Interactions and functions of RNA-binding proteins

Dissertation

for the award of the degree

"Doctor rerum naturalium"

of the Georg-August-Universität Göttingen

within the doctoral program Molecular biology of cells

of the Georg-August University School of Science (GAUSS)

submitted by

Jens Kretschmer

from Frankfurt am Main

Göttingen 2016

Promovierenden-Erklärung

Georg-August-Universität Göttingen

Name: Kretschmer,Jens

(Name, Vorname)

Anschrift: Südring 19B, 37120, Bovenden

(Straße, PLZ, Wohnort)

Ich beabsichtige, eine Dissertation zum Thema "Interactions and functions of RNA-binding proteins" an der Georg-August-Universität Göttingen anzufertigen. Dabei werde ich von Herrn Prof. Dr. Markus T. Bohnsack betreut.

Ich gebe folgende Erklärung ab:

1. Die Gelegenheit zum vorliegenden Promotionsvorhaben ist mir nicht kommerziell vermittelt worden. Insbesondere habe ich keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.

2. Hilfe Dritter wurde bis jetzt und wird auch künftig nur in wissenschaftlich vertretbarem und prüfungsrechtlich zulässigem Ausmaß in Anspruch genommen. Insbesondere werden alle Teile der Dissertation selbst angefertigt; unzulässige fremde Hilfe habe ich dazu weder unentgeltlich noch entgeltlich entgegengenommen und werde dies auch zukünftig so halten.

3. Die Richtlinien zur Sicherung der guten wissenschaftlichen Praxis an der Universität Göttingen werden von mir beachtet.

4. Eine entsprechende Promotion wurde an keiner anderen Hochschule im In- oder Ausland beantragt; die eingereichte Dissertation oder Teile von ihr wurden nicht für ein anderes Promotionsvorhaben verwendet.

Mir ist bekannt, dass unrichtige Angaben die Zulassung zur Promotion ausschließen bzw. später zum Verfahrensabbruch oder zur Rücknahme des erlangten Grades führen.

Göttingen, den 29.11.2016

(Unterschrift)

Table of Contents

List of Figures

List of Tabels

Abbreviation	Meaning
(v/v)	Volume per volume
(w/v)	Weight per volume
bp	Base pair
cDNA	Complementary DNA
CLIP	Cross-linking and immunoprecipitation
CRAC	Cross-linking and analysis of cDNA
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
ETS	External transcribed spacer
FLAG	His_6-PreScission protease site-$Flag_2$
HZZT	His_{10}-ZZ-TEV protease site
ITS	Internal transcribed spacer
lncRNA	Long non-coding RNA
LSU	Large ribosomal subunit
miRNA	Micro RNA
miscRNA	Miscellaneous RNA
mRNA	Messenger RNA
NPC	Nuclear pore complex
nt	Nucleotide
NTP	Nucleoside triphosphate
PAGE	Polyacrylamide gel electrophoresis
PAR-CRAC	Photoactivatable-Ribonucleoside-Enhanced -CRAC
PCR	Polymerase chain reaction
pre-mRNA	Precursor mRNA
pre-rRNA	Precursor rRNA
RNA	Ribonucleic acid
RNAP I/II/III	RNA polymerase I/II/III
RNP	Ribonucleoprotein particle
rRNA	Ribosomal RNA
RT-qPCR	Real-time quantitative PCR
S	Svedberg Units
snoRNA	Small nucleolar RNA
snoRNP	Small nucleolar ribonucleoprotein particle
snRNA	Small nuclear RNA
SSU	Small ribosomal subunit
tRNA	Transfer RNA
U	Units
UTR	Untranslated region
UV	Ultraviolet
Ψ	Pseudouridine

Abstract

Alongside several well-known modifications in DNA and proteins, more than 100 different types of chemical modification are also found in cellular RNAs. RNA modifications can influence the secondary structure and interactions of the RNAs that carry them and they can therefore play important roles in regulating the functions of the RNAs. For many RNA modifications, the enzymes that introduce them are known but the modification targets of several predicted modification enzymes remain to be identified. Interestingly, a particular modification, N^6-methyladenosine (m^6A), was recently found to be reversible and a group of proteins, termed "readers" that can recognise the modification often via a specialised RNA binding domain (YTH domain), have been identified. Such "reader" proteins have been shown to regulate the fate of RNAs according to their modification status, suggesting that this so called "epitranscriptome" is an additional layer of regulation of gene expression.

In this study, cross-linking and analysis of cDNA (CRAC) was used to identify the RNA-interactome of the five human YTH domain-containing proteins, YTHDF1, YTHDF2, YTHDF3, YTHDC1, YTHDC2. To facilitate the mapping of the deep sequencing data obtained from CRAC experiments performed in human cells, a bioinformatic pipeline was adapted and further developed. Analysis of the CRAC data showed that YTHDF1, YTHDF2, YTHDF3 and YTHDC1 predominantly cross-link to mRNAs, which is in line with recently published reports describing functions for these proteins in mRNA degradation, alternative pre-mRNA splicing and enhancing mRNA translation. Interestingly, the CRAC analysis of YTHDC2 revealed a specific cross-linking site on the 18S ribosomal RNA and the association of this protein with ribosomal complexes was confirmed by independent experimental approaches. CRAC analysis using truncated versions of YTHDC2 suggested that the R3H RNA binding domain is required for stable association of this protein with the ribosome and *in vitro* anisotropy experiments demonstrated that the YTH domain of YTHDC2 has a higher affinity for m^6A modifications present in the sequence context found in the ribosomal RNAs than the classical m^6A consensus motif found in many mRNAs. Interestingly, immunoprecipitation experiments followed by mass spectrometry identified the cytoplasmic 5'-3' exonuclease XRN1 as an interaction partner of YTHDC2. These data could suggest a model in which recognition of the m^6A modification(s) on the ribosomal RNA by the YTH domain of YTHDC2 promotes RNA degradation by XRN1. Taken together, this study contributes to the understanding of the diverse functions that modification "reader" proteins can play in regulating RNA metabolism.

1 Introduction

1.1 RNA modifications

Since the 1970s it has been known that RNA can be modified and the knowledge of RNA modifications has continually increased since (Rottman et al., 1974; Schibler et al., 1977; Wei et al., 1975, 1976; Wei and Moss, 1977). So far, about 150 different RNA modifications are known in all kingdoms of life, spanning all major classes of RNA in the cell (Machnicka et al., 2013). The most modified RNAs in the cell are transfer (t)RNAs, followed by the extensively modified ribosomal (r)RNAs. RNA modifications are known to affect the structural stability and folding of RNA, leading to the degradation of misfolded or aberrant RNA if important modifications are missing. In addition, translation can be fine tuned by modifications on tRNAs that influence the decoding of cognate messenger (m)RNA codons, and rRNA modifications that can influence the translation efficiency and fidelity of the ribosome. Modifications in mRNA can influence the stability of mRNAs and thereby also affect translation efficiency. Among the myriad of RNA modifications, the most common modifications are the addition of a methyl-group to various positions of the base and on the ribose, and pseudouridination of uridines. Less common modifications are, for example, acetylations, geranylations, wybutosine or carbamylation, which can be found in tRNAs (and rRNAs in the case of acetylation). RNA modifications have essential, but highly variable functions in the cell, forming an additional layer of regulation in gene expression termed the epitranscriptome.

1.2 tRNA modifications

1.2.1 tRNA biogenesis and function

During translation of the mRNA by the ribosome, tRNAs serve as adapter molecules that translate the mRNA sequence into the amino acid sequence of the protein. Extensive and highly regulated RNA-RNA interactions between the tRNA, the codon triplet of the mRNA and the ribosome ensure the correct selection of the tRNA and incorporation of the correct amino acid into the nascent polypeptide chain by the ribosome (Demeshkina et al., 2010).

However, the delivery of aminoacyl-residues to the ribosome during translation is not the only function of amino-acyl tRNAs and additional roles in the cell have been described. For example, tRNAs are required for the addition of destabilizing amino acids to the N-terminus of proteins to facilitate their turnover in the N-end rule pathway. Targeted

endonucleolytic cleavage at the codon loop divides the tRNA in half, forming stable fragments, which are proposed to have regulatory and signalling functions or are involved in the response to HIV-infection (Banerjee et al., 2010; Nawrot et al., 2011; Yeung et al., 2009).

In *Saccharomyces cerevisiae*, a total of 275 tRNA genes are transcribed by RNA polymerase (RNAP) III (Chan and Lowe, 2009). The nascent transcript includes a 5' leader sequence, a 3' trailer sequence and can contain an intronic sequence, which needs to be removed during biogenesis. Removal of the leader sequence by RNase P precedes the removal of the trailer sequence by RNase Z after nucleotide 73. The 3' end is further processed by the addition of a CCA sequence by the nucleotidyltransferase Cca1 (TRNT1 in human cells) (Aebi et al., 1990). tRNAs are exported by Los1 (XPO-T in human cells), which also serves as a quality control step by sensing correct tertiary structure and end processing of the tRNA (Arts et al., 1998; Lipowsky et al., 1999; Lund and Dahlberg, 1998; Sarkar and Hopper, 1998). 61 tRNA genes contain introns, which are removed (often referred to as tRNA splicing) by the conserved family of Sen proteins either after export at the surface of mitochondria in yeast, or prior to export in the nucleoplasm in vertebrates (Lund and Dahlberg, 1998; Melton et al., 1980; Yoshihisa et al., 2003). tRNAs can undergo retrograde import into the nucleus, either for temporary storage or for further maturation. Finally, aminoacylation of tRNAs by aminoacylsynthetases takes place either in the nucleus after retrograde import or directly in the cytoplasm after maturation (Grosshans et al., 2000; Lund and Dahlberg, 1998; Steiner-Mosonyi and Mangroo, 2004). An overview of the pathway is shown in Figure 1A.

The extensive modification of tRNAs is achieved by the addition of modifications throughout the whole maturation pathway (Figure 1B). The first modifications are directly added to the nascent transcript and further modifications are added in the nucleus and the cytoplasm (Hopper, 2013). The timing of some tRNA modifications is determined by the localisation of the enzymes that introduce them while other tRNA modification enzymes recognise specific features of the tRNA, e.g. introns or the 3'-CAA and can only modify the tRNA at a particular stage during its maturation (Grosjean et al., 1997). Modifications are necessary for the correct folding of tRNAs into the canonical clover leaf secondary structure with the acceptor stem and the D-loop, the TΨC-loop, the anticodon-loop and the variable loop resembling the four parts of the leaf. The tertiary structure of tRNAs is an inverted L-shaped structure with the anticodon-loop at the bottom and the CCA-acceptor stem at the top (Shi and Moore, 2000). In addition, tRNA modifications are involved in codon recognition as well as tRNA stability, as described in the next chapter.
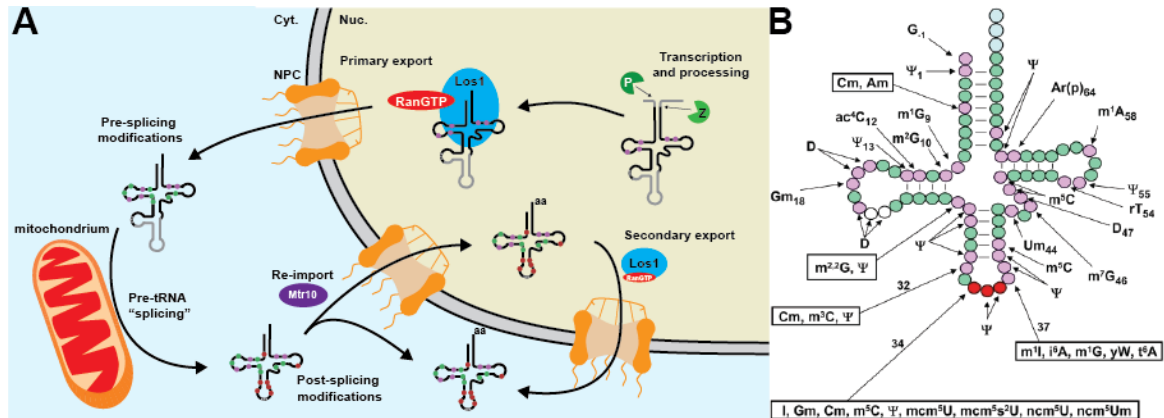
Figure 1: tRNA biogenesis and modifications. A Schematic overview of transfer (t)RNA biogenesis in yeast. tRNA biogenesis starts with the transcription of pre-tRNAs in the nucleus by RNAP III, followed by end processing and modification of the RNA. After export, introns are removed on the surface of mitochondria and further modifications are installed. The tRNAs either undergo aminoacylation and are primed for translation or they undergo retrograde import. Modifications are indicated as coloured circles: Early modifications are shown in pink, "pre-splicing" modifications are displayed in green and "post-splcing" modifications are shown in red. Abbreviations: Cyt, cytoplasm; Nuc. nucleus; NPC, nuclear-pore-complex. Modified from Sloan et al. (2016). B Overview of chemical modifications found in cytoplasmic tRNAs in yeast. A tRNA structure is shown in the cloverleaf representation. Residues that are unmodified in all tRNAs are shown in green, residues that are modified in some or all tRNAs are shown in pink, and white residues represent additional residues that are present in some tRNA species, which can also carry modifications. The anticodon loop is coloured in red and is also sometimes modified. The CCA end is shown in light blue. From Phizicky and Hopper (2010).

1.2.2 Different types of tRNA modifications and their function

Numerous different modifications are found in tRNAs. Together, over 100 chemically unique modifications are found in tRNAs in all three domains of life, of which 18 are universally present (Jackman and Alfonzo, 2013). However, many of these chemical modifications are also found in other types of RNA in the cell, although the enzymes that install them are often different (Phizicky and Hopper, 2010). tRNAs are also the most extensively modified RNA in the cell. Approximately, 17 % of the residues are modified, which is ten times more than in rRNA (Jackman and Alfonzo, 2013).

In general, tRNA modifications can be sorted into two categories based on their position within the tRNA. Modifications in the anti-codon loop often affect codon recognition and therefore synthesis of proteins, whereas modifications in the main body are frequently connected to tRNA stability.

Modifications in the anti-codon loop are often found at position 34, which is called the "wobble position". The genetic code is degenerate meaning that multiple codons code for the same amino acid, because the number of codons exceeds the number of amino

acids. This results in the fact that for the decoding of many codons, the nucleotide at the third position is flexible and the corresponding tRNA is required to recognise multiple different nucleotides at this position. This flexibility can be achieved by modifications at the wobble position of the tRNA. A well-studied example of a wobble position modification that affects translation is the deamination of adenine to inosine by the RNA-dependent adenosine deaminases Tad2 and Tad3 (Gerber and Keller, 1999). The conversion to inosine leads to an increased base-pairing capability with cytidine and adenine in addition to the conventional base-pairing with uracil. The lack of these modification leads to decoding errors during translation. Many other modifications, such as 5-methylcytosine (m^5C) and 5-methoxycarbonylmethyl-2-thiouridine, are also found at position 34 of tRNA and similarly function to influence the decoding capacity of the tRNAs that carry them (reviwed in Ranjan and Rodnina, 2016).

Modifications in the body of the tRNA are commonly connected to structural stability by defining either more loose or rigid parts of the tRNA. Several studies showed that loss of certain modifications can lead to increased instability of mature tRNAs and the generation of tRNA fragments, for example during heat stress (Alexandrov et al., 2006; Dewe et al., 2012; Kotelawala et al., 2008). However, also the initial folding of tRNAs can be affected by modifications and single modifications can promote the correct folding of tRNAs (reviewed in Motorin and Helm, 2010).

Loss of tRNA modifications or mutations in tRNA modifying enzymes have been connected to a variety of human diseases, including neurological and metabolic diseases and cancer (reviewed in Torres et al., 2014). It has been suggested that the molecular connection between these diseases and tRNA modifications can be based on perturbed protein synthesis due to the inefficient reading of certain codons, however, the molecular basis of many diseases is not known. However, it could be shown that the lack of $ms^2t^6A_{37}$ in the $tRNA^{Lys}$ leads to the production of aberrant proinsulin, thus impeding the processing to insulin and causing type II diabetes.

1.3 Ribosomal RNA modifications

1.3.1 Biogenesis and function of ribosomes

The ribosome is the protein synthesis machine of the cell and contains four ribosomal (r)RNAs and approximately 80 ribosomal proteins (Anger et al., 2013; Ben-Shem et al., 2011). These are asymmetrically organised in a large and a small ribosomal subunit (SSU and LSU). The 18S rRNA is part of the SSU, whereas the 25S (yeast)/28S (humans), 5.8S and 5S rRNAs form the core of the LSU. Interestingly, the proteins solely function as a scaffold for stabilising the rRNA and do not have an enzymatic activity (Simonovic and Steitz, 2009). Thus, the ribosome is a large ribozyme because the rRNA is responsible for providing the catalytic enzymatic activity. The ribosome has two major functions, firstly providing a framework for the translation and secondly, enabling formation of the peptide bond. The mRNA decoding centre is located in the SSU providing a scaffold for high fidelity decoding (Demeshkina et al., 2012). The LSU on the other hand contains the peptidyl transferase centre (PTC), which is responsible for the introduction of peptide bonds connecting single amino acids to form nascent peptides (Ben-Shem et al., 2011; Simonovic and Steitz, 2009). Notably, these two functionally important sites are conserved in all kingdoms of life, while other regions of the ribosome are more variable (Armache et al., 2013; Ban et al., 2014; Melnikov et al., 2012; Wilson and Doudna Cate, 2012). These include so called eukaryotic extensions, which are stretches of rRNA specifically found in eukaryotic ribosomes, which are thought to enable the translation of more complex mRNAs in higher eukaryotes and might also have regulatory functions.

Ribosome biogenesis is one of the major energy consuming pathways in the cell (reviewed in Henras et al., 2015; Woolford and Baserga, 2013). In a rapidly dividing yeast cell 2,000 ribosomes are produced per minute leading to the synthesis of 200,000 ribosomes per generation (Warner, 1999). The rRNA represents approximately 80 % of the total cellular RNA and 60 % of mRNA transcripts are related to ribosome biogenesis. All three RNA-Polymerases (RNAP) are involved in this process. It starts with the transcription of a precursor (pre-)rRNA (47S in humans and 35S in yeast) by RNAPI from the rDNA repeats. In human cells, the repeats are localised at the short arms of the five acrocentric chromosomes HSA-13, 14, 15, 21 and 22 (Worton et al., 1988). The 47S pre-rRNA contains the sequences of the 18S, 5.8S and 28S rRNAs, separated by internal transcribed spacers (ITS1 and ITS2) and external transcribed spacers (5' ETS and 3' ETS). Ribosome biogenesis factors bind co-transcriptionally to the 47S rRNA
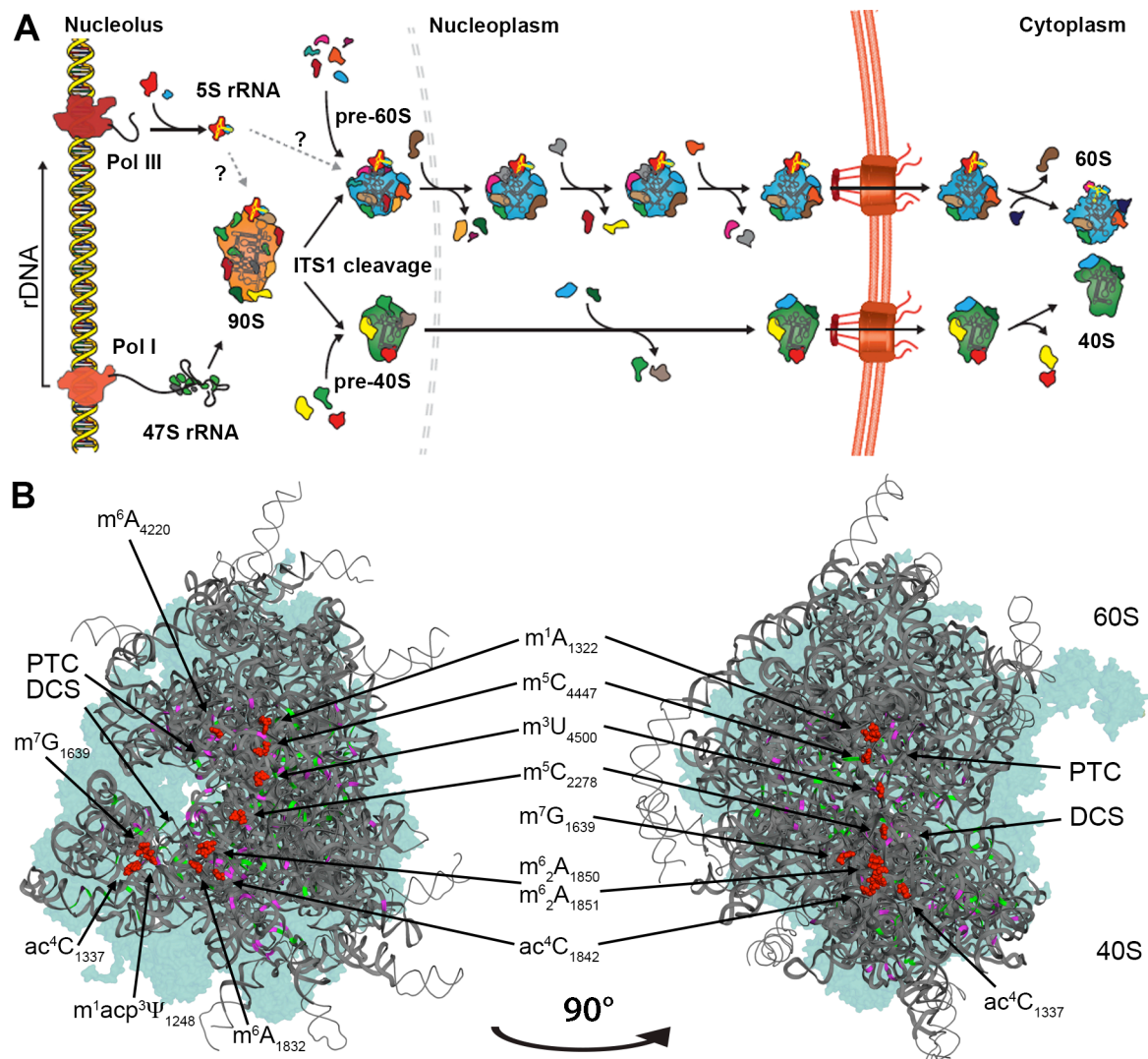
Figure 2: Ribosome biogenesis and rRNA modifications. A Schematic overview of ribosome biogenesis in human cells. Maturation of the ribosomal subunits is shown from left to right, starting with the transcription of the 47S pre-ribosomal (r)RNA from the rDNA repeat by RNAP I and assembly of the 90S pre-ribosomal complexes. The 5S rRNA is transcribed independently by RNAP III and joins the pre-60S (large ribosomal subunit) complex (blue) in the nucleolus as part of the 5S RNP. After a central cleavage step, pre-60S subunit and pre-40S (small ribosomal) subunit (green) maturation continues separately. During this process, numerous ribosome biogenesis factors, indicated by coloured shapes, transiently interact with the pre-ribosomal subunits. The pre-ribosomal subunits are exported and final maturation steps occur in the cytoplasm. Cellular compartments are indicated at the top. Abbreviations: RNAP I, RNA polymerase I; RNAP III, RNA polymerase III; RNP, ribonucleoprotein particle. Adapted from Gerhardy et al. (2014). **B** 3D structure of the human ribosome (PDB 4V6X, Anger et al., 2013). The ribosomal RNA is depicted in grey cartoon model representation and ribosomal proteins are shown as light blue in the background. The positions of the base modifications are shown by red spheres, and the type of modification and the modified residue are indicated. Pseudouridinations and 2'-O-methylations are marked on the rRNA in magenta and green, respectively. The positions of functionally important regions of the ribosome such as the peptidyl transferase centre (PTC) and decoding site (DCS) are indicated.

forming the 90S pre-ribosome in the nucleolus (reviewed in Kornprobst et al., 2016; Tschochner and Hurt, 2003). The 5S rRNA is transcribed separately by RNAPIII and is incorporated into the pre-ribosome in complex with the ribosomal proteins RPL5 (uL18) and RPL11 (uL5) (reviewed in Ciganda and Williams, 2011). After a central pre-rRNA cleavage step, the 90S pre-ribosome is separated into the 60S pre-ribosomal complex (pre-LSU) and the 40S pre-ribosomal complex (pre-SSU) (Figure 2A, Henras et al., 2015; Sloan et al., 2013; Thomson et al., 2013; Woolford and Baserga, 2013). These mature independently while transported through the nucleoplasm constantly swapping ribosome biogenesis factors (reviewed in Gerhardy et al., 2014). Final maturation steps occur in the cytoplasm after the pre-SSU and pre-LSU particles are separately exported to the cytoplasm (Lebaron et al., 2012; Sloan et al., 2016). More than 200 co-factors are involved in processing and modifying of the rRNA and assembly of the ribosomal subunits (Gasse et al., 2015; Sharma and Lafontaine, 2015; Sloan et al., 2015; Sloan et al., 2013; Watkins and Bohnsack, 2012; Woolford and Baserga, 2013). Endo- and exo-nucleases are involved in the removal of the spacer fragments, whereas methyltransferases and pseudouridine synthases introduce the majority of the rRNA modifications. RNA helicases are involved in remodelling of RNA-RNA and RNA-protein interactions in concert with ATPases and GTPases, which introduce conformational changes to the pre-ribosome.

1.3.2 snoRNA-guided modifications

The rRNA modifications can be classified into snoRNA-guided modifications and base-modifications based on their location on the nucleotide. Backbone modifications are the most abundant modifications on the rRNA and can be further sub-divided into 2'-O-methylations and pseudouridinations. 55 2'-O-methylations and 45 pseudouridines are found in yeast rRNA, whereas approximately 100 2'-O-methylations and 100 pseudouridines are found in human rRNA (Birkedal et al., 2015; Krogh et al., 2016; Lestrade and Weber, 2006; Piekna-Przybylska et al., 2008a; Taoka et al., 2016). These modifications are mostly introduced by small nucleolar ribonucleoprotein complexes (snoRNPs), which consist of a small nucleolar (sno)RNA and four core proteins (reviewed in Watkins and Bohnsack, 2012). The snoRNA guides the complex to the site of modification by forming base-paring interactions with the pre-rRNA. The proteins of the complex also establish further pre-rRNA interactions and provide the enzymatic activity of the complex. 2'-O-methylations or pseudouridines are introduced either by Box C/D snoRNPs, which contain the methyltransferase Fibrillarin (yeast Nop1) or H/ACA snoRNPs that contain the pseudouridine synthase Dyskerin (yeast Cbf5). Interestingly, in bacteria 2'-O-methylation and pseudouridine modifications are not introduced by

snoRNPs but by standalone methyltransferases or pseudouridine synthases, meaning a separate enzyme is required for each of the 14 backbone modifications (Popova and Williamson, 2014). The switch to a modular system with constant protein components and a flexible guide-snoRNA makes the system much more efficient for 100-200 different modifications in eukaryotes.

1.3.3 Base modifications

Interestingly, from bacteria to lower eukaryotes and further on to humans the ratio of modifications shifts from primarily base modifications to mainly backbone modifications. In bacterial rRNA, the majority are base modifications, whereas in yeast rRNA the amount drops to approximately 10 % (12 of 112 in total), which further decreases to 5 % (11 of >200 in total) in human rRNA. This reduction is mainly due to an increase in 2'-O-methylations and pseudouridinations suggesting that most base modifications are important and therefore conserved (Piekna-Przybylska et al., 2008a; Popova and Williamson, 2014; Sharma and Lafontaine, 2015).

Seven different types of base modifications are found in yeast rRNA. Half of the 12 base modifications are found in the SSU and half in the LSU in yeast. All six SSU modifications are conserved in humans, while only three of the six LSU base modifications are conserved (Figure 2B, Sharma and Lafontaine, 2015). Notably, except for one, the individual modifications are not essential for cell growth in yeast, however, several of the modifying enzymes are essential or lead to impaired growth, meaning that only the presence of the enzymes is important for ribosome biogenesis (Sharma and Lafontaine, 2015). In yeast, the enzymes responsible for the modifications are all known: nine stand-alone methyltransferases and one aminocarboxypropyl (acp) transferase.

The 18S rRNA of the SSU contains a hypermodified uridine, $m^1acp^3\Psi$, at position 1191. The first step in this modification pathway is a pseudouridylation guided by the snoRNA snR35, which was shown to be not essential but deletion strains show a growth defect (Liang et al., 2009). In the second step, the SPOUT class methyltransferase Emg1/Nep1 introduces the N1-methylation to the base of the pseudouridine (Leulliot et al., 2008; Wurm et al., 2010). SPOUT class methyltransferases use S-adenosyl-methionine (SAM) as the source of the methyl group that is transferred (Tkaczuk et al., 2007). The last step takes place in the cytoplasm and is catalysed by the acp-transferase Tsr3, which interestingly also resembles a SPOUT class methyltransferases. However, instead of the methyl group of SAM, the acp-group is transferred to the substrate (Meyer et al., 2016). This $m^1acp^3\Psi$ hypermodification is conserved in human rRNA and the human orthologue of Emg1 was shown to complement the function in yeast and knockdown of human Tsr3

was shown to reduce the modification *in vivo* (Eschrich et al., 2002; Liu and Thiele, 2001; Meyer et al., 2016). The eukaryotic specific acetylations of 18S-ac^{4}C$_{1280}$ and 18S-ac^{4}C$_{1773}$ are both introduced by the ATP and acetyl-CoA dependent acetylases Kre33 in yeast and NAT10 (ac^{4}C$_{1337}$, ac^{4}C$_{1842}$) in human cells (Ito et al., 2014a; Ito et al., 2014b; Sharma et al., 2015). The yeast modification 18S-m^{7}G$_{1575}$ is installed by Bud23 together with Trm112, and is conserved in humans, where WBSCR22 together with TRMT112 are responsible for this modification (Haag et al., 2015a; White et al., 2008). Despite the fact that the methyltransferase activity of Bud23/WBSCR22 is not essential, the proteins are required for ribosome biogenesis and for the efficient export of the 40S subunits in yeast and humans (Haag et al., 2015a; White et al., 2008; Zorbas et al., 2015). The only di-methylations in rRNA are the two 18S-m$_2^6$A$_{1781}$, 18S-m$_2^6$A$_{1782}$ modifications, which are conserved from bacteria to eukaryotes. In yeast, the modifying di-methyltransferase is Dim1 that joins the pre-ribosome in the nucleus, but installs the modification in the cytoplasm, whereas in human cells DIMT1L stays in the nucleus where also the modification takes place (Lafontaine et al., 1995; Zorbas et al., 2015).

In the LSU, six mono-methylations are reported in yeast. All modifications are introduced by Rossmann-fold methyltransferases that use SAM as the methyl group donor (Sharma and Lafontaine, 2015). The m^{1}A$_{645}$ is conserved in higher eukaryotes and mediated by the ribosome biogenesis factor Rrp8 in yeast (Peifer et al., 2013). The second m^{1}A$_{2142}$ modification is introduced by Bmt2, as could be shown by mutation analysis (Sharma et al., 2013a). Unlike m^{1}A$_{645}$, m^{1}A$_{2142}$ is not conserved in human cells. Two m^{5}C modifications can be found in yeast at position 2278 and 2870 of the 25S rRNA. They are installed by Rcm1 and Nop2, respectively (Sharma et al., 2013b). Both modifications are not essential, however, loss of m^{1}A$_{2142}$ leads to slow growth and Nop2, in contrast to Rcm1, is essential. The modifications are conserved in human and it was shown that human NSUN1 (p120) could complement a *nop2Δ* yeast strain and restore the m^{5}C$_{2870}$ modification, suggesting that it is the methyltransferase for m^{5}C$_{4447}$ in human rRNA (Bourgeois et al., 2015). It is suggested that the human homologue of Rcm1, NSUN5, is responsible for the corresponding human m^{5}C$_{3782}$ modification, but although this has not been directly proven, evidence from fruit flies and worms, as well as homology studies strongly supports this (Schosserer et al., 2015; Sharma et al., 2013b). Finally, the methyltransferases responsible for m^{3}U$_{2634}$ and m^{3}U$_{2843}$ were identified as Bmt5 and Bmt6 in yeast (Sharma et al., 2014). Also, one m^{3}U$_{4500}$ in 28S rRNA of human cells is reported, however, the methyltransferase responsible for this modification remains elusive (Piekna-Przybylska et al., 2008a). Compared to yeast, human rRNA has an additional type of modification; one modified N^6-methyladenosine (m^{6}A) residue at

position 1832 in the 18S rRNA and one at position 4220 in the 28S rRNA (Linder et al., 2015; Piekna-Przybylska et al., 2008a). The modifying methyltransferases still need to be discovered.

1.3.4 Functions of rRNA modifications

In general, RNA modifications expand the chemical properties of nucleotides and thereby influence the functions of the RNAs that carry them. 2'-O-methylations of the ribose increase hydrophobicity and lead to enhanced hydrophobic interactions. This causes increased rigidity of the RNA by additional base stacking capabilities (Prusiner et al., 1974). Pseudouridine is an isomerisation of the uracil ring resulting in additional hydrogen bond formation capabilities compared to uridine. This increases the thermal stability of the RNA by forming additional RNA-RNA interaction (Hayrapetyan, 2009). On the other hand, base modifications can have several effects, depending on their location on the base. They can increase base stacking as well as introduce a charge to the aromatic ring system, abrogate Watson-Crick base paring or introduce non canonical hydrogen bonding (Hayrapetyan, 2009). Introduction of a charge may also extend RNA-protein interaction possibilities (Agris et al., 1986).

In the ribosome, rRNA modifications cluster at functionally important regions, such as the PTC in the LSU, the decoding centre in the SSU and at the inter-subunit contact sites. These modifications are suggested to regulate the stability of the RNA and thereby enhance efficient and accurate translation by the ribosome. To achieve a high fidelity of the ribosome the single modifications work in concert, meaning that deletion of single modifications often does not affect translation, however, if several modifications are deleted, effects in translational fidelity can be detected (Gigova et al., 2014; King et al., 2003). For example frame shifting and stop codon read-through is increased and tRNA incorporation is decreased when snoRNAs guiding clusters of modifications are deleted (Baudin-Baillieu et al., 2009; King et al., 2003; Liang et al., 2007, 2009). For example, an important modification cluster is located on a structure in the LSU called the A-site finger (helix 38) (Piekna-Przybylska et al., 2008b). This helix makes important contacts with the 5S rRNA, tRNAs and also the SSU and is thought to function as an attenuator while moving the tRNA from the A- to the P-site during translation. Six pseudouridines are clustered there and three of them are conserved in eukaryotes. Depletion of two of the three modifications showed no significant effect, however, if all three were depleted, the cells displayed slower growth rates, lower LSU abundance and decreased translation.

As mentioned above, lack of individual base modifications often does not have a significant effect on ribosome biogenesis, but lack of these modifications can have an

effect on translation fidelity. For example, expression of a catalytically inactive version of the RNA methyltransferase Dim1 leads to translation defects *in vitro* (Lafontaine 1998). Another example of a base modification that affects translation is m^5C_{2278}, introduced by Rcm1 (Schosserer et al., 2015). Loss of this modification leads to reduced translational fidelity and increased STOP codon read-through due to structural changes in the vicinity of the modification.

Interestingly, modification sites on the rRNA are not all fully modified under normal growth conditions. Recent 2'-O-methylation profiling of the rRNA revealed that one third of the positions are only partially modified in human cells and studies in yeast similarly identified sites of partial modification (Birkedal et al., 2015; Buchhaupt et al., 2014; Krogh et al., 2016; Taoka et al., 2016). The extent of modification at specific positions may vary in different cells, supporting the concept of ribosome heterogeneity. Specialised ribosomes could translate a subset of mRNAs or are concentrated at different locations in the cytoplasm. Partial modifications might also have a regulatory function under different stress conditions and could also play a role in pathological settings, as several rRNA modifications and modifying enzymes are linked to human diseases.

A variety of human disorders have been linked to defects in rRNA modifications or enzymes that install them. Altered snoRNA levels were found in haematological disorders like leukaemia and dyskeratosis congenital as well as in lung and prostate cancer (McMahon et al., 2015). For example, the Bowen-Conradi syndrome is caused by a mutation in the methyltransferase gene *EMG1* and the genes encoding for the m^7G and m^5C methyltransferases WBSCR22 and NSUN5 are deleted in Williams-Beuren syndrome (Armistead et al., 2009; Doll and Grzeschik, 2001).

1.4 Messenger RNA modifications

The first publications of modifications in messenger (m)RNAs were published in the 1970s with the identification of m^6A and m^5C (Dubin and Taylor, 1975; Schibler et al., 1977). Due to methodical and technical limitations, the extent of modifications could not be detected at that time. With the emergence of new sequencing techniques further modifications could be identified, leading to the term "epitranscriptome" and based on the term epigenome (reviewed in Hoernes and Erlacher, 2016; Soshnev et al., 2016). So far, four modifications have been found in mRNAs; m^6A, m^5C, pseudouridine and N^1-methyladenine (m^1A). Since the m^6A modification is discussed in detail in section 1.5, this section will focus on the other modifications found in mRNAs. In addition to these methylations and pseudouridylations, mRNA can undergo RNA editing, involving insertion or deletion of nucleotides, or alteration of cytosine to uridine or adenine to

inosine. Such RNA editing can change protein sequences by altering nucleotides of codons and potentially introducing additional STOP codons as well as influencing the differential expression of miRNAs (Chawla and Sokol, 2014; Powell et al., 1987).

Beside its occurrence in tRNAs, rRNA and the well characterized function in transcription regulation on DNA, m^5C is also present in mRNA in humans and *archaea* (Edelheit et al., 2013; Squires et al., 2012). Next-generation sequencing together with bisulfide sequencing has allowed the transcriptome-wide mapping of m^5C and its oxidation products 5-hydroxymethylcytidine (hm^5C) and 5-formylcytidine (f^5C) (Booth et al., 2013; Edelheit et al., 2013; Lee and Kim, 2016). In HeLa cells over 10,000 m^5C modification sites were discovered in non-coding (nc)RNAs and mRNAs. Analysis of the distribution on mRNAs showed signals along the mRNA with an enrichment in 5' and 3' UTRs, suggesting a function in regulating protein translation (Squires et al., 2012). This is supported by experiments in fruit flies, which showed impaired brain development upon reduction of the hm^5C modification and a positive correlation of m^5C to hm^5C conversion and translation was exhibited *in vitro* (Delatte et al., 2016). In human cells, the m^5C content can also be regulated by oxidation to hm^5C and f^5C, implying that a similar mechanism may exist in mammals (Huber et al., 2015).

Very recently m^1A was discovered in mRNAs by making use of specific chemical properties of this modified nucleotide. Two parallel studies demonstrated the presence of m^1A in mRNAs by using an antibody-based approach to pull down m^1A containing sequences, followed by next-generation sequencing to map the modified nucleotides (Dominissini et al., 2016; Tserovski et al., 2016). During reverse transcription, m^1A modifications introduce transcription stops represented by accumulation of 3' ends that correlate with the modification sites. The presence of m^1A in mRNAs and ncRNAs is conserved in eukaryotes from yeast to human and it is suggested that the modifications are often embedded in a GC rich sequence (Dominissini et al., 2016). The number of m^1A-containing mRNAs is reported to range from over 800 (Li et al., 2016) to over 4,000 (Dominissini et al., 2016) and the average methylation level of a single m^1A-containing mRNA is approximately 20 % (Dominissini et al., 2016). The distribution of the m^1A modification on mRNAs is still open to debate as one report suggests a bias towards the 5' end of mRNAs with an elevated abundance in 5' UTRs, especially if they contain strong secondary structures, and an increase of m^1A near the first splice site and the start codon (Dominissini et al., 2016), whereas the other report suggests an over representation of m^1A modification in both the 5' and 3' UTRs (Li et al., 2016). However, both reports suggest a role for the modification in translation regulation as the methylation pattern was found to be altered in response to different physiological

conditions and external stresses (Dominissini et al., 2016; Li et al., 2016). In line with such a dynamic function, the modification was also found to be reversible by the alpha-ketoglutarate-dependent dioxygenase ALKBH3 (Li et al., 2016).

Next-generation sequencing has also enabled the transcriptome-wide mapping of pseudouridine. Similar to m^{1}A, pseudouridine introduces a reverse transcription stop after treatment with the molecule CMC (Bakin and Ofengand, 1993; Zaringhalam and Papavasiliou, 2016). Pseudouridine was reported to be present in yeast and human mRNAs and ncRNAs (Carlile et al., 2014; Li et al., 2015a; Lovejoy et al., 2014). By mass spectrometry analysis, a high pseudouridine to uridine ratio of 0.2-0.6 % was found in human mRNA and depletion/deletion experiments showed that these modifications are installed by the conserved pseudouridine synthetases of the Pus family (Carlile et al., 2014; Lovejoy et al., 2014; Schwartz et al., 2014a). Pseudouridylation has a high regulation potential because of its high abundance in mRNAs and indeed, it was shown that pseudouridylation is altered upon starvation induced stress in yeast and human cells (Carlile et al., 2014; Li et al., 2015a; Schwartz et al., 2014a). Different effects of the modification have also been observed on translation. *In vitro* assays using rabbit reticulocyte lysate and *in vivo* assays in mice and human cells show an increased translation rate and stability for pseudouridnated mRNAs (Kariko et al., 2008). The effect is the opposite in plant wheat germ extract and translation is completely abolished for multiple pseudouridinated mRNAs in *E. coli* lysate (Kariko et al., 2008). A second property of pseudouridine is the alteration of nonsense stop codons. Pseudouridination of UAA, UAG or UGA stop codons prevents the ribosome from recognising the stop codon and alternative amino acids are incorporated instead depending on the codon (Hoernes et al., 2016; Karijolich and Yu, 2011).

Recent improvements in modification detection techniques have enabled the detection of modifications in mRNAs but still, of the myriad of possible RNA modifications only four could be identified in the mRNAs so far. Further improvements to detection methods will likely lead to the discovery of even more types of modifications in mRNAs.

1.5 N⁶-methyladenosine

1.5.1 m⁶A in mRNA

A series of reports in the 1970 for the first time reported N⁶-methyladenosine (m⁶A) in RNA (Rottman et al., 1974; Schibler et al., 1977). The invention of new techniques, especially the development of next-generation sequencing, allowed a more sensitive analysis and revealed the extent of this modification in the transcriptome. Antibodies were developed that specifically detect the m⁶A modification and this allowed, in combination with deep sequencing, the detection of m⁶A in cellular mRNA (Chen et al., 2015; Dominissini et al., 2012). However, due to the approach used, the modifications could not be mapped to individual nucleotides, but only enabled the m⁶A modifications to be assigned to a 50-100 nucleotide region of specific mRNAs. Recently, the resolution was improved to single nucleotide level using specific mutational patterns of the m⁶A binding site introduced by the cross-linking method (Linder et al., 2015). Alternative mapping methods have since been developed to detect m⁶A modification
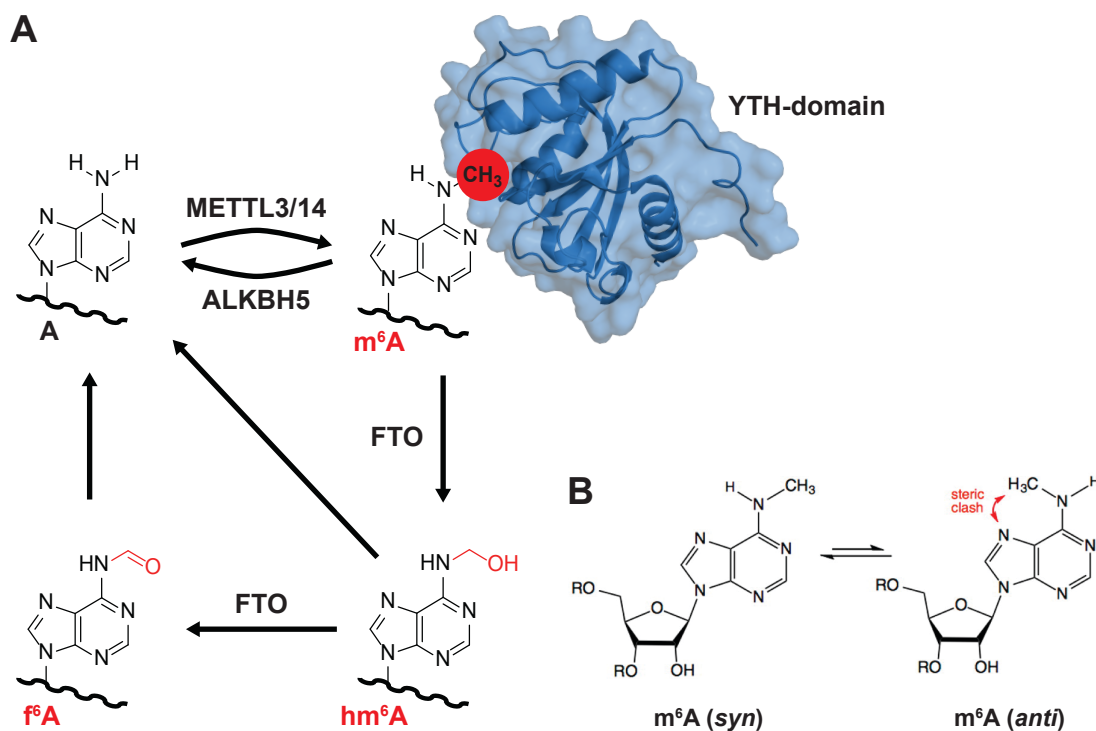
Figure 3: m⁶A is a dynamic modification. A Schematic representation of the methylation and demethylation reactions of the m⁶A modification. Methylation is accomplished in a one-step reaction, while demethylation can include several oxidative intermediates. The enzymes that mediate the different reactions are indicated. m⁶A modifications in RNAs can be specifically recognised by the YTH domains of specific proteins. YTH domain crystal structure (PDB 4RDN, Li et al., 2014). **B** Different possible conformational states of the methyl-group at N⁶. The *syn* conformation is energetically favoured, however, the *anti* conformation is adopted during Watson-Crick base paring in RNA double strands leading to steric interference of the methyl group with N⁷ of the purine ring. Modified from Roost et al., 2015.

sites independent of antibodies. Microarray based methods exploit altered base pairing properties of m⁶A compared to unmodified adenosine to detect the modification, however this approach is only suitable for highly enriched m⁶A sites (Li et al., 2015b). Similarly, site-specific cleavage and radioactive-labelling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET) can be used to detect m⁶As at single nucleotide resolution, however, it is not suitable for high throughput approaches, as it can only be used to confirm known m⁶A sites in RNAs (Liu et al., 2013).

Studies using these methods have revealed over 12,000 m⁶A sites in mRNAs and ncRNAs of over 7,000 human transcripts. Global positional analysis of m⁶A-containing mRNAs revealed an increase in m⁶A modifications around stop-codons, long internal exons and in the 3' UTRs of mRNAs (Chen et al., 2015; Dominissini et al., 2012; Meyer et al., 2012). Furthermore, the GGACU motif was highly enriched in the data, resembling the formerly established consensus motif RRACH, which was already proposed in the 1970's by chromatographically isolation and paper sequencing of m⁶A-containing mRNA oligonucleotides (Chen et al., 2015; Dominissini et al., 2012; Meyer et al., 2012; Schibler et al., 1977). Former studies also found m⁶A in introns of mRNAs (Carroll et al., 1990). Notably, the m⁶A modification is universally present in mRNA of lower and higher eukaryotes like human, mouse, fruit fly and yeast (Dominissini et al., 2012; Hongay and Orr-Weaver, 2011; Schwartz et al., 2013). However, in yeast the m⁶A modification is limited to over 1,000 mRNAs, restricted to meiosis and suggested to be highly regulated (Schwartz et al., 2013). Also, it is reported to influence the translation of certain mRNA transcripts during meiosis (Bodi et al., 2015).

The m⁶A modification can work as a molecular switch by changing the secondary structure of the RNA that can lead to the presentation of RNA binding motifs or structures for certain RNA binding proteins (Liu et al., 2016). The methyl group at position N⁶ of the adenine can accommodate either the *syn* or *anti* conformation (Figure 3B, Roost et al., 2015). The *syn* conformation has a lower energy and is the preferred position, because it avoids steric clashes with the purine ring of the base. However, during Watson-Crick base pairing, the *syn* conformation is not possible because it interferes with the hydrogen bonding network, thus pushing the methyl group in the less favoured *anti* conformation (Roost et al., 2015). This conformation has a higher energy and can destabilise duplexes in short double stranded regions. However, m⁶A also has increased base stacking capabilities leading to more stable single stranded structures, especially next to helices. Based on these findings a so-called 'spring loaded mechanism' is proposed, switching form double stranded to single stranded upon methylation.

On a cellular level the m^{6}A modification has implications in stress response and is part of the circadian rhythm of cells (Fustin et al., 2013; reviewed in Hastings, 2013; Meyer et al., 2015). In mouse, the m^{6}A modification was shown to affect the regulation of embryonic stem cells by keeping the omnipotence of the cells and might be involved in the cell cycle regulation in human cells (Dominissini et al., 2012; Wang et al., 2014b).

1.5.2 m^{6}A methyltransferases

The m^{6}A modification can be installed by a methylation complex comprised of the methyltransferase METTL3, the putative methyltransferase METTL14, and the regulatory proteins WTAP and KIAA1429 (Figure 3A, Liu et al., 2014; Ping et al., 2014; Schwartz et al., 2014b). Notably, other m^{6}A methyltransferases might also be involved in introducing such modifications, because knockdown of individual components of the METTL3/METTL14 complex does not abolish m^{6}A modification completely and the binding sites of the complex on cellular RNAs only partially overlap with the portion of the known m^{6}A modification sites (Chen et al., 2015; Liu et al., 2014; Schwartz et al., 2014b). The regulatory protein WTAP seems to have an influence on the position of the methylation within the mRNA, because WTAP-independent modification sites are mainly found at the 5' cap structure of mRNAs, whereas the installation of internal m^{6}A sites requires WTAP (Schwartz et al., 2014b). Recent studies identified METTL3 as the main subunit responsible for the modification (Wang et al., 2016b; Wang et al., 2016c). Structural and biochemical analyses showed that METTL3 and METTL14 form a heterodimer by forming a large hydrogen bond interaction network, resulting in a positively charged groove for RNA-binding. Mutational analysis of the SAM binding pocket of both methyltransferases revealed that METTL3 is the active, catalytic subunit. This is supported by a crystal structure of the heterodimer, which showed that only the binding pocket of METTL3 contained SAM and was highly conserved among methyltransferases while the binding pocket of METTL14 only showed low conservation (Wang et al., 2016b; Wang et al., 2016c). It is proposed that METTL14 supports the methylation activity of METTL3 by stabilising the conformation of METTL3, leading to a higher activity (Wang et al., 2016b). Identification of the binding sites of METLL3 and METTL14 on cellular RNAs by CLIP and motif analysis has revealed a GGAC motif, which is identical to the m^{6}A motif (Dominissini et al., 2012; Liu et al., 2014; Ping et al., 2014). The proteins are conserved in human, mouse, zebrafish and drosophila, suggesting that this mechanism of m^{6}A modification is conserved (Bokar et al., 1997; Liu et al., 2014; Ping et al., 2014; Schwartz et al., 2014b).

1.5.3   Oxidative demethylation

Interestingly, m^6A has been found to be a reversible modification (reviewed in Fu et al., 2014). It is either removed directly or by oxidative demethylation via N^6-hydroxymethyladenosine (hm^6A) or N^6-formyladenosine (f^6A) (Figure 3A). However, hm^6A and f^6A have a low stability under physiological conditions, displaying a half-life of only 3 h, which is very short compared to the half life of the m^5C oxidation products f^5C and hm^5C, thus the physiological relevance of hm^6A and f^6A has to be confirmed (Fu et al., 2013). The enzymes implicated in demethylation are the human AlkB homolog 5 (ALKBH5) and the fat mass and obesity-associated protein (FTO) (Jia et al., 2011; Zheng et al., 2013). Both proteins belong to the family of non-heme Fe(II)- and α-ketoglutarate-dependent dioxygenases. ALKBH5 directly demethylates m^6A, whereas FTO uses the oxidative demethylation pathway (Figure 3A, Fu et al., 2013; Zheng et al., 2013). In line with this, overexpression of FTO or ALKBH5 decreases the level of m^6A in mRNAs, whereas depletion has the opposite effect (Jia et al., 2011; Zheng et al., 2013). Both proteins posses structural features that enable them to specifically target single stranded nucleic acids and to regulate substrate specificity, which is important to avoid demethylating DNA as they are both localised in the nucleus (Aik et al., 2014; Feng et al., 2014; Han et al., 2010; Jia et al., 2011; Zheng et al., 2013; Zou et al., 2016).

FTO is mainly expressed in neuronal tissue in mice and is associated with several diseases. Overexpression of FTO leads to increased food intake causing obesity in mice, mainly by increasing the body fat mass (Church et al., 2010). Studies with patients affected by a loss of function mutation on the *fto* gene showed severe growth retardation and multiple malformations of the body and defects in several organs, including the central nervous system. Isolated fibroblasts showed a higher senescence, reduced proliferation rates and altered cell morphology, consistent with the death of the patients before the age of three (Boissel et al., 2009). This morphological effect might be the result of a splicing defect, because it was shown that FTO affects splicing of the adipogenic regulatory factor RUNXT1 by specifically targeting m^6As around 3' and 5' splice sites. An increase of m^6A modifications at these positions upon depletion of FTO leads to an elevated level of the splicing regulator SRS2, promoting inclusion of target exons (Zhao et al., 2014).

In contrast to FTO, ALKBH5 is mainly expressed in testes and is connected to spermatogenesis in mice (Zheng et al., 2013). Knockdown of ALKBH5 in mice leads to abnormal expression of spermatogenesis genes and cell apoptosis in testes. This might be due to aberrant mRNA processing because ALKBH5 was shown to localise to nuclear speckles and co-localise with RNA processing factors, i.e. phosphorylated SC35.

Upon depletion of ALKBH5, SC35 phosphorylation is lost in human cell lines and mRNA export is facilitated (Zheng et al., 2013). Beside the function in testes, ALKBH5 was shown to be a target of hypoxia induced transcription factor Hif1-α and plays a role in regulating pluripotency factors in breast cancer stem cells upon exposure to hypoxia (Thalhammer et al., 2011; Zhang et al., 2016).

1.6 Recognition of RNA modifications

1.6.1 The YTH domain and m^6A recognising proteins

In addition to the identification of METTL3/METTL14 as an m^6A methyltransferases complex (also termed m^6A 'writers') and ALKBH5/FTO as m^6A 'erasers', an exciting discovery was the identification of proteins that specifically recognise the m^6A modification in cellular RNA and thereby can influence the fate of the RNA. These proteins are called m^6A 'readers' (reviewed in Wang and He, 2014). The first proteins that were identified as such reader proteins all share a common feature, which is a specific protein domain called the <u>YT</u>521-B <u>h</u>omology (YTH) domain (Zhang et al., 2010), however, more recently, non-YTH domain-containing readers have also been identified (see section 1.6.3).

The YTH domain was first described in the human splicing factor YT521-B and shortly after, was defined as a new class of RNA-binding domain that is exclusively present in eukaryotes (Hartmann et al., 1999; Imai et al., 1998; Stoilov et al., 2002; Zhang et al., 2010). In humans, five YTH domain-containing proteins are known (YTHDF1, YTHDF2, YTHDF3, YTHDC1, YTHDC2) and for some of these proteins, different functions in RNA metabolism have been described (Theler et al., 2014; Wang et al., 2014a; Wang et al., 2015; Xiao et al., 2016; Xu et al., 2015). A crystal structure of the YTH domain of YTHDF2 revealed the recognition mechanism by which the m^6A is identified (Zhu et al., 2014). The YTH domain of YTHDF2 consists of three α-helices and eight β-strands (Figure 4). The six central β-strands (β8- β1- β3- β4- β5- β2) are arranged in an open β-barrel-type fold surrounded by the three α-helices, which together constitute the hydrophobic core (Li et al., 2014). Residues of the α1 α-helix, β2 β-strand and β4-β5 loop form a hydrophobic pocket for m^6A binding. The interactions are established by three highly conserved tryptophan residues, which build an aromatic cage around the m^6A. The adenine moiety is sandwiched between two parallel oriented tryptophans, while the methyl group is pointed towards the third one. Additional hydrogen interactions select for an adenine residue, locking the m^6A into place (Li et al., 2014). The area around the
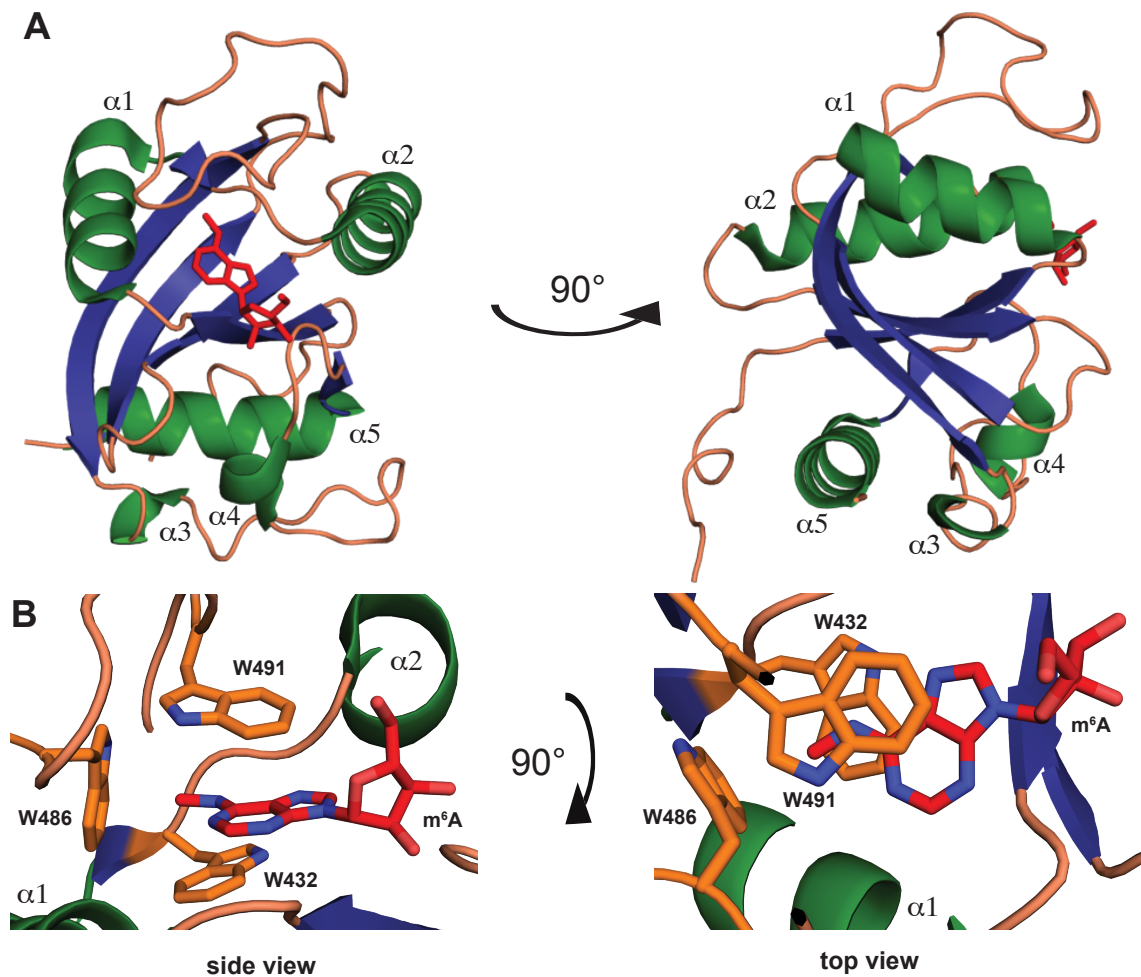
Figure 4: Crystal structure of the YTH domain. A Crystal structure of the YTH domain of YTHDF2 (PDB 4RDN, Li et al., 2014) represented in cartoon mode. α-helical secondary structures are coloured in green, β-strands are depicted in blue and flexible regions are shown in orange. α-helices are numbered from α1 to α5 starting at the N-terminus. The m^6A is represented in stick mode and is coloured in red. **B** Detailed view of the hydrophobic m^6A-binding pocket of the YTH domain. Important tryptophan residues (orange) defining the hydrophobic pocket and the m^6A modification (red) are represented as stick models and labelled accordingly. Nitrogen atoms of the stick models are coloured in blue.

m^6A-binding pocket is positively charged, resembling an RNA interaction surface (Li et al., 2014; Xu et al., 2015). The crystal structures of YTHDC1 and MRB1, which is a homologue of the yeast YTH domain-containing protein Pho92 in *Zygosaccharomyces rouxii*, showed similar structural properties (Luo and Tong, 2014; Xu et al., 2014).

The presence of a YTH domain is, however, not necessarily synonymous with m^6A recognition. Despite sequence and structural similarities, including a potential hydrophobic pocket for accommodating the m^6A residue, the yeast protein Mmi1 does not bind to the m^6A consensus motif (Wang et al., 2016a). It was shown to bind to the DSR motif instead, which is specific for meiotic transcripts in yeast. Mmi1 recognizes the motif via a long positively charged groove opposite of the potential m^6A binding pocket.

The area surrounding the potential m^{6}A binding pocket is, in contrast to the other YTH domains, negatively charged, thus repulsing potential m^{6}A-containing RNAs.

1.6.2 Functions of YTH domain proteins

YTHDF2 was the first protein of this family that was shown to bind to the m^{6}A modification *in vivo* using PAR-CLIP, a protein-RNA cross-linking based immunoprecipitation followed by deep sequencing of the co-purified RNA (Wang et al., 2014a). These data show that YTHDF2 binds to a subset of m^{6}A-containing mRNAs, while 59 % of the YTHDF2 binding sites overlap with m^{6}A sites identified by m^{6}A-Seq in the same cell line. In accordance with the m^{6}A pattern in mRNAs, YTHDF2 binds primarily around stop codons and within long exons. Ribosome profiling of YTHDF2 target mRNAs upon knockdown of YTHDF2 revealed reduced translation efficiency suggesting an increased pool of non-translatable mRNAs and a role for YTHDF2 in degradation of aberrant mRNAs. This is in line with the prolonged lifetime of YTHDF2 target mRNAs and an increase in m^{6}A/A ratio in total mRNAs and in the translatable pool after YTHDF2 depletion (Wang et al., 2014a). The SON mRNA was identified as a specific target of YTHDF2 where the C-terminal YTH domain of the protein specifically recognises m^{6}A modification in the mRNA, while the N-terminal part of the protein is responsible for the localisation of this complex to processing (p)-bodies for degradation of the mRNA (Wang et al., 2014a). A second function of YTHDF2 was reported under heat stress conditions (Zhou et al., 2016). YTHDF2, which is cytoplasmic under normal conditions, re-localises to the nucleus, where it is suggested to interact with m^{6}As in the 5' UTR of specific mRNAs, preventing FTO from demethylation of these sites. The increased methylation in the 5' UTRs is proposed to facilitate cap-independent translation of heat shock response genes (Meyer et al., 2015).

YTHDF1 was the second m^{6}A binding protein that was found to have a regulatory function on the mRNA level (Wang et al., 2015). However, in contrast to YTHDF2, YTHDF1 is involved in regulating translation efficiency of particular mRNAs. Knockdown of YTHDF1 does not alter the overall m^{6}A/A ratio in cells implying that it is not involved in RNA turnover. In contrast, knockdown of YTHDF1 leads to reduced ribosome occupancy of YTHDF1 target mRNAs and a reduced quantity of these mRNAs in the translated pool, suggesting a direct involvement in translation. Consistent with this, co-immunoprecipitation experiments also confirmed the interaction of YTHDF1 with translation initiation factors (Wang et al., 2015). Furthermore, tethering assays displayed a translation promoting effect of the N-terminal domain of YTHDF1, revealing an overall protein structure similar to YTHDF2, with an N-terminal protein-interaction domain and a

C-terminal YTH domain that mediates interactions with m^{6}A-containing mRNAs (Wang et al., 2014a; Wang et al., 2015). The finding that YTHDF1 seems to recognise m^{6}As near STOP codons and affects translation initiation by a direct interaction with the initiation factor eIF3 that primarily binds to the 5' end of mRNAs, led to the proposal that YTHDF1 takes advantage of the loop structure formed by eIF4G binding to the poly-A-binding-protein PABP and the initiation factor eIF4E (Wang et al., 2015). Also, YTHDF1 is suggested to keep the translation initiation complex primed during stress that is reducing translation, which then leads to a shorter recovery time after stress (Wang et al., 2015). Together, YTHDF2 and YTHDF1 form a tight regulation network for m^{6}A modified mRNAs, resulting in short-lived mRNAs with high translation rates, enabling rapid adaptation of gene expression in response to changing environmental conditions.

YTHDF3 is the least studied member of the three YTHDF proteins. Along with YTHDF1 and YTHDF2, it was shown that it binds to viral m^{6}A-containing RNAs of HIV-1 (Tirumuru et al., 2016). By binding to the viral RNA, the proteins block reverse transcription after HIV-1 cell infection. Overexpression of the three YTHDF proteins in different HIV-1 infected cells leads to decreased HIV-Gag protein expression, while knockdown has the opposite effect. Therefore, the three YTHDF proteins can have an influence on HIV-1 at the time of infection of the cell and at the time of virus production after integration of the viral RNA into the genome. Since this function of YTHDF3 is redundant with YTHDF1 and YTHDF2, it is likely that the main function of the protein in the cell remains to be identified.

For several years YTHDC1, formerly called YT521-B, was known to be involved in pre-mRNA splicing, however, the mode of regulation remained elusive and direct evidence was missing (Hartmann et al., 1999; Imai et al., 1998). Recently, it was shown that YTHDC1 functions in alternative splicing by interacting with splicing enhancer-binding SR proteins, specifically SRSF3 and SRSF10 that function to promote inclusion and skipping of their targeted exons, respectively (Xiao et al., 2016). *In vitro* experiments confirmed interactions between the N-terminal domain of YTHDC1 and the C-terminal domains of SRSF3 and SRSF10. However, *in vivo* YTHDC1 seems to interact mainly with SRSF3 and knockdown of YTHDC1 abolishes SRSF3 localisation to nuclear speckles and decreases its RNA binding. Depletion of YTHDC1 has the opposite effect on SRSF10, increasing its RNA-binding ability and promoting its localisation to nuclear speckles. The same is true when reducing the global m^{6}A level by knocking down METTL3, showing that the m^{6}A binding ability of YTHDC1 is important. Together, this suggest a model in which m^{6}A modifications can mark an exon for inclusion by recruiting the

YTHDC1-SRSF3 complex, while if the site is not modified, SRSF10 binds instead, promoting exon skipping.

Little is known about the close homologue of YTHDC1, YTHDC2. Reports connect it to pancreatic cancer and facilitated hepatitis C virus replication (Fanale et al., 2014; Morohashi et al., 2011). Lately, reports suggest that it is involved in cancer metastasis by enhancing the translation of hypoxia-inducible-factor-1alpha (HIF-1α) (Tanabe et al., 2016). Colon-tumour cells injected in mice showed a reduced metastasis rate compared to WT cells if YTHDC2 was knocked down and *in vitro* studies showed less cell mobility. This is supposed to be influenced by the increased translation level of HIF-1α. On a molecular level, YTHDC2 may be involved by unwinding secondary structures in the 5' UTR of HIF-1α, thus enhancing translation initiation (Tanabe et al., 2016).

1.6.3   Other m^{6}A modification readers

In addition to YTH domain-containing proteins, several other cellular proteins have recently been found to recognise m^{6}A. However, more evidence is needed in some cases to prove a direct interaction with the modification. The RNA stabilizing protein HUR was detected in RNA affinity assays using short RNA oligonucleotides containing the m^{6}A modification as a bait (Dominissini et al., 2012). HUR is known to bind to uridine-stretches in the 3' UTR of RNAs, thus its enrichment in m^{6}A pull downs was unexpected (Kishore and Stamm, 2006; Lebedeva et al., 2011). Further experiments lead to the conclusion that HUR is not recognising the modification itself, but instead is binding in close proximity to it. Depending on the presence of the modification, different secondary structures could change the accessibility of the HUR binding motif and extend the lifetime of the mRNA (reviewed in Wang and He, 2014).

Stress was shown to increase m^{6}A abundance in 5' UTRs, which facilitates cap-independent translation of particular mRNAs by increasing binding of translation initiation factors, without the help of YTH domain reader proteins (Meyer et al., 2015). *In vitro* and *in vivo* assays showed enhanced interaction of the eukaryotic initiation factor eIF3 with such mRNAs, dependent on the presence of modification. It has been suggested that this mode of translation initiation could be used during heat shock when the levels of m^{6}As in 5' UTRs are elevated, resulting in increased Hsp70 translation (Meyer et al., 2015; Zhou et al., 2015).

The nuclear protein HNRNPA2B1 is involved in micro (mi)RNA processing and alternative splicing. CLIP experiments showed that the binding sites of HNRNPA2B1 on cellular RNAs overlap known m^{6}A modification sites and m^{6}A-containing RNA could be detected in HNRNPA2B1 immunoprecipitations. 52 out of 53 miRNAs that contain m^{6}A

residues are reduced upon HNRNPA2B1 knockdown. These results, combined with the interaction of HNRNPA2B1 with the microprocessor protein DGCR8, led to the proposal that HNRNPA2B1 functions as a reader of the m^6A modification in miRNAs thereby effecting miRNA processing (Alarcon et al., 2015).

Surprisingly, a new study suggests that the m^6A writer METTL3 can also act as a non-conventional m^6A reader (Lin et al., 2016). METTL3 partially relocates to the cytoplasm in human cancer cells and seems to be involved in the translation of certain oncoproteins for increased cell proliferation, survival and invasion. The other components of the methylation complex WTAP and METTL14 are not relocated to the cytoplasm and the C-terminal methylation domain of METTL3 is dispensable for translation enhancement making it a unique feature of METTL3. The N-terminal domain METTL3 directly interacts with the initiation factor eIF3 recruiting it to the initiation complex in an YTHDF1-independent manner. Overall, translation regulation via the m^6A modification is a complicated process and it seems that several translation regulation mechanisms act in parallel on different mRNA transcripts.

1.7 Aims

RNA modifications are present in many cellular RNAs and can influence the properties and functions of the RNAs that carry them in a variety of different ways. However, the substrates of many modification enzymes are unknown and often the roles of specific modifications remain elusive. Recent findings highlight the N^6-methyladenosine (m^6A) modification as an important regulator of RNA metabolism as it was found to be present in mRNAs and was revealed to be reversible. Furthermore, proteins were identified that can specifically bind to this modification via a special recognition domain termed the YTH domain. A family of five YTH domain-containing proteins have been identified in human cells and YTHDF2 was shown to regulate the levels of specific m^6A-containing mRNAs by targeting them for degradation. At the outset of the project, however, RNA substrates of the other YTH domain-containing proteins and their cellular functions were not known.

The objectives of this study were the further development of computational tools to aid in the identification of RNA substrates of proteins involved in introducing or recognising RNA modifications and the genome-wide identification RNA interaction sites of the YTH domain-containing proteins.

Therefore, this work aimed to:

•   Adapt a bioinformatic pipeline for the genome-wide mapping of next-generation sequencing data derived from cross-linking and analysis of cDNA (CRAC) experiments in human cells

•   Identify the RNA targets of YTH domain-containing proteins using the CRAC approach

•   Analyse recognition of the m^6A modification by the YTH domains of the different proteins

•   Provide insight into the functions and interactions of YTHDC2

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals and enzymes

Chemicals were obtained from Applichem, Roth and Sigma-Aldrich at molecular biology quality or above. Enzymes were generally supplied by ThermoFisher scientific or NEB if not stated otherwise. Antibodies were purchased from Jackson ImmunoResearch, Bethyl or Sigma-Aldrich. Small interfering (si)RNA was ordered from Qiagen or were synthesised at MWG if sequence were available.

2.1.2 Oligonucleotides

Table 1 Oligonucleotides used in this study. Restriction sites are underlined; mutated nucleotides are lowercase.

Name	Sequence (5'-3')	Application
TUBG1_Fw	CGGCTGAATGACAGGTATCCTA	qPCR
TUBG1_Rv	CACCACATCGCTCATCTCGT	qPCR
GAPDH_Fw	CAGCCTCAAGATCATCAGCAATG	qPCR
GAPDH_Rv	GTCTTCTGGGTGGCAGTGATG	qPCR
YTHDC2_Fw	GCCAGAGGCAGCTAGTTTATTG	qPCR
YTHDC2_RV	GAGACCAAGGTTTAGATGGAGC	qPCR
YTHDC1_Fw	ACCAGGAAGTGGACAGACGA	qPCR
YTHDC1_Rv	TTCCTGGGTAAGGGGGCATT	qPCR
YTHDF2_Fw	TGTTAAAAAGGAACGTCAAGGTCG	qPCR
YTHDF2_Rv	GCAAGTCTGCAATCGTCTCTG	qPCR
YTHDC2seq1_Fw	ATTCAAAGATTTCCTGTCACCAA	sequencing
YTHDC2seq2_Fw	GCTTCGTACATTGATGGCAGGAGA	sequencing
YTHDC2seq3_Fw	AATGGATGCTTGCCTTTCTGATA	sequencing
YTHDC2seq4_Fw	CGGTTTGCTGACAGTACACATAGA	sequencing
YTHDC2seq5_Fw	AGATCTGACTGAACTTGGGTAT	sequencing
YTHDC2seq6_Fw	AAACTCTGAGAATTGGGCTGTCGT	sequencing
YTHDC2_BamHI_Fw	ATAT<u>GGATCC</u>ACCATGTCCAGGCCGAGCAG	cloning
YTHDC2_Nhe_Rv	ATAT<u>AGCTAGC</u>ATCAGTTGTGTTTTTTTCTCCCAAGG	cloning

DC2_192_BamHI_Fw	ATAT<u>GGATCC</u>CCAATGTCTTTACCAGTGTTTGAGAAACAGG	cloning
DC2_1287_NheI_Rv	ATAT<u>AGCTAGC</u>AGGCATGTTTGGTCTTGGCG	cloning
YTHDF1_BamHI_Fw	AAAA<u>GGATCC</u>GCCACCATGTCGGCCACCAGCGTGG	cloning
YTHDF1_NheI_Rv	AAAA<u>GCTAGC</u>TTGTTTGTTTCGACTCTGCCGTTCC	cloning
YTHDF2_KpnI_Fw	AAAA<u>GGTACC</u>GCCACCATGTCGGCCAGCAGCCTCTTG	cloning
YTHDF2_BamHI_Rv	AAAA<u>GGATCC</u>TTTCCCACGACCTTGACGTTCC	cloning
YTHDF3_BamHI_Fw	AAAA<u>GGATCC</u>GCCACCATGTCAGCCACTAGCGTGGATCAG	cloning
YTHDF3_NheI_Rv	AAAA<u>GCTAGC</u>TTGTTTGTTTCTATTTCTCTCCCTACGC	cloning
YTHDC1_KpnI_Fw	AAAA<u>GGTACC</u>GCCACCATGGCGGCTGACAGTCGGG	cloning
YTHDC1_NheI_Rv	AAAA<u>GCTAGC</u>TCTTCTATATCGACCTCTCTCCCC	cloning
DF2_380_BahmHI_Fw	ATAT<u>GGATCC</u>TCTACTCCTTCAGAACCCCACC	cloning
DF2_579_XmaI_Rv	ATAT<u>CCCCGGG</u>TTATTTCCCACGACCTTGACG	cloning
DC2_1277_BamHI_Fw	ATAT<u>GGATCC</u>TCAAAATCTCCTTCGCCAAGACC	cloning
DC2_1430_XmaI_Rv	ATAT<u>CCCCGGG</u>TCAATCAGTTGTGTTTTTTTCTC	cloning
DC1_344_BamHI_Fw	ATATAT<u>GGATCC</u>ACCAGTAAACTCAAATATGTGCTTC	cloning
DC1_509_XmaI_Rv	ATATATT<u>CCCGGG</u>CTAGTGACGCATTTTATGAATGACCTG	cloning
DC2_W1310A_Fw	CTCAACAGAAGGGTATcgcGTCTACAACTCCTAGTAATG	site directed mutagenesis
DC2_W1310A_Rv	CATTACTAGGAGTTGTAGACgcgATACCCTTCTGTTGAG	site directed mutagenesis
DC2_W1360A_Fw	GAAGGGAAAAGAGTCAGGAcgcGGGCTCTGCTGGACTAGGAG	site directed mutagenesis
DC2_W1360A_Rv	CTCCTAGTCCAGCAGAGCCCgcgTCCTGACTCTTTTCCCTTC	site directed mutagenesis
oligo(dT)	TTTTTTTTTTTTTTTTTTTTTTTTTVN	reverse transcription

2.1.3 Plasmids used in this study

Table 2 Plasmids used in this study.

ID	Name	Reference
pMB-044	A21-H10zzTevpQE80N(3)	Markus Bohnsack
pMB-187	pcDNA5/FRT/TO/FLAG	Invitrogen
pMB-1048	pcDNA5-FRT-TO-YTHDF1-cHisPrcFlag	This study
pMB-1049	pcDNA5-FRT-TO-YTHDF2-cHisPrcFlag	This study
pMB-1050	pcDNA5-FRT-TO-YTHDF3-cHisPrcFlag	This study
pMB-1047	pcDNA5-FRT-TO-YTHDC1-cHisPrcFlag	This study
pMB-1089	pcDNA5-FRT-TO-YTHDC2-cHisPrcFlag	This study
pMB-1098	pcDNA5-FRT-TO-YTHDC2_192-1430-cHisPrcFlag	This study
pMB-1100	pcDNA5-FRT-TO-YTHDC2_1-1287-cHisPrcFlag	This study
pMB-1099	pcDNA5-FRT-TO-YTHDC2_192-1287-cHisPrcFlag	This study
pMB-1222	A21-YTHDF2_380-579	This study
pMB-1246	A21-YTHDC1_344-509	This study
pMB-1221	A21-YTHDC2_1277-1430	This study
pMB-1241	A21-YTHDC2_1277-1430_W1310A	This study
pMB-1242	A21-YTHDC2_1277-1430_W160A	This study

2.1.4 siRNAs used in this study

Table 3 siRNAs used in this study.

Name	Sequence (5'-3')	Company
siYTHDC2_1	CGAAAUUGUUGGACUGAGA(dTdT)	eurofins MWG
siYTHDC2_2	GAAUUGGGCUGUCGUUAAA(dTdT)	eurofins MWG
siYTHDC2_3	CAUGAAAGGGAUCGAUUUA(dTdT)	eurofins MWG
siNT (control)	UCGUAAGUAAGCGCAACCC(dTdT)	Ambion

2.1.5 Antibodies used in this study

Table 4 Antibodies used in this study. WB, dilution used in western blotting. IF, dilution used in immunofluorescence

Name	Product number	Company	Fold dilution
Flag	F3165	Sigma-Aldrich	1:10,000 for WB; 1:500 for IF
Tubulin	T6199	Sigma-Aldrich	1:5,000 for WB
YTHDC2	A303-026A	Bethyl	1:10,000 for WB
XRN1	A300-443A-3	Bethyl	1:10,000 for WB
Goat-anti-mouse-HRP	115-035-003	Jackson ImmunoResearch	1:10,000
Goat-anti-rabbit-HRP	111-035-003	Jackson ImmunoResearch	1:10,000
Goat-anti-mouse-Alexa488	115-545-003	Jackson ImmunoResearch	1:1,000

2.2 Methods

2.2.1 Molecular cloning

The plasmids used in this study (see 2.1.3) were created using standard cloning techniques. In brief, cDNA was prepared from human cell culture cells (see 2.2.11) and the sequence of interest was amplified using sequence specific primers (see 2.1.2) and Phusion Polymerase (ThermoFisher scientific) according to manufacturer's instructions. 10 % of the reaction was analysed on a 1-2 % DNA agarose gel in 1x TAE (40 mM Tris pH 7.6, 20 mM acetate, 1 mM EDTA) to verify the correct size of the product. The rest of the reaction was purified using the PCR clean up Kit (Macherey-Nagel). To generate compatible ends 1-2 µg of vector or purified PCR product were incubated at 37°C for 2 h in a 15 µl reaction volume using restriction enzymes according to the manufacturer's instructions (ThermoFisher scientific). To increase the ligation efficiency and prevent re-circularisation, the vector was dephosphorylated for 20 min with 1 U Fast-alkaline phosphatase (ThermoFisher scientific). The DNA was separated on a 1-2 % agarose gel and gel purified using the Gel purification Kit (Macherey-Nagel). 50 ng of vector and a 5-fold molar excess of insert were used in a 20 µl ligation reaction using T4 DNA ligase (ThermoFisher scientific). The reaction was incubated for 2 h at 22°C or overnight at 16°C. Chemically competent Top10 cells were transformed with the ligation reaction and plated on LB-agar plates containing 100 µg/ml. After incubation at 37°C overnight, individual colonies were picked and 4 ml overnight cultures in LB-medium were inoculated. The next day plasmid DNA was isolated from the overnight cultures using a Mini-Prep plasmid isolation kit (Macherey-Nagel) according to the manufacturer's

instructions. The identity of the clones was verified by restriction digest and Sanger sequencing at GATC Biotech.

2.2.2 Site-directed mutagenesis

Site directed mutagenesis was used to introduce specific sequence mutations into plasmids. Mutagenic PCR primers were designed in such a way that the forward and reverse primer were fully complementary with the mutation placed in the middle of both primers with at least 15 flanking bases on each side ending with a guanine or cytidine. Three 50 µl reactions were set up with 5 ng, 20 ng or 50 ng of template plasmid. In addition, the reactions contain 125 ng of each primer, 0.2 mM of dNTP mix, PfuT buffer (10 mM Tris pH 8.85, 25 mM KCl, 5 mM $(NH4)_2SO_4$, 2 mM $MgSO_4$) and 2 U of homemade PfuT polymerase. The thermal cycling program was as follows: denaturation at 95°C for 30 sec, then 12-18 cycles of 95°C, 30 sec; 55°C, 1 min; 68°C, 120 sec/kb plasmid length. 12 cycles were used for a point mutation, 16 cycles for a triplet change and 18 cycles for multiple triplet insertions or deletions. After the PCR, the samples were briefly chilled on ice, 1 µl of *DpnI* was added and the reactions were incubated at 37°C for 2 h to fragment the methylated template DNA. Afterwards the three reactions were pooled and the DNA was precipitated by adding 0.1 volume of 3 M sodium acetate pH 5.2 and 3 volumes of 100 % ethanol, followed by incubation at -20°C for at least 1 h. After centrifugation at 20,000 rcf for 20 min at 4°C, the pellet was washed once with 70 % ethanol and resuspended in 10 µl ddH$_2$O. *E. coli Top10* cells were transformed and selected on appropriate agar plates. Plasmid DNA was isolated from single colonies as described above (2.2.1) and the presence of the mutation was confirmed by sequencing at GATC Biotech.

2.2.3 SDS-PAGE and western blotting

SDS-polyacrylamide gel electrophorese was performed according to Laemmli (Laemmli, 1970). Protein samples were pre-incubated with 1x SDS loading dye (60 mM Tris pH 6.8, 2 % SDS, 10 % glycerol, 0.01 % bromphenol blue, 1.25 % β-mercaptoethanol) and heated to 95°C for 5 min. The gels were run on a Biorad system according to the manufacturer's instructions. Subsequently, the gels were either stained using Coomassie brilliant blue (0.1% Coomassie brilliant blue G-250, 0.1 % Coomassie brilliant blue R-250, 10 % acetic acid, 40 % methanol) and destained (10 % acetic acid, 30 % methanol) or subjected to western blotting.

For western blotting, the proteins were transferred onto PVDF or nitrocellulose membrane at 100 V for 60 min in blotting buffer (250 mM Tris, 1.92 M glycine, 20 % methanol (v/v) using a wet blot system (BioRad). The membrane was then blocked using

TBS (50 mM Tris pH 7.6, 150 mM NaCl) supplemented with 5 % milk powder. Incubation with the primary antibody was done in 5 % milk in TBS for 3 h at room temperature (RT) or overnight at 4°C. After three times 10 min washing with TBS supplemented with 0.1 % Triton X-100 (TBS-T) the blots were incubated with the secondary antibody coupled to horseradish peroxidase for 1 h at RT. Three additional washing steps with TBS-T were performed, the blot was developed using Immobilon™ ECL solution (Millipore) and signals were detected by exposure of an x-ray film.

2.2.4 Cell culture

HeLa CCL2 or HEK Flp-In™ T-REx™ 293 (Invitrogen) cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % foetal calf serum (FCS) and 100 U/ml Penicillin-Streptomycin mix (Gibco). The cells were cultivated at 37°C, 5 % CO_2 in a dark and humid environment. To maintain viability, after washing with Phosphate buffered saline (PBS) (Gibco) the cells were detached using Trypsin-EDTA (0.25 %) (Gibco) and were seeded onto a new plate every three days in a 1:10 dilution.

2.2.5 Generation of HEK293 stable cell lines

The Invitrogen Flp-In™ system was used to generate stable cell lines. The gene of interest is cloned into the expression vector pcDNA5 vector generating a fusion protein with a C-terminal $His_6PrecFlag_2$-tag. Following transfection the construct is integrated at a specific location in the genome by a single recombination event utilising the Flp recombinase. Expression of the gene of interest is controlled by the CMV promoter and is inducible by tetracycline or doxycycline. The introduced hygromycin resistance gene is utilised for the selection of positive clones.

300,000 HEK Flp-In™ T-REx™ 293 cells were seeded into a well of a 6-well plate to be 50% confluent on the day of transfection. The next day, 0.6 µg pcDNA5 plasmid conteining the sequence of the ORF to be expressed was mixed with 1.8 µg pOG44 plasmid, which enables expression of the Flp recombinase. In a second tube 9 µl X-treme Gene HP DNA transfection reagent was mixed with 91 µl Opti-MEM reduced serum medium (ThermoFisher scientific) and incubated at RT for 5 min. The two reactions were mixed and incubated at RT for additional 15 min. Fresh DMEM media without penicillin-streptomycin was applied to the cells, the transfection mixture added in a drop wise manner cells were cultivated for two days. The cells were then trypsinised and transferred to a 10 cm dish and selection for positive transformats was started by addition of 100 µg/ml hygromycin B (Applichem) and 10 µg/ml blasticidin S (Applichem).

The cells were washed with PBS and fresh medium with antibiotics was supplied every two to three days until single colonies were visible under the light microscope. At this stage the cells were trypsinised and resuspended to acquire a mixed population. At 80 % density the cells were split onto a new plate and hygromycin B selection was maintained during the cultivation of the cells and blasticidin S was added every third split.

To verify expression of the tagged protein 125,000 cells were seeded into a well of a 24-well plate and induced with 1 µg\ml doxycycline for 24 h. The medium was removed and the cells were washed once with PBS before adding 100 µl of 1x SDS-loading dye. 10 % of the sample was analysed by SDS-PAGE and Western blotting using specific antibodies against the tagged protein.

2.2.6 Immunofluorescence

Sterile poly-L-lysine coated coverslips were placed into 24-well plates and cells were seeded to be 50 % confluent at the time of fixation. The cells were washed three times with PBS before fixation with 4 % paraformaldehyde in PBS at RT for 20 min. The coverslips were washed again three times with PBS and the cells were permeabilised using 0.1 % Triton X-100 diluted in PBS for 15 min at RT. To stop the permeabilisation reaction, four additional PBS washes were performed. Reduction of background signal was accomplished by 1 h incubation in blocking solution (10 % FCS, 0.1 Triton X-100, PBS) at RT. The primary antibody, diluted in blocking solution, was then applied to the coverslips for 1-2 h at RT. Excess primary antibody was removed by three quick washes with PBS followed by three 10 min washes in PBS. The fluorescently labelled secondary antibody was applied to the coverslips in the same way as the primary antibody and after incubation for 1-2 h at RT, the same washing steps were performed. Finally, the coverslips were dipped in ddH2O and ethanol then dried before mounting onto glass slides using 4 µl of Vectrashield mounting medium (Vector laboratories). Confocal microscopy was performed on a LSM 510 META (Zeiss).

2.2.7 Immunoprecipitation

Immunoprecipitation experiments were conducted using HEK293 cell lines expressing a stably integrated His-Prc-2xFlag tagged version of the gene of interest (2.2.5) and αFlag M2 magnetic beads (Sigma Aldrich) for precipitation. Tagged protein expression was induced at the endogenous protein level with doxycycline. Approximately $1.4 \cdot 10^7$ cells were harvested using gentle PBS washing to detach the cells from the plate. The cells were harvested by centrifugation at 200 rcf for 3 min. The supernatant was aspirated and the pellet was snap-frozen in liquid nitrogen and stored at -80°C or kept on ice for immediate use. The pellet was resupended in 1 ml lysis buffer (150 mM KCl, 20 mM

Hepes pH 8.0. 0.1 mM DTT) and sonicated on ice (20 % amplitude, 0.3 sec pulse, 0.7 sec off, 3x 16 pulses, 20 sec intervals). After the sonication step, the lysate was supplemented with 0.2 % Trition X-100, 10 % glycerol and 1.5 mM MgCl$_2$. Cell debris was pelleted by centrifugation at 20,000 rcf, for 10 min, at 4°C. 50 µl of αFlag M2 magnetic beads were equilibrated by washing three times in ice cold IP buffer (150 mM KCl, 20 mM Hepes pH 8.0. 0.1 mM DTT, 0.2 % Trition X-100, 10 % glycerol, 1.5 mM MgCl$_2$). After centrifugation the cleared lysate was directly added to the equilibrated beads or was subjected to RNA digest. For the RNA digest a 1:1,000 dilution of RNase A (4.5 U/µl stock solution) and RNase T1 (1 U/µl stock solution) was added and samples were incubated for 15 min at RT before adding to the equilibrated beads. After incubation for 2 h at 4°C on a rotating wheel none-specifically bound proteins were removed by washing the beads four times with 1 ml ice cold IP buffer. During the last wash the beads were transferred to a new tube. The Flag-tagged proteins and any co-precipitated proteins were then eluted by adding 50 µl of Flag-peptide solution (250 µg/ml Flag-peptide, 20 mM HEPES pH 8.0, 150 mM KCl and 1.5 mM MgCl$_2$, 0.05% Triton X-100) and incubating the suspension shaking for 30 min at 12°C. The elution step was repeated and the two elution fractions were pooled. Precipitation of the proteins was facilitated by trichloroacetic acid (TCA) to a final concentration of 20 % and incubation on ice for 20 min. Finally, the samples were centrifuged at 20,000 rcf for 20 min at 4°C. The supernatants were discarded and the pellets were washed once with ice-cold acetone, briefly centrifuged at 20,000 rcf for 10 min at 4°C and air dried for 5 min. The pellets were resuspended in 4x NuPage loading dye (Invitrogen) supplemented with 50 mM DTT and separated on a 4-12 % NuPage Bis-Tris gel system (Invitrogen) with MES running buffer (50 mM MES, 50 mM Tris, 1 mM EDTA, 0.1 % SDS) if the samples were to be analysed by mas spectrometry. Alternatively, samples were resuspended in 1x SDS-loading dye and analysed by SDS-PAGE and subsequent western blotting.

2.2.8 Pre-ribosome and sucrose density gradients and polysome profiling

Sucrose density gradients were performed to separate ribosomal particles according to their sedimentation coefficient. 10^7 cells were harvested using trypsin digestion and pelleted. After shock frosting in liquid nitrogen the pellet was stored either at -80°C for later use or directly thawed on ice for imminent use. After resuspension of the pellet in 500 µl lysis buffer (50mM Tris pH 7.5, 100 mM NaCl, 5 mM MgCl$_2$, 1 mM DTT) the cells were lysed by sonication (20 % amplitude, 0.3 sec on, 0.7 sec off, 3x 20 pulses, 20 sec intervals). Prior to loading on the sucrose gradient the lysate was cleared by

centrifugation at 20,000 rcf for 15 min at 4°C. To prepare 10-45 % sucrose gradients, 10 % sucrose solution (w/v) and 45 % sucrose solution (w/v), both prepared in lysis buffer, were layered and mixed using the GradientMaster (Biocomp) "short sucrose 10-45 %" program with the following settings: 82.0° angle, 19 speed, rotation time 1 min 25 sec. The prepared gradients were stored at 4°C for 1 h before use. 400 µl of sucrose solution was removed from the top of the gradient and was replaced by 400 µl of cleared lysate. The gradients were centrifuged at 23,500 rpm, 16 h, 4°C in a Beckman Coulter SW40-Ti swinging bucket 6x 14 ml rotor. After centrifugation, 530 µl fractions were taken by hand from the top. The RNA content of the fractions was recorded at 254 nm using the NanoDrop ND-2000c (ThermoFisher). Proteins in each fraction were precipitated using TCA (2.2.7) and analysed by SDS-PAGE and subsequent western blotting (2.2.3).

2.2.9   RNA interference

RNA(i) interference was carried out using small interfering (si)RNAs (Table 3) to transiently knockdown specific genes. Cells were seeded one day before transfection to reach 20-30 % confluence at the time of transfection. For one well of a 6-well plate, 4 µl (HeLa) or 5 µl (HEK Flp-In™ T-REx™ 293) of RNAiMax transfection reagent (Invitrogen) and 500 µl Opti-MEM reduced serum medium (ThermoFisher scientific) were premixed before adding the mixture to 30 nM siRNA (final concentration) and further incubation for 15 min at RT. The medium on the cells was changed for fresh DMEM without penicillin/streptomycin and the transfection mix was added to the cells in a drop-wise manner. The next day the medium was exchanged and 72 h or 96 h after siRNA transfection the cells were harvested. For a transfection in a 10 cm dish the volumes were tripled.

2.2.10 RNA extraction

Total RNA from human cells was extracted using TriReagent (Sigma-Aldrich) according to manufacturer's instructions. In brief, the cells were washed once with PBS and for one well of a 6-well plate 1 ml of TriReagent was added and incubated at RT for 1 min. The lysate was transferred into a tube and 200 µl chloroform were added. After thorough mixing, the samples were incubated at RT for 2 min and then centrifuged at 12,000 rcf, for 15 min, at 4°C for phase separation. The upper aqueous phase was carefully transferred to a fresh tube, 500 µl isopropanol was added to the upper phase and samples were incubated for 5 min at RT to precipitate the RNA. After centrifugation at 20,000 rcf, for 15 min, at 4°C the supernatant was discarded and the RNA pellet was washed once with 70 % ethanol. After centrifugation at 20,000 rcf, for 5 min, at 4°C the

supernatant was removed and the pellet was air-dried for 5 min then resuspended in 20 µl diethylpyrocarbonate (DEPC) treated ddH$_2$O.

2.2.11 Quantitative real time PCR

Total RNA was reverse transcribed using the Superscript III reverse transcriptase (Invitrogen) according to the manufacturer's instructions to produce cDNA. 2 µg to 5 µg total RNA were mixed with 50 µM oligo(dT) primer (Table 1) or 186 ng random hexamers and 1 µl 10 nM dNTPs. ddH$_2$O was added to 13 µl total volume and the sample was incubated at 65°C for 5 min. Afterwards the sample was shortly chilled on ice before adding 1 µl 100 mM DTT, 4 µl 5x Superscript III reaction buffer and 1 µl Superscript III. Reverse transcription was carried out at 55°C for 1 h. Gene specific primers were designed using NCBI Primer-BLAST to achieve high specificity. If applicable, exon-exon spanning primers were designed to avoid the amplification of DNA contaminations. The primers were designed to have a melting temperature between 57°C and 63°C and the product length had to be between 80 nt and 150 nt. To test the primers a serial dilution of human cDNA was prepared from 1:5 to 1:625 with a dilution factor of 5. Samples were analysed using absolute quantification and 2nd derivative max method. C$_T$ values were plotted against the dilution factor and only primers with an amplification efficiency greater than 90 % were used for qPCR (Table 1). Quantitative real time (q)PCR was carried out using a Light cycler 480 system (Roche) with the supplied Light cycler 480 SYBR Green I Master mix (Roche). For a single well 2.5 µl diluted cDNA, 0.3 µl 10 mM primer mix, 3.3 µl SYBR Green Master mix and 3.9 µl ddH$_2$O were used. The following qPCR program was used: 1x (95°C, 5 min); 50x (95°C, 30 sec; 58°C, 30 sec, 72°C, 30 sec) Each reaction was pipetted in triplicates to account for pipetting errors. Two of the triplicates had to be within 0.5 threshold cycles (C$_T$) otherwise the reaction was repeated. The experiments were analysed using the Roche Lightcycler 480 software version 1.5.1.62. C$_T$ values were calculated using the 2nd derivative max method of the software and relative quantification was conducted using the $\Delta\Delta$C$_T$ method (Schmittgen and Livak, 2008). The results were normalised to *GAPDH* (NG_007073.2), COPS6 (NG_046973.1) or TUBG1 (NG_033886.1) to consider different cDNA amounts.

2.2.12 Agarose-glyoxal gel electrophoresis and northern blotting

5 µg of total RNA was mixed with glyoxal loading dye (61 % DMSO (v/v), 20 % glyoxal (v/v), 5 % glycerol (v/v), 1x BPTE) and incubated at 55°C for 1 h. A 1.2 % agarose gel was prepared using ultra pure agarose (Invitrogen). After loading the samples the gel was run at 60 V for 16 h in 1x BPTE (10 mM Pipes, 30 mM Bis-Tris, 1 mM EDTA). Afterwards the following washing steps were applied to denature the RNA and enable

subsequent transfer; 100 mM sodium hydroxide for 20 min, 0.5 M Tris/1.5 M NaCl for 15 min twice and 6x SSC (900 mM NaCl, 100 mM sodium citrate) for 15 min. The RNA was then vacuum blotted onto Hybond N membrane (GE Healthcare) at approximately -300 psi for 2 h. After blotting the RNA was cross-linked to the membrane with two times 180 mJ per cm^2 256 nm UV light. To check for efficient transfer and visualise abundant RNAs, the membrane was stained with methylene blue (0.3 M sodium actetate pH 5.2, 0.1 % methylene blue (w/v)) The membrane was then incubated in pre-hybridisation buffer (SES1; 0.25 M sodium phosphate pH 7.0, 7 % SDS (w/v), 1 mM EDTA) for 30 min at 37°C, rotating. Antisense DNA probes were radioactively labelled in a 20 μl reaction using 1 μl 10 U/μl T4 Polynucleotide kinase (ThermoFisher scientific), 2 μl ^{32}P-γ-ATP (2 μCi, Perkin Elmer), 2 μl 10 mM oligonuclotide, 2 μl 10x PNK buffer, 12.5 μl ddH_2O. After labelling, the probe was diluted in SES1 and incubate with the membrane overnight at 37°C, rotating. The membrane was then washed with 6x SSC for 30 min at 37°C and for 30 min with 2x SSC supplemented with 0.1% SDS at 37°C before drying the membrane. Radioactive signals were visualised by exposing the membrane to a phosphorimager screen (GE Healthcare) and detection using a Typhoon FLA 9500 laser scanner (GE Healthcare). If re-probing was necessary, the membrane was stripped by incubation with 0.1x SSC supplemented with 0.1 % SDS (w/v) at 70°C for 1 h with shaking before blocking and probing as described above.

2.2.13 Pulse-chase labelling of RNA

Cells were transfected with siRNAs in 6-well plates as described in section 2.2.9. After 72 h the media was replaced by phosphate-free DMEM and cells were grown for a further 1 h. ^{32}P-orthophosphate was added to a final concentration of 15 μCi/ml to fresh phosphate-free DMEM and the cells were grown in this media for an additional 1 h. The labelled media was then removed and the cells were washed with PBS. Normal media was added to the cells and they were grown for a further 3 h. Then the medium was removed and the cells were washed with PBS. RNA was extracted using TriReagent (Sigma-Aldrich) and separated on an agarose-glyoxal gel (2.2.12) at 185 V for 3 h. The RNA was capillary blotted onto Hybond N membrane and radioactive signals were visualised using a Typhoon FLA 9500 (GE Healthcare).

2.2.14 Cross-linking and analysis of cDNA (CRAC)

10^8 cells of stably transfected HEK293 cell lines for expression of His$_6$-PreScission protease site-Flag$_2$ tagged proteins were used per CRAC experiment. Protein expression was induced either with 1 μg/μl doxycycline or the optimal concentration to mimic endogenous expression for 36 h. 100 μM 4-thiouridine was added to each plate 6 h

before cross-linking if photoactivatable-ribonucleoside-enhanced (PAR)-CRAC was done. Before cross-linking the media was removed and cells were washed once with 30 ml of PBS and 8 ml of PBS was added. The cells were irradiated with 3 x 800 mJ/cm^2 at 254 nm (UV-CRAC), or 2 x 180 mJ/cm^2 at 365 nm (PAR-CRAC) using a Stratalinker at RT. Subsequently, the PBS was removed and 200 µl of TMN150 (50 mM Tris pH 7.8, 150 mM NaCl, 1.5 mM MgCl$_2$, 0.1 % NP-40 (v/v), 5 mM beta-mercaptoethanol) supplemented with complete mini protease inhibitor EDTA free (Roche) was added before harvesting of the cells by scraping. The cells were then lysed by sonication (40 % amplitude, 0.5 sec pulse, 0.5 sec off, 3 x 15 pulses, 20 sec intervals). The lysate was cleared by centrifugation at 20,000 rcf for 15 min at 4°C. The cleared lysate was added to 50 µl anti-Flag-magnetic beads (Sigma-Aldrich) that had been pre-equilibrated in lysis buffer and the samples were incubated for 3 h rotating at 4°C. The beads were then washed twice with 500µl TNM1000 (50 mM Tris pH 7.8, 1 M NaCl, 1.5 mM MgCl$_2$, 0.1 % NP-40 (v/v), 5 mM beta-mercaptoethanol) and three times with TMN150. The resin was transferred to a new tube during the last washing step. The bound protein was eluted by adding 200 µl of TMN150 supplemented with 0.2 µg/µl Flag peptide (Sigma-Aldrich) and incubation overnight rotating at 4°C. The eluate was transferred to a fresh tube, the volume was adjusted to 600 µl with TMN150 and the sample was subjected to RNase treatment. 0.1 U RNace-IT was added and the samples were incubated for 30 sec at 37°C. The reaction was stopped by transferring the sample to a 1.5 ml reaction tube containing 0.4 g guanidine hydrochloride, 45 µl 3 M NaCl and 3 µl 2.5 M imidazole pH 8. For the denaturing binding step, the samples were incubated with 50 µl of Ni-NTA resin (Qiagen) for 2 h rotating at 4°C that had been pre-equilibrated in wash buffer I (50 mM Tris pH 7.8, 300 mM NaCl, 10 mM imidazole pH 8, 6 M Guanidine hydrochloride, 0.1 % NP-40 (v/v), 5 mM beta-mercaptoethanol). The beads were washed twice with 750 µl wash buffer I and three times with PNK buffer (50 mM Tris pH. 7.8, 10 mM MgCl$_2$, 0.5 % NP-40 (v/v), 5 mM beta-mercaptoethanol). With the last wash the beads were transferred to Mobicol spin columns (Bio-Rad) and incubated with 8 U TSAP (Promega), 60 U of RNasin (Promega) in PNK buffer for 30 min at 37°C to dephosphorylate the RNA. The beads were washed once with wash buffer I to stop the reaction and three times with PNK buffer before ligation of the Illumina RA3 3' adapter to the 3' end of the RNA. The reaction was carried out in PNK buffer with 1 µM RA3 3' adapter, 800 U T4 RNA ligase 2 deletion mutant (Epicentre), 60 U RNasin (Promega), 10 % PEG8000 (Sigma-Aldrich) overnight at 16°C. The beads were washed once with wash buffer I to stop the reaction and three times with PNK buffer before radioactive labelling of the 5' end of the RNA. A mix of 80 U T4 PNK (NEB), 60 U of RNasin (Promega) in PNK buffer was added first before the addition of ^{32}P-γ-ATP and incubation for 40 min at 37°C. 1.25 µM Li-ATP

(Roche) was added and the samples were incubated for an additional 20 min. The samples were washed once with wash buffer I to stop the reaction and three times with PNK buffer before ligation of the RA5 (N5) 5' adapter to the 5' end of the RNA. The reaction was carried out in PNK buffer with 40 U T4 single-strand RNA ligase I (NEB), 1.25 µM RA5 (N5) 5' adapter, 1 mM ATP (Roche) overnight at 16°C. After the ligation reaction the samples were washed three times with 400 µl of wash buffer I and seven times with wash buffer II (50 mM Tris pH 7.8, 50 mM NaCl, 10 mM Imidazol pH 8, 0.1 % NP-40 (v/v), 5 mM beta-mercaptoethanol). The RNA-protein complexes were eluted twice with 200 µl Elution buffer (50 mM Tris pH 7.8, 50 mM NaCl, 150 mM imidazole, 0.1 % NP-40 (v/v), 5 mM beta-mercaptoethanol). The samples were subjected to TCA precipitation before separation in a 4-12 % NuPAGE gel as described in section 2.2.7. The proteins were transferred onto Hybond C membrane (Amersham) in transfer buffer (25 mM Bicine, 25 mM Bis-Tris, 1 mM EDTA, 20 % methanol) and radioactive signals were detected by exposure of the membrane to an x-ray film for 1-16 h. Membrane segments corresponding to the radioactive signals were excised and the RNA was eluted from the membrane by protein digestion with 100 µg Proteinase K, PCR grade (Roche) in 400 µl wash buffer II supplemented with 1 % SDS and 5 mM EDTA overnight shaking at 55°C. 50 µl 3 M sodium acetic acid pH 5.2 and 500 µl phenol:chloroform:isoamylalcohol (25:24:1) were added and centrifuged at 20,000 rcf for 5 min at RT. The upper phase was transferred to a new tube, 20 µg glycogen was added and the RNA was precipitated with 3 vol. of 100 % ethanol overnight at -20°C. The RNA is precipitated at 20,000 rcf for 30 min at 4°C and the pellet was washed with 70°C ethanol once. The RNA pellet was directly resuspended in the following mix: 10 µL ddH$_2$O, 1 µl RTP Primer (10 µM), 2 µl dNTP mix (5 mM, Roche). Reverse transcription was carried out using the Superscript III kit (Invitrogen). After inactivation if the reverse transcriptase by incubation at 65°C for 15 min, 1 µl of the 20 µl reaction was used for the following PCR reaction: 1x LA Taq buffer+MgCl$_2$, 0.2 µM RPI (Illumina forward index primer), 0.2 µl RP1 (Illumina general reverse primer), 0.125 mM dNTPs, 2.5 U La TakaRa Taq (Clontech). (1x 95°C, 2 min; (40x 98°C, 30 sec.; 60°C, 40 sec; 68°C, 40 sec); 72°C, 5 min). The PCR products were extracted using phenol:chloroform:isoamylalcohol as described above and were precipitated by the addition of sodium acetate to a final concentration of 0.3 M and 3 volumes of 100 % ethanol then incubation at -20°C overnight. After centrifugation at 20,000 rcf for 20 min at 4°C, the pellet was washed with 70 % ethanol, dried and resuspended in 2x gel loading dye (Qiagen). The PCR product was separated on a 3 % Metaphore agarose gel (Lonza) in TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA) and gel purified using the Qiagen MINI elute Kit and RNA-binding columns (Qiagen). Afterwards, the DNA concentration

was quantified using the Qubit quantification system (Invitrogen) and the library was sent for Illumina deep sequencing.

2.2.15 Genome-wide mapping of deep sequencing data

The data received from Illumina deep sequencing were mapped onto the human genome version GRCh 37.75 from ensembl (grch73.ensembl.org). In addition one human ribosome repeating unit (NCBI accession number U13369.1) and one 5S rRNA sequence were added as additional chromosomes. The sequences of the mature rRNAs are based on the rRNA sequences shown in the structure of the human 80S ribosome (PDB 4V6X; (Anger et al., 2013). Existing 5S sequences similar to the introduced sequence were masked with "N" to get a specific mapping onto the introduced 5S sequence. A modified version of the ensembl GTF annotation file was used. Wrongly categorised genes were corrected by hand and the tRNA annotation based on tRNAscan-SE from Ensembl and the annotation of the introduced rDNA sequences were added.

The sequences obtained from Illumina were 50 bp single read sequences comprised of an introduced random barcode NNNNNAGC at the 5' end to distinguish PCR artefacts from the accumulation of independent sequences, the actual sequence and the 3' linker in the case of short reads. Thus, several quality-processing steps have to be made before the read can be mapped onto the genome. Firstly, the introduced barcode is removed using pyBarcodeFilter from the pyCRAC utility (Webb, 2014). Afterwards the read is quality based trimmed, a potential 3' linker sequence is removed and reads shorter than 21 nucleotides are discarded using Flexbar 2.7 with the following settings: -q TAIL –qf i1.8 -qt 13 –as TGGAATTCTCGGGTGCCAAGG -ao 2 -u 0 -m 21 -ae RIGHT. Duplicated sequences are removed using "pyFastqDuplicateRemover" script from the pyCRAC utility and the remaining sequences are mapped onto the genome. If PAR-CRAC was applied the sequences are mapped with Bowtie 1.1.2 with one mismatch allowed and only the best alignments with one T to C conversion are filtered by self-written python 2.7 scripts and are considered for further analysis. If UV-CRAC was applied the sequences are mapped with Bowtie2 2.2.6 using the following settings: -D 20 -R 3 -N 1 -L 20 -i S,1,0.50 --rdg 0,11 --mp 10,6 --ignore-quals. Further analysis and normalization of the data were done making use of self-written python 2.7 scripts and scripts from the pyCRAC utility including "pyReadCounters", "pyPileup" and "pyReadAligner" scripts.

For visualization the data were imported into MochiView 1.46 (www.johnsonlab.ucsf.edu/mochi/) and hits on the rRNA were mapped onto the 2D (Petrov et al., 2014) and 3D structure (PDB 4V6X, Anger et al., 2013) of the mature human ribosome using self-written python 2.7 scripts and in addition the 3D protein viewer pymol (www.pymol.org) for the 3D structure.

2.2.16 Recombinant expression of proteins in *E. coli*

For recombinant expression of proteins in *E. coli* BL21 codon plus cells (Agilent) the pQE derived A21 vector was used, which introduces a N-terminal His_{10}-ZZ-TEV-tag to the protein. The main culture was inoculated from an overnight culture and grown to an OD_{600} of 0.5 at 37°C before switching to 16°C and induction with 0.3 mM of IPTG overnight. The cells were harvested by centrifugation at 6,000 rcf for 15 min at 4°C. They were washed once with PBS before the pellet was snap-frozen in liquid nitrogen and stored at -20°C.

2.2.17 Purification of His-tagged proteins

A pellet equivalent to a 1 L culture was resuspended in 20 ml lysis buffer (30 mM Tris pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 1 mM PMSF, 10 % glycerol) and the cells were lysed using sonication (45 % amplitude, 0.7 sec on, 0.3 sec off, 4x 30 pulses with 30 sec pause in between). The lysate was cleared by centrifugation at 25,000 rcf for 30 min at 4°C. 1 ml of cOmplete His-Tag purification resin (Roche) was pre-equilibrated with 10 ml of lysis buffer and the lysate was added three times for binding using the gravity flow technique. The resin was then washed two times with 10 ml wash buffer (30 mM Tris pH 7.4, 300mM NaCl, 15 mM Imidazole pH 8, 1 mM $MgCl_2$, 10 % glycerol) and finally the bound protein was eluted with 8 ml elution buffer (30 mM Tris pH 7.4, 150mM NaCl, 500 mM Imidazole pH 8, 1mM $MgCl_2$, 10% glycerol) and collected in 1 ml fractions. 1µl of each fraction was spotted onto a nitrocellulose membrane (Amersham), which was then stained with 0.1 % amidoblack in 50 % ethanol. Fractions containing protein were pooled and dialysed against 1 L dialysis buffer (30 mM Tris pH 7, 120 mM NaCl, 2 mM $MgCl_2$, 20 % glycerol) at 4°C overnight using Spectrum dialysis tubing that had been briefly boiled in 5 mM EDTA before use. After dialysis, the protein concentration was measured using the Pierce Coomassie plus Bradford assay kit (ThermoFisher scientific) and the protein was aliquoted, snap-frozen in liquid nitrogen and stored at -80°C.

2.2.18 Anisotropy

For anisotropy measurements, proteins were first dialysed against anisotropy buffer (30 mM Tris pH 7.5, 120 mM NaCl) overnight. Increasing concentrations of protein were incubated with 20 nM of fluorescein labelled RNA (Table 1) in anisotropy buffer for 5 min at RT. The samples were then transferred to a Quartz SUPRASIL® 10x2 mm High Precision Cell cuvette (Hellma Analytics) for measurement. Anisotropy measurements were performed on a FluoroMax-4 spectrofluorometer (Horriba Scientific) with the FluorEssenceV3.5 software at 25°C. The excitation and emission wavelengths were set to 470 nm 517 nm, respectively. Excitation slit width was set to 5 nm and emission slit width was set to 10 nm. The integration time was 1 sec, the maximal trials per sample were set to 6 and the target standard error was 2 %. The data were fitted with formula 1 and dissociation constants were calculated using the Origin 8.2 software.

1 $\qquad r = r_0 + \frac{\Delta r_{max}}{[RNA]_{tot}} \cdot \left(\frac{[protein]_{tot}+[RNA]_{tot}+K_d}{2} - \sqrt{\left(\frac{[protein]_{tot}+[RNA]_{tot}+K_d}{2}\right)^2 - [protein]_{tot}[RNA]_{tot}} \right)$

r_0, anisotropy of unbound RNA; r_{max}, amplitude; *[protein]*$_{tot}$, total protein concentration; *[RNA]*$_{tot}$, total RNA concentration

3 Results

3.1 Bioinformatic analysis of high throughput next-generation sequencing data

RNA binding proteins are a major part of the proteasome of the cell. For characterisation of these proteins, it is essential to identify their RNA interaction partners. Therefore, cross-linking methods have been established that allow the purification of RNA-protein complexes with subsequent next-generation sequencing of the RNA. The first method that was established was called CLIP for cross-linking and immunoprecipitation (Darnell, 2012). Therein, UV light is used to establish covalent bonds between proteins and the RNAs they contact, and the resulting RNA-protein complex is purified via an antibody targeting the endogenous protein. However, this relies on the availability of highly specific antibodies for purification and the one step purification might lead to contaminating RNA that is not cross-linked to the protein of interest. To overcome these caveats, new CLIP-based methods were developed such as the cross-linking and analysis of cDNA (CRAC) method, which was first established in yeast (Bohnsack et al., 2009; Granneman et al., 2009). Briefly, CRAC uses a bi-partite terminal-tag, which allows a two-step purification of RNA-protein complexes with a specific first elution step utilising a protease cleavage site in between the tags (Figure 5A). After the first elution the RNA is trimmed to generate a footprint of the protein interaction site on the RNA. The remaining His-tag is then used for a second purification under denaturing conditions to remove all non-covalently cross-linked RNAs and non-specifically bound proteins. During this step, the RNA is radioactively-labelled and sequencing adapters are ligated to the 3' and 5' ends. Finally, the RNA-protein complexes are eluted, separated by PAGE and transferred to a membrane. The radioactively-labelled RNA-protein complex is detected by autoradiography, excised from the membrane and the RNA is eluted by digestion of the protein. The cDNA library is generated by reverse transcription of the RNA and amplification by PCR and is then sent for Illumina next-generation sequencing.

Two alternative cross-linking approaches can be employed resulting in two different protocols called UV-CRAC and photoactivatable-ribonucleoside-enhanced (PAR)-CRAC. The first one developed was RNA cross-linking with high energy UV light at 254 nm, which can also result in unspecific protein-DNA cross-links. To address this issue, PAR-CRAC uses the uridine analogue 4-Thiouridine. This is supplemented to the growth medium several hours before cross-linking to allow incorporation into nascent RNA. Afterwards, lower energy UV light at 365 nm is used, which specifically introduces
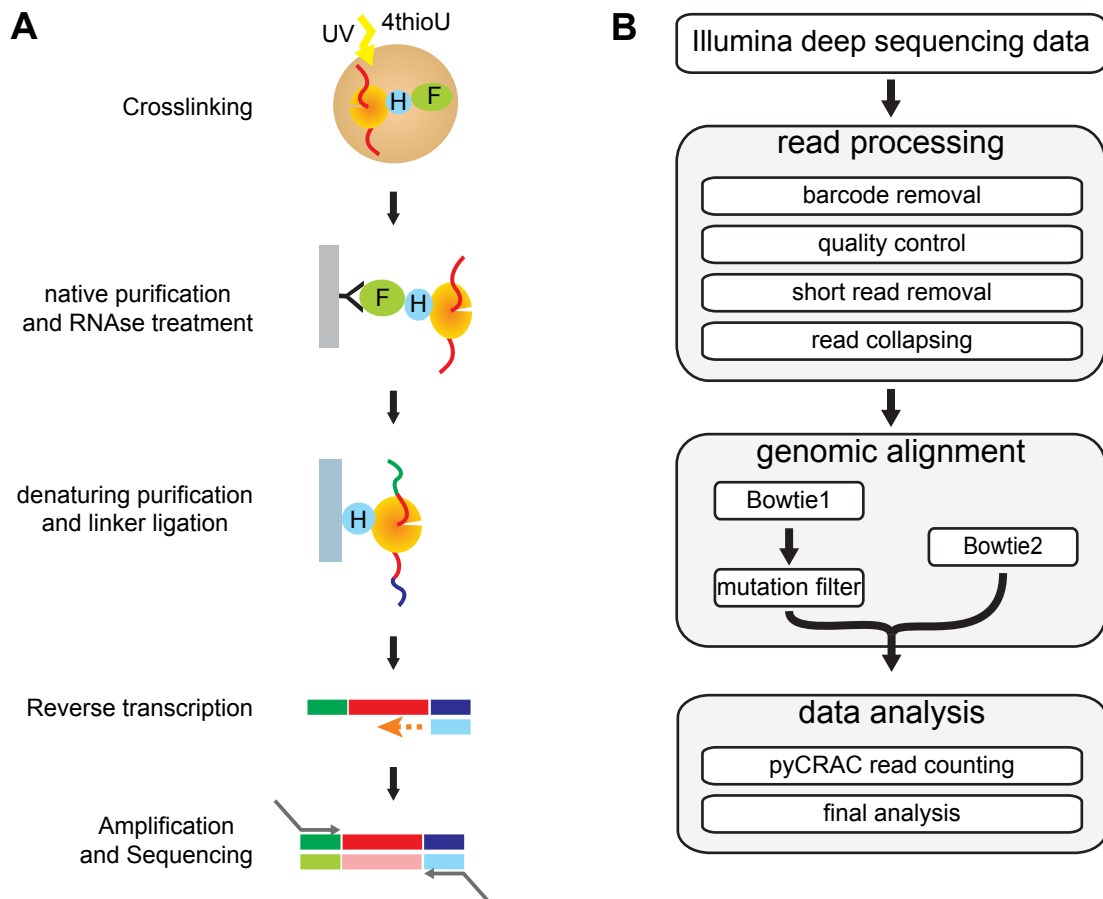
Figure 5: Cross-linking and analysis of cDNA (CRAC). A Schematic overview of the CRAC protocol. Cells expressing the protein of interest fused to a FLAG (His$_6$-PreScission protease site-FLAG$_2$)-tag are grown in the presence of 4-thiouridine (4thioU) and then cross-linked using light at 365 nm (PAR-CRAC) or are directly cross-linked using UV light at 254 nm (UV-CRAC). After a first native purification step of the cross-linked RNA-protein complexes using αFlag antibodies attached to magnetic beads, the complex is eluted and the RNA is trimmed to the binding site of the protein. A second denaturing purification is then carried out and "on-bead" linker ligation for library preparation is performed. The protein is digested and purified RNA is reverse transcribed, amplified by PCR and sequenced by Illumina next-generation sequencing. Abbreviations: 4thioU, 4-thiouridine; F, Flag$_2$-tag; H, His$_6$-tag; PAR, photoactivatable-ribonucleoside-enhanced. Modified from Haag et al., *accepted manuscript*. **B** Schematic overview of the bioinformatic pipeline used for CRAC data analysis. The sequence reads from Illumina next-generation sequencing are first processed, ensuring that only high quality reads are mapped. Barcodes used for the detection of PCR artefacts over representing certain sequences are removed and the sequences are collapsed afterwards. The reads are aligned to a modified version of the human genome and, in the case of PAR-CRAC, only reads with specific T to C mutations are further analysed. Sense-aligned reads for each gene are summed and the numbers are normalised to the total number of sense-aligned reads. The final analysis includes an overview of the hit distribution on different types of RNA and prepares the data for visualisation in common genome viewers using self-written python scripts.

covalent bonds between the thio-group and amino acid side-chains, thereby reducing background compared to the UV$_{254}$ cross-linking approach (Hafner et al., 2011).

Analysis of the sequencing data can only be achieved using bioinformatical methods, because it consists of millions of reads that need to be quality controlled and processed. Therefore, an in-house pipeline was developed for the analysis of human UV-CRAC and PAR-CRAC data, based on the pyCRAC suite developed for yeast (Webb et al., 2014). In general, the pipeline is divided into three major parts: read preparation for the alignment, genome-wide read alignment and analysis of the alignment (Figure 5B).

First, artificial sequences that were added during cDNA library preparation need to be removed. The library preparation includes a PCR step, for which sequences complementary to the 5' and 3' primers are ligated to the RNA. If short RNA sequences are amplified, this can lead to sequencing of the 3' adapter, which then needs to be removed before mapping to the genome. Also, a random barcode is added to detect overrepresented sequences that were artificially amplified during the PCR step and this also needs to be removed. Second, the quality of the sequence read needs to be checked before mapping. Illumina sequencing is based on monitoring the incorporation of fluorescently labelled nucleotides during DNA synthesis and although this process has high accuracy, sometimes bases cannot be identified. Therefore, it is necessary for all reads to undergo quality control before they can be mapped. The program Flexbar was chosen for this step in the pipeline because it is fast and versatile, and it can remove trailing adapter sequences and low quality reads in one step (Dodt et al., 2012). Furthermore, Flexbar was set to discard sequence reads that are shorter than 21 nt, because short reads increase the possibility of false mapping (i.e. short reads can be mapped to multiple places in the genome meaning that their origin cannot be unambiguously determined). In the same part of the pipeline, the random barcode is removed and reads containing the identical barcodes together with identical sequences are consolidated by the "pyBarcodeRemover" and the "pyFastDuplicateRemover" scripts from the pyCRAC suite (Webb et al., 2014), so that only unique reads are left.

The remaining reads can then be aligned to the human genome. For this, an altered version of the human genome GRCh 37.75 was generated as a template for the alignment. The human ribosomal DNA complete repeating unit (NCBI accession number U13369.1) and the 5S rDNA sequence were added as additional chromosomes containing the rDNA sequences of the mature rRNAs published in the 3D structure of the ribosome (PDB 4V6X; Anger, 2013). Additional copies of the 5S rDNA sequence were also masked by 'N' in the genome to avoid mapping of sequence reads to multiple locations within the genome. Also, a modified version of the ensembl annotation of the

genome was generated as the published version contains several incorrectly categorised genes (e.g. long non-coding RNAs as mRNAs and snoRNAs as snRNAs). Such genes were manually assigned to the correct category and the genome annotation was simplified by disregarding the 5' and 3' UTR annotation of protein coding genes.

Various alignment programs can be used for the alignment of the reads to the genome, e.g. Bowtie, Bowtie2, novoalign, etc. and different programmes were used depending on the cross-linking method. UV-CRAC often induces mutations and microdeletions during reverse transcription at the nucleotides that have been cross-linked to amino acids, so Bowtie2 was used because it allows mutations as well as gaps in the read alignments (Langmead and Salzberg, 2012). Bowtie2 is a versatile program, since penalties can be individually set for mutations, the number of gaps in the alignment and the length of the gaps. The threshold score, which determines whether an alignment is considered valid or not, can also be adjusted as a function of the read length. The standard pipeline settings were chosen to allow one mismatch or one gap for short reads below 32 nucleotides, however, only the best alignment is reported in the output file. In contrast to UV-CRAC, due to the incorporation of 4-thiouridine into cellular RNAs, PAR-CRAC introduces specific T to C mutations at the cross-linking site, but no deletions. In this case, the Bowtie alignment programme was used, because it has more stringed mapping criteria, independent of the read length (Langmead et al., 2009). It is an ultrafast short read aligner, which can only detect full alignments containing mutations but no gaps. The maximal number of allowed mutations was set to one, allowing only the mutation introduced by the cross-link. Bowtie, as Bowtie2, only reports the single best alignment and cannot distinguish between different mutations, thus, python scripts were written to select for alignments containing T to C mutations for further analysis. The next step in the pipeline is to check whether alignments overlap with annotated regions of the genome. For this task, the "pyReadCounters" script of the pyCRAC suite was integrated into the pipeline. It counts the number of alignments for each annotated genomic feature, which are summarised in eleven different RNA categories (tRNA, mRNA, rRNA, micro (mi)RNA precursors, mitochondrial (mito)RNA, long non-coding (lnc)RNA, small nucleolar (sno)RNA, small nuclear (sn)RNA, pseudogenes and miscellaneous (misc)RNA). The data mapped to tRNA genes are then further processed by self-made python scripts so that reads aligned with genes encoding for the same type of tRNA, are summed up for simplification. Finally, results are reordered and formatted as excel tables using self-written scripts and the data are normalised to the total number of mapped sense-reads to enable comparison of different samples and sequencing runs.

3.1.1 Verification of the pipeline

The pipeline was initially tested with data from UV cross-linking of the helicase DDX21 that had previously been studied (Sloan et al., 2015). The pipeline was then applied to NSUN6 and NSUN3 CRAC data generated by Dr. Sara Haag. When cross-linking these RNA methyltransferases to generate data for these analyses, HEK293 cell lines expressing genomically integrated version of NSUN6 or NSUN3 with a His_6-PreScission protease site-$Flag_2$ (FLAG)-tag from a tetracycline promoter had been subjected to UV-CRAC. Mapping of the obtained Illumina sequencing data using the developed bioinformatic pipeline showed an over-representation of tRNAs for the NSUN6 sample compared to cells only expressing the FLAG-tag (FLAG, Figure 6A). Closer analysis of the reads aligning to tRNA genes showed that tRNAs for threonine, cysteine and arginine codons were highly enriched compared to the FLAG control (Figure 6B). Having identified $tRNA^{Thr}$ and $tRNA^{Cys}$ as putative RNA substrates of NSUN6, the results of the bioinformatic analysis of the CRAC data were then confirmed by analysing RNAs cross-linked to NSUN6 using Northern blotting (performed by Dr. Haag) and this confirmed that $tRNA^{Thr}$ and $tRNA^{Cys}$ are indeed bound by NSUN6 *in vivo*, thereby verifying the output of the new pipeline (Haag et al., 2015b).

In contrast, analysis of the NSUN3 CRAC data revealed an over-representation of reads aligning to the mitochondrial genome, especially mitochondrial tRNAs (Figure 6C), which is consistent with the mitochondrial localisation of the protein (Haag et al., 2016). Closer analysis of the data showed that among the mitochondrial tRNAs, the mitochondrial $tRNA^{Met}$ (mt-TM, Figure 6D) was highly enriched, suggesting mt-$tRNA^{Met}$ as a target of NSUN3. As for NSUN6, further *in vitro* and *in vivo* experiments (performed by Dr. Haag) confirmed the interaction of NSUN3 with mt-$tRNA^{Met}$ and in this case, showed that NSUN3 specifically methylates C_{34} at the wobble position (Haag et al., 2016).

The finding that the RNA substrates of two proteins that were identified using the human CRAC analysis pipeline could be confirmed by biochemical experiments *in vivo* and *in vitro* shows that the different modules of the pipeline are working correctly together and the parameters for alignment and quality control are chosen properly. This therefore validates the pipeline as a tool for the mapping of next generation sequencing data obtained during CRAC experiments onto the human genome to identify RNA substrates of proteins of interest.
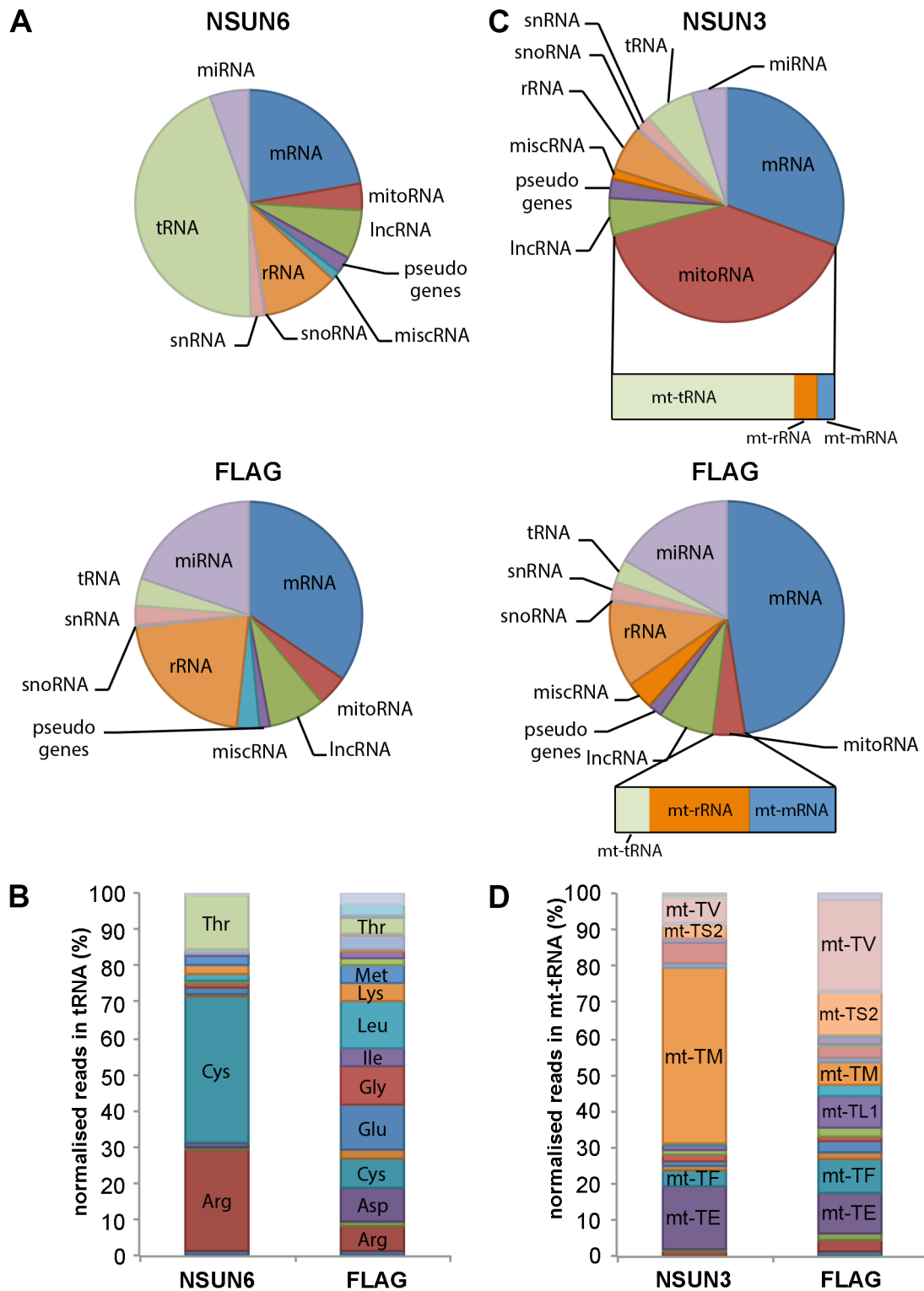
Figure 6: Genome-wide mapping of NSUN3 and NSUN6 CRAC data. A HEK293 cells expressing NSUN6-FLAG or the FLAG-tag alone were UV cross-linked. The RNA-protein complexes were affinity purified and the RNA was trimmed, radioactively labelled and ligated to linkers. RNA-protein complexes were separated by NuPAGE, transferred to a nitrocellulose membrane and the RNA was isolated from the bound RNA-protein complexes by protelytic digest with Proteinase K. The RNA was converted to cDNA for sequencing library production and Illumina next-generation sequencing. The pie charts represent different

(legend continued on next page)

RNA classes and the relative distribution of sequencing reads that was obtained after mapping of the reads to the human genome. Abbreviations: mRNA, messenger RNA; tRNA, transfer RNA; snRNA, small nuclear RNA; snoRNA, small nucleolar RNA; rRNA, ribosomal RNA; mitoRNA, mitochondrial-encoded RNA; miscRNA, miscellaneous RNA; miRNA, microRNA; lncRNA, long non-coding RNA. **B** Relative distribution of tRNA sequences obtained from NSUN6-FLAG or FLAG CRAC experiments described in **A**. Only tRNA genes above 3 % are labelled. **C** HEK293 cells expressing NSUN3-FLAG or FLAG-tag alone were treated and analysed as described in **A**. The bar graphs below the pie charts represent the read distribution of mitochondrial (mt)-tRNA, mt-rRNA and mt-mRNA sequence reads among reads mapped to the mitochondrial genome. **D** Relative distribution of mt-tRNA sequences from NSUN3-FLAG or FLAG CRAC experiments described in **C**. Only mt-tRNA genes above 3 % are labelled. The experiments were performed by Dr. Sara Haag (Haag et al., 2016; Haag et al., 2015b).

3.2 Identification of RNA interactions of the YTH domain-containing proteins

Having established a bioinformatic pipeline for the mapping of CRAC data onto the human genome and verified that this could be used to identify RNA substrates of RNA modifying enzymes, we next applied this approach to identify the RNA interaction partners of other RNA-binding domain-containing proteins. A family of proteins that has raised great interest recently are the YTH domain-containing proteins. The YTH domain was suggested to be an RNA binding domain that can specifically recognise the N^6-methyladenosine (m^6A) modification. Five YTH domain-containing proteins (YTHDF1, YTHDF2, YTHDF3, YTHDC1, YTHDC2) have been identified in humans so far. Interestingly, PAR-CLIP analysis of YTHDF2 enabled detection of YTHDF2-associated mRNAs leading to the identification of a role for this protein in regulating the decay of specific m^6A-containing mRNAs (Wang et al., 2014a). This shows that the identification of interaction targets is crucial and therefore, CRAC was applied to identify RNA interaction partners, which can provide insights into the function of the other YTH domain-containing proteins. To perform the CRAC analysis, HEK293 stable cell lines where the individual YTH domain-containing proteins could be inducibly expressed with a C-terminal FLAG-tag first had to be generated and protein expression verified.

For this, the coding sequences of the YTH domain-containing proteins, YTHDF1, YTHDF2, YTHDF3, YTHDC1 and YTHDC2 were cloned into a pcDNA5 vector encoding a C-terminal FLAG-tag. HEK293 T-Rex Flp-In cells were transfected with the constructs to generate stable cell lines. The cloned coding sequence is genomically integrated at a defined locus in the genome to ensure that no random alterations of the genome can occur and transfected cells are selected using antibiotics. To check proper integration of the coding sequence of interest, the generated stable cell lines were grown in the
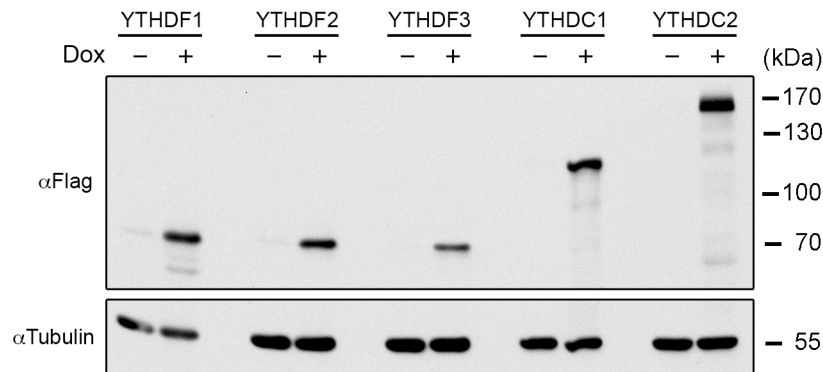
Figure 7: Generation of stable HEK293 cell lines expressing YTH domain-containing proteins. Whole cell lysate from HEK293 cell lines expressing FLAG-tag versions of the indicated proteins was prepared, separated by SDS-PAGE and subjected to western blotting. The cells were induced (+) for 24 h with doxycycline (Dox) or untreated (-). The FLAG-tagged proteins were detected using an αFlag antibody. Tubulin was used as a loading control and detected with an αTubulin antibody.

presence (induced) or absence (non-induced) of doxycycline for 24 h. Whole cell lysates were prepared and separated by SDS-PAGE followed by western blotting (Figure 7). FLAG-tagged proteins were detected using an αFlag antibody. For each of the proteins, a signal at the expected size was detected for all samples grown in the presence of doxycycline, confirming expression of tagged forms of the correct proteins. No signal could be detected for non-induced cells, showing that the expression was inducible. Tubulin was used as a loading control to display equal loading of the associated samples.

After confirming correct expression of the YTH domain-containing proteins, the cell lines were further analysed by immunofluorescence microscopy (IF) to ensure correct localisation of the proteins, if known, as well as normal appearance of the cells (Figure 8). Protein expression was induced for 24 h before fixation of the cells for IF. An αFlag antibody was used for the detection of the FLAG-tagged version of the protein and nuclear material was visualised by DAPI staining. This shows an exclusively cytoplasmic localisation for YTHDF2-FLAG and nuclear localisation for YTHDC1-FLAG, which is in line with the localisations reported for the endogenous proteins and with their proposed function in mRNA decay and pre-mRNA splicing respectively (Hartmann et al., 1999; Wang et al., 2014a). In keeping with their close homology to YTHDF2, YTHDF1-FLAG and YTHDF3-FLAG display cytoplasmic localisation similar to YTHDF2-FLAG. Interestingly, YTHDC2-FLAG shows localisation to two compartments, a dominant cytoplasmic signal and minor nuclear signal.
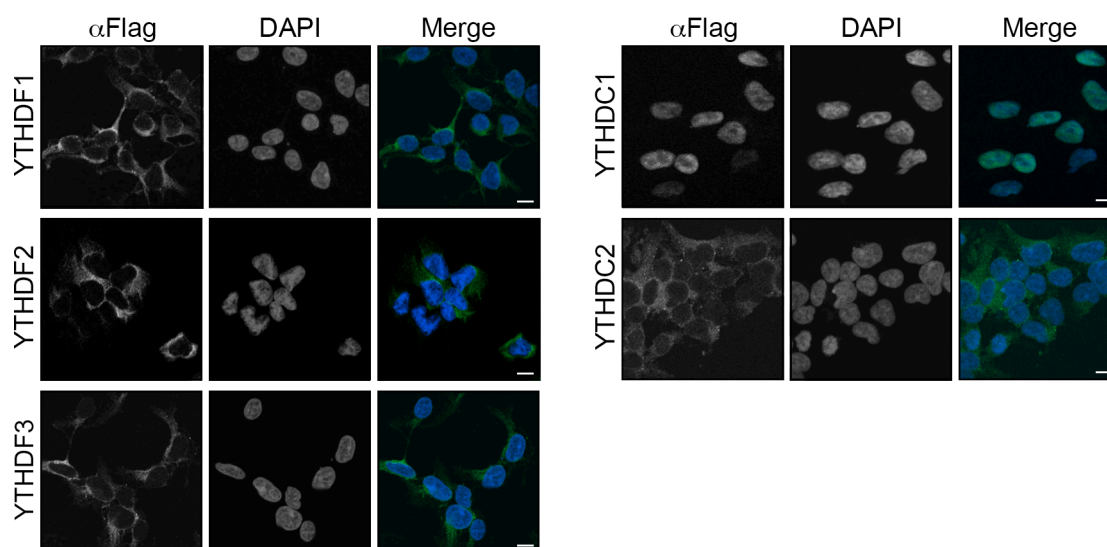
Figure 8: Localisation of YTH domain-containing proteins. Stable cell lines expressing FLAG-tagged versions of the indicated proteins were subjected to immunofluorescence microscopy after 24 h of induction of protein expression. An αFlag antibody was used to detect the FLAG-tagged proteins. DAPI staining indicates the position of the cell nucleus. The scale bar represents 10 µm.

Having established inducible stable cell lines for the expression of FLAG-tagged versions of all five YTH domain-containing proteins, CRAC experiments were performed with these cell lines to obtain an unbiased, genome-wide overview of RNA interaction partners of the YTH domain-containing proteins. For the PAR-CRAC experiments, YTHDF1-FLAG, YTHDF2-FLAG, YTHDF3-FLAG, YTHDC2-FLAG and YTHDC2-FLAG cell lines were induced for 36 h to express the FLAG-tagged version of the protein. Cells expressing the FLAG-tag alone were used as a negative control. After purification of protein-RNA complexes they were then separated by SDS-PAGE, transferred to a membrane and radioactive signals of the RNA were detected by autoradiography (Figure 9). The autoradiogram revealed signals of radioactively-labelled RNA at the expected size of the RNA-protein complex in all YTH domain-containing protein samples, confirming efficient cross-linking of the tagged proteins to cellular RNAs. As expected, no signal was detected in the negative control representing cells only expressing the FLAG-tag (Figure 9). After excision of the regions of the membrane containing the radioactive signals and cDNA library preparation, the library was send for Illumina next-generation sequencing. The resulting data were then processed with the established in-house human CRAC pipeline.
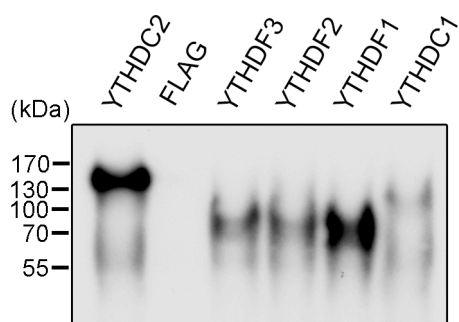
Figure 9: YTH domain-containing proteins cross-link to cellular RNAs. HEK293 cells expressing FLAG-tagged YTH domain-containing proteins were grown in media containing 4-thiouridine and RNAs were cross-linked using light at 365 nm. After affinity purification and trimming, the 5' ends of the co-purified RNAs were radioactively labelled. RNA-protein complexes were separated by NuPAGE, transferred to a nylon membrane and visualised by autoradiography.

The overall distribution of the alignments to the different RNA categories was analysed first (Figure 10). Compared to the FLAG control, YTHDF1, YTHDF2, YTHDF3 and YTHDC1 show an increased number of mRNA alignments (Figure 10A-D, F). For YTHDF1, the proportion of the total reads that aligned to mRNA genes increased from 50 % to 70 % (Figure 10A, F). Similarly, the fraction of alignments to mRNA genes in the YTHDF2 and YTHDF3 samples rose to 75 % and 73 % respectively (Figure 10B, C). Furthermore, YTHDC1 also showed an 11 % increase to 61 % in mRNA alignments compared to the FLAG control (Figure 10D). An increase in the number of reads mapping to tRNA genes was seen for YTHDF1 (14 %), YTHDF2 (8 %), YTHDF3 (10 %) and YTHDC2 (38 %), compared to the FLAG control (6 %) (Figure 10A-F). However, an increase in the proportion of reads mapping to tRNAs has also previously been seen in other PAR-CRAC samples of RNA-binding proteins (data not shown), suggesting that it can be non-specific. Since this is particular significant for YTHDC2, the distribution of the reads among the tRNA genes was analysed and compared to the control. The distribution did not show any substantial changes, further supporting that the increase is unspecific. Interestingly, YTHDC2 was the only sample that showed an increase in the proportion of reads corresponding to rRNA sequences compared to the FLAG control (Figure 10E-F). To determine whether this subtle enrichment of rRNA sequences in the YTHDC2 PAR-CRAC sample was reproducible, a second CRAC experiment was conducted with the YTHDC2-FLAG cell lines using an alternative cross-linking method (UV-CRAC). Mapping and analysis of the data obtained from this experiment demonstrated an even more prominent enrichment of rRNA sequences in the YTHDC2-FLAG sample compared to the FLAG control (Figure 10G-H). Indeed, reads mapping to the rDNA represented 36 % of total reads compared to only 13 % in the corresponding FLAG control.
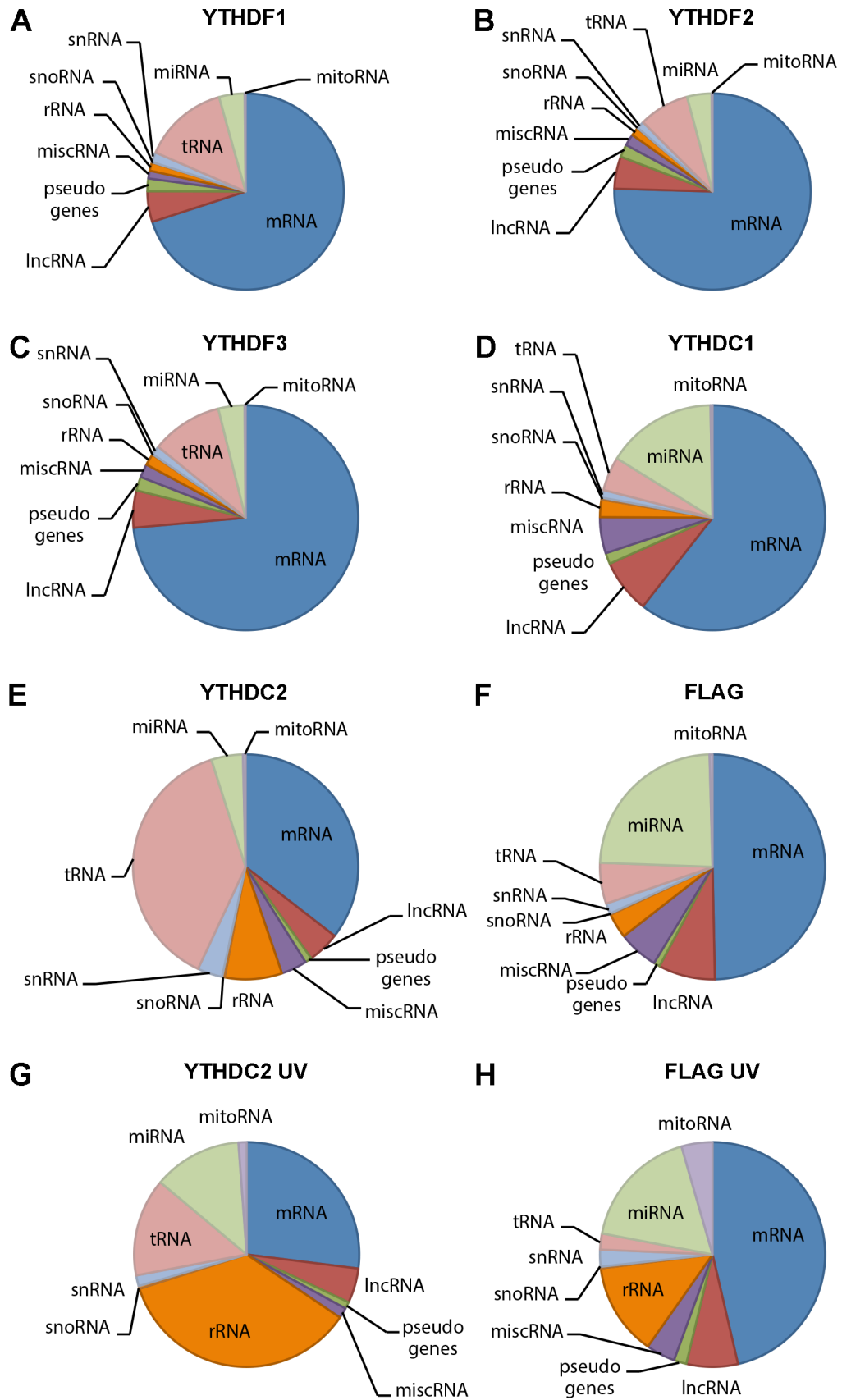
Figure 10: Genome-wide mapping of YTH domain-containing protein CRAC data.

(legend on next page)

A-F HEK293 cells expressing YTHDF1-FLAG (**A**), YTHDF2-FLAG (**B**), YTHDF3-FLAG (**C**), YTHDC1-FLAG (**D**), YTHDC2-FLAG (**E**) or the FLAG-tag alone (**F**) were grown in media containing 4-thiouridine and RNAs were cross-linked using light at 365 nm (see also Figure 9). The RNA-protein complexes were affinity purified, the co-purified RNA was trimmed, the 5' ends of were radioactively labelled and adapters were ligated to both ends of the RNA. RNA-protein complexes were separated by NuPAGE, transferred to a nylon membrane and visualised by autoradiography. The RNA was eluted from the RNA-protein complexes on the nylon membrane and transcribed into cDNA for sequencing library preparation and Illumina next-generation sequencing. The pie charts represent different RNA classes and the relative distribution of sequencing reads that was obtained after mapping of the reads to the human genome. Abbreviations: mRNA, messenger RNA; tRNA, transfer RNA; snRNA, small nuclear RNA; snoRNA, small nucleolar RNA; rRNA, ribosomal RNA; mitoRNA, mitochondrial-encoded RNA; miscRNA, miscellaneous RNA; miRNA, microRNA; lncRNA, long non-coding RNA. **G-H** HEK293 cells expressing YTHDC2-FLAG (**G**) or the FLAG-tag alone (**H**) were UV cross-linked at 254 nm and treated otherwise as in **A-F**.

Taken together, the CRAC data show an overrepresentation of reads mapping to protein-coding genes (mRNAs) for YTHDF1, YTHDF2, YTHDF3 and YTHDC1, which is in line with the earlier publications of YTHDF2 and YTHDC1 (Hartmann et al., 1999; Wang et al., 2014a). During this study, two additional reports were published using the CLIP method to identify genome-wide RNA substrates of YTHDF1 and YTHDC1 (Wang et al., 2015; Xiao et al., 2016). There, YTHDF1 was identified as an mRNA binding translation enhancer and it was confirmed that YTHDC1 is involved in pre-mRNA splicing, consequently supporting the elevated mRNA levels found in the PAR-CRAC experiments seen in Figure 10. Similar to YTHDF1 and YTHDF2, YTHDF3 is also likely to recognise m^6As in mRNAs due to the increased proportion of reads mapping to mRNAs and the cytoplasmic localisation of the protein. This fact makes the rRNA increase of YTHDC2 even more interesting because it is the only YTH domain-containing protein with an elevated level of reads mapping to rRNA, suggesting a specific interaction with ribosomal RNA.

Since YTHDC2 showed an increase in the proportion of reads mapping to the rDNA, the next step was to check for the accumulation of reads at specific sites on the rDNA. A multiple sequence alignment of the FLAG and YTHDC2-FLAG PAR-CRAC data on the human ribosomal DNA complete repeating unit was done and the accumulation of reads was mapped at single nucleotide level within the region of the cDNA repeat from which the 47S pre-rRNA is described (Figure 11). High accumulation of reads (peaks) corresponds to potential protein cross-linking sites on the RNA. YTHDC2 showed distinct and specific peaks in regions coding for the mature rRNAs on the repeat, but not in the spacer regions (Figure 11A). The main cross-linking site of YTHDC2 is located at the 3' end in the mature 18S rRNA (Figure 11B). In accordance, an increased mutation rate
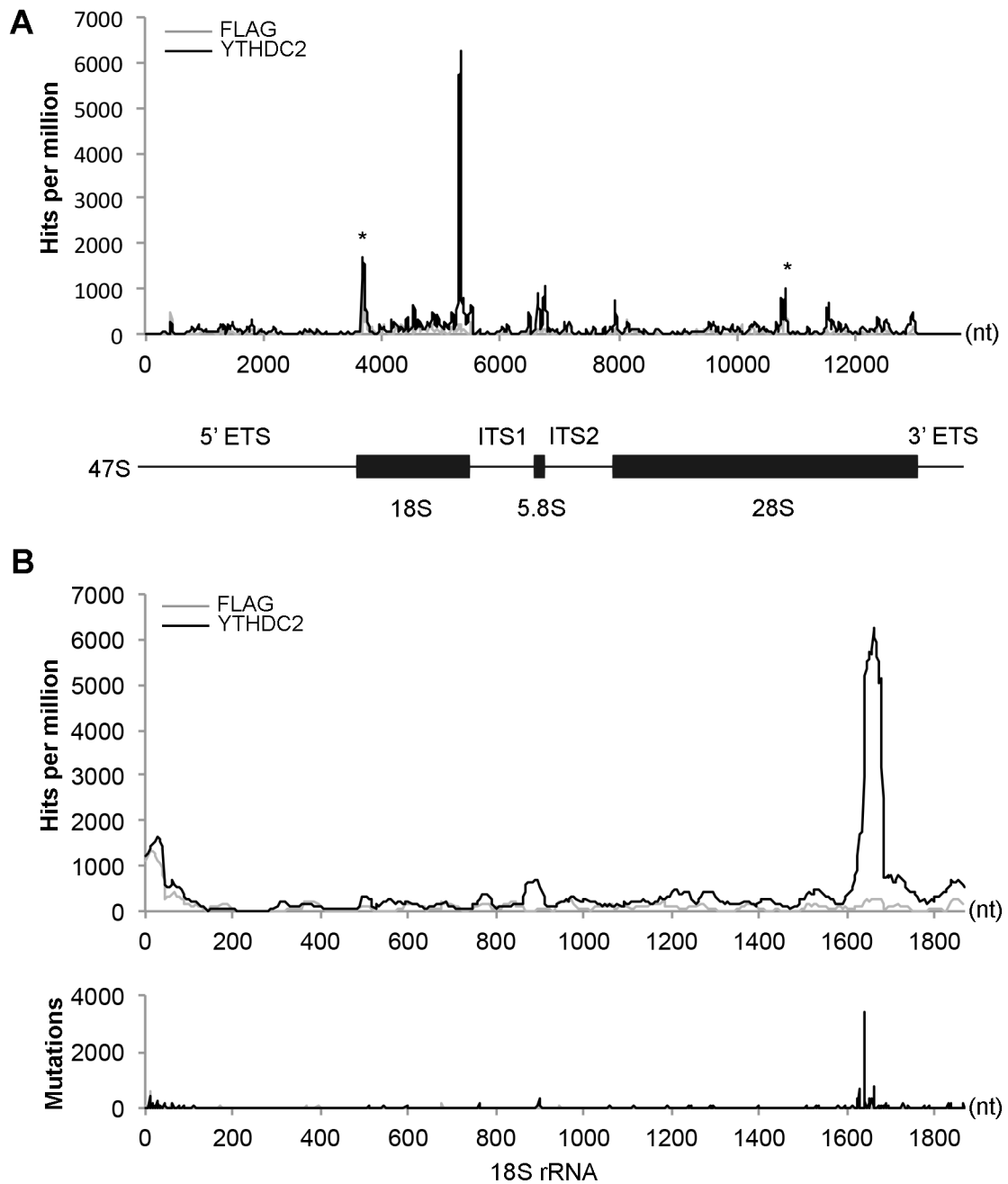
Figure 11: Profile of YTHDC2 PAR-CRAC hits on pre-rRNA. A Sequence read distribution of YTHDC2-FLAG (YTHDC2, black plot) or FLAG-tag (FLAG, grey plot) PAR-CRAC samples mapped to the rDNA gene encoding the 47S pre-rRNA. The accumulation of sequence reads on the y-axis is normalised to the total number of mapped sense-reads (see also Figures 9 and 10E-F). The asterisks mark background peaks also present in samples derived from cells expressing the FLAG-tag only. A representation of the 47S pre-rRNA is shown below indicating the positions of the mature rRNAs as rectangles and externally and internally transcribed spacers as lines on the x-axis. Abbreviations. ETS, external transcribed spacer; ITS, internal transcribed spacer; nt, nucleotides. **B** Magnified view of the 18S rDNA sequence depicted in **A**. The x-axis represents the nucleotide numbering of 18S rDNA. The y-axis of the upper diagram is the same as in **A**. The lower diagram indicates the position of T to C mismatches between sequence reads and the genomic sequence for YTHDC2-FLAG (YTHDC2, black plot) or FLAG-tag alone (FLAG, grey plot).

of distinct thymine residues was detected within the cross-linking site, confirming the specificity of the peak (Figure 11B lower panel). Interestingly, the YTHDC2 cross-linking site is localised close to the 3'-end of the 18S rRNA similar to the location of the m^6A modification. Cross-linking of YTHDC2 close or at this modification site would be in line with the putative function of the YTH domain recognising the m^6A modification, however, the cross-linking site of YTHDC2 and the m^6A modification are separated by about 150 nt in the rRNA sequence. To analyse whether the modification is close to the cross-linking site in 2D or 3D, the YTHDC2-FLAG CRAC data were mapped onto the 2D structure of the 18S rRNA (Petrov et al., 2014) and the 3D structure of the SSU of the human ribosome (PDB 4V6X, Anger et al., 2013). The number of sequence reads mapped per nucleotide is represented by a colour gradient from orange to red, red representing the highest accumulation of sequence reads normalised to 100 % (Figure 12 and 13). Analysis of the 2D mapping revealed that the YTHDC2 cross-linking site covers the stem loop structure of helix 43, which is in the vicinity of the m^6A_{1832} at the 3' end of 18S (Figure 12). Additionally, mapping on the 3D structure confirmed this observation and showed that the cross-linking site is in close proximity to helices 41 and 42, and helices 28 and 29 (Figure 13). Therefore, YTHDC2 cross-links to the "head" region of the SSU and the YTH domain of YTHDC2 could very well recognise the m^6A modification at the end of the 18S rRNA.
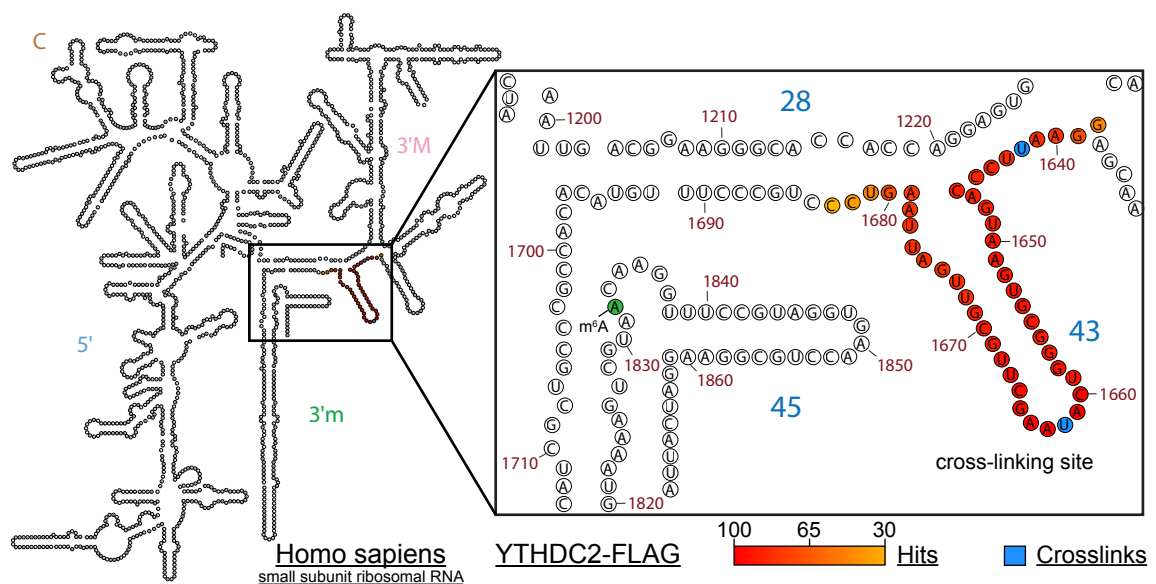
Figure 12: Mapping of YTHDC2 PAR-CRAC data onto the human 18S rRNA secondary structure. The sequence read distribution of YTHDC2-FLAG PAR-CRAC data (see Figure 11) were mapped on the secondary structure of the human 18S rRNA (Petrov et al., 2014). Cross-linking sites are coloured from orange (30 %) to red (100 %) and normalised to the highest peak on the 18S rDNA, the highest peak being 100 %. Residues containing more than 10 % T to C mismatches are coloured in blue. The m^6A-modified residue is coloured in green.
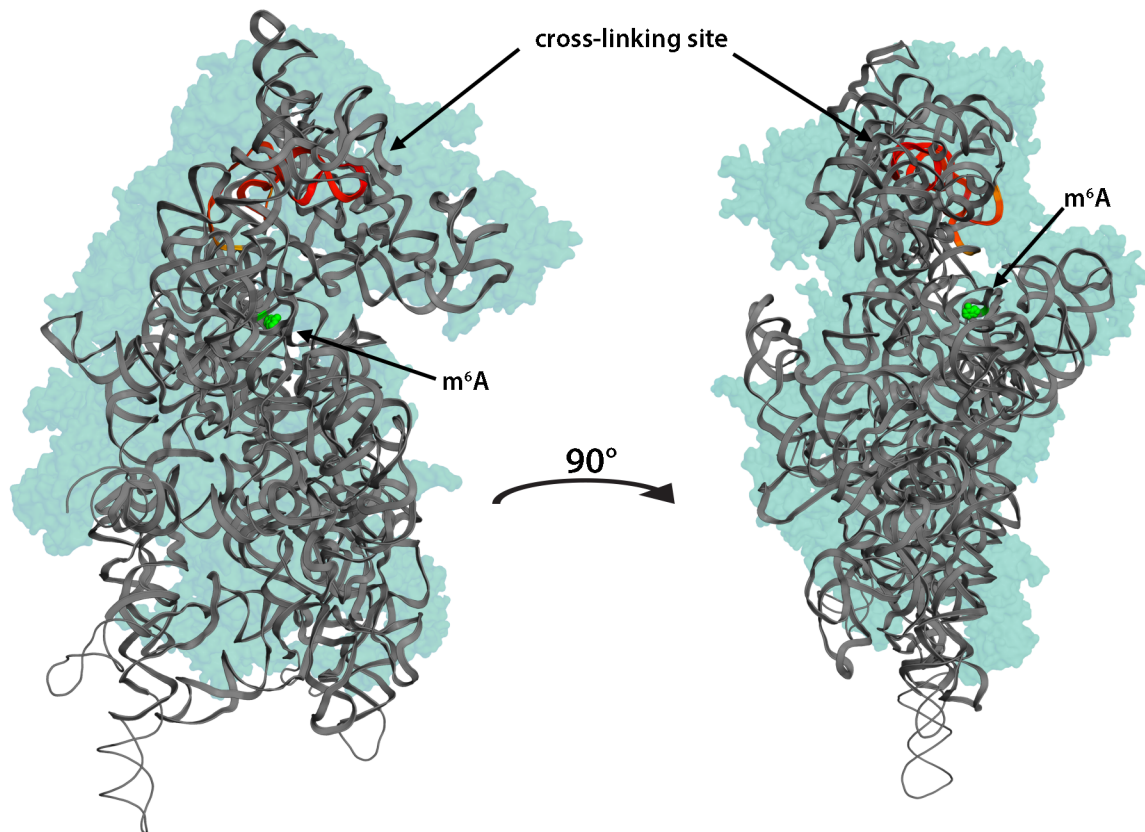
Figure 13: Mapping of YTHDC2 PAR-CRAC data on the 3D structure of the human 18S rRNA. The read distribution of YTHDC2-FLAG (see Figure 11) was mapped on the 3D structure of the human 18S rRNA (PDB 4V6X, Anger et al., 2013). Cross-linking sites are coloured from orange (30 %) to red (100 %) and normalised to the highest peak on the 18S rRNA, the highest peak being 100 %. The A_{1832} residue, which carries an m^6A modification, is highlighted in green.

3.3 YTHDC2 associates with ribosomal complexes

The next step was to verify the association of YTHDC2 with the ribosome suggested by the CRAC analysis with different methods. The large size of ribosomal and pre-ribosomal complexes enables the separation of these complexes from free proteins and small complexes by sucrose gradient density centrifugation. The small (SSU; 40S) and large (LSU; 60S) ribosomal subunits as well as assembled 80S ribosomes have distinct migration patterns within the gradient allowing the differentiation of proteins associating with the different subunits. Furthermore, the availability of an antibody against YTHDC2 made it possible to investigate the association of the endogenous protein with ribosomal complexes in wild type (WT) cells.

Thus, sucrose gradient density centrifugation was performed using HEK293 WT cells and YTHDF2-FLAG cells, in which fusion protein expression had been induced for 36 h before cell harvesting, were used as a control. Whole cell lysates of HEK293 WT and YTHDF2-FLAG cells were prepared and loaded on a 10-45 % sucrose gradient followed

by overnight ultracentrifugation to separate ribosomal and non-ribosomal fractions. After fractionation of the gradient into 23 fractions, the A_{260} of each fraction was measured. This estimate of the RNA content was used to generate a profile and determine which fractions contained the 40S and 60S subunits and 80S monosomes (Figure 14 upper panel). Afterwards, the protein content of each fraction was analysed by SDS-PAGE followed by western blotting (Figure 14 lower panels). The αFlag antibody, which detected the YTHDF2-FLAG, showed a signal exclusively in fractions containing non-ribosomal complexes (Figure 14 bottom western blot). This is in line with the overrepresentation of mRNA alignments seen in the YTHDF2 CRAC, data because mRNPs are enriched in these fractions (see Figure 10B, F). In contrast, the αYTHDC2 antibody shows an extensive signal for YTHDC2 in the fractions containing ribosomal complexes with distinct peaks in fractions 11-13 and fraction 17 (Figure 14 top western blot). Interestingly, these fractions contain the 40S subunit (fractions 11-13) and the 80S ribosome (fraction 17) meaning that all of these fractions contain SSU particles, which is in line with the identification of a cross-linking site for YTHDC2 at the 3' end of 18S rRNA sequence (see Figure 10E-H).
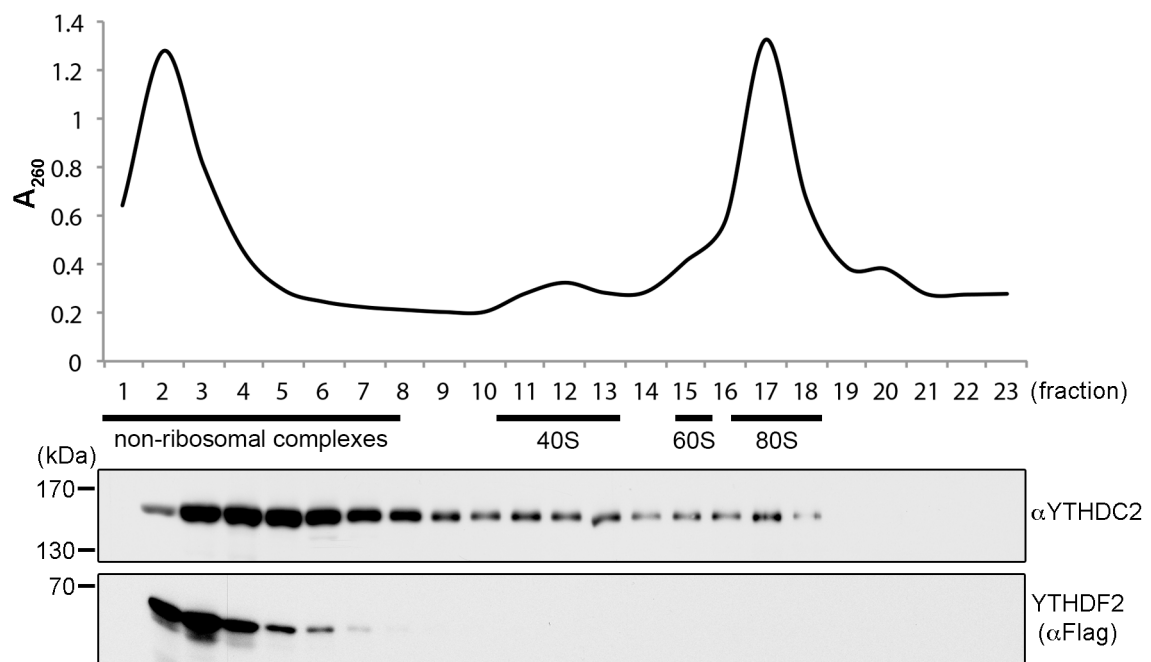
Figure 14: YTHDC2 is associated with ribosomal complexes. Whole cell lysates from wild type HEK293 cells or cells expressing YTHDF2-FLAG were separated by sucrose gradient density centrifugation. Fractions were taken manually and A_{260} measurements of the fractions were used to generate an RNA profile from which the fractions containing different ribosomal complexes could be determined (indicated below the profile). Proteins in the different fractions were TCA precipitated, separated by SDS-PAGE and subjected to western blotting (bottom). Endogenous YTHDC2 was detected using an αYTHDC2 antibody. YTHDF2-FLAG was detected using an αFlag antibody.

Having confirmed the association of YTHDC2 with the ribosome, the next step was to identify the region of the protein that is responsible for this interaction. YTHDC2 is a large protein with a calculated mass of 160 kDa and has a multi-domain structure. *In silico* sequence homology analysis identified three domains, a N-terminal R3H domain, a putative RNA helicase domain and a C-terminal YTH domain (Figure 15A, UniProt, 2015). The R3H domain is a general nucleic acid binding domain, with a preference for binding guanine residues (Jaudzems et al., 2012). RNA helicase domains are generally involved in altering RNA-RNA and RNA-protein interactions (reviewed in Bleichert and Baserga, 2007) and in other proteins, the YTH domain has been shown to recognise m^{6}A modifications in RNAs (Xu et al., 2015). To investigate if either of the two terminal RNA-binding domains is responsible for the interaction of YTHDC2 with the ribosome, constructs for the expression of truncated versions of YTHDC2 were generated. The DNA sequences coding for amino acids 192-1430 (ΔR3H) or amino acids 1-1287 (ΔYTH) of YTHDC2 (Figure 15A) were cloned into the pcDNA5 vector adding the sequence of a C-terminal FLAG-tag to the CDS. The CDSs were genomically integrated into the genome of HEK293 Flp-In T-Rex cells to generate stable cell lines as previously described. To test expression of the truncated proteins, the cell lines were grown in the presence or absence of doxycycline for 24 h, whole cell lysates were prepared and separated by SDS-PAGE and tagged proteins were detected by western blotting using an αFlag antibody. A signal could be detected at the correct size in cells grown in the presence of doxycycline confirming the correct expression of the inserted sequence (Figure 15B). CRAC experiments were therefore conducted with these cell lines after inducing ΔR3H-FLAG or ΔYTH-FLAG expression for 36 h similar to the PAR-CRAC experiments of the YTH domain-containing proteins before and analysed similarly (Figure 15C and D). The stable cell line expressing full-length YTHDC2-FLAG was included as a positive control. Analysis of the distribution of the sequence reads obtained for YTHDC2-FLAG between the different RNA classes showed that the portion of reads mapping to rDNA sequences was similar to the first YTHDC2-FLAG PAR-CRAC experiment. Interestingly, the rRNA portion is reduced to 3.5 % in the ΔR3H sample (Figure 15E). Mapping of this data onto the 18S rRNA sequence of the rDNA repeat confirmed this as, compared to YTHDC2-FLAG sample (black plot), the ΔR3H-FLAG sample (red plot) showed a reduced peak at the end of 18S rRNA (Figure 15C). This suggests that the R3H domain is involved in the interaction of YTHDC2 with the ribosome. In contrast, the ΔYTH-FLAG sample showed a similar fraction of reads mapped to rDNA (8 %) as full-length YTHDC2 (Figure 15E). Also, the profile along the 18S rRNA sequence of
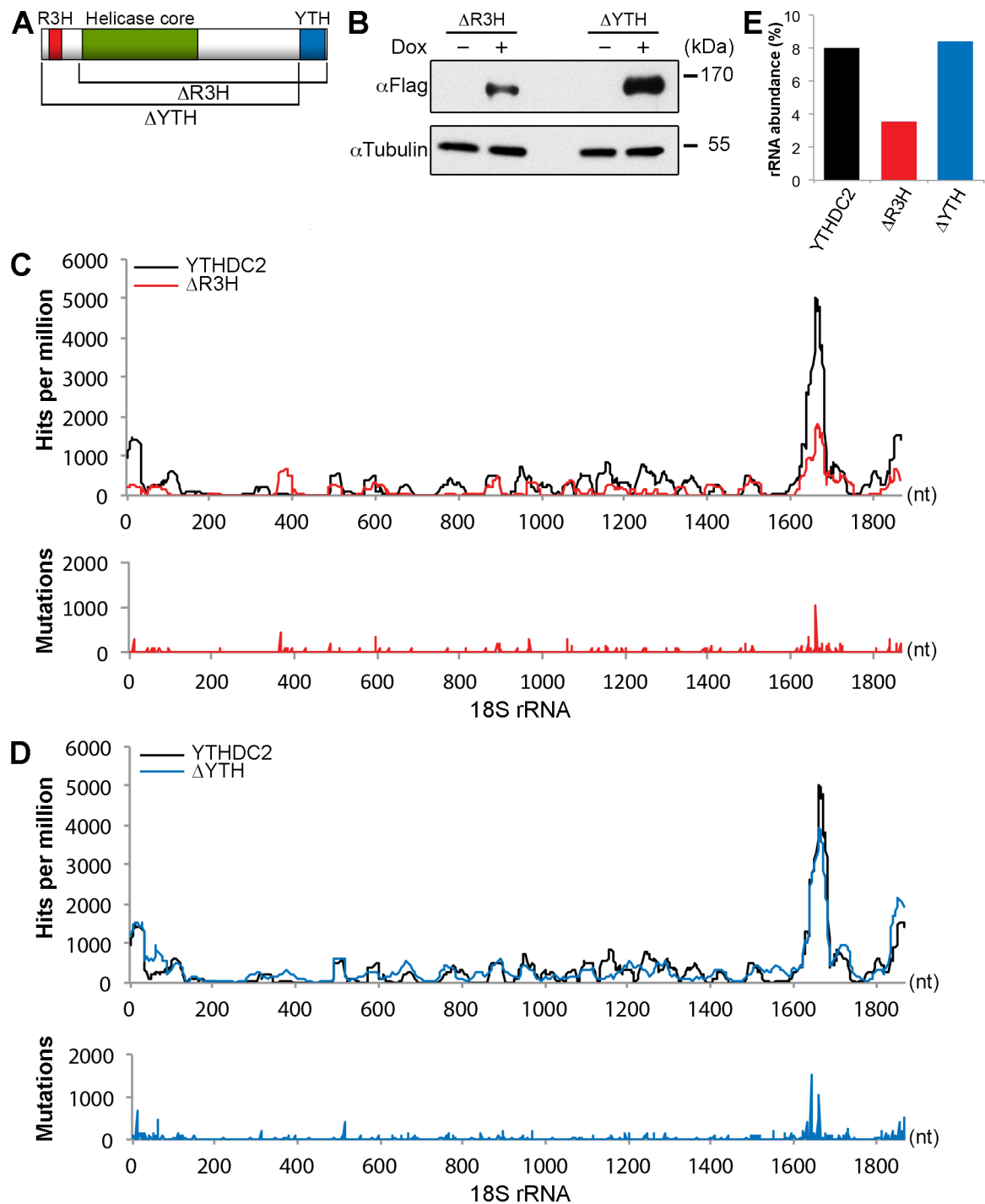
Figure 15: The R3H domain of YTHDC2 contributes to the association with the ribosome. A Schematic representation of the domain structure of YTHDC2. The putative R3H domain (red), the RNA helicase core domain (green) and the YTH domain (blue) are indicated. The bars below represent the boundaries of the truncation constructs. ΔR3H, truncation of the R3H domain (YTHDC2$_{192\text{-}1430}$); ΔYTH, truncation of the YTH domain (YTHDC2$_{1\text{-}1287}$). **B** Whole cell lysate from HEK293 cell lines expressing ΔR3H-FLAG (ΔR3H) or ΔYTH-FLAG (ΔYTH) was prepared, separated by SDS-PAGE and subjected to western blotting. The cells were induced (+) for 24 h with doxycycline (Dox) or untreated (-). The FLAG-tagged proteins were detected using an αFlag antibody. Tubulin was used as a loading control and detected with an αTubulin antibody.

(legend continued on the next page)

D-C HEK293 cells expressing ΔR3H-FLAG (ΔR3H, red plot, **C**), ΔYTH-FLAG (ΔYTH, blue plot, **D**) or full length YTHDC2-FLAG (YTHDC2, black plot, **C** and **D**) were grown in media containing 4-thiouridine and RNAs were cross-linked using light at 365 nm. After trimming, the 5' ends of the co-purified RNAs were radioactively labelled and adapters were ligated to both ends. RNA-protein complexes were separated by NuPAGE, transferred to a nylon membrane and visualised by autoradiography. The RNA was eluted from the RNA-protein complexes on the nylon membrane and transcribed into cDNA for sequencing library preparation and Illumina next-generation sequencing. The plot represents the sequence read distribution on the human 18S rDNA normalised to the total number of mapped sense-reads. The lower diagram indicates the position of T to C changes in sequence reads compared to the genomic sequence for ΔR3H-FLAG (ΔR3H, red plot, **C**) or ΔYTH-FLAG (ΔYTH, blue plot, **D**). Abbreviation: nt, nucleotides. **E** Relative abundance of sequence reads mapped to rDNA for full-length YTHDC2-FLAG (YTHDC2, black), ΔR3H-FLAG (ΔR3H, red) and ΔYTH-FLAG (ΔYTH, blue) PAR-CRAC data shown in **C** and **D**. The data is normalised to the total number of mapped sense-reads.

the ΔYTH sample (blue plot, Figure 15D) is similar to that of the YTHDC2 sample (black plot, Figure 15D) including the peak at the 3' end of the 18S rRNA sequence. The CRAC analyses of the truncated versions of YTHDC2 therefore suggest that the R3H RNA-binding domain facilitates the contact between YTHDC2 and the 18S rRNA, whereas deletion of the YTH domain has no effect on the interaction seen in the CRAC analysis.

3.4 The YTH domain of YTHDC2 recognises the 18S m^{6}A *in vitro*

The interaction of YTHDC2 with the ribosome could be verified *in vivo* by two independent methods, namely CRAC experiments and sucrose density gradients. However, CRAC of the ΔR3H-FLAG and ΔYTH-FLAG cell lines suggests that the R3H domain is primarily responsible for ribosome interaction of YTHDC2 and that the YTH domain does not significantly contribute to the association of this protein with the ribosome. Given that in other proteins, the YTH domain has been shown to specifically bind m^{6}A modifications and that the cross-linking site of YTHDC2 is in close proximity to the 3' end of the 18S rRNA where such a modification has been reported, this raises the possibility that the YTH domain of YTHDC2 "reads" the m^{6}A modification in the 18S rRNA.

To test this possibility, *in vitro* anisotropy experiments were carried out to first test the affinities of the YTH domain of YTHDC2 for N^6-methyadenosine-containing RNA. Therefore, the YTH domain of YTHDC2 (amino acids 1277-1430) was cloned into a pQE80-vector derivative (A21) for expression of this protein domain with an N-terminal His$_{10}$-Zz-TEV protease site (HZZT)-tag. Also, the YTH domains of YTHDF2 (amino acids 380-579) and YTHDC1 (amino acids 344-509) were cloned into the same vector, because they could serve as a positive control, since they have previously been shown
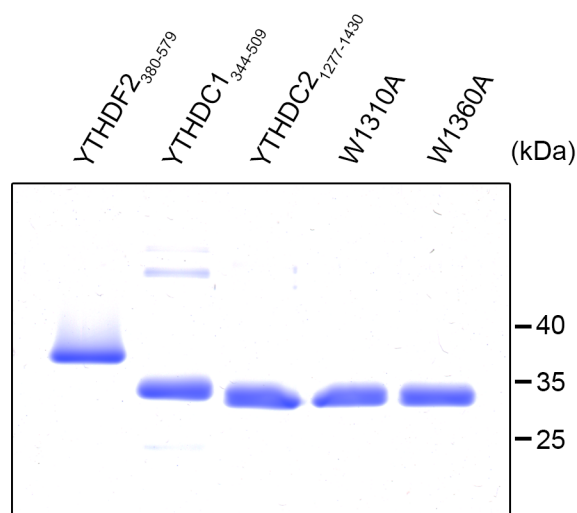
Figure 16: Recombinant expression of YTH domains. Recombinantly expressed, HZZT-tagged fragments of the indicated proteins corresponding to the YTH domains were purified, separated by SDS-PAGE and visualised with Coomassie staining. W1310A and W1360A are versions of YTHDC2$_{1277-1430}$ with tryptophan to alanine substitutions of the amino acids corresponding to positions 1310 or 1360 of the full-length protein.

to specifically bind m^{6}A-containing RNA (Li et al., 2014; Xu et al., 2014). The boundaries were chosen according to the published crystal structures of YTHDC1 and YTHDF2 (Xu et al., 2015; Xu et al., 2014) or homology sequence alignments in the case of YTHDC2. The proteins were expressed in *Escherichia coli* and purification was established (Figure 16).

The majority of the m^{6}A modifications are found in the sequence context GG(m^{6}A)CU (Dominissini et al., 2012), which can be recognised by the YTH domains of YTHDF2 and YTHDC1 (Xu et al., 2015). Anisotropy measurements were carried out to measure binding of the YTH domains to 9 nucleotide long single stranded RNA oligonucleotide labelled at the 5' end with fluorescein containing this sequence motif either with or without m^{6}A modification (Figure 17). Strong binding could be detected for HZZT-YTHDF2$_{380-579}$ (K$_d$ 0.019 ± 0.003 µM) and HZZT-YTHDC1$_{344-509}$ (K$_d$ 0.043 ± 0.004 µM) to the modified GG(m^{6}A)CU. No binding could be detected for HZZT-YTHDF2$_{380-579}$ to the unmodified sequence and only very low binding of HZZT-YTHDC1$_{344-509}$ (K$_d$ 33 ± 8) (Figure 17A, B, D), confirming that the YTH domains of YTHDC1 and YTHDF2 specifically recognise the m^{6}A modification and establishing conditions with which the affinity of the YTH domain of YTHDC2 for the modified and unmodified sequence could be monitored. The K$_d$ of YTHDC2 to the modified sequence was found to be 3.19 ± 0.27 µM while the affinity of the YTH domain of YTHDC2 for the unmodified sequence was so low that a K$_d$ value could not be determined (Figure 17C, D). This demonstrates that similar to the YTH domains of YTHDF2 and YTHDC1, the YTH domain of YTHDC2 also specifically recognises this modification. Notably however, in these experiments, the YTH domain of YTHDC2 displayed a significantly lower binding affinity for the m^{6}A-modified oligonucleotide than the YTH domains of YTHDC1 of YTHDF2.
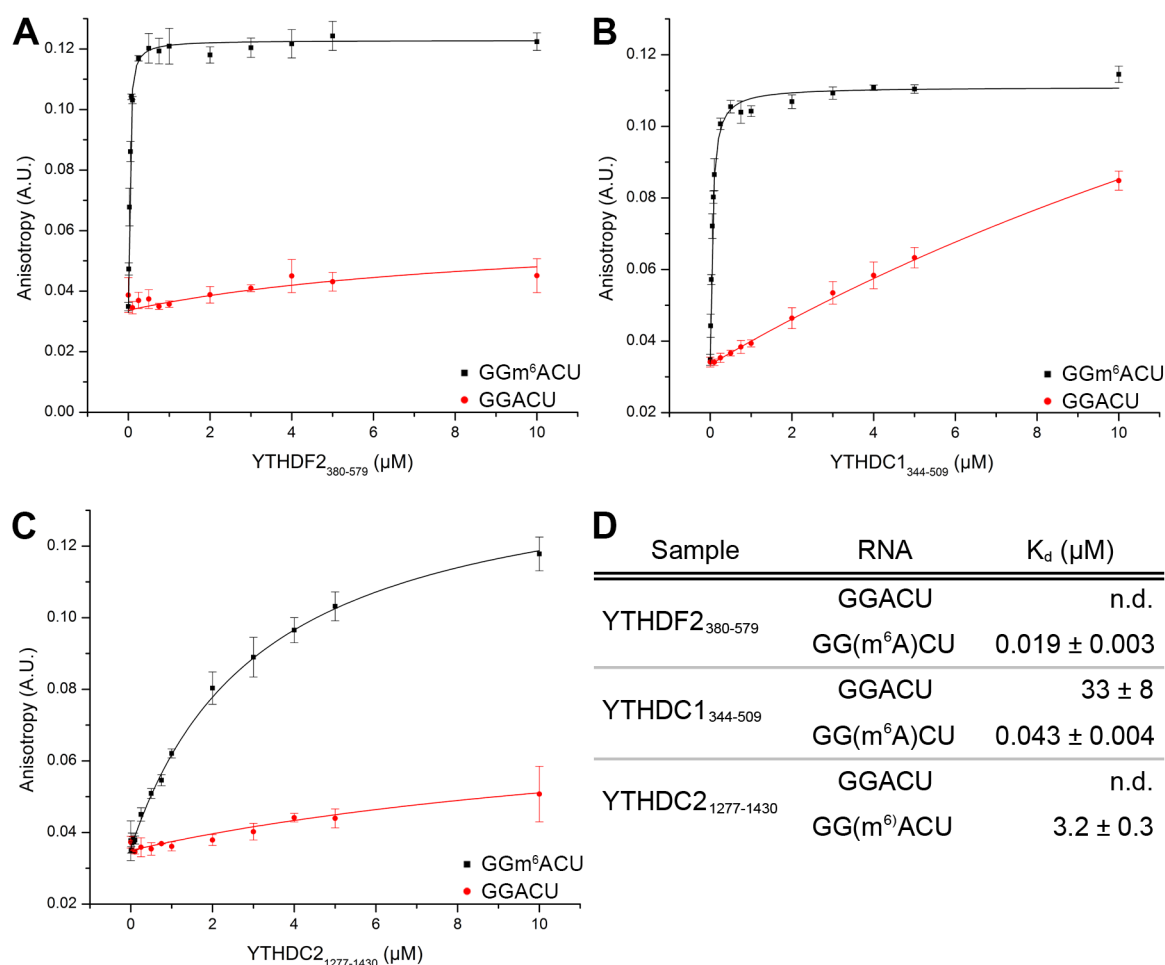
Figure 17: The YTH domain of YTHDC2 recognises the m⁶A modification. A-C A fluorescine-labelled single stranded 9 nucleotide RNA containing the GGACU mRNA sequence in which the "A" was either unmodified (red circles) or carried an m⁶A modification (black squares) was incubated with HZZT-YTHDF2$_{380-579}$ (**A**), HZZT-YTHDC1$_{344-509}$ (**B**) or HZZT-YTHDC2$_{1277-1430}$ (**C**) at the indicated concentrations and fluorescence anisotropy was measured. **D** The dissociation constants (K_d) of the experiments shown in A-C are given. In cases of negligible protein binding to the RNA K_d values could not be determined (n.d.).

The finding that YTHDC2 is associated with the ribosome and cross-links to the 3' end of the 18S rRNA in close proximity to the m⁶A modification present there, suggests that this modification could be recognised by the YTH domain of YTHDC2. Interestingly, the m⁶A modification in the 18S rRNA is present in a different sequence motif UA(m⁶A)CA than the majority of other m⁶A modifications. Therefore, anisotropy measurements were performed to test the relative affinities of the different YTH domains for the m⁶A modification in these two alternative sequence contexts (Figure 18). This showed that, compared to the GG(m⁶A)CU motif, the K_d of HZZT-YTHDF2$_{380-579}$ for the UA(m⁶A)CA motif increased significantly to 0.058 ± 0.003 µM (p<0.0001; Students t test) (Figure 18A, B). Similarly, for the UA(m⁶A)CA motif, the K_d of HZZT-YTHDC1$_{344-509}$ increased to 0.256 ± 0.009 µM (p<0.0001, Students t test) (Figure 18B, D), demonstrating a clear

reduction in binding affinity of these YTH domains to the m^6A modification in the 18S rRNA sequence context, as compared to the classical $GG(m^6A)CU$ motif identified for most m^6A modifications in mRNAs. In contrast, the K_d of HZZT-YTHDC2$_{1277-1430}$ significantly decreased to 2.57 ± 0.18 µM (p=0.0297; Students t test) showing opposite effects when interacting with the modified 18S rRNA sequence (Figure 18C, D). This means that the YTH domain of YTHDC2 is the only protein that shows an increased binding to the m^6A modification in the 18S rRNA sequence context, further supporting a model in which YTHDC2 associates with the ribosome and would recognise this modification. Furthermore, the YTH domain of YTHDC1 strongly selects against this sequence context for m^6A binding suggesting that it is highly specialised for recognition of the m^6A modification in the mRNA/$GG(m^6A)CU$ context.
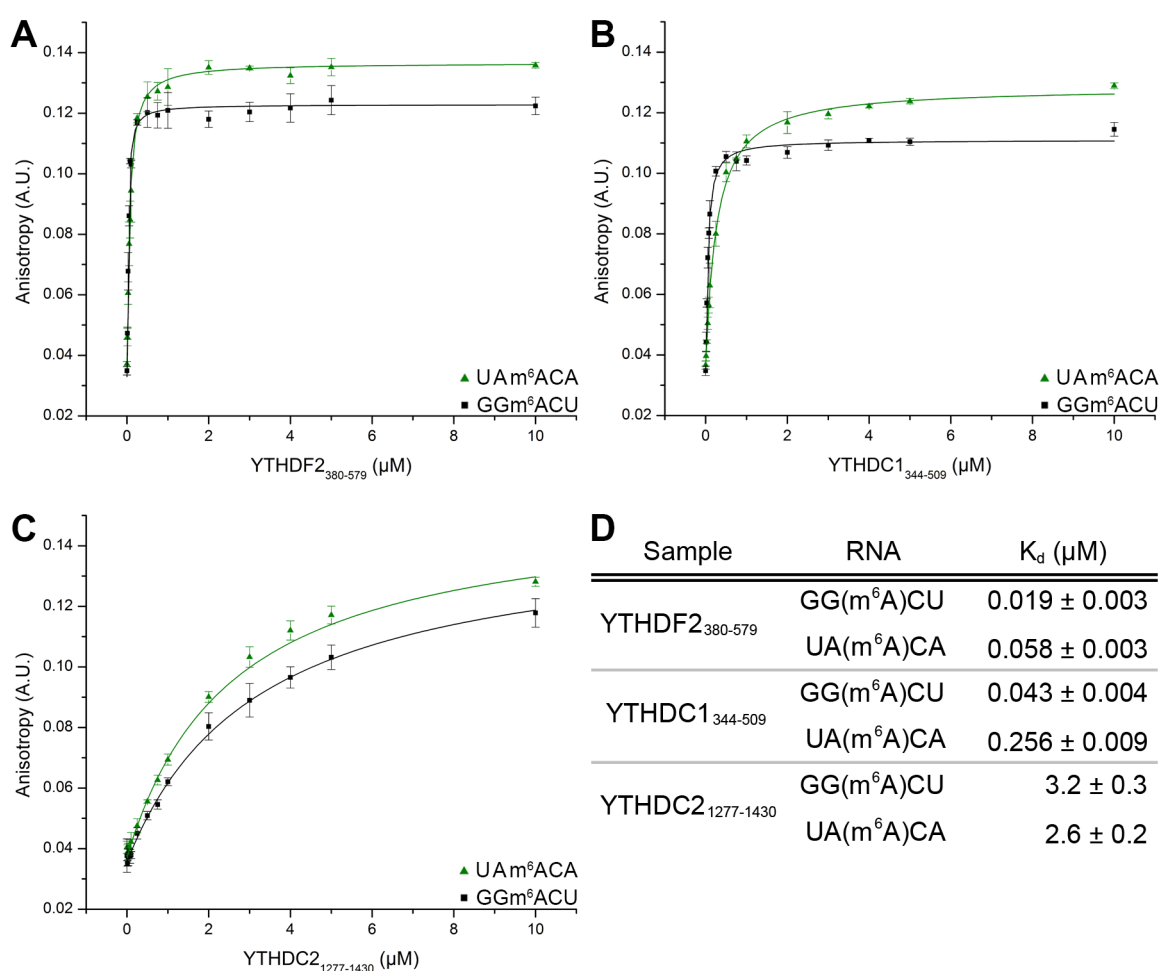
Figure 18: The YTH domain of YTHDC2 recognise the m^6A in the 18S rRNA sequence context with higher affinity than the m^6A in the consensus motif. A-C A fluorescine-labelled single stranded 9 nucleotide RNA containing the modified $UA(m^6A)CA$ 18S rRNA sequence (green triangles) was incubated with HZZT-YTHDF2$_{380-579}$ (**A**), HZZT-YTHDC1$_{344-509}$ (**B**) or HZZT- YTHDC2$_{1277-1430}$ (**C**) at the indicated concentrations and fluorescence anisotropy was measured. For comparison the data of the proteins binding to the modified $GG(m^6A)CU$ mRNA sequence are shown again (see Figure 17, black squares) **D** The dissociation constants (K_d) of the experiments shown in **A-C** are given.

To identify key residues in the YTH domain of YTHDC2 that contribute to m^{6}A recognition, the amino acid sequences of these three YTH domains were aligned and analysed with the multiple sequence aligner MUSCLE (Figure 19, Edgar, 2004). This revealed that the YTH domains of YTHDC2 and YTHDC1 are 49.25 % identical, whereas the YTH domain of YTHDF2 shares only 32.84 % sequence identity. Interestingly, the tryptophan residues W_{432}, W_{486} and W_{499} in YTHDF2 and W_{377}, W_{428} and W_{447} in YTHDC1, which were shown to form the aromatic cage and mediate the m^{6}A interaction (Xu et al., 2014; Zhu et al., 2014) are also conserved in YTHDC2 (W_{1310}, W_{1360} and W_{1373}; arrows, Figure 19). To test whether these residues also contribute to the binding of the YTH domain of YTHDC2 to the m^{6}A in the 18S rRNA sequence context, two of the tryptophan residues, W_{1310} and W_{1360}, were mutated to alanine (Figure 19, red arrows). Proteins containing the mutations were expressed and purified (Figure 16) and additional anisotropy experiments were performed to monitor association with the RNA oligonucleotide containing the m^{6}A modification in the 18S rRNA sequence motif. Interestingly, both mutations abolished the binding to the m^{6}A modification implying that the YTH domain of YTHDC2 interacts with the m^{6}A modification in a similar manner as the other YTH domain-containing proteins (Figure 20).
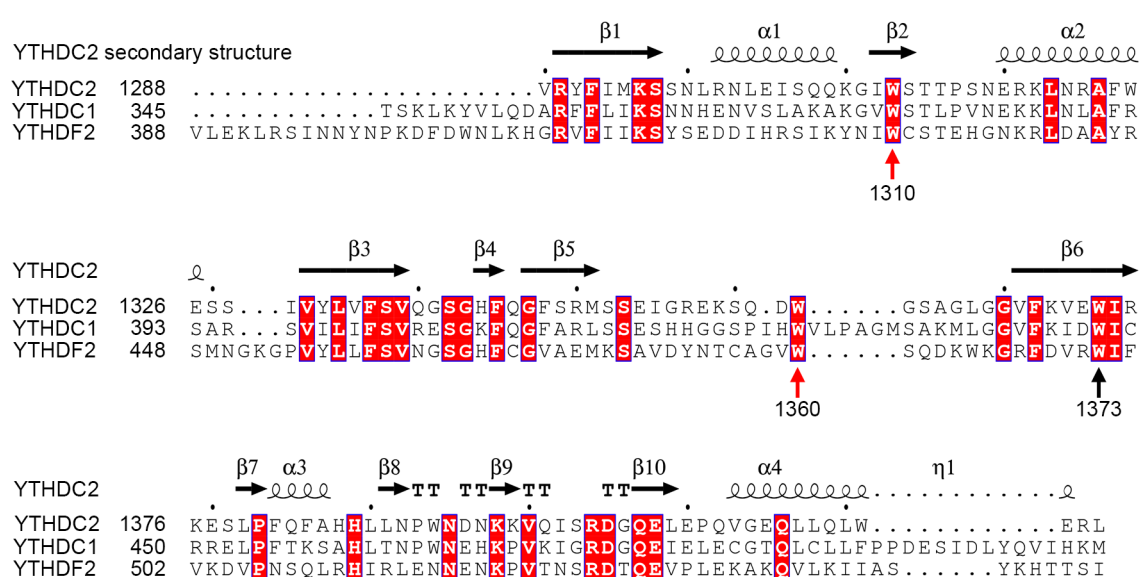
Figure 19: Sequence alignment of the YTH domains found in human proteins. Sequence alignment of the YTH domains of YTHDC2, YTHDC1 and YTHDF2 by the multi sequence aligner MUSCLE visualised with ESPript 3.0. The secondary structural features of the YTH domain of YTHDC2 (PDB 2YU6) are shown above the corresponding amino acids. Amino acids that are conserved in all three protein domains are shown with a red background and the red arrows indicate tryptophan 1310 and 1360 of YTHDC2. The black arrow indicates tryptophan 1373 of YTHDC2.
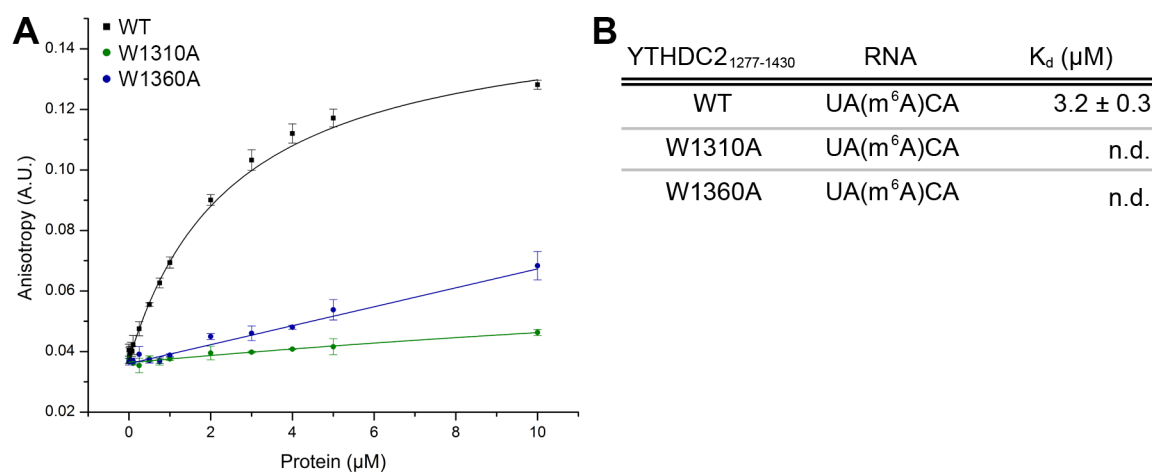
Figure 20: W1310 and W1360 of the YTH domain of YTHDC2 are required for binding to m^{6}A. A A fluorescine-labelled single stranded 9 nucleotide RNA containing the modified UA(m^{6}A)CA 18S rRNA sequence was incubated with HZZT-YTHDC2$_{1277-1430}$ containing either the W1310A substitution (green) or W1360A substitution (blue). Incubation with the wilde type protein (WT, black squares) is shown as a comparison (see Figure 18). **B** The dissociation constant (K$_d$) of the WT YTHDC2 experiment shown in **A** is given.

Together these results confirm that the YTH domain of YHTDC2 can specifically bind m^{6}A modified RNAs and that it has a significantly higher affinity to the modification in the sequence context of the 18S rRNA, implying that it may function as a reader of this modification in the cell. Mutational analysis showed that YTHDC2 has probably a similar binding mechanism for m^{6}A modified RNA as the other YTH domain-containing proteins.

3.5 Analysis of the cellular function of YTHDC2

The data show that YTHDC2 associates with ribosomal complexes, likely via its R3H domain and the YTH domain can specifically recognise the m^{6}A modification, with an increased affinity for modified nucleotide in the sequence context of the rRNA. This raises the question of what the function of such interactions might be in the cell. The immunofluorescence experiments showed that YTHDC2 predominantly localises to the cytoplasm but a fraction of the protein is also present in the nucleus (Figure 8). Together with the partially enclosed localisation of the m^{6}A modification of the 18S rRNA in the mature ribosome, this lead to the hypothesis that it might be involved in ribosome biogenesis and may be recognising the 18S rRNA m^{6}A modification in a pre-ribosomal complex. Sucrose density gradients could not answer the question whether YTHDC2 interacts with pre-ribosomal complexes or the mature ribosome, because both particles are located in similar fractions. Therefore, to determine if YTHDC2 is required for ribosome biogenesis, YTHDC2 was depleted from human cells by RNAi and effects on

the levels of precursor rRNAs and newly synthesised rRNAs were monitored by northern blotting and pulse-chase labelling experiments, respectively.

First, the knockdown efficiency of three independent siRNAs targeted against *YTHDC2* mRNA (siYTHDC2_1, siYTHDC2_2, siYTHDC2_3) was tested. HeLa cells were transfected with 30 nM siRNA and after 96 h, total cellular RNA was isolated. Analysis by RT-qPCR showed that the level of the *YTHDC2* mRNA in cells that had been transfected with siRNAs against YTHDC2 were reduced to 15-25 % of the level in wild-type cells or of cells that had been treated with the control siRNA. GAPDH and Tubulin mRNAs were used for normalisation of different cDNA concentrations (Figure 21A). In parallel, siRNA-treated cells were used to prepare whole cell extracts that were separated by SDS-PAGE and analysed by western blotting using antibodies against YTHDC2 and Tubulin as a loading control. In extracts from cells that had been transfected with siRNAs against YTHDC2, no signal could be detected for endogenous YTHDC2 using the αYTHDC2 antibody, showing that the knockdown leads to efficient protein depletion (Figure 21B).

Having established suitable knockdown conditions, northern blots were prepared to visualise accumulation or depletion of rRNA precursors. Therefore, human cells were treated with the three different siRNAs against *YTHDC2* and a control siRNA. Total cellular RNA was extracted, separated on an agarose-glyoxal gel and transferred to a membrane. Two DNA probes complementary to sequences either in ITS1 or ITS2 (depicted as red asterisks in Figure 22A) were used for monitoring SSU or LSU pre-rRNA processing. A probe recognising the actin mRNA was used as a loading control as well
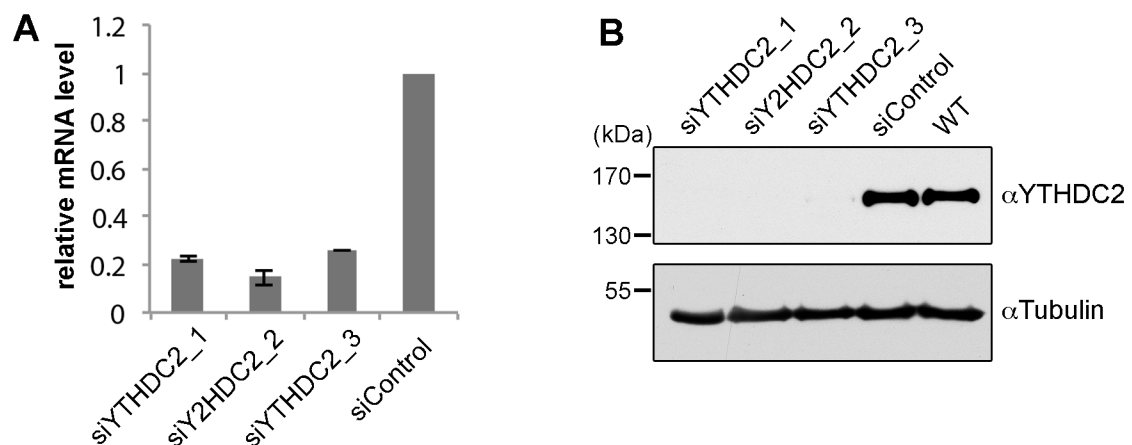
Figure 21: Establishment of RNAi against YTHDC2. A HEK293 cells were transfected with 30 nM of the indicated siRNA and after 96 h, RNA was isolated and used to generate cDNA. The levels of the YTHDC2 mRNA were determined by RT-qPCR and the levels of the GAPDH and Tubulin mRNAs were used for normalisation. **B** HEK293 WT cells were transfected with 30 nM of the indicated siRNA and after 96 h cells were harvested. Whole cell extract was prepared, separated by SDS-PAGE and analysed by western blotting using an αYTHDC2 antibody, or an αTubulin antibody as a loading control.

as methylene blue staining of mature 18S and 28S rRNA (Figure 22B). No changes in the levels of any pre-rRNA species could be detected while comparing knockdown samples (siYTHDC1_1, siYTHDC2_2, siYTHDC2_3) and control (siControl) or wild-type (WT) sample (Figure 22B), suggesting that ribosome biogenesis is not altered upon knockdown of YTHDC2. However, northern blots represent a steady state level of pre-cursor RNAs in which aberrant pre-rRNAs might have already been degraded by the RNA surveillance machinery, therefore, the next step was to investigate pre-RNA processing via pulse-chase experiments.

Pulse-chase experiments have a higher sensitivity, because only nascent rRNA transcripts are detected and this method can enable visualisation of subtle processing defects that effect mature rRNA abundance or the ratio of SSU to LSU maturation. Therefore, human cells were treated with siRNAs as described for the northern blotting. Then, cells were grown in the absence of phosphate to deplete the cells of non-radioactive phosphate. The medium was removed and the cells were grown in the presence of radioactively labelled phosphate (pulse) followed by cultivation in normal cell culture growth medium (chase). During the pulse, nascent RNA chains are labelled radioactively by the incorporation of radioactive phosphate. Then cellular RNA was extracted, separated on an agarose-glyoxal gel, transferred to a membrane and abundant, labelled RNAs were visualised using a phosphorimager (Figure 22C). The top panel of Figure 22C shows the signals of radioactively labelled 47S and 32S pre-rRNAs, and the mature 18S and 28S rRNAs. The ratio between the mature 18S and 28S rRNAs in the samples derived from cells treated with siRNAs against YTHDC2 was not changed compared to the control and WT cells, thus indicating that depletion of YTHDC2 does not specifically effect the maturation of one of the ribosomal subunits. Also, a general defect of rRNA maturation was not detected, because the total amount of radioactively labelled mature rRNAs was not altered by the siRNA treatment considering the UV loading control representing total amounts of 18S and 28S rRNA.

These results show that knockdown of YTHDC2 has no significant effect on pre-rRNA processing implying that this protein is not required for ribosome biogenesis and therefore likely predominantly interacts with mature ribosomes. However, mature ribosomes are present in large quantities in the cell and exceed the amount of YTHDC2. Therefore, it is very likely that YTHDC2 only interacts with a sub-population of cytoplasmic ribosomes. To identify this sub-population and gain more information about the function of YTHDC2, the next aim was to identify additional protein interaction partners.
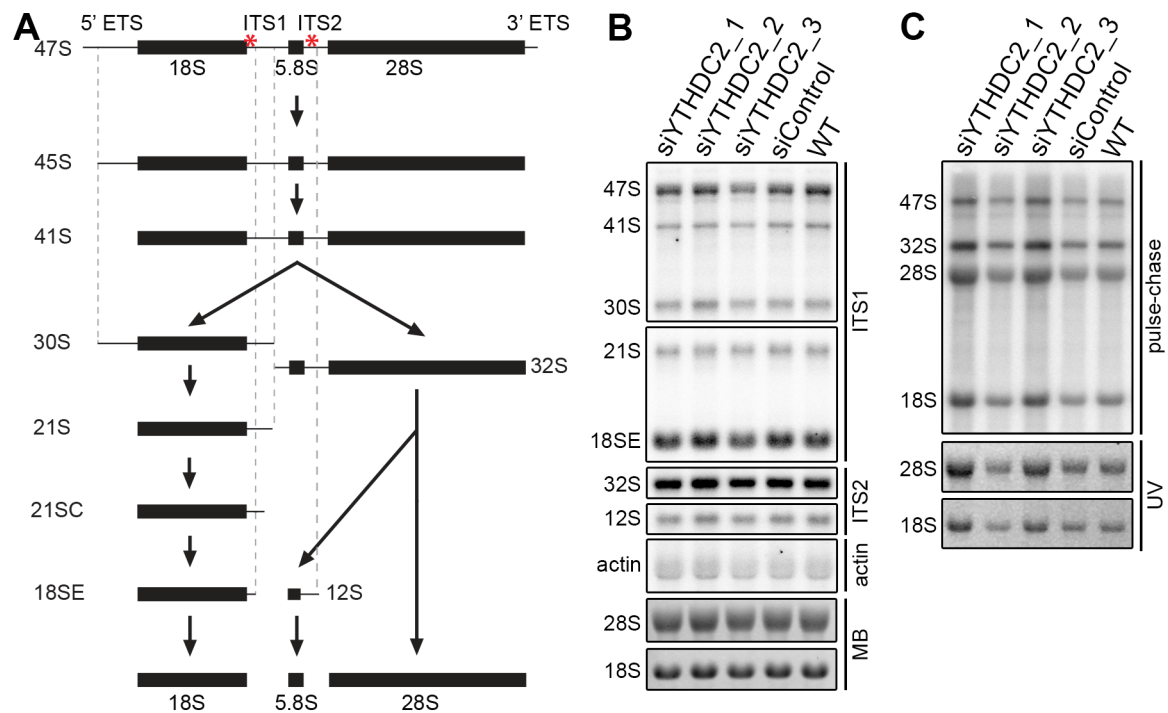
Figure 22: Depletion of YTHDC2 does not affect pre-rRNA processing. A Simplified scheme of pre-rRNA processing in human cells. The mature rRNAs are shown as rectangles and the internal and external transcribed spacers (ITS and ETS, respectively) are shown as lines. The binding sites of the probes used for northern blotting are indicated with red asterisks. **B** HEK293 cells were treated with 30 nM of the indicated siRNAs for 96 h. Total cellular RNA was isolated and subjected to northern blot analysis using the probes indicated to the right of the panel with the binding sites shown in **A**. Northern blotting using a probe against the actin mRNA was used as a loading control and mature rRNAs were visualised by methylene blue staining (MB). **C** HEK293 cells were treated with 30 nM of the indicated siRNAs for 96 h. Then the cells were pulse-labelled with ^{32}P-orthophosphate for 1 h and then grown in unlabelled media for 3 h before isolation of cellular RNA. The upper panel shows the newly synthesised radioactively labelled RNA (pulse-chase), the lower panels show the total amount of mature rRNA visualised by UV light (UV).

To get an overview of the protein-interactome of YTHDC2 immunoprecipitation (IP) of YTHDC2-containing complexes was performed, followed by the identification of the co-precipitated proteins using mass spectrometry. For this stable cell lines expressing YTHDC2-FLAG or the FLAG-tag alone were induced for 36 h, whole cell lysate was prepared and the complexes were purified using immobilised αFlag antibodies. Eluates were then separated by NuPAGE and analysed by mass spectrometry in collaboration with the group of Prof. Dr. Henning Urlaub. Analysis of the data showed that the cytoplasmic 5'-3' exonuclease XRN1 was the most enriched protein co-immunoprecipitating with YTHDC2-FLAG compared to the FLAG control, suggesting a stable interaction between YTHDC2 and XRN1. To confirm this interaction, immunoprecipitation assays of YTHDC2-FLAG were repeated and analysed by western

blotting using an αXRN1-specific antibody (Figure 23). Cell lines expressing FLAG-tagged version of other YTH domain-containing proteins (YTHDC1-FLAG, YTHDF1-FLAG, YTHDF2-FLAG and YTHDF3-FLAG) were included as additional controls to determine if the putative interaction with XRN1 is specific for YTHDC2-FLAG. Immunoprecipitation assays were performed as described above. In addition, a mix of RNase A and T1 was added to half of the sample to abolish RNA-mediated interactions by the digestion of cellular RNAs. The input, and the eluate from samples with and without RNase treatment, were then separated by SDS-PAGE and analysed by western blotting. Only the YTHDC2-FLAG IP shows a specific signal for XRN1 confirming the interaction detected by mass spectrometry and demonstrating the specificity of this interaction (Figure 23A). Also, addition of RNase during the purification did not disrupt the interaction, indicating that it is an RNA-independent protein-protein interaction.
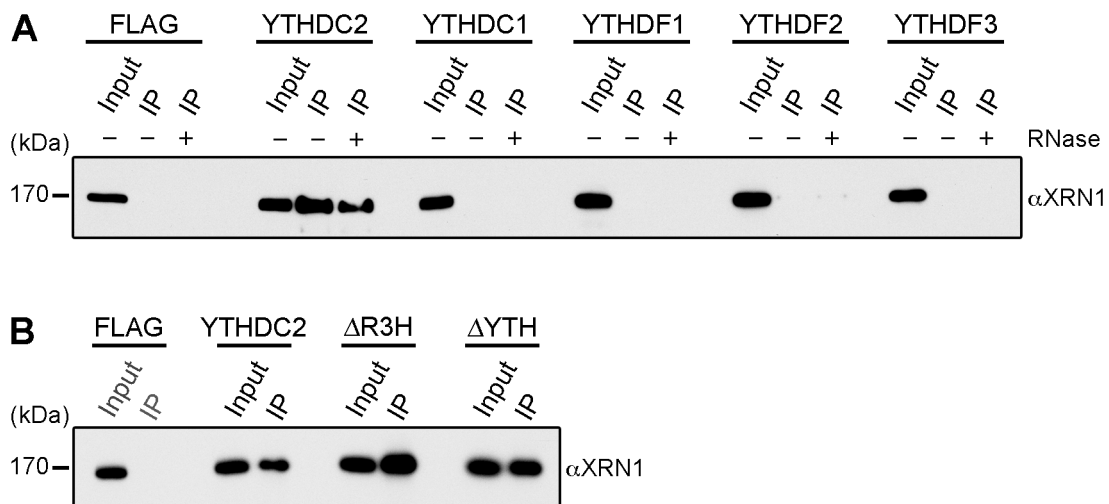
Figure 23: YTHDC2 associates with the cytoplasmic 5'-3' exonuclease XRN1. A Extracts from stable cell lines expressing the FLAG-tag alone (FLAG) or FLAG-tagged version of the indicated proteins were used for immunoprecipitation experiments. Protein-complexes were purified using an immobilised αFlag antibody. Inputs (1%) and elutions (IP) were analysed by western blotting using an αXRN1 antibody. Cell lysates were either treated with RNase (+) or left untreated (-) prior to immunoprecipitation. **B** Immunoprecipitation experiments were performed and analysed as in **A**, without RNase treatment, using extracts prepared from cells expressing the FLAG-tag alone (FLAG), full-length YTHDC2 (YTHDC2) or truncated version of YTHDC2 lacking the R3H-domain (ΔR3H) or the YTH domain (ΔYTH) (see Figure 15).

To gain more information about the YTHDC2 domain that forms contacts with XRN1, additional immunoprecipitation experiments were performed as described above using cell lines expressing the truncated FLAG-tagged versions of YTHDC2, i.e. ΔR3H-FLAG and ΔYTH-FLAG. Input and eluates were separated by SDS-PAGE and analysed by western blotting using an antibody against endogenous XRN1. In addition to the full-length protein, XRN1 was co-precipitated with ΔR3H-FLAG and ΔYTH-FLAG, indicating that XRN1 interacts with the central domain of the protein (Figure 23B). This is in line with the fact that the R3H and YTH domains are RNA-interaction domains, which have so far not been suggested to be involved in forming protein-protein interactions (Jaudzems et al., 2012; Xu et al., 2015). The interaction of YTHDC2 with the ribosome in combination with the exonuclease XRN1 suggest that the function of YTHDC2 and the recognition of the m^{6}A could be in rRNA degradation and quality control.

4 Discussion

4.1 Development of computational tools for the transcriptome-wide analysis of the RNA targets of RNA-modifying enzymes

In addition to the long known "epigenetic" methylations in DNA and various post-translational modifications of proteins (e.g. phosphorylation, ubiquitination, acetylation, etc.), chemical modifications are also found in most cellular RNAs. A wide variety of such RNA modifications exist in nature and in general, they serve to expand the chemical properties of the four basic nucleotides, meaning that they can regulate the functions of the RNAs that carry them. Many enzymes that mediate RNA modifications contain conserved protein domains that harbour their catalytic activity. Although the enzymes that are responsible for introducing some RNA modifications are known, the specific substrate RNAs and target nucleotides of many other putative RNA modifying enzymes remain to be determined.

The human genome encodes seven 5-methylcytosine (m^5C) RNA methyltransferases that belong to the Nol1/Nop2/SUN (NSUN) family (reviewed in Motorin et al., 2010). So far, proteins of this family have been linked to modifications at specific sites in cytoplasmic tRNAs (NSUN2; Brzezicha et al., 2006), mitochondrial and cytoplasmic rRNAs (NSUN4, NSUN1 and NSUN5; Camara et al., 2011; Schosserer et al., 2015; Sharma et al., 2013b) and enhancer RNAs (NSUN7; Aguilo et al., 2016), but the targets of the NSUN6 and NSUN3 m^5C RNA methyltransferases remained elusive. A strategy that can be used for the identification of the RNAs associated with RNA binding proteins is *in vivo* cross-linking followed by isolation of RNA-protein complexes, isolation of RNA and deep sequencing of a corresponding cDNA library (Bohnsack et al., 2012). This approach (CRAC) was employed in the Bohnsack lab to identify RNA-interaction partners of NSUN6 and NSUN3 (Haag et al., 2016; Haag et al., 2015b). However, the identification of the cellular RNAs bound by the proteins in this method requires that the obtained sequence reads are quality controlled and mapped to a well annotated version of the human genome. Therefore, bioinformatic algorithms and mapping tools were employed and further developed to generate a systematic pipeline, specifically adapted for the mapping and analysis of CRAC data derived from human cells.

Many newly developed techniques for the transcriptome-wide mapping of RNA modifications, determining the RNA-interactome of RNA-binding proteins and the analysis of gene expression, are based on the analysis of next generation sequencing

data (Bohnsack et al., 2012; Darnell, 2012; Hafner et al., 2010; Ingolia et al., 2009; Krogh et al., 2016; Nagalakshmi et al., 2008). For example, RNA-Seq, in which total cellular RNA (depleted of ribosomal RNA) is isolated, fragmented and sequenced, is used to investigate the transcriptome of a cell population at a given time under certain conditions (Nagalakshmi et al., 2008), as the relative number of unique sequence reads mapping to individual genes allows a quantitative statement about the expression level of the corresponding mRNAs. Similarly, ribosome profiling (Ingolia et al., 2009) can provide a snapshot of the mRNAs that are being translated in a given cell population by enabling sequencing and identification of ribosome-associated mRNAs. The analysis of CRAC data relies on similar principles as the sub-population of cellular RNA that is attached to the protein of interest is isolated, sequenced and mapped to the genome. The accumulation of multiple sequence reads mapping to a specific region of the genome then indicates binding of the protein to the corresponding RNA transcript. As RNA modifications can occur in the majority of transcripts and are highly abundant in non-coding RNAs, in contrast to approaches for analysis of RNA-Seq and ribosome profiling data, a bioinformatic pipeline for the mapping of CRAC data generated for RNA modification enzymes, requires a well-annotated and complete reference genome or transcriptome. Another difference between the analysis of CRAC data and the analysis of gene expression by RNA-Seq is that in CRAC, the cross-linked RNA-protein complexes are purified on matrices and the non-specific binding of RNAs to such beads could lead to background. Alternatives to the standard UV_{254} cross-linking, such as cross-linking with light at 365 nm after treatment of the cells with 4-thiouridine (PAR-CRAC), can be used to increase the specificity of cross-linking and furthermore, modules were developed within the bioinformatic pipeline to enable sorting and mapping of reads containing only specific mutations that are introduced by the direct cross-linking of the RNA and protein, thereby significantly reducing the non-specific background in the final data output of the analysis pipeline.

Also in contrast to RNA-Seq, in which only the number of reads mapping to the genes coding for individual transcripts is considered, one of the aims of CRAC is to identify the specific binding site of the protein on the RNA transcript. In the case of the RNA methyltransferases NSUN6 and NSUN3, analysis of the read distribution between different classes of RNA transcript and between different tRNA genes suggested that the cytoplasmic tRNAs tRNA[Cys] and tRNA[Thr] are bound by NSUN6 and the mitochondrial tRNA[Met] is associated with NSUN3. These putative target RNAs were confirmed by additional *in vivo* experiments, but close analysis of the distribution of mapped sequence reads on the tRNA sequences also provided the basis for the identification of the

modification target nucleotides of these enzymes. These could subsequently be determined by mutational analysis combined with *in vitro* methylation assays (Haag et al., 2015; Haag et al., 2016). The identification of the specific binding sites of proteins on their target RNAs is especially relevant for characterisation of proteins that contact the (pre-) ribosomal RNAs and for such proteins, additional scripts were added to the basic CRAC pipeline to enable mapping of the obtained sequence reads onto the available 2D structures of the mature rRNAs and the 3D structure of the human 80S ribosome. Such modelling significantly helps the interpretation of the obtained CRAC data, as it allows it to be determined if multiple cross-linking sites that may be distant on the linear sequence of a particular rRNA come in close proximity to each other on the folded RNA. It also enables the identification of other features in close proximity of the protein cross-linking sites, such as RNA modifications and the binding sites of other proteins that need to be considered in the context of the assembled RNP.

In the case of proteins that cross-link to mRNAs, one of the limitations of the current pipeline is the simplification of the annotation of the protein coding genes in the genome version to which sequences are mapped. This means that reads mapping to 5' and 3' UTRs cannot be distinguished from reads that map to the coding sequences. Similarly, the present pipeline is not able to map exon-exon spanning reads and such information can be highly valuable for understanding the functions of proteins involved in mRNA processing/mRNP biogenesis and can also be relevant for the analysis of proteins involved in RNA modification as an asymmetric distribution of RNA modifications is often observed, e.g. m^6A modifications are enriched around stop codons, long internal exons and in 3' UTRs (Chen et al., 2015; Dominissini et al., 2012; Linder et al., 2015) and m^1A modifications are typically clustered in 5' and 3' UTRs (Dominissini et al., 2016; Li et al., 2016). The analysis of such features can be done by using alignment tools, such as HISAT2 (Hierarchical Indexing for Spliced Alignment of Transcripts; Kim et al., 2015) or STAR (Spliced Transcripts Alignment to a Reference; Dobin et al., 2013) instead of the currently used Bowtie sequence alignment tool, as these algorithms are specially designed for the alignment of spliced reads spanning exon-exon junctions in mRNA. An alternative strategy for the mRNA analysis could be to use a dedicated mapping collection, such as HOMER (Hypergeometric Optimization of Motif EnRichment; Heinz et al., 2010) that was originally designed to identify binding motifs within deep sequencing data but which can also be used for genome-wide analysis of next generation sequencing data. Lastly, the recent availability of transcriptome-wide maps of sites of specific RNA modifications, such as m^6A, m^1A, pseudouridine and m^5C, means that it would also be interesting to also extend the CRAC pipeline to enable the overlap

between the cross-linking sites of a particular protein and the known sites of RNA modification to be automatically determined.

4.2 The YTH domain-containing proteins associate with different RNA substrates and perform diverse cellular functions

RNA modifications often influence RNA secondary structure and thereby can exert a stabilising or destabilising effect on the modified RNAs. Alternatively, RNA modifications can contribute to the function of the RNA, for example, RNA modifications in the anticodon of tRNAs can increase the decoding capacity of the tRNAs. Excitingly, as well as these direct influences of modifications on the functions of RNAs, it has recently emerged that RNA modifications can be recognised by specific proteins, termed "readers", and that binding of such proteins to modified RNAs can affect the cellular fate of such RNAs. This concept now extends beyond recognition of the m^7G cap at the 5' end of mRNAs by the cap binding complex of CBC20/CBC80 in the nucleus, which is essential for mRNA stability and export or recognition of this modification by eIF4E in the cytoplasm, which is required for mRNA translation (reviewed in Liu and Jia, 2014; Ramanathan et al., 2016). It was discovered that m^6A modifications in cellular RNAs can be recognised and bound by the YTH protein domain, however, the binding of two other non-YTH domain-containing proteins to RNA substrates has also been found to be dependent on m^6A modifications in the RNAs. HNRNPA2B1 recognises m^6A modification in microRNA precursors and its binding regulates processing events (Alarcon et al., 2015), while eIF3 promotes cap-independent translation and shows altered binding to its RNA substrates depending on the m^6A modification (Meyer et al., 2015).

The YTH domain protein family contains five members in humans and can be further divided into two subfamilies based on protein sequence similarity: YTHDF1, YTHDF2, YTHDF3, and YTHDC1 and YTHDC2. The CRAC approach, followed by bioinformatic analysis of obtain sequence reads, was applied to identify the RNA targets of all these YTH domain-containing proteins. This comparative analysis implied that YTHDF1, YTHDF2 and YTHDC1 predominantly associate with mRNA, which is in line with recently published data from the He, Pan and Yang labs (Wang et al., 2014a; Wang et al., 2015; Xiao et al., 2016), and similarly, YTHDF3 also seems to bind to mRNAs. The exclusively cytoplasmic localisation of all three YTHDF proteins (YTHDF1, YTHDF2 and YTHDF3) suggested that they associate with mature, rather than pre-mRNAs and indeed, YTHDF1 was shown to be involved in enhancing the translation of specific mRNAs (Wang et al., 2015) whereas YTHDF2 facilitates the localisation of certain mRNAs to p-bodies, thus

promoting their degradation (Wang et al., 2014a). The exact function of YTHDF3 is unknown so far, but sequence similarity to YTHDF1 suggests that it might perform a similar function on different mRNA substrates. Notably, YTHDF1 and YTHDF2 both bind to m^6A modifications that are located near mRNA stop codons and interestingly, these proteins share approximately 50 % of their target RNAs (Wang et al., 2015). They are proposed to bind to these mRNAs at different time points to enable tightly regulated translation of these transcripts. It is suggested that first, YTHDF1 binds the m^6A-modified mRNAs in order to promote their translation and then, YTHDF2 is recruited to efficiently deplete of these mRNAs resulting in an abrupt decrease in protein production. In contrast, YTHDC1 is a nuclear protein that was recently shown to act during mRNP biogenesis, by promoting alternative splicing of m^6A containing pre-mRNAs in the nucleus via the interaction with several splicing factors (Xiao et al., 2016).

Interestingly, in addition to interactions with m^6A-modified substrate mRNAs, protein interaction partners have also been identified for YTHDF1, YTHDF2 and YTHDC1. Consistent with its function in promoting mRNA translation, YTHDF1 was shown to interact with the eukaryotic translation initiation factor eIF3 (Wang et al., 2015) and similarly, in line with its role in mRNA degradation, YTHDF2 interacts with components of the CCR4-NOT deadenylation complex (Du et al., 2016; Wang et al., 2015). Likewise, the splicing factors SRSF3 and SRSF10 were found to be protein interaction partners of YTHDC1 (Xiao et al., 2016). YTHDF1, YTHDF2, YTHDF3 and YTHDC1 are all relatively small proteins and sequence analyses suggest that in all cases, the only defined domain they contain is the YTH domain, which mediates interactions with the m^6A modification in mRNAs. This suggests that the specific protein-protein interactions formed by these proteins are mediated by the divergent non-YTH domain regions of these proteins. In contrast to these proteins, YTHDC2 is a much larger protein that in addition to the YTH domain, also contains an R3H RNA-binding domain and a helicase core domain that contains two predicted ANK repeats (UniProt, 2015), a motif that is implicated in mediating protein-protein interactions (reviewed in Sedgwick and Smerdon, 1999). This implies that the RNA/RNP substrate(s), protein-protein interactions and mode of recruitment of YTHDC2 may differ considerably form the other members of the YTH domain protein family.

4.3 YTHDC2 associates with ribosomal complexes via an RNA binding motif

In contrast to YTHDF1, YTHDF2, YTHDF3 and YTHDC1 that predominately associate with mRNAs, ribosomal RNA sequences were found to be enriched in the CRAC data of YTHDC2 compared to the control sample and sucrose density gradient centrifugation followed by western blotting confirmed that YTHDC2 co-migrates with ribosomal complexes. More specifically, a cross-linking site of YTHDC2 was found in the 18S rRNA sequence and notably, similar to the 18S rRNA m^6A modification, this site is close to the 3' end of the rRNA. Further CRAC analysis using truncated forms of YTHDC2 suggested that this cross-linking is formed with the R3H domain of YTHDC2 and that this interaction is important for recruitment of YTHDC2 to the ribosome as lack of this domain reduced the fraction of rRNA sequences present in the CRAC data and decreased the height of the cross-linking peak at this position. However, no alteration in the CRAC profile compared to the full-length protein was observed when the YTH domain was lacking, implying that either this region of the protein cannot be efficiently cross-linked to bound RNAs or that this domain does not significantly contribute to the stable association of YTHDC2 with ribosomal complexes. Although the *in vitro* anisotropy data demonstrate that the YTH domain of YTHDC2 preferentially binds to m^6A modified RNAs, the 18S m^6A modification is likely not accessible on the surface of the mature ribosomal subunit (see Figure 13). This could suggest a model in which YTHDC2 is recruited to ribosomal complexes via its R3H domain and only once this stable association is formed, is the YTH domain positioned to probe for the presence of the m^6A modification. The position of the 18S m^6A modification within the ribosomal complex is likely in contrast to the majority of m^6A modifications present in mRNAs and this alternative substrate may explain why an additional RNA binding domain is specifically required in YTHDC2, but not other YTH domain-containing proteins. Alternatively, it is possible that the association of YTHDC2 with ribosomes via its R3H domain also positions the YTH domain of YTHDC2 for recognition of m^6A modification in other ribosome-associated RNAs, such as mRNAs. However, no specific mRNA targets of YTHDC2 could be identified within the CRAC data and the binding of the YTH domain of YTHDC2 to m^6A modifications in the $GG(m^6A)CU$ sequence context, present in most mRNAs, is relatively weak (see also below).

4.4 The sequence context of m⁶A can affect recognition by the YTH domains

Mapping of the m^6A modifications in mRNAs and identification of an RNA binding motif for the known m^6A methyltransferase complex comprised of METTL3, METTL14 and WTAP demonstrated that the majority of m^6A modifications in mRNAs lie within a $GG(m^6A)CU$ sequence motif. Notably, in contrast to this, the m^6A modifications present in both the 18S and 28S rRNAs are found in a different sequence context, $UA(m^6A)CA/G$ (Dominissini et al., 2012; Linder et al., 2015; Piekna-Przybylska et al., 2008a; Wei and Moss, 1977). The changes of G to U at -2 position, G to A at -1 and U to A or G at +2 considerably alters the hydrogen bond capabilities of the nucleotides in the vicinity of the m^6A modification and the charge of the local environment that, based on the available crystal structures of YTH domains together with m^6A nucleosides, could be expected to have a significant effect on the binding of a YTH domain to the modified residue. Analysis of the affinities of the YTH domain of YTHDC2 for unmodified and m^6A-modified RNAs confirmed that this domain preferentially binds to m^6A-modified RNAs, however, compared to the YTH domains of YTHDF2 and YTHDC1, the affinity of the YTH domain of YTHDC2 for the RNAs containing the m^6A modification in the $GG(m^6A)CU/$mRNA motif was lower. This might be due to the fact that YTHDC2 possess an additional RNA binding domain (R3H) that stabilises interactions with its substrate RNA *in vivo*, meaning that unlike the other YTH domain-containing proteins, a high affinity for the modified residue is not required for the YTH domain-containing protein to exert its effect on the RNA. This idea is supported by comparison of the amino acid sequences of the human YTH domains (see Figure 19) and the available crystal structures of these domains. The RNA-interaction surface of the YTH domains of YTHDF2 (PDB 4RDN) and YTHDC1 (PDB 4R3I) are more positively charged compared to the YTH domain of YTHDC2 (PDB 2YU6), suggesting that the YTH domains of YTHDF2 and YTHDC1 may interact more strongly with the RNA backbone than that of YTHDC2. This may in turn facilitate the initial binding of the YTH domains to RNA and subsequently tighter binding to the N^6-methylated residue once the modification is correctly positioned in the hydrophobic pocket of the YTH domain.

However, the anisotropy data suggest that the sequence context of the m^6A modification also influences the affinity of the different YTH domains. Although the YTH domain of YTHDC2 had a lower affinity for RNAs containing the $GG(m^6A)CU$ sequence than the YTH domains of YTHDF2 and YTHDC1, it bound more strongly to the $UA(m^6A)CA$ sequence found in the 18S rRNA. The higher affinity of the YTH domain of YTHDC2 for the m^6A modification in the rRNA context supports the model in which YTHDC2 is

recruited to the ribosome to enable recognition of the m6A present on the rRNA. Interestingly, the YTH domains of YTHDF2 and YTHDC1 show the opposite effect with a lower affinity for the UA(m6A)CA/rRNA sequence than the GG(m6A)CU/mRNA sequence. This is especially notable for the YTH domain of YTHDC1, which shows a 10-fold increased affinity towards the m6A modification in the mRNA sequence compared to the modified rRNA sequence. An explanation for this differential binding may lie in the different localisations of the YTHDF2 and YTHDC1 proteins. On the one hand, YTHDF2 is localised in the cytoplasm, where it will predominantly encounter mature ribosomes where the m6A modification is not readily accessible, meaning that discrimination between mRNA and rRNA is not necessary. On the other hand, YTHDC1 is localised in the nucleus, where it could interact with pre-ribosomal subunits that have a more open conformation and where the rRNA m6A modification may be exposed, making it advantageous to select against recognition of these modifications by having low affinity for the rRNA sequence context. Interestingly, the crystal structure of the YTH domain of YTHDC1 provides an explanation as to how this sequence discrimination is achieved (Xu et al., 2014). Analysis of the binding interface shows that the carboxyl oxygen of the guanine at the -1 position of the mRNA consensus motif can form hydrogen bonds with the secondary amine of the peptide backbone, stabilising RNA-protein interactions. Substituting this guanine with any other nucleotide would disrupt this hydrogen bonding and thereby decrease the binding to the RNA substrate. In particular, substitution to an adenine, which is present at the -1 position in the rRNA sequence context, would introduce steric clashes with the peptide backbone of YTHDC1, making substrates containing this sequence particularly unfavourable.

4.5 YTHDC2 associates with the cytoplasmic 5'-3' exonuclease XRN1

Ribosomal RNA modifications are found in all kingdoms of life and in general are proposed to help stabilise the tertiary structure of the ribosome, enabling efficient and accurate translation. Consistent with this, rRNA modifications cluster at functionally important sites on the ribosome, such as the peptidyl transferase centre and the decoding site (reviewed in Sharma and Lafontaine, 2015). In yeast, no m6A modifications are present in the rRNAs, while in humans these modifications are present at positions 1832 of the 18S rRNA and 4220 of the 28S rRNA (nucleotide numbering according to the human 80S ribosome structure, PDB 4V6X; Linder et al., 2015). It is not yet known what the specific functions of these modifications are, however, quantitative analysis of the extent of these modifications suggests that these sites are 98% modified in HeLa cells (Liu et al., 2013). This also suggests that the rRNA m6A modifications are not targeted by

the demethylases ALKBH5 and FTO that have been shown to reverse m^6A modification in mRNAs (reviewed in Liu and Jia, 2014).

To gain insight into the role of YTHDC2 on the ribosome and a possible function in recognition of the m^6A modification in the 18S rRNA, immunoprecipitation experiments followed by mass spectrometry were performed to identify protein interaction partners, as in the case of the other YTH domain-containing proteins YTHDF1, YTHDF2 and YTHDC1 such analysis provided functional information. This revealed that YTHDC2 interacts with the cytoplasmic 5'-3' exonuclease XRN1 (see also Figure 24). Further immunoprecipitation experiments suggested that XRN1 interacts with the central region of YTHDC2, as neither the R3H domain nor the YTH domain were required for this interaction. The central region of YTHDC2 contains two predicted ANK repeats, which are protein interaction motifs that may allow YTHDC2 to form protein-protein interactions with XRN1. Interestingly, XRN1 has an extended C-terminus, which is not conserved in the related, nuclear 5'-3' exonuclease XRN2 and which is suggested to serve as a protein-protein interaction domain in yeast (Chang et al., 2011). It is therefore possible that this region of XRN1 is involved in the interaction with YTHDC2 and that XRN1 is recruited to its numerous RNA substrates in the cytoplasm by various adaptor proteins, including YTHDC2.

Since XRN1 is the major exonuclease component of the cytoplasmic RNA degradation machinery, this suggests that recruitment of YTHDC2 to the ribosome or to m^6A-containing RNAs may be linked to RNA surveillance or RNA turnover. In the context of the ribosome, in yeast, Xrn1 is linked to two related RNA decay pathways: no-go decay (NGD) in which aberrant mRNAs that are stalled on the ribosome during translation elongation are degraded and non-functional ribosome decay (NRD) where defective ribosomes are removed (reviewed in Graille and Seraphin, 2012; Lafontaine, 2010). It is possible that YTHDC2 contributes to either of these pathways by recruiting XRN1 to either aberrant ribosomes or to ribosomes that are blocked in translation. In support of a hypothesis that YTHDC2 recruits to aberrant ribosomes for NRD is that the misfolded ribosomes in which the rRNA m^6A modification(s) are readily accessible may be detected by the YTH domain of YTHDC2 and thereby targeted for degradation by XRN1.
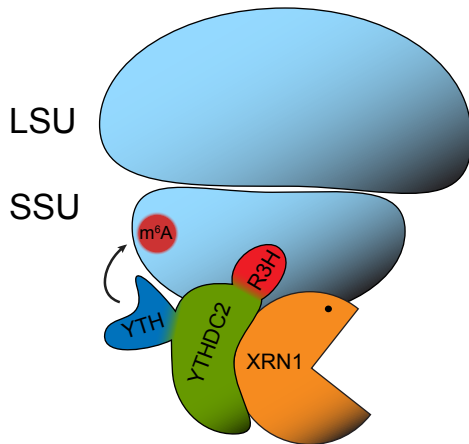
Figure 24 Model of YTHDC2 ribosome association. YTHDC2 in complex with the 5'-3' exonuclease XRN1 interacts with the ribosome via the R3H domain and the YTH domain is probing for the 18S rRNA m^{6}A modification in the SSU. The R3H domain, helicase core domain and YTH domain of YHTDC2 are coloured in red, green and blue, respectively. Abbreviations: LSU, large ribosomal subunit, SSU; small ribosomal subunit

The pathway of NRD is poorly characterised in human cells but in yeast, both NRD and NGD are known to involve the protein Dom34 and Hbs1, which are also conserved in humans (Passos et al., 2009; Shoemaker et al., 2010). Since neither of these proteins was identified in the mass spectrometry analysis of YTHDC2-containing complexes, this could suggest that the YTHDC2-XRN1 complex functions independently in such pathways. In yeast, the pathway of NRD has mostly been studied using rRNA reporters that carry mutations in functional regions of the rRNA and which are expressed from an RNA polymerase II driven promoter (Cole et al., 2009; LaRiviere et al., 2006). However, expression of rRNAs in such a way has additional effects on ribosome assembly and due to the increased complexity of pre-rRNA transcription in human cells, such a system is currently not available to study the NRD pathway in human cells.

Alternatively, YTHDC2 may also function in NRD by monitoring the presence of the m^{6}A modification on the rRNA and directing ribosomes lacking this modification for degradation since the absence of this modification may have negative effects on translation fidelity or efficiency. Notably, another specific feature of the YTHDC2 domain architecture is the presence of an RNA helicase domain. RNA helicases are a ubiquitously expressed family of proteins that can unwind RNA duplexes in an ATP-dependent manner. Most RNA helicases can be classified as either DEAD- or DEAH-box proteins based on the presence of a conserved sequence motif. DEAD- and DEAH-box helicases are distinct in their RNA unwinding mechanism (reviewed in Jarmoskaite and Russell, 2014). DEAD box helicases perform local strand unwinding whereas DEAH box helicases are processive and can move along an RNA strand while unwinding other base paired RNA strands. YTHDC2 is predicted to belong to the family of DEAH box RNA helicases and an earlier study reported an RNA dependent ATPase activity of the C-terminal part of the protein consisting of the amino acids 761-1430 (Morohashi et al.,

2011), implying that YTHDC2 is an active RNA helicase although further experiments using the full-length protein or complete helicase core domain would be required to confirm this. It is possible, therefore, that the helicase activity contributes to rRNA remodelling in the vicinity of the rRNA m⁶A modification(s) to enable the YTH domain of YTHDC2 to detect the presence of these modifications *in vivo*. The helicase domain of YTHDC2 may also be required for efficient RNA degradation by XRN1 as this exonuclease is not very active on highly structured substrates (e.g. ribosomal RNAs) (Poole and Stevens, 1997). A complex containing XRN1 and the helicase DDX6 is used in 5'-3' mRNA degradation, where the helicase might unwind secondary structures prior to degradation by XRN1 (Ozgur et al., 2010) as it is the case in 3'-5' RNA degradation by the exosome. RNA helicases unwind mRNA secondary structures ahead of the exosome and channel the RNA into the exosome for degradation (reviewed in Schneider and Tollervey, 2013).

5 Conclusion

RNA modifications play diverse and important roles in regulating the functions of the RNAs that carry them. Furthermore, RNA modifications in cellular RNAs can be specifically recognised by proteins termed "readers" that bind to the modified nucleotides and regulate the fate of the modified RNA. In humans, the five YTH domain proteins bind to RNAs containing m⁶A modifications and mediate diverse functions. In line with recently published data, CRAC analysis showed that YTHDF1, YTHDF2, YTHDF3 and YTHDC1 were found to predominantly associate with mRNAs, which in contrast YTHDC2 was found to be associated with ribosomal complexes and to cross-linking to the 18S rRNA in proximity to an m⁶A modification. *In vitro* anisotropy experiments demonstrated that, unlike the YTH domains of YTHDF2 and YTHDC1, the YTH domain of YTHDC2 has a higher affinity for m⁶A modifications in the sequence context present in the rRNA rather than the classical m⁶A consensus motif found in many mRNAs implying that it may recognise the 18S rRNA m⁶A modification. Furthermore, YTHDC2 was found to interact with the 5'-3' exonuclease XRN1 suggesting that YTHDC2 may serve as an adaptor protein to target m⁶A-containing RNAs for degradation.

Bibliography

Aebi, M., Kirchner, G., Chen, J.Y., Vijayraghavan, U., Jacobson, A., Martin, N.C., and Abelson, J. (1990). Isolation of a temperature-sensitive mutant with an altered tRNA nucleotidyltransferase and cloning of the gene encoding tRNA nucleotidyltransferase in the yeast Saccharomyces cerevisiae. J Biol Chem *265*, 16216-16220.

Agris, P.F., Sierzputowska-Gracz, H., and Smith, C. (1986). Transfer RNA contains sites of localized positive charge: carbon NMR studies of [13C]methyl-enriched Escherichia coli and yeast tRNAPhe. Biochemistry *25*, 5126-5131.

Aguilo, F., Li, S., Balasubramaniyan, N., Sancho, A., Benko, S., Zhang, F., Vashisht, A., Rengasamy, M., Andino, B., Chen, C.H., Zhou, F., Qian, C., Zhou, M.M., Wohlschlegel, J.A., Zhang, W., Suchy, F.J., and Walsh, M.J. (2016). Deposition of 5-Methylcytosine on Enhancer RNAs Enables the Coactivator Function of PGC-1alpha. Cell Rep *14*, 479-492.

Aik, W., Scotti, J.S., Choi, H., Gong, L., Demetriades, M., Schofield, C.J., and McDonough, M.A. (2014). Structure of human RNA N(6)-methyladenine demethylase ALKBH5 provides insights into its mechanisms of nucleic acid recognition and demethylation. Nucleic Acids Res *42*, 4741-4754.

Alarcon, C.R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S.F. (2015). HNRNPA2B1 Is a Mediator of m(6)A-Dependent Nuclear RNA Processing Events. Cell *162*, 1299-1308.

Alexandrov, A., Chernyakov, I., Gu, W., Hiley, S.L., Hughes, T.R., Grayhack, E.J., and Phizicky, E.M. (2006). Rapid tRNA decay can result from lack of nonessential modifications. Mol Cell *21*, 87-96.

Anger, A.M., Armache, J.P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D.N., and Beckmann, R. (2013). Structures of the human and Drosophila 80S ribosome. Nature *497*, 80-85.

Armache, J.P., Anger, A.M., Marquez, V., Franckenberg, S., Frohlich, T., Villa, E., Berninghausen, O., Thomm, M., Arnold, G.J., Beckmann, R., and Wilson, D.N. (2013). Promiscuous behaviour of archaeal ribosomal proteins: implications for eukaryotic ribosome evolution. Nucleic Acids Res *41*, 1284-1293.

Armistead, J., Khatkar, S., Meyer, B., Mark, B.L., Patel, N., Coghlan, G., Lamont, R.E., Liu, S., Wiechert, J., Cattini, P.A., Koetter, P., Wrogemann, K., Greenberg, C.R., Entian, K.D., Zelinski, T., and Triggs-Raine, B. (2009). Mutation of a gene essential for ribosome biogenesis, EMG1, causes Bowen-Conradi syndrome. Am J Hum Genet *84*, 728-739.

Arts, G.J., Kuersten, S., Romby, P., Ehresmann, B., and Mattaj, I.W. (1998). The role of exportin-t in selective nuclear export of mature tRNAs. EMBO J *17*, 7430-7441.

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

Ban, N., Beckmann, R., Cate, J.H., Dinman, J.D., Dragon, F., Ellis, S.R., Lafontaine, D.L., Lindahl, L., Liljas, A., Lipton, J.M., McAlear, M.A., Moore, P.B., Noller, H.F., Ortega, J., Panse, V.G., Ramakrishnan, V., Spahn, C.M., Steitz, T.A., Tchorzewski, M., Tollervey, D., Warren, A.J., Williamson, J.R., Wilson, D., Yonath, A., and Yusupov, M. (2014). A new system for naming ribosomal proteins. Curr Opin Struct Biol *24*, 165-169.

Banerjee, R., Chen, S., Dare, K., Gilreath, M., Praetorius-Ibba, M., Raina, M., Reynolds, N.M., Rogers, T., Roy, H., Yadavalli, S.S., and Ibba, M. (2010). tRNAs: cellular barcodes for amino acids. FEBS Lett *584*, 387-395.

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 research *37*, 7665-7677.

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 A resolution. Science *334*, 1524-1529.

Birkedal, U., Christensen-Dalsgaard, M., Krogh, N., Sabarinathan, R., Gorodkin, J., and Nielsen, H. (2015). Profiling of ribose methylations in RNA by high-throughput sequencing. Angew Chem Int Ed Engl *54*, 451-455.

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

Bodi, Z., Bottley, A., Archer, N., May, S.T., and Fray, R.G. (2015). Yeast m6A Methylated mRNAs Are Enriched on Translating Ribosomes during Meiosis, and under Rapamycin Treatment. PloS one *10*, e0132090.

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.

Boissel, S., Reish, O., Proulx, K., Kawagoe-Takaki, H., Sedgwick, B., Yeo, G.S., Meyre, D., Golzio, C., Molinari, F., Kadhom, N., Etchevers, H.C., Saudek, V., Farooqi, I.S., Froguel, P., Lindahl, T., O'Rahilly, S., Munnich, A., and Colleaux, L. (2009). Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. Am J Hum Genet *85*, 106-111.

Bokar, J.A., Shambaugh, M.E., Polayes, D., Matera, A.G., and Rottman, F.M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA *3*, 1233-1247.

Booth, M.J., Ost, T.W., Beraldi, D., Bell, N.M., Branco, M.R., Reik, W., and Balasubramanian, S. (2013). Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. Nat Protoc *8*, 1841-1851.

Bourgeois, G., Ney, M., Gaspar, I., Aigueperse, C., Schaefer, M., Kellner, S., Helm, M., and Motorin, Y. (2015). Eukaryotic rRNA Modification by Yeast 5-Methylcytosine-Methyltransferases and Human Proliferation-Associated Antigen p120. PloS one *10*, e0133321.

Brzezicha, B., Schmidt, M., Makalowska, I., Jarmolowski, A., Pienkowska, J., and Szweykowska-Kulinska, Z. (2006). Identification of human tRNA:m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). Nucleic Acids Res *34*, 6034-6043.

Buchhaupt, M., Sharma, S., Kellner, S., Oswald, S., Paetzold, M., Peifer, C., Watzinger, P., Schrader, J., Helm, M., and Entian, K.D. (2014). Partial methylation at Am100 in 18S rRNA of baker's yeast reveals ribosome heterogeneity on the level of eukaryotic rRNA modification. PloS one *9*, e89640.

Camara, Y., Asin-Cayuela, J., Park, C.B., Metodiev, M.D., Shi, Y., Ruzzenente, B., Kukat, C., Habermann, B., Wibom, R., Hultenby, K., Franz, T., Erdjument-Bromage, H., Tempst, P., Hallberg, B.M., Gustafsson, C.M., and Larsson, N.G. (2011). MTERF4 regulates translation by targeting the methyltransferase NSUN4 to the mammalian mitochondrial ribosome. Cell Metab *13*, 527-539.

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

Carroll, S.M., Narayan, P., and Rottman, F.M. (1990). N6-methyladenosine residues in an intron-specific region of prolactin pre-mRNA. Mol Cell Biol *10*, 4456-4465.

Chan, P.P., and Lowe, T.M. (2009). GtRNAdb: a database of transfer RNA genes detected in genomic sequence. Nucleic Acids Res *37*, D93-97.

Chang, J.H., Xiang, S., Xiang, K., Manley, J.L., and Tong, L. (2011). Structural and biochemical studies of the 5'-->3' exoribonuclease Xrn1. Nat Struct Mol Biol *18*, 270-276.

Chawla, G., and Sokol, N.S. (2014). ADAR mediates differential expression of polycistronic microRNAs. Nucleic Acids Res *42*, 5245-5255.

Chen, K., Lu, Z., Wang, X., Fu, Y., Luo, G.Z., Liu, N., Han, D., Dominissini, D., Dai, Q., Pan, T., and He, C. (2015). High-resolution N(6) -methyladenosine (m(6) A) map using photo-crosslinking-assisted m(6) A sequencing. Angew Chem Int Ed Engl *54*, 1587-1590.

Church, C., Moir, L., McMurray, F., Girard, C., Banks, G.T., Teboul, L., Wells, S., Bruning, J.C., Nolan, P.M., Ashcroft, F.M., and Cox, R.D. (2010). Overexpression of Fto leads to increased food intake and results in obesity. Nat Genet *42*, 1086-1092.

Ciganda, M., and Williams, N. (2011). Eukaryotic 5S rRNA biogenesis. Wiley Interdiscip Rev RNA *2*, 523-533.

Cole, S.E., LaRiviere, F.J., Merrikh, C.N., and Moore, M.J. (2009). A convergence of rRNA and mRNA quality control pathways revealed by mechanistic analysis of nonfunctional rRNA decay. Mol Cell *34*, 440-450.

Darnell, R. (2012). CLIP (cross-linking and immunoprecipitation) identification of RNAs bound by a specific protein. Cold Spring Harb Protoc *2012*, 1146-1160.

Delatte, B., Wang, F., Ngoc, L.V., Collignon, E., Bonvin, E., Deplus, R., Calonne, E., Hassabi, B., Putmans, P., Awe, S., Wetzel, C., Kreher, J., Soin, R., Creppe, C., Limbach, P.A., Gueydan, C., Kruys, V., Brehm, A., Minakhina, S., Defrance, M., Steward, R., and Fuks, F. (2016). RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. Science *351*, 282-285.

Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M., and Yusupova, G. (2012). A new understanding of the decoding principle on the ribosome. Nature *484*, 256-259.

Demeshkina, N., Jenner, L., Yusupova, G., and Yusupov, M. (2010). Interactions of the ribosome with mRNA and tRNA. Curr Opin Struct Biol *20*, 325-332.

Dewe, J.M., Whipple, J.M., Chernyakov, I., Jaramillo, L.N., and Phizicky, E.M. (2012). The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. RNA *18*, 1886-1896.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Dodt, M., Roehr, J.T., Ahmed, R., and Dieterich, C. (2012). FLEXBAR-Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. Biology (Basel) *1*, 895-905.

Doll, A., and Grzeschik, K.H. (2001). Characterization of two novel genes, WBSCR20 and WBSCR22, deleted in Williams-Beuren syndrome. Cytogenet Cell Genet *95*, 20-27.

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., Sorek, R., and Rechavi, G. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature *485*, 201-206.

Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M.S., Dai, Q., Di Segni, A., Salmon-Divon, M., Clark, W.C., Zheng, G., Pan, T., Solomon, O., Eyal, E., Hershkovitz, V., Han, D., Dore, L.C., Amariglio, N., Rechavi, G., and He, C. (2016). The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. Nature *530*, 441-446.

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., and Wu, L. (2016). YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat Commun *7*, 12626.

Dubin, D.T., and Taylor, R.H. (1975). The methylation state of poly A-containing messenger RNA from cultured hamster cells. Nucleic Acids Res *2*, 1653-1668.

Edelheit, S., Schwartz, S., Mumbach, M.R., Wurtzel, O., and Sorek, R. (2013). Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m5C within archaeal mRNAs. PLoS Genet *9*, e1003602.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res *32*, 1792-1797.

Eschrich, D., Buchhaupt, M., Kotter, P., and Entian, K.D. (2002). Nep1p (Emg1p), a novel protein conserved in eukaryotes and archaea, is involved in ribosome biogenesis. Curr Genet *40*, 326-338.

Fanale, D., Iovanna, J.L., Calvo, E.L., Berthezene, P., Belleau, P., Dagorn, J.C., Bronte, G., Cicero, G., Bazan, V., Rolfo, C., Santini, D., and Russo, A. (2014). Germline copy number variation in the YTHDC2 gene: does it have a role in finding a novel potential molecular target involved in pancreatic adenocarcinoma susceptibility? Expert Opin Ther Targets *18*, 841-850.

Feng, C., Liu, Y., Wang, G., Deng, Z., Zhang, Q., Wu, W., Tong, Y., Cheng, C., and Chen, Z. (2014). Crystal structures of the human RNA demethylase Alkbh5 reveal basis for substrate recognition. J Biol Chem *289*, 11571-11583.

Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m(6)A RNA methylation. Nat Rev Genet *15*, 293-306.

Fu, Y., Jia, G., Pang, X., Wang, R.N., Wang, X., Li, C.J., Smemo, S., Dai, Q., Bailey, K.A., Nobrega, M.A., Han, K.L., Cui, Q., and He, C. (2013). FTO-mediated formation of N6-hydroxymethyladenosine and N6-formyladenosine in mammalian RNA. Nat Commun *4*, 1798.

Fustin, J.M., Doi, M., Yamaguchi, Y., Hida, H., Nishimura, S., Yoshida, M., Isagawa, T., Morioka, M.S., Kakeya, H., Manabe, I., and Okamura, H. (2013). RNA-methylation-dependent RNA processing controls the speed of the circadian clock. Cell *155*, 793-806.

Gasse, L., Flemming, D., and Hurt, E. (2015). Coordinated Ribosomal ITS2 RNA Processing by the Las1 Complex Integrating Endonuclease, Polynucleotide Kinase, and Exonuclease Activities. Mol Cell *60*, 808-815.

Gerber, A.P., and Keller, W. (1999). An adenosine deaminase that generates inosine at the wobble position of tRNAs. Science *286*, 1146-1149.

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

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.

Graille, M., and Seraphin, B. (2012). Surveillance pathways rescuing eukaryotic ribosomes lost in translation. Nat Rev Mol Cell Biol *13*, 727-735.

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.

Grosjean, H., Szweykowska-Kulinska, Z., Motorin, Y., Fasiolo, F., and Simos, G. (1997). Intron-dependent enzymatic formation of modified nucleosides in eukaryotic tRNAs: a review. Biochimie *79*, 293-302.

Grosshans, H., Simos, G., and Hurt, E. (2000). Review: transport of tRNA out of the nucleus-direct channeling to the ribosome? J Struct Biol *129*, 288-294.

Haag, S., Kretschmer, J., and Bohnsack, M.T. (2015a). WBSCR22/Merm1 is required for late nuclear pre-ribosomal RNA processing and mediates N7-methylation of G1639 in human 18S rRNA. RNA *21*, 180-187.

Haag, S., Kretschmer, J., Sloan, K.E., and Bohnsack, M.T. Crosslinking methods to identify RNA methyltransferase targets in vivo. Methods Mol Biol, *accepted manuscript*.

Haag, S., Sloan, K.E., Ranjan, N., Warda, A.S., Kretschmer, J., Blessing, C., Hubner, B., Seikowski, J., Dennerlein, S., Rehling, P., Rodnina, M.V., Hobartner, C., and Bohnsack, M.T. (2016). NSUN3 and ABH1 modify the wobble position of mt-tRNAMet to expand codon recognition in mitochondrial translation. EMBO J *35*, 2104-2119.

Haag, S., Warda, A.S., Kretschmer, J., Gunnigmann, M.A., Hobartner, C., and Bohnsack, M.T. (2015b). NSUN6 is a human RNA methyltransferase that catalyzes formation of m5C72 in specific tRNAs. RNA *21*, 1532-1543.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.C., Munschauer, M., Ulrich, A., Wardle, G.S., Dewell, S., Zavolan, M., and Tuschl, T. (2010). PAR-CliP--a method to identify transcriptome-wide the binding sites of RNA binding proteins. J Vis Exp.

Hafner, M., Renwick, N., Brown, M., Mihailovic, A., Holoch, D., Lin, C., Pena, J.T., Nusbaum, J.D., Morozov, P., Ludwig, J., Ojo, T., Luo, S., Schroth, G., and Tuschl, T. (2011). RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. RNA *17*, 1697-1712.

Han, Z., Niu, T., Chang, J., Lei, X., Zhao, M., Wang, Q., Cheng, W., Wang, J., Feng, Y., and Chai, J. (2010). Crystal structure of the FTO protein reveals basis for its substrate specificity. Nature *464*, 1205-1209.

Hartmann, A.M., Nayler, O., Schwaiger, F.W., Obermeier, A., and Stamm, S. (1999). The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn). Mol Biol Cell *10*, 3909-3926.

Hastings, M.H. (2013). m(6)A mRNA methylation: a new circadian pacesetter. Cell *155*, 740-741.

Hayrapetyan, A.S.-L., S.; Helm, M. (2009). Function of Modified Nucleosides

in RNA Stabilization. In DNA and RNA Modifi cation Enzymes: Structure, Mechanism, Function and Evolution, H. Grosjean, ed. (Landes Bioscience), pp. 550-563.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Henras, A.K., Plisson-Chastang, C., O'Donohue, M.F., Chakraborty, A., and Gleizes, P.E. (2015). An overview of pre-ribosomal RNA processing in eukaryotes. Wiley Interdiscip Rev RNA *6*, 225-242.

Hoernes, T.P., Clementi, N., Faserl, K., Glasner, H., Breuker, K., Lindner, H., Huttenhofer, A., and Erlacher, M.D. (2016). Nucleotide modifications within bacterial messenger RNAs regulate their translation and are able to rewire the genetic code. Nucleic Acids Res *44*, 852-862.

Hoernes, T.P., and Erlacher, M.D. (2016). Translating the epitranscriptome. Wiley Interdiscip Rev RNA.

Hongay, C.F., and Orr-Weaver, T.L. (2011). Drosophila Inducer of MEiosis 4 (IME4) is required for Notch signaling during oogenesis. Proc Natl Acad Sci U S A *108*, 14855-14860.

Hopper, A.K. (2013). Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast Saccharomyces cerevisiae. Genetics *194*, 43-67.

Huber, S.M., van Delft, P., Mendil, L., Bachman, M., Smollett, K., Werner, F., Miska, E.A., and Balasubramanian, S. (2015). Formation and abundance of 5-hydroxymethylcytosine in RNA. Chembiochem *16*, 752-755.

Imai, Y., Matsuo, N., Ogawa, S., Tohyama, M., and Takagi, T. (1998). Cloning of a gene, YT521, for a novel RNA splicing-related protein induced by hypoxia/reoxygenation. Brain Res Mol Brain Res *53*, 33-40.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218-223.

Ito, S., Akamatsu, Y., Noma, A., Kimura, S., Miyauchi, K., Ikeuchi, Y., Suzuki, T., and Suzuki, T. (2014a). A single acetylation of 18 S rRNA is essential for biogenesis of the small ribosomal subunit in Saccharomyces cerevisiae. J Biol Chem *289*, 26201-26212.

Ito, S., Horikawa, S., Suzuki, T., Kawauchi, H., Tanaka, Y., Suzuki, T., and Suzuki, T. (2014b). Human NAT10 is an ATP-dependent RNA acetyltransferase responsible for N4-acetylcytidine formation in 18 S ribosomal RNA (rRNA). J Biol Chem *289*, 35724-35730.

Jackman, J.E., and Alfonzo, J.D. (2013). Transfer RNA modifications: nature's combinatorial chemistry playground. Wiley Interdiscip Rev RNA *4*, 35-48.

Jarmoskaite, I., and Russell, R. (2014). RNA helicase proteins as chaperones and remodelers. Annu Rev Biochem *83*, 697-725.

Jaudzems, K., Jia, X., Yagi, H., Zhulenkovs, D., Graham, B., Otting, G., and Liepinsh, E. (2012). Structural basis for 5'-end-specific recognition of single-stranded DNA by the R3H domain from human Smubp-2. J Mol Biol *424*, 42-53.

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.G., and He, C. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol *7*, 885-887.

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

Kariko, K., Muramatsu, H., Welsh, F.A., Ludwig, J., Kato, H., Akira, S., and Weissman, D. (2008). Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. Mol Ther *16*, 1833-1840.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods *12*, 357-360.

King, T.H., Liu, B., McCully, R.R., and Fournier, M.J. (2003). Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. Mol Cell *11*, 425-435.

Kishore, S., and Stamm, S. (2006). The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. Science *311*, 230-232.

Kornprobst, M., Turk, M., Kellner, N., Cheng, J., Flemming, D., Kos-Braun, I., Kos, M., Thoms, M., Berninghausen, O., Beckmann, R., and Hurt, E. (2016). Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome. Cell *166*, 380-393.

Kotelawala, L., Grayhack, E.J., and Phizicky, E.M. (2008). Identification of yeast tRNA Um(44) 2'-O-methyltransferase (Trm44) and demonstration of a Trm44 role in sustaining levels of specific tRNA(Ser) species. RNA *14*, 158-169.

Krogh, N., Jansson, M.D., Hafner, S.J., Tehler, D., Birkedal, U., Christensen-Dalsgaard, M., Lund, A.H., and Nielsen, H. (2016). Profiling of 2'-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. Nucleic Acids Res *44*, 7884-7895.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Lafontaine, D., Vandenhaute, J., and Tollervey, D. (1995). The 18S rRNA dimethylase Dim1p is required for pre-ribosomal RNA processing in yeast. Genes Dev *9*, 2470-2481.

Lafontaine, D.L. (2010). A 'garbage can' for ribosomes: how eukaryotes degrade their ribosomes. Trends Biochem Sci *35*, 267-277.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

LaRiviere, F.J., Cole, S.E., Ferullo, D.J., and Moore, M.J. (2006). A late-acting quality control process for mature eukaryotic rRNAs. Mol Cell *24*, 619-626.

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.

Lebedeva, S., Jens, M., Theil, K., Schwanhausser, B., Selbach, M., Landthaler, M., and Rajewsky, N. (2011). Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. Mol Cell *43*, 340-352.

Lee, S., and Kim, J. (2016). NGS-based deep bisulfite sequencing. MethodsX *3*, 1-7.

Lestrade, L., and Weber, M.J. (2006). snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. Nucleic Acids Res *34*, D158-162.

Leulliot, N., Bohnsack, M.T., Graille, M., Tollervey, D., and Van Tilbeurgh, H. (2008). The yeast ribosome synthesis factor Emg1 is a novel member of the superfamily of alpha/beta knot fold methyltransferases. Nucleic Acids Res *36*, 629-639.

Li, F., Zhao, D., Wu, J., and Shi, Y. (2014). Structure of the YTH domain of human YTHDF2 in complex with an m(6)A mononucleotide reveals an aromatic cage for m(6)A recognition. Cell Res *24*, 1490-1492.

Li, X., Xiong, X., Wang, K., Wang, L., Shu, X., Ma, S., and Yi, C. (2016). Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. Nat Chem Biol *12*, 311-316.

Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., and Yi, C. (2015a). Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol *11*, 592-597.

Li, Y., Wang, Y., Zhang, Z., Zamudio, A.V., and Zhao, J.C. (2015b). Genome-wide detection of high abundance N6-methyladenosine sites by microarray. RNA *21*, 1511-1518.

Liang, X.H., Liu, Q., and Fournier, M.J. (2007). rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. Mol Cell *28*, 965-977.

Liang, X.H., Liu, Q., and Fournier, M.J. (2009). Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. RNA *15*, 1716-1728.

Lin, S., Choe, J., Du, P., Triboulet, R., and Gregory, R.I. (2016). The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. Mol Cell *62*, 335-345.

Linder, B., Grozhik, A.V., Olarerin-George, A.O., Meydan, C., Mason, C.E., and Jaffrey, S.R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods *12*, 767-772.

Lipowsky, G., Bischoff, F.R., Izaurralde, E., Kutay, U., Schafer, S., Gross, H.J., Beier, H., and Gorlich, D. (1999). Coordination of tRNA nuclear export with processing of tRNA. RNA *5*, 539-549.

Liu, J., and Jia, G. (2014). Methylation modifications in eukaryotic messenger RNA. J Genet Genomics *41*, 21-33.

Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., Dai, Q., Chen, W., and He, C. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol *10*, 93-95.

Liu, K., Ding, Y., Ye, W., Liu, Y., Yang, J., Liu, J., and Qi, C. (2016). Structural and Functional Characterization of the Proteins Responsible for N(6)-Methyladenosine Modification and Recognition. Curr Protein Pept Sci *17*, 306-318.

Liu, N., Parisien, M., Dai, Q., Zheng, G., He, C., and Pan, T. (2013). Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA *19*, 1848-1856.

Liu, P.C., and Thiele, D.J. (2001). Novel stress-responsive genes EMG1 and NOP14 encode conserved, interacting proteins required for 40S ribosome biogenesis. Mol Biol Cell *12*, 3644-3657.

Lovejoy, A.F., Riordan, D.P., and Brown, P.O. (2014). Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in S. cerevisiae. PloS one *9*, e110799.

Lund, E., and Dahlberg, J.E. (1998). Proofreading and aminoacylation of tRNAs before export from the nucleus. Science *282*, 2082-2085.

Luo, S., and Tong, L. (2014). Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain. Proc Natl Acad Sci U S A *111*, 13834-13839.

Machnicka, M.A., Milanowska, K., Osman Oglou, O., Purta, E., Kurkowska, M., Olchowik, A., Januszewski, W., Kalinowski, S., Dunin-Horkawicz, S., Rother, K.M., Helm, M., Bujnicki, J.M., and Grosjean, H. (2013). MODOMICS: a database of RNA modification pathways--2013 update. Nucleic Acids Res *41*, D262-267.

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

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. Nat Struct Mol Biol *19*, 560-567.

Melton, D.A., De Robertis, E.M., and Cortese, R. (1980). Order and intracellular location of the events involved in the maturation of a spliced tRNA. Nature *284*, 143-148.

Meyer, B., Wurm, J.P., Sharma, S., Immer, C., Pogoryelov, D., Kotter, P., Lafontaine, D.L., Wohnert, J., and Entian, K.D. (2016). Ribosome biogenesis factor Tsr3 is the aminocarboxypropyl transferase responsible for 18S rRNA hypermodification in yeast and humans. Nucleic Acids Res *44*, 4304-4316.

Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.B., and Jaffrey, S.R. (2015). 5' UTR m(6)A Promotes Cap-Independent Translation. Cell *163*, 999-1010.

Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell *149*, 1635-1646.

Morohashi, K., Sahara, H., Watashi, K., Iwabata, K., Sunoki, T., Kuramochi, K., Takakusagi, K., Miyashita, H., Sato, N., Tanabe, A., Shimotohno, K., Kobayashi, S., Sakaguchi, K., and Sugawara, F. (2011). Cyclosporin A associated helicase-like protein facilitates the association of hepatitis C virus RNA polymerase with its cellular cyclophilin B. PloS one *6*, e18285.

Motorin, Y., and Helm, M. (2010). tRNA stabilization by modified nucleotides. Biochemistry *49*, 4934-4944.

Motorin, Y., Lyko, F., and Helm, M. (2010). 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. Nucleic Acids Res *38*, 1415-1430.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. Science *320*, 1344-1349.

Nawrot, B., Sochacka, E., and Duchler, M. (2011). tRNA structural and functional changes induced by oxidative stress. Cell Mol Life Sci *68*, 4023-4032.

Ozgur, S., Chekulaeva, M., and Stoecklin, G. (2010). Human Pat1b connects deadenylation with mRNA decapping and controls the assembly of processing bodies. Mol Cell Biol *30*, 4308-4323.

Passos, D.O., Doma, M.K., Shoemaker, C.J., Muhlrad, D., Green, R., Weissman, J., Hollien, J., and Parker, R. (2009). Analysis of Dom34 and its function in no-go decay. Mol Biol Cell *20*, 3025-3032.

Peifer, C., Sharma, S., Watzinger, P., Lamberth, S., Kotter, P., and Entian, K.D. (2013). Yeast Rrp8p, a novel methyltransferase responsible for m1A 645 base modification of 25S rRNA. Nucleic Acids Res *41*, 1151-1163.

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. (2014). Secondary structures of rRNAs from all three domains of life. PloS one *9*, e88222.

Phizicky, E.M., and Hopper, A.K. (2010). tRNA biology charges to the front. Genes Dev *24*, 1832-1860.

Piekna-Przybylska, D., Decatur, W.A., and Fournier, M.J. (2008a). The 3D rRNA modification maps database: with interactive tools for ribosome analysis. Nucleic Acids Res *36*, D178-183.

Piekna-Przybylska, D., Przybylski, P., Baudin-Baillieu, A., Rousset, J.P., and Fournier, M.J. (2008b). Ribosome performance is enhanced by a rich cluster of pseudouridines in the A-site finger region of the large subunit. J Biol Chem *283*, 26026-26036.

Ping, X.L., Sun, B.F., Wang, L., Xiao, W., Yang, X., Wang, W.J., Adhikari, S., Shi, Y., Lv, Y., Chen, Y.S., Zhao, X., Li, A., Yang, Y., Dahal, U., Lou, X.M., Liu, X., Huang, J., Yuan, W.P., Zhu, X.F., Cheng, T., Zhao, Y.L., Wang, X., Rendtlew Danielsen, J.M., Liu, F., and Yang, Y.G. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res *24*, 177-189.

Poole, T.L., and Stevens, A. (1997). Structural modifications of RNA influence the 5' exoribonucleolytic hydrolysis by XRN1 and HKE1 of Saccharomyces cerevisiae. Biochem Biophys Res Commun *235*, 799-805.

Popova, A.M., and Williamson, J.R. (2014). Quantitative analysis of rRNA modifications using stable isotope labeling and mass spectrometry. J Am Chem Soc *136*, 2058-2069.

Powell, L.M., Wallis, S.C., Pease, R.J., Edwards, Y.H., Knott, T.J., and Scott, J. (1987). A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. Cell *50*, 831-840.

Prusiner, P., Yathindra, N., and Sundaralingam, M. (1974). Effect of ribose O(2')-methylation on the conformation of nucleosides and nucleotides. Biochim Biophys Acta *366*, 115-123.

Ramanathan, A., Robb, G.B., and Chan, S.H. (2016). mRNA capping: biological functions and applications. Nucleic Acids Res *44*, 7511-7526.

Ranjan, N., and Rodnina, M.V. (2016). tRNA wobble modifications and protein homeostasis. Translation (Austin) *4*, e1143076.

Roost, C., Lynch, S.R., Batista, P.J., Qu, K., Chang, H.Y., and Kool, E.T. (2015). Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification. J Am Chem Soc *137*, 2107-2115.

Rottman, F., Shatkin, A.J., and Perry, R.P. (1974). Sequences containing methylated nucleotides at the 5' termini of messenger RNAs: possible implications for processing. Cell *3*, 197-199.

Sarkar, S., and Hopper, A.K. (1998). tRNA nuclear export in saccharomyces cerevisiae: in situ hybridization analysis. Mol Biol Cell *9*, 3041-3055.

Schibler, U., Kelley, D.E., and Perry, R.P. (1977). Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. J Mol Biol *115*, 695-714.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc *3*, 1101-1108.

Schneider, C., and Tollervey, D. (2013). Threading the barrel of the RNA exosome. Trends Biochem Sci *38*, 485-493.

Schosserer, M., Minois, N., Angerer, T.B., Amring, M., Dellago, H., Harreither, E., Calle-Perez, A., Pircher, A., Gerstl, M.P., Pfeifenberger, S., Brandl, C., Sonntagbauer, M., Kriegner, A., Linder, A., Weinhausel, A., Mohr, T., Steiger, M., Mattanovich, D., Rinnerthaler, M., Karl, T., Sharma, S., Entian, K.D., Kos, M., Breitenbach, M., Wilson, I.B., Polacek, N., Grillari-Voglauer, R., Breitenbach-Koller, L., and Grillari, J. (2015). Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan. Nat Commun *6*, 6158.

Schwartz, S., Agarwala, S.D., Mumbach, M.R., Jovanovic, M., Mertins, P., Shishkin, A., Tabach, Y., Mikkelsen, T.S., Satija, R., Ruvkun, G., Carr, S.A., Lander, E.S., Fink, G.R., and Regev, A. (2013). High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. Cell *155*, 1409-1421.

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

Schwartz, S., Mumbach, M.R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G.G., Mertins, P., Ter-Ovanesyan, D., Habib, N., Cacchiarelli, D., Sanjana, N.E., Freinkman, E., Pacold, M.E., Satija, R., Mikkelsen, T.S., Hacohen, N., Zhang, F., Carr, S.A., Lander, E.S., and Regev, A. (2014b). Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. Cell Rep *8*, 284-296.

Sedgwick, S.G., and Smerdon, S.J. (1999). The ankyrin repeat: a diversity of interactions on a common structural framework. Trends Biochem Sci *24*, 311-316.

Sharma, S., and Lafontaine, D.L. (2015). 'View From A Bridge': A New Perspective on Eukaryotic rRNA Base Modification. Trends Biochem Sci *40*, 560-575.

Sharma, S., Langhendries, J.L., Watzinger, P., Kotter, P., Entian, K.D., and Lafontaine, D.L. (2015). Yeast Kre33 and human NAT10 are conserved 18S rRNA cytosine acetyltransferases that modify tRNAs assisted by the adaptor Tan1/THUMPD1. Nucleic Acids Res *43*, 2242-2258.

Sharma, S., Watzinger, P., Kotter, P., and Entian, K.D. (2013a). Identification of a novel methyltransferase, Bmt2, responsible for the N-1-methyl-adenosine base modification of 25S rRNA in Saccharomyces cerevisiae. Nucleic Acids Res *41*, 5428-5443.

Sharma, S., Yang, J., Duttmann, S., Watzinger, P., Kotter, P., and Entian, K.D. (2014). Identification of novel methyltransferases, Bmt5 and Bmt6, responsible for the m3U methylations of 25S rRNA in Saccharomyces cerevisiae. Nucleic Acids Res *42*, 3246-3260.

Sharma, S., Yang, J., Watzinger, P., Kotter, P., and Entian, K.D. (2013b). Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively. Nucleic Acids Res *41*, 9062-9076.

Shi, H., and Moore, P.B. (2000). The crystal structure of yeast phenylalanine tRNA at 1.93 A resolution: a classic structure revisited. RNA *6*, 1091-1105.

Shoemaker, C.J., Eyler, D.E., and Green, R. (2010). Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. Science *330*, 369-372.

Simonovic, M., and Steitz, T.A. (2009). A structural view on the mechanism of the ribosome-catalyzed peptide bond formation. Biochim Biophys Acta *1789*, 612-623.

Sloan, K.E., Gleizes, P.E., and Bohnsack, M.T. (2016). Nucleocytoplasmic Transport of RNAs and RNA-Protein Complexes. J Mol Biol *428*, 2040-2059.

Sloan, K.E., Leisegang, M.S., Doebele, C., Ramirez, A.S., Simm, S., Safferthal, C., Kretschmer, J., Schorge, T., Markoutsa, S., Haag, S., Karas, M., Ebersberger, I., Schleiff, E., Watkins, N.J., and Bohnsack, M.T. (2015). The association of late-acting snoRNPs with human pre-ribosomal complexes requires the RNA helicase DDX21. Nucleic Acids Res *43*, 553-564.

Sloan, K.E., Mattijssen, S., Lebaron, S., Tollervey, D., Pruijn, G.J., and Watkins, N.J. (2013). Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. J Cell Biol *200*, 577-588.

Soshnev, A.A., Josefowicz, S.Z., and Allis, C.D. (2016). Greater Than the Sum of Parts: Complexity of the Dynamic Epigenome. Mol Cell *62*, 681-694.

Squires, J.E., Patel, H.R., Nousch, M., Sibbritt, T., Humphreys, D.T., Parker, B.J., Suter, C.M., and Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res *40*, 5023-5033.

Steiner-Mosonyi, M., and Mangroo, D. (2004). The nuclear tRNA aminoacylation-dependent pathway may be the principal route used to export tRNA from the nucleus in Saccharomyces cerevisiae. Biochem J *378*, 809-816.

Stoilov, P., Rafalska, I., and Stamm, S. (2002). YTH: a new domain in nuclear proteins. Trends Biochem Sci *27*, 495-497.

Tanabe, A., Tanikawa, K., Tsunetomi, M., Takai, K., Ikeda, H., Konno, J., Torigoe, T., Maeda, H., Kutomi, G., Okita, K., Mori, M., and Sahara, H. (2016). RNA helicase YTHDC2 promotes cancer metastasis via the enhancement of the efficiency by which HIF-1alpha mRNA is translated. Cancer Lett *376*, 34-42.

Taoka, M., Nobe, Y., Yamaki, Y., Yamauchi, Y., Ishikawa, H., Takahashi, N., Nakayama, H., and Isobe, T. (2016). The complete chemical structure of Saccharomyces cerevisiae rRNA: partial pseudouridylation of U2345 in 25S rRNA by snoRNA snR9. Nucleic Acids Res *44*, 8951-8961.

Thalhammer, A., Bencokova, Z., Poole, R., Loenarz, C., Adam, J., O'Flaherty, L., Schodel, J., Mole, D., Giaslakiotis, K., Schofield, C.J., Hammond, E.M., Ratcliffe, P.J., and Pollard, P.J. (2011). Human AlkB homologue 5 is a nuclear 2-oxoglutarate dependent oxygenase and a direct target of hypoxia-inducible factor 1alpha (HIF-1alpha). PloS one *6*, e16210.

Theler, D., Dominguez, C., Blatter, M., Boudet, J., and Allain, F.H. (2014). Solution structure of the YTH domain in complex with N6-methyladenosine RNA: a reader of methylated RNA. Nucleic Acids Res *42*, 13911-13919.

Thomson, E., Ferreira-Cerca, S., and Hurt, E. (2013). Eukaryotic ribosome biogenesis at a glance. J Cell Sci *126*, 4815-4821.

Tirumuru, N., Zhao, B.S., Lu, W., Lu, Z., He, C., and Wu, L. (2016). N(6)-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. Elife *5*.

Tkaczuk, K.L., Dunin-Horkawicz, S., Purta, E., and Bujnicki, J.M. (2007). Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases. BMC Bioinformatics *8*, 73.

Torres, A.G., Batlle, E., and Ribas de Pouplana, L. (2014). Role of tRNA modifications in human diseases. Trends Mol Med *20*, 306-314.

Tschochner, H., and Hurt, E. (2003). Pre-ribosomes on the road from the nucleolus to the cytoplasm. Trends Cell Biol *13*, 255-263.

Tserovski, L., Marchand, V., Hauenschild, R., Blanloeil-Oillo, F., Helm, M., and Motorin, Y. (2016). High-throughput sequencing for 1-methyladenosine (m(1)A) mapping in RNA. Methods *107*, 110-121.

UniProt, C. (2015). UniProt: a hub for protein information. Nucleic Acids Res *43*, D204-212.

Wang, C., Zhu, Y., Bao, H., Jiang, Y., Xu, C., Wu, J., and Shi, Y. (2016a). A novel RNA-binding mode of the YTH domain reveals the mechanism for recognition of determinant of selective removal by Mmi1. Nucleic Acids Res *44*, 969-982.

Wang, P., Doxtader, K.A., and Nam, Y. (2016b). Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. Mol Cell *63*, 306-317.

Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., Gong, Z., Wang, Q., Huang, J., Tang, C., Zou, T., and Yin, P. (2016c). Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. Nature *534*, 575-578.

Wang, X., and He, C. (2014). Reading RNA methylation codes through methyl-specific binding proteins. RNA Biol *11*, 669-672.

Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., Ren, B., Pan, T., and He, C. (2014a). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature *505*, 117-120.

Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015). N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell *161*, 1388-1399.

Wang, Y., Li, Y., Toth, J.I., Petroski, M.D., Zhang, Z., and Zhao, J.C. (2014b). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. Nat Cell Biol *16*, 191-198.

Warner, J.R. (1999). The economics of ribosome biosynthesis in yeast. Trends Biochem Sci *24*, 437-440.

Watkins, N.J., and Bohnsack, M.T. (2012). The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. Wiley Interdiscip Rev RNA *3*, 397-414.

Webb, S., Hector, R.D., Kudla, G., and Granneman, S. (2014). PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. Genome Biol *15*, R8.

Wei, C.M., Gershowitz, A., and Moss, B. (1975). Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. Cell *4*, 379-386.

Wei, C.M., Gershowitz, A., and Moss, B. (1976). 5'-Terminal and internal methylated nucleotide sequences in HeLa cell mRNA. Biochemistry *15*, 397-401.

Wei, C.M., and Moss, B. (1977). Nucleotide sequences at the N6-methyladenosine sites of HeLa cell messenger ribonucleic acid. Biochemistry *16*, 1672-1676.

White, J., Li, Z., Sardana, R., Bujnicki, J.M., Marcotte, E.M., and Johnson, A.W. (2008). Bud23 methylates G1575 of 18S rRNA and is required for efficient nuclear export of pre-40S subunits. Mol Cell Biol *28*, 3151-3161.

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.

Worton, R.G., Sutherland, J., Sylvester, J.E., Willard, H.F., Bodrug, S., Dube, I., Duff, C., Kean, V., Ray, P.N., and Schmickel, R.D. (1988). Human ribosomal RNA genes: orientation of the tandem array and conservation of the 5' end. Science *239*, 64-68.

Wurm, J.P., Meyer, B., Bahr, U., Held, M., Frolow, O., Kotter, P., Engels, J.W., Heckel, A., Karas, M., Entian, K.D., and Wohnert, J. (2010). The ribosome assembly factor Nep1 responsible for Bowen-Conradi syndrome is a pseudouridine-N1-specific methyltransferase. Nucleic Acids Res *38*, 2387-2398.

Xiao, W., Adhikari, S., Dahal, U., Chen, Y.S., Hao, Y.J., Sun, B.F., Sun, H.Y., Li, A., Ping, X.L., Lai, W.Y., Wang, X., Ma, H.L., Huang, C.M., Yang, Y., Huang, N., Jiang, G.B., Wang, H.L., Zhou, Q., Wang, X.J., Zhao, Y.L., and Yang, Y.G. (2016). Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol Cell *61*, 507-519.

Xu, C., Liu, K., Ahmed, H., Loppnau, P., Schapira, M., and Min, J. (2015). Structural Basis for the Discriminative Recognition of N6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. J Biol Chem *290*, 24902-24913.

Xu, C., Wang, X., Liu, K., Roundtree, I.A., Tempel, W., Li, Y., Lu, Z., He, C., and Min, J. (2014). Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. Nat Chem Biol *10*, 927-929.

Yeung, M.L., Bennasser, Y., Watashi, K., Le, S.Y., Houzet, L., and Jeang, K.T. (2009). Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence for the processing of a viral-cellular double-stranded RNA hybrid. Nucleic Acids Res *37*, 6575-6586.

Yoshihisa, T., Yunoki-Esaki, K., Ohshima, C., Tanaka, N., and Endo, T. (2003). Possibility of cytoplasmic pre-tRNA splicing: the yeast tRNA splicing endonuclease mainly localizes on the mitochondria. Mol Biol Cell *14*, 3266-3279.

Zaringhalam, M., and Papavasiliou, F.N. (2016). Pseudouridylation meets next-generation sequencing. Methods *107*, 63-72.

Zhang, C., Zhi, W.I., Lu, H., Samanta, D., Chen, I., Gabrielson, E., and Semenza, G.L. (2016). Hypoxia-inducible factors regulate pluripotency factor expression by ZNF217- and ALKBH5-mediated modulation of RNA methylation in breast cancer cells. Oncotarget.

Zhang, Z., Theler, D., Kaminska, K.H., Hiller, M., de la Grange, P., Pudimat, R., Rafalska, I., Heinrich, B., Bujnicki, J.M., Allain, F.H., and Stamm, S. (2010). The YTH domain is a novel RNA binding domain. J Biol Chem *285*, 14701-14710.

Zhao, X., Yang, Y., Sun, B.F., Shi, Y., Yang, X., Xiao, W., Hao, Y.J., Ping, X.L., Chen, Y.S., Wang, W.J., Jin, K.X., Wang, X., Huang, C.M., Fu, Y., Ge, X.M., Song, S.H., Jeong, H.S., Yanagisawa, H., Niu, Y., Jia, G.F., Wu, W., Tong, W.M., Okamoto, A., He, C., Rendtlew Danielsen, J.M., Wang, X.J., and Yang, Y.G. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res *24*, 1403-1419.

Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.M., Li, C.J., Vagbo, C.B., Shi, Y., Wang, W.L., Song, S.H., Lu, Z., Bosmans, R.P., Dai, Q., Hao, Y.J., Yang, X., Zhao, W.M., Tong, W.M., Wang, X.J., Bogdan, F., Furu, K., Fu, Y., Jia, G., Zhao, X., Liu, J., Krokan, H.E., Klungland, A., Yang, Y.G., and He, C. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell *49*, 18-29.

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., and Qian, S.B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. Nature *526*, 591-594.

Zhou, K.I., Parisien, M., Dai, Q., Liu, N., Diatchenko, L., Sachleben, J.R., and Pan, T. (2016). N(6)-Methyladenosine Modification in a Long Noncoding RNA Hairpin Predisposes Its Conformation to Protein Binding. J Mol Biol *428*, 822-833.

Zhu, T., Roundtree, I.A., Wang, P., Wang, X., Wang, L., Sun, C., Tian, Y., Li, J., He, C., and Xu, Y. (2014). Crystal structure of the YTH domain of YTHDF2 reveals mechanism for recognition of N6-methyladenosine. Cell Res *24*, 1493-1496.

Zorbas, C., Nicolas, E., Wacheul, L., Huvelle, E., Heurgue-Hamard, V., and Lafontaine, D.L. (2015). The human 18S rRNA base methyltransferases DIMT1L and WBSCR22-TRMT112 but not rRNA modification are required for ribosome biogenesis. Mol Biol Cell *26*, 2080-2095.

Zou, S., Toh, J.D., Wong, K.H., Gao, Y.G., Hong, W., and Woon, E.C. (2016). N(6)-Methyladenosine: a conformational marker that regulates the substrate specificity of human demethylases FTO and ALKBH5. Sci Rep *6*, 25677.

Publications associated with this thesis

Haag, S., Kretschmer, J., Sloan, K.E., and Bohnsack, M.T. Crosslinking methods to identify RNA methyltransferase targets in vivo. Methods Mol Biol, *accepted manuscript*.

Haag, S., Sloan, K.E., Ranjan, N., Warda, A.S., Kretschmer, J., Blessing, C., Hubner, B., Seikowski, J., Dennerlein, S., Rehling, P., Rodnina, M.V., Hobartner, C., and Bohnsack, M.T. (2016). NSUN3 and ABH1 modify the wobble position of mt-tRNAMet to expand codon recognition in mitochondrial translation. EMBO J *35*, 2104-2119.

Haag, S., Kretschmer, J., and Bohnsack, M.T. (2015). WBSCR22/Merm1 is required for late nuclear pre-ribosomal RNA processing and mediates N7-methylation of G1639 in human 18S rRNA. RNA *21*, 180-187.

Haag, S., Warda, A.S., Kretschmer, J., Gunnigmann, M.A., Hobartner, C., and Bohnsack, M.T. (2015). NSUN6 is a human RNA methyltransferase that catalyzes formation of m5C72 in specific tRNAs. RNA *21*, 1532-1543.

Sloan, K.E., Leisegang, M.S., Doebele, C., Ramirez, A.S., Simm, S., Safferthal, C., Kretschmer, J., Schorge, T., Markoutsa, S., Haag, S., Karas, M., Ebersberger, I., Schleiff, E., Watkins, N.J., and Bohnsack, M.T. (2015). The association of late-acting snoRNPs with human pre-ribosomal complexes requires the RNA helicase DDX21. Nucleic Acids Res *43*, 553-564.

<u>Acknowledgements</u>

First of all, I would like to thank my supervisor Prof. Dr. Markus T. Bohnsack for the opportunity to work on such an interesting, state-of-the-art project that allowed me to extend my skillset considerably. Furthermore, I would like to express my gratitude for his excellent supervision, discussions and guidance throughout this challenging project, as well as helpful comments on this thesis.

I am very thankful for the support and constructive discussions with the members of my thesis committee, Prof. Dr. Peter Rehling and Dr. Heinz Neumann. In addition, I would also like to thank Prof. Dr. Claudia Höbartner, Prof. Dr. Jörg Stülke and Prof. Dr. Ralf Ficner for being a part of my examination board.

Very special thanks go to Dr. Katherine Sloan for the constant support and helpful discussions during my PhD. I would also like to thank her for her useful suggestions and comments on this thesis.

I thank Prof. Dr. Henning Urlaub for mass spectrometry and support regarding analysis of the data.

I would also like to thank the head of the Institute for Molecular Biology Prof. Dr. Blanche Schwappach-Pignataro and the other colleagues in the institute for creating a pleasant and productive working environment.

In this context, I would also like to acknowledge the current and former members of the Bohnsack Lab, especially Dr. Sara Haag, Dr. Carmen Döbele, Dr. Annika Heininger, Dr. Roman Martin, Phillip Hackert, Lukas Brüning, and Ahmed Warda, Indira Memet Priyanka Choudhury and Jimmena Dávila Gallesio, who created a nice working environment and who were always open for discussions.

I especially thank Lena and Katja for the awesome climbing fun we have together - it just Roxx.

Last but not least I like to thank my family for the enduring support during my studies.