Dissertation

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the
Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by
Dipl. Biochemist Birgit Rogell
born in Speyer (Germany)
Oral-examination: 2nd November 2016

Exploring the biology of RNPs: specific capture of RNPs using antisense locked nucleic acids

Referees: 1. Dr. Anne Ephrussi
2. Prof. Dr. Ralf Bartenschlager

Abstract

RNA-binding proteins (RBPs) are central players in cell biology and respond to a multitude of cellular cues and environmental stimuli. Identification of RBPs associated with specific transcripts in a cell is a challenging task; and the available strategies to purify specific transcripts and their bound proteome face numerous limitations. Thus, methods to determine the composition of proteins on a given RNA are required to further understand the regulation and biological function of any given RNA.

Therefore, the focus of my PhD project was to develop a highly specific and selective method, “specific Ribonucleoprotein (RNP) capture”, to isolate a specific RNA species together with its bound proteome. Following irradiation with ultraviolet (UV) light that creates a covalent bond between RNA and protein, RNAs are captured using short LNA (locked nucleic acid)/DNA mixmer antisense probes coupled to a solid support.The proteins covalently linked to the isolated RNA are then identified by quantitative mass spectrometry. First, I successfully established the method for this application *in vitro.* Mass spectrometry data revealed that the protein Sister of Sex lethal (Ssx) has similar binding preferences to a mRNA derived from male-specific lethal (msl)2 mRNA as its paralog Sxl in *Drosophila melanogaster* embryo extracts*.* This demonstrated the specificity and selectivity of the method and provided direct experimental evidence for Ssx-RNA binding.

Following these experiments, I extended the protocol to in-cell applications, focusing on the 18S and 28S ribosomal RNA (rRNA) of HeLa cells. Compared to bacteria, eukaryotic rRNAs possess “expansion segments” with much to be learnt about their bound RBPs and function. The high specificity of the method allowed me to generate distinct proteomic datasets for these two rRNAs. The method’s excellent biochemical performance is reflected by the overlap of these datasets with previous literature information on the cytoplasmic ribosome and system-wide screens of ribosomal biogenesis. Notably, my data revealed a strong connection between heterogeneous ribonucleoproteins (HNRNPs) and ribosomal RNA biogenesis, which is an unexplored area of research. In

summary, “specific RNP capture” allows identification of a given RNAs proteome and can be applied to both *in vitro* and cultured cells systems.

Zusammenfassung

Erforschung der Biologie von Ribonukleoprotein Komplexen: spezifische Aufreinigung von Ribonukleoproteinkomplexen mit komplementären modifizierten Oligonukleotiden (Locked Nucleic Acids)

RNA-Bindeproteine (RBPs) sind wichtige und zentrale Moleküle in der Zellbiologie, da sie dynamisch auf verschiedene Stimuli und zelluläre Stresssituation reagieren können. Die Aufreinigung von einem spezifischen Transkript mit den gebunden RBPs unterliegt vielerlei Limitierungen, sodass zu Beginn meiner Doktorarbeit ein Bedarf an Methoden bestand, um neue biologische Fragestellen zu einer spezifischen RNA zu beantworten. Der Fokus dieser Doktorarbeit lag daher in der Entwicklung einer spezifischen und selektiven Aufreinigungsstrategie von Ribonukleoproteinen („specific Ribonuleoprotein capture“, „specific RNP capture“).
Diese neu entwickelte Methode erlaubt eine Aufreinigung einer spezifischen RNA zusammen mit den gebundenen Protein-Interaktionspartnern. Direkte RNA-Protein Interaktionen werden durch Bestrahlung mit ultravioletten Licht fixiert. Die RNA mit den direkt daran gebunden Proteinen wird dann mit Hilfe von kurzen komplementären Oligonukleotidsequenzen, welche kovalent an magnetische Partikel gekoppelt sind und aus einer Mixtur von Locked Nucleic Acids (LNA) und DNA bestehen, herausgezogen. Proteine, welche an der spezifischen RNA gebunden sind, können schließlich über quantitative Massenspektrometrie analysiert und identifiziert werden.
Zunächst habe ich „specific RNP capture“ für die *in vitro* Anwendung etabliert. Mit der massenspektrometrischen Analyse konnte ich dabei aufdecken, dass das Protein Sister of Sex lethal (Ssx) in *Drosophila melanogaster* Embryo Extrakten ähnliche Sequenzbindungspräferenzen, kommend aus der mRNA von male-specific lethal (msl)2 mRNA, wie dessen Paralog Sxl besitzt. Mit den Ergebnissen dieses Experiments konnte ich die Spezifität und Selektivität der Methode demonstrieren und zugleich experimentellen Beweis für die Ssx-RNA Bindung liefern.

Danach erweiterte ich die Anwendung der Methode auf die Aufreinigung von spezifischen endogenen RNA Spezies in humanen Zellen. Ich fokussierte mich

auf die spezifische Aufreinigung der 18S und 28S ribosomalen RNA (rRNA) von HeLa Zellen. In Vergleich zu Bakterien besitzt eukaryotische rRNA sogenannte „Expansion Segments“ (interne Einlagerung und Evolution von Nukleotiden in die konservierte rRNA Sequenz). Da noch wenig über diese „Expansion Segments“ und deren gebundenen RNA-Bindeproteine bekannt ist, ist eine generelle Studie Bindepartner von rRNA interessant und ermöglicht eine zukünftige Erweiterung der Methode auf Region-spezifische Anreicherung von den rRNA „Expansion Segments“ und deren gebunden Proteinen. Die spezifische Isolation der 18S und 28S rRNA erlaubte mir die Generierung von spezifischen Proteindatensätzen der jeweiligen rRNA. Die exzellente biochemische Leistung der Methode zeigte sich insbesondere in der Re-Identifizierung von Proteinen des zytoplasmatischen Ribosoms und Proteinen, welche in der Biogenese des Ribosoms involviert sind. Darüber hinaus konnte ich neue Proteine identifizieren, welche eine mögliche Rolle in der ribosomalen Funktion oder Biogenese haben könnten. Beachtenswert war die Anreicherung von HNRNP Proteinen (heterogenous ribonucleoproteins) als größere Gruppierung in den rRNA Interaktionsdatensätzen. HNRNP Proteine wurden als RNA Bindeproteine bisher hauptsächlich mit der Reifung von mRNA in Verbindung gebracht.

Zusammengefasst beschreibt meine Arbeit die Etablierung der neuen Methode „specific RNP capture“ (spezifische Aufreinigung von Ribonukleoprotein Komplexen), welche eine Aufreinigung von einer spezifischen RNA Spezies zusammen mit den direkt gebundenen Proteinen erlaubt. Ich konnte zeigen, dass diese Methode sowohl für den *in vitro* Gebrauch als auch für den Gebrauch in Zellen geeignet ist. In den Experimenten in humanen Zellen konnte ich neben bekannten rRNA Interaktionspartnern, Proteine identifizieren, welche eine mögliche neue Funktion in der Biologie der rRNA oder des Ribosoms einnehmen könnten.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 The discovery of RNA as a central player in Biology

In the 1950s, after the discovery of the DNA structure (Watson & Crick, 1953), the interest for RNA and its connection to DNA emerged. The central dogma, the first time postulated by Francis Crick in 1956 (Crick, 1958), states that DNA is transcribed into RNA which is translated into proteins. RNA was in those days rather seen as a supporting player of DNA, which mediates the information transfer from DNA. Therefore, the research focus was limited to the three types of RNA involved in protein synthesis: the mRNA, the transfer RNA (tRNA) and the ribosomal RNA (rRNA). During these early days, RNA was not an easy object of research. Its short-lived and unstable nature and the complexity and variety of mRNAs represented a challenge for RNA biologists. First, the so called “microsomes” could be isolated (Claude, 1944). These “microsomes” turned out later to be the ribosomes, the cellular machinery of protein translation, which associate in multiple copies to mRNA forming polysomes (Slayter et al., 1963). The discovery of the first phages with a RNA genome (f2/MS2/R17) (Loeb & Zinder, 1961; Davis et al., 1963; Parenchynch & Graham, 1962) exemplified that RNA can exhibit multitude of functions, even independent of DNA. With every new finding of novel functions of RNA and novel RNA types, it became clearer that the RNA biology is more complex than previously anticipated and RNA was not just a mediator of the genetic information, but a central player in cell biology.

1.2 The functional diversity of RNA

Discovery of many different RNA types (Figure 1.1) spurred further experimentation and resulted in the establishment of RNA biology as a research field of its own.

After the discovery that the ribosome consists of ribosomal RNA and proteins, the highly structured transfer RNA (tRNA) (Hoagland et al., 1958), with its cloverleaf like structure, was identified as the molecule needed to bring the right amino acids to the ribosome. The mRNA on the other hand was described to be the central molecule for the translation of the genetic code into a protein. The

mRNA molecule is capped at the 5'end and polyadenylated at the 3' end. These modifications allow recognition by translation initiation factors and are determinants of mRNA stability.

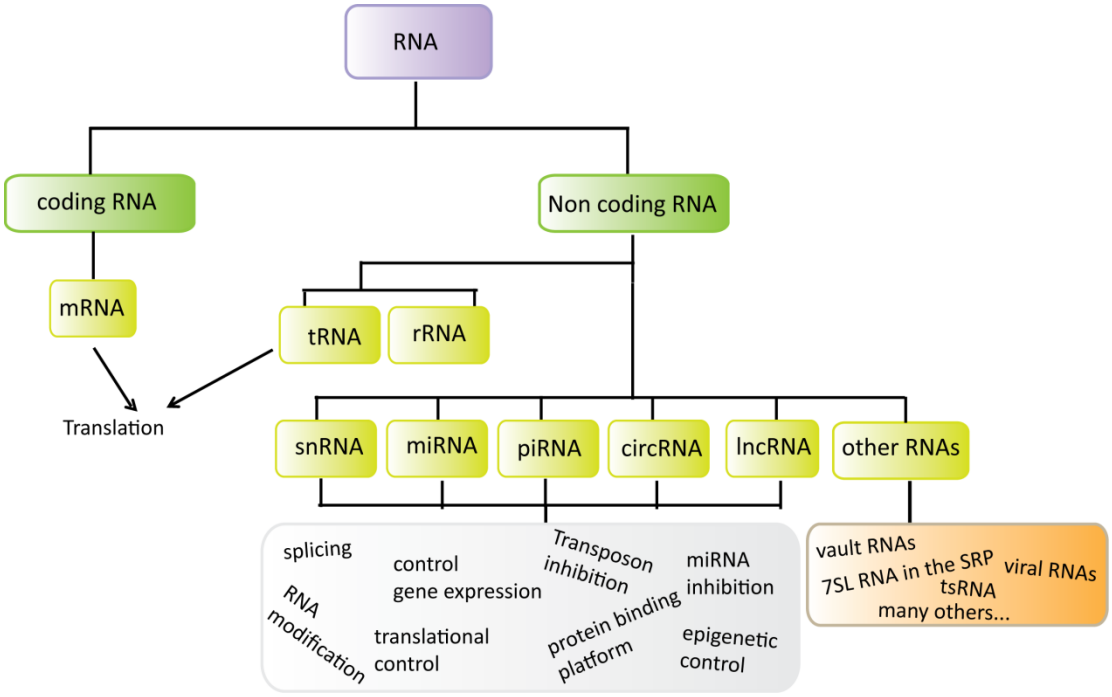

Figure 1.1; The diversity of the RNA world. Besides the coding mRNA there are many types of RNA that are non-coding. Just some examples of the multitude of existing RNA types are displayed here and explained in the text. Other RNAs include e.g. vault RNAs (Stadler et al., 2009) from the vault particle, the 7SL RNA (Pool., 2004) in the signal recognition particle (SRP), viral RNAs (Poltronieri et al., 2015) and tRNA derived small RNAs (tsRNAs) (Haussecker et al., 2010). The non-coding tRNA and rRNA are involved in the translation process. rRNA catalyzes the peptidyl-transferase reaction within the ribosome during translation; The various non-coding RNAs in the cell serve a multitude of different functions within the cell.

After transcription, the mRNA contains intron and exons and is referred to as pre-mRNA. Upon intron removal, alternative exon usage allows variations of the primary mRNA sequence expanding the potential protein sequence outcomes that an mRNA can encode. The responsible process for the variation of the mRNA pool is called splicing. With the discovery of small nuclear RNAs (snRNAs) (Lerner & Steitz, 1979), a novel function of RNA in splicing was described. Deep sequencing revealed that 90% of all transcription units are spliced in more than one pattern (Pan et al., 2008). Different proteins can be translated from these differentially spliced RNAs, including non-functional products (Delaney et al., 1993). Thereby, splicing allows an additional level of regulation in a cell.

The discovery of RNA interference (RNAi) (Ecker & Davis, 1986; Napoli et al., 1990; Van Blockland et al., 1994; Fire et al., 1998) and microRNAs (miRNAs)

(Lee et al., 1993; Wightman et al., 1993) demonstrated that another type of small RNAs can have regulatory function in mRNA fate. Small RNAs can silence gene expression by interacting with a target mRNA and either trigger its degradation or regulating its translation (Jonas & Izaurralde, 2015). Another type of small RNAs, the piwi interacting RNAs (piRNAs), are found in the germline of many animals and are the adaptive defense against transposons in the genome. This involves silencing mechanisms that act on multiple levels, including epigenetic control of the mobile elements (Iwasaki et al., 2015).

Advances in RNA sequencing technologies lead to the discovery of other new types of RNA species, which were initially considered to be rather transcriptional noise or sequencing artefacts. This included the long non coding RNAs (lncRNAs) (Kapranov et al., 2007) and circular RNAs (circRNAs) (Salzman et al., 2012). However, shortly after it could be demonstrated that gene regulation can be controlled by lncRNAs as exemplified by Xist, which associates with one copy of the X-chromosome and mediates its silencing in female mammals in a process called dosage compensation (Wutz, 2011). To date, biological roles of numerous lncRNAs are described, but the function of other members of the family is still poorly understood. Furthermore, it remains unclear, whether all members of the family are truly non-coding, as many lncRNAs can be found associated with ribosomes (Quinn & Chang, 2016).

Similarly, the role of circRNAs in biology initially was enigmatic. Circularization of these RNAs occurs co-transcriptionally and it is still unclear, whether numerous circRNAs are non-functional products of aberrant splicing (Chen, 2016). However, recently proteins were identified, which can actively promote circularization of RNA (Conn et al., 2015).This might be a first indication that circularization is a regulated process of biological relevance. CircRNAs carrying a multitude of miRNA binding sites have already been shown to act as miRNA sponges (Hansen et al., 2013; Memczak et al., 2013). circRNAs do not have 5’ or 3’ ends and are therefore resistant to exonuclease mediated degradation; circRNAs are hence more stable than most linear RNAs. Based on the stability and the possibility to carry several interaction sites, it is possible to envision a role as scavenger complexes targeting molecules such as miRNAs or RNA-binding proteins (Hentze & Preiss, 2013).

There are additional RNA species described, and the repertoire of RNA molecules is still growing, placing this molecule in the epicenter of cellular biology. The complexity of the RNA world with its biological diversity implies that a broad variety of trans-acting protein factors are required to enable RNA functions, the RNA binding proteins (RBPs).

1.3 From RNA-binding proteins to ribonucleoprotein complexes

Proteins associate with RNA from the transcription to the decay of RNA. RNA-binding proteins (RBPs) interact in a very dynamic manner with RNA to form ribonucleoprotein (RNP) complexes. RBPs execute regulatory functions at specific stages of RNA metabolism (e.g. during splicing, processing, transport, translation) (Dreyfuss et al., 2002). Further functions of RBPs can include structure modulation of the RNA by acting as RNA chaperones or helicases (Mohr et al., 2002; Semrad, 2011). Interaction with RNA is mediated by the RNA-binding domains (RBDs). Classical RBPs follow a modular design combining a limited scope of RBDs, including the RNA recognition motif (RRM), the zinc finger motif (Znf), the hnrnp K homology motif (KH), the S1 domain, double stranded RNA-binding motif (dsRBM) and others (see Lunde et al., 2007 for the complete repertoire). Some of these domains can bind also DNA or proteins (Clery & Allain, 2011). RRM is the most common RBD that can be found in about 0.5-1% of human genes (Finn et al., 2005). Several RRMs are often combined in tandem to increase binding affinity and specificity, as in the case of Poly-(A) binding protein (PABP), Sex lethal (Sxl) or Nucleolin (Ncl). Having several RRMs in tandem enables higher RNA binding affinity and sequence specificity (Chen & Varani, 2014).

The broad family of Znf motifs was initially described as DNA binding (Hanas et al., 1983). The Znf domains are classified by the combination of cysteines (C) and histidines (H), which are the amino acids that chelate the Zn^{2+} ion (e.g. CCHH, CCCH, CCCC) (Clery & Allain, 2011).

dsRBMs are often found in multiple copies in RBPs (e.g. Drosophila protein Staufen with 5 copies). dsRBMs recognize the RNA structure but they can also bind sequences specifically by projecting amino acids to the Watson-Crick

paired bases, like Adar2 that converts adenosine to inosine (Barraud & Allain, 2011).

The repertoire of RBPs has substantially grown due to recent genome-wide studies that enabled the identification of novel RBPs. With the increasing knowledge about RBPs many more non classical RBDs have been described (see later).

1.4 The central role of RBPs in post transcriptional gene regulation

Gene expression is a tightly regulated process, because the timely and spatial expression of a protein is important for the survival and adaptation of the organism. The numerous cues and stresses that affect an organism require a highly versatile, fast and flexible toolset of gene expression regulation (Figure 1.2). The extent of gene regulation increases with the complexity of the organism. In prokaryotes, transcription and translation are coupled processes. However, in eukaryotes, these two processes are separated into different compartments, the nucleus and the cytoplasm. The compartmentalization of processes allows regulation based on differential localization. In addition, eukaryotic mRNAs undergo a multitude of processing events including splicing, editing, capping with 7-methylguanosin (m7G) at the 5'end and polyadenylation at the 3' end. These events start already co-transcriptionally and are mediated by RBPs (Moore, 2005; Glisovic et al., 2008.)

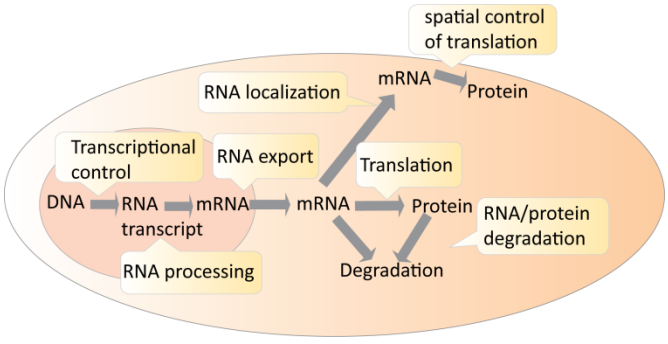

Figure 1.2; Different modes of gene expression control in eukaryotes. From transcription to decay.

Gene expression can be controlled through transcriptional regulation or post-transcriptional regulation. Transcriptional control is more economic, as this early control checkpoint saves energy resources like the nucleotides that are used during transcription. Nevertheless, post-transcriptional control is a favorable

process when it comes to fast responses to stimuli and stresses (Holcic & Sonenberg, 2005; Hershey et al., 2012). Post-transcriptional control is fast because it operates on already transcribed RNA. Translational control is conceptually the fastest way of post-transcriptional control because it bypasses all upstream time consuming processes of RNA production and processing.

RBPs are key players in post-transcriptional regulation and their differential binding specificities, combinatorial binding abilities and differential localizations allow a highly specific and diverse regulatory control of post-transcriptional events (Moore, 2005; Glisovic et al., 2008). One level of post-transcriptional control affects mRNA levels by altering the stability of the transcript. Abundance of an mRNA in the cytoplasm is dependent on its synthesis rate, processing, its nuclear export rates and degradation rate. The interaction of proteins on cis-acting elements to the RNA can fine tune all these processes, for example determining the mRNA degradation rate by recruiting mRNA degradation machineries. These cis-acting elements can be often found in the 3'UTR of mRNAs (Matoulkova et al., 2012). One prominent example is the iron responsive element (IRE), which is found in 5 copies in the 3'UTR of transferrin receptor mRNA. In iron-replete cells, the iron regulatory proteins 1/2 (IRP1/2) (also described later) do not bind the IREs and therefore the transcript becomes a target of endonuclease attack and subsequent degradation. In iron-deficient cells IRP1/IRP2 bind to the IREs and thereby protect the mRNA from degradation by a not further understood mechanism (Casey et al., 1988; Müllner & Kühn, 1988; Binder et al., 1994). Other examples are the AU rich elements (AREs) present in in the 3'UTR of mRNA (Gardiner et al. 2015) of a number of mRNAs encoding cytokines and chemokines. AREs can be bound by trans-acting factors like ARE- binding proteins causing mRNA decay and translational repression. There are also other factors like Hu-antigen D (HuD) that bind AREs and stabilize targeted mRNA (Bolognani et al. 2010; Gardiner et al. 2015).

Another way to control gene expression is on the level of mRNA splicing. Splicing of a pre-mRNA removes the introns, which are noncoding regions of the RNA. Splicing events on one pre-mRNA can then also lead to varying mRNA products that are translated into different proteins (Pan et al., 2008). This

process is in large extent regulated by differential association of RBPs that increase selection or occlude a given splicing donor or acceptor.

Another mechanism to expand or change the protein repertoire relies on RNA editing (Nishikura, 2016). RNA editing involves the conversion of adenosine to inosine catalyzed by the ADAR proteins. RNA editing can alter the protein product and also affect RNA secondary structure, stability and splicing. C to U editing by Apobec1 on apolipoprotein (apoB) mRNA creates a UAA termination codon and thereby results in a shorter protein product encoded by the apoB gene. The apoB gene thereby encodes for two apolipoproteins, the shorter ApoB48 and the longer ApoB100, that both have specific functions in the fatty acid transport in mammals (Davidson & Shelness, 2000).

The spatial control of RNA abundance by locating proteins to particular subcellular regions is an additional important element of gene expression regulation (Gonsalvez & Long, 2012). Localization of transcripts to a destined region of the cell or the organism has been well studied in different systems, including prominently Drosophila (Berleth et al., 1988; Ephrussi et al., 1991; Lécuyer et al., 2007) and yeast (Andoh et al., 2006; Schmid et al., 2006; Aronov et al., 2007). Cell polarization requires asymmetric localization of proteins. For this purpose specific mRNAs are often differentially localized. Localization of *oskar* (*osk*), *nanos* (*nos*), *bicoid* (*bcd*), and *gurken* (*grk*) mRNA is crucial for Drosophila embryonic axis specification (Lasko, 2012). The involvement of the cytoskeleton and the interaction with *cis* acting factors that associate the target mRNAs to molecular motors and microtubules is necessary for the localization of these mRNAs. Other *cis* acting factors prevent translation during transport of these RNPs.

mRNA levels do not always correlate with the protein levels in a cell. Correlation of mRNA levels and protein levels decreases with the complexity of the organism. The correlation of mRNA and protein levels for *E.coli* is R^2 ~0.8 (Taniguchi et al., 2010) whereas it is ~0.7 for yeast (Ghaemmaghami et al., 2003) and 0.4 in human (Schwanhäusser et al., 2011). Translational control is one explanation for the decreased correlation between mRNA and protein levels in the cell. Translational control is a post-transcriptional control mechanism that does not affect mRNA abundance but allows rapid changes of

protein levels in response to environmental cues (Hershey et al., 2012). Translation can be subdivided into initiation, elongation, termination and ribosome recycling. Many translational control mechanisms target the first step, the translational initiation (see below). Some eukaryotic mRNAs (~5-10% of all mRNAs) (Le & Maizel, 1997) or viral RNAs (Kieft et al., 2008) with internal ribosomal entry sites (IRESs) can often be translated by a mechanisms that differ from the canonical cap-dependent translation initiation through direct recruitment of the RNA to the 40S subunit. This makes IRES-containing transcripts less sensitive to downregulation of initiation factors under cellular stresses such as starvation (Gilbert et al., 2007; Riley et al., 2010) or infection (Bonnneau & Sonenberg, 1987; Irurzun et al., 1995). Length and structure of the 5' UTR of an mRNA can affect the efficiency of scanning, acting as a regulatory element. Specific translational regulation at the initiation step occurs often by binding of RBPs to specific sites in the 3'-UTR of an mRNA, often modulating the interactions between the 5' and 3' end of the RNA (closed loop) or altering the activity of initiation factors (Gray & Hentze, 1994). This is exemplified by the action of the Drosophila protein sex lethal (Sxl) on the mRNA of male specific lethal 2 (msl2), as described below (Duncan et al., 2006; Gebauer et al., 2003; Abaza et al., 2006; Beckmann et al., 2005; Duncan et al., 2009; Medenbach et al., 2011).

Besides silencing of their targets by induction of RNA degradation, miRNAs can regulate target RNAs at the translation step. Translational repression through miRNA can occur by shortening of the poly(A) tail (deadenylation), which hampers the closed loop formation and slows down initiation and the re-initiation event by post-terminating ribosomes. There is still no full understanding of how miRNAs regulate translation but they might act by targeting eIF4A, an RNA helicase needed to unwind the secondary structure nearby the start codon during translation initiation (Meijer et al., 2013; Ricci et al., 2013; Fukao et al., 2014; Fukaya et al., 2014). A recent study suggests an additional mechanism independent of eIF4AI activity and 43S scanning mechanism (Kuzuoğlu-Öztürk et al., 2016) that involves CCR4-NOT and DDX6. However, the molecular basis of this mechanism remains obscure.

Translational control events that affect the elongation phase are less frequently observed. Most of the time elongation works at maximal speed and efficiency. It is known that rare codons can slow down this process and thereby affect translational efficiency (Irwin et al., 1995; Yu et al., 2015). During translation of proteins resident in the endoplasmic reticulum the elongation is paused by the signal recognition particle (Wolin & Walter, 1989).

1.4.1 Sxl, a paradigmatic RBP involved in posttranscriptional control during Drosophila development

In Drosophila, the protein Sxl plays a pivotal role at the top of a sex determination and dosage compensation regulation cascade. Dosage compensation is a mechanism that ensures females and males to have equal levels of gene product despite having different number of chromosome X (Cline & Meyer, 1996). Sex determination ensures the correct development of sex characteristics in an organism. The Sxl gene is present on the X-chromosome; its sex specific transcription is regulated by transcription factors that also reside on the X-chromosome (Keyes et al., 1992). Just embryos that harbor 2 X-chromosomes can initiate the Sxl expression. Shortly after the sex specific transcription of Sxl from the early promoter, Sxl transcription starts from a second, “late”, promoter that is present and active in both males and females. The early sex specific Sxl directs then splicing of the late Sxl transcript to generate a translatable version of Sxl (Cline, 1984; Bell et al., 1991). It thereby creates a sex specific activation pulse of active Sxl. The autoregulative loop of splicing by the Sxl protein allows then Sxl production from the late transcript throughout life.

Sxl is a RBP with two conserved RRMs and an N-terminal glycine-rich region (Wang & Bell, 1994). The interaction of Sxl to target RNA works in a cooperative fashion. The glycine-rich region promotes this cooperativity and also allows co-recruitment of proteins with similar glycine-rich regions such as the HNRNP A/B homolog Hrb87F (Wang et al., 1997). The region of Sxl containing residues 122–301 (called dRBD4) shows full biological activity but the two RRMs alone (residues 123–294, called dRBD3) are sufficient for the RNA-binding activity (Grskovic et al., 2003).

Sxl binds to poly(U) stretches of target mRNAs (Inoue et al., 1990; Valcárcel et al., 1993). Target mRNAs include for example Sxl itself (Cline, 1984; Bell et al., 1991), transformer mRNA (tra) (Boggs et al., 1987) and male specific lethal-2 (msl-2) (Bashaw & Baker, 1995). On these target mRNAs Sxl functions as a translational repressor and/or as a regulator of alternative splicing. Tra is downstream of Sxl in the sex determination cascade and its sex specific expression is controlled by splicing regulation by Sxl. The transcription factor msl-2 is part of the male specific lethal complex consisting of four proteins that bind specifically to the X chromosome in the male to upregulate its expression for dosage compensation (Conrad & Akhtar, 2012). The msl-2 transcript carries three U-rich binding sites targeted by Sxl that could be defined to be 70 nt in the 5′ UTR (site B) (Gebauer et al., 1999) and 46 nt in the 3′ UTR (sites E and F) (Gebauer et al., 2003). For effective msl-2 silencing, Sxl acts via several mechanisms. At the level of splicing, Sxl promotes the retention of an intron that is necessary for downstream translational repression steps through Sxl (Gebauer et al., 1998). Binding of Sxl to its binding site displaces the large subunit of the U2 auxiliary factor (U2AF), diverting it to a distal low-affinity poly(Y) tract and inhibits thereby recognition of the 5' splice by the U2 small nuclear recognition particle (U2 snRNP) (Merendino et al., 1999).

Translational repression through Sxl is promoted by a dual-mechanism (Beckmann et al., 2005) (Figure 1.3). When Sxl binds to the 3’ UTR, it co-recruits Upstream of N-ras (UNR) (Duncan et al., 2006) and inhibits ribosome recruitment (Duncan et al., 2009) and thereby translational initiation. Sxl bound to the 5’ binding site interferes with ribosomal scanning (Gebauer et al., 2003) by a mechanism that involves redirecting the ribosome to an upstream AUG (Medenbach et al., 2011). Sxl is an enthralling RBP because of its strong RNA interaction and the multitude of functions it has on its target mRNAs. It also exemplifies the diversity of post transcriptional control mechanisms.

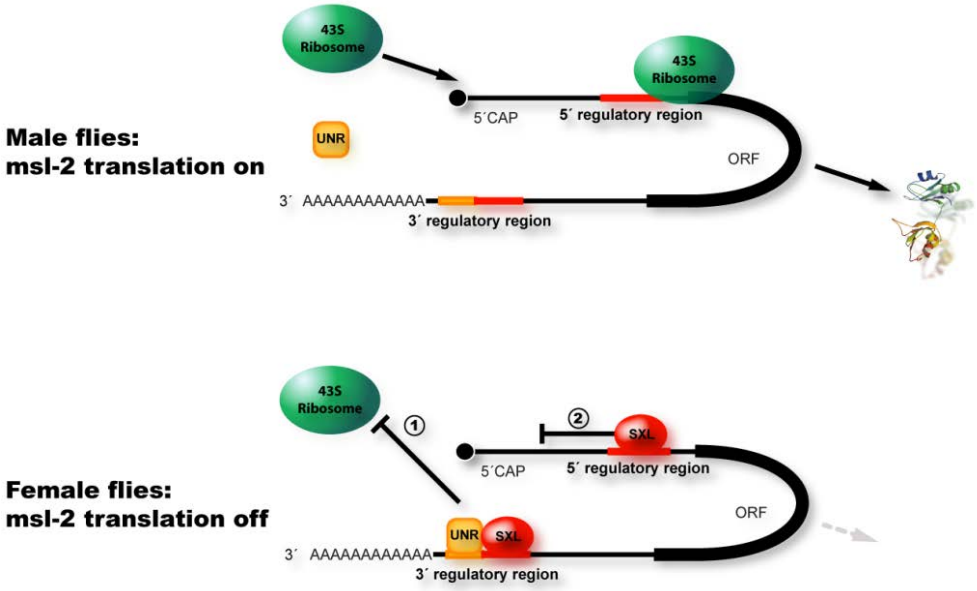

Figure 1.3; Mechanisms of translational control of msl2 mRNA promoted by Sxl. In female *Drosophila* cells, SXL is expressed and it can associate with U rich binding sites, which reside in the 5' and 3' UTRs of *msl-2* mRNA. Bound to the 3' UTR, Sxl can recruit UNR. The SXL–UNR complex then blocks the association of the 43S ribosomal preinitiation complex at the 5' end of *msl-2* mRNA, thereby repressing translation initiation. Sxl bound to the 5'UTR inhibits translation by redirecting the ribosome to an upstream AUG. Male *Drosophila* cells do not express SXL, the translation inhibitory complex is not formed and msl2 can be translated. (Figure kindly provided by Jan Medenbach, University of Regensburg; adapted from Duncan et al., 2006)

1.5 The Ribosome

1.5.1 Ribosomes –macromolecular complexes made of rRNAs and proteins

In all living cells the mRNA derived synthesis of proteins is carried out by a macromolecular machine, the ribosome. The ribosome was first discovered via electron microscopy of plant cells 1953 (Robinson & Brown, 1953). Later it was isolated from animal cells and at this time referred to as RNP particles or microsomes (Palade, 1955). The ribosome (termed 70S in prokaryotes and 80S in eukaryotes based on its Svedberg sedimentation coefficient) is composed of two subunits that have different functions during translation. The small subunit (30S in prokaryotes and 40S in eukaryotes) is the decoding center, where interactions between codons in the mRNA and the anticodons of tRNAs are monitored. Based on these interactions the right order of amino acids will be added to the nascent polypeptide. The large subunit (50S in prokaryotes and 60S in eukaryotes) carries the catalytic active center in which the formation of the peptide bonds is catalyzed. Both subunits consist of ribosomal RNA (rRNA) and ribosomal proteins. In eubacteria, the small subunit contains one 16S ribosomal RNA (rRNA) molecule and 21 small subunit ribosomal proteins (Rps

proteins), the large subunit contains 5S and 23S rRNAs and 33 large subunit ribosomal proteins (Rpl proteins) (Melnikov et al., 2012). In eukaryotes, the small subunit contains 18S rRNA and 32 Rps proteins, the large subunit consists of 28S, 5.8S and 5S rRNA and contains 47 Rpl proteins. Mitochondria also possess ribosomes for the synthesis of highly hydrophobic proteins (13 in human) involved in oxidative phosphorylation. Mitochondrial ribosomes (55S) have diverged from the prokaryotic ribosomes by having reduced RNA content and increased protein content (O'Brien, 2003).

Crystallographic analyses of ribosomes have been very challenging because of their size, complexity and the lack of symmetry, the presence of several flexible components and the instability of rRNA. Therefore it was a breakthrough in the field, when the first high resolution (2.4 Å) crystal structure of the prokaryotic 50S subunit was resolved (Ban et al., 2000) and could give first structural insights into protein synthesis. The peptidyl-transferase center directly attracted attention because of the absence of proteins. It was therefore observed in the structure what was already speculated, that the ribosome is a ribozyme in which the RNA acts as the catalyst (Cech, 2000).

Soon after first structural insights of the ribosome were given by the prokaryotic 50S subunit structure, other structures were resolved by cryo-electron microscopy and crystallography including eukaryotic, prokaryotic and the mitochondrial ribosome (Wimberley et al., 2000; Spahn et al., 2001; Spahn et al., 2004; Boehringer et al., 2005; Chandramouli et al., 2008 Taylor et al., 2009; Armache et al., 2010; Klinge et al., 2011; Rabl et al., 2011; Ben-Shem et al., 2011; Jenner et al., 2012; Anger et al., 2013; Greber et al., 2014; Amunts et al., 2015; Khatter et al., 2015).

The collection of structural data from different organisms in the tree of life allowed studies about structural conservations and differences (Figure 1.4). Bacterial and eukaryotic ribosomes share a common structural core, which includes 34 conserved proteins (15 in the small and 19 in the large subunit) and a core of 4400 nt rRNA. The conserved core harbors the functional centers of the ribosome, the A (Aminoacyl), the P (Peptidyl) and E (Exit) site (Spahn et al., 2001; Ramakrishnan, 2002; Smith et al., 2008). These sites are formed between the small and big subunit of the ribosome. The mRNA interacts with

the small subunit and moves in a step-wise manner from codon to codon during the translation, the catalytic function of the ribosome.

1.5.2 Expansions of rRNA in eukaryotes – the ribosomal expansion segments

Eukaryotes exhibit multiple extensions of the 18S and 28S rRNA in comparison to the prokaryotic rRNA; these insertions are referred to as expansion segments (Gerbi, 1996). Within eukaryotes the ribosomes also differ in their protein and RNA mass, which increases with the complexity of the organism (Melnikov et al., 2012; Anger et al., 2013). The increase of protein mass associated to protein size increase (protein extensions) appears to be modest, whereas the increase of rRNA mass is significant (Figure 1.4).

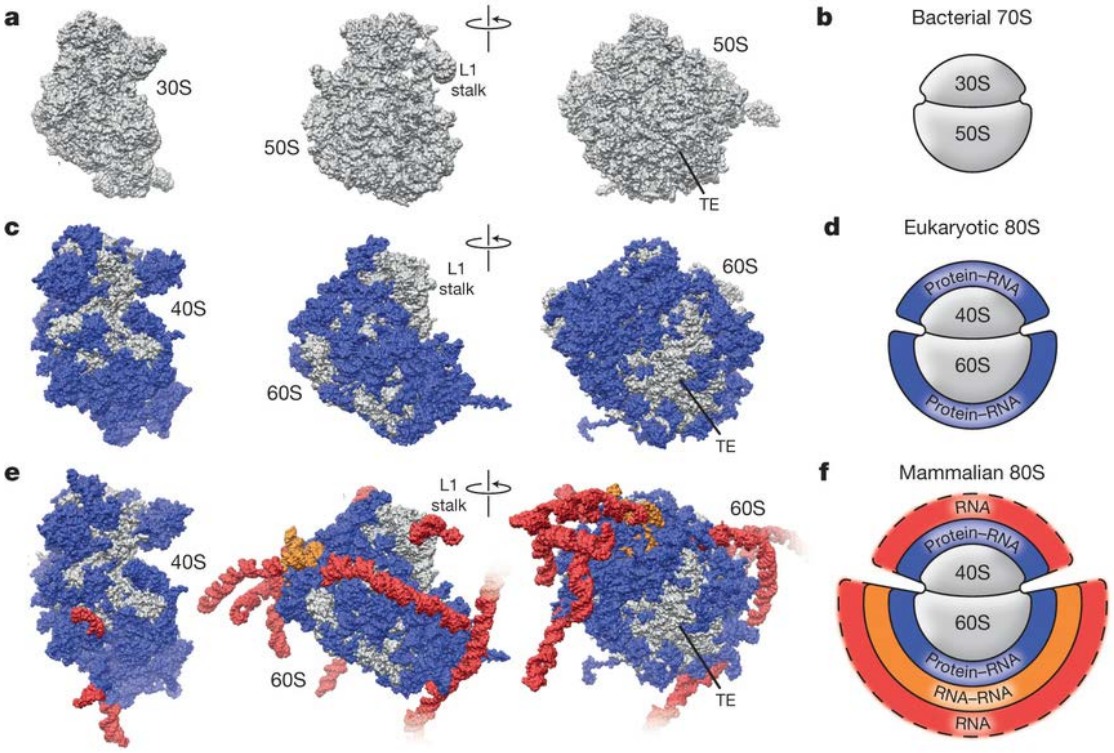

Figure 1.4; Expansion and evolution of the ribosome. Surface representations (a, c, e) and schematics (b, d, f) of the bacterial (*T. thermophiles*) ribosome (Jenner et al. 2010), the *S. cerevisiae* ribosome (Ben-Shem et al. 2011) and the human ribosome (Anger et al. 2013). TE, tunnel exit.(Figure modified from Anger et al. 2013; reprinted with permission; license number 3920180057435 provided by Copyright Clearance Center)

The 30 rRNA expansion elements (28 in yeast and tetrahymena), often protruding out of the ribosomal core, are involved in inter-subunit bridge formation (Armache et al., 2010; Anger et al., 2013; Khatter et al., 2015) and could serve as interaction platform of factors that might regulate translation or interact with the nascent peptide chain (Beckmann et al., 2001). Deletion of

rRNA expansion segments showed a role of some expansion segments in ribosomal biogenesis (Ramesh & Woolford, 2016). The longest expansion segment in humans is ES27L in the 28S rRNA. It builds a strong stem loop structure and might link the ribosome with the ER membrane (Pfeffer et al., 2012). ES27L is also involved in inter-subunit bridge formation and is suggested to recruit chaperones or modifying enzymes to the nascent peptide chain close at the peptide exit channel (Beckmann et al., 2001). In yeast, structural data showed that the biogenesis factor Arx1 (human ortholog Pa2G4/EBP1) interacts with the expansion segment ES27L (Greber et al., 2012) in the premature 60S subunit where it associates close to the polypeptide tunnel exit. Arx1 might thereby lock the premature 60S subunit to prevent premature association of nascent chain processing factors. The expansion segment ES6S is the longest in the human 18S rRNA and is involved in inter-subunit bridge formation. Nevertheless, there is still a lack of knowledge about the different functions of the expansion segments and why the complexity of these expansion segments increases with complexity of the organism. The expansion segments might associate with many translational and maturation regulators and can thereby contribute to the increasing layer of regulation in organisms with increasing complexity.

1.5.3 The ribosome – a molecular machine

The ribosome is required to translate the mRNA into protein. It reads the genetic code, i.e. nucleotide base triplet in the mRNA, and catalyzes the peptide bond formation between amino acids. The mechanism of translation by the ribosome comprises three phases, the initiation, the elongation and the termination and recycling phase (Dever & Green, 2012).

In eukaryotic translation (Hershey et al., 2012), cap-dependent translational initiation is the most frequent initiation mechanism (Figure 1.5) and involves scanning of the 43S pre-initiation complex from the 5’ cap of the mRNA to the start codon. Eukaryotic translation initiation factors (eIF1-6) are involved in the different steps of translation initiation. The GTP binding protein eIF2 specifically recognizes the initiator tRNA, which is loaded with the initiator amino acid Methionine (Met). This ternary complex binds to the small subunit and forms the 43S complex. The cap structure is recognized by the initiation factor complex

eIF4F (composed by eIF4A, E, and G), guiding the 43S complex to the 5' end of the mRNA. Supported by the eIFs and consuming ATP, the 40S ribosome subunit starts the scanning of the mRNA until it reaches the first start codon. Upon reaching the start codon, the 60S subunit binds to the 40S subunit in a process known as "joining". The initiator Met-tRNA then positions to the ribosomal P-site and eIF5 hydrolyzes the eiF2-bound GTP and thereby releases the initiation factors to finish the initiation phase.

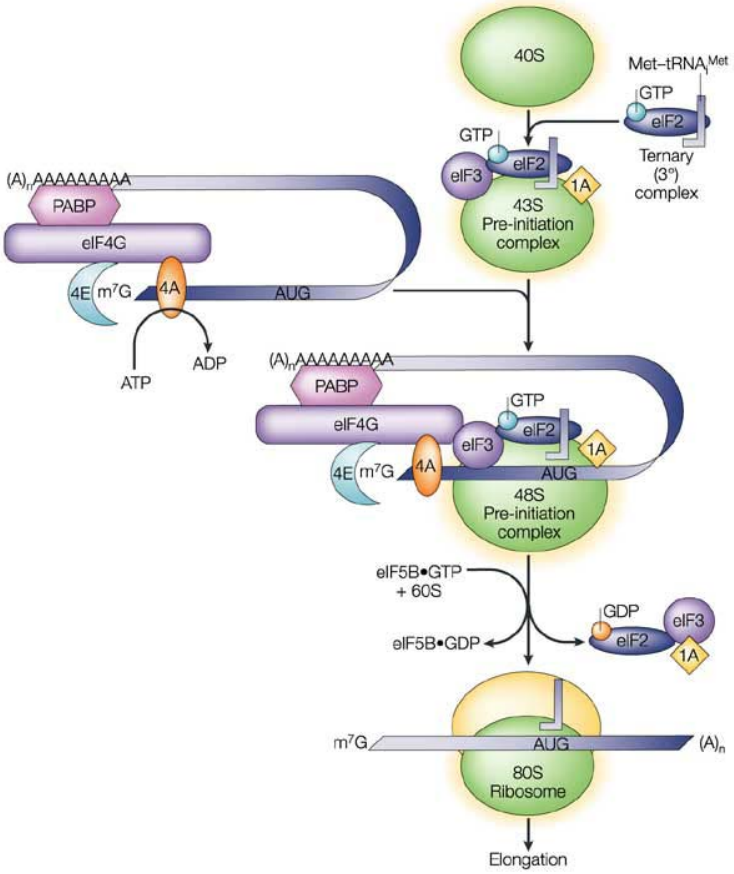

Figure 1.5; Cap dependent translation initiation mechanism in eukaryotes. (Figure from Klann & Dever, 2004; reprinted with permission; license number 3920190031575 provided by Copyright Clearance Center)

The elongation phase of translation (Dever & Green, 2012) involves the action of eukaryotic elongation factors (eEFs). When the first codon resides at the P site of the ribosome, the second codon is present at the A site and is awaiting the respective aminoacyl-tRNA. The elongation factor eEF1A binds the aminoacyl-tRNA in a GTP dependent manner and guides the tRNA to the A site of the ribosome, where the tRNA anticodon base pairs with the mRNA codon. Upon codon recognition and GTP hydrolysis by eEF1A, the aminoacyl-tRNA is positioned into the A site. Peptide bond formation with the P-site aminoacyl-

tRNA or peptidyl-tRNA is a rapid process catalyzed by the large subunit rRNA in the peptidyl transferase center. Conformational changes of the ribosome trigger the movement of the tRNAs into hybrid P/E and A/P states. GTP hydrolysis through eEF2 leads to the conformational changes that allow movement of the mRNA and tRNA. In this post-translocation state a free tRNA resides in the E site and the peptidyl-tRNA in the P site. The A site is free for the arrival of the next tRNA.

By reaching the stop codon during translation, the translational termination mechanism is initiated (Dever & Green, 2012). Translational termination involves the eukaryotic release factors (eRFs). eRF1 has a tRNA shape and can recognize the stop codon. eRF3 is a GTP binding protein that triggers upon GTP hydrolysis the deposition of a domain of eRF1 into the peptidyl transferase center, a process that results in peptide release from the ribosome.

1.5.4 The birth of the ribosome – ribosomal biogenesis

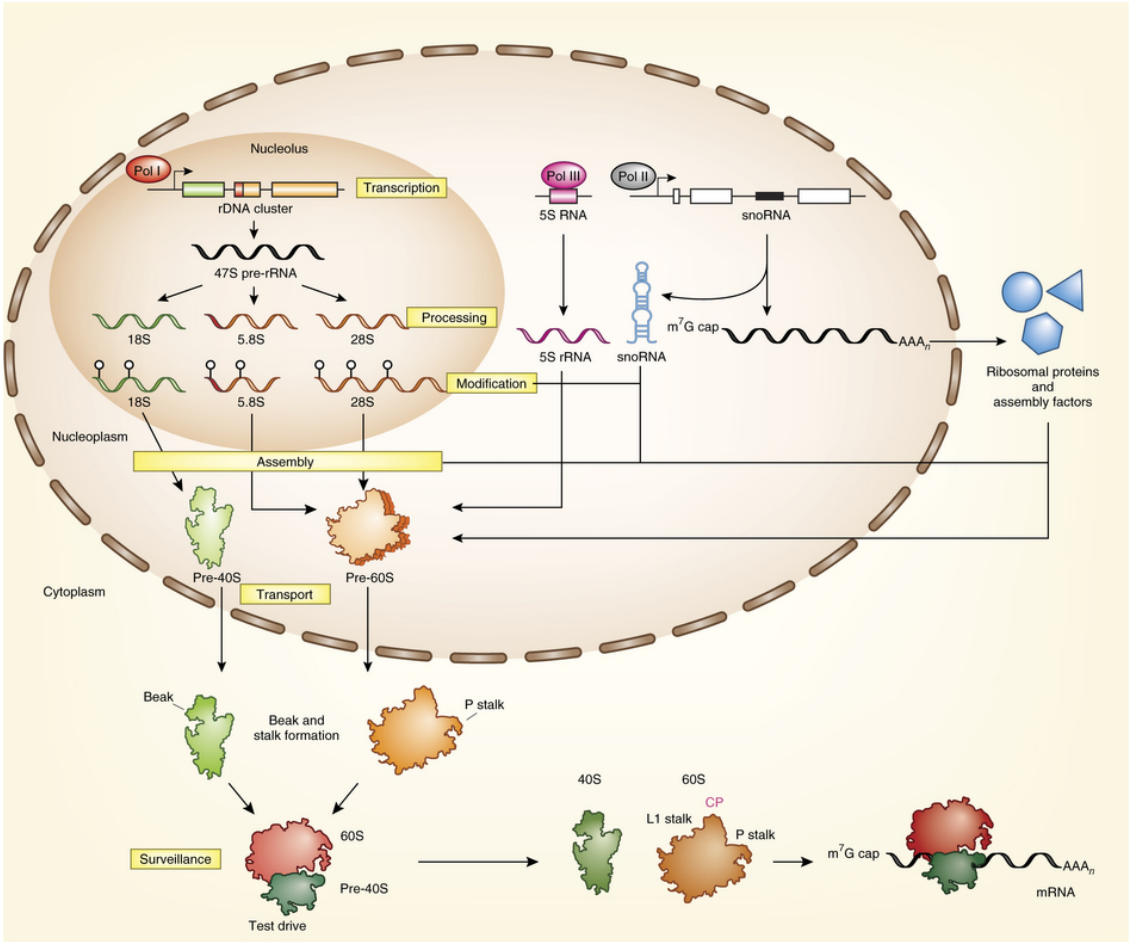

Figure 1.6; Ribosome Biogenesis at a glance. (from Lafontaine, 2015; reprinted with permission; license number 3920240959172 provided by Copyright Clearance Center)

Biogenesis of the ribosomes starts in the nucleolus and ends with the formation of the catalytic ribosome in the cytoplasm (Olson & Dundr, 2001). The process of ribosome biogenesis is a complex and energy demanding mechanism that includes rRNA processing and protein assembly steps (Figure 1.6). The 18S, 5.8S and 28S rRNAs are co-transcribed by Polymerase I (Pol I) as a long primary transcript (35S in yeast, 47S in human) from head to tail tandem repeats of rDNA (~150 in haploid yeast cells and 300-400 in diploid human cells) (Worton et al., 1988). In *S.cerevisiae*, the rDNA repeat also encodes for the 5S rRNA, which is transcribed by Pol III in reverse direction (Srivastasa & Schlesinger, 1991). In human and other eukaryotes, the 5S rRNA is synthesized by Pol III from multiple genes in close proximity to the nucleolus (Fedoriw et al., 2012). In the primary transcript, rRNAs are separated by spacer sequences, the internal transcribed spacers 1 and 2 (ITS1, ITS2) and the external transcribed spacers (5'ETS, 3'ETS). After transcription, a cascade of processing steps follows that include exonucleolytic and endonucleolytic cleavages (Henras et al., 2015). The order of some of the nucleolytic steps can differ between organisms but also from cell type to cell type (Bowman et al., 1983; Savino & Gerbi, 1990; Hadjiolova et al., 1993;). The first cleavage occurs at ITS1 and downstream to this cleavage the pathways for the small and big subunit biogenesis diverge. For biogenesis steps that include base modifications and folding, the primary transcripts associate with ribosomal proteins, pre-ribosomal factors and nucleolar ribonucleoprotein particles (snRNPs). More than 200 factors in addition to ribosomal proteins are involved in the ribosomal biogenesis (Thomson et al., 2013). snRNPs like U17, U14 or U8 function as RNA chaperones and assist the early cleavages (Reichow et al., 2007). Ribose methylation and uridine isomerization to pseudouridine is catalyzed boxC/D and box H/ACA RNPs respectively (Watkins & Bohnsack, 2012). The snoRNAs in these complexes hybridize to the specific positions in the ribosomal RNA and guide the modifications by the core proteins, the methyltransferase Fibrillarin and the pseudouridine synthase Dyskerin. Modification of rRNA is mostly found on functionally relevant positions of the ribosome (Liang et al., 2009). Many helicases of the AAA ATPase family are involved in ribosome biogenesis; they might modulate the RNA interactions within the pre-rRNA (Martin et al., 2013).

Most of the processing steps occur still in the nucleolus but additional processing take place in the nucleoplasm and the cytoplasm. The export to the cytoplasm appears to be specific for each subunit by interaction with specific export receptors. The karyopherin Crm1 mediates export of 40S and 60S particles in a Ran-GTP dependent manner. Adapter proteins that are distinct for 40S and 60S regulate the interaction with Crm1 via a nuclear export signal (NES). Nmd3 was identified to be the adapter protein specific for 60S export (Matsuo et al., 2014), whereas Slx9 was recently discovered to be the adapter protein for 40S export (Fischer et al., 2015). In the cytoplasm, the assembly of the ribosomal subunits is finalized and the translation inhibited till the maturation of the ribosome is complete. An example is the already mentioned binding of the protein Arx1 (human ortholog Pa2G4 or EBP1) to the peptide exit tunnel where it inhibits the association of translation factors in yeast (Greber et al., 2012).

Defective Ribosomes can be problematic for the function of the cell; therefore, surveillance mechanisms exist to eliminate dysfunctional ribosomal subunits. The multi-protein TRAMP complex, a cofactor of the exosomes, specifically targets misprocessed nuclear pre-rRNA and marks it for degradation by adding a 3’ oligo-A-tail (Dez et al., 2006). If the nuclear surveillance mechanism is bypassed, the non-functional rRNA decay (NRD), in which rRNAs that are fully maturated but harbor point mutation at sites important for translation in mature ribosomes are detected and degraded (Séraphin & Graille, 2012). NRD acts co-translationally and involves factors of the “No go” decay (NGD) machinery, which targets mRNAs that lead to stalling of ribosomes during translation. The mechanism of NRD is less clear than the NGD mechanism but one assumes that additional ubiquitination of ribosomal proteins and the additional action of the proteasome is required.

If biogenesis of the ribosome is affected by mutations of genes involved in biogenesis, a group of diseases is caused. The group of diseases associated with ribosome biogenesis is collectively called ribosomopathy (Freed et al., 2010). The most studied ribosomopathy disease is Diamond-Blackfan anemia (DBA) caused by mutation of ribosomal proteins. The phenotype of ribosomopathies is complex because a defect in ribosome biogenesis causes

ribosomal stress, which also involves p53 dependent pathways. DBA patients have malformations of limbs, an abnormal red blood cell physiology, and have an increased risk of cancer. The complexity of the ribosomopathies phenotypes illustrates the central role of ribosomes for cellular function and are not yet fully understood at the molecular level.

1.6 RBP analysis

1.6.1 Genome wide discovery of RBPs

Global RNA centric methods to identify RBPs have substantially expanded the diversity of RNA-protein interactions in the last years. First attempts to identify novel RBP were performed *in vitro*. Proteome-wide microarrays were incubated with fluorescence labeled RNA (Scherrer et al., 2010). After the washing steps, the arrays were scanned and analyzed for proteins that bound labeled RNAs. Unexpectedly, a number of metabolic enzymes were identified in these studies. *In vitro* studies are nevertheless prone to non-specific interactions with basic proteins through interaction with the negative charges at the RNA phosphate backbone. To circumvent this limitation, RNA interactome capture was developed in two independent laboratories to capture native protein-RNA interactions of living cells (Castello et al., 2012; Baltz et al., 2012). Both studies were based on UV crosslinking with 254 nm UV light, or with 365 nm UV using photoactivatable ribonucleotides (PAR), to covalently link the proteins to RNA placed at zero distances in living cells. Poly(A)+ RNA was capture by hybridization with oligod(T) beads . Stringent washes allowed the removal of proteins that were not covalently bound to the captured RNA. Proteins were released by RNase digestion and then analyzed by quantitative mass spectrometry (MS).

For human cells, the study lead to the identification of 1218 high probability RBPs in HeLa, HEK293 and Huh-7 cells (Beckmann et al., 2016). RNA interactome capture opened a new chapter for RBP biology and promoted the discovery of additional RBPs in other cellular systems and species, that were different human cell types (Castello et al., 2012; Baltz et al., 2012; Beckmann et al., 2015), murine ES cells (Kwon et al., 2013), murine macrophages (Liepelt et al., 2016), murine cardiomyocytes (Liao et al., 2016); *C.elegans* (Matia-

Gonzalez et al., 2015), parasites (*Plasmodium*, *Trypanosoma*) (Bunnik et al., 2016; Lueong et al., 2016), *D. melanogaster* (Wessels et al., 2016; Sysoev et al., 2016) *A. thaliana* (Marondedze et al., 2016) and *S.cerevisiae* (Mitchell et al., 2012; Matia-Gonzalez et al., 2015; Beckmann et al., 2015). Data from even more species are expected to become available in the near future. Knowing the repertoire of RBPs of different organisms will offer unprecedented opportunities to study the evolution of RBPs across species.

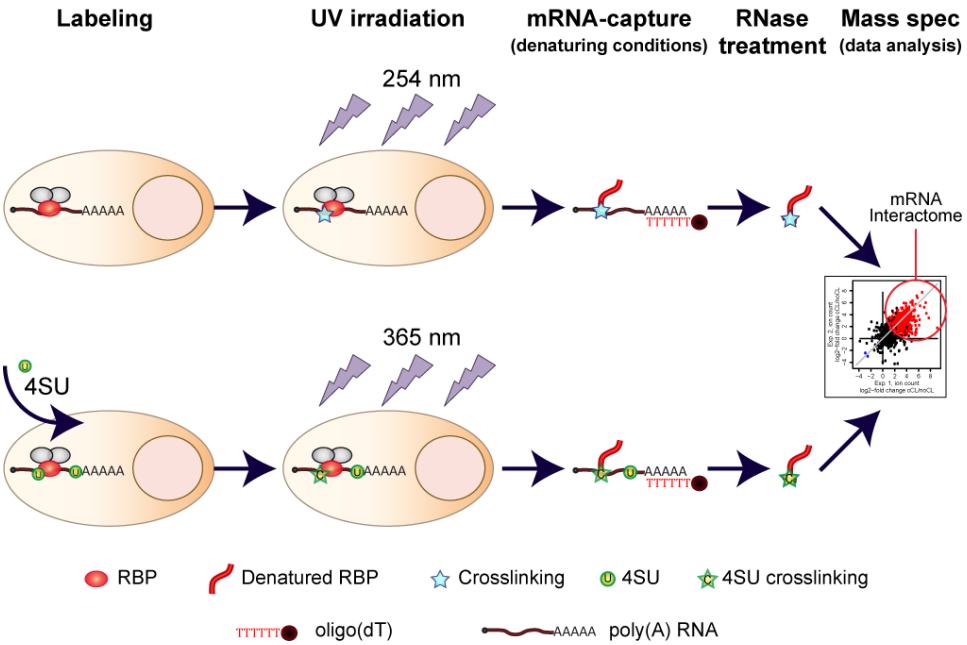

Figure 1.7; Schematic of the (m)RNA interactome capture protocol. Proteins bound to RNA are crosslinked in living cells by either conventional crosslinking at 254 nm (cCL) (top) or via PAR crosslinking at 365 nm (bottom). The mRNA together with the crosslinked proteins are captured on oligo(dT) magnetic beads. After stringent washes, the proteins are released by RNase treatment and analyzed via mass spectrometry (Mass Spec). (from Castello et al. 2013; figure reprinted with permission; license number 3920180833736 provided by Copyright Clearance Center)

The RNA interactome turned out to be more complex than previously expected, with 1218 high probability RBPs identified in HeLa (Castello et al., 2012). Interestingly, the majority of the newly identified RBPs do not contain a classical RBD (~55%) (Beckmann et al. 2016) (Figure 1.8). The proteins that do not contain a known RBD often harbor other motifs or domains associated to other functions. For example, the WD40 domain was previously described to be involved in protein-protein interactions and has now also been experimentally confirmed to also interact specifically with RNA (Lau et al., 2009).

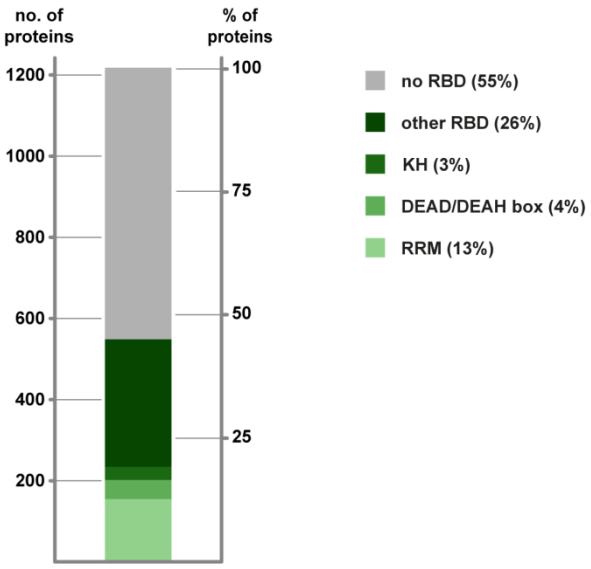

Figure 1.8; Novel RBDs account for majority of RNA binding. RNA interactome capture discovers many RNA-binding proteins that lack identifiable RNA-binding domains. (taken from Beckmann et al., 2016; Open Source, used under the terms of the Creative Commons Attribution 4.0)

To shed light on the possible new RNA-binding motifs or domains, a new method combining RNA interactome capture with differential (partial) protease digestion was developed. This method, named RBDmap (Castello et al., 2016), can define in a global scale the exact region with which the identified protein interacts with RNA. As in RNA interactome capture, RBPs crosslinked to polyadenylated RNA after UV irradiation are isolated using oligo(dT) beads. After elution, the RNA-protein complexes are treated with a protease that digests the proteins at positions that occur less frequently than thos for trypsin. The subsequent tryptic digestion step allowed, after a second round of oligo(dT) selection, the identification of the high-confidence RNA-binding region by MS measurements and bioinformatics analysis. RBDmap allowed the identification of 1174 high-confidence RNA binding sites in 529 RBPs, re-identifying classical RBDs but also discovering as RNA partners domains that have been previously known to be involved in other functions like enzymatic catalysis (e.g. thioredoxin, chaperones), involved in protein-protein interactions (e.g. 14-3-3), mediator of protein localization (cytoskeleton binding protein) or metabolic enzymes (e.g dinucleotide-binding, mononucleotide binding). Interestingly, half of the identified regions that interacted with RNA are disordered protein regions of intrinsically disordered proteins. Intrinsically disordered regions provide the highest flexibility to the interaction potential of a protein sequence allowing association with multitude of binding partners (Järvelin et al., 2016). These

regions are often of low amino acid complexity and specific amino acid combinations are readily identified, including the RGG-box. Previous NMR studies of FMR1 confirmed that the RGG boxes are important for the interaction with RNA (Phan et al., 2011). The arginines of RGG boxes can form electrostatic interactions with the RNA bases; the flexible glycine backbone allows spatial adaptation of to the polypeptide to the RNA shape.

Intriguing among newly discovered RBPs is the presence of metabolic enzymes (Castello et al., 2015). These RNA-binding metabolic enzymes often share mononucleotide or dinucleotide binding domains, which were also confirmed by RBDmap to be the probable docking sites for the interaction with RNA. The previously proposed REM (RNA, enzyme, metabolite) hypothesis describes the existence of a functional link between metabolism and gene expression that is driven by a group of metabolic enzymes with capacity to bind RNA (Hentze & Preiss, 2010). A well-studied example of such a network is the already mentioned Aconitase/ Iron responsive protein 1 (IRP1) function, which loses its cofactor under iron-deficient conditions and switches its enzymatic activity to binding of specific hairpin structures, Iron responsive elements (IREs), in the 5’ or 3’ of mRNA encoding iron homeostasis factors (Hentze et al., 1987; Casey et al., 1988; Müllner & Kühn, 1988). It thereby controls the translation or stability of RNAs involved in iron metabolism. A more recent example is the role of Glyceraldehyde phosphate Dehydrogenase (GAPDH) during the metabolic switch to aerobic glycolysis in T cell activation (Chang et al., 2013) (Figure 1.9). In inactive T cells GAPDH represses Interferon γ (IFNγ) expression by binding to an AU rich element (ARE) in the 3′ UTR of the IFNγ mRNA. Following T cell activation and the switch to aerobic glycolysis, GAPDH fully engages in its metabolic role allowing expression of IFNγ mRNA. As metabolic enzymes are highly susceptible to metabolic changes resulting from different cellular or environmental changes, they represent an attractive choice to regulate gene expression under these conditions. The exact mechanisms by which metabolic enzymes can switch their activities from metabolism to RNA binding remain poorly understood.

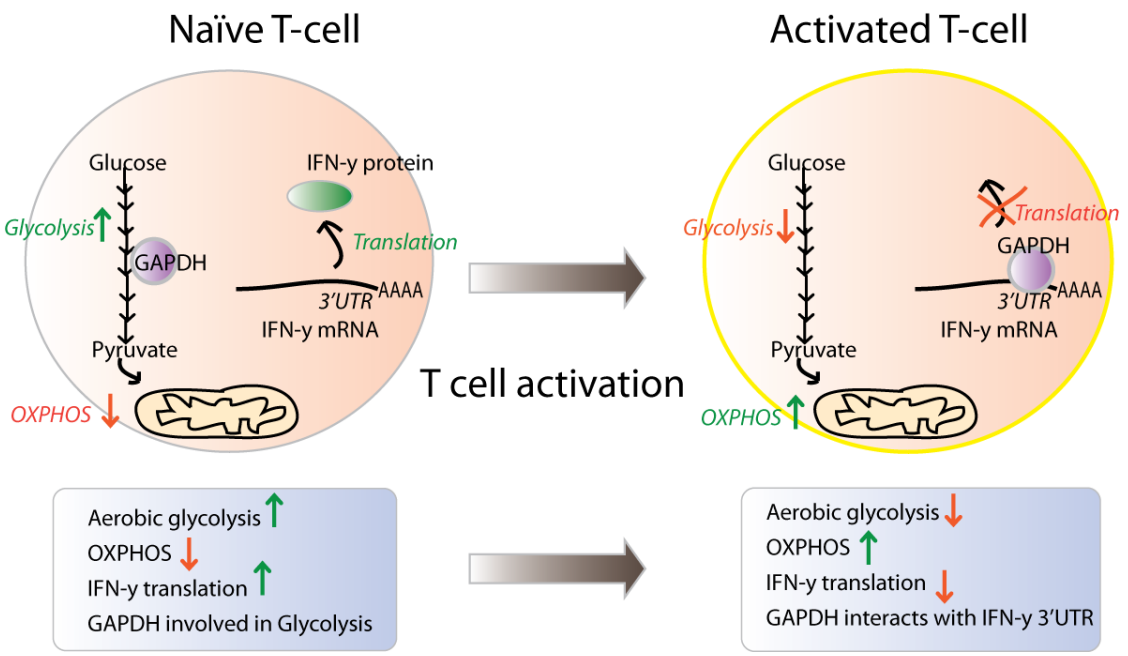

Figure 1.9; A metabolic switch during T cell activation transforms the glycolytic enzyme GAPDH to a RBP. After T cell activation, oxidative phosphorylation (OXPHOS) is active and glycolysis disengaged. After disengaging of glycolysis, GAPDH binds to AU rich elements within the 3'UTR of IFN-γ-mRNA and regulates its translation. (Redrawn and adapted with permission from Chang et al. 2013; license number 3920181256777 provided by Copyright Clearance Center)

Protein-RNA studies were mostly studied from a protein-centric perspective, assuming that binding affinity and specificity is promoted by the protein and the RNA acts as a passive entity in the interaction. As the RNA interactome studies revealed many high confidence RBPs without any known RBD, one can hypothesize that the RNA could also mediate the binding through 3D shape-complementarity (Beckmann et al., 2016). Specific binding mediated by an RNA molecule is well established for RNA aptamers and viral internal ribosomal entry sites (IRESs) (Pelletier & Sonenberg, 1988; Jang et al., 1988; Hashem et al., 2013). It is then possible that many RNAs evolved to specifically interact with a given protein. With increasing knowledge on RNA modifications coming from epitranscriptomics data (Saletore et al., 2012) one can speculate that posttranscriptional modifications on RNA also regulate protein-RNA interactions. One example for an epitranscriptomic switch was recently described for the interaction between HNRNPC and mRNA, where an N6-Methyladenosine (m6A) modification increases the binding (Liu et al., 2015).

In summary, the nature of most RNPs is highly dynamic and versatile. RBPs are therefore the key players of RNA fate through regulation of gene expression upon environmental cues, cellular stresses and signaling.

1.6.2 Studying the posttranscriptional networks of RBPs

Protein centric methods to examine the interaction of a given protein with RNA are often based on protein immunoprecipitation. The RNAs associated with the protein of interest are then identified by sequencing or microarray analysis. RNA immunoprecipitation (RIP) (Steitz, 1989; Tenenbaum et al., 2002; Gilbert & Svejstrup, 2006) and cross-linking immunoprecipitation (CLIP) (Ule et al., 2003; Licatalosi et al., 2008) are the two fundamental protein-centric approaches extensively used in the RNA field (Figure 1.10). The difference between these techniques is the use of UV crosslinking step at 254 nm in the CLIP protocol or 365 nm in the PAR-CLIP protocol (see below) to perform site specific crosslinking. UV crosslinking is analogous to the formaldehyde fixation step for the analysis of chromatin in the chromatin immunoprecipitation (CHIP) protocol (Orlando, 2000). UV light produces an irreversible cross-link of proteins interacting with RNA in “zero” distance (Pashev et al., 1991; Hafner et al., 2010). UV crosslinking has the advantage to allow more stringent washes and thereby addresses possible issues with post-lysis associations of binding partners, which appears to be a major problem in RIP based approaches (Darnell, 2010). Stringent conditions in RIP still have to be mild enough to allow RNA-protein interactions to be maintained. RNA can be sticky and thereby one can get significant backgrounds in RIP experiments. The validation of noisy data can be challenging.

In CLIP, downstream identification of the binding site can be achieved with single nucleotide resolution, whereas using RIP only a low resolution identification of the binding site can be made. Single nucleotide resolution identification was improved with the revised protocol of individual-nucleotide resolution CLIP (iCLIP) (König et al. 2010).

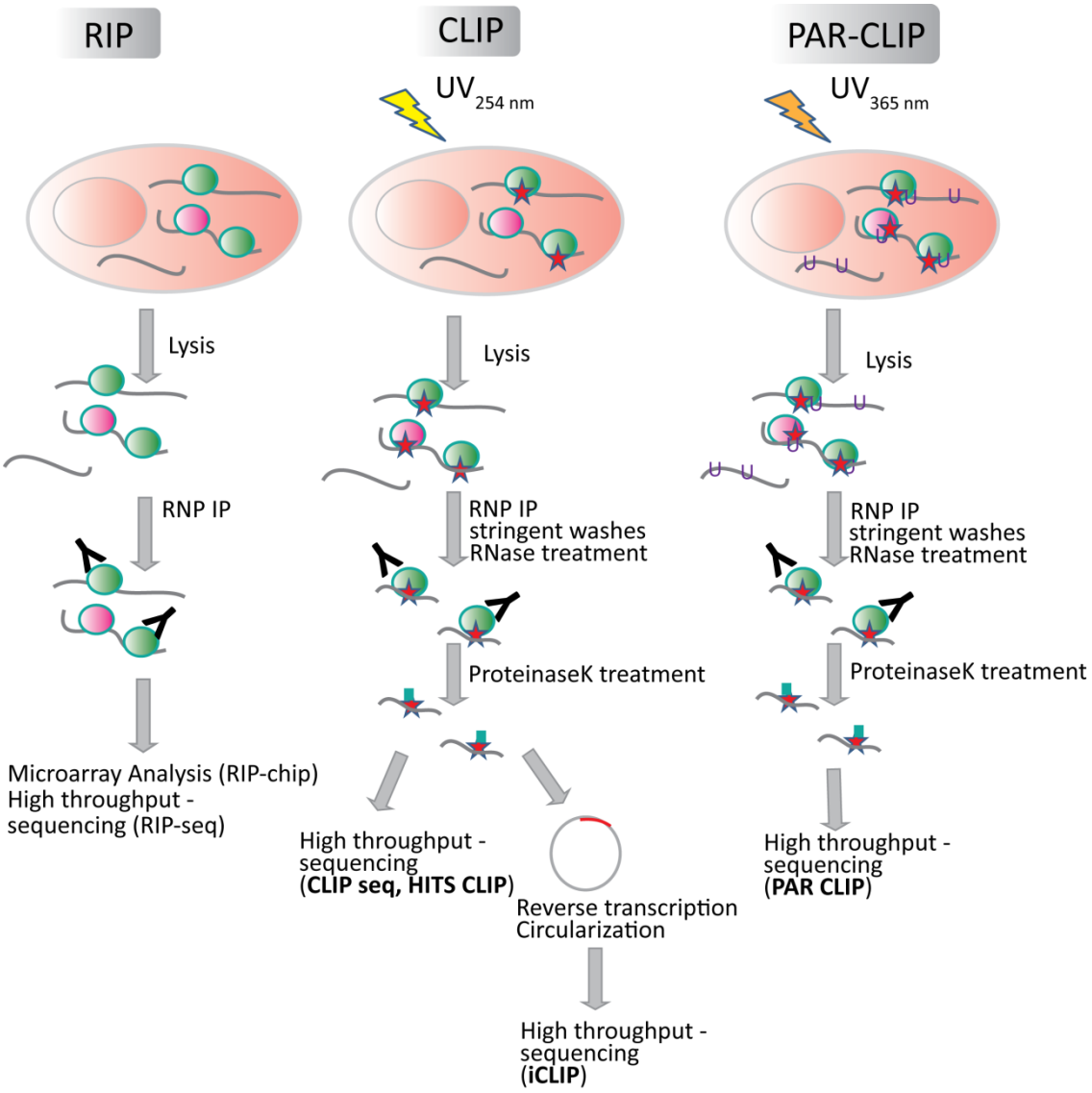

Figure 1.10; Comparison of immunoprecipitation based protein centric methods.

Another protocol based on site specific crosslinking uses photoactivatable ribonucleotides like 4-thiouridine or 6-thioguanosine. These photoactivatable ribonucleotides (PAR) allow crosslinking at a wavelength of 365 nm, the corresponding method is called PAR-CLIP (Hafner et al., 2010; Figure 1.10). Enhanced CLIP (eCLIP) represents a recent improvement of iCLIP, which allows mid-throughput library generation through a highly optimized workflow. This allowed the generation of large initial datasets including 102 eCLIP experiments for 73 diverse RBPs in two human cell lines (van Nostrand et al., 2016). The number of generated eCLIP datasets of RBPs is continuously growing.

1.6.3 RBP identification of specific RNA species

For the specific enrichment of a transcript together with its protein binding partners different strategies have been developed. Current techniques involve the use of different affinity tags inserted into the target RNA or hybridization with antisense oligonucleotides. Tags can be biotin (Lamond et al., 1988; Srisawat & Engelke, 2001), aptamers (Bachler et al., 1999; Leppek & Stoecklin, 2013) or specific protein binding sequences (Youngman & Green, 2005). The biotin tag is widely used because of its high affinity to streptavidin (K_d = 10^{-14} M) (Green, 1975) that is almost as stable as covalent linkage. As there are also biotinylated proteins in a cell, streptavidin-biotin based pull-down of protein complexes have the risk of trapping many contaminants (Wilchek & Bayer, 1988). Insertion of sequences (e.g. MS2 loops) in target RNAs has the disadvantage that the tag can alter the structural properties of the RNA affecting its lifecycle. An early approach used biotin tagged pre-mRNA immobilized via avidin to a solid support to pull-down proteins involved in splicing from HeLa cell nuclear extracts (Lamond et al., 1988).

Bacteriophage-derived tethering systems like the interaction between the MS2 motif with the MS2 coat protein (MCP) and the interaction between the boxB RNA element with the lambda N protein are often used for RNA pull-down experiments. Especially the MS2/MCP system was utilized for several purifications of RNP complexes (Das et al., 2000; Leonov et al., 2003). In MS2-BioTRAP (Tsai et al., 2011) the MS2 system was also applied for the isolation of *in vivo* assembled RNPs. One disadvantage of RNA capture through tethering systems is the incompatibility with highly stringent denaturing and high salt conditions, which does not allow the removal of indirect binders. To limit contaminations due to resin interaction, tandem RNA affinity methods like TRAP (Krause & Simmonds, 2004) or RAT (Hogg & Collins, 2007) make use of two different tags on the RNA to remove unspecific contaminants that are interacting with the resin.

Another possibility to target specific RNAs is the use of aptamers. Aptamers are highly structured RNA motifs that were designed and selected *in vitro* to bind specific proteins or molecules. One example is an aptamer that was modeled to bind streptomycin (Bachler et al., 1999) and was used to purify U1 snRNPs.

Another example is a streptavidin binding aptamer that was termed S1. S1 enabled the purification of RNaseP associated RNPs. Later the S1 aptamer was further improved to have higher affinity than S1 or the MS2 and PP7 system (Leppek & Stoecklin, 2013). An advantage is that the aptamer can be directly immobilized on a resin during the pull-down protocol and no co-expression or *in vitro* expression of tagged protein is necessary.

As already mentioned, the application of tags has its advantages to target specific RNAs, but the required modifications on the RNA structure or the usage of genetic engineering to express a tagged RNA in cells can produce biases in the analysis of the RNA interactome. Structural elements of the RNA that are involved in RNA-protein interactions might not form when the endogenous RNA structure is affected by the tag. In addition, the processing steps, intracellular localization and translation might be changed by the introduced tag. Therefore, also other approaches were developed to target the RNA of interest with its bound RBPs specifically. These approaches are based on CRISPR technology, or antisense oligonucleotide hybridization.

With the discovery of the biological function and mechanism of the clustered interspaced short palindromic repeats (CRISPR) and CRISPR associated protein (CAS) in prokaryotic immunity, the research toolbox could be widely extended (Wright et al., 2016). For the RNA-pulldown the protein Cas9 is biotin tagged and a specially designed protospacer adjacent motif (PAM) present in oligonucleotides (PAMmers) is used to direct Cas9 to the RNA of interest. The complex can then be pulled down via the Biotin tag of Cas9 (O'Connell et al., 2014). The drawback of this elegant pulldown strategy is its complexity, the incompatibility with highly stringent denaturing conditions (due to its dependent on folded Cas protein) and the dependence on genetic engineering.

One early example of the use of antisense oligonucleotides is the isolation of snRNP complexes (Lamond et al., 1989) and the telomerase (Lingner & Cech, 1996) with antisense oligonucleotides made of 2′-*O*-alkyl RNA. The alkylation makes RNA resistant to nucleases and enable it to form specific and stable hybrids with the target RNA. The hybridization of 2′-*O*-alkyl RNA oligonucleotides to the target RNA is more specific and stable than with DNA oligonucleotides. The coupling of biotin residues to the used oligonucleotides

enables affinity purification of the RNP complexes. Another example for an antisense oligonucleotide based protocol is the peptide nucleic acid (PNA) assisted identification of RBPs (PAIR) (Zielinski et al., 2006). PNAs enable the oligonucleotide to penetrate cell membranes of living cells and to hybridize to their targets. The PNA used in PAIR also contains the photoactivatable amino acid adduct p-benzophenylalanine. This adduct covalently crosslinks to proteins after exposure to UV light. The crosslinking has just an effect over a short distance (<4.5 Å), hence a large number of oligonucleotides has to be used for a comprehensive characterization of the binding partners. Another drawback is the fact that the protocol is dependent on two hybridization steps, one within the cell between the PNA and the target mRNA, one for the isolation between biotin tagged oligonucleotides and the target mRNA. Two sequential hybridization steps might limit the efficiency of the protocol.

For the characterization of proteins bound to lncRNAs several labs established a protocol that uses a multitude of biotinylated DNA oligonucleotides hybridizing to the RNA of interest (referred to as “tiling approach”) (West et al., 2014; Chu et al., 2015; McHugh et al., 2015) (Figure 1.11). This protocol is exemplified by the recent determination of the Xist RNA-bound proteome by the methods named CHIRP-MS (Chromatin Isolation by RNA Purification) (Chu et al., 2015) and RAP-MS (RNA Antisense Purification) (McHugh et al., 2015). The advantage is that for the “tiling approach” RNA integrity is not crucial. Nevertheless, the disadvantage of using a broad number of oligonucleotides is that it is impossible to estimate the contribution to noise and signal of each individual probe. Moreover, it would be difficult to target distinct transcript isoforms or particular regions of the RNA of interest, since the probes cover the whole transcript. In addition, some of the “tiling approaches” make use of extensive formaldehyde crosslinking (West et al., 2014; Chu et al., 2015), which leads to the detection of many indirect interacting proteins. The high percentage of formaldehyde (3%) used in these approaches might even lead to a fixation of the whole cell.

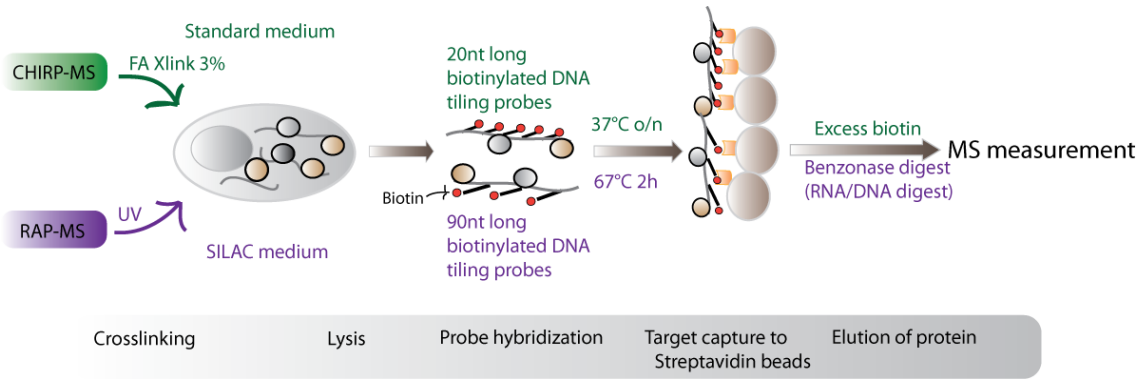

Figure 1.11; Two recent „tiling approaches“ for specific RNP capture. (CHIRP MS from Chu et al. 2015; RAP-MS from McHugh et al. 2015)**;** Both methods make use a pool of biotinylated DNA probes complementary to the RNA of interest. RNA-protein interactions are preserved by UV crosslinking or fixation with formaldehyde. After hybridization, the complexes are captured with streptavidin beads that bind the biotinylated DNA probes. The proteins are finally released and analyzed via mass spectrometry. (Figure redrawn and adapted with permission (Roth & Diederich 2015); license number 3920181058057 provided by Copyright Clearance Center) FA Xlink = Formaldehyde Crosslinking

My PhD thesis project is based on the antisense hybridization strategy and involves UV crosslinking. The difference to other methods is the use of locked nucleic acids (LNAs) that allow stabilization of the oligonucleotide and also increases the hybridization efficiency by forcing the oligonucleotide in a hybridization friendly conformation (Kauppinen et al., 2005) (Figure 1.12) , even better than the already mentioned 2′-O-alkyl RNA oligonucleotides. The conformation of the furanose ring of DNA and RNA exists in equilibrium between two conformations, the C2′-endo and C3′-endo conformation. In contrast to DNA, RNA forms the most stable hybrids in the C3’endo conformation. The methylene bridge in the LNA molecule is locking the conformation of the furanose ring to an extreme C3′-endo conformation, thereby promotes the stable RNA-RNA hybrid formation. LNA oligonucleotides were used before for poly(A) RNA enrichment from guanidinium thiocyanate (Jacobsen et al., 2004) containing cell lysate and a kit for LNA poly(A) mRNA sample preparation is available from the supplier Exiqon. Another difference to the other protocols is the use of covalent linkage of the oligonucleotide to solid support, which allows stringent conditions and avoids contaminations by interaction with resin proteins as it is known for the biotin-streptavidin immobilization (Wilchek & Bayer, 1988).

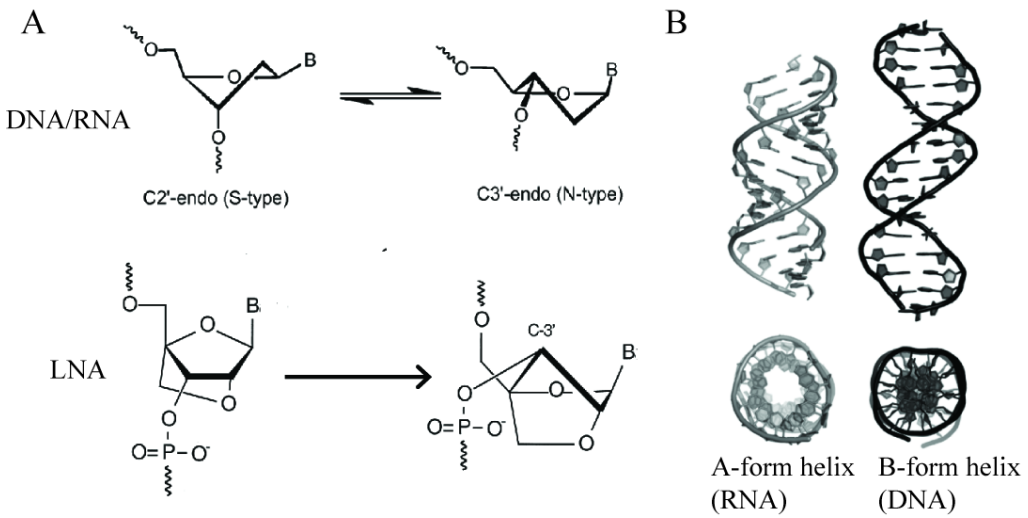

Figure 1.12; Comparison between conformations of nucleotides. A) Conformations of furanose ring in RNA, DNA and LNA (pictures taken and modified from Elmen et al. 2014; Wahlestedt et al. 2000) B) Helix formations of RNA in C3' endo (A-form helix) and DNA in C2' endo (B-form helix) (picture taken and modified from Anosova et al. 2016; Open Source used under the terms of the Creative Commons Attribution 4.0)

Several strategies to identify the proteins bound to the RNA of interest have been developed. The most important feature a method should have is to be specific, effective and robust.

1.7 Aims of this Thesis

RNA-binding proteins (RBPs) associate with RNA from synthesis to decay, forming dynamic ribonucleoprotein (RNP) complexes that enable regulation of diverse biological functions at multiple levels. The impact of RBPs on all components of the cell highlights the importance of studying the function, regulation and biology of RBPs. The knowledge about the number of existing RBPs has expanded substantially thanks to RNA interactome capture (Baltz et al., 2012; Castello et al., 2012), which provided new insights about the biological scope of RBPs and the interconnection of RBPs with different cellular networks. The method employs UV crosslinking that creates a covalent bond between RNA and protein *in vivo*, a stringent oligo(dT) capture protocol and mass spectrometry. The RNA interactome studies revealed both known and novel RBPs bound to polyadenylated RNA pool, but cannot be applied to defined RNA species or to RNAs lacking poly(A) tails. Currently, identification of RBPs associated with specific transcripts is a challenging task; and the available strategies to purify specific transcripts and their bound proteome are faced with numerous limitations. Thus, new methods to determine the

composition of proteins on a given RNA are required to further understand the regulation and biological function of the RNA and its associated RBPs

Aim 1: Development of a method for the capture of a specific RNA transcript and its bound proteome.

The aim of my thesis was to develop a method to isolate a specific RNA transcript and its bound proteome guided by technical knowledge from the established RNA interactome capture protocol. We refer to this protocol as “specific RNP capture”. This approach should aim to be highly specific and still give a reasonable yield, features achieved through the use of UV protein-RNA crosslinking followed by isolation with short LNA (locked nucleic acid)/DNA antisense probes. As the probes are covalently linked to beads and proteins are covalently bound to the target RNA, the stringency of the washes can reach denaturing conditions. These features combined with different specificity controls (i.e. non-irradiated samples and scramble LNA probes) allow high confidence identification of RBPs with Specific RNP capture. The set-up of specific RNP capture should enable RBP purification both in *in vitro* systems and in cell applications, making specific RNP capture a versatile tool for RNA biology research.

Aim 2: Validation of the method with a well-established protein-RNA pair system as positive control.

Using the well-characterized RBP-mRNA interaction, Sxl-msl2 mRNA, I aimed to validate specific RNP capture in HeLa extracts supplemented with recombinant Sxl protein and programmed with an ectopic reporter mRNA bearing the sequences from msl2 known to be recognized by this protein. I tested the protocol further in *D. melanogaster* embryo extracts without supplementation of exogenous proteins and combined it with quantitative mass spectrometry to allow unbiased identification of proteins interacting with the defined sequences in the reporter RNAs.

Aim 3: Characterization of 18S and 28S RNPs

Once specific RNP capture was established in *in vitro* systems, I aimed to challenge it to capture endogenous RBPs. I chose 18S and 28S rRNA as both a

proof-of-principle and biologically relevant target because these transcripts are highly abundant, there are equally abundant “contaminants” and there is incomplete knowledge of rRNA biogenesis, especially in humans. Cross-referencing of the identified proteins to recent ribosomal biogenesis screens (Wild et al., 2010; Tafforeau et al., 2013; Badertscher et al., 2015) and ribosome related literature gave further confidence in the identified proteins. It also allowed me to identify potential novel proteins involved in rRNA biology. This aim should therefore provide new insights into proteins involved in rRNA biology, and provide an example of specific RNP capture as new tool for RNA biology.

2 Results and Discussion

2.1 Specific RNP capture: concept and experimental strategy

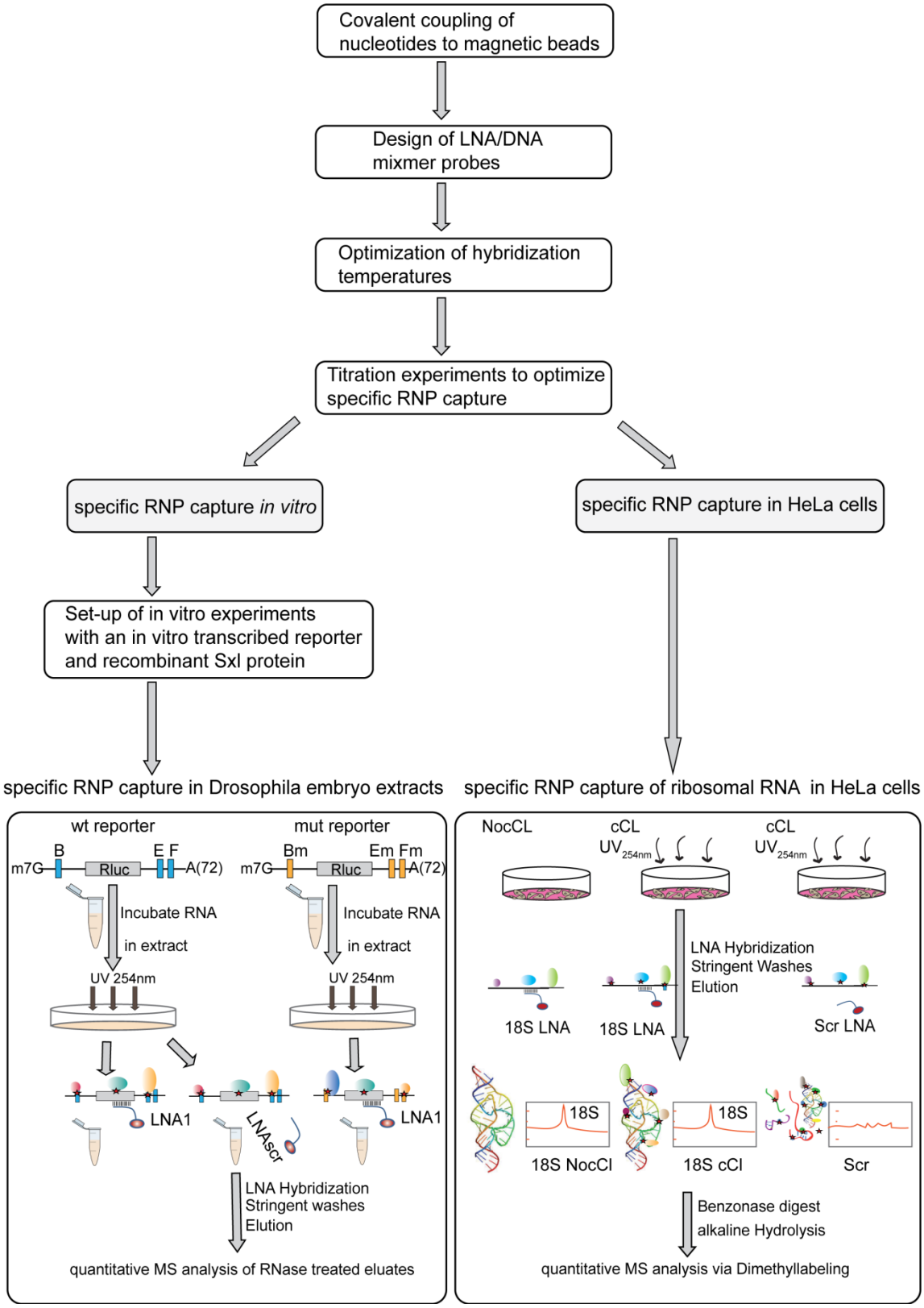

Figure 2.1; Workflow for the set up and experiments of specific RNP capture.

For setting up the protocol for specific RNP capture I followed strategic steps required to maximize the specificity and stringency of the protocol (Figure 2.1). After the general set-up steps, my workfkow is branching in *in vitro* and in-cell Specific RNP capture experiments. The protocol is built on the principles of RNA interactome capture (Castello et al., 2012; Castello et al., 2013), which makes use of ultraviolet (UV) irradiation of living cells to covalently cross-link native RBPs to RNA. Proteins covalently bound to polyadenylated RNA are captured by hybridization of the poly(A) tail with oligo(dT) magnetic beads, replacing the oligo(dT) beads of the interactome capture protocol. After stringent washes to remove indirect binders, the RBPs are detected by quantitative MS. To capture specific RNAs, I modified the current protocol by implementing the use of specific target antisense oligonucleotides covalently coupled to beads. I performed covalent coupling of the oligonucleotide probes to the magnetic beads to allow the use of high stringent washes that cannot be withstood by non-covalent protein-substrate interactions. For the set-up of specific RNP capture I first had to establish the protocol to covalent couple the probes to the beads. Next, I designed oligonucleotide probes composed of a mix of LNA and DNA nucleotides (LNA/DNA mixmers) and determined the optimal hybridization temperature. Titration experiments helped me to identify the optimal UV irradiation energy and the amount of input RNA needed to detect the RNA bound proteins by Specific RNP capture. Finally, I applied the resulting protocol to an *in vitro* system, using the well-established sex lethal-msl2 system, and to an in-cells set up to better understand the composition of 18S and 28S ribosomal RNPs.

2.2 Covalent coupling of oligonucleotide probes to magnetic beads

I first examined different strategies for covalent coupling of oligonucleotides to magnetic beads. As I was interested in a quantitative and simple readout, I established a test system that is based on dual fluorescence measurement on a plate reader (Figure 2.2). For coupling, I used 20mer DNA oligonucleotides that carried a fluorescent Cyanine 5 (Cy5) label at the 5’end and a primary amine with a C6 carbon linker at the 3’end. The carbon linker should allow flexibility for efficient hybridization. I tested two different coupling strategies (i.e. coupling to

epoxy residues or carboxy residues) and different coupling temperatures. After coupling, I used an antisense DNA oligonucleotide carrying a Cyanine 3 (Cy3) label at its 3'end to test the hybridization capacity of the probe-coated beads. Cy3 and Cy5 have distinct excitation and emission wavelengths from one another and can be therefore used together in a dual fluorescent screen. The testing conditions were chosen according to suggestions of different bead suppliers and knowledge of the chemical reaction conditions of the coupling groups. I ensured by the choice of the buffer system and the pH that a reaction targeting a primary amine is favored, in order to prevent coupling to exocyclic amines of the nucleotide bases.

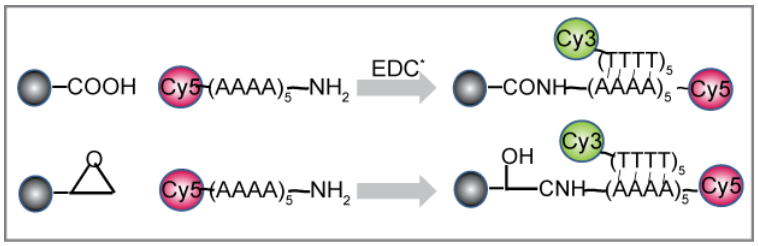

Figure 2.2; Dual Fluorescence Assay. The assay was established to monitor covalent coupling efficiencies and downstream hybridization efficiencies on the same beads. The oligonucleotide that should be coupled to the beads carries a 3' Amino-modifier (includes a primary amine and a flexible carbon C6 linker) and at its 5' a fluorescent Cyanine dye (Cy5). The hybridization oligonucleotide is antisense to the coupling oligonucleotide and contains a 3' fluorescent Cyanine dye (Cy3). Two coupling strategies were tested, coupling to carboxylic or to epoxy groups. Coupling to carboxylic groups needs the auxiliary activating agent N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC).

Comparing the Cy5 fluorescence signal, I observed that the carboxy-beads exhibited higher coupling efficiency, the most efficient coupling reaction being achieved at 50°C (Figure 2.3 A). For the epoxy-beads, increase of temperature improved coupling efficiency but it did not reach the level of coupling of carboxy-beads. Subsequently, I tested the coupled beads for hybridization efficiency at 4°C using the in series the buffers employed in the RNA interactome capture protocol. The highest amount of antisense DNA captured, measured as Cy3 fluorescence signal, was achieved with carboxy-beads coupled with the probe for 2 h at 50°C (Figure 2.3 B).

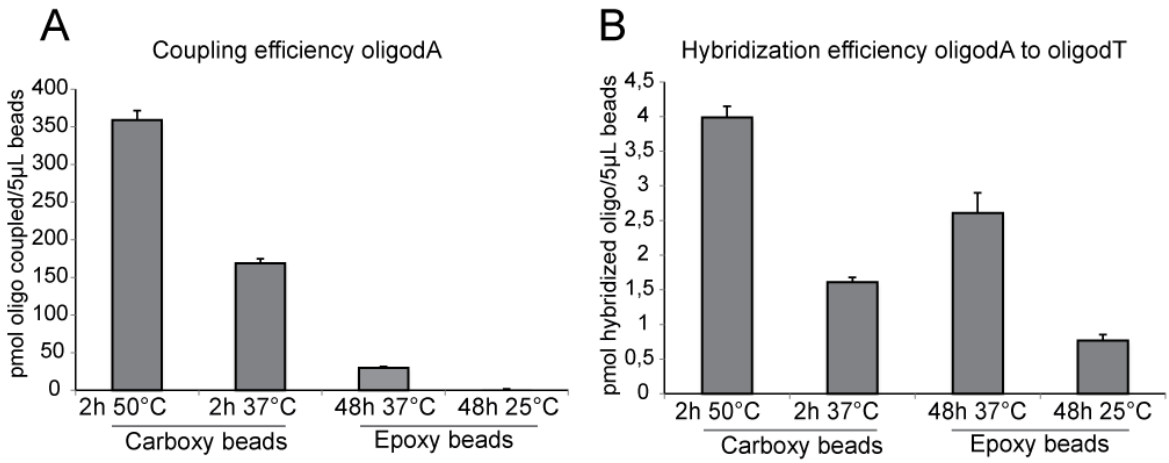

Figure 2.3; Coupling and Hybridization efficiency of carboxy- and epoxy-beads monitored by dual fluorescence assay. Amount of coupled or hybridized oligonucleotide was calculated from a standard. A) Coupling efficiency of dA_{20} with 3' Aminomodifier and Cyanine dye Cy5 at the 5'end; B) Hybridization efficiency of dT_{20} with Cyanine dye Cy3 at the 3'end. n=3 replicates; error bar = standard deviation

To ensure that the different base composition of the oligonucleotide does not change the coupling efficiency to the beads, I also tested the DNA sequences used later to target the *Renilla luciferase* (*Rluc*) open reading frame (ORF) (named LNA1, LNA2) and the scrambled control (LNAscr) for coupling efficiency to carboxy-beads for 2 h and 5 h at 50°C. All probes showed efficient hybridization with their target sequences at 4°C, while they fail to do so with non-complementary sequences (Figure 2.4).

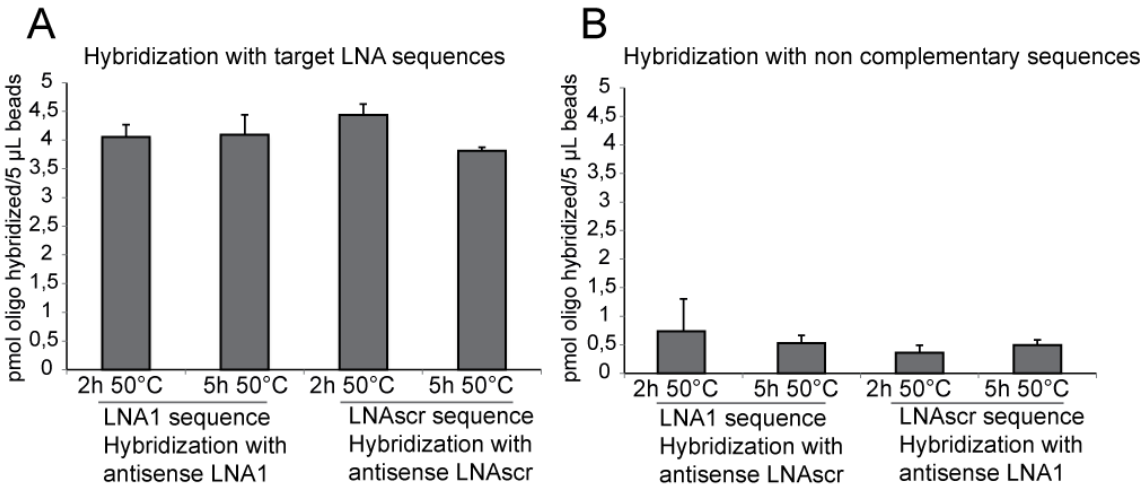

Figure 2.4; Hybridization efficiencies with different LNA probes. Amount of hybridized probe was calculated from a concentration curve of Cy5-labeled target oligo nucleotide. A) Hybridization efficiency of LNA1 and LNAscr with their antisense target sequence; B) Hybridization efficiency of LNA1 and LNAscr with non-complementary sequences; n=3 replicates; error bar represents standard deviation.

2.3 Design of the LNA/DNA mixmer antisense probes

I decided to use a mixture of LNA and DNA nucleotides for a probe of a total length of ~20-22 nucleotides. In the LNA nucleosides the ribose ring is "locked"

by a methylene bridge that connects the 2'O-atom and the 4'-C atom. This bridge forces the conformation equilibrium to an ideal Watson-Crick base pairing conformation. The LNA forms therefore more stable hybrids than RNA or DNA oligonucleotides. For the placement of the LNA nucleotides with respect to the "normal" deoxynucleotides, I followed general design guidelines described in Figure 2.5.

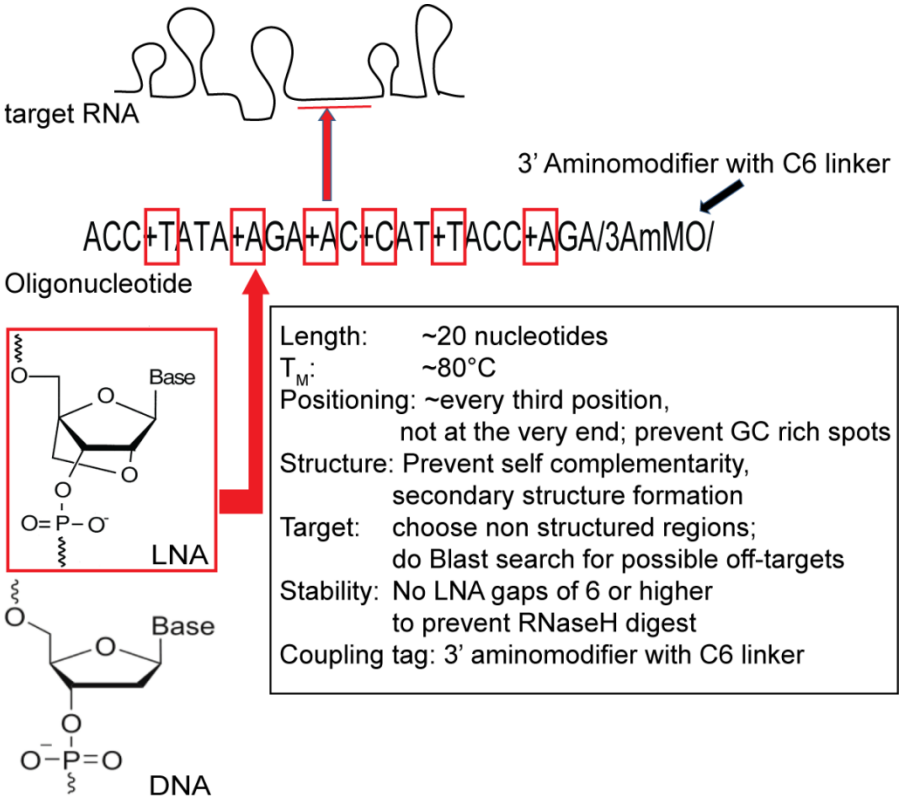

Figure 2.5; Design guidelines for LNA/DNA mixmer oligonucleotide probes. *(Tm:melting temperature)*

Target selection: The seed sequence of the probe on the RNA is selected avoiding sequences present in other transcripts. In the case of few nucleotides exhibiting complementarity with a non-target RNA, I avoided insertion of LNA at these positions. In addition, probes were designed to hybridize to regions lacking stable secondary structures, determined with mfold (Zuker, 2003), a RNA folding prediction software, or published structural data.

Melting Temperature: By incorporating LNA into the oligonucleotide, the melting temperature of the hybrid is increased. Every LNA monomer introduced in the probe can increase the melting temperature from 2 to10°C (Wahlestedt et al., 2000). The increased melting temperature will promote selective hybridization with probes as short as 20 nucleotides. All the probes were

designed to result in a melting temperature of ~80°C (determined by Exiqon Tm prediction tool; https://www.exiqon.com/rna-tm).

Stability: Introducing LNA also enables stabilization of the probe against nucleases. To prevent RNaseH mediated cleavage of DNA/RNA hybrids I ensured to have less than 6 DNA oligonucleotides in a row. A length of 6 or more DNA/RNA hybrids is sufficient for RNaseH cleavage (Kauppinen et al., 2005) as this principle is exploited for the specific degradation of RNA targets using LNA Gapmers (Exiqon). I introduced around 6 LNA monomers within a probe of 20 nucleotides.

Positioning: The LNA was positioned at approximately every third position. As the LNA monomer forces also the neighboring nucleotides to adopt the optimal conformation for base pairing, this design allowed the maximal hybridization competence with the LNA used. Mixing the LNA with DNA nucleotides prevents increased stickiness and excessive rigidity that will result from an oligonucleotide with (too) high LNA content. LNA should not be positioned at the very end to allow better flexibility for hybridization. In addition, GC rich spots should be avoided as these might increase secondary structure or duplex formations of the oligonucleotide.

Coupling tag: To allow covalent coupling, I introduced in the designed probe a 3' amino-modifier which carries a primary amine and flexible C6 linker.

With these principles in mind, I first designed LNA/DNA probes targeting the *Rluc* ORF and, later, targeting human 18S and 28S rRNA.

2.4 Optimization of hybridization temperatures

Next, I determined the optimal hybridization temperature for the *RLuc*-targeting probes (LNA1, LNA2) using the buffer conditions employed in specific RNP capture. I included the scrambled control (LNAscr) in order to see if the experimental noise can be reduced by modulating the hybridization temperature. For this experiment I made use of a cell free system, in which I could easily test several conditions in replicates. I incubated HeLa translationally competent cytoplasmic extracts (Cilbiotech) with *in vitro* transcribed RLuc and control *Firefly Luciferase* (Fluc) mRNAs in equimolar

amounts, adjusted the buffer compositions of specific RNP capture to maximize hybridization and then performed specific RNP capture with the LNA coupled beads. With this set-up I tested a hybridization temperature range for the LNA probes from 40°C - 52.2°C. After pull down, I analyzed RNA purity and relative amounts by RT-qPCR. Beside the target Rluc mRNA, I quantified the specificity control Fluc and the endogenous 18S rRNA. Rluc was strongly enriched in eluates of LNA1 and LNA2 probes, although LNA1 showed the best performance, especially when the hybridization was performed at 41°C (Figure 2.6 A/B). Notably, the efficiency of target RNA capture was not improved by mixing the two oligonucleotide probes (LNA1, LNA2) (Figure 2.6 C). RNA traces derived from LNAscr pulldown resemble the background off-target hybridization of LNA1; LNAscr represents therefore an excellent negative control (Figure 2.6 D) with no specific enrichments (Figure 2.6 E). LNA1 reached the best enrichment of Rluc over the Fluc control of almost 200 fold at 41°C in the test set-up (Figure 2.6 E) and I thus chose LNA1 for the further experiments.

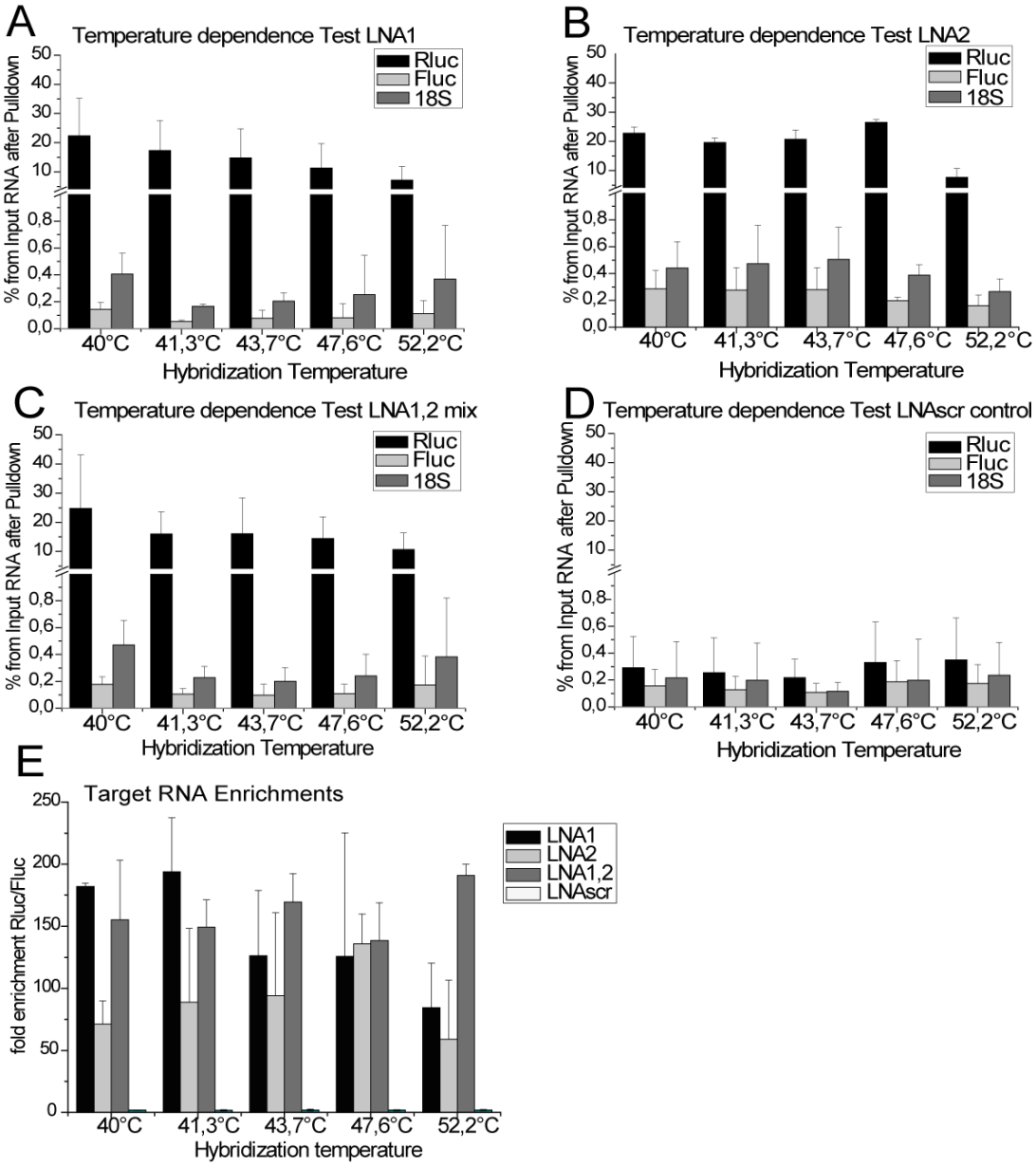

Figure 2.6; Hybridization efficiency and specificity test of LNA probes coupled to beads. Different hybridization temperatures were tested in a gradient PCR machine. LNA1 and LNA2 were designed to target Rluc RNA and LNAscr was used as control. After stringent washes and elution, RNA was analyzed by RT-qPCR. 18S rRNA, Fluc RNA and RLuc RNA were quantified in relation to the Input RNA in %. Error bars represent standard deviation of three experiments. (A) Hybridization of LNA1; (B) Hybridization with LNA2; (C) Hybridization with a mixture of LNA1 and LNA2 D) Hybridization with the control LNAscr. (E) Enrichment of Rluc target RNA over the equimolar added Fluc control for the different applied LNA probes;

Covalent coupling allows re-use of the beads in several rounds of pull-down, since the probe remains attached throughout the whole protocol. I therefore used the hybridization temperature test set-up to monitor the performance of the beads upon recycling the beads for several rounds of pull-down. After adding fresh sample after each round, I observed comparable performance of the beads over three rounds of pull-down (Figure 2.7).

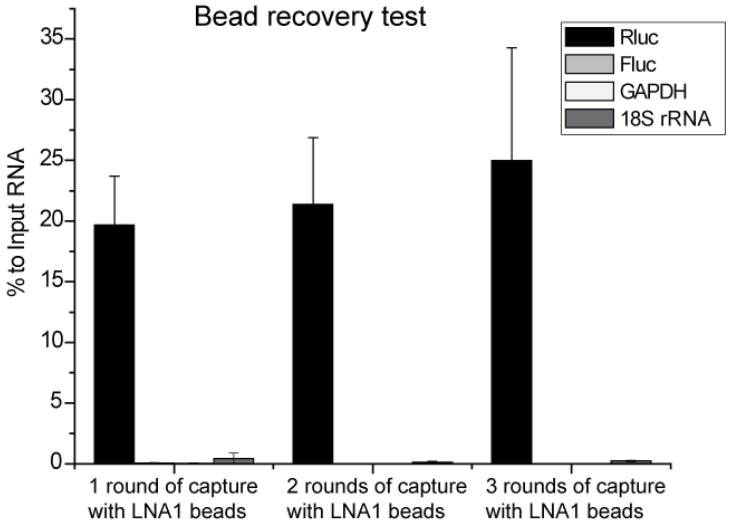

Figure 2.7; Bead recovery test. Beads with probes targeting Rluc (LNA1) were used twice after the first capture. After stringent washes and elution, RNA was analyzed by RT-qPCR. 18S rRNA, Fluc RNA RLuc RNA and GAPDH RNA were quantified in relation to the Input RNA in %. Error bars represent standard deviation of three experiments.

2.5 Titration experiments to optimize specific RNP capture

I next performed titration experiments in order to identify the optimal conditions that allow co-purification of proteins bound to the target RNA. I first optimized the energy of ultraviolet (UV) light at 254 nm required for efficient crosslinking in cell extract. I applied conventional crosslinking (cCl) at 254 nm together with a non-crosslinking (NocCl) control in order to visualize proteins directly bound to the enriched mRNA after RNA interactome capture. The used cell extract cannot be plated akin to a thin monolayer of cells, therefore the UV light penetrance and crosslinking efficiency is different. I observed that UV light energies above the previously used dose of 150 mJ/cm^2 (Castello et al., 2012) increase crosslinking efficiency in HeLa cytoplasmic extracts (Figure 2.8 A). However, RNA integrity is compromised as the UV radiation energy increases (Figure 2.8 B and C. 300 mJ/cm^2 of UV light appears as the optimal dose, because it induces efficient crosslinking with the minimal effect on RNA integrity.

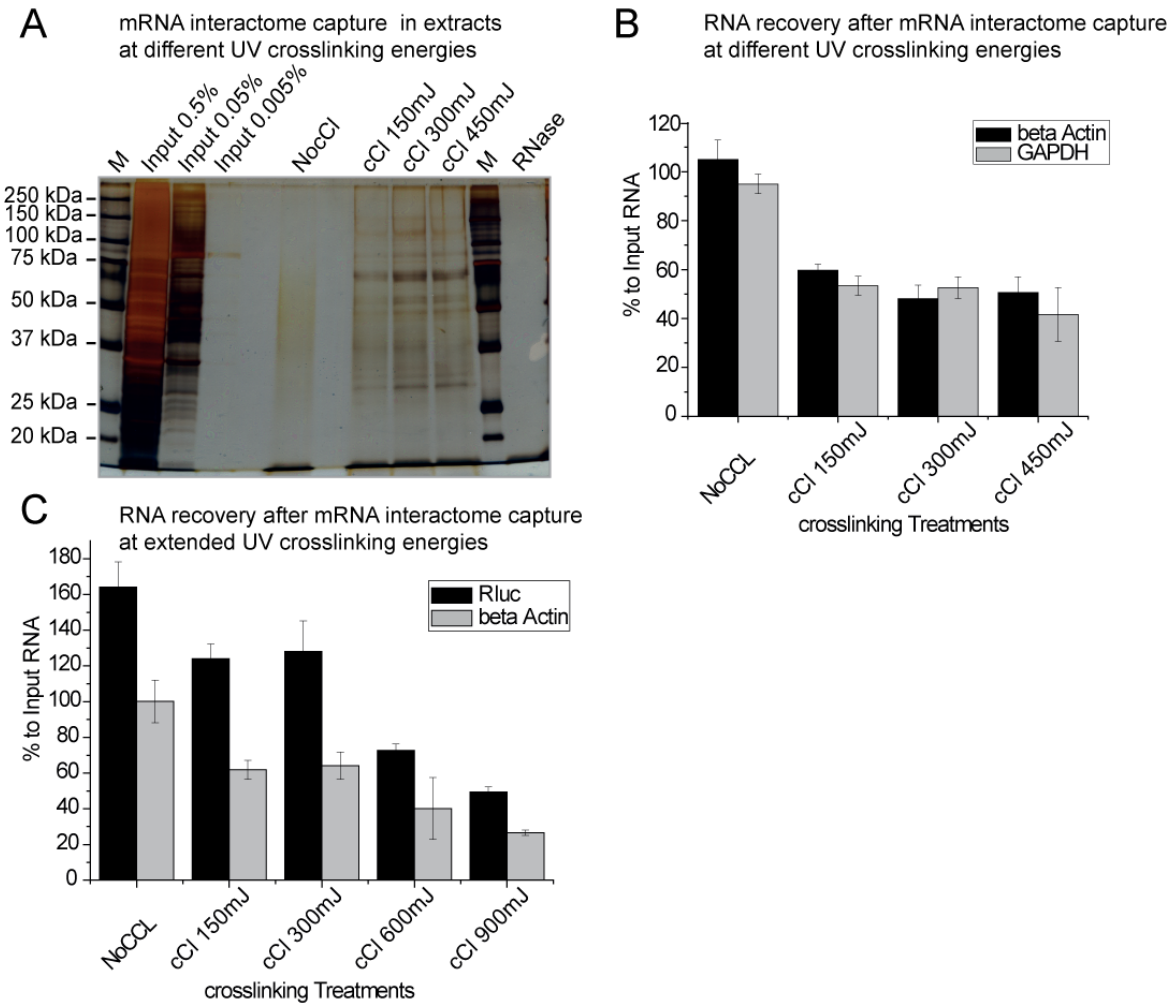

Figure 2.8; Titration of the UV dose required for efficient crosslinking. (A) HeLa cytoplasmic extracts were crosslinked at 254nm using different UV energies. Proteins crosslinked to oligodT-captured RNA were visualized via Silver Staining. (B) RNA integrity upon irradiation with different UV energies; HeLa cytoplasmic extracts were UV irradiated with different energy doses. RNA integrity was tested by analyzing the eluted RNA after oligodT capture by RT-qPCR. Quantified RNAs were displayed relative to the respective Input RNA. (C) as B with extended UV crosslinking energy doses.

Subsequently, I determined the minimal amount of RNA required to detect co-purified RBPs in HeLa cell extract and from living HeLa cells, using RNA interactome capture followed by Silver Staining. After UV irradiation (300 mJ/cm^2) the minimal amount of RNA required to obtain a detectable protein pattern on a silver stained gel was 65 µg of total poly(A)+ RNA in HeLa extracts (Figure 2.9 A). No proteins were detected when lower amounts of RNA were used. In contrast, a complex protein pattern was already detectable after purification of 2-20 µg of total poly(A)+ RNA from living HeLa cells, reflecting a higher UV crosslinking efficiency (Figure 2.9 B). To gain insight into the detection limits of MS in these samples, proteins purified after applying RNA interactome capture to HeLa cells showed that the number of peptide identifications using 0.5 µg RNA was similar to background detection of

peptides from RNA in non-irradiated control cells (Figure 2.9 C). A larger number of peptides was identified with 2 µg of RNA from UV-irradiated cells (~300 peptides), implying that this is a more suitable amount of material to obtain a comprehensive RNA interactome of specific RNA species.

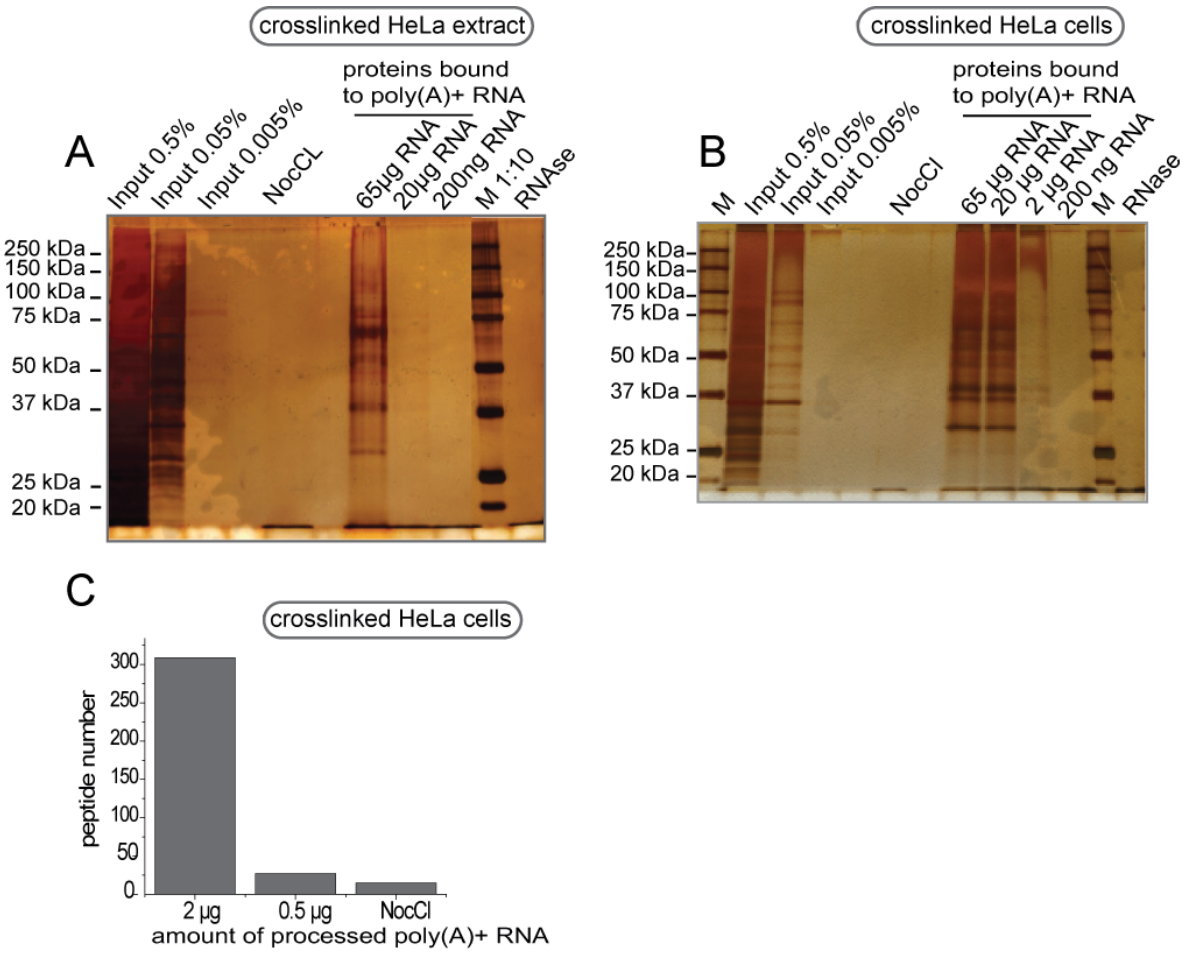

Figure 2.9; Titration of the RNA amount required for detection of RBPs. (A) HeLa cytoplasmic extracts were crosslinked at 254 nm under the chosen energy of 300 mJ/cm^2. Different amounts of RNA eluted from RNA interactome capture applied to HeLa extracts was processed for protein analysis and visualized via Silver Staining. (B-C) HeLa cells were crosslinked at 254 nm with 150 mJ/cm^2 *in vivo*. The amount of eluate after oligo(dT) capture was processed for protein analysis and visualized via Silver Staining; (B) and Mass spectrometry (C). For the NocCl negative control eluate from 14 µg captured RNA was loaded; NocCl = non crosslinked control; cCl = crosslinked sample (conventional crosslinking performed). Mass spectrometry analysis was performed by Christian Frese.

Taken together, these results suggest that UV irradiation of cell lysates *in vitro* is less efficient than on cell monolayers and that at a lower limit of approximately 2µg of isolated RNA is expected to be required for specific RNP capture from HeLa cells.

2.6 Specific RNP capture “rediscovers” Sxl as binding partner of msl-2 mRNA

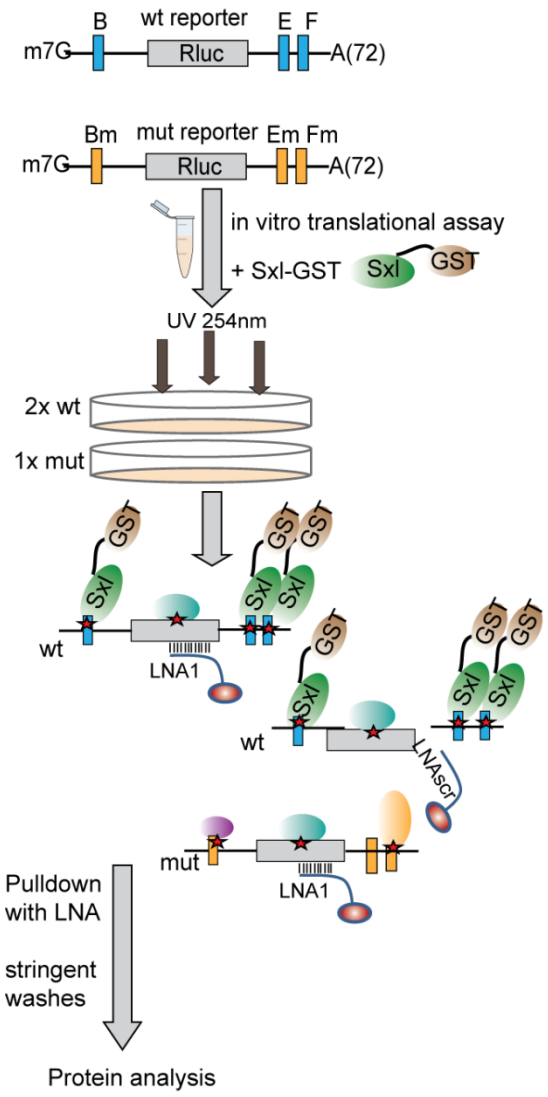

Figure 2.10; Scheme of specific RNP capture *in vitro* with recombinant Sxl protein. HeLa translational active extracts are incubated and UV crosslinked with recombinant GST tagged Sxl protein and Rluc reporter RNAs harbouring Sxl binding sites (wt) or harbouring the mutated sites. Specific RNP capture was performed with LNA1 targeting the Rluc ORF and LNAscr control.

To get further insights into the performance of specific RNP capture, I performed *in vitro* pull-down experiments in a well-defined system derived from *Drosophila*, the interaction between the protein Sex lethal (Sxl) and poly(U) stretches of male specific lethal-2 (msl-2) RNA. I introduced the well-characterized Sxl target sequences upstream and downstream of the Rluc reporter ORF following the binding site architecture of the endogenous msl-2 mRNA (scheme in Figure 2.10; Gebauer et al., 2003). This construct harbors two motifs of 7 nucleotide long U stretches in the 3’UTR, referred to as E and F site. Furthermore, a 16 nucleotide long U stretch is present in the 5’UTR of this RNA, which is called B site. A mutant version of these sites known to prevent Sxl-binding was generated by mutating every second U to a C in all binding

sites of E, F and B (Gebauer et al., 2003). Considering the previous analyses, I used 65 µg (200 pmol) of each of wild-type (wt) and mutant (mut) reporter mRNA in order to detect bound proteins by silver staining and western blot analysis. I added four times molar excess of the recombinant Sxl (dRBD4)-GST protein to the mixture and started *in vitro* translational reactions with either the wt or the mut reporter RNA in HeLa cytoplasmic extracts. Afterwards the extracts were irradiated with 300 mJ/cm^2 of UV light (254 nm) and specific RNP capture was performed with a probe targeting the Rluc ORF (LNA1) or with a control probe (LNAscr).

Notably, silver staining and western blotting analyses revealed that the Rluc-specific LNA1 probe successfully enriched the GST-tagged Sxl from HeLa extracts only when the wt reporter mRNA was used (Figure 2.11 A/B). In contrast, the LNAscr control failed to isolate Sxl, highlighting the selectivity of the LNA/DNA mixmer probes. Moreover, mutations that abrogate Sxl-binding to the reporter also prevented purification of the exogenous protein using LNA1 probe. Notably, I observed Sxl binding to endogenous mRNAs in both wt and mut mRNA supplemented extracts when the complete poly(A)+ RNA pool was purified with oligo (dT). These results reflect that other mRNAs present in the extract do bear analogous U-rich sequences to those present in the reporter and can effectively be bound by Sxl. Moreover, these results demonstrate that the exogenously added Sxl is active in both loaded samples with wt or mut mRNAs. Thus, the lack of binding of Sxl to mut mRNA in pulldowns with LNA1 cannot be explained by a dysfunction of the protein. Taken together these results, I can conclude that Sxl 1) binds to the wt reporter mRNA; 2) does not bind the exogenous RLuc mRNA when mutations are introduced in its binding site and 3) Sxl can interact with other mRNAs within the extract. Therefore, Specific RNP capture successfully recapitulated data generated in the past by orthogonal approaches.

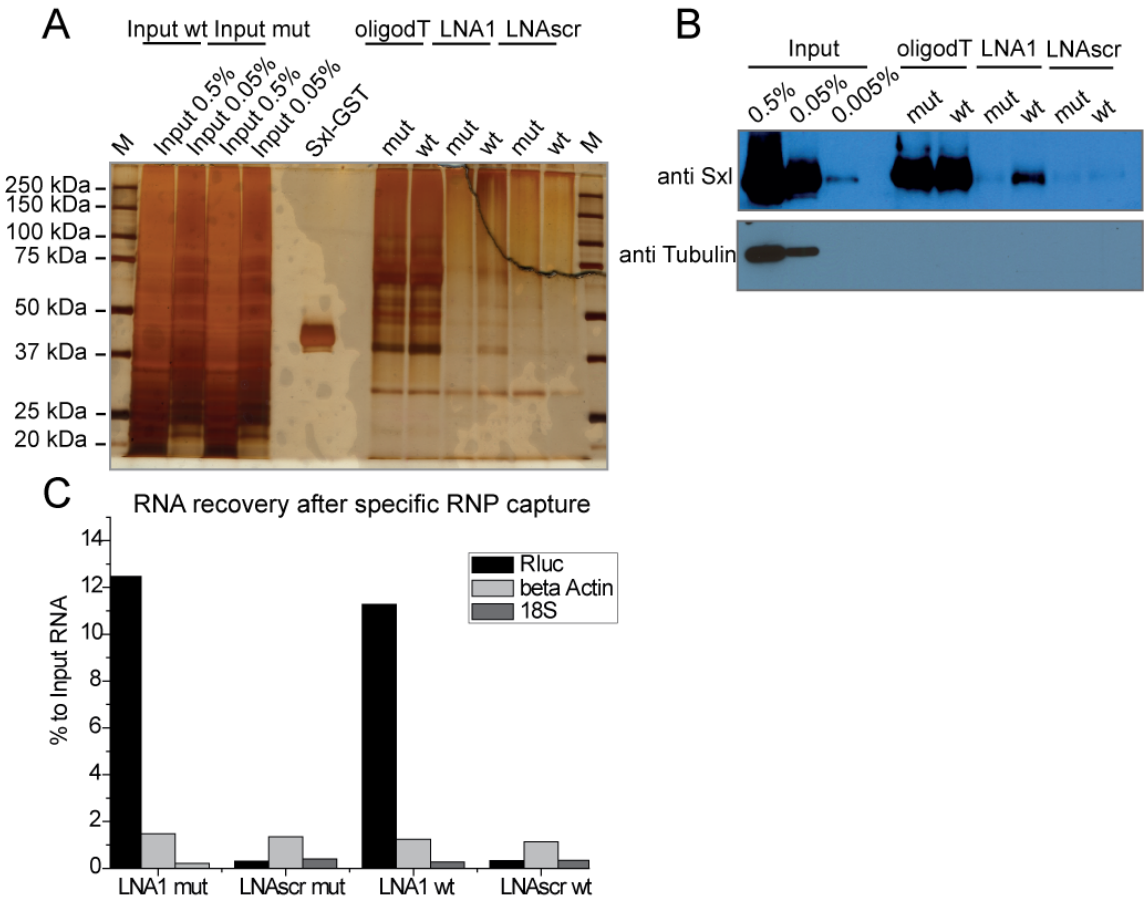

Figure 2.11; Capture of recombinant Sxl bound to the reporter RNA in HeLa cell extracts. (A) Proteins were separated on a SDS-PAGE and detected via Silver Staining. The recombinant protein (Sxl-GST) was loaded on the gel as a running control. (B) Western Blot result confirmed the silver staining results with regard to Sex lethal specific pull-down via the wt reporter RNA with the specific LNA1. The contaminant Tubulin could not be detected in either of the pull-downs.(C) RNA analysis of the specific pulldowns with LNAs. The quantified RNAs are displayed relative to the Input used for the specific pulldown protocol. The pulldown levels of the wt and mut reporter are specific and comparable.

2.7 Identification of proteins bound to an *in vitro* transcribed mRNA in *Drosophila* embryo extracts

To follow up the encouraging results for specific RNP capture with exogenously added protein and mRNA, I wanted to determine if specific RNP capture can be applied to endogenous proteins using *D. melanogaster* embryo extracts supplemented with the previously used reporter mRNA. This time, I combined specific RNP capture with quantitative mass spectrometry to identify the proteins bound to the RNA of interest (Figure 2.12). By using dimethyl labelling, three samples were labelled 1) pulldown of wt or 2) mut RLuc mRNA captured with LNA1 and the 3) control pulldown using wt mRNA and LNAscr probe. The labeling was coupled to a recently developed protocol for efficient sample preparation when protein material is low, called Single-Pot Solid-Phase-enhanced Sample Preparation (SP3) (Hughes et al., 2014). Comparison of

LNA1 vs LNAscr is applied to determine the proteins that bind to the RLuc mRNA irrespectively of Sxl binding sites, while comparison between LNA1 targeted wt reporter RNA and LNA1 targeted mut reporter RNA will serve to identify the alterations in the RNP promoted by the presence of Sxl binding sites. I performed specific RNP capture for these set-ups in four replicates for stringent statistically evaluation of the data.

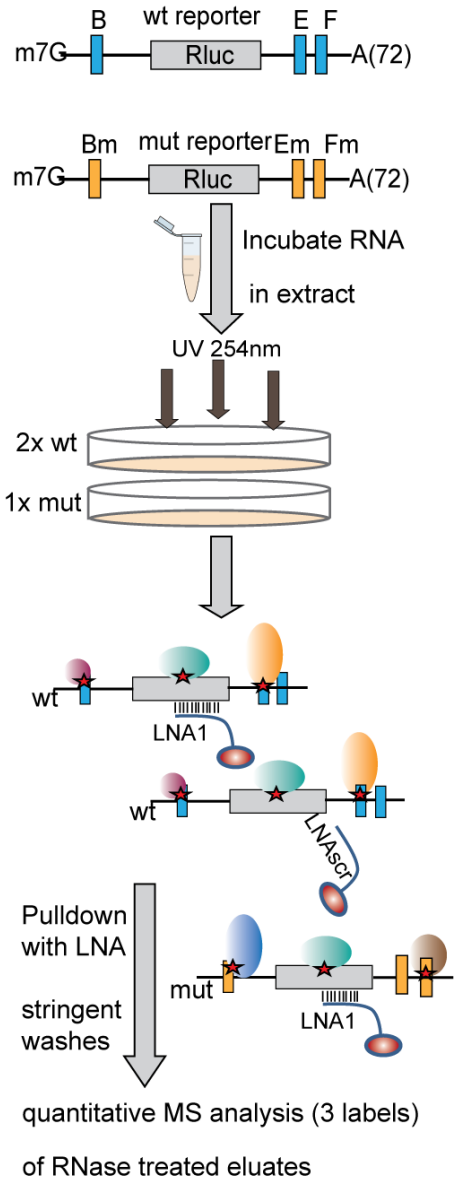

Figure 2.12; Scheme of specific RNP capture *in vitro*. Translational active *Drosophila* embryo extracts are incubated and UV crosslinked with Rluc reporter RNAs harbouring Sxl binding sites (wt) or harbouring the mutated sites. Specific RNP capture was performed with LNA1 targeting the Rluc ORF and LNAscr control. After RNase treatment proteins are analyzed via quantitiative mass spectrometry, that allows the comparison of 3 labels.

LNA1 isolated both wt and mut Rluc mRNAs to a similar extent (Figure 2.13 A). As in the previous experiment, the LNAscr control was not enriching specifically for any RNA and resembled the background capture of RNA performed by LNA1.

Proteomic analysis revealed that capture with LNA1 yields a notable enrichment of peptides over LNAscr (Figure 2.13 B; Figure S 7.1). These results show that peptide purification relies on the capture of the RLuc mRNA. I can therefore conclude that the purification of proteins is not due to unspecific binding of proteins to the bead surface or to the LNA itself. 163 proteins could be identified via specific RNP capture, 109 of them could be quantified in wt/mut samples (Table S 7.1). However, I only observed minor differences between the wt and the mut reporter using LNA1 for specific pull-down (Figure 2.13 C). These alterations very well reflect the small differences in primary sequence between the wt and the mut reporter mRNAs. Strikingly, specific enrichment of sister of sex lethal (Ssx), Heterogeneous nuclear ribonucleoprotein at 98DE (Hrb98DE, Hrp38) and Heterogeneous nuclear ribonucleoprotein at 87F (Hrb87F, Hrp36) in the wt over the mut samples was reproducibly observed. Unexpectedly, Ssx but not its paralog Sxl, which was previously described to bind the U-rich sequences introduced in the reporter RNA, was identified as specific binder of the wt mRNA. This protein has high homology with Sxl, but its RNA-binding activity was only recently reported in the RNA interactome capture studies performed in *D. melanogaster* embryos (Wessels et al., 2016; Sysoev et al., 2016). The identity score between these two proteins is especially high across the two RNA recognition motifs (RRMs) as well as in the glycine rich region in the N-terminus (Figure S 7.2). Supporting the idea of Ssx RNA binding activity, I purified Ssx bound to poly(A) RNA by RNA interactome capture followed by Western blot (Figure 2.13 D). Taken together, these results show that Ssx is an RBP that displays similar specificity for U-rich sequences as its paralog Sxl.

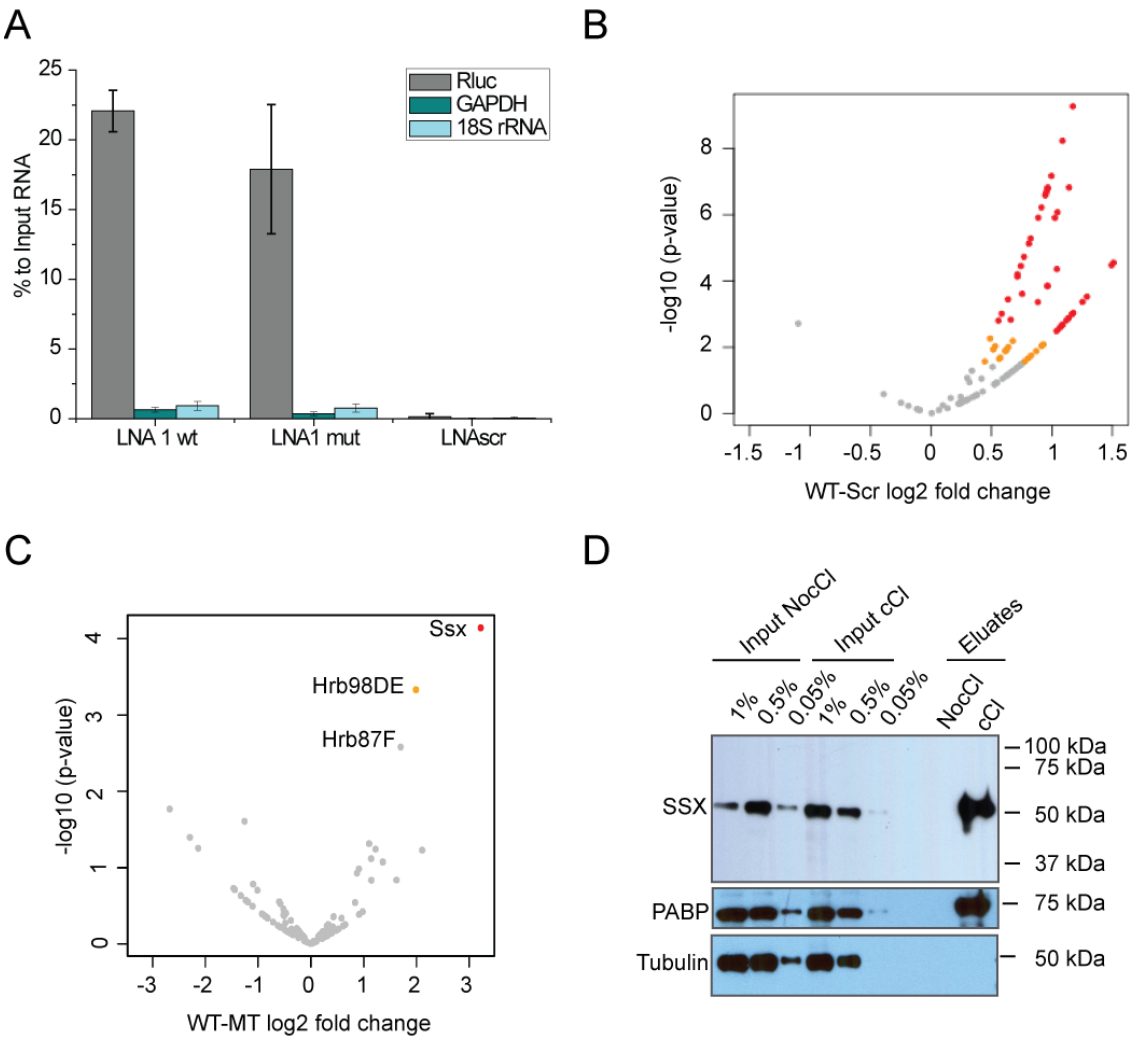

Figure 2.13; Specific RNP capture from *Drosophila* embryo extract. Mass spectrometry measurements were performed by the EMBL proteomics core facility (A) RNA analysis of the specific pulldowns with LNAs. The quantified RNAs are displayed relative to the Input used for the specific pulldown protocol. The pulldown levels of the wt and mut reporter are specific and comparable. The error bars represent standard deviation coming from four biological replicates (B) Vulcano plot comparing the enrichments in the wt sample over the scramble sample in 4 biological replicates. The axes display the log2-fold changes (orange: 5% FDR, red: 1% FDR). Statistical analysis and plot was created by Bernd Fischer (C) Vulcano plot comparing the enrichments in the wt sample over the mut sample in 4 biological replicates. The axes display the log2-fold changes (orange: 5% FDR, red: 1% FDR). Statistical analysis and plot was created by Bernd Fischer (D) Western Blot with antibodies raised against Sister of Sex lethal (Ssx),Tubulin and PABP after oligodT capture of crosslinked Drosophila embryo extract. Ssx was strongly enriched in the crosslinked eluate, the negative control Tubulin was not detected in the eluates.

Sxl and UNR (normally co-recruited by Sxl to perform translational repression) could not be detected in MS and were therefore also not enriched in the wt sample, as expected. In order to understand the absence of Sxl and UNR in the specific RNP interactome, the EMBL proteomics core facility analyzed the whole proteome of the embryo extracts for this study. While Ssx was identified with a medium intensity score, Sxl and UNR were not detected. I can therefore conclude that Sxl and UNR are present at very low concentrations in the embryo extracts used in this work and is, most probably, outcompeted by its

ubiquitously expressed paralog Ssx. This can be due to the use of male/female mixed extracts or the developmental stage of the embryos that were collected.

Hrb87F and Hrb98DE are splicing factors with high homology between them in both sequence and structure (Haynes et al., 1991; Blanchette et al., 2009). They were enriched from wt RLuc reporter over its mutant version. Sxl molecules can interact with Hrb87F, most likely through a common glycine-rich region, and increases the affinity of Hrb87F for U-rich sequences (Wang et al., 1997). Ssx shares the N-terminal glycine rich track with Sxl, suggesting that Ssx can also recruit Hrb87F. These results imply some degree of functional redundancy between the two paralogs. The biological role of Ssx is not well understood, its similarity to Sxl suggests an involvement in translational or splicing regulation. As my results suggest that both proteins recognize similar RNA signatures, it will be of interest to determine if both proteins bind the same mRNA targets. This is especially important since Sxl plays an essential function in sex determination and dosage compensation, whereas there is no evidence that its partially redundant paralog Ssx does.

2.8 Identification of proteins bound to human 18S and 28S rRNA

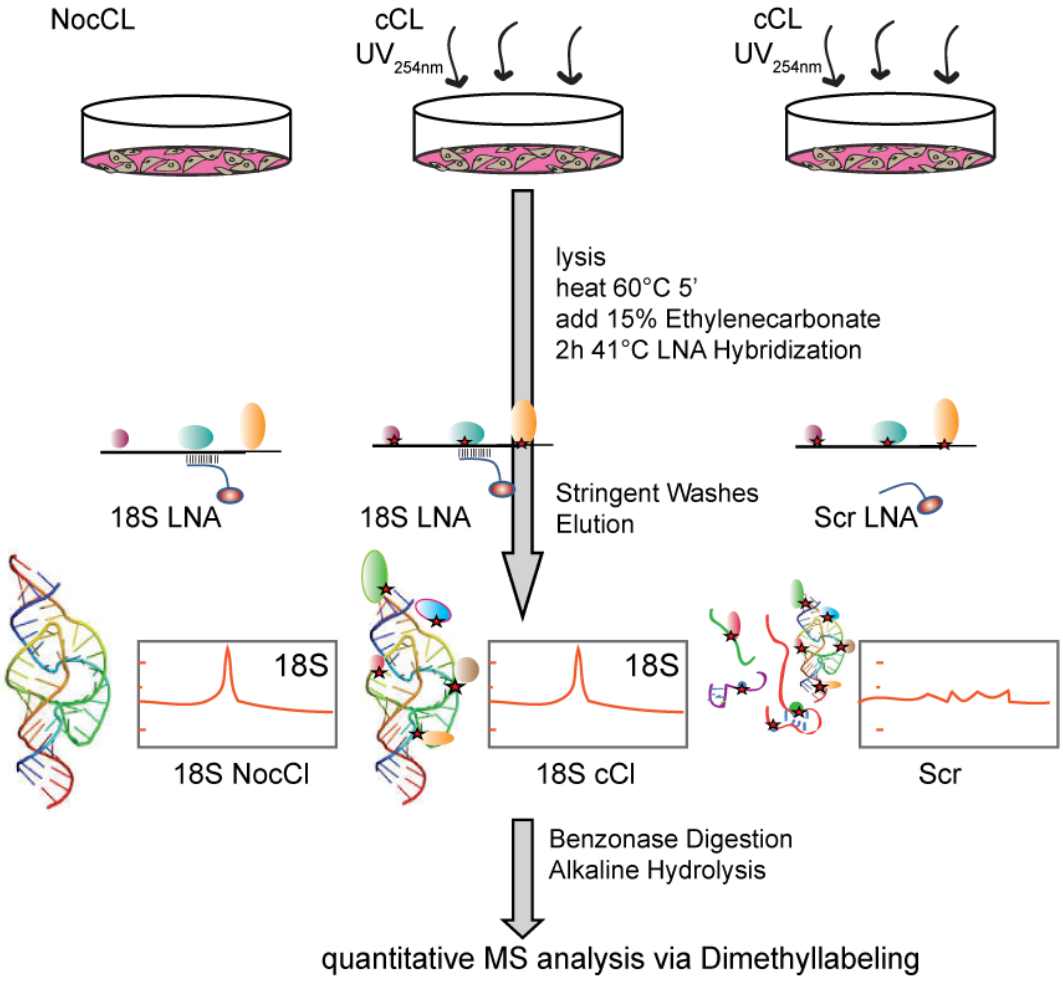

Figure 2.14; Scheme of specific RNP capture of 18S rRNA from HeLa cells. HeLa cell are UV irradiated *in vivo*, followed by lysis and short heat treatment at 60°C to unfold the RNA. Ethylenecarbonate is added as hybridization enhancer. Specific RNP capture is performed with LNA/DNA mixmer probes targeting 18S rRNA (18S LNA) in crosslinked and non crosslinked sample. Another control is the capture with the control probe (LNAscr) in the crosslinked sample. After elution, RNA composition is analysed using bioanalyzer. After Benzonase digestion and alkaline Hydrolysis, released proteins are analyzed via quantitative Mass spectrometry using dymethyl labeling labels. NocCl= non crosslinked control; cCl= crosslinked sample (conventional crsllinking performed)

2.8.1 Protocol adjustment of specific RNP capture protocol for the rRNA interactome

To finally extend this method to in-cell applications, I performed specific RNP capture for human 18S and 28S rRNA (exemplified on 18S rRNA in Figure 2.14). I targeted regions of the ribosomal RNA that were predicted to exhibit a minimal degree of secondary structure. As these probes hybridize specifically to the expansion segment ES6S in 18S rRNA and ES27L in 28S rRNA, it will be

possible to adapt Specific RNP capture to study these expansion segments in more detail.

As dimethyllabeling allows the use of three labels, I could include two controls to the capture of the ribosomal RNP: i) a non-crosslinked sample with the same ribosomal targeting probe that should be devoid of proteins, and ii) a LNAscr probe applied to a crosslinked sample that should not enrich for ribosomal RNA. By including these two controls, results are corrected against both purification of non-covalently bound proteins and the signature of unspecific RNA capture, which increases data quality.

I had to adjust the initial protocol to the pull-down of ribosomal RNAs. As the ribosomal RNAs are highly structured, I tested the chemical ethylenecarbonate that was used previously in in situ hybridizations as a non-toxic substitute for formamide (Matthiesen & Hansen, 2012). This compound allows faster probe-RNA hybridization by lowering the melting temperature of the target RNA and thereby helps to remove RNA structures. Addition of 15% ethylenecarbonate to the lysate increased the hybridization efficiency for ribosomal RNA (Figure S 7.3 A) and the structured *Drosophila* mRNA *oskar* (Figure S 7.3 B). Previous experiments showed that the standard RNase digestion before MS analysis was inefficient in the case of rRNA due to its high and stable secondary structure. Unprocessed rRNA caused precipitation, which produced technical dysfunction of the mass spectrometer by blocking the capillary/column. To circumvent this problem, I added to the RNase digestion a step of alkaline hydrolysis, which enabled efficient processing of the rRNA. The scale of the experiment was another adjustment that was required for the ribosomal RNA interactome capture. Based on previous tests (Figure 2.9 B), I captured around 2 µg RNA in the first experiments; however, the number of identified proteins by mass spectrometry was low. Therefore, I scaled-up the experiment accordingly to capture an amount of 4-9 µg of RNA from the crosslinked samples.

2.8.2 Specific enrichment of 18S and 28S rRNAs.

With the designed LNA/DNA mixmer probes, I achieved specific enrichments of both the 18S and 28S rRNA in a selective way. Bioanalyzer traces of the RNA revealed an enrichment of the rRNA matching the used probe and depletion of the respective other rRNAs (Figure 2.15). The LNAscr control defined the

background signal that was similar to the background detected for the rRNA pull-downs using specific probes. Nevertheless, the capture of rRNA is not very efficicent (1-5% of the total) with the LNA/DNA mixmer probes (Figure S 7.4). A possible explanation is unexpectedly strong secondary structure formation or the induction of RNA-RNA crosslinks that may affect hybridization with the probe.

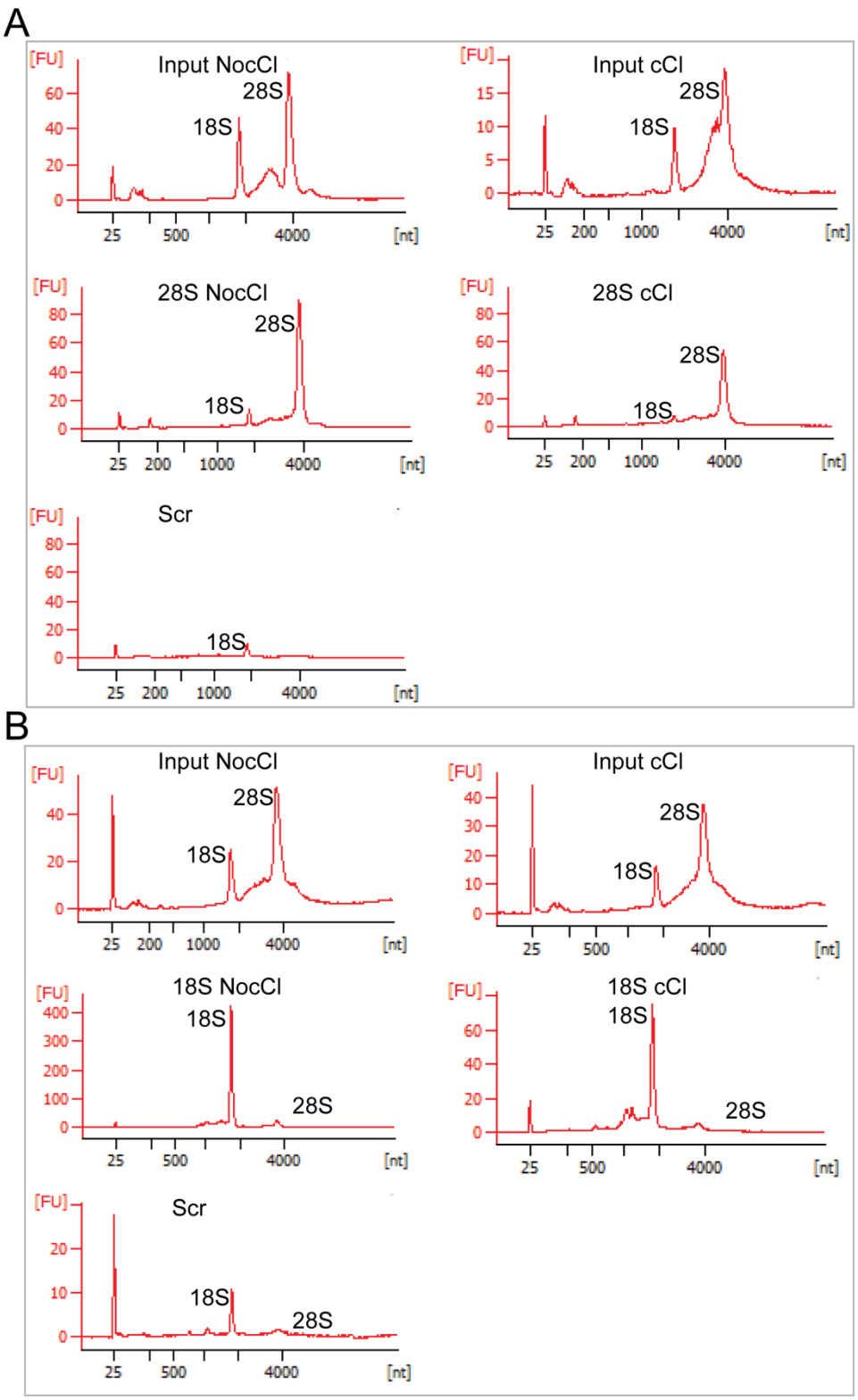

Figure 2.15 RNA analysis after specific RNP capture of 28S rRNA and 18S rRNA. (A) Electropherograms of RNAs coming from Bioanalyzer measurements from one representative 28S rRNP capture replicate; After UV crosslinking (Ccl) or non-crosslinking (NocCl) of Hela cells the specific RNA capture protocol was performed with oligonucleotides targeting 28S or with scrambled probe as control. Released RNAs were analyzed using a 6000 Pico bioanalyzer. The 2 peaks from the Input samples represent 18S (runs at 2000 nt) and 28S (runs at 4000 nt) rRNAs. The y-axes display the fluorescence intensity coming from a dye labeling the RNA quantitatively. (B) Electropherograms of RNAs from one representative 18S rRNP capture replicate (as A).

2.8.3 Protein identification in the rRNA interactomes.

In the quantitative mass spectrometry measurements a reproducible strong enrichment of proteins on the experimental over the controls could be observed. The fact that most proteins could only be detected in the crosslinked specific capture sample, did not allow an analysis that is limited on ratio formation to the controls. Therefore, a semi-quantitative method was implemented that considers the presence or lack of peptide intensity signal for each label across biological replicates. When the MS signal was too low to assign a protein to a specific label, the protein could not be included into the analysis. A stringent cut-off was chosen; all genes were counted that were quantified in at least two more cCL samples than in control samples. Although only a limited amount of proteins could be detected, the data could show the expected enrichments for ribosomal proteins in the 18S (Table 2.1) and 28S (Table 2) rRNA interactomes. Moreover, the 18S and 28S rRNA interactomes display the expected differences regarding the known interaction of each rRNA with the protein components of the large and small ribosomal subunit. Most identified proteins were either part of the cytoplasmic ribosome, were associated with ribosomal function or involved in ribosomal biogenesis. Interestingly, many heterogenous ribonucleoproteins (HNRNPs) were identified in both data sets, suggesting a role of these proteins in the biogenesis of ribosomal RNA. When considering proteins enriched only in one replicate (low confidence binders) the amount of HNRNP proteins increased substantially (Table S 7.2;Table S 7.3). String cluster analysis of high confidence hits (Figure 2.16) showed a strong interconnection between the proteins enriched in both or either datasets. This strong interconnection supports the specific enrichments of protein complexes associated to rRNAs.

Table 2.1 Proteins enriched from 18S rRNA after specific RNP capture. Mass spectrometry measurements were performed by the EMBL proteomics core facility; Bernd Fischer did the semi-quantitative analysis of the three replicates. The amount of enrichments (n) in crosslinked sample (cCl) was compared to the non crosslinked (NocCl) and scrambled control (Scr). The estimated false discovery rate was calculated with following formula: (0+1)/n; being 7.1% for the listed proteins; Proteins were sorted according to the knowledge about relation to ribosome biology including cytoplasmic ribosome constituents, knowledge from human ribosome biogenesis screens or literature reports; Three recent ribosomal biogenesis screens were taken into consideration: Wild et al., 2010 (siRNA depletion based biogenesis screen of the 40S and 60S ribosomal subunit, with microscopic readout); Tafforeau et al., 2013 (siRNA depletion based rRNA maturation screen; with Northern blot readout); Badertscher et al., 2015 (genome wide siRNA depletion based biogenesis screen of the 40S ribosomal subunit; with microscopic readout.) Red: ribosomal proteins from the large ribosomal subunit;

	Symbol	nCCL	nNoCL	nScr	Source Literature
part of the 40S ribosomal subunit	RPSA	2	0	0	
	RPS24	3	0	0	Wild et al. 2010; Badertscher et al. 2015
	RPS15	3	0	1	Badertscher et al. 2015
involved in ribosomal Biogenesis	NCL	3	0	1	Tafforeau et al. 2013
	HNRNPF	3	0	0	Tafforeau et al. 2013
	SERBP1	3	0	1	Tafforeau et al. 2013
	RPL8	3	0	0	Wild et al. 2010; Badertscher et al. 2015
Ribosome association	EIF3J	2	0	0	Valásek et al. 2001; Fraser et al. 2003
	RPL22	3	0	0	
no described cytoplasmic ribosome related function	HNRNPA1	2	0	0	
	HNRNPH1	2	0	0	
	YBX3	2	0	0	
	ELAVL1	2	0	0	
	LRPPRC	2	0	0	

Table 2.2 Proteins enriched from 28S rRNA after specific RNP capture. Mass spectrometry measurements were performed by the EMBL proteomics core facility; Bernd Fischer did the semi-quantitative analysis of the three replicates. The amount of enrichments (n) in crosslinked sample (cCl) was compared to the non crosslinked (NocCl) and scrambled control (Scr). The estimated false discovery rate was calculated with following formula: (0+1)/n; being 5.3% for the listed proteins; Proteins were sorted according to the knowledge about relation to ribosome biology including cytoplasmic ribosome constituents, knowledge from human ribosome biogenesis screens or literature reports; Three recent ribosomal biogenesis screens were taken into consideration: Wild et al., 2010 (siRNA depletion based biogenesis screen of the 40S and 60S ribosomal subunit, with microscopic readout); Tafforeau et al., 2013 (siRNA depletion based rRNA maturation screen; with Northern blot readout); Badertscher et al., 2015 (genome wide siRNA depletion based biogenesis screen of the 40S ribosomal subunit; with microscopic readout.)

	Symbol	nCCL	nNoCL	nScr	Source Literature
part of the 60S ribosomal subunit	RPL22	3	0	0	
	RPL29	2	0	0	
	RPL31	2	0	0	
	RPL5	3	0	0	Wild et al. 2010; Badertscher et al. 2015
involved in ribosomal Biogenesis	RPL4	2	0	0	Wild et al. 2010
	PA2G4	3	0	0	Tafforeau et al. 2013
	NCL	3	0	0	Tafforeau et al. 2013
	HNRNPF	3	0	0	Tafforeau et al. 2013
	HNRNPC	2	0	0	Tafforeau et al. 2013; Badertscher et al. 2015
	HNRNPD	2	0	0	Tafforeau et al. 2013
Ribosome association	BTF3	2	0	0	Beatrix et al. 2000
no described cytoplasmic ribosome related function	HNRNPA1	3	0	0	
	HNRNPH1	2	0	0	
	HNRNPA2B1	2	0	0	
	HNRNPM	2	0	0	
	YBX3	3	0	0	
	ELAVL1	2	0	0	
	PLIN4	2	0	0	
	DSG1	2	0	0	

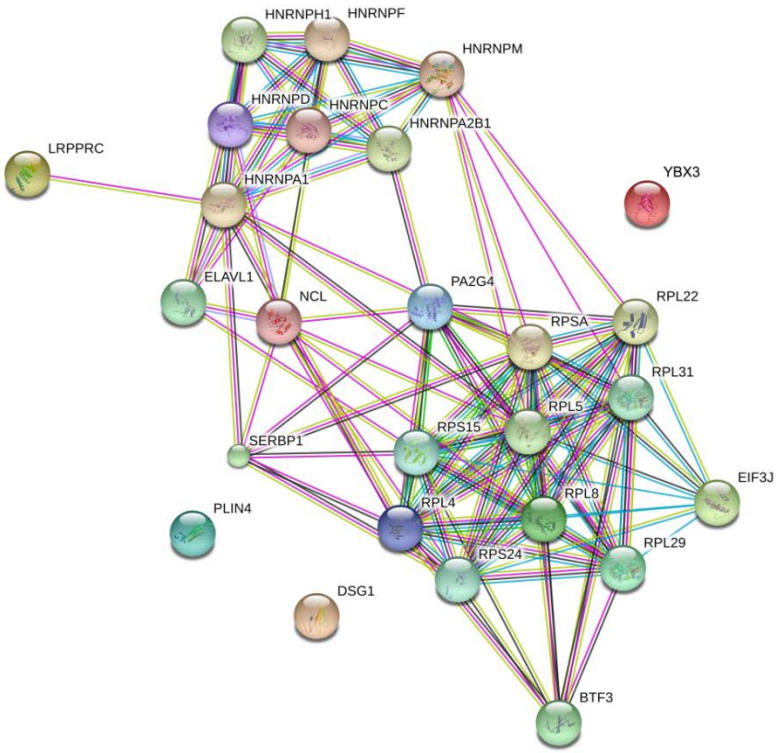

Figure 2.16; String cluster analysis (string-db.org) with the proteins enriched in 18S and/or 28S rRNA specific RNP captures. The enriched proteins are highly interconnected and form clusters of ribosomal proteins and HNRNP proteins.

The two rRNA interactomes also overlap in proteins, which were not described before to be part of the ribosome or being associated with the rRNA (YBX3, ELAVL1, HNRNPA1,HNRNPH). The overlap suggests a general role of these proteins in rRNA biology. These proteins have multifunctional roles as RBPs in the cell; they could therefore be general RNA binders. As some proteins of the HNRNP family were identified in recent ribosomal biogenesis screens (Tafforeau et al., 2013, Badertscher et al., 2015), a functional role of HNRNPA1 and HNRNPH in rRNA biogenesis could also be taken into consideration. YBX3 is interacting with many RNAs and is involved in a multitude of biological functions. YBX3 hasn't been described before to be involved in ribosomal biogenesis or being associated with the ribosome. It would be therefore interesting to see, which role YBX3 would have specifically in ribosome biology. ELAVL1 (also known as HUR) is a RBP known to interact with AU rich elements (AREs) (Herdy et al., 2015). An interesting characteristic of ELAVL1 RNA binding has been described in embryonic stem cells (ESCs) differentiation: In this biological context ELAVL1 preferentially binds mRNAs that do not harbour N6-methyladenosine (m6A) sites and stabilizes those specifically (Wang et al. 2014). As human 18S and 28S rRNAs carry m6A modifications (positions 1832,

1851 for 18S rRNA and 4189 and 4190 for 28S rRNA) (Motorin & Helm 2011), ELAVL1 could also bind to rRNA dependent on the modification status of the rRNA during biogenesis.

2.8.4 Detection of inter-subunit and subunit-specific ribosomal proteins.

After having a closer look on the distribution of the enriched ribosomal proteins on the human ribosomal structure (Figure 2.17 A/B), I could confirm that the identified proteins are covering all regions of the respective ribosomal subunits. Therefore, I could exclude any targeting region biases. Biases could have evolved if the rRNA would have been partially disrupted during specific RNP capture, although I could not observe any rRNA integrity loss on the RNA bioanalyzer traces.

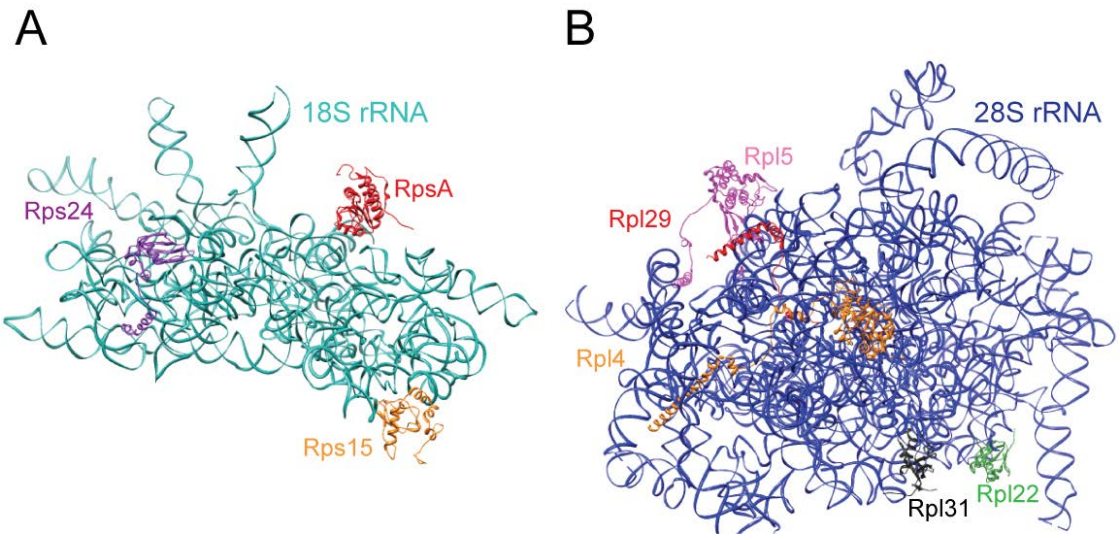

Figure 2.17; Position of enriched ribosomal proteins in the structure of the cytoplasmic ribosome. The structure of the human ribosome is from Anger et al. 2013 (PDB entry 4V6X). The programme Chimera (cgl.ucsf.edu/chimera) was used to display the structure. (A) small ribosomal subunit proteins enriched with the 18S rRNA. (B) large ribosomal subunit proteins enriched with the 28S rRNA. There is no bias in location of the enriched ribosomal proteins.

Two large ribosomal proteins were enriched in the 18S rRNA interactome (Rpl8, Rpl22). Therefore I had a closer look at the position of these proteins in the human ribosomal structure (Figure 2.18). Rpl8 was detected in previous ribosomal biogenesis screens (Wild et al., 2010; Badertscher et al., 2015) and resides close to the inter-subunit space. It could therefore have contact points with both the 18S and the 28S rRNA. Rpl22 on the other hand might be too distant from the 18S rRNA in the mature ribosome. An explanation for the detection of Rpl22 in both the 18S rRNA and 28S rRNA interactome could be that the protein associates during ribosomal biogenesis steps with the

unprocessed pre-rRNA, although a role of Rpl22 during maturation of the rRNA has not been described so far.

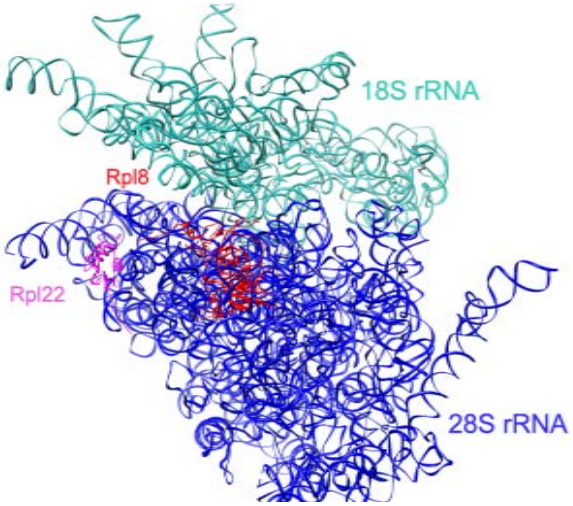

Figure 2.18; Position of enriched ribosomal proteins coming from the respective other subunit of the cytoplasmic ribosome. The structure of the human ribosome is from Anger et al. 2013 (PDB entry 4V6X). The programme Chimera (cgl.ucsf.edu/chimera) was used to display the structure.

2.8.5 Specific differences between the two rRNA interactomes

The 18S rRNA and 28S rRNAs show some distinct differences from each other. Proteins that were specifically enriched with the 18S rRNA but not the 28S rRNA interactome were eiF3J, SERBP1, Rps15, LRPPRC, Rpl8, RPS24 and RPSA. The small subunit proteins RPS15, RPS24 and RPSA were expected to associate specifically with 18S rRNA. The occurrence of Rpl8 in the small subunit rRNA dataset might be there due to its position at the intersubunit space of the ribosome (discussed above).

SERBP1 (also known as plasminogen activator inhibitor 1 RNA-binding protein) was previously identified in the 80S ribosomal structure. It mostly associates with the 40S subunit but reaches to the 60S subunit (Anger et al., 2013), which fits with the enrichment of SERBP1 in the 18S rRNA interactome. SERBP1 is a RBP that simultaneously shifts to cytoplasmic stress granules and nucleoli upon stress. SERBP1 is homologous to the translation repressor Stm1 (Balagopal & Parker, 2011) and might therefore also possess a role in translational repression when localized to stress granules. Its observed localization to nucleoli suggests a possible involvement in ribosomal biogenesis.

The initiation factor eiF3J is a component of the eukaryotic translation initiation factor 3 (eIF3) complex that associates with the 40S ribosomal subunit (Fraser et al., 2003). In *Saccharomyces cerevisiae*, eIF3J has also been shown to be required for the processing of 20S pre-rRNA in 18S rRNA and also binds to 18S

rRNA (Valásek et al., 2001). For LRPPRC (Leucine Rich Pentatricopeptide Repeat Containing) there is no documentation about a biological role associated with the function or biogenesis of the cytoplasmic ribosome. LRPPRC is a RBP that associates with both nuclear and mitochondrial mRNAs (Mili & Piñol-Roma, 2003). For nuclear RNAs, LRPPRC is described to be involved in mRNA export from the nucleus. In mitochondria, LRPPRC controls RNA processing and translational control (Sasarman et al., 2010; Chujo et al., 2012). A possible role of LRPPRC in cytoplasmic ribosomal biogenesis or translational activity of the ribosome has not been described yet and it might be therefore intriguing to see if LRPPRC could be identified as a novel factor for ribosomal RNA processing or ribosomal activity. Potentially it could be involved in rRNP export from the nucleus as this function was already described for the export of mRNA.

Proteins that were specifically enriched with the 28S rRNA but not the 18S rRNA interactome were Rpl29, Rpl31, Rpl5, Rpl4, PA2G4, HNRNPC, HNRNPD, BTF3, HNRNPA2B1, HNRNPM, PLIN4, DSG1. The large subunit proteins were expected to associate specifically with 28S rRNA. HNRNPC,-D and –F could also be identified in the biogenesis screens that were used for cross-referencing (discussed later).

PA2G4 (proliferation-associated 2G4; other name = EBP1) is a double stranded RNA binding protein (Squatrito et al., 2006) that is present in pre-ribosomal ribonucleoprotein complexes and is implicated in ribosome maturation and the regulation of rRNA processing (Squatrito et al., 2004). Its orthologue in yeast, Arx1, is a nuclear export receptor for the 60S ribosomal subunit (Hung et al., 2008). Structural data in yeast show that Arx1 interacts with the expansion segment ES27L (Greber et al., 2012). It positions close to the peptide exit tunnel where it inhibits the association of translation factors before the maturation is complete. In summary, these observations are in line with the significant enrichment of PA2G4 with the 28S rRNA.

BTF3 (also known as Nascent Polypeptide-Associated Complex Subunit Beta) is part of the Nascent polypeptide associated complex and was described to be associated with the ribosome and to prevent inappropriate ribosome binding to ER membranes (Beatrix et al., 2000). As the ribosomal exit tunnel resides in the

large ribosomal subunit, it is an expected finding that BTF3 was identified in the 28S rRNA interactome but not in the 18S rRNA interactome.

The occurrence of Perilipin-4 (Plin4) and Desmoglein-1 (Dsg1) in the 28S rRNA interactome is interesting because both proteins have not been described before to be associated with RNA, to bear RNA binding motifs, or to have a biological function related to the ribosome. Desmoglein is associated with cell junctions (Ohsugi et al., 1997), whereas Plin4 might play a role in the biogenesis of lipid droplets (Heid et al., 2014). It would be interesting, in which biological context these proteins interact with the 28S rRNA and if their association with the ribosomal RNA involves translational control, ribosomal maturation or other functions.

2.8.6 Enrichment of HNRNP proteins and their possible role in rRNA biology

HNRNPs are involved in mRNA maturation. Not so much is known about their possible role in rRNA maturation. In the case of HNRNPK it was shown that it associates with rRNA and is then required for rRNA stability in non-stressed cells (Wen et al., 2012).

The occurrence of some HNRNPs in the ribosomal biogenesis screens are a strong indicator that the here identified may play an important role in the biogenesis of the ribosomal subunits; although this role is so far unexplored. The northern blot readout used in Tafforeau et al., 2013 showed that HNRNPC knockdown affects the early processing of 18S and 28S rRNAs. HNRNPC might therefore interact with the pre-rRNA before cleavage of the internal transcribed spacer 1 (ITS1) that separates 18S and 28S rRNA. HNRNPF is, as HNRNPC, involved in the early pre-rRNA processing. The northern blot results of HNRNPD in the siRNA screen imply an involvement in the large subunit maturation. This complies with the high confidence enrichment of HNRNPD in the 28S rRNA interactome but not in the 18S rRNA interactome.

Interestingly, a recent study revealed that RNA binding of HNRNPC binding is enhanced when mRNAs carry m6A modifications (Liu et al., 2015). As human 18S and 28S rRNAs carry m6A modifications (Motorin & Helm, 2011), it is possible that HNRNPC or other HNRNPs interact with rRNA in an m6A-dependent manner. In addition, there is the possibility that other rRNA

modifications regulate HNRNP binding to rRNA. iCLIP or ECLIP data will provide insights into the exact interaction sites of HNRNP proteins with 18S and 28S rRNAs.

In summary, with the rRNA interactomes I could provide evidence that specific RNP capture in addition to its use *in vitro* can also be applied to purify RNA-binding proteins bound to specific RNA species in cells. The two rRNA interactomes displayed the clean biochemical performance of the protocol. Most of the differences in the rRNA interactomes could be cross-correlated to previous observations in literature. Thus, specific RNP capture again successfully recapitulated data generated in the past by orthogonal approaches. In addition specific RNP capture provided data about potential novel proteins involved in ribosome biology or function.

2.8.7 rRNA capture after cell fractionation

In order to distinguish cytoplasmic interactors of rRNA from interactors of rRNA that are involved in the early processing and assembly steps in the nucleolus and nucleus, I performed rRNA capture with subcellular fractions of HeLa cells. The subcellular fractionation was biochemically clean as shown by the enrichment of cytoplasmic proteins and depletion of the nuclear proteins in the cytoplasm and vice versa. By using the subcellular fractions I could then specifically pull-down 18S rRNA with the LNA/DNA mixmer probes. However, I could observe RNA degradation in the nuclear fraction. This degradation could be due to sonication of the nuclear fraction to disrupt the chromatin. For specific RNP capture of the nuclear fraction, I will have to apply a different strategy to maintain RNA integrity. Prospectively, specific RNP capture of fractionated sample in principle works and can be used in future experiments. It will allow distinguishing better biogenesis factors from factors involved in ribosomal function in protein synthesis or ribosomal degradation.

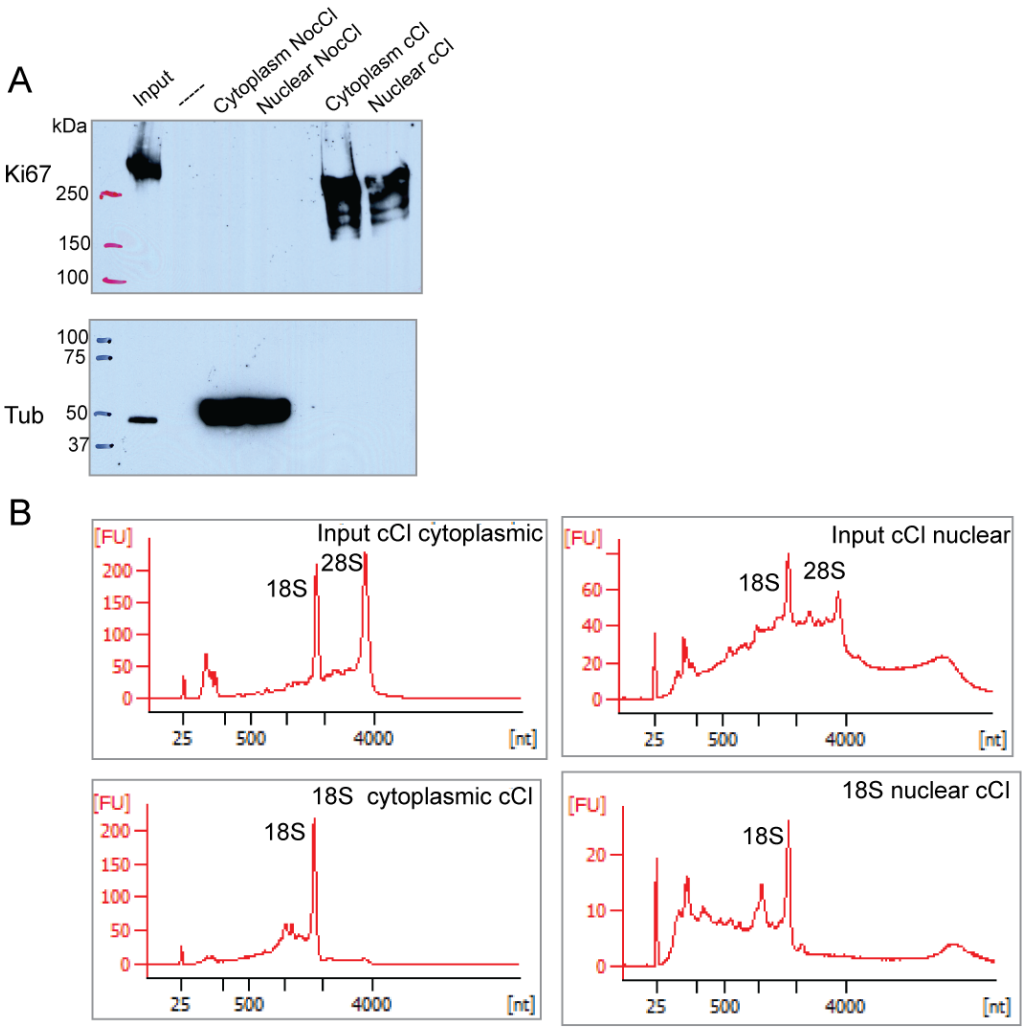

Figure 2.19; Specific RNA capture of 18S rRNA after cellular fractionation of HeLa cells subjected to UV crosslinking. (A) After fractionation, the cytoplasmic and nuclear fractions were analyzed via Western Blot to monitor clean separation of the subcellular fractions. Ki67 was used as a nuclear marker and Tubulin (Tub) was used as the respective cytoplasmic marker. Clean separation of both sub-cellular fractions could be performed. (B) 18S rRNA capture was performed with the respective sub-cellular fractions. The enriched RNA was analyzed using a 6000 Pico bioanalyzer. The 2 peaks from the Input samples come from 18S (runs at 2000 nt) and 28S (runs at 4000 nt). The y-axes display the fluorescence intensity coming from a dye labeling the RNA quantitatively. 18S rRNA could be enriched in both sub-cellular fractions. Degradation of RNA could be monitored in the nuclear fraction.

3 Conclusions

3.1 Specific and stringent isolation of RNPs by Specific RNP capture

I have demonstrated that the protocol of specific RNP capture specifically enriches the targeted RNA and its bound proteome in eluates. I was also able to extend the protocol from *in vitro* to in-cell applications.

Specific RNP capture performed in *Drosophila* embryo extracts with reporter RNAs not only “re-discovered” the Sxl-msl2 U-rich motif interaction, but also uncovered that Ssx exhibits similar binding preferences as its paralog. With these experiments I could also demonstrate that specific RNP capture can be used in comparative binding studies on reporter RNAs as Ssx showed differential association to the wt over the mut reporter RNA. When I used specific RNP capture in cells to isolate human 18S and 28S rRNA, I could specifically purify the targeted rRNA and effectively remove the other (abundant) rRNA. The specificity of the pulldown allowed me to generate two distinct datasets, the 18S rRNA interactome and the 28S rRNA interactome. Although the amount of detected proteins is limited (reasons and solutions discussed below) the good biochemical performance of the method could be confirmed by the nature of the identified proteins. Many of the enriched proteins were identified as components of the cytoplasmic ribosome or playing potential roles in ribosomal biogenesis.

3.2 Limitations of specific RNP capture and how to circumvent them

The stringency included at all levels of the specific RNP capture protocol imposes limitations, such as relatively low level of purified RNA and co-purified proteins. I observed that a substantial amount of RNA is required in order to generate a comprehensive interactome of the RNA of interest, reaching the order of micrograms. Micrograms of RNAs for an *in vitro* experiment are in principle not a problem because exogenous RNA can be used in high quantities by scaling-up proportionally the amounts of cell extracts as needed. However, this represents a challenge when applied to endogenous RNA. A typical mammalian cell contains 10–50 pg total RNA, whereas the majority of RNA

molecules are tRNAs and rRNAs (Boon et al. 2011). mRNA accounts for only 1–5% of the total cellular RNA although the actual amount depends on the cell type and physiological state. During my PhD thesis I applied specific RNP capture to 18S and 28S rRNA, targeting therefore highly abundant transcripts in the cell. I was able to isolate 4-9 µg of ribosomal RNA from around $2*10^8$ cells. Nonetheless, I observed that the recovery of the ribosomal RNA was in the low percentage range (1-5%), which is contrary to the recovery *in vitro* (~20%). Possible explanations could be the nature of the RNA with regard to RNA modifications, RNA degradation during hybridization, strong secondary structure formation or RNA-RNA crosslink formation. Even when using 4-9 µg of crosslinked rRNA, I still had problems in terms of MS detection depth. Enriched proteins were nevertheless different between the 18S and 28S rRNA interactome, implying that even if the datasets is not complete, it was highly specific.

One possible way to overcome these limitations will involve the use of alternative MS measurement strategies following the newest developments in the field of MS. The targeted proteomics technique, based on the processing technique called selected reaction monitoring (SRM), allows highly sensitive protein detection, expanding the limits of the conventional discovery proteomics (Picotti & Aebersold, 2012). But targeted proteomics is also biased towards a pre-selection of proteins of interest. Only peptides will be processed by the MS apparatus that have been pre-selected before. One pre-selection criteria in the case of specific RNP capture could be proteins enriched via RNA interactome capture. Targeted proteomics is a recent development and it will be interesting to explore in the context of Specific RNP capture.

Another possible way to circumvent detection limits are related to the crosslinking strategy. UV crosslinking allowed specific enrichments of proteins that directly contact the RNA of interest. Nevertheless, it is known that the usual efficiency of UV crosslinking at 254 nm and energy of 150 mJ/cm^2 is between 1-5% (Darnell, 2010); although this limit could be pushed to 20% for certain proteins such us CUG-BP using optimized protocols (Castello et al., 2012). One way to address this limitation would be to increase the crosslinking energy. Nevertheless, I observed that the RNA integrity is compromised reducing recovery rates. A multitude of targeting probes covering the whole transcript

length (a strategy referred to as “tiling”) would circumvent the problematics of RNA fragmentation. The tiling strategy has been already used to target lncRNA (West et al., 2014; Chu et al., 2015; McHugh et al., 2015). Nonetheless, as LNA/DNA mixmer probes are more expensive in synthesis than the DNA oligonucleotides used in the “tiling” strategies, specific RNP capture would become an inaccessible protocol.

Another way to address the crosslinking limitations, would be to use 4-thiouridine (4SU). Since 4SU-mediated crosslinking is achieved at 365 nm and conventional bases do not absorb at this wavelength, RNA fragmentation would be reduced upon high energy irradiations. Nonetheless, in the case of the rRNA interactome the use of 4SU can cause rRNA biogenesis defects (Burger et al., 2013) and thereby would result in an interactome dataset which does not reflect the normal physiological state of the cell.

Alternatively, a chemical crosslinker like formaldehyde can be used, although this strategy would promote protein-protein crosslinks that will mediate the isolation of proteins that do not directly contacting the RNA. However, using limited formaldehyde crosslinking would still enrich for proteins that are forming part of RNP complexes. Direct vs indirect RNA binders can be separated a posteriori by cross-referencing the resulting datasets with the repertoire of known RBPs (derived from the RNA interactome studies). In summary, if the protein-RNA crosslinking is enhanced and the downstream analysis and data interpretation are performed carefully, the advantages of chemical crosslinking can overcome its disadvantages.

3.3 Comparison of specific RNP capture to other recent protocols based on antisense oligonucleotide probes

For the identification and characterization of proteins bound to lncRNAs, several methods based on the usage of antisense oligonucleotides have been established (Table 3.1).

Table 3.1; Comparison of recent antisense oligonucleotide approaches to identify RBPs bound to a specific transcript. CHART-MS (West et al., 2014); CHIRP-MS (Chu et al., 2015); RAP-MS (McHugh et al., 2015)

Method	**No of cells**	**Immobilization of oligos**	**crosslinking**	**Length of oligos**	**Hybridization**	**Release from beads**	**MS quantification method**
CHART-MS	10^8	Streptavidin - Biotin	3% Formaldehyde	25 nts	20°C over night	RNaseH	spectral counts
CHIRP-MS	10^8 – $5*10^8$	Streptavidin - Biotin	3% Formaldehyde	20 nts	37°C over night	Biotin elution buffer	peptide counts
RAP-MS	$2*10^8$- $8*10^8$	Streptavidin - Biotin	UV_{254} 800mJ/cm^2	90 nts	67°C 2h	Benzonase	SILAC
Specific RNP capture	$2*10^8$	Covalent coupling to carboxy beads	UV_{254} 150 mJ/cm^2	20 nts	41°C 2h	Heat 3' 90°C	Dimethyllabeling

"Tiling" versus single antisense nucleotide: The recent protocols of CHART-MS (Capture Hybridization Analysis of RNA Targets) (West et al., 2014), CHIRP-MS (Chromatin Isolation by RNA Purification) (Chu et al., 2015) and RAP-MS (RNA Antisense Purification) (McHugh et al., 2015) are all based on the so called "tiling" approach, in which a number of overlapping biotinylated oligonucleotides that cover the whole transcript length are used. The advantage of the tiling approaches is that the RNA integrity is not as crucial as for Specific RNP capture, in which just one oligonucleotide is applied to target the RNA of interest. Nevertheless, the disadvantage of using a broad number of oligonucleotides is that it is very challenging to estimate the contribution to noise and signal of each individual probe. Moreover, it would be difficult to target distinct transcript isoforms or particular regions of the RNA of interest, since the probes cover the whole transcript. Specific RNP capture can for example be utilized to compare the RBPs bound to different splice variants of one transcript. Another application for which specific RNP capture can outperform its competitor methods is enrichment of a specific region within a transcript by combining specific RNP capture with site selective RNaseH digestion (see Outlook; Figure 4.1).

Nucleotide length, composition and hybridization temperatures: An interesting difference between the mentioned methods is the nature and length of the oligonucleotides. All "tiling" approaches employ DNA oligonucleotides

because they are easy to obtain and cheap. CHART-MS and CHIRP-MS use relatively short DNA oligonucleotides and accordingly low hybridization temperatures (20-37°C). This might lead to lower specificity of the hybridization. RAP-MS uses rather long oligonucleotides and accordingly higher hybridization temperatures (67°C). As specific RNP capture uses LNA/DNA mixmer, the oligonucleotides could be as short as 20 oligonucleotides with increased hybridization temperatures for specificity. LNA/DNA mixmers also hybridize better to RNA than DNA oligonucleotides. DNA oligonucleotides could also hybridize to a higher extent to genomic DNA, leading to increased contaminations with DNA and proteins bound to it.

Immobilization of the oligonucleotides to beads: Another difference to specific RNP capture is the way in which the probes are immobilized to beads. In the specific RNP capture protocol the probes are covalently coupled to magnetic beads before the hybridization step. For the “tiling” approaches, the oligonucleotides carry a biotin tag and are immobilized on the beads after the hybridization to the target RNA. An advantage of this approach might be the higher flexibility of the soluble oligonucleotides during hybridization. Nonetheless, the biotin-streptavidin coupling has the disadvantage that several proteins contain biotin molecules and can bind to the streptavidin beads in a RNA-independent manner. As the bead surface of the covalent coupling strategy is in principle inert, the bead surface is less prone for unspecific interactions. Another disadvantage is that the used biotinylated oligonucleotides cannot be applied in several rounds of capture as in the case for the covalently coupled oligonucleotides.

Crosslinking: The methods CHART-MS (West et al., 2014) and CHIRP-MS (Chu et al., 2015) make use of extensive formaldehyde crosslinking, which leads to the detection of many indirectly interacting proteins. The high percentage of formaldehyde (3%) used in these approaches might even lead to a fixation of the whole cell. RAP-MS on the other hand is based on UV crosslinking at high energies (800 mJ/cm^2). This extensive crosslinking leads to perturbations of RNA integrity. In contrast to specific RNP capture it is still possible to capture RNA fragments with the “tiling” approaches. It is important to know when using such high crosslinking energies that fragmentation of the RNA

can affect the downstream RNA quality control steps performed with RT-qPCR or RNA bioanalyzer.

Mass spectrometry quantification methods: The afore mentioned protocols use different mass spectrometry methods to identify the proteome bound to the RNA of interest. CHART-MS and CHIRP-MS use label-free quantification. In label-free quantification, no label is introduced to the samples; the comparison is then done by comparing different MS runs, with the difficulty of aligning peptide peaks in non-identical profiles. A drawback is that technical variations of the Mass Spectrometer runs or properties of the compared peptides can make such analyses difficult and can generate biases due to the analytic software used.

RAP-MS uses SILAC ratios formed between a pulldown of an abundant RNA to the RNA of interest. The advantage of the SILAC label is that the label is introduced at an early point during the protocol and the quantification is therefore less biased by technical variations. The comparison to an abundant RNA can be problematic when the interacting RBPs are also strong interactors of the abundant RNA. Specific RNP capture uses dimethyl labeling which allows comparing up to two different controls with the sample of interest. The strength of dimethyl labeling is used for specific RNP capture in HeLa cells by including the scrambled oligonucleotide and non-crosslinking control. As in SILAC, the three samples are processed together in the same MS run, which makes quantification more accurate than the label-free approach. As a drawback, dimethyl labeling requires multiple processing steps that are known to account for sample losses, thus promoting accuracy over sensitivity. In addition, specific RNP capture gives statistics based on three replicates. No statistical analysis has been performed in any of the “tiling” studies. Including statistical tests or at least comparing replicates directly would increase the confidence of the identified candidates in these studies.

4 Outlook

4.1 The biological role of Ssx

In the *in vitro* experiment performed with *Drosophila* embryo extract it was expected to re-identify Sxl to interact with the established Sxl binding sites introduced in the reporter RNA. Nonetheless, the surprising finding was that instead of Sxl, its paralog Ssx together with the HNRNP proteins Hrb87F and Hrb98DE were identified to interact with the reporter RNA bearing Sxl-binding sites. Sxl was previously shown to interact with Hrb87F, most likely through a common glycine-rich region, and increases the affinity of Hrb87F for U-rich sequences (Wang et al., 1997). As Ssx shares the N-terminal glycine-rich feature with Sxl, the results suggest that Ssx can also recruit Hrb87F, implying functional redundancy between the two paralogs. Co-Immunoprecipitation (Co-IP) experiments could give first insight into a possible interaction of these proteins. In the co-IP experiments one could also test protein fragments to narrow down the interaction site and include RNase digestion to determine if the interaction is RNA dependent or not.
Specific RNP capture could provide evidence that Ssx and Sxl proteins bind similar target sites in the RNA. It might be therefore interesting to compare the RNA targets and binding sites of both proteins by performing CLIP experiments. It would be interesting which differences in the binding site determine that Sxl, in contrast to Ssx, has an essential function in sex determination and dosage compensation.

Based on the similarity of both proteins it would be interesting if Ssx can, as Sxl, act as a splicing and translational regulator. Reporter RNAs bearing Sxl splicing sites could be tested with recombinant Ssx. An ECLIP or iCLIP experiment could show if the Ssx target RNAs share the same splice sites as the Sxl target RNAs.

4.2 The role of HNRNP proteins during ribosomal biogenesis

The occurrence of several HNRNP proteins in the 18S rRNA and 28S rRNA interactome is intriguing. Especially in the context that several HNRNP proteins have also been identified in ribosomal biogenesis screens (Tafforeau et al., 2013; Badertscher et al., 2015). As HNRNP proteins are involved in the

maturation of mRNA, a similar role could be true for the maturation of rRNA. As in Tafforeau et al., 2013, knockdown of HNRNP proteins could be coupled to Northern Blot experiments in order to see if HNRNP proteins influence the pre-rRNA maturation.

It would be interesting to see if HNRNP proteins interact with rRNA and pre-rRNA by analyzing the recent ECLIP and iCLIP datasets under this aspect. Previous ECLIP and iCLIP analysis pipelines exclude rRNA as a contaminant. Additionally, rRNA is excluded with multiple genome mappers because rRNA is transcribed in highly repetitive genomic clusters. It is therefore necessary to adjust the analytical pipeline by mapping the sequences to the transcriptome instead of mapping it to the genome. The enriched crosslink sites could then tell if HNRNP might preferably bind pre-rRNA. A question that could be addressed is if the binding of HNRNP proteins target sites in the rRNA that are modified. It would be intriguing to see if the binding during ribosomal biogenesis is influenced by modifications on the rRNA such as m6A. m6A dependent RNA structural switches were observed previously to modulate binding of HNRNPC to a target mRNA (Liu et al., 2015). Progress in the epitranscriptomic field might give more insight into dynamics and variety of rRNA modifications.

4.3 Future applications of specific RNP capture

When the limitations of specific RNP capture have been overcome by increasing the crosslinking efficiency with the use of limited chemical crosslinking, the specific RNP capture protocol can be expanded from targeting rRNA also to other, less abundant RNA species. Progress with regard to targeted mass spectrometry might additionally increase the sensitivity of the methods. The stringency of the protocol in combination with quantitative mass spectrometry will help to generate lists of proteins that interact with high probability with the RNA of interest. The high specificity of the protocol will facilitate the downstream analyses.

With regard to the rRNA interactome a comparison could be made between different cell types or development. It is now known that specialized ribosomes with specialized translational machineries are important regulators of gene expression (Xue & Barna, 2012). Other experiments based on the rRNA

interactome could be the region specific enrichment of ribosomal RNA by combining specific RNP capture with RNase H digest (Figure 4.1). This type of protocol is dependent on the respective sensitivity improvements of the method. The approach would then allow the specific enrichment of ribosomal expansion segments with the bound proteome. Identification of the expansion segment RBPs would give more insight about the biological role of these expansion segments, for which the function especially in human cells is still poorly understood.

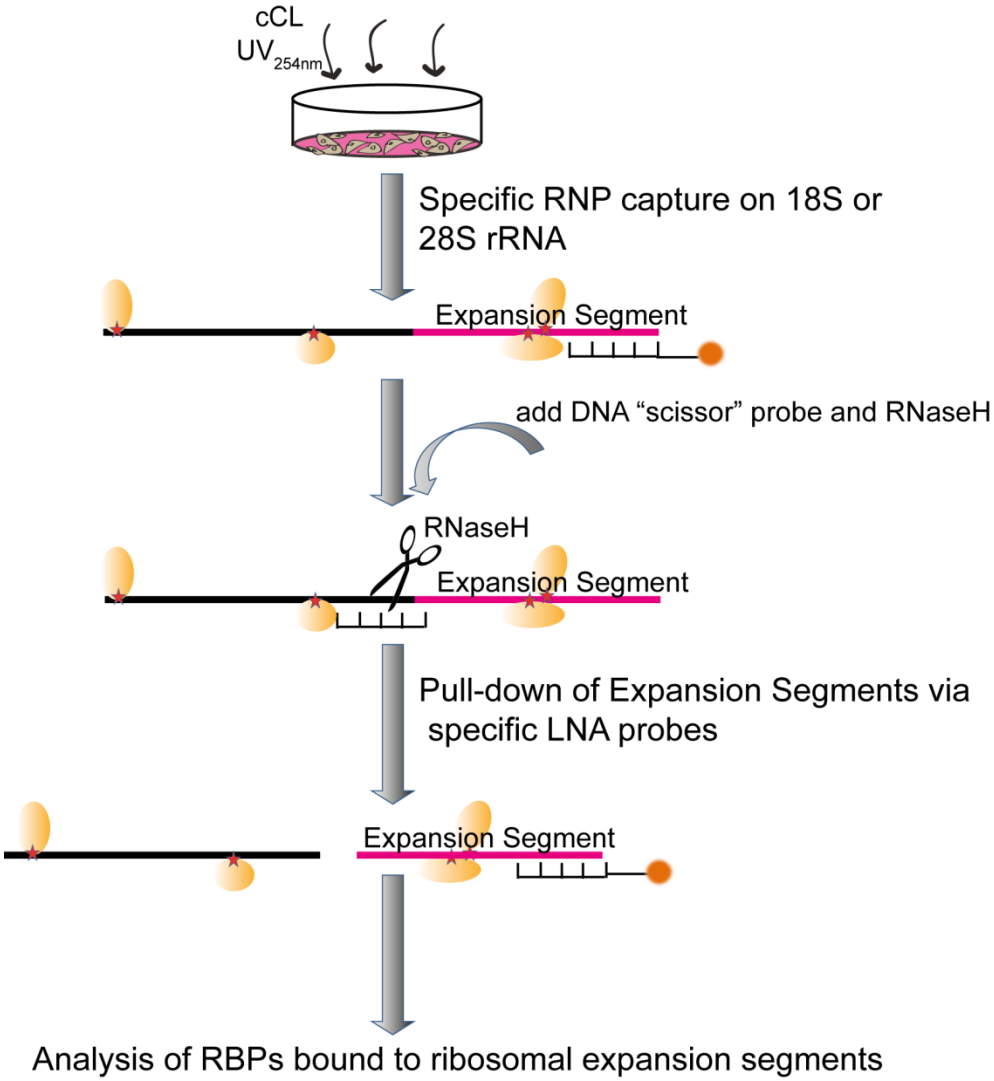

Figure 4.1; Specific RNP capture of ribosomal expansion segments. After UV crosslinking of cells, the specific RNP capture protocol is performed to capture 18S or 28S rRNA. After the enrichment of the respective rRNA, DNA probes (scissor probes) complementary to upstream and downstream regions of the ribosomal expansion segment are added and a digest with RNAseH is performed. The RNAseH digest enables a dissection of the target ribosomal expansion segment. Subsequently, a second specific RNP capture step targeting the expansion segment with complementary LNA/DNA mixmer probe is performed. This step allows an enrichment of the respective expansion segment together with the bound proteins. The proteins are finally analyzed via mass spectrometry.

Another application for “specific RNP capture” could be in the context of viral infection of eukaryotic cells. Eukaryotic RNA viruses utilize host factors for their virus life cycle. For many RNA viruses there is a lack of knowledge of host factors interacting with the viral RNA. A deeper understanding of the role of specific host factors in the viral life cycle will enable the development of new antiviral treatment possibilities (de Chassey et al., 2014). Some RNA viruses can produce a high copy number of viral RNA within the host cell and thereby the sensitivity limits of “specific RNP capture” could be overcome. Specific RNP capture would therefore be a suitable tool for the identification of cellular proteins interacting with high copy number viral RNA during the viral life cycle.

5 Materials and Methods

5.1 Materials

5.1.1 Chemicals/Disposables/Kits

Chemicals / Disposables / Kits	Company
100bp DNA ladder	New England Biolabs
1kb DNA ladder	New England Biolabs
2-(*N*-morpholino)ethanesulfonic acid (MES)	Merck
30% Acrylamide/Bis Solution 37,5:1	Biorad
4-15%™ TGX™ Precast gel	Biorad
5 mL, 10 mL, 25 mL serological pipets	Falcon
500 cm2 cell culture dishes	Corning
Acetonitril	Sigma-Aldrich
Agarose	Sigma-Aldrich
Amino acid mixture, complete	Promega
Ammoniumpersulfate (APS)	Sigma-Aldrich
Ampicilline	Sigma-Aldrich
Bactotryptone	Difco
Benzonase	Novagen
boric acid	Sigma-Aldrich
BSA in solution (20 mg/mL)	NEB
Calf Intestine Phosphatase (CIP)	NEB
Carboxylated magnetic beads (M-PVA C11)	Perkin Elmer
Cell culture dishes	NuncThermo
Cell culture flasks	Falcon
Complete EDTA-free Protease Inhibitor Cocktail	Roche
Creatinekinase	Roche
Creatinephosphate	Sigma-Aldrich
DC protein assay	Biorad
Deoxynucleotides (dNTPs)	Peqlab
Dithiothreitol	Biomol

DMSO	Sigma-Aldrich
DNA/RNA LoBind Tubes	Eppendorf
Dnase 1	Ambion
Dual Luciferase Assay System	Promega
Dual-Luciferase Reporter 1000 Assay System	Promega
Dulbecco's Modified Eagle's Medium	Sigma-Aldrich
EDTA	Sigma-Aldrich
Epoxy magnetic beads M-PVA E01	Perkin Elmer
Ethanolamine	Sigma-Aldrich
Ethidium bromide	MP Biomedicals
Ethylenecarbonate	Sigma-Aldrich
Fetal Bovine Serum (Gold)	GE Healthcare
Fluorescent modified DNA oligonucleotides	MWG
Formic Acid	Sigma-Aldrich
Glycerol	Merck
Glycine	Sigma-Aldrich
HeLa cytoplasmic extracts	Cilbiotech
HEPES	Biomol
Immobilon Western HRP substrate	Millipore
Iodoacetamide (IAA)	Merck
KCl	Sigma-Aldrich
KH2PO4	Sigma-Aldrich
L-Amino Acids Kit, 21 Amino Acids	Sigma-Aldrich
L-glutamine	Gibco
Lithium Chloride	Roth
Lithium Dodecylsulfate (LiDS)	Roth
Locked Nucleic Acids (LNA)	Exiqon
MEGAscript® Transcription Kit	Thermo
Milk Powder	Frema Reform
N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC) E7750	Sigma-Aldrich
N,N,N';N'-tetramethylethylenediamine (TEMED)	Fluka
Na2HPO4	Sigma-Aldrich
NaCl	Merck

Nitrocellulose Membrane	Whatman
NP-40 Igepal CA630	Sigma-Aldrich
NuPAGE® LDS Sample Buffer	Thermo
OASIS 96 well plate	Waters
oligodT beads	NEB
Penicillin-Streptomycin Mix (P4333) with 10,000 units penicillin and 10 mg streptomycin/mL	Sigma-Aldrich
Phusion High Fidelity PCR kit	NEB
Plasmid Maxi Prep Kit	Qiagen
Protein ladder; Precision Plus Protein Dual Color	Biorad
Proteinase K	Roche
QIAquick Gel Extraction Kit	Qiagen
Random hexamer primers	Life Technologies
RNA 6000 Pico Kit	Agilent
RNA Micro Prep kit	Zymo
RNA Mini Prep kit	Qiagen, Zymo
RNase A	Sigma-Aldrich
Rnase free H2O	Ambion
RNAse Inhibitor	protein purification prepared in house
Rnase T1	Sigma-Aldrich
SDS	Sigma-Aldrich
Silvernitrate	Merck
sodium cyanoborohydride (NaBH3CN, NaBD3CN)	Sigma-Aldrich
Sodiumhydroxide	Roth
Spermidine	Sigma-Aldrich
SteriCup Filters	Millipore
SuperScript II	Life Technologies
T4 DNA ligase	NEB
Tri Reagent	Sigma-Aldrich
Trichloracetic Acid (TCA)	Sigma-Aldrich
Tris base (Tris(hydroxymethyl)aminomethane)	Sigma-Aldrich

Trypsin	Gibco
Turbo DNAse	Ambion
Tween-20	Sigma-Aldrich
Western X ray films	Advansta
yeast extract	Difco
yeast trna	Thermo

5.1.2 Buffers and Solutions

All solutions were prepared with double deionized water (ddH2O). Buffers for RNA methods and Cell Culture were autoclaved and filtered. Buffers for Bacteria were autoclaved.

DNA loading buffer (6x)	0.25 % (w/v) Bromphenol blue 0.25 % (w/v) Xylene cyanol FF 40 % (w/v) Sucrose
TBE Buffer (10x) 1L	108 g Trizma base 55 g boric acid 40 mL 0.5 M EDTA pH 8
Semidry Transfer Buffer (1L)	5.82 g Trizma Base 2.93 g Glycine 20 % MeOH
Laemmli Buffer (10x) (1L)	30 g Trizma Base 144.2 g Glycine 10 g SDS
L-Broth (LB) (1L)	10 g Bactotryptone 5 g yeast extract 5 g NaCl
L-Agar (1L)	10 g Bactotryptone 5 g yeast extract 10 g NaCl 15 g Agar

PBS-Dulbecco (1L)	0.2 KCl 0.2 g KH2PO4 1.15 g Na2HPO4 8 g NaCl
TE Buffer pH 7.5	10mM Tris pH 7.5 1mM EDTA
Light labeling reagent for Peptide labeling on OASIS plate	9 mL of 50 mM Sodium Phosphate buffer pH 7.5 500 µL of 600 mM NaBH3CN 500 µL of 4% formaldehyde
Medium labeling reagent for Peptide labeling on OASIS plate	9 mL of 50 mM Sodium Phosphate buffer pH 7.5 500 µL of 600 mM NaBH3CN 500 µL of 4% D-formaldehyde
Heavy labeling reagent for Peptide labeling on OASIS plate	9 mL of 50 mM Sodium Phosphate buffer pH 7.5 500 µL of 600 mM NaBH3CN 500 µL of 4% D-13C-formaldehyde
Nuclear Buffer (for Nuclear Cytoplasmic Fractionation) add 0,5% NP40 for Fractionation; add freshly RNAsIn and protease Inhibitor	10 mM Tris HCl pH 7.5 20 mM KCl 1.5 mM MgCl2 1 mM DTT
200x Dnase Solution (for Dnase digest in Nuclear fraction)	500 mM $MgCl_2$ 100 mM $CaCl_2$

Buffers for RNA interactome captures:

Lysis Buffer	20 mM Tris-HCl (pH 7.5) 500 mM LiCl, 0.5% LiDS (w/v, stock 10%) 1 mM EDTA 5 mM DTT
2x lysis Buffer	40 mM Tris-HCl (pH 7.5) 1 M LiCl 1% LiDS (w/v, stock 10%) 2 mM EDTA
(add DTT and LiDS separately to sample);	10 mM DTT
LiDS precipitates at higher concentrations in cooled buffer	
5x lysis Buffer	100 mM Tris-HCl (pH 7.5) 2.5 M LiCl 2.5% LiDS (w/v, stock 10%) 5 mM EDTA
(add DTT and LiDS separately to sample);	25 mM DTT
LiDS precipitates at higher concentrations in cooled buffer	
Buffer 1	20 mM Tris-HCl (pH 7.5) 500 mM LiCl 0.1% LiDS (w/v, stock 10%) 1 mM EDTA 5 mM DTT

Buffer 2
20 mM Tris-HCl (pH 7.5)
500 mM LiCl
1 mM EDTA
5 mM DTT

Buffer 3
20 mM Tris-HCl (pH 7.5)
200 mM LiCl
1 mM EDTA
5 mM DTT

5x proteinase K buffer
50 mM Tris-HCl (pH 7.5)
750 mM NaCl
1% (w/v) SDS
50 mM EDTA
2.5 mM DTT
25 mM CaCl2

5.1.3 Antibodies

Antigen	Clonality	Species	Company	Dilution used
HRP coupled secondary antibody	rabbit, mouse	goat	Santa Cruz (sc-2005; sc 2004)	1:10,000
Ki67	polyclonal	rabbit	Abcam (ab15580)	1:7000
Drosophila Sxl	polyclonal	rabbit	prepared in house	1:2000
Tubulin	monoclonal	mouse	Sigma (T5168)	1:10,000
Drosophila Ssx	polyclonal	rabbit	kind gift from Jan Medenbach	1:2000
Actin	monoclonal	mouse	Sigma (A5441)	1:10,000
Drosophila PABP	polyclonal	rabbit	prepared in house	1:2000
HRP coupled secondary antibody	rabbit, mouse	sheep, donkey	GE Healthcare (NA931-1ML;NA934-1ML)	1:10,000

5.1.4 Instruments

Instrument	Company
Acrylamide Gel Electrophoresis chamber	EMBL
Agarose Gel Electrophoresis chamber	EMBL
Balance	Scaltec
Bioanlayzer; lab on chip	Agilent
Cell Cryo Freezing Container	Inalgene
Cell Cryo Freezing Container	Biocision
Cell Culture Hood	Heraeus
Cell Incubator	Thermo Scientific Fisher
Cell Incubator Hera cell 240	Heraeus
Centrifuge 5810R	Eppendorf
Cooled tabletop microcentrifuge (5415R)	Eppendorf
Dual Intensity Transilluminator Table	Herolab
Electrophoresis Power Supply	EMBL
Electrophoresis Power Supply	Biometra
Film Developer RP X-OMAT	Kodak
Luminometer Centro XS LB960	Berthold Technologies
Macrofuge	Heraeus Instruments
Magnetic stirrer	Heidolph
Megafuge 1.0R	Heraeus
Micropipette	Gilson
Microscope	Olympus
Microwave	Bosch
Multipette Plus	Eppendorf
pH meter	Inolab
Pipetboy acu	Integra Biosciences
Rocker	EMBL
Rotator	Starlab
Semi Dry Protein Blotter	Biometra
Shaker	Infors AG
Spectrophotometer Ultraspec 2100 pro	Amersham Biosciences

Spectrophotometer, Nanodrop1000	Nanodrop
Step One qPCR machine	LifeTechnologies
T100 Thermal Cycler	Biorad
Tabletop Microcentrifuge (5424)	Eppendorf
Tecan Safire2 M1000	Tecan
Thermomixer	Eppendorf
Thermoprinter for Gel Images	Mitsubishi
Trans Blot Turbo	Biorad
Transilluminator	Peqlab
UV Stratalinker 2400	Stratagene
Vortexer	IKA Labortechnik
Waterbath	EMBL
Waterbath for Cell Culture	GFL
nanoAcquity UPLC system	Waters
LTQ Orbitrap Velos Pro	Thermo Scientific Fisher
Sonifier Cell Disruptor B15	Branson

5.2 Methods

All commercial kits and reagents were used according to the manufacturer's instructions unless otherwise stated.

5.2.1 Molecular Biology Methods

5.2.2 Plasmid constructions

The plasmids pBluescript-BLEF and p-Bluescript-BmL(EF)m were described before (Gebauer et al., 2003). The plasmids p-Bluescript-BREF and p-Bluescript-BmR(EF)m were generated by exchanging the *Firefly Luciferase* (Fluc) ORF of pBluescript-BLEF and p-Bluescript-BmL(EF)m by *Renilla Luciferase* (Rluc) ORF via restriction digest with *XmaI* and *HpaI*. *Rluc* ORF was PCR amplified from pRL-null plasmid (Promega).

Used cloning primers for p-Bluescript-BREF and p-Bluescript-BmR(EF)m:

XmaI Renilla for (5'-3') TCCCCCCGGGATGACTTCGAAAGTTTATGATCCAGAACAAAGG

HpaI Renilla rev (5'-3') AGCTTTGTTAACTTATTGTTCATTTTTGAGAACTCGCTCAACGAA

The plasmid carrying Fluc ORF for *in vitro* transcription was previously described (Iizuka et al., 1994).
The plasmid carrying chloramphenicol acetyl transferase (CAT) ORF for in vitro transcription was described before (Preiss et al., 1998).

5.2.2.1 Polymerase Chain Reaction (PCR)

Phusion High Fidelity (NEB) was used for PCR reactions following the protocol below:

- 5 µL 10 x reaction buffer
- 1 µL dNTPs (each 10 mM)
- 1 µL template Plasmid (100 ng/µL)
- 1 µL forward primer (100 µM)
- 1 µL reverse primer (100 µM)
- 1 µL Phusion HF
- 40 µL H_2O

PCR cycle program:

1. 98°C 30’’
2. 98°C 10’’
3. 55°C 10’’
4. 72°C 1’ per kb
   35 repeats of steps 2.-4.
5. 72°C 10’

5.2.2.2 Digestion of DNA

Restriction digestion enzymes from NEB were used for digestion of DNA. Reaction volumes were either 10 µL with 0.5 µL of each enzyme or 40 µL with 2µL of each enzyme for analytical mini test or preparative reactions respectively. For cloning, linearized vectors were additionally incubated for 1h with 1 µL of calf intestine phosphatase (CIP; NEB). DNA fragments and digested vectors were separated on a 1% agarose gel in 1x TBE Buffer and extracted from the gel using the Qiagen gel extraction kit.

5.2.3 Ligation

DNA ligation of sticky ends was performed using the T4 DNA Ligase (NEB) in its respective buffer for 2h at 25°C in a final volume of 20 µL.

5.2.3.1 Transformation in chemical competent *E.coli* cells

For molecular cloning, plasmids carrying Ampicillin resistance were transformed into chemically competent *Escherichia coli* TOP10 cells. The used TOP10 cells did not need any heat shock for efficient transformation. 50 µL cells were added to the ligation mix and incubated 20' on ice. Afterwards cell were transferred on LB plates with Ampicillin.

5.2.3.2 Isolation of plasmid DNA from *E.coli*

Single colonies were picked and inoculated in 4 mL LB medium with Ampicillin . For Maxi Culture 100-400 mL culture was prepared. Plasmid DNA was isolated using either mini (Sigma; Qiagen) or maxi prep kit (Qiagen) according to the manufacturer's instructions.

5.2.4 Standard Biochemistry Methods

5.2.4.1 In vitro transcription

10 µg Plasmid templates for the Rluc, Fluc and CAT reporter mRNAs were linearized with *HindIII*. 1 µg of the Plasmid was used as template for in vitro transcription using the T7 RNA in vitro transcription kit (MEGAscript®; ThermoAmbion) or T3 RNA in vitro transcription kit. The capping with m7GpppG cap (Darżynkiewicz lab) was included during the in vitro transcription protocol:

- 1 µg linearized template plasmid
- 2 µL ATP (75 mM)
- 2 µL CTP (75 mM)
- 2 µL UTP (75 mM)
- 2 µL CAP (40 mM)
- 2 µL 10x reaction buffer
- 2 µL Enzyme Mix
- Fill up to 10 µL with RNase free H_2O
- Incubate first 5´at 37°C for capping then add 2 µL GTP (75 mM) to perform in vitro transcription for 4 h at 37°C.

After purification using Quick-RNA™ MiniPrep kit (Zymo), aliquots were analyzed by Agarose gel electrophoresis to confirm integrity. For this, the RNA was denatured in Gel Loading Buffer II (Ambion) for 5' at 95°C and loaded on a 1% Agarose Gel and run in 1x TBE Buffer. Fluc and Rluc mRNAs were additionally tested in an in vitro translation assay. Rluc RNA contained a poly(A)

tail of 72 nt, Fluc RNA contained a poly(A) tail of 73 nt CAT RNA contained a poly(A) tail of 15 nt.

5.2.4.2 Translation Assays

Drosophila melanogaster embryo translation extracts were taken from previous lab members and were prepared as described before, without micrococcal nuclease treatment for translation of reporter RNAs to occur in a competitive mode as *in vivo* (Gebauer et al., 1999). For luciferase assays in *Drosophila* embryo extracts, 6 fmol of Fluc and 6 fmol of Rluc reporter mRNA were translated in a final volume of 10 µl containing 16 mM HEPES/KOH (pH 7.4), 50 µM Spermidine, RNase Inhibitor, 100 µM GTP, 800 µM ATP, 80 mM KOAc,2 mM Mg(OAc)2, 60 µM amino acids, 20 mM creatine phosphate, 400 ng creatine kinase, and 40% Drosophila m. embryo extract for 90 min at 25°C. Luciferase activities were assayed with the Dual Luciferase Assay System (Promega) in a microplate luminometer (Berthold) with the following program:

Fluc measurement:

Inject: 10 µL volume Dual-Glo ® Luciferase Buffer + Substrate

Delay: 1.6 seconds

Rluc measurement:

Inject: 10µL volume Dual-Glo ® Stop & Glo ® Buffer + Substrate

Hela cytoplasmic extracts (Cilbiotech) were precleared for 5’ at 13.000g.

For luciferase assays in precleared HeLa cytoplasmic extracts (Cilbiotech), 4 fmol of Fluc and 4 fmol of Rluc reporter mRNA were translated in a final volume of 10µL containing 16 mM HEPES/KOH (pH 7.4), 50 µM Spermidine, RNase Inhibitor (produced in house), 100 µM GTP, 800 µM ATP, 80 mM KOAc,2 mM Mg(OAc)2, 60 µM amino acids, 20 mM creatine phosphate, 400 ng creatine kinase, and 40% HeLa cytoplasmic extract for 30 min at 37°C. Luciferase activities were assayed with the Dual Luciferase Assay System (Promega) in a microplate luminometer (Berthold) (see above).

5.2.4.3 Real-Time PCR

Isolated and proteinase K-digested RNAs as well as total RNA from whole-cell lysate, were purified using Quick-RNA™ MicroPrep kit (Zymo), and then retrotranscribed with random primers (LifeTechnologies) into cDNA (Superscript II, LifeTechnologies) CAT RNA was spiked in as efficiency control. First, 10 µL RNA solution, 1 µL dNTPs (10 mM) and 1 µL random primers (150 ng/µL, LifeTechnologies) were incubated at 65°C for 5’ and then let cool down. Then 2 µL DTT (0.1 M), 4 µL 5x Buffer, 1 µL RNAsIn (prepared in house) and 1 µL Superscript II reverse transcriptase was added. The reaction was performed for 10’ 25°C, 60’ 42°C, 10’ 70°C. For each RNA a reaction without reverse transcriptase was added as control (-RT control). Reverse-transcriptase quantitative PCR (RT-qPCR) was performed by SYBR green (Applied Biosystems) with specific primers against different RNAs in StepOne Real time PCR system (LifeTechnologies).

Used RT-qPCR primers (all from 5’ to 3’; f:forward; r:reverse):

hs18S rRNA: f: GAAACTGCGAATGGCTCATTAAA; r: CACAGTTATCCAAGTGGGAGAGG

hsGAPDH: f: GTGGAGATTGTTGCCATCAACGA; r: CCCATTCTCGGCCTTGACTGT

hsbetaActin: f:TCACCGGAGTCCATCACGAT; r:CGCGAGAAGATGACCCAGAT;

dm18S rRNA: f:CGGAGAGGGAGCCTGAGAA; r:AGCTGGGAGTGGGTAATTTACG

dmGAPDH: f:TTCACCACCATTGACAAGGC; r:CTTCATGTCGGGGCTGTAGG

Rluc: f:GAATTTGCAGCATATCTTGAACCAT; r:GGATTTCACGAGGCCATGATAA

Fluc: f:CCTCTGGATCTACTGGGTTACCTAAG; r:TCTGGCATGCGAGAATCTGA

CAT: f:GCAAGATGTGGCGTGTTACG; r:AAAACGGGGGCGAAGAAGTT

5.2.4.4 Protein quantification with DC protein assay

HeLa cell extracts and *Drosophila melanogaster* embryo extracts were tested for protein content before the application in experiments. For this purpose DC protein assay (Biorad) was used. For every measurement a standard with 7 BSA concentrations between 0-20 µg was prepared. Different amounts of the

test samples were prepared in triplicates. 125 µL of Buffer A and 1 mL of Buffer B were added to the samples. After 15' incubation at RT the samples were measured in cuvettes at 750 nm using a spectrophotometer (Ultraspec 2100 pro, Amersham Biosciences).

5.2.4.5 Silver Staining of Protein Polyacrylamide gels

For protein analysis eluates from RNA pull downs were treated with 0.5 µL RNAse T1 (1 mg/mL; Sigma) and 0.5 µL RNase A (2 mg/mL; Sigma) or 1 µl Benzonase (Novagen, 25 U/µl). Subsequently, proteins were loaded on a 10% Tris-Glycine SDS-polyacrylamide gel or a 4-15% TGX™ Precast gel (Biorad). First the gel was fixed for 20' in fixing solution (50% Methanol, 5% Acetic Acid). Then washing of the gel was performed for 5' in the respective washing solution (washing 1: 50% Ethanol; washing 2: 30% Ethanol). Afterwards, the gel was washed for another 10' in H_2O. The gel was then impregnated for 60'' with freshly prepared sodium thiosulfate solution (0.02%). After three times 20-30'' washings in H_2O, the gel was incubated for 20' in silver solution (6 mM silver nitrate, 0.0185% formaldehyde). Then the gel was rinsed three times for 20-30'' in H_2O. The stain was developed by adding developer solution (2% sodium carbonate (anhydrous), 0.0185% formaldehyde, 0.0004% sodium thiosulfate). The developing reaction was stopped by adding stopping solution (5% acetic acid). Before scanning the silver stained gel the gel was stored in H_2O.

5.2.4.6 Western Blot Analysis

For protein analysis, eluates from RNA pull downs were treated with 0.5 µL RNAse T1 (1 mg/mL; Sigma) and 0.5 µL RNase A (2 mg/mL; Sigma) or 1 µl Benzonase (Novagen, 25 U/µl). Proteins were resolved on a 10% Tris-Glycine SDS-polyacrylamide gel or a 4-15%™ TGX™ Precast gel (Biorad) transferred onto 0.45 µm pore size nitrocellulose membranes (Whatman) with the Semidry Protein Blotter or on a 0.2 µm pore size nitrocellulose membrane with the turbo Blot system (Biorad). Protein loading was assessed by Ponceau Red (Sigma) staining. Membranes were then blocked in PBS containing 0.1% Tween-20 and 5% milk powder, incubated with antibodies diluted in the same solution over-night at 4°C, and washed in PBS containing 0.1% Tween-20. Horseradish Peroxidase (HRP) -coupled secondary antibodies (Santa Cruz or GE

Healthcare) in combination with Immobilon Western HRP substrate (Millipore) and X-ray films (Advansta) were used for detection. Films were documented using Kodak RP-X-OMAT.

5.2.5 Design of LNA/DNA mixmer oligonucleotides

LNA/DNA mixmer oligonucleotides were designed antisense to the RNA of interest. For efficient performance of the oligonucleotide different design guidelines were considered as the incorporation of LNA in an oligonucleotide sequence can strongly affect the properties of the probe. Regions of strong secondary structure in the target RNA were avoided. Secondary structure was either predicted using secondary structure prediction tools (mfold: Zuker, 2003) or making use of published RNA secondary structure (ribosomal RNA). The length of the oligonucleotides was designed to be between 21-23 nucleotides. The specificity of the antisense sequence was checked by performing a BLAST search for possible other RNA targets. Single LNA residue stabilizes not only base pairing with the complementary RNA nucleotide, but also that of immediate DNA neighbors with their complementary RNA bases. Thus, LNA/DNA mixmers confer substantial duplex stabilization, making the use of all-LNA probes unnecessary. The end positions of the oligonucleotide were avoided in order to keep flexibility of the probe for hybridization. Self-complementarity and dimerization can increase when positioning of the LNA is wrong. The general design tools provided by Exiqon (https://www.exiqon.com/oligo-tools) were used to check melting temperature (Tm), self-complementarity and dimerization. At the 3'end a flexible C6 linker carrying a primary Amine was introduced for covalent coupling chemistry. All oligonucleotides were purified by Exiqon via HPLC purification.

LNA scramble: A+CAC+TTAAC+CGTA+TAT+TCC+TA/3AmMO
LNA1: ACC+TATA+AGA+AC+CAT+TACC+AGA/3AmMO/
LNA2: T+TCAA+TCA+CAT+CTA+CTAC+ACTT/3AmMO/
LNA 18S rRNA: T+TAA+TCA+TGG+CCTC+AGTT+CCGA/3AmMO/
LNA 28S rRNA: CC+AA+TCC+TTAT+CCCG+AAGTT+AC/3AmMO/

5.2.6 Oligonucleotide coupling to carboxylated magnetic beads

For the coupling procedure, DNA/RNA LoBind tubes (Eppendorf) were used to reduce the loss of oligonucleotide to the tube surface. During the coupling process, the oligonucleotide was coupled to the carboxy activated surface of the magnetic beads (Perkin Elmer MPVAC11) via its 3' Aminomodifier. Coupling was performed in 100 µL bead slurry portions (50 mg/mL). First, beads were washed three times in 50 mM MES buffer pH 6. The coupling activator N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC-HCl; Sigma) was freshly prepared in a 20 mg/mL solution in MES buffer and cooled on ice before usage. To 100 µL beads 100 µL of oligonucleotide (100 pmol/µL) and 500 µL of EDC-HCl solution was added. Coupling was performed for 5 h at 50°C shaking at 800 rpm. Temperature of 37°C and shorter incubation time for 2 h was tested as well together with the described standard protocol with a bead slurry volume of 5 µL using oligo. For coupling tests, 20 mers of oligodA with a Cy5 label at its 5'end and a primary Amine at its 3'end were used. Coupling performance for the test protocol was controlled by fluorescence measurement and for the standard protocol coupling efficiency was controlled via NanoDrop measurement. For NanoDrop measurement around half of the nucleotide amount was depleted from the supernatant after coupling. After coupling the beads were washed two times with PBS and residual carboxyl residues were inactivated by incubating the beads with 200 mM Ethanolamine ph 8 for 1h at 37°C. Beads were finally washed three times with 1 M NaCl and stored in PBS supplemented with 0.1% Tween.

5.2.7 Oligonucleotide coupling to magnetic Epoxy beads

For the coupling procedure, DNA/RNA LoBind tubes (Eppendorf) were used to reduce the loss of oligonucleotide to the tube surface. During the coupling process the oligonucleotide was coupled to the epoxy activated surface of the magnetic beads (Perkin Elmer MPVA E01) via its 3' Aminomodifier. For the coupling efficiency tests, coupling was performed in 5 µL bead slurry portions (50 mg/mL). First, beads were washed three times in 0.1 M sodiumphosphate buffer pH 8. For coupling tests 20mers of oligodA with a Cy5 label at its 5'end and a primary Amine at its 3'end were used. The coupling was performed in a volume of 100 µL buffered with 0.1 M sodiumphosphate buffer pH 8 for 48 h at 25°C or 37°C respectively. The beads were continuously shaking at 800 rpm

during the coupling. Coupling performance was controlled via Fluorescence measurement. After coupling the beads were washed two times with PBS and residual epoxy residues were inactivated by incubating the beads with 200 mM Ethanolamine ph 8 for 1 h at 37°C. Beads were finally washed two times with 1 M NaCl and stored in PBS supplemented with 0.1% Tween-20.

5.2.8 Fluorescence measurements for oligonucleotide coupling and hybridization tests

Fluorescence was measured in a TECAN Safire II microplate reader, using the following parameters: Cy5: the optimal excitation and emission are 630 nm and 670 nm; Cy3: the optimal excitation and emission are 520 nm and 590 nm. In all cases bandwidth was restricted to 5 nm and gain was set using a well of the prepared standard. A standard was prepared for the fluorescence oligonucleotide in beads and in buffer. Measurements were prepared in triplicates for coupling and hybridization conditions. Fluorescent oligonucleotides were synthesized by MWG. For coupling tests 20mers of oligodA with a Cy5 label at its 5'end and a primary Amine at its 3'end were used. For hybridization tests 20mers of oligodT with a Cy3 label at its 3'end were used. Hybridization was performed in 100 µL volume in the lysis buffer used in interactome capture for 1 h at 4°C rotating. Washings after the hybridization were performed two times with buffers of the interactome capture with a volume of 750 µL per wash. The hybridized oligonucleotide was not released but measured on the bead by fluorescence measurement. Non-hybridized oligonucleotide was measured in the supernatant.

5.2.9 Cell culture

Adherent HeLa cells (American type culture collection (ATCC), cat. no. CCL-2) were cultured in Dulbecco's Modified Eagle's Medium (Sigma-Aldrich) with 4.5 mg/L D-glucose (Sigma-Aldrich) supplemented with 10% heat inactivated fetal bovine serum (Gold, GE Healtcare), 2 mM L-glutamine (Sigma-Aldrich) and 10 mL/L penicillin/streptomycin mix (Sigma-Aldrich; P4333) at 37°C in 5% CO_2. Cells were seeded in 500 cm^2 dishes (Corning) and grown till 70-90% confluency.

5.2.10 RNA interactome captures

The methods used and the buffers for lysis and the washings were based on a published protocol (Castello et al. 2012; Castello et al. 2013).

5.2.10.1 HeLa RNA interactome capture

HeLa cells with a confluency of 80-90% on a 500 cm^2 dish were washed two times with 30 mL of PBS. Afterwards all the liquid was removed and the cells were irradiated on ice with an energy of 150 mJ/cm^2 at 254 nm in a Stratalinker. For the non-crosslinked control the cells were processed in parallel with the same washings and incubation on ice. HeLa cells from 1x500 cm^2 dish were harvested and lysed in 4 mL of lysis buffer supplemented with protease inhibitor (complete EDTAfree, Roche) and RNase Inhibitor (RNAsIn produced in house), and homogenized with five strokes using a narrow gauge needle (0.4 mm diameter). After an Input sample was taken, Poly(A)+ mRNAs and crosslinked proteins were captured with oligo(dT)25 magnetic beads (NE Biolabs). Subsequently, oligo(dT)25 beads were washed two times with 1.9 mL of buffers containing decreasing concentrations of LiCl and LiDS (Buffer 1; Buffer2; Buffer 3). RNAs and crosslinked proteins were eluted with RNase free H_2O at 90°C. For RNA analysis, samples were digested with proteinase K in proteinase K buffer at 55°C for 1h. RNA from Input and eluates was isolated from a 10µL aliquot with RNA Microprep Kit (Zymo Research). For protein analysis, samples were treated for 1 h at 37°C with 0.5 µL RNAse T1 (1 mg/mL; Sigma) and 0.5 µL RNase A (2 mg/mL; Sigma), and released proteins were analyzed by silver staining or Mass Spectrometry.

5.2.10.2 Specific RNP capture in HeLa translational active extracts

Single RNP capture was performed using 3 mL of cleared HeLa cytoplasmic extracts (Cilbiotech) in which translational assays of 7.5 mL total volume with heat denatured 65 µg Rluc reporter RNA (RNA with wt = BREF or mut Sxl binding sites BmR(EF)m, see plasmid construction) and 4xmolar excess of recombinant N-terminal GST-tagged Sxl RNA binding domains (a generous gift from Medenbach lab) were performed at 25°C for 1 h. Afterwards translational extracts were evenly spread on 500 cm^2 cell culture dishes and irradiated with an energy dose of 300 mJ/cm^2 at 254 nm. 5x lysis buffer supplemented with

protease inhibitor and RNasIn (complete EDTAfree, Roche) was added to the extracts. An input sample was taken from the processed lysate. 250 µl LNA/DNA mixmer oligonucleotide coupled magnetic beads targeting the Rluc reporter sequence of the RNA were 3x washed with TE buffer and then blocked with 200 µg yeast tRNA for 15 min at 30°C. Subsequently, beads were washed in lysis buffer and heat denatured for 3' at 95°C. The crosslinked lysate was added to the beads and hybridization was performed for 2 h at 41°C in a waterbath with intermittent shaking for a pulldown with LNA beads and for 2 h at 25°C for a pulldown with oligodT beads. Afterwards, beads were washed two times with buffers containing decreasing concentrations of LiCl and LiDS (Buffer 1; Buffer 2; Buffer 3) as described previously (Castello et al. 2012; Castello et al. 2013). RNAs and crosslinked proteins were eluted with TE buffer at 90°C for 3'. Three rounds of pull down were performed. For RNA analysis, samples were digested with proteinase K in proteinase K buffer at 55°C for 1h. RNA from Input and eluates was isolated from a 10µL aliquot with RNA Microprep Kit (Zymo Research). For protein analysis, samples were treated for 1 h at 37°C with 0.5 µL RNAse T1 (1 mg/mL; Sigma) and 0.5 µL RNase A (2 mg/mL; Sigma), and released proteins were analyzed by silver staining and Western Blotting.

5.2.10.3 Specific RNP capture in Drosophila melanogaster embryo extracts

10 µg (32 pmol) *in vitro* transcribed reporter RNA was heat denatured at 95°C for 3 min in LoBind tubes (Eppendorf). Then 2 mL *Drosophila melanogaster* embryo extract was added to the cooled down RNA and incubated for 30 min at 25°C in a waterbath. Afterwards, the extract was evenly spread on a 15 cm cell culture dish on ice and irradiated with an energy dose of 300 mJ/cm^2 at 254 nm. After crosslinking, 2 mL of a 2x lysis buffer supplemented with RNase Inhibitor (RNAsIn produced in house) and protease inhibitor (complete, Roche) was added. After snap freezing the lysate in liquid nitrogen, the beads were prepared for the pulldown. 200 µL with LNA/DNA mixmer oligonucleotide coupled magnetic beads were 3x washed with TE buffer and then blocked with 200 µg yeast tRNA for 15 min shaking at 30°C. Subsequently, beads were washed in 1x lysis buffer and heat denatured for 3' at 95°C. The crosslinked lysate was added to the beads and hybridization was performed for 2 h at 41°C in a waterbath with intermittent shaking. Afterwards beads were washed two

times with buffers containing decreasing concentrations of LiCl and LiDS (Buffer 1; Buffer 2; Buffer 3) as described previously. RNAs and crosslinked proteins were eluted with TE Buffer pH 7.5 for 3' at 90°C. The pulldown was performed in four replicates and in three different set ups. LNA1 was targeting the wild type or the mutant *in vitro* transcribed Rluc reporter, whereas LNAscr control was used for the wild type Rluc reporter. For RNA analysis, samples were digested with proteinase K in proteinase K buffer at 55°C for 1h. RNA from Input and eluates was isolated from a 10 µL aliquot with RNA Microprep Kit (Zymo Research). For protein analysis via Mass Spectrometry, the eluates were concentrated to 100 µL on a Amicon filter column with a cut off of 3 kDa. At the proteomics core facility RNase digest with Benzonase was performed before starting the sample preparation (see Mass Spectrometry sample preparation).

5.2.10.4 Specific RNP capture of 18S and 28S rRNA in HeLa cells

HeLa cells with a confluency of 70% on a 500 cm^2 dish were washed two times with 30 mL of PBS. Afterwards, all the liquid was removed and the cells were irradiated on ice with an energy dose of 150 mJ/cm^2 at 254 nm in a Stratalinker. For the non-crosslinked control the cells were processed in parallel with the same washings and incubation on ice. HeLa cells from 1x500 cm^2 dish were harvested and lysed in 4 mL of lysis buffer supplemented with protease inhibitor (complete EDTAfree, Roche) and RNase Inhibitor (RNAsIn produced in house), and homogenized with five strokes using a narrow gauge needle (0.4 mm diameter). 4x 500 cm^2 plates were prepared per sample to get a final volume of 16 mL lysate. After snap freezing in liquid nitrogen, lysate was thawed, an Input was taken and the lysate was pre-warmed in the waterbath to 60°C. Then 15% v/v Ethylenecarbonate was added to the lysate as hybridization enhancer. 300 µL LNA/DNA mixmer oligonucleotide coupled magnetic beads were three times washed with TE buffer and then blocked with 300 µg yeast tRNA for 15 min at 30°C. Subsequently, beads were washed in lysis buffer and heat denatured for 3' at 95°C. The beads were added to 16 mL of lysate (coming from 4x 500cm^2 dishes) and hybridization was performed for 2 h at 41°C in a waterbath with intermittent shaking. For each experiment three samples were prepared, one with the specific 18S or 28S LNA with crosslinked lysate, one with the specific 18S or 28S LNA with non-crosslinked lysate and one with a scramble LNA

control on crosslinked lysate. After hybridization, beads were washed three times in 1.9 mL of lysis buffer and then twice in 1.9mL of buffers with decreasing amount of LiDS and LiCl (Buffer 1, Buffer 2 and Buffer 3). Afterwards captured RNA was eluted with 150 µL RNase free H_2O for 3' at 90°C. Six pulldown rounds were performed. Each eluate was checked on Bioanalyzer (Agilent) for specific enrichment of the respective ribosomal RNA, then the six elution rounds were pooled. For RNA analysis RNA from input and eluates was isolated with RNA Microprep Kit (Zymo Research) after proteinase K treatment from a 10 µL aliquot. For Mass Spectrometry, samples were treated for 1 h at 37°C with Benzonase and then alkaline hydrolysis was performed to remove any residual RNA. For alkaline hydrolysis the sample was incubated at pH 12 (adjusted with NaOH) for 1 h at 60°C. Afterwards the pH was neutralized again with HCl. The sample was buffered by adding HEPES buffer to get final 50 mM HEPES pH 7.5. The protein sample was then submitted for Mass Spectrometry to the proteomics core facility.

5.2.10.5 18S rRNA capture after Nuclear-Cytoplasmic Fractionation from HeLa cells

HeLa cells from 1x 500 cm^2 dish were washed and crosslinked as mentioned above in specific capture of 18S rRNA. The cell pellet was re-suspended in three volumes of nuclear lysis buffer (Nuclear buffer + 0.5% NP40). After pipetting 5 times up and down the lysate was incubated for 5' on ice. Then the lysate was centrifuged for 3' at 3000 rpm at 4°C. The supernatant (cytoplasmic fraction) was taken and again spun for 3' at 3000 rpm to remove any residual nuclei. The supernatant was then taken and mixed with 2x LiDS lysis buffer. The Nuclei were washed three times with Nuclear Buffer that does not include NP40. Then two volumes of Nuclear Buffer with 0.5% NP40 was added to the Nuclei and incubated with the Nuclei for 10' on ice. The Nuclear fraction was sonicated with three cycles of 10'' (3 cycles of 10 '', 50% duty, level4, 15'' pause). After sonication 200x DNAse Buffer and 1µL DNaseI (Ambion) and 1µL TurboDNAse (Ambion) were added and incubated for 10' at 37°C. After digestion the Nuclear fraction was mixed with 2x LiDS lysis buffer (added 5 mM more EDTA to chelate Mg Ions). Inputs from the Nuclear and Cytoplasmic fractions were taken to confirm clean fractionation on Western Blot by using antibodies raised against the nuclear marker Ki67 and the cytoplasmic marker

Tubulin. 18S rRNA was performed as described before with 100 µL beads and 2 rounds of hybridization. For RNA analysis RNA from input and eluates was isolated after proteinase K treatment from a 10µL aliquot with RNA Microprep Kit (Zymo Research).

5.2.11 Sample Preparation for Mass Spectrometry

5.2.11.1 Sample Preparation for Mass Spectrometry (specific RNP capture in *Drosophila melanogaster* embryo extracts)

Samples were submitted to EMBL Proteomics Core Facility in TE Buffer. 1x Protease Inhibitor Cocktail (Roche) and 1 µl Benzonase (Novagen, 25 U/µl) was added to the samples and incubated at 37°C for 1 h. Cysteines were reduced with 1 µl of 200 mM DTT (Biomol) in 50 mM HEPES (Biomol) for 30 min. at 57 °C. The sample was cooled to 24 °C and 2 µL of 400 mM IAA (Merck) dissolved in 50 mM HEPES was added and incubated in dark for 30 min. at 24 °C. A novel protocol using paramagnetic beads, termed Single-Pot Solid-Phase-enhanced Sample Preparation (SP3) (Hughes et al. 2014) was used to prepare the samples for LC-MS/MS. The proteins were digested using trypsin (Promega) with an enzyme to protein ratio 1:50 at 37°C overnight. Peptides were labeled on magnetic rack using 1% formaldehyde (CH2O, CD2O, 13CD2O) and 0.2 M sodium cyanoborohydride solution (NaBH3CN, NaBD3CN) for 1 hour at room temperature.

5.2.11.2 Sample Preparation for Mass Spectrometry (human rRNA interactome)

~1 mL RNAse treated samples were submitted to EMBL Proteomics Core Facility in 50 mM HEPES buffer pH 7.5. Proteins were precipitated by adding ice cold trichloractetic acid (TCA) to 4 parts of protein sample. Then the sample was left on ice for 20-30 min. Samples were spun for 20 min at 4°C at 20k g. TCA supernatant was removed and the protein pellet was washed with 1 mL of 10% TCA, vortexed and spun again for 20 min at 4°C at 20k g. The washing step was repeated twice with 1 mL of acetone and spun for 30 min at 4°C at 20k g. Afterwards the pellet was air dried and solubilized in 60 µL of 50 mM HEPES. Cysteines were reduced with 1 µl of 200 mM DTT (Biomol) in 50 mM HEPES (Biomol) for 30 min. at 57 °C. The sample was cooled to 24 °C and 2 µL of 400

mM IAA (Merck) dissolved in 50 mM HEPES was added and incubated in dark for 30 min at 24 °C. Proteins were then digested with 800 ng Trypsin (Promega) at 37°C overnight. Peptides were desalted and labeled with high, medium and low dimethyl-labeling reagent using OASIS 96 well plate (Waters). Afterwards the labeled peptides were eluted from the plate with Buffer B (60% methanol, 1% formaldehyde in water).

5.2.12 LC-MS/MS

Measurements were performed by the EMBL proteomics core facility. Peptides were separated using the nanoAcquity ultra performance liquid chromatography (UPLC) system (Waters) fitted with a trapping (nanoAcquity Symmetry C18, 5 µm, 180 µm x 20 mm) and an analytical column (nanoAcquity BEH C18, 1.7 µm, 75 µm x 200 mm). The outlet of the analytical column was coupled directly to a LTQ Orbitrap Velos Pro (Thermo Fisher Scientific) using the Proxeon nanospray source. The samples were loaded with a constant flow of solvent A (0.1% formic acid) at 5 µL/min onto the trapping column. Trapping time was 6 min. Peptides were eluted via the analytical column with a constant flow of 0.3 µL/min. During the elution step, the percentage of solvent B (acetonitril, 0.1% formic acid) increased in a linear fashion from 3% to 7% within 10 min., then increased to 25% within 100 min (60 min for human ribosomal rna interactome) and finally to 40% within a further 10 min (5 min for human ribosomal rna interactome). The peptides were introduced into the mass spectrometer (Orbitrap Velos, Thermo Fisher Scientific) via a Pico-Tip Emitter 360 µm OD x 20 µm ID; 10 µm tip (New Objective) and a spray voltage of 2.2 kV was applied. The capillary temperature was set at 300°C. Full scan MS spectra with mass range 300-1700 m/z were acquired in profile mode in the FT with resolution of 30000. The filling time was set at maximum of 500 ms with limitation of 1.0 x 10^6 ions. The most intense ions (up to 15) from the full scan MS were selected for sequencing in the LTQ. Normalized collision energy of 40% was used, and the fragmentation was performed after accumulation of 3.0 x10^4 ions or after filling time of 100 ms for each precursor ion (whichever occurred first). MS/MS data was acquired in centroid mode. Only multiply charged (2+, 3+, 4+) precursor ions were selected for MS/MS. The dynamic exclusion list was restricted to 500 entries with maximum retention period of 30 s and relative

mass window of 10 ppm. In order to improve the mass accuracy, a lock mass correction using the ion (m/z 445.12003) was applied.

5.2.13 Proteomic Data Analyses

Analysis was performed by the EMBL proteomic core facility.All data-dependent data were analyzed using MaxQuant (v 1.4.1.2). Carbamidomethylation of cysteine was specified as a fixed modification, acetylation of the protein N-terminus and oxidation of methionine as variable. More variable modifications of the N-terminus and lysine were specified for masses of +28.0313 (light dimethyl), +32.0564 (medium dimethyl), +36.0757 (heavy dimethyl) and oxidation of methionine in all searches. The data were searched against the Uniprot_Drosophila_20140519 database. For all searches, trypsin was specified with 2 missed cleavage allowed.

5.2.14 Bioinformatic Analyses

Performed by Bernd Fischer

5.2.14.1 Bioinformatic Analysis for RNP capture in *Drosophila melanogaster* embryo extracts

Peptides were mapped back to the Uniprot protein database (Version 2014_06). Peptides uniquely mapping to one gene model and to the largest protein translation of this gene model were used for subsequent analysis. Protein log2-intensity ratios were computed by averaging the peptide log2-intensity ratios per protein. Protein ratios were divided by the median deviation to compensate for scaling effects. For comparison of wild type to mutant proteins, the median ratio was subtracted to center the ratios around zero.

Protein ratios were tested against the hypothesis that the ratio is zero by a moderated t-test (Lönnstedt & Speed 2002) implemented in the R/Bioconductor package limma (Gordon K Smyth 2004). The method of Benjamini-Hochberg (Benjamini & Hochberg 1995) that controls for the false discovery rate was used to adjust the p-values for multiple testing.

5.2.14.2 Bioinformatic Analysis for RNP capture of human ribosomal RNA

Peptides were mapped back to the Uniprot protein database (Version 2014_06). Peptides uniquely mapping to one gene model and to the largest protein

translation of this gene model are used for subsequent analysis. To compare the conventional cross-linking fraction to the non-cross linking fraction a semi-quantitative approach was chosen. For each sample and each protein, it was assessed, if the protein was identied for at least one protein. Then it was counted in how many samples each protein is quantied for CCL samples (nCCL), for NoCL samples (nNoCL) and for Scr samples (nScr). We counted and list all genes that are quantified in at least two more CCL samples than in NoCL samples. The false discovery rate was estimated by the reverse events (at least two more quantifications in the NoCL samples). 1 was added to the estimated number of false discoveries to reduce the chance of an underestimation of the false-discovery rate due to small numbers.

6 Abbreviations

4SU	4-thiouridine
6SG	6-thioguanosine
ARE	AU rich element
bcd	bicoid
bp	Basepairs
C	Celsius
CAS	CRISPR associated protein
cCl	Conventional Crosslink (254nm)
cDNA	complementary/copy DNA
CDS	Coding sequence
CHIP	chromatin immunoprecipitation
CIP	calf intestine phosphatase
circRNA	circular RNA
CLIP	cross-linking immunoprecipitation
CLIP-seq	cross-linking immunoprecipitation-high-throughput sequencing
CRISPR	clustered interspaced short palindromic repeats
Da	Dalton
DNA	Desoxyribonucleic acid
DMEM	Dulbecco's Modified Eagle Medium
dsRBM	double stranded RNA binding motif
dsRNA	Double stranded RNA
DTT	Dithiothreitol
E. coli	Escherichia coli
ECL	Enhanced Chemoluminscence
eCLIP	enhanced CLIP
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
eEF	eukaryotic elongation factor
eIF	eukaryotic translation initiation factor
eRF	eukaryotic release factor
ES	ribosomal expansion segment
ES cells	embryonic stem cells
ETS	External transcribed spacer
FA Xlink	Formaldehyde crosslinking
FBS	Fetal Bovine Serum
Fluc	Firefly Luciferase
GAPDH	Glyceraldehyde phosphate Dehydrogenase
GO	Gene Ontology
grk	gurken
GTP	guanosinetriphosphate
h	Hour(s)
HEPES	*4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid*

HRP	Horseradish peroxidase
iCLIP	individual-nucleotide resolution CLIP
IFNy	Interferon γ
IRE	Iron responsive element
IRES	Internal ribosomal entry site
IRP1	Iron regulatory protein 1
ITC	Isothermal Calorimetry
ITS	Internal transcribed spacer
J	Joule
k	kilo
KH	hnrnp K homology motif
LC-MS/MS	Liquid chromoatography coupled with tandem mass spectrometry
LiDS	Lithium Dodecyl Sulfate
LNA	Locked Nucleic Acid(s)
lncRNA	long non coding RNA
m7G	7-methylguanosin
MCP	MS2 coat protein
Met	Methionine
min	Minute(s)
miRNA	microRNA
mm	Millimeter
mRNA	Messenger RNA
MS	Mass spectrometry
msl2	male specific lethal 2
mut	mutant
Ncl	Nucleolin
NGD	“No go” decay
NMR	Nuclear magnetic resonance
nos	nanos
NRD	non-functional rRNA decay
nt	Nucleotides
oligodT	oligonucleotide consisting of deoxy Thymidin
ORF	Open reading frame
osk	oskar
PABP	Poly-(A) binding protein
PAIR	PNA assisted identification of RBPs
PAM	protospacer adjacent motif
PAR-CL	Photoactivatable Ribonucleoside Crosslink
PAR-CLIP	PAR-Crosslinking and Immunoprecipitation
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
piRNA	piwi interacting RNA
PNA	peptide nucleic acid
Pol	Polymerase
qPCR	qunatitative Polymerase Chain Reaction

RBD	RNA Binding Domain
RBP	RNA Binding Protein
REM	RNA, enzme, metabolite
RIP	RNA immunoprecipitation
Rluc	Renilla Luciferase
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
Rpl	large subunit ribosomal protein
rpm	Rotations per minute
Rps	small subunit ribosomal protein
RRM	RNA Recognition Motif
rRNA	Ribosomal RNA
RT	room temperature (25°C)
RT-PCR	Reverse transcription polymerase chain reaction
s	Second(s)
scr	scramble(d)
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SILAC	Stable isotope labeling amino acid in cell culture
snRNP	small nuclear recognition particle
SP3	Single-Pot Solid-Phase-enhanced Sample Preparation
Ssx	Sister of Sex lethal
Sxl	Sex lethal
tra	transformer
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
U2AF	U2 auxiliary factor
UNR	Upstream of N-ras
UPLC	ultra performance liquid chromatography
UTR	Untranslated Region
UV	Ultraviolet
V	Volt
wt	wildtype
Znf	Zinc finger motif

7 Supplemental Data

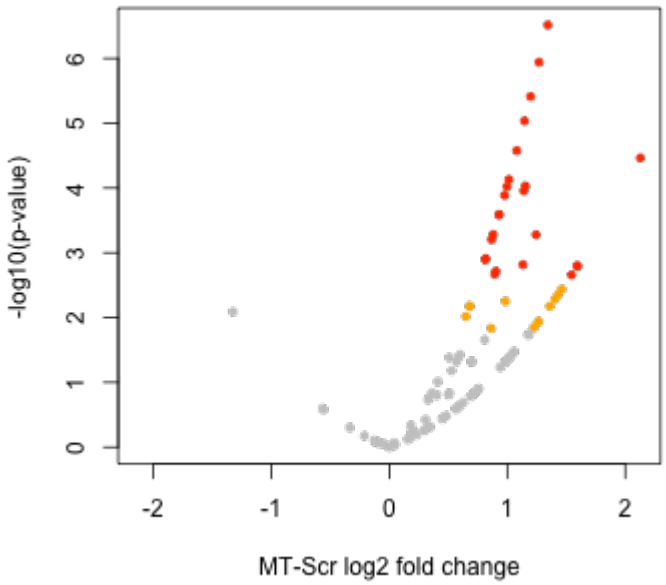

Figure S 7.1 Volcano plot comparing sample enrichments. Enrichments in the mut sample over the scramble sample in 4 biological replicates. Mass spectrometry measurements were performed by the EMBL proteomics core facility Statistical analysis and plot was created by Bernd Fischer The axes display the log2-fold changes (orange: 5% FDR, red: 1% FDR).

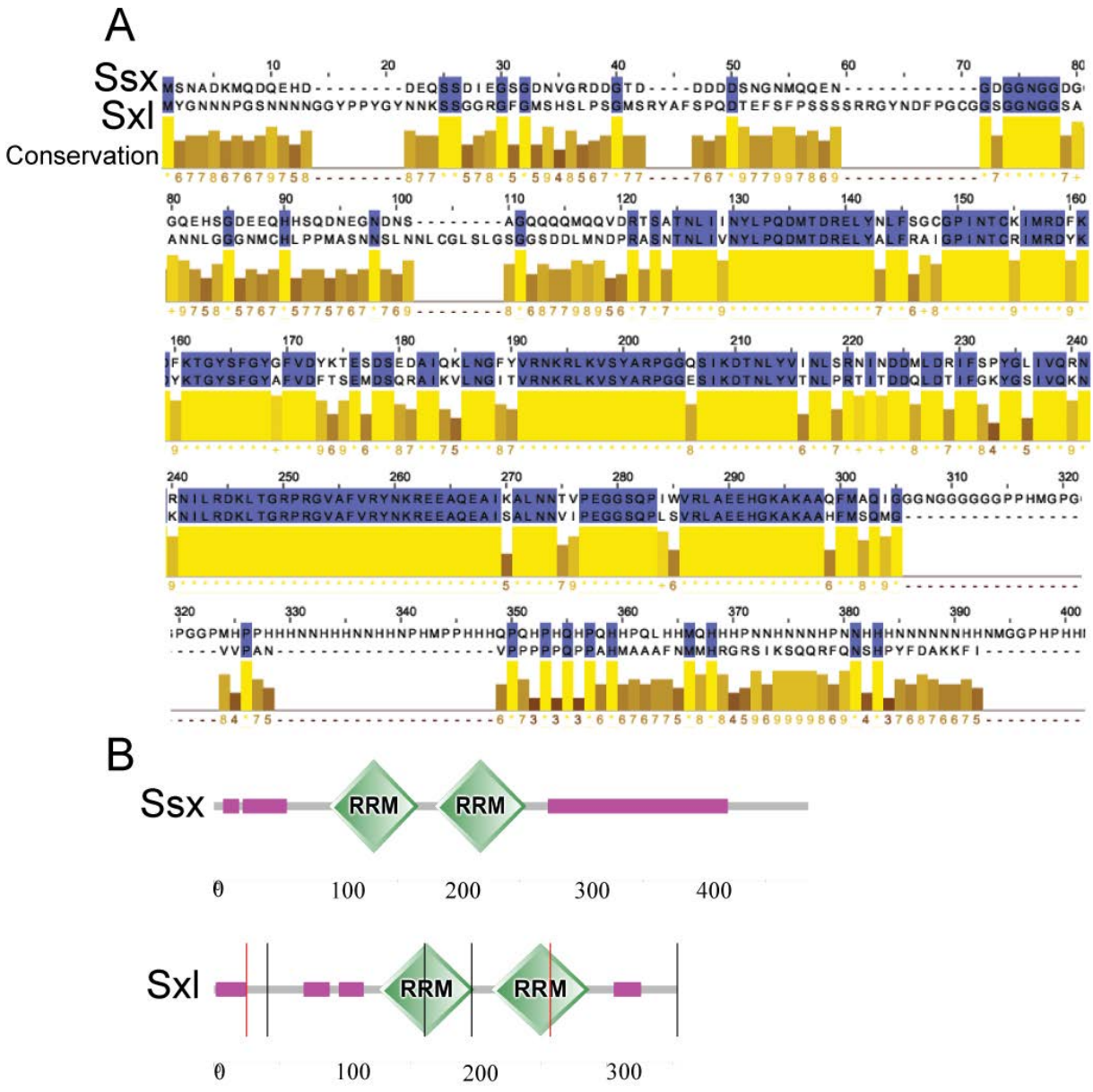

Figure S 7.2; Comparison of the protein sequences of Sxl and Ssx. (A) Alignment of the protein sequences. There is a strong conservation in the two RRMs and part of the N-terminus. (B) Domain architecture s of Sxl and Ssx displayed via SMART (http://smart.embl-heidelberg.de/).

Table S 7.1; log2 fold enrichments from wt/mut RNA in the four biological replicates. Log2FC column gives the mean of the enrichments in the different replicates. ** 1% FDR; * = 5% FDR.

Symbol	WT.MT.rep	WT.MT.rep	WT.MT.rep	WT.MT.rep	log2FC	p.value	p.adj	si
ssx	4,28	NA	NA	2,16	3,22083	0,0000	0,0078	**
bic	2,11	NA	NA	NA	2,11027	0,0592	0,6454	
Hrb98DE	0,43	1,07	5,33	1,13	1,99239	0,0004	0,0254	*
Hrb87F	1,74	1,34	2,52	1,22	1,70213	0,0026	0,0951	
lin-28	1,62	NA	NA	NA	1,62453	0,1452	0,9402	
RpS27A	0,21	NA	2,52	NA	1,36351	0,0843	0,7662	
RpS7	0,21	NA	2,71	0,75	1,22690	0,0574	0,6454	
elav	1,44	NA	NA	0,85	1,14650	0,1459	0,9402	
rump	0,83	2,62	NA	-0,03	1,14280	0,0765	0,7585	
RpS15	-0,31	-2,08	6,42	0,39	1,10330	0,0486	0,6454	
rig	0,98	NA	NA	NA	0,98168	0,3774	0,9402	
RpL32	0,91	NA	NA	NA	0,91407	0,4110	0,9402	
sqd	0,56	1,16	1,02	0,90	0,90845	0,1036	0,8692	
bnb	-0,43	0,42	2,70	0,80	0,87284	0,1177	0,9169	
sp:P21187	1,11	NA	NA	0,56	0,83831	0,2868	0,9402	
RpL4	0,65	NA	NA	NA	0,65186	0,5574	0,9402	
Pabp2	0,61	NA	NA	NA	0,61405	0,5805	0,9402	
larp	0,61	NA	NA	NA	0,61223	0,5816	0,9402	
CG30122	0,80	0,37	NA	NA	0,58123	0,4596	0,9402	
eIF3-S10	0,52	NA	NA	NA	0,52264	0,6380	0,9402	
Ars2	0,48	NA	NA	NA	0,47666	0,6678	0,9402	
vig2	0,60	0,01	0,74	0,37	0,43091	0,4382	0,9402	
Df31	2,51	-0,69	-0,70	0,59	0,42798	0,4413	0,9402	
B52	-0,33	NA	NA	1,16	0,41404	0,5981	0,9402	
vig	0,74	NA	0,06	NA	0,39838	0,6120	0,9402	
nonA	0,39	NA	NA	NA	0,39391	0,7228	0,9402	
RpL30	1,27	0,56	-0,78	NA	0,34773	0,5877	0,9402	
RpS24	1,56	-2,35	2,07	0,03	0,32873	0,5540	0,9402	
CG9915	0,73	0,41	0,25	-0,08	0,32607	0,5572	0,9402	
Imp	0,33	NA	NA	NA	0,32566	0,7693	0,9402	
La	0,33	NA	NA	NA	0,32500	0,7698	0,9402	
RpL15	1,23	0,00	0,32	-0,25	0,32454	0,5591	0,9402	
RpL13	0,32	NA	NA	NA	0,31890	0,7740	0,9402	
RpS12	3,09	0,12	-0,01	-1,94	0,31721	0,5680	0,9402	
lost	-0,37	1,41	-1,06	1,23	0,30241	0,5862	0,9402	
RpS3	-0,02	0,11	NA	0,81	0,30192	0,6378	0,9402	
Nlp	NA	NA	NA	0,29	0,29263	0,7921	0,9402	
Dp1	0,23	NA	NA	NA	0,23383	0,8332	0,9426	
CG3800	-0,51	1,20	-0,49	0,69	0,22092	0,6908	0,9402	
RpL3	0,21	NA	NA	NA	0,20871	0,8509	0,9464	
RpS17	1,09	NA	-0,15	-0,38	0,18552	0,7723	0,9402	
sp:Q9V3T8	1,32	0,72	-1,59	0,28	0,17983	0,7460	0,9402	
RpS2	0,57	-1,30	1,18	0,14	0,14526	0,7936	0,9402	
CG9894	0,21	-0,68	2,77	-1,74	0,13890	0,8024	0,9405	
CG8636	0,10	NA	NA	NA	0,10364	0,9256	0,9836	
RpS4	0,09	NA	NA	NA	0,08584	0,9383	0,9836	
rin	0,16	-0,04	0,31	-0,21	0,05593	0,9197	0,9836	

RpS27	0,04	NA	NA	NA	0,04057	0,9708	0,9878	
Hrb27C	0,95	-0,17	0,01	-0,70	0,02122	0,9695	0,9878	
RpS6	-0,44	NA	NA	0,46	0,01190	0,9878	0,9878	
RpS14a	0,71	NA	NA	-0,74	-	0,9852	0,9878	
Ef2b	-0,07	NA	NA	NA	-	0,9475	0,9836	
RpS16	-0,08	NA	NA	NA	-	0,9438	0,9836	
x16	-0,12	NA	NA	NA	-	0,9165	0,9836	
CG42351	0,68	-0,06	-1,00	-0,11	-	0,8243	0,9426	
eIF-3p66	-0,13	NA	NA	NA	-	0,9094	0,9836	
Tudor-SN	0,55	NA	-0,56	-0,38	-	0,8388	0,9426	
Dek	-0,26	0,91	-0,33	-0,93	-	0,7822	0,9402	
CG1677	0,55	-0,08	-0,90	-0,37	-	0,7171	0,9402	
tral	-0,09	-0,08	NA	-0,48	-	0,7354	0,9402	
yps	-0,81	0,67	-0,51	-0,27	-	0,6809	0,9402	
SF2	-0,24	NA	NA	NA	-	0,8312	0,9426	
Pep	-1,33	1,13	-0,39	-0,43	-	0,6454	0,9402	
RpS11	-0,64	0,32	0,44	-1,20	-	0,6286	0,9402	
lig	0,11	NA	NA	-0,71	-	0,7034	0,9402	
RpS10b	0,48	-1,14	NA	NA	-	0,6738	0,9402	
glo	-0,34	NA	NA	NA	-	0,7622	0,9402	
lark	-0,36	NA	NA	NA	-	0,7485	0,9402	
bel	-0,36	NA	NA	NA	-	0,7460	0,9402	
eIF3-S8	-0,24	NA	-0,49	NA	-	0,6395	0,9402	
Rox8	-0,38	NA	NA	NA	-	0,7340	0,9402	
Fmr1	0,86	-1,45	-0,48	-0,46	-	0,4933	0,9402	
nito	-0,43	NA	NA	NA	-	0,6987	0,9402	
fne	-0,44	NA	NA	NA	-	0,6919	0,9402	
bl	-0,45	NA	NA	NA	-	0,6859	0,9402	
RpL5	-0,46	NA	NA	NA	-	0,6791	0,9402	
asp	-0,46	NA	NA	NA	-	0,6758	0,9402	
CG31813	-0,73	-0,29	-0,65	-0,22	-	0,3959	0,9402	
RpS13	-0,18	-0,80	NA	NA	-	0,5345	0,9402	
RpS25	-0,50	NA	NA	NA	-	0,6558	0,9402	
RpS19a	-0,42	NA	0,31	-1,42	-	0,4269	0,9402	
RpS20	0,04	NA	-0,78	-0,83	-	0,4157	0,9402	
RpL24	1,24	-2,39	0,03	-0,97	-	0,3462	0,9402	
CG5787	-0,57	NA	NA	NA	-	0,6049	0,9402	
RpS9	0,02	-0,10	-1,24	-1,00	-	0,2973	0,9402	
RpS18	0,66	-0,55	-1,72	-0,80	-	0,2795	0,9402	
mod	-0,63	NA	NA	NA	-	0,5738	0,9402	
RpS3A	-0,72	NA	NA	NA	-	0,5180	0,9402	
Capr	-0,82	NA	NA	NA	-	0,4594	0,9402	
nonA-l	-0,84	NA	NA	NA	-	0,4495	0,9402	
RpS26	-0,85	NA	NA	NA	-	0,4458	0,9402	
RpS8	-0,86	NA	NA	NA	-	0,4375	0,9402	
RpL18	-0,88	NA	NA	NA	-	0,4309	0,9402	
Rack1	-0,91	NA	NA	NA	-	0,4121	0,9402	
RpL18A	-0,93	NA	NA	NA	-	0,4043	0,9402	
RpL36	-0,93	NA	NA	NA	-	0,4034	0,9402	

sta	1,45	NA	NA	-3,49	-	0,1971	0,9402	
RpL14	-2,39	NA	NA	0,20	-	0,1640	0,9402	
RpL19	-1,11	NA	NA	NA	-	0,3199	0,9402	
Psi	-1,20	NA	NA	NA	-	0,2829	0,9402	
RpL6	-1,21	NA	NA	NA	-	0,2759	0,9402	
RpL13A	-1,24	NA	NA	NA	-	0,2661	0,9402	
Ref1	-1,10	-0,68	-2,63	-0,64	-	0,0247	0,5384	
CG9775	-1,33	NA	NA	NA	-	0,2320	0,9402	
RpL7	-1,44	NA	NA	NA	-	0,1949	0,9402	
Pdcd4	-1,47	NA	NA	NA	-	0,1869	0,9402	
RpL22	-2,14	NA	NA	NA	-	0,0559	0,6454	
RpL8	-2,30	NA	NA	NA	-	0,0402	0,6454	
CG8108	-2,68	NA	NA	NA	-	0,0170	0,4647	
CG30118	NA	NA	NA	NA	NA	NA	NA	
CG13114	NA	NA	NA	NA	NA	NA	NA	
CoRest	NA	NA	NA	NA	NA	NA	NA	
tko	NA	NA	NA	NA	NA	NA	NA	
RpS15Aa	NA	NA	NA	NA	NA	NA	NA	
RpL27A	NA	NA	NA	NA	NA	NA	NA	
U2af38	NA	NA	NA	NA	NA	NA	NA	
RpS21	NA	NA	NA	NA	NA	NA	NA	
RpL37a	NA	NA	NA	NA	NA	NA	NA	
cib	NA	NA	NA	NA	NA	NA	NA	
sp:P02283	NA	NA	NA	NA	NA	NA	NA	
Yp1	NA	NA	NA	NA	NA	NA	NA	
Yp2	NA	NA	NA	NA	NA	NA	NA	
Cp36	NA	NA	NA	NA	NA	NA	NA	
Cp15	NA	NA	NA	NA	NA	NA	NA	
Ef1alpha48D	NA	NA	NA	NA	NA	NA	NA	
dec-1	NA	NA	NA	NA	NA	NA	NA	
Rm62	NA	NA	NA	NA	NA	NA	NA	
bap	NA	NA	NA	NA	NA	NA	NA	
RpL7A	NA	NA	NA	NA	NA	NA	NA	
Cam	NA	NA	NA	NA	NA	NA	NA	
wibg	NA	NA	NA	NA	NA	NA	NA	
RpS5a	NA	NA	NA	NA	NA	NA	NA	
SmF	NA	NA	NA	NA	NA	NA	NA	
CG14756	NA	NA	NA	NA	NA	NA	NA	
trsn	NA	NA	NA	NA	NA	NA	NA	
CG8635	NA	NA	NA	NA	NA	NA	NA	
CG13326	NA	NA	NA	NA	NA	NA	NA	
CG13083	NA	NA	NA	NA	NA	NA	NA	
Zn72D	NA	NA	NA	NA	NA	NA	NA	
CG32017	NA	NA	NA	NA	NA	NA	NA	
CG31064	NA	NA	NA	NA	NA	NA	NA	
CG3508	NA	NA	NA	NA	NA	NA	NA	
CG32720	NA	NA	NA	NA	NA	NA	NA	
Jafrac1	NA	NA	NA	NA	NA	NA	NA	
Sucb	NA	NA	NA	NA	NA	NA	NA	
Pur-alpha	NA	NA	NA	NA	NA	NA	NA	

CG7911	NA	NA	NA	NA	NA	NA	NA	
CG6255	NA	NA	NA	NA	NA	NA	NA	
eIF-1A	NA	NA	NA	NA	NA	NA	NA	
Bin1	NA	NA	NA	NA	NA	NA	NA	
RpS5b	NA	NA	NA	NA	NA	NA	NA	
CG9684	NA	NA	NA	NA	NA	NA	NA	
eIF4AIII	NA	NA	NA	NA	NA	NA	NA	
CG13084	NA	NA	NA	NA	NA	NA	NA	
CG10341	NA	NA	NA	NA	NA	NA	NA	
ncm	NA	NA	NA	NA	NA	NA	NA	
CG17768	NA	NA	NA	NA	NA	NA	NA	
CG13096	NA	NA	NA	NA	NA	NA	NA	
CG10418	NA	NA	NA	NA	NA	NA	NA	
SRm160	NA	NA	NA	NA	NA	NA	NA	
SmG	NA	NA	NA	NA	NA	NA	NA	
Pkcdelta	NA	NA	NA	NA	NA	NA	NA	

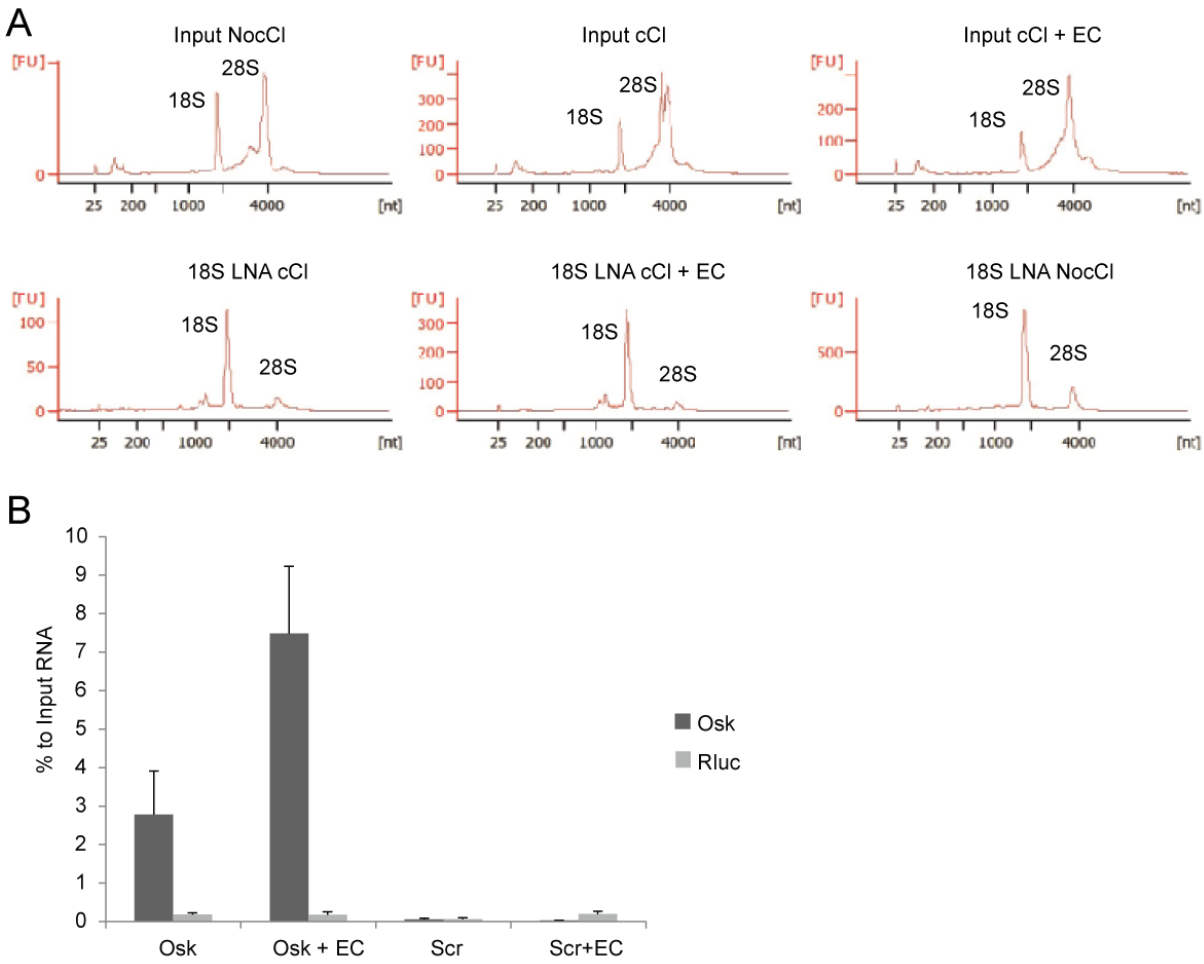

Figure S 7.3; RNA analysis after specific RNP capture with and without ethylenecarbonate (EC) inclusion during hybridization. (A) After specific RNP capture of 18S rRNA RNA was analyzed using a 6000 Pico bioanalyzer. The 2 peaks from the Input samples come from 18S (runs at 2000 nt) and 28S (runs at 4000 nt). The y-axes display the fluorescence intensity coming from a dye labeling the RNA quantitatively. (B) After specific RNP capture of oskar mRNA (Osk) in Drosophila embryo extract, RNA was analyzed by qPCR; Levels of pulled out oskar mRNA and control RLuc RNA were quantified in relation to the Input in %. Error bars display standard deviation coming from 3 replicates.

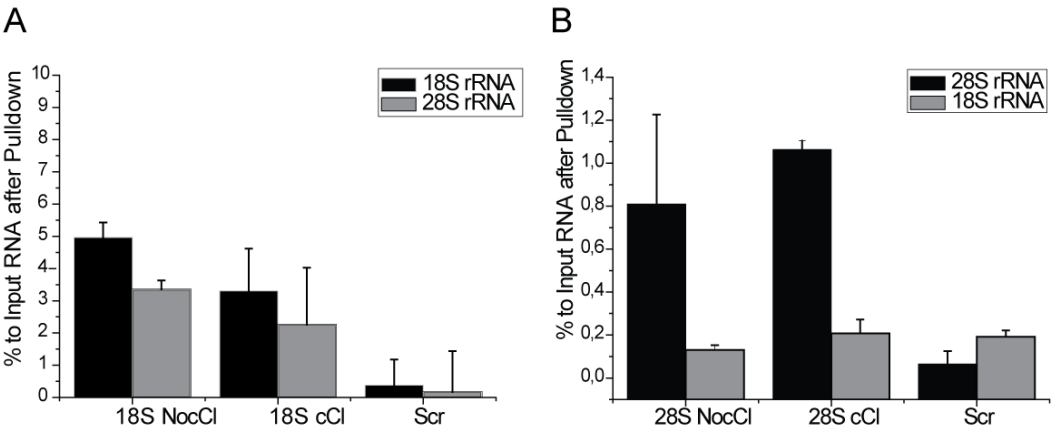

Figure S 7.4; RNA analysis after ribosomal specific RNP capture. (A) After specific RNP capture of 18S rRNA, RNA was analyzed by qPCR; Levels of pulled out18S rRNA and28S rRNA were quantified in relation to the Input in %. Error bars display standard deviation coming from 3 replicates. (B) After specific RNP capture of 28S rRNA, RNA was analyzed by qPCR; Levels of pulled out18S rRNA and28S rRNA were quantified in relation to the Input in %. Error bars display standard deviation coming from 3 replicates.

Table S 7.2; Proteins enriched from 18S rRNA after specific RNP capture. Mass spectrometry measurements were performed by the EMBL proteomics core facility; Bernd Fischer did the semi-quantitative analysis of the three replicates. The amount of enrichments (n) in crosslinked sample (cCl) was compared to the non crosslinked (NocCl) and scrambled control (Scr). The estimated false discovery rate was calculated with following formula: (0+1)/n; being 6.2% (enriched over NocCl), being 7.1% (enriched over Scr) for the listed proteins;

Symbol	nCCL	nNoCL	nScr	sigNoCL	sigScr
YBX3	2	0	0	*	*
ELAVL1	2	0	0	*	*
EIF3J	2	0	0	*	*
NCL	3	0	1	*	*
RPS15	3	0	1	*	*
RPL22	3	0	0	*	*
HNRNPA1	2	0	0	*	*
LRPPRC	2	0	0	*	*
RPS24	3	0	0	*	*
SERBP1	3	0	1	*	*
RPL8	3	0	0	*	*
RPSA	2	0	0	*	*
HNRNPH1	2	0	0	*	*
HNRNPF	3	0	0	*	*
YBX1	3	0	3	*	
RPS2	2	0	2	*	
RPL31	0	0	0		
ZFYVE26	0	0	0		
P2RY10	1	1	1		
TNPO1	0	0	0		
XRN2	0	0	0		
HNRNPC	0	0	0		
HSP90AB1	0	0	0		
HNRNPH3	0	0	0		
HNRNPM	0	0	0		
HCFC1R1	0	0	0		
CASP14	0	0	0		
SSBP1	0	0	1		
COL1A1	0	0	0		
CPE	0	1	0		
FOLR3	0	0	0		
ATP5B	0	0	0		
ARHGDIB	1	1	1		
RPL24	1	0	0		
TAF1B	1	0	0		
SFPQ	0	0	0		
KMT2A	1	1	1		
FKBP15	1	1	1		
CSTA	1	0	0		
RPL5	0	0	0		
HNRNPA2B1	1	0	0		
H2BFWT	0	0	0		
ATG14	1	1	1		

ZSCAN5A	0	0	0		
DSG1	0	0	0		
HNRNPD	1	0	0		
BTF3	0	0	0		
NONO	0	0	0		
RPL7	0	0	0		
MKI67	0	0	0		
OLAH	0	0	0		
HNRNPDL	0	0	0		
TMEM87B	1	1	1		
MCOLN2	0	0	0		
MAGED4	0	0	0		
AFAP1L1	1	1	1		
HIST1H4H	0	0	0		
ATP6V0D1	2	2	2		
TRAPPC10	0	0	0		
LCN1	2	1	1		
SRSF2	0	0	0		
DCD	0	0	0		
FMNL3	0	0	1		
CCDC78	1	1	1		
RPL29	0	0	0		
CFAP221	0	0	0		
SPTA1	0	0	0		
GTPBP8	0	0	0		
HNRNPK	0	0	0		
C19orf33	1	0	0		
PLIN4	1	0	0		
KRT1	3	3	3		
SRRM2	1	0	0		
PA2G4	1	0	0		
JUP	0	0	0		
RPL4	0	0	0		
TUBB6	1	0	0		
COX5A	0	0	0		
CCDC89	0	0	0		
RESP18	0	0	0		
NAA38	0	1	0		
ABAT	1	1	1		
ARAP1	0	0	0		
KCNJ11	1	1	1		
NCOR2	0	0	0		
HNRNPAB	0	0	0		
TGM2	1	1	1		
CFAP44	0	0	0		
PPME1	0	0	0		
C21orf140	0	0	0		

Table S 7.3 Proteins enriched from 28S rRNA after specific RNP capture. Mass spectrometry measurements were performed by the EMBL proteomics core facility; Bernd Fischer did the semi-quantitative analysis of the three replicates. The amount of enrichments (n) in crosslinked sample (cCl) was compared to the non crosslinked (NocCl) and scrambled control (Scr). The estimated false discovery rate was calculated with following formula: (0+1)/n; being 4.3% (enriched over NocCl), being 5.3% (enriched over Scr) for the listed proteins;

Symbol	nCCL	nNoCL	nScr	sigNoCL	sigScr
YBX3	3	0	0	*	*
ELAVL1	2	0	0	*	*
RPL31	2	0	0	*	*
HNRNPC	2	0	0	*	*
HNRNPM	2	0	0	*	*
NCL	3	0	0	*	*
RPL22	3	0	0	*	*
RPL5	3	0	0	*	*
HNRNPA2B1	2	0	0	*	*
DSG1	2	0	0	*	*
HNRNPA1	3	0	0	*	*
HNRNPD	2	0	0	*	*
BTF3	2	0	0	*	*
RPL29	2	0	0	*	*
PLIN4	2	0	0	*	*
HNRNPH1	2	0	0	*	*
HNRNPF	3	0	0	*	*
PA2G4	3	0	0	*	*
RPL4	2	0	0	*	*
YBX1	3	0	2	*	
RPS2	3	0	2	*	
SERBP1	3	1	2	*	
RPL8	3	1	2	*	
ZFYVE26	1	1	1		
P2RY10	0	0	0		
TNPO1	0	0	0		
XRN2	0	0	0		
HSP90AB1	1	0	0		
HNRNPH3	1	0	0		
HCFC1R1	0	0	0		
EIF3J	0	0	0		
CASP14	2	2	2		
SSBP1	0	0	0		
COL1A1	1	1	1		
CPE	0	0	0		
FOLR3	0	0	0		
ATP5B	1	1	1		
ARHGDIB	3	3	3		
RPL24	0	0	0		
RPS15	0	0	1		
TAF1B	0	0	0		
SFPQ	1	0	0		
KMT2A	0	0	0		

FKBP15	0	0	0		
CSTA	2	1	1		
H2BFWT	0	0	0		
ATG14	0	0	0		
ZSCAN5A	3	3	3		
LRPPRC	1	0	0		
RPS24	0	0	0		
NONO	1	0	0		
RPL7	1	0	0		
MKI67	0	0	0		
OLAH	2	2	2		
HNRNPDL	1	0	0		
TMEM87B	0	0	0		
MCOLN2	0	0	0		
MAGED4	0	0	0		
AFAP1L1	0	0	0		
HIST1H4H	1	0	0		
ATP6V0D1	0	0	0		
TRAPPC10	3	3	3		
LCN1	2	1	1		
SRSF2	1	0	0		
DCD	1	1	1		
FMNL3	0	0	0		
CCDC78	3	3	3		
CFAP221	0	1	0		
SPTA1	2	2	2		
GTPBP8	0	0	0		
HNRNPK	0	0	0		
C19orf33	0	0	0		
KRT1	3	3	3		
SRRM2	1	0	0		
RPSA	0	0	0		
JUP	1	0	1		
TUBB6	0	0	0		
COX5A	0	0	0		
CCDC89	0	0	0		
RESP18	3	3	3		
NAA38	0	0	0		
ABAT	1	1	1		
ARAP1	2	2	2		
KCNJ11	1	1	1		
NCOR2	1	1	0		
HNRNPAB	1	0	0		
TGM2	0	0	0		
CFAP44	0	0	0		
PPME1	0	0	0		
C21orf140	0	0	0		

8 Bibliography

Abaza, I., Coll, O., Patalano, S., and Gebauer, F. (2006). Drosophila UNR is required for translational repression of male-specific lethal 2 mRNA during regulation of X-chromosome dosage compensation. Genes Dev. *20*, 380–389.

Amunts, A., Brown, A., Toots, J., Scheres, S.H.W., and Ramakrishnan, V. (2015). Ribosome. The structure of the human mitochondrial ribosome. Science *348*, 95–98.

Andoh, T., Oshiro, Y., Hayashi, S., Takeo, H., and Tani, T. (2006). Visual screening for localized RNAs in yeast revealed novel RNAs at the bud-tip. Biochem. Biophys. Res. Commun. *351*, 999–1004.

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.

Anosova, I., Kowal, E.A., Dunn, M.R., Chaput, J.C., Van Horn, W.D., and Egli, M. (2016). The structural diversity of artificial genetic polymers. Nucleic Acids Res. *44*, 1007–1021.

Armache, J.-P., Jarasch, A., Anger, A.M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S., et al. (2010). Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-Å resolution. PNAS *107*, 19748–19753.

Aronov, S., Gelin-Licht, R., Zipor, G., Haim, L., Safran, E., and Gerst, J.E. (2007). mRNAs encoding polarity and exocytosis factors are cotransported with the cortical endoplasmic reticulum to the incipient bud in Saccharomyces cerevisiae. Mol. Cell. Biol. *27*, 3441–3455.

Bachler, M., Schroeder, R., and von Ahsen, U. (1999). StreptoTag: a novel method for the isolation of RNA-binding proteins. RNA *5*, 1509–1516.

Badertscher, L., Wild, T., Montellese, C., Alexander, L.T., Bammert, L., Sarazova, M., Stebler, M., Csucs, G., Mayer, T.U., Zamboni, N., et al. (2015). Genome-wide RNAi Screening Identifies Protein Modules Required for 40S Subunit Synthesis in Human Cells. Cell Reports *13*, 2879–2891.

Balagopal, V., and Parker, R. (2011). Stm1 modulates translation after 80S formation in Saccharomyces cerevisiae. RNA *17*, 835–842.

Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., et al. (2012). The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. Molecular Cell *46*, 674–690.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science *289*, 905–920.

Barraud, P., and Allain, F.H.-T. (2012). ADAR Proteins: Double-stranded RNA and Z-DNA Binding Domains. Curr Top Microbiol Immunol *353*, 35–60.

Bashaw, G.J., and Baker, B.S. (1995). The msl-2 dosage compensation gene of Drosophila encodes a putative DNA-binding protein whose expression is sex specifically regulated by Sex-lethal. Development *121*, 3245–3258.

Beatrix, B., Sakai, H., and Wiedmann, M. (2000). The alpha and beta subunit of the nascent polypeptide-associated complex have distinct functions. J. Biol. Chem. *275*, 37838–37845.

Beckmann, B.M., Horos, R., Fischer, B., Castello, A., Eichelbaum, K., Alleaume, A.-M., Schwarzl, T., Curk, T., Foehr, S., Huber, W., et al. (2015). The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. Nat Commun *6*, 10127.

Beckmann, B.M., Castello, A., and Medenbach, J. (2016). The expanding universe of ribonucleoproteins: of novel RNA-binding proteins and unconventional interactions. Pflugers Arch - Eur J Physiol *468*, 1029–1040.

Beckmann, K., Grskovic, M., Gebauer, F., and Hentze, M.W. (2005). A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in Drosophila. Cell *122*, 529–540.

Beckmann, R., Spahn, C.M.T., Eswar, N., Helmers, J., Penczek, P.A., Sali, A., Frank, J., and Blobel, G. (2001). Architecture of the Protein-Conducting Channel Associated with the Translating 80S Ribosome. Cell *107*, 361–372.

Bell, L.R., Horabin, J.I., Schedl, P., and Cline, T.W. (1991). Positive autoregulation of sex-lethal by alternative splicing maintains the female determined state in Drosophila. Cell *65*, 229–239.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) Vol. 57, No. 1 (1995), pp. 289-300

Ben-Shem, A., Loubresse, N.G. de, Melnikov, S., Jenner, L., Yusupova, G., and Yusupov, M. (2011). The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution. Science *334*, 1524–1529.

Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M., and Nüsslein-Volhard, C. (1988). The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. EMBO J *7*, 1749–1756.

Binder, R., Horowitz, J.A., Basilion, J.P., Koeller, D.M., Klausner, R.D., and Harford, J.B. (1994). Evidence that the pathway of transferrin receptor mRNA degradation involves an endonucleolytic cleavage within the 3' UTR and does not involve poly(A) tail shortening. EMBO J. *13*, 1969–1980.

Blanchette, M., Green, R.E., MacArthur, S., Brooks, A.N., Brenner, S.E., Eisen, M.B., and Rio, D.C. (2009). Genome-wide Analysis of Alternative Pre-mRNA Splicing and RNA-Binding Specificities of the Drosophila hnRNP A/B Family Members. Molecular Cell *33*, 438–449.

Boehringer, D., Thermann, R., Ostareck-Lederer, A., Lewis, J.D., and Stark, H. (2005). Structure of the hepatitis C virus IRES bound to the human 80S ribosome: remodeling of the HCV IRES. Structure *13*, 1695–1706.

Boggs, R.T., Gregor, P., Idriss, S., Belote, J.M., and McKeown, M. (1987). Regulation of sexual differentiation in D. melanogaster via alternative splicing of RNA from the transformer gene. Cell *50*, 739–747.

Bolognani, F., Contente-Cuomo, T., and Perrone-Bizzozero, N.I. (2010). Novel recognition motifs and biological functions of the RNA-binding protein HuD revealed by genome-wide identification of its targets. Nucl. Acids Res. *38*, 117–130.

Bonneau, A.M., and Sonenberg, N. (1987). Proteolysis of the p220 component of the cap-binding protein complex is not sufficient for complete inhibition of host cell protein synthesis after poliovirus infection. J. Virol. *61*, 986–991.

Boon, W.C., Petkovic-Duran, K., Zhu, Y., Manasseh, R., Horne, M.K., and Aumann, T.D. (2011). Increasing cDNA yields from single-cell quantities of mRNA in standard laboratory reverse transcriptase reactions using acoustic microstreaming. J Vis Exp e3144.

Bowman, L.H., Goldman, W.E., Goldberg, G.I., Hebert, M.B., and Schlessinger, D. (1983). Location of the initial cleavage sites in mouse pre-rRNA. Mol. Cell. Biol. *3*, 1501–1510.

Bunnik, E.M., Batugedara, G., Saraf, A., Prudhomme, J., Florens, L., and Le Roch, K.G. (2016). The mRNA-bound proteome of the human malaria parasite Plasmodium falciparum. Genome Biol. *17*, 147.

Burger, K., Mühl, B., Kellner, M., Rohrmoser, M., Gruber-Eber, A., Windhager, L., Friedel, C.C., Dölken, L., and Eick, D. (2013). 4-thiouridine inhibits rRNA synthesis and causes a nucleolar stress response. RNA Biol 10, 1623–1630.

Casey, J.L., Hentze, M.W., Koeller, D.M., Caughman, S.W., Rouault, T.A., Klausner, R.D., and Harford, J.B. (1988). Iron-responsive elements: regulatory RNA sequences that control mRNA levels and translation. Science *240*, 924–928.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., et al. (2012). Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. Cell *149*, 1393–1406.

Castello, A., Fischer, B., Hentze, M.W., and Preiss, T. (2013a). RNA-binding proteins in Mendelian disease. Trends in Genetics *29*, 318–327.

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L.M., Krijgsveld, J., and Hentze, M.W. (2013b). System-wide identification of RNA-binding proteins by interactome capture. Nat. Protocols *8*, 491–500.

Castello, A., Hentze, M.W., and Preiss, T. (2015). Metabolic Enzymes Enjoying New Partnerships as RNA-Binding Proteins. Trends Endocrinol Metab *26*, 746–757.

Castello, A., Fischer, B., Frese, C.K., Horos, R., Alleaume, A.-M., Foehr, S., Curk, T., Krijgsveld, J., and Hentze, M.W. (2016). Comprehensive Identification of RNA-Binding Domains in Human Cells. Molecular Cell *0*.

Cech, T.R. (2000). The Ribosome Is a Ribozyme. Science *289*, 878–879.

Chandramouli, P., Topf, M., Ménétret, J.-F., Eswar, N., Cannone, J.J., Gutell, R.R., Sali, A., and Akey, C.W. (2008). Structure of the mammalian 80S ribosome at 8.7 A resolution. Structure *16*, 535–548.

Chang, C.-H., Curtis, J.D., Maggi, L.B., Faubert, B., Villarino, A.V., O'Sullivan, D., Huang, S.C.-C., van der Windt, G.J.W., Blagih, J., Qiu, J., et al. (2013). Posttranscriptional control of T cell effector function by aerobic glycolysis. Cell *153*, 1239–1251.

Chen, L.-L. (2016). The biogenesis and emerging roles of circular RNAs. Nat Rev Mol Cell Biol *17*, 205–211.

Chen, Y., and Varani, G. (2013). Engineering RNA-binding proteins for biology. FEBS J *280*, 3734–3754.

Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E., and Chang, H.Y. (2015). Systematic Discovery of Xist RNA Binding Proteins. Cell *161*, 404–416.

Chujo, T., Ohira, T., Sakaguchi, Y., Goshima, N., Nomura, N., Nagao, A., and Suzuki, T. (2012). LRPPRC/SLIRP suppresses PNPase-mediated mRNA decay and promotes polyadenylation in human mitochondria. Nucl. Acids Res. *40*, 8033–8047.

Claude, A. (1944). THE CONSTITUTION OF MITOCHONDRIA AND MICROSOMES, AND THE DISTRIBUTION OF NUCLEIC ACID IN THE CYTOPLASM OF A LEUKEMIC CELL. J. Exp. Med. *80*, 19–29.

Cléry, A., and Allain, F.H.-T. (2013). FROM STRUCTURE TO FUNCTION OF RNA BINDING DOMAINS (Landes Bioscience).

Cline, T.W. (1984). Autoregulatory functioning of a Drosophila gene product that establish es and maintains the sexually determined state. Genetics *107*, 231–277.

Cline, T.W., and Meyer, B.J. (1996). Vive la différence: males vs females in flies vs worms. Annu. Rev. Genet. *30*, 637–702.

Conn, S.J., Pillman, K.A., Toubia, J., Conn, V.M., Salmanidis, M., Phillips, C.A., Roslan, S., Schreiber, A.W., Gregory, P.A., and Goodall, G.J. (2015). The RNA binding protein quaking regulates formation of circRNAs. Cell *160*, 1125–1134.

Conrad, T., and Akhtar, A. (2012). Dosage compensation in Drosophila melanogaster: epigenetic fine-tuning of chromosome-wide transcription. Nat Rev Genet *13*, 123–134.

Crick, F.H. (1958). On protein synthesis. Symp. Soc. Exp. Biol. *12*, 138–163.

Darnell, R.B. (2010). HITS-CLIP: panoramic views of protein–RNA regulation in living cells. WIREs RNA *1*, 266–286.

Das, R., Zhou, Z., and Reed, R. (2000). Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. Mol. Cell *5*, 779–787.

Davidson, N.O., and Shelness, G.S. (2000). APOLIPOPROTEIN B: mRNA Editing, Lipoprotein Assembly, and Presecretory Degradation. Annual Review of Nutrition *20*, 169–193.

Davis, J.E., and Sinsheimer, R.L. (1963). The replication of bacteriophage MS2. 1. Transfer of parental nucleic acid to progeny phage. J. Mol. Biol. *6*, 203–207.

de Chassey, B., Meyniel-Schicklin, L., Vonderscher, J., André, P., and Lotteau, V. (2014). Virus-host interactomics: new insights and opportunities for antiviral drug discovery. Genome Med *6*.

Delaney, S.J., Rich, D.P., Thomson, S.A., Hargrave, M.R., Lovelock, P.K., Welsh, M.J., and Wainwright, B.J. (1993). Cystic fibrosis transmembrane conductance regulator splice variants are not conserved and fail to produce chloride channels. Nat Genet *4*, 426–430.

Dever, T.E., and Green, R. (2012). The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. Cold Spring Harb Perspect Biol *4*.

Dez, C., Houseley, J., and Tollervey, D. (2006). Surveillance of nuclear-restricted pre-ribosomes within a subnucleolar region of Saccharomyces cerevisiae. EMBO J. *25*, 1534–1546.

Dreyfuss, G., Kim, V.N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. Nat Rev Mol Cell Biol *3*, 195–205.

Duncan, K., Grskovic, M., Strein, C., Beckmann, K., Niggeweg, R., Abaza, I., Gebauer, F., Wilm, M., and Hentze, M.W. (2006). Sex-lethal imparts a sex-specific function to UNR by recruiting it to the msl-2 mRNA 3' UTR: translational repression for dosage compensation. Genes Dev. *20*, 368–379.

Duncan, K.E., Strein, C., and Hentze, M.W. (2009). The SXL-UNR Corepressor Complex Uses a PABP-Mediated Mechanism to Inhibit Ribosome Recruitment to msl-2 mRNA. Molecular Cell *36*, 571–582.

Ecker, J.R., and Davis, R.W. (1986). Inhibition of gene expression in plant cells by expression of antisense RNA. Proc Natl Acad Sci U S A *83*, 5372–5376.

Elmen, J, Wahlestedt, E. , Liang, Z., Sorensen, A. M., Orum,H., Koch, T. (2014). Short Interfering RNA (siRNA) Analogues. US20140235844 A1

Ephrussi, A., Dickinson, L.K., and Lehmann, R. (1991). Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. Cell *66*, 37–50.

Fedoriw, A.M., Starmer, J., Yee, D., and Magnuson, T. (2012). Nucleolar association and transcriptional inhibition through 5S rDNA in mammals. PLoS Genet. *8*, e1002468.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806–811.

Fischer, U., Schäuble, N., Schütz, S., Altvater, M., Chang, Y., Faza, M.B., and Panse, V.G. (2015). A non-canonical mechanism for Crm1-export cargo complex assembly. Elife *4*.

Fraser, C.S., Lee, J.Y., Mayeur, G.L., Bushell, M., Doudna, J.A., and Hershey, J.W.B. (2004). The j-Subunit of Human Translation Initiation Factor eIF3 Is Required for the Stable Binding of eIF3 and Its Subcomplexes to 40 S Ribosomal Subunits in Vitro. J. Biol. Chem. *279*, 8946–8956.

Freed, E.F., Bleichert, F., Dutca, L.M., and Baserga, S.J. (2010). When ribosomes go bad: diseases of ribosome biogenesis. Mol Biosyst *6*, 481–493.

Fukao, A., Mishima, Y., Takizawa, N., Oka, S., Imataka, H., Pelletier, J., Sonenberg, N., Thoma, C., and Fujiwara, T. (2014). MicroRNAs trigger dissociation of eIF4AI and eIF4AII from target mRNAs in humans. Mol. Cell *56*, 79–89.

Fukaya, T., Iwakawa, H.-O., and Tomari, Y. (2014). MicroRNAs block assembly of eIF4F translation initiation complex in Drosophila. Mol. Cell *56*, 67–78.

Gardiner, A.S., Twiss, J.L., and Perrone-Bizzozero, N.I. (2015). Competing Interactions of RNA-Binding Proteins, MicroRNAs, and Their Targets Control Neuronal Development and Function. Biomolecules *5*, 2903–2918.

Gebauer, F., Merendino, L., Hentze, M.W., and Valcárcel, J. (1998). The Drosophila splicing regulator sex-lethal directly inhibits translation of male-specific-lethal 2 mRNA. RNA *4*, 142–150.

Gebauer, F., Corona, D.F., Preiss, T., Becker, P.B., and Hentze, M.W. (1999). Translational control of dosage compensation in Drosophila by Sex-lethal: cooperative silencing via the 5’ and 3’ UTRs of msl-2 mRNA is independent of the poly(A) tail. EMBO J. *18*, 6146–6154.

Gebauer, F., Grskovic, M., and Hentze, M.W. (2003). Drosophila Sex-Lethal Inhibits the Stable Association of the 40S Ribosomal Subunit with msl-2 mRNA. Molecular Cell *11*, 1397–1404.

Gerbi SA (1996) Expansion segments: regions of variable size that interrupt the universal core secondary structure of ribosomal RNA. In *Ribosomal RNA: Structure, Evolution, Processing and Function in Protein Synthesis*, eds: R.A. Zimmermann and A.E. Dahlberg. Telford - CRC Press, Boca Raton, FL. pp. 71-87.

Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737–741.

Gilbert, C., and Svejstrup, J.Q. (2006). RNA immunoprecipitation for determining RNA-protein associations in vivo. Curr Protoc Mol Biol *Chapter 27*, Unit 27.4.

Gilbert, W.V., Zhou, K., Butler, T.K., and Doudna, J.A. (2007). Cap-independent translation is required for starvation-induced differentiation in yeast. Science *317*, 1224–1227.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett *582*, 1977–1986.

Gonsalvez, G.B., and Long, R.M. (2012). Spatial regulation of translation through RNA localization. F1000Prime Rep *4*.

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

Gray, N.K., and Hentze, M.W. (1994). Regulation of protein synthesis by mRNA structure. Mol. Biol. Rep. *19*, 195–200.

Greber, B.J., Boehringer, D., Montellese, C., and Ban, N. (2012a). Cryo-EM structures of Arx1 and maturation factors Rei1 and Jjj1 bound to the 60S ribosomal subunit. Nat Struct Mol Biol *19*, 1228–1233.

Greber, B.J., Gerhardy, S., Leitner, A., Leibundgut, M., Salem, M., Boehringer, D., Leulliot, N., Aebersold, R., Panse, V.G., and Ban, N. (2016). Insertion of the Biogenesis Factor Rei1 Probes the Ribosomal Tunnel during 60S Maturation. Cell *164*, 91–102.

Green, N.M. (1975). Avidin. In Advances in Protein Chemistry, J.T.E. and F.M.R. C.B. Anfinsen, ed. (Academic Press), pp. 85–133.

Grskovic, M., Hentze, M.W., and Gebauer, F. (2003). A co-repressor assembly nucleated by Sex-lethal in the 3′UTR mediates translational control of Drosophila msl-2 mRNA. EMBO J *22*, 5571–5581.

Hadjiolova, K.V., Nicoloso, M., Mazan, S., Hadjiolov, A.A., and Bachellerie, J.P. (1993). Alternative pre-rRNA processing pathways in human cells and their alteration by cycloheximide inhibition of protein synthesis. Eur. J. Biochem. *212*, 211–215.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano Jr., M., Jungkamp, A.-C., Munschauer, M., et al. (2010). Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. Cell *141*, 129–141.

Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. Nature *495*, 384–388.

Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H.Y., Grassucci, R.A., Pestova, T.V., Hellen, C.U.T., and Frank, J. (2013). Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. Nature *503*, 539–543.

Haussecker, D., Huang, Y., Lau, A., Parameswaran, P., Fire, A.Z., and Kay, M.A. (2010). Human tRNA-derived small RNAs in the global regulation of RNA silencing. RNA *16*, 673–695.

Haynes, S.R., Johnson, D., Raychaudhuri, G., and Beyer, A.L. (1991). The Drosophila Hrb87F gene encodes a new member of the A and B hnRNP protein group. Nucleic Acids Res *19*, 25–31.

Heid, H., Rickelt, S., Zimbelmann, R., Winter, S., Schumacher, H., Dörflinger, Y., Kuhn, C., and Franke, W.W. (2014). On the formation of lipid droplets in human adipocytes: the organization of the perilipin-vimentin cortex. PLoS ONE *9*, e90386.

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. WIREs RNA *6*, 225–242.

Hentze, M.W., and Preiss, T. (2010). The REM phase of gene regulation. Trends Biochem. Sci. *35*, 423–426.

Hentze, M.W., and Preiss, T. (2013). Circular RNAs: splicing's enigma variations. The EMBO Journal *32*, 923–925.

Hentze, M.W., Caughman, S.W., Rouault, T.A., Barriocanal, J.G., Dancis, A., Harford, J.B., and Klausner, R.D. (1987). Identification of the iron-responsive element for the translational regulation of human ferritin mRNA. Science *238*, 1570–1573.

Herdy, B., Karonitsch, T., Vladimer, G.I., Tan, C.S.H., Stukalov, A., Trefzer, C., Bigenzahn, J.W., Theil, T., Holinka, J., Kiener, H.P., et al. (2015). The RNA-binding protein HuR/ELAVL1 regulates IFN-β mRNA abundance and the type I IFN response. Eur. J. Immunol. 45, 1500–1511.

Hershey, J.W.B., Sonenberg, N., and Mathews, M.B. (2012). Principles of Translational Control: An Overview. Cold Spring Harb Perspect Biol *4*.

Hoagland, M.B., Stephenson, M.L., Scott, J.F., Hecht, L.I., and Zamecnik, P.C. (1958). A soluble ribonucleic acid intermediate in protein synthesis. J. Biol. Chem. *231*, 241–257.

Hogg, J.R., and Collins, K. (2007). RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. RNA *13*, 868–880.

Holcik, M., and Sonenberg, N. (2005). Translational control in stress and apoptosis. Nat. Rev. Mol. Cell Biol. *6*, 318–327.

Hughes, C.S., Foehr, S., Garfield, D.A., Furlong, E.E., Steinmetz, L.M., and Krijgsveld, J. (2014). Ultrasensitive proteome analysis using paramagnetic bead technology. Molecular Systems Biology *10*, 757.

Hung, N.-J., Lo, K.-Y., Patel, S.S., Helmke, K., and Johnson, A.W. (2008). Arx1 Is a Nuclear Export Receptor for the 60S Ribosomal Subunit in Yeast. Mol Biol Cell *19*, 735–744.

Iizuka, N., Najita, L., Franzusoff, A., and Sarnow, P. (1994). Cap-dependent and cap-independent translation by internal initiation of mRNAs in cell extracts prepared from Saccharomyces cerevisiae. Mol. Cell. Biol. *14*, 7322–7330.

Inoue, K., Hoshijima, K., Sakamoto, H., and Shimura, Y. (1990). Binding of the Drosophila sex-lethal gene product to the alternative splice site of transformer primary transcript. Nature *344*, 461–463.

Irurzun, A., Sánchez-Palomino, S., Novoa, I., and Carrasco, L. (1995). Monensin and nigericin prevent the inhibition of host translation by poliovirus, without affecting p220 cleavage. J. Virol. *69*, 7453–7460.

Irwin, B., Heck, J.D., and Hatfield, G.W. (1995). Codon Pair Utilization Biases Influence Translational Elongation Step Times. J. Biol. Chem. *270*, 22801–22806.

Iwasaki, Y.W., Siomi, M.C., and Siomi, H. (2015). PIWI-Interacting RNA: Its Biogenesis and Functions. Annu. Rev. Biochem. *84*, 405–433.

Jacobsen, N., Nielsen, P.S., Jeffares, D.C., Eriksen, J., Ohlsson, H., Arctander, P., and Kauppinen, S. (2004). Direct isolation of poly(A)+ RNA from 4 M guanidine thiocyanate-lysed cell extracts using locked nucleic acid-oligo(T) capture. Nucleic Acids Res. *32*, e64.

Jang, S.K., Kräusslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C., and Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J. Virol. *62*, 2636–2643.

Järvelin, A.I., Noerenberg, M., Davis, I., and Castello, A. (2016). The new (dis)order in RNA regulation. Cell Commun Signal *14*.

Jenner, L., Demeshkina, N., Yusupova, G., and Yusupov, M. (2010). Structural rearrangements of the ribosome at the tRNA proofreading step. Nat Struct Mol Biol *17*, 1072–1078.

Jonas, S., and Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. Nat Rev Genet *16*, 421–433.

Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermüller, J., Hofacker, I.L., et al. (2007). RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. Science *316*, 1484–1488.

Kauppinen, S., Vester, B., and Wengel, J. (2005). Locked nucleic acid (LNA): High affinity targeting of RNA for diagnostics and therapeutics. Drug Discovery Today: Technologies *2*, 287–290.

Keyes, L.N., Cline, T.W., and Schedl, P. (1992). The primary sex determination signal of Drosophila acts at the level of transcription. Cell *68*, 933–943.

Khatter, H., Myasnikov, A.G., Natchiar, S.K., and Klaholz, B.P. (2015). Structure of the human 80S ribosome. Nature *520*, 640–645.

Kieft, J.S. (2008). Viral IRES RNA structures and ribosome interactions. Trends Biochem Sci *33*, 274–283.

Klann, E., and Dever, T.E. (2004). Biochemical mechanisms for translational regulation in synaptic plasticity. Nat Rev Neurosci *5*, 931–942.

Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S., and Ban, N. (2011). Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. Science *334*, 941–948.

König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol *17*, 909–915.

Krause, H. M. & Simmonds, A. J. (2004). Trap-tagging: a novel method for the identification and purification of rna-protein complexes. WO2004033718 A2

Kuzuoğlu-Öztürk, D., Bhandari, D., Huntzinger, E., Fauser, M., Helms, S., and Izaurralde, E. (2016). miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning. EMBO J. *35*, 1186–1203.

Kwon, S.C., Yi, H., Eichelbaum, K., Föhr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013). The RNA-binding protein repertoire of embryonic stem cells. Nat Struct Mol Biol *20*, 1122–1130.

Lafontaine, D.L.J. (2015). Noncoding RNAs in eukaryotic ribosome biogenesis and function. Nat Struct Mol Biol *22*, 11–19.

Lamond, A.I., Konarska, M.M., Grabowski, P.J., and Sharp, P.A. (1988). Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. Proc Natl Acad Sci U S A *85*, 411–415.

Lamond, A.I., Sproat, B., Ryder, U., and Hamm, J. (1989). Probing the structure and function of U2 snRNP with antisense oligonucleotides made of 2′-OMe RNA. Cell *58*, 383–390.

Lasko, P. (2012). mRNA Localization and Translational Control in Drosophila Oogenesis. Cold Spring Harb Perspect Biol *4*.

Lau, C., Bachorik, J.L., and Dreyfuss, G. (2009). Gemin5-snRNA interaction reveals an RNA binding function for WD repeat domains. Nat. Struct. Mol. Biol. *16*, 486–491.

Le, S.Y., and Maizel, J.V. (1997). A common RNA structural motif involved in the internal initiation of translation of cellular mRNAs. Nucleic Acids Res. *25*, 362–369.

Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., and Krause, H.M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. Cell *131*, 174–187.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843–854.

Lenarcic, E.M., Landry, D.M., Greco, T.M., Cristea, I.M., and Thompson, S.R. (2013). Thiouracil Cross-Linking Mass Spectrometry: a Cell-Based Method To Identify Host Factors Involved in Viral Amplification. J. Virol. *87*, 8697–8712.

Leonov, A.A., Sergiev, P.V., Bogdanov, A.A., Brimacombe, R., and Dontsova, O.A. (2003). Affinity purification of ribosomes with a lethal G2655C mutation in 23 S rRNA that affects the translocation. J. Biol. Chem. *278*, 25664–25670.

Leppek, K., and Stoecklin, G. (2014). An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. Nucl. Acids Res. *42*, e13–e13.

Lerner, M.R., and Steitz, J.A. (1979). Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. Proc Natl Acad Sci U S A *76*, 5495–5499.

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.

Liao, Y., Castello, A., Fischer, B., Leicht, S., Föehr, S., Frese, C.K., Ragan, C., Kurscheid, S., Pagler, E., Yang, H., et al. (2016). The Cardiomyocyte RNA-Binding Proteome: Links to Intermediary Metabolism and Heart Disease. Cell Rep.

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., et al. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature *456*, 464–469.

Liepelt, A., Naarmann-de Vries, I.S., Simons, N., Eichelbaum, K., Foehr, S., Archer, S.K., Castello, A., Usadel, B., Krijgsveld, J., Preiss, T., et al. (2016). Identification of RNA-binding proteins in macrophages by interactome capture. Mol. Cell Proteomics.

Lingner, J., and Cech, T.R. (1996). Purification of telomerase from Euplotes aediculatus: requirement of a primer 3’ overhang. Proc Natl Acad Sci U S A *93*, 10712–10717.

Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature *518*, 560–564.

Loeb, T., and Zinder, N.D. (1961). A bacteriophage containing RNA. Proc. Natl. Acad. Sci. U.S.A. *47*, 282–289.

Lönnstedt, I. and Speed, T. (2002). Replicated microarray data. Statistica Sinica 12(1):31-46

Lueong, S., Merce, C., Fischer, B., Hoheisel, J.D., and Erben, E.D. (2016). Gene expression regulatory networks in Trypanosoma brucei: insights into the role of the mRNA-binding proteome. Molecular Microbiology *100*, 457–471.

Lunde, B.M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. Nat Rev Mol Cell Biol *8*, 479–490.

Marondedze, C., Thomas, L., Serrano, N.L., Lilley, K.S., and Gehring, C. (2016). The RNA-binding protein repertoire of Arabidopsis thaliana. Sci Rep *6*, 29766.

Martin, R., Straub, A.U., Doebele, C., and Bohnsack, M.T. (2013). DExD/H-box RNA helicases in ribosome biogenesis. RNA Biol *10*, 4–18.

Matia-González, A.M., Laing, E.E., and Gerber, A.P. (2015). Conserved mRNA-binding proteomes in eukaryotic organisms. Nat. Struct. Mol. Biol. *22*, 1027–1033.

Matoulkova, E., Michalova, E., Vojtesek, B., and Hrstka, R. (2012). The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. RNA Biology *9*, 563–576.

Matsuo, Y., Granneman, S., Thoms, M., Manikas, R.-G., Tollervey, D., and Hurt, E. (2014). Coupled GTPase and remodelling ATPase activities form a checkpoint for ribosome export. Nature *505*, 112–116.

Matthiesen, S.H., and Hansen, C.M. (2012). Fast and Non-Toxic In Situ Hybridization without Blocking of Repetitive Sequences. PLOS ONE *7*, e40675.

McHugh, C.A., Chen, C.-K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature *521*, 232–236.

Medenbach, J., Seiler, M., and Hentze, M.W. (2011). Translational Control via Protein-Regulated Upstream Open Reading Frames. Cell *145*, 902–913.

Meijer, H.A., Kong, Y.W., Lu, W.T., Wilczynska, A., Spriggs, R.V., Robinson, S.W., Godfrey, J.D., Willis, A.E., and Bushell, M. (2013). Translational Repression and eIF4A2 Activity Are Critical for MicroRNA-Mediated Gene Regulation. Science *340*, 82–85.

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.

Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. Nature *495*, 333–338.

Merendino, L., Guth, S., Bilbao, D., Martínez, C., and Valcárcel, J. (1999). Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3′ splice site AG. Nature *402*, 838–841.

Mili, S., and Piñol-Roma, S. (2003). LRP130, a Pentatricopeptide Motif Protein with a Noncanonical RNA-Binding Domain, Is Bound In Vivo to Mitochondrial and Nuclear RNAs. Mol. Cell. Biol. *23*, 4972–4982.

Mitchell, S.F., Jain, S., She, M., and Parker, R. (2013). Global analysis of yeast mRNPs. Nat Struct Mol Biol *20*, 127–133.

Mohr, S., Stryker, J.M., and Lambowitz, A.M. (2002). A DEAD-Box Protein Functions as an ATP-Dependent RNA Chaperone in Group I Intron Splicing. Cell *109*, 769–779.

Moore, M.J. (2005). From Birth to Death: The Complex Lives of Eukaryotic mRNAs. Science *309*, 1514–1518.

Müllner, E.W., and Kühn, L.C. (1988). A stem-loop in the 3′ untranslated region mediates iron-dependent regulation of transferrin receptor mRNA stability in the cytoplasm. Cell *53*, 815–825.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. Plant Cell *2*, 279–289.

Nishikura, K. (2016). A-to-I editing of coding and non-coding RNAs by ADARs. Nat. Rev. Mol. Cell Biol. *17*, 83–96.

O'Brien, T.W. (2003). Properties of human mitochondrial ribosomes. IUBMB Life *55*, 505–513.

O'Connell, M.R., Oakes, B.L., Sternberg, S.H., East-Seletsky, A., Kaplan, M., and Doudna, J.A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature *516*, 263–266.

Ohsugi, M., Larue, L., Schwarz, H., and Kemler, R. (1997). Cell-Junctional and Cytoskeletal Organization in Mouse Blastocysts Lacking E-Cadherin. Developmental Biology *185*, 261–271.

Olson, M.O., and Dundr, M. (2001). Nucleolus: Structure and Function. In eLS, (John Wiley & Sons, Ltd), ISBN: 978-0-470-01590-2

Orlando, V. (2000). Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. Trends in Biochemical Sciences *25*, 99–104.

Palade, G.E. (1955). STUDIES ON THE ENDOPLASMIC RETICULUM. J Biophys Biochem Cytol *1*, 567–582.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet *40*, 1413–1415.

Paranchych, W., and Graham, A.F. (1962). Isolation and properties of an RNA-containing bacteriophage. J Cell Comp Physiol *60*, 199–208.

Pashev, I.G., Dimitrov, S.I., and Angelov, D. (1991). Crosslinking proteins to nucleic acids by ultraviolet laser irradiation. Trends in Biochemical Sciences *16*, 323–326.

Pelletier, J., and Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature *334*, 320–325.

Pfeffer, S., Brandt, F., Hrabe, T., Lang, S., Eibauer, M., Zimmermann, R., and Förster, F. (2012). Structure and 3D Arrangement of Endoplasmic Reticulum Membrane-Associated Ribosomes. Structure *20*, 1508–1518.

Phan, A.T., Kuryavyi, V., Darnell, J.C., Serganov, A., Majumdar, A., Ilin, S., Raslin, T., Polonskaia, A., Chen, C., Clain, D., et al. (2011). Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. Nat. Struct. Mol. Biol. *18*, 796–804.

Picotti, P., and Aebersold, R. (2012). Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nat. Methods *9*, 555–566.

Poltronieri, P., Sun, B., and Mallardo, M. (2015). RNA Viruses: RNA Roles in Pathogenesis, Coreplication and Viral Load. Curr Genomics *16*, 327–335.

Pool, M.R. (2005). Signal recognition particles in chloroplasts, bacteria, yeast and mammals (Review). Molecular Membrane Biology *22*, 3–15.

Preiss, T., Muckenthaler, M., and Hentze, M.W. (1998). Poly(A)-tail-promoted translation in yeast: implications for translational control. RNA *4*, 1321–1331.

Quinn, J.J., and Chang, H.Y. (2016). Unique features of long non-coding RNA biogenesis and function. Nat Rev Genet *17*, 47–62.

Rabl, J., Leibundgut, M., Ataide, S.F., Haag, A., and Ban, N. (2011). Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. Science *331*, 730–736.

Ramakrishnan, V. (2002). Ribosome Structure and the Mechanism of Translation. Cell *108*, 557–572.

Ramesh, M., and Woolford, J.L. (2016). Eukaryote-specific rRNA expansion segments function in ribosome biogenesis. RNA.

Reichow, S.L., Hamma, T., Ferré-D'Amaré, A.R., and Varani, G. (2007). The structure and function of small nucleolar ribonucleoproteins. Nucleic Acids Res. *35*, 1452–1464.

Ricci, E.P., Limousin, T., Soto-Rifo, R., Rubilar, P.S., Decimo, D., and Ohlmann, T. (2013). miRNA repression of translation in vitro takes place during 43S ribosomal scanning. Nucleic Acids Res *41*, 586–598.

Riley, A., Jordan, L.E., and Holcik, M. (2010). Distinct 5′ UTRs regulate XIAP expression under normal growth conditions and during cellular stress. Nucleic Acids Res *38*, 4665–4674.

Robinson, E., and Brown, R. (1953). Cytoplasmic particles in bean root cells. Nature *171*, 313.

Saletore, Y., Meyer, K., Korlach, J., Vilfan, I.D., Jaffrey, S., and Mason, C.E. (2012). The birth of the Epitranscriptome: deciphering the function of RNA modifications. Genome Biology *13*, 175.

Salzman, J., Gawad, C., Wang, P.L., Lacayo, N., and Brown, P.O. (2012). Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. PLOS ONE *7*, e30733.

Sasarman, F., Brunel-Guitton, C., Antonicka, H., Wai, T., Shoubridge, E.A., and Consortium, L. (2010). LRPPRC and SLIRP Interact in a Ribonucleoprotein Complex That Regulates Posttranscriptional Gene Expression in Mitochondria. Mol. Biol. Cell *21*, 1315–1323.

Savino, R., and Gerbi, S.A. (1990). In vivo disruption of Xenopus U3 snRNA affects ribosomal RNA processing. EMBO J. *9*, 2299–2308.

Scherrer, T., Mittal, N., Janga, S.C., and Gerber, A.P. (2010). A Screen for RNA-Binding Proteins in Yeast Indicates Dual Functions for Many Enzymes. PLOS ONE *5*, e15499.

Schmid, M., Jaedicke, A., Du, T.-G., and Jansen, R.-P. (2006). Coordination of Endoplasmic Reticulum and mRNA Localization to the Yeast Bud. Current Biology *16*, 1538–1543.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Semrad, K., and Semrad, K. (2010). Proteins with RNA Chaperone Activity: A World of Diverse Proteins with a Common Task—Impediment of RNA Misfolding, Proteins with RNA Chaperone Activity: A World of Diverse Proteins with a Common Task—Impediment of RNA Misfolding. Biochemistry Research International, Biochemistry Research International *2011*, *2011*, e532908.

Slayter, H.S., Warner, J.R., Rich, A., and Hall, C.E. (1963). THE VISUALIZATION OF POLYRIBOSOMAL STRUCTURE. J. Mol. Biol. *7*, 652–657.

Smith, T.F., Lee, J.C., Gutell, R.R., and Hartman, H. (2008). The origin and evolution of the ribosome. Biol. Direct *3*, 16.

Spahn, C.M., Beckmann, R., Eswar, N., Penczek, P.A., Sali, A., Blobel, G., and Frank, J. (2001). Structure of the 80S ribosome from Saccharomyces cerevisiae--tRNA-ribosome and subunit-subunit interactions. Cell *107*, 373–386.

Spahn, C.M.T., Jan, E., Mulder, A., Grassucci, R.A., Sarnow, P., and Frank, J. (2004). Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. Cell *118*, 465–475.

Squatrito, M., Mancino, M., Donzelli, M., Areces, L.B., and Draetta, G.F. (2004). EBP1 is a nucleolar growth-regulating protein that is part of pre-ribosomal ribonucleoprotein complexes. Oncogene *23*, 4454–4465.

Squatrito, M., Mancino, M., Sala, L., and Draetta, G.F. (2006). Ebp1 is a dsRNA-binding protein associated with ribosomes that modulates eIF2α phosphorylation. Biochemical and Biophysical Research Communications *344*, 859–868.

Srisawat, C., and Engelke, D.R. (2001). Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. RNA *7*, 632–641.

Srivastava, A.K., and Schlessinger, D. (1991). Structure and organization of ribosomal DNA. Biochimie *73*, 631–638.

Stadler, P.F., Chen, J.J.-L., Hackermüller, J., Hoffmann, S., Horn, F., Khaitovich, P., Kretzschmar, A.K., Mosig, A., Prohaska, S.J., Qi, X., et al. (2009). Evolution of Vault RNAs. Mol Biol Evol *26*, 1975–1991.

Steitz, J.A. (1989). Immunoprecipitation of ribonucleoproteins using autoantibodies. Meth. Enzymol. *180*, 468–481.

Sysoev, V.O., Fischer, B., Frese, C.K., Gupta, I., Krijgsveld, J., Hentze, M.W., Castello, A., and Ephrussi, A. (2016). Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in Drosophila. Nat Commun *7*, 12128.

Tafforeau, L., Zorbas, C., Langhendries, J.-L., Mullineux, S.-T., Stamatopoulou, V., Mullier, R., Wacheul, L., and Lafontaine, D.L.J. (2013). The Complexity of Human Ribosome Biogenesis Revealed by Systematic Nucleolar Screening of Pre-rRNA Processing Factors. Molecular Cell *51*, 539–551.

Taylor, D.J., Devkota, B., Huang, A.D., Topf, M., Narayanan, E., Sali, A., Harvey, S.C., and Frank, J. (2009). Comprehensive molecular structure of the eukaryotic ribosome. Structure *17*, 1591–1604.

Tenenbaum, S.A., Lager, P.J., Carson, C.C., and Keene, J.D. (2002). Ribonomics: identifying mRNA subsets in mRNP complexes using antibodies to RNA-binding proteins and genomic arrays. Methods *26*, 191–198.

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

Tsai, B.P., Wang, X., Huang, L., and Waterman, M.L. (2011). Quantitative Profiling of In Vivo-assembled RNA-Protein Complexes Using a Novel Integrated Proteomic Approach. Mol Cell Proteomics *10*.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP Identifies Nova-Regulated RNA Networks in the Brain. Science *302*, 1212–1215.

Valásek, L., Hasek, J., Nielsen, K.H., and Hinnebusch, A.G. (2001). Dual function of eIF3j/Hcr1p in processing 20 S pre-rRNA and translation initiation. J. Biol. Chem. *276*, 43351–43360.

Valcárcel, J., Singh, R., Zamore, P.D., and Green, M.R. (1993). The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. Nature *362*, 171–175.

Van Blokland, R., Van der Geest, N., Mol, J. n. m., and Kooter, J. m. (1994). Transgene-mediated suppression of chalcone synthase expression in Petunia hybrida results from an increase in RNA turnover. The Plant Journal *6*, 861–877.

Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat. Methods *13*, 508–514.

Wahlestedt, C., Salmi, P., Good, L., Kela, J., Johnsson, T., Hokfelt, T., Broberger, C., Porreca, F., Lai, J., Ren, K., et al. (2000a). Potent and Nontoxic Antisense Oligonucleotides Containing Locked Nucleic Acids. Proceedings of the National Academy of Sciences of the United States of America *97*, 5633–5638.

Wahlestedt, C., Salmi, P., Good, L., Kela, J., Johnsson, T., Hökfelt, T., Broberger, C., Porreca, F., Lai, J., Ren, K., et al. (2000b). Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. Proc Natl Acad Sci U S A *97*, 5633–5638.

Wang, J., Dong, Z., and Bell, L.R. (1997a). Sex-lethal interactions with protein and RNA. Roles of glycine-rich and RNA binding domains. J. Biol. Chem. *272*, 22227–22235.

Wang, J., Dong, Z., and Bell, L.R. (1997b). Sex-lethal Interactions with Protein and RNA ROLES OF GLYCINE-RICH AND RNA BINDING DOMAINS. J. Biol. Chem. *272*, 22227–22235.

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.

Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature *171*, 737–738.

Wen, F., Zhou, R., Shen, A., Choi, A., Uribe, D., and Shi, J. (2012). The tumor suppressive role of eIF3f and its function in translation inhibition and rRNA degradation. PLoS ONE *7*, e34194.

Wessels, H.-H., Imami, K., Baltz, A.G., Kolinski, M., Beldovskaya, A., Selbach, M., Small, S., Ohler, U., and Landthaler, M. (2016). The mRNA-bound proteome of the early fly embryo. Genome Res.

West, J.A., Davis, C.P., Sunwoo, H., Simon, M.D., Sadreyev, R.I., Wang, P.I., Tolstorukov, M.Y., and Kingston, R.E. (2014). The Long Noncoding RNAs NEAT1 and MALAT1 Bind Active Chromatin Sites. Molecular Cell *55*, 791–802.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855–862.

Wilchek, M., and Bayer, E.A. (1988). The avidin-biotin complex in bioanalytical applications. Analytical Biochemistry *171*, 1–32.

Wild, T., Horvath, P., Wyler, E., Widmann, B., Badertscher, L., Zemp, I., Kozak, K., Csucs, G., Lund, E., and Kutay, U. (2010). A Protein Inventory of Human Ribosome Biogenesis Reveals an Essential Function of Exportin 5 in 60S Subunit Export. PLOS Biol *8*, e1000522.

Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. Nature *407*, 327–339.

Wolin, S.L., and Walter, P. (1989). Signal recognition particle mediates a transient elongation arrest of preprolactin in reticulocyte lysate. J. Cell Biol. *109*, 2617–2622.

Worton, R.G., Sutherland, J., Sylvester, J.E., Willard, H.F., Bodrug, S., Dubé, 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.

Wright, A.V., Nuñez, J.K., and Doudna, J.A. (2016). Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. Cell *164*, 29–44.

Wutz, A. (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat. Rev. Genet. *12*, 542–553.

Xue, S., and Barna, M. (2012). Specialized ribosomes: a new frontier in gene regulation and organismal biology. Nat Rev Mol Cell Biol *13*, 355–369.

Youngman, E.M., and Green, R. (2005). Affinity purification of in vivo-assembled ribosomes for in vitro biochemical analysis. Methods *36*, 305–312.

Yu, C.-H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M.S., and Liu, Y. (2015). Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. Molecular Cell *59*, 744–754.

Zielinski, J., Kilk, K., Peritz, T., Kannanayakal, T., Miyashiro, K.Y., Eiríksdóttir, E., Jochems, J., Langel, Û., and Eberwine, J. (2006). In vivo identification of ribonucleoprotein-RNA interactions. PNAS *103*, 1557–1562.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406–3415.

9 Presentations and Publications

Presentations

Birgit Schuster; Alfredo Castello; Benedikt Beckmann; Bruno Galy; Matthias W. Hentze; “In vivo capture of RBPs bound to defined RNA species” RNA meeting 2013 in Davos (Switzerland); Poster Presentation

Birgit Schuster; Alfredo Castello; Matthias W. Hentze; “RNP capture of defined RNA species in vivo” Protein Synthesis and Translational Control Meeting 2013 in Heidelberg (Germany); Poster Presentation

Birgit Schuster, Alfredo Castello, Mandy Rettel, Jeroen Krijgsveld, Matthias W. Hentze; “Oligo probe specific RNP capture” PhD retreat 2014 in Pilsen (Czech); Talk

Birgit Schuster; Alfredo Castello; Matthias Hentze; “Interactome capture of defined RNA species” Non coding RNA conference 2014 in Heidelberg (Germany); Poster Presentation

Birgit Schuster; Alfredo Castello; Christian Frese; Mandy Rettel; Jeroen Krijgsveld; Matthias W. Hentze; “Interactome capture of defined RNA species” Complex Life of mRNA meeting 2014 in Heidelberg (Germany); Poster Presentation

Birgit Rogell; Alfredo Castello; Mandy Rettel; Jeroen Krijgsveld; Matthias W. Hentze; “Identification of the proteome bound to a defined RNA” Protein Synthesis and Translational Control Meeting 2015 in Heidelberg (Germany); Poster Presentation

Publications

Most results presented in this dissertation are part of a manuscript in preparation.

10 Acknowledgements

This work was financially supported by the ERC Advanced Grant ERC-2011-ADG_20110310 provided to M.W. Hentze.

I want to thank all the people who contributed to the success of this work:

First of all, I like to thank **Matthias** for giving me the great opportunity to do my PhD in his “postdoc” lab and being always an encouraging and supportive person. I am grateful for his many advices and that he gave good examples as experienced mentor not only on research related topics but also on topics like presentation skills, scientific writing or project planning. I also want to thank him for investing time to critically proof-read this thesis.
I am especially grateful that I had **Alfredo** as lab supervisor, who even as PI in Oxford was always happy to send me his many ideas, advices and feedback. I learnt a lot from him. He always gave me critical feedback for my personal developments. I was happy about his many encouragements and his optimistic attitude he shared with me. Special thanks also for the precious time he invested during the critical proof-reading of this thesis. It was also a pleasure to have Alfredo as colleague in lab, with funny stories to remember my whole life.

Many thanks go also to my Thesis advisory committee members **Anne Ephrussi, Ralf Bartenschlager and Wolfgang Huber**, who took their time to follow the progress of my PhD work and to give me valuable feedback and advice.
I want to acknowledge **Georg Stoecklin**, who kindly agreed to be part of my thesis defense committee.
I am grateful for the statistical analysis from **Bernd Fischer**, an exceptional expert in his field. Thank you for the great discussions we had together.
I want to thank **Mandy Rettel** and the whole **EMBL Proteomics Core facility**, for the MS measurements of my samples, the many advices and the patience when my samples caused problems.
I want to thank **all members of the Hentze lab** (former and current) for being great colleagues, for creating a friendly and supportive atmosphere, for the very

constructive team meetings, for being great scientists and for the fun we had together. Special thanks go to **Amy** for teaching me the *in vitro* transcription and translation protocol and for the critical proof-reading of some parts of this thesis together with **Elli**.

From the previous lab members I particularly want to thank **Jan Medenbach**, who was always great in giving me valuable advice during our discussions. I want to thank him for his help with the input on the *in vitro* part of my thesis and the sharing of recombinant Sxl protein and Ssx antibodies. Many thanks also for the proof-reading of parts of the introduction of this thesis.

From the previous lab members I also want to thank **Christian Frese**, who was always happy to answer questions regarding mass spectrometry and performing some MS measurements for me.

I want to thank **members of the Ephrussi lab** for sharing *Drosophila* antibodies. In particular I want to thank **Frank Wippich** for the discussions we had and some experiments on oskar mRNA we performed together.

I was also always happy to come to the **EMBL Gene Core facility** with well-maintained instrumentation (qPCR, Bioanalyzer) and staff, which was always helpful and super nice.

During my PhD time I was happy to be part of the EMBL PhD programme and the helpful and friendly support I received from the **people working in the PhD office**

I also want to thank EMBL and the **EMBL Kinderhaus** for the great support with childcare. It is really helpful to know your child playing happily nearby when you focus on your PhD work.

Special thanks go to my **parents and my family** who supported me throughout my whole life and gave me the freedom and the opportunity to do what I like. I am also very thankful for the childcare during the preparation of my PhD thesis.
Danke Mama&Papa!