DROSOPHILA UNR REGULATES DOSAGE COMPENSATION THROUGH MODULATION OF RNA-PROTEIN INTERACTIONS

Cristina Militti

Doctoral Thesis UPF / 2013

Thesis Director

Dra. Fátima Gebauer Hernández

GENE REGULATION, STEM CELLS AND CANCER DEPARMENT
CENTER FOR GENOMIC REGULATION (CRG)

To my family

Contents

Acknowledgments

It is good to have an end to journey toward; but it is the journey that matters, in the end…and now that I'm at the end, I like to think that it has been worth to go through this troubled extraordinary trip that my PhD has been. Doing research is among the most exciting adventures and for this I thank Fatima that has given me the opportunity to experience it. And I want to thank all the people that I've met and have been closed to me during these long years. The lab, that has supported me with help and advices, forgiving and accepting all my oddities. All the people inside and outside of the institute with whom I had the chance to share an idea or a smile. Francesca and Camilla, because they've always reminded me that the best way to understand life is enjoying it. Berta, that has guided me through the way to myself. My family, for their unconditional love and blind trust. And Mirko, that from afar, from the beginning to the end, has always helped me to get up every time I fell.

Abstract

In *Drosophila,* the imbalance in X-linked gene content between females (XX) and males (XY) is restored through the 2-fold hypertranscription of the single male X-chromosome. This process, which is called dosage compensation, is mediated by the action of the dosage compensation complex (DCC), a ribonucleoprotein assembly composed of at least five proteins (MSL1, MSL2, MSL3, MLE and MOF) and two long non-coding RNAs (*roX1* and *roX2*). Two features are essential for correct dosage compensation: the specific recognition of the X-chromosome by the DCC and the confinement of the DCC function to the male organism.

The RNA binding protein Upstream of N-ras (UNR) is involved in the regulation of these two processes and we have dissected the molecular mechanisms by which this regulation occurs. We have found that, in male flies, UNR promotes dosage compensation by facilitating the association of *roX2* with MLE, which is required for correct DCC formation and X-chromosome targeting. In female flies, UNR represses dosage compensation in part by enhancing the binding of SXL to the 3'UTR of *msl2* mRNA, thus ensuring tight *msl2* translational repression and subsequent inhibition of DCC formation.

Resúmen

En *Drosophila*, el desequilibrio en cuanto al contenido de genes ligados al cromosoma X entre hembras (XX) y machos (XY) es corregido mediante la duplicación de la transcripción del único cromosoma X del macho. Este proceso, llamado compensación de dosis, es mediado por un ensamblaje molecular compuesto por al menos cinco proteínas (MSL1, MSL2, MSL3, MLE y MOF) y dos RNAs largos no codificantes (*roX1* y *roX2*), llamado complejo de compensación de dosis (DCC). La compensación de dosis requiere dos condiciones fundamentales: el reconocimiento específico del cromosoma X por el DCC, y la restricción del proceso a moscas macho.

La proteína de unión a RNA Upstream-of-N-Ras (UNR) está implicada en la consecución de ambas condiciones, y aquí hemos estudiado los mecanismos moleculares por los que UNR actúa. Hemos encontrado que, en machos, UNR promueve la compensación de dosis facilitando la asociación de *roX2* a MLE, necesaria para una correcta formación del DCC y para su unión al cromosoma X. En

hembras, UNR inhibe la compensación de dosis, al menos en parte, promoviendo la unión de SXL al extremo 3’ UTR del mRNA que codifica para *msl2*, lo que resulta en represión de la traducción de *msl2* e inhibición de la formación del DCC.

Preface

RNAs are essential components of the molecular machinery of the cell, playing fundamental roles not only as messengers of the genetic information, but also as regulatory molecules. Their function and regulation relies on the action of RNA binding proteins, which decide for the fate of the target RNAs by determining when, where and how their activity will be exerted. Investigating the molecular mechanisms by which RNA binding proteins work is of great interest to understand RNA function.

Here we study the way the RNA binding protein UNR acts on its target RNAs to regulate dosage compensation, a process essential for survival.

INTRODUCTION

1. Long non-coding RNAs

The formulation of the central dogma by Crick in 1958, together with the description of the lac operon in 1961 by Jacob and Monod, led to a model for gene function according to which RNA represents an intermediary molecule in the genetic flow that goes from the informational storage of DNA to the executive activity of proteins. For decades the concept that "DNA makes RNA makes protein" has ruled in the field of molecular biology, strongly contributing to the idea that most of the RNAs (with the exclusion of the well known RNAs involved in splicing and translation) function as "messenger" molecules.

Genome wide transcriptome studies performed with advanced and annotation unbiased techniques have recently challenged this idea. In the late 1990s- early 2000s pioneering studies conducted in yeast, mouse and human first revealed the high complexity of eukaryotic transcriptomes (Velculescu et al., 1997; Okazaki et al., 2002; Kapranov et al., 2002). Follow-up studies have found that 63% and 93% of the mouse and human genomes, respectively, can give rise to detectable transcripts (Carninci et al., 2005; Birney et al., 2007). However, only a small proportion encode for proteins, while most of them seem to lack any coding potential falling into the category of non-coding RNAs.

The discovery that eukaryotic genomes are pervasively transcribed has reshaped the way scientists look at genome organization, giving a chance for function to what was previously considered as "junk" DNA (Kapranov et al., 2007). Nonetheless, certain skepticism has arisen towards the potential functionality of pervasive transcription, suggesting the possibility that many of the newly identified transcripts could just represent products of transcriptional noise (Struhl, 2007). Despite these odds, a growing number of studies have reported functional roles for a subset of newly identified non-coding RNAs, demonstrating that at least part of them truly represent a new class of cellular regulators.

1.1. Long non-coding RNAs: features and conservation

Apart from the well-known housekeeping structural RNAs (tRNAs, rRNAs, snRNAs, snoRNAs) and the most recently studied small regulatory RNAs (siRNAs, miRNAs, piRNAs), a big proportion of the non-coding genome is represented by long non-coding RNAs (long ncRNAs/ lncRNAs).

LncRNAs are RNA molecules which lack any coding potential and whose size can greatly vary between 200 nucleotides and several kilobases. They can originate from a diversity of sources like intergenic regions, pseudogenes or coding gene- associated promoters, transcription start sites, introns and terminators. From yeast to humans, lncRNAs have been identified in a variety of eukaryotic organisms.

One of the most prominent features of lncRNAs is the poor conservation of their primary sequence. Initial observations have reported a degree of conservation for lncRNAs as low as for non-constrained intergenic sequences. Following studies, however, have revealed that the rates of nucleotide substitution, insertion and deletion in the transcribed sequences, promoters and splice sites of lncRNAs are actually lower than those associated with intergenic sequences (Ponjavic et al., 2007). Moreover, analysis of sequence conservation in 50 nucleotide (nt) windows have identified a significant presence of patches of conserved sequences that could likely function as binding sites for protein or RNA targets (Pang et al., 2006). Secondary structures experimentally tested for functionality have also been found conserved in lncRNAs such as *Xist* (Zhao et al., 2008) and *roXs* (Park et al., 2007; Park et al., 2008).

These observations suggest that lncRNAs could act by means of functionally conserved domains (sequences or structures) embedded in dispensable regions, whose sequence would not be subjected to evolutionary pressure. In such a scenario, the great sequence flexibility would allow the lncRNA to rapidly evolve and explore new strategies for regulation.

1.2. Molecular mechanisms of lncRNA function

A growing number of studies have reported the involvement of lncRNAs in a variety of cellular and developmental processes. In some cases functional links between lncRNAs and anomalous conditions such as disease and cancer have been uncovered (Taft et al., 2009; Huarte and Rinn, 2010). In fact the high flexibility associated with the RNA molecule, which can potentially sustain single or multiple concurrent interactions

with DNA, RNAs and proteins, finds its paradigm in lncRNAs, which indeed can perform their function through an extremely disparate set of molecular mechanisms (Figure 1).

1.2.1. Regulation of adjacent gene expression by transcription *per se*

It has been reported that the act of transcribing lncRNAs can suffice for the regulation of the expression of nearby genes. In yeast, the gene *SER3*, which encodes for an enzyme involved in serine biosynthesis, is regulated by the transcription of the *SRG1* lncRNA. In conditions of high intracellular serine levels, the Cha4-serine dependent activator binds to the *SRG1* promoter, inducing its expression. Transcription of the lncRNA then leads to increased nucleosome occupancy of the downstream *SER3* promoter, causing transcriptional interference (Martens et al., 2004, Hainer et al., 2011).

Recently it has been shown that also the *Airn* lncRNA regulates genomic imprinting by transcriptional interference, although early initial results had suggested a different mechanism (Santoro et al., 2013). *Airn* is involved in the parental-specific silencing of a 400 kb region containing the *Slc22a3*, *Slc22a2* and *Igf2r* genes. Its transcription on the paternal allele in the antisense orientation to the *Igf2r* gene prevents *Igf2r* expression and blocks the binding of transcriptional activators, thus inhibiting transcription at the imprinted locus. In a second step, the repressed genes in the cluster attract chromatin modifying factors, such as the histone methyltransferase 2 and the PRC complexes, which would tighten gene silencing by inducing the formation of a repressive nuclear domain.

LncRNA transcription can also repress the expression of closely located genes by directly inducing the establishment of a repressive chromatin state. In yeast, indeed, it has been shown that the transcription in antisense direction of the *GAL10-ncRNA* from a promoter located at the 3' end of the *GAL10* ORF contributes to the repression of *GAL1-10* expression at low glucose concentration. After initiation at the *GAL10-ncRNA* promoter, RNA Pol II elongation would induce di- and tri- methylation of H3K36 and H3 deacetylation, leading to the induction of a repressive chromatin state that would secure the transcriptional repression of the *GAL1-10* locus (Houseley et al., 2008). In other cases lncRNA transcription has an activatory role on the expression of closely located genes. In *S. Pombe*, transcription of a series of lncRNAs from multiple sites upstream of the *fbp1*+ promoter induces the progressive opening of the chromatin

toward the *fbp1*+ transcription start site, making it accessible to transcriptional activators and RNA Pol II (Hirota et al., 2008).

1.2.2. *Cis*-tethering of protein complexes

One of the first identified functions of lncRNAs is the capacity to tether epigenetic complexes to target loci *in cis*. The *Kcnq1ot1* lncRNA is among the best studied examples. Like *Airn*, *Kcnq1ot1* is involved in the parental-specific silencing of the associated imprinted gene cluster. Similarly to *Airn*-mediated silencing, formation of a repressive chromatin domain on the imprinted genes seems to occur through the action of chromatin modifying factors. However, unlike *Airn*, the histone methyltransferase 2 and PRC complexes appear to be directly tethered on the imprinted locus by the *Kcnq1ot1* lncRNA, which is able to coat *in cis* the target genes. Some evidences suggest that *Kcnq1ot1* could also mediate the recruitment of DNA methyltransferases that would participate in the maintenance of the silenced state (Kanduri, 2011).

The lncRNA *ANRIL* has also been reported to *cis*-tether chromatin modifying complexes to the *INK4b-ARF-INK4a* locus, which contains genes involved in the regulation of cell proliferation and senescence. Nascent *ANRIL* ncRNAs transcribed in antisense direction to the *INK4b* gene are able to recruit PRC1 and PCR2 complexes and target them to the *INK4b* gene, allowing for silencing of the *INK4b-ARF-INK4a* locus (Aguilo et al., 2011).

A second class of lncRNAs participate in gene activation by recruiting activating complexes to target genes. One example is *Hottip,* a lncRNA transcribed in the opposite direction to the *Hox-a* cluster. *Hottip* co-transcriptionally interacts with the activating complex MLL and targets it to the *Hox-a* genes. Chromosomal looping inside the cluster allows *Hottip* to promote MLL targeting to distal genes in the locus. A similar role has been observed for the lncRNA *Mira* (*Mistral*) in the activation of *Hoxa6* and *Hoxa7* genes.

Another class of lncRNAs involved in the *cis*-activation of target genes are the enhancer RNAs (eRNAs), which are transcribed from enhancer regions and appear to be necessary for the activation potential of the enhancer itself (Wang et al., 2011). Lai and colleagues have recently found that the two eRNAs *ncRNA-a3* and *ncRNA-a7*, previously discovered to act as activators, interact with the Mediator complex (Lai et al., 2013). The Mediator complex is known to activate gene transcription by mediating DNA looping between enhancers and promoters, thus physically linking enhancer-bound

transcription factors with the promoter-bound transcription initiation machinery. By interacting with Mediator, the two activating eRNAs appear to participate in the enhancer-promoter connection, thereby stimulating transcription.

1.2.3. *Trans*-targeting of protein complexes by molecular scaffolding

Evidence for a function of lncRNAs as molecular scaffolds for protein complex targeting date back to 2002, when the lncRNAs *roX1* and *roX2* were found to be necessary for the correct targeting of the DCC to the X-chromosome in *Drosophila* (Meller et al., 2002). However, the ability of lncRNAs to "guide" the binding of protein complexes to distant target sites on the genome has become apparent only in the late 2000s. Indeed in 2007, Rinn and colleagues showed that the lncRNA *HOTAIR*, produced from the *HOX-C* locus in humans, interacts with the PRC2 complex and guides it to the *HOX-D* locus (Rinn et al. 2007). Intriguingly, also a second repressor complex (REST/coREST) has been reported to be targeted by *HOTAIR* to the same locus (Tsai et al 2010). The observation that the interaction with the two chromatin complexes is mediated by different domains of *HOTAIR* suggests that the lncRNA could function as a molecular scaffold, bridging together the two repressive complexes and guiding them for the coordinated and efficient repression of the target locus. However it is still unclear how the *HOTAIR* RNA enables the specific binding of the chromatin complexes to the target genes.

1.2.4. *Trans*-targeting of protein complexes by DNA sequence recognition

LncRNAs can also mediate targeting of protein complexes by the direct recognition of DNA sequences in target genes. Such a mechanism has been suggested for the action of pRNAs (promoter RNAs) in the regulation of rRNA gene silencing. pRNAs are ncRNAs produced from a RNA Pol I promoter located upstream of the rRNA transcription start site and whose sequence matches the rRNA gene promoter. It has been shown that pRNAs can interact with the silencing complex NoRC and recruit it to the rRNA promoter, most probably through DNA recognition and formation of a triple stranded structure. The triple-helix is also thought to mediate the recruitment of a DNA methyltransferase, which would secure the silencing of the rRNA gene through methylation of the CpG islands at the promoter (Schmitz et al., 2010).

1.2.5. Modulation of protein activity

LncRNAs can also function by modulating the activity of the proteins to which they bind. In human cell lines, DNA damage induces the production of ncRNAs from a region upstream of the cyclin D1 (*CCDN1*) promoter. *CCDN1* ncRNAs are able to bind the TLS protein and induce an allosteric change responsible for its activation. Active TLS is then able to interact and inhibit the acetyltransferase CBP present at the *CCND1* promoter, thus inducing *CCND1* silencing (Wang et al., 2008). Similar mechanisms have been suggested for the activation of the steroid receptors by the ncRNA *SRA* (Lanz et al., 1999) and of Dlx2 protein by the ncRNA *Evf-2* (Feng et al., 2006). A function as modulator of protein activity has also been suggested for the mouse B2 and the human Alu RNAs in the transcriptional regulation of Pol II transcription. Indeed, during heat shock the two RNAs are transcriptionally activated through a still unknown mechanism, and then interact with the RNA Pol II present at the promoter of housekeeping genes, thereby inhibiting their transcription (Walters et al., 2009).

LncRNAs can also influence the activity of the target protein by functioning as decoys. One example is the *C.elegans rncs-1* ncRNA. *Rncs-1* is a double stranded ncRNA expressed during starvation, which has been shown to inhibit siRNA processing through the interaction with Dicer. The ncRNA indeed is able to bind to Dicer without being processed, thanks to the presence of terminal branches in the dsRNA structure that protect it from cleavage. In this way it competes with endogenous dsRNA for Dicer binding, thus affecting siRNA production and ultimately mRNA stability (Hellwig et al., 2008).

1.2.6. LncRNAs as structural elements

LncRNAs have also been reported to play a role as structural elements in the formation of subcellular structures. One of these is the *MEN* ε/β (*NEAT 1*) ncRNA, which is thought to have a role in the nucleation and maintenance of paraspeckles, subnuclear structures whose function in RNA processing still needs to be elucidated (Bond et al., 2009). Another non-coding RNA with structural function is *xlsirts*, which is involved in the organization of the cytoskeleton in *Xenopus* oocytes.

1.2.7. Modulation of RNA processing or activity by base-pairing

LncRNAs can affect the processing or function of other RNA molecules by base-pairing. Pseudogenes and natural antisense transcripts (NATs) have been reported to base-pair to target RNA molecules and produce dsRNAs that are then processed to yield endo-siRNAs. The produced siRNAs then target additional molecules contributing to further decrease the levels of target RNAs (Tam et al., 2008; Watanabe et al., 2008).
NATs can also influence RNA processing. The *ZEB2* gene encodes a transcriptional repressor whose translation is prevented by the presence of a sequence at the 5'UTR that inhibits ribosomal scanning. Upon completion of the epithelial to mesenchimal transition, a NAT corresponding to the 5'UTR of the *ZEB2* mRNA is produced. The NAT base-pairs with *ZEB2* and hinders the binding of the spliceosome to the 5'splice site of an intron present at the 5'UTR. This leads to the retention of the intron that contains an IRES from which translation of the *ZEB2* mRNA is initiated, bypassing the scanning inhibitory sequence at the 5'UTR (Beltran et al., 2008).
The involvement of lncRNAs in mRNA decay was reported in the process of SMD (Staufen-mediated decay). LncRNAs containing Alu or SINE elements are able to specifically recognize target mRNAs by base-pairing with similar Alu or SINE elements present in their 3'UTRs. This interaction leads to the formation of a binding site for the Staufen protein, which in turn induces mRNA degradation (Gong et al., 2011; Wang et al., 2013).
LncRNAs may act as miRNA sponges by titrating miRNA molecules through base-pair recognition. By controlling the concentration of available miRNA molecules, miRNA sponges temporally regulate miRNA activity, and preserve off-target mRNAs from being recognized by the miRNA. One of the first non-coding RNAs reported to play such a role is the Arabidopsis *IPS1*. *IPS1* is a 550 nt non-coding RNA which contains a short conserved motif highly complementary to miR-399. miR-399 is up-regulated during inorganic phosphate (Pi) deprivation and down-regulates the expression of PHO2, an enzyme involved in the response to Pi starvation. Once the response has been triggered, PHO2 is stabilized to physiological levels by the expression of *IPS1*, which recognizes and sequesters miR-399, modulating miRNA silencing activity on the *PHO2* mRNA (Franco-Zorrilla et al., 2007). A similar mechanism is employed by the pseudogene *PTENP1* in the regulation of *PTEN* mRNA levels (Poliseno et al., 2010) and by the lncRNA *MD1* in the control of miR-133 activity during muscle differentiation (Cesana et al., 2011).

A recently identified new class of lncRNAs, named circular RNAs (circRNA), has been reported to function as miRNA sponges (Memczack et al., 2013; Hansen et al., 2013). *CDR1as/ciRS-7* is a conserved circular RNA containing about 70 binding sites for miR-7, which appears to be sequestered by the circular RNA in complex with Ago proteins. The observed high stability of the circRNA, together with the high number of miRNA binding sites, contribute to the potent capacity of *CDR1as/ciRS-7* circRNA to down-regulate miR7 function.

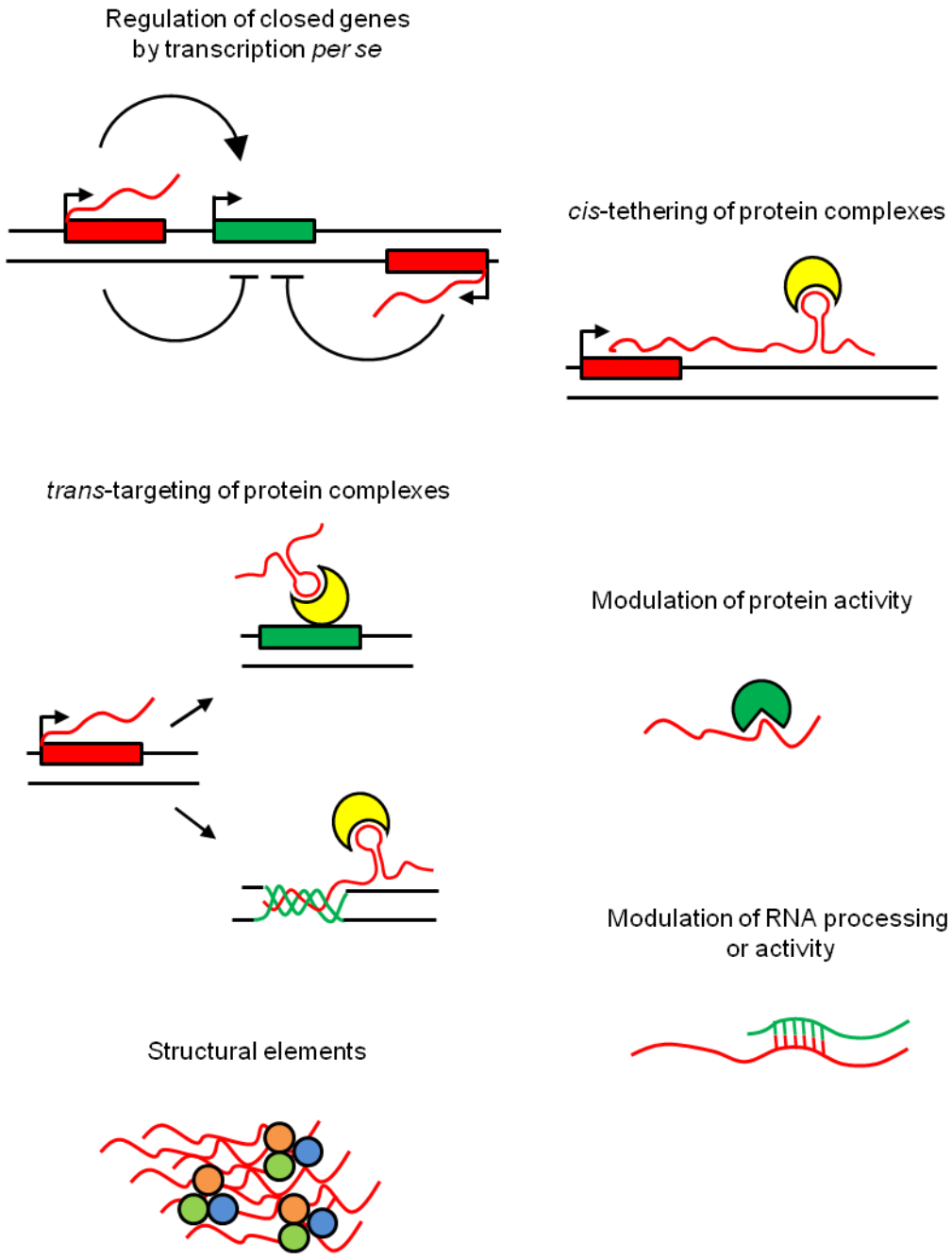


Figure 1. Molecular mechanims of long non-coding RNA function.

LncRNAs (red) can modulate the activity of target genes, proteins or RNA (green) through a variety of mechanisms. Transcription *per se* of lncRNAs can activate or inhibit transcription of closed target genes. LncRNAs can *cis*-tether protein complexes to closed loci or target them to distant loci by acting as molecular scaffolds or through DNA sequence recognition. They can modulate the activity of target proteins or RNAs through binding. LncRNAs can also function as structural elements.

2. Dosage compensation: an overview

2.1. Sex chromosome evolution and the need of dosage compensation

In species where sex is genetically controlled, sex determination mechanisms often rely on heteromorphic sex chromosomes, whose different combination in the two sexes determines the female or male identity. Although sex chromosome systems can be very diverse and complex, the male heterogametic XX/XY (XX/X0) and the female heterogametic ZW/ZZ systems have been so far the best studied, providing insights onto how heteromorphic sex chromosomes have evolved.

The prevailing model for sex chromosome evolution posits that sex chromosomes originate from an ancestral pair of autosomes that, after acquiring a sex-determining locus, have ceased to recombine. Over time, and in the absence of recombination, the two sex chromosomes evolve independently, becoming very different in size and morphology. While recombination with the homologue pair is still possible for the homomorphic chromosome (X or Z) in the homogametic sex, recombination of the heteromorphic sex chromosome (Y or W) is almost completely abolished, leading to the accumulation of deleterious mutations and consequent pseudogenization and gene loss. The degeneration of the heteromorphic chromosome leads to aneuploidy in the heterogametic sex because of the presence of one single active copy of the sex chromosome, a condition that is not tolerated by the organism. For this reason, several times and independently during evolution, dosage compensation mechanisms have evolved to balance the expression of sex chromosome genes with the rest of the genome, and to equalize the levels of sex chromosome transcripts between the two sexes.

2.2. Models of dosage compensation

So far dosage compensation has been extensively studied at the molecular level in three distinct systems: worms (*Caenorhabdtis elegans*), flies (*Drosophila melanogaster*) and mammals (*Mus musculus*). Although these organisms use different strategies to equalize sex chromosome gene expression between females and males, similar molecular solutions have been employed (Figure 2). In the three systems,

global modifications of chromatin states are generated to modify the transcriptional output of most genes on the targeted sex chromosome. The sex chromosome needs to be distinguished from the autosomes, and this is achieved by DNA sequence recognition and lncRNA action. Finally, in all systems, activation of the dosage compensation process is restricted to one of the two sexes by means of sex-specific regulation of the dosage compensation triggering factors.

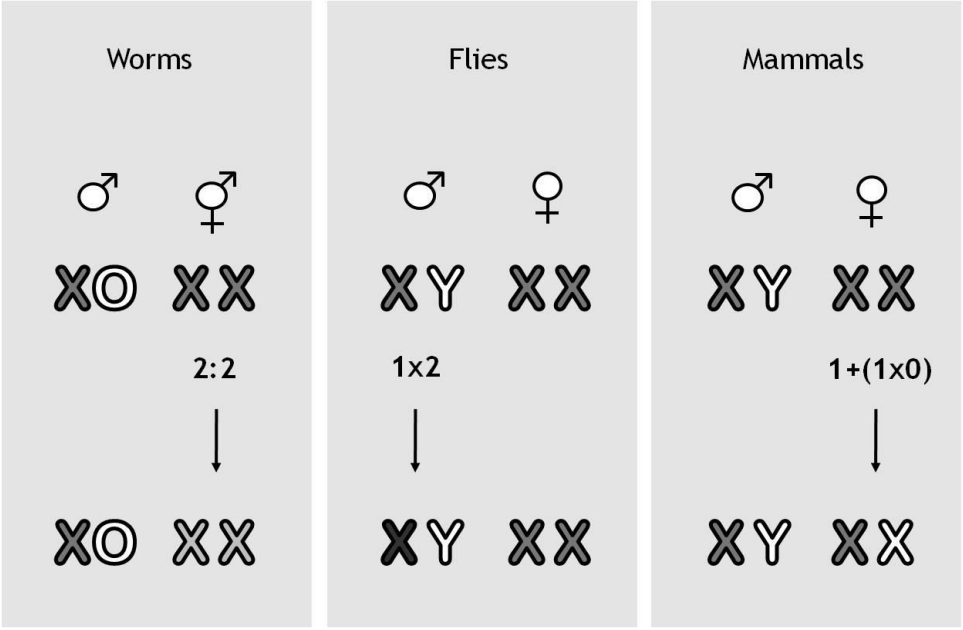


Figure 2. Strategies of dosage compensation.
Equalization of the X-linked transcripts between females and males is achieved through the two-fold down-regulation of both X-chromosomes in hermaphrodite worms, the 2-fold upregulation of the single X-chromosome in male flies and the inactivation of one of the X-chromosome in the female mammals.

2.2.1. Dosage compensation in *Caenorhabditis elegans*

In *C.elegans,* equalization of X-linked gene expression between hermaphrodites (XX) and males (X0) occurs through a two-fold reduction of the transcription of the hermaphrodite X chromosomes (Ercan and Lieb, 2009; Csankovszki, 2009a). Such regulation is exerted by the dosage compensation complex (DCC), which is constituted of at least 8 proteins: SDC-1, SDC-2, SDC-3 (*sex determination and dosage compensation defects*), DPY-21, DPY-26, DPY-27, DPY-28 (*dumpy*) and MIX-1 (*mitosis and X-associated*).

The DCC recognizes and binds to the two X-chromosomes, inducing their down-regulation. However, the molecular mechanisms by which binding and repression are achieved are still unclear. The homology shared by three of the DCC subunits (DPY-26, DPY-28, MIX-1) with subunits from the condensin complex suggests a possible function in chromosome condensation and formation of a higher order chromosomal structure that would result in transcriptional inhibition (Csankovszki, 2009b).

DCC assembly and first X-chromosome recognition seem to rely on the activity of SDC-2, which is expressed only in hermaphrodites and binds to the Xs in the absence of the other complex subunits. Although their function is still not clear, SDC-1 and SDC-3 are thought to participate in X-recognition and DNA binding. Targeting to the X-chromosome seems to occur in a two-step manner (Figure 3). Initial binding of the DCC to the so called *rex* sites (recruitment element on X) is partially dependent on a 10-12 bp motif, which is enriched on the Xs, although it can be found in autosomes and is not a prerequisite of all the *rex* sites. Once bound to the *rex* site, the DCC spreads to the promoters of closely active genes through a still unknown mechanism that could involve H3K4 recognition and the function of DPY-30. In males inhibition of dosage compensation is achieved through the repression of SDC-2 by the action of the male specific factor XOL-1 (Csankovszki, 2009b).

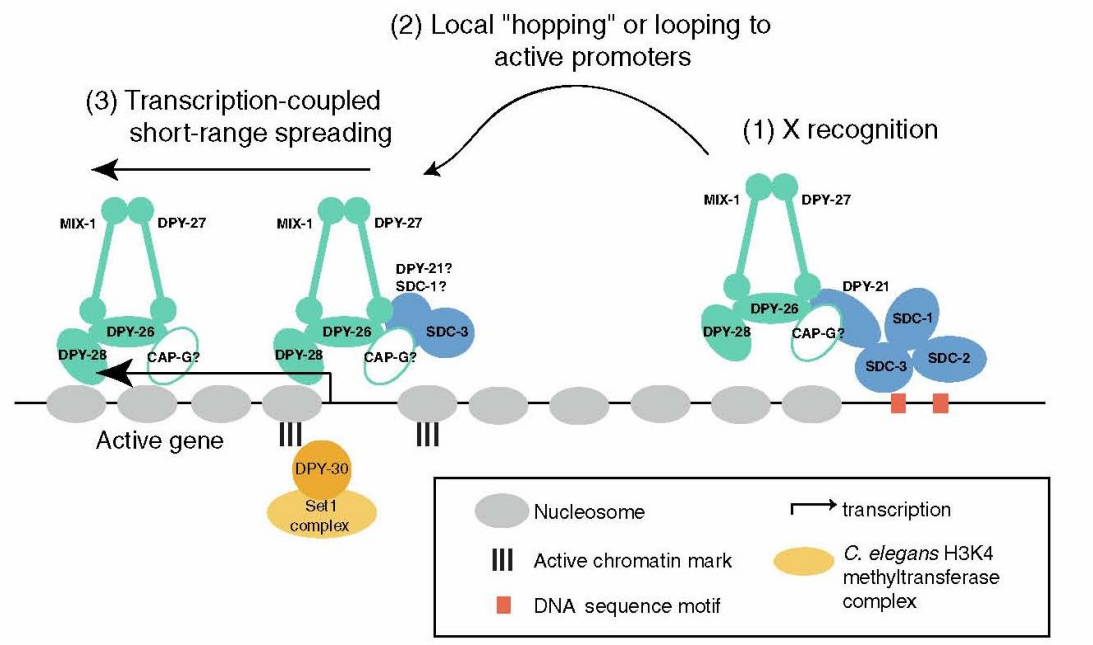


Figure 3. Model for targeting of the DCC to the X-chromosome in *C.elegans*.
First X-chromosome recognition is mediated by the binding of the SDC subunits of the DCC to DNA motifs enriched on the X. Spreading of the complex to the active genes occurs through a still unknown mechanism that could involve local "hopping" or chromosomal looping. Recognition of active chromatin marks on promoters could be mediated by DPY-30 and the Set1-complex (taken from Ercan and Lieb, 2009).

2.2.2. X-chromosome inactivation in mammals

In mammals, dosage compensation occurs through the inactivation of one of the two X-chromosomes in the female organism (Payer and Lee, 2008; Augui et al., 2011; Lee and Bartolomei, 2013). In marsupials and in extra-embryonic tissues of some placental mammals, X-inactivation is imprinted, occurring exclusively on the paternal chromosome, while embryos of placental mammals show random X-inactivation.
Shut-down of X-chromosome expression is mediated by the action of the lncRNA *Xist*, which is transcribed from the X-chromosome to be inactivated and coats it *in cis*. Spreading of *Xist* is followed by an initial step of silencing, characterized by the deposition of chromatin repressive marks, partially through the action of recruited PRC1 and PRC2 complexes. Later on, incorporation of macroH2A and DNA methylation ensures maintenance of X-chromosome silencing (Figure 4).

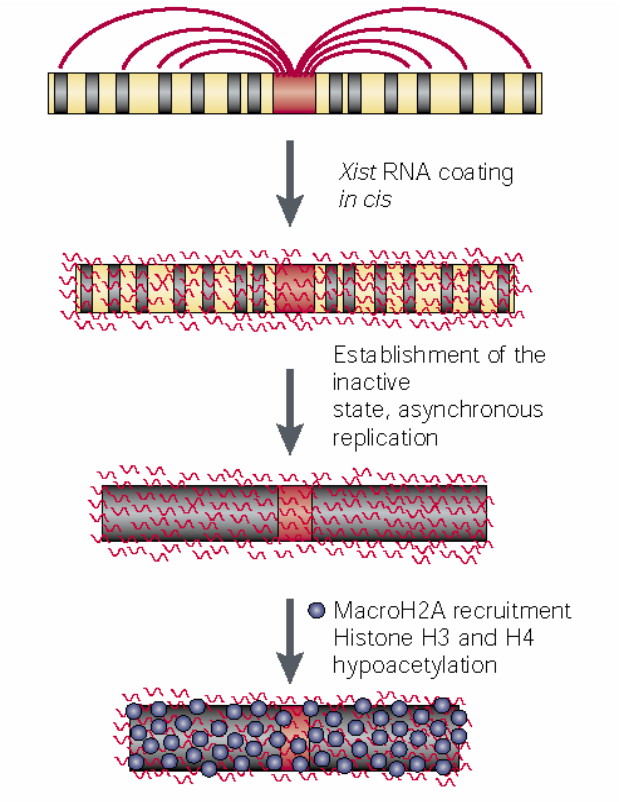


Figure 4. Mammalian X-chromosome inactivation.

Xist is expressed from the future inactive X-chromosome and coats it *in cis*. Initial deposition of chromatin repressive marks starts the process of silencing that is later maintained by macroH2A incorporation, H3 and H4 hypoacetylation and DNA methylation (taken from Avner and Heard, 2001).

Mammalian X-inactivation can be probably considered the best example of a biological process regulated by the interplay between different lncRNAs (Figure 5). Such interplay appears to be crucial during the X:A ratio measurement, that ensures that only one of the two Xs per diploid cell gets inactivated, and in the selection of the future inactive X-chromosome. In fact, several lncRNA genes are located at the *Xic* (X-inactivation center), which is necessary and essential for X-inactivation. One of these, the *Xist* antisense lncRNA *TsiX*, is necessary for repression of *Xist* expression in the active X-chromosome, probably through mechanisms that involve the induction of a heterochromatic state at the *Xist* promoter. On the other hand, *TsiX* expression is enhanced on the active chromosome by the two lncRNAs *Xite* and *DXPas34*. Conversely, *Xist* expression is promoted by two other lncRNAs, *JpX* and *RepA*.

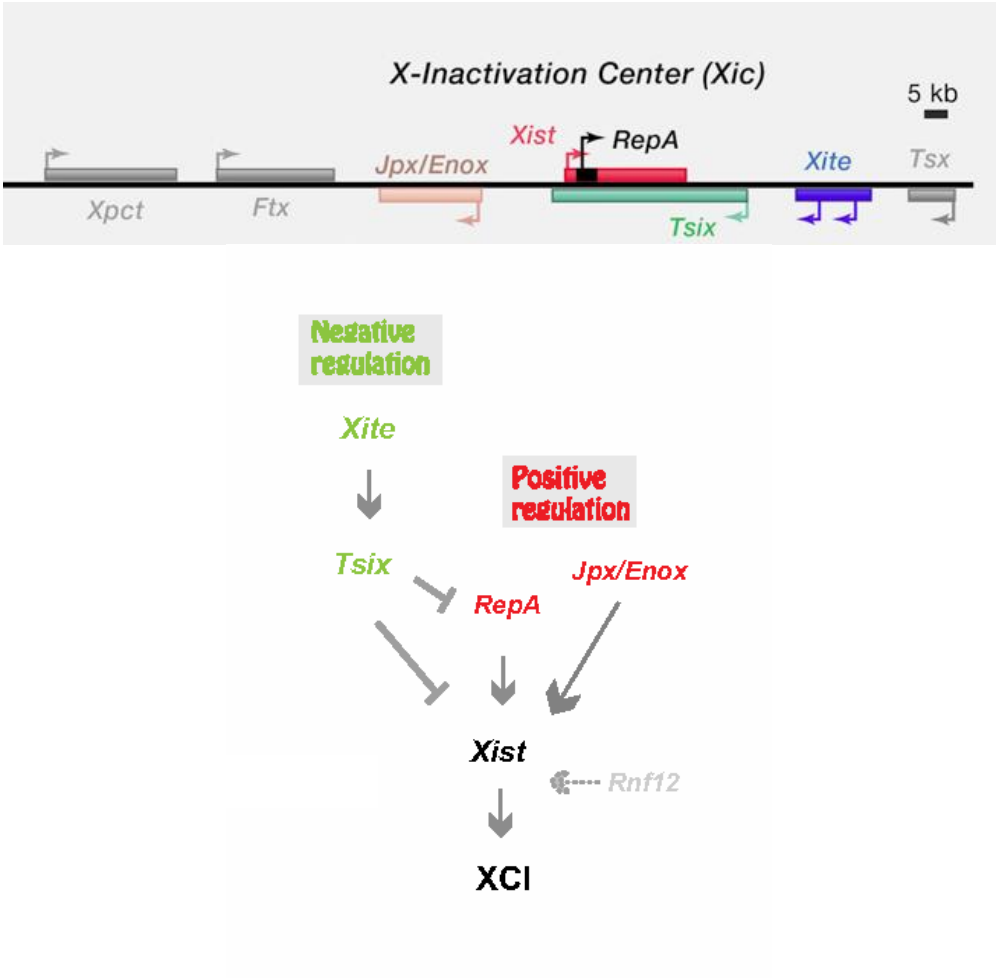


Figure 5. Regulation of X-inactivation by lncRNAs.

Several genes for lncRNAs are located in the X-Inactivation Center (*Xic*). The interplay between the positive (*RepA* and *Jpx/Enox*) and negative (*Xite* and *Tsix*) regulators determines the final expression of *Xist* and the consequent X-inactivation of the target X-chromosome (taken from Lee and Bartolomei, 2013; Di Tian and Lee, 2010)

3. Dosage compensation in *Drosophila*

3.1. Molecular mechanism of dosage compensation in *Drosophila melanogaster*

In *Drosophila,* equalization of X-linked transcript levels between females (XX) and males (XY) is achieved through the 2-fold hyper-transcription of the single X chromosome in the male organism (Gelbart and Kuroda, 2009). The dosage compensation complex (DCC) in *Drosophila* (also named *male-specific lethal* or MSL complex) is a ribonucleoprotein assembly made of the five proteins MSL1, MSL2, MSL3, MOF (*males absent on first*), MLE *(maleless),* and the two lncRNAs *roX1* and *roX2* (*RNA on the X*) (Figure 6).

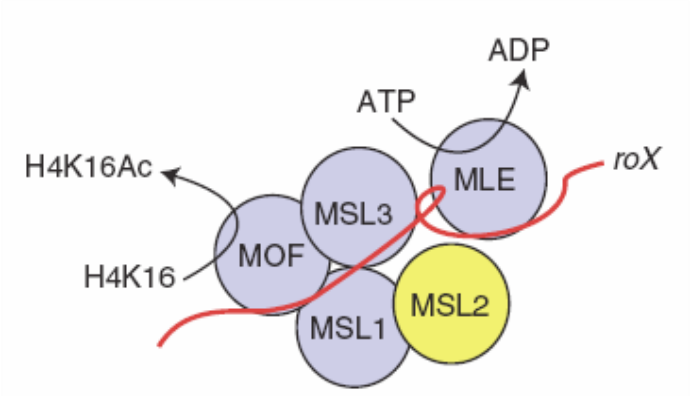


Figure 6. Schematic representation of the dosage compensation complex (DCC).
The protein and RNA components of the DCC are shown, as well as the enzymatic activities associated with some of the complex subunits.

In males, the DCC binds at the middle and 3' ends of hundreds of active genes, where it induces H4 acetylation at lysine 16, a chromatin modification that is thought to loosen DNA-nucleosome association, increasing the processivity of RNA Pol II (Smith et al., 2001; Larschan et al., 2011) and leading to a rough 2-fold increase in X-chromosome transcript levels. H4K16 acetylation is catalyzed by MOF, the acetyl-transferase of the DCC (Akhtar and Becker, 2000; Smith et al., 2000; Smith et al., 2001).
Similar to the DCC in *C.elegans,* recognition and binding of the *Drosophila* DCC to the X-chromosome occurs in a two step- manner (Figure 7). Initial binding occurs at the so called "high affinity sites" (HAS) or "chromatin entry sites" (CES), which are short DNA

fragments (<250 bp) that support DCC recruitment even when artificially transferred to autosomal loci (Dahlsveen et al., 2006; Fagegaltier et al., 2004; Oh et al., 2004). A conserved 21 bp GA-rich motif present at the HAS has been proposed to mediate DCC binding (Straub et al., 2008; Alekseyenko et al., 2008). However, this motif is only enriched 2-fold on the X-chromosome with respect to autosomes and is also present at non-HAS sites on the X, suggesting that other DNA or chromatin features are implicated in first X-chromosome recognition (Kageyama et al., 2001; Park et al.,2003). Binding to the HAS relies on the protein MSL1 and MSL2, the only components of the DCC that are able to bind to the X-chromosome even in the absence of MSL3 and MOF (Lyman et al., 1997). While the role of MSL1 in X-chromosome binding is still unclear, MSL2 has been proposed to recognize the DNA through its CXC domain (Bashaw and Baker, 1995; Fauth et al., 2010).

Once bound to the HAS, the DCC associates with closely active genes through recognition of H3K36me3, a chromatin mark associated with highly transcribed genes (Sural et al., 2008; Bell et al., 2008). These sites are commonly referred to as “low affinity sites” and correct binding requires the MSL-complex components MSL3 and MOF (Sural et al., 2008; Kind et al., 2008). Clustering of the HAS in a specific three-dimensional conformation is also thought to facilitate DCC binding to the low affinity sites through the creation of a local DCC enrichment (Grimaud et al., 2009). Nuclear pore-mediated localization of the X-chromosome at the nuclear periphery has also been suggested to promote X-chromosome hypertranscription, through the formation of a specialized transcriptional domain (Mendjan et al., 2006).

Although necessary, MSL1, MSL2 and MSL3 are not sufficient for correct binding of the DCC to the X-chromosome. Indeed, assembly of a fully functional complex requires the presence of the two lncRNAs *roX1* and *roX2*, and the functionally linked RNA helicase MLE.

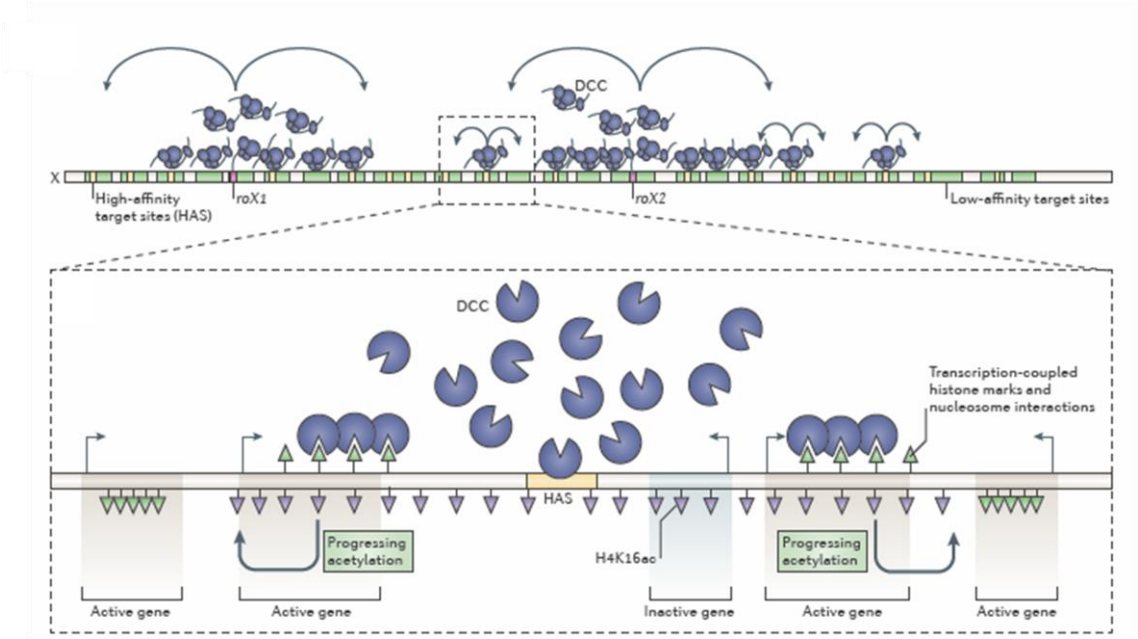


Figure 7. Targeting of the DCC to the X-chromosome.
Initial binding is mediated by the recognition of the high affinity sites (HAS) among which the *roX* loci. Spreading of the complex to the low affinity sites occurs through the interaction with chromatin marks of active transcription (taken from Conrad and Akthar, 2012)

3.1.1. *roX1* and *roX2* non-coding RNAs

3.1.1.1. Gene structure and transcript isoforms

roX1 and *roX2* are redundant lncRNAs expressed in males and essential for male survival (Amrein and Axel, 1997; Meller et al., 1997; Meller and Rattner, 2002). The two *roX* genes are located in the X-chromosome at positions 3F and 10C, which coincide with DCC HAS. Alternative splicing, promoter and 3' end cleavage site usage give rise to a series of different transcript isoforms (Figure 8).

roX1 can be found in two isoforms; the first (3748 nt) is transcribed from the most upstream promoter, contains a small intron and is cleaved at the closest 3' cleavage site. The second isoform (3460 nt) is produced from a downstream promoter, is spliced and is cleaved at the distal 3' cleavage site. (Amrein et Axel, 1997; Meller et al., 1997; Flybase). No difference in the function of the two isoforms has been reported so far.

A cluster of different isoforms results from the processing of *roX2* transcripts (Amrein et Axel, 1997; Park et. al, 2003; Park et al., 2005). The most abundant is 571 nt long,

contains the exons 1 and 3, and ends at the shorter 3' end (Figure 8, asterisk). Two other less abundant transcripts of about 1100 nt and 1400 nt have been observed that contain exons 2 and 3, and finish at different 3' ends. In addition, alternative splicing inside exon 2 can yield 21 different transcript isoforms (Park et al., 2005). The alternative splicing pattern displayed by exon 2 is conserved in different *Drosophila* species and all described isoforms have been found to associate with the DCC, suggesting a functional relevance. Indeed *roX* mutant male flies containing a *roX2* transgene that prevents alternative splicing of exon 2 show decreased DCC binding to the X-chromosome, although they are still viable (Park et al., 2005).

3.1.1.2. Functional domains

The two *roX* RNAs display deep differences in size and very low homology at their primary sequence (Franke et Baker, 1999). However, they are interchangeable for dosage compensation, indicating functional redundancy (Meller et al., 1997; Franke and Baker, 1999; Meller and Rattner, 2002). Low sequence conservation is also observed between *roX* genes within the *Drosophila* lineage. Nevertheless, *roX* RNAs from other *Drosophila* species are able to rescue DCC binding to the X-chromosome and male viability when expressed in *Drosophila melanogaster roX1⁻roX2⁻* mutants, suggesting conservation of *roX* function in *Drosophila* (Park et al., 2007, Park et al., 2008). Similar to other lncRNAs, functional conservation and redundancy between the *roX* RNAs could be explained by the presence of conserved short sequences or secondary structures inserted in an otherwise poorly conserved nucleotide context. In fact, a short conserved motif has been found at the 3' end of *roX1* and *roX2*, in 3 and 2 copies, respectively. This motif is called the roX-box and consists of GUUNUACG (Park et al., 2007; Park et al., 2008). One of the roX-boxes of *roX2* is located at the stem of a conserved stem-loop at the 3' end. A very similar sequence (GUUNUCCG) is also contained in the stem of a different conserved stem-loop structure in *roX1*. Interestingly, a transgene containing six tandem repeats of the *roX2* conserved stem loop was able to rescue the DCC binding and H4K16 acetylation defects of *roX⁻* male flies. Mutation of the roX2-box within the stem-loop led to loss of DCC binding and H4K16 acetylation (Park et al., 2007). These results suggest a functional relevance for the conserved stem-loop and roX-box sequences, although their role in *roX* function is still unclear.

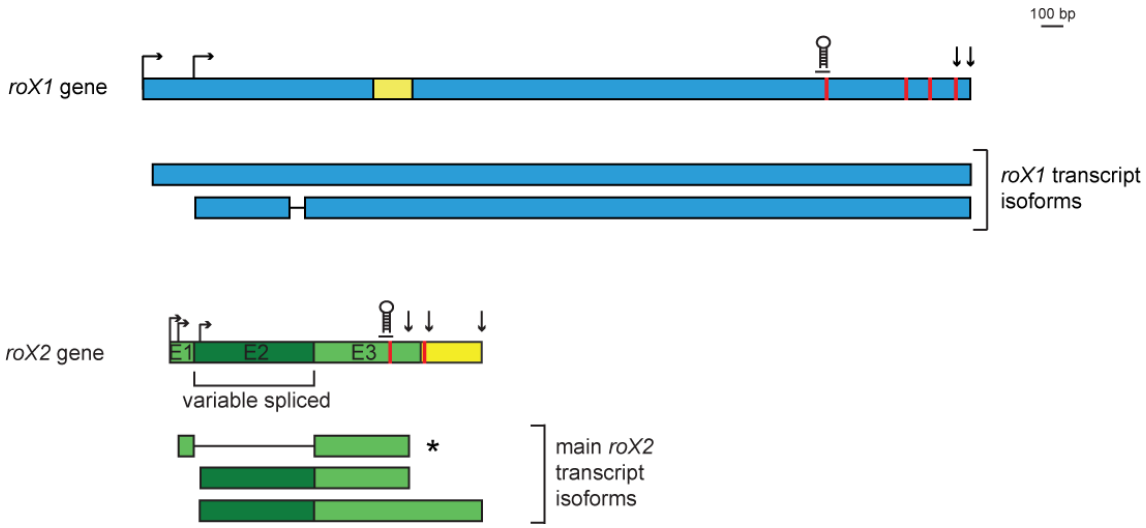


Figure 8. *roX1* and *roX2* gene structure, transcript isoforms and functional domains.
The gene structure and the transcript isoforms produced from *roX1* and *roX2* genes are shown (boxes: exons, lines: introns, horizontal arrows: transcription start sites, vertical arrows: 3' cleavage sites). The functional domains are depicted (yellow boxes: DHS, red lines: roX-boxes, functional stem-loops). The asterisk indicates the major *roX2* transcript isoform.

3.1.1.3. Function in dosage compensation

As mentioned above, *roX1* and *roX2* play a fundamental role in DCC targeting to the X-chromosome. Indeed in *roX*$^{-}$ male flies the WT pattern of MSL binding appears disrupted, with an almost complete displacement of the DCC from the X-chromosome to some sites on the autosomes and to the chromocenter (Meller and Rattner, 2002).
The X-chromosome localization and the physical association with the MSL proteins clearly indicate that the *roXs* are stable components of the DCC (Franke and Baker, 1999; Meller et al., 2000).
Incorporation of *roX* RNAs to the DCC is thought to occur through the recruitment of partial MSL-complexes by virtue of a 110 nt DNA sequence (DNase hypersensitive site or DHS) present at both *roX* loci (Kageyama et al., 2001; Park et al., 2003). The DHS cannot be recognized by DCC-subcomplexes in the absence of *roX* transcription, suggesting that a first initial association of the DCC with *roX* RNAs could be DHS-independent (Kageyama et al., 2001, Park et al., 2003). *roX* RNAs are loaded into the complex by the action of the ATPase/helicase MLE, which is an essential mediator of *roX* function in dosage compensation. Indeed in *mle* mutant males, *roX* RNAs are no longer found associated with the other MSL proteins, and are only detectable at their

site of transcription (Meller et al., 2000). The functional interconnection between MLE and *roX* is also suggested by the RNase sensitivity and the *roX* dependency of the association of MLE with the other components of the DCC complex (Richter et al., 1996; Oh et al., 2003).

Once incorporated, *roX* RNAs enable the DCC to correctly bind the target sites on the X-chromosome. Binding to the low affinity sites is thought to be promoted by *roX* incorporation, even though a high resolution analysis of the sites bound by the MSL-complex in the absence and presence of *roX* is still missing (Meller and Rattner, 2002). Moreover, recent analysis suggest that *roX* could also influence binding to the HAS (Straub et al., 2012).

The molecular switch that makes *roX*-containing DCCs competent for X-chromosome binding is unknown. The current hypothesis posits that *roX* RNAs function as scaffolding molecules and confer the DCC a structural architecture suitable for binding. In addition to MLE, MSL3 and MOF appear to associate to the complex in an RNase-sensitive manner, suggesting that *roX* could be involved. Moreover, all DCC components with the exception of MSL1 contain putative RNA-binding domains.

3.2. Regulation of dosage compensation in male flies

Dosage compensation is triggered in males by the regulated formation of the DCC through mechanisms that range from transcriptional to protein stability control. Mechanisms of homeostasis are employed to fine-tune the levels of each DCC component, maintaining the optimal subunit stoichiometry and ensuring correct DCC function. These mechanisms are briefly explained below.

DCC assembly is triggered by MSL2, the rate-limiting subunit of the complex. MSL2 expression is restricted to males. MSL2 promotes the stabilization of MSL1, which is responsible for the recruitment of two other DCC components, MSL3 and MOF (Chang and Kuroda, 1998; Scott et al., 2000). MSL2 alone or in partial MSL complexes is able to stimulate *roX* expression by increasing *roX* transcription and stability (Amrein and Axel, 1997; Meller et al., 1997; Meller et al., 2000; Bai et al., 2004). MLE, in turn, promotes the incorporation of the two *roX* RNAs into the complex (Meller et al., 2000; Gu et al., 2000; Meller and Rattner, 2002). MSL2 displays ubiquitin-ligase activity and adjusts the levels of MSL1, MSL3, MOF and its own through ubiquitin-mediated protein degradation (Villa et al., 2012). Such mechanism ensures the maintenance of constant protein subunit levels, avoiding over-expression and ectopic binding of the DCC to the

autosomes (Demakova et al., 2003; Park et al., 2002). Soluble MSL2 can also interact with its own mRNA and retain it in the nucleus, suggesting an autoregulatory feedback loop controlling MSL2 protein levels (Johansson et al., 2011). *roX* RNAs also show autoregulation (Lim et al., 2012) and have been proposed to regulate MSL2 levels (Lim et al., 2012).

Non-DCC factors have been shown to influence dosage compensation. The H3S10 kinase JIL-1, the chromatin binding protein SU(VAR)3-7 and the DNA supercoiling factor (SCF) have been proposed to regulate the chromatin organization of the male X-chromosome (Spierer et al., 2005; Furuhashi et al., 2006; Deng et al., 2008). Moreover the nuclear pore components Mtor and NUP153 associate with the DCC and are thought to mediate DCC localization at the nuclear X-territory (Mendjan et al., 2006). At the post-transcriptional level, other factors can affect DCC complex formation. The zinc-finger protein Zn72D and the RNA helicase Belle influence DCC formation by promoting the correct splicing of *mle* pre-mRNA (Worringer et al., 2007, Worringer et al., 2009). In addition, the RNA binding protein Upstream of N-ras (UNR) has been found to promote DCC targeting to the X-chromosome (Patalano et al., 2009). As UNR is the basis of my thesis, in the following I will extend on the roles of this protein.

3.2.1. Upstream of N-ras (UNR)

UNR is an evolutionarily conserved RNA binding protein that contains five cold shock domains (CSD) (Figure 9).

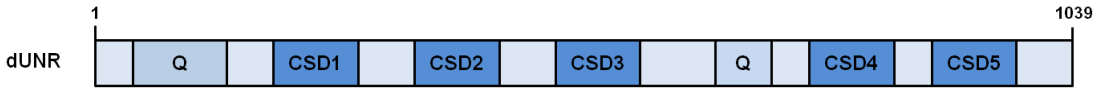


Figure 9. Schematic representation of UNR protein structure.

CSDs are OB-fold structures consisting of a β-barrel composed of five anti-parallel β-strands. Each CSD contains conserved RNP1 and RNP2 motifs that serve to bind single stranded nucleic acids, both RNA and DNA (Arcus, 2002) (Figure 10). CSDs can display a wide range of binding affinities and sequence recognition properties. In addition to nucleic acid binding, CSDs can also mediate protein-protein interactions (Leshkowitz et al., 1996).

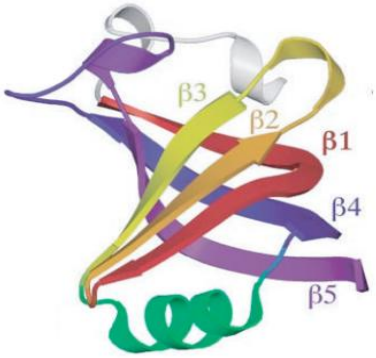


Figure 10. OB-fold structure.

Drawing of a classical OB-fold domain. The five β-sheets are indicated and depicted each with a different color (taken from Theobald et al., 2003).

In vitro selection/amplification experiments (SELEX) using human UNR have identified a consensus motif consisting of AAGUA/G or AACG followed by a purine stretch, to which the protein binds with a Kd of 10 nM (Triqueneaux et al., 1999). In these experiments, the full-length protein showed same sequence specificity as one single CSD; moreover, redundancy in sequence recognition by different CSDs was detected. Redundancy, however, has not been observed for cellular RNA targets (Mihailovich et al., 2010). No consensus binding motif is known for *Drosophila* UNR, even though the protein binds to thousands of transcripts, a large part of which in a sex-specific manner (Mihailovich et al., 2012).

Unr-deficient mice display embryonic lethality, suggesting a role for UNR in development and differentiation (Boussadia et al., 1997); functions in cell proliferation and death have also been reported (Dormoy-Raclet et al., 2007). UNR regulates transcript targets at the post-transcriptional level, concordant with its primarily cytoplasmic localization. In mammals, UNR regulates mRNA stability and translation, and is considered an IRES trans-acting factor (ITAF), as it regulates translation from a series of viral and cellular IRESes. A well-studied case is that of Apaf-1 mRNA; UNR binds to Apaf-1 IRES and modifies the structure of the IRES to allow for nPTB binding and consequent translation initiation (Mitchell et al., 2003).

3.2.2. Regulation of dosage compensation by UNR in male flies

UNR functions as a positive regulator of dosage compensation in male flies (Patalano et al., 2009). Hypomorph mutant UNR males that express a truncated form of UNR consisting of the amino-terminal half of the protein display defects in DCC binding to the X-chromosome. A reduction in binding of all DCC protein components was observed, and was especially strong in the case of MLE. In agreement with a function of UNR in DCC targeting to the X-chromosome, delocalization of MSL2 from the X-territory of the nucleus was also detected in UNR-depleted SL2 cells, which have male properties. Reduction of DCC targeting was not due to a reduced availability of the complex subunits, since total protein levels and nucleo-cytoplasmic distribution of the protein components were not affected by UNR depletion. Interestingly, however, UNR was found to interact with both *roX1* and *roX2*, suggesting that UNR could function through the lncRNAs to regulate dosage compensation.

3.3. Regulation of dosage compensation in female flies

As mentioned above, MSL2 is the limiting subunit of the DCC. Enforced expression of MSL2 in females leads to DCC formation and female lethality (Ref). Females repress dosage compensation by inhibiting the expression of the *msl2* gene. The factor that orchestrates *msl2* repression is a female-specific RNA binding protein that also functions as the master sex-determination switch in *Drosophila*, the protein Sex-lethal (SXL) (reviewed in Penalva et al., 2003). SXL, thus, links female sex-determination with repression of dosage compensation, and it does so by regulating the expression of target genes at the post-transcriptional level. To repress dosage compensation, SXL needs to interact and form a complex with UNR.

3.3.1. Sex-lethal (SXL)

SXL is a ~35 kDa protein containing a glycine/asparagine (GN)-rich amino terminus followed by two RRM- type RNA binding domains (RBDs) (Figure 11A). RBDs are among the most common RNA binding domains and consist of a four-stranded antiparallel β-sheet packed against two α-helices (Nagai et al., 1990). The β-sheets contain the RNA binding motifs RNP1 and RNP2, through which the protein binds to

RNA. In the case of SXL, the two RBDs are necessary for high specificity and affinity of binding (Samuels et al., 1998).

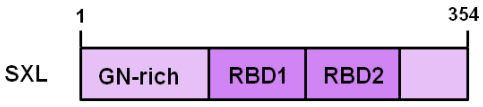
Mutational analysis of natural SXL targets has identified a sequence for binding consisting of a polyuridine tract of eight or more nucleotides that can be interrupted by guanosines, but not by cytosines (Sakamoto et al., 1992; Sosnowski et al., 1989). Stretches of five or seven Us can also be bound by SXL, although with lower affinity (Valcarcel et al., 1993; Samuels et al., 1994; Wang and Bell, 1994). SELEX experiments are consistent with the consensus $U_6GU_2G/UUG/UU_3G/UU_2$ (Sakashita and Sakamoto, 1994; Singh et al., 1995). The presence of an adenosine immediately upstream of the U-stretch increases SXL affinity, while disruption of the polyuridine tract with cytosines greatly decreases SXL binding (Sosnowski et al., 1989; Inoue et al., 1990; Samuels et al., 1994; Singh et al., 1995). Crystal structures of soluble SXL or SXL complexed with RNA have revealed that the two RBDs, which are flexible in solution, establish interdomain contacts upon RNA binding. RNA-bound RBDs create a V-shape cleft, where the RNA is sandwiched in an extended conformation (Crowder et al., 1999; Handa et al., 1999) (Figure 11B).

SXL RBDs have also been proposed to be involved in homo-dimerization (Samuels et al., 1998; Wang and Bell, 1994). In addition, cooperative binding of SXL molecules to multiple sites as well as SXL interaction with other factors requires the N-terminal GN-rich region, at least in some instances (Sakashita and Sakamoto, 1996; Samuels et al., 1998; Wang and Bell, 1994).

Sxl expression is exclusively initiated in the female embryo through a mechanism of "sex-recognition" based on the X-chromosome to autosome ratio (X:A). Transcription factors expressed from the X (numerators) and from the autosomes (denominators) are respectively positive and negative regulators of early *Sxl* expression. In females, where the X:A ratio is equal to 1 (2X:2A), the activation activity of numerators is prevalent, and *Sxl* is transcribed from the early or establishment promoter (PE). In males, where the X:A ratio is 0.5, *Sxl* transcription remains silent (Schütt and Nöthiger, 2000). Later on in development, *Sxl* expression from the PE promoter ceases and a second maintenance promoter (PM) is activated in both male and female flies. However, restriction of *Sxl* expression to the female organism is ensured by a mechanism that allows for production of new *Sxl* products only in the presence of pre-existing SXL protein. Such positive autoregulatory loop relies on the ability of SXL to bind to its own pre-mRNA and induce skipping of an exon that contains stop codons, thereby allowing expression of a full-length protein. In males, the exon is retained due to the absence of

pre-existing SXL, and functional SXL cannot be synthesized (Lallena et al., 2002). A similar splicing regulatory mechanism is employed by SXL to regulate downstream targets in the sex-determination pathway, such as *transformer* (Valcarcel et al., 1993).
SXL has also been proposed to adjust its own protein levels through a negative autoregulatory mechanism that involves translational repression (Yanowitz et al., 1999). A role for SXL in translational regulation has also been proposed for the inhibition of *nanos* during differentiation of germ stem cells into adult ovarian cells (Chau et al., 2012). The best characterized, although not completely understood, example of translational repression by SXL is that of inhibition of *msl2* to block dosage compensation in female flies (reviewed in Graindorge et al, 2011). In this case, SXL coordinates several post-transcriptional regulatory mechanisms to ensure that expression of *msl2* is shut-off in female flies. First, SXL inhibits the splicing of an intron in the 5' UTR of *msl2* pre-mRNA (Gebauer et al., 1998; Forch et al., 2001; Merendino et al., 1999). Second, SXL promotes nuclear *msl2* mRNA retention in concert with HOW, a SXL co-factor that binds to the 5'UTR intron (Graindorge et al., 2013). Finally, SXL bound to both the 5' and 3' UTRs inhibits the translation of *msl2* mRNA (Kelley et al., 1997; Bashaw and Baker, 1997; Gebauer et al., 1998). To inhibit translation, SXL needs to recruit UNR to bind in close proximity in the 3' UTR (Abaza et al., 2006; Duncan et al., 2006).

A



B

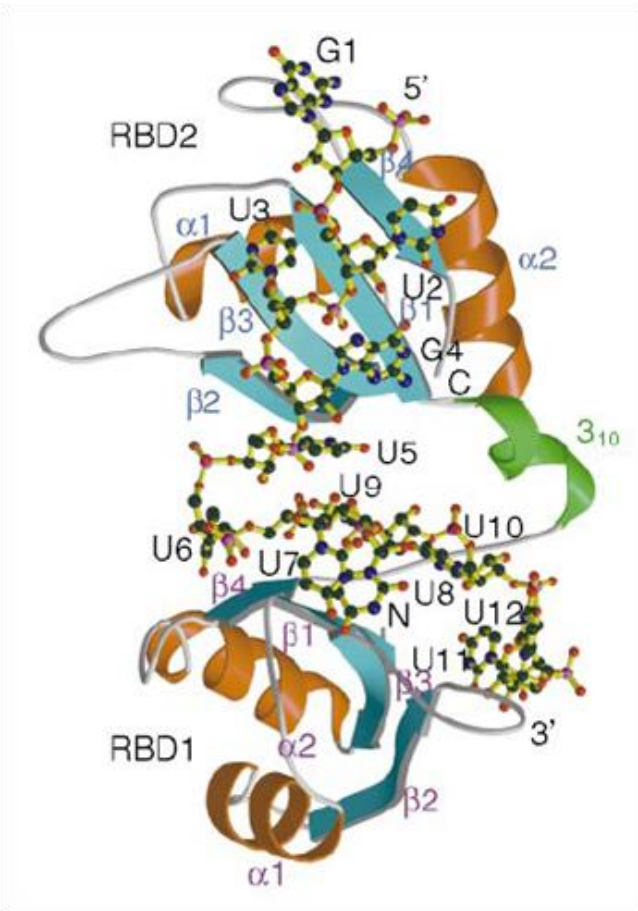


Figure 11. SXL protein structure.

(A) Schematic representation of SXL protein structure. The N-terminal glycine/asparagine (GN)-rich domain and the two RBDs (1 and 2) are indicated.

(B) Structure of SXL protein complexed with a 12 nucleotides RNA derived from *transformer* transcript. The two RBDs are indicated; β-sheets are depicted in cyan, helices in orange, the linker helix in green and random coils in grey. The RNA is represented by a ball-and-stick model. RBDs create a V-shape cleft, where the RNA is sandwiched in an extended conformation (taken from Handa et al.,1999)

3.3.2. Inhibition of *msl2* translation by SXL and UNR

In order to understand how SXL and UNR cooperate to regulate *msl2* mRNA, we must first briefly review the process of translation. mRNA translation can be subdivided in four steps: initiation, elongation, termination and recycling. Initiation is the rate-limiting step of translation and for this reason is the most common target of translational control mechanisms, although regulation can occur at every step of translation. In eukaryotes, translation initiation consists of the positioning of the ribosome at the AUG start codon and requires the concerted action of more than 30 polypeptides known as eukaryotic initiation factors (eIFs) (Sonenberg and Hinnebusch, 2009). Initiation can be divided in: i) formation of the 43S pre-initiation complex, an assembly of the small ribosomal subunit loaded with initiation factors; ii) recruitment of 43S complexes to the 5' UTR of the mRNA in a cap-proximal position; iii) scanning of the 43S complexes along the 5' UTR and positioning at the start codon; and iv) joining of the 60S ribosomal subunit (Figure 12). Each of these steps can be targeted for mRNA specific translational control through the action of proteins bound to the 5' or 3'-UTRs (Abaza and Gebauer, 2008; Szostak and Gebauer, 2013). SXL inhibits two of these steps: 43S complex recruitment and scanning (Gebauer et al., 2003; Beckmann et al., 2005).

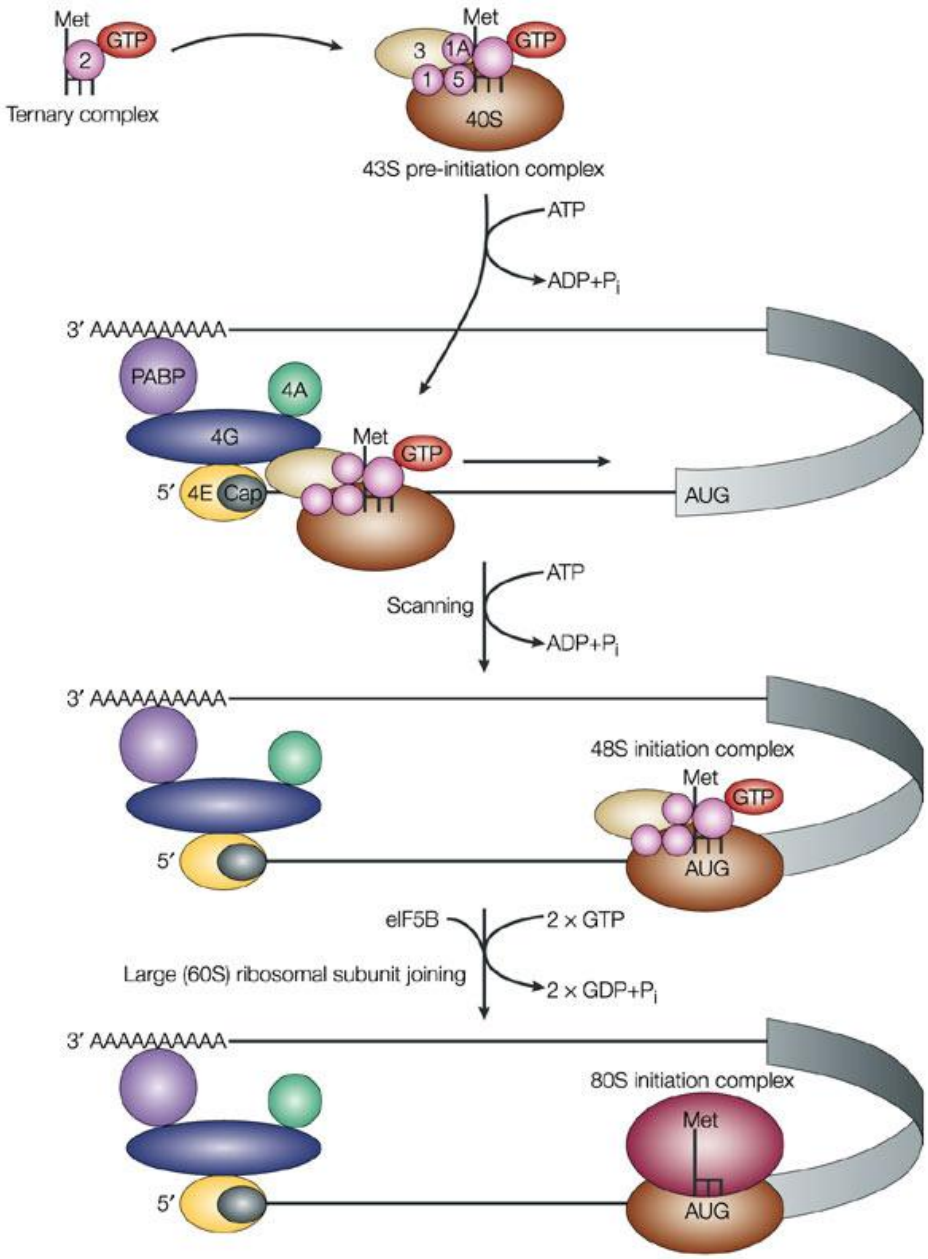


Figure 12. Mechanism of translation initiation.

The 43S pre-initiation complex is composed of the 40S ribosomal subunit associated with the translation factors eIF3, eIF2, 1A and the ternary complex (Met-tRNAi-eIF2-GTP).The first step of translation initiation is the recognition of the cap-binding complex through the interaction of eIF3 with eIF4G. The cap-binding complex also contains 4E and 4A and interacts with the PABP to promote formation of the closed loop. After mRNA binding, the 43S complex scan the 5'UTR until the appropriate initiation codon (AUG) is reached. Stable binding of the 43S complex to the AUG leads to the formation of the 48S initiation complex. Upon GTP hydrolysis

the 60S ribosomal subunit joins the 48S complex and the 80S ribosome assemble to start translation elongation (taken from Gebauer and Hentze, 2004).

Ribosomal 43S complex recruitment is probably the most frequently targeted step for regulation. Ribosomes are attracted to the mRNA by the cap-binding complex, consisting of the cap-binding factor eIF4E, the scaffolding protein eIF4G and the RNA helicase eIF4A. eIF4G also interacts with PABP bound to the poly(A) tail at the 3'end of the mRNA, mediating the formation of what is commonly known as the "closed-loop" structure, a conformation that is believed to enhance translation by facilitating ribosome recycling. In most cases, RNA-binding proteins block the formation of the closed-loop or inhibit the interaction of ribosome-bound eIFs with the cap-binding complex. In the case of *msl2*, however, translational repression occurs without disruption of the cap-binding complex or the closed-loop, suggesting a novel mechanism of regulation (Duncan et al., 2009). SXL bound to the 3' UTR of *msl2* attracts the co-factor UNR to bind in close proximity and, together, SXL and UNR inhibit 43S ribosome recruitment (Gebauer et al., 2003; Abaza et al., 2006). SXL bound to the 5' UTR inhibits the scanning of 43S complexes that have presumably escaped recruitment inhibition. The capacity of SXL to promote ribosome recognition of an AUG located upstream of the SXL binding site in the 5' UTR (uAUG3) contributes to scanning inhibition and leads to diminished translation at the major *msl2* ORF (Medenbach et al., 2012).

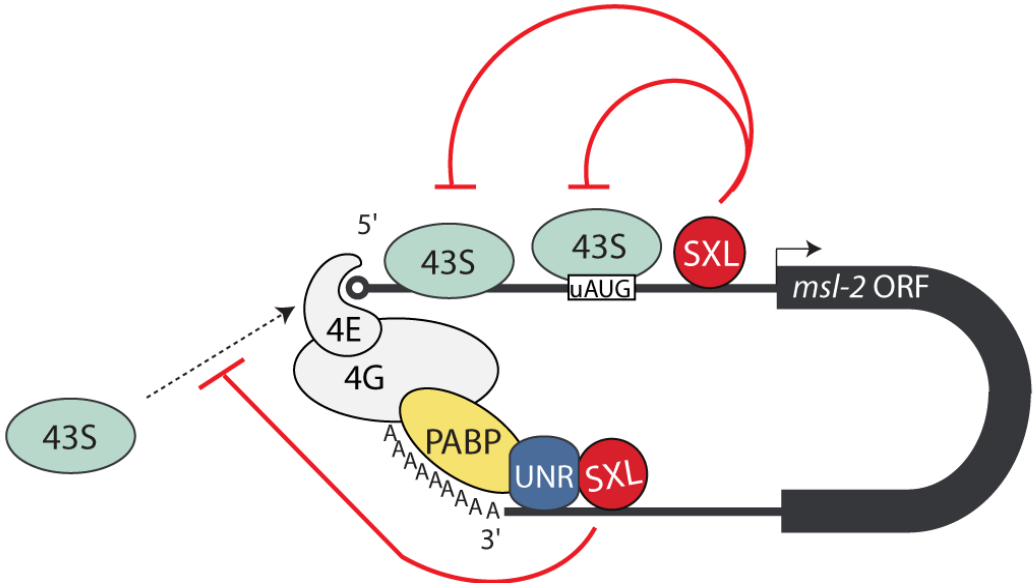


Figure 13. Mechanisms of *msl2* translational repression.

Binding of the SXL-UNR corepressor complex to the *msl2* 3' UTR inhibits the recruitment of the 43S complex. SXL bound to the 5' UTR represses scanning of those 43S complexes that have escaped the 3' UTR mediated repression; in addition it forces 43S complexes to recognize an upstream AUG (uAUG) thus interfering with translation of the major *msl2* ORF.

OBJECTIVES

The objective of this thesis is to understand the function of *Drosophila* UNR in the regulation of dosage compensation in male and female flies.

Specifically we aim to elucidate:

- the molecular mechanism by which UNR promotes dosage compensation in male flies

- the structural elements responsible for the recruitment of UNR to the *msl2* mRNA by SXL.

RESULTS

PART I: REGULATION OF DOSAGE COMPENSATION BY UNR IN MALES

Publication:

Drosophila Upstream of N-Ras acts as an RNA chaperone on *roX2* non coding RNA and facilitates the recruitment of MLE during X-chromosome dosage compensation

Cristina Militti[1], Sylvain Maenner[2], Peter B. Becker[2] and Fátima Gebauer[1*]

(1) Gene Regulation, Stem Cells and Cancer Programme, Centre for Genomic Regulation (CRG), Dr Aiguader 88, 08003-Barcelona, Spain, and Universitat Pompeu Fabra (UPF), 08003-Barcelona, Spain
(2) Adolf-Butenandt-Institute and Center for Integrated Protein Science, Ludwig-Maximilians University, D-80336 Munich, Germany

∗Corresponding author:
Tel: +34-93 3160120
Fax: +34-93 3969983
E-mail: fatima.gebauer@crg.eu

Running title: UNR promotes the association of *roX2* with MLE
Keywords: UNR, *roX2*, non coding RNA, RNA chaperone, dosage compensation

Summary

The *Drosophila* long non-coding RNAs (lncRNA) *roX1* and *roX2* are essential components of the dosage compensation complex (DCC). The *roX* RNAs are required for the targeting and spreading of the DCC along the male X-chromosome, yet their molecular mechanisms of action remain obscure. Here we report that the RNA binding protein Upstream of N-Ras (UNR) binds close to and promotes a conformational change at a conserved, functional stem-loop of *roX2*. UNR also interacts with the RNA helicase MLE, a critical DCC subunit, and enhances MLE association with *roX2* both *in vitro* and *in vivo,* which is required for appropriate DCC targeting to the X-chromosome. Our results uncover a direct role of an RNA binding protein on lncRNA function in *Drosophila* dosage compensation.

Introduction

Long non-coding RNAs (lncRNAs) are RNA species longer than 200 nucleotides with no evident coding potential. Although lncRNAs have been involved in several cellular processes, little is known about the molecular mechanisms underlying their roles (reviewed in Wilusz et al., 2009; Koziol and Rinn, 2010; Nagano and Fraser, 2011; Lee, 2012). The *Drosophila RNA-on-X* (*roX*) *1* and *2* lncRNAs participate in the process of dosage compensation, a mechanism that corrects the imbalance of X-linked gene content between females (XX) and males (XY). In flies, dosage compensation involves the binding of the dosage compensation complex (DCC, also referred to as the male-specific-lethal or MSL complex) to hundreds of sites on the single male X-chromosome and the subsequent 2-fold hypertranscription of active genes (reviewed in Gelbart and Kuroda, 2009; Lucchesi, 2009; Conrad and Akhtar, 2012; Straub and Becker, 2011; Larsson and Meller, 2006). *RoX1* and *roX2* are components of the DCC together with five proteins, some of which display enzymatic activities: MSL1, MSL2, MSL3, the RNA helicase/ATPase Maleless (MLE) and the histone acetylase Males-absent-on-first (MOF). MSL2 is the rate-limiting subunit of the DCC and nucleates complex formation at specific X-chromosomal sites known as the high affinity sites or HAS. High resolution Chip-Seq analyses have revealed that primary contacts of MSL2 and MLE define the HAS (Straub et al., 2013). MSL2 also stabilizes MSL1, which serves as a scaffolding protein for the assembly of a full DCC (Chang and Kuroda, 1998; Scott et al., 2000). From the HAS, the DCC spreads to adjacent active genes via interactions of MSL3 with

chromatin (Sural et al.,2008; Bell et al., 2008; Straub et al., 2013). MOF promotes the acetylation of histone H4 on lysine 16 (H4K16), a modification that specifically marks the compensated X-chromosome leading to hypertranscription of its genes (Akhtar and Becker, 2000; Smith et al., 2000; Smith et al., 2001).

MLE facilitates the incorporation of *roX* into the complex, and this is important for both targeting and spreading of the DCC along the X-chromosome (Meller et al., 2000; Gu et al., 2000; Meller and Rattner, 2002; Li et al, MCB 2008). Although how the *roX* RNAs function is unclear, the finding that four of the DCC components (MSL3, MOF, MLE and MSL2) display RNase-sensitive association and/or contain potential RNA-binding domains suggests a role for the two *roXs* as scaffolding molecules. According to this hypothesis, the incorporation of the two non-coding RNAs would be essential for the assembly of DCC complexes competent for efficient X-chromosomal binding. Genetic and biochemical studies suggest redundant roles for *roX1* and *roX2*. Males that are devoid of both *roX* RNAs display an almost complete displacement of the DCC from the X-chromosome (Meller et al., 2000). Interestingly, a conserved stem-loop structure located at the 3' end of *roX2* and *roX1* is sufficient to restore X-chromosomal targeting of the DCC (Park et al., 2007; Park et al., 2008). Recently, Maenner and colleagues have shown that MLE recognizes and remodels this conserved stem loop, thereby promoting MSL2 binding to *roX* (Maenner et al., 2013).

We have previously shown that the RNA-binding protein Upstream of N-ras (UNR) promotes DCC targeting in male flies (Patalano et al., 2009). Mutant male flies containing a truncated form of UNR show decreased DCC binding to the X-chromosome, and depletion of UNR from male SL2 cells causes loss of DCC localization at the nuclear X-territory. Here we elucidate the molecular mechanism by which UNR regulates dosage compensation. We show that UNR binds proximal to and remodels the conserved stem-loop of *roX2*. UNR also interacts with MLE and facilitates MLE association with *roX2* both *in vitro* and *in vivo*, leading to appropriate DCC targeting to the X-chromosome. These results show that UNR has a direct role on DCC assembly via interactions with the lncRNA *roX2*.

Results and Discussion

UNR does not regulate *roX2* RNA metabolism

UNR is a conserved RNA-binding protein that regulates mRNA translation and stability in mammals and *Drosophila* (reviewed in Mihailovich et al., 2010). A role for UNR in the regulation of mRNAs encoding DCC components was unlikely, because UNR depletion causes no defect in the levels or intracellular distribution of DCC protein components (Patalano et al., 2009). As UNR binds to the lncRNAs *roX1* and *roX2* (Patalano et al., 2009), we set to investigate whether UNR affects *roX* metabolism. We used SL2 cells, which only express *roX2* and display fully functional dosage compensation (Smith et al., 2000; Straub et al., 2005).

The structure and processing of the *roX2* gene have been previously characterized (Park et al., 2003; Park et al., 2005). Alternative splicing and differential promoter and 3' cleavage site usage yield a cluster of transcript isoforms, of which the most abundant contains 571 nt with the functional stem-loop in a most 3' position (Figure 1A). We first tested whether UNR affects *roX2* steady state levels by depleting UNR from SL2 cells and evaluating *roX2* levels using RT-qPCR with oligonucleotides hybridizing to exon 3, common to all *roX2* transcript isoforms. The results showed no significant alteration of *roX2* levels upon UNR depletion (Figure 1B). We next tested whether UNR depletion changed the ratio of *roX2* transcript isoforms, as variable splicing of *roX2* has been reported to influence DCC targeting (Park et al., 2005). We used Northern analysis to visualize the three major isoforms of *roX2*, and found no difference in their relative abundance (Figure 1C). Finally, we tested whether UNR affected the intracellular distribution of *roX2*, a transcript that is primarily located in the nucleus and whose misallocation to the cytoplasm could in principle preclude its interaction with other components of the DCC (Franke and Baker, 1999). Using UNR depletion followed by nucleo-cytoplasmic cell fractionation, we detected no differences between control and depleted cells (Figure 1D). Altogether, these results indicate that UNR does not affect the processing, levels or distribution of *roX2* RNA.

UNR binds to *roX2* in a stem-loop proximal position

To gain insight into a putative role of UNR in *roX2* function, we tested whether UNR binds to *roX2* directly, and mapped its binding site. We performed gel mobility shift assays with recombinant UNR and radiolabelled RNA fragments of similar size spanning almost the entire length of *roX2* (Figure 2A). UNR could bind with low affinity

to most of the fragments, and bound with high affinity (~10 nM) to fragment E (Figure 2B). Consistent with the observation that mammalian UNR binds purine-rich sequences, fragment E is enriched in CAAUA repeats and contains a purine-rich stretch (Figure 2C) (Triqueneaux et al., 1999; Mitchell et al., 2003). Interestingly, fragment E is located just upstream of the functional stem-loop of *roX2*.

UNR is a protein that contains five cold-shock domains and has been shown to modify the structure of the Apaf-1 internal ribosome entry site (IRES) to permit the interaction of nPTB and allow translation initiation (Mitchell et al., 2003). Thus, as bacterial cold-shock proteins, UNR seems to behave as an RNA chaperone in certain molecular scenarios (Graumann and Mahariel, 1998; Rajkowitsch et al., 2007). Given the position of the UNR binding site on *roX2* RNA, we hypothesized that UNR could modify the structure of the functional stem-loop to permit further interactions for correct DCC formation. To test this hypothesis, we performed chemical and enzymatic footprinting of UNR on *roX2* RNA. An example of these footprinting experiments is shown in Figure 3A. In the absence of UNR, fragment E and surrounding sequences fold into a structure containing two loops (loops 1 and 2) that is placed upstream of the functional stem-loop (Figure 3B, left panel) (Maenner et al., 2013). Addition of UNR strongly protected two AC dinucleotides in loop 1 and most nucleotides of loop 2 (Figure 3B, right panel, black lines). Given the gel mobility shift assays, these contacts provide a high affinity binding site for UNR. High affinity binding likely depends on the AC dinucleotides, because UNR interacts with low affinity to a fragment lacking these dinucleotides (fragment F, Figure 2B). Although more weakly, UNR also protected the loop of the functional stem-loop (grey line). Importantly, UNR also increased the accessibility of the nucleotides at the base of the stem (dashed lines). These results suggest a role for UNR as an RNA chaperone of *roX2*.

UNR promotes *roX2* association with MLE

Interactions of the RNA helicase MLE with *roX2* are essential for DCC complex assembly and spreading (Meller et al., 2000; Gu et al., 2000; Meller and Rattner, 2002; Li et al., 2008). Recent data indicate that MLE remodels the functional stem-loop of *roX2* in an ATP-dependent fashion, allowing for specific association of MSL2 (Maenner et al., 2013). RNA helicases can associate with RNA chaperones to promote RNA remodeling (Hunger et al., 2006). We, thus, asked whether UNR interacts with MLE. Although UNR is primarily cytoplasmic (Abaza et al., 2006 and data not shown), a small amount can be found in the nucleus that interacts with MLE in co-

immunoprecipitation assays (Figure 4A). This interaction is specific, as UNR does not interact with MSL3. In addition, the interaction is probably transient, because we could only detect it under mild formaldehyde crosslinking conditions.

We next interrogated the functional consequences of UNR binding to *roX2* and MLE. Intriguingly, the region of the *roX2* functional stem-loop that is remodeled by UNR serves as binding site for MLE (Maenner et al., 2013). MLE helicase activity requires a 3'-tailed single stranded region to efficiently unwind dsRNA (Lee et al., 1997). Thus, melting of the base of the stem by UNR and exposure of the corresponding nucleotides could allow for efficient MLE interactions with *roX2*. To test whether UNR promotes the interaction of MLE with *roX2* we employed *in vitro* RNA pull-down assays. We used full-length *roX2*, or derivatives lacking the high-affinity UNR binding site (roX2-ΔCA) or the functional stem-loop (roX2-ΔSL), and tested the binding of recombinant UNR and MLE to these RNAs. We found that, in the absence of UNR, MLE could not associate to any of the *roX2* isoforms (Figure 4B, lanes 2-4). UNR, on the contrary, could bind in the absence of MLE (lanes 5-7). Deletion of the high-affinity UNR binding site reduced, but did not eliminate UNR binding, consistent with the capacity of UNR to bind with low affinity to other parts of *roX2* (lane 6). Importantly, MLE bound strongly to *roX2* upon addition of UNR (lane 8), and this binding was reduced upon deletion of either the UNR binding site or the stem-loop (lanes 9 and 10). These results indicate that UNR promotes the association of MLE with *roX2 in vitro*, and that this association partly depends on the UNR binding site and the stem-loop. To test whether UNR promotes *roX2* and MLE interactions also *in vivo*, we depleted UNR from SL2 cells and measured the amount of endogenous *roX2* associated to MLE upon MLE immunoprecipitation. The results showed that UNR depletion indeed reduced the amount of *roX2* associated to MLE (Figure 4C, right panel). This difference could not be attributed to variations in MLE or *roX2* amounts upon UNR depletion, or to differences in the efficiency of MLE immunoprecipitation (Figure 4C, left and middle panels). We conclude that UNR promotes the association of *roX2* with MLE both *in vitro* and *in vivo*.

Our data indicates that UNR plays a direct role in dosage compensation by remodeling the structure of the lncRNA *roX2* and promoting MLE association. UNR is the first RNA binding protein shown to directly function in dosage compensation in *Drosophila*, and adds to the small list of RNA binding proteins involved in dosage compensation in other organisms. Such factors include hnRNP U and YY1, which have been proposed to

tether the lncRNA *Xist* to the inactive X-chromosome during mammalian dosage compensation (Hasegawa et al., 2010; Jeon and Lee, 2011). Rather than a molecular tether, the role of UNR is more transient. UNR interacts with MLE but not with MSL3, suggesting that UNR is not stably bound to the DCC. Even if transient, the RNA chaperone function of UNR might be relevant, as UNR hypomorph mutants display a dramatic decrease of DCC localization to the X-chromosome (Patalano et al., 2009) and the DCC does not associate to dosage compensated genes in UNR depleted cells (Figure 4D). RNA structural dynamics is at the basis of many fundamental post-transcriptional processes (Dethoff et al., 2012). Understanding how RNA remodelers participate in the regulation of the activity of lncRNAs may reveal an important role of RNA structural transitions also at the transcriptional level.

Experimental procedures

Cell culture and RNAi treatment

Cell culture and RNAi treatment were performed essentially as described, by incubating $1.5x10^6$ cells with 15 µg of dsRNA corresponding to the UNR coding region (nt 2139-2691 relative to the start codon). Cells incubated with dsRNAs against GFP were carried in parallel as control. Cells were recovered 6 days after plating and the efficiency of UNR depletion was monitored by western blot using anti-UNR (Abaza et al. 2006) and anti-tubulin (Sigma) sera.

Quantitative RT-PCR

Total RNA from control and UNR-depleted SL2 cells was extracted using Trizol and treated with Turbo DNAse (Ambion). First-strand cDNAs were synthesized from 500 ng of total RNA with Superscript II (Invitrogen). Parallel samples without reverse transcriptase were carried as control. The reaction mixture was diluted and amplified by quantitative PCR using the Power SYBR Green kit (Applied Biosystem) and the following gene-specific primers: *roX2* 5'-TTCTCCGAAGCAAAATCAAGC-3' and 5'-ACAAGCGCGTCAACCATGAA-3', *actin* 5'-ACGAGTTGCCCGATGGACAG-3' and 5'-GCACAGTGTTGGCGTACAGA-3'. qPCR was performed on an Applied Biosystem machine and the amplification curves were analyzed using the associated software. Quantitative values were normalized to the internal standard *actin*.

Sub-cellular fractionation and Northern blot

Nucleo-cytoplasmic fractionation was performed by resuspending cells in 3 volumes of hypotonic buffer (10 mM HEPES pH 7.6, 10 mM K-Acetate, 0.5 mM Mg-Acetate, 0.5 % Triton X-100, 5 mM DTT, 1x protease inhibitor cocktail from Roche). Cells were incubated on ice for 5 min, centrifuged and the supernatant recovered as the cytoplasmic fraction. Nuclear pellets were washed with PBS, and RNA was extracted from the nuclear and cytoplasmic fractions using Trizol.

For Northern blot analysis of *roX2* subcellular localization, 10 µg of total RNA, 2 µg of nuclear RNA and an amount of cytoplasmic RNA proportional in volume to the nuclear fraction were resolved in a 2.5% denaturing agarose gel, transferred, and hybridized to the following oligonucleotide probes: *roX2* 5'-ATGTTGCGTTCCAAGACACA-3'; *U3* 5'-GATGCGAGGCACCACAAAGA-3'; *S18* 5'-CCAAGTAACTGTTAACGATCTAAGGAACC-3'. Northern blot of *roX2* transcript isoforms was performed on 25 µg of total RNA using a random-primed probe against *roX2* isoform C.

Immunoprecipitation and Western Blot

UNR was immunoprecipitated from SL2 nuclear extracts prepared as follows. Briefly, SL2 nuclei were isolated as described above and crosslinked with 0.5% formaldehyde at room temperature for 10 min. Crosslinking was quenched by addition of 125 mM glycine. Nuclei were washed with PBS and proteins were extracted with RIPA buffer (150 mM NaCl, 10 mM TRIS pH 7.5, 0.1% SDS, 1% deoxycholate, 5 mM EDTA, 1% Triton X-100, 1x protease inhibitor cocktail from Roche) after incubation for 20' on ice and sonication.

Purified anti-UNR IgGs were bound to Protein A dynabeads (Invitrogen) and UNR immunoprecipitation was performed for 1 h at 4ºC. Beads were washed three times with 10 vol of 1XNET (50 mM TRIS-HCl pH 7.5, 150 mM NaCl, 0.1% NP40, 1 mM EDTA) and resuspended in SDS buffer. Western blots were performed with anti-UNR, anti-MLE and anti-MSL3 sera at dilutions of 1:1000, 1:2000 and 1:1000, respectively.

MLE was immunoprecipitated from extracts obtained by resuspending nuclei in Triton-X buffer (0.5 mM EDTA, 150 mM NaCl, 50 mM Tris-HCl pH 7.5, 1% Triton, 1x protease inhibitor cocktail) and incubating the suspension for 1 hour on ice. After immunoprecipitation, beads were divided in two aliquots, one for detection of MLE by Western blot, the other for detection of *roX2* after proteinase K elution followed by Trizol treatment and quantitative RT-PCR. Prior to proteinase K elution, *Firefly*

luciferase RNA was added as a spike-in control, and was used as a standard in the RT-qPCR analysis. *U3* snoRNA was used to normalize the input samples. *Firefly* luciferase and *U3* were detected using the following oligos: *Luc* 5'-AACACCCCAACATCTTCGAC-3' and 5'- TTTTCCGTCATCGTCTTTCC-3' and *U3* 5'-CCAAGTGCACCCGCGTTG-3' and 5'-TCTATCCGTTTCTACCGAGCGATCA-3'.

In vitro transcription and gel mobility-shift assay

roX2 RNA derivatives were labeled with ^{32}Pα-UTP by *in vitro* transcription using hybridized oligonucleotide templates. One femptomole of RNA was incubated with increasing amounts of recombinant UNR and processed as previously described (Valcárcel et al. 1993).

Protein expression and purification

His-tagged, full-length UNR was purified according to the pET system user's manual (Novagen). The protein was dialyzed against buffer D (20 mM HEPES pH 8.0, 20% glycerol, 1 mM DTT, 0.01% NP40, 0.2 mM EDTA). For UNR footprinting experiments full-length His-FLAG-UNR was further purified by FLAG-affinity chromatography as described (Fauth et al., 2010). MLE protein was expressed in SF21 cells using recombinant baculoviruses and purified by FLAG-affinity chromatography.

Enzymatic and chemical probing of RNA secondary structure and RNP complexes

roX2 RNA 2-D structure was probed according to (Maenner et al., 2010). For probing of *roX2*-UNR complex, RNA was first incubated with recombinant UNR at 15, 30 and 60 protein/RNA molar ratio.

Acknowledgements

C.M. was supported by a fellowship from La Caixa Foundation. This work was supported by grants BFU2009-08243 and Consolider CSD2009-00080 from the Spanish Ministry of Economy and Competitiveness.

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.

Akhtar, A., and Becker, P.B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. Mol Cell *5*, 367-375.

Bell, O., Conrad, T., Kind, J., Wirbelauer, C., Akhtar, A., and Schubeler, D. (2008). Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in Drosophila melanogaster. Mol Cell Biol *28*, 3401-3409.

Chang, K.A., and Kuroda, M.I. (1998). Modulation of MSL1 abundance in female Drosophila contributes to the sex specificity of dosage compensation. Genetics *150*, 699-709.

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

Dethoff, E.A., Chugh, J., Mustoe, A.M., and Al-Hashimi, H.M. Functional complexity and regulation through RNA dynamics. Nature *482*, 322-330.

Franke, A., and Baker, B.S. (1999). The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. Mol Cell *4*, 117-122.

Gelbart, M.E., and Kuroda, M.I. (2009). Drosophila dosage compensation: a complex voyage to the X chromosome. Development *136*, 1399-1410.

Graumann, P.L., and Marahiel, M.A. (1998). A superfamily of proteins that contain the cold-shock domain. Trends Biochem Sci *23*, 286-290.

Gu, W., Wei, X., Pannuti, A., and Lucchesi, J.C. (2000). Targeting the chromatin-remodeling MSL complex of Drosophila to its sites of action on the X chromosome requires both acetyl transferase and ATPase activities. EMBO J *19*, 5202-5211.

Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., and Nakagawa, S. The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. Dev Cell *19*, 469-476.

Hunger, K., Beckering, C.L., Wiegeshoff, F., Graumann, P.L., and Marahiel, M.A. (2006). Cold-induced putative DEAD box RNA helicases CshA and CshB are essential for cold adaptation and interact with cold shock protein B in Bacillus subtilis. J Bacteriol *188*, 240-248.

Jeon, Y., and Lee, J.T. YY1 tethers Xist RNA to the inactive X nucleation center. Cell *146*, 119-133.

Koziol, M.J., and Rinn, J.L. RNA traffic control of chromatin complexes. Curr Opin Genet Dev *20*, 142-148.

Larsson, J., and Meller, V.H. (2006). Dosage compensation, the origin and the afterlife of sex chromosomes. Chromosome Res *14*, 417-431.

Lee, C.G., Chang, K.A., Kuroda, M.I., and Hurwitz, J. (1997). The NTPase/helicase activities of Drosophila maleless, an essential factor in dosage compensation. EMBO J *16*, 2671-2681.

Lee, J.T. Epigenetic regulation by long noncoding RNAs. Science *338*, 1435-1439.

Li, F., Schiemann, A.H., and Scott, M.J. (2008). Incorporation of the noncoding roX RNAs alters the chromatin-binding specificity of the Drosophila MSL1/MSL2 complex. Mol Cell Biol *28*, 1252-1264.

Lucchesi, J.C. (2009). The structure-function link of compensated chromatin in Drosophila. Curr Opin Genet Dev *19*, 550-556.

Maenner, S., Müller, M., Fröhlich, J., Langer, D. and Becker, P. B. (2013). ATP-dependent roX RNA remodeling enables specific association of MSL proteins. Mol Cell (in press)

Meller, V.H., Gordadze, P.R., Park, Y., Chu, X., Stuckenholz, C., Kelley, R.L., and Kuroda, M.I. (2000). Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in Drosophila. Curr Biol *10*, 136-143.

Meller, V.H., and Rattner, B.P. (2002). The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. EMBO J *21*, 1084-1091.

Mihailovich, M., Militti, C., Gabaldon, T., and Gebauer, F. Eukaryotic cold shock domain proteins: highly versatile regulators of gene expression. Bioessays *32*, 109-118.

Mitchell, S.A., Spriggs, K.A., Coldwell, M.J., Jackson, R.J., and Willis, A.E. (2003). The Apaf-1 internal ribosome entry segment attains the correct structural conformation for function via interactions with PTB and unr. Mol Cell *11*, 757-771.

Nagano, T., and Fraser, P. No-nonsense functions for long noncoding RNAs. Cell *145*, 178-181.

Park, S.W., Kang, Y., Sypula, J.G., Choi, J., Oh, H., and Park, Y. (2007). An evolutionarily conserved domain of roX2 RNA is sufficient for induction of H4-Lys16 acetylation on the Drosophila X chromosome. Genetics *177*, 1429-1437.

Park, S.W., Kuroda, M.I., and Park, Y. (2008). Regulation of histone H4 Lys16 acetylation by predicted alternative secondary structures in roX noncoding RNAs. Mol Cell Biol *28*, 4952-4962.

Park, Y., Mengus, G., Bai, X., Kageyama, Y., Meller, V.H., Becker, P.B., and Kuroda, M.I. (2003). Sequence-specific targeting of Drosophila roX genes by the MSL dosage compensation complex. Mol Cell *11*, 977-986.

Park, Y., Oh, H., Meller, V.H., and Kuroda, M.I. (2005). Variable splicing of non-coding roX2 RNAs influences targeting of MSL dosage compensation complexes in Drosophila. RNA Biol *2*, 157-164.

Patalano, S., Mihailovich, M., Belacortu, Y., Paricio, N., and Gebauer, F. (2009). Dual sex-specific functions of Drosophila Upstream of N-ras in the control of X chromosome dosage compensation. Development *136*, 689-698.

Rajkowitsch, L., Chen, D., Stampfl, S., Semrad, K., Waldsich, C., Mayer, O., Jantsch, M.F., Konrat, R., Blasi, U., and Schroeder, R. (2007). RNA chaperones, RNA annealers and RNA helicases. RNA Biol *4*, 118-130.

Scott, M.J., Pan, L.L., Cleland, S.B., Knox, A.L., and Heinrich, J. (2000). MSL1 plays a central role in assembly of the MSL complex, essential for dosage compensation in Drosophila. EMBO J *19*, 144-155.

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*, 31483-31486.

Smith, E.R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R.G., Allis, C.D., and Lucchesi, J.C. (2000). The drosophila MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. Mol Cell Biol *20*, 312-318.

Straub, T., and Becker, P.B. Transcription modulation chromosome-wide: universal features and principles of dosage compensation in worms and flies. Curr Opin Genet Dev *21*, 147-153.

Straub, T., Neumann, M.F., Prestel, M., Kremmer, E., Kaether, C., Haass, C., and Becker, P.B. (2005). Stable chromosomal association of MSL2 defines a dosage-compensated nuclear compartment. Chromosoma *114*, 352-364.

Straub, T., Zabel, A., Gilfillan, G.D., Feller, C., and Becker, P.B. Different chromatin interfaces of the Drosophila dosage compensation complex revealed by high-shear ChIP-seq. Genome Res *23*, 473-485.

Sural, T.H., Peng, S., Li, B., Workman, J.L., Park, P.J., and Kuroda, M.I. (2008). The MSL3 chromodomain directs a key targeting step for dosage compensation of the Drosophila melanogaster X chromosome. Nat Struct Mol Biol *15*, 1318-1325.

Triqueneaux, G., Velten, M., Franzon, P., Dautry, F., and Jacquemin-Sablon, H. (1999). RNA binding specificity of Unr, a protein with five cold shock domains. Nucleic Acids Res *27*, 1926-1934.

Wilusz, J.E., Sunwoo, H., and Spector, D.L. (2009). Long noncoding RNAs: functional surprises from the RNA world. Genes Dev *23*, 1494-1504.

Figure legends

Figure 1. Analysis of *roX2* levels, processing and nucleo-cytoplasmic distribution in UNR-depleted SL2 cells.

(A) Schematic representation of *roX2* gene structure. The *roX2* gene contains three exons (E1-E3) that are alternatively spliced. Alternative promoter and poly(A) site usage are indicated by horizontal and vertical arrows, respectively. The location of the known functional elements is indicated (conserved stem loop; thick vertical lines, *roX2*-boxes; hatched box, DHS). The major transcript isoforms (A-C) are depicted. Additional minor isoforms are produced by alternative splicing of exon 2. The asterisk denotes the most abundant isoform. The position of oligonucleotides used for qPCR is indicated. **(B)** Depletion of UNR does not alter the levels of *roX2*. Control (GFP) and UNR RNAi were performed, and the efficiency of UNR depletion was assessed by Western blot. Total *roX2* RNA levels were measured by RT-qPCR. Error bars represent the standard deviation of three independent experiments. **(C)** Detection of *roX2* isoforms by Northern blot. *S18* rRNA levels were monitored as loading control. **(D)** UNR depletion does not affect the intracellular localization of *roX2*. RNA samples (T, total; N, nuclear; C, cytoplasmic) were separated in a denaturing gel and were visualized by Northern blot. *U3* snoRNA and *S18* rRNA were used to monitor correct nucleo-cytoplasmic separation.

Figure 2. Analysis of UNR binding to *roX2*

(A) Schematic representation of the *roX2* fragments used to map the UNR binding site. The positions of the first and last nucleotides of each fragment are indicated. **(B)** Gel

mobility shift analysis of UNR binding to the *roX2* derivatives depicted in (A). The concentrations of UNR used are indicated at the bottom. **(C)** Sequence of fragment E. The repeats and the adenine-rich stretch are underlined.

Figure 3. Chemical and enzymatic footprinting of UNR on *roX2*

(A) Full length *roX2* was subjected to limited digestion with RNase T2, or modification with DMS in the absence or presence of increasing concentrations of UNR, followed by primer extension analysis. The resulting cDNAs were separated in a denaturing acrylamide/urea gel. A sequencing reaction with the same oligonucleotide was used as marker (UGCA). Nucleotides are numbered relative to nucleotide +1 of *roX2*. The position of the high-affinity UNR binding site and the functional stem-loop are indicated. **(B)** Schematic summary of the experimental footprinting data. The free energy value of the structure containing the high-affinity UNR binding site is indicated, as well as the loops within this structure. UNR binding sites on the right panel are indicated by continuous lines, and the remodeled region at the base of the stem loop is highlighted by a dashed line.

Figure 4. UNR promotes the association of *roX2* with MLE

(A) UNR interacts with MLE. Nuclear extracts were used to immunoprecipitate UNR, and presence of MLE and MSL3 in the pellet was assessed by Western blot. Non-specific IgGs were used as negative control. **(B)** UNR promotes the binding of MLE to *roX2 in vitro*. Left panel, schematic representation of *roX2* and deletion derivatives. RNAs were tagged at the 3' end with three copies of the MS2 coat protein binding site. The functionally relevant *roX2* stem-loop structure is depicted. Right panel, RNA pull-down experiment. RNAs were immobilized on amylose beads bound to MBP-coat protein, and incubated with recombinant MLE and UNR at the indicated concentrations in the presence of ATP. After extensive washes, RNAs were eluted with maltose and analyzed by PAGE. Associated proteins were analyzed by Western blot. The percentage of bound MLE and UNR was quantified relative to the amount of eluted RNA. Values were normalized to *roX2* binding in the presence of both MLE and UNR. **(C)** UNR promotes the binding of MLE to *roX2 in vivo*. UNR was depleted from SL2 cells (RNAi +); depletion of GFP was used as control (RNAi -). The total levels of *roX2* after depletion were monitored by RT-qPCR (middle panel). MLE was immunoprecipitated from control and UNR-depleted cells, and the efficiency of immunoprecipitation assessed by Western blot (left panel). A control IP with empty

beads was carried in parallel. The amount of *roX2* associated to MLE was determined by RT-qPCR (right panel). Values were normalized to MLE binding in control RNAi cells. Error bars represent the standard deviation of three independent biological replicates, each performed with technical duplicates (n= 6 experiments).

Figures

Figure 1

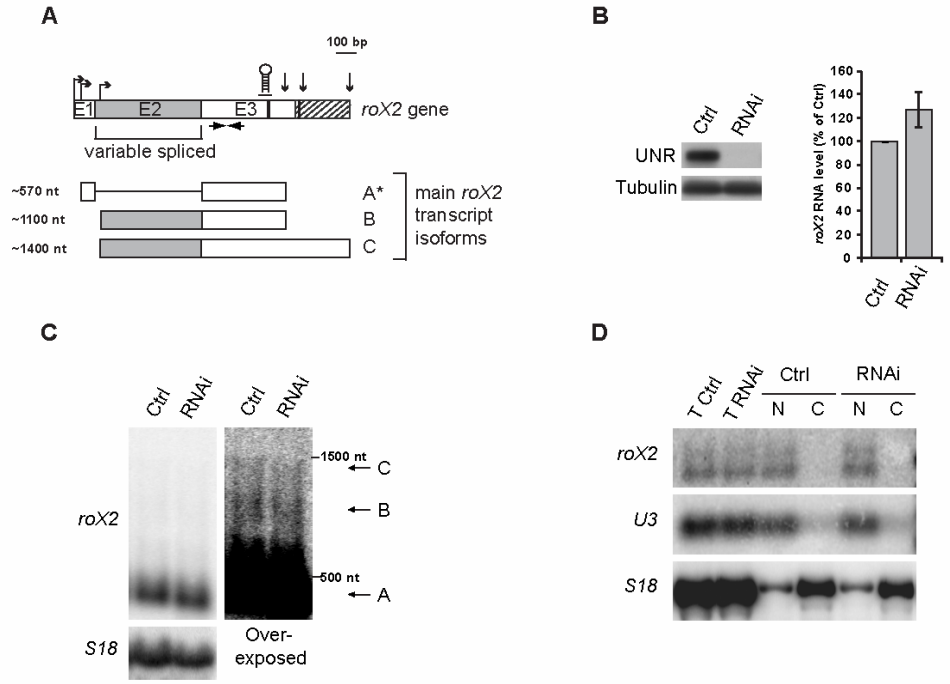
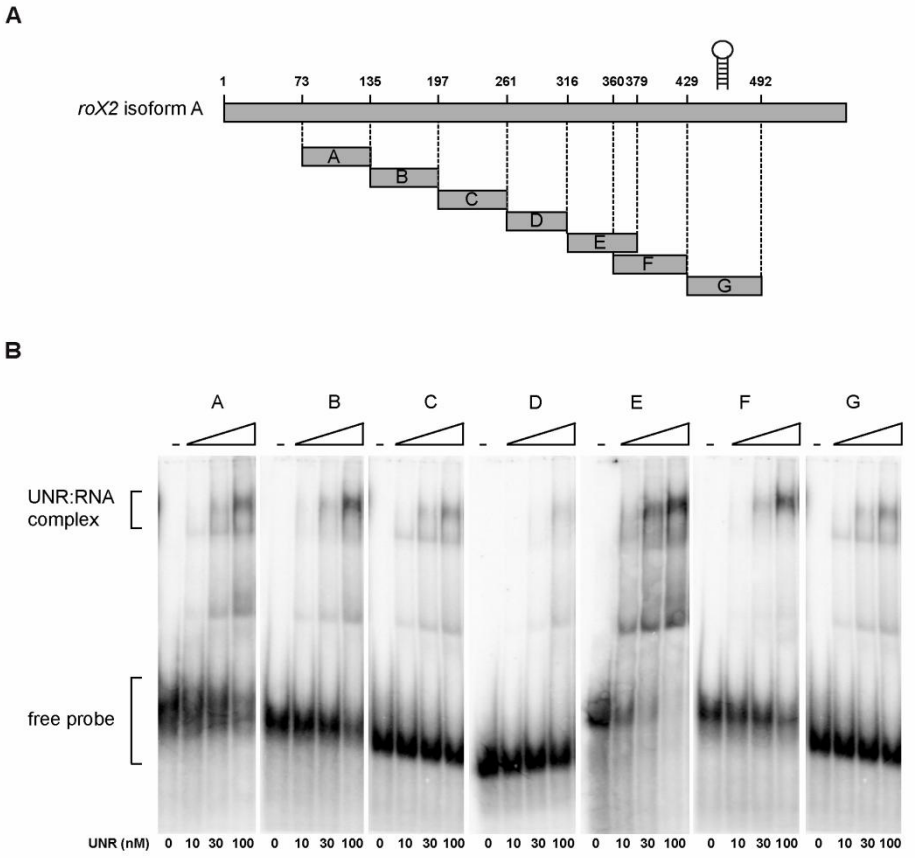


Figure 2

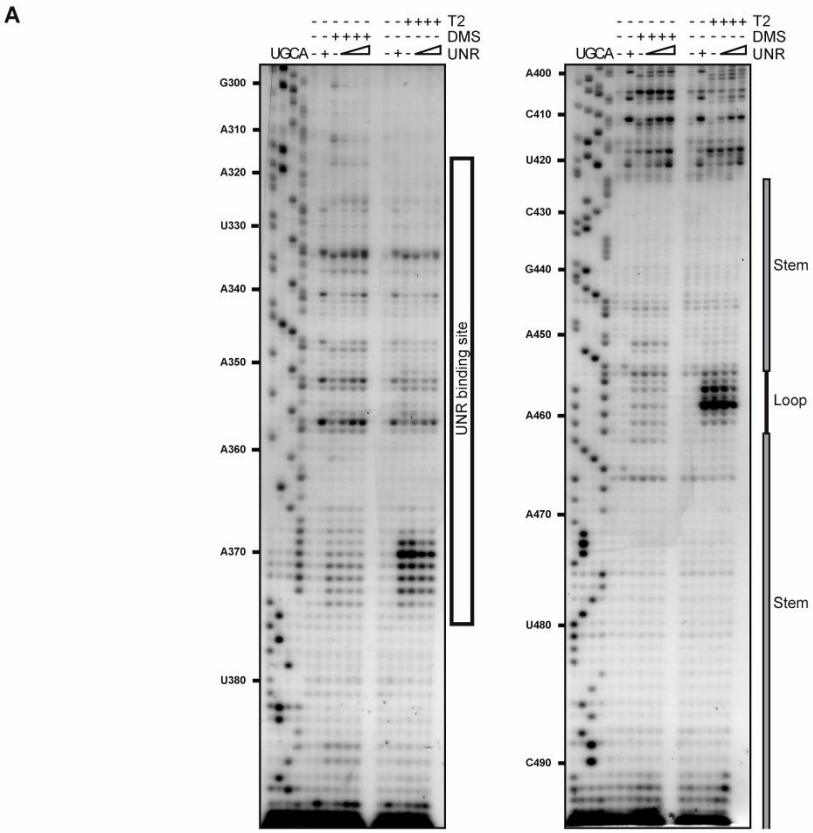


C

Fragment E

5'- AAUGAUAUACAAUAUACAAUAUACAAUAUGCAAUACAAUACAAUACAAGACAAAAAAAUGUGUC-3'

Figure 3


A
T2
DMS
UNR
UGCA
G300
A310
A320
U330
A340
A350
A360
A370
U380
UNR binding site
A400
C410
U420
C430
G440
A450
A460
A470
U480
C490
Stem
Loop
Stem

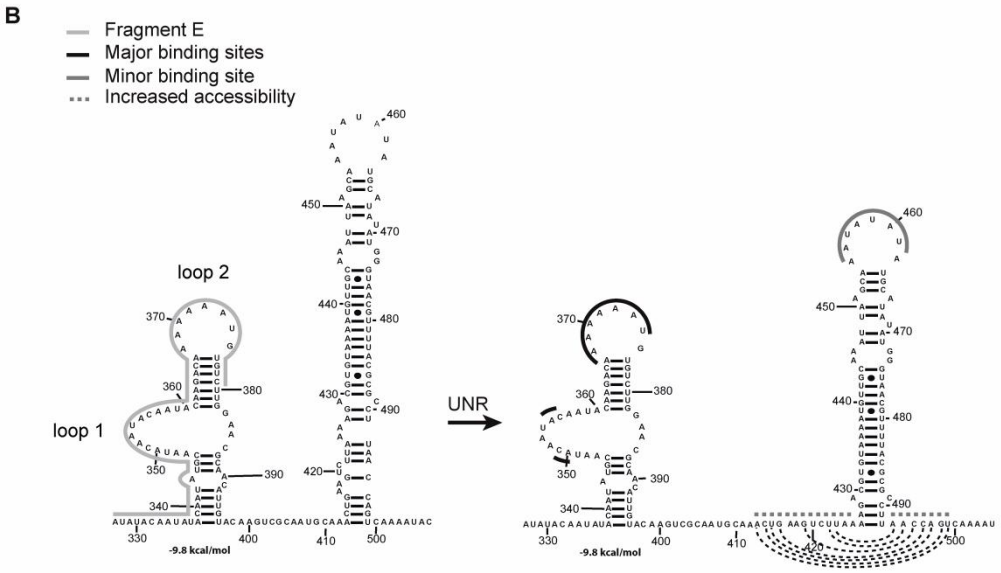
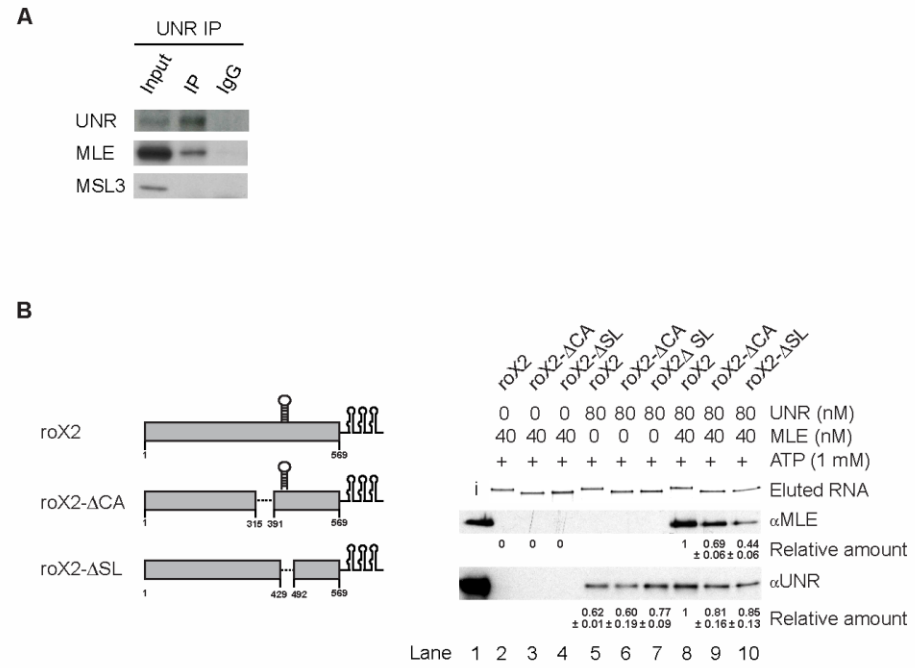

B
Fragment E
Major binding sites
Minor binding site
Increased accessibility
loop 2
loop 1
UNR
-9.8 kcal/mol
-9.8 kcal/mol

Figure 4


A
UNR IP
Input
IP
IgG
UNR
MLE
MSL3
B
roX2
roX2-ΔCA
roX2-ΔSL
1
569
315
391
429
492
roX2
roX2-ΔCA
roX2-ΔSL
roX2
roX2-ΔCA
roX2Δ SL
roX2
roX2-ΔCA
roX2-ΔSL
0 0 0 80 80 80 80 80 80 UNR (nM)
40 40 40 0 0 0 40 40 40 MLE (nM)
+ + + + + + + + + ATP (1 mM)
i
Eluted RNA
αMLE
0 0 0
1 0.69 ± 0.06 0.44 ± 0.06
Relative amount
αUNR
0.62 ± 0.01 0.60 ± 0.19 0.77 ± 0.09 1 0.81 ± 0.16 0.85 ± 0.13
Relative amount
Lane 1 2 3 4 5 6 7 8 9 10
C
Input
IP
Mock
UNR RNAi - + - + - +
MLE
Input
MLE IP
roX2 RNA level (% of Ctrl)
160
140
120
100
80
60
40
20
0
Ctrl
RNAi
IP
Mock
IP
Mock
Ctrl
RNAi

PART II: REGULATION OF DOSAGE COMPENSATION BY UNR IN FEMALES

UNR is a SXL co-factor required for translational repression of *msl2* mRNA in female flies (Abaza et al., 2006; Duncan et al., 2006). SXL provides a female-specific function to UNR, because it recruits UNR to the 3' UTR of *msl2*. In males, there are considerable amounts of UNR protein and *msl2* mRNA, but they do not interact because SXL is absent. Thus, we were intrigued by how SXL promoted UNR:*msl2* recognition. To answer this question, we took advantage of available information on the minimal domains of SXL and UNR required for complex formation, and identified the minimal residues on *msl2* mRNA. This information was then used to determine the three-dimensional structure of the ternary complex in collaboration with the laboratory of Michael Sattler (Technische Universität München, Germany).

A fragment of SXL containing the two RBDs plus the following 7aa (aa 122-301, fragment named dRBD4) is fully functional for *msl2* binding and for translational repression (Grskovic et al., 2003) (Figure 14). Interestingly, substituting the first RBD with that from the conserved SXL homologue in *Musca domestica* leads to loss of repression, suggesting that a critical non-conserved feature of SXL RBD1 is required for inhibition. Regarding UNR, the amino-terminal third (Q-CSD12 region) exerts the translational repression function (Abaza et al., 2008). CSD1 within this region is sufficient for SXL and *msl2* binding, although not for translational repression (Figure 14). The minimal sequences in the 3'UTR of *msl2* necessary for translational repression have been mapped to a 46 nt region containing two SXL binding sites (sites E and F) (Figure 15). This region also binds UNR, although the specific nucleotides necessary for UNR binding have not been mapped. We, thus, set to identify the UNR binding site on *msl2* mRNA.

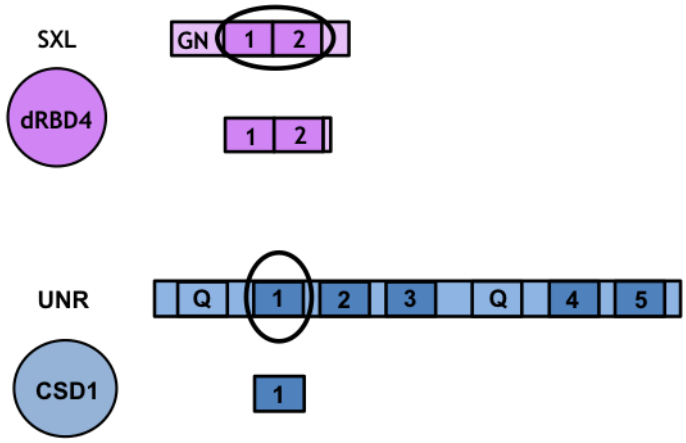


Figure 14. Schematic representation of the minimal domains of SXL and UNR required for complex formation. dRBD4 contains the two SXL RBDs and the first 7aa from the C-term region. CSD1 is the first cold shock domain from UNR.

UNR binds to sequences downstream of the SXL binding sites: defining the stoichiometry of the SXL:UNR:RNA ternary complex

Previous data indicated that the sequence of the EF fragment surrounding the SXL binding sites was necessary for UNR binding, since an RNA in which all nucleotides had been mutated except the SXL binding sites did not bind to UNR (Grskovic et al., 2003; Abaza et al., 2006). In order to narrow-down the sequence requirements for UNR binding, we used gel mobility shift assays (GEMSA) to test the formation of SXL:UNR complexes on EF substitution mutants. Six EF derivatives (Mut1-6) were obtained by substituting sequences of 8 nucleotides with CU repeats, which are known to interfere with SXL and UNR binding (Abaza et al., 2008) (Figure 15).

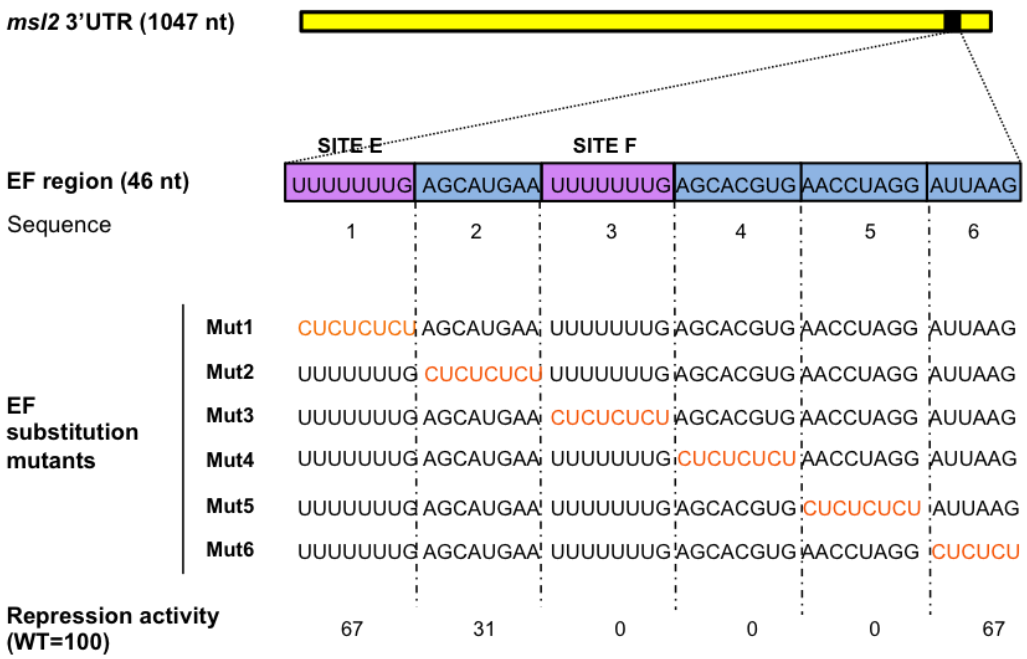


Figure 15. Schematic representation of the EF region of *msl2* 3'UTR and the EF substitution mutants. Nucleotides from the EF region important for SXL and UNR binding are represented in purple and blue, respectively. The sequences of the EF substitution mutants (Mut1-Mut6) are indicated; mutated nucleotides are depicted in orange. Repression activity of each mutant is reported relative to the WT sequence (100).

Radiolabeled WT and Mut1-6 RNAs were incubated with a constant amount of GST-tagged dRBD4 (10 nM) and increasing amounts of full-length His-tagged UNR. dRBD4:RNA and dRBD4:UNR:RNA complex formation were visualized by monitoring the appearance of specific shifts in GEMSA (Figure 16). When dRBD4 was incubated with WT EF RNA (Figure 16, Panels A, B, C lane 2) two complexes could be detected, being the faster-migrating one more prominent (asterisks). Since the WT probe contains two binding sites for SXL, these complexes may correspond to two dRBD4:RNA assemblies with different dRBD4 stoichiometries (one and two molecules of dRBD4, respectively). Indeed, mutation of one or the other SXL binding site results in the appearance of only one complex (Figure 16A, compare lanes 2, 9 and 16). However, the appearance of two complexes is influenced by the length of the probe and is not always detectable (data not shown).

Addition of UNR to a WT probe in presence of dRBD4 resulted in a further shift corresponding to the formation of a dRBD4:UNR:RNA complex (Figure 16A, lanes 3-7). Curiously, also in this case two bands could be visualized (circles), potentially

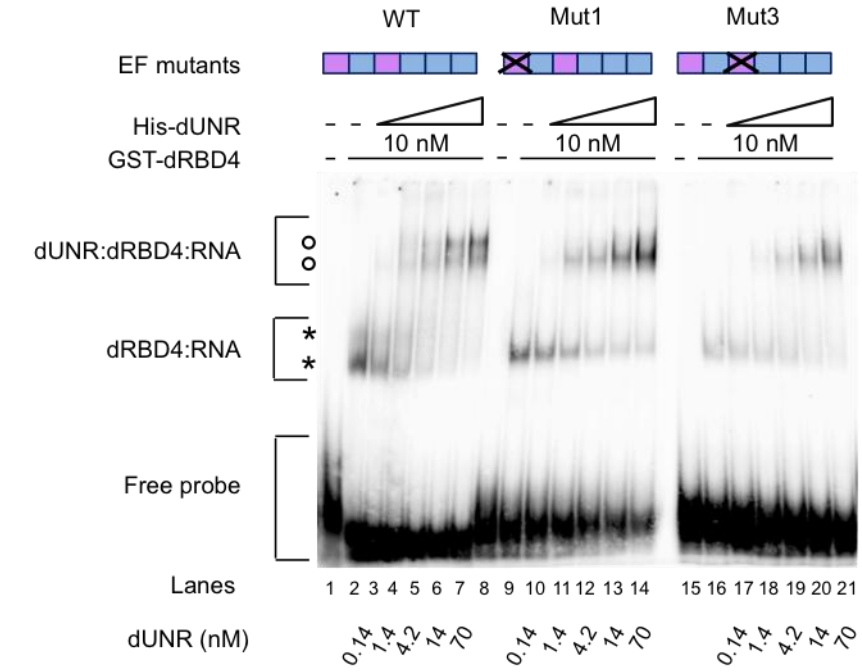
corresponding to two dRBD4:UNR:RNA complexes with different dRBD4 and UNR stoichiometries. Mutation of either SXL binding site resulted in the loss of the slower migrating dRBD4:UNR:RNA complex, while the faster migrating complex formed at an efficiency comparable to wild type (Figure 16A). One possibility to explain these data is a dRBD4:UNR:RNA complex configuration with a stoichiometry of 2:2:1 on the wild type RNA (see Figure 17).

Mutation of the sequences surrounding SXL binding site F (Mut 2 and 4) greatly reduced dRBD4:UNR:RNA complex formation (Figure 16B). Also in this case only one band is visible for the ternary complex. Since for these RNAs the dRBD4:RNA complex appears as a doublet, it is very likely that the two bands for the dRBD4:UNR:RNA complex signal correspond to two complexes that differ in the UNR, rather than the dRBD4, stoichiometry. According to this hypothesis, sequences 2 and 4 could represent two distinct binding sites for UNR. Binding of UNR to each of the sites would be supported by the presence of a SXL molecule binding to the adjacent U-rich stretch. Notably, mutants corresponding to sequence 2 and 4 show reduced or disrupted *msl2* translational repression activity (Figure 15) in accordance with their involvement in UNR binding.

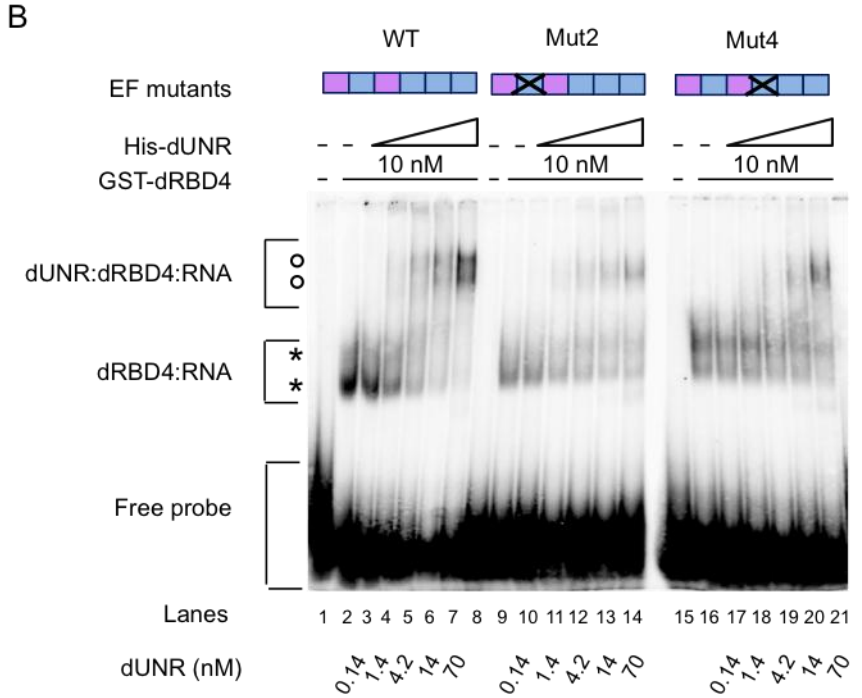
dRBD4:UNR:RNA complex formation does not seem to be affected by the mutation of sequences 5 and 6 (Figure 16C). On the contrary, binding of SXL was enhanced by mutation of these sequences. Interestingly, mutation of sequence 5 abolishes *msl2* translational repression, indicating that this sequence could be recognized by a yet unknown factor involved in 3'UTR-mediated translation inhibition.

Altogether, GEMSA analysis have identified two binding sites for UNR, each downstream of a SXL binding site. The sequence of these sites is highly similar, consistent with binding to the same RNA-binding protein. Furthermore, the data is consistent with a stoichiometry of 2:2:1 for the SXL:UNR:RNA ternary complex. Indeed, further analysis by static light scattering (SLS) of a dRBD4:CSD1:34 mer RNA complex performed by Janosch Henning in the lab of our collaborator have confirmed such 2:2:1 stoichiometry (Figure 17). Thus, the EF RNA supports the formation of two tandem SXL:UNR complexes.

A


WT
Mut1
Mut3
EF mutants
His-dUNR
GST-dRBD4
10 nM
10 nM
10 nM
dUNR:dRBD4:RNA
dRBD4:RNA
Free probe
Lanes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
dUNR (nM)
0.14 1.4 4.2 14 70
0.14 1.4 4.2 14 70
0.14 1.4 4.2 14 70

B


WT
Mut2
Mut4
EF mutants
His-dUNR
GST-dRBD4
10 nM
10 nM
10 nM
dUNR:dRBD4:RNA
dRBD4:RNA
Free probe
Lanes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
dUNR (nM)
0.14 1.4 4.2 14 70
0.14 1.4 4.2 14 70
0.14 1.4 4.2 14 70

C

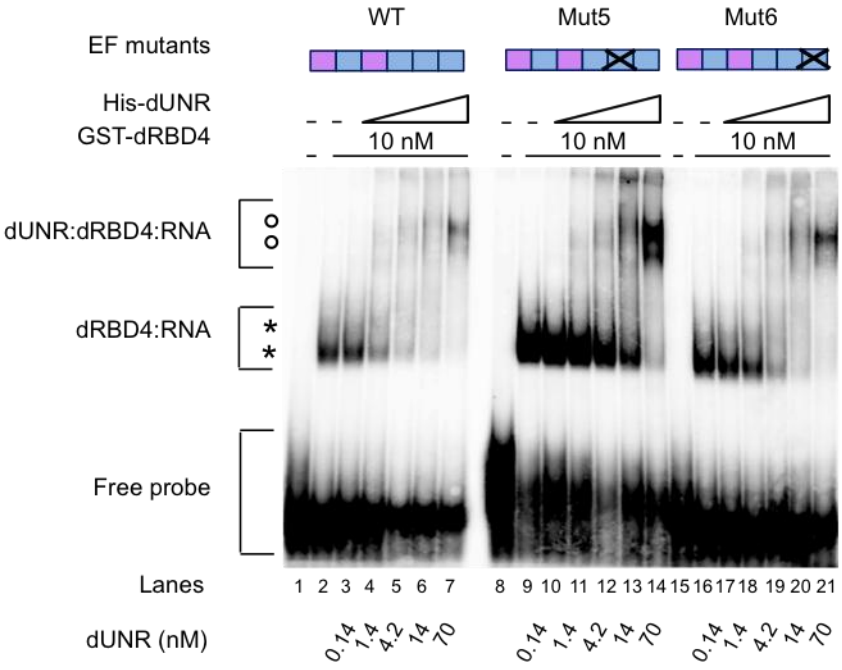


Figure 16. GEMSA analysis of dRBD4 and UNR binding to EF substitution mutants. GEMSAs were conducted by adding a constant amount of GST-dRBD4 (10 nM) and increasing amounts of His-UNR (from 0.14 to 70 nM) to radiolabeled WT and mutant EF probes. Migration of the free probe, dRBD4:RNA and dUNR:dRBD4:RNA complexes is indicated. Doublets for dRBD4:RNA and dUNR:dRBD4:RNA complexes are indicated by asterisks and circles, respectively.

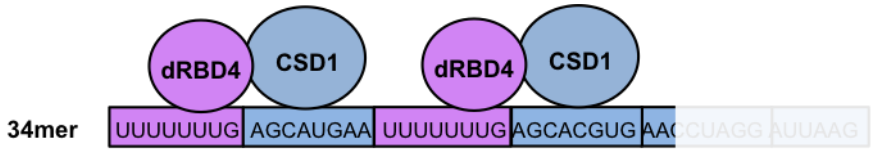


Figure 17. Schematic representation of the SXL:UNR complex on the WT RNA. dRBD4 and CSD1 assemble with a 2:2:1 stoichiometry on the wild type RNA.

SXL and UNR bind EF RNA in a cooperative manner

Having found the minimal elements involved in the SXL:UNR:*msl2* complex formation, we decided to proceed with the complex characterization by using a minimal system consisting of dRBD4, CSD1 and an EF RNA derivative containing one single binding site for SXL and UNR (sequences 3 and 4; 18mer; Figure 18)

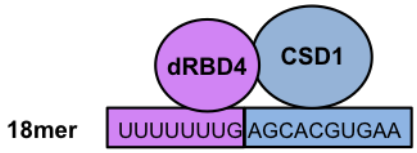


Figure 18. Schematic representation of the SXL:UNR complex on the 18mer RNA

Isothermal titration calorimetric (ITC) experiments conducted by our collaborators showed that dRBD4 and CSD1 cannot interact in the absence of RNA, confirming previous results obtained in our laboratory (Abaza et al., 2006). Moreover, dRBD4 and CSD1 bind the 18mer RNA with a relative low affinity, corresponding to 200 nM and 500 nM Kd respectively. Interestingly, however, when dRBD4 and CSD1 are incubated together in the presence of the 18mer, the affinity of the two proteins for the RNA is greatly enhanced, and binding occurs with a Kd as low as 15 nM, leading to the formation of a stable dRBD4:CSD1:18mer complex (Figure 19). These results indicate that binding to the RNA favors interactions between the proteins, resulting in cooperative binding.

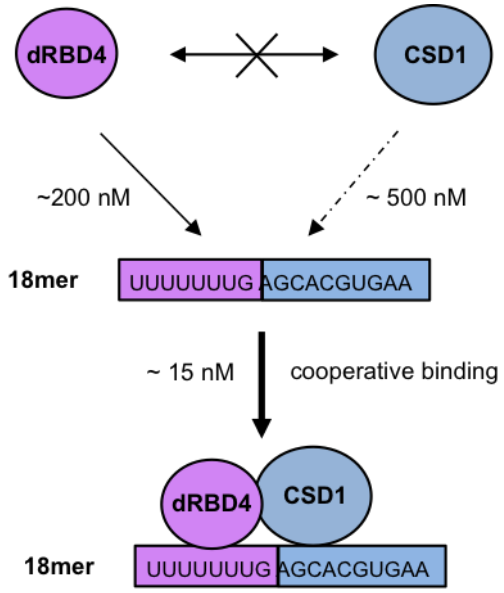


Figure 19. Cooperative binding of dRBD4 and CSD1 to the 18mer RNA. Free soluble dRBD4 and CSD1 proteins cannot interact with each other in the absence of RNA. dRBD4 and CSD1 alone have low affinity for the 18mer RNA (200 nM and 500 nM, respectively). When dRBD4, CSD1 and the 18mer are present together, dRBD4 and CSD1 bind to the RNA with high affinity (15 nM).

We confirmed these results by GEMSA using dRBD4, CSD1 and a 5' end-labeled 18mer RNA (Figure 20).

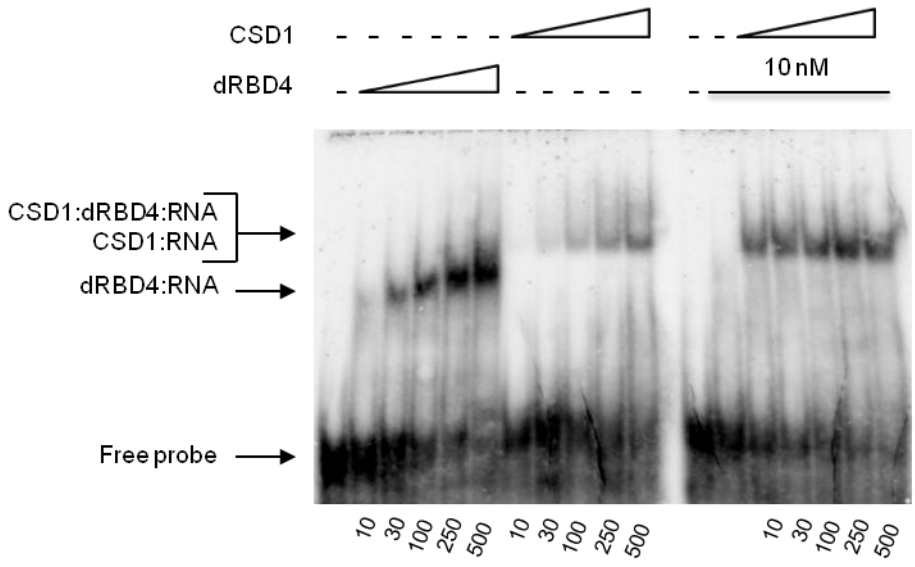


Figure 20. GEMSA analysis of dRBD4 and CSD1 binding to the 18mer RNA.
GEMSA analysis were conducted by adding the indicated amounts of dRBD4 or CSD1 to radiolabeled 18mer RNA probe. Migration of the free probe, dRBD4:RNA, CSD1:RNA and CSD1:dRBD4:RNA complexes is indicated.

Crystal structure of the dRBD4:CSD1:18mer complex

Our collaborators used X-ray crystallography to solve the structure of the dRBD4:CSD1:18 mer complex. The structure was consistent with NMR spectroscopy data. Figure 21 shows a representation of the crystal structure of the complex, obtained at 2.9 A resolution. Consistent with functional data, dRBD4 interacts with CSD1 through RBD1 (Grskovic et al., 2003). Interaction occurs by means of protein-protein contacts established with residues located in unstructured loops of the two proteins. The 5'-half of the 18mer RNA interacts with the two SXL RBDs via canonical interactions with the β-sheets of the two domains. The 3'half of the RNA is "trapped" with its first nucleotides in the dRBD4-CSD1 interaction interface and then allowed to associate with a second binding interface formed by CSD1 and the back helix of RBD1. I will highlight the novel features of this structure below.

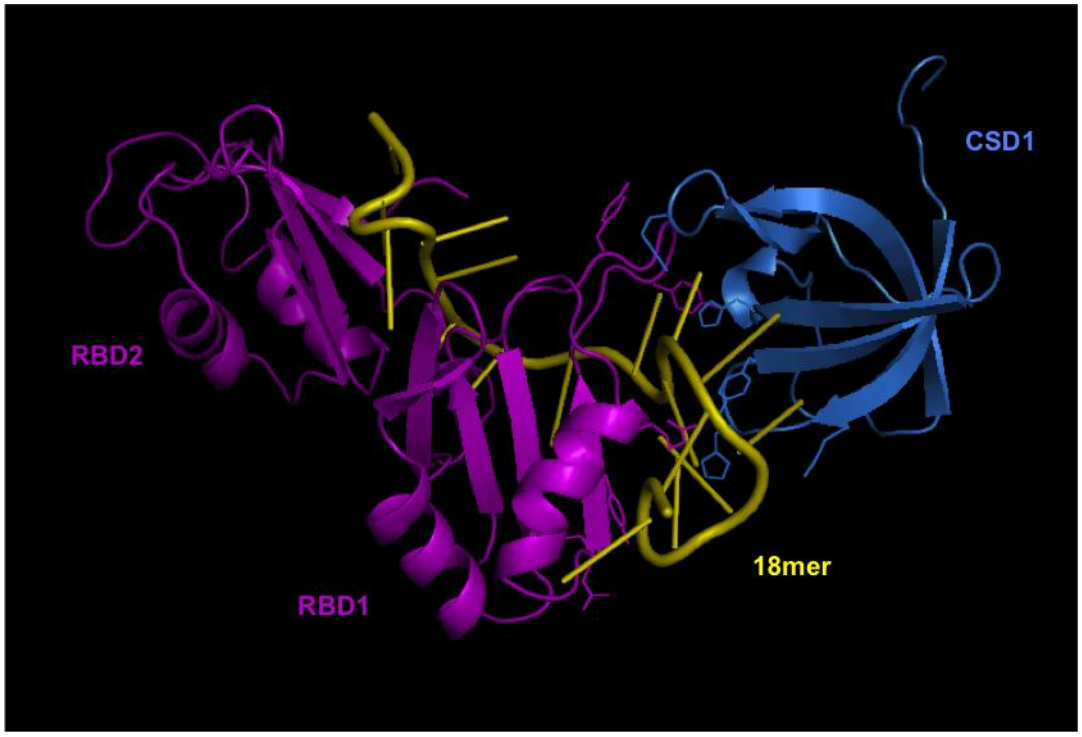


Figure 21. Crystal structure of the dRBD4:CSD1:18mer complex.
Purple: RBD1 and RBD2 domains of SXL. Blue: CSD1 of UNR. Yellow: 18mer RNA from *msl2*.

The dRBD4:CSD1:18mer interaction interface

Association of dRBD4 and CSD1 is mainly directed by the formation of a dRBD4:CSD1:18mer interaction interface where nucleotides and residues from the two proteins are “sandwiched” in an intricate zip-like structure (Figure 22).

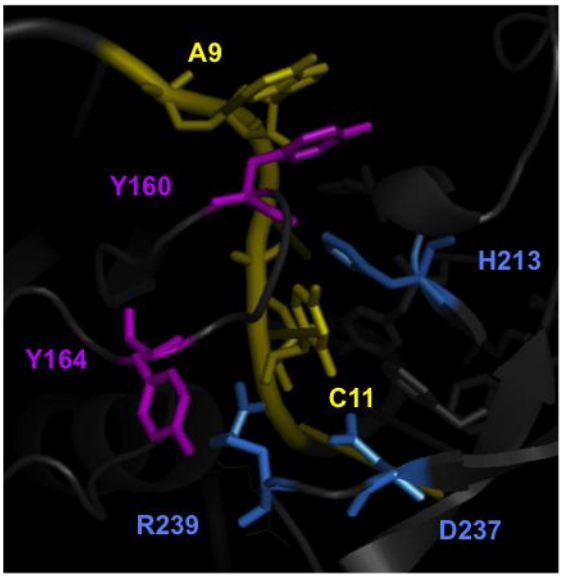


Figure 22. dRBD4:CSD1:18mer interaction interface. The residues from dRBD4 (purple), CSD1 (blue) and the RNA (yellow) are indicated. Residue numbers refer to the full length SXL or UNR proteins.

UNR R239, contacts SXL Y164 on one side and the RNA C11 on the other side. UNR H213 is sandwiched between the RNA C11 and SXL Y160. C11 and H213 are also contacted by UNR D237, which was previously shown to be important for complex formation (Abaza et al., 2008). Finally SXL Y160 makes contacts with UNR H213 and the RNA A9. We validated these interactions by mutating the corresponding residues on dRBD4 and CSD1, and testing complex formation by GEMSA. Mutation of UNR R239 or H213 to alanine lead to complete complex disruption (Figure 23A). Mutation of SXL Y160 and Y164 also decrease complex formation, although the effect appears less dramatic than those for the UNR counterpart (Figure 23B). The presence of A9 and C11 RNA nucleotides in the dRBD4:CSD1 interaction interface is of great relevance for the establishment of a strong protein-protein interaction, explaining why the two proteins do not interact in the absence of RNA.

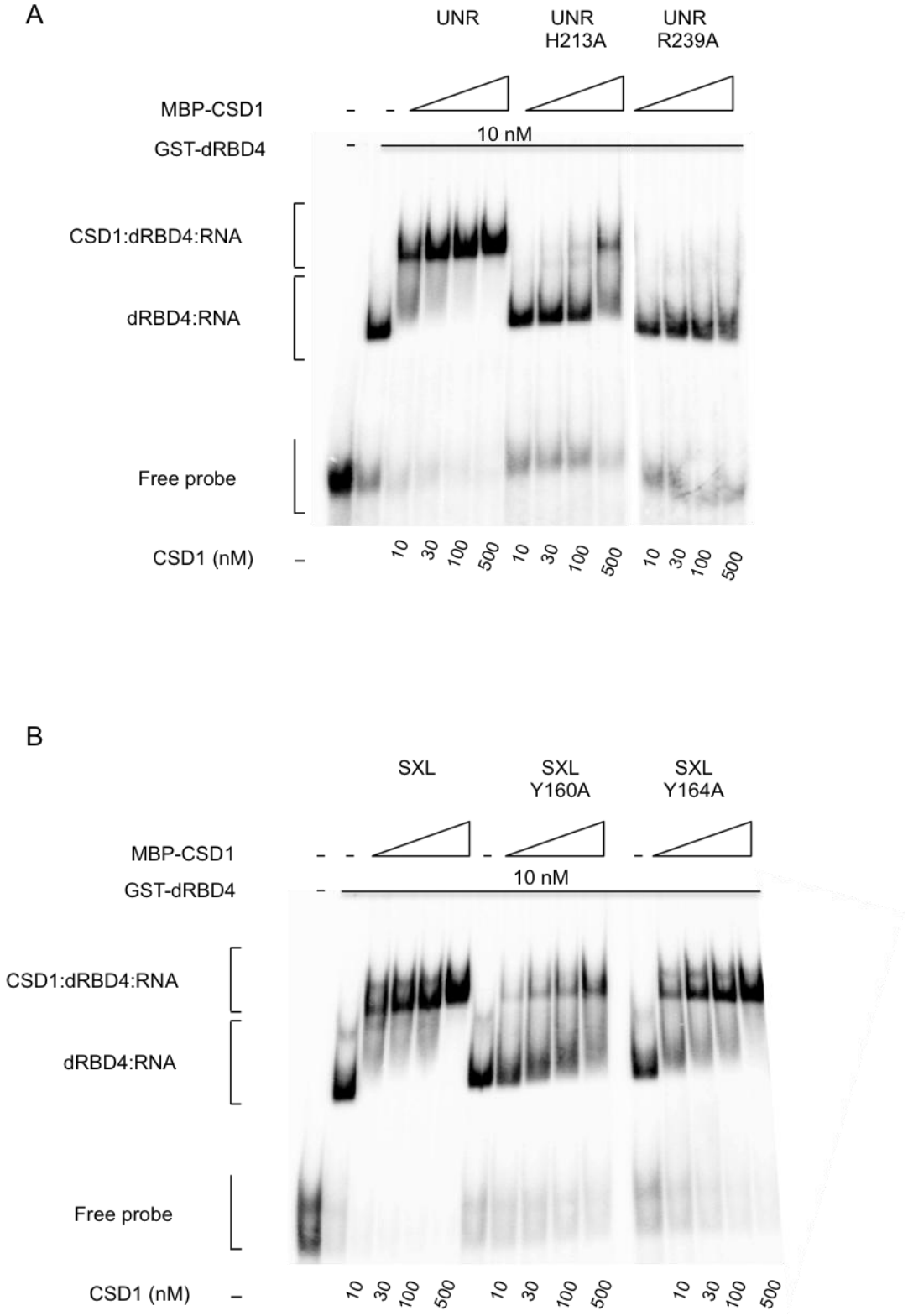


Figure 23. GEMSA analysis of WT and mutant dRBD4 and CSD1 proteins binding to the EF RNA.

The CSD1:18mer interaction interface

The GCAC sequence in the 3'half of the 18mer RNA is contacted by CSD1, through the engagement of residues involved in canonical and non-canonical RNA binding (Figure 24). The residues Y198, F200, F211 and H213 are part of the two CSD RNP motifs (RNP1 and RNP2) normally involved in RNA binding (Max et al., 2006). In particular Y198 from UNR has been already proved to play an essential role in SXL:UNR:msl2 complex formation (Abaza et al., 2008). Non-canonical RNA-interactions are mediated by K193 and H196 (Figure 24). Follow-up GEMSA analysis confirmed the role of K193, but showed little contribution of H196 (Figure 25).,

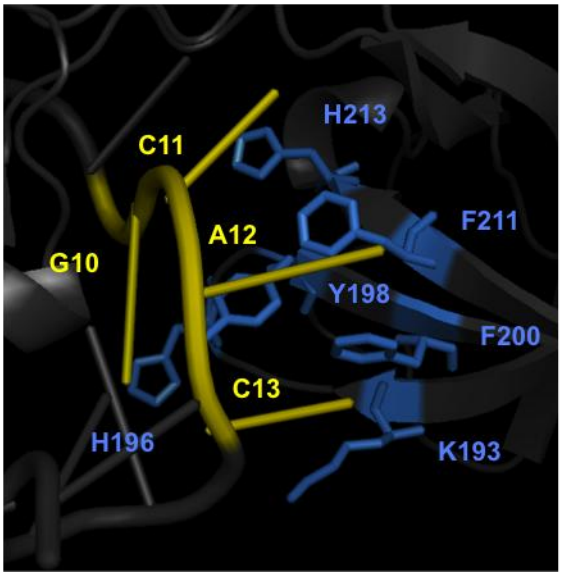


Figure 24. CSD1:18mer interaction interface. The involved residues from CSD1 (blue) and the RNA (yellow) are indicated. Residue numbers refer to the full length UNR protein.

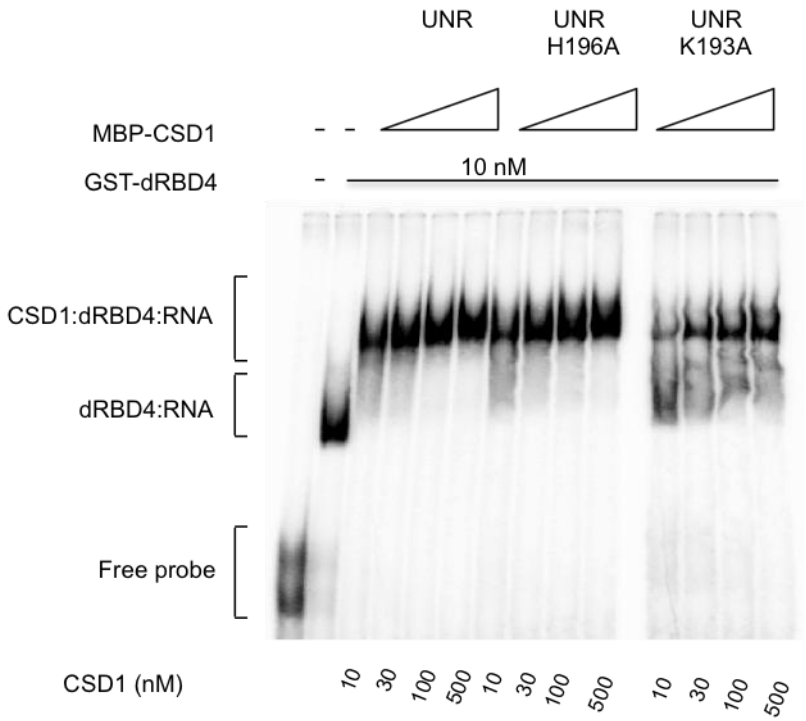


Figure 25. GEMSA analysis of WT and mutant UNR binding to EF RNA in presence of WT dRBD4.

The dRBD4:18mer interaction interface

dRBD4 interacts with the U-stretch and the following G residue through canonical contacts mediated by the RNP motifs located in the β-sheets of the two RBDs (Handa et al.,1996). Interestingly, new contacts are established with the RNA by residues located in one helix of RBD1 (Figure 26). In particular, SXL residues R139, Y142 and R146 are involved in contacts with the U15-G16-A17 RNA segment. Mutation of two of these residues, R139 and R146, lead to mild complex formation defects, which becomes more apparent when the two mutations are coupled (Figure 27). Further mutagenesis is required to confirm the role of Y142.

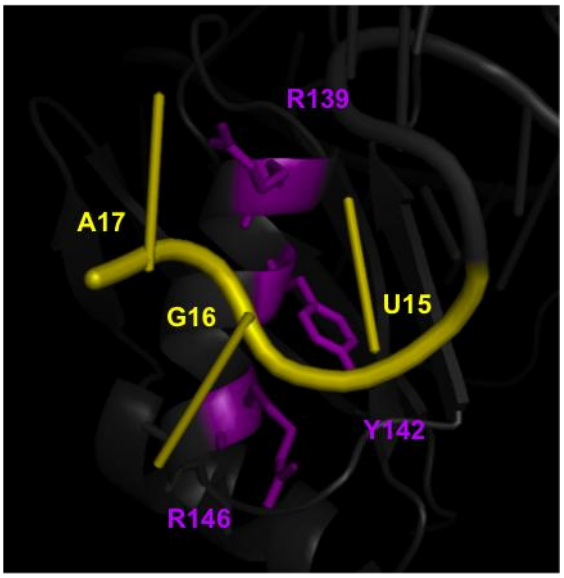


Figure 26. Non-canonical SXL:RNA contacts. The involved residues of dRBD4 (purple) and the RNA (yellow) are indicated. Residue numbers refer to the full length SXL protein.

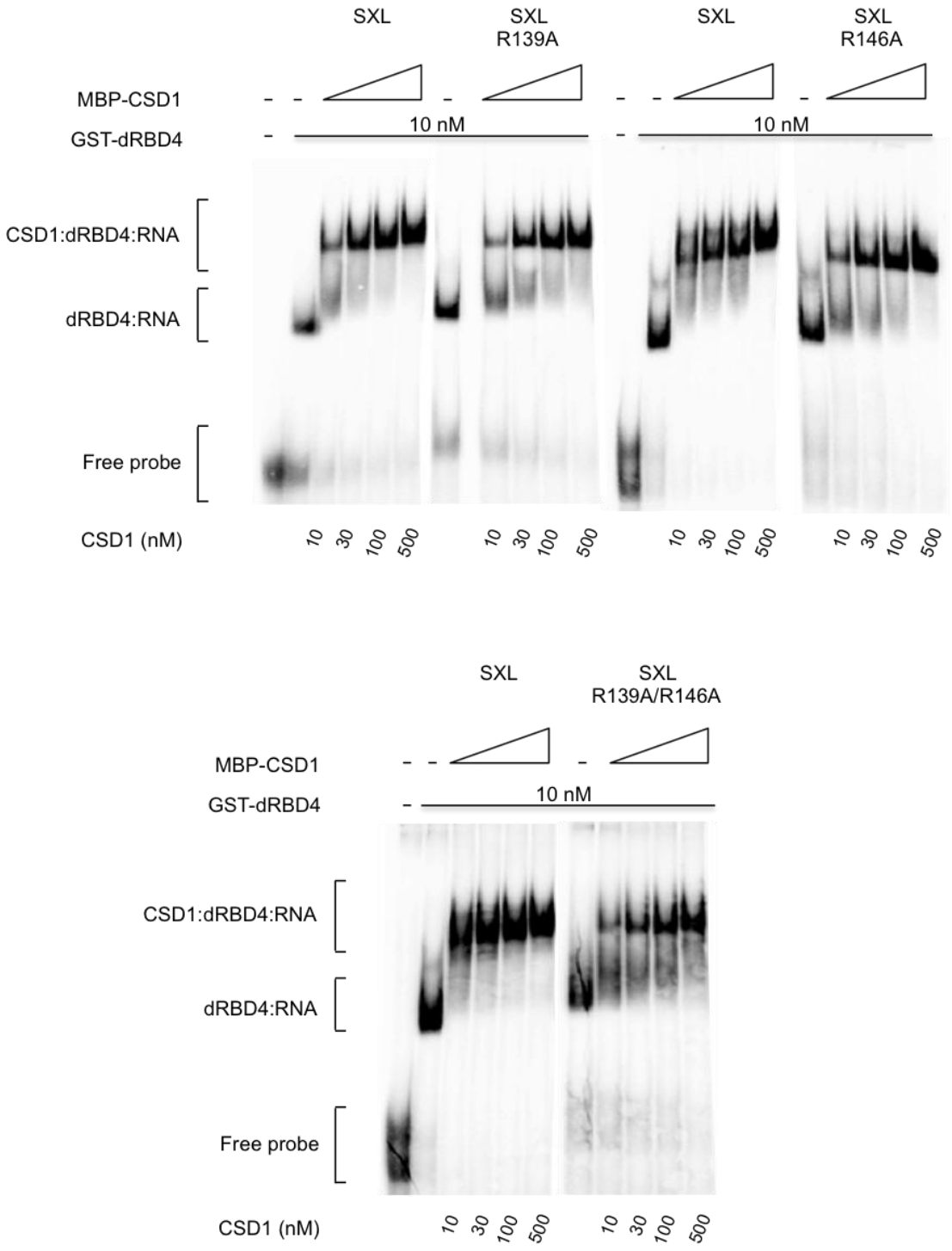


Figure 27. GEMSA analysis of WT and mutant dRBD4 binding to EF RNA in the presence of CSD1.

Comparison between *Drosophila* and *Musca* SXL

Binding and translation assays have revealed that the SXL homologue from *Musca domestica* (mSXL) does not repress *msl2* translation even though it binds to *msl2* 3'UTR with an affinity comparable to *Drosophila* SXL (dSXL) (Grskovic et al., 2003). Failure to repress translation is thought to be due to the inability of mSXL to interact with UNR (Grskovic et al, 2003; Abaza et al., 2006). Having the molecular architecture of the complex, we first checked whether residues involved in complex formation were different between the two SXL proteins. However, when the two proteins were compared, conservation was found for all residues directly involved in complex formation. Why, then, mSXL cannot form a complex with UNR and *msl2* in extracts? To gain insight into this conundrum, we decided to perform GEMSA with *Drosophila* and *Musca* SXL (dSXL and mSXL, respectively) in the presence or absence of CSD1 (Figure 28). As expected, the two proteins were able to interact with EF RNA with similar efficiencies. However, when CSD1 was added to the reaction, ternary complex formation seemed weaker in the case of mSXL. Complex formation did occur, but cooperativity seemed partially lost in the mSXL-directed complex, since addition of CSD1 did not enhance mSXL binding to the RNA to the same extent as with dSXL. Loss in cooperative binding could explain why UNR is not able to interact with mSXL in the competitive conditions of the embryo extract. Why mSXL:CSD1 complexes are less stable is unclear. One possibility is that non-conserved residues not directly involved in binding affect the overall flexibility or conformation of the protein. Alternatively, unknown factors could contribute to modulate the binding between SXL and UNR in embryo extracts. Experiments are underway to test these possibilities.

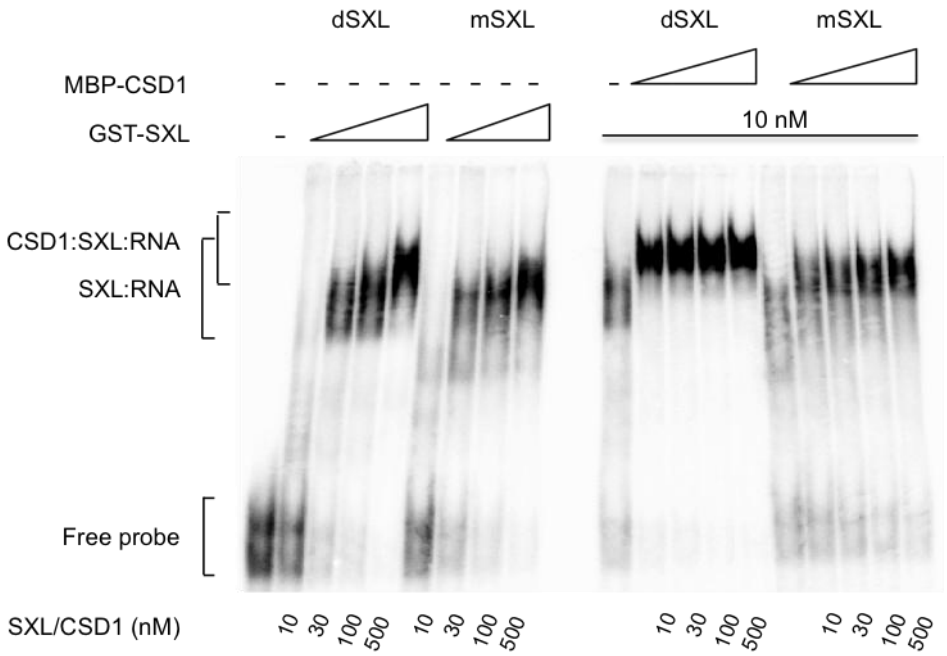


Figure 28. GEMSA analysis of dSXL and mSXL binding to the EF RNA in presence or absence of CSD1.

MATERIALS AND METHODS

This section refers to Results Part II and is intended to complement the experimental procedures of Results Part I.

Probe templates

Hybridized sense (S) and antisense (AS) oligos containing a T7 promoter were used as templates for the *in vitro* transcription of the EF substitution mutants used in GEMSA of Figure 16. For the remaining GEMSA experiments, WT EF was produced by *in vitro* transcription using pBSK-EF as template (Grskovic et al., 2003).

EF RNA		Oligo
WT	S	5'-TAATACGACTCACTATAGGGTTTTTTTGAGCATGAA TTTTTTTGAGCACGTGAACCTAGGATTAAG-3'
	AS	5'-CTTAATCCTAGGTTCACGTGCTCAAAAAAATTCATG CTCAAAAAAACCCTATAGTGAGTCGTATTA-3'
Mut1	S	5'-TAATACGACTCACTATAGGGCTCTCTCTAGCATGA ATTTTTTTGAGCACGTGAACCTAGGATTAAG-3'
	AS	5'-CTTAATCCTAGGTTCACGTGCTCAAAAAAATTCATG CTAGAGAGAGCCCTATAGTGAGTCGTATTA-3'
Mut2	S	5'-TAATACGACTCACTATAGGGTTTTTTTGCTCTCTCC TTTTTTTGAGCACGTGAACCTAGGATTAAG-3'
	AS	5'-CTTAATCCTAGGTTCACGTGCTCAAAAAAAGGA GAGAGCAAAAAAACCCTATAGTGAGTCGTATTA-3'
Mut3	S	5'-TAATACGACTCACTATAGGGTTTTTTTGAGCATGAA CTCTCTCTAGCACGTGAACCTAGGATTAAG-3'
	AS	5'-CTTAATCCTAGGTTCACGTGCTAGAGAGAGTTC ATGCTCAAAAAAACCCTATAGTGAGTCGTATTA-3'
Mut4	S	5'-TAATACGACTCACTATAGGGTTTTTTTGAGCATGAA TTTTTTTGCTCTCTCTAACCTAGGATTAAG-3'
	AS	5'-CTTAATCCTAGGTTAGAGAGAGCAAAAAAATTC ATGCTCAAAAAAACCCTATAGTGAGTCGTATTA-3'
Mut5	S	5'-TAATACGACTCACTATAGGGTTTTTTTGAGCATGAA TTTTTTTGAGCACGTGCTCTCTCTATTAAG-3'
	AS	5'-CTTAATAGAGAGAGCACGTGCTCAAAAAAATTC ATGCTCAAAAAAACCCTATAGTGAGTCGTATTA-3'

Mut6	S	5'-TAATACGACTCACTATAGGGTTTTTTTGAGCATGAA TTTTTTTGAGCACGTGAACCTAGGCTCTCT-3'
	AS	5'-AGAGAGCCTAGGTTCACGTGCTCAAAAAAATTC ATGCTCAAAAAAACCCTATAGTGAGTCGTATTA-3'

DNA constructs for recombinant protein production

Plasmids expressing mutant proteins were obtained by mutagenesis of the following WT plasmids using the Quick-site directed mutagenesis system (Agilent):

Recombinant protein	Plasmid	Reference
His-dUNR	pET15b-dUNR	Abaza et al., 2006
GST-dRBD4	pGEX-dRBD4	Grskovic et al., 2003
MBP-CSD1	pMALc-dCSD1	Abaza et al., 2008
GST-dSXL	pGEX-dSXL	Grskovic et al., 2003
GST-mSXL	pGEX-mSXL	Grskovic et al., 2003

Protein expression and purification

Recombinant proteins used in the GEMSA analysis were expressed and purified from E.Coli. SXL derivatives were expressed as N-terminal GST-tagged fusions and purified as described previously (Grskovic et al. 2003). His-tagged, full-length UNR was purified following the pET system user's manual (Novagen). MBP-CSD1 was purified following the pMALc system user's manual (New England Biolabs). All proteins were dialyzed against buffer D (20 mM HEPES at pH 8.0, 20% glycerol, 1 mM DTT, 0.01% NP-40, 0.2 mM EDTA). Protein preparations were monitored by Coomasie staining.

Probe synthesis

In vitro transcription of WT and mutant EF probes was performed as described (Gebauer et al., 1999). 18mer RNA oligos were 5'-end labeled using ^{32}PγATP and T4 Polynucleotide Kinase (Thermoscientific). Probe integrity was monitored in denaturing acrylamide gels. When appropriate, probes were gel-purified.

Gel mobility shift assays (GEMSA)

GEMSA was performed as described (Abaza et al., 2006). Briefly few femptomoles of ^{32}P-labeled RNAs were incubated for 30min at 4 ºC with the indicated recombinant proteins in buffer D supplemented with tRNA (150 ng/μl) and 100 mM KCl. Samples

were resolved in non-denaturing 4% acrylamide gels; gels were dried, exposed and revealed using a Typhoon Phosphorimager (Amersham).

DISCUSSION

UNR is an evolutionary conserved RNA binding factor belonging to the family of cold shock domain proteins, which are known to play diverse roles in a variety of processes (Mihailovich et al., 2010). In mammals, UNR has been shown to regulate mRNA stability and translation in processes such as proliferation and apoptosis. In *Drosophila,* UNR binds to hundreds of target mRNAs (Mihailovich et al., 2012), although to date a role for this protein has only been reported in the regulation of dosage compensation. Here we dissect the molecular mechanisms by which UNR regulates this process.

In male flies, UNR exerts a positive role in the regulation of dosage compensation by promoting the targeting of the dosage compensation complex to the X-chromosome (Patalano et al., 2009). We found that targeting is stimulated through a mechanism that involves the post-transcriptional regulation of *roX2*, one of the two lncRNAs of the complex. A role for RNA-binding proteins in the regulation of lncRNA function in dosage compensation has been so far reported only for the mammalian X-inactivation process. Indeed the two RNA binding proteins hnRNP U and YY1 have been proposed to have a role in tethering the lncRNA *Xist* to the inactive X-chromosome (Hasegawa et al., 2010; Jeon and Lee, 2011).

Here we report that UNR regulates *roX2* function by facilitating its association with MLE, the RNA helicase of the DCC. MLE is thought to be the first component of the complex to interact with *roX2,* and is known to be necessary for its ultimate incorporation into the DCC (Meller et al., 2000; Gu et al., 2000; Meller and Rattner, 2002). By regulating MLE binding to *roX2* UNR would favor the robust formation of DCCs, thereby promoting their efficient binding to the X-chromosome. We propose that the molecular mechanism by which UNR promotes MLE binding to *roX2* involves the interaction of UNR with a region in the non-coding RNA located upstream of a functional stem-loop. This region, which is particularly enriched in CAAUA repeats, folds into a secondary structure containing two loops, which are both contacted by UNR. CA-dinucleotides located in one of the loops appear necessary for the formation of a specific high affinity interaction. Additional contacts are also established with the second A-rich loop. UNR binding to single stranded regions within secondary structures is not a special feature of *roX2*. Mammalian UNR has been shown to bind such RNA configurations in IRESes, and other cold shock domain- containing proteins display

similar binding properties. For example, the crystal structure of Lin28 complexed with let7 pri-miRNA has revealed binding of the Lin28 cold shock domain to the terminal loop of the pri-miRNA (Nam et al., 2011).

Secondary and tertiary structures endow the RNA molecule of functional properties, by influencing their ability to interact with proteins or to exert catalytic activities. RNA molecules are known to fluctuate between different conformational states; this can cause an RNA to get trapped in an unproductive conformation that could need to be resolved. Moreover, RNA switches can be required to coordinate the activity of the RNA molecule (Dethoff et al., 2012). RNA chaperones and helicases possess the ability to orchestrate such conformational transitions, thus playing a central role in the regulation of RNA function. In mammals, UNR has been shown to act as an RNA chaperone on the Apaf-1 mRNA (Mitchell et al., 2003). UNR binding to Apaf-1 IRES modifies the structure of the RNA and allows for the binding of the nPTB protein. Further conformational rearrangements promoted by both proteins allow the small ribosomal subunit to recognize the IRES, thereby stimulating translation. Similarly, here we show that *Drosophila* UNR acts as an RNA chaperone on the *roX2* RNA thereby facilitating MLE interaction with *roX2*. By binding to its site on *roX2*, UNR produces a conformational change on the neighboring functional stem loop that results in the "melting" of the base of the stem and the exposure of the corresponding nucleotides. Interestingly, it has been shown that MLE requires a 3'-tailed single stranded region to be fully functional in the efficient unwinding of the targeted dsRNA (Lee et al., 1997). Opening of the base of the stem would provide a more extended platform for MLE binding and/or unwinding.

UNR regulates dosage compensation in female flies by acting as a repressor of *msl2* expression. In the female organism, UNR is recruited by the female-specific factor SXL to the 3'UTR of *msl2* mRNA and contributes to its translational repression (Abaza et al., 2006; Duncan et al., 2006). UNR is indeed needed for efficient inhibition of *msl2* since depletion of UNR in translation-competent embryo extracts reduces the efficiency of repression by SXL (Abaza et al., 2006) and UNR hypomorph mutant females show formation of MSL-complexes on the X-chromosome (Patalano et al., 2009). Although expressed in both sexes, UNR does not bind *msl2* mRNA in male flies because binding requires SXL. The requirement of SXL for UNR binding ensures the production of

MSL2 protein in the male organism, which is a prerequisite for the establishment of dosage compensation.

The molecular mechanism by which sex-specific binding of UNR is achieved was unclear at the initial stages of this thesis. By dissecting the structural properties of the SXL:UNR complex on the *msl2* 3'UTR, we gained insights on this question. The minimal functional *msl2* 3' UTR contains two poly-U stretches which serve as binding sites for SXL. We found that UNR binds to two sequences flanking the SXL binding sites. The presence of two binding sites for SXL and UNR allows the assembly of two tandem SXL:UNR complexes on the same *msl2* mRNA molecule. The functional relevance of such arrangement is still unclear; the presence of two different landing platforms for the SXL:UNR corepressor complex could create a high local concentration of the two proteins, ensuring interactions to yield ternary complex formation.

The fact that SXL cannot repress translation when artificially tethered to the mRNA speaks for the importance of correct RNA binding by SXL (Grskovic et al., 2003), SXL binding could expose the RNA in an appropriate conformation for UNR recognition. Supporting this scenario, the nucleotides bound by SXL in the 18mer RNA flank those recognized by UNR (Figure 29). Canonical SXL contacts are established with the U-rich stretch via the two RNP motifs located in the RBD β-sheets of the protein (Figure 11) (Handa et al., 1999). Additional non-canonical contacts between one of the RBD1 α-helices and the UGA sequence of the 18mer may extend the RNA for proper recognition by UNR. Once initial contacts are established, the formation of a stable complex relies on the cooperativity of the binding of SXL and UNR to the RNA. Binding of one protein to the RNA enhances the association of the other. Collectively, cooperativity could be explained by the ability of the two proteins to create a continuous surface with which the RNA could interact. Nucleotides from the RNA appear to play an important role in the SXL:UNR association, explaining why the two proteins are not able to interact in the absence of RNA. An intricate network of RNA-protein interactions results in the formation of a "zipper" that ensures tight complex formation (Figure 22). Moreover, the non-canonical contacts between SXL and the RNA further extend the protein surface available for RNA interaction.

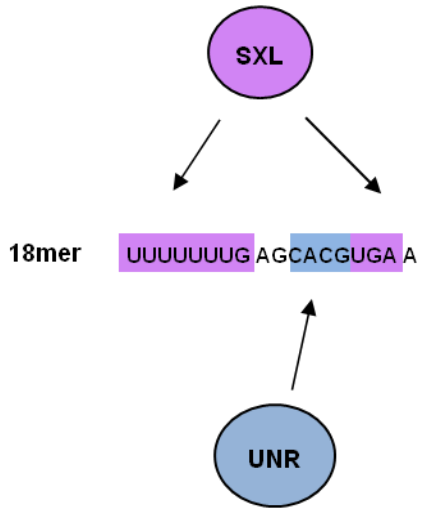


Figure 29. Sequence recognition by SXL and UNR.

The nucleotides of the 18mer RNA recognized by SXL and UNR are depicted in purple and blue respectively.

Overall we show that UNR promotes the formation of RNA-protein interactions that are fundamental for the regulation of dosage compensation in male and female flies. In males, UNR facilitates MLE association with *roX2*, ensuring efficient DCC assembly and function. In females, UNR enhances SXL binding to *msl2* 3'UTR and contributes to tighten *msl2* translational repression (Figure 30).

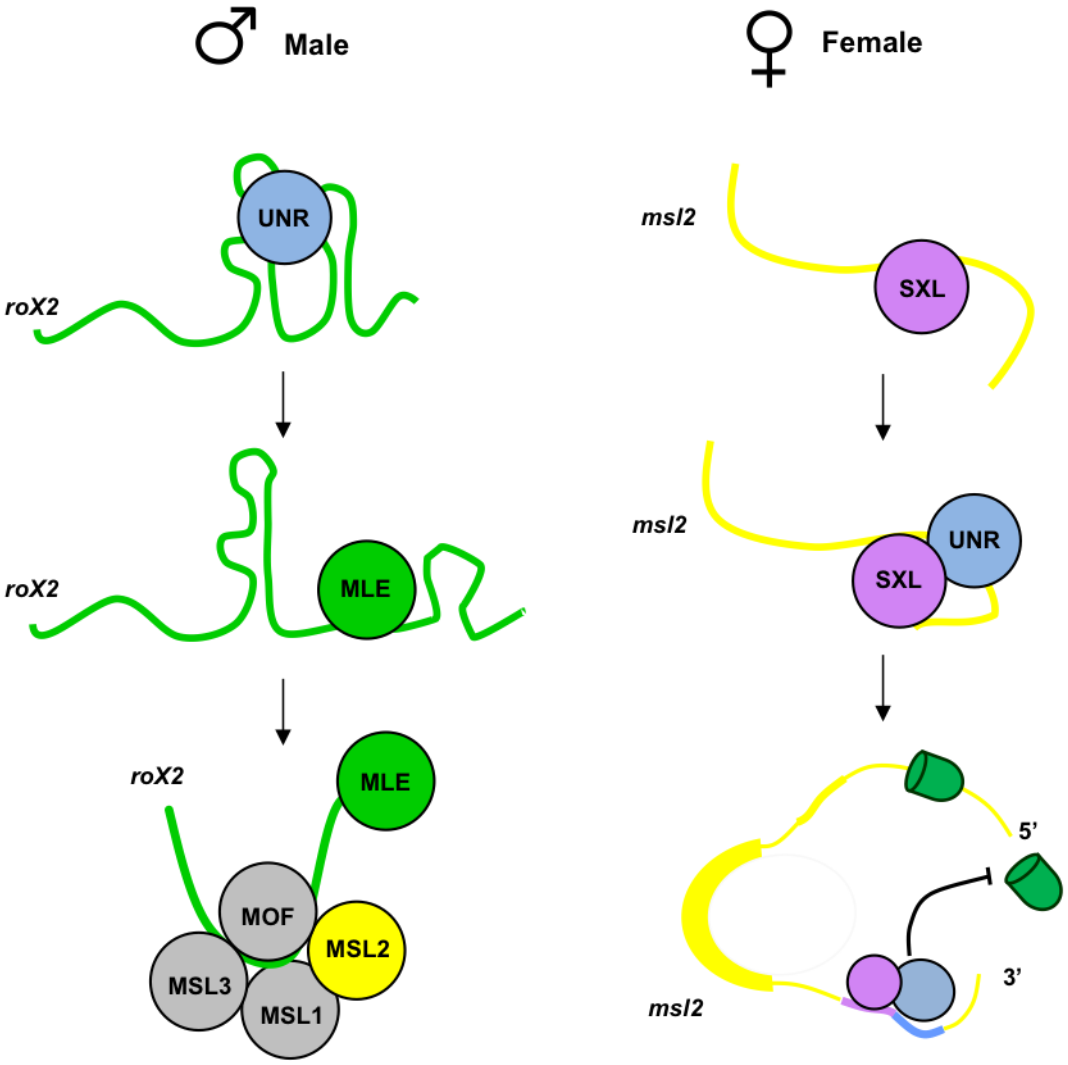


Figure 30. Molecular mechanisms by which UNR regulates dosage compensation.
In male flies, UNR binds to *roX2* and facilitates its association with MLE. This leads to the incorporation of both MLE and *roX2* into a functional DCC. In female flies, UNR enhances SXL binding to *msl2* mRNA. Cooperative interactions between SXL and UNR ensure tight *msl2* translational repression.

CONCLUSIONS

Part I

1. UNR binds to a region of *roX2* located immediately upstream of a functional stem-loop structure.
2. UNR binding to *roX2* induces a conformational change that exposes the base of the functional stem-loop.
3. UNR interacts with MLE.
4. UNR promotes MLE binding to *roX2 in vivo* and *in vitro.*

Part II

1. UNR binds to GCAC/U sequences in the *msl2* 3'-UTR downstream of the SXL binding sites.
2. The EF region of *msl2* 3' UTR supports tandem assembly of two units of the SXL:UNR complex.
3. SXL and UNR bind to the *msl2* 3'-UTR in a cooperative manner.
4. The interaction between SXL and UNR leads to the formation of a continuous surface with which the RNA interacts.
5. Tight complex formation is ensured by a molecular zipper in which RNA, SXL and UNR residues alternate.
6. Binding of UNR to SXL and *msl2* increases the RNA-binding potential of SXL.

BIBLIOGRAPHY

A

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.

Abaza, I., and Gebauer, F. (2008). Trading translation with RNA-binding proteins. RNA *14*, 404-409.

Aguilo, F., Zhou, M.M., and Walsh, M.J.(2011). Long noncoding RNA, polycomb, and the ghosts haunting INK4b-ARF-INK4a expression. Cancer Res *71*, 5365-5369.

Akhtar, A., and Becker, P.B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. Mol Cell *5*, 367-375.

Alekseyenko, A.A., Peng, S., Larschan, E., Gorchakov, A.A., Lee, O.K., Kharchenko, P., McGrath, S.D., Wang, C.I., Mardis, E.R., Park, P.J., and Kuroda, M.I. (2008). A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome. Cell *134*, 599-609.

Amrein, H., and Axel, R. (1997). Genes expressed in neurons of adult male Drosophila. Cell *88*, 459-469.

Arcus, V. (2002). OB-fold domains: a snapshot of the evolution of sequence, structure and function. Curr Opin Struct Biol *12*, 794-801.

Augui, S., Nora, E.P., and Heard, E. (2011). Regulation of X-chromosome inactivation by the X-inactivation centre. Nat Rev Genet *12*, 429-442.

Avner, P., and Heard, E. (2001). X-chromosome inactivation: counting, choice and initiation. Nat Rev Genet *2*, 59-67.

B

Bai, X., Alekseyenko, A.A., and Kuroda, M.I. (2004). Sequence-specific targeting of MSL complex regulates transcription of the roX RNA genes. EMBO J *23*, 2853-2861.

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

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

Bell, O., Conrad, T., Kind, J., Wirbelauer, C., Akhtar, A., and Schubeler, D. (2008). Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in Drosophila melanogaster. Mol Cell Biol *28*, 3401-3409.

Beltran, M., Puig, I., Pena, C., Garcia, J.M., Alvarez, A.B., Pena, R., Bonilla, F., and de Herreros, A.G. (2008). A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. Genes Dev *22*, 756-769.

Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011). The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. Mol Cell *43*, 1040-1046.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E.*, et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

Bond, C.S., and Fox, A.H. (2009). Paraspeckles: nuclear bodies built on long noncoding RNA. J Cell Biol *186*, 637-644.

Boussadia, O., Amiot, F., Cases, S., Triqueneaux, G., Jacquemin-Sablon, H., and Dautry, F. (1997). Transcription of unr (upstream of N-ras) down-modulates N-ras expression in vivo. FEBS Lett *420*, 20-24.

C

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C.*, et al.* (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559-1563.

Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., Tramontano, A., and Bozzoni, I. (2011). A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell *147*, 358-369.

Chang, K.A., and Kuroda, M.I. (1998). Modulation of MSL1 abundance in female Drosophila contributes to the sex specificity of dosage compensation. Genetics 150, 699-709.

Chau, J., Kulnane, L.S., and Salz, H.K. (2012). Sex-lethal enables germline stem cell differentiation by down-regulating Nanos protein levels during Drosophila oogenesis. Proc Natl Acad Sci U S A 109, 9465-9470.

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

Crowder, S.M., Kanaar, R., Rio, D.C., and Alber, T. (1999). Absence of interdomain contacts in the crystal structure of the RNA recognition motifs of Sex-lethal. Proc Natl Acad Sci U S A *96*, 4892-4897.

Csankovszki, G. (2009b). Condensin function in dosage compensation. Epigenetics *4*, 212-215.

Csankovszki, G., Petty, E.L., and Collette, K.S. (2009a). The worm solution: a chromosome-full of condensin helps gene expression go down. Chromosome Res *17*, 621-635.

D

Dahlsveen, I.K., Gilfillan, G.D., Shelest, V.I., Lamm, R., and Becker, P.B. (2006). Targeting determinants of dosage compensation in Drosophila. PLoS Genet *2*, e5.

Demakova, O.V., Kotlikova, I.V., Gordadze, P.R., Alekseyenko, A.A., Kuroda, M.I., and Zhimulev, I.F. (2003). The MSL complex levels are critical for its correct targeting to the chromosomes in Drosophila melanogaster. Chromosoma *112*, 103-115.

Deng, H., Zhang, W., Bao, X., Martin, J.N., Girton, J., Johansen, J., and Johansen, K.M. (2005). The JIL-1 kinase regulates the structure of Drosophila polytene chromosomes. Chromosoma *114*, 173-182.

Dethoff, E.A., Chugh, J., Mustoe, A.M., and Al-Hashimi, H.M. (2012). Functional complexity and regulation through RNA dynamics. Nature *482*, 322-330.

Dormoy-Raclet, V., Markovits, J., Malato, Y., Huet, S., Lagarde, P., Montaudon, D., Jacquemin-Sablon, A., and Jacquemin-Sablon, H. (2007). Unr, a cytoplasmic RNA-binding protein with cold-shock domains, is involved in control of apoptosis in ES and HuH7 cells. Oncogene *26*, 2595-2605.

Duncan, K., Grskovic, M., Strein, C., Beckmann, K., Niggeweg, R., Abaza, I., Gebauer, F., Wilm, M., and Hentze, M.W. (2006). Sex-lethal imparts a sex-specific

function to UNR by recruiting it to the msl-2 mRNA 3' UTR: translational repression for dosage compensation. Genes Dev *20*, 368-379.

Duncan, K.E., Strein, C., and Hentze, M.W. (2009). The SXL-UNR corepressor complex uses a PABP-mediated mechanism to inhibit ribosome recruitment to msl-2 mRNA. Mol Cell *36*, 571-582.

E

Ellegren, H. (2011). Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. Nat Rev Genet *12*, 157-166.

Ercan, S., and Lieb, J.D. (2009). C. elegans dosage compensation: a window into mechanisms of domain-scale gene regulation. Chromosome Res *17*, 215-227.

F

Fagegaltier, D., and Baker, B.S. (2004). X chromosome sites autonomously recruit the dosage compensation complex in Drosophila males. PLoS Biol *2*, e341.

Fauth, T., Muller-Planitz, F., Konig, C., Straub, T., and Becker, P.B. (2010). The DNA binding CXC domain of MSL2 is required for faithful targeting the Dosage Compensation Complex to the X chromosome. Nucleic Acids Res *38*, 3209-3221.

Feng, J., Bi, C., Clark, B.S., Mady, R., Shah, P., and Kohtz, J.D. (2006). The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. Genes Dev *20*, 1470-1484.

Forch, P., Merendino, L., Martinez, C., and Valcarcel, J. (2001). Modulation of msl-2 5' splice site recognition by Sex-lethal. RNA *7*, 1185-1191.

Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-Somoza, I., Leyva, A., Weigel, D., Garcia, J.A., and Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. Nat Genet *39*, 1033-1037.

Franke, A., and Baker, B.S. (1999). The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. Mol Cell *4*, 117-122.

Furuhashi, H., Nakajima, M., and Hirose, S. (2006). DNA supercoiling factor contributes to dosage compensation in Drosophila. Development *133*, 4475-4483.

G

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*, 1397-1404.

Gebauer, F., and Hentze, M.W. (2004). Molecular mechanisms of translational control. Nat Rev Mol Cell Biol *5*, 827-835.

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

Gelbart, M.E., and Kuroda, M.I. (2009). Drosophila dosage compensation: a complex voyage to the X chromosome. Development *136*, 1399-1410.

Gong, C., and Maquat, L.E. (2011). lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. Nature *470*, 284-288.

Graindorge, A., Militti, C., and Gebauer, F. Posttranscriptional 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 (in press).

Grimaud, C., and Becker, P.B. (2009). The dosage compensation complex shapes the conformation of the X chromosome in Drosophila. Genes Dev *23*, 2490-2495.

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.

Gu, W., Wei, X., Pannuti, A., and Lucchesi, J.C. (2000). Targeting the chromatin-remodeling MSL complex of Drosophila to its sites of action on the X chromosome requires both acetyl transferase and ATPase activities. EMBO J *19*, 5202-5211.

H

Hainer, S.J., Pruneski, J.A., Mitchell, R.D., Monteverde, R.M., and Martens, J.A. (2011) Intergenic transcription causes repression by directing nucleosome assembly. Genes Dev *25*, 29-40.

Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. (1999). Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. Nature *398*, 579-585.

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

Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., and Nakagawa, S. (2010). The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. Dev Cell *19*, 469-476.

Hellwig, S., and Bass, B.L. (2008). A starvation-induced noncoding RNA modulates expression of Dicer-regulated genes. Proc Natl Acad Sci U S A *105*, 12897-12902.

Hirota, K., Miyoshi, T., Kugou, K., Hoffman, C.S., Shibata, T., and Ohta, K. (2008). Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. Nature *456*, 130-134.

Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D., and Vogelauer, M. (2008). A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. Mol Cell *32*, 685-695.

Huarte, M., and Rinn, J.L. (2010). Large non-coding RNAs: missing links in cancer? Hum Mol Genet *19*, R152-161.

I

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

J

Jeon, Y., and Lee, J.T. (2011). YY1 tethers Xist RNA to the inactive X nucleation center. Cell *146*, 119-133.

Johansson, A.M., Allgardsson, A., Stenberg, P., and Larsson, J. (2011). msl2 mRNA is bound by free nuclear MSL complex in Drosophila melanogaster. Nucleic Acids Res *39*, 6428-6439.

K

Kageyama, Y., Mengus, G., Gilfillan, G., Kennedy, H.G., Stuckenholz, C., Kelley, R.L., Becker, P.B., and Kuroda, M.I. (2001). Association and spreading of the Drosophila dosage compensation complex from a discrete roX1 chromatin entry site. EMBO J *20*, 2236-2245.

Kanduri, C. (2011). Kcnq1ot1: a chromatin regulatory RNA. Semin Cell Dev Biol *22*, 343-350.

Kapranov, P., Cawley, S.E., Drenkow, J., Bekiranov, S., Strausberg, R.L., Fodor, S.P., and Gingeras, T.R. (2002). Large-scale transcriptional activity in chromosomes 21 and 22. Science *296*, 916-919.

Kapranov, P., Willingham, A.T., and Gingeras, T.R. (2007). Genome-wide transcription and the implications for genomic organization. Nat Rev Genet *8*, 413-423.

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*, 867-877.

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*, 195-199.

Kind, J., and Akhtar, A. (2007). Cotranscriptional recruitment of the dosage compensation complex to X-linked target genes. Genes Dev *21*, 2030-2040.

Kind, J., Vaquerizas, J.M., Gebhardt, P., Gentzel, M., Luscombe, N.M., Bertone, P., and Akhtar, A. (2008). Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in Drosophila. Cell *133*, 813-828.

Kloc, M., Bilinski, S., and Dougherty, M.T. (2007). Organization of cytokeratin cytoskeleton and germ plasm in the vegetal cortex of Xenopus laevis oocytes depends on coding and non-coding RNAs: three-dimensional and ultrastructural analysis. Exp Cell Res *313*, 1639-1651.

Kloc, M., Wilk, K., Vargas, D., Shirato, Y., Bilinski, S., and Etkin, L.D. (2005). Potential structural role of non-coding and coding RNAs in the organization of the cytoskeleton at the vegetal cortex of Xenopus oocytes. Development *132*, 3445-3457.

L

Lai, F., Orom, U.A., Cesaroni, M., Beringer, M., Taatjes, D.J., Blobel, G.A., and Shiekhattar, R. (2013)**.** Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature *494*, 497-501.

Lallena, M.J., Chalmers, K.J., Llamazares, S., Lamond, A.I., and Valcarcel, J. (2002). Splicing regulation at the second catalytic step by Sex-lethal involves 3' splice site recognition by SPF45. Cell *109*, 285-296.

Lanz, R.B., McKenna, N.J., Onate, S.A., Albrecht, U., Wong, J., Tsai, S.Y., Tsai, M.J., and O'Malley, B.W. (1999). A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. Cell *97*, 17-27.

Larschan, E., Alekseyenko, A.A., Gortchakov, A.A., Peng, S., Li, B., Yang, P., Workman, J.L., Park, P.J., and Kuroda, M.I. (2007). MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. Mol Cell *28*, 121-133.

Lee, C.G., Chang, K.A., Kuroda, M.I., and Hurwitz, J. (1997). The NTPase/helicase activities of Drosophila maleless, an essential factor in dosage compensation. EMBO J *16*, 2671-2681.

Lee, J.T., and Bartolomei, M.S. (2013). X-inactivation, imprinting, and long noncoding RNAs in health and disease. Cell *152*, 1308-1323.

Leshkowitz, D., Rozenblatt, O., Nakamura, T., Yano, T., Dautry, F., Croce, C.M., and Canaani, E. (1996). ALL-1 interacts with unr, a protein containing multiple cold shock domains. Oncogene *13*, 2027-2031.

Lim, C.K., and Kelley, R.L. Autoregulation of the Drosophila Noncoding roX1 RNA Gene. PLoS Genet *8*, e1002564.

Lyman, L.M., Copps, K., Rastelli, L., Kelley, R.L., and Kuroda, M.I. (1997). Drosophila male-specific lethal-2 protein: structure/function analysis and dependence on MSL-1 for chromosome association. Genetics *147*, 1743-1753.

M

Martens, J.A., Laprade, L., and Winston, F. (2004). Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature *429*, 571-574.

Max, K.E., Zeeb, M., Bienert, R., Balbach, J., and Heinemann, U. (2006). T-rich DNA single strands bind to a preformed site on the bacterial cold shock protein Bs-CspB. J Mol Biol *360*, 702-714.

Medenbach, J., Seiler, M., and Hentze, M.W. (2011). Translational control via protein-regulated upstream open reading frames. Cell *145*, 902-913.

Meller, V.H., Gordadze, P.R., Park, Y., Chu, X., Stuckenholz, C., Kelley, R.L., and Kuroda, M.I. (2000). Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in Drosophila. Curr Biol *10*, 136-143.

Meller, V.H., and Rattner, B.P. (2002). The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. EMBO J *21*, 1084-1091.

Meller, V.H., Wu, K.H., Roman, G., Kuroda, M.I., and Davis, R.L. (1997). roX1 RNA paints the X chromosome of male Drosophila and is regulated by the dosage compensation system. Cell *88*, 445-457.

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

Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J.*, et al.* (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila. Mol Cell *21*, 811-823.

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

Mihailovic, M., Wurth, L., Zambelli, F., Abaza, I., Militti, C., Mancuso, F.M., Roma, G., Pavesi, G., and Gebauer, F. (2012) Widespread generation of alternative UTRs contributes to sex-specific RNA binding by UNR. RNA *18*, 53-64.

Mihailovich, M., Militti, C., Gabaldon, T., and Gebauer, F. (2010). Eukaryotic cold shock domain proteins: highly versatile regulators of gene expression. Bioessays *32*, 109-118.

Mitchell, S.A., Spriggs, K.A., Coldwell, M.J., Jackson, R.J., and Willis, A.E. (2003). The Apaf-1 internal ribosome entry segment attains the correct structural conformation for function via interactions with PTB and unr. Mol Cell *11*, 757-771.

N

Nagai, K., Oubridge, C., Jessen, T.H., Li, J., and Evans, P.R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. Nature *348*, 515-520.

O

Oh, H., Bone, J.R., and Kuroda, M.I. (2004). Multiple classes of MSL binding sites target dosage compensation to the X chromosome of Drosophila. Curr Biol *14*, 481-487.

Oh, H., Park, Y., and Kuroda, M.I. (2003). Local spreading of MSL complexes from roX genes on the Drosophila X chromosome. Genes Dev *17*, 1334-1339.

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H.*, et al.* (2002). Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature *420*, 563-573.

Orom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q.*, et al.* (2010). Long noncoding RNAs with enhancer-like function in human cells. Cell *143*, 46-58.

P

Pang, K.C., Frith, M.C., and Mattick, J.S. (2006). Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. Trends Genet *22*, 1-5.

Park, S.W., Kang, Y., Sypula, J.G., Choi, J., Oh, H., and Park, Y. (2007). An evolutionarily conserved domain of roX2 RNA is sufficient for induction of H4-Lys16 acetylation on the Drosophila X chromosome. Genetics *177*, 1429-1437.

Park, S.W., Kuroda, M.I., and Park, Y. (2008). Regulation of histone H4 Lys16 acetylation by predicted alternative secondary structures in roX noncoding RNAs. Mol Cell Biol *28*, 4952-4962.

Park, Y., Kelley, R.L., Oh, H., Kuroda, M.I., and Meller, V.H. (2002). Extent of chromatin spreading determined by roX RNA recruitment of MSL proteins. Science *298*, 1620-1623.

Park, Y., Mengus, G., Bai, X., Kageyama, Y., Meller, V.H., Becker, P.B., and Kuroda, M.I. (2003). Sequence-specific targeting of Drosophila roX genes by the MSL dosage compensation complex. Mol Cell *11*, 977-986.

Park, Y., Oh, H., Meller, V.H., and Kuroda, M.I. (2005). Variable splicing of non-coding roX2 RNAs influences targeting of MSL dosage compensation complexes in Drosophila. RNA Biol *2*, 157-164.

Patalano, S., Mihailovich, M., Belacortu, Y., Paricio, N., and Gebauer, F. (2009). Dual sex-specific functions of Drosophila Upstream of N-ras in the control of X chromosome dosage compensation. Development *136*, 689-698.

Patel, G.P., Ma, S., and Bag, J. (2005). The autoregulatory translational control element of poly(A)-binding protein mRNA forms a heteromeric ribonucleoprotein complex. Nucleic Acids Res *33*, 7074-7089.

Payer, B., and Lee, J.T. (2008). X chromosome dosage compensation: how mammals keep the balance. Annu Rev Genet *42*, 733-772.

Penalva, L.O., and Sanchez, L. (2003). RNA binding protein sex-lethal (Sxl) and control of Drosophila sex determination and dosage compensation. Microbiol Mol Biol Rev *67*, 343-359, table of contents.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W.J., and Pandolfi, P.P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature *465*, 1033-1038.

Ponjavic, J., Ponting, C.P., and Lunter, G. (2007). Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. Genome Res *17*, 556-565.

R

Richter, L., Bone, J.R., and Kuroda, M.I. (1996). RNA-dependent association of the Drosophila maleless protein with the male X chromosome. Genes Cells *1*, 325-336.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., and Chang, H.Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

S

Sakamoto, H., Inoue, K., Higuchi, I., Ono, Y., and Shimura, Y. (1992). Control of Drosophila Sex-lethal pre-mRNA splicing by its own female-specific product. Nucleic Acids Res *20*, 5533-5540.

Sakashita, E., and Sakamoto, H. (1994). Characterization of RNA binding specificity of the Drosophila sex-lethal protein by in vitro ligand selection. Nucleic Acids Res *22*, 4082-4086.

Sakashita, E., and Sakamoto, H. (1996). Protein-RNA and protein-protein interactions of the Drosophila sex-lethal mediated by its RNA-binding domains. J Biochem *120*, 1028-1033.

Samuels, M., Deshpande, G., and Schedl, P. (1998). Activities of the Sex-lethal protein in RNA binding and protein:protein interactions. Nucleic Acids Res *26*, 2625-2637.

Samuels, M.E., Bopp, D., Colvin, R.A., Roscigno, R.F., Garcia-Blanco, M.A., and Schedl, P. (1994). RNA binding by Sxl proteins in vitro and in vivo. Mol Cell Biol *14*, 4975-4990.

Santoro, F., and Pauler, F.M. Silencing by the imprinted Airn macro lncRNA: transcription is the answer. Cell Cycle *12*, 711-712.

Santoro, F., and Pauler, F.M. (2013). Silencing by the imprinted Airn macro lncRNA: transcription is the answer. Cell Cycle 12, 711-712.

Schmitz, K.M., Mayer, C., Postepska, A., and Grummt, I. (2010). Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. Genes Dev *24*, 2264-2269.

Schutt, C., and Nothiger, R. (2000). Structure, function and evolution of sex-determining systems in Dipteran insects. Development *127*, 667-677.

Scott, M.J., Pan, L.L., Cleland, S.B., Knox, A.L., and Heinrich, J. (2000). MSL1 plays a central role in assembly of the MSL complex, essential for dosage compensation in Drosophila. EMBO J *19*, 144-155.

Singh, R., Valcarcel, J., and Green, M.R. (1995). Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science *268*, 1173-1176.

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*, 31483-31486.

Smith, E.R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R.G., Allis, C.D., and Lucchesi, J.C. (2000). The drosophila MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. Mol Cell Biol *20*, 312-318.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*, 731-745.

Sosnowski, B.A., Belote, J.M., and McKeown, M. (1989). Sex-specific alternative splicing of RNA from the transformer gene results from sequence-dependent splice site blockage. Cell *58*, 449-459.

Spierer, A., Seum, C., Delattre, M., and Spierer, P. (2005). Loss of the modifiers of variegation Su(var)3-7 or HP1 impacts male X polytene chromosome morphology and dosage compensation. J Cell Sci *118*, 5047-5057.

Straub, T., Grimaud, C., Gilfillan, G.D., Mitterweger, A., and Becker, P.B. (2008). The chromosomal high-affinity binding sites for the Drosophila dosage compensation complex. PLoS Genet *4*, e1000302.

Straub, T., Zabel, A., Gilfillan, G.D., Feller, C., and Becker, P.B. (2013). Different chromatin interfaces of the Drosophila dosage compensation complex revealed by high-shear ChIP-seq. Genome Res *23*, 473-485.

Struhl, K. (2007). Transcriptional noise and the fidelity of initiation by RNA polymerase II. Nat Struct Mol Biol *14*, 103-105.

Sural, T.H., Peng, S., Li, B., Workman, J.L., Park, P.J., and Kuroda, M.I. (2008). The MSL3 chromodomain directs a key targeting step for dosage compensation of the Drosophila melanogaster X chromosome. Nat Struct Mol Biol *15*, 1318-1325.

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

T

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

Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M., and Hannon, G.J. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534-538.

Theobald, D.L., Mitton-Fry, R.M., and Wuttke, D.S. (2003). Nucleic acid recognition by OB-fold proteins. Annu Rev Biophys Biomol Struct *32*, 115-133.

Tian, D., Sun, S., and Lee, J.T. (2010). The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. Cell *143*, 390-403.

Triqueneaux, G., Velten, M., Franzon, P., Dautry, F., and Jacquemin-Sablon, H. (1999). RNA binding specificity of Unr, a protein with five cold shock domains. Nucleic Acids Res *27*, 1926-1934.

Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science *329*, 689-693.

V

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

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Jr., Hieter, P., Vogelstein, B., and Kinzler, K.W. (1997). Characterization of the yeast transcriptome. Cell *88*, 243-251.

Villa, R., Forne, I., Muller, M., Imhof, A., Straub, T., and Becker, P.B. (2012). MSL2 combines sensor and effector functions in homeostatic control of the Drosophila dosage compensation machinery. Mol Cell *48*, 647-654.

W

Walters, R.D., Kugel, J.F., and Goodrich, J.A. (2009). InvAluable junk: the cellular impact and function of Alu and B2 RNAs. IUBMB Life *61*, 831-837.

Wang, D., Garcia-Bassets, I., Benner, C., Li, W., Su, X., Zhou, Y., Qiu, J., Liu, W., Kaikkonen, M.U., Ohgi, K.A.*, et al.* (2011). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature *474*, 390-394.

Wang, J., and Bell, L.R. (1994). The Sex-lethal amino terminus mediates cooperative interactions in RNA binding and is essential for splicing regulation. Genes Dev *8*, 2072-2085.

Wang, J., Gong, C., and Maquat, L.E. (2013). Control of myogenesis by rodent SINE-containing lncRNAs. Genes Dev *27*, 793-804.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A.*, et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature *454*, 126-130.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T.*, et al.* (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature *453*, 539-543.

Worringer, K.A., Chu, F., and Panning, B. (2009). The zinc finger protein Zn72D and DEAD box helicase Belle interact and control maleless mRNA and protein levels. BMC Mol Biol *10*, 33.

Worringer, K.A., and Panning, B. (2007). Zinc finger protein Zn72D promotes productive splicing of the maleless transcript. Mol Cell Biol *27*, 8760-8769.

Z

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J., and Lee, J.T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 322, 750-756.

ANNEXES

I.

Mihailovich M, Militti C, Gabaldón T, Gebauer F. Eukaryotic cold shock domain proteins: highly versatile regulators of gene expression. Bioessays. 2010 Feb;32(2):109-18.

II.

Graindorge A, Militti C, Gebauer F. Posttranscriptional control of X-chromosome dosage compensation. Wiley Interdiscip Rev RNA. 2011 Jul-Aug;2(4):534-45.

III.

Mihailovic M, Wurth L, Zambelli F, Abaza I, Militti C, Mancuso FM, Roma G, Pavesi G, Gebauer F. Widespread generation of alternative UTRs contributes to sex-specific RNA binding by UNR. RNA. 2012 Jan;18(1):53-64.