Molecular mechanisms of *msl2* mRNA translational regulation

Marina García Beyaert

TESI DOCTORAL UPF / 2015

THESIS DIRECTOR

Dra. Fátima Gebauer

GENE REGULATION, STEM CELLS AND CANCER DEPARTMENT

CENTRE FOR GENOMIC REGULATION (CRG)

upf. Universitat Pompeu Fabra Barcelona

To my family…

Aknowledgements

And here comes to its end a long adventure on the land of knowledge. As all adventures, the expectations vary all along the way, with all the uncertainty and ups and downs that this implies. It has overall been a very interesting, intense and enriching trip. I would thus like to thank Fátima, for her expertise, support all along the way and humanity, important aspects to start and continue walking. I would also like to thank all the members of the lab, former and present, for making a nice working atmosphere. Especially, Tanit and Hima, for sharing one or another coffee when it was more needed, and Olga for being always available for any question. Also Jae-Seong, for interesting and fruitful discussions. My friends from the Master, for making me feel surrounded all these years, since the very beginning I arrived to Barcelona. My "mountain adventure" friends, especially Ádria and Laura, for the wonderful, funny and refilling times spent in the outside world adventure.

I would also like to thank Ana, for the fun, confidence and trust. Arantxa, for the enriching conversations, support, laughs and wonderful refilling moments. Marie, with whom I had to chance to share important moments during these years, and who had the splendid capacity to help me making the hard ones much easier and share the fun of life.

Of course, I also would like to thank my parents, for their contribution, since much longer ago than the beginning of my PhD. For their trust and understanding, for their splendid capacity to clarify when it is obscure, and being by far the most undeniable support. And finally, I would like to thank Joaquim, for being by far one of the most precious persons who made me grow and laugh when it was hard, and who gave me confidence to stand up when it was harder.

Abstract

Regulation of *msl2* translation is a key step in the modulation of X-chromosome dosage compensation. MSL2 is the limiting subunit of the dosage compensation complex, an assembly that promotes hyper-transcription of the single male X-chromosome to equalize expression of X-linked genes between males and females. In females, dosage compensation must be repressed for viability, and this is achieved in large part by translational repression of *msl2* mRNA. The female-specific protein Sex-lethal (SXL) binds to both untranslated regions (UTRs) of the *msl2* transcript to inhibit two steps of translation initiation: SXL bound to the 3' recruits the co-factor UNR and inhibits ribosomal recruitment; SXL bound to the 5' UTR inhibits ribosomal scanning by promoting the recognition of an upstream AUG. In the lab, we recently found that eIF3d is a target of the 3'UTR repressor complex. In this thesis, we show that eIF3d can be recruited to the mRNA even in the absence of a cap structure by virtue of its binding to *msl2* 5' UTR. Our results suggest that recruitment of this factor to the mRNA by multiple routes may sensitize translation of *msl2* to inhibition of eIF3d.

In addition, we have identified residues of SXL important for 5' UTR, but not 3' UTR, -mediated repression. Analysis of SXL variants with mutations in these residues has led us to identify two factors in the ribosome, RACK1 and RPS3, as likely mediators of inhibition via the 5' UTR. Interestingly, RACK1, RPS3 and eIF3d are located close to each other on the ribosome, allowing us to propose an integrated model for translational repression that explains coordinated inhibition of ribosome recruitment and scanning by SXL.

Resumen

La regulación de la traducción de *msl2* es un paso crucial en la modulación de la compensación de la dosis del cromosoma X. MSL2 es la subunidad limitante del complejo de compensación de dosis génica, un complejo que promueve la hiper-transcripción del único cromosoma X en machos para igualar la expresión de sus genes a la de hembras. La viabilidad de las hembras require que la compensación de dosis esté reprimida, y esto se consigue en gran parte por la represión de la traducción del ARNm que codifica para MSL2. La proteína específica de hembras Sex-lethal (SXL) se une a ambas regiones no traducidas (UTRs) del mensajero para inhibir dos etapas del inicio de la traducción: SXL unido al 3' UTR recluta al co-factor UNR e inhibe el reclutamiento del ribosoma; SXL unido al 5' UTR inhibe el escaneo del ribosoma al promover el reconocimiento de un uAUG. En el laboratorio, identificamos recientemente eIF3d como una diana del complejo represor unido al 3' UTR. En esta tesis mostramos que eIF3d puede ser reclutado al ARNm incluso en ausencia del *cap*, gracias a su unión al 5' UTR. Nuestros resultados sugieren que el reclutamiento de este factor al ARNm a través de varias rutas sensibiliza la traducción de *msl2* a la inhibición de eIF3d.

Además, hemos identificado resíduos de SXL importantes para la represión mediada por el 5' pero no por el 3' UTR. El análisis de variantes de SXL con mutaciones en estos resíduos nos ha permitido identificar dos factores en el ribosoma, RACK1 y RPS3, como posibles mediadores de la inhibición por el 5' UTR. Curiosamente, RACK1, RPS3 y eIF3d están localizados en zonas próximas del ribosoma, lo que nos permite proponer un modelo integrado para la inhibición coordinada del reclutamiento y escaneo del ribosoma por SXL.

Preface

The information encoded in the mRNA is converted into functional polypeptides during translation. This step of gene expression is essential for cell function and it is tightly regulated to respond to the needs of the cell. The molecular mechanisms of translation, and how translation is regulated, have been a matter of study since many years. It is now clear that exceptions to the standard (cap-dependent) mechanism of translation initiation exist, and that a variety of factors (translation regulators) allow cells to cope with changing external stimuli. However, the molecular mechanisms used by these factors to interfere or intersect with the translation machinery are often obscure. Here we analyse the molecular mechanisms of translational repression of *msl2* mRNA by the protein SXL, and propose an integrated model to explain how SXL binding to both UTRs of the message achieves repression of two initial steps of translation: ribosome recruitment and scanning.

CONTENTS

I. INTRODUCTION

1. Translation

All cellular proteins are generated by mRNA translation, a process occurring in the cytoplasm that requires ribosomes and the sequential function of translation factors that bind to either the mRNA or the ribosome. Although the encoding function of the mRNA has been understood since it was first described in 1961 (BRENNER et al., 1961, SPIEGELMAN, 1961), it was assumed that the regulation of gene expression in response to cellular demands was achieved by adjusting transcription and protein stability. Today, however, we know that mRNA translation provides an important and highly responsive layer of gene expression control. In particular, translation initiation is the step that is more extensively regulated. In eukaryotes, the canonical translation initiation process requires more than 30 polypeptides that interact with each other in a dynamic and coordinated fashion to ensure the fidelity of start codon selection.

1.1. Translation initiation

Translation initiation is the process that enables the decoding of the mRNA start codon AUG, by the methyonyl initiator transfer RNA (Met-tRNAi^{Met}). In other words, translation initiation identifies the correct AUG and, therefore, the reading frame to generate the translated product. During canonical translation initiation in eukaryotes, the small (40S) ribosomal subunit is loaded into the mRNA as a 43S pre-initiation complex (PIC) that scans the 5' UTR until it reaches the AUG start codon (reviewed in Hinnebusch, 2014). The 43S PIC is formed by the association of the 40S with the Ternary Complex (TC: Met-tRNA$_i^{Met}$-eIF2-GTP), in a reaction promoted by eukaryotic initiation factor 1 (eIF1), 1A, 5, and the multi-

subunit complex eIF3 (Figure 1). Loading of the 43S PIC onto the mRNA involves contacts between eIF3 and eIF4F, a complex bound to the m^7GpppN cap at the 5'end of the mRNA which contains the cap binding protein eIF4E, the scaffolding protein eIF4G, and the RNA helicase eIF4A. eIF4G exhibits binding sites for eIF4E, eIF4A, poly(A)-binding protein (PABP), eIF3 (only in mammals), and eIF1 plus eIF5 (only in yeast), and can also bind mRNA directly, although it is normally recruited to the mRNA via cap-bound eIF4E. The contacts of eIG4G with eIF4E and PABP enable the formation of a stable closed loop mRNA structure. This structure is thought to allow recycling of ribosomes after subsequent rounds of initiation in the same mRNA, and also to activate the recruitment of the PIC. Binding of eIF4G to eIF4A allows its recruitment to a cap-proximal position and activates the helicase activity of eIF4A, leading to local unwinding of the mRNA. This process, together with eIF4G interactions with eIF3, eIF5 or eIF1, promotes loading of the 43S PIC - which is in an open and scanning competent conformation - onto the mRNA. The attachment of the 43S PIC is confined to the 5'end of the 5'UTR (PICs cannot be recruited to circular mRNAs), which is scanned base by base until the anticodon (AC) of the Met-tRNAi^{Met} shows complementarity with an AUG. During scanning, eIF5 stimulates hydrolysis of eIF2-GTP, but (Pi) release only occurs after AUG recognition and requires eIF1 dissociation from the 40S. Upon AUG recognition, eIF1 dissociation and (Pi) release are thought to induce a closed, scanning incompetent, conformation of the ribosome. The eIF2•GDP and probably eIF5 as well, then dissociate from the 40S and are replaced by the 60S ribosomal subunit. This process is mediated by the GTPase eIF5B and leads to the formation of the 80S initiation complex (IC) that can now translate the ORF of the mRNA (Figure 1) (Hinnebusch, 2014, Aitken and Lorsch, 2012, Jackson et al., 2010).

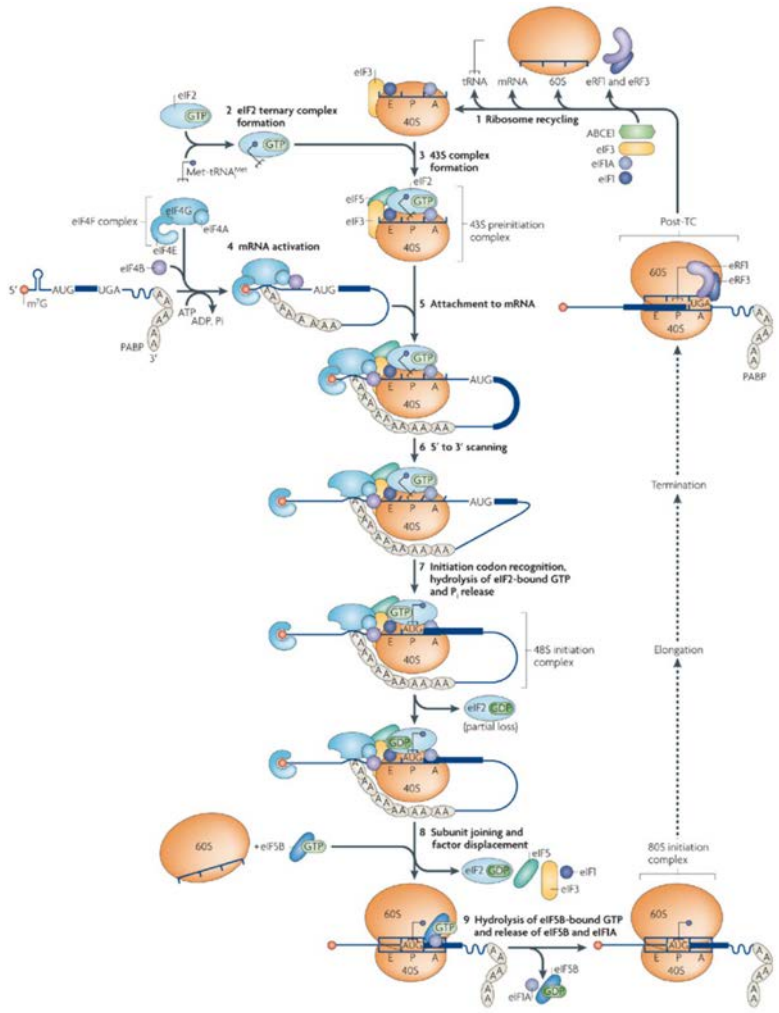


Figure 1. **Schematic illustration of the translation initiation process**. Taken from Jackson et al 2010.

1.2. Ternary complex formation

eIF2 is a heterotrimer composed of α and β subunits that bind to a central γ subunit. In the ternary complex, the γ subunit binds to GTP and tRNA, while subunits α and β stabilize the interaction with the tRNA. The Met-tRNAi^{Met} contains some conserved signature sequences that confer eIF2 specificity for the initiator versus the elongator Met-tRNA. On the other side, Met-tRNA$_i^{Met}$ binds with higher affinity to eIF2-GTP than to eIF2-GDP, a specificity of coupling that depends on the Met moiety (Kapp and Lorsch, 2004). eIF2B recycles the eIF2-GDP resulting from a previous round of translation initiation into eIF2-GTP which can be loaded into a new TC, a function that is used as a powerful target for controlling general translation (Sonenberg and Hinnebusch, 2009). Indeed, this recycling is inhibited by phosphorylation of eIF2α by kinases involved in various types of stress response, leading to general shut down of translation.

1.3. PIC assembly

The factors that promote the assembly of the PIC are eIFs 1, 1A, 3 and 5, which help loading the TC onto the 40S in such a way that the Met-tRNA$_i^{Met}$ is positioned at the P site of the ribosome. eIF3, along with eIFs 1, 5 and the ternary complex constitute the multifactor complex (MFC) that can be formed in the absence of the ribosome both *in vivo* and *in vitro*. In yeast, the MFC is stabilized by:

- eIF5 C-terminal domain (CTD) interactions with eIF2β N-terminal tail (NTT)
- eIF1 interactions with eIF2β –NTT and eIF3c N-terminal domain (NTD)
- eIF3a-CTD interactions with eIF2β

Since eIFs 1, 1A, 3 and 5 interact with the 40S, it has been proposed that the preformed MFC provides a major pathway to TC recruitment *in vivo*. The MFC was also proposed to be a depot to maintain the critical elements of translation together close to sites of translation. Indeed, the mammalian MFC appears to be the major reservoir of eIF3.

The interaction between the ternary complex and the 40S is also mediated by interactions of eIF2γ with helix 44 of the 18S rRNA, interactions of eIF2β with eIF1A, and interaction of eIF2α with eIF3d bound to the 40S in the exit channel (des Georges et al., 2015).

1.4. 43S PIC attachment to the mRNA

The m^7GpppN cap and its bound complex eIF4F are central in this process. Binding of eIF4G to eIF4A is thought to activate its ATP-dependent RNA helicase activity to unwind RNA structures near the cap to provide a landing platform for the PIC. The helicase activity of eIF4A is also enhanced by eIF4B and eIF4H, however how they stimulate eIF4A is unclear.

In addition, eIF4G interacts with eIF3, attracting the PIC to a cap-proximal position. In mammals, this interaction is mediated by eIF4G binding subunits eIF3e, eIF3c and eIF3d (Villa et al., 2013, LeFebvre et al., 2006). The fact that the ribosome engages in the mRNA at its 5' end ensures a preference for translation of the 5'-proximal AUGs. In yeast, however, eIF3 and eIF4G lack their interaction domains and they have not been observed to bind to each other. But yeIF4G can interact with yeIF5, and eIF5-CTD can bridge yeIF4G interaction with yeIF3c-NTD, thus promoting an indirect yeIF4G/yeIF3 interaction. Since eIF5 can also interact with the 40S, the interaction yeIF4G/yeIF5 may also help to recruit the PIC. Mammalian and yeast eIF4B can as well interact with eIF3, and it has also been proposed that eIF4B can help to load the PIC onto the mRNA (Aitken and Lorsch, 2012, Hinnebusch, 2014). Although it seems that the ribosome can be recruited

following different strategies in yeast and mammals, the role of most initiation factors is conserved.

The mRNA closed-loop structure resulting from the interaction of PABP with eIF4G is assumed to be crucial for PIC recruitment. However, the requirement of eIF4G interaction with PABP varies with cell types. In yeast, removal of the PABP binding domain on yeIF4G has an effect on global translation only if both, the mRNA and eIF4E binding sites in eIF4G are also mutated. Thus, in yeast, the closed-loop structure would be incidental to efficient 43S attachment (Hinnebusch, 2014).

1.5. Scanning and AUG recognition

Once the 43S PIC is located at the 5' end of the mRNA, it starts scanning the 5' UTR using the anticodon bases AC of the Met-tRNA$_i^{Met}$ to identify the start codon AUG. The molecular basis of the scanning process are poorly understood. Scanning requires both a specific conformation of the ribosome that allows movement along the 5' UTR, and the unwinding of mRNA structures along the way. In addition to eIF4A, other helicases may participate in the unwinding of the 5' UTR, including Dhx29 and Ded1. Scanning also requires energy to promote the movement towards the forward direction (5' to 3'), accompanied by mechanisms that block excursions in the reverse direction (3' to 5'). Scanning is very processive, and it is thought that multiple PICs can simultaneously scan the 5'UTR.

Once the mRNA is initially unwound, to be threaded through the mRNA channel in the ribosome, the ribosome needs to be in an open conformation, compatible with scanning, that closes once the AUG is recognized. In yeast, the 40S alone contains a latch at the mRNA entry channel, formed by the interaction between helix 18 and 34 of the 18S rRNA that impedes the association of the mRNA with the ribosome. Upon eIF1 and eIF1A binding, the latch of the 40S opens, and the mRNA can be recruited and scanned. When eIF1 is released after AUG recognition, the latch of the ribosome is

though to close again, committing the ribosome to start translation from that AUG (Passmore et al., 2007). However in mammals, the mRNA entry channel appears closed in X-ray data of the 40S bound to eIF1 and eIF1A (Lomakin and Steitz, 2013). To explain this difference, it was proposed that a closed latch may be required to lock the mRNA during scanning, and that the open conformation observed in yeast might be only required for the initial recruitment of the mRNA. An alternative explanation is that, unlike yeast, mammalian eIF3 may be additionally required to open the latch of the 40S-eIF1-eIF1A complex (Hinnebusch, 2014).

In both mammals and yeast, eIF1 impedes $tRNA_i^{Met}$ pairing at non-AUG codons. Only when an AUG is recognized, a conformational change allows eIF1 to be released, leading to a closed (scanning refractory) conformation of the 40S. Indeed, eIF1 seems to clash with Met-$tRNA_i^{Met}$ at the P site, such that when the Met-$tRNA_i^{Met}$ recognizes an AUG and is stably positioned at the P site, this weakens the binding of eIF1 with the 40S. Short repeats at the yeIF1A-CTT (scanning enhancers [SE]) may cooperate with eIF1 in this clashing function. Release of eIF1 is thought to be mediated by the eIF3c-NTD as well, since eIF3c and 40S binding surfaces overlap in eIF1 (Hinnebusch, 2014). eIF5 also helps eIF1 to dissociate from the PIC once the AUG is recognized. Upon eIF1 dissociation, the free Pi resulting from the hydrolysis of eIF2-GTP during scanning is also released, making GTP hydrolysis irreversible and committing initiation at the selected AUG (Algire et al., 2005). At this moment, the 43S PIC forms a stable complex with the mRNA (resistant to sucrose gradient fractionation) called the 48S initiation complex.

Particular sequences surrounding the AUG enhance AUG selection by the PIC (Kozak, 1978). In mammals this optimum context is: 5'-(A/G)NNAUGG-3'. AUGs that deviate enough from this context are bypassed by the ribosome in a process termed leaky scanning. The optimal context seems to participate in stabilizing the closed state of the PIC. Exactly

how the context is recognized is unknown, but eIF2α seems to contact the -3 position (taking as 1 the A of the AUG), while the ribosomal RNA seems to contact the +4 position (Pisarev et al, 2006).

1.6. Subunit joining and 80S initiation complex formation

Once the PIC closes at the AUG start codon, eIF2-GDP and eIF5 dissociate and are replaced by the 60S ribosomal subunit, in a process mediated by the GTPase eIF5B. Hydrolysis of GTP by eIF5B upon subunit joining enables the conformational rearrangements leading to formation of the 80S initiation complex. Since eIF5B-GDP has low affinity for the initiation complex compared to eIF5B-GTP, eIF5B is released after GTP hydrolysis. Subunit joining is also promoted by the interaction of eIF1A-CTT with domain IV of eIF5B. eIF1A only leaves the interface of the ribosome after eIF5B is released (Hinnebusch, 2014, Aitken and Lorsch, 2012).

1.7. Re-initiation and upstream open reading frames (uORFs)

Reinitiation of translation after termination at the stop codon requires reassembly of the TC, and the presence of the translation initiation factors at the 5' end of the mRNA. The closed loop structure stabilizes eIF4F binding to the 5' end of the mRNA and facilitates re-engagement of ribosomes into the 5'-end of the mRNA after they terminated translation (Sonenberg and Hinnebusch, 2009). Re-initiation can also occur in a linear mode after an upstream open reading frame (uORF). After termination at the uORF stop codon, the 40S subunit does not detach from the mRNA and resumes scanning, during which time it is reloaded with Met-tRNA$_i^{Met}$. Efficient reinitiation depends on: i) cis-acting mRNA features surrounding the uORF; ii) the time required for the translation of the uORF (the longer it takes, the

more initiation factors have the opportunity to leave the ribosome, and the less competent is the ribosome for another round of translation); iii) various initiation factors; iv) the distance between the uORF termination codon and the downstream AUG, as the rescanning PICs require a certain time to recruit a new TC (reviewed in Valásek, 2012). At least in yeast, eIF3 remains bound to the 80S after several rounds of elongation, and critically enhances the re-initiation capacity of post-termination 40S ribosomes. In the *GCN4* mRNA, which contains several uORFs, retention of eIF3a at the 40S after termination at uORF1 is mediated by an mRNA enhancer sequence (Valásek, 2012).

1.8. eIF3: a complex translation initiation factor

eIF3 is the largest translation initiation factor, a structure of 804 kDa involved in almost all steps of translation initiation: TC assembly into the 43S PIC, recruitment of the PIC to the mRNA, and scanning. In mammals, eIF3 contains 13 subunits (named from a to m), and in the budding yeast *S. cerevisiae* it is only formed by 6 (a, b, c g, i and j). Subunits are conserved among species, and therefore very likely their function are conserved as well. The six subunits from *S. cerevisiae* are thought to provide the basic functions, whereas subunits only present in mammals may provide regulatory roles. It is thought that each eIF3 contains one copy of each subunit, but in fission yeast (*S. pombe*) non-core subunits might be absent from some eIF3 complexes, and may be only required for translation of some specific mRNAs (Zhou et al., 2005). Six subunits of mammalian eIF3 (a, c, e, k, l and m) contain a PCI domain, a structural motif present in two other complexes: the proteasome lid and the signalosome COP9 (PCI stands for *P*roteasome, *C*OP9 signalosome, and *I*nitiation factor). Subunits f and h contain an MPN (*M*pr1p and *P*AD1p *N*-terminal) domain, as also do some related proteins of the proteasome lid and the signalosome COP9. This suggests that a link may exist between protein synthesis and protein degradation.

A cryo-EM model of human eIF3 showed it resembles a five lobed hand that embraces the 40S subunit from the mRNA entry to the exit channel (Figure 2). The bulk of the eIF3 is located at the mRNA exit channel, as suggested by cross-link experiments (Pisarev et al., 2008), and observed in recent cryo-EM reconstitutions (des Georges et al., 2015). This location fits with the role of eIF3 in recruitment, since eIF4G also locates in this region of the ribosome. Interaction between both factors is mediated by a small internal region of mammalian eIF4G (residues 1015-1118) and subunits eIF3c, eiF3d and eIF3e (Sun et al., 2011, Korneeva et al., 2001). eIF3 also interacts with poly(A)-binding-protein-interacting-protein 1 (Paip1) that binds PABP, suggesting that eIF3 also contributes to the formation of the closed loop structure that stimulates ribosome recruitment, and ribosome recycling (reviewed in Hershey, 2015). In the ribosome, eIF3 interacts with several other initiation factors like eIF1, eIF2 and eIF5. eIF3, thus, seems to contribute to the organization of the initiation factors on the 40S surface.

During scanning, eIF3 has a regulatory role as well. eIF3c-NTD binding eIF1 promotes release of eIF1 from the PIC, since eIF3c and 40S binding surfaces overlap in eIF1(Hinnebusch, 2014). A conserved module of yeIF3 (subunits j, eIF3b RRM, and eIF3a-CTD) has also been implicated in AUG recognition. This module is localized near the entry channel, and mutation of these subunits reduces efficient AUG recognition. However how they exactly contribute to the process is not clear (reviewed in Hinnebusch, 2014, Aitken and Lorsch, 2012).

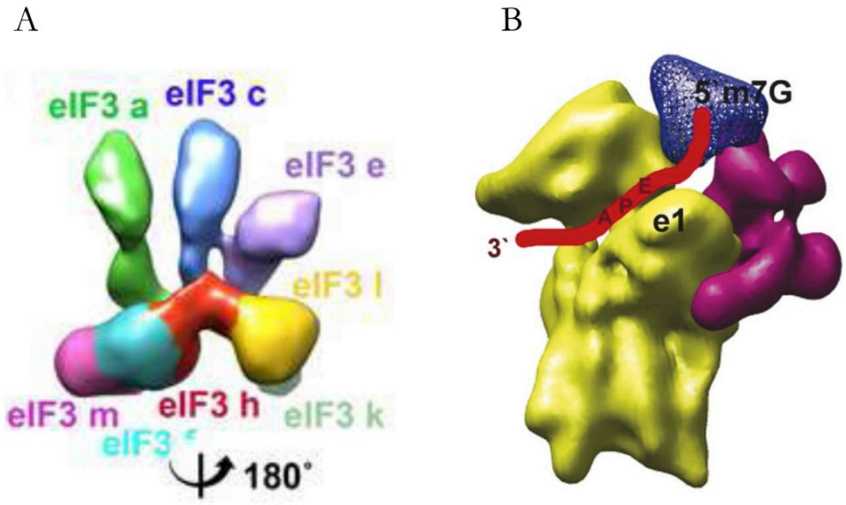


Figure 2. 3D reconstruction of the human eIF3 alone (A) or attached to the 40S ribosomal subunit (B). Taken from Querol-Audi et al. (2013) and Siridechadilok et al (2005).

2. Alternative mechanisms of ribosome recruitment

Ribosome recruitment and translation of most mRNAs in the cell occurs by the standard mechanism exposed above that involves the chain of interactions 7m-G/eIF4E/eIF4G/eIF3/43S. However, recruitment of the ribosome can follow alternative routes that ensure translational activity in cases where specific translation factors are inactivated. Some of the known alternative mechanisms of ribosome recruitment are summarized in this section.

2.1. Internal ribosome entry site (IRES) structures

IRES were first found in viral mRNAs (Pelletier and Sonenberg, 1988, Jang et al., 1988). Viruses use IRESs to ensure viral translation while cap-dependent host translation is inhibited by the virus or by the host itself. IRESs are RNA structures located in the 5' UTR of some mRNAs that have the ability to recruit ribosomes in a cap-independent manner. Circularized mRNAs containing IRES can be translated, which confirms the robustness of the mechanism. To initiate translation, IRESs require non-canonical interactions with eukaryotic initiation factors and/or 40S subunits. Initiation on some IRESs also requires specific RNA binding proteins named IRES trans-acting factors (ITAFs), some of which stabilize the optimal three-dimensional IRES conformation (reviewed in Jackson et al., 2010, Kieft, 2008).

According to their requirements on eIFs and the location of the AUG, IRESs can be classified in 4 types (Figure 3A):

- Type 1: IRESs that bind domain p50 of eIF4G, even when the protein lacks the eIF4E binding site. Translation initiates at a codon somewhat downstream of the IRES (Picarnoviruses, i.e. poliovirus).
- Type 2: As Type1, they bind directly to eIF4G in the absence of eIF4E, but translation starts immediately at the 3' end of the IRES (Picarnovireses, ie. EMCV).

- Type3: IRESs that do not require eIF4F, nor eIFs 1 and 1A, but they do use a subset of eIFs (eIF3 and eIF2) as well as the Met-$tRNA_i^{Met}$ (i.e. HCV).
- Type 4: IRESs that bind the 40S and do not require any initiation factor. They do not need Met-$tRNA_i^{Met}$ neither (i.e. CrPV).

Although still a matter of debate, IRESs seem to be present in cellular mRNAs (Thompson, 2012). For example, recently, RNA structures in the 5'UTR of Hox RNAs were found to behave as IRESs (Xue et al., 2015). Additionally, ribosomal protein Rps25 is required for viral IRES translation, and its function is conserved across IRES types (Landry et al., 2009). Rps25 is not required for cap-dependent translation, suggesting that the existence of this protein may reflect the usage of an IRES-dependent mechanism for at least some cellular mRNAs. Thus, although the methods to assess cellular IRESs have not always been correctly applied, and the list of published cellular IRESs may require revision, it seems reasonable that cellular IRESs might exist.

2.2. Cap Independent Translation Enhancers (CITEs)

Another strategy for cap-independent translation used by plant viruses is the CITEs (Shatsky et al., 2010). As IRESs, they are RNA structures, but most of them localize in the 3' UTR. CITEs are though to recruit components of the translation apparatus and deliver them to the 5' end of the mRNA through long distance base paring between 5' and 3' UTRs (Figure 3B). Most studied CITEs bind to eIF4E or eIF4G, or both, to bring eIF4F to the 5'end of the mRNA, except for the CITE of Turnip crinkle virus (TCV) that binds directly to 60S and 80S ribosomes. It has been proposed that CITEs may also exist in cellular mRNAs; however, to the best of our knowledge, no evidence has been shown so far.

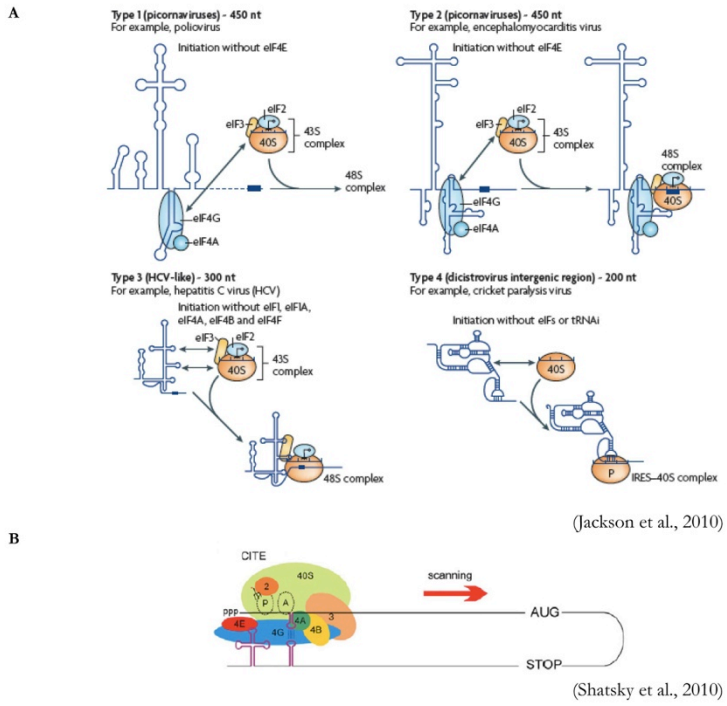


Figure 3. **Cap-independent mechanisms of translation initiation of viral mRNAs**. (A) Internal ribosome entry sites (IRESs) structures classified based on their dependence on eIFs and on the position of the AUG start codon with respect to the IRES. Taken from Jackson et al, 2010. (B) CITE-assisted initiation. 3′ CITEs are shown in violet. In the particular case depicted, eIF4E and eIF4G bind by means of CITEs to the 3′ UTR of the mRNA. Interactions of the CITE with loop structures located in the 5' UTR result in the circularization of the mRNA, positioning the translation factors at the 5' end of the mRNA.

2.3. Other modes of ribosomal recruitment

As mentioned above, in yeast, the ribosome can be recruited following a different chain of interactions than the reported: 7m-G-eIF4E-eIF4G-eIF3-43S. Additionally, some experiments surprisingly suggest that cap- and IRES-independent mechanisms of translation can take place in the cell. For instance, overexpression of a non-phosphorylatable mutant 4E-BP1 (unphosphorylated 4E-BP competes with eIF4G to bind eIF4E) in mammalian cells, inhibits translation of a *Renilla* reporter by only 60%, and overexpression of the WT 4E-BP1 inhibits it by only 20% (Mothe-Satney et al., 2000). This result suggests that translation of this IRES-less reporter is relatively insensitive to eIF4E inhibition, and hence that some mRNAs might not absolutely require this initiation factor for their translation. Indeed, in rabbit reticulocyte lysates (RRL), truncated eIF4G proteins whose eFI4E binding site (N-terminal region) has been separated from the eIF4A and eIF3 sites (central and C-terminal regions) by the L-protease, could stimulate translation of uncapped mRNAs that do not bear IRES structures, but not that of capped transcripts (Ohlmann et al., 1995). The same results were obtained working with an eIF4G truncated protein synthesized in bacteria that lacks the eIF4E binding site (De Gregorio et al., 1998). This truncated protein was then shown to contain an RNA binding motif that is required for its stimulatory effect on translation (Prévôt et al., 2003). Therefore, it seems that direct eIF4G binding to the mRNA may be able to to stimulate translation of uncapped mRNAs without any eIF4E requirement.

Similar to eIF4G, a recent report showed that some eIF3 subunits (a, b, d and g) can bind sequences in the 5'UTR of mammalian mRNAs. These interactions could activate or repress translation, depending on the mRNA target (Lee et al., 2015). Although these interactions occurred in the context of capped mRNAs, one could imagine that direct binding of eIF3 subunits to uncapped transcripts could also stimulate translation.

Thus, mRNAs have developed non-canonical mechanisms to recruit initiation factors or ribosomal subunits without the requirement of the cap, the cap-binding factor eIF4E, or an IRES structure.

3. Translation regulation

Translational control allows to rapidly switch gene expression on and off, or tune its intensity according to biological or metabolic needs. Several biological processes, including embryonic development, neuronal plasticity, cell differentiation and division, senescence and response to cellular stress, rely on the regulation of translation (Sonenberg 2009). In particular, translation initiation is the step that is more frequently regulated. The mechanisms that regulate translation initiation can be divided in two groups: those that affect general translation, by impacting on the eIFs or ribosomes, and those that affect translation of specific mRNA subsets through sequence-specific mRNA binding proteins (RBPs) or microRNAs. We will present examples of both general and mRNA-specific categories; regarding the latter, we will focus on RBP-mediated mechanisms, which is the subject of this thesis.

3.1. Regulation of global translation

Global translational control occurs in situations of cellular stress (i.e. viral infection, starvation, cold shock, etc), or when the cell needs to divide after hormone or growth factor stimuli. The most prevalent pathways of global translational control target the formation of the eIF4F complex or the recycling of eIF2 (Figure 4A).

General regulation via eIF4F is primarily based on 4E binding proteins (4E-BPs). When 4E-BPs are hypo-phosphorylated, they bind eIF4E in the same site recognized by eIF4G, thus blocking the assembly of functional eIF4F. However, upon metabolic stimulation, 4E-BPs are phosphorylated via the

mTOR kinase cascade, leading to release of eIF4E and translational activation (reviewed in Jackson et al., 2010).

Regarding eIF2 recycling, eIF2α phosphorylation at serine 51 blocks GTP recycling on eIF2, due to retention or sequestering of eIF2B. This reduces the availability of TC, and therefore reduces general translational levels. eIF2α is phosphorylated by several stress-responsive kinases: HRI, PERK and PKR (reviewed in Jackson et al., 2010).

Another mechanism of global translation regulation is the assembly of mRNAs into cytoplasmic structures termed mRNP granules. They are present in different cell types and their classification depends on the cellular context and the composition of the granules (P-bodies, stress-granules, germ granules or neuronal transport granules) (reviewed in Buchan, 2014).

3.2. Regulation by 5' UTR binding proteins

Regulation of translation by proteins bound to 5' UTRs is surprisingly rare, and there are only few well-known mechanisms (Figure 4B). One of them pertains the regulation of ferritin mRNAs by iron regulatory proteins (IRP) 1 and 2. In conditions of low iron, IRP binds to the 5' UTR of ferritin and inhibits its translation. The general principle of this mechanism is that strong inhibition requires binding of the IRP at cap-proximal locations, which prevents loading of the 43S complex onto the mRNA, but not eFI4F binding to the cap structure. If moved to a more cap-distal position, 43S scanning ribosomes can displace IRP and translation is not repressed (Muckenthaler et al., 1998).

Another example is the regulation of PABP mRNA translation. PABP protein can bind to clustered oligo(A) motifs in PABP 5' UTR that are 70-130 nucleotides downstream of the cap, thus providing an auto-regulatory mechanism. Bound PABP can block scanning 43S complexes without being displaced by them (de Melo Neto et al., 1995).

A third example is the case of SXL bound to the 5'UTR of *msl2* mRNA in *Drosophila*. In this case, SXL inhibits the scanning of the 43S complex by promoting the recognition of an uAUG, thus reducing translation at the main AUG. Substitution of SXL by an unrelated protein disrupts the repressive mechanism, thus indicating that it is not achieved by simple steric hindrance (Beckmann et al., 2005, Medenbach et al., 2011).

3.3. Regulation by 3' UTR binding proteins

In contrast to above, there are numerous cases of 3'-mediated translational control. This indicates that there might be some benefit on repressing translation by proteins binding to the 3' UTR, rather than the 5' UTR. Certainly, regulators placed in the 3' UTR are not displaced by scanning ribosomal subunits or elongating ribosomes (this would occur in the 5' UTR or the ORF, respectively). Furthermore, 3'UTR sequences are not evolutionarily constrained by having to support ribosome transit or to encode a protein; so their evolutionary freedom has allowed for a higher diversity of regulatory elements.

How can 3' UTR binding proteins regulate translation initiation at the 5' end? The closed-loop model provides a physical framework for the activity of 3' UTR binding factors. Communication between 5' and 3' UTRs can also be brought about by long range structures, or by natural cooperativity between RBPs that recognize opposite UTRs (de Moor et al., 2005, Lin and Bundschuh, 2015).

A variety of mechanisms have been described for proteins that bind to the 3' UTR (reviewed in Szostak and Gebauer, 2013). In many cases, a protein X binds to the 3' UTR in a sequence-specific manner, and interacts with an intermediate bridging protein (Y), which in turn interacts with a cap-binding factor (Z) (Figure 4C). This inhibits the formation of eIF4F or its assembly at the 5' end of the mRNA. The functions of protein X and Y can be embedded in one single protein, or in a group of proteins (Jackson et al.,

2010). In other cases, 3' UTR binding proteins inhibit the recruitment of the 43S PIC by interfering with the interaction of eIF3 to eIF4G, inhibit 60S subunit joining at the initiation codon, or inhibit translation elongation. In addition, they can promote deadenylation of the mRNA (thereby inhibiting closed-loop formation), alter the structure of the 3' UTR to expose or hide miRNA binding sites, or promote mRNA packing into higher order structures refractory to translation. A summary of these mechanisms is shown in Figure 4D (Szostak and Gebauer, 2013).

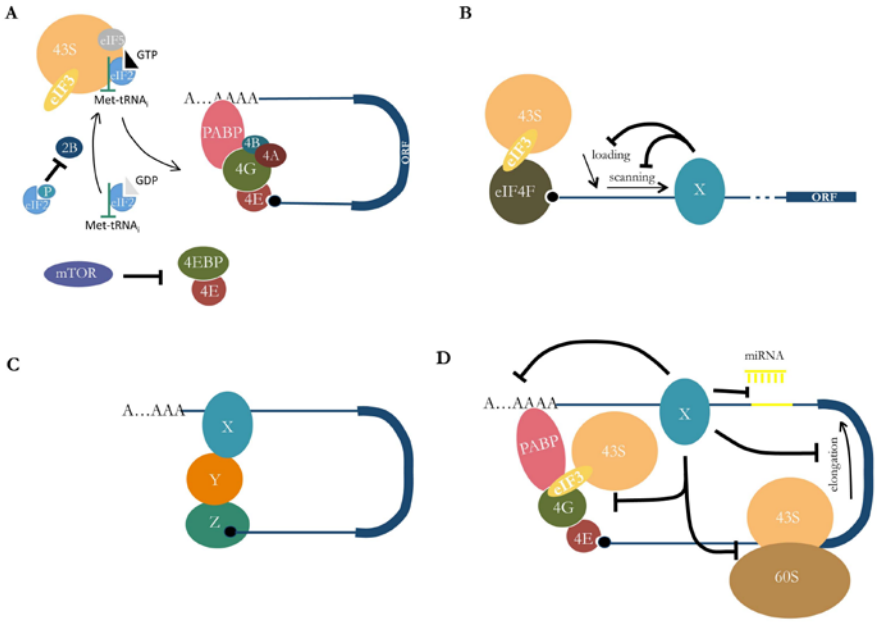


Figure 4. Translational regulation by sequence specific mRNA binding proteins. (A) Principles of global translational regulation (B) Proteins binding to the 5'UTR of the mRNA generally affect loading of the 43S onto the mRNA or the scanning process. (C) Principles of translational regulation mediated by many 3' UTR binding proteins. (D) Other mechanisms of translational regulation by 3' UTR binding proteins.

4. Dosage compensation

Sex can be determined by the environment or be genetically controlled. In species where sex is genetically determined, most commonly either males or females are heterogametic (female/male: XX/XY, XX/XO, ZW/ZZ). The pervading model for the evolution of sex chromosomes posits that they originate from a pair of autosomal chromosomes that acquired a sexual locus and then ceased to recombine. Thus, the two sex chromosomes evolved independently, and acquired different size and morphology. Recombination is still possible for the homomorphic chromosome (X or Z), but not for the heteromorphic one (Y or W), leading to the accumulation of deleterious mutations, pseudogenization and gene loss. The homomorphic gene harbors several genes that are important for both sexes, however one sex contains only one copy of this gene, while the other contains two. This leads to an imbalance of X-linked gene expression between males and females, but also between sex-linked and autosomal genes. Deviation from a balanced gene dosage affects cell physiology and ultimately organism fitness. Thus, a selective pressure exists to ensure balanced X-linked gene expression between males and females, leading to dosage compensation processes.

Dosage compensation has been extensively studied in three model systems: worms (*Caenorhabditis elegans*), flies (*Drosophila melanogaster*) and mammals (*Mus musculus*). In the three systems, males are heterogametic, but they use different strategies to equalize X-linked gene expression between females and males (Figure 5). In *C. elegans* worms, which can be hermaphrodite (XX) or males (X0), and in mammals (XX/XY), the X chromosomes of the homogametic sex are targeted. In worms, transcription of both X chromosomes of the hermaphrodite is repressed by half. In mammals, one of the X chromosomes of the female is randomly inactivated. This would equalize X-linked gene expression between males and females, but it does not solve the X chromosome autosomy of the male. To achieve this balance, in both systems the active X chromosomes are two-fold upregulated

(reviewed in Graindorge et al., 2011). In *D. melanogaster* (XX/XY) the single X chromosome of the male is hyper-transcribed by two-fold. Despite these differences in general strategies used between organisms, the molecular mechanisms employed are quite similar: i) global modification of chromatin states affects the transcriptional output of most genes of the target chromosome, ii) the sex chromosome is distinguished from autosomes by DNA sequence recognition and long non-coding RNA (lncRNA) action, and iii) dosage compensation is restricted to one of the sexes, because the responsible factors are regulated in a sex-specific fashion. Since the molecular mechanism of translational control analyzed in this thesis occurs during dosage compensation in *D. melanogaster*, we will only describe in detail the process in this organism.

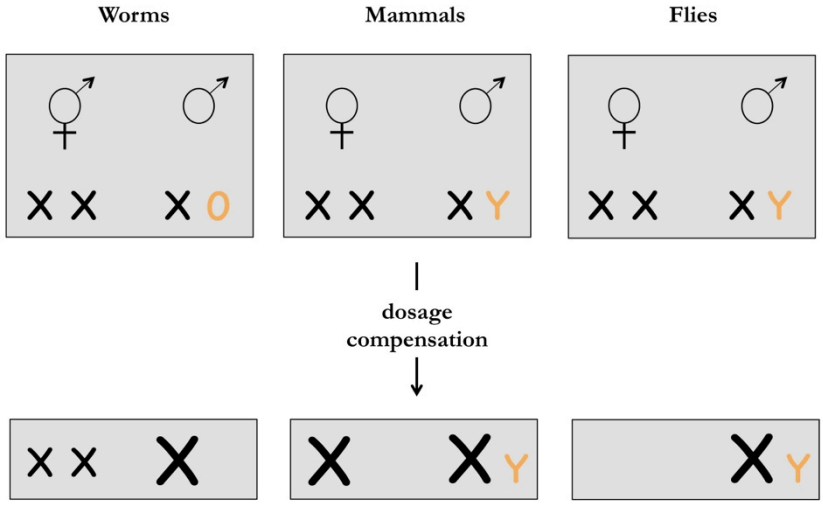


Figure 5. Equalization of X-linked gene transcription between females and males is achieved in worms and mammals by the two-fold overexpression of the X chromosomes, followed in worms by the two-fold down-regulation of both X-chromosomes of the hermaphrodite, and in mammals by the silencing of one X in the female. In flies it is achieved by a 2-fold upregulation of the single X-chromosome in the male.

4.1. Dosage compensation in *Drosophila melanogaster*

In *D. melanogaster*, dosage compensation is achieved by the Dosage Compensation Complex (DCC) (also named male-specific lethal or MSL complex) that binds to hundreds of sites in the single male X-chromosome to induce a two-fold hyper-transcription (Gelbart and Kuroda, 2009). The DCC is formed by five proteins MSL1, MSL2, MSL3, the histone acetylase MOF (males absent on the first), the RNA helicase MLE (maleless) and two long-non-conding RNAs roX1 and roX2 (RNA on the X) (Figure 6). MSL2 is the limiting subunit of the DCC. Binding of MSL2 stabilizes MSL1, which acts as a scaffolding protein to mediate the integration of MSL3 and MOF into the complex. MSL2, in cooperation with MSL1 and MLE, also activates transcription of roX1 and roX2, which are functionally redundant. They are required for DCC formation and spreading on the X chromosome (reviewed in Conrad and Akhtar 2014). RoX1 and roX2 are structurally modified by the RNA chaperone Upstream of N-ras (UNR) which promotes their interaction with MLE (Militti et al., 2014).

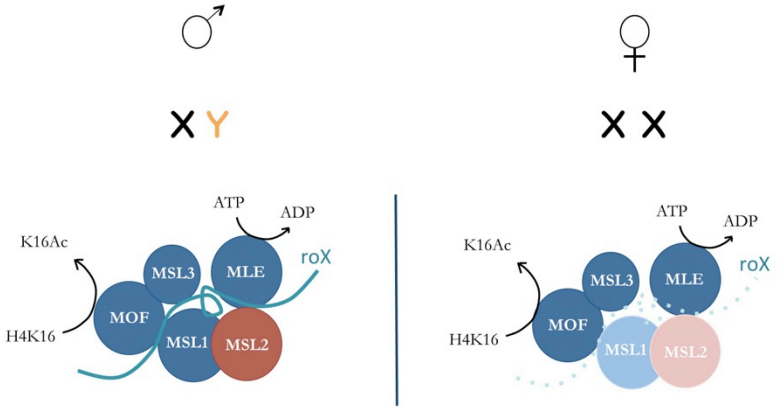


Figure 6. The dosage compensation complex in Drosophila melanogaster. The illustration represents the DCC and the enzymatic activities connected to its components. The complex is formed by 5 proteins and two long non-coding RNAs (MSL1, MSL2, MSL3, MLE, MOF, roX1 and roX2) and covers the entire male X-chromosome. In female flies, the DCC fails to assemble due to the lack of the MSL2 protein. Absence of MSL2 in turn destabilizes MSL1 and the roX RNAs.

The binding of the DCC to the X occurs in a two-step manner. Initial binding occurs at around 200 "high affinity sites" (HAS), or "chromatin entry sites" (CES) distributed uniformly on the X chromosome. These sites are less than 250 bp long, and when transferred into autosomal loci, they still recruit the DCC. A conserved GA-rich motif (21 bp) present at the HAS has been proposed to mediate DCC binding. However, other DNA or chromatin features must be required, since this motif is only enriched 2-fold on the X-chromosome with respect to autosomes, and it is also present in non-HAS sites in the X chromosome. Once bound to the HAS, the DCC associates with active genes in the vicinity by recognition of the H3K36me3 mark, which is associated with highly transcribed genes. These sites are called "low affinity sites". Clustering of the DCC at HAS facilitates, in a three

dimensional conformation, its binding to the low affinity sites (reviewed in Conrad and Akhtar 2012).

Once bound, the DCC induces acetylation of histone H4 at lysine 16, to loosen DNA-nucleosome association and increase the processivity of RNA polymerase II. This modification is catalyzed by the acetyl-transferase MOF, and allows the two-fold hyper-transcription which is stably maintained thereafter (Smith et al., 2001, Larschan et al., 2011, Keller and Akhtar, 2015).

Importantly, in females dosage compensation does not occur because the expression of MSL2 is inhibited. Repression of MSL2 occurs at several post-transcriptional steps of gene expression that are orchestrated by the female specific protein Sex-lethal (SXL).

5. Regulation of MSL2 expression by SXL

Repression of MSL2 expression is crucial for female survival (Kelley et al., 1995). Several super-imposed post-transcriptional mechanisms ensure MSL2 inhibition: splicing, nucleo-cytoplasmic transport and translation. All these mechanisms are orchestrated by SXL, which is itself under a fine control of transcription and splicing that ensures SXL expression only in females flies.

5.1. SXL protein

SXL is a 35 kDa RNA binding protein that functions as the sex determination switch in flies. SXL is expressed very early in development and directs female traits and behavior, and repression of dosage compensation (Salz and Erickson, 2010). Loss of SXL function in XX animals results in female lethality, and aberrant expression of SXL in males leads to male lethality as well. Thus, a fine sex-specific control of SXL expression is required in flies.

SXL is formed by two RNA recognition motifs (RRM1 and RRM2), and an N-terminal domain that is enriched in glycine and asparagine residues (Figure

7A). The RRMs mediate RNA binding and interaction with SXL partners. SXL localizes both in the nucleus and in the cytoplasm, and binds long poly(U) stretches that can be interrupted by guanosines. This consensus sequence is very common in the *Drosophila* genome, suggesting that SXL might require a specific context or protein partners to bind to its target mRNAs.

Figure 7. Schematic representation of SXL (A) and msl-2 mRNA (B). The domains of SXL, and its binding sites on msl2 mRNA (A-F) are indicated.

Regulation of SXL expression can be divided in two phases: initiation and maintenance. The initiation phase leads to transcription of SXL from the "establishment" promoter *SxlPe*, while the maintenance phase is governed by a SXL protein expressed from the "maintenance" promoter *SxlPm*. *SxlPe* promoter is under the control of four X-encoded proteins, collectively called X-linked signal elements (XSE). These four proteins are SCUTE, SISA, RUNT and UNPAIRED. *SxlPe* promoter responds to a threshold of XSE proteins concentration with respect to the concentration of proteins from autosomal and maternal origin. In the case of the female fly that contains two X chromosomes, this threshold is reached at syncytial cycle 12, and then shuts off early in cycle 14. However in males, it is never reached. At the beginning of cycle 14, the maintenance phase starts with expression from the *SxlPm*. The maintenance promoter is also controlled by XSE elements (SCUTE and RUNT) and by a maternally provided protein DA. A low expression of SXL from *SxlPm* also occurs in males, however, the transcripts

are non-functional because they contain a translation-terminating exon (exon 3). Skipping of exon 3 is mediated by SXL itself. Thus, in female embryos, the presence of SXL protein since early stages forces the skipping of exon 3, leading only to protein-encoding mRNAs, which establishes a positive autoregulatory loop that ensures expression of functional SXL all along development. SXL then promotes female development by regulating the alternative splicing of the downstream genes *transformer* (*tra*) and *doublesex* (*dsx*). Thus, the activation of this autoregulatory loop determines the sex of the fly (reviewed in (Salz and Erickson, 2010). SXL also regulates dosage compensation, by inhibiting the expression of MSL2 at several post-transcriptional steps. This important process ensures female development and viability.

5.2. Inhibition of *msl2* expression by SXL

Repression of *msl2* is achieved by SXL binding to several poly(U) stretches located in *msl2* UTRs: sites named A and B in the 5'UTR, and C, D, E and F in the 3'UTR (Figure 7B). As mentioned above, to ensure *msl2* repression, SXL regulates several steps of the gene expression cascade: splicing, nucleo-cytoplasmic transport and translation. During splicing, SXL promotes the retention of a facultative intron in the 5'UTR of *msl2* pre-mRNA by binding to sites A and B that flank this intron. At the 5' splice site, SXL competes with the binding of the splicing activator TIA-1, while at the 3' splice site SXL blocks the interaction with U2AF. Since binding of TIA-1 and U2AF are early steps of intron recognition, SXL inhibits splicing at its first steps (Merendino et al., 1999; Förch et al., 2001; Lallena et al., 2002) (Figure 8, step 1). The intron contains binding sites for the protein Held-out-wings (HOW), thus preservation of the intron allows for HOW binding. HOW belongs to the STAR family of RNA binding proteins. It interacts directly with SXL to facilitate the retention of *msl*2 pre-mRNA in the nucleus (Graindorge et al., 2013) (Figure 8, step 2). Since SXL and the 5' UTR intron are female-

specific, this function of HOW only occurs in female flies. Retention of *msl2* pre-mRNA in the nucleus is, however, not sufficient to ensure MSL2 repression, as a significant fraction of *msl2* transcripts escape nuclear retention. Thus, once in the cytoplasm, translation of *msl2* transcripts must be inhibited. The current model of SXL-mediated *msl2* translational repression poses that by binding to the 5' and the 3'UTR of *msl2*, SXL inhibits two early steps of translation initiation. SXL binding to the 3' UTR inhibits the initial recruitment of the 43S complex to the mRNA, whereas SXL binding to the 5' UTR blocks scanning of 43S complexes in search of the *msl2* AUG (Figure 8, steps 3 and 4) (Kelley et al., 1997, Bashaw and Baker, 1997, Gebauer et al., 1999, Gebauer et al., 2003, Beckmann et al., 2005, Medenbach et al., 2011, Gebauer et al., 1998). SXL binding to only one of the UTRs results in significant translational repression, but binding to both UTRs provides synergistic repression, indicating that although both mechanisms are independent, they are intimately interconnected.

Repression by SXL binding to the 5' UTR involves the recognition of an uORF. How SXL promotes recognition of this uORF is unknown; however the mechanism may involve specific interactions with the translational machinery, since binding of the protein PTB to *msl2* 5' UTR does not support repression. Therefore, the mechanism is unlikely to function by steric hindrance (Medenbach et al., 2011). Repression by SXL binding to the 3' UTR requires the co-factor Upstream-of-Nras (UNR) (Abaza et al., 2006, Duncan et al., 2006). UNR is a conserved, cytoplasmic protein essential for development that contains five beta-barrel structures called cold-shock domains (CSDs) that mediate binding to single-stranded nucleic acids. UNR binding to *msl2* requires interaction with SXL; thus, even though UNR is present in males, it cannot bind *msl2* mRNA nor repress its translation (Abaza et al., 2006; Abaza and Gebauer, 2008). Binding of SXL and UNR to the 3' UTR of *msl2* is cooperative, and the recent three-dimensional structure of the minimal domains involved in complex formation has revealed the

reasons for this cooperativity. First, a molecular zipper is formed in the interphase between SXL, UNR and *msl2* where residues of the three components are extensively intertwined; second, formation of this zipper changes the conformation of the RNA, which is now able to establish non-canonical contacts with the SXL RRM1 α-helix 1, thereby fixing SXL to the mRNA (Hennig et al., 2014). UNR bound to the 3' UTR contacts PABP, but the role of PABP in *msl2* regulation remains unclear (Duncan et al., 2009). The targets of the SXL mRNP repressor complexes in the translational machinery, thus, have remained unknown.

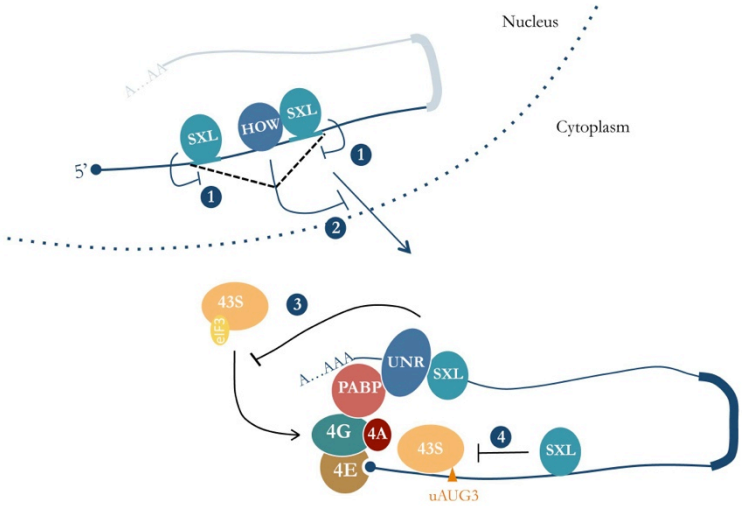


Figure 8. Current model for *msl2* mRNA regulation.

In the nucleus, SXL binding to the poly(U) stretches in the 5'UTR blocks splicing of a facultative intron leading to preservation of SXL and HOW binding sites (1). Subsequently HOW and SXL block the nuclear export of *msl2* transcripts (2). In the cytoplasm, additional factors cooperate with SXL to repress translation from the 5' and 3' UTRs. SXL bound to the 3' UTR inhibits the recruitment of the 43S PIC to the mRNA (3), whereas SXL bound to the 5' UTR stalls the scanning of 43S complexes in a process that involves the recognition of an uAUG (uAUG3) (4).

6. Objectives

The objective of this thesis is to study the mechanisms of translational regulation of *msl2* mRNA, and to identify the SXL targets in the translational machinery. *Msl2* translation requires the cap and the polyA tail, suggesting that *msl2* might follow the standard mechanism of translation initiation that involves the chain of interactions 7mG-eIF4E-eIF4G-eIF3-43S. However, SXL-mediated repression of *msl2* translation is efficient in the absence of both structures, indicating that the targets of SXL in the translational machinery can be recruited to the mRNA by an alternative mechanism as well (Gebauer et al., 1999; Gebauer et al., 2003). To get insight into how translation factors promote translation of *msl2* mRNA and are repressed by SXL, we have set two objectives:

1) Define the alternative, cap-independent recruitment of the ribosome to the mRNA, and identify the responsible factor(s) (Results Part 1, and Annex II).
2) Obtain and characterize new SXL mutants (Results Part 2).

The results of this thesis have allowed us to build an integrated model for translational repression of *msl2* via both the 5' and 3' UTRs.

II. RESULTS

1. eIF3d, a target of the *msl2* 3' UTR repressor complex, binds to *msl2* 5' UTR

eIF3d, a target of the *msl2* 3' UTR repressor complex, binds to *msl2* 5' UTR

Marina García-Beyaert and Fátima Gebauer[1]. (*Manuscript in preparation).*

Gene Regulation, Stem Cells and Cancer Programme, Centre for Genomic Regulation (CRG), Dr Aiguader 88, 08003-Barcelona, Spain

Universitat Pompeu Fabra (UPF), 08003-Barcelona, Spain

(1) Corresponding author:

Tel: +34-93 3160120

Fax: +34-93 3969983

E-mail: fatima.gebauer@crg.eu

Running title: eFI3d binds to *msl2* 5' UTR

Keywords: Translation, msl2, SXL, eIF3d, cap-independent translation

ABSTRACT

Translational repression of *msl2* mRNA is an important step in the regulation of X-chromosome dosage compensation in the fruitfly *Drosophila melanogaster*. Repression is orchestrated by SXL, which binds to both UTRs of the transcript. SXL bound to the 3'UTR inhibits ribosomal recruitment, a process mediated by the m^7G cap structure in most mRNAs. While translation of *msl2* requires the cap, repression by SXL is cap-independent, suggesting that SXL can block alternative modes of ribosomal recruitment to *msl2* mRNA. Mutational analysis indicate that this cap-independent property is conferred by the 5' UTR. In addition, cap-independent repression is sensitive to Hrp48 binding, As Hrp48 is a component of the 3' UTR repressor complex that targets eIF3d, these results suggest that eIF3d recruitment may explain cap-independent repression. Indeed, RNA IP analyses show that eIF3d can be specifically recruited to *msl2* 5' UTR in the absence of the cap structure. Our results reveal alternative pathways of translation initiation factor recruitment to the mRNA, and show that these can be targeted by RNA binding proteins for translational regulation.

INTRODUCTION

Regulation of mRNA translation is crucial for embryonic development and adult cell physiology (reviewed in Sonenberg and Hinnebusch, 2007, de Moor et al., 2005, Gebauer and Hentze, 2004, Kong and Lasko, 2012). A paradigmatic example of translational regulation during development of *Drosophila melanogaster* leads to sex-specific expression of MSL2, a rate limiting component of the X-chromosome dosage compensation complex (DCC). In male flies (XY), the DCC binds to hundreds of sites along the single X-chromosome and induces a 2-fold hypertranscription to equate the dosage of X-linked genes to that of females (XX) (Conrad and Akhtar, 2011, Graindorge et al., 2011). In females, dosage compensation is repressed via the inhibition of *msl2* expression. Both expression of *msl2* in males and its repression in females are essential for fly viability (Belote and Lucchesi, 1980, Kelley et al., 1995). Repression of *msl2* in females is orchestrated by the female-specific RNA binding protein Sex-lethal (SXL) and occurs at several steps of the post-transcriptional regulation cascade: mRNA nucleo-cytoplasmic transport, splicing and translation. SXL binds to uridine stretches on the 5' and 3' Untranslated Regions (UTRs) of *msl2* pre-mRNA. SXL bound to the 5' UTR induces the retention of a facultative intron (Merendino et al., 1999, Förch et al., 2001), and promotes nuclear retention of *msl2* transcripts (Graindorge et al., 2013). Once in the cytoplasm, SXL represses translation following a double block mechanism: SXL bound to the 3'UTR recruits Upstream of N-ras (UNR) and inhibits ribosome recruitment, while SXL bound to the retained intron in the 5' UTR inhibits ribosomal scanning by a mechanism that involves recognition of an upstream AUG (Beckmann et al., 2005, Abaza et al., 2006, Duncan et al., 2009, Duncan et al., 2006). PABP and eIF3 are required for efficient ribosome recruitment to the mRNA (reviewed in Valásek, 2012). During cap-dependent translation initiation, the 5' m^7GpppG cap structure is

recognized by the complex eIF4F, composed of eIF4E (which binds directly to the cap), eIF4G and eIF4A. PABP binds to the poly(A) tail and contacts eIF4G, increasing the affinity of eIF4F for the cap and favoring the formation of a closed-loop mRNA configuration (Kahvejian et al., 2005). eIF3, together with other initiation factors, binds to the small ribosomal subunit, and contacts cap-bound eIF4G to promote the recruitment of the small ribosomal subunit to the mRNA (Valásek, 2012, Sonenberg and Hinnebusch, 2009, Jackson et al., 2010).

Consistent with a cap-dependent mechanism of translation initiation, the cap and the poly(A) tail are necessary for *msl2* translation. Paradoxically, however, both of these structures are largely dispensable for *msl2* translational repression (Gebauer et al., 1999; Gebauer et al., 2003). This suggests that initiation factors different to those that directly recognize the cap and the poly(A) tail are targeted for regulation, and/or that the responsible factors can be recruited to *msl2* mRNA in additional ways. To address this conundrum, we have analyzed the requirements for cap-independent repression of *msl2* mRNA. We show that the 5' UTR of *msl2* confers cap-independent repression. Mutational analyses indicate that a feature in the 5' UTR, but not the length, is responsible for regulation. Furthermore, binding analysis indicate that eIF3d recognizes *msl2* 5' UTR independently of the cap structure. These results suggest that direct eIF3d binding to the 5' UTR of *msl2* supports ribosome recruitment, and that this binding may be targeted for regulation.

RESULTS

The 5' UTR of *msl2* confers cap-independent repression

The *msl2* transcript contains long 5' and 3'UTRs (626 and 1047 nucleotides [nt], respectively) with SXL binding sites designated A-F (Figure 1A). The minimal sequences required for SXL-mediated repression have been delimited to 69 nt in the 5' UTR containing SXL binding site B, and 46 nt in the 3' UTR containing sites E and F (BLEF, Figure 1A) (Gebauer et al., 2003).

A reporter containing the full length *msl2* UTRs fused to the Firefly luciferase open reading frame (mLm) was shown to be repressed by SXL in a cap-independent fashion (Gebauer et al., 2003). To analyze cap-independent repression, we first recapitulated these results and additionally compared repression of the full-length reporter mlm to that of the minimal construct BLEF in the presence or absence of a functional cap-structure. The functional cap consists of m^7GpppG (thereof termed G-cap) and can bind eIF4E, while the non-functional ApppG cap (thereof termed A-cap) fails to bind eIF4E and, therefore, does not support cap-dependent translation. The stability of mRNAs with G- and A- caps in a *Drosophila* cell free translation system is similar (Gebauer et al., 1999). G- or A- capped mRNA reporters were added to a *Drosophila* in vitro translation system in the presence of increasing amounts of a recombinant SXL fragment (dRBD4) which is fully functional in translational repression (Grskovic et al., 2003). As previously reported, repression of mLm was efficient in the presence and absence of a functional cap (Figure 1B, red and pink lines). Surprisingly, however, repression of BLEF with an A-cap structure was impaired compared to that of BLEF with a G-cap (Figure 1B, dark and light blue lines). We define "cap-dependency" as the distance between the A- and G- repression lines for any

given construct (Δ translation (A;G)). This distance is constant at all SXL/RNA ratios along the translation curve (see Figure 1B, left panel). The cap-dependency of BLEF is, thus, higher than that of mLm (Figure 1B, right panel). Importantly, the translation efficiencies of mRNAs containing the same 5'-end structure are similar (Fig 1C), indicating that the effects observed on repression are not due to differences on basal translation. These results suggest that cap-independent repression is conferred by an RNA feature that has been lost in the BLEF construct.

To determine where this RNA feature resides, we performed UTR swapping experiments (Figure 2). For simplicity, we represent the cap-dependency of each construct at SXL/RNA molar ratio 40, referred to that of BLEF. Constructs with full-length *msl2* 5'UTR showed cap-independent repression regardless of the 3' UTR (Figure 2B, compare BLm and mLEF), indicating that the RNA feature responsible for cap-independent repression resides on the 5'UTR.

Notably, cap-independent repression was maintained in the absence of SXL binding sites in the 5' UTR (Figure 2B, compare 5'mutLm, 5'mutLEF and BmutLEF constructs). In these constructs, only the 3' UTR-mediated mechanism is operative. Because 3' UTR-mediated repression involves inhibition of ribosome recruitment (Gebauer et al., 2003, Beckmann et al., 2005), these results imply that a factor required for ribosome recruitment can associate to the full length 5' UTR independently of the cap structure and can be targeted for regulation.

Features of the 5' UTR responsable for cap-independent repression

We next wondered whether the feature that confers cap-independent repression was simply the length of the 5'UTR, as such length has been

reduced about 10-fold in the minimal BLEF construct. Substituting the 5' UTR of *msl2* by one of similar length from an unrelated transcript (Bruno mRNA, 521 nt) abrogated cap-independent repression (Figure 3, BrunoLEF). This result indicates that an RNA sequence or structural feature, but not the length, explains cap-independent repression.

In addition to two SXL binding sites, the 5' UTR of *msl2* harbors the splice sites of a facultative intron and three upstream AUGs (uAUG) (Figure 3A). uAUGs have been shown to modulate translation initiation (Geballe and Sachs 2000) and *msl2* uAUG3 participates in repression mediated by the 5' UTR (Medenbach et al., 2011). To test if these sequence features were responsible for cap-independent repression, we analyzed the repression of constructs containing mutated splice sites or uAUGs. Mutation of these features did not affect cap-independent repression (Figure 3B).

Next, we analyzed the contribution of different regions of the *msl2* 5'UTR by transferring these regions to the unresponsive BruLEF construct. We first transferred the 5'UTR central region, but observed no recovery of cap-independent repression (Figure 3B, mut central construct). Neither additional transfer of the distal region (mut distal) nor transfer of the proximal third alone (mut proximal) restored cap-independent repression (Figure 3B). We conclude that none of those regions alone is responsible for cap-independent repression, but that it is rather a combination of sequences from the *msl2* 5'UTR.

eIF3d binds to the 5' UTR of *msl2* and contributes to cap-independent repression

To gain insight into factors binding to the 5' UTR in a cap-independent fashion which could be targeted by the repressor complex assembled at the 3' UTR of *msl2*, we focused on eIF3d. Our laboratory has recently shown that eIF3d is targeted by a component of the 3' UTR repressor complex, Hrp48, for efficient *msl2* repression (Szóstak et al., *submitted*). Interestingly, eIF3d is a non-core subunit of eIF3 that is not required for general translation in yeast nor *Drosophila*, but is specifically required for *msl2* translation and for efficient SXL-mediated repression (Szóstak et al., *submitted*). To test if eIF3d could play a role in cap-independent repression, we monitored the translational repression profile of a construct lacking the Hrp48 binding site (Figure 4A). Consistent with previous results, G-capped transcripts that could not bind Hrp48 were repressed less efficiently (left panel). Remarkably, a similar effect was observed when the transcripts contained an A-cap (right panel), indicating that the contribution of Hrp48 to the 3'-mediated mechanism persists in the absence of the cap structure. These results suggest that eIF3d contributes to cap-independent repression. A possible explanation for these results is that, in addition to being recruited through the cap, eIF3d can recognize directly the 5' UTR of *msl2* and attract the small ribosomal subunit in a cap-independent fashion, yielding *msl2* particularly dependent on eIF3d for translation. We therefore tested whether eIF3d could bind to the 5' UTR of *msl2*. We expressed FLAG-tagged eIF3d in rabbit reticulocytes, and added radiolabelled A-capped 5' UTRs from mRNAs that were repressed in a cap-independent (5'mutLEF) or cap-dependent (BLEF, BrunoLEF) manner. After immunoprecipitation with anti-FLAG beads, only the 5' UTR from the transcript repressed in a cap-independent fashion was enriched in the pellet (Figure 4B, left panel). Additional experiments using purified recombinant HA-tagged eIF3d

confirmed these results (Figure 4B, right panel). These data indicate that eIF3d can directly recognize the 5' UTR of *msl2*, contributing to translation in the absence of a cap structure, and to translational repression mediated by the 3' UTR.

DISCUSSION

During cap-dependent translation, factors required for translational activation are recruited through the cap structure (Sonenberg and Hinnebusch, 2009, Hinnebusch, 2014). *msl2* mRNA requires the cap structure to be translated efficiently, but its translation can be repressed in the absence of the cap (Gebauer et al., 2003). If *msl2* translation is cap-dependent, and translation factors are recruited through the cap structure, how can *msl2* translation be repressed without the cap? The mechanism underlying this apparent contradiction is analyzed in this work.

Repression mediated by SXL bound only to the 3' UTR occurred efficiently in the absence of the cap (Fig. 2). This result was important, because repression mediated by the 3' UTR SXL complex was shown to inhibit the recruitment of the ribosome to the mRNA (Gebauer et al., 2003, Beckmann et al., 2005). This result implied that inhibition of ribosome recruitment could occur in the absence of the cap and, therefore, an alternative non cap-mediated ribosome recruitment mechanism should operate on *msl2* mRNA. Mutational analysis showed that, for repression to occur in the absence of the cap, an RNA sequence or structural feature in *msl2* 5' UTR must be present (Figures 2 and 3). Despite intense efforts to elucidate this feature, we were unable to nail it down. Only the full length 5' UTR of *msl2*, but not short variants nor a long 5' UTR from an unrelated message, could support SXL-mediated repression in the absence of a cap structure. We hypothesize

that a structure resulting from the interaction between different parts of *msl2* 5' UTR might promote recruitment of a factor targeted during cap-independent translation. Such mechanism of initiation factor recruitment occurs at internal ribosome entry sites (IRESs). IRESs were first discovered as complex structures in the 5' UTRs of viral mRNAs that allowed direct recruitment of initiation factors and ribosomes and efficient translation in the absence of a cap (reviewed in Jackson et al., 2010, Kieft, 2008). IRESs were subsequently discovered in cellular mRNAs efficiently translated under conditions of eIF4E inactivation (Xue et al., 2015). Although the presence of IRES activity in the 5' UTR of *msl2* is in principle possible, we believe it is unlikely because translation of *msl2* is strongly cap-dependent (Gebauer et al., 2003; see also values on the Y axis of Figure 1C, where it can be observed that G-capped mRNAs are translated with a 50-fold higher efficiency than A-capped mRNAs). Other non-IRES, cap-independent modes of translation have been reported, as well as non-conventional forms of cap-dependent translation, suggesting that translation initiation mechanisms are more diverse and complex than previously anticipated (reviewed in Shatsky et al., 2014). For instance, eIF4G lacking its eIF4E binding domain can promote translation of mRNAs that do not bear a cap nor an IRES structure in the 5' UTR (Ohlmann et al., 1995, De Gregorio et al., 1998) in a manner that depends on the RNA binding domains of eIF4G (Prévôt et al., 2003).

Our previous work showed that eIF3d is required for *msl2* translation, and is targeted by the 3' UTR repressor complex to inhibit translation. eIF3d is a non-core subunit of eIF3 whose depletion does not affect global translation in *Drosophila* S2 cells (Szóstak et al., *submitted*). The function of eIF3d in translation is poorly understood. However, the location of eIF3d near the mRNA exit channel in the ribosome (where eIF4G is also localized), and experiments showing its interaction with eIF4G and the relevance of this interaction for mRNA translation, strongly point to a role of eIF3d in the recruitment of the ribosome to the mRNA (Villa et al., 2013, Pisarev et al.,

2008, Sun et al., 2011, Erzberger et al., 2014, des Georges et al., 2015). We therefore tested whether eIF3d was the factor recruited to the 5' UTR of *msl2* and responsible for cap-independent repression. Indeed, eIF3d binds to the 5' UTR of *msl2* in the absence of the cap structure, but does not bind to the 5' UTR of transcripts repressed in a cap-dependent fashion (Figure 4B). Binding of eIF3d to the 5' UTR of *msl2* is consistent with a recent report showing that eIF3d, as well as eIF3 subunits a, b and g have the capacity to bind directly to the 5' UTR of certain transcripts in mammalian cells, although in these cases a cap structure was also present (Lee et al., 2015). The authors also found that eIF3 subunits only bind to ~3% of expressed transcripts, which is consistent with our results showing selective binding of eIF3d to *msl2* mRNA. This pattern differs from that of eIF4G, whose RNA binding capacity is rather general (Prévôt et al., 2003).

Our results are consistent with a model where eIF3d-bound ribosomes can be recruited to *msl2* mRNA in two ways, either through the cap structure or via the 5' UTR (Figure 5). As *msl2* is required for dosage compensation and male fly viability (Kelley et al., 1995), both modes of ribosome recruitment would ensure efficient translation of *msl2* even in situations of stress, when cap-dependent translation is compromised. Both modes of recruitment can be targeted in females by the SXL:UNR:Hrp48 repressor complex assembled at the 3' UTR of *msl2*. Targeting of eIF3d by the 3' UTR repressor complex might not be, however, the whole story, as G- and A-capped mRNAs are repressed to some extent in the absence of Hrp48 binding (Figure 4A). In this regard, PABP has been shown to participate in the 3' UTR-mediated mechanism. Future work may elucidate the interplay between eIF3d, PABP and other initiation factors in translation regulation of *msl2* mRNA.

MATERIALS AND METHODS

Constructs

The mlm and BLEF constructs containing the full-length and minimal *msl2* UTRs, respectively, have been published previously (Gebauer et al. 2003). The mLEF constructs was obtained by digestion of mLm with EcoNI and BglII to release the 3' UTR, and insertion of the minimal UTR in the same sites. The BLm construct was obtained in a similar fashion, by insertion of the full length *msl2* 3' UTR into the EcoNI and BglII sites of BLEF. The 5'mutLEF construct was obtained by exchanging a 453nt 5' UTR fragment of mLEF between restriction sites SmaI and NheI by a similar fragment containing mutated SXL binding sites A and B, obtained by digestion of the previously described 5'A+Bmut plasmid (Gebauer et al 1999). The ssmut construct was obtained by site directed mutagenesis of 5'mutLEF to mutate the splice sites. The aug3mut and aug123mut constructs were obtained by mutating the AUG codons to AUU by site directed mutagenesis of the 5'mutLEF plasmid. The BrunoLEF construct was obtained in several steps. First, the 5' UTR of Bruno mRNA was amplified by primer extension from RNA isolated from *Drosophila* embryos and cloned into the XhoI/BamHI sites of pBSK. This fragment was then inserted into the SacI and BamHI sites of mLEF. The mut proximal, central and distal constructs were obtained by insertion of the corresponding *msl2* 5' UTR fragments into the 5' UTR of BrunoLEF (in the case of mut proximal) or a BrunoLEF construct modified to contain BlpI and SacI sites (BruLEFmod, in the case of central and distal mutants). To obtain mut proximal, the fragment between BlpI and NcoI of 5mLEF was exchanged with the AflII-NcoI fragment of BruLEF. In the case of mut central, the 5'UTR of 5'mLEF was digested between BlpI and SacI and inserted between the same sites in BRuLEFmod. To obtain the mut distal construct we exchanged the BlpI-NCoI fragment of BruLEFmod with the fragment flanked by the same sites in 5mLEF.

The His-2HA-eIf3d plasmid was produced by insertion of the eIF3d cDNA, obtained from the PGADT7-eIF3d plasmid (a generous gift of Jan Medenbach), into the pETM14 vector using the Gibson cloning method (Gibson et al., 2009). Two HA tags were introduced at the N-terminus of the eIf3d sequence, separated by a linker sequence (TTACGCTATGGCCATGTACCCAT).

The FLAG-eIF3d plasmid was obtained by PCR amplification of eIF3d from the same PGADT7-eIF3d mentioned above, inserting a FLAG sequence in the N-terminus. This cDNA was then cloned into a pBSK vector containing a polyA tail, using the Gibson method.

RNA synthesis

RNA was synthesized by in vitro transcription as previously described (Gebauer et al., 1999). During transcription, either anti-reverse 5' m7GpppG (ARCAP) or ApppG cap analogs (KEDAR) were introduced to obtain G-capped or A-capped mRNAs, respectively. All mRNAs contained a poly(A) tail of 73 residues. 5' UTR RNAs used in RNA IP experiments were non-adenylated and contained A-caps. All RNAs used in the same experiment were synthesized and quantified in parallel.

In vitro translation

In vitro translation reactions in Drosophila embryo extracts were performed in a final volume of 10 µl as described (Gebauer et al., 1999). Given their lesser efficiency of translation, for A-capped RNAs the reaction was scaled 2.5 fold. mLm was used at a concentration of 3.2 ng/µl, and equimolar amounts of all other mRNAs were used. As an internal control, 0.2 ng/ul of Renilla luciferase mRNA were added. Increasing amounts of recombinant dRBD4 (a fully functional fragment of SXL) were added. Luciferase activity was measured using the Promega Dual Luciferase kit. Firefly luciferase was

corrected for co-translated Renilla expression, and plotted relative to translation in the absence of SXL.

Protein expression and purification

Recombinant His-dRBD4 (amino acids 122-354 of SXL) was prepared following the pET system user's manual (Novagen). Proteins were dialyzed in buffer D (20mM HEPES at pH 8.0, 20% glycerol, 1mM DTT, 0.01% NP-40, 0.2 mM EDTA). His-2HA-eIf3d was expressed in E. Coli using autoinduction culture medium (Novagen, 71491-5). The protein was purified by two passages in Nickel affinity Trap Fast Flow columns (GE Healthcare) followed by size exclusion chromatography using Superdex 200 columns (GE Healthcare). The protein was stored in a buffer containing 50mM Tris-HCl pH 7.4, 200 mM NaCl, 10 % Glycerol, 2mM DTT. Protein quality was assessed by Coomassie staining.

FLAG-eIF3d was expressed in the Rabbit Reticulocyte Lysate (RRL) Translation System (Promega), following the indications of the manufacturer.

RNA immunoprecipitation

When using the His-2HA-eIF3d protein, first the indicated A-capped and ^{32}P-labelled 5'UTRs were incubated in a typical translation reaction in a final volume of 22.5 µl. When indicated, 400 ng of recombinant HA-eIF3d were added, and the reactions were incubated for 30 min at 25ºC. 30ul of anti-HA Magnetic Beads slurry (Pierce) equilibrated in 1XNET buffer (50mM Tris-HCl pH7.5, 150mM NaCl, 0.1% NP40, 1mM EDTA) were then added. After incubation for 1 h at 4ºC, beads were washed once with 1ml of 1X NET. RNAs were extracted with Trizol (Invitrogen) and resolved in a 2% agarose gel and finally visualized using a phosphorimager.

For the assays with FLAG tagged eIF3d, 20 ul of anti-FLAG magnetic beads (Sigma) were first washed with 5 bead volumes of TBS1X and then

resuspended in 20 ul of TBS1X. Twenty ul of the RRL translation System, either programmed or not to express FLAG-eIF3d, were added to the beads. A-capped 5' UTRs were radiolabeled with ^{32}P and added to the beads. Samples were incubated for 1h at 4ºC in a shaking platform. Beads were then washed 4 times with 1 ml of TBS1X. RNAs were extracted with Trizol (Invitrogen) and visualized following the same procedure as indicated above.

ACKNOWLEDGEMENTS

We thank Jan Medenbach for kindly providing the PGADT7-eIF3d plasmid and the CRG-protein facility for protein production. M.G-B was supported by a fellowship from La Caixa Foundation. This work was supported by Spanish Ministry of Economy and Competitiveness MINECO and the European Regional Development Fund (ERDF) under the grant number BFU2012-37135 and Consolider grant CSD2009-00080. We acknowledge support of the Spanish Ministry of Economy and Competitiveness, 'Centro de Excelencia Severo Ochoa 2013-2017', SEV-2012-0208.

REFERENCES

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(3), pp. 380-9.

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(4), pp. 529-40.

Belote, J. M. and Lucchesi, J. C. (1980) 'Male-specific lethal mutations of Drosophila melanogaster', *Genetics,* 96(1), pp. 165-86.

Conrad, T. and Akhtar, A. (2011) 'Dosage compensation in Drosophila melanogaster: epigenetic fine-tuning of chromosome-wide transcription', *Nat Rev Genet,* 13(2), pp. 123-34.

De Gregorio, E., Preiss, T. and Hentze, M. W. (1998) 'Translational activation of uncapped mRNAs by the central part of human eIF4G is 5' end-dependent', *RNA,* 4(7), pp. 828-36.

de Moor, C. H., Meijer, H. and Lissenden, S. (2005) 'Mechanisms of translational control by the 3' UTR in development and differentiation', *Semin Cell Dev Biol,* 16(1), pp. 49-58.

des Georges, A., Dhote, V., Kuhn, L., Hellen, C. U., Pestova, T. V., Frank, J. and Hashem, Y. (2015) 'Structure of mammalian eIF3 in the context of the 43S preinitiation complex', *Nature.*

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(3), pp. 368-79.

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', *Mol Cell,* 36(4), pp. 571-82.

Erzberger, J. P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C. H., Cimermančič, P., Boehringer, D., Sali, A., Aebersold, R. and Ban, N. (2014) 'Molecular architecture of the 40S·eIF1·eIF3 translation initiation complex', *Cell,* 158(5), pp. 1123-35.

Förch, P., Merendino, L., Martínez, C. and Valcárcel, J. (2001) 'Modulation of msl-2 5' splice site recognition by Sex-lethal', *RNA,* 7(9), pp. 1185-91.

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(21), pp. 6146-54.

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', *Mol Cell,* 11(5), pp. 1397-404.

Gebauer, F. and Hentze, M. W. (2004) 'Molecular mechanisms of translational control', *Nat Rev Mol Cell Biol,* 5(10), pp. 827-35.

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. and Smith, H. O. (2009) 'Enzymatic assembly of DNA molecules up to several hundred kilobases', *Nat Methods,* 6(5), pp. 343-5.

Graindorge, A., Carré, C. and Gebauer, F. (2013) 'Sex-lethal promotes nuclear retention of *msl2* mRNA via interactions with the STAR protein HOW', *Genes Dev,* 27(12), pp. 1421-33.

Graindorge, A., Militti, C. and Gebauer, F. (2011) 'Posttranscriptional control of X-chromosome dosage compensation', *Wiley Interdiscip Rev RNA,* 2(4), pp. 534-45.

Hinnebusch, A. G. (2014) 'The scanning mechanism of eukaryotic translation initiation', *Annu Rev Biochem,* 83, pp. 779-812.

Jackson, R. J., Hellen, C. U. and Pestova, T. V. (2010) 'The mechanism of eukaryotic translation initiation and principles of its regulation', *Nat Rev Mol Cell Biol,* 11(2), pp. 113-27.

Kahvejian, A., Svitkin, Y. V., Sukarieh, R., M'Boutchou, M. N. and Sonenberg, N. (2005) 'Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms', *Genes Dev,* 19(1), pp. 104-13.

Kelley, R. L., Solovyeva, I., Lyman, L. M., Richman, R., Solovyev, V. and Kuroda, M. I. (1995) 'Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in Drosophila', *Cell,* 81(6), pp. 867-77.

Kieft, J. S. (2008) 'Viral IRES RNA structures and ribosome interactions', *Trends Biochem Sci,* 33(6), pp. 274-83.

Kong, J. and Lasko, P. (2012) 'Translational control in cellular and developmental processes', *Nat Rev Genet,* 13(6), pp. 383-94.

Lee, A. S., Kranzusch, P. J. and Cate, J. H. (2015) 'eIF3 targets cell-proliferation messenger RNAs for translational activation or repression', *Nature,* 522(7554), pp. 111-4.

Medenbach, J., Seiler, M. and Hentze, M. W. (2011) 'Translational control via protein-regulated upstream open reading frames', *Cell,* 145(6), pp. 902-13.

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(6763), pp. 838-41.

Ohlmann, T., Rau, M., Morley, S. J. and Pain, V. M. (1995) 'Proteolytic cleavage of initiation factor eIF-4 gamma in the reticulocyte lysate

inhibits translation of capped mRNAs but enhances that of uncapped mRNAs', *Nucleic Acids Res,* 23(3), pp. 334-40.

Pisarev, A. V., Kolupaeva, V. G., Yusupov, M. M., Hellen, C. U. and Pestova, T. V. (2008) 'Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes', *EMBO J,* 27(11), pp. 1609-21.

Prévôt, D., Décimo, D., Herbreteau, C. H., Roux, F., Garin, J., Darlix, J. L. and Ohlmann, T. (2003) 'Characterization of a novel RNA-binding region of eIF4GI critical for ribosomal scanning', *EMBO J,* 22(8), pp. 1909-21.

Shatsky, I. N., Dmitriev, S. E., Andreev, D. E. and Terenin, I. M. (2014) 'Transcriptome-wide studies uncover the diversity of modes of mRNA recruitment to eukaryotic ribosomes', *Crit Rev Biochem Mol Biol,* 49(2), pp. 164-77.

Sonenberg, N. and Hinnebusch, A. G. (2007) 'New modes of translational control in development, behavior, and disease', *Mol Cell,* 28(5), pp. 721-9.

Sonenberg, N. and Hinnebusch, A. G. (2009) 'Regulation of translation initiation in eukaryotes: mechanisms and biological targets', *Cell,* 136(4), pp. 731-45.

Sun, C., Todorovic, A., Querol-Audí, J., Bai, Y., Villa, N., Snyder, M., Ashchyan, J., Lewis, C. S., Hartland, A., Gradia, S., Fraser, C. S., Doudna, J. A., Nogales, E. and Cate, J. H. (2011) 'Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3)', *Proc Natl Acad Sci U S A,* 108(51), pp. 20473-8.

Valásek, L. S. (2012) "Ribozoomin'--translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs)', *Curr Protein Pept Sci,* 13(4), pp. 305-30.

Villa, N., Do, A., Hershey, J. W. and Fraser, C. S. (2013) 'Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to

promote mRNA recruitment to the ribosome', *J Biol Chem,* 288(46), pp. 32932-40.

Xue, S., Tian, S., Fujii, K., Kladwang, W., Das, R. and Barna, M. (2015) 'RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation', *Nature,* 517(7532), pp. 33-8.

FIGURE LEGENDS

Figure 1. An RNA element of *msl2* mRNA is required for cap-independent translational repression. (A) Schematic representation of constructs used in (B). SXL binding sites are designated from A to F. mLm contains the 5' and 3' UTRs of *msl2* mRNA and the Firefly luciferase open reading frame (ORF). BLEF is the minimal sequence required for SXL-mediated translational repression and only contains SXL binding sites B, E and F, as indicated. Numbers indicate the length of the 5' and 3' UTRs.

(B) An RNA element present in mlm and absent in BLEF is required for cap-independent translational repression. SXL-mediated repression was analyzed for construct BLEF and mlm, containing a G-cap or an A-cap. . A schematic representation of the functional G-cap and of the un-functional A-cap and their capacity to interact with the eIF4E-eIF4G complex is depicted. The cap type is indicated by an A- or G- at the beginning of the transcript name in the legend. *In vitro* translation assays were performed with increasing amounts of recombinant His-dRBD4 protein, a recombinant SXL protein that contains the minmal regions required for efficient translational repression. *Renilla* luciferase mRNA was co-translated as an internal control. *Firefly* luciferase was corrected for *Renilla* expression, and the data were plotted as the percentage of translation in the absence of SXL. A representation of the degree of cap-dependent translational repression is shown on the right panel for both mlm and BLEF transcripts. The degree of cap-dependent repression was calculated as the difference of translation at SXL/RNA ratio 40 between a construct containing an A-cap and the same construct with a G-cap at the same SXL/RNA ratio. The degree of cap-dependency for each construct is represented by a line in the translation curve.

(C) Translation is cap-dependent for both mRNAs and the differences in translational repression are not due to differences in basal translation. The

levels of basal translation correspond to the luciferase reads normalized by Renilla in the absence of SXL. Students t-Test Pval indicate no differences between mlm and BLEF translation levels when both containing the same cap-structure.

Figure 2. The 5'UTR of *msl2* confers cap-independent repression. In all figures, error bars represent the standard deviation from at least 3 independent experiments (Students t-Test Pval **<0.01, *** <0.001).
(A) Schematic representation of the constructs used in the translation assay in (B and C). mlm and BLEF are used to produce all the constructs bellow the dashed line.
(B) 3'-mediated repression alone recapitulates cap-independent repression, and relies on a sequence element located in the 5'UTR of *msl2* mRNA. The degree of cap-dependency is represented in percentage, taking BLEF as a reference. BLEF and mlm are used as controls for cap-dependent and cap-independent repression respectively.

Figure3. mRNA sequences in the first 400 nt of *msl2* 5'UTR confer the cap-independent repression. (A) Schematic representation of constructs used in (B). BLEF, 5'mutLEF and BrunoLEF are used to produce the constructs bellow the dashed line. The punctual difference between the 5'UTR sequences of 5'mutlEF and BLEF, besides the SXL binding sites, are indicated: three upstream AUG (uAUG) in orange and 2 splice sites (ss) in blue. None of the constructs contain SXL binding sites in 5'UTR, except BLEF used as control. Punctual elements that were mutated in the constrcuts bellow the dashed line are simply not represented. The length of the 5'UTRs that differs from that of 5'mLEF is indicated. The 3'UTR is the same for all constructs and contains only sites E and F, except for mut prox+central that exceptionally contains the 4 SXL binding sites (C, D, E and F).

(B) The percentage of cap-dependent repression is analyzed as in fig 2. For mut proximal, the degree of cap-dependency is represented relative to BrunoLEF.

Figure 4. eIF3d binds to the 5'UTR of *msl2* and contributes to cap-independent repression. Error bars represent the standard deviation from at least 3 independent experiments.

(A) The contribution of Hrp48 to the 3'-mediated repression persists in the absence of the cap structure. 3'-mediated repression is analyzed with constructs that contain (5'mutLEF) or lack (5'mutLEFHrmut) the Hrp48 binding site. The constructs used in the translation assay are depicted in the upper part of the panel. The analysis was performed with transcripts containing a functional G-cap or an un-functional A-cap as indicated by a G- or an A- at the beginning of the name of each transcript in the legend. The translation assays were performed as in Fig 1 (B).

(B) Schematic representation of the constructs whose 5'UTR was used in (C). The name of each 5'UTR is indicated in bold, followed in brackets by the name of its original construct, previously used in the translation assays. Their capacity to support cap-independent repression is indicated. Numbers refer to the length of the 5'UTR. For simplicity, the 3'UTR of the constructs is not shown.

(C) eIF3d binds to the 5'UTR of constructs that recapitulate the cap-independency but not to the 5'UTRs of those constructs that do not.

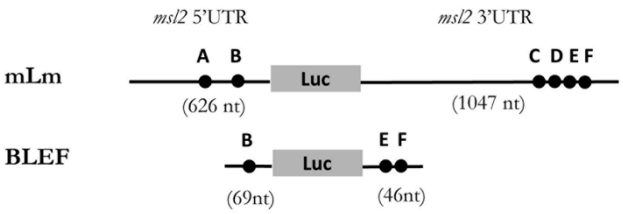
Both panels are RNA IPs against eFI3d, using radiolabeled, A-capped and non-polyadenylated 5'UTRs. RNAs were added on a molar ratio basis, according to their size. In the left panel the protein used is flag tagged eIF3d expressed in rabbit reticulocyte lysates. In the right panel a recombinant HA-eIF3d expressed in bacteria was used.

Figure 5. Schematic representation of the mechanism. The subunit d of the eIF3 complex (black letter *d* in the eIF3 orange circle) can be recruited, alone or as part of the eFI3 complex, to the 5'UTR of *msl2* through an interaction with eIF4G (brown circle), engaged with the cap binding protein eIF4E (green), or by interaction with an mRNA sequence in the 5'UTR (black arrows). The double mode of eIF3d recruitment may ensure ribosome recruitment to *msl2* in a cap-dependent and cap-independent mode. Hrp48 (purple), a co-factor of the 3'UTR SXL RNP repressor complex, targets eIF3d (and likely also other factors) to inhibit the recruitment of the ribosome in all conditions.

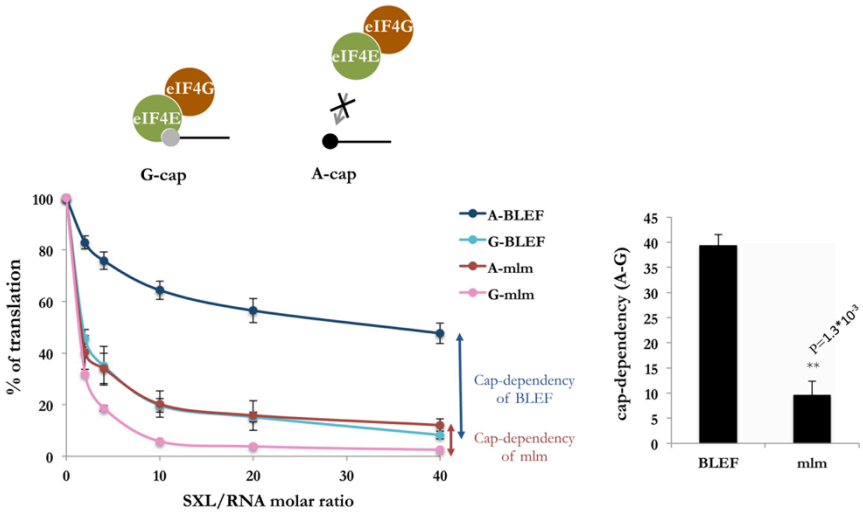
FIGURES

Figure 1

A



B



C

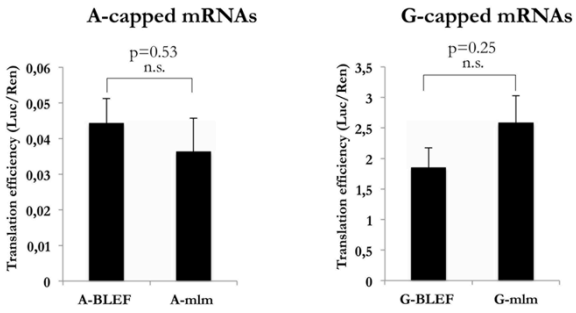
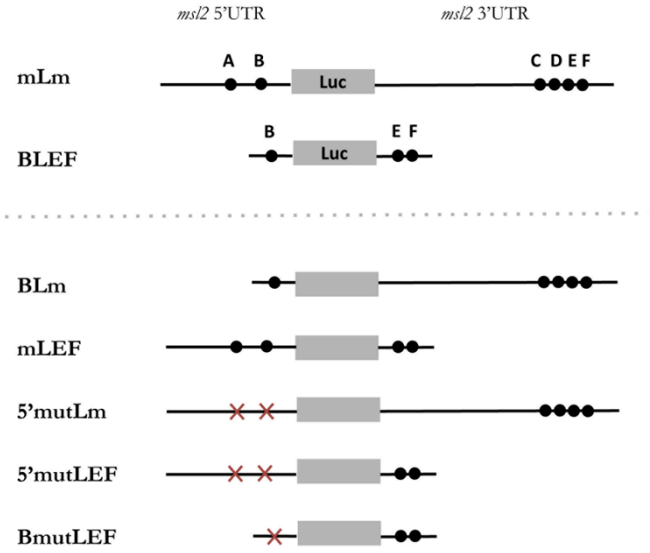


Figure 2

A


msl2 5'UTR
msl2 3'UTR
mLm
A B
Luc
C D E F
BLEF
B
Luc
E F
BLm
mLEF
5'mutLm
5'mutLEF
BmutLEF

B

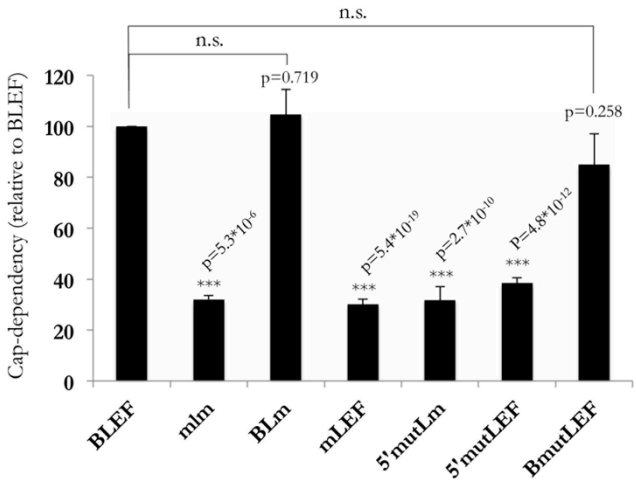
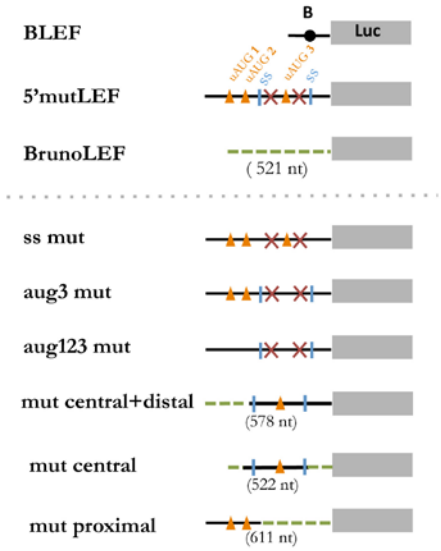

n.s.
n.s.
p=0.719
p=0.258
p=5.3*10-6
p=5.4*10-19
p=2.7*10-10
p=4.8*10-12
***
***
***
***
Cap-dependency (relative to BLEF)
120
100
80
60
40
20
0
BLEF
mlm
BLm
mLEF
5'mutLm
5'mutLEF
BmutLEF

Figure 3

A


B
BLEF
Luc
uAUG1
uAUG2
uAUG3
ss
ss
5'mutLEF
BrunoLEF
(521 nt)
ss mut
aug3 mut
aug123 mut
mut central+distal
(578 nt)
mut central
(522 nt)
mut proximal
(611 nt)

B

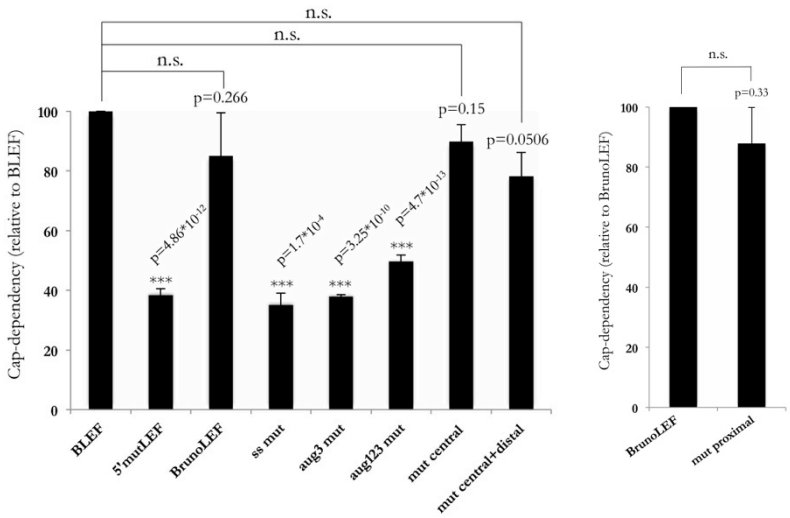
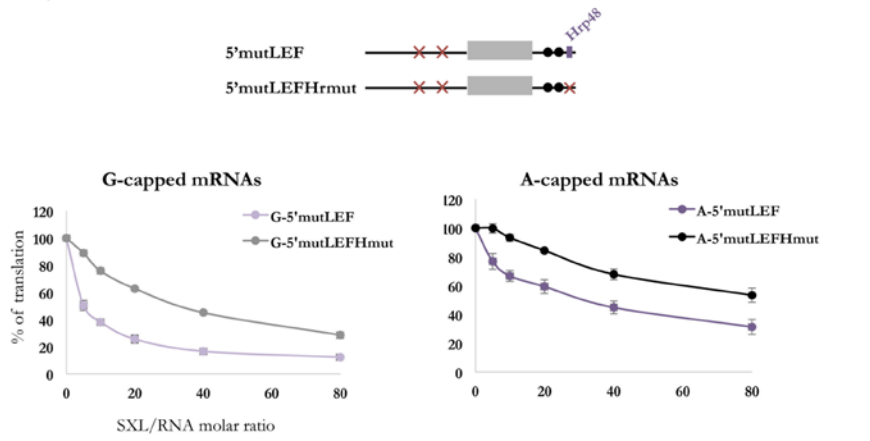
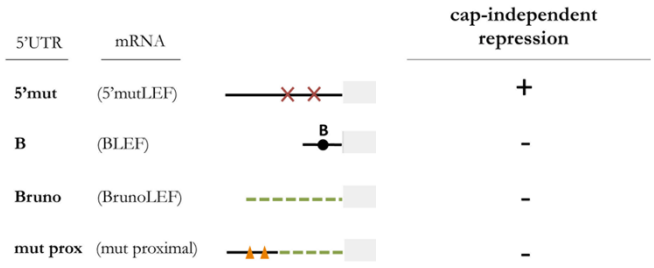

n.s.
n.s.
n.s.
p=0.266
p=0.15
p=0.0506
p=4.86*10-12
p=1.7*10-4
p=3.25*10-10
p=4.7*10-13
***
***
***
***
Cap-dependency (relative to BLEF)
100
80
60
40
20
0
BLEF
5'mutLEF
BrunoLEF
ss mut
aug3 mut
aug123 mut
mut central
mut central+distal
n.s.
p=0.33
Cap-dependency (relative to BrunoLEF)
100
80
60
40
20
0
BrunoLEF
mut proximal

Figure 4

A


Hrp48
5'mutLEF
5'mutLEFHrmut
G-capped mRNAs
G-5'mutLEF
G-5'mutLEFHmut
% of translation
SXL/RNA molar ratio
A-capped mRNAs
A-5'mutLEF
A-5'mutLEFHmut

B


5'UTR
mRNA
cap-independent repression
5'mut (5'mutLEF) +
B (BLEF) -
Bruno (BrunoLEF) -
mut prox (mut proximal) -

C

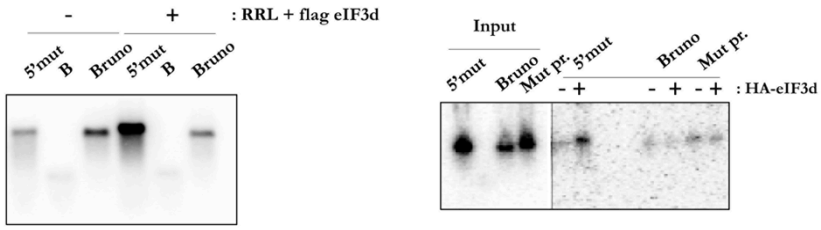
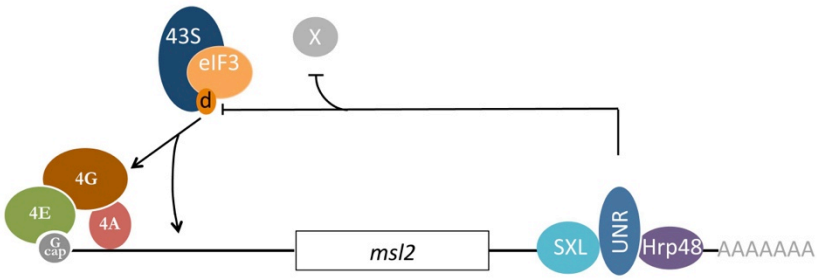

- + : RRL + flag eIF3d
5'mut B Bruno 5'mut B Bruno
Input
5'mut Bruno Mut pr. 5'mut Bruno Mut pr.
- + - + - + : HA-eIF3d

Figure 5



2. Insights into translational repression of *msl2* mRNA by dissection of relevant SXL residues.

In this part of the thesis, I describe my efforts to identify novel residues of SXL important for translational repression using information from the three-dimensional structure of the SXL:UNR:*msl2* complex that was recently described in the laboratory in collaboration with the group of Michael Sattler (TUM, Munich) (Henning et al, 2014). I describe the characterization of an important flexible loop and our fishing expedition to identify factors interacting with this loop. These results led to insights into 5' UTR-mediated repression by SXL, which are briefly discussed here and more extensively commented in the Discussion section.

2.1. Identification of novel SXL residues required for repression

As described in the introduction, SXL contains a GN-rich amino-terminal domain, followed by 2 RNA recognition motifs (RRM1 and RRM2) and a 60 amino-acid C-terminal region with no particular domain content (Figure 1A). A previous report identified RRM1 and RRM2 followed by a 7 amino acid extension as the minimal region required and sufficient for full *msl2* translational repression (Grskovic et al., 2003). This minimal SXL protein was referred to as dRBD4. Residues in the RRMs are highly conserved. Comparison of SXL proteins from *Drosophila melanogaster* (dSXL) and the distant relative *Musca domestica* (mSXL) show 89% amino-acid identity in this region. Despite this conservation, however, mSXL (or a fragment equivalent to dRBD4 referred to as mRBD4) cannot repress *msl2* translation (Grskovic et al, 2003). This implies that residues that are not conserved between the two proteins are important for translational repression. Domain swapping

experiments revealed that non-conserved residues in RRM1 are responsible (Grskovic et al., 2003) (Figure 1A). A structural analysis of dRBD4 bound to *msl2* EF RNA and UNR CSD1 uncovered the identity of residues in RRM1 relevant for *msl2*:SXL:UNR complex formation (Hennig et al., 2014) (Figure 1B). On the other hand, the RRM2 fragments of dSXL and mSXL are functionally interchangeable, indicating that conserved residues in RRM2 contribute to translational repression (Grskovic et al., 2003). Thus, we hypothesized that these conserved RRM2 residues may be important for the interaction of SXL with co-factors distinct from UNR, or with components of the translation machinery targeted during *msl2* translational repression.

Observing the *msl2*:dRBD4:CSD1 structure, we noticed two exposed loops in RRM2 (loop1 and loop2) whose residues are strictly conserved (Figures 1B and C). These loops are far away from the UNR and *msl2* interaction region, suggesting that they could be important for interaction with other factors. We thus mutated loop1 and loop2 as shown in Figure 1B, and tested the capacity of mutant GST-tagged dRBD4 proteins to repress *msl2* translation. Two loop1 mutants (loop1-S1 and S2) and one loop2 mutant (loop2 mut) were generated. Loop1-S1 and S2 were designed to change the sequence but not the overall charge of loop1, so that the binding sequence of a putative factor would be disrupted but the structure and charge of the loop would remain intact. For example, the positively charged arginines [R (+)] on the base of the loop (positions 1 and 7) remained untouched to avoid a possible disruption of the loop structure. The loop2 mutant was obtained by mutating two conserved amino acids: a hydrophobic isoleucine (I1) and a glutamate [E3 (-)] to alanine (A).

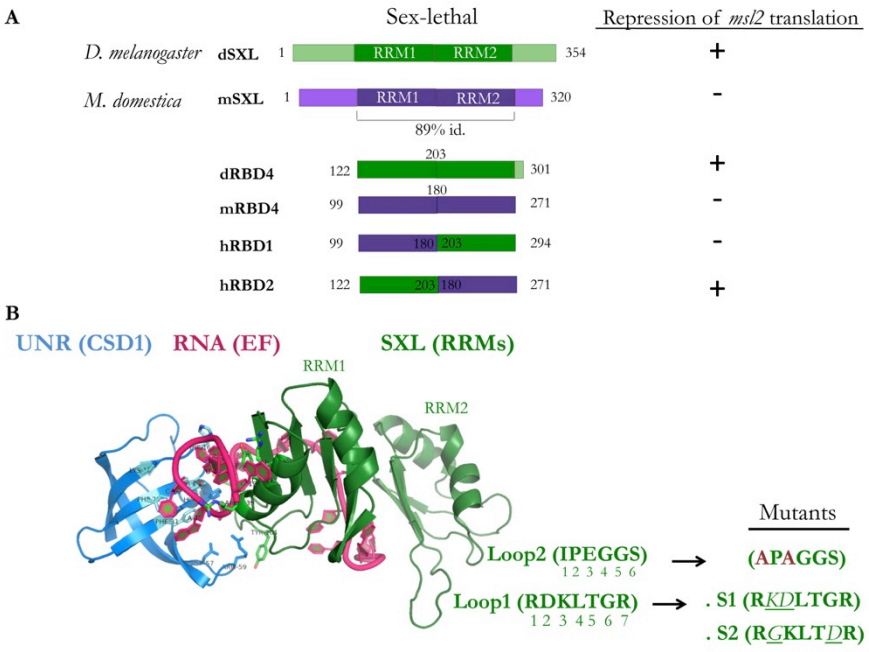

A
Sex-lethal
Repression of msl2 translation
D. melanogaster
dSXL
1
RRM1
RRM2
354
+
M. domestica
mSXL
1
RRM1
RRM2
320
-
89% id.
203
dRBD4
122
301
+
180
mRBD4
99
271
-
hRBD1
99
180
203
294
-
hRBD2
122
203
180
271
+
B
UNR (CSD1)
RNA (EF)
SXL (RRMs)
RRM1
RRM2
Mutants
Loop2 (IPEGGS)
1 2 3 4 5 6
(APAGGS)
Loop1 (RDKLTGR)
1 2 3 4 5 6 7
. S1 (RKDLTGR)
. S2 (RGKLTDR)
C
D. melanogaster
NLYVTNLPRTITDDQLDTIFGKYGSIVQKNILRDKLTGRPRGVAFVRYNKREEAQEAISALNNVIPEGGSQPLSVRLAEE
M. domestica
NLYVTNLPRTITDDELEKIFGKYGNIVQKNILRDKLTGRPRGVAFVRFNKREEAQEAISALNNVIPEGASQPLTVRLAEE
loop1
loop2

Figure 1. Residues of SXL important for translational repression.

A) Schematic representation of *Drosophila* and *Musca* SXL, together with their respective derivatives. The percentage of sequence identity in the RRMs is indicated. Numbers correspond to amino acid positions of the respective proteins. The capacity of each protein to repress *msl2* translation is indicated (see also Groskovic et al., 2003).

B) Crystal structure of the ternary complex of UNR (cold shock domain 1 [CSD1]), *msl2* EF RNA, and SXL (RRM1+RRM2). Loops 1 and 2 in SXL RRM2 and their amino-acid sequence are indicated. Numbers indicate the position of each amino acid in the loop. Mutations are indicated in red or cursiva. Adapted from (Hennig et al., 2014).

C) Sequence alignment of *Drosophila* and *Musca* SXL RRM2. In blue is highlighted loop1 sequence and in orange that of loop2.

We then evaluated the capacity of the loop mutants to repress *msl2* translation. As SXL can bind to both the 5' and 3' UTRs of *msl2* and can repress translation independently from both UTRs targeting different steps of translation initiation (Grskovic et al., 2003, Gebauer et al., 2003, Beckmann et al., 2005), we tested the effect of the loop mutations in both 5' UTR- and 3' UTR-mediated repression. To this end, we used reporter constructs containing SXL binding sites either at the 5' or the 3' UTRs. In addition, we used minimal and full length UTR constructs, since in Results part 1 of this thesis we observed that minimal constructs do not fully recapitulate the repression mechanism (i.e. cap-independent repression) (Figure 2A). Translation repression assays were performed in embryo extracts supplemented with increasing amounts of recombinant GST-dRBD4 variants, and an unrelated *Renilla* luciferase transcript as an internal control. The results of these experiments are shown in Figures 2B-C. The loop2 mutant was able to repress translation as efficiently as the wild type protein, whether 5' UTR- or 3' UTR-mediated. However, both Loop1-S1 and -S2 mutants showed defective 5' UTR-mediated repression. Interestingly, the 3' UTR-mediated mechanism was unaffected by these

mutations, indicating that the structure of the SXL variants was not overtly disturbed, and that loop1 is specifically required for 5' UTR-mediated repression. These results were true for both the minimal and the full UTR reporters (Figure 2).

Noteworthy is the slight reduction of the de-repression exerted by the loop1-S1 and -S2 mutations on the full-length constructs compared to that observed with the minimal transcripts. A possible explanation for this result is that the presence of more SXL binding sites in the 5'UTR (sites A and B) partially compensates for the loss of the repressive capacity of the protein.

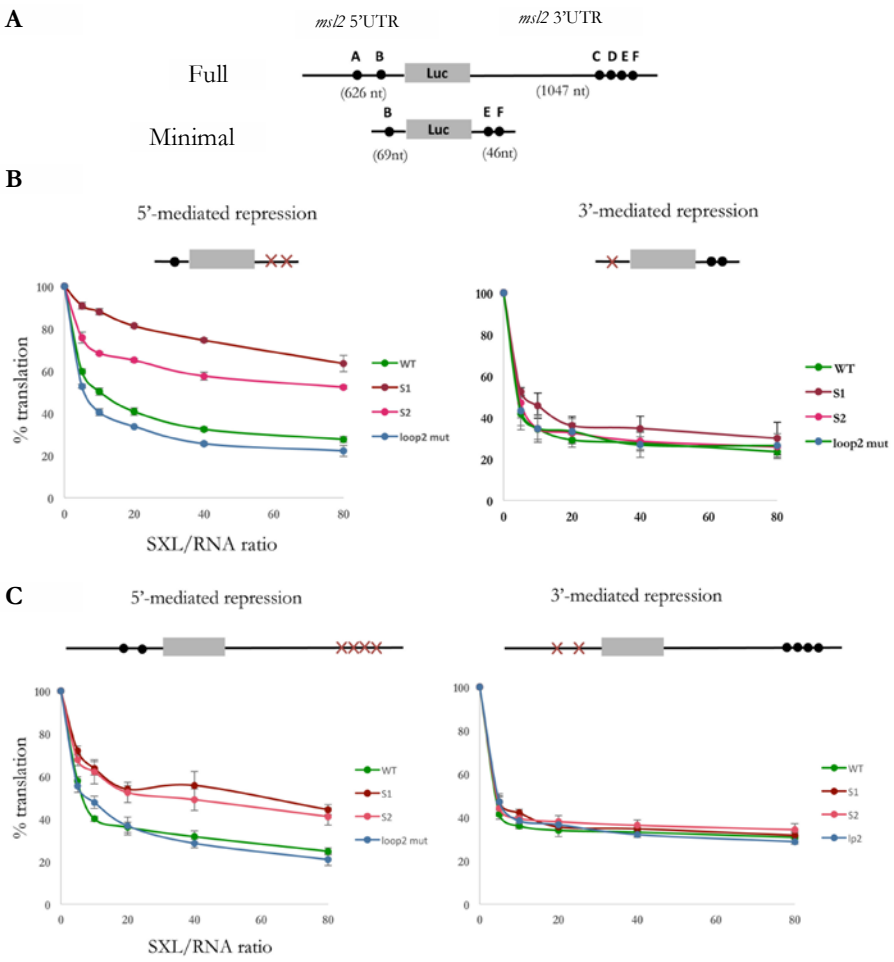


Figure 2. Translational repression ability of the SXL loop mutants.

A) Representation of the full-length and minimal UTR constructs, from were the transcripts used in (B) and (C) were derived. The length of the UTRs and the SXL binding sites (A to F) are indicated.

B) Repression capacity of the mutants using the minimal reporter constructs. A schematic representation of the construct used to analyze 5' UTR- or 3' UTR-mediated repression is indicated in the upper side of the graph. Translational repression assays were performed adding increasing amounts of the WT or mutant GST-dRBD4 derivatives to *Drosophila* embryo extracts, together with the indicated Firefly luciferase reporters. *Renilla* mRNA was co-translated as an internal control. Firefly luciferase values were corrected for Renilla expression and plotted relative to values obtained in the absence of dRBD4. Error bars represent the standard deviation of three biological replicates. C) Repression capacity of the mutants using the full length reporter constructs. Assays were performed as in B).

2.2. RNA binding capacity of the SXL mutants

A possibility to explain the effects of the loop1 mutations in 5'-mediated repression is that the mutants loose capacity to bind to the 5'UTR of *msl2*. We thus tested the RNA binding capacity of these mutants by cross-linking of the recombinant proteins in translation conditions (i.e. mixed with *Drosophila* extracts) to radiolabeled minimal *msl2* 5' UTR followed by immunoprecipitation (crosslink-IP). Loop2 mut protein, and the 3' UTR EF RNA were used as negative controls. As expected from the fact that the loop mutations do not affect 3' UTR-mediated repression, all proteins bound similarly to EF RNA (Figure 3A, lower panel). In addition, all proteins also bound similarly to the 5' UTR (Figure 3A, upper panel). A gel mobility shift assay (EMSA) with the purified recombinant proteins confirmed that the loop mutations did not affect the capacity of the proteins to bind to the 5' UTR (Figure 3B). Thus, we conclude that the loss of 5'-mediated repression observed for the loop1 mutants is not due to defects in RNA binding, but

most probably to defects in binding to (a) novel regulator(s) important for the 5' UTR-mediated mechanism.

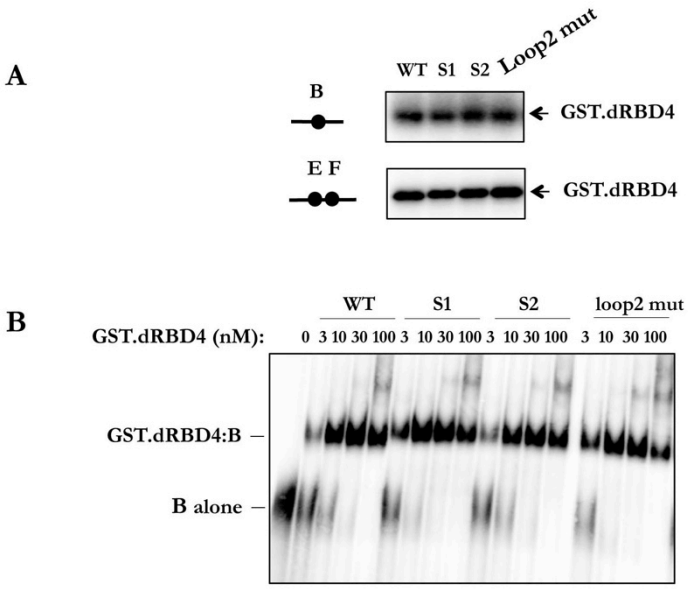
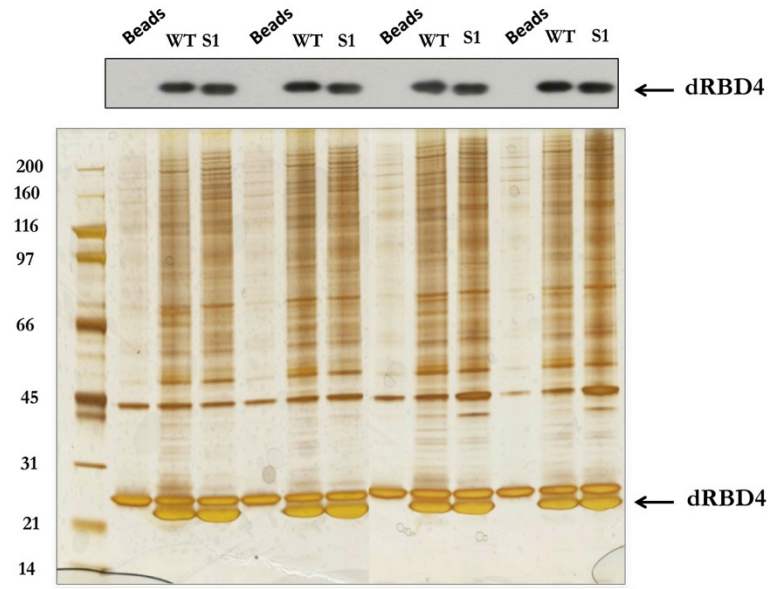


Figure 3. RNA binding capacity of SXL mutants. A) Cross-link IP of GST.dRBD4 derivatives in the presence of *Drosophila* extract and the radiolabeled 5' or 3'UTRs of the minimal reporter (containing B and EF sites, respectively). B) Gel mobility shift assay using radiolabeled minimal 5' UTR and increasing amounts of recombinant GST.dRBD4 derivatives. The amount of protein added (nM) is indicated. The location of the free RNA probe and the GST.dRBD4:RNA complexes are indicated.

2.3. Fishing for factors involved in 5' UTR- mediated regulation

In order to search for factors with potential to mediate repression driven by *msl2* 5' UTR, we performed GST-pull downs with WT GST.dRBD4 and the loop1-S1 mutant. We only used one of the two mutants for simplicity. The proteins were bound to glutathione sepharose beads, and *Drosophila* extracts equilibrated in translation conditions were passed several times through the sepharose bead columns. Empty beads were also carried to assess for background. The pull-down was performed in quadruplicates and the proteins retained in the columns analyzed by mass-spectrometry, as described in Materials and Methods. The quality of the experiment was assessed by silver staining (Figure 4A). The fold enrichment of each factor in the WT compared to the S1 pull-down was determined by measuring the area of the three most intense peptides detected for each protein. A t-student test was used to define the significance of the enrichment. The results of this analysis are shown in the volcano plot of Figure 4B and Table 1 of ANNEX 1. This conservative analysis gives a high degree of confidence to the selected candidates; however, factors only detected in one or none of the four eluates are not considered in the calculations for this method. As the ideal result is precisely a factor detected in none of the S1 eluates but reproducibly present in the WT eluates, this type of calculation eliminated the most interesting candidates. Therefore, we revised the proteomics data to search for these profiles and listed the identified factors in Table1. Table 2 is a classification of the candidates obtained by both methods according to biological function.

A


Beads
WT S1
Beads
WT S1
Beads
WT S1
Beads
WT S1
dRBD4
200
160
116
97
66
45
31
21
14
dRBD4

B

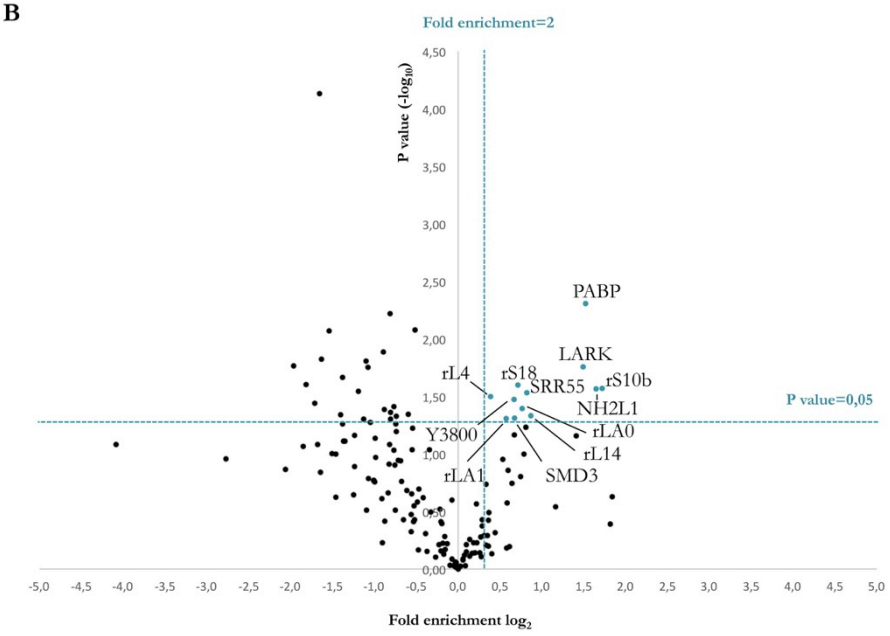

Fold enrichment=2
P value (-log10)
4,50
4,00
3,50
3,00
2,50
2,00
1,50
1,00
0,50
0,00
PABP
LARK
rL4
rS18
SRR55
rS10b
NH2L1
P value=0,05
Y3800
rLA0
rL14
rLA1
SMD3
-5,0 -4,5 -4,0 -3,5 -3,0 -2,5 -2,0 -1,5 -1,0 -0,5 0,0 0,5 1,0 1,5 2,0 2,5 3,0 3,5 4,0 4,5 5,0
Fold enrichment log2

Figure 4. GST-pull down to identify novel regulators of msl2 translation. A) Silver staining of the GST-pull downs using WT or S1 mutant GST.dRBD4 columns. Empty beads were carried as control. The experiment was performed in quadruplicates. A western blot of the eluates with αdRBD4 antibodies is shown at the top. B) Volcano plot of the mass spectrometry results. The blue lines indicate the significance thresholds (Pval=0.05 and fold enrichment 2). Significant factors are depicted in blue and their identities are indicated.

Table 1. Proteins present in at least 3 eluates of the WT GST.dRBD4 pull-down, and absent in at least 3 eluates of the S1 pull-down.

40S ribosomal protein S3(RS3)
40S ribosomal protein S20 (RS20)
40S ribosomal protein S9 (RS9)
Ribosomal L1 domain-containing protein (Y3096)
Ribosome biogenesis protein BOP1 homolog (BOP1)
Ribosome biogenesis protein WDR12 homolog (WDR12)
Ribosomal RNA processing protein 1homolog (RRP1L)
La-related protein (LARP)
Polycomb protein l(1) (U20)
Pre-mRNA-splicing factor Syf2 (SYF2)
Puff-specific protein Bx42(BX42)
RNA-binding protein fusilli (FUSIL)
Small nuclear ribonucleoprotein-associated protein B (RSMB)
Regulator of nonsense transcripts 1 homolog (RENT1/ UPF1)
HEAT repeat-containing protein 1 homolog (HEAT1)

Table 2. Candidate factors organized by their functions

TRANSLATION
Polyadenylation binding protein (PABP)
RNA-binding protein LARK (LARK)
La-related protein (LARP)
RIBOSOME COMPONENT
40S ribosomal protein (rS18)
40S ribosomal protein S10b (rS10B)
60S ribosomal protein L4 (rL4)
60S acidic ribosomal protein P0 (rLA0)
60S riboaomal protein L14 (rL14)
60S ribosomal protein P1 (rLA1)
Ribosomal L1 domain-containing protein (Y3096)
40S ribosomal protein S3(RS3)
40S ribosomal protein S20 (RS20)
40S ribosomal protein S9 (RS9)
SPLICING
Serine-arginine protein 55 (SRR55)
NHP2-like protein 1 homolog (NH2L1)
Small nuclear ribonucleoprotein Sm D3 (SMD3)
RNA-binding protein fusilli (FUSIL)
Small nuclear ribonucleoprotein-associated protein B (RSMB)
Puff-specific protein Bx42(BX42)
Pre-mRNA-splicing factor Syf2 (SYF2)
RIBOSOME BIOGENESIS
Ribosome biogenesis protein WDR12 homolog (WDR12)
Ribosome biogenesis protein BOP1 homolog (BOP1)
HEAT repeat-containing protein 1 homolog (HEAT1)
Ribosomal RNA processing protein 1homolog (RRPL1) (+ other functions)
OTHER
Polycomb protein l(1) (U20)
Regulator of nonsense transcripts 1 homolog (RENT1/ UPF1)
CCHC-type zinc finger protein (Y3800) (unknown function)

2.4. Testing the role of the identified factors on 5'-mediated repression.

From the candidates shown in Table 2, we focused on those shown to have roles in translation regulation: PABP, LARK and LARP. PABP was previously found to participate in repression of *msl2* via the 3'UTR-mediated mechanism (Duncan et al., 2006). Since the loop1 mutations had no effect on this mechanism, and PABP is an abundant protein that is often found as an interactor in pull-down assays, we did not give priority to PABP.

LARK is an RNA binding protein involved in the translational activation of several mRNAs in *Drosophila* (Huang et al., 2007) and is specifically required for translation regulation of circadian clocks (Huang et al., 2014). Its mammalian homolog RBM4 is involved in IRES-dependent translation (Lin et al., 2007) and in circadian clock translational control as well (Kojima et al., 2007). We therefore tested whether LARK participated in SXL-mediated *msl2* translational repression. We depleted LARK from male *Drosophila* S2 cells, which lack endogenous SXL, and tested the ability of transfected SXL to repress the translation of an *msl2* β-gal reporter containing SXL binding sites only in the 5' UTR (Δ3', Figure 5A). Cells were co-transfected with a plasmid encoding Renilla as an internal control. We also performed GFP knock-down as control. β-gal reads were normalized for Renilla expression and for the amount of reporter RNA to reveal true translational control (as opposed to regulation of RNA levels). The results indicated that LARK depletion did not affect the repression of the reporter by SXL (Figure 5B). Surprisingly, basal translation of the reporter was increased upon LARK depletion (Figure 5C), suggesting that LARK may repress *msl2* in a SXL-independent manner. To test this hypothesis, we turned to *Drosophila* female Kc cells, which express endogenous SXL. We used a reporter lacking SXL binding sites, and therefore independent of any SXL-mediated regulation. As expected, depletion of GFP or SXL did not affect translation of the reporter.

However, depletion of LARK augmented translation of the reporter (Figure 5E). These results indicate that LARK represses translation of *msl2* in a SXL-independent manner.

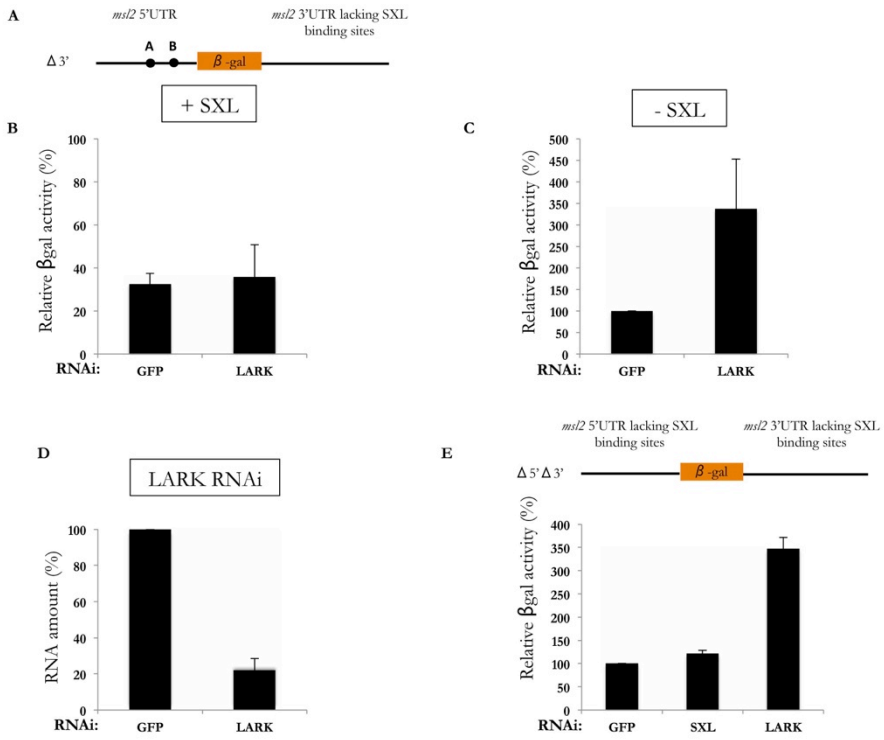


Figure 5. Role of LARK in msl2 translational repression. A) Schematic representation of the *msl2* reporter used in these experiments. The reporter contained the β-gal ORF fused to the full length UTRs of *msl2*, lacking SXL binding sites in the 3' UTR (Δ3'). B) LARK does not affect 5'-mediated repression of *msl2*. *Drosophila* male S2 cells were treated with RNAi against LARK, or GFP as control, and were subsequently transfected with the *msl2* reporter together with a control *Renilla* plasmid, and 75 ng of a SXL-encoding plasmid (pAC-SXL). β-gal reads were normalized for *Renilla* expression and for the amounts of reporter RNA. Repression was calculated as the percentage of translation relative to cells transfected with the pAC empty vector (i.e. lacking SXL). C) Basal translation of the Δ3' reporter upon LARK depletion. Translation in cells transfected with the pAC empty vector was measured as in (B), and the data referred to the values obtained in the control GFP depletion. D) The efficiency of LARK depletion was measured by RT-qPCR, taking *actin* as the transcript of reference. Results represent the average of three experiments. E) *Drosophila* female Kc cells were treated with RNAi against SXL, LARK and GFP, and transfected with a full length *msl2* reporter lacking SXL binding sites (Δ5'Δ3', schematic representation on the top of the graph) as described in (B). Translation was calculated as the percentage of β-gal activity relative to GFP depleted cells.

We next focused on LARP. The LARP family of proteins are conserved in metazoans, with five human LARP subfamilies identified so far (Larp1, Larp2, Larp4, Larp5 and Larp7) (Blagden et al., 2009). They contain two RNA binding motifs, a La motif (LAM) followed by an RNA recognition motif (RRM), arranged in a unique way to bind RNA. *Drosophila* LARP belongs to the LARP1 subfamily (Blagden et al., 2009, Bayfield et al., 2010). Mammalian LARP1 has been shown to activate translation of 5'-oligopyrimidine track (5' TOP) mRNAs, which usually encode proteins of the translational machinery, and to interact with the 3' terminus of poly(A) tails (Tcherkezian et al., 2014, Aoki et al., 2013). In mammalian cells,

LARP4B interacts with PABPC1 and RACK1 to activate translation (Schäffler et al., 2010). *Drosophila* LARP also interacts with PABPC, and therefore was proposed to have a role in translation (Blagden et al., 2009). We, thus, set to assess the function of LARP in SXL-mediated repression of *msl2*.

In parallel to LARP, we assessed the role of RACK1 (Receptor of activated C kinase 1). First, as mentioned above, LARP family members interact with RACK1; second, RACK1 is a ribosomal protein involved in mRNA-specific translational regulation (Baum et al., 2004, Ceci et al., 2003, Majzoub et al., 2014, Wolf and Grayhack, 2015, Kouba et al., 2012); third, RACK1 interacts with the ribosome using an "RDK" amino acid sequence in a loop, identical to the SXL region mutated in the loop1 mutants (Coyle et al., 2009) (Figure 6). Interestingly, mutation of the arginine (R) and the lysine (K) of the "RDK" loop sequence in RACK1 disrupts its interaction with the ribosome and affects translation in yeast (Coyle et al., 2009). As the RDK loops of RACK1 and SXL are similar, we (perhaps wildly) hypothesized that SXL could replace RACK1 in the ribosome through loop1. Replacement of RACK1 could be important either to interfere with a putative role of RACK1 on *msl2* translational activation, or to simply position SXL at a location in the ribosome appropriate to interact with other translation factors required for *msl2* expression. Our GST-pull downs show that RACK1 interacts with SXL, albeit not in a loop1-dependent fashion (ANNEX I), suggesting that RACK1 and dRBD4 are very likely engaged, but not through the "RDK". These data fit with the scenario that SXL binds to RACK1 to land in an appropriate location of the ribosome in order to interfere with other factors.

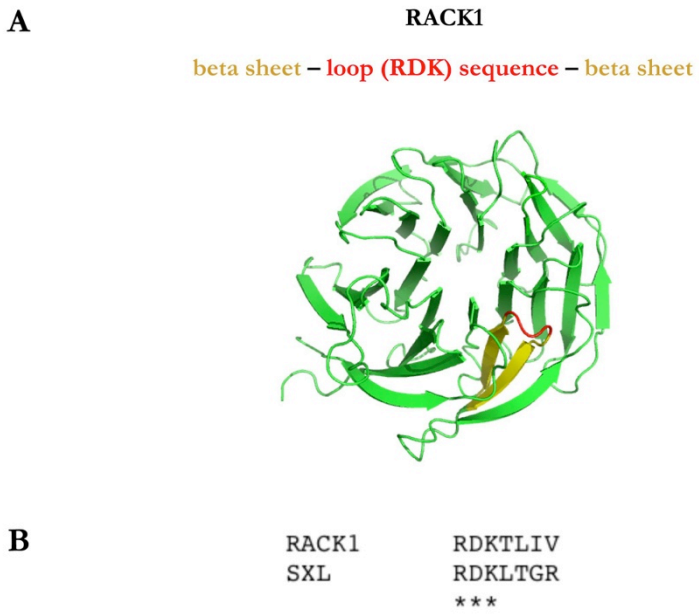


Figure 6. RACK1 structure. A) RACK1 displays a beta-propeller structure. The RDK loop that mediates interaction with the 40S ribosomal subunit is indicated in red. The β-sheet structures allowing the formation of the RDK loop are indicated in yellow. Figure made by Jae-Seong Yang (CRG). B) Amino acid sequence similarity between loop1 of SXL and the RDK sequence of RACK1 invovled in the contact with the ribosome (Coyle et al., 2009).

To test the role of LARP and RACK1 on translational repression of *msl2* via the 5' UTR, we followed an approach identical to that used for LARK. We depleted these proteins from *Drosophila* S2 cells and tested the ability of SXL to repress the translation of the Δ3' reporter. It is important to note that none of the depletions had a significant effect on cell viability (data not shown). The results showed that, while depletion of LARP did not alter SXL-mediated repression, depletion of RACK1 completely abolished it (Figure 7A). Both depletions were efficient (Figure 7C). We conclude that RACK1 is required for translational repression of *msl2* via the 5' UTR.

Basal translation of the reporter was not affected by any of the depletions (Figure 7B), indicating that RACK1 is not required for *msl2* translation, and favoring the option that RACK1 is simply acting as a scaffolding protein to locate SXL at an appropriate position in the ribosome to interact with translation factors necessary for *msl2* regulation (see Discussion).

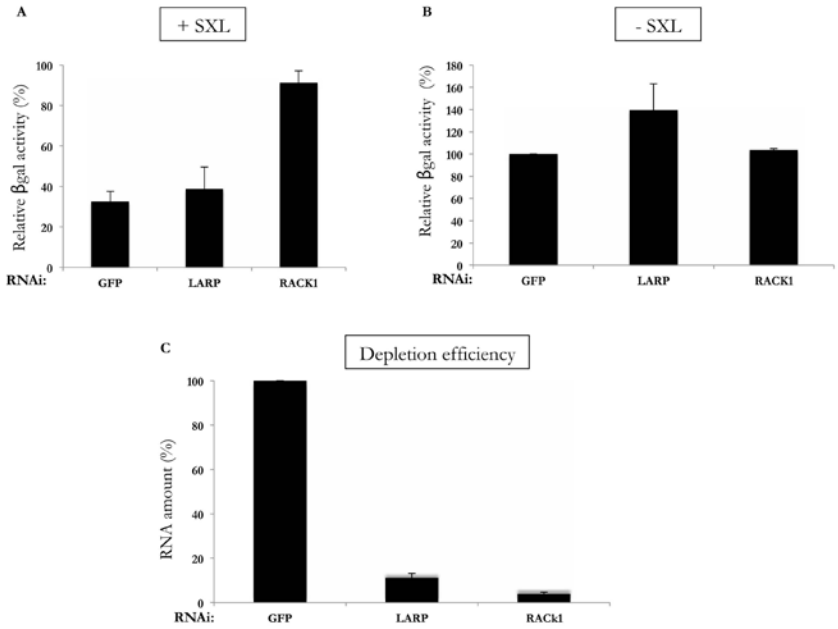


Figure 7. RACK1, but not LARP1, is involved in 5' UTR-mediated regulation of msl2 mRNA. A) Translational repression by SXL in cells depleted of GFP, LARP or RACK1. Assays were performed as described in the legend of Figure 5B. Error bars represent the standard deviation of 3 replicate experiments. B) Basal translation is unaffected by the depletions. Translation of the Δ3' reporter was assessed in cells transfected with the pAC empty vector (i.e. in the absence of SXL). C) Efficiency of LARP and RACK1 depletion, measured by RT-qPCR using *actin* as the transcript of reference. Data represent the average of at least two experiments.

III. MATERIALS AND METHODS

This section refers to Results Part 2 and is intended to complement the experimental procedures of Results Part 1.

Plasmids

SXL dRBD4 plasmid is derived from PGEX6P vector and comprises a GST tag connected via a 3C protease site to SXL residues 122-301 (a generous gift from Jan Medenbach). dRBD4 mutant derivatives loop1-S1, loop1-S2 and Loop2 mutant were obtained using the Quick site directed mutagenesis system (Agilent).

BLEFm and BmLEF plasmids were previously decribed (Abaza et al., 2008).

The EF site, *msl2* βgal reporter, pAc-Renilla and pAc-SXL were previously described (Graindorge et al., 2013; Szostak et al. Submitted). The B probe was obtained by digestion of the BLEF construct already described (Gebauer et al., 2003) with the restriction enzyme NcoI.

Protein preparation

SXL dRBD4 (amino acids 122-301) and its mutant derivatives were expressed in *Escherichia Coli* as N-terminal GST-tagged fusions and purified as previously decribed (Grskovic et al., 2003).

RNA in vitro transcription

In vitro transcription was performed as described (Gebauer et al., 1999) with a 5' m^7GpppG cap structure (KEDAR). Unless indicated otherwise, all mRNAs contained a poly(A) tail of 73 residues. RNAs used in the X-link IP and in the EMSA were non-polyadenilated, A-capped and they were trace-labelled with [α-^{32}P]ATP.

UV-cross-link IP and Gel mobility shift assay (EMSA)

UV-cross-link IPs and EMSAs were performed as described (Abaza et al., 2006). For the Cross-link IP proteins were separated by 12% SDS-PAGE and visualized using PhophorImager. For the EMSA, probes were separated in a 4% acrylamide gel and visualized following the same procedure.

Mass spectrometry analysis

Samples were digested with a combination of LysC and trypsin as described before (Chiva et al., 2014). Peptide mixtures were analyzed using a LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, San Jose, USA) coupled to a nano-LC (Proxeon, Odense, Denmark) equipped with a reversed-phase chromatography 2-cm C18 pre-column (Acclaim PepMap-100, Thermo; 100 μm i.d., 5 μm), and a 12-cm C18 analytical column (Nikkyo Technos, 75 μm i.d., 3 μm) using a 3-35% buffer B in a 60 min gradient (Buffer A: 0.1% formic acid in water. Buffer B: 0.1% formic acid in acetonitrile). The instrument was operated in data-dependent acquisition (DDA) mode. Following each survey scan acquired in the Orbitrap, the 10 most intense ions with multiple charged ions were selected for fragment ion spectra produced via collision-induced dissociation (CID).

Acquired data were analyzed using the Proteome Discoverer software suite (v1.4, Thermo Fisher Scientific), and the Mascot search engine (v2.5, Matrix Science(Perkins et al., 1999)) was used for peptide identification. Data were searched against a *Drosophila melanogaster* protein database derived from SwissProt (total of 3817 sequences). A precursor ion mass tolerance of 7 ppm at the MS1 level was used, and up to three missed cleavages for trypsin were allowed. The fragment ion mass tolerance was set to 0.5 Da. Oxidation of Methionine was defined as variable modification and carbamidomethylation of Cysteines was set as fixed modification. The identified peptides were filtered by *Mascot ion score* higher than 20. The $\log_2$

corresponding to the average area of the three most intense peptides per protein as calculated by Proteome Discoverer was used as quantitation indicator and a Student's t-test was performed between the 4 replicates of each state to pinpoint differentially abundant proteins.

GST pull-down

300 ug of GST.dRBD4 or its mutant derivative S1 were resuspended in 600 ul of buffer T (30 mM HEPES pH 8.0, 0.6mM Mg(OAc)2 , 80 mM KOAc) and added to 200 ul of a (1:1) surry of sepharose glutathione beads (GE Healthcare) previously equilibrated in buffer T. For the negative control sample, 600 ul of buffer T with no protein were added to the beads. Samples were incubated 30 min at 4ºC on a rotating platform. The suspension was loaded on a column and washed 5 times with buffer T supplemented with 3mg/ml heparin (buffer TH). 10 mg per sample of *Drosophila* embryo extract were centrifuged 15 min at 13000 rpm at 4ºC and the supernatant was adjusted to buffer T and supplemented with protease inhibitor cocktail (Boehring). The equilibrated pre-centrifged extract was loaded eather in the negative control, WT or S1 column, and incubated 20 min at RT. The extract was loaded over the same column 4 times. The incubation followed by 4 loadings on the columns was repeated with new extract in the same conditions.

The columns were washed 4 times witth 1ml of buffer TH, and equilibrated in 500 ul of elution buffer (50mM Tris HCl pH 8, 150mM NaCl, 1mM DTT). 100 ul of elution supplemented with 3ug of 3C protease were added to each column that were incubated 2 hours at 4ºC. The columns were spinned and 20ul of the eluate were separated in Bis-Tris Gels 4-12% (NuPAGE Novex) to be analyzed by Silver Staining. 2 ul of the eluate were run on a 12% SDS-PAGE acrylamide gel and analyzed by WB using a 1:1000 dilution of anti-dRBD4 antibody previously desccribed (Graindorge

et al., 2013). The rest of the eluate was analyzed by the Proteomic Facility at the Centre for Genomic Regulation.

RNAi and transfection

RNA intererence and transfections were performed as described previously (Graindorge et al., 2013; Szostak et al., submitted). The βgal Δ3 reporter constrcut was co-transfected with pAc-Renilla control and increasing amounts of pAc-SXL. Renilla luciferase and Bgalactosidaes activities were measured with luciferase (Promega) and Galacto-Star (tropix) kits, respectively. βgal activity was corrected for co-transfected Renilla, and normalized for reporter RNA levels. RNA extraction and quantification was performed as described (Szostak et al., submitted).

IV. DISCUSSION

Expression of *msl2* is required in male flies for the formation of the dosage compensation complex (DCC) to promote a two-fold hypertranscription of the single X chromosome. However, in females, *msl2* must be repressed because its expression leads to the assembly of the DCC on both X chromosomes and to lethality (Kelley et al., 1995). Thus, efficient and tight repression of *msl2* must take place to ensure female viability. Repression occurs at many levels in the post-transcriptional cascade. The female specific protein SXL orchestrates repression by binding to the 5' and 3' UTRs of the *msl2* transcript, resulting in inhibition of splicing, nucleo-cytoplasmic transport and translation. The inhibition of *msl2* translation is itself tightly coordinated. SXL bound to the 3'UTR recruits the co-factor UNR to inhibit ribosome recruitment, while SXL bound to the 5' UTR inhibits the scanning of the ribosomal subunits that presumably have escaped the 3'UTR-mediated control by a mechanism that involves recognition of an uAUG (Gebauer et al., 2003, Beckmann et al., 2005, Duncan et al., 2006, Abaza et al., 2006). The targets of the SXL repressor complexes in the translational machinery have remained elusive to date. We have recently found that the 3' UTR repressive complex targets eIF3d, an initiation factor required for *msl2* translation (Szostak et al, *submitted*). In the first part of this work, studying the contribution of the cap structure to the repressory mechanism, we have found that eIF3d binds to mRNA sequences in the 5' UTR of *msl2* in the absence of a cap structure. We believe that this feature may confer increased eIF3d sensitivity to *msl2* mRNA. In the second part of this work, we have found that the ribosomal protein RACK1 is a potential target of the 5'UTR repressor mechanism. Intriguingly, eIF3d and RACK1 are located close together in the 40S ribosomal subunit.

1. eIF3d, a target of the *msl2* 3' UTR repressor complex, binds to *msl2* 5' UTR

The standard mechanism of translation initiation poses that the small ribosomal subunit is recruited to a cap-proximal position by the chain of interactions 7mGTP-eIF4E-eIF4G-eIF3 (Sonenberg and Hinnebusch, 2009). Our results suggest that initiating ribosomal subunits can also be directly recruited to the 5' UTR mediated by the factor eIF3d, and support the idea that coordination of translation initiation is more complex than it seems at a first glance. As described in the introduction, alternative ways of translation initiation might take place in the cell. Other initiation factors have been shown to stimulate translation of uncapped mRNAs. For instance, a proteolytic product of eIF4G lacking eIF4E binding sites can stimulate translation of uncapped mRNAs that do not contain IRES structures (Ohlmann et al., 1995, De Gregorio et al., 1998). This ability depends on the RNA binding domains of eIF4G (Prévôt et al., 2003). However, the binding ability of eIF4G seems rather general and non-specific, contrary to the binding of eIF3d (Results Part 1, Figure 4C). Our data are consistent with those of Lee and colleagues (2015), which have observed that eIF3 subunits (a, b, d, and g) only bind to ~3% of total expressed transcripts, indicating that eIF3 subunits do not bind mRNA in general. The sequence responsible for cap-independent *msl2* repression seems to be formed by several regions of the 5' UTR (Results Part 1, Figure 3). Thus, it is possible that eIF3d binds to an *msl2* mRNA structure formed by distant sequences in the 5' UTR. It is also conceivable that binding of eIF3d to *msl2* is not direct, but is brought about by factors recruited to different regions of the 5' UTR. Future experiments with recombinant eIF3d will help to elucidate this question.

An important question is whether eIF3d is recruited to the 5' UTR as part of the eIF3 complex or whether it functions in an eIF3-independent manner. On one side, eIF3d is a non-core subunit of the eIF3 complex, and is not

present in eIF3 of all organisms (Masutani et al., 2007, Sun et al., 2011, Smith et al., 2013). The structure of the mammalian eIF3 complex bound to the 40S ribosomal subunit was recently obtained by electron microscopy (des Georges et al., 2015). These authors found that eIF3d is located in the mRNA exit channel, but did not observe interactions of eIF3d with any other eIF3 subunit (although they speculate that unassigned densities at the eIF3 core could perhaps belong to eIF3d). The only eIF3d interactions detected were with eIF2α and with different ribosomal proteins (RACK1, Rps28, Rps7 and Rps9) (des Georges et al., 2015). Our mass-spectrometry analysis of Hrp48 pull-down eluates detected eIF3d but not other eIF3 subunit, suggesting that eIF3d could play a role independent of the core eIF3 in *msl2* translation (Szóstak et al., *submitted*). However, future experiments are needed to answer this question.

Other authors have shown that eIF3d interacts with eIF4G to promote ribosome recruitment (Sun et al., 2011, Villa et al., 2013, des Georges et al., 2015). eIF3d could also promote recruitment of ribosomes to the 5' UTR of *msl2* independently of eIF4G by interactions with mRNA elements in the 5'UTR of *msl2*. This interaction would allow *msl2* mRNA to bypass the requirement of m^7GTP-eIF4E during ribosome recruitment. Supporting this scenario, yeast eIF3 has been shown to stimulate recruitment of ribosomes to uncapped mRNAs in the absence of eIF4F factors (Mitchell et al., 2010). eIF4F-independent binding of eIF3 has also been reported for IRES elements (Cai et al., 2010, Kieft, 2008). However, we do not favor the hypothesis that the 5' UTR of *msl2* contains an IRES because the efficiency of translation of *msl2* reporters drops dramatically in the absence of the cap structure. Thus, should there be an IRES, its activity would be extremely low.

Dissecting eIF3d-dependent translation requires further investigation. eIF3d is required for *msl2* translation because its knock-down in *Drosophila* S2 cells significantly affects *msl2* translation without overtly disturbing general

protein synthesis (Szóstak et al., *submitted*). Nevertheless in our hands, the basal translation of transcripts that recruit eIF3d to the 5' UTR is similar to that of transcripts that cannot recruit it, whether the mRNAs contain a cap or not (e.g. compare A-mlm with A-BLEF in Fig 1C of Part 1). However, since we believe these 5' UTRs follow different mechanisms of translation initiation, it is difficult to compare their degree of eIF3d dependence by simply comparing their levels of translation. Further experiments are required to solve this question.

In summary, our data support a model where efficient recruitment of small ribosomal subunits to *msl2* mRNA (and therefore efficient translation initiation) requires eIF3d. This initiation factor may be attracted to *msl2* mRNA via the cap structure, i.e. as part of the 43S pre-initiation complex in the context of the whole eIF3, or by binding to the 5' UTR (Results Part 1 Figure 5 MODEL). The SXL:UNR:Hrp48 repressor complex assembled on the 3' UTR of *msl2* would then target eIF3d, either to inhibit its productive recruitment to the mRNA or to block its interaction with downstream initiation factors required for translation.

2. Insights into translational repression of *msl2* mRNA by dissection of relevant SXL residues.

Structure-based analysis of the SXL:UNR:*msl2* complex identified a conserved, exposed loop in SXL RRM2 with potential to interact with factors important for translational repression. Unexpectedly, mutation of this loop did not affect 3' UTR- mediated repression but impaired repression via the 5' UTR (Results part 2, Figure 2). Cross-link IP and GEMSA analyses showed that the mutant had intact capacity to bind to *msl2* 5'UTR, indicating that the mutation might have rather disrupted the binding of a co-factor or a target in the translational machinery. 5'UTR-mediated repression has been

shown to inhibit ribosomal scanning by promoting the recognition of an upstream AUG (uAUG3) (Beckmann et al., 2005, Medenbach et al., 2011). However, the factors involved in this mechanism are unknown. Thus, we reasoned that proteins binding to WT SXL, but not to SXL with mutations in the loop, could give us insights about these factors. Pull-down experiments identified several candidates, including LARK and LARP, two proteins previously implicated in translational control (Huang et al., 2007, Huang et al., 2014, Lin et al., 2007, Kojima et al., 2007, Tcherkezian et al., 2014, Aoki et al., 2013, Schäffler et al., 2010, Blagden et al., 2009). Surprisingly, functional analyses did not validate these candidates as co-factors for SXL-mediated repression. LARK did show a translational repressor activity on *msl2* mRNA, but this effect was independent of SXL. Why does then LARK bind to SXL? It is conceivable that there is coordination between different mechanisms acting on *msl2* to ensure its repression e.g. in a temporal or tissue-specific manner, and the SXL:LARK contact reflects this coordination.

We then noticed that the SXL loop was almost identical to a loop that the ribosomal protein RACK1 uses to bind to the ribosome (Coyle et al., 2009) (Results part 2, Figure 6), and that RACK1 interacts strongly with both the WT and loop mutant SXL isoforms in our GST pull-down experiments (Annex I, Table I). Importantly, depletion of RACK1 completely abolished 5' UTR- mediated repression by SXL (Part 2, Figure 7). Depletion of RACK1 did not affect basal translation of the *msl2* reporter compared to a control transcript, suggesting that RACK1 is not necessary for *msl2* translation and, therefore, is not likely a target of SXL in the translational machinery but rather functions as a SXL co-factor.

RACK1 is a highly conserved protein of 36 kDa that was found on a search for receptors of active PKC in rat brains (thereby its name as Receptor for Activated C Kinase 1) (Ron et al., 1994). RACK1 belongs to the tryptophan-aspartate (WD) repeat family of proteins, and behaves as a scaffolding

protein that interacts with multiple partners, participating in diverse physiological processes such as embryonic development, cell migration, circadian rhythm and central nervous system function (Adams et al., 2011). In addition to PKC, RACK1 interacts with other kinases such as MAPK, JNK, FAK, Fyn and Src, and with the cytoplasmic tails of a number of transmembrane receptors, acting as an integrator of signaling cascades in the cell (López-Bergami et al., 2005, Mamidipudi et al., 2004, Yaka et al., 2002, Yaka et al., 2003). The most frequent consequences of these interactions are the relocation of the kinases to appropriate places in the cell, and/or the activation or inhibition of the kinase activity. Consistent with its diverse functions, RACK can be found at several cellular locations (cytoplasm, endoplasmic reticulum, nucleus). A fraction of RACK1 associates with the small ribosomal subunit where it localizes near the mRNA exit channel (des Georges et al., 2015). On the ribosome, RACK1 has been proposed to regulate general and mRNA-specific translation. For instance, RACK1 recruits PKCβII and promotes phosphorylation of the anti-association factor eIF6, leading to the dissociation of this factor from the 40S ribosomal subunit and to the joining of the 60S subunit to start translation (Ceci et al., 2003). RACK1 has also been shown to activate IRES- dependent translation in *Drosophila* cells (Majzoub et al., 2014), and to associate with RNA binding proteins in yeast, with the potential to regulate the translation of mRNA subsets (Baum et al., 2004).

RACK1 adopts a seven-bladed β-propeller structure (Results part 2, Figure 6). The RDK-containing loop in blade 1 of the β-propeller hooks RACK1 to the head of the 40S ribosomal subunit, close to the mRNA exit channel (Coyle et al., 2009). The primary anchoring region for RACK1 in the ribosome is thought to implicate mainly helices 39 and 40 of the 18S rRNA, but RACK1 also interacts with ribosomal proteins RpS17, RpS16 and RpS3. The RDK loop interacts with 18S rRNA via the lysine (K), while the arginine (R) binds to RpS17. RpS3 binds RACK1 through a C-terminal strand that

wraps blade 4 and terminates in contact with blade 5 (reviewed in Adams et al., 2011). RpS3 interacts with RpS20, (throught its N-terminal domain) in addition to RACK1, and similarly to RpS3, RpS20 is enriched in the GST pull-down of WT compared to that of the loop mutant. Interestingly, RpS3 which is located at the mRNA entry channel, actively participates on the conformational changes of the ribosome required for scanning. In the scanning-resistant form, the mRNA entry channel of the 40S subunit is closed by a latch formed by interactions of 18S rRNA h34-h18 that is stabilized by the C-terminal domain of RpS3 (Passmore et al., 2007, Hussain et al., 2014, Zhang et al., 2015, Graifer et al., 2014). During the formation of the pre-initiation complex, binding of eIF1A and eIF1 promote interactions of the N-terminal domain of RpS3 with h16 (Graifer et al., 2014) at the entry channel, leading to dissolution of the h34-h18 latch and to an open, scanning-competent ribosome (Passmore et al., 2007).

Given that SXL inhibits the scanning of 40S subunits on the 5' UTR of *msl2* and promotes recognition of an uAUG (Beckman et al., 2005; Medenbach et al., 2011), it seems conceivable that RpS3 could be a suitable target. In this scenario, the interaction of SXL with RACK1 could guide SXL into the ribosome to interact with RpS3 (Figure 8). Interestingly, RACK1 contains a second RDK loop that locates very close to the C terminal domain of RpS3 (data not shown), suggesting that the loop1 of SXL could mimic this loop to interact with RpS3. Subsequently, SXL could inhibit the RpS3-h16 contact in a loop1-dependent manner leading to a closed conformation of the ribosome and to scanning arrest. Appropriate repression by SXL bound to the 5' UTR requires an optimal distance of 27-28 nucleotides between the SXL binding site and the uAUG, which is precisely the size occupied by a ribosome on the mRNA (28-30 nt), suggesting spatial constraints compatible with the mechanism proposed here (Medenbach et al., 2011). Scanning arrest at this position would lead to recognition of the uAUG and the initiation of translation away from the main ORF.

Although in the above scenario RACK1 would function as a scaffolding protein to guide SXL, given the role of RACK1 as a hub for protein kinases, an alternative- although not mutually exclusive- scenario is that RACK1 promotes phosphorylation of SXL to habilitate its downstream effect on RpS3. Future experiments will be directed to test these hypotheses.

Altogether, the results from this thesis allow us to propose a model where SXL bound to the 5' and 3' UTRs of *msl2* targets different factors on the same region of the ribosome. SXL bound to the 3' UTR, in cooperation with Hrp48 and UNR, targets eIF3d leading to inhibition of ribosome recruitment. SXL bound to the 5' UTR binds RACK1 to mediate targeting of RpS3. SXL interaction with RpS3 may induce conformational changes leading to re-establishment of the mRNA channel latch and intiation of translation at uAUG3. Interestingly, RACK1 also contacts eIF3d in the ribosome. As RACK1, RpS3 and eIF3d are spatially interconnected, their coordinated inhibition by SXL may explain the cooperativity between the 5' and 3' UTRs of *msl2* for translational repression (Figure 8).

MODEL

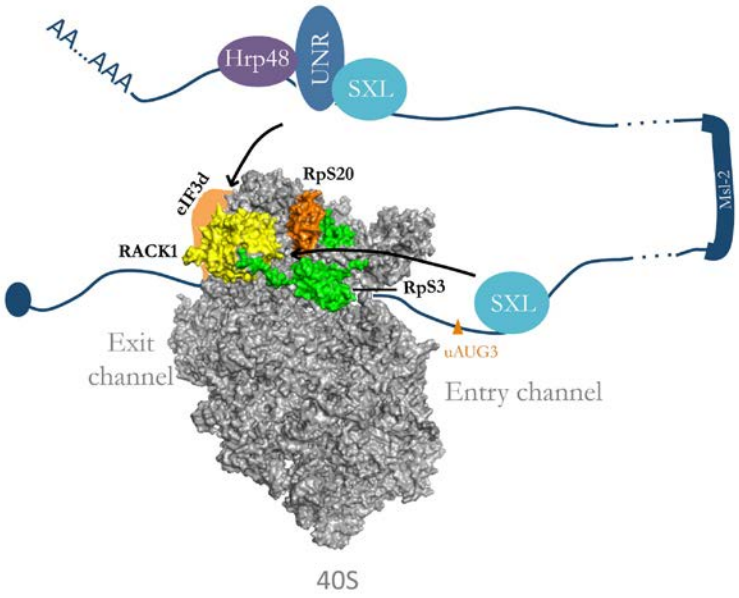


Figure 8. Model of msl2 translational regulation by SXL. Rack1 is depicted in yellow, RpS3 in green, eIF3d in light orange, and RpS20 in dark orange. The 3'UTR repressor complex targets eIF3d, which interacts with RACK1 in the ribosome. SXL bound to the 5'UTR interacts with RACK1 which may mediate targeting of RpS3. Figure made by Jae-Seong Yang (CRG) (PDM accession number 2xZN).

V. CONCLUSIONS

PART 1

1. SXL represses *msl2* translation in a cap-independent manner.
2. The cap-independency relies on features of the 5' UTR of *msl2*.
3. Hrp48, a SXL co-factor binding to the 3'UTR, is required for cap-independent repression.
4. eIF3d, a target of Hrp48 in the translation machinery, binds to the 5'UTR of *msl2* and explains cap-independent repression.

PART 2

5. A loop located in SXL RRM2 is specifically required for *msl2* translational repression mediated by the 5'UTR, but not for the 3'-mediated mechanism.
6. Mutation of this loop does not affect binding of SXL to the 5'UTR of *msl2*, suggesting that it contacts additional regulators or targets in the translation machinery.
7. Pull-down followed by mass-spectrometry analysis revealed potential co-factors and targets. Notably, RpS3 and RACK1.
8. Functional studies indicate that RACK1 is required for the 5' UTR-mediated mechanism. We propose a model whereby SXL is positioned in the ribosome through RACK1 to target RpS3 in order to inhibit ribosomal scanning. Coordinated repression of eIF3d and RpS3 by SXL may explain the cooperativity of *msl2* UTRs for translational repression.

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(3), pp. 380-9.

Adams, D. R., Ron, D. and Kiely, P. A. (2011) 'RACK1, A multifaceted scaffolding protein: Structure and function', *Cell Commun Signal,* 9, pp. 22.

Aitken, C. E. and Lorsch, J. R. (2012) 'A mechanistic overview of translation initiation in eukaryotes', *Nat Struct Mol Biol,* 19(6), pp. 568-76.

Aoki, K., Adachi, S., Homoto, M., Kusano, H., Koike, K. and Natsume, T. (2013) 'LARP1 specifically recognizes the 3' terminus of poly(A) mRNA', *FEBS Lett,* 587(14), pp. 2173-8.

Bashaw, G. J. and Baker, B. S. (1997) 'The regulation of the Drosophila msl-2 gene reveals a function for Sex-lethal in translational control', *Cell,* 89(5), pp. 789-98.

Baum, S., Bittins, M., Frey, S. and Seedorf, M. (2004) 'Asc1p, a WD40-domain containing adaptor protein, is required for the interaction of the RNA-binding protein Scp160p with polysomes', *Biochem J,* 380(Pt 3), pp. 823-30.

Bayfield, M. A., Yang, R. and Maraia, R. J. (2010) 'Conserved and divergent features of the structure and function of La and La-related proteins (LARPs)', *Biochim Biophys Acta,* 1799(5-6), pp. 365-78.

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(4), pp. 529-40.

Belote, J. M. and Lucchesi, J. C. (1980) 'Male-specific lethal mutations of Drosophila melanogaster', *Genetics,* 96(1), pp. 165-86.

Blagden, S. P., Gatt, M. K., Archambault, V., Lada, K., Ichihara, K., Lilley, K. S., Inoue, Y. H. and Glover, D. M. (2009) 'Drosophila Larp associates with poly(A)-binding protein and is required for male fertility and syncytial embryo development', *Dev Biol,* 334(1), pp. 186-97.

BRENNER, S., JACOB, F. and MESELSON, M. (1961) 'An unstable intermediate carrying information from genes to ribosomes for protein synthesis', *Nature,* 190, pp. 576-581.

Buchan, J. R. (2014) 'mRNP granules. Assembly, function, and connections with disease', *RNA Biol,* 11(8), pp. 1019-30.

Cai, Q., Todorovic, A., Andaya, A., Gao, J., Leary, J. A. and Cate, J. H. (2010) 'Distinct regions of human eIF3 are sufficient for binding to the HCV IRES and the 40S ribosomal subunit', *J Mol Biol,* 403(2), pp. 185-96.

Ceci, M., Gaviraghi, C., Gorrini, C., Sala, L. A., Offenhäuser, N., Marchisio, P. C. and Biffo, S. (2003) 'Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly', *Nature,* 426(6966), pp. 579-84.

Chiva, C., Ortega, M. and Sabidó, E. (2014) 'Influence of the digestion technique, protease, and missed cleavage peptides in protein quantitation', *J Proteome Res,* 13(9), pp. 3979-86.

Conrad, T. and Akhtar, A. (2011) 'Dosage compensation in Drosophila melanogaster: epigenetic fine-tuning of chromosome-wide transcription', *Nat Rev Genet,* 13(2), pp. 123-34.

Coyle, S. M., Gilbert, W. V. and Doudna, J. A. (2009) 'Direct link between RACK1 function and localization at the ribosome in vivo', *Mol Cell Biol,* 29(6), pp. 1626-34.

De Gregorio, E., Preiss, T. and Hentze, M. W. (1998) 'Translational activation of uncapped mRNAs by the central part of human eIF4G is 5' end-dependent', *RNA,* 4(7), pp. 828-36.

de Melo Neto, O. P., Standart, N. and Martins de Sa, C. (1995) 'Autoregulation of poly(A)-binding protein synthesis in vitro', *Nucleic Acids Res,* 23(12), pp. 2198-205.

de Moor, C. H., Meijer, H. and Lissenden, S. (2005) 'Mechanisms of translational control by the 3' UTR in development and differentiation', *Semin Cell Dev Biol,* 16(1), pp. 49-58.

des Georges, A., Dhote, V., Kuhn, L., Hellen, C. U., Pestova, T. V., Frank, J. and Hashem, Y. (2015) 'Structure of mammalian eIF3 in the context of the 43S preinitiation complex', *Nature.*

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(3), pp. 368-79.

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', *Mol Cell,* 36(4), pp. 571-82.

Erzberger, J. P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C. H., Cimermančič, P., Boehringer, D., Sali, A., Aebersold, R. and Ban, N. (2014) 'Molecular architecture of the 40S·eIF1·eIF3 translation initiation complex', *Cell,* 158(5), pp. 1123-35.

Förch, P., Merendino, L., Martínez, C. and Valcárcel, J. (2001) 'Modulation of msl-2 5' splice site recognition by Sex-lethal', *RNA,* 7(9), pp. 1185-91.

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(21), pp. 6146-54.

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', *Mol Cell,* 11(5), pp. 1397-404.

Gebauer, F. and Hentze, M. W. (2004) 'Molecular mechanisms of translational control', *Nat Rev Mol Cell Biol,* 5(10), pp. 827-35.

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(2), pp. 142-50.

Gelbart, M. E. and Kuroda, M. I. (2009) 'Drosophila dosage compensation: a complex voyage to the X chromosome', *Development,* 136(9), pp. 1399-410.

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. and Smith, H. O. (2009) 'Enzymatic assembly of DNA molecules up to several hundred kilobases', *Nat Methods,* 6(5), pp. 343-5.

Graifer, D., Malygin, A., Zharkov, D. O. and Karpova, G. (2014) 'Eukaryotic ribosomal protein S3: A constituent of translational machinery and an extraribosomal player in various cellular processes', *Biochimie,* 99, pp. 8-18.

Graindorge, A., Carré, C. and Gebauer, F. (2013) 'Sex-lethal promotes nuclear retention of msl2 mRNA via interactions with the STAR protein HOW', *Genes Dev,* 27(12), pp. 1421-33.

Graindorge, A., Militti, C. and Gebauer, F. (2011) 'Posttranscriptional control of X-chromosome dosage compensation', *Wiley Interdiscip Rev RNA,* 2(4), pp. 534-45.

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(20), pp. 5571-81.

Hennig, J., Militti, C., Popowicz, G. M., Wang, I., Sonntag, M., Geerlof, A., Gabel, F., Gebauer, F. and Sattler, M. (2014) 'Structural basis for the

assembly of the Sxl-Unr translation regulatory complex', *Nature,* 515(7526), pp. 287-90.

Hershey, J. W. (2015) 'The role of eIF3 and its individual subunits in cancer', *Biochim Biophys Acta,* 1849(7), pp. 792-800.

Hinnebusch, A. G. (2014) 'The scanning mechanism of eukaryotic translation initiation', *Annu Rev Biochem,* 83, pp. 779-812.

Hinnebusch, A. G. (2015) 'Cell biology. Blocking stress response for better memory?', *Science,* 348(6238), pp. 967-8.

Huang, Y., Genova, G., Roberts, M. and Jackson, F. R. (2007) 'The LARK RNA-binding protein selectively regulates the circadian eclosion rhythm by controlling E74 protein expression', *PLoS One,* 2(10), pp. e1107.

Huang, Y., McNeil, G. P. and Jackson, F. R. (2014) 'Translational regulation of the DOUBLETIME/CKIδ/ε kinase by LARK contributes to circadian period modulation', *PLoS Genet,* 10(9), pp. e1004536.

Hussain, T., Llácer, J. L., Fernández, I. S., Munoz, A., Martin-Marcos, P., Savva, C. G., Lorsch, J. R., Hinnebusch, A. G. and Ramakrishnan, V. (2014) 'Structural changes enable start codon recognition by the eukaryotic translation initiation complex', *Cell,* 159(3), pp. 597-607.

Jackson, R. J., Hellen, C. U. and Pestova, T. V. (2010) 'The mechanism of eukaryotic translation initiation and principles of its regulation', *Nat Rev Mol Cell Biol,* 11(2), pp. 113-27.

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(8), pp. 2636-43.

Kahvejian, A., Svitkin, Y. V., Sukarieh, R., M'Boutchou, M. N. and Sonenberg, N. (2005) 'Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms', *Genes Dev,* 19(1), pp. 104-13.

Kapp, L. D. and Lorsch, J. R. (2004) 'GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2', *J Mol Biol,* 335(4), pp. 923-36.

Keller, C. I. and Akhtar, A. (2015) 'The MSL complex: juggling RNA-protein interactions for dosage compensation and beyond', *Curr Opin Genet Dev,* 31, pp. 1-11.

Kelley, R. L., Solovyeva, I., Lyman, L. M., Richman, R., Solovyev, V. and Kuroda, M. I. (1995) 'Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in Drosophila', *Cell,* 81(6), pp. 867-77.

Kelley, R. L., Wang, J., Bell, L. and Kuroda, M. I. (1997) 'Sex lethal controls dosage compensation in Drosophila by a non-splicing mechanism', *Nature,* 387(6629), pp. 195-9.

Kieft, J. S. (2008) 'Viral IRES RNA structures and ribosome interactions', *Trends Biochem Sci,* 33(6), pp. 274-83.

Kojima, S., Matsumoto, K., Hirose, M., Shimada, M., Nagano, M., Shigeyoshi, Y., Hoshino, S., Ui-Tei, K., Saigo, K., Green, C. B., Sakaki, Y. and Tei, H. (2007) 'LARK activates posttranscriptional expression of an essential mammalian clock protein, PERIOD1', *Proc Natl Acad Sci U S A,* 104(6), pp. 1859-64.

Kong, J. and Lasko, P. (2012) 'Translational control in cellular and developmental processes', *Nat Rev Genet,* 13(6), pp. 383-94.

Korneeva, N. L., Lamphear, B. J., Hennigan, F. L., Merrick, W. C. and Rhoads, R. E. (2001) 'Characterization of the two eIF4A-binding sites on human eIF4G-1', *J Biol Chem,* 276(4), pp. 2872-9.

Kouba, T., Rutkai, E., Karásková, M. and Valášek, L. (2012) 'The eIF3c/NIP1 PCI domain interacts with RNA and RACK1/ASC1 and promotes assembly of translation preinitiation complexes', *Nucleic Acids Res,* 40(6), pp. 2683-99.

Kozak, M. (1978) 'How do eucaryotic ribosomes select initiation regions in messenger RNA?', *Cell,* 15(4), pp. 1109-23.

Landry, D. M., Hertz, M. I. and Thompson, S. R. (2009) 'RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs', *Genes Dev,* 23(23), pp. 2753-64.

Larschan, E., Bishop, E. P., Kharchenko, P. V., Core, L. J., Lis, J. T., Park, P. J. and Kuroda, M. I. (2011) 'X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila', *Nature,* 471(7336), pp. 115-8.

Lee, A. S., Kranzusch, P. J. and Cate, J. H. (2015) 'eIF3 targets cell-proliferation messenger RNAs for translational activation or repression', *Nature,* 522(7554), pp. 111-4.

LeFebvre, A. K., Korneeva, N. L., Trutschl, M., Cvek, U., Duzan, R. D., Bradley, C. A., Hershey, J. W. and Rhoads, R. E. (2006) 'Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit', *J Biol Chem,* 281(32), pp. 22917-32.

Lin, J. C., Hsu, M. and Tarn, W. Y. (2007) 'Cell stress modulates the function of splicing regulatory protein RBM4 in translation control', *Proc Natl Acad Sci U S A,* 104(7), pp. 2235-40.

Lin, Y. H. and Bundschuh, R. (2015) 'RNA structure generates natural cooperativity between single-stranded RNA binding proteins targeting 5' and 3'UTRs', *Nucleic Acids Res,* 43(2), pp. 1160-9.

Lomakin, I. B. and Steitz, T. A. (2013) 'The initiation of mammalian protein synthesis and mRNA scanning mechanism', *Nature,* 500(7462), pp. 307-11.

López-Bergami, P., Habelhah, H., Bhoumik, A., Zhang, W., Wang, L. H. and Ronai, Z. (2005) 'RACK1 mediates activation of JNK by protein kinase C [corrected]', *Mol Cell,* 19(3), pp. 309-20.

Majzoub, K., Hafirassou, M. L., Meignin, C., Goto, A., Marzi, S., Fedorova, A., Verdier, Y., Vinh, J., Hoffmann, J. A., Martin, F., Baumert, T. F.,

Schuster, C. and Imler, J. L. (2014) 'RACK1 controls IRES-mediated translation of viruses', *Cell,* 159(5), pp. 1086-95.

Mamidipudi, V., Zhang, J., Lee, K. C. and Cartwright, C. A. (2004) 'RACK1 regulates G1/S progression by suppressing Src kinase activity', *Mol Cell Biol,* 24(15), pp. 6788-98.

Masutani, M., Sonenberg, N., Yokoyama, S. and Imataka, H. (2007) 'Reconstitution reveals the functional core of mammalian eIF3', *EMBO J,* 26(14), pp. 3373-83.

Medenbach, J., Seiler, M. and Hentze, M. W. (2011) 'Translational control via protein-regulated upstream open reading frames', *Cell,* 145(6), pp. 902-13.

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(6763), pp. 838-41.

Militti, C., Maenner, S., Becker, P. B. and Gebauer, F. (2014) 'UNR facilitates the interaction of MLE with the lncRNA roX2 during Drosophila dosage compensation', *Nat Commun,* 5, pp. 4762.

Mitchell, S. F., Walker, S. E., Algire, M. A., Park, E. H., Hinnebusch, A. G. and Lorsch, J. R. (2010) 'The 5'-7-methylguanosine cap on eukaryotic mRNAs serves both to stimulate canonical translation initiation and to block an alternative pathway', *Mol Cell,* 39(6), pp. 950-62.

Mothe-Satney, I., Yang, D., Fadden, P., Haystead, T. A. and Lawrence, J. C. (2000) 'Multiple mechanisms control phosphorylation of PHAS-I in five (S/T)P sites that govern translational repression', *Mol Cell Biol,* 20(10), pp. 3558-67.

Muckenthaler, M., Gray, N. K. and Hentze, M. W. (1998) 'IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F', *Mol Cell,* 2(3), pp. 383-8.

Ohlmann, T., Rau, M., Morley, S. J. and Pain, V. M. (1995) 'Proteolytic cleavage of initiation factor eIF-4 gamma in the reticulocyte lysate inhibits translation of capped mRNAs but enhances that of uncapped mRNAs', *Nucleic Acids Res,* 23(3), pp. 334-40.

Passmore, L. A., Schmeing, T. M., Maag, D., Applefield, D. J., Acker, M. G., Algire, M. A., Lorsch, J. R. and Ramakrishnan, V. (2007) 'The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome', *Mol Cell,* 26(1), pp. 41-50.

Pelletier, J. and Sonenberg, N. (1988) 'Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA', *Nature,* 334(6180), pp. 320-5.

Perkins, D. N., Pappin, D. J., Creasy, D. M. and Cottrell, J. S. (1999) 'Probability-based protein identification by searching sequence databases using mass spectrometry data', *Electrophoresis,* 20(18), pp. 3551-67.

Pisarev, A. V., Kolupaeva, V. G., Yusupov, M. M., Hellen, C. U. and Pestova, T. V. (2008) 'Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes', *EMBO J,* 27(11), pp. 1609-21.

Prévôt, D., Décimo, D., Herbreteau, C. H., Roux, F., Garin, J., Darlix, J. L. and Ohlmann, T. (2003) 'Characterization of a novel RNA-binding region of eIF4GI critical for ribosomal scanning', *EMBO J,* 22(8), pp. 1909-21.

Ron, D., Chen, C. H., Caldwell, J., Jamieson, L., Orr, E. and Mochly-Rosen, D. (1994) 'Cloning of an intracellular receptor for protein kinase C: a homolog of the beta subunit of G proteins', *Proc Natl Acad Sci U S A,* 91(3), pp. 839-43.

Salz, H. K. and Erickson, J. W. (2010) 'Sex determination in Drosophila: The view from the top', *Fly (Austin),* 4(1), pp. 60-70.

Schäffler, K., Schulz, K., Hirmer, A., Wiesner, J., Grimm, M., Sickmann, A. and Fischer, U. (2010) 'A stimulatory role for the La-related protein 4B in translation', *RNA,* 16(8), pp. 1488-99.

Shatsky, I. N., Dmitriev, S. E., Andreev, D. E. and Terenin, I. M. (2014) 'Transcriptome-wide studies uncover the diversity of modes of mRNA recruitment to eukaryotic ribosomes', *Crit Rev Biochem Mol Biol,* 49(2), pp. 164-77.

Shatsky, I. N., Dmitriev, S. E., Terenin, I. M. and Andreev, D. E. (2010) 'Cap- and IRES-independent scanning mechanism of translation initiation as an alternative to the concept of cellular IRESs', *Mol Cells,* 30(4), pp. 285-93.

Smith, E. R., Allis, C. D. and Lucchesi, J. C. (2001) 'Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in Drosophila males', *J Biol Chem,* 276(34), pp. 31483-6.

Smith, M. D., Gu, Y., Querol-Audí, J., Vogan, J. M., Nitido, A. and Cate, J. H. (2013) 'Human-like eukaryotic translation initiation factor 3 from Neurospora crassa', *PLoS One,* 8(11), pp. e78715.

Sonenberg, N. and Hinnebusch, A. G. (2007) 'New modes of translational control in development, behavior, and disease', *Mol Cell,* 28(5), pp. 721-9.

Sonenberg, N. and Hinnebusch, A. G. (2009) 'Regulation of translation initiation in eukaryotes: mechanisms and biological targets', *Cell,* 136(4), pp. 731-45.

SPIEGELMAN, S. (1961) 'The relation of informational RNA to DNA', *Cold Spring Harb Symp Quant Biol,* 26, pp. 75-90.

Sun, C., Todorovic, A., Querol-Audí, J., Bai, Y., Villa, N., Snyder, M., Ashchyan, J., Lewis, C. S., Hartland, A., Gradia, S., Fraser, C. S., Doudna, J. A., Nogales, E. and Cate, J. H. (2011) 'Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3)', *Proc Natl Acad Sci U S A,* 108(51), pp. 20473-8.

Szostak, E. and Gebauer, F. (2013) 'Translational control by 3'-UTR-binding proteins', *Brief Funct Genomics,* 12(1), pp. 58-65.

Tcherkezian, J., Cargnello, M., Romeo, Y., Huttlin, E. L., Lavoie, G., Gygi, S. P. and Roux, P. P. (2014) 'Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator of 5'TOP mRNA translation', *Genes Dev,* 28(4), pp. 357-71.

Thompson, S. R. (2012) 'So you want to know if your message has an IRES?', *Wiley Interdiscip Rev RNA,* 3(5), pp. 697-705.

Valásek, L. S. (2012) "Ribozoomin'--translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs)', *Curr Protein Pept Sci,* 13(4), pp. 305-30.

Villa, N., Do, A., Hershey, J. W. and Fraser, C. S. (2013) 'Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome', *J Biol Chem,* 288(46), pp. 32932-40.

Wolf, A. S. and Grayhack, E. J. (2015) 'Asc1, homolog of human RACK1, prevents frameshifting in yeast by ribosomes stalled at CGA codon repeats', *RNA,* 21(5), pp. 935-45.

Xue, S., Tian, S., Fujii, K., Kladwang, W., Das, R. and Barna, M. (2015) 'RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation', *Nature,* 517(7532), pp. 33-8.

Yaka, R., Phamluong, K. and Ron, D. (2003) 'Scaffolding of Fyn kinase to the NMDA receptor determines brain region sensitivity to ethanol', *J Neurosci,* 23(9), pp. 3623-32.

Yaka, R., Thornton, C., Vagts, A. J., Phamluong, K., Bonci, A. and Ron, D. (2002) 'NMDA receptor function is regulated by the inhibitory scaffolding protein, RACK1', *Proc Natl Acad Sci U S A,* 99(8), pp. 5710-5.

Zhang, F., Saini, A. K., Shin, B. S., Nanda, J. and Hinnebusch, A. G. (2015) 'Conformational changes in the P site and mRNA entry channel

evoked by AUG recognition in yeast translation preinitiation complexes', *Nucleic Acids Res,* 43(4), pp. 2293-312.

Zhou, C., Arslan, F., Wee, S., Krishnan, S., Ivanov, A. R., Oliva, A., Leatherwood, J. and Wolf, D. A. (2005) 'PCI proteins eIF3e and eIF3m define distinct translation initiation factor 3 complexes', *BMC Biol,* 3, pp. 14.

ANNEX I

Table 1. List of candidates enriched in the GST-SXL pull-down compared to the GST-loop1-S1 pull-down described in Results part 2.

Accession	DROME annotation	log2 WT area	log2 loop1-S1 area	log2 FC WT/S1	P val
P21187	PABP	32,38	30,86	1,524	0,0049
Q94901	LARK	28,35	26,86	1,493	0,0175
P41094	RS18	27,25	26,54	0,715	0,0251
Q9VWG3	RS10B	28,37	26,65	1,721	0,0269
Q9U3Z7	NH2L1	27,90	26,25	1,649	0,0272
P26686	SRR55	26,84	26,02	0,820	0,0294
P09180	RL4	27,46	27,07	0,388	0,0315
Q8T8R1	Y3800	25,27	24,60	0,670	0,0335
P19889	RLA0	28,61	27,84	0,765	0,0403
P55841	RL14	27,38	26,51	0,871	0,0465
O44437	SMD3	27,66	26,99	0,675	0,0488
P08570	RLA1	29,91	29,34	0,574	0,0494
Q9VXE0	RUXG	27,88	27,07	0,808	0,0588
Q9VEB3	RPF2	26,10	25,43	0,671	0,0680
O61345	PEN	26,02	24,61	1,411	0,0695
P48588	RS25	27,03	26,24	0,786	0,1003
Q9VLV5	RUXE	27,17	26,63	0,535	0,1117
Q9V4M2	WECH	26,82	26,23	0,597	0,1382
Q86B79	UNK	24,89	24,14	0,748	0,1578
Q9VTP4	R10AB	27,37	26,73	0,640	0,1805
P17704	RS17	26,01	25,67	0,336	0,1838
Q24154	RL29	26,50	24,66	1,841	0,2354
Q9VJY6	RL24	26,89	26,30	0,584	0,2677
P31009	RS2	27,14	26,93	0,217	0,2709
P0CG69	UBIQP	26,46	25,29	1,165	0,2887
Q9V4Q8	RU2A	26,60	26,24	0,368	0,3234
O16797	RL3	26,67	26,38	0,285	0,3732

Accession	DROME annotation	log2 WT area	log2 loop1-S1 area	log2 FC WT/S1	P val
Q9VI10	**SMD2**	27,67	27,30	0,364	0,3781
P84051	**H2A**	27,45	25,63	1,817	0,4080
Q9VMA3	**CUP**	26,36	26,08	0,287	0,4231
Q9VZS5	**RL28**	26,08	25,64	0,442	0,4834
P46223	**RL7A**	28,92	28,61	0,310	0,5117
P29310	**1433Z**	25,51	25,16	0,352	0,5123
Q9V3G1	**RL8**	27,73	27,46	0,269	0,5250
Q8MLY8	**RS8**	27,12	26,98	0,140	0,5546
P39018	**RS19A**	26,12	25,93	0,186	0,5888
O18640	**GBLP**	27,75	27,52	0,225	0,5919
P19339	**SXL**	**36,71**	**36,62**	**0,097**	**0,6132**
P80455	**RS12**	27,92	27,59	0,331	0,6174
P36179	**2AAA**	24,32	23,95	0,364	0,6337
Q9V3G3	**PPIE**	24,91	24,29	0,614	0,6375
Q9V3Z6	**MYO7A**	30,06	29,48	0,577	0,6538
Q9VKK1	**EDC4**	27,41	27,31	0,098	0,7091
P06606	**TBA4**	27,13	26,94	0,199	0,7198
P52034	**K6PF**	34,67	34,41	0,257	0,7260
Q7KLV9	**rack1**	25,30	25,13	0,166	0,7282
P06607	**VIT3**	26,66	26,26	0,399	0,7402
P61857	**TBB2**	28,57	28,49	0,077	0,7553
Q9W1V3	**FBRL**	27,15	27,02	0,136	0,7739
Q9W237	**RS16**	26,71	26,43	0,277	0,7796
P29327	**RS6**	27,01	26,95	0,060	0,8227
P04359	**RL32**	26,84	26,78	0,061	0,8316
Q0E9B6	**RS11**	26,15	26,06	0,089	0,9381
Q9VPC0	**KP58**	25,46	25,43	0,030	0,9495
Q05856	**RSMB**	25,91	*	*	*
Q9VKQ3	**WDR12**	25,36	*	*	*
Q9VAW5	**LARP**	26,60	*	*	*
Q7K0Y1	**BOP1**	24,60	*	*	*

Accession	DROME annotation	log2 WT area	log2 loop1-S1 area	log2 FC WT/S1	P val
Q9VLK2	Y3096	25,26	*	*	*
Q9VJZ7	RRP1L	25,27	*	*	*
Q9BJZ5	FUSIL	27,15	*	*	*
Q9VM75	HEAT1	26,84	*	*	*
Q9W3C1	U202	23,87	*	*	*
Q06559	RS3	25,81	*	*	*
Q9VYS3	RENT1	26,14	*	*	*
P39736	BX42	26,38	*	*	*
Q9V5Q4	SYF2	25,15	*	*	*
P55828	RS20	26,57	*	*	*
P55935	RS9	25,78	*	*	*

Table legend:

The amount of protein present in the pull-downs is estimated by the Log2 of the area of the 3 most abundant peptides detected by mass-spectrometry. The values represented here as "log2 WT" or "log2 loop1-S1" correspond to the average of 4 replicates. The p-value was calculated using a t-student test. Candidates are listed from the lower to the higher p-value. The asterisks denote values that could not be calculated because the protein was absent in the loop1-S1 pull-down. These values refer to proteins absent in at least three of the loop1-S1 replicates but present in at least 3 of the WT replicates. Proteins significantly enriched in the GST-WT pull-down are in red letters. Proteins of interest mentioned in this work are highlighted with a light blue background. SXL is written in white letters.

ANNEX II

Manuscript:

Hrp48 targets eIF3d to repress *msl2* mRNA translation

Emilia Szostak, Marina García-Beyaert, Antoine Graindorge and Fátima Gebauer. *Under revision*.

I have contributed to this manuscript with Figures 1B, 4A and 5B

Hrp48 targets eIF3d to repress *msl2* mRNA translation

Emilia Szostak, Marina García-Beyaert, Antoine Graindorge[1] and Fátima Gebauer[2]

Gene Regulation, Stem Cells and Cancer Programme, Centre for Genomic Regulation (CRG), Dr Aiguader 88, 08003-Barcelona, Spain

Universitat Pompeu Fabra (UPF), 08003-Barcelona, Spain

(1) Current address: Institut Curie, INSERM U934/CNRS UMR3215, 26 Rue d'Ulm, 75005 Paris, France

(2) Corresponding author:

Tel: +34-93 3160120

Fax: +34-93 3969983

E-mail: fatima.gebauer@crg.eu

Running title: Hrp48 inhibits translation by targeting eIF3d

Keywords: Translation, Hrp48, eIF3d, msl2, SXL, UNR

SUMMARY

Translational repression of *msl2* mRNA in females of *Drosophila melanogaster* is an essential step in the regulation of X-chromosome dosage compensation. Repression is orchestrated by Sex-lethal (SXL), which binds to both untranslated regions (UTRs) of *msl2* and inhibits translation initiation by poorly understood mechanisms. Here we identify Hrp48 as a SXL co-factor. Hrp48 binds to a critical region in the 3' UTR of *msl2* and interacts with known components of the repressor complex. Depletion of Hrp48 decreases the ability of SXL to repress translation. A proteomics search for factors associated to Hrp48 identified eIF3d, a non-core subunit of the eIF3 translation initiation complex. Reporter assays showed that eIF3d is specifically required for translation of *msl2* mRNA and is a target of SXL-mediated translational repression. These data uncover a novel co-factor in *msl2* regulation, its target in the translational machinery, and illustrate how a general translation initiation factor contributes to mRNA-specific regulation.

INTRODUCTION

Post-transcriptional regulation plays an important role in numerous biological situations. In particular, translational control by RNA binding proteins pervades embryonic development and adult cell homeostasis, yet few mechanisms of translational control have been described to date (Sonenberg and Hinnebusch, 2009; Lasko, 2011; Szostak and Gebauer, 2013; Blackinton and Keene, 2014). Dosage compensation, a mechanism that equalizes the expression of X-linked genes in species where genders carry a different number of X chromosomes, depends on translational regulation (Graindorge et al., 2011). In *Drosophila melanogaster*, males (XY) hyper-transcribe the single X-chromosome about 2-fold by the activity of the chromatin remodeling dosage compensation complex (DCC). While active in males, dosage compensation is repressed in females via a complex set of post-transcriptional events orchestrated by the female-specific RNA binding protein Sex-lethal (SXL). These events are directed to repress the expression of the limiting DCC subunit MSL2. SXL, a primarily nuclear protein, first inhibits the splicing of a facultative intron in the 5' UTR of *msl2* pre-mRNA

and promotes nuclear retention of *msl2* transcripts (Merendino et al., 1999; Forch et al., 2001; Graindorge et al., 2013). SXL then inhibits *msl2* translation in the cytoplasm by binding to both the 5' and 3' UTRs of *msl2* transcripts (Bashaw and Baker, 1997; Kelley et al., 1997; Gebauer et al., 1998). SXL bound to the 3' UTR recruits the co-repressor UNR to inhibit initial ribosome binding (Gebauer et al., 2003; Abaza et al., 2006; Duncan et al., 2006; Hennig et al., 2014). SXL bound to the 5' UTR inhibits ribosomal scanning by a mechanism that promotes the usage of an upstream ORF (Beckmann et al., 2005; Medenbach et al., 2011). The targets in the translation machinery for either 5' or 3'- mediated regulation are unknown.

Here we focus on 3' UTR-mediated regulation of *msl2* mRNA. We found that the 3' UTR minimal region required for translational repression contains binding sites for regulators distinct from SXL and UNR. Using a combination of GST pull-down and RNA affinity binding (GRAB, Graindorge et al., 2013), we have identified Hrp48 as a novel *msl2* regulator. A search for Hrp48 interactors identified eIF3d, a non-core subunit of the translation

initiation factor eIF3. eIF3d is specifically required for *msl2* translation, and is necessary for SXL-mediated translational repression. Our results identify a new co-factor in *msl2* regulation and its target in the translational machinery, and illustrate how general translation initiation factors can contribute to mRNA-specific regulation.

RESULTS

Identification of Hrp48 as a candidate factor for *msl2* translational regulation

The *msl2* transcript contains long 5' and 3' UTRs with multiple uridine stretches that serve as SXL-binding sites (sites A-F, Figure 1A). Sites B, E and F (black boxes) are required for translational repression, while sites A, C and D (grey boxes) are dispensable. The minimal region of the 3' UTR required for regulation consists of 46 nucleotides containing sites E and F, each followed by a UNR-binding site (blue). Sequences downstream of these sites are also important for translational repression (Gebauer et al., 2003). We performed *in vitro* translation assays in *Drosophila* embryo extracts

to thoroughly evaluate the relevance of these sequences. The translation of *msl2* reporters was assessed in extracts supplemented with increasing amounts of a recombinant SXL derivative that is fully competent in translational repression (dRBD4, Grskovic et al., 2003) (Figure 1B). Substitution of the 8-nucleotide segment immediately downstream of the SXL and UNR binding sites (Figure 1A, green) by $(CU)_4$ was equivalent to deletion of the 3' UTR (Figure 1B, left panel, compare 5m and Δ3′ constructs). To evaluate the contribution of this region, hereafter termed "region 5", to 5' or 3' UTR- mediated regulation, we tested the region 5 mutant (5m) in the framework of transcripts that lacked SXL-binding sites in either UTR. While mutation of region 5 was irrelevant for 5' UTR-mediated regulation (Figure 1B, right panel), we observed that 3' UTR-driven repression was less efficient in 5m mutants (middle panel), indicating that region 5 plays an important role in *msl2* silencing mediated by the 3′ UTR.

One possibility to explain these effects is that region 5 contributes to SXL and UNR binding in the 3' UTR. We therefore tested the binding of recombinant UNR and dRBD4 to 3' UTR fragments

containing or lacking region 5 using gel-mobility shift assays (Figure 1C). dRBD4 binds with high affinity to a wild type 3' UTR *msl2* fragment (lanes 1-5). As previously reported, UNR also binds with high affinity after addition of dRBD4 (lanes 6-8) (Abaza et al., 2006; Hennig et al, 2014). Binding of both proteins was not affected by region 5 mutation (lanes 9-16), indicating that region 5 is not required for binding of purified SXL and/or UNR and, thus, may contribute to *msl2* repression by binding to a novel regulator.

To identify factors binding to region 5 that could potentially contribute to translational repression we used a technology previously optimized in our laboratory termed GRAB (Graindorge et al., 2013). GRAB allows for the enrichment of SXL RNPs from a mix containing recombinant GST-SXL derivatives, biotinylated *msl2* mRNA and *Drosophila* embryo extracts following two affinity purification steps: i) GST-pull down and elution with a protease that separates the GST moiety, and ii) RNA affinity purification using streptavidin beads (Figure 2A). We compared the GRAB profiles obtained using wild type (WT) or 5m 3' UTR fragments together with GST-dRBD4 (Figure 2B). A protein of 48 kDa was present in

the WT but absent in the 5m eluates (Figure 2C). Mass spectrometry analysis identified this protein as the Heterogeneous Ribonucleoprotein 48 (Hrp48, also known as Hrb27C). A group of ribosomal proteins (S10b, S17, S18, S19a) was also reproducibly absent in the 5m profile while, as expected, UNR and dRBD4 bound equally to both WT and 5m RNAs. We focused on Hrp48, as this protein has been previously shown to participate in translational regulation of *oskar* and other *Drosophila* mRNAs by mechanisms that are poorly understood (Yano et al., 2004; Nelson et al., 2007; Suissa et al., 2010).

Independent RNA chromatography experiments confirmed that Hrp48 binds to WT but not 5m RNA (Figure 2D). To test whether binding of Hrp48 to *msl2* 3' UTR depends on prior SXL or UNR binding, we performed RNA chromatography experiments using SXL derivatives that either support (dRBD4) or not (hRBD1, mRBD) binding of UNR (see legend of Figure 2B for details) (Grskovic et al, 2003), or an *msl2* construct lacking sites E and F (EFm) which binds neither SXL nor UNR (Gebauer et al., 2003). The results indicated that Hrp48 binding was efficient in all cases

(Figure 2E). This contrasts with UNR, which is strictly dependent on SXL for *msl2* RNA recognition. Therefore, Hrp48 binds to region 5 independently of known *msl2* translational regulators.

Hrp48 is required for *msl2* translational repression

We next performed functional assays to determine whether Hrp48 is involved in *msl2* translational regulation. We used male SL2 cells, which lack endogenous SXL and where transfection of a SXL-encoding plasmid allows for a tight control of SXL levels. We first depleted Hrp48 from SL2 cells and then tested the ability of exogenous SXL to inhibit the translation of a reporter containing full length *msl2* 5' and 3' UTRs. Depletion of Hrp48 indeed reduced the capacity of SXL to inhibit translation of the reporter (Figure 3A). Similar results were obtained *in vitro*, after depletion of Hrp48 from *Drosophila* embryo extracts using an oligonucleotide containing Hrp48 binding sites. Repression of the WT RNA reporter was less efficient in the Hrp48-depleted extract compared to control or untreated extracts (Figure 3B, WT RNA). To evaluate whether Hrp48 functions through region 5, the effect of Hrp48 depletion on 5m RNA was tested. As expected, 5m RNA was less

efficiently repressed in untreated or control extracts compared to WT RNA (Figure 3B, 5m). Importantly, depletion of Hrp48 had no effect on 5m RNA, as 5m RNA was repressed equally less efficiently in all extracts. Altogether, these results indicate that Hrp48 is required for optimal translational repression of *msl2* by SXL, and that the effect of Hrp48 is mediated by region 5.

Hrp48 interacts with components of the repressed *msl2* mRNP

To determine whether Hrp48 interacts with components of the repressed *msl2* mRNP, we first tested binding of Hrp48 to endogenous *msl2* by immunoprecipitation followed by semi-quantitative PCR. The results showed that Hrp48 interacts with *msl2* mRNA (Figure 4A). We then used co-immunoprecipitation assays to test if Hrp48 interacts with the *msl2* mRNP components UNR, SXL and HOW. As mentioned above, UNR and SXL are involved in *msl2* translational regulation, while HOW binds to the *msl2* 5' UTR intron and promotes nuclear pre-mRNA retention (Graindorge et al., 2013). We found that Hrp48 interacts with SXL and HOW in an RNA-dependent manner, whereas Hrp48 interacts

with UNR in an RNA-independent fashion (Figure 4B). These results suggest that Hrp48 is part of the repressed *msl2* mRNP.

Hrp48 represses translation by interfering with eIF3d

To gain insight into the molecular mechanism used by Hrp48 to repress *msl2* translation, we searched for novel Hrp48 interactors in *Drosophila* embryo extracts using oligonucleotide pull-down. The Hrp48 binding oligonucleotide was used in three independent replicates, in parallel with poly(C) as control. Efficient and specific Hrp48 pull-down was assessed by Western blot (Figure 5A, bottom panel). Mass-spectrometry analysis revealed a number of proteins significantly enriched in the Hrp48 oligo eluate (Figure 5A and Table S1). The presence of HOW and Squid in the group of co-purified proteins validate the use of oligo pull-down to reveal *bonafide* interactors of Hrp48. Indeed, Squid is a partner of Hrp48 in the regulation of *gurken* mRNA expression and HOW is a component of the msl2 mRNP that co-immunoprecipitates with Hrp48 (Figure 4B) (Goodrich et al., 2004; Geng and Macdonald, 2006; Graindorge et al, 2013). Interestingly, the subunit d of eIF3 was also enriched in the Hrp48 oligo eluate. This interaction was

confirmed by independent co-immunoprecipitation analysis using epitope-tagged eIF3d (Figure 5B).

eIF3d is a non-core component of the multi-subunit translation initiation factor eIF3, a factor that orchestrates translation initiation (Hinnebusch, Ann Rev Biochem 2014). We thus hypothesized that Hrp48 might interfere with translation initiation by targeting eIF3d. To test this hypothesis, we first examined whether eIF3d was specifically required for *msl2* translation. Depletion of eIF3d from SL2 cells reduced translation of an *msl2* reporter, while depletion of other non-core subunits (eIF3e, eIF3h) had no effect (Figure 5C). Importantly, eIF3d depletion did not cause major defects in cellular translation (Fig 5D) nor affected cell viability (data not shown). These results indicate that eIF3d promotes the translation of specific transcripts like *msl2* mRNA.

To test whether eIF3d was a target for *msl2* regulation, we measured the ability of SXL to repress the translation of the full-length *msl2* reporter after depletion of eIF3d from SL2 cells. Depletion of eIF3d, but not eIF3e or eIF3h, reduced SXL-mediated

repression, indicating that eIF3d is a target of the repressor complex within the translational machinery (Figure 5E).

DISCUSSION

The regulation of *msl2* expression is a prime example of translational control where binding of RNA binding proteins to both UTRs of the transcript coordinate mRNA output. Despite efforts to unravel the *msl2* regulatory mechanism, the full set of regulators and how they interact with the translational machinery have remained elusive. Here we identify Hrp48 as a novel *msl2* regulator, and eIF3d as its target in the protein synthesis machinery.

Hrp48 is an hnRNP A/B family member involved in post-transcriptional regulation at multiple levels, including splicing, mRNA localization and translation (Blanchette et al., 2009; Yano et al., 2004; Huyhn et al., 2004). We show here that Hrp48 inhibits *msl2* translation by interacting with the 3' UTR of the transcript in complex with SXL and UNR (Figures 1-4). Intriguingly, depletion of Hrp48 from female Kc cells augmented SXL expression (data not

shown), consistent with previous reports indicating that Hrp48 inhibits SXL expression in monomorphic tissues of female flies (Suissa et al., 2010). These data suggest the existence of cross-regulatory loops to maintain the balance of the components of the *msl2* repressor complex. These effects compromised the assessment of the role of Hrp48 on *msl2* regulation in the fly. However, it has been shown that Hrp48 loss causes sex-specific defects: hypomorph females display strong lethality while males are less affected (Hammond et al., 1997). We propose that regulation of *msl2* expression by Hrp48 contributes to the observed female-specific lethality.

Hrp48 participates in the localization and translational regulation of *oskar* and *gurken* mRNAs during *Drosophila* early embryogenesis by mechanisms that are poorly understood (Yano et al., 2004; Huynh et al., 2004; Goodrich et al., 2004; Geng and Macdonald, 2006). In the case of *oskar*, Hrp48 binds to both UTRs of the transcript and represses *oskar* translation during transport (Yano et al., 2004). In the case of *msl2*, Hrp48 binds to the 3' UTR, although indirect contacts with the 5' UTR via interactions with HOW –a 5'

UTR binding factor (Graindorge et al., 2013)- could be evoked (Figure 4). Hrp48 and its mammalian ortholog DAZAP1 have also been shown to stimulate translation of certain transcripts (Nelson et al., 2007; Smith et al., 2011). DAZAP1 modulates translation initiation downstream of the recognition of the 5' cap structure by initiation factors (Smith et al, 2011). As Hrp48 interacts with a subunit of eIF3 (Figure 5), our data raise the possibility that DAZAP1 exploits eIF3 interactions to stimulate translation.

eIF3 is involved in practically all steps of translation initiation, as it controls the formation of the 43S pre-initiation complex, the binding of this complex to the mRNA and the stringency of start codon selection (reviewed in Hinnebusch, 2014). The composition of eIF3 varies across species (Smith et al., 2013). In metazoa, eIF3 consists of 13 subunits, named eIF3a to eIF3m, and the specific contributions of these subunits to the variety of eIF3 functions has only started to be elucidated. Although subunit eIF3d is not conserved in budding yeast and is not part of the functional (abcefh) or the structural (acefhklm) eIF3 core, it is essential in some organisms (Smith et al., 2013; Masutani et al., 2007; Sun et al., 2011). eIF3d was found in a screen for factors required to cope with

stress in *S. pombe*, suggesting that it may contribute to translation of important mRNAs (Calvo et al., 2009). In addition, eIF3d is necessary for translation initiation *in vitro*, and is required for efficient binding of eIF3 to eIF4G, an initiation factor that is part of the mRNA cap-binding complex and participates in ribosome recruitment to the mRNA (Sun et al., 2011; Villa et al., 2013). Although a structure of eIF3d on the ribosome is still lacking, a recent model has placed eIF3d close to the mRNA exit channel in the 40S ribosomal subunit, consistent with cross-linking of this factor to mRNA at several positions upstream of the start codon in reconstituted mammalian pre-initiation complexes (Pisarev et al., 2008; Erzberger et al., 2014). This is a suitable location for eIF3d to mediate interactions with eIF4G during recruitment of the ribosome to the mRNA, because eIF4G is also located in the vicinity of the mRNA exit channel (reviewed in Valasek, 2012). Targeting of eIF3d, therefore, fits with the proposed role of the *msl2* 3'UTR complex in inhibiting ribosome recruitment (Gebauer et al., 2003).

A recent report showed that mammalian eIF3d binds directly to the 5' UTR of specific mRNAs during cap-dependent translation (Lee

et al., 2015). Our results show that eIF3d indeed regulates translation in a message-specific manner and that it can be targeted by RNA binding proteins for regulation.

PABP has previously been implicated in the *msl2* inhibitory mechanism (Duncan et al., 2009). PABP interacts with UNR and is required for optimal repression of polyadenylated *msl2* transcripts. The mechanism does not preclude binding of PABP to the mRNA, and it is currently unclear whether PABP is a partner or a target in repression. The observation that non-adenylated transcripts are also repressed efficiently suggests alternative or complementary scenarios for *msl2* regulation (Gebauer et al., 1999). eIF3d provides such an alternative. Further studies are required to analyze the interplay between PABP and eIF3d to repress *msl2* translation in female flies.

MATERIALS AND METHODS

Plasmids

Plasmids used for the synthesis of the WT, Δ3'and 5m *msl2* reporters were described previously (Gebauer et al., 2003). Plasmids WT and 5m have been renamed in this work for simplicity (named 3'EF and mut5 in Gebauer et al. 2003, respectively). Plasmid (AB)m-5m was obtained by replacing the 5′ UTR of 5m with the full-length (626 nt) 5′ UTR of *msl2* containing mutated SXL binding sites. Plasmid (AB)m is as (AB)m-5m but contains wild type 3′ UTR sequences. Plasmid uORF-BL(EF)m is as described in Medenbach et al. (2011), and contains wild type nucleotides 270-339 of *msl2* 5′UTR including AUG_3 . Plasmid uORF-BL(EF5)m is as uORF-BL(EF)m but contains a substitution of region 5 by $(CT)_8$.

Plasmids used for the generation of biotinylated probes were obtained by insertion of hybridized complementary oligonucleotides containing nucleotides 909-954 of the *msl2* 3' UTR or derivatives into pBluescript, as described in Grskovic et al (2003).

The FC *msl2* reporter and pAc-Renilla constructs used for transfection were previously described (Bashaw and Baker, 1997; Graindorge et al., 2013). To obtain pAc-SXL, the SXL ORF was amplified by PCR and cloned into the EcoRI and XhoI sites of pAc5.1B-NheI plasmid.

Plasmids used for expression of UNR and recombinant SXL variants (pGEX-dRBD4, pGEX-mRBD and pGEX-hRBD1) have been described (Grskovic et al. 2003; Abaza et al., 2006). The plasmid used for the expression of recombinant Hrp48 was obtained by cloning the Hrp48 ORF into the NdeI and XhoI sites of pET15b.

Recombinant proteins

SXL derivatives dRBD4 (amino acids 122-301 of *D. melanogaster* SXL), mRBD (amino acids 99-271 of *Musca domestica* SXL) and hRBD1 (a hybrid containing RRM1 of *M. domestica* and RRM2 of *D. melanogaster* SXL) were expressed in *E. coli* as N-terminal GST-tagged fusions and purified as described (Grskovic et al. 2003). His-dRBD4 and His-FLAG-tagged, full length UNR were purified according to the pET systems user's manual, with a second purification step for UNR using FLAG columns (Novagen). All

proteins were dialyzed against buffer D (20mM HEPES pH 8.0, 20% glycerol, 1mM DTT, 0.01% NP-40, 0.2 mM EDTA). Hrp48 was expressed as a His-tagged fusion and used to generate antibodies in rabbits. Hrp48 was largely insoluble, and thus to increase solubility, Hrp48-transformed *E. coli* were induced with 1 mM IPTG for 3h at 30°C, and purification was performed under 6M urea.

RNA synthesis

Biotinylated RNAs were synthesized using the MEGAshort script kit (Ambion), adding bio-14-CTP (Invitrogen) at an equimolar ratio with CTP in the reaction, and were purified using G25 columns (GE Healthcare). Radiolabeled *msl2* probes used in gel-mobility shift assays were prepared by *in vitro* transcription from hybridized oligonucleotide templates containing the T7 promoter followed by the relevant *msl2* sequences. Both biotinylated and radiolabeled probes contained an ApppG cap (KEDAR). dsRNA for depletion experiments was synthesized using the MEGA script kit (Ambion). The synthesis of mRNAs used in *in vitro* translation reactions was

performed as described (Gebauer et al. 1999). All mRNAs contained a ^{7m}GpppG cap and a poly(A) tail of 73 residues.

Gel mobility shift assays

Radiolabelled *msl2* probes were incubated with increasing amounts of recombinant dRBD4 and UNR as described previously (Abaza et al., 2006). RNA-protein complexes were resolved in non-denaturing 4% polyacrylamide gels.

Antibodies and immunoprecipitation

Antibodies against full-length Hrp48 were generated in rabbits and characterized by Western blot and immunoprecipitation of Hrp48 from *Drosophila* embryo extracts. Anti-Hrp48 antibodies used in initial experiments were kindly provided by Anne Ephrussi (Yano et al., 2004) and Marco Blanchette (Blanchette et al., 2005). Anti-UNR, anti-SXL and anti-HOW antibodies were previously described (Abaza et al. 2006; Graindorge et al. 2013). Anti-tubulin antibody (Sigma, T6199) was provided commercially.

Immunoprecipitation was performed using antibodies covalently cross-linked to protein A Dynabeads (Invitrogen). Four milligrams

of *Drosophila* embryo extract were incubated with 120ul of protein A Dynabeads slurry containing the appropriate antibody, complete protease inhibitor cocktail (Roche), and PBS pH8.0 to a final volume of 600ul. After 1h of incubation at 4°C, washes were performed with 15x bead volumes of cold PBS supplemented with 10% glycerol. Beads were then treated with a mix of 40 ug RNase A and 10-20 units RNase ONE (Promega) in RNase ONE buffer, or with buffer alone in a total volume of 100 ul, for 30 min at 37°C. The supernatant was removed and the beads were washed with 15 volumes of cold PBS pH8.0. Proteins were recovered with 2x Laemmli buffer and resolved by SDS-PAGE.

RNAi, transfections, reporter activity assays and ^{35}S-Methionine labeling

RNA interference and transfections were performed as described previously (Graindorge et al. 2013). The βgal FC reporter construct was co-transfected with pAc-Renilla control and increasing amounts of pAc-SXL. Renilla luciferase and βgalactosidase activities were measured with luciferase (Promega) and Galacto-

Star (Tropix) kits, respectively. βgal activity was corrected for co-transfected Renilla, and normalized for reporter RNA levels.

For ^{35}S-Methionine labeling, cells were seeded in 6-well plates and depletion of eIF3 subunits was performed for 3 days. Cells were then washed with fresh Schneider's medium, and incubated for 2h at 25°C with 2 µl ^{35}S-meth 10mCi/ml in the medium. After incubation, cells were washed with cold PBS and lysed with passive lysis buffer (Promega). Samples were quantified, resolved on SDS-PAGE, and visualized by Coomassie staining and exposure to Phosphorimager.

RNA extraction and quantification

Total RNA was extracted from SL2 cells using Trizol (Invitrogen) following the manufacturer's instructions. RNA was treated with Turbo DNase (Ambion) and reverse-transcribed using random primers and SuperScript II (Invitrogen). cDNA was amplified by qPCR with SYBR Green (Roche). Reporter β*gal* RNA levels were normalized for co-transfected *Renilla* luciferase RNA levels.

GRAB and RNA affinity chromatography

GRAB was performed as described previously (Graindorge et al. 2013). For direct RNA affinity chromatography, 120 ul of streptavidin Dynabead slurry (Invitrogen) were pre-blocked for 10 min with 100 ng/µl tRNA in binding buffer (5 mM Tris pH7.4, 0.5 mM EDTA, 1M NaCl). Beads were washed with binding buffer and subsequently incubated with 400 pmol of biotinylated RNA for 30min at room temperature. Beads were then washed with 20 bead volumes of ice-cold TCB (17 mM creatine phosphate, 80 ng/ul creatine kinase, 25 mM HEPES pH 8.0, 0.6 mM $Mg(OAc)_2$, 80 mM KOAc). Washed beads were mixed with 10 mg of *Drosophila* embryo extract and 300 pmol GST-dRBD4 or derivatives, in TCB supplemented with 100 U RNAsin (Promega) and 1x Complete protease inhibitor cocktail (Roche). The mix was incubated for 1h at 4°C. Beads were subsequently washed with 15 bead volumes of cold TCB supplemented with 10% glycerol and 0,01% Triton X-100. RNA bound proteins were recovered by elution with 50 µl RNase mix (10-20U RNase ONE and 40 µg RNase A in RNase ONE buffer) after incubation for 30 min at 37°C. Eluted proteins were resolved by SDS-PAGE and analyzed by immunoblotting.

Extract depletion and *in vitro* translation assays

Two hundred µl of streptavidin Dynabeads were pre-blocked with *Drosophila* embryo extract for 1h at 4°C, washed with binding buffer (see above) and further incubated with a 5' biotinylated RNA oligomer containing two Hrp48 binding sites (bio-ACCACCUAGGAUUAAGACCUAGGAUUAAG) or with bio-polyC as control. After incubation for 30min at room temperature, beads were washed with 15 volumes of 20 mM Hepes pH 7.4, and divided in 15 aliquots. Embryo extract (1 mg diluted with 60 mM Hepes pH 7.4 in a ratio of 2:1) was passed sequentially from one aliquot to the other, after incubation with each aliquot for 10 min at 4°C (i.e. a total of 15 rounds of depletion). The efficiency of depletion was tested by Western blot. Depleted extracts were then used for *in vitro* translation as described (Gebauer et al. 1999).

Oligonucleotide pull-down

Groups of 15 bead aliquots from independent depletion experiments (see above for depletion protocol) were pooled and washed 4 times with 1 ml of ice-cold 20 mM HEPES pH 7.4. Beads were then

resuspended in 100 µl RNase mix (10-20 U RNase ONE and 40 µg RNase A in RNase ONE buffer) and incubated at 37°C for 30 min. The supernatant was recovered and analyzed by quantitative mass spectrometry.

ACKNOWLEDGMENTS

We thank Jan Medenbach for the eIF3d plasmid and for useful discussions. We are greatful to Marco Blanchette and Anne Ephrussi for kindly providing anti-Hrp48 antibodies, and to Marco Blanchette for advice on Hrp48 purification. We thank the CRG Protein Service and the CRG-UPF Proteomics Facility for protein production and identification, and Juan Valcárcel for critically reading this manuscript. M.G-B was supported by a fellowship from La Caixa Foundation. A.G. was supported by a fellowship from the Fondation pour la Recherche Médicale (FRM). This work was supported by Spanish Ministry of Economy and Competitiveness MINECO and the European Regional Development Fund (ERDF) under the grant number BFU2012-37135 and Consolider grant CSD2009-00080. We acknowledge support of the Spanish Ministry

of Economy and Competitiveness, 'Centro de Excelencia Severo Ochoa 2013-2017', SEV-2012-0208.

REFERENCES

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.

Bashaw, G.J., Baker, B.S. 1997. The regulation of the Drosophila msl-2 gene reveals a function for Sex-lethal in translational control. *Cell* **89:** 789-798.

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

Blackinton, J.G., Keene, J.D. 2014. Post-transcriptional RNA regulons affecting cell cycle and proliferation. *Semin Cell Dev Biol* **34:** 44-54.

Blanchette, M., Green, R.E., Brenner, S.E., Rio, D.C. 2005. Global analysis of positive and negative pre-mRNA splicing regulators in Drosophila. *Genes Dev* **19:** 1306-1314.

Blanchette, M., Green, R.E., MacArthur, S., Brooks, A.N., Brenner, S.E., Eisen, M.B., Rio, D.C. 2009. Genome-wide analysis of alternative pre-mRNA splicing and RNA-binding specificities of the Drosophila hnRNP A/B family members. *Mol Cell* **33:** 438-449.

Calvo, I.A., Gabrielli, N., Iglesias-Baena, I., García-Santamarina, S., Hoe, K.L., Kim, D.U., Sansó, M., Zuin, A., Pérez, P., Ayté, J., Hidalgo, E. 2009. Genome-wide screen of genes required for caffeine tolerance in fission yeast. *PLoS One* **4:** e6619.

Duncan, K., Grskovic, M., Strein, C., Beckmann, K., Niggeweg, R., Abaza, I., Gebauer, F., Wilm, M., Matthias, W. 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., Hentze, M.W. 2009. The SXL-UNR corepressor complex uses a PABP-mediated mechanism to inhibit ribosome recruitment to msl-2 mRNA. *Mol Cell* **36:** 571-582.

Erzberger, J.P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C.H., Cimermančič, P., Boehringer, D., Sali, A., Aebersold, R., Ban, N. 2014. Molecular architecture of the 40S.eIF1.eIF3 translation initiation complex. *Cell* **158:** 1123-1135.

Forch, P., Merendino, L., Martinez, C., Valcarcel, J. 2001. Modulation of msl-2 5′ splice site recognition by sex-lethal. *RNA* **7:** 1185–1191.

Gebauer, F., Merendino, L., Hentze, M.W., Valcarcel, 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., 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., Hentze, M.W. 2003. Drosophila sex-lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA. *Mol Cell* **11:** 1397-1404.

Geng, C., Macdonald, P.M. 2006. Imp associates with squid and Hrp48 and contributes to localized expression of gurken in the oocyte. *Mol Cell Biol* **26:** 9508-9516.

Goodrich, J.S., Clouse, K.N., Schüpbach, T. 2004. Hrb27C, Sqd and Otu cooperatively regulate gurken RNA localization and mediate nurse cell chromosome dispersion in Drosophila oogenesis. *Development* **131:** 1949-1958.

Graindorge, A., Militti C., Gebauer F. 2011. Post-transcriptional control of X-chromosome dosage compensation. *Wiley Interdiscip Rev RNA* **2:** 534-545

Graindorge, A., Carre, C., Gebauer, F. 2013. SXL promotes nuclear retention of msl2 mRNA via interactions with the STAR protein HOW. *Genes Dev* **27:** 1421-1433.

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.

Hammond, L.E., Rudner, D.Z., Kanaar, R., Rio, D.C. 1997. Mutations in the hrp48 gene, which encodes a Drosophila heterogeneous nuclear ribonucleoprotein particle protein, cause lethality and developmental defects and affect P-element third-intron splicing in vivo. *Mol Cell Biol* **17:** 7260-7267.

Hennig, J., Militti, C., Popowicz, G.M., Wang, I., Sonntag, M., Geerlof, A., Gabel. F., Gebauer, F., Sattler, M. 2014. Structural basis for the assembly of the Sxl-Unr translation regulatory complex. *Nature* **515:** 287-290.

Hinnebusch, A. G. 2014. The Scanning Mechanism of Eukaryotic Translation Initiation. *Annu Rev Biochem* **83:** 779-812.

Huynh, J.R., Munro, T.P. Smith-Litière L., Lepesant,J.A. 2004. The *Drosophila* hnRNPA/B Homolog, Hrp48, Is Specifically Required for a Distinct Step in *osk* mRNA Localization. *Dev Cell* **6:** 625–635.

Kelley, R.L., Wang, J., Bell, L., Kuroda, M.I. 1997. Sex lethal controls dosage compensation in Drosophila by a non-splicing mechanism. *Nature* **387:** 195-199.

Lasko, P. 2011. Posttranscriptional regulation in Drosophila oocytes and early embryos. *Wiley Interdiscip Rev RNA* **2:** 408-416.

Lee, A.S., Kranzusch, P.J., Cate, J.H. 2015. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature* **522:** 111-114.

Masutani, M., Sonenberg, N., Yokoyama, S., Imataka, H. 2007. Reconstitution reveals the functional core of mammalian eIF3. *EMBO J* **26:** 3373-3383.

Medenbach, J., Seiler, M., Hentze, M. W. 2010. Translational Control via Protein-Regulated Upstream Open Reading Frames. *Cell* **145:** 902–913.

Merendino, L., Guth, S., Bilbao, D., Martinez, C., and Valcarc el, J. 1999. Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature* **402:** 838-841.

Nelson, M.R., Luo, H., Vari, H.K., Cox, B.J., Simmonds, A.J., Krause, H.M., Lipshitz, H.D., Smibert, C.A. 2007. A multiprotein complex that mediates translational enhancement in Drosophila. *J Biol Chem* **282:** 34031-34038.

Pisarev, A.V., Kolupaeva, V.G., Yusupov, M.M., Hellen, C.U., Pestova, T.V. 2008. Ribosomal position and contacts of mRNA in

eukaryotic translation initiation complexes. *EMBO J* **27:** 1609-1621.

Smith, R.W., Anderson, R.C., Smithm J.W., Brook, M., Richardson, W.A., Gray, N.K. 2011. DAZAP1, an RNA-binding protein required for development and spermatogenesis, can regulate mRNA translation. *RNA* **17:** 1282-1295.

Smith, M.D., Gu, Y., Querol-Audí, J., Vogan, J.M., Nitido, A., Cate, J.H. 2013. Human-like eukaryotic translation initiation factor 3 from Neurospora crassa. *PLoS One* **8:** e78715.

Sonenberg, N., Hinnebusch, A.G. 2009. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* **136:** 731–745.

Suissa, Y., Kalifa, Y., Dinur, T., Graham, P., Deshpande, G., Schedl, P., Gerlitz, O. 2010. Hrp48 attenuates Sxl expression to allow for proper notch expression and signaling in wing development. *Proc Natl Acad Sci USA* **107:** 6930-6935.

Sun, C., Todorovic, A., Querol-Audí, J., Bai, Y., Villa, N., Snyder, M., Ashchyan, J., Lewis, C.S., Hartland, A., Gradia, S., Fraser, C.S., Doudna, J.A., Nogales, E., Cate, J.H. 2011. Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3). *Proc Natl Acad Sci USA* **108:** 20473-20478.

Szostak, E., Gebauer, F. 2012. Translational control by 3'UTR binding proteins. *Brief Funct Genomics* **12:** 58-65.

Valásek, L. S. 2012. 'Ribozoomin'-translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs). *Curr Protein Pept Sci* **13:** 305-330.

Villa, N., Do, A., Hershey, J.W., Fraser, C.S. 2013. Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome. *J Biol Chem* **288:** 32932-32940.

Yano, T., López, de Quinto, S., Matsui, Y., Shevchenko, A., Shevchenko, A., Ephrussi, A. 2004. Hrp48, a Drosophila hnRNPA/B homolog, binds and regulates translation of oskar mRNA. *Dev Cell* **6:** 637-648.

FIGURE LEGENDS

Figure 1. "Region 5" is important for translational repression of *msl2* mRNA independent of SXL and UNR binding

(A) Schematic representation of *msl2* mRNA. SXL binding sites are depicted with grey and black boxes (A-F); sites A, C, D (grey) are dispensable for translational repression while sites B, E and F (black) are required. Numbers indicate the length of the 5' and 3' UTRs (626 and 1047 nucleotides, respectively) and the position of the minimal sequences required for translational repression (nt 270-339 in the 5' UTR and nt 909-954 in the 3' UTR). A detail of the minimal functional 3' UTR is shown, with the UNR binding sites in blue and the region 5 in green. *(B)* Region 5 is important for 3'-mediated regulation. *In vitro* translation assays were performed with a series of indicator constructs that support either 5', 3' or 5'+3'-mediated regulation, schematically represented above each graph. Constructs WT, Δ3' and 5m contain a 5' UTR of 354 nt including sites A and B. Constructs (AB)m and (AB)m-5m contain a 5' UTR of 626 nt lacking sites A and B. Constructs

uORF-BL(EF)m and uORF-BL(EF5)m contain the minimal functional 5' UTR. All constructs contain minimal 3' UTR derivatives, as indicated. Mutated region 5 (CU_8) is highlighted in red. *In vitro* translation assays were performed with increasing amounts of recombinant His-dRBD4. Renilla luciferase mRNA was co-translated as an internal control. Firefly luciferase was corrected for Renilla expression, and the data were plotted as the percentage of translation in the absence of SXL. Error bars represent the standard deviation of at least 3 experiments. *(C)* Gel mobility shift assays using the wild type minimal *msl2* 3' UTR (WT), or a derivative lacking region 5 (5m). Increasing amounts of GST-dRBD4 or UNR were added to the reaction, as indicated. The positions of the protein-RNA complexes and the free probe are indicated.

Figure 2. Hrp48 binds to region 5

(A) Schematic representation of the GRAB purification protocol. Recombinant GST-dRBD4 was incubated with biotinylated *msl2* RNA probes and *Drosophila* embryo extract in translation reaction conditions. A first purification step includes GST pull-down and

elution with TEV protease, which separates the GST moiety. In the second purification step, the biotinylated RNA is pulled-down with streptavidin beads, and complexes are eluted with SDS buffer. *(B)* Schematic representation of the *msl2* RNA probes and the SXL derivatives used in this study. WT and 5m RNAs are as described in the legend of Figure 1C. EFm RNA lacks sites E and F, which have been substituted by CU repeats. *D. melanogaster* SXL is a 354 amino acid protein containing two RRM-type RNA binding domains and a glycine/ asparagine (GN)-rich amino-terminal region. The deletion derivative dRBD4 is fully competent for translational repression. mRBD contains the RNA-binding domains of the SXL homolog from *Musca domestica*, sharing 95% identity with *Drosophila* SXL but inactive in translational repression. hRBD1 is a hybrid of *Drosophila* and Musca SXL unable to repress translation, as critical contacts for SXL:UNR complex formation are supported by *Drosophila* SXL RRM1 (Hennig et al., 2014). *(C)* GRAB eluates obtained with WT and 5m RNAs were analyzed by PAGE and silver stained. Selected bands were cut and sent for identification by mass spectrometry. Asterisks denote bands

that were not reproducibly absent in the 5m eluate. *(D)* Hrp48 binds to region 5. RNA affinity chromatography followed by Western blot (WB) was used to confirm the specificity of Hrp48 binding. Mutation of region 5 reduces Hrp48 binding, while binding of SXL and UNR remains unaffected. *(E)* Hrp48 binds to *msl2* independently of SXL and UNR. RNA affinity chromatography was performed with WT and EFm RNAs, using the SXL derivatives described in part B. WB, Western blot.

Figure 3. Hrp48 contributes to *msl2* mRNA translational repression

(A) Depletion of Hrp48 impairs SXL-mediated repression. Hrp48 was depleted from male SL2 cells, which were then transfected with a βGal reporter containing the full length 5' and 3' UTRs of *msl2* (FC, *Top* panel), a control Renilla luciferase plasmid, and increasing amounts of a SXL-encoding plasmid. GFP RNAi was carried as negative control. βGal activity was normalized for Renilla expression and corrected for the levels of the reporter RNA. The data were plotted relative to the βGal activity in the absence of SXL. Error bars represent the standard deviation from five

independent experiments (Student's t-Test Pval *<0.05, **<0.01). The efficiency of depletion was assessed by Western blot (*Bottom* panel). *(B)* Hrp48 functions through region 5. Hrp48 was depleted from embryo extracts after 15 rounds of incubation with an RNA oligomer containing two Hrp48 binding sites. Depletion with poly(C) was carried as control (Ctrl). The repression of *msl2* reporters containing or lacking region 5 (*Top* panel) upon addition of 10 ng GST-dRBD4 was tested. Renilla luciferase was co-translated as internal control. Firefly luciferase activity was corrected for Renilla expression and plotted relative to the activity in the absence of SXL. Error bars represent the standard deviation from at least 11 replicates in 4 independent experiments. The efficiency of depletion was assessed by Western blot (*Bottom* panel).

Figure 4. Hrp48 interacts with components of the *msl2* repressor complex

(A) Hrp48 interacts with endogenous *msl2* mRNA. Hrp48 was immunoprecipitated from *Drosophila* embryo extracts, and the presence of *msl2* mRNA in the pellet was tested by semi-

quantitative PCR. Immunoprecipitation with non-specific IgG was carried as negative control. 18S RNA is shown as a measure of background. *(B)* Hrp48 interacts with *msl2* repressors. SXL, UNR and HOW were immunoprecipitated from *Drosophila* embryo extracts, and the presence of Hrp48 in the pellet was tested by Western blot. A parallel sample was treated with RNase. Non-specific IgG was carried as negative control.

Figure 5. Hrp48 targets eIF3d, an initiation factor specifically required for *msl2* mRNA translation

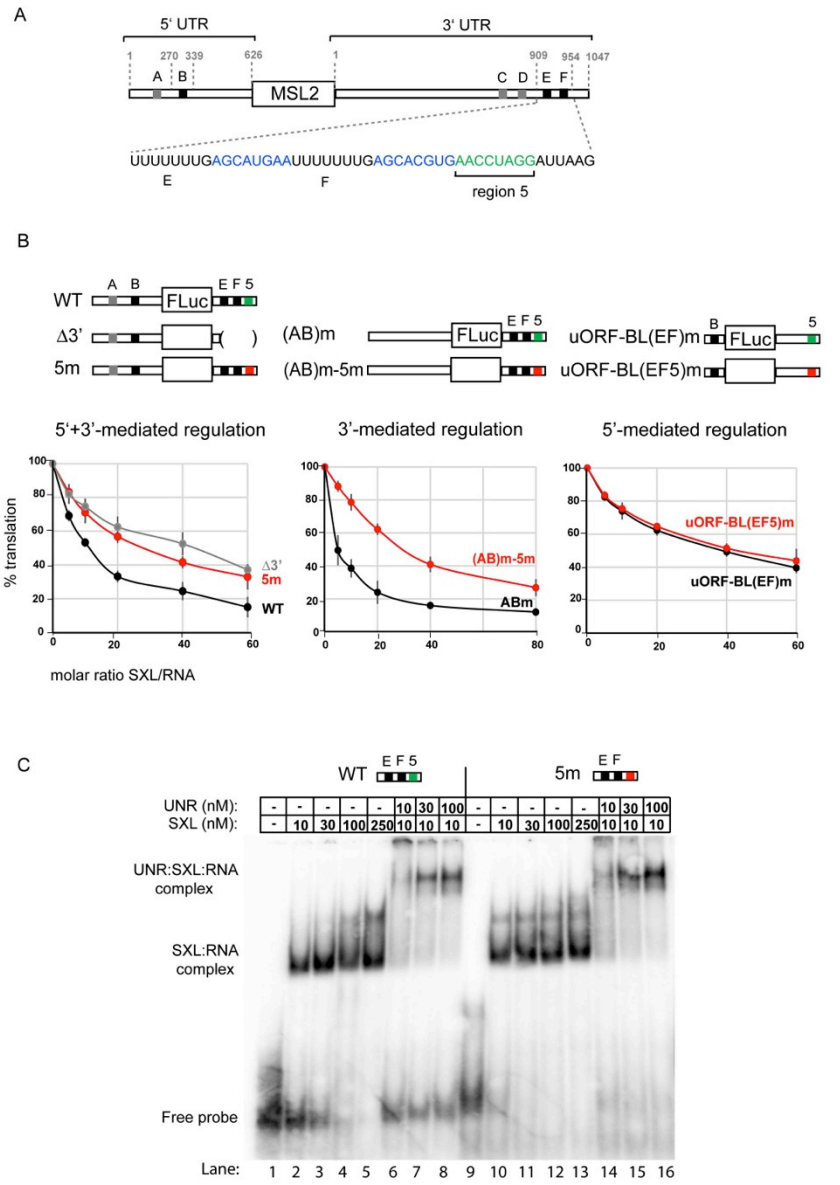
(A) Top, Volcano plot showing the mass spectrometry analysis of triplicate pull-downs of *Drosophila* embryo extracts with an oligomer containing Hrp48 binding sites. An unrelated oligomer (polyC) was used as control. Relevant proteins are marked with color. The red line indicates the significance threshold (Pval=0.05). *Bottom*, Western blot of Hrp48 in the eluates to test the efficiency and specificity of pull-down. H, Hrp48 oligo; C, control oligo. *(B)* Hrp48 co-immunoprecipitates with eIF3d. Recombinant HA-tagged eIF3d was incubated with *Drosophila* embryo extract, captured with αHA beads, and the presence of Hrp48 in the pellet was tested by

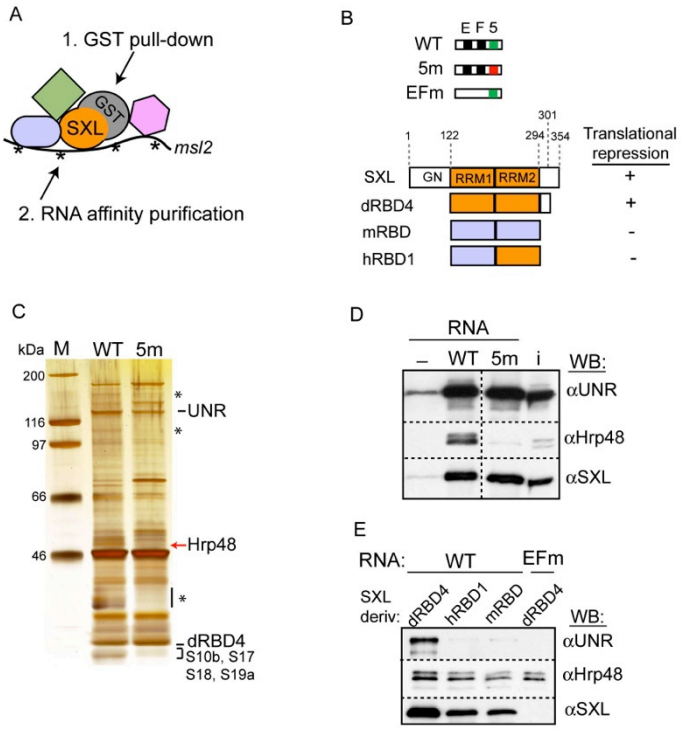
Western blot. *(C)* eIF3d is necessary for *msl2* mRNA translation. eIF3d was depleted from SL2 cells and the efficiency of translation of the FC *msl2* reporter was measured. Depletion of two additional non-core eIF3 subunits (eIF3e, eIF3h) and RNAi against GFP were carried as controls. The depletion efficiency was measured by RT-qPCR, and plotted relative to the amount of the corresponding eIF subunit in GFP RNAi cells (*Left* panel). To obtain the efficiency of FC translation, βGal activity was normalized for co-transfected Renilla expression and corrected for the levels of the reporter RNA. The data were plotted relative to the βGal activity in GFP RNAi cells (*Right* panel). Error bars represent the standard deviation from five experiments. *(D)* Depletion of eIF3d causes a mild defect in global translation. *De novo* protein synthesis was assessed by metabolic labeling with ^{35}S-methionine. A Coomassie stained gel is shown as loading reference. Numbers represent quantification of the ^{35}S signal corrected for loading. *(E)* eIF3d is required for efficient SXL-mediated repression. The ability of SXL to repress translation of the FC *msl2* reporter was measured in eIF3d depleted cells. A Renilla luciferase encoding plasmid was co-transfected as an

internal control. RNAi against eIF3e, eIF3h and GFP were carried as controls. The data were processed as described for Figure 3A.

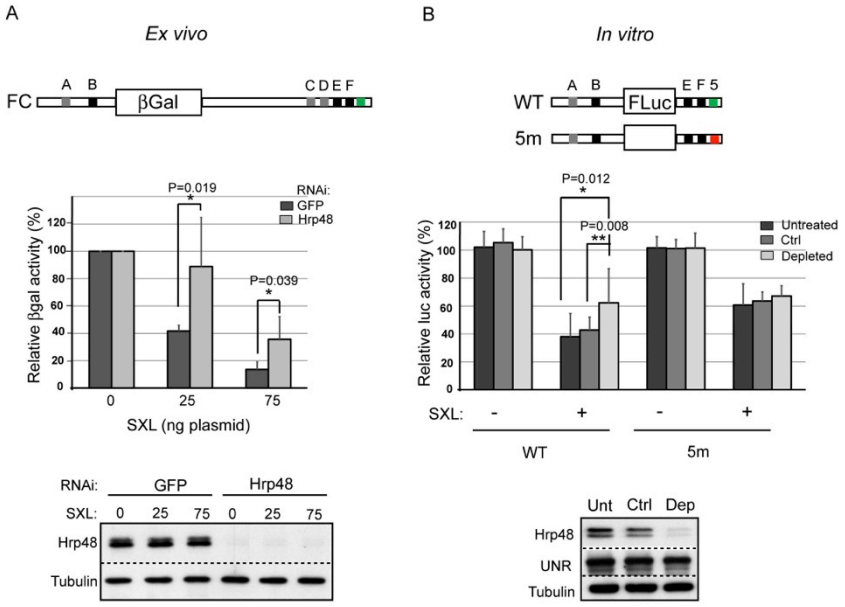
FIGURES

Szostak_Fig1

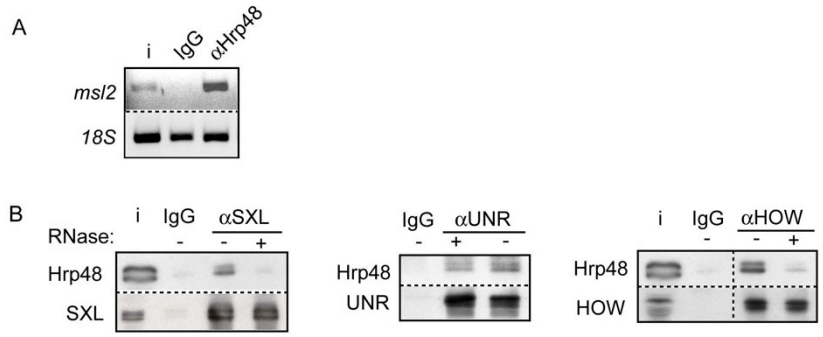


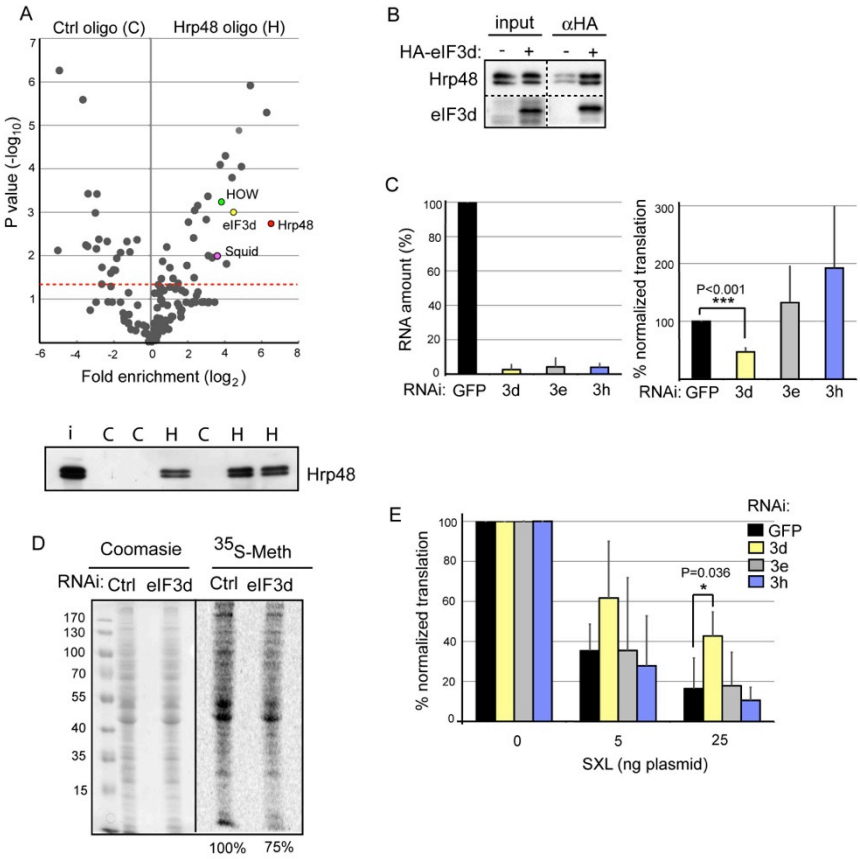

A
1. GST pull-down
GST
SXL
msl2
2. RNA affinity purification
B
E F 5
WT
5m
EFm
1
122
294
301
354
Translational repression
SXL
GN
RRM1
RRM2
+
dRBD4
+
mRBD
-
hRBD1
-
C
kDa
M
WT
5m
200
116
97
66
46
UNR
Hrp48
dRBD4
S10b, S17
S18, S19a
D
RNA
–
WT
5m
i
WB:
αUNR
αHrp48
αSXL
E
RNA:
WT
EFm
SXL deriv:
dRBD4
hRBD1
mRBD
dRBD4
WB:
αUNR
αHrp48
αSXL

Szostak_Fig 3



Szostak_Fig 4




A
Ctrl oligo (C)
Hrp48 oligo (H)
P value (-log10)
Fold enrichment (log2)
HOW
eIF3d
Hrp48
Squid
i C C H C H H
Hrp48
B
input
αHA
HA-eIF3d: - + - +
Hrp48
eIF3d
C
RNA amount (%)
RNAi: GFP 3d 3e 3h
% normalized translation
P<0.001
***
RNAi: GFP 3d 3e 3h
D
Coomasie
35S-Meth
RNAi: Ctrl eIF3d Ctrl eIF3d
170
130
100
70
55
40
35
15
100% 75%
E
% normalized translation
RNAi:
GFP
3d
3e
3h
P=0.036
*
0
5
25
SXL (ng plasmid)