1997) but it is also an important maturation factor for ribosomes (Combs et al., 2006; Lebaron et al., 2005; Leeds et al., 2006). Prp43 has been shown to act in the maturation of both ribosomal subunits (Lebaron et al.,

2005). However, the precise molecular functions of Prp43 and its interaction sites on pre-rRNA have remained unknown. As demonstrated above for the RNA helicase Rok1, use of the CRAC approach allows the identification of potential binding sites and provides the basis for understanding the functions of RNA helicases in the biogenesis of ribosomes. Therefore, this method was also used to identify RNAs that interact with Prp43.

In order to perform CRAC analysis for Prp43, this helicase was genomically HTP-tagged in the BY4741 strain background (2.3.1). Similar to Rok1, the Prp43-C-HTP strain was subjected to *in vivo* and *in culturo* crosslinking. Additionally, *in vitro* crosslinking was performed (2.3.2). In this approach, following native purification of Prp43 containing complexes on IgG sepharose, UV light is applied to crosslink the isolated RNA to the bait protein. Further processing of samples and assignment of Illumina deep sequencing reads to genomic locations was performed as previously described for Rok1.

Classification of the obtained RNA sequences into functional groups revealed that for all crosslinking approaches the majority of sequence reads mapped to rRNA. Additionally, strong crosslinking of Prp43 to box C/D snoRNAs was found, especially for the *in vivo* and *in culturo* samples (Table 7). Prp43 *in vivo* crosslinking yielded an approximately 20-fold enrichment of box C/D snoRNAs compared to the corresponding control strain (Figure S). *In culturo* crosslinking resulted in more than 35-fold enrichment of box C/D snoRNA hits. Remarkably, neither box H/ACA snoRNAs nor small nuclear RNAs (snRNAs) were generally crosslinked above control levels. The strongest enrichment of rRNA sequences was observed for the *in vitro* sample. *In vitro*, however, the box C/D snoRNAs were not significantly crosslinked. This finding suggests that transiently interacting ribosome maturation cofactors, such as snoRNAs, are not effectively crosslinked after cell lysis and native complex purification.

In summary, Prp43 CRAC analysis and genome wide mapping of deep sequencing reads revealed strong enrichment of rRNA sequences and box C/D snoRNAs. These findings are consistent with the known Prp43 function in ribosome biogenesis (Lebaron et al., 2005). However, the significant crosslinking of box C/D snoRNAs by Prp43 further suggests functional interplay of the helicase and these non-coding RNAs during the process of ribosome maturation.

Table 7: Prp43 crosslinks to snoRNAs and pre-rRNA. Mapping statistics for Prp43 and control samples (upper panel). Counts of crosslinking hits shown for the complete subset of genome features of all samples normalised to the amount of mapped reads per million (lower panel).

	Prp43 <i>vitro</i>	Prp43 <i>vivo</i>	Prp43 <i>culturo</i>	Control <i>vitro</i>	Control <i>vivo</i>	Control <i>culturo</i>
mapped reads	3,400,667	4,154,261	10,985,667	6,473,346	4,106,088	10,902,523
discarded reads	879,799	1,610,473	1,655,399	2,587,652	2,761,161	6,459,656
unique mapped reads	2,145,885	2,006,070	8,028,667	3,405,888	1,082,670	3,808,416
Features	Hits per million					
35S rRNA	826,794	550,703	562,672	332,617	448,956	449,874
Intergenic	105,515	211,595	141,371	387,385	296,175	294,331
Genes	64,522	193,499	212,764	261,195	218,662	227,094
snoRNA (C/D)	1,120	15,988	54,409	1,822	742	1,464
5S rRNA	559	1,752	1,689	1,402	1,882	2,732
SRP RNA	415	1,003	1,972	13	1,132	1,147
Rep origin	187	1,028	975	1,686	492	693
Repeat region	185	1,233	850	808	993	892
tRNA	181	17,233	14,532	8,670	26,232	17,699
snRNA	112	2,135	3,211	1,930	3,109	2,487
snoRNA (H/ACA)	112	1,313	2,908	1,498	3,109	2,424
LTR	103	432	490	305	65	191
Telomerase RNA	12	67	76	1	143	27
RNase P RNA	7	231	86	4	339	213
RNase MRP RNA	7	20	376	3	54	194
Misc feature	2	47	49	51	112	244

(SRP: signal recognition particle; Rep origin: replication origin; LTR: long terminal repeats; Misc: miscellaneous)

3.9 Prp43 is required for pre-ribosomal release of box C/D snoRNAs involved in LSU maturation

Most snoRNAs guide post-transcriptional modifications in pre-rRNA by basepairing to their designated target positions (Kiss-Laszlo et al., 1996). Recent publications showed that RNA helicases are able to regulate snoRNA accumulation on pre-ribosomal particles (3.5; Bohnsack et al., 2008; Kos and Tollervey, 2005; Liang and Fournier, 2006). Since a striking enrichment of box C/D snoRNAs was observed in the Prp43 crosslinking experiments, a qPCR-based analysis was employed to identify snoRNAs that accumulate on pre-ribosomes when Prp43 is depleted (2.7).

Depletion of the helicase was achieved in a strain carrying a tetracycline or doxycycline repressible promoter in front of the Prp43 gene locus (Longtine et al., 1998). In order to identify exclusively primary effects caused by the absence of Prp43 from pre-ribosomal particles, cells were harvested 6 hours after doxycycline addition before any growth defects were detectable (Bohnsack et al., 2009). The cell lysate was subjected to density gradient centrifugation to separate free snoRNAs from

bound snoRNAs co-migrating with pre-ribosomal complexes (40-90S). The relative levels of pre-ribosome-bound snoRNAs were quantified by qPCR and normalised to the values of the wild type sample, which was processed in parallel (Figure 18; Bohnsack et al., 2009). A ratio of 1 corresponds to the snoRNA distribution found in wild type. Values of 0.48 and 2.1, indicated by two horizontal lines, represent the borders of the 95% confidence interval. Differences of snoRNA levels on pre-ribosomal particles after Prp43 depletion were observed for several snoRNAs. Interestingly, only the pre-ribosome levels of box C/D snoRNAs involved in maturation of the LSU were affected by Prp43 depletion. The snoRNAs U18, snR39, snR39b, snR59 and snR72 were significantly enriched. The accumulation of these snoRNAs on pre-ribosomes after depletion of Prp43 suggests that the RNA helicase is required for their release. In contrast, the relative levels of the snoRNAs snR64

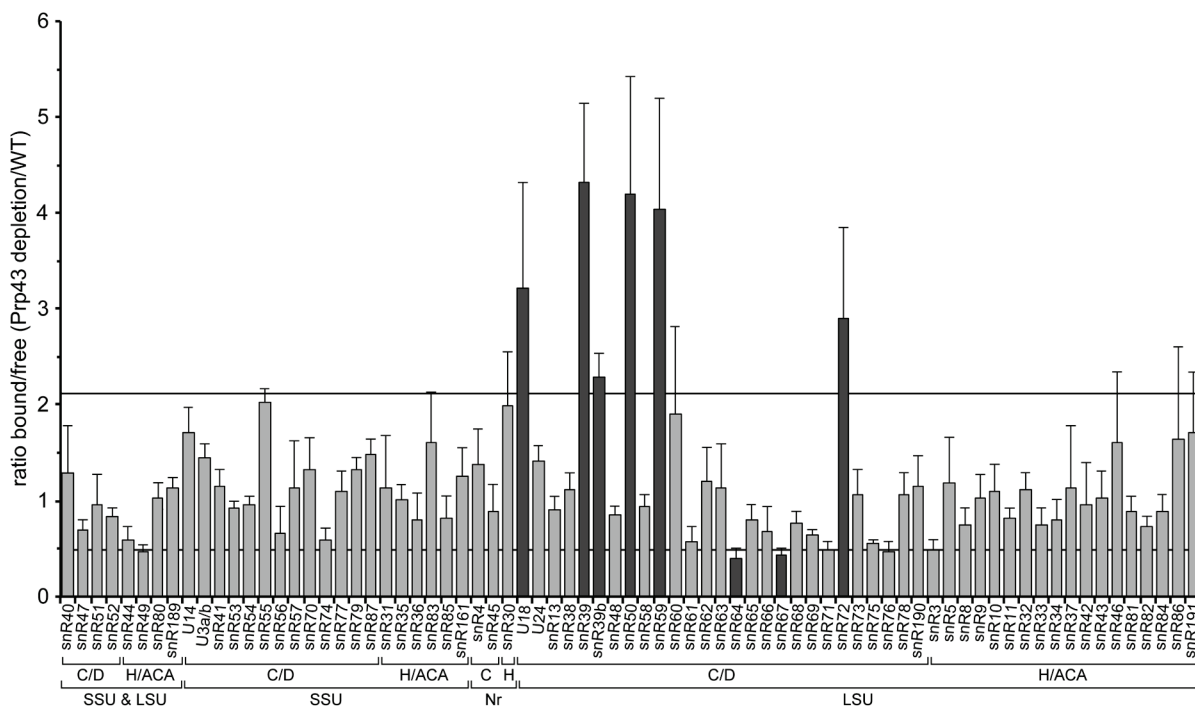


Figure 18: Prp43 depletion alters the accumulation of several box C/D snoRNAs on pre-ribosomal particles. Exponentially growing wild type (WT) and TetO₇-Prp43 strains were cultivated for 48 h in YPD medium. Prp43 was depleted for 6 h by addition of doxycycline and cells were subsequently harvested and lysed. Soluble components of the cell extracts were separated by sucrose density centrifugation (10-45%). The fractions at the top of the gradient containing free RNA were pooled and separated from the fraction pool that includes pre-ribosomal-bound RNA. The RNA was extracted and reverse transcribed. Relative levels of all 75 snoRNAs involved in ribosome biogenesis in each pool were determined by qPCR. The obtained data were normalised to snoRNA levels in WT samples (set as 1) and the average ratios of pre-ribosome bound to free snoRNAs are shown. Error bars indicate standard errors from three independent experiments. The horizontal lines represent the borders of the 95% confidence interval. Values below or above the significance borders are highlighted in black. The snoRNAs are grouped according to their classification and modification targets (C/D or C: C/D box snoRNA; H/ACA or H: H/ACA snoRNA; SSU: small subunit; LSU: large subunit; Nr: no target reported).

and snR67 on pre-ribosomes were decreased after Prp43 depletion indicating that Prp43 is required for the association of these snoRNAs with pre-ribosomal particles.

The observed effects on snoRNA accumulation on pre-ribosomal particles obtained here by qPCR were also confirmed by Northern blotting analysis (Bohnsack et al., 2009). To verify that the changes in the amount of snoRNAs associated with pre-ribosomal complexes did not arise due to alterations in the overall levels of these snoRNAs, total snoRNA levels were determined by Northern blotting. This showed that depletion of Prp43 did not significantly alter the total levels of any of the affected snoRNAs (Figure 19).

In summary, the qPCR-based approach identified a novel role for Prp43 in regulating the association of a particular subset of snoRNAs from pre-LSU complexes. The RNA helicase is required for the release of several snoRNAs from pre-ribosomes but might also regulate the access of other snoRNAs to pre-ribosomal complexes.

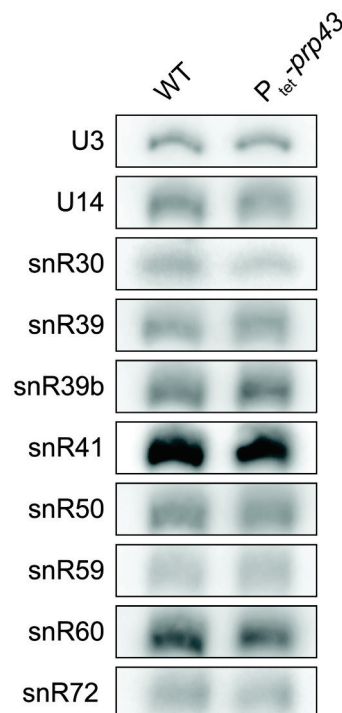


Figure 19: Prp43 depletion does not affect total cellular levels of box C/D snoRNAs.

RNA was extracted from exponentially growing wild type (WT) and Prp43 (P_{tet} -*prp43*) depleted cells. Equal quantities of RNA were separated on a denaturing polyacrylamide gel and subsequently transferred to nitrocellulose membrane by Northern blotting. The levels of several box C/D snoRNAs (indicated on the left) was monitored by hybridisation of radiolabelled probes.

3.10 Prp43 crosslinks 25S rRNA sequences near snoRNA guided modification sites

The CRAC data revealed the affinity of Prp43 for pre-rRNA and box C/D snoRNAs and qPCR analysis showed that the helicase regulates the interaction of several snoRNAs with pre-ribosomal particles. These findings suggest the presence of Prp43 in pre-ribosomal complexes and indicate that the helicase might also interact with pre-rRNA. In order to allocate potential binding sites of Prp43, sequence reads obtained from the CRAC analysis were mapped on the genomic locus of the 35S pre-rRNA transcript, RDN37 (Figure 20 A-C).

Consistent with the known function of Prp43 in the maturation of both ribosomal subunits (Lebaron et al., 2005), strong crosslinking of the helicase to several sites in the sequences of 18S and 25S rRNA was observed. In the 18S rRNA sequence a major peak in H44 and a second less prominent crosslinking site in ES6H2 were consistently identified (3.12). Pre-rRNA of the LSU was found to contain four prominent crosslinking sites located in helix 23 (H23), helix 29 (H29), helix 40 (H40) and helix 61 (H61) of the 25S rRNA sequence. All peaks were found in at least two independent experiments. Additionally, a peak located in helix 84 (H84) was present in the samples crosslinked *in culturo* (Figure 20 C). Crosslinking sites also present in control samples were disregarded (Figure S).

To assess the correlation of snoRNA modification sites and Prp43 crosslinking sites in more detail, the specific crosslinking peaks of the helicase were mapped on the 2D structures of the mature ribosomal LSU rRNAs (Figure 21 A; Petrov et al., 2013). One major crosslinking peak in H29 of pre-25S rRNA is located in close proximity to several nucleotides that are modified by snoRNPs, which were shown to accumulate on pre-ribosomal complexes after the depletion of Prp43 (U18, snR39, snR39b, snR50 and snR72; Figure 21 B). Close to the prominent peak in H29, an additional minor crosslinking signal was located in helix 35a (H35a). This crosslinking site covers one of the modifications guided by snR60, a snoRNA that accumulated on pre-ribosomes with values just below the 95% confidence border. These findings demonstrate that the modifications guided by the snoRNAs accumulating after Prp43 depletion cluster in one region of the 25S rRNA sequence and interestingly, this location contains a prominent Prp43 crosslinking site. This observation provides

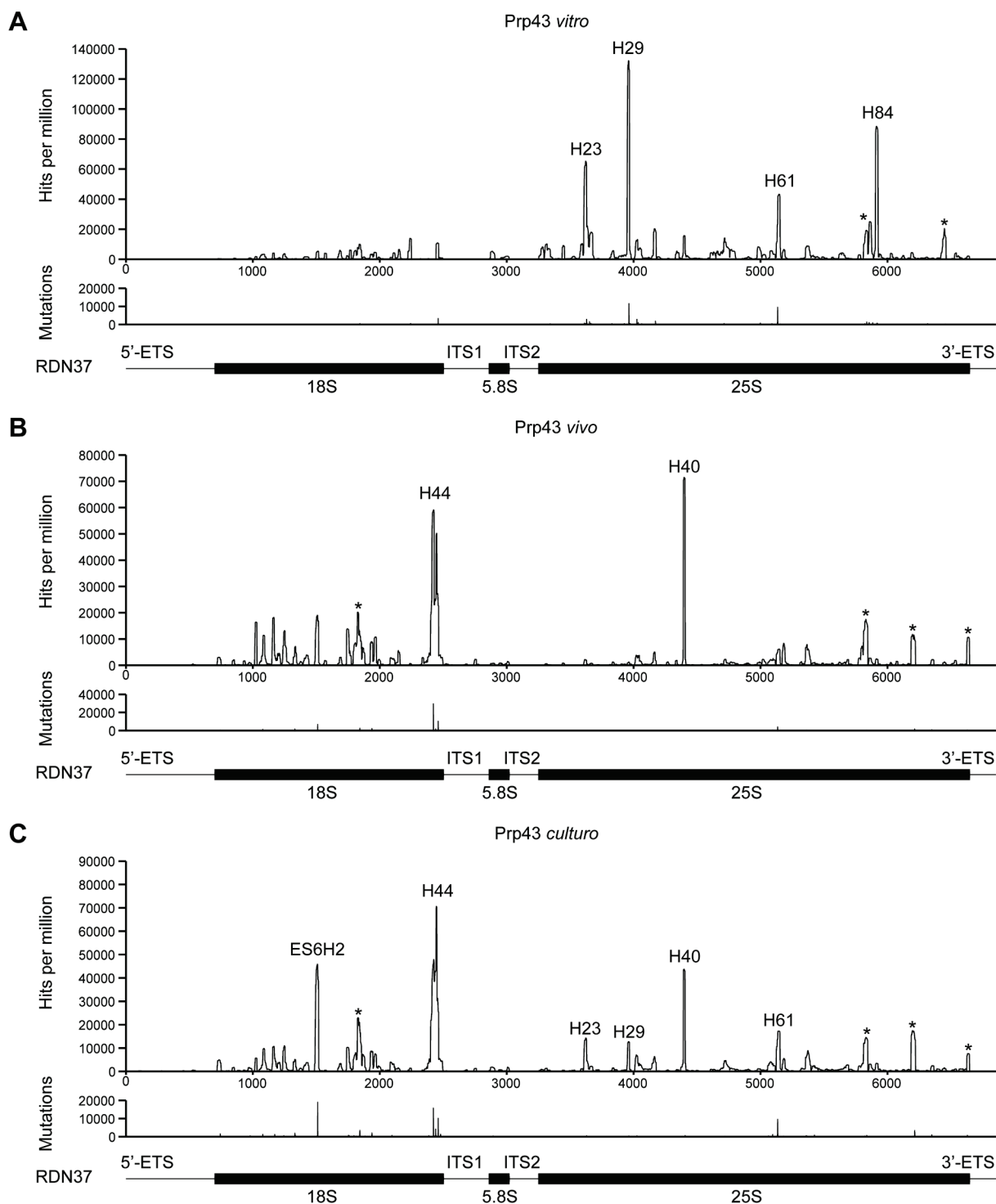


Figure 20: Prp43 crosslinks to rRNA sequences of both ribosomal subunits. Distribution of sequence reads obtained by (A) *in vitro*, (B) *in vivo* and (C) *in culturo* crosslinking of Prp43-C-HTP on the linear RDN37 sequence, which encodes the 35S pre-rRNA, is shown. Number of hits were normalised to total mapped reads per million for each nucleotide. Relative positions of the mature 18S, 5.8S and 25S rRNA sequences, external transcribed spacer regions (ETS) as well as internal transcribed spacers (ITS) are indicated below the graphs. Peaks are labelled according to structural elements to which they correspond (H: helix; ES: expansion segment). Peaks also present in the corresponding control experiments are labelled with asterisks (Figure S).

substantial evidence that the RNA helicase is directly involved in the release of these snoRNAs from pre-ribosomal particles.

Further analysis of the secondary structure did not reveal a functional relevance for the Prp43 crosslinking peak in H40 of 25S pre-rRNA since the two snoRNAs, which guide modifications in close proximity to this potential interaction site, were not affected by depletion of the RNA helicase (Figure 21 C). Interestingly, Prp43 was also found to crosslink between H23 and H25 of the 25S rRNA sequence (Figure 21 D). This region of the pre-rRNA is devoid of any modifications but is involved in basepairing interactions with the 5.8S rRNA in the mature ribosome. Prp43 crosslinking to H23 and H25 as well as a small peak near the 5'-end of 5.8S rRNA sequence (data not shown) suggest a possible role for the helicase in mediating structural rearrangements in these regions of the pre-ribosome.

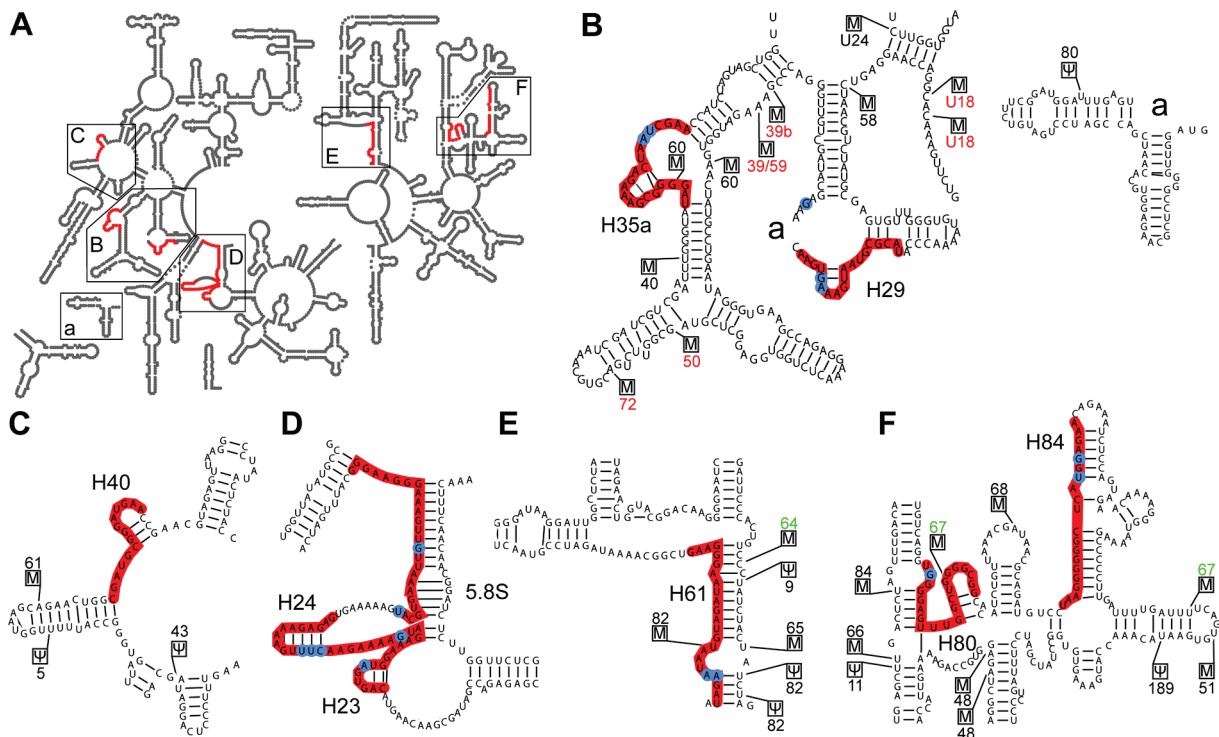


Figure 21: Prp43 crosslinks to sequences of the 25S rRNA in close vicinity to snoRNA modification sites. (A) Overview of the relative position of the individual crosslinking sites. (B-F) Magnifications of structural elements surrounding the crosslinking site: (B) helix 29 (H29) and helix 35a (H35a), (C) helix 40 (H40), (D) helix 23 (H23), helix 24 (H24) and 5.8S rRNA (5.8S), (E) helix 61 (H61), (F) helix 80 (H80) and helix 84 (H84). The prominent crosslinking sites of all independent Prp43 CRAC experiments were merged and mapped on the secondary structure of the mature 25S rRNA. Prp43 crosslinking sites are coloured red and positions with a substantial accumulation of mutations arising from crosslinking induced reverse transcription errors are indicated as blue-circled nucleotides. SnoRNA directed modification sites are indicated by boxes (M: 2'-O-methylation; Ψ: pseudouridylation) and numbers indicate the snoRNA guiding the respective modification. Red numbers indicate snoRNAs that significantly accumulate on pre-ribosomal particles upon depletion of Prp43. Green numbers mark snoRNAs with significantly reduced levels in pre-ribosomal fractions after Prp43 depletion.

In contrast, the Prp43 crosslinking sites at H61 and H80/H84 are adjacent to multiple snoRNA guided modification sites (Figure 21 E and F). Interestingly, snR64, whose association with pre-ribosomal complexes was decreased after Prp43 depletion, guides a modification on the complementary strand of the Prp43 crosslinking site in H61 (Figure 21 E). Similarly, the 2'-O-methylation site of snR67, which also shows decreased association with pre-ribosomal complexes upon Prp43 depletion, is located directly within the Prp43 crosslinking site in H80 and additionally guides a modification reaction about 40 nucleotides downstream of the Prp43 crosslinking site in H84 (Figure 21 F). The close correlation between the Prp43 crosslinking site in H61, H80 and H84 and modifications sites of the snoRNAs snR64 and snR67, whose pre-ribosomal association is reduced upon helicase depletion, strongly suggests that the helicase might be involved in remodelling processes that allow the recruitment of these two snoRNAs to pre-ribosomal complexes.

To include the aspect of tertiary RNA folding in the data analysis, the crosslinking sites of Prp43 were also mapped on the 3D structure of the mature large ribosomal subunit (Ben-Shem et al., 2011). Although the pre-ribosomal particles targeted by Prp43 may differ from the mature LSU structure used for the data mapping, a remarkable clustering of Prp43 crosslinking sites in H23, H40, H61, H80 and modification sites of snoRNAs affected by Prp43 depletion was found in the centre of the LSU (Figure 22). This cluster also contains the Prp43 crosslinking sites in H23 and H40, which seemed to be distant from functional relevant snoRNA modification sites in the secondary structure of 25S rRNA. Additionally, the Prp43 crosslinking sites in H61 and H80 that are located close to snR64 and snR67 in the 2D structure are close to the basepairing sites of snoRNAs that accumulate on pre-ribosomes upon Prp43 depletion.

Together, these data show that Prp43 crosslinks to multiple sequences in precursors of the 18S and 25S rRNAs, validating its role in the maturation of both ribosomal subunits (Lebaron et al., 2005). The modification sites of snoRNAs affected by Prp43 depletion and several crosslinking sites of the RNA helicase are in immediate proximity within the secondary and cluster in the 3D structure of the mature LSU. These findings suggest that Prp43 targets the centre of pre-LSU particles and regulates the association or dissociation of certain snoRNAs.

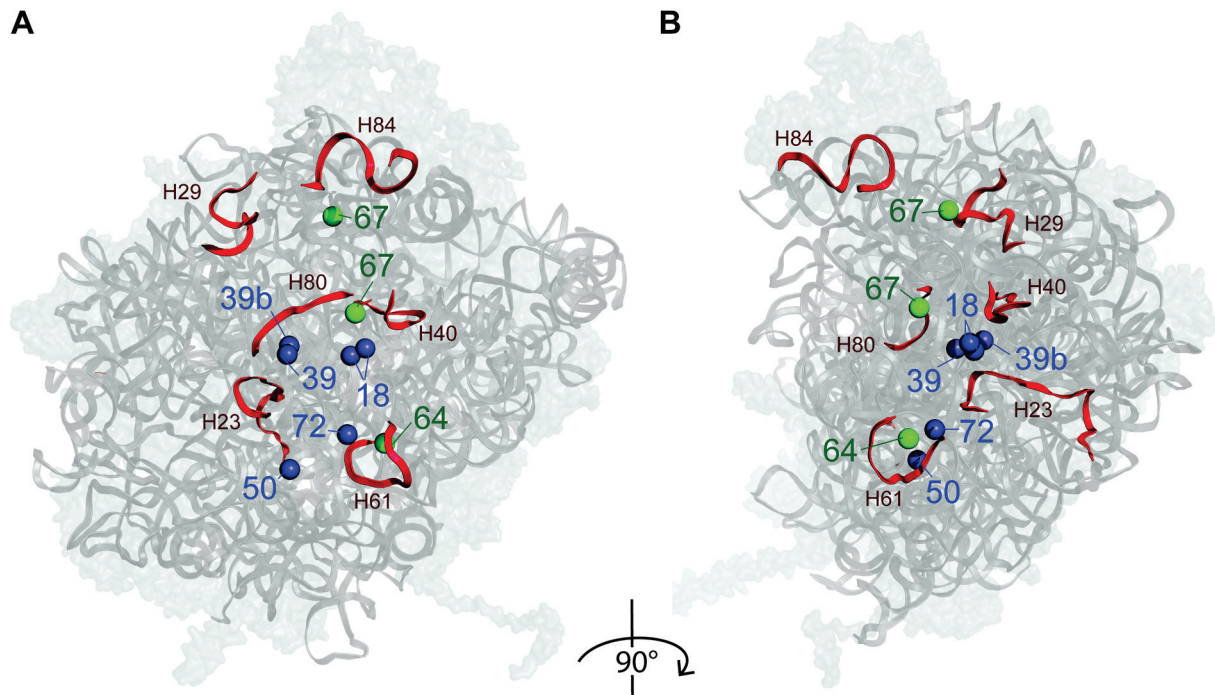


Figure 22: Prp43 crosslinking sites cluster mainly in the centre of the 3D LSU structure. Mapping of Prp43-C-HTP crosslinking sites on the 3D structure of the mature 60S ribosomal subunit (Ben-Shem et al., 2011). The major Prp43 crosslinking sites are shown in red on the RNA backbone. Prp43 crosslinking sites are labelled according to the structural elements that are crosslinked (H: helix, ES: expansion segment). 2'-O-modification sites are mapped and shown in blue (accumulating snoRNAs upon Prp43 depletion) and in green (snoRNAs with decreased association to pre-ribosomal complexes after Prp43 depletion). SnoRNAs are specified by numbers and corresponding colours (blue and green). The surfaces of ribosomal proteins included in the crystal structure are shown in light grey. **(A)** View on the subunit interface. **(B)** Side view rotated 90° clockwise.

3.11 Prp43 crosslinks to the affected snoRNAs and copurifies with snoRNA-rRNA hybrids

The clear correlation of Prp43 crosslinking and modification sites of snoRNAs affected by the depletion of the helicase suggests a direct role of Prp43 in the release of these snoRNAs from pre-ribosomal complexes. Since box C/D snoRNAs guide 2'-O-methylations by direct basepairing to their specific target sites on pre-rRNA, Prp43 might facilitate their release by its local strand unwinding activity, which was previously shown *in vitro* (Tanaka and Schwer, 2006). This model would suggest a direct interaction of Prp43 with both substrate RNA molecules. Therefore, the sequence reads obtained from the three independent crosslinking experiments were mapped to the genomic loci of four box C/D snoRNAs (snR39, snR50, snR64 and snR72), which displayed an altered accumulation on pre-ribosomal particles after depletion of Prp43 (Figure 23 A-D). The number of reads that map to each nucleotide

was determined and normalised on the total count of mapped reads in the particular experiment. The resulting peaks indicate potential Prp43 interaction sites on the snoRNA similar to the previously described mapping on the rDNA locus.

The snoRNAs snR39 and snR50 are crosslinked by Prp43 in their central region and in their 3'-end, which contains the conserved box D', box C' and box D motifs as well as the sequences that guide the nucleotide modifications by basepairing to the 25S pre-rRNA (Figure 23 A and B). In contrast to snR39 and snR50, sequence reads

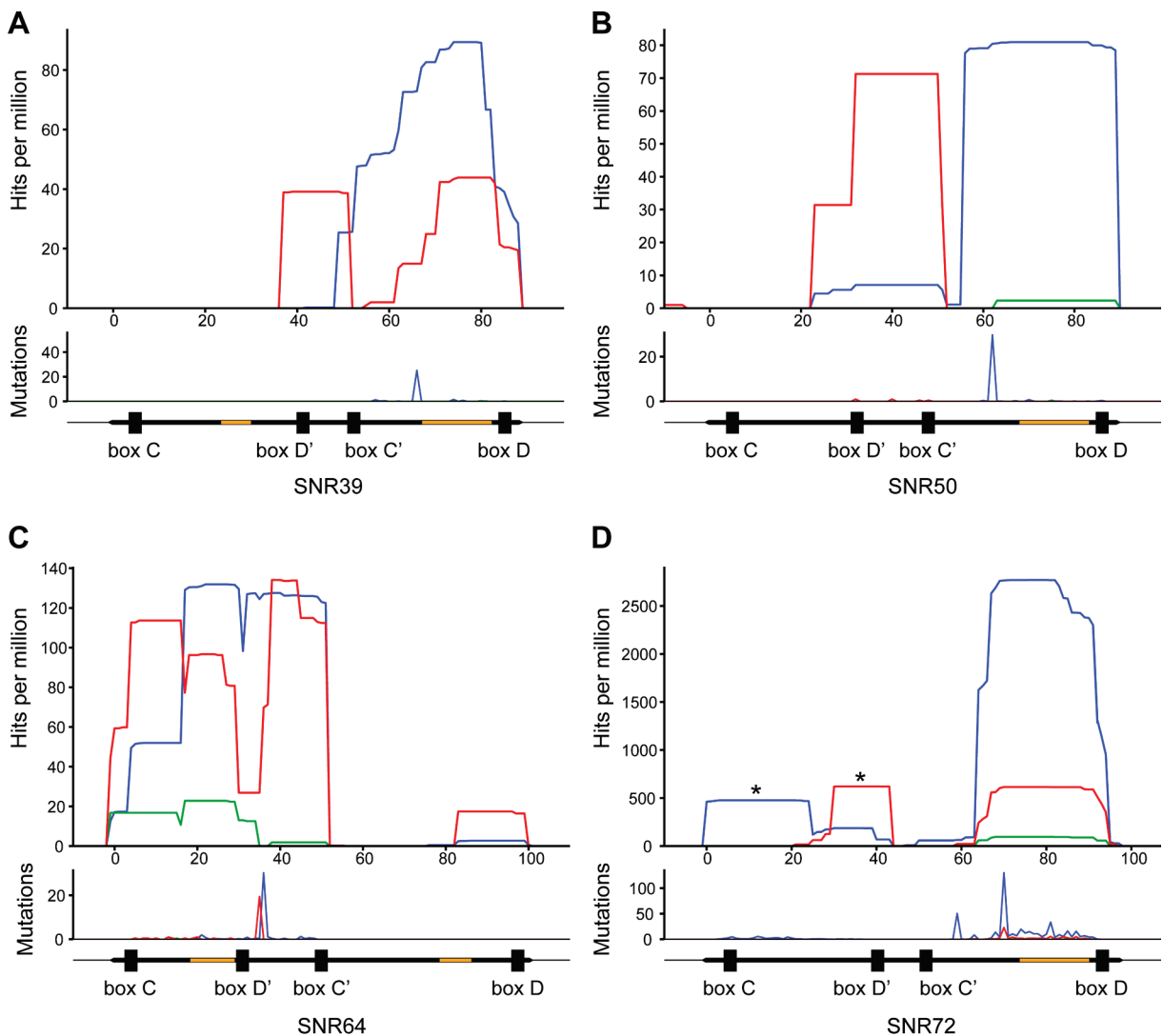


Figure 23: Prp43 crosslinks to box C/D snoRNAs whose pre-ribosome association is affected by depletion of the RNA helicase. Number of Prp43 crosslinking hits is assigned on the linear genomic sequence of (A) snR39, (B) snR50, (C) snR64 and (D) snR72. The hit values were normalised to the total number of mapped reads per million. Colouring of the histograms indicates the crosslinking method (green: *in vitro*; red: *in vivo*; blue: *in culturo*). Relative positions of the mature snoRNA sequences and the location of the conserved motifs (box C, box D', box C' and box D) as well as the sequences guiding the 2'-O-methylation by basepairing (yellow bars) are indicated at the bottom of each graph. Asterisks mark peaks that were also present in control samples. The numbers of point mutations, which arise during reverse transcription and indicate direct Prp43 crosslinks, are plotted in the lower panels of the graphs.

mapping to snR64 are predominantly located in the 5'-part of the snoRNA in all three independent experiments (Figure 23 C). Interestingly, the potential Prp43 interaction site again covers the 2'-O-methylation guide region but also the conserved box C, box D' and box C' motifs. An additional small peak at the 3'-end partially overlaps with a short sequence element that basepairs to pre-rRNA immediately downstream of the modification site and has been reported to stabilise the snoRNA-pre-rRNA interaction (van Nues et al., 2011). The majority of Prp43 crosslinks to snoRNAs were found to map to snR72 (Figure 23 D). Similar to snR39 and snR50, the main Prp43 crosslinking site is located in the 3'-part of the snoRNA. This region of snR72 harbours the conserved box D motif and the 2'-O-methylation guide sequence, which mediates the basepairing to 25S rRNA precursors. Two minor peaks in the 5'-end of snR72 were also found in the control samples and were therefore disregarded. For all analysed snoRNAs the occurrence of several mutations within the mapped sequence reads indicate direct contacts between Prp43 and these snoRNAs and therefore validates these potential interaction sites (Figure 23, lower panels).

The data show that Prp43 crosslinks mainly to the sequences, which mediate basepairing interactions between the snoRNA and the pre-rRNA, providing additional support for the model of local strand unwinding by Prp43 to release snoRNAs from pre-ribosomal complexes.

In contrast to the snoRNAs examined above, the main target sites of Prp43 on 25S pre-rRNA are adjacent but do not directly overlap with the basepairing sites of the affected snoRNAs (Figure 21). However, close examination of the Prp43 crosslinking profile (Figure 20; Figure 24 A) revealed several small peaks encompassing the crosslinking site in H29 and covering several 2'-O-methylation sites of the affected snoRNAs indicating that pre-rRNA sequences involved in snoRNA basepairing are directly contacted by Prp43. The fact that the peaks are small might indicate that Prp43 only interacts transiently with the basepaired snoRNA at these locations of the 25S rRNA sequence and is consistent with the suggested processive activity of helicase (Walbott et al., 2010). The major crosslinking sites may therefore rather reflect pausing or recruitment sites of Prp43. To show that Prp43 simultaneously crosslinks both substrate molecules, a search for sequence hybrids in the crosslinking datasets obtained by CRAC was performed with the CLASH analysis (2.4).

Chimeric reads between pre-rRNA and snoRNAs that accumulate on pre-rRNA after Prp43 depletion were identified in all crosslinking experiments and subsequently merged. The potential basepairing locations of hybrids were mapped on the 25S rRNA sequence and correlated to the 2'-O-methylation sites of the corresponding snoRNAs. Remarkably, nearly all potential basepairing interactions found by CLASH analysis colocalise with the known modification sites of the analysed snoRNAs (Figure 24 A) except for snR39b, for which no chimeric read was discovered. In contrast, hybrid search in the data sets of the control samples revealed only one chimeric read composed of U18 and the 25S rRNA but it is currently not clear whether there is any physiological relevance for this interaction (Figure 24 B).

Taken together, the crosslinking data revealed that Prp43 crosslinks to pre-rRNA and box C/D snoRNAs. Additionally, the RNA helicase was shown to be required for the release of a particular subset of the box C/D snoRNAs from pre-ribosomal particles. The modification sites of the affected snoRNAs and the main crosslinking sites of Prp43 cluster in the centre of the LSU. Crosslinking data and CLASH analysis

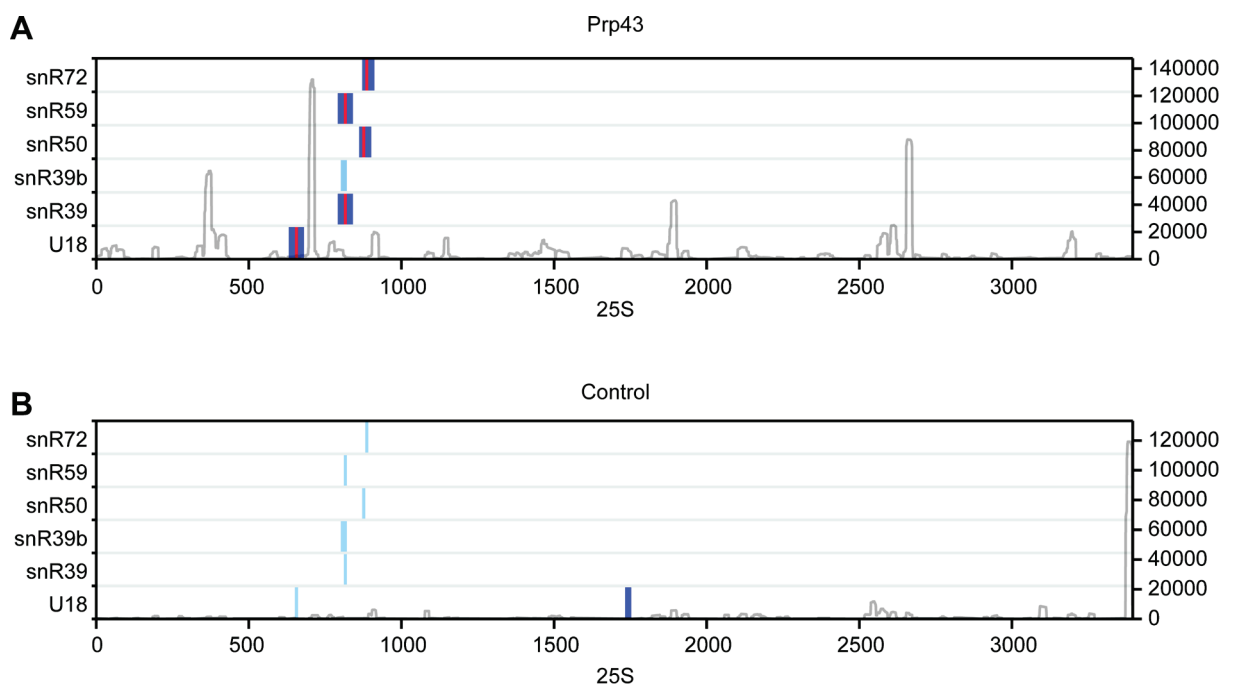


Figure 24: Chimeric reads within the CRAC data validate simultaneous interaction of Prp43 with the affected snoRNAs and pre-ribosomal RNA. Locations of 2'-O-methylations (light blue) and chimeric reads obtained by CLASH analysis (dark blue) between the snoRNAs accumulating upon Prp43 depletion were correlated according to their positions on the linear 25S rRNA sequence. **(A)** Hybrid reads identified in all independent Prp43 crosslinking experiments. **(B)** Chimeric reads identified in all control crosslinking experiments. Overlapping positions are highlighted in red. The distribution of Prp43 or control crosslinking reads on the 25S rRNA sequences are indicated (grey).

revealed that Prp43 directly crosslinks the duplexes formed between snoRNAs and sequences of the 25S rRNA. The data presented here provide strong evidence that Prp43 dissolves the snoRNA-pre-rRNA duplexes and thereby mediates the release of these snoRNAs from the pre-ribosomal complexes.

3.12 Prp43 crosslinks close to the 3'-end of the 18S rRNA sequence

Crosslinking and qPCR data revealed a role for Prp43 in the release of a set of snoRNAs from pre-60S ribosomal particles (Figure 18 - Figure 24). However, Prp43 is also implicated in the maturation of the SSU, where it has been shown to be required for the final processing step of the 20S pre-rRNA to the mature 18S rRNA (Lebaron et al., 2005). Cleavage of the 20S pre-rRNA at site D, which generates the 3'-end of the mature 18S rRNA, is performed by the endonuclease Nob1 (Fatica et al., 2003). Nob1 was shown to genetically interact with Prp43 (Pertschy et al., 2009) and the helicase was therefore suggested to function in remodelling of the 20S pre-rRNA to enable the cleavage reaction catalysed by the endonuclease. Consistent with the proposed function of Prp43 in the biogenesis of both ribosomal subunits, CRAC analysis and mapping of obtained crosslinking reads identified not only the LSU targets (3.10) but also two potential target sites of the helicase in ES6H2 and H44 within the 18S rRNA sequence (Figure 20; Figure 25). A crosslinking peak in H27 was observed in both Prp43 and control samples and was therefore disregarded (Figure S). ES6H2 is located in the central domain of the 18S secondary structure whereas the most prominent crosslinking site of Prp43 was found near the 3'-end of 18S within helix 44 (H44) directly preceding the cleavage site D. This crosslinking site was validated by its appearance in two independent CRAC experiments. Interestingly, the target site of the endonuclease Nob1 is located immediately downstream of this potential Prp43 interaction site. The close proximity of the Prp43 crosslinking site to the Nob1 active site is consistent with their reported genetic interactions (Pertschy et al., 2009) and supports a model of functional interplay.

Taken together, besides its functions in LSU maturation, crosslinking data uncovered a potential interaction site of Prp43 near the 3'-end of the 18S rRNA sequence in H44. The spatial proximity of the Prp43 crosslinking and the Nob1 target site suggest a role of the helicase in structural rearrangement in the vicinity of H44 to enable the processing performed by the endonuclease.

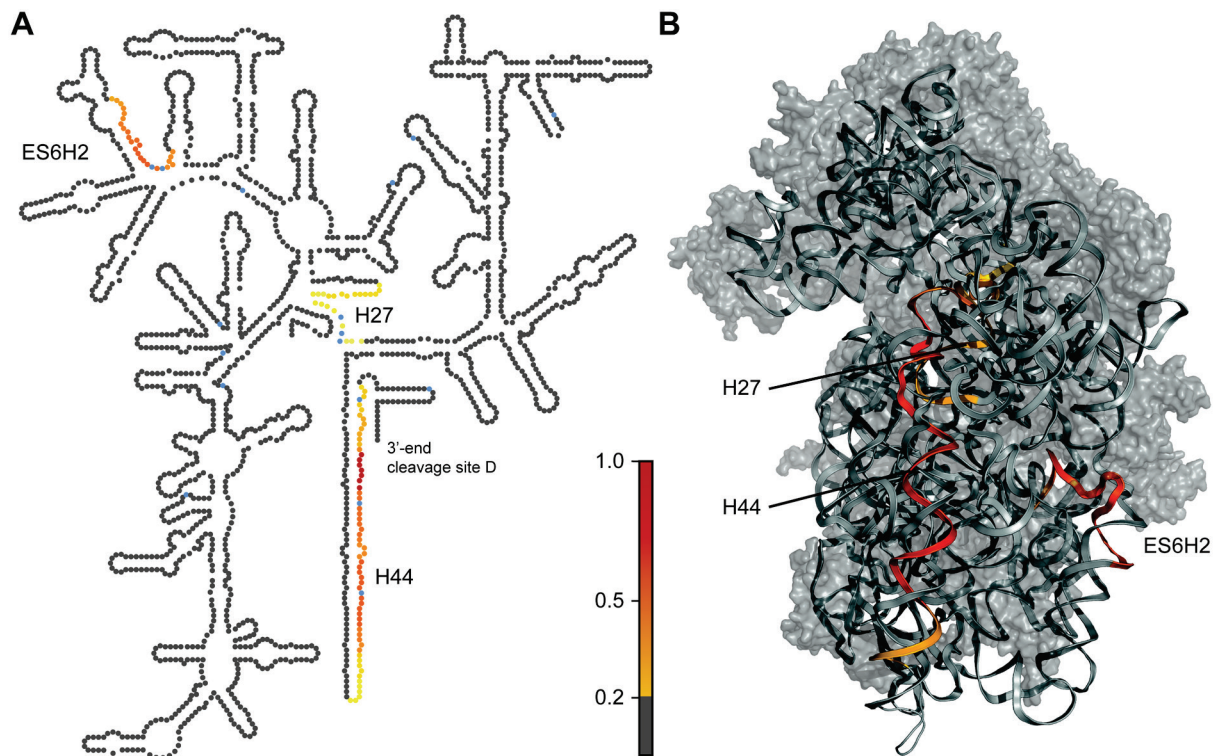


Figure 25: Prp43 crosslinks mainly to helix 44 in the 18S rRNA sequence. (A) Crosslinking hits from Prp43-C-HTP *in culturo* experiments above the threshold of 20% of the highest peak within 35S pre-rRNA were mapped on the secondary structure model of mature 18S rRNA (Petrov et al., 2013). Prp43 crosslinking sites are indicated as coloured circles. The fill colour of the circles indicates the relative hit values as colour gradient from red (100%) to yellow (20%) also indicated in the colour bar. Blue circles mark positions of mutations that arise from crosslinking induced reverse transcription errors. Helices that contain Prp43 crosslinking sites are specified (H: helix, ES: expansion segment). The 3'-end of mature 18S rRNA, which is generated after cleavage at processing site D, is indicated. **(B)** Mapping of Prp43 crosslinking sites from C-HTP *in culturo* experiments on the 3D structure of the mature 40S ribosomal subunit (Ben-Shem et al., 2011). Prp43 crosslinking sites are visualised as colour gradient on the RNA backbone as described for A. Locations of Prp43 crosslinking sites are indicated (H: helix, ES: expansion segment). The surfaces of ribosomal proteins included in the crystal structure are shown in light grey.

3.13 The amino acid sequence of Nob1 is conserved from archaea to eukaryotes

The endonuclease Nob1 was extensively studied by *in vitro* and *in vivo* analysis in *S. cerevisiae*, which led to the discovery of pre-rRNA targets (Fatica et al., 2003), enzymatic cofactors (Pertschy et al., 2009) and regulation mechanisms (Strunk et al., 2011) of the enzyme. However, an atomic resolution structure of the Nob1 protein, which might provide detailed information about the structural organisation and functions of the endonuclease within pre-ribosomal particles, was not available.

The final maturation step of the 18S rRNA performed by Nob1 is assumed to be conserved throughout eukaryotic organisms (Mullineux and Lafontaine, 2012). In

addition to mammals, fungi, protists and archaea were also found to possess many orthologous sequences of yeast Nob1 (Fatica et al., 2003). Thermophile archaea, which possess a highly stable proteome, can effectively be used to determine atomic structures, so a detailed analysis of the sequence conservation between archaeal and eukaryotic Nob1 orthologues was performed.

A sequence alignment of Nob1 homologues from *S. cerevisiae*, *Homo sapiens*, *Arabidopsis thaliana* and the archaeon *Pyrococcus horikoshii* demonstrated that all sequences have a PIN (PiIT N terminus) domain. The PIN domain contains four aspartate residues, which have been shown to be critical for the endonucleolytic cleavage activity of Nob1 (Figure 26; Makarova et al., 1999). Additionally, an N-terminal zinc ribbon domain containing four conserved cysteine residues, which coordinate a central zinc ion, was identified within the Nob1 sequence. This structural motif was characterised as an RNA binding domain and is also found in sequences of several ribosomal proteins (Makarova et al., 2001). Further comparison of archaeal and eukaryotic Nob1 sequences revealed two eukaryotic specific features: a large insertion at the end of the PIN domain and an elongated C-terminus (Figure 26). However, the archaeal Nob1 from *P. horikoshii* (PhNob1) exhibits a sequence similarity of 68 and 66% to Nob1 of *S. cerevisiae* and *H. sapiens*, respectively, in the conserved regions (Veith et al., 2012a).

The high sequence conservation of PhNob1 and eukaryotic Nob1 homologues prompted additional functional and structural studies of the archaeal protein. As a consequence of its modest size and its efficient expression in *E. coli* (performed in collaboration with the group of Prof. Dr. Enrico Schleiff), PhNob1 was perfectly suited for structural analysis by nuclear magnetic resonance spectroscopy (NMR; performed by the group of Prof. Dr. Jens Wöhnert; Veith et al., 2012b). The solved structure of PhNob1 confirmed the presence of a typical PIN domain that is connected by a flexible linker region with a zinc ribbon domain. The two domains were found to be structurally independent. Additionally, the presence of the metal ions manganese and zinc was confirmed for the PIN and the zinc ribbon domain, respectively. Both domains of PhNob1 show affinity for RNA. Consistent with previous crosslinking results obtained in yeast (Granneman et al., 2010), the zinc ribbon domain was found to bind the H40 hairpin of the *P. horikoshii* 18S rRNA sequence with high affinity (Veith et al., 2012a).

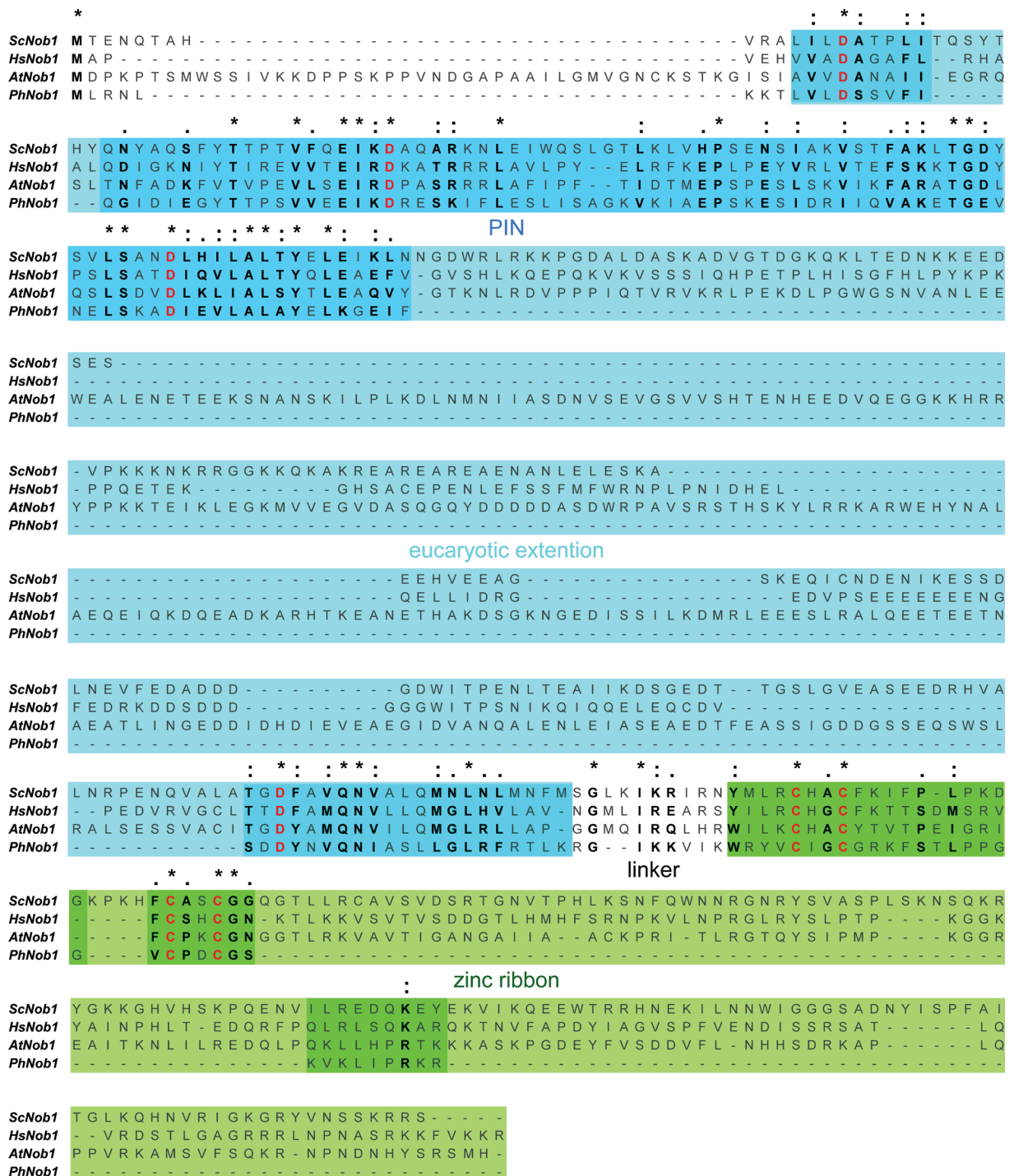


Figure 26: The amino acid sequence of Nob1 is conserved from archaea to eukaryotes. ClustalW alignment of Nob1 sequences of *Saccharomyces cerevisiae* (ScNob1, accession number: NC_001147), *Homo sapiens* (HsNob1, accession number: NC_000016), *Arabidopsis thaliana* (AtNob1, accession number: NC_003076) and *Pyrococcus horikoshii* (PhNob1, accession number: NP_142655) is shown. Conserved residues are highlighted in black. The degree of conservation is indicated by symbols above the alignment (* = high conservation, : = moderate conservation, . = low conservation). Residues of the PIN domain are highlighted in blue, residues of the zinc ribbon domain are marked in green. Light colours indicate residues present only in eukaryotic sequences. The highly conserved aspartate residues (D) of the PIN domain and cysteine residues (C) of the zinc ribbon domain suggested to be important for Nob1 function are highlighted in red.

In summary, the PIN and the zinc ribbon domains of the endonuclease Nob1 are highly conserved from archaea to eukaryotes. The expression of recombinant Nob1 from the thermophilic archaea *P. horikoshii* enabled the determination of the atomic structure of the endonuclease by NMR (Veith et al., 2012b). Further structural and biochemical analyses of our collaborators revealed coordination of metal ions and affinity for RNA of both independent domains, which are separated by a flexible linker (Veith et al., 2012a).

3.14 The function of Nob1 is conserved in archaea and eukaryotes

Based on the high sequence similarity of the eukaryotic and archaeal Nob1 proteins, it has been assumed that the function of the endonuclease is conserved. However, *in vitro* cleavage of model substrates was only demonstrated for ScNob1. In order to substantiate the general conservation of the Nob1 function in endonucleolytic cleavage at site D, the recombinantly expressed archaeal PhNob1 was used in *in vitro* cleavage assays (2.8). An internally radiolabelled, *in vitro* transcribed 90 nucleotide pre-rRNA fragment spanning the 3'-end of the 16S rRNA from *P. horikoshii* was used as a substrate. *In vitro* cleavage assays using recombinant PhNob1 yielded two prominent cleavage products and a decrease of full-length transcript suggesting that PhNob1 possesses a specific endonuclease activity (Figure 27 A, labels: 1 and 2). Increasing the protein concentration demonstrated that the cleavage of the RNA is specific and depends on the amount of PhNob1.

To determine the precise location of the cleavage site, the *in vitro* cleavage reaction was also performed using a 5'-labelled pre-rRNA fragment, so only one 5' cleavage product was detectable (Figure 27 A, label: 3). The amount of cleaved RNA again depended on the concentration of PhNob1 used in the assay. In order to map the 3'-end of the cleavage product, a plasmid containing the sequence of the transcribed pre-rRNA fragment was used as template for the generation of a sequencing ladder. Comparison of the 5'-labelled cleavage product with the sequencing ladder revealed that Nob1 cleaved specifically at site D and generated a fragment that corresponds to the 3'-end of 16S rRNA (Figure 27 B, label: 3; Kawarabayasi et al., 1998).

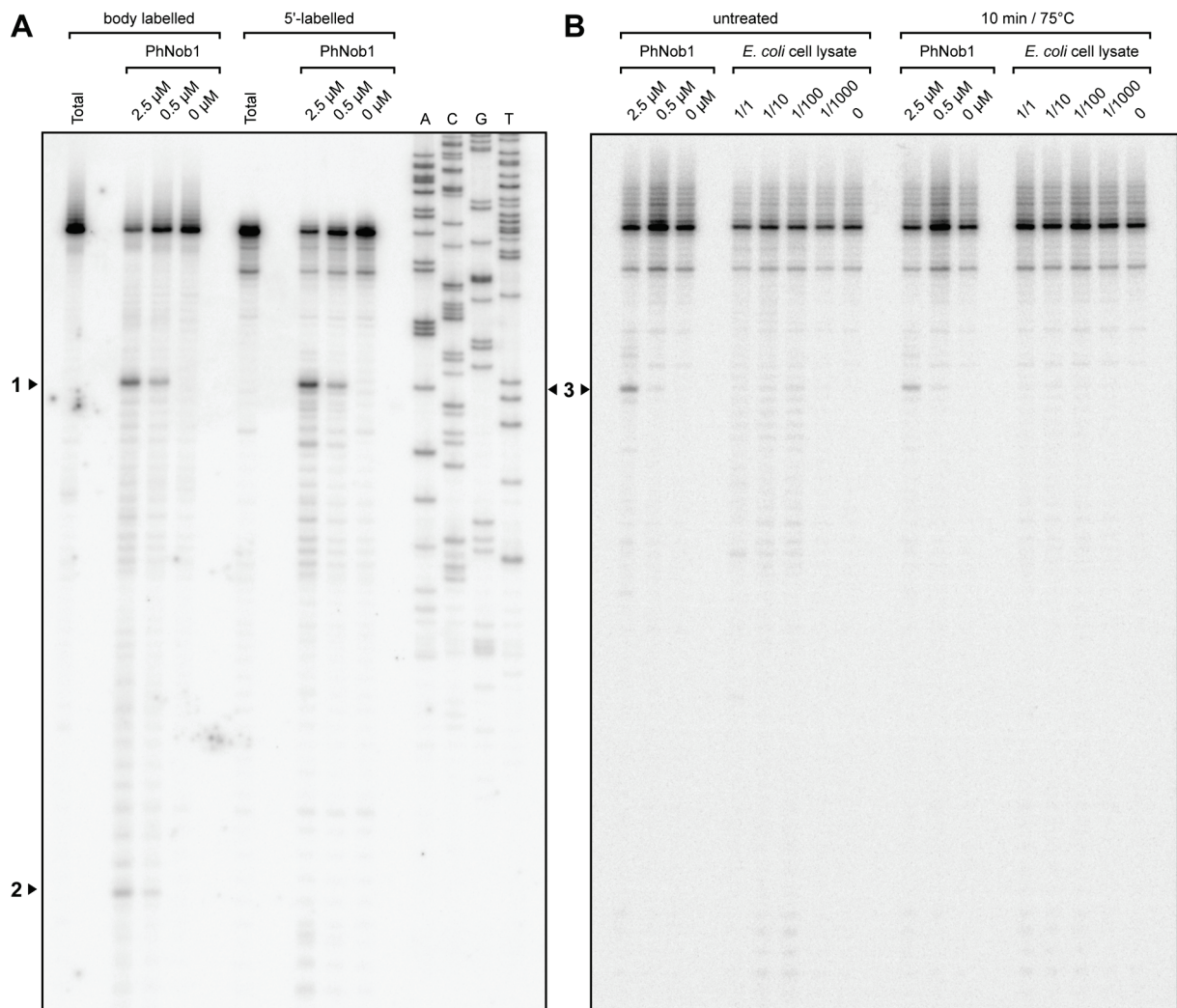


Figure 27: Recombinant Nob1 from *Pyrococcus horikoshii* specifically cleaves pre-rRNA at the processing site D *in vitro*. PhNob1 was recombinantly expressed in *E. coli* cells, purified and incubated with a synthesised precursor fragment of *P. horikoshii* 16S rRNA harbouring the cleavage site D. Samples were separated according to their size on a denaturing polyacrylamide gel. Gels were dried and radioactively labelled RNA was visualised with a phosphorimager. **(A)** Body labelled (left panel) and 5'-end labelled (right panel), *in vitro* transcribed RNA was incubated with different concentrations of PhNob1 as indicated. Samples with untreated substrate RNA were loaded as control and are indicated (Total). Cleavage products are labelled with numbers (1-3). The sequencing ladder of the substrate RNA (A,C,G,T) allows mapping of the precise cleavage position to the 3'-end of 16S rRNA. **(B)** 5'-labelled and *in vitro* transcribed substrate RNA was incubated with PhNob1 or various dilutions of *E. coli* cell lysates without (untreated) or with prior heat treatment (10 min / 75 °C) of the samples. The cleavage product is labelled (3).

To exclude the possibility that the cleavage products arise from copurification of *E. coli* nucleases with PhNob1, 5'-end labelled substrate RNAs were incubated with various concentrations of *E. coli* cell lysate. The treatment did not produce any specific cleavage product and indicates that the fragments generated in presence of PhNob1 are indeed the result of a cleavage reaction performed by the endonuclease. The background cleavage activity of the extract decreased noticeably after heat treatment for 10 min at 75 °C. However, the thermo-stable PhNob1 protein was not

affected by this treatment and retained its specific cleavage activity (Gonzalez et al., 1998). These observations led to the conclusion that the Nob1 cleavage of the synthesised RNA substrates at site D is performed in a highly specific manner (Figure 27 B, label: 3).

The coordination of the divalent cation manganese was shown to be crucial for the catalytic activity of ScNob1 (Pertschy et al., 2009). In order to analyse the conservation of the manganese-dependent cleavage mechanism for PhNob1, *in vitro* cleavage assays were performed in absence or in presence of manganese. The substitution of manganese with another divalent cation (magnesium) completely blocked PhNob1 cleavage (Figure 28 A). In contrast, the presence of 5 mM manganese chloride in the reaction buffer resulted in the presence of the two specific cleavage products (Figure 27; Figure 28, labels: 1 and 2). This finding suggests that the manganese-dependent Nob1 cleavage mechanism is conserved from archaea to eukaryotes.

In yeast, the processing of 20S pre-rRNA to mature 18S rRNA occurs in a single endonucleolytic cleavage event catalysed by Nob1 (Pertschy et al., 2009). To demonstrate that the fragments generated by PhNob1 are not the result of exonucleolytic digestion but rather arise from its ability to perform endonucleolytic cleavage reactions, further *in vitro* experiments were performed. For this, a 17-nucleotide long fragment of the PhNob1 target site covering the 3'-end of 16S rRNA was fused to two artificial RNA stems (Figure 28 B and C). A highly stable secondary structure at the hairpin base was introduced to protect the substrate ends from exonucleolytic digestion (Pertschy et al., 2009). Application of PhNob1 to the artificial substrate RNAs resulted in a site-specific and concentration-dependent cleavage reaction producing two fragments of distinct length (Figure 28 D, labels: 1 and 2). This suggests that PhNob1 is indeed able to perform an endonucleolytic cleavage reaction *in vitro*.

Chemical shifts in NMR spectra upon manganese or substrate RNA addition identified the PIN domain as the catalytic centre of the endonuclease PhNob1 (Veith et al., 2012a). Based on this structure, residues were suggested to be important for mediating the interaction of PhNob1 with RNA (S79, D100 and R115) or proposed to be involved in the coordination of manganese (D12, and D100; Figure 29 A). Experimental validation of these interpretations was accomplished using

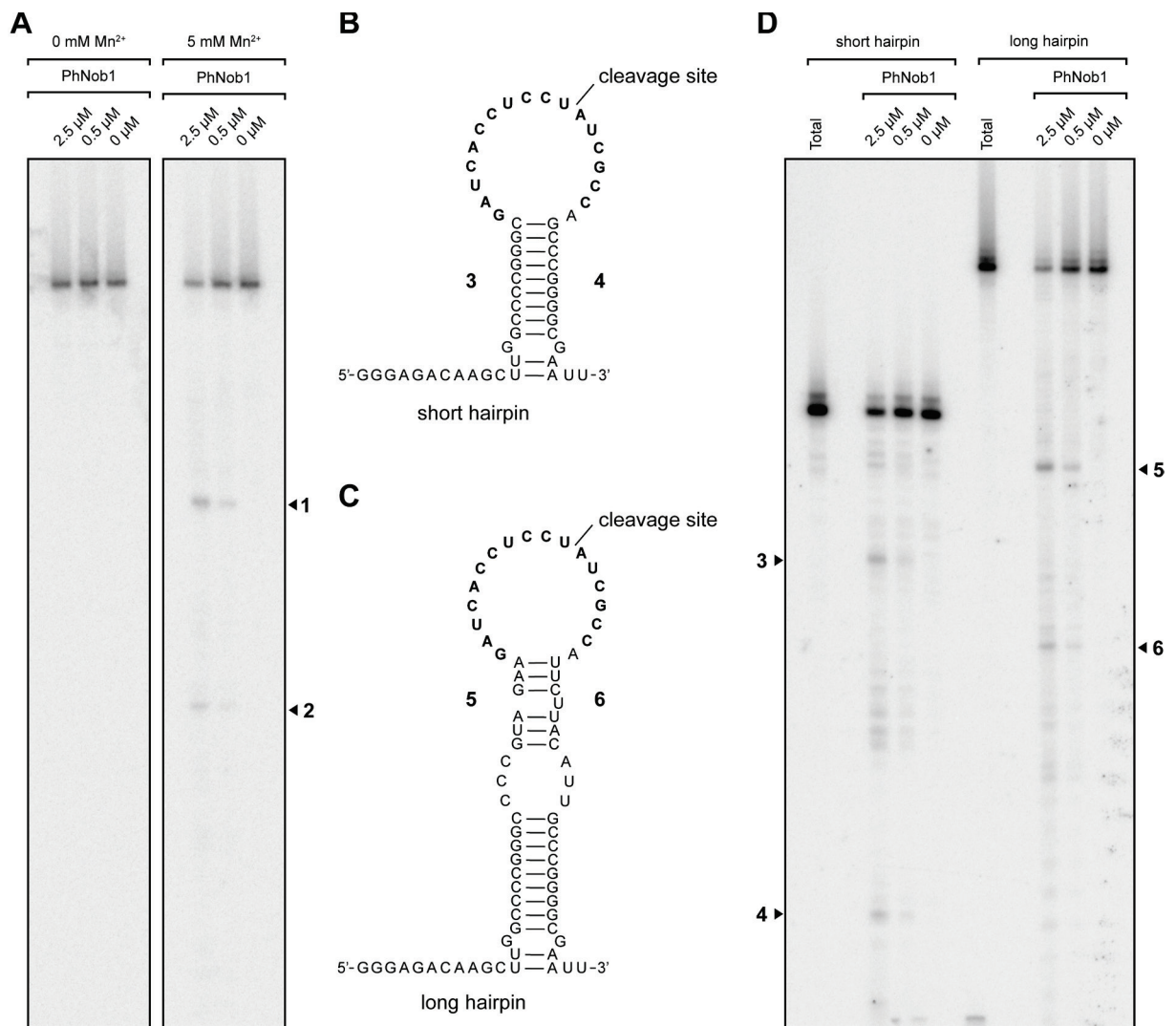


Figure 28: The cleavage mechanism of PhNob1 is conserved from archaea to eukaryotes. PhNob1 was recombinantly expressed in *E. coli* cells, purified and incubated with a synthesised fragment of *P. horikoshii* 16S pre-rRNA harbouring the cleavage site D. Samples were separated according to their size on a denaturing polyacrylamide gel. Gels were dried and radioactively labelled RNA was visualised with a phosphorimager. **(A)** Body labelled substrate RNA was incubated with recombinantly expressed PhNob1 in absence (0 mM Mn²⁺; left panel) or presence (5 mM Mn²⁺; right panel) of manganese chloride. Specific cleavage products are indicated (1-2). **(B and C)** Secondary structure predictions of hairpin constructs used to test the endonucleolytic cleavage activity of PhNob1. Sequences from *P. horikoshii* 16S pre-rRNA including the cleavage site D are indicated as bold letters. The precise position of the PhNob1 cleavage site is indicated. Cleavage products are indicated by bold numbers (3-6). **(D)** Body labelled *in vitro* transcribed substrate hairpin constructs (left: short hairpin, right: long hairpin) were incubated with different concentrations of PhNob1 as specified. Samples with untreated substrate RNA were loaded as control and are indicated (Total). Cleavage products are numbered as in B and C.

PhNob1 mutants carrying single amino acid substitutions at potentially critical sites. Mutants were recombinantly expressed in *E. coli* and tested for their functional activity *in vitro* (Figure 29 B). Replacement of aspartate at position 12 by an asparagine (D12N) resulted in a complete loss of catalytic activity. This residue belongs to a group of aspartate residues (D12, D46, D82 and D100), which are

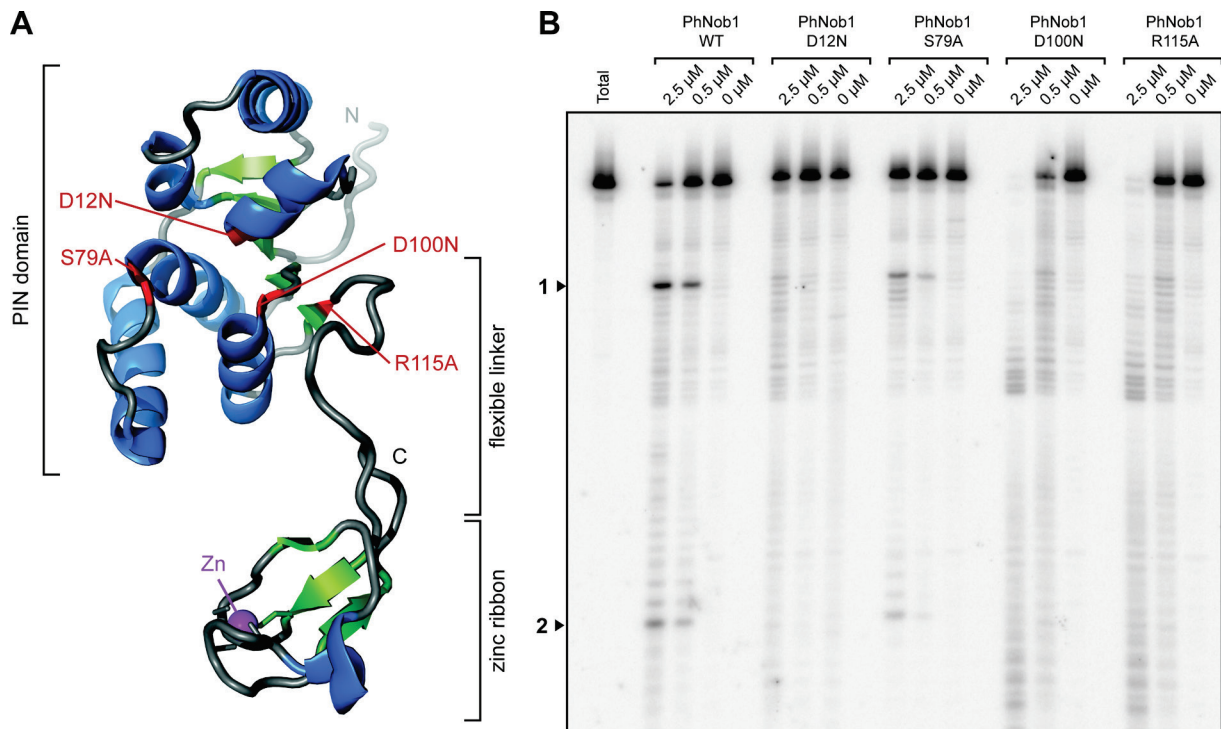


Figure 29: Structural analysis of PhNob1 by NMR identified critical residues for its endonuclease function. (A) The 3D structure of PhNob1 (PDB entry: 2LCQ), which was determined by nuclear magnetic resonance spectroscopy (NMR) in collaboration with the group of Prof. Dr. Jens Wöhnert (Goethe-University, Frankfurt am Main) is shown (Veith et al., 2012a). Secondary structure motifs are coloured (α -helices: blue; β -sheets: green). Critical residues for PhNob1 function, identified by NMR shift analysis are highlighted in red. The domain organisation is illustrated in brackets. The zinc ion (Zn), which is coordinated by residues of the zinc ribbon domain, is shown in purple. The amino- and carboxyl- termini of the amino acid sequence are given (N; C). **(B)** Body labelled substrate RNA was incubated with different concentrations of recombinantly expressed and purified PhNob1 point mutants as indicated. Samples were separated according to their size on a denaturing polyacrylamide gel. Gels were dried and radioactively labelled RNA was visualised with a phosphorimager. Applied concentrations and precise locations of amino acid substitutions are given (D12N, S79A, D100N and R115A). Samples with untreated substrate RNA were loaded as control and are indicated (Total). Cleavage products are numbered (1-2).

highly conserved from archaea to eukaryotes. Consistent with this, mutagenesis of the corresponding residue (D15N) in Nob1 from *S. cerevisiae* was previously reported to abolish the cleavage activity of the endonuclease (Pertschy et al., 2009). A significant decrease in the amount of substrate cleaved by PhNob1 was also observed after replacement of serine at position 79 by alanine (S79A). This finding is consistent with the proposed role of this residue in RNA binding (Veith et al., 2012a). Interestingly, the full-length substrate was cleaved by PhNob1 containing either a substitution of aspartate for asparagine at position 100 (D100N) or replacement of arginine by alanine at position 115 (R115A), but no longer in a site-specific manner. This suggests a function for both amino acids not only in binding of the substrate but also in recognition of the target sequence and the correct positioning of the RNA

substrate with respect to the catalytic centre of the endonuclease (Veith et al., 2012a).

3.15 Nob1 is structurally and functionally conserved from archaea to eukaryotes.

In order to investigate the functional conservation of Nob1 in higher eukaryotes Nob1 from *A. thaliana* was recombinantly expressed and purified (Missbach et al., 2013). AtNob1 was incubated with a radiolabelled fragment of *A. thaliana* 20S pre-rRNA including the cleavage site D. The *in vitro* cleavage reaction produced two RNA fragments, which accumulated in an AtNob1 concentration dependent manner (Figure 30, labels: 1 and 2). However, recombinant AtNob1, in which the aspartate at position 50 was substituted by an asparagine residue (D50N), did not generate any products in the *in vitro* cleavage assay. This finding validates the effect of the corresponding mutation in *S. cerevisiae* (D15N) and *P. horikoshii* (D12N), which also caused a complete loss of the Nob1 cleavage activity. Since recombinant expression and purification of the D50N mutant yielded low amounts of protein, the maximal applied concentration was limited to only 5 μ M AtNob1-D50N. In comparison to the archaeal PhNob1, significantly higher concentrations of AtNob1 were required to achieve detectable amounts of cleavage product. The reduced activity of the recombinant eukaryotic Nob1 homologues might be the result of a less stable protein after the purification procedure but could potentially be caused by additional

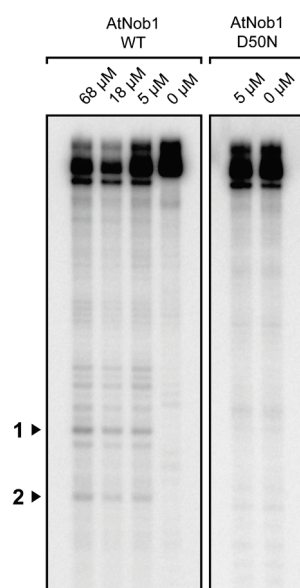


Figure 30: *A. thaliana* Nob1 is able to perform a specific cleavage reaction at site D. Wild type (WT; left panel) and mutant (D50N; right panel) AtNob1 were recombinantly expressed in *E. coli* cells and purified (Missbach et al., 2013). Recombinant AtNob1 proteins were incubated with an *in vitro* transcribed and radioactively body labelled fragment of *A. thaliana* 18S pre-rRNA, harbouring the cleavage site D. Samples were separated according to their size on a denaturing polyacrylamide gel. Gels were dried and radioactively labelled RNA was visualised with a phosphor-imager. The different concentrations of AtNob1 are given. Cleavage products are labelled by numbers (1-2).

regulating mechanisms, which are mediated by the eukaryotic insertions in the PIN domain. Nevertheless, the results strongly suggest, that AtNob1 is able to perform the specific cleavage reaction at site D and that the function of conserved residues in the Nob1 PIN domain is maintained from archaea to higher eukaryotes.

Analysis of PhNob1 activity *in vitro* showed that the enzyme is able to perform a site-specific endonucleolytic cleavage of a pre-rRNA transcript at site D. Conservation of the site-specific cleavage reaction was demonstrated not only for *P. horikoshii* Nob1, but also for Nob1 from *A. thaliana*. Yeast and PhNob1 enzyme activity depends on the complexation of a manganese ion. The NMR structure of PhNob enabled the identification of residues important for substrate binding and coordination of the manganese cation. *In vitro* experiments using PhNob1 variants with single amino acid substitutions of these critical residues confirmed their influence on the endonuclease activity. In conclusion, these findings suggest a general conservation of the Nob1 function and cleavage mechanism from archaea to eukaryotes in processing pre-rRNA at site D and suggest that the solved structure of PhNob1 is structurally and functionally highly comparable to its eukaryotic homologue.

4 Discussion

The functions of RNA helicases in the maturation of ribosomes are poorly characterised mostly since the target sites of these enzymes on pre-rRNA are still unknown. The crosslinking and analysis of cDNA approach (CRAC) used in this study revealed the first RNA helicase binding sites on eukaryotic pre-rRNAs providing the basis for a detailed functional characterisation of Rok1 and Prp43 in ribosome biogenesis.

4.1 Characterising RNA helicase function using CRAC analysis

Next generation sequencing approaches combined with new bioinformatic tools for mapping of the acquired data (Granneman et al., 2009; Simm, 2013) have been important technical innovations and are fundamental for the CRAC method (Bohnsack et al., 2012) used in this study. Moreover, a recently developed computational analysis (CLASH; Kudla et al., 2011; Travis et al., 2014) now enables the identification of RNA-RNA interactions within CRAC data sets. Such additional information about RNA contacts at the binding sites of RNA helicases are especially useful for understanding the roles of these enzymes in the remodelling of RNA-containing complexes.

In the cases of Prp43 and Rok1, genome wide mapping of the obtained CRAC data showed that pre-rRNAs are the primary target of both RNA helicases. However, both RNA helicases were also found to crosslink specifically to snoRNAs, which also participate in the biogenesis of ribosomes by guiding nucleotide modifications or by performing essential structural functions required for pre-rRNA processing. Interestingly, Rok1 associated preferentially to snoRNAs involved in early pre-rRNA processing events (snR10, snR30, U3 and U14; Table 6 and Figure 13), whereas Prp43 crosslinking enriched box C/D snoRNAs, which guide 2'-O-methylations in the LSU (Table 7).

4.2 A pre-ribosomal interaction network regulated by Rok1

CRAC analysis and mapping of crosslinking data on the mature ribosomal structures revealed that the Rok1 targets several sequences that cluster on one face of the SSU body (Figure 11). A major Rok1 target site was identified in ES6H3 (Figure 10)

and verified as bona fide binding site by DMS structure probing analysis (Figure 12). The binding of Rok1 to ES6H3 is also consistent with the known role of the helicase in release of snR30, which basepairs with this part of the 18S rRNA precursors (Bohnsack et al., 2008).

4.2.1 Rok1 cooperates with protein cofactors of ribosome biogenesis

Interestingly, a study performed in parallel showed that like Rok1, the ribosome biogenesis cofactor Rrp7 also interacts with ES6H3 (Lin et al., 2013). Rrp7 is required for the initial rRNA processing steps (A_0 - A_2 ; Baudin-Baillieu et al., 1997) and assembles early to the pre-ribosomal particles as part of the UTP-C SSU processome subcomplex (Grandi et al., 2002; Krogan et al., 2004). Additionally, Rrp7 was found to crosslink to snR30 and requires the snoRNA for its recruitment to 90S particles, but depletion of Rrp7 does not influence the association of snR30 with pre-ribosomes (Lin et al., 2013).

A second common crosslinking site of Rok1 and Rrp7 is ES7, which is positioned between the two ribosomal proteins Rps13 and Rps14 in the mature SSU structure (Figure 11; Ben-Shem et al., 2011). Both ribosomal proteins were found to be important for the early processing events of 18S rRNA precursors (Ferreira-Cerca et al., 2005) and their depletion results in a decreased association of Rok1 and Rrp7 with pre-ribosomal complexes (Jakob et al., 2012). Remarkably, ES7 contains a novel snR30 basepairing site identified by CLASH analysis of the Rok1 crosslinking data (Figure 14).

These data suggest a model in which Rrp7 and Rok1 act sequentially. Initially, snR30 basepairs with ES6H3 and ES7 to bring these two regions of the pre-ribosome together and to enable the recruitment of Rrp7 (Figure 31). Following release of Rrp7, Rok1 is able to interact with ES6H3 and ES7 and mediates the release of snR30 (Bohnsack et al., 2008). The release of snR30 is a prerequisite for the formation of interactions between ES6H3 and ES3 that are found in the mature SSU (Figure 11; Alkemar and Nygard, 2004). Interestingly, Rok1 but not Rrp7 was observed to crosslink to ES3 (Lin et al., 2013), which is consistent with the presence of Rrp7 in earlier pre-ribosomal particles in which ES6H3 is likely associated with ES7 rather than with ES3. Possibly, Rok1 even participates actively in establishing

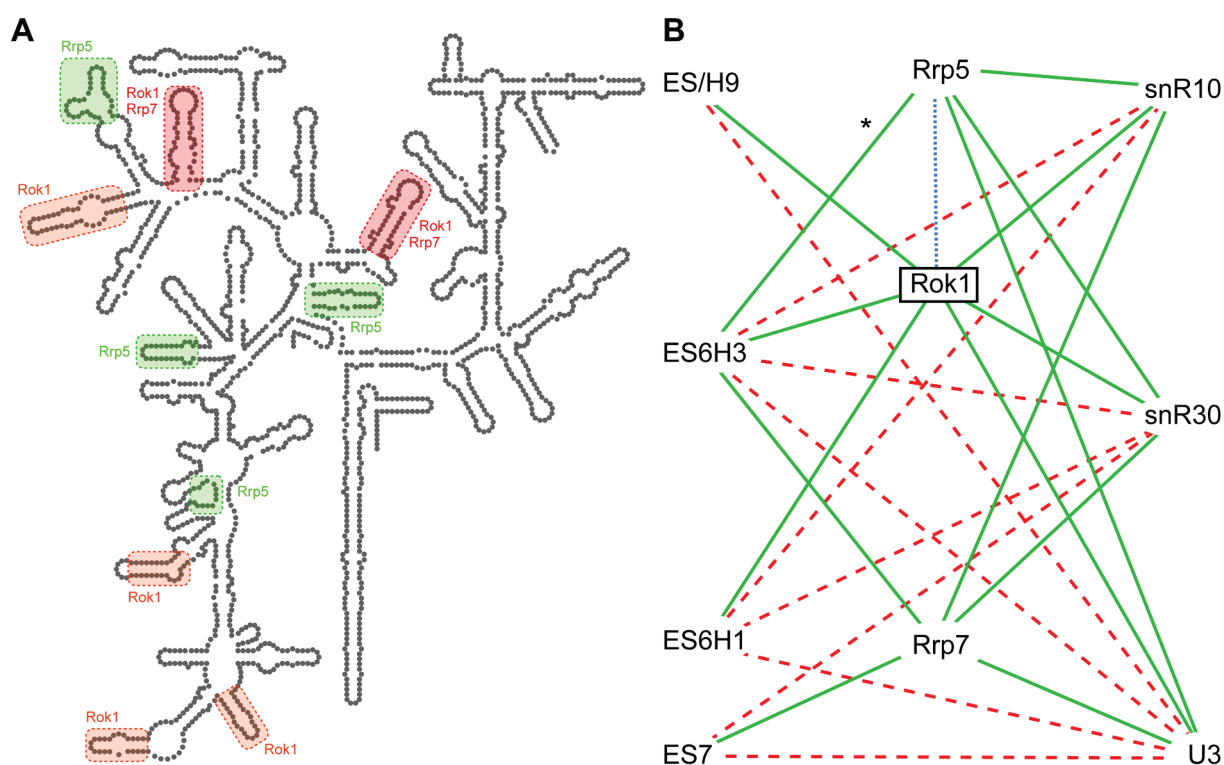


Figure 31: Interaction network of 40S pre-ribosomal maturation cofactors. (A) Binding sites of maturation cofactors Rok1, Prp43, Rrp5 (Lebaron et al., 2013) and Rrp7 (Lin et al., 2013) were mapped on the secondary structure of the 18S rRNA (Petrov et al., 2013). (B) Interactions between pre-rRNA elements (left), protein maturation cofactors (middle) and snoRNAs in early pre-40S particles are shown. The dotted blue line indicates a direct protein-protein interaction; solid green lines represent RNA-protein interactions identified by CRAC and dashed red lines mark snoRNA-pre-rRNA interactions detected by CLASH. The asterisk indicates that Rrp5 interacts with an 18S rRNA sequence adjacent to but not within ES6H3.

the interaction between ES6H3 and ES3 since Rok1 was shown to possess ADP dependent annealing activity *in vitro* (Young et al., 2013).

The *in vitro* annealing activity of Rok1 is stimulated by the ribosome maturation factor Rrp5, which was previously reported to physically and genetically interact with the RNA helicase (Lebaron et al., 2013; Torchet et al., 1998; Venema and Tollervey, 1996; Young et al., 2013). In the pre-ribosome, Rrp5 provides a structural framework that is important for the coordination of pre-rRNA cleavage events, the compaction of early pre-ribosomal particles and the formation of the characteristic terminal knobs seen in chromatin spreads (Lebaron et al., 2013; Miller and Beatty, 1969; Venema and Tollervey, 1996). Recently, Rrp5 was found to crosslink to several sequences within the 35S primary transcript including sequences in the 18S rRNA, ITS1, 25S rRNA and 5.8S rRNA (Lebaron et al., 2013). Interestingly, one Rrp5 crosslinking site is located close to the Rok1 binding site in ES6H3 (Figure 31 A) and Rrp5 and Rok1

crosslink significantly to the same subset of snoRNAs, which are all involved in processing of pre-rRNA (snR10, snR30, U3 and U14; Figure 31 B) The collocation of the Rok1 and Rrp5 crosslinking sites is consistent with the previously reported interactions of Rok1 with Rrp5 (Torchet et al., 1998) and the regulation of the RNA helicase by Rrp5 *in vitro* (Young et al., 2013)

4.2.2 Rok1 is associated with snoRNAs involved in pre-rRNA processing

As well as protein cofactors like Rrp5, snoRNAs have also been proposed to be involved in coordination of pre-ribosome assembly and remodelling events. For example the U3 snoRNA basepairs with sequences of the 5'-ETS and the 18S rRNA sequence and these long-range interactions are thought to facilitate coordinated pre-rRNA processing and prevent premature formation of the central pseudoknot structure in the SSU (Beltrame and Tollervey, 1992; Beltrame and Tollervey, 1995; Dutca et al., 2011; Hughes and Ares, 1991). However, such snoRNA-rRNA basepairing interactions are often short, making their identification challenging (Beltrame and Tollervey, 1992; Beltrame and Tollervey, 1995; Dutca et al., 2011; Kudla et al., 2011; Liang et al., 2010; Morrissey and Tollervey, 1993). Using the CLASH analysis enabled the discovery of such previously unknown basepairing sites for snR30, snR10, U3 and U14 in the 18S rRNA sequence.

At a length of over 600 nucleotides snR30 is the longest snoRNA identified in *S. cerevisiae* (Bally et al., 1988). However, only two short sequence elements (m1 and m2) near the 3'-end of the snoRNA have been proposed to basepair with the 18S rRNA sequence (Atzorn et al., 2004; Beltrame and Tollervey, 1992; Beltrame and Tollervey, 1995; Dutca et al., 2011; Fayet-Lebaron et al., 2009; Hughes, 1996). The data presented here demonstrate that the 5'-part of snR30 can form basepairing interactions with ES7 (Figure 14) implying that the snoRNA might form a bridge between this region and ES6H3 in the pre-ribosome (Figure 31). Evolutionary conservation of both sequences involved in this interaction in other yeasts such as *Saccharomyces pombe* supports ES7 as a novel basepairing site of snR30 (Martin et al., 2014). However, no hybrids between ES3 and snR30 were found in the CLASH analysis supporting the model that snR30 is absent when the ES3-ES6 interaction, found in the mature SSU, is established.

Besides snR30, Rok1 was found to associate also with the snoRNA snR10 (Figure 15). Interestingly, Rrp5 was recently shown to contact sequences directly adjacent to the Rok1 crosslinking site near a 7-nucleotide processing element on snR10 (Lebaron et al., 2013) again suggesting cooperation between these two factors. Additionally, snR10 was reported to strongly crosslink to Rrp7 leading to the suggestion of a novel snR10 basepairing site near the Rrp7 target site, in ES6 (Lin et al., 2013). This hypothesis is supported by the results of the Rok1 CLASH analysis described here, which revealed chimeric snR10-ES6 reads (Figure 15). Similar to the known basepairing interactions of snR30 and U3 (Atzorn et al., 2004; Beltrame and Tollervey, 1992; Beltrame and Tollervey, 1995; Dutca et al., 2011; Fayet-Lebaron et al., 2009; Hughes, 1996) these potential duplexes between snR10 and ES6 appear to be short since their calculated free energy is quite low.

CLASH analysis of Rok1 crosslinking data also uncovered novel basepairing interactions of U3 with sequences in ES6, ES3 and H9 (Figure 16). Interestingly, these interactions are mediated by sequences in the 3'-domain of the U3 snoRNA, indicating that this region might not only constitute a platform for protein binding (Granneman et al., 2009; Mereau et al., 1997; Samarsky and Fournier, 1998). This observation is consistent with a recent study that discovered a basepairing interaction involving sequences close to the 3'-end of U3 and ES7 (Kudla et al., 2011).

The results obtained by CLASH analysis indicate that the amount of long-range interactions in pre-ribosomal complexes formed by snoRNAs and required for pre-rRNA processing is higher than previously assumed (Kudla et al., 2011). Together with the data provided by Lin et al. (2013) and Lebaron et al. (2013) these findings uncovered a dynamic pre-ribosomal interaction network of snoRNAs, protein cofactors and pre-rRNA sequences that coordinate remodelling and processing of the pre-SSU particles. In more detail the data presented here suggest a model in which ES6 and ES7 are structurally linked when Rrp7 is present in the pre-ribosomal particles. Dissociation of Rrp7 allows Rok1 to access its target site and the RNA helicase unwinds the duplex formed by snR30 and ES6H3. After the release of snR30, Rok1 might cooperate with snR10 and U3 in tethering ES3 and ES6 together. As an ATP-dependent RNA helicase, Rok1 is likely to be the active regulator of these events.

Interestingly, eukaryotic expansion segments seem to be important elements of the pre-ribosomal interaction network described here since they contain the major binding sites of protein cofactors (Figure 10; Figure 31 A; Lebaron et al., 2013; Lin et al., 2013) as well as multiple basepairing sites of the snoRNAs involved in pre-rRNA processing (Figure 14; Figure 15; Figure 16). The function of most of these sequences has remained elusive so far, but the data presented here together with recent studies from Bradatsch et al. (2012) and Lin et al. (2013) suggest that eukaryotic expansion segments constitute important elements during ribosome assembly by coordinating pre-ribosome compaction (Lebaron et al., 2013), pre-rRNA folding and pre-rRNA processing (Figure 31 B). It is therefore likely that eukaryotic expansion segments evolved due to the greater complexity of eukaryotic ribosome biogenesis.

4.3 The role of RNA helicases in releasing snoRNAs from pre-ribosomes

The key function of Rok1 is the release of snR30 from pre-ribosomes (Figure 13; Bohnsack et al., 2008). In general, snoRNAs, especially of the box C/D class, form stable basepairing interactions with pre-rRNA and it is thought that the activity of RNA helicases may be required to displace snoRNAs from their pre-rRNA targets.

Indeed, several RNA helicases have been reported to be required for the release of mostly single snoRNAs. Has1 and Dbp4 are required for the release of U14 from pre-ribosomal particles (Kos and Tollervey, 2005; Liang and Fournier, 2006), whereas Has1 and Rok1 are implicated in snR30 dissociation (Bohnsack et al., 2008; Kos and Tollervey, 2005). Both snoRNAs exhibit multiple contacts with the 18S rRNA sequence (Figure 15; Figure 17; Fayet-Lebaron et al., 2009; Liang and Fournier, 1995) suggesting that these RNA helicases might cooperate to mediate release of the snoRNAs by unwinding of the individual basepairing sites. However, sequential targeting or partial redundancy of these RNA helicases is also possible and may explain the relatively low number of snoRNAs that significantly accumulate on pre-ribosomal complexes upon depletion of a single RNA helicase (Bohnsack et al., 2008).

In this study, a qPCR-based analysis of snoRNA levels on pre-ribosomes also uncovered a role for the DEAH box helicase Prp43 in the release of several snoRNAs from pre-LSU complexes (U18, snR39, snR39b, snR50, snR72 and snR60;

Figure 18). Mapping of the crosslinking sites on the mature secondary (Petrov et al., 2013) and tertiary (Ben-Shem et al., 2011) LSU structures showed that these snoRNAs guide modifications clustering next to the Prp43 interaction sites identified by CRAC (Figure 21; Figure 22). However, some snoRNAs that guide modifications within this region of the 25S rRNA sequence do not show retention on pre-ribosomal particles upon Prp43 depletion or mutation of critical residues in its helicase domain indicating that the functions of Prp43 require its catalytic activity and are specific for a particular subset of snoRNAs rather than caused by a general defect in LSU maturation (Bohnsack et al., 2009).

Unlike DEAD box helicases such as Rok1, which perform local strand unwinding close to their target sites, the DEAH box helicase Prp43 was proposed to possess processive activity allowing the helicase to translocate along the substrate RNA (Walbott et al., 2010). However, it is unlikely that Prp43 translocates across the snoRNA cluster in one step since some basepairing sites of the affected snoRNAs partially overlap suggesting that these snoRNAs bind sequentially and need to be released in several rounds by the RNA helicase (Piekna-Przybylska et al., 2007; van Nues et al., 2011; Watkins and Bohnsack, 2012). In between release of particular snoRNAs, an additional RNA binding domain of Prp43, which was found within its C-terminus, might allow stable association to the substrate complexes (Walbott et al., 2010). Thus, the individual pre-rRNA crosslinking peaks might reflect recruitment or pausing sites of Prp43. An overview of the operating range of the RNA helicase was obtained by mapping of these crosslinking sites onto the 3D structure (Ben-Shem et al., 2011) of the LSU (Figure 22).

Although RNA helicases (Rok1, Prp43, Has1 and Dbp4) were found to be required for the release of snoRNAs from pre-ribosomes, a formal demonstration of their participation in snoRNA-pre-rRNA duplex unwinding is still lacking. However, the CLASH analysis applied here provides experimental evidence of direct RNA-RNA interactions close to the protein binding sites. Notably, CLASH hybrids containing sequences of snR30 and ES6H3 were identified for Rok1 (Figure 14). A direct role of Prp43 in snoRNA release was indicated by a detailed analysis of Prp43 crosslinking profiles of four affected snoRNAs (snR39, snR50, snR64 and snR72) showing that Prp43 contacts the sequences of the snoRNAs known to mediate the basepairing interactions with the 25S pre-rRNA (Figure 23). Consistent with this, close

examination of the corresponding sequences in the 25S rRNA sequence revealed also several small crosslinking peaks directly overlapping with the snoRNA basepairing sites (Figure 20). In fact, chimeric reads consisting of snoRNA and their targeted 25S rRNA sequences were found for virtually all snoRNAs that were affected by Prp43 depletion (Figure 24). The identification of CLASH hybrids in both Rok1 and Prp43 CRAC data therefore strongly supports a model of direct snoRNA-pre-rRNA duplex unwinding by these RNA helicases.

Taken together, the identification of RNA helicase binding sites by CRAC combined with functional qPCR analysis allowed the identification of novel snoRNA substrates for the DEAH box helicase Prp43. Furthermore, CLASH analysis provided strong evidence for a direct release of snoRNAs by the unwinding activity of the RNA helicases Prp43 and Rok1.

4.4 Structural remodelling of RNP complexes by RNA helicases

Besides functioning in the release of snoRNAs from pre ribosomal particles, RNA helicases have been shown to catalyse structural rearrangements of the RNP complexes by controlling the folding of RNA, mediating protein dissociation or functioning as RNA clamps (reviewed in Jarmoskaite and Russell, 2014). However, in ribosome biogenesis only a few examples involving RNA helicases in such remodelling tasks have been proposed so far.

Besides the implications of Rok1 in formation of the ES6H3-ES3 basepairing interaction discussed above (4.2), the results obtained by CRAC and qPCR analyses also suggest a role in remodelling of pre-ribosomal particles for Prp43. Mapping of the Prp43 crosslinking sites on the secondary structure model of the mature 25S rRNA demonstrated that the crosslinking sites in H61, H80 and H84 are close to modification sites guided by snR64 and snR67 (Figure 21 E and F), which show decreased association with pre-ribosomal particles after depletion of Prp43 (Figure 18). These findings can be explained by a model in which Prp43 is required for structural rearrangements of the regions targeted by these snoRNAs allowing them to access their basepairing sites. Consistent with this, depletion of Prp43 was previously reported to cause defects in the 2'-O-methylation guided by snR64 (Leeds et al., 2006).

In the LSU, Prp43 was also observed to crosslink to H23 of the 25S rRNA sequence (Figure 21 E). H23 is located within domain I of the 25S rRNA sequence, which has previously been shown to undergo structural rearrangements, enabling the formation of the 25S-5.8S basepairing interaction found in the mature LSU (Ben-Shem et al., 2011; Georgiev et al., 1984). The crosslinking data therefore indicate that Prp43 might also play a role in this remodelling process similar to the RNA helicase Has1 (Dembowski et al., 2013).

Beside its function in LSU biogenesis, Prp43 is additionally implicated in the maturation of the SSU (Combs et al., 2006; Lebaron et al., 2005; Leeds et al., 2006). A prominent crosslinking peak of Prp43 in H44 of the 18S rRNA sequence identified by CRAC analysis (Figure 25) supports a role of the RNA helicase in this pathway. Interestingly, H44 is located close to the 3'-end of the 18S rRNA, which is generated by Nob1 in an endonucleolytic cleavage reaction of the 20S pre-rRNA at site D. This, together with co-immunoprecipitation of 20S pre-rRNA (Combs et al., 2006; Lebaron et al., 2005; Leeds et al., 2006) and genetic interactions with Nob1, the pre-40S maturation factor Ltv1 and the Prp43 cofactor Pfa1 (Pertschy et al., 2009) lead to the suggestion that Prp43 is implicated in restructuring this region to enable the final processing of 20S pre-rRNA by Nob1. Further evidence for the presence of Prp43 in this region was provided by analysis of a Prp43-Ltv1 double mutant, which produces an aberrant 17S rRNA that is generated by 3'-5' exonucleolytic degradation of the 18S 3'-minor domain (Pertschy et al., 2009). Furthermore, both the export adapter Ltv1 (Seiser et al., 2006) and the G-patch protein Pfa1 were shown to directly interact with and stimulate the activity of Prp43 (Lebaron et al., 2009) suggesting that they cooperate with the RNA helicase to promote the processing of 20S rRNA. Interestingly, the upper part of H44, which is crosslinked by Prp43, together with H45 forms the decoding centre of the mature 40S ribosome, which is essential for protein translation (Ben-Shem et al., 2011). Possibly, Prp43 together with Pfa1 and Ltv1 is involved in regulating the maturation of this sub-structure (Pertschy et al., 2009). Additional evidence for a significant reorganisation of this pre-ribosomal region was provided by cryo-EM analysis showing that the upper part of H44 is disordered in late pre-40S complexes when compared to the mature structure of the SSU (Strunk et al., 2011).

The final processing of the 20S pre-rRNA by the endonuclease Nob1 at site D occurs in late cytoplasmic pre-40S particles (Zemp et al., 2009). Structural remodelling of the 3'-minor domain might be a necessary regulation mechanism to prevent the cleavage of 20S pre-rRNA in early pre-ribosomal complexes, to which Nob1 is recruited (Fatica et al., 2003). The 3'-domain of the 18S rRNA sequence was therefore suggested to adopt a structural conformation that blocks access of the Nob1 PIN domain to its target site (Lamanna and Karbstein, 2011; Pertschy et al., 2009). Indeed, the structural conformation of H44 and ITS1, which flank the Nob1 cleavage site, has been reported to influence the binding of the endonuclease (Lamanna and Karbstein, 2011). An additional Nob1 interaction site, which anchors the endonuclease to the pre-SSU particles, was identified in H40 of the 18S rRNA sequence (Granneman et al., 2010). This binding site is conserved from archaea to eukaryotes and potentially constitutes the primary Nob1 recruitment site (Veith et al., 2012a). Cryo-EM analysis of pre-40S complexes from *S. cerevisiae* revealed that ScNob1 is positioned between the platform, which harbours the cleavage site D and the neck of these particles, where H40 is located (Strunk et al., 2011). A model, which illustrates the structure of PhNob1 bound to the mature 40S ribosome of yeast, confirms that the functional core of the endonuclease is sufficient to span the distance between the two interaction sites (Figure 32; Veith et al., 2012a). The conserved aspartate residues (D12, D82 and D100) of the PIN domain point to the 3'-end of the 18S rRNA, whereas the β -sheet surface of the zinc ribbon domain is adjacent to H40 and thereby anchors the endonuclease to the head domain of the ribosomal particle (Figure 32; Veith et al., 2012a). Moreover, the structural model indicates that the linker sequence connecting the two functional domains of Nob1 provides the required structural flexibility to allow the simultaneous binding of Nob1 to H40 and the processing at site D in late pre-ribosomal particles.

Remodelling of pre-SSU and pre-LSU particles add to the role of Prp43 in mediating snoRNA release from pre-ribosomes. Beside its functions in ribosome biogenesis, Prp43 was found implicated in the disassembly of late and aberrant spliceosomal particles (Arenas and Abelson, 1997; Fourmann et al., 2013; Koodathingal et al., 2010; Martin et al., 2002; Mayas et al., 2010), demonstrating the ability of the multifunctional RNA helicase to regulate remodelling events in RNP complexes.

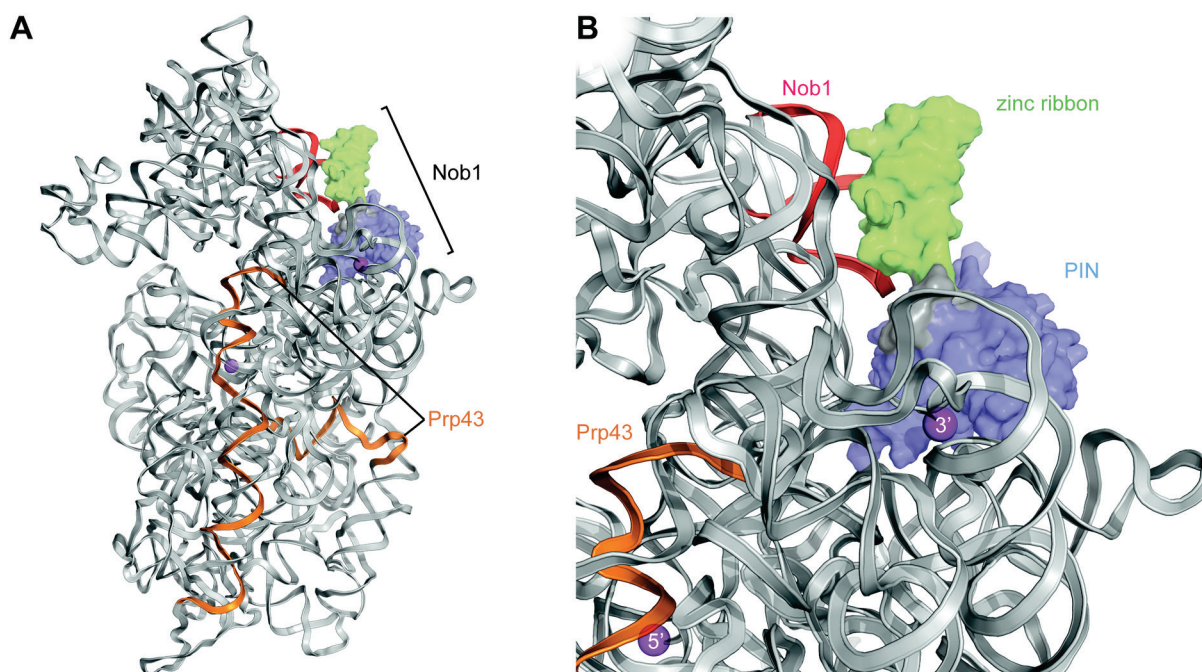


Figure 32: 3D Model of Nob1 binding to the 40S ribosome structure. (A) The 40S 3D structure (Ben-Shem et al., 2010) is shown from the side of the subunit interface. Crosslinking sites of Prp43 (orange) and Nob1 (red; Granneman et al., 2010) are indicated on the back bone of the rRNA. The rRNA ends are highlighted as purple spheres. Nob1 was modelled on the 40S structure and is shown in surface representation. Surface colours indicate the two functional domains of the endonuclease. Green denotes the zinc ribbon domain, the blue surface marks the PIN domain and the flexible linker is shown in grey. **(B)** A detailed section of the 40S platform domain illustrating the potential interaction of Nob1 PIN domain with the 3'-end of 18S rRNA and the zinc ribbon domain with H44 (Nob1 crosslinking site).

Likely, some of the other RNA helicases involved in ribosome biogenesis, whose functions are currently unknown, also participate in remodelling of pre-ribosomal complexes. Investigating the reorganisation of pre-ribosomal particles during their course of maturation is now facilitated by the development of high throughput structure probing techniques such as SHAPE-Seq, Mod-seq or ChemModSeq (Hector et al., 2014; Mortimer et al., 2012; Talkish et al., 2014). These methods are based on the DMS structure probing approach used here, but coupling with next generation sequencing allows screening of entire pre-rRNAs. Such methods have recently been applied to investigate the structures of cellular mRNAs (Rouskin et al., 2014) or to follow structural rearrangements during the late steps of SSU maturation and will in future give substantial insights into rearrangements that occur during ribosome biogenesis and help uncovering the functions of further RNA helicases in these remodelling events.

4.5 RNA helicases are regulated by cofactors

RNA helicases frequently function in large complexes such as pre-ribosomes and many of them require protein cofactors to regulate their activity or aid in target recognition (Silverman et al., 2003). Especially in the complex process of ribosome biogenesis a careful regulation of maturation cofactors, like RNA helicases, seems to be crucial. However, relatively few RNA helicase cofactors have been identified and characterised. As mentioned above (4.2) Rok1 was shown to interact with Rrp5, a large scaffold protein containing several RNA binding domains (Venema and Tollervey, 1996), which was suggested to stimulate the enzymatic activity of Rok1 (Young et al., 2013). Additionally, the close proximity of their pre-rRNA and snoRNA binding sites supports the coordinated function of Rrp5 and Rok1 *in vivo* (Figure 31 A; Lebaron et al., 2013).

Similarly for Dbp8, another DEAD box helicase involved in 40S biogenesis, (Daugeron and Linder, 2001) a ribosome maturation cofactor with a conserved RNA binding domain, Esf2 (Hoang et al., 2005), was shown to bind and stimulate the activity of the RNA helicase (Granneman et al., 2006).

Interestingly, Prp43 has been shown to be regulated by several cofactors, all containing a G-patch domain with several conserved glycine residues (Aravind and Koonin, 1999). Initially, the G-patch protein Ntr1 was identified as cofactor of Prp43 during splicing (Tsai et al., 2005), in which Prp43 facilitates the release of the lariat structure and drives disassembly of the spliceosome (Martin et al., 2002). In biogenesis of the 40S ribosome, Prp43 was shown to interact with the G-patch protein Pfa1 to coordinate structural remodelling processes that facilitate the final processing of 20S pre-rRNA (4.4; Pertschy et al., 2009). Recently, Chen et al. (2014) confirmed the direct interaction of Prp43 with a third G-patch protein, Gno1, and suggested that the protein regulates the functions of the RNA helicase in maturation events occurring in late 90S an early pre-60S ribosomal particles. In all these cases, the interaction of Prp43 with the G-patch proteins is mediated by the G-patch motif of the cofactor (Behrens et al., 2012; Chen et al., 2014; Christian et al., 2014; Walbott et al., 2010) and results in stimulation of the ATPase or helicase activity *in vitro* (Chen et al., 2014; Lebaron et al., 2009; Tanaka et al., 2007). Since Pfa1 and Ntr1 both bind the C-terminal domain of Prp43 (Christian et al., 2014; Walbott et al., 2010) it is likely that the G-patch proteins compete for Prp43. Possibly, G-patch proteins help to

recruit the multifunctional RNA helicase to its diverse substrates throughout the cell. Such regulatory mechanisms could contribute to the synchronisation of ribosome biogenesis with other cellular pathways. However, G-patch proteins might also exhibit local regulation mechanisms, which allow multitasking of Prp43 within pre-ribosomal complexes, since no physical interactions between the G-patch proteins Pfa1 and Gno1 were detected (Chen et al., 2014). Further studies into the regulation of Prp43 and its cofactors will likely provide more insights into the regulation mechanisms that allow participation of the helicase in multiple cellular pathways and potentially lead to the discovery of new Prp43 substrates.

4.6 Conclusions

In summary, the data presented here enabled the identification of the first RNA helicase binding sites on pre-rRNA and allowed the functional characterisation of Rok1 and Prp43. Moreover, the crosslinking analysis confirmed for the first time immediate contacts of RNA helicases with snoRNA-pre-rRNA duplexes *in vivo* providing strong evidence for a direct release of specific snoRNAs by the helicase activity of these enzymes. Together, these findings give new insights in pre-ribosomal structures formed by pre-rRNAs and snoRNAs and uncovered potential RNA elements that undergo helicase mediated rearrangements within these particles.

Literature

- Alexander, R.D., Barrass, J.D., Dichtl, B., Kos, M., Obtulowicz, T., Robert, M.C., Koper, M., Karkusiewicz, I., Mariconti, L., Tollervey, D., Dichtl, B., Kufel, J., Bertrand, E. and Beggs, J.D. **(2010)** RiboSys, a high-resolution, quantitative approach to measure the in vivo kinetics of pre-mRNA splicing and 3'-end processing in *Saccharomyces cerevisiae*. RNA, 16: 2570-2580.
- Alkemar, G. and Nygard, O. **(2003)** A possible tertiary rRNA interaction between expansion segments ES3 and ES6 in eukaryotic 40S ribosomal subunits. RNA, 9: 20-24.
- Alkemar, G. and Nygard, O. **(2004)** Secondary structure of two regions in expansion segments ES3 and ES6 with the potential of forming a tertiary interaction in eukaryotic 40S ribosomal subunits. RNA, 10: 403-411.
- Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E. and Tollervey, D. **(1999)** Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J, 18: 5399-5410.
- Aravind, L. and Koonin, E.V. **(1999)** G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. Trends Biochem Sci, 24: 342-344.
- Arenas, J.E. and Abelson, J.N. **(1997)** Prp43: An RNA helicase-like factor involved in spliceosome disassembly. Proc Natl Acad Sci U S A, 94: 11798-11802.
- Armache, J.P., Jarasch, A., Anger, A.M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S., Marquez, V., Mielke, T., Thomm, M., Berninghausen, O., Beatrix, B., Soding, J., Westhof, E., Wilson, D.N. and Beckmann, R. **(2010)** Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-A resolution. Proc Natl Acad Sci U S A, 107: 19748-19753.
- Arnez, J.G. and Steitz, T.A. **(1994)** Crystal structure of unmodified tRNA(Gln) complexed with glutamyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. Biochemistry, 33: 7560-7567.
- Atzorn, V., Fragapane, P. and Kiss, T. **(2004)** U17/snoR30 is a ubiquitous snoRNA with two conserved sequence motifs essential for 18S rRNA production. Mol Cell Biol, 24: 1769-1778.
- Bachellet, J.P., Cavaille, J. and Huttenhofer, A. **(2002)** The expanding snoRNA world. Biochimie, 84: 775-790.
- Bagni, C. and Lapeyre, B. **(1998)** Gar1p binds to the small nucleolar RNAs snR10 and snR30 in vitro through a nontypical RNA binding element. J Biol Chem, 273: 10868-10873.
- Balakin, A.G., Smith, L. and Fournier, M.J. **(1996)** The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. Cell, 86: 823-834.
- Bally, M., Hughes, J. and Cesareni, G. **(1988)** SnR30: a new, essential small nuclear RNA from *Saccharomyces cerevisiae*. Nucleic Acids Res, 16: 5291-5303.
- Baudin-Baillieu, A., Tollervey, D., Cullin, C. and Lacroute, F. **(1997)** Functional analysis of Rrp7p, an essential yeast protein involved in pre-rRNA processing and ribosome assembly. Mol Cell Biol, 17: 5023-5032.

- Baudin-Baillieu, A., Fabret, C., Liang, X.H., Piekna-Przybylska, D., Fournier, M.J. and Rousset, J.P. **(2009)** Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. Nucleic Acids Res, 37: 7665-7677.
- Behrens, M.A., He, Y., Oliveira, C.L., Andersen, G.R., Pedersen, J.S. and Nielsen, K.H. **(2012)** Structural analysis of RNA helicases with small-angle X-ray scattering. Methods Enzymol, 511: 191-212.
- Beltrame, M. and Tollervey, D. **(1992)** Identification and functional analysis of two U3 binding sites on yeast pre-ribosomal RNA. EMBO J, 11: 1531-1542.
- Beltrame, M., Henry, Y. and Tollervey, D. **(1994)** Mutational analysis of an essential binding site for the U3 snoRNA in the 5' external transcribed spacer of yeast pre-rRNA. Nucleic Acids Res, 22: 5139-5147.
- Beltrame, M. and Tollervey, D. **(1995)** Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. EMBO J, 14: 4350-4356.
- Ben-Shem, A., Jenner, L., Yusupova, G. and Yusupov, M. **(2010)** Crystal structure of the eukaryotic ribosome. Science, 330: 1203-1209.
- Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G. and Yusupov, M. **(2011)** The structure of the eukaryotic ribosome at 3.0 Å resolution. Science, 334: 1524-1529.
- Birnboim, H.C. and Doly, J. **(1979)** A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res, 7: 1513-1523.
- Bleichert, F. and Baserga, S.J. **(2007)** The long unwinding road of RNA helicases. Mol Cell, 27: 339-352.
- Bohnsack, M.T., Kos, M. and Tollervey, D. **(2008)** Quantitative analysis of snoRNA association with pre-ribosomes and release of snR30 by Rok1 helicase. EMBO Rep, 9: 1230-1236.
- Bohnsack, M.T., Martin, R., Granneman, S., Ruprecht, M., Schleiff, E. and Tollervey, D. **(2009)** Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. Mol Cell, 36: 583-592.
- Bohnsack, M.T., Tollervey, D. and Granneman, S. **(2012)** Identification of RNA helicase target sites by UV cross-linking and analysis of cDNA. Methods Enzymol, 511: 275-288.
- Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P. and Boeke, J.D. **(1998)** Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast, 14: 115-132.
- Bradatsch, B., Leidig, C., Granneman, S., Gnadig, M., Tollervey, D., Bottcher, B., Beckmann, R. and Hurt, E. **(2012)** Structure of the pre-60S ribosomal subunit with nuclear export factor Arx1 bound at the exit tunnel. Nat Struct Mol Biol, 19: 1234-1241.
- Briggs, M.W., Burkard, K.T. and Butler, J.S. **(1998)** Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. J Biol Chem, 273: 13255-13263.

- Brimacombe, R., Mitchell, P., Osswald, M., Stade, K. and Bochkariov, D. (1993) Clustering of modified nucleotides at the functional center of bacterial ribosomal RNA. FASEB J, 7: 161-167.
- Cahill, N.M., Friend, K., Speckmann, W., Li, Z.H., Terns, R.M., Terns, M.P. and Steitz, J.A. (2002) Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP. EMBO J, 21: 3816-3828.
- Charette, M. and Gray, M.W. (2000) Pseudouridine in RNA: what, where, how, and why. IUBMB Life, 49: 341-351.
- Chen, Y., Potratz, J.P., Tijerina, P., Del Campo, M., Lambowitz, A.M. and Russell, R. (2008) DEAD-box proteins can completely separate an RNA duplex using a single ATP. Proc Natl Acad Sci U S A, 105: 20203-20208.
- Chen, Y.L., Capeyrou, R., Humbert, O., Mouffok, S., Kadri, Y.A., Lebaron, S., Henras, A.K. and Henry, Y. (2014) The telomerase inhibitor Gno1p/PINX1 activates the helicase Prp43p during ribosome biogenesis. Nucleic Acids Res, 42: 7330-7345.
- Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem, 162: 156-159.
- Chomczynski, P. and Sacchi, N. (2006) The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. Nat Protoc, 1: 581-585.
- Christian, H., Hofele, R.V., Urlaub, H. and Ficner, R. (2014) Insights into the activation of the helicase Prp43 by biochemical studies and structural mass spectrometry. Nucleic Acids Res, 42: 1162-1179.
- Chu, S., Archer, R.H., Zengel, J.M. and Lindahl, L. (1994) The RNA of RNase MRP is required for normal processing of ribosomal RNA. Proc Natl Acad Sci U S A, 91: 659-663.
- Ciganda, M. and Williams, N. (2011) Eukaryotic 5S rRNA biogenesis. Wiley Interdiscip Rev RNA, 2: 523-533.
- Combs, D.J., Nagel, R.J., Ares, M., Jr. and Stevens, S.W. (2006) Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis. Mol Cell Biol, 26: 523-534.
- Daugeron, M.C. and Linder, P. (2001) Characterization and mutational analysis of yeast Dbp8p, a putative RNA helicase involved in ribosome biogenesis. Nucleic Acids Res, 29: 1144-1155.
- Davis, D.R. (1995) Stabilization of RNA stacking by pseudouridine. Nucleic Acids Res, 23: 5020-5026.
- Decatur, W.A. and Fournier, M.J. (2002) rRNA modifications and ribosome function. Trends Biochem Sci, 27: 344-351.
- Deley, J. (1964) Sedimentation Coefficients of Yeast Ribosomes. J Gen Microbiol, 37: 153-156.
- Dembowski, J.A., Kuo, B. and Woolford, J.L., Jr. (2013) Has1 regulates consecutive maturation and processing steps for assembly of 60S ribosomal subunits. Nucleic Acids Res, 41: 7889-7904.

- Dever, T.E. and Green, R. **(2012)** The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb Perspect Biol, 4: a013706.
- Dragon, F., Gallagher, J.E., Compagnone-Post, P.A., Mitchell, B.M., Porwancher, K.A., Wehner, K.A., Wormsley, S., Settlage, R.E., Shabanowitz, J., Osheim, Y., Beyer, A.L., Hunt, D.F. and Baserga, S.J. **(2002)** A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature, 417: 967-970.
- Dunbar, D.A. and Baserga, S.J. **(1998)** The U14 snoRNA is required for 2'-O-methylation of the pre-18S rRNA in *Xenopus* oocytes. RNA, 4: 195-204.
- Dutca, L.M., Gallagher, J.E. and Baserga, S.J. **(2011)** The initial U3 snoRNA:pre-rRNA base pairing interaction required for pre-18S rRNA folding revealed by in vivo chemical probing. Nucleic Acids Res, 39: 5164-5180.
- Faber, A.W., Van Dijk, M., Raue, H.A. and Vos, J.C. **(2002)** Ngl2p is a Ccr4p-like RNA nuclease essential for the final step in 3'-end processing of 5.8S rRNA in *Saccharomyces cerevisiae*. RNA, 8: 1095-1101.
- Fairman-Williams, M.E., Guenther, U.P. and Jankowsky, E. **(2010)** SF1 and SF2 helicases: family matters. Curr Opin Struct Biol, 20: 313-324.
- Fatica, A., Oeffinger, M., Dlakic, M. and Tollervey, D. **(2003)** Nob1p is required for cleavage of the 3' end of 18S rRNA. Mol Cell Biol, 23: 1798-1807.
- Fayet-Lebaron, E., Atzorn, V., Henry, Y. and Kiss, T. **(2009)** 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. EMBO J, 28: 1260-1270.
- Ferreira-Cerca, S., Poll, G., Gleizes, P.E., Tschochner, H. and Milkereit, P. **(2005)** Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. Mol Cell, 20: 263-275.
- Fourmann, J.B., Schmitzova, J., Christian, H., Urlaub, H., Ficner, R., Boon, K.L., Fabrizio, P. and Luhrmann, R. **(2013)** Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. Genes Dev, 27: 413-428.
- Fromont-Racine, M., Senger, B., Saveanu, C. and Fasiolo, F. **(2003)** Ribosome assembly in eukaryotes. Gene, 313: 17-42.
- Gallagher, J.E., Dunbar, D.A., Granneman, S., Mitchell, B.M., Osheim, Y., Beyer, A.L. and Baserga, S.J. **(2004)** RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. Genes Dev, 18: 2506-2517.
- Ganot, P., Bortolin, M.L. and Kiss, T. **(1997a)** Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. Cell, 89: 799-809.
- Ganot, P., Caizergues-Ferrer, M. and Kiss, T. **(1997b)** The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. Genes Dev, 11: 941-956.

- Garcia, I., Albring, M.J. and Uhlenbeck, O.C. (2012) Duplex destabilization by four ribosomal DEAD-box proteins. Biochemistry, 51: 10109-10118.
- Geerlings, T.H., Vos, J.C. and Raue, H.A. (2000) The final step in the formation of 25S rRNA in *Saccharomyces cerevisiae* is performed by 5'→3' exonucleases. RNA, 6: 1698-1703.
- Georgiev, O.I., Dudov, K.P., Hadjiolov, A.A. and Skryabin, K.G. (1984) Evidence for interaction of 5.8S rRNA with the 5'- and 3'- terminal segments of *Saccharomyces cerevisiae* 25S rRNA. Folia Biol (Praha), 30: 1-14.
- Gerbi, S.A. (1986) The evolution of eukaryotic ribosomal DNA. Biosystems, 19: 247-258.
- Gietz, D., St Jean, A., Woods, R.A. and Schiestl, R.H. (1992) Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res, 20: 1425.
- Gigova, A., Duggimpudi, S., Pollex, T., Schaefer, M. and Kos, M. (2014) A cluster of methylations in the domain IV of 25S rRNA is required for ribosome stability. RNA, 20: 1632-1644.
- Gonzalez, J.M., Masuchi, Y., Robb, F.T., Ammerman, J.W., Maeder, D.L., Yanagibayashi, M., Tamaoka, J. and Kato, C. (1998) *Pyrococcus horikoshii* sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. Extremophiles, 2: 123-130.
- Gorbalenya, A.E. and Koonin, E.V. (1993) Helicases: amino acid sequence comparisons and structure-function relationships. Current Opinion in Structural Biology, 3: 419-429.
- Grandi, P., Rybin, V., Bassler, J., Petfalski, E., Strauss, D., Marzioch, M., Schafer, T., Kuster, B., Tschochner, H., Tollervey, D., Gavin, A.C. and Hurt, E. (2002) 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. Mol Cell, 10: 105-115.
- Granneman, S. and Baserga, S.J. (2005) Crosstalk in gene expression: coupling and co-regulation of rDNA transcription, pre-ribosome assembly and pre-rRNA processing. Curr Opin Cell Biol, 17: 281-286.
- Granneman, S., Lin, C., Champion, E.A., Nandineni, M.R., Zorca, C. and Baserga, S.J. (2006) The nucleolar protein Esf2 interacts directly with the DExD/H box RNA helicase, Dbp8, to stimulate ATP hydrolysis. Nucleic Acids Res, 34: 3189-3199.
- Granneman, S., Kudla, G., Petfalski, E. and Tollervey, D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. Proc Natl Acad Sci U S A, 106: 9613-9618.
- Granneman, S., Petfalski, E., Swiatkowska, A. and Tollervey, D. (2010) Cracking pre-40S ribosomal subunit structure by systematic analyses of RNA-protein cross-linking. EMBO J, 29: 2026-2036.
- Guglielmi, B. and Werner, M. (2002) The yeast homolog of human PinX1 is involved in rRNA and small nucleolar RNA maturation, not in telomere elongation inhibition. J Biol Chem, 277: 35712-35719.

- Hector, R.D., Burlacu, E., Aitken, S., Bihan, T.L., Tuijtel, M., Zaplatina, A., Cook, A.G. and Granneman, S. **(2014)** Snapshots of pre-rRNA structural flexibility reveal eukaryotic 40S assembly dynamics at nucleotide resolution. Nucleic Acids Res, 42: 12138-12154.
- Held, W.A., Mizushima, S. and Nomura, M. **(1973)** Reconstitution of Escherichia coli 30 S ribosomal subunits from purified molecular components. J Biol Chem, 248: 5720-5730.
- Helm, M. **(2006)** Post-transcriptional nucleotide modification and alternative folding of RNA. Nucleic Acids Res, 34: 721-733.
- Henn, A., Cao, W., Licciardello, N., Heitkamp, S.E., Hackney, D.D. and De La Cruz, E.M. **(2010)** Pathway of ATP utilization and duplex rRNA unwinding by the DEAD-box helicase, DbpA. Proc Natl Acad Sci U S A, 107: 4046-4050.
- Henras, A.K., Soudet, J., Gerus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A. and Henry, Y. **(2008)** The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell Mol Life Sci, 65: 2334-2359.
- Henras, A.K., Plisson-Chastang, C., O'Donohue, M.F., Chakraborty, A. and Gleizes, P.E. **(2014)** An overview of pre-ribosomal RNA processing in eukaryotes. Wiley Interdiscip Rev RNA,
- Henry, Y., Wood, H., Morrissey, J.P., Petfalski, E., Kearsey, S. and Tollervey, D. **(1994)** The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. EMBO J, 13: 2452-2463.
- Hilbert, M., Karow, A.R. and Klostermeier, D. **(2009)** The mechanism of ATP-dependent RNA unwinding by DEAD box proteins. Biol Chem, 390: 1237-1250.
- Hoang, T., Peng, W.T., Vanrobays, E., Krogan, N., Hiley, S., Beyer, A.L., Osheim, Y.N., Greenblatt, J., Hughes, T.R. and Lafontaine, D.L. **(2005)** Esf2p, a U3-associated factor required for small-subunit processome assembly and compaction. Mol Cell Biol, 25: 5523-5534.
- Hoareau-Aveilla, C., Fayet-Lebaron, E., Jady, B.E., Henras, A.K. and Kiss, T. **(2012)** Utp23p is required for dissociation of snR30 small nucleolar RNP from preribosomal particles. Nucleic Acids Res, 40: 3641-3652.
- Houseley, J. and Tollervey, D. **(2010)** Apparent non-canonical trans-splicing is generated by reverse transcriptase in vitro. PLoS One, 5: e12271.
- Hughes, J.M. and Ares, M., Jr. **(1991)** Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. EMBO J, 10: 4231-4239.
- Hughes, J.M. **(1996)** Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA. J Mol Biol, 259: 645-654.
- Hunter, J.D. **(2007)** Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering, 9: 90-95.
- Jakob, S., Ohmayer, U., Neueder, A., Hierlmeier, T., Perez-Fernandez, J., Hochmuth, E., Deutzmann, R., Griesenbeck, J., Tschochner, H. and Milkereit, P. **(2012)** Interrelationships between yeast

- ribosomal protein assembly events and transient ribosome biogenesis factors interactions in early pre-ribosomes. PLoS One, 7: e32552.
- Jankowsky, E. (2011) RNA helicases at work: binding and rearranging. Trends Biochem Sci, 36: 19-29.
- Jarmolowski, A., Zagorski, J., Li, H.V. and Fournier, M.J. (1990) Identification of essential elements in U14 RNA of *Saccharomyces cerevisiae*. EMBO J, 9: 4503-4509.
- Jarmoskaite, I. and Russell, R. (2014) RNA helicase proteins as chaperones and remodelers. Annu Rev Biochem, 83: 697-725.
- Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K. and Kikuchi, H. (1998) Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3 (supplement). DNA Res, 5: 147-155.
- Kim, J., Jeon, S. and Yang, Y.S. (2004) Posttranscriptional regulation of the karyogamy gene by Kem1p/Xrn1p exoribonuclease and Rok1p RNA helicase of *Saccharomyces cerevisiae*. Biochem Biophys Res Commun, 321: 1032-1039.
- Kiss-Laszlo, Z., Henry, Y., Bachellerie, J.P., Caizergues-Ferrer, M. and Kiss, T. (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell, 85: 1077-1088.
- Kiss-Laszlo, Z., Henry, Y. and Kiss, T. (1998) Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. EMBO J, 17: 797-807.
- Klein, D.J., Schmeing, T.M., Moore, P.B. and Steitz, T.A. (2001) The kink-turn: a new RNA secondary structure motif. EMBO J, 20: 4214-4221.
- Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S. and Ban, N. (2011) Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. Science, 334: 941-948.
- Klinge, S., Voigts-Hoffmann, F., Leibundgut, M. and Ban, N. (2012) Atomic structures of the eukaryotic ribosome. Trends Biochem Sci, 37: 189-198.
- Koodathingal, P., Novak, T., Piccirilli, J.A. and Staley, J.P. (2010) The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. Mol Cell, 39: 385-395.
- Kos, M. and Tollervey, D. (2005) The Putative RNA Helicase Dbp4p Is Required for Release of the U14 snoRNA from Preribosomes in *Saccharomyces cerevisiae*. Mol Cell, 20: 53-64.
- Kos, M. and Tollervey, D. (2010) Yeast pre-rRNA processing and modification occur cotranscriptionally. Mol Cell, 37: 809-820.
- Kressler, D., Hurt, E. and Bassler, J. (2010) Driving ribosome assembly. Biochim Biophys Acta, 1803: 673-683.

- Kressler, D., Bange, G., Ogawa, Y., Stjepanovic, G., Bradatsch, B., Pratte, D., Amlacher, S., Strauss, D., Yoneda, Y., Katahira, J., Sinning, I. and Hurt, E. **(2012)** Synchronizing nuclear import of ribosomal proteins with ribosome assembly. Science, 338: 666-671.
- Krogan, N.J., Peng, W.T., Cagney, G., Robinson, M.D., Haw, R., Zhong, G., Guo, X., Zhang, X., Canadien, V., Richards, D.P., Beattie, B.K., Lalev, A., Zhang, W., Davierwala, A.P., Mnaimneh, S., Starostine, A., Tikuisis, A.P., Grigull, J., Datta, N., Bray, J.E., Hughes, T.R., Emili, A. and Greenblatt, J.F. **(2004)** High-definition macromolecular composition of yeast RNA-processing complexes. Mol Cell, 13: 225-239.
- Kruiswijk, T., Planta, R.J. and Krop, J.M. **(1978)** The course of the assembly of ribosomal subunits in yeast. Biochim Biophys Acta, 517: 378-389.
- Kudla, G., Granneman, S., Hahn, D., Beggs, J.D. and Tollervey, D. **(2011)** Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. Proc Natl Acad Sci U S A, 108: 10010-10015.
- Laemmli, U.K. **(1970)** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227: 680-685.
- Lafontaine, D.L., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. and Tollervey, D. **(1998)** The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. Genes Dev, 12: 527-537.
- Lamanna, A.C. and Karbstein, K. **(2011)** An RNA conformational switch regulates pre-18S rRNA cleavage. J Mol Biol, 405: 3-17.
- Lebaron, S., Froment, C., Fromont-Racine, M., Rain, J.C., Monsarrat, B., Caizergues-Ferrer, M. and Henry, Y. **(2005)** The splicing ATPase prp43p is a component of multiple preribosomal particles. Mol Cell Biol, 25: 9269-9282.
- Lebaron, S., Papin, C., Capeyrou, R., Chen, Y.L., Froment, C., Monsarrat, B., Caizergues-Ferrer, M., Grigoriev, M. and Henry, Y. **(2009)** The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis. EMBO J, 28: 3808-3819.
- Lebaron, S., Schneider, C., van Nues, R.W., Swiatkowska, A., Walsh, D., Bottcher, B., Granneman, S., Watkins, N.J. and Tollervey, D. **(2012)** Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. Nat Struct Mol Biol, 19: 744-753.
- Lebaron, S., Segerstolpe, A., French, S.L., Dudnakova, T., de Lima Alves, F., Granneman, S., Rappsilber, J., Beyer, A.L., Wieslander, L. and Tollervey, D. **(2013)** Rrp5 binding at multiple sites coordinates pre-rRNA processing and assembly. Mol Cell, 52: 707-719.
- Lee, S.J. and Baserga, S.J. **(1999)** Imp3p and Imp4p, two specific components of the U3 small nucleolar ribonucleoprotein that are essential for pre-18S rRNA processing. Mol Cell Biol, 19: 5441-5452.
- Leeds, N.B., Small, E.C., Hiley, S.L., Hughes, T.R. and Staley, J.P. **(2006)** The splicing factor Prp43p, a DEAH box ATPase, functions in ribosome biogenesis. Mol Cell Biol, 26: 513-522.

- Lemay, V., Hossain, A., Osheim, Y.N., Beyer, A.L. and Dragon, F. **(2011)** Identification of novel proteins associated with yeast snR30 small nucleolar RNA. Nucleic Acids Res, 39: 9659-9670.
- Li, B., Nierras, C.R. and Warner, J.R. **(1999)** Transcriptional elements involved in the repression of ribosomal protein synthesis. Mol Cell Biol, 19: 5393-5404.
- Li, L. and Ye, K. **(2006)** Crystal structure of an H/ACA box ribonucleoprotein particle. Nature, 443: 302-307.
- Liang, W.Q. and Fournier, M.J. **(1995)** U14 base-pairs with 18S rRNA: a novel snoRNA interaction required for rRNA processing. Genes Dev, 9: 2433-2443.
- Liang, X.H. and Fournier, M.J. **(2006)** The helicase Has1p is required for snoRNA release from pre-rRNA. Mol Cell Biol, 26: 7437-7450.
- Liang, X.H., Liu, Q., Liu, Q., King, T.H. and Fournier, M.J. **(2010)** Strong dependence between functional domains in a dual-function snoRNA infers coupling of rRNA processing and modification events. Nucleic Acids Res, 38: 3376-3387.
- Lin, J., Lu, J., Feng, Y., Sun, M. and Ye, K. **(2013)** An RNA-binding complex involved in ribosome biogenesis contains a protein with homology to tRNA CCA-adding enzyme. PLoS Biol, 11: e1001669.
- Lindahl, L., Bommankanti, A., Li, X., Hayden, L., Jones, A., Khan, M., Oni, T. and Zengel, J.M. **(2009)** RNase MRP is required for entry of 35S precursor rRNA into the canonical processing pathway. RNA, 15: 1407-1416.
- Linder, P. and Jankowsky, E. **(2011)** From unwinding to clamping - the DEAD box RNA helicase family. Nat Rev Mol Cell Biol, 12: 505-516.
- Liu, F., Putnam, A. and Jankowsky, E. **(2008)** ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. Proc Natl Acad Sci U S A, 105: 20209-20214.
- Loc'h, J., Blaud, M., Rety, S., Lebaron, S., Deschamps, P., Bareille, J., Jombart, J., Robert-Paganin, J., Delbos, L., Chardon, F., Zhang, E., Charenton, C., Tollervey, D. and Leulliot, N. **(2014)** RNA mimicry by the fap7 adenylate kinase in ribosome biogenesis. PLoS Biol, 12: e1001860.
- Lohman, T.M. and Bjornson, K.P. **(1996)** Mechanisms of helicase-catalyzed DNA unwinding. Annu Rev Biochem, 65: 169-214.
- Long, E.O. and Dawid, I.B. **(1980)** Repeated genes in eukaryotes. Annu Rev Biochem, 49: 727-764.
- Longtine, M.S., McKenzie, A., 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P. and Pringle, J.R. **(1998)** Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. Yeast, 14: 953-961.
- Lorenz, R., Bernhart, S.H., Honer Zu Siederdisen, C., Tafer, H., Flamm, C., Stadler, P.F. and Hofacker, I.L. **(2011)** ViennaRNA Package 2.0. Algorithms Mol Biol, 6: 26.
- Lu, J., Sun, M. and Ye, K. **(2013)** Structural and functional analysis of Utp23, a yeast ribosome synthesis factor with degenerate PIN domain. RNA, 19: 1815-1824.

- Lygerou, Z., Allmang, C., Tollervey, D. and Seraphin, B. **(1996)** Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP in vitro. Science, 272: 268-270.
- Makarova, K.S., Aravind, L., Galperin, M.Y., Grishin, N.V., Tatusov, R.L., Wolf, Y.I. and Koonin, E.V. **(1999)** Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. Genome Res, 9: 608-628.
- Makarova, K.S., Ponomarev, V.A. and Koonin, E.V. **(2001)** Two C or not two C: recurrent disruption of Zn-ribbons, gene duplication, lineage-specific gene loss, and horizontal gene transfer in evolution of bacterial ribosomal proteins. Genome Biol, 2: RESEARCH 0033.
- Martin, A., Schneider, S. and Schwer, B. **(2002)** Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. J Biol Chem, 277: 17743-17750.
- Martin, R., Straub, A.U., Doebele, C. and Bohnsack, M.T. **(2013)** DExD/H-box RNA helicases in ribosome biogenesis. RNA Biol, 10: 4-18.
- Martin, R., Hackert, P., Ruprecht, M., Simm, S., Bruning, L., Mirus, O., Sloan, K.E., Kudla, G., Schleiff, E. and Bohnsack, M.T. **(2014)** A pre-ribosomal RNA interaction network involving snoRNAs and the Rok1 helicase. RNA, 20: 1173-1182.
- Mayas, R.M., Maita, H., Semlow, D.R. and Staley, J.P. **(2010)** Spliceosome discards intermediates via the DEAH box ATPase Prp43p. Proc Natl Acad Sci U S A, 107: 10020-10025.
- Mereau, A., Fournier, R., Gregoire, A., Mougou, A., Fabrizio, P., Luhrmann, R. and Branlant, C. **(1997)** An in vivo and in vitro structure-function analysis of the *Saccharomyces cerevisiae* U3A snoRNP: protein-RNA contacts and base-pair interaction with the pre-ribosomal RNA. J Mol Biol, 273: 552-571.
- Miller, O.L., Jr. and Beatty, B.R. **(1969)** Visualization of nucleolar genes. Science, 164: 955-957.
- Missbach, S., Weis, B.L., Martin, R., Simm, S., Bohnsack, M.T. and Schleiff, E. **(2013)** 40S ribosome biogenesis co-factors are essential for gametophyte and embryo development. PLoS One, 8: e54084.
- Mitchell, P., Petfalski, E. and Tollervey, D. **(1996)** The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism. Genes Dev, 10: 502-513.
- Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. and Tollervey, D. **(1997)** The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases. Cell, 91: 457-466.
- Morrissey, J.P. and Tollervey, D. **(1993)** Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. Mol Cell Biol, 13: 2469-2477.
- Morrissey, J.P. and Tollervey, D. **(1997)** U14 small nucleolar RNA makes multiple contacts with the pre-ribosomal RNA. Chromosoma, 105: 515-522.
- Mortimer, S.A., Trapnell, C., Aviran, S., Pachter, L. and Lucks, J.B. **(2012)** SHAPE-Seq: High-Throughput RNA Structure Analysis. Curr Protoc Chem Biol, 4: 275-297.

- Mullineux, S.T. and Lafontaine, D.L. (2012) Mapping the cleavage sites on mammalian pre-rRNAs: where do we stand? Biochimie, 94: 1521-1532.
- Ni, J., Tien, A.L. and Fournier, M.J. (1997) Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. Cell, 89: 565-573.
- Nierhaus, K.H. and Dohme, F. (1974) Total reconstitution of functionally active 50S ribosomal subunits from Escherichia coli. Proc Natl Acad Sci U S A, 71: 4713-4717.
- Nissan, T.A., Bassler, J., Petfalski, E., Tollervey, D. and Hurt, E. (2002) 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. EMBO J, 21: 5539-5547.
- Oeffinger, M., Zenklusen, D., Ferguson, A., Wei, K.E., El Hage, A., Tollervey, D., Chait, B.T., Singer, R.H. and Rout, M.P. (2009) Rrp17p is a eukaryotic exonuclease required for 5' end processing of Pre-60S ribosomal RNA. Mol Cell, 36: 768-781.
- Oh, J.Y. and Kim, J. (1999) ATP hydrolysis activity of the DEAD box protein Rok1p is required for in vivo ROK1 function. Nucleic Acids Res, 27: 2753-2759.
- Perez-Fernandez, J., Roman, A., De Las Rivas, J., Bustelo, X.R. and Dosil, M. (2007) The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. Mol Cell Biol, 27: 5414-5429.
- Perez-Fernandez, J., Martin-Marcos, P. and Dosil, M. (2011) Elucidation of the assembly events required for the recruitment of Utp20, Imp4 and Bms1 onto nascent pre-ribosomes. Nucleic Acids Res, 39: 8105-8121.
- Pertschy, B., Schneider, C., Gnadig, M., Schafer, T., Tollervey, D. and Hurt, E. (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. J Biol Chem, 284: 35079-35091.
- Petrov, A.S., Bernier, C.R., Hershkovits, E., Xue, Y., Waterbury, C.C., Hsiao, C., Stepanov, V.G., Gaucher, E.A., Grover, M.A., Harvey, S.C., Hud, N.V., Wartell, R.M., Fox, G.E. and Williams, L.D. (2013) Secondary structure and domain architecture of the 23S and 5S rRNAs. Nucleic Acids Res, 41: 7522-7535.
- Petrov, A.S., Bernier, C.R., Gulen, B., Waterbury, C.C., Hershkovits, E., Hsiao, C., Harvey, S.C., Hud, N.V., Fox, G.E., Wartell, R.M. and Williams, L.D. (2014a) Secondary structures of rRNAs from all three domains of life. PLoS One, 9: e88222.
- Petrov, A.S., Bernier, C.R., Hsiao, C., Norris, A.M., Kovacs, N.A., Waterbury, C.C., Stepanov, V.G., Harvey, S.C., Fox, G.E., Wartell, R.M., Hud, N.V. and Williams, L.D. (2014b) Evolution of the ribosome at atomic resolution. Proc Natl Acad Sci U S A, 111: 10251-10256.
- Pichon, X., Wilson, L.A., Stoneley, M., Bastide, A., King, H.A., Somers, J. and Willis, A.E. (2012) RNA binding protein/RNA element interactions and the control of translation. Curr Protein Pept Sci, 13: 294-304.
- Piekna-Przybylska, D., Decatur, W.A. and Fournier, M.J. (2007) New bioinformatic tools for analysis of nucleotide modifications in eukaryotic rRNA. RNA, 13: 305-312.

- Proud, C.G. (2007) Signalling to translation: how signal transduction pathways control the protein synthetic machinery. Biochem J, 403: 217-234.
- Pyle, A.M. (2008) Translocation and unwinding mechanisms of RNA and DNA helicases. Annu Rev Biophys, 37: 317-336.
- Qu, G., van Nues, R.W., Watkins, N.J. and Maxwell, E.S. (2011) The spatial-functional coupling of box C/D and C'/D' RNPs is an evolutionarily conserved feature of the eukaryotic box C/D snoRNP nucleotide modification complex. Mol Cell Biol, 31: 365-374.
- Rabl, J., Leibundgut, M., Ataide, S.F., Haag, A. and Ban, N. (2011) Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. Science, 331: 730-736.
- Reichow, S.L., Hamma, T., Ferre-D'Amare, A.R. and Varani, G. (2007) The structure and function of small nucleolar ribonucleoproteins. Nucleic Acids Res, 35: 1452-1464.
- Rio, D.C., Ares, M., Hannon, G.J. and Nilsen, T.W. (2011) RNA a laboratory manual. CSH Laboratory Press, Cold Spring Harbor, NY. XIII, 586 S. pp.
- Rodnina, M.V., Daviter, T., Gromadski, K. and Wintermeyer, W. (2002) Structural dynamics of ribosomal RNA during decoding on the ribosome. Biochimie, 84: 745-754.
- Rodnina, M.V., Beringer, M. and Wintermeyer, W. (2007) How ribosomes make peptide bonds. Trends Biochem Sci, 32: 20-26.
- Rodriguez-Galan, O., Garcia-Gomez, J.J. and de la Cruz, J. (2013) Yeast and human RNA helicases involved in ribosome biogenesis: current status and perspectives. Biochim Biophys Acta, 1829: 775-790.
- Rouskin, S., Zubradt, M., Washietl, S., Kellis, M. and Weissman, J.S. (2014) Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. Nature, 505: 701-705.
- Samarsky, D.A. and Fournier, M.J. (1998) Functional mapping of the U3 small nucleolar RNA from the yeast *Saccharomyces cerevisiae*. Mol Cell Biol, 18: 3431-3444.
- Sambrook, J. and Russell, D.W. (2001) Molecular cloning a laboratory manual.3. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schafer, T., Strauss, D., Petfalski, E., Tollervey, D. and Hurt, E. (2003) The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. EMBO J, 22: 1370-1380.
- Schmitt, M.E. and Clayton, D.A. (1993) Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in *Saccharomyces cerevisiae*. Mol Cell Biol, 13: 7935-7941.
- Schneider, C., Leung, E., Brown, J. and Tollervey, D. (2009) The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. Nucleic Acids Res, 37: 1127-1140.
- Seiser, R.M., Sundberg, A.E., Wollam, B.J., Zobel-Thropp, P., Baldwin, K., Spector, M.D. and Lycan, D.E. (2006) Ltv1 is required for efficient nuclear export of the ribosomal small subunit in *Saccharomyces cerevisiae*. Genetics, 174: 679-691.

- Shuai, K. and Warner, J.R. (1991) A temperature sensitive mutant of *Saccharomyces cerevisiae* defective in pre-rRNA processing. Nucleic Acids Res, 19: 5059-5064.
- Sikorski, R.S. and Hieter, P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. Genetics, 122: 19-27.
- Silverman, E., Edwalds-Gilbert, G. and Lin, R.J. (2003) DExD/H-box proteins and their partners: helping RNA helicases unwind. Gene, 312: 1-16.
- Simm, S. (2013) Pipeline for analyzing next generation sequencing data. Goethe-University Frankfurt am Main, Universitätsbibliothek Johann Christian Senckenberg, Zentralbibliothek, Pages.
- Singleton, M.R., Dillingham, M.S. and Wigley, D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. Annu Rev Biochem, 76: 23-50.
- Song, Y., Kim, S. and Kim, J. (1995) ROK1, a high-copy-number plasmid suppressor of *kem1*, encodes a putative ATP-dependent RNA helicase in *Saccharomyces cerevisiae*. Gene, 166: 151-154.
- Story, R.M., Weber, I.T. and Steitz, T.A. (1992) The structure of the *E. coli* recA protein monomer and polymer. Nature, 355: 318-325.
- Strunk, B.S., Loucks, C.R., Su, M., Vashisth, H., Cheng, S., Schilling, J., Brooks, C.L., 3rd, Karbstein, K. and Skiniotis, G. (2011) Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. Science, 333: 1449-1453.
- Swiatkowska, A., Wlotzka, W., Tuck, A., Barrass, J.D., Beggs, J.D. and Tollervey, D. (2012) Kinetic analysis of pre-ribosome structure in vivo. RNA, 18: 2187-2200.
- Szewczak, L.B., DeGregorio, S.J., Strobel, S.A. and Steitz, J.A. (2002) Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP. Chem Biol, 9: 1095-1107.
- Talkish, J., May, G., Lin, Y., Woolford, J.L., Jr. and McManus, C.J. (2014) Mod-seq: high-throughput sequencing for chemical probing of RNA structure. RNA, 20: 713-720.
- Tanaka, N. and Schwer, B. (2006) Mutations in PRP43 that uncouple RNA-dependent NTPase activity and pre-mRNA splicing function. Biochemistry, 45: 6510-6521.
- Tanaka, N., Aronova, A. and Schwer, B. (2007) Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. Genes Dev, 21: 2312-2325.
- Thiry, M. and Lafontaine, D.L. (2005) Birth of a nucleolus: the evolution of nucleolar compartments. Trends Cell Biol, 15: 194-199.
- Thomson, E., Ferreira-Cerca, S. and Hurt, E. (2013) Eukaryotic ribosome biogenesis at a glance. J Cell Sci, 126: 4815-4821.
- Tollervey, D. and Guthrie, C. (1985) Deletion of a yeast small nuclear RNA gene impairs growth. EMBO J, 4: 3873-3878.

- Tollervey, D. (1987) A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA. EMBO J, 6: 4169-4175.
- Torchet, C., Jacq, C. and Hermann-Le Denmat, S. (1998) Two mutant forms of the S1/TPR-containing protein Rrp5p affect the 18S rRNA synthesis in *Saccharomyces cerevisiae*. RNA, 4: 1636-1652.
- Torchet, C. and Hermann-Le Denmat, S. (2000) Bypassing the rRNA processing endonucleolytic cleavage at site A2 in *Saccharomyces cerevisiae*. RNA, 6: 1498-1508.
- Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci U S A, 76: 4350-4354.
- Travis, A.J., Moody, J., Helwak, A., Tollervey, D. and Kudla, G. (2014) Hyb: a bioinformatics pipeline for the analysis of CLASH (crosslinking, ligation and sequencing of hybrids) data. Methods, 65: 263-273.
- Trinh-Rohlik, Q. and Maxwell, E.S. (1988) Homologous genes for mouse 4.5S hybRNA are found in all eukaryotes and their low molecular weight RNA transcripts intermolecularly hybridize with eukaryotic 18S ribosomal RNAs. Nucleic Acids Res, 16: 6041-6056.
- Tsai, R.T., Fu, R.H., Yeh, F.L., Tseng, C.K., Lin, Y.C., Huang, Y.H. and Cheng, S.C. (2005) Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. Genes Dev, 19: 2991-3003.
- Udem, S.A. and Warner, J.R. (1972) Ribosomal RNA synthesis in *Saccharomyces cerevisiae*. J Mol Biol, 65: 227-242.
- Van Der Kelen, K., Beyaert, R., Inze, D. and De Veylder, L. (2009) Translational control of eukaryotic gene expression. Crit Rev Biochem Mol Biol, 44: 143-168.
- van Dijk, E.L., Jaszczyszyn, Y. and Thermes, C. (2014) Library preparation methods for next-generation sequencing: tone down the bias. Exp Cell Res, 322: 12-20.
- van Hoof, A., Lennertz, P. and Parker, R. (2000) Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. Mol Cell Biol, 20: 441-452.
- van Nues, R.W., Granneman, S., Kudla, G., Sloan, K.E., Chicken, M., Tollervey, D. and Watkins, N.J. (2011) Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. EMBO J, 30: 2420-2430.
- Veith, T., Martin, R., Wurm, J.P., Weis, B.L., Duchardt-Ferner, E., Safferthal, C., Hennig, R., Mirus, O., Bohnsack, M.T., Wohnert, J. and Schleiff, E. (2012a) Structural and functional analysis of the archaeal endonuclease Nob1. Nucleic Acids Res, 40: 3259-3274.
- Veith, T., Wurm, J.P., Duchardt-Ferner, E., Weis, B., Martin, R., Safferthal, C., Bohnsack, M.T., Schleiff, E. and Wohnert, J. (2012b) Backbone and side chain NMR resonance assignments for an archaeal homolog of the endonuclease Nob1 involved in ribosome biogenesis. Biomol NMR Assign, 6: 47-50.

- Venema, J. and Tollervey, D. (1996) RRP5 is required for formation of both 18S and 5.8S rRNA in yeast. EMBO J, 15: 5701-5714.
- Venema, J., Bousquet-Antonelli, C., Gelugne, J.P., Caizergues-Ferrer, M. and Tollervey, D. (1997) Rok1p is a putative RNA helicase required for rRNA processing. Mol Cell Biol, 17: 3398-3407.
- Venema, J. and Tollervey, D. (1999) Ribosome synthesis in *Saccharomyces cerevisiae*. Annu Rev Genet, 33: 261-311.
- Vos, H.R., Bax, R., Faber, A.W., Vos, J.C. and Raue, H.A. (2004) U3 snoRNP and Rrp5p associate independently with *Saccharomyces cerevisiae* 35S pre-rRNA, but Rrp5p is essential for association of Rok1p. Nucleic Acids Res, 32: 5827-5833.
- Walbott, H., Mouffok, S., Capeyrou, R., Lebaron, S., Humbert, O., van Tilbeurgh, H., Henry, Y. and Leulliot, N. (2010) Prp43p contains a processive helicase structural architecture with a specific regulatory domain. EMBO J, 29: 2194-2204.
- Warner, J.R. (1999) The economics of ribosome biosynthesis in yeast. Trends Biochem Sci, 24: 437-440.
- Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M. and Luhrmann, R. (1998) Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. RNA, 4: 1549-1568.
- Watkins, N.J. and Bohnsack, M.T. (2012) The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *In* Wiley Interdiscip Rev RNA. Vol. 3. 397-414.
- Wells, S.E., Hughes, J.M., Igel, A.H. and Ares, M., Jr. (2000) Use of dimethyl sulfate to probe RNA structure in vivo. Methods Enzymol, 318: 479-493.
- Westhof, E. and Fritsch, V. (2000) RNA folding: beyond Watson-Crick pairs. Structure, 8: R55-65.
- Wilson, D.N. and Doudna Cate, J.H. (2012) The structure and function of the eukaryotic ribosome. Cold Spring Harb Perspect Biol, 4:
- Woolford, J.L., Jr. and Baserga, S.J. (2013) Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. Genetics, 195: 643-681.
- Wu, H. and Feigon, J. (2007) H/ACA small nucleolar RNA pseudouridylation pockets bind substrate RNA to form three-way junctions that position the target U for modification. Proc Natl Acad Sci U S A, 104: 6655-6660.
- Yang, Q., Del Campo, M., Lambowitz, A.M. and Jankowsky, E. (2007) DEAD-box proteins unwind duplexes by local strand separation. Mol Cell, 28: 253-263.
- Ye, J., Osborne, A.R., Groll, M. and Rapoport, T.A. (2004) RecA-like motor ATPases--lessons from structures. Biochim Biophys Acta, 1659: 1-18.
- Young, C.L., Khoshnevis, S. and Karbstein, K. (2013) Cofactor-dependent specificity of a DEAD-box protein. Proc Natl Acad Sci U S A, 110: E2668-2676.

Zemp, I., Wild, T., O'Donohue, M.F., Wandrey, F., Widmann, B., Gleizes, P.E. and Kutay, U. **(2009)** Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. Cell Biol, 185: 1167-1180.

Zheng, L., Baumann, U. and Reymond, J.L. **(2004)** An efficient one-step site-directed and site-saturation mutagenesis protocol. Nucleic Acids Res, 32: e115.

Supplementary figure

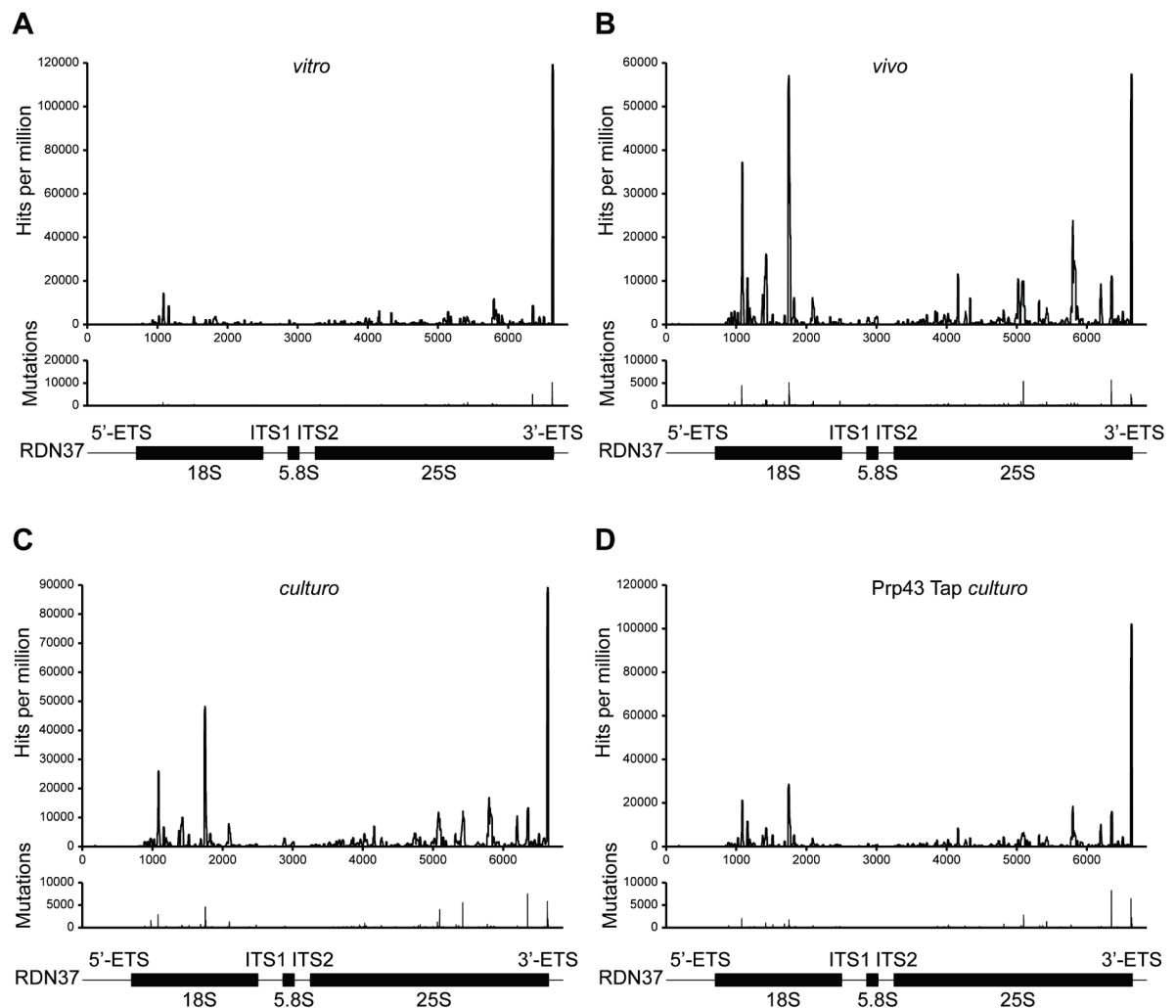


Figure S: CRAC control experiments. Distribution of sequence reads obtained by (A) *in vitro* (B) *in vivo* and (C) *in culturo* crosslinking of untagged control samples on the linear RDN37 sequence, encoding the 35S pre-rRNA. Number of hits were normalised to total mapped reads per million for each nucleotide. Relative positions of the mature 18S, 5.8S and 25S rRNA sequences are indicated at the below the graphs. (D) Mapping of crosslinked reads obtained by CRAC analysis with TAP tagged Prp43 control strains on the primary ribosomal transcript.

Acknowledgements

For the opportunity to undertake this doctoral work I would like to thank the members of my thesis committee Prof. Dr. Jörg Stülke and Prof. Dr. Markus Bohnsack. My thank also go to the external reviewer Prof. Dr. Dagmar Klostermeier and the additional members of the examination board Prof. Dr. Gerhard Braus, Prof. Dr. Ralf Ficner, Prof. Dr. Heinz Neumann, Prof. Dr. Thomas Pieler.

My sincere and special thanks go to Prof. Dr. Markus Bohnsack for the exciting subject of my thesis, for the dedicated supervision and for the chance to develop a diverse range of skills during my time in his group.

I would like to express my gratitude to Prof. Dr. Enrico Schleiff, the head of the Institute of Molecular Bioscience in Frankfurt am Main and Prof. Dr. Blanche Schwappach the head of the Department of Molecular Biology in Göttingen for their support during my work.

Especially, I would like to thank Dr. Katherine Sloan for her valuable comments and discussions while preparing this thesis.

I would like to thank Prof. Dr. Enrico Schleiff and Prof. Dr. Jens Wöhnert for their support and interesting collaborations. In this respect I would also like to thank Stefan Simm for his help with the analysis of bioinformatic data and Maike Ruprecht for performing the UV crosslinking experiments. Thanks also go to Benjamin Weis and Dr. Sandra Missbach for their collaboration in the characterisation of the endonuclease Nob1.

Additionally, I would like to thank the other, present and former members of the AK Bohnsack and AK Schleiff, especially Dr. Ana Ramirez, Annika Straub, Benjamin Weis, Dr. Carmen Doebele, Dr. Matthias Leisegang, Phillip Hackert and Dr. Sandra Missbach, for their support, stimulating discussions and the excellent working atmosphere. Thanks also go to Sunny Sharma for awesome scientific discussions.

I would like to thank my wife Hasret for her caring support and for our harmonious relationship despite the spatial separation.

Finally, I want to express my gratitude to my parents for their patience and ongoing support allowing me to successfully chase my dreams.

Publications associated with this dissertation

- Martin, R.**, Hackert, P., Ruprecht, M., Simm, S., Bruning, L., Mirus, O., Sloan, K.E., Kudla, G., Schleiff, E. and Bohnsack, M.T. (2014) A pre-ribosomal RNA interaction network involving snoRNAs and the Rok1 helicase. RNA, 20: 1173-1182.
- Martin, R.**, Straub, A.U., Doebele, C. and Bohnsack, M.T. (2013) DExD/H-box RNA helicases in ribosome biogenesis. RNA Biol, 10: 4-18.
- Veith, T., **Martin, R.**, Wurm, J.P., Weis, B.L., Duchardt-Ferner, E., Safferthal, C., Hennig, R., Mirus, O., Bohnsack, M.T., Wohnert, J. and Schleiff, E. (2012a) Structural and functional analysis of the archaeal endonuclease Nob1. Nucleic Acids Res, 40: 3259-3274.
- Veith, T., Wurm, J.P., Duchardt-Ferner, E., Weis, B., **Martin, R.**, Safferthal, C., Bohnsack, M.T., Schleiff, E. and Wohnert, J. (2012b) Backbone and side chain NMR resonance assignments for an archaeal homolog of the endonuclease Nob1 involved in ribosome biogenesis. Biomol NMR Assign, 6: 47-50.
- Bohnsack, M.T., **Martin, R.**, Granneman, S., Ruprecht, M., Schleiff, E. and Tollervey, D. (2009) Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. Mol Cell, 36: 583-592.

Curriculum vitae

Roman MARTINStegemuehlenweg 73
37083 GoettingenMartin-Roman@gmx.net
+49(0)176/20691417

Education

- PhD** in the group of Markus T. Bohnsack 07 / 2012 – 11 / 2014
University Medical Center Goettingen, Department of Molecular Biology
- Diploma** in Biology 09 / 2003 – 04 / 2009
Ludwig-Maximilians-University Muenchen
Subjects:
 - Cell biology
 - Genetics
 - Plant physiology
- General qualification for university entrance,** 09 / 1993 – 07 / 2002
Leonhard-Wagner-Gymnasium, Schwabmuenchen
-

Obligatory military service

Luitpold-Kaserne Dillingen an der Donau 07 / 2002 – 04 / 2003

Research Experience

- Research associate** in the group of Markus T. Bohnsack 07 / 2012 – 06 / 2014
University Medical Center Goettingen, Department of Molecular Biology
Research interests:
 - Ribosome biogenesis
 - Structure analysis of RNA *in vivo*
 - Next generation sequencing
- Research associate** in the group of Prof. Dr. Enrico Schleiff 05 / 2009 – 07 / 2012
Goethe-University Frankfurt am Main, Institute of Molecular Bioscience
Research interests:
 - Protein characterisation *in vitro*
 - Quantification of snoRNAs
- Student research assistant** in the group of Prof. Dr. Enrico Schleiff 05 / 2008 – 10 / 2008
Ludwig-Maximilians-University Muenchen
Research interests:
 - Molecular cloning
 - Segregation analysis in plants
- Student research assistant** in the group of Prof. Dr. Enrico Schleiff 12 / 2006 – 11 / 2007
Ludwig-Maximilians-University Muenchen
Research interests:
 - Protein import into chloroplasts
 - Analysis of lipid bilayers *in vitro*

Scholarships

- Research scholarship**, Institute of Molecular Bioscience, Goethe-University Frankfurt am Main 01 / 2011 – 06 / 2011
- Research scholarship**, Frankfurt International Research Graduate School for Translational Biomedicine 05 / 2009 – 12 / 2010

Publications

- **Martin, R.**, Hackert, P., Ruprecht, M., Simm, S., Bruning, L., Mirus, O., Sloan, K.E., Kudla, G., Schleiff, E., and Bohnsack, M.T. (2014) A pre-ribosomal RNA interaction network involving snoRNAs and the Rok1 helicase. [RNA](#)
- **Martin, R.**, Straub, A.U., Doebele, C., and Bohnsack, M.T. (2013) DExD/H-box RNA helicases in ribosome biogenesis. [RNA biology](#)
- Missbach, S., Weis, B.L., **Martin, R.**, Simm, S., Bohnsack, M.T., and Schleiff, E. (2013) 40S ribosome biogenesis co-factors are essential for gametophyte and embryo development. [PloS one](#)
- Veith, T., Wurm, J.P., Duchardt-Ferner, E., Weis, B., **Martin, R.**, Safferthal, C., Bohnsack, M.T., Schleiff, E., and Wohnert, J. (2012) Backbone and side chain NMR resonance assignments for an archaeal homolog of the endonuclease Nob1 involved in ribosome biogenesis. [Biomolecular NMR assignments](#)
- Veith, T., **Martin, R.**, Wurm, J.P., Weis, B.L., Duchardt-Ferner, E., Safferthal, C., Hennig, R., Mirus, O., Bohnsack, M.T., Wohnert, J., *et al.* (2012) Structural and functional analysis of the archaeal endonuclease Nob1. [Nucleic acids research](#)
- Leisegang, M.S., **Martin, R.**, Ramirez, A.S., and Bohnsack, M.T. (2012) Exportin t and Exportin 5: tRNA and miRNA biogenesis - and beyond. [Biological chemistry](#)
- Bohnsack, M.T., **Martin, R.**, Granneman, S., Ruprecht, M., Schleiff, E., and Tollervey, D. (2009) Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. [Molecular cell](#)
- Wunder, T., **Martin, R.**, Loffelhardt, W., Schleiff, E., and Steiner, J.M. (2007) The invariant phenylalanine of precursor proteins discloses the importance of Omp85 for protein translocation into cyanelles. [BMC evolutionary biology](#)