

**Characterization of MLE RNA helicase,  
a subunit of the Dosage Compensation Complex (DCC)  
in *Drosophila melanogaster***

Dissertation der Fakultät für Biologie  
der Ludwig-Maximilians-Universität München

Vorgelegt von  
**Annalisa Izzo**  
Aus Varese, Italy

April 2006

Erklärung:

Ich erkläre hiermit, daß ich die vorliegende Dissertation selbst verfasst und mich dabei keiner anderen Mittel, als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Desweiteren erkläre ich hiermit, daß ich an keiner anderen Stelle ein Prüfungsverfahren beantragt, beziehungsweise die Dissertation in dieser oder anderer Form bereits anderweitig als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe.

München, February 2006

Erstgutachter: Prof. Dr. Peter Becker

Zweitgutachter: Prof. Dr. Charles David

Mündliche Prüfung am: 25.04.06

*„ Do not go where the path may lead;  
go instead where there is no path  
and leave a trail... “*

R. W. Emerson

## *Acknowledgements*

*First of all I would like to thank Peter for giving me the possibility to do my PhD in his lab and for providing an environment of high scientific quality.*

*I would also like to thank Cat, Rogi and Vio for the long discussions about my project and not only...and in particular Cat for helping me with the thesis and with my work in these last months. I learn a lot from you and I would never be grateful enough for your advices and for being always present in the nice and difficult moments. You have become very good friends to me and I enjoy a lot also the time we spent together outside of the lab.*

*I am also very grateful to all the people in the lab and in particular to the “three musketeers” for the scientific and human support and for the fun we had together. I appreciate you a lot as scientists and as persons and I think the lab will never be the same without you.*

*Special thanks go to my friends in Italy, Sara, Filippo and Stella because beside the distance nothing really changed. Thank you for visiting me and for the long telephone calls in the many rainy days of Munich.*

*Of course I cannot forget my Robertino! Thank you Robi for the hard decision you took to come here, for your support and patience and I apologize for all the time I came late at home.*

*Last but not least I would like to thank my family for accepting the choice I made to come here. I miss you a lot, but wherever I will go you will always be my reference point.*

*I apologize for all the people that I did not mention by name, but that help me in many ways during my time here in Munich.*

## Table of contents

<b>1. Zusammenfassung.....</b>	<b>1</b>
<b>2. Summary .....</b>	<b>3</b>
<b><i>I. Introduction .....</i></b>	<b>5</b>
<b>1. Chromatin .....</b>	<b>5</b>
1.1. Chromatin Dynamics.....	7
1.2. Post-translational modifications of histones.....	8
<b>2. Dosage compensation.....</b>	<b>14</b>
2.1. Dosage compensation in mammals.....	15
2.2. Dosage compensation in worms.....	16
2.3. Dosage compensation in flies.....	17
<b>3. DNA and RNA helicases.....</b>	<b>31</b>
3.1. Helicases and their functions.....	31
3.2. DNA and RNA helicase families .....	31
3.3. The Helicase domain .....	32
3.4. Monomeric and multimeric helicase enzymes .....	36
3.5. Unwinding models.....	39
3.6. RNA helicase A, human homolog of MLE .....	41
<b>4. Aim of this work .....</b>	<b>44</b>
<b><i>II. Results .....</i></b>	<b>46</b>
<b>1. MLE monoclonal antibodies.....</b>	<b>46</b>
1.1. Purification of MLE interacting partners.....	48
1.2. Limitations of 6E11 antibodies .....	48
<b>2. MLE and the dosage compensation complex (DCC).....</b>	<b>49</b>
2.1. Effects of MLE depletion in SL2 cells .....	49
2.2. Reconstitution of a DCC complex specifically containing roX2 RNA.....	51
2.3. MLE preferentially associates with MSL1 and MSL2.....	54
2.4. MSL proteins do not influence the ATPase activity of MLE .....	57
<b>3. Biochemical characterization of MLE RNA helicase .....</b>	<b>59</b>
3.1. Effects of RNA binding domain deletions on MLE ATPase activity.....	59
3.2. Effects of RNA binding domain deletions on MLE helicase activity .....	62
3.3. RNA binding properties of recombinant MLE and MLE deletion mutants.....	64
<b>4. <i>In vivo</i> localization of MLE deletion mutants.....</b>	<b>68</b>

<b>5. Strategies to study the function of roX RNAs .....</b>	<b>72</b>
5.1. <i>In vivo</i> localization of the tagged ms2-roX2 RNA.....	72
5.2. Identification of proteins specifically binding ms2-roX2 RNA.....	74
<b><i>III. Discussion.....</i></b>	<b>76</b>
<b>1. Association of MLE with the DCC complex.....</b>	<b>76</b>
1.1. Reconstitution of a complete DCC in Sf9 cells: advantages of the baculosystem.....	76
1.2. MSL proteins specifically bind roX2.....	77
1.3. Integration of MLE into the DCC does not require roX2 RNA.....	78
<b>2. Targeting of MLE to the X chromosome .....</b>	<b>79</b>
2.1. Different properties of MLE RNA binding domains .....	79
2.2. Model of MLE translocation on dsRNA.....	83
2.3. MLE enzymatic activity is not sufficient for proper targeting to the X .....	83
2.4. Correct targeting of the DCC to the X territory in SL2 cells requires MLE.....	85
<b>3. Models for DCC assembly and targeting to the X chromosome in males .....</b>	<b>86</b>
<b><i>IV Materials and Methods.....</i></b>	<b>89</b>
<b>1. Materials.....</b>	<b>89</b>
1.1. Chemicals, reagents and enzymes.....	89
1.2. Oligonucleotides .....	89
1.3. Antibodies .....	90
<b>2. Methods .....</b>	<b>90</b>
2.1. Plasmids .....	91
<b>3. Biochemical methods .....</b>	<b>95</b>
3.1. <i>In vitro</i> transcription and translation.....	95
3.2. Expression and purification of GST fusion proteins from bacteria.....	95
3.3. Expression and purification of proteins and RNA from Sf9 cells .....	96
3.4. Whole SF4 cells extract preparation .....	98
3.5. Immunoprecipitation experiments .....	98
3.6. Pull down experiments .....	99
3.7. GST-MS2 Pull down experiments .....	99
3.8. RNA extraction and RT-PCR.....	100
3.9. ATPase assay .....	100
3.10. Preparation of RNA substrates for the helicase and the gel-mobility shift assays.....	101
3.11. Helicase assay .....	102
3.12. Gel mobility shift assay .....	102
<b>4. Cell biology methods.....</b>	<b>103</b>
4.1. Cell culture.....	103

4.2. Transient transfection in SF4 cells.....	103
4.3. Establishment of the ms2-roX2 stable cell line.....	104
4.4. Immunofluorescence on SF4 and SL2 cells.....	104
4.5. RNAi in SL2 cells.....	105
<b>V. Bibliography .....</b>	<b>106</b>
<b>VI. Appendix .....</b>	<b>1</b>
<b>1. Maps of plasmids.....</b>	<b>1</b>
PfastBac-MLE.....	1
pFastBac-MLE-flag.....	2
pFastBac-MLE <sup>ARB1</sup> -flag.....	3
pFastBac-MLE <sup>ARB2</sup> -flag.....	4
pFastBac-MLE <sup>ARB12</sup> -flag.....	5
pFastBac-MLE <sup>ARGG</sup> -flag.....	6
pFastBac-roX2.....	7
pFastBac-ant-roX2.....	8
pHSP70-MLE-EGFP.....	9
pHSP70-MLE <sup>ARB1</sup> -EGFP.....	10
pHSP70-MLE <sup>ARB2</sup> -EGFP.....	11
pHSP70-MLE <sup>ARB12</sup> -EGFP.....	12
pHSP70-MLE <sup>ARGG</sup> -EGFP.....	13
pHSP70-MLE <sup>ARGG</sup> -EGFP-NLS.....	14
pHSP70-MLE <sup>RGG</sup> -EGFP.....	15
PGEM-Sp6-MSL1.....	16
pGEMT7-MSL2.....	17
pGEX2KG MLE <sup>1-265</sup> .....	18
pGEX2T-MS2.....	19
<b>VII Curriculum vitae.....</b>	<b>20</b>

## **1. Zusammenfassung**

In der Taufliege *Drosophila melanogaster* wird die transkriptionelle Aktivität des männlichen X Chromosoms erhöht, um die (im Vergleich zu weiblichen Zellen) verminderte Gen-Dosis X-chromosomaler Gene auszugleichen. Dieser Prozess wird durch den Dosis-Kompensationskomplex (DCC) vermittelt, einen Ribonucleoproteinkomplex, der aus fünf Proteinen (MSL1, MSL2, MSL3, MOF und MLE) sowie zwei nicht-kodierenden RNAs (roX1 und roX2) besteht. DCC bindet bevorzugt an das X Chromosom, wo er die Transkriptionsrate verdoppelt. Zwei Enzyme sind Teil des DCC: die Acetyltransferase MOF, die spezifisch das Lysin 16 des Histon H4 (H4K16) acetyliert, sowie die RNA/DNA-helicase MLE. Beide Aktivitäten sind für die korrekte Dosiskompensation in männlichen Fliegen verantwortlich. Allerdings erschwert die schwache Assoziation von MLE mit den übrigen MSL Proteinen die biochemische Analyse der Wirkung von MLE. Bislang konnte der Beitrag von MLE zur Dosiskompensation nur genetisch gezeigt werden.

In dieser Arbeit werden mithilfe von verschiedenen experimentellen Ansätzen *in vivo* und *in vitro* die physischen und funktionalen Wechselwirkungen von MLE mit den übrigen MSL Proteinen, sowie mit den roX RNAs untersucht. Monoklonale Antikörper gegen MLE wurden in Ratten induziert, ein wichtiges neues Werkzeug zur Charakterisierung von MLE. Durch Co-Expression von DCC Untereinheiten in sf9 Zellen konnte ein rekombinanter Komplex rekonstituiert und gereinigt werden. Ein präferentieller Einbau von roX RNA konnte beobachtet werden, allerdings nur in Abwesenheit von MLE. *In vitro* konnte nur die unspezifische Bindung von MLE an RNA charakterisiert werden. Der gereinigte DCC hatte keinen Einfluss auf die ATPase Aktivität von MLE, unabhängig von der Anwesenheit von RNA. *In vitro* konnte eine spezifische Assoziation von MLE mit MSL1 und MSL2 beobachtet werden, die direkt und nicht durch RNA vermittelt war. Angesichts dieser Ergebnisse erscheint eine wichtige Rolle der roX2 RNA bei der Koppelung von MLE an den DCC unwahrscheinlich; vielmehr tragen Protein-Protein Wechselwirkungen zur Rekrutierung von MLE an das X Chromosom wesentlich bei.

MLE gehört zur Familie der DEAD-box RNA Helicasen, mit denen es eine ähnliche Domänen-Organisation verbindet. Neben der zentralen ATPase/Helicase Domäne, werden für MLE zwei N-terminale Doppelstrang-RNA (dsRNA) Bindungsmotive (dsRBM1 und dsRBM2), sowie eine C-terminale Einzelstrang RNA/DNA Bindungsdomäne (RGG Box) aus Sequenzvergleichen vorhergesagt. Diese Domänen sind bislang im Kontext von RHA (RNA Helicase A), dem menschlichen Orthologen zu MLE, charakterisiert und ihre RNA Bindungseigenschaften bestätigt worden. Allerdings ist nicht bekannt, wie MLE an die RNA bindet und wie die verschiedenen RNA-bindenden Module zur RNA Stimulierung der ATPase beitragen.

Hier konnte eine bevorzugte Bindung von MLE an doppelsträngige RNA (im Vergleich zu einzelsträngiger) in Bindungsassays beschrieben werden. Die Affinität der Helicase für RNA wird zudem durch Adenin-Nucleotide moduliert. Um den Beitrag einzelner Domänen zu den MLE Funktionen zu bestimmen, wurden eine Reihe von Deletionsmutanten von MLE in Insektenzellen hergestellt und gereinigt. Durch transiente Expression von entsprechenden MLE Derivaten als Fusionen mit GFP (Green Fluorescent Protein) wurde der Einfluss der Domänen auf die Rekrutierung von MLE an das X-Chromosom bestimmt. Im Gegensatz zu den beschriebenen Daten zu RHA erwiesen sich die RB1 und RGG als verzichtbar zur RNA Bindung und deren Entwindung, während dsRB2 eine wichtige Rolle zukommt. Zur korrekten Zielsteuerung von MLE an das X Chromosom ist allerdings die Funktionalität des Enzyms als ATPase alleine nicht ausreichend. Die hier vorgestellte Struktur-Funktionsanalyse leistet einen wichtigen Beitrag zur Beschreibung des Wirkmechanismus der RNA Helicase MLE.

## **2. Summary**

In *Drosophila melanogaster* the transcriptional activity of the male X chromosome is upregulated to compensate for the reduced dosage of X-linked genes as compared to the two X chromosomes in females. This process is mediated by the Dosage Compensation Complex (DCC), a ribonucleoprotein complex consisting of five proteins (MSL1, MSL2, MSL3, MOF and MLE) and two non-coding RNAs (roX1 and roX2). The DCC preferentially localizes on the X chromosomes in males where it doubles its transcription rate. Two enzymes are associated with the DCC: the acetyltransferase MOF, specific for the lysine 16 of H4 (H4-K16), and the DNA/RNA helicase MLE. Genetic experiments demonstrated that both activities are required for dosage compensation in male flies. However, the weak association of MLE to the DCC has complicated its biochemical analysis and, so far, the involvement of MLE RNA helicase in dosage compensation has only been demonstrated genetically.

Using different *in vivo* and *in vitro* approaches the physical and functional interactions of MLE with the other MSL proteins and with the roX RNAs was addressed.

Monoclonal antibodies, specifically recognizing MLE, were raised in rats, offering a new tool for MLE characterization. By coexpression of the DCC subunits in SF9 cells, a recombinant complex containing MSL1-2-3, MOF, MLE and the roX2 RNA was reconstituted and purified. A specific integration of roX2 into the DCC could be observed only in the absence of MLE. Non specific RNA binding properties seemed instead associated to MLE RNA helicase. Moreover, the purified MSL complex did not affect the ATPase activity of MLE in the presence or absence of roX2 RNA. *In vitro*, MLE showed a preferential association with MSL1 and MSL2 and MLE interaction with both MSL proteins were not RNA mediated. In view of these results we suggest that binding to roX2 is not the main determinant for MLE integration into the DCC complex and protein-protein interactions might instead contribute to its proper recruitment to the X chromosome.

MLE is a member of the DEAD-box RNA helicase family and it shares with the other members the same domain organization. In addition to a central ATPase/helicase domain, two predicted N-terminal double strand (ds) RNA-binding motifs (dsRBM1 and dsRBM2) and a predicted C-terminal single strand (ss) RNA/DNA-binding domain (RGG-box) are also present in MLE protein. These domains have been extensively

characterized in RHA, human ortholog of MLE, and their RNA binding properties confirmed. However, it is not known how MLE binds RNA and how the different RNA binding modules contribute to stimulate its enzymatic activities.

A preferential binding of MLE to dsRNA compared to ssRNA was shown by binding assays. In addition, changes in the affinity of MLE for both ssRNA and dsRNA were observed in the presence of different nucleotides. Deletion mutants of MLE were produced and purified from insect cells in order to address the contribution of the different RNA binding domains to MLE enzymatic activities. By transient expression in *Drosophila* cells of the same deletion mutants fused to GFP, the effects of individual domains on MLE recruitment to the X chromosome were also determined. The results show that unlike RHA, the dsRB1 and the RGG domains are dispensable for MLE RNA binding and unwinding, whereas dsRB2 seems to play the major role in coordinating both activities. However, the enzymatic activities alone are not sufficient to properly target MLE to the X chromosome.

These results provide new data on the functional properties of MLE RNA helicase that may help to elucidate its molecular mechanisms of action.

## **I. Introduction**

### **1. Chromatin**

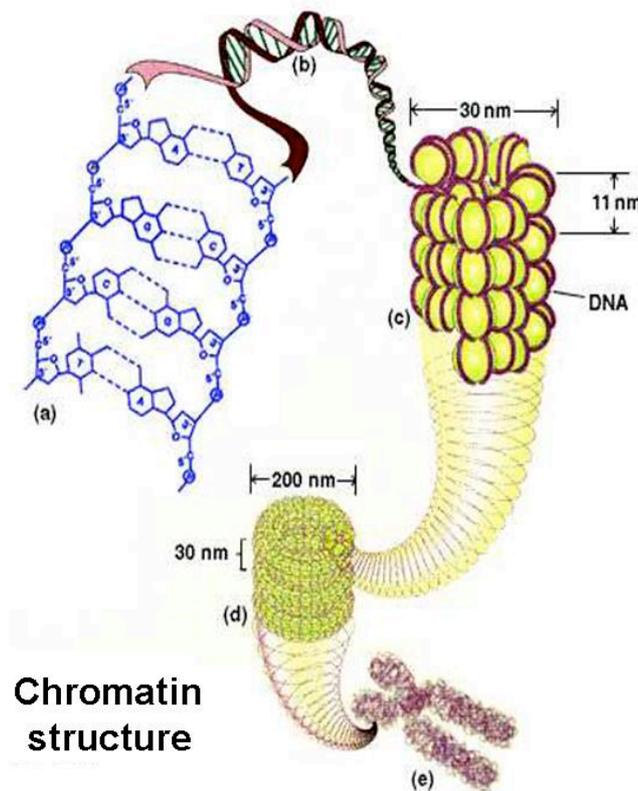
At the end of the 19th century the genetic material was shown to be organized into a complex structure composed of DNA and proteins and localized in a specialized compartment, the nucleus. This structure, detected with basic and acidic dyes, was called chromatin (from the Greek "khroma" meaning coloured) (Flemming, 1882).

In the late 1860's, based on microscopy observations, chromatin has been divided into two distinct domains, distinguished by different levels of dye incorporation. Darkly stained regions were designated as "heterochromatin" whereas less condensed and stained regions were referred to as "euchromatin" (Heitz, 1928; Lyon, 1999). Euchromatin and heterochromatin differ not only at a cytological level but they also represent two different functional states of chromatin itself. Euchromatin is found often associated with transcriptionally active sites, in contrast to heterochromatin consisting predominantly of transcriptionally inactive sequences. Specific chromosomal regions such as telomeres or centromeres are organized constitutively in heterochromatic forms. By contrast, euchromatin is more dynamic and can undergo dramatic changes in compaction levels during specific cellular events (Adkins et al., 2004a; Farkas et al., 2000).

Biochemical and electron microscopy studies have identified the nucleosome as the structural and functional unit of chromatin (Germond et al., 1976). A nucleosome consists of 147 bp of DNA wrapped around an octamer of histones. A tetramer of histones H3 and H4, assembled with two dimers of histones H2A and H2B, forms the nucleosome core particle (Thatcher and Gorovsky, 1994). The inner part of the nucleosome consists of the globular parts of the histones, while their N- and C-terminal domains protrude out of the nucleosome. The N-terminal basic tail region is highly conserved across various species and constitutes the more accessible part of chromatin (Luger et al., 1997; Richmond and Davey, 2003).

Consecutive nucleosomes, lined up on a DNA stretch, generate a fibre with a diameter of 11 nm, named "beads on a string" (Benbow, 1992) (Fig. 1). This configuration is only the first level of packaging of DNA. Several additional folding principles

contribute to a high degree of compaction, necessary to fit, for example, in humans approximately 2 m of DNA into a cell nucleus of about 10  $\mu\text{m}$  of diameter. Association of a fifth histone, H1, allows the formation of a more compact chromatin fibre with a diameter of approximately 30 nm (Adkins et al., 2004b; Thoma and Koller, 1977). Finally, intermolecular interactions between nucleosomes and other not well-characterized non-histone proteins define higher orders of chromatin structure (Felsenfeld and Groudine, 2003; Hayes and Hansen, 2001).



**Fig 1. Different levels of chromatin organizations.** 147 bp of DNA (a) form a double helix (b) wrapped around a histone octamer (yellow) to constitute the nucleosome core particle. Adjacent nucleosomes, connected by short stretches of DNA linker form a fibre of 10 nm, named "beads on a string" (c). The string of nucleosome can be further folded into a 30 nm fibre (d). Association of other not well characterized factors induce higher levels of chromatin compaction (e).

## **1.1. Chromatin Dynamics**

The complex organization of chromatin into higher order structures restricts the access of regulatory proteins to DNA, which mediate essential cellular processes as transcription, DNA replication or DNA repair. The need for a cell to modulate chromatin organization, in response to different stimuli, becomes then a necessity. Three major strategies to specifically and locally remodel chromatin have been described so far (reviewed in (Mellor, 2005)).

Variants of histones H2A, H2B, H3 and H1 have been identified and shown to mediate specific biological processes (Ausio and Abbott, 2002). For instance, the H3 variant Cse4/CENP-A has an important role in centromeres assembly and function (Black et al., 2004). Another H3 variant, H3.3, is enriched at active loci throughout the cell cycle and represents a specific mark for the transcription event (Tagami et al., 2004).

Remodelling complexes act directly on nucleosomes and use the energy of ATP hydrolysis to alter their position on DNA, which may promote either repression or activation of transcription (Becker and Horz, 2002; Lusser and Kadonaga, 2003).

Histone-modifying enzymes post-translationally modify histones to create specific signals and to induce localized alterations of the chromatin structure and function (Cosgrove and Wolberger, 2005). The “histone code” hypothesis postulates that specific factors can act on chromatin by binding particular histone modifications. This hypothesis is so far supported by the discovery of chromodomains and bromodomains present in various proteins, which specifically recognize methylated and acetylated lysines of histones, respectively (de la Cruz et al., 2005). Only this last part will be discussed more in details in this introduction.

## **1.2. Post-translational modifications of histones**

### **1.2.1. Histone acetylation and deacetylation**

Histone acetylation is the transfer of an acetyl group from the acetyl-coenzyme A cofactor to the  $\epsilon$ -amino nitrogen of lysine residues, mainly in histone tails (Fig. 2). This reaction is catalysed by histone acetyltransferase enzymes (HATs), whereas the reverse reaction is performed by histone deacetylases (HDACs) (Legube and Trouche, 2003).

The first histone acetyltransferase, *hat1*, was identified in yeast in 1995. Subsequently, different transcriptional coactivators like Gcn5 or p300/CBP were shown to possess histone acetyltransferase activity, suggesting a role for these enzymes in transcriptional activation (Brownell and Allis, 1996). Further studies in yeast reinforced this hypothesis demonstrating a clear enrichment of acetylated histones in promoter regions (Parekh and Maniatis, 1999). Acetylation of lysines in the histone tails probably weakens the interaction of histones with the nucleosomal DNA, thus opening chromatin and facilitating the access of transcription factors to their recognition elements (Kurdistani and Grunstein, 2003). In addition, acetylated lysines can create a specific signal for regulating factors or chromatin remodelling complexes and contribute to their targeting to a limited, specific region. Consistent with this last hypothesis, bromodomains have been shown to act as acetyl-lysine recognition modules, targeting HATs, which contain such a domain, to particular chromosomal sites and contributing to their substrate specificity (Yang, 2004).

According to their sequence similarities, known HATs have been classified in six major families: the Gcn5/PCAF, the HAT1, the CBP/p300, the SRC, the TAFII250 and the MYST families (Marmorstein and Roth, 2001). Beside the presence of the HAT domain, each family presents unique features. The size of the active domain and the context of different modules in which this domain is present have been also used as parameters for the HATs classification. The MYST HAT enzymes differ from the other family members because of the presence of a chromodomain instead of a bromodomain. Most of the MYST proteins are implicated in a wide range of processes and, in contrast to a general role of HATs in transcriptional activation, some members of this family appear to be involved in gene silencing (Grienenberger et al., 2002; Utley and Cote, 2003).

Histone acetylation is a reversible modification and HDAC enzymes that specifically remove acetyl groups from lysine residues have also been identified. In general, if HATs are associated with transcriptional activation, HDACs negatively regulate gene expression. Based on the sequence similarity to yeast Rpd3 (Reduced potassium dependency), Hda1p (Histone acetylase 1) and Sir2 (Silent information regulator), HDAC proteins have been classified in three families: class I, II, and III. In mammals also a class IV has been described (Ng and Bird, 2000).

### **1.2.2. Histone methylation and demethylation**

Histone methyltransferases (HMT) are the more recently characterized histone modifying enzymes. Their general targets are lysines and arginines on H3 and H4 (Fig. 2). Lysines may be mono-, di- or tri-methylated, whereas arginines may be only mono- or di-methylated (Bannister et al., 2002).

Initially the *Su(var)3-9* gene was identified in *Drosophila* as a suppressor of position effect variegation, suggesting a role for histone methyltransferases enzymes in transcriptional repression (Tschiersch et al., 1994). Later, H3-K9 and H3-K27 methylation have been shown to be enriched at heterochromatic sites and to be required for establishment and maintenance of gene silencing (Fischle et al., 2003; Min et al., 2003). However, H3-K4 methylation correlates with the active transcriptional state both in yeast and human (Strahl and Allis, 2000). This modification, widely present on all autosomes in mammals, is absent from the transcriptionally inactive X chromosome in females that is instead enriched in H3-K9 methylation (Cohen et al., 2005; Khalil et al., 2004).

Methyl groups are relatively small and, in contrast to acetyl groups, they do not neutralize positive charges. Methylation of histone tails does probably not induce structural chromatin changes. However, chromodomains have been identified as methyl-lysine recognition modules in several chromatin associated proteins (Nielsen et al., 2002) as well as in several modifying enzymes (Lachner et al., 2001) and ATPases of different chromatin remodelling complexes (Bannister et al., 2001; Wysocka et al., 2005).

Lysines are not the only target of histone methyltransferases, also arginine methylation occurs within the tails of histones (Strahl et al., 2001). So far two classes of arginine methyltransferases that share a conserved catalytic domain have been described: class I,

that asymmetrically dimethylate arginines and class II that catalyse symmetric dimethylation. Arginine methylation, like acetylation, correlate with active transcription (Cheng et al., 2005). The molecular mechanisms used by arginine methyltransferase enzymes to activate transcription are still poorly understood and up to now a specific methyl-arginine binding module has not yet been identified.

Unlike other modifications, histone methylation has been considered for a long time as a stable mark. The recent discovery of LSD1, a specific lysine histone demethylase, strongly suggests that also this modification is reversible. LSD1 is an amino oxidase and it specifically demethylates H3-K4 in a FAD-dependent way (Shi et al., 2004). Even though an arginine demethylase has not been reported, an alternative pathway to reverse arginine methylation has been proposed. Arginine demethylation might be brought about by the conversion of a methyl-arginine residue to citrulline, catalyzed by the deiminase PADI4 (Cuthbert et al., 2004). This enzyme has been shown to act on arginines or monomethyl-arginines at specific sites on H3 and H4 (Wang et al., 2004).

### 1.2.3. Histone phosphorylation

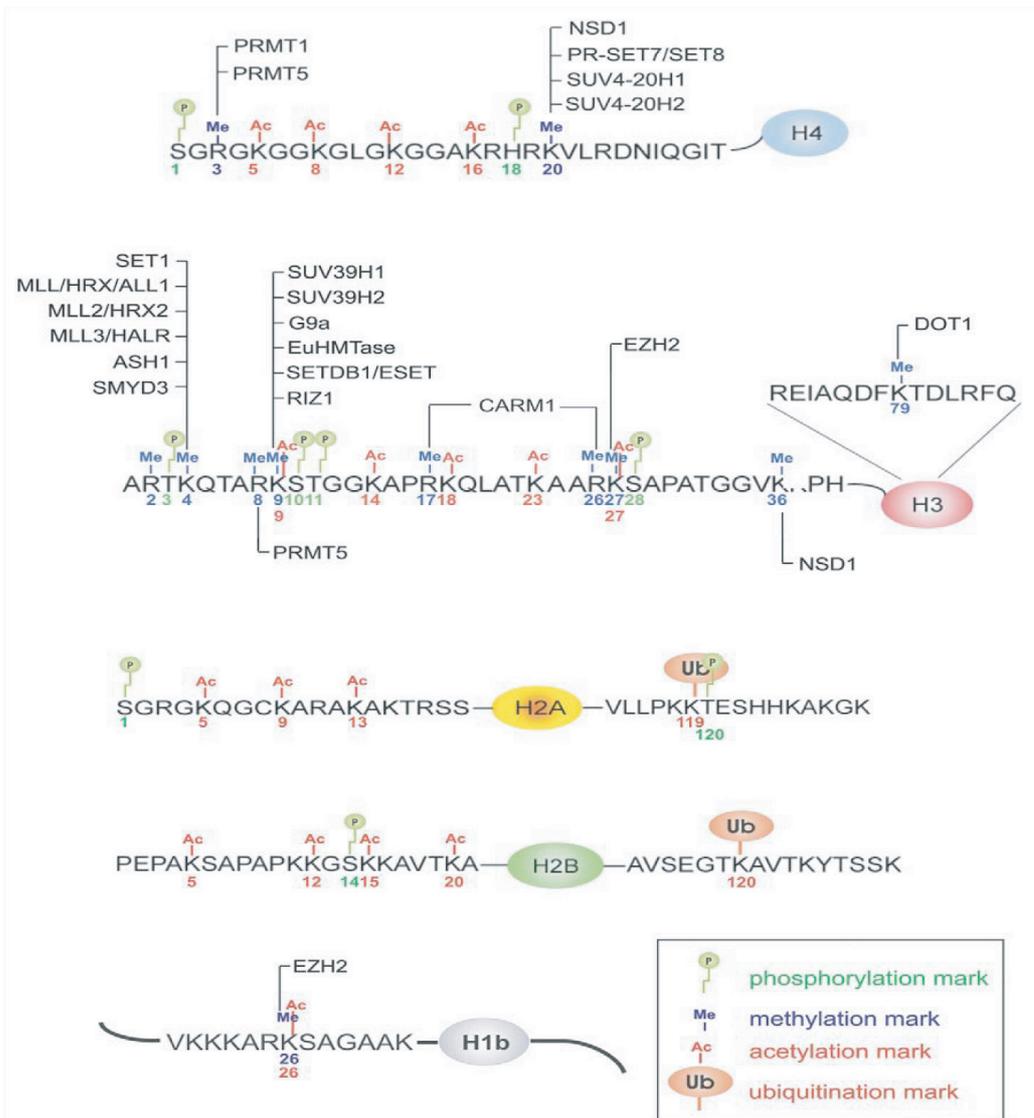
On H3, phosphorylation of serine 10 and serine 28 has been described (Fig. 2). In particular, H3-S10 phosphorylation is linked to chromosome condensation during mitosis and meiosis (Wei et al., 1998). Members of the aurora AIR2-Ipl1 kinase family are responsible for this specific modification, whereas type 1 phosphatases (PP1) catalyze the dephosphorylation reaction (Murnion et al., 2001). More recently, several evidences indicate that H3-S10 phosphorylation, in combination with other modifications on the same histone tails, has also an important role in transcriptional activation in different organisms (Clayton and Mahadevan, 2003; Fischle et al., 2005).

How H3 phosphorylation affects gene expression is not known. One possibility is that phosphorylation, like acetylation, increases the accessibility of chromatin to regulatory proteins (Nowak and Corces, 2004). According to the “histone code” model, phosphorylated serines can serve as a signal for the binding of transcription factors or remodelling complexes. Recently, members of the 14-3-3 family of phosphoprotein-interacting factors have been identified as phosphorylated H3 binding proteins (Liao et al., 2005). These factors exist *in vivo* in a dimeric form, offering two sites for phosphoprotein binding. 14-3-3 isoforms can then act as adaptor between phosphorylated H3 and another phosphorylated protein, recruiting specific effectors to

chromatin. Alternatively, they could bind modified H3 tails on two adjacent nucleosomes (Macdonald et al., 2005).

#### **1.2.4. Histone ubiquitylation**

Histone ubiquitylation has been described for H2A and H2B only in the context of other modifications. In yeast, H2B-K123 ubiquitylation is required prior to H3 methylation to induce gene silencing at telomeric regions (Ng et al., 2002; Sun and Allis, 2002). Although this covalent modification has been known for a long time, the enzyme responsible for it has been identified only recently as the Bre1 RING domain (Hwang et al., 2003). The improvement of techniques to study post-translational modifications, such as mass-spectrometry, has allowed the identification of new histone modifications and in particular of histone variants (Freitas et al., 2004). For example, macroH2A1.2 can be mono-ubiquitylated as well as phosphorylated and methylated at different positions, suggesting that, like canonical H2A, this H2A variant is regulated by combination of covalent modifications (Chu et al., 2005).



**Fig. 2. Histones are subjected to a variety of post-translational modifications.** These modifications include acetylation (Ac, red), methylation (Me, blue), phosphorylation (P, green) and ubiquitylation (Ub, brown). The enzymes responsible for methylation of mammalian histones are listed above or below their target sites (Margueron et al., 2005).

### **1.2.5. The “histone code” hypothesis**

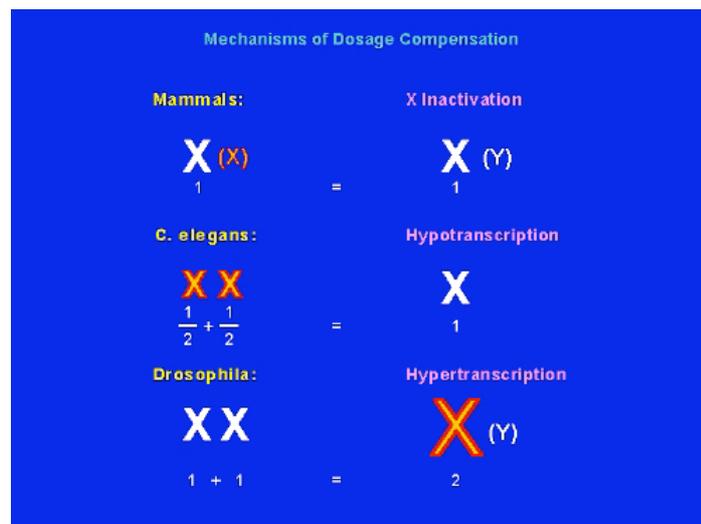
In view of the large number of histone modifications known, a concept has been introduced to explain how post-translational histone modifications can affect chromatin functions. The “histone code” hypothesis proposes that covalent modifications of histone tails constitute a code. The encoded information can then be read by chromatin binding factors specifically recognizing particular modified residues within the histone tails (Turner, 1993). The discovery of chromodomains and bromodomains as methyl-lysine and acetyl-lysine binding modules, respectively, strongly supports this idea (reviewed in (de la Cruz et al., 2005)). In the last years, several experimental results suggested that the “histone code” probably uses combinations of modifications to regulate functional chromatin states. This model can explain how the same modification might lead in some cases to either transcriptional activation or repression (Strahl and Allis, 2000). For example, H3-K9 methylation initiates chromatin condensation and silencing (Lachner et al., 2001). In the context of methylated H3-K4 and H4-K20, H3-K9 methylation contributes to create an active mark, by allowing the binding of Brahma, the catalytic subunit of the *Drosophila* SWI/SNF remodelling complex (Beisel et al., 2002). Another interesting aspect of the “histone code” hypothesis is the idea that combinations of modifications on the same histone or in adjacent ones are interdependent. H3-K4 methylation inhibits H3-K9 methylation, preventing the placement of silencing marks at specific loci (Zegerman et al., 2002). In yeast, H2B-K123 ubiquitylation is also required for efficient H3-K4 methylation (Sun and Allis, 2002).

In addition, the “histone code” hypothesis has been extended to the “nucleosome code” hypothesis, by proposing that chromatin structures and functions are largely dependent on the local concentration and combination of differentially modified nucleosomes (Jenuwein and Allis, 2001).

## 2. Dosage compensation

In many animal species, males and females differ by the number and type of sex chromosomes, with females having often two X chromosomes and males only one X chromosome in addition to the gene-poor Y chromosome. To overcome the unbalanced expression of the X-linked genes, heterogametic organisms have evolved several mechanisms to equalize the level of expression of sex chromosome-linked genes. These strategies have been grouped under the term “dosage compensation” (reviewed in (Lucchesi et al., 2005)) (Fig. 3).

So far, three model organisms are being used to study dosage compensation at the molecular level: mammals, nematodes and flies. Despite some differences, all three systems share a common feature: regulation of transcription through modulation of chromatin structure (Mellor, 2005).



**Fig. 3. Dosage Compensation mechanisms.** In mammals equalization of X linked gene expression is achieved through random inactivation of one X chromosome in females. In worms both X chromosomes in the hermaphrodites are downregulated two fold as compared to males. In *Drosophila melanogaster* the single male X chromosome is activated two fold (Lucchesi et al., 2005).

## 2.1. Dosage compensation in mammals

In mammals, dosage compensation is initiated early in development and requires, as a key step, the transcription of the non-coding RNA *Xist* (Brockdorff et al., 1991). In undifferentiated ES cells both X chromosomes in females are active and X inactivation occurs upon induction of differentiation. At this time, low level of *Xist* expression can be detected on both X chromosomes in females. When differentiation starts, *Xist* expression increases and becomes restricted to the inactive X chromosome (Lee et al., 1999). The gene coding for *Xist* RNA resides in the X inactivation centre (*Xic*), which is located within the X chromosome. The *Xic* regulates all three crucial events necessary for X chromosome inactivation: X chromosome counting, selection of an X chromosome to inactivate and the inactivation process itself (Lyon, 1996). Once transcribed, *Xist* RNA spreads *in cis* along the entire X chromosome that undergoes repression (Chow et al., 2005). This event precedes a multitude of chromatin changes that specifically mark the inactive X chromosome during differentiation (Cohen et al., 2005; Jeppesen and Turner, 1993). The pattern of DNA methylation at the *Xist* locus correlates with its activity: the active *Xist* allele on the inactive X chromosome is unmethylated, whereas the inactive *Xist* allele on the active X chromosome is methylated (Hansen, 2003). DNA methylation is not the only epigenetic mark of the inactive X chromosome. Several histone modifications known to be associated either with expressed or silent chromatin also distinguish the active and inactive X chromosomes respectively. Examples are the hypoacetylation of histones H3 and H4, di-methylation of H3-K9, H3-K27 tri-methylation, a lack of H3-K4 di- and tri-methylation and, as a final event, the incorporation of the histone variant macroH2A. All those marks seem to be required for the late maintenance of a repressive state on the inactive X (Cohen et al., 2005).

*Xist* function is regulated by another locus mapped at the 3' of the *Xic* region, the X controlling element (*Xce*) (Simmler et al., 1993). *Xce* codes another non-coding RNA, *Tsix*, antisense to *Xist*. Like *Xist*, *Tsix* is polyadenylated, does not contain any open reading frame and does not accumulate in the cytoplasm (Lee et al., 1999). *Tsix* is expressed from the active X chromosome in somatic cells where it exerts a repressive effect on *Xist* RNA accumulation before and during the initiation of X inactivation (Mlynarczyk and Panning, 2000).

Even though some progress have been made on the characterization of Xist RNA, open questions regarding the tethering of Xist RNA on the X chromosomes and its exact mechanism of action remain unanswered. Very little is known about the proteins that bind to the different functional regions of Xist transcripts and the enzymes responsible for the chromatin modifications associated with the inactive X chromosome. One important finding has been the discovery that the Polycomb group proteins ESC (extra sex chromosome) and Enx1 (mouse homologue of Enhancer of zeste) may be implicated in X inactivation. The ESC/Enx1 complex transiently associates with the X chromosome shortly after Xist coating. The presence of this complex correlates with H3-K27 methylation consistent with a role for Enx1 as an HMTase (Silva et al., 2003). In a recent study, the histone methyltransferase G9a has been shown to be responsible for the modifications of H3-K9 and H3-K27 on the inactive X chromosome (Rougeulle et al., 2004). The activity of these complexes seems to be implicated in the early maintenance of X inactivation, whereas silencing appears strictly dependent on Xist RNA (Plath et al., 2003).

Identification of new factors that specifically bind the inactive X chromosome will shed new light on the molecular basis of counting and choice in female mammals.

## **2.2. Dosage compensation in worms**

Unlike mammals, worms have no need to distinguish homologous chromosomes. In the nematode *C. elegans*, dosage compensation is mediated by a large protein complex targeted to both X chromosomes of hermaphrodites to reduce their expression by half (Meyer et al., 2004). The *C. elegans* Dosage Compensation Complex (DCC) complex has been extensively characterized both genetically and biochemically. This complex can be separated into two further subcomplexes, SDC and DPY. DPY subcomplex is composed of DPY-26, DPY-27, DPY-28 and MIX1, highly homologous to components of the 13S condensin complex, required in both sexes for DNA compaction in mitosis and meiosis (Chan et al., 2004). The SDC proteins SDC1, SDC2, SDC3 together with DPY21 and DPY30 constitute the SDC subcomplex (Chu et al., 2002). *SDC1* encodes a conserved protein with seven N-terminal zinc finger motifs and expressed in both sexes, suggesting a more general role for this factor not restricted to dosage compensation (Nonet and Meyer, 1991). One important regulatory player of dosage compensation in

nematodes is the SDC2 protein (Dawes et al., 1999). SDC2 is present only in hermaphrodites, whereas its expression is blocked in males by the XOL-1 (XO lethal 1) factor. Recruitment of SDC2 to the X chromosomes in hermaphrodites is independent of the other complex components, but its interaction with SDC-1 and SDC-3 is required for targeting the DPY/condensing-like complex to the X of and for transcriptional downregulation of the X-linked genes (Nusbaum and Meyer, 1989). SDC3 possesses two domains with separable functions: a zinc finger domain required for dosage compensation and a myosin-like ATP-binding domain involved in transcriptional repression at the *her-1* locus. SDC3 does not bind chromatin and its association to the X chromosomes has been shown to require the other DCC members (Davis and Meyer, 1997). In addition to SDC1, SDC2 and SDC3, targeting of the DPY subcomplex to the X chromosomes also required DPY-30. This protein is ubiquitously expressed in the nucleus and has high homology with the H3-K4 specific histone methyltransferase Set1p (Hsu and Meyer, 1994).

The worm DCC also localizes to the autosomal, male sex determining *her-1* gene, repressing its activity around 20 fold (Hodgkin, 1980). In this case, SDC3 seems to play a central role in the complex assembly and it can localize at the *her-1* gene independently of SDC2 (Yonker and Meyer, 2003).

Even though several recognition elements have been identified within the X chromosomes in hermaphrodites, large portions of the X chromosomes do not attract the DCC when translocated to autosomal positions. These regions still contain compensated genes, suggesting that localization of the DCC to both X chromosomes happens presumably by spreading of the complex from flanking binding sites (Csankovszki et al., 2004). It is still not clear how the recognition elements recruit the DCC proteins and how a two fold downregulation of gene expression is triggered. In contrast to mammals and flies, no RNA component involved in compensation of the hermaphrodite X chromosomes has been identified so far.

### **2.3. Dosage compensation in flies**

In *Drosophila melanogaster*, dosage compensation happens in males, where the transcription of the X chromosome is doubled to ensure a similar gene expression level to females that carry two X chromosomes. This hyperactivation is mediated by a

ribonucleoprotein Dosage Compensation Complex (DCC), containing at least six known proteins (MSL1, MSL2, MSL3 (Male Specific Lethal 1, 2, 3), MOF (Male absent on the First), MLE (Maleless) and JIL1) and two non-coding RNA (roX1 and roX2) (Kelley and Kuroda, 2000). Both the MSL proteins and the roX RNAs have been identified in a screen for mutation that confer male-specific lethality and subsequently shown to be required for dosage compensation (Belote and Lucchesi, 1980).

One key regulatory player of dosage compensation in flies is the sex lethal gene product SXL. This factor is absent in males, but expressed in females where it prevents the translation of the MSL2 transcript. The absence of MSL2 and the subsequent decreased stability of the other MSL proteins prevent the formation of the DCC complex in females (Bashaw and Baker, 1995). Unlike MSL2 all the other MSL proteins are also present in females, even though at a lower level as compared to males. The function of these factors in females is not known.

In males, the MSL proteins are specifically targeted to the X chromosome where they are responsible for the acetylation of H4-K16 by the acetyltransferase MOF (Akhtar and Becker, 2000; Parekh and Maniatis, 1999). Phosphorylated H3-S10 is also enriched on the X chromosomes in males. This modification is due to the Kinase JIL1. However, the contribution of this enzyme to dosage compensation is still not clear (Wang et al., 2001).

### **2.3.1. Establishment of the *Drosophila* Dosage Compensation Complex (DCC)**

Dosage compensation in males begins at the late blastoderm/early gastrula stages. In this time window, all the MSL proteins, maternally contributed, are present in the embryos, with the exception of MSL2. Upon MSL2 expression, 3 hours after egg laying (AEL), localization of the DCC to the male X chromosome can be observed in all somatic cells (Franke et al., 1996; Hilfiker et al., 1994). H4-K16 acetylation appears slightly after MSL recruitment to the X, in agreement with the role of MOF as the histone acetyltransferase responsible for this modification (Rastelli et al., 1995). Transcription of roX1 and roX2 non-coding RNAs, identical in expression in male larvae and adults, is different in embryos. RoX1 is initially expressed in both sexes 2 hours AEL and disappears gradually from females during embryogenesis. In contrast, roX2 expression is specifically restricted to males 6 hours AEL. Deletion of the *roX2* gene has no phenotype and localization of the MSLs on the X chromosome happens

normally 3 hours AEL. MSL binding to the X is delayed to 7 hours in embryos lacking roX1 and drastically reduced in double-mutant flies for both RNAs. These results suggest that while roX2 is dispensable, roX1 plays an important role in the initiation of dosage compensation in males (Meller, 2003). RoX1 is expressed in embryos before MSL2 and consequent formation of the complex. However, in embryos mutated for *mle*, it disappears after gastrulation and can be detected again only 4-5 hours AEL. These results suggest that none of the MSL protein is required for initial roX1 transcription, whereas MLE might contribute to stabilize roX1 RNA at early developmental stages. This role of MLE seems dependent on its RNA helicase activity, as a similar lack of roX1 stability is also observed in embryos expressing the MLE helicase inactive mutant, MLE<sup>GET</sup> (Meller, 2003).

A model has been proposed to explain the order of assembly of the DCC components during embryogenesis. A first interaction between roX1 RNA and MLE is initially formed and stabilized by MSL3 and MOF. All these proteins contain RNA binding motives and, maternally provided, are present in the embryos during roX1 transcription (Akhtar et al., 2000; Rastelli et al., 1995). Once expressed, MSL2 together with MSL1 and roX2 can complete the assembly of this initial ribonucleoprotein complex (Meller, 2003).

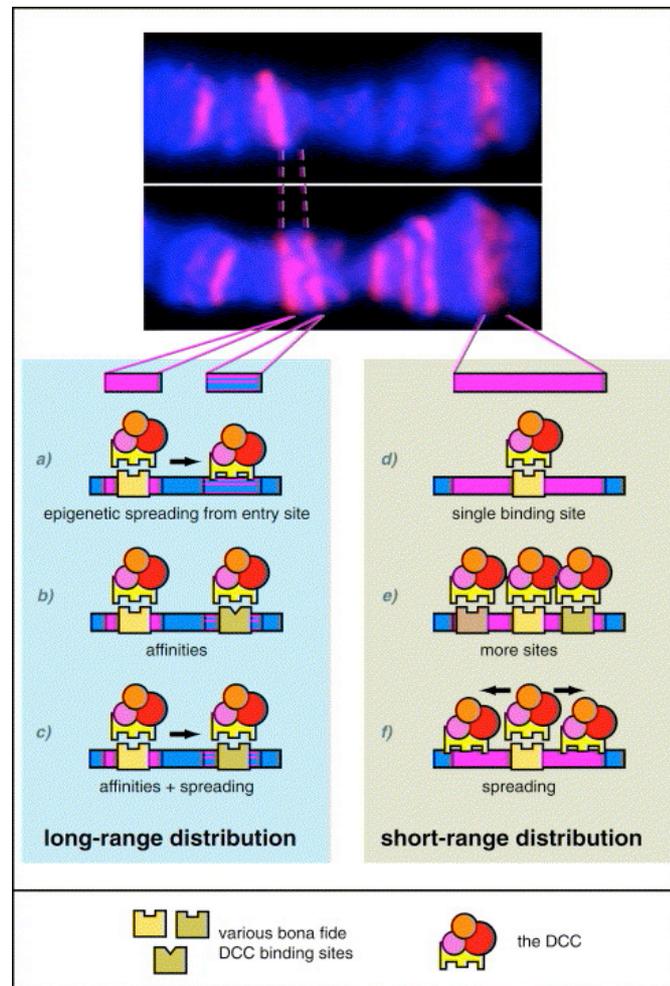
### 2.3.2. Targeting of the DCC to the male X chromosome

Targeting of the MSL proteins to the X chromosome can be visualized by immunostaining on polytene chromosomes from salivary glands of *Drosophila* 3<sup>rd</sup> instar larvae (Zhimulev et al., 2004). In wild type males the binding pattern of the DCC is restricted to interbands on the X chromosome at around hundred sites (Demakova et al., 2003). However, in the absence of MLE, MSL3 or MOF, only MSL1 and MSL2 can be reproducibly observed at a subset of binding sites originally referred to as “chromatin entry sites” (CES) (Kelley et al., 1999; Lyman et al., 1997). These genetic studies allowed proposing the first “spreading model” to explain how the MSL proteins are targeted to the X chromosome in males (Fig. 4, a). This model assumed that the CES are the primary recruitment elements bound initially by MSL1 and MSL2. Additional association of the other members, MLE with roX RNAs, MSL3 and MOF leads to the formation of a stable ribonucleoprotein complex able to spread *in cis* and to cover all hundreds sites along the X chromosome (Meller et al., 2000). The discovery of *roX1*

and *roX2* genes as two of the sites required for recruitment of the DCC to the X and the further characterization of their targeting functions, strongly supported the “spreading” mechanism (Kageyama et al., 2001; Park et al., 2003) (see paragraph 2.3.2.1).

Several recent observations brought up the idea that targeting of the DCC cannot be restricted to a limited number of entry sites and that “spreading” may be only a specific feature of the *roX1* and *roX2* genes. Large DNA fragments, lacking mapped entry sites, have been shown to recruit the DCC when translocated from the X to autosomal locations. “Spreading” to autosomal regions flanking the insertion site was not observed even when a *roX* gene has been inserted close to it (Fagegaltier and Baker, 2004). Moreover a third entry site, mapping to the *18D* locus, recruits the DCC when moved to autosomes both in wild type and in *msl* mutant backgrounds. However, unlike *roX1* and *roX2*, the *18D* locus is not transcribed and, interestingly, it shows autosomal spreading only as a multimer (Oh et al., 2004). These results suggest that all hundred sites on the X are competent for binding and that, at the resolution of polytenes, “spreading” might not be the major mechanism employed by the DCC to coat the X chromosome. In this scenario, chromatin entry sites are not specific recruitment elements but rather the highest affinity sites. In the new proposed “affinities model”, the concentration of the DCC is the major determinant for recruitment. At limiting amount of MSL proteins, partial or full complexes bind only the “high affinity” sites (Fig. 4, b). As complex levels rise, also the “moderate” and the “low affinity” sites will be filled until the entire X chromosome is covered (Fagegaltier and Baker, 2004; Straub et al., 2005a) (Fig. 4, b). However, a 179 kb DNA segment, located between the two compensated genes *white* and *roughest* and normally bound by the DCC on the X, failed to target the complex to autosomes (Bhadra et al., 1999). It is therefore possible that many of the “lowest affinity” sites profit from a high local concentration of MSL brought about by “high affinities” sites on the X. When moved out of this context, these elements are no longer able to recruit the complex. The term “spreading” can then be used to explain how the DCC binds some of the “low affinity” sites (Straub et al., 2005a) (Fig. 4, c). It should be considered that the low resolution of the polytene chromosomes does not allow distinguishing the short-range distribution of the DCC within a single immunofluorescent band (Fig. 4, d-f) and different possibilities have been proposed. Each band could correspond to one single binding site for the DCC. Alternatively, many specific binding sites might reside within it. In addition, each band could represent the result of several binding events: a first binding to a specific recruitment element,

followed by diffusion of the complex *in cis*. The term “spreading” could then be used to describe the occupancy of different sites within a chromosomal domain or even a single gene, a situation not visible on stained polytenes chromosomes.



**Fig. 4. Models for targeting of the DCC to the male X chromosome.** Depicted are two *Drosophila* polytene chromosomes, stained for MSL1 at low (upper panel) and high (lower panel) complex concentration. Models for targeting of the DCC to the male X chromosome. **a)** The “spreading model”. **b)** The affinity model. Binding of the MSL proteins to “high” and “moderate affinity” sites (long-range distribution). **c)** Recruitment of the DCC to low affinity sites. **d), e), f)** MSLs binding within one band (short-range distribution) (Straub et al., 2005a).

### 2.3.2.1. *roX1* and *roX2* “high affinity” sites

*roX1* and *roX2* genes are located on the X chromosome and have two distinct roles in dosage compensation. First, they contain “high affinity” sites, bound by the DCC complex in wild type male flies. Second they are the source of *roX1* and *roX2* non-coding RNAs (Meller and Rattner, 2002). As binding element they have been extensively characterized. Both loci can recruit the MSL proteins when inserted into an autosomal position and in some cases they can support weak “spreading” to neighbouring polytene chromosomal bands (Kelley et al., 1999). *RoX1* and *roX2* genes contain a male-specific DNase hypersensitive site (DHS), within a 110 bp region, conserved among different *Drosophila* species. Both *roX1* and *roX2* DHS are necessary for MSL targeting, however in contrast to *roX1* DHS, no spreading occurs from the *roX2* DHS when inserted in an autosome (Park et al., 2003). Search for a similar sequence on the X chromosome, did not identify any candidate within the remaining postulated “high affinity” sites, failing in the identification of a common feature between the X binding elements for the MSL proteins (Kageyama et al., 2001). *RoX1* and *roX2* seem to be atypical chromatin entry sites. They are the only so far characterized binding elements able to produce non-coding RNAs and differently from the other X chromosome sites, they require MLE binding to attract the remaining components of the DCC complex. In a *mle* background, only the “high affinity” sites on the X chromosome are bound by MSL1 and MSL2 proteins, with the exception of *roX1* and *roX2* genes. In contrast, in the absence of MOF or MSL3, partial complexes can be visualized on both *roX* genes, as well as at a few other “high affinity” sites on the X (Kageyama et al., 2001).

In addition, even though the targeting function of *roX1* and *roX2* DHS is independent of their own transcription, in both cases, recruitment of the DCC complex requires *roX* RNAs, which can be supplied *in trans* (Park et al., 2003).

These results, as well as the RNA-dependency of *roX* genes for MSL binding, provide a strong evidence for an “RNA first” model explaining the MSL complex assembly at the *roX* loci. This model suggests that the affinity of the MSL complex for chromatin is increased by association of *roX* RNAs. On the other hand, stabilization of both RNAs seems to require MLE (Park et al., 2003).

### 2.3.3. How two fold activation is achieved

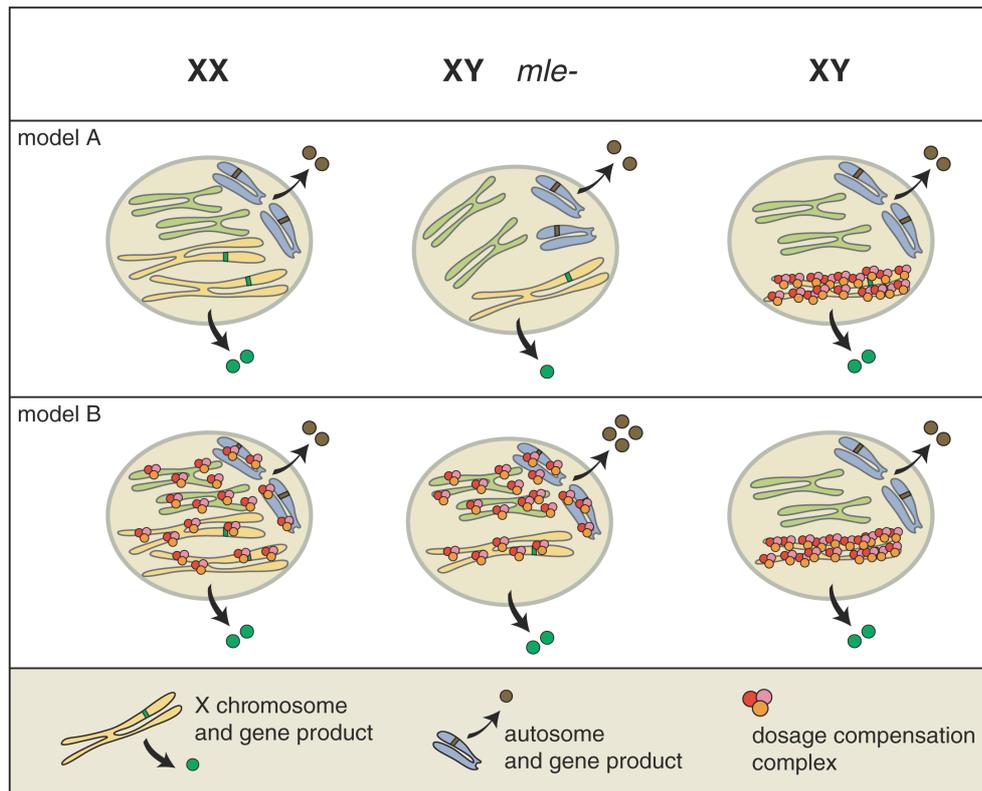
A specific feature of dosage compensation in flies is the two fold activation of X-linked genes in males, to adjust their expression to female cells with two X chromosomes. The molecular mechanism responsible for this compensation process is still debated.

MOF, one of the DCC members, is a histone acetyltransferase specific for H4-K16 acetylation (Akhtar and Becker, 2000). During embryogenesis, H4-K16 acetylation follows targeting of the complex to the X chromosome, suggesting a main role for the MSL proteins to specifically concentrate MOF activity on the X chromosome (Rastelli et al., 1995). The DCC complex seems then to be directly responsible for chromatin changes on the X, which in turn lead to activation of gene expression (Fig. 5, model A). In agreement with this “direct model”, MOF has been shown to relieve chromatin-mediated repression of transcription *in vitro* and *in vivo*, when artificially tethered to a promoter (Akhtar and Becker, 2000). In addition, MSL proteins seem to associate with actively transcribed regions for chromatin remodelling and compensation (Sass et al., 2003).

Such a direct connection between histone acetylation and X-linked genes expression, even though supported by several experimental observations, cannot explain all situations. In yeast, H4 acetylation correlates in general with active transcription, but acetylation at H4-K16 can also mark inactive chromatin regions by preventing binding of the Bdf1 transcription factor (Kurdistani et al., 2004). An “inverse model” proposes that the recruitment of the DCC complex to the X chromosome might not necessarily be related to activation but it is instead necessary to counteract autosomal hyperactivation, due to the presence of only one X chromosome in males (Birchler et al., 2003) (Fig 5, model B). This model was initially developed to explain how aneuploidies could lead to changes in gene expression and in this context monosomy of the X chromosome in males can be considered a special case of aneuploidy. Accumulation of activating factors to the autosomes (Fig. 5, model B, middle), after loss of one X chromosome in males, increases gene expression. MOF, as activating factor, is then sequestered to the X chromosome by the MSLs in order to re-establish on the autosomes a normal expression profile (Fig. 5, model B, right). Finally, hyperactivation on the X due to MOF activity is limited to two fold by other repressive regulatory mechanisms. This model predicts a two-fold activation of the autosomal genes, upon ablation of the DCC, a situation that can be studied in Schneider cells.

Depletion of MSL2 in SL2 cells by RNA interference (RNAi) decreased X-linked genes transcription around 50%, in agreement with a direct function of the DCC in transcriptional activation. In the same experiments the autosomal genes tested did not show any significative change in expression levels, contradicting the “inverse model” (Hamada et al., 2005; Straub et al., 2005b).

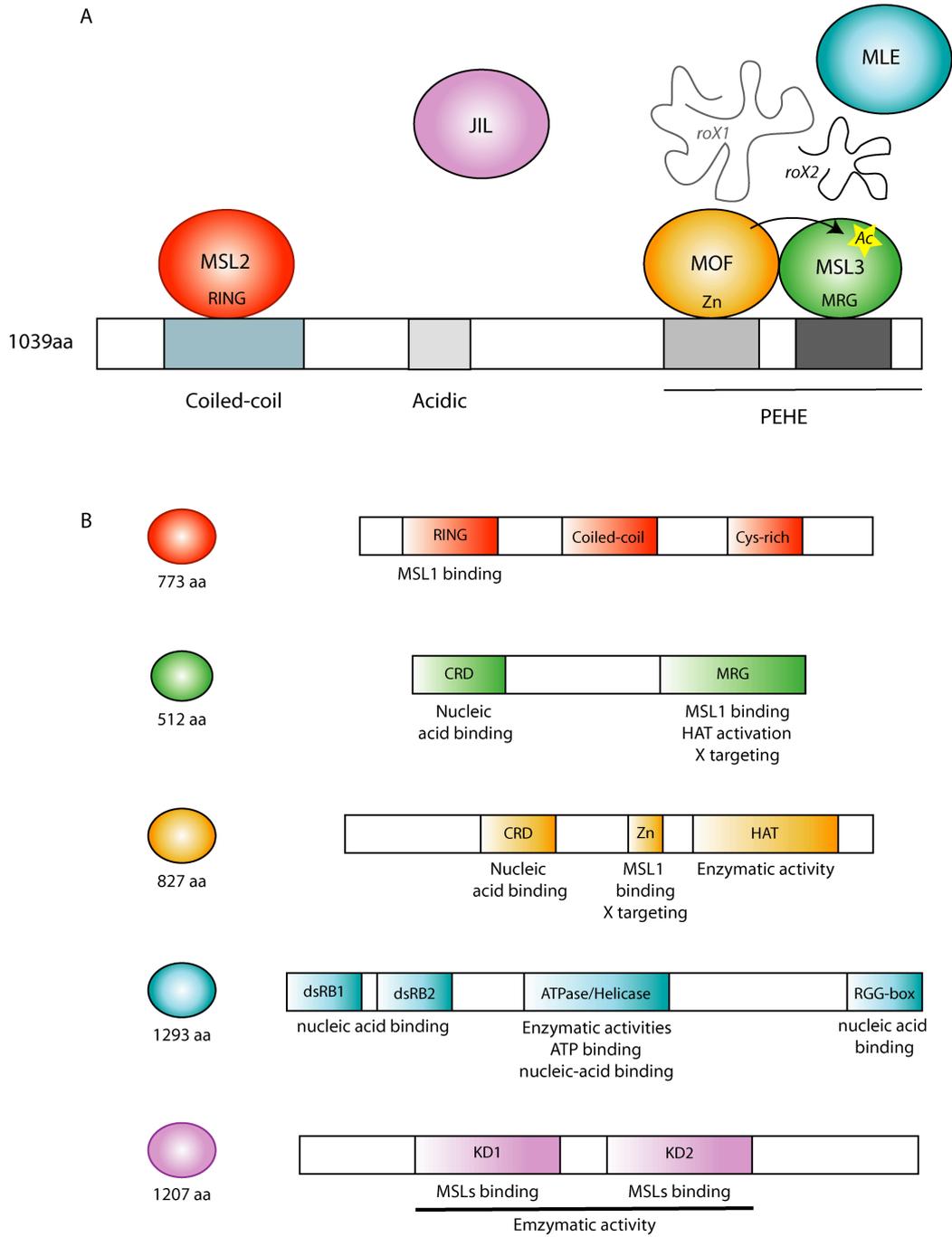
However, even if the majority of genes on the X respond directly to MSL binding, some of them appear to be compensated in an MSL-independent manner (Gergen and Wieschaus, 1986). In addition, the large induction of transcription by artificial targeting of MOF to a promoter in yeast cannot explain alone the precise two fold activation of the X-linked genes observed *in vivo* (Akhtar and Becker, 2000). It is likely that general chromatin regulators with repressive functions are also involved in dosage compensation. On the other hand, a repressive role for the MSLs in limiting MOF HAT activity to a precise two-fold activation should also be taken in account. Compensation *in vivo* can be then seen as a balance between activation and repression. Several factors, not identified in the screening for male-specific lethal mutations, have been also shown to regulate X chromosomes structure and function. Low levels of JIL1 in hypomorph flies severely reduce euchromatic regions of polytene chromosomes in both sexes, with more severe effects on the male X chromosome (Wang et al., 2001). Mutations in ISWI and NURF301, subunits of the NURF remodelling complex, affect both cell viability and gene expression during *Drosophila* development and also cause a characteristic “puffing” of the male X chromosome (Badenhorst et al., 2002; Corona et al., 2002). Recently the level of expression of Suv(var)3-7, a protein involved in heterochromatin formation together with Su(var)3-9 and HP1, has been shown to be critical for fly viability and chromosome integrity. High doses of this protein induce chromosome compaction both in females and in males, with the latter being the most affected. In these transgenic flies the male X chromosome is covered by H3-K9 dimethylation and to a lower extent by HP1. One possible explanation is that the unique chromatin environment determined by the MSL protein renders the male X chromosome more accessible to Su(var)3-7, followed by targeting of Su(var)3-9 and HP1 (Delattre et al., 2004).



**Fig 5. Models to explain how dosage compensation is achieved between male (XY) and female (XX) flies. Model A** suggests that the dosage compensation complex (DCC) directly activates gene expression on the male X. **Model B** proposes that compensation is accomplished by a genome-wide increase in transcription, due to activator imbalance inferred by X chromosome monosomy. Recruitment of the activators onto the X alleviates the transcription enhancement on the autosomes. While the expression status is the same in wild type males for both models, MSL mutants (XY *mle*<sup>-</sup>) are predicted to show major differences characterizing the two hypotheses (Straub et al., 2005a).

#### 2.3.4. The MSL proteins and their physical and functional interactions

MSL1 has a central “scaffold” role in the assembly of the DCC (Fig. 6). It interacts directly with the RING finger domain of MSL2 through its N-terminal region and with the MRG domain of MSL3 and the zinc-finger domain of MOF, respectively, via two distinct region of its C-terminal PEHE domain (Li et al., 2005; Morales et al., 2004; Scott et al., 2000). MSL2 contains also a cysteine-rich region (CXC) at the C-terminus, not essential for dosage compensation (Copps et al., 1998). Association of MOF and MSL3 with MSL1 is required *in vitro* for MOF nucleosomal HAT activity and specificity (Morales et al., 2004). MLE is an ATP-dependent DNA/RNA helicase and according to the prevalent concept is recruited to the X chromosome by binding roX RNAs (Richter et al., 1996) (see paragraph 2.3.6). Also MSL3 and MOF have been shown to bind RNA via their related chromo-barrel domains (Nielsen et al., 2005) and together with MLE are released from the male X chromosome upon treatment of cultured SL2 cell nuclei with RNase (Akhtar et al., 2000). Furthermore, MSL3 is acetylated by MOF both *in vitro* and *in vivo*. This modification has been suggested to lead to a specific loss of interaction between MSL3 and roX2 RNA (Buscaino et al., 2003). However, MSL3 and MOF RNA binding capacities are dispensable for recruitment and, association with MSL1 seems instead determinant for their correct localization on the X chromosome (Morales et al., 2005; Morales et al., 2004). Even though MSL interactions seem necessary for correct complex assembly and recruitment, it is possible that stable maintenance of the DCC on the X chromosome requires association of MSL3, MOF and MLE with roX RNAs (Gu et al., 2000).



**Fig. 6. The MSL proteins and their physical and functional interactions. A)** MSL1 works as a “scaffold” protein for the DCC assembly. **B)** Functional domain and regions of the MSL proteins, identified by protein-protein, protein-RNA/DNA interactions and in vivo localization studies.

### 2.3.5. roX1 and roX2 non-coding RNAs

RoX1 and roX2 non-coding RNAs have been isolated in a RNA screening as specific male transcripts (Amrein and Axel, 1997). Even though as binding element they have been extensively characterized (see paragraph 2.3.2.1), little is known about the function of roX1 and roX2 non-coding RNAs in dosage compensation.

*RoX1* has only one intron and its spliced transcript is 3.1 kb long. *RoX2* contains instead two introns and undergoes alternative splicing. The major roX2 isoform present in males has a length of approximately 0.6 kb (Park et al., 2003). A genetic approach has been used to identify functional domains within roX1 RNA necessary for targeting and functional activity of the DCC complex on the X chromosome in males (Stuckenholz et al., 2003). The results showed that the 3' end of roX1 is required for dosage compensation. However, most roX1 RNA deletions failed to rescue double *roX1-roX2* mutant males, suggesting that different RNA domains, interspersed along the length of roX1, cooperate with the MSL complex to compensate the X chromosome. Considering the redundant function of roX1 and roX2 RNAs in dosage compensation and the lack of an obvious sequence similarity between them it is more probable that the secondary structure instead of the primary sequence of these two non-coding RNAs could be important.

Despite their differences, roX1 and roX2 RNAs seem to have a redundant function in dosage compensation. Transgenic flies mutated either for *roX1* or *roX2* loci do not show any particular phenotype, while deletion of both genes is male lethal. In the double-mutant flies, targeting of the MSL complex to the X chromosome is severely compromised. MSL1 and MSL2 proteins bind only few "high affinity" sites on the X and together with H4-K16 acetylation are mostly relocated to autosomal locations, to the fourth chromosome and to heterochromatin. These results suggest a potential role for roX RNAs in the specific targeting of the DCC to the male X chromosome (Kelley et al., 1999; Meller and Rattner, 2002). In contrast to the absence of escapers from flies mutated in each *mSl* gene, a low percentage of *roX1-roX2* double-mutant males are still viable and overexpression of MSL1 and MSL2 can partly rescue their lethality. One possible explanation is that a partial complex lacking non-coding RNAs can still be assembled and targeted to the X for dosage compensation (Oh et al., 2003). This idea suggests that roX RNAs are not the only components mediating targeting of the complex. Alternatively, targeting of the MSL complex to chromatin required general

not yet identified RNA cofactors, while recruitment to the male X chromosome is roX-dependent.

Several published studies have addressed the transcriptional regulation of the *roX* genes. Genetic experiments pointed to MSL2 as the only DCC factor responsible for roX transcription independently of a functional DCC. In agreement with these data, deletion of the *roX1* DHS did not influence the ability of MSL2 to drive roX expression and transcription was also observed in the absence of MSL1, MSL3, MOF or MLE (Rattner and Meller, 2004). In a parallel, genetic study, the *roX1* and *roX2* DHS have been shown to positively regulate roX expression and all MSL proteins were also required (Bai et al., 2004). Additional contradictory results came from an independent analysis to address the functional role of MLE in dosage compensation. Using a biochemical approach, the authors showed a specific binding of MLE to a small region in the *roX2* promoter, suggesting a possible function for this RNA helicase in roX2 transcription (Lee et al., 2004). In view of all these data, further experiments are needed to elucidate the molecular mechanisms controlling roX expression and function.

### 2.3.6. The DNA/RNA helicase MLE

MLE is a member of the DEAH family and has both ATPase and helicase activities (Gorbalenya et al., 1989). Its human ortholog, RNA helicase A (RHA), has been extensively characterized and showed to be involved in different biological processes (Lee and Hurwitz, 1993) (see chapter 3).

Several observations classify MLE as a component of the dosage compensation complex in *Drosophila melanogaster*. Similar to the other MSL proteins, depletion of MLE is lethal in males whereas it has no phenotype in females. MLE colocalizes with the MSLs on the X chromosome in males and is required early in development for stabilization of roX1 RNA and correct assembly of the complex (Meller, 2003). After establishment, maintenance of MSL3, MOF and roX RNAs with the DCC complex depends on MLE, as in its absence, only MSL1 and MSL2 can be seen at a subset of “high affinity” sites on the X (Meller et al., 2000).

Recruitment of MSL3, MOF and MLE to the X chromosome has been shown in cells to be RNase-sensitive (Richter et al., 1996). Surviving male flies, deleted for both *roX1* and *roX2* genes, show MSL1 and MSL2 at few “high affinity” sites on the X, whereas MSL3, MOF and MLE strongly stained only autosomal loci. These observations are

consistent with a role for roX1 and roX2 RNAs in the specific recruitment of MSL3, MOF and MLE to the male X chromosome (Park et al., 2002). However, the existence of RNA-free MSL complexes has also been documented. Overexpression of MSL2 can rescue in part the lethality of the double-mutant males for *roX1* and *roX2* by inducing further recruitment of MSL3 and MOF to the X (Meller and Rattner, 2002). In addition, the MLE<sup>GET</sup> protein, deficient in ATPase activity, associates with the other MSLs at a certain number of “high affinity” sites in the absence of any detectable roX1 and roX2 RNAs (Gu et al., 2000; Lee et al., 1997).

These results suggest that MLE RNA helicase activity may be required for stable incorporation of the roX RNAs into the DCC. It has been proposed that once the assembly of the complex is accomplished, roX RNAs can be lost and the other complex components could then be held together by protein–protein interactions (Meller and Rattner, 2002). Consistent with this idea, MSL2, MSL3 and MOF strongly interact with MSL1 *in vitro* (Copps et al., 1998; Morales et al., 2004). In addition, partial MSL complexes can be immunoprecipitated from RNase-treated S2 cell extracts in the absence of detectable roX2 RNA (Akhtar et al., 2000).

MLE is the only MSL protein with a demonstrated function outside of dosage compensation. Mutations in an MLE allele, *nap* (non action potential), have been shown to block action potentials at high temperatures. This phenotype is mainly due to decreased expression levels of the *para* genes coding for the Na<sup>+</sup> channels (Kernan et al., 1991). How exactly *mle* (*nap*) affects *para* expression remains uncertain. *Para* transcripts undergo adenosine to inosine (A to I) RNA editing via a mechanism that requires dsRNA secondary structure formation (Valente and Nishikura, 2005). In a *mle* (*nap*) background, more than 80% of *para* transcripts are aberrantly spliced and contain internal deletions including the normally edited regions. One possible explanation is that MLE helicase activity is required in general to resolve secondary structures within RNAs, allowing their correct folding (Kernan et al., 1991). Even if it is not known whether roX RNAs are edited and whether this modification is required for their functional integration into the DCC complex, such a role for MLE helicase activity should not be excluded.

### **3. DNA and RNA helicases**

#### **3.1. Helicases and their functions**

Nucleic acid helicases are proteins that catalyze the transient unwinding of stable DNA or RNA duplex molecules, using NTP hydrolysis as a source of energy (Delagoutte and von Hippel, 2002). These enzymes are involved in all biological processes that require manipulation of nucleic acids, like DNA replication, DNA repair, translation, RNA splicing, RNA editing and ribosome biogenesis (Silverman et al., 2003). Although for a long time helicase enzymes were thought to function mainly by destabilization of DNA or RNA helices, it is becoming clearer that they also have many other activities. Helicases can also facilitate the annealing of single-stranded nucleic acids, acting as chaperones (Lorsch, 2002). Some helicases have been also shown to displace proteins from nucleic acids or to be involved in nucleosomes remodelling, through their translocation along DNA (Byrd and Raney, 2004; Jankowsky et al., 2001). Moreover, a role for helicases as transcriptional coactivator or corepressor has also started to emerge (Bates et al., 2005; Fujita et al., 2003; Lee et al., 2005).

The broad range of functions in which DNA and RNA helicases are involved cannot explain their *in vivo* substrate specificity. The presence of accessory domains flanking the conserved helicase domain, integration of helicases into multisubunits complexes and/or restriction of these enzymes in specialized nuclear compartments might all contribute to their specific target recognition (Delagoutte and von Hippel, 2003).

#### **3.2. DNA and RNA helicase families**

RNA and DNA helicases are characterized by the presence of a conserved helicase domain, containing up to nine motifs (designated Q, I, Ia, Ib, II, III, IV, V and VI), which are required for ATP binding, hydrolysis, nucleic acid binding and unwinding (de la Cruz et al., 1999). Both DNA and RNA helicases enzymes have been grouped into four superfamilies, SF1, SF2, SF3 and SF4, based on the extent of similarity and the internal organization of the conserved motifs within the helicase domain (Gorbalenya et al., 1989).

SF1 and SF2 superfamilies are the largest and closely related groups, containing DNA and RNA helicases (Silverman et al., 2003) from viral, prokaryotic and eukaryotic organisms. Both family members contain up to nine conserved amino acid motifs, whose sequences, arrangements and predicted secondary structures are in general very similar. Most RNA helicases belong to the SF2 superfamily, although some viral RNA helicases are more related to the SF1 superfamily members (Gorbalenya et al., 1989).

SF3 helicases share a small helicase domain with only three conserved motifs (I, II and a novel motif, C). Most members of this family are viral enzymes involved in DNA replication (Hickman and Dyda, 2005).

The SF4 superfamily includes DNA helicases related in sequence to the *Escherichia coli* Dna-B protein. These enzymes share five conserved motifs (I, Ia, II, III and IV) and unwind DNA in the 5' to 3' direction (Oakley et al., 2005).

An additional fifth group, exemplified by the transcription termination factor Rho, has been also recognized as a subgroup of the SF4 superfamily. SF5 helicases show high sequence similarity to the  $\beta$  subunit of proto-translocating ATPase (Walmacq et al., 2004).

### 3.3. The Helicase domain

The conserved helicase domain of SF1 and SF2 helicases contain up to nine characteristic motifs (Q, I, Ia, Ib, II, III IV, V and VI) (Fig. 7). These enzymes can unwind DNA or RNA duplexes either in 3' to 5' or 5' to 3' directions (de la Cruz et al., 1999).

The Q motif has been discovered as a unique feature of the DEAD subclass of RNA helicases (see later). It consists of a nine amino acid sequence, containing an invariant glutamine (Q) and it is generally present upstream of motif I. Extensive analyses of this element in the yeast translation initiation factor Ded1, showed that it regulates not only ATP binding and hydrolysis but also the affinity of the protein for RNA substrates and ultimately its helicase activity (Cordin et al., 2004).

Motif I, also known as Walker A motif, contains the AxxGxGKT consensus sequence, where x could be any amino acid. The lysine within the GKT sequence binds to  $\beta$  and  $\gamma$  phosphates of the ATP molecule. Replacement of this lysine with uncharged amino

acids, as asparagines (GKT to GET), abolishes binding to ATP in several RNA helicases (Pause and Sonenberg, 1992).

Motif Ia seems to be involved in single-stranded DNA or RNA (ssDNA or ssRNA) binding and helicase activities. Specific mutations in motif Ia of UL9, a protein essential for herpes simplex virus type 1 (HSV-1) replication *in vivo*, exhibited wild type levels of ATPase activity and moderate to severe defects in ssDNA-stimulated ATPase activity and ssDNA binding (Marintcheva and Weller, 2003).

Motif Ib participates in RNA binding, in association with motifs Ia, IV and V (Rogers et al., 2002). However, motif Ib has been poorly studied at the biochemical level and its involvement in RNA binding has been only deduced from the structure of the NS3 RNA helicase co-crystallized with poly(dU)<sub>10</sub> (Kim et al., 1998).

Motif II, or Walker B motif, is highly conserved among DNA and RNA helicases and it is also present in many proteins involved in DNA replication and repair (Gorbalenya et al., 1988). The first D residue within this motif has been shown to bind Mg<sup>2+</sup>, required for ATP binding and to be responsible for coupling of ATPase and helicase activities (Pause and Sonenberg, 1992). Variations in motif II have been also used to sub classify SF2 RNA helicases into the DEAD-, DExH- and DEAH-box proteins. The exact number and biological role of these enzymes are unclear. The majority of DEAD-box helicases have a role in ribosome biogenesis and translation, whereas DExD/H enzymes have been shown to be mostly involved in pre-mRNA splicing (Tuteja and Tuteja, 2004).

Motif III is required for both ATPase and unwinding activities. Eukaryotic initiation factor 4A (eIF-4A), mutated within this domain, retains both ATP binding and hydrolysis as well as RNA binding functions, while RNA helicase activity is severely compromised (Pause and Sonenberg, 1992).

Motif IV contributes together with motif Ia to a specific interaction between RNA helicases and oligonucleotides (Caruthers and McKay, 2002).

Motif V, as well as motif III, shows substantial differences in both length and amino acid sequence among the different helicase families. The function of this motif is still not well understood (Caruthers and McKay, 2002).

Sequence of motif VI is less conserved between SF1 and SF2 helicases. However, both families share an arginine residue in the middle of the motif, shown to be important for RNA binding and ATP hydrolysis in several DEAD-box RNA helicases (Pause et al., 1993).

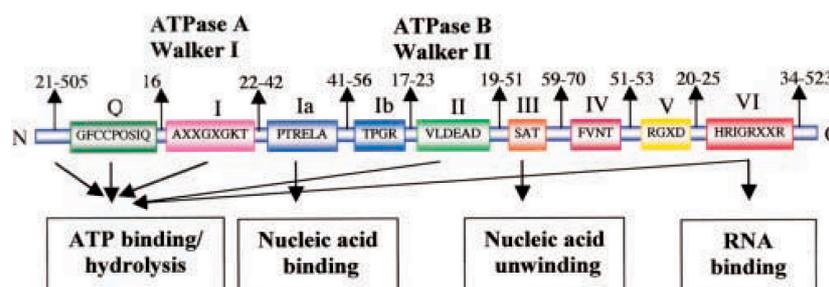
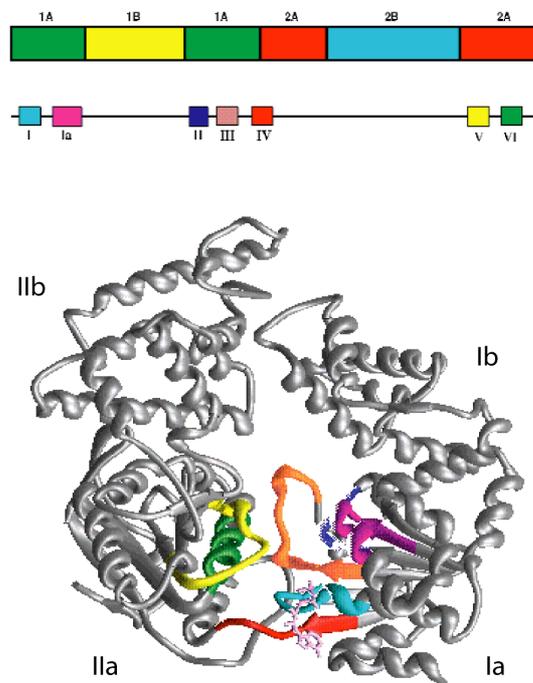


Fig 7. Conserved motifs in the helicase domain of SF1 and SF2 helicases (Tuteja and Tuteja, 2004).

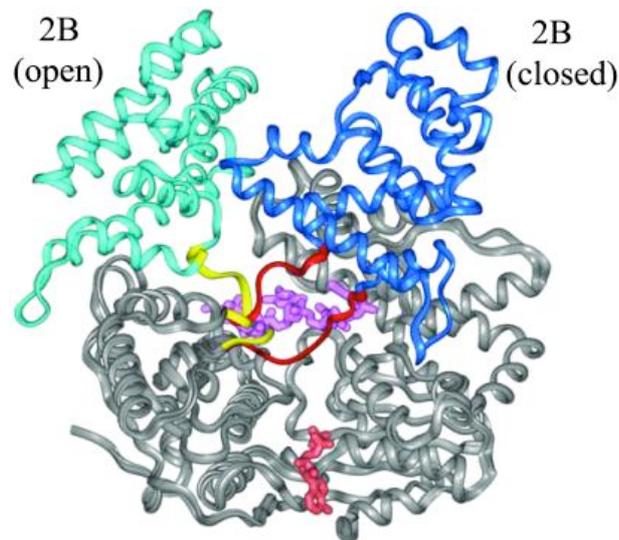
Only few DNA and RNA helicases have been crystallized (Carmel and Matthews, 2004; Subramanya et al., 1996; Velankar et al., 1999; Zhao et al., 2004). These structures are all very similar and to a large extent also superimposable. The first DNA helicase to be crystallized in a complex with ADP was PcrA from a thermophilic bacterium *Bacillus stearothermophilus* (Subramanya et al., 1996). PcrA crystallizes as a monomer, which consists of two parallel domains (1 and 2), separated by a deep cleft. Each domain consists of two additional subdomains (A and B), resembling the central region of the recombinase RecA, which contains both helicase-like strand separating and strand annealing activities (Fig. 8). Subdomain 1A contains both the Walker A and the Walker B motifs. The ADP moiety is located at the bottom of the cleft between subdomain 1A and 2A. However, only subdomain 1A can bind ADP. Subsequent crystal structures of PcrA in complex with ssDNA/dsDNA junction and in the presence or absence of ATP, suggested that PcrA undergoes conformational changes upon NTP binding and hydrolysis. NTP binding determines the closure of a cleft between domain 1A and 2A. The position of domains 1B and 2B is also consequently altered as a function of NTP binding (Velankar et al., 1999).

The *E. coli* Rep SF1 helicase exists in solution as a monomer only in the absence of DNA and binding to ssDNA induces its dimerization (Chao and Lohman, 1990). Although in the crystal structure the two Rep monomers display the same major contacts with ssDNA, they exist in two conformations, “open” and “closed”, which differ in the orientation of the 2B subdomain, relative to the 2A domain (Fig. 9) (Korolev et al., 1997). ADP is present in the cleft, formed by domain 1A, 1B and 2A of only one Rep monomer and induces the formation of the “closed” state. Interestingly,

the 2B domain is the only domain that does not contain any of the conserved aminoacids of the nine helicase motifs and characteristic of the SF1 helicase superfamily (Korolev et al., 1997). Later experiments using recombinant Rep, deleted for the 2B domain did not show any effect on its helicase activity (Cheng et al., 2002). These observations suggest that despite their high structural homology in the helicase core region, the molecular mechanisms by which these enzymes carry out their catalysis is different. This might be related to the different length of the linker region connecting the two helicase domains as well as to the different rates of ATP hydrolysis, observed for these enzymes.



**Fig. 8. The relationship between linear sequence and the position of the conserved helicase motifs in the PcrA helicase structure.** In PcrA, domain A and B in the helicase core region are further divided into two additional subdomain, 1 and 2. The ATP binding site is present in domain 1A. Binding of ATP determines the closure of a cleft between domain 1A and 2A, maintained by interactions between the bound  $Mg^{2+}$  in domain 1A and the  $\gamma$  phosphate of ATP in domain 2A. The domains are colored the same as in the structure image. The bound ATP is colored pale lilac (Velankar et al., 1999).



**Fig. 9. Ribbon diagrams showing the superposition of the two conformations (“open” and “closed”) of the wild type Rep protein in complex with ssDNA and ADP. Six nucleotides are shown in purple, and the ADP is shown in orange (modified from (Korolev et al., 1997)). The two conformations differ in the orientation of the 2B domain, shown in light blue for the open form and deep blue for the closed form. The hinge region connecting the 2B domain to the 2A domain, and about which the 2B domain rotates by  $\approx 130^\circ$  to convert from one form to the other, is shown in yellow for the open form and red for the closed form (Cheng et al., 2002).**

### 3.4. Monomeric and multimeric helicase enzymes

The strong structural conservation of the helicase domain implies an equally strong functional conservation between these enzymes. However, not all proteins containing a helicase domain show unwinding activity, when tested *in vitro*. Examples are members of the Swi/Snf family of remodelling complexes. These factors do not have helicase activity and they use ATP hydrolysis mainly to move nucleosomes along DNA (Havas et al., 2000). These observations suggest that the helicase domain is not sufficient on its own to promote strand separation and that additional nucleic acid binding capacities must be supplied either as adjacent motifs or by associated factors. DsRNA binding motifs (dsRBM) and RGG-boxes have been infact identified outside of the conserved core region of several RNA helicases. The presence or the absence of these domains has been used to divide RNA helicases into monomeric and multimeric enzymes respectively (Gibson and Thompson, 1994). It has been proposed that monomeric helicases can bind non specifically both ssRNA and dsRNA in a forked RNA template. However, only one of the RNA binding domains requires ssRNA specificity to account

for the translocation polarity of the correspondent enzyme. In addition dsRBMs and RGG-boxes seem also to mediate protein-protein interactions and to influence specific substrate recognition (Lee and Hurwitz, 1993).

RNA helicases lacking accessory RNA binding domains are predicted to function in multimeric complexes, by association with external RNA binding factors. This situation has been described for the eukaryotic initiation factor eIF-4A and the Prp16 RNA helicase from *Saccharomyces cerevisiae*. Even though eIF-4A hydrolyses ATP *in vitro* as a monomer, its helicase activity is strongly enhanced by association with eIF-4B, eIF-4H and eIF-4F (Rogers et al., 2001). In yeast, overexpression of helicase domain-flanking regions of Prp16 helicase caused dominant negative phenotypes, suggesting that sequestration of other factors or association with non functional complexes might occur (Wang and Guthrie, 1998).

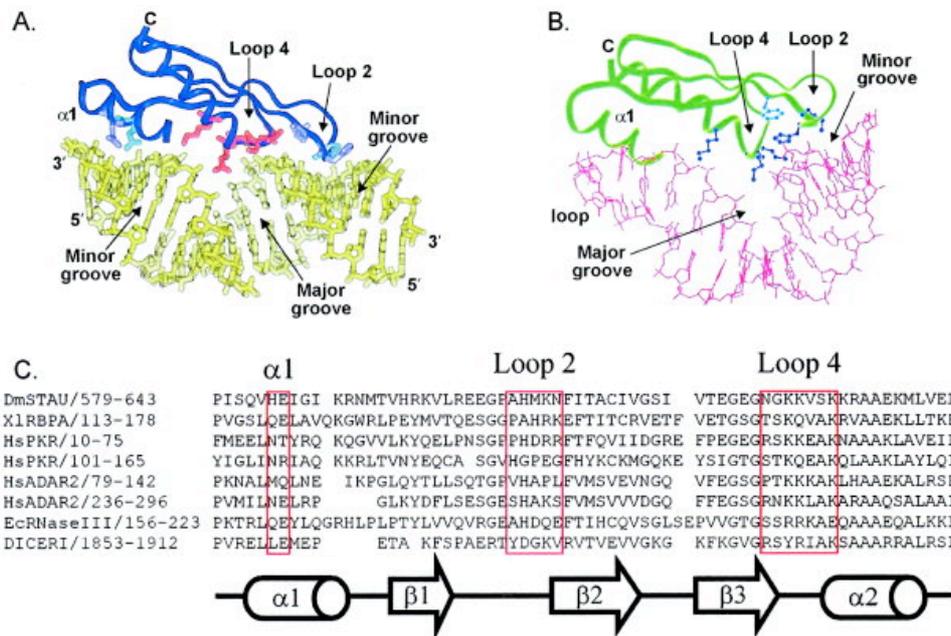
#### 3.4.1. The dsRNA binding motif (dsRBM)

The double-stranded RNA binding motif (dsRBM) has been initially identified in the *Drosophila melanogaster* protein staufer and the *Xenopus laevis* RNA binding protein A, by sequence similarity searches (Fig. 10). Later, several other proteins have been shown to contain this domain as a single motif or in multiple copies (St Johnston et al., 1992).

DsRBMs from several proteins have been solved and revealed a peculiar  $\alpha$ - $\beta$ - $\beta$ - $\beta$ - $\alpha$  fold. High resolution structures of these motifs bound to dsRNA targets showed a striking similarity between them, suggesting a common mode of substrate recognition (Nanduri et al., 1998). However, these structures do not explain how dsRNA binding proteins select specific binding sites on dsRNA ligands. Biochemical data suggested only a general binding affinity role for these domains with little or no selectivity. The context of different domains or regions, flanking dsRBMs, seems instead to play an important role in the specific RNA target recognition. In addition, dsRBMs can also act as protein-protein interaction modules. Isolated dsRBMs of PKR form heterodimeric complexes with full length PKR *in vitro*. Dimerization seems to be responsible for PKR autophosphorylation and enzymatic activities (Zhang et al., 2001).

Even though the dsRNA nucleotide sequence is not essential for specific substrate binding, it seems to contribute to the affinity of dsRBMs for RNA as compared to DNA (Bevilacqua et al., 1998). RNA tends to fold together via base pair interactions into

intricate secondary structures, essential for correct biological function. Common secondary structure motifs include hairpin, loops, stems, bulges and the A-form double helix. This double helix differs from the DNA B-form double helix because of a narrow major groove, virtually inaccessible to proteins. Interactions of proteins with RNA helices occur mainly in the minor groove. This structure allows dsRNA binding proteins to interact non specifically with dsRNA or structured ssRNA regions and to ignore ssRNA, ssDNA and dsDNA species (Seeman et al., 1976).



**Fig. 10. Two high-resolution structures of dsRBMs bound to RNA.** (A) Crystal structure of the *X. laevis* RNA binding protein A (Xlrba) bound to coaxially stacked helices. The structure revealed three regions on the RNA contacted by the dsRBM: two successive minor grooves and the intervening major groove. (B) Structure derived by NMR studies of Staufen's third dsRBM bound to a stem-loop RNA. The stem-loop RNA is composed by a minor groove, major groove and the loop region. These regions are contacted by the same face of the protein as in the crystal structure. (C) Sequence alignment of numerous dsRBMs (*Drosophila* Staufen motif III, *X. laevis* RNA binding protein A motif II, human PKR motifs I and II, human ADAR2 motifs I and II, *E. coli* RNase III, and DICER1). The 1, loop 2, and loop 4 are predicted to contact dsRNA based on the structures above (Carlson et al., 2003).

### 3.4.2. The RGG-box

The RGG domain harbours a consecutive stretch of glycines, regularly interrupted by either aromatic residues or arginines. In addition to RNA helicases, RGG-boxes have also been found in many proteins involved in RNA metabolism (Ghisolfi et al., 1992). Isolated RGG boxes from several RNA helicases show a preferential binding to ssRNA and ssDNA *in vitro* (Cobianchi and Wilson, 1987), (Kiledjian and Dreyfuss, 1992), (Zhang and Grosse, 1997). However, in combination with other RNA binding modules, RGG domains can increase the affinity of a protein for both ssRNA and dsRNA, suggesting that intramolecular interactions can also occur between adjacent RNA binding domains (Zhang and Grosse, 1997). Interestingly, methylation of arginines has been described for several RGG-box containing proteins (Kzhyshkowska et al., 2001; Mears and Rice, 1996; Smith et al., 2004). In some cases, this modification has been shown to modulate both helicases activity and nuclear localization (Smith et al., 2004). In addition, RGG-boxes, like dsRBMs, also mediate protein-protein interactions (Gendra et al., 2004; Smith et al., 2004). The C-terminal RGG-box of RNA helicase A (RHA), the human ortholog of *Drosophila* MLE, forms with the N-terminal double-stranded RNA binding domain II a protein-protein interaction surface required for its association with the Werner syndrome protein (WRN) (Friedemann et al., 2005).

### 3.5. Unwinding models

While DNA helicases are processive enzymes and can unwind thousands of base pairs without dissociating from the substrate, RNA helicases separate mainly short dsRNA regions. However, both enzymes seem to utilize the same mechanisms to unwind double-stranded nucleic acid substrates. Special features of each helicase and experimental data have allowed proposing two major helicase mechanisms: the “active rolling model” and the “inchworm model” (Fig. 11).

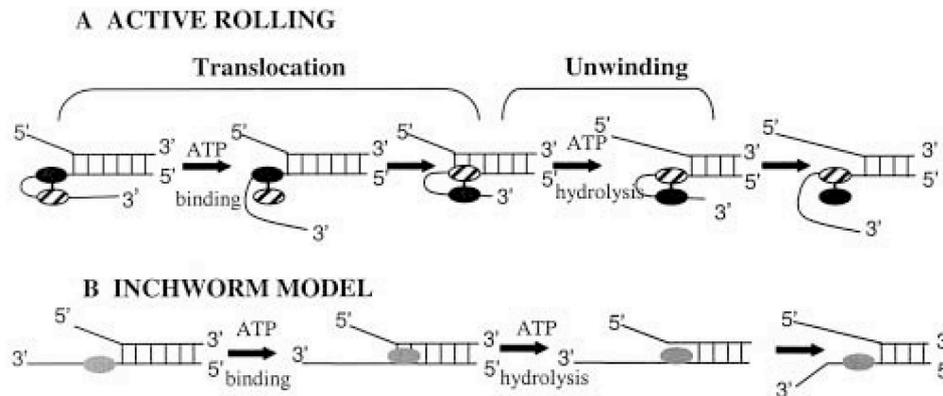
The “inchworm model” seems to better explain the mode of action of monomeric helicases while the “active rolling model” can be easier applied to multimeric enzymes. However, both situations require binding and hydrolysis of NTPs and although in some cases ATP binding seems to increase the affinity of the enzymes for duplex RNA or

DNA, it is not certain at which step ATP hydrolysis takes place (Tuteja and Tuteja, 2004).

In the “active rolling model” two distinct DNA/RNA binding domains from a multimeric enzyme alternatively move along DNA/RNA and unwind double-stranded regions. Each site could bind single- or double-stranded nucleic acids with higher affinity, depending on its conformational state. However, binding to ssDNA/RNA and dsDNA/RNA cannot occur simultaneously in the same subunit. In an initial step both domains interact with ssDNA/RNA. As a consequence of ATP binding, one subunit undergoes conformational changes and acquires higher affinity for the duplex region. Hydrolysis of ATP destabilizes the double helix and induces the release of one of the two strands. At this point a new catalytic event, in which the second subunit acts as a motor, can start (Yarranton et al., 1979).

In the “inchworm model”, two domains of a monomeric enzyme bind respectively single-stranded or double-stranded nucleic acids. In this model one domain is constantly bound to ssDNA/RNA during all the catalytic cycle, while the other one translocates along the double-stranded region in concomitant with ATP binding and hydrolysis. DsRNA/DNA destabilization and release of one of the single strand is not an active event, but happens mainly as a passive consequence of the enzyme translocation (Lohman and Bjornson, 1996).

Even if some helicases seem to fit this simple classification, for others, more complicated mechanisms must be considered. Contradictory results have been documented for the function of the viral RNA helicase NS3 (non structural protein 3) as a monomeric or oligomeric enzyme (Kuang et al., 2004). RNA helicases of the Brr2 subfamily have been shown to contain two helicase domains, although mutational analysis suggested that only one of them is functionally active (Kim and Rossi, 1999). Some replicative DNA helicases form *in vivo* hexameric complexes, characterized by a peculiar ring structure and different unwinding mechanisms have been proposed for these enzymes (Davey and O'Donnell, 2003). In addition, even though an active rolling model was first proposed for Rep unwinding ability, further studies rather supported a dimeric inchworm model, in which the two monomers are alternatively active on the dsDNA substrate (Cheng et al., 2002; Soultanas and Wigley, 2001). The further characterization of other helicases enzymatic activities should help to elucidate the molecular mechanisms that distinguish monomeric and multimeric enzymes.



**Fig. 11. Models for DNA and RNA helicase unwinding and translocation. A)** Active rolling model. The two subunits of dimeric helicase are shown as an oval shape in which one is black. The dimeric helicase unwinds by interacting directly with both dsDNA and ssDNA. Each subunit alternates binding to dsDNA as the dimer translocates when one subunit releases ssDNA and rebinds to dsDNA. In this model, translocation along ssDNA is coupled to ATP binding, whereas ATP hydrolysis drives the unwinding of multiple DNA base pairs for each catalytic event. **B)** Inchworm model. This model is consistent with the monomeric or oligomeric state of the protein (shown as an oval shape). The enzyme monomer first binds to ssDNA and then translocates along the DNA strand. It then binds to the duplex region at the fork. This is followed by unwinding and release of one of the ssDNA strand. If the enzyme is a hexamer, such as the SV40 large T antigen, then one of the subunits (monomer 1) remains associated with the fork through the unwinding cycle (Tuteja and Tuteja, 2004).

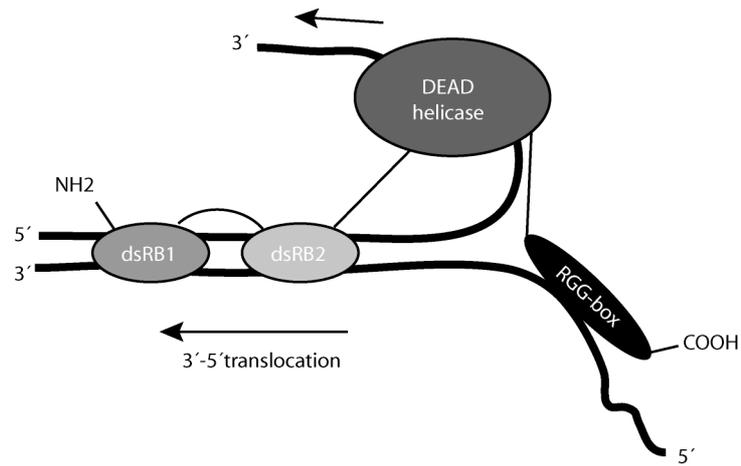
### 3.6. RNA helicase A, human homolog of MLE

RNA helicase A (RHA) has been isolated from HeLa cells by following its RNA unwinding activity (Lee and Hurwitz, 1992). Based on its extensive similarity (49% identity and 85% similarity) with MLE, it has been proposed that RHA is the human counterpart of the *Drosophila* protein (Lee and Hurwitz, 1993). Both RHA and MLE belong to the DEAH family of RNA helicases and they share with the other members the same domain organization. In addition to the central conserved helicase core, two dsRNA binding domains (dsRBM) and an RGG-box are present at the N-terminus and at the C-terminus, respectively, of both enzymes (Gorbalenya et al., 1989). MLE and RHA can unwind dsRNA and dsDNA in the 3' to 5' direction, with a higher efficiency for dsRNA. In both cases, the presence of an overhanging single-stranded nucleic acid is required for efficient unwinding (Lee et al., 1997; Lee and Hurwitz, 1992). Both

enzymes can hydrolyze all common NTPs with similar  $K_M$  values, illustrating their broad substrate specificity (Zhang and Grosse, 2004).

The enzymatic properties of MLE have been not studied in detail, whereas RHA has been extensively characterized. Biochemical studies confirmed that the isolated N-terminal dsRBMs from RHA can bind dsRNA specifically, with only a limited affinity for ssDNA or dsDNA. Moreover, both domains are required *in vitro* for efficient binding to dsRNA substrates. In contrast, the RGG domain preferentially binds single-stranded nucleic acids and in particular ssDNA. Interestingly, neither the deletion of the two dsRBMs nor the deletion of the RGG-box abolishes RHA helicase activity, although in both cases a decrease in the nucleic acid-stimulated ATPase activity of RHA has been observed (Zhang and Grosse, 1997). These results suggest that all three domains might be neighbors in the three-dimensional structure of RHA (Fig. 12). In this model cooperation between the RGG-box bound to ssDNA/RNA and one or both the dsRBMs docked to dsRNA/DNA, are required in order to stimulate RHA ATPase activity and subsequent unwinding of the nucleic acid substrate (Lee and Hurwitz, 1993).

Interestingly, the dsRB1 of RHA has been shown to cooperate with its C-terminal proline-rich region to form an integrated nucleic acid-binding module specific for the consensus dsDNA sequence of the cyclin-AMP responsive element (CRE) (Hung et al., 2003). This result was unexpected, as the crystal structure suggested that dsRNA binding domain cannot accommodate either the A form of RNA or the B form of DNA (Fierro-Monti and Mathews, 2000; St Johnston et al., 1992). However, variations in the common dsRBM fold have been observed in some DNA binding proteins (Connolly et al., 1998). In addition, the RNA binding affinity of dsRBMs is extremely variable and some of them are even RNA binding incompetent (Romano et al., 1995). It has been proposed that the linker region separating the dsRB1 and dsRB2 in RHA, due to its flexibility, might cause conformational changes in RHA and determine the formation of a versatile platform with different ligand binding properties.



**Fig.12. Hypothetical scheme showing how the various modules of RHA may fit to a forked RNA substrate.** The helicase domain is shown binding the 3' single-stranded region, which it pulls away from the other strand, and along which it translocates. Models in which the C-terminal domain binds to the 3' strand instead of the 5' are also plausible (modified from (Gibson and Thompson, 1994)).

#### **4. Aim of this work**

While the other MSL proteins have been studied to some degree, the role of MLE in dosage compensation so far has been only demonstrated genetically. A male-specific phenotype has been observed in transgenic flies deleted for *mle* (Fukunaga et al., 1975). In addition, the MLE<sup>GET</sup> ATPase-inactive mutant localizes only to few “high affinity” sites on the X chromosome when overexpressed in flies, suggesting that MLE enzymatic activities are required for its proper targeting (Lee et al., 1997). Contradictory observations regarding the roX RNA-dependent integration of MLE into the DCC complex are present in the literature. In transgenic flies expressing the MLE<sup>GET</sup> mutant neither roX1, nor roX2 can be detected on the X, suggesting that an MSL complex containing MLE, but lacking RNA can still be assembled *in vivo* and targeted to a limited number of sites (Gu et al., 2000). However, MLE is lost from the X territory upon Rnase treatment of SL2 cells (Akhtar and Becker, 2000). Moreover, MLE can be copurified from embryo extracts with the other MSLs only under low salt condition and when special precautions against RNA degradation are taken (Scott et al., 2000). These results suggest that association of MLE with the DCC members is RNA-dependent. These conclusions are mainly indirect; final proof of a direct interaction of MLE with the other MSL does not exist.

One of the aims of this work is to better understand how MLE interacts with the other DCC members. For this purpose the baculovirus system has been employed to express and purify the different DCC components.

In addition, a proper biochemical characterization of MLE is missing. MLE ATPase activity has been shown to be strongly stimulated by dsRNA and to a lower extent by DNA (Lee et al., 1997). Moreover, although MLE binds better to ssRNA as compared to ssDNA, it is not known if it is able to discriminate between ssRNA and dsRNA substrates. As a member of the DEAH family of RNA helicase MLE harbours two RNA binding domains at the N-terminus and an RGG-box at the C-terminus (Gibson and Thompson, 1994). These regions have been predicted to bind dsRNA and ssRNA respectively, based on the homology of MLE with other enzymes, for which these properties have been described.

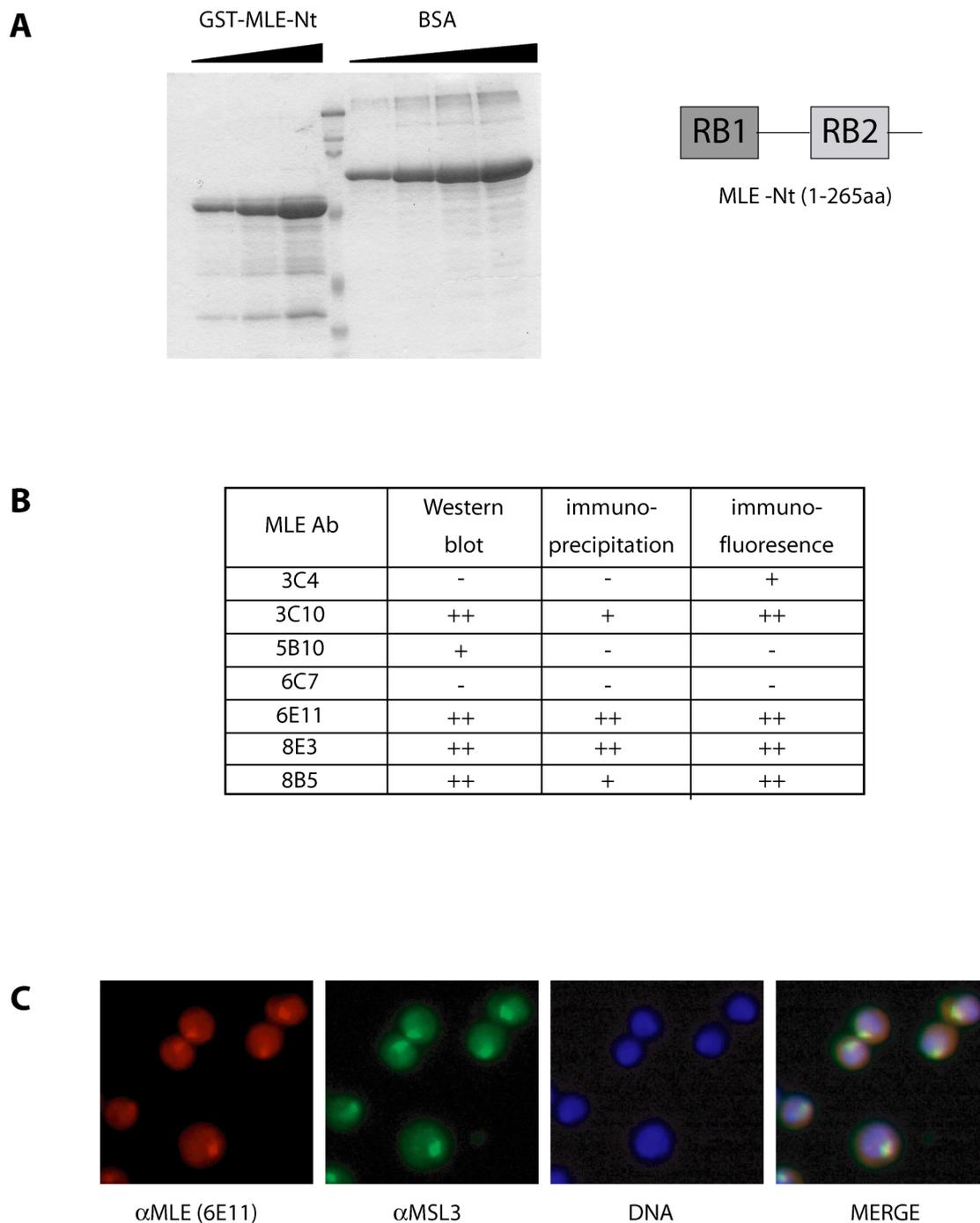
To elucidate the RNA binding properties of the different MLE domains and the mechanism of action of MLE RNA helicase, MLE deletion mutants have been expressed in Sf9 cells and tested in several functional *in vitro* assays. Finally, transient

transfection in the SF4 cells of MLE wild type and the different deletion mutants fused to GFP have been used to describe those parts of MLE required for its correct targeting to the X chromosome.

## II. Results

### 1. MLE monoclonal antibodies

In order to characterize MLE, rat monoclonal antibodies were generated. The first 265 aa of MLE were fused to glutathione S-transferase (GST) and expressed in bacterial cells. The recombinant GST-MLE (1-265) was affinity purified on glutathione beads and the eluted material quantified on SDS-PAGE by Coomassie staining (Fig. 13, A). The following steps were carried out by Dr. E. Kremmer (GSF, Martinsried). Purified GST-MLE (1-265) was injected in rats to induce antibody production. Spleen cells isolated from immunized animals were fused with myeloma cells to generate immortal hybridoma cells. From a first screen of each hybridoma culture, seven supernatants contained antigen-specific antibodies and they were then tested in different assays for their specificity (Fig 13, B). When used in Western blot probing *Drosophila* SF4 cell extracts or in immunoprecipitation assays, supernatants 8B5, 3C10, 5B10, 8E3 and 6E11 recognize a protein with a size that matches that of MLE. However, 5B10 also strongly reacts with additional proteins and 3C10 clone immunoprecipitates MLE less efficiently, when compared to the other antibodies (summary in Fig. 13, B). In immunofluorescence experiments on SF4 cells, antibodies against each of the MSL proteins stain a small area of the nucleus that corresponds to the X chromosome territory. In these assays, 8B5, 3C4, 3C10, 8E3 and 6E11 supernatants recognize a nuclear protein that colocalizes with MSL3 on the X chromosome and corresponds with high probability to MLE (Fig. 13, C). In addition, a distinct punctuate staining in the nucleus can also be observed. These results are in agreement with published immunofluorescence experiments on polytene chromosomes from male larvae, where MLE, in addition to the X chromosome, also localizes at different autosomal positions. The two hybridoma 6E11 and 8E3 were further subcloned by Dr. Kremmer in order to obtain monoclonal antibodies against MLE. The specificity of 6E11 and 8E3 subclones was tested *in vivo* after depletion of endogenous MLE from SF4 cells using RNA interference (RNAi, (Celotto and Graveley, 2004)). After 6 days of MLE-dsRNA treatment (see 2.1 and Fig. 14, B), only background staining in the nucleoplasm can be observed with both antibodies. These results confirmed that the protein recognized by 6E11 and 8E3 monoclonal antibodies and colocalizing with MSL3 on the X chromosome is indeed MLE.



**Fig. 13. Generation of  $\alpha$ MLE monoclonal antibodies.** **A)** Expression and purification of recombinant GST-MLE (1-265). The N-terminal region of MLE (1-265 aa), fused to GST, was expressed and purified from bacterial cells. 0.25  $\mu$ l, 0.5  $\mu$ l and 1  $\mu$ l of the eluted fractions were loaded on a 8% polyacrylamide gel and compared to 125 ng, 250 ng, 0.5  $\mu$ g and 1  $\mu$ g of BSA by Coomassie staining. **B)** Summary of the results from Western blot, immunoprecipitation and immunofluorescence experiments, using the different MLE hybridoma supernatants. **C)** Immunofluorescence staining of SF4 cells. MLE 6E11 monoclonal antibodies were used to detect MLE ( $\alpha$ MLE (6E11)) and MSL3 polyclonal antibodies were used to mark the X territory ( $\alpha$ MSL3). DNA was counterstained with Hoechst 33258 (DNA).

### 1.1. Purification of MLE interacting partners

The 6E11 clone was used in preparative immunoprecipitation experiments, in order to identify new proteins, interacting with MLE (T. Straub and data not shown). Among the DCC members, MLE is the only protein that was shown to have an additional function not related to dosage compensation (Kenan et al., 1991). In order to distinguish male specific proteins, nuclear extracts from *Drosophila* male SF4 cells and female Kc cells were compared under the same conditions. A band of approximately 80 KDa could be specifically coimmunoprecipitated with MLE from SF4 cells but not from Kc cells. By mass-spectrometry this protein was identified as the DEAD-box RNA helicase Belle. Belle had been shown to be required for viability in both male and female flies and may have a role in translation initiation (Johnstone et al., 2005). The fact that another RNA helicase can interact with MLE is not surprising as many of these enzymes are integrated in multimeric complexes, where additional cofactors modulate their activity or target helicases on a specific RNA substrate. RNAi experiments were employed to test the possibility that Belle is directly involved in dosage compensation in males. However, upon ablation of Belle in SL2 cells, no effect on the correct targeting of the MSL proteins to the X territory could be observed.

These results suggest that Belle, unlike MLE (see 2.1.), is not required to assemble the DCC, but it might act downstream of the DCC or be required for an MLE function outside of the context of dosage compensation.

### 1.2. Limitations of 6E11 antibodies

Interestingly, 6E11 monoclonal antibodies specifically recognize an epitope within the N-terminal dsRB1 of MLE, as they fail to detect both MLE<sup>ΔRB1</sup> and MLE<sup>ΔRB1-2</sup> mutants in Western blot analysis (data not shown). In addition, none of the MSL proteins could be identified by mass-spectrometry analysis in the fraction immunoprecipitated by 6E11 clone from SF4 cells. The same antibodies also failed to copurify the DCC complex from Sf9 cells coinfecting with the respective MSL baculoviruses, whereas recombinant MLE-flag, under the same conditions, successfully pulled down all the DCC members. Taken together these observations suggest that the epitope within the dsRB1 is available in free MLE and hidden when MLE is integrated into the DCC complex.

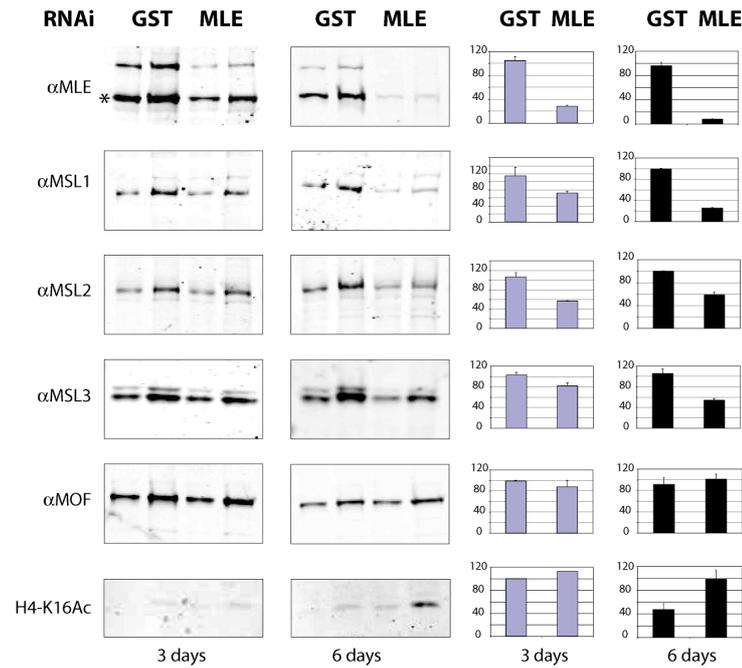
## **2. MLE and the dosage compensation complex (DCC)**

### **2.1. Effects of MLE depletion in SL2 cells**

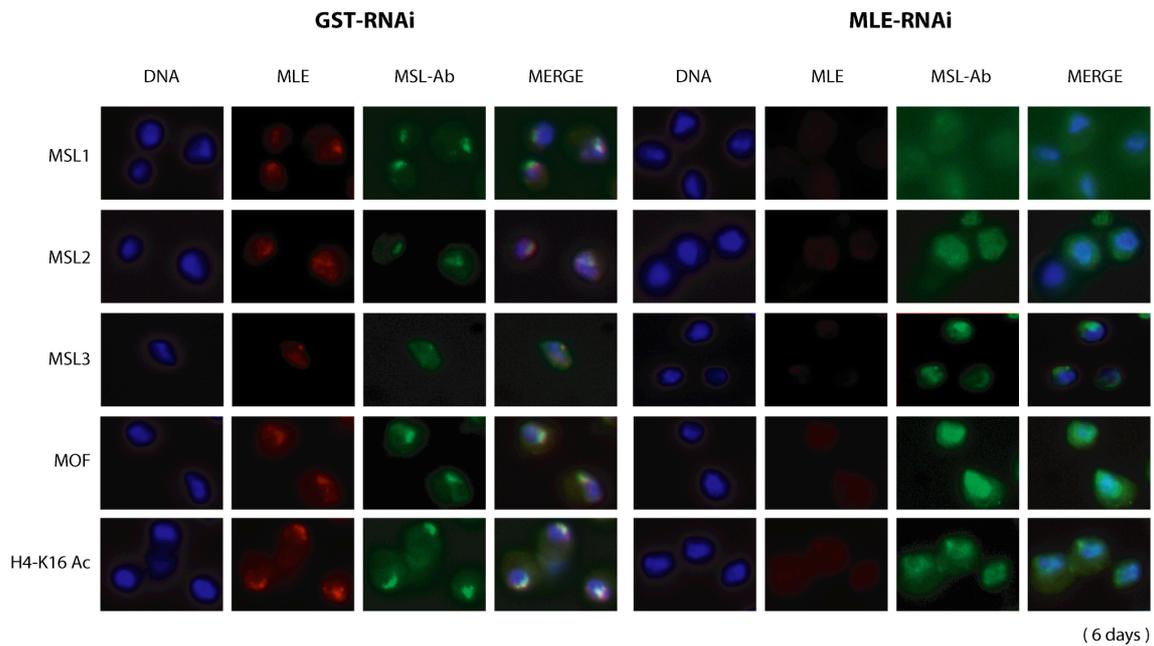
In *mle* null flies, only MSL1 and MSL2 are recruited to few “high affinity” sites on the X chromosome, suggesting a role for MLE in the first steps of the DCC assembly (Meller, 2003). In order to test if the same situation can be reproduced in cells and to generate a tool for studying MLE function *in vivo*, SL2 cells were depleted of endogenous MLE by RNAi.

A dsRNA, spanning the first 700 bp of the *MLE* cDNA was used and as a control for non-specific effects of dsRNA, cells were treated in parallel with a dsRNA corresponding to the sequence of GST. 3 and 6 days after addition of dsRNA, the relative amounts of each protein in total extract from GST- and MLE-dsRNA treated cells were quantified by Western blot (Fig. 14, A). After 3 days of dsRNA treatment, 70% MLE depletion was reached, while only a slight decrease in the amounts of the other MSLs was observed, with MSL1 and MSL2 being the most affected. After 6 days of treatment, also the MSL3 level start to decrease, consistent with its dependency on MSL1 and MSL2 for stability (Chang and Kuroda, 1998; Rastelli et al., 1995). Surprisingly, while almost no differences in the levels of MOF and of H4-K16 acetylation were observed after 3 days of MLE RNAi, after 6 days a reproducible increase in H4-K16 acetylation was observed. The relevance of these data was checked *in vivo* by immunofluorescence experiments (Fig. 14, B). 6 days after dsRNA treatment of the control cells the active MSL complex is on the X chromosome, as shown by the colocalization of the MSL proteins and by the specific presence of H4-K16 acetylation on the X territory. In contrast, in the absence of MLE, none of the MSL proteins could be detected on the X and also H4-K16 acetylation was lost from the X chromosome. The same situation was observed after 3 days of dsRNA treatment, with a lower percentage of cells depleted of endogenous MLE (data not shown). While for MSL1 and MSL2 only a homogeneous background staining in both nucleus and cytoplasm could be observed, MSL3, MOF and H4-K16 acetylation were still detected in the nucleus, although delocalized from the X territory (Fig. 14, B).

A



B



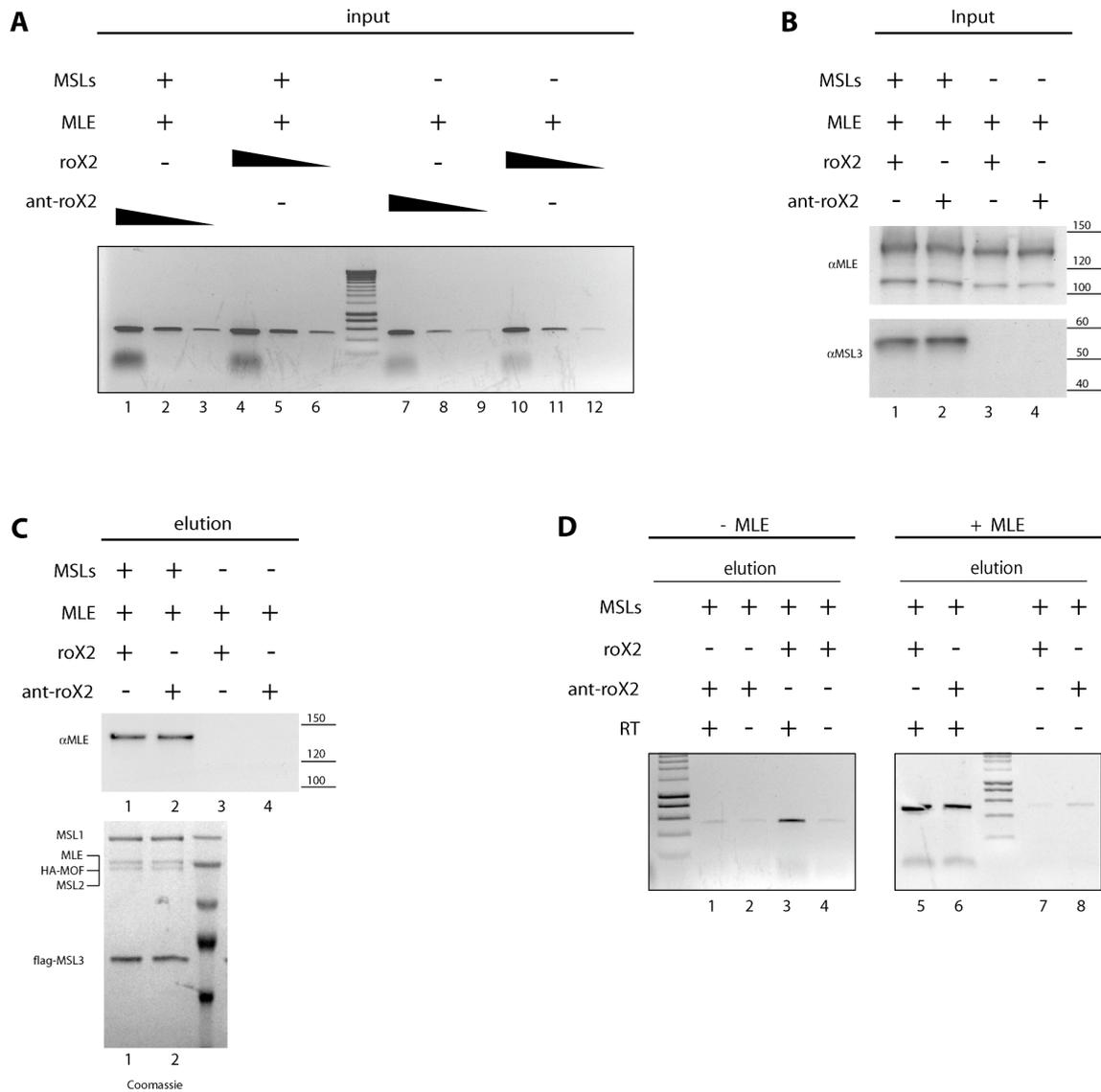
**Fig. 14. Effect of MLE depletion in SL2 cells. A)** Western blots with specific  $\alpha$ MSL and  $\alpha$ H4-K16 acetylated antibodies on total cell extracts after 3 and 6 days of RNAi were quantified using the Odyssey Infrared Imaging system. The amount of each protein in the MLE-RNAi treated cells is expressed as percentage of the corresponding amount in the GST-RNAi control cells. Error bars correspond to the values of two loadings. In  $\alpha$ MLE Western blot the star indicates endogenous MLE protein. **B)** Immunofluorescence on SL2 cells after 6 days of dsRNA treatment against GST (GST-RNAi) or MLE (MLE-RNAi). The cells were stained for all MSL proteins (MSL1, MSL2, MSL3, MOF and MLE) as well as for acetylation of H4-K16 (H4-K16ac). DNA was counterstained with Hoechst 33258.

These results show that like in flies, also in the cells MLE is required for proper targeting of the DCC complex to the male X. However, in the absence of MLE and at limiting amounts of MSL1, it seems that an active MOF is redistributed in the nucleus and H4-K16 acetylation is not anymore confined to the X chromosome.

## 2.2. Reconstitution of a DCC complex specifically containing roX2 RNA

Even though MLE RNA helicase is an essential component of dosage compensation, it cannot be copurified from crude embryo extracts with the other MSL proteins under stringent salt conditions (Scott et al., 2000). In addition, in transgenic flies deleted for both *roX1* and *roX2* genes, MLE is no longer present at any of the X chromosomal sites (Meller and Rattner, 2002). Binding of MLE to the X chromosome is also lost in RNase-treated SL2 nuclei (Akhtar et al., 2000). These results suggest that MLE associates only weakly with the other MSLs and that its recruitment to the X chromosome is likely roX RNA-dependent. However, genetic experiments in flies demonstrated that MLE may still colocalize with MSL proteins in the absence of roX RNAs at a few “high affinity” sites on the X chromosome (Gu et al., 2000).

To better understand how roX2 RNA influences association of MLE with the DCC, all known complex components were coexpressed in Sf9 cells using the baculovirus-expression system. In addition to the already available baculoviruses encoding MSL1, MSL2, flag-MSL3 and HA-MOF, two viruses expressing MLE and roX2 RNA, respectively, were created. As a control for MSL-specific RNA binding, a virus expressing the antisense roX2 RNA (ant-roX2) was also generated. Sf9 cells were coinfecting with up to six viruses, thus expressing all the known subunits of the complex with either roX2 or ant-roX2 RNA. The amount of roX2 and ant-roX2 viruses was accurately titrated in order to express comparable levels of both RNAs (Fig. 15, A). Correct expression of the MSL proteins in the presence of either roX2 or ant-roX2 RNAs was also controlled by Western blot (Fig. 15, B and data not shown).



**Fig 15. MSL proteins specifically interact with roX2 RNA.** Sf9 cells were coinfecting with baculoviruses encoding MSL1, MSL2, flag-MSL3, HA-MOF (MSLs) and MLE, roX2 RNA or ant-roX2 RNA in the combination indicated with (+) above each lane. RNA-protein complexes were purified from total Sf9 extracts using flag affinity beads. **A)** RT PCR on the input. **B)** Western blot on the input with  $\alpha$ MLE and  $\alpha$ MSL3 antibodies. **C)** Western blot with  $\alpha$ MLE antibodies (top) and Coomassie blue stained SDS-PAGE of the eluted material (bottom). **D)** RT PCR on the eluted material.

In order to avoid dissociation of MLE from the other MSLs, low salt conditions were used to prepare total Sf9 extracts, two days after infection. As previously reported, MSL1, MSL2, MSL3 and MOF could be copurified from infected Sf9 cells via flag-MSL3 (Morales et al., 2004) and association of these proteins in a four subunit complex, can be visualized on a Coomassie blue stained gel (Fig. 15, C, bottom). Because MLE has

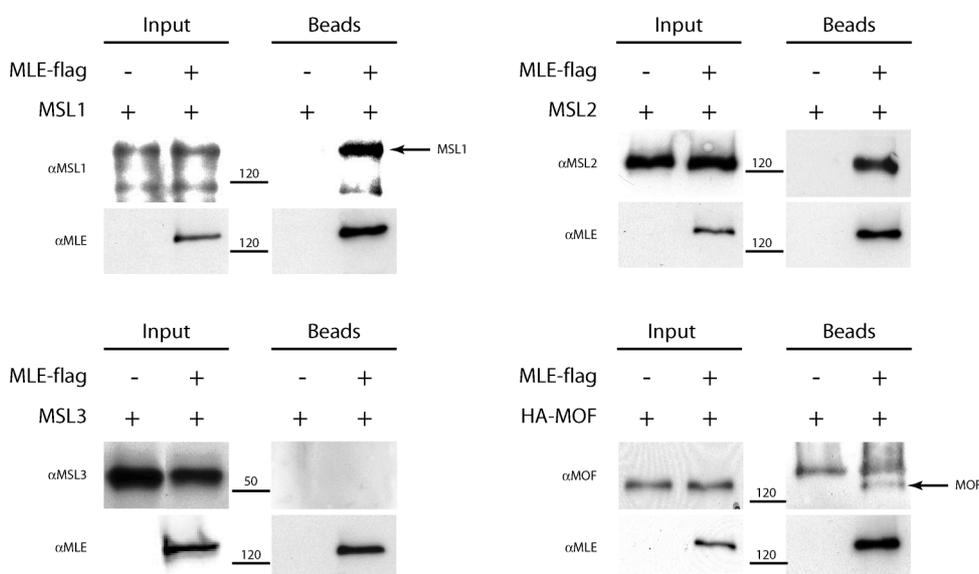
a similar size to HA-MOF, its association with the other MSLs can be only confirmed by Western blot (Fig. 15, C, top).

In the absence of MLE, the purified complex specifically coimmunoprecipitates roX2 RNA. By contrast, in the presence of MLE, both roX2 and ant-roX2 RNAs associate to the complex (Fig. 15, D). These results suggest that MSL proteins can specifically interact with roX2 RNA, while MLE seems to bind RNA non-specifically. RoX2 and ant-roX2 RNAs have the same length and GC content, but they are probably folded differently. Specific features of roX2 RNA, might be then recognized by MSL3 and MOF, the other two RNA binding proteins of the DCC. In addition, even if MSL proteins do not bind ant-rox2 and integration of this RNA into the DCC complex is strictly MLE-dependent, comparable amounts of MLE can be copurified with the other MSL proteins in the presence of either roX2 or ant-roX2 RNAs (Fig. 15, D, top). Taken together, these results suggest that association of MLE with the DCC complex may be possible in the absence of roX2 RNA.

### 2.3. MLE preferentially associates with MSL1 and MSL2

Since MLE can associate with the DCC complex in the absence of roX2 RNA, its direct association with the other MSL proteins was analyzed. In order to have comparable conditions, MLE was flag-tagged at the C-terminus and coexpressed in Sf9 cells with individual untagged MSL proteins. Partial complexes were purified via anti-flag affinity beads and associated proteins were visualized by Western blot (Fig. 16).

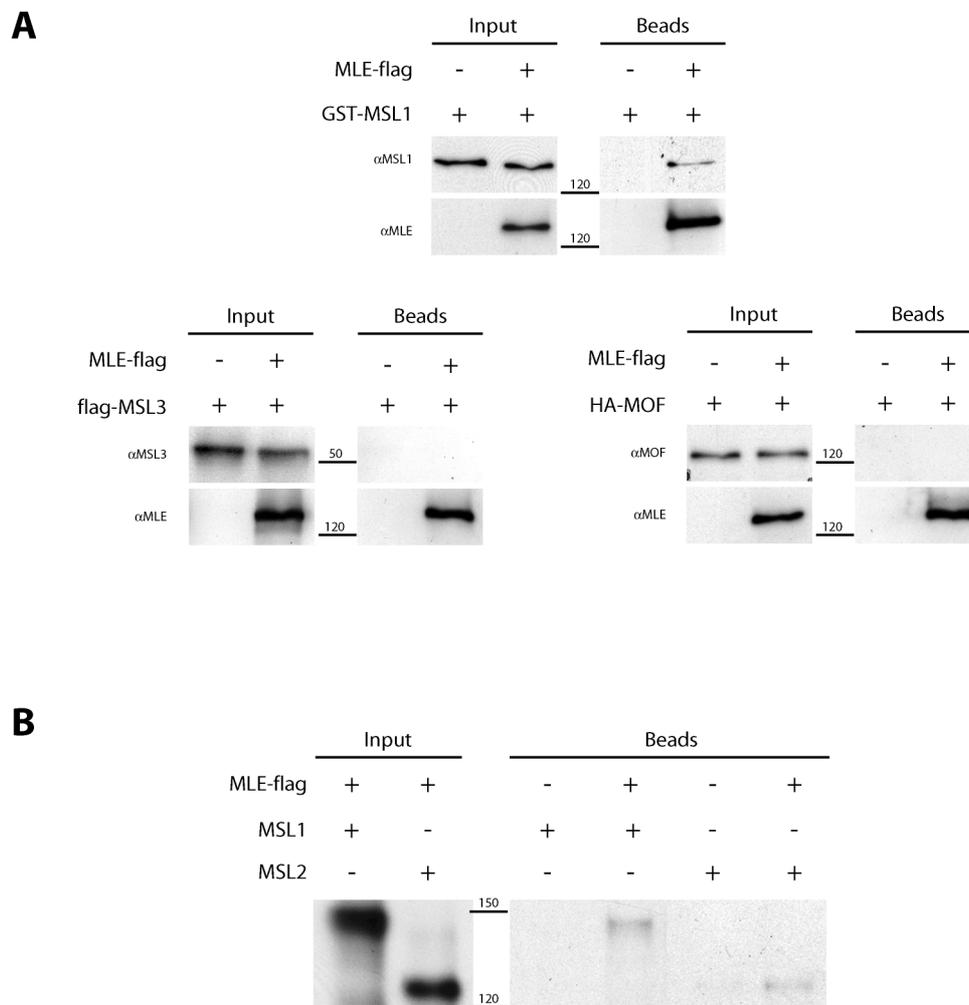
MSL1 and MSL2 associate with MLE independently, whereas a weaker interaction with MOF and no association with MSL3 were observed. However, binding of MSL1 or MSL2 to MLE is rather weak, since all proteins are expressed at comparable levels (data not shown), but only a small fraction (4-5%) associates with MLE.



**Fig. 16. MLE preferentially associates with MSL1 and MSL2.** Sf9 cells were infected with baculoviruses expressing MLE-flag in combination with individual MSL protein (+). MLE-MSL complexes were purified via anti-flag affinity beads. 2,5% of the input (for the MSLs), 0,5% of input (for MLE) and 50% of the beads were loaded on a 8% polyacrylamide gel and subjected to Western blot analysis with specific MSL antibodies ( $\alpha$ MSL1,  $\alpha$ MSL2,  $\alpha$ MSL3,  $\alpha$ MOF,  $\alpha$ MLE).

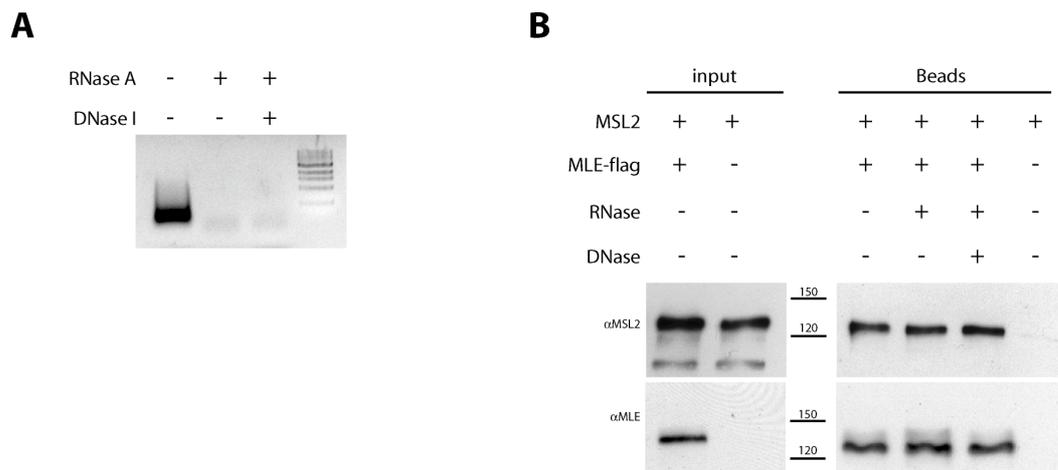
Interaction between MLE and MSL1 was confirmed by *in vitro* pull-down experiments, using recombinant MLE-flag and GST-MSL1, independently expressed and purified from Sf9 cells. No interaction of MLE with both HA-MOF and flag-MSL3 was observed under the same conditions (Fig. 17, A). In addition, *in vitro* translated MSL1

and MSL2 also specifically interact with MLE (Fig. 17, B). These results show a preferential association of MLE with MSL1 and MSL2.



**Fig. 17. MLE directly interacts with MSL1.** **A)** MLE-flag, GST-MSL1, flag-MSL3 and HA-MOF were independently purified from Sf9 cells. 100 ng of each protein were used in *in vitro* pull-down experiments. Due to the presence of the flag tag in both MLE and MSL3, MLE monoclonal antibodies (6E11) were used in the pull-down experiments. 10% of input and 50% of the beads were loaded on an 8% polyacrylamide gel and subjected to Western blot analysis with specific MSL antibodies ( $\alpha$ MSL1,  $\alpha$ MSL3,  $\alpha$ MOF,  $\alpha$ MLE). **B)** *In vitro* pull down experiment using MLE-flag purified from Sf9 and MSL1 and MSL2 translated *in vitro* in the presence of [ $^{35}$ S]. After immunoprecipitation with flag affinity beads, 10% of input and 50% of beads were loaded on an 8% polyacrylamide gel, followed by autoradiography.

Since MSL1 it is not known to bind RNA and stringent salt conditions were used to purify MLE-flag, the interaction between these two proteins is probably direct. In addition, the ATPase activity of MLE is strongly stimulated by RNA (Lee et al., 1997), but no enzymatic activity can be detected when recombinant MLE-flag alone is used in the ATPase assay, suggesting that no RNA was copurified (Fig. 19). The same conclusion cannot be derived for the interaction with MSL2, as RNA is present in Sf9 cells as well as in the *in vitro* translation reaction. To rule out the possibility that non-specific RNA mediates interaction between MLE and MSL2, extracts from cells infected with MLE-flag and MSL2 baculoviruses were treated with RNase A and/or DNaseI. Efficient depletion of nucleic acids was confirmed (Fig. 18, A). However, under these conditions, MSL2 still interacts with MLE and no effects on the amounts of MSL2 protein copurified are observed (Fig. 18, B). These data show that MLE, albeit weakly, also directly interacts with MSL2.



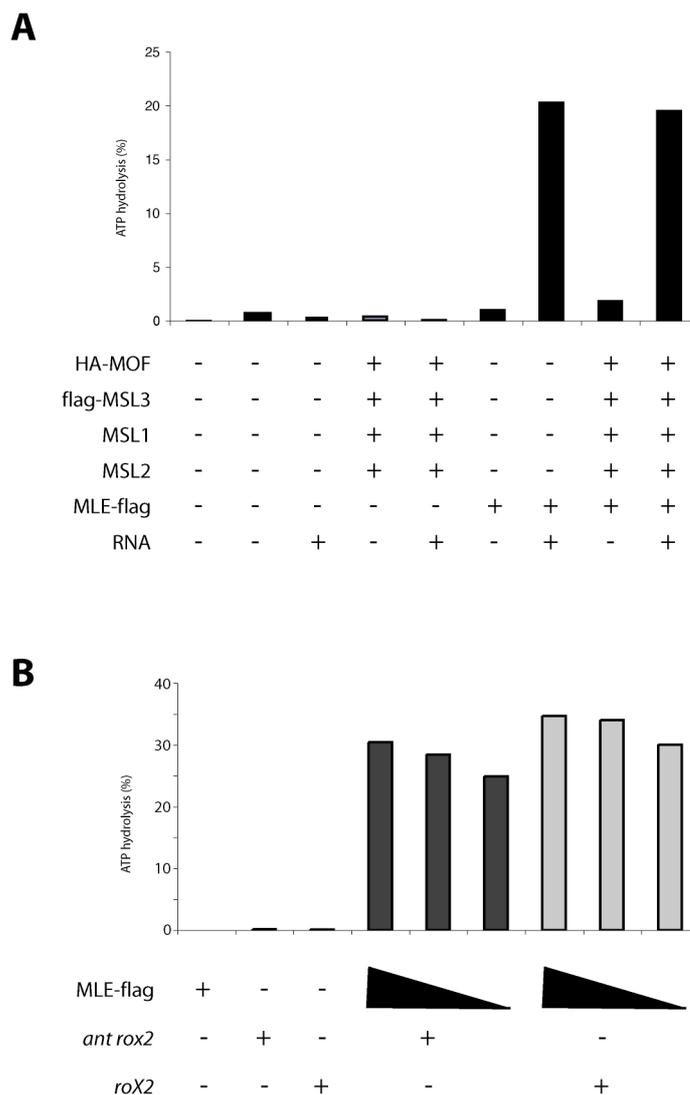
**Fig. 18. MLE directly interacts with MSL2.** **A)** Total nucleic acids extracted from Sf9 cell extracts infected with MLE-flag and MSL2 encoding baculoviruses and treated (+) or not (-) with RNase A and DNaseI were loaded on a 1% agarose gel. **B)** MLE-MSL2 complexes were purified via flag affinity beads and association of MSL2 with MLE was analyzed by Western blot using specific  $\alpha$ MSL2 (upper panel) and  $\alpha$ MLE (lower panel) antibodies.

#### 2.4. MSL proteins do not influence the ATPase activity of MLE

In a previous study, MOF was shown to specifically acetylate H4-K16 on reconstituted nucleosome arrays only when both MSL1 and MSL3 were present (Morales et al., 2004). These results suggest that MSL interactions have an important role in regulating MOF enzymatic activity, confining the H4-K16 acetylation mark specifically to the male X chromosome. MLE possesses both ATPase and helicase activities, which are required for dosage compensation (Gu et al., 2000). Since interaction of MLE with MSL1 and MSL2 was observed, the possibility that these proteins affect MLE enzymatic properties was explored.

Only small amounts of MSL1, MSL2, MSL3 and MOF can be copurified via MLE-flag from Sf9 cells (data not shown). Therefore, HA-MOF and MLE-flag were purified separately whereas MSL1, MSL2 and flag-MSL3 were expressed and copurified. Equimolar amounts of each protein were used in the ATPase assay. To stimulate the ATPase activity of MLE, total RNA, extracted from Sf9 cells, was used as a substrate in the ATPase reaction. The percentage of ATP hydrolyzed by MLE was measured in the presence or absence of the other MSL proteins. As previously shown, the ATPase activity of MLE was strongly stimulated by RNA. However, when all MSLs were present together in the same reaction mixture, no effect on the capacity of MLE to hydrolyze ATP was observed (Fig. 19, A). These results strongly suggest that MSL proteins do not influence the ATPase activity of MLE.

Although MLE binds roX2 and ant-roX2 RNAs indiscriminately, the possibility that specific secondary structures within roX2 RNA differently affect MLE enzymatic activities were also tested in the ATPase assay (Fig. 19, B). Total RNA was prepared from Sf9 cells infected with roX2 and ant-roX2 baculoviruses, respectively, and the same amounts of RNA were used to stimulate MLE ATPase activity. MLE did not show any preference for roX2 versus ant-roX2 RNAs as a substrate, suggesting that it probably binds RNA non specifically. Because MLE helicase activity is strictly dependent on its ATPase activity, the capacity of MLE to unwind double-stranded nucleic acids is not expected to be influenced by its interaction with the other MSL proteins.



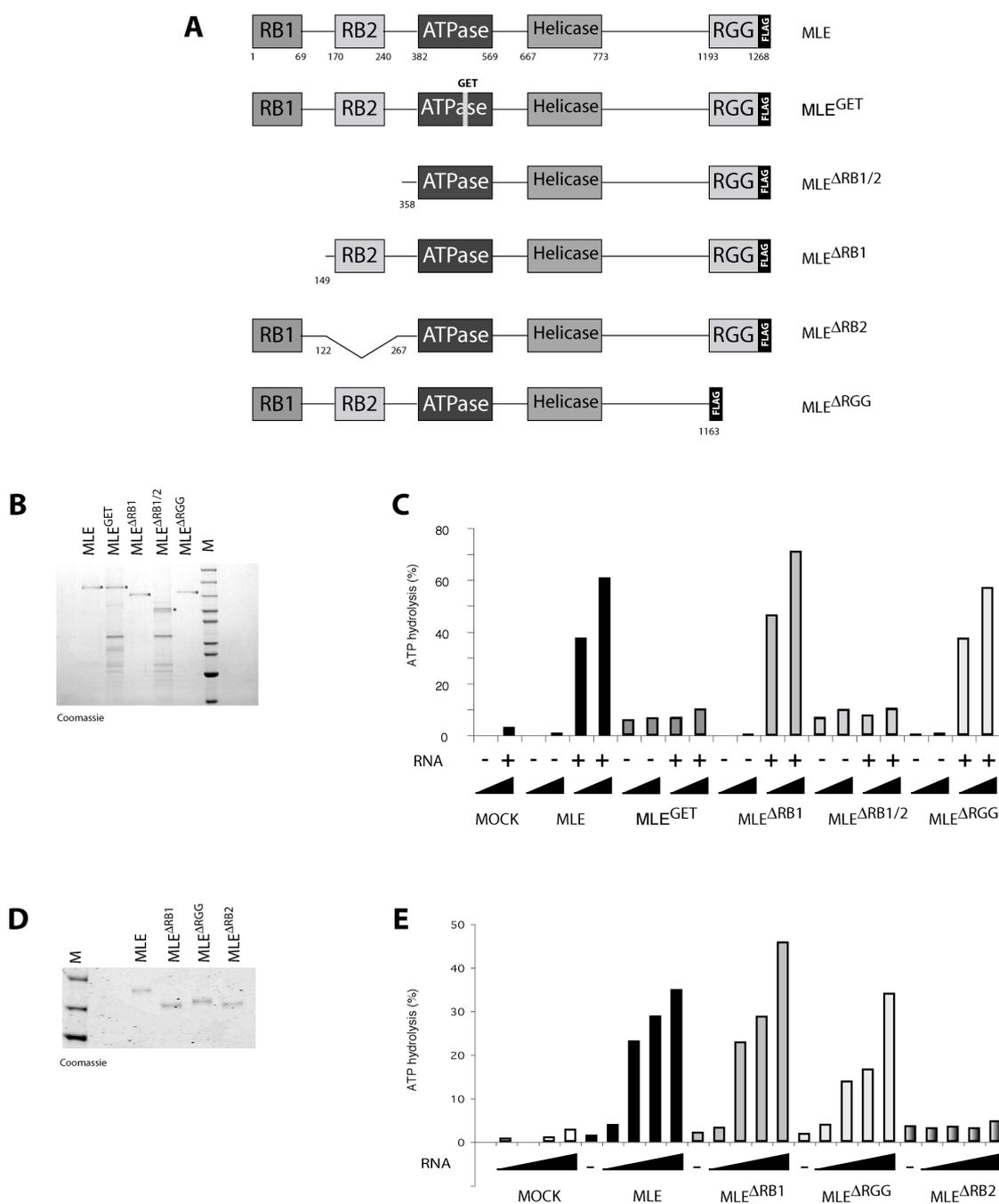
**Fig. 19. Effects of MSL proteins and roX2 RNA on MLE ATPase activity.** **A)** MSL1, MSL2 and flag-MSL3 were copurified from Sf9 cells. HA-MOF and MLE-flag were purified independently from Sf9 cells. Equimolar amounts (30 fmoles) of each protein were used in the ATPase assay in the combination indicated by the (+). 1  $\mu$ g of total RNA extracted from Sf9 cells was used to stimulate MLE ATPase activity in the presence of saturating amounts of  $[\gamma^{32}\text{P}]\text{-ATP}$ . After 30 min of incubation at 26  $^{\circ}\text{C}$ , an aliquot of each reaction was analyzed by thin layer chromatography (TCL) and quantitated by PhosphoImager analysis. **B)** RoX2 and ant-roX2 were expressed in Sf9 cells. Comparable amounts (400ng) of total RNA extracted from these Sf9 cells were used to stimulate MLE ATPase activity.

### **3. Biochemical characterization of MLE RNA helicase**

#### **3.1. Effects of RNA binding domain deletions on MLE ATPase activity**

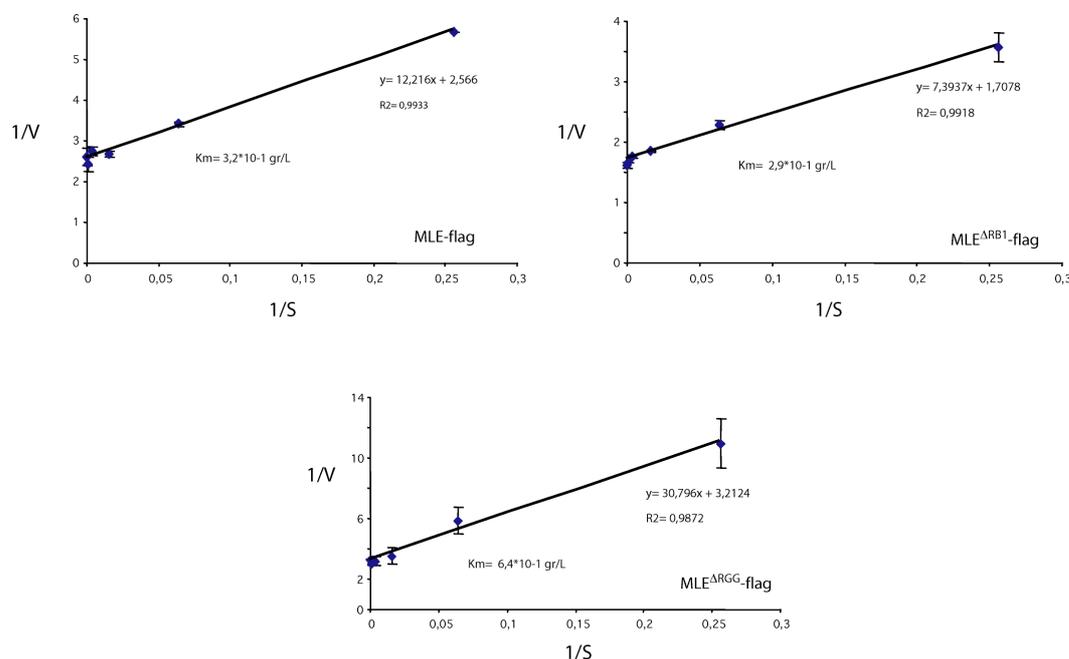
Mutational analysis and RNA binding studies revealed a weak RNA binding capacity by the common motif VI in the conserved DEAD/DExH domain of many RNA helicases (Cordin et al., 2005). However DEAD/DExH enzymes also present additional RNA binding domains outside of the conserved core region (Pause et al., 1993). MLE, as well as other members of the same DEAH-box RNA helicase family, harbors two dsRNA binding domains (dsRB1 and dsRB2) at the N-terminus and a predicted ssRNA/ssDNA binding domain at the very C-terminus. This region, highly enriched in arginines and glycines, was named the RGG-box. Nevertheless the RNA binding properties of these additional domains were characterized only for some of these enzymes.

In order to define how the proximal and distal RNA binding domains of MLE influence its enzymatic activity, viruses expressing different MLE deletion mutants were produced. The corresponding recombinant proteins were expressed and purified from Sf9 cells via the fusion of a flag-tag at their C-terminal ends (Fig. 20, A). Comparable amounts of each protein were tested in the ATPase assay for their ability to hydrolyze ATP in the presence of total RNA extracted from Sf9 cells (Fig. 20 B, D). As previously shown, the MLE ATPase activity is strongly stimulated by RNA, whereas no enzymatic activities are associated to the MLE<sup>GET</sup> ATPase-inactive mutant ((Lee et al., 1997) and fig. 20, C). Deletion of the first dsRNA binding domain (MLE<sup>ARB1</sup>) as well as deletion of the C-terminal RGG domain (MLE<sup>ARGG</sup>) does not affect the ATPase activity, and both proteins hydrolyze ATP with an efficiency comparable to that of the wild type protein. By contrast, the MLE<sup>ARB1-2</sup> mutant protein does not hydrolyze ATP in the presence of RNA (Fig. 20, C). This protein, as well as the MLE<sup>GET</sup> point mutant is expressed in Sf9 cells at low levels and several additional bands can be visualized on Coomassie-stained gels after flag purification (Fig.20, B). However, the presence of equimolar amounts of MLE<sup>GET</sup> or MLE<sup>ARB1-2</sup> do not affect MLE ATPase activity, excluding the possibility that the contaminant factors copurified are responsible for the absence of enzymatic activity observed for both mutant proteins (data not shown). To test if dsRB1 and dsRB2 have a redundant function for MLE RNA-dependent ATPase activity, an additional mutant lacking only the dsRB2 (MLE<sup>ARB2</sup>) was tested in the ATPase assay.



**Fig. 20. MLE RNA binding domains differently affect MLE ATPase activity.** **A)** Schematic representation of MLE and MLE mutants. **B), D)** Two amounts of each purified protein (2.5 ng and 5 ng) were compared on a Coomassie-stained SDS-PAGE. Stars indicate MLE and the respective MLE mutants. **C)** ATPase activity of recombinant MLE and the respective mutants was compared. Two amounts of each protein (25 and 50 fmoles) were incubated in the absence or in the presence of saturating amounts of total Sf9 RNA (1  $\mu$ g) and [ $\gamma$ <sup>32</sup>P]-ATP. After 30 min at 26 °C, ATP hydrolysis was measured as describe in Material and Methods. The ATPase activities are indicated as a percentage of hydrolyzed ATP. **E)** Constant amounts of MLE and the respective mutants (2.5 ng) were incubated in the absence or in the presence of decreasing amounts of RNA (1  $\mu$ g, 250 ng, 62 ng, 16 ng) and saturating amounts of [ $\gamma$ <sup>32</sup>P]-ATP. After 20 min at 26 °C, ATPase activity was measured as above.

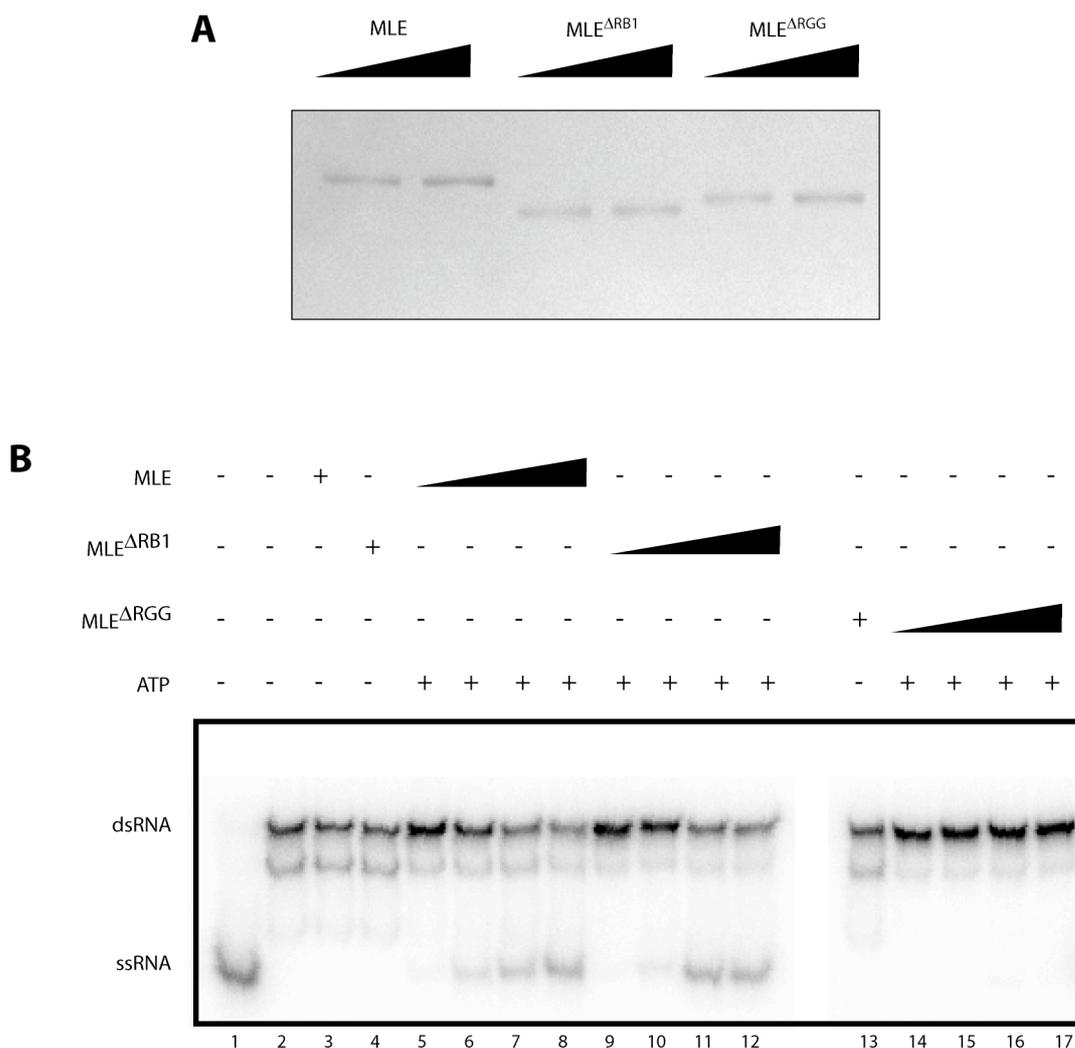
Deletion of the dsRB2 completely abolished the capacity of the corresponding protein to hydrolyze ATP (Fig. 20, E). This result suggests that the dsRB2 is essential for RNA-induced MLE ATPase activity. Deletion of dsRB2 might affect the correct folding of the neighboring ATPase domain and, consequently, the ATPase activity. Interestingly, in the presence of limiting amounts of RNA, the MLE<sup>ΔRGG</sup> mutant seems to hydrolyze ATP less efficiently than wild type MLE and MLE<sup>ΔRB1</sup> (Fig. 20, E). To determine the Michaelis-Menten kinetic parameters of MLE and the different mutant enzymes decreasing amounts of RNA at saturating amounts of [<sup>32</sup>P]-ATP were used (Fig. 21). The Lineweaver-Burk representation enabled the determination of the affinity constant ( $K_M$ ) of MLE and the deletion mutant enzymes for RNA. Similar  $K_M$  values were obtained for MLE and MLE<sup>ΔRB1</sup>, while MLE<sup>ΔRGG</sup> shows a higher  $K_M$ , revealing a lower affinity for RNA. All together these data suggest that while the dsRB2 is essential for the RNA dependent ATPase activity of MLE, the dsRB1 is dispensable and the RGG C-terminal region influences or modulates the affinity of MLE for RNA.



**Fig. 21.** Lineweaver-Burk linear representation for the reaction kinetics of MLE, MLE<sup>ΔRB1</sup> and MLE<sup>ΔRGG</sup> as a function of RNA concentration [S]. V (reaction velocity) corresponds to the ratio of hydrolyzed ATP during the time for each concentration of substrate. Error bars correspond to the average reaction velocity at 10 min and 20 min.

### 3.2. Effects of RNA binding domain deletions on MLE helicase activity

The RNA helicase activity is strictly dependent on the ATPase activity. However, certain mutations in the conserved motifs of some helicases enzymes are known to affect only the helicase activity (Rocak et al., 2005). In view of these results, MLE and all derivatives of MLE, which are able to hydrolyze ATP, were tested in the helicase assay for their capacity to unwind dsRNA. A 75 bp roX2 DNA sequence, predicted not to form any secondary structures, was transcribed *in vitro* in the presence of radioactive [ $\alpha^{32}\text{P}$ ]-ATP. DsRNA hybrids were obtained by annealing the radioactive ssRNA with an excess of the corresponding cold complementary sequence. Both ssRNA and dsRNA substrates can be distinguished on a native 10 % polyacrylamide gel, due to their different mobility. Increasing amounts of MLE, MLE<sup>ARB1</sup> and MLE<sup>ARGG</sup> (Fig. 22 A) were incubated with constant amounts of dsRNA yielding a molar ratio of approximately 1:4, 1:2, 1:1 and 2:1. MLE unwinds dsRNA only in the presence of ATP and the efficiency of unwinding is proportional to the amount of enzymes used (Fig. 22, B and lanes 5, 6, 7, 8). In the control reactions (absence of ATP and absence of the recombinant protein) no helicase activity was observed (Fig. 22, B and lanes 2, 3, 4, 13). MLE<sup>ARB1</sup> also efficiently unwinds dsRNA in an ATP-dependent manner, with no significant differences to MLE (Fig. 22, B and lanes 9, 10, 11, 12). Interestingly, deletion of the RGG domain strongly decreases the helicase ability of MLE (Fig. 22, B and lanes 14, 15, 16, 17). It should be mentioned that in these experiments the dsRNA probe had not been purified. DsRNA helicase activity was therefore measured in the presence of an excess of ssRNA. However, when a purified dsRNA substrate was tested, the MLE<sup>ARGG</sup> mutant shows an ATP-dependent helicase activity comparable to MLE (data not shown). One possible explanation for this result is that the MLE<sup>ARGG</sup> enzyme binds better ssRNA than MLE and the MLE<sup>ARB1</sup> mutant. In the presence of excess ssRNA, less MLE<sup>ARGG</sup> protein may be available for dsRNA binding and unwinding. The dsRB1 seems instead dispensable for the helicase activity of MLE.



**Fig. 22. Helicase activity of MLE, MLE<sup>ARB1</sup> and MLE<sup>ARGG</sup> recombinant proteins.** **A)** MLE, MLE<sup>ARB1</sup> and MLE<sup>ARGG</sup> proteins were purified from Sf9 cells as described in Material and Methods. 25 fmoles and 50 fmoles of each protein were loaded on an 8% polyacrylamide gel and quantified by Coomassie blue staining. **B)** Increasing amounts of each protein (35.25, 75, 150 and 300 fmoles) were mixed with radioactive, unpurified dsRNA substrates (120 fmoles) in the combinations indicated by (+). In the absence of ATP the highest amount of each protein was used. Unwinding reactions were carried out at 37°C in the presence (+) or in the absence (-) of 1 mM ATP. After 30 min, an aliquot (10 µl) from each reaction was analyzed by native 10% PAGE.

### 3.3. RNA binding properties of recombinant MLE and MLE deletion mutants

#### 3.3.1 Binding to ssRNA

MLE has been shown to bind ssRNA with higher affinity than ssDNA. In the same studies MLE ATPase activity was strongly stimulated by dsRNA and MLE unwound dsRNA substrates more efficiently than dsDNA. These data suggest that both, ssRNA and dsRNA, can be bound by MLE (Lee et al., 1997). However, it is not known whether MLE is able to distinguish both RNA substrates. To address the RNA binding properties of MLE, electrophoresis mobility shift assays (EMSA) were employed. Comparable amounts of MLE and MLE derivatives (Fig. 23, A) were mixed with ssRNA at an approximate molar ratio of 1:1.

Only a weak mobility-shifted band could be observed for recombinant MLE and ssRNA, and the shift was ATP independent. However, in the presence of non-hydrolysable ATP (ATP $\gamma$ S), MLE was unable to bind the substrate. Inclusion of ADP permitted binding (Fig. 23, B and lanes 2, 3, 4, 5).

The MLE<sup>GET</sup> protein contains a point mutation (K413E) in the active site of the ATPase domain, shown for other helicases to abolish both ATP binding and hydrolysis. However, in the case of MLE, this mutation impairs only its ATPase activity, but affects less severely its capacity to bind NTPs. A 5 to 10 fold lower affinity for GTP compared to wild type MLE has been estimated for recombinant his-MLE<sup>GET</sup> (Lee et al., 1997). The experiment shows that MLE<sup>GET</sup> bound ssRNA in an ATP-independent manner. In addition, unlike MLE, ATP $\gamma$ S did not affect the capacity of MLE<sup>GET</sup> to bind ssRNA (Fig. 23, B and lanes 6, 7, 8, 9).

Deletion of the dsRB1 had no impact on MLE ssRNA binding ability (Fig. 23, B and lanes 10, 11, 12, 13). In contrast, MLE<sup>ARB2</sup> did not bind ssRNA in all tested conditions (Fig. 23, B and lanes 14, 15, 16, 17).

Interestingly, MLE<sup>ARGG</sup> showed a better capacity to bind ssRNA than MLE, and defined bands can be reproducibly visualized for this mutant. In addition, while ATP did not affect the ssRNA binding ability of the other proteins, binding of MLE<sup>ARGG</sup> to ssRNA was less efficient in the presence of ATP (Fig. 23, B and lanes 18, 19, 20, 21).

### 3.3.2 dsRNA binding of MLE and MLE deletion mutants and helicase activity

The influence of the different RNA binding domains on MLE dsRNA binding capacity was also tested. The same dsRNA substrate used in the helicase assay was further purified and used as probe for the band shift experiments. Due to the strong activities of the helicases, their unwinding capacity could be visualized in the same assay. A small amount of ssRNA was not eliminated upon purification of the dsRNA. However, both RNAs can be discriminated on the agarose gel, allowing comparing the ssRNA and dsRNA binding properties of the MLE derivatives.

In the absence of ATP, two types of MLE-dsRNA complexes were observed (Fig. 23, C and lane 3). With increasing amounts of MLE, the fast-migrating complex disappeared, with the concomitant increase of the slow-migrating complex (data not shown). These results suggest that more than one MLE protein binds one molecule of RNA successively. It is also possible that MLE is able to dimerize and that the shifted bands correspond to different multimeric protein-RNA complexes. In the presence of ATP, MLE helicase activity was strongly stimulated and most of the dsRNA-hybrids were converted into ssRNA (Fig. 23, C and lane 4). Only a weak shift of the corresponding ssRNA was observed. In the presence of ATP $\gamma$ S, MLE helicase activity was inhibited and binding to dsRNA was also impaired (Fig. 23, C and lane 5). When ADP was included, binding of MLE to dsRNA occurred and MLE-RNA complexes were formed as in the absence of ATP (Fig. 23, C and lane 6). Thus, MLE dsRNA binding capacity is different with ATP and ATP $\gamma$ S, suggesting that the affinity to dsRNA is modulated at different stages in the cycle of ATP hydrolysis.

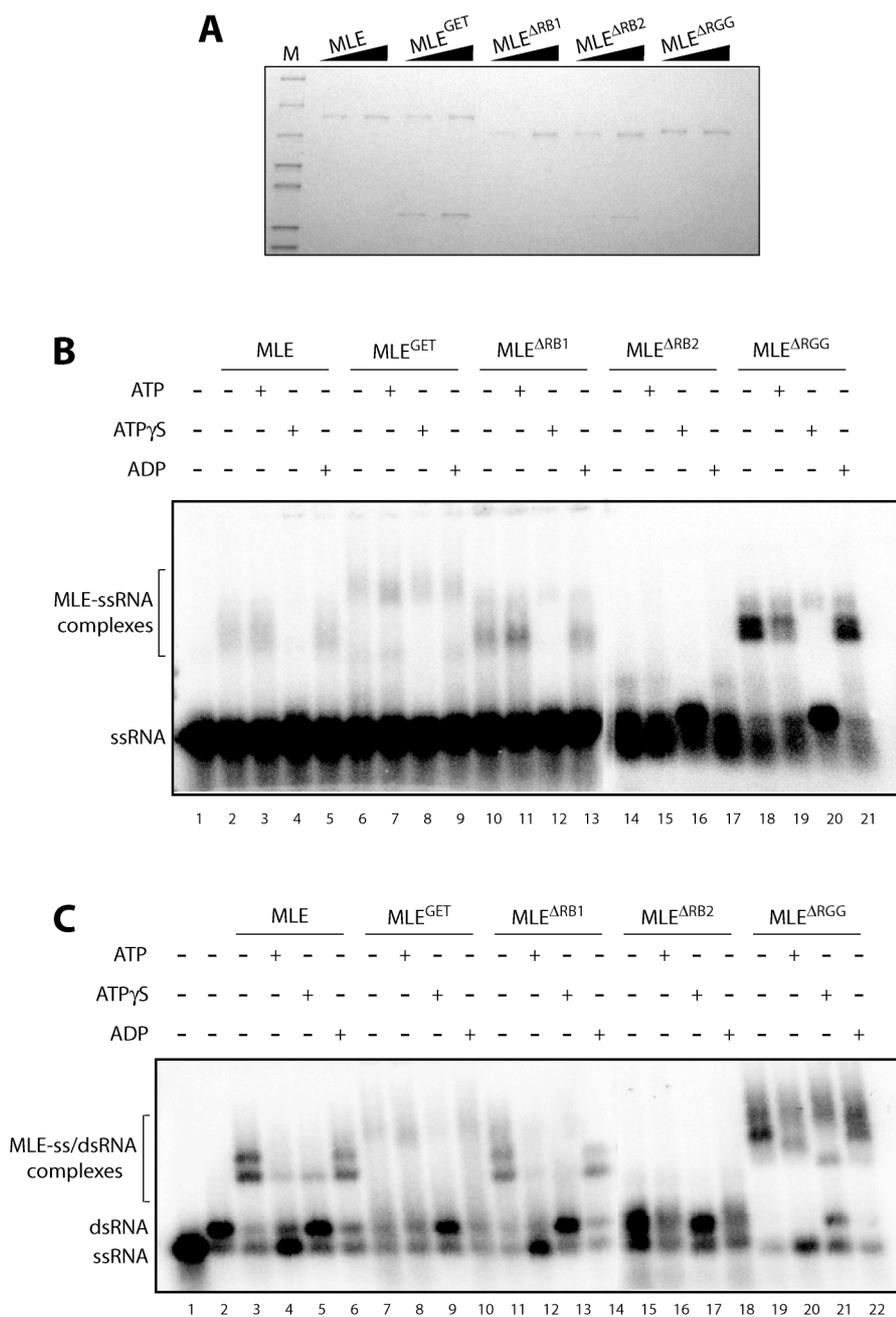
MLE<sup>GET</sup> bound dsRNA less efficiently than MLE (Fig. 23, C and lanes 7, 8, 9, 10) and dsRNA-MLE<sup>GET</sup> complexes were poorly resolved. However MLE<sup>GET</sup>, as expected, has no helicase activity and in the presence of ATP no release of ssRNA was observed. In addition, the binding of MLE<sup>GET</sup> to dsRNA was also inhibited in the presence of ATP $\gamma$ S.

The RNA binding capacity of MLE was not affected by deletion of the dsRB1, in agreement with its similar ATPase and helicase activities to the wild type protein (Fig. 23, C, lane 11, 12, 13, 14).

The MLE<sup>ARB2</sup> mutant did not bind dsRNA (Fig. 23, C and lanes 15, 16, 17, 18). However, in the presence of ATP or ADP a decrease in the dsRNA probe was observed, although no defined RNA-proteins complexes could be resolved.

Deletion of the RGG domain differently affected MLE RNA binding properties (Fig. 23, C and lanes 19, 20, 21, 22). In the presence of ATP, helicase activities could be observed. However, in contrast to the wild type MLE, MLE<sup>ΔRGG</sup> has increased affinity for ssRNA, yielding robust band shift. This result can also explain the higher  $K_M$  for RNA calculated for the MLE<sup>ΔRGG</sup> mutant in the ATPase assay. Due to its better binding to ssRNA the MLE<sup>ΔRGG</sup> mutant releases ssRNA slower than the wild type MLE and the MLE<sup>ΔRB1</sup> mutant and thus the binding of new dsRNA substrate is affected. In addition, in the presence of ATP $\gamma$ S, both dsRNA and ssRNA bands were shifted to slow-migrating bands, whereas under similar conditions almost no binding to RNA could be observed for the other proteins. In the ADP-bound conformation, binding of MLE<sup>ΔRGG</sup> to ssRNA and dsRNA substrates occurred as in the absence of any NTP.

These results show that MLE is able to discriminate between ssRNA and dsRNA, and that it associates preferentially with dsRNA. Deletion of the C-terminal RGG domain increases the affinity of the corresponding mutant for both dsRNA and ssRNA, although like the wild type protein, the mutant binds dsRNA better than ssRNA. The dsRB1, although predicted to bind double stranded RNA, appears not to contribute significantly to MLE RNA binding in these conditions. Instead, the dsRB2 domain seems essential for both ssRNA and dsRNA binding. However, the possibility that the MLE<sup>ΔRB2</sup> mutant is not properly folded cannot be excluded. In addition, MLE RNA-binding capacity is strongly modulated by ATP, with ATP binding and ATP hydrolysis having different effects.



**Fig. 23. RNA binding properties of MLE and MLE derivatives.** **A)** Quantification of MLE and MLE deletion mutants on Coomassie stained polyacrylamide gel. For each protein two amounts were loaded (30 ng and 60 ng). **B)** Electrophoresis mobility shift assay, using 25-50 fmoles of MLE or MLE deletion mutants and 25-50 fmoles of radiolabeled ssRNA. Different NTP substrates (ATP, ATP<sub>γ</sub>S and ADP) were also added (+) or not (-) in each reaction mixture, at a concentration of 1 mM. **C)** Binding of recombinant MLE and MLE mutants to dsRNA. Reaction conditions are the same as for the ssRNA band shift. The position of ssRNA and dsRNA are also indicated.

#### **4. In vivo localization of MLE deletion mutants**

MLE enzymatic activities are required *in vivo* for its stable localization to the X territory and for dosage compensation (Gu et al., 2000; Lee et al., 1997). Since MLE mutant proteins have different enzymatic properties *in vitro*, their ability to selectively associate to the X chromosome was tested. The same constructs used to produce the baculoviruses were cloned in a vector for expression in the male *Drosophila* Schneider cells line (SF4). A sequence encoding the green fluorescent protein (GFP) was fused at their C-terminus in order to distinguish them *in vivo* from endogenous MLE. 2 days after transfection, SF4 cells were subjected to immunofluorescence experiments, where GFP antibodies were used to visualize the different fusion proteins and MSL1 antibodies were chosen to mark the X territory. In order to have a better idea about the recruitment capacity of the different mutants, only those cells where the GFP fusion proteins were expressed at low levels were considered.

Exogenous MLE-GFP reveals a distribution pattern similar to the endogenous MLE protein with a clear enrichment to the X territory as shown by its colocalization with MSL1 (Fig. 24, A)

Surprisingly, the MLE<sup>ARB1</sup>-GFP mutant, even though still active *in vitro* in both ATPase and helicase assays did not preferentially localize to the X territory. Enrichment on the X could be observed only in a very limited number of cells. However, the definition of the territory was not as clear as for the wild type protein and higher staining in the nucleoplasm was seen.

In flies MLE<sup>GET</sup> had been shown to be recruited to the X chromosome at only few sites as well as at some autosomal loci (Lee et al., 1997). In *Drosophila* SF4 cells MLE<sup>GET</sup>-GFP did not preferentially colocalize with MSL1 on the X. However, in some cases, like for the MLE<sup>ARB1</sup> mutant, enrichment on the X chromosome was observed.

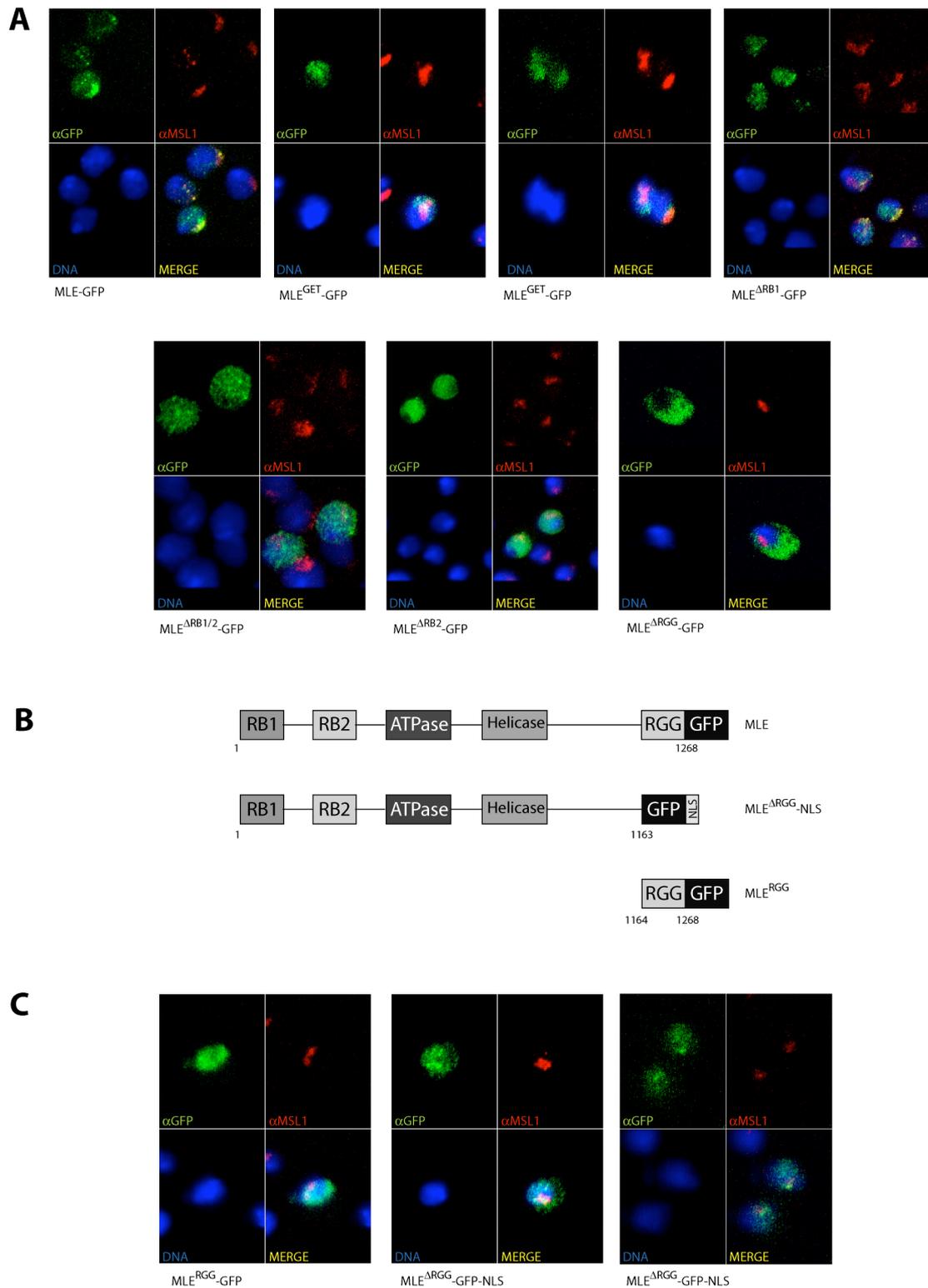
Both MLE<sup>ARB2</sup>-GFP and MLE<sup>ARB1-2</sup>-GFP fusion proteins were distributed over the entire nucleus and no targeting to the X could be visualized. Differently from the other fusion proteins, MLE<sup>ARB1-2</sup> also localizes in the cytoplasm.

Deletion of the RGG-box completely dislocated the MLE<sup>ARGG</sup>-GFP mutant to the cytoplasm, suggesting the presence of a nuclear localization signal (NLS) within this region. To confirm this hypothesis and to verify the effect of the RGG domain on MLE recruitment to the X, two additional mutants were produced. First, the region deleted in

the MLE<sup>ARGG</sup>-GFP protein was fused directly N-terminally to GFP. Second, an artificial nuclear localization signal was placed behind the GFP in the MLE<sup>ARGG</sup>-GFP fusion protein (Fig. 24, B).

The RGG domain alone was sufficient to localize GFP in the nucleus, whereas the MLE<sup>ARGG</sup>-GFP-NLS protein, once artificially translocated to the nucleus, showed a distribution pattern reminiscent of the one seen for MLE<sup>GET</sup> and MLE<sup>ARB1</sup> (Fig. 24, C).

These results confirm that the C-terminal region of MLE, containing the RGG-box, also harbors a nuclear localization signal (NLS), required for its stable maintenance in the nucleus. In addition, the failed localization of the MLE<sup>ARB1-2</sup> and the MLE<sup>ARB2</sup> inactive mutants suggest that MLE enzymatic activities are necessary for the correct targeting of MLE to the X chromosome. However, correct targeting does not follow the activity profile as shown by the different distribution of the active MLE<sup>ARB1</sup> mutant when compared to the wild type protein. Taken together these results also suggest that MLE enzymatic activities, although necessary, are not sufficient for its correct targeting to the X and reveal additional requirements (summary in Table 1).



**Fig. 24. *In vivo* localization of MLE and MLE deletion mutants.** **A, C)** *Drosophila* SF4 cells were transiently transfected with either MLE or the indicated MLE mutants, all fused to GFP at their C-terminus. Localization of the fusion proteins was visualized using GFP antibodies ( $\alpha$ GFP) and the X territory was marked with MSL1 ( $\alpha$ MSL1). DNA was counterstained with TOPRO-3. Images were acquired with a Zeiss LSM 510 META confocal microscope. **B)** Schematic representation of RGG-GFP and MLE<sup>ΔRGG</sup>-GFP-NLS fusion proteins.

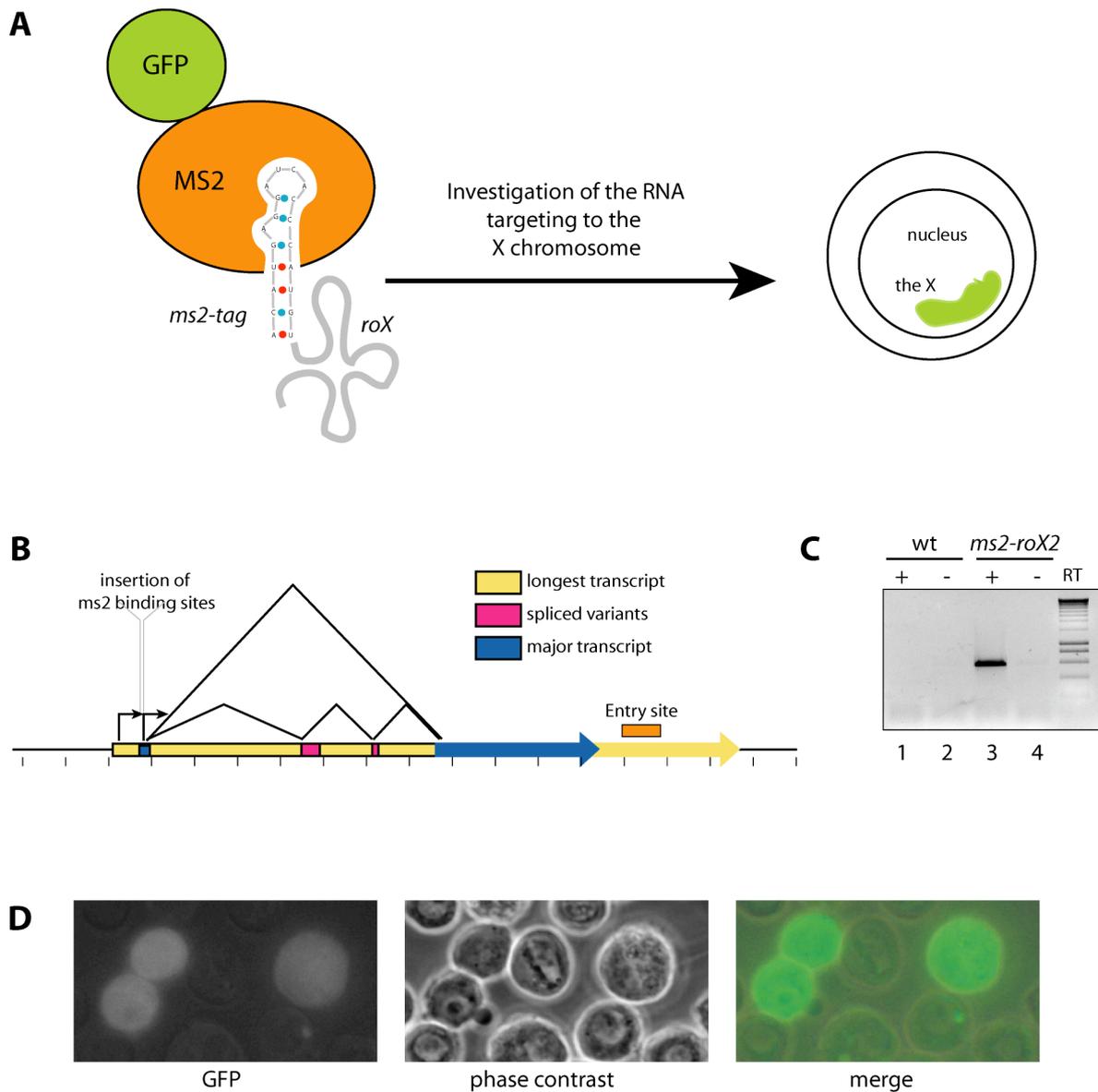
Protein	X chromosome targeting	Nuclear localization
MLE	++	++
MLE <sup>GET</sup>	+-	++
MLE <sup>ARB1</sup>	+-	++
MLE <sup>ARB2</sup>	-	++
MLE <sup>ARB1-2</sup>	-	+-
MLE <sup>ARGG</sup> -NLS	+-	++
MLE <sup>ARGG</sup>		-

**Table 1. Summary of the results from the *in vivo* localization studies of MLE and MLE deletion mutants in SF4 cells.** Degrees of X chromosome targeting are indicated as follows: ++ (defined staining on the X). +- (enrichment on the X with higher nuclear staining). - (no enrichment on the X). Degrees of nuclear localization are indicated as follows: ++ (localization in the nucleus). +- (localization in both nucleus and cytoplasm). - (localization in the cytoplasm).

## **5. Strategies to study the function of roX RNAs**

### **5.1. *In vivo* localization of the tagged ms2-roX2 RNA**

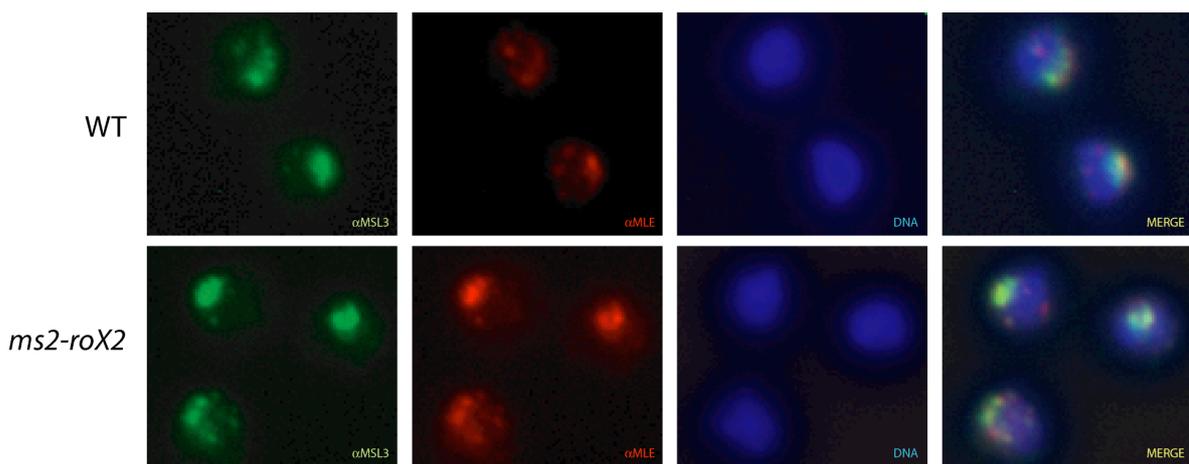
An ms2 tagging approach has been used in different systems to visualize the *in vivo* localization of reporter RNAs (Bertrand et al., 1998). The ms2 sequence codes a 19-nucleotide RNA stem loop, which is specifically recognized by the viral coat protein MS2. Fusion of MS2 to GFP can then be used to follow RNA movements in living cells. The same system was utilized in *Drosophila* cells to characterize important regions of roX2 RNA required for its correct targeting to the X chromosome (Fig. 25, A). Six ms2 sites were inserted at the 5' of the spliced isoform rox2 78.2.2, in order to increase the signal from multiple bound GFPs (Fig. 25, B). The endogenous roX2 promoter was chosen to avoid undesirable effects due to overexpression of the reporter roX2 RNA. The expression levels of the MS2-GFP fusion protein were, however, driven by the inducible metallothionine promoter. The ms2-roX2 construct was stably transfected into SF4 cells and its correct expression analyzed by RT-PCR using specific ms2 and roX2 primers (Fig. 25, C). Quantification showed an approximately 2 fold higher expression of exogenous roX2 compared to endogenous roX2 RNA (data not shown). Recombinant MS2-GFP was then expressed in the stable ms2-roX2 line and 2 days after transfection, cells were observed under the microscope. An overall staining of MS2-GFP protein in both nucleus and cytoplasm was observed in cells expressing MS2-GFP and specific recruitment of ms2-roX2 RNA to the X chromosome was not apparent (Fig. 25, D). A fine titration of MS2-GFP expression levels, avoiding overexpression of the chimera fusion protein, also led to the same negative results (data not shown). These data suggest that the presence of ms2 tags at the 5' end of roX2 RNA may affect its correct folding and as a consequence its proper localization to the X chromosome. This hypothesis is also supported by the fact that the same ms2-roX2 construct did not rescue the lethality of the double knockout roX RNAs male flies (M. Prestel, unpublished results).



**Fig 25. *In vivo* strategy to study the function of roX RNAs.** **A)** Schematic representation of the *in vivo* strategy to visualize roX2 RNA recruitment to the X chromosome in *Drosophila* cells. **B)** Schematic representation of the ms2-roX2 construct. **C)** RT-PCR on total extract from wild type and the ms2-rox2 stable cell lines with specific ms2 and roX2 primers. **D)** A construct expressing recombinant GFP-MS2 fusion protein was transiently transfected into ms2-roX2 expressing cells.

When ectopically expressed from autosomal positions, roX2 RNA was shown to compete for the binding of the MSL proteins to the X chromosome (Park et al., 2002). In this situation, recruitment of the MSLs to the insertion site and limited spreading of the DCC on the X chromosome can be observed.

In view of these data it is possible that the ms2-roX2 RNA does not localize to the X chromosome because the endogenous roX2 competes for proper targeting. However, the tagged RNA might still be integrated into the DCC complex and lead to the delocalization of the MSL proteins from the X territory. To test this hypothesis, the ms2-roX2 stable cell line was subjected to immunofluorescence analysis with specific antibodies against MSL3 and MLE. These proteins did not show any different distribution pattern when compared to wild type SF4 cells (Fig. 26), suggesting that the ms2 tagging strategy compromised both the targeting and the MSL binding properties of roX2 RNA.

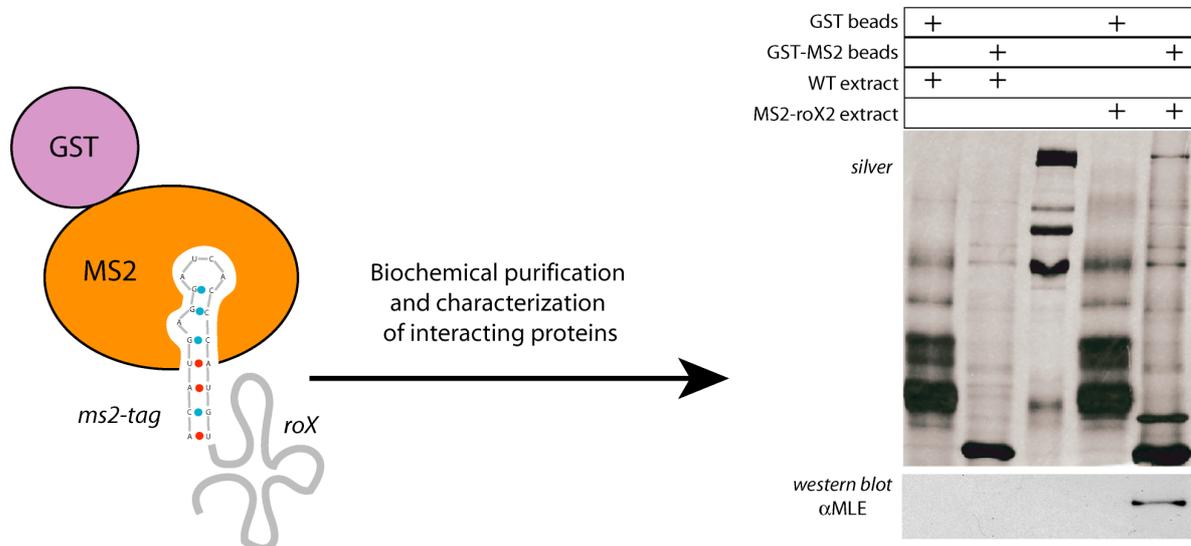


**Fig. 26. Immunofluorescence on wild type SF4 cells and ms2-roX2 stably transfected SF4 cells.**  $\alpha$ MLE and  $\alpha$ MSL3 antibodies were used to detect the corresponding endogenous proteins. DNA was counterstained with Hoechst 33258.

## 5.2. Identification of proteins specifically binding ms2-roX2 RNA

*In vitro* pull down experiments were also attempted to analyze binding of MSLs to ms2-roX2 (Fig. 27, left). The same approach may allow the identification of new candidate proteins involved in dosage compensation.

Nuclear extracts from the *ms2-roX2* stable cell lines were incubated with GST-MS2 or GST proteins bound to affinity glutathione beads. Nuclear extracts from untransfected cells were used as a control. Specific binding of *ms2-roX2* RNA to GST-MS2, but not to GST beads, was confirmed by RT-PCR (data not shown). The presence of the MSL proteins was checked by Western blot analysis with specific MSL antibodies. MLE could be specifically copurified with *ms2-roX2* as shown by its absence in the pull-down fractions from wild type cells incubated with GST or GST-MS2 beads, as well as from *ms2-rox2* cell extracts incubated with GST beads (Fig. 27, right). In contrast, MSL1, MSL2, MSL3 and MOF did not bind *ms2-roX2* (data not shown). Since expression of *ms2-roX2* RNA does not affect the proper localization of MLE on the X chromosome, one possible explanation is that only a small fraction of MLE interacts with this RNA. On the other hand, the capacity of MLE to bind RNA not specifically should also be considered. As shown in the silver stained gel, specific proteins were copurified with *ms2-roX2* (Fig. 27, right). However, since MSL proteins cannot be purified using this approach, it remains doubtful that these factors have a role in dosage compensation.



**Fig. 27. *In vitro* strategy to study the function of *roX* RNAs.** **Left)** Schematic representation of the *in vitro* approach to purify *ms2-roX2* RNA binding proteins. **Right)** Extracts from wild type and *ms2-roX2* expressing cells were incubated with GST or GST-MS2 coated beads. After pull down, the presence of the MSL proteins was checked by Western blot, whereas unknown proteins, binding *ms2-roX2* RNA, were visualized on silver stained gel.

### III. Discussion

#### 1. Association of MLE with the DCC complex

##### 1.1. Reconstitution of a complete DCC in Sf9 cells: advantages of the baculosystem

The baculosystem has been previously used for the purification of a four subunit complex containing an active MOF enzyme (Morales et al., 2004). Recombinant MLE-flag, expressed and purified from Sf9 cells, is also strongly active and helicase activity could be measured even at 4°C (see results). For these reasons the baculosystem was employed for the reconstitution of the complete DCC. In addition to the known protein components, a virus expressing roX2 RNA was generated. roX2 is correctly expressed, showing that the Sf9 cells can also be used to produce and study non-coding RNAs. For the purification of the MSL complex, Sf9 cells were coinfecting with up to six viruses. Although the recovery of the proteins was generally higher using fewer viruses, the MSL proteins were stable and no degradation products were detected. Since MLE has a similar size to MOF its interaction with the other MSL proteins was confirmed by Western blot. However, when fractions of the eluted DCC complex were tested in the ATPase assay, only a weak enzymatic activity could be measured (data not shown). These results show that only a small fraction of MLE is copurified with the MSLs. It is possible that MLE associates only transiently with the DCC because it enzymatically acts on the complex, which may render MLE difficult to copurify with the MSL proteins. Although the amount of baculoviruses expressing the different MSLs had been fine-titrated, substoichiometric levels of MSL2 and MOF were present in the complex relative to MSL1 and MSL3. If association of MLE with the DCC is strictly dependent on MSL2, improving the stoichiometry of the reconstituted complex would also increase the amount of MLE protein copurified.

When compared to the input material, only small amounts of roX2 RNA were integrated into the DCC after purification, and several cycles of amplification were needed for its detection. It is possible that correct incorporation of roX RNA requires a co-transcriptional assembly of the MSL proteins, a condition that may not be met in Sf9 cells. However, the stoichiometry of roX RNAs and MSL proteins in the complex is

currently not known. Moreover, whether each MSL complex, bound to the X, contains roX RNAs cannot be verified with the state of art techniques employed. The possibility that our results may reflect the real *in vivo* situation should not be excluded.

## 1.2. MSL proteins specifically bind roX2

Expression of the DCC components in the Sf9 cells has allowed for the first time to demonstrate a specific interaction of the MSL complex with roX2 RNA. Among the DCC members, MSL3, MOF and MLE bind RNA (Akhtar et al., 2000; Lee et al., 1997) and so far, only MSL3 has been suggested to interact specifically with roX2 RNA *in vitro* (Buscaino et al., 2003). However, this result has not been reproduced in our lab. A possible explanation is that a specific and stable association of the MSLs with roX1 and roX2 non-coding RNAs depends on their faithful integration into the DCC complex. Analysis of the roX RNA binding properties of each single MSL protein should be tested in the same system to confirm this hypothesis.

To address the specificity of the MSL proteins for roX2 RNA, the antisense sequence of roX2 (ant-roX2) was chosen as control. Ant-roX2 RNA also codes for a non-coding RNA with similar GC content and similar propensity to form secondary structures. It should be mentioned that the programs available for RNA structure predictions are not very reliable since for the same sequence they generally give very different results. Therefore these observations should be interpreted with caution. However, among the most stable structures, no identical hairpins were predicted for roX2 and ant-roX2 RNAs (data not shown). The MSL proteins specifically bind roX2 RNA, and no interaction with ant-roX2 RNA could be detected. It is likely that only properly folded roX2 RNA molecules are recognized by the DCC. It is conceivable that MLE, through continuous cycles of winding and unwinding, might be required to maintain correct structures under physiological conditions. However, I was unable to detect the beneficial effects of MLE in the reconstitution system employed. In addition, stable association of roX RNAs into the complex might also depend on MLE.

### 1.3. Integration of MLE into the DCC does not require roX2 RNA

In flies mutated for MOF or MSL3, partial complexes consisting of MSL1, MSL2, MLE and roX RNAs can be visualized at few high affinity sites on the polytene male X chromosome (Kelley et al., 1999; Lyman et al., 1997). Due to the presence of roX RNAs, it was suggested that integration of MLE into the DCC complex is roX RNA-mediated (Meller et al., 2000). This conclusion was also suggested by previous attempts to purify the DCC, which led to dissociation of MLE from remaining MSLs at high stringency (Scott et al., 2000).

In my DCC reconstitution, MLE copurified with the other MSLs in the absence of expressed RNA and in the presence of either roX2 or ant-roX2 RNAs. This result suggests that association of MLE with the other MSL proteins is direct and does not specifically require roX2. This hypothesis was confirmed *in vitro* by several assays, where a preferential association of MLE with MSL1 and MSL2 was observed. In addition, RNase and/or DNase treatment of coinfecting Sf9 cells extracts did not affect the capacity of MLE to pull down MSL2, showing that this interaction does not depend on RNA.

Although roX RNA improves the assembly and the distribution of the DCC over the X chromosome in wild type flies (Meller et al., 2000; Meller and Rattner, 2002; Park et al., 2002), the lethality of *roX1* and *roX2* null males can be partially overcome by overexpressing MSL1 and MSL2 (Oh et al., 2003). Consistent with these observations, interaction of MSL3 and MOF with MSL1, rather than their ability to bind RNA, has been shown to play the major role in the recruitment of both proteins to the X territory (Morales et al., 2005; Morales et al., 2004). MSL3 and MOF bind RNA *in vitro* via their related chromo-barrel domains and point mutations in MOF chromo-barrel domain also affected RNA binding both *in vitro* and *in vivo* (Akhtar et al., 2000). However, structure analyses suggest that, although the MOF-like chromo-barrel domains may have a role in RNA binding, they also share a common fold with chromatin-associated modules (Maurer-Stroh et al., 2003). The MSL3-like chromo-barrel domains are instead predicted to bind methylated residues, a proposition that needs to be tested *in vivo* (Nielsen et al., 2005). Moreover, MLE might have only a general affinity for RNA, as comparable amount of roX2 and ant-roX2 are copurified in its presence. If MLE binds RNA non specifically, considering that roX RNAs represent only a small fraction of the total RNA present in a cell, it is plausible to think that they are also not the main

determinant for MLE targeting to the X. Direct interaction of MLE with the other MSLs might instead play the major role.

First attempts to identify regions within MLE required for MSL1 and MSL2 binding failed. All the MLE derivatives, used for the characterization of MLE enzymatic properties still interact with both MSLs, although with different affinities. A possible explanation is that MLE uses more than one domain to interact with MSL1 or MSL2. In support of this hypothesis, several factors have been shown to utilize dsRNA-binding domains and RGG boxes to be integrated into multiprotein complexes. Recently the RB1 and the RGG domains of RHA have been shown to form together a specific interacting surface for the Werner syndrome protein, WRN (Friedemann et al., 2005).

## **2. Targeting of MLE to the X chromosome**

### **2.1. Different properties of MLE RNA binding domains**

Since for all helicases it is believed that nucleic acid binding plays a critical role in coordinating NTP hydrolysis and the unwinding process (Delagoutte and von Hippel, 2002), the contribution of the different RNA binding domains to MLE enzymatic activity was analyzed in detail.

In general, MLE and the different deletion mutants show a better binding to dsRNA as compared to ssRNA. Interestingly, deleting the RGG domain increased binding of MLE to ssRNA and presumably to dsRNA as well. As a result, a slight decrease in the efficiency of MLE ATPase and helicase activities could be observed. The opposite phenomenon has been described for RHA. The RGG box of RHA hardly had any influence on RNA binding, whereas the ATPase activity of the corresponding RHA mutant was severely affected (Zhang and Grosse, 1997). These results suggest that, while the RGG domain of RHA is required for proper enzymatic activities, in the case of MLE this domain might instead be needed to modulate the affinity of RB1 and/or RB2 domains for RNA. In addition, the MLE<sup>ARGG</sup> mutant is the only one showing dependency on ATP for ssRNA binding.

In the absence of RB2, the RNA binding ability of MLE was strongly impaired and as a consequence its ATPase and helicase activities were affected. The RB2 domain seems,

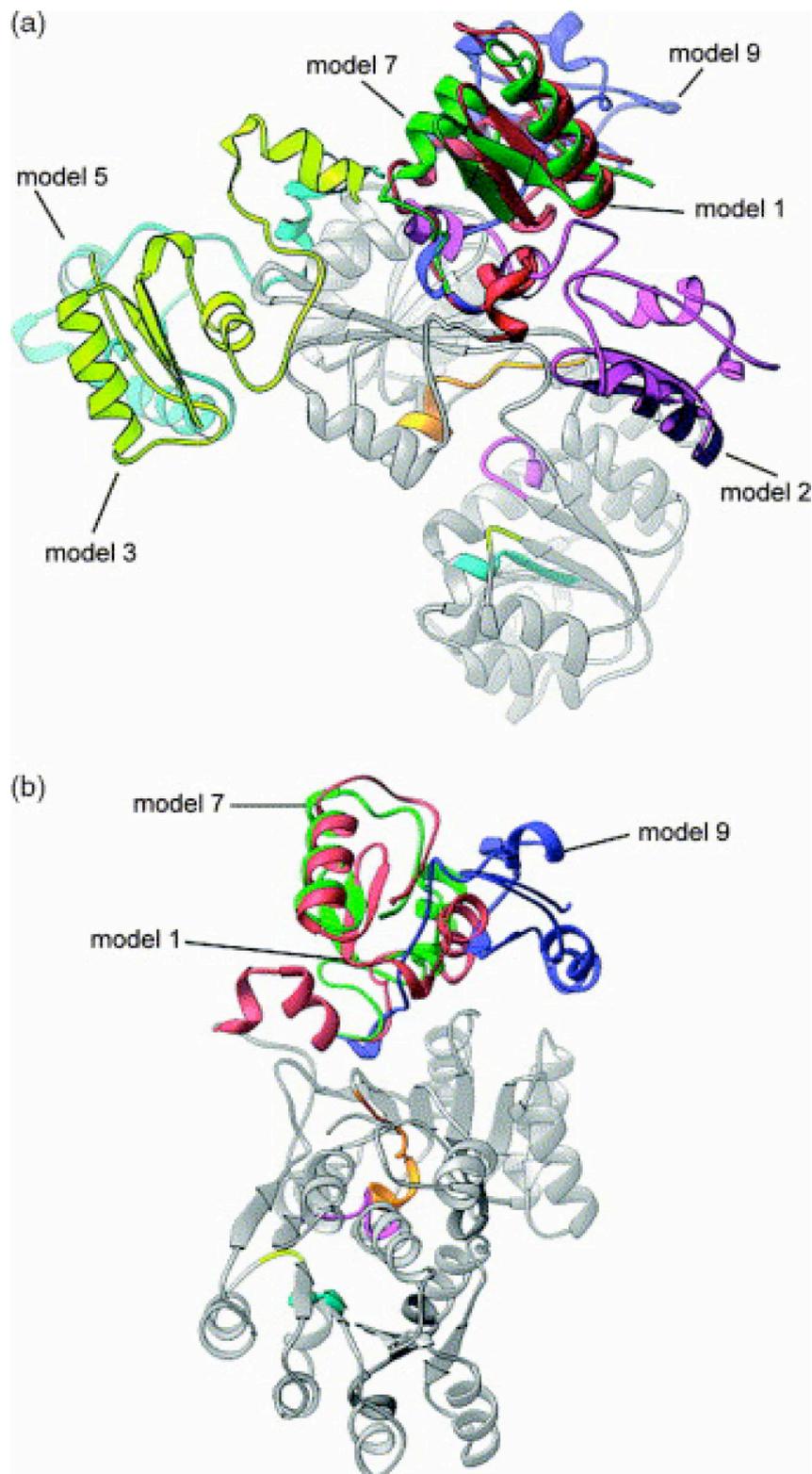
therefore, to play the major role in coordinating RNA binding and enzymatic properties. The same type of analysis for RHA showed that both RB1 and RB2 domains are dispensable for helicase activity, as their deletion did not abolish the capacity of RHA to hydrolyze ATP (Zhang and Grosse, 1997). However, I cannot exclude the possibility that removal of the RB2 domain in MLE affects the correct folding of the ATPase domain, resulting in an inactive protein. The RB1 domain of MLE is clearly dispensable for both RNA binding and enzymatic activities.

The K413E mutation in MLE ATPase domain abolishes its ATPase activity, although the MLE<sup>GET</sup> mutant still retains the capacity to bind NTP. A five to ten fold lower affinity for GTP has been measured for this mutant as compared to the wild type protein (Lee et al., 1997). As expected, MLE<sup>GET</sup> mutant does not have ATPase and helicase activities. As previously shown, it also binds ssRNA less efficiently than wild type MLE and in an ATP-independent manner. We also found a weaker association of MLE<sup>GET</sup> to dsRNA compared to wild type MLE. However, contradictory results have been obtained in the dsRNA binding assays. ATP and ATP $\gamma$ S, supposed to have the same effect, gave instead opposite results. In the presence of ATP a poorly resolved dsRNA/MLE<sup>GET</sup> band shift was observed, but ATP $\gamma$ S completely abolished the dsRNA binding. One possible explanation is that the affinity of the MLE<sup>GET</sup> mutant for ATP and ATP $\gamma$ S is different. However, upon sequencing of the MLE<sup>GET</sup> plasmid, in addition to the expected K413E mutation, a second mutation P590R in the loop connecting the ATPase domain and the helicase domain was identified. When the MLE<sup>GET-KRR</sup> double mutant and the MLE<sup>GET</sup> point mutant were fused to GFP and transfected in SF4 cells to compare their recruitment capacities to the X, clearly different distribution patterns were observed (data not shown). MLE<sup>GET</sup> showed a nuclear distribution and sometimes enrichment on the X, whereas the MLE<sup>GET-KRR</sup> was distributed all over the nucleus and the cytoplasm and no preferential association to the X could be observed. In addition, since the MLE<sup>GET-KRR</sup> mutant is expressed at a low level in Sf9 cells and copurified with additional factors, MLE<sup>GET-KPR</sup>/dsRNA complexes are poorly resolved. The further expression and purification of the MLE<sup>GET</sup> mutant is needed to elucidate its mechanisms of RNA binding.

Beside their high similarity (Lee and Hurwitz, 1993), RHA and MLE RNA binding domains appear to have different roles. The isolated RGG domain from RHA showed a strong affinity for single strand nucleic acids and in particular for ssDNA (Zhang and Grosse, 1997). Based on these findings the same features have been predicted for the

RGG domain of MLE (Gibson and Thompson, 1994). By contrast, removal of the RGG region from MLE increased its affinity for ssRNA. However, it cannot be excluded that the intrinsic properties of the isolated MLE RNA binding modules are different.

In addition, in the context of RHA cooperation between the RNA binding domains have been proposed to change their affinity for ATP as well as for RNA in order to achieve a proper stimulation of enzymatic activities. In view of our results such a model could also be true for MLE. Unfortunately, there are no data available about the orientation of RGG-boxes and dsRNA-binding domains in the tridimensional structure of large helicase enzymes. Recently, the tertiary structure of the *E. coli* DbpA SFII RNA helicase has been predicted by computational analysis (Talavera and De La Cruz, 2005). DbpA is implicated in ribosome biogenesis and is characterized by a central helicase domain and a C-terminal RGG domain, thought to determine the specificity for ribosomal RNA (Tsu and Uhlenbeck, 1998). Using comparative modeling and structure prediction methods, ten models for DbpA have been generated, all consisting of two domains (Fig. 28): the N-terminal domain, containing the conserved helicase core and the C-terminal domain, containing the RGG box. The structure of the helicase core is very similar to other helicases (review in (Delagoutte and von Hippel, 2002)), whereas the predicted C-terminal fold consisted of a  $\beta$ -hairpin and a  $\alpha$ -helix with an overall positive electrostatic potential due to a cluster of basic amino acids residues. These findings are consistent with this domain playing a role in RNA binding (Tsu et al., 2001). Although the predicted models differ in the relative orientation of the N-terminal and the C-terminal domains, in all cases hydrophobic contacts bring the two domains close to each other. This interaction has been proposed to influence the activity of the helicase domain as well as its affinity for the ATP and the RNA substrates (Henn et al., 2002; Polach and Uhlenbeck, 2002).



**Fig. 28. A)** Structural homology model of DbpA illustrating the possible orientations of the C-terminal domain. The ribbon diagrams are colored at the helicase core (residues 1–371 are silver; P-loop motif, residues 47–54 are magenta; DEAD box, residues 151–156 are cyan; SAT motif, residues 184–186 are yellow; motif VI, residues 325–335 are gold) and at the C-terminal domain. Residues 372–456 are colored red (model 1), magenta (model 2), yellow-green (model 3), cyan (model 5), green (model 7), and blue (model 9). **B)** Derived models consistent with hydrodynamic measurements (sedimentation coefficient and diffusion coefficients) (Talavera et al., 2006).

## 2.2. Model of MLE translocation on dsRNA

Different types of nucleotides have been employed in the band-shift assay to elucidate the molecular mechanisms that govern MLE translocation along the dsRNA substrate. Interestingly ATP binding and hydrolysis differently modulate MLE binding to both ss and dsRNA. In the presence of ATP $\gamma$ S binding of MLE to ssRNA and dsRNA is impaired, whereas ADP has the opposite effect. These results suggest that conformational changes occur in MLE during the cycle of ATP hydrolysis (Fig. 22). In the ADP-bound state the RB2 and the RGG domains might be close to each other, forming the “closed” conformation of MLE. This situation allows the RGG domain to constrain binding of the RB2 to the RNA substrate, perhaps decreasing its affinity for RNA. Binding of ATP might instead lead to the formation of the “open” conformation, in which the RB2 and the RGG domains are separated from each other. In this state MLE seems to have a very low affinity for RNA and can then move along the substrate and release ssRNA, in concomitance with ATP hydrolysis. In this model the RB1 has not been considered. Deletion of RB1 did not have any effects on MLE enzymatic activities and whether this domain is really able to bind RNA cannot be addressed in these experiments. Noteworthy, in the MLE<sup>ΔRB2</sup> mutant, RB1 cannot complement the loss of RB2. However, RB1 might be able to bind RNA, but it cannot reach the substrate, because of the presence of RGG interacting with RB2. In this scenario, deletion of RGG enables RB1 and RB2 to bind RNA, explaining the better RNA binding observed for the MLE<sup>ΔRGG</sup> mutant. The corresponding RB1 domain in RHA has been shown to interact preferentially with dsRNA, although more efficient binding could be observed when a mutant consisting of both RB1 and RB2 was tested. A more detailed analysis of the isolated RB1 and RB2 domains of MLE is needed to elucidate the precise function of all domain within MLE.

## 2.3. MLE enzymatic activity is not sufficient for proper targeting to the X

The enzymatic activities of MLE are essential for dosage compensation, as the MLE<sup>GET</sup> ATPase-inactive mutant failed to rescue the lethality of *mle* null flies (Gu et al., 2000; Lee et al., 1997). Since the deletion mutants used in this study affected the helicase

activity of MLE differently, the role of the RNA binding domains on the recruitment of MLE to the X territory was addressed.

The most surprising result was the distribution pattern of the MLE<sup>ARB1</sup> mutant. This protein was expected to localize on the X chromosome, as it shows similar enzymatic activities to the wild type protein *in vitro*. Instead, only a slight enrichment on the X territory could be observed. In addition, similarly to the MLE<sup>GET</sup> and the MLE<sup>ARGG</sup>-NLS mutants, enhanced nuclear staining relative to wild type MLE was observed. These results suggest that MLE enzymatic activities, although necessary for dosage compensation, are not sufficient to target or to maintain MLE on the X chromosome. However the main determinant for proper MLE targeting remains unknown. The RNA binding capacity, protein-protein interactions and proper enzymatic activities might contribute all together to integrate MLE into the DCC complex and to achieve a proper compensation of X linked genes.

One intriguing possibility is that the N-terminal RB1 domain of MLE, which is dispensable for RNA binding, contributes instead to the specific targeting of MLE to the X chromosome. In support of this hypothesis, the dsRB1 of RHA has been shown to be required for specific dsDNA recognition at the p16<sup>INK4a</sup> promoter (Myohanen and Baylin, 2001). Interestingly, a specific interaction of MLE with a short DNA sequence within the *roX2* promoter has been recently observed (Lee et al., 2004).

The nuclear localization signal of MLE was mapped to the RGG domain. The MLE<sup>ARGG</sup> fusion protein is cytoplasmic, whereas the isolated RGG domain alone positively localizes to the nucleus. The presence of an NLS within the RGG domain had already been described for other RGG-containing proteins, including RHA (Nichols et al., 2000; Siomi and Dreyfuss, 1995). Interestingly, methylation of specific arginine residues within this region by the PRMT class of arginine methyltransferases had been shown to control the nuclear import of several RNA binding factors (Liu and Dreyfuss, 1995; McBride et al., 2005; Xu and Henry, 2004). In the case of RHA, methylation is also required for its stable maintenance in the nucleus (Smith et al., 2004). Treatment of SF4 cells with inhibitors of arginine methylation did not affect MLE localization, raising the possibility that the nuclear import of MLE is regulated by a different mechanism (data not shown). However, due to the lack of a proper control for arginine methylation we cannot exclude the possibility that this modification also occurs within the RGG domain of MLE. In this context it should be noted that the MLE<sup>ARB1-2</sup> fusion protein is the only

MLE mutant that shows both nuclear and cytoplasmic staining. Stable maintenance of MLE in the nucleus may rather rely on its N-terminal RNA binding domain than on post-translational modification of the RGG domain.

#### **2.4. Correct targeting of the DCC to the X territory in SL2 cells requires MLE**

Genetic studies in flies led to a model of sequential recruitment of the MSL proteins to the X chromosome in males. Accordingly, it has been suggested that MSL1 and MSL2 are the first DCC members that bind to the high affinity sites on the X. In this scenario, recruitment of MSL1 and MSL2 is independent of the other MSLs, whereas MLE seems required for correct targeting of MOF and MSL3 as well as of roX RNAs. Since the MSLs have been shown to stabilize each other, their absence from the X chromosome in flies mutated for each single *msl* genes, could simply be due to a decreased amount of proteins not detectable by immunofluorescence.

I employed RNAi of MLE in SL2 cells to study the effects of MLE ablation on the targeting of the other DCC components on the X chromosome. Even though the resolution of signals in the small Schneider cell nuclei is not high enough to discriminate X chromosomal binding sites, the RNAi technique offers the advantage to follow the effects of changing the MSL proteins levels on their correct targeting on the X during a time course. After 3 days of RNAi, when most MLE is gone, only a slight reduction in MSL1 protein can be detected by Western blot and normal levels of the other MSLs are still present in the cells. However no recruitment of the DCC on the X territory could be observed and as a consequence H4-K16 acetylation was redistributed all over the nucleus. These results suggest that MLE is required for proper targeting and perhaps maintenance of the DCC on the X chromosome. However, whether MSL1 and MSL2, in the absence of MLE, are still present at the “high affinity” sites on the X could not be addressed. One possibility is that in the absence of MLE, MSL1 and MSL3 can still activate MOF. However, since recruitment of MOF to the X depends on MLE, the acetyltransferase activity cannot be targeted to the X, but it is instead relocated onto all chromosomes.

Interestingly, MLE 6E11 antibodies recognize a slow migrating band in the total SL2 extract, in addition to the band corresponding to endogenous MLE. The specificity of this band for MLE is demonstrated by its loss upon RNAi of MLE. Modifications of

MLE have not been described so far, but unstable post-translational modifications like sumoylation or ubiquitylation might be considered as the slow migrating band is highly unstable and lost upon nuclear extract preparation. E3 ligases, which are required in the last step of the ubiquitylation reaction, generally contain RING finger domains, which are essential for their enzymatic activity. Such a domain is present in MSL2 and I have observed a preferential association between MLE and MSL2. A speculative hypothesis is that MLE might be ubiquitylated by MSL2 in order to be targeted to the X chromosome. Unmodified MLE might instead account for other functions not related to dosage compensation. Recently, DNA-dependent protein kinase (DNA-PK) has been shown to phosphorylate RNA helicase A, in a RNA-dependent manner (Zhang and Grosse, 2004), opening even more complicated possibilities for MLE regulation.

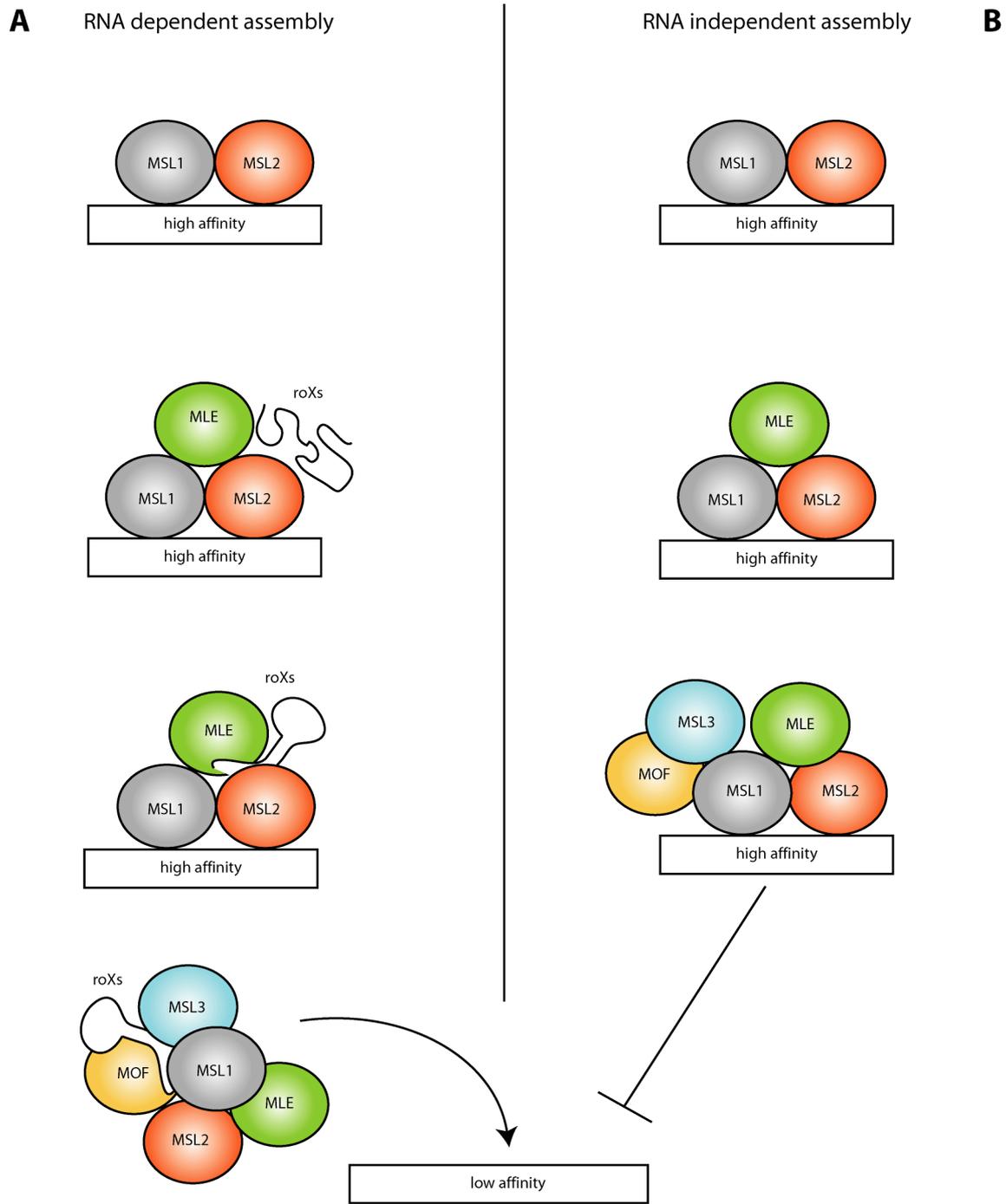
A strong reduction in MSL1 protein level and a mild reduction of MSL3 protein level were observed after 6 days of MLE RNAi. Considering that MOF has been proposed as the only *Drosophila* histone acetyltransferase responsible for H4-K16 acetylation (Akhtar and Becker, 2000) and that the HAT activity of MOF *in vitro* requires its interaction with MSL1 and MSL3 (Morales et al., 2004), a decrease in H4-K16 acetylation levels was expected. In contrast, an increased H4-K16 acetylation was reproducibly observed. These results suggest that while the MSL proteins might have a specific role in restricting H4-K16 acetylation specifically to the X for dosage compensation, unknown factors might activate MOF's activity on the autosomes. The increased level of this modification might be due to the fact that under those circumstances more autosomal genes are targets of MOF. It is also possible that MOF activity is counteracted on the X by repressive factors, whereas different mechanisms might modulate its activity on autosomes.

### **3. Models for DCC assembly and targeting to the X chromosome in males**

The results obtained and published data can be integrated into a model explaining the targeting of the DCC to the X chromosome in males. We propose two main mechanisms of recruitment, which mainly differ by the presence and the absence of the roX RNAs (Fig. 29).

In the former case MSL1 and MSL2 are the first DCC components, binding to few “high affinity sites” on the X. Subsequently, MLE is targeted to the X by association with MSL1 and MSL2. The MLE enzymatic activities might be required for the proper folding of roX RNAs and/or for stabilization, by feeding them into partial MSL complexes. Finally, interaction of MOF and MSL3 to MSL1 allows the formation of a complete DCC. Due to the presence of roX RNAs, this ribonucleoprotein complex might be more stable on chromatin and it can bind also moderate and low affinity sites, covering the entire X.

In the second model the same protein-protein interactions are utilized by the MSLs to assemble a complete complex on the X. However, due to the absence of roX RNA, the DCC is less stable on chromatin and cannot bind moderate and low affinity sites. This scenario is reminiscent of the one observed in flies lacking both roX1 and roX2 genes. In these transgenic males RNA-free complexes can be visualized only at few high affinity sites (Meller and Rattner, 2002). However, overexpressing MSL1 and MSL2 enhances the affinity of the DCC for X chromosome binding (Oh et al., 2003). It is possible that roX RNAs form a scaffold upon which the MSLs are arranged in the correct conformation. The enzymatic action of MLE on these RNAs might induce large conformational changes of the DCC complex, inducing a stronger binding to chromatin. Another possibility, that cannot yet be excluded, is that roX RNA contributes to sequence specificity by annealing to chromosomal DNA on the X to form short heteroduplex. This possibility could also explain the failure of past attempts to identify of common sequences between the different binding sites on the X.



**Fig. 29. Models for targeting of the DCC to the male X chromosome.** A) RNA dependent assembly. B) RNA independent assembly.

## **IV Materials and Methods**

### **1. Materials**

#### **1.1. Chemicals, reagents and enzymes**

Acrylamide (Roth); Agar, (Difco); Agarose high EEO (Biomol); Ampicillin (Roth); ADP (Sigma); ATP (Sigma); ATP $\gamma$ S (Sigma); Aprotinin (Genaxxon); BSA, purified (NEB); Bacto-Tryptone (Difco)  $\beta$ -mercaptoethanol (Sigma); Bromophenol blue (Serva); Coomassie G250 (Serva);  $\alpha$ -[<sup>32</sup>P]-ATP (Amersham Bioscience);  $\gamma$ -[<sup>32</sup>P]-ATP (Amersham Bioscience); NTPs, dNTP mix (Amersham Bioscience); DTT (Roth); EDTA (Sigma); EGTA (Sigma); Ethidium bromide (Sigma); Glycerol (Merck); Glycogen (Roche); Glutathione (Sigma); Goat serum (Gibco/BRL); Hepes (N-(2-hydroxyethyl)piperazine-H<sup>+</sup>-(2-ethanesulfonic acid), (Roth); IPTG (Isopropyl-D-thiogalactopyranoside), (Sigma); Leupeptin (Genaxxon); N-propyl-gallate (Sigma); NP40 (Igepal CA-630), (Sigma); Orange G (Sigma); Paraformaldehyde (Sigma); Pepstatin (Genaxxon); PMSF (Phenylmethanesulfonyl fluoride), (Sigma); SDS (Sodium dodecyl sulfate), (Serva); Temed (N,N,N',N'-Tetramethylethylenediamine), (Roth); Tris (USB Corporation); Xylene cyanol (Sigma).

DNA Polymerase I, Large (DNA Polymerase I (Klenow)) fragment (NEB); Pfu polymerase with 10x PCR buffer (Stratagene); Proteinase K (Boeringer); Restriction enzymes (NEB, Roche, MBI Fermentas); RNase A (Sigma); RNAsin (Promega); RQ1 DNase (Promega); SP6, T7, T3 RNA polymerase with 5x incubation buffer (Promega); SuperscriptII, supplemented with 5x incubation buffer and DTT 100mM (Invitrogen); Taq DNA Polymerase with 10X reaction buffer (Genaxxon); Protein-Marker II and IV (PeqGOLD);

#### **1.2. Oligonucleotides**

All oligonucleotides were purchased by MWG Biotech (Ebersberg, Germany) or Biomer Companies (Ulm, Germany).

### 1.3. Antibodies

Antibody	Host	Type	WB	IF	Source
anti-MSL1	rabbit	Affinity purified	1:1000	1:500	Kuroda M
anti-MSL2	rabbit	Affinity purified	1:1000	1:500	Kuroda M
anti-MSL3	goat	Affinity purified	1:1000	1:500	Kuroda M
anti-MOF	rabbit	Affinity purified	1:1000	1:500	Akhtar A
anti-MLE (6E11)	rat	Culture supernatant	1:1000	1:1000	
anti-H4-K16ac	rabbit	Affinity purified	1:1000	1:1000	Upstate
anti-GFP	mouse	Monoclonal 3E6		1:200	Molecular Probes

**Table 2.** List of antibodies used for the Western blot (WB) and for the immunofluorescence experiments (IF). Dilutions in the respective blocking buffers are indicated

## 2. Methods

Standard molecular biology procedures like polymerase chain reaction (PCR), restriction enzyme digestion, ligation of DNA fragments, preparation and transformation of competent cells, analysis of DNA on agarose gels were performed according to standard protocols (Ausubel, 1999; Sambrook and Rusell, 2001). Plasmid DNA was prepared with the plasmid purification kit (Qiagen). Isolation and purification of DNA fragments used for cloning were performed using the gel extraction kit (Qiagen) according to the manufacture instructions. Further standard methods for protein analysis, like SDS-PAGE and Western blot were done following the protocols described in Sambrook and Rusell 2001.

## 2.1. Plasmids

All plasmid have been sequenced by the MWG Biotech Company, (Ebersberg, Germany)

### 2.1.1. Cloning of *MSL1* and *MSL2* cDNAs into the pGEM7zf<sup>+</sup> vector

*MSL1* cDNA was cut out from the pFastBac-*MSL1* plasmid (Morales et al., 2004) with HindIII and subcloned into the pGEM7zf<sup>+</sup> vector (Promega) under the SP6 RNA Polymerase promoter. *MSL2* cDNA was cut out from the pHSP70-*MSL2*-GFP plasmid (Straub et al., 2005b) with AgeI followed by fill in reaction with the DNA Polymerase I (Klenow) to yield blunt ends and KpnI. The resulting DNA fragment was subcloned into the pGEM7zf<sup>+</sup> vector (Promega) digested SmaI/KpnI under the control of the T7 RNA Polymerase promoter.

### 2.1.2. GST-MLE (1-265 aa)

The NcoI/HindIII fragment, corresponding to the first 265 aa of *MLE* cDNA (M. Kuroda, Howard Hughes Medical Institute, Brigham and Women's Hospital, Boston, Massachusetts, USA) was cloned into the pGEX-2KG expression vector (Amersham), digested with the same enzymes.

### 2.1.3. GST and GST-MS2

GST was expressed using the pGEX-2T expression vector (Amersham). GST-MS2 recombinant protein was produced by cloning the XbaI/NotI blunt DNA fragment, containing the *MS2* sequence, from the pPACE-MS2 plasmid to the pGEX-2T vector, digested with EcoRI and incubated with DNA Polymerase I (Klenow) to yield blunt ends. Correct orientation was checked by analytical digestion.

### 2.1.4. Proteins and RNAs baculovirus expression vectors

*MSL1*, *MSL2*, *MSL3* and *MOF* cDNAs have been cloned into the pFastBac1 (for the expression of *MSL1*, *MSL2*, *MSL3* and *MLE* proteins) or pVL1392.2 (for *MOF*)

vectors as described (Morales et al., 2004). For the cloning of untagged MLE, *MLE* cDNA (M. Kuroda) was digested with NcoI and StuI enzymes, followed by fill in reaction with the DNA Polymerase I (Klenow) enzyme (Biolabs) to yield blunt ends and cloned into the pFastBac1 donor plasmid cut with StuI. Correct orientation was checked by analytic digestion. Sense roX2 RNA expressing baculovirus was produced by cloning of *roX2* 78.2.2 cDNA (M. Kuroda) into the pFastBac1 expression vector. Both plasmids were cut with XbaI and XhoI enzymes. For the preparation of the ant-roX2 RNA expressing baculovirus, the same *roX2*.78.2.2 cDNA was digested with XbaI, followed by incubation with the DNA Polymerase I (Klenow) enzyme to yield blunt ends and XhoI. PfastBac1 plasmid was digested instead with HindIII, followed by incubation with the DNA Polymerase I (Klenow) enzyme to yield blunt ends and XhoI enzymes. Recombinant baculoviruses expressing the MSL proteins, roX2 and ant-roX2 RNAs were produced using the 'Bac-to-Bac' expression system (Invitrogen).

MLE-flag and MLE-flag mutant constructs were produced as follows: the complete *MLE* sequence was amplified from the *MLE* cDNA using the MLE-BamHI.rv primer and the MLE-SacI.fw primer, containing a BamHI and a SacI sites respectively. The PCR product was digested BamHI/SacI and inserted into the pPACK-BACK-flag expression vector (Invitrogen), digested with the same enzymes. Correct sequence was confirmed by sequencing. The BamHI/XbaI DNA fragment containing *MLE-flag* sequence was removed from pPACK-BACK-MLE-flag and inserted into the pFastBac1 Sf9 cells expression vector. The MLE<sup>GET</sup>-flag construct was generated by replacement of the EcoRI/StuI fragment in MLE-flag with the correspondent fragment, containing the GKT-GET mutation, from the His-MLE<sup>GET</sup> plasmid (a gift from C.G. Lee, University of Medicine and Dentistry of New Jersey, Newark).

The MLE<sup>ARB1</sup>-flag construct was obtained by digestion of pFastBac1-MLE-flag plasmid with BamHI and StuI, followed by incubation with the DNA Polymerase I (Klenow) to yield blunt ends and religation.

The MLE<sup>ARB1-2</sup>-flag construct was obtained by digestion of pFastBac-MLE-flag with SphI. The SphI MLE DNA fragment was subsequently subcloned into the pFast-Bac1 vector, SphI digested.

The MLE<sup>ARGG</sup>-flag construct was obtained in a multistep procedure: a BamHI/EcoRI DNA fragment was initially cut out from the pFastBac-MLE-flag construct and inserted into the pFastBac1 vector to obtain the pFastBac-MLE<sup>ARGG</sup> construct. Subsequently, the

pFastBac-MLE<sup>ARGG</sup> vector was digested BamHI/SacI and the correspondent 3543 bp DNA fragment was inserted into the pFastBac-MLE-flag construct, lacking the BamHI/SacI fragment, to obtain the pFastBac-MLE<sup>ARGG</sup>-flag construct.

The MLE<sup>ARB2</sup>-flag construct has been produced as follows: in a first step a 369 bp DNA fragment, containing the first double strand RNA binding domain (dsRB1) of MLE protein, was digested with BamHI and StuI from the pFastBac-MLE-flag construct and cloned into the pFastBac1, cut with the same enzymes, to obtain the pFastBac-MLE<sup>RB1</sup> plasmid. In a second step, pFastBac-MLE-flag construct was digested with HindIII, followed by incubation with DNA Polymerase I (Klenow) to yield blunt ends and XbaI. The resulting HindIII/XbaI 3210 bp DNA fragment was inserted into the pFastBac-MLE<sup>RB1</sup> plasmid digested with StuI and XbaI.

MLE-SacI.rv: 5'-TATGCAGAGCTCTTTCGATTTCGAACCC-3'

MLE-BamHI.fw: 5'-CGCGGGATCC ATGGATATAA AATC-3'

### 2.1.5. Wild type MLE-GFP and MLE-GFP mutants

The complete *MLE* sequence was amplified by PCR using the MLE-XmnI forward primer and the MLE-AgeI reverse primer. The PCR product was cloned into pEGFP-1 plasmid (Clontech) digested with XmnI and AgeI.

The MLE<sup>GET</sup>-GFP construct was obtained by replacing the 3070 bp BstEII/PstI DNA fragment in MLE-GFP with the corresponding DNA fragment containing the K413E mutation, cut out from the pFastBac-MLE<sup>GET</sup>-flag plasmid.

The MLE<sup>ARB1</sup>-GFP and MLE<sup>ARB1-2</sup>-GFP plasmids were obtained by inserting the StuI/AgeI and the DraI/AgeI DNA fragments from MLE-GFP respectively into the pEGFP plasmid AgeI/XmnI digested.

The MLE<sup>ARGG</sup>-GFP plasmid was obtained by digestion of MLE-GFP plasmid with ScaI/AgeI, followed by incubation with the DNA Polymerase I enzyme (Klenow) to yield blunt ends and religation.

The MLE<sup>ARB2</sup>-GFP construct was obtained by insertion of the BglII/StuI DNA fragment from MLE-GFP into MLE-GFP plasmid digested with HindIII/BglII and blunt for HindIII. The MLE<sup>RGG</sup>-GFP plasmid was obtained by PCR amplification of the DNA fragment, containing the RGG box, with XmnI-RGG.fw and AgeI-MLE.rv primers

using MLE-GFP as a template. The PCR product was digested XmnI/AgeI and inserted into the pEGFP plasmid XmnI/AgeI digested.

The MLE<sup>ARGG</sup>-GFP-NLS construct was obtained by insertion of an artificial NLS after the GFP tag into the MLE<sup>ARGG</sup>-GFP plasmid. The DNA fragment containing the NLS sequence from the simian virus large T-antigen (Lanford et al., 1986), was obtained by annealing of two complementary primers (NLS-NotI.fw and NLS-BsRGI.rv), containing a NotI and a BsRGI site respectively.

MLE-AgeI.rv: 5'-TATAACCGGT CGGATCCAAG CTAATCG-3'

MLE-XmnI.fw: 5'-CGCGAACCCCTTCACCCCATGGATATAAAA-3'

MLE-RGG-XmnI.fw: 5'-CGGAATACCT TCCCATGGTG CCGCACCAAT-3'

MLE-RGG-BamHI.fw: 5'-ATGGATCCCC ATGGTGCCGC ACCAATCA-3'

NLS-NotI.fw: 5'-GTACAAGGATCCAAAAAGAAGAGAAAGGTATAAGC-3'

NLS-BsRGI.rv: 5'-GGCCGCTTATACCTTCTCTCTTTTTGGATCCTT-3'

### 2.1.6. Cloning of *ms2-roX2* under the endogenous promoter

Since the promoter region of *roX2* is not well defined, for the cloning of the *ms2-roX2* sequence under the control of the endogenous promoter a big DNA fragment 5' of the *roX2* transcription start site was amplified by PCR using a reverse primer containing the StuI site (*roX2-StuI.rv*) and a forward primer containing the KpnI site (*roX2-KpnI.fw*). Genomic DNA, extracted from SF4 cells was used as template. A reverse primer containing the EcoRI site (*roX2-EcoRI.rv*) and a forward primer containing the StuI site (*rox2-StuI.fw*) were used to amplify the *roX2* sequence. Both amplified products were cloned into pBL-KS plasmid (Stratagene), digested with KpnI and EcoRI enzymes. After cloning, BamHI/AvaI/StuI sites were generated just before the first splice site in *roX2*. Two complementary primers, containing the *ms2* sequence (Bertrand et al., 1998) were annealed and multimerized into the previous construct using the AvaI site. A maximum of six *ms2* sites could be inserted into the *rox2* construct in front of the first splice site.

*roX2-EcoRI.rv*: 5'-CGCGTCAACCAATGAAAACAA-3'

*roX2-StuI.fw*: 5'-AAGGCCTGAGCGTTTAGGTAAGCAAAC-3'

*roX2-StuI.rv*: 5'-AAGGCCTCCCAGGATCCGACTTATGATATAAAAAGTAGCCC-3'

roX2-KpnI.fw: 5'-GGGGTACCGGGCGACATGAATGACTGA-3'

ms2.rv: 5'-TCGGGCACTGTACATGAGGATCACCCATGTCCGTACGGACAGTGC-3'

ms2.fw: 5'-CGTGACATGTACTCCTAGTGGGTACAGGCATGCCTGTCACGAGCC-3'

### 2.1.7. GFP-MS2

The GFP-MS2 plasmid was a gift from E. Gavis (Princeton University, Princeton, NJ, USA). Cloning is described in Forrest et al., 2003.

## 3. Biochemical methods

### 3.1. *In vitro* transcription and translation

*In vitro* transcription and translation (IVT) reactions were performed for 3 h at 30°C with the TNT System (Promega) using 1 µg of plasmid and 10 µCi of [<sup>35</sup>S] methionine in 25 µl reaction volume.

### 3.2. Expression and purification of GST fusion proteins from bacteria

GST, GST-MLE (1-265) and GST-MS2 expressing constructs were used to transform *E. coli* cells BL21. Bacteria were grown in 1 l of LB medium, containing ampicillin (100 mg/ml), to  $A_{600} = 0,6$ . The expression of GST-MLE recombinant protein was induced by 1 mM IPTG for 2 h. Harvested bacteria were subjected to one freeze/thaw cycle at -80°C and resuspended in 20 ml Lysis Buffer A (PBS, 100 mM NaCl, 1 mM DTT, 1% Triton X, 10 % glycerol, protease inhibitor mix (PMSF, leupeptine, aprotinin, pepstatin)). After incubation on ice for 15 min, the cell suspension was sonicated 3 times by pulses of 20 sec at 45% amplitude (Branson digital sonifier model 250-D), and centrifuged for 30 min at 14 krpm at 4°C. The supernatant was incubated 4 h on a rotating wheel at 4°C, with glutathione beads (400 µl, Pharmacia Biotech) pre-equilibrated in a 1:1 ratio in Lysis Buffer A. Beads were washed 4 times for 10 min on a rotating wheel at 4°C with Lysis Buffer B (PBS, 200 mM NaCl, 1 mM DTT, 1% Triton

X, 10 % glycerol, protease inhibitor mix (PMSF, leupeptine, aprotinin, pepstatin)). Beads were applied to a column (Pharmacia Biotech) and washed once with 10 ml of Wash Buffer I (50 mM TrisHCl pH 7.5, 600 mM NaCl, 1 mM DTT) and once with 10 ml of Wash Buffer II (50 mM TrisHCl pH 7.5, 300 mM NaCl, 1 mM DTT) respectively. Recombinant GST-MLE protein was eluted from beads in 0.5 ml fractions, using 3 ml of Elution Buffer (50 mM TrisHCl pH 7.5, 300 mM NaCl, 1 mM DTT, 5 mM glutathione and 50% glycerol). Aliquots of each fraction were separated on 10% SDS polyacrilamide gel and stained by Coomassie Blue. Fractions containing recombinant GST-MLE were collected and protein concentration was estimated on Coomassie Blue stained SDS-PAGE, using BSA (Biolabs) as a standard.

### **3.3. Expression and purification of proteins and RNA from Sf9 cells**

#### **3.3.1. Sf9 cells infection**

Sf9 cells were grown in Sf-900 II medium (Gibco) at 26°C.  $1 \times 10^6$  cells were seeded onto a 35 mm diameter dish and transfected with the respective bacmid plasmids following the manufacture instructions. After 5 days of incubation at 26°C, supernatants containing the respective baculoviruses were collected and amplified up to three times following the manufacture instructions. Viruses were stored at 4°C.

The optimal amount of each virus, used to infect Sf9 cells was determined empirically. The cells were collected after 2 days of incubation at 26°C. In general, the cell pellets were frozen in liquid nitrogen and stored at -80°C prior to protein or RNA purification.

#### **3.3.2. Whole Sf9 cell extracts preparation and single protein purification**

Whole cell extract was prepared by resuspending Sf9 cell pellets in 1 ml Lysis Buffer (50 mM Hepes pH 7.6, 300 mM KCl, 0.5 mM EDTA, 0.1% NP40, 1 mM DTT, protease inhibitor mix (PMSF, leupeptine, aprotinin, pepstatin)) per dish and incubated for 15 min on ice. The cell suspension was sonicated by three pulses of 15 sec at 15% amplitude (Branson digital sonifier model 250-D), and centrifuged for 30 min at 14 krpm at 4°C. Supernatants (0.5 ml) were incubated for 2 h on a rotating wheel with

affinity beads (25  $\mu$ l of a 1:1 slurry in Lysis Buffer) according to the tag of the protein: anti-flag M2 beads (Sigma) for flag-MSL3, MLE-flag and MLE-flag mutants, anti-HA (Roche) beads for HA-MOF and glutathione beads (Amersham) for GST-MSL1. The beads were washed five times with 0.5 ml of each of the three following buffers: Buffer I: 0.3 M KCl, 50 mM Hepes pH 7.6, 0.5 mM EDTA, 1% NP40, 0.5 mM DTT, protease inhibitor mix; Buffer II: Buffer I with 1 M KCl; Buffer III: 0.15 M KCl, 10 mM Hepes pH 7.6, 0.5 mM EDTA, 0.1% NP40, 10% glycerol, 0.5 mM DTT, protease inhibitor mix. Proteins were eluted (two times for 4 h on a rotating wheel at 4°C) in 100  $\mu$ l of Buffer III supplemented with flag peptides (0.25 mg/ml) for flag-MSL3, MLE-flag, MLE<sup>GET</sup>-flag, MLE<sup>ARB1</sup>-flag, MLE<sup>ARB2</sup>-flag, MLE<sup>ARB1-2</sup>-flag and MLE<sup>ARGG</sup>-flag or HA peptides (0.3 mg/ml) for HA-MOF. For GST-MSL1, elution was achieved with 20 mM glutathione at pH 8.0. Eluted proteins were stored at -80°C.

### **3.3.3. Whole Sf9 cell extracts preparation and protein complexes purification**

For the interaction of MLE-flag with individual MSL proteins coexpressed in Sf9 cells, extracts were prepared as in 3.3.2 with the following differences: extracts were not sonicated, but frozen and thawed three times in liquid nitrogen. Before Pull-down experiments, the KCl concentration was adjusted to 0.15 M by addition of Buffer I without KCl. After Pull-down, beads were washed 7 times in Buffer I (0.15 M KCl) and resuspended in Loading Buffer. 2.5% of input and 50% of beads were loaded onto an 8% SDS-PAGE.

### **3.3.4. Whole Sf9 cell extracts preparation and protein/RNA complexes purification**

For the extraction of the MSL complex in the presence of roX2 or ant-roX2 RNAs, RNAsin (Promega) was added in the Lysis Buffer I, at a concentration of 1 U/ml. After 15 min of incubation in ice, the cell suspensions were frozen and thawed three times in liquid nitrogen, prior to centrifugation for 30 min at 14 krpm at 4°C. The respective cell extracts were diluted to 0.15 M KCl by the addition of Buffer I without KCl. 20% of total Sf9 extract (input) was subjected to RNA extraction and RT-PCR or to Western-Blot analysis, to control for correct expression of roX RNAs and MSL proteins, respectively. 0.5 ml of extract was used for the RNA/protein complex purification, by

the addition of affinity flag beads for 2 h at 4°C on a rotating wheel. Beads were then briefly spun down and washed 5 times in buffer I (0.5 M KCl), containing RNasin 1 U/ml. The bound material was eluted from beads using flag peptides, as described above. 10% of the eluted material was boiled for 5 min in denaturing Loading Buffer and analyzed by Coomassie-stained SDS-PAGE or Western-Blot. Total RNA was extracted from half of the eluted material and the presence of roX2 or ant-roX2 RNAs was detected by RT-PCR with specific roX2 primers.

### **3.4. Whole SF4 cells extract preparation**

SF4 cells were grown in 500 ml of Schneider's medium (Gibco) to a density of  $5 \times 10^6$  cells/ml, collected and resuspend in 2 ml of Extraction Buffer (30 mM Hepes, 250 mM KCl, 2 mM  $MgCl_2$ , 1 mM EGTA, 10% glycerol, 1mM PMSF and protease inhibitor mix (leupeptine, aprotinin, pepstatin)). After 15 min on ice, the cell suspension was centrifuged at 14 Krpm for 30 min at 4°C. Total protein concentration was estimated using the Bio-Rad Protein assay according to the manufacture instructions. Total proteins (0.1 mg) were separated on an 8% polyacrilamide gel and transferred onto a Hybond-P PVDF membrane (Amersham Biosciences). Western-Blot analysis was performed following standard protocols. In particular, after blocking, the membrane was cut in slices and each slice was incubated with 50  $\mu$ l of the respective MLE monoclonal antibody, diluted in PBS, 0.1% Tween-20.

### **3.5. Immunoprecipitation experiments**

0.4 ml of total SF4 extract (12 mg/ml) were diluted by addition of 0.5 ml of Binding buffer (30 mM Hepes, 2 mM  $MgCl_2$ , 1 mM EGTA, 10% glycerol, 1 mM PMSF protease inhibitor mix (leupeptine, aprotinin, pepstatin) and 0.1% NP40) to adjust the KCl concentration to 100 mM. The diluted extract was precleared for 1 h at 4°C on a rotating wheel, using 50  $\mu$ l of protein A/protein G mix affinity beads (Pharmacia Biotech), pre-equilibrated in Binding Buffer in a 1:1 ratio. After a brief centrifugation, aliquots (125  $\mu$ l) from the precleared extract were incubated with 50  $\mu$ l of each MLE antibody for 3 h at 4°C on a rotating wheel. 50  $\mu$ l of protein A/protein G affinity beads

were then added for 2 more hours. Beads were briefly centrifuged and washed 5 times in binding buffer. 5% of extract before immunoprecipitation (input), 5% of extract after immunoprecipitation (supernatant) and 15% of beads were separated by 8% SDS-PAGE. Western-Blot analysis was performed following standard protocols using rabbit anti-MLE polyclonal antibodies (a gift from M. Kuroda).

### **3.6. Pull down experiments**

8  $\mu$ l of MSL1 or MSL2 IVT reaction was mixed in 172  $\mu$ l of buffer I (150 mM KCl). 9  $\mu$ l of the mix was removed (input), and the remainder was split into two 85  $\mu$ l aliquots, incubated for 2-4 h at 4°C (rotating wheel) with anti-flag beads as negative controls or anti-flag beads and MLE-flag (100 ng). Beads were washed 7 times with 0.5 ml of Buffer I (150 mM KCl). Bound and unbound materials were analyzed by SDS-PAGE and autoradiography of the corresponding dried gel.

To test the interaction of MLE-flag with GST-MSL1 or HA-MOF, equimolar amounts (100 ng) of the respective proteins were incubated in 100  $\mu$ l of Lysis Buffer I (150 mM KCl) 2 h at 4°C on a rotating wheel, in the presence of flag affinity beads (30  $\mu$ l). After pull-down, beads were washed 7 times in Lysis Buffer I (150 mM KCl) and resuspended in Loading Buffer. 1% of input and 50% of bound material were loaded on a 8% SDS-PAGE and protein were visualized by Western-Blot.

Interaction of MLE-flag and flag-MSL3 was assayed under the same Pull-down condition, but 5  $\mu$ l of monoclonal MLE antibody (6E11) were used to immunoprecipitate MLE.

### **3.7. GST-MS2 Pull down experiments**

500  $\mu$ l of total SF4 extract (12 mg/ml) or *ms2-roX2* extract were incubated with comparable amounts of GST- and GST-MS2-coated glutathione beads (100  $\mu$ l) 3 h at 4°C on a rotating wheel. After 5 washes with Extraction Buffer, half of the beads were

separated on a 10% SDS-PAGE, Silver stained or analyzed by Western-Blot. The other half was used to purify total RNA as described in 3.8.

### **3.8. RNA extraction and RT-PCR**

For the RNA extraction, the respective materials were incubated 30 min at 65°C in the presence of Proteinase K (1 mg/ml) and 0.1% SDS. The protein component was eliminated by Trizol (Invitrogen) extraction following the manufacture instructions. Nucleic acids were resuspended in DEPC-treated water. The SuperScript II Reverse Transcriptase kit (Invitrogen) was used for the cDNA synthesis. Specific roX2 reverse (mut1-XhoI.rv) and forward (mut1-SmaI.fw) primers were used to reverse transcribe roX2 and ant-roX2 RNAs, respectively. Both primers were used together for the PCR reaction (94°C for 5 min; 30 cycles at 94°C for 30 sec, 60°C for 30 sec and 72°C for 2 min; 72°C for 7 min). The presence of the ms2-roX2 RNA in the correspondent stable cell line was determined using a roX2 specific primer (roX2c.rv) for the reverse transcription and the same primer with a specific ms2 primer (ms2-1d.fw), for the PCR reaction (5 min at 94°C; 25 cycles 30 sec at 94°C, 30 sec at 55°C and 1 min at 72°C; 7 min at 72°C). In all cases, PCR products were electrophoresed on a 0.8% agarose gel in TBE.

Mut1-SmaI.fw: 5'-CGGCCCGGGATTGGTGTACATATAGCT-3'

Mut1-XhoI.rv: 5'-GCCCTCGAGGCTCTAAATCGATCTGTA-3'

ms2-1d.fw: 5'-CAGCAGCAAAATCAAGTGAATCAT-3'

roX2c.rv: 5'-TTGTCATCTCGCTCTAAATCG-3'

### **3.9. ATPase assay**

The ATPase activity of MLE-flag (30 fmole) and the corresponding mutant proteins was measured in a reaction mixture (15 µl), containing 6.6 mM HEPES pH 7.6, 0.66 mM EDTA, 0.66 mM 2-mercaptoethanol, 0.033% NP40, 1.1 mM MgCl<sub>2</sub>, 33 mM ATP, 5 Ci [ $\gamma$ -<sup>32</sup>P] ATP (3000 Ci mmol<sup>-1</sup>, NEN). When indicated, 30 fmoles of the respective MSL proteins were added as well. To stimulate the MLE ATPase activity, total RNA

was extracted from Sf9 cells, or Sf9 cells infected with either *roX2* or *ant-roX2* baculoviruses using the Rnase Mini Kit (Qiagen), according to the manufacture instructions. The indicated amounts of RNA were added in the ATPase reaction. After 30 min at 26°C, aliquots (1 µl) of each reaction were spotted onto Thin Layer Chromatography (TCL) plates (Merck). ATP and phosphate were separated by chromatography in 1M formic acid, 0.5 M LiCl for about 10 min. Plates were then dried in a oven at 65°C for 5 min and exposed onto a PhosphorImager screen.

For the kinetic experiments, reaction mixtures were incubated at 26°C for 10 min, 20 min and 40 min, respectively.

### **3.10. Preparation of RNA substrates for the helicase and the gel-mobility shift assays**

The SP6.roX2 reverse primer, containing a binding site for the SP6 RNA Polymerase and the T7.roX2 forward primer, containing a binding site for the T7 RNA Polymerase were used to amplify a 75 bp *roX2* DNA sequence, predicted to not form double strand RNA structures (RNA fold program; <http://rna.tbiunivie.ac.at/cgi-bin/RNAfold.cgi>). The *roX2.78.2.2* cDNA (a gift from M. Kuroda) was used as a template for the PCR reaction. The PCR product was purified using the MinElute Spin Columns (Qiagen). 1 µg of the PCR product was transcribed *in vitro* in the presence of ATP, GTP, UTP (2.5 mM each), CTP (100 mM), [ $\alpha^{32}\text{P}$ ]-CTP 1 mM DTT, 1 µl of T7 RNA Polymerase (Promega), 1 U/ml RNAsin (Promega) and 1x Transcription Buffer (Promega) in 25 µl of final volume. The reaction mixture was incubated at 37°C for 2 h and the radiolabeled ssRNA was purified in 100 µl of DEPC-treated water, using the mini Quick Spin RNA Columns (Roche). The RNA concentration was measured spectrophotometrically. In most cases, the specific activity of the labeled ssRNA substrate was 10<sup>2</sup> cpm/fmole. The dsRNA substrate was prepared by annealing of the radiolabeled ssRNA with the corresponding cold complementary ssRNA, which had been transcribed under standard condition using the PCR product as a template for the SP6 RNA Polymerase. The radioactive ssRNA was incubated with a 5 fold excess of cold complementary RNA in a 200 µl of Hybridization Buffer (20 mM Hepes-KOH, pH 7.6, 0.5 M NaCl, 1 mM EDTA, 0.1% SDS) 5 min at 100°C, 30 min at 65°C and then left cool down to room temperature. The dsRNA was then purified by electrophoresis

on a native 10% polyacrylamide (30:1) gel, 1x TBE. The gel slice, containing the dsRNA transcript (located by autoradiography) were ground with a micropestle and eluted in 0.4 ml of Elution Buffer (0.5 M ammonium acetate, pH 7, 0.1% SDS, 10 mM EDTA) 4 h at room temperature in a shaking incubator. After a brief centrifugation the supernatant was extracted with phenol/chloroform and treated with 2.5 volumes of absolute ethanol. After 1 h at -80°C, the transcript was collected by centrifugation at 4°C for 30 min. The pellet was washed with 0.5 ml of 70% ethanol and resuspend in the appropriate volume of DEPC-treated water.

SP6-roX2.rv: 5'-CGATTTAGGTGACACTATAGAAATATTTGCTTAATTTGC-3'

T7-roX2.fw: 5'-CGTTAATACGACTCACTATAGGGAGACGTGTAATAATGTT-3'

### **3.11. Helicase assay**

RNA helicase activity was measured in a reaction mixture containing 20 mM HEPES-NaOH, pH 7.5, 2 mM DTT, 3 mM MgCl<sub>2</sub>, 1 mM ATP (when indicated), 0.1 mg/ml BSA (NEB) 1 U/ml RNasin (Promega), 120 fmoles of dsRNA substrate and increasing amounts (35.3, 75, 150 and 300 fmoles) of MLE-flag, MLE<sup>ARB1</sup> and MLE<sup>ARGG</sup> purified proteins. Upon annealing, the dsRNA was not purified, but directly used in the helicase assay. Enzymatic activities were then measured in the presence of a 5 fold excess of ssRNA. Mixtures were incubated 30 min at 37°C, then stopped by the addition of 5 ml of 5x Stop Buffer (0.1 M TrisHCl pH 7.5, 20 mM EDTA, 0.5% SDS, 0.1% NP40, 0.1% bromophenol blue, 0.1% xylene cyanol and 50% glycerol). Aliquots (10 µl) of each reaction were loaded onto a 10% polyacrylamide (30:1) gel and electrophoresed at 20 mA for 2-3 h. The dsRNA and the displaced ssRNA products were visualized by autoradiography of the dried gel.

### **3.12. Gel mobility shift assay**

RNA gel mobility shift assays were performed as follows: a mixture containing radiolabeled ssRNA or gel purified dsRNA (25-50 fmoles) in Band-shift Buffer (20 mM HEPES-KOH pH 7.6, 3 mM MgCl<sub>2</sub>, 10% glycerol, 1mM DTT, 0.1 mg/ml BSA (NEB))

was prepared and divided into four aliquots. To the first aliquot 1 mM of ATP was added; to the second 1 mM of ATP $\gamma$ S; to the third 1 mM of ADP and to the fourth only water (control). For each of them, the volume was adjusted to 64  $\mu$ l with water. In parallel, each recombinant protein (25-50 fmoles) was diluted in Buffer I (150 mM KCl) and divided into 4 aliquots of 4  $\mu$ l each. 16  $\mu$ l from the ssRNA/nucleotide mixes, were added to the 4  $\mu$ l, containing the recombinant proteins. After incubation for 5-10 min on ice, 10  $\mu$ l of each reaction were electrophoresed on a 0.6% agarose gel in 0.3X TBE at 20 mA for about 1.5 h. Gels were dried and the RNA was visualized by autoradiography.

#### **4. Cell biology methods**

##### **4.1. Cell culture**

*Drosophila* SF4 cells were obtained from D. Arndt-Jovin (Max Planck Institute for Biophysical Chemistry, Göttingen, Germany). Cells were kept at 26°C in Schneider's *Drosophila* medium (Invitrogen), supplemented with penicillin, streptomycin, glutamine and 10% fetal calf serum (FCS). The cell density was maintained to approximately  $4 \cdot 10^6$  cell/ml by splitting them in 20 ml of Schneider's *Drosophila* medium (Invitrogen) at a 1:10 ratio.

Sf9 cells were kept at 26°C in Sf-900 II medium (Invitrogen) supplemented with penicillin and streptomycin.

##### **4.2. Transient transfection in SF4 cells**

In all cases transient transfections were performed using the Effectine Reagent Kit (Qiagen), following the manufacture instructions.

For the visualization of MLE and deletion mutants fused to GFP, SF4 cells ( $1 \cdot 10^6$ ) were transiently transfected with 0.4  $\mu$ g of the respective DNA plasmid in six-well plates. After 48 h of incubation, cells were immunostained as described previously with the following modifications: anti-MSL1 antibodies (a gift of M. Kuroda) were used to mark

the X territory and mouse anti-GFP antibody (Molecular Probes) to visualize the fusion proteins. DNA was counterstained with TOPRO 3. Images were acquired with a Zeiss LSM 510 META confocal microscope.

For the visualization of the MS2-GFP protein in the *ms2-roX2* stable cell line, different amounts (125 ng, 250 ng, 500 ng and 1  $\mu$ g) of the MS2-GFP construct were used. Cotransfected cells were generally examined after 2-4 days and pictures were taken at 1200 magnification using a Zeiss AxiopHot microscope coupled to a Retiga Exi CCD Camera (Qimaging, Burnaby, Canada). The MS2-GFP expressing cells were visualized using the Zeiss filterset 44.

#### **4.3. Establishment of the *ms2-roX2* stable cell line**

For the preparation of the *ms2-roX2* stable cell lines, SF4 cells ( $4 \times 10^6$ /10 cm diameter dish) were cotransfected with 1  $\mu$ g of the *ms2-roX2* plasmid and 0.2  $\mu$ g pUCHS-NEO (coding for neomycin resistance). After 48 h at 26°C, 1 mg/ml of neomycin was added and selective growth was maintained for 4 weeks. As a control, untransfected SF4 cells were incubated under the same induction conditions.

#### **4.4. Immunofluorescence on SF4 and SL2 cells**

SF4 or SL2 cells ( $1 \times 10^6$ ) were seeded onto coverslips and allowed to attach for 2 h. Subsequently, cells were fixed in 2% paraformaldehyde (PFA) in PBS for 10 min on ice. Following permeabilization in PBS supplemented with 0.25% Triton-X and 1% PFA for 8 min on ice, cells were washed twice in PBS. All subsequent steps were done at room temperature. Coverslips were blocked with 2% BSA and 5% goat serum in PBS for 1 h. Cells were incubated, as needed, with primary antibodies, for 1 h at room temperature in a humid chamber. After washing in PBS, slides were stained for 1 h with the appropriate Cy3 and Cy2 labeled secondary antibodies (Jackson) diluted in Blocking Buffer. Cells were washed four times in PBS. DNA was counterstained with 1 mg/ml bisbenzimidazole (Hoechst 33258). Cells were mounted using 1.5% n-propyl gallate and 60% glycerol in PBS. Cells were examined and pictures were taken at 1200 magnification using a Zeiss AxiopHot microscope coupled to a Retiga Exi CCD

Camera (Qimaging, Burnaby, Canada). Images were processed in Adobe PHotoshop.

#### 4.5. RNAi in SL2 cells

A DNA plasmid containing *MLE* cDNA was used as a template in a PCR reaction, with MLE-RNAi.fw and MLE-RNAi.rv specific primers

GST-RNAi.fw and GST-RNAi.rv primers served to amplify glutathione *S*-transferase (GST). All primers contained a binding site for T7 RNA Polymerase. PCR products were purified using the PCR Clean-Up Kit (Qiagen) according to the manufacturer instructions. The purified DNA fragments were used as templates to transcribe *in vitro* both sense and anti-sense MLE or GST mRNA, respectively, using the MEGAscript T7 kit (Ambion). A yield of approximately 100 µg of total RNA was generally obtained.

To denature secondary structures, RNA were heated for 30 min at 65°C, and slowly cooled down to room temperature to form the dsRNA duplexes. A total of 2\*10<sup>6</sup> SL2 cells were seeded into 6 well plates in 1 ml of Schneider's *Drosophila* medium without fetal calf serum (FCS) just before the addition of 10 µg of dsRNA. Plates were placed onto a shaking platform for 10 min and then for 50 min at 26°C. 2 ml of medium supplemented with FCS were then added to the cells and incubated at 26°C. Cells were collected after 3 or 6 days after dsRNA treatment. For Western-Blot analysis after RNAi, 1-2 \*10<sup>6</sup> cells were centrifuged and resuspended in 50 µl of Urea Buffer (8 M urea, 5% SDS, 200 mM Tris-HCl pH 6.8, 0.1 mM EDTA, 100 mM DTT), incubated at 65°C for 15 min and loaded on a 8% polyacrilamide gel.. Detection and quantification was performed using the Odyssey Infrared Imaging system. For the immunostaining experiments, cells were treated as described above.

RNAiMLE.fw: 5'-TTAATACGACTCACTATAGGGAGAATGGATATAAAAATCTTTTTGTACCAATTTTG-3'

RNAi-MLE.rv: 5'-TTAATACGACTCACTATAGGGAGAACAGGGCGCATGACTTGCT-3'

RNAi-GST.fw: 5'-TTAATACGACTCACTATAGGGAGAATGTCCCCTATACTAGGTTA-3'

RNAi-GST.rv: 5'-TTAATACGACTCACTATAGGGAGAACGCATCCAGGCACATTG-3'

## V. Bibliography

Adkins, M. W., Howar, S. R., and Tyler, J. K. (2004a). Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes. *Mol Cell* *14*, 657-666.

Adkins, N. L., Watts, M., and Georgel, P. T. (2004b). To the 30-nm chromatin fiber and beyond. *Biochim Biophys Acta* *1677*, 12-23.

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

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.

Ausio, J., and Abbott, D. W. (2002). The many tales of a tail: carboxyl-terminal tail heterogeneity specializes histone H2A variants for defined chromatin function. *Biochemistry* *41*, 5945-5949.

Badenhorst, P., Voas, M., Rebay, I., and Wu, C. (2002). Biological functions of the ISWI chromatin remodeling complex NURF. *Genes Dev* *16*, 3186-3198.

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

Bannister, A. J., Schneider, R., and Kouzarides, T. (2002). Histone methylation: dynamic or static? *Cell* *109*, 801-806.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* *410*, 120-124.

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

Bates, G. J., Nicol, S. M., Wilson, B. J., Jacobs, A. M., Bourdon, J. C., Wardrop, J., Gregory, D. J., Lane, D. P., Perkins, N. D., and Fuller-Pace, F. V. (2005). The DEAD box protein p68: a novel transcriptional coactivator of the p53 tumour suppressor. *Embo J* *24*, 543-553.

Becker, P. B., and Horz, W. (2002). ATP-dependent nucleosome remodeling. *Annu Rev Biochem* 71, 247-273.

Beisel, C., Imhof, A., Greene, J., Kremmer, E., and Sauer, F. (2002). Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1. *Nature* 419, 857-862.

Belote, J. M., and Lucchesi, J. C. (1980). Male-specific lethal mutations of *Drosophila melanogaster*. *Genetics* 96, 165-186.

Benbow, R. M. (1992). Chromosome structures. *Sci Prog* 76, 425-450.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H., and Long, R. M. (1998). Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 2, 437-445.

Bevilacqua, P. C., George, C. X., Samuel, C. E., and Cech, T. R. (1998). Binding of the protein kinase PKR to RNAs with secondary structure defects: role of the tandem A-G mismatch and noncontiguous helices. *Biochemistry* 37, 6303-6316.

Bhadra, U., Pal-Bhadra, M., and Birchler, J. A. (1999). Role of the male specific lethal (*msl*) genes in modifying the effects of sex chromosomal dosage in *Drosophila*. *Genetics* 152, 249-268.

Birchler, J. A., Pal-Bhadra, M., and Bhadra, U. (2003). Dosage dependent gene regulation and the compensation of the X chromosome in *Drosophila* males. *Genetica* 117, 179-190.

Black, B. E., Foltz, D. R., Chakravarthy, S., Luger, K., Woods, V. L., Jr., and Cleveland, D. W. (2004). Structural determinants for generating centromeric chromatin. *Nature* 430, 578-582.

Brockdorff, N., Ashworth, A., Kay, G. F., Cooper, P., Smith, S., McCabe, V. M., Norris, D. P., Penny, G. D., Patel, D., and Rastan, S. (1991). Conservation of position and exclusive expression of mouse *Xist* from the inactive X chromosome. *Nature* 351, 329-331.

Brownell, J. E., and Allis, C. D. (1996). Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr Opin Genet Dev* 6, 176-184.

Buscaino, A., Kocher, T., Kind, J. H., Holz, H., Taipale, M., Wagner, K., Wilm, M., and Akhtar, A. (2003). MOF-regulated acetylation of MSL-3 in the *Drosophila* dosage compensation complex. *Mol Cell* 11, 1265-1277.

Byrd, A. K., and Raney, K. D. (2004). Protein displacement by an assembly of helicase molecules aligned along single-stranded DNA. *Nat Struct Mol Biol* *11*, 531-538.

Carlson, C. B., Stephens, O. M., and Beal, P. A. (2003). Recognition of double-stranded RNA by proteins and small molecules. *Biopolymers* *70*, 86-102.

Carmel, A. B., and Matthews, B. W. (2004). Crystal structure of the BstDEAD N-terminal domain: a novel DEAD protein from *Bacillus stearothermophilus*. *Rna* *10*, 66-74.

Caruthers, J. M., and McKay, D. B. (2002). Helicase structure and mechanism. *Curr Opin Struct Biol* *12*, 123-133.

Celotto, A. M., and Graveley, B. R. (2004). RNA interference of mRNA processing factors in *Drosophila* S2 cells. *Methods Mol Biol* *257*, 245-254.

Chan, R. C., Severson, A. F., and Meyer, B. J. (2004). Condensin restructures chromosomes in preparation for meiotic divisions. *J Cell Biol* *167*, 613-625.

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

Chao, K., and Lohman, T. M. (1990). DNA and nucleotide-induced conformational changes in the *Escherichia coli* Rep and helicase II (UvrD) proteins. *J Biol Chem* *265*, 1067-1076.

Cheng, W., Brendza, K. M., Gauss, G. H., Korolev, S., Waksman, G., and Lohman, T. M. (2002). The 2B domain of the *Escherichia coli* Rep protein is not required for DNA helicase activity. *Proc Natl Acad Sci U S A* *99*, 16006-16011.

Cheng, X., Collins, R. E., and Zhang, X. (2005). Structural and sequence motifs of protein (histone) methylation enzymes. *Annu Rev Biophys Biomol Struct* *34*, 267-294.

Chow, J. C., Yen, Z., Ziesche, S. M., and Brown, C. J. (2005). Silencing of the Mammalian X Chromosome. *Annu Rev Genomics Hum Genet*.

Chu, D. S., Dawes, H. E., Lieb, J. D., Chan, R. C., Kuo, A. F., and Meyer, B. J. (2002). A molecular link between gene-specific and chromosome-wide transcriptional repression. *Genes Dev* *16*, 796-805.

Chu, F., Nusinow, D. A., Chalkley, R. J., Plath, K., Panning, B., and Burlingame, A. L. (2005). Mapping post-translational modifications of the histone variant macroH2A1 using tandem mass spectrometry. *Mol Cell Proteomics*.

Clayton, A. L., and Mahadevan, L. C. (2003). MAP kinase-mediated phosphoacetylation of histone H3 and inducible gene regulation. *FEBS Lett* 546, 51-58.

Cobianchi, F., and Wilson, S. H. (1987). Enzymes for modifying and labeling DNA and RNA. *Methods Enzymol* 152, 94-110.

Cohen, H. R., Royce-Tolland, M. E., Worringer, K. A., and Panning, B. (2005). Chromatin modifications on the inactive X chromosome. *Prog Mol Subcell Biol* 38, 91-122.

Connolly, K. M., Wojciak, J. M., and Clubb, R. T. (1998). Site-specific DNA binding using a variation of the double stranded RNA binding motif. *Nat Struct Biol* 5, 546-550.

Copps, K., Richman, R., Lyman, L. M., Chang, K. A., Rampersad-Ammons, J., and Kuroda, M. I. (1998). Complex formation by the Drosophila MSL proteins: role of the MSL2 RING finger in protein complex assembly. *Embo J* 17, 5409-5417.

Cordin, O., Banroques, J., Tanner, N. K., and Linder, P. (2005). The DEAD-box protein family of RNA helicases. *Gene*.

Cordin, O., Tanner, N. K., Doere, M., Linder, P., and Banroques, J. (2004). The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *Embo J* 23, 2478-2487.

Corona, D. F., Clapier, C. R., Becker, P. B., and Tamkun, J. W. (2002). Modulation of ISWI function by site-specific histone acetylation. *EMBO Rep* 3, 242-247.

Cosgrove, M. S., and Wolberger, C. (2005). How does the histone code work? *Biochem Cell Biol* 83, 468-476.

Csankovszki, G., McDonel, P., and Meyer, B. J. (2004). Recruitment and spreading of the *C. elegans* dosage compensation complex along X chromosomes. *Science* 303, 1182-1185.

Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P. D., Tempst, P., Bannister, A. J., and Kouzarides, T. (2004). Histone deimination antagonizes arginine methylation. *Cell* 118, 545-553.

Davey, M. J., and O'Donnell, M. (2003). Replicative helicase loaders: ring breakers and ring makers. *Curr Biol* 13, R594-596.

Davis, T. L., and Meyer, B. J. (1997). SDC-3 coordinates the assembly of a dosage compensation complex on the nematode X chromosome. *Development* 124, 1019-1031.

Dawes, H. E., Berlin, D. S., Lapidus, D. M., Nusbaum, C., Davis, T. L., and Meyer, B. J. (1999). Dosage compensation proteins targeted to X chromosomes by a determinant of hermaphrodite fate. *Science* 284, 1800-1804.

de la Cruz, J., Kressler, D., and Linder, P. (1999). Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *Trends Biochem Sci* 24, 192-198.

de la Cruz, X., Lois, S., Sanchez-Molina, S., and Martinez-Balbas, M. A. (2005). Do protein motifs read the histone code? *Bioessays* 27, 164-175.

Delagoutte, E., and von Hippel, P. H. (2002). Helicase mechanisms and the coupling of helicases within macromolecular machines. Part I: Structures and properties of isolated helicases. *Q Rev Biophys* 35, 431-478.

Delagoutte, E., and von Hippel, P. H. (2003). Helicase mechanisms and the coupling of helicases within macromolecular machines. Part II: Integration of helicases into cellular processes. *Q Rev Biophys* 36, 1-69.

Delattre, M., Spierer, A., Jaquet, Y., and Spierer, P. (2004). Increased expression of *Drosophila* Su(var)3-7 triggers Su(var)3-9-dependent heterochromatin formation. *J Cell Sci* 117, 6239-6247.

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

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

Farkas, G., Leibovitch, B. A., and Elgin, S. C. (2000). Chromatin organization and transcriptional control of gene expression in *Drosophila*. *Gene* 253, 117-136.

Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. *Nature* 421, 448-453.

Fierro-Monti, I., and Mathews, M. B. (2000). Proteins binding to duplexed RNA: one motif, multiple functions. *Trends Biochem Sci* 25, 241-246.

Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Funabiki, H., and Allis, C. D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* 438, 1116-1122.

Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev* 17, 1870-1881.

Flemming, W. (1882). *Zell substanz, Kern und Zelltheilung*.

Franke, A., Dernburg, A., Bashaw, G. J., and Baker, B. S. (1996). Evidence that MSL-mediated dosage compensation in *Drosophila* begins at blastoderm. *Development* 122, 2751-2760.

Freitas, M. A., Sklenar, A. R., and Parthun, M. R. (2004). Application of mass spectrometry to the identification and quantification of histone post-translational modifications. *J Cell Biochem* 92, 691-700.

Friedemann, J., Grosse, F., and Zhang, S. (2005). Nuclear DNA helicase II (RNA helicase A) interacts with Werner syndrome helicase and stimulates its exonuclease activity. *J Biol Chem* 280, 31303-31313.

Fujita, H., Fujii, R., Aratani, S., Amano, T., Fukamizu, A., and Nakajima, T. (2003). Antithetic effects of MBD2a on gene regulation. *Mol Cell Biol* 23, 2645-2657.

Fukunaga, A., Tanaka, A., and Oishi, K. (1975). Maleless, a recessive autosomal mutant of *Drosophila melanogaster* that specifically kills male zygotes. *Genetics* 81, 135-141.

Gendra, E., Moreno, A., Alba, M. M., and Pages, M. (2004). Interaction of the plant glycine-rich RNA-binding protein MA16 with a novel nucleolar DEAD box RNA helicase protein from *Zea mays*. *Plant J* 38, 875-886.

Gergen, J. P., and Wieschaus, E. (1986). Dosage requirements for runt in the segmentation of *Drosophila* embryos. *Cell* 45, 289-299.

Germond, J. E., Bellard, M., Oudet, P., and Chambon, P. (1976). Stability of nucleosomes in native and reconstituted chromatins. *Nucleic Acids Res* 3, 3173-3192.

Ghisolfi, L., Joseph, G., Amalric, F., and Erard, M. (1992). The glycine-rich domain of nucleolin has an unusual supersecondary structure responsible for its RNA-helix-destabilizing properties. *J Biol Chem* 267, 2955-2959.

Gibson, T. J., and Thompson, J. D. (1994). Detection of dsRNA-binding domains in RNA helicase A and *Drosophila* maleless: implications for monomeric RNA helicases. *Nucleic Acids Res* 22, 2552-2556.

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1988). A novel superfamily of nucleoside triphosphate-binding motif containing proteins which

are probably involved in duplex unwinding in DNA and RNA replication and recombination. *FEBS Lett* 235, 16-24.

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1989). Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res* 17, 4713-4730.

Grienenberger, A., Miotto, B., Sagnier, T., Cavalli, G., Schramke, V., Geli, V., Mariol, M. C., Berenger, H., Graba, Y., and Pradel, J. (2002). The MYST domain acetyltransferase Chameau functions in epigenetic mechanisms of transcriptional repression. *Curr Biol* 12, 762-766.

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

Hamada, F. N., Park, P. J., Gordadze, P. R., and Kuroda, M. I. (2005). Global regulation of X chromosomal genes by the MSL complex in *Drosophila melanogaster*. *Genes Dev* 19, 2289-2294.

Hansen, R. S. (2003). X inactivation-specific methylation of LINE-1 elements by DNMT3B: implications for the Lyon repeat hypothesis. *Hum Mol Genet* 12, 2559-2567.

Havas, K., Flaus, A., Phelan, M., Kingston, R., Wade, P. A., Lilley, D. M., and Owen-Hughes, T. (2000). Generation of superhelical torsion by ATP-dependent chromatin remodeling activities. *Cell* 103, 1133-1142.

Hayes, J. J., and Hansen, J. C. (2001). Nucleosomes and the chromatin fiber. *Curr Opin Genet Dev* 11, 124-129.

Heitz, E. (1928). Das heterochromatin der Moose. *Jb Wiss Bot* 69, 728.

Henn, A., Shi, S. P., Zarivach, R., Ben-Zeev, E., and Sagi, I. (2002). The RNA helicase DbpA exhibits a markedly different conformation in the ADP-bound state when compared with the ATP- or RNA-bound states. *J Biol Chem* 277, 46559-46565.

Hickman, A. B., and Dyda, F. (2005). Binding and unwinding: SF3 viral helicases. *Curr Opin Struct Biol* 15, 77-85.

Hilfiker, A., Yang, Y., Hayes, D. H., Beard, C. A., Manning, J. E., and Lucchesi, J. C. (1994). Dosage compensation in *Drosophila*: the X-chromosomal binding of MSL-1 and MLE is dependent on Sxl activity. *Embo J* 13, 3542-3550.

Hodgkin, J. (1980). More sex-determination mutants of *Caenorhabditis elegans*. *Genetics* *96*, 649-664.

Hsu, D. R., and Meyer, B. J. (1994). The *dpy-30* gene encodes an essential component of the *Caenorhabditis elegans* dosage compensation machinery. *Genetics* *137*, 999-1018.

Hung, M. L., Chao, P., and Chang, K. Y. (2003). dsRBM1 and a proline-rich domain of RNA helicase A can form a composite binder to recognize a specific dsDNA. *Nucleic Acids Res* *31*, 5741-5753.

Hwang, W. W., Venkatasubrahmanyam, S., Ianculescu, A. G., Tong, A., Boone, C., and Madhani, H. D. (2003). A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Mol Cell* *11*, 261-266.

Jankowsky, E., Gross, C. H., Shuman, S., and Pyle, A. M. (2001). Active disruption of an RNA-protein interaction by a DExH/D RNA helicase. *Science* *291*, 121-125.

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. *Science* *293*, 1074-1080.

Jeppesen, P., and Turner, B. M. (1993). The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* *74*, 281-289.

Johnstone, O., Deuring, R., Bock, R., Linder, P., Fuller, M. T., and Lasko, P. (2005). Belle is a *Drosophila* DEAD-box protein required for viability and in the germ line. *Dev Biol* *277*, 92-101.

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

Kelley, R. L., and Kuroda, M. I. (2000). The role of chromosomal RNAs in marking the X for dosage compensation. *Curr Opin Genet Dev* *10*, 555-561.

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.

Kenan, D. J., Query, C. C., and Keene, J. D. (1991). RNA recognition: towards identifying determinants of specificity. *Trends Biochem Sci* *16*, 214-220.

Kernan, M. J., Kuroda, M. I., Kreber, R., Baker, B. S., and Ganetzky, B. (1991). *naps*, a mutation affecting sodium channel activity in *Drosophila*, is an allele of *mle*, a regulator of X chromosome transcription. *Cell* 66, 949-959.

Khalil, A. M., Boyar, F. Z., and Driscoll, D. J. (2004). Dynamic histone modifications mark sex chromosome inactivation and reactivation during mammalian spermatogenesis. *Proc Natl Acad Sci U S A* 101, 16583-16587.

Kiledjian, M., and Dreyfuss, G. (1992). Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *Embo J* 11, 2655-2664.

Kim, D. H., and Rossi, J. J. (1999). The first ATPase domain of the yeast 246-kDa protein is required for in vivo unwinding of the U4/U6 duplex. *Rna* 5, 959-971.

Kim, J. L., Morgenstern, K. A., Griffith, J. P., Dwyer, M. D., Thomson, J. A., Murcko, M. A., Lin, C., and Caron, P. R. (1998). Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure* 6, 89-100.

Korolev, S., Hsieh, J., Gauss, G. H., Lohman, T. M., and Waksman, G. (1997). Major domain swiveling revealed by the crystal structures of complexes of *E. coli* Rep helicase bound to single-stranded DNA and ADP. *Cell* 90, 635-647.

Kuang, W. F., Lin, Y. C., Jean, F., Huang, Y. W., Tai, C. L., Chen, D. S., Chen, P. J., and Hwang, L. H. (2004). Hepatitis C virus NS3 RNA helicase activity is modulated by the two domains of NS3 and NS4A. *Biochem Biophys Res Commun* 317, 211-217.

Kurdistani, S. K., and Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. *Nat Rev Mol Cell Biol* 4, 276-284.

Kurdistani, S. K., Tavazoie, S., and Grunstein, M. (2004). Mapping global histone acetylation patterns to gene expression. *Cell* 117, 721-733.

Kzhyshkowska, J., Schutt, H., Liss, M., Kremmer, E., Stauber, R., Wolf, H., and Dobner, T. (2001). Heterogeneous nuclear ribonucleoprotein E1B-AP5 is methylated in its Arg-Gly-Gly (RGG) box and interacts with human arginine methyltransferase HRMT1L1. *Biochem J* 358, 305-314.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116-120.

Lanford, R. E., Kanda, P., and Kennedy, R. C. (1986). Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. *Cell* 46, 575-582.

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

Lee, C. G., and Hurwitz, J. (1992). A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3' to 5' direction. *J Biol Chem* 267, 4398-4407.

Lee, C. G., and Hurwitz, J. (1993). Human RNA helicase A is homologous to the maleless protein of *Drosophila*. *J Biol Chem* 268, 16822-16830.

Lee, C. G., Reichman, T. W., Baik, T., and Mathews, M. B. (2004). MLE functions as a transcriptional regulator of the roX2 gene. *J Biol Chem* 279, 47740-47745.

Lee, J. T., Davidow, L. S., and Warshawsky, D. (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet* 21, 400-404.

Lee, M. B., Lebedeva, L. A., Suzawa, M., Wadekar, S. A., Desclozeaux, M., and Ingraham, H. A. (2005). The DEAD-box protein DP103 (Ddx20 or Gemin-3) represses orphan nuclear receptor activity via SUMO modification. *Mol Cell Biol* 25, 1879-1890.

Legube, G., and Trouche, D. (2003). Regulating histone acetyltransferases and deacetylases. *EMBO Rep* 4, 944-947.

Li, F., Parry, D. A., and Scott, M. J. (2005). The Amino-Terminal Region of *Drosophila* MSL1 Contains Basic, Glycine-Rich, and Leucine Zipper-Like Motifs That Promote X Chromosome Binding, Self-Association, and MSL2 Binding, Respectively. *Mol Cell Biol* 25, 8913-8924.

Liao, W., Wang, S., Han, C., and Zhang, Y. (2005). 14-3-3 proteins regulate glycogen synthase 3beta phosphorylation and inhibit cardiomyocyte hypertrophy. *Febs J* 272, 1845-1854.

Liu, Q., and Dreyfuss, G. (1995). In vivo and in vitro arginine methylation of RNA-binding proteins. *Mol Cell Biol* 15, 2800-2808.

Lohman, T. M., and Bjornson, K. P. (1996). Mechanisms of helicase-catalyzed DNA unwinding. *Annu Rev Biochem* 65, 169-214.

Lorsch, J. R. (2002). RNA chaperones exist and DEAD box proteins get a life. *Cell* 109, 797-800.

Lucchesi, J. C., Kelly, W. G., and Panning, B. (2005). Chromatin Remodeling in Dosage Compensation. *Annu Rev Genet*.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.

Lusser, A., and Kadonaga, J. T. (2003). Chromatin remodeling by ATP-dependent molecular machines. *Bioessays* 25, 1192-1200.

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

Lyon, M. (1999). X chromosome inactivation. *Curr Biol* 9, 235-237.

Lyon, M. F. (1996). X-chromosome inactivation. Pinpointing the centre. *Nature* 379, 116-117.

Macdonald, N., Welburn, J. P., Noble, M. E., Nguyen, A., Yaffe, M. B., Clynes, D., Moggs, J. G., Orphanides, G., Thomson, S., Edmunds, J. W., *et al.* (2005). Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3. *Mol Cell* 20, 199-211.

Margueron, R., Trojer, P., and Reinberg, D. (2005). The key to development: interpreting the histone code? *Curr Opin Genet Dev* 15, 163-176.

Marintcheva, B., and Weller, S. K. (2003). Helicase motif Ia is involved in single-strand DNA-binding and helicase activities of the herpes simplex virus type 1 origin-binding protein, UL9. *J Virol* 77, 2477-2488.

Marmorstein, R., and Roth, S. Y. (2001). Histone acetyltransferases: function, structure, and catalysis. *Curr Opin Genet Dev* 11, 155-161.

Maurer-Stroh, S., Dickens, N. J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., and Ponting, C. P. (2003). The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem Sci* 28, 69-74.

McBride, A. E., Cook, J. T., Stemmler, E. A., Rutledge, K. L., McGrath, K. A., and Rubens, J. A. (2005). Arginine methylation of yeast mRNA-binding protein Npl3 directly affects its function, nuclear export, and intranuclear protein interactions. *J Biol Chem* 280, 30888-30898.

Mears, W. E., and Rice, S. A. (1996). The RGG box motif of the herpes simplex virus ICP27 protein mediates an RNA-binding activity and determines in vivo methylation. *J Virol* 70, 7445-7453.

Meller, V. H. (2003). Initiation of dosage compensation in *Drosophila* embryos depends on expression of the roX RNAs. *Mech Dev* 120, 759-767.

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

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

Mellor, J. (2005). The dynamics of chromatin remodeling at promoters. *Mol Cell Biol* 19, 147-157.

Meyer, B. J., McDonel, P., Csankovszki, G., and Ralston, E. (2004). Sex and X-chromosome-wide repression in *Caenorhabditis elegans*. *Cold Spring Harb Symp Quant Biol* 69, 71-79.

Min, J., Zhang, Y., and Xu, R. M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev* 17, 1823-1828.

Mlynarczyk, S. K., and Panning, B. (2000). X inactivation: Tsix and Xist as yin and yang. *Curr Biol* 10, R899-903.

Morales, V., Regnard, C., Izzo, A., Vetter, I., and Becker, P. B. (2005). The MRG domain mediates the functional integration of MSL3 into the dosage compensation complex. *Mol Cell Biol* 25, 5947-5954.

Morales, V., Straub, T., Neumann, M. F., Mengus, G., Akhtar, A., and Becker, P. B. (2004). Functional integration of the histone acetyltransferase MOF into the dosage compensation complex. *Embo J* 23, 2258-2268.

Murnion, M. E., Adams, R. R., Callister, D. M., Allis, C. D., Earnshaw, W. C., and Swedlow, J. R. (2001). Chromatin-associated protein phosphatase 1 regulates aurora-B and histone H3 phosphorylation. *J Biol Chem* 276, 26656-26665.

Myohanen, S., and Baylin, S. B. (2001). Sequence-specific DNA binding activity of RNA helicase A to the p16INK4a promoter. *J Biol Chem* 276, 1634-1642.

Nanduri, S., Carpick, B. W., Yang, Y., Williams, B. R., and Qin, J. (1998). Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. *Embo J* 17, 5458-5465.

Ng, H. H., and Bird, A. (2000). Histone deacetylases: silencers for hire. *Trends Biochem Sci* 25, 121-126.

Ng, H. H., Xu, R. M., Zhang, Y., and Struhl, K. (2002). Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. *J Biol Chem* 277, 34655-34657.

Nichols, R. C., Wang, X. W., Tang, J., Hamilton, B. J., High, F. A., Herschman, H. R., and Rigby, W. F. (2000). The RGG domain in hnRNP A2 affects subcellular localization. *Exp Cell Res* 256, 522-532.

Nielsen, P. R., Nietlispach, D., Buscaino, A., Warner, R. J., Akhtar, A., Murzin, A. G., Murzina, N. V., and Laue, E. D. (2005). Structure of the chromo barrel domain from the MOF acetyltransferase. *J Biol Chem* 280, 32326-32331.

Nielsen, P. R., Nietlispach, D., Mott, H. R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A. G., Murzina, N. V., and Laue, E. D. (2002). Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* 416, 103-107.

Nonet, M. L., and Meyer, B. J. (1991). Early aspects of *Caenorhabditis elegans* sex determination and dosage compensation are regulated by a zinc-finger protein. *Nature* 351, 65-68.

Nowak, S. J., and Corces, V. G. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet* 20, 214-220.

Nusbaum, C., and Meyer, B. J. (1989). The *Caenorhabditis elegans* gene *sdc-2* controls sex determination and dosage compensation in XX animals. *Genetics* 122, 579-593.

Oakley, A. J., Loscha, K. V., Schaeffer, P. M., Liepinsh, E., Pintacuda, G., Wilce, M. C., Otting, G., and Dixon, N. E. (2005). Crystal and solution structures of the helicase-binding domain of *Escherichia coli* primase. *J Biol Chem* 280, 11495-11504.

Oh, H., Bai, X., Park, Y., Bone, J. R., and Kuroda, M. I. (2004). Targeting dosage compensation to the X chromosome of *Drosophila* males. *Cold Spring Harb Symp Quant Biol* 69, 81-88.

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

Parekh, B. S., and Maniatis, T. (1999). Virus infection leads to localized hyperacetylation of histones H3 and H4 at the IFN-beta promoter. *Mol Cell* 3, 125-129.

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

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

Pause, A., Methot, N., and Sonenberg, N. (1993). The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis. *Mol Cell Biol* *13*, 6789-6798.

Pause, A., and Sonenberg, N. (1992). Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *Embo J* *11*, 2643-2654.

Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., de la Cruz, C. C., Otte, A. P., Panning, B., and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. *Science* *300*, 131-135.

Polach, K. J., and Uhlenbeck, O. C. (2002). Cooperative binding of ATP and RNA substrates to the DEAD/H protein DbpA. *Biochemistry* *41*, 3693-3702.

Rastelli, L., Richman, R., and Kuroda, M. I. (1995). The dosage compensation regulators MLE, MSL-1 and MSL-2 are interdependent since early embryogenesis in *Drosophila*. *Mech Dev* *53*, 223-233.

Rattner, B. P., and Meller, V. H. (2004). *Drosophila* male-specific lethal 2 protein controls sex-specific expression of the roX genes. *Genetics* *166*, 1825-1832.

Richmond, T. J., and Davey, C. A. (2003). The structure of DNA in the nucleosome core. *Nature* *423*, 145-150.

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.

Rocak, S., Emery, B., Tanner, N. K., and Linder, P. (2005). Characterization of the ATPase and unwinding activities of the yeast DEAD-box protein Has1p and the analysis of the roles of the conserved motifs. *Nucleic Acids Res* *33*, 999-1009.

Rogers, G. W., Jr., Komar, A. A., and Merrick, W. C. (2002). eIF4A: the godfather of the DEAD box helicases. *Prog Nucleic Acid Res Mol Biol* *72*, 307-331.

Rogers, G. W., Jr., Richter, N. J., Lima, W. F., and Merrick, W. C. (2001). Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J Biol Chem* *276*, 30914-30922.

Romano, P. R., Green, S. R., Barber, G. N., Mathews, M. B., and Hinnebusch, A. G. (1995). Structural requirements for double-stranded RNA binding, dimerization, and activation of the human eIF-2 alpha kinase DAI in *Saccharomyces cerevisiae*. *Mol Cell Biol* *15*, 365-378.

Rougeulle, C., Chaumeil, J., Sarma, K., Allis, C. D., Reinberg, D., Avner, P., and Heard, E. (2004). Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol Cell Biol* *24*, 5475-5484.

Sass, G. L., Pannuti, A., and Lucchesi, J. C. (2003). Male-specific lethal complex of *Drosophila* targets activated regions of the X chromosome for chromatin remodeling. *Proc Natl Acad Sci U S A* *100*, 8287-8291.

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

Seeman, N. C., Rosenberg, J. M., and Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. *Proc Natl Acad Sci U S A* *73*, 804-808.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* *119*, 941-953.

Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T. B., Webster, Z., Peters, A. H., Jenuwein, T., Otte, A. P., and Brockdorff, N. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev Cell* *4*, 481-495.

Silverman, E., Edwalds-Gilbert, G., and Lin, R. J. (2003). DExD/H-box proteins and their partners: helping RNA helicases unwind. *Gene* *312*, 1-16.

Simmler, M. C., Cattanach, B. M., Rasberry, C., Rougeulle, C., and Avner, P. (1993). Mapping the murine Xce locus with (CA)<sub>n</sub> repeats. *Mamm Genome* *4*, 523-530.

Siomi, H., and Dreyfuss, G. (1995). A nuclear localization domain in the hnRNP A1 protein. *J Cell Biol* *129*, 551-560.

Smith, W. A., Schurter, B. T., Wong-Staal, F., and David, M. (2004). Arginine methylation of RNA helicase a determines its subcellular localization. *J Biol Chem* *279*, 22795-22798.

Soultanas, P., and Wigley, D. B. (2001). Unwinding the 'Gordian knot' of helicase action. *Trends Biochem Sci* *26*, 47-54.

St Johnston, D., Brown, N. H., Gall, J. G., and Jantsch, M. (1992). A conserved double-stranded RNA-binding domain. *Proc Natl Acad Sci U S A* *89*, 10979-10983.

Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* *403*, 41-45.

Strahl, B. D., Briggs, S. D., Brame, C. J., Caldwell, J. A., Koh, S. S., Ma, H., Cook, R. G., Shabanowitz, J., Hunt, D. F., Stallcup, M. R., and Allis, C. D. (2001).

Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr Biol* *11*, 996-1000.

Straub, T., Dahlsveen, I. K., and Becker, P. B. (2005a). Dosage compensation in flies: mechanism, models, mystery. *FEBS Lett* *579*, 3258-3263.

Straub, T., Gilfillan, G. D., Maier, V. K., and Becker, P. B. (2005b). The *Drosophila* MSL complex activates the transcription of target genes. *Genes Dev* *19*, 2284-2288.

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.

Subramanya, H. S., Bird, L. E., Brannigan, J. A., and Wigley, D. B. (1996). Crystal structure of a DExx box DNA helicase. *Nature* *384*, 379-383.

Sun, Z. W., and Allis, C. D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* *418*, 104-108.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* *116*, 51-61.

Talavera, M. A., and De La Cruz, E. M. (2005). Equilibrium and kinetic analysis of nucleotide binding to the DEAD-box RNA helicase DbpA. *Biochemistry* *44*, 959-970.

Talavera, M. A., Matthews, E. E., Eliason, W. K., Sagi, I., Wang, J., Henn, A., and De La Cruz, E. M. (2006). Hydrodynamic characterization of the DEAD-box RNA Helicase DbpA. *J Mol Biol* *355*, 697-707.

Thatcher, T. H., and Gorovsky, M. A. (1994). Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic Acids Res* *22*, 174-179.

Thoma, F., and Koller, T. (1977). Influence of histone H1 on chromatin structure. *Cell* *12*, 101-107.

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. (1994). The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *Embo J* *13*, 3822-3831.

Tsu, C. A., Kossen, K., and Uhlenbeck, O. C. (2001). The *Escherichia coli* DEAD protein DbpA recognizes a small RNA hairpin in 23S rRNA. *Rna* *7*, 702-709.

Tsu, C. A., and Uhlenbeck, O. C. (1998). Kinetic analysis of the RNA-dependent adenosinetriphosphatase activity of DbpA, an *Escherichia coli* DEAD protein specific for 23S ribosomal RNA. *Biochemistry* *37*, 16989-16996.

- Turner, B. M. (1993). Decoding the nucleosome. *Cell* 75, 5-8.
- Tuteja, N., and Tuteja, R. (2004). Unraveling DNA helicases. Motif, structure, mechanism and function. *Eur J Biochem* 271, 1849-1863.
- Utley, R. T., and Cote, J. (2003). The MYST family of histone acetyltransferases. *Curr Top Microbiol Immunol* 274, 203-236.
- Valente, L., and Nishikura, K. (2005). ADAR gene family and A-to-I RNA editing: diverse roles in posttranscriptional gene regulation. *Prog Nucleic Acid Res Mol Biol* 79, 299-338.
- Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S., and Wigley, D. B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* 97, 75-84.
- Walmacq, C., Rahmouni, A. R., and Boudvillain, M. (2004). Influence of substrate composition on the helicase activity of transcription termination factor Rho: reduced processivity of Rho hexamers during unwinding of RNA-DNA hybrid regions. *J Mol Biol* 342, 403-420.
- Wang, Y., and Guthrie, C. (1998). PRP16, a DEAH-box RNA helicase, is recruited to the spliceosome primarily via its nonconserved N-terminal domain. *Rna* 4, 1216-1229.
- Wang, Y., Wysocka, J., Sayegh, J., Lee, Y. H., Perlin, J. R., Leonelli, L., Sonbuchner, L. S., McDonald, C. H., Cook, R. G., Dou, Y., *et al.* (2004). Human PAD4 regulates histone arginine methylation levels via demethylination. *Science* 306, 279-283.
- Wang, Y., Zhang, W., Jin, Y., Johansen, J., and Johansen, K. M. (2001). The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in *Drosophila*. *Cell* 105, 433-443.
- Wei, Y., Mizzen, C. A., Cook, R. G., Gorovsky, M. A., and Allis, C. D. (1998). Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in *Tetrahymena*. *Proc Natl Acad Sci U S A* 95, 7480-7484.
- Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H., and Allis, C. D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121, 859-872.

Xu, C., and Henry, M. F. (2004). Nuclear export of hnRNP Hrp1p and nuclear export of hnRNP Npl3p are linked and influenced by the methylation state of Npl3p. *Mol Cell Biol* 24, 10742-10756.

Yang, X. J. (2004). Lysine acetylation and the bromodomain: a new partnership for signaling. *Bioessays* 26, 1076-1087.

Yarranton, G. T., Das, R. H., and Gefter, M. L. (1979). Enzyme-catalyzed DNA unwinding. Mechanism of action of helicase III. *J Biol Chem* 254, 12002-12006.

Yonker, S. A., and Meyer, B. J. (2003). Recruitment of *C. elegans* dosage compensation proteins for gene-specific versus chromosome-wide repression. *Development* 130, 6519-6532.

Zegerman, P., Canas, B., Pappin, D., and Kouzarides, T. (2002). Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. *J Biol Chem* 277, 11621-11624.

Zhang, F., Romano, P. R., Nagamura-Inoue, T., Tian, B., Dever, T. E., Mathews, M. B., Ozato, K., and Hinnebusch, A. G. (2001). Binding of double-stranded RNA to protein kinase PKR is required for dimerization and promotes critical autophosphorylation events in the activation loop. *J Biol Chem* 276, 24946-24958.

Zhang, S., and Grosse, F. (1997). Domain structure of human nuclear DNA helicase II (RNA helicase A). *J Biol Chem* 272, 11487-11494.

Zhang, S., and Grosse, F. (2004). Multiple functions of nuclear DNA helicase II (RNA helicase A) in nucleic acid metabolism. *Acta Biochim Biophys Sin (Shanghai)* 36, 177-183.

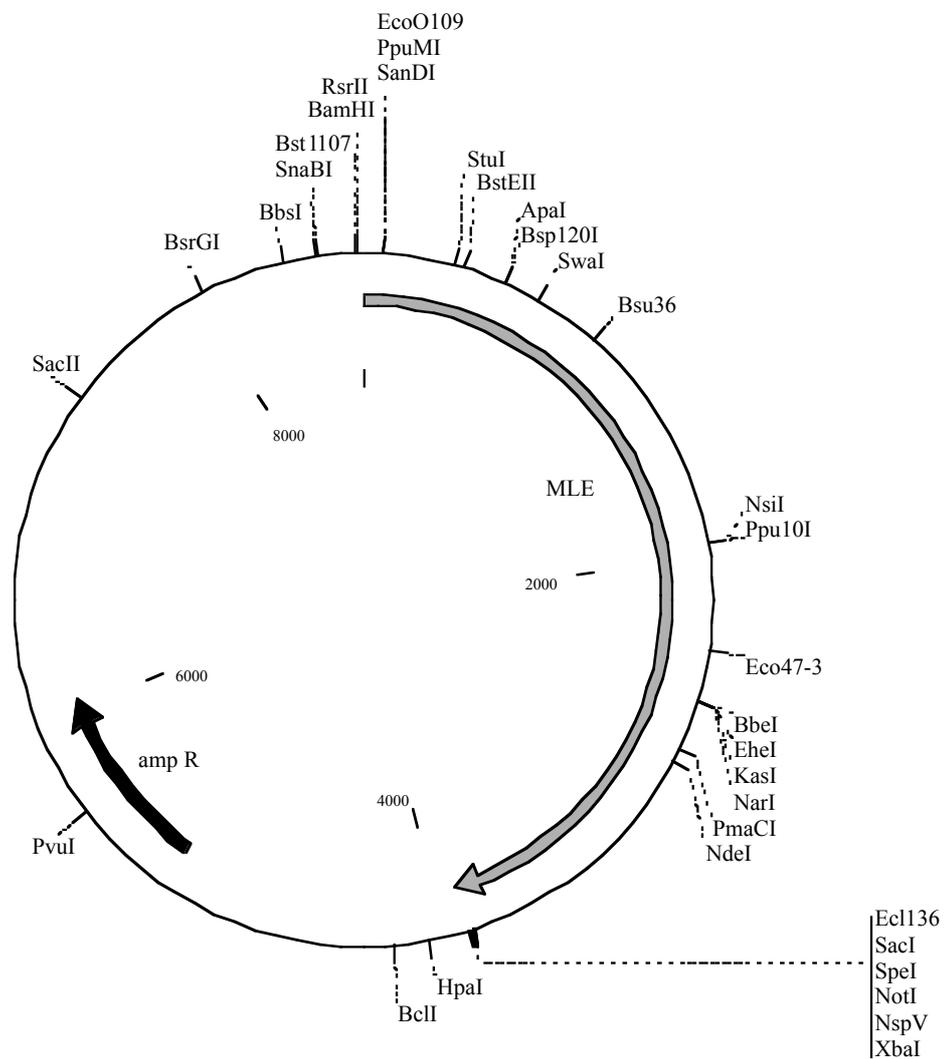
Zhao, R., Shen, J., Green, M. R., MacMorris, M., and Blumenthal, T. (2004). Crystal structure of UAP56, a DExD/H-box protein involved in pre-mRNA splicing and mRNA export. *Structure* 12, 1373-1381.

Zhimulev, I. F., Belyaeva, E. S., Semeshin, V. F., Koryakov, D. E., Demakov, S. A., Demakova, O. V., Pokholkova, G. V., and Andreyeva, E. N. (2004). Polytene chromosomes: 70 years of genetic research. *Int Rev Cytol* 241, 203-275.

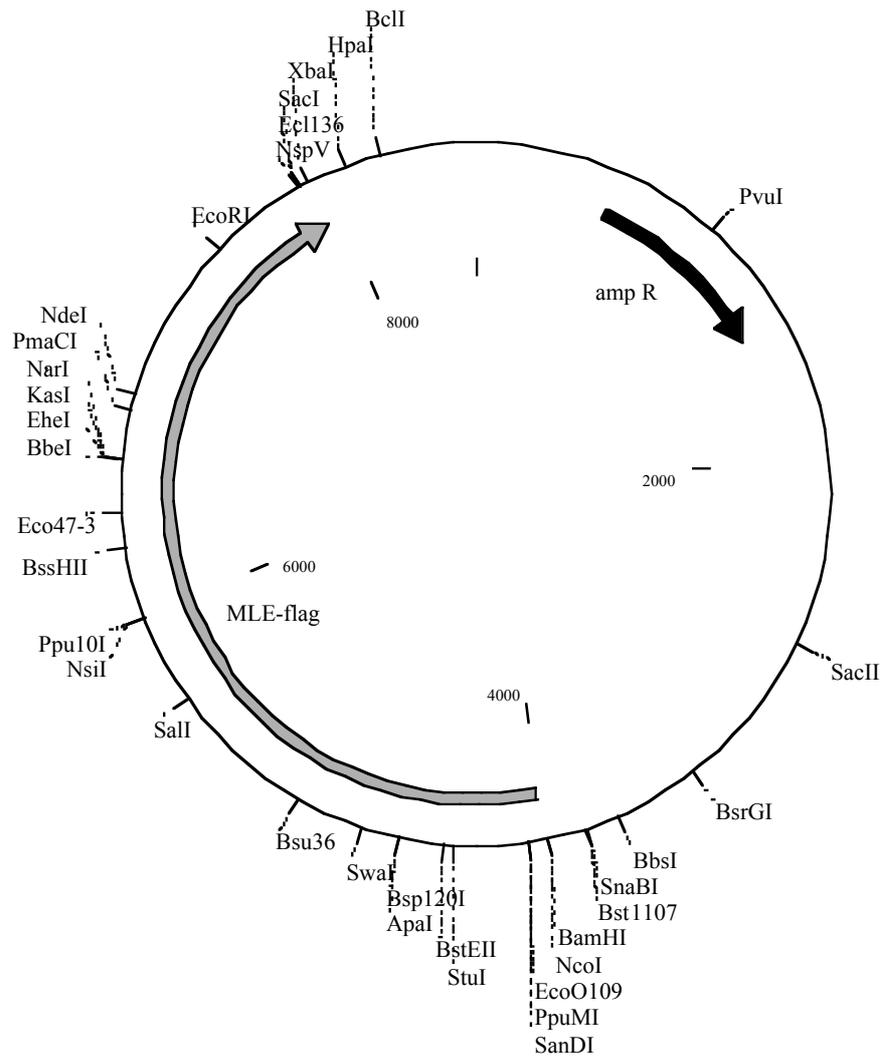
## VI. Appendix

### 1. Maps of plasmids

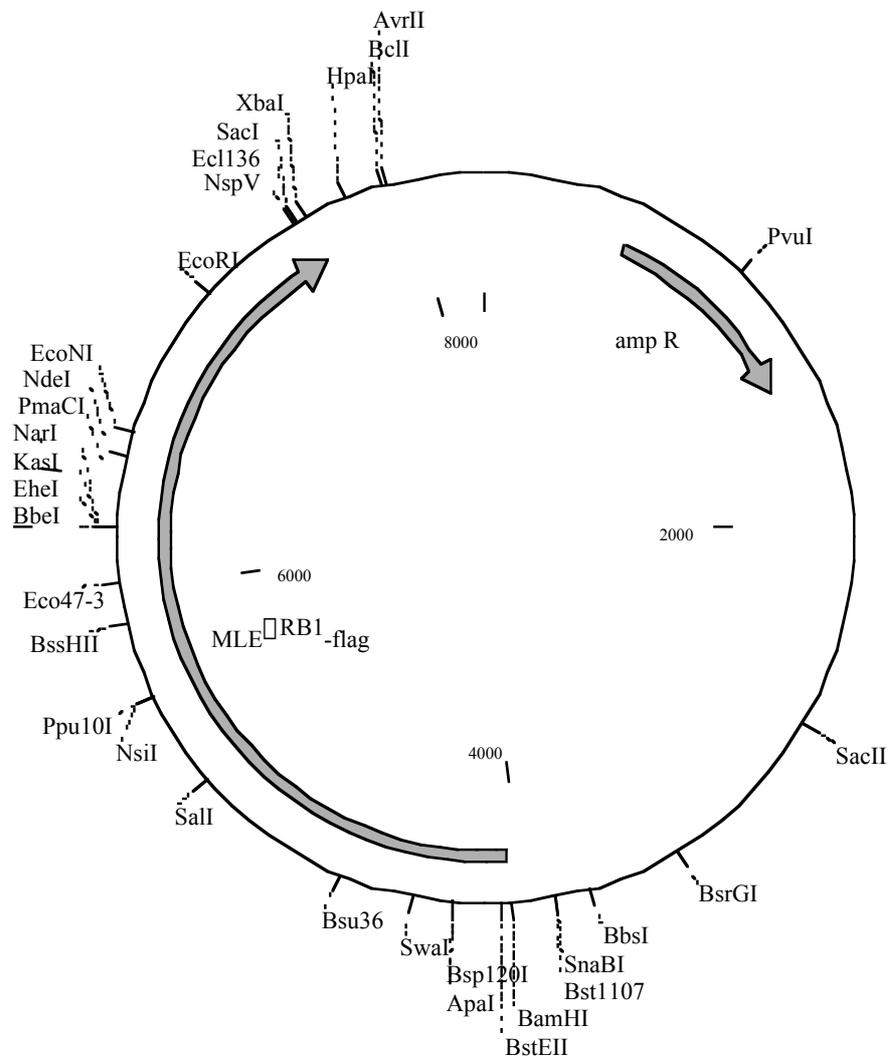
#### PfastBac-MLE



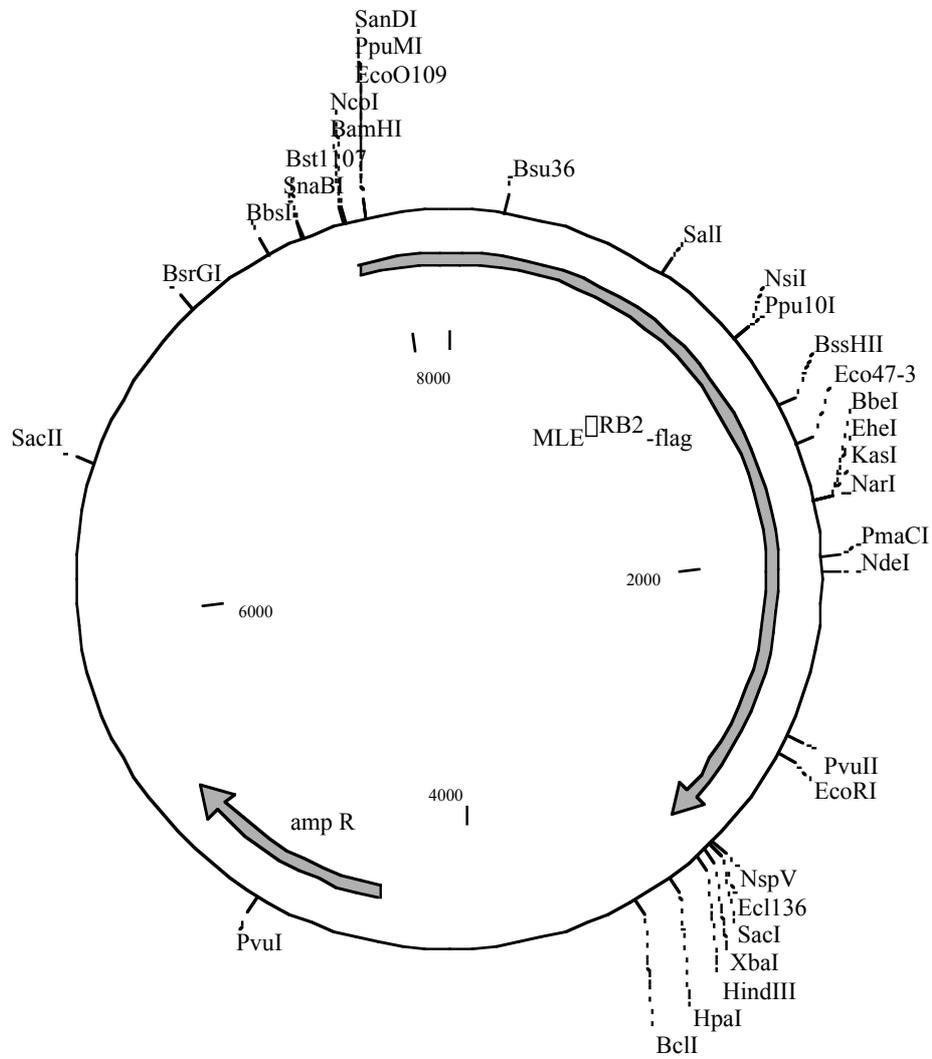
### pFastBac-MLE-flag



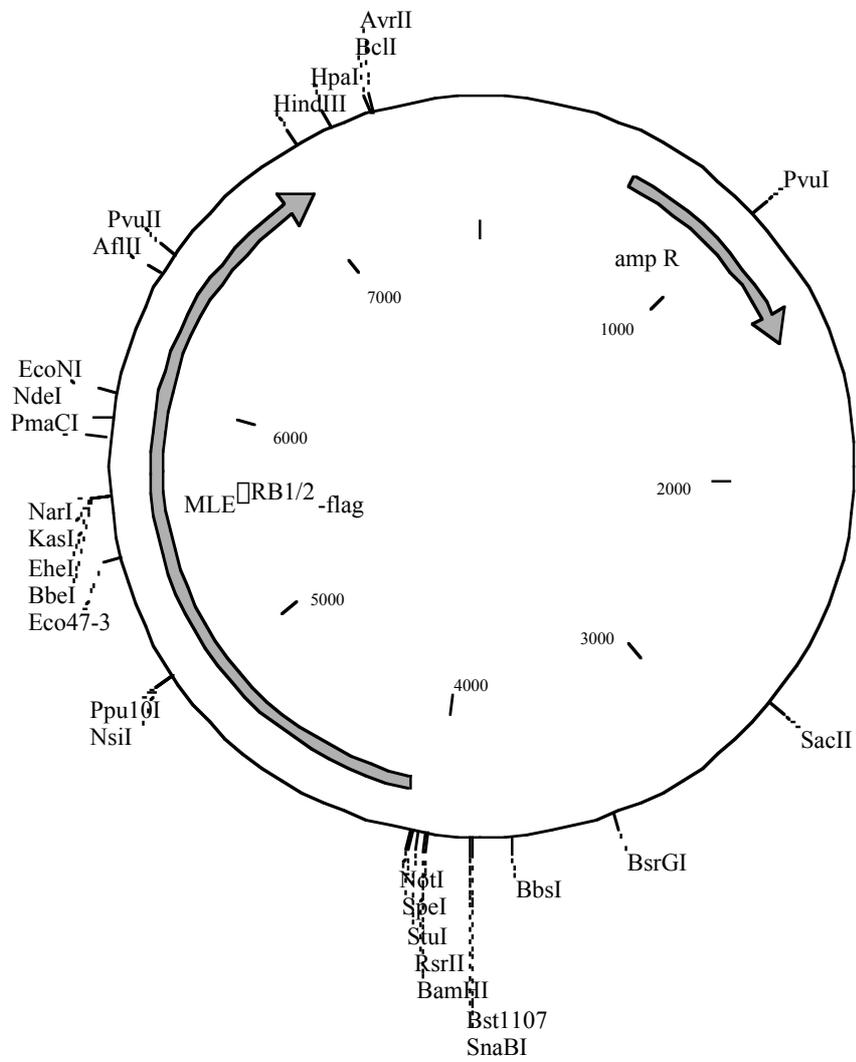
**pFastBac- $MLE^{\Delta RB1}$ -flag**



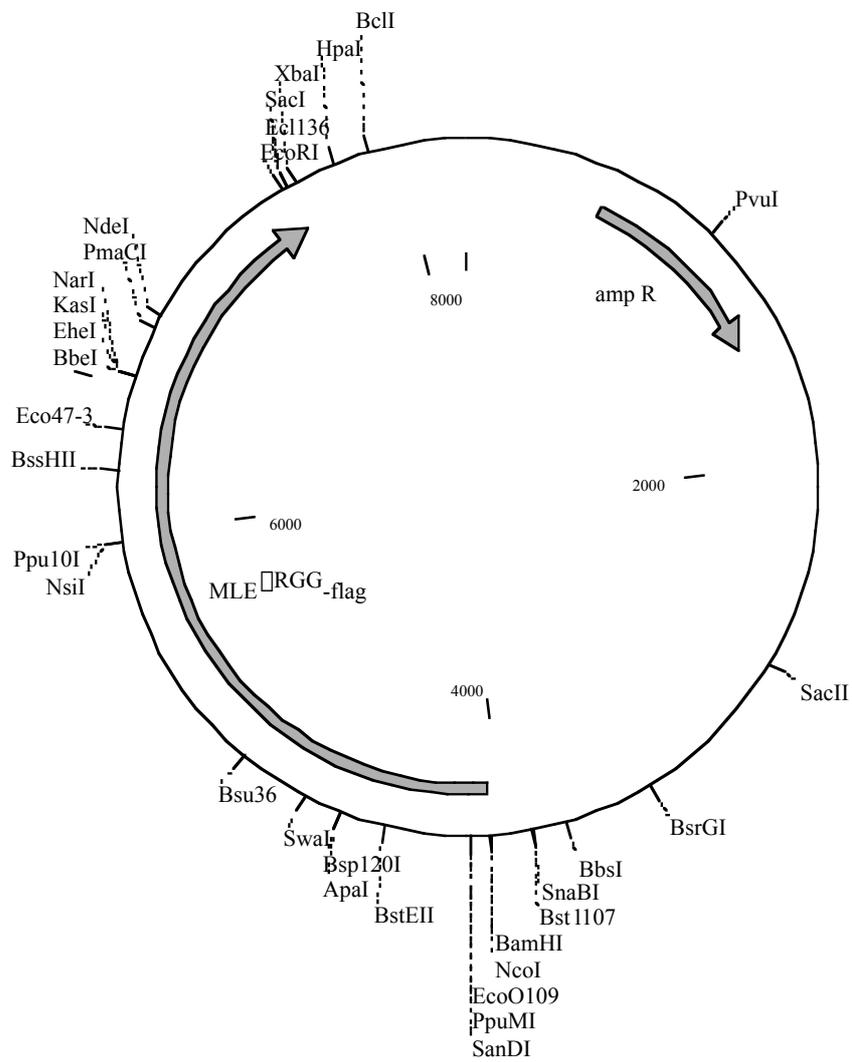
**pFastBac- $MLE^{\Delta RB2}$ -flag**



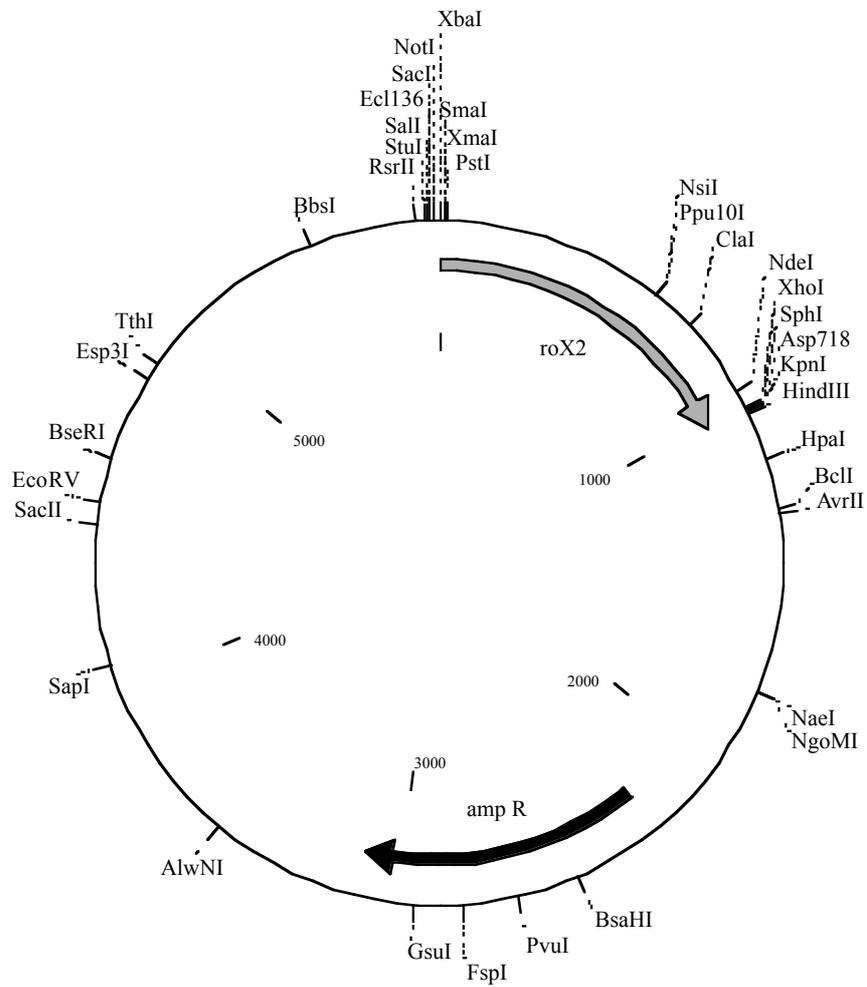
**pFastBac- $MLE^{\Delta RB12}$ -flag**



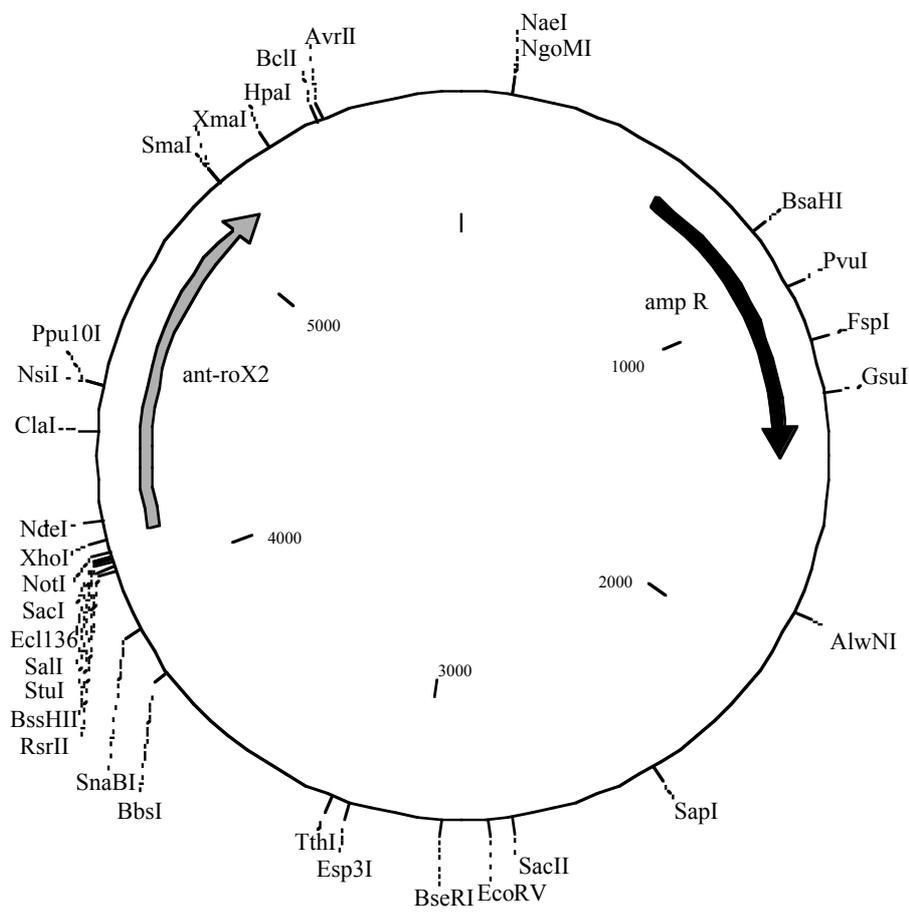
**pFastBac-MLE<sup>ΔRGG</sup>-flag**



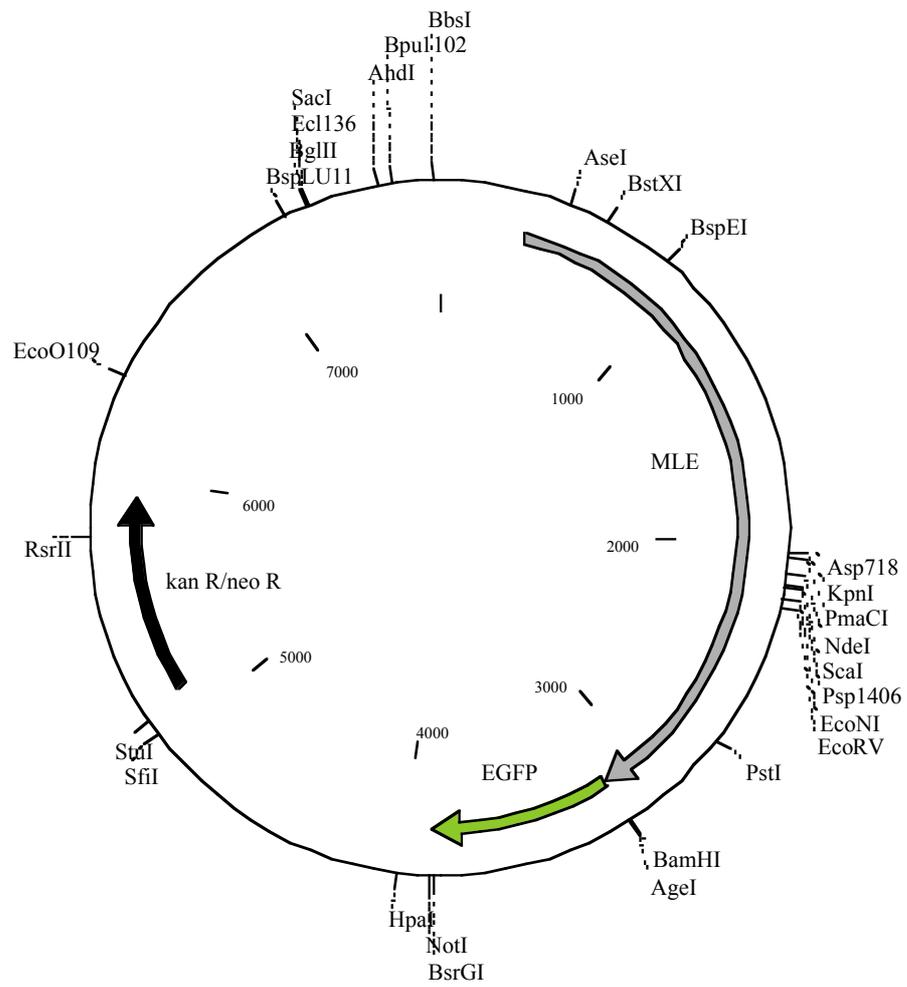
**pFastBac-roX2**



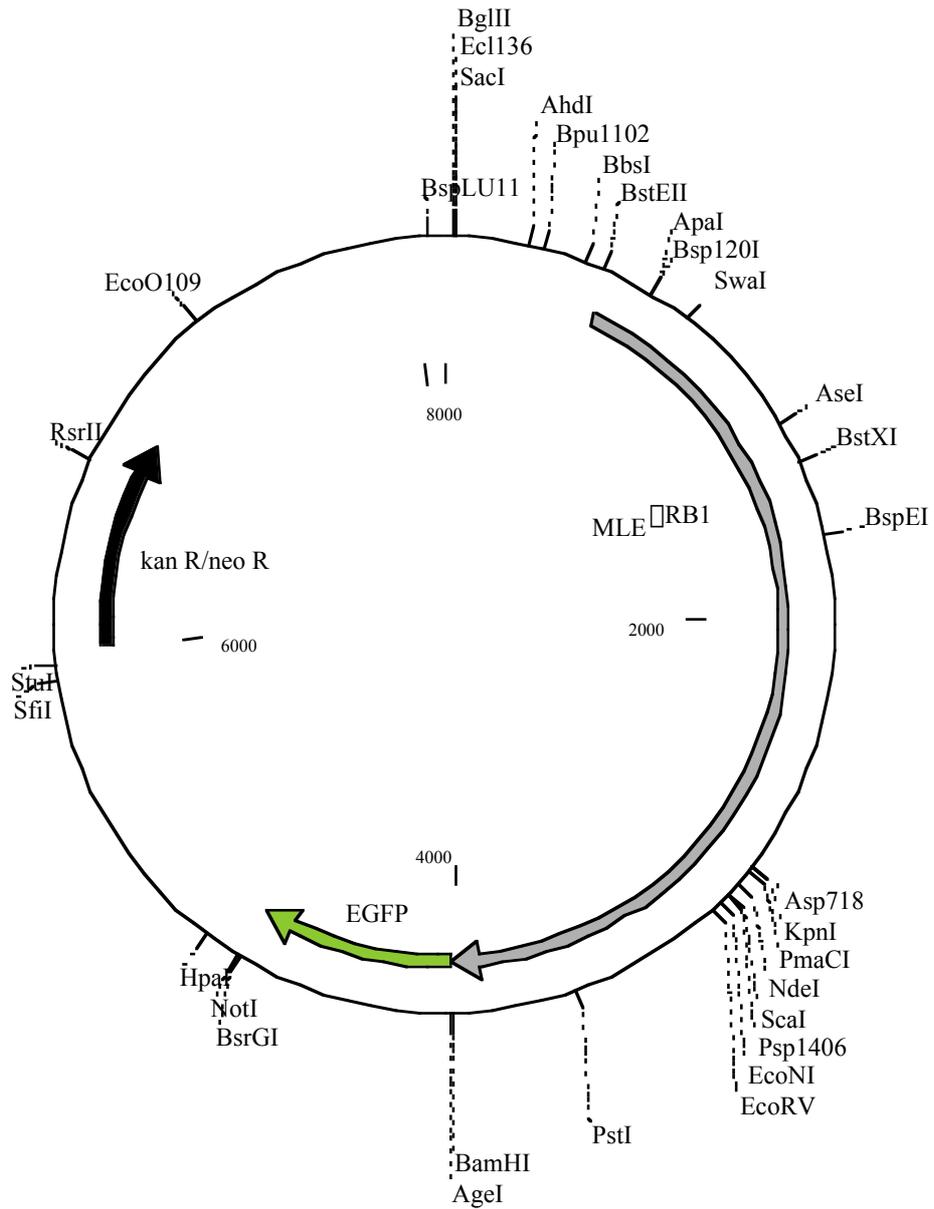
### pFastBac-ant-roX2



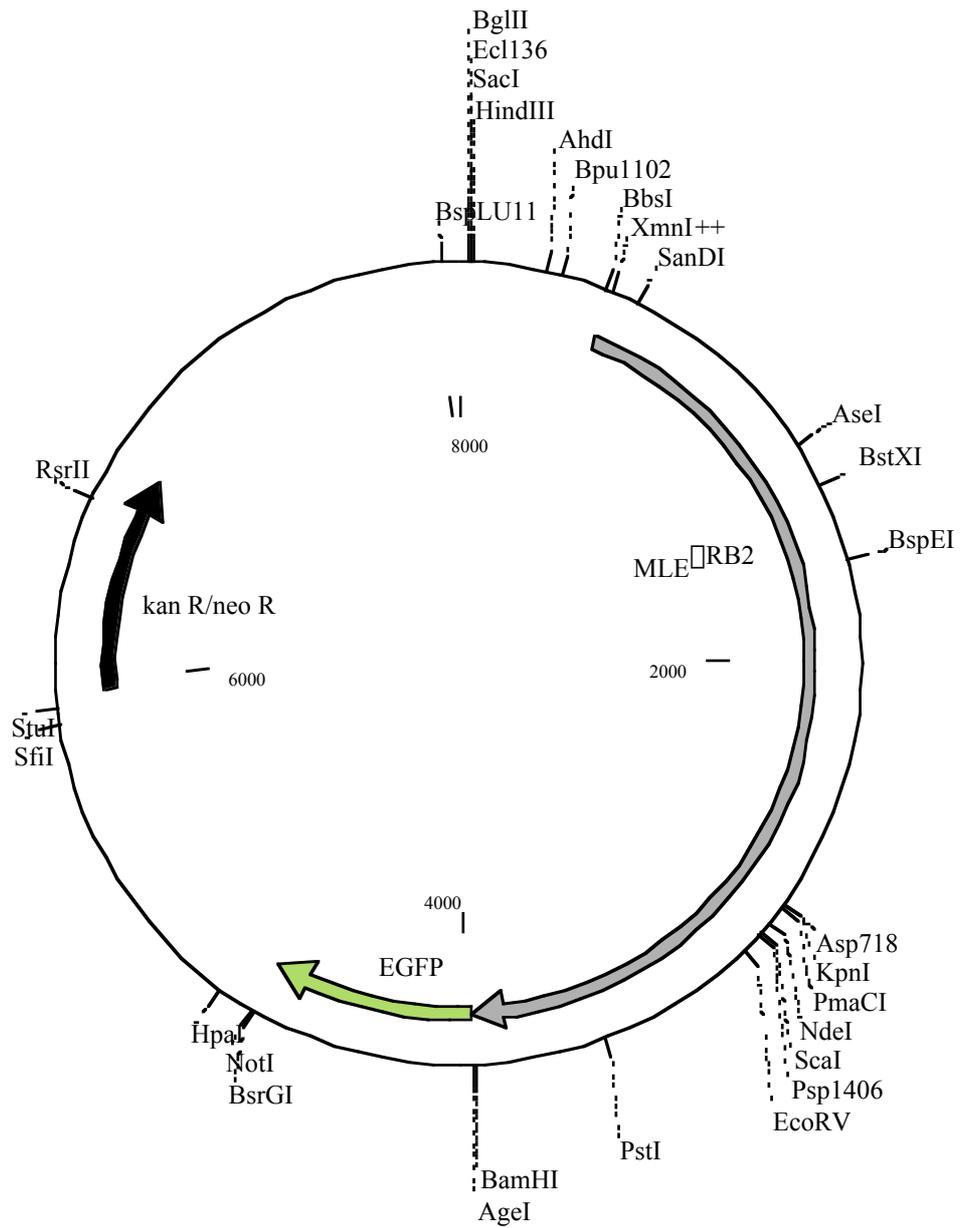
**pHSP70-MLE-EGFP**



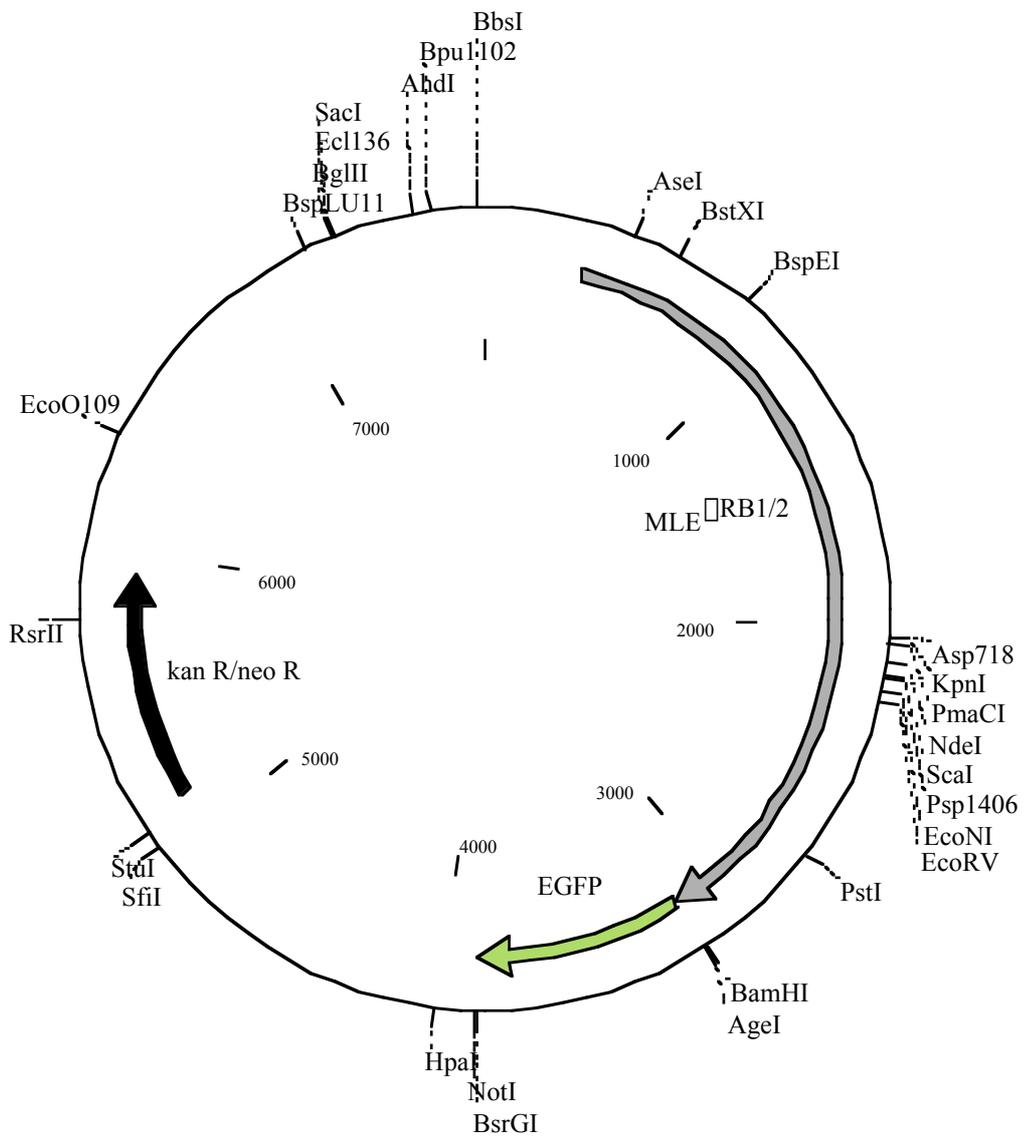
pHSP70-MLE<sup>ARB1</sup>-EGFP



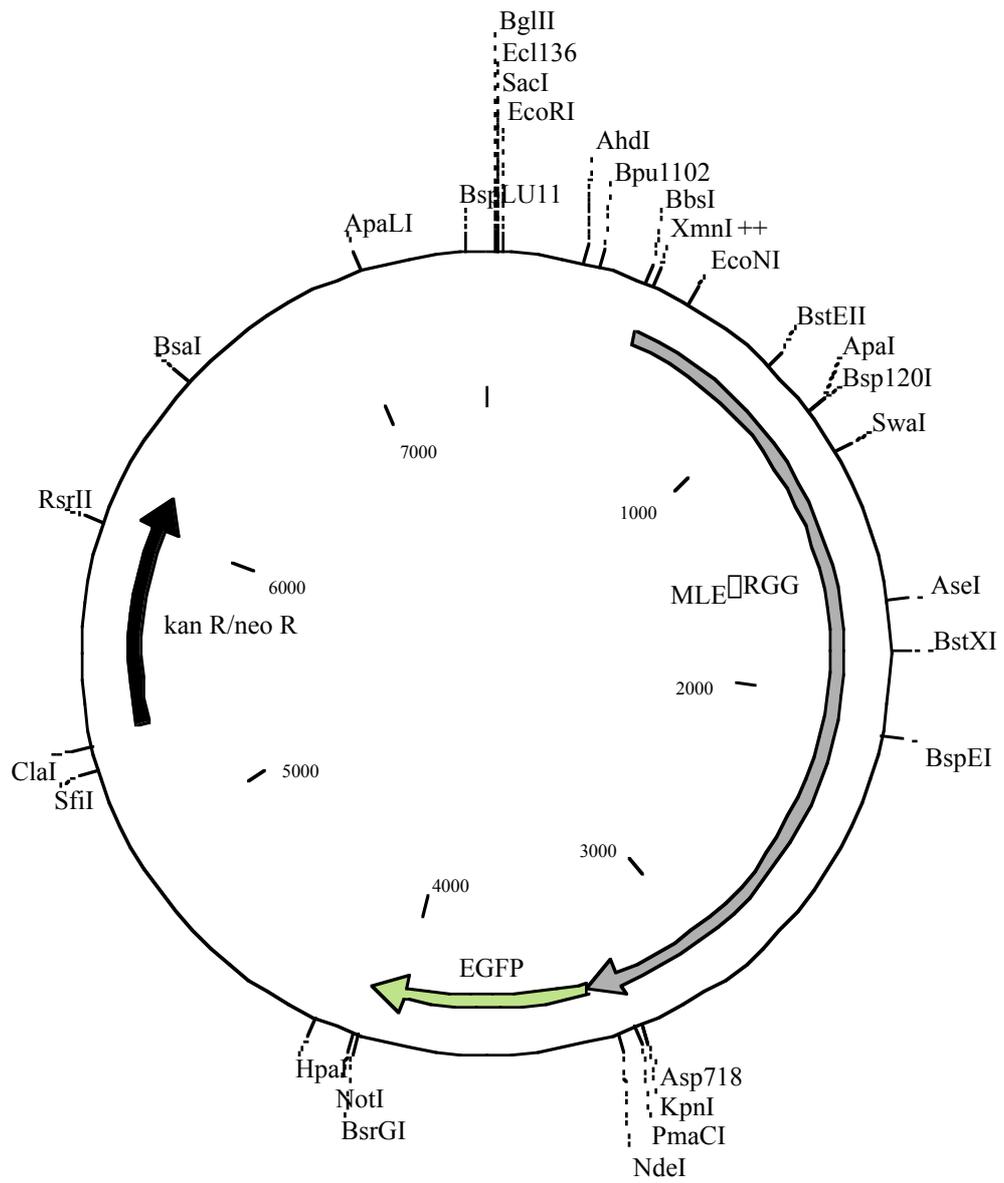
**pHSP70-MLE<sup>ΔRB2</sup>-EGFP**



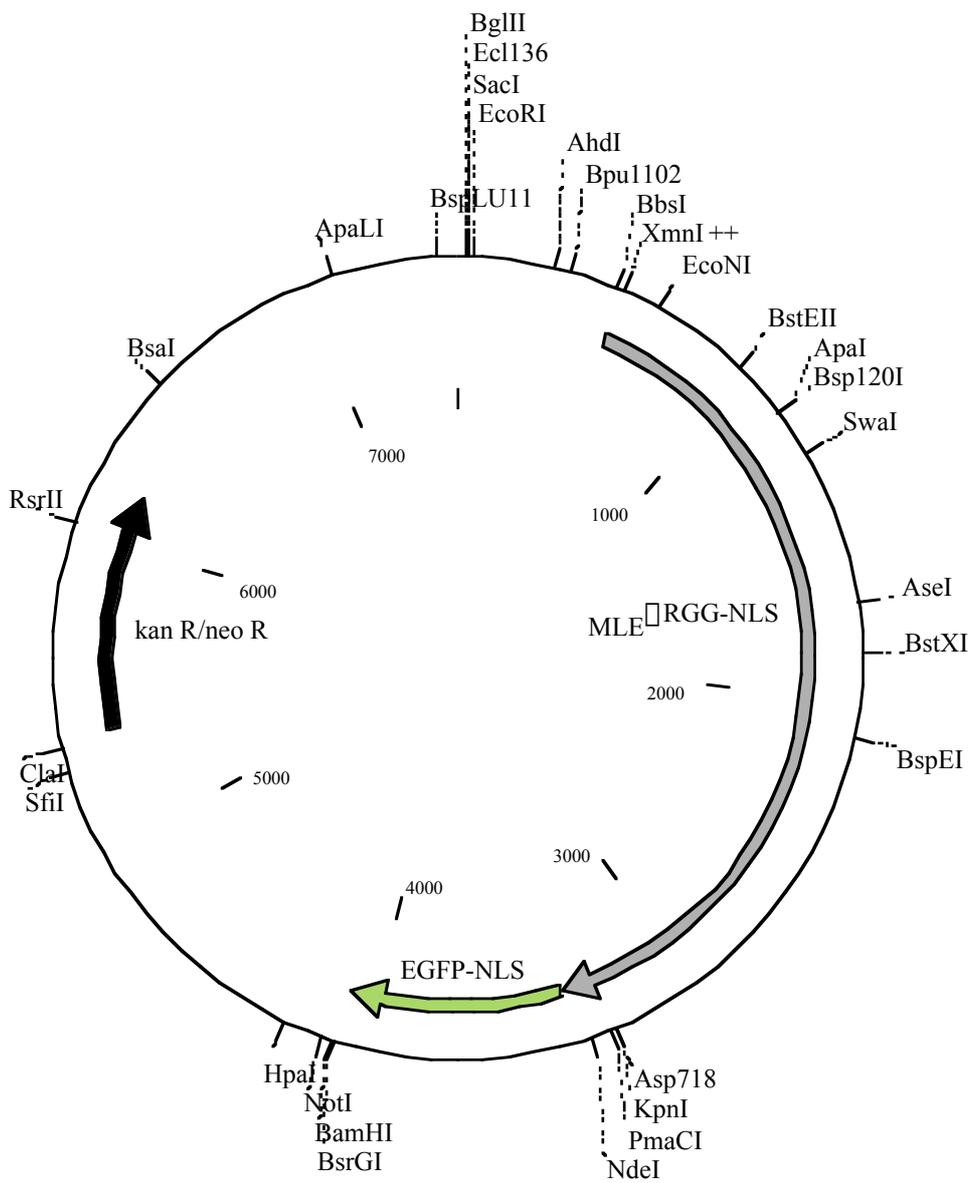
pHSP70-MLE<sup>ΔRB12</sup>-EGFP



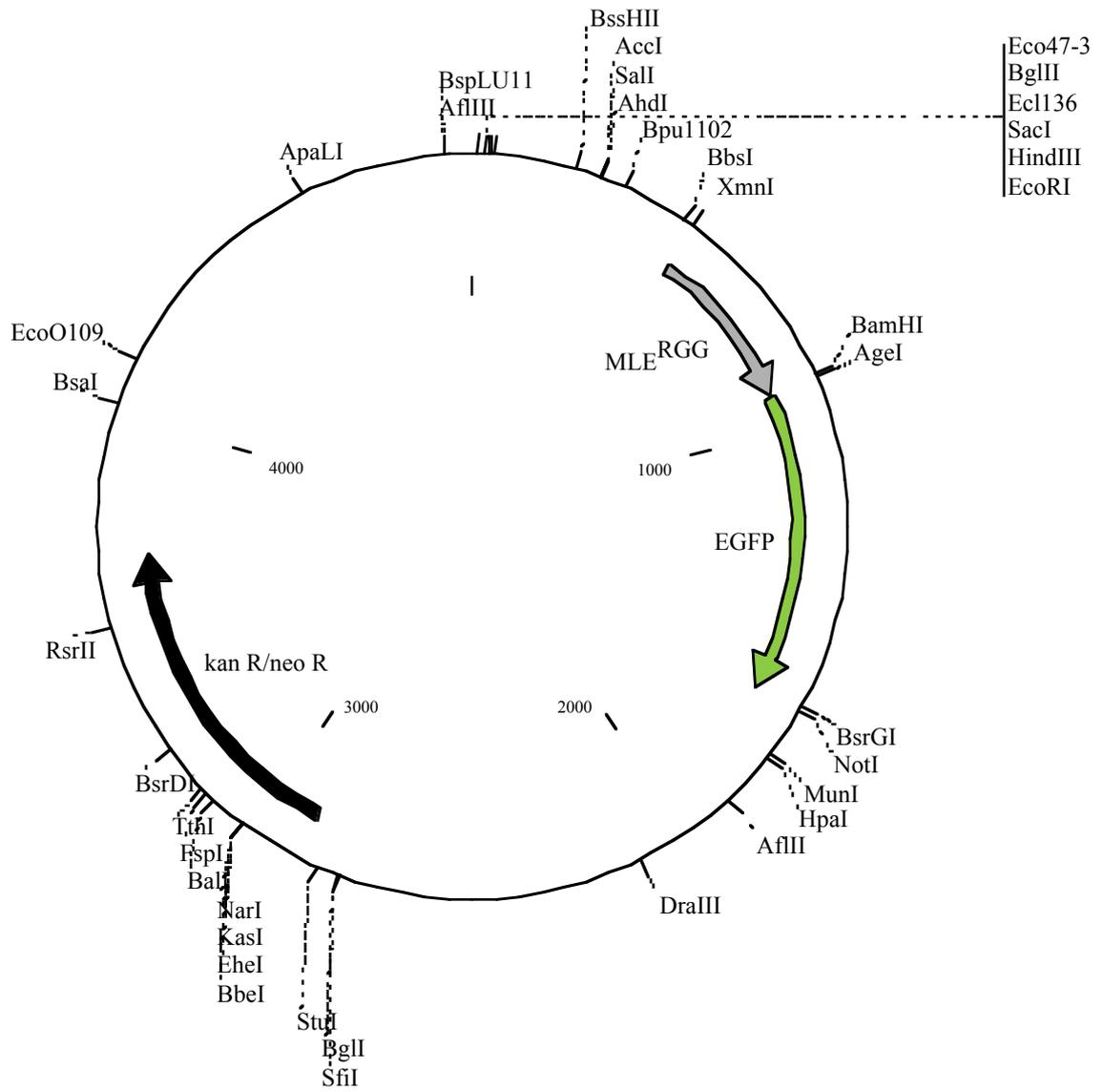
pHSP70-MLE<sup>ΔRGG</sup>-EGFP



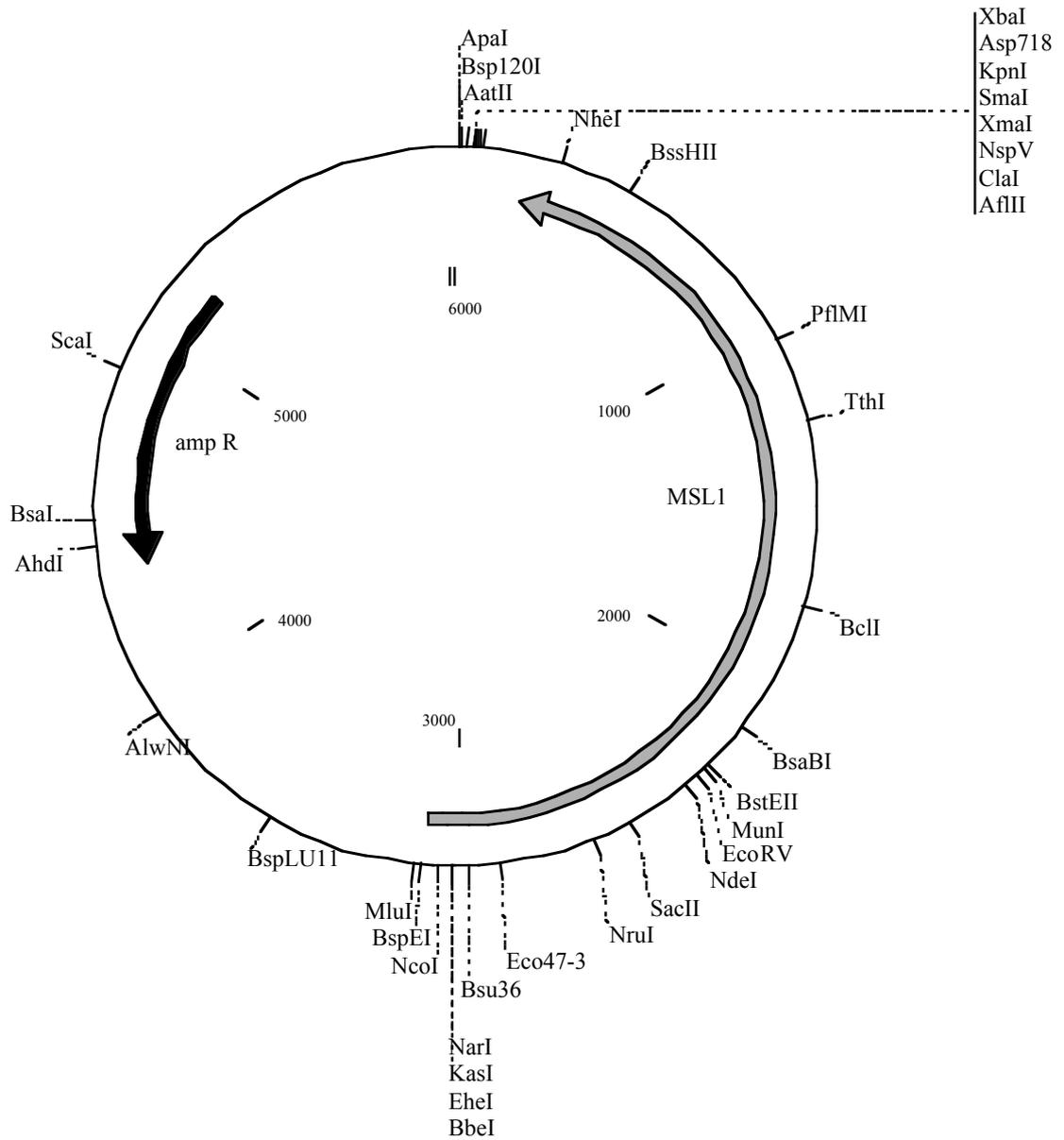
pHSP70-MLE<sup>ΔRGG</sup>-EGFP-NLS



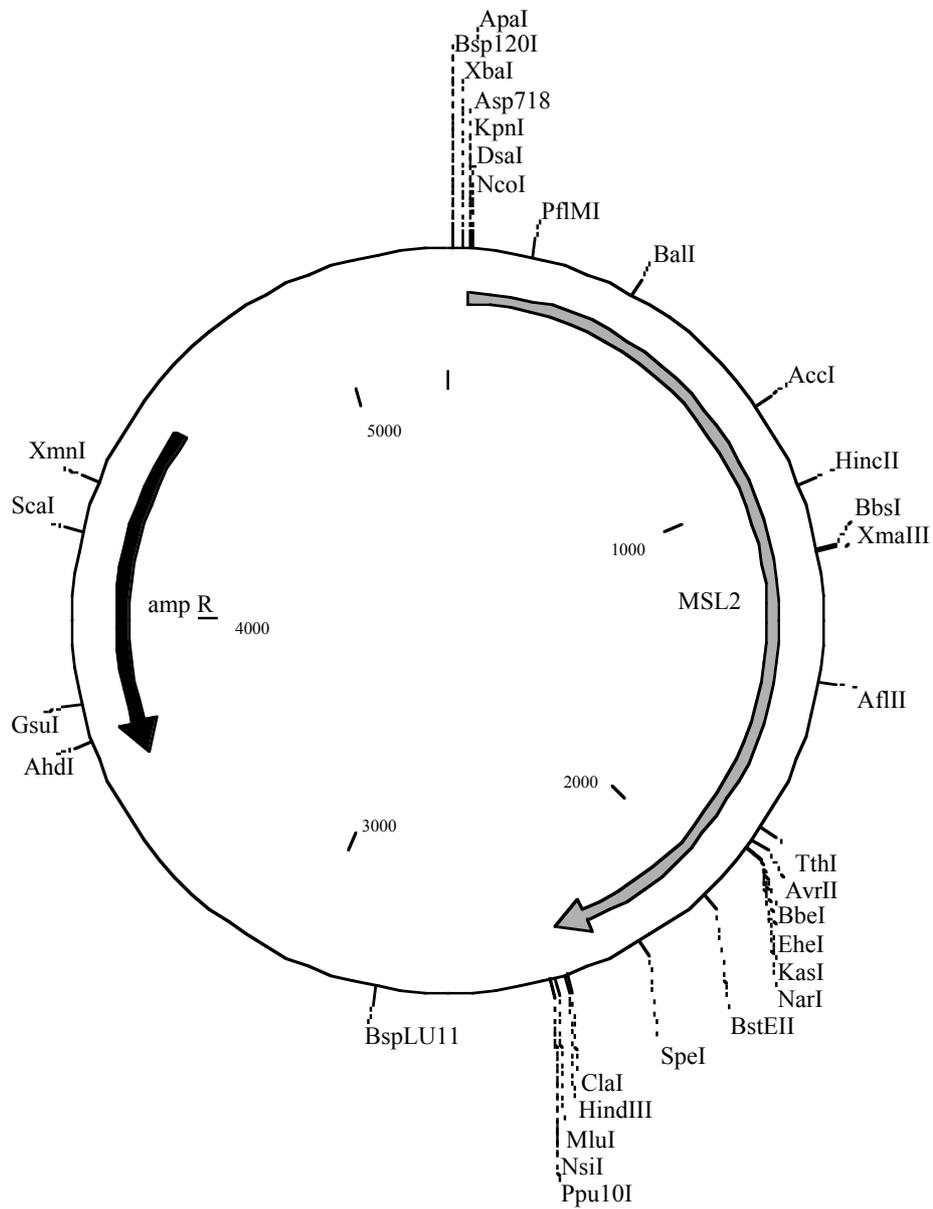
pHSP70-MLE<sup>RGG</sup>-EGFP



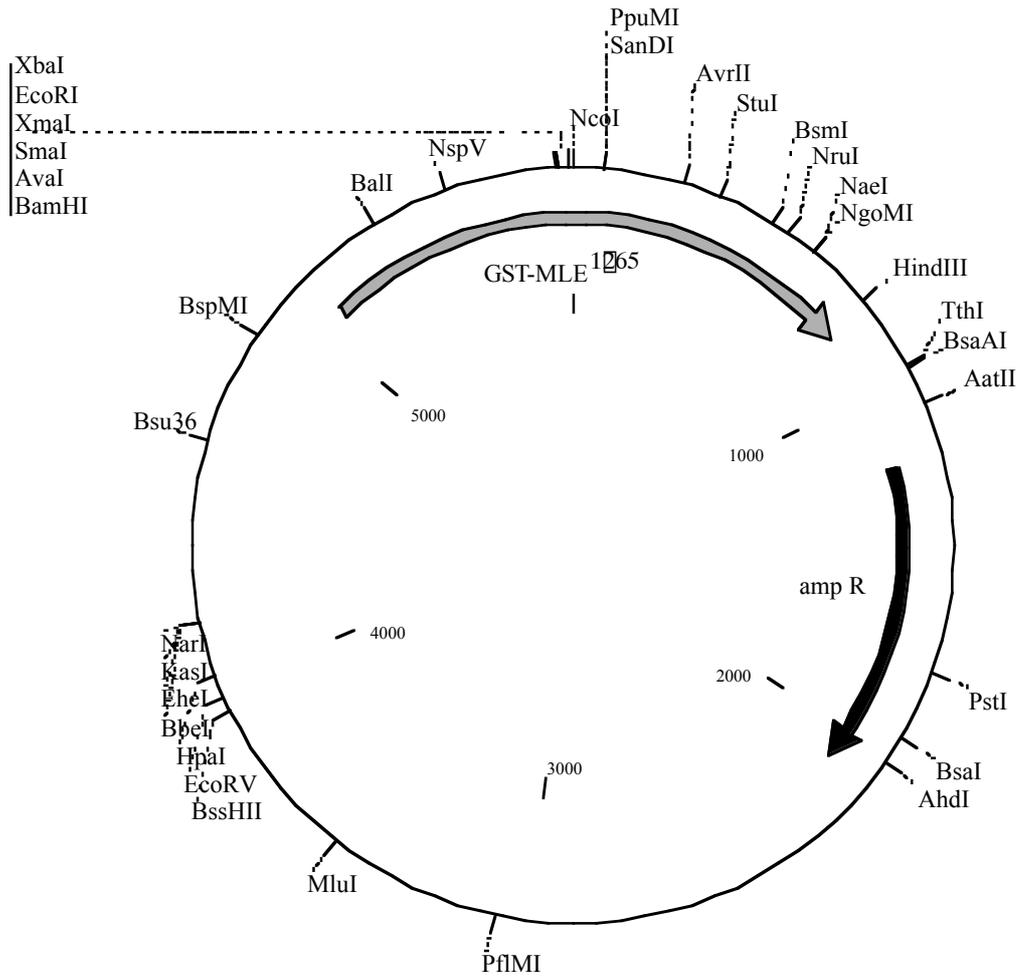
**PGEM-Sp6-MSL1**



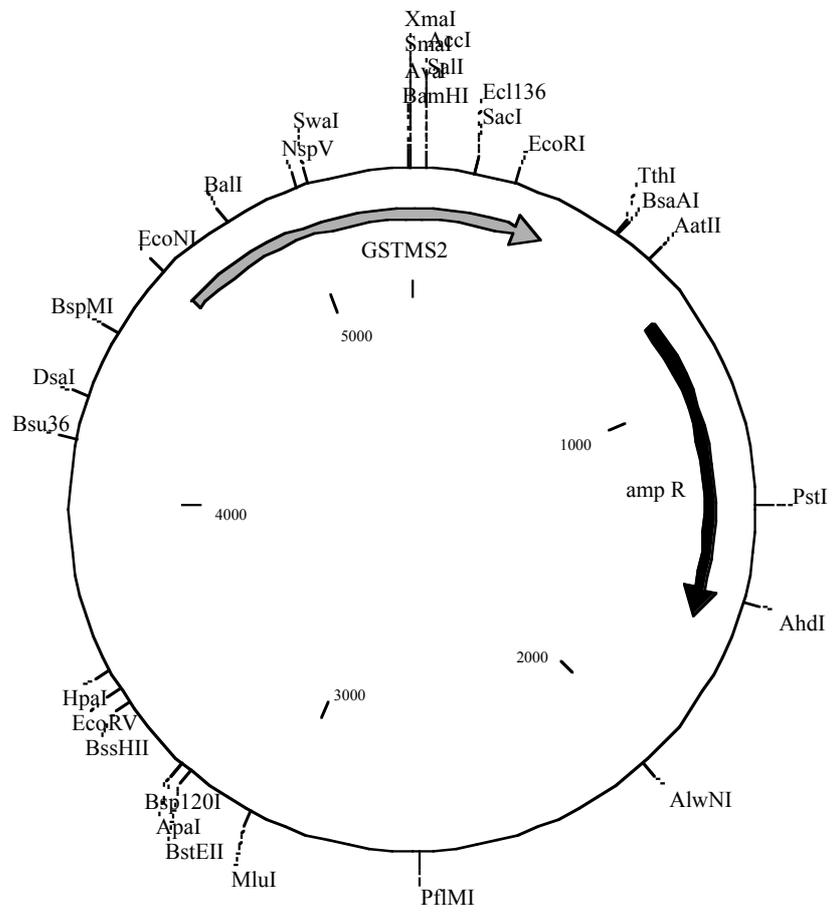
**pGEMT7-MSL2**



pGEX2KG MLE<sup>1-265</sup>



**pGEX2T-MS2**



## VII Curriculum vitae

**Name:** Annalisa Izzo

**Born:** Varese (Italy), November 11, 1974

**Nationality:** Italian

### *Education*

**July 2000:** Biological Science Degree, University of Insubria (Varese), Italy

Votation: 110/110 summa cum laude

**Thesis Title** " DNA methylation induces transcriptional repression diffusing *in cis*"

**July 1993:** High School Diploma Liceo Classico E. Cairoli (Varese), Italy.

### *Research activity*

#### **January 2002-2005: PhD Student**

Ludwig-Maximilian-Universität, Adolf-Butenandt-Institut

Molekularbiologie, Munich, Germany

Dosage Compensation in *Drosophila Melanogaster*

Group leader: Prof. Dr. Peter B. Becker

#### **September 2001-December 2001: Post-Diploma Student**

(Fellowship offered from the Italian Foundation for Cancer Research)

University of Insubria,

Molecular Biology Lab, Varese, Italy

DNA methylation

Group leader: Prof. Dr. G. Badaracco

#### **October 1997-July 2000: Diploma Student**

University of Insubria,

Molecular Biology Lab, Varese, Italy,

DNA methylation

Supervisors: Dott. Nicoletta Landsberger and Prof. Dr. G. Badaracco

***Publication List***

M. Curradi, **A. Izzo**, G. Badaracco and N. Landsberger  
“Molecular mechanisms of gene silencing mediated by DNA methylation.”  
Mol Cell Biol. 2002 May; 22(9): 3157-73.

V. Morales, C. Regnard, **A. Izzo**, I. Vetter and P. B. Becker  
“The MRG domain mediates the functional integration of MSL3 into the Dosage  
Compensation Complex”  
Mol Cell Biol. 2005 Jul;25(14):5947-54