Molecular Cell
Article

Tandem Stem-Loops in roX RNAs Act Together to Mediate X Chromosome Dosage Compensation in *Drosophila*

Ibrahim Avsar Ilik,[1,2] Jeffrey J. Quinn,[3,8] Plamen Georgiev,[1,8] Filipe Tavares-Cadete,[4,8] Daniel Maticzka,[5,8] Sarah Toscano,[1] Yue Wan,[3] Robert C. Spitale,[3] Nicholas Luscombe,[6] Rolf Backofen,[6,7] Howard Y. Chang,[3] and Asifa Akhtar[1,*]

[1]Max-Planck Institute of Immunobiology and Epigenetics, Stübeweg 51, 79108 Freiburg im Breisgau, Germany
[2]European Molecular Biology Laboratory, Meyerhofstraße 1, 69117 Heidelberg, Germany
[3]Howard Hughes Medical Institute and Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA 94305, USA
[4]EMBL European Bioinformatics Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SD, UK
[5]Bioinformatics Group, Department of Computer Science, University of Freiburg, Georges-Köhler-Allee 106, 79110 Freiburg im Breisgau, Germany
[6]UCL Genetics Institute, Department of Genetics, Evolution, and Environment, University College London, Gower Street, London WC1E 6BT, UK
[7]Centre for Biological Signalling Studies (BIOSS), University of Freiburg, Schänzlestrasse 18, 79104 Freiburg im Breisgau, Germany
[8]These authors contributed equally to this work
*Correspondence: akhtar@ie-freiburg.mpg.de
http://dx.doi.org/10.1016/j.molcel.2013.07.001

SUMMARY

Dosage compensation in *Drosophila* is an epigenetic phenomenon utilizing proteins and long noncoding RNAs (lncRNAs) for transcriptional upregulation of the male X chromosome. Here, by using UV cross-linking followed by deep sequencing, we show that two enzymes in the Male-Specific Lethal complex, MLE RNA helicase and MSL2 ubiquitin ligase, bind evolutionarily conserved domains containing tandem stem-loops in roX1 and roX2 RNAs in vivo. These domains constitute the minimal RNA unit present in multiple copies in diverse arrangements for nucleation of the MSL complex. MLE binds to these domains with distinct ATP-independent and ATP-dependent behavior. Importantly, we show that different roX RNA domains have overlapping function, since only combinatorial mutations in the tandem stem-loops result in severe loss of dosage compensation and consequently male-specific lethality. We propose that repetitive structural motifs in lncRNAs could provide plasticity during multiprotein complex assemblies to ensure efficient targeting in *cis* or in *trans* along chromosomes.

INTRODUCTION

Long noncoding RNAs (lncRNAs) are emerging as important regulators of chromatin state and transcription in eukaryotic cells. They can contribute to the regulation of single genes or whole chromosomes and can influence the 3D structure of large genomic regions. Due to their length, which typically is in the range of kilobases, it has been difficult to determine functional domains in these lncRNAs and therefore to understand their exact contributions to transcriptional regulation (for reviews, see Augui et al., 2011; Rinn and Chang, 2012).

In both *Drosophila* and mammals, X-chromosomal dosage compensation is carried out by the concerted action of lncRNAs and protein complexes (Maenner et al., 2012). In mammals, females suppress transcription from one of the two X chromosomes in a process called X chromosome inactivation (XCI) (Augui et al., 2011; Jeon et al., 2012). One of the most prominent lncRNAs that is involved in XCI is a 17 kb long lncRNA called X-inactive specific transcript (Xist), which is transcribed from the X-inactivation centre (*Xic*) and coats the X chromosome in *cis*, leading to its transcriptional repression. Xist is hypothesized to achieve this repression by recruiting the PRC2 complex mostly through a conserved region called "repeat A" toward the 5′ end of the RNA (Zhao et al., 2008). This repeat is one of several in Xist RNA that has been shown to be important for its function (Wutz et al., 2002).

In contrast to mammals, *Drosophila* achieves dosage compensation by transcriptionally upregulating the single X chromosome in males (Conrad and Akhtar, 2011). Although the end result is opposite (activation versus repression), *Drosophila* also utilizes lncRNAs for dosage compensation. Both transcribed from the X chromosome, these RNAs are called RNA on the X 1 and 2 (roX1 and roX2) (Amrein and Axel, 1997; Ilik and Akhtar, 2009; Meller and Rattner, 2002; Meller et al., 1997), and together with five proteins (MSL1, MSL2, MSL3, MOF, and MLE) they form the Male-Specific Lethal (MSL) complex. Once formed, the MSL complex coats the X chromosome and acetylates H4K16 through the acetyltransferase activity of MOF, which is linked to increased transcriptional output of X-chromosomal genes in males (Conrad et al., 2012; Larschan et al., 2011).

Both roX1 and roX2 contain conserved regions that are shared by the two RNAs called roX boxes (henceforth RB or

RB element). Spotted in one of the earliest studies on roX RNAs (Franke and Baker, 1999) as a short stretch of sequence common to both RNAs, the biological significance of these elements is still unknown, although genetic studies have shown that they are important for the function of both roX1 (Kelley et al., 2008) and roX2 (Park et al., 2007, 2008) in dosage compensation.

Immunoprecipitation of various members of the complex from cell extracts, with or without formaldehyde fixation, has shown that roX RNAs are found associated with the MSL complex with or without MLE (Akhtar et al., 2000; Fauth et al., 2010; Izzo et al., 2008; Meller et al., 2000; Smith et al., 2000). However, genetic experiments suggest that MLE is important for the incorporation of the roX RNA into the MSL complex (Meller et al., 2000) and in the absence of MLE, MSL1, and MSL2 binds only to high-affinity sites (HASs) on the X chromosome (Gilfillan et al., 2004). Therefore, studying how roX RNAs interact with MSL complex members is important to gain a better understanding of the mechanism underlying dosage compensation.

Here, by employing individual-nucleotide resolution UV cross-linking and immunoprecipitation (iCLIP), we show that the most prominent targets of MLE and MSL2 are roX1 and roX2 in vivo. Intriguingly, roX1 and roX2 are bound by MLE and MSL2 only at discrete domains that are common for both proteins. These conserved domains contain highly structured tandem stem-loops with a repetitive organization shared by both roX1 and roX2. We further show that MLE binding to the first half of the roX2 RNA occurs in an ATP-independent manner, while the binding to the second half is ATP dependent. The ATP-independent interaction is mediated via the N-terminal double-strand RNA-binding domains of MLE. Importantly, we show that tandem stem-loops in roX2 RNA have partially overlapping functions, as combined mutations cause male-specific lethality. Taken together, these data reveal the critical role of structured domains in roX RNAs in nucleation of the MSL complex for efficient targeting to the X chromosome. We propose that such RNA organization could be a widespread feature of lncRNAs that can facilitate assembly and propagation of multiprotein complexes along chromatin in *cis* and in *trans*. Moreover, MLE's differential binding mode could allow, on one side, stable anchoring on the RNA and at the same time can ensure that roX RNAs act as dynamic platforms for MSL complex assembly.

RESULTS

MLE Is Enriched on HAS, and Its Chromatin Association Is RNase Sensitive

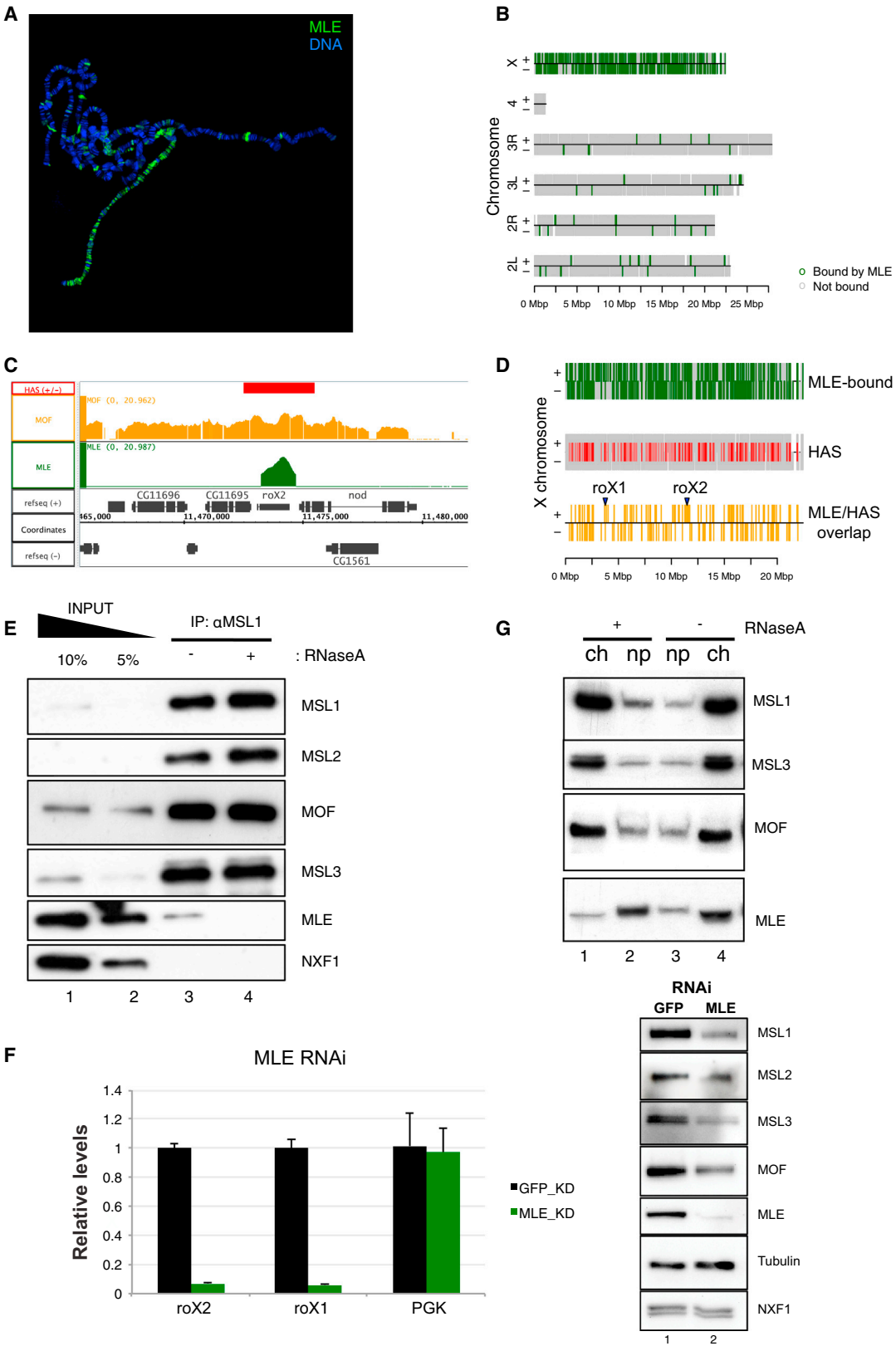
Although MLE is enriched on the X chromosome (Figure 1A; Kuroda et al., 1991), how and where this protein interacts with chromatin has not been thoroughly explored. This knowledge is important for our understanding of X-chromosomal targeting, since recent studies show that the MSL complex members do not have identical binding patterns on X-linked genes: MOF binds both promoters and gene bodies, while MSL3 is only enriched in gene bodies (Hallacli et al., 2012; Kadlec et al., 2011; Kind et al., 2008). Therefore, we first carried out chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) on wild-type Schneider (S2) cells using a specific MLE antibody to identify binding sites for MLE in the *Drosophila* genome. As expected, MLE-bound genes are concentrated on the X chromosome (Figure 1B). However, in contrast to MOF, MLE binding appeared almost restricted to HAS on the X chromosome (Figures 1C and 1D, see Figures S1A and S1E online; 204 [41.7%] of MLE-bound genes overlap a HAS, and 170 [89.9%] of HASs fall within a MLE-bound gene). Interestingly, in biochemical assays MLE is weakly associated with the core MSL complex (MSL1, MSL2, MSL3, MOF) (Figure 1E), and this weak interaction is sensitive to RNase A treatment (Figure 1E, compare lanes 3 and 4). However, despite this apparent substoichiometric association, depletion of MLE leads to reduction of MSL1-3 and MOF protein levels in addition to severely reduced roX1 and roX2 RNAs (Figure 1F).

Since MLE's interaction with the MSL complex members was not robust in our assay conditions, we rationalized that MLE could be part of a different protein complex. We therefore generated flies carrying an HA-FLAG-tagged MLE transgene (which rescues male lethality in a loss-of-function mutant mle^{1} background and localizes to the X chromosome just as the wild-type MLE does, Figures S1B and S1C) and purified it from embryonic nuclear extracts. To our surprise, we could only detect MLE in these purifications (Figure S1D), suggesting that interactions between MLE and other proteins, if any, are too weak or transient and did not withstand our purification conditions.

MLE's RNase-sensitive association with the MSL complex prompted us to address whether bulk MLE-chromatin interactions could also be RNase sensitive. We therefore incubated S2 nuclei with RNase A. Most of MLE is dissociated by this treatment (Figure 1G), which is consistent with MLE's RNase-sensitive association on polytene chromosomes (Richter et al., 1996). Taken together, these observations suggest that MLE-chromatin interactions include RNA intermediates. However, it is possible that protein-protein interactions are also involved in MLE's interaction with chromatin and the MSL complex, which could be dynamic and therefore not easy to score in our assays.

iCLIP Identifies MLE- and MSL2-Interacting RNAs in the *Drosophila* Transcriptome

MLE interaction with chromatin and the MSL complex appears to be RNase sensitive. Since the MSL complex contains roX RNAs, they could constitute the most obvious interaction candidates. However, roX RNAs can also be recovered in immunoprecipitations of the MSL complex where no MLE can be detected (Akhtar et al., 2000; Mendjan et al., 2006), indicating that there has to be at least one, if not more, member of the MSL complex that interacts with these lncRNAs. In previous studies we reported that partial MSL complexes lacking MSL3 and/or MOF can still coimmunoprecipitate roX RNA (Kadlec et al., 2011), whereas complexes that lack MSL2 lack any detectable roX RNA (Hallacli et al., 2012). This makes MSL2 a good candidate for RNA interaction, since it has been shown to interact with both DNA and RNA in vitro (Fauth et al., 2010). Furthermore, it remains unclear whether the MSL complex members only interact with roX RNAs or whether they also have other RNA targets. We systematically addressed this issue by using iCLIP, a technology that provides direct evidence for protein-RNA interaction in vivo (König et al., 2010). For this purpose, antibodies against MLE and MSL2



(legend on next page)

were used to isolate UV-crosslinked RNAs from *Drosophila* clone8 cells, followed by deep sequencing analysis (Figures 2A and 2B, see the Supplemental Information). This allowed us to investigate crosslinking events genome-wide on a nucleotide-scale resolution (Figures 2C and 2D). Genome-wide view of all targets clearly shows that MLE and MSL2 bind a number of RNAs in the *Drosophila* transcriptome. However, roX1 and roX2 stand out as the most significant lncRNA targets for both proteins, which are further characterized in this study. (Figure 2E, Figures S2 and S3).

SHAPE and PARS Analysis Reveals Conserved Secondary Structures within roX1 and roX2

Previous studies have shown that roX RNAs contain conserved elements known as roX boxes, which could play an important role in dosage compensation (Kelley et al., 2008; Park et al., 2007, 2008; Stuckenholz et al., 2003). However, no systematic biochemical analysis of their secondary structure has been reported to date. Therefore, before analyzing the MLE/MSL-roX interaction in detail, we performed **s**elective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) (Wilkinson et al., 2006) and parallel analysis of RNA structure (PARS) (Kertesz et al., 2010) on roX1 and roX2 to study their intrinsic secondary structures, which may contribute to their specific interactions with the MSL complex (Figure 3, Figures S4 and S5).

SHAPE is a chemical technique for probing RNA structure, which resolves unpaired RNA bases of a target RNA (Wilkinson et al., 2006; also see the Experimental Procedures and the Supplemental Information). Given that the 3′ terminus of roX1 has been previously shown to be important for dosage compensation and sensitive to deletion/mutation (Kelley et al., 2008; Stuckenholz et al., 2003), and that MLE and MSL2 directly interact with this region by iCLIP (Figure 4), we first sought to analyze the structure of the 3′-terminal domain of roX1. We subjected full-length in vitro-transcribed roX1 to SHAPE and resolved sites of modification by PAGE (Figure 3). We integrated the band intensities and calculated the "SHAPE reactivity" for each RNA base, a measurement of its single-strandedness. We used the individual bases' SHAPE reactivities to direct a structure model of the interrogated roX1 region. Figure 3A summarizes one such region (roX1 helix 1, R1H1), showing lightly modified bases with low SHAPE reactivity characteristic of paired helices, and heavily modified bases with high SHAPE reactivity characteristic of unpaired loops, bulges, and linkers. Piecing together the SHAPE reactivities for each reverse transcription priming event on roX1, we built a composite structure model that best fits the data (Figure 3C) and identified three stable helices connected by flexible linker regions, supporting previous structure predictions (inset, Figure 3B). The largest, most stable structure identified is R1H1. The stem contains a sequence element that resembles the roX box motif that we called "roX box-like" (RBL from here on, also see below and Figure 5E and Figure S4) motif. We have found other instances of this RBL motif also in helical structures within roX2 RNA (discussed below). Immediately downstream is R1H2, a helix formed by a long-range interaction between the inverted roX box (IRB) and roX box 1 (RB1) (see also Figure S4 and Figure 4B). In the loop formed by R1H2 lies the "P2" stem-loop and A-bulge (Figure 3C). Interestingly, the low-complexity, adenosine-rich A-bulge has low SHAPE reactivity characteristic of paired bases, suggesting that it may pair with one of several distant U-rich regions on roX1. The remaining 3′ terminus of roX1, including two roX boxes, is not predicted to fold into a significant structure (Figure S4).

Following the structural map of roX1, we next analyzed roX2 by SHAPE. Again, we identified support for structures such as R2H1 (for roX2 helix 1), a 21 bp stem-loop at the 5′ end of roX2 exon-3 (Figure 3D). We constructed a global structure model for roX2 and found that roX2 exon-3 is organized into repeated structural domains that are connected by flexible linkers (Figure 3F and Figure S5). The overall structure of roX2 RNA can be split into two clusters of tandem stem-loops. The first cluster is situated at the 5′ end and consists of four stem-loops (R2H1–R2H3 and P3), the first three containing incidences of the RBL motif (R2H1-R2H3) in their stems (Figure 3E and Figure 5B). The second cluster is at the 3′ end of the RNA and consists also of four stem-loops (P4 and R2H4–R2H6), this time the last three containing RB elements instead of RBL elements which fold into stable stem-loops (R2H4–R2H6). Two secondary structures, P3 and P4, lack a recognizable RB/RBL motif. The intervening $(CAATA)_n$ repeat region separating the two hairpin clusters has no significant reconcilable structure. To complement the SHAPE data for roX2, we also performed PARS, an enzymatic RNA structure probing technique that relies on the structure specificity of V1 and S1 ribonucleases (Figure S5),

Figure 1. MLE Is Enriched on High-Affinity Sites, and Its Chromatin Association Is RNase Sensitive

(A) A polytene squash staining of male third-instar larvae shows that MLE (green) coats the X chromosome. DNA is shown in blue (DAPI).

(B) ChIP-seq analysis shows that MLE is enriched on the X chromosome. All genes are represented in their correct chromosomal location (gray). MLE binding sites are shown in green. Top (+) and bottom strand (−) are indicated. Scale at the bottom indicates the size chromosomal length in megabasepair (Mbp).

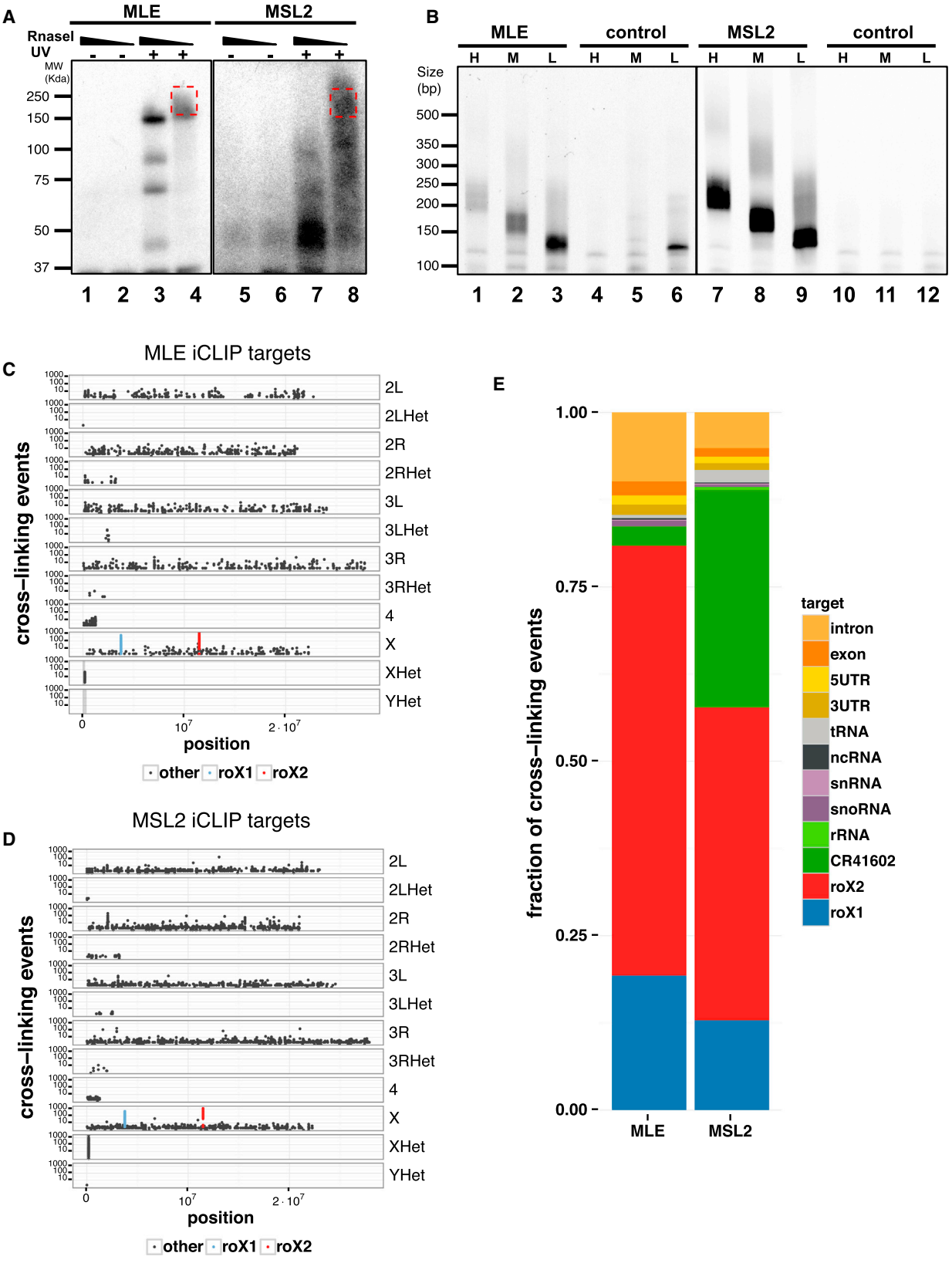
(C) A browser snapshot for roX2 HAS. MLE binding (green) to HAS is more restricted than MOF (orange). The red box on top marks the HAS that overlaps the roX2 gene.

(D) Almost all HASs are bound by MLE. The X chromosome is represented in three tracks: MLE-bound genes are shown in green (as in B), the HASs in red, and the MLE-bound HASs in orange (also see Figure S1E). Two specific HASs, the roX1 and roX2 genes, are specified by arrows.

(E) Immunoprecipitation of MSL1 from an S2 nuclear extract, under mild conditions, results in coprecipitation of small amounts of MLE in comparison to MSL2, MSL3, and MOF (compare lanes 1 and 2 with lane 3). RNase A treatment (lane 4) leads to the loss of this weak MLE interaction. NXF1 served as a negative control.

(F) RNAi-mediated depletion of MLE in S2 cells led to the destabilization of the MSL complex together with the loss of roX RNAs. (Left) Total RNA was isolated from double-stranded RNA (GFP [control] or MLE) treated S2 cells. Expression levels of roX1 and roX2 were determined by RT-qPCR analysis. PGK levels were used as a control. Error bars represent standard deviation (± SD) of three biological replicates. (Right) Western blot analysis of the corresponding experiment using whole-cell extracts. Tubulin and NXF1 served as loading controls.

(G) The solubility of MSL1, MOF, MSL3, and MLE is tested in nuclei isolated from S2 cells, with or without RNase A incubation (ch, chromatin fraction; np, nucleoplasmic fraction). See Figure S1.



(legend on next page)

and found that the SHAPE and PARS independently supported our roX2 structure model. One alternative secondary structure that is partially supported by SHAPE and PARS remodels R2H4 and R2H5, wherein the inverted roX box (IRB) of R2H5 forms a helix with the roX box of R2H4, leaving the roX box of R2H5 exposed (Figure S5).

Overall, our structure model for roX2 exon-3 reveals a similar architecture to that of roX1: stable helices that are formed by paired roX box motifs strung together by flexible single-stranded linkers. We aligned the structure models of roX1 and roX2 to the evolutionary conservation scores to identify which regions of the RNAs are conserved. Most stems (particularly those containing roX box motifs) are conserved throughout evolution, but not loops or linkers (Figures S4 and S5). Taken together this analysis revealed that both roX1 and roX2 contain conserved secondary structures scattered throughout the length of the RNA.

MLE and MSL2 Interact with Conserved and Structured Regions of roX1 and roX2

The strong enrichment of roX1 and roX2 RNAs in iCLIP analysis called for a detailed investigation of the similarities and differences between these RNA structures with respect to MLE and MSL2 binding.

roX1

We identified three domains of 200–500 nucleotides (D1–D3) in roX1 that are bound by both MLE and MSL2 with high iCLIP scores (Figure 4A). MSL2 binding profile on these regions appeared slightly broader compared to MLE (Figure 4A and Figures S3A and S3B). Among these, the third binding domain (D3, also used above in SHAPE and PARS analysis) is highly enriched for MLE binding and harbors most of the top peaks that map to roX1. Although this domain contains many conserved elements and is implicated in roX1 function, it remained unclear whether this region is a protein-docking site for the MSL complex members (Kelley et al., 2008; Stuckenholz et al., 2003). Furthermore, binding to the additional two domains was unexpected. Remarkably, bound regions constitute only approximately 25% of roX1 RNA (3.7 kb), suggesting that the rest of the RNA is flexible and not present in a defined structure that could be scored using the in vivo crosslinking method.

Since MLE binding appeared clustered and restricted to three domains of roX1 in vivo (Figure 4A), we next addressed whether we could also recapitulate this mode of binding in vitro to gain insight into the binding mechanism. For this purpose, we utilized GRNA chromatography (Czaplinski et al., 2005). This method relies on the high-affinity interaction between a 19 nt RNA element called boxB and a 22 amino acid long peptide called λ_{N22} (Figure 4C) and allows relatively long RNAs (up to ~1 kb) to be used as well as endogenously expressed proteins in nuclear extracts made from *Drosophila* embryos, thus placing technique in between in vivo approaches such as iCLIP and complete in vitro experiments such as electrophoretic mobility shift assays (EMSAs). We designed five roX1 RNA derivatives for chromatography: three regions that interact with MLE (D1, D2, and D3) and two regions that show no MLE binding (U1 and U2) in vivo. We observed exquisite binding specificity of MLE to these regions (D1–D3) in comparison to unbound regions (U1 and U2) or to GFP RNA in vitro (Figure 4D). The data show for the first time that the three regions in roX1 have the capacity to autonomously and specifically interact with MLE.

Since roX1 D3 resembles roX2 exon-3 both in length (450 nt versus 504 nt) and in domain architecture (RBL motif containing helical regions at the 5′ end and RB elements at the 3′ end) (Figure 4, also see below), we examined in more detail how this domain interacts with MLE. First we split this domain into two: the 5′ half containing stem-loops R1H1 and P2 and the 3′ half that contains three RB elements (see schematic representation in Figure 4B). Strikingly, we detected robust binding of MLE to the 5′ part of the RNA, but not to the 3′ part (Figure 4E, lanes 5–7). However, when the 3′ half is exten-ded to include P2 and the IRB element, thus allowing the formation of R1H2 (119–450, Figure 4E, lane 8) in addition to the RB elements, we started to regain MLE interaction with roX1. This binding is still weaker than what we observed with the full-length D3 RNA, which only has R1H1 in addition to this fragment, underscoring the importance of this structure in MLE binding. In fact, R1H1 alone (1–118) is sufficient to interact with MLE (Figure 4E lane 8), suggesting that this region makes stable contacts with MLE both in vitro and in vivo.

roX2

The *roX2* gene is much smaller than the *roX1* gene (1.2 kb versus 3.7 kb) and contains a large intron of ~500 bp separating a small first exon of ~70 bp and a large conserved exon of ~500 bp. The "intron" in roX2 undergoes extensive alternative splicing, and is thus called exon-2 rather than intron-1 (Park et al., 2005). However, in contrast to roX1, iCLIP identified only one domain within roX2 (the conserved exon-3) that showed extensive MLE binding and similar to roX1; this domain was also scored for MSL2 binding (Figure 5A).

In order to identify structured regions that are protected by MLE in roX2 RNA, we next carried out a small-scale high-throughput sequencing of RNAs isolated by crosslinking immunoprecipitation (HITS-CLIP) (Licatalosi et al., 2008; a method

Figure 2. iCLIP Reveals that MLE and MSL2 Interact with RNA In Vivo, roX1 and roX2 Being the Major Interactors

(A) MLE or MSL2 immunoprecipitated from nuclear extracts prepared from clone8 cells under stringent conditions, with or without UV-C crosslinking, and treated with low or high concentrations of RNaseI. Bound RNA is radioactively labeled and visualized by autoradiography. RNA molecules above the molecular weight of the protein of interest (red box) are isolated from nitrocellulose membranes in triplicates, cloned into a library, and sequenced.

(B) MLE and MSL2 libraries generated using the iCLIP approach. High-, medium-, and low-range libraries are prepared by gel fractionation of cDNA after reverse transcription and separate amplification with Illumina sequencing-compatible primers.

(C) MLE iCLIP detects 2,447 crosslinked nucleotides. The genomic distribution of crosslinked nucleotides scored by number of crosslinking events shows no particular bias toward any chromosome. roX1 (blue) and roX2 (red) nucleotides score considerably higher than most other nucleotides.

(D) MSL2 iCLIP detects 5,206 crosslinked nucleotides. Scores and distribution of roX1, roX2, and other nucleotides are similar to MLE.

(E) Distribution of crosslinking events. The majority of crosslinking events fall on roX1 and roX2 (Figure S2E). See Figures S2 and S3.

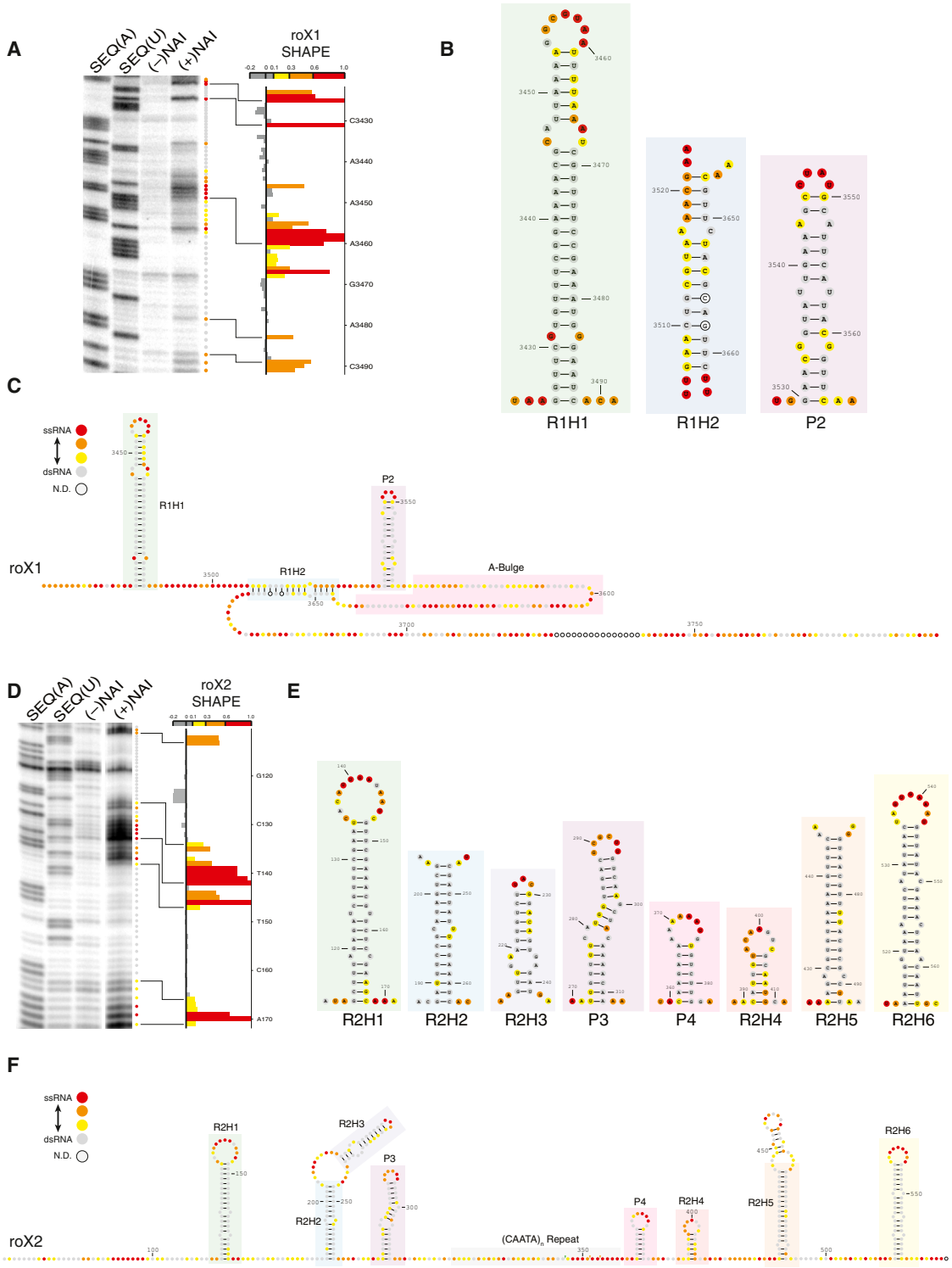


Figure 3. SHAPE Structural Analysis Reveals the Architecture of roX1 and roX2

(A) SHAPE analysis of the R1H1 region of roX1. Full-length roX1 RNA was in vitro transcribed and treated with or without SHAPE reagent, (± NAI). cDNAs were reverse transcribed from the modified RNA to identify modified, flexible bases. Sequencing lanes identify the base position + 1, and the (–) NAI lane identifies spurious reverse transcription stops that are modification independent. The SHAPE reactivity score for each base is calculated by integrating gel band intensities, subtracting background (–NAI), and normalizing to top peaks. Paired bases have low reactivity (gray), and flexible unpaired bases have high reactivity (increasing from yellow to orange to red).

(legend continued on next page)

which only scores for uninterrupted reverse transcriptase read-through cDNAs) experiment in S2 cells. The resulting data agree well with the iCLIP data, as all reads that mapped to roX2 fell on exon-3 (Figure S6A). Under closer inspection, four groups of reads emerge from this data, all of which map to roX2 helices. Interestingly, R2H1 and R2H2/3 account for almost all of the reads that map to roX2 exon-3 (Figure S6A). This observation suggests that this tandem stem-loop structure within the first hairpin cluster of roX2 is the likely tether that interacts with MLE.

iCLIP, HITS-CLIP, and structural analysis of roX RNAs suggest that MLE interacts with two distinct regions of roX2 RNA: the 5′ end of exon-3 that contains tandem stem-loops with RBL elements, and the 3′ end of exon-3 with its RB elements. To understand how these elements contribute to MLE binding, we split roX2 exon-3 into two fragments (Figure 5B; 1–280 nt, the first stem-loop cluster; and 281–504 nt, the second stem-loop cluster containing RB elements) and used them in GRNA chromatography, along with GFP RNA to control for background MLE binding. MLE indeed interacts very specifically with roX2 exon-3 in this assay (Figure 5C, compare lane 5 to lanes 3 and 4), but surprisingly only with the first hairpin cluster and not with the second (Figure 5C, lanes 5–7).

The first roX2 cluster contains four helical structures that are likely candidates for MLE binding. We next tested whether all of these helical regions contribute to MLE-roX interactions and designed point mutations that disrupt the stability of the stem regions of R2H1, R2H2, and P3. Again, consistent with the outcome of the HITS-CLIP experiment (Figure S6A), we observed that the helical region of R2H1 provides the most important contact in roX2 RNA, and destabilizing this stem leads to reduced binding of MLE to full-length exon-3 RNA (Figure 5D, compare lane 5 with lane 6). Interestingly, a double mutant of R2H1 and R2H2 leads to an even more severe loss of binding (Figure 5D compare lane 5 with lane 9), suggesting that these regions can interact with MLE individually but act cooperatively for a more stable interaction.

Taken together, a close inspection of MLE and MSL2 binding data on roX1 and roX2 RNAs revealed two important aspects: first, binding of these proteins is restricted to distinct domains of each RNA; and second, these domains are similar for both MLE and MSL2 (Figures 4 and 5). Furthermore, integrated analysis of roX RNA structure and iCLIP revealed that the roX box sequence motif adopts a common, repeated secondary structure that underlies MLE-roX binding and that both roX RNAs have architectural similarities on sequence and structural level that may explain their redundant nature.

MLE Binding Is Demarcated by Specific Motifs in roX RNAs

Both roX1 and roX2 contain multiple structural elements that seem to operate as targeting cues for MLE and the MSL complex in vivo. In order to define the sequences that are enriched in MLE targets, we first analyzed the top 1% iCLIP peaks of MLE in roX RNAs. This analysis revealed an extended version of the roX box motif, which we have called the roX box/roX box-like (RB/RBL) motif (Figure 5E). We identified three copies of the RBL motif in the first hairpin cluster of roX2 and a single copy in the stem region of R1H1 in roX1. Taken together, 11 of these motifs (RB and RBL) identified in roX1 and roX2 account for the strongest binding of MLE to these RNAs. This consensus sequence is only present in roX RNAs and not present otherwise in the *Drosophila* transcriptome underscoring a sex-specific function of roX RNAs.

MLE Double-Stranded RNA-Binding Domains Mediate Interaction with roX2

MLE has two N-terminal double-stranded RNA-binding domains (dsRBDs) (Figure 6A), and the results from our GRNA experiments show that MLE interacts mainly with two tandem helical structures (R2H1 and R2H2) in the first hairpin cluster of roX2 exon-3 (Figures 5C and 5D). Next we investigated if these domains are indeed able to interact with roX2 exon-3 in a completely in vitro experiment with recombinant proteins and RNA. To this end we expressed the first 254 amino acids of MLE (MLE^{1-254}; Figure 6B, lane 1), which includes dsRBD1, dsRBD2, and the intervening region between them and tested its ability to interact with an RNA probe consisting of the first three stem-loops of roX2 exon-3 (R2H1wt R2H2wt) using EMSA. MLE^{1-254} interacts specifically with this probe, consistent with our GRNA experiments (Figure 6C, lanes 1–5). We introduced three point mutations to this protein predicted to disrupt contacts with the RNA backbone, one in dsRBD1 (K4E) and two in dsRBD2 (H196E, R198E), creating $MLE^{1-254(KHRmut)}$ (Figure 6B, lane 2). This mutant, as predicted, no longer interacts with RNA (Figure 6C, lanes 6–9, right). We then mutated the stem-loops in our RNA probe, individually or at the same time, in order to gauge their contribution toward the observed interactions. In accord with the GRNA experiments, mutating single stem-loops did little to these interactions (Figure S6B), whereas mutating both resulted in an RNA that could no longer interact efficiently with MLE^{1-254} (Figure 6C, left, compare lanes 2–5 with lanes 7–10). To study whether individual dsRBDs would be able to recapitulate the interactions, we purified GST fusions of dsRBD1 and dsRBD2 (Figure 6B) and tested their ability to

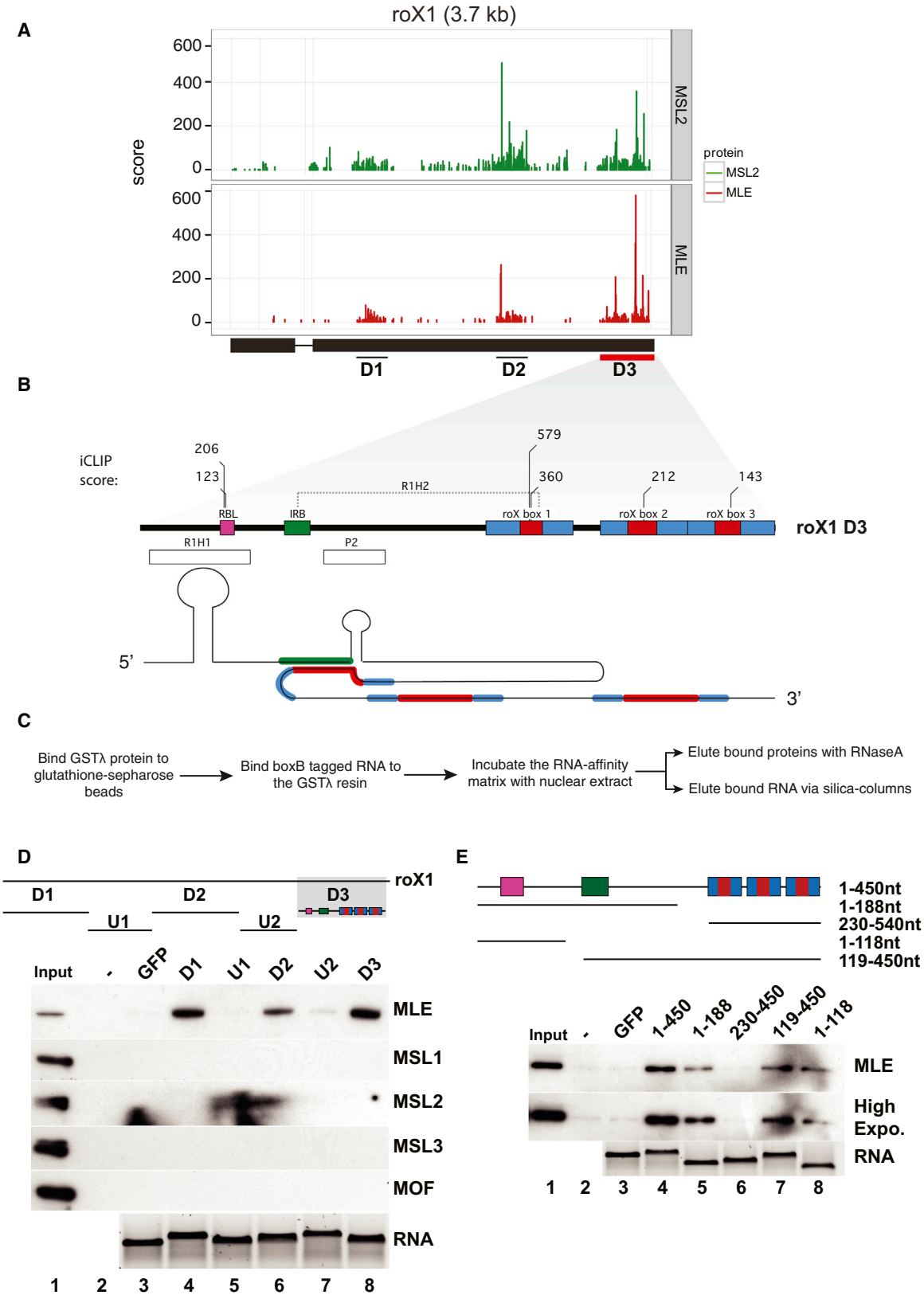
(B) SHAPE-derived structural domains of roX1. From the SHAPE reactivity profile for the 3′-terminal domain of roX1, a structure model is built. The R1H1 structure is derived from the SHAPE gel in (A). For each structure, reactive bases are present in bulges and the terminal loop, whereas the helix of this stem-loop is characterized by stretches of low reactivity.

(C) Structure model of the entire 3′ terminus of roX1. The 3′ terminus of roX1 is organized into three stable helices, two of which contain a core roX box motif (R1H1 and R1H2). The A-bulge likely participates in long-range base pairing with distal U-rich regions of roX1 that were not interrogated here. The very 3′ end lacks any well-defined secondary structure.

(D) The SHAPE profile of the R2H1 region of roX2. The R2H1 domain of roX2 exhibits light-heavy-light modification, characteristic of a stable stem-loop.

(E) SHAPE-derived structural domains of roX2. SHAPE of roX2 exon-3 resolved many stable secondary structures. Three of the helices contain roX box-like motif (R2H1-3), and three contain the roX box motif (R2H4-6).

(F) Structure model for roX2 exon-3. roX2 exon-3 is arranged into stable structural domains that are linked together by a flexible backbone. See Figures S4 and S5.



(legend on next page)

interact with the identical RNA probes that were used for MLE$^{1-254}$. Albeit much weaker than MLE$^{1-254}$, both dsRBD1 and dsRBD2 have the capacity to interact with the wild-type RNA probe (dsRBD1 < dsRBD2) (Figure S6C). We also tested dsRBD1 and dsRBD2 against the helix-destabilized RNA probes. Surprisingly, unlike MLE$^{1-254}$ (Figure S6B), dsRBD2's interaction with the mutant RNA probes was severely affected when the first helix is destabilized, suggesting that this domain prefers certain helical structures to others (Figure S6D, compare lanes 2–5 [left] with lanes 2–5 [right]). In contrast, the weak interaction of dsRBD1 with the wild-type or mutant RNA probes remained unaffected, indicating that the scored interactions are likely not specific to dsRNA (Figure S6E). Finally, we tested if MLE$^{1-254}$ and dsRBD2 interact with the 5′ and 3′ ends of roX2 RNA with equal efficiency. This was indeed the case (Figures S6F and S6G). Taken together, our EMSAs show that the N terminus of MLE (MLE$^{1-254}$) forms a robust dsRNA binding domain which interacts with roX2 RNA more efficiently than either dsRBD alone, suggesting cooperativity between dsRBD1 and dsRBD2.

ATP-Independent and -Dependent Binding of MLE to roX2

Although iCLIP clearly detects MLE binding on both the first and second stem-loop clusters of roX2 in vivo (Figures 5A and 5B), stable binding is observed within the first stem-loop cluster using GRNA chromatography (Figure 5C). MLE binding could be further narrowed down to the first two stem-loops in that cluster (R2H1 and R2H2) (Figure 5D). Since MLE is a robust RNA/DNA helicase (Lee et al., 1997), we next addressed whether this difference could be accounted for by the fact that the GRNA chromatography experiments were performed in the absence of ATP. We therefore repeated the experiment, comparing differences in binding in the presence or absence of ATP (Figure 6D). Interestingly, we observed that the MLE binding to the first stem-loop cluster of roX2 (Figure 6D, 1–280, lanes 2 and 4) is sensitive to ATP and is slightly reduced (same is observed for the nonspecific RNA control GFP and full-length roX2 exon-3, Figure S7A). However, in the presence of ATP, MLE binding to the second stem-loop cluster dramatically increases and can be visualized now in vitro (Figure 6D, 281–504, compare lanes 3 and 5).

In order to further characterize MLE's ATP-dependent and independent interaction with roX2, we next utilized GRNA chromatography and nuclear extracts prepared from S2 cells transfected with three different MLE derivatives (Figure 6E and Figure S7B). We used a wild-type MLE construct (MLE^{wt}), a well-characterized mutant of MLE that cannot bind ATP efficiently due to a point mutation at the conserved ATP-binding motif I (K413E or MLE^{GET}, Lee et al., 1997) and a mutant that is incapable of interacting with dsRNA (MLE^{KHR}) in vitro due to three point mutations introduced into MLE's two N-terminal dsRBDs (Figure 6C). We can detect the transfected transgenic MLE construct alongside the endogenous MLE that serves as a positive control in each experiment because of the size shift imparted by the affinity tag (Figure 6E, left). The transgenic proteins were also detected using an antibody against an epitope that is only found in the transfected constructs (RGS, Figure 6E, right). The RNAs eluted from the beads were equal in all the GRNA experiments, indicating that differences in MLE binding are not a result of differential RNA loading or stability in extracts (Figure S7C). These results clearly show that the wild-type MLE construct behaves identically to the endogenously expressed MLE protein with respect to its interaction with the 5′ and 3′ ends of roX2 exon-3 in the presence or absence of ATP (Figure 6E, MLE^{wt}). Importantly, disrupting MLE's ATP-binding pocket greatly reduces MLE's interaction with the 3′ end of roX2 exon-3 in the presence of ATP (Figure 6E, compare MLE^{wt} with MLE^{GET}, lanes 4–7). Furthermore, the MLE^{KHR} no longer interacts with either the 5′ end or the 3′ end of roX2 with or without ATP, showing that the interactions scored in GRNA experiments are mediated by MLE's N-terminal dsRBDs.

Only One Species of roX RNAs Can Be Detected per MSL Complex

An interesting question is whether roX1 and roX2 can be present within the same MSL complex, since they are redundant in function and localize to the same loci at the level of polytene chromosome stainings (Meller et al., 2000). To directly address this issue, we utilized the chromatin isolation by RNA purification (ChIRP) technique (Chu et al., 2011). We used DNA probes antisense to either roX1 or roX2 (and to LacZ as a negative control) and purified the target RNA and associated biomolecules from

Figure 4. MLE and MSL2 Bind to Three Clustered Regions within roX1 RNA

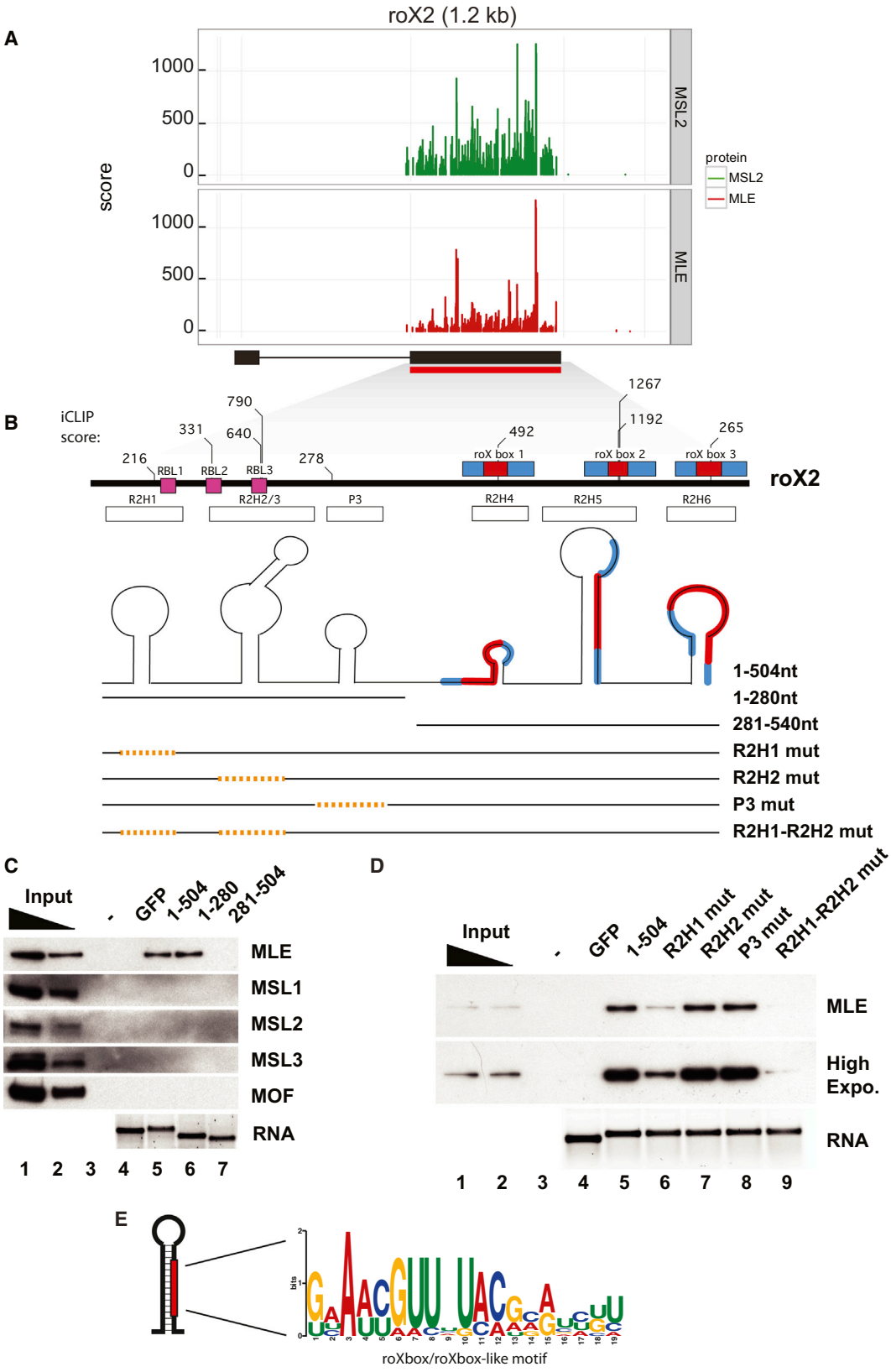
(A) The iCLIP data show that MLE and MSL2 interact with three domains (D1–D3) of roX1. The red box marks roX1 D3, which shows the highest score of MLE binding. See also Figure S2.

(B) A cartoon representation of roX1 RNA domain-3 (D3). roX1 region-3 is similar to roX2 exon-3 in its arrangement of stem-loops and roX boxes (see Figure 5B). R1H1 contains an RBL element in its stem (pink box) and is followed by R1H2, which is formed by a long-range interaction between the inverted roX box (IRB, green box) element and roX1 box1 (RB1, red-in-blue box). Another stem-loop (P2) is predicted to form in the bulge separating IRB and roX1 box1. Some of MLE's top iCLIP scores are indicated on top.

(C) The flow of a GRNA chromatography experiment.

(D) (Top) A cartoon representation of roX1 RNA delineated by MLE/MSL2 binding shows the constructs used in GRNA experiments. Colored boxes represent the elements described in (B). All RNA fragments used were of similar length (428–585 nt). (Bottom) GRNA chromatography shows that MLE interacts with D1, D2, and D3 (lanes 4, 6, and 8, respectively) and not with regions U1 and U2, GFP RNA, or beads (lanes 5, 7, 3, and 2, respectively). Antibodies used for immunoblotting are indicated on the right.

(E) (Top) The full-length roX1 region3 (D3) is split into four fragments for GRNA experiments. Colored boxes represent the elements described in nucleotide is indicated as nt (B). (Bottom) Specific binding of MLE to D3 (1–450) (compare lane 2 [beads] and lane 3 [GFP] with lane 4) could be further reduced to first half of roX1 D3 (1–188) (lane 5). No significant binding was scored for the second half of D3 (230–450) (lanes 6). R1H1 (1–118) seems to be able to interact with MLE (lane 8); however, its deletion does not lead to a complete loss of MLE binding (119–450; lane 7), suggesting that P2 and/or R1H2 can also interact with MLE in this assay. RNA eluted from the beads was run on a 1.2% agarose gel and stained with SYBR Safe and is shown beneath the immunoblots.



(legend on next page)

clone8 cells. We observed that roX1 probes isolated roX1 RNA, but not roX2 RNA, while roX2 probes isolated roX2 RNA, but not roX1 RNA (Figure 6F). These data suggest that although roX1 and roX2 interact with the same proteins and localize to the same loci on the X chromosome, the two RNAs are likely not present in the same complexes.

Combinatorial Mutations in Both Stem-Loop Clusters in roX2 Lead to Severe Male Lethality

Combination of iCLIP and GRNA chromatography experiments suggest that the stem-loops at the 5′ and 3′ ends of roX2 are qualitatively different from each other with respect to their interaction with MLE (Figures 6D and 6E). We next addressed the functional importance of the different roX2 stem-loops in dosage compensation in vivo. Using the ΦC-31 system, we generated 12 transgenic flies expressing various roX2 RNA derivatives under the control of a UAS promoter and tested their ability to rescue male lethality in a mutant that lacks endogenous roX expression ($roX1^{SMC17A}roX2^{\Delta}$, also see Figure 7 for schematic representations of these mutants). In accordance with our biochemical observations, we operationally split roX2 exon-3 (1–504 nt) into two: the first half (1–280 nt) we called "A" and the second half (281–504 nt) "B." We introduced various helix-destabilization point mutations (see the Supplemental Information) in both halves of the RNA independently (A mutations, A1–A3; or B mutations, B1–B5, with the wild-type construct designated as A0B0). As expected, introduction of wild-type roX2 exon-3 (A0B0) led to rescue of male lethality in roX double mutants (Figures 7A–7C). Interestingly, when mutations were introduced into only the first stem-loop cluster (A1B0, A2B0, or A3B0), the transgenic RNAs were as efficient as the wild-type RNA in rescuing male lethality upon tubulin-GAL4-driven expression (Figure 7A, middle). Similarly, when single helices were disrupted at the 3′ end of roX2 exon-3, the transgenic RNAs were again as efficient as the wild-type construct in rescuing male lethality (A0B1, A0B2, and A0B3; Figure 7B, middle). However, destabilizing the last two stem-loops (A0B4, R2H5 and R2H6) significantly reduced the viability of male flies. Destabilization of R2H4, in addition to R2H5 and R2H6 (A0B5), did not enhance the lethality phenotype (Figure 7B, middle; compare A0B5 to A0B4). When mutations in the helical elements of both the first and the second stem-loop clusters are combined (Figure 7C), we observed a severe male lethality phenotype with very few escapers, especially when the last two stem-loops are mutated (A1B4, A2B4, and A3B4) (Figure 7C). Quantitative RT-PCR (qRT-PCR) analysis of the stability of these transgenic RNAs in wandering third-instar male larvae showed that isolated mutations at the 5′ end, or individual mutations at the 3′ end of roX2 exon-3, had little or no adverse effects on the stability of these RNAs (Figures 7A and 7B, right). However, mutating the last two stem-loops concurrently led to destabilization of the transgenic RNA (Figure 7B, right; A0B4 and A0B5), in line with the partially functional RNA leading to low rate of male survival. Interestingly, combining these destabilizing mutations at the 3′ end with mutations at the 5′ end once again results in transgenic roX2 exon-3 RNAs that are as stable as the wild-type roX2 exon-3 RNA, even though these RNAs are effectively not functional in dosage compensation, judged both by their inability to rescue lethality in roX double mutants (Figure 7C, middle) and their inability to support the targeting of the MSL1 and H4K16 acetylation to the X chromosome (Figure 7D).

Tubulin-GAL4-driven expression of the A0B0 construct (wild-type roX2 exon-3) results in approximately eight times higher roX2 RNA levels compared to wild-type roX2 levels in males (Figure S7D). To determine if this overexpression might have masked biologically important contributions of some of our mutants, we used a weaker but ubiquitously expressed GAL4 driver, daughterless-GAL4 (Figure S7E, left), which reduced the expression of our transgenes to approximately three times higher roX2 RNA levels (Figure S7E, right). At this level of expression, the wild-type construct (A0B0) rescued male lethality as efficiently (Figures S7F and S7G). However, isolated mutations at the first (A1BO) or at the second hairpin cluster (A0B1) now resulted in a marked decrease in male viability phenotype (Figures S7F and S7G). Moreover, combining these mutations led to a roX2 exon-3 construct that could only rescue ~1% of males.

Taken together, these experiments suggest that at higher expression levels the first half of roX2 is dispensable for dosage compensation. However, since only combinatorial mutations in "A" and "B" are incompatible with male viability, the two halves of the RNA appear to have partially overlapping yet distinct functions in vivo.

DISCUSSION

Identification of functional domains in lncRNAs is an important step toward understanding how they work in vivo. Here, we characterize roX1 and roX2 as the most significant RNAs associated with MLE and MSL2. roX1 and roX2 contain

Figure 5. MLE and MSL2 Bind Exclusively to roX2 Exon-3

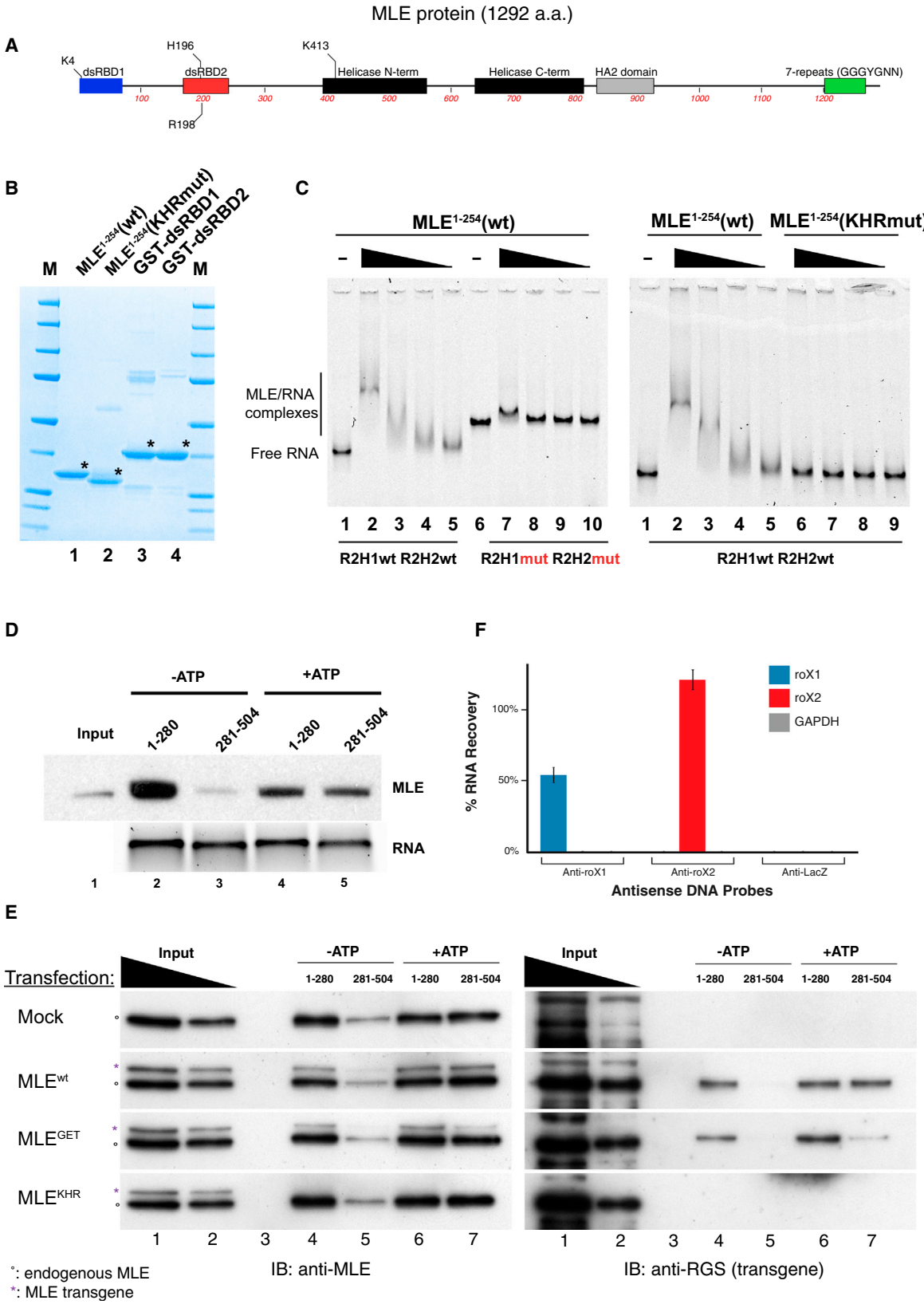
(A) MSL2 and MLE interact with the evolutionarily conserved third exon of roX2 in vivo. See also Figure S2.

(B) A cartoon representation of roX2 exon-3 shows that it has tandem helical regions at its 5′-end (R2H1, R2H2/3, and P3, white boxes), forming the first stem-loop cluster with RBL elements (pink boxes), and three roX-box elements at its 3′ end (RB1-3, red-in-blue boxes, indicated on top), which also resides in helical structures that form the second stem-loop cluster. MLE's top iCLIP scores are indicated on top. A schematic representation of fragments used in (C) and (D) is also shown.

(C) GRNA chromatography shows MLE interaction with the full-length roX2 exon-3 (1–504 nt, lane 5) and not with the beads (lane 3) or GFP RNA (lane 4). Based on SHAPE, iCLIP, and HITS-CLIP data, RNA was split into two halves. MLE binding was retained on the first stem-loop cluster (1–280, lane 6), and not with the second stem-loop cluster (281–504, lane 7). The proteins are probed with the indicated antibodies on the right.

(D) GRNA experiment, similar to (C), where the stem regions of R2H1, R2H2, or P3 are individually disrupted by point mutations in the context of the full-length exon-3 RNA in order to see if these are the regions that are responsible for MLE-roX2 interaction. High Expo. is higher exposure of the MLE blot. (C and D) RNA eluted from the beads was run on a 1.2% agarose gel and stained with SYBR Safe and is shown beneath the immunoblots.

(E) The extended roXbox/roXbox-like motif found at MLE crosslinking sites with very high scores. See Figure S6.



(legend on next page)

common, conserved, and distinct structural domains, which form the binding platform for these proteins. Interestingly, regions of lncRNAs that lie outside of these MLE/MSL2 interaction domains appear to be unstructured and not conserved. Our data also provide evidence on how roX1 and roX2 RNAs can be functionally redundant by showing that MLE-MSL2-interacting regions are present in multiple copies in both RNAs.

Similarities and Differences between roX1 and roX2

Earlier work on roX RNAs has identified short stretches of RNA that are shared between these lncRNAs that are conserved throughout evolution, called the roX boxes. We show that all three roX boxes at the 3′ end of roX2 RNA and one of the three roX boxes in roX1 form stable helices in vitro (Figures 3, 4, and 5). Moreover, these elements represent binding sites for MLE and MSL2 proteins in vivo (Figures 2, 4, and 5). Detailed analysis of MLE iCLIP data revealed that there are other RNA elements in both roX1 and roX2 that resemble the roX box sequence, and also serve as binding sites for MLE. We named these elements roX box-like sequences. Together roX box and roX box-like sequences uncover a consensus binding site for MLE in roX RNAs (Figure 5E). We could not identify other RNAs in the *Drosophila* transcriptome that contain the roX box/roX box-like motif other than the roX RNAs, adding to the evidence that it is likely that there are no more roX-like RNAs that can function in *Drosophila* dosage compensation.

Our results expose the logic of a roX RNA: stem-loops containing RBL elements at the 5′ end and RB containing helical structures joined by a flexible, single-stranded spacer region. In roX2, these elements are repeated on a very small scale with perfect copies of RBL/RB elements. On the other hand, in roX1, these sequence motifs deviate from the ideal consensus more than roX2, but roX1 probably compensates for this by containing multiple, autonomous interaction domains that form a much larger RNA, which is about six times as big as roX2.

MLE Interaction with roX2 RNA Appears to Be Dynamic

Endogenously expressed MLE and affinity-tagged MLE proteins have the ability to differentiate between the two stem-loop clusters in the two halves of roX2 exon-3 (1–280 versus 281–504 in Figure 5C and Figure 6E). Interestingly, addition of ATP led to the specific interaction of endogenously expressed MLE with the second stem-loop cluster of roX2 exon-3 that contains roX box-containing helical structures (Figure 6D). Furthermore, we observed that the N-terminal dsRBDs are not only required for the ATP-independent interactions with the 5′ end of roX2 exon-3 but are also required for the ATP-dependent interactions of MLE with the 3′ end of roX2 exon-3 (Figure 6E). A similar dichotomy was observed in HITS-CLIP and iCLIP experiments. In HITS-CLIP, our reads accumulated around the first stem-loop cluster of roX2 exon-3, concentrating on R2H1, whereas in iCLIP the balance was shifted toward the second cluster (Figure S6A).

MLE is a member of the RHA/DEAH family of RNA helicases, which can remodel RNA and RNPs (Jankowsky, 2011). Interestingly, biochemical work on the closely related DEAD box RNA helicases shows that these enzymes have varying degrees of affinity toward their RNA substrates during their ATPase cycles (Russell et al., 2013). Since the GRNA chromatography experiments reveal that the second half of roX2 gains MLE binding when supplemented with ATP in vitro (Figures 6E and 6F), it is possible that we are scoring for the quaternary complex between MLE, roX2, and ADP-Pi in these experiments and catching enzyme on its way to eventually remodel this part of the RNA (Figure 7E).

The ATP-independent binding of MLE to the first half of roX2 RNA via its N-terminal dsRBDs could be an initial regulatory step as roX RNAs constitute important tethers for MLE within the MSL complex. However, the in vivo analysis (Figures 7A–7C) revealed that at high expression levels this cluster "A" is dispensable for dosage compensation. The second hairpin cluster appears to recruit MLE through the helicase domain in a way that still requires N-terminal dsRBDs, thus providing a dynamic platform for rerecruitment and spreading of MLE along the roX RNA. The function of this cluster "B" can only be partially compensated by higher expression (Figure 7B), suggesting that it is functionally distinct from "A." We propose that the dynamic interaction of MLE with roX RNAs may ensure that different regions of roX RNAs are exposed such that they can be used in a redundant or cooperative manner for the interaction with the MSL complex members providing plasticity, as clearly detected for MSL2 binding on roX RNAs in the iCLIP data in vivo (Figures 4 and 5). However, we cannot exclude the

Figure 6. MLE's Double-Stranded RNA Binding Domains Interact with roX2 RNA

(A) Schematic representation of MLE protein domains. Residues mutated in later experiments are also indicated.

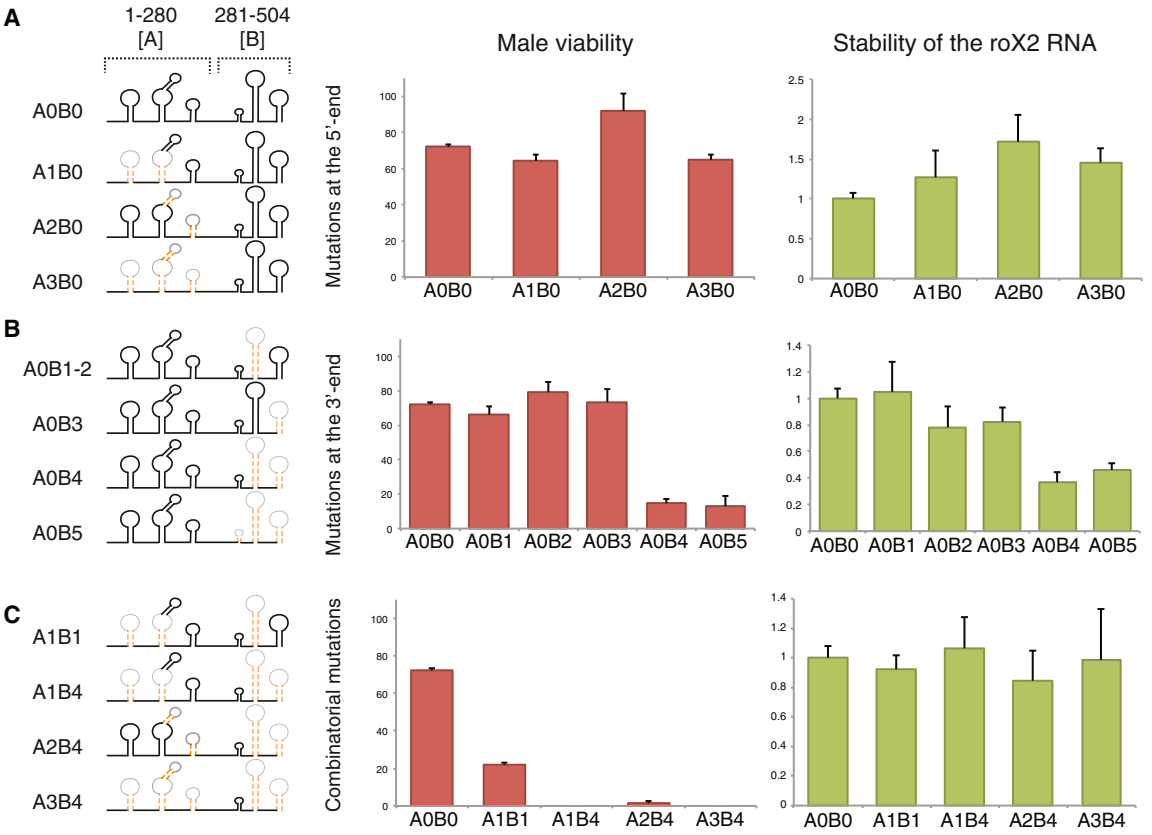
(B) MLE dsRBDs were expressed for EMSA. (Lane 1) His-tagged wild-type MLE^{1-254} (32 kDa) protein includes dsRBD 1 and dsRBD 2; (lane 2) a derivative expressing one point mutation in the dsRBD1 (K4E) and two in the dsRBD2 (H196E, R198E); (lanes 3 and 4) GST-tagged dsRBD1 and dsRBD2, respectively. M, molecular weight marker. Protein gel was stained with Coomassie blue. Asterisks (*) indicate the correct protein band.

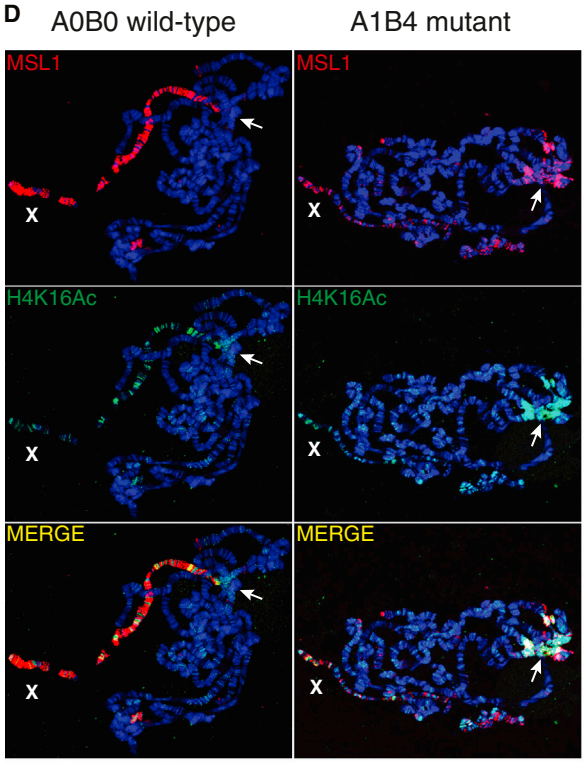
(C) MLE^{1-254} interacts with an RNA probe consisting of R2H1 and R2H2/3 regions of roX2 exon-3 (lanes 2–5). Point mutations that disrupt the stability of the stems in both R2H1 and R2H2 (lanes 7–10, left), or point mutations predicted to render dsRBD1 (K4E) and dsRBD2 (H196E, R198E; lanes 7–10, right) incapable of binding RNA, severely reduce the interaction between $MLE^{1-254(KHR\ mutant)}$ and R2H1wt R2H2/3wt RNA. For each protein derivative, 125 nM, 250 nM, 500 nM, or 1 μM, respectively, was titrated (black triangle). Free RNA probe is shown in lanes 1 and 6 (left gel) and lane 1 (right gel).

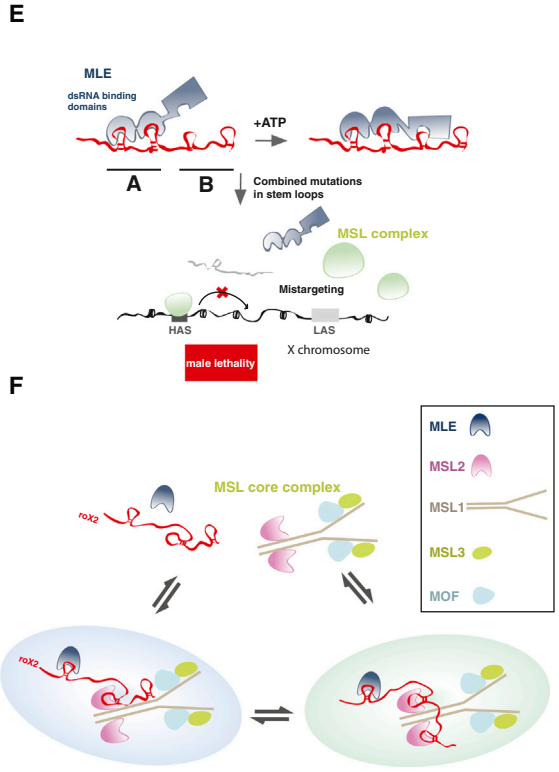
(D) The first (1–280) and the second (281–504) stem-loop clusters of roX2 RNA are used in a GRNA experiment as described in Figures 4 and 5. Here, ATP (3 mM) was added during the incubation of RNA with the nuclear extract (lanes 4 and 5).

(E) GRNA chromatography using transgenic MLE constructs that are either wild-type (MLE^{wt}), compromised in their ATP-binding pocket (MLE^{GET}), or dsRBDs (MLE^{KHR}). Immunoblots (IB) were probed either with an MLE antibody detecting endogenous (°) and ectopically expressed MLE derivatives (*) (left) or with an antibody that only detects the MLE derivatives (right). Lanes 1 and 2 show loading of 4% and 12% of input nuclear extracts.

(F) roX1 and roX2 do not directly interact in vivo. Antisense oligos to roX1, roX2, and LacZ (negative control) were used to isolate RNA-protein-chromatin complexes with the targeted RNA. roX1 ChIRP recovers no roX2 RNA, and roX2 recovers no roX1 RNA. LacZ ChIRP enriches for none of the target RNAs. RNA was quantified by qRT-PCR. Error bars represent ± SD of three independent measurements. See Figure S6.







(legend on next page)

possibility that integration of roX RNAs with the "core" MSL complex may also follow a more direct route involving independent MSL-RNA interactions that need not require MLE as a mediator.

roX RNAs as Targeting Sites and Assembly Platform for the MSL Complex

roX genes have a dual function in *Drosophila* dosage compensation: they are sites of roX transcription, but they also contain two HASs that can recruit the MSL complex independent of the roX RNAs (Kelley et al., 1999). Moreover, roX RNAs can travel from their sites of synthesis to the X chromosome when placed as a transgene to an autosomal site. This suggests that the holo-MSL complex, containing the core components (MSL1-3 and MOF) and MLE together with roX RNAs, can form on chromatin at roX transcription sites and spread in *cis* on the X chromosome, but it may also form in solution and be targeted to X-chromosomal sites in *trans*. We have recently shown that MSL2 interacts with a dimer of MSL1 and is itself present as a dimer within the MSL complex (Hallacli et al., 2012). Interestingly, we find here that roX RNAs do not interact with each other in vivo, suggesting that there could be one roX RNA species per holo-MSL complex (Figure 6F and Figure 7F).

Notably, iCLIP methodology utilizing UV crosslinking provides us with a snapshot of a pool of interactions that are present at the instant of irradiation. Although iCLIP data clearly show that MLE and MSL2 bind to the same domains, it is possible that they occupy different stem-loops on different molecules of roX RNAs rather than occupying the same structure at the same time. roX RNAs, by evolving multiple interaction platforms, can indeed support such combinatorial binding events, thus facilitating spreading along the X chromosome. We propose that roX RNAs, by virtue of being able to interact with MLE and the MSL complex, play a central role in the assembly of the holo-MSL complex containing the "core" (MSL1, MSL2, MSL3, MOF) and MLE. These complexes may form in solution or on chromatin such as on sites of roX transcription. Such configuration thus brings different enzymatic activities together (ATPase/helicase activity of MLE; acetyltransferase activity of MOF and ubiquitin-ligase activity of MSL2) for the X chromosome-specific dosage compensation process (Figure 7F). Directing the assembly of such a complex and by providing plasticity through its multiple stem-loops, roX RNAs could thus facilitate local spreading of the MSL complex from HAS into neighboring chromatin including low-affinity sites (LASs).

lncRNAs Provide Binding Specificity and Plasticity for Epigenetic Regulation

Domains identified in roX1 and roX2 could be defined as "information units" within these lncRNAs, which organizationally resemble stem-loops strung together like "beads on a string" (Dreyfuss et al., 1984). It is tempting to hypothesize that one could compare the "information unit" within roX lncRNAs to be the counterpart of codons which provide the basic information units in mRNAs. For roX lncRNAs, the information unit is a bit larger (11–15 nt) and involves both primary sequence and secondary structure. Another important aspect is that the domains are repeated in the lncRNA but do not appear to have a strict requirement of spacing or order relative to each other (i.e., no polarity and reading frame), so functional versions may be relatively easy to evolve during evolution.

In summary, our data suggest that lncRNAs may utilize discrete repetitive motifs for distinct protein-RNA or even RNA-DNA interactions to achieve functional specificity in vivo. Such a mechanism may also be used for other global epigenetic phenomenon such as the X-inactivation in mammalian cells (Augui et al., 2011; Jeon et al., 2012). Future analysis of other lncRNAs that combine CLIP methods and RNA structural analysis will be crucial in identifying structural domains within these RNAs and understanding their precise roles in the regulation of gene expression.

EXPERIMENTAL PROCEDURES

Detailed experimental procedures are available in the Supplemental Information.

ChIP Analysis

ChIP analysis was performed as described in Lam et al. (2012). MLE antibody used was described in Mendjan et al. (2006). For details, see the Supplemental Information.

Figure 7. Cooperative Interaction of Tandem Stem-Loops Is Important for roX Function In Vivo

(A–C) Schematic representation of wild-type and mutant roX2 exon-3 constructs are depicted (left). "A" and "B" refer to the first and second stem-loop clusters. Mutations introduced to the helical regions are indicated by the dashed orange line. (A) Helix-destabilizing mutations introduced to the first stem-loop cluster (A1B0, A2B0, and A3B0) rescue male lethality (middle, dark orange) and do not affect the stability of these RNAs (right, green). (B) Mutations in individual helices of the second stem-loop cluster also rescue male lethality (A0B1, A0B2, and A0B3). When the two terminal helices are disrupted at the same time (A0B4 and A0B5), male viability and RNA stability are reduced (middle and right). (C) However, combinatorial mutations affecting both the first and second half of roX2 exon-3 (A1B1, A1B4, A2B4, and A3B4) severely affect male viability (middle), without affecting RNA stability in vivo (right). (A–C) Error bars for male viability data represent ± SEM and for expression analysis ± SD of at least three biological replicates.

(D) Polytene chromosomal immunostainings of male third-instar larvae show that the wild-type roX2 exon-3 construct (left) efficiently targets the MSL complex (MSL1, red) and H4K16 acetylation (green) to the X chromosome (X) in males. However, the A1B4 mutant shows a severe reduction in the number of MSL1-bound target sites on the X chromosome, and mistargeting to the chromocenter (arrow) is observed (right). DNA is stained with Hoechst 33342 (blue).

(E) A summary of MLE-roX interactions described in this study. MLE engages with roX2 RNA stem-loops in an ATP-independent and -dependent manner. Tandem stem-loops in roX2 play an important role in dosage compensation as combinatorial mutations lead to severe male-specific lethality due to defective targeting of the MSL complex to the X chromosome.

(F) A summary model using roX2 RNA as an example. We propose that roX RNAs play an important role in integration of MLE with the core MSL complex (MSL1–3, MOF). The dynamic interaction of MLE with different regions of roX RNAs in the absence or presence of ATP could provide a means to compose holo-MSL complexes where multiple molecules of either MLE or MSL core complex could be assembled. Such "heterogenous" assemblies could contribute to efficient targeting and also have the potential to mediate spreading of the MSL complex, enhancing the possibility of short- and long-range interactions on the X chromosome. See Figure S7.

iCLIP

iCLIP was performed essentially as described in König et al. (2010). For details, see the Supplemental Information.

GRNA Chromatography

GRNA protocol is adapted from Czaplinski et al. (2005) and Duncan et al. (2006). For details, see the Supplemental Information.

SHAPE and PARS Analysis

Probing roX1 and roX2 RNA using SHAPE analysis was performed as described in Wilkinson et al. (2006). PARS reactions, library construction, sequencing, and data processing were performed as in Kertesz et al. (2010). For details, see the Supplemental Information.

Fly Culture and Genetics

All transgenic lines were generated through the phiC31 integrase-mediated germline transformation. For details, see the Supplemental Information.

ACCESSION NUMBERS

MLE ChIP seq data are available in the ArrayExpress database under accession number E-MTAB-1705, the MLE HTS CLIP data under accession number E-MTAB-1706, and the MLE and MSL2 iCLIP data under accession number E-MTAB-1669.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, Supplemental Experimental Procedures, and Supplemental References.

ACKNOWLEDGMENTS

We thank K. Duncan, J. Medenbach, and M. Hentze for GRNA chromatography protocol; J. König and J. Ule for the iCLIP protocol; J. Kadlec and G. Pitchai for the MLE dsRBD expression plasmids; E. Izaurralde for the NXF1 antibody; P. Crisalli and E. Kool for the NAI; and V. Meller for the roX null mutant flies. We thank I. De La Rosa for help with deep sequencing. We thank Akhtar lab members for helpful discussions, especially T. Aktas for help with the artwork. We apologize to those whose papers were not cited due to space limitations. This work was supported by EU-funded NOE "EpiGeneSys," awarded to A.A. and N.L.; DFG-SFB 992/1 to A.A. and R.B.; DFG-BA 2168/4-2 to R.B.; and NIH R01-HG004361 (to H.Y.C.). H.Y.C. is an Early Career Scientist of the Howard Hughes Medical Institute.

Received: February 18, 2013
Revised: April 28, 2013
Accepted: June 20, 2013
Published: July 25, 2013

REFERENCES

Akhtar, A., Zink, D., and Becker, P.B. (2000). Chromodomains are protein-RNA interaction modules. Nature *407*, 405–409.

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

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.

Chu, C., Qu, K., Zhong, F.L., Artandi, S.E., and Chang, H.Y. (2011). Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol. Cell *44*, 667–678.

Conrad, T., and Akhtar, A. (2011). Dosage compensation in Drosophila melanogaster: epigenetic fine-tuning of chromosome-wide transcription. Nat. Rev. Genet. *13*, 123–134.

Conrad, T., Cavalli, F.M., Vaquerizas, J.M., Luscombe, N.M., and Akhtar, A. (2012). Drosophila dosage compensation involves enhanced Pol II recruitment to male X-linked promoters. Science *337*, 742–746.

Czaplinski, K., Köcher, T., Schelder, M., Segref, A., Wilm, M., and Mattaj, I.W. (2005). Identification of 40LoVe, a *Xenopus* hnRNP D family protein involved in localizing a TGF-β-related mRNA during oogenesis. Dev. Cell *8*, 505–515.

Dreyfuss, G., Choi, Y.D., and Adam, S.A. (1984). Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. Mol. Cell. Biol. *4*, 1104–1114.

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.

Fauth, T., Müller-Planitz, F., König, 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.

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.

Gilfillan, G.D., Dahlsveen, I.K., and Becker, P.B. (2004). Lifting a chromosome: dosage compensation in Drosophila melanogaster. FEBS Lett. *567*, 8–14.

Hallacli, E., Lipp, M., Georgiev, P., Spielman, C., Cusack, S., Akhtar, A., and Kadlec, J. (2012). Msl1-mediated dimerization of the dosage compensation complex is essential for male X-chromosome regulation in Drosophila. Mol Cell. *48*, 587–600.

Ilik, I., and Akhtar, A. (2009). roX RNAs: non-coding regulators of the male X chromosome in flies. RNA Biol. *6*, 113–121.

Izzo, A., Regnard, C., Morales, V., Kremmer, E., and Becker, P.B. (2008). Structure-function analysis of the RNA helicase maleless. Nucleic Acids Res. *36*, 950–962.

Jankowsky, E. (2011). RNA helicases at work: binding and rearranging. Trends Biochem. Sci. *36*, 19–29.

Jeon, Y., Sarma, K., and Lee, J.T. (2012). New and Xisting regulatory mechanisms of X chromosome inactivation. Curr. Opin. Genet. Dev. *22*, 62–71.

Kadlec, J., Hallacli, E., Lipp, M., Holz, H., Sanchez-Weatherby, J., Cusack, S., and Akhtar, A. (2011). Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1. Nat. Struct. Mol. Biol. *18*, 142–149.

Kelley, R.L., Meller, V.H., Gordadze, P.R., Roman, G., Davis, R.L., and Kuroda, M.I. (1999). Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin. Cell *98*, 513–522.

Kelley, R.L., Lee, O.K., and Shim, Y.K. (2008). Transcription rate of noncoding roX1 RNA controls local spreading of the Drosophila MSL chromatin remodeling complex. Mech. Dev. *125*, 1009–1019.

Kertesz, M., Wan, Y., Mazor, E., Rinn, J.L., Nutter, R.C., Chang, H.Y., and Segal, E. (2010). Genome-wide measurement of RNA secondary structure in yeast. Nature *467*, 103–107.

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.

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

Kuroda, M.I., Kernan, M.J., Kreber, R., Ganetzky, B., and Baker, B.S. (1991). The maleless protein associates with the X chromosome to regulate dosage compensation in *Drosophila*. Cell *66*, 935–947.

Lam, K.C., Mühlpfordt, F., Vaquerizas, J.M., Raja, S.J., Holz, H., Luscombe, N.M., Manke, T., and Akhtar, A. (2012). The NSL complex regulates housekeeping genes in Drosophila. PLoS Genet. *8*, e1002736. http://dx.doi.org/10.1371/journal.pgen.

Larschan, E., Bishop, E.P., Kharchenko, P.V., Core, L.J., Lis, J.T., Park, P.J., and Kuroda, M.I. (2011). X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila. Nature *471*, 115–118.

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.

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

Maenner, S., Müller, M., and Becker, P.B. (2012). Roles of long, non-coding RNA in chromosome-wide transcription regulation: lessons from two dosage compensation systems. Biochimie *94*, 1490–1498.

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.

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.

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.

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.

Park, S.W., Kang, Y.Ie., 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.

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., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. *81*, 145–166.

Russell, R., Jarmoskaite, I., and Lambowitz, A.M. (2013). Toward a molecular understanding of RNA remodeling by DEAD-box proteins. RNA Biol. *10*, 44–55.

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.

Stuckenholz, C., Meller, V.H., and Kuroda, M.I. (2003). Functional redundancy within roX1, a noncoding RNA involved in dosage compensation in Drosophila melanogaster. Genetics *164*, 1003–1014.

Wilkinson, K.A., Merino, E.J., and Weeks, K.M. (2006). Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. Nat. Protoc. *1*, 1610–1616.

Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat. Genet. *30*, 167–174.

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.