

Minireview

Multiple RNA-protein interactions in *Drosophila* dosage compensation

Hubert Amrein

Address: Department of Genetics, Duke University Medical Center, Research Drive, Durham, NC 27710, USA. E-mail: hoa1@duke.edu

Published: 6 December 2000

Genome **Biology** 2000, **1**(6):reviews1030.1–1030.5

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2000/1/6/reviews/1030>

© Genome**Biology**.com (Print ISSN 1465-6906; Online ISSN 1465-6914)

Abstract

From worms to humans, recognizing and modifying a specific chromosome is essential for dosage compensation, the mechanism by which equal X-linked gene expression in males and females is achieved. Recent molecular genetic and biochemical studies have provided new insights into how regulatory factors in *Drosophila* are recruited and assembled on the X chromosome, leading to the essential hypertranscription of its genes.

Distinct mechanisms of dosage compensation

Dosage compensation, which ensures that the expression of X-linked genes is equal in males and females, is an essential process in organisms that have sex chromosomes. Because sex chromosomes appeared relatively late in evolution, dosage compensation evolved independently in different organisms and is accomplished by distinct mechanisms. For example, in *Caenorhabditis elegans*, dosage compensation occurs in the homogametic hermaphrodite (XX) by the down-regulation of virtually all genes on the two X chromosomes by about 50%. In mammals, dosage compensation also occurs in the homogametic sex (the female) by X inactivation, whereby an entire X chromosome forms a distinct, heterochromatic, transcriptionally inactive nuclear structure known as the Barr body. Consequently, each gene on the single active X chromosome in female cells and the corresponding gene on the single X chromosome in male cells are expressed at equal levels. In *Drosophila*, dosage compensation is achieved in the heterogametic male by a twofold chromosome-wide up-regulation (hypertranscription) of essentially all genes on the single X chromosome. In *C. elegans* and *Drosophila*, several proteins with different molecular functions involved in dosage compensation have been identified. Much attention has been dedicated to the elusive question of how the compensated or inactivated chromosomes are recognized by these proteins, and three recent papers investigating some of these aspects in *Drosophila* have provided new insights into this mystery [1-3].

At least one common theme has emerged in *Drosophila* and mammals of how the dosage-compensated chromosomes are marked and eventually recognized by regulatory proteins: the use of non-coding RNAs transcribed from genes located on the X chromosome itself [4-7]. These RNAs, *Xist* (X-inactive specific transcript) in mammals, and *roX1* and *roX2* (RNA on the X) in *Drosophila*, are structurally unrelated, yet they share the intriguing property of remaining tightly associated with the X chromosome. In mammals, *Xist* RNA is transcribed only from the inactive X, with which it associates at its site of synthesis and then spreads over the entire chromosome through an unknown mechanism [8]. It is assumed that *Xist* RNA provides a mark for specific histones (for example, histone macroH2A1.2), as well as proteins deacetylating histones and methylating many of the X-linked genes at GpC dinucleotides [9-11]; all these events are important for maintaining the silenced state of the inactivated X.

The *Drosophila* dosage compensation complex

In *Drosophila*, the dosage compensation complex (DCC) is composed of five proteins encoded by the male-specific lethal genes, *male-specific lethal-1, -2, -3* (*msl1, msl2* and *msl3*), *maleless* (*mle*) and *males absent on the first* (*mof*), and at least two non-coding RNAs, *roX1* and *roX2* (for a review see [12]). The MSL proteins colocalize to hundreds of sites along the single male X chromosome; they all are essential for the hypertranscription of the X-chromosomal genes in males, as

male but not female animals die during development when they are mutant for any one of the five *msl* genes. Hypertranscription is a consequence of at least one chromatin modification: the acetylation of histone H4 at Lys16, which is thought to 'loosen' chromatin structure, thereby allowing the general transcription machinery easier access to the regulatory regions of most X-linked genes [13]. The role of the *roX* genes and *roX* RNAs in this process is still unclear. Localization studies of MSL proteins in male nuclei carrying autosomal *roX1* or *roX2* transgenes showed that the *roX* genes or RNAs recruit the entire set of MSL proteins to their transgenic location and can lead locally to acetylation of histone H4 [1,14]. Moreover, the DCCs can spread, by several hundred kilobases, into neighboring autosomal DNA. These experiments indicated that the *roX* genes might function as nuclear entry sites for the assembly of the MSL proteins on the X chromosome. Little was known, however, about the specific role of the *roX* RNAs during the formation of the DCC. The three recent papers [1-3] have now started to address the role of the *roX* genes in this process and investigated which MSL proteins interact with the *roX* RNAs.

Meller *et al.* [1] took advantage of the fact that dosage compensation can be initiated and analyzed in females ectopically expressing the male-specific MSL2 protein. These females hypertranscribe their two X chromosomes and therefore die during development, but they can be rescued if any other *msl* mutation is present. It has been shown previously, using a similar strategy, that the DCC is assembled in an ordered sequence (Figure 1). For example, MSL1 and MSL2 are the first proteins to bind to the X chromosome; in the absence of all other MSL proteins they associate with about 35 sites, the so-called chromatin entry sites. MSL1 and MSL2 are interdependent, however: they require each other for binding [15]. MLE appears to be the next protein to join the growing DCC as it is found at reduced levels at some of these 35 sites when the remaining two proteins, MSL3 and MOF, are absent [16]. These latter two proteins require all the other MSL proteins to be present to enable them to associate with the X chromosome. Importantly, all MSLs are required to generate the normal DCC distribution to the hundreds of sites throughout the entire X chromosome.

Meller *et al.* [1] have now analyzed and compared *roX* RNA and MSL protein distribution in females ectopically expressing MSL2 but lacking other MSL proteins. They found that *roX1* and *roX2* RNAs are associated with the X chromosome at different stages and sites during the assembly of the DCC. When the MLE protein, which encodes an RNA helicase, was absent, both *roX* RNAs were found only at their site of transcription but not in any other of the 35 entry sites where MSL1 and MSL2 were present, suggesting that neither of the *roX* RNAs can be integrated in a minimal MSL1-MSL2 complex. When MSL3 was absent in these females, however, *roX2* but not *roX1* RNA was found at the entry 35 sites. These findings indicate that *roX2* might be incorporated into

this partial DCC complex in an MLE-dependent manner and that *roX1* RNA is incorporated at a later stage, together with MSL3 and perhaps MOF.

Further evidence that at least *roX2* RNA is an integral part of the DCC comes from independent immunoprecipitation experiments from all three groups [1-3] using extracts from *Drosophila* S2 cells that express all the MSLs and *roX2* RNA. Surprisingly, Akthar *et al.* [2] found that DCC lacking MLE but containing the remaining MSLs still precipitated *roX2* RNA. The ordered assembly data from Meller *et al.* [1] as well as the fact that MLE is the only protein in the DCC featuring a known RNA-binding domain pointed towards the RNA helicase MLE as the primary candidate for a partner for the *roX2* RNA. Akthar *et al.* [2] turned their focus towards the MOF protein as a potential candidate for interaction with *roX2* RNA. MOF is the crucial component that links the DCC to the X-chromosome-specific acetylated form of histone H4. MOF encodes an 827 amino-acid protein containing an amino terminus of unknown function, a chromodomain, a zinc finger and a carboxy-terminal histone acetyltransferase (HAT) domain. The HAT domain of MOF is responsible for histone H4 acetylation, both *in vivo* and *in vitro*, as shown by Smith *et al.* [3]. Interestingly, immunoprecipitation assays using extracts from SL-2 cells transfected with wild-type or mutant *mof* genes indicated that the chromodomain is essential for the specific interaction between MOF and *roX2*, but that the amino-terminal region and HAT domain were not essential for this interaction.

Akthar *et al.* [2] then tested the putative RNA-binding property of MOF directly using electromobility shift assays. MOF appeared to bind to RNA rather non-specifically, but preferred RNA to DNA. In agreement with the immunoprecipitation results, they found that an intact chromodomain is essential for RNA binding, whereas the amino terminus or the HAT domain of MOF was not essential. Furthermore, they also showed that MSL3, which features two chromodomains on its own, also shows *roX* RNA binding properties. Although the specificity of these interactions remains to be investigated, the findings of Akthar and coworkers [2] provide a new possible role for chromodomains involved in RNA binding. Chromodomains are modules of about 50 amino acids found in a number of proteins from yeast to mammals; the function of these proteins vary, but they are most often associated with gene silencing and chromatin remodeling events; significantly, many chromodomain-containing proteins are associated on chromosomes with heterochromatin or heterochromatin-like regions (reviewed in [17]). The specific role of the chromodomain is unknown, but chromodomain swapping experiments in *Drosophila* suggest that they might be protein interaction modules [18]. Thus, the data by Akthar *et al.* [2] suggests a new and rather unexpected role for these modules in RNA binding. It is intriguing to speculate that the chromodomain of other remodeling proteins also exert their activity through RNA

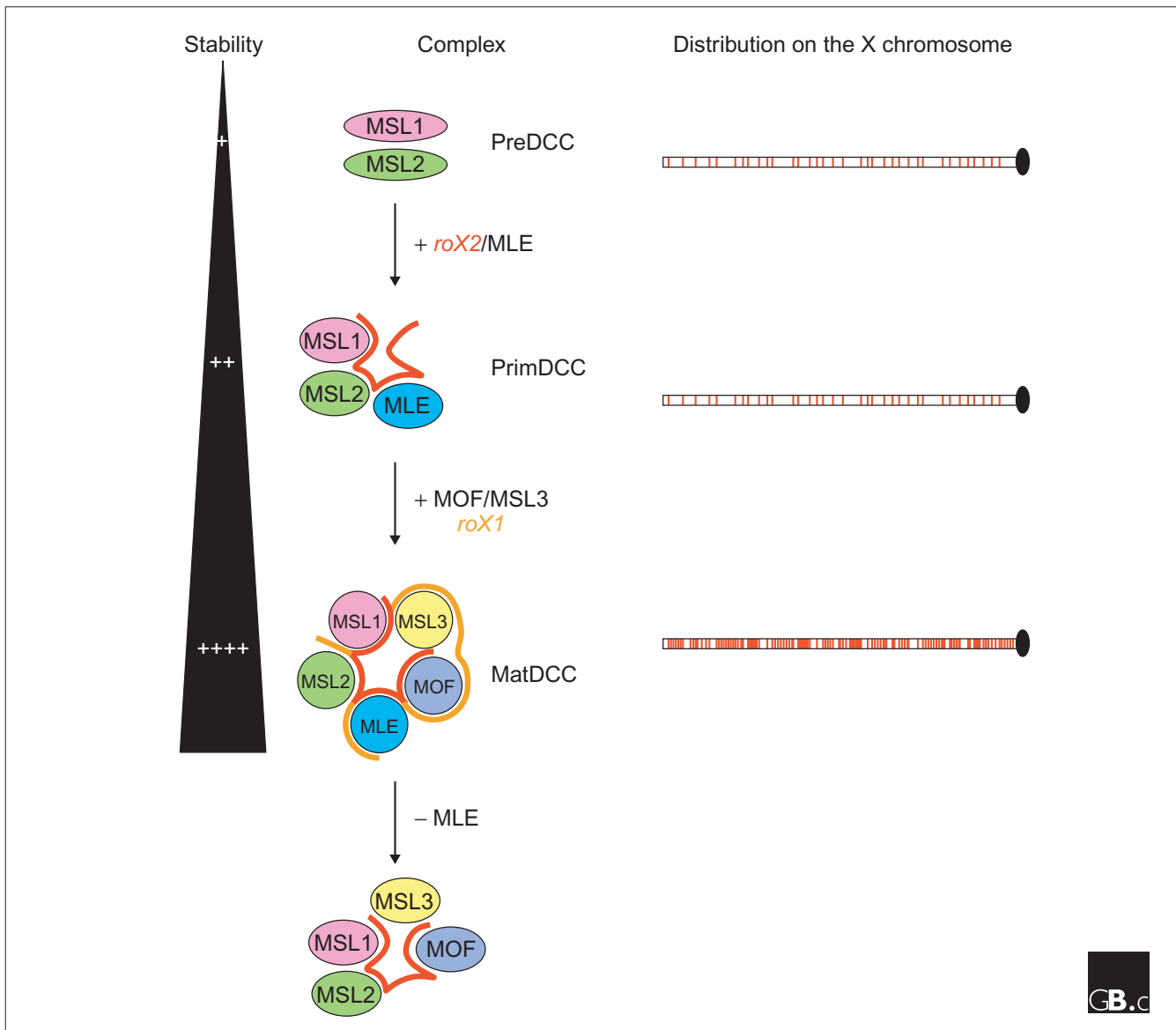


Figure 1

Ordered assembly of MSL proteins and *roX* RNA increases stability of the DCC. The first step of assembly involves the recognition of about 35 chromosomal entry sites by a preDCC consisting of MSL1 and MSL2. This recognition does not require any of the known *roX* RNAs. In the next step, *roX2* and MLE enter the complex to form a more stable primary DCC (PrimDCC) at these 35 sites. Spreading throughout the entire X chromosome requires the formation of the complete complex by addition of MOF and MSL3. The sequential addition of new components, particularly the *roX2* RNA, might induce changes in the structure of already incorporated components (illustrated by ovals becoming circles), increasing the stability of the DCC. MLE might be removed *in vitro* (for example, by elevated ionic strength) without destroying the entire complex, because of the stabilizing presence of MOF-MSL3. The role of *roX1* RNA is not clear, but it is integrated late in this process, together with or after MOF and MSL3; *roX1* might provide additional stability to the mature DCC (MatDCC). So, *roX1* and *roX2* might be partially redundant.

interactions and that RNAs might be more common than generally appreciated in transcriptional regulation.

Multiple RNA-protein interactions within the DCC

How can we reconcile the findings of Akthar *et al.* [2] and Meller *et al.* [1]? First, it might be relevant that the

co-immunoprecipitations and the *in vitro* binding experiments were performed with SL-2 cell extracts, cells in which dosage compensation appears not to be necessary; SL-2 cells, like other *Drosophila* cell lines, can become aneuploid without reduced viability. The soluble DCCs from these extracts are therefore not necessarily functional and might differ from those in males. Second, the interaction between *roX* RNA and MOF protein appears to lack specificity. The

lack of specificity might be attributed to a number of reasons, such as the absence of other MSL components, the presence of other RNAs interacting with MOF, or worse, it might reflect a property of MOF without functional significance: that is, MOF might not contact RNA at all *in vivo*.

Keeping this cautious note in mind, a more attractive alternative could be envisioned. The main proposition of this model is that *roX2* can interact with both MLE and MOF, and perhaps even other MSL proteins, as both *roX* RNAs are certainly large enough to accommodate binding sites for several proteins [6]. If such a scenario is correct, each of the interactions between different components of the DCC might be quite weak, and perhaps not even highly specific, but in the context of additional interactions becomes more stable. The addition of each new component might therefore strengthen existing interactions by optimizing contacts between binding surfaces through adjustments in the higher-order structure of the components (Figure 1). Evidence from co-immunoprecipitation experiments indicates that MSL1 and MSL2 interact directly with each other, a notion supported by the MSL1-MSL2 association at the 35 chromosomal entry sites in the absence of any other MSLs (the preDCC; Figure 1). The addition of both MLE and *roX2* RNA might allow the formation of a primary complex (the primDCC), which might be stabilized by the direct interaction of a structural motif of *roX2* RNA with the MLE protein, as well as interactions of this complex with MSL1 and MSL2. The primDCC might then be reinforced by the addition of the remaining two proteins, MSL3 and MOF, one or both of which might recognize other motifs within the *roX2* RNA, to form the mature complex (matDCC). Finally, matDCC might be further stabilized by the addition of *roX1* RNA, which could interact with several of the MSLs and perhaps *roX2* RNA as well. Significantly, assembly and disassembly of the complex might not necessarily follow the same order. For example, once the complex is formed, removal (for example, by high ionic strength) of a single component brought in relatively early (such as MLE) might be possible without affecting the rest of the complex.

One major task at hand now will be to resolve the specificities of the proposed interactions between the various MSL proteins and the *roX* RNAs. This could be done by carrying out detailed *in vitro* studies, which might not be easy to perform; many MSL proteins are difficult to express stably in bacteria or eukaryotic expression systems. An alternative might be to establish a heterologous *in vivo* system such as yeast, where protein-protein and RNA-protein interactions can be readily detected using sensitive reporter assays. A second important question is whether the *roX* genes have the same, overlapping or complementing functions. Although *roX1* mutant males are fully viable [7], it was suggested that *roX1* mutant males lacking a large number of additional genes, including *roX2*, have a lethal dosage-compensation phenotype because the DCC fails to assemble on the X chromosome [19]. Unfortunately, no *roX2* mutations have been recovered yet; they

would be important tools for validating the significance of the *roX* RNAs *in vivo*. Furthermore, *roX1* and *roX2* are only the first two of about 35 chromatin entry sites, although perhaps they are the most important ones. The nature of the remaining 33 sites is entirely unclear and there is, as yet, no evidence that they contain other *roX*-like RNAs. This brings us to the largest of all mysteries, namely how the DCC is spread along the X chromosome. One possibility is that the remaining 33 or so sites are 'stations' that serve as spreading facilitators in the form of DNA elements. Thus, primDCCs and/or mature (mat) DCCs, originating from the *roX* entry sites, might hop from one 'station' to the next and eventually reach all the entry sites distributed along the entire X chromosome [1]. From these sites, they then reach into the neighboring chromatin regions by yet another mechanism.

One final thought: the dosage compensation machinery in *C. elegans* acts on the two X chromosomes in the hermaphrodite by twofold down-regulation of gene expression. Six proteins are essential for this process and they associate with the X chromosomes; four of them are related to factors involved in chromosome condensation during mitosis in other systems. Not unlike in *Drosophila*, the DCC complex of *C. elegans* is also assembled in an ordered sequence, but no RNAs have been identified that might be involved in this assembly process [20]; either there will be yet another surprising turn in this multifaceted process or the *CroX* (*C. elegans roX*) RNAs wait to be discovered.

Acknowledgements

I thank Joe Nevins and Raquel Sitcheran for critically reading the manuscript. The author is supported by a Grant from N.I.H. (GM60234-01).

References

- Meller VH, Gordadze PR, Park Y, Chu X, Stuckenholz C, Kelley RL, Kuroda MI: **Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in *Drosophila***. *Curr Biol* 2000, **10**:136-43.
- Akhtar A, Zink D, Becker PB: **Chromodomains are protein-RNA interaction modules**. *Nature* 2000, **407**:405-408.
- Smith ER, Pannuti A, Gu W, Steurnagel A, Cook RG, Allis, CD, Lucchesi JC: **The *Drosophila* MSL complex acetylates histone H4 at lysine 16, a chromosome modification linked to dosage compensation**. *Mol Cell Biol* 2000, **20**:312-318.
- Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S: **The product of the mouse *Xist* gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus**. *Cell* 1992, **71**:515-526.
- Brown CJ, Hendrich BD, Rupert JL, Lafreniere RG, Xing Y, Lawrence J, Willard HF: **The human *XIST* gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus**. *Cell* 1992, **71**:527-542.
- Amrein H, Axel R: **Genes expressed in neurons of adult male *Drosophila***. *Cell* 1997, **88**:459-469.
- Meller VH, Wu KH, Roman G, Kuroda MI, Davis RL: ***roX1* RNA paints the X chromosome of male *Drosophila* and is regulated by the dosage compensation system**. *Cell* 1997, **88**:445-457.
- Clemson CM, McNeil JA, Willard HF, Lawrence JB: ***XIST* RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure**. *J Cell Biol* 1996, **132**:259-275.

9. Costanzi C, Pehrson JR: **Histone macroH2AI is concentrated in the inactive X chromosome of female mammals.** *Nature* 1998, **393**:599-601.
10. Jeppesen P, Turner BM: **The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression.** *Cell* 1993, **74**:281-289.
11. Boggs BA, Connors B, Sobel RE, Chinault AC, Allis CD: **Reduced levels of histone H3 acetylation on the inactive X chromosome in human females.** *Chromosoma* 1996, **105**:303-309.
12. Lucchesi JC: **Dosage compensation in flies and worms: the ups and downs of X-chromosome regulation.** *Curr Opin Genet Dev* 1998, **8**:179-184.
13. Turner BM, Birley AJ, Lavender J: **Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei.** *Cell* 1992, **69**:375-384.
14. Kelley RL, Meller VH, Gordadze PR, Roman G, Davis, RL, Kuroda MI: **Epigenetic spreading of the *Drosophila* dosage compensation complex from roX RNA genes into flanking chromatin.** *Cell* 1999, **98**:513-522.
15. Lyman LM, Copps K, Rastelli L, Kelley RL, Kuroda MI: ***Drosophila* male-specific lethal-2 protein: structure/function analysis and dependence on MSL-I for chromosome association.** *Genetics* 1997, **147**:1743-1753.
16. Gu W, Szauter P, Lucchesi JC: **Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of *Drosophila melanogaster*.** *Dev Genet* 1998, **22**:56-64.
17. Jones DO, Cowell IG, Singh PB: **Mammalian chromodomain proteins: their role in genome organisation and expression.** *Bioessays* 2000, **22**:124-137.
18. Platero JS, Hartnett T, Eissenberg JC: **Functional analysis of the chromo domain of HPI.** *EMBO J* 1995, **14**:3977-3986.
19. Franke A, Baker BS: **The roX1 and roX2 RNAs are essential components of the compensasome, which mediates dosage compensation in *Drosophila*.** *Mol Cell* 1999, **4**:117-122.
20. Meyer BJ: **Sex in the worm: counting and compensating X-chromosome dose.** *Trends Genet* 2000, **16**:247-253.